# On Absolute Continuity for Stochastic Partial Differential Equations and an Averaging Principle for a Queueing Network

by

Andrea Cherese Watkins

Department of Mathematics
Duke University

Date: ________________________

Approved:

Jonathan Mattingly, Co-Supervisor

Otis Jennings, Co-Supervisor

James Beale

James Nolen

Michael Reed

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2010

# Abstract
(Mathematics)

# On Absolute Continuity for Stochastic Partial Differential Equations and an Averaging Principle for a Queueing Network

by

Andrea Cherese Watkins

Department of Mathematics
Duke University

Date: ____________________

Approved:

____________________
Jonathan Mattingly, Co-Supervisor

____________________
Otis Jennings, Co-Supervisor

____________________
James Beale

____________________
James Nolen

____________________
Michael Reed

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2010

# Abstract

The connection between elliptic stochastic diffusion processes and partial differential equations is rich and well understood. This connection is not very well understood when the stochastic differential equation takes values in an infinite dimensional space such as a function space. In this case, the diffusion is a stochastic partial differential equation (SPDE) and the notion of ellipticity is ambiguous. We establish a sufficient condition on the diffusion coefficient of a class of nonlinear SPDEs, which is analogous to the nondegeneracy condition in finite dimensions, that allows for the existence of a Markov transition density that is absolutely continuous with respect to an infinite dimensional Gaussian measure.

In the second part of this work, we consider a two-station queueing network that processes $K$ job types. The first station in this network is a polling station, and we assume that the second station is operating under any nonidling service discipline. We consider diffusion-scaled versions of many of the processes governing this system, and we show that the scaled two-dimensional *total* workload process converges to Brownian motion in a wedge. We also show that the scaled immediate workload process for station 2 does not converge, but admits an averaging principle.

*To Bernice Hill (1932-2010)*

# Contents

# List of Abbreviations and Symbols

## Abbreviations

| | |
|---|---|
| SPDE | Stochastic Partial Differential Equation |
| FES | Fractional Exhaustive Service, a family polling policies that can govern a polling station |
| HTLT | Heavy Traffic Limit Theorem |

## Symbols

| | |
|---|---|
| $\mathbb{C}$ | The complex plane |
| $\mathbb{T}$ | The two-dimensional torus. $\mathbb{T}^2 \stackrel{\text{def}}{=} [0, 2\pi]^2$ |
| $\mathbb{Z}^n$ | The $n$-dimensional integer lattice |
| $H$ | Separable Hilbert space |
| $\lVert \cdot \rVert$ | Norm on the Hilbert space $H$ |
| $\mathcal{D}(A)$ | The domain of an operator $A$ |
| $\lVert \cdot \rVert_\alpha$ | Norm on the space $\mathcal{H}^\alpha = \mathcal{D}(A^{\frac{\alpha}{2}})$ for some operator $A$ |
| $\lVert \cdot \rVert_{\text{HS}}$ | The Hilbert-Schmidt norm |
| $B_b(H)$ | The Banach space of all Borel bounded mappings |
| $C(\Omega; W)$ | The space of continuous mappings $f : \Omega \to W$ |
| $L^p(\Omega; W)$ | The space of strongly measurable functions $f : \Omega \to W$ such that $\int_\Omega \lVert f \rVert_W^p dx < \infty, 1 \leq p < \infty$ |
| $\mathcal{B}(X)$ | Borel $\sigma$-algebra generated by all open subsets of space $X$ |
| $\mathcal{L}(X)$ | Banach space of bounded linear operators of space $X$ |

| | |
|---|---|
| $\mathcal{L}(X,Y)$ | Space of bounded linear operators from space $X$ to space $Y$ |
| $\mathcal{L}^+(X)$ | The collection of $T \in \mathcal{L}(X)$ that are symmetric and positive |
| $\mathcal{L}_1(X)$ | The collection of $T \in \mathcal{L}(X)$ that are trace class operators |
| $\mathcal{L}_2(X)$ | The collection of $T \in \mathcal{L}(X)$ that are Hilbert-Schmidt operators |
| $\Delta_\delta(a)$ | Sector in the complex plane. $\Delta_\delta(a) = \{z \in \mathbb{C} : \lvert\arg(z-a)\rvert < \delta, z \neq a\}$ |
| $\Sigma_\sigma(a)$ | Sector in the complex plane. $\Sigma_\sigma(a) = \{z \in \mathbb{C} : \lvert\arg(z-a)\rvert > \sigma, z \neq a\}$ |
| $\sigma(A)$ | The spectrum of the operator $A$ |
| $\rho(A)$ | The resolvent set of the operator $A$ |
| $\ell^2$ | Hilbert space of square-summable sequences |
| $\mu$-a.e. | Except for a set $E$ such that $\mu(E) = 0$ |
| $\mu \ll \nu$ | $\mu$ is absolutely continuous with respect to $\nu$ |
| $\mu \perp \nu$ | $\mu$ and $\nu$ are mutually singular |
| $\mu \sim \nu$ | $\mu$ and $\nu$ are equivalent in the sense of mutually absolutely continuous |
| $s \wedge t$ | The minimum of $t$ and $s$ |
| $s \vee t$ | The maximum of $t$ and $s$ |
| $\mathbf{1}_A$ | The indicator function that equals 1 if $A$ is true and is 0 otherwise |
| $e = e(t)$ | The identity function $e(t) = t$ for all $t$ |
| $X^n \Rightarrow X$ | $X^n$ converges to $X$ in distribution as $n \to \infty$ |
| $X^n \xrightarrow{P} X$ | $X^n$ converges to $X$ in probability as $n \to \infty$ |
| $a\%b$ | $a$ mod $b$ |
| $O_p(g(n))$ | "Order $g(n)$ in probability". See defintion 7.5.3 |
| $o_p(g(n))$ | A stochastic order relationship. See definition 7.5.4 |
| $D(s,t)$ | The Skorokhod space of real valued càdlàg functions on the domain $(s,t)$ |

| | |
|---|---|
| $(f \circ g)(x)$ | The composition of the functions $f$ and $g$, meaning $f(g(x))$ |

# Acknowledgements

I am most grateful to God, whose presence I have felt at every step along this academic journey, and without whom I would be lost. I have a wonderfully supportive family, and I am extremely grateful to my parents Bill and Patricia Watkins who have always given me just the right balance of space and support to allow me to reach my fullest potential. I thank them for their sacrifices and I view this accomplishment as a return on their investment. I also thank my circle of friends who were always calling to check up on me. Though most of them were long distance, they always seemed to be right in the wings cheering me on. And, thank you to Brandon for being there to share in the highs and lows of the last couple of years of this journey. I truly appreciate you.

I would also like to thank my advisors, Jonathan Mattingly and Otis Jennings. Professor Mattingly's generous support has allowed me to see mathematics for the global and dynamic field that it is. From him, I learned to not only have broad interests, but to also be able to take a broad approach to particular problems. I believe these lessons will serve me well in the future. I must thank Professor Jennings for his extreme patience and guidance during my final years of graduate school. He has been exceptionally generous with his time, and without the clarity of his instruction, I would have never been able to collaborate with him in making this contribution to queueing theory. I am also very grateful for the tremendous support and mentoring I received from Dr. Arlie Petters during my time at Duke.

The majority of my graduate work was conducted with the financial support of a graduate fellowship from the National Physical Sciences Consortium. I thank this organization and my fellowship sponsor, the National Security Agency. Many of the employees of these organizations have been very supportive during my graduate school career. I especially thank Dr. Michelle Wagner and Dr. Joretta Joseph.

I would also like to acknowledge the many friends, colleagues, and professors that have aided in making my time at Duke a pleasant experience. Thank you to the entire staff of the Duke University Math Department. Thanks to Alberto, Rachel, Oliver, Amir, Jeff, and a special thank you to Prakash for being willing to talk math at any time. Thank you to the Duke University Bouchet Society for their mission and their support.

I wish to acknowledge Martin J. Carnaghi, my high school math teacher who passed away during my last year of graduate school. I credit him with laying a strong foundation for me mathematically, and I thank him for his continuous support even as I matriculated through graduate school.

Finally, I would like to thank my grandmother, Bernice Hill, who passed away just a few months before this dissertation was complete and to whom this work is dedicated. I thank her for telling me that I was fearless so often that I started to believe it.

# 1

# Introduction

Parabolic partial differential equations provide a rich class of examples that describe a wide family of physical and mathematical systems. It is often the case that adding random forcing terms to these equations provide even better models for these systems. Thus, contributing to the understanding of equations like stochastic reaction diffusion equations, the stochastic Burgers equation, stochastic Navier-Stokes equations, and the stochastic Cahn-Hilliard equation seems like a worthy endeavor. We study the dynamics of these stochastic partial differential equations as evolutionary equations in some Banach space using their mild solutions and semiflow properties.

One dynamical property of interest is the absolute continuity of the solutions to these equations. This interest is motivated by the fascinating interplay of probability theory and partial differential equations. The most famous example of this interplay being the theorem of Kolmogorov stating that the solution to the Cauchy problem, a *deterministic* parabolic partial differential equation, is given by the expected value of the solution to a *stochastic* differential equation. As a result, when the elliptic operator in the Cauchy problem has a fundamental solution, the law of the associated stochastic process can be written as a transition density that is absolutely continuous with respect to Lebesgue measure. This theory is well established in finite dimensions where Lebesgue measure is the natural reference measure. However, in infinite dimensional spaces, Lebesgue measure does not make sense.

In infinite dimensions, the natural reference measures are Gaussian measures. In our case, they will arise from the law of the "linearized" version of our nonlinear stochastic partial differential equations. As in the finite dimensional case, we ask the question when can the law of the solutions to the stochastic partial differential equations that we consider be written as transition densities with respect to a Gaussian measure? In some ways, the question is the same as in the infinite dimensional case, we would like to establish conditions on the infinitesimal generator of the stochastic partial differential equation that would equate to a condition like ellipticity.

This effort has been undertaken in a more complex setting with a view toward ergodicity using Malliavin calculus. Hairer and Mattingly (2008). The question of absolute continuity of the invariant measures for the stochastic Burgers equation and a reaction-diffusion equation was addressed by Da Prato and Debussche (2004) using a very different technique than what is presented here. We actually establish equivalence of the transition densities and the Gaussian measures, not just absolute continuity, which allows us to quickly prove unique ergodicity for these equations. This method gives results on ergodicity that agrees with results of Goldys and Maslowski (2006b) and Goldys and Maslowski (2006a). Some other works related to this thesis are Da Prato and Debussche (1996), Da Prato et al. (1994), and in particular, Mattingly and Suidan (2005).

In this thesis, we establish, among other assumptions, a majorization and minorization condition on the diffusion coefficient of the noise in our stochastic partial differential equations that is reminiscent of the nondegenerate, positivity condition that is necessary for elliptic operators. We show that under these assumptions, we have equivalence of the transition densities of the nonlinear stochastic equation, the linearized equation, and the invariant measure of the linearized equation for all times and initial conditions. An obvious first attempt to prove absolute continuity of these measures would be to apply Girsanov's theorem. However, the nonlinear equation and the linearized equation do not necessarily have equivalent distributions on pathspace, although their particular transition densities may be equivalent at a point in time. Thus, Girsanov's theorem is not applicable. We use a technique that was first introduced in Mattingly and Suidan (2005). It involves introducing an auxiliary process that is equivalent to the linearized equation on pathspace and is equal to the nonlinear equation at a point in time. We then apply Girsanov's theorem to the auxiliary equation and the linear equation to get our result. This technique was used in Mattingly and Suidan (2005), where it was applied to the small scales of the Navier-Stokes equation. We have generalized the technique to a broader class of equations and established sufficient conditions, in particular, regularizing properties, on the nonlinear term and the noise, so that the interaction between the two causes the nonlinear equation to behave like a mere perturbation of the linear equation.

# 2

# The Setting

## 2.1 The Nonlinear Equation

We are primarily concerned with nonlinear stochastic partial differential equations of the form

$$dX_t = -LX_t dt + F(X_t, t)dt + QdW_t \tag{2.1}$$

that take values in a separable infinite dimensional Hilbert space $H$ and generate strong Markov processes $X_t$ on the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$. We assume that $L$ is a strictly positive, self-adjoint, linear operator with compact resolvent on $H$, and the operator $F$ is a map from $H \times [0, T] \to H$ that we will describe later in more detail. We assume that $W$ is a cylindrical Wiener process on a second Hilbert space $\mathcal{U}$, and $Q$ is a bounded linear operator from $\mathcal{U} \to H$. We will also describe $Q$ in more detail in later sections. Some examples of equations that fit the form of (2.1) are the vorticity formulation of the stochastic Navier-Stokes equation

$$dq_t = \nu \Delta q_t dt + B(\mathcal{K}q_t, q_t)dt + QdW_t \tag{2.2}$$

and the stochastic Cahn-Hilliard equation

$$du_t = -\Delta^2 u_t dt + \Delta f(u_t)dt + QdW_t. \tag{2.3}$$

Before we discuss properties of the solution to (2.1), we need to introduce a few more concepts. These concepts will make our discussion more informative and concise.

## 2.2 Analytic Semigroups

As stated earlier, $L$ is a strictly positive, self-adjoint, linear operator with compact resolvent on $H$. In this section, we will show that these properties imply that $-L$ generates an analytic semigroup $S(t) = \exp(-Lt)$, which is a compact operator for all $t > 0$. The properties of $S(t)$ will prove to be crucial in proving some of our main results. We begin with some fundamental definitions.

### *2.2.1 The Generated $C_0$-semigroup*

**Definition 2.2.1.** *Let $W$ denote a given Banach space. A semigroup $T(t)$ of bounded linear operators on $W$ is a $C_0$-semigroup if for all $t \geq 0$ the following hold:*

$$T(0) = I \tag{2.4}$$

$$T(t)T(s) = T(t+s), \quad s, t \in [0, \infty) \tag{2.5}$$

$$\lim_{t \to 0^+} T(t)w = w \quad \textit{for all } w \in W \tag{2.6}$$

If we call $A$ the infinitesimal generator of the $C_0$-semigroup $T(t)$, then we define its domain as

$$\mathcal{D}(A) \stackrel{\text{def}}{=} \left\{ w \in W : \lim_{h \to 0^+} \frac{T(h) - I}{h} w \text{ exists in } W \right\}, \tag{2.7}$$

and for $w \in \mathcal{D}(A)$ we let

$$Aw \stackrel{\text{def}}{=} \lim_{h \to 0^+} \frac{T(h) - I}{h} w = \frac{d^+(T(t)w)}{dt} \Big|_{t=0}. \tag{2.8}$$

Now, we follow the techniques given in Sell and You (2002) and show that $-L$ as described in (2.1) generates a $C_0$-semigroup, $e^{-Lt}$. The properties of $L$ imply that it has a real point spectrum bounded below by some $a > 0$. We can order the eigenvalues of $L$ so that

$$0 < a = \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots. \tag{2.9}$$

Since we assume that $H$ is infinite dimensional, we also have $\lambda_k \to \infty$ as $k \to \infty$. The properties of $L$ further imply that the associated eigenvectors $\{e_k, k \in \mathbb{N}\}$ form an orthonormal basis for $H$. We define the domain of $L$ as

$$\mathcal{D}(L) = \{u \in H : \sum_{i=1}^{\infty} |\lambda_i|^2 | \langle u, e_i \rangle |^2 < \infty\}, \tag{2.10}$$

where $L$ is given by

$$Lu = \sum_{i=1}^{\infty} \lambda_i \langle u, e_i \rangle e_i. \tag{2.11}$$

For any continuous, real-valued function $f$ defined on the spectrum of $L$, the Spectral Mapping Theorem states that $f(L)$ is defined by

$$f(L)u = \sum_{i=1}^{\infty} f(\lambda_i) \langle u, e_i \rangle e_i, \tag{2.12}$$

where the domain of $f(L)$ is given by

$$\mathcal{D}(f(L)) = \{u \in H : \sum_{i=1}^{\infty} |f(\lambda_i)|^2 | \langle u, e_i \rangle |^2 < \infty\}. \tag{2.13}$$

Now, we show that the $C_0$-semigroup generated by $-L$ is given by

$$e^{-Lt}u \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} e^{-\lambda_i t} \langle u, e_i \rangle e_i. \tag{2.14}$$

First, we see that

$$\|e^{-Lt}u\|^2 = \sum_{i=1}^{\infty} e^{-2\lambda_i t} | \langle u, e_i \rangle |^2 \leq \sum_{i=1}^{\infty} e^{-2\lambda_1 t} | \langle u, e_i \rangle |^2 \leq \sum_{i=1}^{\infty} e^{-2at} \|u\|^2, \tag{2.15}$$

where the last inequality is by the boundedness of $\sigma(L)$ and Parseval's equality. Thus, we see that $e^{-Lt} \in \mathcal{L}(H)$ and $\|e^{-Lt}\| \leq e^{-at}$ for all $t \geq 0$. Notice that (2.4) is trivially true for $e^{-Lt}$, and (2.5) follows immediately from the Spectral Mapping Theorem and the orthonormality of the $\{e_k\}$. Now, we verify (2.6). Let $u \in H$ and let $\epsilon > 0$ be given. We can choose some $N \geq 1$ so that $\sum_{i=N+1}^{\infty} | \langle u, e_k \rangle |^2 \leq \epsilon/2$. We can also choose $h_0 > 0$ so that for $0 < h \leq h_0$,

$$(e^{-\lambda_i h} - 1)^2 | \langle u, e_i \rangle |^2 \leq \frac{\epsilon}{2N}, \qquad \text{for all } 1 \leq i \leq N \text{ and } 0 < h \leq h_0. \tag{2.16}$$

Then, for the given $\epsilon$, we have

$$\|e^{-Lh}u - u\|^2 = \sum_{i=1}^{\infty} (e^{-\lambda_i h} - 1)^2 | \langle u, e_i \rangle |^2 \tag{2.17}$$

$$\leq \sum_{i=1}^{N} (e^{-\lambda_i h} - 1)^2 | \langle u, e_i \rangle |^2 + \sum_{i=1}^{\infty} | \langle u, e_i \rangle |^2 \leq \epsilon \tag{2.18}$$

for $0 < h \leq h_0$. Thus, (2.6) is satisfied, and $e^{-Lt}$ is a $C_0$-semigroup on $H$. Finally, we show that $-L$ is indeed the infinitesimal generator of the $C_0$-semigroup $e^{-Lt}$. Consider the continuous orthogonal projection $P_N$ defined as

$$P_N u \stackrel{\text{def}}{=} \sum_{i=1}^{N} \langle u e_i \rangle \, e_i, \tag{2.19}$$

and suppose that some operator $B$ is the infinitesimal generator of $e^{-Lt}$. For $u \in \mathcal{D}(B)$,

$$P_N \left( \lim_{h \to 0^+} \frac{e^{-Lh} - I}{h} u \right) = \lim_{h \to 0^+} P_N \left( \frac{e^{-Lh} - I}{h} u \right) = \sum_{i=1}^{N} -\lambda_i \langle u, e_i \rangle \, e_i. \tag{2.20}$$

If we let $N \to \infty$, we see that $u \in \mathcal{D}(L)$ and $Bu = -Lu$. Now, we suppose $u \in \mathcal{D}(L)$. By definition, $\mathcal{D}(L) \subset \mathcal{D}(e^{-Lt})$, and we want to check that

$$\left. \frac{d\left(e^{-Lt}\right)}{dt} u \right|_{t=0} \tag{2.21}$$

exists in $H$. Using Abel's test, we can show that

$$\sum_{i=1}^{\infty} e^{-\lambda_i t} \langle u, e_i \rangle e_i \text{ converges uniformly to } e^{-Lt}u \tag{2.22}$$

and thus,

$$\frac{d\left(e^{-Lt}\right)}{dt}u\bigg|_{t=0} = \sum_{i=1}^{\infty} -\lambda_i e^{-\lambda_i t} \langle u, e_i \rangle e_i \bigg|_{t=0} = -Lu. \tag{2.23}$$

Therefore, $u \in \mathcal{D}(B)$ and $Bu = -Lu$, and we conclude that $-L$ is the generator of the $C_0$-semigroup, $e^{-Lt}$. To show that $e^{-Lt}$ is compact, we note that for each $t > 0$, the linear operator $e^{-Lt}$ is the uniform limit of linear operators with finite dimensional range. Each of these finite dimensional operators are compact, hence, $e^{-Lt}$ is compact. More details to this argument can be found in Naylor and Sell (1982) (p. 384).

### *2.2.2 Extending to Analytic Semigroups*

Recall that we claimed that the operator $-L$ in (2.1) generates an *analytic* semigroup. This is a special class of $C_0$-semigroups that admit two inequalities that will be useful in studying the long-term behavior of solutions of equations in the form of (2.1). Under the appropriate conditions, a $C_0$-semigroup, $e^{At}$, generated by $A$, on a Banach space $W$ is an analytic semigroup if and only if there are constants $M_0 \geq 1$ and $M_1 > 0$ and an $a \in \mathbb{R}$ such that

$$\|e^{At}w\| \leq M_0 e^{-at}\|w\| \quad \text{and} \quad \|Ae^{At}w\| \leq M_1 t^{-1} e^{-at}\|w\| \tag{2.24}$$

In short, an analytic semigroup is a $C_0$-semigroup, $T(z)$, extended to a sector of the complex plane, $\mathbb{C}$, that is complex differentiable. More precisely, a function $f : D \to W$, where $D$ is an open set of $\mathbb{C}$ and $W$ is a Banach space, is said to be analytic if, for any $z_0 \in D$, the strong limit

$$\lim_{z \to 0} \frac{1}{z} \left[ f(z_0 + z) - f(z_0) \right] \tag{2.25}$$

exists in $W$.

It will be useful for us to define the following open sectors in $\mathbb{C}$: Let $\delta \in (0, \pi)$ and $\sigma \in (0, \pi)$ then,

$$\Delta_\delta(a) \stackrel{\text{def}}{=} \{z \in \mathbb{C} : |\arg(z - a)| < \delta, z \neq a\}, \tag{2.26}$$

$$\Sigma_\sigma(a) \stackrel{\text{def}}{=} \{z \in \mathbb{C} : |\arg(z - a)| > \sigma, z \neq a\}, \tag{2.27}$$

where we write $\Delta_\delta$ and $\Sigma_\sigma$ for $\Delta_\delta(0)$ and $\Sigma_\sigma(0)$, respectively.

**Definition 2.2.2.** *Let $(T(t), A)$ be a $C_0$-semigroup on $W$. We say that $(T(t), A)$ is an analytic semigroup if there exists an extension of $T(t)$ to a mapping $T(z)$ defined for $z$ in some sector $\Delta_\delta \cup \{0\}$ and satisfying the following conditions:*

*(i) The mapping $z \to T(z)$ is a mapping of $\Delta_\delta \cup \{0\}$ into $\mathcal{L}(W)$.*

*(ii) $T(z_1 + z_2) = T(z_1)T(z_2)$ for all $z_1$ and $z_2$ in $\Delta_\delta \cup \{0\}$*

*(iii) For each $w \in W$, one has $T(z)w \to w$, as $z \to 0$ in $\Delta_\delta \cup \{0\}$*

*(iv) For each $w \in W$, the function $z \to T(z)w$ is an analytic mapping from $\Delta_\delta$ into $W$*

It turns out that there are certain properties on an infinitesimal generator $A$ that will guarantee that the $C_0$-semigroup $(T(t), A)$ is an analytic semigroup. This characterization is captured in the next theorem whose proof can be found in Sell and You (2002) and is based on ideas found in Pazy (1983). First, we give the following definition.

**Definition 2.2.3.** *A linear operator $A$ is said to be a sectorial operator on $W$ if $A : \mathcal{D}(A) \to W$, with $\mathcal{D}(A) \subset W$, and:*

*(i) $A$ is densely defined and closed*

*(ii) There exist* $a \in \mathbb{R}, \sigma \in (0, \frac{\pi}{2})$, *and* $M \geq 1$ *such that* $\Sigma_\sigma(a) \subset \rho(A)$, *and*

$$\|(\lambda I - A)^{-1}\| \leq \frac{M}{|\lambda - a|} \qquad \text{for all } \lambda \in \Sigma_\sigma(a) \tag{2.28}$$

**Theorem 2.2.4.** *Let* $(e^{-At}, -A)$ *be a* $C_0$*-semigroup on a Banach space* $W$, *and let* $M \geq 1$ *and* $a \in \mathbb{R}$ *be chosen so that* $\|e^{-At}\| \leq Me^{-at}$, *for all* $t \geq 0$. *Then the following statements are equivalent:*

*(i)* $e^{-At}$ *is an analytic semigroup, and there is an analytic semigroup extension* $e^{-Az}$ *of* $e^{-At}$, *defined on some sector* $\Delta_\delta \cup \{0\}$, *with* $0 < \delta < \pi/2$, *and a constant* $M_1 \geq M$ *such that* $\|e^{-Az}\| \leq M_1 e^{-aRe\, z}$, *for all* $z \in \Delta_\delta \cup \{0\}$

*(ii) The operator* $A$ *is a sectorial operator and one has*

$$\|(\lambda I - A)^{-1}\| \leq \frac{M_2}{|\lambda - a|} \qquad \text{for all } \lambda \in \Sigma_\xi(a), \tag{2.29}$$

*for appropriate constants* $M_2 \geq 1$ *and* $\xi \in (0, \frac{\pi}{2})$

*(iii) The semigroup* $e^{-At}$ *is differentiable for* $t > 0$, *and there is a constant* $M_3 > 0$ *such that* $\|Ae^{-At}\| \leq \begin{cases} M_3 t^{-1} e^{-at}, & \text{for } 0 < t \leq 1, \\ M_3 e^{-at}, & \text{for } t \geq 1. \end{cases}$

Now, for the $C_0$-semigroup $(e^{-Lt}, -L)$ generated by the operator $-L$ in (2.1) and described in the previous subsection, we can replace the time parameter $t$ with the complex variable $z$ to get

$$e^{-Lz}u = \sum_{i=1}^{\infty} e^{-\lambda_i z} \langle u, e_i \rangle e_i \qquad u \in H. \tag{2.30}$$

Notice that

$$\|e^{-Lz}u\|^2 \leq \sum_{i=1}^{\infty} |e^{-\lambda_i z}|^2 |\langle u, e_i \rangle|^2 \leq \sum_{i=1}^{\infty} e^{-2\lambda_i \mathrm{Re}\, z} |\langle u, e_i \rangle|^2 \leq e^{-2\lambda_1 \mathrm{Re}\, z} \|u\|^2 \tag{2.31}$$

for Re $z > 0$. Thus, $e^{-Lz}$ is well-defined for all $z$ in the sector $\Delta_{\frac{\pi}{2}} \cup \{0\}$, and using an argument similar to the one we used to verify (2.6) in the previous subsection, we can show that for all $z \in \Delta_{\frac{\pi}{2}}$

$$\frac{1}{z}\left[e^{-L(z_0+z)}w - e^{-Lz_0}w\right] \to -Le^{-Lz_0}w \qquad \text{as } z \to 0 \text{ in } \Delta_{\pi/2}. \tag{2.32}$$

This implies that the mapping $z \to e^{-Lz}w$ is an analytic mapping from $\Delta_{\frac{\pi}{2}}$ into $W$. Thus, by Theorem 2.2.4, we see that $L$ is a sectorial operator.

## 2.3 Fractional Power Spaces

Now that we have shown that the positive, self-adjoint, linear operator $L$, which has compact resolvent as in (2.1), is a sectorial operator and that $-L$ generates an analytic semigroup $e^{-Lt}$ on $H$, we can introduce some interpolation spaces that will be very useful in describing some of our results. The properties of the spectrum of $L$ allow us to define the function $f(\lambda) = \lambda^\alpha$ on the spectrum of $L$ for any $\alpha$. Recalling Equation (2.13), we let $\mathcal{H}^\alpha$ be the fractional power space associated with $L$ for all $\alpha$. More precisely, we set

$$\mathcal{H}^\alpha \stackrel{\text{def}}{=} \mathcal{D}\left(L^{\frac{\alpha}{2}}\right) \stackrel{\text{def}}{=} \left\{u \in H : \sum_{k=1}^{\infty} \lambda_k^{2\alpha} |\left\langle u, e_k \right\rangle|^2 < \infty \right\}, \tag{2.33}$$

where the $\{\lambda_k\}$ are the eigenvalues of $L$ with the associated eigenbasis $\{e_k\}$ and $\mathcal{H}^0 = H$. As indicated in the previous section, the operator $L^\alpha$ is given by

$$L^\alpha u = \sum_{k+1}^{\infty} \lambda_k^\alpha \left\langle u, e_k \right\rangle e_k \qquad \text{for all } u \in \mathcal{D}(\mathcal{H}^\alpha). \tag{2.34}$$

For $\alpha > 0$, we let $\mathcal{H}^{-\alpha}$ denote the dual space of $\mathcal{H}^\alpha$.

For each $\alpha \geq 0$, the set $\mathcal{H}^\alpha$ is a linear subspace in $H$. In fact, $\mathcal{H}^\alpha$ is itself a

Hilbert space with the $\mathcal{H}^\alpha$-norm defined as

$$\|x\|_\alpha^2 \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \lambda_k^\alpha |x_k|^2 = \|L^{\alpha/2}x\|^2 \quad x \in \mathcal{H}^\alpha, \tag{2.35}$$

where $x_k = \langle x, e_k \rangle$. Notice that

$$\|x\|_2^2 = \sum_{k=1}^{\infty} \lambda_k^2 |x_k|^2 = \|Lx\|^2, \qquad \text{for all } x \in \mathcal{H}^2. \tag{2.36}$$

Thus, for any $\beta \in \mathbb{R}$, $L : \mathcal{H}^{\beta+2} \to \mathcal{H}^\beta$, and $L$ is a linear isometry from $\mathcal{H}^{\beta+2}$ onto $\mathcal{H}^\beta$. Moreover, the family of $\mathcal{H}^\alpha$ spaces, $\alpha \in \mathbb{R}$ are a family of interpolation spaces. The structure of the spectrum of $L$ as given in (2.9), along with (2.35) imply that for $\alpha \geq \beta$, $\mathcal{H}^\alpha$ is continuously embedded in $\mathcal{H}^\beta$ and

$$\|x\|_\beta^2 \leq \lambda_1^{\beta-\alpha} \|x\|_\alpha^2. \tag{2.37}$$

In fact, since $L$ is compact, the embedding is compact. We also have intermediate interpolation spaces for $\mathcal{H}^\alpha$ and $\mathcal{H}^\beta$. That is, for every $\theta, 0 \leq \theta \leq 1$, there exists a constant $C(\alpha, \beta, \theta)$ such that

$$\|x\|_\gamma \leq C \|x\|_\alpha^\theta \|x\|_\beta^{1-\theta}, \qquad \text{for all } x \in \mathcal{H}^\alpha, \tag{2.38}$$

where $\gamma = \theta\alpha + (1-\theta)\beta$. We also note that the analytic semigroup $e^{-Lt}$ on $H$ extends to, or restricts to, an analytic semigroup on each $\mathcal{H}^\alpha, \alpha \in \mathbb{R}$.

Now that we have shown that $L$ is a sectorial operator and that $-L$ generates an analytic semigroup, $e^{-Lt}$, we state a few very important properties of these operators, which are based on ideas found in Sell and You (2002) Theorem 37.5.

**Theorem 2.3.1.** *Let $L$ be a positive, sectorial operator on a Hilbert space $H$, and let $S(t) = e^{-Lt}$ be the analytic semigroup generated by $-L$. Then the following statements hold:*

*(i) For any $r \geq 0$ and $t > 0$, the semigroup $e^{-Lt}$ maps $H$ into $\mathcal{H}^r$ and it is strongly continuous in $t > 0$.*

*(ii) For any $0 \leq \alpha \leq \beta$, there is a constant $C = C(\alpha, \beta) > 0$ such that*

$$\|e^{-Lt}\|_{\mathcal{L}(\mathcal{H}^\alpha, \mathcal{H}^\beta)} \leq Ct^{-(\beta-\alpha)/2} \quad \text{for all } t > 0 \tag{2.39}$$

*Proof.* Let $S^{(n)}(t)$ denote the $n^{th}$ derivative of $S(t)$ for $n = 1, 2, \ldots$. Since $S(t)$ is analytic, it is easy to show that

$$S^{(n)}(t) = -L^n S(t) = \left(-LS\left(\frac{t}{n}\right)\right)^n. \tag{2.40}$$

By Part (iii) of Theorem 2.2.4, we see that $S(t)$ maps $H$ into $\mathcal{H}^n$ for any positive integer $n$. Now, for any $r \geq 0$, we can take an integer $m > r$ and we have $e^{-Lt} : H \to \mathcal{H}^m \subset \mathcal{H}^\alpha$ since the $\mathcal{H}^\alpha$ spaces are interpolation spaces. The strong continuity will follow from (2.39), which we prove now. Let $x \in \mathcal{H}^\alpha$. Then

$$\|e^{-Lt}x\|_\beta^2 = \|L^{\beta/2}e^{-Lt}x\|^2 = \sum_{k=1}^{\infty} e^{-2\lambda_k t}\lambda_k^\beta x_k^2 = \sum_{k=1}^{\infty} e^{-2\lambda_k t}\lambda_k^{\beta-\alpha}\lambda_k^\alpha x_k^2 \leq (K^*)^2\|x\|_\alpha^2 \tag{2.41}$$

where,

$$(K^*)^2 = \sup_k \left(e^{-2\lambda_k t}\lambda_k^{\beta-\alpha}\right) \tag{2.42}$$

If we differentiate with respect to $\lambda$, we see that the function $\left(e^{-2\lambda_k t}\lambda_k^{\beta-\alpha}\right)$ is maximized when $\lambda_k = (\beta - \alpha)/2t$. Substituting this into (2.42) and taking the square root, we see that

$$K^* \leq \left(\frac{\beta - \alpha}{2e}\right)^{(\beta-\alpha)/2} t^{-(\beta-\alpha)/2}, \tag{2.43}$$

and the proof is complete. □

## 2.4 Assumptions on the Nonlinear Equation

Let us return to Equation (2.1)

$$dX_t = -LX_t dt + F(X_t, t)dt + QdW_t. \tag{2.44}$$

In this section we will describe in detail the additional assumptions we make on the operators $F(X_t, t)$ and $Q$.

**Definition 2.4.1.** *A linear operator $A$ on $H$ is called a Hilbert Schmidt operator if, for any orthonormal basis $\{e_k\}$ of $H$*

$$\sum_{k=1}^{\infty} \|Ae_k\|^2 < \infty. \tag{2.45}$$

*The Hilbert-Schmidt norm of $A$ is defined as*

$$\|A\|_{HS} = \left( \sum_{k=1}^{\infty} \|Ae_k\|^2 \right)^{\frac{1}{2}} \tag{2.46}$$

We let $\mathcal{U}$ be a second Hilbert space and assume that $Q$ is a bounded linear operator, $Q : \mathcal{U} \to H$, such that $Q(\mathcal{U}) \subset \mathcal{H}^{\gamma}$ for some $\gamma \geq 0$. Furthermore, we assume that $Q$ is a Hilbert-Schmidt operator from $\mathcal{U}$ into $\mathcal{H}^{\gamma-\rho}$ for some $\rho \in [0, 2)$. Now, suppose $\{g_k, k = 1, 2, \ldots\}$ is an orthonormal basis for $\mathcal{U}$. For $t \geq 0$, let $\{B_t^{(k)}\}$ be a sequence of independent standard Brownian motions in the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$. We assume that $W$ is a cylindrical Wiener process on $\mathcal{U}$. More precisely,

$$W_t = \sum_{k=1}^{\infty} g_k B_t^{(k)}. \tag{2.47}$$

If $Qg_k = f_k$, then the Hilbert-Schmidt assumption on $Q$ implies

$$\sum_{k=1}^{\infty} \|f_k\|_{\gamma-\rho}^2 < \infty. \tag{2.48}$$

This notation also allows us to write

$$QdW_t = \sum_{k=1}^{\infty} f_k dB_t^{(k)}. \tag{2.49}$$

We note that

$$\sum_{k=1}^{\infty} \|f_k\|_{\gamma-\rho}^2 = \sum_{k=1}^{\infty} \|L^{-\rho/2} f_k\|_{\gamma}^2, \tag{2.50}$$

so that the assumption that $Q$ is a Hilbert-Schmidt operator from $\mathcal{U}$ to $\mathcal{H}^{\gamma-\rho}$ is equivalent to assuming that the operator $L^{-\rho/2}Q$ is a Hilbert-Schmidt operator from $\mathcal{U}$ to $\mathcal{H}^{\gamma}$.

Now, we describe the nonlinear operator $F(X_t, t)$. We assume that there exists some $a \in [0, 2-\rho)$ such that the continous (nonlinear) map $F : \mathcal{H}^{\beta+a} \times [0,T] \to \mathcal{H}^{\beta}$ is locally Lipschitz in $\mathcal{H}^{\beta+a}$, Lipschitz continuous in time on closed intervals, and for some $p \geq 0$ satisfies

$$\sup_{s\in[0,T]} \|F(x,s)\|_{\beta} \leq C_{\beta}(T)(1 + \|x\|_{\beta+a}^p) \tag{2.51}$$

for all $\beta \in [-a, \gamma - \rho + 1 - a)$ and some constant $C_{\beta}(T)$.

## 2.5 Existence and Regularity of the Solution

Under the above assumptions, the following local existence result is standard.

**Proposition 2.5.1.** *For every initial $x \in \mathcal{H}^{\alpha}$ with $\alpha \in [0, \gamma - \rho + 1)$ and $s \geq 0$, there exists a positive random time $\tau$ and an adapted stochastic process $X_t$ with $X \in C([s, s+\tau); H)$ so that with probability one, $X_t$ is the unique solution to* (2.1) *and*

$$X_t = S(t-s)x + \int_s^t S(t-r)F(X_r, r)dr + \int_s^t S(t-r)QdW_r$$

*for $t \in [s, s+\tau)$.*

*Assumption of Finite Time Stability:* We assume there exists some time $T$ so that for any $x \in H$ and $s \in [0,T]$, Equation (2.1) has a unique mild solution $X_t$ on $[s,T]$ with $X_s = x$. Furthermore, $X_t$ satisfies

$$\sup_{r\in[s,T]} \|X_r\| < \infty \quad \text{a.s.}\,. \tag{2.52}$$

We now state a regularity result for solutions, which we will need subsequently.

**Proposition 2.5.2.** *Under the above assumptions let $X_t$ be the unique mild solution to* (2.1). *If $X_s \in \mathcal{H}^\alpha$ for some $s \in [0,T)$ and some $\alpha \in [0, \gamma - \rho + 1)$ almost surely, then it almost surely satisfies*

$$X \in C([s,T];\mathcal{H}^\alpha) \cap C((s,T];\mathcal{H}^\beta) \tag{2.53}$$

*for all $\beta \in [0, \gamma - \rho + 1)$.*

*Proof.* This result follows from a standard bootstrap argument. In the general PDE setting, it can be found in Sell and You (2002) for example. In almost exactly this setting it can be found in Hairer and Mattingly (2008). The maximum regularity in time is set by either the regularity of the initial data or the regularity of the stochastic convolution. □

We now outline a few of the ideas used in the proofs of the above propositions using some of the concepts that have already been discussed. First, we determine the regularity for the *stochastic convolution* given by the process

$$W_L(t) = \int_0^t S(t-s)QdW_s. \tag{2.54}$$

To check the regularity of $W_L(t)$ in $\mathcal{H}^\beta$ for some $\beta \geq 0$, we have

$$\mathbb{E}\left\|\int_0^t S(t-s)QdW_s\right\|_\beta^2 = \mathbb{E}\left\|\int_0^t \sum_{k=1}^{\infty} S(t-s)f_k dB_s^{(k)}\right\|_\beta^2 \tag{2.55}$$

$$\leq \sum_{k=1}^{\infty}\int_0^t \|S(t-s)f_k\|_\beta^2\, ds \tag{2.56}$$

$$\leq C\|Q\|_{\text{HS}}^2 \int_0^t (t-s)^{-(\beta-\gamma+\rho)}ds, \tag{2.57}$$

where the first inequality is a result of Ito's isometry, and the second inequality is a consequence of Theorem 2.3.1. The last integral is finite if $\beta-\gamma+\rho<1$. Therefore, for any fixed time $t$, the solution must be in $\mathcal{H}^\beta$ for $\beta \in [0, \gamma-\rho+1)$. Now, we check the continuity of the stochastic convolution in $\mathcal{H}^\beta$ for some $\beta \geq 0$.

$$\mathbb{E}\|W_L(t+s)-W_L(t)\|_\beta^2 \tag{2.58}$$

$$= \mathbb{E}\left\|\int_0^t [S(t+s-r)-S(t-r)]QdW_r + \int_t^{t+s} S(t+s-r)QdW_r\right\|_\beta^2 \tag{2.59}$$

$$\leq \mathbb{E}\left\|\int_0^t [S(t+s-r)-S(t-r)]QdW_r\right\|_\beta^2 + \mathbb{E}\left\|\int_t^{t+s} S(t+s-r)QdW_r\right\|_\beta^2 \tag{2.60}$$

$$\leq C\|Q\|_{\text{HS}}^2 s^{1-(\beta-\gamma+\rho)} \tag{2.61}$$

for $\beta \in [0, \gamma-\rho+1)$. Since $W_L(t+s)-W_L(t)$ is Gaussian, Fernique's Theorem (Da Prato and Zabczyk (1992) Theorem 2.6) tells us that $W_L(t+s)-W_L(t)$ has moments of all orders. Thus,

$$\mathbb{E}\|W_L(t+s)-W_L(t)\|_\beta^{2p} \leq C_p\|Q\|_{\text{HS}}^{2p} s^{p(1-\beta+\gamma-\rho)} \tag{2.62}$$

for any $p \geq 1$ and corresponding constant $C_p$. Then, by the Kolmogorov Continuity Theorem, we see that $W_L(t)$ is almost surely continuous in $\mathcal{H}^\beta$ for $\beta \in [0, \gamma-\rho+1)$. Now, we analyze the second term in the proposed solution $X_t$, the nonlinear operator "convolved" with the analytic semigroup generated by $-L$. First, we assume that we have finite time stability in $\mathcal{H}^\beta$. That is, for some $\beta \in [0, \gamma - \rho + 1)$ assume

$$\sup_{s\in[0,T]} \|X_s\|_\beta < \infty \qquad \text{a.s.} \tag{2.63}$$

Thus, for any $\sigma \geq 0$,

$$\left\| \int_0^t S(t-s)F(X_s,s)ds \right\|_{\beta+\sigma} \leq \int_0^t \|S(t-s)\|_{\mathcal{L}(\mathcal{H}^{\beta-a},\mathcal{H}^{\beta+\sigma})} \|F(X_s,s)\|_{\beta-a} \tag{2.64}$$

$$\leq C_{\beta,a}(T) \int_0^t \|S(t-s)\|_{\mathcal{L}(\mathcal{H}^{\beta-a},\mathcal{H}^{\beta+\sigma})} (1 + \|X_s\|_\beta^p) ds \tag{2.65}$$

$$\leq C(1 + \sup_{s\in[0,T]} \|X_s\|_\beta^p) \int_0^t (t-s)^{-\frac{\sigma+a}{2}} ds. \tag{2.66}$$

The last integral will be finite if $\sigma \in [0, 2-a)$. Since we assume that $0 \leq a < 2$, this is always possible. The fact that $\sigma > 0$ is allowed, show that our solution is indeed regularizing, and for $\sigma \in [0, 2-a)$, it is clear that the integral

$$\int_0^t S(t-s)F(X_s,s)ds \tag{2.67}$$

is continuous as a function of $t$ in the space $\mathcal{H}^{\beta+\sigma}$. We can conclude that for the analytic semigroup $S(t)$ and $x \in \mathcal{H}^\alpha$, the mapping $t \mapsto S(t)x$ is a continuous mapping of $[0, \infty)$ into $\mathcal{H}^\alpha$ and extends to a continuous mapping of $(0, \infty)$ into $\mathcal{H}^\beta$ for $\beta > \alpha$.

To prove the existence of a solution, we use a fixed point method. Define $\Upsilon_{T,x} : C([0,T], \mathcal{H}^\alpha) \to C([0,T], \mathcal{H}^\alpha)$ for any $\alpha \in [0, \gamma - \rho + 1)$ and $x \in \mathcal{H}^\alpha$ by

$$(\Upsilon_{T,x}u)(t) = S(t)x + \int_0^t S(t,s)F(u_s,s)ds + W_L(t)\,. \tag{2.68}$$

Using calculations a long the line of the preceding ones, we have

$$\sup_{t\in[0,T]} \|(\Upsilon_{T,x}u)(t) - (\Upsilon_{T,x}v)(t)\|_\alpha \leq C\big(1 + \sup_{t\in[0,T]} \|u_t\|_\alpha + \|v_t\|_\alpha\big)^p \sup_t \|u_t - v_t\|_\alpha T^{1-a} \tag{2.69}$$

Hence for each typical realization of $W_L$ and initial point $x$, $\Upsilon_{T,x}$ is a contraction in a ball of size $R$ for small enough $T$ about the function $S(t)x + W_L(t)$. Hence, the Banach fixed point theorem ensures the existence of a unique solution in $\mathcal{H}^\alpha$. Our observations on the continuity in time of the various other terms of the solution ensures that the solution is continuous in time in $\mathcal{H}^\alpha$ provided $\alpha \in [0, \gamma - \rho)$. (Otherwise $W_L$ is not continuous in $\mathcal{H}^\alpha$.) We obtain a stochastic flow, because we can choose one exceptional set of $W_L$ for all initial $x$. Thus, we have proven Proposition 2.5.1.

To see that the equation is regularizing, notice that

$$X_t = S(t/2)X_0 + \int_{t/2}^{t} S(t-s)F(X_s, s)ds + W_L(t) - W_L(t/2).$$

For any $\epsilon \in [0, 2-a)$ with $\beta + \epsilon \in [0, \gamma - \rho + 1)$ there exists a deterministic $C > 0$ and $\alpha \geq 0$ so that

$$\|X_t\|_{\beta+\epsilon} \leq Ct^{-\alpha}(1 + \sup_{s\in[t/2,T]} \|X_s\|_\beta^p + \sup_{s\in[t/2,T]} \|W_L(s)\|_\beta)\,. \tag{2.70}$$

Thus, if $X_t$ is bounded in $\mathcal{H}^\beta$ on a closed interval of time $[t, T]$ then we deduce that it is bounded in $\mathcal{H}^\alpha$ on the half-open interval $(t, T]$ for any $\alpha \in [\beta, \gamma - \rho + 1)$ by successively applying the above estimate.

Furthermore, we see that for any $s \in (0, T]$, $q > 1$ and $\alpha \in [0, \gamma - \rho + 1)$, there exist a $C$ and $p_1 > 1$ so that for any $X_0 \in H$

$$\mathbb{E} \sup_{r\in[s,T]} \|X_r\|_\alpha^q \leq C\big(1 + \mathbb{E} \sup_{r\in[0,T]} \|X_r\|^{p_1}\big). \tag{2.71}$$

Moreover, if $X_0 \in \mathcal{H}^\alpha$ then

$$\mathbb{E} \sup_{r\in[0,T]} \|X_r\|_\alpha^q \leq C\big(1 + \mathbb{E} \sup_{r\in[0,T]} \|X_r\|^{p_1}\big) \tag{2.72}$$

# 3

# Main Results

We are interested in comparing the dynamics of the process $X_t$ satisfying

$$dX_t = -LX_t dt + F(X_t, t)dt + QdW_t$$

with the dynamics of the process $Z_t$ on $H$ satisfying

$$dZ_t = -LZ_t dt + QdW_t. \tag{3.1}$$

We think of $Z_t$ as a "linearized" version of $X_t$, and we would like to understand when $X_t$ behaves essentially like a perturbation of $Z_t$.

The process $Z_t$ is called an *Ornstein-Uhlenbeck* process, and the solution to (3.1), for the initial value $Z_0 = x$, is given by

$$Z(t) = S(t)x + \int_0^t S(t-s)QdW_s\,. \tag{3.2}$$

Under the assumptions placed on the nonlinear equation, $X_t$, we see that

$$\int_0^t \|S(r)Q\|^2 dr \leq C\int_0^t \|e^{-Lr}\|^2_{\mathcal{L}(\mathcal{H}^{\gamma-\rho},H)}\|Q\|_{\mathrm{HS}} dr \leq C\|Q\|_{\mathrm{HS}}\int_0^t r^{-(\rho-\gamma)}dr < \infty\,. \tag{3.3}$$

Thus, by Theorem 5.2 of Da Prato and Zabczyk (1992), the law induced by the dynamics of $Z_t$ is a Gaussian measure with a mean $m_t$ and a covariance operator $\mathcal{S}_t$ given respectively by

$$m_t = S(t)x \qquad \text{and} \qquad \mathcal{S}_t = \int_0^t S(r)QQ^*S(r)^*dr\,, \tag{3.4}$$

where $\mathcal{S}_t$ is a trace class operator. We will define these quantities in greater detail in Chapter 5. However, we also note that under our assumptions, $Z_t$ has a unique invariant measure that we will denote by $\nu$. It is also Gaussian with mean 0 and covariance operator $\mathcal{S}_\infty$. Before we state our main results, we give a few definitions.

## 3.1 Definitions and Notation

**Definition 3.1.1.** *A Markov semigroup $P_t$ on the space of real-valued bounded mappings, $B_b(H)$, is itself a mapping*

$$P_t : [0,\infty) \mapsto \mathcal{L}(B_b(H)), \quad t \mapsto P_t \tag{3.5}$$

*such that*

*(i) $P_0 = 1, P_{t+s} = P_tP_s$ for all $t, s \geq 0$*

*(ii) For any $t \geq 0$ and $x \in H$ there exists a probability measure $\pi_t(x,\cdot)$ on $H$ such that*

$$P_t\varphi(x) = \int_H \varphi(y)\pi_t(x,dy) \quad \textit{for all } \varphi \in B_b(H) \tag{3.6}$$

*(iii) For any real-valued, uniformly continous, bounded mapping $\varphi$ and $x \in H$, the mapping $t \mapsto P_t\varphi(x)$ is continuous.*

Now, if we define

$$(\mathcal{R}_t\varphi)(x) = \mathbb{E}_x\varphi(Z_t) = \int_H \varphi(y)\mathbb{P}(Z_t \in dy|X_0 = x), \tag{3.7}$$

then we see that $\mathcal{R}_t$ is a Markov semigroup. Similarly, we can define the time inhomogeneous Markov operator $\mathcal{P}_{s,t}$ associated with the process $X_t$ given in (2.1) to be

$$\mathcal{P}_{s,t}\varphi(x) = \mathbb{E}[\varphi(X_t)|X_s = x] = \int_H \varphi(y)\mathbb{P}(X_t \in dy|X_s = x)\,. \tag{3.8}$$

When $X_t$ is time homogeneous, that is $F(x,s)$ is independent of $s$, then $\mathcal{P}_{s,s+t}$ only depends on $t$ and we will simply write $\mathcal{P}_t\varphi(x) = \mathbb{E}_x\varphi(X_t)$. Notice that $Z_t$ is time homogeneous, and hence $\mathcal{R}_{s,t} = \mathcal{R}_{t-s}$. By duality, both operators define transition kernels denoted by

$$\mathcal{R}_t(x,E) = \mathbb{P}(Z_t \in E|X_0 = x) \quad x \in H, E \in \mathcal{B}(H)$$
$$\mathcal{P}_{s,t}(x,E) = \mathbb{P}(X_t \in E|X_s = x) \quad x \in H, E \in \mathcal{B}(H).$$

Note that for the invariant measure $\nu$ of $Z_t$ described above, we have

$$\int_H \mathcal{R}_t\varphi d\nu = \int_H \varphi d\nu. \tag{3.9}$$

## 3.2 The Main Result

First, we need the following assumption and some of its consequences[1].

**Assumption 3.2.1** (Non-degeneracy of $Q$)**.** *For some $\delta \in [0, 2-a-\rho)$ and $c > 0$,*

$$cL^{-(\gamma+\delta)} \leq QQ^* \leq \frac{1}{c}L^{-\gamma}\,. \tag{3.10}$$

The following lemma, which is proven in Appendix A, provides a useful alternative characterization of Assumption 3.2.1 as well as an important consequence which we will use later.

**Lemma 3.2.2.** *Assumption 3.2.1 is equivalent to*

$$\mathcal{H}^{\gamma+\delta} \subset Q(\mathcal{U}) \subset \mathcal{H}^{\gamma}\,. \tag{3.11}$$

[1] For two linear operators $L$ and $K$ on $H$ we write $L \leq K$ if $\langle Lx, x\rangle \leq \langle Kx, x\rangle$ for all $x \in H$.

*In either case, $Q$ has a bounded pseudo-inverse on $\mathcal{H}^{\gamma+\delta}$. That is, there exists a bounded linear operator $Q^\dagger : \mathcal{H}^{\gamma+\delta} \to \mathcal{U}$ with $QQ^\dagger y = y$ for all $y \in \mathcal{H}^{\gamma+\delta}$.*

Now, we state our main theorem, which we will prove in later chapters.

**Theorem 3.2.3.** *Let Assumption 3.2.1 and the standing assumptions hold. Assume that there exists a positive $\alpha \in (\gamma + a + \delta - 1, \gamma - \rho + 1)$ such that for any $x \in \mathcal{H}^\alpha$*

$$\mathbb{E} \sup_{s\in[0,T]} \|X_s(\omega)\|_\alpha^{2p} < \infty \, . \tag{3.12}$$

*Then for all $x, y \in H$ and $s, t, r \in [0,T]$ with $s \le t$, we have $\mathcal{P}_{s,t}(x, \cdot\,) \sim \mathcal{R}_r(y, \cdot\,) \sim \nu$*

**Remark 3.2.4.** *Notice that by assumption $a+\delta+\rho < 2$, so that $\gamma+a+\delta-1 < \gamma-\rho+1$.*

**Remark 3.2.5.** *As a result of the comments on the regularization of solutions in Chapter 2, particularly Equations* (2.71) *and* (2.72), *we see that it is sufficient that*

$$\mathbb{E} \sup_{s\in[0,T]} \|X_s(\omega)\|^q < \infty$$

*for some sufficiently large $q$ and all $X_0 \in \mathcal{H}^\alpha$ to ensure that the estimate in* (3.12) *holds.*

# 4

# Examples

## 4.1 Navier-Stokes

The setting of this paper covers a number of important equations that model physical phenomena such as reaction diffusion equations and fluid equations. In this section, we will concentrate on the two dimensional Navier-Stokes equation since it is an example that was not covered by previous results and hence makes it a good candidate for the techniques we use throughout this work. For simplicity, we will consider the vorticity formulation of the problem on the two-dimensional torus. The vorticity $q(x,t)$ for $x = (x_1, x_2) \in \mathbb{T}^2 \stackrel{\text{def}}{=} [0, 2\pi]^2$ is defined from the velocity field $(u_1(x,t), u_2(x,t))$ by $q(x,t) = \partial_{x_2} u_1(x,t) - \partial_{x_1} u_2(x,t)$ and evolves according to

$$dq_t = \nu \Delta q_t dt + B(\mathcal{K} q_t, q_t) dt + Q dW_t \tag{4.1}$$

where the viscosity $\nu > 0$, the Biot-Savart $\mathcal{K}$ maps a scalar field $q \in L^2(\mathbb{T})$ to a vector field $u = (u_1, u_2) \in L^2(\mathbb{T}^2) \times L^2(\mathbb{T}^2)$, and $B(u,q) = (u \cdot \nabla) q$ again for a scalar field $q$ and a vector field $u$. If $e_k = \exp(ik \cdot x)$, with $k \in \mathbb{Z}^2$, is the $k$-th Fourier basis, then $\mathcal{K} q$ is defined by $\langle \mathcal{K} q, e_k \rangle = -i \langle q, e_k \rangle k^\perp / |k|^2$ where $k^\perp = (-k_2, k_1)$ and $|k|^2 = k_1^2 + k_2^2$. The following proposition establishes the appropriate conditions on the nonlinearity in this Navier-Stokes example that satisfy the assumptions used in the setting described in Chapter 2.

**Proposition 4.1.1.** *$B(\mathcal{K}\cdot,\cdot)$ is a continuous map from $\mathcal{H}^{\beta+1+\epsilon}\times\mathcal{H}^{\beta+1+\epsilon}\to\mathcal{H}^{\beta}$ for all $\beta\geq -1-\epsilon$*

Before we prove this proposition, we discuss an important property of interpolation spaces. First, for $\alpha>\beta$, we will denote by $[\mathcal{H}^{\alpha},\mathcal{H}^{\beta}]_{\theta}$, $\theta\in[0,1]$, the intermediate spaces between $\mathcal{H}^{\alpha}$ and $\mathcal{H}^{\beta}$. The norm of any element in one of these spaces obeys the standard interpolation inequality given in (2.38). Now, for any two given pairs of spaces, $\mathcal{H}^{x_1},\mathcal{H}^{x_2}$ and $\mathcal{H}^{y_1},\mathcal{H}^{y_2}$, with $x_1>x_2$ and $y_1>y_2$, if $A$ is a linear continuous operator from $\mathcal{H}^{x_1}$ into $\mathcal{H}^{y_1}$ and from $\mathcal{H}^{x_2}$ into $\mathcal{H}^{y_2}$, then $A$ is also a linear continuous operator from $[\mathcal{H}^{x_1},\mathcal{H}^{x_2}]_{\theta}$ into $[\mathcal{H}^{y_1},\mathcal{H}^{y_2}]_{\theta}$ for all $\theta\in[0,1]$. This fact will be useful in the following proof.

*Proof.* (Proof of Proposition 4.1.1) Based on the well known result in Constantin and Foias (1988), we have that for any triple $(s_1,s_2,s_3)$ with $s_i\geq 0$, $\sum_i s_i>1$ there exists a $C$ so that one has

$$|\langle B(\mathcal{K}q,v),w\rangle|\leq C\|q\|_{s_1-1}\|v\|_{s_2+1}\|w\|_{s_3} \tag{4.2}$$

$$|\langle B(\mathcal{K}q,v),w\rangle|\leq C\|q\|_{s_1-1}\|w\|_{s_2+1}\|v\|_{s_3} \tag{4.3}$$

for any $q,v,w$ in the appropriate space. From the second inequality, if we let $(s_1,s_2,s_3)=(1,\epsilon,0)$, we see that $B(\mathcal{K}\cdot,\cdot)$ is continuous from $H\times H\to\mathcal{H}^{-1-\epsilon}$ for any $\epsilon>0$. Also, since $\mathcal{H}^{s}$ is a multiplicative algebra for $s>1$, we have that

$$\|B(\mathcal{K}q,v)\|_{\beta}=\|(\mathcal{K}q)(\partial_1 v)+(\mathcal{K}q)(\partial_2 v)\|_{s}\leq C\|q\|_{\beta-1}\|v\|_{\beta+1} \tag{4.4}$$

for all $\beta>1$, and $B(\mathcal{K}\cdot,\cdot)$ is continuous from $\mathcal{H}^{\beta+1}\times\mathcal{H}^{\beta+1}\to\mathcal{H}^{\beta}$ for all $\beta>1$. Using the properties of interpolation spaces we discussed above, we conclude that $B(\mathcal{K}\cdot,\cdot)$ is continuous from $\mathcal{H}^{\beta+1+\epsilon}\times\mathcal{H}^{\beta+1+\epsilon}\to\mathcal{H}^{\beta}$ for any $\epsilon>0$ and for all $\beta\geq-1-\epsilon$. □

We now let $H^\alpha = \mathcal{H}^\alpha$ be the standard Sobolev space and summarize the implications of the above calculation.

**Theorem 4.1.2.** *Suppose the $f_k$ used to define the forcing in equation* (4.1) *are such that $\sum_k \|f_k\|^2_{H^{(\gamma-\rho)}} < \infty$ for some $\gamma \geq 0$ and $\rho \in [0,1)$. If there exists a $\delta \in [0, 1-\rho)$ such that Assumption 3.2.1 holds then the conditions of Theorem 3.2.3 hold for this 2D Navier-Stokes example.*

*Proof.* Proposition 4.1.1 implies that the assumption on the nonlinearity in our setting for the 2D Navier-Stokes equation is satisfied for $a = 1 + \epsilon$. Immediately we see that $\rho \geq 1$ violates our assumption for Theorem 3.2.3 that $a + \delta + \rho < 2$. Now, given $\delta \in [0, 1-\rho)$ that satisfies Assumption 3.2.1, we can choose $\epsilon > 0$ sufficiently small so that $\alpha$ in Theorem 3.2.3 is positive. Finally, it is well known that the moment condition in Remark 3.2.5 holds for all $q \geq 0$. (See for example E et al. (2001)). Thus, the conditions of Theorem 3.2.3 hold and consequently, we have the absolute continuity result. □

Notice that if we allow $f_k = e_k$, which would correspond to white in space forcing, then our theorem just fails. More specifically, our assumption on the nondegeneracy of $Q$ would imply $\gamma = \delta = 0$. Therefore, the assumption on the forcing would require

$$\sum_{k \in \mathbb{Z}^2} \|e_k\|^2_{\gamma-\rho} = \sum_{k \in \mathbb{Z}^2} |k|^{-2\rho} < \infty, \tag{4.5}$$

which in our two dimensional space is only possible for $\rho > 1$. However, if we take $f_k = e_k/|k|^\kappa$ for any $\kappa > 1$, then we have

$$\sum_{k \in \mathbb{Z}^2} \|e_k\|^2_{\gamma-\rho} = \sum_{k \in \mathbb{Z}^2} |k|^{-2\rho\kappa} < \infty, \tag{4.6}$$

and the conditions of Theorem 3.2.3 are satisfied by taking $\rho < 1$ and very close to 1 and fixing $a = 1 + \epsilon$ with $\epsilon > 0$ sufficiently small.

## 4.2 Cahn-Hilliard

We can also apply our results to the stochastic Cahn-Hilliard equation, an equation that describes the evolution of the phase separation of a binary alloy when the temperature has been quenched. We consider the equation on the three-dimensional cube with Neumann boundary conditions. The concentration of the material, denoted by $u(x,t)$, satisfies the equation

$$du_t = -\Delta^2 u_t dt + \Delta f(u_t) dt + Q dW_t \tag{4.7}$$

and we consider the case where $f(u) = u^3 - u$. We denote the nonlinear term in (4.7) by $G(u) = \Delta f(u)$.

We begin our analysis by establishing regularity properties of $G$. We let $H^m$ denote the Hilbert Sobolev space of functions $f$ in $L^2$ whose distribution derivatives of any order less than or equal to $m$ are in $L^2$. For $f, g \in H^m$ we define the norm and inner product in this space by

$$\|f\|_{H^m} = \sum_{[\alpha] \le m} \|D^\alpha f\|_{L^2} \qquad \langle f, g \rangle_{H^m} = \sum_{[\alpha] \le m} \langle D^\alpha f, D^\alpha g \rangle \,. \tag{4.8}$$

We not that when $m = s \notin \mathbb{N}$, we can still define the $H^s$ Sobolev norm by

$$\|f\|_{H^s}^2 = \sum_{n=1}^{\infty} (1+n^2)^s |\hat{f}(n)|^2 \tag{4.9}$$

where $\hat{f}$ is the Fourier series of $f$. Using standard Sobolev embeddings, we can show that for sufficiently smooth functions $u$ and some constant $C$,

$$\|\Delta(u^3 - u)\|_{H^0} = \left\| 6u \sum_{i=1}^{3} \left( \frac{\partial u}{\partial x_i} \right)^2 + (3u^2 - 1) \sum_{i=1}^{3} \frac{\partial^2 u}{\partial x_i^2} \right\|_{H^0} \le C \|u\|_{H^2}^3 \,. \tag{4.10}$$

In general,

$$\|\Delta f(u)\|_{H^s} = \left\| f^{(2)}(u)(\nabla u \cdot \nabla u) + f^{(1)}(u) \Delta u \right\|_{H^s} \,, \tag{4.11}$$

where $f^{(k)}$ is the $k$th derivative of $f$ with respect to $u$. We see from (4.10) that the nonlinear map $G$ is Lipschitz continuous from $H^2 \to H^0$. Also, using the fact that the Sobolev space $H^s$ is a multiplicative algebra for $s > 3/2$, from equation (4.11) we have $\|\Delta f(u)\|_{H^s} \leq C\|u\|_{H^{s+2}}^{\ell}$ for all $s > 3/2$ and some constants $C$ and $\ell$. Thus, $G$ is also Lipschitz continuous from $H^{s+2} \to H^s$ for all $s > 3/2$.

However, in order to apply the results of this paper to this particular example, we need our nonlinearity to map to fractional power spaces of negative order. This is not true in general. To overcome this, we redefine our base space be the Sobolev space $H^2$. This is made explicit in the proof of the following proposition.

**Proposition 4.2.1.** *$G$ is a continuous map from $\mathcal{H}^{\beta+1} \to \mathcal{H}^{\beta}$ for all $\beta \geq -1$.*

*Proof.* First, we redefine the base fractional power space by letting $\mathcal{H}^0 = H^2$ and $\mathcal{H}^{\alpha} = H^{2\alpha+2}$ for all $\alpha \in \mathbb{R}$. Notice that the operator $L = \Delta^2$ is still a strictly positive, selfadjoint, linear operator with compact resolvent on the space $H^2$, and thus the setting of our paper has not changed. Based on the calculations above, we now have that $G$ is a continuous map from $\mathcal{H}^0 \to \mathcal{H}^{-1}$ and from $\mathcal{H}^{s/2} \to \mathcal{H}^{s/2-1}$ for all $s > 3/2$. Finally, by interpolation we have that $G$ is a continuous map from $\mathcal{H}^{\beta+1} \to \mathcal{H}^{\beta}$ for all $\beta \geq -1$. □

We have a result similar to the Navier-Stokes result concerning the absolute continuity of the transition densities.

**Theorem 4.2.2.** *Suppose the $f_k$ used to define the forcing in equation (4.7) are such that $\sum_k \|f_k\|^2_{\mathcal{H}^{\gamma-\rho}} < \infty$ for some $\gamma \geq 0$ and $\rho \in [0, 1)$. If there exists a $\delta \in [0, 1-\rho)$ such that Assumption 3.2.1 holds, then the conditions of Theorem 3.2.3 hold for this 3D Cahn-Hilliard example.*

*Proof.* Proposition 4.2.1 implies that the assumption of Theorem 3.2.3 on the nonlinear map $G$ is satisfied for $a = 1$. We see that $\rho \geq 1$ would violate our assumption that

$a+\delta+\rho<2$. Now, given $\delta \in [0, 1-\rho)$, the conditions of Theorem 3.2.3 always hold with $a = 1$. We prove in Appendix B that the moment condition in Remark 3.2.5 holds for the stochastic Cahn-Hilliard equation illustrated here. Therefore, we have the absolute continuity result of Theorem 3.2.3 in the case of the stochastic Cahn-Hilliard equation. □

We can illustrate how the theorem works in the case of the Cahn-Hilliard equation on a closed interval of the real line. For simplicity, we let the $f_k = e_k$ and we take $\delta = 0$ and some $\rho < 1$. The Hilbert-Schmidt condition on our forcing would require,

$$\sum_{k\in\mathbb{N}} \|e_k\|^2_{\gamma-\rho} = \sum_{k\in\mathbb{N}} k^{-2\rho} < \infty\,. \tag{4.12}$$

Thus, we can choose any $1/2 < \rho < 1$ and the theorem holds.

# 5

# Absolute Continuity of the Linearized Equation

## 5.1 Gaussian Measures

We begin by formally defining Gaussian measures in a separable Hilbert space. For any probability measure $\mu$ on $(H, \mathcal{B}(H))$, there exists an $m \in H$ such that

$$\langle m, h \rangle = \int_H \langle x, h \rangle \, \mu(dx) \tag{5.1}$$

for all $h \in H$. We call $m$ the *mean* of $\mu$. There is also a bounded linear operator $S \in \mathcal{L}(H)$ such that

$$\langle Sh, k \rangle = \int_H \langle h, x - m \rangle \langle k, x - m \rangle \, \mu(dx) \tag{5.2}$$

for all $h, k \in H$. The operator $S$ is called the *covariance* of $\mu$. We have the following definition:

**Definition 5.1.1.** *Let $a \in H$ and let $S \in \mathcal{L}(H)$ be a positive definite, symmetric, and trace class operator. A Gaussian Measure $\mu \equiv N_{a,S}$ on $(H, \mathcal{B}(H))$ is a probability measure having mean $a$, covariance $S$, and a characteristic function (Fourier* *transform) given by*

$$\widehat{N_{a,S}}(h) := \int_H e^{i\langle h,x\rangle} N_{a,S}(dx) = e^{i\langle a,h\rangle - \frac{1}{2}\langle Sh,h\rangle}, \qquad h \in H. \tag{5.3}$$

We will find that an alternative characterization of Gaussian measures will be more useful than this definition. The symmetric operator $S$ is defined on all of $H$ and is trace class. Therefore, it is a self-adjoint, compact operator, and by the spectral theorem, there exists an orthonormal basis of $H$, which we will denote by $\{e_k, k = 1, 2, \ldots\}$, that are eigenvectors of $S$ with $Se_k = \sigma_k e_k$. We consider the natural isomorphism between $H$ and $\ell^2$ and reference the following theorem in Da Prato (2006):

**Theorem 5.1.2.** *A Gaussian measure $N_{a,S}$ in a separable Hilbert space $H$ is the restriction to $H$ (identified with $\ell^2$) of the product measure*

$$\mu = \bigotimes_{k=1}^{\infty} N_{a_k,\sigma_k}, \tag{5.4}$$

*which is defined on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$.*

Now, we will note a few properties of Gaussian measures. Hellinger distance, which we denote by $H_d^2(\mu, \nu)$, is used to quantify the similarity between two probability distributions.

**Definition 5.1.3.** *Let $\mu, \nu$ be any two probability measures on a measurable space $(X, \mathcal{B})$. Let $\lambda$ be a third probability measure such that $\mu \ll \lambda$ and $\nu \ll \lambda$. Then,*

$$H_d^2(\mu, \nu) := \frac{1}{2} \int \left( \sqrt{\frac{d\mu}{d\lambda}} - \sqrt{\frac{d\nu}{d\lambda}} \right)^2 d\lambda = 1 - \int \sqrt{\frac{d\mu}{d\lambda}\frac{d\nu}{d\lambda}} d\lambda. \tag{5.5}$$

*We call $\int \sqrt{\frac{d\mu}{d\lambda}\frac{d\nu}{d\lambda}} d\lambda$ the Hellinger Integral, denoted by $H(\mu, \nu)$.*

The following properties of the Hellinger integral hold true:

(i) $0 \leq H(\mu,\nu) \leq 1$

(ii) $H(\mu,\nu) = 1 \Leftrightarrow \mu = \nu$

(iii) $H(\mu,\nu) = 0 \Leftrightarrow \mu \perp \nu$

(iv) If $\mu \sim \nu$ then $H(\mu,\nu) > 0$ and $H(\mu,\nu) = \int \sqrt{\frac{d\nu}{d\mu}} d\mu$

Notice that if we let $\{\mu_k\}$ and $\{\nu_k\}$ be sequences of probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and we consider the product measures

$$\mu = \bigotimes_{k=1}^{\infty} \mu_k \quad \text{and} \quad \nu = \bigotimes_{k=1}^{\infty} \nu_k \tag{5.6}$$

on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$, then as a result of Fubini's theorem we have

$$H(\mu,\nu) = \prod_{k=1}^{\infty} H(\mu_k, \nu_k). \tag{5.7}$$

The next theorems are fundamental in proving the results on absolute continuity that we are interested in. The first is a classic theorem of Kakutani on infinite product measures.

**Theorem 5.1.4.** *(Kakutani) Let $\{\mu_k\}$ and $\{\nu_k\}$ be sequences of probability measures on $(\mathbb{R}, \mathcal{B})$ such that $\mu_k \sim \nu_k$ for all $k \in \mathbb{N}$, and let $\mu = \bigotimes_{k=1}^{\infty} \mu_k$ and $\nu = \bigotimes_{k=1}^{\infty} \nu_k$. If $H(\mu,\nu) > 0$, then $\mu$ and $\nu$ are equivalent and we have*

$$\frac{d\nu}{d\mu}(x) = \lim_{n\to\infty} \prod_{i=1}^{n} \frac{d\nu_k}{d\mu_k}(x_k) \quad \textit{in } L^1(\mathbb{R}^\infty, \mu). \tag{5.8}$$

*If $H(\mu, \nu) = 0$, then $\mu$ and $\nu$ are singular.*

An immediate consequence of this theorem is that two Gaussian measures are either equivalent or singular. The next two theorems give us criteria for determining when they are absolutely continuous.

**Theorem 5.1.5.** *(Cameron-Martin Formula) Let $\mu = N_{0,S}$ and $\nu = N_{a,S}$ be two Gaussian measures on $(H, \mathcal{B}(H))$, with $a \in H$ and $S \in \mathcal{L}_1^+(H)$. Then, the following are true:*

*(i) If $a \notin S^{1/2}(H)$, then $\mu$ and $\nu$ are singular*

*(ii) If $a \in S^{1/2}(H)$, then $\mu$ and $\nu$ are equivalent*

*(iii) If $\mu \sim \nu$, then the density $\frac{d\nu}{d\mu}$ is given by*

$$\frac{d\nu}{d\mu}(x) = \exp\left(-\frac{1}{2}\left|S^{1/2}a\right|^2 + \left\langle S^{1/2}a, S^{1/2}x\right\rangle\right) \quad x \in H \tag{5.9}$$

The space $S^{1/2}(H)$ is called the *Cameron-Martin* space.

**Theorem 5.1.6.** *Let $\mu = N_{0,S}$ and $\nu = N_{0,R}$ be two Gaussian measures on $(H, \mathcal{B}(H))$ with $S, R \in \mathcal{L}_1^+(H)$. Then, $\mu$ is equivalent to $\nu$ if and only if $R = S^{1/2}TS^{1/2}$ for some positive definite, bounded, invertible operator $T$ such that $T - I$ is a Hilbert-Schmidt operator on $H$.*

*Proof.* The details of the proof can be found in Kuo (1975), Theorem 3.2. □

## 5.2 $Z_t$ in a Simplified Setting

In this section, we hope to provide some insight into the implications and meaning of some of the assumptions made in section 2. We consider the linear equation given in (3.1)

$$dZ_t = -LZ_t dt + QdW_t \tag{5.10}$$

on the space $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ and recall that the set $\{e_k, k = 1, 2, \ldots\}$ is the orthonormal eigenbasis for the operator $L$ with the associated eigenvalues $\{\lambda_k\}$. Now we simplify the setting even further by assuming that the $f_k = \sigma_k e_k$ with $\sigma_k > 0$ and $\sum_k \sigma_k^2/\lambda_k < \infty$. In this setting, (3.1) diagonalizes into infinitely many uncoupled one dimensional Ornstein-Uhlenbeck processes. That is,

$$dZ_t^{(k)} = -\lambda_k Z_t^{(k)} dt + \sigma_k dB_t^{(k)} \quad k = 1, 2, \ldots \tag{5.11}$$

In this simplified setting, we will investigate the conditions under which $\mathcal{R}_t(x, \cdot) \sim \mathcal{R}_s(y, \cdot)$ for various possible choices of $x$ and $y$ and $t$ and $s$ in this simplified setting to give us better insight into how the structure of the operators affect absolute continuity. If we consider the space of transition measures $\mathcal{R}_t(x, \cdot)$ for all $t > 0$ and $x \in H$. Then, the space foliates into disjoint leaves $H_{x,t}$ defined by

$$H_{x,t} \stackrel{\text{def}}{=} \{(y, s) \in H \times (0, \infty) : \mathcal{R}_s(y, \cdot) \sim \mathcal{R}_t(x, \cdot)\}\,. \tag{5.12}$$

Thus, on each leaf, $H_{x,t}$, $\mathcal{R}_t(x, \cdot)$ is equivalent to the transition density starting from any other point in the leaf. Also, at any finite time $t$, only the points in the leaf $H_{x,t}$, including 0, are equivalent to the invariant measure $\nu$ at any finite time.

For all initial values $y$, $Z_t$ has a unique invariant measure $\nu$, and the transition density $\mathcal{R}_t(y, \cdot)$ converges to $\nu$ as $t \to \infty$. However, the convergence is not necessarily in the total variation norm. For two probability measures, $\mu$ and $\nu$ on a $\sigma$-algebra $\mathcal{F}$, the total variation norm between them is given by

$$\|\mu - \nu\|_{\text{TV}} = \sup\{|\mu(A) - \nu(A)| : A \in \mathcal{F}\}. \tag{5.13}$$

If $\mathcal{R}_t(x, \cdot) \xrightarrow{\text{TV}} \nu$, this would imply that for any finite time, all transition densities are equivalent.

Now, in this simplified setting, we have the following proposition:

**Proposition 5.2.1.** *For any $x, y \in H$ and $t > 0$, $H_{x,t} = H_{y,t}$ if and only if*

$$\sum_k \frac{\lambda_k}{\sigma_k^2}(x_k - y_k)^2 \, e^{-2\lambda_k t} < \infty \,. \tag{5.14}$$

*Similarly, for any $x \in H$ and $t, s > 0$, $H_{x,t} = H_{x,s}$ if and only if*

$$\sum_{k=1}^{\infty} e^{-4\lambda_k (s \wedge t)} < \infty \,. \tag{5.15}$$

*In particular, if for some choice of $\alpha \geq 0$ and $t > 0$*

$$\sup_k \frac{e^{-2\lambda_k t}}{\lambda_k^{\alpha-1}\sigma_k^2} < \infty \,,$$

*then $H_{x,t} = H_{y,t}$ if $x - y \in \mathcal{H}^\alpha$. And if $\sum_k e^{-4\lambda_k t} < \infty$ then $H_{x,t} = H_{x,s}$ for all $x \in H$ and $s > t$.*

*Proof.* We begin by proving the first claim in the above proposition. Without loss of generality, we will show that $\mathcal{R}_t(0, \cdot\,) \sim \mathcal{R}_t(z, \cdot\,)$ where $z = x - y$. Recall that the law of $\mathcal{R}_t(z, \cdot\,)$ is a Gaussian measure with mean $m$ and covariance operator $\mathcal{S}_t$ as defined in (3.4). In our current simple case we have

$$\langle m_t, e_k \rangle = e^{-\lambda_k t} z_k \qquad \text{and} \qquad \langle \mathcal{S}_t e_k, e_j \rangle = \delta_{kj} \frac{\sigma_k^2}{2\lambda_k}\big(1 - e^{-2\lambda_k t}\big) \,.$$

Hence, the Cameron-Martin space associated with this Gaussian measure is

$$\mathcal{S}^{\frac{1}{2}}(H) = \Big\{ x \in H : \sum_k \frac{\lambda_k}{\sigma_k^2} x_k^2 < \infty \Big\} \,.$$

Theorem 5.1.5 implies that $\mathcal{R}_t(0, \cdot\,) \sim \mathcal{R}_t(z, \cdot\,)$ if and only if $m_t \in \Sigma^{\frac{1}{2}}(H)$ and that otherwise $\mathcal{R}_t(0, \cdot\,) \perp \mathcal{R}_t(z, \cdot\,)$. Thus, for equivalence it is necessary and sufficient to require that

$$\sum_k \frac{\lambda_k}{\sigma_k^2} z_k^2 e^{-2\lambda_k t} < \infty \,,$$

and the first claim of Proposition 5.2.1 is proved.

We now show that $\mathcal{R}_t(x,\,\cdot\,) \sim \mathcal{R}_s(x,\,\cdot\,)$ when (5.15) holds. Without loss of generality, we consider the measures $\mathcal{R}_t(0,\,\cdot\,)$ and $\mathcal{R}_s(0,\,\cdot\,)$. Define the operator $T\colon H \to H$ to be

$$T = \mathcal{S}_t^{-1/2}\mathcal{S}_s\mathcal{S}_t^{-1/2} \qquad \text{where} \qquad \langle Te_k, e_j\rangle = \delta_{kj}\frac{(1-e^{-2\lambda_k s})}{(1-e^{-2\lambda_k t})}\,.$$

Recall that the $\{\lambda_k\}$ are such that $0 < \lambda_1 \leq \lambda_2 \leq \cdots$. Therefore, $T$ is a positive, bounded, and invertible operator. Now,

$$\sum_k |(T-I)e_k|^2 = \sum_k e^{-4\lambda_k(s\wedge t)}\frac{(e^{-2\lambda_k|t-s|}-1)^2}{(1-e^{-2\lambda_k t})^2}|e_k|^2$$

is finite if and only if (5.15) holds since $|e_k|^2 = 1$ and

$$\frac{(e^{-2\lambda_k|t-s|}-1)^2}{(1-e^{-2\lambda_k t})^2} \tag{5.16}$$

is bounded from above and below by positive constants uniformly for all $k$ . This shows that $T - I$ is a Hilbert-Schmidt operator and by Theorem 5.1.6 we have $\mathcal{R}_t(x,\,\cdot\,) \sim \mathcal{R}_s(x,\,\cdot\,)$.

To see the last claim in the proposition, recall that since $z \in \mathcal{H}^\alpha$, we know that $\sum \lambda_k^\alpha z_k^2 < \infty$. Therefore,

$$\sum_k \frac{\lambda_k}{\sigma_k^2} z_k^2 e^{-2\lambda_k t} \leq \Big(\sup_k \frac{e^{-2\lambda_k t}}{\lambda_k^{\alpha-1}\sigma_k^2}\Big)\|z\|_\alpha^2\,,$$

which proves the last claim. □

Now, let us analyze the implications of Proposition 5.2.1. We see that it is possible to have mutually singular transition densities from various points $x, y \in \mathbb{R}^\infty$. To illustrate this point, let us a choose an operator $L$ that satisfies our standing assumptions. For example, let $L = -\Delta$ with $\lambda_k = |k|^2$. We will consider three cases in which the $\sigma_k$ decay at different rates and result in very distinct conclusions for the transition densities.

1. *Suppose* $\sigma_k = \exp(-c\lambda_k^b)$.

   If $b > 1$, Proposition 5.2.1 tells us that there are a dense set of points such that $\mathcal{R}_t(x, \cdot) \perp \mathcal{R}_t(y, \cdot)$ for any two points $x$ and $y$ in the dense set.

2. *Suppose* $\sigma_k = \lambda_k^{-b}$ *for* $b > 0$ *or* $\sigma_k = \exp(-c\lambda_k^\gamma)$ *for some* $\gamma \in (0, 1)$.

   In both of these cases, Proposition 5.2.1 tells us that $\mathcal{R}_t(x, \cdot) \sim \mathcal{R}_t(y, \cdot)$ for all $x, y \in \mathbb{R}^\infty$.

3. *Suppose* $\sigma_k = \exp(-c\lambda_k)$.

   In this case, for any fixed $t$, there exists a dense set of points where $\mathcal{R}_t(x, \cdot) \perp \mathcal{R}_t(y, \cdot)$ for for all points in the set. However, at a later time, the measures will be equivalent for the same collection of points.

## 5.3 Equivalence of the Ornstein-Uhlenbeck Process

In this section, we return to our standing assumptions as described in Chapter 2. We will show that the under Assumption 3.2.1, $\mathcal{R}_t(x, \cdot) \sim \mathcal{R}_s(y, \cdot)$ for all $x, y \in H$ and $t, s > 0$. Thus, the result in Theorem 3.2.3 is to be expected from the linear equation.

If we consider the sequence of $\{f_k\}$ as in Equation (2.48). Then, it is reasonable to equate $f_k \in \mathcal{H}^s$ with $\lambda_k^s$ for some $s > 0$. In light of the previous discussion in Section 5.2 concerning the various possibilities for $\{\sigma_k\}$, in particular when $\sigma_k = \lambda_k^{-b}$, the following result is not surprising.

**Lemma 5.3.1.** *Under Assumption 3.2.1, for any* $t, s > 0$ *and* $x, y \in H$ *one has* $\mathcal{R}_t(x, \cdot) \sim \mathcal{R}_s(y, \cdot) \sim \nu$.

*Proof.* We begin by proving that $\mathcal{R}_t(x,\,\cdot\,) \sim \mathcal{R}_t(y,\,\cdot\,)$. As above, without loss of generality, we will prove the lemma for $\mathcal{R}_t(0,\,\cdot\,)$ and $\mathcal{R}_t(z,\,\cdot\,)$, where $z = x - y$. We first show that $\mathcal{S}_t^{1/2}(H) \supset \widehat{\mathcal{S}}_t^{1/2}(H)$, where we have defined

$$\widehat{\mathcal{S}}_t = \int_0^t S(t-s)L^{-(\gamma+\delta)}S(t-s)^*ds \tag{5.17}$$

and $\mathcal{S}_t$ is the covariance operator defined in Section 3. Now, we have

$$\begin{aligned}\left\|\mathcal{S}_t^{\frac{1}{2}}x\right\|^2 = \langle \mathcal{S}_t x, x\rangle &= \int_0^t \langle S(t-s)QQ^*S(t-s)^*x, x\rangle\, ds \\ &= \int_0^t \langle QQ^*S(t-s)^*x, S(t-s)^*x\rangle\, ds \\ &\geq c\int_0^t \left\langle L^{-(\gamma+\delta)}S(t-s)^*x, S(t-s)^*x\right\rangle ds \\ &= c\left\langle \widehat{\mathcal{S}}_t x, x\right\rangle = c\left\|\widehat{\mathcal{S}}_t^{\frac{1}{2}}x\right\|^2\end{aligned}$$

and by Proposition A.1.1, $\mathcal{S}_t^{1/2}(H) \supset \widehat{\mathcal{S}}_t^{1/2}(H)$.

Now, to use the Cameron-Martin theorem, we need to show that there exists some $x \in H$ such that $\mathcal{S}_t^{1/2}x = S(t)z$. For any $x \in H$ with $x = \sum x_k e_k$, we have

$$\widehat{\mathcal{S}}_t^{\frac{1}{2}}x = \sum_k \int_0^t S(t-s)L^{\frac{-(\gamma+\delta)}{2}}x_k e_k = \sum_k (1-e^{-\lambda_k t})\lambda_k^{\frac{-(\gamma+\delta)}{2}-1}x_k e_k\,.$$

Therefore, if we define the norm

$$|x|_{\widehat{\mathcal{S}}} = \|\widehat{\mathcal{S}}_t^{-1/2}x\|_0 \tag{5.18}$$

then $|\cdot|_{\widehat{\mathcal{S}}}$ is an equivalent norm to $\|\cdot\|_{\gamma+\delta+2}$. Thus, $x \in \widehat{\mathcal{S}}_t^{-1/2}(H)$ is a property that is equivalent to $x \in \mathcal{H}^{\gamma+\delta+2}$, and if $S(t)z \in \mathcal{H}^{\gamma+\delta+2}$, then $S(t)z \in \mathcal{S}_t^{\frac{1}{2}}(H)$. Thus, by Theorem 2.3.1 we have $\mathcal{R}_t(x,\,\cdot\,) \sim \mathcal{R}_t(y,\,\cdot\,)$.

Lastly, we show that $\mathcal{R}_t(x,\,\cdot\,) \sim \nu$ for all $x \in H$ and $t > 0$. The final result follows then because equivalence of measures is an equivalence relation. We start by observing that for any measurable $A \subset H$

$$\nu(A) = \int \mathcal{R}_t(y, A)\nu(dy)\,.$$

Hence, if $\nu(A) = 0$ we know that $\mathcal{R}_t(y, A) = 0$ for $\mu$-a.e. $y$. We have just shown that $\mathcal{R}_t(x, \cdot\,) \sim \mathcal{R}_t(y, \cdot\,)$ for all $x, y \in H$, so $\mathcal{R}_t(x, \cdot\,) \ll \nu$. On the other hand, $\mathcal{R}_t(x, \cdot\,) \sim \mathcal{R}_t(y, \cdot\,)$ implies that if $\mathcal{R}_t(x, \cdot\,) = 0$ for some $x \in H$, then it is 0 for all $x \in H$. Therefore, by the above representation of $\nu$, we see that $\nu \ll \mathcal{R}_t(x, \cdot\,)$, which concludes the proof. □

# 6

# Proof of the Main Results

## 6.1 The Auxiliary Process

In this section, we introduce an auxiliary process $Y_t$ associated with the process $X_t$ given in (2.1) that will help us to prove the main theorem of this paper.

**Definition 6.1.1.** *The law of a stochastic process $X_t$ on pathspace is a probability measure on the entire path $X : [0,T] \to W$ of the process $X_t$, with $X_0 = x$, which we denote by $P^X_{[0,T]}(x, \cdot)$. If $M([0,T]; W)$ is the space of mappings from the interval $[0,T]$ to $W$, then $P^X_{[0,T]}(x, E) = \mathbb{P}(X \in E)$ for any $E \in \mathcal{B}(M)$*

Recall that $S(t) = e^{-Lt}$ was the analytic semigroup generated by $-L$. Using the variation of constants formula, we can write the mild solution to (2.1) as

$$X_t(\omega) = S(t)X_0 + \int_0^t S(t-s)F(X_s(\omega), s)ds + \int_0^t S(t-s)QdW_s(\omega)\,.$$

We can rewrite this equation as

$$X_t(\omega) = S(t)X_0 + \int_0^t S\left(\tfrac{t-s}{2}\right) S\left(\tfrac{t-s}{2}\right) F(X_s(\omega), s)ds + \int_0^t S(t-s)QdW_s(\omega)\,.$$

If we do a simple substitution of $s \mapsto 2s - t$ in the first integral, we have

$$X_t = S(t)X_0 + 2\int_{t/2}^{t} S(t-s)S(t-s)F(X_{2s-t}, 2s-t)ds + \int_0^t S(t-s)QdW_s\,. \quad (6.1)$$

Now, fix $t$ and define the stochastic process $G_s$ by

$$G_s(\omega) = \begin{cases} 0 & s < t/2 \\ 2S(t-s)F(X_{2s-t}(\omega), 2s-t) & t/2 \le s \le t \end{cases}.$$

Observe that $G_s$ is adapted to the filtration generated by $\{W_r : r \le s\}$. It is clearly not Markovian. Nonetheless, we can write a system of Itô equations

$$dX_s(\omega) = -LX_s(\omega)ds + F(X_s(\omega), s)ds + QdW_s(\omega) \quad (6.2a)$$

$$dY_s(\omega) = -LY_s(\omega)ds + G_s(\omega)\,ds + QdW_s, \qquad Y_0 = X_0\,, \quad (6.2b)$$

where (6.1) looks almost like the mild solution to (6.2b). Note that for $s < \frac{t}{2}$, the $Y$-component of this system is an Ornstein-Uhlenbeck process, and using the variation of constants formula, the solution to this equation can be written explicitly as

$$Y_s(\omega) = \begin{cases} S(s)X_0 + \int_0^s S(s-r)QdW_r(\omega) & s \in (0, t/2] \\ S(s)X_0 + \int_{t/2}^s S(s-r)G_r(\omega)dr + \int_0^s S(s-r)QdW_r(\omega) & s \in [t/2, t] \end{cases}.$$

Notice that the solution to the $Y$ equation is an adapted functional of the solution to the $X$ equation. Thus, it is clear that the solution exists. The auxiliary process is useful to us because at time $t$, $Y_t = X_t$, and the two processes are equal only at time $t$. To be explicit, observe that the solution to (6.2b) at time $t$ is

$$\begin{aligned} Y_t(\omega) &= S(t)X_0 + \int_{t/2}^t S(t-s)G_s(\omega)\,ds + \int_0^t S(t-s)QdW_s(\omega) \\ &= S(t)X_0 + 2\int_{t/2}^t S(t-s)S(t-s)F(X_{2s-t}(\omega), 2s-t)ds + \int_0^t S(t-s)QdW_s(\omega)\,. \end{aligned}$$

Comparing the last line with (6.1) shows that $Y_t = X_t$.

Now, we will denote by $P^Y_{[0,t]}(x, \cdot)$ and $P^Z_{[0,t]}(x, \cdot)$ the measures induced by the $Y(\omega)$ and $Z(\omega)$ processes, respectively, on the pathspace $C([0,t]; H)$. As just observed, $Y_t = X_t$ only at time $t$, but not for general time. Therefore, if we can show that the law of the $Y$ process, defined by the solution to (6.2b), on pathspace is equivalent to the law of the $Z$ process, defined by the solution to (3.1), on pathspace, then the time $t$ transition probabilities of the process $X_t$ given in (2.1) are equivalent to the time $t$ transition probabilities of $Z_t$. That is if $P^Y_{[0,t]}(x, \cdot) \sim P^Z_{[0,t]}(x, \cdot)$ as measures on $C([0,t]; H)$, then $\mathcal{P}_{0,t}(x, \cdot) \sim \mathcal{R}_t(x, \cdot)$ as measures on $H$. This observation will be critical in proving Theorem 3.2.3.

## 6.2 Absolute Continuity in One Direction

Based on the observation made at the end of Section 6.1, our first step in proving Theorem 3.2.3 will be to show the equivalence of the pathspace measures induced by the Ornstein-Uhlenbeck process $Z_t$ and the auxiliary process $Y_t$, which was defined in Section 6.1. In this section, we will show the $Y$ pathspace measure is absolutely continuous with respect to the $Z$ pathspace measure.

An important tool in proving this absolute continuity is Girsanov's theorem, a version of which is stated below.

**Theorem 6.2.1.** *(Girsanov) Let $X_t$ be a stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a Hilbert space $H$ of the form*

$$dX_t = AX_t dt + F(X_t)dt + QdW_t \tag{6.3}$$

*where for some Hilbert space $U$, $Q^{1/2} : U \to U_0$, and for any $u \in U_0$, $\|u\|_{U_0} = \|Q^{-1/2}u\|_U$. If $\psi(\cdot)$ is a $U_0$-valued $\mathcal{F}_t$-predictable process such that*

$$M_t = \exp\left( -\int_0^t \langle \psi(s), dW_s \rangle_{U_0} - \frac{1}{2} \int_0^t \|\psi(s)\|^2_{U_0} ds \right) \tag{6.4}$$

*is a square integrable martingale, then the process*

$$\widehat{W}_t = W_t + \int_0^t \psi(s)ds \tag{6.5}$$

*is a a cylindrical Wiener process with respect to* $\{\mathcal{F}_t\}_{t\geq 0}$ *on the probability space* $(\Omega, \mathcal{F}, \widehat{\mathbb{P}})$ *where*

$$d\widehat{\mathbb{P}} = \exp\left(-\int_0^t \langle \psi(s), dW_s\rangle_{U_0} - \frac{1}{2}\int_0^t \|\psi(s)\|_{U_0}^2 ds\right) d\mathbb{P}. \tag{6.6}$$

**Remark 6.2.2.** *The condition*

$$\mathbb{P}\left(\int_0^t \|\psi(s)\|_{U_0}^2 ds < \infty\right) = 1 \tag{6.7}$$

*is enough to guarantee that* $M_t$ *in Theorem 6.2.1 is a square integrable martingale. (See Lemma 10.15 in Da Prato and Zabczyk (1992)*

Now, we prove the absolute continuity result, which is stated in the following proposition.

**Proposition 6.2.3.** *Let Assumption 3.2.1 and the standing assumptions hold. Fixing a positive* $\alpha$ *with* $\alpha \in (\gamma + a + \delta - 1, \gamma - \rho + 1)$*, if* $x \in \mathcal{H}^\alpha$ *and* $s, t \in [0, T]$ *with* $s \leq t$*, then* $P^Y_{[s,t]}(x, \cdot\,) \ll P^Z_{[s,t]}(x, \cdot\,)$ *as measures on* $C([s,t]; H)$*. As a consequence,* $\mathcal{R}_{t-s}(x, \cdot\,) \ll \mathcal{P}_{s,t}(x, \cdot\,)$*.*

*Proof.* Without loss of generality, we can assume that $\alpha < \gamma + 1$ and take $s = 0$. Observe that by setting $\beta = \alpha - a$, we have $\beta \in (\gamma + \delta - 1, \gamma - \rho + 1 - a)$ which is a non-empty interval since we assumed $\delta + a + \rho < 2$. Notice that the conditions on the operator $F$ and the conditions on the process $X_s$ together imply that

$$\sup_{0\leq s\leq t} \|F(X_s, s)\|_\beta \leq C(t)(1 + \sup_{0\leq s\leq t} \|X_s\|_\alpha) < \infty \qquad \text{a.s.} \tag{6.8}$$

Fix any positive $r \in \big(1, 1/(\gamma+\delta-\beta)\big)$. This always possible since our assumptions ensure that $\gamma+\delta-\beta < 1$. Now define $q$ by $1/r + 1/q = 1$. Notice that for any $N > 0$, the stopping time

$$\tau_N(\omega) = \inf\left\{\tau : \int_0^\tau \|F(X_s(\omega), s)\|_\beta^{2q} ds > N^q\right\}.$$

is well defined, almost surely finite, and

$$\lim_{N\to\infty} \min\{\tau_N, t\} = t \qquad \text{a.s.} \tag{6.9}$$

Now, consider the truncated $Y_s$ process, $Y_s^N$, which is the mild solution to the stochastic partial differential equation

$$dY_s^N = -LY_s^N ds + G_s^N ds + Q dW_s\,, \tag{6.10}$$

where $G_s^N(\omega)$ is defined to be

$$G_s^N(\omega) = \begin{cases} 0 & s < t/2 \\ 2S(t-s)F(X_{2s-t}(\omega), 2s-t)\mathbf{1}_{\{\frac{\tau_N+t}{2}>s\}} & t/2 \le s \le t \end{cases}.$$

We will first prove that the law of $Y^N$ on path space is equivalent to the law of $Z$ on path space. We will then remove the truncation of $Y^N$ to prove the desired result.

Observe that $G_s^N$ is still adapted to the filtration generated by $\{W_r : r \le s\}$ and that there is no concern about the existence of the process $Y_s^N$ since, as with $Y_s$, it can be written as an explicit adapted functional of the $X$ and $W$ processes.

Next, notice that

$$\begin{aligned}\int_0^t \left\|G_s^N(\omega)\right\|_{\gamma+\delta}^2 ds &= \int_{t/2}^t \left\|S(t-s)F(X_{2s-t}(\omega), 2s-t)\mathbf{1}_{\{\frac{\tau_N+t}{2}>s\}}\right\|_{\gamma+\delta}^2 ds \\ &< \int_{t/2}^t \frac{1}{(t-s)^{\gamma+\delta-\beta}} \|F(X_{2s-t}, 2s-t)\|_\beta^2 \mathbf{1}_{\{\frac{\tau_N+t}{2}>s\}} ds \\ &\le N\Big(\int_{t/2}^t \frac{1}{(t-s)^{(\gamma+\delta-\beta)r}} ds\Big)^{\frac{1}{r}} \le NK \qquad \text{a.s}\end{aligned} \tag{6.11}$$

where $K$ is some finite constant since by construction $r(\gamma + \delta - \beta) < 1$. By Lemma 3.2.2, we see that this bound implies for some constant $C > 0$

$$\int_0^t |Q^\dagger G_s^N(\omega)|_{\mathcal{U}}^2 ds \le CKN \quad \text{a.s.}\,. \tag{6.12}$$

Based on Remark 6.2.2, this estimate implies that the stochastic process

$$\mathcal{E}_t^N(\omega) = \exp\left(-\int_0^t \left\langle Q^\dagger G_s^N(\omega), dW_s(\omega)\right\rangle_{\mathcal{U}} - \frac{1}{2}\int_0^t |Q^\dagger G_s^N(\omega)|_{\mathcal{U}}^2 ds\right) \tag{6.13}$$

is an $L^2$-martingale with respect to the filtration $\mathcal{F}$ and the measure $\mathbb{P}$. If we define a new probability measure

$$d\widetilde{P}^N(\omega) = \mathcal{E}_t^N(\omega)d\mathbb{P}(\omega), \tag{6.14}$$

and let

$$H_t^N(\omega) = \int_0^t Q^\dagger G_s^N(\omega)ds\,,$$

then by Girsanov's theorem, the process $\widetilde{W}_t^N(\omega) = W_t(\omega) + H_t^N(\omega)$ is a cylindrical Wiener process with respect to the measure $\widetilde{P}^N$. Now, observe that we can rewrite the $Y_s^N(\omega)$ process given (6.10) as

$$dY_s^N(\omega) = -LY_s^N(\omega)ds + Qd\widetilde{W}_s^N(\omega)\,. \tag{6.15}$$

Since under the measure $\widetilde{P}^N$, $\widetilde{W}^N$ is a cylindrical Brownian motion, we see from (6.15) that the distribution of $Y_t^N$ with respect to $\widetilde{P}^N$ is exactly the same as the distribution of $Z_t$ with respect to $\mathbb{P}$. We will now show that the measures on pathspace induced by these processes are equivalent. That is, for any $A \in \mathcal{B}(C([0,t];H))$, $P_{[0,t]}^{Y^N}(x,A) = 0$ if and only if $P_{[0,t]}^{Z}(x,A) = 0$.

We begin by defining $P_{[0,t]}^Z(x,\,\cdot\,)$ precisely by,

$$P_{[0,t]}^Z(x,A) = \mathbf{P}\Big(\omega : Z([0,t],x)(\omega) \in A\Big)$$

and defining $P^{Y^N}_{[0,t]}(x,\cdot)$ similarly. Now, considering $Y^N_s$ in the form of equation (6.15), we have

$$\begin{aligned} P^Z_{[0,t]}(x,A) &= \mathbb{P}\Big(\omega : Z([0,t],x)(\omega)\in A\Big) = \widetilde{P}^N\Big(\omega : Y^N([0,t],x)(\omega)\in A\Big) \\ &= \mathbb{E}\Big[\mathbf{1}_A\big(Y^N([0,t],x)(\omega)\big)\mathcal{E}^N_t(\omega)\Big] \\ &\le \Big(\mathbb{E}\mathbf{1}_A\big(Y^N([0,t],x)(\omega)\big)\Big)^{\frac{1}{2}}\Big(\mathbb{E}(\mathcal{E}^N_t(\omega))^2\Big)^{\frac{1}{2}} \\ &= \big(P^{Y^N}_{[0,t]}(x,A)\big)^{\frac{1}{2}}\big(\mathbb{E}(\mathcal{E}^N_t(\omega))^2\big)^{\frac{1}{2}} \end{aligned}$$

Thus, if $P^{Y^N}_{[0,t]}(x,A)=0$, then $P^Z_{[0,t]}(x,A)=0$ for any set $A$ in the pathspace. That is, $P^Z_{[0,t]}(x,\cdot)\ll P^{Y^N}_{[0,t]}(x,\cdot)$.

In fact the two measures are equivalent. To see this, we begin by observing that from (6.12) and similar reasoning as in the preceding calculation, we also know that $(\mathcal{E}^N_t)^{-1}$ is an $L^2$-martingale with respect to $\widetilde{P}$. Therefore, we can write

$$\begin{aligned} P^{Y^N}_{[0,t]}(x,A) &= \mathbb{E}\Big[\mathbf{1}_A\big(Y^N([0,t],x)(\omega)\big)\Big] \\ &= \mathbb{E}\Big[\mathbf{1}_A\big(Y^N([0,t],x)(\omega)\big)\mathcal{E}^N_t(\omega)^{\frac{1}{2}}\mathcal{E}^N_t(\omega)^{-\frac{1}{2}}\Big] \\ &\le \Big(\mathbb{E}\Big[\mathbf{1}_A\big(Y^N([0,t],x)(\omega)\big)\mathcal{E}^N_t(\omega)\Big]\Big)^{\frac{1}{2}}\Big(\mathbb{E}\,\mathcal{E}^N_t(\omega)^{-1}\Big)^{\frac{1}{2}} \\ &= \Big(P^Z_{[0,t]}(x,A)\Big)^{\frac{1}{2}}\Big(\widetilde{\mathbb{E}}\,\mathcal{E}^N_t(\omega)^{-2}\Big)^{\frac{1}{2}} \end{aligned}$$

. Because martingales have finite expectation, we see that $P^{Y^N}_{[0,t]}(x,\cdot)\ll P^Z_{[0,t]}(x,\cdot)$, and when combined with the previous estimate we have $P^{Y^N}_{[0,t]}(x,\cdot)\sim P^Z_{[0,t]}(x,\cdot)$. Though not needed for our current goal, this equivalence will be useful later.

To complete the proof, we need to remove the truncation of the $Y^N_s$ process and show that in fact we have, $P^Y_{[0,t]}(x,\cdot)\ll P^Z_{[0,t]}(x,\cdot)$. From (6.9) we see that for any $A$ in the Borel $\sigma$-algebra generated by the pathspace,

$$\lim_{N\to\infty} P^{Y^N}_{[0,t]}(x,A) = P^Y_{[0,t]}(x,A)$$

Now, if $P^Z_{[0,t]}(x, A) = 0$, then $P^{Y^N}_{[0,t]}(x, A) = 0$ since, as proved above, $P^{Y^N}_{[0,t]}(x, \cdot) \sim P^Z_{[0,t]}(x, \cdot)$. Thus, we also have

$$P^Y_{[0,t]}(x, A) = \lim_{N\to\infty} P^{Y^N}_{[0,t]}(x, A) = \lim_{N\to\infty} 0 = 0$$

and we can conclude that $P^Y_{[0,t]}(x, \cdot) \ll P^Z_{[0,t]}(x, \cdot)$. As a consequence of the discussion at the end of Section 6.1, we can also conclude that $\mathcal{P}_{0,t}(x, \cdot) \ll \mathcal{R}_t(x, \cdot)$, and the proof of the proposition is complete. Removing the truncation in the other direction to obtain equivalence requires additional information and will be carried out in the next section.

□

## 6.3 Absolute Continuity in the Other Direction

In this section, we recall that the *relative entropy* of two measures $\mu$ and $\nu$ is defined as

$$H(\mu|\nu) = \int \frac{d\mu}{d\nu} \log(\frac{d\mu}{d\nu}) d\nu \tag{6.16}$$

with $H(\mu|\nu) < \infty$ if and only if $\mu \ll \nu$. To prove $\mathcal{R}_{0,t}(x, \cdot) \ll \mathcal{P}_t(x, \cdot)$, we will use the following lemma whose proof can be found in Mattingly and Suidan (2005), where it was used for essentially the same purpose.

**Lemma 6.3.1.** *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $(W, \mathcal{W})$ be a measure space. Assume $W$ is a Polish space and $\mathcal{W}$ is the Borel sigma algebra. Let $T : \Omega \to W$ and $T_n : \Omega \to W$ $(n = 1, 2, \cdots)$ be measurable transformations. Let $P = T^*\mu$ and $Q_n = T_n^*\mu$ be the push–forward measures on $W$ induced by the respective transformations. Assume that there is a probability measure $Q$ on $W$ such that for any measurable $A \in \mathcal{W}$, $Q_n(A) \to Q(A)$. If $P \sim Q_n$ and $\limsup_{n\to\infty} |H(P|Q_n)| < M < \infty$, then $P \ll Q$.*

We now use this lemma to prove that in our setting we can remove the truncation of the $Y$ process to also get $P^Z_{[0,t]}(x,\cdot) \ll P^Y_{[0,t]}(x,\cdot)$.

**Proposition 6.3.2.** *Let Assumption 3.2.1 and the standing assumptions hold. Fixing a positive $\alpha$, with $\alpha \in (\gamma + a + \delta - 1, \gamma - \rho + 1)$, if the initial condition $x \in \mathcal{H}^\alpha$ is such that*

$$\mathbb{E} \sup_{s\in[0,T]} \|X_s(\omega)\|_\alpha^{2p} < \infty \ ,$$

*then for any $s, t \in [0,T]$ with $s < t$, we have $P^Y_{[s,t]}(x, \cdot\,) \sim P^Z_{[s,t]}(x, \cdot\,)$ as measures on $C([0,t]; H)$ and as a consequence, $\mathcal{R}_{t-s}(x, \cdot\,) \sim \mathcal{P}_{s,t}(x, \cdot\,)$.*

*Proof.* As before with out loss of generality we take $s = 0$. In light of Lemma 6.3.1, we need only to show that

$$\sup_N H(P^Z|P^{Y^N}) < M < \infty. \tag{6.17}$$

Recall the definitions of $Y^N_s$, $\mathcal{E}^N_s$, $\mathbb{P}$, and $\widetilde{P}$ from the proof of Proposition 6.2.3. Since the distribution of $Y^N_t$ with respect to $\widetilde{P}^N$ is the same as the distribution of $Z_t$ with respect to $\mathbb{P}$, we can write

$$H(P^Z|P^{Y^N}) \le H(\mathbb{P}|\widetilde{P}) = \widetilde{\mathbb{E}}\big\{\big(\mathcal{E}^N\big)^{-1} \log\big[\big(\mathcal{E}^N\big)^{-1}\big]\big\} = \mathbb{E}\log\big[\big(\mathcal{E}^N\big)^{-1}\big]. \tag{6.18}$$

As in the proof of Proposition 6.2.3, we now fix a $\beta = \alpha - a$ where $\beta \in (\gamma + \delta -$

$1, \gamma - \rho + 1 - a)$ and observe that

$$\begin{aligned} H(P^Z|P^{Y^N}) &\le \mathbb{E}\left(\log\left(\mathcal{E}_t^N(\omega)^{-1}\right)\right) \\ &= \mathbb{E}\Big(\int_0^t \langle Q^\dagger G_s^N(\omega), dW_s(\omega)\rangle_{\mathcal{U}} + \frac{1}{2}\int_0^t |Q^\dagger G_s^N(\omega)|_{\mathcal{U}}^2 ds\Big) \\ &= \frac{1}{2}\mathbb{E}\int_0^t |Q^\dagger G_s^N(\omega)|_{\mathcal{U}}^2 ds \\ &\le C\mathbb{E}\int_0^t \|S(t-s)F^N(X_{2s-t}(\omega), 2s-t)\|_{\gamma+\delta}^2 ds \\ &\le C\Big(\int_0^t \frac{1}{(t-s)^{\gamma+\delta-\beta}} ds\Big)\Big(1 + \mathbb{E}\sup_{s\in[0,t]} \|X_s(\omega)\|_\alpha^{2p}\Big) \end{aligned}$$

where $C$ is a positive constant whose definition changes from line to line. Now since $\gamma + \delta - \beta < 1$, we know that the first term in the last line is finite. The second term is finite by assumption.

□

## 6.4 Extension to Initial Data in H

In order to complete the proof of Theorem 3.2.3, we need to show that we can extend the equivalence of the time $t$ transition measures for the process $X_t$ to all initial points in $H$.

**Corollary 6.4.1.** *Under the hypotheses of Theorem 3.2.3,*

$$\mathcal{R}_{t-s}(x,\,\cdot\,) \sim \mathcal{P}_{s,t}(x,\,\cdot\,) \sim \mathcal{P}_{s,t}(y,\,\cdot\,) \sim \mathcal{R}_{t-s}(y,\,\cdot\,) \tag{6.19}$$

*for any* $x, y \in H$.

*Proof.* We will begin by showing that $\mathcal{P}_{s,t}(x,\,\cdot\,) \sim \mathcal{P}_{s,t}(y,\,\cdot\,)$ for any $x, y \in H$. By the semigroup property of the transition densities and the fact that $X$ has a continuous path, as claimed in Proposition 2.5.2, for any measurable set $A \subset H$, we can write

$$\mathcal{P}_{s,t}(x, A) = \int_{\mathcal{H}^\alpha} \mathcal{P}_{r,t}(y, A)\mathcal{P}_{s,r}(x, dy) \tag{6.20}$$

for any $r \in (s, t)$ and any $\alpha \in (\gamma+a+\delta-1, \gamma-\rho+1)$. Therefore, if $\mathcal{P}_{s,t}(x, A) = 0$ then $\mathcal{P}_{r,t}(y, A) = 0$ for $\mathcal{P}_{s,r}(x, \cdot)$-a.e. $y$. Now, we fix $\alpha$ as in Theorem 3.2.3. Let $y \in \mathcal{H}^\alpha$ be such that $\mathcal{P}_{r,t}(y, A) = 0$. Then, by Proposition 6.3.2, $\mathcal{P}_{r,t}(y, A) \sim \mathcal{R}_{t-r}(y, A)$ and by the equivalence of $\mathcal{R}_t$ as stated in Lemma 5.3.1, we conclude that $\mathcal{P}_{r,t}(y, A) = 0$ for all $y \in \mathcal{H}^\alpha$. Because of the fact that for any $z \in H$, $\mathcal{P}_{s,t}(z, A)$ has a representation analogous to (6.20),

$$\mathcal{P}_{s,t}(z, A) = \int_{\mathcal{H}^\alpha} \mathcal{P}_{r,t}(y, A)\mathcal{P}_{s,r}(z, dy)\,,$$

we see that $\mathcal{P}_{s,t}(x, A) = 0 \sim \mathcal{P}_{s,t}(z, A)$ for all $x, z \in H$.

Now, combining the preceding paragraph with Proposition 6.3.2 yields $\mathcal{P}_{s,t}(x, \cdot) \sim \mathcal{P}_{s,t}(y, \cdot) \sim \mathcal{R}_{t-s}(y, \cdot)$ for any $x \in H$ and $y \in \mathcal{H}^\alpha$. Lemma 5.3.1 tells us that $\mathcal{R}_{t-s}(y, \cdot) \sim \mathcal{R}_{t-s}(z, \cdot)$ for any $z \in H$, in particular, $\mathcal{R}_{t-s}(y, \cdot) \sim \mathcal{R}_{t-s}(x, \cdot)$. Therefore, we have

$$\mathcal{P}_{s,t}(x, \cdot) \sim \mathcal{P}_{s,t}(y, \cdot) \sim \mathcal{R}_{t-s}(y, \cdot) \sim \mathcal{R}_{t-s}(x, \cdot) \tag{6.21}$$

and the corollary is proved. Combining this corollary with Lemma 5.3.1 gives us the result in Theorem 3.2.3, and we have proven our main result. □

## 6.5 Consequences of the Main Result

When proving ergodicity results, one may be interested in the existence of a unique invariant measure for the process. The equivalence result of Theorem 3.2.3 allows us to prove very easily, that is, without proving properties like strong Feller or irreducibility, that there is at most one stationary measure on $H$ for the time homogeneous version of 2.1. A stationary probability measure is probability measure $\mu$ on $H$ so that for all measurable $A \subset H$,

$$\int_H \mathcal{P}_t(x, A)\mu(dx) = \mu(A)$$

or in other words $\mathcal{P}_t^* \mu = \mu$. The following facts are also true:

1. Any stationary measure can be written as a convex combination of ergodic stationary measures

2. All distinct ergodic measures are mutually singular

**Corollary 6.5.1.** *In the setting of Theorem 3.2.3, if $F(x, s)$ is independent of $s$ there exists at most one stationary probability measure $\mu$ with $\mu(H) = 1$. Furthermore, if an invariant measure $\mu$ exists, $\mu \sim \nu$.*

*Proof.* Since any stationary measure can be written as a convex combination of ergodic stationary measures, it is sufficient to show that there is at most one ergodic stationary measure. Assume that there exist two such measures $\mu_1$ and $\mu_2$. For $i = 1, 2$ and any measurable $A \subset H$ we have

$$\mu_i(A) = \int_H \mathcal{P}_t(x, A)\mu_i(dx)\,.$$

Hence, if $\mu_1(A) = 0$, then $\mathcal{P}_t(x, A) = 0$ for $\mu_1$-a.e $x$. But since Theorem 3.2.3 implies $\mathcal{P}_t(x, \cdot\,) \sim \mathcal{P}_t(y, \cdot\,)$, we conclude that $\mathcal{P}_t(y, A) = 0$ for all $y \in H$ which in turn implies that $\mu_2(A) = 0$. This further implies that $\mu_1 \sim \mu_2$ and by our remark on ergodic measures $\mu_1 = \mu_2$.

The preceding argument showed that $\mu \stackrel{\text{def}}{=} \mu_1 = \mu_2 \gg \mathcal{P}_t(x, \cdot\,)$ for all $t > 0$ and $x \in H$. Now assume that $\mathcal{P}_t(x, A) = 0$ for some $x \in H$ and $t > 0$. Then we know that $\mathcal{P}_t(x, A) = 0$ for all $x \in H$ since $\mathcal{P}_t(x, A) \sim \mathcal{P}_t(y, A)$ for all $x, y \in H$. By the representation of $\mu$ from the preceding paragraph, this implies that $\mu(A) = 0$. Hence $\mathcal{P}_t(x, \cdot) \sim \mu$ for any $x \in H$ and $t > 0$. Lastly, $\mu(H) = 1$ follows from the regularity result given in Proposition 2.5.2. The proof is completed since $\mu \sim \mathcal{P}_t(x, \cdot) \sim \mathcal{R}_s(y, \cdot) \sim \nu$ for any $t, s > 0$ and $x, y \in H$, the fact that any invariant measure can be decomposed into ergodic invariant measures, and the fact that any two distinct ergodic invariant measure must be singular. $\square$

# 7

# A Heavy-Traffic Queueing Network

## 7.1 Introduction

We consider a network of two multi-class queueing stations, where first station in the network is a polling station. A polling station consists of a single server that visits (or polls) the queues at the station in a cyclic fashion and processes only the jobs at the queue it is currently polling. Usually, the server experiences a setup delay while it is switching from one queue to the next. The *service discipline* for a polling station dictates how long polling of each queue lasts and in what order queues are visited.

Polling stations are often used to model phenomena in computer science, telecommunications, and service. Any example in which a person or thing has to switch tasks and re-focus their efforts on each task seems to be able to fit the model of a polling station. In every example the goal is to complete all tasks, but the conflict occurs in trying to prevent an accumulation of any particular job class. The focus of this work assumes a heavy-traffic setting, one in which the system is critically loaded.

This work is an extension of recent work done by Otis Jennings. In Jennings

(2008), the author seems to be the first to prove a heavy traffic limit theorem (HTLT) for a multi-class queueing network containing a polling station. In this paper, the system had only 2 classes of jobs, and in the case of the service disciplines the author considered, there were always moments in time when only one type of job class was present at the first station. We consider a system with $K \geq 2$ classes, and prove a similar HTLT. That is, we prove the convergence of the total workload processes to a two-dimensional reflected Brownian motion. We also extend the results of Jennings (2010), in which the author proves a heavy traffic averaging principle for the diffusion scaled queue length process of a singe polling station with $K$ job classes. In this work, our system is a network of two stations, and the fluctuating processes are the station 2 upstream workload process as well as the station 2 immediate workload process. These fluctuations are due to the assumed difference in speed between the servers at station 1 and 2 for some classes of jobs. We are able to use techniques similar to those of Jennings (2010) to prove a heavy traffic averaging principle for the workload process that tracks the work residing at station 2. We note that throughout this chapter, we will let $\|\boldsymbol{v}\|$ denote the $\ell_1$ norm of some vector $\boldsymbol{v}$, with $\|\boldsymbol{v}\| = \sum_k |v_k|$.

## 7.2 The Model

The model that we consider is a network of two single-server queueing stations. The server at the first station operates under a fractional exhaustive service (FES) polling policy. There are $K$ job classes indexed by the numbers $1, 2, \cdots, K$, and each station has $K$ queues designated for each class of jobs. Class $k$ jobs enter the network at station 1 according to their own distinct exogenous arrival process, and await processing by the server. The server processes jobs in queue $k$ (which consists only of class $k$ jobs) until the queue length falls below a pre-specified fraction of what the queue length was at the time when the server first began processing jobs there. At this point, the server moves on to process a portion of the class $k+1$ jobs in queue

$k+1$ and continues to poll the queues at station 1 in a cyclic fashion. Jobs that have been completely serviced at station 1 immediately leave the station and enter their respective queue at station 2 to be processed by the station 2 server. Within each class, jobs are processed on a first-come-first served basis. Both the station 1 and the station 2 servers are *nonidling*, i.e., the server works continuously as long as there are jobs present at the station. In addition, the servers experience no setup delay.

We consider a sequence of networks, indexed by $n \geq 1$, that operates in this fashion. The dynamics for each network in the sequence is governed by its own set of parameters. For the $n^{th}$ network in the sequence, we let $\{\xi_k^n(i) : i \geq 1\}$ denote the set of interarrival times for class $k$ jobs. If we exclude the arrival time for the first class $k$ job, then the set $\{\xi_k^n(i) : i \geq 2\}$ is i.i.d. with mean $1/\lambda_k^n > 0$ and standard deviation $\sigma_{a,k}^n > 0$. We assume that $\xi_k^n(1)$ has finite first and second moments and is independent of the other interarrival times. We call the quantity $\lambda_k^n$, the class $k$ arrival rate. At time 0, there are some jobs already present in the system. For $j = 1, 2$, let $\{\hat{\eta}_{j,k}^n(i) : i \geq 1\}$ denote the set of service times of class $k$ jobs present at station $j$ before time 0. These service times are i.i.d. except for perhaps the first ones, $\hat{\eta}_{j,k}^n(1)$, which are independent of the others and has finite first and second moments. Let $\{\eta_{j,k}^n(i) : i \geq 1\}$ denote the i.i.d. set of service times of class $k$ jobs at station $j$ that arrived after time 0. The means and standard deviations of the service times for all jobs are positive and are denoted by $1/\mu_{j,k}^n$ and $\sigma_{s,j,k}^n$, respectively. We call the quantity $\mu_{j,k}^n$ the service rate for class $k$ jobs at station $j$. We assume we have the following convergences as $n \to \infty$ :

$$\lambda_k^n \to \lambda_k, \qquad \mu_{j,k}^n \to \mu_{j.k}, \qquad j = 1, 2, \quad k = 1, 2, \ldots, K \tag{7.1}$$

and

$$\sigma_{a,k}^n \to \sigma_{a,k}, \qquad \sigma_{s,j,k}^n \to \sigma_{s,j,k} \qquad j = 1, 2, \quad k = 1, 2, \ldots, K \tag{7.2}$$

Let $\rho^n_{j,k} \equiv \lambda^n_k / \mu^n_{j,k}$ be the fractional contribution of class $k$ jobs to the total arrival of work for station $j$, which is given by $\sum_k \rho^n_{j,k}$. We have the following *heavy traffic condition*:

$$c^n_j \equiv \sqrt{n}\left(\sum_k \rho^n_{j,k} - 1\right) \to c_j, \quad \text{as } n \to \infty \tag{7.3}$$

for $j = 1, 2$, where the limit point $c_j$ is finite. As an immediate consequence of this condition, we have that the *nominal server utilization* at station $j$, $\sum_k \rho^n_{j,k}$, approaches 1 asymptotically. Furthermore, the vector of contributions of the job classes to the servers' utilization also converges to a probability vector of strictly positive entries:

$$(\rho^n_{j,1}, \rho^n_{j,2}, \cdots, \rho^n_{j,K}) \to (\rho_{j,1}, \rho_{j,2}, \cdots \rho_{j,K}) = \boldsymbol{\rho_j}\,. \tag{7.4}$$

We also assume that the Lindeberg conditions hold. That is, for $j = 1, 2$, $i \geq 1$, $k = 1, \cdots, K$, and any $\epsilon > 0$,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\left[(\xi^n_k(i))^2 \cdot \mathbf{1}_{\{\xi^n_k(i) > \epsilon\sqrt{n}\}}\right] &= \lim_{n\to\infty} \mathbb{E}\left[(\eta^n_{j,k}(i))^2 \cdot \mathbf{1}_{\{\eta^n_{j,k}(i) > \epsilon\sqrt{n}\}}\right] \\ &= \lim_{n\to\infty} \mathbb{E}\left[(\hat{\eta}^n_{j,k}(i))^2 \cdot \mathbf{1}_{\{\hat{\eta}^n_{j,k}(i) > \epsilon\sqrt{n}\}}\right] = 0,\end{aligned} \tag{7.5}$$

where $\mathbf{1}_A$ is the indicator function that equals 1 if $A$ is true and is 0 otherwise. We include the technical assumption in (7.5) to justify Lemma 7.5.1

We now define $K \times K$ diagonal matrices for each station that contain the limiting service times for each class of jobs, $M_j = diag(1/\mu_{j,1}, 1/\mu_{j,2}, \cdots, 1/\mu_{j,K})$. We will assume that the servers at the two stations do not have identical limiting service rates: $M_1 \neq M_2$. Furthermore, due to the heavy traffic condition, we can assume that some classes of jobs *decelerate* as they move through the network. That is, there exists some subset $\mathcal{D} \subset \{1, 2, \cdots, K\}$ such that if $k \in \mathcal{D}$, then $\mu_{2,k} \leq \mu_{1,k}$. Naturally, by (7.3), some classes of jobs must also accelerate. Throughout this paper we will be interested in the ratios of service times for class $k$ jobs at station 1 and station 2.

To this end, let $\boldsymbol{\theta} = \mathbf{1} M_2 M_1^{-1}$ where $\mathbf{1}$ is the $K$-dimensional vector of 1's. The $k^{th}$ element of $\boldsymbol{\theta}$ is $\theta_k = \mu_{1,k}/\mu_{2,k}$, and we define $\theta^* \equiv \max_k \theta_k$. The vector $\boldsymbol{\theta}$ allows us to view work at station 1 from a station 2 perspective. More precisely, $\boldsymbol{\theta}$ is a vector of the limiting station 2 service rates for jobs at station 1.

We use a variety of stochastic and deterministic processes to describe the dynamics of our system. For the $n^{th}$ sequence in the system, the queue length, denoted by $Q^n_{j,k}(t)$, tracks the number of class $k$ jobs there are at station $j$ at time $t$, including the job in service. It is also necessary to track the *remaining* number of class $k$ jobs at station $j$ that were in the system at time 0; this process is denoted by $\hat{Q}^n_{j,k}(t)$. The workload process for station $j$, denoted by $W^n_j(t)$, is a real valued function of time that tracks how much effort is needed for the server at station $j$ to process the jobs in the system at time $t$ that it has not yet processed. For station 1, $W^n_1(t)$ simply accounts for the jobs residing in station 1 at time $t$. However, for station 2, $W^n_2(t)$ must account for both the jobs at station 1 that have not yet arrived to station 2 as well as the jobs presently residing at station 2 at time $t$. Suppose we also had a set of probability vectors $\{\boldsymbol{\pi}^{(1)}, \cdots, \boldsymbol{\pi}^{(K)}\}$, where in the limiting case, $\boldsymbol{\pi}^{(k)}$ indicates how the workload is distributed among job classes at station 1 at the moment immediately before the server starts to process class $k$ jobs. These vectors will be defined in more detail in Section 7.4. However, now we can define the parameter

$$\gamma_k = \langle \boldsymbol{\pi}^{(k)} M_1^{-1} M_2, \mathbf{1} \rangle = \langle \boldsymbol{\theta}, \boldsymbol{\pi}^{(k)} \rangle, \tag{7.6}$$

which is the set of transformations that convert station 1 workload to its contribution to station 2's total workload at the beginning of production runs. Without loss of generality, we choose $\gamma_1 \stackrel{\text{def}}{=} \gamma = \max_k \gamma_k$. If we let $U^n_2(t)$ denote the *upstream workload* process, the time station 2 must dedicate to processing jobs currently residing at station 1, then this numbering convention for $\gamma$ guarantees that once the system is in steady state, the upstream workload for station 2 will always be greatest at the beginning of a cycle.

The idle time process, $I_j^n(t)$, tracks the cumulative amount of time that the server at station $j$ has spent idle by time $t$. Due to the nonidling property of our system, the servers only idle when there are no jobs present at their corresponding stations.

The heavy traffic condition given in (7.3) implies that the workload process at time $nt$ is of order $\sqrt{n}$. Therefore, we use a natural diffusion scaling for many of our processes. This scaling accelerates time at rate $n$ and rescales space by a factor of $\sqrt{n}$. Thus, we define the diffusion scaled processes $\tilde{Q}_{j,k}^n = (\tilde{Q}_{j,k}^n(t), t \geq 0)$, $\tilde{W}_j^n = (\tilde{W}_j^n(t), t \geq 0)$, and $\tilde{I}_j^n = (\tilde{I}_j^n(t), t \geq 0)$, where for each $t \geq 0$,

$$\tilde{Q}_{j,k}^n(t) = \frac{Q_{j,k}^n(nt)}{\sqrt{n}}, \quad \tilde{W}_j^n(t) = \frac{W_j^n(nt)}{\sqrt{n}}, \text{and} \quad \tilde{I}_j^n(t) = \frac{I_j^n(nt)}{\sqrt{n}}. \tag{7.7}$$

## 7.3 Main Results

We will prove two results. The first is similar to results that can be found among the literature on heavy-traffic limit theorems for feed-forward queueing networks. We show that the diffusion scaled total workload process converges to a two-dimensional Brownian motion in a wedge. Heuristically, this is a strong Markov process that satisfies the following three properties:

1. The state space is an infinite two-dimensional wedge, and the process behaves in the interior of the wedge like an ordinary Brownian motion.

2. The process reflects instantaneously at the boundary of the wedge, with the direction of reflection being either vertical or horizontal along each side.

3. The amount of time that the process spends at the corner of the wedge is zero (i.e., the set of times for which the process is at the corner has Lebesgue measure zero).

For a more mathematically precise definition, see Varadhan and Williams (1985).

Now, we define a two-dimensional Brownian motion $\tilde{B} = (\tilde{B}_1, \tilde{B}_2)$ with zero drift, covariance matrix

$$\Gamma \equiv \begin{pmatrix} \sum_{k=1}^{K} \lambda_k(\sigma_{s,1,k}^2 + \rho_{1,k}^2 \sigma_{a,k}^2) & \sum_{k=1}^{K} \lambda_k \rho_{1,k}\rho_{2,k}\sigma_{a,k}^2 \\ \sum_{k=1}^{K} \lambda_k \rho_{1,k}\rho_{2,k}\sigma_{a,k}^2 & \sum_{k=1}^{K} \lambda_k(\sigma_{s,2,k}^2 + \rho_{2,k}^2 \sigma_{a,k}^2) \end{pmatrix}, \tag{7.8}$$

and $\tilde{B}(0) = 0$. Next, we define the process $\tilde{X} \equiv (\tilde{X}_1, \tilde{X}_2)$ expressed as

$$\tilde{X}_j = \tilde{W}_j(0) + \tilde{B}_j + c_j e\,, \quad j = 1, 2, \tag{7.9}$$

where $c_j$ is the limit point given in (7.3) and $e = (e(t), t \geq 0)$ is a process such that $e(t) = t$ for all $t \geq 0$. This process $\tilde{X}$ is a Brownian motion with drift $c = (c_1, c_2)$ and initial condition $\tilde{W}(0) = (\tilde{W}_1(0), \tilde{W}_2(0))$, which is independent of $\tilde{B}$.

Lastly, before we give our main results, we define the reflection map $\Phi$. Given a function $x$, with $x(0) \geq 0$, $\Phi(x)(t) = -\inf_{s \leq t}(x(s) \wedge 0)$. The function $\Phi(x)$ is nondecreasing and provides the reflection of $x$ at 0.

**Theorem 7.3.1.** *Assume that*

$$\left(\tilde{W}_1^n(0), \tilde{W}_2^n(0)\right) \xrightarrow{P} \left(\tilde{W}_1(0), \tilde{W}_2(0)\right) \tag{7.10}$$

*as $n \to \infty$ with the constraint that*

$$\tilde{W}_2(0) \geq \theta^* \tilde{W}_1(0) \geq 0. \tag{7.11}$$

*Under* (7.1)-(7.3), *we have*

$$\tilde{W}^n \Rightarrow \tilde{W} \equiv \tilde{X} + \tilde{I}, \tag{7.12}$$

*where $\tilde{I}_1 \equiv \Phi(\tilde{X}_1)$ and $\tilde{I}_2 \equiv \Phi(\tilde{X}_2 - \gamma \tilde{W}_1)$.*

The next result is an averaging principle for the station 2 immediate workload process. First, we define what is essentially a projection of a $K$-dimensional simplex to 1 dimension. Let $\ell^{(k)}(u) : [0,1] \to [0,\infty)$, be a continuous function with

$$\ell^{(k)}(u) = (1-u)\gamma_k + u\gamma_{k+1}. \tag{7.13}$$

We have the following result.

**Theorem 7.3.2.** *For any continuous function $f : [0,\infty) \to (-\infty,\infty)$ and finite $T > 0$, we have*

$$\int_0^T f(\tilde{V}_2^n(s))ds \Longrightarrow \int_0^T \sum_{k=1}^{K} \rho_{1,k} \int_0^1 f(\tilde{W}_2(s) - \ell^{(k)}(u)\tilde{W}_1(s))duds, \tag{7.14}$$

*where $\tilde{V}_2^n(t)$ denotes the workload process for the jobs currently residing at station 2 at time $t$. We refer to this process as the station 2 diffusion scaled immediate workload process, and $(\tilde{W}_1(s), \tilde{W}_2(s)) \equiv \tilde{W}(s)$ is the limiting process given in* (7.12).

We will prove these theorems in Sections 7.7.1 and 7.7.2, but first we must describe our network in further detail.

## 7.4 Dynamics of the polling policy

As already mentioned, the setting of this paper is a network of two queueing stations, in which the first station is a polling station operating under fractional exhaustive service, FES. Under this polling policy, a *production run* is the length of time it takes for the server to process enough jobs in the queue at station 1 so that the remaining number of jobs in the queue is less than or equal to a pre-specified fraction of the number of jobs present when the production run began. More specifically, our system is governed by FES($\alpha$), where the $k^{th}$ element of the parameter $\alpha \in (0,1]^K$ indicates the fraction to which class $k$ jobs are exhausted during a production run. Without loss of generality, we define a *cycle* to be the interval of time beginning at the start of a class 1 production run and ending at the start of the next class 1 production run. We can describe how the server switches from processing one class of jobs to the next class of jobs in terms of stopping times. We allow the server to begin processing jobs in any of the queues at time 0. Thus, the first cycle is a partial cycle; we refer to it as cycle 0. For the $n^{th}$ system in the sequence, let $\tau^n_{i,1,k}$ be the starting time of the $k^{th}$ production run at station 1 during the $i^{th}$ cycle. We let $\nu^n_{i,1,k}$ denote the time of the end of a class $k$ production run at station 1 during the $i^{th}$ cycle. Under FES($\alpha$), we can write this explicitly as

$$\nu^n_{i,1,k} = \inf\{t > \tau^n_{i,1,k} : Q^n_{1,k}(t) \leq (1-\alpha_k)Q^n_{1,k}(\tau^n_{i,1,k})\} \tag{7.15}$$

Notice that if the server idles, then it is possible for $\nu^n_{i,1,k} < \tau^n_{i,1,k+1}$. The gap between these times, however, is negligibly small. To be explicit, we restate Proposition 4.1 of Jennings (2010):

**Proposition 7.4.1.** *For any $\epsilon > 0$, there exists an $N$ such that for $n \geq N$,*

$$\mathbb{P}\left(\sup_{i\in\Gamma^n(T)} |\tau^n_{i,1,k+1} - \nu^n_{i,1,k}| > \epsilon\sqrt{n}\right) < \epsilon \tag{7.16}$$

*where for any $T \geq 0$, $\Gamma^n(T) \equiv \{i \geq 0 : \tau^n_{i,1,1} \leq nT\}$ is the set of indices for all cycles that start by time $nT$ at station 1.*

Recall that the server processes jobs at station 1 in a cyclic fashion. That is, for any cycle $i \geq 0$, once the server finishes processing class $k$ jobs, it moves on to process class $(k\%K)+1$ jobs, if any are present. If there are no class $(k\%K)+1$ jobs present, the server continues to cycle through the queues until it finds a non-empty queue. If the first non empty queue is a class $\ell$ queue, where $\ell \neq ((k\%K)+1)$, then there are two cases. If $\ell > k+1$, then for each $m \in \{k+1, \cdots, \ell-1\}$ set

$\tau^n_{i,1,m} = \tau^n_{i,1,\ell}$ and let the class $m$ production run have length 0. The second possibility is $\ell \le k$. In this case, the class $\ell$ production run starts at time $\tau^n_{i+1,1,\ell}$ and for each $m \in \{k+1, \cdots K\}$ we set $\tau^n_{i,1,m} = \tau^n_{i+1,1,\ell}$. In the same way, for each $m \in \{1, \cdots, \ell-1\}$ we set $\tau^n_{i+1,1,m} = \tau_{i+1,1,\ell}$. Again, we let the production runs for the class $m$ jobs have length 0. Notice that for $i \ge 0$, a class $k$ production run of length 0 implies that $\nu^n_{i,1,k} = \tau^n_{i,1,k}$.

### *7.4.1 Properties of the Cycles*

Throughout this paper, $T$ is fixed as the endpoint of the time interval on which we consider this system. This time $T$ is implied in the notation $\Gamma^n$, which will hereafter denote the set of all cycle indices at station 1. It will also be useful to define a time change function $T^n(t)$ that tracks the beginning of service cycles at station 1:

$$T^n(t) \equiv \sup_{i \ge 0} \{\tau^n_{i,1,1} : \tau^n_{i,1,1} \le t\}. \tag{7.17}$$

We also define the fluid-scaled version of the function

$$\bar{T}^n(t) \equiv \frac{T^n(nt)}{n}, \quad t \ge 0. \tag{7.18}$$

As we noted earlier, we expect the workload to be of order $\sqrt{n}$. Thus, it is reasonable to expect the length of the cycles to be bounded and also of order $\sqrt{n}$. The following statement is proven in Proposition 4.2 of Jennings (2010):

**Proposition 7.4.2.** *For every $\epsilon > 0$, there exists some constant $C_\epsilon > 0$ such that for sufficiently large $n$,*

$$\mathbb{P}\left(\sup_{i \in \Gamma^n_1} (\tau^n_{i+1,1,1} - \tau^n_{i,1,1}) > C_\epsilon \sqrt{n}\right) < \epsilon\,. \tag{7.19}$$

### *7.4.2 Tracking the workload at station 1*

Due to our convergence result and our heavy-traffic condition, we expect our system to settle into an equilibrium. That is, the workload vector at station one will eventually achieve some cyclic steady state. In Jennings (2010), the author proves that the station 1 workload approaches this steady state geometrically and in less than order $n$ time. The consequences of this fact will be useful in proving our main results.

We use the set of stochastic matrices, $\left\{P^{(k)}\right\}$ to capture how the workload transitions during production runs and cycles. For $\alpha \in (0,1]^K$ and each $k = 1, \ldots, K$, $P^{(k)}$ denotes the class $k$ production run transition matrix where

$$P^{(k)}_{i,j} = \begin{cases} 1 & \text{if } i = j \neq k \\ 0 & \text{if } j \neq i \neq k \\ 1 - \alpha_k & \text{if } i = j = k \\ \alpha_k \rho_{1,j}/(1-\rho_{1,k}) & \text{if } i = k \neq j \end{cases}. \tag{7.20}$$

The matrix $P^{(k)}$ indicates how during a class $k$ production run, the class $k$ workload decreases while the other class workload increases. The matrix

$$P = \prod_k P^{(k)} \tag{7.21}$$

is defined as the cycle transition matrix, and it admits a left eigenvector $\pi$ whose corresponding eigenvalue is 1. That is,

$$\boldsymbol{\pi} P = \boldsymbol{\pi}. \tag{7.22}$$

Now, we can define the lifting operators $\boldsymbol{\pi^{(k)}}$, which describe how the work is distributed at station 1 at times $\tau_{i,1,k}$ once the system has reached steady state. For each $k = 1, \ldots, K$,

$$\boldsymbol{\pi^{(k)}} = \boldsymbol{\pi} \prod_{j=1}^{k-1} P^{(j)}, \tag{7.23}$$

which gives us the relationship

$$\boldsymbol{\pi}^{(\boldsymbol{k+1})} = \boldsymbol{\pi}^{(\boldsymbol{k})} P^{(k)} \tag{7.24}$$

## 7.5 Preliminaries

### *7.5.1 Primitive processes*

We begin by describing some basic processes that we use to explicitly define the workload processes that we are interested in. First, we let $A_k^n = (A_k^n(t), t \geq 0)$ be the class $k$ arrival process for the $n^{th}$ network in the sequence that tracks the number of class $k$ jobs that have arrived to the system by time $t$, $t \geq 0$.

$$A_k^n(t) = \sup\left\{ j \geq 0 : \sum_{i=1}^{j} \xi_k^n(i) \leq t \right\} \tag{7.25}$$

Next we define partial sum processes $S_{j,k}^n = (S_{j,k}^n(t), t \geq 0)$ and $\hat{S}_{j,k}^n = (\hat{S}_{j,k}^n(t), t \geq 0)$ that sum the service times of the first $\lfloor t \rfloor$ class $k$ jobs that arrive to station $j$ after time 0 and those that are present at station $j$ before time 0, respectively. That is, for $t \geq 0$ we have

$$S_{j,k}^n(t) = \sum_{i=1}^{\lfloor t \rfloor} \eta_{j,k}^n(i) \tag{7.26}$$

and

$$\hat{S}_{j,k}^n(t) = \sum_{i=1}^{\lfloor t \rfloor} \hat{\eta}_{j,k}^n(i) \tag{7.27}$$

We have the following diffusion scaled versions of the arrival process, $\tilde{A}_k^n = (\tilde{A}_k^n(t), t \geq 0)$, and the service time process, $\tilde{S}_{j,k}^n = (\tilde{S}_{j,k}^n(t), t \geq 0)$, where

$$\tilde{A}_k^n(t) \equiv \frac{A_k^n(nt) - \lambda_k^n nt}{\sqrt{n}} \quad \text{and} \quad \tilde{S}_{j,k}^n(t) \equiv \frac{S_{j,k}^n(nt) - (1/\mu_{j,k}^n)nt}{\sqrt{n}} \tag{7.28}$$

for each $j = 1, 2$ and $k = 1, 2, \cdots, K$. We also define the fluid-scaled arrival process $\bar{A}_k^n = (\bar{A}_k^n(t), t \geq 0)$, where

$$\bar{A}_k^n(t) \equiv \frac{A_k^n(nt)}{n}. \tag{7.29}$$

The primitive processes described above have the following limiting properties.

**Lemma 7.5.1.** *Assuming* (7.1)-(7.5)*, we have*

*(i)* $\tilde{A}_k^n \Rightarrow \tilde{A}_k \equiv \sigma_{a,k}(\lambda_k)^{3/2}\beta_{a,k}, \quad k = 1, 2, \ldots, K,$

*(ii)* $\bar{A}_k^n(t) \xrightarrow{P} \lambda_k e, \quad k = 1, 2, \ldots, K,$

*(iii)* $\tilde{S}_{j,k}^n \Rightarrow \tilde{S}_{j,k} \equiv \sigma_{s,j,k}\beta_{s,j,k}, \quad j = 1, 2,\ k = 1, 2, \ldots, K,$

*(iv)* $\tilde{S}_{j,k}^n \circ \bar{A}_k^n(t) \Rightarrow \sigma_{s,j,k}\sqrt{\lambda_k}\beta_{s,j,k}, \quad j = 1, 2,\ k = 1, 2, \ldots, K,$

*(v)* $(1/\sqrt{n}) \max_{i \leq bn} \xi_k^n(i) \xrightarrow{P} 0, \quad k = 1, 2, \ldots, K,\ b \geq 0,$

*(vi)* $(1/\sqrt{n}) \max_{i \leq bn} \eta_{j,k}^n(i) \xrightarrow{P} 0, \quad j = 1, 2,\ k = 1, 2, \ldots, K,\ b \geq 0,$

*(vii)* $(1/\sqrt{n}) \max_{i \leq bn} \hat{\eta}_{j,k}^n(i) \xrightarrow{P} 0, \quad j = 1, 2,\ k = 1, 2, \ldots, K,\ b \geq 0,$

*as* $n \to \infty$*, where for each* $j = 1, 2$ *and* $k = 1, 2, \ldots, K$*,* $\beta_{a,k}$ *and* $\beta_{s,j,k}$ *are independent standard Brownian motions.*

*Proof.* This lemma is proved in Jennings (2008) (Lemma 3.1), where the author uses standard results found in Billingsley (1999) and Coffman et al. (1998) □

*7.5.2 The Workload Processes*

For each network in our sequence, we can track the total workload process $W^n = (W_1^n, W_2^n)$. $W_j^n = (W_j^n(t), t \geq 0)$ is a function of time that measures the amount of effort needed for station $j$ to process the jobs currently in the system that it has yet to fully process. In addition, we define the $K$-dimensional workload processes $\boldsymbol{Z_1^n} = (Z_{1,1}^n, Z_{1,2}^n, \cdots, Z_{1,K}^n)$ and $\boldsymbol{Z_2^n} = (Z_{2,1}^n, Z_{2,2}^n, \cdots, Z_{2,K}^n)$, where $Z_{j,k}^n = (Z_{j,k}^n(t), t \geq 0)$ tracks the amount of effort needed to process all class $k$ jobs currently at station $j$. We also define the processes $V_j^n = \sum_k Z_{j,k}^n$, which is the total amount of work currently at station $j$. Notice that $V_1^n = W_1^n$. For station 2, $V_2^n$ is the work that is visible to the station 2 server, not including work that is upstream. We refer to $V_j^n$ as the station $j$ *immediate workload process*. To write the workload processes for the $n^{th}$ network in our sequence explicitly, we begin by constructing a "free" process. This free process $X_j^n$ would track station $j$'s total workload process in the case where the server worked constantly and the actual workload process was allowed to have negative values. We could express $X_j^n(t)$ as

$$X_j^n(t) \equiv W_j^n(0) + \sum_{k=1}^{K} S_{j,k}^n(A_k^n(t)) - t \tag{7.30}$$

and its diffusion-scaled version as

$$\tilde{X}_j^n(t) \equiv \frac{X_j^n(nt)}{\sqrt{n}} = \tilde{W}_j^n(0) + \frac{1}{\sqrt{n}} \sum_{k=1}^{K} S_{j,k}^n(A_k^n(nt)) - \sqrt{n}t \tag{7.31}$$

where

$$W_1^n(0) = \sum_{k=1}^{K} \hat{S}_{1,k}^n(\hat{Q}_{1,k}(0)) \tag{7.32}$$

$$W_2^n(0) = \sum_{k=1}^{K} \hat{S}_{2,k}^n(\hat{Q}_{1,k}(0) + \hat{Q}_{2,k}(0)) \tag{7.33}$$

In reality, however, our system does not allow the workload process to take on negative values. Therefore, we define a station $j$ idle time process, $I_j^n$, that actually tracks the cumulative amount of time that the station $j$ server is idle, and by our assumptions, this occurs only when the immediate workload for station $j$ is zero. We can express $I^n(t) = (I_1^n(t), I_2^n(t))$ as

$$I_1^n(t) = \Phi(X_1^n)(t) \tag{7.34}$$

$$I_2^n(t) = \Phi(X_2^n - U_2^n)(t) \tag{7.35}$$

Notice that the expression for $I_2^n(t)$ contains the station 2 upstream workload process $U_2^n(t)$, which tracks the effort required by the server at station 2 to process the jobs in the system at time $t$ that have not yet reached station 2. We write this process explicitly as:

$$\begin{aligned} U_2^n(t) = \sum_{k=1}^{K} &\left[\hat{S}_{2,k}^n(Q_{1,k}^n(0) + Q_{2,k}^n(0)) - \hat{S}_{2,k}^n(Q_{1,k}^n(0) + Q_{2,k}^n(0) - \hat{Q}_{1,k}^n(t))\right] \\ &+ \sum_{k=1}^{K} \left[S_{2,k}^n(A_k^n(t)) - S_{2,k}^n\big(A_k^n(t) - (Q_{1,k}^n(t) - \hat{Q}_{1,k}^n(t))\big)\right]. \end{aligned} \tag{7.36}$$

Adding the idle time process to the free process results in the true workload process, and we have

$$\tilde{W}_1^n(t) = \tilde{X}_1^n(t) + \tilde{I}_1^n(t) \tag{7.37}$$

$$\tilde{W}_2^n(t) = \tilde{X}_2^n(t) + \tilde{I}_2^n(t). \tag{7.38}$$

Additionally, we have

$$\tilde{V}_2^n(t) = \tilde{W}_2^n(t) - \tilde{U}_2^n(t). \tag{7.39}$$

The following well-known convergence result for the station 1 total workload process immediately follows from the above definitions and Lemma 7.5.1 and is proved in Jennings (2008)

**Lemma 7.5.2.** *Let* $\tilde{X}^n = (\tilde{X}_1^n, \tilde{X}_2^n)$. *Assume* (7.1)-(7.3), (7.5), *and* (7.10). *Then, as* $n \to \infty$

$$(\tilde{X}^n, \tilde{W}_1^n) \Rightarrow (\tilde{X}, \tilde{X}_1 + \Phi(\tilde{X}_1)) \tag{7.40}$$

Lemma 7.5.2 is the first step in proving the convergence result given in (7.12). However, it is more complicated to prove convergence of the station 2 total workload process. As discussed above, we can write $W_2^n(t) = X_2^n(t) + I_2^n(t)$ where $I_2^n(t) \equiv \Phi(X_2^n(t) - U_2^n(t))$. We must show that $\tilde{I}_2^n(t)$ converges despite the divergence of $\tilde{U}_2^n(t)$.

### *7.5.3 C-tightness and asymptotic behavior of processes*

We now formally define what it means for a process to be of order $g(n)$ for some function $g$. We also formalize a notion of boundedness for the processes that are stochastic.

**Definition 7.5.3.** *A sequence of random variables $\{Y^n, n \geq\}$ is said to be of order $g(n)$ in probability if for every $\epsilon > 0$, there exists a constant, $C_\epsilon$, such that for sufficiently large $n$,*

$$\mathbb{P}\left(|Y^n| > C_\epsilon g(n)\right) < \epsilon. \tag{7.41}$$

This property is denoted $Y^n = O_p(g(n))$, and if the $Y^n$ are $K$-dimensional, then $Y^n = O_p(g(n))$ if and only if $Y_k^n = O_p(g(n))$ for each $k = 1, \dots, K$. Notice that with this notation, Proposition 7.4.2 implies that for all $i \in \Gamma_1^n$, $(\tau_{i+1,1,1}^n - \tau_{i,1,1}^n) = O_p(\sqrt{n})$.

**Definition 7.5.4.** *For a given sequence of random variables $\{Y^n, n \geq 1\}$, by $Y^n = o_p(g(n))$ it is meant that for every $\epsilon > 0$ and sufficiently large $n$,*

$$\mathbb{P}\left(|Y^n| > g(n)\epsilon\right) < \epsilon \tag{7.42}$$

If $Y^n = o_p(g(n))$, then we say that $Y^n$ is negligible relative to $\sqrt{n}$.

Brownian motion is an almost surely continuous process. Therefore, any sequence of processes that converges to Brownian motion is *C-tight*

**Definition 7.5.5.** *A family of processes is C-tight if it is tight and all weak limit points of the sequence of their laws are laws of continuous processes.*

The following lemma is a consequence of C-tightness and Lemma 7.5.1. It gives us useful bounds on many of our processes and portions of it along with its proof can be found in Lemma 3.5 of Jennings (2008).

**Lemma 7.5.6.** *For any $\epsilon > 0$, $b \in [0, 1)$, constants $K$ and $T$ and sufficiently large $n$,*

*(i)* $\mathbb{P}\left(\sup_{s\leq T, t\leq K} |A^n_k(sn + tn^b) - A^n_k(sn) - \lambda^n_k tn^b| > \epsilon\sqrt{n}\right) < \epsilon$, $k = 1, \ldots, K$

*(ii)* $\mathbb{P}\left(\sup_{s\leq T, t\leq K} |S^n_{j,k}(sn + tn^b) - S^n_{j,k}(sn) - \mu^n_{j,k} tn^b| > \epsilon\sqrt{n}\right) < \epsilon$, $j = 1, 2$, $k = 1, \ldots, K$

*(iii)* $\mathbb{P}\left(\sup_{s\leq T, t\leq K} |X^n_j(sn + tn^b) - X^n_j(sn)| > \epsilon\sqrt{n}\right) < \epsilon$, $j = 1, 2$

*(iv)* $\mathbb{P}\left(\sup_{s\leq T, t\leq K} |W^n_j(sn + tn^b) - W^n_j(sn)| > \epsilon\sqrt{n}\right) < \epsilon$, $j = 1, 2$

*(v)* $\mathbb{P}\left(\sup_{s\leq T, t\leq K} |\hat{S}^n_{j,k}(sn + tn^b) - \hat{S}^n_{j,k}(sn) - \mu^n_{j,k} tn^b| > \epsilon\sqrt{n}\right) < \epsilon$, $j = 1, 2$, $k = 1, \ldots, K$

### *7.5.4 Convergence to steady state*

Our systems quickly reach an equilibrium when the workload oscillates among the job classes in a predictable cyclic fashion. The system reaches this steady state by time $n^{2/3}$. Due to the existence of a stationary distribution for the workload vector transition matrix at station 1, we can assume that the workload vector is well behaved throughout a production run once the system has reached steady state. Therefore, we introduce a station 1 *approximating workload vector process* $\hat{\boldsymbol{Z}}_{\boldsymbol{1}}^{\boldsymbol{n}}(t)$ that is a piecewise linear function, which takes values at the corresponding eigenvector lifting operator at the beginning of each production run. We can show that this process is a close approximation of the true workload vector process.

The process is identical to the one given in Jennings (2010): For $t \in [\tau^n_{i,1,k}, \tau^n_{i,1,k+1})$, let

$$\hat{\boldsymbol{Z}}_{\boldsymbol{1}}^{\boldsymbol{n}}(t) \equiv W^n_1(\tau^n_{i,1,1}) \left[ \frac{(\tau^n_{i,1,k+1} - t)\boldsymbol{\pi}^{\boldsymbol{(k)}} + (t - \tau^n_{i,1,k})\boldsymbol{\pi}^{\boldsymbol{(k+1)}}}{\tau^n_{i,1,k+1} - \tau^n_{i,1,k}} \right]. \tag{7.43}$$

For any $T \geq 0$, we define

$$\hat{\Gamma}^n_1 = \hat{\Gamma}^n_1(T) = \{i \in \Gamma^n_1 : \tau^n_{i,1,1} > n^{2/3}\} \tag{7.44}$$

to be the set of cycles that end before time $nT$ and start after time $n^{2/3}$, a time when the system will have gotten arbitrarily close to steady state. We can now state the following approximation, which is proven in Proposition 5.6 of Jennings (2010):

$$\sup_{i \in \hat{\Gamma}^n} \sup_{t \in [\tau^n_{i,1,1} - \tau^n_{i+1,1,1})} \|\boldsymbol{Z}^{\boldsymbol{n}}_{\boldsymbol{1}}(t) - \hat{\boldsymbol{Z}}^{\boldsymbol{n}}_{\boldsymbol{1}}(t)\| = o_p(\sqrt{n}). \tag{7.45}$$

We also note two important relationships between the queue-length and workload processes that hold once the system has reached steady state:

**Proposition 7.5.7.** *For any* $T \geq 0$

$$\sup_{s \leq nT} \|\boldsymbol{Q}^n_j(s) - M_j^{-1}\boldsymbol{Z}^{\boldsymbol{n}}_{\boldsymbol{j}}(s)\| = o_p(\sqrt{n}) \tag{7.46}$$

*and*

$$\sup_{s \leq nT} |\langle \boldsymbol{\theta}, \boldsymbol{Z}^{\boldsymbol{n}}_{\boldsymbol{1}}(s)\rangle - U^n_2(s)| = o_p(\sqrt{n}) \tag{7.47}$$

*Proof.* This proposition is a generalization of Proposition 3.1 and Corollary 3.1 of Jennings (2008). The ideas for the proof can be found therein. □

We use the station 1 approximate workload vector process to define two new approximating processes. First, we define the *approximate station 2 upstream workload* *process*, $\hat{U}_2^n(t) \equiv \langle \boldsymbol{\theta}, \hat{\boldsymbol{Z}}_{\boldsymbol{1}}^{\boldsymbol{n}}(t) \rangle$. Recalling equation (7.6), we can expand this definition and we have

$$\begin{aligned}
\hat{U}_2^n(s) &= \langle \boldsymbol{\theta}, \hat{\boldsymbol{Z}}_{\boldsymbol{1}}^{\boldsymbol{n}}(s) \rangle \\
&= \left\langle \theta, W_1^n(\tau_{i,1,1}^n) \left[ \frac{(\tau_{i,1,k+1}^n - s)\boldsymbol{\pi}^{(\boldsymbol{k})} + (s - \tau_{i,1,k}^n)\boldsymbol{\pi}^{(\boldsymbol{k+1})}}{\tau_{i,1,k+1}^n - \tau_{i,1,k}^n} \right] \right\rangle \qquad (7.48) \\
&= W_1^n(\tau_{i,1,1}^n)\ell^{(k)} \left( \frac{s - \tau_{i,1,k}^n}{\tau_{i,1,k+1}^n - \tau_{i,1,k}^n} \right).
\end{aligned}$$

Next, we define the *approximate station 2 immediate workload process*,

$$\hat{V}_2^n(s) = W_2^n(T^n(s)) - \hat{U}_2^n(s), \qquad (7.49)$$

where $W_2^n(T^n(s))$ is the "sampled" station 2 total workload process. We have the following propositions:

**Proposition 7.5.8.** *For any* $T \geq 0$

$$\sup_{T^n(n^{2/3}) \leq s \leq nT} |U_2^n(s) - \hat{U}_2^n(s)| = o_p(\sqrt{n}) \qquad (7.50)$$

*and*

$$\sup_{T^n(n^{2/3}) \leq s \leq nT} |V_2^n(s) - \hat{V}_2^n(s)| = o_p(\sqrt{n}) \qquad (7.51)$$

*Proof.* First, we notice that by (7.47)

$$\mathbb{P}\left( \sup_{T^n(n^{2/3}) \leq s \leq nT} |U_2^n(s) - \langle \boldsymbol{\theta}, \boldsymbol{Z}_{\boldsymbol{1}}^{\boldsymbol{n}}(s) \rangle| > \epsilon\sqrt{n}/2 \right) < \epsilon/2 \qquad (7.52)$$

and by (7.45)

$$\mathbb{P}\left( \sup_{T^n(n^{2/3}) \leq s \leq nT} \left| \langle \boldsymbol{\theta}, \boldsymbol{Z}_{\boldsymbol{1}}^{\boldsymbol{n}}(s) \rangle - \langle \boldsymbol{\theta}, \hat{\boldsymbol{Z}}_{\boldsymbol{1}}^{\boldsymbol{n}}(s) \rangle \right| > \epsilon\sqrt{n}/2 \right) < \epsilon/2\,. \qquad (7.53)$$

Equation (7.50) then follows immediately. Similarly, by part (iv) of Lemma 7.5.6

$$\mathbb{P}\left(\sup_{T^n(n^{2/3})\leq s\leq nT} |W_2^n(T^n(s) - W_2^n(s)| > \epsilon\sqrt{n}/2\right) < \epsilon/2\,. \tag{7.54}$$

This fact along with (7.50) gives us (7.51).

□

Lastly, we note that we can approximate the lengths of steady state cycles and production runs. The following propositions make these approximations precise. The first result, which can be found in Proposition 5.7 of Jennings (2010), states that once the system has reached steady state, the length of a class $k$ production run at station 1 is just a fixed fraction of the length of the entire cycle. Furthermore, that fraction is the fractional contribution of class $k$ jobs to station 1's limiting overall utilization.

**Proposition 7.5.9.** *For any* $k = 1, \ldots, K$*,*

$$\sup_{i\in\hat{\Gamma}_1^n} |\tau^n_{i,1,k+1} - \tau^n_{i,1,k} - \rho_{1,k}(\tau^n_{i+1,1,1} - \tau^n_{i,1,1})| = o_p(\sqrt{n}) \tag{7.55}$$

The next proposition, which can be found in Proposition 5.8 of Jennings (2010) states that the steady state cycle length is proportional to the total workload at station 1.

**Proposition 7.5.10.** *There exists some constant* $\tau$ *such that*

$$\sup_{i\in\hat{\Gamma}_1^n} |\tau W_1^n(\tau^n_{i,1,1}) - (\tau^n_{i+1,1,1} - \tau^n_{i,1,1})| = o_p(\sqrt{n})\,, \tag{7.56}$$

*where*

$$\tau = \left(\sum_k \rho_{1,k}\left(\frac{1-\rho_{1,k}}{1-P^{(k)}_{k,k}} - \sum_{j=1}^{k-1}\rho_{1,j}\right)\right)^{-1}. \tag{7.57}$$

The constant $\tau$ is a normalizing constant used in the components of the eigenvectors $\boldsymbol{\pi}$ given in (7.22).

## 7.6 Convergence of the Station 2 Idle Process

To complete the proof of Theorem 7.3.1, we must show that station 2 diffusion-scaled idle process, $\tilde{I}_2^n(t) \equiv \Phi(\tilde{X}_2^n(t) - \tilde{U}_2^n(t))$, converges to $\tilde{I}_2 \equiv \Phi(\tilde{X}_2 - \gamma\tilde{W}_1)$. In this section, we will introduce two auxiliary idle time process that progressively approximate $\tilde{I}_2^n$. The first of these is $\hat{I}_2^n = \{\hat{I}_2^n(t), t \geq 0\}$, where

$$\hat{I}_2^n(t) = \Phi(X_2^n \circ T_1^n - \hat{U}_2^n)(t). \tag{7.58}$$

The second auxiliary process requires that we introduce a "delayed" version of the approximate upstream workload process, which we denote by $U_2'^n = \{U_2'^n(t), t \geq 0\}$ and define as

$$U_2'^n(t) \equiv \begin{cases} \gamma W_1^n(0) & t < T_1^n(n^{2/3}), \\ \hat{U}_2^n(t) & t \geq T_1^n(n^{2/3}). \end{cases} \tag{7.59}$$

This is a delayed approximate upstream workload process because it doesn't equal the true approximate upstream workload process until the system has reached steady state. Now, we can define the second auxiliary process that we will use to approximate the idle time process. Let $I_2'^n = \{I_2'^n(t), t \geq 0\}$, where

$$I_2'^n(t) = \Phi(\tilde{X}_2^n \circ T_1^n - U_2'^n \circ T_1^n)(t). \tag{7.60}$$

Notice that this process is similar to $\hat{I}_2^n(t)$, except that it uses a sampled and delayed version of the approximate upstream workload process.

**Proposition 7.6.1.** *For all $s \leq nT$,*

$$|\tilde{I}_2^n(s) - \hat{I}_2^n(s)| \xrightarrow{P} 0 \tag{7.61}$$

*and*

$$\hat{I}_2^n \Rightarrow \Phi\left(\tilde{X}_2 - \gamma\tilde{W}_1\right) \tag{7.62}$$

*Proof.* The auxiliary idle time process given in (7.58) contains a version of the station 2 total workload free process that is sampled at the beginning of the station 1 cycles. By part (iii) of Lemma 7.5.6, we have

$$\sup_{s \leq n^{2/3}} |X_2^n(s) - X_2^n(0)| = o_p(\sqrt{n}) \,. \tag{7.63}$$

Equation (7.63) combined with the constraint on the initial condition $\tilde{X}_2(0) \geq \theta^* \tilde{W}_1(0)$, implies

$$\inf_{s \leq n^{2/3}} X_2^n(s) - \theta^* W_1^n(0) \geq o_p(\sqrt{n}) \,. \tag{7.64}$$

We also note that $\langle \boldsymbol{\theta}, Z_1^n(t) \rangle \leq \theta^* W_1^n(t)$ for all $t \geq 0$ by definition. Thus, by Part (iv) of Lemma 7.5.6 and Equation (7.47) of Proposition 7.5.7, we have that

$$\sup_{s \leq n^{2/3}} U_2^n(s) - \theta^* W_1^n(0) \leq o_p(\sqrt{n}). \tag{7.65}$$

It immediately follows that

$$I_2^n(T_1^n(n^{2/3})) = o_p(\sqrt{n}) \,. \tag{7.66}$$

Now, it is clear from (7.63) that we also have

$$\sup_{s \leq n^{2/3}} |X_2^n(T_1^n(s)) - X_2^n(0)| = o_p(\sqrt{n}) \tag{7.67}$$

and thus by (7.64),

$$\inf_{s \leq n^{2/3}} X_2^n(T_1^n(s)) - \theta^* W_1^n(0) \geq o_p(\sqrt{n}) \,. \tag{7.68}$$

We also note that $\hat{U}_2^n(s) = \langle \boldsymbol{\theta}, \hat{Z}_1^n(s) \rangle \leq \theta^* W_1^n(T_1^n(s))$. So, again by part (iv) of Lemma 7.5.6, we have

$$\sup_{s \leq n^{2/3}} \hat{U}_2^n(s) - \theta^* W_1^n(0) \leq o_p(\sqrt{n}), \tag{7.69}$$

and it immediately follows that

$$\hat{I}_2^n(T_1^n(n^{2/3})) = o_p(\sqrt{n}). \tag{7.70}$$

Equations (7.66) and (7.70) allow us to write,

$$I_2^n(nT) = \Phi(X_2^n - U_2^n)(nT) = -\inf_{T_1^n(n^{2/3}) \le s \le nT} [(X_2^n(s) - U_2^n(s)) \wedge 0] + o_p(\sqrt{n}) \tag{7.71}$$

and

$$\hat{I}_2^n(nT) = \Phi(X_2^n \circ T_1^n - \hat{U}_2^n)(nT) = -\inf_{T_1^n(n^{2/3}) \le s \le nT} \left[\left(X_2^n(T_1^n(s)) - \hat{U}_2^n(s)\right) \wedge 0\right] + o_p(\sqrt{n}). \tag{7.72}$$

By Equation (7.50) of Proposition 7.5.8, Proposition 7.4.2, and part (iii) of Lemma 7.5.6,

$$\sup_{s \le nT} \left| I_2^n(s) - \hat{I}_2^n(s) \right| = o_p(\sqrt{n}). \tag{7.73}$$

Now, recall our second auxiliary idle time process

$$I_2'^n(t) = \Phi(\tilde{X}_2^n \circ T_1^n - U_2'^n \circ T_1^n)(t). \tag{7.74}$$

Using similar arguments for $I_2^n(n^{2/3}))$ and $\hat{I}_2(n^{2/3})$, it follows that

$$I_2'^n(n^{2/3}) = o_p(\sqrt{n}), \tag{7.75}$$

so that we have

$$\begin{aligned} I_2'^n(nt) &= \Phi(X_2^n \circ T_1^n - U_2'^n \circ T_1^n)(t) \\ &= -\inf_{T_1^n(n^{2/3}) \le s \le nT} (X_2^n(T_1^n(s)) - U_2'^n(T_1^n(s))) \wedge 0 + o_p(\sqrt{n}). \end{aligned} \tag{7.76}$$

Notice that for any $s \ge T_1^n(n^{2/3})$,

$$\hat{U}_2^n(s) \le \gamma W_1^n(T_1^n(s)) = \hat{U}_2^n(T_1^n(s)) = U_2'^n(T_1^n(s)). \tag{7.77}$$

It follows immediately that

$$\sup_{s \le nT} \left| I_2'^n(s) - \hat{I}_2^n(s) \right| = o_p(\sqrt{n}) \tag{7.78}$$

and

$$\sup_{s \le nT} |I_2'^n(s) - I_2^n(s)| = o_p(\sqrt{n}). \tag{7.79}$$

Finally, we note that

$$\tilde{X}_2^n \circ \bar{T}^n \to \tilde{X}_2 \quad \text{and} \quad \frac{1}{\sqrt{n}} \hat{U}_2^n(n\bar{T}^n) \Rightarrow \gamma \tilde{W}_1 \,. \tag{7.80}$$

Because $\Phi$ is continuous, we have

$$\frac{1}{\sqrt{n}} I_2'^n(nt) \Rightarrow \tilde{I}_2 = \Phi(\tilde{X}_2 - \gamma \tilde{W}_1) \tag{7.81}$$

and the proof is complete. □

## 7.7 Proof of Main Results

We now complete the proof of the fact that the total workload processes for station 1 and station 2 converge to a two-dimensional Brownian motion in a wedge.

### *7.7.1 Proof of Theorem 7.3.1*

*Proof.* Lemma 7.5.2 showed that the station 1 total workload process converges to a reflected Brownian motion. It required more work to prove Station 2's convergence. Section 7.6 gives us the tools that we need to prove that the diffusion scaled station 2 idle time process, $\tilde{I}_2^n = \Phi(\tilde{X}_2^n - \tilde{U}_2^n)$ converges. By Lemma 7.5.2, Proposition 7.6.1, and the continuous mapping theorem, we have

$$\left(\tilde{W}_1^n, \tilde{X}^n, \tilde{I}_2^n\right) \Rightarrow \left(\tilde{W}_1, \tilde{X}, \tilde{I}_2\right) \tag{7.82}$$

where

$$\tilde{I}_2 \equiv \Phi\left(\tilde{X}_2 - \gamma \tilde{W}_1\right) \tag{7.83}$$

It follows that

$$\tilde{W}_2^n = \tilde{X}_2^n + \tilde{I}_2^n \Rightarrow \tilde{X}_2 + \tilde{I}_2 \tag{7.84}$$

together with the convergence of $\tilde{W}_1^n$.

□

As an immediate consequence of Theorem 7.3.1, we also have the following lemma stating that the station 1 and station 2 total workload processes are asymptotically bounded in probability:

**Lemma 7.7.1.** *For any time $T \geq 0$ and any $j = 1, 2$,*

$$\sup_{t \leq nT} W_j^n(t) = O_p(\sqrt{n}). \tag{7.85}$$

*7.7.2 Proof of Theorem 7.3.2*

As previously discussed, the upstream workload process $\tilde{U}_2^n$ does not converge, and thus, the immediate station 2 workload process $\tilde{V}_2^n = \tilde{W}_2^n - \tilde{U}_2^n$ does not converge as $n \to \infty$. However, the process $\tilde{V}_2^n$ obeys an averaging principle. We now prove Theorem 7.3.2:

*Proof.*

$$\begin{aligned}
\int_0^T f(\tilde{V}_2^n(s))ds &= \frac{1}{n}\int_0^{nT} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds \\
&= \frac{1}{n}\int_0^{T_1^n(n^{2/3})} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds + \frac{1}{n}\int_{T_1^n(n^{2/3})}^{T_1^n(nT)} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds \qquad (7.86) \\
&\quad + \frac{1}{n}\int_{T_1^n(nT)}^{nT} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds
\end{aligned}$$

By Lemma 7.7.1, we know that $W_2^n(s)$ is asymptotically bounded in probability. Also, $U_2^n(s) \leq W_2^n(s)$ for all $s \geq 0$. Thus, $V_2^n(s) = W_2^n(s) - U_2^n(s)$ is bounded in probability and the first and last terms on the right hand side of the above equation are asymptotically negligible due to the boundedness of the intervals of integration and the continuity of $f$. That is,

$$\frac{1}{n}\int_0^{T_1^n(n^{2/3})} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds \xrightarrow{P} 0 \tag{7.87}$$

and

$$\frac{1}{n}\int_{T_1^n(nT)}^{nT} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds \xrightarrow{P} 0. \tag{7.88}$$

For the remaining term, we can write:

$$\begin{aligned}&\int_{T_1^n(n^{2/3})}^{T_1^n(nT)} f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) ds \\ &\quad = \int_{T_1^n(n^{2/3})}^{T_1^n(nT)} f\left(\frac{\hat{V}_2^n(s)}{\sqrt{n}}\right) ds + \int_{T_1^n(n^{2/3})}^{T_1^n(nT)} \left[f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) - f\left(\frac{\hat{V}_2^n(s)}{\sqrt{n}}\right)\right] ds.\end{aligned} \tag{7.89}$$

The second term on the right hand side of the above equation is asymptotically negligible. More precisely, by Equation (7.51) of Proposition 7.5.8 and the asymptotic boundedness of $V_2$, we have

$$\int_{T_1^n(n^{2/3})}^{T_1^n(nT)} \left[f\left(\frac{V_2^n(s)}{\sqrt{n}}\right) - f\left(\frac{\hat{V}_2^n(s)}{\sqrt{n}}\right)\right] ds \xrightarrow{P} 0 \tag{7.90}$$

as $n \to \infty$. Now, we have

$$\int_{T_1^n(n^{2/3})}^{T_1^n(nT)} f\left(\frac{\hat{V}_2^n(s)}{\sqrt{n}}\right) ds = \sum_{i\in\hat{\Gamma}_1^n} \sum_k \int_{\tau_{i,1,k}^n}^{\tau_{i,1,k+1}^n} f\left(\frac{W_2^n(T_1^n(s))}{\sqrt{n}} - \frac{\hat{U}_2^n(s)}{\sqrt{n}}\right) ds$$

$$= \sum_{i\in\hat{\Gamma}_1^n} \sum_k \int_{\tau_{i,1,k}^n}^{\tau_{i,1,k+1}^n} f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}\left(\frac{s-\tau_{i,1,k}^n}{\tau_{i,1,k+1}^n - \tau_{i,1,k}^n}\right)\right) ds$$

$$= \sum_{i\in\hat{\Gamma}_1^n} \sum_k (\tau_{i,1,k+1}^n - \tau_{i,1,k}^n) \int_0^1 f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}(u)\right) du \tag{7.91}$$

where we have done a change of variables by letting

$$u = \frac{s-\tau_{i,1,k}^n}{\tau_{i,1,k+1}^n - \tau_{i,1,k}^n}.$$

Notice that Equation 7.91 involves summing over the lengths of each production run in each cycle. We would like to replace the lengths of the production runs with some appropriately chosen fractions of the cycle length. To this end, we let

$$H^n = \sum_{i\in\hat{\Gamma}_1^n} \sum_k (\tau_{i,1,k+1}^n - \tau_{i,1,k}^n) \int_0^1 f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}(u)\right) du$$

$$- \sum_{i\in\hat{\Gamma}_1^n} (\tau_{i+1,1,1}^n - \tau_{i,1,1}^n) \sum_k \rho_{1,k} \int_0^1 f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}(u)\right) du. \tag{7.92}$$

Our goal will be to show that $H^n = o_p(n)$. However, in both terms of $H^n$, we also have to be concerned with the number of cycles over which we are summing. To handle the size of the set $\{i : i \in \hat{\Gamma}_1^n\}$, we will follow the techniques of Jennings (2010). We notice that for sufficiently long cycles, we can easily show that $H^n$ is $o_p(n)$. Therefore, for any $\epsilon > 0$, we split $H^n$ into two pieces:

$$H^n = H_0^n(\epsilon) + H_1^n(\epsilon), \tag{7.93}$$

where for each $j = 0, 1$,

$$H_j^n(\epsilon) = \sum_{i \in \tilde{\Gamma}_j^n(\epsilon)} \sum_k \left(\tau_{i,1,k+1}^n - \tau_{i,1,k}^n - \rho_{1,k}(\tau_{i+1,1,1}^n - \tau_{i,1,1}^n)\right) C_{i,k}^n \tag{7.94}$$

and

$$C_{i,k}^n = \int_0^1 f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}(u)\right) du. \tag{7.95}$$

Each $H_j^n(\epsilon)$ is a summation taken over a distinct subset of cycles, $\tilde{\Gamma}_j^n(\epsilon) \subseteq \hat{\Gamma}_1^n$, where

$$\tilde{\Gamma}_0^n(\epsilon) = \{i \in \hat{\Gamma}_1^n : \tau_{i+1,1,1}^n - \tau_{i,1,1}^n < \epsilon\sqrt{n}\},$$

$$\tilde{\Gamma}_1^n(\epsilon) = \{i \in \hat{\Gamma}_1^n : \tau_{i+1,1,1}^n - \tau_{i,1,1}^n \geq \epsilon\sqrt{n}\},$$

and

$$\tilde{\Gamma}_0^n(\epsilon) \cup \tilde{\Gamma}_1^n(\epsilon) = \hat{\Gamma}_1^n$$

We analyze $H_1^n(\epsilon)$ first. Since the length of the cycles in $\tilde{\Gamma}_1^n$ are greater than $\epsilon\sqrt{n}$, the maximum number of these cycles that can occur in the time interval $[0, nT]$ is bounded above by $2 + \sqrt{n}T/\epsilon$. So, we have

$$|H_1^n(\epsilon)| \leq \left(2 + \frac{\sqrt{n}T}{\epsilon}\right) \left(\sup_{i \in \hat{\Gamma}_1^n} |\tau_{i,1,k+1}^n - \tau_{i,1,k}^n - \rho_{1,k}(\tau_{i+1,1,1}^n - \tau_{i,1,1}^n)|\right) \cdot \left(\sup_{0 \leq z \leq \gamma} \sup_{i \in \hat{\Gamma}_1^n} \left| f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} z\right)\right|\right) \tag{7.96}$$

By Lemma 7.7.1, the station 1 and station 2 diffusion-scaled workload processes are uniformly bounded in probability. Thus, by the continuity of $f$, the last term on the right hand side is $O_p(1)$. By Proposition 7.5.9, we also know that the middle term is $o_p(\sqrt{n})$. Therefore, it follows that for any $\epsilon > 0$, $H_1^n(\epsilon) = o_p(n)$.

Now, we analyze $H_0^n(\epsilon)$. As in the proof of Theorem 6.1 of Jennings (2010), we have the following bound

$$\left|\sum_k \left((\tau^n_{i,1,k+1} - \tau^n_{i,1,k}) - \rho_{1,k}(\tau^n_{i+1,1,1} - \tau^n_{i,1,1})\right) C^n_{i,1,k}\right| \leq (\tau^n_{i+1,1,1} - \tau^n_{i,1,1}) \left(\max_k C^n_{i,1,k} - \min_k C^n_{i,1,k}\right). \quad (7.97)$$

Now, by Equations (7.94) and (7.97), we have

$$\begin{aligned} |H_0^n(\epsilon)| &\leq \sum_{i \in \tilde{\Gamma}_0^n(\epsilon)} \left|\sum_k \left(\tau^n_{i,1,k+1} - \tau^n_{i,1,k} - \rho_{1,k}(\tau^n_{i+1,1,1} - \tau^n_{i,1,1})\right) C^n_{i,1,k}\right| \\ &\leq nT \sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \left(\max_k C^n_{i,1,k} - \min_k C^n_{i,1,k}\right). \end{aligned} \quad (7.98)$$

and we want to show that for any $\delta > 0$,

$$\lim_{\epsilon \downarrow 0} \limsup_{n \to \infty} \mathbb{P}\left(\sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \left[\max_k C^n_{i,1,k} - \min_k C^n_{i,1,k}\right] > \delta\right) = 0. \quad (7.99)$$

First, notice that

$$\begin{aligned} &\sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \left[\max_k \sup_{0 \leq u \leq 1} \left(\frac{W_2^n(\tau^n_{i,1,1})}{\sqrt{n}} - \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} \ell^{(k)}(u)\right)\right. \\ &\qquad\qquad \left. - \min_k \inf_{0 \leq u \leq 1} \left(\frac{W_2^n(\tau^n_{i,1,1})}{\sqrt{n}} - \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} \ell^{(k)}(u)\right)\right] \\ &= \sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \left[\max_k \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} \ell^{(k)}(u) - \min_k \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} \ell^{(k)}(u)\right] \\ &\leq \sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} \gamma_1 \end{aligned}$$

Now, by Proposition 7.5.10 and the fact that $\tau^n_{i+1,1,1} - \tau_{i,1,1} < \epsilon\sqrt{n}$ for $i \in \tilde{\Gamma}_0^n(\epsilon)$, we have that there exists some $N(\epsilon)$ such that

$$\mathbb{P}\left(\sup_{i \in \tilde{\Gamma}_0^n(\epsilon)} \frac{W_1^n(\tau^n_{i,1,1})}{\sqrt{n}} > (1+\tau)\epsilon\right) < \epsilon \quad (7.100)$$

for all $n \geq N(\epsilon)$. It immediately follows that

$$\mathbb{P}\left( \max_{k_1,k_2} \sup_{i,j\in\tilde{\Gamma}_0^n(\epsilon)} \sup_{u,v\in[0,1]} \left| \left( \frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k_1)}(u) \right) \right.\right.$$
$$\left.\left. - \left( \frac{W_2^n(\tau_{j,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{j,1,1}^n)}{\sqrt{n}} \ell^{(k_2)}(v) \right) \right| > 2\gamma_1(1+\tau)\epsilon \right) < \epsilon \tag{7.101}$$

for all $n \geq N(\epsilon)$. By the continuity of $f$, we can always choose $\epsilon \leq \delta/(2\gamma_1(1+\tau))$ in the above inequality so that

$$\mathbb{P}\left( \sup_{i\in\tilde{\Gamma}_0^n(\epsilon)} \left[ \max_k C_{i,1,k}^n - \min_k C_{i,1,k}^n \right] > \delta \right) < \epsilon \tag{7.102}$$

for all $n \geq N(\epsilon)$. Taking limits gives us (7.99), and by (7.98), we have

$$\lim_{\epsilon\downarrow 0} \limsup_{n\to\infty} \frac{1}{n} H_0^n(\epsilon) = 0. \tag{7.103}$$

Therefore, by (7.93) and (7.96), we can conclude that

$$H^n = o_p(n). \tag{7.104}$$

This justifies replacing the production runs as in (7.92).

Now, we notice that we can also replace the summation over cycles with a time integral and rewrite our workload processes as functions of time using our time change function, $T_1^n(t)$. So we have,

$$\sum_{i\in\hat{\Gamma}_1^n} (\tau_{i+1,1,1}^n - \tau_{i,1,1}^n) \sum_k \rho_{1,k} \int_0^1 f\left( \frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}} \ell^{(k)}(u) \right) du$$
$$= \int_{T_1^n(n^{2/3})}^{T_1^n(nT)} \sum_k \rho_{1,k} \int_0^1 f\left( \frac{W_2^n(t)}{\sqrt{n}} - \frac{W_1^n(t)}{\sqrt{n}} \ell^{(k)}(u) \right) du dt + o_p(n), \tag{7.105}$$

where the last equality comes from Part (iv) of Lemma 7.5.6 and the asymptotic boundedness of $\tilde{W}_1$ and $\tilde{W}_2$. We can show that the cycles before steady state and the last partial cycle are negligible relative to $n$. That is,

$$\int_0^{T_1^n(n^{2/3})} \sum_k \rho_{1,k} \int_0^1 f\left(\frac{W_2^n(T_1^n(t))}{\sqrt{n}} - \frac{W_1^n(T_1^n(t))}{\sqrt{n}}\ell^{(k)}(u)\right) dudt = o_p(n)$$

and

$$\int_{T_1^n(nT)}^{nT} \sum_k \rho_{1,k} \int_0^1 f\left(\frac{W_2^n(T_1^n(t))}{\sqrt{n}} - \frac{W_1^n(T_1^n(t))}{\sqrt{n}}\ell^{(k)}(u)\right) dudt = o_p(n),$$

which follow from the facts that $T_1^n(n^{2/3}) \leq n^{2/3}$, $nt - T_1^n(nT) = O_p(\sqrt{n})$ by Proposition 7.4.2, Part (iv) of Lemma 7.5.6, and the continuity of $f$. As an immediate consequence, we have

$$\begin{aligned}
\sum_{i\in\hat{\Gamma}_1^n}(\tau_{i+1,1,1}^n - \tau_{i,1,1}^n) \sum_k \rho_{1,k} \int_0^1 f\left(\frac{W_2^n(\tau_{i,1,1}^n)}{\sqrt{n}} - \frac{W_1^n(\tau_{i,1,1}^n)}{\sqrt{n}}\ell^{(k)}(u)\right) du \\
= \int_0^{nT} \sum_k \rho_{1,k} \int_0^1 f\left(\frac{W_2^n(t)}{\sqrt{n}} - \frac{W_1^n(t)}{\sqrt{n}}\ell^{(k)}(u)\right) dudt + o_p(n) \\
= n\int_0^{T} \sum_k \rho_{1,k} \int_0^1 f\left(\tilde{W}_2^n(t) - \tilde{W}_1^n(t)\ell^{(k)}(u)\right) dudt + o_p(n).
\end{aligned} \tag{7.106}$$

Next, we define the functional $h : D[0,\infty) \times D[0,\infty) \mapsto (-\infty,\infty)$ where for any $w_1, w_2 \in D[0,\infty)$ and $T \geq 0$,

$$h(w_1,w_2)(T) = \int_0^{T} \sum_k \rho_{1,k} \int_0^1 f\left(w_2 - w_1\ell^{(k)}(u)\right) dudt \tag{7.107}$$

is a continuous map. Now, by (7.12) and the Continuous Mapping Theorem,

$$\int_0^{T} \sum_k \rho_{1,k} \int_0^1 f\left(\tilde{W}_2^n(t) - \tilde{W}_1^n(t)\ell^{(k)}(u)\right) dudt \Rightarrow \int_0^{T} \sum_k \rho_{1,k} \int_0^1 f\left(\tilde{W}_2(t) - \tilde{W}_1(t)\ell^{(k)}(u)\right) dudt \tag{7.108}$$

This convergence, along with (7.105) and (7.106), imply

$$\left(\frac{1}{n}\right)\int_{T_1^n(n^{2/3})}^{T_1^n(nT)}\sum_k \rho_{1,k}\int_0^1 f\left(\frac{W_2^n(t)}{\sqrt{n}}-\frac{W_1^n(t)}{\sqrt{n}}\ell^{(k)}(u)\right)dudt$$
$$\Longrightarrow \int_0^T \sum_k \rho_{1,k}\int_0^1 f\left(\tilde{W}_2(t)-\tilde{W}_1(t)\ell^{(k)}(u)\right)dudt. \quad (7.109)$$

The averaging principle in (7.14) follows from (7.86), (7.87)-(7.92), (7.104), (7.105), and (7.109). □

## 7.8 Future Directions

It turns out that we can extend the averaging principle of the station 2 immediate workload to a multidimensional setting. If we consider the $K-$*dimensional* workload vector $Z_2^n$. In this setting, knowing where the workload presides at station 2, even when the system is in steady state, is not immediately obvious. If we consider a simple service policy like first-in-first-out (FIFO), we still must be able to track the arrival of work from station 1 as well as the departure of jobs from station 2, as the servers work simultaneously. We also have to account for the difference in speed between the servers at station 1 and 2. Although the cycle lengths at both stations will be identical, the lengths of the production runs will be different, and stochastic fluctuations in the workload will cause station 2 to become backlogged at some point in time. This will result in station 2 being in a different cycle than station 1 at the same point in time. With the right operators to keep track of station 2 cycles and production runs in terms of station 1 cycles and production runs, we conjecture that the station 2 immediate workload vector obeys an averaging principle much like the averaging principle for the scalar version of the station 2 immediate workload.

# 8

# Conclusion

The first part of this thesis generalizes a very useful result for a class of nonlinear stochastic partial differential equations. We have given a very general result that can be applied to any equation satisfying our hypotheses. Equations like the stochastic Navier-stokes, stochastic Cahn-Hilliard, stochastic Burger's, etc. are very useful in modeling physical phenomena. Our result is just a first, but important, step in being able to fully analyze these equations and others like them. One looming question is the case of the 2-dimensional stochastic Navier-Stokes equation. In this thesis, we have taken a novel approach to trying to determine whether its transition density is equivalent to that of the related Ornstein-Uhlenbeck process. In agreement with other articles concerning this subject, we were not able to conclude equivalence for the 2-D stochastic Navier-Stokes equation with normal viscosity. This does not constitute a proof to the contrary, however, we strongly believe that in fact it is true that the densities are not equivalent.

Showing the equivalence between the transition densities of these nonlinear equations and an Ornstein-Uhlenbeck process, which is analogous to a stochastic heat equation, has many implications for application. One immediate application that the author has recently been made aware of, is the ability to determine the Hausdorff dimension of the level sets of certain stochastic SPDEs. Much work has already been done to determine the Hausdorff dimension of the level sets of the Ornstein-Uhlenbeck process on a separable Hilbert space, and it is not a far reach to conjecture that if a nonlinear SPDE has a transition density equivalent to the O-U process for all initial data, then the Hausdorff dimension of the level sets for the nonlinear SPDE should be the same as that for the Ornstein-Uhlenbeck process. We expect that our result should be useful in many other applications of stochastic analysis as well.

The portion of this work that concerns queueing theory contributes to the literature on heavy-traffic limit theorems by extending some known results to more complex models. The future work of extending the averaging principles to a multi-dimensional setting will introduce some mathematical machinery that is not common in this field. By adding complexity to these models, we are able to more accurately model real world queueing systems. One obvious extension to the model we consider is to allow station 2 to operate under any service discipline besides First-In-First-Out. It will be interesting to see what, if any, averaging principles exist in this case.

# Appendix A

## Proof of Lemma 3.2.2

### A.1 Preliminary Results

We now give a few consequences of the minorization condition given in Assumption 3.2.1. They will be derived from the following two results from the literature.

**Proposition A.1.1.** *Let $L_1 : V_1 \to V$ and $L_2 : V_2 \to V$ be two bounded linear operators on the Hilbert spaces $V_1$ and $V_2$ that map into a third Hilbert space $V$. We denote their corresponding adjoints by $L_1^* : V \to V_1$ and $L_2^* : V \to V_2$. The following statements are equivalent:*

1. $L_1(V_1) \subset L_2(V_2)$
2. *For some $k > 0$, $|L_1^* x|_{V_1} \leq k|L_2^* x|_{V_2}$ for all $x \in V$.*
3. *For some $k > 0$, $L_1 L_1^* \leq k L_2 L_2^*$*

*Proof.* The proofs of these statement seem to originate in Douglas (1966); Embry (1973), but can also be found in our current setting in (Da Prato and Zabczyk, 2002, Proposition B.2.1) □

**Corollary A.1.2.** *If $L_1$ is a linear operator mapping a Hilbert space $V_1$ into another Hilbert space $V$ and let $L_1^\dagger : Dom(L_1^\dagger) \subset V \to V_1$ be the corresponding pseudo-inverse. If we let $T_1$ denote the positive operator on $V$ defined by $T_1 = L_1 L_1^*$, then $T_1^{1/2}(V) = L_1(V_1)$ and $|T_1^{-1/2}x|_V = |L_1^\dagger x|_V$ for all $x \in L_1(V_1)$, where $T_1^{-1/2}$ denotes the pseudo-inverse of $T_1^{1/2}$.*

*Proof.* See (Da Prato and Zabczyk, 2002, Proposition B.2.4) □

With these in hand we deduce the following result which will be a useful tool in our investigation.

## A.2 Proof of Lemma 3.2.2

*Proof.* Since $\Gamma = QQ^*$ is a positive definite, selfadjoint operator, it has a positive definite, selfadjoint square root which we will denote by $\Gamma^{\frac{1}{2}}$. Since by assumption

$$L^{-(\gamma+\delta)/2}(L^{-(\gamma+\delta)/2})^* = L^{-(\gamma+\delta)} \leq \frac{1}{c}\Gamma = \frac{1}{c}\Gamma^{\frac{1}{2}}(\Gamma^{\frac{1}{2}})^*$$

Proposition A.1.1 implies that

$$\mathcal{H}^{\gamma+\delta} = L^{-(\gamma+\delta)/2}(H) \subset \Gamma^{\frac{1}{2}}(H) = Q(\mathcal{U}), \tag{A.1}$$

where the last equality is a result of Corollary A.1.2. To see the second conclusion, we let $Q^\dagger$ be the pseudo-inverse of $Q$. We first show that $L^{-(\gamma+\delta)} \leq \frac{1}{c}\Gamma$ implies $L^{(\gamma+\delta)} \geq c\Gamma^\dagger$. Notice that $\Gamma^{-\frac{1}{2}}L^{-(\gamma+\delta)}(\Gamma^{-\frac{1}{2}})^*$ is a self-adjoint, bounded operator on $\mathcal{U}$, and by the spectral theorem it is unitarily equivalent to a real-valued multiplication operator. That is, there exists a unitary operator $U$ on $\mathcal{U}$ such that,

$$M \stackrel{\text{def}}{=} U\Gamma^{-\frac{1}{2}}L^{-(\gamma+\delta)}(\Gamma^{-\frac{1}{2}})^*U^{-1} \tag{A.2}$$

is a multiplication operator. Now, define an operator $B = U\Gamma^{\frac{1}{2}}$. Then, $\Gamma = B^*IB$ and $L^{-(\gamma+\delta)} = B^*MB$. Therefore,

$$\frac{1}{c}\Gamma \geq L^{-(\gamma+\delta)} \iff \frac{1}{c}I \geq M \iff \frac{1}{c} \geq \rho > 0 \tag{A.3}$$

where $\rho$ is the spectral radius of the operator $\Gamma^{-\frac{1}{2}}L^{-(\gamma+\delta)}(\Gamma^{-\frac{1}{2}})^*$. Furthermore, because the operator is strictly positive, the spectral mapping theorem tells us that $\Gamma^\dagger = B^{-1}I(B^*)^{-1}$ and $L^{(\gamma+\delta)} = B^{-1}M^{-1}(B^*)^{-1}$. Thus, we have

$$\frac{1}{c} \geq \rho \Longleftrightarrow c \leq \frac{1}{\rho} \Longleftrightarrow cI \leq M^{-1} \Longleftrightarrow c\Gamma^\dagger \leq L^{(\gamma+\delta)} \tag{A.4}$$

and we can write

$$L^{(\gamma+\delta)/2}(L^{(\gamma+\delta)/2})^* = L^{(\gamma+\delta)} \geq c\Gamma^\dagger = c\Gamma^{-\frac{1}{2}}(\Gamma^{-\frac{1}{2}})^* \tag{A.5}$$

Thus, by Proposition A.1.1 and Corollary A.1.2, we conclude

$$\|Q^\dagger x\|_{\mathcal{U}}^2 \leq \frac{1}{c^2}\|x\|_{\gamma+\delta}^2 . \tag{A.6}$$

□

# Appendix B

## Proof of a Moment Condition for Stochastic Cahn Hilliard

Let

$$du_t = -\Delta^2 u_t dt - \Delta f(u_t) dt + Q dW_t \tag{B.1}$$

as in Equation (4.7) on some bounded domain $D \subset \mathbb{R}^3$, and let

$$F'(u) = -f(u) = u^3 - u\,. \tag{B.2}$$

**Lemma B.0.1.** *There exists some sufficently large constant $q$ such that*

$$\mathbb{E} \sup_{s \in [0,T]} \|u_s(\omega)\|^q < \infty\,, \tag{B.3}$$

*where $\|\cdot\|$ is the $L^2$-norm.*

*Proof.* The energy functional for equation (B.1) is defined as

$$J(u) \to \frac{1}{2}\|\nabla u\|^2 + \int_D F(u) dx\,. \tag{B.4}$$

Applying Ito's formula to this functional gives

$$dJ(u_t) = \langle \nabla u, \nabla du\rangle + \langle -f(u), du_t\rangle + \frac{1}{2}\langle \nabla QdW_t, \nabla QdW_t\rangle + \frac{1}{2}\langle -f'(u)QdW_t, QdW_t\rangle$$

$$= \langle -\Delta u, du_t\rangle + \langle -f(u), du_t\rangle + \frac{1}{2}\langle \nabla QdW_t, \nabla QdW_t\rangle + \frac{1}{2}\langle -f'(u)QdW_t, QdW_t\rangle$$

where the first term comes from integrating by parts.

Now, if we let $K(u) = \nu\Delta u + f(u)$, then

$$dJ(u_t) = \langle -K(u), -\Delta K(u)\rangle dt + \langle -K(u), QdW_t\rangle + \frac{1}{2}\langle \nabla QdW_t, \nabla QdW_t\rangle + \frac{1}{2}\langle -f'(u)QdW_t, QdW_t\rangle$$

$$= -\|\nabla K(u)\|^2 dt - \sum_{k=1}^{\infty}\langle K(u), f_k\rangle dB_t^{(k)} + \frac{1}{2}\sum_{k=1}^{\infty}\|\nabla f_k\|^2 dt + \frac{1}{2}\sum_{k=1}^{\infty}\langle -f'(u), f_k^2\rangle dt$$

Define

$$\mathcal{E}_0 = \sum_k \|f_k\|^2 \quad \text{and} \quad \mathcal{E}_1 = \sum_k \|\nabla f_k\|^2 \quad \text{and} \quad M_t = -\sum_k \int_0^t \langle K(u), f_k\rangle dB_s^{(k)} \tag{B.5}$$

Notice that the quadratic variation of $M_t$ is given by

$$\langle M_t\rangle = \sum_k \int_0^t \langle K(u), f_k\rangle^2 ds \le \frac{\mathcal{E}_0}{\lambda_1}\int_0^t \|K(u)\|_1^2 ds \tag{B.6}$$

which follows from Cauchy-Schwartz and the Poincare inequalities. Also, notice that

$$\langle -f'(u), f_k^2\rangle = \int_D (1-3u^2) f_k^2 \le \|f_k\|^2 \tag{B.7}$$

So, we have

$$J(u_t) \le \|u_0\|^2 + \frac{\mathcal{E}_0 + \mathcal{E}_1}{2}t + M_t - \int_0^t \|\nabla K(u)\|^2 ds \tag{B.8}$$

and

$$J(u_t) - \frac{\mathcal{E}_0 + \mathcal{E}_1}{2}t \le \|u_0\|^2 + M_t - \int_0^t \|K(u)\|_1^2 ds \tag{B.9}$$

Now since

$$M_t - \frac{1}{2}\int_0^t \|K(u)\|_1^2 ds \le M_t - \frac{\lambda_1}{2\mathcal{E}_0}\langle M\rangle_t \tag{B.10}$$

we have,

$$\mathbf{P}\left(\sup_{t\ge 0} J(u_t) - \frac{\mathcal{E}_0+\mathcal{E}_1}{2}t > \|u_0\|^2 + \beta\right) \le \exp\left(-\frac{\lambda_1}{\mathcal{E}_0}\beta\right) \tag{B.11}$$

We can define

$$X_t = \sup_{t\ge 0} J(u_t) - \frac{\mathcal{E}_0+\mathcal{E}_1}{2}t \tag{B.12}$$

and let $B_0 = \|u_0\|^2 + \beta$. Then we can write

$$\mathbb{E}\exp(\gamma X_t) \le \int_0^\infty \exp(\gamma y)\mathbf{P}(X_t \ge y)dy \tag{B.13}$$

$$\le \int_0^{B_0} \exp(\gamma y)dy + \int_0^\infty \exp(\gamma B_0)\exp(\gamma x)\exp(-\frac{\lambda_1}{\mathcal{E}}\beta)dx \tag{B.14}$$

where $x = y - B_0$.This implies that there exists some constant $C_\gamma$ such that

$$\mathbb{E}\exp\left(\gamma \sup_{t\ge 0} J(u_t) - \frac{\mathcal{E}_0+\mathcal{E}_1}{2}t\right) \le C_\gamma \exp(\gamma\|u_0\|^2) \tag{B.15}$$

for all $\gamma \in (0, \frac{\lambda_1}{\mathcal{E}_0}\beta)$. □

# Bibliography


Billingsley, P. (1999), *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, second edn., A Wiley-Interscience Publication.

Coffman, Jr., E. G., Puhalskii, A. A., and Reiman, M. I. (1998), "Polling systems in heavy traffic: a Bessel process limit," *Math. Oper. Res.*, 23, 257–304.

Constantin, P. and Foias, C. (1988), *Navier-Stokes equations*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL.

Da Prato, G. (2006), *An introduction to infinite-dimensional analysis*, Universitext, Springer-Verlag, Berlin, Revised and extended from the 2001 original by Da Prato.

Da Prato, G. and Debussche, A. (1996), "Stochastic Cahn-Hilliard equation," *Nonlinear Anal.*, 26, 241–263.

Da Prato, G. and Debussche, A. (2004), "Absolute Continuity of the Invariant Measures for Some Stochastic PDEs," *Journal of Statistical Physics*, 115, 451–468.

Da Prato, G. and Zabczyk, J. (1992), *Stochastic equations in infinite dimensions*, vol. 44 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge.

Da Prato, G. and Zabczyk, J. (2002), *Second order partial differential equations in Hilbert spaces*, vol. 293 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge.

Da Prato, G., Debussche, A., and Temam, R. (1994), "Stochastic Burgers' equation," *NoDEA Nonlinear Differential Equations Appl.*, 1, 389–402.

Douglas, R. G. (1966), "On majorization, factorization, and range inclusion of operators on Hilbert space," *Proc. Amer. Math. Soc.*, 17, 413–415.

E, W., Mattingly, J. C., and Sinai, Y. (2001), "Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation," *Communications in Mathematical Physics*, 224, 83106, Dedicated to Joel L. Lebowitz.

Embry, M. R. (1973), “Factorization of operators on Banach space,” *Proc. Amer. Math. Soc.*, 38, 587–590.

Goldys, B. and Maslowski, B. (2006a), “Exponential ergodicity for stochastic reaction-diffusion equations,” in *Stochastic partial differential equations and applications—VII*, vol. 245 of *Lect. Notes Pure Appl. Math.*, pp. 115–131, Chapman & Hall/CRC, Boca Raton, FL.

Goldys, B. and Maslowski, B. (2006b), “Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's,” *Ann. Probab.*, 34, 1451–1496.

Hairer, M. and Mattingly, J. C. (2008), “A Theory of Hypoellipticity and Unique Ergodicity for Semilinear Stochastic PDEs,” http://www.citebase.org/abstract?id=oai:arXiv.org:0808.1361.

Jennings, O. B. (2008), “Heavy-Traffic Limits of Queueing Networks with Polling Stations: Brownian Motion in a Wedge,” *Mathematics of Operations Research*, 33, 12–35.

Jennings, O. B. (2010), “Averaging Principles for a Diffusion-Scaled, Heavy-Traffic Polling Station with K Job Classes,” *Mathematics of Operations Research*, 35.

Kuo, H. H. (1975), *Gaussian measures in Banach spaces*, Lecture Notes in Mathematics, Vol. 463, Springer-Verlag, Berlin.

Mattingly, J. C. and Suidan, T. M. (2005), “The small scales of the stochastic Navier-Stokes equations under rough forcing,” *J. Stat. Phys.*, 118, 343–364.

Mattingly, J. C. and Watkins, A. C. (2010), “On Truly Elliptic Stochastic Partial Differential Equations,” In Preparation.

Naylor, A. W. and Sell, G. R. (1982), *Linear operator theory in engineering and science*, vol. 40 of *Applied Mathematical Sciences*, Springer-Verlag, New York, second edn.

Øksendal, B. (2003), *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, sixth edn., An introduction with applications.

Pazy, A. (1983), *Semigroups of linear operators and applications to partial differential equations*, vol. 44 of *Applied Mathematical Sciences*, Springer-Verlag, New York.

Reed, M. and Simon, B. (1972), *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York.

Sell, G. R. and You, Y. (2002), *Dynamics of evolutionary equations*, vol. 143 of *Applied Mathematical Sciences*, Springer-Verlag, New York.

Temam, R. (1997), *Infinite-dimensional dynamical systems in mechanics and physics*, vol. 68 of *Applied Mathematical Sciences*, Springer-Verlag, New York, second edn.

Varadhan, S. R. S. and Williams, R. J. (1985), "Brownian motion in a wedge with oblique reflection," *Comm. Pure Appl. Math.*, 38, 405–443.

# Biography

Andrea Cherese Watkins was born on November 12, 1983 in Detroit, MI. After completing high school in Detroit, she attended Howard University in Washington, DC. In May 2004, Andrea graduated Phi Beta Kappa, Summa Cum Laude from Howard with a B.S. in Mathematics and a minor in Philosophy. She then matriculated to Duke University, with the support of a National Physical Scieneces Consortium Graduate Fellowship. Andrea earned an M.A. and Ph.D. in Mathematics from Duke in 2006 and 2010, respectively.