J. Eur. Math. Soc. 27, 4083–4158 (2025) DOI 10.4171/JEMS/1450

Nicolas Fournier · Yoan Tardy

Collisions of the supercritical Keller–Segel particle system

Received January 10, 2022; revised February 8, 2023

Abstract. We study a particle system naturally associated to the 2-dimensional Keller–Segel equation. It consists of N Brownian particles in the plane, interacting through a binary attraction in $\theta/(Nr)$, where r stands for the distance between two particles. When the intensity θ of this attraction is greater than 2, this particle system explodes in finite time. We assume that $N > 3\theta$ and study in detail what happens near explosion. There are two slightly different scenarios, depending on the values of N and θ, here is one: at explosion, a cluster consisting of precisely k_0 particles emerges, for some deterministic $k_0 \geq 7$ depending on N and θ. Just before explosion, there are infinitely many $(k_0 - 1)$-ary collisions. There are also infinitely many $(k_0 - 2)$-ary collisions before each $(k_0 - 1)$-ary collision. And there are infinitely many binary collisions before each $(k_0 - 2)$-ary collision. Finally, collisions of subsets of $3, \dots, k_0 - 3$ particles never occur. The other scenario is similar except that there are no $(k_0 - 2)$-ary collisions.

Keywords: Keller–Segel equation, stochastic particle systems, Bessel processes, collisions.

1. Introduction and main results

1.1. Informal definition of the model

We consider some scalar parameter $\theta > 0$ and a number $N \geq 2$ of particles with positions $X_t = (X_t^1, \dots, X_t^N) \in (\mathbb{R}^2)^N$ at time $t \geq 0$. Informally, we assume that the dynamics of these particles are given by the system of SDEs

$$\mathrm{d}X_t^i = \mathrm{d}B_t^i - \frac{\theta}{N} \sum_{j \neq i} \frac{X_t^i - X_t^j}{\|X_t^i - X_t^j\|^2} \mathrm{d}t, \quad i \in [\![1, N]\!], \tag{1}$$

where the 2-dimensional Brownian motions $((B_t^i)_{t \geq 0})_{i \in [\![1,N]\!]}$ are independent. In other words, we have N Brownian particles in the plane interacting through a (Coulombian) attraction in $1/r$, where r stands for the distance between two particles. Actually, this SDE does not clearly make sense, due to the singularity of the drift, and we will use,

Nicolas Fournier: LPSM, Sorbonne Université, 75005 Paris, France;
nicolas.fournier@sorbonne-universite.fr

Yoan Tardy: CMAP, UMR 7641, École Polytechnique, 91120 Palaiseau, France;
yoan.tardy@polytechnique.edu

Mathematics Subject Classification 2020: 60H10 (primary); 60K35 (secondary).

as suggested by Cattiaux–Pédèches [4], the theory of Dirichlet spaces (see Fukushima–Oshima–Takeda [11]).

1.2. Brief motivation and informal presentation of the main results

This particle system is very natural from the physical point of view, because, as we will see, there is a tight competition between the Brownian excitation and the Coulombian attraction. It can also be seen as an approximation of the famous Keller–Segel equation [16]; see also Patlak [20]. This nonlinear PDE has been introduced to model the collective motion of cells, which are attracted by a chemical substance that they emit. It is well-known that a phase transition occurs: if the intensity of the attraction is small, then there exist global solutions, while if the attraction is large, the solution explodes in finite time.

We will show that this phase transition already occurs at the level of the particle system (1): there exist global (very weak) solutions if $\theta \in (0,2)$ (*subcritical* case, see Proposition 3 below), but solutions must explode in finite time if $\theta \geq 2$ (*supercritical* case).

To our knowledge, the supercritical case has not been studied in detail, and we aim to describe precisely the explosion phenomenon. Informally, we will show the following (see Theorem 5 below). We assume that $\theta \geq 2$ and $N > 3\theta$, and we set $k_0 = \lceil 2N/\theta \rceil \in [\![7, N]\!]$. There exists a (very weak) solution $(X_t)_{t\in[0,\zeta)}$ to (1), with $\zeta < \infty$ a.s. and such that $X_{\zeta-} = \lim_{t\to\zeta-} X_t$ exists. Moreover, there is a cluster containing precisely k_0 particles in the configuration $X_{\zeta-}$, and no cluster containing strictly more particles. Such a cluster containing k_0 particles is inseparable, so that (1) is meaningless (even in a very weak sense) after ζ. Just before explosion, there are infinitely many k_1-ary collisions, where $k_1 = k_0 - 1$. If $(k_0 - 3)(2 - (k_0 - 2)\theta/N) < 2$, we set $k_2 = k_1 - 2$ and just before each k_1-ary collision, there are infinitely many k_2-collisions. Else, we set $k_2 = k_1$. In any case, there are infinitely many binary collisions just before each k_2-ary collision. During the whole time interval $[0, \zeta)$, there are no k-ary collisions, for any $k \in [\![3, k_2 - 1]\!]$.

This phenomenon seems surprising and original, in particular because of the gap between binary and k_2-ary collisions.

1.3. Sets of configurations

We introduce, for all $K \subset [\![1, N]\!]$ and all $x = (x^1, \dots, x^N) \in (\mathbb{R}^2)^N$,

$$S_K(x) = \frac{1}{|K|} \sum_{i\in K} x^i \in \mathbb{R}^2,$$

$$R_K(x) = \sum_{i\in K} \|x^i - S_K(x)\|^2 = \frac{1}{2|K|} \sum_{i,j\in K} \|x^i - x^j\|^2 \geq 0.$$

Here $|K|$ is the cardinality of K and $\|\cdot\|$ stands for the Euclidean norm in $\mathbb{R}^2$. Observe that $R_K(x) = 0$ if and only if all the particles indexed in K are at the same place. We also set, for $k \geq 2$,

$$E_k = \{x \in (\mathbb{R}^2)^N : \forall K \subset [\![1, N]\!] \text{ with } |K| = k,\ R_K(x) > 0\},$$

which represents the set of configurations with no cluster of k (or more) particles. Observe that $E_k = (\mathbb{R}^2)^N$ for all $k > N$.

1.4. Bessel processes

We recall that a *squared Bessel process* $(Z_t)_{t\geq 0}$ of dimension $\delta \in \mathbb{R}$ is a nonnegative solution, killed when it reaches 0 if $\delta \leq 0$, of the equation

$$Z_t = Z_0 + 2\int_0^t \sqrt{Z_s}\,\mathrm{d}W_s + \delta t,$$

where $(W_t)_{t\geq 0}$ is a 1-dimensional Brownian motion. We then say that $(\sqrt{Z_t})_{t\geq 0}$ is a *Bessel process* of dimension δ. This process has the following property (see Revuz–Yor [21, Chapter XI]):

- if $\delta \geq 2$, then a.s., $Z_t > 0$ for all $t > 0$;
- if $\delta \in (0, 2)$, then a.s., Z is reflected infinitely often at 0;
- if $\delta \leq 0$, then Z a.s. hits 0 and is then killed.

Applying informally the Itô formula, one finds that $Y_t = \sqrt{Z_t}$ should solve

$$Y_t = Y_0 + W_t + \frac{\delta - 1}{2}\int_0^t \frac{\mathrm{d}s}{Y_s},$$

which resembles (1) in that we have Brownian excitation in competition with attraction by 0, or repulsion by 0, depending on the value of δ, proportional to $1/r$. This formula rigorously holds true only when $\delta > 1$ (see [21, Chapter XI]).

1.5. Some important quantities

Consider a (possibly very weak) solution $(X_t)_{t\geq 0}$ to (1). As we will see, when fixing a subset $K \subset [\![1, N]\!]$ and neglecting the interactions between the particles indexed in K and the other ones, one finds that the process $(R_K(X_t))_{t\geq 0}$ behaves like a squared Bessel process of dimension $d_{\theta,N}(|K|)$, where

$$d_{\theta,N}(k) = (k-1)\left(2 - \frac{k\theta}{N}\right). \tag{2}$$

Similar computations already appear in Haškovec–Schmeiser [12]; see also [9]. A little study (see Appendix A; see also Figure 1 and Section 1.8 for numerical examples) shows the following facts. For $r \in \mathbb{R}_+$, we set $\lceil r \rceil = \min\{n \in \mathbb{N} : n \geq r\}$.

Lemma 1. *Fix $\theta > 0$ and $N \geq 2$ such that $N > \theta$. For $k_0 = \lceil 2N/\theta \rceil \geq 3$, we have*

$$d_{\theta,N}(k) > 0 \quad \text{if } k \in [\![2, k_0 - 1]\!] \quad \text{and} \quad d_{\theta,N}(k) \leq 0 \quad \text{if } k \geq k_0. \tag{3}$$

Also define $k_1 = k_0 - 1$, and

$$k_2 = \begin{cases} k_0 - 2 & \text{if } d_{\theta,N}(k_0 - 2) < 2, \\ k_0 - 1 & \text{if } d_{\theta,N}(k_0 - 2) \geq 2. \end{cases}$$

If $\theta \geq 2$ and $N > 3\theta$, then $k_0 \in [\![7, N]\!]$ and

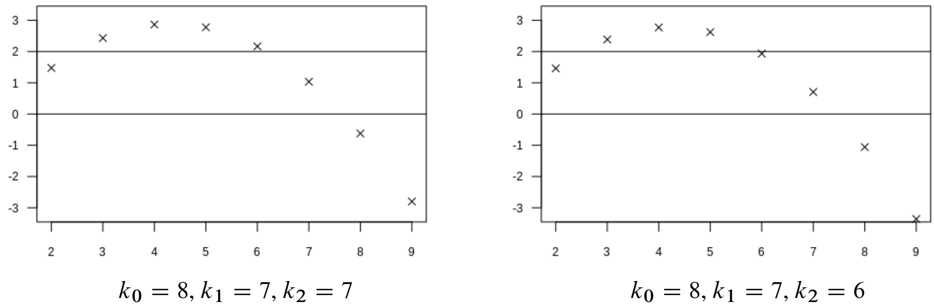

Fig. 1. Plot of $d_{\theta,N}(k)$ as a function of $k \in [\![2, N]\!]$ with $N = 9$ and with $\theta = 2.35$ (left) and $\theta = 2.42$ (right).

- $d_{\theta,N}(2) \in (0, 2)$*;*
- $d_{\theta,N}(k) \geq 2$ *if* $k \in [\![3, k_2 - 1]\!]$*;*
- $d_{\theta,N}(k) \in (0, 2)$ *if* $k \in \{k_2, k_1\}$*;*
- $d_{\theta,N}(k) \leq 0$ *if* $k \geq k_0$.

We thus expect that there may be some nonsticky k-ary collisions for $k \in \{2, k_2, k_1\}$, some sticky k-ary collisions when $k \geq k_0$, but no k-ary collision for $k \in [\![3, k_2 - 1]\!]$.

1.6. Generator and invariant measure

As we will see in Section 3.13, the SDE (1) cannot have a solution in the classical sense, at least when $d_{\theta,N}(k_1) \in (0, 1)$, because the drift term cannot be integrable in time. We will thus define a solution through the theory of Dirichlet spaces.

For $x = (x^1, \dots, x^N) \in (\mathbb{R}^2)^N$ and for dx the Lebesgue measure on $(\mathbb{R}^2)^N$, we set

$$\mathbf{m}(x) = \prod_{1 \leq i \neq j \leq N} \|x^i - x^j\|^{-\theta/N} \quad \text{and} \quad \mu(\mathrm{d}x) = \mathbf{m}(x)\mathrm{d}x, \tag{4}$$

where the product is over the set $\{(i, j) \in [\![1, N]\!]^2 : i \neq j\}$.

Informally, the generator of the solution to (1) is given by $\mathcal{L}^X$, where for $\varphi \in C^2((\mathbb{R}^2)^N)$,

$$\begin{aligned} \mathcal{L}^X \varphi(x) &= \frac{1}{2}\Delta\varphi(x) - \frac{\theta}{N} \sum_{1 \leq i \neq j \leq N} \frac{x^i - x^j}{\|x^i - x^j\|^2} \cdot \nabla_{x^i}\varphi(x) \\ &= \frac{1}{2\mathbf{m}(x)} \operatorname{div}[\mathbf{m}(x)\nabla\varphi(x)]; \end{aligned} \tag{5}$$

see (11) for the last equality. The generator is well-defined for all $x \in E_2$ and μ-symmetric. Indeed, integration by parts shows that

$$\begin{aligned} \forall \varphi, \psi \in C_c^2(E_2), \quad \int_{(\mathbb{R}^2)^N} \varphi \mathcal{L}^X \psi \,\mathrm{d}\mu &= -\frac{1}{2} \int_{(\mathbb{R}^2)^N} \nabla\varphi \cdot \nabla\psi \,\mathrm{d}\mu \\ &= \int_{(\mathbb{R}^2)^N} \psi \mathcal{L}^X \varphi \,\mathrm{d}\mu. \end{aligned} \tag{6}$$

As we will see in Proposition A.1, the measure μ is Radon on $(\mathbb{R}^2)^N$ in the subcritical case $\theta \in (0,2)$, while it is Radon on E_{k_0} (and not on E_{k_0+1}) in the supercritical case $\theta \geq 2$. This will allow us to use some results found in Fukushima–Oshima–Takeda [11] and to obtain the following existence result.

Proposition 2. *Fix $N \geq 2$ and $\theta > 0$ such that $N > \theta$ and recall that $k_0 = \lceil 2N/\theta \rceil$. Set $\mathcal{X} = E_{k_0}$ and $\mathcal{X}_\triangle = \mathcal{X} \cup \{\triangle\}$, where $\triangle$ is a cemetery point. There exists a diffusion $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}_\triangle})$ with values in $\mathcal{X}_\triangle$, which is μ-symmetric, with regular Dirichlet space $(\mathcal{E}^X, \mathcal{F}^X)$ on $L^2((\mathbb{R}^2)^N, \mu)$ with core $C_c^\infty(\mathcal{X})$ defined by*

$$\mathcal{E}^X(\varphi,\varphi) = \frac{1}{2}\int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \,\mathrm{d}\mu = -\int_{(\mathbb{R}^2)^N} \varphi \mathcal{L}^X \varphi \,\mathrm{d}\mu \quad \text{for all } \varphi \in C_c^\infty(\mathcal{X}),$$

*and such that for all $x \in E_2$ and $t > 0$, the law of X_t under $\mathbb{P}_x$ has a density with respect to the Lebesgue measure on $(\mathbb{R}^2)^N$. We call such a process a KS(θ, N)-*process *and denote by $\zeta = \inf\{t \geq 0 : X_t = \triangle\}$ its life-time.*

We refer to Section B.1 for a quick summary of the notions used in this proposition: diffusion (i.e. continuous Hunt process), link between its generator, semigroup and Dirichlet space, definition of the one-point compactification topology on $\mathcal{X}_\triangle$, etc. Let us mention that by definition, $\triangle$ is absorbing, i.e. $X_t = \triangle$ for all $t \geq \zeta$. Also, $t \mapsto X_t$ is *a priori* continuous on $[0,\infty)$ only for the one-point compactification topology on $\mathcal{X}_\triangle$, which precisely means that it is continuous for the usual topology of $(\mathbb{R}^2)^N$ during $[0,\zeta)$, and $\zeta = \lim_{n\to\infty} \inf\{t \geq 0 : X_t \notin \mathcal{K}_n\}$ for any increasing sequence $(\mathcal{K}_n)_{n\geq 1}$ of compact subsets of E_{k_0} such that $\bigcup_{n\geq 1} \mathcal{K}_n = E_{k_0}$.

As we will see in Remark 29, for all $x \in E_2$, under $\mathbb{P}_x^X$, X_t solves (1) during $[0,\sigma)$, where $\sigma = \inf\{t \geq 0 : X_t \notin E_2\}$. By the Markov property, this implies X_t solves (1) during any open time interval on which it does not visit $\mathcal{X} \setminus E_2$.

When $\theta < 2$, we have $k_0 > N$ and thus $E_{k_0} = (\mathbb{R}^2)^N$. We will easily prove the following nonexplosion result, which is almost contained in Cattiaux–Pédèches [4], who treat the case where $\theta \in (0, 2(N-2)/(N-1))$.

Proposition 3. *Fix $\theta \in (0,2)$ and $N \geq 2$. Consider the KS(θ, N)-process $\mathbb{X}$ introduced in Proposition 2. For all $x \in E_2$, we have $\mathbb{P}_x(\zeta = \infty) = 1$.*

When $\theta \geq 2$, we will see that there is explosion. Note that any collision of a set of $k \geq k_0$ particles makes the process leave E_{k_0} and thus explode. However, it is not clear at all at this point that the explosion is due to a precise collision: the process could explode because it tends to infinity (which is not hard to exclude) or to the boundary of E_{k_0} with possibly many oscillations.

1.7. Main result

To avoid any confusion, let us define precisely what we call a collision.

Definition 4. (i) For $K \subset [\![1,N]\!]$, we say that there is a *K-collision in the configuration* $x \in (\mathbb{R}^2)^N$ if $R_K(x) = 0$ and $R_{K\cup\{i\}}(x) > 0$ for all $i \in [\![1,N]\!] \setminus K$.

(ii) For an $(\mathbb{R}^2)^N$-valued process $(X_t)_{t\in[0,\zeta)}$, we say that there is a *K-collision at time* $s \in [0, \zeta)$ if there is a K-collision in the configuration X_s.

The main result of the present paper is the following description of the explosion phenomenon.

Theorem 5. *Assume that $\theta \geq 2$ and $N > 3\theta$, and recall that $k_0 \in [\![7, N]\!]$, $k_1 = k_0 - 1$ and $k_2 \in \{k_0 - 1, k_0 - 2\}$ were defined in Lemma 1. Consider the KS(θ, N)-process $\mathbb{X}$ introduced in Proposition 2. For all $x \in E_2$, the following properties hold $\mathbb{P}_x$-a.s.:*

(i) *ζ is finite and $X_{\zeta-} = \lim_{t\to\zeta-} X_t$ exists for the usual topology of $(\mathbb{R}^2)^N$;*

(ii) *there is $K_0 \subset [\![1, N]\!]$ with $|K_0| = k_0$ such that there is a K_0-collision in the configuration $X_{\zeta-}$, and for all $K \subset [\![1, N]\!]$ such that $|K| > k_0$, there is no K-collision in $X_{\zeta-}$;*

(iii) *for all $t \in [0, \zeta)$ and all $K \subset K_0$ with $|K| = k_1$, there are infinitely many K-collisions during (t, ζ) and none of these instants of K-collision is isolated;*

(iv) *if $k_2 = k_0 - 2$, then for all $L \subset K \subset K_0$ such that $|L| = k_2$ and $|K| = k_1$, for all instants $t \in (0, \zeta)$ of K-collision and all $s \in [0, t)$, there are infinitely many L-collisions during (s, t) and none of these instants of L-collision is isolated;*

(v) *for all $K \subset [\![1, N]\!]$ with $|K| \in [\![3, k_2 - 1]\!]$, there is no K-collision during $[0, \zeta)$;*

(vi) *for all $L \subset K \subset K_0$ such that $|L| = 2$ and $|K| = k_2$, all instants $t \in (0, \zeta)$ of K-collision and all $s \in [0, t)$, there are infinitely many L-collisions during (s, t) and none of these instants of L-collision is isolated.*

The condition $\theta \geq 2$ is crucial to guarantee that $k_0 \leq N$. On the other hand, we impose $N > 3\theta$ just for simplicity, because Lemma 1 does not hold true without this assumption. The other cases may also be studied, but we believe this is not very restrictive: N is thought of to be very large compared to θ, at least as far as the approximation of the Keller–Segel equation is concerned.

1.8. Comments

Let us mention that the very precise values of N and θ influence the value k_2:

(a) If $N = 200$ and $\theta = 4.04$, we have $k_0 = 100$, $k_1 = 99$ and $k_2 = 98$.

(b) If $N = 200$ and $\theta = 4.015$, we have $k_0 = 100$ and $k_1 = k_2 = 99$.

Let us describe informally, in the chronological order, what happens e.g. in case (b) above. We start with 200 particles at 200 different places. During the whole story, there is no k-ary collision for $k = 3, \dots, 98$. Here and there, two particles meet, they collide an infinite number of times, but manage to separate. Then at some time, we have 98 particles close to each other and there are many binary collisions. Then, if a 99-th particle arrives in the same zone (and this eventually occurs), there are infinitely many 99-ary collisions, with infinitely many binary collisions of all possible pairs before each. These 99 particles may manage to separate forever, or for a large time, but if a 100-th particle arrives in the zone (and this situation eventually occurs), then there are infinitely many 99-ary collisions of all the possible subsets and, finally, a 100-ary collision producing explosion, and the story is finished. Informally, the resulting cluster is not able to separate, because attraction dominates Brownian excitation, since a Bessel process of dimension $d_{\theta,N}(100) \leq 0$ is absorbed when it reaches 0. We hope to be able, in future work, to propose and justify a model describing what happens after explosion.

1.9. References

In many papers about the Keller–Segel equation, the parameter $\chi = 4\pi\theta$ is used, so that the transition at $\theta = 2$ corresponds to the transition at $\chi = 8\pi$. As already mentioned, this nonlinear PDE has been introduced to model the collective motion of cells, which are attracted by a chemical substance that they emit. It describes the density $f_t(x)$ of particles (cells) with position $x \in \mathbb{R}^2$ at time $t \geq 0$ and reads, in the so-called parabolic-elliptic case,

$$\partial_t f_t(x) + \theta \operatorname{div}_x((K \star f_t)(x) f_t(x)) = \tfrac{1}{2}\Delta_x f_t(x), \quad \text{where} \quad K(x) = -\frac{x}{|x|^2}. \tag{7}$$

Informally, this solution should be the mean-field limit of the particle system (1) as $N \to \infty$.

We refer to the recent review paper on (7) by Arumugam–Tyagi [1]. The best result on existence of a global solution to (7), including all the subcritical parameters $\theta \in (0, 2)$, is due to Blanchet–Dolbeault–Perthame [2]. The blow-up of solutions to (7), in the supercritical case $\theta > 2$, have been studied e.g. by Fatkullin [7] and Velázquez [24,25]. Closer to our study, Suzuki [23] has shown, still in the supercritical case, the appearance of a Dirac mass with a precise (critical) weight, at explosion. This is the equivalent, in the limit $N \to \infty$, to the fact that $\lim_{t\to\zeta-} X_t$ exists and corresponds to a K-collision, for some $K \subset [\![1, N]\!]$ with precise cardinality k_0. Let us finally mention Dolbeault–Schmeiser [6], who propose a post-explosion model in the supercritical case.

Concerning particle systems associated with (7), Stevens [22] studies a physically more complete particle system with two types of particles, for cells and chemo-attractant particles, with a regularized attraction kernel. Haškovec–Schmeiser [12, 13] study a particle system closer to (1), but with, again, a regularized attraction kernel.

Cattiaux–Pédèches [4], as well as [9], study the system (1) without regularization in the subcritical case: existence of a global solution to (1) has been shown in [9] when $\theta \in (0, 2(N-2)/(N-1))$, and uniqueness of this solution has been established in [4]. Also, the theory of Dirichlet spaces has been used in [4] to build a solution to (1). Finally, the limit as $N \to \infty$ to a solution of (7) is proved in [9] in the very subcritical case where $\theta \in (0, 1/2)$, up to extraction of a subsequence. This last result has been improved by Bresch–Jabin–Wang [3], who remove the necessity of extracting a subsequence and consider the (still very subcritical) case where $\theta \in (0, 1)$. Olivera–Richard–Tomašević [18] have recently established the $N \to \infty$ convergence of a smoothed version of (1), for all the subcritical cases $\theta \in (0, 2)$. Informally, in view of the mean distance between particles, the regularization used in [18] is not far from being physically reasonable. There is also a related paper of Jabir–Talay–Tomašević [14] about a 1-dimensional but more complicated parabolic-parabolic model.

Let us finally mention the seminal paper of Osada [19] (see also [8] for a more recent study), which concerns the vortex model: this is very close to (1), but the attraction $-x/|x|^2$ is replaced by a rotating interaction $x^\perp/|x|^2$, so that particles never encounter.

1.10. Originality and difficulties

To our knowledge, this is the first study of the supercritical Keller–Segel particle system near explosion. We hope that this model, which makes compete diffusion and Coulomb interactions, is very natural from the physical point of view, beyond the Keller–Segel community. The phenomenon we discovered seems surprising and original, in particular because of the gap between binary and k_2-ary collisions. We are not aware of other works, possibly dealing with other models, showing such behavior.

In Section 3, we give the main arguments of the proofs, with a fairly high level of precision, but ignoring the technical issues. While it is rather clear, intuitively, that the process explodes in finite time when $\theta \geq 2$ and that no K-collisions may occur for $|K| \in [\![3, k_2 - 1]\!]$, the continuity at explosion is delicate, and some rather deep arguments are required to show that each k_2-ary collision is preceded by many binary collisions, that each k_1-ary collision is preceded by many k_2-ary collisions, that explosion is preceded by many k_1-ary collisions, and that explosion is due to the emergence of a cluster with precise size k_0 (which more or less says that a possible $(k_0 + 1)$-ary collision would necessarily be preceded by a k_0-collision).

Actually, the rigorous proofs are made technically much more involved than those presented in Section 3, because we have to use the theory of Dirichlet spaces. Due to the singularity of the interactions and to the occurrence of many collisions near explosion, we unfortunately cannot, as already mentioned, deal at the rigorous level directly with the SDE (1). We thus have to use suitable heavy versions of some usual tools such as Itô's formula, Girsanov's theorem, time-change, etc.

1.11. Plan of the paper

In Section 2, we introduce some notation of constant use. In Section 3, we explain the main ideas of the proofs, with a high level of precision, but without speaking of the heavy technical issues related to the use of the theory of Dirichlet spaces. Section 4 is devoted to the existence of a first version of the Keller–Segel process, namely without the property that $\mathbb{P}_x^X \circ X_t^{-1}$ has a density, and we introduce a spherical Keller–Segel process. In Section 5, we show that the Keller–Segel process enjoys a crucial and remarkable decomposition in terms of a 2-dimensional Brownian motion, a squared Bessel process and a spherical process. Section 6 consists in building some smooth approximations of some indicator functions that behave well under the action of the generator $\mathcal{L}^X$. In Section 7, we make use of the Girsanov theorem to prove that when two sets of particles of a *KS*-process are not too close to each other, they behave as two independent smaller *KS*-processes. In

Section 8, we study explosion and continuity (in the usual sense) at the explosion time. Section 9 is devoted to establishing some parts of Theorem 5 for some particular ranges of values of N and θ. Using the results of Section 7, we reduce the general study to the special cases of Section 9 and we prove, in Section 10, that the conclusions of Theorem 5 hold true quasi-everywhere. Finally, in Section 11, we remove the "quasi-everywhere" restriction and conclude the proofs of Propositions 2 and 3 and of Theorem 5.

Appendix A contains a few elementary computations: proof of Lemma 1, proof that μ is Radon on E_{k_0}, and study of a similar measure on a sphere. We end the paper with Appendix B that summarizes all the notions and results about Dirichlet spaces and Hunt processes we shall use.

2. Notation

We introduce the spaces

$$H = \{x \in (\mathbb{R}^2)^N : S_{[\![1,N]\!]}(x) = 0\}, \quad S = \Big\{x \in (\mathbb{R}^2)^N : \sum_{i=1}^N \|x^i\|^2 = 1\Big\},$$
$$\mathbb{S} = H \cap S.$$

For $u \in \mathbb{S}$, we have $S_{[\![1,N]\!]}(u) = 0$ and $R_{[\![1,N]\!]}(u) = 1$. We consider the (unnormalized) Lebesgue measure σ on $\mathbb{S}$, as well as (recall (4)),

$$\beta(\mathrm{d}u) = \mathbf{m}(u)\sigma(\mathrm{d}u). \tag{8}$$

We define $\gamma : \mathbb{R}^2 \to (\mathbb{R}^2)^N$ by $\gamma(z) = (z, \dots, z)$ and $\Psi : \mathbb{R}^2 \times \mathbb{R}_+^* \times \mathbb{S} \to E_N \subset (\mathbb{R}^2)^N$ by

$$\Psi(z, r, u) = \gamma(z) + \sqrt{r}u, \quad \text{i.e.} \quad (\Psi(z, r, u))^i = z - \sqrt{r}u^i \quad \text{for } i \in [\![1, N]\!]. \tag{9}$$

We have $S_{[\![1,N]\!]}(\Psi(z, r, u)) = z$ and $R_{[\![1,N]\!]}(\Psi(z, r, u)) = r$.

The orthogonal projection $\pi_H : (\mathbb{R}^2)^N \to H$ is given by

$$\pi_H(x) = x - \gamma(S_{[\![1,N]\!]}(x)), \quad \text{i.e.} \quad (\pi_H(x))^i = x^i - S_{[\![1,N]\!]}(x) \quad \text{for } i \in [\![1, N]\!],$$

and we introduce $\Phi_{\mathbb{S}} : E_N \to \mathbb{S}$ defined by

$$\Phi_{\mathbb{S}}(x) = \frac{\pi_H x}{\|\pi_H x\|}, \quad \text{i.e.} \quad (\Phi_{\mathbb{S}}(x))^i = \frac{x^i - S_{[\![1,N]\!]}(x)}{\sqrt{R_{[\![1,N]\!]}(x)}} \quad \text{for } i \in [\![1, N]\!]. \tag{10}$$

For $x \in (\mathbb{R}^2)^N \setminus \{0\}$, the projections $\pi_{x^\perp} : (\mathbb{R}^2)^N \to x^\perp$ and $\pi_x : (\mathbb{R}^2)^N \to \mathrm{span}(x)$ are given by

$$\pi_{x^\perp}(y) = y - \frac{x \cdot y}{\|x\|^2}x \quad \text{and} \quad \pi_x(y) = \frac{x \cdot y}{\|x\|^2}x,$$

where $x \cdot y = \sum_{i=1}^N x^i \cdot y^i$.

We denote by $b : E_2 \to (\mathbb{R}^2)^N$ the drift coefficient of (1): for $x = (x^1, \dots, x^N) \in E_2$,

$$b(x) = \frac{\nabla \mathbf{m}(x)}{2\mathbf{m}(x)} = \frac{\nabla \log \mathbf{m}(x)}{2} \in (\mathbb{R}^2)^N, \quad \text{i.e.} \quad b^i(x) = -\frac{\theta}{N} \sum_{j \neq i} \frac{x^i - x^j}{\|x^i - x^j\|^2} \in \mathbb{R}^2 \tag{11}$$

for $i \in [\![1, N]\!]$. Indeed, since $\log \mathbf{m}(x) = -\frac{\theta}{2N} \sum_{1 \le i \neq j \le N} \log \|x^i - x^j\|^2$, we have e.g.

$$\frac{\nabla_{x^1} \log \mathbf{m}(x)}{2} = -\frac{\theta}{4N} \nabla_{x^1} \Big[\sum_{i=2}^N \log \|x^i - x^1\|^2 + \sum_{j=2}^N \log \|x^1 - x^j\|^2 \Big]$$

$$= -\frac{\theta}{2N} \nabla_{x^1} \sum_{j=2}^N \log \|x^1 - x^j\|^2,$$

whence

$$\frac{\nabla_{x^1} \log \mathbf{m}(x)}{2} = -\frac{\theta}{N} \sum_{j=2}^N \frac{x^1 - x^j}{\|x^1 - x^j\|^2}.$$

Finally, we introduce the natural operators defined for $\varphi \in C^1(\mathbb{S})$ and $u \in \mathbb{S}$ by

$$\nabla_{\mathbb{S}} \varphi(u) = \nabla[\varphi \circ \Phi_{\mathbb{S}}](u) \in (\mathbb{R}^2)^N \quad \text{and} \quad \Delta_{\mathbb{S}} \varphi(u) = \Delta[\varphi \circ \Phi_{\mathbb{S}}](u) \in \mathbb{R}, \tag{12}$$

where ∇ and Δ stand for the usual gradient and Laplacian in $(\mathbb{R}^2)^N$. Since $\mathbb{S} \subset E_N \subset (\mathbb{R}^2)^N$, with E_N open, and since $\Phi_{\mathbb{S}}$ is smooth on E_N, we can indeed define $\nabla[\varphi \circ \Phi_{\mathbb{S}}](u)$ and $\Delta[\varphi \circ \Phi_{\mathbb{S}}](u)$ for all $u \in \mathbb{S}$. Similarly, for $\varphi \in C^1(\mathbb{S}, (\mathbb{R}^2)^N)$ and $u \in \mathbb{S}$, we set

$$\operatorname{div}_{\mathbb{S}} \varphi(u) = \operatorname{div}[\varphi \circ \Phi_{\mathbb{S}}](u) \in \mathbb{R}. \tag{13}$$

To conclude this subsection, we note that for all $\varphi \in C^\infty((\mathbb{R}^2)^N)$ and all $u \in \mathbb{S}$,

$$\nabla_{\mathbb{S}}(\varphi|_{\mathbb{S}})(u) = \pi_H(\pi_{u^\perp}(\nabla \varphi(u))). \tag{14}$$

Indeed, it suffices to observe that setting $G(x) = x/\|x\|$ for all $x \in (\mathbb{R}^2)^N \setminus \{0\}$, we have $\Phi_{\mathbb{S}} = G \circ \pi_H$, $\mathrm{d}_x G = \pi_{x^\perp}/\|x\|$ and $\mathrm{d}_x \pi_H = \pi_H$, and also for $u \in \mathbb{S}$, we have $\pi_H(u) = u$ and $\|\pi_H(u)\| = 1$.

3. Main ideas of the proofs

Here we explain the main ideas of the proofs of Proposition 3 and Theorem 5. The arguments below are completely informal. In particular, we act as if our $KS(\theta, N)$-process $(X_t)_{t \in [0,\zeta)}$ was a true solution to (1) until explosion and we apply Itô's formula without care. We always assume at least that $N \ge 2$, $\theta > 0$ and $N > \theta$, which implies that $k_0 = \lceil 2N/\theta \rceil \ge 3$.

3.1. Existence

The existence of the $KS(\theta, N)$-process $(X_t)_{t \in [0,\zeta)}$, with values in E_{k_0}, is an easy application of [11, Theorem 7.2.1]. The only difficulty is to show that the invariant measure μ is

Radon on E_{k_0}; see Proposition A.1. The process may explode, i.e. get out of any compact subset of E_{k_0} in finite time. Observe that a typical compact subset of E_{k_0} is of the form, for $\varepsilon > 0$,

$$\mathcal{K}_\varepsilon = \{x \in (\mathbb{R}^2)^N : \|x\| \leq 1/\varepsilon \text{ and for all } K \subset [\![1, N]\!] \text{ such that } |K| = k_0,\ R_K(x) \geq \varepsilon\}.$$

3.2. *Center of mass and dispersion process*

One can verify, using Itô's formula, that the center of mass $S_{[\![1,N]\!]}(X)$ is a 2-dimensional Brownian motion with diffusion constant $N^{-1/2}$, the dispersion process $R_{[\![1,N]\!]}(X)$ is a squared Bessel process of dimension $d_{\theta,N}(N)$ (recall (2)), and these two processes are independent.

Consequently, if $\zeta < \infty$, the limits $\lim_{t\to\zeta-} S_{[\![1,N]\!]}(X_t)$ and $\lim_{t\to\zeta-} R_{[\![1,N]\!]}(X_t)$ a.s. exist, and this implies that $\limsup_{t\to\zeta-} \|X_t\| < \infty$: the process cannot explode to infinity, it can only explode because it tends to the boundary of E_{k_0}. If moreover $k_0 > N$ (i.e. $\theta < 2$), this is sufficient to show that $\zeta = \infty$, since then $E_{k_0} = (\mathbb{R}^2)^N$.

3.3. *Behavior of distant subsets of particles*

Consider a partition $K_1, \dots, K_p$ of $[\![1, N]\!]$. If we neglect interactions between particles whose indices are not in the same subset, we have, for each $\ell \in [\![1, p]\!]$, setting $\tilde{\theta}_\ell = \theta|K_\ell|/N$,

$$\mathrm{d}X_t^i = \mathrm{d}B_t^i - \frac{\tilde{\theta}_\ell}{|K_\ell|} \sum_{j\in K_\ell\setminus\{i\}} \frac{X_t^i - X_t^j}{\|X_t^i - X_t^j\|^2}\mathrm{d}t, \quad i \in K_\ell,$$

and we recognize a $KS(\tilde{\theta}_\ell, |K_\ell|)$-process.

During time intervals where particles indexed in different subsets are far enough from each other, we can indeed bound the interaction between those particles, so that the Girsanov theorem tells us that $(X_t^i)_{i\in K_1}, \dots, (X_t^i)_{i\in K_p}$ behave similarly, in the sense of trajectories, as independent $KS(\tilde{\theta}_1, |K_1|), \dots, KS(\tilde{\theta}_p, |K_p|)$-processes.

3.4. *Brownian and Bessel behaviors of isolated subsets of particles*

Consider $K \subset [\![1, N]\!]$. As seen just above, during time intervals where the particles indexed in K are far from all the other ones, the system $(X_t^i)_{i\in K}$ behaves, in the sense of trajectories, like a $KS(\theta|K|/N, |K|)$-process. Hence (see Section 3.2), $S_K(X_t)$ behaves like a 2-dimensional Brownian motion with diffusion constant $|K|^{-1/2}$, while $R_K(X_t)$ behaves like a squared Bessel process of dimension $d_{\theta|K|/N,|K|}(|K|)$, which is nothing but $d_{\theta,N}(|K|)$ (recall (2)).

3.5. *Continuity at explosion*

Here we assume that $N > \theta \geq 2$, so that $k_0 \in [\![2, N]\!]$, and we explain why a.s., $\zeta < \infty$ and $X_{\zeta-} = \lim_{t\to\zeta-} X_t$ exists, in the usual sense of $(\mathbb{R}^2)^N$.

(a) We first show that $\zeta < \infty$ a.s. On the event where $\zeta = \infty$, the squared Bessel process $R_{[\![1,N]\!]}(X)$ is defined for all times. Recall that $d_{\theta,N}(N) \le 0$ (because $\theta \ge 2$) and a squared Bessel process of negative dimension can be defined on the whole time half-line and a.s. becomes negative in finite time. Since $R_{[\![1,N]\!]}(X) \ge 0$ by definition, this contradicts the fact that $\zeta = \infty$.

Similarly, one can show that a $KS(\theta, N)$-process has no chance to be defined after the first hitting time τ_K of 0 by $R_K(X_t)$, where $|K| = k_0$: this makes the choice of the space E_{k_0} very natural. Indeed, assume that X is defined during $[0, \zeta')$ with $\zeta' > \tau_K$. Consider the maximal subset L of $[\![1, N]\!]$ containing K and such that $R_L(X_{\tau_K}) = 0$. Then there is $\varepsilon > 0$ such that during $[\tau_K, \tau_K + \varepsilon] \subset [0, \zeta')$, the particles labeled in L are far from the ones labeled outside L. By Section 3.4, $(R_L(X_{\tau_K+t}))_{t\in[0,\varepsilon]}$ behaves like a squared Bessel process of dimension $d_{\theta,N}(|L|)$, issued from 0. But such a process is instantaneously negative, because $d_{\theta,N}(|L|) \le 0$ (since $|L| \ge k_0$). Since $R_L(X) \ge 0$, this contradicts the fact that $\tau_K \in [0, \zeta')$.

(b) We next show by reverse induction that a.s. for all $K \subset [\![1, N]\!]$ with $|K| \ge 2$,

$$\text{either} \quad \lim_{t\to\zeta-} R_K(X_t) = 0 \quad \text{or} \quad \liminf_{t\to\zeta-} R_K(X_t) > 0. \tag{15}$$

If $K = [\![1, N]\!]$, $\lim_{t\to\zeta-} R_K(X_t)$ exists by continuity of the (true) squared Bessel process $R_K(X_t)$ and this implies the result. We now fix $n \in [\![3, N]\!]$ and assume that (15) holds true for all K such that $|K| \ge n$. We consider $K \subset [\![1, N]\!]$ with $|K| = n - 1$; by induction assumption, either there is $i \notin K$ such that $\lim_{t\to\zeta-} R_{K\cup\{i\}}(X_t) = 0$ and then $\lim_{t\to\zeta-} R_K(X_t) = 0$, or for all $i \in [\![1, N]\!] \setminus K$, $\liminf_{t\to\zeta-} R_{K\cup\{i\}}(X_t) > 0$. In this last case, and when

$$\limsup_{t\to\zeta-} R_K(X_t) > 0 \quad \text{and} \quad \liminf_{t\to\zeta-} R_K(X_t) = 0$$

(which is the negation of (15)), there are $\alpha, \varepsilon > 0$ such that (i) $R_K(X_t)$ upcrosses $[\varepsilon/2, \varepsilon]$ infinitely often during $[\zeta - \alpha, \zeta)$ and (ii) for all $t \in [\zeta - \alpha, \zeta)$ such that $R_K(X_t) < \varepsilon$, the particles indexed in K are far from all the other ones (because $R_K(X_t)$ is then small and $R_{K\cup\{i\}}(X_t)$ is large for all $i \notin K$), so that $R_K(X_t)$ behaves like a squared Bessel process of dimension $d_{\theta,N}(|K|)$; see Section 3.4. Points (i) and (ii) are in contradiction, since a squared Bessel process is continuous and thus cannot upcross $[\varepsilon/2, \varepsilon]$ infinitely often during a finite time interval.

(c) We now show that $\lim_{t\to\zeta-} X_t$ exists. Using (b) and the (random) equivalence relation on $[\![1, N]\!]$ defined by $i \sim j$ if and only if $\lim_{t\to\zeta-} R_{\{i,j\}}(X_t) = 0$, one can build a partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1, N]\!]$ such that for all $p \in [\![1, \ell]\!]$, $\lim_{t\to\zeta-} R_{K_p}(X_t) = 0$ and $\liminf_{t\to\zeta-} \min_{i\notin K_p} R_{K_p\cup\{i\}}(X_t) > 0$. Hence, there is $\alpha \in [0, \zeta)$ such that for all $p \ne q$, the particles labeled in K_p are far from the ones labeled in K_q during $[\alpha, \zeta)$. As seen in Section 3.4, we conclude that for all $p \in [\![1, \ell]\!]$, $S_{K_p}(X_t)$ behaves like a Brownian motion during $[\alpha, \zeta)$, and thus $M_p = \lim_{t\to\zeta-} S_{K_p}(X_t)$ exists. Since moreover $\lim_{t\to\zeta-} R_{K_p}(X_t) = 0$, we deduce that for all $i \in K_p$, $\lim_{t\to\zeta-} X_t^i = M_p$. In conclusion, $\lim_{t\to\zeta-} X_t^i$ exists for all $i \in [\![1, N]\!]$.

3.6. A spherical process

We recall that $\mathbb{S}$, π_H, $\pi_{u^\perp}$ and b were defined in Section 2 and introduce the possibly exploding (with life-time ξ) process $(U_t)_{t\in[0,\xi)}$ with values in $\mathbb{S} \cap E_{k_0}$, informally solving (we will also use here the theory of Dirichlet spaces), for some given $U_0 \in \mathbb{S} \cap E_{k_0}$ and some $(\mathbb{R}^2)^N$-valued Brownian motion $(B_t)_{t\ge 0}$, the equation

$$U_t = U_0 + \int_0^t \pi_{U_s^\perp}\pi_H \,\mathrm{d}B_s + \int_0^t \pi_{U_s^\perp}\pi_H b(U_s)\,\mathrm{d}s - \frac{2N-3}{2}\int_0^t U_s\,\mathrm{d}s.$$

We call it an *SKS*(θ, N)*-process*.

One can check that this process is β-symmetric, where β is defined in (8), and that β is Radon on $\mathbb{S} \cap E_{k_0}$; see Proposition A.3. And we will see that if $k_0 \ge N$, then $\beta(\mathbb{S}) < \infty$, so that the process $(U_t)_{t\ge 0}$ is nonexploding and positive recurrent.

3.7. Decomposition of the process

We assume that $N \ge 2$ and $\theta > 0$ are such $d_{\theta,N}(N) < 2$ and, as usual, $N > \theta$. We consider a 2-dimensional Brownian $(M_t)_{t\ge 0}$ with diffusion constant $N^{-1/2}$, a squared Bessel process $(D_t)_{t\in[0,\tau_D)}$ of dimension $d_{\theta,N}(N)$ killed when it hits 0, with life-time τ_D, and an *SKS*(θ, N)-process $(U_t)_{t\in[0,\xi)}$, these three processes being independent. We introduce the time-change

$$A_t = \int_0^t \frac{\mathrm{d}s}{D_s}, \quad t \in [0, \tau_D).$$

Since $\tau_D < \infty$ (because $d_{\theta,N}(N) < 2$) and $D_{\tau_D} = 0$ and since, roughly, the paths of $(\sqrt{D_t})_{t\in[0,\tau_D)}$ are 1/2-Hölder continuous, we have $A_{\tau_D} = \infty$ a.s. We introduce the inverse function $\rho : [0,\infty) \to [0, \tau_D)$ of $A : [0, \tau_D) \to [0,\infty)$.

We also set $\zeta' = \rho_\xi$ and observe that $\zeta' \le \tau_D$, since ρ is $[0, \tau_D)$-valued, and that $\zeta' < \tau_D$ if and only if $\xi < \infty$. A fastidious but straightforward computation shows that, recalling (9), the process

$$X_t = \Psi(M_t, D_t, U_{A_t}), \quad \text{i.e.} \quad X_t^i = M_t + \sqrt{D_t}U^i_{A_t},\ i \in [\![1, N]\!],$$

which is well-defined during $[0, \zeta')$, solves (1).

This decomposition of the KS(θ, N)*-process, which is remarkable in that U satisfies an autonomous SDE and thus is Markov, is at the basis of our analysis.*

In other words, $(X_t)_{t\in[0,\zeta')}$ is the restriction to the time interval $[0, \zeta')$ of a *KS*(θ, N)-process $(X_t)_{t\in[0,\zeta)}$. Moreover, we have $\zeta' = \zeta \wedge \tau_D$: if ξ is finite, then U gets out of $\mathbb{S} \cap E_{k_0}$ at time ξ, so that X gets out of E_{k_0} at time $\zeta' = \rho_\xi < \tau_D$, whence $\zeta = \zeta' = \zeta \wedge \tau_D$; and if $\xi = \infty$, then $\zeta' = \tau_D$ and U remains in E_{k_0} for all times, so that X remains in E_{k_0} during $[0, \tau_D)$, whence $\zeta \ge \tau_D$.

We have $S_{[\![1,N]\!]}(X_t) = M_t$ and $R_{[\![1,N]\!]}(X_t) = D_t$ for all $t \in [0, \zeta \wedge \tau_D)$, because U is $\mathbb{S}$-valued. By definition of $\mathbb{S}$, the process U cannot have any $[\![1, N]\!]$-collision. But for any $K \subset [\![1, N]\!]$ of cardinality at most $N-1$,

$$U \text{ has a } K\text{-collision at } t \in [0, \xi) \text{ if and only if } X \text{ has a } K\text{-collision at } \rho_t \in [0, \zeta \wedge \tau_D). \tag{16}$$

Moreover, as seen a few lines above, $\xi < \infty$ is equivalent to $\zeta < \tau_D$. In other words, since $R_{[\![1,N]\!]}(X_t) = D_t$ for all $t \in [0, \zeta \wedge \tau_D)$ and $\tau_D = \inf\{t > 0 : D_t = 0\}$, we have

$$\xi < \infty \quad \text{if and only if} \quad \inf_{t \in [0,\zeta)} R_{[\![1,N]\!]}(X_t) > 0. \tag{17}$$

3.8. Some special cases

Using the Girsanov theorem (see Section 3.4), we will manage to reduce a large part of the study to the special cases that we examine in the present subsection. Here we explain the following facts, for $N \geq 2$ and $\theta > 0$ with $N > \theta$:

(a) if $d_{\theta,N}(N-1) \in (0, 2)$, then a.s., $\tau_D = \inf\{t > 0 : R_{[\![1,N]\!]}(X_t) = 0\} \leq \zeta$ and for all $r \in [0, \tau_D)$ and all $K \subset [\![1, N]\!]$ with $|K| = N-1$, $(X_t)_{t \in [0,\zeta)}$ has infinitely many K-collisions during $[r, \tau_D)$;

(b) if $d_{\theta,N}(N-1) \leq 0$ (whence $k_0 \leq N-1$), then a.s., $\inf_{t \in [0,\zeta)} R_{[\![1,N]\!]}(X_t) > 0$.

We keep the same notation as in the previous subsection.

(i) We first verify that in (a), $\tau_D \leq \zeta$. Since $d_{\theta,N}(N-1) \in (0, 2)$, we have $k_0 \geq N$. If $k_0 > N$, then $\zeta = \infty$ by Section 3.2 and we are done. If $k_0 = N$, then $\zeta < \infty$ and $X_{\zeta-}$ exists by Section 3.5. Moreover, $X_{\zeta-}$ cannot belong to $E_{k_0} = E_N$ by definition of ζ and thus has its N particles at the same place, i.e. $R_{[\![1,N]\!]}(X_{\zeta-}) = 0$: we have $\zeta = \tau_D$.

(ii) In (b), $\zeta < \infty$ by Section 3.5 because $d_{\theta,N}(N-1) \leq 0$ implies that $\theta \geq 2$.

(iii) We consider, in any case, the spherical process $(U_t)_{t \in [0,\xi)}$ and assume that $\xi = \infty$. An Itô computation shows that for $K \subset [\![1, N]\!]$, for some 1-dimensional Brownian motion $(W_t)_{t \geq 0}$,

$$\begin{aligned} \mathrm{d}R_K(U_t) &= 2\sqrt{R_K(U_t)(1 - R_K(U_t))}\mathrm{d}W_t + d_{\theta,N}(|K|)\mathrm{d}t - d_{\theta,N}(N)R_K(U_t)\mathrm{d}t \\ &\quad - \frac{2\theta}{N} \sum_{i \in K, j \notin K} \frac{U_t^i - U_t^j}{\|U_t^i - U_t^j\|^2} \cdot (U_t^i - S_K(U_t))\mathrm{d}t. \end{aligned}$$

We fix $\varepsilon > 0$ to be chosen later. During time intervals where $\min_{i \in K, j \notin K} \|U_t^i - U_t^j\| \geq \varepsilon$, we thus have, for some constant C_ε,

$$\mathrm{d}R_K(U_t) \leq 2\sqrt{R_K(U_t)(1 - R_K(U_t))}\mathrm{d}W_t + d_{\theta,N}(|K|)\mathrm{d}t + C_\varepsilon\sqrt{R_K(U_t)}\mathrm{d}t, \tag{18}$$

where we have used the Cauchy–Schwarz inequality and the fact that $R_K(U_t)$ is uniformly bounded (because U is $\mathbb{S}$-valued). Hence, still during time intervals where $\min_{i \in K, j \notin K} \|U_t^i - U_t^j\| \geq \varepsilon$, by comparison, $R_K(U_t)$ is smaller than S_t, the solution to

$$\mathrm{d}S_t = 2\sqrt{S_t(1 - S_t)}\mathrm{d}W_t + d_{\theta,N}(|K|)\mathrm{d}t + C_\varepsilon\sqrt{S_t}\mathrm{d}t. \tag{19}$$

And an examination involving scale functions/speed measures shows that this process hits zero in finite time if and only if $d_{\theta,N}(|K|) < 2$, exactly as a squared Bessel process of dimension $d_{\theta,N}(|K|)$.

(iv) We end the proof of (a). In this case $k_0 \geq N$, so that U is nonexploding, as seen in Section 3.6. Hence $\xi = \infty$ and we can use (iii). Moreover, U is recurrent, still by Section 3.6. We fix K with $|K| = N - 1$ and we choose $\varepsilon > 0$ small enough that

$$\beta\Big(\Big\{u \in \mathbb{S} : \min_{i \in K,\, j \notin K} \|u^i - u^j\| \geq \varepsilon\Big\}\Big) > 0,$$

where β is the invariant measure (8) of U. Hence the process $\min_{i \in K, j \notin K} \|U_t^i - U_t^j\|$ visits the zone (ε, ∞) infinitely often, and each time $R_K(U)$ has a (uniformly) positive probability of hitting 0 by (iii) and since $d_{\theta,N}(|K|) = d_{\theta,N}(N-1) < 2$. Consequently, for any $s > 0$, $(U_t)_{t \geq 0}$ has infinitely many K-collisions during $[s, \infty)$. Recalling (16) and that $\zeta \wedge \tau_D = \tau_D$ by (i), we conclude that for any $r \in [0, \tau_D)$, $(X_t)_{t \in [0,\zeta)}$ has infinitely many K-collisions during $[r, \tau_D)$.

(v) We finally complete the proof of (b). By (17), it is sufficient to show that $\xi < \infty$ a.s. Assume that U is recurrent (and thus nonexploding). Then we take $K = [\![2, N]\!]$ and apply the same reasoning as in (iv): since $d_{\theta,N}(|K|) \leq 0 < 2$, $R_K(U)$ hits zero in finite time and this makes U get out of E_{N-1} and thus explode, since U is $(E_{k_0} \cap \mathbb{S})$-valued and $k_0 \leq N - 1$. We thus have a contradiction.

Hence U is transient and it eventually gets out of the compact subset of $E_{k_0} \cap \mathbb{S}$ given by

$$\mathcal{K} = \{u \in \mathbb{S} : \forall K \subset [\![1, N]\!] \text{ with } |K| = k_0, \text{ we have } R_K(u) \geq \varepsilon\},$$

for any fixed $\varepsilon > 0$. Hence on the event that $\xi = \infty$, $\lim_{t\to\infty} \min_{|K|=k_0} R_K(U_t) = 0$ a.s. Recalling now that $k_0 \leq N - 1$ and U is $\mathbb{S}$-valued (whence $R_{[\![1,N]\!]}(U_t) = 1$), we can a.s. find K with $|K| \in [\![k_0, N-1]\!]$ satisfying $\liminf_{t\to\infty} R_K(U_t) = 0$ but such that $\liminf_{t\to\infty} \min_{i \notin K} R_{K\cup\{i\}}(U_t) > 0$. It is then not too hard to find $\alpha, \varepsilon > 0$ such that each time $R_K(U_t) < \alpha$ (which happens often), all the particles indexed in K are farther than $\varepsilon > 0$ from all the other ones. We conclude from (iii), since $d_{\theta,N}(|K|) \leq 0$ (because $|K| \geq k_0$), that each time $R_K(U_t) < \alpha$, U has a (uniformly) positive probability to hit zero. On the event $\xi = \infty$, this will eventually happen, so that the process U will have a K-collision and thus will leave E_{k_0} in finite time. Hence U will explode, so that $\xi < \infty$.

3.9. *Size of the cluster*

We assume that $N > 3\theta \geq 6$. Hence $\zeta < \infty$ and $X_{\zeta-}$ exists, by Section 3.5. Moreover, by definition of ζ, we know that $X_{\zeta-} \notin E_{k_0}$. We want now to show that $X_{\zeta-} \in E_{k_0+1}$, i.e. the cluster causing explosion is composed of precisely k_0 particles. If $k_0 = N$, there is nothing to do, since then $E_{k_0+1} = (\mathbb{R}^2)^N$. Now if $k_0 \leq N - 1$, we assume for contradiction that there is $K \subset [\![1, N]\!]$ with $|K| \geq k_0 + 1$ such that $R_K(X_{\zeta-}) = 0$ and

$\min_{i \notin K} R_{K \cup \{i\}}(X_{\zeta-}) > 0$. Then there is $\alpha > 0$ such that during $[\zeta - \alpha, \zeta)$, the particles indexed in K are far from the other ones, so that the process $(X_t^i)_{t \in [0,\zeta), i \in K}$ behaves like a $KS(\theta|K|/N, |K|)$-process by Section 3.3. Observe now that $d_{\theta|K|/N,|K|}(|K| - 1) = d_{\theta,N}(|K| - 1) \le 0$ because $|K| - 1 \ge k_0$, and $|K| > \theta|K|/N$ because $N > \theta$. We thus know from the special case (b) of Section 3.8 that $\inf_{t \in [\zeta-\alpha,\zeta)} R_K(X_t) > 0$, which contradicts $R_K(X_{\zeta-}) = 0$.

3.10. Collisions before explosion

We fix again $N > 3\theta \ge 6$. We recall that $k_1 = k_0 - 1$ and we show that there are infinitely many k_1-ary collisions just before explosion. By the previous subsection, there exists $K_0 \subset [\![1, N]\!]$ such that $|K_0| = k_0$ and $R_{K_0}(X_{\zeta-}) = 0$ and $\min_{i \notin K_0} R_{K_0 \cup \{i\}}(X_{\zeta-}) > 0$. Then there is $\alpha > 0$ such that during $[\zeta - \alpha, \zeta)$, the particles indexed in K_0 are far from the other ones, so that $(X_t^i)_{i \in K_0}$ behaves like a $KS(\theta k_0/N, k_0)$-process by Section 3.3. Observe now that $d_{\theta k_0/N,k_0}(k_0 - 1) = d_{\theta,N}(k_0 - 1) \in (0, 2)$ thanks to Lemma 1 and that $k_0 > \theta k_0/N$ because $N > \theta$. We thus know from the special case (a) of Section 3.8 that $(X_t^i)_{i \in K_0}$ has infinitely many $(K_0 \setminus \{i\})$-collisions just before ζ, for all $i \in K_0$.

When $k_2 = k_1 - 1$, one can show in the same way that for all K with $|K| = k_1$, and all $i \in K$, there are infinitely many $(K \setminus \{i\})$-collisions just before each K-collision. We may also use Section 3.8 (a), since $d_{\theta k_1/N,k_1}(k_1 - 1) = d_{\theta,N}(k_2) \in (0, 2)$; see Lemma 1.

3.11. Absence of other collisions

We want to show that when $N > 3\theta \ge 6$, for $K \subset [\![1, N]\!]$ with $|K| \in [\![3, k_2 - 1]\!]$ there is no K-collision during $(0, \zeta)$. Suppose for contradiction that there is $K \subset [\![1, N]\!]$ with $|K| \in [\![3, k_2 - 1]\!]$ and $t \in (0, \zeta)$ such that $R_K(X_t) = 0$ and $R_{K \cup \{i\}}(X_t) > 0$ for all $i \notin K$. Then there is $\alpha > 0$ such that during $[t - \alpha, t]$, the particles indexed in K are far from the other ones, so that $R_K(X_t)$ behaves like a squared Bessel process of dimension $d_{\theta|K|/N,|K|}(|K|)$ (see Section 3.4). Since $d_{\theta|K|/N,|K|}(|K|) = d_{\theta,N}(|K|) \ge 2$ because $|K| \in [\![3, k_2 - 1]\!]$ (see Lemma 1), such a Bessel process cannot hit zero, which is a contradiction.

3.12. Binary collisions

We still assume that $N > 3\theta \ge 6$, we suppose that there is a K-collision for some $K \subset [\![1, N]\!]$ such that $|K| = k_2$ at some time $t \in (0, \zeta)$, and we want to show that there are infinitely many binary collisions just before t. There is $\alpha > 0$ such that the particles indexed in K are far from all the other ones during $[t - \alpha, t]$, so that Section 3.3 tells us that $(X_t^i)_{i \in K}$ behaves like a $KS(\theta k_2/N, k_2)$-process. We observe that $k_2 \ge 5$, $d_{\theta k_2/N,k_2}(k_2 - 1) = d_{\theta,N}(k_2 - 1) \ge 2$ and $d_{\theta k_2/N,k_2}(k_2) = d_{\theta,N}(k_2) \in (0, 2)$ by Lemma 1.

We are reduced to showing that a $KS(\theta, N)$-process, still denoted by $(X_t^i)_{i \in [\![1,N]\!], t \ge 0}$, such that $N \ge 5$, $d_{\theta,N}(N - 1) \ge 2$ and $d_{\theta,N}(N) \in (0, 2)$, a.s. has infinitely many binary collisions before the first instant τ_D of $[\![1, N]\!]$-collision. Such a process does not explode, because $k_0 > N$ (since $d_{\theta,N}(N) > 0$; see Section 3.2). Hence using (16) (which is licit since $d_{\theta,N}(N) < 2$), we only have to show that e.g. U^1 collides infinitely often with U^2 during $[0, \infty)$.

First, one easily sees that the probability that e.g. X^1 collides with X^2 before τ_D is positive, because the probability that all the particles are pairwise far from each other, except X^1 and X^2, during the time interval [0, 1], is positive. On this kind of event, by Section 3.4, $R_{\{1,2\}}(X_t)$ behaves like a squared Bessel process of dimension $d_{\theta,N}(2) \in (0, 2)$ and thus hits zero during [0, 1] (and thus before τ_D) with positive probability.

Using again (16), we conclude that the probability that U^1 collides with U^2 in finite time is positive. Since now U is positive recurrent (recall Section 3.6 and that $k_0 > N$, because $d_{\theta,N}(N) > 0$), we conclude that U^1 collides infinitely often with U^2 during $[0, \infty)$ as desired.

3.13. Nonintegrability of the drift term

Here we check that when $d_{\theta,N}(k_1) \in (0, 1)$, the SDE (1) cannot have a solution in the classical sense, because the drift term is not integrable in time. More precisely, recall that there is some K-collision at some time τ strictly before explosion, for some $K \subset [\![1, N]\!]$ of cardinality k_1. We now show that a.s., for $a > 0$,

$$\int_{\tau-a}^{\tau+a} \sum_{i=1}^{N} \left\| \sum_{j\neq i} \frac{X_s^i - X_s^j}{\|X_s^i - X_s^j\|^2} \right\| \mathrm{d}s = \infty,$$

which indeed shows the nonintegrability of the drift term. Since τ is an instant of K-collision, there exists $a > 0$ small enough that during $[\tau - a, \tau + a] \subset [0, \zeta)$, the particles labeled in K are far from the particles labeled in K^c. It clearly suffices to show that $Z = \infty$ a.s., where

$$Z = \int_{\tau-a}^{\tau+a} \sum_{i\in K} \left\| \sum_{j\in K,\, j\neq i} \frac{X_s^i - X_s^j}{\|X_s^i - X_s^j\|^2} \right\| \mathrm{d}s.$$

But

$$Z = \int_{\tau-a}^{\tau+a} \frac{f(V_s)}{\sqrt{R_K(X_s)}} \mathrm{d}s, \quad \text{where} \quad V_s = (V_s^i)_{i\in K} \text{ is defined by } V_s^i = \frac{X_s^i - S_K(X_s)}{\sqrt{R_K(X_s)}},$$

so that V_s a.s. belongs to $\mathbb{S}_K = \{(v^i)_{i\in K} \in (\mathbb{R}^2)^{|K|} : \sum_{i\in K} v^i = 0,\ \sum_{i\in K} \|v^i\|^2 = 1\}$, and where

$$f(v) = \sum_{i\in K} \left\| \sum_{j\in K,\, j\neq i} \frac{v^i - v^j}{\|v^i - v^j\|^2} \right\|$$

for each $v \in \mathbb{S}_K$. Since the invariant measure $\mathbf{m}$ of X satisfies $\mathbf{m}(E_2^c) = 0$, it is a.s. true that $X_s \in E_2$ for a.e. $s \in [0, \zeta)$ (at least for a.e. initial condition), so that a.s., $f(V_s)$ is well-defined for a.e. $s \in [0, \zeta)$. We now show that f is bounded from below on $\mathbb{S}_K$. We have

$$f(v) \geq \max_{i \in K} \Big\| \sum_{j \in K, j \neq i} \frac{v^i - v^j}{\|v^i - v^j\|^2} \Big\| \geq \sqrt{\frac{1}{|K|} \sum_{i \in K} \Big\| \sum_{j \in K, j \neq i} \frac{v^i - v^j}{\|v^i - v^j\|^2} \Big\|^2}.$$

Using now the Cauchy–Schwarz inequality and the fact that $\sum_{i \in K} \|v^i\|^2 = 1$, we find that

$$\begin{aligned} f(v) &\geq \frac{1}{\sqrt{|K|}} \sum_{i \in K} \sum_{j \in K, j \neq i} \frac{v^i - v^j}{\|v^i - v^j\|^2} \cdot v^i \\ &= \frac{1}{2\sqrt{|K|}} \sum_{i,j \in K, j \neq i} \frac{v^i - v^j}{\|v^i - v^j\|^2} \cdot (v^i - v^j) = \frac{|K|(|K|-1)}{2\sqrt{|K|}}. \end{aligned}$$

To conclude that $Z = \infty$ a.s., it remains to verify that $\int_{\tau-a}^{\tau+a} [R_K(X_s)]^{-1/2} \, \mathrm{d}s = \infty$ a.s. By Section 3.4, $R_K(X)$ behaves like a squared Bessel process of dimension $d_{\theta,N}(k_1)$ during $[\tau - a, \tau + a]$. Since $d_{\theta,N}(k_1) \in (0, 1)$ and $R_K(X_\tau) = 0$, we conclude that indeed, $\int_{\tau-a}^{\tau+a} [R_K(X_s)]^{-1/2} \, \mathrm{d}s = \infty$ a.s.: this can be shown by comparison with the 1-dimensional Brownian motion.

4. Construction of the Keller–Segel particle system

The aim of this section is to build a first version of the Keller–Segel particle system using the book of Fukushima–Oshima–Takeda [11]. We also build an $\mathbb{S}$-valued process for later use.

Proposition 6. *Fix $N \geq 2$ and $\theta > 0$ such that $N > \theta$, recall that $k_0 = \lceil 2N/\theta \rceil$ and that μ and β were defined in (4) and (8). Set $\mathcal{X} = E_{k_0}$ and $\mathcal{X}_\triangle = \mathcal{X} \cup \{\triangle\}$, as well as $\mathcal{U} = \mathbb{S} \cap E_{k_0}$ and $\mathcal{U}_\triangle = \mathcal{U} \cup \{\triangle\}$, where $\triangle$ is a cemetery point.*

(i) *There exists a unique diffusion $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t \geq 0}, (\mathbb{P}_x^X)_{x \in \mathcal{X}_\triangle})$ with values in $\mathcal{X}_\triangle$, which is μ-symmetric, with regular Dirichlet space $(\mathcal{E}^X, \mathcal{F}^X)$ on $L^2((\mathbb{R}^2)^N, \mu)$ with core $C_c^\infty(\mathcal{X})$ defined by*

$$\mathcal{E}^X(\varphi, \varphi) = \frac{1}{2} \int_{(\mathbb{R}^2)^N} \|\nabla \varphi\|^2 \, \mathrm{d}\mu \quad \text{for all } \varphi \in C_c^\infty(\mathcal{X}).$$

We call such a process a $QKS(\theta, N)$-process *and denote by $\zeta = \inf\{t \geq 0 : X_t = \triangle\}$ its life-time.*

(ii) *There exists a unique diffusion $\mathbb{U} = (\Omega^U, \mathcal{M}^U, (U_t)_{t \geq 0}, (\mathbb{P}_u^U)_{u \in \mathcal{U}_\triangle})$ with values in $\mathcal{U}_\triangle$, which is β-symmetric, with regular Dirichlet space $(\mathcal{E}^U, \mathcal{F}^U)$ on $L^2(\mathbb{S}, \beta)$ with core $C_c^\infty(\mathcal{U})$ defined by*

$$\mathcal{E}^U(\varphi, \varphi) = \frac{1}{2} \int_{\mathbb{S}} \|\nabla_{\mathbb{S}} \varphi\|^2 \, \mathrm{d}\beta \quad \text{for all } \varphi \in C_c^\infty(\mathcal{U}).$$

We call such a process a $QSKS(\theta, N)$-process *and denote by* $\xi = \inf\{t \geq 0 : U_t = \triangle\}$ *its life-time.*

The proof that we can build a $KS(\theta, N)$-process, i.e. a $QKS(\theta, N)$-process such that $\mathbb{P}_x^X \circ X_t^{-1}$ has density for all $x \in E_2$ and all $t > 0$, will be given in Section 11.

We refer to Section B.1 for some explanations of the notions used in this proposition: link between a diffusion (i.e. a continuous Hunt process), its generator, semigroup and its Dirichlet space, definition of the one-point compactification topology, i.e. the topology endowing $\mathcal{X}_\triangle$ and $\mathcal{U}_\triangle$, and about the "quasi-everywhere" notion. The state $\triangle$ is absorbing, i.e. $X_t = \triangle$ for all $t \geq \zeta$ and $U_t = \triangle$ for all $t \geq \xi$.

Remark 7. By definition of the one-point compactification topology, for any increasing sequence $(\mathcal{K}_n)_{n\geq 1}$ of compact subsets of $\mathcal{X}$ such that $\bigcup_{n\geq 1} \mathcal{K}_n = \mathcal{X}$, we have $\zeta = \lim_{n\to\infty} \inf\{t \geq 0 : X_t \notin \mathcal{K}_n\}$.

Similarly, for any increasing sequence $(\mathcal{L}_n)_{n\geq 1}$ of compact subsets of $\mathcal{U}$ such that $\bigcup_{n\geq 1} \mathcal{L}_n = \mathcal{U}$, we have $\xi = \lim_{n\to\infty} \inf\{t \geq 0 : U_t \notin \mathcal{L}_n\}$.

The uniqueness stated e.g. in Proposition 6 (i) has to be understood in the following sense (see [11, Theorem 4.2.8, p. 167)]: if we have another diffusion process $\mathbb{Y} = (\Omega^Y, \mathcal{M}^Y, (Y_t)_{t\geq 0}, (\mathbb{P}_x^Y)_{x\in\mathcal{X}})$ enjoying the same properties, then quasi-everywhere, the law of $(Y_t)_{t\geq 0}$ under $\mathbb{P}_x^Y$ equals the law of $(X_t)_{t\geq 0}$ under $\mathbb{P}_x^X$. The quasi-everywhere notion depends on the Hunt process under consideration but, as recalled in Section B.1, two Hunt processes with the same Dirichlet space share the same quasi-everywhere notion.

Proof of Proposition 6. We start with (i). We consider the bilinear form $\mathcal{E}^X$ on $C_c^\infty(\mathcal{X})$ defined by $\mathcal{E}^X(\varphi, \varphi) = \frac{1}{2} \int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \,\mathrm{d}\mu$. It is well-defined, since μ is Radon on $\mathcal{X} = E_{k_0}$ by Proposition A.1.

We first show that it is closable (see [11, p. 2]), i.e. if $(\varphi_n)_{n\geq 1} \subset C_c^\infty(\mathcal{X})$ is such that $\lim_n \varphi_n = 0$ in $L^2((\mathbb{R}^2)^N, \mu)$ and $\lim_{n,m} \mathcal{E}^X(\varphi_n - \varphi_m, \varphi_n - \varphi_m) = 0$, then $\lim_n \mathcal{E}^X(\varphi_n, \varphi_n) = 0$. Since $\nabla\varphi_n$ is a Cauchy sequence in $L^2((\mathbb{R}^2)^N, \mu)$, it converges to a limit g and it suffices to prove that $g = 0$ a.e. For $\psi \in C_c^\infty(E_2, (\mathbb{R}^2)^N)$, we have $\int_{(\mathbb{R}^2)^N} g \cdot \psi \,\mathrm{d}\mu = \lim_n \int_{(\mathbb{R}^2)^N} \nabla\varphi_n \cdot \psi \,\mathrm{d}\mu$. But, recalling (4),

$$\begin{aligned}\int_{(\mathbb{R}^2)^N} \nabla\varphi_n \cdot \psi \,\mathrm{d}\mu &= \int_{(\mathbb{R}^2)^N} \nabla\varphi_n(x) \cdot \psi(x)\mathbf{m}(x) \,\mathrm{d}x \\ &= -\int_{(\mathbb{R}^2)^N} \varphi_n(x) \operatorname{div}(\mathbf{m}(x)\psi(x)) \,\mathrm{d}x.\end{aligned}$$

Thus by the Cauchy–Schwarz inequality,

$$\left|\int_{(\mathbb{R}^2)^N} \nabla\varphi_n \cdot \psi \,\mathrm{d}\mu\right| \leq \left(\int_{(\mathbb{R}^2)^N} \varphi_n^2 \,\mathrm{d}\mu\right)^{1/2} \left(\int_{(\mathbb{R}^2)^N} \frac{|\operatorname{div}(\mathbf{m}(x)\psi(x))|^2}{\mathbf{m}(x)} \,\mathrm{d}x\right)^{1/2},$$

which tends to 0 since $\lim_n \varphi_n = 0$ in $L^2((\mathbb{R}^2)^N, \mu)$ and $\psi \in C_c^\infty(E_2, (\mathbb{R}^2)^N)$, and since $\mathbf{m}$ is smooth and positive on E_2. Thus $\int_{(\mathbb{R}^2)^N} g \cdot \psi \,\mathrm{d}\mu = 0$ for all $\psi \in C_c^\infty(E_2, (\mathbb{R}^2)^N)$, so that $g = 0$ a.e.

We can thus consider the extension of $\mathcal{E}^X$ to $\mathcal{F}^X = \overline{C_c^\infty(\mathcal{X})}^{\mathcal{E}_1^X}$, where we have set $\mathcal{E}_1^X(\varphi,\varphi) = \int_{(\mathbb{R}^2)^N} (\varphi^2 + \frac{1}{2}\|\nabla\varphi\|^2)\,\mathrm{d}\mu$ for $\varphi \in C_c^\infty(\mathcal{X})$.

Next, $(\mathcal{E}^X, \mathcal{F}^X)$ is obviously regular with core $C_c^\infty(\mathcal{X})$ (see [11, p. 6]), because $C_c^\infty(\mathcal{X})$ is dense in $\mathcal{F}^X$ for the norm associated to $\mathcal{E}_1^X$ by definition of $\mathcal{F}^X$, and $C_c^\infty(\mathcal{X})$ is dense, for the uniform norm, in $C_c(\mathcal{X})$. It is also strongly local (see [11, p. 6]), i.e. $\mathcal{E}^X(\varphi,\psi) = \frac{1}{2}\int_{(\mathbb{R}^2)^N} \nabla\varphi\cdot\nabla\psi\,\mathrm{d}\mu = 0$ if $\varphi,\psi \in C_c^\infty(\mathcal{X})$ with φ constant on a neighborhood of supp ψ.

Then [11, Theorems 7.2.2, p. 380, and 4.2.8, p. 167] imply the existence and uniqueness of a Hunt process $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\ge 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}_\triangle})$, with values in $\mathcal{X}_\triangle$, which is μ-symmetric, has Dirichlet space is $(\mathcal{E}^X, \mathcal{F}^X)$, and $t \mapsto X_t$ is $\mathbb{P}_x^X$-a.s. continuous on $[0,\zeta)$ for all $x \in \mathcal{X}$, where $\zeta = \inf\{t \ge 0 : X_t = \triangle\}$.

Furthermore, since $\mathcal{E}^X$ is strongly local, we know from [11, Theorem 4.5.3, p. 186] that we can choose $\mathbb{X}$ (modifying $\mathbb{P}_x^X$ only on a properly exceptional set) in such a way that $\mathbb{P}_x(\zeta < \infty, X_{\zeta-} = \triangle) = 1$ for all $x \in \mathcal{X}$. This implies that for all $x \in \mathcal{X}$, $\mathbb{P}_x$-a.s., the map $t \mapsto X_t$ is continuous from $[0,\infty)$ to $\mathcal{X}_\triangle$ endowed with the one-point compactification topology recalled in Section B.1. Hence $\mathbb{X}$ is a diffusion.

For (ii), the same strategy applies. The only difference is the integration by parts to be used for closability: for $\varphi \in C_c^1(\mathcal{U})$ and $\psi \in C_c^1(\mathbb{S}\cap E_2, (\mathbb{R}^2)^N)$, classically,

$$\int_{\mathbb{S}} (\nabla_{\mathbb{S}}\varphi)\cdot\psi\,\mathrm{d}\beta = \int_{\mathbb{S}} (\nabla_{\mathbb{S}}\varphi(u))\cdot\psi(u)\mathbf{m}(u)\,\sigma(\mathrm{d}u) = -\int_{\mathbb{S}} \varphi(u)\,\mathrm{div}_{\mathbb{S}}(\mathbf{m}(u)\psi(u))\,\sigma(\mathrm{d}u). \tag{20}$$

This can be shown naively using Lemma A.2. ∎

We now make explicit the generators of $\mathbb{X}$ and $\mathbb{U}$ when applied to some functions enjoying a few properties. See Section B.1 for a precise definition of the generator of a Hunt process. We have to introduce some notation.

For $\varphi \in C^\infty((\mathbb{R}^2)^N)$, $\alpha \in (0,1]$ and $x \in (\mathbb{R}^2)^N$, we set

$$\mathcal{L}_\alpha^X\varphi(x) = \frac{1}{2}\Delta\varphi(x) - \frac{\theta}{N}\sum_{1\le i\ne j\le N}\frac{x^i - x^j}{\|x^i - x^j\|^2 + \alpha}\cdot(\nabla\varphi(x))^i = \frac{1}{2\mathbf{m}_\alpha(x)}\,\mathrm{div}[\mathbf{m}_\alpha(x)\nabla\varphi(x)], \tag{21}$$

where

$$\mathbf{m}_\alpha(x) = \prod_{1\le i\ne j\le N}(\|x^i - x^j\|^2 + \alpha)^{-\theta/(2N)}.$$

This is in accordance with (4), in the sense that $\mathbf{m}_0 = \mathbf{m}$. Formula (21) makes sense for $x \in E_2$ when $\alpha = 0$ (with $\mathbf{m}_\alpha$ replaced by $\mathbf{m}$) and we recall that for $\varphi \in C^\infty((\mathbb{R}^2)^N)$ and $x \in E_2$, $\mathcal{L}^X\varphi(x)$ was defined in (5) by $\mathcal{L}^X\varphi(x) = \mathcal{L}_0^X\varphi(x)$. We will often use the fact that for all $\varphi,\psi \in C^\infty((\mathbb{R}^2)^N)$, and all $x \in (\mathbb{R}^2)^N$ and $\alpha \in (0,1]$,

$$\mathcal{L}_\alpha^X(\varphi\psi)(x) = \varphi(x)\mathcal{L}_\alpha^X\psi(x) + \psi(x)\mathcal{L}_\alpha^X\varphi(x) + \nabla\varphi(x)\cdot\nabla\psi(x). \tag{22}$$

For $\varphi \in C^\infty(\mathbb{S})$, $\alpha \in (0,1]$ and $u \in \mathbb{S}$, we set

$$\mathcal{L}^U_\alpha \varphi(u) = \frac{1}{2}\Delta_{\mathbb{S}}\varphi(u) - \frac{\theta}{N} \sum_{1\le i\ne j\le N} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (\nabla_{\mathbb{S}}\varphi(u))^i$$
$$= \frac{1}{2\mathbf{m}_\alpha(u)} \operatorname{div}_{\mathbb{S}}[\mathbf{m}_\alpha(u)\nabla_{\mathbb{S}}\varphi(u)]. \tag{23}$$

This formula makes sense for $u \in \mathbb{S} \cap E_2$ when $\alpha = 0$ (with $\mathbf{m}_\alpha$ replaced by $\mathbf{m}$) and we set $\mathcal{L}^U\varphi(u) = \mathcal{L}^U_0\varphi(u)$ for $\varphi \in C^\infty(\mathbb{S})$ and $u \in \mathbb{S} \cap E_2$.

Remark 8. (i) Denote by $(\mathcal{A}^X, \mathcal{D}_{A^X})$ the generator of the process $\mathbb{X}$ of Proposition 6 (i). If $\varphi \in C^\infty_c(\mathcal{X})$ satisfies $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}^X_\alpha \varphi(x)| < \infty$, then $\varphi \in \mathcal{D}_{A^X}$ and $\mathcal{A}^X\varphi = \mathcal{L}^X\varphi$.

(ii) Denote by $(\mathcal{A}^U, \mathcal{D}_{A^U})$ the generator of the process $\mathbb{U}$ of Proposition 6 (ii). If $\varphi \in C^\infty_c(\mathcal{U})$ satisfies $\sup_{\alpha\in(0,1]} \sup_{u\in\mathbb{S}} |\mathcal{L}^U_\alpha \varphi(u)| < \infty$, then $\varphi \in \mathcal{D}_{A^U}$ and $\mathcal{A}^U\varphi = \mathcal{L}^U\varphi$.

Proof. To check (i), it suffices by (B.1) to verify that (a) $\varphi \in \mathcal{F}^X$, (b) $\mathcal{L}^X\varphi \in L^2(\mathcal{X}, \mu)$ and (c) for all $\psi \in \mathcal{F}^X$, we have $\mathcal{E}^X(\varphi, \psi) = -\int_{\mathcal{X}} (\mathcal{L}^X\varphi)\psi \mathrm{d}\mu$.

Point (a) is clear, since $\varphi \in C^\infty_c(\mathcal{X})$. Point (b) follows from the facts that μ is Radon on $\mathcal{X}$, that φ is compactly supported in $\mathcal{X}$ and that $\mathcal{L}^X\varphi \in L^\infty((\mathbb{R}^2)^N, \mathrm{d}x)$, because $\mathcal{L}^X\varphi(x) = \lim_{\alpha\to 0} \mathcal{L}^X_\alpha\varphi(x)$ for all $x \in E_2$. Concerning (c) it suffices, by definition of $(\mathcal{E}^X, \mathcal{F}^X)$ and since $\mathcal{L}^X\varphi \in L^2(\mathcal{X}, \mu)$, to show that for all $\psi \in C^\infty_c(\mathcal{X})$, we have $\frac{1}{2}\int_{(\mathbb{R}^2)^N} \nabla\varphi \cdot \nabla\psi \,\mathrm{d}\mu = -\int_{(\mathbb{R}^2)^N} (\mathcal{L}^X\varphi)\psi \,\mathrm{d}\mu$. But for $\alpha \in (0,1]$, by a standard integration by parts, since φ, ψ and $\mathbf{m}_\alpha$ are smooth,

$$\frac{1}{2}\int_{(\mathbb{R}^2)^N} \nabla\varphi(x) \cdot \nabla\psi(x)\mathbf{m}_\alpha(x)\,\mathrm{d}x = -\frac{1}{2}\int_{(\mathbb{R}^2)^N} \operatorname{div}(\mathbf{m}_\alpha(x)\nabla\varphi(x))\psi(x)\,\mathrm{d}x$$
$$= -\int_{(\mathbb{R}^2)^N} [\mathcal{L}^X_\alpha\varphi(x)]\psi(x)\mathbf{m}_\alpha(x)\,\mathrm{d}x.$$

Letting $\alpha \to 0$ we conclude the proof by dominated convergence, since $\mathbf{m}_\alpha \to \mathbf{m}$ and $\mathcal{L}^X_\alpha\varphi \to \mathcal{L}^X\varphi$ a.e., since $|\nabla\varphi(x) \cdot \nabla\psi(x)\mathbf{m}_\alpha(x)| + |[\mathcal{L}^X_\alpha\varphi(x)]\psi(x)\mathbf{m}_\alpha(x)| \le C\mathbf{1}_{\{x\in\mathcal{K}\}}\mathbf{m}(x)$ for some constant C and $\mathcal{K} = \operatorname{supp}\psi$ which is compact in $\mathcal{X}$, and since $\mu(\mathcal{K}) = \int_{\mathcal{K}} \mathbf{m}(x)\,\mathrm{d}x < \infty$.

The proof of (ii) is exactly the same, using the fact that if $\varphi, \psi \in C^\infty(\mathbb{S})$, then

$$\frac{1}{2}\int_{\mathbb{S}} \nabla_{\mathbb{S}}\varphi \cdot \nabla_{\mathbb{S}}\psi \;\mathbf{m}_\alpha\,\mathrm{d}\sigma = -\frac{1}{2}\int_{\mathbb{S}} \operatorname{div}_{\mathbb{S}}(\mathbf{m}_\alpha\nabla_{\mathbb{S}}\varphi)\psi\,\mathrm{d}\sigma = -\int_{\mathbb{S}} [\mathcal{L}^U_\alpha\varphi]\psi\mathbf{m}_\alpha\,\mathrm{d}\sigma,$$

which can be shown naively using the projection $\Phi_{\mathbb{S}}$ (see (10)) and Lemma A.2. ■

We end the section with an irreducibility/recurrence/transience study of the spherical process; see Section B.1 again for definitions.

Lemma 9. *Fix $N \ge 2$ and $\theta > 0$ such that $N > \theta$ and consider the process $\mathbb{U}$ and its Dirichlet space $(\mathcal{E}^U, \mathcal{F}^U)$ as in Proposition* 6 (ii).

(i) $(\mathcal{E}^U, \mathcal{F}^U)$ *is irreducible and either*

- $(\mathcal{E}^U, \mathcal{F}^U)$ *is recurrent and in particular nonexploding and for all measurable* $A \subset \mathcal{U}$ *such that* $\beta(A) > 0$*, we have* $\mathbb{P}_u^U(\limsup_{t\to\infty}\{U_t \in A\}) = 1$ *quasi-everywhere; or*
- $(\mathcal{E}^U, \mathcal{F}^U)$ *is transient and in particular for all compact subsets* $\mathcal{K}$ *of* $\mathcal{U}$*, we have* $\mathbb{P}_u^U(\liminf_{t\to\infty}\{U_t \in \mathcal{K}\}) = 0$ *quasi-everywhere.*

(ii) *If* $d_{\theta,N}(N-1) > 0$*, then* $(\mathcal{E}^U, \mathcal{F}^U)$ *is recurrent.*

In the transient case, one might also prove that $\mathbb{P}_u^U(\limsup_{t\to\infty}\{U_t \in \mathcal{K}\}) = 0$, but this would be useless for our purpose.

Proof of Lemma 9. (i) We first show that in any case, $(\mathcal{E}^U, \mathcal{F}^U)$ is irreducible. By [11, Corollary 4.6.4, p. 195] and since $\mathcal{E}^U(\varphi,\varphi) = \frac{1}{2}\int_{\mathbb{S}} \|\nabla_{\mathbb{S}}\varphi\|^2 \mathbf{m}\, d\sigma$ with $\mathbf{m}$ bounded from below by a constant (on $\mathbb{S}$), it suffices to prove that the σ-symmetric Hunt process with regular Dirichlet space $(\mathcal{E}, \mathcal{F})$ on $L^2(\mathcal{U}, \sigma)$ with core $C_c^\infty(\mathcal{U})$ such that $\mathcal{E}(\varphi,\varphi) = \frac{1}{2}\int_{\mathbb{S}} \|\nabla_{\mathbb{S}}\varphi\|^2 d\sigma$ for all $\varphi \in C_c^\infty(\mathcal{U})$ is irreducible. But this Hunt process is nothing but an $\mathbb{S}$-valued Brownian motion. This Brownian motion is *a priori* killed when it gets out of $\mathcal{U}$, but a.s. this never occurs since such a Brownian motion never has two (bi-dimensional) coordinates equal. This $\mathbb{S}$-valued Brownian motion is of course irreducible. We conclude from [11, Lemma 1.6.4, p. 55] that $(\mathcal{E}^U, \mathcal{F}^U)$ is either recurrent or transient.

When $(\mathcal{E}^U, \mathcal{F}^U)$ is recurrent, [11, Theorem 4.7.1 (iii), p. 202] gives us the result.

When $(\mathcal{E}^U, \mathcal{F}^U)$ is transient, we fix a compact set $\mathcal{K}$ of $\mathcal{U}$ and we know from Lemma A.3 that $\beta(\mathcal{K}) < \infty$, so that by definition of transience, for β-a.e. $u \in \mathcal{U}$, we have $\mathbb{E}_u^U[\int_0^\infty \mathbf{1}_{\mathcal{K}}(U_s)\, ds] < \infty$. Setting $\tau_{\mathcal{K}^c} = \inf\{t \geq 0 : U_t \notin \mathcal{K}\}$, we see in particular that $\mathbb{P}_u^U(\tau_{\mathcal{K}^c} < \infty) = 1$ for β-a.e. $u \in \mathcal{U}$. But, by [11, (4.1.9), p. 155], $u \mapsto \mathbb{P}_u^U(\tau_{\mathcal{K}^c} < \infty)$ is finely continuous. Using [11, Lemma 4.1.5, p. 155], we deduce that $\mathbb{P}_u^U(\tau_{\mathcal{K}^c} < \infty) = 1$ quasi-everywhere. The Markov property allows us to conclude the proof.

(ii) We recall from Proposition A.3 that $\beta(\mathbb{S}) < \infty$, because $d_{\theta,N}(N-1) > 0$ implies that $k_0 \geq N$ (see Lemma 1). Moreover, $k_0 \geq N$ implies that $E_{k_0} \supset E_N \supset \mathbb{S}$, whence $\mathcal{U} = E_{k_0} \cap \mathbb{S} = \mathbb{S}$ is compact; the process cannot explode, i.e. $\xi = \infty$. Consequently, $(\mathcal{E}^U, \mathcal{F}^U)$ is recurrent, since $\varphi \equiv 1$ belongs to $L^1(\mathcal{U}, \beta)$ and $\mathbb{E}_u^U[\int_0^\infty \varphi(U_s)\, ds] = \mathbb{E}_u^U[\xi] = \infty$. Indeed, as recalled in Section B.1, if $(\mathcal{E}^U, \mathcal{F}^U)$ was transient, we would have $\mathbb{E}_u^U[\int_0^\infty \varphi(U_s)\, ds] < \infty$ for all $\varphi \in L^1(\mathcal{U}, \beta)$, with the convention that $\varphi(\triangle) = 0$. ∎

5. Decomposition

The goal of this section is to prove the following decomposition of the Keller–Segel particle system defined in Proposition 6 (i). This decomposition is remarkable and crucial for our purpose.

Proposition 10. *Fix* $N \geq 2$ *and* $\theta > 0$ *such that* $N > \theta$*, and recall that* $k_0 = \lceil 2N/\theta \rceil$*,* $\mathcal{X} = E_{k_0}$ *and* $\mathcal{U} = \mathbb{S} \cap E_{k_0}$*.*

For $x \in E_N$, *set* $r = R_{[\![1,N]\!]}(x) > 0$, $z = S_{[\![1,N]\!]}(x) \in \mathbb{R}^2$ *and* $u = (x - \gamma(z))/\sqrt{r} \in \mathbb{S}$ *and consider three independent processes:*

- $(M_t)_{t\geq 0}$, *a* 2*-dimensional Brownian motion with diffusion constant* $N^{-1/2}$ *starting from* z,
- $(D_t)_{t\geq 0}$, *a squared Bessel process of dimension* $d_{\theta,N}(N)$ *starting from* r *and killed when it gets out of* $(0,\infty)$, *with life-time* $\tau_D = \inf\{t \geq 0 : D_t = \triangle\}$,
- $(U_t)_{t\geq 0}$, *a QSKS*(θ, N)*-process starting from* u, *with life-time* $\xi = \inf\{t \geq 0 : U_t = \triangle\}$.

Set $A_t = \int_0^{t\wedge\tau_D} D_s^{-1}\,\mathrm{d}s$, *with generalized inverse* $\rho_t = \inf\{s > 0 : A_s > t\}$. *Define* $Y_t = \Psi(M_t, D_t, U_{A_t})$, *where we recall from* (9) *that* $\Psi(z,r,u) = \gamma(z) + \sqrt{r}u \in E_N$ *when* $(z,r,u) \in \mathbb{R}^2 \times (0,\infty) \times \mathbb{S}$, *and* $\Psi(z,r,u) = \triangle$ *when* $r = \triangle$ *or* $u = \triangle$. *Observe that the life-time of* Y *equals* $\zeta' = \rho_\xi \wedge \tau_D$.

Consider also a QKS(θ,N)*-process* $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}_\triangle})$, *with life-time* ζ, *and* $\mathbb{X}^* = (\Omega^X, \mathcal{M}^X, (X_t^*)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in(\mathcal{X}\cap E_N)\cup\{\triangle\}})$, *where* $X_t^* = X_t\mathbf{1}_{\{t<\tau\}} + \triangle\mathbf{1}_{\{t\geq\tau\}}$ *and where* $\tau = \inf\{t \geq 0 : R_{[\![1,N]\!]}(X_t) \notin (0,\infty)\}$. *In other words,* $\mathbb{X}^*$ *is the version of* $\mathbb{X}$ *killed when it gets out of* E_N. *The life-time of* $\mathbb{X}^*$ *is* τ.

The law of $(Y_t)_{t\geq 0}$ *is the same as that of* $(X_t^*)_{t\geq 0}$ *under* $\mathbb{P}_x^X$, *quasi-everywhere in* $\mathcal{X} \cap E_N$.

We make the convention that $R_{[\![1,N]\!]}(\triangle) = 0$, so that $\tau \in [0,\zeta]$. Since $R_{[\![1,N]\!]}(Y_t) = D_t$ and $S_{[\![1,N]\!]}(Y_t) = M_t$ for all $t \in [0,\zeta')$, Proposition 10 in particular implies that $(R_{[\![1,N]\!]}(X_t))_{t\geq 0}$ and $(S_{[\![1,N]\!]}(X_t))_{t\geq 0}$ are a squared Bessel process and a Brownian motion, independent until the first time $(R_{[\![1,N]\!]}(X_t))_{t\geq 0}$ vanishes. This actually holds true until explosion, as shown in Lemma 11 below. The quasi-everywhere notion refers to the Hunt process $\mathbb{X}$. Observe that when $\theta \geq 2$, we have $k_0 \leq N$, so that $\mathcal{X} \cap E_N = \mathcal{X}$ and $\mathbb{X} = \mathbb{X}^*$.

Proof of Proposition 10. We divide the proof into several steps. The first two steps are more or less classical, even if we give all the details: we determine the Dirichlet spaces of the three processes $(M_t)_{t\geq 0}$, $(D_t)_{t\geq 0}$ and $(U_t)_{t\geq 0}$ involved in the construction of $(Y_t)_{t\geq 0}$; then we compute the Dirichlet space of $(D_{\rho_t})_{t\geq 0}$; next we identify the Dirichlet space of $(D_{\rho_t}, U_t)_{t\geq 0}$, which allows us to find the one of $(D_t, U_{A_t})_{t\geq 0}$ by a second time-change; by concatenation, we deduce the Dirichlet space of $(M_t, D_t, U_{A_t})_{t\geq 0}$. The main computations are handled in Steps 3 and 4, where we find the Dirichlet space of $(Y_t)_{t\geq 0}$, which allows us to conclude the proof in Step 5 by uniqueness.

Step 1. First, take $\mathbb{U} = (\Omega^U, \mathcal{M}^U, (U_t)_{t\geq 0}, (\mathbb{P}_u^U)_{u\in\mathcal{U}_\triangle})$ as in Proposition 6(ii). Next, consider a 2-dimensional Brownian motion $\mathbb{M} = (\Omega^M, \mathcal{M}^M, (M_t)_{t\geq 0}, (\mathbb{P}_z^M)_{z\in\mathbb{R}^2})$ with diffusion constant $N^{-1/2}$. We know from [11, Example 4.2.1, p. 167] that $\mathbb{M}$ is a $\mathrm{d}z$-symmetric (here $\mathrm{d}z$ is the Lebesgue measure on $\mathbb{R}^2$) diffusion with regular Dirichlet space $(\mathcal{E}^M, \mathcal{F}^M)$ on $L^2(\mathbb{R}^2, \mathrm{d}z)$ with core $C_c^\infty(\mathbb{R}^2)$, and for all $\varphi \in C_c^\infty(\mathbb{R}^2)$,

$$\mathcal{E}^M(\varphi,\varphi) = \frac{1}{2N}\int_{\mathbb{R}^2} \|\nabla_z\varphi(z)\|^2\,\mathrm{d}z. \tag{24}$$

Finally, let $\mathbb{D} = (\Omega^D, \mathcal{M}^D, (D_t)_{t\geq 0}, (\mathbb{P}_r^D)_{r\in\mathbb{R}_+^*\cup\{\triangle\}})$ be a squared Bessel process of dimension $d_{\theta,N}(N)$ killed when it gets out of $\mathbb{R}_+^* = (0,\infty)$, and set $\nu = d_{\theta,N}(N)/2 - 1$ (see Revuz–Yor [21, p. 443]). Fukushima [10, Theorem 3.3] tells us that $\mathbb{D}$ is an $r^\nu \mathrm{d}r$-symmetric diffusion (here $\mathrm{d}r$ is the Lebesgue measure on $\mathbb{R}_+^*$) with regular Dirichlet space $(\mathcal{E}^D, \mathcal{F}^D)$ on $L^2(\mathbb{R}_+, r^\nu \mathrm{d}r)$ with core $C_c^\infty(\mathbb{R}_+^*)$, where for all $\varphi \in C_c^\infty(\mathbb{R}_+^*)$,

$$\mathcal{E}^D(\varphi,\varphi) = 2\int_{\mathbb{R}_+} |\varphi'(r)|^2 r^{\nu+1}\,\mathrm{d}r. \tag{25}$$

Together with [10, Theorem 3.3], this uses the fact that the scale function and the speed measure of $(D_t)_{t\geq 0}$ are respectively $r \mapsto r^{-\nu}$ and $-[r^\nu/(2\nu)]\mathrm{d}r$. Actually, we do not take the speed measure as reference measure but $r^\nu \mathrm{d}r$, which is the same up to a constant.

Step 2. We apply Lemma B.3 to $\mathbb{D}$ with $g(r) = 1/r$, i.e. with $A_t = \int_0^t D_s^{-1}\,\mathrm{d}s = \int_0^{t\wedge\tau_D} D_s^{-1}\,\mathrm{d}s$ thanks to the convention $\triangle^{-1} = 0$ and recall that ρ is its generalized inverse; setting $D_{\rho_t} = D_{\rho_t}\mathbf{1}_{\{\rho_t<\infty\}} + \triangle\mathbf{1}_{\{\rho_t=\infty\}}$, we find that

$$\mathbb{D}_\rho = (\Omega^D, \mathcal{M}^D, (D_{\rho_t})_{t\geq 0}, (\mathbb{P}_r^D)_{r\in\mathbb{R}_+^*})$$

is an $r^{\nu-1}\mathrm{d}r$-symmetric $(\mathbb{R}_+^*\cup\{\triangle\})$-valued diffusion with Dirichlet space $(\mathcal{E}^{D_\rho}, \mathcal{F}^{D_\rho})$ on $L^2(\mathbb{R}_+, r^{\nu-1}\mathrm{d}r)$, regular with core $C_c^\infty(\mathbb{R}_+^*)$ such that for all $\varphi \in C_c^\infty(\mathbb{R}_+^*)$,

$$\mathcal{E}^{D_\rho}(\varphi,\varphi) = \mathcal{E}^D(\varphi,\varphi) = 2\int_{\mathbb{R}_+} |\varphi'(r)|^2 r^{\nu+1}\,\mathrm{d}r = 2\int_{\mathbb{R}_+} |r\varphi'(r)|^2 r^{\nu-1}\,\mathrm{d}r. \tag{26}$$

We use Lemma B.5 and the notation therein: recalling that by definition, $\mathcal{M}^{(D,U)} = \sigma((D_{\rho_t}, U_t) : t \geq 0)$, with the convention that $(r,\triangle) = (\triangle, u) = (\triangle,\triangle) = \triangle$, and $\mathbb{P}_{(r,u)}^{(D,U)} = \mathbb{P}_r^D \otimes \mathbb{P}_u^U$ if $(r,u) \in \mathbb{R}_+^* \times \mathcal{U}$ and $\mathbb{P}_\triangle^{(D,U)} = \mathbb{P}_\triangle^D \otimes \mathbb{P}_\triangle^U$, we find that

$$(\mathbb{D},\mathbb{U}) = \big(\Omega^D\times\Omega^U, \mathcal{M}^{(D,U)}, (D_{\rho_t}, U_t)_{t\geq 0}, (\mathbb{P}_{(r,u)}^{(D,U)})_{(r,u)\in(\mathbb{R}_+^*\times\mathcal{U})\cup\{\triangle\}}\big)$$

is an $r^{\nu-1}\mathrm{d}r\beta(\mathrm{d}u)$-symmetric $(\mathbb{R}_+^*\times\mathcal{U})\cup\{\triangle\}$-valued diffusion with regular Dirichlet space $(\mathcal{E}^{(D_\rho,U)}, \mathcal{F}^{(D_\rho,U)})$ on $L^2(\mathbb{R}_+\times\mathbb{S}, r^{\nu-1}\mathrm{d}r\beta(\mathrm{d}u))$ with core $C_c^\infty(\mathbb{R}_+^*\times\mathcal{U})$, and for all $\varphi\in C_c^\infty(\mathbb{R}_+^*\times\mathcal{U})$,

$$\mathcal{E}^{(D_\rho,U)}(\varphi,\varphi) = \int_{\mathbb{R}_+} \mathcal{E}^U(\varphi(r,\cdot),\varphi(r,\cdot))r^{\nu-1}\,\mathrm{d}r + \int_{\mathbb{S}} \mathcal{E}^{D_\rho}(\varphi(\cdot,u),\varphi(\cdot,u))\,\beta(\mathrm{d}u).$$

We now apply Lemma B.3 to $(\mathbb{D},\mathbb{U})$ with $g(r,u) = r$ for all $r\in\mathbb{R}_+^*$ and all $u\in\mathcal{U}$. We consider the time-change $\alpha_t = \int_0^t g(D_{\rho_s}, U_s)\,\mathrm{d}s$, with the convention that $g(r,u) = 0$ as long as $(r,u) = \triangle$. We also set $B_t = \inf\{s > 0 : \alpha_s > t\}$. As we will see in a few lines,

$$(D_{\rho_{B_t}}, U_{B_t}) = (D_t, U_{A_t}) \quad \text{for all } t \geq 0. \tag{27}$$

Hence Lemma B.3 tells us that

$$(\mathbb{D},\mathbb{U}_A) = \big(\Omega^D\times\Omega^U, \mathcal{M}^{(D,U)}, (D_t, U_{A_t})_{t\geq 0}, (\mathbb{P}_{(r,u)}^{(D,U)})_{(r,u)\in(\mathbb{R}_+^*\times\mathcal{U})\cup\{\triangle\}}\big)$$

is an $r^\nu\mathrm{d}r\beta(\mathrm{d}u)$-symmetric $(\mathbb{R}_+^*\times\mathcal{U})\cup\{\triangle\}$-valued diffusion with regular Dirichlet space $(\mathcal{E}^{(D,U_A)}, \mathcal{F}^{(D,U_A)})$ on $L^2(\mathbb{R}_+\times\mathbb{S}, r^\nu\mathrm{d}r\beta(\mathrm{d}u))$ with core $C_c^\infty(\mathbb{R}_+^*\times\mathcal{U})$ and for all $\varphi \in C_c^\infty(\mathbb{R}_+^* \times \mathcal{U})$,

$$\mathcal{E}^{(D,U_A)}(\varphi,\varphi) = \mathcal{E}^{(D_\rho,U)}(\varphi,\varphi)$$
$$= \int_{\mathbb{R}_+} \mathcal{E}^U(\varphi(r,\cdot),\varphi(r,\cdot))r^{\nu-1}\,\mathrm{d}r + \int_{\mathbb{S}} \mathcal{E}^{D_\rho}(\varphi(\cdot,u),\varphi(\cdot,u))\,\beta(\mathrm{d}u). \tag{28}$$

We now check (27). Recall that D explodes at time τ_D, $A_t = \int_0^{t\wedge\tau_D} D_s^{-1}\,\mathrm{d}s$, and ρ is the generalized inverse of A. Hence $(\rho_t)_{t\in[0,A_{\tau_D})}$ is the true inverse of $(A_t)_{t\in[0,\tau_D)}$ and we have $\rho_t' = D_{\rho_t}$, whence $\rho_t = \int_0^t D_{\rho_s}\mathrm{d}s$ for $t \in [0, A_{\tau_D})$. We also have $\rho_t = \infty$ for $t \geq A_{\tau_D}$. Next, $\alpha_t = \int_0^t D_{\rho_s}\mathrm{d}s = \rho_t$ for $t \in [0, A_{\tau_D} \wedge \xi)$, because $g(D_{\rho_s}, U_s) = D_{\rho_s}$ if $(D_{\rho_s}, U_s) \neq \triangle$, i.e. if $s < A_{\tau_D} \wedge \xi$. Hence B, the generalized inverse of α, equals A during $[0, \tau_D \wedge \rho_\xi)$, thus in particular $\rho_{B_t} = t$ for $t \in [0, A_{\tau_D} \wedge \xi)$. In conclusion, (27) holds true for $t \in [0, A_{\tau_D} \wedge \xi)$. If now $t \geq \tau_D \wedge \rho_\xi$, then $B_t = \infty$, because B is the generalized inverse of α and because for all $t \geq 0$,

$$\alpha_t \leq \alpha_{A_{\tau_D}\wedge\xi} = \rho_{A_{\tau_D}\wedge\xi} = \tau_D \wedge \rho_\xi.$$

Hence, still if $t \geq \tau_D \wedge \rho_\xi$, we have $(D_{\rho_{B_t}}, U_{B_t}) = \triangle$, while $(D_t, U_{A_t}) = \triangle$ because either $t \geq \tau_D$ and thus $D_t = \triangle$, or $t \geq \rho_\xi$ and thus $A_t \geq \xi$ so that $U_{A_t} = \triangle$. We have proved (27).

Thanks to Lemma B.5, setting $\mathcal{M}^{(M,D,U)} = \sigma((M_t, D_t, U_{A_t}) : t \geq 0)$ with the convention that $(z, \triangle) = \triangle$ and setting $\mathbb{P}^{(M,D,U)}_{(z,r,u)} = \mathbb{P}^M_z \otimes \mathbb{P}^{(D,U)}_{(r,u)}$ in the case where $(z, r, u) \in \mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U}$ and $\mathbb{P}^{(M,D,U)}_\triangle = \mathbb{P}^M_\triangle \otimes \mathbb{P}^{(D,U)}_\triangle$, we conclude that

$$(\mathbb{M}, \mathbb{D}, \mathbb{U}_{\mathbb{A}}) = \big(\Omega^M \times \Omega^D \times \Omega^U, \mathcal{M}^{(M,D,U)}, (M_t, D_t, U_{A_t})_{t\geq 0},$$
$$(\mathbb{P}^{(M,D,U)}_{(z,r,u)})_{(z,r,u)\in(\mathbb{R}^2\times\mathbb{R}_+^*\times\mathcal{U})\cup\{\triangle\}}\big)$$

is a $\mathrm{d}z\, r^\nu \mathrm{d}r\beta(\mathrm{d}u)$-symmetric $(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U}) \cup \{\triangle\}$-valued diffusion with regular Dirichlet space $(\mathcal{E}^{(M,D,U_A)}, \mathcal{F}^{(M,D,U_A)})$ on $L^2(\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{S}, \mathrm{d}z\, r^\nu \mathrm{d}r\beta(\mathrm{d}u))$, and its core is $C_c^\infty(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U})$. Moreover, for all $\varphi \in C_c^\infty(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U})$,

$$\mathcal{E}^{(M,D,U_A)}(\varphi,\varphi)$$
$$= \int_{\mathbb{R}_+\times\mathbb{S}} \mathcal{E}^M(\varphi(\cdot,r,u),\varphi(\cdot,r,u))r^\nu\,\mathrm{d}r\,\beta(\mathrm{d}u) + \int_{\mathbb{R}^2} \mathcal{E}^{(D,U_A)}(\varphi(z,\cdot,\cdot),\varphi(z,\cdot,\cdot))\,\mathrm{d}z$$
$$= \int_{\mathbb{R}_+\times\mathbb{S}} \mathcal{E}^M(\varphi(\cdot,r,u),\varphi(\cdot,r,u))r^\nu\,\mathrm{d}r\,\beta(\mathrm{d}u)$$
$$+ \int_{\mathbb{R}^2\times\mathbb{S}} \mathcal{E}^{D_\rho}(\varphi(z,\cdot,u),\varphi(z,\cdot,u))\,\mathrm{d}z\,\beta(\mathrm{d}u)$$
$$+ \int_{\mathbb{R}^2\times\mathbb{R}_+} \mathcal{E}^U(\varphi(z,r,\cdot),\varphi(z,r,\cdot))\,\mathrm{d}z\,r^{\nu-1}\,\mathrm{d}r$$
$$= \int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} \Big[\frac{1}{2N}\|\nabla_z\varphi(z,r,u)\|^2 + 2r|\partial_r\varphi(z,r,u)|^2$$
$$+ \frac{1}{2r}\|\nabla_{\mathbb{S}}\varphi(z,r,u)\|^2\Big]\,\mathrm{d}z\,r^\nu\,\mathrm{d}r\,\beta(\mathrm{d}u). \tag{29}$$

For the second line, we have used (28). For the last line, we have used (24), (26) and the expression of $\mathcal{E}^U$ (see Proposition 6 (ii)).

Step 3. We recall that $Y_t = \Psi(M_t, D_t, U_{A_t})$, where $\Psi(z,r,u) = \gamma(z) + \sqrt{r}u$ for $(z,r,u) \in \mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U}$ and $\Psi(z,r,u) = \triangle$ for $(z,r,u) = \triangle$. One easily checks that Ψ is a bijection from $(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U}) \cup \{\triangle\}$ to $(\mathcal{X} \cap E_N) \cup \{\triangle\}$; recall that $\mathcal{X} = E_{k_0}$ and $\mathcal{U} = E_{k_0} \cap \mathbb{S}$.

We now study

$$\mathbb{Y} = (\Omega^Y, \mathcal{M}^Y, (Y_t)_{t \geq 0}, (\mathbb{P}_y^Y)_{y \in (\mathcal{X} \cap E_N) \cup \{\triangle\}}),$$

where $\Omega^Y = \Omega^M \times \Omega^D \times \Omega^U$, $\mathcal{M}^Y = \mathcal{M}^{(M,D,U)}$ and $\mathbb{P}_y^Y = \mathbb{P}_{(z,r,u)}^{(M,D,U)}$ for $(z,r,u) = \Psi^{-1}(y)$.

First, $\mathbb{Y}$ is an $(\mathcal{X} \cap E_N) \cup \{\triangle\}$-valued diffusion, because the bijection Ψ, which maps $(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U}) \cup \{\triangle\}$ onto $(\mathcal{X} \cap E_N) \cup \{\triangle\}$, is continuous, both sets being endowed with the one-point compactification topology; see Section B.1.

Next, we prove that $\mathbb{Y}$ is μ-symmetric: if φ, ψ are nonnegative measurable functions on $\mathcal{X} \cap E_N$ and $t \geq 0$, we have, thanks to Lemma A.2 (recall that $\nu = d_{\theta,N}(N)/2 - 1$),

$$\int_{(\mathbb{R}^2)^N} [P_t^Y \varphi(y)] \psi(y) \, \mu(\mathrm{d}y) = \frac{1}{2} \int_{\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{S}} [(P_t^Y \varphi)(\Psi(z,r,u))] \psi(\Psi(z,r,u)) r^\nu \, \mathrm{d}z \, \mathrm{d}r \, \beta(\mathrm{d}u).$$

But $(P_t^Y \varphi)(\Psi(z,r,u)) = \mathbb{E}_{(z,r,u)}[\varphi(\Psi(M_t, D_t, U_{A_t}))] = P_t^{(M,D,U_A)}(\varphi \circ \Psi)(z,r,u)$, so that

$$\int_{(\mathbb{R}^2)^N} [P_t^Y \varphi(y)] \psi(y) \, \mu(\mathrm{d}y) = \frac{1}{2} \int_{\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{S}} [P_t^{(M,D,U_A)}(\varphi \circ \Psi)(z,r,u)] [(\psi \circ \Psi)(z,r,u)] \, r^\nu \, \mathrm{d}z \, \mathrm{d}r \, \beta(\mathrm{d}u).$$

Using the fact that $(\mathbb{M}, \mathbb{D}, \mathbb{U}_{\mathbb{A}})$ is $\mathrm{d}z r^\nu \mathrm{d}r \beta(\mathrm{d}u)$-symmetric and then the same computation in reverse order, one concludes that $\int_{(\mathbb{R}^2)^N} [P_t^Y \varphi] \psi \, \mathrm{d}\mu = \int_{(\mathbb{R}^2)^N} \varphi [P_t^Y \psi] \, \mathrm{d}\mu$ as desired.

Thus $\mathbb{Y}$ has a Dirichlet space $(\mathcal{E}^Y, \mathcal{F}^Y)$ on $L^2((\mathbb{R}^2)^N, \mu)$ that we now determine. For $\varphi \in L^2((\mathbb{R}^2)^N, \mu)$, using as above Lemma A.2 and the fact that $(P_t^Y \varphi)(\Psi(z,r,u)) = P_t^{(M,D,U_A)}(\varphi \circ \Psi)(z,r,u)$, we find

$$\frac{1}{t} \int_{(\mathbb{R}^2)^N} (P_t^Y \varphi - \varphi) \varphi \, \mathrm{d}\mu = \frac{1}{2t} \int_{\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathbb{S}} [P_t^{(M,D,U_A)}(\varphi \circ \Psi)(z,r,u) - (\varphi \circ \Psi)(z,r,u)] [\varphi \circ \Psi(z,r,u)] r^\nu \, \mathrm{d}z \, \mathrm{d}r \, \beta(\mathrm{d}u).$$

Since Ψ is bijective, we deduce (see [11, Lemma 1.3.4, p. 23]) that

$$\mathcal{F}^Y = \{\varphi \in L^2((\mathbb{R}^2)^N, \mu) : \varphi \circ \Psi \in \mathcal{F}^{(M,D,U_A)}\} \tag{30}$$

and

$$\mathcal{E}^Y(\varphi,\varphi) = \tfrac{1}{2}\mathcal{E}^{(M,D,U_A)}(\varphi\circ\Psi,\varphi\circ\Psi) \quad \text{for } \varphi\in\mathcal{F}^Y. \tag{31}$$

Step 4. We now compute $\mathcal{E}^Y(\varphi,\varphi)$ for $\varphi\in C_c^\infty(\mathcal{X}\cap E_N)$. Observe that $\varphi\circ\Psi\in C_c^\infty(\mathbb{R}^2\times\mathbb{R}_+^*\times\mathcal{U})$. Thanks to (29) and (31), we have

$$\mathcal{E}^Y(\varphi,\varphi) = \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} I(z,r,u)\,\mathrm{d}z\,r^\nu\,\mathrm{d}r\,\beta(\mathrm{d}u), \tag{32}$$

where

$$\begin{aligned} I(z,r,u) = {} & \frac{1}{2N}\|\nabla_z(\varphi\circ\Psi)(z,r,u)\|^2 + 2r|\partial_r(\varphi\circ\Psi)(z,r,u)|^2 \\ & + \frac{1}{2r}\|\nabla_{\mathbb{S}}(\varphi\circ\Psi)(z,r,u)\|^2. \end{aligned}$$

We recall that for $\varphi:(\mathbb{R}^2)^N\to\mathbb{R}$, we call $\nabla\varphi(x) = ((\nabla\varphi(x))^1,\dots,(\nabla\varphi(x))^N)\in(\mathbb{R}^2)^N$ the *total gradient* of φ at $x\in(\mathbb{R}^2)^N$, and we have $(\nabla\varphi(x))^i\in\mathbb{R}^2$ for each $i\in[\![1,N]\!]$. And for $\phi: O\to\mathbb{R}^p$, where O is open in $\mathbb{R}^n$, we denote by $\mathrm{d}_z\phi$ the differential of ϕ at $z\in O$.

We start with the study of $\Psi(z,r,u) = \gamma(z)+\sqrt{r}u$, where we recall that γ was introduced in Section 2 and $\Phi_{\mathbb{S}}(x) = \pi_H x/\|\pi_H x\|$ is defined on a neighborhood of $\mathbb{S}$ in $(\mathbb{R}^2)^N$ (see (10)). For all $(z,r,u)\in\mathbb{R}^2\times\mathbb{R}_+^*\times\mathbb{S}$ and all $h\in\mathbb{R}^2$, $k\in\mathbb{R}$ and $\ell\in(\mathbb{R}^2)^N$, we have

$$\begin{aligned} & \mathrm{d}_z\Psi(\cdot,r,u)(h) = \gamma(h), \quad \mathrm{d}_r\Psi(z,\cdot,u)(k) = \frac{k}{2\sqrt{r}}u, \\ & \mathrm{d}_u[\Psi(z,r,\Phi_{\mathbb{S}}(\cdot))](\ell) = \sqrt{r}\pi_{u^\perp}(\pi_H(\ell)). \end{aligned}$$

For the first equality, it suffices to use the fact that γ is linear, so that $\mathrm{d}_z\Psi(\cdot,r,u)(h) = \mathrm{d}_z\gamma(h) = \gamma(h)$. The second equality is obvious. For the third equality, which is the differential at $u\in\mathbb{S}$ of the function $F(x) = \gamma(z)+\sqrt{r}\Phi_{\mathbb{S}}(x)$ defined for x in E_N (which is open in $(\mathbb{R}^2)^N$ and contains $\mathbb{S}$), we write $\mathrm{d}_u F = \sqrt{r}\mathrm{d}_u\Phi_{\mathbb{S}}$. But $\Phi_S = G\circ\pi_H$, where $G(x) = x/\|x\|$, and we have $\mathrm{d}_u\pi_H = \pi_H$ and $\mathrm{d}_{\pi_H(u)}G = \mathrm{d}_u G = \pi_{u^\perp}$ for $u\in\mathbb{S}$. All in all, $\mathrm{d}_u F = \sqrt{r}\pi_{u^\perp}\circ\pi_H$.

First, $\nabla_z(\varphi\circ\Psi)(z,r,u) = \sum_{i=1}^N[\nabla\varphi(\Psi(z,r,u))]^i$. Indeed, for all $h\in\mathbb{R}^2$,

$$\begin{aligned} \mathrm{d}_z(\varphi\circ\Psi(\cdot,r,u))(h) &= (\mathrm{d}_{\Psi(z,r,u)}\varphi)[(\mathrm{d}_z\Psi(\cdot,r,u))(h)] \\ &= (\mathrm{d}_{\Psi(z,r,u)}\varphi)(\gamma(h)) = \nabla\varphi(\Psi(z,r,u))\cdot\gamma(h), \end{aligned}$$

which, by definition of γ, equals $h\cdot\sum_{i=1}^N[\nabla\varphi(\Psi(z,r,u))]^i$.

This implies that

$$\frac{1}{2N}\|\nabla_z(\varphi\circ\Psi(z,r,u))\|^2 = \frac{1}{2N}\left\|\sum_{i=1}^N[\nabla\varphi(\Psi(z,r,u))]^i\right\|^2 = \frac{1}{2}\|\pi_{H^\perp}(\nabla\varphi(\Psi(z,r,u)))\|^2. \tag{33}$$

Indeed, recalling the expression of π_H (see Section 2), it suffices to note that for every $x \in (\mathbb{R}^2)^N$, $\|\pi_{H^\perp}(x)\|^2 = \|\gamma(S_{[\![1,N]\!]}(x))\|^2 = N\|S_{[\![1,N]\!]}(x)\|^2 = N^{-1}\|\sum_{i=1}^N x^i\|^2$.

Next, $\partial_r(\varphi \circ \Psi)(z,r,u) = (\nabla\varphi)(\Psi(z,r,u)) \cdot u/(2\sqrt{r})$. Indeed, for $k \in \mathbb{R}$,

$$\mathrm{d}_r(\varphi \circ \Psi(z,\cdot,u))(k) = (\mathrm{d}_{\Psi(z,r,u)}\varphi)[(\mathrm{d}_r\Psi(z,\cdot,u))(k)] = (\mathrm{d}_{\Psi(z,r,u)}\varphi)(u) \times \frac{k}{2\sqrt{r}},$$

which is nothing but $(\nabla\varphi)(\Psi(z,r,u)) \cdot u \times k/(2\sqrt{r})$.

This implies, recalling that π_u is the orthogonal projection on $\mathrm{span}(u) \subset (\mathbb{R}^2)^N$, that

$$2r|\partial_r(\varphi \circ \Psi)(z,r,u)|^2 = \tfrac{1}{2}\|\pi_u((\nabla\varphi)(\Psi(z,r,u)))\|^2 = \tfrac{1}{2}\|\pi_H(\pi_u((\nabla\varphi)(\Psi(z,r,u))))\|^2 \tag{34}$$

since $u \in \mathbb{S}$, so that $\|u\| = 1$ and $u \in H$.

Finally, $\nabla_{\mathbb{S}}(\varphi \circ \Psi)(z,r,u) = \sqrt{r}\pi_H(\pi_{u^\perp}(\nabla\varphi(\Psi(z,r,u))))$. Indeed, for all ℓ in $(\mathbb{R}^2)^N$,

$$\begin{aligned} d_u((\varphi \circ \Psi)(z,r,\Phi_{\mathbb{S}}(\cdot)))(\ell) &= (d_{\Psi(z,r,u)}\varphi)(d_u[\Psi(z,r,\Phi_{\mathbb{S}}(\cdot))](\ell)) \\ &= \sqrt{r}(d_{\Psi(z,r,u)}\varphi)(\pi_{u^\perp}(\pi_H(\ell))) \\ &= \sqrt{r}\nabla\varphi(\Psi(z,r,u)) \cdot \pi_{u^\perp}(\pi_H(\ell)) \\ &= \sqrt{r}\pi_H(\pi_{u^\perp}(\nabla\varphi(\Psi(z,r,u)))) \cdot \ell \end{aligned}$$

as desired, since $\nabla_{\mathbb{S}}(\varphi \circ \Psi)(z,r,u) = \nabla_x((\varphi \circ \Psi)(z,r,\Phi_{\mathbb{S}}(\cdot)))(u)$ by definition of $\nabla_{\mathbb{S}}$ (see (12)).

This implies that

$$\frac{1}{2r}\|\nabla_{\mathbb{S}}(\varphi \circ \Psi)(z,r,u)\|^2 = \frac{1}{2}\|\pi_H(\pi_{u^\perp}(\nabla\varphi(\Psi(z,r,u))))\|^2. \tag{35}$$

Gathering (33)–(35), we see that $I(z,r,u) = \frac{1}{2}\|\nabla\varphi(\Psi(z,r,u))\|^2$, since for $x \in (\mathbb{R}^2)^N$,

$$\|\pi_{H^\perp}(x)\|^2 + \|\pi_H(\pi_u(x))\|^2 + \|\pi_H(\pi_{u^\perp}(x))\|^2 = \|x\|^2$$

because $u \in \mathbb{S} \subset H$.

Injecting the value of I in (32) and using Lemma A.2, we obtain

$$\mathcal{E}^Y(\varphi,\varphi) = \frac{1}{4}\int_{\mathbb{R}^2\times\mathbb{R}_+^*\times\mathbb{S}} \|\nabla\varphi(\Psi(z,r,u))\|^2\,\mathrm{d}z\,r^\nu\,\mathrm{d}r\,\beta(\mathrm{d}u) = \frac{1}{2}\int_{(\mathbb{R}^2)^N}\|\nabla\varphi\|^2\,\mathrm{d}\mu.$$

Step 5. As a last technical step, we verify that $(\mathcal{E}^Y, \mathcal{F}^Y)$ is a regular Dirichlet space on $L^2((\mathbb{R}^2)^N,\mu)$ with core $C_c^\infty(\mathcal{X} \cap E_N)$, i.e. for all $\varphi \in \mathcal{F}^Y$, there is $\varphi_n \in C_c^\infty(\mathcal{X} \cap E_N)$ such that $\lim_n \|\varphi_n - \varphi\|_{L^2((\mathbb{R}^2)^N,\mu)} + \mathcal{E}^Y(\varphi_n - \varphi, \varphi_n - \varphi) = 0$.

Since $(\mathcal{E}^{(M,D,U_A)}, \mathcal{F}^{(M,D,U_A)})$ on $L^2(\mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{S}, \mathrm{d}zr^\nu\mathrm{d}r\beta(\mathrm{d}u))$ is regular with core $C_c^\infty(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U})$, and recalling (30), there is $g_n \in C_c^\infty(\mathbb{R}^2 \times \mathbb{R}_+^* \times \mathcal{U})$ such that

$$\|g_n - \varphi \circ \Psi\|_{L^2(\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S},\mathrm{d}zr^\nu\mathrm{d}r\beta(\mathrm{d}u))} + \mathcal{E}^{(M,D,U_A)}(g_n - \varphi \circ \Psi, g_n - \varphi \circ \Psi) \to 0.$$

Setting $\varphi_n = g_n \circ \Psi^{-1}$, we have $\varphi_n \in C_c^\infty(\mathcal{X} \cap E_N)$, and by (31),

$$\mathcal{E}^Y(\varphi_n - \varphi, \varphi_n - \varphi) = \tfrac{1}{2}\mathcal{E}^{(M,D,U_A)}(g_n - \varphi \circ \Psi, g_n - \varphi \circ \Psi) \to 0,$$

as well as, by Lemma A.2,

$$\|\varphi_n - \varphi\|_{L^2((\mathbb{R}^2)^N,\mu)} = \tfrac{1}{2}\|g_n - \varphi \circ \Psi\|_{L^2(\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S},\mathrm{d}z r^\nu \mathrm{d}r\beta(\mathrm{d}u))} \to 0.$$

Step 6. By Steps 3–5, $\mathbb{Y}$ is a μ-symmetric $(\mathcal{X} \cap E_N) \cup \{\triangle\}$-valued diffusion with regular Dirichlet space $(\mathcal{E}^Y, \mathcal{F}^Y)$ with core $C_c^\infty(\mathcal{X} \cap E_N)$ and with $\mathcal{E}^Y(\varphi, \varphi) = \frac{1}{2}\int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \,\mathrm{d}\mu$ for $\varphi \in C_c^\infty(\mathcal{X} \cap E_N)$.

Now, applying Lemma B.6 to $\mathbb{X}$ defined in Proposition 6 (i) with the set $\mathcal{X} \cap E_N$, we see that $\mathbb{X}^*$, i.e. $\mathbb{X}$ killed when getting outside $\mathcal{X} \cap E_N$, is a μ-symmetric $(\mathcal{X} \cap E_N) \cup \{\triangle\}$-valued diffusion process with regular Dirichlet space $(\mathcal{E}^{X^*}, \mathcal{F}^{X^*})$ with core $C_c^\infty(\mathcal{X} \cap E_N)$ and with $\mathcal{E}^{X^*}(\varphi, \varphi) = \frac{1}{2}\int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \,\mathrm{d}\mu$ for $\varphi \in C_c^\infty(\mathcal{X} \cap E_N)$.

This implies, as recalled in Section B.1, that $(\mathcal{E}^{X^*}, \mathcal{F}^{X^*}) = (\mathcal{E}^Y, \mathcal{F}^Y)$. The conclusion follows by uniqueness (see [11, Theorem 4.2.8, p. 167]). ∎

Actually, $(R_{[\![1,N]\!]}(X_t))_{t\geq 0}$ and $(S_{[\![1,N]\!]}(X_t))_{t\geq 0}$ are a squared Bessel process and a Brownian motion independent *until explosion* (and not only until the first time that $R_{[\![1,N]\!]}(X_t) = 0$, as shown in Proposition 10), a fact that we shall often use.

Lemma 11. *Fix $N \geq 2$ and $\theta > 0$ such that $N > \theta$ and consider a QKS(θ, N)-process $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}_\triangle})$. Quasi-everywhere, there are a 2-dimensional Brownian motion $(M_t)_{t\geq 0}$ with diffusion constant $N^{-1/2}$ issued from $S_{[\![1,N]\!]}(x)$ and a squared Bessel process $(D_t)_{t\geq 0}$ of dimension $d_{\theta,N}(N)$ issued from $R_{[\![1,N]\!]}(x)$ (killed when it leaves $(0,\infty)$ if $d_{\theta,N}(N) \leq 0$) independent of $(M_t)_{t\geq 0}$ such that $\mathbb{P}_x^X$-a.s., $S_{[\![1,N]\!]}(X_t) = M_t$ and $R_{[\![1,N]\!]}(X_t) = D_t$ for all $t \in [0, \zeta)$.*

Proof. If $\theta \geq 2$, this is a consequence of Proposition 10: introducing the stopping time $\tau = \inf\{t > 0 : R_{[\![1,N]\!]}(X_t) \notin (0,\infty)\}$, we have $\tau = \zeta$. Indeed, on $\{\tau < \zeta\}$, we have $X_\tau \notin E_N$, whence $X_\tau \notin \mathcal{X}$ since $\mathcal{X} = E_{k_0}$ with $k_0 \leq N$ (because $\theta \geq 2$), which contradicts $\tau < \zeta$.

We now suppose that $\theta < 2$, so that $k_0 > N$ and thus $\mathcal{X} = (\mathbb{R}^2)^N$. We introduce the shortened notation $R(x) = R_{[\![1,N]\!]}(x)$, $S(x) = S_{[\![1,N]\!]}(x)$ and split the proof into three parts.

Step 1. First, one can show similarly to the proof of Proposition 10 (but much more easily) that there exists a 2-dimensional Brownian motion $(M_t)_{t\geq 0}$, which is independent of $(X_t - \gamma(S(X_t)))_{t\geq 0}$, such that $S(X_t) = M_t$ for all $t \in [0, \zeta)$. This moreover shows that $(M_t)_{t\geq 0}$ is independent of $(R(X_t))_{t\geq 0}$, because $R(X_t) = \|X_t - \gamma(S(X_t))\|^2$.

Step 2. We consider some function $g_m \in C_c^\infty((\mathbb{R}^2)^N)$ such that $g_m = 1$ on $B(0, m)$ and $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}_\alpha^X g_m(x)| < \infty$. Such a function exists by Remark 14. For $\varphi \in C_c^\infty(\mathbb{R}_+)$, we set $\psi(x) = \varphi(R(x))$ and show that $\psi g_m \in \mathcal{D}_{\mathcal{A}^X}$ and for all $x \in B(0, m)$,

$$\mathcal{A}^X(\psi g_m)(x) = 2R(x)\varphi''(R(x)) + d_{\theta,N}(N)\varphi'(R(x)). \tag{36}$$

To this end, we apply Remark 8. Since $\psi g_m \in C_c^\infty((\mathbb{R}^2)^N)$ and $\mathcal{X} = (\mathbb{R}^2)^N$, we have to show that $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}_\alpha^X(\psi g_m)(x)| < \infty$, and we will deduce that $\mathcal{A}^X(\psi g_m) = \mathcal{L}^X(\psi g_m)$. By (22), we have $\mathcal{L}_\alpha^X(\psi g_m) = g_m \mathcal{L}_\alpha^X \psi + \psi \mathcal{L}_\alpha^X g_m + \nabla\psi \cdot \nabla g_m$. The only difficulty consists in showing $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}_\alpha^X \psi(x)| < \infty$. Using $\nabla_{x^i} R(x) = 2(x^i - S(x))$, we find $\nabla_{x^i}\psi(x) = 2(x^i - S(x))\varphi'(R(x))$. Hence, by symmetry,

$$\frac{\theta}{N} \sum_{1\le i\ne j\le N} \frac{x^i - x^j}{\|x^i - x^j\|^2 + \alpha} \cdot \nabla_{x^i}\psi(x) = \frac{2\theta}{N}\varphi'(R(x)) \sum_{1\le i\ne j\le N} \frac{x^i - x^j}{\|x^i - x^j\|^2 + \alpha} \cdot x^i$$
$$= \frac{\theta}{N}\varphi'(R(x)) \sum_{1\le i\ne j\le N} \frac{\|x^i - x^j\|^2}{\|x^i - x^j\|^2 + \alpha}. \quad (37)$$

Moreover, $\Delta_{x^i}\psi(x) = 4(1 - 1/N)\varphi'(R(x)) + 4\|x^i - S(x)\|^2\varphi''(R(x))$, whence

$$\Delta\psi(x) = 4(N-1)\varphi'(R(x)) + 4R(x)\varphi''(R(x)). \quad (38)$$

By combining (37) and (38) we conclude that

$$\mathcal{L}_\alpha^X\psi(x) = 2R(x)\varphi''(R(x)) + \Big(2(N-1) - \frac{\theta}{N} \sum_{1\le i\ne j\le N} \frac{\|x^i - x^j\|^2}{\|x^i - x^j\|^2 + \alpha}\Big)\varphi'(R(x)).$$

Since φ is compactly supported, we deduce that $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}_\alpha^X\psi(x)| < \infty$, whence $\sup_{\alpha\in(0,1]} \sup_{x\in(\mathbb{R}^2)^N} |\mathcal{L}_\alpha^X(\psi g_m)(x)| < \infty$. Hence $\psi g_m \in \mathcal{D}_{\mathcal{A}^X}$ and $\mathcal{A}^X(\psi g_m) = \mathcal{L}^X(\psi g_m)$. Moreover, recalling that $\mathcal{L}^X\psi = \mathcal{L}_\alpha^X\psi$ with $\alpha = 0$, and $g_m = 1$ on $B(0,m)$, we conclude that $\mathcal{A}^X(\psi g_m)(x) = \mathcal{L}_0^X\psi(x)$ for $x \in B(0,m)$. This implies (36), because $2(N-1) - \theta(N-1) = d_{\theta,N}(N)$.

Step 3. We define $\zeta_m = \inf\{t > 0 : X_t \notin B(0,m)\}$. By Lemma B.2 and Step 1, for all $\varphi \in C_c^\infty(\mathbb{R}_+)$, quasi-everywhere in $B(0,m)$,

$$\varphi(R(X_{t\wedge\zeta_m})) - \varphi(R(x)) - \int_0^{t\wedge\zeta_m} \mathcal{L}^X\varphi(R(X_s))\,\mathrm{d}s$$

is a $\mathbb{P}_x^X$-martingale. Recalling (36), we classically conclude that there is a Brownian motion W such that $R(X_t) = R(x) + 2\int_0^t \sqrt{R(X_s)}\mathrm{d}W_s + d_{\theta,N}(N)t$ during $[0,\zeta_n]$. We recognize the SDE of a squared Bessel process of dimension $d_{\theta,N}(N)$ (see Revuz–Yor [21, Chapter XI]). Since we know from Remark 7 that $\zeta = \lim_m \zeta_m$, the proof is complete. ■

6. Some cutoff functions

We will need several times to approximate indicator functions by smooth functions, on which the generator $\mathcal{L}^X$ (or $\mathcal{L}^U$) is bounded. This does not seem obvious, due to the singularity of $\mathcal{L}^X$. We recall that $\mathcal{L}_\alpha^X$ and $\mathcal{L}_\alpha^U$ were defined in (21) and (23).

Lemma 12. *Fix $N \geq 2$ and $\theta > 0$, and recall that $k_0 = \lceil 2N/\theta \rceil$ and $\mathcal{X} = E_{k_0}$. Consider a partition $\mathbf{K} = (K_p)_{p \in [\![1,\ell]\!]}$ of $[\![1, N]\!]$ and define, for $\varepsilon \in [0, 1]$ (with the convention that $B(0, 1/0) = (\mathbb{R}^2)^N$),*

$$G_{\mathbf{K},\varepsilon} = \Big\{x \in \mathcal{X} : \min_{1 \leq p \neq q \leq \ell} \min_{i \in K_p,\, j \in K_q} \|x^i - x^j\|^2 > \varepsilon\Big\} \cap B(0, 1/\varepsilon).$$

(i) *For all $\varepsilon \in (0, 1]$, there is a family of open relatively compact subsets $G^n_{\mathbf{K},\varepsilon}$ of $G_{\mathbf{K},0}$ such that*

$$\bigcup_{n \geq 1} G^n_{\mathbf{K},\varepsilon} \supset G_{\mathbf{K},\varepsilon} \quad \textit{and} \quad G^n_{\mathbf{K},\varepsilon} \subset G^{n+1}_{\mathbf{K},\varepsilon} \quad \textit{for each } n \geq 1$$

and a family of $[0, 1]$-valued functions $\Gamma^n_{\mathbf{K},\varepsilon} \in C^\infty_c(G_{\mathbf{K},0})$ such that for some $\eta \in (0, 1]$ and all $n \geq 1$,

$$\operatorname{supp} \Gamma^n_{\mathbf{K},\varepsilon} \subset G_{\mathbf{K},\eta}, \quad \Gamma^n_{\mathbf{K},\varepsilon} = 1 \textit{ on } G^n_{\mathbf{K},\varepsilon}, \quad \sup_{\alpha \in (0,1]} \sup_{x \in (\mathbb{R}^2)^N} |\mathcal{L}^X_\alpha \Gamma^n_{\mathbf{K},\varepsilon}(x)| < \infty.$$

(ii) *With the same sets $G^n_{\mathbf{K},\varepsilon}$ as in* (i)*, there exist functions $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon} \in C^\infty_c(\mathbb{S} \cap G_{\mathbf{K},0})$ with values in $[0, 1]$ such that for all $n \geq 1$,*

$$\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon} = 1 \quad \textit{on } \mathbb{S} \cap G^n_{\mathbf{K},\varepsilon} \quad \textit{and} \quad \sup_{\alpha \in (0,1]} \sup_{u \in \mathbb{S}} |\mathcal{L}^U_\alpha \Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)| < \infty.$$

The section is devoted to the proof of this lemma. We start with the following technical result.

Lemma 13. *Define a family $(c_\ell)_{\ell \in [\![1,N]\!]}$ by $c_0 = 1$ and $c_{\ell+1} = (2 + 4\ell)c_\ell$ for all $\ell \in [\![1, N - 1]\!]$. For all $K \subsetneq [\![1, N]\!]$, and all $\varepsilon \in (0, 1]$ and $x \in (\mathbb{R}^2)^N$ such that*

$$R_K(x) \leq 2c_{|K|}\varepsilon \quad \textit{and} \quad \min_{j \notin K} R_{K \cup \{j\}}(x) \geq c_{|K|+1}\varepsilon,$$

we have $\|x^i - x^j\|^2 \geq c_{|K|}\varepsilon$ for all $i \in K$ and $j \notin K$.

Proof. We fix $K \subsetneq [\![1, N]\!]$, $\varepsilon \in (0, 1]$ and $x \in (\mathbb{R}^2)^N$ as in the statement and assume for contradiction that there are $i_0 \in K$ and $j_0 \notin K$ such that $\|x^{i_0} - x^{j_0}\|^2 < c_{|K|}\varepsilon$. Then for all $i \in K$,

$$\begin{aligned}\|x^{j_0} - x^i\|^2 &\leq 2\|x^{i_0} - x^{j_0}\|^2 + 2\|x^{i_0} - x^i\|^2 \leq 2\|x^{i_0} - x^{j_0}\|^2 + 2|K|R_K(x) \\ &< (2 + 4|K|)c_{|K|}\varepsilon.\end{aligned}$$

This implies that

$$\begin{aligned}R_{K \cup \{j_0\}}(x) &= \frac{1}{2(|K| + 1)}\Big(2|K|R_K(x) + 2\sum_{i \in K} \|x^{j_0} - x^i\|^2\Big) \\ &\leq R_K(x) + \frac{1}{|K| + 1} \sum_{i \in K} \|x^{j_0} - x^i\|^2,\end{aligned}$$

whence

$$R_{K\cup\{j_0\}}(x) < 2c_{|K|}\varepsilon + \frac{2+4|K|}{|K|+1}|K|c_{|K|}\varepsilon < (2+4|K|)c_{|K|}\varepsilon = c_{|K|+1}\varepsilon,$$

which is a contradiction. ∎

Proof of Lemma 12. We introduce a nondecreasing C^∞ function $\varrho : \mathbb{R}_+ \to [0,1]$ such that $\varrho = 0$ on $[0,1/2]$ and $\varrho = 1$ on $[1,\infty)$. We divide the proof into three steps.

Step 1. We fix $n \geq 1$ and define, for $K \subset [\![1,N]\!]$, using the family $(c_\ell)_{\ell\in[\![1,N]\!]}$ of Lemma 13,

$$\tilde{E}_{K,n} = \{x \in (\mathbb{R}^2)^N : \forall L \supset K,\ R_L(x) > c_{|L|}/n\}, \quad \tilde{\Gamma}_{K,n}(x) = \prod_{L\supset K} \varrho(nR_L(x)/c_{|L|}),$$

where $\{L \supset K\} = \{L \subset [\![1,N]\!] : K \subset L\}$. We have

$$\tilde{\Gamma}_{K,n} \in C^\infty((\mathbb{R}^2)^N), \quad \operatorname{supp}\tilde{\Gamma}_{K,n} \subset \tilde{E}_{K,2n}, \quad \tilde{\Gamma}_{K,n} = 1 \quad \text{on } \tilde{E}_{K,n}. \tag{39}$$

Since $R_K(x) > 0$ implies that $R_L(x) > 0$ for all $L \supset K$, we also have

$$\bigcup_{n\geq 1} \tilde{E}_{K,n} = \tilde{E}_K, \quad \text{where} \quad \tilde{E}_K = \{x \in (\mathbb{R}^2)^N : R_K(x) > 0\}. \tag{40}$$

We now show, and this is the main difficulty of this step, that for all $A > 0$ and $K \subset [\![1,N]\!]$ with $|K| \geq 2$, we have $\sup_{\alpha\in(0,1]} \sup_{x\in B(0,A)} |\mathcal{L}^X_\alpha \tilde{\Gamma}_{K,n}(x)| < \infty$. Observe that since $\sup_{x\in B(0,A)} |\Delta\tilde{\Gamma}_{K,n}(x)| < \infty$, it suffices that $\sup_{\alpha\in(0,1]} \sup_{x\in B(0,A)} |I_{K,n,\alpha}(x)| < \infty$, where

$$\begin{aligned} I_{K,n,\alpha}(x) &= \sum_{1\leq i\neq j\leq N} \frac{x^i - x^j}{\|x^i - x^j\|^2} \cdot \nabla_{x^i}\tilde{\Gamma}_{K,n}(x) \\ &= \sum_{L\supset K} f_{K,L,n}(x) \sum_{1\leq i\neq j\leq N} \frac{x^i - x^j}{\|x^i - x^j\|^2} \cdot \nabla_{x^i} R_L(x), \end{aligned}$$

with

$$f_{K,L,n}(x) = \frac{n}{c_{|L|}}\varrho'\Big(\frac{nR_L(x)}{c_{|L|}}\Big) \prod_{M\supset K,\, M\neq L} \varrho\Big(\frac{nR_M(x)}{c_{|M|}}\Big).$$

Using $\nabla_{x^i} R_L(x) = 2(x^i - S_L(x))\mathbf{1}_{\{i\in L\}}$, we now write

$$I_{K,n,\alpha}(x) = 2\sum_{L\supset K} f_{K,L,n}(x)(A_{L,\alpha}(x) + B_{L,\alpha}(x)),$$

where

$$\begin{aligned} A_{L,\alpha}(x) &= \sum_{i,j\in L,\, i\neq j} \frac{(x^i - x^j)\cdot(x^i - S_L(x))}{\|x^i - x^j\|^2 + \alpha}, \\ B_{L,\alpha}(x) &= \sum_{i\in L,\, j\in L^c} \frac{(x^i - x^j)\cdot(x^i - S_L(x))}{\|x^i - x^j\|^2 + \alpha}. \end{aligned}$$

We have $\sup_{\alpha\in(0,1]}\sup_{x\in B(0,A)}|f_{K,L,n}(x)A_{L,\alpha}(x)|<\infty$ because $f_{K,L,n}$ is bounded and

$$A_{L,\alpha}(x)=\sum_{i,j\in L,i\neq j}\frac{(x^i-x^j)\cdot x^i}{\|x^i-x^j\|^2+\alpha}$$
$$=\frac{1}{2}\sum_{i,j\in L,i\neq j}\frac{\|x^i-x^j\|^2}{\|x^i-x^j\|^2+\alpha}\in\Big[0,\frac{|L|(|L|-1)}{2}\Big].$$

Next, assume that $L\subsetneq[\![1,N]\!]$ (else $B_{L,\alpha}(x)=0$) and observe that $f_{K,L,n}(x)\neq 0$ implies that $R_L(x)<c_{|L|}/n$ (since $\varrho'=0$ on $[1,\infty)$) and that $\min_{i\notin L}R_{L\cup\{i\}}(x)>c_{|L|+1}/(2n)$ (as $\varrho=0$ on $[0,1/2]$). By Lemma 13, this gives $\min_{i\in L,j\in L^c}\|x^i-x^j\|^2\geq c_{|L|}/(2n)$. We immediately conclude that $\sup_{\alpha\in(0,1]}\sup_{x\in B(0,A)}|f_{K,L,n}(x)B_{L,\alpha}(x)|<\infty$.

Step 2. We can now prove (i). We fix $\varepsilon\in(0,1]$ and a partition $\mathbf{K}=(K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1,N]\!]$. For some $m\geq 1$ to be chosen later (as a function of ε), for each $n\geq 1$ we set

$$G^n_{\mathbf{K},\varepsilon}=B(0,m)\cap\Big(\bigcap_{K\subset[\![1,N]\!]:|K|=k_0}\tilde{E}_{K,n}\Big)\cap\Big(\bigcap_{1\leq p\neq q\leq\ell}\bigcap_{i\in K_p,j\in K_q}\tilde{E}_{\{i,j\},m}\Big),$$
$$\Gamma^n_{\mathbf{K},\varepsilon}(x)=g_m(x)\Big(\prod_{K\subset[\![1,N]\!]:|K|=k_0}\tilde{\Gamma}_{K,n}(x)\Big)\Big(\prod_{1\leq p\neq q\leq\ell}\prod_{i\in K_p,j\in K_q}\tilde{\Gamma}_{\{i,j\},m}(x)\Big),$$

where $g_m(x)=\varrho(m/\|x\|)$ with the extension $g_m(0)=1$.

First, $G^n_{\mathbf{K},\varepsilon}$ is clearly included in $G^{n+1}_{\mathbf{K},\varepsilon}$ and relatively compact in $G_{\mathbf{K},0}$. We deduce from (40) that, setting $H_{\mathbf{K},m}=B(0,m)\cap\bigcap_{1\leq p\neq q\leq\ell}\bigcap_{i\in K_p,j\in K_q}\tilde{E}_{\{i,j\},m}$,

$$\bigcup_{n\geq1}G^n_{\mathbf{K},\varepsilon}=\Big(\bigcap_{K\subset[\![1,N]\!]:|K|=k_0}\tilde{E}_K\Big)\cap H_{\mathbf{K},m}=E_{k_0}\cap H_{\mathbf{K},m}=\mathcal{X}\cap H_{\mathbf{K},m}.$$

By (40) again, we can choose m large enough that $H_{\mathbf{K},m}$ contains $G_{\mathbf{K},\varepsilon}$. Next, by (39), $\Gamma^n_{\mathbf{K},\varepsilon}\in C^\infty((\mathbb{R}^2)^N)$, $\Gamma^n_{\mathbf{K},\varepsilon}=1$ on $G^n_{\mathbf{K},\varepsilon}$ and

$$\operatorname{supp}\Gamma^n_{\mathbf{K},\varepsilon}\subset B(0,2m)\cap\Big(\bigcap_{K\subset[\![1,N]\!]:|K|=k_0}\tilde{E}_{K,2n}\Big)\cap\Big(\bigcap_{1\leq p\neq q\leq\ell}\bigcap_{i\in K_p,j\in K_q}\tilde{E}_{\{i,j\},2m}\Big),$$

which is compact in $G_{\mathbf{K},0}$. Moreover, $\operatorname{supp}\Gamma^n_{\mathbf{K},\varepsilon}\subset H_{\mathbf{K},2m}$. Since there exists $\eta\in(0,1]$ such that $H_{\mathbf{K},2m}\subset G_{\mathbf{K},\eta}$, we conclude that $\operatorname{supp}\Gamma^n_{\mathbf{K},\varepsilon}\subset G_{\mathbf{K},\eta}$.

It remains to show that $\sup_{\alpha\in(0,1]}\sup_{x\in(\mathbb{R}^2)^N}|\mathcal{L}^X_\alpha\Gamma^n_{\mathbf{K},\varepsilon}(x)|<\infty$. Introducing

$$\chi^n_{\mathbf{K},\varepsilon}(x)=\Big(\prod_{K\subset[\![1,N]\!]:|K|=k_0}\tilde{\Gamma}_{K,n}(x)\Big)\Big(\prod_{1\leq p\neq q\leq\ell}\prod_{i\in K_p,j\in K_q}\tilde{\Gamma}_{\{i,j\},m}(x)\Big),$$

which belongs to $C^\infty((\mathbb{R}^2)^N)$ by Step 1, we have $\Gamma^n_{\mathbf{K},\varepsilon}=g_m\chi^n_{\mathbf{K},\varepsilon}(x)$ (with the chosen value of m), and thus by (22),

$$\mathcal{L}^X_\alpha\Gamma^n_{\mathbf{K},\varepsilon}(x)=g_m(x)\mathcal{L}^X_\alpha\chi^n_{\mathbf{K},\varepsilon}(x)+\chi^n_{\mathbf{K},\varepsilon}\mathcal{L}^X_\alpha g_m(x)+\nabla g_m(x)\cdot\nabla\chi^n_{\mathbf{K},\varepsilon}(x).$$

The first term is uniformly bounded because g_m is bounded and supported in $B(0,2m)$ and because $\sup_{\alpha\in(0,1]}\sup_{x\in B(0,2m)}|\mathcal{L}^X\chi^n_{\mathbf{K},\varepsilon}(x)|<\infty$ by Step 1 and (22). The third term is also uniformly bounded, since $\chi^n_{\mathbf{K},\varepsilon}\in C^\infty((\mathbb{R}^2)^N)$ and ∇g_m is bounded and supported in $B(0,2m)$. Finally, the middle term is bounded because $\chi^n_{\mathbf{K},\varepsilon}$ is bounded by 1 and $\mathcal{L}^X_\alpha g_m$ is uniformly bounded, as we now show: Δg_m is obviously bounded since $g_m\in C^\infty_c((\mathbb{R}^2)^N)$, and since $\nabla_{x^i}g_m(x)=-m\varrho'(m/\|x\|)x^i/\|x\|^3$, we have

$$\sum_{1\le i,j\le N}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\cdot\nabla_{x^i}g_m(x)=-\frac{m\varrho'(m/\|x\|)}{\|x\|^3}\sum_{1\le i,j\le N}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\cdot x^i$$
$$=-\frac{m\varrho'(m/\|x\|)}{2\|x\|^3}\sum_{1\le i,j\le N}\frac{\|x^i-x^j\|^2}{\|x^i-x^j\|^2+\alpha}.$$

This last quantity is uniformly bounded, since ϱ' is bounded and vanishes on $[1,\infty)$.

Step 3. We now prove (ii), by showing that the restriction $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}=\Gamma^n_{\mathbf{K},\varepsilon}|_{\mathbb{S}}$ satisfies the required conditions. Obviously $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}\in C^\infty_c(\mathbb{S}\cap G_{\mathbf{K},0})$ and $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}=1$ on $\mathbb{S}\cap G^n_{\mathbf{K},\varepsilon}$. It remains to show that $\sup_{\alpha\in(0,1]}\sup_{u\in\mathbb{S}}|\mathcal{L}^U_\alpha\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}|<\infty$ (recall (23)). Since $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}\in C^\infty(\mathbb{S})$, $\Delta_{\mathbb{S}}\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}$ is bounded. We thus only have to verify that $\sup_{\alpha\in(0,1]}\sup_{u\in\mathbb{S}}|T_\alpha(u)|<\infty$, where

$$T_\alpha(u)=-\frac{\theta}{N}\sum_{1\le i,j\le N}\frac{u^i-u^j}{\|u^i-u^j\|^2+\alpha}\cdot(\nabla_{\mathbb{S}}\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u))^i.$$

Setting $b^i_\alpha(u)=-\frac{\theta}{N}\sum_{j=1}^N\frac{u^i-u^j}{\|u^i-u^j\|^2+\alpha}$ and using (14), we obtain

$$T_\alpha(u)=b_\alpha(u)\cdot\nabla_{\mathbb{S}}\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)=b_\alpha(u)\cdot\pi_H(\pi_{u^\perp}(\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u))).$$

Since now $b(u)\in H$ and since π_H and $\pi_{u^\perp}$ are self-adjoint, being orthogonal projections, we get

$$T_\alpha(u)=\pi_{u^\perp}(b_\alpha(u))\cdot\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)=b_\alpha(u)\cdot\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)-(b_\alpha(u)\cdot u)(u\cdot\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)).$$

But $b_\alpha(u)\cdot\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)=\mathcal{L}^X_\alpha\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)-\frac{1}{2}\Delta\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)$ is uniformly bounded by (i) and since $\Delta\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)$ is bounded on $\mathbb{S}$. Next, $u\cdot\nabla\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)$ is smooth and thus bounded on $\mathbb{S}$. Finally,

$$b_\alpha(u)\cdot u=-\frac{\theta}{N}\sum_{1\le i,j\le N}\frac{(u^i-u^j)\cdot u^i}{\|u^i-u^j\|^2+\alpha}=-\frac{\theta}{2N}\sum_{1\le i,j\le N}\frac{\|u^i-u^j\|^2}{\|u^i-u^j\|^2+\alpha}$$

is also uniformly bounded. ■

Remark 14. We have proved in Step 2 that for each $m>0$, $g_m\in C^\infty_c((\mathbb{R}^2)^N)$ satisfies $g_m=1$ on $B(0,m)$ and $\sup_{\alpha\in(0,1]}\sup_{x\in(\mathbb{R}^2)^N}|\mathcal{L}^X_\alpha g_m(x)|<\infty$.

7. A Girsanov theorem for the Keller–Segel particle system

In this section, we prove a rigorous version of the intuitive argument presented in Section 3.4.

For $x \in (\mathbb{R}^2)^N$ and $K \subset [\![1, N]\!]$, we set $x|_K = (x^i)_{i \in K}$. For $\mathbf{K} = (K_p)_{p \in [\![1,\ell]\!]}$ a partition of $[\![1, N]\!]$, for $y_1 \in (\mathbb{R}^2)^{|K_1|}, \dots, y_\ell \in (\mathbb{R}^2)^{|K_\ell|}$, we abusively denote by $(y_p)_{p \in [\![1,\ell]\!]}$ the element y of $(\mathbb{R}^2)^N$ such that $y|_{K_i} = y_i$ for all $i \in [\![1, \ell]\!]$.

We adopt the convention that for any $\theta > 0$, a *QKS*$(\theta, 1)$-process is a 2-dimensional Brownian motion. This is natural in view of (1).

Proposition 15. *Let $N \geq 2$ and $\theta > 0$ be such that $N > \theta$ and set $k_0 = \lceil 2N/\theta \rceil$. Fix some partition $\mathbf{K} = (K_p)_{p \in [\![1,\ell]\!]}$ of $[\![1, N]\!]$ with $\ell \geq 2$. Consider the state spaces $\mathcal{X} = E_{k_0}$ and, for each $p \in [\![1, \ell]\!]$,*

$$\mathcal{Y}_p = \Big\{ y \in (\mathbb{R}^2)^{|K_p|} : \forall K \subset [\![1, |K_p|]\!] \text{ with } |K| \geq k_0, \sum_{i,j=1}^{|K_p|} \|y^i - y^j\|^2 > 0 \Big\}.$$

Consider

- *a QKS(θ, N)-process $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t \geq 0}, (\mathbb{P}_x^X)_{x \in \mathcal{X}_\triangle})$,*
- *a QKS$(\theta|K_p|/N, |K_p|)$-process $\mathbb{Y}^p = (\Omega^p, \mathcal{M}^p, (Y_{p,t})_{t \geq 0}, (\mathbb{P}_y^p)_{y \in \mathcal{Y}_\triangle^p})$ for all $p \in [\![1, \ell]\!]$.*

Set $\Omega^Y = \prod_{p=1}^\ell \Omega^p$ and $Y_t = (Y_{p,t})_{p \in [\![1,\ell]\!]}$, with the convention that $Y_t = \triangle$ as long as $Y_{p,t} = \triangle$ for some $p \in [\![1, \ell]\!]$. Also set $\mathcal{M}^Y = \sigma(Y_t : t \geq 0)$ and $\mathbb{P}_y^Y = \otimes_{p=1}^\ell \mathbb{P}_{y_p}^p$ for all $y = (y_p)_{p \in [\![1,\ell]\!]} \in (\mathbb{R}^2)^N$.

Fix $\varepsilon \in (0, 1]$, recall that

$$G_{\mathbf{K},\varepsilon} = \Big\{ x \in \mathcal{X} : \min_{1 \leq p \neq q \leq \ell} \; \min_{i \in K_p, j \in K_q} \|x^i - x^j\|^2 > \varepsilon \Big\} \cap B(0, 1/\varepsilon),$$

and set

$$\tau_{\mathbf{K},\varepsilon} = \{t \geq 0 : X_t \notin G_{\mathbf{K},\varepsilon}\} \quad \text{and} \quad \tilde{\tau}_{\mathbf{K},\varepsilon} = \{t \geq 0 : Y_t \notin G_{\mathbf{K},\varepsilon}\}.$$

Fix $T > 0$. Quasi-everywhere in $G_{\mathbf{K},\varepsilon}$, there is a probability measure $\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}}$ on $(\Omega^X, \mathcal{M}^X)$, equivalent to $\mathbb{P}_x^X$, such that the law of the process $(X_{t \wedge T \wedge \tau_{\mathbf{K},\varepsilon}})_{t \geq 0}$ under $\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}}$ is the same as that of $(Y_{t \wedge T \wedge \tilde{\tau}_{\mathbf{K},\varepsilon}})_{t \geq 0}$ on $(\Omega^Y, \mathcal{M}^Y)$ under $\mathbb{P}_x^Y$.

Furthermore, the Radon–Nikodym density $\frac{\mathrm{d}\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}}}{\mathrm{d}\mathbb{P}_x^X}$ is $\mathcal{M}^X_{T \wedge \tau_{\mathbf{K},\varepsilon}}$-measurable, where as usual $\mathcal{M}_t^X = \sigma(X_s, s \leq t)$, and there is a deterministic constant $C_{T,\varepsilon,\mathbf{K}} > 0$ such that quasi-everywhere in $G_{\mathbf{K},\varepsilon}$,

$$C_{T,\varepsilon,\mathbf{K}}^{-1} \leq \frac{\mathrm{d}\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}}}{\mathrm{d}\mathbb{P}_x^X} \leq C_{T,\varepsilon,\mathbf{K}}.$$

The quasi-everywhere notion refers to the process $\mathbb{X}$. Let us mention that for ζ being the life-time of $\mathbb{X}$, we have $\tau_{\mathbf{K},\varepsilon} \in [0, \zeta]$ when $\zeta < \infty$ because $\triangle \notin G_{\mathbf{K},\varepsilon}$. Although this is not clear at this point of the paper, the event $\{\tau_{\mathbf{K},\varepsilon} = \zeta\}$ has a positive probability if $\max_{p=1,\dots,\ell} |K_p| \geq k_0$.

Proof of Proposition 15. We only consider the case where $\ell = 2$. The general case is heavier in terms of notation but contains no additional difficulty. We fix a nontrivial partition $\mathbf{K} = (K_1, K_2)$ of $[\![1, N]\!]$. The main idea is to apply Lemma B.7 to $\mathbb{X}$ with the function

$$\varrho(x) = \exp(u(x)), \quad \text{where} \quad u(x) = \frac{\theta}{N} \sum_{i \in K_1, j \in K_2} \log \|x^i - x^j\|. \tag{41}$$

Unfortunately, this is not licit because $u \notin \mathcal{F}^X$.

Step 1. Set $\mathbb{Y} = (\Omega^Y, \mathcal{M}^Y, (Y_t)_{t\geq 0}, (\mathbb{P}_y^Y)_{y\in(\mathcal{Y}_1\times\mathcal{Y}_2)\cup\{\triangle\}})$ and fix $\varepsilon \in (0, 1]$ and $n \geq 1$. We first compute the Dirichlet space of $\mathbb{Y}$ killed when it outside $G^n_{\mathbf{K},\varepsilon}$ (recall Lemma 12). Consider the measures

$$\mu_1(\mathrm{d}y) = \prod_{i,j\in K_1, i\neq j} \|y^i - y^j\|^{-\theta/N} \mathrm{d}y \quad \text{and} \quad \mu_2(\mathrm{d}y) = \prod_{i,j\in K_2, i\neq j} \|y^i - y^j\|^{-\theta/N} \mathrm{d}y$$

on $(\mathbb{R}^2)^{|K_1|}$ and $(\mathbb{R}^2)^{|K_2|}$, with $\mu_i(\mathrm{d}y) = \mathrm{d}y$ if $|K_i| = 1$. Recall that $\mu(\mathrm{d}x) = \mathbf{m}(x)\mathrm{d}x$ (see (4)) and by definition (see (41)), $\varrho(x) = \prod_{i\in K_1, j\in K_2} \|x^i - x^j\|^{\theta/N}$; we deduce that

$$\mu_1 \otimes \mu_2 = \varrho^2 \mu.$$

By Proposition 6, for $p = 1, 2$, $\mathbb{Y}^p$ is a $\mathcal{Y}^p_\triangle$-valued μ_p-symmetric diffusion (since $(\theta|K_p|/N)/|K_p| = \theta/N$) with regular Dirichlet space $(\mathcal{E}_p, \mathcal{F}_p)$ with core $C_c^\infty(\mathcal{Y}_p)$, and $\mathcal{E}_p(\varphi, \varphi) = \frac{1}{2}\int_{(\mathbb{R}^2)^{|K_p|}} \|\nabla\varphi\|^2 \,\mathrm{d}\mu_p$ for $\varphi \in C_c^\infty(\mathcal{Y}_p)$. This also holds true if e.g. $|K_1| = 1$ (see [11, Example 4.2.1, p. 167]), since then μ_1 is nothing but the Lebesgue measure on $\mathbb{R}^2$. Since now $\mu_1 \otimes \mu_2 = \varrho^2\mu$, by Lemma B.5, $\mathbb{Y}$ is a $\varrho^2\mu$-symmetric $\mathcal{X}_\triangle$-valued diffusion with regular Dirichlet space $(\mathcal{E}^Y, \mathcal{F}^Y)$ on $L^2(\mathcal{Y}_1 \times \mathcal{Y}_2, \varrho^2\mathrm{d}\mu)$ with core $C_c^\infty(\mathcal{Y}_1 \times \mathcal{Y}_2)$ and, for $\varphi \in C_c^\infty(\mathcal{Y}_1 \times \mathcal{Y}_2)$,

$$\begin{aligned}\mathcal{E}^Y(\varphi, \varphi) &= \int_{(\mathbb{R}^2)^{|K_1|}} \mathcal{E}_2(\varphi(y, \cdot), \varphi(y, \cdot))\, \mu_1(\mathrm{d}y) + \int_{(\mathbb{R}^2)^{|K_2|}} \mathcal{E}_1(\varphi(\cdot, z), \varphi(\cdot, z))\, \mu_2(\mathrm{d}z) \\ &= \frac{1}{2}\int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \varrho^2 \,\mathrm{d}\mu.\end{aligned}$$

Finally, we apply Lemma B.6 to $\mathbb{Y}$ with the open set $G^n_{\mathbf{K},\varepsilon} \subset \mathcal{X} \subset \mathcal{Y}_1 \times \mathcal{Y}_2$, to find that the resulting killed process

$$\mathbb{Y}^{n,\varepsilon} = (\Omega^Y, \mathcal{M}^Y, (Y_t^{n,\varepsilon})_{t\geq 0}, (\mathbb{P}^Y_y)_{y\in G^n_{\mathbf{K},\varepsilon}\cup\{\triangle\}})$$

is a $\varrho^2\mu|_{G^n_{\mathbf{K},\varepsilon}}$-symmetric $G^n_{\mathbf{K},\varepsilon} \cup \{\triangle\}$-valued diffusion, and its regular Dirichlet space $(\mathcal{E}^{Y,n,\varepsilon}, \mathcal{F}^{Y,n,\varepsilon})$ has core $C_c^\infty(G^n_{\mathbf{K},\varepsilon})$ and is such that for all $\varphi \in C_c^\infty(G^n_{\mathbf{K},\varepsilon})$,

$$\mathcal{E}^{Y,n,\varepsilon}(\varphi, \varphi) = \frac{1}{2}\int_{G^n_{\mathbf{K},\varepsilon}} \|\nabla\varphi\|^2 \varrho^2 \,\mathrm{d}\mu.$$

Step 2. We now fix $\varepsilon \in (0, 1]$ and introduce, for each $n \geq 1$, $u_{n,\varepsilon}(x) = u(x)\Gamma^n_{\mathbf{K},\varepsilon}(x)$ (recall (41) and Lemma 12), and $\varrho_{n,\varepsilon} = \exp(u_{n,\varepsilon})$. We check here that the functions $u_{n,\varepsilon}$ and $\varrho_{n,\varepsilon}$

satisfy the assumptions of Lemma B.7 (to be applied to $\mathbb{X}$), that $\mathcal{A}^X[\varrho_{n,\varepsilon}-1]=\mathcal{L}^X\varrho_{n,\varepsilon}$ and that

$$\sup_{n\geq 1}\sup_{x\in\mathcal{X}}|u_{n,\varepsilon}(x)|<\infty \quad\text{and}\quad \sup_{n\geq 1}\sup_{x\in G^n_{\mathbf{K},\varepsilon}}|\mathcal{L}^X\varrho_{n,\varepsilon}(x)|<\infty. \tag{42}$$

First, $u_{n,\varepsilon}\in\mathcal{F}^X$ because $u_{n,\varepsilon}\in C^\infty_c(\mathcal{X})$, and $|u_{n,\varepsilon}|$ is bounded, uniformly in $n\geq 1$, because $\Gamma^n_{\mathbf{K},\varepsilon}$ is bounded by 1 and vanishes outside $G_{\mathbf{K},\eta}$ (see Lemma 12), while u is smooth on $G_{\mathbf{K},\eta}$. To show that $\mathcal{A}^X[\varrho_{n,\varepsilon}-1]=\mathcal{L}^X\varrho_{n,\varepsilon}$, it suffices by Remark 8 to verify that $\varrho_{n,\varepsilon}-1\in C^\infty_c(\mathcal{X})$, which is clear, and $\sup_{\alpha\in(0,1]}\sup_{x\in(\mathbb{R}^2)^N}|\mathcal{L}^X_\alpha\varrho_{n,\varepsilon}(x)|<\infty$. We have

$$\mathcal{L}^X_\alpha\varrho_{n,\varepsilon}(x)=e^{u_{n,\varepsilon}(x)}\mathcal{L}^X_\alpha u_{n,\varepsilon}(x)+\tfrac{1}{2}e^{u_{n,\varepsilon}(x)}\|\nabla u_{n,\varepsilon}(x)\|^2.$$

Since $u_{n,\varepsilon}\in C^\infty_c((\mathbb{R}^2)^N)$, it suffices to show that $\sup_{\alpha\in(0,1]}\sup_{x\in(\mathbb{R}^2)^N}|\mathcal{L}^X_\alpha u_{n,\varepsilon}(x)|<\infty$. By (22),

$$\mathcal{L}^X_\alpha u_{n,\varepsilon}(x)=\Gamma^n_{\mathbf{K},\varepsilon}(x)\mathcal{L}^X_\alpha u(x)+u(x)\mathcal{L}^X_\alpha\Gamma^n_{\mathbf{K},\varepsilon}(x)+\nabla\Gamma^n_{\mathbf{K},\varepsilon}(x)\cdot\nabla u(x).$$

Again, the only difficulty is in the first term, because $\mathcal{L}^X_\alpha\Gamma^n_{\mathbf{K},\varepsilon}$ is uniformly bounded by Lemma 12 and vanishes outside $G_{\mathbf{K},\eta}$, while u is smooth on $G_{\mathbf{K},\eta}$. Since $\operatorname{supp}\Gamma^n_{\mathbf{K},\varepsilon}\subset G_{\mathbf{K},\eta}$, the task is reduced to showing that $\sup_{\alpha\in(0,1]}\sup_{x\in G_{\mathbf{K},\eta}}|\mathcal{L}^X_\alpha u(x)|<\infty$. But

$$\mathcal{L}^X_\alpha u=\frac{1}{2}\Delta u-\frac{\theta}{N}S_\alpha,\quad\text{where}\quad S_\alpha(x)=\sum_{1\leq i,j\leq N}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\cdot\nabla_{x^i}u(x),$$

and we only have to verify that

$$\sup_{\alpha\in(0,1]}\sup_{x\in G_{\mathbf{K},\eta}}|S_\alpha(x)|<\infty.$$

For $k\in K_1$ and $\ell\in K_2$, we have

$$\nabla_{x^k}u(x)=\sum_{j\in K_2}\frac{\theta}{N}\frac{x^k-x^j}{\|x^k-x^j\|^2}\quad\text{and}\quad\nabla_{x^\ell}u(x)=\sum_{i\in K_1}\frac{\theta}{N}\frac{x^\ell-x^i}{\|x^\ell-x^i\|^2}.$$

Hence $S_\alpha=S_{1,\alpha}+S_{2,\alpha}+S_{3,\alpha}+S_{4,\alpha}$, where

$$S_{1,\alpha}(x)=\frac{\theta}{N}\sum_{i,j\in K_1}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\cdot\sum_{k\in K_2}\frac{x^i-x^k}{\|x^i-x^k\|^2},$$

$$S_{2,\alpha}(x)=\frac{\theta}{N}\sum_{i\in K_2,j\in K_1}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\cdot\sum_{k\in K_1}\frac{x^i-x^k}{\|x^i-x^k\|^2},$$

and $S_{3,\alpha}$ (resp. $S_{4,\alpha}$) is defined as $S_{1,\alpha}$ (resp. $S_{2,\alpha}$) with the roles of K_1 and K_2 exchanged. First, $S_{2,\alpha}$ (and $S_{4,\alpha}$) is obviously uniformly bounded on $G_{\mathbf{K},\eta}$. Next, by symmetry,

$$S_{1,\alpha}(x)=\frac{\theta}{2N}\sum_{i,j\in K_1}\frac{x^i-x^j}{\|x^i-x^j\|^2+\alpha}\sum_{k\in K_2}\left(\frac{x^i-x^k}{\|x^i-x^k\|^2}-\frac{x^j-x^k}{\|x^j-x^k\|^2}\right).$$

Moreover, there is $C_\eta > 0$ such that for all $x \in G_{\mathbf{K},\eta}$, all $i, j \in K_1$ distinct, and all $k \in K_2$,

$$\left\| \frac{x^i - x^k}{\|x^i - x^k\|^2} - \frac{x^j - x^k}{\|x^j - x^k\|^2} \right\| \le C_\eta \|x^i - x^j\|,$$

so that $S_{1,\alpha}$ (and $S_{3,\alpha}$) is bounded on $G_{\mathbf{K},\eta}$, uniformly in $\alpha \in (0, 1]$, as desired.

Finally, the above computations, together with the facts that $\Gamma^n_{\mathbf{K},\varepsilon} = 1$ on $G^n_{\mathbf{K},\varepsilon}$, also show that for $x \in G^n_{\mathbf{K},\varepsilon}$,

$$\mathcal{L}^X \varrho_{n,\varepsilon}(x) = e^{u(x)} \left(\frac{1}{2} \Delta u(x) - \frac{\theta}{N} S_\alpha(x) \right) + \frac{1}{2} e^{u(x)} \|\nabla u(x)\|^2,$$

which is bounded on $G_{\mathbf{K},\eta}$. Since $G^n_{\mathbf{K},\varepsilon} \subset G_{\mathbf{K},\eta}$, this gives $\sup_{n\ge1} \sup_{x\in G^n_{\mathbf{K},\varepsilon}} |\mathcal{L}^X \varrho_{n,\varepsilon}(x)| < \infty$ and finishes the step.

Step 3. We apply Lemma B.7 to the process $\mathbb{X}$ with $u_{n,\varepsilon}$ and $\varrho_{n,\varepsilon}$ defined in Step 2. Recalling that $\mathcal{A}^X[\varrho_{n,\varepsilon} - 1] = \mathcal{L}^X \varrho_{n,\varepsilon}$ and using the conventions $\varrho_{n,\varepsilon}(\triangle) = 1$ and $\mathcal{L}^X \varrho_{n,\varepsilon}(\triangle) = 0$, we set

$$L^{n,\varepsilon}_t = \frac{\varrho_{n,\varepsilon}(X_t)}{\varrho_{n,\varepsilon}(X_0)} \exp\left(- \int_0^t \frac{\mathcal{L}^X \varrho_{n,\varepsilon}(X_s)}{\varrho_{n,\varepsilon}(X_s)} \,\mathrm{d}s \right). \tag{43}$$

Set $\mathcal{M}^X_t = \sigma(X_s, s \le t)$. By Lemma B.7, there is a family $(\mathbb{Q}^{n,\varepsilon}_x)_{x\in\mathcal{X}\cup\{\triangle\}}$ of probability measures such that

$$\mathbb{Q}^{n,\varepsilon}_x = L^{n,\varepsilon}_t \cdot \mathbb{P}^X_x \quad \text{on } \mathcal{M}^X_t$$

for all $t \ge 0$ and quasi-everywhere in $\mathcal{X} \cup \{\triangle\}$, and such that

$$\mathbb{X}^{n,\varepsilon} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\ge0}, (\mathbb{Q}^{n,\varepsilon}_x)_{x\in\mathcal{X}_\triangle})$$

is a $\varrho^2_{n,\varepsilon}\mu$-symmetric $\mathcal{X} \cup \{\triangle\}$-valued diffusion with regular Dirichlet space $(\mathcal{E}^{n,\varepsilon}, \mathcal{F}^{n,\varepsilon})$ with core $C^\infty_c(\mathcal{X})$ such that for all $\varphi \in C^\infty_c(\mathcal{X})$,

$$\mathcal{E}^{n,\varepsilon}(\varphi,\varphi) = \frac{1}{2} \int_{(\mathbb{R}^2)^N} \|\nabla\varphi\|^2 \varrho^2_{n,\varepsilon} \,\mathrm{d}\mu.$$

Next, we apply Lemma B.6 to $\mathbb{X}^{n,\varepsilon}$ with the open set $G^n_{\mathbf{K},\varepsilon}$; the resulting killed process

$$\mathbb{X}^{*,n,\varepsilon} = \left(\Omega^X, \mathcal{M}^X, (X^{*,n,\varepsilon}_t)_{t\ge0}, (\mathbb{Q}^{n,\varepsilon}_x)_{x\in G^n_{\mathbf{K},\varepsilon}\cup\{\triangle\}}\right)$$

is a $\varrho^2_{n,\varepsilon}\mu|_{G^n_{\mathbf{K},\varepsilon}}$-symmetric $G^n_{\mathbf{K},\varepsilon} \cup \{\triangle\}$-valued diffusion, its regular Dirichlet space $(\mathcal{E}^{*,n,\varepsilon}, \mathcal{F}^{*,n,\varepsilon})$ has core $C^\infty_c(G^n_{\mathbf{K},\varepsilon})$, and for all $\varphi \in C^\infty_c(G^n_{\mathbf{K},\varepsilon})$,

$$\mathcal{E}^{*,n,\varepsilon}(\varphi,\varphi) = \frac{1}{2} \int_{G^n_{\mathbf{K},\varepsilon}} \|\nabla\varphi\|^2 \varrho^2_{n,\varepsilon} \,\mathrm{d}\mu.$$

Comparing this Dirichlet space with the one found in Step 1, using $\varrho_{n,\varepsilon} = \varrho$ on $G^n_{\mathbf{K},\varepsilon}$ and a uniqueness argument (see [11, Theorem 4.2.8, p. 167]), we conclude that quasi-everywhere in $G^n_{\mathbf{K},\varepsilon}$, the law of $X^{*,n,\varepsilon}$ under $\mathbb{Q}^{n,\varepsilon}_x$ equals the law of $Y^{n,\varepsilon}$ under $\mathbb{P}^Y_x$.

Step 4. We fix $T > 0$ and $\varepsilon \in (0, 1]$ and complete the proof. Since $\mathbb{Q}_x^{n,\varepsilon} = L_T^{n,\varepsilon} \cdot \mathbb{P}_x^X$ on $\mathcal{M}_T^X$, we know from Step 3 that for all $n \geq 1$, quasi-everywhere in $G_{\mathbf{K},\varepsilon}^n$, for all continuous bounded $\Phi : C([0, T], \mathcal{X}_\triangle) \to \mathbb{R}$ (observe that $\bar{G}_{\mathbf{K},\varepsilon}^n \subset \mathcal{X} \subset \mathcal{X}_\triangle$),

$$\mathbb{E}_x^X[\Phi(X_{\cdot\wedge\tau_{\mathbf{K},n,\varepsilon}\wedge T})L_T^{n,\varepsilon}] = \mathbb{E}_x^Y[\Phi(Y_{\cdot\wedge\tilde{\tau}_{\mathbf{K},n,\varepsilon}\wedge T})],$$

where $\tau_{\mathbf{K},n,\varepsilon} = \inf\{t > 0 : X_t \notin G_{\mathbf{K},\varepsilon}^n\} \wedge \tau_{\mathbf{K},\varepsilon}$ and $\tilde{\tau}_{\mathbf{K},n,\varepsilon} = \inf\{t > 0 : Y_t \notin G_{\mathbf{K},\varepsilon}^n\} \wedge \tilde{\tau}_{\mathbf{K},\varepsilon}$. Since $(L_t^{n,\varepsilon})_{t\geq 0}$ is a $\mathbb{P}_x^X$-martingale by Lemma B.7, we deduce that quasi-everywhere in $G_{\mathbf{K},\varepsilon}^n$,

$$\mathbb{E}_x^X[\Phi(X_{\cdot\wedge\tau_{\mathbf{K},n,\varepsilon}\wedge T})L_{\tau_{\mathbf{K},n,\varepsilon}\wedge T}^{n,\varepsilon}] = \mathbb{E}_x^Y[\Phi(Y_{\cdot\wedge\tilde{\tau}_{\mathbf{K},n,\varepsilon}\wedge T})]. \tag{44}$$

Recall that $G_{\mathbf{K},\varepsilon} \subset \bigcup_{n\geq 1} G_{\mathbf{K},\varepsilon}^n$ (see Lemma 12). Hence $\lim_n \tau_{\mathbf{K},n,\varepsilon} = \tau_{\mathbf{K},\varepsilon}$, $\lim_n \tilde{\tau}_{\mathbf{K},n,\varepsilon} = \tilde{\tau}_{\mathbf{K},\varepsilon}$, and for each $x \in G_{\mathbf{K},\varepsilon}$, there is $n_x \geq 1$ such that $x \in G_{\mathbf{K},\varepsilon}^n$ for all $n \geq n_x$. We deduce from (44) that q.e. in $G_{\mathbf{K},\varepsilon}$, the process $(L_{\tau_{\mathbf{K},n,\varepsilon}\wedge T}^{n,\varepsilon})_{n\geq n_x}$ is an $(\mathcal{M}_{\tau_{\mathbf{K},n,\varepsilon}\wedge T}^X)_{n\geq n_x}$-martingale under $\mathbb{P}_x^X$. Moreover, recalling the expression (43) of $L^{n,\varepsilon}$, the equality $\varrho_{n,\varepsilon} = \exp(u_{n,\varepsilon})$ and the bound (42), we conclude that there is a constant $C_{T,\varepsilon,\mathbf{K}} > 0$ such that quasi-everywhere in $G_{\mathbf{K},\varepsilon}$,

$$C_{T,\varepsilon,\mathbf{K}}^{-1} \leq L_{\tau_{\mathbf{K},n,\varepsilon}\wedge T}^{n,\varepsilon} \leq C_{T,\varepsilon,\mathbf{K}} \quad \mathbb{P}_x^X\text{-a.s., for all } n \geq n_x.$$

Hence the martingale $(L_{\tau_{\mathbf{K},n,\varepsilon}\wedge T}^{n,\varepsilon})_{n\geq n_x}$ is closed by some $\mathcal{M}_{\tau_{\mathbf{K},\varepsilon}\wedge T}$-measurable random variable $J_{T,\varepsilon,\mathbf{K}}$ that satisfies $C_{T,\varepsilon,\mathbf{K}}^{-1} \leq J_{T,\varepsilon,\mathbf{K}} \leq C_{T,\varepsilon,\mathbf{K}}$, and (44) implies that for $n \geq n_x$,

$$\mathbb{E}_x^X[\Phi(X_{\cdot\wedge\tau_{\mathbf{K},n,\varepsilon}\wedge T})J_{T,\varepsilon,\mathbf{K}}] = \mathbb{E}_x^Y[\Phi(Y_{\cdot\wedge\tilde{\tau}_{\mathbf{K},n,\varepsilon}\wedge T})].$$

Letting $n \to \infty$, we find that quasi-everywhere in $G_{\mathbf{K},\varepsilon}$, for $\Phi \in C_b(C([0, T], \mathcal{X}_\triangle), \mathbb{R})$,

$$\mathbb{E}_x^X[\Phi(X_{\cdot\wedge\tau_{\mathbf{K},\varepsilon}\wedge T})J_{T,\varepsilon,\mathbf{K}}] = \mathbb{E}_x^Y[\Phi(Y_{\cdot\wedge\tilde{\tau}_{\mathbf{K},\varepsilon}\wedge T})].$$

Setting $\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}} = J_{T,\varepsilon,\mathbf{K}} \cdot \mathbb{P}_x^X$ completes the proof. ∎

8. Explosion and continuity at explosion

In this section we consider a $QKS(\theta, N)$-process $\mathbb{X}$ with life-time ζ. We show that $\zeta = \infty$ when $\theta \in (0, 2)$, and $\zeta < \infty$ when $\theta \geq 2$. In the latter case, we also prove that $\lim_{t\to\zeta-} X_t$ a.s. exists, *for the usual topology of* $(\mathbb{R}^2)^N$: the Keller–Segel process is continuous at explosion. This is not clear at all at first sight: we know that $\lim_{t\to\zeta-} X_t = \triangle$ a.s. for the one-point compactification topology, which means that the process escapes from every compact subset of $\mathcal{X}$, but it could either go to infinity, which is not difficult to exclude, or tend to the boundary of $\mathcal{X}$ without converging, e.g. because it could alternate very fast between having its particles labeled in $[\![1, k_0]\!]$ very close and having its particles labeled in $[\![2, k_0 + 1]\!]$ very close. The goal of this section is to prove the following result.

Proposition 16. *Fix $\theta > 0$ and $N \geq 2$ such that $N > \theta$, set $k_0 = \lceil 2N/\theta \rceil$ and $\mathcal{X} = E_{k_0}$ and consider a $QKS(\theta, N)$-process $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}\cup\{\triangle\}})$ with life-time ζ.*

(i) *If* $\theta < 2$*, then quasi-everywhere,* $\mathbb{P}_x^X(\zeta = \infty) = 1$*.*

(ii) *If* $\theta \geq 2$*, then quasi-everywhere,* $\mathbb{P}_x^X$*-a.s.,* $\zeta < \infty$ *and* $X_{\zeta-} = \lim_{t\to\zeta} X_t$ *exists for the usual topology of* $(\mathbb{R}^2)^N$ *and does not belong to* E_{k_0}*.*

We first show that the process does not explode in the subcritical case and cannot go to infinity at explosion in the supercritical case.

Lemma 17. (i) *If* $\theta < 2$ *and* $N \geq 2$*, then quasi-everywhere,* $\mathbb{P}_x^X(\zeta = \infty) = 1$*.*

(ii) *If* $\theta \geq 2$ *and* $N > \theta$*, then quasi-everywhere,*

$$\mathbb{P}_x^X\Big(\zeta < \infty \text{ and } \sup_{t\in[0,\zeta)} \|X_t\| < \infty\Big) = 1.$$

Proof. The arguments below only apply quasi-everywhere, since we use Proposition 10. In both cases, for all $i \in [\![1, N]\!]$ and all $t \in [0, \zeta)$ we have

$$\begin{aligned}\|X_t\|^2 &\leq 2\sum_{i=1}^N \big(\|X_t^i - S_{[\![1,N]\!]}(X_t)\|^2 + \|S_{[\![1,N]\!]}(X_t)\|^2\big) \\ &= 2R_{[\![1,N]\!]}(X_t) + 2N\|S_{[\![1,N]\!]}(X_t)\|^2.\end{aligned}$$

By Lemma 11, there are a Brownian motion $(M_t)_{t\geq 0}$ and a squared Bessel process $(D_t)_{t\geq 0}$ of dimension $d_{\theta,N}(N)$ (killed when it gets out of $(0,\infty)$ if $d_{\theta,N}(N) \leq 0$) such that $S_{[\![1,N]\!]}(X_t) = M_t$ and $R_{[\![1,N]\!]}(X_t) = D_t$ for all $t \in [0, \zeta)$. These processes being locally bounded, we conclude that

$$\sup_{t\in[0,\zeta\wedge T)} \|X_t\| < \infty \quad \text{a.s., for all } T > 0. \tag{45}$$

(i) When $\theta < 2$ and $N \geq 2$, we have $k_0 = \lceil 2N/\theta \rceil > N$, so that $\mathcal{X} = (\mathbb{R}^2)^N$. Hence on the event $\{\zeta < \infty\}$, we necessarily have $\limsup_{t\to\zeta-} \|X_t\| = \infty$, and this is incompatible with (45) with $T = \zeta$.

(ii) When $N > \theta \geq 2$, we have $d_{\theta,N}(N) \leq 0$, so that $(D_t)_{t\geq 0}$ is killed at some finite time τ. Moreover, $\zeta \leq \tau$. Indeed, on the event that $\tau < \zeta$, we have $R_{[\![1,N]\!]}(X_\tau) = \lim_{t\to\tau-} R_{[\![1,N]\!]}(X_t) = \lim_{t\to\tau-} D_t = 0$, so that $X_\tau \notin E_{k_0}$ (since $k_0 \leq N$), which is not possible since $\tau < \zeta$. Hence ζ is also a.s. finite and $\sup_{t\in[0,\zeta)} \|X_t\| < \infty$ a.s. by (45) with the choice $T = \zeta$. ∎

To show the continuity at explosion in the supercritical case, we need to prove the following delicate lemma.

Lemma 18. *Assume that* $N > \theta \geq 2$*. Quasi-everywhere, for all* $K \subset [\![1, N]\!]$ *with* $|K| \geq 2$*,*

$$\mathbb{P}_x^X\text{-a.s.,} \quad \lim_{t\to\zeta-} R_K(X_t) = 0 \quad \text{or} \quad \liminf_{t\to\zeta-} R_K(X_t) > 0.$$

Proof. We proceed by reverse induction on the cardinality of K. If $K = [\![1, N]\!]$, the result is clear because $(R_{[\![1,N]\!]}(X_t))_{t\in[0,\zeta)}$ is a (killed) squared Bessel process on $[0, \zeta)$

by Lemma 11 (and since $\zeta \le \tau$ exactly as in the proof of Lemma 17 (ii)), hence it has a limit in $\mathbb{R}_+$ as $t \to \zeta$. Now, we assume that the property is proved if $|K| \ge n$ where $n \in [\![3, N]\!]$, we take $K \subset [\![1, N]\!]$ such that $|K| = n-1$, and we show in several steps that a.s., either $\lim_{t\to\zeta-} R_K(X_t) = 0$ or $\liminf_{t\to\zeta-} R_K(X_t) > 0$.

Step 1. We fix $\varepsilon \in (0,1]$ and introduce $\tilde\sigma_0^\varepsilon = 0$ and, for $k \ge 1$,

$$\sigma_k^\varepsilon = \inf\{t \in (\tilde\sigma_{k-1}^\varepsilon, \zeta) : R_K(X_t) \le \varepsilon\} \quad \text{and} \quad \tilde\sigma_k^\varepsilon = \inf\{t \in (\sigma_k^\varepsilon, \zeta) : R_K(X_t) \ge 2\varepsilon\},$$

with the convention that $\inf\emptyset = \zeta$. We show in this step that for all deterministic $A > 0$, there exists a constant $p_{A,\varepsilon} > 0$ such that for all $k \ge 1$, quasi-everywhere, on $\{\sigma_k^\varepsilon < \zeta\}$,

$$\mathbb{P}_x^X\big(\{\tilde\sigma_k^\varepsilon \ge (\sigma_k^\varepsilon + A)\wedge\zeta\} \cup B_{k,\varepsilon} \,\big|\, \mathcal{M}^X_{\sigma_k^\varepsilon}\big) \ge p_{A,\varepsilon},$$

where $\mathcal{M}_t^X = \sigma(X_s : s \in [0,t])$ and, with $a_\varepsilon = c_{|K|+1}\varepsilon/c_{|K|}$ (recall Lemma 13),

$$B_{k,\varepsilon} = \Big\{ \sup_{t\in[\sigma_k^\varepsilon, \tilde\sigma_k^\varepsilon)} \|X_t\| \ge 1/\varepsilon \text{ or } \inf_{t\in[\sigma_k^\varepsilon,\tilde\sigma_k^\varepsilon)} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \le a_\varepsilon \Big\}.$$

By the strong Markov property of $\mathbb{X}$, on $\{\sigma_k^\varepsilon < \zeta\}$,

$$\mathbb{P}_x^X\big(\{\tilde\sigma_k^\varepsilon \ge (\sigma_k^\varepsilon + A)\wedge\zeta\} \cup B_{k,\varepsilon} \,\big|\, \mathcal{M}^X_{\sigma_k^\varepsilon}\big) = g(X_{\sigma_k^\varepsilon}),$$

where

$$g(y) = \mathbb{P}_y^X\big(\{\tilde\sigma_1^\varepsilon \ge (\sigma_1^\varepsilon + A)\wedge\zeta\} \cup B_{1,\varepsilon}\big) = \mathbb{P}_y^X\big(\{\tilde\sigma_1^\varepsilon \ge A\wedge\zeta\} \cup C_{1,\varepsilon}\big)$$

and

$$C_{1,\varepsilon} = \Big\{ \sup_{t\in[0,\tilde\sigma_1^\varepsilon)} \|X_t\| \ge 1/\varepsilon \text{ or } \inf_{t\in[0,\tilde\sigma_1^\varepsilon)} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \le a_\varepsilon \Big\}.$$

We have used the fact that $R_K(X_{\sigma_k^\varepsilon}) \le \varepsilon$ on $\{\sigma_k^\varepsilon < \zeta\}$ by definition of σ_k^ε, so that $\sigma_1^\varepsilon = 0$ under $\mathbb{P}^X_{X_{\sigma_k^\varepsilon}}$. Using again $R_K(X_{\sigma_k^\varepsilon}) \le \varepsilon$ on $\{\sigma_k^\varepsilon < \zeta\}$, it suffices to show that there is a constant $p_{A,\varepsilon} > 0$ such that $g(y) \ge p_{A,\varepsilon}$ quasi-everywhere in $\{y \in \mathcal{X} : R_K(y) \le \varepsilon\}$.

If $\|y\| \ge 1/\varepsilon$ or $\min_{i\notin K} R_{K\cup\{i\}}(y) \le a_\varepsilon$, then clearly $g(y) = 1$.

Otherwise, $y \in G_{\mathbf{K},\varepsilon}$, where

$$G_{\mathbf{K},\varepsilon} = \{x \in \mathcal{X} : \|x^i - x^j\|^2 > \varepsilon \text{ for all } i \in K \text{ and } j \notin K\} \cap B(0, 1/\varepsilon)$$

as in Proposition 15 with $\mathbf{K} = (K, K^c)$, because $\|y\| < 1/\varepsilon$ and because $R_K(y) \le \varepsilon < 2\varepsilon$ and $\min_{i\notin K} R_{K\cup\{i\}}(y) > a_\varepsilon = c_{|K|+1}\varepsilon/c_{|K|}$ imply that $\|x^i - x^k\|^2 > \varepsilon$ for all $i \in K$ and $j \notin K$ by Lemma 13. For the same reasons and by definition of $\tilde\sigma_1^\varepsilon$,

$$C_{1,\varepsilon}^c \subset \{X_t \in G_{\mathbf{K},\varepsilon} \text{ for all } t \in [0, \tilde\sigma_1^\varepsilon)\}. \tag{46}$$

We now apply Proposition 15 with $T = A$ (and ε) and we find that quasi-everywhere in $G_{\mathbf{K},\varepsilon}$,

$$\begin{aligned} g(y) &\ge C^{-1}_{A,\varepsilon,\mathbf{K}} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(\{\tilde\sigma_1^\varepsilon \ge A\wedge\zeta\} \cup C_{1,\varepsilon}) \\ &= C^{-1}_{A,\varepsilon,\mathbf{K}} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(\{\tilde\sigma_1^\varepsilon \ge A\wedge\zeta\} \cap C_{1,\varepsilon}^c) + C^{-1}_{A,\varepsilon,\mathbf{K}} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(C_{1,\varepsilon}). \end{aligned} \tag{47}$$

But we know from Proposition 15 and Lemma 11 that under $\mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}$, $(R_K(X_t))_{t\in[0,\tau_{K,\varepsilon}\wedge A]}$ is a squared Bessel process of dimension $d_{\theta|K|/N,|K|}(|K|) = d_{\theta,N}(|K|)$, issued from $R_K(y) \le \varepsilon$, stopped at time $\tau_{\mathbf{K},\varepsilon} \wedge A$, where $\tau_{\mathbf{K},\varepsilon} = \inf\{t > 0 : X_t \notin G_{\mathbf{K},\varepsilon}\}$. Hence there exists, under $\mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}$, a squared Bessel process $(S_t)_{t\ge 0}$ of dimension $d_{\theta,N}(|K|)$ such that $S_t = R_K(X_t)$ for all $t \in [0, \tau_{\mathbf{K},\varepsilon} \wedge A]$. We introduce $\kappa_\varepsilon = \inf\{t > 0 : S_t \ge 2\varepsilon\}$ and we observe that

$$\{\kappa_\varepsilon \ge A \wedge \zeta\} \cap C_{1,\varepsilon}^c = \{\tilde\sigma_1^\varepsilon \ge A\} \cap C_{1,\varepsilon}^c.$$

Indeed, we have used the fact that on $C_{1,\varepsilon}^c$, we have $\tau_{\mathbf{K},\varepsilon} \ge \tilde\sigma_1^\varepsilon$ by (46) so that $R_K(X_t) = S_t$ for all $t \in [0, \tilde\sigma_1^\varepsilon \wedge A)$, from which we deduce that $\kappa_\varepsilon \ge A \wedge \zeta$ if and only if $\tilde\sigma_1^\varepsilon \ge A \wedge \zeta$. Coming back to (47), we get

$$\begin{aligned} g(y) &\ge C_{A,\varepsilon,\mathbf{K}}^{-1} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(\{\kappa_\varepsilon \ge A \wedge \zeta\} \cap C_{1,\varepsilon}^c) + C_{A,\varepsilon,\mathbf{K}}^{-1} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(C_{1,\varepsilon}) \\ &= C_{A,\varepsilon,\mathbf{K}}^{-1} \mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(\kappa_\varepsilon \ge A \wedge \zeta). \end{aligned}$$

The step is complete, since $\mathbb{Q}_y^{A,\varepsilon,\mathbf{K}}(\kappa_\varepsilon \ge A)$ is the probability that a squared Bessel process of dimension $d_{\theta,N}(|K|)$ issued from $R_K(y) \le \varepsilon$ remains below 2ε during $[0, A]$ and is thus strictly positive, uniformly in y (such that $y \in G_{\mathbf{K},\varepsilon}$ and $R_K(y) \le \varepsilon$).

Step 2. We prove here that for all $\varepsilon \in (0, 1]$ all $A > 0$, quasi-everywhere,

$$\mathbb{P}_x^X\Big(\limsup_{t\to\zeta-} \|X_t\| \ge 1/\varepsilon \text{ or } \liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \le a_\varepsilon \text{ or } \exists k \ge 1,\ \sigma_k^\varepsilon \ge \zeta \wedge A\Big) = 1.$$

All the arguments below only hold quasi-everywhere, even if we do not mention it explicitly during this step. For $k \ge 1$, we introduce, with $B_{k,\varepsilon}$ defined in Step 1,

$$\Omega_{k+1} = \{\sigma_{k+1}^\varepsilon < \zeta \wedge A\} \cap B_{k,\varepsilon}^c,$$

and we first show that $\mathbb{P}_x^X(\liminf_k \Omega_k) = 0$. To this end, it suffices to check that for all $\ell \ge 1$, $\mathbb{P}_x^X(\bigcap_{k=\ell}^\infty \Omega_k) = 0$. Since Ω_k is $\mathcal{M}_{\sigma_k^\varepsilon}$-measurable, for all $m \ge \ell \ge 1$,

$$\mathbb{P}_x^X\Big(\bigcap_{k=\ell}^{m+1} \Omega_k\Big) = \mathbb{E}_x^X[\mathbf{1}_{\bigcap_{k=\ell}^m \Omega_k} \mathbb{P}_x^X(\Omega_{m+1} \mid \mathcal{M}_{\sigma_m^\varepsilon})].$$

Since moreover $\bigcap_{k=\ell}^m \Omega_k \subset \{\sigma_m^\varepsilon < \zeta\}$ and $\sigma_{m+1}^\varepsilon \ge \tilde\sigma_m^\varepsilon \ge \tilde\sigma_m^\varepsilon - \sigma_m^\varepsilon$, we deduce that on $\bigcap_{k=\ell}^m \Omega_k$,

$$\begin{aligned} \mathbb{P}_x^X(\Omega_{m+1} \mid \mathcal{M}_{\sigma_m^\varepsilon}) &= 1 - \mathbb{P}_x^X(\{\sigma_{m+1}^\varepsilon \ge \zeta \wedge A\} \cup B_{m,\varepsilon} \mid \mathcal{M}_{\sigma_m^\varepsilon}) \\ &\le 1 - \mathbb{P}_x^X(\{\tilde\sigma_m^\varepsilon \ge (\sigma_m^\varepsilon + A) \wedge \zeta\} \cup B_{m,\varepsilon} \mid \mathcal{M}_{\sigma_m^\varepsilon}), \end{aligned}$$

so that $\mathbb{P}_x^X(\Omega_{m+1} \mid \mathcal{M}_{\sigma_m^\varepsilon}) \le 1 - p_{A,\varepsilon}$ by Step 1. Hence we conclude that

$$\mathbb{P}_x^X\Big(\bigcap_{k=\ell}^{m+1} \Omega_k\Big) \le (1 - p_{A,\varepsilon}) \mathbb{P}_x^X\Big(\bigcap_{k=\ell}^{m} \Omega_k\Big) \quad \text{for all } m \ge \ell \ge 1,$$

so that $\mathbb{P}_x^X(\bigcap_{k=\ell}^\infty \Omega_k) = 0$ as desired.

Hence $\mathbb{P}_x^X(\liminf_k \Omega_k) = 0$, so that a.s., an infinite number of Ω_k^c are realized. Recalling that

$$\Omega_{k+1}^c = \Big\{\sigma_{k+1}^\varepsilon \geq \zeta \wedge A \text{ or } \inf_{t\in[\sigma_k^\varepsilon,\tilde\sigma_k^\varepsilon)} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \leq a_\varepsilon \text{ or } \sup_{t\in[\sigma_k^\varepsilon,\tilde\sigma_k^\varepsilon)} \|X_t\| \geq 1/\varepsilon\Big\},$$

we find the following alternative: either

- there is $k \geq 1$ such that $\sigma_k^\varepsilon \geq \zeta \wedge A$; or
- for all $k \geq 1$, we have $\sigma_k^\varepsilon < \zeta$ and $\inf_{t\in[\sigma_k^\varepsilon,\tilde\sigma_k^\varepsilon)} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \leq a_\varepsilon$ for infinitely many k's, which implies that $\liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \leq a_\varepsilon$ because $\lim_\infty \sigma_k^\varepsilon = \zeta$ by definition of $(\sigma_k^\varepsilon)_{k\geq1}$ and by continuity of $t \mapsto R_K(X_t)$ on $[0,\zeta)$; or
- for all $k \geq 1$, we have $\sigma_k^\varepsilon < \zeta$ and there are infinitely many k's for which $\sup_{t\in[\sigma_k^\varepsilon,\tilde\sigma_k^\varepsilon)} \|X_t\| \geq 1/\varepsilon$; note that this implies that $\limsup_{t\to\zeta-} \|X_t\| \geq 1/\varepsilon$, because $\lim_\infty \sigma_k^\varepsilon = \zeta$ as previously.

Step 3. We conclude the proof. Applying Step 2, we find that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $A \in \mathbb{N}$ and all $\varepsilon \in \mathbb{Q} \cap (0,1]$,

$$\limsup_{t\to\zeta-} \|X_t\| \geq 1/\varepsilon \quad \text{or} \quad \liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \leq a_\varepsilon \quad \text{or} \quad \exists k \geq 1,\ \sigma_k^\varepsilon \geq \zeta \wedge A.$$

By Lemma 17 (ii), we know that $\zeta < \infty$, so that choosing $A = \lceil\zeta\rceil$, we conclude that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $\varepsilon \in \mathbb{Q} \cap (0,1]$,

$$\limsup_{t\to\zeta-} \|X_t\| \geq 1/\varepsilon \quad \text{or} \quad \liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) \leq a_\varepsilon \quad \text{or} \quad \exists k \geq 1,\ \sigma_k^\varepsilon = \zeta. \tag{48}$$

And by Lemma 17 (ii) again, $\limsup_{t\to\zeta-} \|X_t\| \leq 1/\varepsilon_0$ for some (random) $\varepsilon_0 \in (0,1]$.

On the event where $\liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) = 0$, there exists some (random) $i_0 \notin K$ such that $\liminf_{t\to\zeta-} R_{K\cup\{i_0\}}(X_t) = 0$, whence $\lim_{t\to\zeta-} R_{K\cup\{i_0\}}(X_t) = 0$ by induction assumption, and this obviously implies that $\lim_{t\to\zeta-} R_K(X_t) = 0$.

On the complementary event, we consider $\varepsilon_1 \in (0,\varepsilon_0]$ enjoying the property that $\liminf_{t\to\zeta-} \min_{i\notin K} R_{K\cup\{i\}}(X_t) > a_{\varepsilon_1}$ and we conclude from (48) and the inequality $\limsup_{t\to\zeta-} \|X_t\| \leq 1/\varepsilon_0$ that for all $\varepsilon \in \mathbb{Q} \cap (0,\varepsilon_1]$, there exists $k_\varepsilon \geq 1$ such that $\sigma_{k_\varepsilon}^\varepsilon = \zeta$. Recalling the definition of $(\sigma_k^\varepsilon)_{k\geq1}$, we deduce that for all $\varepsilon \in \mathbb{Q} \cap (0,\varepsilon_1]$, $R_K(X_t)$ upcrosses the segment $[\varepsilon, 2\varepsilon]$ a finite number of times during $[0,\zeta)$. Hence for all $\varepsilon \in (0,\varepsilon_1] \cap \mathbb{Q}$, there exists $t_\varepsilon \in [0,\zeta)$ such that either $R_K(X_t) > \varepsilon$ for all $t \in [t_\varepsilon,\zeta)$ or $R_K(X_t) < 2\varepsilon$ for all $t \in [t_\varepsilon,\zeta)$. If there is $\varepsilon \in \mathbb{Q} \cap (0,\varepsilon_1]$ such that $R_K(X_t) > \varepsilon$ for all $t \in [t_\varepsilon,\zeta)$, then $\liminf_{t\to\zeta-} R_K(X_t) \geq \varepsilon > 0$. If for all $\varepsilon \in \mathbb{Q} \cap (0,\varepsilon_1]$, we have $R_K(X_t) < 2\varepsilon$ for all $t \in [t_\varepsilon,\zeta)$, then $\lim_{t\to\zeta-} R_K(X_t) = 0$.

Hence in any case, we have either

$$\lim_{t\to\zeta-} R_K(X_t) = 0 \quad \text{or} \quad \liminf_{t\to\zeta-} R_K(X_t) > 0. \qquad \blacksquare$$

Proof of Proposition 16. Point (i), which concerns the subcritical case, has already been checked in Lemma 17 (i). As for (ii), which concerns the supercritical case $\theta \geq 2$, we already know that quasi-everywhere, $\mathbb{P}_x^X(\zeta < \infty) = 1$ by Lemma 17 (ii), and it remains to prove that $\mathbb{P}_x^X$-a.s., $\lim_{t\to\zeta-} X_t$ exists and does not belong to E_{k_0}. We divide the proof into four steps.

Step 1. For a partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1,N]\!]$ and $\varepsilon \in (0,1]$ we consider, as in Proposition 15

$$G_{\mathbf{K},\varepsilon} = \Big\{x \in \mathcal{X} : \min_{1\le p\neq q\le \ell} \min_{i\in K_p, j\in K_q} \|x^i - x^j\|^2 > \varepsilon\Big\} \cap B(0, 1/\varepsilon)$$

and $\tau_{\mathbf{K},\varepsilon} = \inf\{t \ge 0 : X_t \notin G_{\mathbf{K},\varepsilon}\} \in [0,\zeta]$. We show here that for each $T > 0$, q.e. in $G_{\mathbf{K},\varepsilon}$, $\mathbb{P}_x^X$-a.s., for all $T > 0$ all $p \in [\![1,\ell]\!]$, $S_{K_p}(X_t)$ has a limit in $\mathbb{R}^2$ as $t \to (\tau_{\mathbf{K},\varepsilon} \wedge T)^-$.

If $\ell = 1$, the result is obvious since $S_{[\![1,N]\!]}(X_t)$ is a Brownian motion during $[0,\zeta)$ by Lemma 11. If $\ell \ge 2$, Proposition 15 and Lemma 11 tell us that under $\mathbb{Q}_x^{T,\varepsilon,\mathbf{K}}$, which is equivalent to $\mathbb{P}_x^X$, the processes $S_{K_p}(X_t)$ are Brownian motions on $[0, \tau_{\mathbf{K},\varepsilon} \wedge T)$, and thus have limits as $t \to (\tau_{\mathbf{K},\varepsilon} \wedge T)-$.

Step 2. For $\varepsilon \in (0,1]$ and a partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1,N]\!]$, we set $\tilde{\eta}_0^{\mathbf{K},\varepsilon} = 0$, and for $k \ge 0$,

$$\eta_{k+1}^{\mathbf{K},\varepsilon} = \inf\{t \ge \tilde{\eta}_k^{\mathbf{K},\varepsilon} : X_t \in G_{\mathbf{K},2\varepsilon}\} \quad \text{and} \quad \tilde{\eta}_{k+1}^{\mathbf{K},\varepsilon} = \inf\{t \ge \eta_{k+1}^{\mathbf{K},\varepsilon} : X_t \notin G_{\mathbf{K},\varepsilon}\},$$

with the convention that $\inf \emptyset = \zeta$. Using Step 1 and the strong Markov property, we conclude that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $\varepsilon \in (0,1] \cap \mathbb{Q}$, and all $k \ge 1$ and $T \in \mathbb{N}_+$, on $\{\eta_k^{\mathbf{K},\varepsilon} < \zeta\}$, for all $p \in [\![1,\ell]\!]$, $S_{K_p}(X_t)$ admits a limit in $\mathbb{R}^2$ as $t \to (\tilde{\eta}_k^{\mathbf{K},\varepsilon} \wedge T)-$. Choosing $T = \lceil \zeta \rceil$, we conclude that quasi-everywhere, $\mathbb{P}_x^X$-a.s., on $\{\eta_k^{\mathbf{K},\varepsilon} < \zeta\}$, for all $\varepsilon \in (0,1] \cap \mathbb{Q}$, and all $k \ge 1$ and $p \in [\![1,\ell]\!]$,

$$S_{K_p}(X_t) \text{ has a limit in } \mathbb{R}^2 \text{ as } t \to \tilde{\eta}_k^{\mathbf{K},\varepsilon}-.$$

Step 3. We now check that q.e., $\mathbb{P}_x^X$-a.s., there is a partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1,N]\!]$, some $\varepsilon \in (0,1] \cap \mathbb{Q}$ and some $k \ge 1$ such that (i) $\eta_k^{\mathbf{K},\varepsilon} < \zeta$ and $\tilde{\eta}_k^{\mathbf{K},\varepsilon} = \zeta$ and (ii) $\lim_{t\to\zeta-} R_{K_p}(X_t) = 0$ for all $p \in [\![1,\ell]\!]$.

By Lemma 18, we know that for all $K \subset [\![1,N]\!]$, we have either $\lim_{t\to\zeta-} R_K(X_t) = 0$ or $\liminf_{t\to\zeta-} R_K(X_t) > 0$. Hence the partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ of $[\![1,N]\!]$ consisting of the classes of the equivalence relation defined by $i \sim j$ if and only if $\lim_{t\to\zeta} R_{\{i,j\}}(X_t) = 0$ satisfies, for all $p \in [\![1,\ell]\!]$,

$$\lim_{t\to\zeta-} R_{K_p}(X_t) = 0 \quad \text{and} \quad \liminf_{t\to\zeta-} \min_{i\notin K_p} R_{K_p\cup\{i\}}(X_t) > 0.$$

Since moreover $\limsup_{t\to\zeta-} \|X_t\| < \infty$ according to Lemma 17, we deduce that there are $\alpha \in (0,\zeta)$ and $\varepsilon \in (0,1] \cap \mathbb{Q}$ such that $X_t \in G_{\mathbf{K},2\varepsilon}$ for all $t \in [\alpha,\zeta)$. Finally, we consider $k = \max\{m \ge 1 : \eta_m^{\mathbf{K},\varepsilon} \le \alpha\}$, which is finite by continuity of $t \mapsto X_t$ on $[0,\alpha]$, and we have $\eta_k^{\mathbf{K},\varepsilon} \le \alpha < \zeta$ and $\tilde{\eta}_k^{\mathbf{K},\varepsilon} = \zeta$.

Step 4. We consider the (random) partition $\mathbf{K} = (K_p)_{p\in[\![1,\ell]\!]}$ introduced in Step 3. By Step 2 and since $\eta_k^{\mathbf{K},\varepsilon} < \zeta$ and $\tilde{\eta}_k^{\mathbf{K},\varepsilon} = \zeta$, we know that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $p \in [\![1,\ell]\!]$, $M_p = \lim_{t\to\zeta-} S_{K_p}(X_t)$ exists in $\mathbb{R}^2$. By Step 3, we know that

$\lim_{t\to\zeta} R_{K_p}(X_t) = 0$ for all $p \in [\![1, \ell]\!]$. We easily conclude that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $p \in [\![1, \ell]\!]$ and $i \in K_p$, $\lim_{t\to\zeta-} X_t^i = M_p$. This shows that quasi-everywhere, $\mathbb{P}_x^X$-a.s., $X_{\zeta-} = \lim_{t\to\zeta-} X_t$ exists in $(\mathbb{R}^2)^N$. Moreover, $X_{\zeta-}$ cannot belong to $\mathcal{X} = E_{k_0}$, because $\lim_{t\to\zeta-} X_t = \triangle$ when $E_{k_0} \cup \{\triangle\}$ is endowed with the one-point compactification topology (see Section B.1). ■

9. Some special cases

During a K-collision, the particles labeled in K are isolated from the other ones. Thanks to Proposition 15, it will thus be possible to describe what happens in a neighborhood of the instant of this K-collision, by studying a $QKS(\theta|K|/N, |K|)$-process. In other words, we may assume that $|K| = N$, so that the following special cases, which are the purpose of this section, will be crucial.

Proposition 19. *Let $N \geq 4$ and $\theta > 0$ be such that $N > \theta$. Consider a $QKS(\theta, N)$-process $\mathbb{X}$ as in Proposition 6. Recall that $\zeta = \inf\{t \geq 0 : X_t = \triangle\}$ and introduce the stopping time $\tau = \inf\{t \geq 0 : R_{[\![1,N]\!]}(X_t) \notin (0, \infty)\}$ with the convention that $R_K(\triangle) = 0$, so that $\tau \in [0, \zeta]$.*

(i) *If $d_{\theta,N}(N-1) \leq 0$ and $d_{\theta,N}(N) < 2$, then quasi-everywhere,*

$$\mathbb{P}_x^X\Big(\inf_{t\in[0,\zeta)} R_{[\![1,N]\!]}(X_t) > 0\Big) = 1.$$

(ii) *If $d_{\theta,N}(N-1) \in (0, 2)$ and $d_{\theta,N}(N) < 2$, then quasi-everywhere, $\mathbb{P}_x^X$-a.s, for all $K \subset [\![1, N]\!]$ with $|K| = N - 1$, there is $t \in [0, \tau)$ such that $R_K(X_t) = 0$.*

(iii) *If $0 < d_{\theta,N}(N) < 2 \leq d_{\theta,N}(N-1)$, then quasi-everywhere, $\mathbb{P}_x^X$-a.s, for all $K \subset [\![1, N]\!]$ with $|K| = 2$, there is $t \in [0, \tau)$ such that $R_K(X_t) = 0$.*

The proof of this proposition is very long. First, we recall some notation about the decomposition of $\mathbb{X}$ obtained in Proposition 10 and we study the time-change involved. We then derive a formula describing $R_K(U_t)$, valid on certain time intervals, for any $K \subset [\![1, N]\!]$. This formula is of course not closed, but it allows us to compare $R_K(U_t)$, when it is close to 0, to some process resembling a squared Bessel process, whose behavior near 0 is easily established. Finally, we prove Proposition 19, unifying points (i) and (ii) somewhat and treating point (iii) separately.

9.1. Notation and preliminaries

We recall the decomposition of Proposition 10, which holds true quasi-everywhere in $\mathcal{X} \cap E_N$. Consider a Brownian motion $(M_t)_{t\geq 0}$ with diffusion coefficient $N^{-1/2}$ starting from $S_{[\![1,N]\!]}(x)$, a squared Bessel process $(D_t)_{t\geq 0}$ starting from $R_{[\![1,N]\!]}(x) > 0$ killed when leaving $(0, \infty)$ with life-time $\tau_D = \inf\{t \geq 0 : D_t = \triangle\}$, and a $QSKS(\theta, N)$-process $(U_t)_{t\geq 0}$ starting from $\Phi_{\mathbb{S}}(x)$ with life-time $\xi = \inf\{t \geq 0 : U_t = \triangle\}$, all these processes being independent. For $t \in [0, \tau_D)$, we put $A_t = \int_0^t \frac{\mathrm{d}s}{D_s}$. We also consider the inverse $\rho : [0, A_{\tau_D}) \to [0, \tau_D)$ of A.

Lemma 20. *If $d_{\theta,N}(N) < 2$, then $\tau_D < \infty$ and $A_{\tau_D} = \infty$ a.s.*

Proof. Since $(D_t)_{t\geq 0}$ is a (killed) squared Bessel process of dimension $d_{\theta,N}(N) < 2$, we have $\tau_D < \infty$ a.s according to Revuz–Yor [21, Chapter XI]. Moreover, there is a Brownian motion $(B_t)_{t\geq 0}$ such that $D_t = r + 2\int_0^t \sqrt{D_s}\,\mathrm{d}B_s + d_{\theta,N}(N)t$ for all $t \in [0, \tau_D)$, where $r = R_{[\![1,N]\!]}(x) > 0$. A simple computation shows the existence of a Brownian motion $(W_t)_{t\geq 0}$ such that for all $t \in [0, A_{\tau_D})$,

$$D_{\rho_t} = r + 2\int_0^t D_{\rho_s}\,\mathrm{d}W_s + d_{\theta,N}(N)\int_0^t D_{\rho_s}\,\mathrm{d}s.$$

Hence $D_{\rho_t} = r\exp(2W_t + (d_{\theta,N}(N) - 2)t)$ for all $t \in [0, A_{\tau_D})$. On the event $A_{\tau_D} < \infty$, we have

$$0 = D_{\tau_D -} = \lim_{t \to A_{\tau_D}} D_{\rho_t} = \exp\big(2W_{A_{\tau_D}} + (d_{\theta,N}(N) - 2)A_{\tau_D}\big) > 0.$$

Hence $A_{\tau_D} = \infty$ a.s. ∎

From now on, we assume that $d_{\theta,N}(N) < 2$. Hence $A : [0, \tau_D) \to [0, \infty)$ is an increasing bijection, as also is $\rho : [0, \infty) \to [0, \tau_D)$. By Proposition 10, quasi-everywhere in $\mathcal{X} \cap E_N$, we can find a triple $(M_t, D_t, U_t)_{t\geq 0}$ as above such that for our $QKS(\theta, N)$-process $\mathbb{X}$ starting from x, for all $t \in [0, \tau_D \wedge \rho_\xi)$, and actually for all $t \in [0, \rho_\xi)$ because $\rho_\xi \leq \tau_D$ since ρ is $[0, \tau_D)$-valued,

$$X_t = \Psi(M_t, D_t, U_{A_t}), \ \text{ i.e. } \ M_t = S_{[\![1,N]\!]}(X_t), \ \ D_t = R_{[\![1,N]\!]}(X_t) \ \text{ and } \ U_{A_t} = \Phi_{\mathbb{S}}(X_t).$$

We recall that $\Psi(m, r, u) = \gamma(m) + \sqrt{r}u$ if $(m, r, u) \in \mathbb{R}^2 \times (0, \infty) \times \mathcal{U}$, and $\Psi(m, r, u) = \triangle$ if $(m, r, u) = \triangle$. Observe that $\tau = \tau_D \wedge \rho_\xi = \rho_\xi$, where

$$\tau = \inf\{t \geq 0 : R_{[\![1,N]\!]}(X_t) \notin (0, \infty)\} \in [0, \zeta].$$

We note that if $\xi < \infty$, then $\rho_\xi < \tau_D$, because ρ is an increasing bijection from $[0, \infty)$ into $[0, \tau_D)$. Hence, still if $\xi < \infty$, then X explodes at time ρ_ξ strictly before τ_D, whence

$$\{\xi < \infty\} \subset \Big\{\inf_{t\in[0,\zeta)} R_{[\![1,N]\!]}(X_t) > 0\Big\}. \tag{49}$$

Finally, note that since U is $\mathbb{S}$-valued, it cannot have a $[\![1, N]\!]$-collision. But for any $K \subset [\![1, N]\!]$ with $|K| \leq N - 1$,

$$U \text{ has a } K\text{-collision at } t \in [0, \xi) \text{ if and only if } X \text{ has a } K\text{-collision at } \rho_t \in [0, \tau), \tag{50}$$

which follows from the facts that

- for all $(m, r, u) \in \mathbb{R}^2 \times (0, \infty) \times \mathcal{U}$, $R_K(\Psi(m, r, u)) = 0$ if and only if $R_K(u) = 0$;
- ρ is an increasing bijection from $[0, \xi)$ into $[0, \tau)$, because $\rho_\xi = \tau$.

We conclude this subsection with a remark about the quasi-everywhere notions of $\mathbb{X}$ and $\mathbb{U}$, in the case where they are related as above. See Section B.1 for a short reminder on this notion.

Remark 21. Fix $B \in \mathcal{M}^U$ such that $\mathbb{P}_u^U(B) = 1$ quasi-everywhere (here q.e. refers to the Hunt process $\mathbb{U}$). Then $\mathbb{P}^U_{\Phi_{\mathbb{S}}(x)}(B) = 1$ quasi-everywhere (here q.e. refers to the Hunt process $\mathbb{X}^*$, which is $\mathbb{X}$ killed when it leaves E_N).

Proof. By definition, there exists a properly exceptional set $\mathcal{N}^U$ relative to $\mathbb{U}$ such that $\mathbb{P}_u^U(B) = 1$ for all $u \in \mathcal{U} \setminus \mathcal{N}^U$. Thus $\mathbb{P}^U_{\Phi_{\mathbb{S}}(x)}(B) = 1$ for all $x \in \Phi_{\mathbb{S}}^{-1}(\mathcal{U} \setminus \mathcal{N}^U)$.

By Proposition 10, there exists a properly exceptional $\mathcal{N}^X$ set relative to $\mathbb{X}^*$ such that for all $x \in (\mathcal{X} \cap E_N) \setminus \mathcal{N}^X$, the law of $(X_t)_{t\geq 0}$ under $\mathbb{P}_x^X$ is equal to the law of $(Y_t = \Psi(M_t, D_t, U_{A_t}))_{t\geq 0}$ under $\mathbb{Q}_x^Y = \mathbb{P}^M_{\pi_{H^\perp}(x)} \otimes \mathbb{P}^D_{\|\pi_H(x)\|^2} \otimes \mathbb{P}^U_{\Phi_{\mathbb{S}}(x)}$, with obvious notation.

Hence we only have to prove that $\mathcal{N} = \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U) \cup \mathcal{N}^X$ is properly exceptional for $\mathbb{X}^*$.

First, $\mathbb{P}_x^X(X_t^* \notin \mathcal{N}$ for all $t \geq 0) = 1$ for all $x \in \mathcal{X} \setminus \mathcal{N}$. Indeed, since $x \in \mathcal{X} \setminus \mathcal{N}$, the law of $(X_t^*)_{t\geq 0}$ under $\mathbb{P}_x^X$ equals the law of $(Y_t)_{t\geq 0}$ under $\mathbb{Q}_x^Y$. Since $\mathbb{P}_u^U(U_t \notin \mathcal{N}^U$ for all $t \geq 0) = 1$ for all $u \in \mathcal{U} \setminus \mathcal{N}^U$, and since $\Phi_{\mathbb{S}}(Y_t) = U_{A_t}$, we have $\mathbb{Q}_x^Y(Y_t \notin \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U)$ for all $t \geq 0) = 1$ for all $x \in \mathcal{X} \setminus \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U)$. Consequently, $\mathbb{P}_x^X(X_t^* \notin \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U)$ for all $t \geq 0) = 1$ for all $x \in \mathcal{X} \setminus (\Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U) \cup \mathcal{N}^X)$. Finally, $\mathbb{P}_x^X(X_t^* \notin \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U) \cup \mathcal{N}^X$ for all $t \geq 0) = 1$ for all $x \in \mathcal{X} \setminus (\Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U) \cup \mathcal{N}^X)$ because $\mathcal{N}^X$ is properly exceptional for $\mathbb{X}^*$.

We have $\mu(\mathcal{N}) = 0$. Indeed, $\mu(\mathcal{N}^X) = 0$ by definition, and using Lemma A.2,

$$\begin{aligned}\mu(\Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U)) &= \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+^*\times\mathbb{S}} \mathbf{1}_{\{\Psi(z,r,u)\in\Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U)\}} r^\nu \,\mathrm{d}z\,\mathrm{d}r\,\beta(\mathrm{d}u) \\ &= \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+^*} \beta(\mathcal{N}^U) r^\nu \,\mathrm{d}z\,\mathrm{d}r = 0,\end{aligned}$$

since $\beta(\mathcal{N}^U) = 0$. We have used the fact that $\Psi(z, r, u) \in \Phi_{\mathbb{S}}^{-1}(\mathcal{N}^U) \Leftrightarrow u \in \mathcal{N}^U$, since $\Phi_{\mathbb{S}}(\Psi(z, r, u)) = u$. ∎

9.2. *An expression of dispersion processes on the sphere*

We now study the dispersion process $(R_K(U_t))_{t\geq 0}$ for $K \subset [\![1, N]\!]$. The equation below can be informally established assuming that (1) rigorously holds true, after a time-change and several Itô computations.

Lemma 22. *Fix $N \geq 2$ and $\theta > 0$ such that $N > \theta$ and recall that $k_0 = \lceil 2N/\theta \rceil$. Consider a QSKS$(\theta, N)$-process $\mathbb{U}$ with life-time ξ, fix $K \subset [\![1, N]\!]$ such that $|K| \geq 2$, and set $\mathbf{K} = (K, K^c)$. Recall that $G_{\mathbf{K},\varepsilon}$ was introduced in Lemma 12, and observe that*

$$G_{\mathbf{K},0} \cap \mathbb{S} = \Big\{u \in \mathcal{U} : \min_{i\in K, j\notin K} \|u^i - u^j\| > 0\Big\}.$$

Quasi-everywhere in $G_{\mathbf{K},0} \cap \mathbb{S}$*, enlarging the probability space* $(\Omega^U, \mathcal{M}^U, (\mathcal{M}^U_t)_{t\geq 0}, \mathbb{P}^U_u)$ *if necessary, there exists a* 1*-dimensional* $(\mathcal{M}^U_t)_{t\geq 0}$*-Brownian motion* $(W_t)_{t\geq 0}$ *under* $\mathbb{P}^U_u$ *such that*

$$R_K(U_t) = R_K(u) + 2\int_0^t \sqrt{R_K(U_s)(1-R_K(U_s))}\,\mathrm{d}W_s + d_{\theta,N}(|K|)t$$
$$- d_{\theta,N}(N)\int_0^t R_K(U_s)\,\mathrm{d}s - \frac{2\theta}{N}\sum_{i\in K,\, j\notin K}\int_0^t \frac{U^i_s - U^j_s}{\|U^i_s - U^j_s\|^2}\cdot (U^i_s - S_K(U_s))\,\mathrm{d}s \qquad (51)$$

for all $t \in [0, \kappa_K)$*, where* $\kappa_K = \inf\{t \geq 0 : U_t \notin G_{\mathbf{K},0}\}$.

As usual, $\kappa_K \leq \xi$ because $\triangle \notin G_{\mathbf{K},0}$. Note also that if $K = [\![1, N]\!]$, then $R_K(U_t) = 1$ for all $t \in [0, \xi)$, and that the constant process 1 indeed solves (51).

Proof of Lemma 22. We divide the proof into several steps. The main idea is to compute $\mathcal{L}^U R_K$ and $\mathcal{L}^U (R_K)^2$ and to use $R_K(U_t) = R_K(u) + \int_0^t \mathcal{L}^U R_K(U_s)\,\mathrm{d}s + M_t$ for some martingale $(M_t)_{t\geq 0}$ whose bracket we can compute. However, we need to regularize R_K and to localize space in a zone where the last term of (51) is bounded.

Step 1. We fix $n \geq 1$ and $\varepsilon \in (0, 1]$ and recall that $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon} \in C^\infty(\mathbb{S})$, compactly supported in $G_{\mathbf{K},0} \cap \mathbb{S}$, was defined in Lemma 12. We want to apply Remark 8 to $R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}$ and $(R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon})^2$. We thus have to show that $R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}$ and $(R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon})^2$ belong to $C^\infty_c(\mathcal{U})$ for all $n \geq 1$, which is clear, and that

$$\sup_{\alpha\in(0,1]}\sup_{u\in\mathbb{S}}\big(|\mathcal{L}^U_\alpha[R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}](u)| + |\mathcal{L}^U_\alpha[(R_K\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon})^2](u)|\big) < \infty$$

for all $n \geq 1$. Since

$$\mathcal{L}^U_\alpha(fg) = f\mathcal{L}^U_\alpha g + g\mathcal{L}^U_\alpha f + \nabla_{\mathbb{S}} f\cdot\nabla_{\mathbb{S}} g \qquad (52)$$

for all $f, g \in C^\infty(\mathbb{S})$ and recalling that $\sup_{\alpha\in(0,1]}\sup_{u\in\mathbb{S}}|\mathcal{L}^U_\alpha \Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}(u)| < \infty$ by Lemma 12, and $\Gamma^{\mathbb{S},n}_{\mathbf{K},\varepsilon}$ is compactly supported in $G_{\mathbf{K},0} \cap \mathbb{S}$, the only issue is to verify that, for A compact in $G_{\mathbf{K},0} \cap \mathbb{S}$,

$$\sup_{\alpha\in(0,1]}\sup_{u\in A}|\mathcal{L}^U_\alpha R_K(u)| < \infty. \qquad (53)$$

Step 2. Here we prove that

$$\mathcal{L}^U_\alpha R_K(u) = 2(|K|-1) - 2(N-1)R_K(u) + \frac{\theta}{N}R_K(u)\sum_{1\leq i,j\leq N}\frac{\|u^i-u^j\|^2}{\|u^i-u^j\|^2+\alpha}$$
$$- \frac{\theta}{N}\sum_{i\in K,\, j\in K}\frac{\|u^i-u^j\|^2}{\|u^i-u^j\|^2+\alpha} - \frac{2\theta}{N}\sum_{i\in K,\, j\notin K}\frac{u^i-u^j}{\|u^i-u^j\|^2+\alpha}\cdot(u^i - S_K(u)), \qquad (54)$$

and this will imply (53): the first four terms are obviously uniformly bounded on $\mathbb{S}$, and the last one is uniformly bounded on A (because A is compact in $G_{\mathbf{K},0} \cap \mathbb{S}$).

This will also imply, taking $\alpha = 0$ and observing that $2(|K|-1) - \frac{\theta}{N}|K|(|K|-1) = d_{\theta,N}(|K|)$ and $2(N-1) - \frac{\theta}{N}N(N-1) = d_{\theta,N}(N)$, that for all $u \in \mathbb{S} \cap E_2$,

$$\mathcal{L}^U R_K(u) = d_{\theta,N}(|K|) - d_{\theta,N}(N)R_K(u) - \frac{2\theta}{N}\sum_{i\in K,\, j\notin K}\frac{u^i - u^j}{\|u^i - u^j\|^2}\cdot(u^i - S_K(u)). \tag{55}$$

Step 2.1. We first verify that for all $u \in \mathbb{S}$,

$$(\nabla_{\mathbb{S}} R_K(u))^i = 2(u^i - S_K(u))\mathbf{1}_{\{i\in K\}} - 2R_K(u)u^i, \quad i \in [\![1, N]\!], \tag{56}$$

$$\Delta_{\mathbb{S}} R_K(u) = 4(|K|-1) - 4(N-1)R_K(u). \tag{57}$$

First, a simple computation shows that for $x \in (\mathbb{R}^2)^N$ and $i \in [\![1, N]\!]$,

$$\nabla_{x^i} R_K(x) = 2(x^i - S_K(x))\mathbf{1}_{\{i\in K\}} \quad \text{and} \quad \Delta_{x^i} R_K(x) = \frac{4(|K|-1)}{|K|}\mathbf{1}_{\{i\in K\}}, \tag{58}$$

so that in particular $\nabla R_K(x) \in H$ and

$$\begin{aligned}\nabla R_K(x)\cdot x &= 2\sum_{i\in K}(x^i - S_K(x))\cdot x^i \\ &= 2\sum_{i\in K}(x^i - S_K(x))\cdot(x^i - S_K(x)) = 2R_K(x).\end{aligned} \tag{59}$$

Next, proceeding as in (14), we find that for all $x \in E_N$,

$$\begin{aligned}\nabla[R_K \circ \Phi_{\mathbb{S}}](x) &= \|\pi_H(x)\|^{-1}\pi_H(\pi_{(\pi_H(x))^\perp}(\nabla R_K(\Phi_{\mathbb{S}}(x)))) \\ &= \frac{\pi_H\left(\nabla R_K(\Phi_{\mathbb{S}}(x)) - \frac{\pi_H(x)\cdot\nabla R_K(\Phi_{\mathbb{S}}(x))}{\|\pi_H(x)\|^2}\pi_H(x)\right)}{\|\pi_H(x)\|} \\ &= \frac{\nabla R_K(x) - 2R_K(x)\frac{\pi_H(x)}{\|\pi_H(x)\|^2}}{\|\pi_H(x)\|^2}.\end{aligned}$$

We have used the facts that $\nabla R_K(\Phi_{\mathbb{S}}(x)) = \nabla R_K(x)/\|\pi_H(x)\|$ thanks to (58), $\nabla R_K(x) \in H$ by (58), and $\pi_H(x)\cdot\nabla R_K(x) = x\cdot\nabla R_K(x) = 2R_K(x)$ by (59).

We first conclude that for $u \in \mathbb{S}$, since $\pi_H(u) = u$ and $\|u\| = 1$,

$$\nabla_{\mathbb{S}} R_K(u) = \nabla[R_K \circ \Phi_{\mathbb{S}}](u) = \nabla R_K(u) - 2R_K(u)u, \tag{60}$$

which implies (56) by (58).

Second, we deduce that for $x \in E_N$,

$$\begin{aligned}&\Delta[R_K \circ \Phi_{\mathbb{S}}](x) \\ &= \frac{1}{\|\pi_H(x)\|^2}\left(\Delta R_K(x) - 2\nabla R_K(x)\cdot\frac{\pi_H(x)}{\|\pi_H(x)\|^2} - 2R_K(x)\frac{\operatorname{div}\pi_H(x)}{\|\pi_H(x)\|^2} + \frac{4R_K(x)}{\|\pi_H(x)\|^2}\right) \\ &\quad - \frac{2\pi_H(x)}{\|\pi_H(x)\|^4}\cdot\left(\nabla R_K(x) - 2R_K(x)\frac{\pi_H(x)}{\|\pi_H(x)\|^2}\right).\end{aligned}$$

Using $\operatorname{div} \pi_H(x) = 2(N-1)$, we conclude that for $u \in \mathbb{S}$, since $\pi_H(u) = u$, $\|u\| = 1$ and $u \cdot \nabla R_K(u) = 2R_K(u)$ by (59), we have

$$\Delta_{\mathbb{S}} R_K(u) = \Delta[R_K \circ \Phi_{\mathbb{S}}](u) = \Delta R_K(u) - 4R_K(u) - 4(N-1)R_K(u) + 4R_K(u).$$

Since finally $\Delta R_K(u) = 4(|K|-1)$ by (58), this leads to (57).

Step 2.2. We fix $u \in \mathbb{S}$ and set

$$I_\alpha(u) = -\frac{\theta}{N} \sum_{1 \le i,j \le N} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (\nabla_{\mathbb{S}} R_K(u))^i.$$

Here we show that

$$I_\alpha(u) = -\frac{\theta}{N} \sum_{i,j \in K} \frac{\|u^i - u^j\|^2}{\|u^i - u^j\|^2 + \alpha} + \frac{\theta}{N} R_K(u) \sum_{1 \le i,j \le N} \frac{\|u^i - u^j\|^2}{\|u^i - u^j\|^2 + \alpha}$$
$$- \frac{2\theta}{N} \sum_{i \in K,\, j \notin K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u)). \tag{61}$$

By (56), we may write $I_\alpha = I_{1,\alpha} + I_{2,\alpha}$, where

$$I_{1,\alpha}(u) = -\frac{2\theta}{N} \sum_{i \in K,\, j \in [\![1,N]\!]} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u)),$$
$$I_{2,\alpha}(u) = \frac{2\theta}{N} R_K(u) \sum_{1 \le i,j \le N} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot u^i.$$

First, by symmetry,

$$I_{1,\alpha}(u) = -\frac{2\theta}{N} \sum_{i,j \in K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u))$$
$$- \frac{2\theta}{N} \sum_{i \in K, j \notin K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u))$$
$$= -\frac{2\theta}{N} \sum_{i,j \in K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot u^i - \frac{2\theta}{N} \sum_{i \in K,\, j \notin K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u))$$
$$= -\frac{\theta}{N} \sum_{i,j \in K} \frac{\|u^i - u^j\|^2}{\|u^i - u^j\|^2 + \alpha} - \frac{2\theta}{N} \sum_{i \in K,\, j \notin K} \frac{u^i - u^j}{\|u^i - u^j\|^2 + \alpha} \cdot (u^i - S_K(u)).$$

Second, by symmetry,

$$I_{2,\alpha}(u) = \frac{\theta}{N} R_K(u) \sum_{1 \le i,j \le N} \frac{\|u^i - u^j\|^2}{\|u^i - u^j\|^2 + \alpha}.$$

Step 2.3. Since $\mathcal{L}_\alpha^U R_K(u) = \frac{1}{2} \Delta_{\mathbb{S}} R_K(u) + I_\alpha(u)$, (54) follows from (57) and (61).

Step 3. By Steps 1 and 2, we can apply Remark 8 and Lemma B.2: quasi-everywhere, for all $n \geq 1$, there exist $(\mathcal{M}_t^U)_{t\geq 0}$-martingales $(M_t^{1,n,\varepsilon})_{t\geq 0}$ and $(M_t^{2,n,\varepsilon})_{t\geq 0}$ under $\mathbb{P}_u^U$ such that

$$(R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})(U_t) = (R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})(u) + M_t^{1,n,\varepsilon} + \int_0^t \mathcal{L}^U(R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})(U_s)\,\mathrm{d}s,$$

$$(R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})^2(U_t) = (R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})^2(u) + M_t^{2,n,\varepsilon} + \int_0^t \mathcal{L}^U(R_K\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n})^2(U_s)\,\mathrm{d}s,$$

for all $t \geq 0$. We recall that $\kappa_K = \inf\{t \geq 0 : U_t \notin G_{\mathbf{K},0}^n\}$ and introduce

$$\kappa_{K,n,\varepsilon} = \inf\{t \geq 0 : U_t \notin G_{\mathbf{K},\varepsilon}^n\} \wedge \kappa_K.$$

Since $\bigcup_{n\geq 1} G_{\mathbf{K},\varepsilon}^n \supset G_{\mathbf{K},\varepsilon}$ and $G_{\mathbf{K},\varepsilon}$ increases to $G_{\mathbf{K},0}$ as $\varepsilon \to 0$ (see Lemma 12), we conclude that $\lim_{\varepsilon\to 0}\lim_{n\to\infty}\kappa_{K,n,\varepsilon} = \kappa_K$. Next, since $\Gamma_{\mathbf{K},\varepsilon}^{\mathbb{S},n} = 1$ on $G_{\mathbf{K},\varepsilon}^n \cap \mathbb{S}$, we have, for all $t \in [0, \kappa_{K,n,\varepsilon}]$,

$$R_K(U_t) = R_K(u) + M_t^{1,n,\varepsilon} + \int_0^t \mathcal{L}^U R_K(U_s)\,\mathrm{d}s, \tag{62}$$

$$(R_K(U_t))^2 = (R_K(u))^2 + M_t^{2,n,\varepsilon} + \int_0^t \mathcal{L}^U(R_K^2)(U_s)\,\mathrm{d}s. \tag{63}$$

Applying the Itô formula to compute $(R_K(U_t))^2$ from (62), recalling from (52) that $\mathcal{L}^U(R_K^2) = 2R_K\mathcal{L}^U R_K + \|\nabla_{\mathbb{S}} R_K\|^2$ and comparing to (63), we see that for $t \in [0, \kappa_{K,n,\varepsilon}]$,

$$\langle M^{1,n,\varepsilon}\rangle_t = \int_0^t \|\nabla_{\mathbb{S}} R_K(U_s)\|^2\,\mathrm{d}s.$$

Hence, enlarging the probability space if necessary, we can find a Brownian motion $(W_t)_{t\geq 0}$, which is defined by $W_t = \int_0^t \|\nabla_{\mathbb{S}} R_K(U_s)\|^{-1}\,\mathrm{d}M_s^{1,n,\varepsilon}$ for $t \in [0, \kappa_{K,n,\varepsilon}]$ and then extended to $\mathbb{R}_+$, such that $M_t^{1,n,\varepsilon} = \int_0^t \|\nabla_{\mathbb{S}} R_K(U_s)\|\,\mathrm{d}W_s$ during $[0, \kappa_{K,n,\varepsilon}]$. Hence, still for $t \in [0, \kappa_{K,n,\varepsilon}]$,

$$R_K(U_t) = R_K(u) + \int_0^t \|\nabla_{\mathbb{S}} R_K(U_s)\|\,\mathrm{d}W_s + \int_0^t \mathcal{L}^U R_K(U_s)\,\mathrm{d}s. \tag{64}$$

But $\nabla_{\mathbb{S}} R_K(u) = \nabla R_K(u) - 2R_K(u)u$ by (60), whence

$$\|\nabla_{\mathbb{S}} R_K(u)\|^2 = \|\nabla R_K(u)\|^2 - 4R_K(u)\nabla R_K(u)\cdot u + 4(R_K(u))^2.$$

Since $\|\nabla R_K(u)\|^2 = 4R_K(u)$ by (58) and $\nabla R_K(u)\cdot u = 2R_K(u)$ by (59), we have

$$\|\nabla_{\mathbb{S}} R_K(u)\|^2 = 4R_K(u) - 4(R_K(u))^2 = 4R_K(u)(1 - R_K(u)).$$

Inserting this, as well as the expression (55) of $\mathcal{L}^U R_K$, in (64) shows that $R_K(U_t)$ satisfies the desired equation on $[0, \kappa_{K,n,\varepsilon}]$. Since $\lim_{\varepsilon\to 0}\lim_{n\to\infty}\kappa_{K,n,\varepsilon} = \kappa_K$ a.s., the proof is complete. ∎

9.3. A squared Bessel-like process

The equation obtained in the previous lemma will be studied by comparison with the process we now introduce. This process behaves, near 0, like a squared Bessel process.

Lemma 23. *Fix $\delta \in \mathbb{R}$ and $a, b > 0$ such that $\delta + a\sqrt{b} < 2$. For a 1-dimensional Brownian motion $(W_t)_{t\geq 0}$ and $x \in [0, 1)$, consider the unique solution $(S_t)_{t\geq 0}$ of*

$$S_t = x + \int_0^t 2\sqrt{|S_s(1 - S_s)|}\,\mathrm{d}W_s + \delta t + a\int_0^t \sqrt{b + |S_s|}\,\mathrm{d}s. \tag{65}$$

For $z \in \mathbb{R}$, set $\tau_z = \inf\{t > 0 : S_t = z\}$. For all $y \in (x, 1)$, we have $\mathbb{P}(\tau_0 < \tau_y) > 0$.

Proof. This equation is classically well-posed: the diffusion coefficient is 1/2-Hölder continuous and the drift coefficient is Lipschitz continuous (see Revuz–Yor [21, Theorem 3.5, p. 390]). As in Karatzas–Shreve [15, (5.42), p. 339], we introduce the scale function

$$f(z) = \int_{1/2}^z \exp\Big(-\int_{1/2}^u \frac{\delta + a\sqrt{b + |v|}}{2|v(1 - v)|}\,\mathrm{d}v\Big)\,\mathrm{d}u.$$

This function is obviously continuous on (0, 1), and for example approximating $(\delta + a\sqrt{b + |v|})/(2|v(1 - v)|)$ by $(\delta + a\sqrt{b})/(2|v|)$, we see that it is also continuous at 0 because $\delta + a\sqrt{b} < 2$. By [15, (5.61), p. 344], we have

$$\mathbb{P}(\tau_0 < \tau_y) = \frac{f(y) - f(x)}{f(y) - f(0)} \tag{66}$$

for all $y \in (x, 1)$. This last quantity is nonzero (which would not be the case if $\delta + a\sqrt{b} \geq 2$, since then $f(0) = -\infty$). ■

9.4. Collisions of large clusters

Proof of Proposition 19 (i, ii). We fix $N \geq 4$ and $\theta > 0$ such that $N > \theta$. We always assume that $d_{\theta,N}(N) < 2$ and we use the notation of Section 9.1.

Step 1. We consider $\varepsilon \in (0, 1]$ and $K \subset [\![1, N]\!]$ such that $|K| \in [\![2, N - 1]\!]$ and $d_{\theta,N}(|K|) < 2$. We introduce the constant $a_K = c_{|K|+1}/(2c_{|K|})$ with $(c_\ell)_{\ell\in[\![1,N]\!]}$ defined in Lemma 13. We prove in this step that there are constants $p_{K,\varepsilon}, T_{K,\varepsilon} > 0$ such that, if we set

$$\tilde\sigma^{K,\varepsilon} = \inf\Big\{t > 0 : R_K(U_t) \geq \varepsilon \text{ or } \min_{i\notin K} R_{K\cup\{i\}}(U_t) \leq a_K\varepsilon\Big\} \wedge T_{K,\varepsilon},$$

with the convention that $\inf\emptyset = \xi$, then q.e. on $\{u \in \mathcal{U} : R_K(u) \leq \varepsilon/2\}$,

$$\mathbb{P}_u^U\Big(\tilde\sigma^{K,\varepsilon} = \xi \text{ or } \inf_{t\in[0,\tilde\sigma^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t) \leq 2a_K\varepsilon \\ \text{or } R_K(U_t) = 0 \text{ for some } t \in [0, \tilde\sigma^{K,\varepsilon})\Big) \geq p_{K,\varepsilon}.$$

We introduce $Z_{K,\varepsilon} = \inf_{t\in[0,\tilde{\sigma}^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t)$. For all $t \in [0, \tilde{\sigma}^{K,\varepsilon})$, we have $R_K(U_t) \leq \varepsilon$ and $Z_{K,\varepsilon} \geq a_K\varepsilon$ so that $\min_{i\in K,\, j\notin K} \|U_t^i - U_t^j\| \geq \varepsilon/2$ thanks to the definition of a_K and Lemma 13. This implies that $\tilde{\sigma}^{K,\varepsilon} \leq \kappa_K$, where $\kappa_K = \inf\{t \geq 0 : U_t \notin G_{\mathbf{K},\mathbf{0}}\}$ was defined in Lemma 22, and $G_{\mathbf{K},\mathbf{0}} \cap \mathbb{S} = \{u \in \mathcal{U} : \min_{i\in K,\, j\notin K} \|u^i - u^j\| > 0\}$.

By the Cauchy–Schwarz inequality, and since R_K is bounded on $\mathcal{U}$, there exists a deterministic constant $C_{K,\varepsilon} > 0$, allowed to change from line to line, such that for all $t \in [0, \tilde{\sigma}^{K,\varepsilon})$,

$$
\begin{aligned}
- d_{\theta,N}(N) R_K(U_t) - \frac{2\theta}{N} \sum_{i\in K, j\notin K} \frac{U_t^i - U_t^j}{\|U_t^i - U_t^j\|^2} \cdot (U_t^i - S_K(U_t)) \\
\leq C_{K,\varepsilon}\sqrt{R_K(U_t)} + C_{K,\varepsilon}\Big(\sum_{i\in K} \|U_t^i - S_K(U_t)\|^2\Big)^{1/2} \\
\leq C_{K,\varepsilon}\sqrt{R_K(U_t)} \leq C_{K,\varepsilon}\sqrt{b + R_K(U_t)},
\end{aligned}
$$

where $b > 0$ is chosen small enough that $d_{\theta,N}(|K|) + C_{K,\varepsilon}\sqrt{b} < 2$. Actually, b is only introduced to make the drift coefficient of (65) Lipschitz continuous.

Recalling that $R_K(U_0) \leq \varepsilon/2$ and the formula describing $R_K(U_t) \in [0,1]$ for $t \in [0, \kappa_K) \supset [0, \tilde{\sigma}^{K,\varepsilon})$ in Lemma 22, considering the process $(S_t)_{t\geq 0}$ solving (65) with $x = \varepsilon/2$, $\delta = d_{\theta,N}(|K|)$, $a = C_{K,\varepsilon}$ and with b introduced a few lines above, driven by the same Brownian motion $(W_t)_{t\geq 0}$, and using the comparison theorem, we find that $R_K(U_t) \leq S_t$ for all $t \in [0, \tilde{\sigma}^{K,\varepsilon})$.

Setting $\tau_z = \inf\{t \geq 0 : S_t = z\}$ for $z \in \mathbb{R}$ and recalling the definition of $\tilde{\sigma}^{K,\varepsilon}$, we conclude that $\{Z_{K,\varepsilon} > 2a_K\varepsilon\} \subset \{\tilde{\sigma}^{K,\varepsilon} \geq \tau_\varepsilon \wedge T_{K,\varepsilon}\}$. Indeed, note that on the event that $\inf_{t\in[0,\tilde{\sigma}^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t) > 2a_K\varepsilon$, either $\tilde{\sigma}^{K,\varepsilon} = T_{K,\varepsilon}$, or $(R_K(U_t))_{t\geq 0}$ reaches ε at time $\tilde{\sigma}^{K,\varepsilon}$ and then $\tau_\varepsilon \leq \tilde{\sigma}^{K,\varepsilon}$. In both cases, $\tilde{\sigma}^{K,\varepsilon} \geq \tau_\varepsilon \wedge T_{K,\varepsilon}$. Hence, using again $R_K(U_t) \leq S_t$ for all $t \in [0, \tilde{\sigma}^{K,\varepsilon})$, we have

$$
\begin{aligned}
&\{\tilde{\sigma}^{K,\varepsilon} < \xi \text{ and } Z_{K,\varepsilon} > 2a_K\varepsilon \text{ and } S_t = 0 \text{ for some } t \in [0, \tau_\varepsilon \wedge T_{K,\varepsilon}]\} \\
&\qquad \subset \{\tilde{\sigma}^{K,\varepsilon} < \xi \text{ and } Z_{K,\varepsilon} > 2a_K\varepsilon \text{ and } R_K(U_t) = 0 \text{ for some } t \in [0, \tilde{\sigma}^{K,\varepsilon})\}.
\end{aligned}
$$

But $A^c \cap B' \subset A^c \cap B$ gives $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(A^c \cap B) \geq \mathbb{P}(A) + \mathbb{P}(A^c \cap B') = \mathbb{P}(A \cup B')$. Hence

$$
\begin{aligned}
&\mathbb{P}_u^U\big(\tilde{\sigma}^{K,\varepsilon} = \xi \text{ or } Z_{K,\varepsilon} \leq 2a_K\varepsilon \text{ or } R_K(U_t) = 0 \text{ for some } t \in [0, \tilde{\sigma}^{K,\varepsilon})\big) \\
&\qquad \geq \mathbb{P}_u^U\big(\tilde{\sigma}^{K,\varepsilon} = \xi \text{ or } Z_{K,\varepsilon} \leq 2a_K\varepsilon \text{ or } S_t = 0 \text{ for some } t \in [0, \tau_\varepsilon \wedge T_{K,\varepsilon})\big) \\
&\qquad \geq \mathbb{P}_u^U(S_t = 0 \text{ for some } t \in [0, \tau_\varepsilon \wedge T_{K,\varepsilon})).
\end{aligned}
$$

This last quantity equals $\mathbb{P}(\tau_0 < \tau_\varepsilon \wedge T_{K,\varepsilon})$ and does not depend on u such that $R_K(u) \leq \varepsilon/2$. But $\mathbb{P}(\tau_0 < \tau_\varepsilon) > 0$ by Lemma 23 and since $d_{\theta,N}(|K|) + C_{K,\varepsilon}\sqrt{b} < 2$. Hence there exists $T_{K,\varepsilon} > 0$ with $\mathbb{P}(\tau_0 < \tau_\varepsilon \wedge T_{K,\varepsilon}) > 0$, and this finishes this step.

Step 2. We prove (ii), i.e. if $d_{\theta,N}(N-1) \in (0,2)$, then for any $K \subset [\![1, N]\!]$ with $|K| = N - 1$, quasi-everywhere, $\mathbb{P}_x^X$-a.s., $R_K(X_t)$ vanishes during $[0, \zeta)$. By (50) and

Remark 21, and since $\mathbb{P}_u^U(\xi = \infty) = 1$ quasi-everywhere by Lemma 9 (ii), it suffices to check that quasi-everywhere, $\mathbb{P}_u^U$-a.s., $(R_K(U_t))_{t\ge 0}$ vanishes at least once during $[0,\infty)$.

We fix $K \subset [\![1,N]\!]$ with $|K| = N-1$, set $\varepsilon_0 = 1/(4a_K)$ and introduce $\tilde\tau_0^K = 0$ and for all $k \ge 0$,

$$\tau_{k+1}^K = \inf\{t \ge \tilde\tau_k^K : R_K(U_t) \le \varepsilon_0/2\},$$
$$\tilde\tau_{k+1}^K = \inf\{t \ge \tau_{k+1}^K : R_K(U_t) \ge \varepsilon_0\} \wedge (\tau_{k+1}^K + T_{K,\varepsilon_0}),$$

with T_{K,ε_0} defined in Step 1. All these stopping times are finite since $(\mathcal{E}^U, \mathcal{F}^U)$ is recurrent by Lemma 9 (ii). We also put, for $k \ge 1$,

$$\Omega_k^K = \{R_K(U_t) = 0 \text{ for some } t \in [\tau_k^K, \tilde\tau_k^K]\}.$$

We now prove that $\mathbb{P}_u^U(\bigcap_{k\ge 1}(\Omega_k^K)^c) = 0$ quasi-everywhere, and this will complete the proof of (ii).

For $\ell \ge 1$, since $\bigcap_{k=1}^{\ell}(\Omega_k^K)^c$ is $\mathcal{M}^U_{\tau^K_{\ell+1}}$-measurable, the strong Markov property tells us that

$$\mathbb{P}_u^U\Big(\bigcap_{k=1}^{\ell+1}(\Omega_k^K)^c\Big) = \mathbb{E}_u^U\Big[\Big(\prod_{k=1}^{\ell}\mathbf{1}_{(\Omega_k^K)^c}\Big)\mathbb{P}^U_{U_{\tau^K_{\ell+1}}}((\Omega_1^K)^c)\Big].$$

We now prove that $\mathbb{P}_u^U(\Omega_1^K) \ge p_{K,\varepsilon_0}$ quasi-everywhere on $\{u \in \mathcal{U} : R_K(u) \le \varepsilon_0/2\}$. For such a u, we have $\tau_1^K = 0$. Moreover, for all $i \notin K$, we have $R_{K\cup\{i\}}(u) = R_{[\![1,N]\!]}(u) = 1 > 2a_K\varepsilon_0$ thanks to our choice of ε_0. Hence $\tilde\tau_1^K = \tilde\sigma^{K,\varepsilon_0}$ (recall Step 1). Since finally $\tilde\sigma^{K,\varepsilon_0} < \infty = \xi$ and since $R_{K\cup\{i\}}(U_t) = R_{[\![1,N]\!]}(U_t) = 1 > 2a_K\varepsilon_0$ for all $t \ge 0$ and all $i \notin K$, we have

$$\begin{aligned}\Omega_1^K &= \{R_K(U_t) = 0 \text{ for some } t \in [0,\tilde\sigma^{K,\varepsilon_0}]\}\\ &= \Big\{\tilde\sigma^{K,\varepsilon_0} = \xi \text{ or } \inf_{t\in[0,\tilde\sigma^{K,\varepsilon_0})}\min_{i\notin K} R_{K\cup\{i\}}(U_t) \le 2a_K\varepsilon_0 \\ &\qquad \text{or } R_K(U_t) = 0 \text{ for some } t \in [0,\tilde\sigma^{K,\varepsilon_0}]\Big\}.\end{aligned}$$

Hence Step 1 tells us that $\mathbb{P}_u^U(\Omega_1^K) \ge p_{K,\varepsilon_0}$ quasi-everywhere on the set $\{u \in \mathcal{U} : R_K(u) \le \varepsilon_0/2\}$.

Since $R_K(U_{\tau^K_{\ell+1}}) \le \varepsilon_0/2$, we have proved that for all $\ell \ge 1$,

$$\mathbb{P}_u^U\Big(\bigcap_{k=1}^{\ell+1}(\Omega_k^K)^c\Big) \le (1-p_{K,\varepsilon_0})\mathbb{P}_u^U\Big(\bigcap_{k=1}^{\ell}(\Omega_k^K)^c\Big).$$

This allows us to conclude that indeed $\mathbb{P}_u^U(\bigcap_{k=1}^{\infty}(\Omega_k^K)^c) = 0$.

Step 3. We prove (i), i.e. if $d_{\theta,N}(N-1) \le 0$, then $\mathbb{P}_x^X(\inf_{[0,\zeta)} R_{[\![1,N]\!]}(X_t) > 0) = 1$ quasi-everywhere. By Remark 21 and (49), it suffices to show that quasi-everywhere, $\mathbb{P}_u^U(\xi < \infty) = 1$.

For all $K \subset [\![1, N]\!]$ and $\varepsilon \in (0, 1]$, we introduce $\tilde{\sigma}_0^{K,\varepsilon} = 0$ and for all $k \geq 0$,

$$\sigma_{k+1}^{K,\varepsilon} = \inf\Big\{t \geq \tilde{\sigma}_k^{K,\varepsilon} : R_K(U_t) \leq \varepsilon/2 \text{ and } \min_{i \notin K} R_{K\cup\{i\}}(U_t) \geq 2a_K\varepsilon\Big\},$$

$$\tilde{\sigma}_{k+1}^{K,\varepsilon} = \inf\Big\{t \geq \sigma_{k+1}^{K,\varepsilon} : R_K(U_t) \geq \varepsilon \text{ or } \min_{i \notin K} R_{K\cup\{i\}}(U_t) \leq a_K\varepsilon\Big\} \wedge (\sigma_{k+1}^{K,\varepsilon} + T_{K,\varepsilon}),$$

with $T_{K,\varepsilon}$ defined in Step 1 and with the convention that $\inf \emptyset = \xi$.

Step 3.1. We fix $\varepsilon \in (0, 1]$ and assume that $|K| \geq k_0$, so that $d_{\theta,N}(|K|) \leq 0$ by Lemma 1. We prove here that quasi-everywhere, $\mathbb{P}_u^U$-a.s., either there is $t \in [0, \xi)$ such that $R_K(U_t) = 0$ or there is $k \geq 1$ such that either $\sigma_{k+1}^{K,\varepsilon} = \xi$ or there is $k \geq 1$ such that that $\inf_{t\in[\sigma_k^{K,\varepsilon},\tilde{\sigma}_k^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t) \leq 2a_K\varepsilon$.

It suffices to prove that $\mathbb{P}_u^U(\bigcap_{k\geq 1}(\Omega_k^{K,\varepsilon})^c) = 0$, where

$$\Omega_k^{K,\varepsilon} = \Big\{\sigma_{k+1}^{K,\varepsilon} = \xi \text{ or } \inf_{t\in[\sigma_k^{K,\varepsilon},\tilde{\sigma}_k^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t) \leq 2a_K\varepsilon \\ \text{ or } R_K(U_t) = 0 \text{ for some } t \in [\sigma_k^{K,\varepsilon}, \tilde{\sigma}_k^{K,\varepsilon})\Big\}.$$

But for all $\ell \geq 1$, $\bigcap_{k=1}^{\ell}(\Omega_k^{K,\varepsilon})^c$ is $\mathcal{M}^U_{\sigma_{\ell+1}^{K,\varepsilon}}$-measurable, whence, by the strong Markov property,

$$\mathbb{P}_u^U\Big(\bigcap_{k=1}^{\ell+1}(\Omega_k^{K,\varepsilon})^c\Big) = \mathbb{E}_u^U\Big[\Big(\prod_{k=1}^{\ell}\mathbf{1}_{(\Omega_k^{K,\varepsilon})^c}\Big)\mathbb{P}^U_{U_{\sigma_{\ell+1}^{K,\varepsilon}}}((\Omega_1^{K,\varepsilon})^c)\Big]$$
$$\leq (1 - p_{K,\varepsilon})\mathbb{P}_u^U\Big(\bigcap_{k=1}^{\ell}(\Omega_k^{K,\varepsilon})^c\Big).$$

We have used Step 1, the fact that $R_K(U_{\sigma_{\ell+1}^{K,\varepsilon}}) \leq \varepsilon/2$ on the event $(\Omega_\ell^{K,\varepsilon})^c \subset \{\sigma_{\ell+1}^{K,\varepsilon} < \xi\}$, as well as the inclusion $\{\tilde{\sigma}_k^{K,\varepsilon} = \xi\} \subset \{\sigma_{k+1}^{K,\varepsilon} = \xi\}$. One easily concludes the proof.

Step 3.2. For all $K \subset [\![1, N]\!]$ such that $|K| \geq k_0$, quasi-everywhere, $\mathbb{P}_u^U$-a.s., there is no $t \in [0, \xi)$ such that $R_K(U_t) = 0$. Indeed, on the contrary event, there is $t \in [0, \xi)$ such that $U_t \notin E_{k_0}$, whence $U_t \notin \mathcal{U}$, which contradicts the fact that $t \in [0, \xi)$.

Step 3.3. We show by decreasing induction that for every $n \in [\![k_0, N]\!]$,

$\mathcal{P}(n)$: q.e., $\mathbb{P}_u^U$-a.s. on the event $\{\xi = \infty\}$, $b_n = \min_{\{|K|=n\}} \inf_{t\geq 0} R_K(U_t) > 0$.

The result is clear when $n = N$, because for all $t \in [0, \xi)$, $R_{[\![1,N]\!]}(U_t) = 1$.

We next assume $\mathcal{P}(n)$ for some $n \in [\![k_0 + 1, N]\!]$ and we show that $\mathcal{P}(n - 1)$ is true. We fix $K \subset [\![1, N]\!]$ with $|K| = n - 1$ and we apply Step 3.1 with K and some $\varepsilon \in (0, b_n/(4a_K))$ (b_n is random but we may apply Step 3.1 simultaneously for all $\varepsilon \in \mathbb{Q}_+^* \cap (0, 1]$) and Step 3.2. We find that on the event $\{\xi = \infty\}$, either there exists $k \geq 1$ such that $\sigma_{k+1}^{K,\varepsilon} = \infty$ or there exists $k \geq 1$ such that $\inf_{t\in[\sigma_k^{K,\varepsilon},\tilde{\sigma}_k^{K,\varepsilon})} \min_{i\notin K} R_{K\cup\{i\}}(U_t) \leq 2a_K\varepsilon$. This second choice is not possible, since by induction assumption, $R_{K\cup\{i\}}(U_t) \geq b_n$ for all $t > 0$ and all $i \notin K$. Hence there is $k \geq 1$ such that $\sigma_{k+1}^{K,\varepsilon} = \infty$.

By definition of $\sigma_{k+1}^{K,\varepsilon}$, this implies that, still on the event $\{\xi = \infty\}$, there exists $t_0 \geq 0$ such that for all $t \geq t_0$, either $R_K(U_t) \geq \varepsilon/2$ or $\min_{i\in K} R_{K\cup\{i\}}(U_t) \leq 2a_K\varepsilon$. Using again the induction assumption, we see that the second choice is never possible, so that actually $R_K(U_t) \geq \varepsilon/2$ for all $t \geq t_0$. Since $(R_K(U_t))_{t\geq 0}$ is continuous and positive on $[0, t_0]$ according to Step 3.2, this finishes this step.

Step 3.4. We conclude from Step 3.3 that quasi-everywhere, $\mathbb{P}_u^U$-a.s. on the event $\{\xi = \infty\}$, $U_t \in \mathcal{K}$ for all $t \geq 0$, where

$$\mathcal{K} = \{u \in \mathcal{U} : R_K(u) \geq b_n \text{ for all } n \in [\![k_0, N]\!] \text{ and all } K \subset [\![1, N]\!] \text{ with } |K| = n\}.$$

This (random) set is compact in $\mathcal{U}$, so that Lemma 9 (i) tells us, both when $(\mathcal{E}^U, \mathcal{F}^U)$ is recurrent and when $(\mathcal{E}^U, \mathcal{F}^U)$ is transient, that this happens with probability 0. Hence quasi-everywhere, $\mathbb{P}_u^U(\xi = \infty) = 0$ as desired. ∎

9.5. Binary collisions

Proof of Proposition 19 (iii). Assume that $N \geq 4$ and $0 < d_{\theta,N}(N) < 2 \leq d_{\theta,N}(N-1)$ and observe that $\theta < 2$ and $k_0 > N$, so that $\mathcal{X} = (\mathbb{R}^2)^N$ and $\mathcal{U} = \mathbb{S}$. The $QKS(\theta, N)$-process $\mathbb{X}$ is nonexploding by Proposition 16 (i), and the $QSKS(\theta, N)$-process $\mathbb{U}$ is irreducible recurrent by Lemma 9 (ii). In particular, $\zeta = \xi = \infty$ a.s. We divide the proof into four steps. First, we prove that $\mathbb{X}$ may have some binary collisions with positive probability. Then we check that this implies that $\mathbb{U}$ may also have some binary collisions with positive probability. Since $\mathbb{U}$ is recurrent, it will then necessarily be a.s. subjected to (infinitely many) binary collisions. Finally, we conclude the proof using (50).

Step 1. We set $\mathbf{K} = (\{1, 2\}, \{3\}, \dots, \{N\})$ and

$$\mathcal{K} = \Big\{x \in B(0, C) : \|x^1 - x^2\| < 1 \text{ and } \min_{i\in[\![1,N]\!],\, j\in[\![3,N]\!],\, i\neq j} \|x^i - x^j\| > 10\Big\},$$

with C large enough so that $\mu(\mathcal{K}) > 0$. We show in this step that $\mathbb{P}_x^X(A) > 0$ quasi-everywhere in $\mathcal{K}$, where

$$A = \Big\{X_t^1 = X_t^2 \text{ for some } t \in [0, 1] \text{ and } \min_{t\in[0,1]} R_{[\![1,N]\!]}(X_t) > 0\Big\}.$$

To this end, we fix $x \in \mathcal{K}$ and introduce the set

$$O = \Big\{y \in (\mathbb{R}^2)^2 : R_{\{1,2\}}(y) < 2,\ \Big\|\frac{y^1 + y^2}{2} - \frac{x^1 + x^2}{2}\Big\| < 1\Big\},$$

and $B_i = \{y \in \mathbb{R}^2 : \|y - x^i\|^2 < 1\}$ for $i \in [\![3, N]\!]$. Clearly, there is some $\varepsilon \in (0, 1]$ such that

$$L = \{y \in (\mathbb{R}^2)^N : (y^1, y^2) \in O \text{ and } y^i \in B_i \text{ for all } i \in [\![3, N]\!]\} \subset G_{\mathbf{K},\varepsilon},$$

where as usual $G_{\mathbf{K},\varepsilon} = \{y \in B(0, 1/\varepsilon) : \forall i \in [\![1, N]\!],\ \forall j \in [\![3, N]\!] \setminus \{i\},\ \|y^i - y^j\|^2 > \varepsilon\}$; recall that $\mathcal{X} = (\mathbb{R}^2)^N$ because $k_0 > N$.

Since $G_{\mathbf{K},\varepsilon}$ is obviously included in $\{y \in (\mathbb{R}^2)^N : R_{[\![1,N]\!]}(y) > 0\}$, we conclude that

$$\begin{aligned}\mathbb{P}^X_x(A) &\geq \mathbb{P}^X_x(X^1_t = X^2_t \text{ for some } t \in [0,1] \text{ and } X_t \in L \text{ for all } t \in [0,1]) \\ &\geq C^{-1}_{1,\varepsilon,\mathbf{K}} \mathbb{Q}^{1,\varepsilon,\mathbf{K}}_x(X^1_t = X^2_t \text{ for some } t \in [0,1] \text{ and } X_t \in L \text{ for all } t \in [0,1])\end{aligned}$$

by Proposition 15 with $T = 1$. We now set $\tau_{\mathbf{K},\varepsilon} = \inf\{t > 0 : X_t \notin G_{\mathbf{K},\varepsilon}\}$. Proposition 15 tells us that, quasi-everywhere in $\mathcal{K} \subset G_{\mathbf{K},\varepsilon}$, the law of $(X_t)_{t\in[0,\tau_{\mathbf{K},\varepsilon}]}$ under $\mathbb{Q}^{1,\varepsilon,\mathbf{K}}_x$ equals the law of $Y_t = (Y^1_t, \ldots, Y^N_t)_{t\in[0,\tilde{\tau}_{\mathbf{K},\varepsilon}]}$ where $(Y^1_t, Y^2_t)_{t\geq 0}$ is a $QKS(2\theta/N, 2)$-process issued from (x^1, x^2); for all $i \in [\![3, N]\!]$, $(Y^i_t)_{t\geq 0}$ is a $QKS(\theta/N, 1)$-process, i.e. a 2-dimensional Brownian motion, issued from x^i; and all these processes are independent. We have set $\tilde{\tau}_{\mathbf{K},\varepsilon} = \inf\{t > 0 : Y_t \notin G_{\mathbf{K},\varepsilon}\}$. This implies, together with the fact that $\{X_t \in L \text{ for all } t \in [0,1]\} \subset \{\tau_{\mathbf{K},\varepsilon} > 1\}$, that

$$\mathbb{P}^X_x(A) \geq C^{-1}_{1,\varepsilon,\mathbf{K}}\, p \prod_{i=3}^N q_i$$

quasi-everywhere in $\mathcal{K}$, where

$$p = \mathbb{P}\Big(\min_{s\in[0,1]} R_{\{1,2\}}((Y^1_s, Y^2_s)) = 0 \text{ and } (Y^1_t, Y^2_t) \in O \text{ for all } t \in [0,1]\Big),$$

and $q_i = \mathbb{P}(Y^i_t \in B_i$ for all $t \in [0,1])$. Of course, $q_i > 0$ for all $i \in [\![3, N]\!]$, since $(Y^i_t)_{t\geq 0}$ is a Brownian motion issued from x^i. Moreover, we know from Lemma 11 that $(M_t = (Y^1_t + Y^2_t)/2)_{t\geq 0}$ is a 2-dimensional Brownian motion with diffusion coefficient $2^{-1/2}$ issued from $m = (x^1 + x^2)/2$, $(R_t = R_{\{1,2\}}((Y^1_t, Y^2_t)))_{t\geq 0}$ is a squared Bessel process of dimension $d_{2\theta/N,2}(2) = d_{\theta,N}(2)$ issued from $r = \|x^1 - x^2\|^2/2 \in (0, 1/2)$, and these processes are independent. Hence, recalling the definition of O,

$$p = \mathbb{P}\Big(\min_{s\in[0,1]} R_s = 0 \text{ and } \max_{s\in[0,1]} R_s < 2\Big)\mathbb{P}\Big(\max_{s\in[0,1]} \|M_t - m\| < 1\Big).$$

This last quantity is clearly positive, because a squared Bessel process of dimension $d_{\theta,N}(2) \in (0, 2)$ (see Lemma 1) does hit zero (see Revuz–Yor [21, Chapter XI]).

Step 2. We now deduce from Step 1 that the set $F = \{u \in \mathcal{U} : u^1 = u^2\}$ is not exceptional for $\mathbb{U}$. Indeed, if it was exceptional, we would have $\mathbb{P}^U_u(\exists t \geq 0 : U_t \in F) = 0$ q.e. By (50) and Remark 21, this would imply that q.e., $\mathbb{P}^X_x(\exists t \in [0,\tau) : X_t \in G) = 0$, where $G = \{x \in \mathcal{X} : x^1 = x^2\}$ and $\tau = \inf\{t > 0 : R_{[\![1,N]\!]}(X_t) = 0\}$. But on the event A defined in Step 1, there is $t \in [0,1]$ such that $X_t \in G$ and $\tau > 1$. In conclusion, we have $\mathbb{P}^X_x(\exists t \in [0,\tau) : X_t \in G) > 0$ q.e. in $\mathcal{K}$, which is a contradiction, since $\mu(\mathcal{K}) > 0$.

Step 3. Since $(\mathcal{E}^U, \mathcal{F}^U)$ is irreducible-recurrent and F is not exceptional, we know from Fukushima–Oshima–Takeda [11, Theorem 4.7.1 (iii), p. 202] that quasi-everywhere,

$$\mathbb{P}^U_u(\forall r > 0, \exists t \geq r : U_t \in F) = 1.$$

Step 4. Using again (50) and Remark 21 and recalling that $\xi = \infty$ and ρ is an increasing bijection from $[0,\infty)$ to $[0,\tau)$, we conclude that quasi-everywhere, $\mathbb{P}^X_x$-a.s., X_t visits F (an infinite number of times) during $[0,\tau)$. Of course, the same arguments apply on replacing $\{1, 2\}$ by any 2-element subset of $[\![1, N]\!]$, and the proof is complete. ∎

10. Quasi-everywhere conclusion

Here we prove that the conclusions of Theorem 5 hold quasi-everywhere.

Partial proof of Theorem 5. We assume that $\theta \geq 2$ and $N > 3\theta$, so that $k_0 = \lceil 2N/\theta \rceil \in [\![7, N]\!]$, and consider an $\mathcal{X}_\triangle$-valued $QKS(\theta, N)$-process $\mathbb{X}$ with life-time ζ as in Proposition 6, where $\mathcal{X} = E_{k_0}$.

Preliminaries. For $K \subset [\![1, N]\!]$ and $\varepsilon \in (0, 1]$, we use the simplified notation $\tau_{K,\varepsilon} = \inf\{t > 0 : X_t \notin G_{K,\varepsilon}\} \in [0, \zeta]$ and $G_{K,\varepsilon} = \{x \in \mathcal{X} : \min_{i\in K, j\notin K} \|x^i - x^j\|^2 > \varepsilon\} \cap B(0, 1/\varepsilon)$ instead of $\tau_{\mathbf{K},\varepsilon}$ and $G_{\mathbf{K},\varepsilon}$ with $\mathbf{K} = (K, K^c)$ as in Proposition 15. We also write $\mathbb{Q}_x^{T,\varepsilon,K}$ instead of $\mathbb{Q}^{T,\varepsilon,\mathbf{K}}$ and recall that it is equivalent to $\mathbb{P}_x^X$ on $\mathcal{M}_T^X = \sigma(X_s : s \in [0, T])$.

Setting $X_t^K = (X_t^i)_{i\in K}$ and $X_t^{K^c} = (X_t^i)_{i\in K^c}$, we know that q.e. in $G_{K,\varepsilon}$, the law of $(X_t^K, X_t^{K^c})_{t\in[0,\tau_{K,\varepsilon}\wedge T]}$ under $\mathbb{Q}_x^{T,\varepsilon,K}$ is the same as the law of $(Y_t, Z_t)_{t\in[0,\tilde{\tau}_{K,\varepsilon}\wedge T]}$, where $(Y_t)_{t\geq 0}$ is a $QKS(|K|\theta/N, |K|)$-process issued from $x|_K$, $(Z_t)_{t\geq 0}$ is a $QKS(|K^c|\theta/N, |K^c|)$-process issued from $x|_{K^c}$, these two processes being independent, and $\tilde{\tau}_{K,\varepsilon} = \inf\{t > 0 : (Y_t, Z_t) \notin G_{K,\varepsilon}\}$. We denote by ζ^Y and ζ^Z the life-times of $(Y_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$. The life-time of $(Y_t, Z_t)_{t\geq 0}$ is given by $\zeta' = \zeta^Y \wedge \zeta^Z$ and we have $\tilde{\tau}_{K,\varepsilon} \in [0, \zeta']$.

No isolated points. Here we prove that for all $K \subset [\![1, N]\!]$ with $d_{\theta,N}(|K|) \in (0, 2)$, quasi-everywhere, we have $\mathbb{P}_x^X(A_K) = 0$, where $A_K = \{\mathcal{Z}_K$ has an isolated point$\}$ and

$$\mathcal{Z}_K = \{t \in (0, \zeta) : \text{there is a } K\text{-collision in the configuration } X_t\}.$$

On A_K, we can find $u, v \in \mathbb{Q}_+$ such that $u < v < \zeta$ and there is a unique $t \in (u, v)$ with $R_K(X_t) = 0$ and $\min_{i\notin K} R_{K\cup\{i\}}(X_t) > 0$. By continuity, we deduce that on A_K, there exist $r, s \in \mathbb{Q}_+$ and $\varepsilon \in \mathbb{Q} \cap (0, 1]$ such that $r < s < \zeta$, $X_t \in G_{K,\varepsilon}$ for all $t \in [r, s]$, and $\{t \in (r, s) : R_K(X_t) = 0\}$ has an isolated point. It thus suffices that for all $r < s$ and $\varepsilon \in (0, 1]$, fixed from now on, quasi-everywhere, $\mathbb{P}_x^X(A_{K,r,s,\varepsilon}) = 0$, where

$$A_{K,r,s,\varepsilon} = \Big\{X_t \in G_{K,\varepsilon} \text{ for all } t \in (r, s) \text{ and } \{t \in (r, s) : R_K(X_t) = 0\} \text{ has an isolated point}\Big\}.$$

By the Markov property, it suffices that $\mathbb{P}_x^X(A_{K,0,s,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$ and, by equivalence, that $\mathbb{Q}_x^{s,\varepsilon,K}(A_{K,0,s,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$. We write, recalling the preliminaries,

$$\begin{aligned}\mathbb{Q}_x^{s,\varepsilon,K}(A_{K,0,s,\varepsilon}) &= \mathbb{Q}_x^{s,\varepsilon,K}\big(\tau_{K,\varepsilon} \geq s \text{ and } \{t \in (0, s) : R_K(X_t) = 0\} \text{ has an isolated point}\big)\\ &= \mathbb{P}\big(\tilde{\tau}_{K,\varepsilon} \geq s \text{ and } \{t \in (0, s) : R_K(Y_t) = 0\} \text{ has an isolated point}\big)\\ &\leq \mathbb{P}\big(\{t \in (0, s) : R_K(Y_t) = 0\} \text{ has an isolated point}\big).\end{aligned}$$

But $(Y_t)_{t\geq 0}$ is a $QKS(|K|\theta/N, |K|)$-process, so that by Lemma 11, $(R_K(Y_t))_{t\geq 0}$ is a squared Bessel process of dimension $d_{|K|\theta/N,|K|}(|K|) = d_{\theta,N}(|K|) \in (0, 2)$. Such a process has no isolated zero (see Revuz–Yor [21, Chapter XI]).

Proof of (i). We have already seen in Proposition 16 (ii) that quasi-everywhere, $\mathbb{P}_x^X$-a.s., $\zeta < \infty$ and $X_{\zeta-} = \lim_{t\to\zeta-} X_t$ exists in $(\mathbb{R}^2)^N$ and does not belong to E_{k_0}.

Proof of (ii). We want to show that quasi-everywhere, $\mathbb{P}_x^X$-a.s., there is $K_0 \subset [\![1, N]\!]$ with $|K_0| = k_0$ such that there is a K_0-collision and no K-collision with $|K| > k_0$ in the configuration $X_{\zeta-}$. We already know that $X_{\zeta-} \notin E_{k_0}$, so that there is $K \subset [\![1, N]\!]$ with $|K| \geq k_0$ such that there is a K-collision in $X_{\zeta-}$. Hence the goal is to verify that quasi-everywhere, for all $K \subset [\![1, N]\!]$ with $|K| > k_0$, $\mathbb{P}_x^X(B_K) = 0$, where

$$B_K = \{\text{there is a } K\text{-collision in the configuration } X_{\zeta-}\}.$$

On B_K, there is $\varepsilon \in \mathbb{Q} \cap (0, 1]$ such that $X_{\zeta-} \in G_{K,2\varepsilon}$. By continuity, there also exists, still on B_K, some $r \in \mathbb{Q}_+ \cap [0, \zeta)$ such that $X_t \in G_{K,\varepsilon}$ for all $t \in [r, \zeta)$. Hence we only have to prove that for all $\varepsilon \in \mathbb{Q} \cap (0, 1]$, all $t \in \mathbb{Q}_+$, and all $T \in \mathbb{Q}_+$ such that $T > r$, quasi-everywhere, $\mathbb{P}_x^X(B_{K,r,T,\varepsilon}) = 0$, where

$$B_{K,r,T,\varepsilon} = \{\zeta \in (r, T],\ X_t \in G_{K,\varepsilon} \text{ for all } t \in [r, \zeta) \text{ and } R_K(X_{\zeta-}) = 0\}.$$

By the Markov property, it suffices that $\mathbb{P}_x^X(B_{K,0,T,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$ for all $\varepsilon \in \mathbb{Q} \cap (0, 1]$ and $T \in \mathbb{Q}_+^*$. We now fix $\varepsilon \in \mathbb{Q} \cap (0, 1]$ and $T \in \mathbb{Q}_+^*$. By equivalence, it suffices to prove that $\mathbb{Q}_x^{T,\varepsilon,K}(B_{K,0,T,\varepsilon}) = 0$. Using the notation introduced in the preliminaries, we write

$$\begin{aligned}\mathbb{Q}_x^{T,\varepsilon,K}(B_{K,0,T,\varepsilon}) &= \mathbb{Q}_x^{T,\varepsilon,K}\big(\zeta \leq T,\ \tau_{K,\varepsilon} = \zeta \text{ and } R_K(X_{\zeta-}) = 0\big)\\ &= \mathbb{P}\big(\zeta' \leq T,\ \tilde{\tau}_{K,\varepsilon} = \zeta' \text{ and } R_K(Y_{\zeta'-}) = 0\big)\\ &\leq \mathbb{P}\Big(\inf_{t\in[0,\zeta^Y)} R_K(Y_t) = 0\Big).\end{aligned}$$

But $(Y_t)_{t\geq 0}$ is a $QKS(|K|\theta/N, |K|)$-process with $|K| > k_0 \geq 7$ and $d_{|K|\theta/N,|K|}(|K| - 1) = d_{\theta,N}(|K| - 1) \leq 0$ by Lemma 1 because $|K| - 1 \geq k_0$. Also $d_{|K|\theta/N,|K|}(|K|) = d_{\theta,N}(|K|) \leq 0$. Hence Proposition 19 (i) tells us that $\mathbb{P}(\inf_{t\in[0,\zeta^Y)} R_K(Y_t) = 0) = 0$.

Proof of (iii). We recall that $k_1 = k_0 - 1$ and we fix $L \subset K \subset [\![1, N]\!]$ with $|K| = k_0$ and $|L| = k_1$. We want to prove that quasi-everywhere, $\mathbb{P}_x^X$-a.s., if $R_K(X_{\zeta-}) = 0$, then for all $t \in [0, \zeta)$, the set $\mathcal{Z}_L \cap (t, \zeta)$ is infinite and has no isolated point. But since $d_{\theta,N}(k_1) \in (0, 2)$ (see Lemma 1), we already know that $\mathcal{Z}_L$ has no isolated point. It thus suffices to check that quasi-everywhere, for all $r \in \mathbb{Q}_+$, we have $\mathbb{P}_x^X(C_{K,L,r}) = 0$, where

$$C_{K,L,r} = \{\zeta > r,\ R_K(X_{\zeta-}) = 0, \text{ and } R_L(X_t) > 0 \text{ for all } t \in (r, \zeta)\}.$$

We have used the fact that since $|L| = k_1 = k_0 - 1$, for all $x \in \mathcal{X} = E_{k_0}$ there is an L-collision in the configuration x if and only if $R_L(x) = 0$.

On $C_{K,L,r}$, thanks to (ii) , there are $\varepsilon \in \mathbb{Q} \cap (0, 1]$, $T \in \mathbb{Q}_+$ and $s \in \mathbb{Q}_+^* \cap [r, \zeta)$ such that $\zeta \in (s, T]$ and $X_t \in G_{K,\varepsilon}$ for all $t \in [s, \zeta)$. Thus it suffices to prove that for all $s < T$ and $\varepsilon \in (0, 1]$, now fixed, quasi-everywhere, $\mathbb{P}_x^X(C_{K,L,s,T,\varepsilon}) = 0$, where

$$C_{K,L,s,T,\varepsilon} = \{\zeta \in (s, T],\ R_K(X_{\zeta-}) = 0,\ X_t \in G_{K,\varepsilon} \text{ and } R_L(X_t) > 0 \text{ for all } t \in [s, \zeta)\}.$$

By the Markov property, it suffices that $\mathbb{P}_x^X(C_{K,L,0,T,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$ and, by equivalence, that $\mathbb{Q}_x^{T,\varepsilon,K}(C_{K,L,0,T,\varepsilon}) = 0$. Recalling the preliminaries, we write

$$\begin{aligned}\mathbb{Q}_x^{T,\varepsilon,K}&(C_{K,L,0,T,\varepsilon})\\ &= \mathbb{Q}_x^{T,\varepsilon,K}\big(\zeta \le T,\ R_K(X_{\zeta-}) = 0,\ \tau_{K,\varepsilon} = \zeta \text{ and } R_L(X_t) > 0 \text{ for all } t \in [0,\zeta)\big)\\ &= \mathbb{P}\big(\zeta' \le T,\ R_K(Y_{\zeta'-}) = 0,\ \tilde\tau_{K,\varepsilon} = \zeta' \text{ and } R_L(Y_t) > 0 \text{ for all } t \in [0,\zeta')\big).\end{aligned}$$

Setting $\sigma_K = \inf\{t > 0 : R_K(Y_t) = 0\}$, we observe that $\sigma_K = \zeta^Y$. Indeed, $|K| = k_0$ and $(Y_t)_{t\ge0}$ is a $QKS(|K|\theta/N, |K|)$-process with state space $\mathcal{Y}_\triangle = \mathcal{Y} \cup \{\triangle\}$, where

$$\mathcal{Y} = \{y \in (\mathbb{R}^2)^{|K|} : R_M(y) > 0 \text{ for all } M \subset [\![1,N]\!] \text{ with } |M| \ge k_0\},$$

because $\lceil 2|K|/(|K|\theta/N)\rceil = \lceil 2N/\theta\rceil = k_0$. Hence $\{R_K(Y_{\zeta'-}) = 0\} \subset \{\zeta' = \sigma_K\}$, so that

$$\mathbb{Q}_x^{T,\varepsilon,K}(C_{K,L,0,T,\varepsilon}) \le \mathbb{P}\big(R_L(Y_t) > 0 \text{ for all } t \in [0,\sigma_K)\big).$$

This last quantity equals zero by Proposition 19 (ii), since we have $d_{|K|\theta/N,|K|}(|K|-1) = d_{\theta,N}(|K|-1) = d_{\theta,N}(k_0-1) \in (0,2)$ by Lemma 1 and $|L| = k_1 = |K|-1$ and since $d_{|K|\theta/N,|K|}(|K|) = d_{\theta,N}(|K|) = d_{\theta,N}(k_0) \le 0 < 2$.

Proof of (iv). We assume that $k_2 = k_0 - 2$, i.e. $d_{\theta,N}(k_0-2) \in (0,2)$. We fix $L \subset K \subset [\![1,N]\!]$ with $|K| = k_1$ and $|L| = k_2$. We want to prove that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $t \in [0,\zeta)$, if there is a K-collision in the configuration X_t, then for all $r \in [0,t)$, the set $\mathcal{Z}_L \cap (r,t)$ is infinite and has no isolated point. We already know that $\mathcal{Z}_L$ has no isolated point. It thus suffices to check that quasi-everywhere, for all $r \in \mathbb{Q}_+$, we have $\mathbb{P}_x^X(D_{K,L,r}) = 0$, where

$$D_{K,L,r} = \{\zeta > r \text{ and there is } t \in (r,\zeta) \text{ such that there is a } K\text{-collision at time } t \text{ but no } L\text{-collision during } (r,t)\}.$$

We set $\sigma_{K,r} = \inf\{t > r : \text{there is a } K\text{-collision in } X_t\}$. Then

$$D_{K,L,r} = \{\zeta > r,\ \sigma_{K,r} < \zeta \text{ and there is no } L\text{-collision during } u \in [r,\sigma_{K,r})\}.$$

On $D_{K,L,r}$, there exists $\varepsilon \in \mathbb{Q} \cap (0,1]$ such that $X_{\sigma_{K,r}} \in G_{K,2\varepsilon}$, so that by continuity, there exists $v \in \mathbb{Q}_+ \cap [r,\sigma_{K,r})$ such that $X_u \in G_{K,\varepsilon}$ for all $u \in [v,\sigma_{K,r}]$. Observe that $\sigma_{K,v} = \sigma_{K,r}$ and for all $t \in [v,\sigma_{K,v})$, there is an L-collision at time t if and only if $R_L(X_t) = 0$, by definition of $\sigma_{K,v}$ and since $X_t \in G_{K,\varepsilon}$. All in all, it suffices to prove that for all $v \in \mathbb{Q}_+$, and all $\varepsilon \in \mathbb{Q} \cap (0,1]$ and $T \in \mathbb{Q}_+^*$, $\mathbb{P}_x^X(D_{K,L,v,T,\varepsilon}) = 0$ quasi-everywhere, where

$$D_{K,L,v,T,\varepsilon} = \{\zeta \in (v,T],\ \sigma_{K,v} < \zeta,\ X_u \in G_{K,\varepsilon} \text{ and } R_L(X_u) > 0 \text{ for all } u \in [v,\sigma_{K,v})\}.$$

By the Markov property, it suffices to prove that $\mathbb{P}_x^X(D_{K,L,0,T,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$ and, by equivalence, we may use $\mathbb{Q}_x^{T,\varepsilon,K}$ instead of $\mathbb{P}_x^X$. But recalling the preliminaries,

$$\begin{aligned}\mathbb{Q}_x^{T,\varepsilon,K}(D_{K,L,0,T,\varepsilon})\\ &= \mathbb{Q}_x^{T,\varepsilon,K}\big(\zeta \le T,\ \sigma_{K,0} < \zeta,\ \tau_{K,\varepsilon} \ge \sigma_{K,0} \text{ and } R_L(X_t) > 0 \text{ for all } t \in [0,\sigma_{K,0})\big)\\ &= \mathbb{P}\big(\zeta' \le T,\ \tilde\sigma_{K,0} < \zeta',\ \tilde\tau_{K,\varepsilon} \ge \tilde\sigma_{K,0} \text{ and } R_L(Y_t) > 0 \text{ for all } t \in [0,\tilde\sigma_{K,0})\big)\\ &\le \mathbb{P}\big(R_L(Y_t) > 0 \text{ for all } t \in [0,\tilde\sigma_{K,0})\big),\end{aligned}$$

where $\tilde\sigma_{K,0} = \inf\{t > 0 : R_K(Y_t) = 0\}$. Finally, $\mathbb{P}(R_L(Y_t) > 0 \text{ for all } t \in [0,\tilde\sigma_{K,0})) = 0$ by Proposition 19 (ii), because $(Y_t)_{t\ge0}$ is a $QKS(|K|\theta/N, |K|)$-process and $|L| = k_2 = |K| - 1$ and $d_{|K|\theta/N,|K|}(|K| - 1) = d_{\theta,N}(|K| - 1) = d_{\theta,N}(k_2) \in (0, 2)$ and $d_{|K|\theta/N,|K|}(|K|) = d_{\theta,N}(|K|) = d_{\theta,N}(k_1) \in (0, 2)$.

Proof of (v). We fix $K \subset [\![1, N]\!]$ with $|K| \in [\![3, k_2 - 1]\!]$, so that $d_{\theta,N}(|K|) \ge 2$. We want to prove that quasi-everywhere, $\mathbb{P}_x^X$-a.s., for all $t \in [0, \zeta)$, there is no K-collision in the configuration X_t. We introduce $\sigma_K = \inf\{t > 0 : \text{there is a } K\text{-collision in } X_t\}$, with the convention that $\inf\emptyset = \zeta$, and we have to verify that quasi-everywhere, $\mathbb{P}_x^X(\sigma_K < \zeta) = 0$.

On the event $\{\sigma_K < \zeta\}$, there exist $\varepsilon \in \mathbb{Q} \cap (0, 1]$ and $r \in \mathbb{Q}_+^* \cap [0, \sigma_K)$ such that $X_t \in G_{K,\varepsilon}$ for all $t \in [r, \sigma_K]$. Hence it suffices to check that for all $\varepsilon \in \mathbb{Q} \cap (0, 1]$, all $r \in \mathbb{Q}_+^*$ and all $T \in \mathbb{Q}_+^* \cap (r,\infty)$, now fixed, quasi-everywhere, $\mathbb{P}_x^X(F_{K,r,T,\varepsilon}) = 0$, where

$$F_{K,r,T,\varepsilon} = \{\sigma_K \in (r, \zeta \wedge T) \text{ and } X_t \in G_{K,\varepsilon} \text{ for all } t \in [r, \sigma_K]\}.$$

By the Markov property, it suffices that $\mathbb{P}_x^X(F_{K,0,T,\varepsilon}) = 0$ quasi-everywhere in $G_{K,\varepsilon}$ and, by equivalence, that $\mathbb{Q}_x^{T,\varepsilon,K}(F_{K,0,T,\varepsilon}) = 0$. Recalling the preliminaries, we write

$$\begin{aligned}\mathbb{Q}_x^{T,\varepsilon,K}(F_{K,0,T,\varepsilon}) &= \mathbb{Q}_x^{T,\varepsilon,K}\big(\sigma_K \in (0, \zeta \wedge T) \text{ and } \tau_{K,\varepsilon} \ge \sigma_K\big)\\ &= \mathbb{P}\big(\tilde\sigma_K \in (0, \zeta' \wedge T) \text{ and } \tilde\tau_{K,\varepsilon} \ge \tilde\sigma_K\big)\\ &\le \mathbb{P}\Big(\inf_{t\in[0,T]} R_K(Y_t) = 0\Big),\end{aligned}$$

where we have set $\tilde\sigma_K = \inf\{t > 0 : \text{there is a } K\text{-collision in the configuration } (Y_t, Z_t)\}$. Since $(Y_t)_{t\ge0}$ is a $QKS(|K|\theta/N, |K|)$-process, we know from Lemma 11 that $(R_K(Y_t))_{t\ge0}$ is a squared Bessel process of dimension $d_{|K|\theta/N,|K|}(|K|) = d_{\theta,N}(|K|) \ge 2$. Such a process a.s. never reaches 0.

Proof of (vi). The proof is exactly the same as that of (iv), replacing everywhere k_1 by k_2 and k_2 by 2, and using Proposition 19 (iii) instead of Proposition 19 (ii), which is licit because $0 < d_{k_2\theta/N,k_2}(k_2) < 2 \le d_{k_2\theta/N,k_2}(k_2 - 1)$, $d_{k_2\theta/N,k_2}(k_2) = d_{\theta,N}(k_2)$ and $d_{k_2\theta/N,k_2}(k_2 - 1) = d_{\theta,N}(k_2 - 1)$ and by Lemma 1. ∎

11. Extension to all initial conditions without two particles at the same place

We first prove Proposition 2: we can build a $KS(\theta, N)$-process, i.e. a $QKS(\theta, N)$-process such that $\mathbb{P}_x^X \circ X_t^{-1}$ is absolutely continuous for all $x \in E_2$ and all $t > 0$. We next conclude the proofs of Proposition 3 and Theorem 5.

11.1. Construction of a Keller–Segel process

We fix $\theta > 0$ and $N \geq 2$ such that $N > \theta$ during the whole subsection. For each $n \in \mathbb{N}^*$, we introduce $\phi_n \in C^\infty(\mathbb{R}_+, \mathbb{R}_+^*)$ such that $\phi_n(r) = r$ for all $r \geq 1/n$ and we set, for $x \in (\mathbb{R}^2)^N$,

$$\mathbf{m}_n(x) = \prod_{1 \leq i \neq j \leq N} [\phi_n(\|x^i - x^j\|^2)]^{-\theta/N} \quad \text{and} \quad \mu_n(\mathrm{d}x) = \mathbf{m}_n(x)\mathrm{d}x.$$

We then consider the $(\mathbb{R}^2)^N$-valued SDE

$$X_t^n = x + B_t + \int_0^t \frac{\nabla \mathbf{m}_n(X_s^n)}{2\mathbf{m}_n(X_s^n)} \mathrm{d}s, \tag{67}$$

which is strongly well-posed, for every initial condition, since the drift coefficient is smooth and bounded. We denote by $\mathbb{X}^n = (\Omega^n, \mathcal{M}^n, (X_t^n)_{t \geq 0}, (\mathbb{P}_x^n)_{x \in (\mathbb{R}^2)^N})$ the corresponding Markov process.

Lemma 24. *For all $n \geq 1$, $\mathbb{X}^n$ is a μ_n-symmetric $(\mathbb{R}^2)^N$-valued diffusion with regular Dirichlet space $(\mathcal{E}^n, \mathcal{F}^n)$ with core $C_c^\infty((\mathbb{R}^2)^N)$ such that for all $\varphi \in C_c^\infty((\mathbb{R}^2)^N)$,*

$$\mathcal{E}^n(\varphi, \varphi) = \frac{1}{2} \int_{(\mathbb{R}^2)^N} \|\nabla \varphi\|^2 \, \mathrm{d}\mu_n.$$

Moreover, $\mathbb{P}_x^n \circ (X_t^n)^{-1}$ has a density with respect to the Lebesgue measure on $(\mathbb{R}^2)^N$ for all $t > 0$ and all $x \in (\mathbb{R}^2)^N$.

Proof. Classically, $\mathbb{X}^n$ is a μ_n-symmetric diffusion and its (strong) generator $\mathcal{L}^n$ has the property that for all $\varphi \in C_c^\infty((\mathbb{R}^2)^N)$ and $x \in (\mathbb{R}^2)^N$,

$$\mathcal{L}^n \varphi(x) = \frac{1}{2} \Delta \varphi(x) + \frac{\nabla \mathbf{m}_n(x)}{2\mathbf{m}_n(x)} \cdot \nabla \varphi(x).$$

Hence (see Section B.1), one easily shows that for the Dirichlet space $(\mathcal{E}^n, \mathcal{F}^n)$ of $\mathbb{X}^n$, we have $C_c^\infty((\mathbb{R}^2)^N) \subset \mathcal{F}^n$ and, for $\varphi \in C_c^\infty((\mathbb{R}^2)^N)$, $\mathcal{E}^n(\varphi, \varphi) = \frac{1}{2} \int_{(\mathbb{R}^2)^N} \|\nabla \varphi\|^2 \mathrm{d}\mu_n$. Since $(\mathcal{E}^n, \mathcal{F}^n)$ is closed, we deduce that

$$\overline{C_c^\infty((\mathbb{R}^2)^N)}^{\mathcal{E}_1^n} \subset \mathcal{F}^n,$$

where $\mathcal{E}_1^n(\cdot, \cdot) = \mathcal{E}^n(\cdot, \cdot) + \|\cdot\|_{L^2((\mathbb{R}^2)^N, \mu_n)}^2$. But thanks to [11, Lemma 3.3.5, p. 136],

$$\mathcal{F}^n \subset \{\varphi \in L^2((\mathbb{R}^2)^N, \mu_n) : \nabla \varphi \in L^2((\mathbb{R}^2)^N, \mu_n)\},$$

where ∇ is understood in the sense of distributions. Since finally

$$\overline{C_c^\infty((\mathbb{R}^2)^N)}^{\mathcal{E}_1^n} = \{\varphi \in L^2((\mathbb{R}^2)^N, \mu_n) : \nabla \varphi \in L^2((\mathbb{R}^2)^N, \mu_n)\},$$

$\mathbb{X}^n$ has the announced Dirichlet space. Finally, the absolute continuity of $\mathbb{P}_x^n \circ (X_t^n)^{-1}$, for $t > 0$ and $x \in (\mathbb{R}^2)^N$, immediately follows from the (standard) Girsanov theorem, since the drift coefficient is bounded. ■

For all $x \in E_2$ we set $d_x = \min_{i \neq j} \|x^i - x^j\|^2$. For $n \geq 1$, we introduce the open set

$$E_2^n = \{x \in (\mathbb{R}^2)^N : d_x > 1/n \text{ and } \|x\| < n\}. \quad (68)$$

We also fix a $QKS(\theta, N)$-process $\mathbb{X} = (\Omega^X, \mathcal{M}^X, (X_t)_{t\geq 0}, (\mathbb{P}_x^X)_{x\in\mathcal{X}_\triangle})$ for the whole subsection.

Lemma 25. *There exists an exceptional set $\mathcal{N}_0 \subset E_2$ with respect to $\mathbb{X}$ such that for all $n \geq 1$ and all $x \in E_2^n \setminus \mathcal{N}_0$, the law of $(X^n_{t\wedge\tau_n})_{t\geq 0}$ under $\mathbb{P}_x^n$ equals the law of $(X_{t\wedge\sigma_n})_{t\geq 0}$ under $\mathbb{P}_x^X$, where*

$$\tau_n = \inf\{t > 0 : X_t^n \notin E_2^n\} \quad \text{and} \quad \sigma_n = \inf\{t > 0 : X_t \notin E_2^n\}.$$

Proof. We fix $n \geq 1$. Applying Lemma B.6 to $\mathbb{X}^n$ and $\mathbb{X}$ with the open set E_2^n, using the fact that $\mathbf{m}_n = \mathbf{m}$ on E_2^n and Lemma 24, we find that the processes $\mathbb{X}^n$ and $\mathbb{X}$ killed when leaving E_2^n have the same Dirichlet space. By uniqueness (see [11, Theorem 4.2.8, p. 167]), there exists an exceptional set $\mathcal{N}_n$ such that for all $x \in E_2^n \setminus \mathcal{N}_n$, the law of $(X_t^n)_{t\geq 0}$ killed when leaving E_2^n under $\mathbb{P}_x^n$ equals the law of $(X_t)_{t\geq 0}$ killed when leaving E_2^n under $\mathbb{P}_x^X$. We conclude the proof by setting $\mathcal{N}_0 = \bigcup_{n\geq 1} \mathcal{N}_n$. ∎

Lemma 26. *For all exceptional sets $\mathcal{N}$ with respect to $\mathbb{X}$, and all $n \geq 1$ and $x \in E_2^n$, we have $\mathbb{P}_x^n(X^n_{\tau_n} \notin \mathcal{N}) = 1$.*

Proof. We fix $\mathcal{N}$ an exceptional set with respect to $\mathbb{X}$, $n \geq 1$ and $x \in E_2^n$. For $\varepsilon \in (0, 1]$, we write

$$\begin{aligned}\mathbb{P}_x^n(X^n_{\tau_n} \in \mathcal{N}) &\leq \mathbb{P}_x^n(\tau_n \leq \varepsilon) + \mathbb{P}_x^n(\tau_n > \varepsilon, X^n_{\tau_n} \in \mathcal{N}) \\ &= \mathbb{P}_x^n(\tau_n \leq \varepsilon) + \mathbb{E}_x^n[\mathbf{1}_{\{\tau_n>\varepsilon\}}\mathbb{P}^n_{X^n_\varepsilon}(X^n_{\tau_n} \in \mathcal{N})]\end{aligned}$$

by the Markov property. But by Lemma 25, for all $y \in E_2^n \setminus \mathcal{N}_0$, the law of $(X^n_{t\wedge\tau_n})_{t\geq 0}$ under $\mathbb{P}_y^n$ is equal to the law of $(X_{t\wedge\sigma_n})_{t\geq 0}$ under $\mathbb{P}_y^X$. Since $\mathcal{N}_0 \cup \mathcal{N}$ is exceptional for $\mathbb{X}$, we can find $\mathcal{N}' \supset \mathcal{N}_0 \cup \mathcal{N}$ properly exceptional for $\mathbb{X}$ (see Section B.1). Hence for all $y \in E_2^n \setminus \mathcal{N}'$,

$$\mathbb{P}_y^n(X^n_{\tau_n} \in \mathcal{N}) \leq \mathbb{P}_y^n(X^n_{\tau_n} \in \mathcal{N}') = \mathbb{P}_y^X(X_{\sigma_n} \in \mathcal{N}') = 0.$$

Since $\mathbb{P}_x^n \circ (X_\varepsilon^n)^{-1}$ has a density by Lemma 25, we conclude that $\mathbb{P}_x^n(X_\varepsilon^n \in \mathcal{N}') = 0$ and thus $\mathbb{P}_x^n$-a.s., we have $\mathbb{P}^n_{X^n_\varepsilon}(X^n_{\tau_n} \in \mathcal{N}) = 0$. All in all, we have proved the inequality $\mathbb{P}_x^n(X^n_{\tau_n} \in \mathcal{N}) \leq \mathbb{P}_x^n(\tau_n \leq \varepsilon)$, and it suffices to let $\varepsilon \to 0$, since $\mathbb{P}_x^n(\tau_n > 0) = 1$ by continuity and $x \in E_2^n$. ∎

Using Lemmas 25 and 26, it is slightly technical but not difficult to build from $\mathbb{X}$ and the family $(\mathbb{X}^n)_{n\geq 1}$ an $\mathcal{X}_\triangle$-valued diffusion $\tilde{\mathbb{X}} = (\tilde{\Omega}^X, \tilde{\mathcal{M}}^X, (\tilde{X}_t)_{t\geq 0}, (\tilde{\mathbb{P}}_x^X)_{x\in\mathcal{X}_\triangle})$ such that

- for all $x \in \mathcal{X}_\triangle \setminus \mathcal{N}_0$, the law of $(\tilde{X}_t)_{t\geq 0}$ under $\tilde{\mathbb{P}}_x^X$ equals the law of $(X_t)_{t\geq 0}$ under $\mathbb{P}_x^X$,
- for all $x \in \mathcal{N}_0$, setting $n = 1 + \lfloor \max(1/d_x, \|x\|) \rfloor$ (so that $x \in E_2^n$), the law of $(\tilde{X}_{t\wedge\tilde{\sigma}_n})_{t\geq 0}$ under $\tilde{\mathbb{P}}_x^X$ is the same as that of $(X^n_{t\wedge\tau_n})_{t\geq 0}$ under $\mathbb{P}_x^n$, and the law of

$(\tilde X_{\tilde\sigma_n+t})_{t\geq 0}$ under $\tilde{\mathbb{P}}_x^X$ conditioned on $\tilde{\mathcal{M}}_{\tilde\sigma_n}^X$ equals the law of $(X_t)_{t\geq0}$ under $\mathbb{P}^X_{\tilde X_{\sigma_n}}$. We have used the notation $\tilde\sigma_n = \inf\{t > 0 : \tilde X_t \notin E_2^n\}$ and $\tilde{\mathcal{M}}_t^X = \sigma(\tilde X_s : s \in [0,t])$.

Remark 27. For all $x \in E_2$, setting $n = 1 + \lfloor \max(1/d_x, \|x\|)\rfloor$, the law of $(\tilde X_{t\wedge\tilde\sigma_n})_{t\geq0}$ under $\tilde{\mathbb{P}}_x^X$ is the same as that of $(X^n_{t\wedge\tau_n})_{t\geq0}$ under $\mathbb{P}^n_x$.

Proof. This follows from Lemma 25 when $x \in E_2 \setminus \mathcal{N}_0$ and from the definition of $\tilde{\mathbb{X}}$ otherwise. ■

Proof of Proposition 2. We fix $N \geq 2$ and $\theta > 0$ such that $N > \theta$ and we prove that $\tilde{\mathbb{X}}$ defined above is a $KS(\theta, N)$-process. First, it is clear that $\tilde{\mathbb{X}}$ is a $QKS(\theta, N)$-process because $\tilde{\mathbb{X}}$ is an $\mathcal{X}_\triangle$-valued diffusion and for all $x \in \mathcal{X}_\triangle \setminus \mathcal{N}_0$, the law of $(\tilde X_t)_{t\geq0}$ under $\tilde{\mathbb{P}}_x^X$ equals the law of $(X_t)_{t\geq0}$ under $\mathbb{P}_x^X$, with $\mathcal{N}_0$ exceptional for $\mathbb{X}$. It remains to prove that for all $x \in E_2$, all $t > 0$ and all Lebesgue-null $A \subset (\mathbb{R}^2)^N$, we have $\tilde{\mathbb{P}}_x^X(\tilde X_t \in A) = 0$. We set $n = 1 + \lfloor \max(1/d_x, \|x\|)\rfloor$ and write, for any $\varepsilon \in (0,t)$,

$$\begin{aligned}\tilde{\mathbb{P}}_x^X(\tilde X_t \in A) &\leq \tilde{\mathbb{P}}_x^X(\tilde\sigma_n > \varepsilon,\ \tilde X_t \in A) + \tilde{\mathbb{P}}_x^X(\tilde\sigma_n \leq \varepsilon)\\ &= \tilde{\mathbb{E}}_x^X[\mathbf{1}_{\{\tilde\sigma_n>\varepsilon\}}\tilde{\mathbb{P}}^X_{\tilde X_\varepsilon}(\tilde X_{t-\varepsilon} \in A)] + \tilde{\mathbb{P}}_x^X(\tilde\sigma_n \leq \varepsilon).\end{aligned}$$

Since $\tilde{\mathbb{X}}$ is μ-symmetric (because it is a $QKS(\theta, N)$-process) and $\tilde P_{t-\varepsilon}1 \leq 1$, where $\tilde P_t$ is the semigroup of $\tilde{\mathbb{X}}$, and since A is Lebesgue-null, we have

$$\int_{(\mathbb{R}^2)^N} \tilde{\mathbb{P}}_y(\tilde X_{t-\varepsilon} \in A)\,\mu(\mathrm{d}y) \leq \mu(A) = 0.$$

Hence there is a Lebesgue-null subset B of $(\mathbb{R}^2)^N$ (depending on $t-\varepsilon$) such that we have $\tilde{\mathbb{P}}_y(\tilde X_{t-\varepsilon} \in A) = 0$ for every $y \in (\mathbb{R}^2)^N \setminus B$. We conclude that

$$\begin{aligned}\tilde{\mathbb{P}}_x^X(\tilde X_t \in A) &\leq \tilde{\mathbb{P}}_x^X(\tilde\sigma_n > \varepsilon, \tilde X_\varepsilon \in B) + \tilde{\mathbb{P}}_x^X(\tilde\sigma_n \leq \varepsilon)\\ &= \mathbb{P}_x^n(\tau_n > \varepsilon, X^n_\varepsilon \in B) + \tilde{\mathbb{P}}_x^X(\tilde\sigma_n \leq \varepsilon),\end{aligned}$$

where we have used Remark 27. Since B is Lebesgue-null, we deduce from Lemma 24 that $\mathbb{P}_x^n(\tau_n > \varepsilon, X^n_\varepsilon \in B) = 0$. Thus $\tilde{\mathbb{P}}_x^X(\tilde X_t \in A) \leq \tilde{\mathbb{P}}_x^X(\tilde\sigma_n \leq \varepsilon)$, which tends to 0 as $\varepsilon \to 0$ because $\tilde{\mathbb{P}}_x^X(\tilde\sigma_n > 0) = 1$ by continuity. ■

11.2. Final proofs

We fix $\theta > 0$ and $N \geq 2$ such that $N > \theta$, and a $KS(\theta, N)$-process $\mathbb{X}$, which exists thanks to Section 11.1. We recall that E_2^n was introduced in (68) and define, for all $n \geq 1$, $\sigma_n = \inf\{t \geq 0 : X_t \notin E_2^n\}$, as well as the σ-field

$$\mathcal{G} = \bigcap_{n\geq1} \sigma(\{X_{\sigma_n+t}, t \geq 0\}).$$

Lemma 28. *Fix $A \in \mathcal{G}$. If $\mathbb{P}_x^X(A) = 0$ quasi-everywhere, then $\mathbb{P}_x^X(A) = 0$ for all $x \in E_2$.*

Proof. We fix $A \in \mathcal{G}$ such that $\mathbb{P}_x^X(A) = 0$ quasi-everywhere. There is an exceptional set $\mathcal{N}$ such that $\mathbb{P}_x^X(A) = 0$ for all $x \in E_2 \setminus \mathcal{N}$. We now fix $x \in E_2$ and set $n = 1 + \lfloor \max(1/d_x, \|x\|) \rfloor$. For any $\varepsilon \in (0, 1]$,

$$\mathbb{P}_x^X(A) \le \mathbb{P}_x^X(\sigma_n \le \varepsilon) + \mathbb{P}_x^X[\sigma_n > \varepsilon, A].$$

By the Markov property and since $A \in \mathcal{G} \subset \sigma(X_{\sigma_n+t}, t \ge 0)$, we get

$$\mathbb{P}_x^X[\sigma_n > \varepsilon, A] = \mathbb{E}_x^X[\mathbf{1}_{\{\sigma_n > \varepsilon\}} \mathbb{P}_{X_\varepsilon}^X(A)].$$

But the law of X_ε under $\mathbb{P}_x^X$ has a density, whence $\mathbb{P}_x^X(X_\varepsilon \in \mathcal{N}) = 0$, and consequently $\mathbb{P}_x^X(\mathbb{P}_{X_\varepsilon}^X(A) = 0) = 1$. Hence $\mathbb{P}_x^X[\sigma_n > \varepsilon, A] = 0$ and we end up with $\mathbb{P}_x^X(A) \le \mathbb{P}_x^X(\tau_n \le \varepsilon)$. As usual, we conclude that $\mathbb{P}_x^X(A) = 0$ by letting $\varepsilon \to 0$. ∎

Proof of Proposition 3. Let $\theta \in (0, 2)$ and $N \ge 2$. Since our $KS(\theta, N)$-process $\mathbb{X}$ is a $QKS(\theta, N)$-process, we know from Proposition 16 (i) that $\mathbb{P}_x^X(\zeta = \infty) = 1$ q.e. We want to prove that $\mathbb{P}_x^X(\zeta = \infty) = 1$ for all $x \in E_2$. By Lemma 28, it suffices to check that $\{\zeta = \infty\}$ belongs to $\mathcal{G}$, which is not hard since for each $n \ge 1$,

$$\{\zeta = \infty\} = \{X_t \in \mathcal{X} \text{ for all } t \ge 0\} = \{X_t \in \mathcal{X} \text{ for all } t \ge \sigma_n\} \in \sigma(\{X_{\sigma_n+t}, t \ge 0\}).$$

For the second equality, we have used the fact that $X_t \in \bar{E}_2^n \subset \mathcal{X}$ for all $t \in [0, \sigma_n]$ by definition. ∎

Proof of Theorem 5. Let us fix $\theta \ge 2$ and $N > 3\theta$. Since our $KS(\theta, N)$-process $\mathbb{X}$ is a $QKS(\theta, N)$-process, we know from Section 10 that all the conclusions of Theorem 5 hold quasi-everywhere. In other words, $\mathbb{P}_x^X(A) = 1$ quasi-everywhere, where A is the event on which we have $\zeta < \infty$, $X_{\zeta-} = \lim_{t \to \zeta-} X_t \in (\mathbb{R}^2)^N$, there is $K_0 \in [\![1, N]\!]$ with $|K_0| = k_0$ such that there is a K_0-collision in the configuration $X_{\zeta-}$, etc. We want to prove that $\mathbb{P}_x^X(A) = 1$ for all $x \in E_2$. By Lemma 28, it suffices to check that A belongs to $\mathcal{G}$. But for each $n \ge 1$, A indeed belongs to $\sigma(X_{\sigma_n+t}, t \ge 0)$, because no collision (nor explosion) may happen before leaving E_2^n. ∎

We end this section with the following remark (which we will not use anywhere).

Remark 29. Fix $\theta \ge 0$ and $N \ge 2$ such that $N > \theta$. Consider a $KS(\theta, N)$ process $\mathbb{X}$ and define $\sigma = \inf\{t \ge 0 : X_t \notin E_2\}$. For all $x \in E_2$, there is some $(\mathcal{M}_t^X)_{t \ge 0}$-Brownian motion $((B_t^i)_{t \ge 0})_{i \in [\![1,N]\!]}$ (of dimension $2N$) under $\mathbb{P}_x^X$ such that for all $t \in [0, \sigma)$ and $i \in [\![1, N]\!]$,

$$X_t^i = x^i + B_t^i - \frac{\theta}{N} \sum_{j \ne i} \int_0^t \frac{X_s^i - X_s^j}{\|X_s^i - X_s^j\|^2} \, ds. \tag{69}$$

Proof. It suffices to prove the result on $[0, \sigma_n)$, where $\sigma_n = \inf\{t \ge 0 : X_t \notin E_2^n\}$. For any $x \in E_2^n$ and for a given Brownian motion, the solutions to (69) and (67) classically coincide while they remain in E_2^n, because their drift coefficients coincide and are smooth inside E_2^n. Hence, recalling the notation of Section 11.1, it suffices to prove that the semigroups $P_t(x,\cdot)$ and $P_t^n(x,\cdot)$ of the Markov processes $\mathbb{X}$ and $\mathbb{X}^n$ killed when getting out of E_2^n coincide for all $x \in E_2^n$.

By Lemma 25, there is an exceptional set $\mathcal{N}_0$ such that $P_t(x,\cdot) = P_t^n(x,\cdot)$ for all $x \in E_2^n \setminus \mathcal{N}_0$. We next fix $x \in E_2^n$. For any $\varepsilon \in (0,t)$, since $P_\varepsilon(x,\cdot)$ has a density and $\mathcal{N}_0$ is Lebesgue-null, we easily deduce that $P_t(x,\cdot) = (P_\varepsilon P_{t-\varepsilon})(x,\cdot) = (P_\varepsilon P_{t-\varepsilon}^n)(x,\cdot)$. It is then not difficult, as P_t^n is Feller, to let $\varepsilon \to 0$ and conclude that indeed $P_t(x,\cdot) = P_t^n(x,\cdot)$. ∎

Appendix A. A few elementary computations

We recall that $d_{\theta,N}(k) = (k-1)(2-\theta k/N)$ for $k \geq 2$.

Proof of Lemma 1. First, (3), which says that $d_{\theta,N}(k) > 0$ if and only if $k < k_0 = \lceil 2N/\theta \rceil$, is clear. We next fix $N > 3\theta \geq 6$ so that $k_0 \in [\![7,N]\!]$ and $d_{\theta,N}(2) = 2 - 2\theta/N \in (4/3, 2)$. By concavity of $x \mapsto (x-1)(2-\theta x/N)$, it only remains to check that (i) $d_{\theta,N}(3) \geq 2$, (ii) $d_{\theta,N}(k_0-3) \geq 2$, and (iii) $d_{\theta,N}(k_0-1) < 2$. We introduce $a = 2N/\theta > 6$ and observe that $d_{\theta,N}(k) = 2a^{-1}(k-1)(a-k)$ and $k_0 = \lceil a \rceil$.

For (i), we write $d_{\theta,N}(3) = 4a^{-1}(a-3) = 4 - 12a^{-1} > 2$ since $a > 6$.

Concerning (ii), we have $d_{\theta,N}(k_0-3) = 2a^{-1}(\lceil a \rceil - 4)(a - \lceil a \rceil + 3)$ and we need $(\lceil a \rceil - 4)(a - \lceil a \rceil + 3) \geq a$. Writing $a = n + \alpha$ with an integer $n \geq 6$ and $\alpha \in (0,1]$, we need $(n-3)(2+\alpha) \geq n + \alpha$, and this holds true because $2(n-3) \geq n$ and $(n-3)\alpha \geq \alpha$.

For (iii), we write

$$d_{\theta,N}(k_0-1) = 2a^{-1}(\lceil a \rceil - 2)(a - \lceil a \rceil + 1) \leq 2a^{-1}(\lceil a \rceil - 2) < 2. \quad ∎$$

We next study the reference measure of the Keller–Segel particle system.

Proposition A.1. *Let $N \geq 2$ and $\theta > 0$ be such that $N > \theta$. Recall that $k_0 = \lceil 2N/\theta \rceil$ and the definition (4) of $\mu(\mathrm{d}x) = \mathbf{m}(x)\mathrm{d}x$.*

(i) *The measure μ is Radon on E_{k_0}.*

(ii) *If $k_0 \leq N$, then μ is not Radon on E_{k_0+1}.*

Proof. (i) To show that μ is Radon on E_{k_0}, we have to check that for all $x = (x^1, \dots, x^N) \in E_{k_0}$, which we now fix, there is an open set $O_x \subset E_{k_0}$ such that $x \in O_x$ and $\mu(O_x) < \infty$. We choose $O_x = \prod_{i=1}^N B(x^i, d_x)$, where the balls are subsets of $\mathbb{R}^2$ and where

$$d_x = 1 \wedge \min\{\|x^i - x^j\|/3 : i,j \in [\![1,N]\!] \text{ such that } x^i \neq x^j\} > 0.$$

We consider the partition $K_1, \dots, K_\ell$ of $[\![1,N]\!]$ such that for all $p \neq q$ in $[\![1,\ell]\!]$, all $i,j \in K_p$ and all $k \in K_q$, we have $x^i = x^j$ and $x^i \neq x^k$. Since $x \in E_{k_0}$, it follows that $\max_{p \in [\![1,\ell]\!]} |K_p| \leq k_0 - 1$. By definition of O_x and d_x, we see that for all $y \in O_x$, all $p \neq q$ in $[\![1,\ell]\!]$ and all $i \in K_p$ and $j \in K_q$,

$$\|y^i - y^j\| \geq \|x^i - x^j\| - \|x^i - y^i\| - \|x^j - y^j\| \geq \|x^i - x^j\| - 2d_x \geq d_x.$$

This implies that for some finite constant C depending on x, for all $y \in O_x$,

$$\mathbf{m}(y) = \prod_{1\le i\ne j\le N} \|y^i - y^j\|^{-\theta/N} \le C \prod_{p=1}^{\ell}\Big(\prod_{i,j\in K_p, i\ne j} \|y^i - y^j\|^{-\theta/N}\Big).$$

Recall now that $\mu(\mathrm{d}y) = \mathbf{m}(y)\mathrm{d}y$ and we want to show that $\mu(O_x) < \infty$. Since $x^i = x^j$ for all $i, j \in K_p$ and all $p \in [\![1,\ell]\!]$, and since $|K_p| \le k_0 - 1$, $d_x \le 1$, a translation argument reduces our task to showing that for any $n \in [\![2, k_0 - 1]\!]$ (when $k_0 > N$, one could study only $n \in [\![2, N]\!]$),

$$I_n = \int_{(B(0,1))^n} \Big(\prod_{1\le i\ne j\le n} \|y^i - y^j\|^{-\theta/N}\Big) \mathrm{d}y^1 \dots \mathrm{d}y^n < \infty.$$

We fix $n \in [\![2, k_0 - 1]\!]$ and show that $I_n < \infty$. Since $\|u\|^2 \ge |u_1 u_2|$ for all $u = (u_1, u_2) \in \mathbb{R}^2$, we have $I_n \le J_n^2$, where

$$J_n = \int_{[-1,1]^n} \Big(\prod_{1\le i\ne j\le n} |t^i - t^j|^{-\theta/(2N)}\Big) \mathrm{d}t^1 \dots \mathrm{d}t^n.$$

But for all $t^1, \dots, t^n \in \mathbb{R}$,

$$\begin{aligned}\prod_{1\le i\ne j\le n} |t^i - t^j|^{-\theta/(2N)} &= \prod_{i=1}^{n}\Big(\prod_{j=1, j\ne i}^{n} |t^i - t^j|^{-\theta/(2N)}\Big)\\ &\le \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1, j\ne i}^{n} |t^i - t^j|^{-\theta n/(2N)}\end{aligned}$$

by the inequality of arithmetic and geometric means. Thus by symmetry,

$$\begin{aligned}J_n &\le \int_{[-1,1]^n}\Big(\prod_{j=2}^{n} |t^1 - t^j|^{-\theta n/(2N)}\Big) \mathrm{d}t^1 \dots \mathrm{d}t^n\\ &= \int_{-1}^{1}\Big(\int_{-1}^{1} |t^1 - t^2|^{-\theta n/(2N)} \mathrm{d}t^2\Big)^{n-1} \mathrm{d}t^1.\end{aligned}$$

Consequently,

$$J_n \le \int_{-1}^{1}\Big(\int_{-2}^{2} |s|^{-\theta n/(2N)} \,\mathrm{d}s\Big)^{n-1} \mathrm{d}t^1.$$

Since $n \le k_0 - 1 = \lceil 2N/\theta\rceil - 1 < 2N/\theta$, we have $\theta n/(2N) < 1$, so that $J_n < \infty$, whence $I_n < \infty$.

(ii) We next assume that $k_0 \in [\![2, N]\!]$. To prove that μ is not Radon on E_{k_0+1}, we show that $\mu(K) = \infty$ for the compact subset

$$K = \prod_{i=1}^{k_0} \overline{B}(0,1) \times \prod_{k=k_0+1}^{N} \overline{B}((2k,0), 1/2)$$

of E_{k_0+1}. All the balls in the previous formula are balls in $\mathbb{R}^2$. For $x = (x^1, \dots, x^N) \in K$, the points $x^{k_0+1}, \dots, x^N$ are far apart and far from $x^1, \dots, x^{k_0}$, which explains that K is indeed compact in E_{k_0+1}. There is a positive constant $c>0$ such that for all $x\in K$,

$$\mathbf{m}(x)=\prod_{1\le i\ne j\le N}\|x^i-x^j\|^{-\theta/N}\ge c\prod_{1\le i\ne j\le k_0}\|x^i-x^j\|^{-\theta/N},$$

whence, the value of $c>0$ being allowed to vary,

$$\mu(K)\ge c\int_{(B(0,1))^{k_0}}\Big(\prod_{1\le i\ne j\le k_0}\|x^i-x^j\|^{-\theta/N}\Big)\,\mathrm{d}x^1\dots\mathrm{d}x^{k_0}.$$

We now observe that

$$A=\{x=(x^1,\dots,x^{k_0}):x^1,x^2\in B(0,1/3),\ \forall i\notin\{1,2\},\ x^i\in B(x^1,\|x^1-x^2\|)\}\\ \subset(B(0,1))^{k_0}$$

and for $x\in A$, we have $\|x^i-x^j\|\le\|x^i-x^1\|+\|x^j-x^1\|\le 2\|x^1-x^2\|$ for all $i,j=1,\dots,k_0$, which yields

$$\prod_{1\le i\ne j\le k_0}\|x^i-x^j\|^{-\theta/N}\ge c\|x^1-x^2\|^{-k_0(k_0-1)\theta/N}.$$

Consequently,

$$\begin{aligned}\mu(K)&\ge c\int_{B(0,1/3)^2}\|x^1-x^2\|^{-k_0(k_0-1)\theta/N}\,\mathrm{d}x^1\,\mathrm{d}x^2\int_{B(x_1,\|x^1-x^2\|)^{k_0-2}}\mathrm{d}x^3\dots\mathrm{d}x^{k_0}\\ &\ge c\int_{B(0,1/3)^2}\|x^1-x^2\|^{-k_0(k_0-1)\theta/N+2(k_0-2)}\,\mathrm{d}x^1\,\mathrm{d}x^2\\ &\ge c\int_{B(0,1/3)}\|u\|^{-k_0(k_0-1)\theta/N+2(k_0-2)}\,\mathrm{d}u,\end{aligned}$$

where we have used the change of variables $u=x^1-x^2$ and $v=x^1+x^2$. The last integral diverges, because $-k_0(k_0-1)\theta/N+2(k_0-2)=d_{\theta,N}(k_0)-2\le-2$; recall that $d_{\theta,N}(k_0)=(k_0-1)(2-k_0\theta/N)\le 0$ by definition of k_0. ■

We need a similar result on the sphere $\mathbb{S}$ defined in Section 2, where $\gamma:\mathbb{R}^2\to(\mathbb{R}^2)^N$ and $\Psi:\mathbb{R}^2\times\mathbb{R}_+^*\times\mathbb{S}\to E_N\subset(\mathbb{R}^2)^N$ were also introduced. First, we show an explicit link between $\mu(\mathrm{d}x)=\mathbf{m}(x)\mathrm{d}x$ and $\beta(\mathrm{d}u)=\mathbf{m}(u)\sigma(\mathrm{d}u)$ defined in (4) and (8), which we use several times.

Lemma A.2. *Fix $N\ge2$, $\theta>0$ and set $\nu=d_{\theta,N}(N)/2-1$. For all Borel $\varphi:(\mathbb{R}^2)^N\to\mathbb{R}_+$,*

$$\int_{(\mathbb{R}^2)^N}\varphi(x)\,\mu(\mathrm{d}x)=\frac12\int_{\mathbb{R}^2\times\mathbb{R}_+^*\times\mathbb{S}}\varphi(\Psi(z,r,u))r^\nu\,\mathrm{d}z\,\mathrm{d}r\,\beta(\mathrm{d}u).$$

Proof. Since $H=\{y=(y^1,\dots,y^N)\in(\mathbb{R}^2)^N:\sum_{i=1}^N y^i=0\}$ and $\mathbf{m}$ is translation invariant,

$$\int_{(\mathbb{R}^2)^N}\varphi(x)\,\mu(\mathrm{d}x)=\int_{(\mathbb{R}^2)^N}\varphi(x)\mathbf{m}(x)\,\mathrm{d}x=\int_{\mathbb{R}^2\times H}\varphi(\gamma(z)+y)\mathbf{m}(y)\,\mathrm{d}z\,\mathrm{d}y.$$

We next note that $\mathbb{S}$ is the (true) unit sphere of the $(2N-2)$-dimensional Euclidean space H and make the substitution $(\ell, u) = (\|y\|, y/\|y\|)$:

$$\int_{(\mathbb{R}^2)^N} \varphi(x)\, \mu(\mathrm{d}x) = \int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} \varphi(\gamma(z)+\ell u)\mathbf{m}(\ell u)\ell^{2N-3}\, \mathrm{d}z\, \mathrm{d}\ell\, \sigma(\mathrm{d}u).$$

We substitute $\ell = \sqrt{r}$ to obtain

$$\int_{(\mathbb{R}^2)^N} \varphi(x)\, \mu(\mathrm{d}x) = \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} \varphi(\gamma(z)+\sqrt{r}u)\mathbf{m}(\sqrt{r}u)r^{N-2}\, \mathrm{d}z\, \mathrm{d}r\, \sigma(\mathrm{d}u).$$

But $\mathbf{m}(\sqrt{r}u)r^{N-2} = r^{N-2-\theta(N-1)/2}\mathbf{m}(u)$ by (4) and $\beta(\mathrm{d}u) = \mathbf{m}(u)\sigma(\mathrm{d}u)$, whence

$$\int_{(\mathbb{R}^2)^N} \varphi(x)\, \mu(\mathrm{d}x) = \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} \varphi(\Psi(z,r,u))r^{N-2-\theta(N-1)/2}\, \mathrm{d}z\, \mathrm{d}r\, \beta(\mathrm{d}u).$$

Since finally $\nu = d_{\theta,N}(N)/2 - 1 = N-2-\theta(N-1)/2$, the conclusion follows. ∎

We can now study the measure β on $\mathbb{S}$.

Proposition A.3. *Let $N \geq 2$ and $\theta > 0$ such that $N > \theta$. Recall that $k_0 = \lceil 2N/\theta \rceil$.*

(i) *The measure β is Radon on $\mathbb{S} \cap E_{k_0}$.*

(ii) *If $k_0 \geq N$, then $\beta(\mathbb{S}) < \infty$.*

Proof. (i) For $\varepsilon \in (0,1]$, we introduce

$$\mathcal{K}_\varepsilon = \{x \in (\mathbb{R}^2)^N : \forall K \subset [\![1,N]\!] \text{ with } |K| \geq k_0,\ \text{we have } R_K(x) \geq \varepsilon\}$$

and $\mathcal{L}_\varepsilon = \mathcal{K}_\varepsilon \cap \mathbb{S}$. Since $\mathcal{K}_\varepsilon \cap \bar{B}(0,1)$ is compact in E_{k_0}, where $B(0,1)$ is the unit ball of $(\mathbb{R}^2)^N$, we know from Proposition A.1 (i) that $\mu(\mathcal{K}_\varepsilon \cap B(0,1)) < \infty$. Now by Lemma A.2,

$$\mu(\mathcal{K}_\varepsilon \cap B(0,1)) = \frac{1}{2}\int_{\mathbb{R}^2\times\mathbb{R}_+\times\mathbb{S}} \mathbf{1}_{\{\gamma(z)+\sqrt{r}u\in\mathcal{K}_\varepsilon\cap B(0,1)\}} r^\nu\, \mathrm{d}z\, \mathrm{d}r\, \beta(\mathrm{d}u).$$

But for $(z,r,u) \in \mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{S}$,

$$\gamma(z)+\sqrt{r}u \in \mathcal{K}_\varepsilon \cap B(0,1) \quad \text{if and only if} \quad u \in \mathcal{L}_{\varepsilon/r} \text{ and } N\|z\|^2 + r < 1.$$

Indeed, $R_K(\gamma(z)+\sqrt{r}u) = rR_K(u)$ for all $K \subset [\![1,N]\!]$, while $\|\gamma(z)+\sqrt{r}u\|^2 = \sum_{i=1}^N \|z+\sqrt{r}u^i\|^2 = N\|z\|^2 + r$ because $\sum_{i=1}^N u^i = 0$ and $\sum_{i=1}^N \|u^i\|^2 = 1$. Thus

$$\mu(\mathcal{K}_\varepsilon \cap B(0,1)) = \int_{\mathbb{R}^2\times\mathbb{R}_+} \mathbf{1}_{\{N\|z\|^2+r<1\}} r^\nu \beta(\mathcal{L}_{\varepsilon/r})\, \mathrm{d}z\, \mathrm{d}r.$$

All this implies that $\beta(\mathcal{L}_{\varepsilon/r}) < \infty$ for all $\varepsilon \in (0,1]$ and almost all $r \in (0,1)$. Since $\varepsilon \to \mathcal{L}_\varepsilon$ is monotone, we conclude that $\beta(\mathcal{L}_\varepsilon) < \infty$ for all $\varepsilon \in (0,1]$. Since finally $\bigcup_{\varepsilon\in(0,1]} \mathcal{L}_\varepsilon = \mathbb{S} \cap E_{k_0}$ and since $\mathcal{L}_\varepsilon$ is compact in $\mathbb{S} \cap E_{k_0}$ for each $\varepsilon \in (0,1]$, we conclude as desired that β is Radon on $\mathbb{S} \cap E_{k_0}$.

(ii) We have $\mathbb{S} \subset E_N$, because $R_{[\![1,N]\!]}(u) = 1$ for $u \in \mathbb{S}$. Hence if $k_0 \geq N$, then $\mathbb{S} \subset E_N \subset E_{k_0}$, whence $\mathbb{S} = \mathbb{S} \cap E_{k_0}$ and thus β is Radon on $\mathbb{S}$ by (i). Since finally $\mathbb{S}$ is compact, we conclude that $\beta(\mathbb{S}) < \infty$. ∎

Appendix B. Markov processes and Dirichlet spaces

In a first subsection, we recall some classical definitions and results about Hunt processes, diffusions and Dirichlet spaces found in Fukushima–Oshima–Takeda [11]. In a second subsection, we mention a few results about martingales, time-changes, concatenation, killing and Girsanov transformation of Hunt processes found in [11] and elsewhere.

B.1. Main definitions and properties

Let E be a locally compact separable metrizable space endowed with a Radon measure α such that $\operatorname{supp}\alpha = E$. We set $E_\triangle = E \cup \{\triangle\}$, where $\triangle$ is a cemetery point. See [11, Section A2] for the definition of a Hunt process $\mathbb{Y} = (\Omega, \mathcal{M}, (Y_t)_{t\geq 0}, (\mathbb{P}_y)_{y\in E_\triangle})$: it is a strong Markov process in its canonical filtration, $\mathbb{P}_y(Y_0 = y) = 1$ for all $y \in E_\triangle$, $\triangle$ is an absorbing state, i.e. $Y_t = \triangle$ for all $t \geq 0$ under $\mathbb{P}_\triangle$, and a few more technical properties are satisfied. The life-time of $\mathbb{Y}$ is defined by $\zeta = \inf\{t \geq 0 : Y_t = \triangle\}$.

Let us denote by $P_t(y, \mathrm{d}z)$ its transition kernel. Our Hunt process is said to be α-*symmetric* if $\int_E \varphi P_t\psi \,\mathrm{d}\alpha = \int_E \psi P_t\varphi \,\mathrm{d}\alpha$ for all measurable $\varphi, \psi : E \to \mathbb{R}_+$ and all $t \geq 0$ (see [11, p. 30]). The Dirichlet space $(\mathcal{E}, \mathcal{F})$ of our Hunt process on $L^2(E, \alpha)$ is then defined (see [11, p. 23]) by

$$\mathcal{F} = \Big\{\varphi \in L^2(E,\alpha) : \lim_{t\to 0} \frac{1}{t}\int_E \varphi(P_t\varphi - \varphi)\,\mathrm{d}\alpha \text{ exists}\Big\},$$

$$\mathcal{E}(\varphi,\psi) = -\lim_{t\to 0}\frac{1}{t}\int_E \varphi(P_t\psi - \psi)\,\mathrm{d}\alpha \quad \text{for all } \varphi,\psi \in \mathcal{F}.$$

The generator $(\mathcal{A}, \mathcal{D}_A)$ of $\mathbb{Y}$ is defined as follows:

$$\mathcal{D}_{\mathcal{A}} = \Big\{\varphi \in L^2(E,\alpha) : \lim_{t\to 0}\frac{1}{t}(P_t\varphi - \varphi) \text{ exists in } L^2(E,\alpha)\Big\},$$

and for $\varphi \in \mathcal{D}_{\mathcal{A}}$, we denote by $\mathcal{A}\varphi \in L^2(E,\alpha)$ the above limit. By [11, pp. 20–21],

$$\mathcal{D}_{\mathcal{A}} = \Big\{\varphi \in \mathcal{F} : \exists h \in L^2(E,\alpha) \text{ such that } \forall \psi \in \mathcal{F}, \text{ we have } \mathcal{E}(\varphi,\psi) = -\int_E h\psi\,\mathrm{d}\alpha\Big\}, \tag{B.1}$$

and in that case $\mathcal{A}\varphi = h$.

The one-point compactification $E_\triangle = E \cup \{\triangle\}$ of E is endowed with the topology consisting of all the open subsets of E and of all the sets of the form $K^c \cup \{\triangle\}$ with K compact in E (see [11, p. 69]). Observe that for an $E_\triangle$-valued sequence $(x_n)_{n\geq 0}$, we have $\lim_n x_n = x$ if and only if either

- $x, x_n \in E$ for all n large enough, and $\lim_n x_n = x \in E$ in the usual sense; or
- $x = \triangle$ and for every compact subset K of E, there is $n_K \in \mathbb{N}$ such that $x_n \notin K$ for all $n \geq n_K$.

We say that our Hunt process is *continuous* if $t \mapsto Y_t$ is continuous from $\mathbb{R}_+$ into $E_\triangle$, where $E_\triangle$ is endowed with the one-point compactification topology. A continuous Hunt process is called a *diffusion*.

A Dirichlet space $(\mathcal{E}, \mathcal{F})$ on $L^2(E, \alpha)$ is said to be *regular* if it has a core (see [11, p. 6]), i.e. a subset $\mathcal{C} \subset C_c(E) \cap \mathcal{F}$ which is dense in $C_c(E)$ for the uniform norm and dense in $\mathcal{F}$ for the norm $\|\varphi\| = [\int_E \varphi^2 \, \mathrm{d}\alpha + \mathcal{E}(\varphi, \varphi)]^{1/2}$.

Two regular Dirichlet spaces $(\mathcal{E}, \mathcal{F})$ and $(\mathcal{E}', \mathcal{F}')$ such that $\mathcal{E}(\varphi, \varphi) = \mathcal{E}'(\varphi, \varphi)$ for all φ in a common core $\mathcal{C}$ are necessarily equal, i.e. $\mathcal{F} = \mathcal{F}'$ and $\mathcal{E} = \mathcal{E}'$. This follows from the fact that by definition (see [11, p. 5]), a Dirichlet space is closed.

We say that a Borel set A of E is $(P_t)_{t\geq 0}$*-invariant* if for all $\varphi \in L^2(E, \alpha)$ and all $t > 0$ we have $P_t(\mathbf{1}_A \varphi) = \mathbf{1}_A P_t \varphi$ α-a.e. (see [11, p. 53]). Following [11, p. 55], we say that $(\mathcal{E}, \mathcal{F})$ is *irreducible* if for every $(P_t)_{t\geq 0}$-invariant set A, we have either $\alpha(A) = 0$ or $\alpha(E \setminus A) = 0$.

We say that $(\mathcal{E}, \mathcal{F})$ is *recurrent* if for all nonnegative $\varphi \in L^1(E, \alpha)$ and α-a.e. $y \in E$, we have $\mathbb{E}_y[\int_0^\infty \varphi(Y_s)\mathrm{d}s] \in \{0, \infty\}$ (see [11, p. 55]).

We finally say that $(\mathcal{E}, \mathcal{F})$ is *transient* if for all nonnegative $\varphi \in L^1(E, \alpha)$ and α-a.e. $y \in E$, we have $\mathbb{E}_y[\int_0^\infty \varphi(Y_s) \, \mathrm{d}s] < \infty$, with the convention that $\varphi(\triangle) = 0$ (see [11, p. 55]).

By [11, Lemma 1.6.4, p. 55], if $(\mathcal{E}, \mathcal{F})$ is irreducible, then it is either recurrent or transient.

A Borel set $\mathcal{N} \subset E$ is *properly exceptional* if $\alpha(\mathcal{N}) = 0$ and $\mathbb{P}_y(\exists t \geq 0 : Y_t \in \mathcal{N}) = 0$ for all $y \in E \setminus \mathcal{N}$ (see [11, p. 153]). A property is said to hold true *quasi-everywhere* if it holds true outside a properly exceptional set.

Remark B.1. Two Hunt processes with the same Dirichlet space share the same quasi-everywhere notion, up to the restriction that the capacity of every compact set is finite, which is always the case in the present work.

Proof. We fix a Hunt process $\mathbb{Y}$ and explain why its quasi-everywhere notion depends only on its Dirichlet space. A set $\mathcal{N} \subset E$ is exceptional (see [11, p. 152]) if there exists a Borel set $\tilde{\mathcal{N}}$ such that $\mathcal{N} \subset \tilde{\mathcal{N}}$ and $\mathbb{P}_y(\exists t \geq 0 : Y_t \in \tilde{\mathcal{N}}) = 0$ for α-a.e. $y \in E$. A properly exceptional set is clearly exceptional, and [11, Theorem 4.1.1, p. 155] tells us that any exceptional set is included in a properly exceptional set. Thus, a property is true quasi-everywhere if and only if it holds true outside an exceptional set. Next, [11, Theorem 4.2.1 (ii), p. 161] tells us that a set $\mathcal{N}$ is exceptional if and only if its capacity is 0, where the capacity of $\mathcal{N} \subset E$ is entirely defined from the Dirichlet space. And for [11, Theorem 4.2.1 (ii), p. 161] to apply, one needs that the capacity of all compact sets is finite. ■

B.2. Toolbox

We start with martingales.

Lemma B.2. *Let E be a locally compact separable metrizable space endowed with a Radon measure α such that* $\operatorname{supp} \alpha = E$, *and* $(\Omega, \mathcal{M}, (Z_t)_{t\geq 0}, (\mathbb{P}_z)_{z\in E_\triangle})$ *an α-symmetric*

$E_\triangle$-valued diffusion with regular Dirichlet space $(\mathcal{E},\mathcal{F})$ on $L^2(E,\alpha)$ and generator $(\mathcal{A},\mathcal{D}_\mathcal{A})$. Assume that $\varphi: E \to \mathbb{R}$ belongs to $\mathcal{D}_\mathcal{A}$, and both φ and $\mathcal{A}\varphi$ are bounded. Define

$$M_t^\varphi = \varphi(Z_t) - \varphi(Z_0) - \int_0^t \mathcal{A}\varphi(Z_s)\,\mathrm{d}s,$$

with the convention $\varphi(\triangle) = \mathcal{A}\varphi(\triangle) = 0$. Quasi-everywhere, $(M_t^\varphi)_{t\geq 0}$ is a $\mathbb{P}_z$-martingale in the canonical filtration of $(Z_t)_{t\geq 0}$.

This can be found in [11, p. 332]. There the assumption on φ is that there is f bounded and measurable such that $\varphi = R_1 f$, i.e. $\varphi = (I-\mathcal{A})^{-1} f$, which simply means that $\varphi - \mathcal{A}\varphi$ is bounded. Also, the conclusion is that $(M_t^\varphi)_{t\geq 0}$ is a MAF, which indeed implies that $(M_t^\varphi)_{t\geq 0}$ is a martingale (see [11, p. 243]).

Next, we deal with time-changes.

Lemma B.3. *Let E be a C^∞-manifold, α a Radon measure on E such that* $\operatorname{supp}\alpha = E$, *and $(\Omega, \mathcal{M}, (Z_t)_{t\geq 0}, (\mathbb{P}_z)_{z\in E_\triangle})$ an α-symmetric $E_\triangle$-valued diffusion with regular Dirichlet space $(\mathcal{E},\mathcal{F})$ on $L^2(E,\alpha)$ with core $C_c^\infty(E)$. Fix $g: E \to (0,\infty)$ continuous and let $g(\triangle) = 0$ by convention. Consider the time-change $A_t = \int_0^t g(Z_s)\,\mathrm{d}s$ and its generalized inverse $\rho_t = \inf\{s > 0 : A_s > t\}$. Set*

$$Y_t = Z_{\rho_t}\mathbf{1}_{\{\rho_t<\infty\}} + \triangle\mathbf{1}_{\{\rho_t=\infty\}}.$$

Then $(\Omega, \mathcal{M}, (Y_t)_{t\geq 0}, (\mathbb{P}_y)_{y\in E_\triangle})$ is a $g\alpha$-symmetric $E_\triangle$-valued diffusion with Dirichlet space $(\mathcal{E},\mathcal{F}')$ on $L^2(E, g\alpha)$, regular with core $C_c^\infty(E)$, i.e. $\mathcal{F}'$ is the closure of $C_c^\infty(E)$ with respect to the norm $[\int_E \varphi^2 g\,\mathrm{d}\alpha + \mathcal{E}(\varphi,\varphi)]^{1/2}$.

Remark B.4. If we apply the preceding result to the simple case where E is an open subset of $\mathbb{R}^d$ and where $\mathcal{E}(\varphi,\varphi) = \int_{\mathbb{R}^d} \|\nabla\varphi\|^2\,\mathrm{d}\alpha$ for all $\varphi \in C_c^\infty(E)$, then when $\mathcal{E}$ is seen as the Dirichlet form of a $g\alpha$-symmetric process, it may be better understood as $\mathcal{E}(\varphi,\varphi) = \int_{\mathbb{R}^d} \|g^{-1/2}\nabla\varphi\|^2 g\,\mathrm{d}\alpha$.

Lemma B.3 is nothing but a particular case of [11, Theorem 6.2.1, p. 316] (see also the few pages before). We only have to check that the Revuz measure in our case is $g\alpha$, i.e. (see [11, (5.1.13), p. 229]) for all bounded nonnegative measurable functions φ, ψ on E and all $t > 0$,

$$\int_E \mathbb{E}_x\Big[\int_0^t \varphi(Z_s)g(Z_s)\,\mathrm{d}s\Big]\psi(x)\,\alpha(\mathrm{d}x) = \int_0^t \int_E (P_s^Z\psi)\varphi g\,\mathrm{d}\alpha,$$

where P_t^Z is the semigroup of Z. The left hand side equals $\int_0^t\int_E P_s^Z(\varphi g)\psi\,\mathrm{d}\alpha$, so that the claim is obvious since Z is α-symmetric.

The following concatenation result can be found in Li–Ying [17, Proposition 3.2].

Lemma B.5. *Let E_V, E_W be C^∞-manifolds and let α_V, α_W be Radon measures on E_V and E_W such that* $\operatorname{supp}\alpha_V = E_V$ *and* $\operatorname{supp}\alpha_W = E_W$. *Consider an α_V-symmetric $(E_V \cup \{\triangle\})$-valued diffusion $(\Omega^V, \mathcal{M}^V, (V_t)_{t\geq 0}, (\mathbb{P}_v^V)_{v\in E_V\cup\{\triangle\}})$ with* *regular Dirichlet space* $(\mathcal{E}^V, \mathcal{F}^V)$ *on* $L^2(E_V, \alpha_V)$ *with core* $C_c^\infty(E_V)$*, and an* α_W*-symmetric* $(E_W \cup \{\triangle\})$*-valued diffusion* $(\Omega^W, \mathcal{M}^W, (W_t)_{t\geq 0}, (\mathbb{P}_w^W)_{w\in E_W\cup\{\triangle\}})$ *with regular Dirichlet space* $(\mathcal{E}^W, \mathcal{F}^W)$ *on* $L^2(E_W, \alpha_W)$ *with core* $C_c^\infty(E_W)$*. Introduce the measure* $\alpha = \alpha_V \otimes \alpha_W$ *on* $E = E_V \times E_W$*. By convention,* $(v, \triangle) = (\triangle, w) = (\triangle, \triangle) = \triangle$ *for all* $v \in E_V$ *and* $w \in E_W$*. Moreover, set* $\mathcal{M}^{(V,W)} = \sigma(\{(V_t, W_t) : t \geq 0\})$ *and define* $\mathbb{P}^{(V,W)}_{(v,w)} = \mathbb{P}_v^V \otimes \mathbb{P}_w^W$ *if* $(v, w) \in E_V \times E_W$*, and* $\mathbb{P}^{(V,W)}_\triangle = \mathbb{P}_\triangle^V \otimes \mathbb{P}_\triangle^W$*. The process*

$$\big(\Omega^V \times \Omega^W, \mathcal{M}^{(V,W)}, (V_t, W_t)_{t\geq 0}, (\mathbb{P}^{(V,W)}_{(v,w)})_{(v,w)\in(E_V\times E_W)\cup\{\triangle\}}\big)$$

is an $E_\triangle$*-valued* α*-symmetric diffusion, with regular Dirichlet space* $(\mathcal{E}, \mathcal{F})$ *on* $L^2(E, \alpha)$ *with core* $C_c^\infty(E)$ *and, for* $\varphi \in C_c^\infty(E)$*,*

$$\mathcal{E}(\varphi, \varphi) = \int_{E_V} \mathcal{E}^W(\varphi(v, \cdot), \varphi(v, \cdot))\, \alpha_V(\mathrm{d}v) + \int_{E_W} \mathcal{E}^V(\varphi(\cdot, w), \varphi(\cdot, w))\, \alpha_W(\mathrm{d}w).$$

Observe that $\mathcal{M}^{(V,W)}$ may be strictly smaller than $\mathcal{M}^V \otimes \mathcal{M}^W$ due to the identification of all the cemetery points. Also, actually $\mathbb{P}_\triangle^V \otimes \mathbb{P}_w^W = \mathbb{P}_v^V \otimes \mathbb{P}_\triangle^W = \mathbb{P}_\triangle^V \otimes \mathbb{P}_\triangle^W$ on $\mathcal{M}^{(V,W)}$ so that the choice $\mathbb{P}^{(V,W)}_\triangle = \mathbb{P}_\triangle^V \otimes \mathbb{P}_\triangle^W$ is arbitrary but legitimate.

The following killing result is a summary, adapted to our context, of [11, Theorems 4.4.2 and 4.4.3 (i), pp. 173–174].

Lemma B.6. *Let* E *be a* C^∞*-manifold,* α *be a Radon measure on* E *with* $\operatorname{supp}\alpha = E$*, and* $(\Omega, \mathcal{M}, (Z_t)_{t\geq 0}, (\mathbb{P}_z)_{z\in E_\triangle})$ *be an* α*-symmetric* $E_\triangle$*-valued diffusion with regular Dirichlet space* $(\mathcal{E}, \mathcal{F})$ *on* $L^2(E, \alpha)$ *with core* $C_c^\infty(E)$*. Let* O *be an open subset of* E *and consider* $\tau_O = \inf\{t \geq 0 : X_t \notin O\}$*, with the convention that* $\inf \emptyset = \infty$*. Set*

$$Z_t^O = Z_t \mathbf{1}_{\{t<\tau_O\}} + \triangle \mathbf{1}_{\{t\geq\tau_O\}}.$$

Then $(\Omega, \mathcal{M}, (Z_t^O)_{t\geq 0}, (\mathbb{P}_z)_{z\in O\cup\{\triangle\}})$ *is an* $\alpha|_O$*-symmetric* $O \cup \{\triangle\}$*-valued diffusion with regular Dirichlet space* $(\mathcal{E}_O, \mathcal{F}_O)$ *on* $L^2(O, \alpha|_O)$ *with core* $C_c^\infty(O)$ *and for* $\varphi \in \mathcal{F}_O$*,*

$$\mathcal{E}_O(\varphi, \varphi) = \mathcal{E}(\varphi, \varphi).$$

Note that since O is an open subset of the manifold E and since the Hunt process is continuous, the regularity condition (4.4.6) of [11, Theorem 4.4.2, p. 173] is obviously satisfied.

We finally give an adaptation of the Girsanov theorem in the context of Dirichlet spaces, which is a particular case of Chen–Zhang's [5, Theorem 3.4].

Lemma B.7. *Let* E *be an open subset of* $\mathbb{R}^d$ *with* $d \geq 1$*,* α *be a Radon measure on* E *with* $\operatorname{supp}\alpha = E$*, and* $(\Omega, \mathcal{M}, (Z_t)_{t\geq 0}, (\mathbb{P}_z)_{z\in E_\triangle})$ *be an* α*-symmetric* $E_\triangle$*-valued diffusion with regular Dirichlet space* $(\mathcal{E}, \mathcal{F})$ *on* $L^2(E, \alpha)$ *with core* $C_c^\infty(E)$ *such that for all* $\varphi \in C_c^\infty(E)$*,*

$$\mathcal{E}(\varphi, \varphi) = \int_E \|\nabla\varphi\|^2\, \mathrm{d}\alpha.$$

Let $(\mathcal{A}, \mathcal{D}_{\mathcal{A}})$ *stand for its generator. Let* $u \in \mathcal{F}$ *be bounded and such that for* $\varrho = e^u$*, we have* $\varrho - 1 \in \mathcal{D}_{\mathcal{A}}$ *with* $\mathcal{A}[\varrho - 1]$ *bounded. Set*

$$L_t^\varrho = \frac{\varrho(Z_t)}{\varrho(Z_0)} \exp\left(-\int_0^t \frac{\mathcal{A}[\varrho - 1](Z_s)}{\varrho(Z_s)}\,\mathrm{d}s\right),$$

with the conventions that $\varrho(\triangle) = 1$ *and* $\mathcal{A}[\varrho - 1](\triangle) = 0$.

Assume that ϱ *is continuous on* $E_\triangle$*. Then quasi-everywhere,* $(L_t^\varrho)_{t\geq 0}$ *is a bounded* $(\mathcal{M}_t)_{t\geq 0}$*-martingale under* $\mathbb{P}_z$*, where we have set* $\mathcal{M}_t = \sigma(\{Z_s : s \in [0,t]\})$*, and there exists a probability measure* $\tilde{\mathbb{P}}_z$ *on* $(\Omega, \mathcal{M})$ *such that for all* $t > 0$*,* $\tilde{\mathbb{P}}_z = L_t^\varrho \cdot \mathbb{P}_z$ *on* $\mathcal{M}_t$.

Moreover, $(\Omega, \mathcal{M}, (Z_t)_{t\geq 0}, (\tilde{\mathbb{P}}_z)_{z\in E_\triangle})$ *is a* $\varrho^2\alpha$*-symmetric* $E_\triangle$*-valued diffusion with regular Dirichlet space* $(\tilde{\mathcal{E}}, \mathcal{F})$ *on* $L^2(E, \varrho^2\alpha)$ *such that for all* $\varphi \in \mathcal{F}$,

$$\tilde{\mathcal{E}}(\varphi, \varphi) = \frac{1}{2}\int_E \|\nabla\varphi\|^2 \varrho^2\,\mathrm{d}\alpha.$$

Actually, Chen and Zhang speak of *right processes* in [5], but this is not an issue since we only consider continuous Hunt processes. Also, they assume that L^ϱ is bounded from above and from below by some deterministic constants, on each compact time interval, but this is obvious under our assumptions on u and $\mathcal{A}\varrho$. Finally, their expression of L^ϱ is different (see [5, pp. 485–486]): First, they define M_t^ϱ as the martingale part of $\varrho(X_t)$. By Lemma B.2 (applied to $\varrho - 1$), we see that

$$M_t^\varrho = \varrho(Z_t) - \varrho(Z_0) - \int_0^t \mathcal{A}[\varrho - 1](Z_s)\,\mathrm{d}s.$$

Then they put $M_t = \int_0^t [\varrho(Z_s)]^{-1}\,\mathrm{d}M_s^\varrho$ and define L^ϱ as $L_t^\varrho = \exp(M_t - \frac{1}{2}\langle M\rangle_t)$. But by Itô's formula, we have

$$\log\varrho(Z_t) = \log\varrho(Z_0) + \int_0^t [\varrho(Z_s)]^{-1}\,\mathrm{d}M_s^\varrho + \int_0^t [\varrho(Z_s)]^{-1}\mathcal{A}[\varrho - 1](Z_s)\,\mathrm{d}s$$
$$- \frac{1}{2}\int_0^t [\varrho(Z_s)]^{-2}\,\mathrm{d}\langle M^\varrho\rangle_s,$$

whence $\log\varrho(Z_t) = \log\varrho(Z_0) + M_t + \int_0^t [\varrho(Z_s)]^{-1}\mathcal{A}[\varrho - 1](Z_s)\mathrm{d}s - \frac{1}{2}\langle M\rangle_t$, so that

$$L_t^\varrho = \exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right) = [\varrho(Z_0)]^{-1}\varrho(Z_t)\exp\left(-\int_0^t \varrho(Z_s)^{-1}\mathcal{A}[\varrho - 1](Z_s)\,\mathrm{d}s\right)$$

as desired.

Acknowledgements. The fruitful comments of the two referees helped us to substantially improve the presentation of the paper.

References

[1] Arumugam, G., Tyagi, J.: Keller–Segel chemotaxis models: a review. Acta Appl. Math. **171**, article no. 6, 82 pp. (2021) Zbl 1464.35001 MR 4188348

[2] Blanchet, A., Dolbeault, J., Perthame, B.: Two-dimensional Keller–Segel model: optimal critical mass and qualitative properties of the solutions. Electron. J. Differential Equations **2006**, article no. 44, 32 pp. Zbl 1112.35023 MR 2226917

[3] Bresch, D., Jabin, P.-E., Wang, Z.: On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak–Keller–Segel model. C. R. Math. Acad. Sci. Paris **357**, 708–720 (2019) Zbl 1428.35617 MR 4018082

[4] Cattiaux, P., Pédèches, L.: The 2-D stochastic Keller–Segel particle model: existence and uniqueness. ALEA Latin Amer. J. Probab. Math. Statist. **13**, 447–463 (2016) Zbl 1343.60122 MR 3519253

[5] Chen, Z.-Q., Zhang, T.-S.: Girsanov and Feynman–Kac type transformations for symmetric Markov processes. Ann. Inst. H. Poincaré Probab. Statist. **38**, 475–505 (2002) Zbl 1004.60077 MR 1914937

[6] Dolbeault, J., Schmeiser, C.: The two-dimensional Keller–Segel model after blow-up. Discrete Contin. Dynam. Systems **25**, 109–121 (2009) Zbl 1180.35131 MR 2525170

[7] Fatkullin, I.: A study of blow-ups in the Keller–Segel model of chemotaxis. Nonlinearity **26**, 81–94 (2013) Zbl 1322.35057 MR 3001763

[8] Fournier, N., Hauray, M., Mischler, S.: Propagation of chaos for the 2D viscous vortex model. J. Eur. Math. Soc. **16**, 1423–1466 (2014) Zbl 1299.76040 MR 3254330

[9] Fournier, N., Jourdain, B.: Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. Ann. Appl. Probab. **27**, 2807–2861 (2017) Zbl 1447.65106 MR 3719947

[10] Fukushima, M.: From one dimensional diffusions to symmetric Markov processes. Stochastic Process. Appl. **120**, 590–604 (2010) Zbl 1221.60113 MR 2603055

[11] Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet forms and symmetric Markov processes. De Gruyter Stud. Math. 19, Walter de Gruyter, Berlin (2011) Zbl 1227.31001 MR 2778606

[12] Haškovec, J., Schmeiser, C.: Stochastic particle approximation for measure valued solutions of the 2D Keller–Segel system. J. Statist. Phys. **135**, 133–151 (2009) Zbl 1173.82021 MR 2505729

[13] Haškovec, J., Schmeiser, C.: Convergence of a stochastic particle approximation for measure solutions of the 2D Keller–Segel system. Comm. Partial Differential Equations **36**, 940–960 (2011) Zbl 1229.35111 MR 2765424

[14] Jabir, J.-F., Talay, D., Tomašević, M.: Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller–Segel model without smoothing. Electron. Commun. Probab. **23**, article no. 84, 14 pp. (2018) Zbl 1402.60129 MR 3873791

[15] Karatzas, I., Shreve, S. E.: Brownian motion and stochastic calculus. 2nd ed., Grad. Texts in Math. 113, Springer, New York (1991) Zbl 0734.60060 MR 1121940

[16] Keller, E. F., Segel, L. A.: Initiation of slime mold aggregation viewed as an instability. J. Theoret. Biol. **26**, 399–415 (1970) Zbl 1170.92306 MR 3925816

[17] Li, L., Ying, J.: Regular subspaces of Dirichlet forms. In: Festschrift Masatoshi Fukushima, Interdiscip. Math. Sci. 17, World Scientific Publishing, Hackensack, NJ, 397–420 (2015) Zbl 1336.31028 MR 3379351

[18] Olivera, C., Richard, A., Tomašević, M.: Quantitative particle approximation of nonlinear Fokker–Planck equations with singular kernel. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **24**, 691–749 (2023) Zbl 07741032 MR 4630042

[19] Osada, H.: Propagation of chaos for the two-dimensional Navier–Stokes equation. Proc. Japan Acad. Ser. A Math. Sci. **62**, 8–11 (1986) Zbl 0679.76033 MR 0839794

[20] Patlak, C. S.: Random walk with persistence and external bias. Bull. Math. Biophys. **15**, 311–338 (1953) Zbl 1296.82044 MR 0081586

[21] Revuz, D., Yor, M.: Continuous martingales and Brownian motion. 3rd ed., Grundlehren Math. Wiss. 293, Springer, Berlin (2005) Zbl 1087.60040 MR 1725357

[22] Stevens, A.: The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. SIAM J. Appl. Math. **61**, 183–212 (2000) Zbl 0963.60093 MR 1776393

[23] Suzuki, T.: Exclusion of boundary blowup for 2D chemotaxis system provided with Dirichlet boundary condition for the Poisson part. J. Math. Pures Appl. (9) **100**, 347–367 (2013) Zbl 1285.35121 MR 3095205

[24] Velázquez, J. J. L.: Point dynamics in a singular limit of the Keller–Segel model. I. Motion of the concentration regions. SIAM J. Appl. Math. **64**, 1198–1223 (2004) Zbl 1058.35021 MR 2068667

[25] Velázquez, J. J. L.: Point dynamics in a singular limit of the Keller–Segel model. II. Formation of the concentration regions. SIAM J. Appl. Math. **64**, 1224–1248 (2004) Zbl 1058.35022 MR 2068668