

## THE EMPIRICAL MEAN POSITION OF A BRANCHING LÉVY PROCESS

DAVID CHEEK,\* *Harvard University*  
SEVA SHNEER,\*\* *Heriot-Watt University*

### Abstract

We consider a supercritical branching Lévy process on the real line. Under mild moment assumptions on the number of offspring and their displacements, we prove a second-order limit theorem on the empirical mean position.

*Keywords:* Branching Lévy process; branching random walk; branching Brownian motion; empirical mean position; empirical mean depth

2010 Mathematics Subject Classification: Primary 60J80  
Secondary 60J25

### 1. Introduction

A branching Lévy process describes a population of particles undergoing spatial movement, death, and reproduction. It can be defined informally as follows (for a formal definition, see Section 2). Initially there is one particle located at the origin of the real line. The particle lives for an exponentially distributed time. During this time it moves according to a Lévy process. At the time of death, the particle is replaced by a random number of new particles, displaced from the parent particle's death position according to a point process. All particles move, die, and reproduce in a statistically identical manner, independently of every other particle. We are only concerned with the supercritical case. That is, each particle gives birth to more than one particle on average, and thus the total number of particles grows to infinity with positive probability.

The particle positions' empirical distribution has received much attention, especially for branching random walks and branching Brownian motion, which are special cases of the model. There are many results on the empirical distribution's maximum [4, 5, 10], as well as on large deviations [12] and on the almost-sure weak convergence to a Gaussian distribution [3, 11].

The empirical mean position, which is simple and important for applications, has received relatively little attention. For specific branching random walks, [13] shows that the empirical mean position almost surely grows asymptotically linearly with time, while [7] shows that the empirical mean position's variance converges. These results combined raise the question of characterising a second-order limit term.

For branching Lévy processes, under some mild moment assumptions on the number of offspring and their displacements, we prove a second-order limit theorem for the empirical

---

Received 2 February 2020; revision received 7 May 2020.

\* Postal address: Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts, USA. Email address: [dmcheek@g.harvard.edu](mailto:dmcheek@g.harvard.edu)

\*\* Postal address: School of MACS, Heriot-Watt University, Edinburgh EH14 4AS, UK. Email address: [v.shneer@hw.ac.uk](mailto:v.shneer@hw.ac.uk)

mean position. Namely, we show that the difference between the empirical mean position at time  $t$  and  $rt$ , for some constant  $r$ , converges almost surely to a random variable.

Before proceeding with the remainder of the paper, we discuss some special cases of the model and applications.

First, consider that particles do not move during their lifetime and that each particle is displaced by  $+1$  from its parent. A particle’s position is its generation. Our result describes the average generation, complementing results of [13] and [6]. Second, consider instead that displacement sizes are Poisson-distributed. This is a popular model for cancer evolution [8]. Here particles are cells, and a cell’s position is its number of mutations. Our result gives the average number of mutations per cell. Third, consider that particles are not displaced from their parent but move as a random walk during their lifetime. This model is seen in phylogenetics. The branching process represents speciation [1], while the positions are lengths of a particular DNA segment [9].

The remainder of the paper is organised as follows. We introduce the model in Section 2, formulate our main result in Section 3, and prove it in Section 4.

### 2. Model

Initially there is a single particle named  $\emptyset$  which moves according to a Lévy process  $(Z_{\emptyset,s})_{s \geq 0}$ , with  $Z_{\emptyset,0} = 0$  and  $\mathbb{E}[Z_{\emptyset,1}^2] < \infty$ . After an exponentially distributed waiting time  $A_\emptyset$ , the particle dies and is replaced by a random number  $N_\emptyset$  of new particles with  $\mathbb{E}[N_\emptyset] > 1$  and  $\mathbb{E}[N_\emptyset^2] < \infty$ . The new particles are born at positions  $(Z_{\emptyset,A_\emptyset} + D_i)_{i=1}^{N_\emptyset}$ . The  $D_i$  are  $\mathbb{R}$ -valued random variables with

$$\mathbb{E} \left[ \left( \sum_{i=1}^{N_\emptyset} D_i \right)^2 \right] < \infty \quad \text{and} \quad \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i^2 \right] < \infty.$$

Independence is assumed between  $(Z_{\emptyset,s})_{s \geq 0}$  and  $A_\emptyset$  and  $(D_i)_{i=1}^{N_\emptyset}$  (but the  $D_i$  need not be independent of each other or of  $N_\emptyset$ ). All particles independently follow the initial particle’s behaviour.

To denote particles we follow standard notation. Let

$$\mathcal{T} = \bigcup_{n \in \mathbb{N} \cup \{0\}} \mathbb{N}^n.$$

Here  $\mathbb{N}^0 = \{\emptyset\}$  contains the initial particle. For  $v = (v_1, \dots, v_n) \in \mathcal{T}$  and  $i \in \mathbb{N}$ , write  $vi = (v_1, \dots, v_n, i)$ , where  $v$  is the parent of  $vi$ . To describe genealogical relationships, the set  $\mathcal{T}$  is endowed with a partial ordering  $<$ , defined by

$$(u_i)_{i=1}^m < (v_i)_{i=1}^n \iff m < n \text{ and } (u_i)_{i=1}^m = (v_i)_{i=1}^m.$$

Write  $\leq$  for  $<$  or  $=$ .

Now let

$$[(Z_{v,s})_{s \geq 0}, A_v, (D_{vi})_{i=1}^{N_v}]$$

for  $v \in \mathcal{T}$  be independent and identically distributed copies of

$$[(Z_{\emptyset,s})_{s \geq 0}, A_\emptyset, (D_i)_{i=1}^{N_\emptyset}].$$

The set of all particles that ever exist is

$$\mathcal{T}^* = \{(v_i)_{i=1}^n \in \mathcal{T} : v_{m+1} \leq N_{(v_i)_{i=1}^m}, \text{ for } m = 0, 1, \dots, n - 1\}.$$

The particles alive at time  $t \geq 0$  are

$$\mathcal{T}_t = \left\{ v \in \mathcal{T}^* : \sum_{u < v} A_u \leq t < \sum_{u \leq v} A_u \right\}.$$

Particle  $v$  at time  $t$ , if it is alive, has position

$$X_{v,t} = \sum_{\emptyset < u \leq v} D_u + \sum_{\emptyset \leq u < v} Z_{u,A_u} + Z_{v,t - \sum_{\emptyset \leq u < v} A_u}.$$

For further notation, the branching rate is

$$\lambda = \mathbb{E}[A_\emptyset]^{-1},$$

the effective branching rate is

$$\hat{\lambda} = \lambda \mathbb{E}[N_\emptyset - 1],$$

and the movement rate is

$$r = \mathbb{E}[Z_{\emptyset,1}] + \lambda \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i \right].$$

### 3. Main result

**Theorem 3.1.** *Conditional on the event  $\{\lim_{t \rightarrow \infty} |\mathcal{T}_t| = \infty\}$ , the limit*

$$\lim_{t \rightarrow \infty} \frac{1}{|\mathcal{T}_t|} \left( \sum_{v \in \mathcal{T}_t} X_{v,t} - rt \right)$$

*exists and is finite almost surely.*

### 4. Proof of Theorem 3.1

Our proof will involve conditioning on whether branching occurs during the time interval  $[0, h]$  for some small  $h > 0$ . Write

$$J_{0,h} = \{A_\emptyset > h\}$$

for the event that the first branching occurs after time  $h$ . Write

$$J_{1,h} = \left\{ A_\emptyset \leq h < A_\emptyset + \min_{i=1, \dots, N_\emptyset} A_i \right\}$$

for the event that the first branching occurs before time  $h$  and the second branching occurs after time  $h$ . Write

$$J_{2,h} = \left\{ A_\emptyset + \min_{i=1, \dots, N_\emptyset} A_i \leq h \right\}$$

for the event that the second branching occurs before time  $h$ . Note the probabilities

$$\begin{cases} \mathbb{P}[J_{0,h}] = 1 - h\lambda + o(h), \\ \mathbb{P}[J_{1,h}] = h\lambda + o(h), \\ \mathbb{P}[J_{2,h}] = o(h), \end{cases}$$

as  $h \downarrow 0$ . Observe the conditional distribution

$$\left( \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \mid J_{0,h} \right) \stackrel{d}{=} \sum_{v \in \mathcal{T}'_t} (Z_{\emptyset,h} + X'_{v,t} - r(t+h)), \tag{4.1}$$

where  $(X'_{v,t})_{v \in \mathcal{T}'_t} \stackrel{d}{=} (X_{v,t})_{v \in \mathcal{T}_t}$ , and  $(X'_{v,t})_{v \in \mathcal{T}'_t}$  is independent of  $Z_{\emptyset,h}$ . Meanwhile

$$\left( \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \mid J_{1,h} \right) \stackrel{d}{=} \sum_{i=1}^{N_\emptyset} \sum_{v \in \mathcal{T}'_t} (D_i + X_{v,t}^i - rt) + \eta_h, \tag{4.2}$$

where  $(X_{v,t}^i)_{v \in \mathcal{T}'_t} \stackrel{d}{=} (X_{v,t})_{v \in \mathcal{T}_t}$  for  $i = 1, \dots, N_\emptyset$ , the  $(X_{v,t}^i)_{v \in \mathcal{T}'_t}$  are independent of each other and of  $(D_i)_{i=1}^{N_\emptyset}$ , and

$$\eta_h = \left( \sum_{i=1}^{N_\emptyset} |\mathcal{T}'_t| (Z_{\emptyset, A_\emptyset} + Z_{i, h-A_\emptyset} - rh) \mid J_{1,h} \right).$$

Straightforward calculations show that the first and second moments of  $\eta_h$  converge to 0 as  $h \downarrow 0$ .

**Lemma 4.1.** For  $t \geq 0$ ,

$$\mathbb{E} \left[ \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right] = 0.$$

*Proof.* From (4.1),

$$\mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \mid J_{0,h} \right] = \mathbb{E} \left[ \sum_{v \in \mathcal{T}'_t} (X_{v,t} - rt) \right] + h(\mathbb{E}[Z_{\emptyset,1}] - r)\mathbb{E}|\mathcal{T}'_t|.$$

From (4.2),

$$\begin{aligned} & \mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \mid J_{1,h} \right] \\ &= \mathbb{E}[N_\emptyset] \mathbb{E} \left[ \sum_{v \in \mathcal{T}'_t} (X_{v,t} - rt) \right] + \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i \right] \mathbb{E}|\mathcal{T}'_t| + o(1). \end{aligned}$$

Taking the unconditional expectation,

$$\begin{aligned} & \mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \right] \\ &= (1 - h\lambda) \mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \mid J_{0,h} \right] \end{aligned}$$

$$\begin{aligned}
 &+ h\lambda \mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) | J_{1,h} \right] + o(h) \\
 &= \mathbb{E} \left[ \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right] (1 + h\hat{\lambda}) + o(h).
 \end{aligned}$$

Rearranging and taking  $h \downarrow 0$ ,

$$\frac{d}{dt} \mathbb{E} \left[ \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right] = \hat{\lambda} \mathbb{E} \left[ \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right].$$

The statement of the lemma for any  $t$  now follows from the above and the fact that it clearly holds for  $t = 0$ . □

Next we determine second moments.

**Lemma 4.2.** For  $t \geq 0$ ,

$$\mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right)^2 \right] = c_1 e^{2\hat{\lambda}t} - c_2 t e^{\hat{\lambda}t} - c_1 e^{\hat{\lambda}t},$$

where

$$c_1 = \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]^2} \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i^2 \right] + \frac{1}{\mathbb{E}[N_\emptyset - 1]} \mathbb{E} \left[ \left( \sum_{i=1}^{N_\emptyset} D_i \right)^2 \right]$$

and

$$c_2 = \hat{\lambda} \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]^2} \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i^2 \right].$$

*Proof.* From (4.1),

$$\begin{aligned}
 \mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \right)^2 \mid J_{0,h} \right] &= \mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right)^2 \right] \\
 &+ 2h(\mathbb{E}[Z_{\emptyset,1}] - r) \mathbb{E} \left[ |\mathcal{T}_t| \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right] \\
 &+ h(\mathbb{E}[Z_{\emptyset,1}^2] - \mathbb{E}[Z_{\emptyset,1}]^2) \mathbb{E}[|\mathcal{T}_t|^2] \\
 &+ o(h).
 \end{aligned}$$

From (4.2) and Lemma 4.1,

$$\begin{aligned}
 \mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h)) \right)^2 \mid J_{1,h} \right] &= \mathbb{E} \left[ \left( \sum_{i=1}^{N_\emptyset} \sum_{v \in \mathcal{T}_t^i} (X_{v,t} - rt) \right)^2 \right] \\
 &+ 2 \mathbb{E} \left[ \sum_{i=1}^{N_\emptyset} D_i |\mathcal{T}_t^i| \sum_{v \in \mathcal{T}_t^i} (X_{v,t} - rt) \right]
 \end{aligned}$$

$$\begin{aligned}
 &+ 2\mathbb{E}\left[\sum_{\substack{i,j=1 \\ i \neq j}}^{N_\emptyset} D_i |\mathcal{T}_t^i| \sum_{v \in \mathcal{T}_t^j} (X_{v,t} - rt)\right] \\
 &+ \mathbb{E}\left[\left(\sum_{i=1}^{N_\emptyset} D_i |\mathcal{T}_t^i|\right)^2\right] + o(1) \\
 &= \mathbb{E}[N_\emptyset] \mathbb{E}\left[\left(\sum_{v \in \mathcal{T}_t} (X_{v,t} - rt)\right)^2\right] \\
 &+ 2\mathbb{E}\left[\sum_{i=1}^{N_\emptyset} D_i\right] \mathbb{E}\left[|\mathcal{T}_t| \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt)\right] \\
 &+ \mathbb{E}\left[\sum_{i=1}^{N_\emptyset} D_i^2\right] \left(\mathbb{E}\left[|\mathcal{T}_t|^2\right] - \left(\mathbb{E}|\mathcal{T}_t|\right)^2\right) \\
 &+ \mathbb{E}\left[\left(\sum_{i=1}^{N_\emptyset} D_i\right)^2\right] \left(\mathbb{E}|\mathcal{T}_t|\right)^2 + o(1).
 \end{aligned}$$

But  $\mathbb{E}|\mathcal{T}_t|$  and  $\mathbb{E}[|\mathcal{T}_t|^2]$  are standard knowledge [2]:

$$\mathbb{E}|\mathcal{T}_t| = e^{\hat{\lambda}t}$$

and

$$\mathbb{E}[|\mathcal{T}_t|^2] = \left(1 + \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]}\right) e^{2\hat{\lambda}t} - \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]} e^{\hat{\lambda}t}.$$

Therefore

$$\begin{aligned}
 &\mathbb{E}\left[\left(\sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h))\right)^2\right] \\
 &= (1 - \lambda h) \mathbb{E}\left[\left(\sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h))\right)^2 \mid J_{0,h}\right] \\
 &\quad + h\lambda \mathbb{E}\left[\left(\sum_{v \in \mathcal{T}_{t+h}} (X_{v,t+h} - r(t+h))\right)^2 \mid J_{1,h}\right] + o(h) \\
 &= 1 + h\hat{\lambda} \mathbb{E}\left[\left(\sum_{v \in \mathcal{T}_t^i} (X_{v,t} - rt)\right)^2\right] \\
 &\quad + hae^{2\hat{\lambda}t} + hbe^{\hat{\lambda}t} + o(h),
 \end{aligned}$$

where

$$a = \lambda \left( \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]} \mathbb{E}\left[\sum_{i=1}^{N_\emptyset} D_i^2\right] + \mathbb{E}\left[\left(\sum_{i=1}^{N_\emptyset} D_i\right)^2\right] \right)$$

and

$$b = -\lambda \frac{\mathbb{E}[(N_\emptyset - 1)^2]}{\mathbb{E}[N_\emptyset - 1]} \mathbb{E}\left[\sum_{i=1}^{N_\emptyset} D_i^2\right].$$

Rearranging and taking  $h \downarrow 0$ ,

$$\frac{d}{dt} \mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_t^i} (X_{v,t} - rt) \right)^2 \right] = \hat{\lambda} \mathbb{E} \left[ \left( \sum_{v \in \mathcal{T}_t^i} (X_{v,t} - rt) \right)^2 \right] + ae^{2\hat{\lambda}t} + be^{\hat{\lambda}t}.$$

The statement of Lemma 4.2 now follows directly from the differential equation above. □

Next we present a martingale result for which a filtration  $(\mathcal{F}_t)_{t \geq 0}$  needs to be defined:

$$\mathcal{F}_t = \sigma((X_{v,s})_{v \in \mathcal{T}_s} : 0 \leq s \leq t).$$

**Lemma 4.3.**

$$\left( e^{-\hat{\lambda}t} \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \right)_{t \geq 0}$$

is a martingale with respect to  $(\mathcal{F}_t)_{t \geq 0}$ .

*Proof.* Write

$$\mathcal{T}_{u,t} = \{v \in \mathcal{T}_t : u \leq v\}$$

for the particles alive at time  $t$  which are descendants of  $u \in \mathcal{T}$ . Let  $0 \leq s \leq t$ . Then

$$e^{-\hat{\lambda}t} \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) = e^{-\hat{\lambda}t} \sum_{u \in \mathcal{T}_s} \sum_{v \in \mathcal{T}_{u,t}} (X_{v,t} - X_{u,s} - r(t-s)) + e^{-\hat{\lambda}t} \sum_{u \in \mathcal{T}_s} |\mathcal{T}_{u,t}| (X_{u,s} - rs).$$

Taking conditional expectations,

$$\begin{aligned} & \mathbb{E} \left[ e^{-\hat{\lambda}t} \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) \mid \mathcal{F}_s \right] \\ &= e^{-\hat{\lambda}t} |\mathcal{T}_s| \mathbb{E} \left[ \sum_{v \in \mathcal{T}_{t-s}} (X_{v,t-s} - r(t-s)) \right] + e^{-\hat{\lambda}t} \sum_{u \in \mathcal{T}_s} e^{\hat{\lambda}(t-s)} (X_{u,s} - rs) \\ &= e^{-\hat{\lambda}s} \sum_{u \in \mathcal{T}_s} (X_{u,s} - rs), \end{aligned}$$

where the last equality is due to Lemma 4.1. □

*Proof of Theorem 3.1.* By Lemmas 4.2 and 4.3 and the martingale convergence theorem, there is a  $\mathbb{R}$ -valued random variable  $V$  with

$$\lim_{t \rightarrow \infty} e^{-\hat{\lambda}t} \sum_{v \in \mathcal{T}_t} (X_{v,t} - rt) = V \tag{4.3}$$

almost surely. But conditioned on the event  $\{\lim_{t \rightarrow \infty} |\mathcal{T}_t| = \infty\}$ , there is a positive random variable  $W$  with

$$\lim_{t \rightarrow \infty} e^{-\hat{\lambda}t} |\mathcal{T}_t| = W \tag{4.4}$$

almost surely [2]. Combine (4.3) and (4.4) to conclude the proof. □

### Acknowledgements

The authors are grateful to Ken Duffy for a number of useful discussions and to the anonymous referee for a number of useful comments on the first version of our paper.

## References

- [1] ALDOUS, D. J. (2001). Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statist. Sci.* **16**, 23–34.
- [2] ATHREYA, K. AND NEY, P. (1972). *Branching Processes*. Springer.
- [3] BIGGINS, J. (1990). The central limit theorem for the supercritical branching random walk, and related results. *Stoch. Process. Appl.* **34**, 255–274.
- [4] BIGGINS, J. (1995). The growth and spread of the general branching random walk. *Ann. Appl. Prob.* **5**, 1008–1024.
- [5] BRAMSON, M. D. (1978). Maximal displacement of branching Brownian motion. *Commun. Pure Appl. Math.* **31**, 531–581.
- [6] CHAUVIN, B., KLEIN, T., MARCKERT, J.-F. AND ROUAULT, A. (2005). Martingales and profile of binary search trees. *Electron. J. Prob.* **10**, 420–435.
- [7] DUFFY, K. R., MELI, G. AND SHNEER, S. (2019). The variance of the average depth of a pure birth process converges to 7. *Statist. Prob. Lett.* **150**, 88–93.
- [8] DURRETT, R. (2013). Population genetics of neutral mutations in exponentially growing cancer cell populations. *Ann. Appl. Prob.* **23**, 230–250.
- [9] FELSENSTEIN, J. (2004). *Inferring Phylogenies*. Sinauer Associates.
- [10] GANTERT, N. AND HÖFELSAUER, T. (2018). Large deviations for the maximum of a branching random walk. *Electron. Commun. Prob.* **23**, 34.
- [11] GAO, Z. AND LIU, Q. (2018). Second and third orders asymptotic expansions for the distribution of particles in a branching random walk with a random environment in time. *Bernoulli* **24**, 772–800.
- [12] LOUIDOR, O. AND TSAIRI, E. (2017). Large deviations for the empirical distribution in the general branching random walk. Available at [arXiv:1704.02374](https://arxiv.org/abs/1704.02374).
- [13] MELI, G., WEBER, T. S. AND DUFFY, K. R. (2019). Sample path properties of the average generation of a Bellman–Harris process. *J. Math. Biol.* **79**, 673–704.