Convergence of jump processes with stochastic intensity to Brownian motion with inert drift

CLAYTON BARNES

1E-mail: cbarnes@campus.technion.ac.il

Consider a random walker on the nonnegative lattice, moving in continuous time, whose positive transition intensity is proportional to the time the walker spends at the origin. In this way, the walker is a jump process with a stochastic and adapted jump intensity. We show that, upon Brownian scaling, the sequence of such processes converges to Brownian motion with inert drift (BMID). BMID was introduced by Frank Knight in 2001 and generalized by White in 2007. This confirms a conjecture of Burdzy and White in 2008 in the one-dimensional setting.

Keywords: Brownian motion, Discrete Approximation, Random Walk, Local Time.

1. Introduction

Brownian motion with inert drift (BMID) is a process $X$ that satisfies the SDE

$$dX = dB + dL + KLdt,$$  \hspace{1cm} (1)

where $K \geq 0$ is a constant and $L$ is the local time of $X$ at zero. This process behaves as a Brownian motion away from the origin but has a drift proportional to its local time at zero. Note that such a process is not Markovian because its drift depends on the past history. See Figure 1 for sample path comparisons between reflected Brownian motion and BMID. BMID can be constructed path-by-path from a standard Brownian motion via the employment of a Skorohod map. This is discussed in more detail in Section 2.

We consider continuous time processes $(X_n, V_n)$ on $2^{-n} \mathbb{N} \times \mathbb{R}$ such that for $K \geq 0$,

(i) $V_n(t) := K2^n \cdot \text{Leb}(0 < s < t : X_n(s) = 0)$ is the scaled time $X_n$ spends at the origin.

(ii) $X_n$ is a jump process with positive jump intensity $2^{2n} + 2^n V_n(t)$ and downward jump intensity $2^{2n}$, modified appropriately so $X_n$ does not transition below zero.

The existence of such a process and its rigorous definition is presented in Section 3. Intuitively, $X_n$ is a random walker on the lattice $2^{-n} \mathbb{N}$ whose transition rates depend linearly on the amount of time the walker spends at zero. In other words, the positive jump rate of $X_n$ increases each time $X_n$ reaches zero. We show that as the lattice size shrinks to zero, i.e. as $n \to \infty$, $(X_n, V_n)$ converges in distribution to $(X, V)$, where $X$ is BMID and $V = KL$ is its velocity. See Theorem 4.4 and Corollary 4.5 for precise statements. By setting $K = 0$ we recover the classical result that random walk on the nonnegative lattice converges to reflected Brownian motion.
1.1. Outline

In Section 2 we introduce BMID and its construction using Skorohod maps. We also give an equivalent formulation of the process \((X,V)\). In Section 3 we introduce the necessary background on jump process with stochastic intensity and introduce the setting used by Burdzy and White in [6]. Section 4 contains the statement and proof of the main results, Theorem 4.4 and Corollary 4.5. We conclude by briefly discussing BMID in a multidimensional setting in Section 5.

1.2. Background

The study of BMID began in 2001 when Knight [11] described a Brownian particle reflecting above a particle with Newtonian dynamics. This two-particle system of Knight is equivalent in some sense to BMID in that the gap between the Brownian particle and the Newtonian particle is BMID. See Section 2. For more background on BMID see [13], where White constructs a multidimensional analog to BMID; [3], where Bass, Burdzy, Chen and Hairer study the stationary distribution; [1], where Barnes describes the hydrodynamic behavior of systems of Brownian motions with inert drift.

Burdzy and White studied similar processes from a discrete state point of view [6]. They consider a pair of processes \((X,L)\) with state space \(\mathcal{L} \times \mathbb{R}^d\), where \(\mathcal{L}\) is a finite set, and where the transition rate of \(X\) depends on \(L\), the scaled time \(X\) has spent.
on previous states. See subsection 3.2 for definitions. The authors find necessary and sufficient conditions for such a process \((X,L)\) to have stationary distribution \(\mu \times \gamma\), where \(\mu\) is uniform on space and \(\gamma\) is Gaussian. Burdzy and White make many conjectures involving approximating BMID, and its variants, and suggest the results of Bass, Burdzy, Chen, and Hairer [3] concerning a multidimensional analog of BMID stem from a discrete approximation scheme where the continuous process of BMID is a limit of these processes whose values take place in discretized space. The main result of this article confirms the discrete approximation scheme converges to the continuous model in the one dimensional setting.

BMID is just one example where a process with memory has a Gaussian stationary distribution. Gauthier [8] studies diffusions whose drift is also dependent on the process history through a linear combination of sine and cosine functions. He shows the average displacement across time obtains a Gaussian stationary distribution as time approaches infinity. In [2], Barnes, Burdzy, and Gauthier use this discrete approximation scheme, taking limits of Markov processes in the same class considered here, to demonstrate billiards with certain Markovian reflection laws have \(\mu \times \gamma\) as the stationary measure for space and velocity, where as above \(\mu\) is the uniform stationary measure in the spatial component and \(\gamma\) the Gaussian stationary measure in the velocity component.

In Section 5 we briefly discuss a multidimensional analog of BMID that inspired conjectures of Burdzy and White.

2. An equivalent formulation of BMID

In this section we describe the process \((X,Y,V)\), where \(X\) is Brownian motion reflecting from the inert particle \(Y\) and where \(Y\) has velocity \(V\). We begin with a probability space \((\Omega,F,\mathbb{P},(\mathcal{F}_t))_{t\geq 0}\), with the filtration \((\mathcal{F}_t))_{t\geq 0}\) satisfying the usual conditions, supporting a Brownian motion \(B\).

Theorem 2.1 (Existence and Uniqueness, Knight [11], White [13]). Choose \(K \geq 0\) and \(v \in \mathbb{R}\). There exists a unique strong solution of continuous \(\mathcal{F}_t\)-adapted processes \((X,Y,V)\) satisfying:

\[
\begin{align*}
X(t) &= B(t) + L(t), \text{ for all } t \geq 0, \text{ almost surely,} \\
X(t) &\geq Y(t), \text{ for all } t \geq 0, \text{ almost surely,} \\
dY &= V(t)dt := (v - KL(t))dt, \text{ for all } t \geq 0, \text{ almost surely,} \\
L &\text{ is nondecreasing, and is flat away from the set } \{s : X(s) = Y(s)\}. 
\end{align*}
\]

(2)

Remark 2.2. Flatness of \(L\) off \(\{s : X(s) = Y(s)\}\) means \(\int_{0}^{\infty} 1(X(s) \neq Y(s)) dL(s) = 0\). One can use the Ito-Tanaka formula to show that \(L\) is the local time of \(X-Y\) at zero; see [13, Th. 2.7]. That is,

\[
L(t) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{0}^{t} 1(|X(s) - Y(s)| < \epsilon) ds,
\]
where the right hand side is the local time of $X - Y$ at zero.

BMID together with its velocity is equivalent, in a certain sense, to the process $(X, Y, V)$ whose existence is given in Theorem 2.1. We will refer to the following result by Skorohod.

**Lemma 2.3** (Skorohod, see [10]). Let $f \in C([0, T], \mathbb{R})$ with $f(0) \geq 0$. There is a unique, continuous, nondecreasing function $m_f(t)$ such that

$$x_f(t) = f(t) + m_f(t) \geq 0,$$

$$m_f(0) = 0, \text{ } m_f(t) \text{ is flat off } \{s : x_f(s) = 0\},$$

that is given by

$$m_f(t) = \sup_{0 < s < t} [-f(s)] \vee 0.$$  

**Remark 2.4.** The classical Lévy’s theorem states that for a Brownian motion $B$, $x_B$ is distributed as $|B|$. See [10, Section 3.6C].

To see the equivalence between BMID and the process $(X, Y, V)$ from Theorem 2.1, consider the gap process $G(t) = X(t) - Y(t)$. Obviously $G \geq 0$, almost surely, and from (2) it follows that (when $v = 0$)

$$dG = dB + dL + KLdt,$$

where $L$ is continuous, nondecreasing, and flat off $U^{-1}(0)$. From the comment in Remark 2.2 on local time, $G$ is a reflected diffusion whose drift is proportional to its local time at zero. Consequently, the gap process $G$ is BMID as it satisfies (1).

Assume we have a pair of processes $(U, V)$ adapted to a continuous filtration $\mathcal{F}_t$ that supports a Brownian motion $B$, and that for fixed $K \geq 0, v \in \mathbb{R}$

$$U(t) = B(t) + \int_0^t V(x) \, dx,$$

$$V(t) = -v + KM_U(t),$$

$$M_U(t) = \sup_{0 < s < t} [-U(s)] \vee 0.$$  

In the system (4), it is clear from the definition of $M_U$ that $U + M_U \geq 0$ and it follows from the Skorohod Lemma 2.3 that $M_U$ is flat off the set $\{s : U(s) + M_U(s) = 0\}$. Therefore

$$B(t) + M_U \geq -\int_0^t V(s) \, ds,$$

and $M_U$ is flat off of $\{s : B(t) + M_U = -\int_0^t V(s) \, ds\}$. Consequently,

$$\left( B(t) + M_U(t), -\int_0^t V(s) \, ds, -V(t) \right)$$
Approximating Brownian motion with inert drift by jump processes

satisfies the original equation (2) with respect to the filtration $\mathcal{F}_t$. Similarly, one can use the uniqueness statement in Skorohod’s Lemma to go from a solution of (4) to a solution of (2). This demonstrates the equivalence of the two systems (2) and (4) in the sense that if one solution exists for a given probability space $(\Omega, \mathbb{P}, \mathcal{F}_t)$, where $\mathcal{F}_t$ supports a given Brownian motion, then the other solution can be given by a path-by-path transformation.

Existence of a solution to (2) was first shown by Knight in [11]. A strong solution to a more general process was attained via the employment of a Skorohod map by David White [13] in a more general version of Theorem 2.5 given below.

**Theorem 2.5** (White, [13]). For every $f \in C([0, T], \mathbb{R}), K \geq 0, v \in \mathbb{R}$ there is a unique pair of continuous functions $(I, V)$ such that

$$
\begin{align*}
x(t) := f(t) + I(t), \\
V(t) = v + Km(t), \\
I(t) &= \int_0^t V(s) \, ds, \\
m(t) &= \sup_{0 \leq s \leq t} [-x(s) \vee 0].
\end{align*}
$$

**Remark 2.6.** In Remark 2.4 it is mentioned that replacing the function $f$ with a Brownian motion in the formulation of Skorohod’s Lemma 2.3 gives rise to a representation of reflected Brownian motion. Similarly, when replacing $f$ in Theorem 2.5 pathwise by Brownian motion, the corresponding process $(x, -I, V)$ is a solution to (4). Note that Skorohod’s Lemma 2.3 implies that $m(t)$ in Theorem 2.5 is the unique monotonically increasing, continuous, function which is flat off of the level set $\{s : x(s) + m(s) = 0\}$ such that $x + m$ is nonnegative.

**Remark 2.7.** One can also see from the above arguments that $(U, V)$ of (4) solves

$$
dU(t) = dB(t) + dL(t) + V(t)dt, \quad dV(t) = KL(t)dt, \quad V(0) = -v,
$$

where $L$ is the local time of $U$ at zero.

### 3. Markov Processes with Memory

The title of this section seems contradictory because Markov processes lose their memory when conditioning on their current location. The processes considered are pairs of processes, one process taking values in “space,” and the other process storing the history of the space-valued process. The transition rate of the space-valued process depends on this stored history. We let $\mathcal{C}$ denote the class of such processes which we introduce more formally in this section. We will later construct a sequence of processes in $\mathcal{C}$ that will approximate BMID. First, we review well known facts of Poisson processes and point process with stochastic intensity. For reference, see Brémaud’s description of a doubly-stochastic point process in [5, Chapter 2].
3.1. Non-homogeneous Poisson processes

A non-homogeneous Poisson process with a nonnegative locally integrable rate (or intensity) function \( \lambda(t) \) is a process \( N \) such that

(i) \( N(0) = 0 \), a.s.
(ii) \( N \) has independent increments,
(iii) \( N \) is RCLL, a.s.
(iv) \( N(a,b] = N(b) - N(a) \overset{d}{=} \text{Poisson}(\int_a^b \lambda(s)) \).

If we let \( T = \inf\{ t : N(t) > 0 \} \) be the first jump time of \( N \), then
\[
\mathbb{P}(T > t) = \mathbb{P}(N(t) = 0) = e^{-\int_0^t \lambda(s) \, ds}.
\]

**Lemma 3.1.** Let \( \lambda_1, \lambda_2 \) be two rate functions such that \( \lambda_1(t) \leq \lambda_2(t) \) for all \( t \geq 0 \) and let \( T_1, T_2 \) be the first jump time of their corresponding Poisson process. Then \( T_1 \) stochastically dominates \( T_2 \).

In [5, Chapter II], Brémaud discusses the notion of point processes adapted to a filtration \( \mathcal{F}_t \) whose intensity \( \lambda(s) \) is not a deterministic function but rather a process adapted to \( \mathcal{F}_t \) with certain conditions.

**Definition 3.2.** [5, II] Let \( N_t \) be a point process adapted to the filtration \( \mathcal{F}_t \) and let \( \lambda_t \) be a nonnegative \( \mathcal{F}_t \)-progressive process such that \( \int_0^t \lambda_s \, ds < \infty \) almost surely for each \( t \in [0,T] \). If
\[
\mathbb{E}
\left( \int_0^\infty C_s \, dN_s \right) = \mathbb{E}
\left( \int_0^\infty C_s \lambda_s \, ds \right),
\]
for all nonnegative \( \mathcal{F}_t \)-predictable processes \( C_t \) then we say \( N_t \) has stochastic intensity \( \lambda_t \).

**Remark 3.3.** In the proofs of later results we will refer to point processes with a given intensity or jump/step size. By a point process of jump/step size \( a > 0 \) and (stochastic) intensity \( \lambda \) we mean a process \( aN_t \) where \( N_t \) is a point process with (stochastic) intensity \( \lambda \). By the positive (resp. negative) jump process for a process we mean the process \( aN_t \) (resp. \( -aN_t \)). For example, a process with jump size \( 2^{-n} \) with positive jump rate \( \lambda_1(t) \), and negative jump rate \( \lambda_2(t) \), is \( 2^{-n}(N_1 - N_2) \) where \( N_i \) is a point process with (stochastic) rate \( \lambda_i \).

Some well known facts of Poisson processes have analogous results for Poisson processes with stochastic intensities, which we list below.

**Lemma 3.4.** Let \( N_1, N_2 \) be two independent point processes with stochastic intensities \( \lambda_1, \lambda_2 \) adapted to filtrations \( \mathcal{F}^1, \mathcal{F}^2 \) respectively. Then \( N_1 + N_2 \) is a point process with stochastic intensity \( \lambda_1 + \lambda_2 \), adapted to \( \mathcal{F}_t := \sigma(\mathcal{F}^1_t, \mathcal{F}^2_t) \).
Sketch. The fact that $N_1, N_2$ are independent implies the two processes do not have common jumps, so that $(N_1, N_2)$ is a multivariate point process. The result follows from [5, T15, Chapter II.2].

Lemma 3.5. Let $N_1$ be a point process with stochastic intensity $\lambda^1(t) \geq \lambda$, almost surely, for some $\lambda \in \mathbb{R}_+$. Then we can enlarge the probability space to support a Poisson point process $N_2$ with constant intensity $\lambda$ and a point process $N_3$ of stochastic intensity $\lambda^3 = \lambda^1 - \lambda$ such that $N_3$ is independent of $N_2$ and $N_2 + N_3$ has stochastic intensity $\lambda^1$.

Remark 3.6. It is clear that one can generate a Poisson point process $N_2$ with constant intensity $\lambda$ which is independent of $N_1$. Lemma 3.5 could be generalized to include more general lower bounds than a constant, however, we don’t require this and sketch the proof only in the case when $N_2$ has constant intensity.

Sketch. Enlarge the probability space to support two independent processes $N'_2, N'_3$ where $N'_2$ is a Poisson point process of rate $\lambda$ and $N'_3$ is a point process with stochastic rate $\lambda^3 = \lambda^1 - \lambda$. By Lemma 3.4, $N_1 \overset{d}{=} N_2 + N_3$, and the processes are adapted to the filtration generated by $(N_1, N_2, N_3)$.

3.2. Class C of Markov Processes with Memory

As mentioned, Burdzy and White [6] study continuous time Markov processes $(X, L)$ on $\mathcal{L} \times \mathbb{R}$ where $\mathcal{L} = \{1, 2, \ldots, N\}$ is a finite set. For each $j \in \mathcal{L}$ we associate a vector $v_j \in \mathbb{R}$ and define $L_j(t) = \text{Leb}(0 < s < t : X(s) = j)$ as the time $X$ has spent at location $j$ until time $t$. We also define

$$L(t) = \sum_{j \in \mathcal{L}} v_j L_j(t)$$

as the accumulated time $X$ spends at each location, weighting the time spent at location $j$ by the factor $v_j$. The transition rates of $X$ will depend on $L(t)$. More precisely, we are given RCLL functions $a_{ij} : \mathbb{R} \to \mathbb{R}$ where $a_{ij}$ is the rate function for the Poisson process defining the transition of $X$ from $i$ to $j$. Conditional on $X(t_0) = i, L(t_0) = l$, the jump rate of $X$ transitioning from $i$ to $j$ is $a_{ij}(l + [t - t_0]v_i)$ with $t \geq t_0$. To construct such a process, for each $i$ we create independent random variables $(T_{i,j})_{j \in \mathcal{L}}$ which represents the jump time from $i$ to $j$. Since this jump has intensity $a_{ij}(l + [t - t_0]v_i)$ with $t \geq t_0$,

$$\mathbb{P}(T_{i,j} > t + t_0 | X(t_0) = i, L(t_0) = l) = \exp \left( - \int_0^t a_{ij}(l + s v_i) \text{d}s \right),$$

(7)

for all $t > 0$. Pick $j'$ such that $T_{i,j'} = \min_{j \neq i} T_{i,j}$, and define the first transition of $X$ after time $t_0$ to be location $j'$ and occur at time $T_{i,j'}$. 
These dynamics can be produced from a collection of independent exponential random variables \((E_{i,j})_{i,j \in \mathbb{N}}\) of rate one. Set \(T_0 = 0\) and recursively define
\[
T_{i+1}^j = \inf \left( t > T_i : \int_0^t a_{X(T_j)}(L(T_i) + v_{X(T_j)}(s-T_i))ds > E_{i,j} \right) \tag{8}
\]
\[
T_{i+1} = \min_j T_{i+1}^j. \tag{9}
\]

We use the convention \(\inf \emptyset = \infty\). Define
\[
L(s) = L(T_i) + v_{X_n}(T_i)(s-T_i), \text{ for } s \in [T_i, T_{i+1}] \tag{10}
\]
\[
X(s) = X(T_i), \text{ for } s \in [T_i, T_{i+1}] \tag{11}
\]
\[
X(T_{i+1}) = \arg\min_j T_{i+1}^j. \tag{12}
\]

The pair \((X,L)\) is a strong Markov process with generator
\[
Af(j,l) = v_j \cdot \nabla f(j,l) + \sum_{i \neq j} a_{ji}(l)[f(i,l) - f(j,l)], \quad j = 1, \ldots, N, \quad l \in \mathbb{R}.
\]

Burdzy and White assume \((X,L)\) is irreducible in the sense that there is some \(\{j_0\} \times U \subset \mathcal{L} \times \mathbb{R}\) such that
\[
\mathbb{P}( (X(t), L(t)) \in \{j_0\} \times U | X(0) = i, L(0) = l ) > 0, \quad \text{for all } (i, l) \in \mathcal{L} \times \mathbb{R}.
\]

It should be noted that although they consider \(\mathcal{L}\) to be a finite set, their main results hold assuming that \(\sup_{ij} a_{ij}(l)\) is bounded on compact sets of \(l\) and \(\sup_{i} |v_i| < \infty\). We denote \(\mathcal{C}\) as the class of such processes with these conditions, allowing \(\mathcal{L} = \mathbb{N}\).

4. Discrete Approximation

4.1. Definition of Processes

The reflected diffusion (1) describing BMID is a process whose drift depends on the local time of the diffusion at zero. Intuitively, to approximate this diffusion with a Markov process on the lattice \(2^{-n} \mathbb{N}\) one would want the “velocity” to depend on the accumulated time spent at zero. This is modeled as a jump process whose intensity function is stochastic and depends linearly on the accumulated time the process spends at zero. These jump processes need to converge, as \(n \to \infty\), to a process whose drift is the appropriate local time.

In Section 2 we introduced an equivalent formulation for BMID given by \((U,V)\) in (4). In this subsection we will describe two equivalent discrete processes that mirror the equivalence of the continuous processes described earlier; see Proposition 4.3. We do this because in order to prove the convergence result described in the introduction we actually prove the convergence result for the equivalent formulation.
Jump processes whose intensity depends linearly on the accumulated time at zero are described by the class $C$ in subsection 3.2. Consider a process $(X_n, V_n^X)$ on the state space $2^{-n}\mathbb{N} \times \mathbb{R}$ where $v_j = 0$ for all $j \neq 0$, $v_0 = K 2^n$ as given in the notation in that subsection. (We may hide the dependence on $n$ for convenience.) For an initial “velocity” $v \in \mathbb{R}$, we define

$$V_n^X(t) = -v + K 2^n L_n(t) = -v + K 2^n \cdot \text{Leb}(0 < s < t : X_n(s) = 0).$$

The rate functions $a_{ij} : \mathbb{R} \to \mathbb{R}$ are

$$a_{i(+\text{sign}(l))2^{-n}}(l) = 2^{2n} + 2^n |l| = 2^{2n} + 2^n |V_n^X(t)|,$$

$$a_{i(-\text{sign}(l))2^{-n}}(l) = 2^{2n},$$

where $l = V_n^X(t)$, except when $i = 0$ where we do not allow a downward transition. By Lemma 3.4 the jump process $X_n$ can be decomposed into a sum of independent processes, $S_n$ and $Z_n$, whose rate functions sum to that of $X_n$. The following definition will be used throughout the paper.

**Definition 4.1.** For a process $Q(t)$ we define $M^Q(t)$ as the signed running minimum below zero of $Q$. That is,

$$M^Q(t) = \sup_{0 < s < t} [-Q(s)] \vee 0.$$

**Definition 4.2.** Consider the processes $(S_n, Z_n, V_n)$ on $(2^{-n}\mathbb{Z})^2 \times \mathbb{R}$ where

(i) $S_n$ is a continuous time simple random walk on $2^{-n}\mathbb{Z}$ with positive (and negative) jumps of size $2^{-n}$ and rate $2^{2n}$.

(ii) $Z_n$ is a point process with jump size $2^{-n}$ and with positive (resp. negative) jumps having stochastic and adapted rate $2^n |V_n|$ when $V_n > 0$ (resp. $V_n < 0$).

(iii) We have

$$V_n(t) = -v + K 2^n \cdot \text{Leb}(0 < s < t : U_n(s) = -M_n(s)),$$

$$U_n = S_n + Z_n,$$

$$M_n(t) := M^{U_n}(t).$$

That is, $S_n$ and $Z_n$ are point processes with adapted intensity functions as discussed in [5, Chapter 2].

Note the similarity to the equivalent formulation of BMID given by $(U,V)$ in (4) to $(U_n,V_n)$ given above. Existence of $(S_n, Z_n, V_n)$ follows from the fact that it is of class $C$, or, equivalently, one can construct the processes via the dynamics given in (8) by using the intensity functions (13).

**Proposition 4.3.** The processes $(U_n + M_n, V_n)$ and $(X_n, V_n^X)$ have the same law.
Proof. With these definitions $(U_n + M_n, V_n)$ has the same law as $(X_n, V_n^X)$ because it is of class $C$ and satisfies (13). To see this, note that $V_n$ is adapted to the right continuous filtration $\mathcal{F}_t$ generated by the pair $(S_n, Z_n)$. Also note that $U_n + M_n = S_n + Z_n + M_n$ is a nonnegative process on $2^{-n}\mathbb{N}$. By Lemma 3.4, $S_n + Z_n$ has a jump rate function of $2^n + 2^n V_n(t)$ where
\[ V_n(t) = -v + K2^n \cdot \text{Leb}\{0 < s < t : U_n(s) = -M_n(s)\} \]
\[ = -v + K2^n \cdot \text{Leb}\{0 < s < t : U_n(s) + M_n(s) = 0\}. \]
Consequently, $V_n^X = V_n$ if we define $X_n := U_n + M_n$. Therefore $(U_n + M_n, V_n)$ is one realization of the process $(X_n, V_n^X)$ given by (13).

We will work with $(S_n, Z_n, V_n)$ as an equivalent formulation of $(X_n, V_n^X)$ defined by (13).

4.2. Theorem Statement

The main result of this article is that $(S_n, V_n, Z_n, U_n)$ converges in an appropriate sense to $(B, V, \int_0^t V, U)$.

**Theorem 4.4.** For $K \geq 0$ and $v \in \mathbb{R}$, let $(S_n, Z_n, V_n, U_n)$ be given as in Definition 4.2 in subsection 4.1. Then
\[ (S_n, V_n, Z_n, U_n) \xrightarrow{d} (B, V, \int_0^t V, U), \]
in the Skorohod topology on $D([0, T], \mathbb{R}^4)$, where $(B, V, \int_0^t V, U)$ is a quadruple of continuous processes adapted to the Brownian filtration of the first coordinate $B$ with the following holding for all $t \in [0, T]$, almost surely:
\[ U(t) = B(t) + \int_0^t V(x) \, dx, \]
\[ V(t) = -v + KM_U(t), \]
and where $M_U$ is the running minimum given in Definition 4.1.

Theorem 4.4 has the following corollary.

**Corollary 4.5.** Let $(X_n, V_n^X)$ be the process defined by (13) in subsection 4.1. Then
\[ (X_n, V_n^X) \xrightarrow{d} (X, V), \]
in the Skorohod topology on $D([0, T], \mathbb{R})$. Here $(X, V)$ is BMID together with its velocity. That is,
\[ dX(t) = dB(t) + dL(t) + V(t) \, dt, \quad dV(t) = KL(t) \, dt, \quad V(0) = -v, \]
where $L$ is the local time of $X$ at zero.
Proof of Corollary. By Proposition 4.3 we set $X_n := U_n + M_n$ and $V_n^X := V_n$. Now Theorem 4.4 implies

$$(X_n, V_n^X) \xrightarrow{d} (U, -v + KM^U),$$

and where $(U, -v + KM^U)$ solves (14) as mentioned in Remark 2.7.

Remark 4.6. Let $D([0,T],\mathbb{R})$ denote the space of RCLL paths equipped with the Skorohod metric $d$ [7, Chapter 3 Section 5]. If a process $W_n$ with paths in $D([0,T],\mathbb{R})$ converges weakly to $W$, then according to the Skorohod representation, [7, Theorem 3.1.8], we can place $W_n, W$ on the same probability space such that $d(W_n, W) \rightarrow 0$, almost surely. If the limiting process $W$ is continuous almost surely, then

$$\|W_n - W\|_{[0,T]} := \sup_{0 \leq s \leq T} |W_n(s) - W(s)| \rightarrow 0,$$

almost surely on this probability space, and, in fact, uniform convergence and convergence in the Skorohod metric become equivalent. See Ethier and Kurtz, [7, Chapter 3 Section 5] and [7, Chapter 3 Section 10], and Billingsley [4, Chapter 3].

4.3. Proof of Theorem 4.4

In this section we prove Theorem 4.4 assuming the two lemmas below, one for tightness and the other for classifying the subsequential limits.

Lemma 4.7 (Tightness). The collection of processes $\{(S_n, Z_n, V_n) : n \in \mathbb{N}\}$ is tight in $D([0,T],\mathbb{R}^2) \times C[0,T] \subset D([0,T],\mathbb{R}^3)$. Because $U_n = S_n + Z_n$, it follows that $\{(S_n, V_n, Z_n, U_n) : n \in \mathbb{N}\}$ is also tight in $D([0,T],\mathbb{R}) \times C[0,T] \times D([0,T],\mathbb{R}^2)$. Furthermore, all limiting processes are continuous.

We prove Lemma 4.7 in Section 4.4. Assuming Lemma 4.7 holds, it remains to show there is a unique limit.

Lemma 4.8 (Classification of Limits). Consider a subsequence $n_k$ with processes $(S_{n_k}, Z_{n_k}, V_{n_k}, U_{n_k})$ converging to $(S, Z, V, U)$ in $D([0,T],\mathbb{R}^4)$ with the Skorohod topology. Then $(S, Z, V, U)$ is continuous and satisfies the equivalent formulation for BMID given in (4). That is,

(i) $U(t) = S(t) + Z(t),$

(ii) $S(t)$ is a Brownian motion,

(iii) $V(t) = KL(t) - v$, where $L(t) = \max_{0 \leq s \leq t} [-U(s)] \lor 0$,

(iv) $Z(t) = \int_0^t V(s) \, ds.$
Lemma 4.8 is proved in Section 4.5.

**Proof of Theorem 4.4.** Since the formulation of BMID described by \((B, V, U)\) in the statement of Theorem 4.4 is unique in law, Lemmas 4.7 and 4.8 characterize the subsequential limits of \((S_n, Z_n, V_n)\). See [13] where existence (and uniqueness) of such a system is proved. Consequently, we have convergence of the entire sequence to this equivalent formulation of BMID.

4.4. Lemma 4.7: Tightness of \((S_n, Z_n, V_n)\)

Recall that our process \((S_n, Z_n, V_n)\) is in \(D([0, T], \mathbb{R}^3)\), the space of RCLL paths with the Skorohod topology defined by the product metric \(d \times d \times d\) where \(d\) is the Skorohod metric, see Billingsley [4]. The following definition is taken from Jacod and Shiryaev [9].

**Remark 4.9.** In general, it is not true that if \(\alpha_n \to \alpha, \beta_n \to \beta\) in the Skorohod topology, then \(\alpha_n + \beta_n \to \alpha + \beta\). However, this does hold if either \(\alpha\) or \(\beta\) is continuous. See Jacod and Shiryaev [9, Proposition VI.1.23]. Similarly, by Remark 4.6 one can assume a sequence \((S_{nk}, Z_{nk}, V_{nk})\) that converges in distribution on \(D([0, T], \mathbb{R}^3)\) in fact converges almost surely to a continuous process, say \((S', Z', V')\) in the uniform metric if the limit is continuous; at least on some probability space. This immediately implies \(U_{nk} = S_{nk} + Z_{nk}\) converges almost surely to \(S' + Z'\).

**Definition 4.10.** [9, Definition VI.3.25] A sequence of processes \(\{X_i : i \in \mathbb{N}\}\) in \(D([0, T], \mathbb{R})\) is said to be \(C\)-tight if \(\{X_i : i \in \mathbb{N}\}\) is tight on \(D([0, T], \mathbb{R})\) and all limiting processes are continuous.

See Remark 4.6, which mentions the Skorohod metric when the limit processes are continuous. The proof that \((S_n, Z_n, V_n)\) is tight is broken into multiple lemmas. Recall that \(L_n(t) := 2^n \text{Leb}(0 < s < t : U_n(s) = -M_n(s))\) where \(U_n := S_n + Z_n, M_n := M_{U_n}\), and \(V_n := -\nu + KL_n\).

**Definition 4.11.** For \(f \in D([0, T], \mathbb{R})\), let

\[\omega(f, \delta) := \sup_{0 < s < t < T, |t-s| < \delta} |f(t) - f(s)|\]

be the modulus of continuity of \(f\).

Recall that \(\|f\|_{[a,b]} = \sup_{a < x < b} |f(x)|\).

**Lemma 4.12.** [9, Proposition VI.3.26] A sequence of processes \(X_n\) in \(D([0, T], \mathbb{R})\) is \(C\)-tight if and only if for every \(\epsilon > 0\),

\[\|X_n - X\|_{[a,b]} < \epsilon.\]
Lemma 4.15. \( \lim_{C \to \infty} \limsup_{n \to \infty} \mathbb{P}(\|X_n\|_{[0,T]} \geq C) = 0. \)

(ii) \( \lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}(\omega(X_n, \delta) > \epsilon) = 0. \)

Remark 4.13. Notice that (i) follows from (ii) and \( \lim_{C \to \infty} \limsup_{n \to \infty} \mathbb{P}(|X_n(0)| > C) = 0. \) To see this, take \( \delta = 1 \) and \( \epsilon = C/2 \) so that by definition of the modulus of continuity

\[
\|X_n\|_{[0,T]} \leq |X_n(0)| + \sum_{i=1}^{T} \omega(X_n, 1).
\]

Consequently, the triangle inequality gives

\[
\mathbb{P}(\|X_n\|_{[0,T]} > C) \leq \mathbb{P}(\|X_n(0)\| + |T| \omega(X_n, 1) > C)
\leq \mathbb{P}(\|X_n(0)\| > C/2) + \mathbb{P}(\omega(X_n, 1) > C/(2|T|)),
\]

where \([r]\) is the smallest integer larger than \(r\), from which it is clear that (ii) and \( \lim_{C \to \infty} \limsup_{n \to \infty} \mathbb{P}(|X_n(0)| > C) = 0 \) imply (i).

Lemma 4.14. Assume that the sequences of processes \((X_n), (X'_n), (X''_n)\) in \(D([0,T], \mathbb{R})\) satisfy

\[
X'_n(t) - X'_n(s) \leq X_n(t) - X_n(s) \leq X''_n(t) - X''_n(s), \quad 0 \leq s \leq t,
\]

almost surely. If both \((X'_n)\) and \((X''_n)\) are \(C\)-tight, then \((X_n)\) is also \(C\)-tight.

Proof. For any \(C > 0\), the triangle inequality gives

\[
\mathbb{P}(\|X_n\|_{[0,T]} > C) \leq \mathbb{P}(\|X'_n\|_{[0,T]} > C/2) + \mathbb{P}(\|X''_n\|_{[0,T]} > C/2),
\]

which verifies condition (i) in the statement of Lemma 4.12 by taking \( \lim_{C \to \infty} \limsup_{n \to \infty} \) of both sides. Similarly, for every \(\delta, \epsilon > 0\),

\[
\mathbb{P}(\omega(X_n, \delta) > \epsilon) \leq \mathbb{P}(\omega(X'_n, \delta) > \epsilon/2) + \mathbb{P}(\omega(X''_n, \delta) > \epsilon/2),
\]

and taking \( \lim_{\delta \to 0} \limsup_{n \to \infty} \) on both sides verifies condition (ii) in the statement of Lemma 4.12. \( \square \)

The following two lemmas are classical and we omit proofs.

Lemma 4.15. Let \(X \overset{d}{=} \text{Exp}(\lambda), Y \overset{d}{=} \text{Exp}(\mu)\) be independent. Then \(X \wedge Y \overset{d}{=} \text{Exp}(\lambda + \mu)\) is independent from \(W = 1_{\{X \wedge Y = X\}}\). In other words, the minimum of two independent exponential random variables is independent from which exponential r.v. occurred first.

Lemma 4.16. Fix \(a > 0\), and for each \(n \in \mathbb{N}\) let \(N_n\) be a Poisson process with intensity \(\alpha 2^n\). Then \(\{2^{-n}N_n(t) : t \in [0,T]\}\) converges in distribution to the line \(g(t) = at\), in the space \(D([0,T], \mathbb{R})\). In particular \(\{2^{-n}N_n(t) : t \in [0,T], n \in \mathbb{N}\}\) is \(C\)-tight.
Lemma 4.17. There is a filtered probability space \((\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) satisfying the usual conditions, supporting the \(\mathcal{F}_t\)-adapted process \((S_n, Z_n, V_n)\) given in Definition 4.2, also supporting the \(\mathcal{F}_t\)-adapted process \((U'_n, Z'_n, L'_n)\), such that

\[
U'_n = S_n + Z'_n, \\
M'_n = M^U_n, \\
L'_n(t) = 2^n \text{Leb}(0 < s < t : U'_n = -M'_n).
\]

Here \(Z'_n\) is a Poisson point process of intensity \(|v2^n|\) and jump size \(\text{sign}(v)2^{-n}\). Furthermore,

\[
Z'_n(t) - Z'_n(s) \leq Z_n(t) - Z_n(s),
\]

for all \(0 \leq s \leq t \leq T\), almost surely, and

\[
0 \leq L_n(t) - L_n(s) \leq L'_n(t) - L'_n(s),
\]

for all \(0 \leq s \leq t \leq T\), almost surely. The construction will yield independence between \(Z'_n\) and \(S_n\).

**Proof.** By definition \(V_n \geq -v\) almost surely. Recall that \(Z_n\) is a point process with stochastic intensity \(|2^nV_n|\) and jump size \(\text{sign}(V_n)2^{-n}\), so by Lemma 3.5 we assume the probability space included a process \(Z'_n\) with downward stochastic jump intensity \(|v2^n|\) and step size \(2^{-n}\) such that

\[
Z'_n(t) - Z'_n(s) \leq Z_n(t) - Z_n(s)
\]

for all \(0 \leq s \leq t \leq T\), almost surely. This inequality holds because the negative transitions of \(Z\) will have a rate less than \(|v|2^n\), which is the transition rate for negative jumps of \(Z'_n\). (And by definition \(Z'_n\) makes negative jumps only.) The jump times of \(Z'_n\) are independent of \(S_n\) hence the processes are independent. This demonstrates (15). It follows that

\[
U_n := S_n + Z_n \geq S_n + Z'_n =: U'_n,
\]

almost surely. Notice both processes \(U_n + M_n, U'_n + M'_n\) transition as a continuous time (nonnegative) random walk but with an additional “drift” process of \(Z_n, Z'_n\) respectively. For instance, a transition of \(U_n\) beginning from its running minimum corresponds to a transition from zero for the walk \(U_n + M_n\). By (15), the process \(U_n + M^U_n\) dominates that of \(U'_n + M'_n\). That is,

\[
U_n + M_n \geq U'_n + M'_n \geq 0, \text{ almost surely.}
\]

Hence, \(U'_n + M'_n\) is zero whenever \(U_n + M_n\) is zero. Consequently, \(\{s < z < t : U'_n(z) = -M'_n(z)\} \subset \{z : s < z < t, U_n(z) = -M_n(z)\}\) for every \((s, t) \subset [0, T]\), almost surely. Therefore,

\[
0 \leq L_n(t) - L_n(s) \leq L'_n(t) - L'_n(s),
\]

for every \((s, t) \subset [0, T]\), almost surely, demonstrating (16). \(\square\)
Lemma 4.18. For every $T > 0$, $\mathbb{E}(M_n(T)) \leq \mathbb{E}(M'_n(T)) \leq 2\sqrt{2T + T |v| \sqrt{2T} + 2|v|T}$ where $v$ is the initial value of $V_n$ and $M'_n$ is defined in Lemma 4.17.

**Proof.** According to Lemma 4.17 we assume our probability space supports $(S_n, Z_n, V_n)$ as well as the $Z'_n$ given in that lemma’s statement. Consequently,

$$U_n := S_n(t) + Z_n(t) \geq S_n(t) + Z'_n(t) =: U'_n,$$

for all $t \in [0, T]$, almost surely, implying

$$M_n(T) := M^{U_n}(T) \leq M^{U'_n}(T) =: M'_n(T), \quad \text{(19)}$$

almost surely. We can express the continuous time random walk $S_n$ as $2^{-n}(N_1(t) - N_2(t))$ where $N_i$ are independent Poisson processes of rate $2^{2n}$, and consequently $\mathbb{E}(S_n(T)^2) = \text{Var}(S_n(T)) = 2^{-2n}(\text{Var}(N_1(T)) + \text{Var}(N_2(T))) = 2T$. By Cauchy-Schwarz this yields $\mathbb{E}|S_n(T)| \leq \sqrt{2T}$. By Doob’s Martingale inequality, the fact that $Z'_n(T)$ is distributed as $-2^{-n}$ Poisson($|v2^nT|$), and independence of $Z'_n$ and $S_n$, we compute

$$\mathbb{E}(M_n(T)) \overset{(19)}{\leq} \mathbb{E}(M'_n(T)) \leq \sqrt{\mathbb{E}(M'_n(T)^2)}, \text{ Cauchy-Schwarz}$$

$$\leq \sqrt{4\mathbb{E}(U'_n(T)^2)} = 2\sqrt{\mathbb{E}(S_n(T) + Z'_n(T)^2)}, \text{ by Doob’s Maximal Inequality}$$

$$\leq 2\sqrt{\mathbb{E}(S_n^2(T)) + \mathbb{E}(S_n(T)Z'_n(T)) + \mathbb{E}(Z'_n(T)^2)} \leq 2\sqrt{2T + T |v| \sqrt{2T} + 2|v|T}.$$ 

\[ \square \]

Lemma 4.19. In the notation of Lemma 4.17, $L'_n$ converges in distribution to $M^{B - v}$ in the space $C([0, T], \mathbb{R})$ with the uniform norm. Here, $B^{-v}(t) := B(t) - vt$ where $B$ is a Brownian motion. In particular, $\{L'_n : n \in \mathbb{N}\}$ is $C$-tight. Furthermore, $\sup_n \mathbb{E}L_n(T) \leq \sup_n \mathbb{E}L'_n(T) < \infty.$

**Proof.** We begin by showing the weak convergence for which we use a similar technique in the proof of Lemma 4.8 (iii). We record the amount of time $U'_n$ spends at each level of the running minimum, and express $L'_n(t)$ as the sum of these times. By Lemma 4.16, $Z'_n(t) = -2^{-n}N_n(t)$ converges in distribution to $g(t) = -vt$ in the space $D([0, T], \mathbb{R})$. By Donsker’s theorem, $S_n$ converges in distribution to a Brownian motion $B$ and consequently $U'_n := S_n + Z'_n$ converges in distribution to $B + g :=: U'$. This implies $M'_n := M^{U'_n}$ converges to $M^{U'}$ in distribution because $f \mapsto M^f$ is continuous in the uniform norm and the limiting processes are continuous. See Remark 4.9. Note that $2^nM'_n(t) + 1$ is the number of levels the running minimum of $U'_n$ has reached by time $t$. Let $\tau^{(j)} := \inf \{t > 0 : U'_n(t) = -j2^{-n}\}$, so that $\tau^{(j)}$ is the first time $M'_n$ reaches $j2^{-n}$. Define

$$T_j := \text{Leb}(s \geq \tau^{(j)} \mid -U'_n(s) = M'_n(s) = j2^{-n}).$$
Then,
\[
2^n \sum_{j=0}^{2^n M'_n(s)} T_j \leq L'_n(s) \leq 2^n \sum_{j=0}^{2^n M'_n(s)+1} T_j, \tag{20}
\]
for all \( s \in [0, T] \), almost surely. When \( M'_n = j2^{-n} \), after the process \( U'_n \) arrives at \(-j2^{-n}\) for the \( k \)th time, it makes a positive jump upon the arrival of an \( \text{Exp}(2^n) \) random variable, call it \( \mu_k^{(j)} \), while it makes a negative jump upon the arrival of an \( \text{Exp}(2^n + |v2^n|) \) random variable \( \mu^{(j)'}_k \). Consider the pair \((\mu_k^{(j)} \wedge \mu^{(j)'}_k, I_j)\) where \( I_j = 1_{\{\mu_k^{(j)} \wedge \mu^{(j)'}_k = \mu_k^{(j)}\}} \).

By Lemma 4.15, \( I_j \) is independent from the i.i.d. sequence \( (\mu_k^{(j)} \wedge \mu^{(j)'}_k : k \geq 1) \). Then, \( W_j := \inf\{k : I_k = 0\} \) is the number of times \( U'_n \) visits \(-j2^{-n}\) while the signed running minimum \( M'_n \) is \( j2^{-n} \). Because \( I_j \overset{d}{=} \text{Bernoulli}(p) \) with
\[
p = \frac{q2^n}{2^{2n+1} + v2^n},
\]
and \( W_j \) is \text{Geometric}(p) (since it is the first time this sequence of Bernoulli random variables is zero), and Lemma 4.15 implies that \( W_j \) is independent of \( \mu_k^{(j)}, \mu^{(j)'}_i \). Thus,
\[
T_j = \sum_{i=1}^{W_j} \mu_i^{(j)} \wedge \mu_i^{(j)'}
\]
is a Geometric sum of i.i.d. exponential random variables of rate \( \lambda = 2^{2n+1} + v2^n \) that are independent of the number of the sums \( W_j \). Such a sum is exponential of rate \( p \lambda \). That is, \( T_j \overset{d}{=} \text{Exp}(p \lambda) = \text{Exp}(2^n) \). Each \( T_j \) is measurable with respect to \( \sigma \{ \{U'_n(s) \wedge M'_n(s) : s \in (\tau^{(j)}, \tau^{(j)+1})\} \}, \) and \( (T_j : j \geq 1) \) is independent of \( (U'_n(\tau^{(j)}), M'_n(\tau^{(j)})) \). (In other words, \( T_j \) depends only on the excursions between the stopping times and not the initial position.) Thus \( (T_j : j \geq 1) \) is a sequence of i.i.d. \( \text{Exp}(2^n) \) random variables. We show the left hand side of (20) converges in probability to \( M'(s) \) in the uniform norm for \( s \in [0, T] \). The proof for the right hand side is essentially identical. Without loss of generality, we may assume \( M'_n \) converges almost surely to \( M' \) by the Skorohod representation theorem and the fact shown above that \( M'_n \) converges to \( M' \) in distribution. Therefore
\[
\sup_{s \in [0, T]} \left| \sum_{j=0}^{2^n M'_n(s)} 2^{-n} - M'(s) \right| \rightarrow 0, \quad \text{almost surely.}
\]
To show the sequence \( 2^n \sum_{j=0}^{2^n M'_n(s)} T_j, n \geq 1 \), converges in probability to \( M'(s) \) uniformly for \( s \in [0, T] \), it suffices to show
\[
\sup_{s \in [0, T]} \left| \sum_{j=0}^{2^n M'_n(s)} T_j - \sum_{j=0}^{2^n M'_n(s)} 2^{-n} \right| \overset{p}{\rightarrow} 0.
\]
Approximating Brownian motion with inert drift by jump processes

Because \( T_j \overset{\text{d}}{=} \text{Exp}(2^{2n}) \), we know \( z_j := 2^n T_j \overset{\text{d}}{=} \text{Exp}(2^n) \). Then,
\[
2^n \sum_{j=0}^{2^n M_n'(s)} T_j - 2^n \sum_{j=0}^{2^n M_n'(s)} 2^{-n} = \sum_{j=0}^{2^n M_n'(s)} (z_j - 2^{-n}),
\]
where \( z_j - 2^{-n} \) are i.i.d. mean zero random variables with variance \( 2^{-2n} \). By Kolmogorov’s maximal inequality, for each \( C, \epsilon > 0 \)
\[
P \left( \sup_{s \in [0,T]} \left| \sum_{j=0}^{2^n M_n'(s)} (z_j - 2^{-n}) \right| > \epsilon \right) \leq P \left( \sup_{1 \leq k \leq 2^n C} \left| \sum_{j=0}^{k} (z_j - 2^{-n}) \right| > \epsilon \right) + P(M_n'(T) > C)
\]
\[
\leq \epsilon^{-2} \text{Var} \left( \sum_{j=0}^{2^n C} (z_j - 2^{-n}) \right) + P(M_n'(T) > C)
\]
\[
\leq \epsilon^{-2} 2^n C 2^{-2n} + P(M_n'(T) > C).
\]
Because \( M_n' \) converges to \( M' \) almost surely, this implies
\[
\limsup_{n \to \infty} P \left( \sup_{s \in [0,T]} \left| \sum_{j=0}^{2^n M_n'(s)} (z_j - 2^{-n}) \right| > \epsilon \right) \leq \limsup_{n \to \infty} P(M_n'(T) > C)
\]
\[
= P(M'(T) > C),
\]
where \( C > 0 \) is arbitrary. Since \( M'(T) \) is finite a.s., \( P(M'(T) > C) \) can be made arbitrarily small with a large choice for \( C \). Hence
\[
\sup_{s \in [0,T]} \left| \sum_{j=0}^{2^n M_n'(s)} (z_j - 2^{-n}) \right| \overset{P}{\longrightarrow} 0,
\]
and the left hand side of (20) converges in probability to \( M' \). That is,
\[
\sup_{s \in [0,T]} \left| \sum_{j=0}^{2^n M_n'(s)} T_j - M'(s) \right| \overset{P}{\longrightarrow} 0.
\]
(21)
The convergence in probability for the right hand side is similar,
\[
\sup_{s \in [0,T]} \left| \sum_{j=0}^{2^n M_n'(s)+1} T_j - M'(s) \right| \overset{P}{\longrightarrow} 0.
\]
(22)
Because \((20)\) is an almost sure bound,
\[
\mathbb{P}\left( \sup_{s \in [0,T]} |L_n'(s) - M'(s)| > \epsilon \right) \\
\leq \mathbb{P}\left( \sup_{s \in [0,T]} \left| 2^n M'(s) \right| > \epsilon/2 \right) + \mathbb{P}\left( \sup_{s \in [0,T]} \left| 2^n M'(s) + 1 \right| > \epsilon/2 \right),
\]
and taking \(\lim \sup_{n \to \infty}\) on both sides, \((21)\) and \((22)\) imply
\[
\lim \sup_{n \to \infty} \mathbb{P}\left( \sup_{s \in [0,T]} |L_n'(s) - M'(s)| > \epsilon \right) = 0,
\]
for every \(\epsilon > 0\). This complete the proof that \(L_n'\) converges in probability to \(M'\) in the uniform norm. Since \(M'\) is a continuous process, the sequence \(\{L_n : n \in \mathbb{N}\}\) is \(C\)-tight; see Definition 4.10.

To demonstrate the uniform moment bound, note that \(L_n(T) \leq L_n'(T)\), equation \((20)\), and Wald’s lemma imply
\[
\mathbb{E}(L_n(T)) \leq \mathbb{E}(L_n'(T)) = \mathbb{E}(M_n'(T)) + 2^{-n}.
\]
Applying the uniform moment bound on \(M_n'(T)\) given in Lemma 4.18, we see
\[
\sup_n \mathbb{E}(L_n(T)) \leq \sup_n \mathbb{E}(L_n'(T)) \leq \sup_n \mathbb{E}(M_n'(T)) + 1 < \infty.
\]

\[\square\]

**Corollary 4.20.** The collection of processes \(\{L_n : n \in \mathbb{N}\}\) is \(C\)-tight.

**Proof.** This follows directly from \((18)\), the fact that \(\{L_n' : n \in \mathbb{N}\}\) is \(C\)-tight by Lemma 4.19 (and that the zero process is trivially \(C\)-tight), and Lemma 4.12. \[\square\]

**Lemma 4.21.** The collection of processes \(\{Z_n : n \in \mathbb{N}\}\) is \(C\)-tight.

**Proof.** We will use a localization argument by stopping the stochastic intensity of \(Z_n\) when it becomes large. Recall that \(2^n V_n\) is the stochastic intensity of \(Z_n\). For \(C > v \geq 0\) set
\[
T_C^n = \inf\{t > 0 : V_n(t) > C\} = \inf\left\{ t > 0 : L_n(t) > \frac{v + C}{K} \right\}.
\]
Define a process \(\tilde{Z}_n\) such that \(\tilde{Z}_n|_{[0,T_C^n]} = Z_n|_{[0,T_C^n]}\), almost surely, while after time \(T_C^n\) let \(\tilde{Z}_n\) have positive jump intensity \(C2^n\) and jump size \(2^{-n}\) (and make only positive jumps). By Lemma 3.5 we can stochastically dominate the number of transitions made by \(\tilde{Z}_n\) in a given time interval by the number of transitions made by a point process \(Z_1^n(t)\) of intensity \(C2^n\) and jump size \(2^{-n}\) (in the same time interval). More precisely, we may
assume there exists a process $Z_n^1(t) = 2^{-n} N(C2^n t)$ on our probability space, where $N$ is Poisson process of unit intensity, and

$$0 \leq |Z_n(t) - Z_n(s)| \leq |Z_n^1(t) - Z_n^1(s)|$$  \hspace{1cm} (23)

for every interval $(s, t) \subset [0, T_C^{(n)})$, almost surely. By monotonicity of $Z_n^1$ and Doob’s maximal inequality, for fixed $\epsilon, \delta, C > 0$ we have

$$\mathbb{P}\left( \sup_{t \in [0, T - \delta]} \sup_{t < u, v < t + \delta} |Z_n(u) - Z_n(v)| > \epsilon \right)$$

$$\leq \mathbb{P}\left( \sup_{t \in [0, T - \delta]} \sup_{t < u, v < t + \delta} |\tilde{Z}_n(u) - \tilde{Z}_n(v)| > \epsilon, T_C^{(n)} > T \right) + \mathbb{P}(T_C^{(n)} < T)$$

$$\leq \mathbb{P}\left( \sup_{t \in [0, T - \delta]} |Z_n^1(u) - Z_n^1(v)| > \epsilon \right) + \mathbb{P}(T_C^{(n)} < T)$$

$$= \mathbb{P}\left( \sup_{t \in [0, T - \delta]} |Z_n^1(t + \delta) - Z_n^1(t)| > \epsilon \right) + \mathbb{P}(V_n(T) > C)$$

$$\leq \mathbb{P}\left( \sup_{t \in [0, T - \delta]} |Z_n^1(t + \delta) - Z_n^1(t)| > \epsilon \right) + \frac{|v| + \sup_n \mathbb{E}(L_n(T))}{C}.$$

By Lemma 4.16 and Lemma 4.12 we know

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\left( \sup_{t \in [0, T - \delta]} |Z_n^1(t + \delta) - Z_n^1(t)| > \epsilon \right) = 0.$$

By this and the uniform moment bound of $\sup_n \mathbb{E}(L_n(T))$ in Lemma 4.19, there exists a constant $A$ independent of $C, \epsilon, n$ such that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\left( \sup_{t \in [0, T - \delta]} \sup_{t < u, v < t + \delta} |Z_n(u) - Z_n(v)| > \epsilon \right) \leq \frac{A}{C}.$$

By choosing $C$ arbitrarily large, we see

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\left( \sup_{t \in [0, T - \delta]} \sup_{t < u, v < t + \delta} |Z_n(u) - Z_n(v)| > \epsilon \right) = 0.$$

Hence $Z_n$ satisfies (ii) of Lemma 4.12. Condition (i) follows from Remark 4.13 and the fact that $Z_n(0) = 0$ almost surely. This completes our verification of conditions (i)-(ii) of Lemma 4.12 sufficient for $C$-tightness of \{\{Z_n : n \in \mathbb{N}\}.

**Corollary 4.22.** The collection of processes \{(S_n, Z_n, V_n) : n \in \mathbb{N}\} is $C$-tight in $D([0, T], \mathbb{R}^3)$ with the Skorohod topology.

**Remark 4.23.** As topological spaces, $D([0, T], \mathbb{R}^d)$ with the Skorohod topology is not equivalent to $D([0, T], \mathbb{R})^d$ with the product topology. However, because the marginals are $C$-tight this a non-issue essentially because uniform convergence to a continuous function becomes the same in both spaces. See the comment in Jacod and Shiryaev [9, VI.1.21].
Note that $S_n$ is $C$-tight by Donsker’s theorem, while $C$-tightness of $V_n = v - KL_n$, and $Z_n$, follow from Corollary 4.20, and Lemma 4.21, respectively. By Skorohod’s representation theorem, every subsequence of $(S_n, Z_n, V_n)$ converging to a limiting process $(S, Z, V)$ can be assumed to converge almost surely in the product metric $d \times d \times d$, the product metric on $D([0, T], \mathbb{R}^3)$. By $C$-tightness of the marginals, $S, Z, V$ are all continuous, so that $(S, Z, V)$ is a continuous process. As in Remark 4.6, this implies the subsequence of $(S_n, Z_n, V_n)$ converges to $(S, Z, V)$ almost surely in the uniform norm, which implies almost sure convergence in $D([0, T], \mathbb{R}^3)$ under the Skorohod metric. Thus, $(S_n, Z_n, V_n)$ is $C$-tight as a collection of processes with paths in $D([0, T], \mathbb{R}^3)$.

4.5. Lemma 4.8: Characterization of subsequential limits

We prove items (i)-(iv) in Lemma 4.8 separately. The proof of (iii) was inspired by the proof of Lévy’s theorem given in [12, Chapter 6], where the authors essentially note the equivalence of the processes $(U_n + M_n, V_n)$ and $(X_n, V_n^X)$, which we described in subsection 4.1, for the case $K = 0$. Lévy’s theorem is the statement that $(L, |B|)$ and $(M^B, B + M^B)$ yield the same distribution on $C([0, T], \mathbb{R}^2)$. Here $L$ is the local time of $|B|$ at zero. See [10, Chapter 3.6] for a detailed statement.

**Proof of (i).** This follows trivially from the definition of $U_n = S_n + Z_n$.

**Proof of (ii).** Recall $S_n$ is a continuous time scaled simple random walk. Since $S_{nk}$ converges to $S$, $S$ is a Brownian motion by Donsker’s theorem.

We give a brief heuristic for the proof of (iii). Recall $L_n = 2^n \cdot \text{Leb}(0 < s < t : U_n(s) = -M_n(s))$ and $V_n = -v + KL_n$. Each time $M_n$ increases, $U_n$ will make approximately a Geometric(1/2) number of visits to this new minimum value before $M_n$ increases again. Also, $U_n$ will spend approximately an $\text{Exp}(2^n)$ amount of time at each one of these visits. Therefore, $L_n$, which is the total amount of time $U_n$ spends on $-M_n$ scaled by $2^n$, is approximately

$$\sum_{i=1}^{2^n M_n} 2^n \text{Exp}(2^n) = \sum_{i=1}^{2^n M_n} \text{Exp}(2^n),$$

where $\text{Exp}(2^n)$ indicates independent exponential random variables of rate $2^n$. If you suppose this sum is concentrated around its expectation conditional on $M_n$, then

$$L_n(t) \approx \sum_{i=1}^{2^n M_n(t)} 2^{-n} = M_n(t).$$

Furthermore, if $M_n$ converges almost surely to a process $M$ in the uniform norm, one expects $L_n$ to converge to $M$ as well.
Approximating Brownian motion with inert drift by jump processes

Proof of (iii). By tightness, and without loss of generality, assume $L_{n_k} \rightarrow L$ and $U_{n_k} \rightarrow U$ almost surely in the uniform norm of continuous functions. See Remark 4.6. We will use a localization argument by stopping $L_n$ after it reaches a large value. For a positive constant $C > |v|$, define

$$T_C^{(n_k)} = \inf\{ t > 0 : L_{n_k}(t) > C \}.$$ 

For each $n_k$, consider a modification of $(S_{n_k}, Z_{n_k}, U_{n_k})$, denoted $(S^C_{n_k}, Z^C_{n_k}, U^C_{n_k})$, solving the system (i)-(iii) in subsection 4.1 but replacing (iii) with

$$V^C_n(t) = -v + K[2^n \text{Leb}(0 < s < t : U^C_n(s) = -M^C_n(s)) \land C],$$

$$U^C_n = S^C_n + Z^C_n,$$

$$M^C_n(t) := M^U_n(t).$$

In other words we are stopping $L_n$ when it reaches $C$ while keeping the other dynamics of the system the same. Therefore $(S^C_{n_k}, Z^C_{n_k}, U^C_{n_k})$ is equal to $(S_{n_k}, Z_{n_k}, U_{n_k})$ on the interval $[0, T_C^{(n_k)}]$, a.s., while on $[T_C^{(n_k)}, \infty) \times$ the process $L^C_{n_k}$ is constantly $C$, and $U_{n_k}$ is the sum of a (scaled) continuous time random walk and an independent jump process of rate $-v + KC2^{n_k}$ and jump size $2^{-n}$. Fix $m \in 0^{-n_k} \mathbb{N}$. We bound the number of positive excursions of $U^C_{n_k}$ above $-m$ conditional on $-m$ being the running minimum. Let

$$\tau_{m,i+1} = \inf\{ t > \tau_{m,i} : U^C_{n_k}(t) = -m = -M^C_{n_k} \land T,$$

be the consecutive times $U^C_{n_k}$ visits $-m$ when $m$ is the current value of $M^C_{n_k}$, up until time $T$. (Set $\tau_{m,0} = 0$). That is, the consecutive times $U^C_{n_k}$ visits its running minimum at $-m$. Let

$$z^+_{m,j} = \inf\{ s > 0 : U^C_{n_k}(\tau_{m,j} + s) > U^C_{n_k}(\tau_{m,j}) \}$$

and

$$z^-_{m,j} = \inf\{ s > 0 : U^C_{n_k}(\tau_{m,j} + s) < U^C_{n_k}(\tau_{m,j}) \}$$

denote the amount of time until the next positive and negative, respectively, jump of $U^C_{n_k}$ after the $j^{th}$ excursion starting at $-m$. Then $z^+_{m,j} \land z^-_{m,j}$ is the time $U^C_{n_k}$ spends on $-m$ during its $j^{th}$ visit to its running minimum $-m$. Since $|L^C_{n_k}| \leq C$, the stochastic intensity of $Z^C_{n_k}$ is $2^{n_k} |V^C_n|$, which is bounded below by $2^{2n_k} - v$ and above by $2^{2n_k} + C2^{n_k}$. We set $v = 0$ in the remaining computations for convenience. Therefore the positive jump times of $U^C_{n_k} = S^C_{n_k} + Z^C_{n_k}$ have intensity $2^{2n_k} + 2^{n_k} |V^C_n|$ and the negative jump times arrive with intensity $2^{2n_k}$. In other words, $z^-_{m,j} \stackrel{d}{=} \text{Exp}(2^{2n_k})$.

By Lemma 3.1 we assume the probability space contains two independent sequences of i.i.d. exponential random variables $\{\nu_{i,j}\}_{i,j \in \mathbb{N}}, \{\epsilon_{i,j}\}_{i,j \in \mathbb{N}}$ that are also independent of $z^-_{m,j}$ and of the position $m$, with rates $2^{2n_k}$ and $C2^{n_k}$, respectively, and where

$$\nu_{m,j} \land \epsilon_{m,j} \leq z^+_{m,j} \leq \nu_{m,j},$$
That is, almost surely. Consequently,
\[ z_{m,j}^- \land \nu_{m,j} \land \epsilon_{m,j} \leq z_{m,j}^\land \leq z_{m,j}^- \land \nu_{m,j}, \text{ almost surely.} \] (24)

Define
\[ A_j^m := 1\{z_{m,j}^- \land \nu_{m,j} = \nu_{m,j}\}, \quad B_j^m := 1\{z_{m,j}^- \land z_{m,j}^\land = z_{m,j}^\land\}, \quad C_j^m := 1\{z_{m,j}^- \land \nu_{m,j} \land \epsilon_{m,j} = \nu_{m,j} \land \epsilon_{m,j}\}. \]

Note that \( B_j^m \) is the indicator for whether \( U_{n_k}^C \) jumped in the positive direction during its \( j^{th} \) visit to its running minimum \( -m \). By construction \( A_j^m, B_j^m, C_j^m \) are Bernoulli random variables and are coupled so that
\[ A_j^m \leq B_j^m \leq C_j^m, \text{ almost surely.} \] (25)

The definition of \( A_j^m, B_j^m, C_j^m \) depend on \( n_k \), which is hidden from notation. While the sequence \( (B_j^m : j \geq 1) \) is not an i.i.d. sequence, and is not a sequence of independent random variables since the jump rate changes with time, both \( (A_j^m : j \geq 1) \) and \( (C_j^m : j \geq 1) \) are i.i.d. sequences of Bernoulli(1/2), Bernoulli\([1 + \frac{C2^{-nk}}{2 + C2^{-nk}}]\), respectively. For each \( i \in \mathbb{N} \), denote \( Q_i \) as the number of visits to \(-i2^{-nk}\) by \( U_{n_k} \), while \( M_{n_k}^C = i2^{-nk} \).

This is the number of visits \( U_{n_k}^C \) makes to \(-i2^{-nk}\) when \(-i2^{-nk}\) is the running minimum. We will use (25) to sandwich \( Q_i \) above and below by geometric random variables. Denote \( m_i = i2^{-nk} \). Then
\[ Q_i = \inf\{j \geq 1 : B_j^m = 0\}. \]

That is, \( Q_i \) is the number of visits \( U_{n_k}^C \) makes to its running minimum \(-m_i\) because once a negative jump occurs, i.e. \( B_j^m \) reaches 0, the running minimum decreases to \(-(i + 1)2^{-nk}\). Consider
\[ W_i = \inf\{j \geq 1 : A_j^m = 0\}, \quad V_i = \inf\{j \geq 1 : C_j^m = 0\}. \]

Because \( (A_j^m), (C_j^m) \) are each i.i.d. sequences of Bernoulli random variables, \( W_i, V_i \) are geometric random variables, and
\[
\begin{align*}
W_i \leq Q_i \leq V_i, \text{ almost surely,} \\
P(W_i = k) &= \left(\frac{1}{2}\right)^k, \\
P(V_i = k) &= \left(\frac{1}{2 + C/2^{nk}}\right) \left(\frac{1 + C/2^{nk}}{2 + C/2^{nk}}\right)^{k-1}.
\end{align*}
\]

That is, \( W_i, V_i \) are geometrically distributed with parameters \( 1/2, (1 + C2^{-nk})/(2 + C2^{-nk}) \) respectively.

Now that we have sandwiched the number of steps \( U_{n_k}^C \) makes at a certain level of its running minimum, we will analyze the Lebesgue time the process spends at its running minimum. Since the size of each step is \( 2^{-nk} \), \( M_{n_k}^C \) has visited between \( 2^{nk}m - 1 \) and \( 2^{nk}m + 1 \) sites up until time \( \tau_{m,1} \).
Let $T_i$ be the time that $M^C_{n,k}$ spends at the site $m_i$, for $0 \leq i 2^{-n k} =: m_i < M^C_{n,k}(s)$ and a given $s \in [0, T]$. By definition of $\nu_{m,j}, e_{m,j}, z_{m,j}$ and the inequality (24), for $0 \leq i 2^{-n k} < M^C_{n,k}(s)$ we have

\[
\Phi^i_{nk} := \sum_{j=1}^{W_i} z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}} \leq T_i \leq \sum_{j=1}^{W_i} z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} + \sum_{j=W_i}^{V_i} z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}
\]

(26)

almost surely. This is because $U^C_n$ spends at least $W_i$ steps at the running minimum $-m_i$, each step spending at least $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}}$ Lebesgue amount of time for each $1 \leq j \leq W_i$, giving the lower bound. The upper bound is the same reasoning. By Lemma 4.15, $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}$ is an Exponential$(2^{n+1})$ random variable independent from $A_2^{n,k}$. Because $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}}$ is a measurable function of $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}$ and $e_{m_{i,j}}$, both of which are independent from $A_2^{n,k}$, $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}}$ is independent from $A_2^{n,k}$ as well. Consequently $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}}$, for $1 \leq j \leq W_i$, are independent from $V_i$. Similarly $V_i$ is independent of $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}$ for $1 \leq j \leq V_i$. Since by definition,

\[
L^C_{nk}(s) = \sum_{i=1}^{2^{n k} M^C_{n,k}(s)} 2^{n k} T_i,
\]

(26) implies we can sandwich $L^C_{nk}(s)$ by summing these upper bounds and lower bounds of times spent at each intermediate level. That is,

\[
\sum_{i=1}^{2^{n k} M^C_{n,k}(s)-1} 2^{n k} \Phi^i_{nk} \leq L^C_{nk}(s) \leq \sum_{i=1}^{2^{n k} M^C_{n,k}(s)+1} 2^{n k} \Phi^i_{nk} + \sum_{j=W_i}^{V_i} \sum_{i=1}^{2^{n k} M^C_{n,k}(s)+1} z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}
\]

(27)

and this inequality holds for all $s \in [0, T]$, almost surely.

Now we will apply the squeeze theorem to (27) and show the left hand and right hand of that inequality, and hence $L_{nk}$, converge to $M^U$ on $[0, T]$ in probability, hence for some subsequence of $n_k$ the convergence holds almost surely. The sum of a Geometric$(p)$ number of independent exponentials of rate $\lambda$ is exponential with rate $p \lambda$ provided the number of exponential random variables being summed is independent of the exponential random variables themselves. Therefore, since $W_i$ (resp. $V_i$) is geometric and independent of $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}} \land e_{m_{i,j}}$ for $1 \leq j \leq W_i$ (resp. $z_{m_{i,j}}^{-} \land \nu_{m_{i,j}}$ for $1 \leq j \leq V_i$), $\Phi^i_{nk}$ is distributed as an exponential of rate $(2^{2 n k} + C^{2 n k}) / 2$. Similarly $\Phi^i_{nk}$ has exponential rate $2^{2 n k}$. We
think of $2^{n_k}\tilde{\Phi}_{n_k}^i$ as an exponential random variable with rate approximately $2^{n_k}$, while in fact $2^{n_k}\Phi_{n_k}^i$ is an exponential with rate exactly $2^{2n_k}$. Because $\Phi_{n_k}^i$ is measurable with respect to $\mathcal{F}_{(\tau_{m_i+1},\tau_{m_i+1}]}$ and independent from $\mathcal{F}_{\tau_{m_i+1}}$, $(\tilde{\Phi}_{n_k}^i : 1 \leq i \leq W_i)$ is a collection of independent exponential random variables by the strong Markov property. In other words, $\tilde{\Phi}_{n_k}^i$ depends only on the excursion between these two hitting times and does not depend on the initial position of these excursions. Define

$$T_C := \inf\{t > 0 : L(t) > C\}$$

where $L = \lim_{n_k} L_{n_k}$. For any $0 < \epsilon < C$, it is clear that $\lim inf_{n_k} T_{C-\epsilon}^{(n_k)} \geq T_{C-\epsilon}$, and as a result

$$L_{n_k}^C = L_{n_k}, U_{n_k}^C = U_{n_k}$$

on $[0, T \wedge T_{C-\epsilon}]$ for large enough $n_k$, almost surely. Because $U_{n_k}^C(\cdot)$ converges uniformly on $[0, T]$ to the continuous process $U$, almost surely, we know $M_{n_k}^C(\cdot)$ converges uniformly on $[0, T \wedge T_{C-\epsilon}]$ to $M_U$. We will show that the left and right hand sides of (27) converge almost surely to $M_U(s \wedge T_{C-\epsilon})$ for each fixed $s \in [0, T]$. We go through the details for the left hand side, and the right hand is similar. For ease of notation we denote $s' = s \wedge T_{C-\epsilon}$.

Note

$$|2^{n_k}M_{n_k}^C(s') - 2^{n_k}M_U(s')| \leq 2^{n_k}[M_{n_k}^C(s') - M_U(s')] \leq 2^{n_k}\sum_{i=1}^{2^{n_k}[M_{n_k}^C(s') - M_U(s')]} e_i,$$

almost surely, where $e_i$ are i.i.d. Exp(1), and that are independent from $M_{n_k}^C$. The last inequality comes from Lemma 3.1 and the fact that $2^{n_k}\Phi_{n_k}^i \overset{d}{=} \exp(2^{n_k} + C/2)$ are i.i.d. Since $|M_{n_k}^C(s') - M_U(s')| \to 0$, almost surely, the strong law of large numbers implies that

$$\frac{1}{2^{n_k}} \sum_{i=1}^{2^{n_k}|M_{n_k}^C(s') - M_U(s')|} e_i \to 0,$$

almost surely.

We can express $2^{n_k}\tilde{\Phi}_{n_k}^i$ as $2^{-n_k}u_i^k$ where $u_i \overset{d}{=} \exp(1 + C2^{-(n_k+1)})$ are i.i.d., so

$$\sum_{i=1}^{2^{n_k}M_U(s')} 2^{n_k}\tilde{\Phi}_{n_k}^i = \frac{1}{2^{n_k}} \sum_{i=1}^{2^{n_k}M_U(s')} u_i^k.$$
We condition on $M^U(s')$ to compute
\[
\text{Var} \left( \frac{1}{2n_k} \sum_{i=1}^{2^n_k M^U(s')} u^k_i \bigg| M^U(s') \right) = \frac{1}{2^{2n_k} 2^{n_k} M^U(s')} \frac{1}{(1 + C^2 - (n_k + 1))^2} \\
= \frac{M^U(s')}{2^n_k + C + C^2 2^{-(n_k + 2)}},
\]
which approaches zero. The expectation conditional on $M^U(s')$ is
\[
E \left( \frac{1}{2n_k} \sum_{i=1}^{2^n_k M^U(s')} u^k_i \bigg| M^U(s') \right) = M^U(s') \frac{1}{1 + C^2 - (n_k + 1)}.
\]
Consequently,
\[
\sum_{i=1}^{2^n_k M^U(s')} 2^n_k \Phi^i_{nk} = \frac{1}{2^n_k} \sum_{i=1}^{2^n_k M^U(s')} u^k_i \longrightarrow M^U(s'),
\]
in probability, and by (28),
\[
\sum_{i=1}^{2^n_k M^U(s')} 2^n_k \Phi^i_{nk} \longrightarrow M^U(s'),
\]
in probability. Similarly for the right side of (27), one can show
\[
\sum_{i=1}^{2^n_k M^U(s')} 2^n_k \Phi^i_{nk} \longrightarrow M^U(s'),
\]
in probability. Now the term $R_{nk}(s')$ converges to $M^U(s')$ once we demonstrate
\[
\sum_{i=1}^{2^n_k M^U(s') + 1} V_i \sum_{j=W_i}^{z_{m_{i,j}}} \wedge \nu_{m_{i,j}}
\]
converges to zero in probability. But this follows from two applications of Wald’s lemma, the second of which uses the filtration generated (for fixed $i$) by $z_{m_{i,j}}, \nu_{m_{i,j}}$ and $e_{m_{i,j}}$ to compute
\[
E \left( \sum_{j=W_i}^{V_i} z_{m_{i,j}} \wedge \nu_{m_{i,j}} \right) = (E(V_i - W_i) + 1) E(z_{m_{i,j}} \wedge \nu_{m_{i,j}})
= \left( 2 - \frac{2 + C^2 - n_k}{1 + C^2 - n_k} \right) \frac{1}{2^{2n_k + 1} + C^2 n_k} =: a_{nk},
\]
which clearly approaches zero. Using Lemma 4.18, the moment bound hypothesis of Wald’s equation is satisfied since we have \( E(M_n^C(s')) \leq E(M_n^C(T)) \leq 2(2T+T^2)^{1/2} + 2|v(T)|^{1/2} \), a uniform bound with respect to \( n \). Thus, the second application of Wald’s lemma gives

\[
E \left( \sum_{i=1}^{2^n k M_n^C(s')} + |V_i| Z_{m,i,j} \right) = E(2^n k M_n^C(s') + 1) \cdot a_{n_k},
\]

which approaches zero as \( n_k \to \infty \). Consequently,

\[
\sum_{i=1}^{2^n k M_n^C(s')} + |V_i| Z_{m,i,j}
\]

does indeed converge to zero in probability, and therefore \( R_{n_k}(s') \) converges to \( M^U(s') \) in probability.

Because convergence in probability implies almost sure convergence for some subsequence, we can find a common subsequence \( n'_k \) where both (29) and (30) occur almost surely, for fixed \( s \), where \( s' = s \land T_{C-\epsilon} \). We relabel \( n'_k \) as \( n_k \). Similarly we can use a cantor diagonalization to find a further subsequence where (29) and (30) occur for all rationals in \([0, T \land T_{C-\epsilon}], \) almost surely. Applying the squeeze theorem to the inequality (27) then yields

\[
0 = \lim_{n_k \to \infty} |L_{n_k}^C(s') - M^U(s')| = \lim_{n_k \to \infty} |L_{n_k}(s') - M^U(s')| = |L(s') - M^U(s')| \quad (31)
\]

where \( s' = s \land T_{C-\epsilon} \), for any \( s \in \mathbb{Q} \cap [0, T] \), almost surely. Therefore \( L = \lim_{n_k} L_{n_k} = \lim_{n_k} L_{n_k}^C = M^{UC} = M^U \) for all rational numbers in \([0, T_{C-\epsilon} \land T] \subset [0, \lim inf n_k T_{n_k}^C \land T] \), almost surely. So \( L = M^U \) on \([0, T_{C-\epsilon} \land T] \), almost surely, since both processes are continuous.

Letting \( C \) approach infinity, \( P(T_C \geq T) \to 1 \), since \( L \) is a finite process, which yields \( M^U = L \) on \([0, T] \), almost surely, completing the proof of (iii).

**Proof of (iv).** As in the other proofs, Remark 4.6 and Lemma 4.7 allow us to assume without loss of generality that for the subsequence \((S_{n_m}, Z_{n_m}, V_{n_m}, L_{n_m}), \)

\[
(S_{n_m}, Z_{n_m}, V_{n_m}, L_{n_m}) \to (S, Z, V, L)
\]

almost surely, in the uniform norm on \( C([0, T], \mathbb{R}) \); we have \( U_{n_m} \to U \) on \( C([0, T], \mathbb{R}) \) as well. In the previous proof of (iii) we showed \( L(t) = M^U(t) \) for each \( t \in [0, T] \), almost surely. In this proof we wish to show

\[
Z(t) = \int_0^t V(x) \, dx, \text{ for } t \in [0, T], \quad (33)
\]
almost surely, where $V = KM^U - v$. We take $v \leq 0$ for the time being and reduce to this case at the end. It suffices to demonstrate that for each $s \in [0, T]$ there is a subsequence $n'_m$ such that

$$Z_{n'_m}(s) \rightarrow \int_0^s V(x) \, dx, \text{ almost surely.} \quad (34)$$

By a Cantor diagonalization $Z(\cdot)$ and $\int_0^s M^U(s) \, ds$ will agree for all rationals in $[0, T]$, almost surely. The two processes will then agree on $[0, T]$, almost surely, because both processes are continuous. For a given $n$,

$$\tilde{Z}_n(s) := 2^n Z_n(s)$$

counts the number of jumps of $Z_n$ by time $s$. Equivalently, this counts the number of arrival times $\{u_k : k \geq 1\}$ of jumps by the process $Z_n$. (We hide the dependence of $u_k$ on $n$ for convenience). For $C > 0$, let

$$\tau_C^{(n)} = \inf \{ s > 0 : V_n(s) > C \},$$

$$\bar{\alpha}_k = \sup_{s \in [u_k, u_{k+1}]} V_n(s),$$

$$\underline{\alpha}_k = \inf_{s \in [u_k, u_{k+1}]} V_n(s).$$

Assume for the time being that for every fixed $\delta > 0$,

$$\sup\{u_{i+1} - u_i : \tilde{Z}_{n_m}(\delta \wedge s) \leq i \leq \tilde{Z}_{n_m}(s)\} \rightarrow 0, \text{ in probability.} \quad (35)$$

Then there exists a subsequence $n'_m$, which we relabel as $n_m$, such that $\sup\{u_{i+1} - u_i : \tilde{Z}_{n_m}(\delta \wedge s) \leq i \leq \tilde{Z}_{n_m}(s)\} \rightarrow 0$, almost surely. We use the time between jumps, $u_{k+1} - u_k$, as the time step in a Riemann sum approximation of the integral in $(34)$. By the definition of $\bar{\alpha}_k, \underline{\alpha}_k$ and the exponential representation of the gap times given in $(8)$, there is a sequence $\mu_k$ of i.i.d. $\text{Exp}(2^n)$ random variables such that

$$\underline{\alpha}_k (u_{k+1} - u_k) \leq \mu_k \leq \bar{\alpha}_k (u_{k+1} - u_k),$$

almost surely. Therefore,

$$\sum_{k=\tilde{Z}_{n_m}(\delta \wedge s)}^{\tilde{Z}_{n_m}(s)} \underline{\alpha}_k (u_{k+1} - u_k) \leq \sum_{k=\tilde{Z}_{n_m}(\delta \wedge s)}^{\tilde{Z}_{n_m}(s)} \mu_k \leq \sum_{k=\tilde{Z}_{n_m}(\delta \wedge s)}^{\tilde{Z}_{n_m}(s)} \bar{\alpha}_k (u_{k+1} - u_k) \quad (36)$$

where we define the left and right sums to be zero should the set of such indices $\tilde{Z}_{n_m}(\delta \wedge s) \leq k \leq \tilde{Z}_{n_m}(s)$ be empty.
From (35) together with (32) and Riemann integrability of the limiting function $V$, we have
\[
\lim_{{n_m^* \to \infty}} \sum_{{k=Z_{{n_m}}(\delta \wedge s)}} Z_{{n_m}}(s) = \int_{\delta \wedge s}^{s} V(x) \, dx = \lim_{{n_m^* \to \infty}} \sum_{{k=Z_{{n_m}}(\delta \wedge s)}} Z_{{n_m}}(s),
\]
where convergence holds uniformly on $[0, T]$, almost surely. By the squeeze theorem,
\[
\sum_{{k=Z_{{n_m}}(\delta \wedge s)}} Z_{{n_m}}(s) \mu_k \to \int_{\delta \wedge s}^{s} V(x) \, dx,
\]
almost surely, as well. Since the $\mu_k$ are i.i.d. exponential r.v.’s of rate $2^{n_m}$ and $Z_{{n_m}}(s) = 2^{n_m} Z_{n_m}(\cdot)$ with $Z_{n_m}(\cdot) \to Z(\cdot)$ almost surely, the law of large numbers implies
\[
\sum_{{k=Z_{{n_m}}(\delta \wedge s)}} Z_{{n_m}}(s) \mu_k \xrightarrow{a.s.} Z(s) - Z(\delta \wedge s),
\]
for each $s \in [0, T]$. Therefore
\[
Z(s) - Z(\delta \wedge s) = \int_{\delta \wedge s}^{s} V(x) \, dx
\]
for each $s$ in $[0, T]$, almost surely. Since $Z(\delta) \to 0$ almost surely and $\int_{0}^{\delta} V(x) \, dx \to 0$, as $\delta \to 0$, this gives
\[
Z(s) = \int_{0}^{s} V(x) \, dx
\]
as desired.

To demonstrate (35), recall the jump process $Z_{U_k}$ determining the gap between jump times $u_{i+1} - u_i$ has an intensity process $2^{n_k} |V_{U_k}|$ that is bounded below by $\epsilon 2^{n_k}$ on the interval $[\tau_{U_k}, \infty)$. Heuristically, on this interval the intensity cannot be too small so the inter-arrival times are not too large. (This is where we use the fact that $\nu \leq 0$, so that $|V_{U_k}| = V_{U_k}$. In the case that $\nu > 0$ the intensity $2^{n_k} |V_{U_k}|$ will cross zero, which we handle at the end.) By Lemma 3.1 there exists an i.i.d. sequence $v_i$ of exponential random variables with rate $\epsilon 2^{n_k}$ that stochastically dominate $u_{i+1} - u_i$. We have
\[
\left\{\tau_{e^{(n_k)}} > \delta \right\} = \left\{V_{n_k}(\delta) \leq \epsilon \right\}.
\]
For $0 < \eta \ll 1, C > 0$,
\[
\mathbb{P}(\sup\{ (u_{i+1} - u_i) : \tilde{Z}_{nm}(\delta) \leq i \leq \tilde{Z}_{nm}(t) \} > \eta)
\leq \mathbb{P}(\sup\{ (u_{i+1} - u_i) : \tilde{Z}_{nm}(\tau^{(nm)}_\epsilon) \leq i \leq \tilde{Z}_{nm}(t) \} > \eta, \tau^{(nm)}_\epsilon \leq \delta) + \mathbb{P}(\tau^{(nm)}_\epsilon > \delta)
\leq \mathbb{P}(\sup\{ v_i : 1 \leq i \leq \tilde{Z}_{nm}(t) \} > \eta) + \mathbb{P}(\tau^{(nm)}_\epsilon > \delta)
\leq \mathbb{P}(\sup\{ v_i, 1 \leq i \leq C2^{nm} \} > \eta, \tilde{Z}_{nm}(t) \leq C2^{nm}) + \mathbb{P}(\tilde{Z}_{nm}(t) > C2^{nm}) + \mathbb{P}(\tau^{(nm)}_\epsilon > \delta)
\leq \mathbb{P}(v_i > \eta : \text{some } 1 \leq i \leq C2^{nm}) + \mathbb{P}(Z_{nm}(t) > C) + \mathbb{P}(\tau^{(nm)}_\epsilon > \delta)
\leq C2^{nm} \mathbb{P}(v_i > \eta) + \mathbb{P}(Z_{nm}(t) > C) + \mathbb{P}(\tau^{(nm)}_\epsilon > \delta),
\]
Taking lim sup with respect to $n_m$ on both sides and applying the assumption that $Z_n \to Z$ and $V_n \to V$ almost surely, we have
\[
\limsup_{n_m \to \infty} \mathbb{P}(\sup\{ (u_{i+1} - u_i) : \tilde{Z}_{nm}(\delta) \leq i \leq \tilde{Z}_{nm}(t) \} > \eta) \leq \mathbb{P}(Z(t) > C) + \mathbb{P}(V(\delta) \leq \epsilon).
\]
Since $C, \epsilon > 0$ are arbitrary and $V(0) \geq 0$ on our assumption $v \leq 0$,
\[
\limsup_{n_m \to \infty} \mathbb{P}(\sup\{ (u_{i+1} - u_i) : \tilde{Z}_{nm}(\delta) \leq i \leq \tilde{Z}_{nm}(t) \} > \eta) = 0,
\]
for every fixed $\delta > 0$, proving (35).

To show the case $v > 0$ reduces to $v = 0$, notice that
\[
T_{-3\epsilon/2} \leq \liminf_{n_m \to \infty} \tau^{(nm)}_{-\epsilon} \leq \limsup_{n_m \to \infty} \tau^{(nm)}_{\epsilon} \leq T_{3\epsilon/2},
\]
where
\[
T_a := \inf\{ t > 0 : V(t) > a \}.
\]
For almost each $\omega$ in our probability space there is an $N(\omega)$ such that $|V_{nm}|$ is monotone and bounded away from zero on the intervals $[0, \liminf_{n_m \to \infty} \tau^{(nm)}_{-\epsilon}] \supset [0, T_{-3\epsilon/2} \wedge T]$ and $[\limsup_{n_m \to \infty} \tau^{(nm)}_{\epsilon}, T] \supset [T_{3\epsilon/2} \wedge T, T]$, for all $n_m \geq N(\omega)$. With this fact and (41) we can apply the proof thus far to show
\[
Z(t) - Z(s) = \int_s^t V(x) \, dx \text{ for } s, t \in [0, T_{-3\epsilon/2} \wedge T], \text{ or } s, t \in [T_{3\epsilon/2} \wedge T, T].
\]
In addition to this, an $L_\infty$ bound gives
\[
\int_{T_{-3\epsilon/2} \wedge T}^{T_{3\epsilon/2} \wedge T} |V(x)| \, dx \leq (3\epsilon/2)T, \text{ almost surely,}
\]
which goes to zero as $\epsilon \to 0$. It follows that $Z(s) = \int_0^s V(x) \, dx$ for $s \in [0, T]$ in the case $v < 0$ as well.
5. Mutldimensional Analog of BMID

In 2007, White constructed a multidimensional analog, see [13], whose stationary distribution was found by Bass, Burdzy, Chen and Hairer [3]. This multidimensional analog is a pair of processes \((Z, V)\) where \(Z\) is a diffusion reflecting inside a sufficiently smooth domain \(D \subset \mathbb{R}^n\), and \(V\) is its drift. This drift is the inward normal integrated against the local time \(Z\) spends on \(\partial D\). That is,

\[
Z(t) = B(t) + \int_0^t \eta(Z(s)) \, dL(s) + \int_0^t V(s) \, ds,
\]

\[
V(t) = V_0 + \int_0^t \eta(Z(s)) \, dL(s),
\]

where \(\eta(x)\) is the inward unit normal for \(x \in \partial D\) and \(t \to L(t)\) is a nondecreasing continuous function flat off of \(\partial D\). By this we mean \(L\) increases only on \(Z^{-1}(\partial D)\). The authors show \((Z, V)\) has a stationary distribution of \(\mu \times \gamma\), where \(\mu\) is the uniform distribution on \(D\) and \(\gamma\) is the Gaussian distribution on \(\mathbb{R}^n\). This is interesting in part because the stationary distribution of the drift is always Gaussian and does not depend on \(D\), and also because the stationary distribution is always a product form. When \(Z\) is one dimensional, and \(D = [0, \infty)\), the process \(Z\) is one dimensional reflected BMID which is the process introduced by Knight.

Acknowledgements

CB is a Zuckerman Postdoctoral scholar at Technion-Israel’s Institute of Technology, Industrial Engineering and Management, Haifa, Israel, 32000. The preparation of this manuscript was partially supported by FNS 200021_175728/1. During this research the author was graduate student at the University of Washington and visited Universidad de Chile.

References


