CONCENTRATION INEQUALITIES FOR FLUCTUATIONS IN CLASSICAL AND QUANTUM MARKOV PROCESSES

GEORGE BAKEWELL-SMITH

THESIS SUBMITTED TO THE UNIVERSITY OF NOTTINGHAM
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN MATHEMATICS

School of Mathematical Sciences
University of Nottingham

Supervisors:

Dr. Madalin Guta
Prof. Juan P. Garrahan

January 2025

ABSTRACT

Markov processes describe a plethora of physical systems in nature, including the mechanics of molecular motors, chemical reactions, and even financial markets. In the framework of continuous-time Markov processes, there are many important quantities of interest. One example are dynamical observables: time-integrated functionals of stochastic trajectories. Another are first passage times (FPTs), which are the time taken for an observable to reach a fixed threshold. The fluctuations of these two quantities are bounded from below by the thermodynamic uncertainty relations (TURs), a fundamental result which captures the tradeoff between their precision, and physical parameters such as entropy production or activity. Whilst these lower bounds are well-known, less attention has been given to finding upper bounds on fluctuations.

We first prove the existence of general upper bounds, at any time, on the variance of any linear combination of fluxes for classical, continuous-time Markov processes. These are derived by considering perturbed dynamics and applying techniques in concentration theory, in particular the Cramér-Chernoff method. We call these bounds "inverse thermodynamic uncertainty relations". Spectral methods allow us to express the bounds in terms of parameters of the dynamics which include the symmetrised spectral gap of the generator and max/min escape rates, alongside observable-dependent quantities. Afterwards we provide a concentration inequality for dynamical observables, which upper bounds the probability distribution for finite time. We then extend these results to FPTs for the dynamical activity, as well as a tail bound for general counting observables. Finally we generalise our FPT results to the quantum framework for general quantum Markov processes and reset processes, including a tail bound for FPTs of general counts, i.e. counts of a subset of emissions. We also prove a large deviation principle for FPTs of classical and quantum counting processes.

Our results have several consequences and applications. By providing upper bounds on the relative uncertainty, the range of estimation errors is bounded from *both* sides. These findings suggest that spectral quantities limit the range of fluctuations. The observable-dependent parameters in the bounds offer an advantage over the traditional TURs for precision estimation. Inverse TURs provide bounds on the accuracy of classical or quantum clocks or the efficiency of heat engines. The concentration inequalities can be used in finite-regime parameter estimation of classical and quantum processes, with open quantum dynamics being particularly relevant in experimental situations.

Affiliations:

School of Mathematical Sciences, University of Nottingham

Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems (CQNE), University of Nottingham

PUBLICATIONS

[1] G. Bakewell-Smith, F. Girotti, M. Guţă, and J. P. Garrahan. General Upper Bounds on Fluctuations of Trajectory Observables. *Physical Review Letters* 131.19 (Nov. 2023), p. 197101. DOI: 10.1103/PhysRevLett.131.197101.

[2] G. Bakewell-Smith, F. Girotti, M. Guţă, and J. P. Garrahan. Bounds on Fluctuations of First Passage Times for Counting Observables in Classical and Quantum Markov Processes (May 2024). DOI: 10.48550/arXiv.2405.09669.

ACKNOWLEDGMENTS

I would first like to thank my supervisor, Madalin Guta, for his guidance and support during my PhD, and for the many fruitful discussions over coffee. Secondly, I am grateful to Juan Garrahan, my other supervisor, for his expertise and advice. My PhD has given me the opportunity to begin my research journey, and to travel and present my work, which would not have been possible without them. Both Madalin and Juan, as well as Adam Gammon-Smith in physics, were also instrumental in me securing a fellowship whilst I was undertaking my final year.

Special thanks go to Federico Girotti, for his mentorship during the first part of my PhD, which has made me a better mathematician, and for giving me another excuse to visit Italy.

None of this would have started without Ioannis Rousochatzakis at Loughborough University, whose final-year project motivated me to apply for a PhD.

I would like to thank Niamh Martin, for putting up with me as an office-mate, and for making it a cheery place. Thanks also to Alfie Godley, who has been part of this process over the past few years and helped me with the thesis template.

Thanks go to University of Nottingham Triathlon Club for giving me something to do outside of academic studies.

Finally, I would like to thank my family: Mum, Dad, Beth and Emily for all their support and encouragement during my studies.

CONTENTS

LIST OF FIGURES

ACRONYMS

ATP	adenosine triphosphate
CLT	central limit theorem
CUR	clock uncertainty relation
DMRG	density matrix renormalisation group
FPT	first passage time
KMS	Kubo-Martin-Schwinger
KUR	kinetic uncertainty relation
LDP	large deviation principle
MCMC	Markov chain Monte Carlo
MGF	moment generating function
NESS	non-equilibrium stationary state
PEPS	projected entangled pair states
SLLN	strong law of large numbers
TUR	thermodynamic uncertainty relation

Freedom is the freedom to say that two plus two make four.
If that is granted, all else follows.

— George Orwell, *1984*

1

INTRODUCTION

At equilibrium, macroscopic systems exhibit small and predictable fluctuations of observable quantities around their expected value, a feature fundamental to statistical physics [1–3]. These fluctuations, along with quantities like thermodynamic averages can be quantitatively described using the Boltzmann distribution and standard ensemble methods used in statistical mechanics.

The introduction of a dissipative environment exchanging heat, work or matter etc. in an irreversible manner, disturbs equilibrium. The resulting complexity leads to stronger fluctuations than those experienced in contact with an equilibrium environment [4–6]. Alternatively, a small system even at equilibrium can experience large fluctuations of observed quantities. When a system experiences time-independent driving (e.g. a constant supply of particles) it will eventually approach a non-equilibrium stationary state (NESS) [7–10]. Non-equilibrium systems are commonplace [11], appearing in contexts such as stock markets [12, 13], chemical reactions [14–16], protein folding [17, 18], virus delivery [19, 20] and population dynamics [21–23].

The interaction of a system with the environment can be characterised by observing fluctuations of only the system. In non-equilibrium it is the systems dynamics as opposed to static observations which play a crucial role in studying its behaviour [7, 10]. Analysing the dynamics of a process gives insight into dynamic quantities and phenomena such as self-assembly [24] and dynamical heterogeneity in glasses [25–28].

A large class of these systems can, via a coarse-grained approach, be mapped to continuous-time Markov dynamics over a discrete state space [9]. In the Markovian description an initial probability measure ν_0 of the system evolves according to a Markov semigroup with generator $\mathbf{L}$, such that $\nu_t = \nu_0 e^{t\mathbf{L}}$ is the distribution at time t. Under irreducibility assumptions this converges to a unique and strictly positive NESS $\hat{\pi}$. Outcomes of the process are stochastic trajectories, which are histories of all configurations and transition times. When studying stochastic dynamics it is common to do so using ensemble methods analogous to that in equilibrium. This thermodynamics of trajectories approach [29–32] quantifies the statistics of observables which are time-integrated functionals of trajectories. These *dynamical* observables include currents — which characterise the net flow and degree of reversibility [30, 33–35], and activities — which relate to the relaxation of the system [25, 27]. It is the fluctuations of these dynamical observables about their asymptotic average which are of interest in characterising non-equilibrium. Quantifying the degree of fluctuations has consequences in classifying the dynamics, biological motor efficiency [9, 36, 37], clock per-

formance [38–40] and the estimation of physical quantities [41] (for example, entropy production).

The thermodynamics of trajectories approach formulates stochastic dynamics in the language of large deviations. Dynamical observables obey a large deviation principle (LDP) [42–44]. This quantifies the exponential decay — in the long time limit — of the probability of fluctuations, with rate determined by a *rate function*. The rate function describes the Gaussian fluctuations about their expected value, as well as non-Gaussian fluctuations seen in the tails of these distributions. Rate functions are generally difficult to compute, even when using numerical methods [45–47].

One of the most well-known results in non-equilibrium statistical mechanics are the thermodynamic uncertainty relations (TURs) [34, 37]. These are a class of lower bounds on the relative uncertainty (variance over mean squared) of dynamical observables and correspond to an upper bound on the large deviation rate function. Their significance lies in that the several versions of the TURs are formulated in terms of the quantities which characterise the dynamics, such as entropy production [34], dynamical activity (total number of configuration changes per unit time) [48] or persistence time [38]. General bounds of this type, which are independent of the observable in question, provide a link between the precision of observed quantities and the reversibility or activity of the process.

Dynamical ensembles can alternatively be described in terms of first passage times (FPTs). These are the time it takes an observable to reach a fixed threshold. The statistical properties in this "fixed-observable ensemble" are connected to that of the "fixed-time ensemble" [32, 49] and TURs have been formulated for FPTs [48, 50]. First passage time statistics are widely used [51–56], and in many cases more practical [50].

Some of the notions for classical dynamics can be generalised to the non-commutative framework. Quantum Markov semigroups can describe the evolution of a quantum system interacting with an environment. This evolution is characterised by a GKLS generator $\mathcal{L}_*$ and it is described by the semigroup $\mathcal{T}_{t*} = e^{t\mathcal{L}_*}$, a completely positive trace preserving map acting on quantum states. The state ρ at time 0 is mapped onto the state $\mathcal{T}_{t*}(\rho)$ at time t. In the Markovian description the system undergoes a series of continuous evolutions with intermittent instantaneous jumps. These jumps correspond to a detection in the environment (say a photon) [57–60]. The sequence of emission types and corresponding jump times describe a *quantum trajectory*, for which a trajectory ensemble method and TURs have been established [61–64].

Until recently, the problem of finding bounds has only been considered from one side. An explicit class of bounds instead *limiting* the size of fluctuations has not yet been found. This would restrict the range of the precision of observables, which currently can be arbitrarily large. In the large deviation formalism this type of bound corresponds to a lower bound on the rate function. For quantum Markov processes, in [65, 66] concentration inequalities were found for simple counting measurements. These are bounds on

the whole distribution and correspond to the upper bound on fluctuations we are interested in. It would be useful to obtain results for more general statistics (such as linear combinations of currents or activities) in the classical case, and ideally formulated in terms of simpler quantities. Indeed, in the fixed-observable ensemble, there lack similar bounds for either classical and quantum processes.

In this thesis we tackle the inverse problem of finding upper bounds on fluctuations. Particularly this includes the derivation of general upper bounds on the relative uncertainty of dynamical observables which complement the original formulation of the TURs. We refer to these upper bounds on the uncertainty as "inverse thermodynamic uncertainty relations". They are valid for all currents, activities and any linear combination of fluxes. These bounds are not derived via large deviations and hold for finite time in the stationary state. In addition to bounds on the relative uncertainty we also provide a concentration inequality which upper bounds the whole probability distribution and is valid for any initial state.

We extend these results to first passage times and derive inverse TURs and concentration inequalities for FPTs of counting observables. These bounds we then generalise to the non-commutative setting. Although these results hold for a finite number of jumps k, we also prove a LDP for first passage times in the classical and quantum regimes.

Our results were obtained via well-established methods in concentration theory, namely the Cramér-Chernoff approach. These methods were used in [67–69] to derive concentration inequalities for empirical measures for discrete and continuous-time Markov processes.

Although our investigations are motivated by problems in statistical mechanics, we aim to present our results in a more mathematical framework. The content in Chapter 6 has appeared in [70] whilst the results presented in Chapters 7 and 8 are contained in [71]. These results have been derived in collaboration with Federico Girotti, Madalin Guta and Juan P. Garrahan.

We begin in Chapter 2 by giving a description of classical continuous-time Markov processes, including how trajectories can be separated into a jump trajectory governed by a discrete-time transition matrix, and a collection of holding times (exponentially distributed conditional on the jump process). Then we introduce dynamical observables and first passage times, the quantities of interest in the fixed-time and fixed-observable ensembles respectively.

Dynamical observables and FPTs in the asymptotic regime of large t or large k both satisfy a LDP. We provide an overview of large deviations theory in Chapter 3, working our way from the basic definition, to the i.i.d. case and Cramér's theorem, to weakly dependent processes and Gärtner-Ellis theorem. This leads us to large deviations in the context of dynamical observables and lower bounds on the fluctuations of these observables in the form of the celebrated TURs.

In the finite regime, concentration inequalities can be a useful tool in bounding fluctuations. In Chapter 4 we state basic notions in concentration theory, e.g. the Cramér-Chernoff method, operator perturbation theory and how these are applied to Markov processes.

In Chapter 5 we describe the classical concepts in the more general non-commutative framework. In particular, we outline how we think of trajectories in quantum processes and the statistics of counting measurements.

Chapter 6 focuses on the fixed-time ensemble. We consider trajectories of fixed length t and derive upper bounds on fluctuations of dynamical observables for classical Markov processes in the form of inverse TURs and concentration inequalities. This complements the TURs and are valid for any linear combination of fluxes and for all times.

The latter part of this thesis considers FPT fluctuations in the ensemble conditioned on a fixed number of jumps k. This can either be the total dynamical activity, or the activity within a *subset* of some allowed transitions. In Chapter 7, for classical processes, we prove the LDP holds for FPTs. The main results of this chapter are concentration inequalities and inverse TURs for FPTs. Building on these classical results, Chapter 8 generalises the analysis to the non-commutative setting, including a LDP and a concentration inequality for FPTs of counting measurements in general quantum Markov processes. We also provide similar bounds for quantum reset processes, which are a type of semi-Markov process. Owing to the fact they can be thought of as a classical process, just with a different distribution for holding times, our results for quantum reset processes require fewer assumptions on the dynamics.

2

CLASSICAL MARKOV PROCESSES

This chapter concerns the classical theory of one of the most well-known classes of stochastic process: Markov processes. Informally speaking, they refer to so called "memoryless" dynamics in that the history of the process has no effect on the future distribution. There are two distinct types of Markov process: discrete-time and continuous-time. We begin by introducing discrete-time, before moving on to continuous-time. The object of interest in this thesis are all continuous-time Markov processes; however, we will see in Section 2.3 that these can be decomposed into a discrete-time jump process and a sequence of conditional holding times. We then outline the mathematical formalism of our approach, in the context of linear algebra. Finally, we discuss how Markov processes are studied in statistical mechanics in the form of dynamical observables and first passage times.

2.1 DISCRETE-TIME

A discrete-time Markov process is a sequence of random variables $X := (X_t)_{t\in\mathbb{N}_0}$ taking values in a finite state space E and satisfying the Markov property:

$$\mathbb{P}_\nu\left(X_{t+1} = x | X_0 = x_0, \ldots, X_t = x_t\right) = \mathbb{P}_\nu\left(X_{t+1} = x | X_t = x_t\right)$$

where $\mathbb{P}_\nu$ denotes the law of the process with initial measure $\nu := \nu_0$. The process jumps from configuration $x \to y$, with transition probability $\mathbf{P}_{xy}$ for $x, y \in E$. The n-step evolution of a probability measure ν_t on E, i.e. $\nu_{t+n} = \mathbf{P}^{(n)}\nu_t$, is given by the Chapman-Kolmogorov equation

$$\mathbf{P}^{(n)}_{xy} = \sum_{i_1,\ldots,i_n,\, i_j\in E} \mathbf{P}_{xi_1}\ldots\mathbf{P}_{i_n y}, \quad n \in \mathbb{N}_0$$

or more compactly $\mathbf{P}^{(n)} = \mathbf{P}^n$. We adopt the convention used in mathematics that the transition matrix $\mathbf{P}$ acts on probability measures from the right. If the Markov process is irreducible, there exists a unique and fully supported stationary measure π which satisfies $\pi\mathbf{P} = \pi$, corresponding to the eigenvalue 1 (with multiplicity 1). Furthermore, if the process is aperiodic, any initial measure will converge to π and all other eigenvalues will have magnitude $|\lambda| < 1$. Discrete-time Markov processes themselves have wide application in statistics, machine learning, random walk theory and economics [67–69, 72, 73]. In the next section we will introduce continuous-time Markov processes, the principal stochastic process we consider in this thesis.

2.2 CONTINUOUS-TIME

Let $X := (X_t)_{t\geq 0}$ be a continuous-time Markov process taking values in a finite configuration space E. For an initial distribution $\nu := \nu_0$ over configurations we denote by $\mathbb{P}_\nu$ the corresponding law of the process X and $\mathbb{E}_\nu$ its integral. In continuous-time instead of transition probabilities it is transition *rates* w_{xy} as the rate of transitions from configuration $x \to y$ which characterise the process (note that the rate $w_{xx} = 0$, $\forall x \in E$). The dynamics evolves according to the Kolmogorov forward equation

$$\frac{d}{dt}\nu_t(x) = \sum_{y\in E} \nu_t(y)w_{yx} - \nu_t(x)w_{xy}$$

for some probability measure ν_t. In matrix form this reads

$$\frac{d}{dt}\nu_t = \nu_t \mathbf{L}.$$

The matrix $\mathbf{L}$ is the stochastic generator of X and can be decomposed as

$$\mathbf{L} = \mathbf{W} - \mathbf{R},$$

where the off-diagonal part, $\mathbf{W} = \sum_{x\neq y} w_{xy} |x\rangle \langle y|$, encodes the rates of jumps and the diagonal part, $\mathbf{R} = \sum_x R_x |x\rangle \langle x|$, the escape rates (with $R_x = \sum_y w_{xy}$ the escape rate from configuration x). The generator $\mathbf{L}$ acts on complex valued functions $f : E \to \mathbb{C}$ via right multiplication, i.e.,

$$f(x) \mapsto (\mathbf{L}f)(x) = \sum_{y\in E} \mathbf{L}_{xy} f(y) = \frac{d}{dt}\mathbb{E}_{\delta_x}[f(X_t)],$$

which is similar to the "Heisenberg picture" in quantum mechanics. The natural norm to consider in this setting is the ∞-norm, i.e.

$$\|f\|_\infty := \max_{x\in E} |f(x)|, \quad \|\mathbf{L}\|_{\infty\to\infty} := \max_{\|f\|_\infty = 1} \|\mathbf{L}f\|_\infty.$$

By duality, $\mathbf{L}$ acts also on complex valued measures on E (which can be identified with their density $\nu : E \to \mathbb{C}$) via left multiplication (cf. Schrödinger picture in quantum mechanics):

$$\nu(x) \mapsto (\mathbf{L}_*\nu)(x) := (\nu\mathbf{L})(x) = \sum_{y\in E} \nu(y)\mathbf{L}_{yx} = \frac{d}{dt}\mathbb{P}_\nu(X_t = x).$$

Here the natural norm is the dual norm with respect to the ∞-norm, which is the 1-norm:

$$\|\nu\|_1 := \sum_{x\in E} |\nu(x)|, \quad \|\mathbf{L}_*\|_{1\to 1} := \max_{\|\nu\|_1 = 1} \|\mathbf{L}_*\nu\|_1 = \|\mathbf{L}\|_{\infty\to\infty}.$$

We will often use the notation $\langle \nu, f\rangle$ to denote the integral of f with respect to ν, i.e. $\sum_{x\in E} \nu(x)f(x)$. Next we state our main assumption for classical dynamics.

Hypothesis 2.1 (Irreducibility of **L**). *There exists a unique fully supported measure $\hat{\pi}$ satisfying $\hat{\pi}\mathbf{L} = 0$.*

This unique measure we refer to as the stationary distribution or stationary measure of the classical continuous-time generator **L**. Irreducibility is equivalent to the statement that there is only one communicating class of the Markov process. In simpler terms, in an irreducible process any state can be reached in finite time from any other state.

2.3 JUMPS AND HOLDING TIMES

The process X can be equivalently described in terms of the corresponding discrete-time jump process and holding times in each configuration: indeed any trajectory with k jumps takes the form of a sequence

$$\omega = \{(x_0, t_0), (x_1, t_1), (x_2, t_2), \cdots, (x_k, t_k), (t_{k+1})\},$$

where x_i is the state of the system after the i-th jump and t_i is the time between the $(i-1)$-th and the i-th jump (we set $t_0 = 0$). In other words, the system is in state x_i at time t' with $\sum_{l=1}^{i} t_i \leq t' < \sum_{l=1}^{i+1} t_i$. The final holding time t_{k+1} refers to the time between arriving in the final state x_k, and the finish time t, so that $\sum_{i=1}^{k} t_i \leq t$. For the purpose of these works we can omit this final holding time when writing out a trajectory, in fact when we consider first passage times at the end of this chapter $t_{k+1} = 0$. The process describing the different states of the system along time (*jump process*) is a discrete time Markov process with transition matrix given by

$$\mathbf{P} = \mathbf{R}^{-1}\mathbf{W} \tag{2.1}$$

(notice that due to irreducibility, $R_x > 0$ for every $x \in E$). If **L** is irreducible, then **P** is irreducible too. Indeed

$$\hat{\pi}\mathbf{L} = 0 \Leftrightarrow \hat{\pi}\mathbf{W} = \hat{\pi}\mathbf{R} \Leftrightarrow \hat{\pi}\mathbf{R}\mathbf{P} = \hat{\pi}\mathbf{R}$$

and $\hat{\pi} \mapsto \hat{\pi}\mathbf{R}$ is a positive linear bijection. Therefore the unique invariant measure of **P** (denoted π) is related to the invariant measure of the continuous-time generator by

$$\pi = \frac{\hat{\pi}\mathbf{R}}{\langle \hat{\pi}, \mathbf{R}\underline{1} \rangle} \tag{2.2}$$

where $\underline{1}$ stands for the function identically equal to 1. Irreducibility of the dynamics means that π has full support. Conditional to the jump process, the holding times t_i are independent and $t_i \sim \mathrm{Exp}(R_{x_{i-1}})$ for $1 \leq i \leq k$.

2.4 HILBERT SPACE

We introduce here some Hilbert space notions which will be needed in formulating our classical results and will be used in their proofs. The space

of complex functions on E can be turned into a Hilbert space $L^2_\pi(E)$ using the inner product $\langle \cdot, \cdot \rangle_\pi$ with respect to the invariant measure π defined in Equation (2.2)

$$\langle f, g \rangle_\pi := \sum_{x \in E} \pi(x) \bar{f}(x) g(x).$$

We use the notation $\|f\|_\pi$ for the corresponding norm. In Chapter 6 we make use of $L^2_{\hat{\pi}}(E)$, a similar Hilbert space but instead with respect to the invariant state $\hat{\pi}$ of $\mathbf{L}$. The adjoint $\mathbf{A}^\dagger$ of an operator $\mathbf{A}$ on $L^2_\pi(E)$ has matrix elements

$$\mathbf{A}^\dagger_{xy} := \frac{\pi(y)}{\pi(x)} \mathbf{A}_{yx}$$

with the same expression replaced with $\hat{\pi}$ for the adjoint of an operator $\mathbf{A}$ on $L^2_{\hat{\pi}}(E)$. From the above it follows that $\mathbf{P}^\dagger$ is a transition operator in its own right. An important quantity in this work is the *absolute spectral gap* of $\mathbf{P}$, which we denote by ε and is defined as the spectral gap of $\mathbf{P}^\dagger \mathbf{P}$ (the multiplicative symmetrisation of $\mathbf{P}$)

$$\varepsilon := 1 - \max\{\|\mathbf{P} f\|_\pi : \langle \pi, f \rangle = 0,\ \|f\|_\pi = 1\}. \tag{2.3}$$

We note that the absolute spectral gap is different to the *additive* spectral gap used in Chapter 6.

Using the above we introduce the *León-Perron operator* $\hat{\mathbf{P}}$ associated to $\mathbf{P}$ ([69, Definition 3]).

Definition 2.2 (León-Perron operator)**.** *A Markov operator $\hat{\mathbf{P}}_\lambda$ on $L^2_\pi(E)$ is said to be León-Perron if it is a convex combination of operators $\mathbf{1}$ and Π_π with some coefficient $\lambda \in [0, 1]$, that is*

$$\hat{\mathbf{P}}_\lambda := \lambda \mathbf{1} + (1 - \lambda) \Pi_\pi,$$

where Π_π is the map $\Pi_\pi : f \mapsto \langle f, \underline{1} \rangle_\pi \underline{1}$.

$\hat{\mathbf{P}}_\lambda$ is a self-adjoint transition operator which is simple to handle and will allow us to derive upper bounds for the fluctuations of first passage times of $\mathbf{P}$. In the proofs where we make use of the León-Perron operator, we set $\lambda = 1 - \varepsilon$, and denote as

$$\hat{\mathbf{P}} := (1 - \varepsilon) \mathbf{1} + \varepsilon \Pi_\pi. \tag{2.4}$$

2.5 DYNAMICAL OBSERVABLES

Consider the stochastic process $K_t(x, y)$, the time-integrated process which counts the number of transitions $x \to y$ up to time t. A second quantity defined on the process is $M_t \in \mathbb{R}^d_{\geq 0}$, the *total residence time*. Each element $M_t(x)$ is the total time spent in state x until time t. Linear combinations of counting processes and residence times are called dynamical observables which we define formally.

Definition 2.3 (Dynamical observable)**.** *A dynamical observable (also known in literature as trajectory observable) has the general form*

$$A(t) = \sum_{x \neq y \in E} a_{xy} K_t(x,y) + \sum_{x \in E} b_x M_t(x), \quad a_{xy}, b_x \in \mathbb{R}.$$

The a_{xy} are arbitrary transition weights with $\sum |a_{xy}| > 0$, whilst the b_x are weights associated with the time between transitions. In statistical mechanics, $K_t(x,y)$ is referred to as the *flux* of the transition $x \to y$ and the observable $A(t)$ the *generalised* flux. The scaled version of M_t, denoted $m = M_t/t \in M_1(E)$ is the *empirical measure* belonging to the space $M_1(E)$ of probability measures on E. We are interested in functions of the trajectory of the above form with $b_x = 0$, $\forall x \in E$, so dynamical observables of only the jumps. Observables of the form in Definition 2.3 with $b_x = 0$ satisfy the following [44]:

- strong law of large numbers (SLLN)

$$\frac{1}{t} A(t) \xrightarrow[t \to +\infty]{} \langle a \rangle_{\hat{\pi}} := \sum_{x \neq y} \hat{\pi}(x) w_{xy} a_{xy} \quad \text{a.s.}$$

  where $\langle a \rangle_{\hat{\pi}}$ is the *asymptotic average* of $A(t)$. The convergence is almost sure convergence (denoted a.s.). This means that

$$\mathbb{P}_\nu \left(\lim_{t \to +\infty} \frac{A(t)}{t} = \langle a \rangle_{\hat{\pi}} \right) = 1.$$

- central limit theorem (CLT)

$$\frac{1}{\sqrt{t}} (A(t) - \langle a \rangle_{\hat{\pi}} t) \xrightarrow[t \to +\infty]{} \mathcal{N}(0, V_{\infty,A}) \quad \text{in law.}$$

  In other words, the probability distribution of $\frac{1}{\sqrt{t}}(A(t) - \langle a \rangle_{\hat{\pi}} t)$ converges to a normal distribution with mean 0 and variance $V_{\infty,A}$.

The *asymptotic variance* $V_{\infty,A}$ is

$$V_{\infty,A} := \lim_{t \to +\infty} \frac{V_{\nu,A}(t)}{t}$$

and $V_{\nu,A}(t)$ is the variance of $A(t)$ if ν is the initial distribution (notice, however, that the limit does not depend on ν). When we talk about fluctuations of $A(t)$, we mean a deviation of the process $A(t)/t$ from $\langle a \rangle_{\hat{\pi}}$. There are a few common observables depending on the properties of the a_{xy}, which we will discuss in this section.

2.5.1 *Counting observables*

In practical applications one may wish to count only certain jumps of the trajectory. More formally, given a non-empty subset $\mathfrak{A}$ of the set of possible jumps $\mathfrak{E} := \{(x,y) : x, y \in E : w_{xy} > 0\}$, we denote

$$K_{\mathfrak{A},t} := \sum_{(x,y) \in \mathfrak{A}} K_t(x,y) \tag{2.5}$$

the stochastic process that counts the number of jumps in $\mathfrak{A}$ up to time t, i.e. the observable where $a_{xy} = 1$, $\forall (x, y) \in \mathfrak{A}$. Observables of this form are known as *counting observables*. The *dynamical activity*, or total number of jumps, is the observable corresponding to $\mathfrak{A} = \mathfrak{E}$. Counting observables, in particular the dynamical activity, characterise the system's liveliness. An example of a counting observable could be the number of steps a molecular motor takes in a certain region [36]. Dynamical activity is strongly linked to the relaxation of a system [11, 31, 48] and is used in the study of kinetically constrained models to explain the glass transition [25–28]. The average dynamical activity is denoted $\langle k \rangle_{\hat{\pi}}$ and is defined as follows

$$\langle k \rangle_{\hat{\pi}} := \sum_{x \neq y \in E} \hat{\pi}(x) w_{xy} = \sum_{x \in E} \hat{\pi}(x) R_x. \tag{2.6}$$

We will see in Section 7.5.2 that activity can be used to study models with metastable phases, and how these can give rise to large fluctuations.

2.5.2 *Currents*

Currents are another widely used class of observables, that are antisymmetric under time reversal, i.e. $a_{xy} = -a_{yx}$. We denote generalised currents by

$$J_t := \sum_{x<y} J_t(x, y). \tag{2.7}$$

The individual currents $J_t(x, y) = a_{xy}(K_t(x, y) - K_t(y, x))$ describe the net accumulation of the observable from the contributing (x, y) pair. Typically currents are used to describe physical systems interacting with some environment [74] which drives transitions between states. One example is the electrical current in a quantum dot [75] where the weights correspond to the electron charge. Currents are also used in biochemistry, particularly in the description of cell energy metabolism [76]. In contrast with counting observables which relate to the relaxation of stochastic systems, currents characterise the reversibility [30, 33]. This link between driving transitions and reversibility takes the form of the thermodynamic uncertainty relations (TURs) which we will discuss in Chapter 3.

We demonstrate the behaviour of dynamical observables by considering the following process. Consider a four-site random walk with periodic boundary conditions. The system jumps clockwise with rate γ and anticlockwise with rate ω. Define the current such that $a_{xy} = 1$ for clockwise transitions and $a_{xy} = -1$ for anticlockwise transitions, this type of current is often known as the *particle current*. The particle current will converge to $\langle a \rangle_{\hat{\pi}} = \sum_{x<y} (\hat{\pi}(x)\gamma - \hat{\pi}(y)\omega)$. In the cases where $\gamma = \omega$ the system is an unbiased random walk and the net average accumulation will be 0. Figure 2.1 shows the particle current and for comparison the dynamical activity, which is non-decreasing and has a positive asymptotic average (cf. Equation (2.6)).

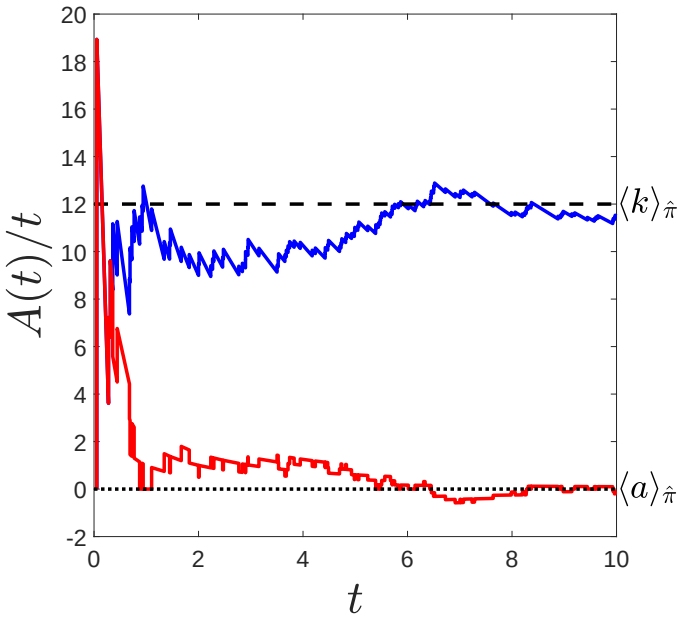

Figure 2.1: Trajectories of an unbiased random walk in a four-state system: in the case where $\gamma = \omega = 6$ the dynamical activity (full/blue) converges to its asymptotic average (dashed/black), whilst the particle current (full/red) converges to its average value of zero (dotted/black).

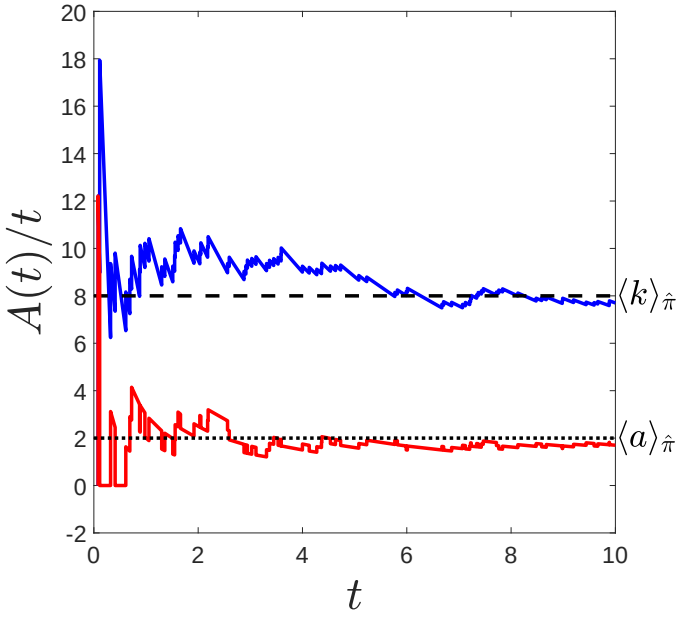

Figure 2.2: Trajectories of a biased random walk in a four-state system: in the case where $\gamma = 6 > \omega = 2$ the dynamical activity (full/blue) and the particle current (full/red) converge to their respective non-zero averages (dashed/black) and (dotted/black).

Note also that contrary to counting observables, currents can accumulate negative values. To observe a non-zero current, we introduce a bias such that there is a tendency for the system to favour clockwise jumps, i.e. $\gamma > \omega$. As we see in Figure 2.2 in addition to the dynamical activity having a positive average, so does the particle current, which reflects the bias in the system.

2.6 FIRST PASSAGE TIMES

So far we have only focused on the fixed-time ensemble, that is we consider the process of accumulating some observable as trajectories of length t evolve. There is another. Instead we can consider trajectories with a fixed threshold value of the observable in question and study the process of the time taken to first pass this threshold. This "first passage time" fluctuates about some average and has much of the same behaviour as the fixed-time ensemble. In fact, the fixed-time and "fixed-observable" ensembles are equivalent [32, 49] in the sense that the statistics of one ensemble can be obtained from the other. Naturally, given a choice of multiple approaches it is logical to choose that most suitable to the problem in hand. In many cases it is more practical to time some process until the threshold is reached than the other way round (computing some observable level for trajectories of fixed t). One example is the enzyme-mediated reaction described in [50] where measuring the first passage time (FPT) for the reaction to reach some desired fluorescence bypasses a more complicated scheme of measuring reaction product concentration levels for a range of trajectories. First passage problems feature in other fields such as economics [51] — computing the likelihood of an investment being lost; search processes [52, 77]; epidemiology [53, 78]; and in various physical systems [54–56, 79]. It is therefore important to study first passage times in the same way we do in the fixed-time ensemble. The study of FPTs is more involved than of time-integrated observables due to the fact that for more complicated observables like currents, each transition can contribute negatively to the accumulated value. For that reason we consider FPTs for the subset of all trajectory functions known as counting observables (which include fundamental quantities such as the dynamical activity), defined in Equation (2.5).

Definition 2.4 (First passage time (counting observable))**.** *The first passage time (FPT), denoted $T_{\mathfrak{A}}(k)$ for a counting observable $K_{\mathfrak{A},t}$ corresponding to the value $k \in \mathbb{N}$ is defined as*

$$T_{\mathfrak{A}}(k) := \inf_{t \geq 0} \{t : K_{\mathfrak{A},t} = k\}.$$

In particular, the first passage time for the total activity corresponding to the level k is given by the sum of the first k holding times

$$T_{\mathfrak{E}}(k) = \sum_{i=1}^{k} t_i.$$

We have seen some of the notions used when working with classical Markov processes in both discrete-time and continuous-time. Furthermore, we have defined the stochastic quantities of interest in this thesis: dynamical observables and first passage times. In the next chapter we will describe how one would begin to study these using large deviations theory.

3

LARGE DEVIATIONS

Anyone who has entered The National Lottery or who lives in proximity to a fault line may be concerned with the probability of rare events. Fortunately there is an entire branch of mathematics, *large deviations theory*, dedicated to addressing these concerns. These tools allow us to make statements about how the probability of events which are atypical decays with respect to some parameter (e.g. system size). This is particularly useful when — by definition of events being "rare" — sample sizes of the events may be very small and usual statistical techniques such as Monte Carlo methods, maximum likelihood estimation and Bayesian inference are no longer effective. The more familiar strong law of large numbers (SLLN) gives us the limiting statement for partial sums $S_n = X_1 + \cdots + X_n$ of i.i.d. random variables with mean μ and variance σ^2

$$\frac{1}{n} S_n \xrightarrow[n\to+\infty]{} \mu \quad \text{a.s.}$$

Let Z be the standard normal distribution. The central limit theorem (CLT)

$$\frac{1}{\sigma\sqrt{n}}(S_n - \mu n) \xrightarrow[n\to+\infty]{} Z \quad \text{in law}$$

captures the fact that the empirical average concentrates around μ with standard deviation of order $1/\sqrt{n}$. Large deviations are any event which differs from this average by *at least* an amount of order n. The SLLN guarantees that the probability of this deviation will vanish and large deviations theory provides us with the rate at which this vanishing occurs. This rate is given to us in the form of the large deviation rate function, which describes both rare and typical statistics of the distribution. Large deviations theory is applicable to other objects than the empirical average, such as the empirical and pair empirical measures, and goes beyond i.i.d. with results for Markov processes and other dependent sequences.

There are many books, reviews and other works comprehensively formalising large deviations theory. This chapter aims to present the main ideas used in this thesis. Mathematically formalised over the course of the 20$^{\text{th}}$ century, large deviations theory constitutes the large deviation principle (LDP) followed by a plethora of results calculating rate functions and proving LDPs in specific situations. We proceed in a similar fashion. We begin in Section 3.1 with the definition of the LDP. In Section 3.2 we state one of the original results, Cramér's theorem, for i.i.d. random variables and demonstrate with an example. Cramér's theorem can be generalised to weakly dependent sequences in $\mathbb{R}^d$ via the Gärtner-Ellis theorem (Section 3.3).

Large deviations is the mathematical basis of much of statistical mechanics [44] in both equilibrium and non-equilibrium physics. In Section 3.4 we go down the non-equilibrium route where the tools and methods used to compute rate functions in the case of dynamical observables of Markov processes is stated. We then finish by introducing the thermodynamic uncertainty relations (TURs) which are key results in this field and were first proven using large deviations.

3.1 LARGE DEVIATION PRINCIPLE

The large deviation principle uses upper and lower bounds to characterise the exponential decay of a sequence of probability measures with respect to their index n. We use the definition in [42, p. 5].

Definition 3.1 (Large Deviation Principle)**.** *Let $\mathcal{X}$ be a Polish space and let $I : \mathcal{X} \to [0, \infty]$ be a lower semicontinuous mapping. A sequence of probability measures (μ_n) on $\mathcal{X}$ satisfies a large deviation principle (LDP) with rate function I if for any Borel measurable set $\Gamma \subseteq \mathcal{X}$*

$$-\inf_{x \in \overset{\circ}{\Gamma}} I(x) \leq \liminf_{n \to +\infty} \frac{1}{n} \log \mu_n(\Gamma) \leq \limsup_{n \to +\infty} \frac{1}{n} \log \mu_n(\Gamma) \leq -\inf_{x \in \overline{\Gamma}} I(x),$$

where $\overset{\circ}{\Gamma}$ denotes interior and $\overline{\Gamma}$ the closure of Γ. Furthermore, I is called good if it has compact level sets.

Qualitatively, Definition 3.1 outlines that the decay of the measure μ_n on the event Γ occurs with rate according to the least unlikely outcome in Γ [42].

Remark 3.2. If for some Γ the rate function I is continuous $\forall x \in \Gamma$, the upper and lower bound in the definition coincide and reduce to

$$\lim_{n \to +\infty} \frac{1}{n} \log \mu_n(\Gamma) = -\inf_{x \in \Gamma} I(x).$$

Note that this is only in general the case for such continuity sets of I.

Remark 3.3. The parameter n need not be a member of $\mathbb{N}$, the sequence μ_n can be replaced with a a probability measure μ_t, indexed with a continuous-time parameter $t \geq 0$ and the limits replaced by a limit with respect to t.

Remark 3.4. Rate functions are unique, and a good rate function implies that there is at least one point $x \in \mathcal{X}$ such that $I(x) = 0$. Furthermore, if the rate function is strictly convex, the zero of the rate function is also unique. Throughout this thesis, and in many other cases this zero corresponds to the expected value obtained from the SLLN. However, this is not always the case; we refer to [44, 80] for details.

Definition 3.1 is rarely used directly; generally, large deviation statements rely on other results which tell us whether a LDP holds and the form of the corresponding rate function. The first of these we discuss is Cramér's theorem.

3.2 CRAMÉR'S THEOREM

Originally derived in [81], we state the formulation in den Hollander [43, Theorem I.4].

Theorem 3.5 (Cramér). *Let (X_i) be i.i.d. random variables taking values in $\mathbb{R}$, with law $\mathbb{P}$ and expectation $\mathbb{E}$ with respect to $\mathbb{P}$ satisfying*

$$\mathbb{E}[e^{uX_1}] < \infty, \quad \forall u \in \mathbb{R}.$$

Let $\mathbb{P}_n$ denote the law of $S_n = \sum_{i=1}^n X_i$. Then the sequence $(\mathbb{P}_n)$ satisfies the LDP and for all $a > \mathbb{E}[X_1]$

$$\lim_{n\to+\infty} \frac{1}{n} \log \mathbb{P}_n \left(S_n \geq an\right) = -I(a)$$

where

$$I(x) = \sup_{u\in\mathbb{R}} \left\{ xu - \log \mathbb{E}[e^{uX_1}] \right\}.$$

Consider the following example. Take (X_i) to be i.i.d. Bernoulli random variables with law $\mathbb{P}(X_1 = 0) = p$ and $\mathbb{P}(X_1 = 1) = 1 - p$. Let $S_n = \sum_{i=1}^n X_i$. The moment generating function of X_1 is

$$\mathbb{E}[e^{uX_1}] = p + (1-p)e^u$$

which is finite $\forall u \in \mathbb{R}$. The SLLN ensures that the sample mean $\frac{S_n}{n}$ converges to $\mathbb{E}[X_1] = 1 - p$. Cramér's theorem then tells us that

$$\lim_{n\to+\infty} \log \mathbb{P}_n \left(\frac{S_n}{n} \geq a \right) = -I(a)$$

where the rate function $I(x)$ is

$$I_p(x) = \begin{cases} x \log px + (1-x)\log(1-p)(1-x) - \log p(1-p) & \text{if } x \in [0,1] \\ +\infty & \text{otherwise.} \end{cases}$$

A real-world application is a bag with 100 coloured marbles, 70 of which are red and 30 blue. Marbles are drawn at random and replaced. We seek the probability after 100 draws of at least 60% being blue. We have $\mathbb{E}[X_1] = 0.3$ so after 100 draws we expect this probability to be very small. The above rate function gives

$$\mathbb{P}_n \left(\frac{S_{100}}{100} \geq 0.6 \right) \approx e^{-100 I_{0.7}(0.6)}. \tag{3.1}$$

This gives a probability of approximately 4.6×10^{-9} of drawing at least 60% blue. As expected, this event is extraordinarily rare. It would not be practical to calculate this via Monte Carlo sampling and the power of large deviations lies in being able to quantify these events despite them being extremely unlikely.

Figure 3.1 shows the whole rate function for the marble setup. Note that in the region outside the dashed lines the rate function exists and is equal to ∞. Recall that Cramér's theorem is valid for $a > \mathbb{E}[X_1]$ and gives us the rate function for $\mathbb{P}_n(S_n \geq an)$. The left deviations, i.e. $\mathbb{P}_n(S_n \leq an)$ in the figure are also calculated from Equation (3.1) because of the symmetry of the problem, since $I_p(x) = I_{1-p}(1-x)$. The "typical" behaviour of S_n corresponds to the minimum of I at $x = 0.3$.

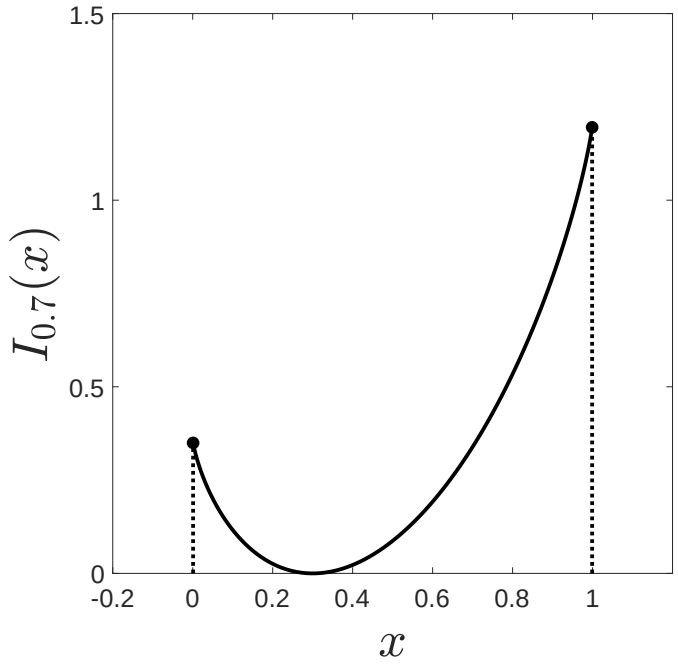

Figure 3.1: Rate function for sum of i.i.d. Bernoulli random variables

3.3 GÄRTNER-ELLIS THEOREM

If we wish to apply large deviation results to Markov processes we require more powerful results than for i.i.d. sequences. The Gärtner-Ellis theorem [82, 83] extends Cramér's theorem to dependent sequences requiring some assumptions on the limiting properties of the moment generating function. Denote by $\mathcal{D}_\Lambda$ the set $\mathcal{D}_\Lambda = \{u \in \mathbb{R}^d : \Lambda(u) < \infty\}$ and its interior by $\overset{\circ}{\mathcal{D}}_\Lambda$. To state the theorem we first require the following [42, Definition 2.3]:

Definition 3.6 (Essentially smooth convex function)**.** *A convex function* $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ *is essentially smooth if*

a) $\overset{\circ}{\mathcal{D}}_\Lambda$ *is non-empty*

b) $\Lambda(\cdot)$ *is differentiable throughout* $\overset{\circ}{\mathcal{D}}_\Lambda$

c) $\Lambda(\cdot)$ *is steep, namely* $\lim_{n\to+\infty} |\nabla\Lambda(u_n)|$ *wherever* $(u_n)_{n\in\mathbb{N}}$ *is a sequence in* $\overset{\circ}{\mathcal{D}}_\Lambda$ *converging to a boundary point in* $\overset{\circ}{\mathcal{D}}_\Lambda$.

Using Definition 3.6 we state [42, Theorem 2.3.6].

Theorem 3.7 (Gärtner-Ellis). *Let $Z_n \in \mathbb{R}^d$ be a sequence of random vectors with probability law μ_n. Assume the limiting logarithmic moment generating function*

$$\Lambda(u) = \lim_{n\to+\infty} \frac{1}{n} \log \mathbb{E}\left[e^{\langle u, Z_n\rangle}\right]$$

exists and whose origin is contained in $\overset{\circ}{\mathcal{D}}_\Lambda$. Furthermore, if Λ is an essentially smooth, lower semicontinuous function, then μ_n satisfies a LDP with good rate function

$$I(z) = \sup_{u\in\mathbb{R}^d} \{\langle u, z\rangle - \Lambda(u)\}.$$

Theorem 3.7 allows us to prove a LDP for dependent sequences from properties of the limiting logarithmic moment generating function, which we do in Theorems 7.4 and 8.5. Furthermore it does not demand that the limiting logarithmic moment generating function is finite everywhere, just that it converges in a neighbourhood of the origin. We also note that there are cases where the Gärtner-Ellis theorem does not hold, but a large deviation principle does; however, such cases are not considered here.

3.4 RATE FUNCTIONS FOR DYNAMICAL OBSERVABLES

There are two ways of studying the large deviations of observables on Markov processes. The first, a more dynamical approach, is the "level 2.5" of large deviations [44, 84]. The other is closer to the thermodynamic formalism for equilibrium physics [26, 31] and uses a "tilted" (also known as perturbed) generator. Recall from Section 3.4 a dynamical observable has the form

$$A(t) = \sum_{x\neq y} a_{xy} K_t(x, y)$$

where $K_t(x, y)$ counts the number of $x \to y$ transitions up to time t, and each transition contributes weight $a_{xy} \in \mathbb{R}$. We seek to calculate the rate function of the time-averaged quantity $a = A(t)/t$.

3.4.1 *via level 2.5 of large deviations*

The next theorem allows us to derive the LDP from transformations of the state space via a *contraction* [42, Theorem 4.2.1].

Theorem 3.8 (Contraction Principle). *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces and $f : \mathcal{X} \to \mathcal{Y}$ a continuous function. Consider a good rate function $I : \mathcal{X} \to [0, \infty]$. For each $y \in \mathcal{Y}$ define*

$$I'(y) = \inf\{I(x) : x \in \mathcal{X}, y = f(x)\}$$

where the infimum of the empty set is taken to be ∞.

1. *Then I' is a good rate function on $\mathcal{Y}$.*
2. *If the family of probability measures $(\mu_n)_{n\in\mathbb{N}}$ on $\mathcal{X}$ satisfy a LDP with rate function I, then the family of probability measures $(\mu_n \circ f^{-1})$ on $\mathcal{Y}$ satisfy a LDP with rate function I'.*

Theorem 3.8 can be used to compute the rate function for the observables in Section 2.5. Large deviations in the context of Markov processes can be categorised into three "levels" [44, 85]. Level 3 is that of trajectories ω, with rate function $I(\omega)$, the LDP of which was proven in [86–89]. Level 2 is that of the empirical measure $m = M_t/t$, and can be obtained via the contraction

$$I(m) = \inf_{\omega}\{I(\omega) : m_\omega = m\}$$

where m_ω denotes empirical measure of the trajectory ω. Level 1 is the level of time-extensive observables. To study observables which are functions of fluxes, level 2 is not sufficient [90]. There is an intermediary level between 2 and 3 known as "level 2.5" which has an explicit form and contains enough information to contract down to functions of jumps of the process. Define the *empirical rate* k be

$$k(x,y) = \frac{K_t(x,y)}{M_t(x)}.$$

The joint rate function for m and k is [84, 91]

$$I(m,k) = \sum_{x,y\in E} m(x)\left[k(x,y)\log\left(\frac{k(x,y)}{w_{xy}}\right) - k(x,y) + w_{xy}\right].$$

The tuple (m,k) belong to the space

$$\mathbb{B} = \left\{ M_1(E) \times \mathbb{R}_{\geq 0}^{|E|(|E|-1)} : \sum_{y\in E}\left(m(y)k(y,x) - m(x)k(x,y)\right)\right\}$$

where $\sum_{y\in E}\left(m(y)k(y,x) - m(x)k(x,y)\right)$ is the stationary condition, which ensures m,k are such that there is no net flow to any particular state. Once we have the level 2.5 rate function, we can contract to level 1. Using Theorem 3.8, the rate function of a dynamical observable of the form in Definition 2.3 is obtained variationally

$$I(a) = \inf\left\{ I(m,k) : (m,k)\in\mathbb{B},\, a = \sum_{x,y\in E} a_{xy}m(x)k(x,y)\right\}.$$

3.4.2 *via tilted generator*

This is the typical method for studying large deviations of dynamical systems. The moment generating function of a dynamical observable is of the form [31, 44, 48]

$$\mathbb{E}_\nu\left[e^{uA(t)}\right] = \langle \nu, e^{t\mathbf{L}_u}\underline{1}\rangle \tag{3.2}$$

where $\mathbf{L}_u$ is the *tilted generator*

$$\mathbf{L}_u = \sum_{x \neq y} (e^{ua_{xy}} - 1) w_{xy} |x\rangle \langle y| + \mathbf{L}. \tag{3.3}$$

The $e^{ua_{xy}}$ terms introduces a bias on the jumps which contribute to the observables. For example, if the observable of interest is the dynamical activity (the case where $a_{xy} = 1, \forall x, y \in E$), the generator reads

$$\mathbf{L}_{u,\mathfrak{E}} = e^u \mathbf{W} - \mathbf{R}$$

where we recall that the set of all possible jumps is denoted $\mathfrak{E}$. Using Perron-Frobenius theorem (cf. Theorem 5.5), the limiting logarithmic moment generating function for a general observable is then

$$\Lambda(u) = \max \{ \Re(z) : z \in \mathrm{Sp}(\mathbf{L}_u) \}.$$

The rate function for the dynamical observable can then be found by performing the Fenchel-Legendre transform (Theorem 3.7).

As we can see, there are many theorems involved in the calculation of rate functions; but, the problem for our situation still reduces to either an optimisation or an eigenvalue problem. The tilted generator can only be exactly diagonalised for small system sizes; however, in many-body physics there are numerical techniques designed to find extremal eigenvalues of such objects, such as density matrix renormalisation group (DMRG) [45, 47, 92] and projected entangled pair states (PEPS) [46, 93]. These techniques can reveal phase transitions for different stochastic models [7, 28].

Although we have considered dynamical observables here, large deviations can be applied to other functionals as well, in particular, to first passage times (FPTs). The rate function of a FPT for a given observable can be computed from the rate function $I(a)$ of the observable itself, and transformed via an equivalence of ensembles relation [48–50]. The large deviations of first passage times is the subject of chapters 7 and 8 where we prove a large deviation principle for FPTs of counting observables (cf. Equation (2.5)) and provide an explicit form of the rate function without having to compute that of the observable. Naturally, this approach still has the same scaling drawbacks as for dynamical observables.

Since it is rarely possible to find an analytical expression for a rate function — at least one which is not in some variational form — one may wish to find bounds which provide an analytical expression whilst still capturing the physics of the system. In the next section we will discuss *upper* bounds on rate functions and their consequences.

3.5 THERMODYNAMIC UNCERTAINTY RELATIONS

Dynamical Observables of the form in Definition 2.3 obey a large deviation principle with a continuous rate function. Definition 3.1 shows that the probability of rare fluctuations decays exponentially with time t at rate $I(a)$.

An upper bound $\bar{I}$ such that $\forall a, \bar{I}(a) \geq I(a)$ implies that the probability of fluctuations decays slower than that given by $\bar{I}(a)$.

$$\lim_{n\to+\infty} -\frac{1}{n} \log \mathbb{P}_{\hat{\pi}} \left(\frac{A(t)}{t} \geq a \right) = I(a) \geq \bar{I}(a).$$

Let $\bar{I}(a)$ have some of the properties that the rate function $I(a)$ has, i.e. that they share the same root at $\langle a \rangle_{\hat{\pi}}$ and have a first derivative of 0 (which is the case for rate functions when they are obtained via Gärtner-Ellis [44]). The upper bound $\bar{I}$ then gives a lower bound on the variance of the observable. This is more often expressed as relative error or relative uncertainty — the variance divided by the mean squared. We define the relative error ϵ_A^2 of A as the ratio between the variance of A and its average squared multiplied by t:

$$\epsilon_A^2 := t \frac{V_{\hat{\pi},A}(t)}{\langle A \rangle_{\hat{\pi}}^2} = \frac{V_{\hat{\pi},A}(t)}{t \langle a \rangle_{\hat{\pi}}^2}. \tag{3.4}$$

Postulated in the context of biomolecular processes [37] and proven via level 2.5 [34], the TUR provides a lower bound on the relative error

$$\frac{2}{\Sigma_{\hat{\pi}}} \leq \epsilon_J^2$$

where $\Sigma_{\hat{\pi}} = \sum_{x \neq y} \hat{\pi}(x) w_{xy} \log(\hat{\pi}(x) w_{xy} / \hat{\pi}(y) w_{yx})$ is the average entropy production rate in the non-equilibrium stationary state (NESS) and ϵ_J^2 is the relative error specific to currents.

What is notable about the TUR is that the variance bound (and the rate function bound from which it is derived) are formulated in terms of $\Sigma_{\hat{\pi}}$, which is easier to compute than the methods for calculating rate functions explicitly. Secondly, the relationship between the average entropy production and size of fluctuations has a physical meaning: better precision requires more dissipation. For example, in living cells, the maintenance of the concentration of adenosine triphosphate (ATP) is a non-equilibrium system [76] where ATP is both produced and removed. The TUR states that the entropy production of the system must be large enough to ensure the concentration level stays within the required range. The entropy production of a process can be used to estimate the value of the observable's precision [37]. An alternative setup is that dissipation can be inferred from observations of accessible currents [9]; furthermore, observations of such systems as molecular motors [9, 36] can provide bounds on the thermodynamic efficiency of the process.

This formulation in terms of entropy production is a bound only for currents. There are TURs for more general dynamical observables. The activity TUR demonstrates the tradeoff between fluctuations and the average dynamical activity $\langle k \rangle_{\hat{\pi}}$ (cf. Equation (2.6)). Because of this, it is sometimes referred to as the kinetic uncertainty relation (KUR) [48]. The recently discovered persistence-time TUR [38] is expressed in terms of the average persistence time $\langle \tau \rangle_{\hat{\pi}} = \sum_{x \in E} \frac{\hat{\pi}(x)}{R_x}$. This is also known as the clock uncertainty

relation (CUR) and was formulated via the Cramér-Rao bound. It can alternatively be derived by large deviations methods [40]. The CUR was shown to always be at least as tight as the KUR:

$$\frac{1}{\langle k \rangle_{\hat{\pi}}} \leq \langle \tau \rangle_{\hat{\pi}} \leq \epsilon_A^2.$$

Furthermore, the CUR is saturated by the observable $a_{xy} = R_y^{-1}$.

These lower bounds on fluctuations are often collectively referred to as TURs or TUR-type bounds as they are all one-sided lower bounds on the same object — the relative error. There is now a large literature on the TUR [33, 74, 94–99] and they have been extended to finite time [100, 101], discrete time [102] and for quantum Markov processes [61, 63, 64, 103]. The TURs can also be derived for FPTs [48, 50]. Note that the quantities used in the TURs do not depend on the observable in question. Whilst this has an interesting physical and useful meaning, it is not possible to infer solely from the TUR how the observable in question affects fluctuations. Moreover, since they are lower bounds, they do not limit the size of fluctuations, which could be arbitrarily worse, so by themselves are not informative when estimating the size of the relative error. We will provide in Chapter 6 "inverse thermodynamic uncertainty relations" which are upper bounds on the relative error, and are formulated in terms of quantities which do depend on the observable. In chapters 7 and 8 we extend this work to inverse TUR-type bounds for FPTs for classical and quantum Markov processes respectively.

4

CONCENTRATION INEQUALITIES

Concentration inequalities are upper bounds on the probability of a random variable deviating from a certain value (e.g. its average). They offer finite-regime guarantees on deviation probabilities and are typically used in situations with no analytical result available. In these cases, we wish our upper bound to contain simpler quantities than the distribution. These are usually quantities such as the expectation or variance of the random variable itself. Whilst the law of large numbers gives guarantees of convergence for sums of random variables, concentration inequalities quantify the rate of this convergence for finite sample averages. Concentration inequalities have broad application in statistics, machine learning, economics and risk management [68, 69, 104].

In this chapter we will outline some of the common techniques used in literature and this thesis for deriving such bounds. Sections 4.1 and 4.2 are an accelerated view of chapter 2 in [104] which contains many more details, examples and other results. Concentration inequalities have their roots in the Markov inequality, and we will see by imposing various restrictions a range of bounds can be derived. We will build up to the Cramér-Chernoff method (Section 4.1.3) which forms part of all the concentration results in this thesis. By applying the Cramér-Chernoff method to sums of random variables, in Section 4.2 we will state famous results in these cases, notably the Bernstein bound, since the bounds we derive in chapters 6–8 are of Bernstein-type.

Although valid for finite samples, in the large deviation regime concentration inequalities provide lower bounds on the rate function (Section 4.3). In deriving these lower bounds we use operator perturbation theory and give an overview of this in Section 4.4. To highlight the main ideas, we will introduce concentration inequalities for random variables with the assumptions they were originally derived for: independent random variables, but in Section 4.5 we lift this independence assumption, specifically for the case of Markov processes.

We begin by stating that a concentration inequality denotes an upper bound with respect to a random variable Z of the form

$$\mathbb{P}\left(Z - \mathbb{E}[Z] \geq \gamma\right), \quad \forall \gamma > 0,$$

which we refer to as *right deviations*. We assume that the expectation $\mathbb{E}[Z]$ exists. Bounds on *left deviations* are upper bounds on

$$\mathbb{P}\left(\mathbb{E}[Z] - Z \geq \gamma\right), \quad \forall \gamma > 0.$$

Although our later results also provide bounds for left deviations, for simplicity we will consider only right deviations in this chapter.

4.1 ROADMAP TO THE CRAMÉR-CHERNOFF METHOD

4.1.1 *Markov's inequality*

As mentioned above, the first step in understanding concentration theory begins with the Markov inequality. Let $Y \in \mathbb{R}_{\geq 0}$ be a non-negative random variable. Then, $\forall \gamma > 0$, Markov's inequality states that

$$\mathbb{P}(Y \geq \gamma) \leq \frac{\mathbb{E}[Y]}{\gamma}.$$

Note that by applying to $Y = |Z - \mathbb{E}[Z]|$ we obtain our first concentration inequality. However, the upper bound in the above only contains the expectation of Y, so will miss influences from the tails of the distribution. As such, the basic Markov inequality is pessimistic. We can improve the situation by instead considering functions of Y. Let $Y \in X \subseteq \mathbb{R}$ and $\phi : X \to \mathbb{R}_{\geq 0}$ be a non-negative, non-decreasing function, then $\forall \gamma \in X$ with $\phi(\gamma) > 0$

$$\mathbb{P}(Y \geq \gamma) \leq \mathbb{P}(\phi(Y) \geq \phi(\gamma)) \leq \frac{\mathbb{E}[\phi(Y)]}{\phi(\gamma)}.$$

In practice, we then choose a ϕ which is convenient to the problem in hand. for example, if we now take $\phi(\gamma) = \gamma^q$ where $q > 0$ and $Y = |Z - \mathbb{E}[Z]|$, then $\forall \gamma > 0$

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \gamma) \leq \frac{\mathbb{E}\left[|Z - \mathbb{E}[Z]|^q\right]}{\gamma^q} \tag{4.1}$$

and we arrive at the generalised Markov inequality. By considering higher order moments of Y, we can control the bound by choosing an optimal q, if $\mathbb{E}\left[|Z|^q\right] < \infty$, $\forall q \geq 0$. This results in an upper bound containing quantities which better represent the shape of the distribution of Z. We state the specific case of $q = 2$ in the section below.

4.1.2 *Chebyshev's inequality*

If in Equation (4.1) we take $q = 2$ we arrive at the special case of Chebyshev's inequality

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \gamma) \leq \frac{V(Z)}{\gamma^2}.$$

We note Chebyshev's inequality as the bound is in terms of the variance $V(\cdot)$ of the random variable, which is often known or of interest. Indeed, if Z is the sum of independent random variables the variance of Z is just the sum of the individual variances.

4.1.3 *Cramér-Chernoff method*

One can consider different functions instead. Conveniently if we choose ϕ to be of the form $\phi(x) = e^{ux}$, this gives for real valued Z, for all $u \geq 0$

$$\mathbb{P}(Z \geq x) \leq e^{-ux}\mathbb{E}[e^{uZ}] = \exp\left(-(ux - \psi_Z(u))\right) \tag{4.2}$$

where $\psi_Z(u) = \log \mathbb{E}[e^{uZ}]$. Let

$$\psi_Z^*(x) = \sup_{u \geq 0} \{ux - \psi_Z(u)\}, \tag{4.3}$$

then Equation (4.2) can be optimised in u by

$$\mathbb{P}(Z \geq x) \leq \exp(-\psi_Z^*(x)).$$

As is remarked in [104], if $\mathbb{E}[Z]$ exists then the supremum in Equation (4.3) can be taken over all $u \in \mathbb{R}$. The Cramér-Chernoff approach is central to many results in this thesis. In the concentration inequalities we derive, we introduce an extra step: we upper bound the moment generating function (MGF) with another function of exponential form. By doing this we obtain an exponentially decaying upper bound in terms of the parameter u which we then optimise to obtain the final concentration inequality. We detail this specific approach more in the proofs of our results in chapters 6–8.

4.2 SUMS OF INDEPENDENT RANDOM VARIABLES

The above Cramér-Chernoff method was invented for the study of sums of independent random variables. It takes advantage of the product rule for the MGF: if $X_1, \ldots, X_n$ are independent random variables with $\mathbb{E}[X_i] < \infty$ and $\mathbb{E}[e^{uX_i}] < \infty$ for some non-empty interval I, for $i \leq n$ and $u \in I$, then

$$\psi_{S_n}(u) = \sum_{i=1}^{n} \log \mathbb{E}\left[e^{u(X_i - \mathbb{E}[X_i])}\right], \quad \forall u \in I.$$

this allows questions about the sum $S_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])$ to be turned into questions about the individual X_i. We state the first example of these in the next section.

4.2.1 *Hoeffding's inequality*

For the simple assumption that the X_i are bounded, the Cramér-Chernoff method is used to derive Hoeffding's inequality [104, Theorem 2.8].

Theorem 4.1 (Hoeffding's inequality)**.** *Let $X_1, \ldots, X_n$ be independent random variables where X_i takes values in $[a_i, b_i]$ almost surely. Let*

$$S_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])$$

and denote the law of S_n as $\mathbb{P}_n$. Then for every $\gamma > 0$,

$$\mathbb{P}_n(S_n \geq \gamma) \leq \exp\left(-n^2 \frac{2\gamma^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Note that the exponent decays $O(n)$, this bound quantifies how fast deviation probability decays as the sample size increases, with few other assumptions. Whilst widely applicable, the distribution may be concentrated in a region much smaller than $\sum_{i=1}^{n}(b_i - a_i)$. Other concentration inequalities improve on this (for more bounds see [104]). We will look at one of these: Bernstein's inequality.

4.2.2 *Bernstein's inequality*

For simplicity we will consider the i.i.d. and centred case, although Bernstein's inequality is valid for independent and not necessarily identically distributed random variables. This is a generalisation of Hoeffding's inequality to unbounded random variables, but instead requires the moments to be bounded. We state the theorem below ([104, Theorem 2.10]).

Theorem 4.2 (Bernstein's inequality)**.** *Let $X_1, \ldots, X_n$ be i.i.d. random variables with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] \leq b^2$ and $|X_1| \leq c$ almost surely. Furthermore, let*

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

with law $\mathbb{P}_n$. Then for every $\gamma > 0$,

$$\mathbb{P}_n\left(S_n \geq \gamma\right) \leq \exp\left(-n\frac{\gamma^2 b^2}{c^2} h\left(\frac{2\gamma c}{b^2}\right)\right),$$

where $h(x) := (\sqrt{1+x} + \frac{x}{2} + 1)^{-1}$.

Quantities in concentration theory such as b^2 are sometimes referred to in literature as the *variance proxy* [68, 69] in that they contain information on the shape of the distribution, and are used in upper bounding the variance. Throughout these works we will see similar quantities in their respective situations. The function h is a result of the Fenchel-Legendre transform done to optimise the bound in the Cramér-Chernoff method (see Equation (4.2)). In cases where the distribution has a wide maximal range but heavy concentration about its mean, the Bernstein inequality will likely outperform Hoeffding's inequality.

4.3 CONNECTION TO LARGE DEVIATIONS

It should be clear that the exponential upper bound in Theorems 4.1 and 4.2 appears to have a similar form to what one would expect from the definition of the large deviation principle (Definition 3.1). Furthermore, the use of the logarithmic moment generating function in the Fenchel-Legendre transform in Equation (4.3) suggest a strong link between the two. This should not be a surprise since in both large deviations and concentration theory we are talking about the same objects: the deviations of some random process from

its average. The difference is with a large deviation result we have equality only in the limit, whereas concentration inequalities provide a bound in the finite regime. We can however obtain a lower bound on the large deviation rate function directly from concentration inequalities. Consider the deviations of a random variable Z_n which is a function of n random variables and obeys a strong law of large numbers (SLLN) so that $\lim_{n\to+\infty} \frac{Z_n}{n} = \mu$. Denote the law of Z_n by $\mathbb{P}_n$ and its expectation $\mathbb{E}_n$. We consider right deviations although similar arguments can be made for left deviations. In this case we are interested in the values $z = \mu + \gamma$, where $\gamma > 0$. The Chernoff bound for Z_n is, $\forall u \geq 0$

$$\mathbb{P}_n\left(\frac{Z_n}{n} \geq z\right) \leq e^{-nuz}\mathbb{E}_n[e^{uZ_n}]. \tag{4.4}$$

If the sequence of measures on Z_n satisfy the condition needed for Gärtner-Ellis theorem (Theorem 3.7) then a large deviation principle (LDP) holds

$$\lim_{n\to+\infty} \frac{1}{n} \log \mathbb{P}_n\left(\frac{Z_n}{n} \geq z\right) = -I(z) = -\sup_{u\in\mathbb{R}}\{zu - \Lambda(u)\}$$

where $I(z)$ is the rate function of a value z and $\Lambda(u)$ the corresponding limiting logarithmic moment generating function (cf. Theorem 3.7). Taking the same limit of the R.H.S. in (4.4) gives

$$\lim_{n\to+\infty} \frac{1}{n} \log e^{-nuz}\mathbb{E}_n[e^{uZ}] = -\left(zu - \Lambda(u)\right).$$

The Chernoff bound in (4.4) is optimised by taking the following supremum (recall from Section 4.1.3 that the supremum can be taken over all $u \in \mathbb{R}$)

$$-\sup_{u\in\mathbb{R}}\{zu - \Lambda(u)\} \leq -\left(zu - \Lambda(u)\right).$$

The Fenchel-Legendre transform in the above inequality is equal to $I(z)$, the rate function obtained from Gärtner-Ellis.

The above shows that if a large deviation principle holds, then the optimised bound from the Cramér-Chernoff method coincides with the large deviation result. If $\Lambda(u)$ is inaccessible, but we can obtain a function $\tilde{\Lambda}(u)$ such that $\tilde{\Lambda}(u) \geq \Lambda(u)$, $\forall u$, then

$$I(z) = \sup_{u\in\mathbb{R}}\{zu - \Lambda(u)\} \geq \sup_{u\in\mathbb{R}}\{zu - \tilde{\Lambda}(u)\} := \tilde{I}(z).$$

and the function $\tilde{I}(z)$ is a *lower* bound on $I(z)$, $\forall z$.

4.4 PERTURBATION THEORY IN FINITE-DIMENSIONAL SPACES

Perturbation theory as a general term concerns the treatment of systems near to some known and exactly solved system. Problems tackled with perturbation theory are found ubiquitously in physics and mathematics. In our context, this form of perturbation theory involves the computation of

eigenvalues of a linear operator which deviates slightly from a simpler one. For more details on the background of this theory and more general cases see [105]. Here we provide some of the ideas relevant to Markov processes; in particular, where the convergence requirements of the perturbation expansion come from. We are interested in the case for a linear operator T on a finite-dimensional vector space X with spectrum $\mathrm{Sp}(T)$. Consider a perturbed operator $T(x)$ of the form

$$T(x) = T + xT^{(1)} + x^2T^{(2)} + \dots$$

and an eigenvalue $\lambda \in \mathrm{Sp}(T)$ of the unperturbed operator $T = T(0)$, with multiplicity $m_\lambda = 1$. We can ask whether the eigenvalues $\lambda(x)$ of $T(x)$ have a similar form, i.e. that they can be written as

$$\lambda(x) = \lambda + \sum_{n=1}^{\infty} x^n \lambda^{(n)} \tag{4.5}$$

and if so, when the series converges. Let $R(\xi, x) := (T(x) - \xi)^{-1}$ be the resolvent of $T(x)$, with $R(\xi, 0) := R(\xi)$, which is analytic on $\mathbb{C} \setminus \mathrm{Sp}(T(x))$. Therefore we can write

$$R(\xi, x) = R(\xi) + \sum_{n=1}^{\infty} x^n R^{(n)}(\xi), \quad \xi \in \mathbb{C} \setminus \mathrm{Sp}(T(x)) \tag{4.6}$$

for some coefficients $R^{(n)}(\xi)$. The ability to write (4.5) depends on the series in Equation (4.6) being uniformly convergent for all $|x| < r_0$ for some r_0, and for all ξ in a disc Γ, containing only one eigenvalue λ. The projection $P(x)$ corresponding to the eigenvalues of $T(x)$ inside Γ can be written as

$$P(x) = -\frac{1}{2\pi i} \int_\Gamma R(\xi, x) d\xi$$

and it can be shown that $\dim P(x) = m_\lambda = 1$, so the only eigenvalue of $T(x)$ inside Γ is $\lambda(x)$. The difference between this perturbed eigenvalue and λ can be written as

$$\lambda(x) - \lambda = \mathrm{tr}\left((T(x) - \lambda)P(x)\right). \tag{4.7}$$

The series (4.6) is uniformly convergent in x if

$$\sum_{n=1}^{\infty} |x|^n \left\| T^{(n)} R(\xi) \right\| < 1 \tag{4.8}$$

with $\| \cdot \|$ the norm associated with the vector space X. We can construct the required disc by denoting $r(\xi)$ the radius of convergence of the power series in the above condition (4.8), for some $\xi \in \Gamma$. Then to guarantee uniform convergence for all $\xi \in \Gamma$ we take the minimum

$$|x| < r_0 = \min_{\xi \in \Gamma} r(\xi).$$

This theory is very general, but there are some simplifications. In our case, the vector space X is a Hilbert space, and the operator T is normal (in fact

it is self-adjoint). This means we can apply perturbation theory for x in a disc of radius r_0 which is calculated as [67]

$$r_0 = (2ad^{-1} + c)^{-1}$$

where $a = \|T^{(1)}\|$, c is such that $\|T^{(n)}\|$ can be bounded by a term in a geometric series in the form $\|T^{(n)}\| \leq ac^{n-1}$ and $d = \min_{\mu \in \mathrm{Sp}(T)\setminus\{\lambda\}}\{\lambda - \mu\}$. Once we have r_0, Equations (4.7), (4.5) and (4.6) can be combined and with further calculations [105] the terms in Equation (4.5) can be written as

$$\lambda^{(n)} = \sum_{p=1}^{n} \frac{(-1)^p}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=l,\, \nu_i \geq 1 \\ \mu_1+\cdots+\mu_p=p-1,\, \mu_j \geq 0}} \mathrm{tr}\left(T^{\nu_1} S^{(\mu_1)}, \cdots, T^{\nu_p} S^{(\mu_p)}\right).$$

where $S^{(0)} = -P, S^{(n)} = S^n$, $S = (T - I + P)^{-1} - P$ and I is the identity.

4.5 EXTENSIONS TO MARKOV PROCESSES

So far we have only looked at concentration inequalities for independent random variables; however, there are numerous examples in statistics where Markov dependence is required. These include Markov chain Monte Carlo (MCMC), a numerical technique used to sample from complex distributions [68, 106]. Another case is in epidemic modelling, where finite regime analysis is used to determine the required sample size for success of "respondent-driven sampling" [107] to sample from a hidden population. Markov processes form the basis of some stock market forecasting models [108]; clearly, guarantees in finite time are of interest here. There have been numerous works extending Hoeffding and Bernstein's inequalities to Markov processes. Notably Glynn [109] and Lezaud [67] have concentration inequalities on functions of states, whilst Bernstein-type [69] and Hoeffding-type [68] further extend these to general state space Markov chains. These bounds are mostly for irreducible, discrete-time Markov processes, with the exception of [67, Theorem 3.4] which is a bound for irreducible continuous-time Markov processes.

Although we are interested in the continuous-time case, and in bounds for functions of transitions as opposed to states, we will see that techniques used in the above literature can be extended to our situation as well. With the exception of the Hoeffding inequality in [109] — which was derived in a different way — the bounds in [67–69] were derived via operator perturbation theory, a background for which can be found in Section 4.4. Below we summarise the method used in [67] to derive a concentration inequality for functions of the empirical measure of reversible discrete-time Markov processes. Recall a Markov process has finite state space E and expectation $\mathbb{E}_\nu$ with respect to some initial state ν. Let $f : E \to \mathbb{R}$, with $\|f\|_\infty \leq 1$. It can be taken that the asymptotic average of f, $\pi(f) = \sum_{x \in E} \pi(x) f(x) = 0$. As we discussed in Section 4.1.3, finding an upper bound on the MGF is enough to derive simple bounds via the Cramér-Chernoff method. Lemma

3.1 in [67] states that the moment generating function in this case can be written as the inner product

$$\mathbb{E}_\nu \left[\exp\left(\frac{u}{n} \sum_{i=1}^{n} (f(X_i)) \right) \right] = \langle \nu, \mathbf{P}(u)^n \underline{1} \rangle.$$

The perturbed transition operator $\mathbf{P}(u) = \mathbf{P}\mathbf{E}_u$ where $\mathbf{E}_u$ is the diagonal matrix with entries $\mathbf{E}_u(x) = e^{uf(x)}$. This inner product can be upper bounded using Cauchy-Schwarz inequality

$$\begin{aligned} \langle \nu, \mathbf{P}(u)^n \underline{1} \rangle &= \left\langle \nu, \mathbf{E}_u^{-\frac{1}{2}} \left(\mathbf{E}_u^{\frac{1}{2}} \mathbf{P} \mathbf{E}_u^{\frac{1}{2}} \right)^n \mathbf{E}_u^{\frac{1}{2}} \underline{1} \right\rangle \\ &\leq C(\nu) \left\| \mathbf{E}_u^{-\frac{1}{2}} \right\|_\pi \left\| \mathbf{E}_u^{\frac{1}{2}} \right\|_\pi \left\| \mathbf{E}_u^{\frac{1}{2}} \mathbf{P} \mathbf{E}_u^{\frac{1}{2}} \right\|_\pi^n \\ &\leq C(\nu) e^u r(u)^n. \end{aligned} \tag{4.9}$$

In the above steps we switch to the $L^2_\pi(E)$ inner product which gives $C(\nu) := \max_{x \in E} \{\nu(x)/\pi(x)\}$. Furthermore, the matrix $\mathbf{E}_u^{\frac{1}{2}} \mathbf{P} \mathbf{E}_u^{\frac{1}{2}}$ in the reversible, finite state space case is self-adjoint, therefore its spectral radius $r(u)$ coincides with its norm. It is also similar to the original perturbed transition operator $\mathbf{P}(u)$ so $r(u)$ is the spectral radius of $\mathbf{P}(u)$, and because of Perron-Frobenius theorem is also an eigenvalue of $\mathbf{P}(u)$. The spectral radius can be expanded in u (cf. Section 4.4 for details) where each coefficient of u^l is denoted $r^{(l)}$.

$$r(u) = 1 + \sum_{l=1}^{\infty} u^l r^{(l)}.$$

In the derivation of Lemma 3.2 in [67] it is found that $r^{(1)} = 0$ (this corresponds to $\pi(f)$ which we took to be zero). Each coefficient $r^{(l)}$ is upper bounded by some $\alpha^{(l)}$ such that $r^{(l)} \leq \alpha^{(l)}$.

$$\begin{aligned} r(u) &\leq 1 + \sum_{l=2}^{\infty} u^l \alpha^{(l)} \\ &\leq \exp\left(\tilde{\Lambda}(u) \right). \end{aligned}$$

The function $\tilde{\Lambda}(u) := \sum_{l=2}^{\infty} u^l \alpha^{(l)}$ is as an upper bound on the limiting logarithmic moment generating function. The final inequality was obtained using that $1 + x \leq \exp(x)$ and collapsing the bounded power series where we are left with an exponential bound on the spectral radius, and therefore also the moment generating function. When we combine the above with Equation (4.9) we are able to derive a concentration inequality, and optimising in u completes the Cramér-Chernoff method.

For non-reversible Markov processes, $\mathbf{P}$ and therefore $\mathbf{P}(u)$ is not self-adjoint in $L^2_\pi(E)$, which is a requirement of the proof. This can be solved by the introduction of symmetrised transition operators as in [67], or by the usage of the Léon-Perron operator (Lemma 7.6) which is used to further upper bound the Cauchy-Schwarz inequality with a self-adjoint operator and allows the perturbative method to be applied. We also make use of the

Léon-Perron version of transition operators in the classical and quantum chapters (Definitions 2.2 and 5.7 respectively). The approach outlined in this section can be applied to continuous-time Markov processes, and the majority of our results consist of extending this further to derive concentration inequalities for both dynamical observables and first passage times.

5

QUANTUM MARKOV PROCESSES

In this chapter we introduce the basic concepts and tools related to quantum Markov processes, in particular, quantum counting processes. Quantum counting processes are used to model detector clicks when an open quantum system is continuously monitored via the environment [110]. Examples of such systems are optomechanical devices in magnetic field sensing [111, 112], gravitational wave detectors [113] and atomic clocks [114]. In these examples the sensitivity of the quantum devices are exploited to obtain extremely precise measurements.

Quantum Markov processes differ to classical Markov processes in that whilst the system state undergoes an evolution, what is observed are the emissions into the environment Throughout this thesis, the quantum system will be finite-dimensional and its state space will be the Hilbert space $\mathbb{C}^d$. Quantum states are represented by positive semi-definite matrices with unit trace, that is $\rho \in M_d(\mathbb{C})$ such that $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$. Observables correspond to self-adjoint operators on the state space, i.e. $x \in M_d(\mathbb{C})$ such that $x = x^*$. As in the classical case, there is a duality between states and observables expressed by the fact that the expectation of an observable x in the state ρ is $\mathrm{tr}(\rho x)$. We can endow $M_d(\mathbb{C})$ with the operator and trace norms which provide natural distances between observables and states, respectively:

$$\|x\|_\infty := \max_{u \in \mathbb{C}^d \setminus \{0\}} \frac{\|xu\|}{\|u\|}, \quad \|x\|_1 := \mathrm{tr}(|x|).$$

We will state the formal definition of a quantum channel in Section 5.1, which describes how states evolve in open quantum dynamics. In Section 5.2 we will describe quantum Markov semigroups and their unravelling by counting measurements. Sections 5.3, 5.4 and 5.5 refer to mathematical notions used in these works, including the non-commuting version of the Perron-Frobenius theorem. Finally, Section 5.6 formalises counting observables in an analogous way to the classical setting.

5.1 QUANTUM CHANNELS

For any linear map Φ on $M_d(\mathbb{C})$ describing the evolution of observables (Heisenberg picture), the unique corresponding evolution Φ_* on states (Schrödinger picture) is characterised by

$$\mathrm{tr}(x\Phi(y)) = \mathrm{tr}(\Phi_*(x)y)$$

where Φ_* denotes the *predual* of the map Φ. As in the previous section, we denote by $\|\Phi\|_{\infty\to\infty}$ the operator norm on Φ induced by $\|\cdot\|_\infty$ and analo-

gously for the trace norm. Every physical evolution of a quantum system is given by a quantum channel.

Definition 5.1 (Quantum channel)**.** *A linear map $\Phi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ is a quantum channel if the following conditions are satisfied:*

1. $\Phi(\mathbf{1}) = \mathbf{1}$ *(unitality),*
2. $\Phi \otimes \mathbf{I}_{M_n(\mathbb{C})}$ *is positive for every $n \in \mathbb{N}$ (complete positivity)*

where $\mathbf{1}$ is the identity matrix and $\mathbf{I}_{M_n(\mathbb{C})}$ is the identity map on $M_n(\mathbb{C})$. These conditions are equivalent in the Schrödinger picture to Φ_* being trace preserving and completely positive. These conditions mirror those satisfied by classical channels/transition operators, but in the quantum setting complete positivity is a stronger requirement that usual positivity, and we refer to [115] for more details on the theory of quantum channels and the physical interpretation.

5.2 QUANTUM MARKOV SEMIGROUPS

5.2.1 *GKLS generator*

A quantum Markov semigroup is a family of channels $(\mathcal{T}_t)_{t\geq 0}$ acting on $M_d(\mathbb{C})$ such that $\mathcal{T}_0 = \mathbf{I}_{M_d(\mathbb{C})}$, $\mathcal{T}_t \circ \mathcal{T}_s = \mathcal{T}_{t+s}$ for all $s, t \geq 0$ and $t \mapsto \mathcal{T}_t$ is continuous. Such a semigroup describes (in the Heisenberg picture) the dissipative evolution of a d-dimensional open quantum system, in physical situations where certain Markov approximations pertaining to the interaction with the environment apply. A fundamental result [116, 117] shows that the generator $\mathcal{L} : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ (known as GKLS generator) of such a semigroup takes the form

$$\mathcal{L} : x \mapsto -i[H, x] + \sum_{i\in I} L_i^* x L_i - \frac{1}{2}\sum_{i\in I}(L_i^* L_i x + x L_i^* L_i), \tag{5.1}$$

where $H \in M_d(\mathbb{C})$ is self-adjoint and $L_i \in M_d(\mathbb{C})$ with indices belonging to a finite set I. Physically, H is interpreted as being the system hamiltonian while L_i describe the coupling to separate "emission channels" in the environment. If the system is prepared in a state ρ and evolves together with the environment for a time period t, then its reduced state is given by $\mathcal{T}_{t*}(\rho)$. On the other hand, if the environment is probed by performing continuous-time counting measurements in each of the emission channels, then one observes stochastic trajectories $\omega = \{(i_1, t_1), (i_2, t_2), \cdots\}$ which record the labels of the jumps together with the times between jumps.

Remark 5.2. A classical Markov chain can be embedded into a quantum Markov process, by setting $H = 0$, $I = \mathfrak{E}$, and the jump operators $L_{ij} = \sqrt{w_{ij}}\,|x_j\rangle\langle x_i|$ for an orthonormal basis $\{|x_j\rangle\}_{j=1}^d$.

5.2.2 *Counting measurements*

In this case one would like to know the probability of observing such a trajectory and the conditional state of the system given this observation. This is the subject of quantum filtering theory which plays an important role in quantum technology and quantum control theory [57–60]. While a full account of the system-environment unitary evolution and subsequent counting measurement goes beyond the scope of this thesis, we employ the Dyson series to convey an intuitive answer to the questions formulated above. For this we decompose the generator as

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{J} = \mathcal{L}_0 + \sum_{i\in I} \mathcal{J}_i, \tag{5.2}$$

where

$$\mathcal{L}_0(x) = G^*x + xG \quad \text{with} \quad G := iH - \frac{1}{2}\sum_{i\in I} L_i^* L_i,$$

and

$$\mathcal{J}(x) = \sum_{i\in I} \mathcal{J}_i(x) \quad \text{with} \quad \mathcal{J}_i(x) = L_i^* x L_i.$$

Note that $\mathcal{J}_i$ is completely positive and $\mathcal{L}_0$ is the generator of the completely positive semigroup $e^{\mathcal{L}_0}(x) = e^{G^*t} x e^{Gt}$. The Dyson expansion of $\mathcal{T}_{*t}$ (Schrödinger picture) corresponding to the split (5.2) is

$$\mathcal{T}_{t*} = e^{t\mathcal{L}_{0*}} + \sum_{k=1}^{+\infty} \int_{\sum_{i=1}^k t_i \le t} e^{\left(t-\sum_{i=1}^k t_i\right)\mathcal{L}_{0*}} \mathcal{J}_{i_k *} e^{t_k \mathcal{L}_{0*}} \cdots \mathcal{J}_{i_1 *} e^{t_1 \mathcal{L}_{0*}} dt_1 \cdots dt_k.$$

By applying both sides to the initial state ρ we find that the evolved system state $\rho_t = \mathcal{T}_{*t}(\rho)$ is a mixture of states corresponding to different counting trajectories. Indeed, let us denote

$$\Omega_t = \{\emptyset\} \cup \bigcup_{k=1}^{+\infty} I^k \times \left\{ (t_1, \ldots, t_k) \in [0,t]^k : \sum_{i=1}^k t_i \le t \right\}$$

the space of counting trajectories up to time t, and let us endow Ω_t with the natural σ-field and denote by $d\mu$ the unique measure such that $\mu(\{\emptyset\}) = 1$ and $\mu(\{(i_1, \ldots, i_k) \times B)$ is the Lebesgue measure of B for every $(i_1, \ldots, i_k) \in I^k$, $B \subseteq \{(t_1, \ldots, t_k) \in [0,t]^k : \sum_{i=1}^k t_i \le t\}$. Then we can write

$$\mathcal{T}_{t*}(\rho) = \int_{\Omega_t} \tilde{\varrho}_t(\omega)\mu(d\omega) = \int_{\Omega_t} \frac{d\mathbb{P}_t}{d\mu}(\omega)\varrho_t(\omega)\mu(d\omega)$$

where for each counting trajectory $\omega = \{(i_1, t_1), \ldots (i_k, t_k)\} \in \Omega_t$. The *unnormalised* system state conditional on observing ω is given by

$$\tilde{\varrho}_t(\omega) = e^{\left(t-\sum_{i=1}^k t_i\right)\mathcal{L}_{0*}} \mathcal{J}_{i_k *} e^{t_1 \mathcal{L}_{0*}} \cdots \mathcal{J}_{i_1 *} e^{t_1 \mathcal{L}_{0*}}(\rho) \tag{5.3}$$

while

$$\frac{d\mathbb{P}_t}{d\mu}(\omega) = \mathrm{tr}(\tilde{\varrho}_t(\omega)), \quad \varrho_t(\omega) = \frac{\tilde{\varrho}_t(\omega)}{\mathrm{tr}(\tilde{\varrho}_t(\omega))} \tag{5.4}$$

represent the probability density, and the *normalised* conditional state, respectively.

With this interpretation, the Dyson expansion expresses the fact that by averaging over all the conditional states $\varrho_t(\omega)$ one obtains the reduced system state ρ_t. Note that in order to avoid confusion, we use different symbols for the conditional and reduced system states.

Based on Equations (5.3) and (5.4) we deduce that during time periods with no jumps the conditional state evolves continuously as

$$\varrho_t \mapsto \varrho_{t+s} := \frac{e^{s\mathcal{L}_{0*}}(\varrho_t)}{\mathrm{tr}[e^{s\mathcal{L}_{0*}}(\varrho_t)]},$$

and at the time of a count with index i the state has an instantaneous jump

$$\varrho_t \mapsto \frac{\mathcal{J}_{i*}(\varrho_t)}{\mathrm{tr}[\mathcal{J}_{i*}(\varrho_t)]}.$$

In addition, the probability density for the time of the first jump after t is

$$w(s) = \mathrm{tr}[\mathcal{J}_* e^{s\mathcal{L}_{0*}}(\varrho_t)] = -\mathrm{tr}[\mathcal{L}_{0*} e^{s\mathcal{L}_{0*}}(\varrho_t)]$$

where we use the fact that $\mathcal{J}(\mathbf{1}) + \mathcal{L}_0(\mathbf{1}) = \mathcal{L}(\mathbf{1}) = 0$. We will now show how to generate the count trajectories in a recursive manner which is reminiscent of the generation of trajectories of classical Markov processes. Given a trajectory $\omega = \{(i_1, t_1), (i_2, t_2), \dots\}$, we denote by ϱ_k the state immediately after the k^{th} count, with $\varrho_0 = \rho$ denoting the initial state.

5.2.3 *Generating quantum trajectories*

The interarrival times and quantum trajectories can be generated recursively with respect to $k = 0, 1, \dots$: given ϱ_k we draw $(\varrho_{k+1}, i_{k+1}, t_{k+1})$ as follows:

1. the $(k+1)$th interarrival time t_{k+1} is drawn from the density
$$w(t) = -\mathrm{tr}(\mathcal{L}_{0*} e^{t\mathcal{L}_{0*}}(\varrho_k))$$
2. given t_{k+1}, the label i_{k+1} is sampled from the following distribution:
$$p(j) = \frac{\mathrm{tr}(\Phi_{j*}\mathcal{L}_{0*} e^{t_{k+1}\mathcal{L}_{0*}}(\varrho_k))}{\mathrm{tr}(\Phi_*\mathcal{L}_{0*} e^{t_{k+1}\mathcal{L}_{0*}}(\varrho_k))}, \quad \Phi_j := -\mathcal{L}_0^{-1}\mathcal{J}_j, \quad \Phi := \sum_{j\in I}\Phi_j$$
3. we define
$$\varrho_{k+1} = \frac{\mathcal{J}_{i_{k+1}*}(e^{t_{k+1}\mathcal{L}_{0*}}(\varrho_k))}{\mathrm{tr}(\mathcal{J}_{i_{k+1}*}(e^{t_{k+1}\mathcal{L}_{0*}}(\varrho_k)))}.$$

The map $\Phi := \sum_{j\in I}\Phi_j = -\mathcal{L}_0^{-1}\mathcal{J}$ appearing in step 2. is the analogue of $\mathbf{P}$ in Chapter 2 and will play a central role in the following. Lemma 8.2 in Chapter 8 shows that $\mathcal{L}_0^{-1}$ is well defined and is equal to $-\int_0^{+\infty} e^{t\mathcal{L}_0}dt$, hence Φ is a completely positive map. Moreover, using that $\mathcal{L}(\mathbf{1}) = 0$, one has

$$\Phi(\mathbf{1}) = -\mathcal{L}_0^{-1}\mathcal{J}(\mathbf{1}) = \mathcal{L}_0^{-1}\mathcal{L}_0(\mathbf{1}) = \mathbf{1},$$

hence Φ is a quantum channel.

5.3 ERGODICITY

As in the classical case, establishing results on law of large numbers, large deviations, or concentration bounds, requires some type of assumption on the ergodicity of the dynamics. We now introduce two irreducibility assumptions which will later be invoked in separate occasions in our results.

Hypothesis 5.3 (Irreducibility of $\mathcal{L}$)**.** *The generator $\mathcal{L}$ is irreducible. This means that there is no non-trivial projection P such that $\mathcal{L}(P) \geq 0$ or equivalently, there exists a unique strictly positive state $\hat{\sigma}$ satisfying $\mathcal{L}_*(\hat{\sigma}) = 0$.*

Hypothesis 5.4 (Irreducibility of Φ)**.** *The channel Φ is irreducible. This means that there is no non-trivial projection P such that $\Phi(P) \geq P$ or equivalently, there exists a unique strictly positive state σ satisfying $\Phi_*(\sigma) = \sigma$.*

As in the classical case, there is a close connection between the generator $\mathcal{L}$ and the channel Φ. We will clarify this connection in Lemma 8.1, Chapter 8.

5.4 NON-COMMUTATIVE PERRON-FROBENIUS THEOREM

We refer to [42, Theorem 3.1.1] for the commutative case and to [118] for the general case of C^*-algebras (which includes the cases of our interest). Let $\mathcal{A}$ be either $L^\infty(E)$ or $M_d(\mathbb{C})$ and let Ψ be a positivity preserving map acting on $\mathcal{A}$. Note that in the quantum case (in our case $\mathcal{A} = M_d(\mathbb{C})$), $\mathcal{A}$ is not a commuting algebra hence Theorem 5.5 is referred to as "non-commutative". In the classical case, Φ acts on an algebra of functions, whose pointwise product is commutative. Ψ is said to be irreducible if there exists no non-trivial projection $p \in \mathcal{A}$ such that

$$\Psi(p) \leq \alpha p \tag{5.5}$$

for some positive constant α.

Theorem 5.5 (Perron-Frobenius)**.** *Let r be the spectral radius of Ψ. The following statements hold true.*

1. *$r \in \mathrm{Sp}(\Psi)$ and there exists $x \in \mathcal{A}$, $x \geq 0$ such that $\Psi(x) = rx$.*

If Ψ is irreducible, then one has some further results.

1. *r is a geometrically simple eigenvalue;*
2. *x is strictly positive and is the unique positive eigenvector.*

Another important result is the Russo-Dye theorem.

Theorem 5.6 (Russo-Dye)**.** *Ψ attains its norm at the identity of $\mathcal{A}$, i.e.*

$$\|\Psi\|_{\infty\to\infty} = \|\Psi(e)\|_\infty$$

where e is the identity element ($\underline{1}$ or $\mathbf{1}$ for $L^\infty(E)$ and $M_d(\mathbb{C})$ respectively).

5.5 KUBO-MARTIN-SCHWINGER INNER PRODUCT

There are infinitely many inner products induced by the stationary state σ of Φ; the choice we adopt is known as the Kubo-Martin-Schwinger (KMS) inner product. We endow the following inner product to the Hilbert space structure of $M_d(\mathbb{C})$:

$$\langle x, y\rangle_\sigma := \mathrm{tr}(\sigma^{\frac{1}{2}} x^* \sigma^{\frac{1}{2}} y), \quad x, y \in M_d(\mathbb{C})$$

and we denote it by L^2_σ. The norm with respect to this inner product will be denoted by $\|x\|_\sigma$. The KMS inner product allows us to define the trace of a map $\mathcal{E} : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ by

$$\mathrm{TR}(\mathcal{E}) = \sum_{i=1}^{d^2} \langle x_i, \mathcal{E}(x_i)\rangle_\sigma,$$

for an orthonormal basis $\{x_i\}$ of $M_d(\mathbb{C})$. The adjoint of an operator $\mathcal{E}$ with respect to this inner product can be expressed in terms of the predual map $\mathcal{E}_*$ as

$$\mathcal{E}^\dagger(x) = \Gamma^{-\frac{1}{2}} \circ \mathcal{E}_* \circ \Gamma^{\frac{1}{2}}(x) \tag{5.6}$$

where $\Gamma^a(x) = \sigma^a x \sigma^a$ for every $a \in \mathbb{R}$.

Given a quantum channel Φ with invariant state σ, its absolute spectral gap ε is defined as 1 minus the square root of the second largest eigenvalue of the multiplicative symmetrisation of Φ, namely $\Phi^\dagger\Phi$. As with its equivalent in the classical case, the proof of the concentration bound in Theorem 8.6 will make use of the following quantum version of Definition 2.2.

Definition 5.7 (Quantum León-Perron operator)**.** *A quantum channel $\hat{\Phi}_\lambda$ on L^2_σ is said to be León-Perron if it is a convex combination of operators $\mathbf{I}_{M_d(\mathbb{C})}$ and Π_σ with coefficient $\lambda \in [0, 1]$, that is*

$$\hat{\Phi}_\lambda = \lambda \mathbf{I}_{M_d(\mathbb{C})} + (1 - \lambda)\Pi_\sigma,$$

where $\mathbf{I}_{M_d(\mathbb{C})}$ is the identity map, and Π_σ is the map $\Pi_\sigma x \mapsto \mathrm{tr}(\sigma x)\mathbf{1}$ for $x \in M_d(\mathbb{C})$.

$\hat{\Phi}_\lambda$ is a quantum channel with unique invariant state σ and which is self-adjoint with respect to the KMS inner product induced by σ. As with the classical León-Perron operator we set $\lambda = 1 - \varepsilon$ and denote as follows

$$\hat{\Phi} := (1 - \varepsilon)\mathbf{I}_{M_d(\mathbb{C})} + \varepsilon\Pi_\sigma.$$

5.6 QUANTUM COUNTING OBSERVABLES

We can define analogous processes to the counting observables in the classical case (Section 2.5.1 Chapter 2). Consider the counting process described in Section 5.2 and let $N_t(i)$ be the stochastic process given by the number of

counts with label $i \in I$ up to time t in the measurement trajectory ω. More generally, for any subset $\mathfrak{A} \subseteq I$ we define the counting observable

$$N_{\mathfrak{A},t} = \sum_{i \in \mathfrak{A}} N_t(i).$$

When $\mathfrak{A} = I$, $N_{I,t}$ is referred to as the total number of counts. The corresponding first passage times (FPTs) are defined in the same way as in the classical case (Definition 2.4):

$$T_{\mathfrak{A}}(k) := \inf_{t \geq 0} \{t : N_{\mathfrak{A},t} = k\}. \tag{5.7}$$

Quantum counting observables also obey a large deviation principle (LDP) [61–64] and for finite time concentration inequalities (cf. Chapter 4) have been derived for empirical measures [66] and emission counts [65]. In Chapter 8 we extend these and derive concentration inequalities for FPTs of emission counts.

6

GENERAL UPPER BOUNDS ON FLUCTUATIONS OF TRAJECTORY OBSERVABLES

Recall that the thermodynamic uncertainty relations (TURs) are limited to providing lower bounds on the size of fluctuations: except in the few cases where they are tight, inference on the observable of interest is hindered by the absence of a corresponding upper bound. In this chapter we correct this issue by introducing a class of *general upper bounds* for fluctuations of trajectory observables consisting of linear combination of fluxes of a continuous-time Markov chain, which includes all currents and activities. For lack of a better name, we call these "inverse thermodynamic uncertainty relations". The inverse TURs are valid for all times and bound fluctuations at all levels. Figure 6.1 illustrates our results: the large deviation rate function $I(A/t)$ for a current A is upper bounded by the TUR, as known, and lower bounded by our inverse TUR.

In this chapter we prove these general relations using spectral and perturbation techniques that are widely applied in the field of quantum and classical Markov processes (see Section 4.4, or [65–69, 109] and references therein). First we provide an analytical expression (Lemma 6.1) for the variance of generalised fluxes, valid for all times. We then derive our main results: an upper bound on this variance (Theorem 6.3), and a concentration inequality for generalised fluxes. Theorem 6.3 can be used in upper bounding precision estimation of time-averaged currents (Section 6.4). We analyse in Section 6.5 the behaviour of the additive spectral gap ε, the spectral gap of $\Re(\mathbf{L}) = (\mathbf{L} + \mathbf{L}^\dagger)/2$, which appears in the upper bounds. In Section 6.5 we show that $\varepsilon \to 0$ as the system approaches a phase transition, causing the inverse TUR to diverge and suggesting an increase in fluctuations. By considering some examples, in Section 6.6, we compare the inverse TURs with the TURs and discuss another model, illustrating these fluctuation patterns near a phase transition.

6.1 EXPLICIT FORM OF THE VARIANCE

Although the main results in this chapter concern bounding the size of fluctuations in the form of variance upper bounds, it is possible to compute the variance analytically. The expression can be derived directly from the moment generating function (MGF). We first define the *static variance*:

$$\langle a^2 \rangle_{\hat{\pi}} = \sum_{x \neq y} \hat{\pi}(x) w_{xy} a_{xy}^2$$

(corresponding to the variance of $\sum_{x \neq y} a_{xy} \tilde{K}(x, y)$, where $\tilde{K}_{xy}$ are independent Poisson variables with intensity $\hat{\pi}(x) w_{xy}$). It is important to note that

the static variance $\langle a^2 \rangle_{\hat{\pi}} \neq \lim_{t\to+\infty} \frac{\mathbb{E}_{\hat{\pi}}[A(t)^2]}{t}$ and so the asymptotic variance $V_{\infty,A}$ (cf. Section 2.5) is not simply $\langle a^2 \rangle_{\hat{\pi}} - \langle a \rangle_{\hat{\pi}}^2$. In fact, the correct limit to compute $V_{\infty,A}$ is

$$V_{\infty,A} = \lim_{t\to+\infty} \frac{\mathbb{E}_{\hat{\pi}}[A(t)^2] - \mathbb{E}_{\hat{\pi}}[A(t)]^2}{t}$$

which we will use in Lemma 6.1. Recall from Equation (3.2) Chapter 3 that the moment generating function of any dynamical observable (cf. Definition 2.3) is

$$\mathbb{E}_\nu \left[e^{uA(t)} \right] = \left\langle \nu, e^{t\mathbf{L}_u} \underline{1} \right\rangle$$

where the the tilted generator is defined as

$$\mathbf{L}_u = \sum_{x\neq y} (e^{ua_{xy}} - 1) w_{xy} \left| x \right\rangle \left\langle y \right| + \mathbf{L}.$$

We denote $\mathbf{L}'_u$ the derivative of $\mathbf{L}_u$ with respect to the parameter u and $\mathbf{L}^{-1}$ the pseudoinverse $\mathbf{L}^{-1} := (\mathbf{L} + \Pi_{\hat{\pi}})^{-1} - \Pi_{\hat{\pi}}$. This pseudoinverse has the property that $\mathbf{L}^{-1}\mathbf{L}f = f$ if $\langle f, \underline{1} \rangle_{\hat{\pi}} = 0$. The operator $\Pi_{\hat{\pi}}$ is the map $\Pi_{\hat{\pi}} : f \mapsto \langle f, \underline{1} \rangle_{\hat{\pi}} \underline{1}$.

Lemma 6.1. *The variance $V_{\hat{\pi},A}(t)$ of any time-integrated current or flux observable $A(t)$ in the stationary state has the explicit form, for all $t \geq 0$*

$$V_{\hat{\pi},A}(t) = t\langle a^2 \rangle_{\hat{\pi}} + 2 \left\langle \hat{\pi}, \mathbf{L}'_0((e^{t\mathbf{L}} - \mathbf{1})\mathbf{L}^{-1} - t\mathbf{1})\mathbf{L}^{-1}\mathbf{L}'_0\underline{1} \right\rangle,$$

with the asymptotic variance given by

$$V_{\infty,A} = \lim_{t\to+\infty} \frac{V_{\hat{\pi},A}(t)}{t} = \langle a^2 \rangle_{\hat{\pi}} - 2 \left\langle \hat{\pi}, \mathbf{L}'_0\mathbf{L}^{-1}\mathbf{L}'_0\underline{1} \right\rangle.$$

Proof. Applying the equation for the derivative of a matrix exponential to $e^{t\mathbf{L}_u}$

$$\frac{de^{t\mathbf{L}_u}}{du} = \int_0^t e^{(t-s)\mathbf{L}_u} \mathbf{L}'_u e^{s\mathbf{L}_u} ds.$$

We obtain the first moment of $A(t)$

$$\left. \frac{d\mathbb{E}_{\hat{\pi}} \left[e^{uA(t)} \right]}{du} \right|_{u=0} = t\langle \hat{\pi}, \mathbf{L}'_0\underline{1} \rangle = t\langle a \rangle_{\hat{\pi}}.$$

The second moment can be broken into three parts by applying Leibniz rule when taking the second derivative. The first part comes from the $\mathbf{L}'_u$ in the product

$$Z''^{(1)}_{\hat{\pi},t}(0) = \int_0^t \left\langle \hat{\pi}, e^{(t-s)\mathbf{L}} \mathbf{L}''_0 e^{s\mathbf{L}} \underline{1} \right\rangle ds = t\langle a^2 \rangle_{\hat{\pi}}.$$

The remaining two terms from the $e^{(t-s)\mathbf{L}_u}$ and the $e^{s\mathbf{L}_u}$ are in fact equal, giving

$$Z''^{(2)}_{\hat{\pi},t}(0) = \int_0^t \int_0^{t-s} \langle \hat{\pi}, \mathbf{L}'_0 e^{r\mathbf{L}} \mathbf{L}'_0 \underline{1} \rangle \, drds.$$

We write $\mathbf{L}_0'\underline{1} = \langle a\rangle_{\hat{\pi}}\underline{1} + (\mathbf{L}_0'\underline{1} - \langle a\rangle_{\hat{\pi}}\underline{1})$ (recall that $\langle \hat{\pi}, \mathbf{L}_0'\underline{1}\rangle = \langle a\rangle_{\hat{\pi}}$) and break the integral into two parts, the first of which is

$$\langle a\rangle_{\hat{\pi}} \int_0^t \int_0^{t-s} \left\langle \hat{\pi}, \mathbf{L}_0' e^{r\mathbf{L}}\underline{1}\right\rangle dr ds = \frac{t^2}{2}\langle a\rangle_{\hat{\pi}}^2.$$

The second part is then obtained from evaluating the integral on the subspace orthogonal to $\underline{1}$, and we can bring in the pseudoinverse $\mathbf{L}^{-1}$

$$\int_0^t \int_0^{t-s} \left\langle \hat{\pi}, \mathbf{L}_0' e^{r\mathbf{L}} f\right\rangle dr ds = \left\langle \hat{\pi}, \mathbf{L}_0'((e^{t\mathbf{L}} - \mathbf{1})(\mathbf{L}^{-1})^2 - t\mathbf{L}^{-1})\mathbf{L}_0'\underline{1}\right\rangle,$$

where $f = \mathbf{L}_0'\underline{1} - \langle a\rangle_{\hat{\pi}}\underline{1}$. Combining all parts gives

$$\begin{aligned}\left.\frac{d^2\mathbb{E}_{\hat{\pi}}\left[e^{uA(t)}\right]}{du^2}\right|_{u=0} &= Z_{\hat{\pi},t}^{\prime\prime(1)} + 2Z_{\hat{\pi},t}^{\prime\prime(2)} \\ &= t\langle a^2\rangle_{\hat{\pi}} + t^2\langle a\rangle_{\hat{\pi}}^2 + 2\left\langle \hat{\pi}, \mathbf{L}_0'((e^{t\mathbf{L}} - \mathbf{1})\mathbf{L}^{-1} - t\mathbf{1})\mathbf{L}^{-1}\mathbf{L}_0'\underline{1}\right\rangle.\end{aligned}$$

Therefore we can write $V_{\hat{\pi},A}(t)$ as

$$\begin{aligned}V_{\hat{\pi},A}(t) &= \mathbb{E}_{\hat{\pi}}\left[A(t)^2\right] - \mathbb{E}_{\hat{\pi}}\left[A(t)\right]^2 \\ &= t\langle a^2\rangle_{\hat{\pi}} + 2\left\langle \hat{\pi}, \mathbf{L}_0'((e^{t\mathbf{L}} - \mathbf{1})\mathbf{L}^{-1} - t\mathbf{1})\mathbf{L}^{-1}\mathbf{L}_0'\underline{1}\right\rangle.\end{aligned}$$

We remark that the asymptotic variance is then obtained by taking the limit

$$V_{\infty,A} = \lim_{t\to+\infty}\frac{V_{\hat{\pi},A}(t)}{t} = \langle a^2\rangle_{\hat{\pi}} - 2\left\langle \hat{\pi}, \mathbf{L}_0'\mathbf{L}^{-1}\mathbf{L}_0'\underline{1}\right\rangle$$

which agrees with that obtained from the central limit theorem.

□

We can also recover the static variance by taking the zero time limit

$$\lim_{t\to 0}\frac{V_{\hat{\pi},A}(t)}{t} = \langle a^2\rangle_{\hat{\pi}}.$$

Lemma 6.1 is used to compare the TURs and the inverse TURs in Section 6.4; however, large system sizes render it impractical for direct computation. This justifies the need for both upper and lower bounds in terms of simpler quantities of the system.

6.2 UPPER BOUND ON THE VARIANCE

In this section we provide our first main result for this chapter: the inverse TUR. It is inverse in the sense that whilst the TURs correspond to lower bounds on fluctuations and upper bounds on the large deviation rate function, the "inverse" TURs correspond to upper bounds on fluctuations and lower bounds on the rate function. To state the remaining results in this chapter we need the following quantities.

1. Maximum escape rate:
$$q := \max_{x \in E}\{R_x\}; \tag{6.1}$$
2. Maximum amplitude of observable coefficients:
$$c = \max_{x \neq y} |a_{xy}|;$$
3. Additive spectral gap: ε the spectral gap of the symmetrisation $\Re(\mathbf{L}) = (\mathbf{L} + \mathbf{L}^\dagger)/2$, where the adjoint is taken with respect to $\hat{\pi}$. The additive spectral gap is a measure of how fast the equilibrium process with generator $\Re(\mathbf{L})$ reaches $\hat{\pi}$.

The following lemma is required for the next theorem. Let us introduce the notation $\mathbf{J}_{xy} = w_{xy} |x\rangle \langle y|$, we write
$$\mathbf{D}^{(k)} := \sum_{x \neq y} a_{xy}^k \Re(\mathbf{J}_{xy}).$$
First we state a lemma required for the proof.

Lemma 6.2. *The following estimates hold true for $k \geq 1$:*

1. $\|\mathbf{D}^{(k)}\|_{\hat{\pi}} \leq q c^k$,
2. $\|\mathbf{D}^{(k)}(\mathbf{1})\|_{\hat{\pi}} \leq \langle a^2 \rangle_{\hat{\pi}}^{1/2} q^{1/2} c^{k-1}$,
3. $2q/\varepsilon \geq 1$.

Proof. Notice that
$$\left(\sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy} \right)^\dagger = \sum_{x \neq y} \frac{\hat{\pi}(y)}{\hat{\pi}(x)} w_{yx} a_{yx}^k |x\rangle \langle y| .$$
Since $\mathbf{D}^{(k)} = \sum_{x \neq y} a_{xy}^k \Re(\mathbf{J}_{xy})$, we have that
$$\|\mathbf{D}^{(k)}\|_{\hat{\pi}} \leq \left\| \sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy} \right\|_{\hat{\pi}} = \left\| \left(\sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy} \right)^\dagger \right\|_{\hat{\pi}}$$
and
$$\|\mathbf{D}^{(k)} \underline{1}\|_{\hat{\pi}} \leq \max \left\{ \left\| \sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy} \underline{1} \right\|_{\hat{\pi}}, \left\| \sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy}^\dagger \underline{1} \right\|_{\hat{\pi}} \right\}.$$

1. Let $f : E \to \mathbb{R}$ be a function. Then
$$\begin{aligned} \left\| \sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy}(f) \right\|_{\hat{\pi}}^2 &= \sum_{x \in E} \hat{\pi}(x) \left(\sum_{y:\, y \neq x} w_{xy} a_{xy}^k f_y \right)^2 \\ &\leq \sum_{x \neq y} \hat{\pi}(x) R_x w_{xy} a_{xy}^{2k} f_y^2 \\ &\leq q c^{2k} \sum_{x \neq y} \hat{\pi}(x) w_{xy} f_y^2 \\ &= q c^{2k} \sum_{y \in E} \hat{\pi}(y) R_y f_y^2 \\ &\leq q^2 c^{2k} \|f\|_{\hat{\pi}}^2. \end{aligned}$$

In the first inequality we applied Jensen's inequality, while in the last equality we made use of the fact that $\hat{\pi}$ is invariant for $\mathbf{L}$.

2. With analogous tricks, one obtains the following:

$$\begin{aligned}\left\|\sum_{x\neq y} a_{xy}^k \mathbf{J}_{xy}\underline{1}\right\|_{\hat{\pi}}^2 &= \sum_{x\in E}\hat{\pi}(x)\left(\sum_{y:\,y\neq x} w_{xy}a_{xy}^k\right)^2\\ &\leq \sum_{x\neq y}\hat{\pi}(x)R_x w_{xy}a_{xy}^{2k}\\ &\leq qc^{2(k-1)}\langle a^2\rangle_{\hat{\pi}}\end{aligned}$$

and

$$\begin{aligned}\left\|\sum_{x\neq y} a_{xy}^k \mathbf{J}_{xy}^\dagger\underline{1}\right\|_{\hat{\pi}}^2 &= \sum_{x\in E}\hat{\pi}(x)\left(\sum_{y:\,y\neq x}\frac{\hat{\pi}(y)}{\hat{\pi}(x)} w_{yx}a_{yx}^k\right)^2\\ &\leq \sum_{x\neq y}\hat{\pi}(y)R_x w_{yx}a_{yx}^{2k}\\ &\leq qc^{2(k-1)}\langle a^2\rangle_{\hat{\pi}}.\end{aligned}$$

3. For notation convenience, let us identify E with the ordered set $\{1,\dots,|E|\}$. Let us consider the diagonalisation of $\Re(\mathbf{L}) = \mathbb{U}^*\Lambda\mathbb{U}$, where $\mathbb{U}$ is the matrix having as rows the coordinates of a orthonormal basis of eigenvectors (we can pick $\mathbb{U}$ having real entries) and Λ is the diagonal matrix of real eigenvalues $\lambda_1 = 0 \geq \lambda_2 = -\varepsilon \geq \cdots \geq \lambda_{|E|}$ in decreasing order. Notice that $\Re(\mathbf{L})_{xx} = -R_x$ and $u_{1x} = \sqrt{\hat{\pi}(x)}$ for every $x \in E$ ($\underline{1}$ the unique eigenvector corresponding to the eigenvalue 0 and $u_{1x} = \langle \delta_x/\sqrt{\hat{\pi}(x)}, \underline{1}\rangle_{\hat{\pi}} = \sqrt{\hat{\pi}(x)}$). Therefore we have

$$\begin{aligned}-R_x = \Re(\mathbf{L})_{xx} = (\mathbb{U}^*\Lambda\mathbb{U})_{xx} &= \sum_{y=1}^{|E|}\lambda_y u_{yx}^2\\ &\leq \lambda_2\sum_{y\neq 1}u_{yx}^2 = \lambda_2(1-\hat{\pi}(x)),\end{aligned}$$

hence

$$\frac{q}{\varepsilon} \geq 1 - \min_x \hat{\pi}(x) \geq 1 - \frac{1}{|E|} \geq \frac{1}{2}.$$

□

We now state the theorem.

Theorem 6.3. *The variance $V_{\hat{\pi},A}(t)$ of any time-integrated current or flux observable $A(t)$ in the stationary state has the general upper bound, for all t*

$$V_{\hat{\pi},A}(t) \leq \langle a^2\rangle_{\hat{\pi}}\left(1 + \frac{2q}{\varepsilon}\right)t.$$

The proof involves upper bounding the moment generating function in terms of the spectral radius of $\Re(\mathbf{L})$. Using operator perturbation theory (cf. Section 4.4) this can be expanded in a power series, where each term in the series is subsequently bounded in terms of the quantities defined at the start of Section 6.2. This power series is then contracted back into an expression which corresponds to an upper bound on the limiting moment generating function, which in turn proves Theorem 6.3.

Proof. Using Cauchy-Schwarz inequality, one can write

$$\mathbb{E}_\nu\left[e^{uA(t)}\right] = \left\langle \frac{\nu}{\hat{\pi}}, e^{t\mathbf{L}_u}\underline{1}\right\rangle_{\hat{\pi}} \leq \left\|\frac{\nu}{\hat{\pi}}\right\|_{\hat{\pi}} \|e^{t\mathbf{L}_u}\|, \quad \|e^{t\mathbf{L}_u}\| := \sup_{\|f\|_{\hat{\pi}}=1} \|e^{t\mathbf{L}_u} f\|_{\hat{\pi}}.$$

Lumer-Phillips theorem ([119, Corollary 3.20, Proposition 3.23]) implies the following further upper bound

$$\|e^{t\mathbf{L}_u}\| \leq e^{t\lambda(u)}$$

where $\lambda(u) := \max\{z : z \in \mathrm{Sp}(\Re(\mathbf{L}_u))\}$.

In order to upper bound $\lambda(u)$, first we find an alternative expression for small u using perturbation theory [105] (cf. Section 4.4). Since $\mathbf{L}$ is irreducible, so is $\Re(\mathbf{L})$ and Perron-Frobenius theory implies that $\lambda(0) = 0$ is an algebraically simple eigenvalue. Recalling the definition of $\mathbf{D}^{(k)}$ we can write the real part of the tilted generator as

$$\Re(\mathbf{L}_u) = \Re(\mathbf{L}) + \sum_{k\geq 1} \frac{u^k}{k!}\mathbf{D}^{(k)}, \quad \mathbf{D}^{(k)} := \sum_{x\neq y} a_{xy}^k \Re(\mathbf{J}_{xy}).$$

With $\Re(\mathbf{L}_u)$ being an analytic perturbation of $\Re(\mathbf{L})$, perturbation theory (see [67] and references therein for details) ensures that if we can find $\alpha, \beta > 0$ such that $\|\mathbf{D}^{(k)}\| \leq \alpha\beta^{k-1}$ for $k \geq 1$, then for $|u| \leq (2\alpha\varepsilon^{-1} + \beta)^{-1}$ (ε is the spectral gap of $\Re(\mathbf{L})$) we can write

$$\lambda(u) = \sum_{k\geq 1} u^k \lambda^{(k)} \tag{6.2}$$

with

$$\lambda^{(k)} = \sum_{p=1}^{k} \lambda_p^{(k)},$$

and

$$\lambda_p^{(k)} := \frac{(-1)^p}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=k,\, \nu_i\geq 1 \\ \mu_1+\cdots+\mu_p=p-1,\, \mu_j\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \mathrm{tr}\left(\mathbf{D}^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\mathbf{D}^{(\nu_p)}\mathbf{S}^{(\mu_p)}\right)$$

where $\mathbf{S}^{(0)} = -\Pi_{\hat{\pi}}$ and $\mathbf{S}^{(\mu)} = (\Re(\mathbf{L}) + \Pi_{\hat{\pi}})^{-\mu} - \Pi_{\hat{\pi}}$. Notice that $\|\mathbf{S}^{(\mu)}\|_2 \leq \varepsilon^{-\mu}$.

Lemma 6.2 provides the estimates required for bounding the convergence radius and the coefficients of the series in Equation (6.2). We focus on $p = 1$:

$$\lambda_1^{(k)} = \frac{1}{k!} \left\langle \underline{1}, \sum_{x \neq y} a_{xy}^k \mathbf{J}_{xy} \underline{1} \right\rangle_{\hat{\pi}} = \frac{1}{k!} \sum_{x \neq y} \hat{\pi}(x) w_{xy} a_{xy}^k,$$

hence

$$\sum_{k \geq 1} u^k \lambda_p^{(k)} = \sum_{x \neq y} \hat{\pi}(x) w_{xy} (e^{u a_{xy}} - 1).$$

Let us consider $k \geq 2$ and $p \geq 2$, then

$$\begin{aligned} \left| \frac{\operatorname{tr}(\mathbf{D}^{(\nu_1)} \mathbf{S}^{(\mu_1)} \cdots \mathbf{D}^{(\nu_p)} \mathbf{S}^{(\mu_p)})}{\nu_1! \cdots \nu_p!} \right| &\leq \frac{\langle a^2 \rangle_{\hat{\pi}}}{c} \left(\frac{2q}{\varepsilon} \right)^{p-1} \left(\frac{c}{2} \right)^{k-1} \\ &\leq \frac{q \langle a^2 \rangle_{\hat{\pi}}}{\varepsilon} \left(\frac{qc}{\varepsilon} \right)^{k-2}. \end{aligned}$$

Above we used that at least one of the μ_i's is equal to 0, so that the product under the trace contains a factor $\mathbf{S}^{(0)} = -\Pi_{\hat{\pi}}$ and applied the bounds in Lemma 6.2. We further used the fact that $\nu_1! \cdots \nu_p! \geq 2^{k-p}$. One can show ([67, Section 3]) that for $k \geq 3$,

$$\sum_{p=1}^{k} \sum_{\substack{\nu_1 + \cdots + \nu_p = k, \, \nu_i \geq 1 \\ \mu_1 + \cdots + \mu_p = p-1, \, \mu_j \geq 0}} \frac{1}{p} \leq 5^{k-2} \tag{6.3}$$

and we obtain that for $k \geq 2$

$$\left| \sum_{p=2}^{k} \lambda_p^{(k)} \right| \leq \frac{q \langle a^2 \rangle_{\hat{\pi}}}{\varepsilon} \left(\frac{5qc}{\varepsilon} \right)^{k-2}.$$

Therefore for every $u \geq 0$ such that the R.H.S. of the following inequality is finite, we have

$$\begin{aligned} \lambda(u) &\leq \sum_{x \neq y} \hat{\pi}(x) w_{xy} (e^{u a_{xy}} - 1) + \frac{q \langle a^2 \rangle_{\hat{\pi}} u^2}{\varepsilon} \sum_{k \geq 0} \left(\frac{5qcu}{\varepsilon} \right)^k \\ &= \sum_{x \neq y} \hat{\pi}(x) w_{xy} (e^{u a_{xy}} - 1) + \frac{q \langle a^2 \rangle_{\hat{\pi}} u^2}{\varepsilon} \left(1 - \frac{5qc}{\varepsilon} \right)^{-1} \\ &:= \tilde{\Lambda}(u). \end{aligned}$$

Notice that the upper bound of $|\lambda(u)|$ diverges before u exits from the convergence radius of the expression in (6.2): indeed,

$$\frac{1}{5cq/\varepsilon} \leq \frac{1}{c(2q/\varepsilon + 1)} \Leftrightarrow \frac{5q}{\varepsilon} \geq \frac{2q}{\varepsilon} + 1 \Leftrightarrow \frac{2q}{\varepsilon} \geq \frac{2}{3}.$$

So far we have showed that for every $u \geq 0$ for which the R.H.S. is finite, we have

$$\begin{aligned}\mathbb{E}_\nu[e^{uA(t)}] &\leq \left\|\frac{\nu}{\hat{\pi}}\right\|_{\hat{\pi}} \exp\left(t\left(\sum_{x\neq y}\hat{\pi}(x)w_{xy}(e^{ua_{xy}}-1)\right.\right.\\ &\qquad\left.\left.+\frac{q\langle a^2\rangle_{\hat{\pi}}u^2}{\varepsilon}\left(1-\frac{5cqu}{\varepsilon}\right)^{-1}\right)\right) \qquad (6.4)\\ &= \left\|\frac{\nu}{\hat{\pi}}\right\|_{\hat{\pi}} e^{t\tilde{\Lambda}(u)},\end{aligned}$$

We can extend the bound to every non-negative u putting the R.H.S. equal to $+\infty$ after it blows up. For getting a bound also for negative u, we can repeat the same reasoning for $\tilde{A}(t) := -A(t)$ and we arrive at the expression

$$\begin{aligned}\mathbb{E}_\nu[e^{uA(t)}] \leq \left\|\frac{\nu}{\hat{\pi}}\right\|_{\hat{\pi}} \exp\left(t\left(\sum_{x\neq y}\hat{\pi}(x)w_{xy}(e^{ua_{xy}}-1)\right.\right.\\ \left.\left.+\frac{q\langle a^2\rangle_{\hat{\pi}}u^2}{\varepsilon}\left(1-\frac{5cq|u|}{\varepsilon}\right)^{-1}\right)\right)\end{aligned}$$

for every $u \in \mathbb{R}$ (again we extend the R.H.S. beyond the blow up putting it equal to $+\infty$).

Theorem 6.3 is proved by observing that since $\tilde{\Lambda}(0) = 0$ and $\tilde{\Lambda}'(0) = \langle a\rangle_{\hat{\pi}}$

$$\begin{aligned}\log(\mathbb{E}_{\hat{\pi}}[e^{uA(t)}]) &= \langle a\rangle_{\hat{\pi}}tu + \frac{1}{2}V_{\hat{\pi},A}(t)u^2 + o(u^2)\\ &\leq \langle a\rangle_{\hat{\pi}}tu + \frac{1}{2}\tilde{\Lambda}''(0)tu^2 + o(u^2).\end{aligned}$$

Hence $V_{\hat{\pi},A}(t) \leq \tilde{\Lambda}''(0)t = \langle a^2\rangle_{\hat{\pi}}\left(1+\frac{2q}{\varepsilon}\right)t$.

□

Theorem 6.3 is an upper bound on the size of fluctuations in the form of an upper bound on the variance of $A(t)$. It shows that the variance grows slower than $\langle a^2\rangle_{\hat{\pi}}(1+2q/\varepsilon)$ and that this limit is controlled by both dynamical quantities such as the maximum escape rate q as well as the additive spectral gap ε. This implies that systems where $\Re(\mathbf{L})$ reaches its steady state faster will have a smaller limit on the size of fluctuations. The upper bound on the moment generating function in Equation (6.4) can be used to upper bound the entire distribution, for finite time, which we will discuss in the next section.

6.3 CONCENTRATION INEQUALITY

Recall that we are interested in the finite-time fluctuations of the process $A(t)/t$ about $\langle a\rangle_{\hat{\pi}}$. By upper bounding the moment generating function $\forall\ u \in \mathbb{R}$ with the bound finite $\forall\ |u| < \frac{\varepsilon}{5cq}$, we can prove a concentration result with a non trivial rate function.

Theorem 6.4. *The distribution of $A(t)/t$ starting from an initial measure ν obeys a concentration bound*

$$\mathbb{P}_\nu\left(\frac{A(t)}{t} \geq \langle a\rangle_{\hat\pi} + \gamma\right) \leq C(\nu)e^{-t\tilde I(\langle a\rangle_{\hat\pi}+\gamma)}, \quad t>0,$$

where $\gamma > 0$ is the fluctuation of A away from the stationary average, and $C(\nu) := (\sum_x \nu(x)^2/\hat\pi(x))^{1/2}$ accounts for the difference between ν and the stationary $\hat\pi$, with $C(\hat\pi) = 1$. The bounding rate function can be written explicitly as

$$\tilde I\left(\langle a\rangle_{\hat\pi}+\gamma\right) = \frac{\gamma^2}{2\left(\langle k\rangle_{\hat\pi}c^2 + \frac{2q\langle a^2\rangle_{\hat\pi}}{\varepsilon} + \frac{5cq\gamma}{\varepsilon}\right)}.$$

The proof of Theorem 6.4 begins by weakening slightly the upper bound on the limiting moment generating function in Equation (6.4). This allows the Fenchel-Legendre transform of this upper bound to be performed which optimises the bound in terms of u and completes the Chernoff method (cf. Section 4.1.3).

Proof. With a further elementary estimate for the moment generating function, we derive an explicit concentration bound. Notice that, using that $\hat\pi(x)w_{xy} \geq 0$, $u > 0$ and $c = \max_{x\neq y}|a_{xy}|$, one of the terms appearing in the upper bound $\tilde\Lambda(u)$ can be upper bounded in the following way:

$$\begin{aligned}\sum_{x\neq y}\hat\pi(x)w_{xy}(e^{ua_{xy}} - ua_{xy} - 1) &= \sum_{x\neq y}\hat\pi(x)w_{xy}\sum_{k\geq 2}\frac{(ua_{xy})^k}{k!}\\ &\leq \sum_{x\neq y}\hat\pi(x)w_{xy}\sum_{k\geq 2}\frac{(uc)^k}{k!}\\ &= \underbrace{\left(\sum_{x\neq y}\hat\pi(x)w_{xy}\right)}_{=\langle k\rangle_{\hat\pi}}(e^{cu}-cu-1).\end{aligned} \tag{6.5}$$

Hence for every $u \geq 0$ we can write

$$\mathbb{E}_\nu[e^{u(A(t)-t\langle a\rangle_{\hat\pi})}] \leq \left\|\frac{\nu}{\hat\pi}\right\|_{\hat\pi}\exp\left(t\left(\langle k\rangle_{\hat\pi}(e^{cu}-cu-1) + \frac{q\langle a^2\rangle_{\hat\pi}u^2}{\varepsilon}\left(1-\frac{5cqu}{\varepsilon}\right)^{-1}\right)\right).$$

Recall that the Fenchel-Legendre transform of a function $h : \mathbb{R} \mapsto \mathbb{R}\cup\{-\infty,+\infty\}$ is given by

$$h^*(\gamma) := \sup_{u\in\mathbb{R}}\{u\gamma - h(u)\}. \tag{6.6}$$

Using Chernoff's bound, we obtain that for every $\gamma > 0$

$$\mathbb{P}_\nu\left(\frac{A(t)}{t} \geq \langle a\rangle_{\hat\pi}+\gamma\right) \leq \left\|\frac{\nu}{\hat\pi}\right\|_{\hat\pi}\exp\left(-t\underbrace{\sup_{u\in\mathbb{R}}\{\gamma u - h_1(u) - h_2(u)\}}_{=:(h_1+h_2)^*(\gamma)}\right),$$

where

$$h_1(u) := \begin{cases} 0 & u < 0 \\ \langle k \rangle_{\hat{\pi}} (e^{cu} - cu - 1) & u \geq 0 \end{cases}$$

and

$$h_2(u) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{q \langle a^2 \rangle_{\hat{\pi}} u^2}{\varepsilon} \left(1 - \frac{5qcu}{\varepsilon}\right)^{-1} & \text{if } 0 \leq u < \frac{\varepsilon}{5cq} \\ +\infty & \text{o.w.} \end{cases}.$$

In order to simplify notation, we derive the result for the general function

$$h_2(u) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{u^2}{\zeta - \xi u} & \text{if } 0 \leq u < \zeta/\xi \\ +\infty & \text{o.w.} \end{cases}$$

for $\zeta, \xi > 0$. From the definition of the Fenchel-Legendre transform (Equation (6.6)) it is easy to see that

$$h_1^*(\gamma) = \begin{cases} +\infty & \text{if } \gamma < 0 \\ \langle k \rangle_{\hat{\pi}} g_1 \left(\frac{\gamma}{\langle k \rangle_{\hat{\pi}} c} \right) & \text{o.w.} \end{cases},$$

$$h_2^*(\gamma) = \begin{cases} +\infty & \text{if } \gamma < 0 \\ \frac{\zeta \gamma^2}{2} g_2 \left(\xi \gamma \right) & \text{o.w.} \end{cases},$$

where $g_1(\gamma) = (1+\gamma)\log(1+\gamma) - \gamma \geq \gamma^2/2(1+\gamma/3)$ and $g_2(\gamma) = (1 + \gamma/2 + \sqrt{\gamma+1})^{-1} \geq (2+\gamma)^{-1}$. We can use Moreau-Rockafellar formula ([120, Theorem 16.4]) to obtain

$$\begin{aligned} (h_1 + h_2)^*(\gamma) &= \inf\{h_1^*(\gamma_1) + h_2^*(\gamma_2) : \gamma_1 + \gamma_2 = \gamma,\ \gamma_1, \gamma_2 \in \mathbb{R}\} \\ &\geq \inf \left\{ \frac{\gamma_1^2}{2\left(\langle k \rangle_{\hat{\pi}} c^2 + \frac{c\gamma_1}{3}\right)} + \frac{\zeta \gamma_2^2}{2(2 + \xi\gamma_2)} : \gamma_1 + \gamma_2 = \gamma,\ \gamma_1, \gamma_2 \geq 0 \right\}. \end{aligned}$$

We can use the fact that $\gamma_1^2/a + \gamma_2^2/d \geq (\gamma_1 + \gamma_2)^2/(a+d)$ for non-negative γ_1, γ_2 and positive a, d to obtain

$$\begin{aligned} (h_1 + h_2)^*(\gamma) &\geq \inf \left\{ \frac{\gamma^2}{2\left(\langle k \rangle_{\hat{\pi}} c^2 + \frac{c\gamma_1}{3} + \frac{2}{\zeta} + \frac{\xi}{\zeta}\gamma_2\right)} \right. \\ &\qquad \left. : \gamma_1 + \gamma_2 = \gamma,\ \gamma_1, \gamma_2 \geq 0 \right\} \\ &= \frac{\gamma^2}{2\left(\langle k \rangle_{\hat{\pi}} c^2 + \frac{2}{\zeta} + \max\left\{\frac{c}{3}, \frac{\xi}{\zeta}\right\}\gamma\right)}. \end{aligned}$$

Hence we obtain the following:

$$\mathbb{P}_\nu\left(\frac{A(t)}{t} \geq \langle a\rangle_{\hat{\pi}} + \gamma\right) \leq \left\|\frac{\nu}{\hat{\pi}}\right\|_{\hat{\pi}} \exp\left(-t\frac{\gamma^2}{2}\left(\langle k\rangle_{\hat{\pi}} c^2 + \frac{2c\langle a^2\rangle_{\hat{\pi}}}{\varepsilon} + \frac{5cq\gamma}{\varepsilon}\right)^{-1}\right). \tag{6.7}$$

Let us introduce the notation

$$\tilde{I}(x) = \frac{(x - \langle a\rangle_{\hat{\pi}})^2}{2\left(\langle k\rangle_{\hat{\pi}} c^2 + \frac{2c\langle a^2\rangle_{\hat{\pi}}}{\varepsilon} + \frac{5cq(x-\langle a\rangle_{\hat{\pi}})}{\varepsilon}\right)}$$

where $x := \langle a\rangle_{\hat{\pi}} + \gamma$. It is easy to see that the bound in Equation (6.7) implies that

$$\tilde{I}(x) \leq I(x), \quad \forall x > \langle a\rangle_{\hat{\pi}}$$

(we recall that I is the rate function of $A(t)/t$): indeed, together with the definition of large deviation principle, it allows us to write

$$\begin{aligned}\inf_{x'>x} I(x') &\geq \liminf_{t\to+\infty} -\frac{1}{t}\log \mathbb{P}_\nu\left(\frac{A(t)}{t} \geq x\right) \\ &\geq \tilde{I}(x).\end{aligned} \tag{6.8}$$

The inequality holds for $x' \geq x$ due to the continuity of $\tilde{I}$ and $\inf_{x'\geq x} I(x') = I(x)$ because I is non-decreasing for $x \geq \langle a\rangle_{\hat{\pi}}$.

□

We remark the upper bound on the moment generating function was made weaker in Equation (6.5) by replacing the weights a_{xy} with their maximal value c, so we could perform the optimisation. If one wishes to numerically Fenchel-Legendre transform rather than obtain an explicit concentration inequality this step can be skipped for a tighter bound.

Theorem 6.4 bounds the distribution of $A(t)$ for finite t in terms of the simpler quantities q, ε, average dynamical activity $\langle k\rangle_{\hat{\pi}}$ and maximum amplitude of observable coefficients c. These would ideally be known or easier to estimate than a computation of the probability distribution itself.

6.4 INVERSE THERMODYNAMIC UNCERTAINTY RELATIONS

The most direct use of TURs is in bounding the precision for estimating a current from its time-average over a trajectory in a non-equilibrium stationary state (NESS) $\hat{\pi}$. From the standard application of the TUR together with the "inverse TUR" Theorem 6.3 we can bound the relative error (Equation (3.4)) from *below and above*

$$\frac{2}{\Sigma_{\hat{\pi}}} \leq \epsilon_J^2 \leq \frac{\langle a^2\rangle_{\hat{\pi}}}{\langle a\rangle_{\hat{\pi}}^2}\left(1 + \frac{2q}{\varepsilon}\right) \tag{6.9}$$

with an equivalent statement for the activity TUR upper bounding the relative error ϵ_A^2 of observables which are not necessarily currents. An example

where the relation in Equation (6.9) could be exploited are the molecular motors described in [8, 36, 37]. If one is interested in estimating the entropy production of such a system, the inverse TUR could be used to compute a lower bound for $\Sigma_{\hat{\pi}}$, which would be useful if the currents used to measure the relative error directly are inaccessible. The physical meaning of our new upper bound can be understood from its two factors. In contrast to the TUR, the first factor in the R.H.S. of Equation (6.9) contains information about the current A of interest via the static variance, $\langle a^2 \rangle_{\hat{\pi}}$. This is measurable, being the stationary mean of $S_A(t) = \sum_{x \neq y} a_{xy}^2 K_t(x, y)$: given a (time-asymmetric) current A, there is an associated *symmetric flux* S_A whose stationary average encodes the interplay between the localisation properties of the dynamics (which jumps $x \leftrightarrow y$ have larger rates $\hat{\pi}(x)w_{xy}$ and $\hat{\pi}(y)w_{yx}$, and how these are spread among all possible transitions), and the transitions relevant for A (which $x \leftrightarrow y$ have a larger $|a_{xy}|$ giving rise to larger variations in A). The first factor therefore quantifies the intuition that if larger variations of A are produced by the most (resp. least) active jumps, we can expect A to have large (resp. small) fluctuations.

The second factor in the R.H.S. of Equation (6.9) encodes overall properties of the dynamics via the ratio q/ε. The symmetrised generator corresponds to the unique equilibrium dynamics that shares key relevant features (steady state and dynamical activity) with the original dynamics [121], and its spectral structure is able to upper bound fluctuations *at all times*. A relevant case is that of dynamics with several mesostates (phases) with frequent jumps within and rare jumps between, implying metastable behaviour with large fluctuations for empirical fluxes. This is captured by q/ε, with q large due to the speed of the intra-state dynamics, and ε small (and vanishing at a first-order phase transition as shown in Section 6.6.2). Thus the second factor in the R.H.S. of Equation (6.9) quantifies the fact that fluctuations in a time-integrated current are limited by the degree of separation of timescales in the dynamics.

6.5 CLOSURE OF SPECTRAL GAP

Since the upper bound in Theorem 6.3 requires the spectral gap of $\Re(\mathbf{L})$, when studying the behaviour of the process near a phase transition, we need to check that $\varepsilon \to 0$ when the gap of the original process described by $\mathbf{L}$ closes. Physically, if the gap of $\mathbf{L}$ vanishes, i.e. the process becomes reducible, this section shows that the equivalent equilibrium process with generator $\Re(\mathbf{L})$ also becomes reducible.

Let $\tilde{\lambda}_2$ (λ_2) be the eigenvalue of $\mathbf{L}$ (resp. $\Re(\mathbf{L})$) which is non-zero and has the biggest real part; we recall that $\varepsilon = -\lambda_2$. If one shows that

$$\varepsilon \leq -\Re(\tilde{\lambda}_2) \Leftrightarrow \lambda_2 \geq \Re(\tilde{\lambda}_2), \tag{6.10}$$

then if $\mathbf{L}$ undergoes a first-order phase transition, i.e. $\tilde{\lambda}_2 \to 0$, one has that $\varepsilon \to 0$ too. Let us show Equation (6.10): it is easy to see that the linear space of centred functions on the state space, i.e. $V = \{1\}^{\perp} = \{f : E \to \mathbb{C}\}$, is

invariant for both $\mathbf{L}$ and $\mathbf{L}^\dagger$, hence also for $\Re(\mathbf{L})$. With an abuse of notation, all the operators appearing in the following equation will be considered as operators acting on V: one has

$$e^{\Re(\tilde{\lambda}_2)} = |e^{\tilde{\lambda}_2}| \leq \|e^{\mathbf{L}}\|_{\hat{\pi}} \leq e^{\|\Re(\mathbf{L})\|_{\hat{\pi}}} = e^{\lambda_2}. \tag{6.11}$$

The first inequality follows from the fact that the spectral radius is less or equal than the norm of an operator and the second inequality follows from Lumer-Phillips theorem. Equations (6.10) and (6.11) are equivalent.

6.6 EXAMPLES

6.6.1 *Comparison to thermodynamic uncertainty relations*

As an illustration of Theorems 6.3 and 6.4 we consider the fluctuations of currents in the 4-state model of Ref. [34]. This allows us to compare our upper bound with the known TUR.

The network of elementary transitions is shown in the inset of Figure 6.1. The rates are as in Ref. [34], $w_{12} = 3$, $w_{13} = 10$, $w_{14} = 9$, $w_{21} = 10$, $w_{23} = 1$, $w_{24} = 2$, $w_{31} = 6$, $w_{32} = 4$, $w_{34} = 1$, $w_{41} = 7$, $w_{42} = 9$ and $w_{43} = 5$. A current is defined by the values of the six coefficients $a_{x>y}$ which we take in the range $a_{xy} \in [-1, 1]$. To perform the analysis, we construct a mesh across the space of current observables $\mathbb{T} = [-1, 1]^6$ discretised with spacing 10^{-1}, with each point corresponding to a different current.

Figure 6.1 shows the bounds for the long-time limit rate function $I(A/t)$ for one such current $A \in \mathbb{T}$. The (full/black) curve is the exact rate function. It is calculated from the "tilted" generator $\mathbf{L}_u$ (cf. Equation (3.3)) as follows (cf. Section 3.4.2):

1. the MGF of A is
$$\mathbb{E}_{\hat{\pi}}\left[e^{uA(t)}\right] = \langle \hat{\pi}, e^{t\mathbf{L}_u}, \underline{1} \rangle,$$
where $\underline{1}$ is the "flat state" vector with each entry 1;
2. at long times $\mathbb{E}_{\hat{\pi}}\left[e^{uA(t)}\right] \asymp e^{t\Lambda(u)}$, where the limiting logarithmic moment generating function $\Lambda(u)$ is the largest eigenvalue of $\mathbf{L}_u$;
3. the rate function is obtained via Legendre transform
$$I(a) = \sup_{u \in \mathbb{R}} \{ua - \Lambda(u)\}.$$

The (dotted/blue) curve in Figure 6.1 is the usual TUR using the entropy production [34]. The (dot-dashed/pink) curve is the alternative TUR (also known as the KUR) which instead of $\Sigma_{\hat{\pi}}$ uses the average dynamical activity, $\langle k \rangle_{\hat{\pi}} = \sum_{x \neq y} \hat{\pi}(x) w_{xy}$ [48, 122]. Both these curves are above the true rate function, thus providing the usual lower bounds on the size of the fluctuations of A. The (dashed/red) curve represents an inverse TUR which upper bounds the size of fluctuations of A at all orders, cf. Theorem 6.4.

Figure 6.2 shows the bounds (6.9) on the precision error Equation (3.4), for all currents in $\mathbb{T}$, both at finite and infinite t. The (full/black) curves are the exact error ϵ_A^2, obtained from Lemma 6.1. The errors are plotted rank ordered by their value at $t = \infty$. The (dotted/blue) lines are the lower bounds from the TUR at either finite [100] or infinite [37] times. The (dot-dashed/pink) lines are the activity TUR, where in the L.H.S. of (6.9) $\Sigma_{\hat{\pi}}$ is replaced by $2\langle k \rangle_{\hat{\pi}}$. As the TURs do not depend on the details of the current that they bound, these curves are constant. Figure 6.2 also shows the inverse TUR from the R.H.S. of (6.9) as (full/red) curves. This gives an upper bound to the error. The inverse TUR contains information about the specific current through its static average and second moment and it tracks the change in shape of the exact error: in many instances the ratio of the relative value of the upper bound to the error is smaller than that of the error to the lower bound.

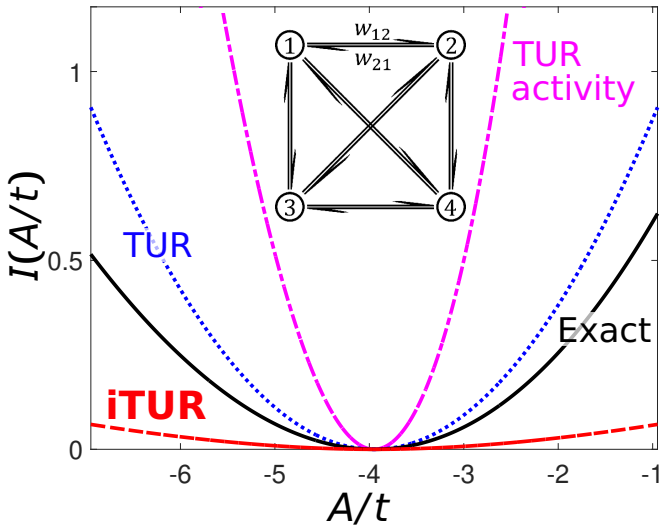

Figure 6.1: Upper bound on current fluctuations: the (full/black) curve shows the exact rate function $I(A/t)$ for the current defined by $a_{12} = 0.9$, $a_{13} = -0.9$, $a_{14} = -0.9$, $a_{23} = 0.9$, $a_{24} = -0.9$, $a_{34} = 0.9$. The rate function is upper bounded by the TURs: the (dotted/blue) curve is the standard TUR using the entropy production, while the (dot-dashed/pink) is the TUR with the dynamical activity. The (dashed/red) curve is the inverse TUR: it lower bounds the rate function, corresponding to an upper bound on fluctuations at all orders. [We plot the iTUR from a parametric Legendre transform of $\tilde{\Lambda}(u)$ in Equation (6.4) to avoid the approximation used to obtain Theorem 6.4; however, the explicit expression for the rate function given in the Theorem gives a near-identical bound.] Inset: sketch of the 4-state model.

6.6.2 *Analysis of model with phases*

As explained above, the inverse TUR captures the increase of fluctuations close to a dynamical phase transition via its dependence on q/ε, see Equa-

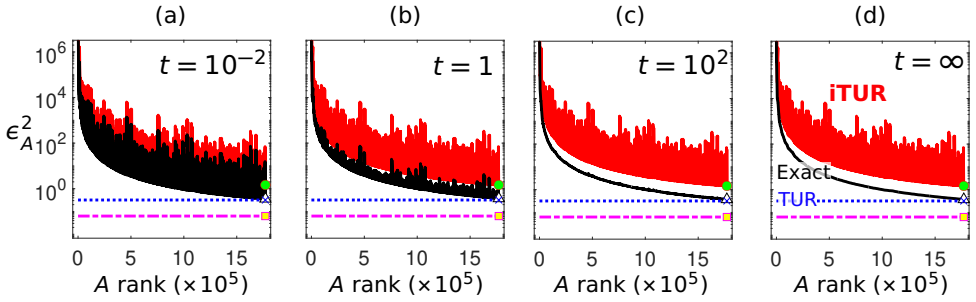

Figure 6.2: Lower and upper bounds on the estimation error: (a) Relative error ϵ_A^2 for estimating a current A from a trajectory of length $t = 10^{-2}$ in the NESS of the model of Figure 6.1. We show results for 20^6 different currents $A \in \mathbb{T}$. The (full/black) curve is the exact error. The standard TUR, (dotted/blue) line, and the activity TUR, (dot-dashed/pink) line, provide lower bounds to the error which are independent of A. The inverse TUR, (dashed/red) curve, gives an upper bound to the error which varies with A. (b-d) Same for times $t = 1, 10^2, \infty$, respectively. The data in all panels is ranked according to decreasing values of the error at $t = \infty$. For comparison, the A corresponding to entropy production is shown according to the same ranking: TUR bound (green circle), exact (white triangle), TUR (blue cross), activity TUR (yellow square).

tion (6.9). Figure 6.3 illustrates this in a six-state model with two competing meso-states: as the spectral gap closes with decreasing ω, the system gets trapped for longer times in each metastable phase, giving rise to larger fluctuations of currents with different mean values at stationarity in the two phases. The inverse TUR tracks this growth in the estimation error, while the TURs do not.

In the second example of this chapter, the set of states can be split into two subsets corresponding to the two phases at the phase transition: $E_1 = \{1,2,3\}$ and $E_2 = \{4,5,6\}$. Such a splitting, induces the following block decomposition of the generator:

$$\mathbf{L} = \begin{pmatrix} \widetilde{\mathbf{L}} & 0 \\ 0 & \widetilde{\mathbf{L}} \end{pmatrix} + \omega \begin{pmatrix} -\mathbf{1}_3 & \mathbf{1}_3 \\ \mathbf{1}_3 & -\mathbf{1}_3 \end{pmatrix},$$

$$\widetilde{\mathbf{L}} = \begin{pmatrix} -(\lambda+\kappa) & \kappa & \lambda \\ \lambda & -(\lambda+\kappa) & \kappa \\ \kappa & \lambda & -(\lambda+\kappa) \end{pmatrix}.$$

$\mathbf{1}_3$ denotes the 3×3 identity matrix, while $\widetilde{\mathbf{L}}$ is the generator corresponding to the dynamic in the two phases at the phase transition. By the symmetry of the model one can easily see that $\widetilde{\mathbf{L}}$ has the uniform distribution on three states as unique invariant measure. Notice that the perturbation in ω annihilates every measure which is the same on E_1 and E_2 (with the correspondence $1 \leftrightarrow 4$, $2 \leftrightarrow 5$ and $3 \leftrightarrow 6$), hence the uniform distribution on six states is the unique invariant measure, independent from ω. As a

consequence, $\langle a \rangle_{\hat{\pi}}$ and $\langle a^2 \rangle_{\hat{\pi}}$ for the current we decided to study ($A(t) = K_t(4,5) - K_t(5,4)$) do not depend on ω as well:

$$\langle a \rangle_{\hat{\pi}} = (\kappa - \lambda)/6, \qquad \langle a^2 \rangle_{\hat{\pi}} = (\kappa + \lambda)/6.$$

This simplifies the analysis, reducing the study of the upper bound to the description of the behaviour of q/ε. Furthermore, notice that the expected value per unit of time at stationarity of the restriction of the current to E_1 is equal to 0, while restricted to E_2 is equal to $(\kappa - \lambda)/6$: this is the reason why its fluctuations are going to increase approaching the phase transition ($\omega \to 0$).

The maximum escape rate is equal to $q = \lambda + \kappa + \omega$. In order to compute the additive spectral gap ε, first we need to find the symmetrised generator, which reads

$$\Re(\mathbf{L}) = \begin{pmatrix} \Re(\widetilde{\mathbf{L}}) & 0 \\ 0 & \Re(\widetilde{\mathbf{L}}) \end{pmatrix} + \omega \begin{pmatrix} -\mathbf{1}_3 & \mathbf{1}_3 \\ \mathbf{1}_3 & -\mathbf{1}_3 \end{pmatrix}.$$

Using Schur determinant lemma [123, p. 4], one obtains that

$$\begin{aligned} \det(\Re(\mathbf{L}) - t\mathbf{1}_6) &= \det((\Re(\widetilde{\mathbf{L}}) - (t+\omega)\mathbf{1}_3)^2 - \omega^2 \mathbf{1}_3) \\ &= \det(\Re(\widetilde{\mathbf{L}}) - (t+2\omega)\mathbf{1}_3) \det(\Re(\widetilde{\mathbf{L}}) - t\mathbf{1}_3), \end{aligned}$$

which means that the spectrum of $\Re(\mathbf{L})$ is equal to the union of the spectrum of $\Re(\widetilde{\mathbf{L}})$ and its translation of -2ω. Therefore, if ω is small enough, ε is equal to 2ω. Hence $q/\varepsilon \approx (\lambda + \kappa)/\omega$ for $\omega \ll 1$.

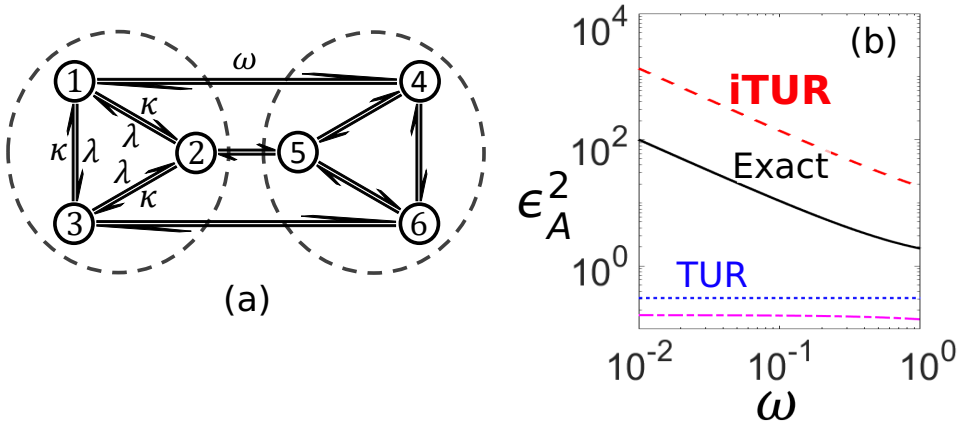

Figure 6.3: Inverse TUR near metastability: (a) Markov network of 6-state model. The dashed circles indicate the two competing metastable phases at small ω. We consider $\lambda = 5, \kappa = 1$ while varying ω. (b) Estimation error ϵ_A^2 for the current between states 4 and 5, $A(t) = K_t(4,5) - K_t(5,4)$, as a function of ω. The (full/black) curve is the exact value, upper bounded by the inverse TUR (dashed/red), and lower bounded by the standard TUR (dotted/blue) and the activity TUR (dot-dashed/pink). Fluctuations increase with increasing metastability, which is tracked by the inverse TUR but not by the standard TUR.

We have proven a general class of upper bounds on the size of fluctuations of dynamical observables which complement the lower bounds provided by TURs. Lemma 6.1 and Theorems 6.3 and 6.4 apply to fluctuations of all orders at all times. In contrast to the standard TUR, our bounds encode details of the current of interest. Having *both* upper and lower bounds is necessary to limit the range of estimation errors.

There are many possible extensions and refinements. We focused on continuous-time Markov chains, but analogous bounds should be obtainable for discrete time dynamics. Our bounds have as input the spectral gap of the (symmetrised) generator, which for many-body systems can be estimated from time correlations [124, 125]. Further approximations may also allow to formulate the inverse TURs in terms of operationally accessible quantities. The classical results here will have a corresponding generalisation for open quantum dynamics by exploiting generalisations of concentration bounds to the quantum case, see e.g. [65].

7

CLASSICAL UPPER BOUNDS ON FLUCTUATIONS OF FIRST PASSAGE TIMES

The Cramér-Chernoff approach discussed in Chapter 4 can be used not just in the fixed-time ensemble as we did in Chapter 6 but also in the fixed-observable ensemble to obtain bounds for first passage times. We refer the reader to Section 2.6 Chapter 2 for background on this dual ensemble when one is instead interested in first passage times (FPTs). Many of the results in this chapter have a counterpart for quantum Markov processes which will be discussed in Chapter 8. In particular, we consider FPTs corresponding to counting observables, which we recall from Section 2.5.1, are time-additive observables of trajectories which are non-decreasing (in contrast to currents), cf. [48].

We begin in Section 7.1 by by computing the form of the moment generating function. Here we state Lemma 7.2 which gives us the analytical form of the variance of FPTs for counting observables. After that, in Section 7.2, we prove that the sequence of FPTs satisfies a large deviation principle [44] and we provide an expression for the rate function. Then, we derive in Section 7.3 a concentration inequality of the FPT for the dynamical activity [11, 27], and a tail bound for FPTs for generic counting observables (Section 7.4). As subsequent propositions, with each concentration inequality we also provide an upper bound on the variance of the FPT. These complement the thermodynamic uncertainty relations (TURs) for first passage times [48, 50] and are the fixed-observable equivalent of our inverse TURs (cf. Chapter 6). We illustrate our results with simple models in Section 7.5; in particular, we discuss the behaviour of the bound for the FPT corresponding to the dynamical activity when the system is at conditions of metastability, i.e., near a first-order phase crossover.

7.1 MOMENT GENERATING FUNCTION

Recall from Section 2.6 that we denote $T_{\mathfrak{E}}(k)$ to be the first passage time for the dynamical activity of the process to reach k jumps. Using the properties of the holding times described in Section 2.6 Chapter 2, one finds that the moment generating function (MGF) of $T_{\mathfrak{E}}(k)$ is well defined for $u < d := \min_x R_x$ and is given by

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{E}}(k)}] = \left\langle \nu, \left(\frac{\mathbf{R}}{\mathbf{R}-u}\mathbf{P}\right)^k \underline{1} \right\rangle = \left\langle \nu, \left(\frac{\mathbf{1}}{\mathbf{R}-u}\mathbf{W}\right)^k \underline{1} \right\rangle. \tag{7.1}$$

An analogous formula can be found for every FPT of the type in Definition (2.4). First of all, it is useful to consider the following splitting of the evolution generator $\mathbf{L}$ [48]

$$\mathbf{L} = \mathbf{W}_1 + \underbrace{\mathbf{W}_2 - \mathbf{R}}_{\mathbf{L}_\infty}, \tag{7.2}$$

where $\mathbf{W}_1$ holds the rates of transitions in $\mathfrak{A}$ and $\mathbf{W}_2$ the rates of transitions not in $\mathfrak{A}$. Notice that we can always write the first passage time corresponding to the level k as a sum of times between subsequent jumps in $\mathfrak{A}$:

$$T_{\mathfrak{A}}(k) = \sum_{i=1}^{k} s_i, \quad s_i := T_{\mathfrak{A}}(i) - T_{\mathfrak{A}}(i-1).$$

The process $Y = (y_0, \dots, y_k, \dots)$ determined by the state of the system at the sequence of times $\{T_{\mathfrak{A}}(k)\}_{k=0}^{+\infty}$ is a discrete-time Markov process with transition matrix given by

$$\mathbf{Q} := -\frac{\mathbf{1}}{\mathbf{L}_\infty}\mathbf{W}_1. \tag{7.3}$$

The state space of Y is $\{y \in E : \exists x \in E : (x,y) \in \mathfrak{A}\}$. Physically, Y is the process which describes the state of the system after competing a jump in $\mathfrak{A}$. Indeed, using that $\mathbf{L}_\infty = \mathbf{W}_2 - \mathbf{R}$, we can write

$$-\frac{\mathbf{1}}{\mathbf{L}_\infty} = \frac{\mathbf{1}}{\mathbf{R} - \mathbf{W}_2} = \frac{\mathbf{1}}{\mathbf{1} - \mathbf{R}^{-1}\mathbf{W}_2}\frac{\mathbf{1}}{\mathbf{R}} = \sum_{k\geq 0}\left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^k \frac{\mathbf{1}}{\mathbf{R}}, \tag{7.4}$$

and therefore

$$\mathbf{Q} = \sum_{k\geq 0}\left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^k \frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_1. \tag{7.5}$$

Since $\mathbf{R}^{-1}\mathbf{W}_1$ and $\mathbf{R}^{-1}\mathbf{W}_2$ are the sub-Markov operators that encode the probabilities of jumps which do and do not, respectively, belong to $\mathfrak{A}$, Equation (7.5) expresses the fact that the probability of the jump $x \to y$ for the process Y is obtained by summing up the probabilities of all possible trajectories of the jump process associated to X that start in x, arrive in a state z such that $(z,y) \in \mathfrak{A}$ by using only jumps in $\mathfrak{A}^C$, and then jump from z to y. Integrating over all such possible paths, we show in Lemma 7.1 that for every $u < \overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_\infty)\}$, the MGF of $T_{\mathfrak{A}}(k)$ can be written as

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}] = \left\langle \nu, \left(\frac{\mathbf{L}_\infty}{u + \mathbf{L}_\infty}\mathbf{Q}\right)^k \underline{1}\right\rangle. \tag{7.6}$$

We remark that for suitable choices of initial distributions and for finite k's, $\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}]$ might be well defined even for some values of u bigger or equal than $\overline{\lambda}$; nevertheless we will see in Theorem 7.4 that in the large k limit, the only values which play a non-trivial role are $u < \overline{\lambda}$.

For the case of dynamical activity, we have already mentioned that $\mathbf{Q} = \mathbf{P}$ is irreducible; more generally, $\mathbf{Q}$ is only irreducible on the subspace $\{y \in E : \exists x \in E : (x, y) \in \mathfrak{A}\}$, the complement of which is transient. Indeed, $\mathbf{Q}$ admits as unique invariant measure

$$\varphi = \frac{\hat{\pi}\mathbf{W}_1}{\langle \hat{\pi}, \mathbf{W}_1 \underline{1} \rangle}, \tag{7.7}$$

which, in general, is not fully supported.

From the expression of the moment generating function, using standard theory (see Theorem 7.4) one obtains that under $\mathbb{P}_\nu$ (for every initial law ν) the following convergence holds true almost surely:

$$\frac{1}{k} T_{\mathfrak{A}}(k) \xrightarrow[k \to +\infty]{} \langle t_{\mathfrak{A}} \rangle := \left\langle \varphi, -\frac{\mathbf{1}}{\mathbf{L}_\infty} \underline{1} \right\rangle \quad \text{a.s..} \tag{7.8}$$

where $t_{\mathfrak{A}}$ denotes the process $T_{\mathfrak{A}}(k)/k$ and $\langle t_{\mathfrak{A}} \rangle$ its asymptotic average.
Lemma 7.1 below ensures that the expressions appearing in (7.3), (7.5) and (7.6) are well defined and that the identities are true. Before stating the lemma, we need to recall a few notions that will also be useful in the rest of the chapter. Given a matrix $\mathbf{A} \in M_d(\mathbb{C})$, the spectral radius of $\mathbf{A}$ is defined as

$$r(\mathbf{A}) := \max\{|z| : z \in \mathrm{Sp}(\mathbf{A})\}.$$

The spectral radius is fundamental in studying the convergence of the geometric series $\sum_{k \geq 0} \mathbf{A}^k$, since Gelfand's formula states that for an operator norm $\| \cdot \|$

$$\lim_{k \to +\infty} \|\mathbf{A}^k\|^{\frac{1}{k}} = r(\mathbf{A}).$$

Therefore, if $r(\mathbf{A}) < 1$, the series converges.

Lemma 7.1. *The following statements hold true:*

1. $\overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_\infty)\} > 0$, *hence* $\mathbf{L}_\infty$ *is invertible;*
2. $r(\mathbf{R}^{-1}\mathbf{W}_2) < 1$, *therefore* $\sum_{k \geq 0} \mathbf{S}^k$ *is well defined with* $\mathbf{S} = \mathbf{R}^{-1}\mathbf{W}_2$ *and one has*
$$-\frac{\mathbf{1}}{\mathbf{L}_\infty} = \sum_{k \geq 0} \left(\frac{\mathbf{1}}{\mathbf{R}} \mathbf{W}_2\right)^k \frac{\mathbf{1}}{\mathbf{R}};$$
3. *for every* $u < \overline{\lambda}$, *one has*
$$\mathbb{E}_\nu[e^{u T_{\mathfrak{A}}(k)}] = \left\langle \nu, \left(\frac{\mathbf{L}_\infty}{u + \mathbf{L}_\infty} \mathbf{Q}\right)^k \underline{1} \right\rangle;$$
4. $\left\|\mathbf{L}_\infty^{-1}\right\|_{\infty \to \infty}^{-1} \leq \overline{\lambda}$.

Proof. 1. Note that $\mathbf{L}_\infty$ generates a sub-Markov semigroup $e^{t\mathbf{L}_\infty}$. By Perron-Frobenius theory (cf. Theorem 5.5) we know that

$$r(e^{t\mathbf{L}_\infty}) = e^{-t\overline{\lambda}} \Leftrightarrow \overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_\infty)\}.$$

We prove this by contradiction. Suppose $\overline{\lambda} = 0$. Then there exists a non-zero non-negative function $f : E \to [0, +\infty)$ such that $\mathbf{L}_\infty f = 0$. We then have

$$\mathbf{L}f = \mathbf{L}_\infty f + \mathbf{W}_1 f = \mathbf{W}_1 f \geq 0.$$

Therefore one has

$$0 = \langle \hat{\pi}, \mathbf{L}f \rangle = \langle \hat{\pi}, \mathbf{W}_1 f \rangle,$$

which implies that $\mathbf{L}f = \mathbf{W}_1 f = 0$ because $\hat{\pi}$ has full support. Since $\mathbf{L}$ is irreducible, $f = \alpha \underline{1}$ for some non-negative α, however $\alpha \mathbf{W}_1 \underline{1} = 0$ implies that α and therefore f are 0 (it follows from the positivity of $\mathbf{W}_1$). We came to a contradiction, which proves that $\overline{\lambda} > 0$. Therefore $\mathrm{Sp}(\mathbf{L}_\infty) \subset \{z \in \mathbb{C} : \Re(z) \leq -\overline{\lambda} < 0\}$ and $\mathbf{L}_\infty$ is invertible.

2. The proof is similar to that of point 1. Notice that $\mathbf{R}^{-1}\mathbf{W}_2$ is a sub-Markov transition kernel and let $r = r(\mathbf{R}^{-1}\mathbf{W}_2)$ such that $r \in [0, 1]$, and the corresponding non-zero non-negative eigenvector $f : E \to [0, +\infty)$. Suppose that $r = 1$, then we can write

$$(\mathbf{P} - \mathbf{1})f = \mathbf{R}^{-1}\mathbf{W}_1 f + (\mathbf{R}^{-1}\mathbf{W}_2 - \mathbf{1})f = \mathbf{R}^{-1}\mathbf{W}_1 f,$$

therefore

$$0 = \langle \pi, (\mathbf{P} - \mathbf{1})f \rangle = \langle \pi, \mathbf{R}^{-1}\mathbf{W}_1 f \rangle,$$

which implies that $\mathbf{R}^{-1}\mathbf{W}_1 f = (\mathbf{P} - \mathbf{1})f = 0$ from which it follows that $f = 0$ as before, hence a contradiction. What we have proved so far shows that the derivation of Equation (7.5) is correct.

3. Using Dyson expansion, one can see that

$$\mathbb{P}_\nu(T_{\mathfrak{A}}(k) \leq t) = \int_{\sum_{i=1}^k t_i \leq t} \langle \nu, e^{t_1 \mathbf{L}_\infty} \mathbf{W}_1 e^{t_2 \mathbf{L}_\infty} \cdots e^{t_k \mathbf{L}_\infty} \mathbf{W}_1 \underline{1} \rangle dt_1 \cdots dt_k.$$

Therefore, if $u < \overline{\lambda}$, then

$$-\frac{\mathbf{1}}{u + \mathbf{L}_\infty} = \int_0^{+\infty} e^{t(u + \mathbf{L}_\infty)} dt \tag{7.9}$$

and we get

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}] = \left\langle \nu, \left(-\frac{\mathbf{1}}{u + \mathbf{L}_\infty} \mathbf{W}_1 \right)^k \underline{1} \right\rangle.$$

4. The spectral mapping theorem implies that $\mathrm{Sp}(\mathbf{L}_\infty^{-1}) = \{z^{-1} : z \in \mathrm{Sp}(\mathbf{L}_\infty)\}$, therefore one has that

$$\left\| \frac{\mathbf{1}}{\mathbf{L}_\infty} \right\|_{\infty \to \infty} \geq r(\mathbf{L}_\infty^{-1}) \geq \frac{1}{\overline{\lambda}} \quad \Leftrightarrow \quad \left\| \frac{\mathbf{1}}{\mathbf{L}_\infty} \right\|_{\infty \to \infty}^{-1} \leq \overline{\lambda}.$$

□

Loosely speaking, items 1 and 2 hold true because $\mathbf{L}_\infty$ and $\mathbf{R}^{-1}\mathbf{W}_2$ are the counterparts of $\mathbf{R}$ and $\mathbf{P}$, respectively, obtained by considering a restricted set of jumps in the original irreducible Markov process.

The next lemma gives us the form of the variance of first passage times for counting observables.

Lemma 7.2. *Let φ be the invariant measure of $\mathbf{Q}$, cf. Equation (7.7), and let Π_φ be the map $\Pi_\varphi : f \mapsto \langle \varphi, f \rangle \underline{1}$. The variance $V_{\varphi,T_{\mathfrak{A}}}(k)$ of the first passage time for counting observables is, $\forall k \geq 0$:*

$$\frac{V_{\varphi,T_{\mathfrak{A}}}(k)}{k} = \left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1} \right\rangle^2 + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{1}}{\mathbf{1}-\mathbf{Q}}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1} \right\rangle - \frac{2}{k}\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{Q}-\mathbf{Q}^{k+1}}{(\mathbf{1}-\mathbf{Q})^2}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1} \right\rangle.$$

Proof. We recall the explicit expression for the moment generating function from Lemma 7.1

$$\mathbb{E}_\varphi[e^{uT_{\mathfrak{A}}(k)}] = \left\langle \varphi, \left(\frac{\mathbf{L}_\infty}{u+\mathbf{L}_\infty}\mathbf{Q}\right)^k \underline{1} \right\rangle, \quad u < \overline{\lambda}.$$

Define

$$\mathbf{F}_u := \frac{\mathbf{L}_\infty}{u+\mathbf{L}_\infty}.$$

We can write the first moment as

$$\mathbb{E}_\varphi[e^{uT_{\mathfrak{A}}(k)}]' = \left\langle \varphi, \sum_{i=1}^{k} (\mathbf{F}_u\mathbf{Q})^{i-1} (\mathbf{F}_u'\mathbf{Q}) (\mathbf{F}_u\mathbf{Q})^{k-i} \underline{1} \right\rangle. \tag{7.10}$$

At $u = 0$ this gives us the form of the asymptotic mean.

$$\mathbb{E}_\varphi[T_{\mathfrak{A}}(k)] = -k\left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1} \right\rangle. \tag{7.11}$$

Differentiating Equation (7.10) at $u = 0$ gives us the second moment

$$\begin{aligned}
\mathbb{E}_\varphi[T_{\mathfrak{A}}(k)^2] &= 2\left\langle \varphi, \mathbf{L}_\infty^{-2}\underline{1} \right\rangle k + \sum_{i=1}^{k}\left\langle \varphi, \mathbf{L}_\infty^{-1}\left(\sum_{j=1}^{i-1}\mathbf{Q}^{i-j} + \sum_{j=1}^{k-i}\mathbf{Q}^j\right)\mathbf{L}_\infty^{-1}\underline{1} \right\rangle \\
&= 2\left\langle \varphi, \mathbf{L}_\infty^{-2}\underline{1} \right\rangle k + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}\sum_{1\leq j<i\leq k}\mathbf{Q}^j\mathbf{L}_\infty^{-1}\underline{1} \right\rangle \\
&= 2\left\langle \varphi, \mathbf{L}_\infty^{-2}\underline{1} \right\rangle k + 2\sum_{i=2}^{k}\sum_{j=1}^{i-1}\left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1} \right\rangle^2 \\
&\qquad + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}\sum_{i=2}^{k}\sum_{j=1}^{i-1}\mathbf{Q}^j(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1} \right\rangle
\end{aligned}$$

where to arrive at the third line, after $\mathbf{Q}^j$ we have inserted $\Pi_\varphi + \mathbf{1} - \Pi_\varphi$, where Π_φ is the projection onto $\underline{1}$. Using the fact that

$\sum_{i=2}^{k}\sum_{j=1}^{i-1} 1 = \frac{k}{2}(k-1)$, and recalling Equation (7.11) for the expression for the first moment, hence

$$\frac{V_{\varphi,T_{\mathfrak{A}}}(k)}{k} = 2\left\langle \varphi, \mathbf{L}_\infty^{-2}\underline{1}\right\rangle - \left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1}\right\rangle^2 + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{Q}}{\mathbf{1}-\mathbf{Q}}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle - \frac{2}{k}\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{Q}-\mathbf{Q}^{k+1}}{(\mathbf{1}-\mathbf{Q})^2}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle.$$

Finally, we can again place $\Pi_\varphi + \mathbf{1} - \Pi_\varphi$ in the first term, in between the two $\mathbf{L}_\infty^{-1}$. Rearranging this gives the final result

$$\frac{V_{\varphi,T_{\mathfrak{A}}}(k)}{k} = \left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1}\right\rangle^2 + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{1}}{\mathbf{1}-\mathbf{Q}}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle - \frac{2}{k}\left\langle \varphi, \mathbf{L}_\infty^{-1}\frac{\mathbf{Q}-\mathbf{Q}^{k+1}}{(\mathbf{1}-\mathbf{Q})^2}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle.$$

□

As with the equivalent expression in the fixed-time ensemble, Lemma 7.2 is not practical for large system sizes. Nevertheless, it gives us the form of the variance, which can be upper bounded directly to obtain inverse TUR-type bounds which we do in Proposition 7.8.1.

7.2 LARGE DEVIATION PRINCIPLE FOR GENERAL COUNTING OBSERVABLES

The large deviations of FPTs for counting observables were studied in [48]. There the equivalency of the fixed-time and fixed-observable ensembles were used to derive bounds on the limiting logarithmic moment generating function, rate function and a lower bound on the variance of the FPT in terms of the average activity. In this section we rigorously prove the large deviation principle (LDP) for first passage times of not just the dynamical activity but for general counting observables.

The following lemma is required for the theorem. It establishes an equivalence between the irreducibility of $\mathbf{L}$ (Hypothesis 2.1) and the irreducibility of the tilted generator $\mathbf{L}_u$ (Equation (3.3)). In the case of a classical Markov chain, the equivalence of irreducibility and primitivity for a Markov chain is a consequence of Levy's theorem [126, Theorem 8]. The lemma requires the generator $\mathbf{L}$ of a classical Markov chain and its perturbations have the form

$$\mathbf{L}_s = \sum_{x\neq y} e^{sa_{xy}} w_{xy} - \mathbf{R}, \quad a_{xy}, s \in \mathbb{R}.$$

We can state the following.

Lemma 7.3. *If* $\mathbf{L}$ *satisfies Hypothesis 2.1, then* $\mathbf{L}_s$ *generates a primitive semigroup.*

We omit the proof here, as it follows the same method as the equivalent lemma in the quantum case, Lemma 8.4 Chapter 8, where the proof is stated in full. Recall the definition of the large deviation principle (Section 3.1). We can now prove the LDP for generic counting observables of the form in Equation (2.5).

Theorem 7.4. *Let us consider any non-empty subset $\mathfrak{A}$ of the set of possible jumps. The collection of corresponding first passage times $\{T_{\mathfrak{A}}(k)/k\}$ satisfies a LDP with good rate function given by*

$$I_{\mathfrak{A}}(t) := \sup_{u\in\mathbb{R}}\{ut - \log(r(u))\}$$

where

$$r(u) = \begin{cases} r\left(\mathbf{Q}_u\right) & \text{if } u < \overline{\lambda} \\ +\infty & \text{otherwise} \end{cases}$$

where $\mathbf{Q}_u := -(u + \mathbf{L}_\infty)^{-1}\mathbf{W}_1$ *and* $\overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_\infty)\}$.

Proof. The proof of Lemma 7.1 shows that if $u < \overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_\infty)\}$, then

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}] = \left\langle \nu, \mathbf{Q}_u^k \underline{1} \right\rangle, \tag{7.12}$$

where $\mathbf{Q}_u := -(u + \mathbf{L}_\infty)^{-1}\mathbf{W}_1$. From the expression

$$\mathbf{Q}_u = \int_0^{+\infty} e^{(u+\mathbf{L}_\infty)t}\mathbf{W}_1 dt$$

one sees that $\mathbf{Q}_u$ is a positivity preserving map for every $u < \overline{\lambda}$. From Perron-Frobenius theorem (see Theorem 5.5 Chapter 5), we know that $r(u) := r(\mathbf{Q}_u)$ is an eigenvalue of $\mathbf{Q}_u$ that admits a positive eigenvector $x(u)$. With simple algebraic manipulations one can see that

$$\mathbf{Q}_u x(u) = r(u)x(u) \Leftrightarrow \mathbf{L}_{s(u)}x(u) = -ux(u)$$

where $\mathbf{L}_{s(u)} := \mathbf{L} + (e^{s(u)} - 1)\mathbf{W}_1$ and $s(u) = -\log(r(u))$.

The perturbations of $\mathbf{L}$ given by $\mathbf{L}_s$ for $s \in \mathbb{R}$ are irreducible (see Lemma 7.3 Chapter 7), hence they admit a unique positive eigenvector, which is actually strictly positive and corresponds to the the eigenvalue given by $\max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_s)\}$. Therefore, $-u = \max\{\Re(z) : z \in \mathrm{Sp}(\mathbf{L}_{s(u)})\}$ and $x(u) > 0$ is the unique eigenvector for $\mathbf{Q}_u$ corresponding to $r(u)$. One can also show that $r(u)$ is in fact algebraically simple for $\mathbf{Q}_u$ as in the proof of [127, Lemma 5.3].

Summing up, one has that for $u < \overline{\lambda}$ the function $u \mapsto r(u)$ is smooth (actually analytic in a complex neighbourhood of the values we are considering) and

$$\lim_{k\to+\infty} \frac{1}{k}\log(\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}]) = \log(r(u)). \tag{7.13}$$

Indeed,

$$\frac{1}{k}\log(\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}]) \leq \frac{1}{k}\log(\|\mathbf{Q}_u^k\|_{\infty\to\infty}) \xrightarrow[k\to+\infty]{} \log(r(u))$$

thanks to Gelfand's formula. On the other hand, using (7.12) and assuming that $\|x(u)\|_\infty \leq 1$, one has

$$\frac{1}{k}\log(\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}]) \geq \log(r(u)) + \frac{1}{k}\log(\langle \nu, x(u)\rangle) \xrightarrow[k\to+\infty]{} \log(r(u)),$$

since $x(u) > 0$ and $\langle \nu, x(u)\rangle > 0$.

In order to apply Gärtner-Ellis theorem (cf. Theorem 3.7), we only need to show that

$$\lim_{u\to\overline{\lambda}^-} \log(r(u)) = \lim_{u\to\overline{\lambda}^-} \log(r(u))' = +\infty. \tag{7.14}$$

Notice that $r(u)$ and $\log(r(u))' = r'(u)/r(u)$ are both monotone non-decreasing (they are limits of monotone non-decreasing functions cf. Equation (7.13)): the limits in Equation (7.14) exist and we only need to show that they cannot be finite.

Let $\mathbf{T}$ be the spectral projection of $\mathbf{L}_\infty$ corresponding to $-\overline{\lambda}$; we remark that $\mathbf{L}_\infty$ restricted to the range of $\mathbf{T}$ is diagonalisable. To show this, let us assume that this is not the case. If $-\overline{\lambda}$ is an eigenvalue of $\mathbf{L}_\infty$, then 0 is an eigenvalue of $\mathbf{L}' := \mathbf{L}_\infty + \overline{\lambda}\mathbf{1}$. The matrix $\mathbf{L}'$ restricted to the range $\mathbf{T}$ is also not diagonalisable. However, this means that the restriction contains a Jordan block, in which case the norm of $e^{t\mathbf{L}'}$ explodes for large t which contradicts the fact that $e^{t\mathbf{L}'}$ generates a contraction semigroup by Lumer-Phillips theorem (see, for instance, Theorem 3.15 and the following corollaries in [119]).

Let us first show that

$$\lim_{u\to\overline{\lambda}^-} r(u) = +\infty.$$

Notice that $\mathbf{TQ} \neq 0$, since $\mathbf{TQ}(\underline{1}) = \mathbf{T}\underline{1} \neq 0$. Therefore

$$\mathbf{Q}_u = \frac{\mathbf{L}_\infty}{\mathbf{L}_\infty + u}\mathbf{Q} = \frac{\overline{\lambda}}{\overline{\lambda} - u}\mathbf{TQ} + (\mathbf{1} - \mathbf{T})\frac{\mathbf{L}_\infty}{u + \mathbf{L}_\infty}\mathbf{Q}$$

has a norm that explodes for $u \to \overline{\lambda}^-$. By contradiction, let us assume that for $u \to \overline{\lambda}^-$, $r(u) \to r(\overline{\lambda})^- < +\infty$. This implies that we can choose $x(u)$ such that it converges to the unique strictly positive Perron-Frobenius eigenvector of $\mathbf{L}_{s(\overline{\lambda})}$ and that $\min \mathrm{Sp}(x(u)) \not\to 0$. Therefore we have for every $0 \leq u < \overline{\lambda}$

$$\begin{aligned}
\|\mathbf{Q}_u\|_{\infty\to\infty} &= \|\mathbf{Q}_u(\underline{1})\|_\infty \\
&\leq \frac{1}{\min \mathrm{Sp}(x(u))}\|\mathbf{Q}_u(x(u))\|_\infty \\
&= r(u)\frac{\|x(u)\|_\infty}{\min \mathrm{Sp}(x(u))}.
\end{aligned}$$

Since the right side remains finite as $u \to \bar{\lambda}^-$, this contradicts the fact that $\|\mathbf{Q}_u\|_{\infty\to\infty}$ diverges. Notice that the first equality in the previous equation is due to Theorem 5.6 in Chapter 5.

Let $l(u)$ be the left eigenvector of $\mathbf{Q}_u$; we can assume that $\langle l(u), x(u)\rangle \equiv 1$, therefore one has

$$\begin{aligned} r(u) &= \langle l(u), -(u+\mathbf{L}_\infty)^{-1}\mathbf{W}_1 x(u)\rangle \\ &= \underbrace{\langle l(u), -(u+\mathbf{L}_\infty)^{-1}\mathbf{T}\mathbf{W}_1 x(u)\rangle}_{(I)} \\ &+ \underbrace{\langle l(u), -(u+\mathbf{L}_\infty)^{-1}(\mathbf{1}-\mathbf{T})\mathbf{W}_1 x(u)\rangle}_{(II)}. \end{aligned}$$

Since (II) stays bounded, for $u \to \overline{\lambda}^-$ one has

$$r(u) \asymp \frac{\overline{\lambda}}{\overline{\lambda}-u}\langle l(u), \mathbf{T}\mathbf{Q}x(u)\rangle$$

with both sides diverging as $u \to \bar{\lambda}^-$. Differentiating the previous expression for $r(u)$ and dividing for $r(u)$ one gets

$$\begin{aligned} \frac{r'(u)}{r(u)} &= \frac{\langle l(u), (u+\mathbf{L}_\infty)^{-2}\mathbf{W}_1 x(u)\rangle}{r(u)} \\ &= \underbrace{\frac{\overline{\lambda}}{(\overline{\lambda}-u)^2}\frac{\langle l(u), \mathbf{T}\mathbf{Q}x(u)\rangle}{r(u)}}_{(I)} + \underbrace{\frac{\langle l(u), (u+\mathbf{L}_\infty)^{-2}(\mathbf{1}-\mathbf{T})\mathbf{W}_1 x(u)\rangle}{r(u)}}_{(II)}. \end{aligned}$$

When $u \to \overline{\lambda}^-$, $(II) \to 0$, while $(I) \asymp (\overline{\lambda}-u)^{-1}$ and we are done.

□

The proof highlights some properties of $r(u)$, which imply (as one would expect) that $I_{\mathfrak{A}}(t) = +\infty$ for $t \leq 0$ and that

$$\lim_{t\to 0^+} I_{\mathfrak{A}}(t) = +\infty, \quad \lim_{t\to+\infty} I_{\mathfrak{A}}(t) = +\infty, \quad \lim_{t\to+\infty} I'_{\mathfrak{A}}(t) = \overline{\lambda}.$$

Moreover, $I_{\mathfrak{A}}(t)$ has a unique minimum in $\langle t_{\mathfrak{A}}\rangle$, where it is equal to 0. The strong law of large numbers (SLLN) is a consequence of the smoothness of $r(u)$ around 0; see for instance [80, Theorem II.6.3 and Theorem II.6.4]. We refer to [32, 48] for a more in depth discussion of the physical meaning of this result.

7.3 CONCENTRATION BOUND FOR DYNAMICAL ACTIVITY

Although we have derived a large deviation principle which is valid for an asymptotic number of jumps k, in this section we provide a concentration inequality for the first passage time for dynamical activity which is valid $\forall$ $k \in \mathbb{N}$. Recall that we consider a classical continuous time Markov process with generator $\mathbf{L}$ whose jumps can be described by a discrete-time process

with transition matrix $\mathbf{P}$, cf. Equation (2.1). The dynamical activity $K_{\mathfrak{E},t}$ is the total number of configuration changes (referred to also as jumps) occurring in a trajectory up to time t [11, 26, 31]. The corresponding first passage time $T_{\mathfrak{E}}(k)$, is the time of the k-th jump, cf. Definition 2.4. The first concentration inequality of this chapter is an upper bound on the probability that the average jump time $T_{\mathfrak{E}}(k)/k$ deviates from its asymptotic or stationary mean (cf. Equation (7.8))

$$\langle t_{\mathfrak{E}} \rangle = \sum_{x \in E} \pi(x) \frac{1}{R_x}.$$

We now introduce two quantities which appear in the bounds of Theorem 7.7 below:

1. the second moment at stationarity:

$$2b_c^2 := \sum_{x \in E} \pi(x) \frac{2}{R_x^2}; \tag{7.15}$$

2. the longest expected holding time:

$$c_c := \max_{x \in E} \left\{ \frac{1}{R_x} \right\} = \frac{1}{\min_{x \in E}\{R_x\}} := \frac{1}{d}. \tag{7.16}$$

Below we report in our notation two technical lemmas that were proved in [68] and which are used in the proofs of some of the bounds obtained in this chapter.

Lemma 7.5 (Lemma 21 (i), [68]). *Let $\mathbf{M}_f : L^2_\pi(E) \to L^2_\pi(E)$ be the multiplication operator associated to a real valued function f, i.e. $\mathbf{M}_f g = fg$ for every $g \in L^2_\pi(E)$. Let $\hat{\mathbf{P}}$ be the León-Perron operator defined in Equation (2.4). Then the following statement holds:*

$$\|f\|_\pi^2 \leq \|\mathbf{M}_f \hat{\mathbf{P}} \mathbf{M}_f\|_\pi.$$

The following result is a slight generalisation of [68, Lemma 2.1 (iii)]. This lemma is also used in the quantum case so we use dual notation for both chapters. In the following, we will use Ψ to denote either a transition matrix $\mathbf{P}$ or a quantum channel Φ, χ to indicate their invariant state, i.e. π and σ, and we denote by $\mathcal{H}$ the Hilbert space corresponding to their invariant state, i.e. $L^2_\pi(E)$ and L^2_σ, respectively. Furthermore, denote Π to be map Π_π (cf. Definition 2.2) or Π_σ (cf. Definition 5.7) in the appropriate case. We recall that the notation $\hat{\Psi}$ stands for the León-Perron version of Ψ (again cf. Definition 2.2 or Definition 5.7).

Lemma 7.6. *For any operators $\mathbf{A}, \mathbf{B}$ acting on $\mathcal{H}$, the following holds true:*

$$\|\mathbf{A}\Psi\mathbf{B}\|_\chi \leq \left\|\mathbf{B}^\dagger \hat{\Psi} \mathbf{B}\right\|_\chi^{\frac{1}{2}} \left\|\mathbf{A}\hat{\Psi}\mathbf{A}^\dagger\right\|_\chi^{\frac{1}{2}}.$$

Proof. Let us consider $h_1, h_2 \in \mathcal{H}$, then

$$\begin{aligned}
|\langle \Psi h_1, h_2\rangle_\chi| &= |\langle (\mathbf{I}_{M_d(\mathbb{C})} - \Pi)(\Psi - \Pi)(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_1, h_2\rangle_\chi + \langle \Pi h_1, h_2\rangle_\chi| \\
&= |\langle (\Psi - \Pi)(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_1, (\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_2\rangle_\chi + \langle \Pi h_1, h_2\rangle_\chi| \\
&\leq |\langle (\Psi - \Pi)(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_1, (\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_2\rangle_\chi| + |\langle \Pi h_1, h_2\rangle_\chi| \\
&\leq (1-\varepsilon)\|(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_1\|_\chi \|(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_2\|_\chi + |\langle h_1, \chi\rangle\langle \chi, h_2\rangle| \\
&\leq \sqrt{g_1}\sqrt{g_2}
\end{aligned}$$

where

$$g_i = (1-\varepsilon)\|(\mathbf{I}_{M_d(\mathbb{C})} - \Pi)h_i\|_\chi^2 + |\langle \chi, h_i\rangle|^2$$

for $i \in \{1,2\}$. Therefore

$$|\langle \Psi h_1, h_2\rangle_\chi| = \langle \hat{\Psi} h_1, h_1\rangle_\chi^{\frac{1}{2}} \langle \hat{\Psi} h_2, h_2\rangle_\chi^{\frac{1}{2}}.$$

We can then proceed to complete the lemma:

$$\begin{aligned}
\|\mathbf{A}\Psi\mathbf{B}\|_\chi &= \sup_{h_1,h_2:\|h_i\|_\chi=1} |\langle \mathbf{A}\Psi\mathbf{B}h_1, h_2\rangle_\chi| \\
&= \sup_{h_1,h_2:\|h_i\|_\chi=1} |\langle \Psi\mathbf{B}h_1, \mathbf{A}^\dagger h_2\rangle_\chi| \\
&\leq \sup_{h_1,h_2:\|h_i\|_\chi=1} \langle \hat{\Psi}\mathbf{B}h_1, \mathbf{B}h_1\rangle_\chi^{\frac{1}{2}} \langle \hat{\Psi}\mathbf{A}^\dagger h_2, \mathbf{A}^\dagger h_2\rangle_\chi^{\frac{1}{2}} \\
&= \sup_{h_1:\|h_1\|_\chi=1} \langle \mathbf{B}^\dagger\hat{\Psi}\mathbf{B}h_1, h_1\rangle_\chi^{\frac{1}{2}} \sup_{h_2:\|h_2\|_\chi=1} \langle \mathbf{A}\hat{\Psi}\mathbf{A}^\dagger h_2, h_2\rangle_\chi^{\frac{1}{2}} \\
&= \left\|\mathbf{B}^\dagger\hat{\Psi}\mathbf{B}\right\|_\chi^{\frac{1}{2}} \left\|\mathbf{A}\hat{\Psi}\mathbf{A}^\dagger\right\|_\chi^{\frac{1}{2}}.
\end{aligned}$$

□

The following theorem states the first concentration inequality of this chapter.

Theorem 7.7. *Suppose Hypothesis 2.1 holds (*$\mathbf{L}$ *is irreducible) and let* ε *be the spectral gap of* $\mathbf{P}^\dagger\mathbf{P}$*, cf. Equation* (2.3)*. For every* $\gamma > 0$ *and* $k \in \mathbb{N}$ *the following holds true:*

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \geq \langle t_{\mathfrak{E}}\rangle + \gamma\right) \leq C(\nu)\exp\left(-k\frac{\gamma^2\varepsilon}{4b_c^2}h\left(\frac{5c_c\gamma}{2b_c^2}\right)\right)$$

and

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \leq \langle t_{\mathfrak{E}}\rangle - \gamma\right) \leq C(\nu)\exp\left(-k\frac{\gamma^2\varepsilon}{4b_c^2}h\left(\frac{5c_c\gamma}{2b_c^2}\right)\right),$$

where $h(x) := (\sqrt{1+x} + \frac{x}{2} + 1)^{-1}$, $C(\nu) := \max_{x\in E}\{\nu(x)/\pi(x)\}$ *and* b_c^2, c_c *are as defined in Equations* (7.15) *and* (7.16) *respectively.*

The proof of Theorem 7.7 is given below and follows the same line as in [67, Theorem 3.3]. From the proof, one can see that if $\mathbf{P}$ is self-adjoint, one can derive an upper bound with a slightly different expression which contains the spectral gap of $\mathbf{P}$ instead of its absolute spectral gap.

First, let us make few considerations regarding the quantities appearing in the bound. $C(\nu)$ accounts for the difference between the initial measure and the stationary one, in particular $C(\pi) = 1$. The absolute spectral gap ε controls the speed at which an arbitrary density ν converges to the invariant measure π under iterations of the transition operator $\mathbf{P}_*$: indeed, for every $k \geq 1$

$$\begin{aligned}\|\mathbf{P}_*^k(\nu - \pi)\|_1 &\leq \left\| \mathbf{P}^{\dagger k} \left(\frac{\nu^{1/2}}{\pi^{1/2}} - \underline{1} \right) \right\|_\pi \\ &\leq 2\varepsilon^{\frac{k}{2}} \left(1 - \sum_{x \in E} \nu(x)^{1/2} \pi(x)^{1/2} \right).\end{aligned}$$

This enables one to upper bound the deviation probability of $T_{\mathfrak{E}}(k)$ using stationary properties of the system. We remark that the use of the spectral gap of $\mathbf{P}^\dagger\mathbf{P}$ instead of the one of $\mathbf{P}$ allows to bound the fluctuations of the first passage time for every $k \geq 1$ and not only asymptotically in k. Small values of ε can correspond in some models to big fluctuations of the first passage time (cf. Section 7.5.2 below and Chapter 6 Section 6.6.2).

The second moment at stationarity b_c^2 encodes the variance of $T_{\mathfrak{E}}(k)$ in the stationary regime. Indeed, the distribution of the interarrival times t_i at stationarity is the same as the random variable obtained drawing a state x from the invariant distribution π and then sampling from an independent exponential random variable with parameter R_x. Such a random variable has a variance equal to

$$2 \sum_{x \in E} \pi(x) \frac{1}{R_x^2} - \left(\sum_{x \in E} \pi(x) \frac{1}{R_x} \right)^2.$$

Notice that the following inequalities hold true:

$$b_c^2 \leq 2 \sum_{x \in E} \pi(x) \frac{1}{R_x^2} - \left(\sum_{x \in E} \pi(x) \frac{1}{R_x} \right)^2 \leq 2 b_c^2,$$

hence the variance of the interarrival times at stationarity and b_c^2 (see Equation (7.15)) differ at most by a factor 2. The bigger b_c^2, the bigger the fluctuations of the first passage time. Finally, as one might reasonably expect, the dependence of the bound on c_c is such that the bigger c_c, the heavier the right tail. Notice that the ratio between b_c^2 and c_c that appears in the bound can be controlled by the average at stationarity:

$$\frac{d}{q} \langle t_{\mathfrak{E}} \rangle \leq \frac{b_c^2}{c_c} = \sum_{x \in E} \pi(x) \frac{d}{R_x^2} \leq \langle t_{\mathfrak{E}} \rangle,$$

where we recall that q (resp. d) are the maximum (resp. minimum) escape rates. On the other hand, ε and b_c^2 are quite independent from each other.

For example, if one modifies uniformly the speed of the Markov process X, i.e. $\mathbf{L} \to \lambda\mathbf{L}$ for some positive λ, one has that the jump process does not change and therefore ε remains the same, while $b_c^2 \to \lambda^{-2} b_c^2$. Notice that the bound has the right scaling with respect to this group of transformations: indeed, the bound becomes

$$C(\nu) \exp\left(-k \frac{(\lambda\gamma)^2 \varepsilon}{4 b_c^2} h\left(\frac{5 c_c \lambda \gamma}{2 b_c^2}\right)\right),$$

which corresponds to the upper bound for deviations of the order $\lambda\gamma$ for the original dynamics.

Proof of Theorem 7.7. We begin by using the Chernoff bound to upper bound the probability of $T_{\mathfrak{E}}(k)/k$ right deviating from $\langle t_{\mathfrak{E}} \rangle$ by more than $\gamma > 0$, using the moment generating function:

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \geq \langle t_{\mathfrak{E}} \rangle + \gamma\right) \leq e^{-ku(\langle t_{\mathfrak{E}} \rangle + \gamma)} \mathbb{E}_\nu[e^{u T_{\mathfrak{E}}(k)}], \quad u \geq 0. \tag{7.17}$$

We now focus on upper bounding the moment generating function using the definition in Equation (7.1). Introducing the notation

$$\mathbf{F}_u := \frac{\mathbf{R}}{\mathbf{R} - u}, \quad u < d,$$

one has that for every $0 \leq u < d$ the following holds true:

$$\begin{aligned}
\mathbb{E}_\nu[e^{u T_{\mathfrak{E}}(k)}] &= \left\langle \nu, (\mathbf{F}_u \mathbf{P})^k \underline{1} \right\rangle \\
&= \left\langle \frac{\nu}{\pi}, (\mathbf{F}_u \mathbf{P})^k \underline{1} \right\rangle_\pi \\
&= \left\langle \mathbf{F}_u^{\frac{1}{2}} \frac{\nu}{\pi}, (\mathbf{F}_u^{\frac{1}{2}} \mathbf{P} \mathbf{F}_u^{\frac{1}{2}})^{k-1} \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\rangle_\pi \\
&\leq \left\| \mathbf{F}_u^{\frac{1}{2}} \frac{\nu}{\pi} \right\|_\pi \left\| \mathbf{F}_u^{\frac{1}{2}} \mathbf{P} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi^{k-1} \left\| \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\|_\pi,
\end{aligned}$$

where $\frac{\nu}{\pi}(x) = \frac{\nu(x)}{\pi(x)}, \forall x \in E$ and $\|\cdot\|_\pi$ the $L^2_\pi(E)$ norm defined in Section 2.4. We use the notation $\mathbf{M}_{\frac{\nu}{\pi}}$ to denote the multiplication operator corresponding to $\frac{\nu}{\pi}$. We can write the following:

$$\begin{aligned}
\left\| \mathbf{F}_u^{\frac{1}{2}} \frac{\nu}{\pi} \right\|_\pi &= \left\| \mathbf{F}_u^{\frac{1}{2}} \mathbf{M}_{\frac{\nu}{\pi}} \underline{1} \right\|_\pi \\
&= \left\| \mathbf{M}_{\frac{\nu}{\pi}} \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\|_\pi \\
&\leq \left\| \mathbf{M}_{\frac{\nu}{\pi}} \right\|_\pi \left\| \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\|_\pi \\
&= \left\| \frac{\nu}{\pi} \right\|_\infty \left\| \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\|_\pi.
\end{aligned}$$

Note that $\|\mathbf{M}_{\frac{\nu}{\pi}}\|_\pi = \|\frac{\nu}{\pi}\|_\infty$ since we know that $\mathbf{M}_{\frac{\nu}{\pi}}$ is a diagonal matrix. Applying Lemma 7.5 with $f = \mathbf{F}_u^{\frac{1}{2}} \underline{1}$, Lemma 7.6 with $\mathbf{A} = \mathbf{B} = \mathbf{F}_u^{\frac{1}{2}}$ and remembering that $\mathbf{F}_u^{\frac{1}{2}}$ is self-adjoint one can derive the following inequalities:

1. $\left\| \mathbf{F}_u^{\frac{1}{2}} \underline{1} \right\|_\pi \leq \left\| \mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi^{\frac{1}{2}}$,
2. $\left\| \mathbf{F}_u^{\frac{1}{2}} \mathbf{P} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi \leq \left\| \mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi$,

where we recall that $\hat{\mathbf{P}}$ is the León-Perron matrix associated to $\mathbf{P}$, cf. Definition 2.2. Therefore, we get:

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{E}}(k)}] \leq \left\| \frac{\nu}{\pi} \right\|_\infty \left\| \mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi^k, \quad 0 \leq u < d. \tag{7.18}$$

If we set $C(\nu) := \left\| \frac{\nu}{\pi} \right\|_\infty$, the problem is now reduced to finding an upper bound on $\left\| \mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}} \right\|_\pi$. Notice that $\mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}}$ is self-adjoint, therefore its norm coincide with its spectral radius $r(u)$. Moreover, $\mathbf{F}_u^{\frac{1}{2}} \hat{\mathbf{P}} \mathbf{F}_u^{\frac{1}{2}}$ is similar to $\hat{\mathbf{P}} \mathbf{F}_u$, hence they share the same spectral radius; finally, Perron-Frobenius theory ensures that $r(u)$ is an eigenvalue of $\mathbf{P}(u) := \hat{\mathbf{P}} \mathbf{F}_u$.

We can write $\mathbf{P}(u)$ as a power series:

$$\mathbf{P}(u) = \hat{\mathbf{P}} + \sum_{l=1}^{\infty} u^l \hat{\mathbf{P}} \left(\frac{\mathbf{1}}{\mathbf{R}} \right)^l, \quad 0 \leq u < d. \tag{7.19}$$

For conciseness of notation, we denote $\mathbf{D} := \mathbf{R}^{-1}$. Perturbation theory (cf. [105] or Section 4.4) implies that if we can bound $\|\hat{\mathbf{P}}\mathbf{D}^l\|_\pi \leq \delta \zeta^{l-1}$ for some $\delta, \zeta > 0$ and $l > 1$, within the range $|u| < (2\delta\varepsilon^{-1} + \zeta)^{-1}$, with ε the spectral gap of $\hat{\mathbf{P}}$ (which is equal to the absolute spectral gap of $\mathbf{P}$), then the spectral radius $r(u)$ can be expressed in the following way:

$$r(u) = 1 + \sum_{l=1}^{\infty} u^l r^{(l)}, \tag{7.20}$$

where:

$$r^{(l)} = \sum_{p=1}^{l} \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \cdots + \nu_p = l,\, \nu_i \geq 1 \\ \mu_1 + \cdots + \mu_p = p-1,\, \mu_j \geq 0}} \operatorname{tr}\left(\hat{\mathbf{P}} \mathbf{D}^{\nu_1} \mathbf{S}^{(\mu_1)} \cdots \hat{\mathbf{P}} \mathbf{D}^{\nu_p} \mathbf{S}^{(\mu_p)} \right), \tag{7.21}$$

with $\mathbf{S}^{(0)} = -\Pi_\pi$, $\mathbf{S}^{(1)} = (\hat{\mathbf{P}} - \mathbf{1} + \Pi_\pi)^{-1} - \Pi_\pi = -\varepsilon^{-1}(\mathbf{1} - \Pi_\pi)$ and $\mathbf{S}^{(\mu)}$ the μ^{th} power of $\mathbf{S}^{(1)}$. Note that $\mathbf{S}^{(1)}$ is equal to the Moore-Penrose inverse of $-(\mathbf{1} - \hat{\mathbf{P}})$ and $\|\mathbf{S}^{(\mu)}\|_\pi = \varepsilon^{-\mu}$ for $\mu \geq 1$. If we set our estimates $\delta = \zeta = c_c := d^{-1}$ (recall that c_c is the longest expected holding time), we can indeed bound $\|\hat{\mathbf{P}}\mathbf{D}^l\|_\pi$ by:

$$\|\hat{\mathbf{P}}\mathbf{D}^l\|_\pi \leq c_c^l,$$

which gives a radius of convergence $|u| < \frac{\varepsilon}{c_c(2+\varepsilon)} < \frac{1}{c_c} = d$. Using Equation (7.21), we can explicitly determine that

$$\begin{aligned} r^{(1)} &= \langle t_{\mathfrak{E}} \rangle, \\ r^{(2)} &= b_c^2 - \left\langle \mathbf{D}\underline{1}, \mathbf{S}^{(1)} \hat{\mathbf{P}} \mathbf{D} \underline{1} \right\rangle_\pi. \end{aligned}$$

We then seek to bound $r^{(l)}$ for $l \geq 3$. For $p = 1$:

$$-\mathrm{tr}(\hat{\mathbf{P}}\mathbf{D}^l(-\Pi_\pi)) = \left\langle \mathbf{D}^l\underline{1}, \underline{1}\right\rangle_\pi = \sum_{x\in E}\pi(x)\frac{1}{R_x^l}.$$

For the $p \geq 2$ cases, we get:

$$\begin{aligned}-\mathrm{tr}(\hat{\mathbf{P}}\mathbf{D}^{\nu_1}\mathbf{S}^{(\mu_1)}\cdots\hat{\mathbf{P}}\mathbf{D}^{\nu_p}\mathbf{S}^{(\mu_p)}) &= \Big\langle \mathbf{D}\underline{1}, \mathbf{D}^{\nu_1-1}\mathbf{S}^{(\mu_1)}\hat{\mathbf{P}}\mathbf{D}^{\nu_2}\mathbf{S}^{(\mu_2)}\cdots \\ &\qquad \cdots\hat{\mathbf{P}}\mathbf{D}^{\nu_{p-1}}\mathbf{S}^{(\mu_{p-1})}\hat{\mathbf{P}}\mathbf{D}^{\nu_p-1}\mathbf{D}\underline{1}\Big\rangle_\pi \\ &\leq \|\mathbf{D}\underline{1}\|_\pi^2\|\mathbf{D}\|_\pi^{l-2}\|\mathbf{S}^{(1)}\|_\pi^{p-1},\end{aligned}$$

where we have taken $\mu_p = 0$, which is justified since $\mu_1 + \cdots + \mu_p = p - 1$, there is at least one $\mu_j = 0$, and the trace is cyclic. Again using Cauchy-Schwarz we can bound the terms as follows: $\|\mathbf{S}^{(1)}\|_\pi = \frac{1}{\varepsilon}$, $\|\mathbf{D}\|_\pi = c_c$, $\|\mathbf{D}\underline{1}\|_\pi^2 = b_c^2$. We also have that $\varepsilon \leq 1$. For $p \geq 2$, each term in the inner sum of Equation (7.21) is then bounded by

$$b_c^2\frac{c_c^{l-2}}{\varepsilon^{l-1}}.$$

From [67] the number of terms $N(l)$ in Equation (7.21) is bounded by

$$N(l) = \sum_{p=1}^{n}\binom{l-1}{p-1}\binom{2(p-1)}{p-1}\frac{1}{p} \leq 5^{l-2}, \tag{7.22}$$

which is valid for $l \geq 3$. Combining everything together, the bound on each $r^{(l)}$ becomes

$$|r^{(l)}| \leq \left\langle \mathbf{D}^l\underline{1}, \underline{1}\right\rangle_\pi + 5^{l-2}b_c^2\frac{c_c^{l-2}}{\varepsilon^{l-1}} = \left\langle \mathbf{D}^l\underline{1}, \underline{1}\right\rangle_\pi + \frac{b_c^2}{5c_c}\left(\frac{5c_c}{\varepsilon}\right)^{l-1}.$$

Which, through a simple computation is in fact valid for $l = 2$ as well. Therefore, the eigenvalue $r(u)$ can be bounded above by

$$\begin{aligned}r(u) &\leq 1 + \langle t_{\mathfrak{E}}\rangle u + \sum_{l=2}^{\infty}\left\langle \mathbf{D}^l\underline{1}, \underline{1}\right\rangle_\pi u^l + \frac{b_c^2 u}{5c_c}\left(\frac{5c_c u}{\varepsilon}\right)^{l-1} \\ &\leq \exp\left(\langle t_{\mathfrak{E}}\rangle u + \sum_{l=2}^{\infty}\left\langle \mathbf{D}^l\underline{1}, \underline{1}\right\rangle_\pi u^l + \frac{b_c^2 u}{5c_c}\left(\frac{5c_c u}{\varepsilon}\right)^{l-1}\right),\end{aligned}$$

where we have used the fact that $1 + x \leq e^x$. We can further bound this by focusing on the latter two terms inside the exponential:

1.
$$\sum_{l=2}^{\infty}\langle \mathbf{D}^l\underline{1}, \underline{1}\rangle_\pi u^l \leq \sum_{l=2}^{\infty} b_c^2 c_c^{l-2}u^l = b_c^2\frac{u^2}{1-c_c u}$$

2.
$$\begin{aligned}\sum_{l=2}^{\infty}\frac{b_c^2 u}{5c_c}\left(\frac{5c_c u}{\varepsilon}\right)^{l-1} &= \sum_{l=2}^{\infty}\frac{b_c^2 u^2}{\varepsilon}\left(\frac{5c_c u}{\varepsilon}\right)^{l-2} \\ &= \frac{b_c^2 u^2}{\varepsilon}\left(1 - \frac{5c_c u}{\varepsilon}\right)^{-1}.\end{aligned}$$

The power series for point 1. again gives a radius of convergence of $0 \leq u < \frac{1}{c_c}$. Point 2. gives a radius of convergence of $0 \leq u < \frac{\varepsilon}{5c_c} < \frac{\varepsilon}{c_c(2+\varepsilon)} < \frac{1}{c_c}$. Combining these terms together and using the upper bound on the Laplace transform in Equation (7.18), we have, for $0 < u < \frac{\varepsilon}{5c_c}$

$$\mathbb{E}_\nu[e^{(uT_{\mathfrak{E}}(k))}] \leq C(\nu) \exp\left(k\left(\langle t_{\mathfrak{E}}\rangle u + b_c^2 u^2\left(\frac{1}{1-c_c u} + \frac{1}{\varepsilon - 5c_c u}\right)\right)\right). \tag{7.23}$$

Since $1 - c_c u > \varepsilon - 5c_c u$, we can relax slightly the bound on the moment generating function, such that when we apply the Chernoff bound in Equation (7.17), we get that

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \geq \langle t_{\mathfrak{E}}\rangle + \gamma\right) \leq C(\nu) \exp\left(-k\left(\gamma u - \frac{2b_c^2 u^2}{\varepsilon}\left(1 - \frac{5c_c u}{\varepsilon}\right)^{-1}\right)\right). \tag{7.24}$$

Consider the more general expression below, with $\alpha, \beta > 0$

$$\gamma u - \alpha u^2 (1 - \beta u)^{-1}.$$

Elementary calculations show that

$$\sup_{u < \frac{1}{\beta}}\left\{\gamma u - \alpha u^2 (1-\beta u)^{-1}\right\} = \frac{\gamma^2}{2\alpha} h\left(\frac{\beta\gamma}{\alpha}\right), \tag{7.25}$$

where $h(x) := (1 + \frac{x}{2} + \sqrt{1+x})^{-1}$. In our case, $\alpha = \frac{2b_c^2}{\varepsilon}$, $\beta = \frac{5c_c}{\varepsilon}$. Therefore taking the infimum on the right hand side of the bound in (7.24), which is valid $\forall u \in [0, \frac{\varepsilon}{5c_c})$, yields the final result for right deviations

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \geq \langle t_{\mathfrak{E}}\rangle + \gamma\right) \leq C(\nu) \exp\left(-k\frac{\gamma^2\varepsilon}{4b_c^2} h\left(\frac{5c_c\gamma}{2b_c^2}\right)\right).$$

To prove the concentration inequality for left deviations, we write the Chernoff bound for this case, this time with $u \leq 0$

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \leq \langle t_{\mathfrak{E}}\rangle - \gamma\right) \leq e^{-ku(\langle t_{\mathfrak{E}}\rangle - \gamma)} \mathbb{E}_\nu[e^{uT_{\mathfrak{E}}(k)}].$$

We can repeat the proof we did for right deviations, due to the fact we are upper bounding the absolute value of the terms in the expansion of (7.21) for $l \geq 2$. We obtain an upper bound on the moment generating function

$$\mathbb{E}_\nu[e^{(uT_{\mathfrak{E}}(k))}] \leq C(\nu) \exp\left(k\left(\langle t_{\mathfrak{E}}\rangle u + b_c^2 u^2\left(\frac{1}{1-c_c|u|} + \frac{1}{\varepsilon - 5c_c|u|}\right)\right)\right),$$

which is valid for $0 \leq |u| < \frac{\varepsilon}{5c_c}$. One obtains a concentration bound in terms of u of a similar form to Equation (7.24). To optimise this, we set $u \to -u$ and note that the resulting expression has the same form as Equation (7.25). Using the same values of α and β, this gives the final result for left deviations and concludes the proof of Theorem 7.7

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{E}}(k)}{k} \leq \langle t_{\mathfrak{E}}\rangle - \gamma\right) \leq C(\nu) \exp\left(-k\frac{\gamma^2\varepsilon}{4b_c^2} h\left(\frac{5c_c\gamma}{2b_c^2}\right)\right).$$

□

We remark that to optimise in u, we relaxed the bound on the moment generating function, albeit in a different way to the derivation of the concentration inequality in Chapter 6.

As a consequence of the proof of Theorem 7.7, we obtain an upper bound on the variance at stationarity of the FPT corresponding to the dynamical activity. This upper bound is derived from the perturbation expansion of the spectral radius $r(u)$, in Equation (7.21). The result complements the lower bound (or TUR) for the FPT of the activity obtained in [48]:

Proposition 7.7.1. *The variance $V_{\pi,T_{\mathfrak{E}}}(k)$ of the first passage time for the total activity at stationarity is bounded from above by:*

$$\frac{V_{\pi,T_{\mathfrak{E}}}(k)}{k} \leq \left(1+\frac{2}{\varepsilon}\right) b_c^2.$$

Proof. Notice that for $u \geq 0$ small enough, one has

$$\begin{aligned}\log(\mathbb{E}_\pi[e^{uT_{\mathfrak{E}}(k)}]) &= \langle t_{\mathfrak{E}}\rangle ku + \frac{1}{2}V_{\pi,T_{\mathfrak{E}}}(k)u^2 + o(u^2)\\ &\leq k\log(r(u))\\ &= kr'(0)u + \frac{k}{2}(r''(0) - (r'(0))^2)u^2 + o(u^2),\end{aligned}$$

where $r(u)$ is given by Equation (7.20). We recall that

$$r'(0) = r^{(1)} = \langle t_{\mathfrak{E}}\rangle = \langle \mathbf{D}\underline{1}, \underline{1}\rangle_\pi,$$

$$r''(0) = 2r^{(2)} = 2\langle \mathbf{D}\underline{1}, \mathbf{D}\underline{1}\rangle_\pi + 2\left\langle \mathbf{D}\underline{1}, \frac{\hat{\mathbf{P}}}{\mathbf{1}-\hat{\mathbf{P}}}\mathbf{D}\underline{1}\right\rangle_\pi,$$

therefore

$$r''(0) - (r'(0))^2 = \langle \mathbf{D}\underline{1}, \underline{1}\rangle_\pi^2 + 2\left\langle \mathbf{D}\underline{1}, \frac{\mathbf{1}}{\mathbf{1}-\hat{\mathbf{P}}}\mathbf{D}\underline{1}\right\rangle_\pi \leq \left(1+\frac{2}{\varepsilon}\right) b_c^2,$$

where the last inequality was obtained using Cauchy-Schwarz and recalling that $\|\mathbf{D}\underline{1}\|_\pi^2 = b_c^2$ and from Equation (7.21) that $\| (\mathbf{1}-\mathbf{P})^{-1} \|_\pi = \varepsilon^{-1}$. Hence $V_{\pi,T_{\mathfrak{E}}}(k) \leq \left(1+\frac{2}{\varepsilon}\right) b_c^2 k$.

□

We could have derived an upper bound on the variance simply from the second derivative of the upper bound on the entire moment generating function in Equation (7.23). We recall that this upper bound on the moment generating function was derived by upper bounding every coefficient in the expansion of the spectral radius (Equation (7.20)). This would have given a weaker upper bound (by an amount b_c^2) so instead we derived the variance bound using the method in the above proof: taking the expansion of the spectral radius and upper bounding *only* the terms corresponding to the variance.

7.4 TAIL BOUND FOR GENERAL COUNTING OBSERVABLES

Our second main result of this chapter is a concentration bound on the tails of the distribution of the FPT for general counting observables, $T_{\mathfrak{A}}(k)/k$. Similarly to the above result, this bounds the probability that $T_{\mathfrak{A}}(k)/k$ deviates from $\langle t_{\mathfrak{A}} \rangle$. Recall that $\mathbf{L}_\infty$ is a sub-Markov generator describing the jumps in $\mathfrak{A}^{\mathsf{C}}$, cf. Equation (7.2). We introduce the following notation

$$\beta := \left\| \frac{\mathbf{1}}{\mathbf{L}_\infty} \right\|_{\infty \to \infty}.$$

In the case of the dynamical activity it is simply given by $\beta = \max_x 1/R_x$. In general, β satisfies $\beta \geq \langle t_{\mathfrak{A}} \rangle$ by Equation (7.8), and as we show below, it can be interpreted as the longest timescale of the system. Indeed, since $-\mathbf{L}_\infty^{-1}$ is a positivity preserving map, Russo-Dye theorem (cf. Theorem 5.6) gives

$$\|\mathbf{L}_\infty^{-1}\|_{\infty \to \infty} = \|\mathbf{L}_\infty^{-1}\underline{1}\|_\infty = \max_{x \in E} \sum_{y \in E} |\mathbf{L}_{\infty,xy}^{-1}|,$$

and since

$$\|\mathbf{L}_\infty^{-1}\underline{1}\|_\infty = \max_\nu -\langle \nu, \mathbf{L}_\infty^{-1}\underline{1} \rangle = \max_\nu \mathbb{E}_\nu[T_{\mathfrak{A}}(1)],$$

we obtain

$$\beta = \max_\nu \mathbb{E}_\nu[T_{\mathfrak{A}}(1)],$$

where ν is a probability density on the state space. We can now state our second main result:

Theorem 7.8. *Let* $\mathbf{L}$ *be irreducible and* $\mathfrak{A} \subseteq \mathfrak{E}$ *be non-empty. For every* $k \in \mathbb{N}$ *and* $\gamma > \beta - \langle t_{\mathfrak{A}} \rangle$

$$\mathbb{P}_\nu \left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}} \rangle + \gamma \right) \leq \exp\left(-k \left(\frac{\gamma + \langle t_{\mathfrak{A}} \rangle - \beta}{\beta} - \log\left(\frac{\gamma + \langle t_{\mathfrak{A}} \rangle}{\beta} \right) \right) \right).$$

Here we comment briefly on the rather simple idea the proof of Theorem 7.8. Let Z be the sum of k independent exponential random variables with parameter β^{-1}, then by applying the Chernoff bound one obtains that for every $0 \leq u < \beta^{-1}$

$$\mathbb{P}(Z/k \geq \beta + \gamma') \leq \exp\left(-k \left(u(\beta + \gamma') + \log(\beta) + \log(\beta^{-1} - u) \right) \right).$$

Optimising in u in the previous equation, one gets

$$\mathbb{P}(Z/k \geq \beta + \gamma') \leq \exp\left(-k \left(\gamma' \beta^{-1} - \log(1 + \gamma' \beta^{-1}) \right) \right). \tag{7.26}$$

As the interarrival times are distributed according to the matrix-exponential distribution ([128]) with rate matrix $-\mathbf{L}_\infty$, and $\beta = \|\mathbf{L}_{\infty *}^{-1}\|_{1 \to 1}$, the moment generating function of $T_{\mathfrak{A}}(k)$ is bounded from above by that of Z. Equation (7.26) then provides the bound in Theorem 7.8.

Proof. The proof begins with the same procedure as Theorems 7.7 and 8.6, but differs in that we do not use the $L^2_\pi(E)$ Hilbert space. We again begin by applying Chernoff bound:

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}}\rangle + \gamma\right) \leq e^{-uk(\langle t_{\mathfrak{A}}\rangle+\gamma)}\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}], \quad u > 0. \tag{7.27}$$

The next step is to upper bound the Laplace transform: for $0 \leq u < \|\mathbf{L}_\infty^{-1}\|^{-1}_{\infty\to\infty}$ one has

$$\begin{aligned}
\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}] &= \left\langle \nu, \left(\frac{\mathbf{1}}{\mathbf{1}+\frac{u}{\mathbf{L}_\infty}}\mathbf{Q}\right)^k \underline{1}\right\rangle \\
&= \left\langle \nu, \left(\sum_{i=0}^\infty u^i \left(-\frac{\mathbf{1}}{\mathbf{L}_\infty}\right)^i \mathbf{Q}\right)^k \underline{1}\right\rangle \\
&\leq \underbrace{\|\nu\|_1}_{=1}\left(\sum_{i=0}^\infty u^i \left\|\frac{\mathbf{1}}{\mathbf{L}_\infty}\right\|^i_{\infty\to\infty}\right)^k \underbrace{\|\underline{1}\|_\infty}_{=1}.
\end{aligned}$$

If we denote $\beta := \left\|\mathbf{L}^{-1}_{\infty *}\right\|_{1\to 1}$, then for $u < \frac{1}{\beta}$

$$\mathbb{E}_\nu[e^{uT_{\mathfrak{A}}(k)}] \leq \left(\frac{1}{1-\beta u}\right)^k = \exp\left(-k\log(1-\beta u)\right).$$

Placing this back into Equation (7.27) we get

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}}\rangle + \gamma\right) \leq \exp\left(-k\left(u(\gamma+\langle t_{\mathfrak{A}}\rangle) + \log(1-\beta u)\right)\right), \tag{7.28}$$

for $0 \leq u < \beta^{-1}$. The minimum of the R.H.S. is achieved at

$$u^* = \frac{1}{\beta} - \frac{1}{\gamma + \langle t_{\mathfrak{A}}\rangle} = \frac{\gamma + \langle t_{\mathfrak{A}}\rangle - \beta}{\beta(\gamma + \langle t_{\mathfrak{A}}\rangle)}.$$

We have $u^* < \beta^{-1}$, and $u^* > 0$ if $\gamma > \beta - \langle t_{\mathfrak{A}}\rangle$. Substituting u^* into Equation (7.28) gives the final result:

$$\mathbb{P}_\nu\left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}}\rangle + \gamma\right) \leq \exp\left(-k\left(\frac{\gamma+\langle t_{\mathfrak{A}}\rangle - \beta}{\beta} - \log\left(\frac{\gamma+\langle t_{\mathfrak{A}}\rangle}{\beta}\right)\right)\right).$$

This concludes the proof of Theorem 7.8.

□

Theorem 7.8 provides a tail bound for a more general class of observables than Theorem 7.7. Unlike the case of Theorem 7.7, the bound in Theorem 7.8 does not cover small fluctuations and this makes it impossible to use to derive any bound on the variance of $T_{\mathfrak{A}}(k)$ in the spirit of Proposition 7.7.1. Nevertheless, using the explicit expression of the variance (see Lemma 7.2), one can derive the following FPT inverse TUR.

Proposition 7.8.1. *Given any non-empty set of jumps $\mathfrak{A}$, the variance $V_{\varphi,T_{\mathfrak{A}}}(k)$ of the corresponding first passage time at stationarity is bounded from above by:*

$$\frac{V_{\varphi,T_{\mathfrak{A}}}(k)}{k} \leq \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2,$$

where

$$\tilde{\varepsilon} := 1 - \max\{\|\mathbf{Q}f\|_\infty : \|f\|_\infty = 1,\ \langle\varphi, f\rangle = 0\}.$$

We recall that φ is the unique invariant law for $\mathbf{Q}$ and was defined in Equation (7.7).

Proof. From the proof of Lemma 7.2 one can see that

$$\begin{aligned}
\frac{V_{\varphi,T_{\mathfrak{A}}}(k)}{k} &= \left\langle \varphi, \mathbf{L}_\infty^{-1}\underline{1}\right\rangle^2 + 2\left\langle \varphi, \mathbf{L}_\infty^{-1}(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle \\
&\quad + \frac{2}{k}\left\langle \varphi, \mathbf{L}_\infty^{-1}\sum_{i=2}^{k}\sum_{j=1}^{i-1}\mathbf{Q}^j(\mathbf{1}-\Pi_\varphi)\mathbf{L}_\infty^{-1}\underline{1}\right\rangle \\
&\leq \left(1+2\left(1+\frac{1}{k}\sum_{i=2}^{k}\sum_{j=1}^{i-1}(1-\tilde{\varepsilon})^j\right)\right)\beta^2 \\
&= \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2 - \frac{2((1-\tilde{\varepsilon})-(1-\tilde{\varepsilon})^{k+1})}{k\tilde{\varepsilon}^2}\beta^2 \\
&\leq \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2.
\end{aligned}$$

□

We remark that Proposition 7.8.1 together with Chebyshev inequality provides bounds on small deviations as well.

The constant β may be difficult to compute, especially for large systems. However, it is not hard to check that Theorem 7.8 remains true if we replace β with any $\tilde{\beta} \geq \beta$. The following proposition shows it is possible to upper bound β (and obtain alternative concentration bounds for the FPT) in terms of the following simpler quantities of the system:

1. maximum escape rate (recall from Equation (6.1)):

$$q := \max_{x\in E}\{R_x\};$$

2. minimum transition rate:

$$h := \min_{x,y\in E}\{w_{xy} : w_{xy} > 0\}; \tag{7.29}$$

3. minimax jump distance $\tilde{k}$: the minimum $k \in \mathbb{N}$ such that for any initial state $i \in E$ there exists a trajectory $(i_0 = i, i_1, \ldots, i_l)$ with $l \leq k$ such that $w_{i_j,i_{j+1}} > 0$ for all $j = 0, \ldots l-1$ and the trajectory ends with a jump in $\mathfrak{A}$, i.e. $(i_{l-1}, i_l) \in \mathfrak{A}$.

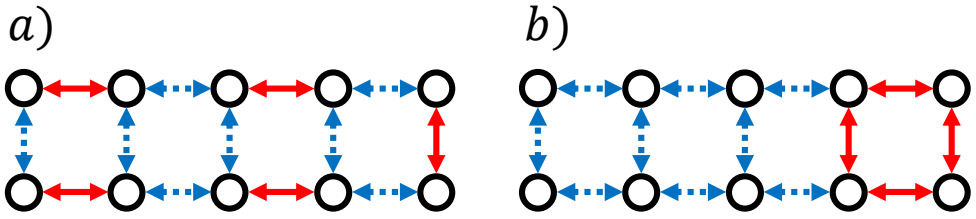

Figure 7.1: Minimax Jump Distance: configurations of a discrete system are represented by circles, and allowed transitions between them by arrows. Jumps in $\mathfrak{A}$ (full/red) contribute to the observable, whilst jumps in $\mathfrak{A}^C$ (dotted/blue) do not. (a) System where red jumps are distributed throughout the graph, in this case $\tilde{k} = 1$. (b) Uneven distribution of jumps in $\mathfrak{A}$, in this case $\tilde{k} = 4$.

While q and h can be computed easily in terms of the transition rates, the minimax jump distance $\tilde{k}$ can be read off the graph of the process, see Figure 7.1. The fact that it is finite follows from the irreducibility of the Markov process. The proposition is stated below:

Proposition 7.8.2. *For general counting observables, the norm* $\beta := \left\|\mathbf{L}_\infty^{-1}\right\|_{\infty\to\infty}$ *is bounded from above by:*

$$\beta \leq c_c \tilde{k} \max_{(x,y)\notin\mathfrak{A}} \left\{\frac{R_x}{w_{xy}}\right\} \max_{(x,y)\in\mathfrak{A}} \left\{\frac{R_x}{w_{xy}}\right\}^{\tilde{k}-1} \leq c_c \tilde{k} \left(\frac{q}{h}\right)^{\tilde{k}} =: \tilde{\beta},$$

with c_c, q, h *defined in Equations* (7.16), (6.1) *and* (7.29) *respectively. The concentration bound in Theorem 7.8 holds with* β *replaced by any of the two upper bounds above.*

Proof. From the expression (7.4) for $-\mathbf{L}_\infty^{-1}$ we obtain

$$-\frac{\mathbf{1}}{\mathbf{L}_\infty} = \sum_{k=0}^{\infty} \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^k \frac{\mathbf{1}}{\mathbf{R}}.$$

Let $\tilde{k} > 0$ be the minimax jump distance as it has been defined before Proposition 7.8.2, i.e. the maximum over all states of the minimum number of jumps which suffices to get from that state to a final jump between states in $\mathfrak{A}$. The previous sum can be written in the following way

$$\sum_{l=0}^{\infty} \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^l = \sum_{m=0}^{\tilde{k}-1} \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^m \sum_{n=0}^{\infty} \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{\tilde{k}n},$$

where we break up l into multiples of $\tilde{k}$ and a remainder term, since $\mathbb{N}_0 = \cup_{m=0}^{\tilde{k}-1} m + \tilde{k}\mathbb{N}_0$. We can upper bound as

$$\beta = \left\|\frac{\mathbf{1}}{\mathbf{L}_\infty}\right\|_{\infty\to\infty} \leq \left\|\frac{\mathbf{1}}{\mathbf{R}}\right\|_{\infty\to\infty} \sum_{m=0}^{\tilde{k}-1} \left\|\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right\|_{\infty\to\infty}^m \sum_{n=0}^{\infty} \left\|\left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{\tilde{k}}\right\|_{\infty\to\infty}^n.$$

Notice that $\left\|\mathbf{R}^{-1}\right\|_{\infty\to\infty} = c_c$. As $\mathfrak{A}$ is non-empty, the spectral radius of $\mathbf{R}^{-1}\mathbf{W}_2$ is strictly less than 1 (see Lemma 7.1 ii)); therefore, there exists a k

such that $\left\| \left(\mathbf{R}^{-1}\mathbf{W}_2\right)^k \right\|_{\infty\to\infty} \leq r < 1$. We will now show that k can be taken equal to $\tilde{k}$; in this case, we can write the upper bound of β in terms of r as

$$\beta \leq c_c \tilde{k} \frac{1}{1-r}. \tag{7.30}$$

So we just need to find r. Note that for any matrix $\mathbf{G}$ with positive entries, one has $\|\mathbf{G}\|_{\infty\to\infty} = \|\mathbf{G}\underline{1}\|_\infty = \max_{x\in E}\langle \delta_x, \mathbf{G}\underline{1}\rangle$. Therefore

$$\left\| \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{\tilde{k}} \right\|_{\infty\to\infty} = \left\langle \delta_{x_0}, \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{\tilde{k}} \underline{1} \right\rangle, \tag{7.31}$$

where x_0 is the state which attains the norm. From the structure of the generator, we know that

$$\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\underline{1} = \underline{1} - \frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_1\underline{1}. \tag{7.32}$$

By the definition of the minimax jump distance, we know that there exists a path $x_0, \ldots, x_l$ that happens with positive probability and such that $(x_{l-1}, x_l) \in \mathfrak{A}$. We can rewrite the right side of (7.31) as

$$\left\langle \delta_{x_0}, \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{\tilde{k}} \underline{1} \right\rangle \leq \left\langle \delta_{x_0}, \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{l} \underline{1} \right\rangle =$$
$$\left\langle \delta_{x_0}, \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{l-1} \underline{1} \right\rangle - \left\langle \delta_{x_0}, \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_2\right)^{l-1} \left(\frac{\mathbf{1}}{\mathbf{R}}\mathbf{W}_1\right) \underline{1} \right\rangle =: g_+ - g_-.$$

In the first inequality we used the fact that $\mathbf{R}^{-1}\mathbf{W}_1$ is sub-Markovian, while in the second equality we made use of Equation (7.32). It is easy to see that $g_+ \leq 1$. Moreover, we know that

$$\begin{aligned} g_- &\geq \frac{w_{x_0x_1}}{R_{x_0}} \cdots \frac{w_{x_{l-1}x_l}}{R_{x_{l-1}}} \geq \min_{(x,y)\notin\mathfrak{A}} \left\{\frac{w_{xy}}{R_x}\right\}^{\tilde{k}-1} \min_{(x,y)\in\mathfrak{A}} \left\{\frac{w_{xy}}{R_x}\right\} \\ &\geq \min_{(x,y)\in\mathfrak{E}} \left\{\frac{w_{xy}}{R_x}\right\}^{\tilde{k}} \geq \left(\frac{h}{q}\right)^{\tilde{k}}. \end{aligned}$$

Hence we can take

$$r = 1 - \min_{(x,y)\notin\mathfrak{A}} \left\{\frac{w_{xy}}{R_x}\right\}^{\tilde{k}-1} \min_{(x,y)\in\mathfrak{A}} \left\{\frac{w_{xy}}{R_x}\right\} \leq 1 - \left(\frac{h}{q}\right)^{\tilde{k}}.$$

Applying this to Equation (7.30), we get that

$$\beta \leq c_c \tilde{k} \max_{(x,y)\notin\mathfrak{A}} \left\{\frac{R_x}{w_{xy}}\right\}^{\tilde{k}-1} \max_{(x,y)\in\mathfrak{A}} \left\{\frac{R_x}{w_{xy}}\right\} \leq c_c \tilde{k} \left(\frac{q}{h}\right)^{\tilde{k}}.$$

□

In the case of total activity, i.e. when $\mathfrak{A} = \mathfrak{E}$, one can easily see that $\beta = c_c$.

7.5 EXAMPLES

7.5.1 *Dynamical activity in a three-level system*

We illustrate the results of Theorem 7.7 with the model of a simple three-level system as sketched in Figure 7.2(a): the set of configurations is $E = \{0, 1, 2\}$, with reversible transitions $w_{01} = w_{10} = \omega$, $w_{02} = w_{20} = \upsilon$ and $w_{12} = w_{21} = \kappa$. Assuming that ω is the largest rate, the longest expected holding time is $c_c = \frac{1}{\kappa+\upsilon}$, whilst $\langle t_{\mathfrak{E}} \rangle$, b_c^2 and ε can easily be determined from the three-dimensional generator $\mathbf{L}$. In addition, we have $\beta = c_c$, for $\mathfrak{A} = \mathfrak{E}$. In Figure 7.2(b) we show the exact long time rate function of the activity (full/black) for a particular set of values of the transitions rates, together with the lower bound from Theorem 7.7 (dashed/blue),

$$\tilde{I}_{\mathfrak{E}}(\langle t_{\mathfrak{E}} \rangle + \gamma) = \frac{\gamma^2 \varepsilon}{4 b_c^2} h\left(\frac{5 c_c \gamma}{2 b_c^2}\right),$$

and the general lower bound from Theorem 7.8 (dotted/red),

$$\hat{I}_{\mathfrak{E}}(\langle t_{\mathfrak{E}} \rangle + \gamma) = \frac{\gamma + \langle t_{\mathfrak{E}} \rangle}{c_c} - 1 - \log\left(\frac{\gamma + \langle t_{\mathfrak{E}} \rangle}{c_c}\right).$$

We see that the bound from Theorem 7.8, $\hat{I}_{\mathfrak{E}}(\langle t_{\mathfrak{E}} \rangle + \gamma)$ is closer to the exact result than that from Theorem 7.7 for large enough deviations. Indeed, for $\gamma \gg 1$, one has $\hat{I}_{\mathfrak{E}}(\langle t_{\mathfrak{E}} \rangle + \gamma) \asymp \frac{\gamma}{c_c}$ while $\tilde{I}_{\mathfrak{E}}(\langle t_{\mathfrak{E}} \rangle + \gamma) \asymp \frac{\gamma\varepsilon}{5c_c} < \frac{\gamma}{c_c}$. For comparison, in Figure 7.2(b) we also show the *upper* bound to the rate function, the so-called TUR, from [48] (dot-dashed/magenta): the combination of the TUR and the "inverse TUR" from Theorems 7.7, 7.8 upper and lower bound the true rate function thus restricting the range of probabilities of rare events of the activity.

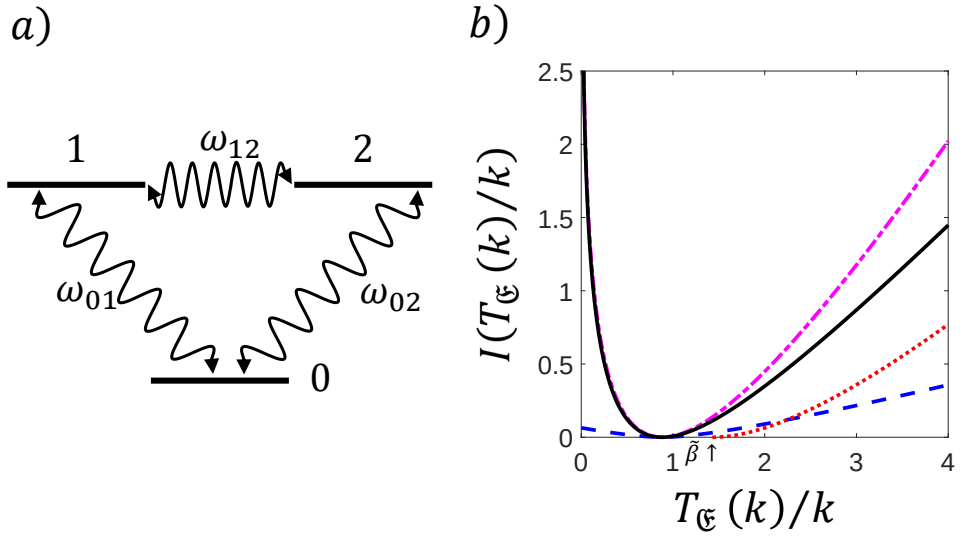

Figure 7.2: Bounds on the rate function of the FPT of the activity in a classical three-level system: (a) Sketch of the three-level system. (b) Rate function $I(T_{\mathfrak{E}}(k)/k)$ of the FPT for the dynamical activity, for the case with rates $w_{01} = w_{10} = \omega = 1$, $w_{02} = w_{20} = \upsilon = 0.5$, $w_{12} = w_{21} = \kappa = 0.2$. We show the exact rate function (full/black) and the lower bound specific to the activity from Theorem 7.7 (dashed/blue). We also show the the generic tail bound for counting observables from Theorem 7.8 (dotted/red) which is valid in the region $T_{\mathfrak{E}}(k)/k > \tilde{\beta} = 1/d = 1/(\kappa + \upsilon)$ (indicated by the arrow). For comparison we include the upper bound on the rate function (dot-dashed/magenta), known as the TUR [48].

7.5.2 *Metastability and the absolute spectral gap*

In the following example we show how closing the absolute spectral gap of the discrete-time generator $\mathbf{P}$ leads to an increase of the fluctuations of the total activity FPT in a simple model, as predicted by the concentration bound in Theorem 7.7. We consider the six-state system introduced in Section 6.6.2 composed of two three-state subsystems connected by edges controlled by a rate parameter ω, see Figure 7.3. For $\omega \to 0$ the spectral gap of the real part of the generator $\Re(\mathbf{L})$ vanishes and the configuration space E breaks up into two disconnected components, $E_1 = \{1,2,3\}$ and $E_2 = \{4,5,6\}$. When ω is non-zero but much smaller than the other rates, the system is metastable, with E_1 and E_2 becoming long-lived metastable "phases", since relaxation within $E_{1,2}$ will be much faster than relaxation in the whole of E.

We now study the statistics of the FPT of the activity in this model. We consider the case where the internal rates in E_1 are much larger than those in E_2, while maintaining the metastability condition, $\lambda_1, \kappa_1 \gg \lambda_2, \kappa_2 \gg \omega$. We call E_1 and E_2 the "active phase" and "inactive phase", respectively, as

the activity in stationary trajectories is much larger while the system is in E_1 than in E_2. The rate matrix can be written as,

$$\mathbf{W} = \begin{pmatrix} \tilde{\mathbf{W}}_1 & 0 \\ 0 & \tilde{\mathbf{W}}_2 \end{pmatrix} + \omega \begin{pmatrix} 0 & \mathbf{1}_3 \\ \mathbf{1}_3 & 0 \end{pmatrix},$$

where $\tilde{\mathbf{W}}_{1,2}$ are the 3×3 rate matrices for internal transitions in $E_{1,2} = \{1, 2, 3\}$, and the three-dimensional identity is denoted $\mathbf{1}_3$. In Theorem 7.7 we require the discrete-time operator $\mathbf{P} = \mathbf{R}^{-1}\mathbf{W}$ and its adjoint $\mathbf{P}^\dagger$, to form the multiplicative symmetrisation

$$\mathbf{P}^\dagger\mathbf{P} = \begin{pmatrix} \frac{\tilde{\mathbf{W}}_1{}^\dagger \tilde{\mathbf{W}}_1}{R_1^2} & 0 \\ 0 & \frac{\tilde{\mathbf{W}}_2{}^\dagger \tilde{\mathbf{W}}_2}{R_2^2} \end{pmatrix} + \omega^2 \begin{pmatrix} \frac{1}{R_2^2}\mathbf{1}_3 & 0 \\ 0 & \frac{1}{R_1^2}\mathbf{1}_3 \end{pmatrix} + \omega \begin{pmatrix} 0 & \frac{\tilde{\mathbf{W}}_1^\dagger}{R_1^2} + \frac{\tilde{\mathbf{W}}_2}{R_2^2} \\ \frac{\tilde{\mathbf{W}}_2^\dagger}{R_2^2} + \frac{\tilde{\mathbf{W}}_1}{R_1^2} & 0 \end{pmatrix},$$

where $R_{1,2} = \lambda_{1,2} + \kappa_{1,2}$. At $\omega = 0$, the spectrum of $\mathbf{P}^\dagger\mathbf{P}$ is equal to the union of the spectra of $\tilde{\mathbf{P}}_1^\dagger\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2^\dagger\tilde{\mathbf{P}}_2$, where $\tilde{\mathbf{P}}_1 = \frac{\tilde{\mathbf{W}}_1}{R_1}$ is the discrete-time transition matrix on each E_1 and $\tilde{\mathbf{P}}_2$ is that of E_2; hence, the algebraic multiplicity of the eigenvalue 1 is 2. By continuity of the spectrum for analytic perturbation, the absolute spectral gap vanishes as $\omega \to 0$. Proposition 7.7.1 then implies that the upper bound on the variance of the FPT will explode as this "phase transition" point is approached. Fluctuations of $T_{\mathfrak{E}}(k)$'s get bigger as well: since ω is much less than either of the escape rates within each metastable phase, the system gets "stuck" in either phase, resulting in larger fluctuations of the observed FPT.

The behaviour of the fluctuations of $T_{\mathfrak{E}}(k)$ as $\omega \to 0$ is not immediately apparent based on the form of the expression for the variance given by Lemma 7.2. We remark that for finite k the variance remains finite even as the gap closes. This can be seen by rewriting $V_{\pi,T_{\mathfrak{E}}}(k)$ as

$$\frac{V_{\pi,T_{\mathfrak{E}}}(k)}{k} = \langle \pi, \mathbf{D}\underline{1} \rangle^2 + 2 \left\langle \pi, \mathbf{D} \left(\mathbf{1} + \frac{\sum_{i=1}^{k-1}\sum_{j=1}^{i} \mathbf{P}^j}{k} \right) (\mathbf{1} - \Pi_\pi)\mathbf{D}\underline{1} \right\rangle$$

which is uniformly bounded in ω. Recall that in the case of total activity $\varphi = \pi$, $\mathbf{L}_\infty^{-1} = -\mathbf{R}^{-1} = -\mathbf{D}$ and $\mathbf{Q} = \mathbf{P}$. In the limit $k \to +\infty$ the expression reduces to the first two terms of Lemma 7.2 and the behaviour depends solely on $(\mathbf{1} - \mathbf{P})^{-1}$ and whether this causes a divergence as $\omega \to 0$. From Figure 7.4 one can see that for this model, the asymptotic variance does diverge and for finite k the fluctuations remain finite as expected. We can however see fingerprints of the asymptotic behaviour for intermediate k, which is captured by the upper bound in Proposition 7.7.1.

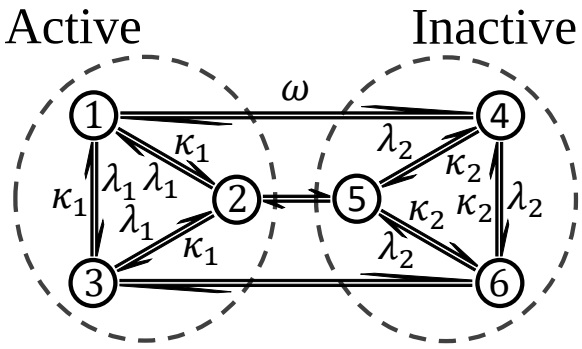

Figure 7.3: Six-state dynamical system: sketch of a six-state system with two phases, the active phase E_1 in the left circle and the inactive phase E_2 in the right circle. The phases are separated by edges controlled by ω. For small ω, each phase is metastable, and for $\omega \to 0$ the size of FPT fluctuations increases. This increase is captured by ε.

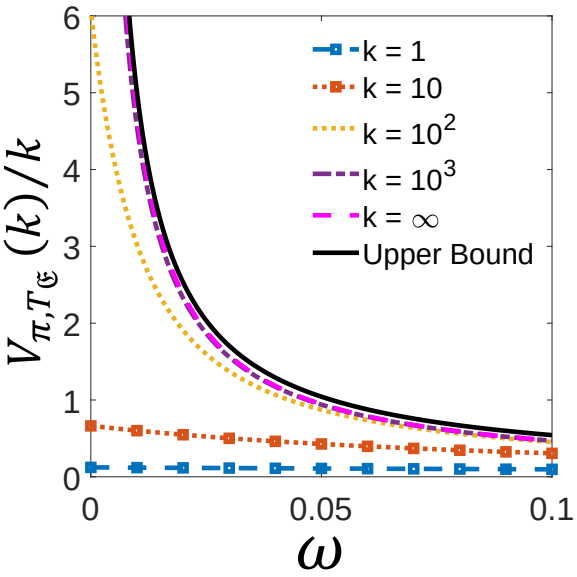

Figure 7.4: Upper bound on the variance of the FPT for activity in a six-state system: upper bound (full/black) on the scaled variance of $V_{\pi,T_\mathfrak{e}}(k)/k$ given by Proposition 7.7.1. This is valid for all k. We compare with the exact variance (cf. Lemma 7.2) for several values of k: 1 (dashed-marked/blue), 10^1 (dotted-marked/orange), 10^2 (dotted/yellow), 10^3 (dot-dashed/purple) and for $k = \infty$ (dashed/magenta). We compare these quantities as the controlling parameter $\omega \to 0$ and the system approaches a phase transition. The system is the model given in Figure 7.3 with rates $\lambda_1 = 30$, $\mu_1 = 10$ and $\lambda_2 = 0.3$, $\mu_2 = 0.1$.

7.5.3 *Three-level system counting subset of jumps*

To illustrate the results of Theorem 7.8, we use the same three-level model as in Section 7.5.1 but we consider the observable that only counts "clockwise" jumps, that is, the $0 \to 1$, $1 \to 2$ and $2 \to 0$ jumps but not their reverses. With this setup, the minimax jump distance is $\tilde{k} = 1$ since it is possible to perform a jump in $\mathfrak{A}$ which begins at any state. The lower bound to the rate function provided by Theorem 7.8 is illustrated in Figure 7.5.

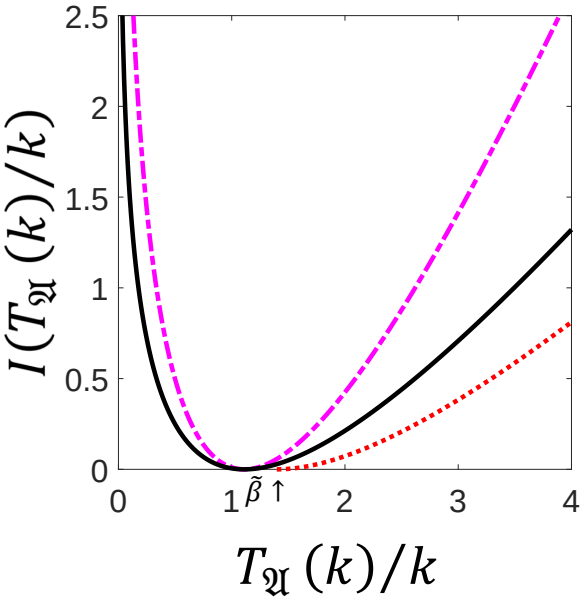

Figure 7.5: Bounds on the rate function of the FPT of a counting observable for a classical three-level system: exact rate function $I(T_{\mathfrak{A}}(k)/k)$ (full/black) for the FPT of the total count of jumps $0 \to 1$, $1 \to 2$ and $2 \to 0$ in the three-level system of Figure 7.2, with rates $w_{01} = w_{10} = \omega = 1$, $w_{02} = w_{20} = v = 0.9$, $w_{12} = w_{21} = \kappa = 0.8$. The tail bound from Theorem 7.8 (dotted/red) bounds deviations in the region $T_{\mathfrak{A}}(k)/k > \tilde{\beta}$ with $\beta \leq \tilde{\beta} = (\omega + v)/[\kappa(\kappa + v)]$. The rate function is bounded from above using the same method as in Figure 7.2 (dot-dashed/magenta) [48].

We demonstrated in the previous chapter, Chapter 6, complementary bounds to the large number of works on the thermodynamic uncertainty relations, often expressed as lower bounds on fluctuations in the fixed-time ensemble, where instead we limited fluctuations from the upper side. In the alternative description with fluctuating time it is natural to also consider bounds from this same side, which has been the contents of this chapter. Our concentration inequalities in Chapter 7 provide an upper bound on the probability of seeing fluctuations in FPTs. These inequalities are valid for all initial distributions and all values of the observable threshold that defines the FPT, and not only in the large threshold limit where large deviations theory applies. The bounds are written in terms of relatively simple quantities which describe the overall properties of the dynamics (and which in an

ideal setting can be determined by observation), in particular the longest expected holding time between events, and the spectral gap of the symmetrised generator.

So far, we have focused exclusively on classical Markov processes. In the next chapter, we will extend these first passage time results to the non-commutative (quantum) setting. In the classical case, where the operators are represented by matrices — some of which are diagonal — we were able to obtain tighter and more intuitive bounds compared to those derived directly in the quantum framework.

8

QUANTUM UPPER BOUNDS ON FLUCTUATIONS OF FIRST PASSAGE TIMES

We now move away from fluctuations of classical processes to fluctuations of quantum Markov processes. Recall that we consider a quantum trajectory as the outcomes of continuous-time counting measurements on the environment, with this environment consisting of emission channels described by the jump operators L_i. As observables we consider integer numbers of counts $N_{\mathfrak{A},t}$ from all or a subset $\mathfrak{A}$ of channels. The first passage times (FPTs) for $N_{\mathfrak{A},t}$ are defined in the same way as the classical process, cf. Equation (5.7) and we are interested in the deviations of the FPT away from some average.

In the quantum case we have less knowledge about the structure of the operators involved. As we will see, in the concentration inequalities some quantities are left in terms of operator norms. In the classical case, these norms can either be calculated analytically, or can be bounded in terms of simpler quantities. Physically, the quantities used in the bounds for classical processes could in principle be estimated (e.g. b_c^2 used in Theorem 7.7) whereas it would be non-trivial to estimate the size of the operator norms used in this chapter without evaluating them numerically.

The quantum setup further differs from the classical one in that the irreducibility of the transition operator Φ and generator $\mathcal{L}$ are not in general equivalent. Recall the ergodicity assumptions we made in Chapter 5 (Hypotheses 5.3 and 5.4). The following lemma proves that Hypothesis 5.4 is strictly stronger than Hypothesis 5.3.

Lemma 8.1. *The generator $\mathcal{L}$ has a unique invariant state if and only if Φ does. If Φ is irreducible then $\mathcal{L}$ is irreducible, but the converse is generally not true.*

Proof. Indeed if $\mathcal{L}_*(\nu) = 0$ for some state ν then

$$\Phi_*(\mathcal{J}_*(\nu)) = -\mathcal{J}_*\mathcal{L}_{0*}^{-1}(-\mathcal{L}_{0*}(\nu)) = \mathcal{J}_*(\nu)$$

so $\nu' := \mathcal{J}_*(\nu)/\mathrm{tr}[\mathcal{J}_*(\nu)]$ is a stationary state for Φ. Notice that if ν is strictly positive, then this is not necessarily true for ν', depending on the form of the jump operators. Conversely, if $\Phi_*(\nu) = \nu$ for some state ν then

$$\mathcal{L}_*[\mathcal{L}_{0*}^{-1}(\nu)] = (\mathcal{L}_{0*} + \mathcal{J}_*)[\mathcal{L}_{0*}^{-1}(\nu)] = \nu + \mathcal{J}_*\mathcal{L}_{0*}^{-1}(\nu) = \nu - \Phi_*(\nu) = 0$$

so $\nu' = \mathcal{L}_{0*}^{-1}(\nu)/\mathrm{tr}[\mathcal{L}_{0*}^{-1}(\nu)]$ is a stationary state for $\mathcal{L}$. Here we used the fact that $-\mathcal{L}_0^{-1} = \int_0^\infty e^{t\mathcal{L}_0}$ is completely positive. If Φ is irreducible then $\nu > 0$ and $-\mathcal{L}_{0*}^{-1}(\nu) > 0$, therefore $\mathcal{L}$ is irreducible, and Hypothesis 5.4 implies Hypothesis 5.3.

□

First we set up the form of the moment generating function (Section 8.1) in preparation for the Cramér-Chernoff method. As we have done in chapters 6 and 7, it is possible even for quantum systems to compute an explicit form of the variance of the first passage times, which we do in Lemma 8.3. Using irreducible properties of the tilted generator we can prove the large deviation principle (LDP) for first passage time of quantum counting observables (Section 8.2). Our first concentration inequality for quantum systems, in Section 8.3, is for the FPT of the total counts of Markov processes where the L_i are of any rank. Despite the additional restriction imposed by Hypothesis 5.4 for cases with these general jump operators, we derive in Section 8.4 a concentration inequality for FPTs of total counts of quantum reset processes — processes with L_i of rank one — which only require the weaker assumption Hypothesis 5.3. We then provide a tail bound for general counting observables in Section 8.5. Alongside each concentration inequality, propositions in the form of inverse TUR-type bounds on the variance follow. We demonstrate our results in Section 8.6 by computing lower bounds of the large deviation rate function of observables on simple quantum systems.

8.1 MOMENT GENERATING FUNCTION

The following splitting of the generator is relevant in order to study the properties of the stochastic process $T_{\mathfrak{A}}(k)$:

$$\mathcal{L} = \mathcal{J}_{\mathfrak{A}} + \mathcal{L}_{\infty},$$

where $\mathcal{J}_{\mathfrak{A}}(x) = \sum_{i\in\mathfrak{A}} L_i^* x L_i$ accounts for the change of state after a jump in $\mathfrak{A}$ and $\mathcal{L}_\infty$ for the average evolution between jumps in $\mathfrak{A}$. We denote as Ψ the transition operator analogous to $\mathbf{Q}$ in the classical case:

$$\Psi(x) = -\mathcal{L}_\infty^{-1} \circ \mathcal{J}_{\mathfrak{A}}(x). \tag{8.1}$$

If Hypothesis 5.3 holds, then Ψ admits a unique invariant state ς (which might have a non-trivial kernel). Note that if $\mathfrak{A} = I$ then $\mathcal{L}_\infty = \mathcal{L}_0$ and $\Psi = \Phi$.

The following lemma shows that all the objects introduced so far are well defined and allows us to write the Laplace transform for general counting observables.

Lemma 8.2. *Assume that Hypothesis 5.3 ($\mathcal{L}$ is irreducible) holds. Then the following statements are true:*

1. $\overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathcal{L}_\infty)\} > 0$, *hence $\mathcal{L}_\infty$ is invertible;*
2. *for every $u < \overline{\lambda}$, one has*

$$\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}] = \mathrm{tr}\left(\rho\left((u+\mathcal{L}_\infty)^{-1}\mathcal{L}_\infty\Psi\right)^k(\mathbf{1})\right);$$

3. $\|\mathcal{L}_\infty^{-1}\|_{\infty\to\infty}^{-1} \leq \overline{\lambda}$.

Proof. 1. $\mathcal{L}_\infty$ generates a sub-Markov semigroup $e^{t\mathcal{L}_\infty}$. From the spectral mapping theorem, one has that

$$\mathrm{Sp}(e^{t\mathcal{L}_\infty}) = e^{t\mathrm{Sp}(\mathcal{L}_\infty)}, \quad \forall t \geq 0.$$

Moreover, since $e^{t\mathcal{L}_\infty}$ is completely positive, by Perron-Frobenius theorem (Theorem 5.5) the spectral radius and largest eigenvalue of $e^{t\mathcal{L}_\infty}$ coincide. Hence

$$r(e^{t\mathcal{L}_\infty}) = e^{t\lambda}, \lambda := \max\{\Re(z) : z \in \mathrm{Sp}(\mathcal{L}_\infty)\},$$

and $\exists x \in M_d(\mathbb{C})$ with $x \geq 0$

$$\mathcal{L}_\infty(x) = \lambda x.$$

By contradiction, suppose that $\lambda = 0$, then

$$\mathcal{L}(x) = \mathcal{J}_{\mathfrak{A}}(x) + \mathcal{L}_\infty(x) = \mathcal{J}_{\mathfrak{A}}(x).$$

Therefore

$$\mathrm{tr}(\hat{\sigma}\mathcal{L}(x)) = \mathrm{tr}(\hat{\sigma}\mathcal{J}_{\mathfrak{A}}(x)) = 0,$$

since $\mathcal{L}_*(\hat{\sigma}) = 0$. From the irreducibility assumption (Hypothesis 5.3), we have that $\hat{\sigma} > 0$, and so $\mathcal{L}(x) = \mathcal{J}_{\mathfrak{A}}(x) = 0$. Moreover, irreducibility also implies that $x = \alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}$. Therefore $\mathcal{J}_{\mathfrak{A}}(x) = \alpha\mathcal{J}_{\mathfrak{A}}(\mathbf{1}) = 0$. Since $\mathcal{J}_{\mathfrak{A}}(\mathbf{1})$ is the sum of positive operators, we have that $\alpha = 0$, consequently $x = 0$ and we reach a contradiction.

2. Integrating over all trajectories, one can write

$$\mathbb{P}(T_{\mathfrak{A}}(k) \leq t) = \int_{\sum_{i=1}^k t_i \leq t} \mathrm{tr}\left(\mathcal{J}_{\mathfrak{A}*}e^{t_k\mathcal{L}_{\infty*}} \cdots e^{t_2\mathcal{L}_{\infty*}}\mathcal{J}_{\mathfrak{A}*}e^{t_1\mathcal{L}_{\infty*}}(\rho)\right) dt_1 \cdots dt_k.$$

For $u < \overline{\lambda}$, one has

$$-(u + \mathcal{L}_\infty)^{-1} = \int_0^{+\infty} e^{(u+\mathcal{L}_\infty)t}dt,$$

hence one can write the Laplace transform of $T_{\mathfrak{A}}(k)$ as

$$\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}] = \mathrm{tr}\left(\left(-\mathcal{J}_{\mathfrak{A}*}(u + \mathcal{L}_{\infty*})^{-1}\right)^k(\rho)\right),$$

and by the definition of Υ in Equation (8.1), we obtain the statement.

3. The spectral mapping theorem implies that $\mathrm{Sp}(\mathcal{L}_\infty^{-1}) = \{z^{-1} : z \in \mathrm{Sp}(\mathcal{L}_\infty)\}$, therefore one has that

$$\left\|\mathcal{L}_\infty^{-1}\right\|_{\infty\to\infty} \geq r(\mathcal{L}_\infty^{-1}) \geq \frac{1}{\overline{\lambda}} \quad \Leftrightarrow \quad \left\|\mathcal{L}_\infty^{-1}\right\|_{\infty\to\infty}^{-1} \leq \overline{\lambda}.$$

□

For the sake of completeness, and for the interested reader, we point out that an alternative expression for the probability density of FPTs from the one we use in the proof of Lemma 8.2 has been recently presented in [129]. This alternative expression can be written for general time-integrated observables of continuously indirectly monitored quantum systems, and it is particularly useful in explicit computations. From the expression of the moment generating function given in point 2 of the previous lemma, one can use standard techniques to show that

$$\frac{1}{k}T_{\mathfrak{A}}(k) \xrightarrow[k\to+\infty]{} \langle t_{\mathfrak{A}} \rangle := -\mathrm{tr}\left(\varsigma \mathcal{L}_\infty^{-1}(\mathbf{1})\right) \quad \text{a.s..} \tag{8.2}$$

where ς is the unique invariant state of Ψ defined in Equation (8.1). Our goal will be to investigate the probability of $T_{\mathfrak{A}}(k)/k$ deviating from $\langle t_{\mathfrak{A}} \rangle$. Using the expression of the moment generating function, we can compute the analytical form of the variance of the FPT for any process counting subsets of emissions.

Lemma 8.3. *Let ς be the invariant state of Ψ, cf. Equation* (8.1)*, and let Π_ς be the map $\Pi_\varsigma : x \mapsto \mathrm{tr}\,(\varsigma x)\,\mathbf{1}$. The variance of the first passage time for counting observables $\forall k \geq 0$ is given by:*

$$\begin{aligned}\frac{V_{\varsigma,T_{\mathfrak{A}}}(k)}{k} &= \mathrm{tr}\left(\varsigma \mathcal{L}_\infty^{-1}(\mathbf{1})\right)^2 \\ &+ 2\mathrm{tr}\left(\varsigma \mathcal{L}_\infty^{-1} \frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathbf{I}_{M_d(\mathbb{C})} - \Psi}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right) \\ &- \frac{2}{k}\mathrm{tr}\left(\varsigma \mathcal{L}_\infty^{-1} \frac{\Psi - \Psi^{k+1}}{(\mathbf{I}_{M_d(\mathbb{C})} - \Psi)^2}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right).\end{aligned}$$

Proof. We recall the explicit expression for the moment generating function from Lemma 8.2

$$\mathbb{E}_\varsigma[e^{uT_{\mathfrak{A}}(k)}] = \mathrm{tr}\left(\varsigma\left((u + \mathcal{L}_\infty)^{-1}\mathcal{L}_\infty \Psi\right)^k(\mathbf{1})\right), \quad u < \overline{\lambda}.$$

Define

$$\mathcal{F}_u := \frac{\mathcal{L}_\infty}{u + \mathcal{L}_\infty}.$$

We can write the first moment as

$$\mathbb{E}_\varsigma[e^{uT_{\mathfrak{A}}(k)}]' = \sum_{i=1}^{k} \mathrm{tr}\left(\varsigma\left(\mathcal{F}_u\Psi\right)^{i-1}\left(\mathcal{F}_u'\Psi\right)\left(\mathcal{F}_u\Psi\right)^{k-i}(\mathbf{1})\right). \tag{8.3}$$

At $u = 0$ this gives us the form of the asymptotic mean.

$$\mathbb{E}_\varsigma[T_{\mathfrak{A}}(k)] = -k\mathrm{tr}\left(\varsigma \mathcal{L}_\infty^{-1}(\mathbf{1})\right). \tag{8.4}$$

Differentiating Equation (8.3) at $u = 0$ gives us the second moment

$$\begin{aligned}
\mathbb{E}_\varsigma[T_{\mathfrak{A}}(k)^2] &= 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-2}(\mathbf{1})\right)k \\
&\quad + \sum_{i=1}^{k}\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\left(\sum_{j=1}^{i-1}\Psi^{i-j}+\sum_{j=1}^{k-i}\Psi^{j}\right)\mathcal{L}_\infty^{-1}(\mathbf{1})\right) \\
&= 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-2}(\mathbf{1})\right)k + 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\sum_{1\le j<i\le k}\Psi^{j}\mathcal{L}_\infty^{-1}(\mathbf{1})\right) \\
&= 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-2}(\mathbf{1})\right)k \\
&\quad + 2\sum_{i=2}^{k}\sum_{j=1}^{i-1}\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}(\mathbf{1})\right)^2 \\
&\quad + 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\sum_{i=2}^{k}\sum_{j=1}^{i-1}\Psi^{j}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right)
\end{aligned}$$

where to arrive at the third line, after Ψ^j we have inserted $\Pi_\varsigma + \mathbf{I}_{M_d(\mathbb{C})} - \Pi_\varsigma$, where Π_ς is the projection onto $\mathbf{1}$. Using the fact that $\sum_{i=2}^{k}\sum_{j=1}^{i-1} 1 = \frac{k}{2}(k-1)$, and recalling Equation (8.4) for the expression for the first moment, hence

$$\begin{aligned}
\frac{V_{\varsigma,T_{\mathfrak{A}}}(k)}{k} &= 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-2}(\mathbf{1})\right) \\
&\quad - \mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}(\mathbf{1})\right)^2 \\
&\quad + 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\frac{\Psi}{\mathbf{I}_{M_d(\mathbb{C})}-\Psi}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right) \\
&\quad - \frac{2}{k}\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\frac{\Psi-\Psi^{k+1}}{(\mathbf{I}_{M_d(\mathbb{C})}-\Psi)^2}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right).
\end{aligned}$$

Finally, we can again place $\Pi_\varsigma + \mathbf{I}_{M_d(\mathbb{C})} - \Pi_\varsigma$ in the first term, in between the two $\mathcal{L}_\infty^{-1}$. Rearranging this gives the final result, $\forall k \geq 0$

$$\begin{aligned}
\frac{V_{\varsigma,T_{\mathfrak{A}}}(k)}{k} &= \mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}(\mathbf{1})\right)^2 \\
&\quad + 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathbf{I}_{M_d(\mathbb{C})}-\Psi}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right) \\
&\quad - \frac{2}{k}\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\frac{\Psi-\Psi^{k+1}}{(\mathbf{I}_{M_d(\mathbb{C})}-\Psi)^2}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right).
\end{aligned}$$

□

8.2 LARGE DEVIATION PRINCIPLE FOR GENERAL COUNTING OBSERVABLES

The fixed-time and fixed-observable ensembles are equivalent in the quantum framework [49]. In this section we rigorously prove a large deviation

principle for first passage times of quantum counting observables. This next lemma will show that when $\mathcal{L}$ generates an irreducible quantum Markov semigroup (Hypothesis 5.3), then the semigroup generated by any perturbation of the form

$$\mathcal{L}_s = \mathcal{L}_0 + \sum_{i \in I} e^{a_i s} \mathcal{J}, \quad a_i, s \in \mathbb{R} \tag{8.5}$$

generates an irreducible semigroup as well, in the sense that for every $t > 0$, $e^{t\mathcal{L}_s}$ is irreducible according to the definition in Equation (5.5). In fact, what we will prove is even stronger: it is well known (see for instance [130, Proposition 7.5]) that the irreducibility of $e^{t\mathcal{L}}$ for $t > 0$ is equivalent to the (a priori stronger) property that $e^{t\mathcal{L}_*}(\rho) > 0$ for every $t > 0$ and every initial state ρ; we will show that such property, often called primitivity, is owned by the semigroup generated by $\mathcal{L}_s$ as well.

Lemma 8.4. *If $\mathcal{L}$ satisfies Hypothesis 5.3, then $\mathcal{L}_s$ generates a primitive semigroup.*

Proof. Let us consider a vector $v \in \mathbb{C}^d$ and a state ρ such that $v \in \ker(e^{t\mathcal{L}_{s*}}(\rho))$ for some $t > 0$. Then, using the Dyson series, one can easily see that $\langle v, e^{t\mathcal{L}_{s*}}(\rho)v\rangle = 0$ implies that $\langle v, e^{t\mathcal{L}_{0*}}(\rho)v\rangle = 0$ and for every $k \geq 1$ and $i_1, \dots, i_k \in I$

$$e^{s\sum_{j=1}^k a_{i_j}} \int_{\sum_{j=1}^k t_j \leq t} \langle v, e^{\left(t-\sum_{j=1}^k t_j\right)\mathcal{L}_{0*}} \mathcal{J}_{i_k} \cdots \mathcal{J}_{i_1} e^{t_1 \mathcal{L}_{0*}}(\rho)v\rangle dt_1 \cdots dt_k = 0,$$

which is equivalent (since $e^{s\sum_{j=1}^k a_{i_j}} > 0$) to

$$\int_{\sum_{j=1}^k t_j \leq t} \langle v, e^{\left(t-\sum_{j=1}^k t_j\right)\mathcal{L}_{0*}} \mathcal{J}_{i_k} \cdots \mathcal{J}_{i_1} e^{t_1 \mathcal{L}_{0*}}(\rho)v\rangle dt_1 \cdots dt_k = 0.$$

However, using now the Dyson series for $e^{t\mathcal{L}_{0*}}(\rho)$, one sees that the equations above imply that $\langle v, e^{t\mathcal{L}_{0*}}(\rho)v\rangle = 0$. Since $e^{t\mathcal{L}}$ is primitive, this implies that $v = 0$ and we are done.

□

Theorem 8.5. *Consider a non-empty subset $\mathfrak{A}$ of the emission channels. The FPT $T_{\mathfrak{A}}(k)/k$ satisfies a large deviation principle with good rate function given by*

$$I_{\mathfrak{A}}(t) := \sup_{u \in \mathbb{R}} \{ut - \log(r(u))\}$$

where

$$r(u) = \begin{cases} r\left(\Psi_u\right) & \text{if } u < \overline{\lambda} \\ +\infty & \text{o.w.} \end{cases}$$

where $\Psi_u(x) := -(u + \mathcal{L}_\infty)^{-1} \mathcal{J}_{\mathfrak{A}}(x)$ and $\overline{\lambda} := -\max\{\Re(z) : z \in \mathrm{Sp}(\mathcal{L}_\infty)\}$.

Proof. The proof follows the same method as the proof of the classical case, Theorem 7.4; for completeness we write the quantum proof in full. Lemma 8.4 states that in the domain $u < \overline{\lambda}$, then

$$\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}] = \operatorname{tr}\left(\rho\Psi_u^k(\mathbf{1})\right),$$

where $\Psi_u(x) := -(u + \mathcal{L}_\infty)^{-1}\mathcal{J}_{\mathfrak{A}}(x)$. Writing Ψ_u as the integral

$$\Psi_u(x) = \int_0^\infty e^{(u+\mathcal{L}_\infty)t}\mathcal{J}(x)dt$$

which is the composition of two completely positive maps, hence Ψ_u is completely positive as well. Therefore, Perron-Frobenius theorem tells us that $r(u) := r(\Psi_u)$ is an eigenvalue of Ψ_u, with a positive eigenvector $x(u)$. We can relate this operator with the generator of a quantum dynamical semigroup:

$$\Psi_u(x(u)) = r(u)x(u) \Leftrightarrow \mathcal{L}_{s(u)}(x(u)) = -ux(u),$$

where $\mathcal{L}_{s(u)} := \mathcal{L} + (e^{s(u)} - 1)\mathcal{J}_{\mathfrak{A}}$ and $s(u) := -\log(r(u))$. Notice that $\mathcal{L}_{s(u)}$ has the form in Equation (8.5). Therefore, by Lemma 8.4 it is irreducible and $x(u)$ is in fact a unique and strictly positive eigenvector of $\mathcal{L}_{s(u)}$ corresponding to the eigenvalue $-u$.

We first need to show that for $u < \overline{\lambda}$

$$\lim_{k\to+\infty}\frac{1}{k}\log(\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}]) < \infty, \tag{8.6}$$

Using Holder's inequality, we have that $\operatorname{tr}\left(\rho\Psi_u^k(\mathbf{1})\right) \leq \|\Psi_u^k\|_{\infty\to\infty}$, hence due to Gelfand's formula

$$\frac{1}{k}\log(\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}]) \leq \frac{1}{k}\log(\|\Psi_u^k\|_{\infty\to\infty}) \xrightarrow[k\to+\infty]{} \log(r(u)).$$

Furthermore, we can take $\|x(u)\|_{\infty\to\infty} \leq 1$ and bound from below:

$$\frac{1}{k}\log(\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}]) \geq \log(r(u)) + \frac{1}{k}\log(\operatorname{tr}(\rho x(u))) \xrightarrow[k\to+\infty]{} \log(r(u)),$$

which we can do since $x(u)$ is strictly positive, ρ is positive semidefinite so $\operatorname{tr}(\rho x(u)) > 0$. Therefore, we have shown that the limit in Equation (8.6) converges to $\log(r(u)) < \infty$ in the range $u < \overline{\lambda}$.

We now use the same version of the Gärtner-Ellis theorem (Theorem 3.7) as in the proof of Theorem 7.4. All that remains is to show that $\log(r(u))$ is steep, i.e. as u approaches the boundary $\overline{\lambda}$, both $\log(r(u))$ and $\log(r(u))'$ diverge to $+\infty$. Denote by $\mathbf{T}$ the spectral projection of $\mathcal{L}_\infty$ with respect to the eigenvalue $-\overline{\lambda}$. We can show that $\mathcal{L}_\infty\mathbf{T}$ is diagonalisable, i.e. the restriction does not feature any Jordan blocks or equivalently, the algebraic and geometric multiplicity of $\overline{\lambda}$ coincide. To do this, assume $\mathcal{L}_\infty\mathbf{T}$ is not diagonalisable. Then the map $\mathcal{L}'\mathbf{T} = (\mathcal{L} + \overline{\lambda}\mathbf{I}_{M_d(\mathbb{C})})\mathbf{T}$ — corresponding to the eigenvalue 0 — is not diagonalisable. One can then always choose $x, y \in M_d(\mathbb{C})$ with $\mathcal{L}'(y) = 0$ and $\mathcal{L}'(x) = y$. Therefore one has

$$e^{t\mathcal{L}'}(x) = x + ty.$$

The semigroup $e^{t\mathcal{L}'}$ is in fact a contraction semigroup, which contradicts the above equation. Therefore, we have

$$\mathcal{L}_\infty \mathbf{T} = -\overline{\lambda}\mathbf{T}.$$

Firstly, we show that $\lim_{u\to\overline{\lambda}^-} r(u) = +\infty$. Note that $\mathbf{T}\Psi \neq 0$ since $\mathbf{T}\Psi(\mathbf{1}) \neq 0$. The map Ψ_u can be written as

$$\Psi_u = \frac{\mathcal{L}_\infty}{\mathcal{L}_\infty + u} = \frac{\overline{\lambda}}{\overline{\lambda} - u}\mathbf{T}\Psi + (\mathbf{I}_{M_d(\mathbb{C})} - \mathbf{T})\frac{\mathcal{L}_\infty}{\mathcal{L}_\infty + u}\Psi$$

and one can see it has a norm which explodes as $u \to \overline{\lambda}^-$. Let us assume by contradiction that $r(u) \to r(\overline{\lambda}^-) < \infty$ as $u \to \overline{\lambda}^-$. Then the eigenvector $x(u)$ can be chosen to converge to the Perron-Frobenius eigenvector of $\mathcal{L}_{s(\overline{\lambda})}$ and $\min \mathrm{Sp}(x(u)) \not\to 0$. We then have for $0 \leq u < \overline{\lambda}$

$$\begin{aligned}
\|\Psi_u\|_{\infty\to\infty} &= \|\Psi_u(\mathbf{1})\|_\infty \\
&\leq \frac{1}{\min \mathrm{Sp}(x(u))}\|\Psi_u(x(u))\|_\infty \\
&= r(u)\frac{\|x(u)\|_\infty}{\min \mathrm{Sp}(x(u))} \\
&< +\infty
\end{aligned}$$

which is a contradiction. Let us show that $\lim_{u\to\overline{\lambda}^-} \log(r(u))' = +\infty$. For each u, we can choose the left eigenvector of Ψ_u, $l(u)$, to be such that $\mathrm{tr}(l(u)x(u)) \equiv 1$. Then we can write $r(u)$ as

$$\begin{aligned}
r(u) &= -\mathrm{tr}\left(l(u)(u+\mathcal{L}_\infty)^{-1}\mathcal{L}_\infty\Psi(x(u))\right) \\
&= \underbrace{-\mathrm{tr}\left(l(u)(u+\mathcal{L}_\infty)^{-1}\mathcal{L}_\infty\mathbf{T}\Psi(x(u))\right)}_{(I)} \\
&+ \underbrace{-\mathrm{tr}\left(l(u)(u+\mathcal{L}_\infty)^{-1}\mathcal{L}_\infty(\mathbf{I}_{M_d(\mathbb{C})} - \mathbf{T})\Psi(x(u))\right)}_{(II)}.
\end{aligned}$$

The term (II) remains bounded, hence as $u \to \overline{\lambda}^-$

$$r(u) \asymp \frac{\overline{\lambda}}{\overline{\lambda} - u}\mathrm{tr}\left(l(u)\mathbf{T}\Psi(x(u))\right) \to +\infty.$$

Since $\log(r(u))' = \frac{r'(u)}{r(u)}$, we can differentiate the expression for $r(u)$ to obtain

$$\frac{r'(u)}{r(u)} = \frac{\operatorname{tr}\left(l(u)(u+\mathcal{L}_\infty)^{-2}\mathcal{L}_\infty\Psi(x(u))\right)}{r(u)}$$
$$= \underbrace{\frac{\overline{\lambda}}{(\overline{\lambda}-u)^2}\frac{\operatorname{tr}\left(l(u)\mathbf{T}\Psi(x(u))\right)}{r(u)}}_{(I)} + \underbrace{\frac{\operatorname{tr}\left(l(u)(\mathbf{I}_{M_d(\mathbb{C})}-\mathbf{T})\Psi(x(u))\right)}{r(u)}}_{(II)}.$$

We have again that part (II) is bounded, but for $u \to \overline{\lambda}^-$, $(I) \asymp (\overline{\lambda}-u)^{-1}$ which completes the proof.

□

In addition to obeying a large deviation principle, we are interested in the fluctuations for a finite number of counts. We will derive bounds for the finite regime in the rest of this chapter.

8.3 CONCENTRATION INEQUALITY FOR TOTAL COUNTS

The FPT $T_I(k)$ for the total number of counts $N_{I,t}$ is the time it takes to observe k counts of any kind on the system. Our first main result for quantum Markov processes is a quantum version of Theorem 7.7: a bound on the fluctuations of the FPT $T_I(k)$ for total jumps. We note that in the quantum framework, "activity" is usually referred to as total "counts" or "jumps" [62] (but other definitions exist [131]). From Equation (8.2), the asymptotic mean in this case is

$$\langle t_I \rangle := -\operatorname{tr}(\sigma\mathcal{L}_0^{-1}(\mathbf{1})).$$

We define

$$c_q := \|\mathcal{L}_0^{-1}\|_{\sigma\to\sigma} \tag{8.7}$$

and note that this is the non-commutative counterpart of c_c, cf. Equation (7.16).

Theorem 8.6. *Assume that Hypothesis 5.4 holds (Φ be irreducible) and let ε be the absolute spectral gap of Φ. Then, for every $\gamma > 0$:*

$$\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \geq \langle t_I\rangle + \gamma\right) \leq C(\rho)\exp\left(-k\frac{\gamma^2\varepsilon}{8c_q^2}h\left(\frac{5\gamma}{2c_q}\right)\right)$$

and

$$\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \leq \langle t_I\rangle - \gamma\right) \leq C(\rho)\exp\left(-k\frac{\gamma^2\varepsilon}{8c_q^2}h\left(\frac{5\gamma}{2c_q}\right)\right), \quad k\in\mathbb{N},$$

where $h(x) := (\sqrt{1+x}+\frac{x}{2}+1)^{-1}$, $C(\rho) := \left\|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\right\|_\sigma$ and c_q is defined in Equation (8.7).

Proof. Applying Chernoff bound, we get, for $u \geq 0$

$$\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \geq \langle t_I \rangle + \gamma\right) \leq e^{-uk(\langle t_I \rangle + \gamma)} \mathbb{E}_\rho\left[e^{uT_I(k)}\right]. \tag{8.8}$$

For $u < \overline{\lambda}$, we define the tilted operator $\Phi_u(x) := \mathcal{F}_u \Phi(x)$, with $\mathcal{F}_u(x) := (u\mathbf{I}_{M_d(\mathbb{C})} + \mathcal{L}_0)^{-1}\mathcal{L}_0(x)$. Using Lemma 8.2, for $0 \leq u < \overline{\lambda}$ we can write

$$\begin{aligned}\mathbb{E}_\rho[e^{uT_I(k)}] = \operatorname{tr}\left(\rho \Phi_u^k(\mathbf{1})\right) &= \left\langle \sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}, \Phi_u^k(\mathbf{1})\right\rangle_\sigma \\ &\leq \underbrace{\|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\|_\sigma}_{:=C(\rho)} \underbrace{\|\mathbf{1}\|_\sigma}_{=1} \|\Phi_u^k\|_\sigma,\end{aligned}$$

where with a small abuse of notation $\|\Phi_u^k\|_\sigma$ denotes the operator norm of the map Φ_u^k with respect to the Kubo-Martin-Schwinger (KMS) inner product associated to σ. We can further break up this operator norm

$$\|\Phi_u^k\|_\sigma \leq \|\Phi_u\|_\sigma^k = \|\mathcal{F}_u\Phi\|_\sigma^k.$$

We now seek to upper bound $\|\mathcal{F}_u\Phi_u\|_\sigma$. Conversely to the classical case, $\mathcal{F}_u\Phi$ is not self-adjoint, but we can upper bound its norm with an operator which is. This can be done using Lemma 7.6 with $\mathbf{A} = \mathcal{F}_u, \mathbf{B} = \mathbf{I}_{M_d(\mathbb{C})}$ to get $\|\mathcal{F}_u\Phi\|_\sigma \leq \|\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger\|_{\hat{\sigma}}^{\frac{1}{2}}$, with $\mathcal{F}_u^\dagger(x) = \Gamma_\sigma^{-\frac{1}{2}} \circ (\mathbf{I}_{M_d(\mathbb{C})} + u\mathcal{L}_{0*}^{-1})^{-1} \circ \Gamma_\sigma^{\frac{1}{2}}(x)$, cf. Equation (5.6). Since $\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger$ is a positive, self-adjoint, irreducible map, operator Perron-Frobenius theory (cf. Theorem 5.5) says $\|\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger\|_\sigma = r(u)$, where $r(u) = \sup\{|\lambda| : \lambda \in \mathrm{Sp}(\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger)\}$ is the spectral radius of $\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger$. Hence, the Laplace transform is upper bounded by

$$\mathbb{E}_\rho[e^{uT_I(k)}] \leq C(\rho) r(u)^{\frac{k}{2}}. \tag{8.9}$$

For u small enough, we can expand $\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger$ as the power series

$$\begin{aligned}\mathcal{F}_u\hat{\Phi}\mathcal{F}_u^\dagger &= \sum_{j\geq 0} u^j \left(-\frac{1}{\mathcal{L}_0}\right)^j \circ \hat{\Phi} \circ \sum_{l\geq 0} u^l \left(-\frac{1}{\mathcal{L}_0^\dagger}\right)^l \\ &= \sum_{l\geq 0} u^l \underbrace{\sum_{j=0}^{l} \left(-\frac{1}{\mathcal{L}_0}\right)^{l-j} \circ \hat{\Phi} \circ \left(-\frac{1}{\mathcal{L}_0^\dagger}\right)^j}_{\Phi^{(l)}},\end{aligned}$$

which is the quantum analogue of Equation (7.19) from the classical case. Again we bound each term in the power series in order to use operator perturbation theory [105] (cf. Section 4.4). Using the definition of c_q in Equation (8.7) and Cauchy-Schwarz, we can upper bound $\Phi^{(l)}$ in the form

$$\|\Phi^{(l)}\|_\sigma \leq (l+1)c_q^l \leq (2c_q)^l.$$

Using perturbation theory, we find that for $u < \frac{\varepsilon}{2c_q(2+\varepsilon)}$, the spectral radius $r(u)$ can then be expressed as

$$r(u) = 1 + \sum_{l=1}^{\infty} u^l r^{(l)}, \tag{8.10}$$

where

$$r^{(l)} = \sum_{p=1}^{l} \frac{(-1)^p}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=l,\, \nu_i\geq 1 \\ \mu_1+\cdots+\mu_p=p-1,\, \mu_j\geq 0}} \mathrm{TR}\left(\Phi^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\Phi^{(\nu_p)}\mathbf{S}^{(\mu_p)}\right). \tag{8.11}$$

We have $\mathbf{S}^{(0)} = -\Pi_\sigma$, $\mathbf{S}^{(1)} = (\hat{\Phi} - \mathbf{I}_{M_d(\mathbb{C})} + \Pi_\sigma)^{-1} - \Pi_\sigma = -\epsilon^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\sigma)$, and $\mathbf{S}^{(\mu)}$ the μ^{th} power of $\mathbf{S}^{(1)}$, and also $\|\mathbf{S}^{(\mu)}\|_\sigma = \varepsilon^{-\mu}$ for $\mu \geq 1$. We now want to bound the terms in the expression for $r^{(l)}$. For $p = 1$

$$\begin{aligned}
\left|\mathrm{TR}\left(\Phi^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\Phi^{(\nu_p)}\mathbf{S}^{(\mu_p)}\right)\right| &= \left|\mathrm{TR}\left(\Phi^{(l)}\mathbf{S}^{(0)}\right)\right| \\
&= \left|\left\langle \mathbf{1}, \Phi^{(l)}(\mathbf{1})\right\rangle_\sigma\right| \\
&\leq \|\mathbf{1}\|_\sigma \|\Phi^{(l)}\|_\sigma \\
&\leq (2c_q)^l \\
&\leq 2c_q \left(\frac{2c_q}{\varepsilon}\right)^{l-1}
\end{aligned} \tag{8.12}$$

since $\varepsilon \leq 1$. For $p \geq 2$, using the fact that one of μ_i is zero and using trace cyclicity

$$\begin{aligned}
\left|\mathrm{TR}\left(\Phi^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\Phi^{(\nu_p)}\mathbf{S}^{(\mu_p)}\right)\right| &= \left|\left\langle \mathbf{1}, \Phi^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\mathbf{S}^{(\mu_{p-1})}\Phi^{(\nu_p)}(\mathbf{1})\right\rangle_\sigma\right| \\
&\leq \frac{1}{\varepsilon^{p-1}}(2c_q)^l \\
&\leq \frac{1}{\varepsilon^{l-1}}(2c_q)^l \\
&\leq 2c_q \left(\frac{2c_q}{\varepsilon}\right)^{l-1}.
\end{aligned} \tag{8.13}$$

We can explicitly calculate $r^{(1)}$

$$\begin{aligned}
r^{(1)} &= -\mathrm{TR}\left(\Phi^{(1)}(-\Pi_\sigma)\right) = \left\langle \mathbf{1}, \Phi^{(1)}(\mathbf{1})\right\rangle_\sigma \\
&= \left\langle \mathbf{1}, \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)\hat{\Phi}(\mathbf{1})\right\rangle_\sigma + \left\langle \mathbf{1}, \hat{\Phi}\left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0^\dagger}\right)(\mathbf{1})\right\rangle_\sigma = 2\langle t_I\rangle.
\end{aligned}$$

The term $|r^{(2)}|$ can be bounded using Equations (8.12) and (8.13)

$$\begin{aligned}
|r^{(2)}| &= \left|\mathrm{TR}\left(\Phi^{(2)}\mathbf{S}^{(0)}\right) - \left\langle \mathbf{1}, \Phi^{(1)}\mathbf{S}^{(1)}\Phi^{(1)}(\mathbf{1})\right\rangle_\sigma\right| \\
&\leq \frac{(2c_q)^2}{\varepsilon} + \frac{(2c_q)^2}{\varepsilon} = \frac{8c_q^2}{\varepsilon}.
\end{aligned}$$

Using the bound on the number of terms in (8.11), given by Equation (7.22), we can bound the $r^{(l)}$ by $|r^{(l)}| \leq \frac{2c_q}{5}\left(\frac{10c_q}{\varepsilon}\right)^{l-1}$. Now we have all the ingredients in place, we can finish bounding $r(u)$

$$
\begin{aligned}
r(u) &\leq 1 + \sum_{l \geq 1} u^l |r^{(l)}| \\
&\leq 1 + 2\langle t_I \rangle u + \frac{8c_q^2}{\varepsilon} u^2 + \sum_{l \geq 3} u^l \frac{2c_q}{5} \left(\frac{10c_q}{\varepsilon} \right)^{l-1} \\
&= 1 + 2\langle t_I \rangle u + \frac{8c_q^2}{\varepsilon} u^2 + u^2 \frac{4c_q^2}{\varepsilon} \sum_{l \geq 1} u^l \left(\frac{10c_q}{\varepsilon} \right)^l \\
&\leq 1 + 2\langle t_I \rangle u + \frac{8c_q^2}{\varepsilon} u^2 \left(1 - \frac{10c_q u}{\varepsilon} \right)^{-1} \\
&\leq \exp\left(2\langle t_I \rangle u + \frac{8c_q^2}{\varepsilon} u^2 \left(1 - \frac{10c_q u}{\varepsilon} \right)^{-1} \right),
\end{aligned}
$$

which is valid for $u < \frac{\varepsilon}{10c_q} < \frac{\varepsilon}{2c_q(2+\varepsilon)}$. Putting this upper bound on $r(u)$ back into Equation (8.9) gives the upper bound on the Laplace transform of $T_I(k)$

$$
\mathbb{E}_\rho[e^{uT_I(k)}] \leq C(\rho) \exp\left(k \left(\langle t_I \rangle u + \frac{4c_q^2 u^2}{\varepsilon} \left(1 - \frac{10c_q u}{\varepsilon} \right)^{-1} \right) \right).
$$

Applying the Chernoff bound in Equation (8.8) gives, $\forall u \in [0, \frac{\varepsilon}{10c_q})$

$$
\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \geq \langle t_I \rangle + \gamma \right) \leq C(\rho) \exp\left(-k \left(\gamma u - \frac{4c_q^2 u^2}{\varepsilon} \left(1 - \frac{10c_q u}{\varepsilon} \right)^{-1} \right) \right).
$$

Optimisation over the allowed u, using Equation (7.25), with $\alpha = \frac{4c_q^2}{\varepsilon}$, $\beta = \frac{10c_q}{\varepsilon}$ gives the final concentration inequality for right deviations

$$
\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \geq \langle t_I \rangle + \gamma \right) \leq C(\rho) \exp\left(-k \frac{\gamma^2 \varepsilon}{8c_q^2} h\left(\frac{5\gamma}{2c_q} \right) \right).
$$

By considering $u \leq 0$ we can prove the bound for left deviations

$$
\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \leq \langle t_I \rangle - \gamma \right) \leq e^{-ku(\langle t_I \rangle - \gamma)} \mathbb{E}_\rho[e^{uT_I(k)}].
$$

Repeating the process we get a similar bound on the Laplace transform, valid for $0 \leq |u| < \frac{\varepsilon}{10c_q}$

$$
\mathbb{E}_\rho[e^{uT_I(k)}] \leq C(\rho) \exp\left(k \left(\langle t_I \rangle u + \frac{4c_q^2 u^2}{\varepsilon} \left(1 - \frac{10c_q |u|}{\varepsilon} \right)^{-1} \right) \right).
$$

One obtains the upper bound on left deviations, a symmetric bound to right deviations

$$
\mathbb{P}_\rho\left(\frac{T_I(k)}{k} \leq \langle t_I \rangle - \gamma \right) \leq C(\rho) \exp\left(-k \frac{\gamma^2 \varepsilon}{8c_q^2} h\left(\frac{5\gamma}{2c_q} \right) \right).
$$

□

Theorem 8.6 provides a concentration inequality in terms of the absolute spectral gap ε and the operator norm c_q. As long as Hypothesis 5.4 is satisfied, it is applicable with any jump operators L_i. As in the classical case, the following proposition follows from the proof of Theorem 8.6, where we bound only the terms in the expression for $r(u)$ which correspond to the variance.

Proposition 8.6.1. *The variance $V_{\sigma,T_I}(k)$ of the first passage time for total counts is bounded from above by*

$$\frac{V_{\sigma,T_I}(k)}{k} \leq \left(\frac{4}{\varepsilon} - (1-\varepsilon)\right) c_q^2.$$

Proof. Notice that for $u \geq 0$ small enough, one has

$$\begin{aligned}\log(\mathbb{E}_\sigma[e^{uT_I(k)}]) &= \langle t_I\rangle ku + \frac{1}{2}V_{\sigma,T_I}(k)u^2 + o(u^2)\\ &\leq \frac{k}{2}\log(r(u)) = \frac{k}{2}r'(0)u + \frac{k}{4}(r''(0) - (r'(0))^2)u^2 + o(u^2)\end{aligned}$$

where $r(u)$ is given by Equation (8.10). We recall that

$$r'(0) = r^{(1)} = 2\langle t_I\rangle = -2\left\langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{1})\right\rangle_\sigma$$

and

$$\begin{aligned}\frac{r''(0)}{2} &= r^{(2)} = \langle \mathbf{1}, \Phi^{(2)}(\mathbf{1})\rangle_\sigma + \left\langle \mathbf{1}, \Phi^{(1)}(\mathbf{I}_{M_d(\mathbb{C})} - \hat{\Phi})^{-1}\Phi^{(1)}(\mathbf{1})\right\rangle_\sigma\\ &= 2\langle \mathbf{1}, \mathcal{L}_0^{-2}(\mathbf{1})\rangle_\sigma + \langle \mathbf{1}, \mathcal{L}_0^{-1}\hat{\Phi}(\mathcal{L}_0^\dagger)^{-1}(\mathbf{1})\rangle_\sigma\\ &\quad + \langle \mathbf{1}, ((\mathcal{L}_0^\dagger)^{-1} + \mathcal{L}_0^{-1}\hat{\Phi})(\mathbf{I}_{M_d(\mathbb{C})} - \hat{\Phi})^{-1}(\hat{\Phi}(\mathcal{L}_0^\dagger)^{-1} + \mathcal{L}_0^{-1})(\mathbf{1})\rangle_\sigma\\ &= 2\langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{1})\rangle_\sigma^2 + 2\langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\sigma)\mathcal{L}_0^{-1}(\mathbf{1})\rangle_\sigma\\ &\quad + \langle \mathbf{1}, \mathcal{L}_0^{-1}\hat{\Phi}(\mathcal{L}_0^\dagger)^{-1}(\mathbf{1})\rangle_\sigma + \varepsilon^{-1}\langle \mathbf{1}, (\mathcal{L}_0^\dagger)^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \sigma)\mathcal{L}_0^{-1}(\mathbf{1})\rangle_\sigma\\ &\quad + (1-\varepsilon)\varepsilon^{-1}(\langle \mathbf{1}, (\mathcal{L}_0^\dagger)^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\sigma)(\mathcal{L}_0^\dagger)^{-1}(\mathbf{1})\rangle_\sigma\\ &\quad + \langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\sigma)\mathcal{L}_0^{-1}(\mathbf{1})\rangle_\sigma)\\ &\quad + (1-\varepsilon)^2\varepsilon^{-1}\langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{I}_{M_d(\mathbb{C})} - \Pi_\sigma)(\mathcal{L}_0^\dagger)^{-1}(\mathbf{1})\rangle_\sigma\\ &\leq 2\langle \mathbf{1}, \mathcal{L}_0^{-1}(\mathbf{1})\rangle_\sigma^2 + \left(\frac{4}{\varepsilon} - (1-\varepsilon)\right)c_q^2.\end{aligned}$$

Therefore one has

$$\frac{r''(0) - (r'(0))^2}{2} \leq \left(\frac{4}{\varepsilon} - (1-\varepsilon)\right)c_q^2$$

and we proved the statement.

□

Recall from Lemma 8.1 that if Hypothesis 5.3 holds ($\mathcal{L}_*$ admits a unique and strictly positive invariant state), then the uniqueness of the invariant state of Φ_* is guaranteed, but not its strict positivity, hence the need for Hypothesis 5.4. One can show however, that if Hypothesis 5.3 holds, then the invariant state of Φ_* is strictly positive if and only if $\cap_{i=1}^{|I|}\ker(L_i^*) = \emptyset$.

8.4 CONCENTRATION INEQUALITY FOR TOTAL COUNTS OF RESET PROCESSES

In Theorem 8.6 we proved a concentration bound for the FPT corresponding to the total number of counts, under the assumption that Hypothesis 5.4 holds. In this section we consider *quantum reset processes* which are characterised by jump operators that have rank one, and we derive a FPT concentration bound using the weaker Hypothesis 5.3.

Let us assume that the jump operators are of the form

$$L_i = |y_i\rangle \langle x_i| \quad x_i, y_i \in \mathbb{C}^d \setminus \{0\}. \tag{8.14}$$

Without any loss of generality, we can assume that $\|y_i\| = 1$. After observing a click of the i-th detector, the state of the system is known and is equal to $|y_i\rangle \langle y_i|$. In this case, by applying step 2 of the iterative procedure in Section 5.2.3 we find that the sequence of click indices is a classical Markov chain on I with transition matrix $\mathbf{P} := (p_{ij})$

$$p_{ij} = -\langle x_j|\mathcal{L}_{0*}^{-1}(|y_i\rangle \langle y_i|)|x_j\rangle. \tag{8.15}$$

We remark that Hypothesis 5.3 is sufficient to imply the irreducibility of the classical transition operator $\mathbf{P}$. Indeed since $\mathcal{L}_*(\hat{\sigma}) = 0$ with stationary state $\hat{\sigma} > 0$, we have

$$\mathcal{L}_{0*}(\hat{\sigma}) = -\mathcal{J}_*(\hat{\sigma}) = -\sum_{i\in I}\langle x_i|\hat{\sigma}|x_i\rangle \cdot |y_i\rangle\langle y_i|$$

which implies

$$\sum_{i\in I}\langle x_i|\hat{\sigma}|x_i\rangle p_{ij} = -\sum_{i\in I}\langle x_i|\hat{\sigma}|x_i\rangle\langle x_j|\mathcal{L}_{0*}^{-1}(|y_i\rangle \langle y_i|)|x_j\rangle =$$
$$-\left\langle x_j \middle| \mathcal{L}_{0*}^{-1}\left(\sum_{i\in I}\langle x_i|\hat{\sigma}|x_i\rangle |y_i\rangle \langle y_i|\right)\middle| x_j\right\rangle = \langle x_j|\hat{\sigma}|x_j\rangle$$

so the stationary state of $\mathbf{P}$ is

$$\pi(i) := \frac{\langle x_i|\hat{\sigma}|x_i\rangle}{\sum_{j\in I}\langle x_j|\hat{\sigma}|x_j\rangle}$$

which is fully supported since $\hat{\sigma} > 0$. The holding times are not exponentially distributed as in the case of a classical continuous time Markov process, instead their probability density function after observing a click of the type i is given by

$$f_i(t) := -\mathrm{tr}(\mathcal{L}_{0*}e^{t\mathcal{L}_{0*}}(|y_i\rangle \langle y_i|). \tag{8.16}$$

In fact, quantum reset processes fall under the more general class of semi-Markov processes. These are processes with a discrete jump component driven by a transition matrix, but with arbitrary holding time distributions.

We now introduce the quantities used in the result of this section:

1. asymptotic value of $T_I(k)/k$:

$$\langle t_I \rangle := -\sum_{i\in I} \pi(i) \mathrm{tr}\left(\mathcal{L}_{0*}^{-1}(|y_i\rangle\langle y_i|)\right);$$

2. average of 1-norm of $\mathcal{L}_{0*}^2$ in stationarity:

$$b_r^2 := \sum_{i\in I} \pi(i) \left\|\mathcal{L}_{0*}^{-2}(|y_i\rangle\langle y_i|)\right\|_1;$$

3. superoperator norm of $\mathcal{L}_0^{-1}$:

$$c_r := \left\|\mathcal{L}_0^{-1}\right\|_{\infty\to\infty} = \|\mathcal{L}_0^{-1}(\mathbf{1})\|_\infty.$$

The last equality in the expression of c_r is due to Theorem 5.6 and makes the superoperator norm analytically computable.

Theorem 8.7. *Assume that Hypothesis 5.3 holds ($\mathcal{L}$ be irreducible) the jump operators are of the form* (8.14) *(reset process). Let ε be the spectral gap of $\mathbf{P}^\dagger\mathbf{P}$. For every $\gamma > 0$*

$$\mathbb{P}_\nu\left(\frac{T_I(k)}{k} \geq \langle t_I\rangle + \gamma\right) \leq C(\nu)\exp\left(-k\frac{\gamma^2\varepsilon}{4b_r^2}h\left(\frac{5c_r\gamma}{2b_r^2}\right)\right)$$

and

$$\mathbb{P}_\nu\left(\frac{T_I(k)}{k} \leq \langle t_I\rangle - \gamma\right) \leq C(\nu)\exp\left(-k\frac{\gamma^2\varepsilon}{4b_r^2}h\left(\frac{5c_r\gamma}{2b_r^2}\right)\right), \quad k\in\mathbb{N},$$

where $h(x) := (\sqrt{1+x} + \frac{x}{2} + 1)^{-1}$ and $C(\nu) := \|\frac{\nu}{\pi}\|_\infty$. Here, $\mathbb{P}_\nu$ is the probability measure induced by the initial state given by

$$\sum_{i\in I}\nu(i)\,|y_i\rangle\langle y_i|, \quad \sum_{i\in I}\nu(i) = 1, \quad \nu(i)\geq 0.$$

Proof. The general formula for the Laplace transform of $T_{\mathfrak{e}}(k)$ in the case of reset processes reads

$$\mathbb{E}_\nu[e^{uT_I(k)}] = \left\langle \nu, (\mathbf{F}_u\mathbf{P})^k\underline{1}\right\rangle, \quad \text{for } u < \overline{\lambda}.$$

$\mathbf{P}$ is given by Equation (8.15) and $\mathbf{F}_u$ is a diagonal matrix whose entries are $(\mathbf{F}_u)_{ii} = \mathrm{tr}(|y_i\rangle\langle y_i|\,(\mathbf{I}_{M_d(\mathbb{C})} + u\mathcal{L}_0^{-1})^{-1}(\mathbf{1}))$, $y_i \in \mathbb{C}_d$. $\mathbf{F}_u$ is self-adjoint and we can use Lemmas 7.5 and 7.6 as in the classical case to upper bound the norm $\|\mathbf{P}\mathbf{F}_u\|_\pi$. We seek as before to expand $\hat{\mathbf{P}}\mathbf{F}_u$ ($\hat{\mathbf{P}}$ given by Definition 2.2)

$$\hat{\mathbf{P}}\mathbf{F}_u = \hat{\mathbf{P}} + \sum_{l=1}^{\infty} u^l\hat{\mathbf{P}}\mathbf{D}^{(l)},$$

where $\mathbf{D}^{(l)}$ is a diagonal matrix with entries $\mathrm{tr}\left(|y_i\rangle\langle y_i|\left(-\mathcal{L}_0^{-1}\right)^l(\mathbf{1})\right)$. Note that contrary to the classical case this is not simply some diagonal

matrix $\mathbf{D}$ raised to a power l. The $L^2_\pi(E)$ norm $\|\mathbf{D}^{(l)}\|_\pi = \mathbf{D}^{(l)}_{\max}$, where $\mathbf{D}^{(l)}_{\max}$ is the maximum absolute element in $\mathbf{D}^{(l)}$. Therefore, $\mathbf{D}^{(l)}$ can be upper bounded in $L_2(\pi)$ by

$$\begin{aligned}\|\mathbf{D}^{(l)}\|_\pi &= \sup_{i\in I}\left\{ \operatorname{tr}\left(|y_{\tilde{i}}\rangle \langle y_{\tilde{i}}| \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^l (\mathbf{1})\right)\right\} \\ &\leq \left\| \frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right\|^l_{\infty\to\infty} = c_r^l.\end{aligned} \tag{8.17}$$

Similarly, if we set $b_r^2 = \sum_{i\in I}\pi(i)\left\|\mathcal{L}_{0*}^{-2}(|y_i\rangle\langle y_i|)\right\|_1$, we can upper bound $\|\mathbf{D}^{(l)}\underline{1}\|_\pi$

$$\begin{aligned}\|\mathbf{D}^{(l)}\underline{1}\|^2_\pi &= \sum_{i\in I}\pi(i)\operatorname{tr}\left(|y_i\rangle\langle y_i| \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^l (\mathbf{1})\right)^2 \\ &\leq \sum_{i\in I}\pi(i)\operatorname{tr}\left(|y_i\rangle\langle y_i| \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^l (\mathbf{1})\right) c_r^l \\ &\leq \sum_{i\in I}\pi(i)\operatorname{tr}\left(\left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_{0*}}\right)^2 (|y_i\rangle\langle y_i|) \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^{l-2} (\mathbf{1})\right) c_r^l \\ &\leq b_r^2 c_r^{2(l-1)}.\end{aligned} \tag{8.18}$$

Therefore $\|\mathbf{D}^{(l)}\underline{1}\|_\pi \leq b_r c_r^{l-1}$. The first inequality in the above is achieved since

$$\begin{aligned}\operatorname{tr}\left(|y_i\rangle\langle y_i| \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^l (\mathbf{1})\right) &\leq \sup_i\left\{ \operatorname{tr}\left(|y_i\rangle\langle y_i| \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_0}\right)^l (\mathbf{1})\right)\right\} \\ &\leq c_r^l, \quad \forall i \in I.\end{aligned}$$

The final step which differs slightly from the classical case is in calculating the $r^{(l)}$ from Equation (7.21). We get the same results for $r^{(0)}$, $r^{(1)}$ and $r^{(2)}$, but for $l \geq 3$ and $p = 1$

$$-\operatorname{tr}(\hat{\mathbf{P}}\mathbf{D}^{(l)}(-\Pi_\pi)) = \left\langle \underline{1}, \mathbf{D}^{(l)}\underline{1}\right\rangle_\pi \leq b_r^2 c_r^{l-2},$$

in which we used the same trick as when bounding $\|\mathbf{D}^{(l)}\underline{1}\|_\pi$. For $p \geq 2$

$$\begin{aligned}-\operatorname{tr}\left(\hat{\mathbf{P}}\mathbf{D}^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\mathbf{S}^{(\mu_p)}\right) &= \left\langle \underline{1}, \mathbf{D}^{(\nu_1)}\mathbf{S}^{(\mu_1)}\cdots\hat{\mathbf{P}}\mathbf{D}^{(\nu_p)}\underline{1}\right\rangle_\pi \\ &\leq \|\mathbf{D}^{(\nu_1)}\underline{1}\|_\pi \|\mathbf{D}^{(\nu_p)}\underline{1}\|_\pi \|\mathbf{S}\|_\pi^{p-1}\prod_{i=2}^{p-1}\|\mathbf{D}^{(\nu_i)}\|_\pi \\ &\leq b_r^2 \frac{c_r^{l-2}}{\varepsilon^{l-1}}.\end{aligned}$$

From here the proof for the reset and classical process are identical and so we obtain Theorem 8.7. Note that for left deviations we can indeed repeat the steps in the classical proof.

□

Theorem 8.7 provides an analogous result to the classical dynamical activity result of Theorem 7.7, but instead features quantities which are derived from the operators used in the quantum framework. We now state our upper bound on the variance for reset processes, which is obtained in a similar fashion to Propositions 7.8.1 and 8.8.1.

Proposition 8.7.1. *The variance $V_{\pi,T_I}(k)$ of the first passage time for total counts of reset processes is bounded from above by:*

$$\frac{V_{\pi,T_I}(k)}{k} \leq \left(1+\frac{2}{\varepsilon}\right) b_r^2.$$

Proof. Notice that for $u \geq 0$ small enough, one has

$$\begin{aligned}\log(\mathbb{E}_\pi[e^{uT_I(k)}]) &= \langle t_I\rangle ku + \frac{1}{2}V_{\pi,T_I}(k)u^2 + o(u^2)\\ &\leq k\log(r(u)) = kr'(0)u + \frac{k}{2}(r''(0)-(r'(0))^2)u^2 + o(u^2)\end{aligned}$$

where $r(u)$ is given by Equation (7.20). We recall using the definition of $\mathbf{D}$ from the proof of Theorem 8.7 that for reset processes

$$r'(0) = r^{(1)} = \langle t_I\rangle = \langle \mathbf{D}\underline{1}, \underline{1}\rangle_\pi,$$

$$r''(0) = 2r^{(2)} = 2\langle \mathbf{D}\underline{1}, \mathbf{D}\underline{1}\rangle_\pi + 2\left\langle \mathbf{D}\underline{1}, \frac{\hat{\mathbf{P}}}{\mathbf{1}-\hat{\mathbf{P}}}\mathbf{D}\underline{1}\right\rangle_\pi$$

and remind ourselves that $\|\mathbf{D}\underline{1}\|_\pi \leq b_r$. Therefore

$$r''(0)-(r'(0))^2 = \langle \mathbf{D}\underline{1}, \underline{1}\rangle_\pi^2 + 2\left\langle \mathbf{D}\underline{1}, \frac{\mathbf{1}}{\mathbf{1}-\hat{\mathbf{P}}}\mathbf{D}\underline{1}\right\rangle_\pi \leq \left(1+\frac{2}{\varepsilon}\right) b_r^2.$$

Hence $V_{\pi,T_I}(k) \leq \left(1+\frac{2}{\varepsilon}\right) b_r^2 k$.

□

8.5 TAIL BOUND FOR GENERAL COUNTING OBSERVABLES

Our final result provides the quantum analogue to the bound of Theorem 7.8. We consider the FPT $T_{\mathfrak{A}}(k)$ for the observable $N_{\mathfrak{A},t}$ which counts the number of jumps with label in the subset $\mathfrak{A} \subseteq I$, cf. Equation (5.7). The next result gives an upper bound to the tails of the FPT distribution under the assumption that Hypothesis 5.3 holds.

Recall that we introduced the generator decomposition

$$\mathcal{L} = \mathcal{J}_{\mathfrak{A}} + \mathcal{L}_\infty,$$

where $\mathcal{J}_{\mathfrak{A}}(x) = \sum_{i\in\mathfrak{A}} L_i^* x L_i$. We denote

$$\beta := \left\|\mathcal{L}_{\infty *}^{-1}\right\|_{1\to 1}.$$

Theorem 8.8. *Assume that Hypothesis 5.3 holds ($\mathcal{L}$ be irreducible), and let $\mathfrak{A} \subseteq I$ be non-empty. For every $\gamma > \beta - \langle t_{\mathfrak{A}} \rangle$ and $k \in \mathbb{N}$*

$$\mathbb{P}_\rho \left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}} \rangle + \gamma \right) \leq \exp\left(-k \left(\frac{\gamma + \langle t_{\mathfrak{A}} \rangle - \beta}{\beta} - \log\left(\frac{\gamma + \langle t_{\mathfrak{A}} \rangle}{\beta} \right) \right) \right).$$

As mentioned above, and in contrast to the classical case, in general there is no explicit expression for the $1 \to 1$ norm of a superoperator, but thanks to Theorem 5.6, we know that

$$\beta = \|\mathcal{L}_\infty^{-1}\|_{\infty\to\infty} = \|\mathcal{L}_\infty^{-1}(\mathbf{1})\|_\infty. \tag{8.19}$$

Despite being in general computational demanding, at least there exists an explicit formula for the new expression for β. We can also derive an upper bound on the variance, in terms of β, stated below.

Proof. We once again use Chernoff bound to upper bound the probability in terms of the Laplace transform

$$\mathbb{P}_\rho \left(\frac{T_{\mathfrak{A}}(k)}{k} \geq \langle t_{\mathfrak{A}} \rangle + \gamma \right) \leq e^{-uk(\langle t_{\mathfrak{A}} \rangle + \gamma)} \mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}], \quad u > 0.$$

Next, we write the Laplace transform in terms of the Hilbert-Schmidt inner product

$$\begin{aligned}
\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}] &= \left\langle \rho, \left(\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathbf{I}_{M_d(\mathbb{C})} + \frac{u}{\mathcal{L}_\infty}} \Psi \right)^k (\mathbf{1}) \right\rangle_{\mathrm{HS}} \\
&= \left\langle \rho, \left(\sum_{i=0}^{\infty} u^i \left(-\frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_\infty} \right) \Psi \right)^k (\mathbf{1}) \right\rangle_{\mathrm{HS}}, \quad u < \|\mathcal{L}_{\infty *}^{-1}\|_{1\to 1}^{-1} \\
&\leq \|\rho\|_{1\to 1} \left(\sum_{i=0}^{\infty} u^i \left\| \frac{\mathbf{I}_{M_d(\mathbb{C})}}{\mathcal{L}_{\infty *}} \right\|_{1\to 1} \|\Psi\|_{\infty\to\infty} \right)^k \|\mathbf{1}\|_{\infty\to\infty}.
\end{aligned}$$

We can denote $\beta := \|\mathcal{L}_{\infty *}^{-1}\|_{1\to 1}$ to obtain

$$\mathbb{E}_\rho[e^{uT_{\mathfrak{A}}(k)}] \leq \left(\frac{1}{1-\beta u} \right)^k = \exp\left(-k \log(1-\beta u) \right). \tag{8.20}$$

From here on, the proof is identical to that of Theorem 7.8.

□

Theorem 8.8 provides a bound in terms of only the operator norm β. The case for general counting observables is important as it encapsulates many physical quantum devices, and includes the class of processes where imperfect photon detection (dark counts) is taken in to account. To obtain the variance bound, the exact expression of the variance (Lemma 8.3) can be used to derive an inverse thermodynamic uncertainty relation (TUR)-type bound.

Proposition 8.8.1. *Given any non-empty set of jumps $\mathfrak{A}$, the variance $V_{\varsigma,T_{\mathfrak{A}}}$ of the corresponding first passage time at stationarity is bounded from above by*

$$\frac{V_{\varsigma,T_{\mathfrak{A}}}(k)}{k} \leq \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2,$$

where

$$\tilde{\varepsilon} := 1 - \max\{\|\Psi(x)\|_{\infty\to\infty} : \|x\|_{\infty\to\infty} = 1,\ \mathrm{tr}(\varsigma x) = 0\}.$$

Proof. From the proof of Lemma 8.3 one can see that

$$\begin{aligned}
\frac{V_{\varsigma,T_{\mathfrak{A}}}(k)}{k} &= \mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}(\mathbf{1})\right)^2 + 2\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right)\\
&\quad + \frac{2}{k}\mathrm{tr}\left(\varsigma\mathcal{L}_\infty^{-1}\sum_{i=2}^{k}\sum_{j=1}^{i-1}\Psi^j(\mathbf{I}_{M_d(\mathbb{C})}-\Pi_\varsigma)\mathcal{L}_\infty^{-1}(\mathbf{1})\right)\\
&\leq \left(1+2\left(1+\frac{1}{k}\sum_{i=2}^{k}\sum_{j=1}^{i-1}(1-\tilde{\varepsilon})^j\right)\right)\beta^2\\
&= \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2 - \frac{2((1-\tilde{\varepsilon})-(1-\tilde{\varepsilon})^{k+1})}{k\tilde{\varepsilon}^2}\beta^2\\
&\leq \left(1+\frac{2}{\tilde{\varepsilon}}\right)\beta^2.
\end{aligned}$$

□

8.6 EXAMPLES

8.6.1 *Total counts in three-level emitter with dephasing*

We consider a three-level system, with one dissipative jump and one dephasing channel, as sketched in Figure 8.1 (a). The system has Hamiltonian

$$H = \Omega_{01}(|0\rangle\langle 1| + |1\rangle\langle 0|) + \Omega_{12}(|1\rangle\langle 2| + |2\rangle\langle 1|)$$

and jump operators

$$L_1 = \omega_{12}\,|1\rangle\langle 2|\,, \qquad L_2 = \omega_{02}(|0\rangle\langle 0| - |2\rangle\langle 2|).$$

We count the total number of jumps of both the emission channel and the dephasing channel and compare the lower bounds on the large deviation rate function obtained from Theorems 8.6 and 8.8 with the exact rate function, see Figure 8.1(b). The exact rate function (full/black) is bounded in the entire region by Theorem 8.6 (dashed/blue). Theorem 8.8 for general counting observables allows one to bound (dotted/red) the right tail of the rate function: as in the classical case, cf. Figure 7.2, this tail bound is tighter than the activity bound for large enough deviations.

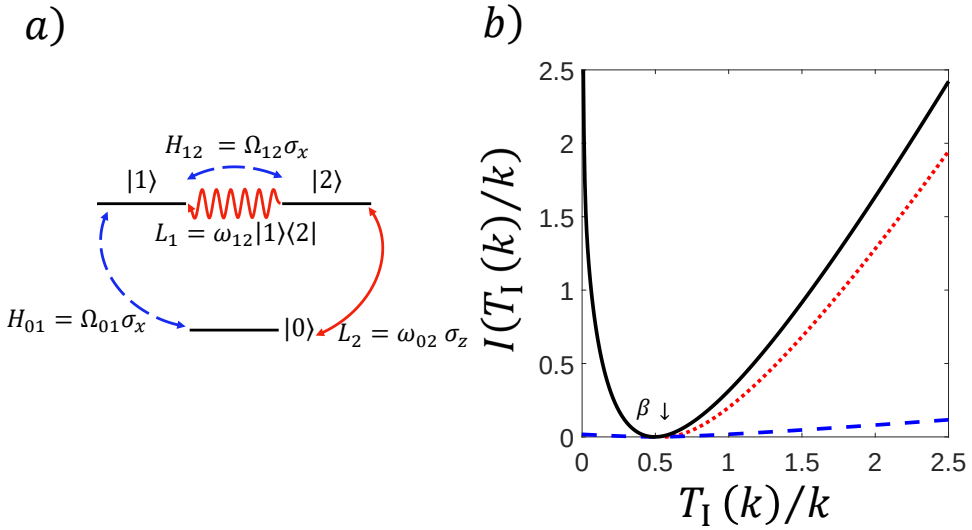

Figure 8.1: Bounds on the rate function of the FPT of the total number of emissions for a quantum three-level system: (a) Sketch of quantum three-level system. The Hamiltonian (dashed/blue) drives the evolution coherently while the jump operators (full/red) give rise to dissipative transitions. (b) Exact rate function $I(T_I(k)/k)$ (full/black) of the FPT for the total number of quantum jumps, for the case $\Omega_{01} = 10$, $\Omega_{12} = 1$, $\omega_{12} = \Omega_{01}$, $\omega_{02} = \frac{1}{5}\Omega_{01}$. Theorem 8.6 gives a lower bound on the entire rate function (dashed/blue). Theorem 8.8 bounds the tail (dotted/red) in the region $T_I(k)/k > \beta$.

8.6.2 *Total counts in two-level emitter*

We illustrate the results of Theorem 8.7 by considering a two level emitter with driving Hamiltonian $H = \Omega_{01}(|0\rangle\langle 1| + |0\rangle\langle 1|)$ and jump operator representing the emitted photon $L = \omega_{01}\,|0\rangle\,\langle 1|$, see Figure 8.2(a). Since L is a rank-one operator, the system jumps to the same state $|0\rangle$ every time there a count. Therefore, the counts process is a renewal process with holding time distribution computed using Equation (8.16). In Figure 8.2(b) we plot the exact rate function (full/black) and two lower bounds, obtained from our reset process bound of Theorem 8.7 (dashed/blue) and the counting observable bound Theorem 8.8 (dotted/red). For comparison to known literature we plot the upper bound on the rate function (dot-dashed/magenta) of reset processes obtained via large deviations [61]. As in the classical example of Section 7.5.1, the bound of Theorem 8.8 outperforms that of Theorem 8.7 for larger deviations. This is because both the expressions for the bounds on the rate function become linear, but the bound from Theorem 8.8 has the steeper slope in this regime. For the two-level emitter, the absolute spectral gap $\varepsilon = 1$, so at least in terms of the spectral quantities, our concentration inequalities and variance upper bound are as small as they can be. To reduce the maximum size of fluctuations further, the system would have to be designed which has a minimal b_r^2.

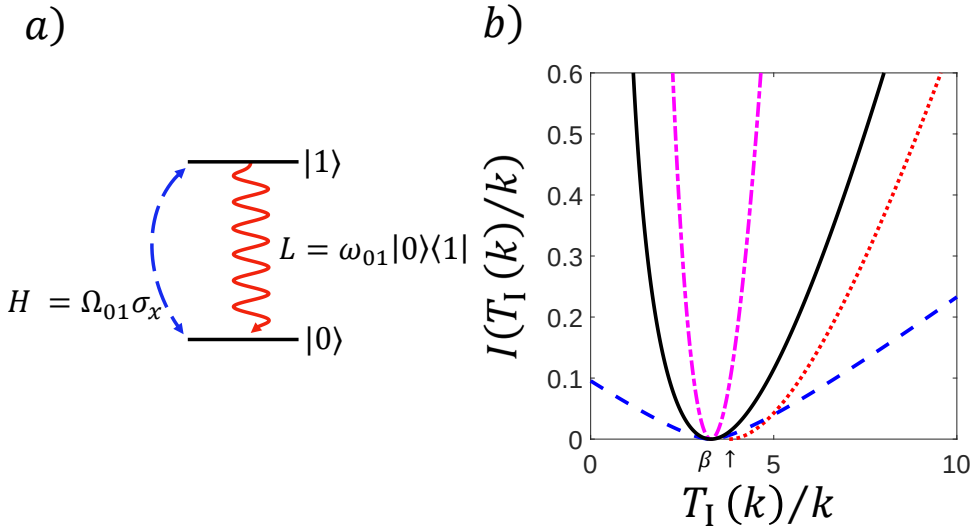

Figure 8.2: Bounds on the rate function of the FPT of the total number of emissions for a two-level emitter: (a) Sketch of two-level emitter model. The Hamiltonian (dashed/blue) drives the evolution coherently while the jump operators (full/red) give rise to dissipative transitions. (b) Exact rate function $I(T_I(k)/k)$ (full/black) of the FPT for the total number of quantum jumps, for the case $\Omega_{01} = 1$, $\omega_{01} = 0.8\Omega_{01}$. As this is a quantum reset process, Theorem 8.7 gives a lower bound on the entire rate function (dashed/blue). Theorem 8.8 bounds the tail (dotted/red) in the region $T_I(k)/k > \beta$. The result from [61] gives an upper bound on the rate function (dash-dotted/magenta).

8.6.3 *Dephasing jumps in three-level emitter*

For our final example we consider a system in which we are only interested in a subset of jumps. We use the same setup as in Section 8.6.1 but this time we only count the number of dephasing jumps (jump operator L_2). In Figure 8.3 we show the exact rate function (full/black) and a lower bound on its right tail from Theorem 8.8 (dotted/red).

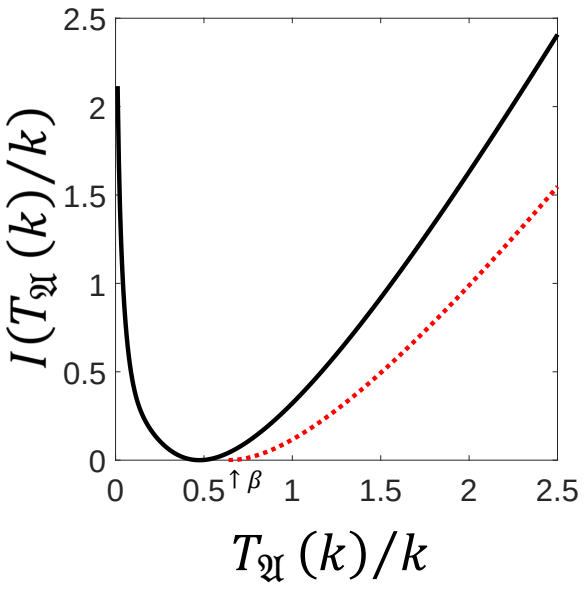

Figure 8.3: Lower bound on the FPT rate function for a counting observable of a quantum three-level system: exact rate function $I(T_{\mathfrak{A}}(k)/k)$ (full/black) of the FPT only counting the dephasing jumps (jump operator by L_2), for the same model of Figure 8.1(a). Theorem 8.8 gives a lower bound on right deviations (dotted/red) in the region $T_{\mathfrak{A}}(k)/k > \beta$.

The models discussed in Section 8.6.2 and 8.6.3 have direct application in parameter estimation. First passage time statistics are used in [112] to test a novel algorithm for fast Bayesian estimation in optomechanical devices. These devices can be used to measure, for example, the signal strength in a magnetic field. Optomechanical devices feature both observed and unobserved detectors, so Theorem 8.8 would be required. The two-level atom is used in [112] to demonstrate the Bayesian estimation algorithm and to derive intuitive results. In the two-level atom case perfect detection is assumed, so Theorem 8.7 could be applied. The variance bounds in Propositions 8.7.1and 8.8.1 correspond to a bound on the error bars of the FPTs in Figure 5 of [112], whilst our concentration inequalities in Theorems 8.7 and 8.8 give a bound on the distribution. Our results could therefore be used to design a setup which best constructs the distribution of the parameter in question, for example, the bounds tell us that systems with a larger absolute spectral gap will give a smaller upper bound on the variance of the FPT.

We generalised the results of Chapter 7 to include quantum Markov processes. Because in this framework we deal with more general operators as opposed to matrices, some of these quantum results are not as practical. For example, there is no quantum equivalent to Proposition 7.8.2 which upper bounds the quantity β used in the variance bound with quantities of the dynamics. Therefore, to compute the upper bound for general counts we must compute β directly from Equation (8.19) and are at the mercy of dimension scaling.

As was the case in the classical setting, the spectral gap of the symmetrised generator also features in the quantum bounds. These spectral quantities have been shown to be relevant in other recent works such as [132, 133]. Our results complement the quantum TUR literature [61, 63, 64, 103] as well as the concentration inequalities for time-integrated quantities for classical (Chapter 6) and quantum dynamics [65].

9

CONCLUSIONS

By employing techniques from probability theory, we have derived a series of concentration results in the study of stochastic dynamics. In Chapter 2 we introduced continuous-time classical Markov processes, which are ubiquitous in this field. The dynamics of Markov processes are typically studied through two distinct but equivalent perspectives: (1) by examining dynamical observables as functions of trajectories over a fixed time t, or (2) by focusing on trajectories where a specific observable reaches a threshold, with interest in the time taken to do so. It is therefore important to tackle problems in both cases, which we have done in this thesis.

We introduced large deviations in Chapter 3 and presented fundamental results necessary to make useful statements in this field. Furthermore, we discussed their relevance to Markov processes, and how bounds on the large deviation rate function correspond to bounds on the fluctuations of the dynamical observables we are interested in. In Chapter 4 we discussed concentration inequalities and outlined the Cramér-Chernoff method which is central to several of our theorems. Chapter 5 outlined how Markov processes and trajectories are described in the non-commutative setting.

By upper bounding the moment generating function of time-integrated observables, in Chapter 6 we derived a direct complement to the thermodynamic uncertainty relations (TURs). Furthermore, we provided a concentration inequality for these quantities. These results hold for all time, and in the large deviation regime (large t) correspond to a lower bound on the rate function. The bounds are in terms of simpler quantities than would be required to compute the probability distribution explicitly and include the maximum escape rate q and the spectral gap of the additive symmetrisation of the generator.

We extended in Chapter 7 these methods to the fixed-observable regime. We began by proving a large deviation principle for first passage times of counting observables, followed by first passage time (FPT) concentration inequalities for dynamical activity and observables counting a subset of jumps. As subsequent propositions, we derived inverse TUR-type upper bounds on the variance of these first passage times.

In our final chapter, Chapter 8, we considered first passage times of quantum Markov processes. We generalised our proof of the large deviation principle of counting observable FPTs to the non-commutative setting. After this we provided concentration inequalities for first passage times of total and subsets of counts. In the case of reset processes, which do not in general satisfy our ergodicity assumptions for general quantum processes, we derived an extra concentration inequality. Again these were followed by upper bounds on their variances.

The most prominent consequence of our results, of the inverse TUR-type, ensure we have both upper *and* lower bounds limiting the range of uncertainty. Therefore, the fluctuations of dynamical observables (and if a counting observable their respective first passage times) cannot be arbitrarily large. A two-sided uncertainty relation is crucial for precision estimation. Instead of computing precision explicitly (although one could do this with our Lemmas 6.1, 7.2 and 8.3), the upper and lower bounds are expressed in terms of operationally accessible quantities. In our upper bounds these quantities are either spectral (spectral gap of additive/multiplicative symmetrisations) or dynamical (such as minimum/maximum escape rates), which could be estimated by observation. The inverse TUR-type bounds also depend on the observable in question, whereas the TURs do not.

The spectral quantities in the upper bounds have recently been found to be relevant in [132, 133] and clearly play a role in fluctuations, but they are not as accessible as those found in the TUR. However, we showed in Figure 6.2 that in many cases our upper bounds on the precision are tighter and therefore can have an advantage in estimation.

Since they are applicable in the finite regime (finite t for fixed-time and finite k for fixed-observable), our inequalities are of interest in experimental setups. The results on first passage times provide upper bounds on fluctuations in this alternative and often more practical ensemble. The quantum case has implications in bounding the accuracy of quantum clocks [39]. Concentration inequalities also have application in constructing confidence intervals for parameter estimation [65].

There are many possible directions of further study. Similar bounds can be derived for discrete dynamics using the same perturbative techniques described here. It would also be useful to bound the spectral quantities in terms of other physical parameters which would provide more accessible bounds and give a better intuition into the physics of the dynamics. Although we obtain inequalities for general fluxes in the fixed-time case, for first passage times we only consider bounds for counting observables. One possible extension would be to find FPT bounds for more general observables such as currents, which are of interest in non-equilibrium physics.

Additionally in the first passage time case, even for the more general counting observables which count subsets of jumps, a better understanding of the transition operators in Equations (7.5) and (8.1) may provide FPT bounds on the whole probability distribution and not just the tails. Finally, recall how we can decompose the continuous-time trajectory in the fixed-observable ensemble into the process which considers jumps and holding times separately. It should therefore be possible to consider similar problems for semi-Markov processes, which would generalise both the classical results and our result for quantum reset processes, as this would include jump processes over discrete configurations with arbitrary holding time distributions.

BIBLIOGRAPHY

[1] L. D. Landau and E. M. Lifshitz. *Statistical Physics*. 3rd. Vol. 5. Oxford: Pergamon Press, 1980.

[2] D. Chandler. *Introduction to Modern Statistical Mechanics*. New York: Oxford University Press, 1987.

[3] C. Jarzynski. Equalities and Inequalities: Irreversibility and the Second Law of Thermodynamics at the Nanoscale. *Annual Review of Condensed Matter Physics* 2.Volume 2, 2011 (Mar. 2011), pp. 329–351. DOI: 10.1146/annurev-conmatphys-062910-140506.

[4] C. Bustamante, J. Liphardt, and F. Ritort. The Nonequilibrium Thermodynamics of Small Systems. *Physics Today* 58.7 (July 2005), pp. 43–48. DOI: 10.1063/1.2012462.

[5] H. Qian. Open-System Nonequilibrium Steady State: Statistical Thermodynamics, Fluctuations, and Chemical Oscillations. *The Journal of Physical Chemistry B* 110.31 (Aug. 2006), pp. 15063–15074. DOI: 10.1021/jp061858z.

[6] C. V. d. Broeck. The many faces of the second law. *Journal of Statistical Mechanics: Theory and Experiment* 2010.10 (Oct. 2010), P10009. DOI: 10.1088/1742-5468/2010/10/P10009.

[7] R. L. Jack. Ergodicity and large deviations in physical systems with stochastic dynamics. *The European Physical Journal B* 93.4 (Apr. 2020), p. 74. DOI: 10.1140/epjb/e2020-100605-3.

[8] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics* 75.12 (Nov. 2012), p. 126001. DOI: 10.1088/0034-4885/75/12/126001.

[9] U. Seifert. From Stochastic Thermodynamics to Thermodynamic Inference. *Annual Review of Condensed Matter Physics* 10.Volume 10, 2019 (Mar. 2019), pp. 171–192. DOI: 10.1146/annurev-conmatphys-031218-013554.

[10] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Macroscopic fluctuation theory. *Reviews of Modern Physics* 87.2 (June 2015), pp. 593–636. DOI: 10.1103/RevModPhys.87.593.

[11] C. Maes. Frenesy: Time-symmetric dynamical activity in nonequilibria. *Physics Reports* 850 (Mar. 2020), pp. 1–33. DOI: 10.1016/j.physrep.2020.01.002.

[12] A. Dmitriev, V. Silchev, and V. Dmitriev. A Simple Econophysics Model of the Stock Market as a Nonequilibrium Open System. *Applied Physics, System Science and Computers II*. Ed. by K. Ntalianis and A. Croitoru. Cham: Springer International Publishing, 2019, pp. 237–243. DOI: 10.1007/978-3-319-75605-9_33.

[13] M. Gligor and M. Ignat. Non-Equilibrium Patterns in the Space of the Stock Market Prices. 27.4 (Nov. 2002), pp. 367–378. DOI: 10.1515/JNETDY.2002.022.

[14] T. Schmiedl and U. Seifert. Stochastic thermodynamics of chemical reaction networks. *The Journal of Chemical Physics* 126.4 (Jan. 2007), p. 044101. DOI: 10.1063/1.2428297.

[15] H. Ge, M. Qian, and H. Qian. Stochastic theory of nonequilibrium steady states. Part II: Applications in chemical biophysics. *Physics Reports* 510.3 (Jan. 2012), pp. 87–118. DOI: 10.1016/j.physrep.2011.09.001.

[16] R. Rao and M. Esposito. Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics. *Physical Review X* 6.4 (Dec. 2016), p. 041064. DOI: 10.1103/PhysRevX.6.041064.

[17] G. R. Bowman and V. S. Pande. Protein folded states are kinetic hubs. *Proceedings of the National Academy of Sciences* 107.24 (June 2010), pp. 10890–10895. DOI: 10.1073/pnas.1003962107.

[18] J. Stigler, F. Ziegler, A. Gieseke, J. C. M. Gebhardt, and M. Rief. The Complex Folding Network of Single Calmodulin Molecules. *Science* 334.6055 (Oct. 2011), pp. 512–516. DOI: 10.1126/science.1207598.

[19] D. Holcman. Modeling DNA and Virus Trafficking in the Cell Cytoplasm. *Journal of Statistical Physics* 127.3 (May 2007), pp. 471–494. DOI: 10.1007/s10955-007-9282-4.

[20] P. C. Bressloff and J. M. Newby. Stochastic models of intracellular transport. *Reviews of Modern Physics* 85.1 (Jan. 2013), pp. 135–196. DOI: 10.1103/RevModPhys.85.135.

[21] E. L. Carroll et al. Incorporating non-equilibrium dynamics into demographic history inferences of a migratory marine species. *Heredity* 122.1 (Jan. 2019), pp. 53–68. DOI: 10.1038/s41437-018-0077-y.

[22] D. Oro and A. Martínez-Abraín. Ecological non-equilibrium and biological conservation. *Biological Conservation* 286 (Oct. 2023), p. 110258. DOI: 10.1016/j.biocon.2023.110258.

[23] T. J. Clark and A. D. Luis. Nonlinear population dynamics are ubiquitous in animals. *Nature Ecology & Evolution* 4.1 (Jan. 2020), pp. 75–81. DOI: 10.1038/s41559-019-1052-6.

[24] S. Whitelam and R. L. Jack. The statistical mechanics of dynamic pathways to self-assembly. *Annual Review of Physical Chemistry* 66 (Apr. 2015), pp. 143–163. DOI: 10.1146/annurev-physchem-040214-121215.

[25] M. Merolle, J. P. Garrahan, and D. Chandler. Space–time thermodynamics of the glass transition. *Proceedings of the National Academy of Sciences* 102.31 (Aug. 2005), pp. 10837–10840. DOI: 10.1073/pnas.0504820102.

[26] J. P. Garrahan. Aspects of non-equilibrium in classical and quantum systems: Slow relaxation and glasses, dynamical large deviations, quantum non-ergodicity, and open quantum dynamics. *Physica A: Statistical Mechanics and its Applications* 504 (Aug. 2018), pp. 130–154. DOI: 10.1016/j.physa.2017.12.149.

[27] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland. Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses. *Physical Review Letters* 98.19 (May 2007), p. 195702. DOI: 10.1103/PhysRevLett.98.195702.

[28] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. v. Duijvendijk, and F. v. Wijland. First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories. *Journal of Physics A: Mathematical and Theoretical* 42.7 (Jan. 2009), p. 075007. DOI: 10.1088/1751-8113/42/7/075007.

[29] D. Ruelle. *Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics*. 2nd ed. Cambridge: Cambridge University Press, 2004. DOI: 10.1017/CBO9780511617546.

[30] J. L. Lebowitz and H. Spohn. A Gallavotti–Cohen-Type Symmetry in the Large Deviation Functional for Stochastic Dynamics. *Journal of Statistical Physics* 95.1 (Apr. 1999), pp. 333–365. DOI: 10.1023/A:1004589714161.

[31] V. Lecomte, C. Appert-Rolland, and F. van Wijland. Thermodynamic Formalism for Systems with Markov Dynamics. *Journal of Statistical Physics* 127.1 (Apr. 2007), pp. 51–106. DOI: 10.1007/s10955-006-9254-0.

[32] A. A. Budini, R. M. Turner, and J. P. Garrahan. Fluctuating observation time ensembles in the thermodynamics of trajectories. *Journal of Statistical Mechanics: Theory and Experiment* 2014.3 (Mar. 2014), P03012. DOI: 10.1088/1742-5468/2014/03/P03012.

[33] P. Pietzonka, A. C. Barato, and U. Seifert. Universal bounds on current fluctuations. *Physical Review E* 93.5 (May 2016), p. 052145. DOI: 10.1103/PhysRevE.93.052145.

[34] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England. Dissipation Bounds All Steady-State Current Fluctuations. *Physical Review Letters* 116.12 (Mar. 2016), p. 120601. DOI: 10.1103/PhysRevLett.116.120601.

[35] M. Kaiser, R. L. Jack, and J. Zimmer. Acceleration of Convergence to Equilibrium in Markov Chains by Breaking Detailed Balance. *Journal of Statistical Physics* 168.2 (July 2017), pp. 259–287. DOI: 10.1007/s10955-017-1805-z.

[36] P. Pietzonka, A. C. Barato, and U. Seifert. Universal bound on the efficiency of molecular motors. *Journal of Statistical Mechanics: Theory and Experiment* 2016.12 (Dec. 2016), p. 124004. DOI: 10.1088/1742-5468/2016/12/124004.

[37] A. C. Barato and U. Seifert. Thermodynamic Uncertainty Relation for Biomolecular Processes. *Physical Review Letters* 114.15 (Apr. 2015), p. 158101. DOI: 10.1103/PhysRevLett.114.158101.

[38] K. Prech, G. T. Landi, F. Meier, N. Nurgalieva, P. P. Potts, R. Silva, and M. T. Mitchison. Optimal time estimation and the clock uncertainty relation for stochastic processes. June 2024. DOI: 10.48550/arXiv.2406.19450.

[39] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber. Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time? *Physical Review X* 7.3 (Aug. 2017), p. 031022. DOI: 10.1103/PhysRevX.7.031022.

[40] K. Macieszczak. Ultimate Kinetic Uncertainty Relation and Optimal Performance of Stochastic Clocks. July 2024. DOI: 10.48550/arXiv.2407.09839.

[41] E. Roldán and J. M. R. Parrondo. Estimating Dissipation from Single Stationary Trajectories. *Physical Review Letters* 105.15 (Oct. 2010), p. 150607. DOI: 10.1103/PhysRevLett.105.150607.

[42] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Vol. 38. Berlin, Heidelberg: Springer, 2010. DOI: 10.1007/978-3-642-03311-7.

[43] F. den Hollander. Large Deviations. June 2008. DOI: 10.1090/fim/014.

[44] H. Touchette. The large deviation approach to statistical mechanics. *Physics Reports* 478.1 (July 2009), pp. 1–69. DOI: 10.1016/j.physrep.2009.05.002.

[45] M. C. Bañuls and J. P. Garrahan. Using Matrix Product States to Study the Dynamical Large Deviations of Kinetically Constrained Models. *Physical Review Letters* 123.20 (Nov. 2019), p. 200601. DOI: 10.1103/PhysRevLett.123.200601.

[46] L. Causer, M. C. Bañuls, and J. P. Garrahan. Optimal Sampling of Dynamical Large Deviations in Two Dimensions via Tensor Networks. *Physical Review Letters* 130.14 (Apr. 2023), p. 147401. DOI: 10.1103/PhysRevLett.130.147401.

[47] M. Gorissen, J. Hooyberghs, and C. Vanderzande. Density-matrix renormalization-group study of current and activity fluctuations near nonequilibrium phase transitions. *Physical Review E* 79.2 (Feb. 2009), p. 020101. DOI: 10.1103/PhysRevE.79.020101.

[48] J. P. Garrahan. Simple bounds on fluctuations and uncertainty relations for first-passage times of counting observables. *Physical Review E* 95.3 (Mar. 2017), p. 032134. DOI: 10.1103/PhysRevE.95.032134.

[49] J. Kiukas, M. Guţă, I. Lesanovsky, and J. P. Garrahan. Equivalence of matrix product ensembles of trajectories in open quantum systems. *Physical Review E* 92.1 (July 2015), p. 012132. DOI: 10.1103/PhysRevE.92.012132.

[50] T. R. Gingrich and J. M. Horowitz. Fundamental Bounds on First Passage Time Fluctuations for Currents. *Physical Review Letters* 119.17 (Oct. 2017), p. 170601. DOI: 10.1103/PhysRevLett.119.170601.

[51] S. Redner. *A Guide to First-Passage Processes*. Cambridge: Cambridge University Press, 2001. DOI: 10.1017/CBO9780511606014.

[52] S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez, and J. Klafter. First-passage times in complex scale-invariant media. *Nature* 450.7166 (Nov. 2007), pp. 77–80. DOI: 10.1038/nature06201.

[53] A. L. Lloyd and R. M. May. How Viruses Spread Among Computers and People. *Science* 292.5520 (May 2001), pp. 1316–1317. DOI: 10.1126/science.1061076.

[54] T. D. Swinburne, D. Kannan, D. J. Sharpe, and D. J. Wales. Rare events and first passage time statistics from the energy landscape. *The Journal of Chemical Physics* 153.13 (Oct. 2020), p. 134115. DOI: 10.1063/5.0016244.

[55] A. Szabo, K. Schulten, and Z. Schulten. First passage time approach to diffusion controlled reactions. *The Journal of Chemical Physics* 72.8 (Apr. 1980), pp. 4350–4357. DOI: 10.1063/1.439715.

[56] D. S. Grebenkov, R. Metzler, and G. Oshanin. Towards a full quantitative description of single-molecule reaction kinetics in biological cells. *Physical Chemistry Chemical Physics* 20.24 (June 2018), pp. 16393–16401. DOI: 10.1039/C8CP02043D.

[57] V. P. Belavkin. A stochastic posterior Schrödinger equation for counting nondemolition measurement. *Letters in Mathematical Physics* 20.2 (Aug. 1990), pp. 85–89. DOI: 10.1007/BF00398273.

[58] L. Bouten, M. Guta, and H. Maassen. Stochastic Schrödinger equations. *Journal of Physics A: Mathematical and General* 37.9 (Feb. 2004), p. 3189. DOI: 10.1088/0305-4470/37/9/010.

[59] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control*. Cambridge: Cambridge University Press, 2009. DOI: 10.1017/CBO9780511813948.

[60] J. Gough and M. James. The Series Product and Its Application to Quantum Feedforward and Feedback Networks. *Automatic Control, IEEE Transactions on* 54 (Dec. 2009), pp. 2530–2544. DOI: 10.1109/TAC.2009.2031205.

[61] F. Carollo, R. L. Jack, and J. P. Garrahan. Unraveling the Large Deviation Statistics of Markovian Open Quantum Systems. *Physical Review Letters* 122.13 (Apr. 2019), p. 130605. DOI: 10.1103/PhysRevLett.122.130605.

[62] J. P. Garrahan and I. Lesanovsky. Thermodynamics of Quantum Jump Trajectories. *Physical Review Letters* 104.16 (Apr. 2010), p. 160601. DOI: 10.1103/PhysRevLett.104.160601.

[63] G. Guarnieri, G. T. Landi, S. R. Clark, and J. Goold. Thermodynamics of precision in quantum nonequilibrium steady states. *Physical Review Research* 1.3 (Oct. 2019), p. 033021. DOI: 10.1103/PhysRevResearch.1.033021.

[64] Y. Hasegawa. Quantum Thermodynamic Uncertainty Relation for Continuous Measurement. *Physical Review Letters* 125.5 (July 2020), p. 050601. DOI: 10.1103/PhysRevLett.125.050601.

[65] F. Girotti, J. P. Garrahan, and M. Guţă. Concentration Inequalities for Output Statistics of Quantum Markov Processes. *Annales Henri Poincaré* 24.8 (Aug. 2023), pp. 2799–2832. DOI: 10.1007/s00023-023-01286-1.

[66] T. Benoist, L. Hänggli, and C. Rouzé. Deviation bounds and concentration inequalities for quantum noises. *Quantum* 6 (Aug. 2022), p. 772. DOI: 10.22331/q-2022-08-04-772.

[67] P. Lezaud. Chernoff-type bound for finite Markov chains. *The Annals of Applied Probability* 8.3 (Aug. 1998), pp. 849–867. DOI: 10.1214/aoap/1028903453.

[68] J. Fan, B. Jiang, and Q. Sun. Hoeffding's Inequality for General Markov Chains and Its Applications to Statistical Learning. *Journal of Machine Learning Research* 22.139 (2021), pp. 1–35.

[69] B. Jiang, Q. Sun, and J. Fan. Bernstein's inequalities for general Markov chains. Jan. 2024. DOI: 10.48550/arXiv.1805.10721.

[70] G. Bakewell-Smith, F. Girotti, M. Guţă, and J. P. Garrahan. General Upper Bounds on Fluctuations of Trajectory Observables. *Physical Review Letters* 131.19 (Nov. 2023), p. 197101. DOI: 10.1103/PhysRevLett.131.197101.

[71] G. Bakewell-Smith, F. Girotti, M. Guţă, and J. P. Garrahan. Bounds on Fluctuations of First Passage Times for Counting Observables in Classical and Quantum Markov Processes (May 2024). DOI: 10.48550/arXiv.2405.09669.

[72] W.-K. Ching, X. Huang, M. K. Ng, and T.-K. Siu. *Markov Chains: Models, Algorithms and Applications*. Vol. 189. Boston, MA: Springer US, 2013. DOI: 10.1007/978-1-4614-6312-2.

[73] R. Serfozo. *Basics of applied stochastic processes*. Berlin: Springer, 2009.

[74] J. M. Horowitz and T. R. Gingrich. Thermodynamic uncertainty relations constrain non-equilibrium fluctuations. *Nature Physics* 16.1 (Jan. 2020), pp. 15–20. DOI: 10.1038/s41567-019-0702-6.

[75] M. Esposito, K. Lindenberg, and C. V. d. Broeck. Thermoelectric efficiency at maximum power in a quantum dot. *Europhysics Letters* 85.6 (Apr. 2009), p. 60010. DOI: 10.1209/0295-5075/85/60010.

[76] H. Qian and D. A. Beard. Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium. *Biophysical Chemistry* 114.2-3 (Apr. 2005), pp. 213–220. DOI: 10.1016/j.bpc.2004.12.001.

[77] S. Condamin, O. Bénichou, and M. Moreau. Random walks and Brownian motion: A method of computation for first-passage times and related quantities in confined geometries. *Physical Review E* 75.2 (Feb. 2007), p. 021111. DOI: 10.1103/PhysRevE.75.021111.

[78] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences* 101.42 (Oct. 2004), pp. 15124–15129. DOI: 10.1073/pnas.0308344101.

[79] S. Sabhapandit and S. N. Majumdar. Freezing Transition in the Barrier Crossing Rate of a Diffusing Particle. *Physical Review Letters* 125.20 (Nov. 2020), p. 200601. DOI: 10.1103/PhysRevLett.125.200601.

[80] R. S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics*. Berlin, Heidelberg: Springer, 2006. DOI: 10.1007/3-540-29060-5.

[81] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. Ed. by A. Martin-Löf. Berlin, Heidelberg: Springer Berlin Heidelberg, 1994, pp. 895–913. DOI: 10.1007/978-3-642-40607-2_8.

[82] J. Gartner. On Large Deviations from the Invariant Measure. *Theory of Probability and its Applications* 22.1 (1977), p. 16. DOI: 10.1137/1122003.

[83] R. S. Ellis. Large Deviations for a General Class of Random Vectors. *The Annals of Probability* 12.1 (Feb. 1984), pp. 1–12. DOI: 10.1214/aop/1176993370.

[84] C. Maes and K. Netočný. Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *Europhysics Letters* 82.3 (Apr. 2008), p. 30003. DOI: 10.1209/0295-5075/82/30003.

[85] J. Hoppenau, D. Nickelsen, and A. Engel. Level 2 and level 2.5 large deviation functionals for systems with and without detailed balance. *New Journal of Physics* 18.8 (July 2016), p. 083010. DOI: 10.1088/1367-2630/18/8/083010.

[86] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations for large time, I. *Communications on Pure and Applied Mathematics* 28.1 (Jan. 1975), pp. 1–47. DOI: 10.1002/cpa.3160280102.

[87] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations for large time, II. *Communications on Pure and Applied Mathematics* 28.2 (Mar. 1975), pp. 279–301. DOI: 10.1002/cpa.3160280206.

[88] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time—III. *Communications on Pure and Applied Mathematics* 29.4 (July 1976), pp. 389–461. DOI: 10.1002/cpa.3160290405.

[89] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain markov process expectations for large time. IV. *Communications on Pure and Applied Mathematics* 36.2 (Mar. 1983), pp. 183–212. DOI: 10.1002/cpa.3160360204.

[90] A. C. Barato and R. Chetrite. A Formal View on Level 2.5 Large Deviations and Fluctuation Relations. *Journal of Statistical Physics* 160.5 (Sept. 2015), pp. 1154–1172. DOI: 10.1007/s10955-015-1283-0.

[91] L. Bertini, A. Faggionato, and D. Gabrielli. Large deviations of the empirical flow for continuous time Markov chains. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 51.3 (Aug. 2015), pp. 867–900. DOI: 10.1214/14-AIHP601.

[92] S. R. White. Density matrix formulation for quantum renormalization groups. *Physical Review Letters* 69.19 (Nov. 1992), pp. 2863–2866. DOI: 10.1103/PhysRevLett.69.2863.

[93] F. Verstraete, J. I. Cirac, J. I. Latorre, E. Rico, and M. M. Wolf. Renormalization-Group Transformations on Quantum States. *Physical Review Letters* 94.14 (Apr. 2005), p. 140601. DOI: 10.1103/PhysRevLett.94.140601.

[94] M. Polettini, A. Lazarescu, and M. Esposito. Tightening the uncertainty principle for stochastic currents. *Physical Review E* 94.5 (Nov. 2016), p. 052104. DOI: 10.1103/PhysRevE.94.052104.

[95] T. Koyuk and U. Seifert. Operationally Accessible Bounds on Fluctuations and Entropy Production in Periodically Driven Systems. *Physical Review Letters* 122.23 (June 2019), p. 230601. DOI: 10.1103/PhysRevLett.122.230601.

[96] A. Dechant and S.-I. Sasa. Fluctuation-response inequality out of equilibrium. *Proceedings of the National Academy of Sciences of the United States of America* 117.12 (Mar. 2020), pp. 6430–6436. DOI: 10.1073/pnas.1918386117.

[97] L. P. Fischer, H.-M. Chun, and U. Seifert. Free diffusion bounds the precision of currents in underdamped dynamics. *Physical Review E* 102.1 (July 2020), p. 012120. DOI: 10.1103/PhysRevE.102.012120.

[98] H.-M. Chun, L. P. Fischer, and U. Seifert. Effect of a magnetic field on the thermodynamic uncertainty relation. *Physical Review E* 99.4 (Apr. 2019), p. 042128. DOI: 10.1103/PhysRevE.99.042128.

[99] A. Dechant and S.-i. Sasa. Entropic bounds on currents in Langevin systems. *Physical Review E* 97.6 (June 2018), p. 062101. DOI: 10.1103/PhysRevE.97.062101.

[100] P. Pietzonka, F. Ritort, and U. Seifert. Finite-time generalization of the thermodynamic uncertainty relation. *Physical Review E* 96.1 (July 2017), p. 012101. DOI: 10.1103/PhysRevE.96.012101.

[101] J. M. Horowitz and T. R. Gingrich. Proof of the finite-time thermodynamic uncertainty relation for steady-state currents. *Physical Review E* 96.2 (Aug. 2017), p. 020103. DOI: 10.1103/PhysRevE.96.020103.

[102] K. Proesmans and C. Broeck. Discrete-time thermodynamic-like uncertainty relation. *EPL (Europhysics Letters)* 119 (Sept. 2017). DOI: 10.1209/0295-5075/119/20001.

[103] K. Brandner, T. Hanazato, and K. Saito. Thermodynamic Bounds on Precision in Ballistic Multiterminal Transport. *Physical Review Letters* 120.9 (Mar. 2018), p. 090601. DOI: 10.1103/PhysRevLett.120.090601.

[104] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Feb. 2013. DOI: 10.1093/acprof:oso/9780199535255.001.0001.

[105] T. Kato. *Perturbation Theory for Linear Operators*. Vol. 132. Berlin, Heidelberg: Springer, 1995. DOI: 10.1007/978-3-642-66282-9.

[106] C. J. Geyer. Practical Markov Chain Monte Carlo. *Statistical Science* 7.4 (Nov. 1992), pp. 473–483. DOI: 10.1214/ss/1177011137.

[107] S. Goel and M. J. Salganik. Respondent-driven sampling as Markov chain Monte Carlo. *Statistics in Medicine* 28.17 (July 2009), pp. 2202–2229. DOI: 10.1002/sim.3613.

[108] A. Wilinski. Time series modeling and forecasting based on a Markov chain with changing transition matrices. *Expert Systems with Applications* 133 (Nov. 2019), pp. 163–172. DOI: 10.1016/j.eswa.2019.04.067.

[109] P. W. Glynn and D. Ormoneit. Hoeffding's inequality for uniformly ergodic Markov chains. *Statistics & Probability Letters* 56.2 (Jan. 2002), pp. 143–146. DOI: 10.1016/S0167-7152(01)00158-4.

[110] C. Gardiner and P. Zoller. *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. Springer Science & Business Media, Aug. 2004.

[111] L. A. Clark, B. Markowicz, and J. Kołodyński. Exploiting non-linear effects in optomechanical sensors with continuous photon-counting. *Quantum* 6 (Sept. 20, 2022), p. 812. DOI: 10.22331/q-2022-09-20-812.

[112] L. A. Clark and J. Kolodynski. Efficient inference of quantum system parameters by Approximate Bayesian Computation. Feb. 23, 2025. DOI: 10.48550/arXiv.2407.00724.

[113] M. Tse et al. Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy. *Physical Review Letters* 123.23 (Dec. 5, 2019), p. 231107. DOI: 10.1103/PhysRevLett.123.231107.

[114] S. Colombo, E. Pedrozo-Peñafiel, and V. Vuletić. Entanglement-enhanced optical atomic clocks. *Applied Physics Letters* 121.21 (Nov. 21, 2022), p. 210502. DOI: 10.1063/5.0121372.

[115] M. A. Nielsen and I. L. Chuang. Quantum Computation and Quantum Information: 10th Anniversary Edition. Dec. 2010. DOI: 10.1017/CBO9780511976667.

[116] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics* 17.5 (May 1976), pp. 821–825. DOI: 10.1063/1.522979.

[117] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics* 48.2 (June 1976), pp. 119–130. DOI: 10.1007/BF01608499.

[118] D. E. Evans and R. Høegh-Krohn. Spectral Properties of Positive Maps on C * -Algebras. *Journal of the London Mathematical Society* s2-17.2 (Apr. 1978), pp. 345–355. DOI: 10.1112/jlms/s2-17.2.345.

[119] K.-J. Engel and R. Nagel. One-parameter semigroups for linear evolution equations. *Semigroup Forum* 63.2 (June 2001), pp. 278–280. DOI: 10.1007/s002330010042.

[120] R. T. Rockafellar. *Convex Analysis: (PMS-28)*. Princeton University Press, Apr. 2015. DOI: 10.1515/9781400873173.

[121] A. Kolchinsky, N. Ohga, and S. Ito. Thermodynamic bound on spectral perturbations, with applications to oscillations and relaxation dynamics. *Physical Review Research* 6.1 (Jan. 2024), p. 013082. DOI: 10.1103/PhysRevResearch.6.013082.

[122] I. D. Terlizzi and M. Baiesi. Kinetic uncertainty relation. *Journal of Physics A: Mathematical and Theoretical* 52.2 (Dec. 2018), 02LT03. DOI: 10.1088/1751-8121/aaee34.

[123] S. Puntanen and G. P. H. Styan. Historical Introduction: Issai Schur and the Early Development of the Schur Complement. *The Schur Complement and Its Applications*. Ed. by F. Zhang. Boston, MA: Springer US, 2005, pp. 1–16. DOI: 10.1007/0-387-24273-2_1.

[124] F. Noé and C. Clementi. Collective variables for the study of long-time kinetics from molecular trajectories: theory and methods. *Current Opinion in Structural Biology* 43 (Apr. 2017), pp. 141–147. DOI: 10.1016/j.sbi.2017.02.006.

[125] L. Bonati, G. Piccini, and M. Parrinello. Deep learning the slow modes for rare events sampling. *Proceedings of the National Academy of Sciences* 118.44 (Nov. 2021), e2113533118. DOI: 10.1073/pnas.2113533118.

[126] D. Freedman. *Approximating Countable Markov Chains*. New York, NY: Springer, 1983. DOI: 10.1007/978-1-4613-8230-0.

[127] R. Carbone and Y. Pautrat. Homogeneous Open Quantum Random Walks on a Lattice. *Journal of Statistical Physics* 160.5 (Sept. 2015), pp. 1125–1153. DOI: 10.1007/s10955-015-1261-6.

[128] N. G. Bean, M. Fackrell, and P. Taylor. Characterization of Matrix-Exponential Distributions. *Stochastic Models* 24.3 (Aug. 2008), pp. 339–363. DOI: 10.1080/15326340802232186.

[129] M. J. Kewming, A. Kiely, S. Campbell, and G. T. Landi. First passage times for continuous quantum measurement currents. *Physical Review A* 109.5 (May 2024), p. L050202. DOI: 10.1103/PhysRevA.109.L050202.

[130] M. Wolf. Quantum Channels & Operations Guided Tour. *Online Lecture Notes* (2012).

[131] T. Nishiyama and Y. Hasegawa. Exact solution to quantum dynamical activity. *Physical Review E* 109.4 (Apr. 2024), p. 044114. DOI: 10.1103/PhysRevE.109.044114.

[132] T. Mori and T. Shirai. Symmetrized Liouvillian Gap in Markovian Open Quantum Systems. *Physical Review Letters* 130.23 (June 2023), p. 230404. DOI: 10.1103/PhysRevLett.130.230404.

[133] A. Dechant. Thermodynamic constraints on the power spectral density in and out of equilibrium. June 2023. DOI: 10.48550/arXiv.2306.00417.