Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 42, 1–35.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-4233

CLT for Ornstein-Uhlenbeck branching particle system*

Radosław Adamczak† Piotr Miłoś‡

Abstract

In this paper we consider a branching particle system consisting of particles moving according to an Ornstein-Uhlenbeck process in $\mathbb{R}^d$ and undergoing binary, supercritical branching with a constant rate $\lambda > 0$. This system is known to fulfil a law of large numbers (under exponential scaling). In the paper we prove the corresponding central limit theorem. The limit and the CLT normalization fall into three qualitatively different classes. In what we call the small branching rate case the situation resembles the classical one. The weak limit is Gaussian and normalization is the square root of the size of the system. In the critical case the limit is still Gaussian, but the normalization requires an additional term. Finally, when branching has a large rate the situation is completely different. The limit is no longer Gaussian, the normalization is substantially larger than the classical one and the convergence holds in probability. We also prove that the spatial fluctuations are asymptotically independent of the fluctuations of the total number of particles (which is a Galton-Watson process).

Keywords: Supercritical branching particle systems; Central limit theorem; Galton-Watson process.
AMS MSC 2010: 60J80; 60F05.
Submitted to EJP on November 22, 2011, final version accepted on September 13, 2014.
Supersedes arXiv:1111.4559.

1 Introduction

We consider a branching particle system $\{X_t\}_{t\geq 0}$ as follows. The system starts off at time $t = 0$ from a single particle located at $x \in \mathbb{R}^d$. The particle moves according to an Ornstein-Uhlenbeck process in $\mathbb{R}^d$ and branches after an exponential time with parameter $\lambda > 0$. Before describing the branching mechanism, let us recall that the Ornstein-Uhlenbeck process with parameters $\sigma, \mu > 0$ is a time homogenous Markov process with the infinitesimal operator

$$L := \frac{1}{2}\sigma^2\Delta - \mu x \circ \nabla, \tag{1.1}$$

where $\circ$ denotes the standard scalar product in $\mathbb{R}^d$ and $\nabla = (\partial/\partial x_1, \ldots, \partial/\partial x_d)$, $\Delta = \sum_{i=1}^d \partial^2/\partial x_i^2$, stand for the gradient and Laplacian operators respectively.

*Support: MNiSW grant N N201 397437 & N N201 397537; Foundation for Polish Science.
†University of Warsaw, Poland. E-mail: radamcz@mimuw.edu.pl.
‡University of Warsaw, Poland. E-mail: pmilos@mimuw.edu.pl (corresponding author).

We consider binary branching with parameter $p > 1/2$, i.e. at each splitting time the particle has two or zero offspring with probability p and $1-p$ respectively. We recall that for binary branching mechanisms the condition $p > 1/2$ is equivalent to supercriticality. Thus with positive probability our system does not become extinct.

The offspring particles follow the same dynamics independently. This system will be referred to as the OU branching process. Formally, the system is identified with the empirical process, i.e. X is a measure-valued process such that for a Borel set A, $X_t(A)$ is the (random) number of particles at time t in A. The existence of such a process is guaranteed by general results in [11, 12]; see also further comments in Section 4.1.

Systems of this type may be regarded as consisting of two components, namely the genealogy part and the diffusion part (for this reason they are sometimes called "branching diffusions"). The genealogy part, being the celebrated Galton-Watson process, is well-studied. In our case ($p > 1/2$), the expected total number of particles grows exponentially at rate

$$\lambda_p := (2p-1)\lambda. \tag{1.2}$$

The Ornstein-Uhlenbeck process has a unique equilibrium measure φ (to be described later). After a long time the positions of two "randomly picked" particles are "almost independent" random variables distributed approximately according to φ, which suggests the following law of large numbers

$$|X_t|^{-1} \langle X_t, f\rangle \to \langle \varphi, f\rangle 1_{Ext^c}, \quad a.s. \tag{1.3}$$

where Ext^c is the event that the system does not become extinct, $|X_t|$ denotes the number of particles at time t and f is a bounded, continuous function. This is indeed the case as follows from [13]. In Theorem 3.1 we obtain (1.3) in probability for a slightly more general class of functions. This is, however, only a preparatory step towards our main goal which is the corresponding central limit theorem. It turns out that the second order behaviour depends qualitatively on the sign of $\lambda_p - 2\mu$. Roughly speaking, this condition reflects the interplay of two antagonistic forces, the growth which is local and makes the system more coarse and the movement which tends to smooth the system (higher μ implies "a stronger attraction of particles towards 0"). Now we describe the behaviour of the spatial fluctuations:

$$F_t^{-1}\left(\langle X_t, f\rangle - |X_t|\langle \varphi, f\rangle\right), \tag{1.4}$$

where F_t is some, not necessarily deterministic, normalising function.

Our results are quite involved from the notational point of view, therefore we postpone their rigorous formulation to Section 3 and at this point we will only provide their somewhat informal description. We present the situation on the set of non-extinction Ext^c.

Small branching rate: $\lambda_p < 2\mu$. Our main result is contained in Theorem 3.3. In this case "the movement part prevails" and the result resembles the standard CLT. For large t,

$$\frac{\langle X_t, f\rangle - |X_t|\langle \varphi, f\rangle}{\sqrt{|X_t|}} \xrightarrow{d} \sigma_f G,$$

where $\xrightarrow{d}$ denotes convergence in distribution, G is a standard Gaussian variable and σ_f is a parameter depending on the function f, but independent of the initial conditions.

Let us note that a random normalization $F_t = |X_t|^{1/2}$ is quite natural as it "filters" out the fluctuation of the total number of particles and thus allows for a Gaussian limit. We remark that almost surely for large t, $|X_t|\exp(-\lambda_p t) \to V_\infty$ for a certain random variable V_∞, so one can easily replace our random normalization with a deterministic one. However the limit in this case will be no longer Gaussian; by using the full strength of Theorem 3.3 one can show that it is a mixture of Gaussian distributions.

Critical branching rate: $\lambda_p = 2\mu$**.** Our main result is given in Theorem 3.8. In this case "the branching prevails". The behaviour of the fluctuations slightly diverges from the classical setting, namely

$$\frac{\langle X_t, f\rangle - |X_t|\langle \varphi, f\rangle}{\sqrt{t|X_t|}} \xrightarrow{d} \sigma_f G,$$

where G is again a standard Gaussian variable and σ_f a constant depending on the function f (different from σ_f in the subcritical regime). Let us briefly comment on the differences between the critical and subcritical cases. First of all, in the critical case the normalization is bigger: $F_t = t^{1/2}|X_t|^{1/2}$. Moreover, although the limit still does not depend on the starting condition and is Gaussian, its variance depends on the distributional derivatives of f (contrary to the small branching case). To explain this at a heuristic level we note that in the critical case the branching is so fast that the fluctuations are no longer smoothened by the motion and become essentially local.

Large branching rate: $\lambda_p > 2\mu$**.** Our main result is presented in Theorem 3.12. In this case not only does the branching "prevail" but also "the motion badly fails to make any smoothing". In this case it is more natural to use deterministic normalization $F_t = e^{(\lambda_p-\mu)t}$, which is even bigger than in the critical case (note that in this range of λ_p, $\exp((\lambda_p - \mu)t) \gg \sqrt{t}\exp(\lambda_p t/2)$ while the later quantity is of the same order as $\sqrt{t|X_t|}$).

We have

$$\frac{\langle X_t, f\rangle - |X_t|\langle \varphi, f\rangle}{e^{(\lambda_p-\mu)t}} \to \langle \varphi, \nabla f\rangle \circ J,$$

where ∇f is the distributional gradient of f and J is a certain d-dimensional random vector (obtained as the limit of some martingale related to the system), depending on the starting position of the system. It turns out that the limit is no longer Gaussian. What is perhaps surprising, the convergence holds in probability. The first term, $\langle \varphi, \nabla f\rangle$, means that like in the critical situation the branching is fast enough so that the fluctuations are local. Even more, it is so fast that the limit depends on the starting condition and in fact, up to some extent, the system "remembers its whole evolution", which is encoded in J.

In either case we also prove that the spatial fluctuations become independent of fluctuations of the total number of particles as time increases.

Until recently most of the research activity was concentrated on proving laws of large numbers for more and more general branching systems. Before our work the only CLT result we are aware of was contained in [6, Proposition 6.4]. Their setting is somewhat different, see Remark 3.17. During the review process of the current paper an extension of our results appeared in [24] . We shortly comment on their result in Remark 3.16 and in Remark 3.18 we discuss a few open questions natural in view of our results.

At this moment we would like to announce our parallel paper [1]. In [1] we consider U-statistics of the OU branching system, namely expressions of the form

$$U_t^n(f) := \sum_{\substack{i_1,i_2,\ldots i_n=1 \\ i_k\neq i_j \text{ if } k\neq j}}^{|X_t|} f(X_t(i_1), X_t(i_2), \ldots, X_t(i_n)). \tag{1.5}$$

We obtain both the law of large numbers (i.e. an analogue of (1.3)) and CLTs, which also fall into three categories corresponding to the cases described above. Moreover, at this point we also advertise a forthcoming work of the second named author [21], which is devoted to studies of the CLT for superprocesses based on the Ornstein-Uhlenbeck process (see also related [25]). Qualitatively the results of [21] are the the same as the ones presented in this paper.

Our proofs utilise a mixture of techniques used for branching systems (e.g. the Laplace transform and the log-Laplace equation, coupling and decoupling). Although the proof schemes loosely resemble known techniques (e.g. are similar to proofs in [5, Section 1.13]) they required many improvements. In our proofs we used also the fact that the Ornstein-Uhlenbeck process has a particularly explicit and tractable structure.

The study of branching models of various types has a long history, we refer the reader to [12, 14, 5, 9] (the list is by no means exhaustive). It has been known for a long time that the behaviour of branching systems differs qualitatively for (sub)critical and supercritical cases. The former become extinct almost surely, their limit properties are studied after conditioning on non-extinction event (e.g. Yaglom type theorems) or as a part of larger infinite structures (e.g. Galton-Watson forests, random snakes, continuum trees – see e.g. [17]). The latter grow exponentially fast (on the set of non-extinction), which makes it possible to study them using laws of large numbers, starting with the celebrated Kesten-Stigum theorem ([5, Theorem I.10.1]). Such theorems were also proved for branching particle systems, they go back to [3, 4] and more recently [13] (the latter being the main inspiration for our paper). This paper presents a law of large numbers for a large class of supercritical branching diffusions and admits unbounded space-dependent branching intensity. Also we would like to mention again the paper [6] in which systems with non-local branching (i.e. particles may jump upon an event of branching) were studied. The article [6] presents a law of large numbers as well as a central limit theorem, though in a different spirit than ours (see comparison in Remark 3.17). Due to its extensive introduction [6] is also an excellent resource of biological motivations for study of branching diffusions.

The article is organised as follows. The next section presents notation and basic facts required further. Section 3 is devoted to the presentation of the results. Proofs are deferred to Section 4 and the Appendix.

2 Notation and preliminaries

Notation For a branching system $\{X_t\}_{t\geq 0}$, we denote by $|X_t|$ the number of particles at time t, and by $X_t(i)$ - the position of the i-th (in a certain ordering) particle at time t. Typically, we use $\mathbb{E}_x$ or $\mathbb{P}_x$ to stress the fact that we consider the system starting from a particle located at x. Sometimes we use also $\mathbb{E}$ and $\mathbb{P}$ when this location is not relevant (e.g. if we calculate the number of particles in the system). We will refer to the system starting from a single particle at time $t = 0$ located at $x \in \mathbb{R}^d$ shortly, as the OU branching system starting from $x \in \mathbb{R}^d$.
For a function $f\colon \mathbb{R}^d \to \mathbb{R}$ and a measure μ on $\mathbb{R}^d$ we denote $\langle \mu, f\rangle := \int_{\mathbb{R}^d} f(x)\mu(\mathrm{d}x)$. Thus in particular we have

$$\langle X_t, f\rangle = \sum_{i=1}^{|X_t|} f(X_t(i)).$$

By $\to^d$ we denote convergence in law. We use $\lesssim$ to denote the situation when an inequality holds up to a (multiplicative) constant $c > 0$ not depending on running parameters, which is irrelevant for calculations. E.g. $f(x) \lesssim g(x)$ means that there exists a constant $c > 0$

such that $f(x) \leq cg(x)$.
Let $x \circ y = \sum_{i=1}^d x_i y_i$ denote the standard scalar product of $x, y \in \mathbb{R}^d$. Moreover, $\|x\| = \sqrt{x \circ x}$ is the standard Euclidean norm in $\mathbb{R}^d$.

In the paper we use the space

$$\mathcal{P} = \mathcal{P}(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \mapsto \mathbb{R} : f \text{ is continuous and } \exists n \text{ such that } \lim_{\|x\| \to +\infty} |f(x)|/\|x\|^n = 0 \right\},$$

that is the space of continuous functions which grow at most polynomially.
Given a function $f \in \mathcal{P}(\mathbb{R}^d)$ we implicitly understand its derivatives (e.g. $\frac{\partial f}{\partial x_i}$) in the space of tempered distributions (see e.g. [26, p. 173]).

For a probability measure μ and a random variable X we write $X \sim \mu$ to express the fact that X is distributed according to μ.

Basic facts on the Galton-Watson process The number of particles $\{|X_t|\}_{t \geq 0}$ is the celebrated Galton-Watson process. We present basic properties of this process used in the paper. The main reference in this section is [5]. We already introduced the growth rate (1.2) (e.g. [5, Section 1.6]). The process becomes extinct with probability $p_e := \frac{1-p}{p}$. (see [5, Theorem I.5.1]). We denote the extinction and non-extinction events by Ext and Ext^c respectively. The process

$$V_t := e^{-\lambda_p t}|X_t|$$

is a positive martingale. Therefore it converges (see also [5, Theorem 1.6.1])

$$V_t \to V_\infty, \quad a.s. \text{ as } t \to +\infty. \tag{2.1}$$

Proposition 2.1. *We have $\{V_\infty = 0\} = Ext$ and conditioned on non-extinction V_∞ has the exponential distribution with parameter $\frac{2p-1}{p}$. We have $\mathbb{E}(V_\infty) = 1$ and $\mathrm{Var}(V_\infty) = \frac{1}{2p-1}$. $\mathbb{E}e^{-4\lambda_p t}|X_t|^4$ is uniformly bounded, i.e. there exists $C > 0$ such that for any $t \geq 0$ we have $\mathbb{E}e^{-4\lambda_p t}|X_t|^4 \leq C$. Moreover, all moments are finite, i.e. for any $n \in \mathbb{N}$ and $t \geq 0$ we have $\mathbb{E}|X_t|^n < +\infty$.*

The proof is deferred to the Appendix. We denote the variable V_∞ conditioned on non-extinction by W.

Basic facts on the Ornstein-Uhlenbeck process The particle movement in our model is governed by the Ornstein-Uhlenbeck process with parameters $\sigma, \mu > 0$. Let us recall again that this process is a time homogenous Markov process with the infinitesimal operator

$$L := \frac{1}{2}\sigma^2 \Delta - \mu x \circ \nabla. \tag{2.2}$$

The invariant measure of the Ornstein-Uhlenbeck process with parameters σ, μ is a mean zero Gaussian distribution on $\mathbb{R}^d$ with covariance matrix $\frac{\sigma^2}{2\mu}\mathrm{Id}$. We denote this measure by φ. Since in the proofs we will use extensively the exact form of its density, we recall that it is equal to

$$\varphi(x) := \left(\frac{\mu}{\pi\sigma^2}\right)^{d/2} \exp\left(-\frac{\mu}{\sigma^2}\|x\|^2\right). \tag{2.3}$$

For simplicity we denote the measure and its density with the same letter, which should not lead to a misunderstanding.

The semigroup of the Ornstein-Uhlenbeck process is denoted by $\mathcal{T}$. It is well known that $\mathcal{T}$ is a strongly continuous semigroup on $L^2(\varphi)$ (see e.g. Sections 8.3. and 9.4. in [8] for a much more general exposition in the context of infinite-dimensional spaces). It can be calculated with the following formula (see formula (8.13) in [8] or Section 5.2. of [2])

$$\mathcal{T}_t f(x) = (g_t * f)(x_t), \quad x_t := e^{-\mu t} x, \tag{2.4}$$

where

$$g_t(x) = \left(\frac{\mu}{\pi\sigma_t^2}\right)^{d/2} \exp\left\{-\frac{\mu}{\sigma_t^2}x^2\right\}, \quad \sigma_t^2 := \sigma^2(1-e^{-2\mu t}).$$

We denote $ou(t) := \sqrt{1-e^{-2\mu t}}$ and let $G \sim \varphi$. Similarly to (2.4) we can express $\mathcal{T}$ using

$$\mathcal{T}_t f(x) = \int_{\mathbb{R}^d} f(x_t - y) g_t(y) \mathrm{d}y = \int_{\mathbb{R}^d} f\left(xe^{-\mu t} + ou(t)y\right)\varphi(y)\mathrm{d}y = \mathbb{E} f(xe^{-\mu t} + ou(t)G). \tag{2.5}$$

For a function $f \in \mathcal{P}(\mathbb{R}^d)$ we will denote

$$\tilde{f}(x) := f(x) - \langle \varphi, f \rangle . \tag{2.6}$$

3 Results

This section is devoted to the presentation of our results. The proofs are deferred to Section 4. Our first aim is to present a central limit theorem corresponding to the following law of large numbers (closely related to [13, Theorem 6] or [6, Theorem 4.2]).

Theorem 3.1. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system starting from $x \in \mathbb{R}^d$. Let us assume that $f \in \mathcal{P}(\mathbb{R}^d)$. Then*

$$\lim_{t\to+\infty} e^{-\lambda_p t} \langle X_t, f\rangle = \langle \varphi, f\rangle V_\infty \quad \textit{in probability},$$

and thus (by the definition of V_∞ and Proposition 2.1) on the set of non-extinction, Ext^c, we have

$$\lim_{t\to+\infty} |X_t|^{-1} \langle X_t, f\rangle = \langle \varphi, f\rangle \quad \textit{in probability}. \tag{3.1}$$

Moreover, if f is bounded then almost sure convergence holds.

Remark 3.2. We believe that the almost sure convergence holds above also for $f \in \mathcal{P}(\mathbb{R}^d)$. However, as in this paper we concentrate on the CLT, we do not pursue this question here. We expect that methods used in the proof of [13, Lemma 18] might be of use.

3.1 Small branching rate: $\lambda_p < 2\mu$

We denote

$$\sigma_f^2 := \left\langle \varphi, \tilde{f}^2 \right\rangle + 2\lambda p \int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s}\mathcal{T}_s \tilde{f}\right)^2 \right\rangle \mathrm{d}s, \tag{3.2}$$

where $\tilde{f}$ is defined by (2.6).

Let us also recall (2.1) and that W is V_∞ conditioned on Ext^c. The main result of this section is

Theorem 3.3. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system starting from $x \in \mathbb{R}^d$. Let us assume $\lambda_p < 2\mu$ and $f \in \mathcal{P}(\mathbb{R}^d)$. Then $\sigma_f^2 < +\infty$ and conditionally on the set of non-extinction Ext^c there is the convergence*

$$\left(e^{-\lambda_p t}|X_t|, \frac{|X_t| - e^{t\lambda_p}V_\infty}{\sqrt{|X_t|}}, \frac{\langle X_t, f\rangle - |X_t|\langle\varphi, f\rangle}{\sqrt{|X_t|}}\right) \to^d (W, G_1, G_2), \tag{3.3}$$

where $G_1 \sim \mathcal{N}(0, 1/(2p-1)), G_2 \sim \mathcal{N}(0, \sigma_f^2)$ and W, G_1, G_2 are independent random variables.

Remark 3.4. As we already mentioned in the Introduction this is the most classical case. By conditioning one can check that the theorem is still valid when the first particle is distributed according to φ. With this assumption $X_t(i) \sim \varphi$ for any i. If these random variables were independent conditionally on $|X_t|$, then the third term, written in the form $|X_t|^{-1/2} \sum_{i=1}^{|X_t|}(f(X_t(i)) - \langle \varphi, f\rangle)$, would converge to $\mathcal{N}(0, \tilde{\sigma}_f^2)$, where $\tilde{\sigma}_f^2 = \left\langle \varphi, \tilde{f}^2 \right\rangle$ (the random number of elements in the sum is only a minor obstacle). Therefore, the additional integral term in (3.2) reflects the dependence between $X_t(i)$'s.

Remark 3.5. An important feature of our result is the factorisation of the fluctuations of the total mass process $\{|X_t|\}_t$ and the spatial fluctuations process i.e. $\{\langle X_t, f\rangle\}_t$. Using this fact we can easily prove a central limit theorem corresponding to (3.1) with deterministic normalization. Namely

$$e^{-\lambda_p t/2} \left(\langle X_t, f\rangle - |X_t| \langle \varphi, f\rangle\right) \to^d G_2 \sqrt{V_\infty},$$

where G_2 are the same as in Theorem 3.3 and V_∞ is given by (2.1).

Remark 3.6. The convergence of the spatial fluctuations can also be regarded as convergence of random fields. It is technically convenient to embed the space of point measures into the space of tempered distributions $\mathcal{S}'(\mathbb{R}^d)$ (i.e. the dual of the space of rapidly decreasing functions $\mathcal{S}(\mathbb{R}^d)$). The following $\mathcal{S}'(\mathbb{R}^d)$-valued random variable:

$$M_t := \frac{X_t - |X_t|\varphi(x)\mathrm{d}x}{|X_t|^{1/2}} = |X_t|^{1/2} \left(\frac{X_t}{|X_t|} - \varphi\right), \tag{3.4}$$

converges to a Gaussian random field M with covariance structure given by

$$\mathrm{Cov}(\langle M, f_1\rangle, \langle M, f_2\rangle) = \left\langle \varphi, \tilde{f}_1 \tilde{f}_2 \right\rangle + 2\lambda p \int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s} \mathcal{T}_s \tilde{f}_1\right) \left(e^{(\lambda_p/2)s} \mathcal{T}_s \tilde{f}_2\right) \right\rangle \mathrm{d}s, \tag{3.5}$$

where $\tilde{f}_i(x) = f_i(x) - \langle \varphi, f_i\rangle$. The convergence holds in $\mathcal{S}'(\mathbb{R}^d)$ space. In order to justify it one first applies the Mitoma theorem [22]. This theorem states that the tightness of $\{M_t\}_{t\geq 0}$ is equivalent to tightness of $\langle M_t, f\rangle$ for all $f \in \mathcal{S}(\mathbb{R}^d)$, which obviously holds by Theorem 3.3. Let us thus consider any convergent subsequence M_{t_n} and denote by M its weak limit. Note that by Theorem 3.3 and the Cramer-Wold device for all $f_1, \ldots, f_m \in \mathcal{S}(\mathbb{R}^d)$, the finite dimensional marginals $(\langle M_{t_n}, f_1\rangle, \ldots, \langle M_{t_n}, f_m\rangle)$ converge to a Gaussian vector with the covariance structure given by (3.5). This implies that the distribution of M on the cylindrical σ-field of $\mathcal{S}'(\mathbb{R}^d)$ is uniquely determined. However, due to the separability of the Schwartz class $\mathcal{S}(\mathbb{R}^d)$, the cylindrical σ-field of $\mathcal{S}'(\mathbb{R}^d)$ coincides with the Borel σ-field (see the remark following Definition 2.3.1 in [20]) and so the distribution of M is uniquely determined. Together with tightness this ends the proof of convergence of M_t as random elements of $\mathcal{S}'(\mathbb{R}^d)$.

Remark 3.7. The formula (3.2) for the variance of the limiting Gaussian variable becomes perhaps less cryptic when expressed in the basis of the Hermite polynomials. Let $\{h_n\}_{n\geq 0}$ be the probabilistic Hermite polynomials (see. e.g. [19, page 5]). We recall that in the case of $d = 1$ and $\sigma^2 = 2$, $\mu = 1$ they are eigenfunctions of the Ornstein-Uhlenbeck semigroup, more precisely $\mathcal{T}_t h_n = \exp(-nt) h_n$. Moreover they form an orthonormal basis in the space $L^2(\varphi)$. Still for $d = 1$, but for general σ^2 and μ, let $\tilde{h}_n(x) = h_n(\sqrt{2\mu}x/\sigma)$. By an appropriate change of variables one easily obtains that $\{\tilde{h}_n\}_{n\geq 0}$ is an orthonormal basis in $L^2(\varphi)$ and $\mathcal{T}_t \tilde{h}_n = \exp(-\mu n t)\tilde{h}_n$. Passing to general d and using the product structure of the Ornstein-Uhlenbeck process in $\mathbb{R}^d$ one obtains that the family $\{\tilde{h}_{i_1,\ldots,i_d}\}_{i_1,\ldots,i_d=0}^{\infty}$ given by

$$\tilde{h}_{i_1,\ldots,i_d} = (x_1, \ldots, x_d) = \tilde{h}_{i_1}(x_1) \cdots \tilde{h}_{i_d}(x_d), \tag{3.6}$$

is an orthonormal basis in $L^2(\varphi)$ and $\mathcal{T}_t\tilde{h}_{i_1,\ldots,i_d} = \exp(-\mu(i_1+\ldots+i_d)t)\tilde{h}_{i_1,\ldots,i_d}$. The elements of this basis are often referred to as multivariate Hermite polynomials. If we now define $f_{i_1,i_2,\ldots,i_d} := \int_{\mathbb{R}^d} \tilde{f}(x)\tilde{h}_{i_1,\ldots,i_d}(x)\varphi(x)\mathrm{d}x$, then (3.2) together with basic Hilbert space theory yield

$$\sigma_f^2 := \sum_{i_1=0,i_2=0,\ldots,i_d=0}^{+\infty} f^2_{i_1,i_2,\ldots,i_d}\left(1+\frac{2\lambda p}{2(i_1+i_2+\ldots+i_d)\mu-\lambda_p}\right). \tag{3.7}$$

3.2 Critical branching rate: $\lambda_p = 2\mu$

We denote

$$\sigma_f^2 := \frac{\lambda p\sigma^2}{\mu}\sum_{i=1}^{d}\left\langle \varphi, \frac{\partial f}{\partial x_i}\right\rangle^2. \tag{3.8}$$

Note that the same symbol σ_f^2 has already been used to denote the asymptotic variance in the small branching case. However, since these cases will always be treated separately, this should not lead to ambiguity. Analogously as in (3.7) we can expresses σ_f^2 nicely using the Hermite expansion:

$$\sigma_f^2 = 2\lambda p\left(f^2_{1,0,\ldots,0}+f^2_{0,1,\ldots,0}+f^2_{0,0,\ldots,1}\right). \tag{3.9}$$

Let us also recall (2.1) and that W is V_∞ conditioned on Ext^c. The main result of this section is

Theorem 3.8. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system starting from $x\in\mathbb{R}^d$. Let us assume that $\lambda_p = 2\mu$ and $f\in\mathcal{P}(\mathbb{R}^d)$. Then $\sigma_f^2 < +\infty$ and conditionally on the set of non-extinction Ext^c there is the convergence*

$$\left(e^{-\lambda_p t}|X_t|, \frac{|X_t|-e^{t\lambda_p}V_\infty}{\sqrt{|X_t|}}, \frac{\langle X_t, f\rangle - |X_t|\langle\varphi, f\rangle}{t^{1/2}\sqrt{|X_t|}}\right)\to^d (W, G_1, G_2),$$

where $G_1\sim\mathcal{N}(0, 1/(2p-1)), G_2\sim\mathcal{N}(0,\sigma_f^2)$ and W, G_1, G_2 are independent random variables.

Remark 3.9. We continue the discussion from Remark 3.4. This time the theorem is less classical as the normalization is larger. We interpret this as the fact that $X_t(i)$'s become more dependent as λ_p increases relatively to μ. In other words the branching is so fast that any particle has many relatives which are still close to it. This also explains, at least on an intuitive level, the appearance of the derivative in (3.8).

3.3 Large branching rate: $\lambda_p > 2\mu$

Let us introduce the process $\{H_t\}_{t\geq 0}$

$$H_t := e^{(-\lambda_p+\mu)t}\sum_{i=1}^{|X_t|} X_t(i). \tag{3.10}$$

Let us also recall the martingale $V_t := e^{-\lambda_p t}|X_t|$. We have

Proposition 3.10. *The process $\{H_t\}_{t\geq 0}$ is a martingale with respect to the filtration of the OU branching system. Moreover, if $\lambda_p > 2\mu$, then we have $\sup_{t\geq 0}\mathbb{E}_x\|H_t\|^2 < +\infty$. Thus, there exists*

$$H_\infty := \lim_{t\to+\infty} H_t,$$

where the convergence holds a.s. and in L^2. When the OU branching system starts from 0, the martingales V_t and H_t are orthogonal.

The distribution of H_∞ depends on the starting conditions.

Proposition 3.11. *Let $\{X_t\}_{t\geq 0}$ and $\{\tilde{X}_t\}_{t\geq 0}$ be two OU branching systems, the first one starting from 0 and the second one from x. Let us denote the limit of the corresponding martingales by $H_\infty, \tilde{H}_\infty$ respectively. Then*

$$\tilde{H}_\infty =^d H_\infty + xV_\infty,$$

where V_∞ is given by (2.1) for the system X.

Let us denote by J the random variable H_∞ conditioned on Ext^c.

Theorem 3.12. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system starting from $x \in \mathbb{R}^d$. Let us assume that $\lambda_p > 2\mu$ and $f \in \mathcal{P}(\mathbb{R}^d)$. Then conditionally on the set of non-extinction Ext^c there is the convergence*

$$\left(e^{-\lambda_p t}|X_t|, \frac{|X_t| - e^{t\lambda_p} V_\infty}{\sqrt{|X_t|}}, \frac{\langle X_t, f\rangle - |X_t| \langle \varphi, f\rangle}{\exp\left((\lambda_p - \mu)t\right)} \right) \to^d (W, G, \langle \varphi, \nabla f\rangle \circ J), \tag{3.11}$$

where $G \sim \mathcal{N}(0, 1/(2p-1))$ and $(W, J), G$ are independent. Moreover

$$\left(e^{-\lambda_p t}|X_t|, \frac{\langle X_t, f\rangle - |X_t| \langle \varphi, f\rangle}{\exp\left((\lambda_p - \mu)t\right)} \right) \to (V_\infty, \langle \varphi, \nabla f\rangle \circ H_\infty), \quad \textit{in probability.} \tag{3.12}$$

Remark 3.13. As we noted in the Introduction this case hardly resembles the classical CLT. The convergence in probability is perhaps its most unexpected feature. This phenomenon seems to be closely related to the fact that the branching is so fast that the system is “not able to forget” the starting condition and in fact, up to some degree, it “remembers” its whole evolution (encoded in the martingale H).

Remark 3.14. We were not able to derive any explicit formula for the law of H_∞. However we calculated some of its moments (for $d = 1$). As the formulas become lengthy, we assume that $p = 1$ and present only:

$$\mathbb{E}_0 H_\infty^2 = 2\gamma, \quad \mathbb{E}_0 H_\infty^4 = \frac{96\gamma^2\left(16 + 39\gamma + 30\gamma^2 + 8\gamma^3\right)}{9 + 27\gamma + 26\gamma^2 + 8\gamma^3},$$

$$\mathbb{E}_0 H_\infty^6 = \frac{1440\gamma^3\left(36847 + 285675\gamma + 948012\gamma^2 + 1760420\gamma^3 + 2005408\gamma^4 + 1441120\gamma^5 + 642112\gamma^6 + 163584\gamma^7 + 18432\gamma^8\right)}{(1+\gamma)^2(3+2\gamma)(5+4\gamma)(5+6\gamma)(5+8\gamma)\left(6 + 17\gamma + 12\gamma^2\right)},$$

where $\gamma = (\lambda_p/\mu - 2)^{-1}$. The odd moments are 0 as the distribution is symmetric. One can now check that H_∞ is never Gaussian. Moreover, V_∞ and H_∞ are not independent (even though uncorrelated). Their dependence is not trivial. By similar calculations of the fourth moment of $xV_\infty + H_\infty$ we also checked that H_∞ is not of the form $V_\infty G$, where G is some random variable (not necessarily normal) independent of V_∞.

Remark 3.15. We suspect that the convergence in (3.12) is in fact almost sure. Similarly as in the case of the law of large numbers we expect that the methods of [13, Lemma 18] might be of use. Let us note however that in this case situation is much more delicate since the norming is much smaller. This potentially gives raise to additional technical problems.

General remarks

Now we will present general remarks common for all cases.

Remark 3.16. We would like to draw attention to an extension of our work with appeared in [24]. The authors consider a branching system with particles moving according to a Markov process fulfilling certain spectral properties. The particles branch according to a space-dependant branching law (with uniformly bounded second moment) and bounded branching intensity. Under this quite general conditions they provide CLT result akin to ours. Notably, they use the spectral theory to describe the limits. This, in particular allows, to treat cases for which our results are not sharp (e.g. in the case $\lambda_p \geq 2\mu$ the limit in our result is degenerated for functions orthogonal to Hermite polynomials of degree at most one and in this case one can find a different normalization yielding a non-trivial limit).

Remark 3.17. A result close to ours was presented in [6]. The authors consider a model more general then ours in dimension $d = 1$. Particles move on a real line according to some diffusion process. Moreover, each particle carries some mass. Upon a split event the offspring are assigned new positions and masses in some, possibly random, fashion depending on the position and mass of the ancestor.

For the sake of notational simplicity, we skip the full generality, and sketch the results of [6] on the example of the Ornstein-Uhlenbeck branching process X studied in our paper. We define a measure-valued processes $\left\{\eta_t^T\right\}_{t\geq 0}$ by

$$\left\langle \eta_t^T, f\right\rangle := e^{\lambda_p(t+T)/2}\left(\frac{\langle X_{t+T}, f\rangle}{e^{\lambda_p(t+T)}} - \frac{\langle X_T, \mathcal{T}_t f\rangle}{e^{\lambda_p T}}\right).$$

In a carefully chosen functional space they prove [6, Proposition 6.4] that $\eta^T \to^d \eta$, as $T \to \infty$, where η is a solution of the following stochastic equation

$$\langle \eta_t, f\rangle = \int_0^t \int_{\mathbb{R}} (Lf(x) + K\eta_s(\mathrm{d}x))\mathrm{d}s + \sqrt{W}\mathcal{W}_t(f),$$

where W is an analogue of our V_∞, $\mathcal{W}$ is a certain Gaussian martingale with values in a functional space and K is a certain operator related to the branching. As already mentioned in the Introduction their approach differs from ours. They always consider two finite times $t, t+T$ while we examine fluctuations between some finite time t and the limit at infinity (recall the third coordinate in (3.3) and that by Theorem 3.1 for T large $\langle X_T, f\rangle$ is of the same order as $|X_T|\langle\varphi, f\rangle$). Our result is more classical in a sense that it studies fluctuations from the limiting object. While theirs diverges from this schema, it captures also some temporal aspects of the convergence.

Remark 3.18. Now we list other possible extensions of our results. The first has already been obtained, a parallel paper [1] contains the corresponding results for U-statistics (as it was described in the Introduction). Secondly, corresponding results were also obtained for superprocesses in [21]. Another natural extension would be to consider more general branching laws. The first problem to consider is a law of large numbers. The statement of the Kesten-Stigum theorem suggest the weakest possible conditions and the Seneta-Heyde theorem suggests further extensions (we refer the reader to [7] and references therein).

The paper [24] treated a very general case of systems allowing space-dependent branching law under the natural condition of the existence of the second moment. Moreover, this paper assumed that the branching intensity is bounded. We suspect that the last condition can be relaxed (e.g. it is known, see [13], that the law of large numbers holds for certain unbounded branching intensity functions).

It is natural to expect that there are analogues of our results for systems with the branching laws with infinite variance. One should expect different normalisation and a stable law in the limit.

Other interesting lines of research are the large deviation principle and functional convergence.

4 Proofs

4.1 Laplace transform and moments

Following [11] the branching system model is formalised as a Markov process in the space $\mathcal{M}$ of integer-valued measures on $\mathbb{R}^d$, endowed with the σ-field Σ generated by all the functions of the form $\mathcal{M} \ni \nu \mapsto \nu(A)$, where A is a Borel subset of $\mathbb{R}^d$. By the assumption that the evolution of distinct particles is independent, to describe the transition probabilities it is enough to consider systems starting from a single particle. By [11, (1.3)], its Laplace transform

$$w(x,t,\theta) := \mathbb{E}_x \exp\left(-\langle X_t, \theta f\rangle\right), \tag{4.1}$$

where $\theta \geq 0$, $f\colon \mathbb{R}^d \to \mathbb{R}_+$ is bounded and measurable, satisfies the following equation

$$w(x,t,\theta) = e^{-\lambda t}\mathcal{T}_t e^{-\theta f}(x) + \lambda \int_0^t e^{-\lambda(t-s)}\mathcal{T}_{t-s}F(w(\cdot,s,\theta))(x)\mathrm{d}s, \tag{4.2}$$

where F is the generating function of the branching law, given by

$$F(s) := ps^2 + (1-p) \tag{4.3}$$

and λ is the branching intensity. We note that we consider a time homogenous transition probability hence our equation is a little bit simpler that the one in [11] (e.g. we do not need the parameter r). Moreover, the additive functional K of [11] is simply $K(s,t) = \lambda(t-s)$.

Transition probabilities for a general point measure ν are determined by the following rule, which 'encodes' the independence of subsystems starting from distinct particles

$$\mathbb{E}_\nu \exp(-\langle \theta f, X_t\rangle) = \prod_{i=1}^n \mathbb{E}_{\nu_i} \exp(-\langle \theta f, X_t\rangle), \tag{4.4}$$

whenever $\nu = \sum_{i=1}^n \nu_i$ (see equation (1.1) in [11]).

We will now use the trivial identity $e^{-\lambda t}\mathcal{T}_t f = \mathcal{T}_t f - \int_0^t \mathcal{T}_s(\lambda e^{-\lambda(t-s)}\mathcal{T}_{t-s}f)\mathrm{d}s$ and the Fubini theorem to transform (4.2). We have

$$\begin{aligned}
w(x,t,\theta) &= e^{-\lambda t}\mathcal{T}_t e^{-\theta f}(x) + \lambda \int_0^t e^{-\lambda s}\mathcal{T}_s F(w(\cdot,t-s,\theta))(x)\mathrm{d}s \\
&= e^{-\lambda t}\mathcal{T}_t e^{-\theta f}(x) + \lambda \int_0^t \mathcal{T}_s F(w(\cdot,t-s,\theta))(x)\mathrm{d}s \\
&\quad - \lambda \int_0^t \left(\int_0^s \mathcal{T}_u \lambda e^{-\lambda(s-u)}\mathcal{T}_{s-u}F(w(\cdot,t-s,\theta))(x)\mathrm{d}u\right)\mathrm{d}s.
\end{aligned}$$

Applying the Fubini theorem to the third summand and using (4.2) we get

$$\begin{aligned}
&\lambda \int_0^t \mathcal{T}_u \left(\lambda \int_u^t e^{-\lambda(s-u)}\mathcal{T}_{s-u}F(w(\cdot,t-s,\theta))(x)\mathrm{d}s\right)\mathrm{d}u \\
&\quad = \lambda \int_0^t \mathcal{T}_u \left(\lambda \int_0^{t-u} e^{-\lambda s}\mathcal{T}_s F(w(\cdot,t-u-s,\theta))(x)\mathrm{d}s\right)\mathrm{d}u \\
&\quad = \lambda \int_0^t \mathcal{T}_u \left(w(\cdot,t-u,\theta) - e^{-\lambda(t-u)}\mathcal{T}_{t-u}e^{-\theta f}\right)(x)\mathrm{d}u \\
&\quad = e^{-\lambda t}\mathcal{T}_t e^{-\theta f}(x) - \mathcal{T}_t e^{-\theta f}(x) + \lambda \int_0^t \mathcal{T}_u w(\cdot,t-u,\theta)(x)\mathrm{d}u.
\end{aligned}$$

Putting the last two expressions together we conclude that w fulfils

$$w(x,t,\theta) = \mathcal{T}_t e^{-\theta f}(x) + \lambda \int_0^t \mathcal{T}_{t-s}\left[p w^2(\cdot,s,\theta) - w(\cdot,s,\theta) + (1-p)\right](x)\mathrm{d}s, \tag{4.5}$$

Let now $w^{(k)}$ denote the k-th derivative of w with respect to θ. By differentiating (4.5) we get for $k \geq 1$,

$$\begin{aligned} w^{(k)}(x,t,\theta) =& (-1)^k \mathcal{T}_t\left[f^k(\cdot)e^{-\theta f}(\cdot)\right](x) \\ &+ \lambda \int_0^t \mathcal{T}_{t-s}\left[p\sum_{l=0}^{k}\binom{k}{l} w^{(l)}(\cdot,s,\theta)w^{(k-l)}(\cdot,s,\theta) - w^{(k)}(\cdot,s,\theta)\right](x)\mathrm{d}s. \end{aligned}$$

Note that this differentiation is valid by Proposition 2.1 and properties of the Laplace transform (e.g. [16, Chapter XIII.2]). We evaluate this expression at $\theta = 0$ and write simply $w^{(k)}(x,t)$ instead of $w^{(k)}(x,t,0)$. Further we use the fact that $w(x,t) = 1$ to get

$$\begin{aligned} w^{(k)}(x,t) =& (-1)^k \mathcal{T}_t f^k(x) \\ &+ \lambda \int_0^t \mathcal{T}_{t-s}\left[p\sum_{l=1}^{k-1}\binom{k}{l} w^{(l)}(\cdot,s)w^{(k-l)}(\cdot,s) + (2p-1)w^{(k)}(\cdot,s)\right](x)\mathrm{d}s. \end{aligned}$$

We recall that $\lambda_p = (2p-1)\lambda$. By a simple variation of the argument leading to (4.5) one checks that

$$w^{(k)}(x,t) = (-1)^k e^{\lambda_p t}\mathcal{T}_t f^k(x) + \lambda \int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[p\sum_{l=1}^{k-1}\binom{k}{l} w^{(l)}(\cdot,s)w^{(k-l)}(\cdot,s)\right](x)\mathrm{d}s. \tag{4.6}$$

By the properties of the Laplace transform the moments are given by

$$\mathbb{E}_x\left(\langle X_t, f\rangle\right)^n = (-1)^n w^{(n)}(x,t). \tag{4.7}$$

The equation (4.2), and therefore our equations so far, are valid for f being positive and bounded. Using standard measure-theoretic techniques one can show that (4.6) still holds for $f \in \mathcal{P}$ if we use (4.7) as the definition of $w^{(n)}(x,t)$ (note that for general $f \in \mathcal{P}$ the Laplace transform may not be well defined). In order to do that one has to know that the moments are finite. This follows by Hölder's inequality, conditioning and Proposition 2.1, viz.

$$\mathbb{E}_x|\langle X_t, f\rangle|^n \leq \mathbb{E}_x|X_t|^{n-1}\langle X_t, |f|^n\rangle = (\mathbb{E}|X_t|^n)\mathcal{T}_t|f|^n(x) < \infty,$$

4.2 Weak convergence facts

In this section we gather simple facts concerning weak convergence. Let us denote by $\|\cdot\|_{TV}$ the total variation norm on the set of probability measures. We have a simple lemma

Lemma 4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A_1 \subset A_2 \in \mathcal{F}$ be such that $\mathbb{P}(A_1) > 0$. Let X be a random variable and ν_1, ν_2 be its law conditioned on A_1, A_2 respectively. Then*

$$\|\nu_1 - \nu_2\|_{TV} \leq 2\frac{\mathbb{P}(A_2) - \mathbb{P}(A_1)}{\mathbb{P}(A_1)}.$$

Proof. Let B be a Borel set, then

$$|\mathbb{P}(X \in B|A_1) - \mathbb{P}(X \in B|A_2)| = \Big|\frac{\mathbb{P}(\{X \in B\} \cap A_1)}{\mathbb{P}(A_1)} - \frac{\mathbb{P}(\{X \in B\} \cap A_2)}{\mathbb{P}(A_2)}\Big|$$
$$= \Big|\frac{\mathbb{P}(\{X \in B\} \cap A_1)(\mathbb{P}(A_2) - \mathbb{P}(A_1)) + (\mathbb{P}(\{X \in B\} \cap A_1) - \mathbb{P}(\{X \in B\} \cap A_2))\mathbb{P}(A_1)}{\mathbb{P}(A_1)\mathbb{P}(A_2)}\Big|$$
$$\leq 2\frac{\mathbb{P}(A_2) - \mathbb{P}(A_1)}{\mathbb{P}(A_1)}.$$

Taking supremum over B we obtain the lemma. □

Let μ_1, μ_2 be two probability measures on $\mathbb{R}$, and $\mathrm{Lip}(1)$ be the space of continuous functions $\mathbb{R} \mapsto [-1, 1]$ with the Lipschitz constant smaller or equal to 1. We define

$$m(\mu_1, \mu_2) := \sup_{g \in \mathrm{Lip}(1)} |\langle \mu_1, g\rangle - \langle \mu_2, g\rangle|. \tag{4.8}$$

It is well known that m is a distance metrizing the weak convergence (see e.g. [10, Theorem 11.3.3]). One easily checks that when μ_1, μ_2 are, respectively, the laws of random variables X_1, X_2 on the same probability space then we have

$$m(\mu_1, \mu_2) \leq \|X_1 - X_2\|_1 \leq \sqrt{\|X_1 - X_2\|_2}. \tag{4.9}$$

4.3 Couplings

The aim of this section is to define a coupling of two OU-branching processes, which will be the main tool in proofs of the limit theorems.

We begin by defining a coupling of two Ornstein-Uhlenbeck processes. Let us first recall that the Ornstein-Uhlenbeck process starting from $x \in \mathbb{R}^d$ is a solution to the stochastic differential equation

$$\mathrm{d}\eta_t = \sigma \mathrm{d}\beta_t - \mu\eta_t \mathrm{d}t, \quad \eta_0 = x,$$

where $(\beta_t)_{t\geq 0}$ is a standard d-dimensional Brownian motion. This property (sometimes taken as the definition of the Ornstein-Uhlenbeck process) is an immediate consequence of the relation between the infinitesimal operator of a diffusion and the coefficients of the defining SDE (see e.g. Theorem 7.3.3. in [23]). The following coupling will be very useful for further analysis.

Proposition 4.2. *Let $x, y \in \mathbb{R}^d$ and let $\{\gamma_t\}_{t\geq 0}$ be an $\mathbb{R}^{2d}$-valued process defined by*

$$\gamma_t := (\eta_t^1, \eta_t^2),$$

where η^1, η^2 are Ornstein-Uhlenbeck processes starting from x, y respectively and satisfying $d\eta_t^i = \sigma d\beta_t - \mu\eta_t^i dt$ with the same d-dimensional Brownian motion β. Then γ is a strong Markov process, moreover

$$\eta_t^1 - \eta_t^2 = (x - y)e^{-\mu t}, \quad \text{a.s.}$$

Proof. The strong Markov property follows from general facts about diffusions (see e.g. Theorem 7.2.4 in [23]). The second part of the proposition follows by subtraction. Namely,

$$\mathrm{d}(\eta_t^1 - \eta_t^2) = -\mu(\eta_t^1 - \eta_t^2)\mathrm{d}t, \quad \eta_0^1 - \eta_0^2 = x - y,$$

which implies the equality in question. □

Now we transfer the coupling to the level of branching systems.

Proposition 4.3. *Let $\{X_t\}_{t\geq 0}$ be an OU-branching particle system starting from $x \in \mathbb{R}^d$. Let $y \in \mathbb{R}^d$ and define a new system $\{Y_t\}_{t\geq 0}$ such that $|Y_t| = |X_t|$ and*

$$Y_t(i) := (y - x)e^{-\mu t} + X_t(i),$$

for any particle i of the system. Then Y is an OU-branching particle system starting from y.

Although the above proposition is intuitively clear, its formal proof is quite lengthy and hence is deferred to the Appendix.

4.4 Rate of convergence to the invariant measure and approximations

We will need also estimates of the speed of convergence to the invariant measure. Recall the definition of $\tilde{f}$ given by (2.6). In proofs below it will be convenient to have some additional regularity conditions. Therefore given $f \in \mathcal{P}$ and $u > 0$ we define

$$l_u(x) = \mathcal{T}_u \tilde{f}(x). \tag{4.10}$$

The reason for introducing the function l_u is that it is smooth and for $u \to 0$ it approximates the function $\tilde{f}$. In the following simple lemma we gather a couple of properties of the function l_u which later will be used to show that the approximation carries on to the level of branching systems.

Lemma 4.4. *If $f \in \mathcal{P}(\mathbb{R}^d)$ and $u > 0$, then l_u given by (4.10) is a $C^\infty(\mathbb{R}^d)$ function. Moreover, $\langle \varphi, l_u \rangle = 0$, and there exist $C, n > 0$ such that*

$$\left| \frac{\partial l_u}{\partial x_i} \right| \leq C(1 + \|x\|)^n, \quad \left| \frac{\partial^2 l_u}{\partial x_i \partial x_j} \right| \leq C(1 + \|x\|)^n,$$

for any $i, j \in \{1, 2, \ldots, d\}$.

Proof. The first two statements are trivial and left to the reader. To prove the bound for derivatives we recall (2.4) and (2.5) and write

$$\left| \frac{\partial l_u}{\partial x_i}(x) \right| = \left| \frac{\partial}{\partial x_i} \int_{\mathbb{R}^d} f(y) g_u(x_u - y) \mathrm{d}y \right| \leq \int_{\mathbb{R}^d} |f(y)| \left| \frac{\partial}{\partial x_i} g_u(x_u - y) \right| \mathrm{d}y.$$

One checks that $\left| \frac{\partial}{\partial x_i} g_u(x_u - y) \right| \leq C_1 \exp(-C_2 \|x_u - y\|)$ for some $C_1, C_2 > 0$. Using this together with the assumption $f \in \mathcal{P}$ yields the first estimate. The second one goes along the same lines. □

Lemma 4.5. *Let $f \in \mathcal{P}(\mathbb{R}^d)$ and $\tilde{f}$ be defined as above. Then there exist constants $C, n > 0$ such that for any $t \geq 0$ we have*

$$|\mathcal{T}_t \tilde{f}(x)| \leq C(1 + \|x\|^n)e^{-\mu t}, \quad |\mathcal{T}_t \tilde{f}(0)| \leq Ce^{-2\mu t}. \tag{4.11}$$

Moreover

$$\lim_{t \to +\infty} e^{\mu t} \mathcal{T}_t \tilde{f}(x) = x \circ \langle \varphi, \nabla f \rangle,$$

where ∇f is understood in a weak sense. There exist $\tilde{C}, \tilde{n} > 0$ such that for any $t \geq 0$ we have

$$\left| e^{\mu t} \mathcal{T}_t \tilde{f}(x) - x \circ \langle \varphi, \nabla f \rangle \right| \leq \tilde{C}(1 + \|x\|^{\tilde{n}})e^{-\mu t}. \tag{4.12}$$

Moreover, there exists a function $c : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that $c(u) \to 0$ as $u \searrow 0$ and for any $t \geq 0$ we have

$$\left| \mathcal{T}_t (l_u - \tilde{f})(x) \right| \leq c(u)(1 + \|x\|^n)e^{-\mu t}. \tag{4.13}$$

Proof. Let us recall equation (2.5) and that $\langle \varphi, l_u \rangle = 0$ and $|1 - ou(t)| \lesssim e^{-2\mu t}$. We write

$$e^{\mu(t+u)}\mathcal{T}_{t+u}\tilde{f}(x) = e^{\mu u}e^{\mu t}\mathcal{T}_t l_u(x) = e^{\mu u}\int_{\mathbb{R}^d} e^{\mu t}(l_u(xe^{-\mu t} + ou(t)y) - l_u(y))\varphi(y)\mathrm{d}y. \quad (4.14)$$

Using the mean value theorem we get

$$h(x, y, t) := e^{\mu t}(l_u(xe^{-\mu t} + ou(t)y) - l_u(y)) = (x + e^{\mu t}(ou(t) - 1)y) \circ \nabla l_u(x_0), \quad (4.15)$$

where x_0 is some point on the interval joining y and $xe^{-\mu t} + ou(t)y$. From this representation and Lemma 4.4 we obtain $|h(x, y, t)| \lesssim 1 + \max(\|x\|^n, \|y\|^n)$ and $|h(0, y, t)| \lesssim e^{-\mu t}(1 + \|y\|^n)$. This is enough to show (4.11) for $t \geq 1$ by setting e.g. $u = 1/2$. Note that for $t < 1$, (4.11) follows easily from the fact that $\mathcal{T}_t f$ is again of polynomial growth.

We notice that $h(x, y, t) \to x \circ \nabla l_u(y)$ point-wise as $t \to +\infty$. This, together with the Lebesgue dominated convergence yields

$$I := \lim_{t\to+\infty} e^{\mu t}\mathcal{T}_t\tilde{f}(x) = e^{\mu u}\sum_{i=1}^d x_i \left\langle \varphi, \frac{\partial l_u}{\partial y_i} \right\rangle = -e^{\mu u}\sum_{i=1}^d x_i \left\langle l_u, \frac{\partial \varphi}{\partial y_i} \right\rangle.$$

The calculations above are valid for any choice of u, in particular, letting $u \searrow 0$ we obtain

$$I = -\sum_{i=1}^d x_i \left\langle \tilde{f}, \frac{\partial \varphi}{\partial y_i} \right\rangle = \sum_{i=1}^d x_i \left\langle \frac{\partial f}{\partial y_i}, \varphi \right\rangle = x \circ \langle \varphi, \nabla f \rangle.$$

For the sake of notational simplicity (4.12) will be proved for $d = 1$. By the above calculations

$$e^{\mu(t+u)}\mathcal{T}_{t+u}\tilde{f}(x) - x\langle \varphi, f' \rangle = e^{\mu u}\left(e^{\mu t}\mathcal{T}_t l_u(x) - x\langle \varphi, l'_u \rangle\right).$$

Let us now treat the inner expression using (2.5):

$$\begin{aligned} e^{\mu t}\mathcal{T}_t l_u(x) - x\langle \varphi, l'_u \rangle &= \mathbb{E}\left(e^{\mu t}\left(l_u(xe^{-\mu t} + ou(t)G) - l_u(G)\right) - xl'_u(G)\right) \\ &= \mathbb{E}\left(e^{\mu t}\int_G^{xe^{-\mu t}+ou(t)G} l'_u(y)\mathrm{d}y - xl'_u(G)\right) \\ &= \mathbb{E}\left(e^{\mu t}\int_G^{xe^{-\mu t}+ou(t)G} (l'_u(y) - l'_u(G))\mathrm{d}y\right) + e^{\mu t}(ou(t) - 1)\mathbb{E}Gl'_u(G). \end{aligned}$$

By Lemma 4.4 the second term is easily bounded from above by $Ce^{-\mu t}$, for some $C > 0$. The first one can be estimated as

$$II := \left|\mathbb{E}e^{\mu t}\int_G^{xe^{-\mu t}+ou(t)G}\int_G^y l''_u(z)\mathrm{d}z\mathrm{d}y\right| \leq \mathbb{E}e^{\mu t}(xe^{-\mu t}+(ou(t)-1)G)^2 \sup_{z\in[G,xe^{-\mu t}+ou(t)G]} |l''_u(z)|.$$

By Lemma 4.4 we know that l''_u grows polynomially, so there exists n such that

$$II \lesssim e^{-\mu t}\mathbb{E}(x + e^{\mu t}(ou(t) - 1)G)^2(\|x\| + \|G\|)^n \lesssim e^{-\mu t}(1 + \|x\|^{n+2}).$$

This is enough to prove (4.12), again by setting e.g. $u = 1/2$ (as with (4.11), the case $t < 1$ is straightforward).

Now, we need an estimate on $e^{\mu t}\mathcal{T}_t(l_u - \tilde{f})(x)$ which takes into account u. For $t > 1$, using the same argument as before it is enough to prove an estimate for $e^{\mu t}\mathcal{T}_t(l_{u+1} - l_1)(x)$. We denote $k_u(x) := l_{u+1}(x) - l_1(x)$ and notice that $\langle \varphi, k_u \rangle = 0$. Further, we recall that $\langle \varphi, \mathcal{T}_t f \rangle = \langle \varphi, f \rangle$ and thus

$$e^{\mu t}\mathcal{T}_t k_u(x) = \int_{\mathbb{R}^d} e^{\mu t}(k_u(xe^{-\mu t} + ou(t)y) - k_u(y))\varphi(y)\mathrm{d}y,$$

where we used that $\langle \varphi, k_u \rangle = 0$. Using the mean value theorem we get

$$h(x,y,t) := e^{\mu t}(k_u(xe^{-\mu t} + ou(t)y) - k_u(y)) = (x + e^{\mu t}(ou(t)-1)y) \circ \nabla k_u(x_0),$$

where x_0 is some point on the interval joining $xe^{-\mu t} + ou(t)y$ and y. Further, following arguments used in (4.14) and (4.15) and after them we conclude that in order to obtain (4.13) we need to estimate $\nabla(l_{u+1} - l_1)$. For simplicity we will provide the details for $d = 1$. We have

$$III := ((l_{u+1} - l_1)(x))' = \left(\int_{\mathbb{R}} (l_1(x-w) - l_1(x)) g_u(w) \mathrm{d}w \right)' = \int_{\mathbb{R}} (l_1'(x-w) - l_1'(x)) g_u(w) \mathrm{d}w.$$

Using the mean value theorem, there exists a function $x_0(x,w)$ such that $x_0(x,w) \in [x-w, x] \cup [x, x-w]$ and

$$III = \int_{\mathbb{R}} w l_1''(x_0(x,w)) g_u(w) \mathrm{d}w.$$

The function l_1'' grows at most polynomially. Therefore, for some c, n we have

$$|III| \leq c \int_{\mathbb{R}^d} |w|(1 + \|x\|^n + \|w\|^n) g_u(w) \mathrm{d}w \leq c_1(u)(1 + \|x\|^n),$$

where $c_1(u)$ is some function fulfilling $c_1(u) \to 0$ as $u \searrow 0$. It remains to prove (4.13) for $t \leq 1$. Since the semigroup $\mathcal{T}_t$ preserves polynomial growth of a function, it is enough to consider the case $t = 0$. Using the easy fact that l_u converges to $\tilde{f}$ uniformly on compact sets and considering balls of increasing radia, one can easily construct a function $c_2(u)$ such that $c_2(u) \to 0$ as $u \to 0$ and $|l_u(x) - \tilde{f}(x)| \leq c_2(u)(1 + \|x\|^n)$ for some $n \in \mathbb{N}$. ☐

4.5 LLN and CLT for small branching rate

First we prove the following

Proposition 4.6. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system and $\lambda_p < 2\mu$ and let $f \in \mathcal{P}$. Then*

$$\mathbb{E}_x \left(e^{-(\lambda_p/2)t} \left\langle X_t, \tilde{f} \right\rangle \right) \to 0, \quad \text{as } t \to +\infty, \tag{4.16}$$

$$\mathbb{E}_x \left(e^{-(\lambda_p/2)t} \left\langle X_t, \tilde{f} \right\rangle \right)^2 \to \sigma_f^2, \quad \mathrm{Var}_x \left(e^{-(\lambda_p/2)t} \left\langle X_t, \tilde{f} \right\rangle \right) \to \sigma_f^2, \quad \text{as } t \to +\infty, \tag{4.17}$$

where σ_f^2 is the same as in (3.2). *Moreover,*

$$\sup_{t\geq 0} \mathbb{E}_x \left(e^{-(\lambda_p/2)t} \left(\left\langle X_t, \tilde{f} \right\rangle - \mathbb{E}_x \left\langle X_t, \tilde{f} \right\rangle \right) \right)^4 < +\infty. \tag{4.18}$$

Proof. We will use the notation from Section 4.1, in particular equations (4.6) and (4.7) (with $\tilde{f}$ instead of f). First we note that by (4.6) and Lemma 4.5 (ineq. (4.11)),

$$|w'(x,t)| \lesssim e^{(\lambda_p - \mu)t}(1 + \|x\|^n) \leq e^{(\lambda_p/2)t}(1 + \|x\|^n),$$

which, by (4.7), implies the first assertion. Using (4.6) and (4.7) again we calculate the second moment

$$\begin{aligned} \mathbb{E}_x \left(e^{-(\lambda_p/2)t} \left\langle X_t, \tilde{f} \right\rangle \right)^2 &= \mathcal{T}_t \tilde{f}^2(x) + 2\lambda p e^{-\lambda_p t} \int_0^t e^{\lambda_p (t-s)} \mathcal{T}_{t-s} \left[\left(e^{\lambda_p s} \mathcal{T}_s \tilde{f}(\cdot) \right)^2 \right](x) \mathrm{d}s \\ &= \mathcal{T}_t \tilde{f}^2(x) + 2\lambda p \int_0^t \mathcal{T}_{t-s} \left[\left(e^{(\lambda_p/2)s} \mathcal{T}_s \tilde{f}(\cdot) \right)^2 \right](x) \mathrm{d}s. \end{aligned}$$

By (4.11) in Lemma 4.5 the integrand in the last expression can be estimated as follows

$$\mathcal{T}_{t-s}\left[\left(e^{(\lambda_p/2)s}\mathcal{T}_s\tilde{f}(\cdot)\right)^2\right](x) \lesssim e^{(\lambda_p-2\mu)s}\mathcal{T}_{t-s}\left[(1+\|\cdot\|^n)^2\right](x).$$

Using (2.4) it can be checked that for any $t \geq 0$ we have $\mathcal{T}_t\left[(1+\|\cdot\|^n)^2\right](x) \lesssim (1+\|x\|^{2n})$. The dominated Lebesgue theorem implies (4.17) (note that the second assertion is an immediate consequence of the first one and (4.16)). We also conclude that for any $t \geq 0$,

$$|w''(x,t)| \lesssim e^{\lambda_p t}(1+\|x\|^{2n}).$$

Similarly we investigate $w'''(x,t)$. By (4.6) we have

$$|w'''(x,t)| \lesssim e^{\lambda_p t}\mathcal{T}_t|\tilde{f}|^3(x) + \int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[w''(\cdot,s)w'(\cdot,s)\right](x)\mathrm{d}s.$$

Using the estimates obtained already and the fact that $\mathcal{T}_t\left[(1+\|\cdot\|^n)^3\right](x) \lesssim (1+\|x\|^{3n})$

$$|w'''(x,t)| \lesssim (1+\|x\|^{3n})e^{\lambda_p t} + e^{\lambda_p t}\int_0^t e^{\lambda_p/2s}\mathcal{T}_{t-s}\left[1+\|\cdot\|^{3n}\right](x)\mathrm{d}s \lesssim e^{(3/2)\lambda_p t}(1+\|x\|^{3n}).$$

Finally, we will also need the fourth moment. By (4.6) and the estimates above we get

$$\begin{aligned}
&\mathbb{E}_x\left(e^{-(\lambda_p/2)t}\left\langle X_t,\tilde{f}\right\rangle\right)^4 \\
&\lesssim e^{-\lambda_p t}\mathcal{T}_t\tilde{f}^4(x) + e^{-2\lambda_p t}\int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[w''(\cdot,s)^2 + w'''(\cdot,s)w'(\cdot,s)\right](x)\mathrm{d}s \\
&\lesssim e^{-\lambda_p t}(1+\|x\|^{4n}) + e^{-\lambda_p t}\int_0^t e^{\lambda_p s}\mathcal{T}_{t-s}\left[(1+\|\cdot\|^{4n})\right](x)\mathrm{d}s \lesssim (1+\|x\|^{4n}).
\end{aligned}$$

It is now easy to get the last assertion of the proposition. □

Now we will prove the representation (3.7). We recall (3.6) and that

$$\mathcal{T}_s\tilde{h}_{i_1,i_2,\ldots,i_d}(x) = e^{-(i_1+i_2+\ldots+i_d)\mu t}\tilde{h}_{i_1,i_2,\ldots,i_d}(x).$$

Therefore

$$\begin{aligned}
\sigma_f^2 &= \left\langle \varphi, \left(\sum_{i_1=0,i_2=0,\ldots,i_d=0}^{+\infty} f_{i_1,i_2,\ldots,i_d}\tilde{h}_{i_1,i_2,\ldots,i_d}\right)^2\right\rangle \\
&+ 2\lambda p\int_0^{+\infty}\left\langle \varphi, \left(e^{(\lambda_p/2)s}\sum_{i_1=0,i_2=0,\ldots,i_d=0}^{+\infty} e^{-\mu(i_1+i_2+\ldots+i_d)s}f_{i_1,i_2,\ldots,i_d}\tilde{h}_{i_1,i_2,\ldots,i_d}(\cdot)\right)^2\right\rangle\mathrm{d}s \\
&= \sum_{i_1=0,i_2=0,\ldots,i_d=0}^{+\infty} f^2_{i_1,i_2,\ldots,i_d} + 2\lambda p\sum_{i_1=0,i_2=0,\ldots,i_d=0}^{+\infty} f^2_{i_1,i_2,\ldots,i_d}\int_0^{+\infty} e^{(\lambda_p-2(i_1+i_2+\ldots+i_d)\mu)s}\mathrm{d}s.
\end{aligned}$$

Now (3.7) follows easily. We are ready for

Proof of Theorem 3.1 (sketch). The proof of [13, Theorem 6] can be checked to hold also for the branching mechanism introduced in our paper. We will show now that the almost sure convergence holds also for bounded functions which are not compactly supported. We decompose $f = f^+ - f^-$, where $f^+(x) = f(x)1_{\{f(x)\geq 0\}}$ and $f^+(x) = -f(x)1_{\{f(x)<0\}}$. It is enough to prove the claim separately for f^+ and f^-, hence we assume that $f \geq 0$. For $n \in \mathbb{N}$ we consider functions $h_n(x) := \min(\max(n-x,0),1)$ and $g_n : \mathbb{R}^d \mapsto [0,1]$ given by

$g_n(x) := h_n(\|x\|)$. By [13, Theorem 6] we know that $e^{-\lambda_p t}\langle X_t, g_n\rangle \to V_\infty \langle \varphi, g_n\rangle\ a.s.$, we know also that $e^{-\lambda_p t}|X_t| \to V_\infty\ a.s.$ Therefore $e^{-\lambda_p t}\langle X_t, (1-g_n)\rangle \to V_\infty \langle \varphi, (1-g_n)\rangle\ a.s.$ Now we estimate

$$\langle X_t, fg_n\rangle \le \langle X_t, f\rangle \le \langle X_t, fg_n\rangle + \|f\|_\infty \langle X_t, (1-g_n)\rangle .$$

Using the previous considerations and the fact that fg_n has compact support we get with probability one

$$\langle \varphi, fg_n\rangle V_\infty \le \liminf_{t\nearrow+\infty} e^{-\lambda_p t}\langle X_t, f\rangle \\ \le \limsup_{t\nearrow+\infty} e^{-\lambda_p t}\langle X_t, f\rangle \le \langle fg_n, \varphi\rangle V_\infty + \|f\|_\infty \langle \varphi, (1-g_n)\rangle V_\infty .$$

To conclude, we observe that $\langle \varphi, fg_n\rangle \to \langle \varphi, f\rangle$ and $\langle \varphi, fg_n\rangle + \|f\|_\infty \langle \varphi, (1-g_n)\rangle \to \langle \varphi, f\rangle$ as $n \to +\infty$.
Let now $f \in \mathcal{P}$ and $\tilde{f} := f - \langle \varphi, f\rangle$. By (4.17) we have

$$\mathbb{E}_x \left(e^{-\lambda_p t}\left\langle X_t, \tilde{f}\right\rangle\right)^2 \to 0.$$

Therefore $e^{-\lambda_p t}\left\langle X_t, \tilde{f}\right\rangle \to^P 0$, and further $e^{-\lambda_p t}\langle X_t, f\rangle - e^{-\lambda_p t}|X_t|\langle \varphi, f\rangle \to^P 0$, which concludes the proof. □

In the proofs below we say "X is asymptotically equivalent to Y" meaning that $X_t - Y_t \to 0$ in probability (equivalently in law) as $t \to +\infty$. Now we are ready for

Proof of Theorem 3.3. Before presenting technical details let us briefly describe the overall strategy of the proof. We will explore the fact that for $t_1 < t_2$, when conditioned on X_{t_1}, the random measure X_{t_2} is a sum of independent random measures corresponding to subsystems originating from distinct particles alive at time t_1. When t_2 is large enough with respect to t_1, thanks to the coupling construction given in Proposition 4.3, we can approximate each such subsystem by a system starting from zero, thus obtaining an approximation of X_{t_2} by a sum of conditionally i.i.d. random variables, which allows for an application of Lindeberg's CLT. Although formally we do not condition on Ext^c but on $\{|X_{t_1}| \neq 0\}$, for $t_1 \to \infty$ these two events become close, which results in a limit theorem with respect to $\mathbb{P}_x(\cdot|Ext^c)$, by Lemma 4.1.

Let us also remark that throughout the proof we will often change the normalization of the random vector we consider. As already mentioned in the Introduction, since $X_t/\exp(\lambda_p t) \to V_\infty$ a.s., weak convergence results with the normalization by $\sqrt{|X_t|}$ and $\exp(\lambda_p t/2)$ can be translated into one another. However for some parts of the proof, due to algebraic or probabilistic reasons one of the normalisations will be more convenient. We will therefore start with the following random vector

$$Z_1(t) := \left(e^{-\lambda_p t}|X_t|, e^{-(\lambda_p/2)t}(|X_t| - e^{\lambda_p t}V_\infty), e^{-(\lambda_p/2)t}\left\langle X_t, \tilde{f}\right\rangle\right). \tag{4.19}$$

Consider $n \in \mathbb{N}$ to be fixed later and let us write

$$Z_1(nt) = \left(e^{-n\lambda_p t}\langle X_{nt}, 1\rangle, e^{-(n\lambda_p/2)t}(|X_{nt}| - e^{n\lambda_p t}V_\infty), e^{-(n\lambda_p/2)t}\sum_{i=1}^{|X_t|}\left\langle X^{i,t}_{(n-1)t}, \tilde{f}\right\rangle\right),$$

where $\left\{X^{i,s}_t\right\}_t$ denotes the subsystem originating from the particle $X_s(i)$.

We know that with probability one, $e^{-\lambda_p t}|X_t| \to V_\infty$. Therefore

$$e^{-n\lambda_p t}|X_{nt}| - e^{-\lambda_p t}|X_t| \to 0\ a.s. \tag{4.20}$$

as $t \to +\infty$. Let us consider the second term

$$
\begin{aligned}
|X_{nt}| - e^{n\lambda_p t} V_\infty &= |X_{nt}| - e^{n\lambda_p t} \lim_{s\to+\infty} e^{-\lambda_p(nt+s)} |X_{nt+s}| \\
&= |X_{nt}| - e^{n\lambda_p t} \lim_{s\to+\infty} e^{-\lambda_p(s+nt)} \sum_{i=1}^{|X_{nt}|} |X_s^{i,nt}| \\
&= |X_{nt}| - \sum_{i=1}^{|X_{nt}|} \lim_{s\to+\infty} e^{-\lambda_p s} |X_s^{i,nt}| = \sum_{i=1}^{|X_{nt}|} \left(1 - V_\infty^i\right),
\end{aligned}
$$

where V_∞^i are independent (conditionally on X_{nt}) copies of V_∞ (note that formally they depend on t, however we will suppress this fact in the notation).

For any $i \le |X_t|$ we define

$$
\tilde{X}_s^{i,t}(j) := X_s^{i,t}(j) - X_t(i) e^{-\mu s}, \quad s \ge 0, \tag{4.21}
$$

where j iterates over all particles of $X_s^{i,t}$ alive at time s.

By Proposition 4.3, the Markov property and the fact that the evolution of subsystems stemming from different particles is conditionally (on X_t) independent we obtain that $\tilde{X}^{i,t}$'s are conditionally i.i.d. OU-branching particle systems starting from 0.

We are going to show the asymptotic behaviour of $Z_1(nt)$ does not change when we replace $X^{i,t}$'s by $\tilde{X}^{i,t}$'s. To this end we write

$$
\begin{aligned}
I(t) &:= \left| e^{-(n\lambda_p/2)t} \sum_{i=1}^{|X_t|} \left\langle X_{(n-1)t}^{i,t}, \tilde{f} \right\rangle - e^{-(n\lambda_p/2)t} \sum_{i=1}^{|X_t|} \left\langle \tilde{X}_{(n-1)t}^{i,t}, \tilde{f} \right\rangle \right| \\
&\le \left| e^{-(n\lambda_p/2)t} \sum_{i=1}^{|X_t|} 1_{\|X_t(i)\| < t} \sum_{j=1}^{|X_{(n-1)t}^{i,t}|} \left(\tilde{f}(X_{(n-1)t}^{i,t}(j)) - \tilde{f}(\tilde{X}_{(n-1)t}^{i,t}(j)) \right) \right| \\
&\qquad + e^{-(n\lambda_p/2)t} \sum_{i=1}^{|X_t|} 1_{\|X_t(i)\| \ge t} \left| \left\langle X_{(n-1)t}^{i,t}, \tilde{f} \right\rangle + \left\langle \tilde{X}_{(n-1)t}^{i,t}, \tilde{f} \right\rangle \right|.
\end{aligned}
$$

At this moment we make an additional assumption that f and consequently $\tilde{f}$ belong to C^1 and that

$$
\|\nabla \tilde{f}\| \lesssim (1 + \|x\|)^k, \tag{4.22}
$$

for some $k \in \mathbb{N}$. By (4.21) and the mean value theorem we have $|\tilde{f}(X_{(n-1)t}^{i,t}(j)) - \tilde{f}(\tilde{X}_{(n-1)t}^{i,t}(j))| 1_{\|X_t(i)\| \le t} \lesssim t e^{-(n-1)\mu t} (1 + \|X_{(n-1)t}^{i,t}(j)\|)^k$. Using the conditional expectation with respect to X_t, (4.7) and $\mathcal{T}_t(1 + \|\cdot\|^k)(x) \lesssim (1 + \|x\|^k)$ we get

$$
\begin{aligned}
\mathbb{E}_x I(t) \lesssim& e^{-(n\lambda_p/2)t} \mathbb{E}_x \sum_{i=1}^{|X_t|} (1 + \|X_t(i)\|)^k t e^{-(n-1)\mu t} e^{(n-1)\lambda_p t} \\
&+ e^{-(n\lambda_p/2)t} \mathbb{E}_x \sum_{i=1}^{|X_t|} 1_{\|X_t(i)\| \ge t} (1 + \|X_t(i)\|)^k e^{(n-1)\lambda_p t} \\
\lesssim& e^{(n\lambda_p/2)t} t e^{-(n-1)\mu t} + e^{(n\lambda_p/2)t} \mathcal{T}_t((1 + \|\cdot\|)^k 1_{\|\cdot\| > t})(x).
\end{aligned}
$$

There exists $n_0 > 0$ such that for any $n > n_0$ we have $n\lambda_p < 2\mu(n-1)$. Using the Schwarz inequality, the second term can be estimated by $\sqrt{\mathcal{T}_t(1 + \|\cdot\|^k)^2(x)} \sqrt{\mathcal{T}_t 1_{\|\cdot\| > t}(x)}$. The Ornstein-Uhlenbeck process has Gaussian marginals with bounded mean and variance therefore $\mathcal{T}_t 1_{\|\cdot\|>t}(x) \lesssim e^{-ct^2}$ for a certain $c > 0$. We may conclude that for $n > n_0$ we have

$$\lim_{t\to+\infty} \mathbb{E}_x I(t) = 0. \tag{4.23}$$

Let us denote $Z_t^i := e^{-((n-1)\lambda_p/2)t} \left\langle \tilde{X}^{i,t}_{(n-1)t}, \tilde{f} \right\rangle$ and $z_t^i := \mathbb{E}_0 Z_t^i$ (we write $\mathbb{E}_0$ to underline the fact that $\tilde{X}^{i,t}$ starts from 0).

By (4.7) and Lemma 4.5 one checks that for $n > n_0$ we have

$$\mathbb{E}_x e^{-(\lambda_p/2)t} \sum_{i=1}^{|X_t|} |z_t^i| = e^{(\lambda_p/2)t} |\mathbb{E}_0 Z_t^1| \lesssim e^{(n\lambda_p/2)t} e^{-2(n-1)\mu t} \to 0, \quad \text{as } t \to +\infty. \tag{4.24}$$

Let us now fix some $n > n_0$. By (4.20), (4.23) and (4.24) and Slutsky's lemma we conclude that $Z_1(nt)$ is asymptotically equivalent to

$$Z_2(t) := \left(e^{-\lambda_p t}|X_t|, e^{-(n\lambda_p/2)t} \sum_{i=1}^{|X_{nt}|} \left(1 - V_\infty^i\right), e^{-(\lambda_p/2)t} \sum_{i=1}^{|X_t|} (Z_t^i - z_t^i) \right).$$

Our next aim is to apply the CLT for independent summands. To this end we define Z_3, which is a "version" of Z_2 with a different (random) normalization. This will be useful when applying standard CLT results since the new normalization corresponds more directly to the number of summands. Let thus

$$Z_3(t) := \left(e^{-\lambda_p t}|X_t|, |X_{nt}|^{-1/2} \sum_{i=1}^{|X_{nt}|} \left(1 - V_\infty^i\right), |X_t|^{-1/2} \sum_{i=1}^{|X_t|} (Z_t^i - z_t^i) \right),$$

which we will consider conditionally on the event $\{|X_t| \neq 0\}$ (here and below we adopt the convention that $a/0 = 0$, the second component of Z_3 is well defined then). The corresponding expected value is denoted by $\mathbb{E}_{x,|X_t|\neq 0}$.

Let us denote the characteristic function of Z_3

$$\chi_1(\theta_1, \theta_2, \theta_3; t) :=$$
$$\mathbb{E}_{x,|X_t|\neq 0} \exp\left\{ i\theta_1 e^{-\lambda_p t}|X_t| + i\theta_2 |X_{nt}|^{-1/2} \sum_{i=1}^{|X_{nt}|} \left(1 - V_\infty^i\right) + i\theta_3 |X_t|^{-1/2} \sum_{i=1}^{|X_t|} (Z_t^i - z_t^i) \right\}.$$

Conditioning on X_{nt} and using the Markov property we check that variables $1 - V_\infty^i$ are i.i.d, moreover they are independent of the system before time nt. We denote their common characteristic function by h.

We have

$$\chi_1(\theta_1, \theta_2, \theta_3; t)$$
$$= \mathbb{E}_{x,|X_t|\neq 0} \exp\left\{ i\theta_1 e^{-\lambda_p t}|X_t| + i\theta_3 |X_t|^{-1/2} \sum_{i=1}^{|X_t|} (Z_t^i - z_t^i) \right\} h\left(\theta_2/\sqrt{|X_{nt}|}\right)^{|X_{nt}|}.$$

By Proposition 2.1 and the central limit theorem for any θ_2 we have $h(\theta_2/\sqrt{m})^m \to e^{-\sigma_V^2 \theta_2^2/2}$ as $m \to \infty$, where $\sigma_V^2 = \frac{1}{2p-1}$. Further we will work with

$$\chi_2(\theta_1, \theta_2, \theta_3; t) :=$$
$$\mathbb{E}_{x,|X_t|\neq 0} \exp\left\{ i\theta_1 e^{-\lambda_p t}|X_t| + i\theta_3 |X_t|^{-1/2} \sum_{i=1}^{|X_t|} (Z_t^i - z_t^i) \right\} h\left(\theta_2/\sqrt{e^{(n-1)\lambda_p t}|X_t|}\right)^{e^{(n-1)\lambda_p t}|X_t|}.$$

The limit of χ_1 is the same as the one of χ_2 providing that any of them exists. Indeed

$$\begin{aligned}&|\chi_2(\theta_1,\theta_2,\theta_3;t)-\chi_1(\theta_1,\theta_2,\theta_3;t)|\\ &\leq \mathbb{E}_{x,|X_t|\neq 0}\left|h\left(\theta_2/\sqrt{e^{(n-1)\lambda_p t}|X_t|}\right)^{e^{(n-1)\lambda_p t}|X_t|}-h\left(\theta_2/\sqrt{|X_{nt}|}\right)^{|X_{nt}|}\right|\\ &=\mathbb{E}_{x,|X_t|\neq 0}|\ldots|\,1_{Ext}+\mathbb{E}_{x,|X_t|\neq 0}|\ldots|\,1_{Ext^c}.\end{aligned}$$

The sequence of events $\{|X_t|\neq 0\}$ decreases to Ext^c and $\mathbb{P}(Ext^c)>0$. We have $|h|\leq 1$ hence the first summand converges to 0. By Proposition 2.1 on Ext^c we have $\frac{|X_{nt}|}{|X_t|e^{(n-1)\lambda_p t}}\to 1$ $a.s.$, therefore the second summand converges to 0 as well. Now we recall that conditionally on X_t, the processes $\tilde{X}^{i,t}_{(n-1)t}$ and thus also the variables Z^i_t are i.i.d. Conditioning with respect to $|X_t|$ we get

$$\chi_2(\theta_1,\theta_2,\theta_3;t)=$$

$$\mathbb{E}_{x,|X_t|\neq 0}\exp\left\{i\theta_1 e^{-\lambda_p t}|X_t|\right\}h_{Z^1_t-z^1_t}\left(\theta_3/|X_t|^{1/2}\right)^{|X_t|}h\left(\theta_2/\sqrt{e^{(n-1)\lambda_p t}|X_t|}\right)^{e^{(n-1)\lambda_p t}|X_t|}, \tag{4.25}$$

where $h_{Z^1_t-z^1_t}$ denotes the characteristic function of $Z^1_t-z^1_t$. We will now investigate the second term of the product above. To this end we fix two sequences $\{t_m\}$ and $\{a_m\}$ such that $t_m\to+\infty, a_m\to+\infty,$ as $m\to+\infty$ and consider

$$S_m:=\frac{1}{s_m}\sum_{i=1}^{a_m}(Z^i_{t_m}-z_{t_m}),$$

where $Z^i_{t_m}$ are i.i.d. copies as described above and $s_m^2:=a_m\operatorname{Var}_0(Z^1_{t_m})$. From Proposition 4.6 we easily see that s_m^2 is of the same order as a_m. In order to use CLT for triangular arrays, we will now verify the Lindeberg condition. Let us calculate

$$L_m(r):=\frac{1}{s_m^2}\sum_{i=1}^{a_m}\mathbb{E}_0\left((Z^i_{t_m}-z^i_{t_m})^2 1_{|Z^i_{t_m}-z^i_{t_m}|>rs_m}\right)=\frac{a_m}{s_m^2}\mathbb{E}_0\left((Z^1_{t_m}-z^1_{t_m})^2 1_{|Z^i_{t_m}-z^i_{t_m}|>rs_m}\right).$$

Further using inequality $1_{x>a}\leq x^2/a^2$ and Proposition 4.6 we get

$$L_m(r)\lesssim\frac{a_m}{s_m^2}\frac{\mathbb{E}_0(Z^i_{t_m}-z^i_{t_m})^4}{(rs_m)^2}\lesssim\frac{a_m}{s_m^2}\frac{1}{(rs_m)^2}\to 0,$$

as $m\to+\infty$. Thus the CLT holds, i.e. $S_m\to^d\mathcal{N}(0,1)$. By Proposition 4.6 we obtain $\frac{s_m^2}{a_m}\to\sigma_f^2$, where σ_f is given by (3.2). We conclude that

$$h_{Z_{t_m}-z_{t_m}}(\theta_3/\sqrt{a_m})^{a_m}\to e^{-\sigma_f^2\theta_3^2/2}. \tag{4.26}$$

Now we go back to χ_2

$$\chi_2(\theta_1,\theta_2,\theta_3;t)=\mathbb{E}_{x,|X_t|\neq 0}(\ldots)1_{Ext}+\mathbb{E}_{x,|X_t|\neq 0}(\ldots)1_{Ext^c}.$$

Arguing as before we get that the first summand converges to 0. The second one is

$$\mathbb{E}_{x,|X_t|\neq 0}(\ldots)1_{Ext^c}=\mathbb{P}_{x,|X_t|\neq 0}(Ext^c)\mathbb{E}_{x,Ext^c}(\ldots),$$

where $\mathbb{E}_{x,Ext^c}$ denotes the conditional probability with respect to Ext^c. We observe that $\mathbb{P}_{|X_t|\neq 0}(Ext^c)\to 1$. We proved (4.26) for arbitrary sequences, therefore the second factor in (4.25) converges (on the set of non-extinction) as follows

$$h_{Z^1_t-z^1_t}\left(\theta_3/|X_t|^{1/2}\right)^{|X_t|}\to e^{-\sigma_f^2\theta_3^2/2},\quad a.s.,\quad \text{as } t\to+\infty.$$

We know that $e^{-\lambda_p t}|X_t| \to W$ a.s. (recall that W denotes V_∞ conditioned on Ext^c). By the Lebesgue dominated convergence theorem we get

$$\chi_2(\theta_1, \theta_2, \theta_3; t) \to e^{-\sigma_V^2 \theta_2^2/2} e^{-\sigma_f^2 \theta_3^2/2} \mathbb{E}_x \exp\{i\theta_1 W\}, \text{ as } t \to +\infty.$$

We know that $\{|X_t| \neq 0\}$ is a decreasing sequence of events and $\bigcap_{t=1}^{\infty} \{|X_t| \neq 0\} = Ext^c$. Moreover $\mathbb{P}(Ext^c) > 0$, therefore by Lemma 4.1 with $A_1 = Ext^c$ and $A_2 = \{|X_t| \neq 0\}$, the limit of $Z_3(t)$ is the same if considered conditioned with respect to $\{|X_t| \neq 0\}$ or conditioned with respect to Ext^c. In this way we have proved that conditionally on Ext^c

$$Z_3(t) \to^d (W, G_1, G_2),$$

where (W, G_1, G_2) is the same as in the statement of Theorem 3.3.

Now, by standard arguments we can "transfer" the convergence of Z_3 to Z_2 (and thus to Z_1, which as we proved, is asymptotically equivalent to Z_2). Indeed, consider the continuous functions

$$g_1, g_2 : \mathbb{R}_+ \times \mathbb{R}^2 \mapsto \mathbb{R}_+ \times \mathbb{R}^2,\ g_1(z_1, z_2, z_3) := (z_1, z_2, z_1^{1/2} z_3),\ g_2(z_1, z_2, z_3) := (z_1, z_1^{1/2} z_2, z_3).$$

We have

$$Z_2(t) = g_2\Big(g_1(Z_3(t)) + (e^{-\lambda_p nt}|X_{nt}| - e^{-\lambda_p t}|X_t|, 0, 0)\Big) + \Big(e^{-\lambda_p t}|X_t| - e^{-\lambda_p nt}|X_{nt}|, 0, 0\Big),$$

and so convergence of Z_3 together with (4.20), Slutsky's lemma and the continuous mapping theorem yield

$$Z_2(t) \to^d \left(W, \sqrt{W} G_1, \sqrt{W} G_2\right) \text{ and thus also } Z_1(t) \to^d \left(W, \sqrt{W} G_1, \sqrt{W} G_2\right). \tag{4.27}$$

We recall that so far we have been working under the additional assumption (4.22), which we will now dispense of. Recall the definition (4.10). By Lemma 4.4 for any $u > 0$ the assumption (4.22) holds for l_u (instead of f) hence so does the convergence of (4.27) (with an obvious modification of the law of G_2). Functions l_u approximate $\tilde{f}$ as $u \searrow 0$. Using this fact together with some metric theoretic considerations we will now extend the convergence of Z_1 to the general case. To this end we make the dependence on the test function explicit. Let $Y_t(l_u)$ be the third coordinate of Z_1 (see (4.19)) with l_u instead of $\tilde{f}$. Further, let $L_t(l_u)$ be the law of $Y_t(l_u)$ and $L(l_u)$ be the law of the limit. By (4.17) and (3.2) we have

$$\begin{aligned} e(u) := \limsup_{t \to +\infty} \mathbb{E}\left(Y_t(\tilde{f}) - Y_t(l_u)\right)^2 &\leq \left\langle \varphi, (\tilde{f} - l_u)^2 \right\rangle \\ &+ 2\lambda p \int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s} \mathcal{T}_s(\tilde{f} - l_u)\right)^2 \right\rangle \mathrm{d}s =: I_1(u) + I_2(u). \end{aligned}$$

Using the dominated Lebesgue converge theorem we easily get $I_1(u) \to 0$ as $u \searrow 0$. Now, using (4.13) we obtain that for some $k > 0$

$$I_2(u) \leq c(u)^2 \int_0^{+\infty} e^{(\lambda_p - 2\mu)s} \left\langle \varphi, \left(1 + \|x\|^k\right)^2 \right\rangle \mathrm{d}s \to 0, \quad \text{as } u \searrow 0.$$

Thus $e(u) \to 0$ as $u \to 0$.

Let us recall (3.2). We have

$$\begin{aligned} &\sigma_{l_u}^2 - \sigma_f^2 = \\ &\left\langle \varphi, (l_u^2 - \tilde{f}^2) \right\rangle + 2\lambda p \int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s} \mathcal{T}_s l_u\right)^2 - \left(e^{(\lambda_p/2)s} \mathcal{T}_s \tilde{f}\right)^2 \right\rangle \mathrm{d}s =: I_3(u) + I_4(u). \end{aligned}$$

By (4.11) and (4.13) and the Cauchy-Schwarz inequality we get

$$I_4(u)^2 \lesssim \left(\int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s}\mathcal{T}_s(l_u - \tilde{f})\right)^2 \right\rangle \mathrm{d}s\right) \left(\int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s}\mathcal{T}_s(l_u + \tilde{f})\right)^2 \right\rangle \mathrm{d}s\right)$$
$$\lesssim c(u)^2 \left(\int_0^{+\infty} \left\langle \varphi, \left(e^{(\lambda_p/2)s}e^{-\mu s}(1+\|x\|)^n\right)^2 \right\rangle \mathrm{d}s\right)^2,$$

for some n. One can now show that $I_4(u) \to 0$ as $u \searrow 0$. One also easily checks that $I_3(u) \to 0$ as $u \searrow 0$. Using the fact that the limiting distributions $L(l_u)$ and $L(f)$ are centred Gaussians we get

$$m(L(l_u), L(\tilde{f})) \to 0, \quad \text{as } u \searrow 0.$$

Let us fix $\epsilon > 0$ and choose $u > 0$ such that $e(u) \leq \epsilon^2$ and $m(L(l_u), L(\tilde{f})) \leq \epsilon$. Since $L_t(l_u) \to^d L(l_u)$ for $t \to \infty$, we can choose $T > 0$ such that for any $t > T$ we have $m(L(l_u), L_t(l_u)) \leq \epsilon$ and (by the definition of $e(u)$)

$$\mathbb{E}\left(Y_t(\tilde{f}) - Y_t(l_u)\right)^2 \leq 4\epsilon^2.$$

Now, for $t > T$ we have $m(L(\tilde{f}), L_t(\tilde{f})) \leq 4\epsilon$ and thus the convergence (4.27) holds for any $f \in \mathcal{P}$. By the continuous mapping theorem this ends the proof. □

Remark 4.7. We remark that the proof of convergence of the fluctuations of the total number of particles $|X_t|$ given above did not use the assumption $\lambda_p < 2\mu$ and is valid for arbitrary λ. Clearly this is also true for the convergence of $e^{-\lambda_p t}|X_t|$. For this reason and to avoid unnecessary repetitions in the proofs for the critical and supercritical branching rate cases we will focus on spatial fluctuations and only indicate how they can be extended to joint convergence asserted in corresponding theorems.

4.6 CLT for large branching rate

In this section we present the proof of Theorem 3.12, deferring to the next section the proof of the CLT in the critical case, which will combine arguments used for small and large branching rates. We start with

Proof of Proposition 3.10. The fact that H is a martingale with respect to the filtration of the OU branching system follows by its branching property and easy calculations using (4.6), (4.7) and (2.5). Consider now $f_i(x) = x_i$ (for any $1 \leq i \leq d$). Notice that $\mathcal{T}_t f_i = e^{-\mu t} f_i$. Using (4.6) and (4.7) we get

$$e^{2(-\lambda_p+\mu)t}\mathbb{E}_x\langle X_t, f_i\rangle^2 \lesssim e^{2(-\lambda_p+\mu)t}e^{\lambda_p t}\mathcal{T}_t f^2(x)$$
$$+ e^{2(-\lambda_p+\mu)t}\int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[e^{2\lambda_p s}(T_s f_i(\cdot))^2\right](x)\mathrm{d}s$$
$$\lesssim e^{(-\lambda_p+2\mu)t}(1+\|x\|^2) + e^{(-\lambda_p+2\mu)t}\int_0^t e^{\lambda_p s}\mathcal{T}_{t-s}\left[e^{-2\mu s}f_i^2\right](x)\mathrm{d}s$$
$$\lesssim (1+\|x\|^2)\left(1 + \int_0^t e^{(-\lambda_p+2\mu)(t-s)}\mathrm{d}s\right) \leq C(1+\|x\|^2).$$

This proves that $\sup_{t\geq 0}\mathbb{E}_x\|H_t\|^2 = \sup_{t\geq 0}\sum_{i=1}^d \mathbb{E}_x\langle X_t, f_i\rangle^2 < \infty$.

To check that V and H are orthogonal one can use (4.7) and the standard polarization formula for inner products in Hilbert spaces. We leave this task to the reader, just mentioning that this result is somehow expected as $1, x$ are orthogonal with respect to densities g_t defining (2.4). □

Proof of Proposition 3.11. The proof follows easily by Proposition 4.3. Indeed it is enough to notice that, if we couple X and $\tilde{X}$ we may write

$$\tilde{H}_t - H_t = e^{(-\lambda_p+\mu)t} x e^{-\mu t}|X_t| = e^{-\lambda_p t} x |X_t|,$$

where $\tilde{H}, H$ are defined according to (3.10). The proof is concluded by (2.1). □

It will be useful to know that the range of the OU branching system grows at most linearly. This is a well-known fact for the branching Brownian motion. It is rather obvious that it also holds for the OU branching motion as the Ornstein-Uhlenbeck process is "better concentrated" than the Brownian motion. However, we were not able to find a proof in the literature hence we provide one.

Lemma 4.8. *Let X be the OU branching system starting from x. There exists a constant $C > 0$ and a random variable T such that with probability one*

$$\forall_{t>T} \forall_{i\in\{1,2,\ldots,|X_t|\}} X_t(i) \in B(xe^{-\mu t}, Ct),$$

where $B(x, r)$ denotes a ball of radius r centred at x.

Proof. By the coupling argument (Proposition 4.3) it is enough to prove the lemma for $x = 0$. We can also assume that $p = 1$. Let us denote by A_n the event that X up to time n is contained in $B(0, C(n-1))$ for some constant C. By the Borel–Cantelli lemma to show our claim it is enough to prove $\sum_{n\geq 1} \mathbb{P}(A_n^c) < +\infty$. By the result from [18] and the Gaussian concentration inequality we know that the supremum of the Ornstein-Uhlenbeck process on the interval $[0, n]$ can be stochastically dominated by $C_1(\sqrt{\log(n)} + |G|)$, where $C_1 > 0$ is a certain constant and G is a standard normal random variable. Let us fix any $\gamma > \lambda$ and estimate

$$\begin{aligned}
\mathbb{P}(A_n') &\leq \mathbb{P}(A_n' \cap \{|X_n| \leq e^{\gamma n}\}) + \mathbb{P}(|X_n| > e^{\gamma n}) \\
&= \mathbb{P}\left(\exists_{i\in\{1,2,\ldots,|X_n|\}} \sup_{s\in[0,n]} \|X_s(i)\| > C(n-1) \cap \{|X_n| \leq e^{\gamma n}\}\right) + \mathbb{P}(|X_n| > e^{\gamma n}) \\
&\leq \mathbb{P}\left(C_1(\sqrt{\log(n)} + |G|) > C(n-1)\right) e^{\gamma n} + \mathbb{P}(|X_n| > e^{\gamma n}),
\end{aligned}$$

where $X_s(i)$ denotes the position of the particle i or its ancestor at time $s \in [0, n]$. The proof is concluded by using the estimates of the Gaussian tail in the first term (giving an upper bound of the order $\exp(\gamma n - cn^2)$) and Proposition 2.1 together with the Chebyshev inequality for handling the second one (yielding an upper bound $\exp((\lambda - \gamma)n) \leq \exp(-cn)$). □

We will also need

Proposition 4.9. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system starting from x and $\lambda_p > 2\mu$. Moreover, let $f \in \mathcal{P}(\mathbb{R}^d)$. Then*

$$\sup_{t\geq 0} \mathbb{E}_x \left(e^{-(\lambda_p-\mu)t} \left\langle X_t, \tilde{f} \right\rangle\right)^2 < +\infty.$$

Moreover there exists $c : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that $c(u) \to 0$ as $u \searrow 0$ and

$$\limsup_{t\to+\infty} \mathbb{E}_x \left(e^{-(\lambda_p-\mu)t} \left(\left\langle X_t, \tilde{f} \right\rangle - \langle X_t, l_u \rangle\right)\right)^2 < c(u), \tag{4.28}$$

where l_u is given by (4.10).

Proof. We recall (4.6) and (4.7), which we will apply for $\tilde{f}$ instead of f. By Lemma 4.5 (formula (4.11)) one sees that $w'(x,t) \lesssim (1+\|x\|^n)e^{(\lambda_p-\mu)t}$. Now we check that

$$|w''(x,t)| \lesssim e^{\lambda_p t}\mathcal{T}_t\tilde{f}^2(x) + \int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[\left((1+\|\cdot\|^n)e^{(\lambda_p-\mu)s}\right)^2\right](x)\mathrm{d}s$$
$$\lesssim e^{\lambda_p t}\mathcal{T}_t\tilde{f}^2(x) + e^{\lambda_p t}\int_0^t e^{(\lambda_p-2\mu)s}\mathcal{T}_{t-s}\left[(1+\|\cdot\|^{2n})\right](x)\mathrm{d}s \lesssim e^{2(\lambda_p-\mu)t}(1+\|x\|^{2n}),$$

which implies the first assertion of the proposition. Inequality (4.28) follows by (4.13) and the above calculations. Indeed, repeating the above argument for $l_u - \tilde{f}$ instead of $\tilde{f}$ one easily checks that

$$\mathbb{E}_x\left(e^{-(\lambda_p-\mu)t}\left(\left\langle X_t,\tilde{f}\right\rangle - \langle X_t,l_u\rangle\right)\right)^2 = e^{-2(\lambda_p-\mu)t}w''(x,t)$$
$$\lesssim e^{(2\mu-\lambda_p)t}\mathcal{T}_t(l_u-\tilde{f})^2(x) + e^{-2(\lambda_p-\mu)t}\int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}[e^{\lambda_p s}\mathcal{T}_s(l_u-\tilde{f})(\cdot)]^2(x)ds.$$

Since $\lambda_p > 2\mu$ and $\mathcal{T}_t(l_u-\tilde{f})^2(x) \lesssim 1+\|x\|^n$ for some positive integer n, the first summand on the right hand side above converges to 0 as $t\to\infty$. Moreover, the integrand in the second term is bounded by $Ce^{\lambda_p(t+s)}(1+\|x\|^n)$ for some positive integer n, so

$$e^{-2(\lambda_p-\mu)t}\int_0^1 e^{\lambda_p(t-s)}\mathcal{T}_{t-s}[e^{\lambda_p s}\mathcal{T}_s(l_u-\tilde{f})(\cdot)]^2(x)ds \lesssim e^{(2\mu-\lambda_p)t}\to 0$$

as $t\to\infty$. On the other hand, for $s>1$, by (4.13), we have $|\mathcal{T}_s(l_u-\tilde{f})(\cdot)| \lesssim c(u)(1+\|\cdot\|^n)e^{-\mu s}$ and thus

$$e^{-2(\lambda_p-\mu)t}\int_1^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}[e^{\lambda_p s}\mathcal{T}_s(l_u-\tilde{f})(\cdot)]^2(x)ds \lesssim e^{-(\lambda_p-2\mu)t}c(u)^2\int_1^t e^{(\lambda_p-2\mu)s}(1+\|x\|^{2n})ds$$
$$\lesssim Cc(u)^2(1+\|x\|^{2n}),$$

which together with the previous estimates proves (4.28). □

Proof of Theorem 3.12. Our first aim will be to prove the convergence of the spatial fluctuations. To this end we denote

$$Y_1(t) := e^{-(\lambda_p-\mu)t}\left(\langle X_t,f\rangle - |X_t|\langle\varphi,f\rangle\right) = e^{-(\lambda_p-\mu)t}\left\langle X_t,\tilde{f}\right\rangle. \tag{4.29}$$

We recall that $\tilde{f}(x) = f(x) - \langle\varphi,f\rangle$. We have

$$Y_1(t+s) = e^{-(\lambda_p-\mu)(t+s)}\sum_{i=1}^{|X_{t+s}|}\tilde{f}(X_{t+s}(i)) = e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|}e^{-(\lambda_p-\mu)s}\sum_{j=1}^{|X_s^{i,t}|}\tilde{f}(X_s^{i,t}(j))$$
$$= e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|}e^{-(\lambda_p-\mu)s}\left(\sum_{j=1}^{|X_s^{i,t}|}\left(\tilde{f}(X_s^{i,t}(j)) - \tilde{f}(\tilde{X}_s^{i,t}(j))\right) + \tilde{f}(\tilde{X}_s^{i,t}(j))\right),$$

where as in the proof of Theorem 3.3, $\{X_s^{i,t}\}_s$ denotes the subsystem starting from $X_t(i)$ and $\{\tilde{X}_s^{i,t}\}_s$ is defined by (4.21). We write

$$Y_1(t+s) = e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|}e^{-(\lambda_p-\mu)s}\sum_{j=1}^{|X_s^{i,t}|}(\tilde{f}(X_s^{i,t}(j))-\tilde{f}(\tilde{X}_s^{i,t}(j)))+e^{-(\lambda_p-\mu)(t+s)}\sum_{i=1}^{|X_t|}\left\langle\tilde{X}_s^{i,t},\tilde{f}\right\rangle. \tag{4.30}$$

Let us first deal with the second term

$$Y_2(t+s) := e^{-(\lambda_p-\mu)(t+s)} \sum_{i=1}^{|X_t|} \left\langle \tilde{X}_s^{i,t}, \tilde{f}_s \right\rangle + e^{-(\lambda_p-\mu)(t+s)} \gamma_s |X_{t+s}|,$$

where $\tilde{f}_s(x) := f(x) - \mathcal{T}_s f(0)$ and $\gamma_s := \mathcal{T}_s f(0) - \langle \varphi, f \rangle$. By Lemma 4.5 we know that $|\gamma_s| \lesssim e^{-2\mu s}$ and therefore

$$\mathbb{E}_x |e^{-(\lambda_p-\mu)(t+s)} \gamma_s |X_{t+s}|| \lesssim e^{\mu(t+s)-2\mu s}.$$

From now on, whenever we prove convergence, we will assume that $s = 2t$. With this convention the above expression converges to 0. We denote the first summand of Y_2 by Y_3. As we explained in the proof of Theorem 3.3 the systems $\tilde{X}^{i,t}$ are i.i.d. conditionally on X_t. Moreover their only connection with X_t is via the number of particles $|X_t|$. Below, we will use the conditional expectation given X_t which is denoted as $\mathbb{E}_{X_t}$. To ease the notation we will sometimes write simply $\mathbb{E}_0 \tilde{X}^{i,t}$ instead of $\mathbb{E}_{X_t} \tilde{X}^{i,t}$ (recall that the systems $\tilde{X}^{t,i}$ start from a single particle located at 0). Using the conditional expectation and the fact that thanks to the particular choice of $\tilde{f}_s$ we have $\mathbb{E}_0 \langle \tilde{X}_s^{i,t}, f_s \rangle = 0$, we calculate

$$\begin{aligned}
\mathbb{E}_x (Y_3(t+s))^2 &= e^{-2(\lambda_p-\mu)(t+s)} \mathbb{E}_x \sum_{i=1}^{|X_t|} \sum_{j=1}^{|X_t|} \mathbb{E}_{X_t} \left(\left\langle \tilde{X}_s^{i,t}, \tilde{f}_s \right\rangle \left\langle \tilde{X}_s^{j,t}, \tilde{f}_s \right\rangle \right) \\
&= e^{-2(\lambda_p-\mu)(t+s)} \mathbb{E}_x \sum_{i=1}^{|X_t|} \mathbb{E}_0 \left\langle \tilde{X}_s^{i,t}, \tilde{f}_s \right\rangle^2.
\end{aligned}$$

By Proposition 4.9 and Proposition 2.1 we have

$$\mathbb{E}_0 \left\langle \tilde{X}_s^{i,t}, \tilde{f}_s \right\rangle^2 = \mathbb{E}_0 \left(\left\langle \tilde{X}_s^{i,t}, \tilde{f} \right\rangle - \gamma_s |\tilde{X}_s^{i,t}| \right)^2 \leq 2\mathbb{E}_0 \left\langle \tilde{X}_s^{i,t}, \tilde{f} \right\rangle^2 + 2\gamma_s^2 \mathbb{E}_0 |\tilde{X}_s^{i,t}|^2 \leq e^{2(\lambda_p-\mu)t}.$$

Therefore

$$\mathbb{E}_x (Y_3(t+s))^2 \lesssim e^{(-\lambda_p+2\mu)t} \to 0.$$

In this way we proved that $Y_2(s+t) \to 0$ in probability. Thus the second term on the right hand side of (4.30) is negligible.

From now on let us assume additionally that f has the second derivative which is bounded by a polynomial. At the end of the proof we will argue that the result can be easily extended to the whole class $\mathcal{P}(\mathbb{R}^d)$.

Now we decompose the first term of (4.30) into $Y_5(t+s) + Y_6(t+s)$ where

$$Y_5(t+s) := e^{-(\lambda_p-\mu)(t+s)} \sum_{i=1}^{|X_t|} \sum_{j=1}^{|X_s^{i,t}|} E(i,j,t,s), \tag{4.31}$$

with $E(i,j,t,s) := \tilde{f}(X_s^{i,t}(j)) - \tilde{f}(\tilde{X}_s^{i,t}(j)) - \nabla f(\tilde{X}_s^{i,t}(j)) \circ (X_s^{i,t}(j) - \tilde{X}_s^{i,t}(j))$ and

$$Y_6(t+s) := e^{-(\lambda_p-\mu)(t+s)} \sum_{i=1}^{|X_t|} \sum_{j=1}^{|X_s^{i,t}|} \nabla f(\tilde{X}_s^{i,t}(j)) \circ (X_s^{i,t}(j) - \tilde{X}_s^{i,t}(j)). \tag{4.32}$$

Let $A_{t,s}$ denote the event that $X_s^{i,t}(j), \tilde{X}_s^{i,t}(j)$ belong to the ball $B(0,R)$ of radius $R = 2(C(t+s) + \|x\|e^{-\mu t})$ for all $i \leq |X_t|$, $j \leq |\tilde{X}_s^{i,t}|$, where C is the constant from Lemma 4.8. On the event $A_{t,s}$, using (4.21) and the assumptions on the second derivative of f, we obtain

$$|E(i,j,t,s)| \lesssim R^n (\|X_t(i)\| e^{-\mu s})^2.$$

By Proposition 2.1 and the convention $s = 2t$ we thus prove

$$\mathbb{E}_x|Y_5(t+s)|1_{A_{t,s}} \lesssim e^{\mu(s+t)}((t+s)+e^{-\mu t}\|x\|)^{3n}e^{-2\mu s} \to 0.$$

By Lemma 4.8 and (4.21) we have $\mathbb{P}(A_{t,s}) \to 1$, hence it follows that $Y_5(t+s) \to 0$ in probability (recall the convention $s = 2t$). Now, using (4.21) we decompose Y_6 as follows

$$\begin{aligned} Y_6(t+s) =& e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|} X_t(i)\circ\left(Z_s^i - z_s\right) + z_s \circ \left(e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|} X_t(i)\right) \\ &+ \langle\varphi,\nabla f\rangle \circ \left(e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|} X_t(i)(e^{-\lambda_p s}|X_s^{i,t}|-1)\right) \\ &+ \langle\varphi,\nabla f\rangle \circ \left(e^{-(\lambda_p-\mu)t}\sum_{i=1}^{|X_t|} X_t(i)\right) \\ =:& Y_7(t+s)+Y_8(t+s)+Y_9(t+s)+Y_{10}(t+s), \end{aligned} \tag{4.33}$$

where $h(x) = \nabla f(x) - \langle\varphi,\nabla f\rangle$, $Z_s^i := e^{-\lambda_p s}\left\langle \tilde{X}_s^{i,t}, h\right\rangle$ and $z_s := \mathbb{E}_0 Z_s^i = \mathcal{T}_s(\nabla f)(0) - \langle\varphi,\nabla f\rangle$ (this does not depend on i since Z_s^i are i.i.d.). Simple calculations using (4.6) and Lemma 4.5 (second estimate in (4.11)) reveal that $\|z_s\| \lesssim e^{-2\mu s}$. Moreover by Proposition 4.9 the covariance matrix $\mathrm{Cov}_0(Z_s^i)$ is bounded by some C (in a sense that each entry is bounded). Using conditioning with respect to $|X_t|$ and the fact that Z_s^i are i.i.d. conditionally on X_t we have

$$\begin{aligned} \mathbb{E}_x Y_7(t+s)^2 &= e^{-2(\lambda_p-\mu)t}\mathbb{E}_x \sum_{i=1}^{|X_t|} X_t(i)^T\,\mathrm{Cov}_0\,(Z_s^i)X_t(i) \\ &\lesssim e^{-2(\lambda_p-\mu)t}\mathbb{E}_x \sum_{i=1}^{|X_t|} \|X_t(i)\|^2 \lesssim e^{(-\lambda_p+2\mu)t}(1+|x|^2) \to 0, \end{aligned}$$

where ^{T} denotes the transposition and the convergence holds by the assumption that $\lambda_p > 2\mu$. The convergence: $\mathbb{E}_x(Y_9(t+s))^2 \to 0$ follows in a very similar fashion with use of Proposition 2.1 and is left to the reader.

Now by Proposition 3.10 and the estimate $\|z_s\| \lesssim e^{-2\mu s}$ one easily checks that $Y_8(t+s) \to 0$ a.s. and Y_{10} converges to $\langle\varphi,\nabla f\rangle \circ H_\infty$. By the already established convergence to zero of $Y_i(t+s)$, $i = 5,7,8,9$ together with (4.31), (4.32) and (4.33) this shows that the first summand on the right hand side of (4.30) converges to the same limit. Since the second summand has already been shown to be negligible, this gives the convergence in probability

$$Y_1(t) \to \langle\varphi,\nabla f\rangle \circ H_\infty$$

and together with (2.1) implies the second part of the theorem.

Clearly the above convergence gives also weak convergence of the third component on the left hand side of (3.11) to $\langle\varphi,\nabla f\rangle \circ J$ (conditionally on Ext^c).

We recall that so far the proof works with the additional assumption that $\tilde{f}$ is C^2 and its second derivative is bounded by a polynomial. We recall also the definition (4.10). By Lemma 4.4 for any $u > 0$ the function l_u satisfies our additional smoothness assumptions, moreover $l_0 = \tilde{f}$. As in the proof of Theorem 3.3, this suffices to reduce the general case to the one of smooth functions. Let $Y_t(l_u)$ be $Y_1(t)$ (see (4.29)) with l_u instead of $\tilde{f}$. Further, let $Y(l_u)$ be its limit (its existence follows from the smooth case considered above and smoothness properties of l_u asserted by Lemma 4.4). Let m be any metric which metrises convergence in probability and such that for any random variables X, Y, $m(X,Y) \leq \|X-Y\|_2$. We recall (3.11). Using integration by parts it is easy to check that $\langle \varphi, \nabla l_u \rangle \to \langle \varphi, \nabla f \rangle$ as $u \searrow 0$, therefore

$$m(Y(l_u), Y(\tilde{f})) \to 0, \quad \text{as } u \searrow 0.$$

Let us fix $\epsilon > 0$ and choose $u > 0$ such that $m(Y(l_u), Y(\tilde{f})) \leq \epsilon$. By (4.28), decreasing u if necessary we can find T such that for any $t > T$

$$m(Y_t(l_u), Y_t(\tilde{f})) \leq \epsilon.$$

Finally, we choose t large enough so that $m(Y_t(l_u), Y(l_u)) < \varepsilon$. Applying the triangle inequality, for these t's we get $m(L(\tilde{f}), L_t(\tilde{f})) \leq 4\epsilon$ and so the convergence holds for any $f \in \mathcal{P}$.

We have thus proved the convergence in probability of the spatial fluctuations. The convergence in probability of the first coordinate of the vector considered in Theorem 3.12 is just (2.1). As for the joint convergence of the whole triple, it can be obtained in the same way as in Theorem 3.3, i.e. by considering the characteristic function of the triple and consecutive conditioning. Since the only difference from the slow branching rate case is the already established convergence of the third summand (see Remark 4.7 following the proof of Theorem 3.3), we leave the details to the reader.

4.7 CLT for critical branching rate

Proposition 4.10. *Let $\{X_t\}_{t\geq 0}$ be the OU branching system and $\lambda_p = 2\mu$. Moreover, let $f \in \mathcal{P}(\mathbb{R}^d)$, then there exists a constant $C > 0$ such that*

$$\mathbb{E}_x \left(e^{-(\lambda_p/2)t} t^{-1/2} \left\langle X_t, \tilde{f} \right\rangle \right) \to 0 \quad \text{as } t \to +\infty,$$

$$\mathbb{E}_x \left(e^{-(\lambda_p/2)t} t^{-1/2} \left\langle X_t, \tilde{f} \right\rangle \right)^2 \to \sigma_f^2, \quad \mathrm{Var}_x \left(e^{-(\lambda_p/2)t} t^{-1/2} \left\langle X_t, \tilde{f} \right\rangle \right) \to \sigma_f^2, \quad \text{as } t \to +\infty, \tag{4.34}$$

where σ_f^2 is the same as in (3.8)*. Moreover,*

$$\sup_{t\geq\delta} \mathbb{E}_x \left(e^{-(\lambda_p/2)t} t^{-1/2} \left\langle X_t, \tilde{f} \right\rangle \right)^4 < +\infty \tag{4.35}$$

for any $\delta > 0$.

Proof. The first convergence follows easily by (4.11) in Lemma 4.5. Using (4.6) and (4.7) we calculate the second moment

$$\begin{aligned} &\mathbb{E}_x \left(e^{-(\lambda_p/2)t} t^{-1/2} \left\langle X_t, \tilde{f} \right\rangle \right)^2 \\ &\quad = t^{-1} \mathcal{T}_t \tilde{f}^2(x) + 2\lambda p e^{-\lambda_p t} t^{-1} \int_0^t e^{\lambda_p (t-s)} \mathcal{T}_{t-s} \left[\left(e^{\lambda_p s} \mathcal{T}_s \tilde{f}(\cdot) \right)^2 \right] (x) \mathrm{d}s \\ &\quad = t^{-1} \mathcal{T}_t \tilde{f}^2(x) + 2\lambda p t^{-1} \int_0^t \mathcal{T}_{t-s} \left[\left(e^{(\lambda_p/2)s} \mathcal{T}_s \tilde{f}(\cdot) \right)^2 \right] (x) \mathrm{d}s. \end{aligned} \tag{4.36}$$

We recall that $\lambda_p/2 = \mu$. Using Lemma 4.5 (equation (4.12)) and elementary considerations, we obtain that the limit of the above expression is the same as the one of

$$2\lambda p t^{-1}\int_0^t \mathcal{T}_{t-s}\left[\left(\sum_{i=1}^d x_i\left\langle\varphi,\frac{\partial f}{\partial x_i}\right\rangle\right)^2\right](x)\mathrm{d}s = \\ 2\lambda p\sum_{i=1}^d\left(\left\langle\varphi,\frac{\partial f}{\partial x_i}\right\rangle^2 t^{-1}\int_0^t \mathcal{T}_{t-s}\left[x_i^2\right](x)\mathrm{d}s\right) \\ +2\lambda p\sum_{i\neq j}\left(\left\langle\varphi,\frac{\partial f}{\partial x_i}\right\rangle\left\langle\varphi,\frac{\partial f}{\partial x_j}\right\rangle t^{-1}\int_0^t \mathcal{T}_{t-s}\left[x_ix_j\right](x)\mathrm{d}s\right).$$

By (2.4) one easily checks that $\mathcal{T}_t[x_ix_j](x)\to 0$, hence the second term disappears in the limit. We also have $\mathcal{T}_t[x_i^2](x)\to\frac{\sigma^2}{2\mu}$ and so the whole expression converges to σ_f given by (3.8).

Obviously the limit of variances is the same. We also conclude that for any $t\geq 0$

$$w''(x,t)\lesssim(1+\|x\|^{2n})e^{\lambda_p t}t.$$

Similarly we investigate $w'''(x,t)$. By (4.6) we have

$$|w'''(x,t)|\lesssim e^{\lambda_p t}\mathcal{T}_t|\tilde{f}|^3(x)+\int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[w''(\cdot,s)w'(\cdot,s)\right](x)\mathrm{d}s.$$

Using the fact that by (4.11) $|w'(x,t)|\lesssim e^{(\lambda_p-\mu)t}(1+\|x\|^n)$, together with the above estimate on w'' and the fact that $\mathcal{T}_t\left[(1+\|\cdot\|^n)^3\right](x)\lesssim(1+\|x\|^{3n})$, we get

$$|w'''(x,t)|\lesssim(1+\|x\|^{3n})e^{\lambda_p t}+e^{\lambda_p t}\int_0^t e^{(\lambda_p/2)s}s\mathcal{T}_{t-s}\left[1+\|\cdot\|^{3n}\right]\mathrm{d}s\lesssim e^{((3/2)\lambda_p)t}t(1+\|x\|^{3n}).$$

Finally, we will also need the fourth moment. By (4.6) and the estimates above we get

$$\mathbb{E}_x\left(e^{-(\lambda_p/2)t}t^{-1/2}\left\langle X_t,\tilde{f}\right\rangle\right)^4\lesssim e^{-\lambda_p t}t^{-2}\mathcal{T}_t\tilde{f}^4(x) \\ +e^{-2\lambda_p t}t^{-2}\int_0^t e^{\lambda_p(t-s)}\mathcal{T}_{t-s}\left[w''(\cdot,s)^2+w'''(\cdot,s)w'(\cdot,s)\right](x)\mathrm{d}s \\ \lesssim e^{-\lambda_p t}t^{-2}(1+\|x\|^{4n})+e^{-\lambda_p t}t^{-2}\int_0^t e^{\lambda_p s}s^2\mathcal{T}_{t-s}\left[(1+\|\cdot\|^{4n})\right](x)\mathrm{d}s\lesssim(1+\|x\|^{4n}).$$

It is now easy to check that for $t>\delta$,

$$\mathbb{E}_x\left(e^{-(\lambda_p/2)t}t^{-1/2}\left(\left\langle X_t,\tilde{f}\right\rangle-\mathbb{E}_x\left\langle X_t,\tilde{f}\right\rangle\right)\right)^4\lesssim(1+\|x\|^{4n}).$$

□

We have yet to prove (3.9). By (2.3) we have $\frac{\partial}{\partial x_i}\varphi(x)=-2(\mu/\sigma^2)x_i\varphi(x)$. Therefore,

$$\left\langle\varphi,\frac{\partial f}{\partial x_1}\right\rangle=-\left\langle\frac{\partial\varphi}{\partial x_1},f\right\rangle=\left\langle 2(\mu/\sigma^2)x_1\varphi,f\right\rangle=(\sqrt{2\mu}/\sigma)f_{1,0,\ldots,0}.$$

Other coordinates can be treated in the same way, giving (3.9).

Now we are ready for

Proof of Theorem 3.8. In this proof we will use both ideas of the proof of Theorem 3.3 and Theorem 3.12. We start with the following random vector

$$Z_1(t) := \left(e^{-\lambda_p t}|X_t|, e^{-(\lambda_p/2)t}(|X_t| - e^{\lambda_p t}V_\infty), e^{-(\lambda_p/2)t}t^{-1/2}\left\langle X_t, \tilde{f}\right\rangle\right).$$

Let $n \in \mathbb{N}$. This parameter will be used in a different way than in the previous proofs. We will comment on it later on. We observe that

$$Z_1(nt) = \left(e^{-n\lambda_p t}\left\langle X_{nt}, 1\right\rangle, e^{-(n\lambda_p/2)t}(|X_{nt}| - e^{n\lambda_p t}V_\infty), e^{-(n\lambda_p/2)t}(nt)^{-1/2}\sum_{i=1}^{|X_t|}\left\langle X^{i,t}_{(n-1)t}, \tilde{f}\right\rangle\right),$$

where $\left\{X^{i,s}_t\right\}_t$ denotes the subsystem originating from the particle $X_s(i)$. Analogously as in the proof of Theorem 3.3 the second term is equal to $e^{-(n\lambda_p/2)t}\sum_{i=1}^{|X_{nt}|}\left(1 - V^i_\infty\right)$, where V^i_∞ are independent (conditionally on X_{nt}) copies of V_∞ arising from the i-th particle alive at time nt. Next, we couple each $X^{i,t}$ with the branching system $\tilde{X}^{i,t}$ starting from one particle located at 0 using (4.21) and Proposition 4.3. We write

$$I_n(t) := e^{-(n\lambda_p/2)t}(nt)^{-1/2}\left(\sum_{i=1}^{|X_t|}\left\langle X^{i,t}_{(n-1)t}, \tilde{f}\right\rangle - \sum_{i=1}^{|X_t|}\left\langle \tilde{X}^{i,t}_{(n-1)t}, \tilde{f}\right\rangle\right).$$

Our immediate aim is to prove that

$$\lim_{n\to\infty}\limsup_{t\to\infty}\mathbb{E}\frac{|I_n(t)|}{1+|I_n(t)|} = 0.$$

Recall that $\rho(X,Y) = \mathbb{E}\frac{|X-Y|}{1+|X-Y|}$ is a metric on the space of all random variables, which metrizes convergence in probability. Thus the above convergence will give us control on the distance between the original process and the one obtained via coupling, which will enable us to establish the limit theorem for the original process. Contrary to previous cases, it will not be enough to consider a fixed parameter n, instead we will prove a limit theorem for the process obtained via coupling evolving through time nt for each n and then we will let $n \to \infty$.

We will assume additionally that f has a second derivative which is bounded by a polynomial. At the end of the proof we will argue that the result can be easily extended to the whole class $\mathcal{P}(\mathbb{R}^d)$.

The expression $I_n(t)$ is harder to analyse compared to the case of the small branching rate. To deal with it we will use the methods and notation of the proof of the large branching rate case (i.e. Theorem 3.12). Recalling that $\lambda_p = 2\mu$ one notices that $I_n(t)$ is the same expression as

$$(nt)^{-1/2}\left(Y_5(t + (n-1)t) + Y_6(t + (n-1)t)\right)$$

when one puts $s = (n-1)t$ (see (4.31) and (4.32)). Y_5 can be then handled in the same way as in the proof of Theorem 3.12, if only $n > 2$ (we skip the corresponding estimates which are completely analogous). Thus we obtain $Y_5(nt) \to 0$ in probability as $t \to \infty$. Next, we denote $s := (n-1)t$. Using the coupling property (4.21) and the equality $\lambda_p = 2\mu$, we decompose Y_6 in the following way

$$(nt)^{-1/2}Y_6(nt) = (nt)^{-1/2}e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}X_t(i)\circ K^i_s = (nt)^{-1/2}e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}X_t(i)\circ(K^i_s - k_s)$$
$$+ k_s \circ \left((nt)^{-1/2}e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}X_t(i)\right), \quad (4.37)$$

where this time $K_s^i := e^{-\lambda_p s}\left\langle \tilde{X}_s^{i,t}, \nabla f\right\rangle$ and $k_s := \mathbb{E}_0 K_s^i$ (this is independent of i as K_s^i are i.i.d). Note that $k_s = \mathcal{T}_s \nabla f(0)$ and so, by the assumption on the derivatives of f, it is uniformly bounded in s. Also $\mathbb{E}_0\|K_s^i - k_s\|^2$ are uniformly bounded in s which follows easily by (4.7). Using conditioning in a similar manner as in the estimation of Y_7 in the proof of Theorem 3.12 we obtain

$$\mathbb{E}_x\left((nt)^{-1/2}e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|} X_t(i)\circ(K_s^i - k_s)\right)^2 = (nt)^{-1}e^{-\lambda_p t}\mathbb{E}_x\sum_{i=1}^{|X_t|} X_t(i)^T \operatorname{Cov}_0(K_s^i)X_t(i) \tag{4.38}$$

$$\lesssim (nt)^{-1}e^{-\lambda_p t}\mathbb{E}_x\sum_{i=1}^{|X_t|}\|X_t(i)\|^2 \lesssim (nt)^{-1}(1+\|x\|^2)\to 0,$$

where again we used (4.6) and (4.7) and Proposition 4.10. Further, we notice also that for all $i = 1,\dots,d$, $\langle \varphi, x_i\rangle = 0$. Therefore by (4.34) of Proposition 4.10 (applied to each coordinate separately) and the fact that k_s is bounded in s we obtain that the second moment of the second term of (4.37) is bounded by $c/(2\sqrt{n})$, where $c > 0$ is a certain constant. Thus, using (4.37), (4.38) and the convergence of Y_5, we may conclude that for any n there exists t_n such that for any $t \geq t_n$,

$$\mathbb{E}\frac{|I_n(t)|}{1+|I_n(t)|} \leq 2c/\sqrt{n}. \tag{4.39}$$

We recall that conditionally on X_t, $\tilde{X}^{i,t}$'s are i.i.d. branching particle systems. Let us denote

$$\begin{aligned} Z_t^i &:= e^{-((n-1)\lambda_p/2)t}(nt)^{-1/2}\left\langle \tilde{X}^{i,t}_{(n-1)t}, \tilde{f}\right\rangle \\ &= \left(\frac{n-1}{n}\right)^{1/2}\left(e^{-((n-1)\lambda_p/2)t}((n-1)t)^{-1/2}\left\langle \tilde{X}^{i,t}_{(n-1)t}, \tilde{f}\right\rangle\right) \end{aligned} \tag{4.40}$$

and $z_t := \mathbb{E}_0 Z_t^i$. By Lemma 4.5 one checks that $|z_t| = e^{((n-1)\lambda_p/2)t}(nt)^{-1/2}|\mathcal{T}_t\tilde{f}(0)| \lesssim e^{((n-1)(\lambda_p/2-2\mu))t}$. Therefore for $n > 2$, using $\lambda_p = 2\mu$ we get

$$\mathbb{E}e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}|z_t^i| = e^{(\lambda_p/2)t}|z_t^1| \lesssim e^{(\lambda_p/2)t}e^{((n-1)(\lambda_p/2-2\mu))t}\to 0, \quad \text{as } t\to+\infty.$$

Using the definitions and considerations above we conclude that for any fixed $n > 2$ the asymptotic behaviour of $Z_1(nt)$ is the same as the one of

$$Z_2^n(t) := \left(e^{-\lambda_p t}|X_t|, e^{-(n\lambda_p/2)t}\sum_{i=1}^{|X_{nt}|}\left(1-V_\infty^i\right), I_n(t) + e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}(Z_t^i - z_t^i)\right).$$

This expression differs from the analogous one in the proof of Theorem 3.3 only by the additional "error" term $I_n(t)$. We note that unlike the previous proofs we do not prove that this error, stemming from replacing $X^{i,t}$ with $\tilde{X}^{i,t}$, converges to 0. The proof strategy is to ignore it for a moment. After having proved the convergence of the modified variables we will be able to cope with it using (4.39). We thus define

$$\tilde{Z}_2^n(t) := \left(e^{-\lambda_p t}|X_t|, e^{-(n\lambda_p/2)t}\sum_{i=1}^{|X_{nt}|}\left(1-V_\infty^i\right), e^{-(\lambda_p/2)t}\sum_{i=1}^{|X_t|}(Z_t^i - z_t^i)\right). \tag{4.41}$$

Returning to the flow of the proof of Theorem 3.3, one can prove that

$$\tilde{Z}_2^n(t) \to^d \left(W, \sqrt{W}G_1, \left(\frac{n-1}{n}\right)^{1/2} \sqrt{W}G_2\right), \tag{4.42}$$

conditionally on Ext^c. In order to achieve it one considers the characteristic function of $\tilde{Z}_2^n$ and follows consecutive steps of the proof of Theorem 3.3, using (4.35) of Proposition 4.10 instead of (4.18) of Proposition 4.6 to check the Lindeberg condition. The additional term $\left(\frac{n-1}{n}\right)^{1/2}$ originates from the same term in (4.40).

We are going to show now how (4.42) implies the convergence of $Z_1(t)$. We recall (4.8) and denote the law of the triple on the right hand side of (4.42) by $\mathcal{L}_n$ and the one of $(W, \sqrt{W}G_1, \sqrt{W}G_2)$ by $\mathcal{L}_\infty$. For any $\epsilon > 0$, there exists n such that $m(\mathcal{L}_n, \mathcal{L}_\infty) \leq \epsilon$ and $c/\sqrt{n} \leq \varepsilon$, where c is the constant in (4.39). Now we choose T large enough to have $m(\mathcal{L}(\tilde{Z}_2^n(t)),\ \mathcal{L}_n) \leq \epsilon$, $m(\mathcal{L}(Z_2^n(t)), \tilde{Z}_2^n(t))) \leq \epsilon$ and $m(\mathcal{L}(Z_2^n(t)), \mathcal{L}(Z_1(nt))) \leq \epsilon$ for any $t \geq T$. By applying the triangle inequality we get $m(\mathcal{L}_\infty, \mathcal{L}(Z_1(nt))) \leq 4\epsilon$ for any $t \geq T$. The proof (in the case of smooth function f) can be now concluded by the continuous mapping theorem and Slutsky's lemma, analogously as in the proof of Theorem 3.3.

We note that the proof so far is valid for f with additional assumption, that the second order derivatives exist and are bounded by some polynomial. Using exactly the same technique as in proof of Theorem 3.3 (i.e. approximating $\tilde{f}$ by l_u with $u \to 0$ and using (4.13) of Lemma 4.5 to control the distance between corresponding processes) one can easily extend the result to any $f \in \mathcal{P}$. We omit the standard details.

Appendix

Proof of Proposition 2.1. Using (4.2) with $f(x) = 1$ we get (we drop the argument x)

$$\frac{d}{dt}w(t,\theta) = (\lambda p)w(t,\theta)^2 - \lambda w(t,\theta) + \lambda(1-p), \quad w(0,\theta) = e^{-\theta}.$$

The solution of this equation is (see e.g. [5, Section III.5])

$$w(t,\theta) = \frac{\lambda(1-p)(e^{-\theta}-1) - e^{-\lambda_p t}(\lambda p e^{-\theta} - \lambda(1-p))}{\lambda p(e^{-\theta}-1) - e^{-\lambda_p t}(\lambda p e^{-\theta} - \lambda(1-p))}.$$

Now we want to investigate convergence of $e^{-\lambda_p t}|X_t|$. Its Laplace transform is

$$L(t,\theta) := \frac{\lambda(1-p)(e^{-\theta e^{-\lambda_p t}}-1) - e^{-\lambda_p t}(\lambda p e^{-\theta e^{-\lambda_p t}} - \lambda(1-p))}{\lambda p(e^{-\theta e^{-\lambda_p t}}-1) - e^{-\lambda_p t}(\lambda p e^{-\theta e^{-\lambda_p t}} - \lambda(1-p))}. \tag{4.43}$$

Using the first order Taylor expansion and dropping terms of lower order we get

$$\begin{aligned} L(t,\theta) &= \frac{\lambda(1-p)\left(-\theta e^{-\lambda_p t} + o(\theta e^{-\lambda_p t})\right) - e^{-\lambda_p t}\left(-\lambda p\theta e^{-\lambda_p t} + o(e^{-\lambda_p t}) + \lambda_p\right)}{\lambda p\left(-\theta e^{-\lambda_p t} + o(\theta e^{-\lambda_p t})\right) - e^{-\lambda_p t}\left(-\lambda p\theta e^{-\lambda_p t} + o(e^{-\lambda_p t}) + \lambda_p\right)} \\ &\approx \frac{-\lambda(1-p)\theta e^{-\lambda_p t} - \lambda_p e^{-\lambda_p t}}{-\lambda p\theta e^{-\lambda_p t} - \lambda_p e^{-\lambda_p t}}. \end{aligned}$$

Therefore

$$L(t,\theta) \to \frac{\lambda(1-p)\theta + \lambda_p}{\lambda p\theta + \lambda_p} = \frac{\theta - p(-2+\theta) - 1}{+p(2+\theta) - 1} =: L(\theta) \quad \text{as} \quad t \to +\infty.$$

Taking $\theta \to +\infty$ it is easy to check that $\mathbb{P}(V_\infty = 0) = p_e$ (we recall that $p_e = \frac{1-p}{p}$), therefore $V_\infty > 0$ on the set of non-extinction Ext^c. Let us now calculate the law of V_∞ on the set of non-extinction

$$L(\theta) = \mathbb{E}e^{-\theta V_\infty} = \mathbb{E}e^{-\theta V_\infty}1_{Ext} + \mathbb{E}e^{-\theta V_\infty}1_{Ext^c} = p_e + (1-p_e)\mathbb{E}\left(e^{-\theta V_\infty}|Ext^c\right).$$

Therefore

$$\mathbb{E}\left(e^{-\theta V_\infty}|Ext^c\right) = \frac{\lambda_p}{\lambda p\theta + \lambda_p} = \frac{(2p-1)}{p\theta + (2p-1)}.$$

Further

$$\begin{aligned}\mathbb{E}|X_t|^4 &= \frac{\partial^4 w(x,t,0)}{\partial \theta^4} \\ &= \frac{e^{t\lambda_p}\left(-1 + 2\left(-4 + 7e^{t\lambda_p}\right)p + \left(8 + 16e^{t\lambda_p} - 36e^{2t\lambda_p}\right)p^2 + 8e^{t\lambda_p}\left(-2 + 3e^{2t\lambda_p}\right)p^3\right)}{(-1+2p)^3}.\end{aligned}$$

Now the second part of the proposition follows. In order to prove that all moments are finite we notice that derivatives of (4.43) are of the form

$$\frac{\mathrm{d}^n}{\mathrm{d}\theta^n} L(t,\theta) = \frac{l(t,\theta)}{(\lambda p(e^{-\theta} - 1) - e^{-\lambda_p t}(\lambda p e^{-\theta} - \lambda(1-p)))^{2n}},$$

where $l(t,\theta)$ is a certain expression. Obviously the denominator is finite for $\theta = 0$ hence the proof is concluded by the properties of the Laplace transform (e.g. [16, Chapter XIII.2]). □

Proof of Proposition 4.3. Recall from Section 4.1 that we consider branching systems as Markov processes in the space $(\mathcal{M}, \Sigma)$ of integer valued measures. Since in what follows we will deal with branching processes in $\mathbb{R}^d$ and $\mathbb{R}^{2d}$ we will sometimes use the notation $(\mathcal{M}(\mathbb{R}^d), \Sigma(\mathbb{R}^d))$ or $(\mathcal{M}(\mathbb{R}^{2d}), \Sigma(\mathbb{R}^{2d}))$ to distinguish the corresponding spaces of measures.

We define projections $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^d$ by $\pi^1((x,y)) := x$ and $\pi^2((x,y)) := y$. Let us consider a branching particle system $\{\mathbf{X}_t\}_{t\geq 0}$ with particles evolving in $\mathbb{R}^{2d}$ according to the Markov process defined in Proposition 4.2 and starting from $\mathbf{x} = (x,y) \in \mathbb{R}^{2d}$ (we denote its semigroup by $\mathbf{T}$).

Its existence, just like in the case of the OU-branching particle system follows from the general theory in [11, 12]. The branching intensity is λ and the branching mechanism given by (4.3). By discussion in Section 4.1 the transition probabilities of $\{\mathbf{X}_t\}_{t\geq 0}$ as a measure-valued Markov process are determined by the solutions of equations

$$\mathbf{w}(\mathbf{x},t,\theta) = e^{-\lambda t}\mathbf{T}_t e^{-\theta f}(\mathbf{x}) + \lambda \int_0^t e^{-\lambda(t-s)}\mathbf{T}_{t-s}F(\mathbf{w}(\cdot,s,\theta))(\mathbf{x})\mathrm{d}s, \qquad (4.44)$$

where $\mathbf{x} \in \mathbb{R}^{2d}$, $f : \mathbb{R}^{2d} \mapsto \mathbb{R}$ is a smooth, bounded function and

$$\mathbf{w}(\mathbf{x},t,\theta) = \mathbb{E}_{\mathbf{x}} \exp(-\langle \mathbf{X}_t, \theta f\rangle). \qquad (4.45)$$

Further, we define new branching systems $\left\{X_t^1\right\}_{t\geq 0}$, $\left\{X_t^2\right\}_{t\geq 0}$, $\left\{X_t^3\right\}_{t\geq 0}$ by

$$X_t^1(i) := \pi^1(\mathbf{X}_t(i)), \quad X_t^2(i) := \pi^2(\mathbf{X}_t(i)), \quad X_t^3(i) := X_t^1(i) - X_t^2(i).$$

where i runs over all particles of $\mathbf{X}$ at time t. By the construction and Proposition 4.2 we have $X_t^3(i) = (x-y)e^{-\mu t}$.

Consider now an arbitrary smooth, bounded function $g \colon \mathbb{R}^d \to \mathbb{R}_+$ and let $w(x,t,\theta) := \mathbb{E}_x \exp(-\langle X_t, \theta g\rangle)$. Set also $f(\mathbf{x}) = g(x)$ for $\mathbf{x} = (x,y)$. Note that the function $\tilde{w}(\mathbf{x},t,\theta) = w(x,t,\theta)$ satisfies

$$\tilde{w}(\mathbf{x},t,\theta) = e^{-\lambda t}\mathbf{T}_t e^{-\theta f}(\mathbf{x}) + \lambda \int_0^t e^{-\lambda(t-s)}\mathbf{T}_{t-s}F(\tilde{w}(\cdot,s,\theta))(\mathbf{x})\mathrm{d}s.$$

Indeed, by definition the semigroup $\mathbf{T}$ acts on functions depending only on the variable x as the semigroup $\mathcal{T}_t$, so the above equation is just (4.2), written in the language of $\mathbf{T}$. Together with (4.44) this gives (via a standard application of Gronwall's inequality, see e.g. Appendix 5.1. in [15]) $\tilde{w}(x,t,\theta) = \mathbf{w}(\mathbf{x},t,\theta) = \mathbb{E}_{\mathbf{x}} \exp(-\langle X^1, \theta g\rangle)$. This implies that for all initial conditions x, y and times t, the distribution of the branching system X_t^1 is the same as the distribution of X_t and in particular does not depend on the coordinate y, (a fact which is intuitively obvious, when considered at the level of particle movements).

Using (4.4) we now obtain that the same is true for general starting measures, i.e. for any $A \in \Sigma(\mathbb{R}^d)$ and any integer valued measure $\mathbf{X}_0$ on $\mathbb{R}^{2d}$ we have

$$\mathbb{P}_{\mathbf{X}_0}(X_t^1 \in A) = \mathbb{P}_{\mathbf{X}_0^1}(X_t \in A). \tag{4.46}$$

Now, using the Markov property of the system $\mathbf{X}$ combined with the fact that $\mathbb{P}_{\mathbf{x}}$-a.e. for every t, $X_t^2(i) = X_t^1(i) + (y-x)e^{-\mu t}$, we get that X^1 has the Markov property and for every measurable set A we have $\mathbb{P}_{\mathbf{x}}$-a.e.

$$\mathbb{P}_{\mathbf{x}}(X_t^1 \in A | (X_u^1)_{u \le s}) = P_{t-s}(\mathbf{X}_s, A),$$

where for $\nu \in \mathcal{M}(\mathbb{R}^{2d})$, $P_t(\nu, A) := \mathbb{P}_\nu(X_t^1 \in A)$. By (4.46) this implies that

$$\mathbb{P}_{\mathbf{x}}(X_t^1 \in A | (X_u^1)_{u \le s}) = \mathbb{P}_{X_s^1}(X_{t-s} \in A),$$

i.e. X^1 is an OU-branching particle system.

An analogous argument proves the same property for X^2. This ends the proof since the fact that X is distributed as X^1 together with the equalities $X^2(i) = X_t^1(i) + (y-x)e^{-\mu t}$ and $Y_t(i) = X_t(i) + (y-x)e^{-\mu t}$ implies that Y is distributed as X^2, i.e. is an OU-branching system starting from y. □

References

[1] Radosław Adamczak and Piotr Miłoś, *U-statistics of Ornstein-Uhlenbeck branching particle system*, J. Theoret. Probab. **27** (2014) December 2014, Volume 27, no. 4, pp 1071-1111, arXiv:1111.4560

[2] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses [Panoramas and Syntheses], vol. 10, Société Mathématique de France, Paris, 2000, With a preface by Dominique Bakry and Michel Ledoux. MR-1845806

[3] S. Asmussen and H. Hering, *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*, Z. Wahrsch. und Verw. Gebiete **36** (1976), no. 3, 195–212. MR-0420889

[4] S. Asmussen and H. Hering, *Strong limit theorems for supercritical immigration-branching processes*, Math. Scand. **39** (1976), no. 2, 327–342. MR-0438498

[5] Krishna B. Athreya and Peter E. Ney, *Branching processes*, Springer-Verlag, New York, 1972. MR-0373040

[6] Vincent Bansaye, Jean-François Delmas, Laurence Marsalle, and Viet Chi Tran, *Limit theorems for Markov processes indexed by continuous time Galton-Watson trees*, Ann. Appl. Probab. **21** (2011), no. 6, 2263–2314. MR-2895416

[7] J. D. Biggins and A. E. Kyprianou, *Seneta-Heyde norming in the branching random walk*, Ann. Probab. **25** (1997), no. 1, 337–360. MR-1428512

[8] Giuseppe Da Prato, *An introduction to infinite-dimensional analysis*, Universitext, Springer-Verlag, Berlin, 2006, Revised and extended from the 2001 original by Da Prato. MR-2244975

[9] D. Dawson, *Measure-valued Markov processes*, Ecole d'Ete de Probabilites de Saint-Flour XXI-1991 (1993), 1–260. MR-1242575

[10] R. M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002, Revised reprint of the 1989 original. MR-1932358

[11] E. B. Dynkin, *Branching particle systems and superprocesses*, Ann. Probab. **19** (1991), no. 3, 1157–1194. MR-1112411

[12] Eugene B. Dynkin, *An introduction to branching measure-valued processes*, CRM Monograph Series, vol. 6, American Mathematical Society, 1994. MR-1280712

[13] Janos Englander, Simon C Harris, and Andreas E Kyprianou, *Strong law of large numbers for branching diffusions*, Ann. Inst. H. Poincaré **46** (2010), 279–298. MR-2641779

[14] A. Etheridge, *An introduction to superprocesses*, American Mathematical Society, 2000. MR-1779100

[15] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986, Characterization and convergence. MR-838085

[16] William Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971. MR-0270403

[17] J. F. Le Gall, *Spatial branching processes, random snakes, and partial differential equations*, Birkhauser, 1999. MR-1714707

[18] S. E. Graversen and G. Peskir, *Maximal inequalities for the Ornstein-Uhlenbeck process*, Proc. Amer. Math. Soc. **128** (2000), no. 10, 3035–3041. MR-1664394

[19] Simon C. Harris, *Convergence of a "Gibbs-Boltzmann" random measure for a typed branching diffusion*, Séminaire de Probabilités, XXXIV, Lecture Notes in Math., vol. 1729, Springer, Berlin, 2000, pp. 239–256. MR-1768067

[20] Gopinath Kallianpur and Jie Xiong, *Stochastic differential equations in infinite dimensional spaces*, IMS, 1995. MR-1465436

[21] Piotr Miłoś, *CLT for Ornstein-Uhlenbeck superprocess*, submitted to J. Th. Probab. arXiv:1203.6661

[22] I. Mitoma, *Tightness of probabilities on* $\mathcal{C}([0,1],\mathcal{S}')$ *and* $\mathcal{D}([0,1],\mathcal{S}')$, Ann. Probab. **11** (1983), 989–999. MR-714961

[23] Bernt Øksendal, *Stochastic differential equations*, fifth ed., Universitext, Springer-Verlag, Berlin, 1998, An introduction with applications. MR-1619188

[24] Yan-Xia Ren, Renming Song, Rui Zhang, *Central limit theorems for supercritical branching Markov processes*, J. Funct Anal, **266**, no. 3, (2014), 1716–1756. MR-3146834

[25] Yan-Xia Ren, Renming Song, Rui Zhang, *Central limit theorems for supercritical superprocesses*, Stoch. Proc. Appl., **125**, no. 2, (2015), 428–457.

[26] Walter Rudin, *Functional analysis*, McGraw-Hill Book Co., New York, 1973, McGraw-Hill Series in Higher Mathematics. MR-0365062

Acknowledgments. We would like to thank Dr Simon Harris for introducing to the topic. We are also grateful to the reviewers whose comments improved considerably the clarity and readability of this paper.