NEAR EQUILIBRIUM FLUCTUATIONS FOR SUPERMARKET MODELS WITH GROWING CHOICES

SHANKAR BHAMIDI, AMARJIT BUDHIRAJA, AND MIHEER DEWASKAR

Abstract. We consider the supermarket model in the usual Markovian setting where jobs arrive at rate \( n\lambda \) for some \( \lambda_n > 0 \), with \( n \) parallel servers each processing jobs in its queue at rate 1. An arriving job joins the shortest among \( d_n \) randomly selected service queues. We show that when \( d_n \to \infty \) and \( \lambda_n \to \lambda \in (0, \infty) \), under natural conditions on the initial queues, the state occupancy process converges in probability, in a suitable path space, to the unique solution of an infinite system of constrained ordinary differential equations parametrized by \( \lambda \). Our main interest is in the study of fluctuations of the state process about its near equilibrium state in the critical regime, namely when \( \lambda_n \to 1 \). Previous papers e.g. [32] have considered the regime \( d_n/\sqrt{n \log n} \to \infty \) while the objective of the current work is to develop diffusion approximations for the state occupancy process that allow for all possible rates of growth of \( d_n \). In particular we consider the three canonical regimes (a) \( d_n/\sqrt{n} \to 0; \) (b) \( d_n/\sqrt{n} \to c \in (0, \infty) \) and, (c) \( d_n/\sqrt{n} \to \infty \). In all three regimes we show, by establishing suitable functional limit theorems, that (under conditions on \( \lambda_n \)) fluctuations of the state process about its near equilibrium are of order \( n^{-1/2} \) and are governed asymptotically by a one-dimensional Brownian motion. The forms of the limit processes in the three regimes are quite different; in the first case we get a linear diffusion; in the second case we get a diffusion with an exponential drift; and in the third case we obtain a reflected diffusion in a half space. In the special case \( d_n/(\sqrt{n} \log n) \to \infty \) our work gives alternative proofs for the universality results established in [32].

1. Introduction

In this work we study the asymptotic behavior of a family of randomized load balancing schemes for many server systems. Consider a processing system with \( n \) parallel queues in which each queue’s jobs are processed by the associated server at rate 1. Jobs arrive at rate \( n\lambda \) and join the shortest queue amongst \( d_n \) randomly selected queues (without replacement), with \( d_n \in [n] \equiv \{1, \ldots, n\} \). The interarrival times and service times are mutually independent exponential random variables. This queuing system with the above described ‘join-the-shortest-queue amongst chosen queues’ discipline is often denoted as \( JSQ(d_n) \) and frequently referred to as the supermarket model (cf. [14,24,26,28,32] and references therein). Note that when \( d_n = n \) the above description corresponds to a policy where an incoming job joins the shortest of all queues in the system (see e.g. [10]). The case \( d_n = 1 \) is the other extreme corresponding to incoming jobs joining a randomly chosen queue in which case the system is equivalent to one with \( n \) independent \( M/M/1 \) queues with arrival rate \( \lambda_n \) and service rate 1. The case \( d_n = d \) where \( d > 1 \) is a fixed positive integer is sometimes also referred to as the power-of-\( d \) scheme. The analysis of \( JSQ(d_n) \) schemes has been a focus of much recent research motivated by problems from large scale service centers, cloud computing platforms, and data storage and retrieval systems (see e.g. [1,3,8,15,27,33,34,36]). The influential works of Mitzenmacher [30,31] and Vvedenskaya et al. [37] showed by considering a fluid scaling
that increasing \( d \) from 1 to 2 leads to significant improvement in performance in terms of steady-state queue length distributions in that the tails of the asymptotic steady-state distributions decay exponentially when \( d = 1 \) and super-exponentially when \( d = 2 \). Limit theorems under a diffusion scaling for the \( JSQ(d) \) system, with a fixed \( d \), can be found in \citen{7,9}. Although \( JSQ(d) \) for a fixed \( d \geq 2 \) leads to significant improvements over \( JSQ(1) \), as observed in \citen{12,13}, no fixed value of \( d \) provides the optimal waiting time properties of the join-the-shortest-queue system (i.e. \( JSQ(n) \)). See the survey \citen{36} for an overview of the progress in this general area. This motivates the study of asymptotic behavior of a \( JSQ(d) \) system in which the number of choices \( d \) increase with system size, namely \( n \). Such an asymptotic study is the goal of this work.

The paper \citen{32} studied the law of large numbers (LLN) behavior of a \( JSQ(d_n) \) system, under a standard scaling, when \( d_n \to \infty \). The precise result of \citen{32} is as follows. For \( i \in \mathbb{N}_0 = \{ 0, 1, 2, \ldots \} \) and \( t \in [0, \infty) \), let \( G_{n,i}(t) \) denote the fraction of queues with at least \( i \) customers at time \( t \) in the \( n \)-th system. Note that \( G_{n,0}(t) = 1 \) for all \( t \geq 0 \). We will call \( G_n(t) = \{ G_{n,i}(t) : i > 0 \} \) the state occupancy process. This process has sample paths in the space of summable nonnegative sequences. More precisely, for \( p \geq 1 \), let \( \ell_p \) be the space of real sequences \( x = (x_1, x_2, \ldots) \) such that \( \| x \|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p} < \infty \). Let

\[
\ell_1^+ = \left\{ x \in \ell_1 : x_i \geq x_{i+1} \text{ and } x_i \in [0, 1] \text{ for all } i \in \mathbb{N} \right\}
\]

be the space of non-increasing sequences in \( \ell_1 \) with values in \([0, 1] \), equipped with the topology generated by \( \| \cdot \|_1 \). Note that \( \ell_1^+ \) is a closed subset of \( \ell_1 \) and hence is a Polish space. Then, whenever \( \| G_n(0) \|_1 < \infty \) a.s., it can be shown that \( \{ G_n(t) : t \geq 0 \} \) is a stochastic process with sample paths in \( \mathbb{D}([0, \infty) : \ell_1^+) \) (the space of right continuous functions with left limits from \([0, \infty) \) to \( \ell_1^+ \) equipped with the usual Skorohod topology); see Section 3. The paper \citen{32} shows the following two facts under the assumption that \( G_n(0) \) converges in probability to some \( r \in \ell_1^+ \):

(a) When \( d_n = n \) and \( \lambda_n \to \lambda \in (0, \infty) \), \( G_n \) is a tight sequence in \( \mathbb{D}([0, \infty) : \ell_1^+) \) and every weak limit point satisfies a certain set of “fluid limit equations” (see \citen{32} Theorem 5), and equations \ref{2.4}-\ref{2.5} in the current work);

(b) When \( d_n \) is an arbitrary sequence growing to \( \infty \) and \( \lambda_n \to \lambda \in (0, 1) \), then the statements in (a) hold once more for \( G_n \).

The current work begins by revisiting the above LLN results from \citen{32}. In Theorem 2.1 of this work, we show that, when \( G_n(0) \) converges in probability to \( r \), for arbitrary sequences \( d_n \to \infty \) and \( \lambda_n \to \lambda \in (0, \infty) \), \( G_n \) converges in probability in \( \mathbb{D}([0, \infty) : \ell_1^+) \) to a continuous trajectory \( g \) in \( \ell_1^+ \) that is characterized as the unique solution of an infinite system of constrained ordinary differential equations (ODE) (see \ref{2.2} in Proposition 2.1). Using standard properties of the Skorohod reflection map we observe in Remark 2.3 that a continuous trajectory in \( \ell_1^+ \) solves the fluid limit equations of \citen{32} if and only if it solves \ref{2.2}. This together with Proposition 2.1 proves that the fluid limit equations in \citen{32} in fact have a unique solution. In this manner we complete and strengthen the result from \citen{32}. Our proof of the LLN result is quite different from the arguments in \citen{32}. The latter are based on sophisticated ideas of separation of time scales and weak convergence of measure valued processes from \citen{10} to handle the convergence for \( d_n = n \), and certain coupling techniques to treat the general case when \( d_n < n \) and \( d_n \to \infty \). In contrast, our approach is more direct and uses martingale estimates and well known characterization properties of solutions of Skorohod problems (see e.g. proof of Lemma 4.7).

Our main goal in this work is to study diffusion approximations for \( G_n \) in the heavy traffic regime, namely when \( \lambda_n \to 1 \). In the case when \( d_n = n \) (\( JSQ(n) \) system), this problem has been studied in \citen{10}. Their basic result is as follows. Suppose \( d_n = n \) and \( \sqrt{n} (1 - \lambda_n) \to \beta > 0 \). Consider the unit vector \( e_1 = (1, 0, \ldots) \) in \( \ell_2 \). Then under conditions on \( G_n(0) \), the process \( Y_n(\cdot) = \sqrt{n} (G_n(\cdot) - e_1) \) converges in distribution in \( \mathbb{D}([0, \infty) : \ell_2) \) to a continuous stochastic process \( Y = (Y_1, Y_2, \ldots) \), described in terms of a one-dimensional Brownian motion, for which \( Y_i = 0 \) for \( i > r \) for some
$r \in \mathbb{N}$ (which depends on the conditions assumed on $G_n(0)$). Specifically, when $r = 2$, the pair $Y_1, Y_2$ is given as a two-dimensional diffusion in the half space $(-\infty, 0) \times \mathbb{R}$ with oblique reflection in the direction $(-1,1)^t$ at the boundary $\{0\} \times \mathbb{R}$. (For the form of the limit in the general case see Corollary 2.7). In [32] this result is extended to the case where $d_n < n$ and $\frac{d_n}{\sqrt{n \log n}} \to \infty$. Under the same assumptions on the initial condition as in [10], it is shown in [32] that $Y_n$ converges to the same limit process as for the case $d_n = n$. The proof, as for the LLN result, proceeded by constructing a suitable coupling between a JSQ($d_n$) and JSQ($n$) system. The paper [32] also argued that when $\frac{d_n}{\sqrt{n \log n}} \to 0$, the process $Y_n$ cannot be tight and thus in this regime the above diffusion approximation cannot hold.

Our objective in this work is to develop diffusion approximations for $G_n$ in the critical regime (i.e. when $\lambda_n \to 1$ in a suitable manner) that allow for possibly a slower growth of $d_n$ than that permitted by the results in [32]. In fact, in contrast to [10,32], we will prove diffusion limits when $d_n \to \infty$ in an arbitrary manner for choices of $\lambda_n \to 1$ constrained by the exact growth rate of $d_n$. See Table 2.1 for an overview of the regimes of $(\lambda_n, d_n)$ that we cover, along with those covered by previous work. In the special case that $\frac{d_n}{\sqrt{n \log n}} \to \infty$, we will recover the results of [32] with a different proof. In order to motivate the type of limit theorems we seek, we begin by observing that the centering $G_1$ introduced in Definition 2. The fixed point $\mu_1$ which is a stationary point of the fluid limit given in (2.2) with $\lambda = 1$ and thus the results of [10] and [32] give information on fluctuations of the state process $G_n$ about this stationary point. However $e_1$ is not the only stationary point of (2.2) (when $\lambda = 1$) and in fact this ODE has uncountably many fixed points given by $f_k = \sum_{j=1}^k e_j + \gamma e_{k+1} = (1,\ldots,1,\gamma,0,0,\ldots) \in \ell^1_k$ for $k \in \mathbb{N}$ and $\gamma \in [0,1]$, where $e_j$ is the $j$-th unit vector in $\ell_2$ (with 1 at the $j$-th coordinate and zeroes elsewhere). All of these stationary points arise in a natural fashion. Indeed, it turns out that the evolution of the state process $G_n$ can be described via the equation (see Remark 3.1)

$$G_n(t) = G_n(0) + \int_0^t [a_n(G_n(s)) - b(G_n(s))] ds + M_n(t),$$

where $M_n$ is a (infinite dimensional) martingale converging to zero in probability (see Lemma 4.1) and $a_n$, $b$ are certain maps from $\ell^1_k$ to $\ell_2$ (see Remark 3.1 for details). Thus for large $n$, trajectories of $G_n$ will be close to solutions of the infinite dimensional ODE

$$\dot{g}_n = a_n(g_n) - b(g_n)$$

where $g_n$ denotes the derivative of $g_n$. This equation has a unique stationary point $\mu_n$ which is introduced in Definition 2.1. The fixed point $\mu_n$ corresponds to the point in the state space $\ell^1_k$ at which the inflow rate equals the outflow rate in the $n$-th system and thus it is of interest to explore system behavior in the neighborhood of this point. Since $G_n$ is approximated by $g_n$ (over any compact time interval), one can loosen interpret $\mu_n$ as a near fixed point of the state process $G_n$. Furthermore, it can be shown (see Remark 2.3[v]) that, if $d_n \to \infty$ and $\lambda_n \to 1$ in a suitable manner, $\mu_n$ can converge to any specified fixed point $f_1$ of (2.2) and thus every fixed point of (2.2) arises from $\mu_n$ in a suitable asymptotic regime. In order to explore fluctuations of $G_n$ close to different fixed points of (2.2) it is then natural to study the asymptotic behavior of

$$Z_n(t) = \sqrt{n}(G_n(t) - \mu_n), \quad t \geq 0. \quad (1.2)$$

We note that in the regime considered in [32] where $\frac{d_n}{\sqrt{n \log n}} \to \infty$ and $\sqrt{n}(1 - \lambda_n) \to \alpha > 0$, $\sqrt{n}(e_1 - \mu_n) \to \alpha e_1$ and so in this case the asymptotic behavior of $Z_n$ can be read off from that of $Y_n$ (see Corollary 2.7 and Remark 2.8[v]). However in general $\sqrt{n}(e_1 - \mu_n)$ (and more generally, $\sqrt{n}(f_k - \mu_n)$) may not be bounded and so the asymptotic behavior of $Z_n$ and $Y_n$ may be very different.

In this work we obtain limit theorems for $Z_n$ as $d_n \to \infty$ in an arbitrary fashion and $\lambda_n \to 1$ in a suitable manner. Specifically in Theorems 2.2, 2.3 and 2.4 we consider the three cases:
(a) $d_n/\sqrt{n} \to 0$; (b) $d_n/\sqrt{n} \to c \in (0, \infty)$ and, (c) $d_n/\sqrt{n} \to \infty$, respectively. In all three regimes we consider initial conditions $G_n(0)$ such that for some $r \in \mathbb{N}$, $\sqrt{n}(G_n,j(0) - \mu_{n,j})$ converge to zero in probability for all $j > r$ and in each case (under conditions on $\lambda_n$) we obtain a limit process driven by a one-dimensional Brownian motion with continuous sample paths in $\ell_2$ which has all but finitely many coordinates 0. In particular, when $r = 2$ in the second and the third case and $r = k + 2$ for some $k \in \mathbb{N}$ in the first case (and $d_n$, $\lambda_n$ depend on $k$ in a suitable fashion), one can describe the limit through a two-dimensional diffusion driven by a one-dimensional Brownian motion. The form of this two-dimensional process in the three regimes is quite different; in the first case we get a linear diffusion (i.e. the drift is of the form $b(y) = Ay$ for, $y \in \mathbb{R}^2$ and some $2 \times 2$ matrix $A$); in the second case we get a diffusion with an exponential drift; and in the third case we obtain a reflected diffusion in the half space $(-\infty, \alpha] \times \mathbb{R}$ for some $\alpha \geq 0$.

Although the limit processes in Theorems 2.2 and 2.3 are quite different from those obtained in [9] and [32], the limit in Theorem 2.2 has a similar form (in that it is a reflected diffusion in a half space) as in the above papers. However here as well there are some differences. In particular, depending on how $\lambda_n$ approaches 1, the reflection occurs at a different barrier $\alpha \in (0, \infty)$; in fact $\alpha = \infty$ is possible as well in which case there is no reflection. Furthermore, recall that $Z_n$ is defined by centering about $\mu_n$. In general $\sqrt{n}(\mu_n - e_1)$ will diverge and thus the process $Y_n$ considered in the above cited papers may not converge in this regime. However, as noted previously, when $d_n$ grows sufficiently fast, namely $d_n/\sqrt{n} \log n \to \infty$ the process $Y_n$ will indeed converge and in that case we recover the result in [32] (in fact a slight strengthening in that the drift parameter in Corollary 2.7 is allowed to be 0). In addition Theorem 2.4 also covers the case $d_n/\sqrt{n} \log n \to c \in (0, \infty)$ and situations where $\lambda_n = 1 + O(n^{-1/2})$ (see Remark 2.8 (iv)). In such settings, once more both $Z_n$ and $Y_n$ converge and the limit of the latter has the same form as in [10,32].

As is observed in Remarks 2.6 and 2.8 under conditions of Theorem 2.3 or Theorem 2.4 $\mu_n$ must converge to the fixed point $e_1 = f_1^\dagger$. In contrast, Theorem 2.2 allows for a range of asymptotic behavior for $\mu_n$. In particular, under the conditions of the theorem, with suitable $\lambda_n, d_n$, $\mu_n$ can converge to the fixed point $f_k^\dagger$ for an arbitrary $k \in \mathbb{N}$ (see [5] for a similar observation). Here $k$ may then be considered as the average time spent by a job in the system, since asymptotically almost all (c.f. [5]) queues will have length $k$ under these conditions. In such a setting the first $k - 1$ coordinates of the limit process are essentially 0 (see Theorem 2.2 for a precise statement) and the $k$-th coordinate is the first one to exhibit stochastic variability. Thus a rather novel asymptotic behavior for the $JSQ(d_n)$ system emerges when $d_n$ approaches $\infty$ at significantly slower rates than those considered in [32] and $\lambda_n$ approach 1 in a suitable manner (in relation to $d_n$).

1.1. Organization of the paper. Section 2 contains all our main results. The remaining Sections starting with Section 3 contain proofs of the main results.

We now make some comments on the proofs of Theorems 2.2 - 2.4. The starting point is a convenient semimartingale representation for the centered state process $Z_n$ in (6.1). In the study of the behavior of the drift term in this decomposition, an important ingredient is an analysis of the asymptotic properties of the near fixed point $\mu_n$, and the asymptotic behavior of the function $\beta_n$ (see Definition 1) in $O(n^{-1/2})$ sized neighborhoods around the coordinates of $\mu_n$. This behavior, which is different in the three regimes considered above, determines the asymptotics of the drift $A_n(Z_n(s)) - b(Z_n(s))$, where $A_n(z) = \sqrt{n}(a_n(\mu_n + n^{-1/2})z - a_n(\mu_n))$. Properties of $\mu_n$ are also key in arguing that, in all three cases, under our conditions, $(Z_n,r+1, \ldots)$ converges to 0 in probability in $\mathbb{D}([0, \infty) : \ell_2)$ (see Lemma 3.4). The rest of the work is in characterizing the asymptotics of the finite dimensional process $(Z_{n,1}, \ldots, Z_{n,r})$. For this study, the three regimes require different approaches. In particular, Theorem 2.2 hinges on a detailed understanding of the asymptotic behavior of a tridiagonal matrix function $Q_n(s)$ (see e.g. Lemmas 7.4 and 7.6); Theorem 2.3 requires an analysis of a stochastic differential equation with an exponential drift term (in particular the drift does not satisfy the usual growth conditions); and Theorem 2.4 is based on a careful study of excursions.
of the prelimit processes above the limiting reflecting barrier and properties of Skorohod maps in order to characterize the reflection properties of the limit process.

1.2. Notation and setup. For \( m \geq 1 \), let \([m] = \{1, 2, \ldots, m\}\). We will denote finite-dimensional vectors in \( \mathbb{R}^m \) as \( \vec{x}, \vec{y} \), etc. and \( \langle \vec{x}, \vec{y} \rangle \) will denote the standard inner-product. Transpose of a vector \( \vec{v} \) will be written as \( \vec{v}^T \). The standard basis vectors in \( \mathbb{R}^m \) will be denoted by \( e_i \) for \( i = 1, 2, \ldots, m \). Also, \( \| \vec{x} \| = \sqrt{\langle \vec{x}, \vec{x} \rangle} \) will denote the usual Euclidean norm.

We will often use bold symbols such as \( \mathbf{x} = (x_1, x_2, \ldots) \) to denote a infinite dimensional vector or function. For \( p \in \{1, 2, \ldots, \infty\} \), let \( \| \mathbf{x} \|_p = \left( \sum_{i=1}^\infty |x_i|^p \right)^{1/p} \) denote the \( p \)-norm on the space of infinite sequences and \( \ell_p = \{ \mathbf{x} \in \mathbb{R}^\infty \mid \| \mathbf{x} \|_p < \infty \} \). Let \( \ell_1 \) be as in \((1.1)\), which is a Polish space under \( \| \cdot \|_1 \).

For \( k \in \mathbb{N} \), let \( f_k = (1, 1, \ldots, 1, 0, 0 \ldots) \in \ell_1^k \) denote the vector with first \( k \) indices equal to 1, and \( e_k = (0, \ldots, 0, 1, 0 \ldots) \in \ell_1 \) denote the vector with 1 in the \( k \)th coordinate. For any \( k \in \mathbb{N} \) and \( \gamma \in (0, 1) \) write \( f_k^\gamma = f_k + \gamma e_{k+1} \in \ell_1^k \). For \( z = (z_1, z_2, \ldots) \in \mathbb{R}^\infty \) and \( r \in \mathbb{N} \), let \( z_{r+} = (z_{r+1}, z_{r+2}, \ldots) \in \mathbb{R}^\infty \) denote the vector shifted by \( r \) steps. Similar notation will be used for functions and processes with values in \( \mathbb{R}^\infty \).

For a Polish space \( S \) and the interval \( I = [0, T] \) for \( T > 0 \) or \( I = [0, \infty) \), denote by \( \mathcal{C}(I : S) \) (resp. \( \mathcal{D}(I : S) \)) the space of functions (resp. right continuous functions with left limits) from \( I \) to \( S \), endowed with the topology defined by uniform convergence on compact sets (resp. Skorokhod topology). For \( h \in \mathcal{D}([0, T] : \mathbb{R}) \), \( g \in \mathcal{D}([0, T] : \ell_p) \) and \( t \in [0, T] \), denote the size of the largest jump up to time \( t \) by \( J_t(h) = \sup_{s \in [0, t]} |h(s) - h(s^-)| \) and \( J_t(g) = \sup_{s \in [0, t]} \| g(s) - g(s^-) \|_p \), and the supremum norms up to time \( t \) by \( |h|_{t^*} = \sup_{s \in [0, t]} |h(s)| \) and \( |g|_{p,t} = \sup_{s \in [0, t]} \| g(s) \|_p \). If \( h \) is absolutely continuous on \([0, T]\), then \( \dot{h}(t) \) (or sometimes \( dh(t)/dt \)) will denote the derivative of \( h \) at \( t \in [0, T] \) (defined almost everywhere).

We will use \( \mathbb{I}_{\{\text{cond}\}} \) to denote the indicator function that takes the value 1 if \( \text{cond} \) is true, otherwise it takes the value 0. We will denote by \( \mathbf{id} \) the identity map, \( \mathbf{id}(t) = t \), on \([0, T] \) or \([0, \infty) \). We use \( \mathbb{P} \) and \( \mathbf{E} \) to denote the probability and expectation operators, respectively. For \( x, y \in \mathbb{R} \), \( x \wedge y \) denotes the minimum and \( x \vee y \) the maximum of \( x \) and \( y \) respectively. For any \( x \in \mathbb{R} \), \( x^+ = x \vee 0 \) and \( x^- = (-x) \vee 0 \). We use \( \Rightarrow \) and \( \Rightarrow_\mathbb{P} \) to denote convergence in probability and convergence in distribution respectively on an appropriate Polish space which will depend on the context. For a sequence of real valued random variables \( \{X_n, n \geq 1\} \), we write \( X_n = o_p(b_n) \) when \( |X_n|/b_n \Rightarrow 0 \) as \( n \to \infty \). For non-negative functions \( h(\cdot), g(\cdot) \), we write \( h(n) = O(g(n)) \) when \( h(n)/g(n) \) is uniformly bounded, and \( h(n) = o(g(n)) \) (or \( h(n) \ll g(n) \)) when \( \lim_{n \to \infty} h(n)/g(n) = 0 \). We write \( h(n) \sim g(n) \) if \( h(n)/g(n) \to 1 \) as \( n \to \infty \).

2. Main Results

Recall the process \( G_n \) from Section 1. Our first result gives a law of large numbers (LLN) for the process \( G_n \) as \( n \to \infty \). In order to state this result we begin by recalling the one-dimensional Skorohod map (cf. [18] Section 3.6.C], [19]) with reflecting barrier at \( \alpha \in \mathbb{R} \). For \( \alpha \in \mathbb{R} \) and \( h \in \mathcal{D}([0, \infty) : \mathbb{R}) \) with \( \bar{h}(0) \leq \alpha \), define \( \Gamma_\alpha(h), \hat{\Gamma}_\alpha(h) \in \mathcal{D}([0, \infty) : \mathbb{R}) \) as

\[
\Gamma_\alpha(h)(t) = h(t) - \sup_{s \in [0, t]} (h(s) - \alpha)^+, \quad \hat{\Gamma}_\alpha(h)(t) = \sup_{s \in [0, t]} (h(s) - \alpha)^+.
\]  

The map \( \Gamma_\alpha \) (and sometimes the pair \((\Gamma_\alpha, \hat{\Gamma}_\alpha)\)) is referred to as the one-dimensional Skorohod map (with reflection at \( \alpha \)). We note that the above map is a modification of the usual definition to account for the fact that in our case reflection occurs from above (in order to prevent \( h \) from exceeding the level \( \alpha \)). The following well-posedness result, which is proved in Section 4, will be used to characterize the LLN limit of \( G_n \).
**Proposition 2.1.** Fix $r \in \ell_1^d$. Then there is a unique $(g, v) \in \mathbb{C}([0, \infty) : \ell_1^d \times \ell_\infty)$ that solves the following system of equations

$$(g_i(t), v_i(t)) = (\Gamma_1, \hat{\Gamma}_1) \left( r_i - \int_0^t (g_i(s) - g_{i+1}(s)) ds + v_{i-1}(t) \right)(t) \quad \forall i \geq 1, t \geq 0 \quad (2.2)$$

where $v_0(t) = \lambda t$ for all $t \geq 0$.

**Remark 2.2.** Using the well known characterization of a one-dimensional Skorohod map, one can alternatively characterize $(g, v)$ as the unique pair in $\mathbb{C}([0, \infty) : \ell_1^d \times \ell_\infty)$ such that $v_i$ is nondecreasing,

$$g_i(t) = r_i - \int_0^t (g_i(s) - g_{i+1}(s)) ds + v_{i-1}(t) - v_i(t) \quad \forall i \geq 1 \quad (2.3)$$

and $v_0(t) = \lambda t$, for all $t > 0$ and $v_i(0) = 0$ for each $i \geq 0$.

We can now present the LLN result. The proof is given in Section 4.

**Theorem 2.1.** Let $r \in \ell_1^d$. Suppose that $G_n(0) \xrightarrow{P} r$ in $\ell_1^d$, $\lambda_n \rightarrow \lambda$ and $d_n \rightarrow \infty$, as $n \rightarrow \infty$. Then $G_n \rightarrow g$ in probability in $\mathbb{D}([0, \infty) : \ell_1^d)$ as $n \rightarrow \infty$, where $(g, v) \in \mathbb{C}([0, \infty) : \ell_1^d \times \ell_\infty)$ is the unique solution of (2.2).

**Remark 2.3.** Note that Theorem 2.1 allows $d_n \rightarrow \infty$ in an arbitrary manner. The Skorokhod reflection term $v_i$ in (2.3), which increases only at time instants $t$ when $g_i(t) = 1$, prevents $g_i$ from exceeding the level 1. It arises as a result of the simple fact that an arriving job cannot join a queue of length $i - 1$ when all the queues in the selection are of length $i$ or more. In [32, Theorem 1] it is shown that, under the assumptions of Theorem 2.1, $G_n$ is a tight sequence of $\mathbb{D}([0, \infty) : \ell_1^d)$ valued random variables and that every subsequential weak limit $\hat{g}$ satisfies a system of equations given as

$$\hat{g}_i(t) = r_i - \int_0^t (\hat{g}_i(s) - \hat{g}_{i+1}(s)) ds + \int_0^t p_{i-1}(\hat{g}(s)) ds \quad \text{for } i \geq 1 \quad (2.4)$$

where

$$p_j(\hat{g}(s)) = \begin{cases} 
\lambda - (\lambda - 1 + \hat{g}_{j+2}(s))^+ & \text{if } j = m(\hat{g}(s)) - 1 \\
(\lambda - 1 + \hat{g}_{j+1}(s))^+ & \text{if } j = m(\hat{g}(s)) > 0 \\
\lambda & \text{if } j = m(\hat{g}(s)) = 0 \\
0 & \text{otherwise,} 
\end{cases} \quad (2.5)$$

and for $x \in \ell_1^d$, $m(x) = \inf\{i \mid x_{i+1} < 1\}$. (Note that $m(G_n(t))$ is the length of the smallest queue at time $t$.) The uniqueness of the above system of equations was not shown in [32].

From (2.2) and the definition in (2.1) it follows that each $v_i$ is absolutely continuous and, for a.e. $t$,

$$\frac{dv_i(t)}{dt} = \left( \frac{dv_{i-1}(t)}{dt} - g_i(t) + g_{i+1}(t) \right)^+ \mathbb{I}_{\{g_i(t) = 0\}}$$

for any $i \geq 1$. From this we see that, for a.e. $t$,

$$\frac{dv_i(t)}{dt} = \begin{cases} 
\lambda & \text{if } i = 0 \\
\frac{dv_{i-1}(t)}{dt} - 1 + g_{i+1}(t) & \text{if } i = m(g(t)) \text{ and } i \geq 1, \\
0 & \text{if } i > m(g(t)) 
\end{cases} \quad (2.6)$$
and consequently $p_j(g(s)) = \frac{dv_j(s)}{ds} - \frac{dv_{j+1}(s)}{ds}$ for a.e. $s$. Substituting this back in (2.3) shows that $g$ solves the system of equations in (2.4). Conversely, for any solution $\tilde{g}$ of (2.4), defining $\tilde{\psi}$ by the right side of (2.6) by replacing $g$ with $\tilde{g}$, we see that $(\tilde{g}, \tilde{\psi})$ solves (2.3). From the uniqueness result in Lemma 2.1 it then follows that in fact there is only one solution to the system of equations in (2.4) and this solution equals $g$ given in (2.2).

Consider now the time asymptotic behavior of $g$ given in (2.2). When $\lambda < 1$, $(\lambda, 0, 0 \ldots) \in \ell_\infty$ is the unique fixed point of (2.2), as can be seen by setting the derivative of the right side of (2.4) to 0. In the critical case, i.e. when $\lambda = 1$, the situation is very different and in fact there are uncountably many fixed points given by the collection $\{f \in \ell_1^\infty \mid m(f) > 0, f_m(f) + 2 = 0\} = \{f_k^g \mid k \in \mathbb{N}, \gamma \in [0, 1]\}$, which once more is seen by checking that the derivative on the right side of (2.4) is 0 at exactly these points when $\lambda = 1$. In this work we are interested in the fluctuations of $G_n$ in the critical case when the system starts suitably close to one of the fixed points of (2.3). Thus for the remaining section we will assume that $\lambda_n < 1$ for every $n$ and $\lambda_n \to 1$ as $n \to \infty$. In order to formulate precisely what we mean by ‘suitably close to the fixed point’ we need some definitions and notation. The functions $\beta_n$ in the next definition will play a central role.

**Definition 1.** Given $d_n \in [n]$, define the function $\beta_n : [0, 1] \to [0, 1]$ by

$$\beta_n(x) = \prod_{i=0}^{d_n-1} \left(1 - \frac{i}{n}\right)^{x \cdot \beta_n(i)}.$$  (2.7)

The function $\beta_n(\cdot)$ arises when sampling $d_n$ random servers without replacement. Specifically, when $nx \in \mathbb{N}$, $\beta_n(x) = P(\hat{A}_{n,d_n} \subseteq [nx]) = (\binom{n}{d_n})/\binom{n}{x}$, where $\hat{A}_{n,d_n}$ is a randomly chosen subset (without replacement) from $[n]$ of size $d_n$. Here we adopt the convention that $\binom{n}{m} = 0$ when $m < r$. An alternative is to perform sampling with replacement, which corresponds to the simpler function $\gamma_n(x) = x^{d_n}$ in place of $\beta_n$.

We now introduce the notion of a ‘near fixed point’ of $G_n$.

**Definition 2.** For $n \in \mathbb{N}$, the near fixed point $\mu_n$ of $G_n$ is the vector in $\ell_1^\infty$ given as $\mu_n = (\mu_{n,1}, \mu_{n,2}, \ldots)$ where $\mu_{n,i}$ are defined recursively as $\mu_{n,1} = \lambda_n$ and $\mu_{n,i+1} = \lambda_n \beta_n(\mu_{n,i})$ for $i \geq 1$.

Using $\beta_n(x) \leq x^{d_n} \leq x$ and $\lambda_n < 1$, it is easy to check that $\mu_n \in \ell_1$. The reason $\mu_n$ is referred to as a near fixed point of $G_n$ is discussed in Remark 3.1. To study the fluctuations of the process around the near fixed point $\mu_n$, we define the centered and scaled process, $Z_n$ as in (1.2). We now present our three main results on fluctuations which correspond to the three cases $d_n/\sqrt{n} \to 0$, $d_n/\sqrt{n} \to c \in (0, \infty)$, and $d_n/\sqrt{n} \to \infty$ respectively. In each of these cases we will assume that the initial configuration starts sufficiently close to the near equilibrium point $\mu_n$.

**Assumption 2.4.** Suppose that $\{\|Z_n(0)\|\}_{n \in \mathbb{N}}$ is tight and $Z_n(0) \overset{P}{\to} z$ in $\ell_2$, where $z_{r+} = 0$ for some $r \in \mathbb{N}$.

In the following, $\tilde{\beta}_n(x)$ is as defined in (5.1) and in the convention noted below (5.1). In particular, for $x \in (0, 1) \setminus \{\frac{d_n-1}{n}\}$, $\tilde{\beta}_n(x)$ is the derivative of $\beta_n$ at $x$.

**Theorem 2.2.** Suppose that, as $n \to \infty$, $1 < d_n < \sqrt{n}$, $\lambda_n \not\to 1$, and there is a $k \in \mathbb{N}$ so that $\mu_{n,k} \to 1$ and $\tilde{\beta}_n(\mu_{n,k}) \to \alpha \in [0, \infty)$ as $n \to \infty$. Further suppose that Assumption 2.4 holds for some $r > k$. Then for any $T \in (0, \infty)$,

$$\lim_{M \to \infty} \sup_n P\left(\|Z_n\|_{2,T} > M\right) = 0.$$  \hspace{1cm} (2.8)

Furthermore, if $k > 1$, then $\sup_{t \in \epsilon, T} |Z_{n,i}(t)| \overset{P}{\to} 0$ as $n \to \infty$ for any $T < \infty$, $0 < \epsilon \leq T$ and $i \in [k-1]$. 


Consider the shifted process \( Y_n(t) \equiv (\sum_{i=1}^{k} Z_{n,i}(t), Z_{n,k+1}(t), Z_{n,k+2}(t), \ldots) \) and \( y \equiv (\sum_{i=1}^{k} z, z_{k+1}, z_{k+2}, \ldots) \). Then \( Y_n \Rightarrow Y \) in \( \mathbb{D}([0, \infty) : \ell_2) \), where \( Y \in \mathbb{C}([0, \infty) : \ell_2) \) is the unique pathwise solution to

\[
Y_1(t) = y_1 - (\alpha + \bar{1}_{(k=1)}) \int_0^t Y_1(s)ds + \int_0^t Y_2(s)ds + \sqrt{2}B(t)
\]

\[
Y_2(t) = y_2 + \alpha \int_0^t Y_1(s)ds - \int_0^t Y_2(s)ds + \int_0^t Y_3(s)ds
\]

\[
Y_i(t) = y_i - \int_0^t Y_i(s)ds + \int_0^t Y_{i+1}(s)ds \quad \text{for } i \in \{3, \ldots, r - k + 1\}
\]

\[
Y_i(t) = 0 \quad \text{for } i > r - k + 1,
\]

and \( B(\cdot) \) is a one-dimensional standard Brownian motion.

**Remark 2.5.**

(i) Note that the convergence \( \sup_{t \in [k, T]} |Z_{n,i}(t)| \xrightarrow{P} 0 \) as \( n \to \infty \) for any \( 0 < \epsilon < T \) is equivalent to the statement that \( Z_{n,i} \to 0 \) in probability in \( \mathbb{D}((0, T] : \mathbb{R}) \) where the latter space is equipped with the topology of uniform convergence on compacts. Note also that, since Theorem 2.2 allows \( Z_{n,i}(0) \) to converge to a non-zero limit, the above convergence to 0 cannot be strengthened to a convergence in probability in \( \mathbb{D}([0, T] : \mathbb{R}) \).

(ii) By Corollary 5.3 in Section 5 when \( \mu_{n,k} \) is away from 0,

\[\hat{\beta}_n(\mu_{n,k}) = (1 + o(1)) \frac{d_n \mu_{n,k+1}}{\lambda_n \mu_{n,k}}\]

as \( n \to \infty \). Hence the assumptions \( d_n \to \infty \), \( \lambda_n \to 1 \), \( \mu_{n,k} \to 1 \) and \( \hat{\beta}_n(\mu_{n,k}) \to \alpha < \infty \) in Theorem 2.2 say that \( \mu_{n,k+1} \to 0 \). Since \( \mu_{n,k} \to 1 \), this in fact shows that \( \mu_n \to f_k \) in \( \ell_1^k \), where recall that \( f_k \) is one of the fixed points of the fluid-limit (2.2) when \( \lambda = 1 \). The fact that the convergence happens in \( \ell_1^k \) can be seen on observing that if \( \mu_{n,k+1} \leq \epsilon \) then, by (5.2), \( \mu_{n,k+1+i} \leq \epsilon d_n \).

We also note that in general \( \sqrt{n}(\mu_n - f_k) \) will diverge, and thus \( \sqrt{n}(G_n - f_k) \) will typically not be tight, in this regime. Nevertheless it may still be interesting to study the behavior of \( n^\alpha(G_n - f_k) \) for some \( \alpha \in (0, 1/2) \) and appropriate choices of \( d_n \to \infty \) and \( \lambda_n \to 1 \). Note however that when \( \alpha \in (0, 1/2) \), the martingale term in the semimartingale decomposition of \( n^\alpha(G_n - f_k) \) will converge to zero (as can be seen from the convergence observed below (6.6)) and thus the limit behavior is expected to be quite different. We leave this for future work.

(iii) In the special case when the system starts sufficiently close to the near fixed point \( \mu_n \) so that \( z_i = 0 \) for \( i > k + 1 \), the limit process \( Y \) simplifies to an essentially two-dimensional process given as, \( Y_i(t) = 0 \) for \( i > 2 \), and

\[
Y_1(t) = y_1 - (\alpha + \bar{1}_{(k=1)}) \int_0^t Y_1(s)ds + \int_0^t Y_2(s)ds + \sqrt{2}B(t)
\]

\[
Y_2(t) = y_2 + \alpha \int_0^t Y_1(s)ds - \int_0^t Y_2(s)ds
\]

(iv) The convergence behavior of \( Z_n \) is governed by the sequence of parameters \( (d_n, \lambda_n) \). In Corollary 5.5 from Section 5 we show that if \( 1 \ll d_n^{k+1} \ll n \) and \( 1 - \lambda_n = \frac{\xi_n + \log d_n}{d_n^2} \) with \( \xi_n \to -\log(\alpha) \in (-\infty, \infty) \) and \( \frac{c^2}{d_n} \to 0 \), then the conditions \( \mu_{n,k} \to 1 \) and \( \hat{\beta}_n(\mu_{n,k}) \to \alpha \in [0, \infty) \) of Theorem 2.2 are satisfied. Using this fact we make the following observations. For simplicity, consider \( z = 0 \).
(a) Suppose that $d_n = \log n$, $1 - \lambda_n = \frac{\log \log n}{(\log n)^c}$. In this case the assumptions of Theorem 2.2 are satisfied and one essentially sees non-zero fluctuations only in the $k$-th and $k + 1$-th coordinates. Note that as $k$ becomes large, the traffic intensity increases and one sees more and more coordinates of the near fixed point approach 1.

(b) With the same $d_n$ as in (a) but a somewhat lower traffic intensity given as $1 - \lambda_n = \frac{\log \log n}{(\log n)^{1/2 - \epsilon}}$ for some $\epsilon \in (0, 1/2)$, one sees that condition of the theorem are satisfied with $\alpha = 0$ (i.e. $\beta_n(\mu_n, k) \to 0$). Thus the limit process $Y_i$, in the case $k > 1$, simplifies to $Y_i = 0$ for $i > 1$ and $Y_1(t) = \sqrt{2}B(t)$. When $k = 1$, $Z_1 = Y_1$ is instead given as the following Ornstein-Uhlenbeck (OU) process

$$Z_1(t) = -\int_0^t Z_1(s)ds + \sqrt{2}B(t). \quad (2.10)$$

(c) With higher values of $d_n$, using Theorem 2.2 one can analyze fluctuations for systems with higher traffic intensity. For example, suppose that $d_n = \frac{\log n}{\log \log n}$. Then the conditions of the theorem are satisfied with $k = 1$ and $1 - \lambda_n \sim (\log n)^2/\sqrt{n}$. In fact in this case $\alpha = 0$ and the limit process is described by the one-dimensional OU process (2.10). With a slightly higher traffic intensity given as $1 - \lambda_n = ((\log n)^2 - 2\log n \log \log n)/2\sqrt{n}$ one obtains a two-dimensional limit diffusion.

(d) The theorem allows for traffic intensity in the Halfin-Whitt scaling regime (i.e. $\sqrt{n}(1 - \lambda_n) \to 0$) as well. Specifically, for $k \geq 2$, if $d_n = (\sqrt{n} \log n)^{-2}$ and $(1 - \lambda_n) = \frac{\beta - \alpha(1)}{\sqrt{n}}$ for some $\beta > \beta_0 = 1/2k$, the conditions of the theorem are satisfied with $\alpha = 0$. With slightly higher traffic intensity (e.g. $\beta + o(1)$ replaced by $\beta_0 + (\frac{1}{k} \log \log n - \log \alpha)/\log n$) conditions of the theorem are met with a non-zero $\alpha$.

(e) More generally, suppose we are interested in studying the fluctuation behavior when the traffic intensity is $\lambda_n = 1 - \gamma n^{-a}$ for some $a \in (0, 1)$ and $\gamma > 0$. The cases $a < 1/2$ and $a > 1/2$ correspond to the so-called sub and super Halfin-Whitt regimes, respectively. The asymptotic behavior of $JSQ(d_n)$ schemes in steady state in these regimes has been studied in [5, 6, 22, 23]. In [5, 6], the authors prove the following: suppose $d_n = n^b$ for some $b \in (0, 1]$ that satisfies $a/b \notin \mathbb{N}$ and $2a < 1 + b(k - 1)$ where $k = \lceil a/b \rceil$; then with high probability in equilibrium, the largest queue will have length $k$ and a vanishingly small fraction of queues have length smaller than $k$. In [22, 23], the authors consider the case $a \in (0, 1/2)$ and show that for the $JSQ(d_n)$ system with buffer size $b_n = O(\log n)$, in equilibrium, both the expected waiting time per job and the probability that a job is routed to an non-idle server are $O(b_n n^{-r(1/2-a)})$, whenever $d_n \geq \frac{\log n}{\gamma n^a \log n}$ for any positive integer $r \leq \frac{\log n}{\gamma n^a \log n}$.

In the current work we study the behavior of $JSQ(d_n)$ over finite intervals of time. Our results, including Theorem 2.2 allow for both sub and super Halfin-Whitt regimes. To see this, choose any $a \in (0, 1)$, $\nu > 1$, an integer $k > a/(1 - a)$, and let $b = a/k$. (In other words, $a$ satisfies $2a < 1 + b(k - 1)$, which is the same condition as in [5]). Then Theorem 2.2 holds with $k$, $d_n = n^b(\nu/\gamma \log n)^{1/k}$ and $\alpha = 0$.

(f) Recall that a fixed point of (2.2) when $\lambda = 1$ takes the form $f^k = f_k + \gamma e_{k+1} = (1, \ldots, \gamma, 0, \ldots) \in \ell_1^k$, where $k \in \mathbb{N}$ and $\gamma \in [0,1)$. Although Theorem 2.2 only considers settings where the near fixed point $\mu_n$ converges to $f_k = f_k$ for some $k$, it is possible to give conditions under which $\mu_n$ converges to a different fixed point. Specifically, suppose that $1 < d_n^{k+1} < n$ and $1 - \lambda_n = \frac{a}{\alpha \gamma}$ for some $a > 0$. Then it can be checked using Lemma 5.4 that $\mu_n \to f^k_n$ with $\gamma = e^{-a}$. However proving fluctuation results in this regime appears to be technically more involved, and we leave it for future work.

The next theorem describes the fluctuations of $Z_n$ when $d_n$ is of order $\sqrt{n}$.
Theorem 2.3. Suppose that \( \frac{d_n}{\sqrt{n}} \to c \in (0, \infty) \) and \( \lambda_n = 1 - \left( \frac{\log d_n}{d_n} + \frac{\alpha_n}{\sqrt{n}} \right) \) with \( \alpha_n \to \alpha \in (-\infty, \infty] \) and \( \alpha_n = o(n^{1/4}) \). Then, \( \mu_n \to f_1 \) in \( \ell_2^1 \). Suppose further that Assumption 2.4 holds for some \( r \geq 2 \). Then, as \( n \to \infty \), \( Z_n \Rightarrow Z \) in \( \mathcal{D}([0, \infty) : \ell_2) \), where \( Z \) is the unique pathwise solution to:

\[
Z_1(t) = z_1 - \int_0^t (Z_1(s) - Z_2(s))ds - (ce^{c\alpha})^{-1} \int_0^t (e^{cZ_1(s)} - 1)ds + \sqrt{2}B(t),
\]
\[
Z_2(t) = z_2 - \int_0^t (Z_2(s) - Z_3(s))ds + (ce^{c\alpha})^{-1} \int_0^t (e^{cZ_1(s)} - 1)ds,
\]
\[
Z_i(t) = z_i - \int_0^t (Z_i(s) - Z_{i+1}(s))ds \quad \text{for each } i \in \{3 \ldots r\},
\]
\[
Z_i(t) = 0 \quad \text{for each } i > r,
\]
and \( B \) is standard Brownian motion.

Remark 2.6.

(i) Note that the coefficients in the above system of equations are only locally Lipschitz and have an exponential growth. However since \( c \) is positive, the system of equations has a unique pathwise solution as is shown in Lemma 8.2.

(ii) Once more, when \( z_i = 0 \) for all \( i \geq 2 \), the system of equations simplifies to a two-dimensional system given as \( Z_i = 0 \) for all \( i \geq 2 \), and

\[
Z_1(t) = z_1 - \int_0^t (Z_1(s) - Z_2(s))ds - (ce^{c\alpha})^{-1} \int_0^t (e^{cZ_1(s)} - 1)ds + \sqrt{2}B(t),
\]
\[
Z_2(t) = z_2 - \int_0^t Z_2(s)ds + (ce^{c\alpha})^{-1} \int_0^t (e^{cZ_1(s)} - 1)ds.
\]

(iii) In the regime considered in Theorem 2.3, the near fixed point \( \mu_n \) can converge to only one particular fixed point of (2.2), namely \( f_1 \). As before, the term \( \sqrt{n}(\mu_n - f_1) \) may diverge and thus \( \sqrt{n}(G_n(\cdot) - f_1) \) will in general not be tight.

(iv) Suppose that \( d_n = c\sqrt{n} \) for some \( c > 0 \), \( z = 0 \) and \( 1 - \lambda_n = (\beta + o(1)) \log n/\sqrt{n} \) for some \( \beta > \beta_0 = 1/2c \). Then the assumptions of the above theorem are satisfied with \( \alpha = \infty \) and the limit system simplifies to a one-dimensional OU process given as \( Z_i = 0 \) for all \( i > 1 \), and \( Z_1 \) satisfies (2.10). If \( (\beta + o(1)) \log n \) is replaced by \( \beta_0 \log n + \gamma \) for some \( \gamma \in \mathbb{R} \), we instead obtain a two-dimensional limit system given as \( Z_i = 0 \) for all \( i > 2 \), and

\[
Z_1(t) = -\int_0^t (Z_1(s) - Z_2(s))ds - e^{-c\gamma} \int_0^t (e^{cZ_1(s)} - 1)ds + \sqrt{2}B(t),
\]
\[
Z_2(t) = -\int_0^t Z_2(s)ds + e^{-c\gamma} \int_0^t (e^{cZ_1(s)} - 1)ds.
\]

Finally we consider the fluctuation behavior when \( d_n \gg \sqrt{n} \). This time the limit system will involve reflected diffusion processes. Recall from (2.1) the definition of the Skorohod maps \( \Gamma_\alpha \) and \( \hat{\Gamma}_\alpha \) associated with a reflection barrier at \( \alpha \in \mathbb{R} \). We will extend the definition of these maps to \( \alpha = \infty \) by setting

\[
\Gamma_\infty(h) = h, \quad \hat{\Gamma}_\infty(h) = 0 \quad \text{for } h \in \mathcal{D}([0, \infty) : \mathbb{R}).
\]

Theorem 2.4. Suppose that \( \sqrt{n} \ll d_n \) and

\[
\lambda_n = 1 - \left( \frac{\log d_n}{d_n} + \frac{\alpha_n}{\sqrt{n}} \right), \quad \text{where } \alpha_n \to \alpha \in [0, \infty], \text{ with } \alpha_n^- = O(\sqrt{n}/d_n), \text{ and } \alpha_n = O(n^{1/6}).
\]

\[
(2.12)
\]
Corollary 2.7. As strengthening in that, unlike [32], we allow \( \alpha \rightarrow \) this Theorem, we obtain the specific regime considered in [32] (in fact we provide a slight
\[\ell \] with values in
sequence of random variables \( \{\|B\| \) and \( r \) for some \( \ell \)
Then
Remark 2.8. (i) The existence and uniqueness of solutions to the stochastic integral equations in (2.13) follows by standard fixed point arguments on using the Lipschitz property of the map \( \Gamma \) on
\[D((0, \infty) : \ell_2)\), where \( D \) is the first component of the pair \( (Z, \eta) \) which is a
\( \ell_2 \times R_+ \) valued continuous process given as the unique solution to:
\[(Z(t), \eta(t)) = (\Gamma_\alpha, \Gamma_\alpha) \left( z_1 - \int_0^t (Z_1(s) - Z_2(s))ds + \sqrt{2}B(t) \right)(t), \]
\[Z_2(t) = z_2 - \int_0^t (Z_2(s) - Z_3(s))ds + \eta(t), \tag{2.13} \]
\[Z_i(t) = z_i - \int_0^t (Z_i(s) - Z_{i+1}(s))ds \quad \text{for each } i \in \{3 ... r\}, \]
\[Z_i(t) = 0 \quad \text{for each } i > r, \]
and \( B \) is a standard Brownian motion.
We note that given a standard Brownian motion \( B \), there is a unique continuous process \( (Z, \eta) \) with values in \( \ell_2 \times \mathbb{R}_+ \), adapted to the filtration generated by \( B \) (See Remark 2.8(i)). As a corollary to this Theorem, we obtain the specific regime considered in [32] (in fact we provide a slight strengthening in that, unlike [32], we allow \( \alpha = 0 \)). See Remark 2.8(v) for further discussion.

Corollary 2.7. As \( n \rightarrow \infty \), suppose that \( d_n \gg \sqrt{n} \log n \) and \( \sqrt{n}(1 - \lambda_n) \rightarrow \alpha \in [0, \infty), \) along with
\[\sqrt{n}(1 - \lambda_n) \geq (\sqrt{n} \log n)/d_n \] for large \( n \) if \( \alpha = 0 \). Let \( Y_n(\cdot) \sim \sqrt{n}(G_n(\cdot) - f_1) \) and assume that the
sequence of random variables \( \{\|Y_n(0)\|_1\} \) is tight, and as \( n \rightarrow \infty \), \( Y_n(0) \xrightarrow{P} y \in \ell_2 \) with \( y_{r+} = 0 \)
for some \( r \geq 2 \). Then \( Y_n \Rightarrow Y \) in \( D((0, \infty) : \ell_2)\), where \( (Y, \eta) \) is the \( \ell_2 \times [0, \infty) \) valued continuous process given by the unique solution to
\[(Y_1(t), \tilde{\eta}(t)) = (\Gamma_0, \Gamma_0) \left( y_1 - \alpha \, \text{id}(\cdot) - \int_0^t (Y_1(s) - Y_2(s))ds + \sqrt{2}B(\cdot) \right)(t), \]
\[Y_2(t) = y_2 - \int_0^t (Y_2(s) - Y_3(s))ds + \tilde{\eta}(t), \]
\[Y_i(t) = y_i - \int_0^t (Y_i(s) - Y_{i+1}(s))ds \quad \text{for each } i \in \{3 ... r\}, \]
\[Y_i(t) = 0 \quad \text{for each } i > r, \]
and \( B \) is a standard Brownian motion.

Remark 2.8.
(i) The existence and uniqueness of solutions to the stochastic integral equations in (2.13) follows by standard fixed point arguments on using the Lipschitz property of the map \( \Gamma_\alpha \) on
\[D((0, \infty) : \mathbb{R}), \) This system of equations can equivalently be written as
\[Z_1(t) = z_1 - \int_0^t (Z_1(s) - Z_2(s))ds + \sqrt{2}B(t) - \eta(t), \]
\[Z_2(t) = z_2 - \int_0^t (Z_2(s) - Z_3(s))ds + \eta(t), \tag{2.14} \]
\[Z_i(t) = z_i - \int_0^t (Z_i(s) - Z_{i+1}(s))ds \quad \text{for each } i \in \{3 ... r\}, \]
\[Z_i(t) = 0 \quad \text{for each } i > r, \]
where $\eta = 0$ when $\alpha = \infty$, and when $\alpha \in \mathbb{R}$, it satisfies
\[
\eta(0) = 0 \text{ and } \eta \text{ is a non-decreasing function.}
\]
\[
Z_1(t) \leq \alpha
\]
\[
\int_0^\infty (\alpha - Z_1(s))d\eta(s) = 0
\]

The system of equations (2.14) describes a constrained multi-dimensional diffusion driven by a one-dimensional Brownian motion. Existence and uniqueness for a similar system of equations and the convergence of $Y_n$ to that system when $d_n = n$ is shown in [10]. However note that unlike in [10] (where the reflection is at 0), the reflection in (2.14) occurs at a barrier $\alpha \in [0, \infty]$.

(ii) The convergence $\mu_n \to f_1$ along with tightness of $\{Z_n\}_{n \in \mathbb{N}}$ shows that, under the conditions of Theorems 2.3 or 2.4, most queues will be of length 1 on any fixed interval $[0, T]$.

(iii) The limit system in Theorem 2.4 simplifies when $z_i = 0$ for $i > 2$ and is given as $Z_i = 0$ for all $i > 2$, and
\[
Z_1(t) = z_1 - \int_0^t (Z_1(s) - Z_2(s))ds + \sqrt{2B(t)} - \eta(t),
\]
\[
Z_2(t) = z_2 - \int_0^t Z_2(s)ds + \eta(t),
\]
where $\eta$ is as in the statement of the theorem.

(iv) Suppose that $d_n = \sqrt{n} \log n / 2a$ for some $a > 0$ and $1 - \lambda_n = \frac{a}{\sqrt{n}} + \frac{2a(\log \log n + O(1))}{\sqrt{n} \log n}$. Then the assumptions in Theorem 2.4 are satisfied with $\alpha = 0$. In this case the reflection barrier is at 0, namely $Z_1(t) \leq 0$ for all $t$. Also note that since $\sqrt{n}(1 - \lambda_n) \to a$, we have that $\mu_{n,1} = \lambda_n \to 1$. Since $d_n/\sqrt{n} \to \infty$, this shows that for $k \geq 2$
\[
\sqrt{n} \mu_{n,2} = \sqrt{n} \lambda_n \beta_n(\lambda_n) \leq \sqrt{n} \lambda_n \lambda_n^{d_n} = \sqrt{n}(1 - (1 - \lambda_n))^{d_n+1} \to 0.
\]

Using $\mu_{n,i+1} \leq \mu_{n,i}$, see that $\sqrt{n}(\mu_n - f_1) \to -ae_1 \in \ell_1$ and hence the fluctuations of $G_n$ about the fixed point $f_1$ can be characterized as well. Specifically, letting $Y_n(\cdot) = \sqrt{n}(G_n(\cdot) - f_1) = Z_n(\cdot) + \sqrt{n}(\mu_n - f_1)$, we see that, under the condition of the above theorem, $Y_n \Rightarrow Y$ in $\mathbb{D}([0, \infty) : \ell_2)$, where $Y = Z - ae_1$ and hence, assuming $z_i = 0$ for $i > 2$, $(Y, \tilde{\eta}) \in \mathbb{C}([0, \infty) : \ell_2 \times \mathbb{R}_+)$ is the unique solution to (2.15) with $(Z_1, \eta, \alpha)$ replaced with $(Y_1, \tilde{\eta}, -a)$, and the equations
\[
Y_1(t) = y_1 - at - \int_0^t (Y_1(s) - Y_2(s))ds + \sqrt{2B(t)} - \tilde{\eta}(t),
\]
\[
Y_2(t) = y_2 - \int_0^t Y_2(s)ds + \tilde{\eta}(t),
\]
where $y = z - ae_1$ and $B$ is a standard Brownian motion. In particular, the limit $Y$ takes the same form as in [10,32] with a stronger constraint that $Y_1(t) \leq -a < 0$ for each $t > 0$.

(v) Suppose that $d_n \gg \sqrt{n} \log n$. Then it is easy to see that (2.12) holds with some $\alpha > 0$ if and only if $\sqrt{n}(1 - \lambda_n) \to \alpha > 0$. This regime was studied in [32]. Using the arguments as in (iv) above, it is easy to check that $\sqrt{n}(\mu_n - f_1) \to -ae_1$ in $\ell_1$ (and hence $\ell_2$). Corollary 2.7 is immediate from this and Theorem 2.4. In particular we recover [32, Theorem 3]. However the proof techniques in the current paper are different from the stochastic coupling techniques employed in [32].

(vi) Suppose $\sqrt{n} \ll d_n \ll \sqrt{n} \log n$ and that (2.12) holds with $\alpha < \infty$. Then, as observed in the proof of Theorem 4 in [32], $Y_n$ will not be tight in this regime. But since $\sqrt{n}(1 - \mu_{n,1}) = (\sqrt{n} \log d_n)/d_n + \alpha_n \to \infty$, this does not preclude the convergence of $Z_n = Y_n - \sqrt{n}(\mu_n - f_1)$. 
Indeed, Theorem 2.4 shows that the process \( \mathbf{Z}_n \) converges in distribution and the limit process has a reflecting barrier at \( \alpha \), i.e. \( Z_1 \leq \alpha \). In particular, unlike the case \( d_n \gg \sqrt{n \log n} \), the barrier in this case does not come from the constraint \( G_{n,1} \leq 1 \).

(vii) Theorem 2.4 allows for a slower approach to criticality than \( n^{-1/2} \), e.g. \( \lambda_n \) such that \( n^{1/3}(\lambda_n - 1) \to \gamma > 0 \). In this case \( \alpha = \infty \) and there is no reflection. When \( z_i = 0 \) for all \( i \geq 1 \), this system reduces to the one-dimensional OU process given by (2.10) with \( Z_i = 0 \) for \( i > 1 \).

Table 2.1 below summarizes some of the key regimes of \((d_n, \lambda_n)\) that are covered by Theorems 2.2, 2.3 and places them in the context of previous work on JSQ\((d_n)\) systems in heavy traffic.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Regimes of ( a, b_n ) and ( k )</th>
<th>Analysis type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braverman [4]</td>
<td>( a = 0.5, b_n = 1, k = 1 )</td>
<td>Convergence of stationary distribution</td>
</tr>
<tr>
<td>Eschenfeldt &amp; Gamarnik [10]</td>
<td>( a = 0.5, b_n = 1, k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Mukherjee et al. [32]</td>
<td>( a = 0.5, b_n \in (0.5 + \log \log \frac{n}{n}, 1], k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Theorem 2.4 (( \alpha = 0 ))</td>
<td>( a \in (1/2, 1), b_n \in [a + \log \log \frac{n}{n}, 1], k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Theorem 2.4 (( \alpha \in (0, \infty) ))</td>
<td>( a = 0.5, b_n \in [0.5 + \log \log \frac{n}{n}, 1], k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Theorem 2.4 (( \alpha = \infty ))</td>
<td>( a \in [1/3, 1/2), b_n \in [0.5, 1], k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Theorem 2.3 (( \alpha = \infty ))</td>
<td>( a \in (1/4, 1/2), b_n = 0.5, k = 1 )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Theorem 2.3 (( \alpha = 0 ))</td>
<td>( a \in (0, 1), b_n = (a + \log \log \frac{n}{n})/k, k &gt; n/(1 - a), k \in \mathbb{N} )</td>
<td>Functional central limit theorem</td>
</tr>
<tr>
<td>Brightwell et al. [8]</td>
<td>( a \in (0, 1), b_n \to b \in (0, 1), 2a &lt; 1 + b(k - 1), a/b \notin \mathbb{N} )</td>
<td>Equilibrium queue lengths</td>
</tr>
<tr>
<td>Liu &amp; Ying [22, 23]</td>
<td>( a \in (0, 1/2), b_n \in [a + \log \log \frac{n}{n}, 1], k = 1 )</td>
<td>Equilibrium performance</td>
</tr>
</tbody>
</table>

In order to make comparison with [5], note that the regime in * can equivalently be written as \( a \in (0, 1), b_n = (a + \log \log \frac{n}{n})/k \to b, \ 2a < 1 + b(k - 1), a/b \notin \mathbb{N} \).

3. Poisson Representation of State Processes

We now embark on the proofs of the main results. We start with a brief overview of the organization of the proofs. In this Section we describe a specific construction of the state process. Proof of the law of large numbers (Theorem 2.1) is given in Section 4. Section 5 describes fine-scaled (deterministic) properties of the function \( \beta_n \) and the near fixed points \( \mu_n \) which play a key technical role in the proofs of our diffusion approximations. Section 6 derives preliminary estimates required to prove all the main results for the fluctuations of the state process. Sections 7, 8 and 9 complete the proofs of Theorems 2.2, 2.3 and 2.4 respectively.

We start with a specific construction of the state process through time changed Poisson processes (cf. [11, 20]). A similar representation has been used in previous work on JSQ\((d)\) systems (cf. [9, 10, 32]). Let \( \{ N_{i,+}, N_{i,-} : i \geq 1 \} \) be a collection of mutually independent rate one Poisson processes given on some probability space \((\Omega, \mathcal{F}, \mathbf{P})\). Then \( \mathbf{G}_n \) has the following (equivalent in distribution) representation. For \( i \geq 1 \) and \( t \geq 0 \)

\[
G_{n,i}(t) = G_{n,i}(0) - \frac{1}{n} N_{i,-} \left( n \int_0^t [G_{n,i}(s) - G_{n,i+1}(s)] \, ds \right)
+ \frac{1}{n} N_{i,+} \left( \lambda_n n \int_0^t [\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))] \, ds \right),
\]

where \( G_{n,0}(t) = 1 \) for all \( t \geq 0 \). Denoting

\[
A_{n,i}(t) \doteq N_{i,+} \left( \lambda_n n \int_0^t [\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))] \, ds \right),
D_{n,i}(t) \doteq N_{i,-} \left( n \int_0^t [G_{n,i}(s) - G_{n,i+1}(s)] \, ds \right),
\]

the above evolution equation can be rewritten as

\[
G_{n,i}(t) = G_{n,i}(0) - \frac{1}{n} D_{n,i}(t) + \frac{1}{n} A_{n,i}(t), \quad i \in \mathbb{N}, t \geq 0.
\]
Using these martingales, the evolution of $G_t$ is given by
\begin{equation}
G_t = G_t(0) - \int_0^t (G_n(s) - G_{n,i+1}(s))ds + \lambda_n \int_0^t \beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))ds + M_{n,i}(t), \quad i \geq 1
\end{equation}
and
\begin{equation}
\langle M_{n,i} \rangle_t = \frac{1}{n} \left( \int_0^t (G_n(s) - G_{n,i+1}(s))ds + \lambda_n \int_0^t (\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s)))ds \right) .
\end{equation}

We will assume throughout that $G_n(0) \in \ell_1^\beta$ a.s. Then it follows that, for every $t \geq 0$, $\|G_n(t)\|_1 < \infty$ a.s. Indeed, since $n \|G_n(t)\|_1$ equals the total number of jobs in the system at time $t$, and over any time interval $[0, t]$ finitely many jobs enter the system a.s., denoting by $A_t$ the number of jobs that arrive over $[0, t]$, we see that $\|G_n(t)\|_1 \leq \|G_n(0)\|_1 + A_t/n < \infty$ a.s. Thus $G_n$ is a stochastic process with sample paths in $\mathbb{D}([0, \infty) : \ell_1^\beta)$. Note that, for any $t > 0$, $\|G_n(t) - G_n(t^-)\|_1 \leq \frac{1}{n}$.

**Remark 3.1.** Let $a_n, b : \ell_1^\beta \to \ell_1$ be given by
\begin{equation}
a_n(x)_i = \lambda_n(\beta_n(x_{i-1}) - \beta_n(x_i)), \quad b(x)_i = x_i - x_{i+1}, \quad x \in \ell_1^\beta, \quad i \geq 1,
\end{equation}
where, by convention, for \( x \in \ell^1_1 \), \( x_0 = 1 \). Then (3.5) can be rewritten as an evolution equation in \( \ell_1 \) as,
\[
G_n(t) = G_n(0) \left( \sum_{i=0}^{t} \left[ a_n(G_n(s)) - b(G_n(s)) \right] ds + M_n(t), \right)
\]
(3.7)
where \( M_n(t) = (M_n(t))_{t \geq 1} \) is a stochastic process with sample paths in \( \mathbb{D}([0, \infty) : \ell_1) \) and the integral is a Bochner-integral \[38\]. Note that the near fixed point \( \mu_n \) for any \( \mu \) under conditions, \( \mu \) converges to one of the fixed points of the fluid limit (2.2) when \( \lambda = 1 \).

Another reason for this terminology comes from the results in Theorems 2.2–2.4 which show that, taking \( (g, v) \) and \( (\ell^1_1) \) to stay close to \( \mu_n \) (over any compact time interval) as \( n \to \infty \). In this sense \( \mu_n \) can be viewed as a ‘near fixed point’ of \( G_n(\cdot) \) and the terminology in Definition 2 is justified. Another reason for this terminology comes from the results in Theorems 2.2–2.4 which show that, under conditions, \( \mu_n \) converges to one of the fixed points of the fluid limit (2.2) when \( \lambda = 1 \).

4. The Law of Large Numbers

In this section we prove Proposition 2.1 and Theorem 2.1.

4.1. Uniqueness of Fluid Limit Equations. In this subsection we show that there is at most one solution of (2.2) in \( \mathbb{C}([0, \infty) : \ell_1^2 \times \ell_\infty) \). Results of Section 4.2 will provide existence of solutions to this equation. Suppose \( (g, v) \) and \( (g', v') \) are two solutions to (2.2) in \( \mathbb{C}([0, \infty) : \ell_1^2 \times \ell_\infty) \). We will now argue that the two solutions are equal.

We claim that that \( v_i' \) and \( v_i \) are non-zero for only finitely many \( i \)'s. Indeed, since \( g, g' \in \mathbb{C}([0, T] : \ell^2_1) \), there is a constant \( C \in (0, \infty) \) so that \( \sup_{s \leq T} \|g(s)\|_1 \vee \sup_{s \leq T} \|g'(s)\|_1 \leq C \). Since
\[
x_i \leq \|x\|_1 / i \quad \text{for any } x \in \ell^1_1,
\]
taking \( M = [C + 1] \in \mathbb{N} \) shows that \( \sup_{s \leq T} g_i(s) \vee g_i'(s) < 1 \) for any \( i \geq M \). But then by the equivalent representation of (2.2) given in (2.3) (in particular the second line), we must have \( v_i = v_i' = 0 \) for any \( i \geq M \). This proves the claim.

Since \( v_i = v_i' = 0 \) for \( i \geq M \), the first line of the equivalent formulation in (2.3) shows that both \( x = g \) and \( x = g' \) satisfy the integral equations
\[
x_i(t) = r_i - \int_0^t (x_i(s) - x_{i+1}(s)) ds \quad \text{for } i \geq M + 1 \text{ and } t \in [0, T].
\]
By standard arguments using Gronwall’s lemma [11, Appendix 5], we then must have \( g_i = g_i' \) for each \( i \geq M + 1 \).

Indeed, letting \( z_i(\cdot) = g_i(\cdot) - g_i'(\cdot) \) for \( i \geq M + 1 \) and \( v(t) = \sum_{i=M+1}^{\infty} |z_i(t)| \) for \( t \in [0, T] \), we have that
\[
|z_i(t)| \leq \int_0^t (|z_i(s)| + |z_{i+1}(s)|) ds \quad \text{for all } i \geq M + 1 \text{ and } t \in [0, T]
\]
and so
\[
v(t) \leq 2 \int_0^t v(s) ds, \quad t \in [0, T],
\]
which implies that \( v(t) = 0 \) for \( t \in [0, T] \).

We now show that \( g_i = g_i' \) for \( i \leq M \). From the definition of the Skorohod map in (2.1) we see that for \( h_1, h_2 \in \mathbb{D}([0, \infty) : \mathbb{R}) \) with \( h_i(0) \leq 1 \), \( i = 1, 2 \), and \( t \geq 0 \),
\[
\|\Gamma_1(h_1) - \Gamma_1(h_2)\|_{s,t} \leq 2 \|h_1 - h_2\|_{s,t}, \quad \|\hat{\Gamma}_1(h_1) - \hat{\Gamma}_1(h_2)\|_{s,t} \leq \|h_1 - h_2\|_{s,t}.
\]
Thus, since \((g, ν)\) and \((g', ν')\) solve (2.2),
\[
\|g_i - g'_i\|_{s,t} \leq 2 \left( \int_0^t \|g_i - g'_i\|_{s,s} \, ds + \int_0^t \|g_{i+1} - g'_{i+1}\|_{s,s} \, ds + \|v_{i-1} - v'_{i-1}\|_{s,t} \right),
\]
and
\[
\|v_i - v'_i\|_{s,t} \leq 2 \int_0^t H_s ds + \|v_{i-1} - v'_{i-1}\|_{s,t}
\]
for any \(i \geq 1\). Let \(H_t = \max_{i \in \{1, \ldots, M\}} \|g_i - g'_i\|_{s,t}\). Note \(g_{M+1} = g'_{M+1}\) and hence \(H_t = \max_{i \in \{1, \ldots, M+1\}} \|g_i - g'_i\|_{s,t}\). Then from (4.3), we have
\[
\|v_i - v'_i\|_{s,t} \leq 2 \int_0^t H_s ds + \|v_{i-1} - v'_{i-1}\|_{s,t}
\]
for any \(i \leq M\).

Repeatedly using (4.4) along with \(v_0 = v'_0\) shows that \(\|v_i - v'_i\|_{s,t} \leq 2i \int_0^s H_s ds\) for any \(i \leq M\).

Using this bound in (4.2) shows for \(1 \leq i \leq M\):
\[
\|g_i - g'_i\|_{s,t} \leq 2 \left( 2 \int_0^t H_s ds + 2(i - 1) \int_0^t H_s ds \right) = 4i \int_0^t H_s ds.
\]

Hence considering the maximum of \(\|g_i - g'_i\|_{s,t}\) over \(1 \leq i \leq M\) we get
\[
0 \leq H_t \leq 4M \int_0^t H_s ds
\]
for each \(t \in [0, T]\).

Gronwall’s Lemma now shows that \(H_T = 0\), and hence \(g_i = g'_i\) for \(i = 1 \ldots M\). Finally, since \(v_0 = v'_0\), we see recursively from the second equation in (2.2) that \(v_i = v'_i\) for all \(i \geq 0\).

### 4.2. Tightness and Limit Point Characterization

Some of the arguments in this section are similar to [32] however in order to keep the presentation self-contained we provide details in a concise manner. The next result establishes the convergence of the martingale term \(M_n\) in the semimartingale decomposition in (3.7). Throughout this subsection and the next we assume that the conditions of Theorem 2.1 are satisfied, namely, \(G_n(0) \overset{P}{\to} r\) in \(\ell_1^n\), \(\lambda_n \to \lambda\) and \(d_n \to \infty\), as \(n \to \infty\).

**Lemma 4.1.** For any \(T > 0\), \(\sup_{s \leq T} \|M_n(s)\|_2 \overset{P}{\to} 0\).

**Proof.** It suffices to show that for any \(T > 0\), \(\lim_n E \sup_{s \leq T} \|M_n(s)\|_2^2 = 0\). Applying Doob’s maximal inequality we have that
\[
E \sup_{s \leq T} \|M_n(s)\|_2^2 \leq 4E \|M_n(T)\|_2^2 = 4E \sum_{i \geq 1} M_{n,i}(T)^2.
\]
Since \(E M_{n,i}^2(T) = E \langle M_{n,i}\rangle_T\), using the monotone convergence theorem in (4.5) shows,
\[
E \sup_{s \leq T} \|M_n(s)\|_2^2 \leq 4E \sum_{i \geq 1} \langle M_{n,i}\rangle_T \leq 4 \frac{T(1 + \sup_n \lambda_n)}{n},
\]
where the last inequality is from (3.6) on observing that
\[
\sum_{i=1}^{\infty} \langle M_{n,i}\rangle_T \leq \frac{1}{n} \int_0^T G_{n,1}(s) + \frac{\lambda_n}{n} \int_0^T \beta_n(G_{n,0}(t)) \leq \frac{T(1 + \lambda_n)}{n}.
\]
Sending \(n \to \infty\) in (4.6) completes the proof of the lemma.

The next proposition characterizes compact sets in \(\ell_1^n\). The proof is standard and can be found for example in [32].

**Proposition 4.2.** A subset \(C \subseteq \ell_1^n\) is precompact if and only if the following two conditions hold:
(1) (norm-bounded) \( \sup_{x \in C} \|x\|_1 < \infty \), and
(2) (uniformly decaying tails) \( \lim_{M \to \infty} \sup_{x \in C} \sum_{i > M} |x_i| = 0 \).

Lemma 4.3. For each \( n \in \mathbb{N} \) there is a square integrable \( \{F^n_t\} \)-martingale \( \{L_n(t)\} \) such that, for any \( t \geq 0 \),
\[
\sup_{s \in [0,t]} \|G_n(s)\|_1 \leq \|G_n(0)\|_1 + \lambda_n t + L_n(t).
\]
Furthermore, \( \langle L_n \rangle_t \leq \frac{\lambda_n t}{n} \), for all \( t \geq 0 \).

Proof. For \( i = 1, \ldots, n \), let \( X_i(t) \) denote the number of jobs in the \( i \)-th server’s queue at time \( t \). Then
\[
\|G_n(t)\|_1 = \sum_{j=1}^{\infty} G_{n,j}(t) = \sum_{j=1}^{\infty} \sum_{i=1}^{n} \frac{\mathbb{I}(X_i(t) \geq j)}{n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{\infty} \mathbb{I}(X_i(t) \geq j) = \frac{1}{n} \sum_{i=1}^{n} X_i(t).
\]
Hence \( \|G_n(t)\|_1 \) is the total number of jobs in the system at time \( t \), divided by \( n \).
Since the total number of jobs in the system at time \( t \) is bounded above by the sum of the number of job arrivals by time \( t \) and the initial number of jobs, \( \sup_{s \in [0,t]} \|G_n(s)\| \leq \|G_n(0)\|_1 + \frac{A_n(t)}{n} \), where \( A_n(t) \) is the total number of arrivals to the system by time \( t \). Since, \( A_n \) is a Poisson process with arrival rate \( \lambda n \), the result follows on setting \( L_n(t) = \frac{A_n(t)}{n} - \lambda n t, t \geq 0 \).

The estimate in the next lemma will be useful when applying Aldous-Kurtz tightness criteria [11] for proving tightness of \( \{G_n\} \).

Lemma 4.4. Fix \( n \in \mathbb{N} \) and \( \delta \in (0, \infty) \). Let \( \tau \) be a bounded \( \{F^n_t\} \)-stopping time. Then
\[
E \|G_n(\tau + \delta) - G_n(\tau)\|_1 \leq (\lambda_n + 1) \delta
\]

Proof. From (3.2), for any \( i \in \mathbb{N} \),
\[
|G_{n,i}(\tau + \delta) - G_{n,i}(\tau)| \leq \frac{1}{n} (A_{n,i}(\tau + \delta) - A_{n,i}(\tau) + D_{n,i}(\tau + \delta) - D_{n,i}(\tau)). \tag{4.7}
\]
From (3.3) and (3.4) we see that
\[
E \frac{1}{n} (A_{n,i}(\tau + \delta) - A_{n,i}(\tau)) = \lambda_n E \int_{\tau}^{\tau + \delta} (\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))) \, ds
\]
\[
E \frac{1}{n} (D_{n,i}(\tau + \delta) - D_{n,i}(\tau)) = E \int_{\tau}^{\tau + \delta} (G_{n,i}(s) - G_{n,i+1}(s)) \, ds.
\]
Using the above identities in (4.7)
\[
E |G_{n,i}(\tau + \delta) - G_{n,i}(\tau)| \leq \lambda_n E \int_{\tau}^{\tau + \delta} (\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))) \, ds + E \int_{\tau}^{\tau + \delta} (G_{n,i}(s) - G_{n,i+1}(s)) \, ds \tag{4.8}
\]
Adding (4.8) over various values of \( i \in \mathbb{N} \), we have
\[
E \|G_n(\tau + \delta) - G_n(\tau)\|_1 \leq \lambda_n \sum_{i=1}^{\infty} E \int_{\tau}^{\tau + \delta} (\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s))) \, ds
\]
\[
+ \sum_{i=1}^{\infty} E \int_{\tau}^{\tau + \delta} (G_{n,i}(s) - G_{n,i+1}(s)) \, ds
\]
\[
\leq E \int_{\tau}^{\tau + \delta} (\lambda_n \beta_n(G_{n,0}(s)) + G_{n,1}(s)) \, ds \leq (\lambda_n + 1) \delta.
\]
The following lemma will be useful in verifying the tightness of \( \{G_n(t)\} \) in \( l^1_t \) for each fixed \( t \geq 0 \).

**Lemma 4.5.** For every \( n, m \in \mathbb{N} \) there is a square integrable \( \{\mathcal{F}_t^n\} \) martingale \( L_{n,m} (\cdot) \) so that, for all \( t \geq 0 \),

\[
\sup_{s \leq t} \sum_{i > m} G_{n,i}(s) \leq \sum_{i > m} G_{n,i}(0) + \frac{\lambda_n t}{m} \|G_n\|_{1,t} + L_{n,m}(t)
\]

and \( \langle L_{n,m} \rangle_t \leq \frac{\lambda_n t}{m} \|G_n\|_{1,t} \).

**Proof.** From \((3.1)\), for any \( i \in \mathbb{N} \) and \( t \geq 0 \):

\[
G_{n,i}(t) \leq G_{n,i}(0) + \frac{1}{n} N_{+,i} \left( n \lambda_n \int_0^t \beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s)) ds \right)
\]

Consider the point-process given by

\[
B_{n,m}(t) = \sum_{i > m} N_{+,i} \left( n \lambda_n \int_0^t \beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s)) ds \right).
\]

Adding over \( i > m \) in \((4.9)\) we get

\[
\sup_{s \leq t} \sum_{i > m} G_{n,i}(s) \leq \sum_{i > m} G_{n,i}(0) + \frac{1}{n} B_{n,m}(t)
\]

It is easy to see that, with

\[
b_{n,m}(t) = n \lambda_n \sum_{i > m} \int_0^t \beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s)) ds, \ t \geq 0,
\]

\( \tilde{L}_{n,m}(t) = B_{n,m}(t) - b_{n,m}(t) \) is a \( \mathcal{F}_t^n \)-martingale and

\[
\langle \tilde{L}_{n,m} \rangle_t = b_{n,m}(t) = n \lambda_n \int_0^t \beta_n(G_{n,m}(s)) ds \leq n \lambda_n \int_0^t G_{n,m}(s) ds \leq n \lambda_n t \left( \sup_{s \leq t} G_{n,m}(s) \right) \leq \frac{n \lambda_n t}{m} \|G_n\|_{1,t},
\]

where, for the last inequality we have used \((4.1)\). The lemma now follows on setting \( L_{n,m}(t) = \tilde{L}_{n,m}(t)/n \) and and using \((4.10)\).

Recall that under our assumptions, \( \lambda_n \to \lambda \) and \( d_n \to \infty \) as \( n \to \infty \).

**Lemma 4.6.** Suppose that \( \{G_n(0)\}_{n \geq 1} \) is a tight sequence of \( l^1_t \) valued random variables. Then for any \( T > 0 \), \( \{G_n\}_{n \geq 1} \) is a tight sequence of \( \mathbb{D} \left( [0,T]: \mathbb{R}_+ \right) \) valued random variables.

**Proof.** To show that \( \{G_n\}_{n \geq 1} \) is tight it suffices to show that (cf. \cite[Theorem 2.7]{[21]} and \cite[Theorem 23.11]{[17]})

1. For any \( t \in [0, T] \) and \( \epsilon > 0 \), there is a compact set \( \Gamma \subset l^1_t \) so that \( \inf_{n \in \mathbb{N}} P(G_n(t) \in \Gamma) \geq 1 - \epsilon \).

2. For every sequence of nonnegative numbers \( \delta_n \) converging to 0 as \( n \to \infty \) and every sequence of \( \mathcal{F}_t^n \)-stopping times \( \tau_n \) such that \( \tau_n \leq T \), \( \limsup_{n \to \infty} E \|G_n(\tau_n + \delta_n) - G_n(\tau_n)\|_1 = 0 \).

The second condition is immediate from Lemma 4.4. Now consider (1). Fix \( \epsilon > 0 \). Let \( \lambda = \sup_{n \geq 1} \lambda_n \). Since \( G_n(0) \) is tight, there is a compact \( K_1 \subset l^1_t \) such that

\[
P(G_n(0) \in K_1) \geq 1 - \frac{\epsilon}{8} \text{ for all } n \in \mathbb{N}.
\]

From Proposition 4.2 there is a \( \kappa_1 \in (0, \infty) \) such that \( \sup_{x \in K_1} \|x\|_1 \leq \kappa_1 \). For the martingale sequence \( \{L_n\}_{n \in \mathbb{N}} \) defined in Lemma 4.3 we can find \( \kappa_2 \in (0, \infty) \) so that

\[
P(|L_n|_{1,T} > \kappa_2) \leq \frac{E \langle L_n \rangle_T}{\kappa_2^2} \leq \frac{\lambda T}{\kappa_2^2} \leq \frac{\epsilon}{8}
\]
for each \( n \in \mathbb{N} \). Then, using the above estimates in Lemma 4.3 with \( \kappa = \bar{\lambda}T + \kappa_1 + \kappa_2 \), we see

\[
P(\| G_n \|_{1,T} \geq \kappa) \leq \frac{\epsilon}{4}.
\]

Let \( m_k \uparrow \infty \) be a sequence such that \( \frac{4\bar{\lambda}T}{m_k} \leq \frac{\epsilon}{2^{k+1}} \) for all \( k \in \mathbb{N} \). Define

\[
K_2 = \left\{ y \in \ell^1_1 : \| y \|_1 \leq \kappa \text{ and for some } x \in K_1, \sum_{j > m_k} y_j \leq \sum_{j > m_k} x_j + \frac{\bar{\lambda}T \kappa}{m_k} + \frac{1}{m_k^{1/4}}, \forall k \in \mathbb{N} \right\}.
\]

Since \( K_1 \) is compact, it is immediate from Proposition 4.2 that \( K_2 \) is precompact in \( \ell^1_1 \). Also, using Lemma 4.5 for any \( t \in [0, T] \),

\[
P(G_n(t) \in K_2) \leq P(\| G_n \|_{1,T} \geq \kappa) + P(G_n(0) \in K_1^c) + P(\| L_{n,m_k} \|_{s,T} > \frac{1}{m_k^{1/4}} \text{ for some } k \in \mathbb{N})
\]

\[
\leq \frac{\epsilon}{4} + \frac{\epsilon}{8} + 4\kappa \bar{\lambda} T \sum_{k=1}^{\infty} m_k^{1/2} \frac{1}{m_k} \leq \epsilon,
\]

where the second inequality follows from Doob’s maximal inequality and from the expression of \( \langle L_{n,m_k} \rangle \) in Lemma 4.5 and the third inequality follows from the choice of \( \{ m_k \} \). This proves (1) and completes the proof of the lemma.

The following lemma gives a characterization of the limit points of \( G_n \).

**Lemma 4.7.** Fix \( T \in (0, \infty) \). Suppose that, along some subsequence \( \{ n_k \}_{k \geq 1} \), \( G_{n_k} \Rightarrow G \) in \( \mathbb{D}( [0, T] : \ell^1_1 ) \) as \( k \to \infty \). Then \( G \in \mathbb{C}( [0, T] : \ell^1_1 ) \) a.s., and (2.2) is satisfied with \( (g_i, v_i) \) replaced with \( (G_i, V_i) \), where \( V_i \) are defined recursively using the second equation in (2.2) with \( V_0(t) = \lambda t \) for \( t \geq 0 \).

**Proof.** From Lemma 4.1 we see that \( M_{n_k} \overset{P}{\to} 0 \), in \( \mathbb{D}( [0, T] : \ell_2 ) \). By Skorohod embedding theorem, let us assume that \( G_{n_k}, M_{n_k}, G \) are all defined on the same probability space and

\[
(G_{n_k}, M_{n_k}) \to (G, 0), \text{ a.s.}
\]

in \( \mathbb{D}( [0, T] : \ell^1_1 \times \ell_2 ) \). Since the jumps of \( G_n \) have size at most \( 1/n \), \( G \) is continuous and \( \| G(s) - G_{n_k}(s) \|_{1,T} \to 0 \) a.s. Similarly, \( \| M_{n_k}(s) \|_{2,T} \to 0 \) a.s. To simplify notation from now on we will take \( n_k = n \).

Let \( V_{n,i}(t) = \lambda_n \int_0^t \beta_n(G_{n,i}(s))ds \) for \( i \geq 1 \) and \( V_{n,0}(t) = \lambda_n t \). From (3.5), for any \( i \geq 1 \)

\[
G_{n,i}(t) = G_{n,i}(0) - \int_0^t (G_{n,i}(s) - G_{n,i+1}(s))ds + V_{n,i-1}(t) - V_{n,i}(t) + M_{n,i}(t).
\]  

(4.11)

For \( i \in \mathbb{N} \), \( \sup_{s \leq T} |G_{n,i}(s) - G_i(s)| \leq \sup_{s \leq T} \| G_n - G \|_{1} \to 0 \) and \( \sup_{s \leq T} \| M_{n,i}(s) \|_2 \to 0 \), a.s. as \( n \to \infty \). We now show that, for each \( i \in \mathbb{N} \), \( V_{n,i} \) converges uniformly on \( [0, T] \) (a.s.) to some limit process \( V_i \). Clearly this is true for \( i = 0 \) and in fact \( V_0(t) = \lambda t, t \geq 0 \). Proceeding recursively, suppose now that \( V_{n,i-1} \to V_{i-1} \) on \( [0, T] \) for some \( i \geq 1 \). Then, since all the terms in (4.11), except \( V_{n,i} \), converge uniformly, \( V_{n,i} \) must converge uniformly as well to some limit process \( V_i \). Sending \( n \to \infty \) in (4.11), we get, for every \( t \leq T \) and \( i \geq 1 \):

\[
G_i(t) = G_i(0) - \int_0^t (G_i(s) - G_{i+1}(s))ds + V_{i-1}(t) - V_i(t), \text{ a.s.}
\]

This shows the first line in (2.3) is satisfied with \( (g_i, v_i) \) replaced with \( (G_i, V_i) \).

We now show that the second line in (2.3) is satisfied as well. Since \( V_i \) is the limit of \( \{ V_{n,i} \} \), the following properties hold:
as a limit point of an arbitrary weakly convergent subsequence of by giving some results on the asymptotic behavior of\[G_i(t)\] was shown in Section 4.1. The result follows.

The fact that this equation can have at most one solution \(G_i\) is continuous, non-decreasing and bounded function and \(V_{n,i} \to V_i\) uniformly on \([0,T]\); the second equality uses the definition of \(V_{n,i}\), the third is from the dominated convergence theorem, and the fourth follows since \(\beta_n(x) \leq x^{d_n}\), for \(x \in [0,1]\) and \(d_n \to \infty\), \(\beta_n(x) \to 0\) for every \(x \in [0,1]\).

Thus we have verified that the second line in (2.3) is satisfied with \((G_i,V_i)\) as well. The result is now immediate from Remark 2.2.

### 4.3. Completing the Proof of LLN.

We can now complete the proofs of Proposition 2.1 and Theorem 2.1.

**Proof of Proposition 2.1.** Fix \(r \in \ell_1^1\), \(\lambda > 0\) and choose a sequence \(r_n \in \ell_1^1\) such that \(r_n \to r\) in \(\ell_1^1\) and for each \(i\), \(nr_{n,i} \in \mathbb{N}_0\). Consider parameters \(\lambda_n = \lambda\), \(d_n = n\) and a JSQ(d) system initialized at \(G_n(0) = r_n\). From Lemma 4.7 we have that there is at least one solution of (2.2) which is given as a limit point of an arbitrary weakly convergent subsequence of \(G_n\) (such a sequence exists in view of the tightness shown in Lemma 4.6). The fact that this equation can have at most one solution was shown in Section 4.1. The result follows.

**Proof of Theorem 2.1.** Since \(G_n(0) \overset{P}{\to} r\) in \(\ell_1^1\), the hypothesis of Lemma 4.6 is satisfied, and thus the sequence \(\{G_n\}_{n \geq 1}\) is tight in \(D([0,T] : \ell_1^1)\) for any fixed \(T > 0\). The result is now immediate from Lemma 4.7 and unique solvability of (2.2) shown in Proposition 2.1.

**Remark 4.8.** We note that the proofs of Lemma 4.7 and Theorem 2.1 also show that, under the conditions of Theorem 2.1 for each \(i \geq 1\),

\[
\sup_{t \leq T} \left| \sum_{i \geq 1} \beta_i(G_n,i) - V_i(t) \right| \overset{P}{\to} 0,
\]

where \((g_i, v_i)\) is the unique solution of (2.2).

### 5. Properties of the Near Fixed Point

In this section we give some important properties of the near fixed point \(\mu_n\) that will be needed in the proofs of fluctuation theorems. Since \(\mu_n\) is defined in terms of the function \(\beta_n\), we begin by giving some results on the asymptotic behavior of \(\beta_n\) and its derivatives. Proofs follow via elementary algebra and Taylor’s approximation and can be found in Appendix A. Roughly speaking, these results control the error between sampling with and without replacement of \(d_n\) servers from a collection of \(n\) servers. We first note that the function \(\beta_n\) is differentiable on \((0,1) \setminus \{d_n-1\}\) and the derivative is given as

\[
\hat{\beta}_n(x) = \sum_{j=0}^{d_n-1} (1 - j/n)^{-1} \prod_{i=0}^{d_n-1} \frac{x - i/n}{1 - i/n} \quad \text{for } x \in (\frac{d_n-1}{n}, 1] \text{ and } \hat{\beta}_n(x) = 0 \quad \text{for } x \in (0, \frac{d_n-1}{n}).
\]
As a convention, we set $\hat{\beta}_n(x) = 0$ for $x = \frac{d_n - 1}{n}$.

Note that $h(t) = \frac{a + t}{b + t}$ is an increasing function of $t$ on $(-b, \infty)$ when $b > a$. Using this fact in (2.7) shows that, when $d_n \leq n$,

$$0 \leq \beta_n(x) \leq a^{d_n} = \gamma_n(x), \quad x \in [0,1]. \quad (5.2)$$

Using the same fact in (5.1) shows that, for $d_n < n$,

$$0 \leq \hat{\beta}_n(x) \leq \frac{d_n a^{d_n - 1}}{1 - \frac{d_n}{n}}, \quad x \in (0,1). \quad (5.3)$$

The following lemma estimates the ratio between $\beta_n$ and $\gamma_n$ and its derivatives.

**Lemma 5.1.** Assume $d_n \ll n$. Then for any $\epsilon \in (0,1)$, as $n \to \infty$,

$$\sup_{x \in [\epsilon,1]} \left| \frac{\beta_n(x)}{\gamma_n(x)} - 1 \right| \to 0. \quad (5.4)$$

Furthermore, if $d_n \ll \sqrt{n}$, then

$$\sup_{x \in [\epsilon,1]} \left| \frac{\beta_n(x)}{\gamma_n(x)} - 1 \right| \to 0 \quad \text{and} \quad \sup_{x \in [\epsilon,1]} \left| \frac{\hat{\beta}_n(x)}{\hat{\gamma}_n(x)} - 1 \right| \to 0. \quad (5.5)$$

The next corollary follows from the proof of Lemma 5.1 (specifically the estimate (A.3) in the proof of the lemma).

**Corollary 5.2.** Assume $d_n \ll n$. Then for any $\epsilon \in (0,1)$

$$\sup_{x \in [\epsilon,1]} |\log \beta_n(x) - \log \gamma_n(x)| = O\left(\frac{d_n^2}{n}\right).$$

Recall the near fixed points $\mu_n = (\mu_n,i)_{i \geq 1}$ introduced in Definition 2.

**Corollary 5.3.** Suppose that $d_n \ll n$. Let $i \in \mathbb{N}$ be such that $\lim \inf_n \mu_{n,i} > 0$. Then

$$\lim_{n \to \infty} \lambda_n \mu_{n,i} \hat{\beta}_n(\mu_{n,i}) = 1.$$  

**Lemma 5.4.** Assume $d_n \ll n$ and fix $\epsilon \in (0,1)$. Then there is a $C \in (0,\infty)$ and $n_0 \in \mathbb{N}$ such that, if for some $k \in \mathbb{N}$ and $n_1 \in \mathbb{N}$, $\mu_{n,k} \geq \epsilon$ for all $n \geq n_1$, then for all $n \geq n_1 \vee n_0$

$$\left| \log \mu_{n,k+1} - (\log \lambda_n) \left( \sum_{i=0}^{k} d_i^i \right) \right| \leq \frac{C}{n} \sum_{i=1}^{k} d_i^{i+1}.$$  

**Corollary 5.5.** Suppose for some $k \in \mathbb{N}$, $1 \ll d_n^{k+1} \ll d_n$. Suppose also that $1 - \lambda_n = \frac{\xi_n - \log d_n}{d_n^k}$ where $\xi_n \to -\log(\alpha) \in (-\infty, \infty]$ and $\frac{\xi_n^2}{d_n^k} \to 0$ as $n \to \infty$. Then $\mu_{n,k} \to 1$ and $\hat{\beta}_n(\mu_{n,k}) \to \alpha$ as $n \to \infty$.

**Lemma 5.6.** Suppose that $\lambda_n \not\to 1$, and $1 \ll d_n \ll n$. Suppose also that, for some $k \geq 2$, $\mu_{n,k} \to 1$ and $\hat{\beta}_n(\mu_{n,k}) \to 0$ in $[0,\infty)$ as $n \to \infty$. Then as $n \to \infty$, $\hat{\beta}_n(\mu_{n,1}) \to \infty$ and for any $i \in [k-1]$

$$\frac{\hat{\beta}_n(\mu_{n,i})}{\hat{\beta}_n(\mu_{n,1})} \to 1.$$
Lemma 5.7. Suppose that $d_n \ll n^{2/3}$. Let $\{\epsilon_n\}$ be a sequence in $[0, 1]$ such that $\epsilon_n^2 \ll d_n^{-1}$. Then as $n \to \infty$:

$$\sup_{x \in [1-\epsilon_n, 1]} \left| \frac{\beta_n(x)}{\gamma_n(x)} - 1 \right| \to 0, \quad (5.6)$$

and

$$\sup_{x \in [1-\epsilon_n, 1]} \left| \frac{\tilde{\beta}_n(x)}{\tilde{\gamma}_n(x)} - 1 \right| \to 0. \quad (5.7)$$

The next result shows that if $d_n \to \infty$, then the behavior of $\beta_n(x)$ is interesting only when $x$ is sufficiently close to 1.

Lemma 5.8. Suppose that $d_n \gg 1$, and let $\epsilon_n = 2 \log_{d_n} d_n$. Then as $n \to \infty$, $\sup_{x \in [0, 1]} |\beta_n(x)| \to 0$. Furthermore, if $\lim \sup_{n} \frac{d_n}{n} < 1$ then we also have $\sup_{x \in [0, 1]} |\tilde{\beta}_n(x)| \to 0$.

6. Preliminary estimates under diffusion scaling

Recall the near fixed point $\mu_n$ from Definition 2 and the process $Z_n$ introduced in (1.2). Also, recall the maps $a_n$ and $b$ from Remark 3.1. We will extend the definition of $\beta_n$ and $\tilde{\beta}_n$ to $\mathbb{R}$ by setting $\beta_n(x) = \tilde{\beta}_n(x) = 0$ for $x < 0$. Further, in what follows, for $z < 0$ and real valued integrable function $h(\cdot)$, the integral $\int_{[0, z]} h(u)du = -\int_{[z, 0]} h(u)du$. We start by giving a semimartingale decomposition for $Z_n$. The quantity $A_n(z)$ defined in the following lemma can be viewed as a discrete derivative of $a_n$ at $\mu_n$ in the direction $z$. The function $A_n$ is asymptotically linear under conditions of Theorem 2.2 (see Lemma 7.1), and is asymptotically non-linear under conditions of Theorems 2.3 and 2.4 (see Lemma 6.7). The asymptotic analysis of this map and the resulting system $Z_n$ is a key ingredient in our proofs.

Lemma 6.1. For $t \geq 0$, $Z_n(t)$ satisfies

$$Z_n(t) = Z_n(0) + \int_0^t A_n(Z_n(s))ds - \int_0^t b(Z_n(s))ds + \sqrt{n}M_n(t), \quad (6.1)$$

where $A_n : \ell_\infty \to \ell_\infty$, is defined as $A_n(z) = \sqrt{n}\{a_n(\mu_n + n^{-1/2}z) - a_n(\mu_n)\}$. Moreover

$$A_n(z)i = q_{n,i-1}(z_{i-1}) - q_{n,i}(z_{i}), \quad i \in \mathbb{N} \quad (6.2)$$

where

$$q_{n,i}(z) \doteq \lambda_n \int_{[0,z]} \tilde{\beta}_n(\mu_{n,i} + y/\sqrt{n})dy, \quad z \in \mathbb{R}, \quad i \in \mathbb{N} \quad (6.3)$$

$$q_{n,0}(z) \doteq 0, \quad z \in \mathbb{R}. \quad (6.3)$$

Proof. From (3.7) and since $a_n(\mu_n) = b(\mu_n)$,

$$\sqrt{n}(G_n(t) - \mu_n) = \sqrt{n}(G_n(0) - \mu_n) + \int_0^t \sqrt{n}\{a_n(G_n(s)) - a_n(\mu_n)\}ds$$

$$- \int_0^t \sqrt{n}\{b(G_n(s)) - b(\mu_n)\}ds + \sqrt{n}M_n(t)$$

Now (6.1) follows by using the the definition of $Z_n$ and $A_n$, and the linearity of $b$. Further, using the definition of $a_n$, we see that (6.2) holds where

$$q_{n,i}(z) \doteq \begin{cases} \lambda_n \sqrt{n}\{\beta_n(\mu_{n,i} + z/\sqrt{n}) - \beta_n(\mu_{n,i})\} & \text{for } i \geq 1 \\ 0 & \text{if } i = 0 \end{cases} \quad (6.4)$$

Clearly, the $q_{n,i}$ defined in (6.4) is same as that given in (6.3). The result follows. \[\Box\]
Lemma 6.2. Suppose that $d_n \to \infty$, $\lambda_n \to 1$, and for some $k \geq 1$, $G_n(0) \xrightarrow{P} f_k$ in $\ell_1$. Then there is a standard Brownian motion $B$ so that $\sqrt{n}M_n \Rightarrow \sqrt{2}Be_k$ in $D([0, \infty) : \ell_2)$.

Proof. Fix $T > 0$. Since $G_n(0) \rightarrow f_k$ and $f_k$ is a fixed point of (2.2), by Theorem 2.1 $G_n \xrightarrow{P} f_k$ in $D([0, T] : \ell_1)$, where $f_k$ here is viewed as the function on $[0, T]$ that takes the constant value $f_k \in \ell_1^d$. Moreover, by Remark 4.5, for every $i \geq 1$ $V_{n,i}(t) \doteq \lambda_n \int_0^t \beta_n(G_{n,i}(s))ds$ converges uniformly on $[0, T]$ in probability to $v_i(t)$, where $v_i$ solves

$$
v_i = \bar{1}_k(f_{k,i} - (f_{k,i} - f_{k,i+1})id + v_{i-1}(.)), \quad i \geq 1,
$$

and $v_0(t) \doteq t$, where recall that ‘id’ denotes the identity map on $[0, T]$. Recalling the definition of $f_k$ we see by a recursive argument that

$$
v_i(t) \doteq \begin{cases} 
  t & \text{if } i < k \\
  0 & \text{if } i \geq k.
\end{cases}
$$

Combining this with (3.6), we have for each $i \geq 1$

$$
\langle \sqrt{n}M_{n,i} \rangle = \int_0^t (G_{n,i}(s) - G_{n,i+1}(s))ds + \lambda_n \int_0^t (\beta_n(G_{n,i-1}(s)) - \beta_n(G_{n,i}(s)))ds
\rightarrow (f_{k,i} - f_{k,i+1})id + v_{i-1}(-) - v_i(-) = H(-),
$$

in probability in $C([0, T] : \mathbb{R})$ where

$$
H(t) \doteq \begin{cases} 
  2t & \text{if } i = k \\
  0 & \text{if } i \neq k,
\end{cases}
\quad t \in [0, T].
$$

Adding (3.6) over $i$, we have for $t \in [0, T]$

$$
\sum_{i > k} \langle \sqrt{n}M_{n,i} \rangle_t \leq \int_0^t G_{n,k+1}(s)ds + \lambda_n \int_0^t \beta_n(G_{n,k})(s)ds.
$$

The process on the right side converges in probability in $C([0, T] : \mathbb{R})$ to $f_{k,k+1}id + v_k(-) = 0$ and thus $\sum_{i > k} \langle \sqrt{n}M_{n,i} \rangle_T$ converges to 0 in probability. By Doob’s maximal inequality,

$$
nE \sup_{t \leq T} \sum_{i > k} M_{n,i}^2(t) \leq 4E \sum_{i > k} \langle \sqrt{n}M_{n,i} \rangle_T \rightarrow 0, \quad \text{as } n \to \infty,
$$

where the last convergence follows by the dominated convergence theorem on noting that the right side of (6.7) is bounded above by $\sup_n (1 + \lambda_n) < \infty$. The result now follows on using the martingale central limit theorem (cf. [11] Theorem 7.1.4) for the $k$-dimensional martingale sequence $(\sqrt{n}M_{n,1}, \cdots, \sqrt{n}M_{n,k})$.

Recall the functions $q_{n,i}$ from Lemma 6.1

Lemma 6.3. Assume that for some $r \in \mathbb{N}$, $\lim \sup_{n \to \infty} \mu_{n,r} < 1$. Then for any $L > 0$

$$
\lim \sup_{n \to \infty} \sup_{0 < |z| \leq L} \left| \sup_{i \geq r} \left| \frac{q_{n,i}(z)}{z} \right| \frac{\lambda_n}{\sqrt{n}} \right| = 0.
$$

Proof. By (6.3):

$$
\sup_{i \geq r} \sup_{0 < |z| \leq L} \left| \frac{q_{n,i}(z)}{z} \right| \leq \lambda_n \sup_{i \geq r} \sup_{0 < |z| \leq L} \left| \frac{\beta_n \left( \mu_{n,i} + \frac{y}{\sqrt{n}} \right)}{z} \right|
= \lambda_n \sup_{0 < |z| \leq L} \left| \frac{\beta_n \left( \mu_{n,i} + \frac{z}{\sqrt{n}} \right)}{z} \right|
\leq \lambda_n \sup_{0 \leq x \leq \mu_{n,r} + \frac{L}{\sqrt{n}}} \frac{\lambda_n}{\sqrt{n}} \beta_n(x).
$$
which converges to 0 by Lemma 5.8 since \( \limsup_{n \to \infty} \left( \mu_{n,r} + \frac{L}{\sqrt{n}} \right) < 1. \)  

For \( L \in (0, \infty) \) define the stopping time

\[
\tau_{n,L} = \inf \left\{ t \mid \|Z_n(t)\|_2 \geq L - \frac{1}{\sqrt{n}} \right\}. \tag{6.8}
\]

Since the jumps of \( Z_n \) are of size \( \frac{1}{\sqrt{n}} \), we see that, for any \( T > 0 \)

\[
\|Z_n\|_{2,T \land \tau_{n,L}} \leq L. \tag{6.9}
\]

Recall from Section 1.2 the vector \( z_{r+} \in \mathbb{R}^\infty \) associated with a vector \( z \in \mathbb{R}^\infty \).

**Lemma 6.4.** Suppose that as \( n \to \infty, G_n(0) \xrightarrow{P} f_k \) in \( \ell_1^+ \) and \( Z_{n,r+}(0) \xrightarrow{P} 0 \) in \( \ell_2 \) for some \( r > k \).

Then for any \( T, L > 0 \), \( \|Z_{n,r+}\|_{2,T \land \tau_{n,L}} \xrightarrow{P} 0. \)

**Proof.** For \( i > k \) and \( z \in \mathbb{R} \), let \( \Delta_{n,i}(z) = \frac{q_n(z)}{\sqrt{z} \mathbb{1}_{z \neq 0}}. \) Then, since \( \lim_{n \to \infty} \mu_{n,k+1} = 0 \), by Lemma 6.3

\[
\delta_{n,L} = \sup_{i \geq k+1} \sup_{|z| \leq L} |\Delta_{n,i}(z)| \to 0, \text{ as } n \to \infty. \tag{6.10}
\]

Next, from (6.1), for \( i \geq r + 1 > k + 1 \)

\[
Z_{n,i}(t \land \tau_n) = Z_{n,i}(0) + \int_0^{t \land \tau_n} \Delta_{n,i-1}(Z_{n,i-1}(s))Z_{n,i-1}(s)ds - \int_0^{t \land \tau_n} \Delta_{n,i}(Z_{n,i}(s))Z_{n,i}(s)ds
\]

\[
- \int_0^{t \land \tau_n} (Z_{n,i}(s) - Z_{n,i+1}(s))ds + \sqrt{n}M_{n,i}(t \land \tau_n)
\]

where we use \( \tau_n \) instead of \( \tau_{n,L} \) for notational simplicity. Then, observing from (6.10) that

\[
\sup_{i \geq k+1} \sup_{t \in [0,\tau_n]} |\Delta_{n,i}(Z_{n,i}(t))| \leq \delta_{n,L}, \text{ we have}
\]

\[
|Z_{n,i}(t \land \tau_n)| \leq |Z_{n,i}(0)| + \delta_{n,L} \int_0^{t \land \tau_n} (|Z_{n,i-1}(s)| + |Z_{n,i}(s)|)ds
\]

\[
+ \int_0^{t \land \tau_n} (|Z_{n,i}(s)| + |Z_{n,i+1}(s)|)ds + \sqrt{n}M_{n,i}(t \land \tau_n). \tag{6.11}
\]

Define maps \( F_1, F_2 : \mathbb{R}^\infty \to \mathbb{R}^\infty \) by

\[
(F_1 x)_i = \begin{cases} 
  x_1 & i = 1 \\
  x_{i-1} + x_i & i \geq 2
\end{cases}
\]

\[
(F_2 x)_i = x_i + x_{i+1}, \quad i \in \mathbb{N}.
\]

Then by collecting (6.11) over all \( i \geq r + 1 \) we get

\[
|Z_{n,r+}(t \land \tau_n)| \leq |Z_{n,r+}(0)| + \delta_{n,L} \int_0^{t \land \tau_n} F_1 |Z_{n,r+}(s)|ds + \delta_{n,L} \int_0^{t \land \tau_n} |Z_{n,r}(s)|e_1ds
\]

\[
+ \int_0^{t \land \tau_n} F_2 |Z_{n,r+}(s)|ds + \sqrt{n}M_{n,r+}(t \land \tau_n). \tag{6.12}
\]

where the absolute values \( |z| \in \mathbb{R}^\infty \) and integrals are interpreted as being coordinate-wise for infinite dimensional vectors \( z \in \mathbb{R}^\infty \). Now noting that the maps \( F_i \), when considered from \( \ell_2 \to \ell_2 \), are bounded linear operators with norm bounded by 2, we have for \( i = 1, 2 \),

\[
\left\| \int_0^{t \land \tau_n} F_i |Z_{n,r+}(s)|ds \right\|_2 \leq \int_0^{t \land \tau_n} 2 \|Z_{n,r+}(s)\|_2 ds.
\]
Using the triangle inequality in (6.12) shows for any \( t \leq T \)
\[
\| Z_{n,r}^+ (t \wedge \tau_n) \|_2 \leq \| Z_{n,r}^+ (0) \|_2 + \left( \sqrt{n} M_{n,r^+} \| e (t, T, r) \|_2 + \delta_{n,L} L T + 2 (1 + \delta_{n,L}) \right) \int_0^{\tau_n \wedge T} \| Z_{n,r}^+(s) \|_2 \, ds
\]
where we have used that \( \int_0^{\tau_n \wedge T} |Z_{n,r}(s)| \, ds \leq L \). Hence, using Gronwall’s inequality
\[
\| Z_{n,r}^+ \|_{2,T \wedge \tau_n} \leq \left( \| Z_{n,r}^+ (0) \|_2 + \delta_{n,L} L T + \| \sqrt{n} M_{n,r^+} \|_{2,T} \right) e^{2 (1 + \delta_{n,L}) T}
\]
Now, as \( n \to \infty \), \( \| Z_{n,r}^+ (0) \|_2 \overset{P}{\to} 0 \) by assumption, \( \delta_{n,L} \to 0 \) by (6.10), and \( \| \sqrt{n} M_{n,r^+} \|_{2,T} \overset{P}{\to} 0 \) by Lemma 6.2. The result follows.

The following elementary lemma will allow us to replace \( \tau_{n,L} \wedge T \) with \( T \) in various convergence results. The proof is omitted.

**Lemma 6.5.** Fix \( T \in [0, \infty) \). Suppose for each \( n \in \mathbb{N} \) and \( L > 0 \) that \( \tau_{n,L} \) is a \([0,T] \)-valued random variable such that \( \lim_{L \to \infty} \sup_n P(\tau_{n,L} < T) \to 0 \) for some \( T > 0 \). Suppose that there is a sequence of stochastic processes \( \{ F_n \}_{n \in \mathbb{N}} \) with sample paths in \( \mathbb{D}([0,T]:\mathbb{R}) \) such that for each \( L > 0 \)
\( |F_n|_{s,T \wedge \tau_{n,L}} \overset{P}{\to} 0 \) as \( n \to \infty \). Then in fact \( |F_n|_{s,T} \overset{P}{\to} 0 \) as \( n \to \infty \).

The next lemma gives conditions under which the near fixed point \( \mu_n \) converges to \( f_1 \).

**Lemma 6.6.** Let \( 0 \leq \epsilon_n \equiv 1 - \lambda_n \) be such that \( \epsilon_n \to 0 \) and \( \epsilon_n d_n \to \infty \). Then \( \mu_n \to f_1 \) in \( \ell_1 \) as \( n \to \infty \).

**Proof.** Using Definition 2 and (5.2) note that \( 0 \leq \mu_{n,i+1} = \lambda_n \beta_n(\mu_{n,i}) \leq \mu_{n,i}^d \) for each \( i \geq 1 \). Hence in order to show \( \mu_n \to f_1 \) in \( \ell_1 \), it suffices to show that (1) \( \mu_{n,1} \to 1 \), and (2) \( \mu_{n,2} \to 0 \). This convergence is immediate on observing for (1) that \( \mu_{n,1} = \lambda_n = 1 - \epsilon_n \to 1 \), and for (2) that \( \mu_{n,2} = \mu_{n,1}^d = (1 - \epsilon_n)^d_n \leq e^{-\epsilon_n d_n} \to 0 \).

The following lemma gives a convenient approximation of the term \( q_{n,1} \) introduced in (6.3) in terms of certain exponentials.

**Lemma 6.7.** Suppose \( d_n \to \infty \) and \( d_n \ll n^{2/3} \). Let \( \lambda_n = 1 - \left( \frac{\log d_n}{d_n} + \frac{\alpha_n}{\sqrt{n}} \right) \) for some real sequence \( \{ \alpha_n \}_{n \geq 1} \) satisfying \( \frac{d_n \alpha_n^2}{n} \to 0 \). Then, for any \( L > 0 \),
\[
\lim_{n \to \infty} \sup_{0 < |z| \leq L} \left| \frac{\exp\left( \frac{d_n}{\sqrt{n}} (z - \alpha_n) \right) - \exp\left( - \frac{d_n}{\sqrt{n}} \alpha_n \right)}{q_{n,1}(z) d_n / \sqrt{n}} - 1 \right| = 0. \tag{6.13}
\]

**Proof.** We only consider the case \( 0 < z \leq L \). The case \(-L \leq z < 0 \) is treated similarly. Recall that \( \mu_{n,1} = \lambda_n \). Noting that \( d_n (1 - \lambda_n + \frac{L}{\sqrt{n}})^2 \leq 4d_n (\log d_n)^2 + \alpha_n^2 / n + L^2 / n \) \( \to 0 \) we have on applying Lemma 5.7 with \( \epsilon_n = (1 - \lambda_n + \frac{L}{\sqrt{n}}) \) that, for any \( |z| \leq L \),
\[
q_{n,1}(z) = (1 + o(1)) \int_0^z \gamma_n \left( \lambda_n + \frac{y}{\sqrt{n}} \right) dy
= (1 + o(1)) \int_0^z \exp\left( (d_n - 1) \log \left( \lambda_n + \frac{y}{\sqrt{n}} \right) \right) dy
= (1 + o(1)) \int_0^z \exp\left( d_n \log \left( \lambda_n + \frac{y}{\sqrt{n}} \right) + \log d_n \right) dy
\]
Using expansion for $\log(1 + h)$ around $h = 0$ and once more the fact that $d_n \left( 1 - \lambda_n + \frac{t}{\sqrt{n}} \right)^2 \to 0$,
\[
q_n,1(z) = (1 + o(1)) \int_0^z \exp \left( d_n \left( \lambda_n - 1 + \frac{y}{\sqrt{n}} \right) + \log d_n \right) dy \\
= (1 + o(1)) \int_0^z \exp \left( \frac{d_n}{\sqrt{n}} (y - \alpha_n) \right) dy \\
= (1 + o(1)) \frac{\exp \left( \frac{d_n}{\sqrt{n}} (z - \alpha_n) \right) - \exp \left( - \frac{\alpha_n}{\sqrt{n}} \right)}{d_n/\sqrt{n}}
\]
which proves (6.13).

Proof of the following lemma proceeds by standard arguments but we provide details in Appendix B.

**Lemma 6.8.** Fix $T > 0$. Let $g,h,M$ be three bounded measurable functions from $[0,T] \to \mathbb{R}$ and assume further that $M$ is a right continuous bounded variation function. Suppose that $m = \inf_{s \in [0,T \wedge \tau]} h(s) > 0$ for some $\tau \geq 0$. Let $z : [0,T] \to \mathbb{R}$ be a bounded measurable function that satisfies for every $t \in [0,T]$
\[
z(t) = z(0) - \int_0^t h(s)z(s)ds + \int_0^t g(s)ds + M(t). \tag{6.14}
\]
Then for any $t \in [0,T \wedge \tau]$
\[
|z(t)| \leq \frac{|g|_{s,T \wedge \tau}}{m} + 2 |M|_{s,T \wedge \tau} + e^{-mt} |z(0)|.
\]

**Lemma 6.9.** Fix $T \in (0, \infty)$. For each $n$, let $V_n$ be a martingale with respect to some filtration $\{G^n_t\}$ such that $V_n(0) = 0$. Let $(r_n)_{n=1}^\infty$ be a positive sequence so that $\lim_{n \to \infty} r_n = +\infty$. Suppose that there is a $C \in (0, \infty)$ such that for all $n \in \mathbb{N}$ and $t \in [0,T]$, $\langle V_n \rangle_t \leq Ct$. Then for any $\epsilon > 0$
\[
\mathbb{P} \left( \sup_{0 \leq t \leq T} (V_n(t) - r_n t) > \epsilon \right) \to 0
\]
as $n \to \infty$.

Proof. Let $\delta_n = \frac{1}{\sqrt{n}}$. Then
\[
\mathbb{P} \left( \sup_{0 \leq t \leq T} |V_n(t) - r_n t| > \epsilon \right) \leq \mathbb{P} \left( \sup_{0 \leq t \leq \delta_n} |V_n(t)| > \epsilon \right) + \mathbb{P} \left( \sup_{\delta_n < t \leq T} |V_n(t)| > r_n \delta_n \right)
\]
\[
\leq \frac{4E V_n(\delta_n)^2}{\epsilon^2} + \frac{4E V_n(T)^2}{r_n^2 (\delta_n)^2}
\]
\[
= \frac{4E \langle V_n \rangle_{\delta_n}}{\epsilon^2} + \frac{4E \langle V_n \rangle_{T}}{r_n^2 (\delta_n)^2} \leq \frac{4C \delta_n}{\epsilon^2} + \frac{4CT}{r_n^2 (\delta_n)^2} \to 0
\]
where the inequality on the second line is from Doob’s maximal inequality.

7. Proof of Theorem 2.2

Now we start with some preliminary lemmas. Recall from Remark 2.5 (b) that under the hypothesis of Theorem 2.2 we have $\mu_n \to f_k \in \ell_1^q$ as $n \to \infty$. Along with the tightness of $\{\|Z_n(0)\|_1\}_{n \in \mathbb{N}}$ this shows that $G_n(0) \to f_k \in \ell_1^q$ as $n \to \infty$. 
Lemma 7.1. Let $d_n \to \infty$, $\frac{d_n}{\sqrt{n}} \to 0$, and $\lambda_n \not\to 1$. Assume that for some $k \in \mathbb{N}$, $\lim \inf_n \mu_{n,k} = \delta > 0$. Then for any $L > 0$ and $1 \leq i \leq k$, as $n \to \infty$

$$\sup_{0<|z|\leq L} \left(\frac{\beta_n(\mu_{n,i})}{\sqrt{n}}\right)^{-1} \lambda_n \int_0^z \frac{\dot{\beta}_n(\mu_{n,i} + y/\sqrt{n})}{\dot{\beta}_n(\mu_{n,i})} dy - 1 \to 0 \quad (7.1)$$

Proof. To prove (7.1), we will approximate $\dot{\beta}_n(\mu_{n,i})$ by $\dot{\gamma}_n(\mu_{n,i})$. Using Lemma 5.1

$$\epsilon_n = \sup_{z \in [\delta/2,1]} \left|\frac{\dot{\beta}_n(\mu_{n,i})}{\dot{\gamma}_n(\mu_{n,i})} - 1\right| \to 0.$$ 

Since $\lim \inf_n \mu_{n,k} > \delta/2$ and $j \mapsto \mu_{n,j}$ is decreasing, there is an $N_0$ so that for $n \geq N_0$, $\mu_{n,i} + \frac{y}{\sqrt{n}} \geq \frac{\delta}{2}$, for any $i \leq k$ and $y \in \mathbb{R}$ with $|y| \leq L$. Hence uniformly in $0 < |z| \leq L$ and $i \leq k$:

$$\frac{\lambda_n}{z} \int_0^z \frac{\dot{\beta}_n(\mu_{n,i} + y/\sqrt{n})}{\dot{\beta}_n(\mu_{n,i})} dy = \frac{1 + o(1)}{z} \int_0^z \frac{\dot{\gamma}_n(\mu_{n,i} + y/\sqrt{n})}{\dot{\gamma}_n(\mu_{n,i})} dy$$

$$= \frac{1 + o(1)}{z} \int_0^z \left(1 + \frac{y}{\sqrt{n}\mu_{n,i}}\right)^{d_n-1} dy$$

$$= \frac{1 + o(1)}{z} \int_0^z \exp\left((d_n - 1) \log \left(1 + \frac{y}{\sqrt{n}\mu_{n,i}}\right)\right) dy$$

$$= \frac{1 + o(1)}{z} \int_0^z \exp\left(O\left(\frac{d_nL}{\sqrt{n}\delta}\right)\right) dy \to 1$$

This shows (7.1).

Remark 7.2. Suppose that the hypothesis of Lemma 7.1 hold. Recall the definition of $\Delta_{n,i}$ for $i > k$ from the proof of Lemma 6.4. We extend this definition by setting

$$\Delta_{n,i}(z) \doteq q_{n,i}(z)/(\dot{\beta}_n(\mu_{n,i})I_{\{z \neq 0\}}) - 1 \quad (i \leq k) \quad (7.2)$$

where $q_{n,i}$ is defined by (6.3). With this extension

$$q_{n,i}(z) = \begin{cases} 
\dot{\beta}_n(\mu_{n,i})(1 + \Delta_{n,i}(z))z & \text{if } 1 \leq i \leq k \\
\Delta_{n,i}(z) & \text{if } i > k
\end{cases} \quad (7.3)$$

Using this notation, Lemma 7.1 and Lemma 6.3 show that, for any $L > 0$

$$\gamma_{n,L} \doteq \sup_{i \in \mathbb{N}} \sup_{0<|z|\leq L} |\Delta_{n,i}(z)| \to 0 \text{ as } n \to \infty. \quad (7.4)$$

The following corollary is an immediate consequence of Remark 7.2 and Lemma 6.1

Corollary 7.3. Under the hypothesis of Lemma 7.1, $Z_n$ satisfies the following integral equations.

For $i = 1$

$$Z_{n,1}(t) = Z_{n,1}(0) - \int_0^t \dot{\beta}_n(\mu_{n,1})(1 + \Delta_{n,1}(Z_{n,1}(s)))Z_{n,1}(s) ds - \int_0^t (Z_{n,1}(s) - Z_{n,2}(s)) ds + \sqrt{n}M_{n,1}(t)$$

For $i \in \{2, \ldots, k\}$

$$Z_{n,i}(t) = Z_{n,i}(0) + \int_0^t \dot{\beta}_n(\mu_{n,i-1})(1 + \Delta_{n,i-1}(Z_{n,i-1}(s)))Z_{n,i-1}(s) ds$$

$$- \int_0^t \dot{\beta}_n(\mu_{n,i})(1 + \Delta_{n,i}(Z_{n,i}(s)))Z_{n,i}(s) ds - \int_0^t (Z_{n,i}(s) - Z_{n,i+1}(s)) ds + \sqrt{n}M_{n,i}(t).$$
For \( i = k + 1 \)
\[
Z_{n,k+1}(t) = Z_{n,k+1}(0) + \int_0^t \dot{\beta}_n(\mu_{n,k})(1 + \Delta_{n,k}(Z_{n,k}(s)))Z_{n,k}(s)ds \\
- \int_0^t \Delta_{n,k+1}(Z_{n,k+1}(s))Z_{n,k+1}(s) - \int_0^t (Z_{n,k+1}(s) - Z_{n,k+2}(s))ds + \sqrt{n}M_{n,k+1}(t),
\]
For \( i > k + 1 \)
\[
Z_{n,i}(t) = Z_{n,i}(0) + \int_0^t \dot{\beta}_n(\mu_{n,k})(1 + \Delta_{n,i}(Z_{n,i}(s)))Z_{n,i}(s)ds \\
- \int_0^t \Delta_{n,i-1}(Z_{n,i-1}(s))Z_{n,i-1}(s)ds - \int_0^t \Delta_{n,i}(Z_{n,i}(s))Z_{n,i}(s)ds \\
- \int_0^t (Z_{n,i}(s) - Z_{n,i+1}(s))ds + \sqrt{n}M_{n,i}(t),
\]
where \( \Delta_{n,i} \) is as in Remark 7.2

Finally, if \( Y_{n,1} = \sum_{i=1}^{k} Z_{n,i} \), then
\[
Y_{n,1}(t) = Y_{n,1}(0) - \int_0^t \dot{\beta}_n(\mu_{n,k})(1 + \Delta_{n,k}(Z_{n,k}(s)))Z_{n,k}(s)ds \\
- \int_0^t (Z_{n,1}(s) - Z_{n,k+1}(s))ds + \sum_{i=1}^{k} \sqrt{n}M_{n,i}(t)
\]
(7.5)

**Lemma 7.4.** Suppose \( \lambda_n \not\to 1 \) and \( 1 \ll d_n \ll n \). Assume that for some \( k \geq 2 \), \( \mu_{n,k} \to 1 \) and \( \dot{\beta}_n(\mu_{n,k}) \to \alpha \in [0, \infty) \) as \( n \to \infty \). Define the \( k-1 \times k-1 \) tridiagonal matrix \( Q_n(s) \) as
\[
Q_n(s)[j,j] = \dot{\beta}_n(\mu_{n,j})(1 + \Delta_{n,j}(Z_{n,j}(s))) + 1, \quad 1 \leq j \leq k-1, \\
Q_n(s)[j,j+1] = -1, \quad 1 \leq j \leq k-2, \\
Q_n(s)[j,j-1] = -\dot{\beta}_n(\mu_{n,j-1})(1 + \Delta_{n,j-1}(Z_{n,j-1}(s))), \quad 2 \leq j \leq k-1,
\]
and for all other \( j,k \), \( Q_n(s)[j,k] = 0 \). Then for any \( T,L \in (0, \infty) \)
\[
\lim_{n \to \infty} \inf_{s \in [0,T \wedge \tau_{n,L}]} \inf_{\|x\|^2} \beta^T Q_n(s) x = +\infty \quad \text{a.s.}
\]

**Proof.** Let \( h_{n,i}(s) \overset{\triangle}{=} \dot{\beta}_n(\mu_{n,i})(1 + \Delta_{n,i}(Z_{n,i}(s))) + 1 \) and \( H_n(s) \overset{\triangle}{=} Q_n(s) + Q_n(s)^T \). Then \( H_n(s) \) is a symmetric tridiagonal matrix with entries
\[
H_n(s)[j,j] = 2h_{n,j}(s), \quad 1 \leq j \leq k-1, \\
H_n(s)[j,j+1] = -h_{n,j}(s), \quad 1 \leq j \leq k-2, \\
H_n(s)[j,j-1] = -h_{n,j-1}(s), \quad 2 \leq j \leq k-1,
\]
(7.7)

Let \( h_n \overset{\triangle}{=} \dot{\beta}_n(\mu_{n,1}) \). By Lemma 5.6, \( h_n \to \infty \) and by the uniform convergence in [7.4] and Lemma 5.6 once more
\[
\max_{i \leq k-1} \sup_{s \in [0,T \wedge \tau_{n,L}]} \left| \frac{h_{n,i}(s)}{h_n} - 1 \right| \to 0 \quad \text{as} \quad n \to \infty \quad \text{a.s.}
\]
This in particular shows that
\[
\sup_{s \in [0,T \wedge \tau_{n,L}]} \left\| \frac{1}{h_n} H_n(s) - H \right\|_F \to 0 \quad \text{a.s.,}
\]
(7.8)
where \( \|\cdot\|_F \) is the Frobenius norm and \( H \) is the \( k-1 \times k-1 \) tridiagonal matrix given as
\[
H[j,j] = 2, \quad 1 \leq j \leq k-1, \\
H[j,j+1] = -1, \quad 1 \leq j \leq k-2, \\
H[j,j-1] = -1, \quad 2 \leq j \leq k-1,
\]
Note for any $\vec{x} = (x_1, x_2, \ldots, x_{k-1}) \in \mathbb{R}^{k-1}$ by completing squares
\[ \vec{x}^t H \vec{x} = x_1^2 + (x_2 - x_1)^2 + (x_3 - x_2)^2 + \cdots + (x_{k-2} - x_{k-1})^2 + x_{k-1}^2, \]
which is strictly positive if $\vec{x} \neq 0$. Let $c = \inf_{\|\vec{x}\|=1} \vec{x}^t H \vec{x}$. Since the unit sphere is compact, the infimum is attained and hence $c > 0$. This shows that $H$ is a positive definite matrix.

Finally, note that for any $s \geq 0$
\[ \vec{x}^t \frac{1}{h_n} H_\mathcal{n}(s) \vec{x} = \vec{x}^t H \vec{x} + \vec{x}^t \left( \frac{1}{h_n} H_\mathcal{n}(s) - H \right) \vec{x} \]
\[ \geq \vec{x}^t H \vec{x} - \|h_n^{-1} H_\mathcal{n}(s) - H\|_F \|\vec{x}\|^2 \geq (c - \|h_n^{-1} H_\mathcal{n}(s) - H\|_F) \|\vec{x}\|^2. \]

On taking infimum and using $\vec{x}^t H_\mathcal{n}(s) \vec{x} = 2 \vec{x}^t Q_n(s) \vec{x}$, this shows
\[ \inf_{s \in [0, T]} \inf_{x \in \mathbb{R}^k \setminus \{0\}} \frac{\|2 \vec{x}^t Q_n(s) \vec{x}\|}{\|\vec{x}\|^2} \geq \left( c - \sup_{s \in [0, T]} \|h_n^{-1} H_\mathcal{n}(s) - H\|_F \right) h_n. \]
As $n \to \infty$, the convergence in (7.8) and the divergence $h_n \to +\infty$ now completes the proof.

**Remark 7.5.** For every $s > 0$, the $k - 1 \times k - 1$ matrix $Q_n(s)$ appearing in the previous lemma is the drift operator that appears in the right hand side of the first $k - 1$ coordinates in Corollary 7.3. More precisely, for each $t \geq 0$:
\[ \vec{X}_n(t) = \vec{X}_n(0) - \int_0^t Q_n(s) \vec{X}_n(s) ds + \vec{c}_{k-1} \cdot \int_0^t Z_{n,k}(s) ds + \vec{W}_n(t). \]
where $\vec{X}_n \equiv (Z_{n,1}, Z_{n,2}, \ldots, Z_{n,k-1})$, $\vec{W}_n \equiv (\sqrt{n}M_{n,1}, \ldots, \sqrt{n}M_{n,k-1})$ and $\vec{c}_{k-1}$ is the vector $(0, 0, \ldots, 0, 1)^t \in \mathbb{R}^{k-1}$.

**Lemma 7.6.** Suppose that the hypothesis of Theorem 2.2 holds with $k \geq 2$ and let $\vec{X}_n \equiv (Z_{n,1}, Z_{n,2}, \ldots, Z_{n,k-1})$. Then for $L, T, \epsilon \in (0, \infty)$
\[ P \left( \sup_{s \in [0, T]} \|\vec{X}_n(s)\| > \|\vec{X}_n(0)\| + \epsilon \right) \to 0, \]
and
\[ \sup_{s \in [\epsilon, T]} \|\vec{X}_n(s)\| \overset{P}{\to} 0, \]
as $n \to \infty$.

**Proof.** Applying Itô’s formula (see [35, Section II.7]) to the function $h(\vec{x}) = \|\vec{x}\|^2$ and the semi-martingale representation of $\vec{X}_n$ from (7.9) in Remark 7.5, we get
\[ \|\vec{X}_n(t)\|^2 = \|\vec{X}_n(0)\|^2 + 2 \int_0^t \langle \vec{X}_n(s-), d\vec{X}_n(s) \rangle + [\vec{W}_n]_t \]
\[ = \|\vec{X}_n(0)\|^2 - 2 \int_0^t \langle \vec{X}_n(s), Q_n(s) \vec{X}_n(s) \rangle ds + 2 \int_0^t Z_{n,k}(s) \langle \vec{X}_n(s), \vec{c}_{k-1} \rangle ds \]
\[ + 2 \int_0^t \langle \vec{X}_n(s-), d\vec{W}_n(s) \rangle + [\vec{W}_n]_t, \]
(7.12)
where $[\bar{W}_n]^*_T \doteq \sum_{i=1}^{k-1} [\sqrt{n}M_{n,i}]_T$. Let

$$h_n(s) \doteq \frac{\overline{X}_n(s)Q_n(s)\overline{X}_n(s)}{\|\overline{X}_n(s)\|^2} \{\overline{X}_n(s) \neq 0\} + n\{\overline{X}_n(s) = 0\}$$

$$g_n(s) \doteq 2Z_{n,k}(s)Z_{n,k-1}(s)$$

$$R_n(s) \doteq 2\int_0^t \langle \overline{X}_n(s), d\overline{W}_n(s) \rangle + [\overline{W}_n]^*_T$$

then (7.12) becomes

$$\left\| \overline{X}_n(t) \right\|^2 = \left\| \overline{X}_n(0) \right\|^2 - 2 \int_0^t h_n(s) \left\| \overline{X}_n(s) \right\|^2 ds + \int_0^t g_n(s) ds + R_n(t). \quad (7.13)$$

Further by Lemma 7.4

$$m_n \doteq \inf_{s \in [0,T \wedge \tau_n,L]} h_n(s) \to +\infty \text{ a.s. as } n \to \infty,$$  

(7.14)

and by Doob’s inequality and Itô’s isometry (see e.g. [35, Corollary 3, Section II.7]), for $i \leq k - 1$

$$E \sup_{s \in [0,T \wedge \tau_n,L]} \left| \int_0^t Z_{n,i}(s-)d(\sqrt{n}M_{n,i})_s \right|^2 \leq 4E \int_0^{T \wedge \tau_n,L} Z_{n,i}^2(s-)d[\sqrt{n}M_{n,i}]_s$$

$$\leq 4L^2 E[\sqrt{n}M_{n,i}]_T = 4L^2 E \left\langle \sqrt{n}M_{n,i} \right\rangle_T.$$

where the second to last inequality is obtained by using $\|Z_n\|_{2,T \wedge \tau_n,L} \leq L$. From the proof of Lemma 6.2 we see that for any $i \leq k - 1$, $E \left\langle \sqrt{n}M_{n,i} \right\rangle_T = E[\sqrt{n}M_{n,i}]_T \to 0$ as $n \to \infty$. Along with the above display, this shows that the two terms appearing in the definition of $R_n$ are converging to zero, and hence

$$\left| R_n \right|_{s,T \wedge \tau_n,L} \xrightarrow{P} 0 \text{ as } n \to \infty. \quad (7.15)$$

Applying Lemma 6.8 to (7.13) with $z(t) = \left\| \overline{X}_n(t) \right\|^2$, $h = 2h_n$, $g = g_n$, $M = R_n$, and $\tau = \tau_{n,L}$ shows for any $t \in [0,T \wedge \tau_{n,L}]

$$\|\overline{X}_n(t)\|^2 \leq \frac{|g_n|_{s,T \wedge \tau_{n,L}}}{2m_n} + 2|R_n|_{s,T \wedge \tau_{n,L}} + e^{-2m_n t} \left\| \overline{X}_n(0) \right\|^2.$$

Taking $t = \epsilon_n \doteq 1/\sqrt{m_n}$ and using (7.14), (7.15), $g_n|_{s,T \wedge \tau_{n,L}} \leq 2L^2$ and $\overline{X}_n(0) \xrightarrow{P} (z_1, \ldots, z_{k-1})^T$, we see that

$$\sup_{t \in [\epsilon_n,T \wedge \tau_{n,L}]} \left\| \overline{X}_n(t) \right\| \xrightarrow{P} 0. \quad (7.16)$$

Since $\epsilon_n \to 0$, this shows (7.11) for any fixed $\epsilon > 0$. Finally, from (7.13), we see that

$$\sup_{t \in [0,\epsilon_n \wedge \tau_{n,L} \wedge T]} \left\| \overline{X}_n(t) \right\|^2 \leq \left\| \overline{X}_n(0) \right\|^2 + |g_n|_{s,T \wedge \tau_{n,L}} \epsilon_n + |R_n|_{s,T \wedge \tau_{n,L}}.$$

Since we have already shown (7.16), the convergence in (7.10) is now immediate on using that $\epsilon_n \to 0$, $|g_n|_{s,T \wedge \tau_{n,L}} \leq 2L^2$ and that (7.15) holds.

Corollary 7.7. Under the assumptions of Lemma 7.6, for each $i < k$, $\int_0^{T \wedge \tau_{n,L}} |Z_{n,i}(s)| ds \xrightarrow{P} 0$, as $n \to \infty$.

Proof. For any $\epsilon > 0$

$$\int_0^{T \wedge \tau_{n,L}} |Z_{n,i}(s)| ds \leq \int_{[0,\epsilon \wedge \tau_{n,L}]} |Z_{n,i}(s)| ds + \int_{[\epsilon,T \wedge \tau_{n,L}]} |Z_{n,i}(s)| ds \leq L \epsilon + \sup_{s \in [\epsilon,T \wedge \tau_{n,L}]} |Z_{n,i}(s)| T.$$
Now fix $\delta > 0$ and let $\epsilon = \frac{\delta}{2T}$. Then for any $i < k$

$$P\left(\int_0^{T \wedge \tau_{n,L}} |Z_{n,i}(s)| ds > \delta \right) \leq P\left( \sup_{s \in [0,T \wedge \tau_{n,L}]} |Z_{n,i}(s)| > \frac{\delta}{2T} \right),$$

(7.17)

which from (7.11) converges to 0 as $n \to \infty$. Since $\delta > 0$ was arbitrary, this completes the proof. \hfill \blacksquare

**Proof of Theorem 2.2** Recall the conditions in the theorem. By Remark 2.5(ii) and the tightness of $\{\|Z_n(0)\|_1\}_{n \in \mathbb{N}}$, the hypothesis of Lemma 6.2 holds. Hence by Skorokhod’s embedding theorem, we can assume that $\{(Z_n(0), M_n)\}_{n \in \mathbb{N}}$ and a standard Brownian motion $B$ are defined on a common probability space such that for any $T > 0$

$$\sup_{t \leq T} \left\| \sqrt{n}M_n(t) - \sqrt{2}B(t)e_k \right\|_2 \to 0$$

(7.18)

and

$$\|Z_n(0) - z\|_2 \to 0 \quad \text{a.s.,}$$

(7.19)

as $n \to \infty$. Let $Y$ and $Y_n$ be as in the statement of the theorem. Taking $m = r - k + 1$, let $Y_n = (\sum_{i=1}^k Z_{n,i}, Z_{n,k+1}, \ldots, Z_{n,r})$ be the stochastic process with sample paths in $\mathbb{D}([0,T] : \mathbb{R}^m)$ corresponding to the first $m$ coordinates of $Y_n$. Note $Y_{n,m+} = Z_{n,r+}$, $Z_{n,k} = Y_{n,1} - \sum_{i=1}^{k-1} Z_{n,i}$, and for $k = 1, Y_{n,1} = Z_{n,1}$. Hence by Corollary 7.3 $\tilde{Y}_n$ satisfy

$$Y_{n,1}(t) = Y_{n,1}(0) - \int_0^t a_{n,k}(s)Y_{n,1}(s)ds - \|_{\{k=1\}} \int_0^t Y_{n,1}(s)ds + \int_0^t Y_{n,2}(s)ds + \sqrt{n}M_{n,k}(t)$$

$$+ \sum_{i=1}^{k-1} \int_0^t a_{n,k}(s)Z_{n,i}(s)ds - \|_{\{k>1\}} \int_0^t Z_{n,1}(s)ds + \sum_{i=1}^{k-1} \sqrt{n}M_{n,i}(t),$$

(7.20)

$$Y_{n,2}(t) = Y_{n,2}(0) + \int_0^t a_{n,k}(s)Y_{n,1}(s)ds - \int_0^t Y_{n,2}(s)ds + \int_0^t Y_{n,3}(s)ds$$

$$- \sum_{i=1}^{k-1} \int_0^t a_{n,k}(s)Z_{n,i}(s)ds - \int_0^t \delta_{n,k+1}(s)Y_{n,2}(s)ds + \sqrt{n}M_{n,k+1}(t),$$

(7.21)

and for $i \in \{3,4 \ldots m\}$

$$Y_{n,i}(t) = Y_{n,i}(0) - \int_0^t Y_{n,i}(s)ds + \int_0^t Y_{n,i+1}(s)ds$$

$$+ \int_0^t \delta_{n,k+i-2}(s)Y_{n,i+1}(s)ds - \int_0^t \delta_{n,k+i-1}(s)Y_{n,i}(s)ds + \sqrt{n}M_{n,k+i-1}(t).$$

(7.22)

where $a_{n,k}(s) \equiv \beta_n(\mu_{n,k})(1 + \Delta_{n,k}(Z_{n,k}(s)))$ and $\delta_{n,i}(s) \equiv \Delta_{n,i}(Z_{n,i}(s))$ for $i \in \mathbb{N}$.

Since $\|Z_n\|_{2,T \wedge \tau_{n,L}} \leq L$, we have by (7.14) that, for any $i \in \mathbb{N}$,

$$|\delta_{n,i}|_{s,T \wedge \tau_{n,L}} \leq \gamma_{n,L} \to 0 \quad \text{a.s. as } n \to \infty.$$  

(7.23)

Moreover since $\beta_n(\mu_{n,k}) \to \alpha \in [0,\infty)$, this also shows that

$$\sup_{s \in [0,T \wedge \tau_{n,L}]} |a_{n,k}(s) - \alpha| \to 0 \quad \text{a.s. as } n \to \infty.$$  

(7.24)

We now show that

$$\|Y_n - Y\|_{2,T \wedge \tau_{n,L}} \overset{P}{\to} 0 \quad \text{as } n \to \infty.$$  

(7.25)
To see this, note that, by Remark $[2.5][i]$, the hypothesis of Lemma $[6.4]$ is satisfied, and hence $\|Z_{n,r+}\|_{2,T \wedge \tau_{n,L}} \overset{P}{\to} 0$. Since $Y_{n,m+} = Z_{n,r+}$ and $Y_{m+} = 0$, this shows that
\[ \|Y_{n,m+} - Y_{m+}\|_{2,T \wedge \tau_{n,L}} \overset{P}{\to} 0. \]  
(7.26)

Thus in order to prove $(7.25)$ it suffices to show that $\sum_{n=1}^{\infty} |Y_{n,i} - Y_{i}|_{s,T \wedge \tau_{n,L}} \overset{P}{\to} 0$ as $n \to \infty$. To show this we consider $U_{n,i} = Y_{n,i} - Y_{i}$. Subtracting $(2.9)$ from $(7.20)$, $(7.21)$ and $(7.22)$, we see
\[ U_{n,1}(t) = U_{n,1}(0) - (\alpha + \|_{k=1}) \int_0^t U_{n,1}(s)ds + \int_0^t U_{n,1}(s)ds + \sqrt{n}M_{n,k}(t) - \sqrt{2}B(t) + W_{n,1}(t) \]
\[ U_{n,2}(t) = U_{n,2}(0) + \alpha \int_0^t U_{n,1}(s)ds + \int_0^t U_{n,2}(s)ds + \int_0^t U_{n,3}(s)ds + W_{n,2}(t) \]
\[ U_{n,i}(t) = U_{n,i}(0) - \int_0^t U_{n,i}(s)ds + \int_0^t U_{n,i+1}(s)ds + W_{n,i}(t) \quad \text{for } i \in \{3, 4, \ldots, m\} \]
(7.27)

where
\[ W_{n,1}(t) \overset{P}{\to} 0 \quad \text{as } n \to \infty \]
(7.28)

for each $i \in [m]$. Let $\|U_{n}\|_{1,t} = \sup_{s \in [0,t]} \sum_{i=1}^{m} |U_{n,i}(t)|$. Then, from $(7.27)$, for any $t \in [0, T \wedge \tau_{n,L}]$
\[ \|U_{n}\|_{1,t} \leq \sum_{i=1}^{m} \left( |U_{n,i}(0)| + |W_{n,i}|_{s,T \wedge \tau_{n,L}} \right) + \left| \sqrt{n}M_{n,k} - \sqrt{2}B \right|_{s,T} + R \int_0^t \|U_{n}\|_{1,s} ds \]
with $R = \max(2\alpha + \|_{k=1}, 2)$. Hence by Gronwall’s inequality
\[ \|U_{n}\|_{1,T \wedge \tau_{n,L}} \leq \left( \left| \sqrt{n}M_{n,k} - \sqrt{2}B \right|_{s,T} + \sum_{i=1}^{m} \left( |U_{n,i}(0)| + |W_{n,i}|_{s,T \wedge \tau_{n,L}} \right) \right) e^{RT}. \]

By our hypothesis, as $n \to \infty$, $|U_{n,i}(0)| = |Z_{n,k,i-1}(0) - Z_{n,k+i-1}| \overset{P}{\to} 0$ for each $i \in [m]$. Hence by $(7.28)$ and $(7.18)$, $\|U_{n}\|_{1,T \wedge \tau_{n,L}} \overset{P}{\to} 0$ as $n \to \infty$. Combined with $(7.26)$, this completes the proof of $(7.25)$.

Next we prove $(2.8)$. Fix $\delta > 0$. Since $Y$ has sample paths in $C([0,T] : \ell_2)$, we can find $L_1 \in (0, \infty)$ so that
\[ P\left( \|Y\|_{2,T} > L_1 \right) \leq \frac{\delta}{2}. \]  
(7.29)

Also, since $Z_{n}(0) \overset{P}{\to} z$, we can find a $L_2 \in (0, \infty)$ so that
\[ \sup_{n} P\left( \|Z_{n}(0)\|_{2} > L_2 \right) \leq \frac{\delta}{2}. \]  
(7.30)
Lemma 8.1. Suppose for $Z$ then taking $\lim \inf_{n} X_{2,T \wedge \tau_{n,L}} \leq 0$. Then, $k \| \langle\{k > \| X_{n} \rangle \|_{2,T \wedge \tau_{n,L}} + \| Y_{n} \|_{2,T \wedge \tau_{n,L}}$.

Hence for each $n \in \mathbb{N}$ $P(\tau_{n,L} \leq T) \leq P(\| Z_{n} \|_{2,T \wedge \tau_{n,L}} > L - 1) \leq P(\| Y_{n} \|_{2,T \wedge \tau_{n,L}} > L + 1) + P(\| X_{n} \|_{2,T \wedge \tau_{n,L}} > L + 1)$, $\leq \delta + P(\| Y_{n} - Y \|_{2,T \wedge \tau_{n,L}} > 1) + P(\| X_{n} \|_{2,T \wedge \tau_{n,L}} > \| X_{n}(0) \| + 1)$, where the last inequality uses (7.29) and (7.30). From Lemma 7.6 and (7.25) we see $\limsup_{n \to \infty} P(\| Z_{n} \|_{2,T} \geq L) \leq \limsup_{n \to \infty} P(\tau_{n,L} \leq T) \leq \delta$.

Since $\delta > 0$ is arbitrary, the convergence in (2.8) is now immediate.

This convergence in particular says that $\lim_{L \to \infty} \sup_{n} P(\tau_{n,L} \leq T) = 0$. Using Lemma 6.5 with $F_{n}(t) = \| Y_{n} - Y \|_{2,T}$ we now see from (7.25) that $\| Y_{n} - Y \|_{2,T} \to 0$ as $n \to \infty$. Similarly, if $k > 1$, then taking $F_{n}(t) = \sup_{s \in [t]} |Z_{n,i}(s)|$ in Lemma 6.5 we conclude from Lemma 7.6 that for each $i \in [k - 1]$ and $\epsilon > 0 \sup_{s \in [t]} |Z_{n,i}(s)| \to 0$ as $n \to \infty$. This completes the proof of Theorem 2.3.

8. Proof of Theorem 2.3

In this section we give the proof of Theorem 2.3. We begin by giving a convenient representation for $Z_{n}$ under the assumptions of Theorem 2.3 and establishing some apriori convergence properties.

Lemma 8.1. Suppose $c_{n} = \frac{d_{n}}{\sqrt{n}} \to c \in (0, \infty)$ and $\lambda_{n} = 1 - \left( \frac{\log d_{n}}{d_{n}} + \frac{\alpha_{n}}{\sqrt{n}} \right)$ where $\alpha_{n} \in \mathbb{R}$, $\liminf_{n \to \infty} \alpha_{n} > -\infty$ and $\frac{\alpha_{n}}{\sqrt{n}} \to 0$. Suppose also that $\{\| Z_{n}(0) \|_{1}\}_{n \in \mathbb{N}}$ is a tight sequence of random variables and $Z_{n,r+1}(0) \to 0$ in $\ell_{2}$ for some $r \geq 2$. Then there are stochastic processes $\delta_{n}, \{W_{n,i}\}_{i=2}$ with sample paths in $\mathbb{D}([0, \infty) : \mathbb{R})$ such that for any $t \geq 0$

$$Z_{n,1}(t) = Z_{n,1}(0) - \int_{0}^{t} Z_{n,1}(s)ds + \int_{0}^{t} Z_{n,2}(s)ds + \sqrt{n}M_{n,1}(t)$$

$$- (c_{n}e^{\sqrt{n}\alpha_{n}})^{-1} \int_{0}^{t} (1 + \delta_{n}(s)) \left( e^{c_{n}Z_{n,1}(s)} - 1 \right)ds$$

$$Z_{n,2}(t) = Z_{n,2}(0) - \int_{0}^{t} Z_{n,2}(s)ds + \int_{0}^{t} Z_{n,3}(s)ds + W_{n,2}(t)$$

$$+ (c_{n}e^{\sqrt{n}\alpha_{n}})^{-1} \int_{0}^{t} (1 + \delta_{n}(s)) \left( e^{c_{n}Z_{n,1}(s)} - 1 \right)ds$$

$$Z_{n,i}(t) = Z_{n,i}(0) - \int_{0}^{t} Z_{n,i}(s)ds + \int_{0}^{t} Z_{n,i+1}(s)ds + W_{n,i}(t) \quad \text{for } i \in \{3, \ldots, r\}$$

and for any fixed $L, T \in (0, \infty)$,

(1) $\sqrt{n}M_{n,1} \Rightarrow \sqrt{2}B$ in $\mathbb{D}([0, \infty) : \mathbb{R})$ where $B$ is a standard Brownian motion,
(2) $|\delta_{n}|_{s,T_{n}} \to 0$ a.s.
(3) $|W_{n,i}|_{s,T_{n}} \to 0$ for $i \in \{2, \ldots, r\}$,
Lemma 8.2. Suppose differential equations in which the drift fails to satisfy a linear growth condition. From Lemma 6.1 it follows that (8.1) is satisfied. Lemma 6.6 shows that $Z$ is defined as in (6.8).

Proof. Recall the definition of $q_{n,i}$ from Lemma 6.1 Define

$$\delta_n(s) = q_{n,1}(Z_{n,1}(s))c_n \left( e^{c_n|Z_{n,1}(s)|} - e^{-c_n\alpha_n} \right)^{-1} - 1$$

so that

$$q_{n,1}(Z_{n,1}(s)) = (1 + \delta_n(s))c_n^{-1} \left( e^{c_n|Z_{n,1}(s)|} - e^{-c_n\alpha_n} \right).$$

Since $\sup_{s \in T \cap \tau_{n,L}} |Z_{n,1}(s)| \leq L$, Lemma 6.7 shows that $|\delta_n(s)| \to 0$ a.s.

$$W_{n,2}(t) = -\int_0^t q_{n,2}(Z_{n,2}(s))ds + \sqrt{n}M_{n,2}(t)$$

$$W_{n,i}(t) = \int_0^t q_{n,i-1}(Z_{n,i-1}(s))ds - \int_0^t q_{n,i}(Z_{n,i}(s))ds + \sqrt{n}M_{n,i}(t) \quad \text{for } i \in \{3, \ldots, r\}.$$

From Lemma 6.1 it follows that (8.1) is satisfied. Lemma 6.6 shows that $\mu_n \to f_1 \in \ell_1^2$. Along with the assumed tightness of $\{\|Z_n(0)\|\}_{n \in \mathbb{N}}$, this shows $G_n(0) = \mu_n + \frac{Z_n(0)}{\sqrt{n}} \to f_1$ in $\ell_1^2$. Hence by Lemma 6.2 and Lemma 6.4

$$\sqrt{n}M_n \Rightarrow \sqrt{2}B_e_1 \text{ in } \mathbb{D}([0, \infty) : \ell_2)$$

and $\|Z_{n,r+}\|_{2,T \cap \tau_{n,L}} \to 0$ as $n \to \infty$. Since $|Z_{n,i,s,T \cap \tau_{n,L}} \leq L$ and $\mu_{n,2} \to 0$, Lemma 6.3, together with (8.2), shows that $|W_{n,i,s,T_n} \to 0$ for each $i \in \{2, \ldots, r\}$, as $n \to \infty$.

The next lemma gives pathwise existence and uniqueness of solutions to a system of stochastic differential equations in which the drift fails to satisfy a linear growth condition.

Lemma 8.2. Suppose $c \in (0, \infty)$, $\alpha \in (0, \infty]$ and $B$ is a standard Brownian motion. Then for any $r \geq 2$ the system of equations

$$Z_1(t) = z_1 - \int_0^t Z_1(s)ds + \int_0^t Z_2(s)ds + \sqrt{2}B(t) - (ce^{\alpha})^{-1} \int_0^t \left( e^{cZ_1(s)} - 1 \right)ds$$

$$Z_2(t) = z_2 - \int_0^t Z_2(s)ds + \int_0^t Z_3(s)ds + (ce^{\alpha})^{-1} \int_0^t \left( e^{cZ_1(s)} - 1 \right)ds$$

$$Z_i(t) = z_i - \int_0^t Z_i(s)ds + \int_0^t Z_{i+1}(s)ds \quad \text{for } i \in \{3, \ldots, r\}$$

$$Z_i(t) = 0 \quad \text{for } i > r$$

has a unique pathwise solution $Z$ with sample paths in $C([0, \infty) : \ell_2)$ for any $(z_1, \ldots, z_r) \in \mathbb{R}^r$.

Proof. The case when $\alpha = \infty$ is standard and is thus omitted. Consider now the case $\alpha < \infty$. It is straightforward to see that there is a unique $Z_{2+} = (Z_3, Z_4, \ldots)$ in $C([0, \infty) : \ell_2)$ that solves the last two equations in (8.3). Hence it suffices to show that, the system of equations

$$Z_1(t) = z_1 - (ce^{\alpha})^{-1} \int_0^t \left( e^{cZ_1(s)} - 1 \right)ds + \int_0^t (Z_2(s) - Z_1(s))ds + \sqrt{2}B(t)$$

$$Z_2(t) = z_2 + (ce^{\alpha})^{-1} \int_0^t \left( e^{cZ_1(s)} - 1 \right)ds - \int_0^t Z_2(s)ds + \int_0^t h(s)ds$$

has a unique pathwise solution $(Z_1, Z_2)$ with sample paths in $C([0, \infty) : \mathbb{R}^2)$ where $h = Z_3 \in C([0, \infty) : \mathbb{R})$ is a given (non-random) continuous trajectory and $(z_1, z_2) \in \mathbb{R}^2$. 


Define $y_1 = z_1$, $y_2 = z_1 + z_2$ and consider the equation:

\[
Y_1(t) = y_1 - (ce^{\alpha t})^{-1} \int_0^t e^{Y_1(s)} ds + \int_0^t (Y_2(s) - 2Y_1(s)) ds + \sqrt{2}B(t)
\]
\[
Y_2(t) = y_2 - \int_0^t Y_1(s) ds + \int_0^t h(s) ds + \sqrt{2}B(t).
\]

Note that $(Z_1, Z_2)$ solve (8.4) if and only if $(Y_1, Y_2)$, with $Y_1 = Z_1$ and $Y_2 = Z_1 + Z_2$ solve (8.5). Thus it suffices to prove existence and uniqueness of solutions for (8.5).

For $L \in (0, \infty)$, let $\eta_L : \mathbb{R} \to [0, 1]$ be such that $\eta_L$ is smooth, $\eta_L(x) = 1$ for $|x| \leq L$ and $\eta_L(x) = 0$ for $|x| \geq L + 1$. Consider the equation

\[
Y_1^L(t) = y_1 - (ce^{\alpha t})^{-1} \int_0^t e^{Y_1^L(s)} \eta_L(Y_1^L(s)) ds + (ce^{\alpha t})^{-1} t + \int_0^t (Y_2^L(s) - 2Y_1^L(s)) ds + \sqrt{2}B(t)
\]
\[
Y_2^L(t) = y_2 - \int_0^t Y_1^L(s) ds + \int_0^t h(s) ds + \sqrt{2}B(t).
\]

Since for each $L$ (8.6) is an equation with (globally) Lipschitz coefficients, by standard results, it has a unique pathwise continuous solution.

Fix $T \in (0, \infty)$ and let $\tau_L = \inf\{t \geq 0 : |Y_1^L(t)| \geq L\} \wedge T$ for any $L > 0$. Then by pathwise uniqueness of (8.6) for $0 \leq t \leq \tau_L \wedge \tau_{L+1}$,

\[
Y^L(t) = Y^{L+1}(t).
\]

This in particular shows that, $\tau_L \leq \tau_{L+1}$ a.s.

We now estimate the second moment of $|Y_1^L(t)|$. By Itô’s formula

\[
(Y_1^L(t))^2 = (y_1)^2 - 2(ce^{\alpha t})^{-1} \int_0^t Y_1^L(s) e^{Y_1^L(s)} \eta_L(Y_1^L(s)) ds + 2(ce^{\alpha t})^{-1} \int_0^t Y_1^L(s) ds
\]
\[
+ 2 \int_0^t Y_1^L(s)(Y_2^L(s) - 2Y_1^L(s)) ds + 2\sqrt{2} \int_0^t Y_1^L(s) dB(s) + 2t
\]
\[
(Y_2^L(t))^2 = (y_2)^2 - 2 \int_0^t Y_1^L(s) Y_2^L(s) ds + 2 \int_0^t Y_2^L(s) h(s) ds + 2\sqrt{2} \int_0^t Y_2^L(s) dB(s) + 2t.
\]

Thus

\[
(Y_1^L(t))^2 + (Y_2^L(t))^2 = (y_1)^2 + (y_2)^2 - 2(ce^{\alpha t})^{-1} \int_0^t Y_1^L(s) e^{Y_1^L(s)} \eta_L(Y_1^L(s)) ds
\]
\[
+ 2(ce^{\alpha t})^{-1} \int_0^t Y_1^L(s) ds + 2 \int_0^t Y_2^L(s) h(s) ds
\]
\[
- 4 \int_0^t (Y_1^L(s))^2 ds + 2\sqrt{2} \int_0^t (Y_1^L(s) + Y_2^L(s)) dB(s) + 4t.
\]

Since $c > 0$, we have on using the inequality $|x| \leq 1 + |x|^2$ that $-xe^{\alpha t}\eta_L(x) \leq (1 + |x|^2)$ for all $x \in \mathbb{R}$. Thus with $\|Y^L\|_{s,t} = \sup_{s \leq t} \|Y^L(s)\|:

\[
\|Y^L\|^2_{s,t} \leq (y_1)^2 + 4(ce^{\alpha t})^{-1} \int_0^t (1 + \|Y^L\|^2_{s,s}) ds + 2 \int_0^t (1 + \|Y^L\|^2_{s,s}) |h(s)| ds
\]
\[
+ 2\sqrt{2} \left( 1 + \sup_{0 \leq s \leq t} \int_0^s (Y_1^L(u) + Y_2^L(u)) dB(u) \right) + 4t.
\]
Taking expectations and using Doob’s inequality and Itô’s isometry to compute the expectation over the supremum:

\[
E\|Y^L\|^2_{s,T} \leq \|y\|^2 + (4e^{c\alpha})^{-1} + 2|h|_{s,T} \int_0^t (1 + E\|Y^L\|^2_{s,s}) ds \\
+ 2\sqrt{2} \left( 1 + 4E \int_0^t Y_1^L(u) + Y_2^L(u)^2 du \right) + 4t \\
\leq (\|y\|^2 + K(T + 1)) + K \int_0^t E\|Y^L\|^2_{s,s} ds.
\]

with \( K \equiv 4(e^{c\alpha})^{-1} + 2|h|_{s,T} + 16\sqrt{2} \) for any \( t \in [0,T] \). By Gronwall lemma, for every \( L \in \mathbb{N} \)

\[
E\|Y^L\|^2_{s,T} \leq (\|y\|^2 + K(T + 1))e^{KT} = c_1.
\]

Thus, as \( L \to \infty \)

\[
P(\tau_L < T) \leq P(\|Y^L\|_{s,T} \geq L) \leq c_1/L^2 \to 0
\]

and consequently \( \tau_L \uparrow T \) a.s. as \( L \to \infty \). Now define \( Y(t) \equiv Y^L(t) \) for \( 0 \leq t \leq \tau_L \). Then \( Y \) is a solution of (8.5) on \([0,T)\). The same argument as before shows that this is the unique pathwise solution on \([0,T)\). Since \( T \) is arbitrary we get a unique pathwise solution of (8.5) on \([0,\infty)\). This completes the proof of the lemma.

**Lemma 8.3.** Suppose the assumptions of Theorem 2.3 hold. Suppose further that \( Z_n(0), M_n \) and a standard Brownian motion \( B \) are given on a common probability space such that \( Z_n(0) \to z \) in \( \ell_1^\top \) and \( M_n \to \sqrt{2}Be_1 \) in \( \mathbb{D}([0,\infty) : \ell_2) \) a.s. Let \( Z \) be as defined in Lemma 8.2. Then for any \( T,L \in (0,\infty) \)

\[
\|Z_n - Z\|_{2,T \land \tau_n, L \land \tau_L} \to 0 \quad \text{as } n \to \infty.
\]

(8.7)

where \( \tau_L \equiv \inf\{ t \mid \|Z(t)\|_{2,t} > L \} \).

**Proof.** Fix \( L, T \in (0,\infty) \) and let \( T_n \equiv T \land \tau_n, \land \tau_L \). Using the estimate \( |e^{ax} - e^{ay}| \leq ae^{a|x-y|} |x - y| \) for \( x, y \in \mathbb{R}, a \geq 0 \) and since \( |Z_{n,1}(s)|, |Z_1(s)| \leq L \) for any \( s \in [0,T_n] \), note

\[
\left| a_n(s)e^{c_nZ_{n,1}(s)} - ae^{cZ_1(s)} \right| \\
\leq \left| a_n(s)e^{c_nZ_{n,1}(s)} - a_n(s)e^{c_nZ_1(s)} \right| + \left| a_n(s)e^{c_nZ_1(s)} - a_n(s)e^{cZ_1(s)} \right| + \left| e^{cZ_1(s)} \right| |a_n(s) - a| \\
\leq |a_n(s)| c_n e^{cL} \left| U_{n,1}(s) \right| + |a_n(s)| Le^{L(c_n\vee c)} |c_n - c| + e^{cL} |a_n(s) - a|,
\]

where \( a_n(s) \equiv (c_n e^{c\alpha})^{-1}(1 + \delta_n(s)), c_n \equiv d_n/\sqrt{n} \to c, \delta_n \) is as in Lemma 8.1 \( a \equiv (ce^{\alpha})^{-1} \), and \( U_{n,i} \equiv Z_{n,i} - Z_i \) for \( i \in \mathbb{N} \). Since \( c_n \to c \) and \( |\delta_n|_{s,T_n} \to 0 \) a.s. by Lemma 8.1 \( |a_n - a|_{s,T_n} \to 0 \) a.s. Hence for any \( s \in [0,T_n] \)

\[
|a_n e^{c_nZ_{n,1}} - ae^{cZ_1}|_{s,s} \leq K |U_{n,1}|_{s,s} + r_n
\]

(8.8)

where \( K \equiv \sup_n \left( c_n e^{cL} |a_{n,1}|_{s,T_n} \right) \leq \infty \) a.s. and

\[
r_n \equiv |a_n|_{s,T_n} Le^{L(c_n\vee c)} |c_n - c| + e^{cL} |a_n - a|_{s,T_n} \to 0 \text{ a.s.}
\]
Subtracting \(8.3\) from \(8.1\), for any \(t > 0\),

\[
U_{n,1}(t) = U_{n,1}(0) - \int_0^t (U_{n,1}(s) - U_{n,2}(s)) ds + \sqrt{n} M_{n,1}(t) - \sqrt{2} B(t) \\
- \int_0^t (a_{n,1}(s) e^{c_n Z_{n,1}(s)} - ae^{Z_1(s)}) ds + \int_0^t (a_{n}(s) - a) ds
\]

\[
U_{n,2}(t) = U_{n,2}(0) - \int_0^t (U_{n,2}(s) - U_{n,3}(s)) ds + W_{n,2}(t) \\
+ \int_0^t (a_{n,1}(s) e^{c_n Z_{n,1}(s)} - ae^{Z_1(s)}) ds - \int_0^t (a_{n}(s) - a) ds
\]

\[
U_{n,i}(t) = U_{n,i}(0) - \int_0^t (U_{n,i}(s) - U_{n,i+1}(s)) ds + W_{n,i}(t) 	ext{ for } i \in \{3, \ldots, r\}.
\]

Let \(H_t = \sup_{s \in [0,t]} \sum_{i=1}^r |U_{n,i}(s)|\). Then from \(8.8\) and \(8.9\), for any \(t \in [0, T_n]\),

\[
H_t \leq H_0 + \left| \sqrt{n} M_{n,1} - \sqrt{2} B \right|_{*T} + 2T (|a_n - a|_{*T} + r_n) + \sum_{i=2}^r |W_{n,i}|_{*T_n} + |U_{n,r+1}|_{*T_n} + 2(1+K) \int_0^t H_s ds.
\]

Hence by Gronwall’s lemma

\[
H_{T_n} \leq \left( H_0 + \left| \sqrt{n} M_{n,1} - \sqrt{2} B \right|_{*T} + 2T (|a_n - a|_{*T} + r_n) + \sum_{i=2}^r |W_{n,i}|_{*T_n} + |U_{n,r+1}|_{*T_n} \right) e^{2(1+K)T_n}.
\]

Note \(U_{n,r+1} = Z_{n,r+} + U_{n,i}(0) = Z_{n,i}(0) - z_i\) for \(i \leq r\); hence using Lemma \(8.1\) and the assumed convergences, it follows that \(\|U_{n,r+}\|_{2,T_n} \overset{P}{\to} 0\) and, based on the above display, that \(H_{T_n} \overset{P}{\to} 0\).

Together these show \(\|U\|_{2,T_n} = \|Z_n - Z\|_{2,T_n} \overset{P}{\to} 0\) as \(n \to \infty\).

**Corollary 8.4.** Under assumptions of Lemma \(8.3\), \(\{\|Z_n\|_{2,T}\}_{n \in \mathbb{N}}\) is a tight sequence of random variables and

\[
\lim_{L \to \infty} \sup_n P(\tau_{n,L} \leq T) = 0. \tag{8.10}
\]

**Proof.** Fix \(\delta > 0\). Since \(Z\) has sample paths in \(C([0,T] : \ell_2)\), we can find \(L \in (0, \infty)\) so that \(P\left(\|Z\|_{2,T} > L\right) \leq \delta\). With \(\tau_{L+2} = \inf\{t \mid \|Z(t)\|_2 > L + 2\}\), note the inclusion \(\left\{\|Z\|_{2,T} \leq L\right\} \subset \{\tau_{L+2} > T\}\) which will be used in the next display. Now, by the right continuity of \(Z_n\), note for each \(n \in \mathbb{N}\)

\[
P(\tau_{n,L+2} \leq T) \leq P\left(\|Z_n\|_{2,T \wedge \tau_{n,L+2}} > L + 1\right) \leq P\left(\|Z_n - Z\|_{2,T \wedge \tau_{n,L+2}} > 1 \text{ or } \|Z\|_{2,T} > L\right)
\]

\[
\leq P\left(\|Z_n - Z\|_{2,T \wedge \tau_{n,L+2}} > 1\right) + P\left(\|Z\|_{2,T} > L\right)
\]

\[
\leq P\left(\|Z_n - Z\|_{2,T \wedge \tau_{n,L+2}} > 1\right) + \delta.
\]

Sending \(n \to \infty\) and using Lemma \(8.3\) shows \(\limsup_n P(\tau_{n,L+2} \leq T) \leq \delta\). Therefore,

\[
\limsup_n P\left(\|Z_n\|_{2,T} > L + 2\right) \leq \limsup_n P(\tau_{n,L+2} \leq T) \leq \delta.
\]

Since \(\delta > 0\) is arbitrary, this shows that \(\{\|Z_n\|_{2,T}\}_{n \in \mathbb{N}}\) is tight. The convergence in \(8.10\) now follows since \(\{\tau_{n,L+1} \leq T\} \subset \{\|Z_n\|_{2,T} > L\}\).
Proof of Theorem 2.3 Using Lemma 8.1 and Skorohod embedding theorem we can assume without loss of generality that $Z_n(0), M_n$ and $B$ are given on a common probability space, $Z_n(0) \to z$ in $\ell^1_1$, and $M_n \to \sqrt{2B}e_1$ in $\mathbb{D}([0, \infty) : \ell_2)$ a.s. From Lemma 8.3 we now have that for every $T, L \in (0, \infty)$ (8.7) holds. In fact, this shows loss of generality that $Z_n \to z$. From Lemma 6.1 it then follows that (9.1) is satisfied. Recall from (6.4) that $\lambda_n \to \infty$ a.s., and by establishing some useful convergence properties.

9. Proof of Theorem 2.4

In this section we give the proof of Theorem 2.4. As for the proof of Theorem 2.3 we begin with a convenient representation for $Z_n$ and by establishing some useful convergence properties.

Lemma 9.1. Let $\lambda_n, \alpha_n, d_n$ be as in the statement of Theorem 2.4. Suppose that $\{\|Z_n(0)\|_{1, \infty}\}_{n \in \mathbb{N}}$ is a tight sequence of random variables and $Z_n, \eta_n$ with sample paths in $\mathbb{D}([0, \infty) : \mathbb{R})$ so that, $W_{n,1}, \eta_n$ have absolutely continuous paths a.s., $W_{n,1}(0) = \eta_n(0) = 0$, and for any $t \geq 0$

\[
Z_n(1) = Z_n(0) - \int_0^t Z_n(1(s))ds + \int_0^t \lambda_n M_n(1) + W_{n,1}(t) - \eta_n(t) \quad (9.1)
\]

\[
Z_n(2) = Z_n(0) - \int_0^t Z_n(2(s))ds + \int_0^t \lambda_n M_n(2) + W_{n,2}(t) + \eta_n(t) \quad (9.2)
\]

\[
Z_n(i) = Z_n(i) - \int_0^t Z_n(i(s))ds + \int_0^t \lambda_n M_n(i) + W_{n,i}(t) + \eta_n(t) \quad \text{for } i \in \{3, \ldots, r\}. \quad (9.3)
\]

Furthermore, $\eta_n$ is non-decreasing process with $\eta_n(0) = 0$ that satisfies

\[
\eta_n(t) = \int_0^t \mathbb{I}_{[Z_n(\tau) \geq \theta_n]} d\eta_n(s) \quad (9.4)
\]

for some constants $\theta_n = \alpha_n + O(\sqrt{n/d_n}) \geq 0$ as $n \to \infty$. Also for any $L, T \in (0, \infty)$, as $n \to \infty$

1. $\lambda_n M_n \to \sqrt{2B}$ in $\mathbb{D}([0, T] : \mathbb{R})$

2. $t v(W_{n,1}; [0, T]) = \int_0^T |W_{n,1}(s)|ds \to 0$

3. $|W_{n,i}; [0, T]) \to 0$ for $i \in \{2, \ldots, r\}$

4. $\|Z_n, \eta_n\|_{2, T} \to 0$.

Here $B$ is a standard Brownian motion and $T_n = T \wedge \tau_{n,L}$.

Proof. By our assumptions on $\alpha_n$, we can find a $\kappa \in (0, \infty)$ such that $\theta_n = \alpha_n + \frac{\kappa \sqrt{n}}{d_n} \geq 0$ for every $n$. Note that $\theta_n \to \alpha$ as $n \to \infty$. Recall the functions $q_{n,i}$ defined in (6.4). Define

\[
W_{n,1}(t) = \int_0^t q_{n,1}(Z_n(1(s))) \mathbb{I}_{[Z_n(1(s)) \geq \theta_n]} ds
\]

\[
\eta_n(t) = \int_0^t q_{n,1}(Z_n(1(s))) \mathbb{I}_{[Z_n(1(s)) \geq \theta_n]} ds
\]

so that $\eta_n(t) = \int_0^t \mathbb{I}_{[Z_n(\tau) \geq \theta_n]} d\eta_n(s)$, and

\[
\int_0^t q_{n,1}(Z_n(1(s)))ds = \eta_n(t) - W_{n,1}(t).
\]

From Lemma 6.1 it then follows that (9.1) is satisfied. Recall from (6.4) that $q_{n,1}(z) = \lambda_n \sqrt{n}\{\beta_n(\lambda_n + z/\sqrt{n}) - \beta_n(\lambda_n)\}$. Then, by monotonicity of $\beta_n$, $q_{n,1}(z) \geq 0$ whenever $z \geq 0$. The
condition \( \theta_n \geq 0 \) shows that \( \eta_n \) is non-decreasing and
\[
\sup_{z \leq \theta_n} |q_n(1)| = \sqrt{n} \beta_n (\lambda_n + \theta_n/\sqrt{n}) \leq \sqrt{n} (\lambda_n + \theta_n/\sqrt{n})^d_n \\
= \sqrt{n} (1 - ((\log d_n)/d_n + (\alpha_n - \theta_n)/\sqrt{n}))^d_n \\
\leq \exp \left( - \log \frac{d_n}{\sqrt{n}} + \kappa \right) \to 0 \text{ as } n \to \infty.
\]
This shows that \( \text{tv}(W_n; [0, T]) \to 0 \text{ a.s.} \)

Next, since \( d_n(1 - \lambda_n) \to \infty \), Lemma 6.4 shows that
\[
\mu_n \to f_1 \in \ell^1_i \text{ as } n \to \infty.
\]
Therefore \( \mathbf{G}_n(0) = \theta_n + \frac{Z_n(0)}{\sqrt{n}} \to f_1 \text{ in } \ell^1_i \).

Then by Lemma 6.2
\[
\sqrt{n} M_n \Rightarrow \sqrt{2} B e_1 \text{ in } \mathbb{D}([0, \infty) : \ell_2),
\]
and by Lemma 6.4 \( \|Z_{n,r}^+\|_{2,T \land n,L} \mathbb{P} \to 0 \text{ as } n \to \infty. \)

Define
\[
W_{n,2}(t) = - \int_0^t q_n,2(Z_{n,2}(s))ds + \sqrt{n} M_{n,2}(t) - W_{n,1}(t).
\]

Using (9.5) and Lemma 6.1 once more, we see that (9.2) is satisfied. Finally, for \( i \in \{3, \ldots, r\} \), define
\[
W_{n,i}(t) = \int_0^t q_n,i-1(Z_{n,i-1}(s))ds - \int_0^t q_n,i(Z_{n,i}(s)) + \sqrt{n} M_{n,i}(t).
\]

Then, from Lemma 6.1 again, it follows that (9.3) is satisfied with the above choice of \( W_{n,i} \).

Lemma 6.3 along with (9.6), (9.7), and \( |Z_{n,i}|_{s,T \land n,L} \leq L \) show that, as \( n \to \infty \), and \( |W_{n,i}|_{s,T_n} \mathbb{P} \to 0 \) for each \( i \in \{2, \ldots, r\} \).

**Corollary 9.2.** Suppose that the assumptions in Lemma 9.1 are satisfied. Assume further that \( d_n \ll n^{2/3} \). Then, the conclusions of Lemma 9.1 hold with \( \theta_n = \alpha_n \) and
\[
\eta_n(t) = \int_0^t \gamma_n^{-1}(1 + \delta_n(s))e^{\gamma_n Z_{1,s}(s) - \alpha_n} I_{\{Z_{1,s}(s) \geq \alpha_n\}}ds,
\]
where \( \gamma_n = \frac{d_n}{\sqrt{n}} \) and \( \delta_n \) is a process with sample paths in \( \mathbb{D}([0, \infty), \mathbb{R}) \) such that \( |\delta_n|_{s,T \land n,L} \to 0 \text{ a.s.} \)

for each \( L > 0 \).

**Proof.** Since \( d_n \ll n^{2/3} \) and \( \alpha_n = O(n^{1/6}) \), the hypothesis of Lemma 6.7 is satisfied. Define
\[
\delta_n(s) = q_n,1(Z_{1,s}(s)) \gamma_n \left( e^{\gamma_n Z_{1,s}(s) - \alpha_n} - e^{-\gamma_n \alpha_n} \right)^{-1} - 1.
\]

Since \( \sup_{s \leq T \land n,L} |Z_{1,s}(s)| \leq L \), Lemma 6.7 shows that \( |\delta_n|_{s,T_n} \to 0 \text{ a.s. as } n \to \infty. \)

Next define
\[
W_{n,1}(t) = \gamma_n^{-1} \int_0^t (1 + \delta_n(s))\left( e^{-\gamma_n \alpha_n} - e^{\gamma_n Z_{1,s}(s) - \alpha_n} \right) I_{\{Z_{1,s}(s) < \alpha_n\}} ds \\
+ \int_0^t \gamma_n^{-1}(1 + \delta_n(s))e^{\gamma_n Z_{1,s}(s) - \alpha_n} I_{\{Z_{1,s}(s) \geq \alpha_n\}} ds
\]

Then \( W_{n,1}(0) = 0 \), \( W_{n,1} \) is absolutely continuous and, with \( \kappa = \sup_n \frac{d_n}{\sqrt{n}} \alpha_n < \infty \),
\[
\text{tv}(W_{n,1}; [0, T]) I_{\{|\delta_n|_{s,T_n} < 1\}} = \gamma_n^{-1} \int_0^T |1 + \delta_n(s)| \left| e^{-\gamma_n \alpha_n} - e^{\gamma_n Z_{1,s}(s) - \alpha_n} \right| I_{\{Z_{1,s}(s) < \alpha_n\}} ds \\
\leq \frac{2(1 + e^\kappa)T}{\gamma_n} \to 0 \text{ as } n \to \infty.
\]
Hence, since $|\delta_n|_{s,T_n} \to 0$, we have that $\text{tv}(W_{n,1}; [0, T_n]) \stackrel{P}{\to} 0$ as $n \to \infty$. By rearranging terms we see that, with the above definitions of $W_{n,1}$ and $\eta_n$, (9.5) is satisfied. The result follows. \[\blacksquare\]

Since $\gamma_n \to \infty$ and $\theta_n \to \alpha$ as $n \to \infty$, the previous lemma suggests a connection to the Skorokhod map $\Gamma_\alpha$ defined in (2.1). In order to make this connection precise, we begin with the following lemma.

**Lemma 9.3.** Under the assumptions of Theorem 2.4 for any $L \in (0, \infty)$

$$
\sup_{t \in [0, T \wedge \tau_n, L]} (Z_{n,1}(t) - \alpha_n) \xrightarrow{P} 0 \text{ as } n \to \infty. \tag{9.9}
$$

**Proof.** Consider first the case when $d_n \gg \sqrt{n} \log n$. For this case $\epsilon_n \xrightarrow{n \rightarrow \infty} 0$, and since

$$
Z_{n,1}(t) = \sqrt{n}(G_{n,1}(t) - \lambda_n) \leq \sqrt{n}(1 - \lambda_n) = \frac{\sqrt{n} \log d_n}{d_n} + \alpha_n,
$$

we have that (9.9) holds. Now consider the complementary case, namely $d_n \gg \sqrt{n}$ but $d_n \gg \sqrt{n} \log n$ does not hold. In this case, we may find an increasing sub-sequence $\{n_k\}_{k \in \mathbb{N}} \subseteq \mathbb{N}$ so that $\sqrt{n} \ll d_n \ll n^{2/3}$ holds when $n \in \{n_k\}_{k \in \mathbb{N}}$ and $d_n \gg \sqrt{n} \log n$ holds when $n \notin \{n_k\}_{k \in \mathbb{N}}$ (e.g. take $\{n_k\}_{k \in \mathbb{N}} = \{n \in \mathbb{N} \mid d_n \leq n^{0.6}\}$). The argument above shows the convergence of (9.9) along the latter sub-sequence. Therefore it suffices to show the convergence of (9.9) along the sub-sequence $\{n_k\}_{k \in \mathbb{N}}$ where $\sqrt{n} \ll d_n \ll n^{2/3}$.

We will use Corollary 9.2. Since $Z_{n,1}(0) \xrightarrow{P} z_1 \in \mathbb{R}$ with $z_1 \leq \alpha$, we have $(Z_{n,1}(0) - \alpha_n)^{+} \xrightarrow{P} 0$ as $n \to \infty$. It now suffices to show that for any $\epsilon > 0$

$$
P \left( \sup_{t \in [0, T \wedge \tau_n, L]} Z_{n,1}(t) > \alpha_n + 6\epsilon \right) \to 0 \text{ as } n \to \infty.
$$

Let $\delta_n \doteq \inf \{ t \geq 0 \mid Z_{n,1}(t) > \alpha_n + 6\epsilon \}$ and, as before, $T_n = T \wedge \tau_n, L$. It is then enough to show that $P(\delta_n \leq T_n) \to 0$ as $n \to \infty$. For this inductively define stopping times, $\sigma_{n,0} = 0$,

$$
\sigma_{n,2k-1} = \inf \{ t > \sigma_{n,2k-2} \mid Z_{n,1}(t) > \alpha_n + 3\epsilon \}
$$

$$
\sigma_{n,2k} = \inf \{ t > \sigma_{n,2k-1} \mid Z_{n,1}(t) < \alpha_n + 2\epsilon \}, \quad k \in \mathbb{N}.
$$

Note that for each $n \in \mathbb{N}$, $\sigma_{n,r} \to \infty$ as $r \to \infty$, almost surely. Also, henceforth, without loss of generality we consider only $n$ that are large enough so that $1/\sqrt{n} < \epsilon$. Hence on the set $\{\delta_n < \infty\}$, $\delta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k})$ for some $k \in \mathbb{N}$. Then for every $K \in \mathbb{N}$

$$
P(\delta_n \leq T_n) \leq \sum_{k=1}^{K} P(\delta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k} \wedge T_n]) + P(\sigma_{n,2K+1} \leq T_n).
$$

Hence to complete the proof it is enough to show that

1. For each $k \in \mathbb{N}$, $\lim_{n \to \infty} P(\delta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k} \wedge T_n]) = 0$,
2. $\lim_{K \to \infty} \limsup_{n \to \infty} P(\sigma_{n,2K+1} \leq T_n) = 0$.

Consider (1) first. Note that on the set $C_{n,1} \doteq \{ Z_{n,1}(0) \leq \alpha_n + 3\epsilon \}$, for any $k \in \mathbb{N}$,

$$
\alpha_n + 3\epsilon \leq Z_{n,1}(\sigma_{n,2k-1}) = Z_{n,1}(\sigma_{n,2k-1}^-) + Z_{n,1}(\sigma_{n,2k-1}) - Z_{n,1}(\sigma_{n,2k-1}^-) \leq Z_{n,1}(\sigma_{n,2k-1}^-) + \epsilon \leq \alpha_n + 4\epsilon. \tag{9.10}
$$

Similarly,

$$
Z_{n,1}(t) \geq \alpha_n + \epsilon \quad \text{for each } t \in [\sigma_{n,2k-1}, \sigma_{n,2k}]. \tag{9.11}
$$

Let $H_n(t) \doteq \sqrt{n}M_{n,1}(t + \sigma_{n,2k-1}) - \sqrt{n}M_{n,1}(\sigma_{n,2k-1})$ for $t \geq 0$ and consider the sets

$$
C_{n,2} \doteq \{ \delta_n \wedge \sigma_{n,2k-1} \leq T_n \}, \quad C_{n,3} \doteq \left\{ |W_{n,1}|_{s,T_n} \leq \epsilon/2, |\delta_n|_{s,T_n} \leq \frac{1}{2} \right\}.
$$
Then on the set $C_n = \cap_{i=1}^3 C_{n,i}$, using Corollary 9.2, for any $t \in [0, (T_n \wedge \sigma_{n,2k}) - \sigma_{n,2k-1}]$,

\[
Z_{n,1}(t + \sigma_{n,2k-1}) - Z_{n,1}(\sigma_{n,2k-1}) = -\int_{\sigma_{n,2k-1}}^{\sigma_{n,2k-1}+t} (Z_{n,1}(s) - Z_{n,2}(s)) ds + H_n(t) + W_{n,1}(t + \sigma_{n,2k-1}) - W_{n,1}(\sigma_{n,2k-1}) - \int_{\sigma_{n,2k-1}}^{\sigma_{n,2k-1}+t} \gamma_n^{-1}(1 + \delta_n(s))e^\gamma_n(Z_{n,1}(s) - \alpha_n)I\{Z_{n,1}(s) \geq \alpha_n\} ds
\]

Since for $t$ in the above interval $\sigma_{n,2k-1} + t \leq T_n \leq \tau_{n,L}$, $|Z_{n,1}(s)| + |Z_{n,2}(s)| \leq 2L$ for any $s \leq \sigma_{n,2k-1} + t$. Also, since $\sigma_{n,2k-1} + t \leq \sigma_{n,2k}$, by (9.11), $Z_{n,1}(s) - \alpha_n \geq \epsilon$ for any $s \in [\sigma_{n,2k-1}, \sigma_{n,2k-1} + t]$. Thus on $C_n$, we have

\[
Z_{n,1}(t + \sigma_{n,2k-1}) - Z_{n,1}(\sigma_{n,2k-1}) \leq 2Lt + H_n(t) + \epsilon - \frac{t}{2\gamma_n} \exp(\gamma_n \epsilon) = Y_n(t).
\]

Using (9.10), on $C_n$, $Z_{n,1}(\vartheta_n) - Z_{n,1}(\sigma_{n,2k-1}) \geq \alpha_n + 6\epsilon - \alpha_n - 4\epsilon = 2\epsilon$. Hence

\[
P(\vartheta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k} \wedge T_n]) \leq P(\vartheta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k} \wedge T_n], C_n) + P(C_{n,1}) + P(C_{n,3})
\]

\[
\leq P\left( \sup_{t \in [0,T]} Y_n(t) \geq 2\epsilon \right) + P(C_{n,1}) + P(C_{n,3}),
\]

where the second inequality is on observing that on the set $\{\vartheta_n \in [\sigma_{n,2k-1}, \sigma_{n,2k} \wedge T_n]\}$, (9.12) holds with $t$ replaced by $\vartheta_n - \sigma_{n,2k-1}$. Next note that $H_n$ is a $\{G_t^n\}$ martingale, where $G_t^n = F_{t+\sigma_{n,2k-1}}$ and

\[
\langle H_n \rangle_t = \langle \sqrt{n}M_{n,1} \rangle_{t+\sigma_{n,2k-1}} - \langle \sqrt{n}M_{n,1} \rangle_{\sigma_{n,2k-1}}
\]

\[
= \int_{\sigma_{n,2k-1}}^{\sigma_{n,2k-1}+t} [G_{n,1}(s) - G_{n,2}(s) + \lambda_n - \lambda_n \beta_n(G_{n,1}(s))] ds
\]

\[
\leq 2t,
\]

where the second equality is from (3.6).

Since $\gamma_n \to \infty$ we can apply Lemma 6.9 to conclude

\[
P\left( \sup_{t \in [0,T]} Y_n(t) \geq 2\epsilon \right) \to 0
\]

as $n \to \infty$. We also have $\lim_n P(C_{n,i}^c) = 0$ for $i = 1, 3$ since, as noted earlier $(Z_{n,1}(0) - \alpha_n)^+ \overset{P}{\to} 0$, and by Corollary 9.2 respectively. From these observations it follows that the right side of (9.13) converges to 0 as $n \to \infty$, which completes the proof of (1).

Now we prove (2). Let $\rho_{n,i} = \sigma_{n,i} \wedge \tau_{n,L}$ and define

\[
Y_{n,K}(t) = \sum_{i=0}^{K} (Z_{n,1}(t \wedge \rho_{n,2i+1}) - Z_{n,1}(t \wedge \rho_{n,2i})).
\]

Note that $\{\sigma_{n,2K+1} \leq T_n\} \subseteq \{Y_{n,K}(T) \geq K\epsilon\}$ and hence to prove (2) it is sufficient to show that

\[
\lim_{n \to \infty} P(Y_{n,K}(T) \geq K\epsilon) \to 0 \text{ as } K \to \infty.
\]
From Corollary 9.2 we have that on the set \(C_{n,4} = \{\text{tv}(W_{n,1}; [0, T_n]) \leq 1\},
\)

\[
Y_{n,K}(T) = \sum_{i=0}^K \int_{T \cap \rho_{n,2i+1}}^{T \cap \rho_{n,2i+1}} (Z_{n,2}(s) - Z_{n,1}(s)) ds + \sum_{i=0}^K \sqrt{n}M_{n,1}(T \cap \rho_{n,2i+1}) - \sqrt{n}M_{n,1}(T \cap \rho_{n,2i})
+ \sum_{i=0}^K W_{n,1}(T \cap \rho_{n,2i+1}) - W_{n,1}(T \cap \rho_{n,2i}) - \sum_{i=0}^K \int_{T \cap \rho_{n,2i+1}}^{T \cap \rho_{n,2i+1}} \gamma_n^{-1}(1 + \delta_n(s)) + e^{-\gamma_n(Z_{n,1}(s) - \alpha_n)} \mathbb{1}_{\{Z_{n,1}(s) \geq \alpha_n\}} ds
\]

\[
\leq 2LT + \sum_{i=0}^K (\sqrt{n}M_{n,1}(T \cap \rho_{n,2i+1}) - \sqrt{n}M_{n,1}(T \cap \rho_{n,2i})) + \text{tv}(W_{n,1}; [0, T])
\]

\[
\leq 2LT + 1 + H_{n,K}(T)
\]

where we have used the facts that \(\sup_{s \leq \tau_{n,L}} |Z_{n,1}(s)| \leq L\), and that the rightmost term in the third line is non-positive. Also, here

\[
H_{n,K}(t) = \sum_{i=0}^K (\sqrt{n}M_{n,1}(t \cap \rho_{n,2i+1}) - \sqrt{n}M_{n,1}(t \cap \rho_{n,2i}))
\]

Using (3.6), we see that \(H_{n,K}\) is a \(\mathcal{F}_t^\rho\)-martingale with quadratic variation given by

\[
\langle H_{n,K} \rangle_t = \sum_{i=0}^K \left( (\sqrt{n}M_{n,1})_{T \cap \rho_{n,2i+1}} - (\sqrt{n}M_{n,1})_{T \cap \rho_{n,2i}} \right)
\]

\[
= \sum_{i=0}^K \int_{T \cap \rho_{n,2i}}^{T \cap \rho_{n,2i+1}} (G_{n,1}(s) - G_{n,2}(s) + \lambda_n - \lambda_n\beta_n(G_{n,1}(s))) ds \leq 2t.
\]

Hence

\[
P(Y_{n,K}(T) \geq K\epsilon) \leq P(Y_{n,K}(T) \geq K\epsilon, C_{n,4}) + P(C_{n,4}^c) \leq P(H_{n,K}(T) > K\epsilon - (2LT + 1)) + P(C_{n,4}^c)
\]

\[
\leq \frac{Eh_{n,K}^2(T)}{(K\epsilon - (2LT + 1))^2} + P(C_{n,4}^c) \leq \frac{2T}{(K\epsilon - (2LT + 1))^2} + P(C_{n,4}^c).
\]

From Corollary 9.2 \(P(C_{n,4}^c) \to 0\) as \(n \to \infty\). This together with the above display shows

\[
\lim_{K \to \infty} \limsup_{n \to \infty} P(Y_{n,K}(T) \geq K\epsilon) = 0.
\]

Thus we have shown (9.14) and the proof of (2) is complete. The result follows.

**Lemma 9.4.** Suppose the hypothesis of Theorem 2.4 holds, then for each \(n \in \mathbb{N}\), there is a real constant \(\theta_n = \alpha_n + O(\sqrt{n}/d_n) \geq 0\) and processes \(W_{n,1}, W_{n,2}\) with sample paths in \(D([0, \infty) : \mathbb{R})\) such that with \(\tilde{Z}_{n,1} \equiv Z_{n,1} \circ \theta_n\)

\[
\tilde{Z}_{n,1}(t) = \Gamma_{\theta_n} \left( \tilde{Z}_{n,1}(0) - \int_0^t (\tilde{Z}_{n,1}(s) - Z_{n,2}(s)) + \sqrt{n}M_{n,1}(\cdot) + \tilde{W}_{n,1}(\cdot) \right)(t), \quad \text{and} \quad (9.15)
\]

\[
Z_{n,2}(t) = Z_{n,2}(0) - \int_0^t (Z_{n,2}(s) - Z_{n,3}(s)) ds + W_{n,2}(t) + \eta_n(t) \quad \text{for all } t > 0,
\]

where

\[
\eta_n = \Gamma_{\theta_n} \left( \tilde{Z}_{n,1}(0) - \int_0^t (\tilde{Z}_{n,1}(s) - Z_{n,2}(s)) + \sqrt{n}M_{n,1}(\cdot) + \tilde{W}_{n,1}(\cdot) \right).
\]

Furthermore, for any \(L, T \in (0, \infty)\), \(\left| W_{n,1} \right|_{\mathcal{F}_T \cap \tau_{n,L}} \prec W_{n,2} \circ \mathcal{T}_T \cap \tau_{n,L}\) and \(\left| (Z_{n,1} - \theta_n)^+ \right|_{\mathcal{F}_T \cap \tau_{n,L}} \overset{P}{\to} 0\) as \(n \to \infty\).

**Proof.** Let \(\theta_n\) be as in Lemma 9.1. Since \(d_n \gg \sqrt{n}\), \(\theta_n = \alpha_n + o(1)\) and Lemma 9.3 shows

\[
\left| (Z_{n,1} - \theta_n)^+ \right|_{\mathcal{F}_T \cap \tau_{n,L}} \to 0.
\]
Note that $\tilde{Z}_{n,1} = Z_{n,1} - (Z_{n,1} - \theta_n)^+$. Hence we can rewrite (9.1) and (9.2) as

$$
\tilde{Z}_{n,1}(t) = \tilde{Z}_{n,1}(0) - \int_0^t \tilde{Z}_{n,1}(s) ds + \int_0^t Z_{n,2}(s) ds + \sqrt{n} M_{n,1}(t) + \tilde{W}_{n,1}(t) - \eta_n(t) \tag{9.18}
$$

$$
Z_{n,2}(t) = Z_{n,2}(0) - \int_0^t (Z_{n,2}(s) - Z_{n,3}(s)) ds + W_{n,2}(t) + \eta_n(t),
$$

where

$$
\tilde{W}_{n,1}(t) \doteq W_{n,1}(t) - \int_0^t (Z_{n,1}(s) - \theta_n)^+ ds - (Z_{n,1}(t) - \theta_n)^+ + (Z_{n,1}(0) - \theta_n)^+.
$$

The properties of $\eta_n$ from Lemma 9.1 (and Corollary 9.2) say that $\eta_n$ is a non-decreasing process, with $\eta_n(0) = 0$ and $\eta_n(t) = \int_0^t 1_{\{\tilde{Z}_{n,1}(s) = \theta_n\}} d\eta(s)$. Since $\tilde{Z}_{n,1} \leq \eta_n$, (9.18) and the characterizing properties of the Skorokhod map show (9.15) and (9.16). Finally, by Lemma 9.1, Corollary 9.2, and Lemma 9.3, we have

$$|W_{n,1}|_{*,T \wedge \tau_{n,L}} \xrightarrow{P} 0, \text{ and } |W_{n,2}|_{*,T \wedge \tau_{n,L}} \xrightarrow{P} 0$$

as $n \to \infty$. Hence, using (9.17), $|\tilde{W}_{n,1}|_{*,T \wedge \tau_{n,L}} \xrightarrow{P} 0$ as $n \to \infty$, and the result follows.

The following lemma will be needed in order to prove the tightness of $Z_n$.

**Lemma 9.5.** Under the hypothesis of Theorem 2.4, the collection of random variables $\{\|Z_n\|_{2,T}\}_{n \in \mathbb{N}}$ is tight for any $T \in (0, \infty)$.

**Proof.** Fix $T \in (0, \infty)$. In Lemma 9.4, using the definition of the Skorokhod map $\Gamma_{\theta_n}$ for $\theta_n \geq 0$ (see (2.1)), we have, for any $t > 0$ that

$$\eta_n(t) \leq \left|\tilde{Z}_{n,1}(0)\right| + \int_0^t \left|\tilde{Z}_{n,1}(s)\right| ds + \int_0^t |Z_{n,2}(s)| ds + |\sqrt{n} M_{n,1}|_{*,t} + |\tilde{W}_{n,1}|_{*,t}.$$

This shows that for any $t \geq 0$

$$|\tilde{Z}_{n,1}|_{*,t} \leq 2 \left( |\tilde{Z}_{n,1}(0)| + \int_0^t |\tilde{Z}_{n,1}|_{*,s} ds + \int_0^t |Z_{n,2}|_{*,s} ds + |\sqrt{n} M_{n,1}|_{*,t} + |\tilde{W}_{n,1}|_{*,t} \right)$$

and

$$|Z_{n,2}|_{*,t} \leq |\tilde{Z}_{n,1}(0)| + |Z_{n,2}(0)| + \int_0^t |\tilde{Z}_{n,1}|_{*,s} ds + \int_0^t (2 |Z_{n,2}|_{*,s} + |Z_{n,3}|_{*,s}) ds$$

$$+ |\sqrt{n} M_{n,1}|_{*,t} + |\tilde{W}_{n,1}|_{*,t} + |W_{n,2}|_{*,t}.$$

and

$$|Z_{n,i}|_{*,t} \leq |Z_{n,0}|(0) + \int_0^t |Z_{n,i}|_{*,s} ds + \int_0^t |Z_{n,i+1}|_{*,s} ds + |W_{n,i}|_{*,t} \quad \text{for } i \in \{3, \ldots, r\}$$

where the last line is from Lemma 9.1. Let $H_t \doteq |\tilde{Z}_{n,1}|_{*,t} + |Z_{n,2}|_{*,t} + \ldots + |Z_{n,r}|_{*,t}$. By adding over equations in the above display, we have for $t \in [0, \tau]$ and $\tau \in [0, T]$ that

$$0 \leq H_t \leq 4 \left( H_0 + |\sqrt{n} M_{n,1}|_{*,\tau} + |\tilde{W}_{n,1}|_{*,\tau} + \sum_{i=2}^r |W_{n,i}|_{*,\tau} + \int_0^t H_s ds \right).$$

By Gronwall’s inequality, for all $\tau \in [0, T]$,

$$H_\tau \leq 4 \left( H_0 + |\sqrt{n} M_{n,1}|_{*,\tau} + |\tilde{W}_{n,1}|_{*,\tau} + \sum_{i=2}^r |W_{n,i}|_{*,\tau} \right) e^{4\tau}. \tag{9.19}$$
Let $\tilde{Z}_n \doteq (\tilde{Z}_{n,1}, Z_{n,2}, \ldots, Z_{n,r})$. Since $\tilde{Z}_n(0) \xrightarrow{P} (z_1, \ldots, z_r)$, and $\sqrt{n}M_n \Rightarrow Be_1$, for every $\epsilon > 0$ there is a $L_1 \in (0, \infty)$ such that for every $n \in \mathbb{N}$

$$P(C_{n,1}) \leq \frac{\epsilon}{2^n},$$

where $C_{n,1} = \{H_0 + |\sqrt{n}M_{n,1}|_{s,T} \geq L_1\}$.

Applying Lemmas 9.1 and Lemma 9.4 with $L = 4(L_1 + 1)e^{4T} + 2$ we can find an $n_0 \in \mathbb{N}$ so that $P(C_{n,2}) \leq \frac{\epsilon}{2}$ for $n \geq n_0$, where

$$C_{n,2} = \left\{ \left| \tilde{W}_{n,1} \right|_{s,T_n} + \sum_{i=2}^{r} |W_{n,i}|_{s,T_n} + |(Z_{n,1} - \theta_n)^+|_{s,T_n} + \|Z_{n,r+}\|_{2,T_n} \geq 1 \right\}$$

and $T_n \doteq T \wedge \tau_{n,L}$. On the event $(C_{n,1} \cup C_{n,2})^c$

$$\left\| \tilde{Z}_n \right\|_{1,T_n} = H_{T_n} < 4(L_1 + 1)e^{4T}$$

by (9.19), and hence by triangle inequality (and noting $\|\tilde{x}\|_2 \leq \|\tilde{x}\|_1$)

$$\left\| Z_n \right\|_{2,T_n} \leq \left\| \tilde{Z}_n \right\|_{1,T_n} + |(Z_{n,1} - \theta_n)^+|_{s,T_n} + \|Z_{n,r+}\|_{2,T_n} \leq 4(L_1 + 1)e^{4T} + 1 = L - 1. \quad (9.20)$$

Also, by the definition of $\tau_{n,L}$, $\|Z_n(\tau_{n,L})\|_2 \geq L - \frac{1}{\sqrt{n}}$ on the set $\tau_{n,L} < T$. Hence we must have that $\tau_{n,L} > T$ whenever (9.20) holds, and hence

$$\|Z_n\|_{2,T} < L - 1$$

on the event $(C_{n,1} \cup C_{n,2})^c$.

This shows that

$$P\left( \left\| Z_n \right\|_{2,T} \geq L \right) \leq P(C_{n,1} \cup C_{n,2}) \leq \epsilon \quad \forall n \geq n_0$$

Since $\epsilon > 0$ is arbitrary, the result follows.

The following result is immediate from Lemmas 6.5, 9.1, 9.4, and 9.5.

**Corollary 9.6.** Under the hypothesis of Theorem 2.4 for any $T > 0$, $\lim_{L \to \infty} \sup_n P(\tau_{n,L} \leq T) = 0$. In particular the processes $\tilde{W}_{n,i}, \{W_{n,i}\}_{i=2}^{r}, \|Z_{n,r+}\|_2, (Z_{n,1} - \theta_n)^+$ converge in probability to zero in $\mathbb{D}(\{0, \infty\} : \mathbb{R})$ as $n \to \infty$.

**Corollary 9.7.** Under the hypothesis of Theorem 2.4, the sequence of processes $\{Z_n\}_{n \in \mathbb{N}}$ is tight in $\mathbb{D}(\{0, \infty\} : \ell_2)$.

**Proof.** Let $\theta_n$ be as in Lemma 9.4. For the sequence $\{\tilde{Z}_n\}_{n \in \mathbb{N}}$ introduced in the proof of Lemma 9.5 note that

$$Z_n = P \tilde{Z}_n + (Z_{n,1} - \theta_n)^+e_1 + S_rZ_{n,r+}, \quad (9.21)$$

where $P : \mathbb{R}^r \to \ell_2$ is given by $P(x_1, \ldots, x_r) = (x_1, \ldots, x_r, 0)$ while $S_r : \ell_2 \to \ell_2$ is given by $S_r y = (0, y)$ where $0$ is the zero vector in $\mathbb{R}^r$. Since these maps are continuous, the tightness of $\{\tilde{Z}_n\}_{n \in \mathbb{N}}$ in $\mathbb{D}(\{0, T\} : \mathbb{R}^k)$, the tightness of $\{Z_{n,r+}\}_{n \in \mathbb{N}}$ in $\mathbb{D}(\{0, T\} : \ell_2)$, and the tightness of $\{(Z_{n,1} - \theta_n)^+\}_{n \in \mathbb{N}}$ in $\mathbb{D}(\{0, T\} : \mathbb{R})$ will show the tightness of the sequence $\{Z_n\}_{n \in \mathbb{N}}$ in $\mathbb{D}(\{0, T\} : \ell_2)$.

Note by Corollary 9.6 for each fixed $T < \infty$, $\|Z_{n,r+}\|_{2,T} \xrightarrow{P} 0$ and $|(Z_{n,1} - \theta_n)^+| \xrightarrow{P} 0$. Hence it is sufficient to show that $\{\tilde{Z}_n\}_{n \in \mathbb{N}}$ is tight in $\mathbb{D}(\{0, T\} : \mathbb{R}^r)$. From Lemma 9.5 the convergence of $W_{n,i}$ in Corollary 9.6 and equations for $Z_{n,j}, j = 3, \ldots, r$ in Lemma 9.1 it is immediate that $(Z_{n,3}, \ldots, Z_{n,r})$ is tight in $\mathbb{D}(\{0, \infty\} : \mathbb{R}^{r-2})$. Finally consider the pair $(\tilde{Z}_{n,1}, Z_{n,2})$. Note that

$$|\tilde{Z}_{n,1} - Z_{n,2}|_{s,T} \leq |(Z_{n,1} - \theta_n)^+|_{s,T} + |Z_{n,1}|_{s,T} + |Z_{n,2}|_{s,T} \leq |(Z_{n,1} - \theta_n)^+|_{s,T + 2} + \|Z_n\|_{2,T}$$
and the right side in the above display is tight in $\mathbb{R}_+$. This shows the tightness of
\[ \int_0^1 (\tilde{Z}_{n,1}(s) - \tilde{Z}_{n,2}(s)) ds \]
in $C([0, \infty) : \mathbb{R})$. Combining this observation with Lemma 9.5, the convergence of $\sqrt{n}M_{n,1}$ in Lemma 9.1 and the convergence of $\tilde{W}_{n,1}$ in Corollary 9.6 it follows that
\[ R_n(\cdot) = \tilde{Z}_{n,1}(0) - \int_0^1 (\tilde{Z}_{n,1}(s) - \tilde{Z}_{n,2}(s)) ds + \sqrt{n}M_{n,1}(\cdot) + \tilde{W}_{n,1}(\cdot) \quad (9.22) \]
is tight in $\mathbb{D}([0, \infty) : \mathbb{R})$. Using the identity
\[ \Gamma_{\theta_n}(R_n)(t) = \Gamma_{\theta_n}(\Gamma_{\theta_n}(R_n)(s) + R_n(\cdot + s) - R_n(s)) (t - s) \]
for $0 \leq s \leq t$, we see from the definition of the Skorohod map that
\[ |\Gamma_{\theta_n}(R_n)(t) - \Gamma_{\theta_n}(R_n)(s)| \leq 2 \sup_{s \leq u \leq t} |R_n(u) - R_n(s)|. \]
Together with the tightness of $R_n$ this immediately implies the tightness of $\tilde{Z}_{n,1} = \Gamma_{\theta_n}(R_n)$ and of $\Gamma_{\theta_n}(R_n)$. Finally the tightness of $Z_{n,2}$ is now immediate from Lemma 9.5 the convergence of $W_{n,2}$ in Corollary 9.6 and the tightness of $\Gamma_{\theta_n}(R_n)$ noted above. The result follows.

**Proof of Theorem 2.4** From Lemma 9.6 and from the tightness of $\{\|Z_n(0)\|_1\}_{n \in \mathbb{N}}$, it follows under the conditions of the theorem that $\mu_n \xrightarrow{P} f_1$ and $G_n(0) \xrightarrow{P} f_1$ in $\ell_2^1$. This proves the first statement in the theorem. Now consider the second statement. Fix $T < \infty$. From Corollary 9.7 $\{Z_n\}_{n \in \mathbb{N}}$ is tight in $\mathbb{D}([0, \infty) : \ell_2)$. Also from Lemma 9.1 $\sqrt{n}M_{n,1}$ converges in distribution to $\sqrt{2}B$ where $B$ is a standard Brownian motion and from Corollary 9.6
\[ (\tilde{W}_{n,1}, \{W_{n,i}\}_{i=2}^\infty, (Z_{n,1} - \theta_n)^+) \xrightarrow{P} 0 \text{ in } \mathbb{D}([0, T] : \mathbb{R}^{r_1+1}) \]
Suppose that along a subsequence
\[ (Z_n, \sqrt{n}M_{n,1}, \tilde{W}_{n,1}, \{W_{n,i}\}_{i=2}^\infty, (Z_{n,1} - \theta_n)^+) \Rightarrow (Z, \sqrt{2}B, 0) \]
in $\mathbb{D}([0, \infty) : \ell_2 \times \mathbb{R}^{r_1+2})$ and for notational simplicity label the subsequence once more as $\{n\}$. Also by appealing to Skorohod embedding theorem we assume that all the processes in the above display are given on a common probability space and the above convergence holds a.s. Since $J_T(Z_n) \Rightarrow \sup_{0 \leq t \leq T} \|Z_n(t) - Z_n(t^-)\|_2$ is at most $\frac{1}{\sqrt{n}}$ and $Z_n(0) \xrightarrow{P} z$, we have $J_T(Z) = 0$ and $Z(0) = z$ a.s.
In particular $Z$ has sample paths in $C([0, \infty) : \ell_2)$ and $(Z_n, \sqrt{n}M_{n,1}) \Rightarrow (Z, \sqrt{2}B)$ uniformly over compact time intervals in $\ell_2 \times \mathbb{R}$. Since by Corollary 9.6, for every $T < \infty$, $\|Z_{n,r}\|_{2,T} \xrightarrow{P} 0$, it suffices to show that $(Z_1, \ldots, Z_r)$ along with $B$ satisfy (2.13).

From the equations of $(Z_{n,3}, \ldots, Z_{n,r})$ in Lemma 9.1 uniform convergences of $Z_n$ to $Z$, and the uniform convergences of $\{W_{n,i}\}_{i=3}^r$ to 0, it is immediate that $(Z_3, \ldots, Z_r)$ satisfy (2.13). Finally consider the equations for $(Z_1, Z_2)$. From (9.22) and uniform convergences observed above it is immediate that $R_n$ converges uniformly, a.s., to $R$ given as
\[ R(\cdot) = Z_1(0) - \int_0^1 (Z_1(s) - Z_2(s)) + \sqrt{2}B(\cdot). \]
Since $\theta_n = \alpha_n + O(\sqrt{n}/d_n) \rightarrow \alpha$, this shows that, for every $T < \infty$,
\[ \Gamma_{\theta_n}(R_n)(t) = R_n(t) - \sup_{s \in [0, t]} (R_n(t) - \theta_n)^+ \]
\[ \Rightarrow R(t) - \sup_{s \in [0, t]} (R(t) - \alpha)^+ = \Gamma_{\alpha}(R)(t) \]
uniformly for $t \in [0,T]$, a.s., where $(R(t) - \alpha)^+$ is taken to be 0 when $\alpha = \infty$. Similarly,
\[ \hat{\Gamma}_n(R_n)(t) \to \hat{\Gamma}_\alpha(R)(t) \]
uniformly for $t \in [0,T]$, a.s. Here, when $\alpha = \infty$, $\Gamma_\alpha$ and $\hat{\Gamma}_\alpha$ are as introduced in (2.11). The fact that $(Z_1, Z_2)$ solve the first two equations in (2.13) is now immediate from Lemma 9.4, the convergence $\tilde{Z}_{n,1} - Z_{n,1} \to 0$, and the uniform convergence of $W_{n,2}$ to 0 noted previously. The result follows.

Acknowledgements
We thank the two anonymous referees for a very careful review of the manuscript that led to numerous improvements in the work. Research of SB is supported in part by NSF grants DMS-1613072, DMS-1606839 and ARO grant W911NF-17-1-0010. Research of AB is supported in part by the National Science Foundation (DMS-1814894 and DMS-1853968). Research of MD is supported by the NSF grant DMS-1613072 and NIH R01 grant HG009125-01.

References
Differentiating \( \Delta \)

Let \( n \)


\[ E. J. McShane and T. A. Botts, \text{Real Analysis}, \text{Courier Corporation}, 2013.

\[ M. Mitzenmacher, \text{The power of two choices in randomized load balancing}, \text{IEEE Transactions on Parallel and Distributed Systems} \textbf{12} (2001), no. 10, 1094–1104.


\textbf{Appendix A. Proofs of results in Section \text{5}}

\textbf{A.1. Proof of Lemma 5.1}

Proof. Fix \( \epsilon \in (0, 1) \). First suppose \( \frac{d_n}{n} \to 0 \). Consider \( x \in (\epsilon, 1] \). Let \( \Delta_n(x) = \log \beta_n(x) - \log \gamma_n(x) \).

Let \( n_0 \in \mathbb{N} \) be such that for all \( n \geq n_0 \), \( \frac{d_n}{n} < \epsilon/2 \). Then, for \( n \geq n_0 \),

\[
\Delta_n(x) = \sum_{i=0}^{d_n-1} \log \left( \frac{x - i/n}{1 - i/n} \right) - \log x = \sum_{i=0}^{d_n-1} \left\{ \log \left( \frac{x - i/n}{1 - i/n} \right) - \log x \right\},
\]

\[
= \sum_{i=0}^{d_n-1} \log \left( \frac{1 - i/(nx)}{1 - i/n} \right) = \sum_{i=0}^{d_n-1} \log \left( 1 - (i/n) \frac{1}{1 - i/n} \right). \tag{A.1}
\]

Differentiating \( \Delta_n \) gives,

\[
\dot{\Delta}_n(x) = \sum_{i=0}^{d_n-1} \left( \frac{1}{x - i/n} - \frac{1}{x} \right) = \sum_{i=0}^{d_n-1} \frac{i/n}{x(x - i/n)}.
\]

Since \( n \geq n_0 \) and \( x \in [\epsilon, 1] \) we have \( x(x - i/n) \geq \epsilon^2/2 \) for \( i \leq d_n - 1 \). Hence,

\[
\left| \dot{\Delta}_n(x) \right| \leq \frac{2}{\epsilon^2} \sum_{i=0}^{d_n-1} (i/n) \leq \frac{1}{\epsilon^2} \frac{d_n^2}{n}.
\]

From the definition of \( \Delta_n \), we also have,

\[
\dot{\Delta}_n(x) = \frac{\dot{\beta}_n(x)}{\beta_n(x)} - \frac{\dot{\gamma}_n(x)}{\gamma_n(x)} = \frac{\dot{\gamma}_n(x)}{\gamma_n(x)} \left( \frac{\dot{\beta}_n(x)}{\beta_n(x)} \frac{\gamma_n(x)}{\gamma_n(x)} - 1 \right). \tag{A.2}
\]
Hence (A.6) holds for each \( j \), from (A.2) we have,
\[
\sup_{x \in [\epsilon, 1]} \left| \frac{\hat{\beta}_n(x)}{\gamma_n(x)} - 1 \right| \leq \frac{1}{d_n} \sup \frac{1}{x \in [\epsilon, 1]} \left| \Delta_n(x) \right| \leq \frac{1}{\epsilon} \to 0.
\]
This proves (5.4).

Now assume \( \frac{d_n}{\sqrt{n}} \to 0 \). Once more consider \( x \in (\epsilon, 1] \) and \( n \geq n_0 \). Let \( C = \sup_{n \geq n_0} \frac{1/\epsilon - 1}{1-d_n/n} < \infty \) and let \( n_1 > n_0 \) be such that \( d_n C/n < 1/2 \) for all \( n \geq n_1 \). Then for \( n \geq n_1 \) and \( x \in [\epsilon, 1] \):
\[
|\Delta_n(x)| \leq \sum_{i=0}^{d_n-1} 2 \left| (i/n) \right| x^{-1} \left| 1 - i/n \right| \leq 2C \sum_{i=0}^{d_n-1} i/n \leq C \frac{d_n^2}{n},
\]
where the first inequality is from (A.1) and the inequality \( |\log(1 + h)| \leq 2|h| \) for \( |h| \leq 1/2 \). This shows \( \sup_{x \in [\epsilon, 1]} |\Delta_n(x)| \to 0 \), hence showing the first convergence in (5.5). Finally the second convergence (5.4) is immediate on combining the first convergence with (5.4).

**A.2. Proof of Corollary 5.2:** This is an immediate consequence of the estimate in (A.3).

**A.3. Proof of Corollary 5.3:**

**Proof.** Let \( \epsilon > 0 \) and \( n_0 \in \mathbb{N} \) be such that \( \mu_{n,i} > \epsilon \) for all \( n \geq n_0 \). By Lemma 5.1 as \( n \to \infty \)
\[
\frac{\hat{\beta}_n(\mu_{n,i})}{\beta_n(\mu_{n,i})} = (1 + o(1)) \frac{\hat{\gamma}_n(\mu_{n,i})}{\gamma_n(\mu_{n,i})}.
\]
Recall that \( \mu_{n,i+1} = \lambda_n \beta_n(\mu_{n,i}) \) and \( \gamma_n(x) = x^{d_n} \). Hence (A.4) gives
\[
\frac{\hat{\beta}_n(\mu_{n,i})}{\mu_{n,i+1}/\lambda_n} = (1 + o(1)) \frac{d_n}{\mu_{n,i}}
\]
completing the proof.

**A.4. Proof of Lemma 5.4**

**Proof.** From Corollary 5.2 there is a \( n_0 \in \mathbb{N} \) and \( C \in (0, \infty) \) such that for all \( n \geq n_0 \)
\[
\sup_{x \in [\epsilon, 1]} |\log \beta_n(x) - \log \gamma_n(x)| \leq C \frac{d_n^2}{n}.
\]
Thus, if for \( n \geq n_0 \) and \( i \in \mathbb{N} \), \( \mu_{n,i} \geq \epsilon \), then
\[
\log \mu_{n,i+1} = \log \lambda_n + \log \beta_n(\mu_{n,i}) = \log \lambda_n + \log \gamma_n(\mu_{n,i}) + \gamma_n \leq \log \lambda_n + d_n \log \mu_{n,i} + \gamma_n.
\]

where \( |\gamma_n| \leq C_0 d_n^2 \). Now let \( k \in \mathbb{N} \) and \( n_1 \in \mathbb{N} \) be such that for all \( n \geq n_1, \mu_n \geq \epsilon \). We will show that for \( n \geq n_0 \lor n_1 \) and \( j \in \{1, \ldots, k\} \) that
\[
\log \mu_{n,j+1} = (\log \lambda_n) \left( \sum_{i=0}^{j} d_n^i \right) + \beta_n \cdot j,
\]
where \( |\beta_n j| \leq \frac{C}{2} \sum_{i=1}^{j} d_n^{i+1} \). Note the the lemma is immediate from (A.7) on taking \( j = 1 \). To prove (A.7) we argue inductively. First note that since \( \mu_n \in l_{\epsilon}^1 \), \( \mu_{n,i} \geq \mu_{n,k} \geq \epsilon \) for each \( i \leq k \) and \( n \geq n_1 \). Hence (A.6) holds for each \( i \leq k \) and \( n \geq n_0 \lor n_1 \). Taking \( i = 1 \) in (A.6) and noting that \( \mu_{n,1} = \lambda_n \) proves (A.7) for the case \( j = 1 \).

Suppose now (A.7) holds for some \( j \leq k - 1 \). Then, using \( i = j + 1 \), in (A.6)
\[
\log \mu_{n,j+2} = \log \lambda_n + d_n \log \mu_{n,j+1} + \gamma_{n,j+1},
\]
where \(|\gamma_{n,j+1}| \leq \frac{Cd^2_n}{n}\). By the induction hypothesis, \((A.7)\) holds for \(j\). Hence

\[
\log \mu_{n,j+2} = \log \lambda_n + d_n \left\{ (\log \lambda_n) \left( \sum_{i=0}^{j} d_n^i \right) + \beta_{n,j} \right\} = (\log \lambda_n) \left( \sum_{i=0}^{j+1} d_n^i \right) + d_n \beta_{n,j} + \gamma_{n,j}
\]

and hence \(\beta_{n,j+1} = d_n \beta_{n,j} + \gamma_{n,j}\). This shows

\[
|\beta_{n,j+1}| = |d_n \beta_{n,j} + \gamma_{n,j}| \leq d_n \left( \frac{C}{n} \sum_{i=1}^{j} d_n^{i+1} + \frac{Cd_n^2}{n} \sum_{i=1}^{j+1} d_n^{i+1} \right)
\]

which shows that \((A.7)\) holds for \(j+1\). This completes the proof.

A.5. Proof of Corollary 5.5:

Proof: Since \(d_n \to \infty\), the assumption \(\frac{\varepsilon^2}{d_n} \to 0\) shows that \(\frac{\left|\xi_n\right|}{d_n} \leq \frac{1+\xi_n^2}{d_n} \to 0\). This shows that \(\varepsilon_n \to 1 - \lambda_n = \frac{\xi_n + \log d_n}{d_n}\) also converges to 0.

We first show that \(\mu_{n,i} \to 1\) for each \(i \in \{1, \ldots, k\}\). We will argue inductively. Since \(\mu_{n,1} = \lambda_n = 1 - \varepsilon_n\), we have \(\mu_{n,1} \to 1\). Suppose now that \(\mu_{n,i} \to 1\) for some \(i \leq k - 1\). Hence eventually \(\mu_{n,i} \geq \frac{1}{2}\). Applying Lemma 5.4 with \(k = i\) and \(\epsilon = \frac{1}{2}\) and simplifying the resulting expression, we get

\[
\log \mu_{n,i+1} = (\log \lambda_n) \frac{d_n^{i+1} - 1}{d_n - 1} + O \left( \frac{d_n^2 (d_n^{i+1} - 1)}{n(d_n - 1)} \right) \quad (A.8)
\]

\[
= O(\epsilon_n) \frac{d_n^{i+1} - 1}{d_n - 1} + O \left( \frac{d_n^2 (d_n^{i+1} - 1)}{n(d_n - 1)} \right) = O \left( \frac{\xi_n + \log d_n}{d_n^{k-i}} \right) + O \left( \frac{d_n^{i+1}}{n} \right). \quad (A.9)
\]

where the second equality uses \(\log \lambda_n = \log(1 - \varepsilon_n) = O(\varepsilon_n)\) and the third follows on recalling that \(d_n \to \infty\). Since \(i \leq k - 1\), \(\frac{\left|\xi_n\right|}{d_n} \leq \frac{1+\xi_n^2}{d_n} \to 0\). Using this along with \(d_n^{k+1} \ll n\) in \((A.9)\) shows that \(\mu_{n,i+1} \to 1\). Hence, by induction, \(\mu_{n,i} \to 1\) for \(i \leq k\).

Next we argue that \(\hat{\beta}_n(\mu_{n,k}) \to \alpha\). Since \(\lambda_n \to 1\) and \(\mu_{n,k} \to 1\), from Corollary 5.3 we have that

\[
\lim_{n \to \infty} \frac{\hat{\beta}_n(\mu_{n,k})}{d_n \mu_{n,k+1}} = 1
\]

Hence it suffices to show that \(d_n \mu_{n,k+1} \to \alpha\). For this note that

\[
\log(d_n \mu_{n,k+1}) = \log \mu_{n,k+1} + \log d_n = \log(1 - \varepsilon_n) \left( \frac{d_n^{k+1} - 1}{d_n - 1} \right) + O \left( \frac{d_n^2 (d_n^{k} - 1)}{d_n - 1} \right) + \log d_n
\]

\[
= (-\varepsilon_n + O(\varepsilon_n^2))d_n^k (1 + O(1/d_n)) + \log d_n + O \left( \frac{d_n^{k+1}}{n} \right),
\]

where the second equality is from \((A.8)\) and last equality is by using Taylor’s expansion for \(\log(1-\varepsilon_n)\). Using \(d_n^{k+1} \ll n\) and \(\left|\varepsilon_n d_n^k\right| \leq 2(\xi_n + \log d_n)^2 \to 0\), we now have

\[
\log(d_n \mu_{n,k+1}) = (-\varepsilon_n d_n^k + o(1))(1 + O(1/d_n)) + \log d_n + o(1) = (-\xi_n - \log d_n)(1 + O(1/d_n)) + \log d_n + o(1)
\]

\[
= -\xi_n - \log d_n + \log d_n + O \left( \frac{\xi_n + \log d_n}{d_n} \right) + o(1) = -\xi_n + o(1) \to \log(1)
\]

where the last equality once more uses the observation that \(\frac{\left|\xi_n\right|}{d_n} \to 0\). Thus we have \(d_n \mu_{n,k+1} \to \alpha\) as \(n \to \infty\) which completes the proof. □
A.6. Proof of Lemma 5.6

Proof. Since \( \mu_{n,k} \to 1 \) and \( j \mapsto \mu_{n,j} \) is non-increasing, we have \( \mu_{n,i} \to 1 \) for each \( i \leq k \). Additionally, since \( \lambda_n \to 1 \), Corollary 5.3 shows that for any \( i \in [k] \) \( \lim_{n \to \infty} \frac{\hat{\beta}_n(\mu_{n,i})}{d_{n}^2 \mu_{n,i+1}} = 1 \). As a consequence, \( \hat{\beta}_n(\mu_{n,k-1}) \to \infty \) as \( n \to \infty \), and for any \( j \in [k-2] \)

\[
\lim_{n \to \infty} \frac{\hat{\beta}_n(\mu_{n,j})}{\beta_n(\mu_{n,j+1})} = \lim_{n \to \infty} \frac{d_{n}^2 \mu_{n,j+1}}{d_{n} \mu_{n,j+2}} = \lim_{n \to \infty} \frac{\mu_{n,j+1}}{\mu_{n,j+2}} = 1.
\]

This completes the proof of the lemma.

A.7. Proof of Lemma 5.7

Proof. By the first part of Lemma 5.1 (5.7) is immediate from (5.6). Now consider (5.6). Taking logarithms in (2.7), for \( x > d_n/n \),

\[
\log \beta_n(x) = \sum_{i=0}^{d_n-1} \left( \log \left( x - \frac{i}{n} \right) - \log \left( 1 - \frac{i}{n} \right) \right) = \sum_{i=0}^{d_n-1} \left( \log \left( 1 - \frac{i}{n} - (1 - x) \right) - \log \left( 1 - \frac{i}{n} \right) \right).
\]

Let \( \delta_n = \epsilon_n + \frac{d_n}{n} \). For large \( n \), \( \delta_n \leq \frac{1}{2} \), and hence, using the expansion \( \log(1-h) = -h + O(h^2) \) for \( |h| \leq \frac{1}{2} \), for any \( x \in [1 - \epsilon_n, 1] \):

\[
\log \beta_n(x) = \sum_{i=0}^{d_n-1} \left\{ -\frac{i}{n} - (1 - x) + \frac{i}{n} + O(\delta_n^2) \right\} = -d_n(1 - x) + O(d_n\delta_n).
\]

Note that \( \delta_n^2 = (\epsilon_n + d_n/n)^2 \leq 2 \left( \frac{\epsilon_n}{n} + \frac{d_n^2}{n^2} \right) \). Hence by our assumptions \( d_n\delta_n^2 \to 0 \). This proves (5.6) and completes the proof of the lemma.

A.8. Proof of Lemma 5.8

Proof. By (5.2)

\[
\sup_{x \in [0, 1 - \epsilon_n]} |\beta_n(x)| \leq (1 - \epsilon_n)^{d_n} = e^{-d_n\epsilon_n + o(1)} \to 0.
\]

Similarly, by (5.3), under the assumption \( \lim sup_n \frac{d_n}{n} < 1 \), for large \( n \),

\[
\sup_{x \in [0, 1 - \epsilon_n]} |\hat{\beta}_n(x)| \leq (1 - d_n/n)^{-1} d_n(1 - \epsilon_n)^{d_n - 1} = e^{-d_n\epsilon_n + \log d_n + O(1)} \to 0.
\]

Appendix B. Proof of Lemma 6.8

For a right continuous bounded variation function \( F : [0, T] \to \mathbb{R} \), let \( dF \) denote the signed measure on \((0, T)\) given by \( dF(a, b] = F(b) - F(a) \) for \( 0 \leq a < b \leq T \), and \( d\lambda \) denote the Lebesgue measure on \((0, T]\). Bounded measurable functions \( h : [0, T] \to \mathbb{R} \) act on signed measure \( d\mu \) on \((0, T]\) on the left as follows: \( hd\mu \) denotes the signed measure \( A \mapsto \int_A h(x)d\mu(x) \), \( A \in B(0, T] \).

Let \( H(t) = \int_0^t h(s)d\lambda(s) \) for \( t \in [0, T] \). Note that \( z \) defined in (6.14) is a right continuous function with bounded variations. The corresponding measure \( dz \) on \((0, T] \) satisfies the identity

\[
dz = -hzd\lambda + gd\lambda + dM,
\]

namely

\[
 dz + hzd\lambda = gd\lambda + dM.
\]
Acting on the left in the above identity by the bounded continuous function \( e^H(t) = e^{H(t)} \) we get
\[
e^H dz + e^H h z d\lambda = e^H g d\lambda + e^H dM.
\]
Since \( dH = hd\lambda \), by the change of variable formula (cf. [29, Theorem VI.8.3]) \( de^H = he^H d\lambda \). Hence
\[
e^H dz + z de^H = e^H g d\lambda + e^H dM.
\]
Two applications of the integration by parts formula (cf. [2, Theorem 18.4]) show that
\[
d(e^H z) = e^H g d\lambda + d(e^H M) - M de^H.
\]
Computing the total measure on \((0, t]\) for \( t \leq T \):
\[
e^H(t) z(t) - z(0) = \int_0^t e^{H(s)} g(s) d\lambda(s) + e^{H(t)} M(t) - M(0) - \int_0^t M(s) de^H(s)
\]
Rearranging terms and multiplying by \( e^{-H(t)} \) on both sides and noting, from (6.14), that \( M(0) = 0 \):
\[
z(t) = \int_0^t e^{H(s) - H(t)} g(s) d\lambda(s) + M(t) - e^{-H(t)} \int_0^t M(s) de^H(s) + e^{-H(t)} z(0).
\]
We now estimate the various terms on the right hand side of (B.1). The first term on the right hand side of (B.1) satisfies for \( t \in [0, T \wedge \tau] \)
\[
\left| \int_0^t e^{H(s) - H(t)} g(s) d\lambda(s) \right| \leq |g|_{*, T \wedge \tau} \int_0^t e^{-m(t-s)} d\lambda(s) \leq |g|_{*, T \wedge \tau} \frac{1 - e^{-mt}}{m} \frac{1}{m}.
\]
Next we estimate the third term in the right hand side of (B.1). Since \( h \) is non-negative on \([0, T \wedge \tau] \), \( de^H \) in a positive measure on \((0, T \wedge \tau] \). Hence for \( t \in [0, T \wedge \tau] \)
\[
\left| e^{-H(t)} \int_0^t M(s) de^H(s) \right| \leq |M|_{*, T \wedge \tau} e^{-H(t)} \int_0^t de^H(s) \leq |M|_{*, T \wedge \tau}.
\]
Finally, the last term in the right hand side of (B.1) for any \( t \in [0, \tau \wedge T] \) can be bounded as
\[
\left| e^{-H(t)} z(0) \right| \leq |z(0)| e^{-mt}.
\]
Using (B.2), (B.3) and (B.4) in (B.1) completes the proof of the lemma.