The Annals of Applied Probability
2001, Vol. 11, No. 4, 1242–1262

EXACT CONVERGENCE RATES FOR THE DISTRIBUTION OF PARTICLES IN BRANCHING RANDOM WALKS

BY XIA CHEN

University of Tennessee

The exact convergence rates of the particle distributions in supercritical branching random walks and supercritical branching Wiener processes are obtained and a conjecture of Révész is confirmed.

1. Introduction. Consider a branching particle system starting from one ancestor at the origin in a d-dimensional space. Independently, each particle moves to a new site after one time unit after its birth, gives birth to a random number of offsprings and dies. The same procedure is repeated by all generations. Throughout, the migration is governed either by a d-dimensional simple symmetric random walk or by a d-dimensional Wiener process, and the reproduction by a Galton–Watson tree whose offspring distribution has the mean $m > 1$ and finite variance. This model is called *branching random walk* (when migration is executed by a random walk), or *branching Wiener process* (when migration is executed by a Wiener process). Under our assumptions, the random sequence $\{B(t)\}_{t\geq 0}$ with $B(t)$ being given as the total population in generation t ($t \geq 0$) is a supercritical branching chain. It is well known [see, cf., Athreya and Ney (1972)] that

$$\lim_{t\to\infty} \frac{B(t)}{m^t} = B \quad \text{a.s.} \tag{1.1}$$

for some random variable B which is not constantly zero.

In addition to their obvious background in the study of population growth and migration, the models of branching random walks (Wiener processes) had their origins in the theory of cascade processes. The study of branching random walks as a probability problem was initiated by Kolmogorov (1941). [The reader is referred to a survey by Ney (1991) for a historical account and for general information on this field.] A central limit theorem conjectured by Harris [(1963), page 75] states that

$$\frac{1}{m^T} \sum_{y \leq x\sqrt{T}} \lambda(y, T) \longrightarrow BG(x) \quad \text{a.s.,} \tag{1.2}$$

where $\lambda(x, T)$ is the population of the particles located at x at time T and $G(x)$ is the d-dimensional normal distribution function attracting the migration random walk through the classic central limit theorem. See, for example, Stam (1966), Asmussen and Kaplan (1976a, b), Athreya and Kaplan (1978),

Received August 2000; revised November 2000.
AMS 2000 *subject classifications*. 60F05, 60I15, 60F25.
Key words and phrases. Supercritical branching random walk, supercritical branching Wiener process, local limit theorem, central limit theorem.

Klebaner (1982), Joffe (1987), Biggins (1990), Bramson, Ney and Tao (1992) and Révész (1994) for the developments on this subject. Concerning the speed of above convergence, Révész (1994) proves that for each $\varepsilon > 0$,

$$T^{1/2-\varepsilon}\left(\frac{1}{m^T}\sum_{y\le x\sqrt{T}}\lambda(y,T)-BG(x)\right)\longrightarrow 0 \quad \text{a.s.} \tag{1.3}$$

Like the classic central limit theorem, the central limit theorem for branching random walks yields its local version [see also Watanabe (1965), Athreya and Kang (1998a, b) for the local central limit theorems for a variety of branching Markov processes]. In the case of branching random walk, Révész (1994) shows that

$$T^{1-\varepsilon}\left(\frac{1}{2}\left(\frac{4\pi T}{d}\right)^{d/2}\frac{\lambda(0,2T)}{m^{2T}}-B\right)\longrightarrow 0 \quad \text{a.s.} \tag{1.4}$$

Naturally, one wonders if (1.3) and (1.4) suggest exact rates of convergence. Indeed, a counterpart of (1.4) given in Theorem 4.9 of Révész (1994) says that for each $C > 0$, there is a $\delta = \delta(C) > 0$ such that

$$P\left\{\left|\frac{1}{2}\left(\frac{4\pi T}{d}\right)^{d/2}\frac{\lambda(0,2T)}{m^{2T}}-B\right|\ge\frac{C}{T}\right\}\ge\delta \tag{1.5}$$

for sufficiently large T. This observation makes him conjecture [Révész (1994), page 79] that the sequence

$$T\left(\frac{1}{2}\left(\frac{4\pi T}{d}\right)^{d/2}\frac{\lambda(0,2T)}{m^{2T}}-B\right),\qquad T=1,2,\dots \tag{1.6}$$

weakly converges to some nondegenerate random variable as $T\to\infty$.

This paper proposes to find the exact convergence rates for these limit theorems, and to settle the conjecture raised by Révész in particular. Instead of the weak convergence proposed by Révész, we shall prove his conjecture in terms of almost sure convergence as well as L_2-convergence. Our tools are some decompositions given in Révész (1994) and martingale approximations.

The rest of the paper is organized as follows: in Section 2, we give our results (Theorems 2.1 and 2.2 and Corollary 2.3) for branching random walks. In Section 3, we point out their analogues (Theorems 3.1 and 3.2) in the case of branching Wiener processes. Theorems 2.1 and 2.2 and Corollary 2.3 are proved in Section 4. Due to similarity, only a sketch is given for the proofs of Theorems 3.1 and 3.2 in Section 5.

The following notations and assumptions will be kept throughout the article. For $x=(x_1,\dots,x_d)$, $y=(x_1,\dots,y_d)\in\mathbf{R}^d$, $x\cdot y$ and $\|x\|$ will be used, respectively, for the inner product between x, y and for the Euclidean norm of x. The partial order "$x\le y$" is defined by the relation $x_1\le y_1,\dots,x_d\le y_d$.

Given a measurable $A \subset \mathbf{R}^d$, $|A|$ denotes its Lebesgue measure. Write

$$\Phi_d(x) = \left(\frac{1}{2\pi}\right)^{d/2} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \exp\left\{-\frac{\|y\|^2}{2}\right\} dy$$

and let $\Phi(x) = \Phi_1(x)$.

We use the nonnegative integer valued random variable Z to represent the distribution of the number of children of each individual in our particle system and assume

$$m \equiv EZ > 1 \quad \text{and} \quad \sigma^2 \equiv \operatorname{Var}(Z) < \infty. \tag{1.7}$$

2. Results for branching random walks. We begin with a formal definition of the local population $\lambda(x, t)$. Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ be the orthogonal unit vectors in the d-dimensional lattice $\mathbf{Z}^d$ and let X be a $\mathbf{Z}^d$-valued random variable independent of Z with

$$P\{X = \mathbf{e}_j\} = \frac{1}{2d}, \qquad j = 1, 2, \dots, d$$

and let

$$\{(X(x, t, k), Z(x, t, k));\ x \in \mathbf{Z}^d,\ t = 0, 1, 2, \dots, k = 1, 2, \dots\}$$

be an array of i.i.d. random vectors with

$$(X(\mathbf{0}, 0, 1), Z(\mathbf{0}, 0, 1)) = (X, Z).$$

Intuitively, we coordinate each individual in our particle system by the 3-tuple (x, t, k), where x represents the birth site, t represents the generation (so the original ancestor belongs to generation 0) and k is the order number as one of the members born at x in that generation. For a given individual (x, t, k), $X(x, t, k)$ is interpreted as the migration and $Z(x, t, k)$ is the number of the individual's children. The local population $\lambda(x, t)$ at $x \in \mathbf{Z}^d$ in the generation t is defined as follows:

$$\lambda(x, 0) = \begin{cases} 1, & \text{if } x = \mathbf{0}, \\ 0, & \text{if } x \neq \mathbf{0}, \end{cases}$$

$$\lambda(x, t) = \sum_{y \in \mathbf{Z}^d} \sum_{k=1}^{\lambda(y, t-1)} I_{x-y}(X(y, t-1, k)) Z(y, t-1, k),$$

where $x = (x_1, \dots, x_d) \in \mathbf{Z}^d$ and $t = 1, 2, \dots$. Clearly, $\lambda(x, t) = 0$ if $t \not\equiv x_1 + \cdots + x_d \bmod(2)$.

Write

$$B(t) = \sum_{x \in \mathbf{Z}^d} \lambda(x, t), \qquad t = 0, 1, 2, \dots .$$

Then $\{B(t)\}_{t \geq 0}$ is a supercritical Branching chain starting with $B(0) = 1$ and having the offspring distribution $\mathscr{L}(Z)$ [see, e.g., Athreya and Ney (1972) for details of branching chains]. It is well known that when $m > 1$, $\{B(t)\}_{t\geq 0}$ survives with positive probability.

Let

$$\mathscr{F}(t) = \mathscr{F}\{\lambda(x, s);\ x \in \mathbf{Z}^d,\ s = 0, 1, \dots, t\}$$

be the σ-algebra generated by the array

$$\{\lambda(x, s);\ x \in \mathbf{Z}^d,\ s = 0, 1, \dots, t\}.$$

THEOREM 2.1. *There exist a real random variable M and a $\mathbf{R}^d$-valued random variable N such that for each $x = (x_1, \dots, x_d) \in \mathbf{Z}^d$,*

$$T\left[\frac{1}{2}\left(\frac{2\pi T}{d}\right)^{d/2}\frac{\lambda(x, T)}{m^T} - B\exp\left\{-\frac{d\|x\|^2}{2T}\right\}\right] \longrightarrow d\left(\frac{1}{2}M + x\cdot N\right) \tag{2.1}$$

almost surely as well as in L_2-norm, as $T \to \infty$ with $T \equiv x_1 + \cdots + x_d \bmod(2)$, where the random variable B is given in (1.1).

In addition, the random variables M and N satisfy the following:

$$EM = 0 \quad \text{and} \quad EM^2 = \frac{4(m^2+\sigma^2)}{d(m-1)^3}, \tag{2.2}$$

$$EN = \mathbf{0} \quad \text{and} \quad \mathbf{cov}(N, N) = \frac{m^2+\sigma^2}{d(m-1)^2}\mathbf{I}_d, \tag{2.3}$$

$$(B, M, N) \overset{d}{=} (B, M, -N), \tag{2.4}$$

$$E[M|\mathscr{F}(t)] = t\frac{B(t)}{m^t} - \frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\|y\|^2\lambda(y, t), \qquad t = 0, 1, \dots, \tag{2.5}$$

$$E[N|\mathscr{F}(t)] = \frac{1}{m^t}\sum_{y\in\mathbf{Z}^d} y\lambda(y, t), \qquad t = 0, 1, \dots, \tag{2.6}$$

where $\mathbf{I}_d$ is the $d\times d$ identity matrix.

Further, if $\{(B_k, M_k, N_k)\}_{k\geq 1}$ are independent copies of (B, M, N) and if they are independent of (X, Z) then

$$(B, M, (N - BX)) \overset{d}{=} \frac{1}{m}\sum_{k=1}^{Z}(B_k, (M_k - 2X\cdot N_k), N_k). \tag{2.7}$$

THEOREM 2.2. *For each $x \in \mathbf{Z}^d$,*

$$\sqrt{T}\left[\frac{1}{m^T}\sum_{y\leq x\sqrt{T}}\lambda(y, T) - BP\{S_T \leq x\sqrt{T}\}\right] \longrightarrow -\nabla\Phi_d(\sqrt{d}x)\cdot N \tag{2.8}$$

almost surely as well as in L_2-norm, provided $T \to \infty$, where B is given in (1.1), *N is given in Theorem* 2.1 *and $\{S_t\}$ is the symmetric simple random walk generated by X.*

REMARK. Taking $x = 0$ in (2.1) we see that the sequence in (1.6) converges almost surely as well as in L_2-norm. So the conjecture made by Révész (1994) is proved. From (1.5) one can also see that the random variable M in Theorem 2.1 is unbounded. By Proposition 1.2.5 of Lawler (1991),

$$P^T(x) \equiv P\{S_T = x\} = 2\left(\frac{d}{2\pi T}\right)^{d/2} \exp\left\{-\frac{d\|x\|^2}{2T}\right\} + O\big(T^{-2-d/2}\big)$$

as $T \to \infty$ with $T \equiv x_1 + \cdots + x_d \bmod(2)$. Therefore (2.1) is equivalent to

$$T^{1+d/2}\left[\frac{\lambda(x,T)}{m^T} - BP^T(x)\right] \longrightarrow d\left(\frac{d}{2\pi}\right)^{d/2}\big(M + 2x\cdot N\big). \tag{2.9}$$

Nevertheless, $P\{S_T \le x\sqrt{T}\}$ in Theorem 2.2 can not be replaced by $\Phi_d(\sqrt{d}x)$. Indeed, we have the following corollary.

COROLLARY 2.3. *For each $x \in \mathbf{Z}^d$,*

$$\begin{aligned} &\sqrt{T}\left(\frac{1}{m^T}\sum_{y\le x\sqrt{T}} \lambda(y,T) - B\Phi_d(\sqrt{d}x)\right) \\ &\quad = \nabla\Phi_d(\sqrt{d}x)\cdot\big(B\mathbf{F}(\sqrt{T}x) - 2N\big) + O(1), \qquad T \to \infty \end{aligned} \tag{2.10}$$

almost surely as well as in L_2-norm, where B is given in (1.1), *N is given in Theorem* 2.1, *$\mathbf{F}(x) = \big(f(x_1), \dots, f(x_d)\big)$ and $f\colon (-\infty,\infty) \longrightarrow (-\frac{1}{2}, \frac{1}{2})$ is a periodic function with period* 1, *$f(k) = 0$ ($k = 0, \pm1, \pm2, \dots$) and*

$$f(\theta) = \frac{1-2\theta}{2}, \qquad 0 < \theta < 1.$$

Corollary 2.3 shows that asymptotically, the sequence

$$\sqrt{T}\left(\frac{1}{m^T}\sum_{y\le x\sqrt{T}} \lambda(y,T) - B\Phi_d(\sqrt{d}x)\right), \qquad T = 1, 2, \dots,$$

oscillates in a finite random interval. So the exact rate for the global central limit theorem is established.

3. Results for branching Wiener processes. The construction of the branching Wiener process is similar. Let $W(t)$ be a standard d-dimensional Wiener process independent of Z and write $W = W(1)$. Let

$$\big\{\big(W(x,t,k), Z(x,t,k)\big);\ x \in \mathbf{R}^d,\ t = 0, 1, 2, \dots, k = 1, 2, \dots\big\}$$

be a set of i.i.d. random vectors such that

$$\big(W(\mathbf{0},0,1), Z(\mathbf{0},0,1)\big) = (W, Z).$$

Define

$$\lambda(x,0)=\begin{cases}1, & \text{if } x=\mathbf{0},\\ 0, & \text{if } x\neq\mathbf{0},\end{cases}$$

$$\lambda(x,t)=\sum_{y\in\mathbf{R}^d}\sum_{k=1}^{\lambda(y,t-1)} I_x\big(y+W(y,t-1,k)\big)Z(y,t-1,k),$$

where $x\in\mathbf{R}^d$ and $t=1,2,\dots$. Clearly, $\lambda(x,t)=0$ for all but finitely many $x\in\mathbf{R}^d$. Define the random measure

$$\psi(A,t)=\sum_{x\in A}\lambda(x,t)=\sum_{y\in\mathbf{R}^d}\sum_{k=1}^{\lambda(y,t-1)} I_A\big(y+W(y,t-1,k)\big)Z(y,t-1,k)$$

for all measurable $A\subset\mathbf{R}^d$. Let

$$\begin{aligned}\mathscr{F}(t)&=\mathscr{F}\{\lambda(x,s);\ x\in\mathbf{R}^d, s=0,1,\dots,t\}\\ &=\mathscr{F}\{\psi(A,s);\ A\subset\mathbf{R}^d, s=0,1,\dots,t\}\end{aligned}$$

be the σ-algebra generated by

$$\{\lambda(x,s);\ x\in\mathbf{Z}^d,\ s=0,1,\dots,t\}.$$

THEOREM 3.1. *There exist a real random variable M and a $\mathbf{R}^d$-valued random variable N such that for each $A\subset\mathbf{R}^d$ with $|A|>0$ and $\int_A\|x\|dx<+\infty$,*

$$T\left[(2\pi T)^{d/2}\frac{\psi(A,T)}{m^T}-B\int_A\exp\left\{-\frac{\|x\|^2}{2T}\right\}dx\right]\longrightarrow|A|\left(\frac{1}{2}M+\bar{x}_A\cdot N\right)\tag{3.1}$$

almost surely as well as in L_2-norm, as $T\to\infty$, where B is given in (1.1) *and*

$$\bar{x}_A=\frac{1}{|A|}\int_A x\,dx.$$

In addition, the random variables M and N satisfy the following:

$$EM=0\quad\textit{and}\quad EM^2=\frac{2d(m^2+\sigma^2)(m+1)}{(m-1)^3},\tag{3.2}$$

$$EN=\mathbf{0}\quad\textit{and}\quad\mathbf{cov}(N,N)=\frac{(m^2+\sigma^2)}{(m-1)^2}\mathbf{I}_d,\tag{3.3}$$

$$(B,M,N)\stackrel{d}{=}(B,M,-N),\tag{3.4}$$

$$E\big[M|\mathscr{F}(t)\big]=dt\frac{B(t)}{m^t}-\frac{1}{m^t}\int\|y\|^2\psi(dy,t),\qquad t=0,1,\dots,\tag{3.5}$$

$$E\big[N|\mathscr{F}(t)\big]=\frac{1}{m^t}\int y\psi(dy,t),\qquad t=0,1,\dots.\tag{3.6}$$

Further, if $\{(B_k, M_k, N_k)\}$ *are independent copies of* (B, M, N) *and if they are independent of* (W, Z) *then*

$$
\begin{aligned}
&\big(B, \big(M - (d - \|W\|^2)B\big), (N - BW)\big) \\
&\quad \stackrel{d}{=} \frac{1}{m} \sum_{k=1}^{Z} \big(B_k, (M_k - 2W \cdot N_k), N_k\big).
\end{aligned}
\tag{3.7}
$$

THEOREM 3.2. *For each* $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$,

$$
\sqrt{T}\left[\frac{\psi\big(\{y : y \le x\sqrt{T}\}, T\big)}{m^T} - B\Phi_d(x)\right] \longrightarrow -\nabla\Phi_d(x) \cdot N \tag{3.8}
$$

almost surely as well as in L_2*-norm, as* $T \to \infty$, *where* B *is given in* (1.1) *and* N *is given in Theorem* 3.1.

4. Proof of Theorems 2.1 and 2.2 and Corollary 2.3.

PROOF OF THEOREM 2.1. To prove (2.1) we need only to verify (2.9). Define, for $0 \le t \le T$ and $x \in \mathbf{Z}^d$,

$$
f(x, T, t) = m^{T-t} \sum_{y \in \mathbf{Z}^d} \lambda(y, t) P^{T-t}(x - y).
$$

According to Lemma 4.3, page 67 in Révész (1994),

$$
E\big[\lambda(x, T) | \mathscr{F}(t)\big] = f(x, T, t), \qquad 0 \le t \le T, \;\; x \in \mathbf{Z}^d. \tag{4.1}
$$

We first follow the decomposition given in Révész (1994). Fix a number $\varepsilon > 0$ (which is sufficiently small to satisfy all the needs in the later argument) and choose $t \sim T^{\varepsilon}$. For all $x \in \mathbf{Z}^d$,

$$
\begin{aligned}
\frac{\lambda(x, T)}{m^T} - P^T(x)B = {} & \left(\frac{\lambda(x, T)}{m^T} - \frac{f(x, T, t)}{m^T}\right) \\
& + \left(\frac{f(x, T, t)}{m^T} - P^T(x)\frac{B(t)}{m^t}\right) \\
& + P^T(x)\left(\frac{B(t)}{m^t} - B\right).
\end{aligned}
\tag{4.2}
$$

In the proof of Theorem 2.1, we assume that $t \equiv x_1 + \cdots + x_d \bmod(2)$. As shown in Révész (1994), the first and the third terms are negligible since [Lemma 4.8, Révész (1994)]

$$
E\left(\sum_{y \in \mathbf{Z}^{\mathbf{d}}} \left(\frac{\lambda(y, T)}{m^T} - \frac{f(y, T, t)}{m^T}\right)^2\right) \le C \cdot \frac{1}{m^t (T - t)^{d/2}} \tag{4.3}
$$

for some constant $C > 0$, and since [Theorem 4.8, Révész (1994)]

$$
E\left(\frac{B(t)}{m^t} - B\right)^2 = O(m^{-t}). \tag{4.4}
$$

So we need to show that

$$T^{1+d/2}\left(\frac{f(x,T,t)}{m^T}-P^T(x)\frac{B(t)}{m^t}\right)\longrightarrow d\left(\frac{d}{2\pi}\right)^{d/2}(M+x\cdot N) \tag{4.5}$$

almost surely as well as in L_2-norm. Notice that

$$\frac{f(x,T,t)}{m^T}-P^T(x)\frac{B(t)}{m^t}=\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\lambda(y,t)\big(P^{T-t}(x-y)-P^T(x)\big). \tag{4.6}$$

By a formula given in Lawler [(1991), page 14],

$$\begin{aligned}P^{T-t}(x-y)&=2(2\pi)^{-d}\int_A e^{-i(x-y)\cdot\lambda}\phi^{T-t}(\lambda)\,d\lambda\\&=2(2\pi)^{-d}\int_A\cos\big((x-y)\cdot\lambda\big)\phi^{T-t}(\lambda)\,d\lambda,\\P^T(x)&=2(2\pi)^{-d}\int_A\cos(x\cdot\lambda)\phi^T(\lambda)\,d\lambda,\end{aligned}$$

where

$$\phi(\lambda)=\frac{1}{d}\sum_{j=1}^d\cos\lambda_j,\qquad \lambda=(\lambda_1,\dots,\lambda_d)\in\mathbf{R}^d$$

is the characteristic function of X and $A=[-\pi/2,\pi/2]\times[-\pi,\pi]^{d-1}$. By variable substitution,

$$\begin{aligned}&P^{T-t}(x-y)-P^T(x)\\&\quad=2(2\pi)^{-d}(T-t)^{-d/2}\int_{\sqrt{T-t}A}\left[\cos\left(\frac{(x-y)\cdot\lambda}{\sqrt{T-t}}\right)-\cos\left(\frac{x\cdot\lambda}{\sqrt{T-t}}\right)\phi\left(\frac{\lambda}{\sqrt{T-t}}\right)^t\right]\\&\qquad\times\phi\left(\frac{\lambda}{\sqrt{T-t}}\right)^{T-t}d\lambda\end{aligned}$$

and, by Taylor's expansion

$$\begin{aligned}&\cos\left(\frac{(x-y)\cdot\lambda}{\sqrt{T-t}}\right)-\cos\left(\frac{x\cdot\lambda}{\sqrt{T-t}}\right)\phi\left(\frac{\lambda}{\sqrt{T-t}}\right)^t\\&\qquad=\frac{1}{2(T-t)}\Big[(x\cdot\lambda)^2+td^{-1}\|\lambda\|^2-\big((x-y)\cdot\lambda\big)^2+o(1)\Big]\end{aligned}$$

uniformly for all $\|y\|\le t$ as $T\to\infty$. Notice that

$$\phi\left(\frac{\lambda}{\sqrt{T-t}}\right)^{T-t}\longrightarrow\exp\left\{-\frac{\|\lambda\|^2}{2d}\right\},\qquad T\to\infty$$

and that $|\phi(\lambda)|<1$ for all $\lambda\in A\setminus\{0\}$. Hence the dominated convergence theorem applies [see, e.g., the proof of Theorem 1.2.1 in Lawler (1991)], which, combined with the above observations, gives that

$$
\begin{aligned}
&P^{T-t}(x-y)-P^{T}(x)\\
&\quad=(2\pi)^{-d}T^{-1-d/2}\bigg[\int\Big(((x\cdot\lambda)^2+td^{-1}\|\lambda\|^2-\big((x-y)\cdot\lambda\big)^2\Big)\\
&\qquad\qquad\times\exp\bigg\{-\frac{\|\lambda\|^2}{2d}\bigg\}d\lambda+o(1)\bigg]\\
&\quad=(2\pi)^{-d/2}d^{1+d/2}T^{-1-d/2}\big[t-\|y\|^2+2x\cdot y+o(1)\big],\qquad T\to\infty
\end{aligned}
\tag{4.7}
$$

uniformly on $\|y\|\le t$. Since $\lambda(y,t)=0$ for all $\|y\|>t$, from (4.6) we have

$$
\begin{aligned}
&T^{1+d/2}\bigg(\frac{f(x,T,t)}{m^T}-P^T(x)\frac{B(t)}{m^t}\bigg)\\
&\quad=(2\pi)^{-d/2}d^{1+d/2}\bigg[t\frac{B(t)}{m^t}-\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\|y\|^2\lambda(y,t)\\
&\qquad\qquad+\frac{2}{m^t}x\cdot\sum_{y\in\mathbf{Z}^d}y\lambda(y,t)+o(1)\bigg]
\end{aligned}
\tag{4.8}
$$

almost surely and in L_2-norm as well. Let

$$
M_t=t\frac{B(t)}{m^t}-\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\|y\|^2\lambda(y,t)
$$

and

$$
N_t=\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}y\lambda(y,t),\qquad t=0,1,\dots.
$$

We claim that $\{M_t\}_{t\ge0}$ and $\{N_t\}_{t\ge0}$ are martingales w.r.t. the filtration $\{\mathscr{F}(t)\}_{t\ge0}$. Indeed,

$$
\begin{aligned}
E\big[M_t|\mathscr{F}(t-1)\big]&=E\bigg[t\frac{B(t)}{m^t}|\mathscr{F}(t-1)\bigg]\\
&\quad-\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\|y\|^2E\big[\lambda(y,t)|\mathscr{F}(t-1)\big]\\
&=t\frac{B(t-1)}{m^{t-1}}-\frac{1}{m^{t-1}}\sum_{y\in\mathbf{Z}^d}\|y\|^2\sum_{z\in\mathbf{Z}^d}\lambda(z,t-1)P(y-z)\\
&=t\frac{B(t-1)}{m^{t-1}}-\frac{1}{m^{t-1}}\sum_{z\in\mathbf{Z}^d}\lambda(z,t-1)E\|z+X\|^2\\
&=t\frac{B(t-1)}{m^{t-1}}-\frac{1}{m^{t-1}}\sum_{z\in\mathbf{Z}^d}\lambda(z,t-1)\big\{\|z\|^2+1\big\}=M_{t-1}.
\end{aligned}
\tag{4.9}
$$

The proof for $\{N_t\}_{t\ge0}$ being a martingale is similar.

To apply the martingale convergence theorem to $\{M_t\}_{t\geq 0}$ and $\{N_t\}_{t\geq 0}$, we compute their second moments. Note that

$$\begin{aligned}
M_t &= \frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\{t-\|y\|^2\}\lambda(y,t)\\
&= \frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\{t-\|y\|^2\}\sum_{z\in\mathbf{Z}^d}\sum_{k=1}^{\lambda(z,t-1)} I_{y-z}\big(X(z,t-1,k)\big)Z(z,t-1,k)\\
&= \frac{1}{m^t}\sum_{z\in\mathbf{Z}^d}\sum_{k=1}^{\lambda(z,t-1)} Z(z,t-1,k)\sum_{y\in\mathbf{Z}^d}\{t-\|y\|^2\}I_{y-z}\big(X(z,t-1,k)\big)\\
&= \frac{1}{m^t}\sum_{z\in\mathbf{Z}^d}\sum_{k=1}^{\lambda(z,t-1)} Z(z,t-1,k)\big\{t-\|z+X(z,t-1,k)\|^2\big\}.
\end{aligned}$$

Thus for each $t\geq 1$,

$$\begin{aligned}
M_t - M_{t-1} &= M_t - E\big[M_t|\mathscr{F}(t-1)\big]\\
&= \frac{1}{m^t}\sum_{z\in\mathbf{Z}^d}\sum_{k=1}^{\lambda(z,\,t-1)}\Big\{Z(z,t-1,k)\{t-\|z+X(z,t-1,k)\|^2\}\\
&\qquad -E\Big(Z(z,t-1,k)\big\{t-\|z+X(z,t-1,k)\|^2\big\}\Big)\Big\}.
\end{aligned}$$

Hence,

$$\begin{aligned}
E\big[M_t - M_{t-1}\big]^2 &= \frac{1}{m^{2t}}\sum_{z\in\mathbf{Z}^d}E\big(\lambda(z,t-1)\big)\cdot \mathrm{Var}\big\{Z\cdot\big[t-\|z+X\|^2\big]\big\}\\
&= \frac{1}{m^{t+1}}\sum_{z\in\mathbf{Z}^d}P^{t-1}(z)\mathrm{Var}\big\{Z\cdot\big[t-\|z+X\|^2\big]\big\}\\
&= \frac{1}{m^{t+1}}\cdot\mathrm{Var}\big\{Z\cdot\big[t-\|S_t\|^2\big]\big\} = \frac{1}{m^{t+1}}EZ^2\cdot E\big[t-\|S_t\|^2\big]^2\\
&= \frac{1}{m^{t+1}}(m^2+\sigma^2)\cdot 2(t^2-t)d^{-1}.
\end{aligned}$$

Therefore,

$$\sum_{t=1}^{\infty}E\big[M_t - M_{t-1}\big]^2 = \frac{4(m^2+\sigma^2)}{d(m-1)^3} < \infty. \tag{4.10}$$

Similarly,

$$N_t = \frac{1}{m^t} \sum_{z\in\mathbf{Z}^d} \sum_{k=1}^{\lambda(z,t-1)} Z(z, t-1, k)\big(z + X(z, t-1, k)\big),$$

$$N_t - N_{t-1} = \frac{1}{m^t} \sum_{z\in\mathbf{Z}^d} \sum_{k=1}^{\lambda(z,t-1)} \big\{Z(z, t-1, k)\big(z + X(z, t-1, k)\big) - E\big(Z(z, t-1, k)\big(z + X(z, t-1, k)\big)\big)\big\}.$$

Thus for each $t \ge 1$,

$$\begin{aligned}\mathbf{cov}(N_t - N_{t-1}, N_t - N_{t-1}) &= \frac{1}{m^{t+1}} \sum_{z\in\mathbf{Z}^d} \sum_{z\in\mathbf{Z}^d} P^{t-1}(z)\mathbf{cov}\big(Z(z+X), Z(z+X)\big)\\ &= \frac{1}{m^{t+1}}\mathbf{cov}\big(ZS_t, ZS_t\big) = \frac{1}{m^{t+1}}(m^2 + \sigma^2)d^{-1}t\mathbf{I}_d.\end{aligned}$$

Consequently,

$$\sum_{t=1}^{\infty} \mathbf{cov}(N_t - N_{t-1}, N_t - N_{t-1}) = \frac{(m^2 + \sigma^2)}{d(m-1)^2}\mathbf{I}_d. \tag{4.11}$$

By the martingale convergence theorem [see, e.g., page 2, Hall and Heyde (1980)], $\{M_t\}_{t\ge0}$ and $\{N_t\}_{t\ge0}$ converge almost surely, as well as in L_2-norm, to a real valued random variable M and a $\mathbf{R}^d$-valued random variable N, respectively. Since $M_0 = 0$ and $N_0 = \mathbf{0}$, we have

$$EM = 0 \quad \text{and} \quad EN = \mathbf{0}.$$

By above computation,

$$EM^2 = \frac{4(m^2 + \sigma^2)}{d(m-1)^3} \quad \text{and} \quad \mathbf{cov}(N, N) = \frac{(m^2 + \sigma^2)}{d(m-1)^2}\mathbf{I}_d.$$

In view of (4.2), (4.3), (4.4) and (4.8) we have (2.1)–(2.3), (2.5) and (2.6) in Theorem 2.1.

Replace $\{X(x, t, k)\}$ by $\{-X(x, t, k)\}$ and introduce the notations $\lambda'(x, t)$, M'_t, N'_t, M', N' for the replacements of $\lambda(x, t)$, M_t, N_t, M, N respectively, in our new particle system. By symmetry of migration we have

$$\big(B, M', N'\big) \overset{d}{=} \big(B, M, N\big). \tag{4.12}$$

On the other hand, $\lambda'(x, t) = \lambda(-x, t)$ for all $t \ge 0$ and $x \in \mathbf{Z}^d$. Hence we have $M'_t = M_t$ and $N'_t = -N_t$, which leads to $M' = M$ and $N' = -N$. Therefore, (2.4) follows from (4.12).

We now prove (2.7). Let $\lambda^*(x, T-1, k)$ be the population of the particles located at x at time T who are descended from the original ancestor's kth child; let

$$B_k(t) = \sum_{x \in \mathbf{Z}^d} \lambda^*(x, T-1, k)$$

and write

$$B_k = \lim_{t \to \infty} \frac{B_k(t)}{m^{t-1}}.$$

Then

$$\lambda(x, T) = \sum_{k=1}^{Z} \lambda^*(x, T-1, k) = \sum_z I_z(X) \sum_{k=1}^{Z} \lambda^*(x, T-1, k)$$

and, $\{B_k\}$ are i.i.d. random variables independent of (X, Z) and distributed as $\mathscr{L}(B)$. Notice that

$$\sum_z I_z(X) \sum_{k=1}^{Z} P^{T-1}(x-z) B_k = mB \sum_z I_z(X) P^{T-1}(x-z).$$

Similar to (4.7), for $\|z\| = 1$,

$$\begin{aligned} P^{T-1}(x-z) - P^T(x) &= d\left(\frac{d}{2\pi}\right)^{d/2} T^{-1-d/2}\big[1 - \|z\|^2 + 2x \cdot z + o(1)\big] \\ &= d\left(\frac{d}{2\pi}\right)^{d/2} T^{-1-d/2}\big(2x \cdot z + o(1)\big) \end{aligned}$$

as $T \to \infty$. Therefore,

$$\begin{aligned} &\sum_z I_z(X) \sum_{k=1}^{Z} \left[\frac{\lambda^*(x, T-1, k)}{m^{T-1}} - B_k P^{T-1}(x-z)\right] \\ &\quad = m\left(\frac{\lambda(x, T)}{m^T} - BP^T(x)\right) - mBd\left(\frac{d}{2\pi}\right)^{d/2} T^{-1-d/2} \sum_z I_z(X)\big(2x \cdot z + o(1)\big) \\ &\quad = m\left(\frac{\lambda(x, T)}{m^T} - BP^T(x)\right) - mBd\left(\frac{d}{2\pi}\right)^{d/2} T^{-1-d/2}\big(2x \cdot X + o(1)\big) \quad \text{a.s.} \end{aligned}$$

Applying (2.9) to the above relation we have

$$M + 2x \cdot (N - BX) = \frac{1}{m} \sum_{k=1}^{Z} \big\{(M_k - 2X \cdot N_k) + 2x \cdot N_k\big\} \quad \text{a.s.,}$$

where $\{(M_k, N_k)\}$ are independent copies of (M, N) and $\{(M_k, N_k)\}$ are independent of (X, Z). Since $x \in \mathbf{Z}^d$ is arbitrary and since

$$B = \frac{1}{m}\sum_{k=1}^{Z} B_k,$$

we have (2.7). □

PROOF OF THEOREM 2.2. Consider the following decomposition:

$$\begin{aligned}
&\frac{1}{m^T}\sum_{y\le x\sqrt{T}}\lambda(y,T)-BP\{S_T\le x\sqrt{T}\}\\
&\quad=\frac{1}{m^T}\sum_{y\le x\sqrt{T}}\big(\lambda(y,T)-f(y,T,t)\big)+\sum_{y\le x\sqrt{T}}\frac{f(y,T,t)}{m^T}\\
&\qquad-P\{S_T\le x\sqrt{T}\}\frac{B(t)}{m^t}+P\{S_T\le x\sqrt{T}\}\bigg(\frac{B(t)}{m^t}-B\bigg).
\end{aligned}$$

Again, we let $t \sim T^{\varepsilon}$ for some sufficiently small $\varepsilon > 0$. In view of (4.4), the third term in the above decomposition is negligible. Since $\lambda(y, T) = 0$ for all $\|y\| > T$,

$$\bigg(\sum_{y\le x\sqrt{T}}\big(\lambda(y,T)-f(y,T,t)\big)\bigg)^2\le KT^{d/2}\sum_{y\in\mathbf{Z}^d}\big(\lambda(y,T)-f(y,T,t)\big)^2,$$

where $K > 0$ is a constant. By (4.3), the first term is also negligible.

Therefore, we only need to deal with the second term, that is, to prove that

$$\sqrt{T}\bigg[\sum_{y\le x\sqrt{T}}\frac{f(y,T,t)}{m^T}-P\{S_T\le x\sqrt{T}\}\frac{B(t)}{m^t}\bigg]\longrightarrow -\nabla\Phi_d(\sqrt{d}x)\cdot N \tag{4.13}$$

almost surely as well as in L_2-norm.

Notice that

$$\begin{aligned}
&\frac{1}{m^T}\sum_{y\le x\sqrt{T}}f(y,T,t)-P\{S_T\le x\sqrt{T}\}\frac{B(t)}{m^t}\\
&\quad=\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d}\Big[P\{S_{T-t}\le x\sqrt{T}-y\}-P\{S_T\le x\sqrt{T}\}\Big]\lambda(y,t).
\end{aligned}\tag{4.14}$$

Write

$$\begin{aligned}
&P\{S_{T-t}\le x\sqrt{T}-y\}-P\{S_T\le x\sqrt{T}\}\\
&\quad=\Big(P\{S_T\le x\sqrt{T}-y\}-P\{S_T\le x\sqrt{T}\}\Big)\\
&\qquad+\Big(P\{S_{T-t}\le x\sqrt{T}-y\}-P\{S_T\le x\sqrt{T}-y\}\Big)\\
&\quad=(I)+(II)\quad\text{(say)}.
\end{aligned}\tag{4.15}$$

Uniformly on $\|y\| \le t$,

$$(I) = -\big(1+o(1)\big)\sum_{j=1}^{d}\operatorname{sgn}(y_j)P\Big\{x_j\sqrt{T}-|y_j| < S_T^{(j)} \le x_j\sqrt{T},\ S_T^{(k)} \le x_k\sqrt{T}, k \ne j, 1 \le k \le d\Big\} \tag{4.16}$$

as $T \to \infty$, where $S_T^{(1)}, \dots, S_T^{(d)}$ are the components of S_T. We claim that for each $1 \le j \le d$, uniformly on $\|y\| \le t$,

$$P\big\{x_j\sqrt{T}-|y_j| < S_T^{(j)} \le x_j\sqrt{T}, S_T^{(k)} \le x_k\sqrt{T}, k \ne j, 1 \le k \le d\big\}$$
$$= \begin{cases} \big(1+o(1)\big)T^{-1/2}\dfrac{d}{dx}\big(\Phi(x)\big) \\ \quad\times\big[|y| + \big(1-(-1)^{|y|}\big)(-1)^{T+[x\sqrt{T}]}\big], & \text{as } d = 1, \\ \big(1+o(1)\big)T^{-1/2}|y_j|\dfrac{\partial}{\partial x_j}\big(\Phi_d(\sqrt{d}x)\big), & \text{as } d \ge 2. \end{cases} \tag{4.17}$$

We only prove (4.16) in the case $d \ge 2$, as the proof for the case $d = 1$ is similar but much simpler. Without loss of generality, we only consider the case $j = 1$. By a combinatorial argument we can see that for any measurable $A_1, \dots, A_d \subset \mathbf{R}$,

$$P\big\{S_T^{(1)} \in A_1, \dots, S_T^{(d)} \in A_d\big\} = \frac{1}{d^T}\sum_{k_1+\cdots+k_d=T}\frac{T!}{k_1!\cdots k_d!}\prod_{j=1}^{d}P\big\{\overline{S}_{k_j} \in A_j\big\}, \tag{4.18}$$

where $\overline{S}_t$ is a one-dimensional symmetric random walk. In particular,

$$P\Big\{x_1\sqrt{T}-|y_1| < S_T^{(1)} \le x_1\sqrt{T}, S_T^{(k)} \le x_k\sqrt{T}, 2 \le k \le d\Big\}$$
$$= \frac{1}{d^T}\sum_{k_1+\cdots+k_d=T}\frac{T!}{k_1!\cdots k_d!}\Bigg(\sum_{x_1\sqrt{T}-|y_1|<z\le x_1\sqrt{T}}P\big\{\overline{S}_{k_1} = z\big\}\Bigg)$$
$$\times\prod_{j=2}^{d}P\big\{\overline{S}_{k_j} \le x_j\sqrt{T}\big\}.$$

Notice that $P\big\{\overline{S}_{k_1} = z\big\} = 0$ when $k_1 \not\equiv z \bmod(2)$, and

$$P\big\{\overline{S}_{k_1} = z\big\} \sim 2T^{-1/2}\sqrt{\frac{d}{2\pi}}\exp\bigg\{-\frac{dx_1^2}{2}\bigg\},$$

when $k_1 \equiv z \bmod(2)$ and $k_1 \sim d^{-1}T$ as $T \to \infty$. On the other hand,

$$\#\big\{z \in \mathbf{Z} : x_1\sqrt{T}-|y_1| < z \le x_1\sqrt{T} \text{ and } z \equiv k_1 \bmod(2)\big\}$$
$$= \tfrac{1}{2}\big[|y_1| + \big(1-(-1)^{|y_1|}\big)(-1)^{k_1+[x_1\sqrt{T}]}\big].$$

Consider Cramér's large deviation [cf. Theorem 2.2.30 of Dembo and Zeitouni (1992)] which gives in our case

$$\frac{1}{d^T}\sum_{(k_1,\dots,k_d)\in A_T(\delta)}\frac{T!}{k_1!\cdots k_d!}=O(e^{-\alpha T}),\qquad T\to\infty \tag{4.19}$$

for any $\delta>0$, where $\alpha=\alpha(\delta)>0$ is a constant and

$$\begin{aligned}A_T(\delta)=\big\{(k_1,\dots,k_d)\in\mathbf{Z}^d;\,k_1+\cdots+k_d=T,\,k_j\ge 0\ (j=1,\dots,d)\\ \text{and } |k_j-d^{-1}T|\ge\delta T \text{ for some } 1\le j\le d\big\}.\end{aligned}$$

[We only need that the left-hand side of (4.19) tend to zero here, which is referred to as the law of large numbers. We give (4.19) for the needs of the later development]. Therefore,

$$\begin{aligned}
&P\Big\{x_1\sqrt{T}-|y_1|<S_T^{(1)}\le x_1\sqrt{T}+|y_1|,\,S_T^{(k)}\le x_k\sqrt{T},\,2\le k\le d\Big\}\\
&\quad=\big(1+o(1)\big)\frac{1}{d^T}\sum_{k_1+\cdots+k_d=T}\frac{T!}{k_1!\cdots k_d!}\Big[|y_1|+\big(1-(-1)^{|y_1|}\big)(-1)^{k_1+[x_k\sqrt{T}]}\Big]\\
&\qquad\times T^{-1/2}\sqrt{\frac{d}{2\pi}}\exp\Big\{-\frac{dx_1^2}{2}\Big\}\prod_{j=2}^{d}\Phi(\sqrt{d}x_j)\\
&\quad=\big(1+o(1)\big)T^{-1/2}\frac{\partial}{\partial x_1}(\Phi_d(\sqrt{d}x))\\
&\qquad\times\bigg[|y_1|+\big(1-(-1)^{|y_1|}\big)(-1)^{[x_1\sqrt{T}]}\bigg(1-\frac{2}{d}\bigg)^T\bigg]\\
&\quad=\big(1+o(1)\big)T^{-1/2}\frac{\partial}{\partial x_1}\Big(\Phi_d(\sqrt{d}x)\Big)|y_1|,\qquad T\to\infty.
\end{aligned}$$

By (4.16) and (4.17),

$$(I)=\begin{cases}-T^{-1/2}\dfrac{d}{dx}(\Phi(x))\\ \quad\big[y+(1-(-1)^{|y|})\operatorname{sgn}(y)(-1)^{T+[x\sqrt{T}]}+o(1)\big], & \text{as } d=1,\\ -T^{-1/2}\big[\nabla\Phi_d(\sqrt{d}x)\cdot y+o(1)\big], & \text{as } d\ge 2\end{cases} \tag{4.20}$$

holds uniformly on $\|y\|\le t$ as $T\to\infty$.

Note that when $d\ge 2$,

$$\begin{aligned}
(II)&=\sum_{z\in\mathbf{Z}^d}P^t(z)\big[P\{S_{T-t}\le x\sqrt{T}-y\}-P\{S_{T-t}\le x\sqrt{T}-y-z\}\big]\\
&=-\sum_{z\in\mathbf{Z}^d}P^t(z)T^{-1/2}\big[\nabla\Phi_d(\sqrt{d}x)\cdot z+o(1)\big] \qquad (4.21)\\
&=-T^{-1/2}\big[E(\nabla\Phi_d(\sqrt{d}x)\cdot S_t)+o(1)\big]=o(T^{-1/2}),\qquad T\to\infty
\end{aligned}$$

holds uniformly on $\|y\|\le t$, where the second equality follows from an obvious modification of (4.20). Similarly, (4.21) is also true when $d=1$.

Combining (4.14), (4.15), (4.20), (4.21) we can see that as $d \geq 2$,

$$\begin{aligned}\frac{1}{m^T}\sum_{y\leq x\sqrt{T}} f(y,T,t) - P\{S_T \leq x\sqrt{T}\}\frac{B(t)}{m^t}\\ = -T^{-1/2}\left[\nabla\Phi_d(\sqrt{d}x)\cdot\left(\frac{1}{m^t}\sum_{y\in\mathbf{Z}^d} y\lambda(y,t)\right)+o(1)\right]\end{aligned}$$

almost surely as well as in L_2-norm. This also holds in the case $d = 1$ as

$$\begin{aligned}&\left|\frac{1}{m^T}\sum_{y\in\mathbf{Z}}\bigl(1-(-1)^{|y|}\bigr)\operatorname{sgn}(y)\lambda(y,t)\right|\\ &\quad = \left|\sum_{y\in\mathbf{Z}}\bigl(1-(-1)^{|y|}\bigr)\operatorname{sgn}(y)\left[\frac{\lambda(y,t)}{m^t}-P^t(y)B\right]\right|\\ &\quad \leq \sum_{y\in\mathbf{Z}}\left|\frac{\lambda(y,t)}{m^t}-P^t(y)B\right| \longrightarrow 0, \qquad T\to\infty,\end{aligned}$$

where the last step follows from Theorem 4.3 and Theorem 4.5 in Révész (1994).

So (4.13) follows from (2.6) and martingale convergence theorem. □

PROOF OF COROLLARY 2.3. By Theorem 2.2, it is enough to prove the following multivariate version of the Edgeworth expansion:

$$P\{S_T \leq x\sqrt{T}\} = \Phi_d(\sqrt{d}x) + T^{-1/2}\nabla\Phi_d(\sqrt{d}x)\cdot\mathbf{F}(\sqrt{T}x) + o\bigl(T^{-1/2}\bigr). \tag{4.22}$$

When $d = 1$, by Theorem 6, page 171 of Petrov (1975) we have

$$\begin{aligned}P\{S_T \leq x\sqrt{T}\} &= \Phi(x) + \frac{1}{\sqrt{T}}\Phi'(x)\sum_{l=1}^{\infty}\frac{\sin\bigl(2l\pi\sqrt{T}x\bigr)}{\pi l} + o\bigl(T^{-1/2}\bigr)\\ &= \Phi(x) + T^{-1/2}\Phi'(x)f(\sqrt{T}x) + o\bigl(T^{-1/2}\bigr),\end{aligned} \tag{4.23}$$

where the last step follows from the Fourier expansion.

Consider the case when $d > 1$. From (4.18),

$$P\bigl\{S_T \leq x\sqrt{T}\bigr\} = \frac{1}{d^T}\sum_{k_1+\cdots+k_d=T}\frac{T!}{k_1!\cdots k_d!}\prod_{j=1}^{d}P\bigl\{\overline{S}_{k_j} \leq x_j\sqrt{T}\bigr\}. \tag{4.24}$$

By a treatment similar to the one used in the proof of (4.17), the expansion (4.22) follows from (4.19), (4.23) and (4.24). □

5. Proofs of Theorems 3.1 and 3.2. Let $P_t(x, A)$ be the transition probability of $\{W(t)\}$. Then

$$P_t(x, A) = \left(\frac{1}{2\pi t}\right)^{d/2} \int_A \exp\left\{-\frac{\|y-x\|^2}{2t}\right\} dy, \qquad t > 0.$$

Define, for each $t \le T$,

$$F(A, T, t) = m^{T-t} \sum_{x \in \mathbf{R}^d} P_{T-t}(x, A)\lambda(x, t) = m^{T-t} \int P_{T-t}(x, A)\psi(dx, t).$$

By Lemma 6.3 in Révész (1994), for each $0 \le t < T$,

$$E\big[\psi(A, T)|\mathscr{F}(t)\big] = F(A, T, t). \tag{5.1}$$

The proof here is similar to the one given in Section 4, except that the range of branching Wiener process is no longer bounded in probability. We shall overcome such a difficulty by truncation. From this motive we need the following lemma.

Lemma 5.1. *There exists a constant $C > 0$ depending only on m and σ^2 such that*

$$E\Big[\int f(x)\Psi(dx, t)\Big]^2 \le Cm^{2t}Ef^2\big(W(t)\big) \tag{5.2}$$

for every function f on $\mathbf{R}^d$ and t.

Proof. By (5.1),

$$E\Big[\int f(x)\Psi(dx, t)\Big] = \int f(x)F(dx, t, 0) = m^t Ef\big(W(t)\big). \tag{5.3}$$

Notice that

$$\int f(x)\Psi(dx, t) - E\Big[\int f(x)\Psi(dx, t)\Big]$$
$$= \sum_{s=1}^{t}\Big[\int f(x)F(dx, t, s) - \int f(x)F(dx, t, s-1)\Big],$$

$$\int f(x)F(dx, t, s)$$
$$= m^{t-s} \sum_{x \in \mathbf{R}^d} P_{t-s}f(x)\lambda(x, t)$$
$$= m^{t-s} \sum_{x \in \mathbf{R}^d} P_{t-s}f(x) \sum_{y \in \mathbf{R}^d} \sum_{k=1}^{\lambda(y,t-1)} I_x\big(y + W(y, t-1, k)\big)Z(y, t-1, k)$$
$$= m^{t-s} \sum_{y \in \mathbf{R}^d} \sum_{k=1}^{\lambda(y,s-1)} P_{t-s}f\big(y + W(y, s-1, k)\big)Z(y, s-1, k).$$

In view of (5.1) we have

$$\begin{aligned}
&\int f(x)F(dx,t,s) - \int f(x)F(dx,t,s-1)\\
&\quad = m^{t-s}\sum_{y\in\mathbf{R}^d}\sum_{k=1}^{\lambda(y,s-1)}\Big[P_{t-s}f\big(y+W(y,s-1,k)\big)Z(y,s-1,k)\\
&\qquad\qquad - E\big(P_{t-s}f\big(y+W(y,s-1,k)\big)Z(y,s-1,k)\big)\Big].
\end{aligned}$$

Hence,

$$\begin{aligned}
&E\Big[\int f(x)F(dx,t,s) - \int f(x)F(dx,t,s-1)\Big]^2\\
&\quad = m^{2t-2s}\sum_{y\in\mathbf{R}^d} E\big(\lambda(y,s-1)\big)\mathrm{Var}\big(P_{t-s}f(y+W)Z\big)\\
&\quad = m^{2t-s-1}\int P_{s-1}(0,dy)\,\mathrm{Var}\big(P_{t-s}f(y+W,A)Z\big)\\
&\quad = m^{2t-s-1}\,\mathrm{Var}\big(P_{t-s}f\big(W(s)\big)Z\big)\\
&\quad \le m^{2t-s-1}(m^2+\sigma^2)E\big(P_{t-s}f\big(W(s)\big)\big)^2\\
&\quad \le m^{2t-s-1}(m^2+\sigma^2)E\big(P_{t-s}f^2\big(W(s)\big)\big)\\
&\quad = m^{2t-s-1}(m^2+\sigma^2)Ef^2\big(W(t)\big).
\end{aligned}$$

Therefore, by orthogonality from (5.1) we have

$$\begin{aligned}
\mathrm{Var}\Big[\int f(x)\Psi(dx,t)\Big] &= \sum_{s=1}^{t} m^{2t-s-1}(m^2+\sigma^2)Ef^2\big(W(t)\big)\\
&\le m^{2t}\cdot\frac{m^2+\sigma^2}{m(m-1)}Ef^2\big(W(t)\big).
\end{aligned}\tag{5.4}$$

Finally, the desired conclusion follows from (5.3) and (5.4). □

To prove Theorem 3.1 and 3.2, we now follow Révész's decomposition again

$$\begin{aligned}
&\frac{\psi(A,T)}{m^T} - B\Big(\frac{1}{2\pi T}\Big)^{d/2}\int_A \exp\Big\{-\frac{\|x\|^2}{2T}\Big\}dx\\
&\quad = \Big(\frac{\psi(A,T)}{m^T} - \frac{F(A,T,t)}{m^T}\Big)\\
&\qquad + \Big(\frac{F(A,T,t)}{m^T} - P_T(0,A)\frac{B(t)}{m^t}\Big)\\
&\qquad + P_T(0,A)\Big(\frac{B(t)}{m^t} - B\Big),
\end{aligned}$$

$$\frac{\psi(\{y; y \le x\sqrt{T}\}, T)}{m^T} - B\Phi_d(x)$$
$$= \left(\frac{\psi(\{y; y \le x\sqrt{T}\}, T)}{m^T} - \frac{F(\{y; y \le x\sqrt{T}\}, T, t)}{m^T}\right)$$
$$+ \left(\frac{F(\{y; y \le x\sqrt{T}\}, T, t)}{m^T} - \Phi_d(x)\frac{B(t)}{m^t}\right) + \Phi_d(x)\left(\frac{B(t)}{m^t} - B\right),$$

where we let $t = T^\varepsilon$ for a small constant $\varepsilon > 0$. By (4.4) and the inequality given in Lemma 6.11 of Révész (1994), one can see that only the second term in each of these two decompositions contributes to the limit behaviors given in our theorems. In other words, (3.1) and (3.8) are equivalent to

$$T^{1+d/2}\left(\frac{F(A, T, t)}{m^T} - P_T(0, A)\frac{B(t)}{m^t}\right) \longrightarrow |A|\left(\frac{1}{2\pi}\right)^{d/2}\left(\frac{1}{2}M + \bar{x}_A \cdot N\right) \tag{5.5}$$

and

$$\sqrt{T}\left(\frac{F(\{y; y \le x\sqrt{T}\}, T, t)}{m^T} - \Phi_d(x)\frac{B(t)}{m^t}\right) \longrightarrow -\nabla\Phi_d(x) \cdot N, \tag{5.6}$$

respectively.

We first prove (5.5). Notice that

$$\begin{aligned}\frac{F(A, T, t)}{m^T} - P_T(0, A)\frac{B(t)}{m^t} &= \frac{1}{m^t}\int [P_{T-t}(x, A) - P_T(0, A)]\psi(dx, t)\\ &= \frac{1}{m^t}\int_{|x|\le t}[P_{T-t}(x, A) - P_T(0, A)]\psi(dx, t)\\ &\quad + \frac{1}{m^t}\int_{|x|>t}[P_{T-t}(x, A) - P_T(0, A)]\psi(dx, t)\end{aligned}$$

and that, by Lemma 5.1,

$$E\left[\frac{1}{m^t}\int_{|x|>t}[P_{T-t}(x, A) - P_T(0, A)]\psi(dx, t)\right]^2 = O(P\{|W(t)| > t\}) = O(e^{-ct}).$$

So (5.5) is equivalent to

$$\begin{aligned}&T^{1+d/2}\left[\frac{1}{m^t}\int_{|x|\le t}[P_{T-t}(x, A) - P_T(0, A)]\psi(dx, t)\right]\\ &\qquad\longrightarrow |A|\left(\frac{1}{2\pi}\right)^{d/2}\left(\frac{1}{2}M + \bar{x}_A \cdot N\right).\end{aligned} \tag{5.7}$$

On the other hand,

$$\begin{aligned}&P_{T-t}(x, A) - P_T(0, A)\\ &\quad = \frac{1}{2}\left(\frac{1}{2\pi}\right)^{d/2}|A|T^{-1-d/2}\left[dt - \|x\|^2 + \frac{2}{|A|}\int_A (x \cdot y)\, dy + o(1)\right],\end{aligned}$$

where the error term $o(1)$ tends to 0 uniformly over $|x| \le t$ as $T \to \infty$. Hence

$$\begin{aligned}
&T^{1+d/2}\frac{1}{m^t}\int_{|x|\le t}\big[P_{T-t}(x, A) - P_T(0, A)\big]\psi(dx, t)\\
&= \frac{1}{2}\left(\frac{1}{2\pi}\right)^{d/2}\frac{|A|}{m^t}\int_{\|x\|\le t}\left[dt - |x|^2 + \frac{2}{|A|}\int_A (x\cdot y)\,dy\right]\psi(dx, t) + o(1)\\
&= \frac{1}{2}\left(\frac{1}{2\pi}\right)^{d/2}|A|\left[dt\frac{B(t)}{m^t} - \frac{1}{m^t}\int |x|^2\psi(dx, t) + 2\bar{x}_A\frac{1}{m^t}\int x\psi(dx, t)\right] + o(1)
\end{aligned}$$

almost surely as well as in L_2-norm, where the last step follows from a truncation estimate via Lemma 5.1.

Similarly as in Section 4, we can show that

$$M_t \equiv dt\frac{B(t)}{m^t} - \frac{1}{m^t}\int \|x\|^2\psi(dx, t) \quad \text{and} \quad N_t \equiv \frac{1}{m^t}\int x\psi(dx, t)$$

are two martingales with

$$\begin{aligned}
E\big[M_t - M_{t-1}\big]^2 &= \frac{2}{m^{t+1}}(m^2 + \sigma^2)dt^2,\\
\mathbf{cov}\big(N_t - N_{t-1}, N_t - N_{t-1}\big) &= \frac{1}{m^{t+1}}(m^2 + \sigma^2)\mathbf{I}_d,
\end{aligned}$$

for all $t \ge 1$. By the martingale convergence theorem $\{M_t\}$ and $\{N_t\}$ converge, almost surely as well as in L_2-norm, to some random variables M and N, respectively. Further,

$$\begin{aligned}
EM^2 &= \sum_{t=1}^{\infty} E\big[M_t - M_{t-1}\big]^2 = \frac{2d(m^2 + \sigma^2)(m + 1)}{(m - 1)^3},\\
\mathbf{cov}(N, N) &= \sum_{t=1}^{\infty}\mathbf{cov}\big(N_t - N_{t-1}, N_t - N_{t-1}\big) = \frac{(m^2 + \sigma^2)}{(m - 1)^2}\mathbf{I}_d.
\end{aligned}$$

Hence we have (5.7) [and therefore (3.1)–(3.3), (3.5), (3.6)]. We omit the proofs of (3.4) and (3.7), as they are analogous to that of (2.4)and (2.7), respectively.

We now come to the proof of (5.6). Note that for all $|y| \le t$, uniformly we have

$$\Phi_d\left(\sqrt{\frac{T}{T-t}}x - \frac{y}{\sqrt{T-t}}\right) - \Phi_d(x) = -T^{-1/2}\big[\nabla\Phi_d(x)\cdot y + o(1)\big].$$

Hence, by truncation (Lemma 5.1) we have

$$\begin{aligned}
&\sqrt{T}\left(\frac{F\big(\{y; y \le x\sqrt{T}\}, T, t\big)}{m^T} - \Phi_d(x)\frac{B(t)}{m^t}\right)\\
&= \sqrt{T}\frac{1}{m^t}\int\big[P\big\{W(T-t) \le x\sqrt{T} - y\big\} - \Phi_d(x)\big]\psi(dy, t)\\
&= -\nabla\Phi_d(x)\cdot\left(\frac{1}{m^t}\int y\psi(dy, t)\right) + o(1)
\end{aligned}$$

almost surely as well as in L_2-norm. So (5.6) follows from (3.6) and the martingale convergence theorem. □

Acknowledgments. Part of this work was done when the author was at the University of Utah and stimulating discussions with Professor Davar Khoshnevisan are gratefully acknowledged. Thanks also go to Professor Jan Rosinski who brought the author's attention to the Edgeworth expansion that led to Corollary 2.3 in its present form.

REFERENCES

ASMUSSEN, S. and KAPLAN, N. (1976a). Branching random walks I. *Stochastic Process. Appl.* **4** 1–13.

ASMUSSEN, S. and KAPLAN, N. (1976b). Branching random walks II. *Stochastic Process. Appl.* **4** 15–31.

ATHREYA, K. B. and KANG, H.-J. (1998a). Some limit theorems for positive recurrent branching Markov chains I. *Adv. in Appl. Probab.* **30** 693–710.

ATHREYA, K. B. and KANG, H.-J. (1998b). Some limit theorems for positive recurrent branching Markov chains II. *Adv. in Appl. Probab.* **30** 711–722.

ATHREYA, K. B. and KAPLAN, N. (1978). The additive property and its application in branching processes. *Adv. in Appl. Probab.* **5** 27–60.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

BIGGINS, J. D. (1990). The central limit theorem for the supercritical branching random walk, and related results. *Stochastic Process. Appl.* **34** 255–274.

BRAMSON, M., NEY, P. E. and TAO, J. (1992). The population composition of a multitype branching random walk. *Ann. Appl. Probab.* **2** 575–596.

DEMBO, A. and ZEITOUNI, O. (1992). *Large Deviations Techniques and Applications*. Jones & Bartlett, Boston.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.

HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

JOFFE, A. (1987). Diffusion approximation of multitype branching processes. In *Probability Theory and Mathematical Statistics* **2** 31–36. VNU Science Press, Utrecht.

KLEBANER, C. F. (1982). Branching random walk in varying environments. *Adv. in Appl. Probab.* **14** 359–367.

KOLMOGOROV, A. N. (1941). Über das logarithmisch normale Vertilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *Doklady* **31** 99–101.

LAWLER, G. (1991). *Intersections of Random Walks*. Birkhäuser, Boston.

NEY, P. E. (1991). Branching random walks. In *Spatial Stochastic Process* (K. S. Alexander and J. C. Watkins, eds.) Birkhäuser, Boston. 3–22.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, Berlin.

RÉVÉSZ, P. (1994). *Random Walks of Infinitely Many Particles*. World Scientific, Singapore.

STAM, A. J. (1966). On a conjecture of Harris. *Z. Wahrsch. Verw. Gebiete*. **5** 202–206.

WATANABE, S. (1965). On the branching process for Brownian particles with an absorbing boundary. *J. Math. Kyoto Univ.* **4** 185–198.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TENNESSEE
KNOXVILLE, TENNESSEE 37996
E-MAIL: xchen@math.utk.edu