Exciton mobility and trapping in europium doped vanadate phosphors described by Markov processes

Alexander Enthoven

10th of September, 2025

Supervisor Applied Mathematics	Prof.dr. F.H.J. Redig
Supervisor Applied Physics	Dr. E. van der Kolk
Daily supervisor	Drs. J. Zom
Committee member AM	Dr. A.B.T. Barbaro
Committee member AP	Dr.ir. D. Lathouwers

TU Delft

Abstract

Diffusion-like behavior of excitons within europium doped vanadate phosphors is studied, with emphasis on the trapping potential induced by Eu^{3+} and VO_4^{3-} luminescence. We describe the hopping, energy transfer, and decay of excitons by means of a continuous-time Markov chain distinguished by a local and delocalized trapping potential. Our model compares luminescence from decay of excitons with the trapping problem, where VO_4^{3-} traps are present everywhere, while Eu^{3+} ions are placed randomly. We distinguish two types of trapping potentials induced by europium traps. In the localized model, simplify the lattice to $\mathbb{Z}^d$ where the trapping potential of europium traps is only present at randomized sites in the lattice. By applying the large deviation principle for the occupation times measure, as done by Donsker and Varadhan previously, we find probability mass asymptotics of the form $\exp\left(t^{\frac{d}{d+2}}\right)$. For a simulation-like approach, we describe the delocalized model, where exciton movement and trapping in crystal lattices of $Rb_3LuV_2O_8$:Eu^{3+} and YVO_4:Eu^{3+} and is modeled using parameters such as temperature, migration rates, emission lifetimes, and europium concentration to estimate luminescent properties of the materials. Experimental data from vanadate lifetimes of $Rb_3LuV_2O_8$:Eu^{3+} and VO_4^{3-}/Eu^{3+} emission ratio's of YVO_4:Eu^{3+} are compared to the model, which agree relatively well within the confidence of the model. For YVO_4:Eu^{3+}, we have found an activation energy of $E_a = (101 \pm 4)\,\text{meV}$. We discuss how our simplified model can be extended to incorporate thermal and concentration quenching, and how to account for defects in the lattice. While we discuss our findings for europium doped vanadate phosphors, the results we have found are applicable for a range of other luminescent materials.

Contents

Acknowledgements

I would like to express my gratitude towards Frank Redig for his guidance and feedback during my bachelor thesis, especially regarding semigroups, large deviation theory and derivation of total mass asymptotics for Brownian motion. But also towards Jeffrey Zom for the idea for this project, feedback and discussion during the lunch and coffee breaks. Finally, I want to thank Erik van der Kolk for his guidance, feedback and support during the project.

1 Introduction

Luminescent materials play a key role within our society, for example the colors of LEDs [Dan+20], but also luminescent solar concentrators (LSC) which are in summary glass panes coated such that some light is reflected sideways to generate electricity [CB23]. Another application is the absorption of UV-light from solar radiation and re-emitting this as photosynthetically active radiation (PAR) which benefits plant growth [Mül+24]. In our case, we consider europium doped vanadate phosphors. We break down what we mean by this term. First, phosphors describe a material which emit 'long'-lasting light, which we call emission, when excited with higher-intensity light. This phenomenon, where a material emits light, is called luminescence. Now, we explain what the constituents are within these phosphors. Europium is an element belonging to the lanthanides, classified as a rare earth element known for its use as a fluorescent materials, emitting a reddish color. For example, the Eu^{3+} ion we study is also found in the 5 euro note [SM02]. The other ion we discuss is the vanadate group VO_4^{3-}, which accommodates so called self-trapped excitons, which are localized around the vanadium atom. This group can absorb ultraviolet light and can also emit light in the visible spectrum, characterized by blue light. Our thesis will study the phosphors $Rb_3Lu_{1-x}V_2O_8{:}Eu_x^{3+}$ and $Y_{1-x}VO_4{:}Eu_x^{3+}$, where x denotes the europium concentration. This means that some lutetium (Lu) or yttrium (Y) atoms are replaced by Eu^{3+} ions, a process which is called doping. Throughout this thesis, we will use the notation $Rb_3LuV_2O_8{:}Eu^{3+}$ and $YVO_4{:}Eu^{3+}$ to describe these materials for different europium concentrations.

The mechanism within these phosphors will be described by so-called excitons, which are pseudo particles consisting of an electron and a hole (the absence of an electron). How this exciton interacts with its surroundings on the unit-cell scale is generally very random, but we can measure and theoretically predict at which rate some interactions occur. This is where we apply the continuous-time Markov chain, which is in essence a sort of model for describing memory-less random processes. In other words, what a particle will do as given by the Markov chain, will not depend on its state history. Markov chains, discrete or continuous, are widely used in for instance population prediction (birth-death processes), financial models and queueing theory (computer networking). A special case of a continuous-time Markov chain is so-called Brownian motion, which is named after Robert Brown, which studied how small pollen particles randomly moved when suspended in water. While Brownian motion is a theoretical concept and is continuous in nature, it can be approximated using simulations and rescaling using Donsker's theorem.

While on the topic of probability theory, we will study large deviation theory and apply this to Markov chains. However, the large deviation principle is used in broad range of fields, such as statistics, statistical physics and queueing theory [Tou11]. Generally speaking, it studies how a general phenomenon in sample means, which is essentially $\frac{1}{m}(X_1+X_2+\cdots+X_n)$, are distributed in the form $e^{-tI(s)}$ (for continuous random variables). A key mathematician in large deviation theory is Varadhan, which together with Donsker (of which, Donsker's Theorem is named after), published many articles on applying the large deviation principle on Brownian motion, which we will study and apply throughout this thesis. In more recent years, Wolfgang König published 'The Parabolic Anderson Model,' which expands on the concept of a randomized potentials and how this changes the total mass asymptotics. This area of research where this is applicable is statistical physics, but is also connected to solving the Schrödinger equation for random potentials.

To assist in reading this thesis, we have provided an overview of the notation and symbols used in appendix A.1. This comes in handy when reading certain sections of the thesis or when a certain notation or abbreviation is not clear. While the topic of this thesis spans multiple fields (mathematics, physics and chemistry), the goal is to combine the knowledge of these fields to further expand the knowledge on this topic. And lastly, we note that a very similar model which we discuss, was given in a research article by C. Hsu and R. C. Powell in 1975 [HP75].

2 Exciton theory

Our first actual chapter begins with the particles we will be modeling, which we call **excitons**. We will give the physical foundation of these particles, and how these behave compared to the crystal structure they interact with. The life of an exciton can be describes as follows. This is also illustrated in figure 1.

1. **(Creation)** We create an exciton in the crystal by emitting light at a specific wavelength, usually in the ultraviolet light spectrum. This wavelength is called the excitation wavelength, since we excite the crystal with some energy which will be used to create the exciton. In our thesis, we assume the exciton starts somewhere randomly, therefore we leave out how these excitons are spawned within the crystal. While this is possible to model, this is not the goal of the thesis.

2. **(Movement)** We distinguish two types of excitons. The first type are mobile excitons, which are free to move through the crystal structure. Another possibility is the self-trapped exciton, which is 'trapped' at some site in the crystal structure. Although this type of exciton is 'trapped,' it can move to other sites within the crystal, called hopping or migration. We will describe what exactly this self-trapped exciton is in the following section.

3. **(Trapping)** The excitons are present at, aside when moving, at so-called VO_4-group sites (the sensitizer ions). These groups host the exciton, but in this excited state of the VO_4-group, the exciton can be annihilated and create blue/white-like light. This is not the only method of emitting light, the exciton can also transfer its energy to an europium ion, specifically Eu^{3+}, more generally referred to as the activator ion. This creates an excited state within the europium ion, emitting red-like light. These phenomena are described in the last two sections of this chapter.

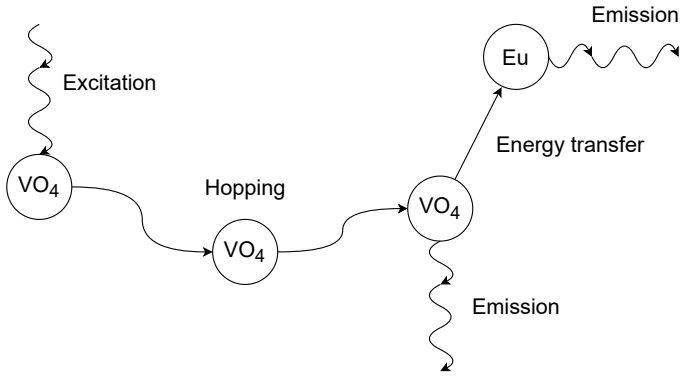


Figure 1: Life of an exciton. The exciton is created when we shine light onto the crystal, so called *excitation*. When created, the exciton can *hop* to other VO_4-sites, also called migration. At any point in time, the exciton can decay as ${VO_4}^{3-}$ emission. When nearby an Eu^{3+} ion, it can *transfer its energy* to the europium atom, which then also emits light, but different compared to ${VO_4}^{3-}$ emission.

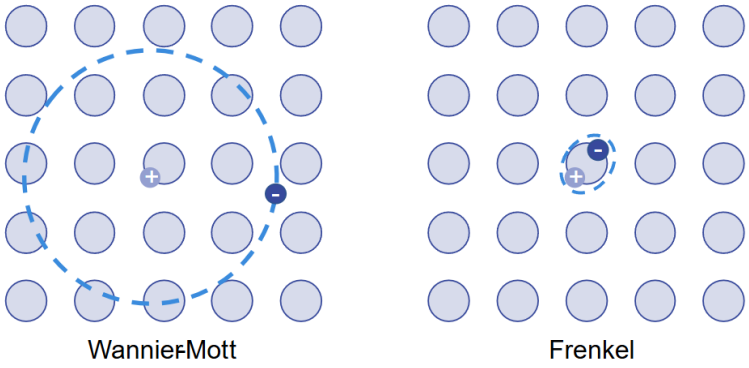
2.1 Self-trapped exciton

Excitons are quasiparticles within crystals which are neutrally charged and may move through the lattice to transfer energy. More generally, an exciton can be described as an excited electron in the conduction band combined with the hole it left in the valence band, which are in relative close proximity of each other such that they are bound by Coulomb forces. This exciton can, depending on the material, move to another atom where the exchange of the hole in the valence band is equivalent to moving an electron from the valence band from one atom to another. Since the exciton moves as a pair of an electron and hole, there is effectively no net flow of charge.

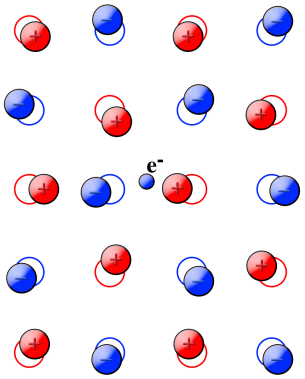
We know by Coulomb's law that negative and positive charges attract and hence there is an electric potential between the electron and hole. However, due to the charges from the nuclei or other electrons in the crystal, the effective charge of the electron and hole is 'screened', which reduces the electric potential between the electron and hole. We describe this potential as follows

$$V(r) = \frac{-e^2}{\varepsilon r}$$

Here, r is the distance between the electron and hole, and ε is the dielectric constant depending on the screening. For example, $\varepsilon \approx 2$ for ionic crystals [Lia70], so that the potential is reduced. In this manner, we consider two different excitons depending on the radius of the exciton. When the electron-hole pair is spread out over multiple lattice sites, we have a 'large' exciton. This is also called the 'Wannier-Mott' exciton, named after Gregory Wannier and Nevill Francis Mott. Due to the screening and thus higher dielectric constant, Wannier-Mott excitons have a low binding energy. On the other hand, when the electron-hole pair is localized to one site, in other words the distance between the electron and hole is smaller than the distance between the atoms, we have a 'small' exciton. Similarly, this is a Frenkel exciton, named after Yakov Frenkel. In this case, there is a small dielectric constant such that the potential of Frenkel excitons is typically greater than Wannier-Mott excitons. See figure 2a for an illustration of the aforementioned exciton types.



(a) Simplified illustration of Wannier-Mott (left) and Frenkel (right) excitons. (Source: [OKa])



(b) Overview of a polaron, distorting the nearby lattice. (Source: [KEL09], CC BY-SA 4.0)

Regarding the energy states of the exciton, since we consider the Coulomb electric potential, the exciton has similar energy levels to the Bohr model for electrons around a nucleus. More specifically, the energy levels can be described by

$$E_n = E_\infty - \frac{R}{n^2}$$

Here, R is the binding energy such that R energy is required to ionize the exciton (in the sense that we get a free electron), and $n = 1, 2, \ldots$ is the state of energy level E_n. When ionized, we set the energy to E_∞, similar to $n \to \infty$.

In this manner, the exciton disappears when it is ionized. The energy of the exciton can also be lost due to collisions with other particles within the lattice. Another possibility is that the exciton is reduced to the ground state, where the energy is transferred by emission of light called luminescence.

Let us consider a specific type of trapping of excitons, namely when the exciton-phonon is strong enough to self-trap the exciton. By exciton-phonon interaction or exciton-phonon coupling we mean how the crystal is distorted due to the charge of the excited electron and hole. For example, consider the excited electron within the lattice. Then the electrons around atoms are repelled and the positive nuclei are attracted towards the excited electron. Due to this deformation of atoms localized around the exciton, the exciton is trapped

within the deformation. For this self-trapped exciton (STE) to exist, we thus require strong exciton-phonon coupling. Also, the exciton must have a radius similar to the lattice constant ('small' excitons), since local charge neutrality increases the exciton-phonon coupling. An electron can distort its local crystal structure due to the charge, but if a hole is far away (in the sense that is some atomic sites away from the electron), then the hole does not influence the crystal structure localized around the electron as much due to screening and distance. We note that if the exciton-phonon coupling is weak, then the movement of the exciton is 'scattered,' similar to if excitons and phonons are particles. In this case, the electron-hole pair interacts weakly with the surrounding crystal, which influences the trajectory of the exciton.

The case of self-trapping excitons is similar to *polarons*, which are, similar to excitons, quasiparticles describing the distortion of the lattice caused by a charge carrier (electron or hole) [Fra+21]. For instance, see figure 2b. Regarding polarons, we will refer to the electron which distorts the crystal structure. Since a hole has the same properties as an electron beside the charge, we have similar properties for holes. Due to this, we speak of electron-phonon coupling or interaction instead of exciton-phonon coupling for excitons.

Again, there exist two types of polarons which depend on the size of the polaron. The 'large' polaron is described by the Fröhlich model. This polaron encompasses multiple lattice sites, similar to the Wannier-Mott exciton, relying on long-range electron-phonon coupling. In this case, we speak of a continuum polaron, as the wave function of the large polaron is spread out over sites, and as such discrete lattice points are not taken into account. Similarly, there exists a 'small' polar also called Holstein polarons, named after the Holstein model. This model does take discrete lattice sites into account, and assumes a short-range electron-phonon coupling takes place. Another distinction between Fröhlich and Holstein polarons is the method of mobility. For Holstein (small) polarons, we may model the movement of polarons as moving from site to site, in a sense that the polaron is 'hopping'. To model the rate of hopping, we consider the Marcus theory of electron transfer reactions, as will be clear in the next section. As for Fröhlich (large) polarons, since the lattice is regarded as a continuum, the polaron can freely move, and can be modeled to an extent as if it were a free particle.

2.2 Holstein model

Since the scope of the thesis is about modelling hopping of self-trapped excitons in a discrete space, we consider 'small' excitons which behave as so-called Holstein polarons. In this section, we examine the model of small polarons as described by T. Holstein in 1959 [Hol59a]. The model of Holstein considers a lattice of molecules. While we mathematically derive the Hamiltonian, the important aspect will be to consider which approximations are made within the model. This derivation also is meant to built upon the derivation of Holstein in the appendix of his aforementioned paper.

To suit our model, we replace molecules as stated in Holstein's model by atoms since we expect hopping of excitons to occur between atoms, which may exist within molecules. For simplicity, we consider a one-dimensional chain of N atoms as a lattice living in a 3-dimensional space, see for example figure 3. The approach of solving for the wave function and energy of the bound electron will be by describing the wave function of the electron by the *tight-binding model*. The position of the electron will be denoted by $\mathbf{r}$, the vector between nearest-neighbor is $\mathbf{a}$ (in other words, the lattice vector). The equilibrium position of the n atom is $n\mathbf{a}$ having a displacement x_n along $\mathbf{a}$. In our situation, we will describe the Hamiltonian of a system with one electron and multiple atoms. The Hamiltonian can be described as the sum of the following components.

$$H = H_e + H_{\text{lat}} + H_{\text{int}} \tag{2.1}$$

We explain the following components below.

1. H_e consists of the kinetic and potential energy of the electron. This potential U is due to the electric force between the electron and atoms. This Hamiltonian can be written as

$$H_e = -\frac{\hbar^2}{2m_e}\nabla^2 + \sum_{n=1}^{N} U(\mathbf{r} - n\mathbf{a}, x_n) \tag{2.2}$$

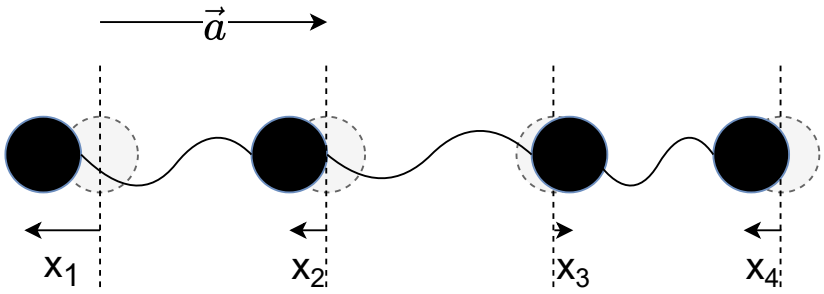


Figure 3: Illustration as example for $N = 4$ atoms. Every atom has an equilibrium position, as drawn by a dotted circle, which is displaced by a length x_n below. The atoms are connected by a 'spring'-model, such that every atom has a harmonic oscillator vibrational energy levels, as illustrated by connecting the atoms by 'strings'. The distance between every equilibrium position is consistent, as given by the lattice vector $\mathbf{a} = \vec{a}$. The equilibrium positions are equivalent to the lattice sites.

2. H_{lat} is the lattice contribution to the Hamiltonian. It consists of the kinetic energy of all N atoms and the sum of all harmonic oscillator potentials of the N atoms. For an atom of mass M (or effective mass of the molecule, as per the paper of Holstein), we have an individual potential of $\frac{1}{2}M\omega_0^2 x_n^2$. This is similar to the approach used in the Einstein model, so we have vibrational states for every atom, having energy level $E_n = \hbar\omega_0(n + \frac{1}{2})$

   Lastly, the Hamiltonian is given as.

$$H_{\text{lat}} = \sum_{n=1}^{N} \left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x_n^2} + \frac{1}{2}M\omega_0^2 x_n^2 \right) \tag{2.3}$$

3. The contribution H_{int} describes the interaction term between the electron and lattice. In the previous section, this was referred to as the electron-phonon or exciton-phonon interaction. This term is usually considered to be linear term, for example Ax_n where A is the electron-phonon coupling constant and x_n is the displacement from the lattice. For now, we ignore this component of the Hamiltonian and introduce it later. It is important to note that, especially in self-trapped excitons, this term is usually large compared to H_e or H_{lat}.

To this extent, we apply the tight-binding model with some peculiarities. First, let $|\phi_n(\mathbf{r}, x_n)\rangle$ be the tight-binding wave function of the electron around atom n. Since we consider identical atoms, we can write $|\phi_m(\mathbf{r}, x_n)\rangle = |\phi(\mathbf{r} - n\mathbf{a}, x_n)\rangle$. We note that $|\phi_n\rangle$ also depends on x_n since the potential is displaced by x_n. In Holstein's model, the collection of wave functions ϕ_n does not necessarily have to be pairwise orthogonal. It is essential for the derivation of hopping to other sites that there exists some overlap between the integrals, which will be discussed later in this section. In the fashion of the tight-binding model, we describe the wave function of the electron as ψ, given by

$$|\psi(\mathbf{r})\rangle = \sum_{n=1}^{N} a_n(x_1, \ldots, x_n)|\phi_n(\mathbf{r}, x_n)\rangle \tag{2.4}$$

Here, $a_n(x_1, \ldots, x_n)$ represents the coefficient of the wave function $|\phi_n(\mathbf{r}, x_n)\rangle$. Since we assume that every coefficient depends on all displacements $x_1, \ldots, x_n$, we write $a_n(\mathbf{x})$ where $\mathbf{x} = (x_1, \ldots, x_n)^T$.

Now, our goal is to rewrite the tight-binding Hamiltonian $H_0 = H_e + H_{\text{lat}}$, which will motivate some assumptions of the model and definitions used throughout the section. For reference, we will rewrite $H_0|\psi\rangle = i\hbar\partial_t|\psi\rangle$, where we use $\partial_i f = \frac{\partial f}{\partial i}$. Combining equations (2.2) and (2.3) gives

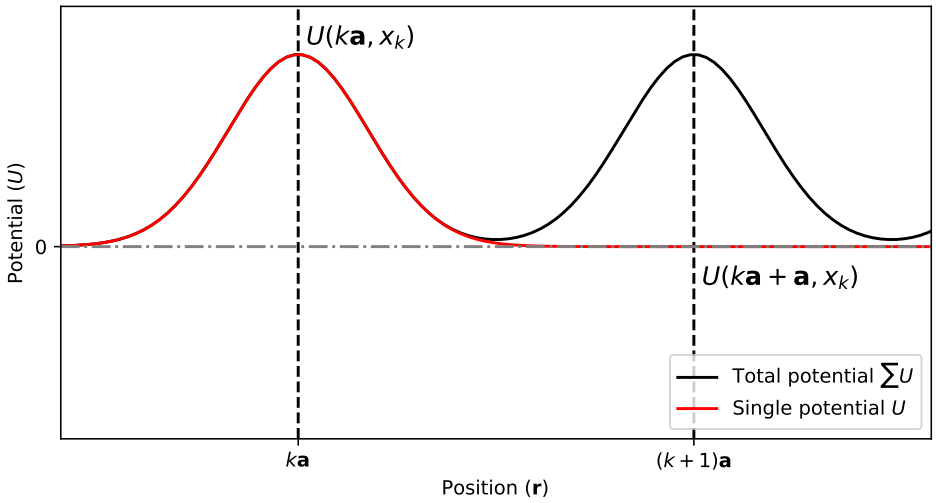


Figure 4: Example where the short-range approximation ($U(\mathbf{a}, x_n) \ll U(\mathbf{0}, x_n)$) is illustrated. At the position $\mathbf{r} = k\mathbf{a}$ we have an atom and we expect that the potential due to this atom is negligible at $\mathbf{r} = (k+1)\mathbf{a}$, which is the neighboring atom. In this drawing, we assume that the displacements x_k and x_{k+1} are zero.

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \Bigg[\underbrace{-\frac{\hbar^2}{2m_e}\nabla^2 + \sum_{n=1}^{N} U(\mathbf{r}-n\mathbf{a}, x_n)}_{H_e} + \underbrace{\sum_{n=1}^{N}\left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x_n^2} + \frac{1}{2}M\omega_0^2 x_n^2\right)}_{H_{\text{lat}}}\Bigg]|\psi\rangle \tag{2.5}$$

The approximation used in the tight-binding model is that $|\phi_n\rangle$ represents an orbital of atom n. We may do this because we make the assumption that U is short-range, this can be stated as

$$U(\mathbf{a}, x_n) \ll U(\mathbf{0}, x_n) \tag{2.6}$$

This concept is illustrated in figure 4.
We do not completely neglect the potential at distances of the order of $|\mathbf{a}|$, but we may 'approximate' that the Schrödinger equation for $|\phi_n\rangle$ holds solely for the potential $U(\mathbf{r} - n\mathbf{a}, x_n)$. In the Holstein model, there are additional approximations made, which mostly consist of ignoring small terms between site wavefunctions and the atomic potential.

After a long derivation, the Holstein Hamiltonian can be written in a simplified way. First, we introduce the following notation.

$$W_l(\mathbf{x}) = \langle\phi_l| \sum_{p\neq l} U(\mathbf{r}-p\mathbf{a})|\phi_l\rangle \tag{2.7}$$

$$J(x_l, x_n) = \langle\phi_l|U(\mathbf{r}-l\mathbf{a})|\phi_n\rangle \tag{2.8}$$

Here, W_l represent the perturbation of other molecular orbitals and J is the overlap integral between its neighbors. Using these definitions, Holstein found the Schrödinger equation as follows.

$$\left[i\hbar\partial_t - E(x_l) - \sum_p \left(-\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial x_p^2 + \frac{1}{2}M\omega_0^2 x_p^2}\right) - W_l(\mathbf{x})\right] a_l(\mathbf{x}) = \sum_{\pm} J(x_l, x_{l\pm1}) a_l(\mathbf{x}) \pm 1$$

At this point, the energy can be found by applying the ansatz $a_n^{(k)} = e^{ikn}$ as is common for tight-binding models. This provides a basis for further derivation of variables regarding excitons. For these derivations, we refer to his paper [Hol59a].

2.3 Hopping of excitons

One of the important results derived from Holstein's model is the hopping rate between sites. The derivation is done in Holstein's second paper [Hol59b]. A similar result can be found in [BB85] and [AD09] [1].

While we spare the effort of reproducing the steps to derive the hopping rate, we do cover how this rate is found. The derivation of the hopping rate is based on Fermi's Golden Rule. In general, the rate between two quantum states i and j with a perturbed Hamiltonian H' is $\frac{2\pi}{\hbar}|\langle j|H'|i\rangle|^2\rho(E_j)$. The term $\rho(E_j)$ represents the density of states at the energy E_j of state j. The energy $\langle j|H'|i\rangle$ is similar to how we have defined J with respect to the perturbation of atomic potentials.

We note that the temperature T influences the vibrational states of the atoms. Implementing the partition function and thermally averaging the hopping rate will eventually give a temperature dependence proportional to $e^{-\frac{E}{k_B T}}$, where k_B is the Boltzmann constant and E the energy. Therefore, we expect Arrhenius-like temperature behavior.

The hopping or migration rate that was found for excitons is given by

$$k_m(T) = \frac{J^2}{\hbar^2}\sqrt{\frac{\pi}{4k_B T E_a}}\exp\left(-\frac{E_a}{k_B T}\right) \tag{2.9}$$

$$E_a = \frac{1}{\pi}\int_0^\pi \frac{A^2}{4M\omega_k^2}(1-\cos k)\,\mathrm{d}k \tag{2.10}$$

In the above equation, k_m is the temperature-dependent migration rate for excitons. The factor J is the function $J(x_l, x_l \pm 1)$ evaluated at the neighboring sites. It is assumed that this is the same for all points, therefore J is assumed to be a constant. The term E_a refers to the **activation energy** for the exciton, of which the definition is also given. Physically, the activation energy E_a determines the difficulty for the excitons to overcome the barrier to migrate to other sites. A higher activation energy will decrease the migration rate. The activation energy is determined by the electron-phonon coupling constant A, mass of site M and ω_k^2 which is related to vibration frequency ω_0. We note that stronger electron-phonon coupling increases the activation energy and therefore decreases the mobility of excitons.

In the model, we do not calculate k_m directly using equation (2.9), but we do want to take temperature dependence into account. By defining a migration rate $k_m(T_1)$ as a temperature T_1, we find the migration rate at a different temperature T_2 as follows.

$$k_m(T_2) = k_m(T_1)\cdot\sqrt{\frac{T_1}{T_2}}\cdot\exp\left(-\frac{E_a}{k_B}\left(\frac{1}{T_2}-\frac{1}{T_1}\right)\right) \tag{2.11}$$

We will explore this temperature dependence in chapter 6.

Note 2.1. *(Marcus-Hush theory)* Until now, we have derived the hopping rate from the tight-binding model of excitons. An analogue of this migration rate can be found by examining diabatic (non-adiabatic) charge transfer from donor to acceptor molecules [Ste+14; Wan+24]. The theory describing this model is called the Marcus-Hush theory named after Rudolph A. Marcus and Noel Hush. In this theory, we have two distinguish two states: (1) the donor and acceptor in a neutral state and (2) the donor donates an electron to the acceptor, creating a positively-charged donor and negatively-charged acceptor. In this diabatic case, the hopping rate between the states (1) and (2) is given by

$$k_m(T) = \frac{V_{ij}^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\exp\left(-\frac{\lambda}{4k_B T}\right)$$

[1] For Bottger: page 55, section 3.2. For Alexandrov: page 84, section 3.8.1.

Here, V_{ij} is similar to the overlapping integral J. The reorganization energy λ is similarly related to the activation energy E_a. The Marcus-Hush theory is very similar to the exciton theory of Holstein, due to the similarity in how the energy of the transitioning state is defined. In both models, we have two parabolic potentials of which have an energy difference of their respective minima. Therefore, Marcus-Hush theory is also applicable to the world of excitons.

Note 2.2. *(Migration between VO_4-groups)* While we have derived exciton migration between atom sites, we will model exciton movement between VO_4-groups and other atoms. This not only simplifies the model for mathematical analysis and reduces computational cost, the tight binding model can also be applied to groups of atoms, for example the VO_4-group, rather than solely atoms as Holstein model originally describes. We do model exciton hopping frm VO_4-groups to europium atoms, due to the energy transfer mechanism between VO_4 ions and lanthanides, as discussed in the next section.

2.4 Energy transfer

In the crystals $Rb_3LuV_2O_8$:Eu^{3+} and YVO_4:Eu^{3+}, luminescence occurs from both VO_4-groups and from europium emission. When the exciton is localized at a VO_4-site, it can transfer energy to an europium site using dipole-dipole interactions between the VO_4-ion and the Eu^{3+} ion. While other energy transfer mechanisms exist, for example energy transfer by exchange interactions[2], we assume this mechanism can be modelled by the Förster (dipole-dipole) energy transfer mechanism. This is due to its relatively large range and the fact we have charged groups and atoms. In summary, the underlying mechanism of Förster energy transfer is the electrostatic interaction between the two ions, which we describe as dipoles. In essence, the transfer can be characterized as follows.

$$\mathrm{VO}_4^* + \mathrm{Eu} \to \mathrm{VO}_4 + \mathrm{Eu}^*$$

Where * denotes the excited state of an ion.

For this type of energy transfer, it is shown [Ron07] that the rate is proportional to the inverse of the distance to the power of 6. Therefore, we will use the following relation for our model.

$$k_{\mathrm{ET}} \propto \frac{1}{(R_{\mathrm{VO_4-Eu}})^6} \tag{2.12}$$

Here, $R_{\mathrm{VO_4-Eu}}$ is the distance between the (dipoles of the) VO_4-ion and the europium ion.

Why this relation is $1/r^6$ can be explained as follows. It is known that the potential energy between two dipoles has the relation $V(r) \propto 1/r^3$. Since the transition rate depends on the related energy squared, see the $|\langle i|H'|j\rangle|^2$ term in Fermi's Golden Rule, we obtain that the transition rate between the two dipoles should be $1/r^6$.

2.5 Emission

At last, we also consider the emission of the VO_4-group and europium atoms. When either ion is in an excited state, it can emit light at a certain wavelength depending on the possible states of the ion. In a photoluminescence spectrum, the intensity of this light is given per wavelength. For now, we will review the theoretical aspect of luminescence of ${VO_4}^{3-}$ and Eu^{3+} ions.

Experimental results indicate a single broad photoluminescence peak from VO_4 in the region of 400 nm to 700 nm, having a maximum intensity at 500 nm to 520 nm [Dan+20; Zho+15]. This broad band emission is attributed to the charge-transfer luminescence of the bond $O^{2-} - V^{5+}$ in the VO_4-group. In this type of luminescence, emission occurs due to interaction between orbitals or electric states of different ions [Ron07], giving rise to broad emission spectra. Charge-transfer between oxygen and vanadium takes place since the exciton is localized within the VO_4-group. Charge-transfer luminescence will generally have a broad photoluminescence spectrum. This light emitted by VO_4-group emission in $Rb_3LuV_2O_8$:Eu^{3+} andYVO_4:Eu^{3+} is described to have a white/blue color.

[2]Also called 'Dexter' energy transfer, named after D.L. Dexter, uses interactions due to wave function overlap. This energy transfer mechanism has the relation e^{-R} with R the distance between the two ions.

We come back to the Eu^{3+} ion, having an electronic configuration of [Xe] $4f^6$. The last filled shell is then $4f$, having exactly 6 electrons due to the 3+ charge. Luminescence originating from Eu^{3+} comes from electronic transitions within the excited states of the orbital. We use term symbols, denoted by $^{2S+1}L_J$, to describe the states of Eu^{3+}. Here, S is the total spin quantum number, L is the total orbital angular momentum quantum number, and J is the total electronic angular momentum. The symbol used for L is S, P, D, F, ... for $L = 0, 1, 2, 3$, ... respectively, similar to s, p, d, f used for ℓ in the electronic configuration. To understand what these term symbols mean, we give two examples for the term symbols $^7\mathrm{F}_J$ and $^5\mathrm{D}_J$

Example 2.1. *(Term symbols ^7F_J, ^5D_J in Eu^{3+})* In the 4f-shell, we have $\ell = 3$ and so $m_l = -3, -2, -1, 0, 1, 2, 3$ are allowed. According to Hund's rules, we have the lowest energy when S is the largest (first rule) and have the lowest energy when L is the largest. So for the ground state, we will have our 6 electrons in the $\uparrow$ spin state, and have $m_l = 3, 2, 1, 0, -1, -2$ respectively. At this moment, we have a total spin of $6 \times \frac{1}{2} = 3$ and total orbital angular momentum $L = 3 + 2 + 1 + 0 - 1 - 2 = 3$ represented by the letter F. Therefore, the lowest energy states have the notation ^7F_j. According to Hund's third rule, the lowest energy goes to $J = |L - S|$ since is not more than half full. So, the ground state will be $^7\mathrm{F}_0$ having excited states $^7\mathrm{F}_1, \ldots, ^7\mathrm{F}_6$.
But what if we do not care about the lowest possible energy state. Instead, if we move an electron from $m_l = -2$ to $m_l = -3$ and switch the spin state, we find $L = 3$ (denoted by D) and $S = 2$. This is then an example of a state ^5D_J.

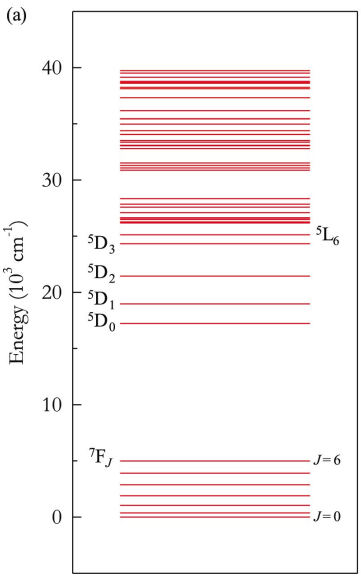


Figure 5: Energy levels of Eu^{3+} denoted by term symbols. Unit cm^{-1} represents the energy of a photon with a 1 cm wavelength, equal to approximately 1.986×10^{-23} J. (Source: [SBC23], CC BY-NC-ND 4.0)

But what are the energy levels of Eu^{3+}? If we only account for spin-orbit coupling, and ignore other finer splitting phenomena, we obtain figure 5. As stated in example 2.1, the level $^7\mathrm{F}_0$ is the ground state, with nearby excited states $^7\mathrm{F}_1$ through $^7\mathrm{F}_6$. Since the energy difference for the transitions between $^7\mathrm{F}_J$ are small (wavelengths greater than approximately 2000 nm), we ignore these transitions. The luminescence primarily comes from transitions from $^5\mathrm{D}_0$ and $^5\mathrm{D}_1$ to levels $^7\mathrm{F}_J$ for $J = 0, \ldots, 6$, having a red-like color [SBC23]. For example, the transition $^5\mathrm{D}_0 \rightarrow ^7\mathrm{F}_0$ has wavelength 580 nm. The transitions from level $^5\mathrm{D}_0$ are the most dominant compared to other levels from $^5\mathrm{D}_J$. Since the electrons stay in the 4f orbital during this transition, these are also called 4f-4f transitions.

There is another difference which sets transitions from levels $^5\mathrm{D}_0$ and $^5\mathrm{D}_1$ apart, aside from the emission wavelength peaks. Namely, the lifetime of transitions from level $^5\mathrm{D}_1$ are relatively quick, being around 10 μs. Meanwhile, the lifetime of transitions from level $^5\mathrm{D}_0$ are slow, being approximately 1.5 ms [Sev+17; CRC09]. To find the lifetime, we measure the intensity at a specific wavelength (the wavelength corresponding to the energy difference for a certain transition) with respect to time. During the transition, spontaneous emission occurs, resulting in a decaying curve for the intensity with respect to time [3]. Let this intensity be given by $A\exp(-t/\tau)$ where A is some constant, then τ is the lifetime, having units s. After τ time has passed, the intensity is reduced to 36.8% of its original intensity.

While the exciton may reach the VO_4-group and europium atom, it does not have to necessarily decay into light. This is called non-radiative decay, where the energy is transferred into heat or other forms of energy except light. In general, the rate at which non-radiative decay occurs is given by ([Ron07])

$$k_{nr} = A\exp\left(-\frac{E}{k_B T}\right) \tag{2.13}$$

[3] Given that $N(t)$ is the population size of a given excited state, we have $\frac{\partial N(t)}{\partial t} = -A_{21}N(t)$, resulting in a function $N(t) \propto e^{-t/\tau}$. Since intensity is proportional to energy which is then proportional to the amount of photons emitted, this will also decay similarly. As will become clear in the later chapters, this motivates the fact we can model (spontaneous) emission by some decay rate using a Markov process.

Where E is the energy difference between the two states of a transition. While we do not incorporate this into our model, it is an interesting discussion point since this influences the efficiency of our luminescent material.

3 Background continuous-time Markov processes

Our approach to the exciton problem will be to model the movement using a continuous-time Markov chain (CTMC). First, we will define what we mean by Markov chain and how we may calculate the transition probability of different timesteps. We will also give theory about semigroups regarding Markov chains. Our approach to analyze the model will require using semigroups and the Feynman-Kac formula. At the end of this chapter, we explore Brownian motion. This is also a Markov process living in a region $\mathbb{R}^d$.

3.1 Markov chain

3.1.1 Definitions

Markov chains represent how certain random processes evolve given a time space T and a state space S. We may think that a particle lives in the state space and may move (jump) to other states in the state space after a given time. The time space T represents which points of time the model considers. Then for any $t \in T$, let $X_t \in S$ be the path (or trajectory) of the particle in the state space. Since we describe random processes, X_t is a random variable for all $t \in T$. So, we describe the Markov process as a family of X_t for which we use the notation $(X_t)_{t\in T} = \{X_t \,|\, t \in T\}$.
In general, a Markov process can be divided into two types. We distinguish two possible time spaces for a Markov chain below.

1. T is countable, for example $T = \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$, then we have a discrete-time Markov chain. The set $(X_t)_{t\in\mathbb{N}_0}$ is then a discrete set.
2. T is uncountable, for instance $T = \mathbb{R}^+ \cup \{0\}$. Then $(X_t)_{t\geq 0}$ is called a continuous-time Markov chain. Whereas the countable state space describes jumps at fixed time steps, the continuous case must also take into account when the jumps take place. As will become clear in this section, both the jumps and the time are randomly distributed.

In this thesis, we will often encounter finite or countable state spaces such as $\mathbb{Z}^d$, where d refers to the dimension. Another important Markov process used throughout this thesis is Brownian motion, having an uncountable state space $\mathbb{R}^d$, which we examine in the last section of this chapter.

We describe a stochastic process in the discrete time by assigning a probability p_{ij} going to state $j \in S$ given it is currently in state i. This leads to an important assumption of Markov chains, which is the Markov property: the probability p_{ij} is independent of the history of a path X_t. This property is what defines a Markov process, as becomes clear when we give the definition of a discrete-time Markov chain.

Definition 3.1. (Discrete-time Markov chain) Let $(X_n)_{n\in\mathbb{N}_0}$ describe a stochastic process on a countable state space S. Let $n \in \mathbb{N}_0$ be arbitrary. If for all $i, j \in S$ and for all $x_0, \dots, x_n \in S$ we have

$$\mathbb{P}(X_{n+1} = j \,|\, X_n = i, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_{n+1} = j \,|\, X_n = i) = p_{ij} \tag{3.1}$$

Then we say $(X_n)_{n\in\mathbb{N}_0}$ is a discrete-time Markov chain.

In equation 3.1, the Markov property is described using conditional probabilities on the history of the trajectory X_t. We assign the transition probability from i to j such that the history before arriving at i does not matter for this probability.

For continuous-time Markov chains, we similarly condition on the history of the trajectory. This time, we use a more general notation by conditioning on a sub-σ-algebra which will represent all possible histories up to a time s. Given the history of a trajectory X_t up until a time s, we will use the σ-algebra generated by the history $\{X_t \,|\, t \leq s\}$, which we call $\mathcal{F}_s = \sigma(\{X_t : t \leq s\})$.

Definition 3.2. (Continuous-time Markov chain) Let $(X_t)_{t\geq 0}$ be a stochastic process with a measurable state space S. If for any $s < t$ such that $t + s \geq 0$, and for any measurable subset A of S we have

$$\mathbb{P}(X_{t+s} \in A \,|\, \mathcal{F}_s) = \mathbb{P}(X_{t+s} \in A \,|\, X_s) \tag{3.2}$$

Then we say $(X_t)_{t\geq 0}$ is a continuous-time Markov chain.

Note 3.1. An equivalent requirement for equation 3.2 is if for any bounded measurable function f we have

$$\mathbb{E}[f(X_{t+s})|\mathcal{F}_s] = \mathbb{E}[f(X_{t+s})|X_s] \tag{3.3}$$

3.1.2 Markov transition matrix

In the discrete-time case, the probabilities p_{ij} can be seen as a probability matrix P for a finite state space S. We call matrix P the (Markov) transition matrix, as it contains the probability of transition from i to j for a single time step. If we wish to know the transition probability for multiple time steps, we can apply powers of matrix P. For example, $\mathbb{P}(X_{n+3} = j \,|\, X_n = i) = \mathbb{P}(X_{n+3} = j \,|\, X_n = i)(P^3)_{ij}$ due to the Markov property. This we prove in the following proposition.

Proposition 3.3. For any discrete-time Markov chain $(X_n)_{n\in\mathbb{N}_0}$ with state space S and for any $m \in \mathbb{N}_0$, the following holds

$$\mathbb{P}(X_{n+m} = j \,|\, X_n = i) = (P^m)_{ij} \tag{3.4}$$

If $(P^m)_{ij}$ is given by

$$(P^m)_{xz} = \sum_{y\in S} P_{xy}(P^{m-1})_{yz} \tag{3.5}$$

$$(P^0)_{xz} = \delta_{xz} \tag{3.6}$$

Proof. We will prove the statement by induction. The base case $m = 0$ is trivial. Assume $\mathbb{P}(X_{n+m-1} = k \,|\, X_n = i) = (P^{m-1})_{ij}$, then by $(P^m)_{xz} = \sum_y P_{xy} \cdot (P^{m-1})_{yz}$ we have

$$\begin{aligned}
\mathbb{P}(X_{n+m} = j \,|\, X_n = i) &= \sum_{s\in S} \mathbb{P}(X_{n+m} = j \,|\, X_{n+m-1} = s)\mathbb{P}(X_{n+m-1} = s \,|\, X_n = i) \\
&= \sum_{s\in S} P_{js} \cdot (P^{m-1})_{si} \\
&= (P^m)_{ji}
\end{aligned}$$

□

Note 3.2. Equation (3.5) is satisfied when S is finite and so P is a matrix.

In the continuous-time case we have a similar transition matrix $P(t)$ describing the transition probabilities going from state i to state j in time t. The family of matrices $(P(t))_{t\geq 0}$ should also satisfy $P(t+s) = P(t)P(s)$. This means we expect that $P(t)$ has to be of the form e^{tL} where L is called the generator matrix. We know that this property has to hold due to the Markov property, and can prove the matrix $P(t)$ has the form e^{tL} if we make this assumption. However, it is also possible to show $P(t)$ has the form e^{tL} if the Markov process has a fixed transition rate. The rate is a given parameter for how often a transition jump is done, in essence describing how fast a continuous-time Markov chain evolves in time. In the proposition below, we prove that the time between jumps is exponentially distributed with a parameter λ, describing the rate.

Proposition 3.4. (Jump rates are exponentially distributed) Let $x \in S$ be arbitrary. Define T_x as the time it takes to jump for the first time from x. Then T_x is exponentially distributed by some parameter $\lambda \in (0, \infty)$.

Proof. Applying the Markov property twice, we get the following result.

$$\begin{aligned}\mathbb{P}(T_x > t+s | T_x > s) &= \mathbb{P}(X_r = x, \forall r : s \le r \le t+s \,|\, X_u, \forall u : 0 \le u \le s) \\ &= \mathbb{P}(X_r = x, \forall r : s \le r \le t+s \,|\, X_s) \\ &= \mathbb{P}(T_x > t)\end{aligned}$$

In the last line, the Markov chain can be interpreted by shifting time $-s$. By conditional probabilities, we have $\mathbb{P}(T_x > t+s) = \mathbb{P}(T_x > t)\mathbb{P}(T_x > s)$. Hence $\mathbb{P}(T_x > t) \sim e^{-\lambda t}$ or $T_x \sim \text{Exp}(\lambda)$, where $\mathbb{E}[T_x] = \frac{1}{\lambda}$ is the average time between jumps so the parameter λ represents the rate of jumps. □

The parameter λ is allowed to depend on x. In the continuous-time case, we now have two variables which describe the process: the transition rate and the transition probability. We can combine these variables into what we call the **rate**, which can be thought of as the transition probability per unit of time in the limit of small time. For a countable state space S, we call $c(x,y) \ge 0$ the rate going from state $x \in S$ to state $y \in S$. We assume $c(x,x) = 0$. To describe the continuous-time Markov chain with a countable state space and rates c, we define the components of the generator L to be

$$L_{xy} = \begin{cases} c(x,y) & \text{if } x \neq y \\ -\sum_{s \in S} c(x,y), & \text{if } x = y \end{cases} \tag{3.7}$$

Note that the sum each row of L is equals to zero.

The process with rates c is then described as follows. When starting from x, the particle waits an exponential time with parameter

$$\lambda_x = \sum_{z \in S} c(x,z)$$

This λ_x is called the total exit rate from x. After this waiting time, the particle jumps to y with probability P_{xy}, given by

$$P_{xy} = \frac{c(x,y)}{\sum_{z \in S} c(x,z)}$$

One may think how we can find the transition rate and probability from only the rate $c(x,y)$. The probability of jumping from state x to state y is the ratio between the rate $c(x,y)$ and the total rate $\sum_{z \in S} c(x,z)$, in other words
Then, the transition rate is simply the sum of all outgoing rates from x.

The rate $c(x,y)$ can be intuitively understood as a 'measure' of how much a particle wants to travel from state x to a state y. For a single particle, the rate describes the probability transition per time. For example, if we consider a small time δt and assume the transition matrix has form e^{tL}, then $\mathbb{P}(X_{\delta t} = y \,|\, X_0 = x) = (e^{\delta t L})_{xy} \approx (1 + \delta t L)_{xy} = \delta t c(x,y)$ if $x \neq y$. As such, the unit of $c(x,y)$ is per unit time, for example s^{-1}. If considering multiple particles, $c(x,y)$ can be thought of as how many particles are being transferred from x to y per unit time.

In the following example, we consider a fixed total exit rate λ_x which does not depend on $x \in S$.

Example 3.1. *(Constant rate case)* If the total exit rate λ_x is constant (in other words, not depending on x and equal to λ), then the number of jumps N_t within time interval $[0,t]$ is Poisson distributed with rate λ. This means that the probability that n jumps have been made in the interval $[0,t]$ equals

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

This is due to the time between jumps T_x being exponentially distributed. We now define P to be the *time-independent* probability matrix, where P_{ij} is the probability going from $i \in S$ to $j \in S$. Connecting this to the discrete case, we consider how many jumps are necessary to get from state i to state j. Similar to Proposition 3.3, for n jumps we have the probability $(P^n)_{ij}$. If we let the parameter λ from Proposition 3.4

be fixed, we can derive the transition matrix. We can then find the continuous-time transition probability as follows.

$$\begin{aligned}
\mathbb{P}(X_t = j \,|\, X_0 = i) &= \sum_{n=0}^{\infty} \mathbb{P}(N_t = n)\mathbb{P}(X_t = j, N_t = n \,|\, X_0 = i) \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} (P^n)_{ij} \\
&= e^{-\lambda t} \left[\sum_{n=0}^{\infty} \frac{(\lambda t P)^n}{n!} \right]_{ij} \\
&= e^{\lambda t (P - I)}
\end{aligned}$$

Where I is the identity matrix. The generator matrix is then $L = \lambda(P - I)$, such that $\mathbb{P}(X_t = j \,|\, X_0 = i) = (e^{tL})_{ij}$. The matrix λP represents the rates to travel between states, so $c(x, y) = \lambda P_{xy}$. Since the sum of a row of P is 1, $-\lambda I$ will be a diagonal matrix having the value $-\sum_{s \in S} c(x, s)$ for row x.

Example 3.2. *(Random walk)* An example of a continuous-time Markov chain with constant rates is the continuous-time random walk on $\mathbb{Z}^d$, which jumps at rate one to each neighbor. For this process, the total exit rate is then $\lambda = 2d$. If the probability of jumping to each neighbor, we say this is a symmetric random walk. As an example of a symmetric random walk, let $S = \mathbb{Z}$ and define the rates

$$c(x, y) = \begin{cases} 1, & \text{if } y = x \pm 1, \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

While S is countable and therefore $L = \lambda(P - I) = 2d(P - I)$ cannot be described as a matrix, we can still model this process by exponentially distributed waiting times. Another option is to represent $\mathbb{Z}$ as a 'chain', for example $S = \{1, \ldots, N\}$ where similarly

$$c(x, y) = \begin{cases} 1, & \text{if } y = (x \pm 1) \mod N, \\ 0, & \text{otherwise} \end{cases}$$

This can be represented as a matrix such that e^{tL} can be computed numerically.

From now on, we only consider the rates $c(x, y)$ and change the definition of $P(t)$ to be the *time-dependent* transition matrix of a continuous-time Markov chain with finite state space. In other words, $(P(t))_{x,y} = \mathbb{P}(X_t = y \,|\, X_0 = x)$. We get a family of probability matrices $\{P(t) \,|\, t \geq 0\}$ instead of a single matrix. We now state three properties which $P(t)$ must satisfy:

1. $P(0) = I$, since at $t = 0$ there cannot have been a jump. Therefore $P(0)_{xy} = \delta_{xy}$ which is equivalent to $P(0) = I$.
2. For any $t \geq 0$, we have that the rows of $P(t)$ sum to 1. This requirement is due to normalization of probabilities.
3. For any $s, t \geq 0$ we have $P(t + s) = P(t)P(s)$ as defined by usual matrix multiplication. This is the continuous-time equivalent of Proposition 3.3.

From these properties, we will have e^{tL} as is shown next. We will assume that $P(t)$ is continuous and differentiable for all $t \geq 0$. It will appear that the derivative of $P(t)$ at $t = 0$ will be the same as L_{xy} as given in equation (3.7). We define L to be the derivative of $P(t)$ at $t = 0$.

$$L = \lim_{h \to 0} \frac{P(h) - P(0)}{h} \tag{3.9}$$

This is done element-wise, therefore $L_{xy} = \lim_{h \to 0} \frac{P(h)_{xy} - P(0)_{xy}}{h}$. In the following proposition, we derive $P'(t)$ for $t \neq 0$ in general.

Proposition 3.5. (Derivative of $P(t)$) Define the generator matrix L as per equation (3.9). Then we have

$$P'(t) = LP(t) = P(t)L \tag{3.10}$$

Proof. Using property (3) we have

$$\begin{aligned} P'(t) &= \lim_{h\to 0} \frac{P(t+h) - P(t)}{h} \\ &= \lim_{h\to 0} \frac{P(t)P(h) - P(t)}{h} \\ &= \left(\lim_{h\to 0} \frac{P(h) - I}{h}\right) P(t) \\ &= \left(\lim_{h\to 0} \frac{P(h) - P(0)}{h}\right) P(t) \end{aligned} \tag{3.11}$$

Last line of the above equation (equation (3.11)) gives $P'(t) = LP(t)$, switching $P(t)$ to the left will yield the result $P(t)L$. □

Corollary 3.6. The transition matrix $P(t)$ will have the following form

$$P(t) = e^{tL} \tag{3.12}$$

Proof. Applying property $\frac{d}{dt}e^{At} = Ae^{At}$ for matrix A to Proposition 3.5 yields the wanted result. □

To give a clear image on what has just been shown, we give an example which derives $P(t)$ from some matrix L.

Example 3.3. Consider the case where we have a continuous-time Markov chain with the generator

$$L = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \tag{3.13}$$

One way to find $P(t)$ is by Corollary 3.6 and using the series of e^{tL}. The diagonalization of L is

$$L = PDP^{-1} = \frac{1}{3}\begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}$$

After using the series expansion of e^x, we can take out P and P^{-1} out the sum. Since D is a diagonal matrix, we can find the D^n by simply applying the power to the diagonals. This gives the following result.

$$\begin{aligned} P(t) = e^{tL} &= \sum_{n=0}^{\infty} \frac{t^n L^n}{n!} \\ &= P\left(\sum_{n=0}^{\infty} \frac{t^n D^n}{n!}\right) P^{-1} \\ &= P\begin{pmatrix} e^{-3t} & 0 \\ 0 & 1 \end{pmatrix} P^{-1} \\ &= \frac{1}{3}\begin{pmatrix} 2 + e^{-3t} & 1 - e^{-3t} \\ 2 - 2e^{-3t} & 2e^{-3t} + 1 \end{pmatrix} \end{aligned}$$

We can now see explicitly that all three properties are satisfied. For the first property,

$$P(0) = \frac{1}{3}\begin{pmatrix} 2+1 & 1-1 \\ 2-2 & 2+1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

For the second property, the first row sums to $\frac{1}{3}(2+e^{-3t}+1-e^{-3t})=1$ and the second row has sum $\frac{1}{3}(2-2e^{-3t}+2e^{-3t}+1)=1$. For the last property, we have

$$\begin{aligned}
P(t)P(s) &= \frac{1}{9}\begin{pmatrix} 2+e^{-3t} & 1-e^{-3t} \\ 2-2e^{-3t} & 2e^{-3t}+1 \end{pmatrix}\begin{pmatrix} 2+e^{-3s} & 1-e^{-3s} \\ 2-2e^{-3s} & 2e^{-3s}+1 \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} (2+e^{-3t})(2+e^{-3s})+(1-e^{-3t})(2-2e^{-3s}) & (2+e^{-3t})(1-e^{-3s})+(1-e^{-3t})(2e^{-3s}+1) \\ (2-2e^{-3t})(2+e^{-3s})+(2e^{-3t}+1)(2-2e^{-3s}) & (2-2e^{-3t})(2+e^{-3s})+(2e^{-3t}+1)(2-2e^{-3s}) \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 6+3e^{-3(t+s)} & 3-3e^{-3(t+s)} \\ 6-6e^{-3(t+s)} & 3+6e^{-3(t+s)} \end{pmatrix} \\
&= \frac{1}{3}\begin{pmatrix} 2+e^{-3(t+s)} & 1-e^{-3(t+s)} \\ 2-2e^{-3(t+s)} & 1+2e^{-3(t+s)} \end{pmatrix} \\
&= P(t+s)
\end{aligned}$$

This then shows that we indeed satisfy property 3.

3.2 Semigroups and generators

In the previous section, we saw that the transition matrix $P(t)$ has the form e^{tL}, where L is the generator matrix, due to the property $P(t+s)=P(t)P(s)$. To recall, we have $\mathbb{P}(X_t=y\,|\,X_0=x)=(P(t))_{xy}$. Let us continue with the finite state space, so we may represent $P(t)$ as a matrix. Then if we wish to calculate the expected value of some bounded function $f:S\to\mathbb{R}$ given we start at $X_0=x$, we may express this as

$$\mathbb{E}[f(X_t)\,|\,X_0=x]=\sum_{y\in S}f(y)\mathbb{P}(X_t=y\,|X_0=x)=\sum_{y\in S}(P(t))_{xy}f(y) \tag{3.14}$$

Let $S=\{x_1,\dots,x_n\}$ be the state space. Let $\vec{f}$ denote the vectorized form of the function f in the sense of $\vec{f}=(f(x_1),\dots,f(x_n))^T$ where similarly $\mathbb{P}(X_t=x_j\,|\,X_0=x_i)=(P(t))_{i,j}=p_t(x_i,x_j)$. Then we can express equation (3.14) using $\vec{f}$ and $P(t)$ as follows.

$$P(t)\vec{f}=P(t)\begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n)\end{pmatrix}=\begin{pmatrix} \sum_{i=1}^n p_t(x_1,x_i)f(x_i) \\ \vdots \\ \sum_{i=1}^n p_t(x_n,x_i)f(x_i)\end{pmatrix}=\begin{pmatrix} \mathbb{E}[f(X_t)\,|\,X_0=x_1] \\ \vdots \\ \mathbb{E}[f(X_t)\,|\,X_0=x_n]\end{pmatrix}$$

Component-wise we have $(P(t)\vec{f})_i=\mathbb{E}[f(X_t)\,|\,X_0=x_i]$.

In general, we will work on state spaces which are infinite. As such, we define the semigroup as an operator working on bounded measurable functions which evaluate the expected value of f at the trajectory X_t given we start at some initial position x. We give the definition of the semigroup below.

Definition 3.7. (Semigroup S_t) The semigroup S_t works on a bounded measurable function f as follows.

$$(S_tf)(x)=\mathbb{E}[f(X_t)\,|\,X_0=x]=\sum_{y\in S}f(y)\mathbb{P}(X_t=y\,|\,X_0=x) \tag{3.15}$$

In essence, S_t can be thought of being $P(t)$ in matrix form. Similarly, we have the same properties which $P(t)$ has, most notably $P(t+s)=P(t)P(s)$, for the semigroup S_t. Most often we are interested in the properties of S_t instead of its actual definition, therefore we introduce a new definition of a Markov semigroup with similar properties.

Definition 3.8. (Markov Semigroup) Let S_t be an operator working on bounded measurable functions. If the following properties are satisfied, we state that S_t is a **(Markov) semigroup**.

(S1) Semigroup property: $\forall s,t>0$:

$$S_{t+s}f=S_t(S_sf)=S_s(S_tf)$$

(S2) Identity at time zero: $S_0 = I$.

(S3) Right continuity: map $t \mapsto S_t f$ is right continuous.

(S4) Positivity: $f \geq 0$ implies $S_t f \geq 0$

(S5) Normalization: $S_t 1 = 1$

(S6) Contraction: $\max_x |(S_t f)(x)| \leq \max_x |f(x)|$

Note 3.3. To show the connection between Definitions 3.7 and 3.8, we check that the properties (S1)-(S6) are satisfied by the definition $(S_t f)(x) = \mathbb{E}[f(X_t) \,|\, X_0 = x]$. For property (S1), we remark that

$$(S_t(S_s f))(x) = \mathbb{E}[(S_s f)(X_t) \,|\, X_0 = x] = \mathbb{E}[\mathbb{E}[f(Y_s) \,|\, Y_0 = X_t] \,|\, X_0 = x]$$

By the Markov property, we may view Y_s as a continuation of X_t up to a time $t+s$, in other words $Y_s = X_{t+s}$. This makes clear that property (S1) is satisfied. For (S2), by definition $X_0 = x$, so $(S_0 f)(x) = \mathbb{E}[f(X_0) \,|\, X_0 = x] = f(x)$, so we say S_0 is the identity function. For (S3), we know by assumption that $P(t)$ is continuous. Property (S4) is also satisfied, since if $f \geq 0$ then $f(y)\mathbb{P}(X_t = y \,|\, X_0 = x) \geq 0$ so $(S_t f)(x) \geq 0$. For (S5), if we choose $f(x) = 1$ then we have

$$(S_t 1)(x) = \sum_{y \in S} 1 \cdot \mathbb{P}(X_t = y \,|\, X_0 = x) = 1$$

Since all probabilities here must add to 1. And for property (S6), let $M = \max_x |f(x)|$ denote the upper bound. Then we have

$$\begin{aligned}
|(S_t f)(x)| &= \left| \sum_{y \in S} f(y)\mathbb{P}(X_t = y \,|\, X_0 = x) \right| \\
&\leq \sum_{y \in S} |f(y) \cdot \mathbb{P}(X_t = y \,|\, X_0 = x)| \\
&\leq \sum_{y \in S} |f(y)| \cdot \mathbb{P}(X_t = y \,|\, X_0 = x) \\
&\leq \sum_{y \in S} \max_x |f(x)| \cdot \mathbb{P}(X_t = y \,|\, X_0 = x) \\
&\leq \max_x |f(x)|
\end{aligned}$$

Therefore taking the maximum on $|(S_t f)(x)|$ also gives the upper bound $\max_x |f(x)|$. In conclusion, we have that (S1)-(S6) are satisfied given the Definition 3.7.

Again, due to these properties we expect the semigroup S_t to have the form e^{tL} where L is some operator we will call the **generator** of the semigroup S_t. The generator can be found by evaluating the series of e^{tL} to the first order, which is $1 + tL + \frac{1}{2}t^2L^2 + \mathcal{O}(t^3)$. If we subtract 1 and then divide by t, we have $L + \frac{1}{2}tL^2 + \mathcal{O}(t^2)$. If we then take the limit $t \to 0$, we have found L as follows.

$$(Lf)(x) = \lim_{t \to 0} \frac{(S_t f)(x) - f(x)}{t} \tag{3.16}$$

As an example, we will find the generator of the semigroup S_t.

Example 3.4. *(Generator of semigroup S_t)* If we assume to have a finite state space, we can rewrite the probability to use $(e^{tL})_{xy}$.

$$\begin{aligned}(S_tf)(x) - f(x) &= \sum_{y\in S} f(y)\mathbb{P}(X_t = y \,|\, X_0 = x) - f(x)\\ &= \sum_{y\in S} f(y)\mathbb{P}(X_t = y \,|\, X_0 = x) - \sum_{y\in S} f(x)\mathbb{P}(X_t = y \,|\, X_0 = x)\\ &= \sum_{y\in S}[f(y) - f(x)]\mathbb{P}(X_t = y \,|\, X_0 = x)\\ &= \sum_{y\in S}[f(y) - f(x)](e^{tL})_{xy}\end{aligned}$$

Now, only e^{tL} depends on t. Therefore if we divide by t and take the limit $t \to 0$, we should obtain L. Since $(e^{tL})_{xy} = \sum_{n=0}^{\infty} \left(\frac{t^n L^n}{n!}\right)_{xy} = \sum_{n=0}^{\infty} \frac{t^n (L^n)_{xy}}{n!}$, we obtain similarly L_{xy} for $\lim_{t\to 0} \frac{1}{t}((S_tf)(x) - f(x))$, so we have

$$(Lf)(x) = \sum_{y\in S}[f(y) - f(x)]L_{xy} = \sum_{y\in S} c(x,y)[f(y) - f(x)] \tag{3.17}$$

Where L_{xx} is ignored due to the $f(y) - f(x)$ term.
Another method is by considering how many jumps the process takes. We know from Proposition 3.4 that the jump time is exponentially distributed and so the amount of jumps is Poisson distributed. Let λ be the total outgoing rate from x, so $\lambda = \sum_{y\in S} c(x,y)$. Let N_t denote the amount of jumps in the interval $[0,t]$. The random variable N_t then has parameter λt. We evaluate $\mathbb{P}(N_t = 0) = e^{-\lambda t}$, $\mathbb{P}(N_t = 1) = \lambda t e^{-\lambda t}$ and finally $\mathbb{P}(N_t = 2) = \frac{(\lambda t)^2}{2} e^{-\lambda t}$. Since we divide by t eventually, we denote the $N_t \geq 2$ case by a bound $\mathcal{O}(t^2)$ which will vanish. Also, we can write $\mathbb{P}(N_t = 1) = 1 - \mathbb{P}(N_t = 0) - \mathbb{P}(N_t \geq 2) = 1 - e^{-\lambda t} + \mathcal{O}(t^2)$ instead of $\lambda t e^{-\lambda t}$ as this simplifies the derivation. Conditioning on N_t gives

$$\begin{aligned}(S_tf)(x) - f(x) &= \mathbb{E}[f(X_t) \,|\, X_0 = x, N_t = 0]\mathbb{P}(N_t = 0) + \mathbb{E}[f(X_t) \,|\, X_0 = x, N_t = 1]\mathbb{P}(N_t = 1) + \mathcal{O}(t^2) - f(x)\\ &= f(x)e^{-\lambda t} + (1 - e^{-\lambda t})\sum_{y\in S} f(y)\cdot p(x,y) - f(x) + \mathcal{O}(t^2)\end{aligned}$$

The probability of going from x to y in one jump is $c(x,y)/\lambda$. Substituting this gives

$$\begin{aligned}(S_tf)(x) - f(x) &= (1 - e^{-\lambda t})\sum_{y\in S} \frac{c(x,y)}{\lambda} f(y) + (1 - e^{-\lambda t})f(x) + \mathcal{O}(t^2)\\ &= (1 - e^{-\lambda t})\sum_{y\in S} \frac{c(x,y)}{\lambda}[f(y) - f(x)] + \mathcal{O}(t^2)\end{aligned}$$

If we divide by t and let $t \to 0$, the term $\mathcal{O}(t^2)$ vanishes. However, $\lim_{t\to 0} \frac{1}{t}(1 - e^{-\lambda t}) = \lim_{t\to 0} \frac{1}{t}(\lambda t + \mathcal{O}(t^2)) = \lambda$. Therefore, we have found the generator L as

$$(Lf)(x) = \lambda \sum_{y\in S} \frac{c(x,y)}{\lambda}[f(y) - f(x)] = \sum_{y\in S} c(x,y)[f(y) - f(x)] \tag{3.18}$$

Now that we have introduced semigroups, we will show that the Kolmogorov (forward and) backwards equations hold. This statement will be useful when working with the Feynman-Kac formula as will be shown later.

Theorem 3.9. (Kolmogorov's backward equation) Let S_t be a semigroup with L the generator as defined by equation (3.15). If f is a bounded measurable function, we then have the following property:

$$\frac{d}{dt}(S_t f(x)) = L(S_t f(x)) = LS_t f(x) \tag{3.19}$$

Proof. By the definition of the generator, $S_t f$ can be linearized with respect to t to obtain tLf, such that $\frac{d}{dt}(S_t f)|_{t=0} = Lf$. If we shift the time by s (given that $t - s \geq 0$), we get $\frac{d}{dt}(S_{t-s} f)|_{t=s} = Lf$. Substituting $f = S_s g$ yields

$$\frac{d}{dt}(S_{t-s} S_s g)|_{t=s} = \frac{d}{dt}(S_t g)|_{t=s} = \frac{d}{ds}(S_s g) = LS_s g$$

□

We remark that equation (3.19) represents the ordinary differential equation $y'(t) = Ly(t)$ which shares the same solution $y(t) = C \cdot e^{tL}$ similar to $S_t = e^{tL}$ ($C = 1$ due to property (S2)). Also, from $S_t = e^{tL} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n$, we also have that the operators S_t and L commute.

Depending on what is chosen for S_t, we can intuitively view the semigroup as the operator that evolves some function f. For instance, let $g(t, x) = S_t f(x)$, then $g(0, x) = f(x)$ is the initial condition. Now, $g(t, x)$ evolves $f(x)$ for $t > 0$ given a semigroup S_t and its defining character governed by the generator L. We will see in the upcoming section that Kolmogorov's backward equation can used to write

$$\frac{\partial}{\partial t} g(t, x) = Lg(t, x) \tag{3.20}$$

3.3 Feynman-Kac formula

From Kolmogorov's backward equation (Theorem 3.9), if we substitute $g(t, x) = S_t f(x)$ we get $\frac{\partial}{\partial t} g(t, x) = Lg(t, x)$ as stated in equation (3.20). If a function f is given, then g is easily derived from $g = S_t f$. However, it is more interesting to solve for g given the equation (3.20). First, $S_0 = I$ thus $f(x) = S_0 f(x) = g(0, x)$. Then, by the definition of S_t, we get

$$g(t, x) = (S_t f)(x) = \mathbb{E}[f(X_t) \mid X_0 = x] = \mathbb{E}_x[g(0, X_t)] \tag{3.21}$$

Where $\mathbb{E}_x[\cdot]$ is the expectation for a path X_t starting at x, in other words $\mathbb{E}_x[\cdot] = \mathbb{E}[\cdot \mid X_t = x]$

From this, we find that g can be found as the expectation of different paths given the initial conditions. Now, we will derive a similar result but for a generalized semigroup with generator $M = L + V$, where L is the generator of the semigroup S_t and $V : S \to \mathbb{R}$ is a multiplicative operator, in other words $Vh = V(x)h(t, x)$. This result is named the Feynman-Kac formula and is given below.

Theorem 3.10. (Feynman-Kac) Let $V : \mathbb{R} \to \mathbb{R}$ be bounded and measurable. If we assume that g satisfies the following partial differential equation

$$\frac{\partial}{\partial t} g(t, x) = (L + V)g(t, x) = Lg(t, x) + V(x)g(t, x) \tag{3.22}$$

Where L is the generator of a continuous-time Markov process, then g admits the following solution.

$$g(t, x) = \mathbb{E}_x \left[g(0, X_t) e^{\int_0^t V(X_s)\, \mathrm{d}s} \right] \tag{3.23}$$

Proof. We will give a sketch of proof of this. Define $f(x) = g(0, x)$ and let $M = L + V$. Let $g(t, x) = K_t f(x)$ be the time evolution of f where K_t is a relaxed semigroup satisfying (S1)-(S3). We define K_t as

$$(K_t f)(x) = \mathbb{E}_x \left[f(X_t) e^{\int_0^t V(X_s)\, \mathrm{d}s} \right] \tag{3.24}$$

We will prove that K_t satisfies (S1)-(S3) having the generator M. After this has been shown, we can apply Kolmogorov's backward equation[4] to obtain

[4]Showing (S1)-(S3) is sufficient enough instead of (S1)-(S6) for Kolmogorov's backward equation to hold.

$$\frac{\partial}{\partial t}g(t,x) = (L+V)g(t,x)$$

To fully prove the solution of equation (3.22) is given by (3.23), one may for instance show that g is unique. *(Generator)* We use the definition of the generator, as given by

$$\lim_{t\to 0}\frac{K_t f - f}{t} = Mf$$

First, we use the series expansion of $e^{\int_0^t V(X_s)\,\mathrm{d}s} = 1 + \int_0^t V(X_s)\,\mathrm{d}s + \mathcal{O}(t^2)$ where we assume $\int_0^t V(X_s)\,\mathrm{d}s$ has a term t (otherwise it will be zero when divided by t). Substituting equation (3.23) into the generator definition yields

$$\begin{aligned}\lim_{t\to 0}\frac{K_t f(x) - f(x)}{t} &= \lim_{t\to 0}\frac{1}{t}\left(\mathbb{E}_x\left[f(X_t)(1+\int_0^t V(X_s)\,\mathrm{d}s + \mathcal{O}(t^2))\right] - f(x)\right)\\ &= \lim_{t\to 0}\frac{\mathbb{E}_x[f(X_t)] - f(x)}{t} + \lim_{t\to 0}\frac{\mathbb{E}_x\left[f(X_t)\int_0^t V(X_s)\,\mathrm{d}s\right]}{t}\end{aligned}$$

The first part $\lim_{t\to 0}\frac{1}{t}(\mathbb{E}_x[f(X_t)] - f(x))$ is by definition the semigroup S_t since we defined L this way. For the second part, we will first approximate the integral using a right Riemann sum and then by linearity acquire the potential V. We approximate the integral $\int_0^t V(X_s)\,\mathrm{d}s = t\cdot V(X_t) + \mathcal{O}(t^2)$, note that the error is $\mathcal{O}(t^2)$ as the error of the right Riemann sum is given by $\left|\int_a^b f\,\mathrm{d}x - S\right| \le M(b-a)^2$. This error will vanish due to $\lim_{t\to 0}\frac{1}{t}$. Thus, we have the term $\lim_{t\to 0}\frac{1}{t}\mathbb{E}_x[t\cdot V(X_t)f(X_t)]$. Now assume that we may take the limit into the integral. For example, if the potential is bounded then the dominated convergence theorem can be applied: $\mathbb{E}_x\lim_{t\to 0}(Vf)(X_t) = (Vf)(x)$. Finally, the limit simplifies to $(Lf)(x)+(Vf)(x) = (L+V)f = Mf$, thus we have shown that this does indeed have the correct generator.

(Property S1) We will show that $K_{t+s}f = K_tK_sf$ using the Markov property. We can write $K_{t+s}f = \mathbb{E}_x[f(X_{t+s})e^{\int_0^{t+s}V(X_p)\,\mathrm{d}p}]$. For K_sK_tf, we get:

$$\begin{aligned}K_sK_tf &= K_s\mathbb{E}[f(X_t)e^{\int_0^t V(X_r)\,\mathrm{d}r}\,|\,X_0 = x]\\ &= \mathbb{E}\{\mathbb{E}[f(X_t)e^{\int_0^t V(X_r)\,\mathrm{d}r}\,|\,X_0 = Y_s]e^{\int_0^s V(Y_r)\,\mathrm{d}r}\,|\,Y_0 = x\}\\ &= \mathbb{E}\{f(Y_{t+s})e^{\int_0^t V(Y_{r+s})\,\mathrm{d}r}e^{\int_0^s V(Y_r)\,\mathrm{d}r}\,|\,Y_0 = x\}\\ &= \mathbb{E}\{f(Y_{t+s})e^{\int_0^{t+s} V(Y_r)\,\mathrm{d}r}\,|\,Y_0 = x\}\\ &= K_{t+s}f\end{aligned}$$

We used the Markov property to combine the expected values and used a time shift to sum the integrals together.

(Property S2) Fixing $t = 0$ gives

$$(K_0f)(x) = \mathbb{E}\left[f(X_0)e^{\int_0^0 V(X_s)\,\mathrm{d}s}\,|\,X_0 = x\right] = \mathbb{E}[f(X_0)\,|\,X_0 = x] = f(x) \tag{3.25}$$

So indeed K_0 is the identity operator.

(Property S3) We know that $h(t) = \int_0^t V(X_s)\,\mathrm{d}s$ is continuous, and so $\lim_{t\downarrow 0}h(t) = h(0) = 0$. Therefore $\exp(\int_0^t V(X_s)\,\mathrm{d}s) = \exp(h(t))$ has $\lim_{t\downarrow 0}\exp(h(t)) = \exp(0) = 1$ since exp is a continuous function. Let $m(t,X_t) = f(X_t)\exp(\int_0^t V(X_s)\,\mathrm{d}s)$. Since g is bounded, f must also be bounded, for example $|f(x)| \le C$ for all $x\in\mathbb{R}$. Similarly, assuming that V is also a bounded function, we have $\left|\int_0^t V(X_s)\,\mathrm{d}s\right| \le \int_0^t |V(X_s)|\,\mathrm{d}s \le Mt$ where $|V(x)| \le M$ for all $x\in\mathbb{R}$. By monotonicity of the function exp, we have

$$\exp\left(\int_0^t V(X_s)\,\mathrm{d}s\right) \le \exp\left(\left|\int_0^t V(X_s)\,\mathrm{d}s\right|\right) \le \exp(Mt)$$

We then have that $m(t, X_t) = f(X_t)\exp(\int_0^t V(X_s)\,\mathrm{d}s)$ must also be bounded from above by $C\exp(Mt)$. Since f and $C\exp(Mt)$ are measurable functions, so is $m(t, X_t)$. Also, $C\exp(Mt)$ does not depend on X_t and is finite, therefore it is integrable with respect to the probability measure of the Markov process. Then, by the dominated convergence theorem, we have

$$\lim_{t\downarrow 0} \mathbb{E}_x[m(t, X_t)] = \mathbb{E}_x\left[\lim_{t\downarrow 0} m(t, X_t)\right]$$

We assume the limit $\lim_{t\downarrow 0} f(X_t)$ exists, then since $\lim_{t\downarrow 0}\exp(\int_0^t V(X_s)\,\mathrm{d}s) = 1$, we can split the limit into two parts.

$$\lim_{t\downarrow 0} \mathbb{E}_x[m(t, X_t)] = \mathbb{E}_x\left[\lim_{t\downarrow 0} f(X_t)\cdot 1\right] = (S_t f)(x)$$

Since S_t also satisfies (S3), then K_t must also satisfy (S3). □

Note 3.4. The fact that we have an expected value in equation (3.23) can be understood from the case $V = 0$, which is

$$\frac{\partial}{\partial t} g(t, x) = Lg(t, x) \tag{3.26}$$

$$g(t, x) = \mathbb{E}_x[g(0, X_t)\cdot \exp(0)] = \mathbb{E}_x[g(0, X_t)] \tag{3.27}$$

which is identical to equation (3.20) where we stated $g(t, x) = S_t g(0, x)$. An exponential term $\exp(\int_0^t V(X_s)\,\mathrm{d}s)$ is added to compensate for the potential.

Although Feynman-Kac formula describes the full solution of a partial differential equation with a random potential, it is most often used for a theoretical approach for solving this type of differential equation. It can however be used for numerical simulations, by taking a certain initial condition $f(x)$ we may obtain $g(t, x)$ by simulating multiple random walks as described by L and taking the mean of $f(x)e^{\int_0^t V(X_p)\,\mathrm{d}p}$ over an ensemble of random walks.

3.4 Brownian motion

3.4.1 Definition and generator

This section will discuss a continuous-time Markov process with a state space $\mathbb{R}^d$. Let us consider the random walk given in Example 3.2, which described a continuous-time Markov chain with state space $\mathbb{Z}^d$ and rates $c(x, y) = 1$ if $y = x \pm 1$ and zero otherwise. As a result, the generator L of this process is

$$Lf(x) = (f(x+1) + f(x-1)) - 2f(x) \tag{3.28}$$

One might note that this is similar to the difference scheme for the second order derivative.

$$f''(x) = \lim_{h\to 0} \frac{f(x-h) - 2f(x) + f(x-h)}{h^2}$$

Of course, since the state space is $\mathbb{Z}^d$ we cannot take $h \to 0$ since there are no points between x and $x+1$. However, this does suggest that there might be an equivalent to the random walk with the same behavior having state space $\mathbb{R}^d$.

Let us assume for now that such a process exists, which we denote by $(B_t)_{t\geq 0}$ taking values in $\mathbb{R}$ and indexed by $[0, \infty)$ representing time. If we would construct such a process with discrete timesteps Δt and discrete spatial steps Δx and then taking the limits $\Delta t, \Delta x \to 0$, the distribution would converge to a normal distribution as per the central limit theorem (see [Sch21]). This will motivate the fact that we will use the normal distribution in the definition of Brownian motion. However, similarly to that of the random walker, we expect the future of Brownian motion to be independent of its history, due to the Markov property. So, we can describe the 'jumps' of Brownian motion to be normally distributed, such that we describe this process only by certain time intervals s to t. We give its definition below.

Definition 3.11. (Brownian motion, one dimensional) Let $B = (B_t)_{t\geq 0}$ take values in $\mathbb{R}$ being indexed by $[0, \infty)$ representing time. Then if the following requirements are satisfied, B is Brownian motion.

1. $B_0 = 0$, in other words the process starts at position 0.
2. The increments of B between times $0 \leq s < t$ is normally distributed with mean 0 and variance $t - s$, in other words

$$B_t - B_s \sim \mathcal{N}(0, t - s) \tag{3.29}$$

3. The increments $B_t - B_s$ are independent for different time intervals. In other words, for $0 \leq t_0 < t_1 < \cdots < t_n$ the following increments are independent of each other:

$$B_{t_0} - B_{t_1}, B_{t_2} - B_{t_3}, \ldots, B_{t_{n-1}} - B_{t_n} \tag{3.30}$$

4. The function B_t is continuous in time.

Definition 3.12. (Brownian motion, d dimensions) Let W_t^i be one-dimensional Brownian motion and let this be independent for $i = 1, \ldots, d$. Then d-dimensional Brownian motion is described by d individual Brownian motion processes, in other words $B_t = (W_t^1, \ldots, W_t^d)^T$. We retain the same properties as given by Definition (3.11), but property (2) now is given by

$$B_t - B_s \sim \mathcal{N}(0, t - s)^d \tag{3.31}$$

Where $\mathcal{N}(0, t-s)^d$ describes the normal distribution in d dimensions which are pairwise independent of each other.

Given the definition of Brownian motion above, we are now interested in the generator as stated at the beginning of the section. For this, we use equation (3.16) from the previous section.

Proposition 3.13. (Generator Brownian motion) The generator for Brownian motion as described by Definition 3.11, for functions $f \in C_0^2(\mathbb{R})$ is given by

$$(Lf)(x) = \frac{1}{2} f''(x) = \frac{1}{2} \frac{d^2 f(x)}{dx^2} \tag{3.32}$$

Proof. First, we note that $S_t f$ requires the expected value from a process X_t where $X_0 = x$. Then $X_t = B_t + x$ where we shift the Brownian motion process by x, since $B_0 = 0$. We keep in mind that $B_t = B_t - B_0 \sim \mathcal{N}(0, t)$. Now, using equation (3.16) and (3.15),

$$\begin{aligned}(Lf)(x) &= \lim_{t\to 0} \frac{\mathbb{E}[f(X_t) \,|\, X_0 = x] - f(x)}{t} \\ &= \lim_{t\to 0} \frac{\mathbb{E}[f(B_t + x) - f(x)]}{t}\end{aligned}$$

Now, using a Taylor expansion on f around x gives

$$f(x + B_t) = f(x) + B_t f'(x) + \frac{1}{2} B_t^2 f''(x) + \mathcal{O}(B_t^3)$$

Using $\mathbb{E}[B_t] = 0$ and $\mathbb{E}[B_t^2] = t$, we have

$$\begin{aligned}(Lf)(x) &= \lim_{t\to 0} \frac{\mathbb{E}[f(x) + B_t f(x) + \frac{1}{2} B_t^2 f''(x) + \mathcal{O}(B_t^3) - f(x)]}{t} \\ &= \lim_{t\to 0} \frac{\mathbb{E}[\frac{1}{2} B_t^2 f''(x) + \mathcal{O}(B_t^3)]}{t} \\ &= \frac{1}{2} f''(x) + \lim_{t\to 0} \frac{\mathbb{E}[\mathcal{O}(B_t^3)]}{t}\end{aligned}$$

For $\mathcal{O}(B_t^3)$, consider the nth moment of B_t, which we calculate using the integral

$$\begin{aligned}
\mathbb{E}[B_t^n] &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2t}} \, \mathrm{d}x, && \text{let } u = \frac{1}{\sqrt{2t}} x \\
&= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} u^n \sqrt{2t}^n e^{-u^2} \cdot \sqrt{2t} \, \mathrm{d}u, && \mathrm{d}u = \frac{1}{\sqrt{2t}} \, \mathrm{d}x \\
&= \sqrt{2t}^n \cdot \underbrace{\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} u^n e^{-u^2} \, \mathrm{d}u}_{\text{does not depend on } t}
\end{aligned}$$

For $n \geq 3$, we have the following limit

$$\lim_{t \to 0} \frac{(2t)^{n/2}}{t} = 0$$

Since in the numerator we always have a greater power than in the denominator. Then, the remainder term $\lim_{t \to 0} \mathbb{E}[\mathcal{O}(B_t^3)]/t$ vanishes. As a consequence, we obtain that the action of the generator of Brownian motion is given by $(Lf)(x) = \frac{1}{2} f''(x)$. □

Note 3.5. We require $f \in C_0^2(\mathbb{R})$ since f we use a Taylor expansion to the second order and require that $\mathbb{E}[f(x)] < \infty$ which holds if and only if $f \in C_0(\mathbb{R})$ (f vanishes at infinity).

Note 3.6. *(Higher dimensional)* For d-dimensional Brownian motion, it can be shown (Example 7.9 of [Sch21]) that the generator is

$$L = \frac{1}{2} \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \tag{3.33}$$

For example, $d = 3$ gives $L = \frac{1}{2} \nabla^2$.

3.4.2 Donsker's theorem

It is clear that the random walk and Brownian motion are similar in terms of generator. We are interested if we can 'rescale' the random walk such that we obtain Brownian motion. Let $(\varepsilon_k)_{k \geq 0}$ be a sequence of i.i.d. Bernoulli random variables such that $\mathbb{P}(\varepsilon_k = -1) = \mathbb{P}(\varepsilon_k = 1) = \frac{1}{2}$. Define the partial sum S_n of the sequence $(\varepsilon_k)_{k \geq 0}$ as follows.

$$S_n = \sum_{k=1}^{n} \varepsilon_k \tag{3.34}$$

Let us now consider $t \in [0, 1]$ which will represent the scaling from 0 to n (such that the time is nt). To acquire a continuous function of t, we interpolate between $S_{\lfloor nt \rfloor}$ and $S_{\lfloor nt \rfloor + 1}$. Then, we scale by $1/\sqrt{n}$ similar to that of $(X - \mu)/\sqrt{\sigma^2} \to N(\mu, \sigma^2)$ in distribution. This yields

$$X_n(t) = \frac{1}{\sqrt{n}} \left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) \varepsilon_{\lfloor nt \rfloor + 1} \right) \tag{3.35}$$

With this given, we now present Donsker's theorem below.

Theorem 3.14. (Donsker's Theorem) Let S_n be the partial sum, given by equation (3.34), of a sequence of i.i.d Bernoulli random variables as stated above. Define the rescaled version as $X_n(t)$, given by equation (3.35). Then we have

$$X_n(t) \underset{n \to \infty}{\longrightarrow} B_t, \text{ in distribution, for } t \in [0, 1] \tag{3.36}$$

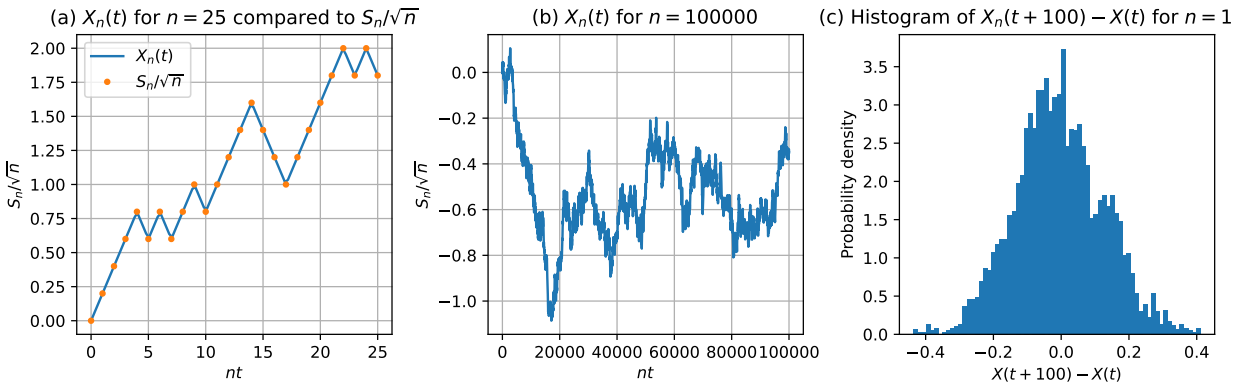


Figure 6: Illustration of Donsker's theorem by means of a simulation of n Bernoulli random variables. Part (a) on the left side illustrates the transformation of $X_n(t)$ as given by equation (3.35). Part (b) in the middle is a sample of $X_n(t)$ for $n = 10^6$. Part (c) on the right is a histogram of $X_n(t+100) - X_n(t)$ acquired from part (b). As $X_n(t) \to B_t$ per Donsker's theorem, we expect part (c) to be similar to a normal distribution. We refer to Example 3.5 for details.

Proof. For proof of Donsker's Theorem we refer to [Sch21]. We remark the discrepancy between the literature and the definition of $X_n(t)$ by $\pm(nt - \lfloor nt \rfloor)$, the key requirement is that $X_n(t)$ is a piecewise continuous function equal to $S_j/\sqrt{n}$ if $t = j/n$. □

Example 3.5. *(Simulation example)* To illustrate the concept of Donsker's theorem, we can sample some n Bernoulli random variables by means of simulation, and then calculate S_n and $X_n(t)$ for different t (and thus nt). We will refer to the simulation results in Figure 6. We can numerically understand the transformation of equation (3.35) by examining part (a) of said figure. We linearly interpolate between the points of $S_n/\sqrt{n}$. If we then increase n to 10^6 as in part (b), the piecewise part becomes unobservable. As one of the properties of Brownian motion, we expect that $B_t - B_s \sim \mathcal{N}(0, t-s)$. This is calculated for $t - s = 100$ for different positions of part (b), given in part (c). As expected, we see a distribution somewhat similar to the normal distribution.

4 Large deviation theory

The theory of large deviations deals with exponentially small probabilities, the parameters describing the distribution are the so called rate and rate function. This theory is applicable to many fields, such as statistics, statistical physics, and queuing theory [Tou11]. In this section, we will work towards applying the large deviation theory for continuous-time Markov processes, in particular the local and occupation time. First, we give an introduction to large deviation theory for sums of i.i.d. random variables and give examples for normally and exponentially distributed variables. In combination with the examples, we will state the Cramér's theorem. In the second section, we expand the notion of sums of i.i.d. random variables to families of random variables. Similar to Cramér's theorem, we end the second section with the Gärtner-Ellis theorem. This will be necessary since Markov processes are not independent of each other, as will be discussed in the third section. In this last section, we apply a similar statement to Gärtner-Ellis for the occupation times measure of a continuous-time Markov process, which will be used in the delocalized model.

Most definitions and theorems regarding large deviation theory in this section are derived from [DZ10] and [Den08]. Furthermore, we refer to these sources for proof of these theorems.

4.1 Sums of independent and identically distributed random variables

4.1.1 Introduction

A basic example of random variables with exponentially small probability densities are certain sums of i.i.d. random variables. Let $X_1, \dots, X_n$ be n i.i.d. random variables, then the sample mean of these random variables, denoted $\overline{X}_n$, is defined as

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{4.1}$$

We also define the partial sums $S_n = \sum_{i=1}^{n} X_i$ such that $\overline{X}_n = \frac{S_n}{n}$.

To find the probability density function (pdf) of $\overline{X}_n$, defined as $p_{\overline{X}_n}(s)$, we first fix $X_1 + \cdots + X_n = ns$. As an example, let $n = 2$ and let X_1, X_2 be discrete independent random variables, such that we have $X_1 + X_2 = ns$. The probability of this event is $\mathbb{P}(X_1 + X_2 = ns) = \sum_x \mathbb{P}(X_1 = x)\mathbb{P}(X_2 = ns - x)$. As more variables are introduced, we have to sum over multiple variables, or integrate if we are dealing with continuous random variables.

An easier way to calculate the sample mean $\overline{X}_n$ from the samples $X_1, \dots, X_n$ is to use the moment generating function (mgf) $M_X(t)$ of the random variable X. The moment generating function has the property that sums of independent random variables give a product of individual moment generating functions. For example, $M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t)$ for X_1, X_2 independent. First we define the moment generating function and then show this statement holds.

Definition 4.1. (Moment generating function) The moment generating function (abbreviated mgf) $M_X(t)$ for a random variable X is defined as

$$M_X(t) = \mathbb{E}[e^{tX}], \quad t \in \mathbb{R} \tag{4.2}$$

given that this expectation exists.

Theorem 4.2. Let $X_1, \dots, X_n$ be independent random variables, then the moment generating function of $S_n = X_1 + \cdots + X_n$ is the product of the moment generating functions of $X_1, \dots, X_n$. In other words

$$M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(t) \tag{4.3}$$

Assuming that each moment generating function of all $X_1, \dots, X_n$ exists for t.

Proof. As X_i and X_j are independent for $i \neq j$, we may separate the expectation in the following way

$$\mathbb{E}[e^{tX_1+tX_2}] = \mathbb{E}[e^{tX_1} \cdot e^{tX_2}] = \mathbb{E}[e^{tX_1}]\mathbb{E}[e^{tX_2}]$$

As becomes clear by induction, we have

$$M_S(t) = \mathbb{E}[e^{t(X_1+\cdots+X_n)}] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] = \prod_{i=1}^{n} M_{X_i}(t)$$

□

A special property of the moment generating function is that it is unique for a given probability distribution and may be used to identify a probability distribution from its moment generating function. For example, if a random variable has the moment generating function of a normal distribution, we may state that the random variable is normally distributed.

With this knowledge in mind, we provide examples of $\overline{X}_n$ for $X_i \sim \mathcal{N}(\mu, \sigma^2)$ and $X_i \sim \text{Exp}(\lambda)$. Also, we compare these results to calculating sample means by means of simulations, as can be seen in Figure 7.

Example 4.1. *(Normal distribution)* Let $X_1, \ldots, X_n$ be i.i.d. normally distributed with mean μ and variance σ^2. Similarly, $X_1, \ldots, X_n$ share the same moment generating function. For X_1, its probability density function is

$$p_{X_1}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, x \in \mathbb{R}$$

Then, the moment generating function of X_1 is

$$\begin{aligned}
M_{X_1}(t) = \mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \, \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x^2 - (2\mu + 2\sigma^2 t)x + \mu^2)\right] \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left((x - (\sigma^2 t + \mu))^2 - (\sigma^2 t + \mu)^2 + \mu^2\right)\right] \mathrm{d}x \\
&= \exp\left[-\frac{\mu^2 - (\sigma^2 t - \mu)^2}{2\sigma^2}\right] \cdot \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x - (\sigma^2 t + \mu))^2\right] \mathrm{d}x}_{= 1, \text{ since } \mathcal{N}(\mu + \sigma^2 t, \sigma^2)} \\
&= e^{\mu t + \frac{1}{2}\sigma^2 t^2}
\end{aligned}$$

The mgf of $X_1 + \cdots + X_n$ will be

$$\prod_{i=1}^{n} e^{t\mu + \frac{1}{2}\sigma^2 t^2} = e^{t(n\mu) + \frac{1}{2}t^2 n\sigma^2}$$

As $X_1 + \cdots + X_n$ has the mgf-form of a normal, we can state $X_1 + \cdots + X_n \sim \mathcal{N}(n\mu, n\sigma^2)$. Since $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_n$ also scales by $\frac{1}{n}$, we have $\mathbb{E}[\overline{X}_n] = \frac{1}{n}\mathbb{E}[X_1 + \cdots + X_n] = \mu$ and $\text{Var}(\overline{X}_n) = \frac{1}{n^2}\text{Var}(X_1 + \cdots + X_n) = \frac{n\sigma^2}{n^2} = \sigma^2/n$. So, we may conclude $\overline{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$ which has a pdf of the form

$$p_{\overline{X}_n}(s) = \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{n}{2\sigma^2}(x-\mu)^2}$$

Its pdf is then exponentially small with the form $\exp(-nI(s))$ where $I(s) = \frac{(s-\mu)^2}{2\sigma^2}$ if we ignore the normalization constant. This general form is a reoccurring trend of which large deviation theory deals with.

Example 4.2. *(Exponential distribution)* Let $X_1, \ldots, X_n$ be exponentially distributed with the parameter λ, such that X_i has the probability density function

$$p_{X_i}(x) = \lambda e^{-\lambda x}, x \in [0, \infty)$$

We use the same strategy as outlined in the example above (normal distribution). Evaluating the mgf for X_1 gives

$$\begin{aligned} M_{X_1}(t) = \mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} \lambda e^{-\lambda x} \, \mathrm{d}x \\ &= \lambda \int_0^{\infty} e^{-(\lambda - t)x} \, \mathrm{d}x \\ &= \lambda \left[\frac{-1}{\lambda - t} e^{-(\lambda - t)x} \right]_{x=0}^{x=\infty} \\ &= \frac{1}{1 - t\lambda^{-1}} \end{aligned}$$

The mgf of the sum $X_1 + \cdots + X_n$ will be $(1 - \lambda^{-1}t)^{-n}$ as it is a n times the product of $(1 - \lambda^{-1}t)^{-1}$. The distribution associated with this mgf is the gamma distribution with parameters n and λ, we will denote this by $\Gamma(n, \lambda)$. This distribution has the following pdf:

$$p_{X_1 + \cdots + X_n}(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}, x \in [0, \infty)$$

Here, $\Gamma(x)$ refers to the gamma function, defined as $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} \, \mathrm{d}t$. The gamma function can be seen as an expansion of the factorial to the real numbers, similarly $\Gamma(x+1) = x\Gamma(x)$ and for $x \in \mathbb{N}$ we have $\Gamma(x) = (x-1)!$.
Now, we only have to rescale $X_1 + \cdots + X_n$ by $1/n$ to find the mgf of $\overline{X}_n$. Rescaling a gamma distribution by a factor b changes the parameter λ by λ/b, this can be shown using the mgf of the gamma distribution. Let $Y \sim \Gamma(n, \lambda)$, then $M_{bY}(t) = \mathbb{E}[e^{bYt}] = \mathbb{E}[e^{Ybt}] = M_Y(bt) = (1 - ct/\lambda)^{-n} = (1 - t(\lambda/c)^{-1})^{-n}$ which is equivalent to $\Gamma(n, \lambda/c)$. We can derive the rate function as follows

$$\begin{aligned} I(s) = -\frac{1}{n} \log p_{\overline{X}_n}(s) &= \frac{1}{n} \log \left(\frac{(n\lambda)^n}{\Gamma(n)} s^{n-1} e^{-\lambda n s} \right) \\ &= \lambda s - \frac{1}{n} \log \left(\frac{(n\lambda s)^n}{s\Gamma(n)} \right) \\ &= \lambda s - \log(\lambda s) - \underbrace{\log(n) + \frac{1}{n} \log(\Gamma(n)) + \frac{1}{n} \log(s)}_{\text{limit case for } n \to \infty} \end{aligned}$$

The term 'limit case' will approach -1 for $n \to \infty$. To show this roughly, let $\Gamma(n) \approx n!$ and use Stirling's approximation $n! = n\log(n) - n + \mathcal{O}(\log(n))$. Since $s \neq 0$ is a constant, $\frac{1}{n}\log(s) \to 0$. As for $-\log(n) + \frac{1}{n}\log(\Gamma(n))$, we get

$$-\log(n) + \frac{1}{n}\log(\Gamma(n)) \approx -\log(n) + \frac{1}{n}(n\log(n) - n) = -1$$

As such, for large n, $\overline{X}_n$ has the following rate function

$$I(s) = \lambda s - \log(\lambda s) - 1$$

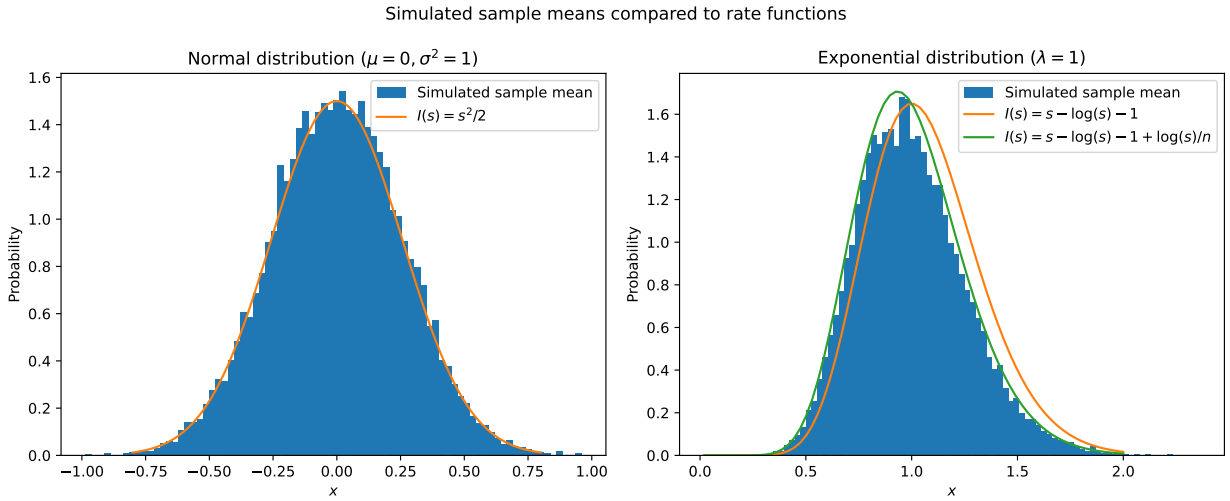


Figure 7: Simulated distribution of sample means $\overline{X}_n$ of $n = 15$ random variables, either $\mathcal{N}(0,1)$ or Exp(1), compared to their rate functions as found in examples 4.1 and 4.2. The rate functions are scaled to the maximum of the simulated distributions. We note that for the exponential distribution, the tail is also taken into account by the rate function. This would not be the case if we were to approximate $\overline{X}_n$ by the central limit theorem, which states that it converges to a normal distribution. Also, the rate function is somewhat shifted since we ignore the term $\frac{1}{n}\log(s)$.

4.1.2 Cramér's theorem

As becomes apparent, the sample mean of i.i.d. continuous random variables will have a pdf of the form.

$$p_{\overline{X}_n}(s) = e^{-nI(s)+o(n)} \tag{4.4}$$

We will make this statement more precise in the form of Cramér's theorem below, which then holds in a broader setting, including discrete random variables. We call n the rate and $I(s)$ the rate function. We remark the usage of $o(n)$, by which we mean all sublinear terms. In other words, all functions $f(n)$ such that $\lim_{n\to\infty} \frac{f(n)}{n} = 0$, take for example $\log(n), \sqrt{n}$ and $\frac{1}{n}$. Before we examine the properties of the rate function $I(s)$ in this case, we are interested when random variables show the phenomena of having exponentially small sample means. As it turns out, we can find the rate function $I(s)$ of the sample mean of any $\mathbb{R}$-valued random variable, up to sublinear terms of the rate function (hence we denote $o(n)$). This important result in large deviation theory is named Cramér's theorem.

Theorem 4.3. (Cramér's Theorem) Let $X_1, \ldots, X_n$ be i.i.d. $\mathbb{R}$-valued random variables which satisfy the following

$$\varphi(t) = \mathbb{E}[e^{tX_1}] < \infty, \quad \forall t \in \mathbb{R} \tag{4.5}$$

Then, for all $a > \mathbb{E}[X_1]$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\overline{X}_n \geq a) = -I(a) \tag{4.6}$$

where

$$I(s) = \sup_{t\in\mathbb{R}} \{ts - \log \varphi(t)\} \tag{4.7}$$

Proof. For proof of Cramér's Theorem, we refer to Theorem I.4 of [Den08]. □

First, we recognize that $\varphi(t)$ is simply the moment generating function of X_1, as defined in Definition 4.1, which we require to exist. Then, Cramér's Theorem states we may find $I(s)$ of the sample mean $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ for deviations greater than the mean $\mathbb{E}[X_1]$ by the *Legendre transform* of $\log \varphi(t)$. Before we derive the rate function using Cramér's Theorem for examples 4.1 and 4.2, we go more in depth into the Legendre transform. This transform is important when we extend Cramér's Theorem to general vector spaces and when we calculate the rate function for the occupation times. This we examine in greater detail in the following sections. First, we give the definition of the Legendre transform below.

Definition 4.4. (Legendre Transform) For a given interval $I \subseteq \mathbb{R}$, let $f : I \to \mathbb{R}$ be a convex function. Then $f^*(s)$ is defined as the Legendre transform and is given as follows

$$f^*(s) = \sup_{x \in I}\{xs - f(x)\} \tag{4.8}$$

We restrict f^* to only take values when the supremum is finite.

Intuitively, the Legendre transform is a mapping between I and a family of tangent values. For example, let f be differentiable, then we find $f^*(s)$ when $\frac{d}{dx}(xs - f(x)) = s - f'(x) = 0$, hence $s = f'(x)$. This motivates the requirement that f must be convex, as otherwise multiple values of I may be associated with a single tangent. Similarly, the bijective map property is used to show that the Legendre transformation is an involution, in other words $(f^*)^* = f$.

For the Legendre transform to be well-defined, we require f to be a convex function. For $I(s)$ to be well-defined as well, we require $\log \varphi(t)$ to be convex. We show this in the following statement.

Proposition 4.5. The moment generating function $\varphi(t)$ of a $\mathbb{R}$-valued random variable X_1, as stated in Cramér's Theorem (Theorem 4.3), is log-convex. By this we mean for any $\theta \in [0, 1]$ and any $t_1, t_2 \in \mathbb{R}$, we have

$$\log \varphi(\theta t_1 + (1-\theta)t_2) \leq \theta \log \varphi(t_1) + (1-\theta)\log \varphi(t_2) \tag{4.9}$$

Proof. As stated earlier, $\varphi(t)$ is the mgf of X_1 (see equation (4.5) of Theorem 4.3). We will show that equation (4.9) holds. For this proof, we will use Hölder's inequality, which states

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{1/p} \cdot \mathbb{E}[|Y|^q]^{1/q}$$

where $1/p + q/1 = 1$ for $p, q \in [1, \infty]$. Let $p = \frac{1}{\theta}$ such that $q = \frac{1}{1-\theta}$, choose $X = e^{\theta t_1 X_1}$ and similarly $Y = e^{(1-\theta)t_2 X_1}$. This yields:

$$\mathbb{E}[e^{(\theta t_1 + (1-\theta)t_2)X_1}] \leq \mathbb{E}[e^{t_1 X_1}]^{\theta} \cdot \mathbb{E}[e^{t_2 X_2}]^{1-\theta}$$

Taking log on both sides gives

$$\begin{aligned}
\log \mathbb{E}[e^{(\theta t_1 + (1-\theta)t_2)X_1}] &\leq \theta \log \mathbb{E}[e^{t_1 X_1}] + (1-\theta)\log \mathbb{E}[e^{t_2 X_1}] \\
\log \varphi(\theta t_1 + (1-\theta)t_2) &\leq \theta \log \varphi(t_1) + (1-\theta)\log \varphi(t_2)
\end{aligned}$$

We conclude that the Legendre transform in Cramér's Theorem is well-defined. □

As stated earlier, Cramér's theorem will only provide the rate function up to sublinear terms. To illustrate this, let $\mathbb{P}(\overline{X}_n \geq a) = e^{-nI(s)+o(n)}$ where $o(n)$ represents some sublinear function. Substituting this into equation (4.6) gives:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\overline{X}_n \geq a) = \lim_{n\to\infty} \frac{1}{n} \log e^{-nI(a)+o(n)} \tag{4.10}$$

$$= \lim_{n\to\infty} \frac{1}{n}(-nI(a) + o(n)) \tag{4.11}$$

$$= -I(a) + \lim_{n\to\infty} \frac{o(n)}{n} \tag{4.12}$$

The term $\lim_{n\to\infty}\frac{o(n)}{n}$ then ignores all sublinear terms present in $\mathbb{P}(\overline{X}_n \geq a)$. Similarly we define $a_n \simeq b_n$ if the share the same rate function, so

$$a_n \simeq b_n \iff \lim_{n\to\infty}\frac{1}{n}(\log a_n - \log b_n) = 0 \tag{4.13}$$

For example, $e^{n^2+\log(n)} \simeq e^{n^2+\sqrt{n}+1}$, as $\lim_{n\to\infty}\frac{1}{n}([n^2+\log(n)]-[n^2+\sqrt{n}]) = \lim_{n\to\infty}\frac{1}{n}(\log(n)-\sqrt{n}) = 0$.

We will now derive the results from examples 4.1 and 4.2 using Cramér's Theorem.

Example 4.3. *(Normal distribution, Cramér)* Let $X_1,\ldots,X_n$ be i.i.d normally distributed with mean μ and variance σ^2. As derived in example 4.1, the mgf of X_1 is $\varphi(t) = \mathbb{E}[e^{tX_1}] = e^{\mu t+\frac{1}{2}\sigma^2t^2}$. To find $I(s)$, we calculate equation (4.7), which is the Legendre transform of $\log\varphi(t) = \mu t+\frac{1}{2}\sigma^2t^2$. As such, we are interested in maximizing

$$g(t) = ts-\mu t-\frac{1}{2}\sigma^2t^2 = (s-\mu)t+\frac{1}{2}\sigma^2t^2$$

The maximum of g is at $\bar{t} = \frac{s-\mu}{\sigma^2}$, so $g(\bar{t}) = \frac{(s-\mu)^2}{\sigma^2}-\frac{(s-\mu)^2}{2\sigma^2} = \frac{(s-\mu)^2}{2\sigma^2}$. Since $I(s) = \sup_{t\in\mathbb{R}} g(t)$, we get $I(s) = \frac{(s-\mu)^2}{2\sigma^2}$ as found in example 4.1. We get the same result since $I(s)$ does not contain sublinear terms.

Example 4.4. *(Exponential distribution, Cramér)*
Let $X_1,\ldots,X_n$ be exponentially distributed with parameter λ. From example 4.2, the mgf of X_1 is $M_{X_1}(t) = (1-t\lambda^{-1})^{-1}$, so we have

$$\varphi(t) = \frac{1}{1-t\lambda^{-1}}$$

Let $g(t) = ts-\log\varphi(t)$ such that $I(s) = \sup_{t\in\mathbb{R}} g(t)$. Our approach is to derive $g(t)$ and find an optimizer $\bar{t}$ which maximizes g, then finally $I(s) = g(\bar{t})$. First, we calculate $g'(t)$ from $g(t) = ts-\log\left[(1-t\lambda^{-1})^{-1}\right] = ts+\log(1-t\lambda^{-1})$ as follows.

$$\begin{aligned} g'(t) &= \frac{d}{dt}\left(ts+\log(1-t\lambda^{-1})\right) \\ &= s+\frac{-\lambda^{-1}}{1-t\lambda^{-1}} \\ &= s-\frac{1}{\lambda-t} \end{aligned}$$

Set $g'(t) = 0$ gives $s = \frac{1}{\lambda-t}$ so $t = \lambda-\frac{1}{s}$. Fix $\bar{t} = \lambda-\frac{1}{s}$. To show that $\bar{t}$ maximizes g, we have shown $g'(\bar{t}) = 0$. Similarly, for the second derivative of g,

$$g''(t) = \frac{d}{dt}\left(s-\frac{1}{\lambda-t}\right) = -\frac{d}{dt}(\lambda-t)^{-1} = -\frac{1}{(\lambda-t)^2} < 0$$

Since we assume $\bar{t}\neq\lambda$, we have $g''(t) < 0$ so this shows that g has a maximum at $\bar{t}$. However, there function might have an upper bound greater than $g(\bar{t})$ for asymptotes $t\to\pm\infty$. Considering $g(t)$, if $t\to-\infty$ then $ts-\log(1-t\lambda^{-1})$ will asymptotically approach ts since $\log(1-t\lambda^{-1})$ is sublinear at this scale. So, g approaches to $-\infty$ if $t\to-\infty$. For $t\to\lambda$, the term $\log(1-t\lambda^{-1})$ will diverge to $-\infty$. The function $g(t)$ does not exist for $t>\lambda$ since then $1-t\lambda^{-1} < 0$ and so $\log(1-t\lambda^{-1})$ does not have a value in $\mathbb{R}$. With this in mind, we have considered all asymptotes for g. Since $g(t)$ is continuous except at $t=\lambda$, we expect $I(s) = g(\bar{t})$ as g does not diverge to $+\infty$ for $t\to\pm\infty$. Therefore, we have the rate function

$$
\begin{aligned}
I(s) &= g(\bar{t}) \\
&= \left(\lambda - \frac{1}{s}\right) s + \log(1 - \lambda^{-1}(\lambda - s^{-1})) \\
&= \lambda s - 1 + \log(\lambda^{-1} s^{-1}) \\
&= \lambda s - 1 - \log(\lambda s)
\end{aligned}
$$

This is the same rate function as found in example 4.2 for $n \to \infty$. This demonstrates that Cramér's theorem only provides $I(s)$ up to $o(n)$ and as such loses some information about the actual pdf.

4.1.3 Connection to Central Limit Theorem

In this subsection we demonstrate how the rate function $I(s)$ is connected to the Central Limit Theorem (CLT). For $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $X_1, \ldots, X_n$ i.i.d. random variables, we have two key convergence properties of the sample mean $\overline{X}_n$:

(1) Law of large numbers:

$$\overline{X}_n \to \mathbb{E}[X_1] \text{ in probability} \tag{4.14}$$

(2) Central Limit Theorem: for $\mu = \mathbb{E}[X_1]$ and $\sigma^2 = \mathrm{Var}(X_1)$

$$\overline{X}_n \to \mathcal{N}(\mu, \sigma^2/n) \text{ in distribution} \tag{4.15}$$

If we assume that the probability density function has the form $p_{\overline{X}_n}(s) = e^{-nI(s)}$, then $s = \mu$ should yield $I(\mu) = 0$ as this is where $\overline{X}_n$ converges to in probability. From this, if we use a Taylor series of $I(s)$ around $s = \mu$, we get $I(s) = 0 + (s-\mu)I'(\mu) + \frac{1}{2}(s-\mu)^2 I''(\mu) + \mathcal{O}(s^3) = \frac{1}{2}(s-\mu)^2 I''(\mu) + \mathcal{O}(s^3)$. Since $\mathcal{O}(s^3)$ becomes a less dominant contributor due to the exponential scaling, we expect $I(s) \approx \frac{(s-\mu)^2}{2} I''(\mu)$, suggesting that $\overline{X}_n$ is similar to a normal distribution, as is expected from the central limit theorem. As shown in Lemma I.14 of [Den08], calculation of $I''(\mu)$ gives $I''(\mu) = \frac{1}{\sigma^2}$. From this result, the distribution takes the form $\exp\left(-n\frac{(s-\mu)^2}{2\sigma^2}\right)$, which states that $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$. This is the same result from the Central Limit Theorem.

4.2 General random variables

We now wish to generalize the idea of exponentially small probability distributions to other random variables instead of sample means. We start by defining when a family of probability measures satisfies the large deviation principle (LDP), which is equivalent to having the form $e^{-nI(s)+o(n)}$ from the previous section, but now formalized.

Definition 4.6. (Large Deviation Principle) Let $\mathcal{X}$ be a topological space and let $\mathcal{B}$ be the Borel σ-algebra of $\mathcal{X}$. Let (μ_n) be a sequence of probability measures on $(\mathcal{X}, \mathcal{B})$. This sequence satisfies the large deviation principle (LDP) with a rate n and rate function I if

(L1) For all closed sets $C \subset \mathcal{X}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(C) \leq - \inf_{x \in C} I(x) \tag{4.16}$$

(L2) For all open sets $O \subset \mathcal{X}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq - \inf_{x \in O} I(x) \tag{4.17}$$

Note 4.1. *(Connection to previous section)* The given definition above is a weaker variant of the following statement:

$$\lim_{n\to\infty} \frac{1}{n} \log \mu_n(A) = -\inf_{x\in A} I(x) \tag{4.18}$$

Which is similar to the statement of Cramér's Theorem (Theorem 4.3) but for probability measures. We note that $-\inf_{x\in A} I(x)$ is used since we now use sets for probability measures and recognize the minima of I in the region of A since large terms in I contribute less due to the exponential scaling.

Note 4.2. *(Weak convergence)* Besides having a weaker definition, the requirements (L1) and (L2) are similar to weak convergence in probability. For context, a definition of weak convergence of probability measures (μ_n) to μ is when

$$\text{(L1')} \quad \limsup_{n\to\infty} P_n(C) \leq P(C), \ \forall C \text{ closed} \tag{4.19}$$

$$\text{(L2')} \quad \liminf_{n\to\infty} P_n(O) \geq P(O), \ \forall O \text{ open} \tag{4.20}$$

Another definition of weak convergence is as follows. Let (μ_n) be a sequence of measures, then μ_n converges weakly to a measure μ if for any continuous bounded function f the following limit converges.

$$\lim_{n\to\infty} \int f(x)\mu_n(\mathrm{d}x) = \int f(x)\mu(\mathrm{d}x) \tag{4.21}$$

Another connection is $X_n \to X$ weakly if and only if $X_n \to X$ in distribution. For related theorems regarding weak convergence, we refer to chapter 18 of [JP04].

We now look to an analogue of Cramér's Theorem but for general topological spaces. In the following section, we will work towards the Gärtner-Ellis Theorem, which does restrict the chosen topological space to be a (real) Hausdorff vector space. To allow for more general spaces, the moment generating function is redefined to depend on a functional λ instead of t, as will be explained in a moment. A simple example would be $\mathbb{R}^2$ where $\mathbf{X} = (X_1, X_2)$ may depend on each other. Then a suitable mgf will be

$$M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[e^{t_1X_1+t_2X_2}] = \mathbb{E}[e^{\mathbf{t}^T\mathbf{X}}] \tag{4.22}$$

If X_1 and X_2 are independent, the mgf of $\mathbf{X}$ will be the product of the moment generating functions of X_1 and X_2. To generalize the idea of the transpose of $\mathbf{t}$ used in equation (4.22) for general vector spaces, we define the concept of a dual space.

Definition 4.7. (Dual space) Let V be a vector space of some field F. We define a *linear functional* on V to be a linear map from V to F. The dual space of V, denoted V^*, is the set of all continuous linear functionals of V.

To illustrate this concept, consider the vector space $V = \mathbb{R}^3$. Any linear map from V to $\mathbb{R}$ will be in the form $\varphi(x, y, z) = ax + by + cz$. Similarly, let $\mathbf{u} = (a, b, c)^T$ and $\mathbf{v} = (x, y, z)^T$ be vectors in $\mathbb{R}^3$, then φ can be written as $\varphi(\mathbf{v}) = \mathbf{u}^T\mathbf{v} = \langle \mathbf{u}, \mathbf{v}\rangle$. Hence the dual space of $V = \mathbb{R}^3$ can be represented as all transposed vectors of $\mathbb{R}^3$. More generally, $\varphi \in V^*$ will be some function if V is more complex, for example if we let V be some function space. We use the notation $\langle \varphi, x\rangle = \varphi(x)$ similar to the inner product in $\mathbb{R}^d$.
For $V = \mathbb{R}^d$, we have that any functional $\varphi \in V^*$ can be written in the form

$$\varphi_\xi(\mathbf{x}) = \langle \varphi_\xi, \mathbf{x}\rangle = \sum_{n=1}^{d} \xi_n \mathbf{x}_n$$

For ℓ^p spaces, a similar representation holds (see chapter 4 of [Nee22]). As will be discussed in the next section, we are interested when V is a function space. Let $C_0(X)$ be the space of all continuous functions

$f : X \to F$ which vanish at infinity. Then let $V = C_0(X)$ [5]. Finally, functionals of the dual space of V can be represented by an integral for different measures of μ.

$$\varphi_\mu(f) = \int_X f \, \mathrm{d}\mu \tag{4.23}$$

With this representation in mind, let us now define the moment generating function for general topological spaces.

Definition 4.8. (Logarithmic moment generating function) For some Hausdorff topological vector space $\mathcal{X}$, let $\mathcal{X}^*$ denote the dual space of $\mathcal{X}$. Let (μ_n) be a family of probability measures associated with random variables (X_n), in other words $\mu_n(\cdot) = \mathbb{P}(X_n \in \cdot)$. Then we define the logarithmic moment generating function $\Lambda_{\mu_n} : \mathcal{X}^* \to (-\infty, \infty]$ in the following way.

$$\Lambda_{\mu_n}(\lambda) = \log \mathbb{E}_{\mu_n}[e^{\langle \lambda, X_n \rangle}] = \log \int_{\mathcal{X}} e^{\lambda(x)} \mu_n(\mathrm{d}x) \tag{4.24}$$

$$\mathcal{D}_\Lambda = \{\lambda \in \mathcal{X}^* : \Lambda(\lambda) < \infty\} \tag{4.25}$$

Note 4.3. As stated earlier, λ is now a functional instead of the scalar constant t.

Since we consider general topological spaces, we also have to redefine the Legendre transformation to suit $\mathcal{X}$. Regarding Definition (4.4), we substitute I for $\mathcal{X}$ and change xs for $\langle \lambda, x \rangle$ similar to $\mathbf{s}^T\mathbf{x}$, if $\mathbf{s}$ and $\mathbf{x}$ are vectors in $\mathbb{R}^d$. If f is a linear map of $\mathcal{X}$, for example $f(x, y) = 2x + 3y$ for $\mathbb{R}^2$, note that f^* is simply the sum of two Legendre transforms, namely f_x^* and f_y^* given that $f_x(x) = 2x$ and $f_y = 3y$ such that $f(x, y) = f_x(x) + f_y(y)$. For general f, the suprema cannot be separated per component as any maximum can depend on multiple components. In this case, we define the modified Legendre transform as follows:

Definition 4.9. (Legendre-Fenchel transform) Let $\mathcal{X}$ be a locally convex Hausdorff topological vector space. Let $f : \mathcal{X} \to (-\infty, \infty]$ be a convex function, then

$$f^*(\lambda) = \sup_{x \in \mathcal{X}} \{\langle \lambda, x \rangle - f(x)\} \tag{4.26}$$

Finally, $f^* : \mathcal{X}^* \to (-\infty, \infty]$ is the Fenchel-Legendre transform of f.

We now provide the Gärtner-Ellis theorem for general topological spaces. For $\mathcal{X} = \mathbb{R}^d$, it is enough to require $0 \in \operatorname{int} \mathcal{D}_\Lambda$ to state that Λ is convex and Λ^* is a rate function. In this situation, we substitute this requirement for exponential tightness. We give its definition below.

Definition 4.10. (Exponentially tight) A family of probability measures $\{\mu_n\}$ on $\mathcal{X}$ is exponentially tight if for all $\alpha < \infty$, there exists a compact set $K_\alpha \subset \mathcal{X}$ such that

$$\limsup_{n \to \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \setminus K_\alpha) < -\alpha \tag{4.27}$$

Note 4.4. The concept of exponentially tightness is connected to (regular) tightness of measures. If a family of measures (μ_n) is right, then for any μ_n, we can bound $\mu_n(\mathcal{X} \setminus K_\alpha)$ by α. For exponential tightness, the definition is similar.

Definition 4.11. (Exposed point) A point $x \in \mathcal{X}$ is exposed for Λ^* if there exists a $\lambda \in \mathcal{X}^*$ such that

$$\langle \lambda, x \rangle - \Lambda^*(x) > \langle \lambda, z \rangle - \Lambda^*(z), \ \forall z \neq x \tag{4.28}$$

[5] Given how this is defined in functional analysis, X is a locally compact space. We state that a function vanishes at infinity if for every $\varepsilon > 0$ there exists a compact set $K \subseteq X$ such that $|f(x)| < \varepsilon$.

Theorem 4.12. (Abstract Gärtner-Ellis) Let (μ_n) be a family of exponentially-tight probability measures. Define $\overline{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n}\Lambda_{\mu_n}(n\lambda)$. Then the following statements hold.

1. For every closed set $C \subset \mathcal{X}$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(C) \leq - \inf_{x\in C} \overline{\Lambda}^*(x) \tag{4.29}$$

2. Let $\mathcal{F}$ be the set of all exposed points of $\overline{\Lambda}^*$ for a given λ, for which $\Lambda(\lambda) = \overline{\Lambda}(\lambda)$ (limit exists) and $\Lambda(\gamma\lambda) < \infty$ for some $\gamma > 1$. Then, for every open set $O \subset \mathcal{X}$

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq - \inf_{x\in O\cap\mathcal{F}} \overline{\Lambda}^*(x) \tag{4.30}$$

3. Let $O \subset \mathcal{X}$ be an open set, if

$$\inf_{x\in O\cap\mathcal{F}} \overline{\Lambda}^*(x) = \inf_{x\in O} \overline{\Lambda}^*(x) \tag{4.31}$$

then (μ_n) satisfies the LDP with rate function $\overline{\Lambda}^*$

Proof. For the proof, we refer to Theorem 4.5.20 from [DZ10]. □

It is clear that if we may 'ignore' the exposed points holds from Theorem 4.12, we have shown that LDP properties (L1) and (L2) are satisfied by (1) and (2) from said theorem. To recover the rate function of the family (μ_n), we have to find $\overline{\Lambda}^*$. If $\overline{\Lambda}(\lambda) = \limsup_{n\to\infty} \frac{1}{n}\Lambda_{\mu_n}(n\lambda)$ is finite, then we have

$$\lim_{n\to\infty} \frac{1}{n}\Lambda_{\mu_n}(n\lambda) = \Lambda(\lambda) = \sup_{x\in\mathcal{X}}\{\langle\lambda, x\rangle - I(x)\} \tag{4.32}$$

We can reverse the Legendre-Fenchel transform to acquire the rate function $I(x)$ *if* I is a convex function. Then we calculate I as follows:

$$I(x) = \Lambda^*(x) = \sup_{\lambda\in\mathcal{X}^*}\{\langle\lambda, x\rangle - \Lambda(\lambda)\} \tag{4.33}$$

For details, we refer to Theorem 5.4.10 of [DZ10].

To summarize, we can find $I(x)$ using the Legendre transform using equation (4.33) if the following requirements are satisfied:

1. The limit supremum of $\frac{1}{n}\Lambda_{\mu_n}(n\lambda)$ exists, referred to by $\overline{\Lambda}(\lambda)$. The stronger requirement which we also mention is that the limit of $\frac{1}{n}\Lambda_{\mu_n}(n\lambda)$, noted as $\Lambda(\lambda)$, exists when the limit supremum is finite (Theorem 5.4.10(a) of [DZ10]).
2. The family of $\{\mu_n\}$ is exponentially tight as defined by Definition 4.10. For $\mathcal{X} = \mathbb{R}^d$, having $0 \in \operatorname{int}\mathcal{D}_\Lambda$ as defined by equation (4.24) also satisfies this requirement.

4.3 Large deviations for Markovian local times

4.3.1 Definitions

In this section, we will discuss how we can apply the knowledge of large deviation theory to Markov processes. We will use a similar theorem to the Gärtner-Ellis theorem (see Theorem 4.12) for probability measures called the local times and occupation times. We give their definitions as follows.

Definition 4.13. (Local times) Consider a continuous-time Markov process $(X_t)_{t\geq 0}$. Define the *local time* ℓ_t as the time the process is contained within some set, in other words

$$\ell_t(\cdot) = \int_0^t \mathbf{1}_{X_s}(\cdot)\,\mathrm{d}s \tag{4.34}$$

The normalized local times, called the *occupation time* L_t, is given below.

$$L_t(\cdot) = \frac{1}{t}\ell_t = \frac{1}{t}\int_0^t \mathbf{1}_{X_s}\,\mathrm{d}s \tag{4.35}$$

The occupation time $L_t(A)$ represents the fraction of which the process X_s spends in the set A up to a time t.

Note 4.5. For simplicity, we may also define ℓ_t and L_t for singleton sets, in other words as a function of the state space. For example, $\ell_t(x) = \int_0^t \mathbf{1}(X_s = x)\,\mathrm{d}s$ or $L_t = \frac{1}{t}\int_0^t \mathbf{1}(X_s = x)\,\mathrm{d}s$.

Note 4.6. We motivate by choosing L_t for the large deviation principle as follows. We can represent L_t by discretizing time into steps n of δt such that $t = n\delta t$. In discrete form, we have

$$L_t(\cdot) = \lim_{\delta t\to 0}\frac{1}{n\delta t}\sum_{i=0}^{n}\mathbf{1}_{X_{i\delta t}}(\cdot)$$

This is similar to the sample mean of a sequence of variables $X_{\delta t},\ldots,X_{n\delta t}$. However, since the position of X_t will depend on its history, we cannot describe this Markov process as a sample mean of i.i.d. random variables. Luckily, we do not have this restriction of independent variables for large deviation theory in generalized topological spaces. With speaking of large deviation theory for occupation times L_t, we speak of the large deviation principle with rate t instead of rate n.

We will consider L_t as this normalized variant represents a probability measure on the state space S. First, $L_t(S) = \frac{1}{t}\int_0^t \mathbf{1}_{X_s}(S)\,\mathrm{d}s = \frac{1}{t}\cdot t = 1$ since $X_s \in S$ always. Secondly, if $A, B \subset S$ are disjoint, then $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$. This also holds for countable families of subsets, thus $L_t(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} L_t(A_n)$ if all A_n are pairwise disjoint. From this, we consider that L_t is a probability measure.

We have to take into account that L_t has two parameters: time t and position x (see note 4.5). Let us assume that S is a finite state space, then we may represent L_t as a vector $\mathbf{L}_t = (L_t(x_1),\ldots,L_t(x_n))^T$ where $x_1,\ldots,x_n$ are the elements of S. For infinite state spaces, we will represent L_t as a function instead of a vector, such that $L_t(x)$ lives in the function space of $\mathbb{R}^d \to [0,\infty)$ or $\mathbb{Z}^d \to [0,\infty)$. Then, $\{L_t(x)\,|\,x\}$ is a family of random variables (in a topological vector space). With this in mind, we now consider finding a rate function $I(g)$ for $L_t(x)$, which we discuss in the next subsection.

4.3.2 Rate function for local times of Brownian motion

Our goal is to derive the rate function $I(g)$ of L_t. Note that g is now a function also living in the same space $\mathcal{X}$ of L_t. In our approach to acquire $I(g)$, let $\mathcal{X}$ be the space of all probability density functions. This may not seem a problem, since L_t is a probability measure. However, we assumed that $\mathcal{X}$ is a topological vector space, which the space of probability density functions unfortunately is not. For now, consider the state space $\mathbb{R}$. We will describe this set of probability density functions as the set $\mathcal{F}_0$.

$$\mathcal{F}_0 = \left\{f \in C^1(\mathbb{R}) : f > 0, \int_{\mathbb{R}} f\,\mathrm{d}x = 1\right\} \tag{4.36}$$

$$\mathcal{F}_0^d = \left\{f \in C^1(\mathbb{R}^d) : f > 0, \int_{\mathbb{R}} f\,\mathrm{d}x = 1\right\} \tag{4.37}$$

$$\tag{4.38}$$

In other words, all functions that are continuously differentiable on $\mathbb{R}$ for which $f(y) > 0$ for all $y \in \mathbb{R}$ and $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$.

It has been shown by Donsker and Varadhan ([DV75a] [6]) that if $(X_t)_{t\geq 0}$ is Brownian motion in one dimension, we may apply a similar form of the Gärtner-Ellis theorem. First, we consider the form of $\langle \Phi, L_t \rangle$ where $\Phi : \mathcal{F}_0 \to \mathbb{R}$ is some functional of a dual space of $\mathcal{F}_0$. If we assume that for measure μ which describes Φ, we let $\mathrm{d}\mu = \varphi(x)\,\mathrm{d}x$ where φ is a bounded continuous function, then we may derive

$$\begin{aligned}\langle \Phi, L_t \rangle &= \int_{\mathbb{R}} L_t \,\mathrm{d}\mu && (4.39)\\ &= \frac{1}{t}\int_{\mathbb{R}}\int_0^t \mathbf{1}(X_s = x)\varphi(x)\,\mathrm{d}s\,\mathrm{d}x && (4.40)\\ &= \frac{1}{t}\int_0^t \varphi(X_s)\,\mathrm{d}s && (4.41)\end{aligned}$$

In the second line of the above equation (equation 4.40), the term $\mathbf{1}(X_s = x)\varphi(x)$ is simply $\varphi(x)$ accounted for the fraction of which X_s is at x. Since we integrate over $\mathbb{R}$ and thus all possibilities of x, we can write $\varphi(X_s)$.

Now, we present the analogue of the Gärtner-Ellis theorem stated by Donsker and Varadhan [DV75a]. We have the following statement.

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{-t\langle \Phi, L_t\rangle}] = -\inf_{f\in\mathcal{F}_0} \left\{ \langle \Phi, f\rangle + \frac{1}{8}\int_{-\infty}^{\infty} \frac{[f'(y)]^2}{f(y)}\,\mathrm{d}y \right\} \tag{4.42}$$

If we let exchange Φ for $-\Phi$ we may use that $\inf(-f) = -\sup(f)$ in equation (4.42). This yields the following equation

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{t\langle \Phi, L_t\rangle}] = \sup_{f\in\mathcal{F}_0} \left\{ \langle \Phi, f\rangle - \frac{1}{8}\int_{-\infty}^{\infty} \frac{[f'(y)]^2}{f(y)}\,\mathrm{d}y \right\} \tag{4.43}$$

This is similar to the Gärtner-Ellis theorem with the rate function $I(f) = \frac{1}{8}\int_{-\infty}^{\infty} \frac{[f'(y)]^2}{f(y)}\,\mathrm{d}y$.

4.3.3 Derivation of rate function

We now know the rate function of the occupation times for Brownian motion. However, we are interested in how we may derive the found result from the left hand side of equation (4.43). Then we can derive the function for $\mathbb{R}^d$ which will be useful if we want to consider movement of particles in 3 dimensions. Our approach is to compute the left hand side of equation (4.43) using the Feynman-Kac formula. We first rewrite $t\langle \Phi, L_t\rangle$ into an integral form.

$$\mathbb{E}[e^{t\langle \Phi, L_t\rangle}] = \mathbb{E}[e^{\int_0^t \varphi(X_s)\,\mathrm{d}s}] \tag{4.44}$$

As was found via equation (4.39). Then, by the Feynman-Kac (see Theorem 3.10), using $f = 1$ (henceforth 1 will denote the constant function $f(x) = 1$) and $V = \varphi$, we obtain

$$g(t,x) = \mathbb{E}_x[e^{\int_0^t \varphi(X_s)\,\mathrm{d}s}] \tag{4.45}$$

Let L be the generator of Brownian motion. Since $g(t,x) = K_t 1 = e^{t(L+\varphi)}1$, our result becomes $\mathbb{E}[e^{t\langle \Phi, L_t\rangle}] = e^{t(L+\varphi)}1$. The left hand side of equation (4.43) becomes

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{t\langle \Phi, L_t\rangle}] = \lim_{t\to\infty} \frac{1}{t} \log(e^{t(L+\varphi)}1) \tag{4.46}$$

Let $\{\nu_i\}$ denote the set of all eigenvalues of the operator $L + \varphi$. As it turns out, we can rewrite equation (4.46) using the maximal eigenvalue $\sup_i \nu_i$, also called the *principal* eigenvalue.

[6] Donsker and Varadhan have published four papers with this title, we refer to the first paper they published in 1975.

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[e^{t\langle\Phi,L_t\rangle}]=\sup_i \nu_i \tag{4.47}$$

We show this statement holds for finite state spaces in Theorem 4.15.

The supremum over all eigenvalues can be generalized to the set over all normalized (L^2) functions. Again, we do not directly show this for functions, but prove a simpler case for the finite state space, see corollary 4.16. In general, since $L+\varphi$ will be a self-adjoint operator [7] , we can state the following.

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[e^{t\langle\Phi,L_t\rangle}]=\sup_{h:\int h^2\,\mathrm{d}x=1}\langle h,(L+\varphi)h\rangle \tag{4.48}$$

$$=\sup_{h:\int h^2\,\mathrm{d}x=1}\{\langle h^2,\psi\rangle+\langle h,Lh\rangle\} \tag{4.49}$$

We now apply the generator of 1-dimensional Brownian motion, which is $L=\frac{1}{2}\frac{d^2}{dx^2}$. If $[hh']_{-\infty}^{\infty}=0$ we obtain the equation

$$\langle h,Lh\rangle=\frac{1}{2}\int_{-\infty}^{\infty}h(x)h''(x)\,\mathrm{d}x=-\frac{1}{2}\int_{-\infty}^{\infty}(h'(x))^2\,\mathrm{d}x \tag{4.50}$$

This is reasonable since we expect $\int_{-\infty}^{\infty}h^2(x)\,\mathrm{d}x<\infty$, and as such h and h' should vanish at infinity. Let $g=h^2$ and $g>0$ such that $g\in\mathcal{F}_0$. From $\int_{-\infty}^{\infty}(h'(x))^2\,\mathrm{d}x=\frac{1}{4}\int_{-\infty}^{\infty}\frac{[g'(x)]^2}{g(x)}\,\mathrm{d}x$, we obtain the equation

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{E}[e^{t\langle\Phi,L_t\rangle}]=\sup_{g\in\mathcal{F}_0}\left\{\langle g,\varphi\rangle-\frac{1}{8}\int\frac{[g'(x)]^2}{g(x)}\,\mathrm{d}x\right\} \tag{4.51}$$

Thus yielding the rate function $I(g)=-\frac{1}{8}\int\frac{[g'(x)]^2}{g(x)}\,\mathrm{d}x$ for occupation times measure L_t with Brownian motion, as was found by Donsker and Varadhan.

We can expand this notion to d-dimensional Brownian motion. As it turns out, this can be decomposed as the sum of d rate functions of the different dimensions. This we show in the proposition below.

Proposition 4.14. (Rate function for occupation times) Given the d-dimensional Brownian motion generator

$$L=\frac{1}{2}\Delta^d=\frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2} \tag{4.52}$$

We obtain the rate function for L_t as

$$I(g)=\frac{1}{8}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{[\partial_{x_i}g]^2}{g}\,\mathrm{d}\mathbf{x} \tag{4.53}$$

Where $g:\mathbb{R}^d\to\mathbb{R}$, $g\in\mathcal{F}_0^d$ and $\mathbf{x}=(x_1,\ldots,x_d)^T$.

Proof. We start at equation (4.48) and use the generator L as given above. Substituting this into $\langle h,Lh\rangle$ gives

[7] φ is real so it is self-adjoint. For $L=\partial_x^2$, we know

$$\langle Lf,g\rangle=\int\partial_x^2 f\cdot g\,\mathrm{d}x=-\int\partial_x f\cdot\partial_x g\,\mathrm{d}x=\int f\cdot\partial_x^2 g\,\mathrm{d}x=\langle f,Lg\rangle$$

For finite state spaces, the self-adjoint requirement is equivalent to $L+\varphi$ being a Hermitian matrix. Furthermore, since $L+\varphi$ is a real matrix, it symmetric, which is a requirement for both Theorem 4.15 and corollary 4.16. For proving the statement for self-adjoint operators in a more rigorous manner, one might find the min-max theorem useful, see theorem 2.19 of [Tes14].

$$\langle h, Lh\rangle = \frac{1}{2}\int_{\mathbb{R}^d} h(\mathbf{x})\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}h(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{4.54}$$

$$= -\frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}(\partial_{x_i}h)^2\,\mathrm{d}\mathbf{x} \tag{4.55}$$

Where in the last line we used a result which we will prove later on in the proof of Lemma 5.2 (see equation (5.32), we similarly assume h and $\partial_{x_i}h$ vanish at $x_i \to \pm\infty$ for all $i = 1, \ldots, d$). Substitute $h^2 = g$ so that $h = \sqrt{g}$. Then $\partial_{x_i}h = \partial_{x_i}\sqrt{g} = \frac{\partial_{x_i}g}{2\sqrt{g}}$. Substitute this derivative into equation (4.54) gives

$$\langle h, Lh\rangle = -\frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\left(\frac{\partial_{x_i}g}{2\sqrt{g}}\right)^2\mathrm{d}\mathbf{x} = -\frac{1}{8}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{[\partial_{x_i}g]^2}{g}\,\mathrm{d}\mathbf{x} \tag{4.56}$$

Assuming convexity of $I(g)$ we can read the rate function by equation (4.48) as $-\langle h, Lh\rangle$. We then change h^2 for g, so we optimize over $g \in \mathcal{F}_0^d$. We then obtain the rate function

$$I(g) = -\frac{1}{8}\sum_{i=1}^{d}\int_{\mathbb{R}^d}\frac{[\partial_{x_i}g]^2}{g}\,\mathrm{d}\mathbf{x} \tag{4.57}$$

□

4.3.4 Largest eigenvalues and tools

In this section, show two statements used in the previous section hold regarding largest eigenvalues. Furthermore, we show some properties of the eigenvalues of the operator L in the $\mathbb{Z}^d$ and $\mathbb{R}^d$ case, which will be used later on.

First, we come back to the assumption made about the largest eigenvalue for a finite state space. For this setting, both L and V are representable by a matrix where the canonical basis vectors represent elements of the state space. We prove the statement below. Afterwards, we prove a corollary which generalizes the maximum over all eigenvalues to an expression which uses the supremum over all normalized vectors.

Theorem 4.15. (Largest eigenvalue for finite state space) Assume $(X_t)_{t\geq 0}$ is a continuous time Markov chain with a finite state space S and a real generator L. Let $\varphi : S \to \mathbb{R}$ be an arbitrary function which acts as a multiplicative operator. For the matrix $L + \varphi$, let ψ_i be the eigenvector associated to eigenvalue ν_i. If L is symmetric and ψ_i has positive entries for all eigenvectors, we have

$$\lim_{t\to\infty}\frac{1}{t}\log\left(e^{t(L+\varphi)}\vec{1}\right)_j = \max_i \nu_i \tag{4.58}$$

Where $\vec{1}$ is the vector $(1, \ldots, 1)^T$.

Proof. Let L and φ be representable by a $n \times n$ matrix. Since L is symmetric and therefore $L + \varphi$ also, we have real eigenvalues and, without loss of generality, let $\psi_1, \ldots, \psi_n$ orthogonal eigenvalue basis where $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_n$. We can rewrite $e^{t(L+\varphi)}\vec{1}$ using this orthogonal basis as follows.

$$\begin{aligned} e^{t(L+\varphi)}\vec{1} &= e^{t(L+\varphi)}\sum_{i=1}^{n}\langle\psi_i, \vec{1}\rangle\psi_i \\ &= \sum_{i=1}^{n}\langle\psi_i, \vec{1}\rangle\sum_{k=0}^{\infty}\frac{t^k(L+\varphi)^k\psi_i}{k!} \\ &= \sum_{i=1}^{n} e^{t\nu_i}\langle\psi_i, \vec{1}\rangle\psi_i \end{aligned}$$

Define $\mathrm{e}_1, \ldots, \mathbf{e}_n$ as the canonical basis vectors in $\mathbb{R}^n$. Then we can rewrite the ith component by taking the inner product.

$$\begin{aligned}\langle \mathbf{e}_j, e^{t(L+\varphi)}\vec{1}\rangle &= \sum_{i=1}^{n} e^{t\nu_i}\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle \\ &= e^{t\nu_1}\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle \left(1 + \sum_{i=2}^{n} e^{t(\nu_i-\nu_1)}\frac{\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle}{\langle \mathbf{e}_j, \psi_1\rangle\langle \psi_1, \vec{1}\rangle}\right)\end{aligned}$$

Taking the log and applying $\lim_{t\to\infty}\frac{1}{t}$ yields

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\log(\langle \mathbf{e}_j, e^{t(L+\varphi)}\vec{1}\rangle) &= \nu_1 + \lim_{t\to\infty}\frac{1}{t}\log\left(1 + \sum_{i=2}^{n} e^{t(\nu_i-\nu_1)}\frac{\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle}{\langle \mathbf{e}_j, \psi_1\rangle\langle \psi_1, \vec{1}\rangle}\right) \\ &= \nu_1 + \lim_{t\to\infty}\frac{\sum_{i=2}^{\infty}(\nu_i-\nu_1)e^{t(\nu_i-\nu_1)}\frac{\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle}{\langle \mathbf{e}_j, \psi_1\rangle\langle \psi_1, \vec{1}\rangle}}{1 + \sum_{i=2}^{\infty} e^{t(\nu_i-\nu_1)}\frac{\langle \mathbf{e}_j, \psi_i\rangle\langle \psi_i, \vec{1}\rangle}{\langle \mathbf{e}_j, \psi_1\rangle\langle \psi_1, \vec{1}\rangle}} \\ &= \nu_1\end{aligned}$$

Where in the last limit, we used L'Hôpital's rule. □

Note 4.7. Matrix L is symmetric if for any two points $x \in S$ and $y \neq x$, we have $c(x, y) = c(y, x)$. For instance, the symmetric continuous-time random walk has this property.

Corollary 4.16. (Generalized largest eigenvalue) Assume the same requirements stated in Theorem 4.15 hold (the finite state space setting). We then have

$$\max_i \nu_i = \sup_{v\in\mathbb{R}^n, \|v\|=1} \langle v, (L+\varphi)v\rangle \tag{4.59}$$

Proof. Without loss of generality, assume the set of orthogonal eigenvectors $\psi_1, \ldots, \psi_n$ is normalized, in other words $\langle \psi_i, \psi_i\rangle = 1$ for all $i = 1, \ldots, n$. Let $v \in \mathbb{R}^n$ with $\|v\| = 1$ be arbitrary. Using spectral decomposition and applying the operator $L + \varphi$ yields

$$(L+\varphi)v = (L+\varphi)\sum_{i=1}^{n}\langle \psi_i, v\rangle\psi_i = \sum_{i=1}^{n}\nu_i\langle \psi_i, v\rangle\psi_i$$

Evaluating $\langle v, (L+\varphi)v\rangle$ gives

$$\begin{aligned}\langle v, (L+\varphi)v\rangle &= \left\langle \sum_{i=1}^{n}\langle \psi_i, v\rangle\psi_i, \sum_{i=1}^{n}\nu_i\langle \psi_i, v\rangle\psi_i\right\rangle \\ &= \sum_{i=1}^{n}\nu_i\langle \psi_i, v\rangle^2 \cdot \langle \psi_i, \psi_i\rangle && \text{(orthogonality)} \\ &= \sum_{i=1}^{n}\nu_i\langle \psi_i, v\rangle^2 && \text{(normalization)} \\ &\leq \nu_1\sum_{i=1}^{n}\langle \psi_i, v\rangle^2 \\ &\leq \nu_1 && \text{(using } \|v\| = 1)\end{aligned}$$

From the above inequality, we obtain $\max_i \nu_i \geq \sup_{\|v\|=1}\langle v, (L+\varphi)v\rangle$. Since ψ_1 is normalized, we find $\langle \psi_i, (L+\varphi)\psi_i\rangle = \nu_i$. Therefore, $\sup_{\|v\|=1}\langle v, (L+\varphi)v\rangle \geq \max_i \nu_i$. From these inequalities, we have shown that equation (4.59) holds. □

While we can find the largest eigenvalue using theorem 4.15 in the finite state space setting, it is also useful to consider what other properties the eigenvalues have. While φ is arbitrary and therefore we cannot state much about its eigenvalues, we find that the eigenvalues of L for the symmetric continuous-time random walk are negative or zero (nonpositive). This is shown for the 1-dimensional case below, we assume this holds for d-dimensions.

Proposition 4.17. (Nonpositive eigenvalues, $\mathbb{Z}$) Define the random walk as a continuous-time Markov process with state space $S = \mathbb{Z} \cap [0, N]$ with the following generator.

$$(Lf)(x) = f(x+1) + f(x-1) - 2f(x) \tag{4.60}$$

Given Dirichlet boundary conditions due to a bounded state space, the eigenvalues of L are nonpositive.

Proof. Let us denote the solution by u_n where we have $u_0, \dots, u_n$. By definition, we have

$$Lu = u_{n-1} + u_{n+1} - 2u_n$$

We try the ansatz $u_n = \sin(na)$, substituting gives:

$$\begin{aligned} Lu_n &= \lambda_a u_n \\ u_{n-1} + u_{n+1} - 2u_n &= \lambda_a u_n \\ \sin((n-1)a) + \sin((n+1)a) &= (\lambda_a - 2)\sin(na) \\ 2\cos(a) &= \lambda_a - 2 \\ \lambda_a &= 2(\cos(a) - 1) \end{aligned}$$

Since we consider $a \in \mathbb{R}$, by $\cos(a) \leq 1$ we have $\lambda_a = 2(\cos(a) - 1) \leq 0$. □

Proposition 4.18. (Nonpositive eigenvalues, $\mathbb{R}$) Let $(B_t)_{t \geq 0}$ be 1-dimensional Brownian motion, having generator (see Proposition 3.13)

$$(Lf)(x) = \frac{1}{2}\frac{d^2}{dx^2} f(x) \tag{4.61}$$

Assuming L is only defined on bounded measurable functions, all eigenvalues of L are nonpositive.

Proof. Let ψ be an eigenfunction of L, then

$$\begin{aligned} \frac{1}{2}\psi''(x) &= \lambda(x) \\ \psi''(x) &= 2\lambda(x) \end{aligned}$$

Assume that $\lambda > 0$, then the solution to the above equation is

$$\psi(x) = Ae^{x\sqrt{2\lambda}} + Be^{-x\sqrt{2\lambda}}$$

Since we consider the space $\mathbb{R}$, both solutions will be unbounded if $A \neq 0$ or $B \neq 0$. Since we do not consider unbounded functions, eigenfunctions with eigenvalues $\lambda > 0$ do not occur. As such, we can state that $\lambda \leq 0$. □

Note 4.8. For d dimensions, we can extend Propositions 4.17 and 4.18 by splitting the d-dimensional generator L up into d parts. For $\mathbb{Z}^d$, the generator is

$$Lf = \sum_{y \sim x} (f(y) - f(x)) = \sum_{i=1}^{d} (f(\dots, x_i - 1, \dots) + f(\dots, x_i + 1, \dots) - 2f(\dots, x_i, \dots)) \tag{4.62}$$

Where we have $f(x_1, \ldots, x_d)$. For $\mathbb{R}^d$, the generator is

$$Lf = \frac{1}{2} \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} \tag{4.63}$$

Then by solving Lu for one dimension, we may suggestable the eigenvalues of the d-dimensional case can be described by the sum of the eigenvalues of the 1-dimensional cases, similar to method of separable variables.

5 Localized model

The localized model will be the first model we develop with regards to describing movement and trapping of excitons. In this thesis, we are interested in two methods of trapping:

1. **(Constant trapping)** We consider that an exciton may decay at any time and thus at any place. So, for a 'living' exciton, the particle can be 'killed' by a trap denoted as A. This will simulate, as we discuss in the delocalized model, the emission of vanadium.

2. **(Randomly distributed traps)** The lattice where the exciton lives in also contains traps which are randomly distributed. We refer to this trap as B and this trap simulates the emission due to trapping by europium atoms (again, we refer to chapter 6).

We first start with this simple model to understand the behavior of excitons within crystals. Our simplification consists of the following.

1. The exciton, also called the particle, is governed by a continuous-time Markov chain $(X_t)_{t\geq 0}$. The state space of the exciton will be $\mathbb{Z}^d$ where d denotes the dimension, plus the cemetery states.

2. An exciton can be killed by either trap A or trap B. We let $\{A\}$, $\{B\}$ denote the killed state of an exciton of A and B respectively. The cemetery states are $\{A, B\}$.

3. The rate of migration between states within $\mathbb{Z}^d$ is independent of the location. In other words, the exciton will behave similar to a symmetric random walk if no traps are present.

4. Trap B is present at positions in $\mathbb{Z}^d$ and will be randomly distributed by a Bernoulli distribution with a probability p of a site having a B-trap. The trapping influence of a B-type trap is only experienced by an exciton when the particle is present at the site of the trap. In other words, an exciton cannot be trapped by a B-type trap is there is no B-type trap at its location. Hence, we have a **localized** trapping potential.

5. We assume that the excitons do not interact with each other. Instead, we assume the behavior of each exciton can be described independently.

This chapter is divided in the following sections. In the first section, named 'Description of the model', we go further in detail how to mathematically describe this simplified model. For example, how 'probability mass flow' is governed by the generator of a continuous symmetric simple walk. It will become clear after the first section that most equations for the probability of (not) trapping depend on a 'total mass'. In the section 'Hard traps' we consider an extreme case where we derive asymptotics for the total mass at a simple level. This serves as an introduction for the next section, 'soft traps', which covers the traps we have in the localized model. In this section, the subsection 'Approach' describes how we solve asymptotics of the total mass with every step a short explanation. The next two subsections, 'Box expansion' and 'Potential contribution' are more rigorous, in the sense they contain more details and proofs about deriving the total mass asymptotics. We also discuss the relation between this model and the parabolic Anderson model, which provides an alternative method to acquire the results from the soft-trap approach. The final section 'Many particle setting' studies how we can expand the one-particle case of probability mass into a more appropriate case of multiple particles.

5.1 Description of model

Let $(X_t)_{t\geq 0}$ be a continuous-time Markov chain which describes the trajectory of a single exciton. The state space of this Markov chain is $S = \mathbb{Z}^d \cup \{A, B\}$, where A and B refer to the trapped states as given at the beginning of this chapter. For every point $x \in \mathbb{Z}^d$ we have a random potential $V(x)$ which is independently and identically distributed with respect to x which will represent the trapping potential. For simple cases, we assume V has taken some value according to its distribution and then calculate with V. At that point, we know $V(x)$ for all $x \in \mathbb{Z}^d$. This is what we call the *quenched* setting. Similarly, if we take into account

that V is randomly distributed, so we do not know $V(x)$, we describe the *annealed* setting. It is often easier to analyze the quenched setting and later relaxing the condition that V is fixed.

Let L be the generator associated with this continuous-time Markov chain. For this instance, let α denote the rate of the exciton moving to a neighbor state (for example, x to $x+1$). Similarly, γ_1 and γ_2 are the trapping rates of traps A and B respectively. Define the function $\xi(x)$ to be 1 if there is a B-type trap at x and 0 if there is no B-type trap at x. Then as said earlier, ξ is Bernoulli-distributed with $\mathbb{P}(\xi(x)=0)=1-p$ and $\mathbb{P}(\xi(x)=1)=p$ with $p\in[0,1]$ the probability of placing a B-trap at some site. Note that $\xi(x)$ is identically and independently distributed with respect to x. From equation 3.18, we find that the generator L equals

$$(Lf)(x)=\sum_{y\sim x}\alpha[f(y)-f(x)]+\gamma_1[f(A)-f(x)]+\gamma_2\xi(x)[f(B)-f(x)] \tag{5.1}$$

Define $\mu(t,x)=\mu_t(x)$ as the probability of the exciton being at the position $x\in S$ at a time t. We are interested when the exciton is trapped, so we want to describe $\mu_t(x)$ when the particle enters the cemetery states $\{A,B\}$. To find how $\mu_t(x)$ changes with respect to t and x, we can acquire partial differential equations from Kolmogorov's backward[8] equation (Theorem 3.9). By using $\mu(t,x)=S_t\mu(0,x)$ and $\frac{\partial}{\partial t}(S_tf)=LS_tf$, we find $\frac{\partial\mu_t(x)}{\partial t}=L(S_t\mu(0,x))$ for $x\in S$. Let $\mu(0,x)=0$ if $x\in\{A,B\}$ since we do not start at a killed state at time zero. Then we have the following partial differential equations for $\mu(t,x)$.

$$\frac{\partial\mu_t(x)}{\partial t}=(L^{RW}\mu_t)(x)-\gamma_1\mu_t(x)-\gamma_2\xi(x)\mu_t(x) \qquad (x\in\mathbb{Z}^d) \tag{5.2}$$

$$\frac{\partial\mu_t(A)}{\partial t}=\sum_{x\in\mathbb{Z}^d}\gamma_1\mu_t(x) \tag{5.3}$$

$$\frac{\partial\mu_t(B)}{\partial t}=\sum_{x\in\mathbb{Z}^d}\gamma_2\xi(x)\mu_t(x) \tag{5.4}$$

Here, L^{RW} is the symmetric random walk on the lattice $\mathbb{Z}^d$ with migration rate α, for simplicity we assume 1 as this will only scale time t. We define the potential V as $V(x)=-\gamma_1-\gamma_2\xi(x)$ as an multiplicative operator. Also, we remark that the sum of $\partial_t^2\mu_t(x)$ (for all $x\in\mathbb{Z}^d$), $\partial_t^2\mu_t(A)$ and $\partial_t^2\mu_t(B)$ equals 0 since the particle always remains in the state space S.
For now we consider only one particle, so let the initial condition be a peak at position x':

$$\mu(0,x)=\mathbf{1}(x=x')$$

If we return to equation (5.2), we may write $\frac{\partial\mu_t(x)}{\partial t}$ as the sum of two operators on $\mu_t(x)$, namely:

$$\frac{\partial\mu_t(x)}{\partial t}=L^{RW}\mu_t(x)+V\mu_t(x)$$

It is clear that the solution for $\mu_t(x)$ is described by the following Feynman-Kac formula (equation (3.23)).

$$\mu_t(x)=e^{-\gamma_1 t}\mathbb{E}_x\left[e^{-\gamma_2\int_0^t\xi(X_s)\,\mathrm{d}s}\mathbf{1}(X_t=x')\right]=e^{-\gamma_1 t}\mathbb{E}_{x'}\left[e^{-\gamma_2\int_0^t\xi(X_s)\,\mathrm{d}s}\mathbf{1}(X_t=x)\right] \tag{5.5}$$

Since the starting position does not matter, we choose $x'=0$. Due to the indicator function $\mathbf{1}(X_t=0)$, we are interested in paths $x\to 0$ since only these matter. By symmetry of the random walk, we may change x for 0 and vice versa, by which we obtain the term to the right of equation (5.5).

For our solution of $\mu(t,x)$, we used a random walk on $\mathbb{Z}^d$ instead of S, therefore we cannot directly calculate $\mu(t,A)$ or $\mu(t,B)$ using the Feynman-Kac formula (equation (5.5)). However, we defined the

[8]We show these equations using Kolmogorov's backward equation, this is equivalent to the forward equation since S_t and L commute. However, when considering the approach taken, it is more logical to apply the forward equation (which is obtained by switching the operators S_t and L) since it evolves $\mu(0,x)$ from time zero, and not backwards in time.

transfer of $\mathbb{Z}^d \to A$ as $\gamma_1 \mu(t,x)$ and for $\mathbb{Z}^d \to B$ the transfer is $\gamma_2 \xi_x \mu(t,x)$. So for the total mass we consider all of $x \in \mathbb{Z}^d$.

$$\sum_{x\in\mathbb{Z}^d} \mu(x,t) = e^{-\gamma_1 t} \sum_{x\in\mathbb{Z}^d} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} \mathbf{1}(X_t = x) \right] = e^{-\gamma_1 t} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} \right] \tag{5.6}$$

$$\frac{\partial \mu(t,A)}{\partial t} = \gamma_1 \sum_{x\in\mathbb{Z}^d} \mu(t) = \gamma_1 e^{-\gamma_1 t} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} \right] \tag{5.7}$$

$$\tag{5.8}$$

Integrating over time t gives

$$\mu(t,A) = \gamma_1 \int_0^t e^{-\gamma_1 v} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^v \xi(X_s)\,\mathrm{d}s} \right] \mathrm{d}v \tag{5.9}$$

To find $\mu(t,B)$, use $\sum_{x\in\mathbb{Z}} \mu(t,x) + \mu(t,A) + \mu(t,B) = 1$ since the probability of remaining in S is constant.

As becomes clear, both $\mu(t,A)$ and $\mu(t,B)$ will depend on the so called 'total mass' $\sum_{x\in\mathbb{Z}^d} \mu(t,x)$. In the upcoming sections, we discuss how to find the total mass for different parameters of γ_1, γ_2 and p. This name can be attributed to the fact that we call $\mu_t(x)$ the probability mass flow. It is as if we split the mass of the particle (with total mass 1) and distribute it over multiple positions. Then, the probability is essentially the partial mass since the probability also sums to 1. For now, let $U(t)$ denote the total mass as given by the equation below.

$$U(t) = \sum_{x\in\mathbb{Z}^d} \mu(t,x) = e^{-\gamma_1 t} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} \right] \tag{5.10}$$

5.2 Hard-trap case

We will, before considering the general case, first consider when $\gamma_1 \to 0$ and $\gamma_2 \to \infty$, by which we mean the *hard*-trap case. Instead of an exciton under influence of a potential, we can describe the exciton as being 'killed' when stepping on a trap, hence the namesake. This can be seen when we rewrite the total mass $U(t)$ in this situation.

$$U(t) = \mathbb{E}_0 \left[e^{-\infty \cdot \int_0^t \xi(X_s)\,\mathrm{d}s} \right] \tag{5.11}$$

This is abuse of notation, but does describe what happens in the limit $\gamma_2 \to \infty$. If the particle with trajectory X_t finds a trap, such that $\xi(X_t) = 1$, we 'kill' the particle where the total mass becomes zero. This is due to $\int_0^t \xi(X_s)\,\mathrm{d}s \neq 0$ so

$$\lim_{\gamma_2\to\infty} e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} = 0$$

Otherwise, if the particle has not 'touched' a trap until time t, we have $\int_0^t \xi(X_s)\,\mathrm{d}s = 0$, which gives

$$\lim_{\gamma_2\to\infty} e^{-\gamma_2 \cdot 0} = 1$$

There is a special technique applicable in this extreme scenario. In essence, since $e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s}$ only takes values 0 or 1 whether it has found a trap, we can think of this random variable having a Bernoulli distribution. Again, using the abuse of notation earlier, we can intuitively understand $e^{-\infty \int_0^t \xi(X_s)\,\mathrm{d}s}$ as

$$\mathbb{P}(e^{-\infty \int_0^t \xi(X_s)\,\mathrm{d}s} = 0) = 1 - q \qquad (X_t \text{ hit trap})$$

$$\mathbb{P}(e^{-\infty \int_0^t \xi(X_s)\,\mathrm{d}s} = 1) = q \qquad (X_t \text{ still alive})$$

Where q is the probability of surviving and similarly $1-q$ is the probability of dying. This probability q can be described to depend on t, but in this case we will ignore t and consider the spatial region of the particle. Let us assume the particle is placed in some bounded region which is trap-free. Then for any time it spends within this boundary, it will always survive for any time t. Since it is very unlikely that a particle will remain in some bounded region for very large t, we will have to increase the size of the region as a function of time. For a random walk and Brownian motion, the spread of the trajectory will roughly depend on $\sqrt{t}$ due to a diffusion-like process[9].

The first step is to define this boundary. Let this boundary be $[-R,R]^d$ where $R \geq 0$ describes the size of the boundary. The probability of not finding a trap within $[-R,R]^d$ is

$$\mathbb{P}(\text{no trap in } [-R,R]^d) = (1-p)^{(2R)^d} = e^{(2R)^d \log(1-p)} \tag{5.12}$$

Recall that p is the probability of placing a trap at some site in $\mathbb{Z}^d$.

With the region defined, we go to the second step. Some trajectory X_t will survive when it is in the box $[-R,R]^d$ and the box is trap-free. So, we are interested in the probability of being within the box $[-R,R]^d$. As will be shown in the next section 'Box expansion', this probability can be described as

$$\mathbb{P}(X_t \text{ stays in } [-R,R]^d) \simeq e^{-c_1 t R^{-2}} \tag{5.13}$$

We remark by $\simeq$ we mean logarithmic equivalence, see section 4.1.2.

Example 5.1. *(Brownian motion)* As stated earlier, the standard deviation of B_t is $\sqrt{t}$. So the boundary where we will find the Brownian motion will be some scaled $[-\sigma(t),\sigma(t)]$. If we then choose $R(t) = \sqrt{t}$ for equation (5.13), we expect it to be constant. Since B_t is normally distributed with variance t, we can state that

$$\mathbb{P}(B_t \in (-\sqrt{t},\sqrt{t})) = \text{const}\,(\approx 0.68) \tag{5.14}$$

Substituting $R(t) = \sqrt{t}$ gives $\exp(\frac{-t}{t}) = \exp(-1)$, which indeed does not depend on t.

As the final step, we find that the probability of surviving will be

$$\begin{aligned}
\mathbb{P}(X_t \text{ survives}) &\approx \mathbb{P}(X_t \in [-R(t),R(t)]^d) \cdot \mathbb{P}([-R(t),R(t)]^d \text{ trap free}) && (5.15)\\
&\simeq e^{\log(1-p)R(t)^d} \cdot e^{c_1 t R(t)^{-2}} && (5.16)\\
&\simeq \exp(c_2 R(t)^d + c_1 t R(t)^{-2}) && (5.17)
\end{aligned}$$

Where we applied equations (5.12) and (5.13).

To make an accurate guess for the survival probability, we need to make sure $R(t)$ is chosen appropriately. If $R(t)$ increases to fast, we 'generate' a too large trap-free zone where most of the trajectories will not be present. Otherwise, if $R(t)$ increases too slowly, most trajectories will have escaped the trap-free box. So, we require that both contributions by equations (5.12) and (5.13) have similar behavior for t. This is equivalent to

$$c_2 R(t)^d = c_1 t R(t)^{-2}$$

We ignore constants c_1 and c_2. The ansatz $R(t) = t^k$ gives

$$\begin{aligned}
t^{kd} &= t \cdot t^{-2k}\\
t^{kd} &= t^{1-2k}\\
kd &= 1-2k\\
k &= \frac{1}{d+2}
\end{aligned} \tag{5.18}$$

[9]For Brownian motion, this is clear since we have $B_t - B_s \sim \mathcal{N}(0, t-s)$. So the 'spread' of the Brownian motion is equivalent to the standard deviation, which is $\sqrt{t-s}$.

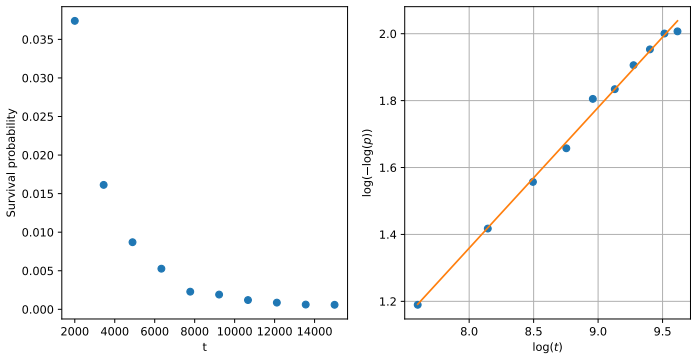


Figure 8: Survival probability calculated for different times for 5000 to 37500 simulations (increasing with t). This is done by simulating a random walk in 1 dimension until a time t is approached. The survival probability is found by calculating which simulations survived and which hit a trap ($p = 0.05$). Left figure shows the averaged survival probability for different values of t. The right figure shows the expected behavior of $\exp(t^\alpha)$ for some constant α, which is determined by the slope.

As such, we have found that $R(t) = t^{\frac{1}{d+2}}$ and that the survival probability takes the form

$$\mathbb{P}(X_t \text{ survives}) \simeq \exp\left(c_3 t^{\frac{d}{d+2}}\right) \tag{5.19}$$

Where c_3 is some constant originating from c_1 and c_2. We remark that $R(t) = t^{\frac{1}{d+2}}$ increases slower than $\sqrt{t} = t^{\frac{1}{2}}$, while we expect Brownian motion and the random walk to act diffusely. This can be attributed to the fact that the hard traps 'slow down' the overall mass spread. Every trap a trajectory comes into contact with is annihilated, so the total mass loses some mass at certain points which would otherwise have spread out further.

Example 5.2. *(Simulation)* We can simulate a 1-dimensional random walk, so it is possible numerically calculate the survival probability equivalent to $\mathbb{E}_0\left[e^{-\gamma \int_0^t \xi(X_s)\,\mathrm{d}s}\right]$. From equation (5.19), we expect that the survival probability, which we will denote by P, follows the law

$$P(t) \approx \exp\left(-c_3 \cdot t^{\frac{1}{3}}\right)$$

Given that we know $P(t)$, we check if the subexponential formula above holds by calculating

$$\log(-\log[P(t)]) = \log\left(c_3 t^{1/3}\right) = \frac{1}{3}\log t + c_4$$

We can analyze $\log(-\log P)$ by applying a least-squares fit for a linear function with respect to $\log t$. An example of a simulation is visible in figure 8. The obtained slope is approximately 0.4204. While the calculated slope deviates somewhat from the expected value of $\frac{1}{3}$, it becomes clear that we do obtain a relatively straight line. That the slope is not equal to $\frac{1}{3}$ has multiple arguments.

1. We state that $P(t) \simeq \exp(c_3 t^{1/3})$, so there can exist other smaller terms that are ignored.
2. We optimized $R(t)$ for the same growth rate by ignoring any constants. It is possible that the 'box expansion' contribution, which is $\exp(-tR(t)^{-2})$, provides more heavily than the 'potential' contribution.

In the next section, we will discuss the case for γ_2 finite (the 'soft'-trap case) where we give more detail about how the probabilities given by equations (5.12) and (5.13) are derived.

5.3 Soft-trap case

5.3.1 Approach

We now build upon the previous section 'Hard traps' to consider the more general case where $\gamma_2 < \infty$. We can ignore γ_1 for now, since this is simply an extra factor $e^{-\gamma_1 t}$, as becomes clear given the equation for the total mass.

$$U(t) = \sum_{x \in \mathbb{Z}^d} \mu(t,x) = e^{-\gamma_1 t} \mathbb{E}_0 \left[e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s} \right] \tag{5.20}$$

This is the case when we consider the quenched setting, where the traps are fixed in place. We are interested in averaging over all possible trap combinations. So in the annealed setting, we define the annealed total mass $\langle U(t) \rangle = \mathbb{E}_\xi[U(t)]$ where $U(t)$ is given above (equation (5.19)).

The approach is similar as stated in the 'hard traps' section. In these sections, we define $Z_t = e^{-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s}$ which is a random variable depending on the process X_t. We list the steps below.

1. **(Approximation by bounded region)** By our approach, we consider the spatial element extensively. As such, for a ball of radius R, which is

$$B(x,R) = \{y \in \mathbb{F}^d, \|x-y\|_2 < R\} \qquad\qquad (\text{for } \mathbb{F} = \mathbb{R}, \mathbb{Z})$$

We distinguish between trajectories $(X_s)_{s\in[0,t]}$ which leave the ball $B(0,R(t))$ and stay within the ball. Conditioning between these cases gives the following expression.

$$\mathbb{E}_0[Z_t] = \mathbb{E}_0[Z_t \,|\, X_t \in B(0,R)]\mathbb{P}(X_t \in B(0,R)) + \mathbb{E}_0[Z_t \,|\, X_t \notin B(0,R)]\mathbb{P}(X_t \notin B(0,R)) \tag{5.21}$$

It depends on how $R = R(t)$ is chosen, but we can assume $\mathbb{P}(X_t \notin B(0,R))$ to be small if R is large. This is due to the high cost for a trajectory to travel to the boundary in time t. For certain values of R, we refer to Lemma 5.1 at the end of this section which gives an upper bound on the probability of leaving the ball [10]. As such, we ignore the second term $\mathbb{P}(X_t \notin B(0,R))$. Then, there are two remaining parts, which we label under the terms.

$$\mathbb{E}_0[Z_t] \approx \underbrace{\mathbb{E}_0[Z_t \mathbf{1}_{B(0,R)}(X_t)]}_{\text{Potential contribution}} \times \underbrace{\mathbb{P}(X_t \in B(0,R))}_{\text{Box expansion}} \tag{5.22}$$

The term $\mathbb{E}_0[Z_t \mathbf{1}_{B(0,R)}(X_t)]$ is the part which contributes the potential to the total mass. In the hard trap section, we simplified this to be the survival probability $(1-p)^{(2R)^d}$, see equation (5.19). The other term, $\mathbb{P}(X_t \in B(0,R))$, represents the probability of remaining in the ball $B(0,R)$. To be consistent in naming, we call this the 'box expansion' term while we technically use a sphere.

2. **(Box expansion)** The approach to find the 'box expansion' term in equation (5.22) will be to use the large deviation principle. From the occupation time $L_t(\cdot) = \frac{1}{t}\int_0^t \mathbf{1}(X_s \in \cdot)\,\mathrm{d}s$, defined by Definition 4.13, we can state that $L_t(B(0,R)) = 1$ if and only if $X_t \in B(0,R)$. Of course, if $X_t \in B(0,R)$ then $\mathbf{1}_{B(0,R)}(X_s) = 1$ and so $L_t(B(0,R)) = \frac{1}{t}\int_0^t 1\,\mathrm{d}s = 1$. For the converse, if $L_t(A) = 1$ then $\int_0^t \mathbf{1}_A(X_s)\,\mathrm{d}s = t$. Since we have a indicator function with height 1, we only obtain this equality if $X_s \in A$ for $s \in [0,t]$ (this can be shown by contradiction).

[10] In our results, we get $R(t) = t^{1/d}$. If we choose a large factor for R, then it will give a meaningful bound up to some certain time t, as t increases faster than $t^{1/(d+2)}$ for large t. If we assume that $R(t)$ roughly moves with the total mass, we can choose a large factor since we are only interested in the asymptotics (behavior of $t^{1/(d+2)}$). Even for large R and large t, by our chosen R, we can still assume the probability of escaping is negligible.

We then continue with the rate function $I(g) = \frac{1}{8}\int \frac{(g'(x))^2}{g(x)}\,\mathrm{d}x$ as given by equation (4.51) for Brownian motion. Next, we will relate this rate function to the principal (largest) eigenvalue λ_1 of $L = \frac{1}{2}\frac{d^2}{dx^2}$. Then, we only consider functions which exist within $B(0,R)$ since we ignore all trajectories outside the ball. As it turns out, using this Dirichlet boundary condition gives the relation $\lambda_1 \propto \frac{1}{R^2}$, which we call the eigenvalue rescaling factor. This will give the result of $\mathbb{P}(X_t \in B(0,R)) \simeq e^{c_1 t R(t)^{-2}}$ as found in the hard-trap case (see equation (5.13)).

3. **(Potential contribution)** For the 'potential contribution' term in equation (5.22), we do not solve this analytically. Instead, we first consider γ_2 in extreme cases, when $\gamma_2 \to \infty$ (hard traps) and $\gamma_2 \to 0$ (no traps). These cases give different solutions for $R(t)$, so we are interested in when this happens specifically. We use an approximation to find a transition equation, which we compare to simulations of this contribution case.

Regarding the approximation of (1) given above, we substantiate this bound as follows.

Lemma 5.1. (Bound X_t) Let $(X_t)_{t\geq 0}$ be a continuous symmetric random walk in $\mathbb{Z}^d$ where $X_0 = 0$. Then we have the upper bound

$$\mathbb{P}_0(X_{[0,t]} \not\subset B(0,R)) \leq 2^{d+1}\exp\left\{-R\log\frac{R}{td} + R\right\} \tag{5.23}$$

Proof. The proof is derived from Lemma 4.3 of [GM98]. An outline of the proof for a box $(-R,R)^d$ is as follows (so this statement also holds for a ball $B(0,R) \subset (-R,R)^d$). Let $X_t = (x^1(t),\ldots,x^d(t))$ be the vector of d components labelled $x^n(t)$. We have

$$\begin{aligned}
\mathbb{P}_0(X_{[0,t]} \not\subset (-R,R)^d) &= \mathbb{P}(\max_{s\in[0,t]} \|X_t\|_1 \geq R) && (5.24)\\
&= \mathbb{P}(\max_{s\in[0,t]} (|x^1(s)| + \cdots + |x^d(s)|) \geq R) && (5.25)\\
&\leq 2^d \mathbb{P}(\max_{s\in} (x^1(s) + \cdots + x^d(s)) \geq R), && \text{(symmetry)} \quad (5.26)\\
&\leq 2^{d+1}\mathbb{P}(x^1(s) + \ldots x^d(s) \geq R), && \text{(reflection principle)} \quad (5.27)\\
&\leq 2^{d+1} e^{-\beta R}\mathbb{E}_0\left[e^{\beta x^1(t)}\right]^d, && \text{(Chebyshev's exp. inequality)} \quad (5.28)\\
&\leq 2^{d+1} e^{-\beta R} e^{2td(\cosh(\beta)-1)} && (5.29)
\end{aligned}$$

Assume $R > td$, if $R \leq td$ then $-R\log\frac{R}{td} + R > 0$ so $2^{d+1}\exp(-R\log\frac{R}{td} + R) \geq 1$, then the statement is always true. Choose $\beta = \log\frac{R}{td}$. Then we have

$$2td[\cosh(\beta) - 1] = td\left(\frac{R}{td} + \frac{td}{R} - 1\right) = R + td\left(\frac{td}{R} - 1\right) \leq R$$

This yields the inequality

$$\mathbb{P}_0(X_{[0,t]} \not\subset (-R,R)^d) \leq 2^{d+1}\exp\left(-R\log\frac{R}{td} + R\right)$$

□

5.3.2 Box expansion

This section is devoted to proving Theorem 5.4. This result relies heavily on the Lemma 5.2 and 5.3 which do the heavy work for said theorem. While the theorem is proven for Brownian motion, we do motivate as to how this statement will hold for $\mathbb{Z}^d$. For example, the eigenvalue rescaling in $\mathbb{Z}$ can be shown to be $1/R^2$

by approximation, and for Lemma 5.3 we motivate the argument for $\mathbb{Z}^d$. Where possible, we give the proof of the d-dimensional case.

First, we show how the principal eigenvalue scales when it support changes depending on a radius R from radius 1.

Lemma 5.2. (Eigenvalue scaling, $\mathbb{R}^d$) Consider the operator

$$L = \Delta^d = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} \tag{5.30}$$

Let Λ_1 be its largest eigenvalue for eigenfunctions having support on the d-dimensional ball $B(0,1)$, and similarly Λ_R for eigenfunctions having support on $B(0,R)$ where $R > 0$. Then we have

$$\Lambda_R = \frac{\Lambda_1}{R^2} \tag{5.31}$$

Proof. Let $u : \mathbb{R}^d \to \mathbb{R}$ where $\mathbf{x} = (x_1, \ldots, x_d)^T$ denotes the d-dimensional real vector consisting of components $x_1, \ldots, x_d \in \mathbb{R}$. For clarity, if $Lu = \Lambda u$ then $\int_{\mathbb{R}^d} u(\mathbf{x}) \cdot (Lu)(\mathbf{x})\,\mathrm{d}\mathbf{x} = \Lambda \int_{\mathbb{R}^d} u^2(\mathbf{x})\,\mathrm{d}\mathbf{x}$, so if $\int_{\mathbb{R}^d} u^2(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$ we have $\int_{\mathbb{R}^d} u(\mathbf{x}) \cdot (Lu)(\mathbf{x})\,\mathrm{d}\mathbf{x} = \Lambda$. By the condition $\int_{\mathbb{R}^d} u^2\,\mathrm{d}\mathbf{x} = 1$, u (and u') must vanish at infinity. Let us assume u is measurable and $\int_{\mathbb{R}^d} u^2\,\mathrm{d}\mathbf{x} = 1$. Then we can use Fubini's theorem to switch individual components since we integrate over $\mathbb{R}^d$. Let $\mathrm{d}\mathbf{x} = \mathrm{d}x_1 \cdots \mathrm{d}x_d$ and $\mathrm{d}\tilde{x}^i = \cdots \mathrm{d}x_{i-1}\,\mathrm{d}x_{i+1} \cdots$ (in other words, we do not integrate over x_i). We can rewrite this integral as follows.

$$\begin{aligned}
\int_{\mathbb{R}^d} u(\mathbf{x})(Lu)(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \int_{\mathbb{R}^d} u(\mathbf{x}) \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} u(\mathbf{x})\,\mathrm{d}\mathbf{x} && (5.32)\\
&= \sum_{i=1}^{d} \int_{\mathbb{R}^d} u(\mathbf{x}) \frac{\partial^2}{\partial x_i^2} u(\mathbf{x})\,\mathrm{d}\mathbf{x} && \text{(finite sum)} \quad (5.33)\\
&= \sum_{i=1}^{d} \int_{\mathbb{R}^{d-1}} \int_{-\infty}^{\infty} u(\mathbf{x}) \frac{\partial^2}{\partial x_i^2} u(\mathbf{x})\,\mathrm{d}x_i\,\mathrm{d}\tilde{x}^i && \text{(apply Fubini)} \quad (5.34)\\
&= \sum_{i=1}^{d} \int_{\mathbb{R}^{d-1}} \left([u \partial_{x_i} u]_{x_i=-\infty}^{x_i=\infty} - \int_{-\infty}^{\infty} (\partial_{x_i} u)^2\,\mathrm{d}x_i \right) \mathrm{d}\tilde{x}^i && \text{(integration by parts)} \quad (5.35)\\
&= \sum_{i=1}^{d} \int_{\mathbb{R}^d} -(\partial_{x_i} u)^2\,\mathrm{d}\mathbf{x} && \text{(vanishing limit)} \quad (5.36)\\
&= -\int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} u)^2\,\mathrm{d}\mathbf{x} && (5.37)
\end{aligned}$$

So for $Lu = \Lambda u$ we obtain $\int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} u)^2\,\mathrm{d}\mathbf{x} = -\Lambda = \lambda$ if $\int_{\mathbb{R}^d} u^2\,\mathrm{d}\mathbf{x} = 1$. Let us the define λ_1 and λ_R.

$$\begin{aligned}
\lambda_1 &= \inf_u \left\{ \int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} u)^2\,\mathrm{d}\mathbf{x} \;\middle|\; \int_{\mathbb{R}^d} u^2\,\mathrm{d}\mathbf{x} = 1, \mathrm{supp}(u) \subseteq B(0,1) \right\}\\
\lambda_R &= \inf_v \left\{ \int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} v)^2\,\mathrm{d}\mathbf{x} \;\middle|\; \int_{-\infty}^{\infty} v^2\,\mathrm{d}x = 1, \mathrm{supp}(v) \subseteq B(0,R) \right\}
\end{aligned}$$

First, we show that $\lambda_R \geq \frac{\lambda_1}{\lambda_R}$. Let v be some function with $\mathrm{supp}(v) \subseteq B(0,R)$. Define the rescaled version of v as

$$u_R(x) = R^{d/2} v(Rx)$$

Then $\text{supp}(u_R) \subseteq B(0,1)$. We will find $\partial_{x_i} v$ and rewrite $\int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{x_i} v)^2 \, \mathrm{d}\mathbf{x}$ as a function of $\int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{x_i} u_R)^2 \, \mathrm{d}\mathbf{x}$, which we can bound by λ_R. From its definition, $v(\mathbf{x}) = R^{-d/2} u_R(\mathbf{x}/R)$, so from the chain rule

$$\partial_{x_i} v(\mathbf{x}) = R^{-d/2} \frac{1}{R} \partial_{x_i/R} u_R(\mathbf{x}/R)$$

Now, we will integrate $\sum_i (\partial_{x_i} v)^2$ to find a bound using λ_1.

$$\begin{aligned}
\int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{x_i} v)^2 \, \mathrm{d}\mathbf{x} &= \int_{\mathbb{R}^d} \sum_{i=1}^d \frac{1}{R^2} R^{-d} (\partial_{x_i/R} u_R(\mathbf{x}/R))^2 \, \mathrm{d}\mathbf{x} \\
&= \frac{1}{R^2} \int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{y_i} u_R(\mathbf{y})) \, \mathrm{d}\mathbf{y} && (\text{let } y_i = x_i/R) \\
&\geq \frac{\lambda_1}{R^2}
\end{aligned}$$

Taking the infimum on both sides gives the equality $\lambda_R \geq \frac{\lambda_1}{R^2}$. Now, we show that $\lambda_R \leq \frac{\lambda_1}{R^2}$ using the same technique. Let u be some function with $\text{supp}(u) \subseteq B(0,1)$, then the rescaled v_R version has $\text{supp}(v_R) \subseteq B(0,R)$ as defined by

$$\begin{aligned}
v_R(\mathbf{x}) &= R^{-d/2} u(\mathbf{x}/R) \\
u(\mathbf{x}) &= R^{d/2} v_R(R\mathbf{x})
\end{aligned}$$

Then its derivative with respect to x_i will be $\partial_{x_i} u(\mathbf{x}) = R^{d/2} R \partial_{Rx_i} v_R(Rx)$. Integrating $\sum_i (\partial_{x_i} u)^2$ over $\mathbb{R}^d$ yields

$$\begin{aligned}
\int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{x_i} u)^2 \, \mathrm{d}\mathbf{x} &= R^2 \int_{\mathbb{R}^d} \sum_{i=1}^d R^d (\partial_{Rx_i} v_R(Rx)) \, \mathrm{d}\mathbf{x} \\
&= R^2 \int_{\mathbb{R}^d} \sum_{i=1}^d (\partial_{y_i} v_R(\mathbf{y})) \, \mathrm{d}\mathbf{y} && (\text{let } y_i = Rx_i) \\
&\leq R^2 \lambda_R
\end{aligned}$$

Again, taking the infimum on both sides gives $\lambda_1 \leq R^2 \lambda_R$ or $\frac{\lambda_1}{R^2} \leq \lambda_R$. Since $\Lambda_R = -\lambda_R$ and $\Lambda_1 = -\lambda_1$, we have shown $\Lambda_R = \frac{\Lambda_1}{R^2}$.

□

Note 5.1. We use the operator $L = \Delta^d$ instead of $L = \frac{1}{2}\Delta^d$ for Brownian motion. However, the factor $\frac{1}{2}$ only rescales the eigenvalues by $\frac{1}{2}$, so the statement still holds.

Secondly, we derive how we can find the probability of Brownian motion staying in a d-dimensional sphere. We apply the large deviation principle and its results we found in the section to the occupation times to find the wanted probability.

Lemma 5.3. (Brownian motion in sphere) Let $(B_s)_{s \geq 0}$ be d-dimensional Brownian motion. Then for $R > 0$ it holds

$$\mathbb{P}(B_{[0,t]} \subseteq B(0,R)) \simeq \exp\left(-t \inf_{g \in \mathcal{F}_0^d} \{ I(g) \,|\, \text{supp}(g) \subseteq B(0,R) \} \right) \tag{5.38}$$

Where $I(g)$ is the rate function given by Proposition 4.14.

Proof. We do not give a complete proof but give an outline of the proof. In equation (4.51) it has been shown that L_t has a rate function $I(g) = \frac{1}{8}\int_{-\infty}^{\infty} \frac{[g'(x)]^2}{g(x)}\,\mathrm{d}x$ where g is a probability measure (in the sense that $\int_{-\infty}^{\infty} g(x)\,\mathrm{d}x = 1$) and $g(x) > 0$ for all $x \in \mathbb{R}$ as required by set $\mathcal{F}_0$ (equation (4.36)). Then for the d-dimensional case we have the sum of all individual rate functions as given by Proposition 4.14.

$$I(g) = \frac{1}{8}\sum_{i=1}^{d}\int_{\mathbb{R}^d} \frac{[\partial_{x_i} g]^2}{g}\,\mathrm{d}\mathbf{x} \tag{5.39}$$

Let us assume that both L_t and g are measures that take singleton sets, in other words $L_t(\mathbf{x}) = g(\mathbf{x})$. Now, we consider all possible $g : \mathbb{R}^d \to \mathbb{R}$ under the family $\mathcal{F}_0^d$. Then in general we have

$$\mathbb{P}(L_t(\mathbf{x}) = g(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^d) \simeq e^{-tI(g)} \tag{5.40}$$

Note that $L_t(B(0,R)) = 1$ and $B_{[0,t]} \subseteq B(0,R)$ are equivalent. Since we consider $L_t(B(0,R)) = 1$, we require that $\int_{B(0,R)} g(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$ so $\mathrm{supp}(g) \subseteq B(0,R)$. The requirement $L_t(B(0,R)) = 1$ can be satisfied by many different $g \in \mathcal{F}_0^d$ and $\mathrm{supp}(g) \subseteq B(0,R)$. We then postulate

$$\mathbb{P}(L_t(B(0,R) = 1) \simeq \exp\left(\sup_{g\in\mathcal{F}_0}\{-tI(g) \,|\, \mathrm{supp}(g) \subseteq B(0,R)\}\right) \tag{5.41}$$

By Laplace's principle, which we explain as follows. If we consider some function $f : A \to (0,\infty)$ where A is some arbitrary set, then if

$$\sum_{x\in A} e^{-tf(x)} < \infty$$

We expect that for large t, we get

$$\sum_{x\in A} e^{-tf(x)} \underset{t \text{ large}}{\approx} e^{-t\sup_{x\in A} f(x)}$$

That we assume by the same reasoning of Theorem 4.15, which essentially proved this statement for A countable (given that there was an element $y \in A$ where $\sup_{x\in A} f(x) = f(y)$). However, the set $\mathcal{F}_0^d$ can be uncountable, so the theorem is not directly applicable. We remark that if we consider a continuous symmetric random walk on $\mathbb{Z}$ (or similarly $\mathbb{Z}^d$), the set $B(0,R)$ will be bounded, so a function $g(x) : \mathbb{Z} \to \mathbb{R}_+$ with $\mathrm{supp}(g) \subseteq B(0,R)$ will only exist on a finite amount of points. As such, if $B(0,R)$ has $n < \infty$ points, we can represent g as a vector of $\mathbb{R}^n$. It has been shown[11] that Laplace's principle does hold when integrating over $\mathbb{R}^n$, so in the case of $\mathbb{Z}^d$ the argument of applying Laplace's method as described here is stronger than for Brownian motion (however, we do assume it to be true).
By equation (5.41), taking $-t$ out of the supremum (changing sup for inf) and substituting $L_t(B(0,R)) = 1$ for $B_{[0,t]} \subseteq B(0,R)$ gives the wanted result. □

With the above lemmas shown, we now combine both results in the following final theorem. We remark the distinction between $\Lambda_{\max}$ and $\lambda_{\max}$ used in the following theorem. By Proposition 4.18, the eigenvalue $\Lambda_{\max}$ should be nonpositive, hence $\lambda_{\max} = -\Lambda_{\max} \geq 0$. This results in a negative coefficient in the exponent. Also, we define $\Lambda_{\max}$ and $\lambda_{\max}$ in terms of $B(0,1)$ so that these values will be constants instead of being a function of R.

Theorem 5.4. (Box expansion term) Let $(B_t)_{t\geq 0}$ be d-dimensional (for now, let $d = 1$) Brownian motion having generator

$$L = \frac{1}{2}\Delta^d = \frac{1}{2}\sum_{j=1}^{d}\frac{\partial^2}{\partial x_j^2} \tag{5.42}$$

Define the principal eigenvalue $\Lambda_{\max}$ of L on $B(0,1)$ as

[11] See [DZ10], section 4.3 about Varadhan's lemma.

$$\Lambda_{\max} = \sup_u \left\{ \int_{\mathbb{R}^d} u(\mathbf{x}) \cdot (Lu)(\mathbf{x}) \, \mathrm{d}\mathbf{x} \,\middle|\, \int_{\mathbb{R}^d} u^2 \, \mathrm{d}\mathbf{x} = 1, \mathrm{supp}(u) \subseteq B(0,1) \right\} \tag{5.43}$$

$$-\Lambda_{\max} = \lambda_{\max} = \inf_u \left\{ -\int_{\mathbb{R}^d} u(\mathbf{x}) \cdot (Lu)(\mathbf{x}) \, \mathrm{d}\mathbf{x} \,\middle|\, \int_{-\infty}^{\infty} u^2 \, \mathrm{d}\mathbf{x} = 1, \mathrm{supp}(u) \subseteq B(0,1) \right\} \tag{5.44}$$

Then for $R > 0$ the following holds.

$$\mathbb{P}(B_{[0,t]} \subseteq B(0,R)) \simeq e^{-tR^{-2}\lambda_{\max}} \tag{5.45}$$

Proof. Our process will be as follows. First, we rewrite $\lambda_{\max}$ to an intermediate form using the first derivatives (in other words, sums of $(\partial_{x_i} h)^2$) and then apply Lemma 5.2 to scale the eigenvalue to the ball $B(0,R)$. The second step revolves around rewriting $I(g)$ to the same intermediate form, then by Lemma 5.3 we get the wanted relation.
As has been shown using equation (5.32) in the proof of Lemma 5.2, we can write

$$\lambda_{\max} = \inf_u \left\{ \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} u)^2 \, \mathrm{d}\mathbf{x} \,\middle|\, \int_{\mathbb{R}^d} u^2 \, \mathrm{d}\mathbf{x} = 1, \mathrm{supp}(u) \subseteq B(0,1) \right\} \tag{5.46}$$

Then, applying Lemma 5.2 to $-\lambda_{\max}$, we get a rescaled version of $\lambda_{\max,R}$ which we define below.

$$\lambda_{\max,R} = \inf_u \left\{ \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} u)^2 \, \mathrm{d}\mathbf{x} \,\middle|\, \int_{\mathbb{R}^d} u^2 \, \mathrm{d}\mathbf{x} = 1, \mathrm{supp}(u) \subseteq B(0,R) \right\}$$

$$\lambda_{\max,R} = \frac{\lambda_{\max}}{R^2} \qquad \text{(Lemma 5.2)}$$

Now, we move to the second step. We know from Proposition 4.14 and Lemma 5.3 that the relevant rate function $I(g)$ is given by

$$I(g) = \frac{1}{8} \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{[\partial_{x_i} g]^2}{g} \, \mathrm{d}\mathbf{x} \tag{5.47}$$

For some $g \in \mathcal{F}_0^d$. Since $g > 0$ and $\int_{\mathbb{R}^d} g \, \mathrm{d}\mathbf{x} = 1$, if we let $h^2 = g$ or $h = \sqrt{g}$ then $\int_{\mathbb{R}^d} h^2 \, \mathrm{d}\mathbf{x} = 1$. Using this substitution in the above rate function (equation (5.47)) yields

$$I(g) = \frac{1}{8} \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{[\partial_{x_i} g]^2}{g} \, \mathrm{d}\mathbf{x} \tag{5.48}$$

$$I(h) = \frac{1}{8} \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{[2h\partial_{x_i} h]^2}{h^2} \, \mathrm{d}\mathbf{x} \tag{5.49}$$

$$= \frac{1}{2} \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{h^2[\partial_{x_i} h]^2}{h^2} \, \mathrm{d}\mathbf{x} \tag{5.50}$$

$$= \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i=1}^{d} (\partial_{x_i} h)^2 \, \mathrm{d}\mathbf{x} \tag{5.51}$$

Relating this result to $\lambda_{\max,R}$ gives

$$\lambda_{\max,R} = \inf_h \left\{ I(h) \,\middle|\, \int_{\mathbb{R}^d} h^2 \, \mathrm{d}\mathbf{x} = 1, \mathrm{supp}(h) \subseteq B(0,R) \right\}$$

This is the same infimum we have for Lemma 5.3, so applying this lemma gives

$$\mathbb{P}(B_{[0,t]} \subseteq B(0,R)) \simeq \exp\left(-t \inf_{g \in \mathcal{F}_0^d}\{I(g) \,|\, \mathrm{supp}(g) \subseteq B(0,R)\}\right) = \exp(-t\lambda_{\max,R})$$

Then, by equation (5.46) we find the wanted result

$$\mathbb{P}(B_{[0,t]} \subseteq B(0,R)) \simeq \exp(-tR^{-2}\lambda_{\max})$$

□

5.3.3 Potential contribution

With the box expansion term covered, we now analyze the 'potential contribution' term, which we define as

$$\mathbb{E}_0\left[Z_t \mathbf{1}_{B(0,R)}(X_t)\right] \tag{5.52}$$

In our situation, we must take into account the presence of traps, so for the total mass we consider the annealed setting. The box expansion term is independent of the annealed or quenched setting (and so is the γ_1-term), so we get

$$\langle U(t)\rangle = \langle e^{-\gamma_1 t} \times \mathbb{E}_0[Z_t]\rangle \approx \underbrace{e^{-\gamma_1 t}}_{\gamma_1 \text{ contribution}} \times \underbrace{\langle \mathbb{E}_0[Z_t \mathbf{1}_{B(0,R)}(X_t)]\rangle}_{\text{Potential contribution}} \times \underbrace{\mathbb{P}(X_t \in B(0,R))}_{\text{Box expansion}} \tag{5.53}$$

Let us consider the situation where t and γ_2 are relatively large. Then the exponential part in Z_t will be drastically reduced to zero if the trajectory visits a trap. Given $t \to \infty$ and/or $\gamma_2 \to \infty$, the value $Z_t \mathbf{1}_{B(0,R)}(X_t)$ will be approximately zero if there exists a trap within $B(0,R)$ since for such large time the trajectory will eventually find the trap. So, the random variable $Z_t \mathbf{1}_{B(0,R)}(X_t)$ will take 0 if there is at least one trap within $B(0,R)$ and be 1 if there is no trap, approximately. This is equal to the hard trap case, where we approximated the potential contribution to be of the form $\log(1-p)^{(2R)^d}$. Similarly, we get

$$\langle \mathbb{E}_0[Z_t \mathbf{1}_{B(0,R)}(X_t)]\rangle \approx \exp(R^d \log(1-p)), \quad \text{(large } t \text{ approximation)} \tag{5.54}$$

Since the probability of finding a trap in a ball $B(0,R)$ will be approximately $(1-p)^{\mathrm{Vol}(B(0,R))} \approx (1-p)^{R^d}$. In this approximation, we find the following solution.

Proposition 5.5. (Total mass asymptotics, large t) Assume that the approximation given by equation (5.54) is true, which is when $t \to \infty$ and γ_2 large. The we have the following total mass asymptotics in the annealed setting for $\gamma_1 = 0$ and a certain constant $c_1 > 0$.

$$\langle U(t)\rangle \simeq \exp\left(-c_1 \cdot t^{\frac{d}{d+2}}\right) \tag{5.55}$$

Where $\simeq$ represents logarithmic equivalence with respect to the rate $t^{\frac{d}{d+2}}$.

Proof. We will give an outline of the proof using the tools given to us so far, but this is not sufficient enough to prove the statement completely. For a detailed proof, see refer to the article 'Asymptotics for the Wiener Sausage' [DV75b] by M.D. Donsker and S.R.S. Varadhan.
By Theorem 5.4 and equation (5.54), we obtain the following expression for the total mass $\langle U(t)\rangle$.

$$\langle U(t)\rangle \simeq \exp(-a \cdot R^d) \times \exp(-tR^{-2}\lambda_{\max})$$

We require both contributions to have the same speed. That is, we let $R = R(t)$ and make sure the exponentials have the same general form of $\exp(-c_1 f(t))$. Otherwise, the box expansion term will over- or underestimate the actual probability mass, delivering unrealistic asymptotic results. Using this ansatz, we find the following.

$$R(t)^d = tR(t)^{-2}$$

Let $R(t) = t^k$, solving for k yields

$$t^{kd} = t^{1-2k}$$

Thus $k = \frac{1}{d+2}$ thus $R(t) = t^{\frac{1}{d+2}}$. We then find $\langle U(t) \rangle \simeq \exp(-c_1 \cdot R(t)^d) = \exp(-c_1 \cdot t^{\frac{d}{d+2}})$. □

Now, we consider the other scenario where t is very small. In fact, we choose $t \ll \frac{1}{\gamma_2 + 2d}$, such that the likelihood of jumping from $x = 0$ to another place is unlikely. At this point, we are always within the ball $B(0, R)$ so we ignore the term $\mathbf{1}_{B(0,R)}(X_t)$. If there is a trap placed at $x = 0$, then the integral $\int_0^t -\gamma_2 \xi(X_s)\,\mathrm{d}s \approx -\gamma_2 t$. If there is no trap, then the integral is simply zero. In short, the integral will be small such that $Z_t \approx 1$. For this scenario, we experience no actual 'potential contribution' due to the small time scale. As such, we expect a diffusion-like process to occur. This is shown in the following statement.

Proposition 5.6. (Total mass asymptotics, small t) Assume for t that

$$t \ll \frac{1}{\gamma_2 + 2d} \tag{5.56}$$

Then we have diffusion-like behavior, in other words

$$R(t) = \sqrt{t} \tag{5.57}$$

Proof. Using the same procedure as in proposition 5.5, but approximate the potential contribution as 1. Then, the factor $\exp(-tR(t)^{-2}\lambda_{\max})$ should decrease as fast as the constant, which in turn requires $R(t)^2 = t$, yielding $R(t) = \sqrt{t}$. □

Now, consider the general case. Let us assume $b = \langle \mathbb{E}_0[Z_t \mathbf{1}_{B(0,R)}(X_t)] \rangle$. Since $Z_t = \exp(-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s)$, the range of b is $[0, 1]$. The particle will spends approximately pt time on a trap, so $\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s \approx -\gamma_2 pt$. Then we can consider time regions $t = \frac{-\log(b)}{\gamma_2 p}$ for different values of b.

For large t we assumed that the potential part was $(1-p)^{R^d}$ due to exponential decay from traps. However, as small γ_2 or t, we found that this approximation does not hold. As is visible in figure 9, for small amount of traps there is still some significant contribution. Let us consider an extreme example where $t \ll 1/(\gamma_2 + 2d)$ such there has not been a jump, then we have $\int_0^t -\gamma_2 \xi(X_s)\,\mathrm{d}s = -\gamma_2 t$ if a trap is present or 0 for no trap for the exponent. In short, we may simplify $e^{\int_0^t -\gamma_2 \xi(X_s)\,\mathrm{d}s}$ as 1, then the equating-powers approach would give $t^0 + t \cdot t^{-2p}$ such that $p = \sqrt{t}$ which is similar to diffusion.

Similarly, we can approximate for small t the time spent at traps as the fraction p (since the random walk is not affected by traps). Then $-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s \approx -\gamma_2 pt$. Since the expected value takes values $[0, 1]$, we may choose some $t = \frac{-\log(b)}{\gamma_2 p}$ where $b \in [0, 1]$ to find where the regime of $\mathbb{E}\left[e^{\int_0^t -\gamma_2 \xi(X_s)\,\mathrm{d}s}\right]$ lies. We visualize this using Monte-Carlo simulations as can be seen in figure 9.

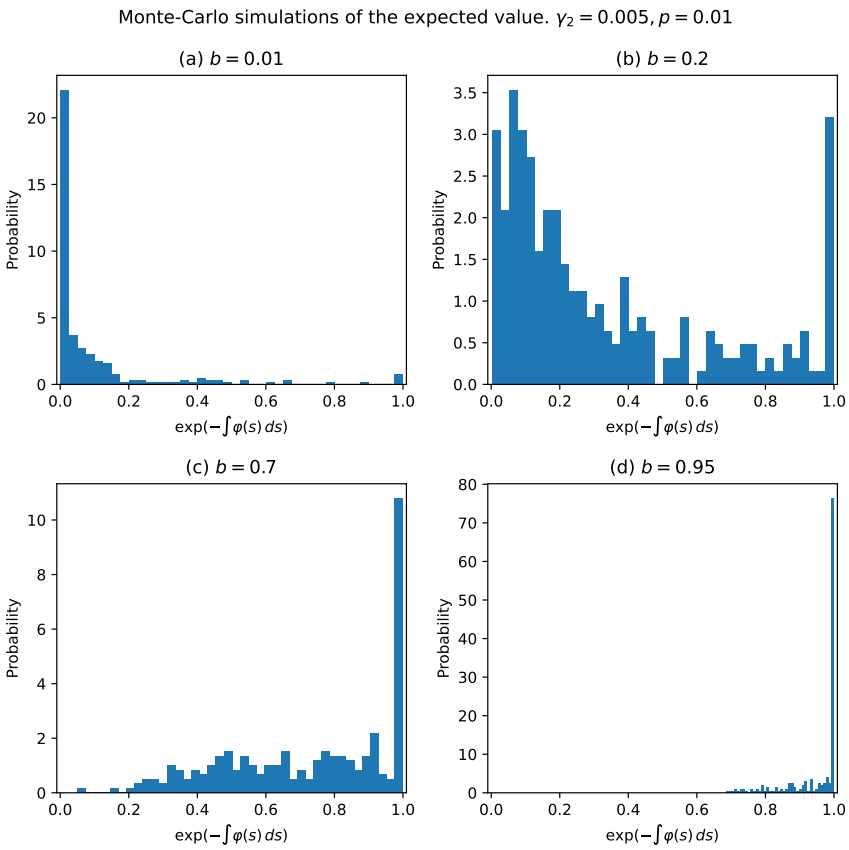


Figure 9: Distributions of $e^{\int_0^t -\gamma_2 \xi(X_s)\, \mathrm{d}s}$ (denoted $\exp(-\int_0^t \varphi(s)\, \mathrm{d}s)$) calculated from Monte-Carlo simulations. The max time of a simulation is set by $t = -\log(b)/(\gamma_2 p)$, where $\gamma_2 = 0.005, p = 0.01$. **(a,** $b = 0.01$**)** This is the $t \to \infty$ case, where $\mathbb{E}[e^{\int_0^t -\gamma_2 \xi(X_s)\, \mathrm{d}s}] \approx (1-p)^{R_d}$, which can be seen as the ratio of the distribution for the values 0 and 1. Since the values will become increasingly smaller, only the value 1 will remain, hence we can make the approximation. This is the $R(t) = t^{1/d}$ time regime as derived earlier. **(b,** $b = 0.2$**)** The value is now roughly distributed, and both approximations for small and large t are not applicable. This represents a "transition state" between $R(t) = t^{1/2}$ and $R(t) = t^{1/d}$. **(c,** $b = 0.7$**)** Slowly, it becomes more apparent that the value 1 is being taken, but still has a small spread-out distribution. Slowly as time increases, we start at the distribution (d) and arrive at (a), as this distribution suggests. **(d,** $b = 0.95$**)** This is the $t \ll 1/\gamma_2$-case as mentioned earlier where we approximated the expected value to be equal to 1. This is the $R(t) = \sqrt{t}$ (regular diffusion) time regime.

5.4 Connection to parabolic Anderson model

While we do not use the parabolic Anderson model, we mention this concept as reference material for future research. The parabolic Anderson model (PAM) describes a similar partial differential equation used in our model, using a random potential V and on $\mathbb{Z}^d$. In essence, it generalizes the concept of a random walk within a random potential. For this section, we refer to the book 'The Parabolic Anderson Model' by Wolfgang König [Kön16], which contains many cases of this model applied and gives details about total mass asymptotics. Let us define the parabolic Anderson model according König's book.

Definition 5.7. (Parabolic Anderson Model) Let $u(t,x) : [0,\infty) \times \mathbb{Z}^d \to [0,\infty)$ be the solution to the following problem:

$$\frac{\partial}{\partial t}u(t,x) = \Delta^d u(t,x) + \xi(t,x)u(t,x) \tag{5.58}$$
$$u(0,x) = \mathbf{1}(x=0) \tag{5.59}$$

The operator Δ^d is used to denote the Laplacian in the discrete d-dimensional case, which we define as follows:

$$\Delta^d f(x) = \sum_{y\in\mathbb{Z}^d, y\sim x} [f(y) - f(x)] \tag{5.60}$$

We remark that Δ^d coincides with the generator L for a symmetric random walk in $\mathbb{Z}^d$.

In a similar procedure as we have done, we can find a solution for u using the Feynman-Kac formula (Theorem 3.10). König provides a generalized result for total mass asymptotics when the potential is independent of position, which we give below. Using the cumulant generating function $H(t) = \log\langle e^{t\xi(0)}\rangle$, the total mass can also be described using the following theorem.

Theorem 5.8. (Moment asymptotics, König) In the case of $R(t) \to \infty$, we have

$$\langle U(t)\rangle = \exp\left(R(t)^d H(tR(t)^{-d})\right) \times \exp\left(-c\cdot tR(t)^{-2}\right) \tag{5.61}$$

Where c is a constant.

Proof. This is an adapted case of theorem 3.13 of [Kön16]. □

Since $H(t) = \log(pe^{-\gamma_2 t} + (1-p)) = \log(1 + p(e^{-\gamma_2 t} - 1))$, for large t we find $H(t) \approx \log(1-p)$. Substituting this in the above theorem yields $\exp(R(t)\log(1-p)) \times \exp(-c\cdot tR(t)^{-2})$, similar to what was found in proposition 5.5.

5.5 Many particle setting

Here we want to explain how to treat the setting of many independent particles. It will turn out, that when the particles start with a Poisson distribution, then at later times this Poisson distribution will persist, and the associated parameters evolve according to the Kolmogorov backward equation of a single particle. Therefore, for independent particles, our study of the single particle distribution in (5.6) is sufficient to understand the evolution in the multiparticle setting.

Up until now, we have studied the behavior of a single particle as a probability density. As will be in the real case, we have multiple particles moving along the lattice instead of one particle. In this section, we discuss how the distribution of particles at some site of the lattice is found. We assume that the particles do not interact with each other.

Assume that the amount of particles at a position $x \in \mathbb{Z}$ is Poisson distributed with parameter $\rho_x(0)$ at time $t = 0$. From point process theory, it is known that the distribution of particles always remains Poisson distributed for any time t, hence we consider this distribution. We will show that this distribution holds for any $t \geq 0$, given we only consider the random walk in 1-dimension without traps.

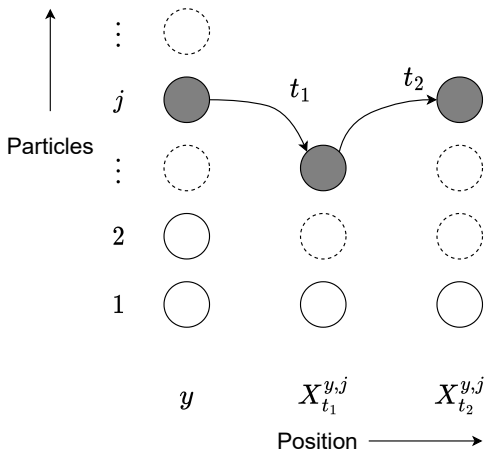


Figure 10: Visualization of $X_t^{y,j}$. Some particle starts at place y and is indexed as the jth particle. Then this particle moves to $X_{t_1}^{y,j}$ at some time $t_1 > t_0 = 0$. After that, it moves to $X_{t_2}^{y,j}$ at time $t_2 > t_1$.

Theorem 5.9. (Distribution of $\eta_x(t)$) Let the trajectory of each particle be described by a continuous random walk in $\mathbb{Z}$, being independent of each other. Define $\eta_x(t)$ as the amount of particles at a position $x \in \mathbb{Z}$ at a time t. Assume that $\eta_x(0) \sim \text{Poisson}(\rho_x(0))$ with parameter $\rho_x(t)$. Then $\eta_x(t)$ is also Poisson distributed, namely

$$\eta_x(t) \sim \text{Poisson}(\rho_x(t)) \tag{5.62}$$

The parameter $\rho_x(t)$ is then given by the semigroup of the random walk,

$$\rho_x(t) = S_t \rho_x(0) \tag{5.63}$$

Proof. To show this is the correct distribution, we need more detail about the parameter $\rho_x(t)$ for different times of t. For a Markov random walk, we use the notation $X_t^{y,j}$ for the path of the j-th particle starting at y (then at time zero), now at a time t. Furthermore, let $p_t(x,y)$ be the probability of $x \to y$ for a time t. Our approach to prove Theorem 5.9 is by evaluating the moments of $\eta_x(t)$. Let us consider the first moment, which is just the expected value of $\eta_x(t)$. We rewrite $\eta_x(t)$ as the sum of all paths of other starting particles, now currently at x.

$$\mathbb{E}\eta_x(t) = \mathbb{E}\left[\sum_{y\in\mathbb{Z}} \sum_{j=1}^{\eta_y(0)} \mathbf{1}(X_t^{y,j} = x)\right] \tag{5.64}$$

$$= \sum_{y\in\mathbb{Z}} \mathbb{E}\left[\sum_{j=1}^{\eta_y(0)} \mathbf{1}(X_t^{y,j} = x)\right] \tag{5.65}$$

Then use $\mathbb{E}(\mathbf{1}(X_t^{y,j} = x)) = p_t(x,y)$. Since we ultimately sum over $\eta_y(0)$ independent trajectories, this gives the expression for $\rho_x(t)$ as follows:

$$\mathbb{E}\eta_x(t) = \sum_y \rho_y(0) p_t(x,y) = \rho_x(t) \tag{5.66}$$

Proving $\eta_x(t) \sim \text{Poisson}(\rho_x(t))$ can be done by showing that the moments of $\eta_x(t)$ are powers of the parameters. More specifically, if its moment generating function (see Definition 4.1) is known. We can check if it is the same as the moment generating function of a Poisson distribution, since the moment generating function is unique. In this section, fix $x \in \mathbb{Z}$, $t \geq 0$ and let $M(l)$ be the moment generating function of $\eta_x(t)$ as given by

$$M(l) = \mathbb{E}[e^{l\eta_x(t)}]$$

For $X \sim \text{Poisson}(\lambda)$, we get the moment generating function

$$M_X(l) = \sum_{k=0}^{\infty} e^{lk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^l)^k}{k!} = e^{-\lambda + \lambda e^l} = e^{\lambda(e^l - 1)} \tag{5.67}$$

We will now show that $\eta_x(t) \sim \text{Poisson}(\rho_x(t))$ by showing that $M(l) = e^{\rho_x(t)(e^l-1)}$. Rewrite $M(l)$ in the same way as for equation (5.64) by considering the trajectory $X_t^{y,j}$.

$$M(l) = \mathbb{E}\left[e^{\eta_x(t)}\right] = \mathbb{E}\left[e^{l\sum_{y\in\mathbb{Z}}\sum_{j=1}^{\eta_y(0)} \mathbf{1}(X_t^{y,j}=x)}\right] \tag{5.68}$$

$$= \prod_{y\in\mathbb{Z}} \mathbb{E}\left(\prod_{j=1}^{\eta_y(0)} \underbrace{e^{l\mathbf{1}(X_t^{y,j}=x)}}_{A_j}\right) \tag{5.69}$$

Let $A_j = e^{l\mathbf{1}(X_t^{y,j}=x)}$, then A_j is independent of A_i if $i \neq j$ since the paths are independent of each other. So we have $\mathbb{E}[A_1 \cdot A_2 \cdots A_{\eta_y}] = \mathbb{E}[A_1] \cdots \mathbb{E}[A_{\eta_y}]$. Again, since the trajectories do not influence each

other, the expected value of A_j must be the same for any $j = 1, \ldots, \eta_y$. We take for example $j = 1$: $\mathbb{E}\prod_{j=1}^{\eta_y(0)} A_j = \mathbb{E}(\mathbb{E}(A_1)^{\eta_y(0)})$ [12] which can be substituted in equation (5.68):

$$M(l) = \prod_{y\in\mathbb{Z}} \mathbb{E}\left[\mathbb{E}\left(e^{l\mathbf{1}(X_t^{y,1}=x)}\right)^{\eta_y(0)}\right] \tag{5.70}$$

$$= \mathbb{E}\exp\left(\sum_y \eta_y(0)\underbrace{\log \mathbb{E}e^{l\mathbf{1}(X_t^{y,1}=x)}}_{m}\right) \tag{5.71}$$

Use $\eta_y(0) \sim \text{Poisson}(\rho_y(0))$ and the moment generating function of $\eta_y(0)$ to rewrite $\mathbb{E}e^{\eta_y(0)m} = \exp\left(\rho_y(0)(e^l - 1)\right)$ where $m = \log \mathbb{E}e^{lI(X_t^{y,1}=x)}$. From independence of $\eta_y(0)$ we can substitute this result into equation (5.70).

$$M(l) = \exp\left(\sum_y \rho_y(0)(e^m - 1)\right) = \exp\left(\sum_y \rho_y(0)(\mathbb{E}e^{l\mathbf{1}(X_t^{y,1}=x)} - 1)\right) \tag{5.72}$$

Take -1 into the expected value: $\mathbb{E}[e^{lI(X_t^{y,1}=x)} - 1]$. Since $X_t^{y,1}$ can be described by a Bernoulli random variable with probability $p_t(x, y)$, we get with probability $p_t(y, x)$ the value $e^{l\cdot 1} - 1 = e^l - 1$ and likewise with probability $1 - p(x, y)$ the value $e^{l\cdot 0} - 1 = 0$. In conclusion, $\mathbb{E}[e^{l\mathbf{1}(X_t^{y,1}=x)} - 1] = p(x, y)(e^l - 1)$. Using this gives the final result as stated by the following equation.

$$M(l) = \exp\left(\sum_y \rho_y(0)p_t(x, y)(e^l - 1)\right) = \exp\left(\left[\sum_y \rho_y(0)p_t(x, y)\right](e^l - 1)\right) = \exp\left(\rho_x(t)(e^l - 1)\right) \tag{5.73}$$

By uniqueness of the moment generating function, we have shown $\eta_x(t) \sim \text{Poisson}(\rho_x(t))$.

The parameter of the Poisson distribution of $\eta_x(t)$ is given by $\rho_x(t) = \sum_{y\in\mathbb{Z}} p_t(y, x)\rho_x(0)$. From what has been stated earlier regarding semigroups, we can rewrite $\rho_x(t)$ as the time evolution of $\rho_x(0)$ using the semigroup S_t (see Definition 3.7), in other words $\rho_x(t) = S_t\rho_x(0)$. By the first property of semigroups, we have get the following partial differential equation for $\rho_x(t)$:

$$\frac{d}{dt}\rho_x(t) = L\rho_x(t) \overset{\text{RW\_1D}}{=} (\rho_{x-1}(t) - \rho_x(t)) + (\rho_{x+1}(t) - \rho_x(t))$$

Here, RW 1D is given as an example for the 1-dimensional random walk. In a way this can be explained as follows. Here, $\rho_x(t)$ represents the average of particles at position x and at a time t, while $\mu(t, x)$ represents the probability that a particle is at a position x at time t. Similar to quantum mechanics, $\mu(t, x)$ can be seen as the probability amplitude (meaning $|\psi|^2$). Scaling $\mu(t, x)$ by the amount of particles roughly represents the amount of particles at some position and time. However, the scale does not matter for the partial differential equation of μ (equation (5.2)) as both sides cancel out the amount of particles. Thus we expect that $\rho_x(t)$ has the same partial differential equation as $\mu(x, t)$, as they describe the same mass/probability flow due to the system. □

If we expand this to the model with traps, we notice that symmetry of $p_t(x, y)$ in space (in other words $p_t(x, y) = p_t(y, x)$) is not required. From this, we can replace L for another generator which represents the movement in this model. We can conclude that the multiple particle model behaves the same as the single particle model, under the assumption that the particles do not interact with one another.

[12] We have double expectation values since both A_1 and η_y are random variables.

6 Delocalized model

In this chapter, we take into consideration a trapping potential that is delocalized. In other words, trapping may occur when the particle nearby the trap and not necessarily on the trap. This allows us to model more realistic trapping potentials in real crystals. Within the delocalized model, we now consider two types of lattices. First, we have a **mobility lattice** (denoted ML) where the excitons move. Secondly, there exists a **trapping lattice** (denoted TL) which consists of atoms that may or may not trap the excitons. For example, in the material YVO_4:Eu^{3+} the mobility lattice consists of solely vanadium (V) atoms, whereas the trapping lattice consists of yttrium (Y) atoms and europium (Eu) atoms. We leave out the oxygen atoms since the excitons are localized around VO_4-groups. The europium atoms are the dopant and replace some of the yttrium atoms in the trapping lattice. In this model, we assume that only the europium atoms and VO_4-groups can trap excitons, while the yttrium atoms do not influence the excitons. The europium traps are randomly distributed throughout the trapping lattice, having probability p of replacing an yttrium atom. We call p the concentration of europium in the material. In contrast to the localized model, the europium traps do not have to be placed on the mobility lattice. To determine the rate of trapping by europium traps, we use some relation as a function of distance to the europium atom. The vanadate traps are placed on the mobility lattice on every site with a fixed trapping rate.

In comparison with the localized model, this section will be less theoretical and contain more simulations to describe a more complex structure with a spatially-dependent trapping potential. The model describes the movement of excitons in a crystal by a continuous-time Markov chain. The state space will be the mobility lattice as discussed above, and the trapped states. The crystals for which we perform the simulations will be YVO_4:Eu^{3+} and $Rb_3LuV_2O_8$:Eu^{3+}. These materials have similar vanadate emission since they both contain VO_4 groups. Similarly, both materials will exhibit europium emission by choosing europium as dopant.

This chapter consists of the following sections. First, we describe the model and the lattice structure of the crystals $Rb_3LuV_2O_8$:Eu^{3+} and YVO_4:Eu^{3+}. In the following section, named 'Numerical calculation of the model,' we provide the mathematical tools required to perform the simulations. After applying these numerical methods, we discuss what results we acquire from the simulations in the last section.

6.1 Model description

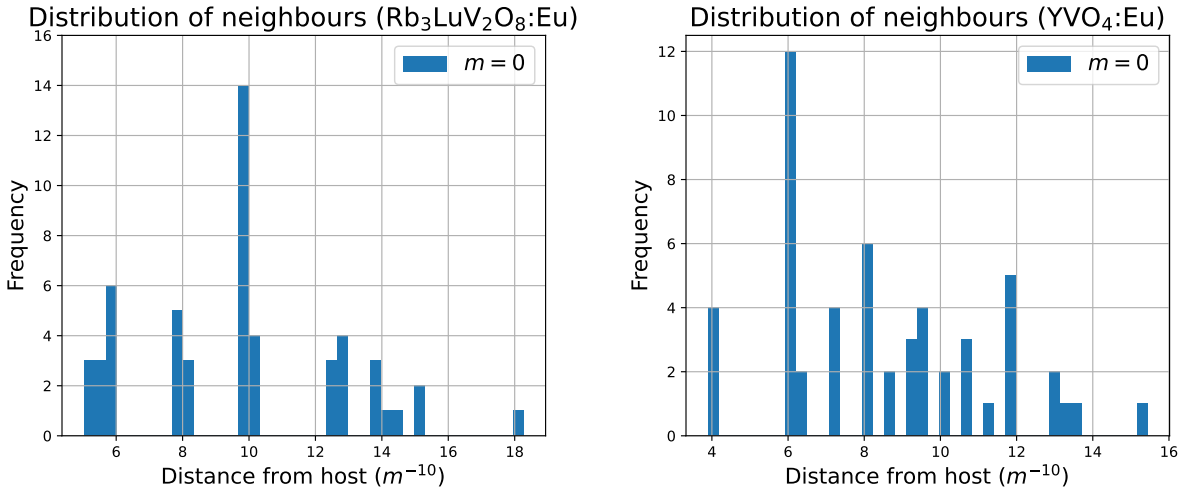
6.1.1 Lattice structure

To find the mobility and trapping lattice, we use the primitive unit cell of a crystal. For the material YVO_4:Eu, we obtain the primitive unit cell by the cell parameters from The Materials Project [Jai+13], using the material `mp-19133`. Using package ASE (Atomic Simulation Environment, [Lar+17]) we find the lattice vectors of the unit cell and fractional coordinates of the atoms. However, since $Rb_3LuV_2O_8$:Eu is not a material on The Materials Project, we use PDF-5+ [KDB24] to acquire a `cif`-file to calculate the lattice vectors and fractional coordinates. In YVO_4:Eu, the primitive unit cell has 2 atoms of V and 2 atoms of Y. For $Rb_3LuV_2O_8$:Eu, the primitive unit cell contains only 1 Lu atom and 2 V atoms. The lattice lengths of $Rb_3LuV_2O_8$:Eu^{3+} are $a = b = 5.944\,\text{Å}$ and $c = 7.866\,\text{Å}$. Those for YVO_4:Eu^{3+} are $a = b = 7.13\,\text{Å}$ and $c = 6.30\,\text{Å}$. We do not account for different lattice lengths for different concentrations of europium or temperature.

We create a finite lattice by extending the lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ by an amount N_1, N_2, N_3 respectively, which we refer to as a $N_1 \times N_2 \times N_3$ lattice. We use (n_1, n_2, n_3, m) to denote a lattice sites of vanadium, lutetium, yttrium or europium. The coordinates n_1, n_2 and n_3 under conditions $0 \leq n_1 \leq N_1, 0 \leq n_2 \leq N_2, 0 \leq n_3 \leq N_3$ refer to the unit cell, while m refers to which fractional coordinate is used.

Example 6.1. *(Lattice structure in $Rb_3LuV_2O_8$:Eu^{3+})* In $Rb_3LuV_2O_8$:Eu^{3+}, we are interested in the location of a vanadium lattice site with coordinates $(n_1, n_2, n_3, m) = (2, 0, 3, 1)$. For $m = 1$ we have the fractional coordinates $(0.3333, 0.6667, 0.2338)$ rounded to 4 digits. The location of this vanadium site will then be

$$2\mathbf{a}_1 + 0\mathbf{a}_2 + 3\mathbf{a}_3 + \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3\end{pmatrix} \begin{pmatrix}0.3333\\0.6667\\0.2338\end{pmatrix}$$



(a) $Rb_3LuV_2O_8$:Eu^{3+} (b) YVO_4:Eu^{3+}

Figure 12: Histogram of the distances of the neighbors for the mobility lattice of a given crystal. Using different m yields the same result since they have same neighborhood of vanadium atoms.

We now know the coordinates of the sites of the mobility and trapping lattice. For the mobility lattice, we assume the exciton can only hop to *nearest neighbor* lattice sites, which we define as being within a certain radius of the host site. How this assumption can be expanded to other lattice sites, we discuss in section 7.2.1. We analyze the amount of neighbors compared to the distance for these materials, as is shown in figure 12.

For the crystal $Rb_3LuV_2O_8$:Eu^{3+}, we find that the nearest neighbors must have a distance smaller than at least $6\,\mathrm{m}^{-10}$, as is visible in figure 12a. In this region, the first bin is distance $5.03\,\mathrm{m}^{-10}$, the second is $5.41\,\mathrm{m}^{-10}$ and the last bin is $5.94\,\mathrm{m}^{-10}$. While the first bin has 3 neighbors, which is enough to travel in all directions x, y and z, choosing only these 3 neighbors will divide the crystal into plates which do not interact with each other. To illustrate this, see figure 11. We expect the exciton to move relatively freely in all directions, so we also consider adding the second bin for the nearest neighbors. Therefore we set the maximum distance to be a nearest neighbor to $5.5\,\mathrm{m}^{-10}$ for $Rb_3LuV_2O_8$:Eu^{3+}.

For YVO_4:Eu^{3+}, the case is simpler compared to $Rb_3LuV_2O_8$:Eu^{3+}. We consider only the first bin of figure 12b to be the nearest neighbors. The next largest bin is at distance $6.0\,\mathrm{m}^{-10}$ which represents 12 neighbors. Since we want to represent roughly equidistant neighbors, we ignore this bin and all distances above it. Thus, we set the maximum distance for nearest neighbors to $5.5\,\mathrm{m}^{-10}$ to be the same as for $Rb_3LuV_2O_8$:Eu^{3+}.

Using the maximum distance $5.5\,\mathrm{m}^{-10}$ for nearest neighbors, we have found the state space and which states to connect.

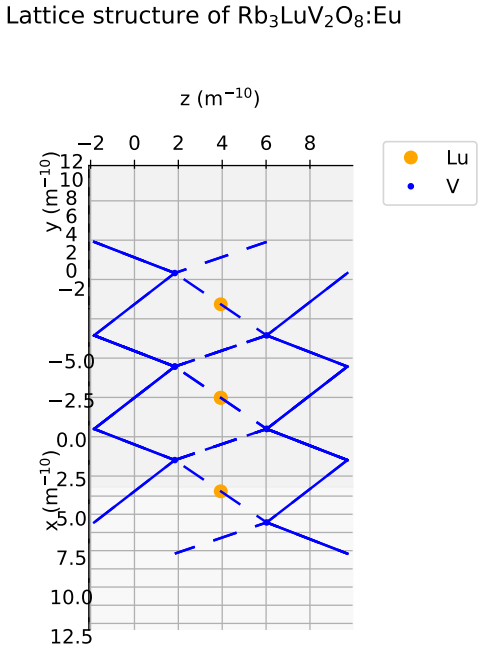


Figure 11: Crystal structure of $Rb_3LuV_2O_8$:Eu^{3+} from given perspective where there 'plate' structure can be seen. Solid lines are connections between vanadium atoms which have a distance smaller than $5.1\,\mathrm{m}^{-10}$ which represent the first bin. The dashed lines indicate the second bin (distances smaller than $5.5\,\mathrm{m}^{-10}$). We remark that without the dashed lines, the crystal does not connect fully.

6.1.2 Transfer rates and boundary conditions

To create the Markov chain of this model, we also need to know the rates between states. Within this lattice, the rate between nearest neighbor lattice sites will be k_m, which we define to be the **migration rate**. Throughout this chapter, we assume the distance to the nearest neighbors to be roughly the same, so we let k_m be a constant independent of the distance. We do consider that migration rate is a function of temperature, as we discuss in the upcoming sections.

Additionally, there is also a rate from every mobility lattice point to the trapped state A due to VO_4-group emission. Let k_{ev} denote the **VO_4 emission rate**, as the rate $c(x, A)$ for all x in the mobility lattice. We assume k_{ev} is independent of temperature.

For trapping by europium, we use the trapped state B. We distinguish two trapped states which encapsulate the B trap.

- **(Excited)** Let B_{exc} denote the excited population of excitons trapped by europium sites. In this state, the exciton can decay due to europium emission represented by the killed state.
- **(Killed)** In the state B_{killed}, the exciton has emitted a photon due to the europium emission. The rate from B_{exc} to B_{killed} is given as k_{ee}, which is the **europium emission rate**.

The rate $c(x, B_{\text{exc}})$ from any x in the mobility lattice depends on the distance to europium sites. Using the Förster energy transfer behavior per equation (2.12), we can define it as follows.

$$c(x, B_{\text{exc}}) = k_t \cdot \sum_{y \in \text{TL}} \frac{1}{d(x, y)^6} \mathbf{1}(y \text{ is Eu}) \tag{6.1}$$

In the above equation, k_t is the **europium energy transfer rate**, $d(x, y)$ is the distance between points states x and y, and TL represents the set of all trapping lattice points. In summary, this equation sums all distances from the host site to all other europium atoms by the power of 6, multiplied by a scalar k_t. We remark that k_t is a model parameter with unit $\text{m}^6\,\text{s}^{-1}$ as this depends on the distance between vanadium and europium sites. For an overview of the model so far, we have illustrated the rates in figure 13.

For YVO_4:Eu^{3+}, the closest distance between a vanadium atom and europium atom is $3.15\,\text{m}^{-10}$. For example, if we have only one trap, then the closest vanadium atom has a rate of approximately $k_t \cdot 1.0235 \cdot 10^{57}\,\text{m}^{-6}$. Due to this extreme order, we can choose the unit angstrom $\text{Å} = \text{m}^{-10}$ instead of converting to m. In this example, we instead have $c(x, B) = k_t(\text{in Å}) \times 1.0236 \cdot 10^{-3}\,\text{Å}^{-6}$. In simulated examples, we often express k_t as a function of another known variable, for example $k_t = K \cdot k_{\text{ev}}$. Then K implicitly obtains the unit Å^6. As an example, let $k_{\text{ev}} = 10^4\,\text{s}^{-1}$. If we state $k_t = 10 \cdot k_{\text{ev}}$ we have $k_t = 10\,\text{Å}^6 \cdot 10^4\,\text{s}^{-1} = 10^5\,\text{Å}^6\,\text{s}^{-1}$ which gives a rate $c(x, B) = k_t \times 1.0236 \times 10^{-3}\,\text{Å}^{-6} = 102.36\,\text{s}^{-1}$.

Our model now depends on the rate parameters k_m, k_t, k_{ev} and k_{ee}. While k_m and k_t are not yet known, we can find k_{ev} and k_{ee} by the lifetime from the photoluminescence (PL) intensity decay. For this approximation,

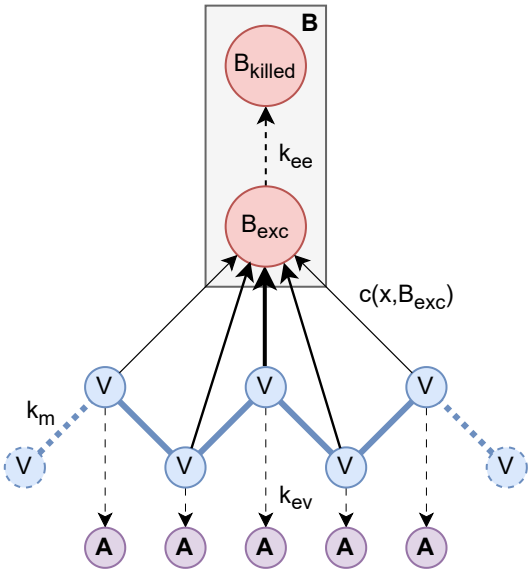


Figure 13: Overview of the rates discussed so far. The V states represent the vanadium atoms, A refers to the vanadate emission trapped state, B similarly is the europium trapped state. The labels k_{ee}, $c(x, B_{\text{exc}}), k_m, k_{\text{ev}}$ are the europium emission rate, europium trapping rate, migration rate and vanadate emission rate respectively. Transition V→A emits vanadate light, while $B_{\text{exc}} \to B_{\text{killed}}$ emits europium light.

we refer to section A.2 in the appendix. In [Dan+20], the lifetime of broad band (420 nm to 700 nm) VO_4-group emission at 500 nm from $Rb3LuV_2O_8$ was measured at $\tau = 23.75\,\mu s$. Then, k_{ev} is calculated by

$$k_{ev} = 1/\tau_{VO_4} = 4.211 \cdot 10^4\,s^{-1} \tag{6.2}$$

Since YVO_4 shares a similar crystal structure and the same vanadate luminescence as $Rb3LuV_2O_8$, we fix $k_{ev} = 4.211 \cdot 10^4\,s^{-1}$ for both crystals. We remark that this is the rate for undoped crystals. It will become clear in the next section that the lifetime of the VO_4-group is a function of the europium concentration, due to the higher trapping potential from europium atoms.

For the europium emission decay, there are multiple sharp peaks in the region of 400nm to 700 nm [Dan+20]. In comparison to VO_4-group decay, within said region there is only a broad single peak. Therefore, when an exciton is trapped by europium, there are multiple ways for the energy to be transferred into light. In the papers [Sev+17] and [CRC09], it becomes clear that there are two different intensity curves originating from europium decay. For decay from level $^5D_1$ we have fast decay where the lifetime lies in the $10\,\mu s$ region, and there is slow decay from the $^5D_0$ level having a lifetime in the region of $1500\,\mu s = 1.5\,ms$. Since transitions from the $^5D_0$ level have a higher intensity, we approximate all transitions from this level as having a lifetime $\tau = 1.5\,ms$ [Sev+17; CRC09] which yields the following europium emission rate

$$k_{ee} = 1/\tau_{Eu} = 6.70 \cdot 10^2\,s^{-1} \tag{6.3}$$

Since $Rb_3LuV_2O_8$:Eu^{3+} and YVO_4:Eu^{3+} both share the same charge of Eu^{3+}, we use the same rate for both models. We remark how there is a considerable order difference between k_{ev} and k_{ee}, however k_{ee} does not influence the transfer between the mobility lattice and the europium sites. Therefore, to expect some europium luminescence, we expect the order of k_t to be nearby the order of k_{ev}. Then k_{ee} only prolongs when a trapped exciton decays into a photon by an europium transition.

To simulate that the crystal is practically 'infinite', we apply periodic boundary conditions to the lattice structures. For example, if an exciton leaves the lattice at the boundary on the right (from xy-perspective), the exciton will reappear at the left side of the lattice. Using these conditions and the above nearest neighbor distances, we have defined the Markov chain to simulate $Rb_3LuV_2O_8$:Eu^{3+} and YVO_4:Eu^{3+}. We give the mobility and trapping lattices in figure 14.

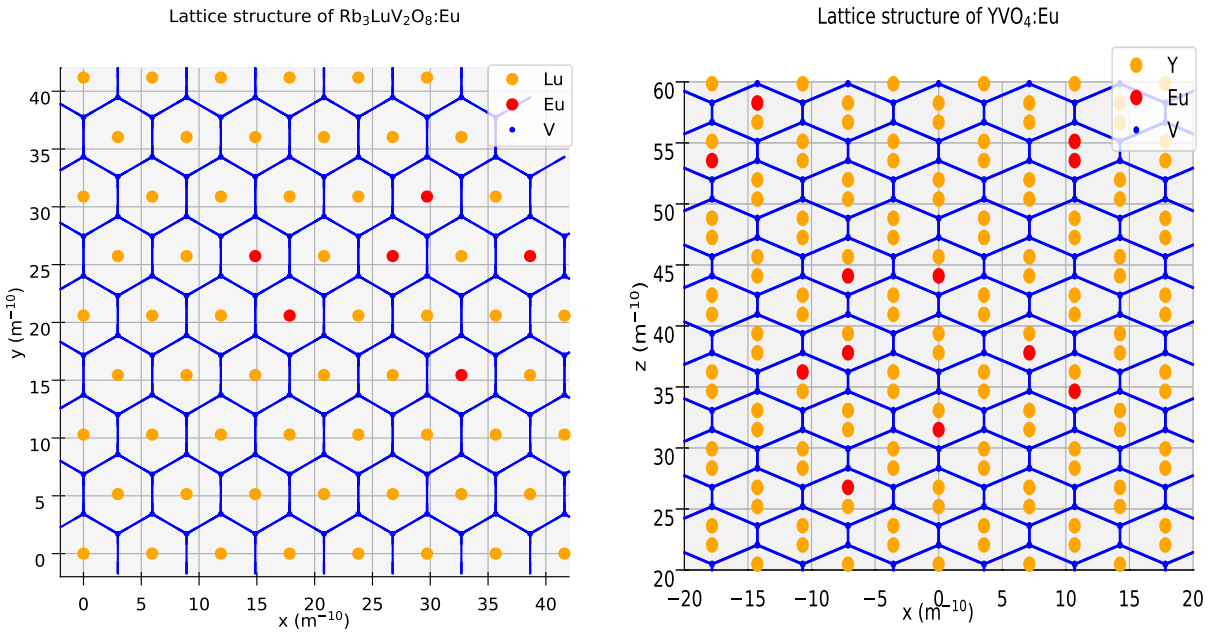
6.2 Numerical calculation of the model

With the lattice structure given, we now continue to discuss how we simulate certain parameters of this model. We divide these two methods into either the 'spatial' or 'temporal' methods. For the spatial method, we discuss the situation when $t \to \infty$. This is useful is we are interested in behavior which is not time-dependent, for example the ratio between the total light intensity of vanadium and europium. We remark that the light intensity is time-dependent, for which we essentially measure the ratio between the integrals of light intensity until $t \to \infty$. When taking into account the temporal aspect, we can find the rise time of the light intensity of vanadium or europium emission.

6.2.1 Spatial methods

To find how much europium emission we have compared to vanadium emission, we first have to define when an exciton is trapped by either trap A or B. This is done by the *hitting time* τ which denotes the first time when the trajectory enters a certain set. If the exciton is trapped by A, we have $\tau_A < \infty$ and $\tau_B = \infty$. For trapping by B we similarly have $\tau_A = \infty$ and $\tau_B < \infty$, but then $\tau_A > \tau_B$. So we can classify how the exciton is trapped based on if $\tau_A < \tau_B$. In conclusion, $\mathbb{P}(\tau_A < \tau_B)$ is the probability of emission by vanadium and $\mathbb{P}(\tau_A > \tau_B) = 1 - \mathbb{P}(\tau_A < \tau_B)$ is [13] the probability of europium emission.
However, we remark that this probability should depend on the position where the exciton starts. For example, if we have a low migration rate, then an exciton which is close to an europium atom should have a lower probability of vanadium emission than an exciton without nearby europium atoms. As such, we refer

[13] It holds $\mathbb{P}(\tau_A = \tau_B) = 0$ since an exciton cannot be trapped by A and B at the same time.



(a) xy-plane of the lattice structure of $Rb_3LuV_2O_8$:Eu^{3+}. (b) xz-plane of lattice structure of YVO_4:Eu^{3+}.

Figure 14: Vanadium atoms (blue dots in figure) are connected to nearest neighbors which have at most $5.5\,\mathrm{m}^{-10}$ from the center atom. In this case, $p = 0.1$ so on average 10% of the Y/Lu atoms (yttrium/lutetium atoms indicated in orange) will be an europium atom (as indicated in red).

to the probability as $\mathbb{P}_x(\tau_A < \tau_B)$ where the exciton starts at $X_0 = x$. In the next theorem, we prove a condition this probability satisfies to solve our problem given a lattice structure.

First, we define the generator for our continuous-time Markov chain. Let L^{RW} denote the generator for the random walk on the mobility lattice, in other words

$$(L^{RW}f)(x) = \sum_{\substack{y\in \mathrm{ML} \\ 0<d(x,y)<5.5\text{Å}}} k_m(f(y) - f(x))$$

Then, combine the states B_{exc} and B_{killed} into the general state B. We can do this since at $t \to \infty$ the buffer state B_{exc} should be zero. As such, the excited population can be ignored in the context of spatial methods.

Using the above definition of L^{RW}, we obtain the following generator for the delocalized model as

$$(Lf)(x) = (L^{RW}f)(x) + c(x, A)(f(A) - f(x)) + c(x, B)(f(B) - f(x)) \tag{6.4}$$

Note that $c(x, A) = k_{\text{ev}}$ which does not depend on x. In contrast, $c(x, B)$ does depend on x.

Given this generator, we find the following condition for the probability $\mathbb{P}_x(\tau_A < \tau_B)$.

Theorem 6.1. (Emission probability condition) Let $(X_t)_{t\geq 0}$ be the continuous-time Markov chain described by the delocalized model. The function $f(x) = \mathbb{P}_x(\tau_A < \tau_B)$ then satisfies the following condition.

$$(Lf)(x) = 0, \qquad x \in \mathrm{ML} \setminus \{A, B\} \tag{6.5}$$
$$f(A) = 1, \tag{6.6}$$
$$f(B) = 0 \tag{6.7}$$

When $k_{\text{ev}} > 0, c(x, B) > 0$ and at least one trap type A or B is present, the converse also holds.

Proof. We first prove the conditions stated by equation (6.5) are true when $f(x) = \mathbb{P}_x(\tau_a < \tau_B)$. If $x = A$ then we know $\mathbb{P}_x(\tau_A < \tau_B) = 1$ since $\tau_A = 0$. If $x = B$, then x is in a trapped state so X_t will never enter the set A. So we know $\tau_B < \tau_A$ always, hence $\mathbb{P}_x(\tau_A < \tau_B) = 0$.
Now, we show the equality $(Lf)(x) = 0$ holds for $x \in \mathrm{ML} \setminus \{A, B\}$. Let us condition on the next step which the process X_t takes when starting at position x. Then we obtain the following.

$$\mathbb{P}_x(\tau_A < \tau_B) = \sum_{y\sim x} \mathbb{P}(x \to y)\mathbb{P}_y(\tau_A < \tau_B) \tag{6.8}$$

$$0 = \sum_{y\sim x} \frac{c(x,y)}{\lambda(x)}\mathbb{P}_y(\tau_A < \tau_B) - \mathbb{P}_x(\tau_A < \tau_B) \tag{6.9}$$

$$0 = \sum_{y\sim x} \frac{c(x,y)}{\lambda(x)}[\mathbb{P}_y(\tau_A < \tau_B) - \mathbb{P}_x(\tau_A < \tau_B)] \tag{6.10}$$

$$0 = \sum_{y\sim x} c(x,y)[f(y) - f(x)] \tag{6.11}$$

$$0 = (Lf)(x) \tag{6.12}$$

Where we used $\lambda(x) = \sum_y c(x,y)$. We conclude that $(Lf)(x) = 0$ for $f(x) = \mathbb{P}_x(\tau_A < \tau_B)$.
We now show the converse holds if $\tilde{L} = L^{RW} - k_{\mathrm{ev}} - c(x,B)$ is invertible, see note 6.1. By conditions $f(A) = 1$ and $f(B) = 0$, the operator $(Lf)(x)$ can be written as

$$(Lf)(x) = \underbrace{\sum_{\substack{y\sim x \\ y\in \mathrm{LT}}} c(x,y)[f(y) - f(x)]}_{L^{RW}f} + c(x,A)[1 - f(x)] + c(x,B)[0 - f(x)]$$

Since $(Lf)(x) = 0$, we can rewrite the above equation as

$$(L^{RW}f)(x) - [c(x,A) + c(x,B)]f(x) = (\tilde{L}f)(x) = -c(x,A)$$

We can substitute $c(x,A) = k_{\mathrm{ev}}$. Given that $\tilde{L}$ is invertible and can find a solution for f. By reversing the steps of equation (6.8), we see that solving this system of equations is the same as of $\mathbb{P}_x(\tau_A < \tau_B)$. Since the model describes a finite state space, if $\tilde{L}$ is invertible then we have a unique solution for f. We conclude that $f(x) = \mathbb{P}_x(\tau_A < \tau_B)$. □

Note 6.1. We remark that $\tilde{L}$ has a zero eigenvalue if and only if $c(x,A) + c(x,B) = 0$ for all $x \in \mathrm{ML}$. By Proposition 4.17 (and extension to 3 dimensions), the eigenvalues of L^{RW} are non-positive. Since $c(x,A) \geq 0$ and $c(x,B) \geq 0$, the operator $\tilde{L} = L^{RW} - [c(x,A) + c(x,B)]$ has strictly negative eigenvalues if $c(x,A) + c(x,B) > 0$. If $c(x,A) + c(x,B) = 0$, the eigenvalues of $\tilde{L}$ are those of L^{RW}. Since the rows of L^{RW} sum to zero by definition, $(1,\ldots,1)^T$ is the eigenvector for eigenvalue zero and thus also for L^{RW}. So, $\tilde{L}$ is not invertible if $c(x,A) + c(x,B) = 0$. This requirement makes sense since if there is no trapping by A or B we have $f(x) = \infty$ for all $x \in \mathrm{LT}$. In summary, $\tilde{L}$ is invertible (and thus we can find $\mathbb{P}_x(\tau_A < \tau_B)$ through equation (6.8)) if we have at least one A or B trap.

We can solve equation (6.5) with the given boundary conditions numerically. To do this, substitute these conditions into equation (6.4). We obtain the following equation.

$$\begin{aligned} 0 = (Lf)(x) &= (L^{RW}f)(x) + k_{\mathrm{ev}}(1 - f(x)) - c(x,B)f(x) \\ -k_{\mathrm{ev}} &= (L^{RW}f)(x) - (k_{\mathrm{ev}} + c(x,B))f(x) \end{aligned}$$

Since ML is finite, L^{RW} is a matrix. Rewriting L and f to some basis $\{x_1,\ldots,x_n\} = \mathrm{ML}$ gives the following vectorized form. Given we assign the probability for an exciton to start at some position $\mathbf{p} = (\mathbb{P}(X_0 = x_1),\ldots,\mathbb{P}(X_0 = x_n))^T$, the probability of vanadium emission is then

$$\mathbb{P}(\text{vanadium emission}) = \mathbf{p} \cdot \mathbf{f}$$

Assuming an uniform distribution for $\mathbf{p}$ (in other words, an exciton is equally likely to appear on any vanadium site), we just average over $\mathbf{f}$. But in the case we create excitons on the unit-cell scale (smaller than the finite lattice), we can also simulate the behavior by changing $\mathbf{p}$. Using this vectorized form gives the equation

$$\mathbf{f} = \left[L^{RW} - (k_{\text{ev}} + \text{diag}(\mathbf{c}))I\right]^{-1} \begin{pmatrix} -k_{\text{ev}} \\ \vdots \\ -k_{\text{ev}} \end{pmatrix} \tag{6.13}$$

Where $\mathbf{c} = (c(x_1, B), \ldots, c(x_n, B))^T$. The vector $\mathbf{f} = (f(x_1), \ldots, f(x_n))^T$ contains the information $(\mathbb{P}_{x_1}(\tau_A < \tau_B), \ldots, \mathbb{P}_{x_n}(\tau_A < \tau_B))^T$. Note that $L^{RW} - (k_{\text{ev}} + \text{diag}(\mathbf{c})) = \tilde{L}$ as defined in Theorem 6.1.

Similar to $\mathbb{P}_x(\tau_A < \tau_B)$, representing which emission the exciton will choose, we also can find $\tau_{A,B}$ which represents the time it takes to be trapped. Intuitively, $\tau_{A,B}$ equals the lifetime of a certain exciton. Of course, this depends on how the exciton travels, as such, we define $\mathbb{E}_x[\tau_{A,B}]$ as the mean lifetime. We can find a similar condition to that of Theorem 6.1 for $\mathbb{E}_x[\tau_{A,B}]$.

Theorem 6.2. (Mean lifetime condition) Let $(X_t)_{t\geq 0}$ be the finite continuous-time Markov chain described by the delocalized model. If $g(x) = \mathbb{E}_x[\tau_{A,B}]$ the following condition holds.

$$(Lg)(x) = -1, \qquad x \in \text{ML} \setminus \{A, B\} \tag{6.14}$$

$$g(x) = 0, \qquad x \in \{A, B\} \tag{6.15}$$

When $k_{\text{ev}} > 0, c(x, B) > 0$ and at least one trap type A or B is present, the converse also holds.

Proof. This proof is similar to that of Theorem 6.1. We first show that $g(x) = \mathbb{E}_x[\tau_{A,B}]$ fulfills the condition stated by equation (6.14). It is clear that $g(x) = 0$ when $x = A$ or $x = B$.
Assume we start at $X_0 = x$. Let $\lambda(x)$ represent the total rate outgoing from x. Then the exciton has to wait $\mathbb{E}[T_X] = 1/\lambda(x)$ time on average to jump to another site. When arrived at the other site, we add $\sum_{y\sim x} \mathbb{E}_y[\tau_{A,B}]\mathbb{P}(x \to y)$ by conditioning on the first jump. This gives the following equation.

$$\mathbb{E}_x[\tau_{A,B}] = \frac{1}{\lambda(x)} + \sum_{y\sim x} \mathbb{P}(x \to y)\mathbb{E}_y[\tau_{A,B}] \tag{6.16}$$

$$\mathbb{E}_x[\tau_{A,B}] = \frac{1}{\lambda(x)} + \sum_{y\sim x} \frac{c(x,y)}{\lambda(x)}\mathbb{E}_x[\tau_{A,B}] \tag{6.17}$$

$$0 = \frac{1}{\lambda(x)} + \sum_{y\sim x} \frac{c(x,y)}{\lambda(x)} \left(\mathbb{E}_y[\tau_{A,B}] - \mathbb{E}_x[\tau_{A,B}]\right) \tag{6.18}$$

$$-1 = \sum_{y\sim x} c(x,y)[g(y) - g(x)] \tag{6.19}$$

$$-1 = (Lg)(x) \tag{6.20}$$

Clearly, $g(x) = 0$ when $x \in \{A, B\}$, which proves the condition if $g(x) = \mathbb{E}_x[\tau_{A,B}]$. To show the converse holds, assume $\tilde{L} = L^{RW} - k_{\text{ev}} - c(x, B)$ is invertible (see note 6.1). By applying the condition $g(A) = g(B) = 0$, we obtain the following equation.

$$-1 = (Lg)(x) = (L^{RW}g)(x) + c(x, A)[0 - g(x)] + c(x, B)[0 - g(x)]$$

Since $c(x, A) = k_{\text{ev}}$, we have $(\tilde{L}g)(x) = (L^{RW}g)(x) - (k_{\text{ev}} + c(x, B))g(x)$. As such, if $\tilde{L}$ is invertible, $g(x) = \mathbb{E}_x[\tau_{A,B}]$ by reversing the steps of equation (6.16). □

Calculating the mean lifetime numerically is done in the same way as with the emission ratio. We can find $\mathbf{g} = (g(x_1), \ldots, g(x_n))^T$ by solving

$$\mathbf{g} = \tilde{L}^{-1} \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} \tag{6.21}$$

We remark that this is the same as $\mathbf{f}$ per equation (6.13) but rescaled by k_{ev}. In other words,

$$\mathbf{g} = \mathbf{f}/k_{\text{ev}} \tag{6.22}$$

This relation makes sense if we consider that $f(x) = \mathbb{P}_x(\tau_A < \tau_B)$. If the exciton at x is near an europium atom (B-type trap), the trapping potential is higher since this potential is the sum of vanadium (A-type) trapping and europium trapping. Thus, the life expectancy of this exciton should be shorter than when there are little europium atoms nearby. Similarly, the probability $\mathbb{P}_x[\tau_A < \tau_B]$ is smaller since we expect more emission by europium. Therefore $f(x)$ decreases and similarly $g(x)$ also decreases.

We can also consider the limit cases for $k_{\text{ev}} \to \infty$ and $k_{\text{ev}} \to 0$. When $k_{\text{ev}} \to \infty$, we essentially place a hard trap everywhere on the lattice, thus the mean lifetime should reduce to zero, as expected by equation (6.22). For the case $k_{\text{ev}} \to 0$, we ignore trapping by vanadium atoms. We can still solve for $\mathbf{g}$ by equation (6.21), but the relation $\mathbf{g} = \mathbf{f}/k_{\text{ev}}$ suggests $\mathbf{g}$ approaches infinity. However, we remark that when $k_{\text{ev}} \to 0$, the probability $f(x) = \mathbb{P}_x(\tau_A < \tau_B)$ becomes zero since there are no A-type traps, as is clear by substituting $k_{\text{ev}} = 0$ into equation (6.13). In conclusion, the relation breaks down due to approaching the limit 0/0 but $\mathbf{g}$ can still be calculated by equation (6.21).
Finally, we note that the units of $\mathbf{g} = \mathbf{f}/k_{\text{ev}}$ are correct, since $\mathbf{g}$ represents a time, $\mathbf{f}$ is unitless (represents a probability) and k_{ev} has units s^{-1}.

6.2.2 Temporal methods

An important aspect of choosing a continuous-time Markov chain, compared to the discrete-time version, is to be able to simulate excitons within a continuous time frame. While we model the entire state space with a continuous-time Markov chain, it is difficult to directly measure the movement in excitons within materials. Another method to compare our model to experimental data related to the light intensity emitted by trapping. Therefore our goal will be to analyze exciton trapping in the finite time setting.

In section 5.5 we discussed how we may relate the behavior of one particle to that of multiple particles. In summary, the average amount of excitons on a lattice site, denoted $\rho_x(t)$, follows the same law as $\mathbb{P}(X_t = x)$, which represents the probability of an exciton being at a position x at a time t. To recall, from Theorem 5.9 we have found

$$\rho_x(t) = S_t \rho_x(0) \tag{6.23}$$

We note that $\rho_x(0)$ simply represents the initial position of the excitons. Since we do not know how many excitons are present at time zero, we cannot exactly determine $\rho_x(t)$ and more importantly $\frac{\partial}{\partial t}\rho_x(t)$ by some scale [14]. Therefore, we will use one exciton, which is equivalent to representing the average amount of excitons by the probability of the exciton being at that location.

To numerically calculate this probability, we need to find e^{tL} and thus first have to define L. As we work with a finite state space and continuous-time Markov chain, the generator L will be a matrix defined using equation (3.18). The matrix will be square with a length of $\#\text{ML} + 3$, where the first elements of the matrix are $x_1, \ldots, x_n$ which are the elements of the mobility lattice. Then we also have the elements A, B_{exc} and B_{killed}.

[14] By Kolmogorov's backward equation (theorem 3.9), we have the relation

$$\frac{\partial}{\partial t}\rho(t,x) = LS_t\rho(0,x) = L\rho(t,x)$$

Regarding the structure of L, we already found the nearest neighbors so $L_{x_i,x_j} = k_m$ if x_i and x_j are nearest neighbors. For all x_i, there exists VO_4 emission with rate k_{ev}, so $L_{x_i,A} = k_{\text{ev}}$. Then for europium trapping, we use equation (6.1). Finally, the transfer from B_{exc} to B_{killed} has a rate $c(B_{\text{exc}}, B_{\text{killed}}) = k_{\text{ee}}$. Afterwards, the diagonal of the matrix is calculated as

$$L_{xx} = \sum_{y \neq x} -L_{xy}$$

Such that the sum of each row is zero.

With L given, we can find the transition matrix for any time $t \geq 0$ with the formula $P = e^{tL}$. This can be computed numerically by for example the SciPy function `scipy.linalg.expm(t*L)`, which uses an advanced algorithm. A simple algorithm would be by diagonalizing L as PDP^{-1} with $D = \text{diag}(D_1, \ldots, D_m)$, then we find

$$P = e^{tL} = P \cdot \text{diag}(e^{tD_1}, \ldots, e^{tD_m}) \cdot P^{-1}$$

When calculating with linear timesteps $0, \Delta t, 2\Delta t, \ldots$, we can simplify calculation by using matrix powers of $e^{L\Delta t}$. For example, $P(n\Delta t) = P(\Delta t)^n = \left(e^{L\Delta t}\right)^n$.

By calculating e^{tL} where L is given by and indexing this matrix, we obtain the probability of an exciton getting trapped in some state. To find the probability of being trapped by trap A leading to vanadate emission, we assume the probability of an exciton spawning on the mobility lattice is equal everywhere. The probability of being in trapped state A conditioned on the starting position will be

$$\mathbb{P}(X_t = A) = \frac{1}{\#\text{ML}} \sum_{x \in \text{ML}} (e^{tL})_{xA} \tag{6.24}$$

We expect an exponential decay curve for light intensity, which is explained in section A.2 in the appendix. This simple model discussed in the appendix can be expanded to the complete lattice instead of two states. Since we distinguish the europium trapped state into the excited state B_{exc} and killed state B_{killed}, we obtain a rise time in the light intensity. This rise time comes from the fact that the excitons get trapped can get trapped sooner if they are closer to an europium trap. This property comes from the Förster $1/r^6$-relation we use for europium trapping.

6.3 Simulation behavior of the model

6.3.1 Spatial distribution of vanadium and europium emission

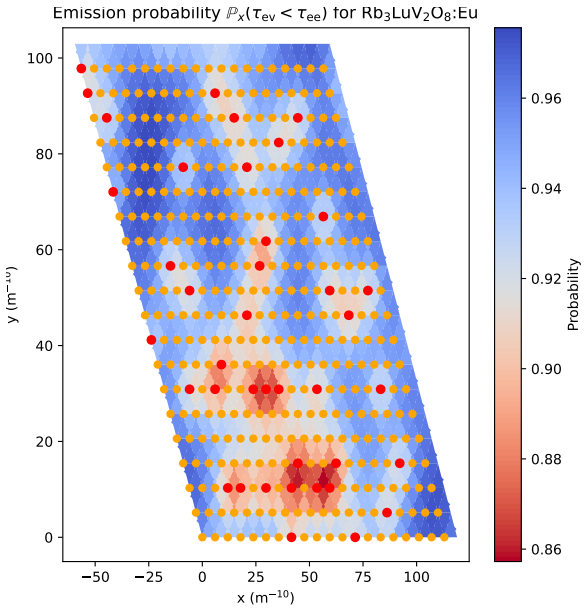
To begin analyzing how the model behaves, we first look to numerically calculating $\mathbb{P}_x(\tau_A < \tau_B)$ and $\mathbb{E}_x[\tau_{A,B}]$ as given section 6.2.1. We will analyze and visualize these values for $Rb_3LuV_2O_8$:Eu^{3+}. We then average $\mathbb{P}_x(\tau_A < \tau_B)$ over all points and try to describe how this average behaves for different parameters.

In figure 15, we can examine per lattice site what the probability will be to decay via the VO_4-group or via europium, and how the mean lifetime differs spatially. We first remark that both plots are essentially the same but rescaled, as found by equation (6.22). In the case of $k_m = 10 \cdot k_{\text{ev}}$, $k_t = 1 \cdot k_{\text{ev}}$ and $p = 0.1$, the probability is approximately 0.9 so the mean lifetime can be estimated as follows.

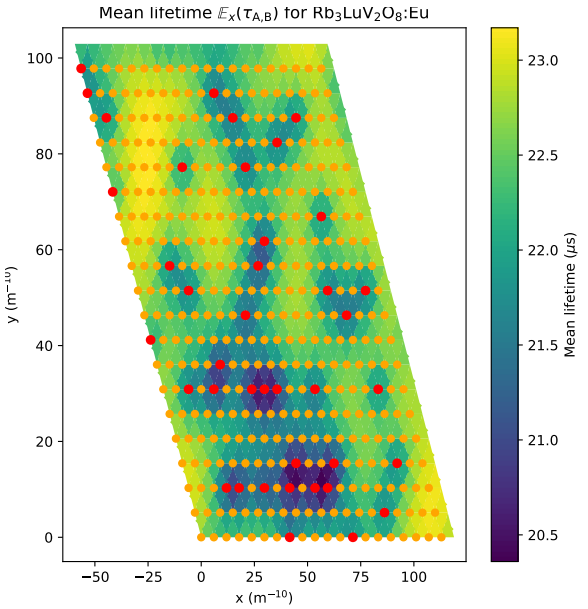
$$\frac{0.9}{k_{\text{ev}}} \approx 0.9 \cdot 0.25 \cdot 10^{-4} \approx 2.2 \cdot 10^{-5}\,\text{s} = 22\,\mu\text{s}$$

Averaging over all points, we obtain a mean lifetime of $\langle \mathbb{E}_x[\tau_{A,B}] \rangle = 22.25\,\mu$s. As is visible in the colorbar in figure 15b, the mean lifetime ranges from approximately $19.5\,\mu$s to $23.0\,\mu$s. For figure 15a, we obtain an average probability $\langle \mathbb{P}_x(\tau_A < \tau_B) \rangle$ of 0.937 if we weigh every lattice point equally. Intuitively, we can understand this probability more as the ratio between vanadate and europium emission. This ratio relates the amount of total photons emitted by the VO_4-group and europium sites.

For the next results, we average over all mobility sites evenly. Since the mean lifetime is a rescaled version of the emission probability, we consider now only the probability $\mathbb{P}_x(\tau_A < \tau_B)$. At this moment, we can configure the following parameters:



(a) Probability of vanadium emission as opposed to europium emission per vanadium lattice site in $Rb_3LuV_2O_8$:Eu^{3+}, illustrated on the xy-plane.



(b) Mean lifetime of excitons given per vanadium lattice site in $Rb_3LuV_2O_8$:Eu^{3+} for the xy-plane.

Figure 15: The vanadium decay probability and mean lifetime as calculated per equation (6.13) and (6.21) respectively for $Rb_3LuV_2O_8$:Eu^{3+}, using a $20\times20\times1$ lattice with parameters $p = 0.1, k_m = 10\cdot k_{ev}, k_t = 1\cdot k_{ev}$. Orange dots represent lutetium atoms, red dots represent europium atoms, and the colored triangles represent the nearby regions of the vanadium atoms.

1. **(Concentration p)** If we increase the concentration p, we expect more europium emission and thus a lower probability. At a certain point, the crystal is saturated with europium and expect the probability to increase slower for higher concentrations, for example $p = 0.2$.

2. **(Migration rate k_m)** For a low migration rate, excitons have difficulty moving, which results in the probability of VO_4 or europium emission being determined by where the exciton is placed. Since the trapping region of the europium trap is small compared to a large lattice, it will benefit europium emission if the migration rate is higher. Excitons outside this region can also come into contact with europium trapping sites, which increases the average exposure with all europium traps.

3. **(Transfer rate k_t)** Increasing k_t will increase the amount of europium emission and thus lower the probability. Depending on how many traps are placed, we expect the probability to be around 0.5 when k_{ev} is in the same order of $c(x, B_{exc})$ for high concentrations of p.

We will come back to the migration rate k_m, for now we set $k_m = 10^5 \cdot k_{ev}$. This is motivated by the assumption that excitons can move relatively freely throughout the lattice. If we then vary the concentration p and transfer rate k_t, we obtain figure 16 for the material $Rb_3LuV_2O_8$:Eu^{3+}. For YVO_4:Eu^{3+}, see figure 29 in the appendix. In this figure, we observe some key details about the behavior of trapping of excitons. Firstly, a higher concentration p decreases the probability and since logically more europium increases the amount of trapping by europium. We note that p is chosen logarithmically, therefore increasing p to obtain a higher emission rate for europium becomes increasingly harder.

While k_m is fixed, this does not have to be the case. Figure 28 in the appendix examines how the graph changes for different values of k_m, this can be summarized as shifting the k_t/k_m axis. We then discuss how k_t influences the probability, we note that the x-axis is a ratio of k_t to k_m displayed logarithmically. It is clear from the figure that k_t changes the probability with asymptotic behavior. When $k_t \to 0$ we neglect europium trapping and, as such, the probability approaches 1. For $k_t \to \infty$ the trapping potential from

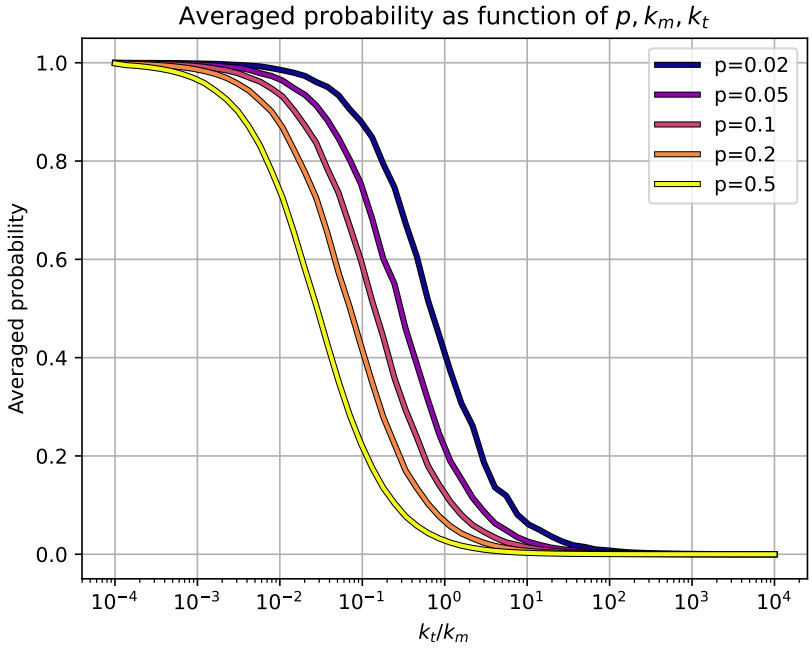


Figure 16: Averaged probability $\langle \mathbb{P}_x(\tau_A < \tau_B) \rangle$ for parameters k_t/k_m and for different concentrations p (e.g. $p = 0.1$ equals 10% europium concentration) for a $12 \times 12 \times 12$ crystal of $Rb_3LuV_2O_8$:Eu^{3+}. Migration rate is $k_m = 10^5 \cdot k_{\text{ev}} \approx 4 \cdot 10^9\,\text{s}^{-1}$. For low p, there are low amounts of europium traps creating jitter.

europium becomes strong enough to win over VO_4-group emission, therefore we only find europium emission. For values $p = 0.02$ to $p = 0.05$ we find that the transition between primarily VO_4-emission and europium emission occurs at $k_t/k_m = 10^{-3}$ and ends at $k_t/k_m = 10^1$. In the middle, when $k_t/k_m = 10^{-1}$, we find that $k_t = 10^{-1} \cdot k_m = 10^4 \cdot k_{\text{ev}}$ therefore $c(x, B) \approx 10^5\,\text{s}^{-1}$, which is just higher than k_{ev} itself. Remarkable, at $k_t/k_m = 10^{-1}$ we find extremely different values for the probability for different concentrations p. For example, $p = 0.5$ gives probability 0.2 (80% is europium trapping), meanwhile $p = 0.02$ gives probability 0.9 (10% europium trapping).

6.3.2 Averaged europium emission as function of temperature

Previously, we considered k_m fixed and k_t, p as variables. Now, we let $p = 0.1$ and let k_t, k_m be variables and for now consider how much europium emission the crystal emits. So, by 'probability' we mean the probability of europium emission, which is $1 - \langle \mathbb{P}_x(\tau_A < \tau_B) \rangle$. Using the same method, we can find the probability of europium emission will be given by the parameters k_m and k_t. In figure 17, we compute for every lattice point this probability of europium emission in YVO_4:Eu^{3+}. Similarly to before, we show how the average probability changes as a function of k_m and k_t. We also show the standard deviation of the probability for all lattice points. This then shows how uniform the probability is spread in the lattice.

As stated above, we calculate the average and standard deviation of the probability of europium emission in YVO_4:Eu^{3+} for concentration $p = 0.1$. In the top figure, we observe that the emission of europium increases as k_m increases for all k_t. This can be explained since for small k_m/k_{ev}, we have a high standard deviation. Thus, the probability of europium emission strongly depends on which lattice site the exciton starts. If the excitons have greater mobility, they can move into regions with a higher region of influence by europium trapping, which benefits europium emission.

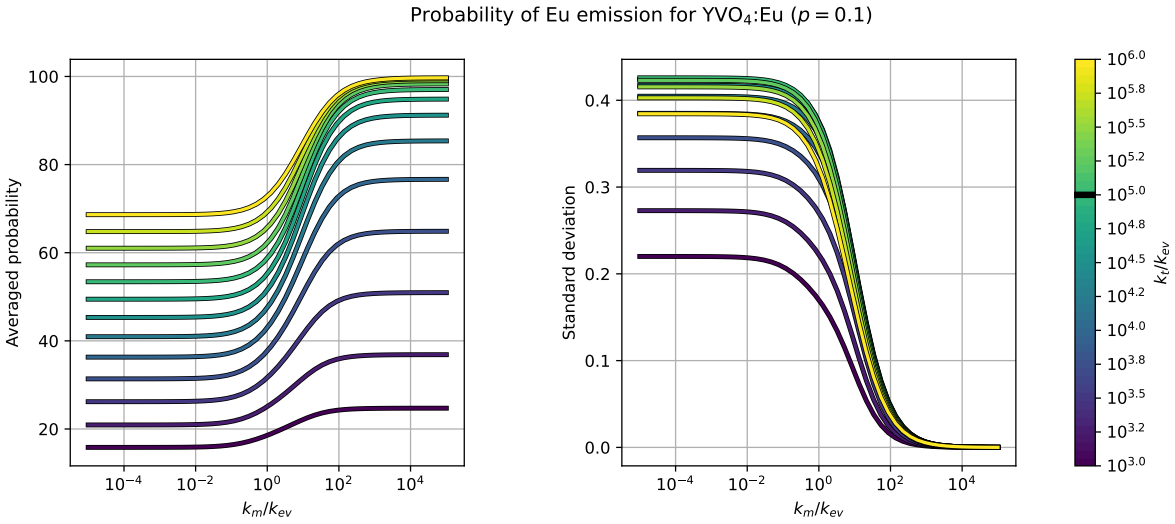


Figure 17: Average and standard deviation of the probability of europium emission for all lattice sites in YVO_4:Eu^{3+} ($10 \times 10 \times 10$ lattice) with europium concentration $p = 0.1$. The top and bottom figure show the average probability of europium emission and the standard deviation respectively. In both the top and bottom figure, the x-axis represents the migration rate by k_m/k_{ev}. The level curves represent the parameter k_t as shown in the color bar to the right, represented by k_t/k_{ev}. In the color bar, the black line indicates $k_t/k_{\text{ev}} = 10^5$ which is equivalent to $k_t/k_m = 1$ where $k_m = 10^5 \cdot k_{\text{ev}}$ as in figure 16.

Also, we note that there exist two asymptotes with respect to k_m with a turning point around $k_m \approx 10 \cdot k_{\text{ev}}$ [15]. The probability of europium emission in these asymptotes depends on the value of k_t for fixed p. For higher k_t, it is clear that there will be more europium emission for any k_m, as is visible in the top figure of 17. This is as expected, since k_t is proportional to the trapping potential/rate of europium. In contrast with the average, in the bottom figure we see that the standard deviation will in most cases be higher with an increasing k_t, but not consistently. In general, for low k_m, modifying k_t will change how strong the europium emission is compared to VO_4-emission in the localized regions around europium atoms. Therefore, higher values of k_t increase the difference of europium emission between the outside and inside of localized sites around europium atoms. For higher k_m, the excitons are able to move relatively freely and thus whether an exciton decays by europium emission or vanadate emission will not depend on the starting position. In conclusion, the standard deviation converges to zero for higher k_m regardless of k_t.

Our next plan is to change k_m for a temperature relation. As derived in the theory about exciton hopping, the migration rate will follow the law per equation (2.11), which states that the migration rate $k_m(T_2)$ at temperature T_2 is given by the migration rate $k_m(T_1)$ at temperature T_1 with activation energy E_a as follows.

$$k_m(T_2) = k_m(T_1) \cdot \sqrt{\frac{T_1}{T_2}} \cdot \exp\left(-\frac{E_a}{k_B}\left[\frac{1}{T_2} - \frac{1}{T_1}\right]\right) \tag{6.25}$$

Let us define $k_{m,0}$ as the reference migration rate at temperature $T_0 = 300\,\text{K}$ so $k_{m,0} = k_m(300\,\text{K})$. Plotting the above equation for a given activation energy $E_a = 0.50\,\text{eV}$ can be seen in figure 18.
We expect the activation energy to be in the $0.1\,\text{eV}$ to $1.0\,\text{eV}$ range due to the strong effect from the exponential function, see figure 30 in the appendix. For our simulations, we set $E_a = 0.50\,\text{eV}$ as a constant, chosen such that this lies within the middle of the region we expect the activation energy to be.

We can now replace the parameter k_m with the reference migration rate $k_{m,0}$ and the activation energy E_a. Using equation (6.25), we obtain the results which can be seen in figure 19.
From the figure, we see that the emission of europium increases as function of temperature T. As the migration rate increases with higher temperatures, this characteristic is to be expected. Also, the terms

[15] In figure 28, the transition region where the level curves broaden is also when $k_m = 10 \cdot k_{\text{ev}}$. While this is for different crystals, parameter k_m is similar in both materials.

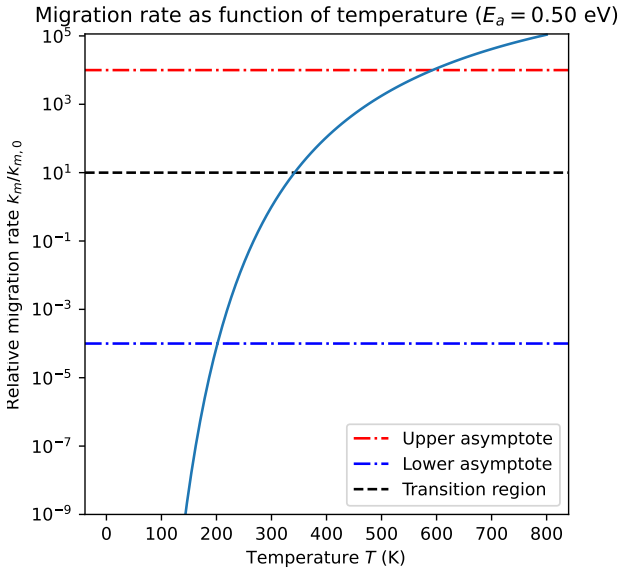


Figure 18: The migration rate as calculated using equation (6.25) divided by $k_{m,0}$ with activation energy $E_a = 0.50\,\text{eV}$. We denote the asymptote regions $k_m/k_{\text{ev}} < 10^{-2}$ and $k_m/k_{\text{ev}} > 10^2$ by the lower and upper asymptote respectively if $k_{m,0} = k_{\text{ev}}$. If the migration rate is outside of these bounds, we europium emission probability can be approximated to be constant, in other words, there will be no transition behavior.

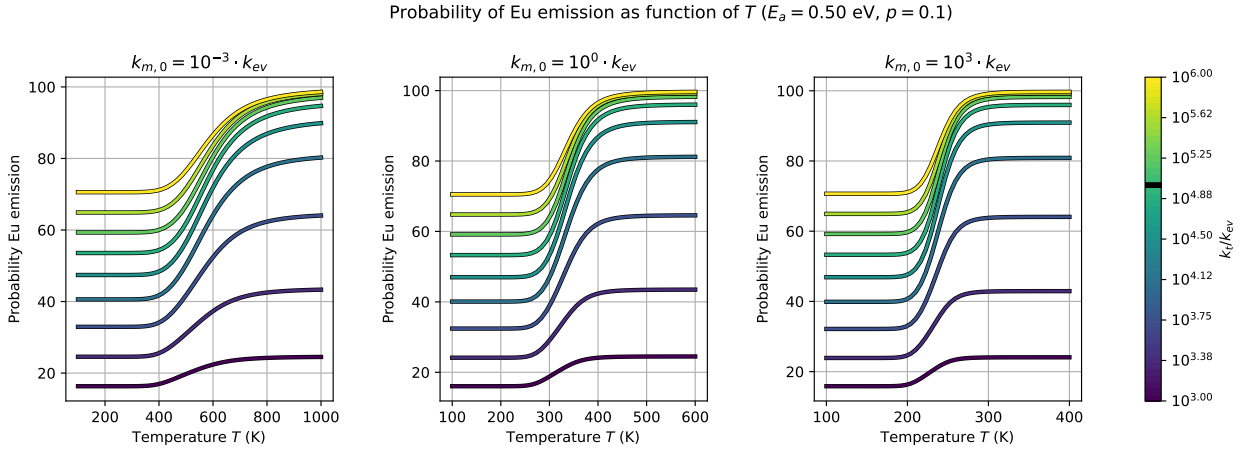


Figure 19: Probability of europium emission as function of temperature T with activation energy $E_a = 0.50\,\text{eV}$ for different $k_{m,0}$ and k_t (color bar). This is simulated for a $10 \times 10 \times 10$ lattice of YVO_4:Eu^{3+}. The temperature axis is shortened for higher $k_{m,0}$ since higher values shift the transition region from $T \approx 500\,\text{K}$ for $k_{m,0} = 10^{-3} \cdot k_{\text{ev}}$ (left plot) to $T \approx 250\,\text{K}$ for $k_{m,0} = 10^{3} \cdot k_{\text{ev}}$ (right plot). We note that since the migration rate is an increasing function for this activation energy and temperature range, the probability of europium emission increases with higher temperatures with asymptotic limits.

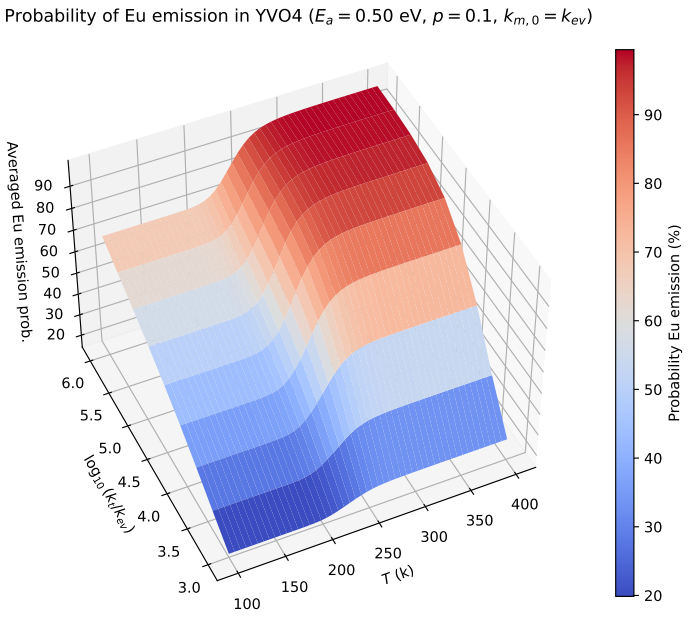


Figure 20: Surface plot of the probability of europium emission in YVO_4:Eu^{3+} as function of temperature T and europium transfer rate k_t. Simulation uses $10 \times 10 \times 10$ lattice with parameters $E_a = 0.50\,\text{eV}$, $p = 0.1$ and $k_{m,0} = k_{\text{ev}}$. We remark the transition region around $T \approx 250\,\text{K}$ which divides the two asymptotes. For highter temperatures, the difference in asymptotes is increases compared to lower temperatures.

'upper' and 'lower asymptote' from when we defined $k_{m,0}$ are now visible in the middle plot of the said figure. For $k_m = k_{\text{ev}}$, we see that the transition region lies between 200 K and 600 K, which are the asymptote regions we mentioned for $E_a = 0.50\,\text{eV}$.

Another way to view the transition region is as a surface as function of temperature T and k_t, see figure 20. In this figure, the transition region is clearly visible which is in between the two asymptotes with respect to temperature. We remark how the difference between asymptotes for different k_t is higher than for lower temperatures. In other words, at higher temperatures the difference in europium emission probability is greater for different k_t than for lower temperatures.

6.3.3 Time-dependent intensity curve

To conclude this chapter, we analyze how trapping occurs for different points of time. Within this section, we examine specifically how excitons are trapped by vanadate and europium trapping. From section 6.2.2, we can compute the intensity of vanadate and europium emission for certain time steps.

First, we begin with vanadate emission due to A trapping. We come back to the comment where we define the vanadate emission rate k_{ev}. This value is determined experimentally using the intensity curve at a wavelength of vanadate emission, however this was measured without europium present. From the delocalized model, we find that the lifetime of VO_4-group emission is a function of the concentration p, the migration rate k_m and the europium transfer rate k_t. We show this in figure 21 for 3 different pairs of parameters k_m and k_t and a range of concentrations of europium. In this figure, we first remark that for 0% europium, we obtain the lifetime $23.75\,\mu\text{s}$ as was experimentally defined. Then, when increasing the europium concentration, the lifetime decreases. Introducing europium trapping reduces vanadate emission [16], but this occurs at a different rate than by A trapping. Therefore, the decay curve of vanadate emission stops 'earlier' which decreases the lifetime.

Secondly, we consider how europium trapping is modelled for different time steps. In comparison to vanadate emission, the europium intensity curve will also have a rise time. This occurs since we distinguish

[16] Metaphorically speaking, europium traps 'steal' excitons from VO_4-groups.

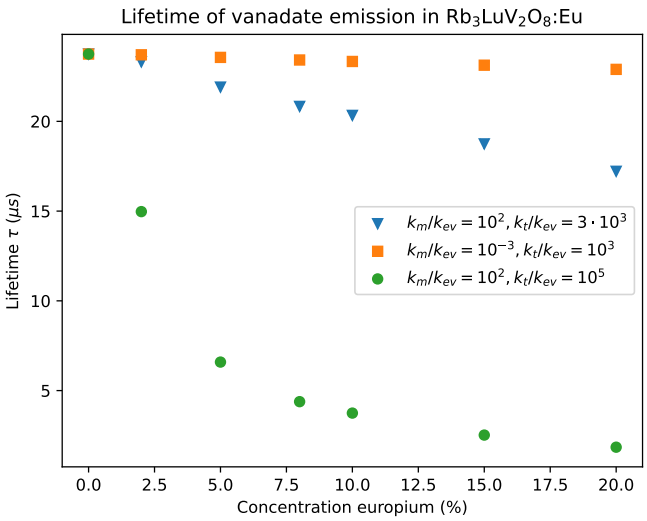


Figure 21: Lifetime of vanadate emission in $Rb_3LuV_2O_8$:Eu^{3+} as a function of europium concentration p. Increasing the strength of the europium traps by parameters k_m and k_t decreases this lifetime. The parameters $k_m = 10^2 \cdot k_{\text{ev}}, k_t = 3 \cdot 10^3 \cdot k_{\text{ev}}$ are specifically chosen to show linear behavior, while $k_m = 10^2 \cdot k_{\text{ev}}, k_t = 10^5 \cdot k_{\text{ev}}$ shows this is not always linear.

both the excited state B_{exc} and the killed state B_{killed}. This is the case when $c(x, B_{\text{exc}})$ is larger than k_{ee}, then the influx for state B_{exc} will be high for short times (on order of $1/c(x, B_{\text{exc}})$). Since we have a finite amount of excitons, the rise of influx will eventually stop. At this point, the excited excitons within the state B_{exc} will slowly decay into photons presented as an exponential decay curve in the intensity [17].

For our purposes, we define the rise time as the time required for the intensity to increase from 10% to 90% of its maximum. Using the same approach as before, we calculate can calculate the rise time from the intensity curve of europium emission. See figure 22 for the rise times in YVO_4:Eu^{3+} for different concentrations. We remark that in this figure, k_t was chosen such that the rise time would be concentration-dependent. For values of k_t lower than approximately $10^3 \cdot k_{\text{ev}}$ (in context of $k_m = 10^5 \cdot k_{\text{ev}}$), the rise time will be independent of the concentration p. We refer to figure 31 in the appendix for an example.

[17] An analogue of this phenomenon is radioactive decay in certain isotopes. For example, uranium-235 will decay into thorium-231 with a lifetime of 25.5 hours. Then, thorium-231 decays into protactinium-231 with a lifetime of approximately 32760 years. If we then plot the amount of thorium-231 in the material, we obtain a rise and decay curve similar to the europium intensity curve.

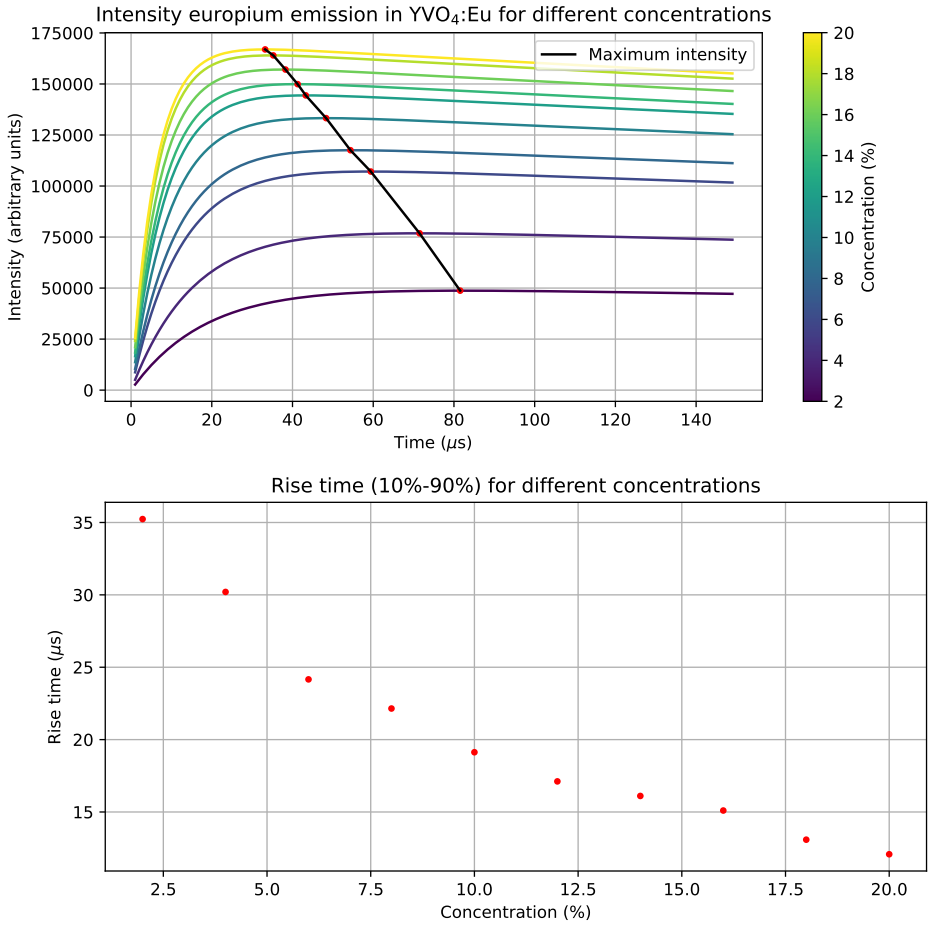


Figure 22: Analysis of the rise time of europium emission in YVO_4:Eu^{3+}. The top figure illustrates the intensity curve for different concentrations, where the 'maximum' line shows how the rise time shrinks as the concentration is increased. The bottom figure expresses the rise time (time it takes for 10% to 90% of maximum intensity) as a function of the concentration. Simulation uses $7 \times 7 \times 7$ lattice with parameters $k_m = 10^5 \cdot k_{\text{ev}}$ and $k_t = 5 \cdot 10^3 \cdot k_{\text{ev}}$

7 Results and Discussion

7.1 Rate level curve for vanadium lifetime in $Rb_3LuV_2O_8$:Eu^{3+}

In the article [Dan+20], the lifetime of VO_4-group in $Rb_3LuV_2O_8$:Eu^{3+} was measured using the luminescence intensity curve for europium concentrations $p = 0\%, 2\%, 5\%, 8\%, 10\%, 15\%, 20\%$. We have simulated the lifetime of VO_4-decay for a $Rb_3LuV_2O_8$:Eu^{3+} lattice of size $7 \times 7 \times 5$ for different model parameters of k_m and k_t. Since these parameters are in a way interchangeable, we cannot find an exact value of both k_m and k_t. We have found an optimal level curve for these parameters such that error between the model and the experimental data is minimal. For a set of experimental lifetimes $\{\tau_{\text{exp,p}} \,|\, p = 0\%, \ldots, 20\%\}$ and simulated model lifetimes $\{\tau_{\text{model,p}} \,|\, p = 0\%, \ldots, 20\%\}$, we define the root square error (RSE) as

$$\text{RSE} = \sqrt{\sum_p (\tau_{\text{model,p}} - \tau_{\text{exp,p}})^2} \tag{7.1}$$

Using the above definition, the simulated root square error for parameters $k_m = 10^{-2} \cdot k_{\text{ev}}$ to $k_m = 10^2 \cdot k_{\text{ev}}$ and $k_t = 10^3 \cdot k_{\text{ev}}$ to $k_t = 10^5 \cdot k_{\text{ev}}$ are given as a contour plot in figure 23.

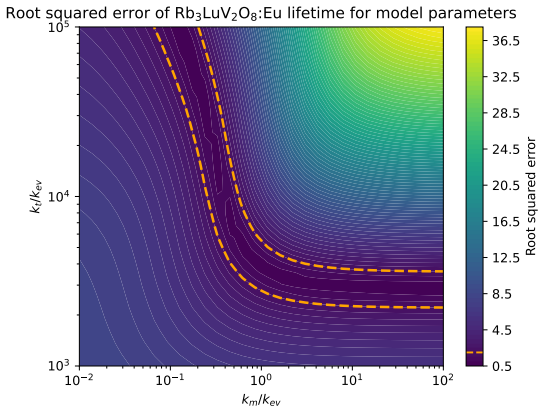


Figure 23: Root squared error between VO_4-group lifetimes in $Rb_3LuV_2O_8$:Eu^{3+} measured experimentally and acquired from simulations. The x-axis represents the migration rate k_m expressed in units of the VO_4emission rate k_{ev} at zero europium concentration. Likewise, the y-axis is the europium energy transfer rate expressed in units of k_{ev}. The orange dotted line encloses the surface for which RSE ≤ 2.

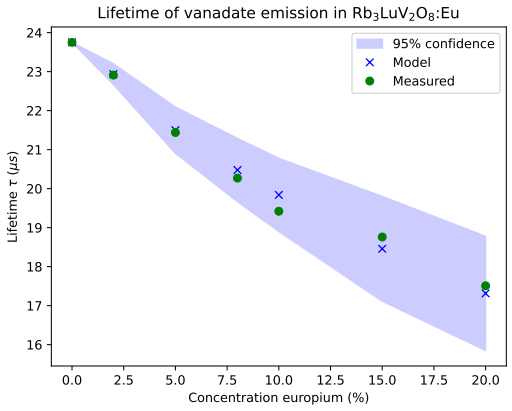


Figure 24: Lifetime of VO_4-group emission for different europium concentrations, comparing experimental data (Measured) and simulated values (Model) for smallest root square error ($k_m = 9.237 \cdot 10^{-2} \cdot k_{\text{ev}}, k_t = 10^5 \cdot k_{\text{ev}}$) for a $7 \times 7 \times 5$ $Rb_3LuV_2O_8$:Eu^{3+} lattice. Confidence interval represents 2 standard deviations of the simulated values, calculated for 200 simulations.

7.1.1 Discussion

The simulation uses a fixed random seed to generate consistent results for different parameters. While it can be argued that having a consistent trap layout does not represent fully the randomness of the crystal, the size is large enough that simulation results lie close enough to averaged values. In figure 24, the confidence interval is given for the model acquired from 200 simulations which have different trap layouts. The parameters chosen in said figure minimize the root square error in figure 23 to the left. The model gives a relatively large confidence interval for the $7 \times 7 \times 5$ lattice, therefore the model has a low precision. Remarkably, the model does have high accuracy, as the averaged simulation values lie close to the measured values. We note that these experimentally obtained values also have some error. For example, this can be due to errors in the europium concentration or in how the material is synthesized, but also from uncertainty of the equipment used. While we do consider this error, we assume that on this scale we state the model is accurate enough.

Our choice for a $7\times7\times5$ lattice with 200 simulations is as follows. We observed that increasing the lattice compared to $7\times7\times5$, for example $10\times10\times8$ did not improve the accuracy or precision (confidence) of the model much. Since the time required for such a N lattice scales non-linearly due to matrix computations, we have chosen for a $7\times7\times5$ lattice. For the amount of simulations, in our case we found that in the range of 100 to 200 simulations would suffice for a consistent result. Within the time frame of minutes to hours, we found that the confidence region does not improve much due to a bigger lattice or from more simulations. For future research, it might be interesting to run these simulations on large lattices (for example $100\times100\times100$, which was limited due to memory and time issues) to find how the confidence region converges (or does not).

Aside from discussion on model assumptions, due to the randomness involved we do not expect a root square error close to zero (for example ≤ 0.1). Therefore it has been chosen that the most optimal parameters which come close to simulating is a level curve, or rather a level surface, to represent which pairwise model parameters (k_m, k_t) are most suitable.

While this is a case of overfitting, another dataset of luminescence measurements on $Rb_3LuV_2O_8$:Eu^{3+} is necessary to find a region of parameters which intersects all regions of suitable parameters of different measurement types. In other words, comparing other properties of the model and/or experimental data obtained from $Rb_3LuV_2O_8$:Eu^{3+}, can narrow the most suitable parameters (k_m, k_t) down. For example, measuring rise times of Eu^{3+} emission and ratio between total light emitted by Eu^{3+} and VO_4^{3-} can bridge the problem of overfitting. A counterargument to the overfitting case is the decreasing linear trend found in the experimental data, which is also found in the model as discussed in section 6.3.3.

7.2 $VO_4{}^{3-}$/Eu^{3+}-intensity ratio as function of temperature in YVO_4:Eu^{3+}

The material YVO_4:Eu^{3+} with 1% europium concentration is synthesized by daily supervisor Jeffrey Zom, who also provided the PL spectrum dataset for temperatures $T = 15\,\mathrm{K}$ to $T = 300\,\mathrm{K}$ with excitation wavelength $\lambda_{\mathrm{exc}} = 300\,\mathrm{nm}$. Those results are given in figure 25. For low temperatures, the emission of the $VO_4{}^{3-}$ is visible, as a broad spectrum around 450 nm. When the temperature is increased, this emission decreases, while the light from Eu^{3+} increases in comparison.

From this data, we calculate the total counts received by the sensor, corrected for background light intensity, for emission from the $VO_4{}^{3-}$ and Eu^{3+} ions. Since the total count is proportional to the amount of photons emitted, the probability of europium emission can be calculated as

$$\text{Probability Eu emission} = \frac{\text{Counts Eu}}{\text{Counts Eu} + \text{Counts VO}_4} \tag{7.2}$$

The amount of counts per ion is found by summing over the corrected count multiplied by its wavelength interval. The two emission sources are distinguished by minimizing an Eu^{3+} template (wavelengths $520\,\mathrm{nm} \leq \lambda \leq 750\,\mathrm{nm}$ at $T = 300\,\mathrm{K}$) and using a moving average to remove the noise at the minimization region. These total counts are given in the appendix (32).

While there are more parameters in this model than datasets, we try to find suitable parameters due to how these parameters shape the model. For example, the activation energy determines the slope at the transition temperature T_1. The parameter $k_{m,0}$ determines the shift compared to the transition temperature T_1, as was seen in figure 19 (previous chapter). Meanwhile, k_t largely determines the the probability at the asymptotic boundaries.

From the experimental data, the slope of the ratio is the steepest for $T = 109.86\,\mathrm{K}$, therefore the transition temperature is chosen as $T_0 = 109.86\,\mathrm{K}$. To find for parameters E_a, $k_{m,0}$ and k_t, we use an iterative process which minimizes the root square error in a certain region when one parameters is variable and the other two parameters are fixed. For E_a and $k_{m,0}$, this is the transition region $0.4 < \text{ratio} < 0.95$. For the transfer rate k_t, the region $T < 100\,\mathrm{K}$ was chosen. The initial parameters for the iterative process are $E_a = 95\,\mathrm{meV}$, $k_{m,0} = 5\times10^1 \cdot k_{\mathrm{ev}}$ and $k_t = 2.5\times10^6 \cdot k_{\mathrm{ev}}$. After the process, we calculated the activation energy to be $E_a = (101 \pm 4)\,\mathrm{meV}$ for a migration rate $k_{m,0} = (42 \pm 4)\times10^4\,\mathrm{s}^{-1}$. The transfer rate was found as $k_t = (3.2 \pm 0.5)\times10^5 \cdot k_{\mathrm{ev}}$, where $(3.2 \pm 0.5)\times10^5$ has units Å. This process yielded a mean square error for the transition region of $(7 \pm 3)\times10^{-4}$ and a mean square error of for the low temperature region. These results are given in figure 26.

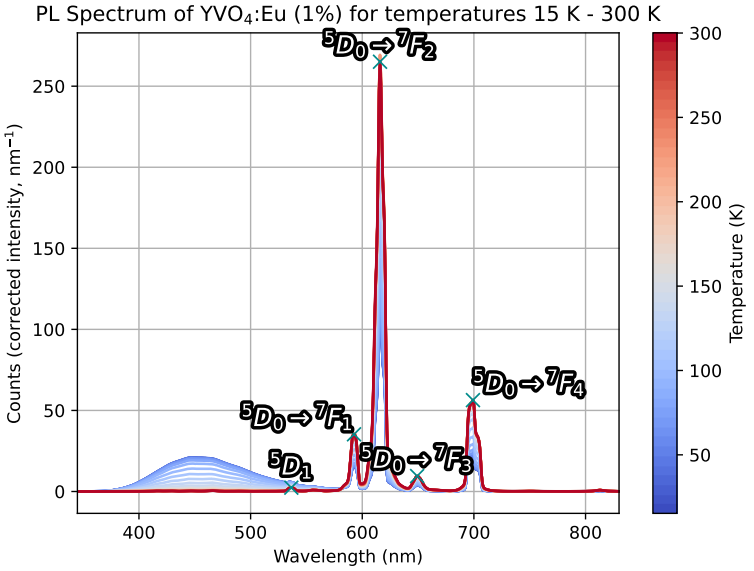


Figure 25: Photoluminescence spectrum ($\lambda_{\text{exc}} = 300\,\text{nm}$) of YVO_4:Eu^{3+} with 1% europium concentration in the temperature range of 15 K to 300 K. The y-axis represents is the counts per some interval wavelength corrected for the background light intensity. The peaks in counts are given by an X, and have the corresponding transition level of Eu^{3+}.

These results acquired from the model are averaged over an ensemble of multiple simulations. For the iterative process for $E_a, k_{m,0}$, we calculated 15 simulations over 30 iterations, and 10 to 110 simulations (depending on the iteration count) for 15 iterations. The uncertainty in the parameters $E_a, k_{m,0}$ and k_t are given by the standard deviation over from the 2nd iteration.

7.2.1 Discussion

We first discuss the data analysis method, then we examine the results regarding the model. A point of criticism is that the dataset used for finding the Eu/VO_4-ratio does not incorporate uncertainty from measurement. Secondly, since there is no measurement with VO_4-only emission and Eu-only emission, we have to acquire 'templates' to approximate roughly this ratio. And finally, during this analysis, sometimes the minimization is too strong, which creates a negative value of counts for certain wavelengths. We correct for this using a moving average to smooth out the noise, but also slightly underestimates the area of VO_4. This is partially negated by setting negative values at zero. We find that the counts of Eu^{3+} is overestimated meanwhile counts of $VO_4{}^{3-}$ is underestimated.

For the model, there is the issue of overfitting. In our results, we hand-picked the initial values k_t, E_a and $k_{m,0}$, which we then optimized from to find a set of parameters which suits best. Of course, we can choose many other initial values to optimize, thus we cannot speak that these parameters describe the reality completely. During the iteration process, we find that this process does not converge. Instead, we find a range of values acceptable for E_a and $k_{m,0}$. This can be helped by increasing the lattice size or increasing the amount of simulations, which makes the confidence interval smaller. If this is done, we must also account for the error of the experimental data. Also, we experimentally find an asymptote of 100% Eu emission at high temperatures. In the model, this would mean extremely high values of k_t. This is likely due to thermal quenching and other phenomena unaccounted for in the model.

Another aspect of the model is the accuracy of the values given by the model. For high temperatures $T > 100\,\text{K}$, we see in figure 25 that the confidence interval is relatively small, so the model yields predictable results. While increasing the lattice size and amount of simulations should decrease the confidence, our expe-

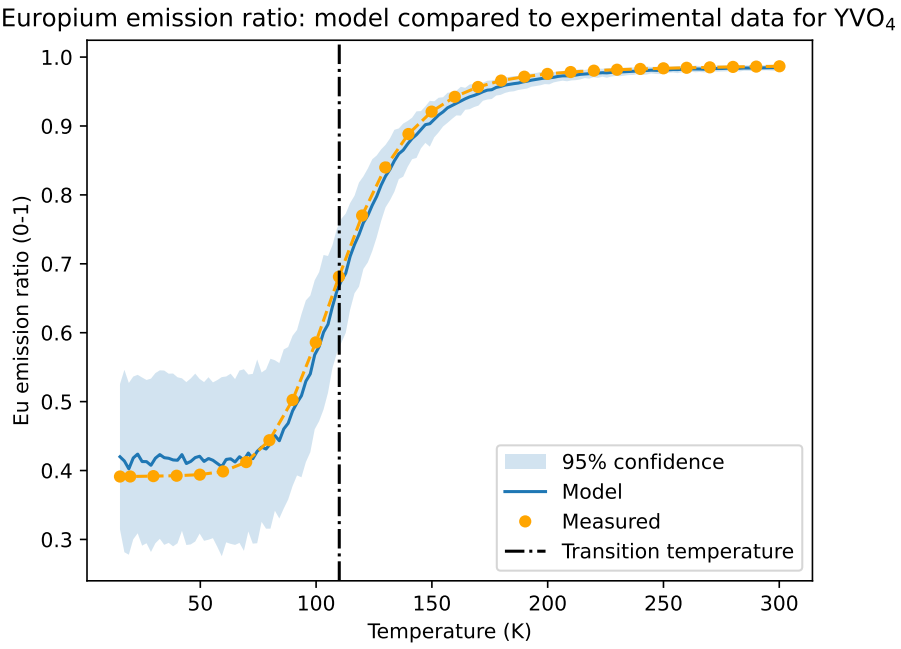


Figure 26: Europium (Eu^{3+}) emission ratio, as found experimentally (Measured) and computed using the model with a given confidence interval. Parameters are $E_a = 101\,\text{meV}$, $k_{m,0} = 42 \times 10^4\,\text{s}^{-1}$ and $k_t = 3.2 \times 10^5 \cdot k_{\text{ev}}$, lattice has size $5 \times 5 \times 5$ and is ensemble is taken over 150 simulations. The transition temperature $T_0 = 109.86\,\text{K}$ is also given, this is where the ratio has the greatest slope.

rience finds that increasing these parameters marginally improves the confidence interval. Other parameters are likely to dictate the confidence interval as well. For example, the concentration of 1% is relatively low, increasing the concentration to 10% could improve the confidence level. Finally, for these high temperatures where the model is 'confident', we do see that the model and experimental data almost agree.
The reason why the confidence interval is large for low temperatures, is likely due to high standard deviation in emission probability as was given in figure 17 (right figure). For low temperatures, the migration rate is relatively low (at $T = 109.86\,\text{K}$: $k_{m,0} = 42 \times 10^4\,\text{s}^{-1}$ is approximately $10 \cdot k_{\text{ev}}$) creating isolated europium centers which increases the standard deviation in the emission probability. Therefore, repeating the same simulation for different trap layouts can yield different results, which gives a broad confidence interval (in other words, the model is not confident). This confidence level can be improved by choosing a material which has a lower transfer rate k_t.

We try to compare the activation energy to those found in other literature and articles. In [HP75], an activation energy for YVO_4:Eu^{3+} was found of 48.0 meV and 19.3 meV by fitting the intensity and lifetime curves respectively. We remark that the authors of [HP75] have chosen a different temperature dependence for migration (of the form $[\exp(E_a/k_BT) + 1]^{-1}$), so the results may not be directly comparable. An article which does use the same temperature relation is by H. Nishimura and S. Nagata [NN88], but was applied to the material NaI instead of YVO_4:Eu^{3+}. They found an activation energy of 70 meV, which is in the same range as we found. In conclusion, the model gives an activation energy in the same range as found by others, but more research is required to sufficiently show the model is valid.

7.3 Further discussion about delocalized model

While the delocalized model converges to some emission ratio for high temperatures, depending on the parameters k_t, $k_{m,0}$ and E_a, and thus does not change. As discussed in the exciton theory chapter, if the temperature is increased, the non-radiative transfer rate also increases. We expect that for higher temper-

atures, the non-radiative transfer rate dominates over the radiative rate [18] , which reduces the efficiency at which light is emitted by the $VO_4{}^{3-}$ and Eu^{3+} ions, which we call thermal quenching. This can be implemented into the model in a simple manner, for example we can expand the state space of B to (1) the excited population B_{exc}, (2) radiative killed state B_{killed} and (3) non-radiative killed state B_{nr}. The rate would be $c(B_{\text{exc}}, B_{\text{killed}}) = k_{\text{ee}}$ and $c(B_{\text{exc}}, B_{\text{nr}}) = A_{\text{Eu}} \exp\left(-\frac{E}{k_B T}\right)$ where E is the energy of the transition. This would require us to define new parameters for the non-radiative transfer rate of VO_4 and Eu.

As was mentioned in the delocalized model chapter, increasing the concentration p marginally improved the emission probability of europium, see figure 16. As it turns out, for relatively high concentrations (for example $p = 0.2$), the probability of europium emission will decrease [19] due to Eu^{3+}-Eu^{3+} interactions, which is also called concentration quenching. If we know the mechanism for these Eu^{3+}-Eu^{3+} interactions, it is possible to incorporate concentration quenching into the model. First, we increase the state space to include an excited trapped state $B_{\text{exc},j}$ for the jth europium atom in the lattice. As before, we have a rate from $B_{\text{exc},j}$ to B_{killed} (and B_{nr} for the non-radiative state due to thermal quenching), but also a rate from $B_{\text{exc},j}$ to a certain non-radiative state B_{nr} for $j \neq k$ depending on, for instance, the distances between the jth europium atom and all other europium atoms in the lattice (or nearby ones). See figure 27 for an illustration of the proposed trapped states.

If we wish to expand this model to other materials, one might find an activator ion which has transitions which have similar but distinguishable lifetimes. In the case of Eu^{3+}, the life time has a relatively large difference between the fast and slow transitions (ranges $10\,\mu$s and 1 ms), so this was not a problem in our case. To account for other materials, we can expand the excited state to the energy levels for the activator or sensitizer. By measuring the lifetime at specific wavelengths for these transitions, we obtain the transition rate between the two levels which can be used in the model. In the case when two transitions of different lifetimes have a similar transition energy, a curve fit of two exponential decaying curves can be used.

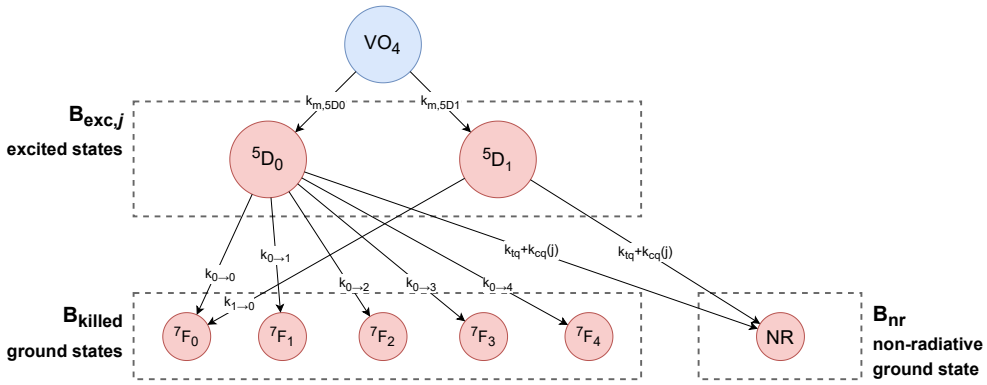


Figure 27: Example for suggested expanded trapped state for Eu^{3+} emission. We simplify the ${}^7F_j$ as the ground states ground states, while the ${}^5D_i$ represent the possible excited states of the jth Eu^{3+} ion. Here, $k_{m,5Di}$ denote energy transfer rate to excited state ${}^5D_i$, in our model the sum of these is $c(x, B_{\text{exc}})$. Similarly, $k_{i\to j}$ represents the transfer rate from excited state ${}^5D_i$ to state ${}^7F_j$. Finally, k_{tq} represents the thermal quenching rate, and k_{cq} is the concentration quenching rate (depends on location).

Another aspect of the model which could be improved is the lattice structure of the atoms. When we change the temperature the lattice lengths can change slightly, which can impact results of the model where the model is comparably confident. However, we expect this modification to not impact the results heavily. Similarly, a high concentration of europium might change the lattice structure which we do not account for,

[18] This is the case if the radiative emission rate increases slower than the non-radiative emission rate. To motivate this, in an article by F.L. Bregolin et al. [Bre+14], it was found that the PL decay time of Eu nanoparticles (both Eu^{2+} and Eu^{3+}) embedded in SiO_2 for the 618 nm wavelength (possibly ${}^5D_0 \to {}^7F_2$ transition in Eu^{3+}) was relatively constant for temperatures 12 K to 300 K.

[19] For an example of concentration quenching in $Rb_3LuV_2O_8$:Eu^{3+}, see figure 8c of [Dan+20].

which likely does not change the results much. At the lattice level, the material in reality consists of grains with boundaries and/or certain defects, such as holes or shifts within the lattice. To model this accurately, sufficient experimental methods are required to estimate sizes of grains and how excitons migrate at grain boundaries. But also on which defects are present and how they influence luminescent properties. Since estimating these properties can be difficult, it can be difficult to implement these defects from experimental data. Conversely, it might be possible to implement certain defects and derive from luminescence measures certain defect concentrations.

While the model can be improved, it can be used for a plethora of settings. For example, we do not have to assume hopping to nearest neighbors to be at a fixed rate. For instance, an anisotropic material can also be used, where the migration rate can depend on the distance to the nearest neighbor, creating directional hopping rates. Another example is as follows. If a crystal is small enough, for example within 100 nm to 100 nm (for lattice lengths of 10 Å this is a lattice of $100 \times 100 \times h$), we can adjust the boundary conditions to model the complete nanocrystal. It might be of interest how the exciton moves when it is nearby a boundary of the nanocrystal, and how luminescent properties are affected by this.
In addition to anisotropic hopping rates, we can also consider hopping rates between next nearest neighbors or all other possible sites. Since $k_m \propto |V_{ij}|^2$ where V_{ij} represents the overlapping integral of the potential between some atom i and j, it is possible to determine how the migration rate k_m changes over distance if the wave functions of atoms i and j are known [20].

On the topic of small lattices, we suggest that the model can be useful for narrow excitation laser beams. For example, for an excitation wavelength of 300 nm, the diffraction limit for a numerical aperture of 1.4 would be approximately 107 nm. Assuming that the probability of exciton creation is proportional to its light intensity, we can change the initial probability vector $\mathbf{p}$ (see section 6.2.1) to change depending on the intensity of the location, which can be described by the Airy disk for diffraction. Given enough computational power, a lattice size of the order $100 \times 100 \times h$ to $200 \times 200 \times h$ should be possible.

From a computational standpoint, it might be interesting to implement how numerical calculation of the model can be optimized. Currently, calculations are done at relatively large matrices, for example a $5 \times 5 \times 5$ lattice has size 253, and solving tens or hundreds of 253×253 matrices can be computationally expensive. The setting is similar for calculating e^{tL}. Other interesting aspects are numerical convergence rate and accuracy of the model.

At last, we also highlight our assumptions regarding the underlying theory used for this model. Our assumptions made for the r^{-6}-relation regarding the energy transfer or Markovian hopping rate (the assumptions made in the Holstein model) can be flawed and different results compared to experimental data. For future research, we suggest a comparison between Förster and Dexter energy transfer relations and relating this to experimental data. We also mention that we only considered a hopping rate as described by equation 2.9. However, at low temperatures[21] it has been shown that polaron hopping does approach zero[Asa+13], which is important for modeling at these low temperatures.

7.4 Discussion about localized model

We discuss some possible future research about the localized model. For a more extendable localized model, we suggest to make it 'delocalized' also, by implementing a r^{-6}-type potential. Also, one might further apply the parabolic Anderson model to find $R(t)$ exactly, or in a simulated manner. Another possibility is changing $\mathbb{Z}^d$ to a more realistic crystal lattice by adapting the generator L.

[20] As a suggestion, if the potential is approximated by delta peaks, one can find that $|V_{ij}|^2 \propto \exp(\alpha \cdot r_{ij})$, where r_{ij} is the distance between the atoms and α is some constant. Similarly, [Lu+15] also presents an exponential relation.
[21] For example, less than 50 K for YVO_4:Eu^{3+}.

8 Conclusions

We found that, in theory, exciton hopping, energy transfer, and decay can be described using a continuous-time Markov chain since the transfer between these states is given by rates. In general, connecting mathematical concepts found in Markov chains, such as semigroups and Feynman-Kac, was very useful in both the localized and delocalized models. While we discussed large deviation theory mainly due to Donsker and Varadhan, it also lead to interesting properties for the generators, which we used in both models.

For the localized model, we have found interesting results which describe exciton diffusion in a neighborhood of traps. Similarly, we can extend this model to one of many particles without many assumptions. It is very theoretical in nature, but if asymptotics are found for a delocalized potential, one might apply these results to real experimental data. Nonetheless, it presents an interesting case of soft traps of the trapping problem, which connects a theoretical mathematical model to a real world example of physics and chemistry.

In the delocalized setting, we found that the model, when described by a Markov chain, can adequately describe luminescence within its confidence level, compared to experimental data. We found that there exist many mathematical methods to model certain kinds of random variables, such as mean lifetimes of excitons or probability ratios. There is a lot of room for adjusting and expanding this model for evaluating luminescence phenomena for not only europium doped vanadate phosphors. Not only is the model flexible for the related parameters, such as migration rates, temperature or dopant concentration, but also can provide smooth (but sometimes inconsistent) curves for many luminescent phenomena.

We hope that the insight gained within this thesis can be expanded to more settings, such as different materials or different luminescence mechanisms. The results build a foundation as reference work for future research on exciton movement and trapping, both in a theoretical and simulation fashion.

9 References

References

[Hol59a] T Holstein. "Studies of polaron motion: Part I. The molecular-crystal model". In: *Annals of Physics* 8.3 (Nov. 1959), pp. 325–342. ISSN: 0003-4916. DOI: 10.1016/0003-4916(59)90002-8. URL: http://dx.doi.org/10.1016/0003-4916(59)90002-8.

[Hol59b] T. Holstein. "Studies of polaron motion: Part II. The "small" polaron". In: *Annals of Physics* 8.3 (Nov. 1959), pp. 343–389. ISSN: 0003-4916. DOI: 10.1016/0003-4916(59)90003-x. URL: http://dx.doi.org/10.1016/0003-4916(59)90003-X.

[Lia70] W Y Liang. "Excitons". In: *Physics Education* 5.4 (July 1970), pp. 226–228. ISSN: 0031-9120. DOI: 10.1088/0031-9120/5/4/003. URL: http://dx.doi.org/10.1088/0031-9120/5/4/003.

[DV75a] M. D. Donsker and S. R. S. Varadhan. "Asymptotic evaluation of certain markov process expectations for large time, I". In: *Communications on Pure and Applied Mathematics* 28.1 (Jan. 1975), pp. 1–47. ISSN: 1097-0312. DOI: 10.1002/cpa.3160280102. URL: http://dx.doi.org/10.1002/cpa.3160280102.

[DV75b] M. D. Donsker and S. R. S. Varadhan. "Asymptotics for the wiener sausage". In: *Communications on Pure and Applied Mathematics* 28.4 (July 1975), pp. 525–565. ISSN: 1097-0312. DOI: 10.1002/cpa.3160280406. URL: http://dx.doi.org/10.1002/cpa.3160280406.

[HP75] Chang Hsu and Richard C. Powell. "Energy transfer in europium doped yttrium vanadate crystals". In: *Journal of Luminescence* 10.5 (June 1975), pp. 273–293. ISSN: 0022-2313. DOI: 10.1016/0022-2313(75)90051-4. URL: http://dx.doi.org/10.1016/0022-2313(75)90051-4.

[BB85] Harald Böttger and Valerij V Bryksin. *Hopping conduction in solids*. en. Weinheim, Germany: Wiley-VCH, 1985.

[NN88] Hitoshi Nishimura and Shunji Nagata. "Hopping motion of self-trapped excitons in NaI". In: *Journal of Luminescence* 40–41 (Feb. 1988), pp. 429–430. ISSN: 0022-2313. DOI: 10.1016/0022-2313(88)90266-9. URL: http://dx.doi.org/10.1016/0022-2313(88)90266-9.

[GM98] J. Gärtner and S. A. Molchanov. "Parabolic problems for the Anderson model: II. Second-order asymptotics and structure of high peaks ". In: *Probability Theory and Related Fields* 111.1 (June 1998), pp. 17–55. ISSN: 1432-2064. DOI: 10.1007/s004400050161. URL: http://dx.doi.org/10.1007/s004400050161.

[SM02] J.F. Suyver and A. Meijerink. *Europium safeguards the Euro*. 2002. URL: https://dspace.library.uu.nl/handle/1874/25831.

[JP04] Jean Jacod and Philip Protter. *Probability Essentials*. Springer Berlin Heidelberg, 2004. ISBN: 9783642556821. DOI: 10.1007/978-3-642-55682-1. URL: http://dx.doi.org/10.1007/978-3-642-55682-1.

[Ron07] Cornelis R Ronda, ed. *Luminescence*. en. Weinheim, Germany: Wiley-VCH Verlag, Oct. 2007.

[Den08] Frank Den Hollander. *Large Deviations*. Fields Institute Monographs. American Mathematical Society, June 2008.

[AD09] Alexandre S Alexandrov and Jozef T Devreese. *Advances in polaron physics*. en. 2010th ed. Springer Series in Solid-State Sciences. Berlin, Germany: Springer, Oct. 2009. ISBN: 9783642018961. DOI: 10.1007/978-3-642-01896-1. URL: https://doi.org/10.1007/978-3-642-01896-1.

[CRC09] M.D. Chambers, P.A. Rousseve, and D.R. Clarke. "Decay pathway and high-temperature luminescence of Eu3+ in Ca2Gd8Si6O26". In: *Journal of Luminescence* 129.3 (Mar. 2009), pp. 263–269. ISSN: 0022-2313. DOI: 10.1016/j.jlumin.2008.10.008. URL: http://dx.doi.org/10.1016/j.jlumin.2008.10.008.

[KEL09] Olivier d'ALLIVY KELLY. *File:Polaron scheme1.svg - Wikimedia Commons — commons.wikimedia.org*. https://commons.wikimedia.org/wiki/File:Polaron_scheme1.svg. [Accessed 01-08-2025]. 2009.

[DZ10] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Springer Berlin Heidelberg, 2010. ISBN: 9783642033117. DOI: 10.1007/978-3-642-03311-7. URL: http://dx.doi.org/10.1007/978-3-642-03311-7.

[Tou11] Hugo Touchette. *A basic introduction to large deviations: Theory, applications, simulations*. 2011. DOI: 10.48550/ARXIV.1106.4146. URL: https://arxiv.org/abs/1106.4146.

[Asa+13] Kamal Asadi et al. "Polaron hopping mediated by nuclear tunnelling in semiconducting polymers at high carrier density". In: *Nature Communications* 4.1 (Apr. 2013). ISSN: 2041-1723. DOI: 10.1038/ncomms2708. URL: http://dx.doi.org/10.1038/ncomms2708.

[Jai+13] Anubhav Jain et al. "The Materials Project: A materials genome approach to accelerating materials innovation". In: *APL Materials* 1.1 (2013), p. 011002. ISSN: 2166532X. DOI: 10.1063/1.4812323. URL: http://link.aip.org/link/AMPADS/v1/i1/p011002/s1%5C&Agg=doi.

[Bre+14] F.L. Bregolin et al. "Low temperature and decay lifetime photoluminescence of Eu and Tb nanoparticles embedded into SiO2". In: *Journal of Luminescence* 153 (Sept. 2014), pp. 144–147. ISSN: 0022-2313. DOI: 10.1016/j.jlumin.2014.03.015. URL: http://dx.doi.org/10.1016/j.jlumin.2014.03.015.

[Ste+14] Vera Stehr et al. "Singlet Exciton Diffusion in Organic Crystals Based on Marcus Transfer Rates". In: *Journal of Chemical Theory and Computation* 10.3 (Feb. 2014), pp. 1242–1255. ISSN: 1549-9626. DOI: 10.1021/ct500014h. URL: http://dx.doi.org/10.1021/ct500014h.

[Tes14] Gerald Teschl. *Mathematical methods in quantum mechanics*. 2nd ed. Graduate studies in mathematics. Providence, RI: American Mathematical Society, Oct. 2014.

[Lu+15] Nianduan Lu et al. "Charge carrier hopping transport based on Marcus theory and variable-range hopping theory in organic semiconductors". In: *Journal of Applied Physics* 118.4 (July 2015). ISSN: 1089-7550. DOI: 10.1063/1.4927334. URL: http://dx.doi.org/10.1063/1.4927334.

[Zho+15] Jiangcong Zhou et al. "Luminescence study of a self-activated and rare earth activated Sr3La(VO4)3phosphor potentially applicable in W-LEDs". In: *J. Mater. Chem. C* 3.13 (2015), pp. 3023–3028. ISSN: 2050-7534. DOI: 10.1039/c4tc02783c. URL: http://dx.doi.org/10.1039/C4TC02783C.

[Kön16] Wolfgang König. *The Parabolic Anderson Model: Random Walk in Random Potential*. Springer International Publishing, 2016. ISBN: 9783319335964. DOI: 10.1007/978-3-319-33596-4. URL: http://dx.doi.org/10.1007/978-3-319-33596-4.

[Lar+17] Ask Hjorth Larsen et al. "The atomic simulation environment—a Python library for working with atoms". In: *Journal of Physics: Condensed Matter* 29.27 (2017), p. 273002. URL: http://stacks.iop.org/0953-8984/29/i=27/a=273002.

[Sev+17] D. Sevic et al. "Characterization and luminescence kinetics of Eu 3+ doped YVO 4 nanopowders". In: *Materials Research Bulletin* 88 (Apr. 2017), pp. 121–126. ISSN: 0025-5408. DOI: 10.1016/j.materresbull.2016.12.021. URL: http://dx.doi.org/10.1016/j.materresbull.2016.12.021.

[Dan+20] Peipei Dang et al. "Highly Efficient Cyan-Green Emission in Self-Activated Rb3RV2O8 (R = Y, Lu) Vanadate Phosphors for Full-Spectrum White Light-Emitting Diodes (LEDs)". In: *Inorganic Chemistry* 59.9 (Apr. 2020), pp. 6026–6038. ISSN: 1520-510X. DOI: 10.1021/acs.inorgchem.0c00015. URL: http://dx.doi.org/10.1021/acs.inorgchem.0c00015.

[Fra+21] Cesare Franchini et al. "Polarons in materials". In: *Nature Reviews Materials* 6.7 (Mar. 2021), pp. 560–586. ISSN: 2058-8437. DOI: 10.1038/s41578-021-00289-w. URL: http://dx.doi.org/10.1038/s41578-021-00289-w.

[Sch21] René L Schilling. *Brownian motion*. en. 3rd ed. De Gruyter Textbook. Berlin, Germany: De Gruyter, Sept. 2021. ISBN: 9783110741278. DOI: 10.1515/9783110741278. URL: https://doi.org/10.1515/9783110741278.

[Nee22] Jan van Neerven. *Functional Analysis*. Cambridge University Press, June 2022. ISBN: 9781009542463. DOI: 10.1017/9781009232487. URL: http://dx.doi.org/10.1017/9781009232487.

[CB23] Stefania Castelletto and Alberto Boretti. "Luminescence solar concentrators: A technology update". In: *Nano Energy* 109 (May 2023), p. 108269. ISSN: 2211-2855. DOI: 10.1016/j.nanoen.2023.108269. URL: http://dx.doi.org/10.1016/j.nanoen.2023.108269.

[SBC23] Pablo Serna-Gallén, Héctor Beltrán-Mir, and Eloísa Cordoncillo. "Practical guidance for easily interpreting the emission and physicochemical parameters of Eu3+ in solid-state hosts". In: *Ceramics International* 49.24 (Dec. 2023), pp. 41078–41089. ISSN: 0272-8842. DOI: 10.1016/j.ceramint.2023.01.141. URL: http://dx.doi.org/10.1016/j.ceramint.2023.01.141.

[KDB24] Soorya N. Kabekkodu, Anja Dosen, and Thomas N. Blanton. "PDF-5+: a comprehensive Powder Diffraction File™ for materials characterization". In: *Powder Diffraction* 39.2 (Apr. 2024), pp. 47–59. ISSN: 1945-7413. DOI: 10.1017/s0885715624000150. URL: http://dx.doi.org/10.1017/S0885715624000150.

[Mül+24] Rosa Müller et al. "Transparent, Sprayable Plastic Films for Luminescent Down-Shifted-Assisted Plant Growth". In: *Advanced Materials Technologies* 10.6 (Nov. 2024). ISSN: 2365-709X. DOI: 10.1002/admt.202400977. URL: http://dx.doi.org/10.1002/admt.202400977.

[Wan+24] Zi Wang et al. "On the application of Marcus–Hush theory to small polaron chemical dynamics in oxides: its relationship to the Holstein model and the importance of lattice–orbital symmetries". In: *Physical Chemistry Chemical Physics* 26.6 (2024), pp. 4812–4827. ISSN: 1463-9084. DOI: 10.1039/d3cp05218d. URL: http://dx.doi.org/10.1039/d3cp05218d.

[OKa] Mary O'Kane. *Exciton: An Introduction — ossila.com*. https://www.ossila.com/pages/what-is-an-exciton. [Accessed 01-08-2025].

A Appendix

A.1 Notation overview

Here we list an overview of symbols and notation, consisting of simple definitions not directly explained or frequently used notation, used throughout the thesis.

Notation or symbol	Meaning
$\mathbf{1}_A(x)$	indicator function of set A, $\mathbf{1}_A(x) = 1$ if $x \in A$, otherwise $\mathbf{1}_A(x) = 0$.
$\mathbf{1}(y = x)$	indicator for case $y = x$, i.e. $\mathbf{1}(y = x) = 1$ if $y = x$ and 0 otherwise.
δ_{ab}	Kronecker delta, equivalent to $\mathbf{1}(a = b)$.
$\partial_x f$	derivative of f with respect to x, i.e. $\partial_x f = \frac{\partial}{\partial x} f$
$\partial_x^2 f$	second derivative of f with respect to x, i.e. $\partial_x^2 = \frac{\partial^2}{\partial x^2} f$
Δ^d (context $\mathbb{Z}^d$)	Laplacian operator in $\mathbb{Z}^d$, defined as $(\Delta^d f)(x) = \sum_{y \sim x}(f(y) - f(x))$ where $y \sim x$ if y is a nearest neighbor to x (i.e. $y = x \pm 1$ for $\mathbb{Z}$).
$\mathbb{E}_x[\cdot]$	$\mathbb{E}[\cdot \mid X_t = x]$ for some given Markov process X_t.
$a \simeq b$	Logarithmic equivalence (see equation 4.13). We have $a \simeq b$ if and only if $\lim_{n\to\infty} \frac{1}{n}(\log(a_n) - \log(b_n)) = 0$.
i.i.d.	independent and identically distributed
pdf	probability density function
mgf	moment generating function
CLT	Central Limit Theorem
LDP	Large Deviation Principle (Definition 4.6)
PAM	Parabolic Anderson Model
CTMT	continuous-time Markov chain
$X \sim \mathcal{N}(\mu, \sigma^2)$	X is normally distributed with mean μ and variance σ^2
$X \sim \text{Exp}(\lambda)$	X is exponentially distributed with rate λ (pdf is $\lambda e^{-\lambda x}$).
$X \sim \text{Poisson}(\lambda)$	X is Poisson distributed with parameter λ (pdf is $\frac{\lambda^k}{k!} e^{-\lambda}$)
f^* with f a function	f^* is the Legendre/Legendre-Fenchel transformation, as given by Definitions 4.4 and 4.9 of f.
$\mathcal{X}^*$ with $\mathcal{X}$ a set	$\mathcal{X}^*$ is the dual space of $\mathcal{X}$, as given by Definition 4.7.
$C_0(X)$	space of all continuous functions $f : X \to F$ (F being a field) which vanish at infinity.
$C^n(X)$	space of all n continuously differentiable functions $f : X \to F$ (F being a field).
$\mathcal{F}_0$, $\mathcal{F}_0^d$	set of probability density functions which are continuously differentiable, strictly positive and are in $L^1(\mathbb{R})$ (see equation 4.36). Then $\mathcal{F}_0^d$ extends the notion to $\mathbb{R}^d$.
$\langle \varphi_\mu, f \rangle$ with φ a functional on $C_0(X)$ with measure μ	$\langle \varphi_\mu, f \rangle = \varphi_\mu(f) = \int_X f \, \mathrm{d}\mu$
$B(x, R)$	d-dimensional ball (depends on context) consisting of all points y such that $\|x - y\|_2 < R$. Can be used for $\mathbb{R}^d$ or $\mathbb{Z}^d$.

Notation or symbol	Meaning
$\operatorname{supp}(f)$, f having support on ...	$\operatorname{supp}(f) = \{x \in X \mid f(x) \neq 0\}$
B_t, $(B_t)_{t\in T}$	Brownian motion (dimensions depending on context) at time t or time in set T.
Z_t (section Localized model)	$Z_t = \exp\left(-\gamma_2 \int_0^t \xi(X_s)\,\mathrm{d}s\right)$
$\mathcal{O}(f(x))$	Big-O notation, $\exists N \geq 0$ such that $\lvert g(x)\rvert = \mathcal{O}(f(x)) \geq \lvert f(x)\rvert$ for all $x \geq N$.
$o(f(x))$	Little-o notation
Å, angstrom	unit representing m^{-10}
ML	mobility lattice
TL	trapping lattice
A	trapping state of vanadate emission
B, B_{exc}, B_{killed}	europium trapping state, where B_{exc} is the excited population and B_{killed} is the killed population
k_m	migration rate, units s^{-1}
k_{ev}	vanadate emission rate, units s^{-1}
k_{ee}	europium emission rate, units s^{-1}
k_t	europium energy transfer rate, units $\mathrm{m}^6\,\mathrm{s}^{-1}$ or $\mathrm{Å}^6\,\mathrm{s}^{-1}$
root squared error, RSE	given two datasets $x[i]$ and $y[i]$, the root squared error is $\sqrt{\sum_{i=1}^{n}(x[i]-y[i])^2}$; the mean square error MSE is RSE^2/n.
k_B	Boltzmann's constant, is $1.380649 \times 10^{-23}\,\mathrm{J}\cdot\mathrm{K}^{-1}$

A.2 Rate in single-transfer state space

For decay curves, we describe the intensity with respect to time of a certain light source. In this work, we refer to luminescence of excitons, which we describe by a Markov process. In this manner, assume we have a continuous-time Markov chain with a state space $\{1, 2\}$, where we have the rate $c(1, 2) = k$ and $c(2, 1) = 0$. Here k is the decay rate in the sense of the Markov chain.
We know the transition matrix P is given by e^{tL}. For this approximation, assume every exciton decays into one photon with a fixed energy. Then the intensity measured is proportional to how many photons are emitted per unit of time. Therefore, $I \propto \frac{d}{dt}P_2(t) = P_2'(t)$ where $P_2(t)$ denotes the probability $\mathbb{P}(X_t = 2 \mid X_0 = 1)$.

To solve for k, we may write $P'(t) = \frac{d}{dt}e^{tL} = Le^{tL}$ which gives

$$\begin{pmatrix} P_1'(t) \\ P_2'(t) \end{pmatrix} = \begin{pmatrix} -k & k \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1(t) \\ P_2(t) \end{pmatrix} \tag{A.1}$$

From this we obtain the differential equation $P_1'(t) = -kP_1(t)$. Since we start at 1, we have $P_1(0) = 1$ and $P_2(0) = 0$. So, the solution for $P_1(t)$ is

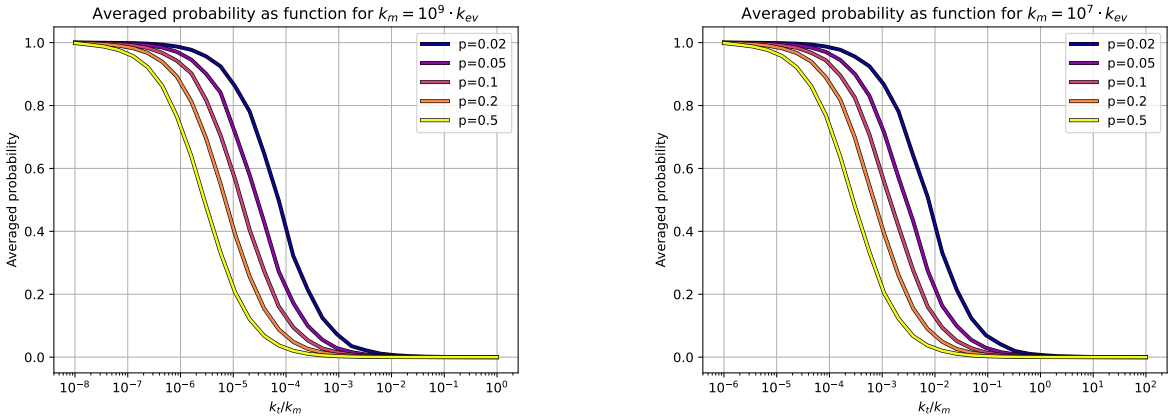
$$P_1(t) = e^{-kt} \tag{A.2}$$

Since the sum of the probability is 1, we find

$$P_2(t) = 1 - P_1(t) = 1 - e^{-kt} \tag{A.3}$$

Differentiating gives $P_2'(t) = ke^{-kt}$. We conclude that the intensity curve decays according to e^{-kt}. Often the parameter for this decay is named the lifetime τ where $I \propto e^{-t/\tau}$. So we find that $k = 1/\tau$.

A.3 Figures



(a) $k_m = 10^9 \cdot k_{\text{ev}}$

(b) $k_m = 10^7 \cdot k_{\text{ev}}$

(c) $k_m = 10 \cdot k_{\text{ev}}$

(d) $k_m = 10^{-7} \cdot k_{\text{ev}}$

Figure 28: This is the same as figure 16 with different values of k_m but with the same crystal structure ($10 \times 10 \times 10$ lattice). This figure therefore skips $k_m = 10^5 \cdot k_{\text{ev}}$. The above figures are shifted by the same order k_m is shifted, hence the figures look the same. For extreme cases for k_m very small, we examine that the distance between the level curves from p increases. This occurs roughly when $k_m = 10 \cdot k_{\text{ev}} \approx 10^5\,\text{s}^{-1}$. It is likely that due to the low migration rate, it is harder for the excitons to be trapped by europium atoms it is more difficult to find these atoms. Therefore, especially for low concentrations, the probability will be higher since we expect less europium emission.

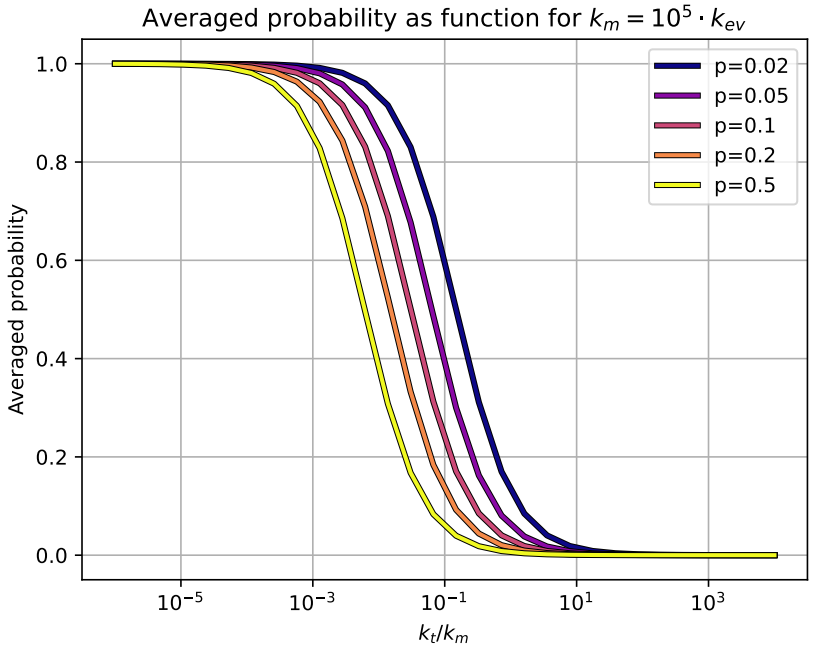
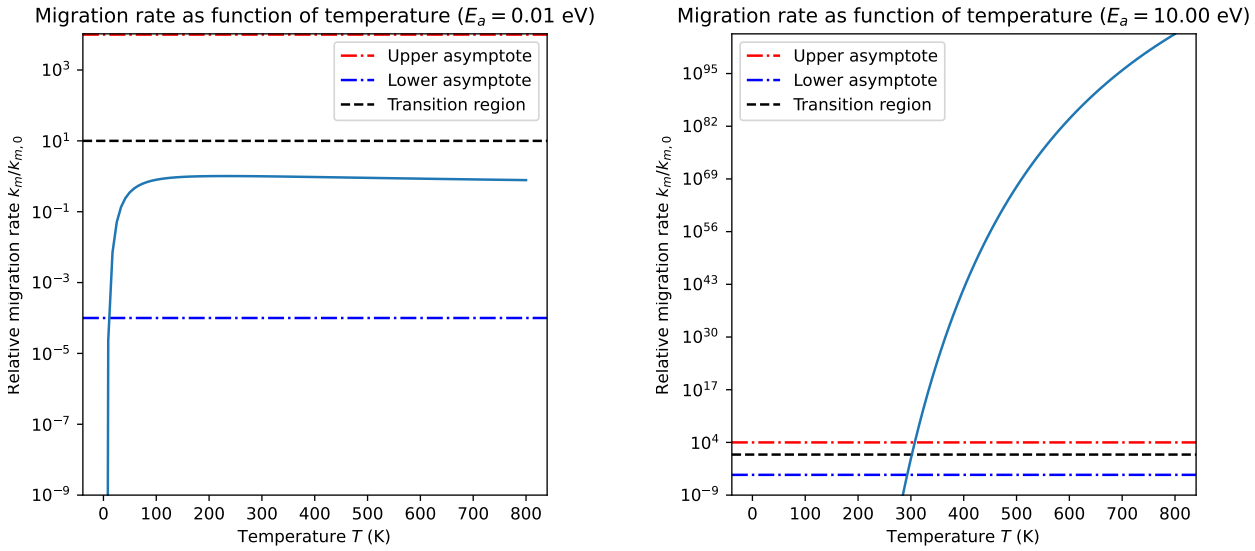


Figure 29: Averaged probability $\langle \mathbb{P}_x(\tau_A < \tau_B) \rangle$ for $k_m = 10^5 \cdot k_{\text{ev}}$ for the crystal YVO_4:Eu^{3+} ($10 \times 10 \times 10$ lattice). This is similar to the figure of 16, but shifted a little bit to the right. A possible explanation is that in YVO_4:Eu^{3+}, the vanadium atoms have less nearest neighbors and therefore the migration rate should be higher (and k_t/k_m lower) to compensate.



(a) $E_a = 0.01\,\text{eV}$

(b) $E_a = 10.0\,\text{eV}$

Figure 30: The migration rate as calculated using equation (6.25) divided by $k_{m,0}$ for different activation energy. For reference, see figure 18. While a low activation energy is feasible ($E_a \leq 0.01\,\text{eV}$) if we choose $k_{m,0}$ carefully, we will only change in behavior caused by the migration rate in low temperatures ($T < 100\,\text{K}$). In literature, we expect drastic change in intensity between 0 K to at least 300 K, which this does not reflect. For high activation energies ($E_a \geq 10\,\text{eV}$), the change is too drastic and focused on one specific temperature. If we choose $k_{m,0} = k_{\text{ev}}$, then our asymptotic range would be in the range of tenths of kelvins if not less.

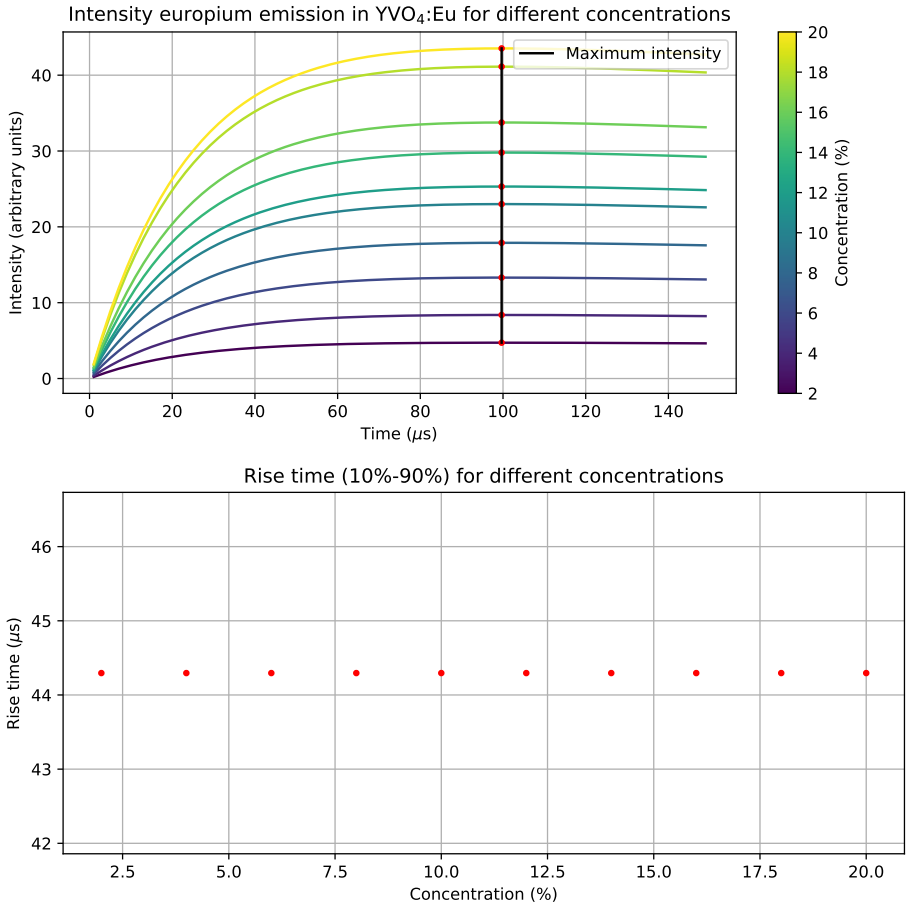


Figure 31: Rise time in YVO_4:Eu^{3+} ($5 \times 5 \times 5$ lattice) for $k_m = 10^5 \cdot k_{\text{ev}}, k_t = 1 \cdot k_{\text{ev}}$. For reference, see figure 22. While for low k_t we still experience a rise time (here $\tau = 4.4295 \cdot 10^{-5}$ s), this value is independent of the europium concentration. Changing k_m for higher or lower values (e.g. $10^8 \cdot k_{\text{ev}}$ or $10^{-5} \cdot k_{\text{ev}}$) does not change this value.

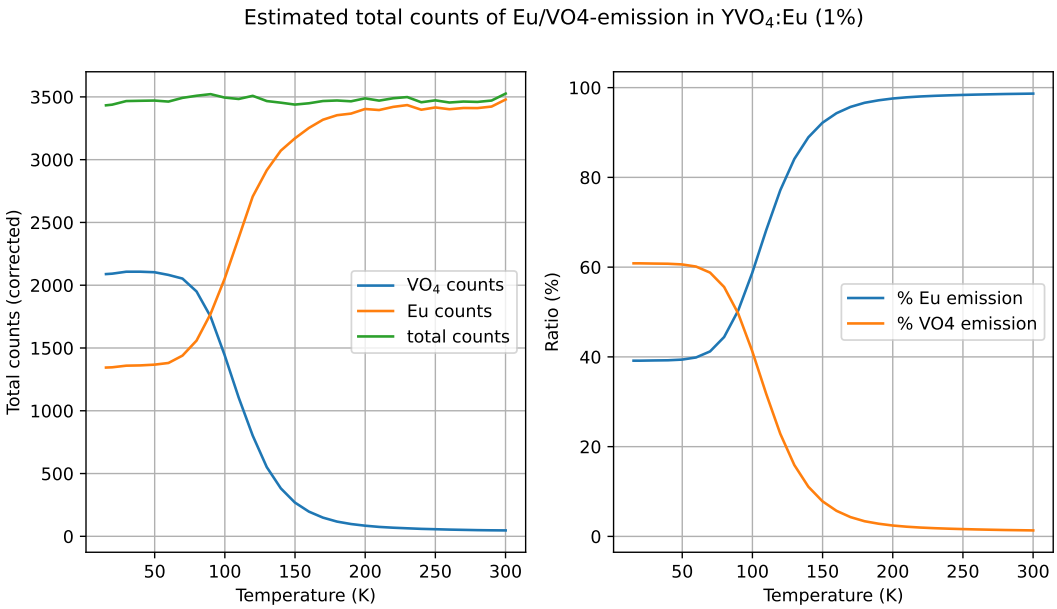


Figure 32: Left figure: calculated total counts (or area) of the PL spectrum, separated for VO_4-emission and Eu-emission, given for temperatures 15 K to 300 K. Right figure: Counts expressed as a ratio of the total counts, per temperature.