

INVARIANT MEASURES OF CRITICAL SPATIAL BRANCHING PROCESSES IN HIGH DIMENSIONS

BY MAURY BRAMSON,¹ J. T. COX² AND ANDREAS GREVEN²

*University of Wisconsin, Syracuse University
and Universität Erlangen–Nürnberg*

We consider two critical spatial branching processes on \mathbb{R}^d : critical branching Brownian motion, and the critical Dawson–Watanabe process. A basic feature of these processes is that their ergodic behavior is highly dimension dependent. It is known that in low dimensions, $d \leq 2$, the only invariant measure is δ_0 , the unit point mass on the empty state. In high dimensions, $d \geq 3$, there is a family $\{\nu_\theta, \theta \in [0, \infty)\}$ of extremal invariant measures; the measures ν_θ are translation invariant and indexed by spatial intensity. We prove here, for $d \geq 3$, that all invariant measures are convex combinations of these measures.

1. Introduction and main results. Critical branching Brownian motion and the critical Dawson–Watanabe process are two closely related models of random motion and branching on \mathbb{R}^d . In our previous paper [2], we proved that in low dimensions, $d \leq 2$, the only invariant measure for either process is the unit point mass on the empty configuration, thus extending previous work on this question. Here, we consider the high-dimensional case, $d \geq 3$, where it is known that there is a one-parameter family $\{\nu_\theta, \theta \in [0, \infty)\}$ of extremal invariant measures that are translation invariant and indexed by spatial intensity. We show that *all* invariant measures are obtained as convex combinations of this collection. We begin by defining our processes.

Critical branching Brownian motion η_t is a system of particles which undergo random motion and branching on \mathbb{R}^d according to the following rules.

1. Each particle lives an exponentially distributed lifetime with parameter $2b$.
2. At the end of its lifetime, a particle disappears and is replaced by zero or two particles, each possibility occurring with equal probability.
3. During its lifetime, a particle moves according to standard Brownian motion.
4. All particle lifetimes, motions and branching are independent of one another.

The parameter b is a positive, finite constant.

Received August 1994.

¹Research supported in part by the NSF.

²Research supported in part by the MSI and NSF.

AMS 1991 *subject classifications*. Primary 60K35; secondary 60J80.

Key words and phrases. Critical branching Brownian motion, critical Dawson–Watanabe process, invariant measures.

It is convenient to view η_t as a measure on \mathbb{R}^d , and to adopt the following notation. Let \mathcal{M} be the set of Borel measures η on \mathbb{R}^d such that $\eta(K) < \infty$ for all compact $K \subset \mathbb{R}^d$ (i.e., the set of Radon measures on \mathbb{R}^d). We endow \mathcal{M} with the topology of vague convergence: $\eta_n \rightarrow_v \eta$ if and only if $\langle \eta_n, f \rangle \rightarrow \langle \eta, f \rangle$ for all $f \in C_c^+(\mathbb{R}^d)$. Here $C_c^+(\mathbb{R}^d)$ is the collection of continuous nonnegative functions on \mathbb{R}^d with compact support, and $\langle \eta, f \rangle = \int_{\mathbb{R}^d} f(x)\eta(dx)$. Let δ_x denote the unit point mass at x , and let \mathcal{M}_0 be the collection of all $\eta \in \mathcal{M}$ of the form $\eta = \sum_i \delta_{x_i}$, where $\{x_i\}$ is a finite or countably infinite sequence of points in \mathbb{R}^d .

For $x \in \mathbb{R}^d$, let η_t^x denote the process starting from a single particle at x at time zero, that is, $\eta_0 = \delta_x$. A construction of these “single ancestor” processes η_t^x can be found in [9]. Systems of infinitely many branching Brownian motions are constructed by superposition. We refer the reader to [5] for an excellent account of the general theory (see also [6], [7] and [13]). As in [2], the following informal description will suffice for our purposes. Given $\zeta = \sum_i \delta_{x_i} \in \mathcal{M}_0$ and a family of independent single ancestor processes $\{\eta_t^{x_i}\}$, we define η_t^ζ , the process with initial configuration ζ , by

$$(1.1) \quad \eta_t^\zeta = \sum_i \eta_t^{x_i}.$$

This also works if ζ is random: we simply require that ζ and the family $\{\eta_t^{x_i}\}$ be independent. Letting \mathcal{L} denote law, we assume that $\mathcal{L}(\eta_0) \in \mathcal{F}(\mathcal{M}_0)$, the space of probability measures on \mathcal{M}_0 .

REMARK. \mathcal{M}_0 is the largest class of initial states that is sensible to consider, since if $\eta_0(K) = \infty$ for some compact K , then $P(\eta_t(U) = \infty) = 1$ for all $t > 0$ and open sets U . [Of course, $P(\eta_t(U) = \infty) = 1$ for all $t > 0$ and open sets U also holds for certain $\eta_0 \in \mathcal{M}_0$.]

The *critical Dawson–Dawson–Watanabe process* $\hat{\eta}_t$ of [4] is the “diffusion limit” of critical branching Brownian motion obtained by speeding up the branching rate, decreasing the mass of particles, and packing more particles into less space. If we let ${}^n\eta_t$ denote branching Brownian motion with lifetime parameter $2bn$, and suppose that $(1/n)({}^n\eta_0)$ converges to a measure $\hat{\eta}_0$ as $n \rightarrow \infty$, then $(1/n){}^n\eta_t$ converges to a measure-valued diffusion $\hat{\eta}_t$, the critical Dawson–Watanabe process. We will not give the details of this construction, but instead refer the interested reader to [5], [12], [17] and [19]. Although it is necessary to restrict $\hat{\eta}_0$ to certain subsets \mathcal{M}_p of \mathcal{M} in order that $\hat{\eta}_t$ be a well-defined right continuous process (see [12]), this restriction plays no role in our analysis. All we will need is the fact that the spaces $\mathcal{F}(\mathcal{M}_p)$, like $\mathcal{F}(\mathcal{M}_0)$, are Polish. (See Chapter 3 of [5] for more detail on \mathcal{M}_p .) Except for this point, we rely exclusively on the Laplace functional of $\hat{\eta}_t$, given in (1.5).

Much is known regarding the ergodic theory of η_t and $\hat{\eta}_t$. We will summarize certain relevant results in Theorem 0 below before stating our results. (See [6], [7], [10] and [11] for branching Brownian motion, and [4], [8] and [11] for the Dawson–Watanabe process.) First, some additional notation. In order to treat η_t and $\hat{\eta}_t$ simultaneously, we will use ζ_t to denote both

processes. We let \mathcal{I} be the set of invariant probability measures, that is, all $\mu \in \mathcal{A} \mathcal{M}_0$ [respectively, all $\mu \in \mathcal{A} \mathcal{M}_p$], such that $\mathcal{L}(\eta_0) = \mu$ implies $\mathcal{L}(\eta_t) = \mu$ for all $t > 0$ [respectively, $\mathcal{L}(\hat{\eta}_0) = \mu$ implies $\mathcal{L}(\hat{\eta}_t) = \mu$]. (Of course, \mathcal{I} for η_t is not the same set as \mathcal{I} for $\hat{\eta}_t$.) \mathcal{I} is convex, and we will let \mathcal{I}_e be the set of extreme points of \mathcal{I} . Let \mathcal{T} denote those $\mu \in \mathcal{A} \mathcal{M}$ which are translation invariant. That is, if ζ has law μ , and $x + \zeta$ is defined by setting $(x + \zeta)(U) = \zeta(x + U)$, then $\mu \in \mathcal{T}$ if and only if $x + \zeta$ has law μ for all $x \in \mathbb{R}^d$. For $\mu \in \mathcal{A} \mathcal{M}$, the Borel measure π on \mathbb{R}^d defined by $\pi(B) = \int \zeta(B) d\mu(\zeta)$ is called the *mean measure* of μ . The mean measure π for $\mu \in \mathcal{T}$ is a multiple of Lebesgue measure m , $\pi = \theta m$; we refer to θ as the *spatial intensity* of μ . We will use \Rightarrow to denote weak convergence.

THEOREM 0. *Assume $d \geq 3$. Both η_t and $\hat{\eta}_t$ have one-parameter families of invariant measures $\{\nu_\theta, 0 \leq \theta < \infty\}$. The measures ν_θ are translation invariant, shift ergodic and indexed by spatial intensity. If $\mathcal{L}(\zeta_0)$ is translation invariant and shift ergodic, with spatial intensity θ , then*

$$\mathcal{L}(\zeta_t) \Rightarrow \nu_\theta \quad \text{as } t \rightarrow \infty.$$

Furthermore, if $\mu \in \mathcal{I}$ has a σ -finite mean measure, then there is a unique probability measure F on $[0, \infty)$ such that

$$(1.2) \quad \mu = \int_{[0, \infty)} \nu_\theta F(d\theta).$$

One can inquire as to whether all extremal invariant measures are given by $\{\nu_\theta, 0 \leq \theta < \infty\}$, and whether (1.2) holds without the assumption that the mean measure of μ is σ -finite. Note that it is easy to construct a “natural” class of $\mu \in \mathcal{I}$ with non- σ -finite mean measures by employing F with infinite mean in (1.2). Then $\mu \in \mathcal{I}$, and

$$E^\mu \zeta(K) = \int_0^\infty E^{\nu_\theta} \zeta(K) F(d\theta) = \int_0^\infty \theta m(K) F(d\theta) = \infty,$$

provided $m(K) > 0$. Our problem, then, is to determine whether or not there exist elements of \mathcal{I} which cannot be represented via (1.2), even allowing F with infinite mean. We answer this with the following theorem.

THEOREM 1. *If $d \geq 3$, then for both η_t and $\hat{\eta}_t$, $\mathcal{I}_e = \{\nu_\theta, 0 \leq \theta < \infty\}$. If $\mu \in \mathcal{I}$, then there is a unique probability measure F on $[0, \infty)$ such that (1.2) holds.*

A key step in our proof of Theorem 1 is to rule out the possibility of an invariant measure which is not also translation invariant. Given that the dynamics of our processes are translation invariant, it does not seem unreasonable for an invariant measure also to be translation invariant. But this does not follow from any “soft” argument, and simply isn’t true in general. There are examples in the literature of infinite particle systems with transla-

tion-invariant dynamics which nonetheless possess nontranslation-invariant invariant measures. These include the shower process example in [20], the exclusion process with asymmetric jump matrix of Example 2.8 of [15], and the Ising model in three or more dimensions (see [1]). Nevertheless, we prove the following result. (The same statement should hold for a wide class of critical spatial branching systems.) The $d \leq 2$ version is contained in [2].

PROPOSITION 1. *For $d \geq 3$, $\mathcal{I} \subset \mathcal{T}$.*

With Proposition 1 in hand, our strategy is this. Assume $\mu \in \mathcal{I}$. Then since $\mu \in \mathcal{T}$, we can employ the ergodic theorem to decompose μ into a mixture of measures μ_θ with “spatial density” θ (see Proposition 5 in Section 5). A convergence result in [11] tells us that if $\mathcal{L}(\zeta_0) = \mu_\theta$, then $\mathcal{L}(\zeta_t) \Rightarrow \nu_\theta$ as $t \rightarrow \infty$. Thus, if $\mathcal{L}(\zeta_0) = \mu$, we find that $\mathcal{L}(\zeta_t)$ must converge to a mixture of the ν_θ . But this means μ is a mixture of the ν_θ , since $\mu \in \mathcal{I}$ implies $\mathcal{L}(\zeta_t) \equiv \mathcal{L}(\zeta_0) = \mu$.

Both critical branching Brownian motion η_t and the critical Dawson–Watanabe process $\hat{\eta}_t$ are intimately connected with the following partial differential equation. For a function $u(x, t)$, $x \in \mathbb{R}^d$, $t > 0$, consider the semilinear heat equation

$$(1.3) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \Delta u - bu^2,$$

and the two initial conditions

$$(1.4a) \quad u(x, 0) = 1 - e^{-f(x)}, \quad x \in \mathbb{R}^d,$$

$$(1.4b) \quad u(x, 0) = f(x), \quad x \in \mathbb{R}^d,$$

where f is a bounded, nonnegative measurable function on \mathbb{R}^d with compact support. It is well known that there is always a unique bounded solution of (1.3) satisfying either (1.4a) or (1.4b). The Laplace functionals of η_t and $\hat{\eta}_t$ can be given in terms of these solutions (see [5]). Namely, if $u(x, t)$ is the solution of (1.3) which satisfies (1.4a) in the branching Brownian motion case and (1.4b) in the Dawson–Watanabe case, then

$$(1.5a) \quad E \exp(-\langle \eta_t, f \rangle) = E \exp(\langle \eta_0, \log(1 - u(\cdot, t)) \rangle),$$

$$(1.5b) \quad E \exp(-\langle \hat{\eta}_t, f \rangle) = E \exp(-\langle \hat{\eta}_0, u(\cdot, t) \rangle).$$

In order to treat η_t and $\hat{\eta}_t$ simultaneously, we define

$$v(x, t) = \begin{cases} -\log(1 - u(x, t)), & \text{if } \zeta_t = \eta_t, \\ u(x, t), & \text{if } \zeta_t = \hat{\eta}_t. \end{cases}$$

With this notation, we can rewrite (1.5) as

$$(1.6) \quad E \exp(-\langle \zeta_t, f \rangle) = E \exp(-\langle \zeta_0, v(\cdot, t) \rangle).$$

We will sometimes write E^μ to denote expectation when $\mathcal{L}(\zeta_0) = \mu$.

The remainder of our paper is organized as follows. In Section 2, we obtain several estimates on the solutions of $u(x, t)$ of (1.3) using partial differential

equations methods. In Section 3, we present the precise convergence results we need to implement the argument sketched in the paragraph below Proposition 1. The proof of Proposition 1, which relies heavily on the p.d.e. estimates of Section 2, is given in Section 4. Finally, we complete the proof of Theorem 1 in Section 5.

REMARK. Although we have considered only the binary case of critical branching Brownian motion, Proposition 1 and Theorem 1 hold just as well whenever the branching mechanism is critical and has a finite second moment. The explicit form of the Gaussian kernel is used heavily in the proofs of these results. It is therefore not immediate how to extend the setting to that where the motion of particles is given by a particular stable law.

2. PDE estimates. As in our previous paper [2], we make use of the method of sub- and super-solutions, and of the maximum principle. Proposition 2 gives upper and lower bounds for $u(x, t)$ in terms of the heat kernel which are uniform in x and t . The proposition is the $d \geq 3$ version of Proposition 1 in [2]. Proposition 3, which compares $u(x, t)$ at “nearby” points x_1 and x_2 , has no counterpart in [2].

Let $p_t(x, dy) = p_t(x, y) dy$ denote the transition kernel of standard Brownian motion, $p_t(x, y) = (2\pi t)^{-d/2} \exp\{-(x-y)^2/2t\}$. We will sometimes write $p_t(x)$ for $p_t(x, 0)$. For $t \geq 1$, put $G(t) = \int_1^t p_s(0) ds$.

PROPOSITION 2. Assume $u(x, t)$ satisfies (1.3), in $d \geq 3$, for $u(x, 0) \in C_c^+(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} u(x, 0) dx > 0$. Then there are finite positive constants a, A , such that

$$(2.1) \quad ap_{t/2}(x) \leq u(x, t) \leq Ap_{2t}(x)$$

for all $x \in \mathbb{R}^d$ and $t \geq 4$.

REMARK. The estimate (2.1) differs slightly from the corresponding estimate in Proposition 1 of [2] for the $d \leq 2$ case, where the function $\phi(x, t)$ is used. Up to constants, $\phi(x, t)$ is in all dimensions of the form $p_t(x)/G(t)$.

PROOF. The upper bound is simple and holds in all dimensions. To see this, set $\bar{u}(x, t) = A'p_t(x)$. Clearly,

$$\frac{\partial \bar{u}}{\partial t} - \frac{1}{2}\Delta \bar{u} + b\bar{u}^2 = b\bar{u}^2 > 0.$$

Choose A' so that $u(x, 0) \leq \bar{u}(x, 1)$ for all x . This can be done since $\text{supp}(u(\cdot, 0))$ is bounded. A standard maximum principle (see Chapter 3 of [18]) implies $u(x, t) \leq \bar{u}(x, t+1)$ for all $x \in \mathbb{R}^d$ and $t \geq 1$. Since $\bar{u}(x, t+1) \leq 2^{d/2}\bar{u}(x, 2t)$ for $t \geq 1$, we have for large enough A ,

$$(2.2) \quad u(x, t) \leq Ap_{2t}(x), \quad x \in \mathbb{R}^d, t \geq 2.$$

Next, take $a' \in (0, 1/b)$, and for $t \geq 1$, set

$$\underline{u}(x, t) = \frac{a' p_t(x)}{G(t)}.$$

Direct calculation shows

$$\frac{\partial \underline{u}}{\partial t} - \frac{1}{2} \Delta \underline{u} + b \underline{u}^2 = \frac{\underline{u}(x, t)}{G(t)} [a' b p_t(x) - p_t(0)] \leq 0, \quad x \in \mathbb{R}^d, t > 1.$$

If we can choose a' small enough so that

$$(2.3) \quad u(x, 4) \geq \underline{u}(x, 2), \quad x \in \mathbb{R}^d,$$

it will follow from the maximum principle that $u(x, t) \geq \underline{u}(x, t-2)$ for all $x \in \mathbb{R}^d$ and $t \geq 4$. Since $G(t)$ converges to a finite limit, in $d \geq 3$, as $t \rightarrow \infty$, this implies that there is a positive a such that

$$(2.4) \quad u(x, t) \geq a p_{t/2}(x), \quad x \in \mathbb{R}^d, t \geq 4.$$

The verification of (2.3) is straightforward. Put $f(x) = u(x, 0)$, and choose $c \in (0, 1)$ such that $cf(x) \leq 1$ for all $x \in \mathbb{R}^d$. If we let

$$\sigma(x, t) = ce^{-bt} \int_{\mathbb{R}^d} p_t(x, y) f(y) dy,$$

then $\sigma(x, 0) \leq u(x, 0)$ and $\sigma(x, t) \leq 1$ for all $x \in \mathbb{R}^d$ and $t \geq 0$. Thus,

$$\frac{\partial \sigma}{\partial t} - \frac{1}{2} \Delta \sigma + b \sigma^2 = b \sigma (\sigma - 1) a \leq 0.$$

By the maximum principle, $u(x, t) \geq \sigma(x, t)$, and consequently

$$u(x, t) \geq ce^{-bt} \left(\int_{\mathbb{R}^d} f(y) dy \right) \left(\inf_{y \in \text{supp}(f)} p_t(x, y) \right), \quad x \in \mathbb{R}^d, t \geq 0.$$

Since $\text{supp}(f)$ is bounded, it is easy to see that we can choose $\varepsilon > 0$ such that

$$\inf_{y \in \text{supp}(f)} p_4(x, y) \geq \varepsilon p_2(x, 0), \quad x \in \mathbb{R}^d.$$

By combining the last two inequalities, we obtain (2.3). \square

For $r > 0$, let $B(r) = \{x \in \mathbb{R}^d: |x| < r\}$. The following result is a key ingredient in the proof of $\mathcal{I} \subset \mathcal{T}$, which is given in Section 4.

PROPOSITION 3. *Assume $u(x, t)$ satisfies (1.3) for $u(x, 0) \in C_c^+(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} u(x, 0) dx > 0$, any $d \geq 1$. Fix $\varepsilon > 0$ and $M > 0$ such that $\varepsilon M < 1/4$. For all sufficiently large t and for all $x_1, x_2 \in B(\varepsilon t)$ such that $|x_1 - x_2| \leq M$,*

$$(2.5) \quad u(x_2, t) \leq \exp(8\sqrt{\varepsilon M}) u(x_1, t).$$

PROOF. Our strategy for deriving (2.5) is as follows. For an appropriately small $\delta > 0$ we will estimate $u(x, t)$ by “turning off” the branching term in (1.3) over the time period $[(1 - \delta)t, t]$. This does not affect the value of $u(x, t)$

very much, but allows us to make use of the smoothing properties of the Brownian transition kernel when comparing $u(x_1, t)$ and $u(x_2, t)$. We will make repeated use of sub- and super-solutions in the context of the maximum principle in the comparisons that follow. Let $\psi(x, t)$ and $w(x, t)$ satisfy

$$(2.6) \quad \frac{\partial \psi}{\partial t} = \frac{1}{2} \Delta \psi$$

and

$$\frac{\partial w}{\partial t} = \frac{1}{2} \Delta w + k(t) w,$$

with k continuous and bounded above, where $w(x, 0) = \psi(x, 0)$. We will also employ the standard representation

$$(2.7) \quad w(x, t) = \psi(x, t) \exp \left\{ \int_0^t k(s) ds \right\}.$$

Step 1. Let $\psi(x, 0) = u(x, 0)$ and $K = \sup_{x \in \mathbb{R}^d} u(x, 0)$. We claim that

$$(2.8) \quad \frac{1}{1 + bKt} \psi(x, t) \leq u(x, t) \leq \psi(x, t), \quad x \in \mathbb{R}^d, t \geq 0.$$

The right-hand inequality follows immediately from the maximum principle. For the left-hand inequality, let $U(x, t)$ be the solution of

$$\frac{\partial U}{\partial t} = -bU^2, \quad U(x, 0) = K \quad \text{for all } x \in \mathbb{R}^d.$$

Then $U(x, t) = K/(1 + bKt)$, and again by the maximum principle, $u(x, t) \leq U(x, t)$ for $x \in \mathbb{R}^d$ and $t \geq 0$. Plugging this bound into (1.3), we have

$$(2.9) \quad \frac{\partial u}{\partial t} \geq \frac{1}{2} \Delta u - \frac{bK}{1 + bKt} u, \quad x \in \mathbb{R}^d, t > 0.$$

Let $w(x, t)$ satisfy

$$(2.10) \quad \frac{\partial w}{\partial t} = \frac{1}{2} \Delta w - \frac{bK}{1 + bKt} w, \quad w(\cdot, 0) = u(\cdot, 0).$$

By (2.7),

$$\begin{aligned} w(x, t) &= \psi(x, t) \exp \left\{ - \int_0^t \frac{bK}{1 + bKs} ds \right\} \\ &= \frac{1}{1 + bKt} \psi(x, t). \end{aligned}$$

Another application of the maximum principle implies $u(x, t) \geq w(x, t)$ for all $x \in \mathbb{R}^d$ and $t \geq 0$, and we have established (2.8).

Step 2. Put $\delta = \sqrt{\varepsilon M} < 1/2$. Fix $t > 0$, and define $\hat{u}(x, s)$ by $\hat{u}(x, s) = u(x, s)$ for $s \leq (1 - \delta)t$, and

$$\frac{\partial \hat{u}}{\partial s} = \frac{1}{2} \Delta \hat{u}, \quad s > (1 - \delta)t.$$

In this step, we establish

$$(2.11) \quad e^{-2\delta} \hat{u}(x, t) \leq u(x, t) \leq \hat{u}(x, t), \quad x \in \mathbb{R}^d, t \geq 0.$$

The right-hand inequality in (2.11) is an immediate application of the maximum principle. For the left-hand inequality, we have, by (2.9),

$$(2.12) \quad \frac{\partial u}{\partial s} \geq \frac{1}{2} \Delta u - \frac{2}{t} u, \quad s > \frac{t}{2}.$$

Letting $\tilde{u}(x, s) = u(x, s)$ for $s \leq (1 - \delta)t$, and

$$\frac{\partial \tilde{u}}{\partial s} = \frac{1}{2} \Delta \tilde{u} - \frac{2}{t}, \quad s > (1 - \delta)t,$$

it follows from the maximum principle and (2.7), that

$$u(x, t) \geq \tilde{u}(x, t) = \hat{u}(x, t) \exp\left\{-\int_0^{\delta t} \frac{2}{t} ds\right\} = e^{-2\delta} \hat{u}(x, t).$$

Step 3. We will now demonstrate (2.5). Assume $x_1, x_2 \in B(\varepsilon t)$ and $|x_1 - x_2| \leq M$. Also, take t large enough so that $\text{supp}(u(\cdot, 0)) \subset B(\varepsilon t)$. By (2.11),

$$u(x_2, t) \leq \hat{u}(x_2, t) = \int_{\mathbb{R}^d} u(y, (1 - \delta)t) p_{\delta t}(y, x_2) dy = I_1 + I_2,$$

where I_1 denotes the integral over $B(4\varepsilon t)$, and I_2 denotes the integral over $B^c(4\varepsilon t)$. We will show that we may replace x_2 with x_1 in I_1 with little error, and that I_2 is negligible.

To calculate I_1 , we first note that for $y \in B(4\varepsilon t)$, the triangle inequality implies

$$\left| |y - x_1|^2 - |y - x_2|^2 \right| = \left| |y - x_1| + |y - x_2| \right| \cdot \left| |y - x_1| - |y - x_2| \right| \leq 10\varepsilon tM.$$

This implies

$$\frac{p_{\delta t}(y, x_2)}{p_{\delta t}(y, x_1)} \leq \exp\left(\frac{5\varepsilon M}{\delta}\right), \quad y \in B(4\varepsilon t).$$

With this estimate, we obtain

$$\begin{aligned} I_1 &= \int_{B(4\varepsilon t)} u(y, (1 - \delta)t) p_{\delta t}(y, x_1) \frac{p_{\delta t}(y, x_2)}{p_{\delta t}(y, x_1)} dy \\ &\leq \exp\left(\frac{5\varepsilon M}{\delta}\right) \int_{B(4\varepsilon t)} u(y, (1 - \delta)t) p_{\delta t}(y, x_1) dy \\ &\leq \exp\left(\frac{5\varepsilon M}{\delta}\right) \hat{u}(x_1, t). \end{aligned}$$

Using (2.11) again, it follows that

$$(2.13) \quad I_1 \leq \exp\{2\delta + 5\varepsilon M/\delta\} u(x_1, t).$$

Turning now to I_2 , for $x \in B(\varepsilon t)$ and $y \in B^c(4\varepsilon t)$, $|y - x|^2 \geq 9\varepsilon^2 t^2$. Thus, using (2.8),

$$\begin{aligned} u(y, (1 - \delta)t) &\leq \psi(y, (1 - \delta)t) \\ &= \int_{B(\varepsilon t)} u(x, 0) p_{(1-\delta)t}(x, y) dx \\ &\leq Km(B(\varepsilon t)) \frac{1}{(2\pi(1 - \delta)t)^{d/2}} \exp\left(-\frac{9\varepsilon^2 t}{2}\right) \\ &\leq \exp(-4\varepsilon^2 t) \end{aligned}$$

for large enough t . This implies

$$(2.14) \quad I_2 \leq \exp(-4\varepsilon^2 t) \int_{\mathbb{R}^d} p_{\delta t}(y, x_2) dy = \exp(-4\varepsilon^2 t).$$

To see that I_2 is negligible compared to $u(x_1, t)$, we use (2.8) again to obtain

$$\begin{aligned} u(x_1, t) &\geq \frac{\psi(x_1, t)}{1 + bKt} \\ &= \frac{1}{1 + bKt} \int_{B(\varepsilon t)} u(y, 0) p_t(y, x_1) dy \\ &\geq \frac{1}{1 + bKt} (2\pi t)^{-d/2} \exp(-2\varepsilon^2 t) \int_{B(\varepsilon t)} u(y, 0) dy \\ &\geq \exp(-3\varepsilon^2 t) \end{aligned}$$

for sufficiently large t . This estimate and (2.14) imply

$$(2.15) \quad I_2 \leq \exp(-\varepsilon^2 t) u(x_1, t).$$

Putting together the estimates (2.13) and (2.15),

$$\begin{aligned} u(x_2, t) &\leq I_1 + I_2 \\ &\leq \exp\{2\delta + 5\varepsilon M/\delta\} u(x_1, t) + \exp(-\varepsilon^2 t) u(x_1, t) \\ &\leq \exp\{2\delta + 6\varepsilon M/\delta\} u(x_1, t) \end{aligned}$$

for large enough t . Since $\delta = \sqrt{\varepsilon M}$, this implies

$$u(x_2, t) \leq \exp\{8\sqrt{\varepsilon M}\} u(x_1, t),$$

and we are done. \square

3. Convergence criteria. We need more information concerning the convergence of $\mathcal{L}(\zeta_t)$ as $t \rightarrow \infty$ than is given in Theorem 0. The results we need are stated in Proposition 4 below. We say that ζ_t is *stable* if for every ball B of finite positive radius, $\zeta_t(B)$ is stochastically bounded as $t \rightarrow \infty$. That is, ζ_t is stable if for any given B and $\varepsilon > 0$, there exists $M < \infty$ such that $\limsup_{t \rightarrow \infty} P(\zeta_t(B) \geq M) \leq \varepsilon$. Otherwise, ζ_t is *unstable*. We say that $\mu \in \mathcal{A}(\mathcal{M})$ has *spatial density* $\theta \in [0, \infty]$, if

$$\langle \zeta, p_t f \rangle \rightarrow_p \theta \langle m, f \rangle \quad \text{as } t \rightarrow \infty$$

for every $f \in C_c^+(\mathbb{R}^d)$ when ζ has law μ . [Or, equivalently, $\int p_t(x, 0)\zeta(dx) \rightarrow_p \theta$, although we only use the first formulation.] Here \rightarrow_p denotes convergence in probability, and $(p_t f)(x) = \int_{\mathbb{R}^d} p_t(x, y)f(y) dy$. Thus, we determine the spatial density of μ not by averaging over large boxes in \mathbb{R}^d , but by averaging with respect to $p_t(x, \cdot)$ for large t . This method of averaging has been used in the study of several interacting particle systems, including the voter model and certain “linear” systems (see [15]). Finally, recall the collection $\{\nu_\theta, 0 \leq \theta < \infty\}$ of invariant measures from Theorem 0.

PROPOSITION 4. Assume $d \geq 3$. (a) If ζ_0 has spatial density $\theta \in [0, \infty)$, then

$$\mathcal{L}(\zeta_t) \Rightarrow \nu_\theta \quad \text{as } t \rightarrow \infty.$$

(b) If ζ_t is stable, then $\langle \zeta_0, p_t(\cdot) \rangle$ is stochastically bounded as $t \rightarrow \infty$.

Part (a) corresponds to one direction of Theorem 1 of [2], although the argument for the case $d \geq 3$ is simpler. For $\zeta_t = \eta_t$, part (a) is a special case of Theorem 2.2 of [11], which concerns a general class of critical branching particle systems whose underlying spatial motion mechanisms are transient. The proof in [11] uses Laplace functionals and works equally well for $\hat{\eta}_t$. A similar convergence result was obtained in [3] for branching Markov chains, where the proof is somewhat easier. Part (a) will be employed in (5.5). It is a useful variant of Theorem 0, since ergodicity of $\mathcal{L}(\eta_0)$ is not assumed here. Part (b) of Proposition 4 is similar to the assertion in Step 1 in the proof of Theorem 1 of [2]. Before starting the proof of (b), we derive a pair of inequalities that will be useful below.

Let $f \in C_c^+(\mathbb{R}^d)$, and let $u(x, t)$ satisfy (1.3), and (1.4a) for $\zeta_t = \eta_t$ and (1.4b) for $\zeta_t = \hat{\eta}_t$. Recall the definition of $v(x, t)$ given just before (1.6). By Proposition 2, for some finite constant C , $u(x, t) \leq C/t^{d/2}$ for all $x \in \mathbb{R}^d$ and $t \geq 4$. This fact and the elementary inequality $s \leq -\log(1 - s) \leq 2s$ for small positive s imply that for large t ,

$$(3.1) \quad u(x, t) \leq v(x, t) \leq 2u(x, t), \quad x \in \mathbb{R}^d.$$

Furthermore, since $\lim_{s \rightarrow 0} \log(1 - s)/s = -1$,

$$(3.2) \quad \lim_{t \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \left| \frac{v(x, t)}{u(x, t)} - 1 \right| = 0.$$

PROOF OF PROPOSITION 4(b). For $M > 0$, put $\Gamma_t(M) = \{\langle \zeta_0, p_t(\cdot) \rangle \geq M\}$. By (2.1) and (3.1), for $f \in C_c^+(\mathbb{R}^d)$ with $\langle m, f \rangle > 0$, there exists $a > 0$ such that $v(x, t) \geq ap_{t/2}(x)$ for large t . For such t ,

$$\begin{aligned} E \exp(-\langle \zeta_t, f \rangle) &= E \exp(-\langle \zeta_0, v(\cdot, t) \rangle) \\ &\leq E \exp(-a \langle \zeta_0, p_{t/2}(\cdot) \rangle) \\ &\leq 1 - P(\Gamma_{t/2}(M)) + \exp(-aM) P(\Gamma_{t/2}(M)), \end{aligned}$$

that is,

$$P(\Gamma_t(M)) \leq \frac{1 - E \exp(-\langle \zeta_{2t}, f \rangle)}{1 - \exp(-aM)}.$$

Fix $\varepsilon > 0$. Since ζ_t is stable, we can choose f so that $E \exp(-\langle \zeta_t, f \rangle) \geq 1 - \varepsilon$ for all large t . If we now choose M large enough so that $\exp(-aM) < 1/2$, then

$$P(\Gamma_t(M)) \leq 2\varepsilon$$

for all large t . This shows $\langle \zeta_0, p_t(\cdot) \rangle$ is stochastically bounded as $t \rightarrow \infty$. \square

4. Proof of Proposition 1. Fix $\mu \in \mathcal{I}$. For a function f on \mathbb{R}^d and $z \in \mathbb{R}^d$, let f_z be the translate $f_z(x) = f(x+z)$. To prove $\mu \in \mathcal{I}$, we will show that if $\mathcal{L}(\zeta_0) = \mu$, then for all $f \in C_c^+(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$,

$$(4.1) \quad E \exp(-\langle \zeta_t, f_z \rangle) - E \exp(-\langle \zeta_t, f \rangle) \rightarrow 0$$

at $t \rightarrow \infty$. This suffices since both $E \exp(-\langle \zeta_t, f_z \rangle)$ and $E \exp(-\langle \zeta_t, f \rangle)$ do not depend on t . To prove (4.1), we invoke (1.6) and compare $\int v(x, t) \zeta_0(dx)$ with $\int v(x+z, t) \zeta_0(dx)$. In Step 1, we use Proposition 4(b) to show that we may neglect the contribution of “large” x (which depend on t) in these integrals. In Step 2, we apply Proposition 3 to handle all remaining x .

Step 1. Fix $\delta > 0$ and let A_t be the event

$$A_t = \left\{ \int_{B^c(t^{3/4})} v(x, t) \zeta_0(dx) > \delta \right\}.$$

We show that

$$(4.2) \quad P(A_t) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

For $x \in B^c(t^{3/4})$,

$$p_t(x) \leq 2^{d/2} \exp\{-\frac{1}{4}t^{1/2}\} p_{2t}(x),$$

and thus

$$\int_{B^c(t^{3/4})} p_t(x) \zeta_0(dx) \leq 2^{d/2} \exp\{-\frac{1}{4}t^{1/2}\} \int_{\mathbb{R}^d} p_{2t}(x) \zeta_0(dx).$$

Because μ is invariant, ζ_t must be stable, and thus by Proposition 4(b), the right-hand side above must tend to 0 as $t \rightarrow \infty$. Consequently,

$$\int_{B^c(t^{3/4})} p_t(x) \zeta_0(dx) \rightarrow_p 0 \quad \text{as } t \rightarrow \infty.$$

By Proposition 2 and (3.1), this implies

$$\int_{B^c(t^{3/4})} v(x, t) \zeta_0(dx) \rightarrow_p 0 \quad \text{as } t \rightarrow \infty,$$

which proves (4.2).

Step 2. We now prove (4.1). By (1.6),

$$\begin{aligned}
 (4.3) \quad E \exp(-\langle \zeta_t, f_z \rangle) &= E \exp(-\langle \zeta_0, v_z(\cdot, t) \rangle) \\
 &= E \mathbf{1}_{A_t} \exp(-\langle \zeta_0, v_z(\cdot, t) \rangle) + E \mathbf{1}_{A_t^c} \exp(-\langle \zeta_0, v_z(\cdot, t) \rangle) \\
 &\leq P(A_t) + E \mathbf{1}_{A_t^c} \exp\left\{-\int_{B(t^{3/4})} v(x+z, t) \zeta_0(dx)\right\}.
 \end{aligned}$$

Set

$$\rho_t = \inf_{x \in B(t^{3/4})} \frac{v(x+z, t)}{v(x, t)}.$$

It follows from (3.2) and Proposition 3 that $\rho_t \rightarrow 1$ as $t \rightarrow \infty$. Since $v(x+z, t) \geq \rho_t v(x, t)$ for $x \in B(t^{3/4})$, and since $\delta \geq \int_{B^c(t^{3/4})} v(x, t) \zeta_0(dx)$ on A_t^c ,

$$\begin{aligned}
 (4.4) \quad &E \mathbf{1}_{A_t^c} \exp\left\{-\int_{B(t^{3/4})} v(x+z, t) \zeta_0(dx)\right\} \\
 &\leq E \mathbf{1}_{A_t^c} \exp\left\{-\rho_t \int_{B(t^{3/4})} v(x, t) \zeta_0(dx) \right. \\
 &\quad \left. + \rho_t \left[\delta - \int_{B^c(t^{3/4})} v(x, t) \zeta_0(dx) \right] \right\} \\
 &= \exp(\delta \rho_t) E \mathbf{1}_{A_t^c} \exp(-\rho_t \langle \zeta_0, v(\cdot, t) \rangle) \\
 &\leq \exp(\delta \rho_t) E \exp(-\rho_t \langle \zeta_0, v(\cdot, t) \rangle).
 \end{aligned}$$

The elementary inequality

$$|e^{-\rho s} - e^{-s}| \leq 1 - \rho, \quad s \geq 0, \quad 0 \leq \rho \leq 1,$$

implies

$$(4.5) \quad E \exp(-\rho_t \langle \zeta_0, v(\cdot, t) \rangle) \leq E \exp(-\langle \zeta_0, v(\cdot, t) \rangle) + 1 - \rho_t.$$

By (1.6),

$$(4.6) \quad E \exp(-\langle \zeta_0, v(\cdot, t) \rangle) = E \exp(-\langle \zeta_t, f \rangle).$$

Taken together, (4.3)–(4.6) imply

$$(4.7) \quad E \exp(-\langle \zeta_t, f_z \rangle) \leq P(A_t) + \exp(\delta \rho_t) [E \exp(-\langle \zeta_t, f \rangle) + 1 - \rho_t].$$

Since $P(A_t) \rightarrow 0$ by (4.2), $\rho_t \rightarrow 1$ as $t \rightarrow \infty$ and $\delta > 0$ is fixed but arbitrary, (4.7) implies

$$\limsup_{t \rightarrow \infty} E \exp(-\langle \zeta_t, f_z \rangle) \leq \liminf_{t \rightarrow \infty} E \exp(-\langle \zeta_t, f \rangle).$$

The inequality in the opposite direction follows analogously by translating by $-z$. This proves (4.1). \square

5. Proof of Theorem 1. We are almost ready to prove Theorem 1. First, we need a general decomposition result for translation invariant measures. The following version will suffice.

PROPOSITION 5. *Let $\mu \in \mathcal{T}$, any d . There is a Borel probability measure F on $[0, \infty]$ and probability measures $\mu_\theta \in \mathcal{T}$, $\theta \in [0, \infty]$, such that each μ_θ has spatial density θ , and*

$$(5.1) \quad \mu = \int_{[0, \infty]} \mu_\theta F(d\theta).$$

PROOF. We give the main steps of the argument, which in its entirety is rather tedious. Put $\Lambda_N = [-N, N]^d$ and $h_N(x) = 1_{\Lambda_N}(x)/m(\Lambda_N)$. (Note that h_N is just the average of N^d translations of h_1 .) Let ζ have law μ . By a standard ergodic theorem (see Chapter 6 of [14] or Section 4 of [16]), if $E\langle \zeta, h_1 \rangle < \infty$, then

$$\langle \zeta, h_N \rangle \rightarrow D \quad \text{a.s. and in } L^1 \text{ as } N \rightarrow \infty,$$

where D is a nonnegative shift-invariant random variable, with $ED = E\langle \zeta, h_1 \rangle$. A truncation argument can be used to show that even if $E\langle \zeta, h_1 \rangle = \infty$,

$$\langle \zeta, h_N \rangle \rightarrow D \quad \text{a.s. as } N \rightarrow \infty,$$

where D is nonnegative and shift invariant, but may be infinite with positive probability. Set F equal to the distribution function of D . Recall that $\mathcal{F}(\mathcal{M})$ and $\mathcal{F}(\mathcal{M}_\rho)$ are Polish spaces. By standard results on regular conditional probabilities (see pages 13–17 of [21]), we can define probability measures

$$\mu_\theta(\cdot) = \mu(\cdot | D = \theta), \quad \theta \in [0, \infty],$$

such that (5.1) holds. Let ζ^θ have law μ_θ . Except for θ in an F -null set, ζ^θ will be translation invariant, with

$$(5.2) \quad \langle \zeta^\theta, h_N \rangle \rightarrow \theta \quad \text{a.s. as } N \rightarrow \infty.$$

Of course, μ_θ can be redefined on the F -null set so that translation invariance and (5.2) still hold [without affecting (5.1)].

It remains to show that μ_θ always has spatial density θ . By approximating $p_t(x)$ uniformly above and below with sums of the form $\sum c_N(t)h_N(x)$, it is not difficult to show from (5.2) that

$$\langle \zeta^\theta, p_t(\cdot) \rangle \rightarrow \theta \quad \text{a.s. as } t \rightarrow \infty.$$

From this and the translation invariance of ζ^θ , one can check that $\langle \zeta^\theta, p_t f \rangle \rightarrow \theta \langle m, f \rangle$ a.s. as $t \rightarrow \infty$ for all $f \in C_c^+(\mathbb{R}^d)$. That is, μ_θ has spatial density θ . \square

PROOF OF THEOREM 1. Fix $\mu \in \mathcal{I}$, and let ζ_0 have law μ . By Proposition 1, $\mu \in \mathcal{T}$, and by Proposition 5, μ may be represented as in (5.1). Since $\mu \in \mathcal{I}$, μ is stable, and so by Proposition 4(b), $\langle \zeta_0, p_t(\cdot) \rangle$ is stochastically bounded as $t \rightarrow \infty$. It follows that $F(\{\infty\}) = 0$. To see this, choose $f \in C_c^+(\mathbb{R}^d)$ such that $\langle m, f \rangle > 0$ and $f(y) \leq 1_{B(1)}(y)$. For large t , $p_t(x, y) \leq 2^d p_{s_t}(x, 0)$ for $y \in B(1)$ and all x . So for large t ,

$$\langle \zeta_0, p_t f \rangle \leq 2^d \langle \zeta_0, p_{s_t}(\cdot) \rangle.$$

Since the right-hand side is stochastically bounded as $t \rightarrow \infty$, the left-hand side must be too. Consequently, $F(\infty) = 0$, or equivalently,

$$(5.3) \quad \mu = \int_{[0, \infty)} \mu_\theta F(d\theta).$$

For given $f \in C_c^+(\mathbb{R}^d)$ and $t \geq 0$, (5.3) implies

$$(5.4) \quad E^\mu \exp(-\langle \zeta_t, f \rangle) = \int_{[0, \infty)} E^{\mu_\theta} \exp(-\langle \zeta_t, f \rangle) F(d\theta).$$

By Proposition 4(a),

$$(5.5) \quad E^{\mu_\theta} \exp(-\langle \zeta_t, f \rangle) \rightarrow E^{\nu_\theta} \exp(-\langle \zeta, f \rangle) \quad \text{as } t \rightarrow \infty.$$

Using the bounded convergence theorem, (5.4) and (5.5), we obtain

$$(5.6) \quad E^\mu \exp(-\langle \zeta_t, f \rangle) \rightarrow \int_{[0, \infty)} E^{\nu_\theta} \exp(-\langle \zeta, f \rangle) F(d\theta) \quad \text{as } t \rightarrow \infty.$$

Since $\mu \in \mathcal{I}$, the left-hand side above does not depend on t , and thus

$$E^\mu \exp(-\langle \zeta, f \rangle) = \int_{[0, \infty)} E^{\nu_\theta} \exp(-\langle \zeta, f \rangle) F(d\theta), \quad f \in C_c^+(\mathbb{R}^d),$$

which is equivalent to (1.2).

It remains to check that $\mathcal{I}_e = \{\nu_\theta, 0 \leq \theta < \infty\}$; the uniqueness of F then follows. Clearly, (1.2) implies $\mathcal{I}_e \subset \{\nu_\theta, 0 \leq \theta < \infty\}$. On the other hand, each ν_θ is shift ergodic (recall Theorem 0). This fact and a standard argument (see Proposition I.4.13 of [15]) imply each ν_θ is extremal, so we are done. \square

REFERENCES

- [1] AIZENMAN, M. (1980). Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.* **73** 83–94.
- [2] BRAMSON, M., COX, J. T. and GREVEN, A. (1993). Ergodicity of critical spatial branching processes in low dimensions. *Ann. Probab.* **21** 1946–1957.
- [3] COX, J. T. (1994). On the ergodic theory of critical branching Markov chains. *Stochastic Process. Appl.* **50** 1–20.
- [4] DAWSON, A. (1977). The critical measure diffusion process. *Z. Wahrsch. Verw. Gebiete* **40** 125–145.
- [5] DAWSON, D. A. (1993). Measure-valued Markov processes. *Ecole d'été de Probabilités de Saint-Flour, Lecture Notes in Math.* **1541** 1–260. Springer, New York.
- [6] DAWSON, D. A. and IVANOFF, B. G. (1978). Branching diffusions and random measures. *Advances in Probability: Branching Processes* 61–109.
- [7] DURRETT, R. (1979). An infinite particle system with additive interactions. *Adv. in Appl. Probab.* **11** 355–383.
- [8] DYNKIN, E. B. (1989). Three classes of infinite dimensional diffusions. *J. Funct. Anal.* **86** 75–110.
- [9] ETHIER, S. and KURTZ, T. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.
- [10] FLEISCHMAN, J. (1978). Limiting distributions for branching random fields. *Trans. Amer. Math. Soc.* **239** 353–389.
- [11] GOROSTIZA, L. G. and WAKOLBINGER, A. (1994). Long time behavior of critical branching particle systems and applications. *CRM Proceedings and Lecture Notes* **5** 119–138.

- [12] ISCOE, I. (1986). A weighted occupation time for a class of measure-valued branching processes. *Probab. Theory Related Fields* **71** 85–116.
- [13] KERSTAN, J., MATTHES, K. and MECKE, J. (1978). *Infinitely Divisible Point Processes*, Wiley, New York.
- [14] KRENGEL, U. (1985). *Ergodic Theorems*. de Gruyter, Berlin.
- [15] LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.
- [16] NGUYEN, X. X. and ZESSIN, H. (1979). Ergodic theorems for spatial processes. *Z. Wahrsch. Verw. Gebiete* **48** 133–158.
- [17] PERKINS, E. A. (1988). A space-time property of a class of measure-valued branching diffusions. *Trans. Amer. Math. Soc.* **305** 743–794.
- [18] PROTTER, M. H. and WEINBERGER, H. F. (1967). *Maximum Principles in Differential Equations*. Prentice-Hall, Englewood Cliffs, NJ.
- [19] ROELLY-COPPOLETTA, S. (1986). A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics* **17** 43–65.
- [20] SIEGMUND-SCHULTZE, RA. (1991). A subcritical spatially homogeneous branching process may have an equilibrium. *Math. Nachr.* **151** 101–120.
- [21] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.

DEPARTMENT OF MATHEMATICS
 UNIVERSITY OF WISCONSIN
 MADISON, WISCONSIN 53706
 E-MAIL: bramson@math.wisc.edu

DEPARTMENT OF MATHEMATICS
 SYRACUSE UNIVERSITY
 SYRACUSE, NEW YORK 13244
 E-MAIL: jtcox@gumby.syr.edu

MATHEMATICS INSTITUT
 UNIVERSITÄT ERLANGEN–NÜRNBERG
 BISMARCKSTR. 1 1 / 2
 D-91054 ERLANGEN
 GERMANY
 E-MAIL: zech@mi.uni-erlangen.de