Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part II, below the critical regime

C. Erignoux, P. Gonçalves and G. Nahum

Clément Erignoux, Dip. Matematica e Fisica, Univ. Degli Studi Roma Tre, Largo San Leonardo Murialdo, 1, 00146 Roma, Italy
E-mail address: `clement.erignoux@gmail.com`

Patrícia Gonçalves, Center for Mathematical Analysis, Geometry and Dynamical Systems, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.
E-mail address: `pgoncalves@tecnico.ulisboa.pt`

Gabriel Nahum, Center for Mathematical Analysis, Geometry and Dynamical Systems, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.
E-mail address: `gabriel.nahum@tecnico.ulisboa.pt`

Abstract. The purpose of this article is to provide a simple proof of the hydrodynamic and hydrostatic behavior of the SSEP in contact with reservoirs which inject and remove particles in a finite size windows at the extremities of the bulk. More precisely, the reservoirs inject/remove particles at/from any point of a window of size K placed at each extremity of the bulk and particles are injected/removed to the first open/occupied position in that window. The reservoirs have slow dynamics, in the sense that they intervene at speed $N^{-\theta}$ w.r.t. the bulk dynamics. In the first part of this article, we treated the case $\theta > 1$ for which the entropy method can be adapted. We treat here the case where the boundary dynamics is too fast for the Entropy Method to apply. We prove using duality estimates inspired by previous work that the hydrodynamic limit is given by the heat equation with Dirichlet boundary conditions, where the density at the boundaries is fixed by the parameters of the model.

1. Introduction

We consider a 1-dimensional boundary driven lattice gas, whose boundary dynamics is non-reversible with respect to product measures, and which is slowed down by an extra factor $N^{-\theta}$ with respect to the diffusive scaling N^2 of the bulk SSEP dynamics. In the first part of this article Erignoux et al. (2020+), the constant θ is assumed to be larger or equal to 1, which allowed us to adapt the classical entropy method to our non-reversible dynamics, to derive both the hydrodynamic and hydrostatic limits under suitable technical assumptions. In the case $\theta = 1$, we show that both scaling limits exhibit so-called non-linear Robin boundary conditions, whereas in the $\theta > 1$ regime, the boundary dynamics is to slow to be visible in the diffusive time scale and the hydrodynamic limit is ruled by the heat equation with Neumann boundary conditions.

In the case $0 < \theta < 1$, however, the non-reversible boundary dynamics generates entropy w.r.t. equilibrium product measure at a fast rate, so that the classical entropy estimates are not sharp enough to derive the hydrodynamics nor the hydrostatic limit. For this reason, we instead adapt the tools exploited in Erignoux et al. (2018); Erignoux (2018) (non-reversible dynamics, diffusive scaling of the boundary dynamics) to our non-reversible slowed down case, to prove that the macroscopic behavior of the system started close to a given profile f_0 is ruled by the heat equation with Dirichlet boundary conditions

$$
\begin{cases}
\partial_t \rho_t(u) = \Delta\, \rho_t(u) & \text{for } t \geqslant 0,\ u \in (0,1) \\
\rho_t(0) = \alpha & \text{for } t > 0 \\
\rho_t(1) = \alpha' & \text{for } t > 0 \\
\rho_0(\cdot) = f_0(\cdot)
\end{cases}
,
$$

and that the macroscopic stationary profile is given by the linear interpolation $\rho^*(u) = \alpha + u(\alpha' - \alpha)$. The general strategy, both for the hydrostatics and the hydrodynamics described above, is to directly estimate the discrete density field ρ^N defined in (2.7) and the 2-point correlation field φ^N defined in (2.8), which both solve discrete difference equations. Due to the non-reversible dynamics, however, these equations are not closed. In Erignoux et al. (2018); Erignoux (2018) (in the regime $\theta = 0$), this difficulty is solved by artificially closing the equations with unknown boundary terms, and then estimating the missing terms by using duality. In De Masi et al. (2012), De Masi et al. treated a non-reversible boundary dynamics in the slow scaling, corresponding here to the case $\theta = 1$, and instead completely estimate the cascading n-points correlation field. They prove that for n large enough, those correlation field vanish, and this is sufficient to backtrack those estimates to the initial 2-points correlation field. This involves significant technical and phenomenological difficulties which can be overcome in the case $\theta \geqslant 1$ without so much effort. In fact, their proof can likely be extended to the case $\theta > 1/2$ with a bit more work, but we do not pursue this issue here since our proof can cover all the cases $\theta \in (0,1)$.

In the present article, we bridge the gap between Erignoux et al. (2018); Erignoux (2018) (case $\theta = 0$), and De Masi et al. (2012) (case $\theta = 1$), and consider the case $0 < \theta < 1$. In order to avoid to try and obtain the same type of delicate n-points correlation estimates used in De Masi et al. (2012), we adapt the strategy of Erignoux et al. (2018); Erignoux (2018), and under the assumption (H1) below, we are

able to artificially close the discrete difference equations for the density and for the 2-points correlation field. More precisely, we consider a boundary dynamics with two distinct mechanisms, one representing the system being put in contact with a infinite equilibrium reservoir (occurring at rate r), and a second non-reversible creation/annihilation process whose rates $b(\eta)$ depend on the local configuration at the boundary. The key point of the proof is the study of a branching process, representing the interaction of a site whose value is currently unknown with the rest of the system. Whenever a reservoir is queried (rate r), one of its branches dies, whereas when the non-reversible event occurs, the process branches out. Assumption (H1) allows us, roughly speaking, to close the equation, by ensuring that this dual branching process eventually dies out, thus allowing us to determine the value of the unknown site. Note that the techniques developed here are not specific to the choice of the boundary dynamics made in this article, and can be applied to any local non-reversible dynamics under an assumption analogous to (H1).

Because of the slow boundary, unlike in Erignoux et al. (2018); Erignoux (2018) the dual branching process determining the value of the density filed at the boundary self-decorrelates. As a consequence, one obtains explicit formulae (cf. (2.3)) for the macroscopic Dirichlet boundary conditions α and α' appearing in the hydrodynamic limit (1). Because of this self-decorrelation property, one could expect that, unlike in the case $\theta = 0$, assumption (H1) could possibly be dismissed. This, however, would involve fundamental change in the dual process at the core of our method, and therefore significant further difficulties in the current proof, it is thus left as an open problem at this point. Another natural question is that of fluctuations around equilibrium for non-reversible boundary dynamics. Within the current state of the art however, deriving the equilibrium fluctuations requires sharp estimates on the 2-point space-time correlation field at the boundaries of the system, and it is not clear whether such estimates actually hold. In any case, such an estimate is not achievable with the technique employed in the current work.

As does the one laid out in De Masi et al. (2012), our approach works in the case $\theta \geqslant 1$, although technical changes would be required to account for the modified boundary conditions. Let us comment briefly on the the boundary densities derived in (2.3). In the case of a single reservoir with $0 < \theta < 1$, the current due to reservoir interaction is on a faster time scale than the particle current from the boundary to the bulk, so that the boundary of the system thermalizes immediately at the density which makes the reservoir current vanish, i.e. that of the reservoir. In our case, site 2 is affected by a non-reversible dynamics, but the overall boundary dynamics is still faster than the bulk dynamics, so that the system immediately enforces a boundary density at which the reservoir and non-reversible current cancel each other out. This condition is precisely given by (2.3). In the case $\theta = 1$, the boundary currents operate on the same time scale as the bulk current, so that the boundary density takes the value as with the stirring current cancels out the two others. Finally, for $\theta > 1$, the boundary dynamics is slow, so that in the hydrodynamic limit, the boundary stirring currents cancel out, thus yielding Neumann boundary conditions.

2. Model and main result

Let N be a scaling parameter and denote by $\Lambda_N = \{1, \dots, N-1\}$ the bulk of the system. We consider a continuous time Markov process $\{\eta_t : t \geqslant 0\}$, with state

space $\Omega_N := \{0,1\}^{\Lambda_N}$. We denote η the configurations, i.e. elements of Ω_N where $\eta(x) = 0$ means that the site x is vacant while $\eta(x) = 1$ means that the site x is occupied. The process $\{\eta_t : t \geqslant 0\}$ is driven by the infinitesimal generator

$$\mathcal{L}_N = N^2 \mathcal{L}_{N,0} + N^{2-\theta} \left(\mathcal{L}_{N,l} + \mathcal{L}_{N,r} \right), \tag{2.1}$$

where $\theta \in (0,1)$. The parts of this generator act on functions $f : \Omega_N \to \mathbb{R}$ as

$$(\mathcal{L}_{N,0}f)(\eta) = \sum_{x=1}^{N-2} \left(f(\eta^{x,x+1}) - f(\eta) \right) \tag{2.2}$$

and

$$(\mathcal{L}_{N,l}f)(\eta) = c_{l,1}(\eta)\left(f(\eta^1) - f(\eta) \right) + c_{l,2}(\eta)\left(f(\eta^2) - f(\eta) \right),$$

$$(\mathcal{L}_{N,r}f)(\eta) = c_{r,1}(\eta)\left(f(\eta^{N-1}) - f(\eta) \right) + c_{r,2}(\eta)\left(f(\eta^{N-2}) - f(\eta) \right)$$

with rates

$$\begin{aligned}
c_{l,1}(\eta) &= r \left[\bar{\rho}(1 - \eta(1)) + (1 - \bar{\rho})\eta(1) \right], \\
c_{r,1}(\eta) &= r' \left[\bar{\rho}'(1 - \eta(N-1)) + (1 - \bar{\rho}')\eta(N-1) \right], \\
c_{l,2}(\eta) &= c \left[\eta(1)(1 - \eta(2)) + (1 - \eta(1))\eta(2) \right] + b\eta(1)(1 - \eta(2)), \\
c_{r,2}(\eta) &= c' \left[\eta(N-1)(1 - \eta(N-2)) + (1 - \eta(N-1))\eta(N-2) \right] \\
&\quad + b'\eta(N-1)(1 - \eta(N-2)).
\end{aligned}$$

In the formulation above, we shortened

$$\eta^{x,y}(z) = \begin{cases} \eta(z), \ z \neq x, y \\ \eta(y), \ z = x \\ \eta(x), \ z = y \end{cases} \quad , \quad \text{and} \quad \eta^x(z) = \begin{cases} \eta(z), \ z \neq x \\ 1 - \eta(x), \ z = x \end{cases} \quad .$$

For convenience, we reparametrized the boundary generators w.r.t. Part I of this article (Erignoux et al., 2020+). To get back to the notations from Erignoux et al. (2020+), one can define

$$\begin{aligned}
\alpha_1 &= r\bar{\rho}, \quad \gamma_1 = r(1 - \bar{\rho}), \quad \alpha_2 = b + c, \quad \gamma_2 = c, \\
\beta_1 &= r'\bar{\rho}', \quad \delta_1 = r'(1 - \bar{\rho}'), \quad \beta_2 = b' + c', \quad \delta_2 = c'.
\end{aligned}$$

We assumed (using Part 1's notations) that $\alpha_2 \geqslant \gamma_2$ and $\beta_2 > \delta_2$. This is purely for convenience: if for example, $\alpha_2 \leqslant \gamma_2$, one would merely switch $\alpha_2 = c$, $\gamma_2 = b + c$, for which our entire proof holds as well. Throughout this article, r, b, c, r', b', c' and θ are constant, and even without specific mention, all quantities can depend on them. With these notations, we can reinterpret the boundary generator as follows: at rate r (resp. r'), site 1 (resp. $N-1$) is replaced by a Bernoulli with parameter $\bar{\rho}$ (resp. $\bar{\rho}'$). At rate c (resp. c'), site 2 (resp. $N-2$) becomes a copy of site 1 (resp. $N-1$), and finally, at rate b (resp. b'), site 2 (resp. $N-2$) is filled by a particle if site 1 (resp. $N-1$) is occupied, otherwise it is left unchanged.

The case $\theta \geqslant 1$ was treated in Erignoux et al. (2020+). We consider here the case where $\theta < 1$, i.e. the case where the reservoirs are strong. The boundary dynamics we chose does not admit product measures as stationary states, and because $\theta < 1$ they are not slowed down enough for the usual entropy estimates to hold. Instead, as in Erignoux et al. (2018), Erignoux (2018), we exploit duality estimates on random

walks, to write both the discrete profile and the two-point correlation function as solutions of "artificially" closed equations, and then estimate the boundary terms needed to close them. Our main assumption is

$$b < r \quad \text{and} \quad b' < r', \tag{H1}$$

which is analogous to assumption (2.13) in Erignoux et al. (2018). As briefly described in the introduction, this assumption is made in order for the dual branching process defined in Section 5.2 to eventually die. Define α and α' as the unique solutions in $[0,1]$ of the equations

$$r(\overline{\rho} - \alpha) + b\alpha(1 - \alpha) = 0 \quad \text{and} \quad r'(\overline{\rho}' - \alpha') + b'\alpha'(1 - \alpha') = 0. \tag{2.3}$$

Note that we have the explicit expressions

$$\alpha = f(r, b, \overline{\rho}), \quad \alpha' = f(r', b', \overline{\rho}'), \quad \text{where} \quad f(r, b, \overline{\rho}) = \frac{\sqrt{(r-b)^2 + 4br\overline{\rho}} + b - r}{2b},$$

and that the parameters c, c' play no role in these definitions.

Fix an initial profile $f_0 : [0,1] \to [0,1]$ in $C^1([0,1])$, and start the process $\{\eta_t : t \geqslant 0\}$ from the product measure

$$\mu_N(\eta) = \prod_{x \in \Lambda_N} \Big[f_0(x/N)\eta(x) + (1 - f_0(x/N))(1 - \eta(x)) \Big]$$

fitting the initial profile f_0. We denote by $\mathbb{P}_{\mu_N}$ the distribution of the process $\{\eta_t : t \geqslant 0\}$ started from the distribution μ_N and with infinitesimal generator given by (2.1). We are now ready to state our main result.

Theorem 2.1 (Hydrodynamic limit). *Under assumption* (H1), *for any $t \geqslant 0$, any Riemann-integrable function G, and any $\delta > 0$*

$$\lim_{N \to \infty} \mathbb{P}_{\mu_N} \left(\left| \frac{1}{N-1} \sum_{x \in \Lambda_N} G\left(\tfrac{x}{N}\right) \eta_t(x) - \int_0^1 G(u)\rho_t(u)du \right| > \delta \right) = 0,$$

where $\rho_t(\cdot)$ is the unique classical solution of the heat equation with Dirichlet boundary conditions α and α' solutions of (2.3):

$$\begin{cases} \partial_t \rho_t(u) = \Delta \rho_t(u) & \text{for } t \geqslant 0, \ u \in (0,1) \\ \rho_t(0) = \alpha & \text{for } t \geqslant 0 \\ \rho_t(1) = \alpha' & \text{for } t \geqslant 0 \\ \rho_0(\cdot) = f_0(\cdot) \end{cases} . \tag{2.4}$$

Remark 2.2 (Initial distribution). We chose the initial distribution to be a product measure fitting a smooth initial profile. This is mainly not to burden the proof, which would hold as well assuming that the correlations of the initial distribution decay uniformly,

$$\limsup_{\ell \to \infty} \limsup_{N \to \infty} \sup_{\substack{x,y \in \Lambda_N \\ |x-y| \geqslant \ell}} \mu_N \left[(\eta(x) - f_0(x))(\eta(y) - f_0(y)) \right] = 0.$$

The regularity assumption on the initial profile f_0 could also be weakened, but that would entail significant extra technical difficulties to prove Lemma 3.3 below.

Remark 2.3 (On assumption (H1)). As was already mentioned, Assumption (H1) is analogous to assumption (2.13) in Erignoux et al. (2018). The reason for this assumption is the following. In order to determine the value of η_t at the boundaries, we are able to write $\eta_t(3)$ as a function of its past, by following the sites that had an influence over its value. This allows us to write $\eta_t(3)$ as a function of a branching process, which branches at the boundaries at rate b, b' and dies at rate r, r'. Assumption (H1) ensures that this branching process ultimately dies out completely, and that we are able to determine the value of $\eta_t(3)$.

Remark 2.4 (The case $\theta = 0$). Note that under an assumption analogous to (2.13) in Erignoux et al. (2018), the case $\theta = 0$ is a consequence of Erignoux et al. (2018). However, unlike in the present article, in the case $\theta = 0$ correlations are introduced in the boundary dynamics, so that we no longer have an explicit expression for the boundary densities α and α'. This is due to the fact that the distribution of the determination tree (cf. Section 5.4) is no longer close to a Galton-Watson distribution, i.e. Lemma 5.5 no longer holds.

Remark 2.5 (Heuristics on (2.3)). Roughly speaking, (2.3) is an equilibrium formula for the boundary dynamics, in which the stirring jump rate is infinite (which is formally the case in the limit $N \to \infty$, since the boundary dynamics is slowed down). Unlike in the case $\theta = 0$, the boundary sites decorrelate, and formally letting $\alpha = \mathbb{E}(\eta_1) = \mathbb{E}(\eta_2) = \mathbb{E}(\eta_3) = \ldots$ in the limit $N \to \infty$, and then using $\mathbb{E}(\mathcal{L}_N \eta_1) = \mathbb{E}(\mathcal{L}_N \eta_2) = 0$ yields (2.3). This decorrelation at the boundary, due to the slowed down dynamics, is the reason why an explicit formula for the boundary density can be achieved.

Remark 2.6 (General jump rates). The proof we present for Theorem 2.1 is not specific to these jump rates. It would actually hold for any perturbation of the flipping dynamics considered in Erignoux et al. (2018), (under an analogous assumption to Equation (2.13) in Erignoux et al., 2018). However, in order not to burden the proof and for consistency w.r.t. Erignoux et al. (2020+), we choose the boundary generator given in (2.2). Note that unlike in Erignoux et al. (2018), because of the slowed down boundary dynamics, an explicit expression for the boundary densities can be derived, however due to the construction of our dual process and its associated tree (cf. Section 5.4), obtaining an explicit formula in the general case would prove burdensome.

The Markov chain induced by the generator $(\mathcal{L})_N$ defined in (2.1) is irreducible on Ω_N. We will denote by μ_N^{ss} its unique stationary state. Given the duality techniques used to prove Theorem 2.1, the hydrostatic limit for this model is a straightforward adaption of the hydrodynamic limit, we state it but will not prove it. Instead, we refer the reader to Erignoux et al. (2018) for more details.

Theorem 2.7 (Hydrostatic limit). *Under assumption H1, for any Riemann integrable function G and any $\delta > 0$*

$$\lim_{N \to \infty} \mu_N^{ss} \left[\left| \frac{1}{N-1} \sum_{x \in \Lambda_N} G\left(\tfrac{x}{N}\right) \eta(x) - \int_0^1 G(u)\rho^*(u)du \right| > \delta \right] = 0,$$

where $\rho^*(\cdot)$ is harmonic with boundary conditions α and α' solutions of (2.3), i.e.:

$$\begin{cases} \Delta \rho^* = 0, \\ \rho^*(0) = \alpha \quad and \quad \rho^*(1) = \alpha', \end{cases} \tag{2.5}$$

so that $\rho^*(u)$ is linear connecting α and α'.

Throughout, we shorten

$$\widehat{\theta} := (1 - \theta)/2 \quad > 0. \tag{2.6}$$

For $x, y \in \Lambda_N$, $t \geqslant 0$, we further introduce the discrete density profile

$$\rho_t^N(x) = \mathbb{E}_{\mu_N}[\eta_t(x)] \tag{2.7}$$

and the two-point correlation function

$$\varphi_t^N(x, y) = \mathbb{E}_{\mu_N}\left[\left(\eta_t(x) - \rho_t^N(x)\right)\left(\eta_t(y) - \rho_t^N(y)\right)\right]. \tag{2.8}$$

The article is organized as follows. To derive the hydrodynamic equation, as in previous works (Erignoux et al., 2018; Erignoux, 2018), we estimate the discrete density profile (Section 3) and the two-point correlation (Section 4) function, namely ρ_t^N and φ_t^N by writing each one of them as a solution of a discrete difference equation with unknown boundary terms. One of the main difficulties w.r.t. Erignoux et al. (2018); Erignoux (2018) is to estimate those boundary terms, which require to refine the construction because of the slowed down boundary dynamics. Once this is done, proving the hydrodynamic limit is straightforward (cf. Section 4.1). The estimation of the boundary term is the purpose of Section 5. We will refer to Erignoux et al. (2018); Erignoux (2018) when the results are analogous, and detail the new contributions of this article.

3. Density field

In this section, we write the discrete profile ρ_t^N defined in (2.7) as an approximation of the solution of (2.4). More precisely, we have the following:

Proposition 3.1. *For any $t \geqslant 0$, and any Riemann integrable function $G \in C^2([0,1])$,*

$$\limsup_{N \to \infty}\left(\frac{1}{N}\sum_{x=3}^{N-3} G(x/N)\rho_t^N(x) - \int_{[0,1]} G(u)\rho_t(u)du\right)^2 = 0.$$

The main ingredient to prove Proposition 3.1 is the following Lemma, whose proof is postponed to Section 5.5 because it requires significant work.

Lemma 3.2. *Recall from Equation 2.3 the definition of α and α', there exists $\varepsilon > 0$ such that*

$$\sup_{t \geqslant N^{-\widehat{\theta}}}|\rho_t^N(3) - \alpha| = O(N^{-\varepsilon}) \quad and \quad \sup_{t \geqslant N^{-\widehat{\theta}}}|\rho_t^N(N-3) - \alpha'| = O(N^{-\varepsilon}).$$

The first identity is a consequence of Corollary 5.6 below. The second identity being strictly analogous, we will not prove it. Note that this result fixes the Dirichlet boundary condition of the discretized density profile ρ_t^N. We assume for now that Lemma 3.2 holds, and with it we will prove Proposition 3.1.

Proof of Proposition 3.1: We claim that for any $t > 0$, there exists $\varepsilon > 0$ such that

$$\sup_{x \in \{3,\ldots,N-3\}}|\rho_t^N(x) - \rho_t(x/N)| = O(N^{-\varepsilon}). \tag{3.1}$$

Applying Dynkin's identity $\partial_t \mathbb{E}_{\mu_N}[f(\eta_t)] = \mathbb{E}_{\mu_N}[\mathcal{L}_N f(\eta_t)]$, to $f(\eta) = \eta(x)$ for any $x \in \{4, \cdots, N-4\}$, elementary computations yield that $\rho_t^N(\cdot)$ is solution of the discrete difference diffusion equation

$$\begin{cases} \partial_t \rho_t^N = \Delta_N \rho_t^N, & \text{for } 4 \leqslant x \leqslant N-4 \\ \rho_t^N(3) = \alpha_{t,N}, \\ \rho_t^N(N-3) = \alpha'_{t,N}, \\ \rho_0^N = \rho_0(\cdot/N) \end{cases} \quad , \tag{3.2}$$

where we defined $\alpha_{t,N} = \rho_t^N(3)$ and $\alpha'_{t,N} = \rho_t^N(N-3)$, and

$$\Delta_N \rho^N(x) = N^2 \left[\rho^N(x+1) + \rho^N(x-1) - 2\rho^N(x) \right]$$

is the discrete Laplacian. Note that the two identities at the space boundaries are trivial, and "artificially" close the equation satisfied by ρ_t^N.

Denote by X_s (resp. B_s) a continuous time random walk on $\mathbb{Z}$ jumping at rate N^2 to each of its neighbors (resp. standard Brownian Motion on $\mathbb{R}$). We denote by H^X (resp. by H^B) the hitting time of the boundary $\{3, N-3\}$ (resp. $\{0,1\}$), and shorten $H_t^X = t \wedge H^X$ (resp. $H_t^B = t \wedge H^B$). We denote by $\mathbb{E}_x^X$, and $\mathbb{E}_u^B$ the expectations w.r.t. the distributions of both X and B starting from x and u, respectively. By Feynman-Kac's formula, we can write for $x \in \{3, \ldots, N-3\}$, $u \in [0,1]$

$$\rho_t^N(x) = \mathbb{E}_x^X \left[\rho_{t-H_t^X}^N(X_{H_t^X}) \right] \quad \text{and} \quad \rho_t(u) = \mathbb{E}_u^B \left[\rho_{t-H_t^B}(B_{H_t^B}) \right].$$

We can therefore write $|\rho_t^N(x) - \rho_t(x/N)|$ as the sum of three contributions, the first one corresponding to the case where $X_{H_t^X} = 3$ (resp. $B_{H_t^B} = 0$), the second $X_{H_t^X} = N-3$ (resp. $B_{H_t^B} = 1$), and the last one to the case $H_t^X = t$ (resp. $H_t^B = t$). We write these three contributions as

$$C_l(t,x) := \left| \mathbb{E}_x^X \left[\rho_{t-H_t^X}^N(3) \mathbf{1}_{\{X_{H_t^X}=3\}} \right] - \alpha \mathbb{P}_{x/N}^B \left(B_{H_t^B} = 0 \right) \right|,$$

$$C_r(t,x) := \left| \mathbb{E}_x^X \left[\rho_{t-H_t^X}^N(N-3) \mathbf{1}_{\{X_{H_t^X}=N-3\}} \right] - \alpha' \mathbb{P}_{x/N}^B \left(B_{H_t^B} = 1 \right) \right|,$$

$$C^*(t,x) := \left| \mathbb{E}_x^X \left[\rho_0^N(X_t) \mathbf{1}_{\{H_t^X=t\}} \right] - \mathbb{E}_{x/N}^B \left(\rho_0^N(B_t) \mathbf{1}_{\{H_t^B=t\}} \right) \right|.$$

Given theses notations, we have

$$|\rho_t^N(x) - \rho_t(x/N)| \leqslant C_l + C_r + C^*.$$

Recall (2.6). To estimate C_l and C_r, we consider two cases, depending on whether H_t^X is in $[0, t-N^{-\widehat{\theta}}]$ or in $(t-N^{-\widehat{\theta}}, t)$. In the first case, thanks to Lemma 3.2, $\rho_{t-H_t^X}^N(3) = \alpha + O(N^{-\varepsilon})$. More precisely, since ρ^N is less than 1 we can write

$$C_l \leqslant \left| \mathbb{P}_x^X \left(X_{H_t^X} = 3 \right) - \mathbb{P}_{x/N}^B \left(B_{H_t^B} = 0 \right) \right| + 2\mathbb{P}_x^X \left(H_t^X \in (t-N^{-\widehat{\theta}}, t) \right) + O(N^{-\varepsilon}).$$

The second term on the right-hand side above can be estimated according to Lemma 3.3, p.12 of Erignoux (2018), which yields

$$\mathbb{P}_x^X \left(H_t^X \in (t-N^{-\widehat{\theta}}, t) \right) = O \left(N^{-\widehat{\theta}} t^{-3/2} \right).$$

Finally, the first contribution , as well as C^*, can be estimated by using the Berry Essen inequality, i.e. approximating the random walk by a Brownian motion as in the proof of Lemma 3.3 of Erignoux (2018), to obtain for a positive constant $\varepsilon' > 0$

$$\left| \mathbb{P}_x^X \left(X_{H_t^X} = 3 \right) - \mathbb{P}_{x/N}^B \left(B_{H_t^B} = 0 \right) \right| = O(N^{-\varepsilon'})$$

and

$$C^*(t, x) = O(N^{-\varepsilon'}).$$

The bounds above can be taken independent of x. This proves (3.1) and the proposition. $\qquad\square$

Before turning to the two-point correlation function, we state the following technical Lemma, which will be needed in what follows.

Lemma 3.3. *For any* $\delta > 0$,

$$\sup_{\substack{\delta N \leqslant x \leqslant (1-\delta)N \\ t \geqslant 0}} |\rho_t^N(x+1) - \rho_t^N(x)| \leqslant O\left((\delta N)^{-1} \log N \right).$$

Proof of Lemma 3.3: Fix $\delta > 0$, and consider once again a random walk X_s on $\mathbb{Z}$, for which we denote $H_{\{3,N-3\}}$ the hitting time of the boundaries x=3, N-3. Recall that $\rho_t^N(x)$ is solution of (3.2), choose a sequence a_N that vanishes as $N \to \infty$, and define the stopping time

$$\tau = \tau_{\delta,N} := \delta^2 N^2 a_N \wedge H_{\{3,N-3\}}.$$

By the Feynman-Kac formula, we can write for any $3 \leqslant x \leqslant N - 3$

$$\rho_t^N(x) = \mathbb{E}_x \left[\rho_{t-t\wedge\tau}^N(X_{t\wedge\tau}) \right].$$

A symmetric random walk would typically require a time of order $\delta^2 N^2$ to move on a distance δN, therefore there exists a universal constant K such that for any $\delta N \leqslant x \leqslant (1-\delta)N$,

$$\mathbb{P}(\tau = H_{\{3,N-3\}}) \leqslant e^{-K/a_N}.$$

In particular, one can write from the identity above

$$\rho_t^N(x) = \sum_{y=3}^{N-3} \mathbb{E}_x \left[\rho_{t-t\wedge\tau}^N(y) \mathbf{1}_{\{X_{t\wedge\tau}=y, \tau=\delta^2 N^2 a_N\}} \right] + O(e^{-K/a_N}).$$

In particular, one obtains for $t \leqslant \delta^2 N^2 a_N$

$$|\rho_t^N(x+1) - \rho_t^N(x)| = \frac{1}{N} ||\partial_u f_0||_\infty + O(e^{-K/a_N}).$$

and for $t \geqslant \delta^2 N^2 a_N$

$$|\rho_t^N(x+1) - \rho_t^N(x)|$$

$$\leqslant \sum_{y=3}^{N-3} |\mathbb{P}_{x+1}[X_{\delta^2 N^2 a_N} = y] - \mathbb{P}_x[X_{\delta^2 N^2 a_N} = y]| + O(e^{-K/a_N})$$

$$= \sum_{y=3}^{N-3} |\mathbb{P}_y[X_{\delta^2 N^2 a_N} = x+1] - \mathbb{P}_y[X_{\delta^2 N^2 a_N} = x]| + O(e^{-K/a_N})$$

$$= \sum_{y=x+3-N}^{x-3} |\mathbb{P}_0[X_{\delta^2 N^2 a_N} = y+1] - \mathbb{P}_0[X_{\delta^2 N^2 a_N} = y]| + O(e^{-K/a_N}).$$

Since $\mathbb{P}_0\left[X_{\delta^2 N^2 a_N} = y\right]$ is maximal in 0 and decreasing in $|y|$, most of the terms in the estimate above cancel out, and we obtain

$$|\rho_t^N(x+1) - \rho_t^N(x)| \leqslant 2\mathbb{P}_0\left[X_{\delta^2 N^2 a_N} = 0\right] + O(e^{-K/a_N}).$$

By approximating the random walk by a Brownian motion, one shows straightforwardly (cf. Equations (3.2) and (3.5) in De Masi et al., 2012) $\mathbb{P}_0\left[X_{\delta^2 N^2 a_N} = 0\right] \leqslant \left(\sqrt{2\pi a_N}\delta N\right)^{-1}$, which, letting $a_N = K/\log N$, proves the result. $\qquad\square$

4. Estimation on the bulk two-point correlation function

We now estimate the two-point correlation function defined in (2.8).

Proposition 4.1. *For any positive t, and any fixed $\delta > 0$*

$$\limsup_{N\to\infty} \sup_{\substack{3\leqslant x\leqslant(1-\delta)N-1 \\ \delta N+1\leqslant y\leqslant N-3}} |\varphi_t(x,y)| = 0.$$

Once again, the main ingredient to prove this result is an estimate of the correlation function at the boundary of the domain above, given by the following Lemma.

Lemma 4.2. *There exists ε such that for any $\delta > 0$,*

$$\sup_{\substack{\delta N\leqslant y\leqslant N-1 \\ t\geqslant 0}} |\varphi_t^N(3,y)| = O(N^{-\varepsilon}) \quad and \quad \sup_{\substack{1\leqslant x\leqslant N(1-\delta) \\ t\geqslant 0}} |\varphi_t^N(x,N-3)| = O(N^{-\varepsilon}),$$

where the constant in $O(N^{-\varepsilon})$ can depend on δ.

The proof of this Lemma is also postponed Section 5.5. For now, we assume it holds and prove Proposition 4.1.

Proof of Proposition 4.1: Define the set, represented in Figure 4.1

$$B_N^\delta = \left\{(x,y) \in [\![3,(1-\delta)N-1]\!] \times [\![\delta N+1,N-3]\!], \text{ such that } x < y-1\right\}.$$

We also define the boundary sets

$$L_1^\delta = \left\{(3,y) \text{ for } y \in [\![\delta N, N-3]\!]\right\},$$
$$L_{2,N}^\delta = \left\{(x,N-3) \text{ for } x \in [\![3,(1-\delta)N]\!]\right\},$$
$$S_{1,N}^\delta = \left\{((1-\delta)N,y) \text{ for } y \in [\![(1-\delta)N+1,N-4]\!]\right\},$$
$$S_{2,N}^\delta = \left\{(x,\delta N) \text{ for } x \in [\![4,\delta N-1]\!]\right\},$$
$$D_N^\delta = \left\{(x,x+1) \text{ for } x \in [\![\delta N,(1-\delta)N]\!]\right\}.$$

Finally, shorten $L_N^\delta = L_1^\delta \cup L_{2,N}^\delta$, $S_N^\delta = S_{1,N}^\delta \cup S_{2,N}^\delta$, and $\partial B_N^\delta = L_N^\delta \cup S_N^\delta$. Using once again Dynkin's identity $\partial_t \mathbb{E}_{\mu_N}[f(\eta_t)] = \mathbb{E}_{\mu_N}[\mathcal{L}_N f(\eta_t)]$ to $f(\eta) = \eta(x)\eta(y)$ for any $(x,y) \in B_N^\delta$, and since the initial measure for the process is a product measure, elementary computations yield that φ_t^N is solution of the discrete difference diffusion
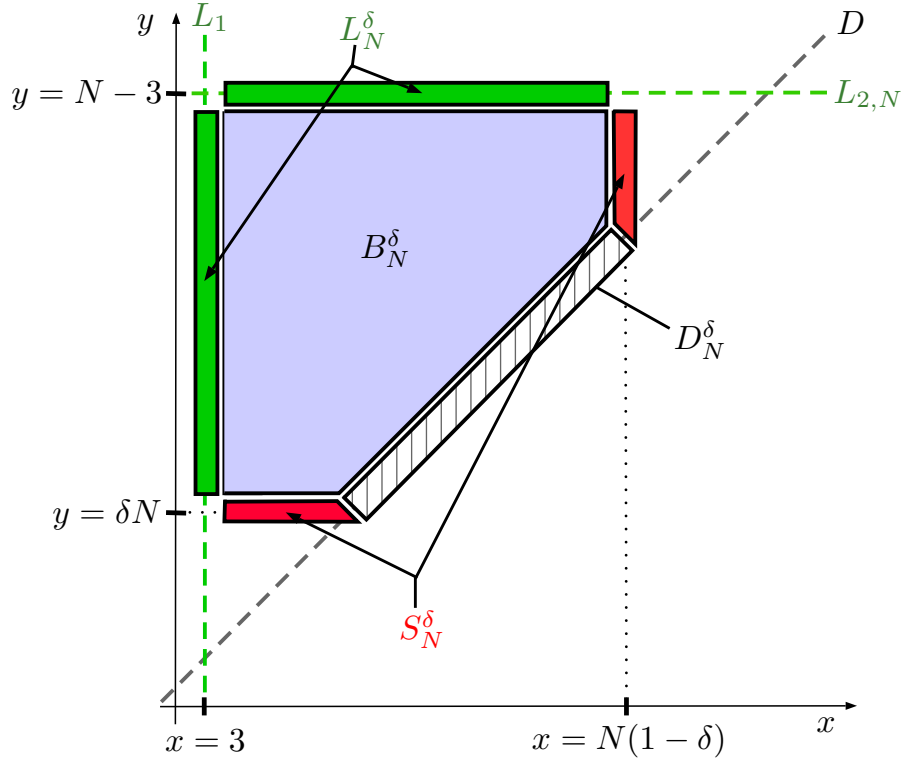
FIGURE 4.1.

equation

$$\begin{cases} \partial_t \varphi_t^N = \Delta_{2,N} \varphi_t^N & \text{for } (x,y) \in B_N^\delta \\ \partial_t \varphi_t^N = \nabla_{2,N} \varphi_t^N + g_{t,N} & \text{for } (x,y) \in D_N^\delta \\ \varphi_t^N = \phi_{t,N} & \text{for } (x,y) \in \partial B_N^\delta \\ \varphi_0^N \equiv 0 \end{cases} \tag{4.1}$$

The operator $\Delta_{2,N}$ stands for the two-dimensional discrete Laplacian

$$\Delta_{2,N}\varphi_t^N(x,y)$$
$$= N^2[\varphi_t^N(x+1,y) + \varphi_t^N(x-1,y) + \varphi_t^N(x,y+1) + \varphi_t^N(x,y-1) - 4\varphi_t^N(x,y)],$$

the operator $\nabla_{2,N}$ represents the reflection at the diagonal D_N^δ

$$\nabla_{2,N}\varphi_t^N(x,x+1) = N^2[\varphi_t^N(x-1,x+1) + \varphi_t^N(x,x+2) - 2\varphi_t^N(x,x+1)],$$

and the function $g_{t,N}$ is defined on D_N^δ as

$$g_{t,N}(x,x+1) = -N^2(\rho_t^N(x+1) - \rho_t^N(x+1))^2.$$

We also defined $\phi_{t,N}(x,y) = \varphi_t^N(x,y)$ on ∂B_N^δ. Note that once again, the boundary identity is trivial, and "artificially" closes the equation satisfied by φ_t^N.

Consider now a random walk X_t on $\mathbb{Z}^2$ with generator $\nabla_{2,N}$ on

$$D := \{(x,x+1), x \in \mathbb{Z}\}$$

and $\Delta_{2,N}$ on $\mathbb{Z}^2 \setminus D$. Defined in this way, X_t is a simple symmetric random walk on $\mathbb{Z}^2$ reflected at the diagonal D. For any set $E \subset \mathbb{Z}^2$, we denote by $H(E)$ the hitting time of the set E, $H(E) = \inf\{t \geqslant 0, \ X_t \in E\}$, and shorten $H_t(E) = t \wedge H(E)$. Finally, we denote by $\mathbb{P}_{x,y}$ the distribution of the random walk started from (x,y) and $\mathbb{E}_{x,y}$ the corresponding expectation. Then, as a consequence of (4.1) and Feynman-Kac's formula, one obtains for any $t \geqslant 0$ and $(x,y) \in B_N^\delta \cup D_N^\delta$

$$\varphi_t(x,y) = \mathbb{E}_{x,y}\left[\varphi_{t-H_t(\partial B_N^\delta)}(X_{H_t(\partial B_N^\delta)}) + \int_0^{H_t(\partial B_N^\delta)} g_{t-s,N}(X_s)\mathbf{1}_{\{X_s \in D_N^\delta\}}ds\right].$$
(4.2)

Before $H_t(\partial B_N^\delta)$, the random walk X_t cannot reach $D \setminus D_N^\delta$, and $H_t(\partial B_N^\delta) \leqslant t$, so that according to Lemma 3.3 the absolute value of the second term above is less or equal than

$$\mathbb{E}_{x,y}\left[\int_0^{H_t(\partial B_N^\delta)} g_{t-s}^N(X_s)\mathbf{1}_{\{X_s \in D_N^\delta\}}ds\right] \leqslant \left(\frac{C\log N}{\delta}\right)^2 \mathbb{E}_{x,y}\left[\int_0^{H_t(\partial B_N^\delta)} \mathbf{1}_{\{X_s \in D_N^\delta\}}ds\right]$$

for some constant C. The right hand side can be estimated using the same steps in Section 4.4 of Erignoux (2018), after which one obtains straightforwardly that

$$\mathbb{E}_{x,y}\left[\int_0^{H_t(\partial B_N^\delta)} \mathbf{1}_{\{X_s \in D_N^\delta\}}ds\right] = O(N^{-1}\log N).$$

We can now write, using (4.2), that

$$\varphi_t(x,y) = \mathbb{E}_{x,y}\left[\varphi_{t-H_t(\partial B_N^\delta)}(X_{H_t(\partial B_N^\delta)})\right] + O(N^{-1}(\log N)^3\delta^{-2}).$$
(4.3)

The function φ is uniformly bounded by 1, vanishes at time 0, and according to Lemma 4.2

$$\sup_{\substack{(x,y) \in L_N^\delta \\ t \geqslant 0}} \varphi_t(x,y) = O(N^{-\varepsilon}),$$

therefore (4.3) yields for any $t \geqslant 0$

$$|\varphi_t(x,y)| \leqslant \mathbb{P}_{x,y}\left(H_t(\partial B_N^\delta) = H(S_N^\delta)\right) + O(N^{-\varepsilon}) + O(N^{-1}(\log N)^3\delta^{-2})$$
$$\leqslant \mathbb{P}_{x,y}\left(H(L_N^\delta) > H(S_N^\delta)\right) + O(\delta^{-2}N^{-\varepsilon}).$$
(4.4)

Define $L_N = L_1 \cup L_{2,N}$, where $L_1 = \{(3,y), \text{ for } y \in \mathbb{Z}\}$ and $L_{2,N} = \{(x, N - 3) \text{ for } x \in \mathbb{Z}\}$ are represented in Figure 4.1. Then one easily checks that

$$\mathbb{P}_{x,y}\left(H(L_N^\delta) > H(S_N^\delta)\right) \leqslant \sum_{i=1}^2 \mathbb{P}_{x,y}\left(H(L_N) > H(S_{i,N}^\delta)\right).$$
(4.5)

Both terms in the right hand side above are treated in the same way. Assume, for example, that $i = 1$. Then,

$$\mathbb{P}_{x,y}\left(H(L_N) > H(S_{1,N}^\delta)\right) \leqslant \widetilde{\mathbb{P}}_{x,y}\left(H(L_N) > H(S_{1,N}^\delta)\right),$$

where $\widetilde{\mathbb{P}}_{x,y}$ is the distribution of a random walk $\widetilde{X}$ is *reflected* both at D and at $L_{2,N}$. Note that a reflected random walk can be mapped into a non-reflecting one by, each time the random walks hits the reflecting line, choosing with probability $1/2$ a side of the line to carry on with the random walk. Consider $\delta' > \delta$ and assume now that the starting point is in $(x,y) \in B_N^{\delta'} \cup D_N^{\delta'}$. Since the random walk is reflected both at the horizontal boundary $L_{2,N}$, as illustrated in Figure 4.2, thanks to the
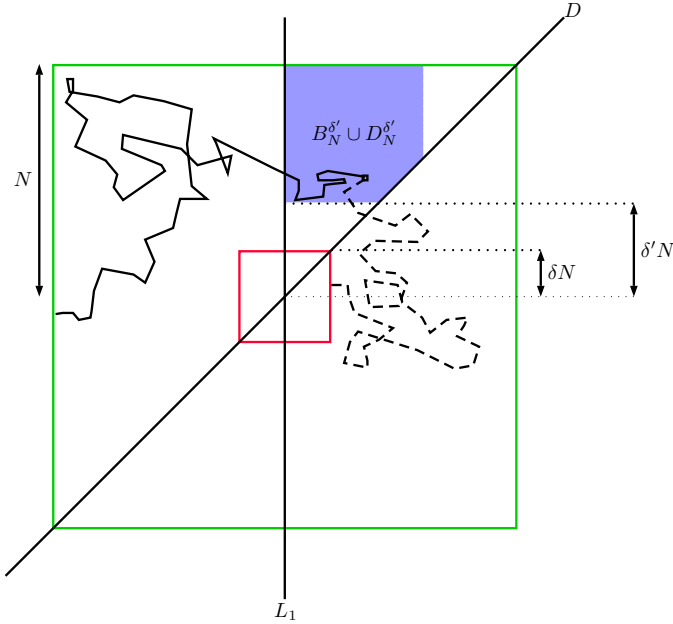
FIGURE 4.2. A random walk started in $B_N^{\delta'}$ (blue area) has a probability of order $\log(\delta')/\log(\delta)$ of hitting $\Lambda_{\delta N}$ (red boundary) before Λ_N^c (green boundary) (dashed trajectory). In particular, as $\delta \to 0$, it will most likely hit the green boundary before the red one (solid trajectory).

previous observation, the probability $\widetilde{\mathbb{P}}_{x,y}\left(H(L_1) > H(S_{1,N}^\delta)\right)$ is less or equal than the probability that a random walk starting at a distance $\delta'N$ of the origin hits $\Lambda_{\delta N} = \{(x,y), x, y \leqslant \delta_N\}$ before hitting Λ_N^c, so that for any $(x,y) \in B_N^{\delta'} \cup D_N^{\delta'}$

$$\widetilde{\mathbb{P}}_{x,y}\left(H(L_1) > H(S_{1,N}^\delta)\right) \leqslant C\frac{\log(\delta'N) - \log N}{\log(\delta N) - \log N} = C\frac{\log \delta'}{\log \delta}.$$

An analogous bound holds for the second term in (4.5), i.e. for any $(x,y) \in B_N^{\delta'} \cup D_N^{\delta'}$,

$$\mathbb{P}_{x,y}\left(H(L_N^\delta) > H(S_{2,N}^\delta)\right) \leqslant C\frac{\log \delta'}{\log \delta}.$$

Plugging this bound into (4.4) yields that for any $(x,y) \in B_N^{\delta'}$, $t \geqslant 0$,

$$|\varphi_t(x,y)| \leqslant C(\delta)O(N^{-\varepsilon}) - \frac{C(\delta')}{\log(\delta)}.$$

so that letting $N \to \infty$ and then $\delta \to 0^+$ proves the result. $\square$

4.1. *Proof of Theorem 2.1.* Propositions 3.1 and 4.1 are enough to prove the hydrodynamic limit. Indeed, fix a smooth test function G and a time $t > 0$. We

write

$$\mathbb{E}_{\mu^N}\left(\left[\frac{1}{N}\sum_{x=1}^{N-1}G(x/N)\eta_t(x)-\int_{[0,1]}G(u)\rho_t(u)du\right]^2\right)$$

$$\leqslant 2\mathbb{E}_{\mu^N}\left(\left[\frac{1}{N}\sum_{x=1}^{N-1}G(x/N)(\eta_t(x)-\rho_t^N(x)\right]^2\right)$$

$$+2\left[\frac{1}{N}\sum_{x=1}^{N-1}G(x/N)\rho_t^N(x)-\int_{[0,1]}G(u)\rho_t(u)du\right]^2. \quad (4.6)$$

Since G is bounded, the first term on the right-hand side is, for any fixed $\delta' > 0$, according to Proposition 4.1 bounded by $2\delta' + O(1/N) + o_N(1)$, so that letting $N \to \infty$ and then $\delta' \to 0$, it vanishes. The second term vanishes as well according to Proposition 3.1. Using Markov's inequality concludes the proof.

5. Branching process and estimation of the density at the boundary

5.1. *Graphical construction and dual branching process.* We now turn to Lemmas 3.2 and 4.2, whose proofs were postponed and are presented here. We follow analogous arguments and constructions to Sections 5.1 and 5.2 in Erignoux et al. (2018), although significant adaptations need to be made to take into account the time dependency and the slowed down boundary dynamics.

We start with a graphical construction of the dynamics. We consider $N + 6$ independent Poisson point processes on $\mathbb{R}$ defined as follows :

- $N - 2$ processes $\mathscr{N}_{x,x+1}$, $x \in [\![1, N-2]\!]$ with intensity N^2, representing the exclusion dynamics.
- Two processes $\mathscr{N}_1^+$, $\mathscr{N}_{N-1}^+$ with respective intensities $N^{2-\theta}\bar{\rho}r$ and $N^{2-\theta}\bar{\rho}'r'$, representing putting a particle at a boundary site 1, $N-1$ regardless of its previous value.
- Two processes $\mathscr{N}_1^-$, $\mathscr{N}_{N-1}^-$ with respective intensities $N^{2-\theta}(1-\bar{\rho})r$ and $N^{2-\theta}(1-\bar{\rho}')r'$ representing emptying a boundary site 1, $N-1$ regardless of its previous value.
- Two processes $\mathscr{N}^c$, $\mathscr{N}^{c'}$ with respective intensities $N^{2-\theta}c$ and $N^{2-\theta}c'$ representing the copy mechanism.
- Two processes $\mathscr{N}^b$, $\mathscr{N}^{b'}$ with respective intensities $N^{2-\theta}b$ and $N^{2-\theta}b'$ representing the boundary sites $2, N-2$ being filled iff the other boundary site 1, $N-1$ is occupied.

In order to build the process starting from a given configuration η_0, we order $0 < t_1 < t_2, \cdots$ the points

$$\{t_i, i \in \mathbb{N}\backslash\{0\}\} = \left(\mathscr{N}_1^\pm \cup \mathscr{N}^c \cup \mathscr{N}^b \cup \mathscr{N}_{N-1}^\pm \cup \mathscr{N}^{c'} \cup \mathscr{N}^{b'} \cup \left(\cup_{x=1}^{N-2}\mathscr{N}_{x,x+1}\right)\right)\cap\mathbb{R}_+$$

of these processes, taken in increasing order starting from 0. Note that since it has probability 0, we discounted the possibility that two of these processes contained the same point. We refer to the points of the processes together with their type as *marks*. For example, if $t \in \mathscr{N}_{N-1}^\pm$, (resp. $t \in \mathscr{N}^b \cup \mathscr{N}^{b'}$) we say that the *mark* at t is of type $(\pm, N-1)$ (resp. of *branching* type). If $t \in \mathscr{N}_{x,x+1}$, we say that the *mark*

at t is of type $(x, x+1)$. For convenience, we store in a variable $\mathcal{M}$ the collection of all the marks up to some arbitrary time horizon T.

With these Poisson point processes, given an initial configuration η_0, we are able to give a graphical construction of our Markov process $(\eta_t)_{t \geqslant 0}$. Define $t_0 = 0$ and $\eta_{t_0} = \eta_0$, and "solve" the marks $t_1, t_2, \cdots$ one by one in the following way. For any $i > 0$, the configuration remains constant on $[t_{i-1}, t_i)$. At time t_i, the configuration is changed depending on the type of the mark t_i.

- If $t_i \in \mathcal{N}_{x,x+1}$, we let $\eta_{t_i} = \eta_{t_{i-1}}^{x,x+1}$.
- If $t_i \in \mathcal{N}_1^+$ (resp. $t_i \in \mathcal{N}_{N-1}^+$), at time t_i, a particle is put at site 1 (resp. $N-1$). We let $\eta_{t_i}(x) = \eta_{t_{i-1}}(x)$ for any $x \neq 1$ (resp. $x \neq N-1$), and $\eta_{t_i}(1) = 1$ (resp. $\eta_{t_i}(N-1) = 1$).
- If $t_i \in \mathcal{N}_1^-$ (resp. $t_i \in \mathcal{N}_{N-1}^-$), at time t_i, site 1 (resp. $N-1$) is emptied. We let $\eta_{t_i}(x) = \eta_{t_{i-1}}(x)$ for any $x \neq 1$ (resp. $x \neq N-1$), and $\eta_{t_i}(1) = 0$ (resp. $\eta_{t_i}(N-1) = 0$).
- If $t_i \in \mathcal{N}^c$ (resp. $t_i \in \mathcal{N}^{c'}$), at time t_i, site 2 (resp. $N-2$) becomes a copy of site 1 (resp. $N-1$), so that we let $\eta_{t_i}(x) = \eta_{t_{i-1}}(x)$ for any $x \neq 2$ (resp. $x \neq N-2$), and $\eta_{t_i}(2) = \eta_{t_{i-1}}(1)$ (resp. $\eta_{t_i}(N-2) = \eta_{t_{i-1}}(N-1)$).
- Finally, if $t_i \in \mathcal{N}_2^b$ (resp. $t_i \in \mathcal{N}^{b'}$), at time t_i, we let $\eta_{t_i}(x) = \eta_{t_{i-1}}(x)$ for any $x \neq 2$ (resp. $x \neq N-2$), and if $\eta_{t_{i-1}}(1) = 1$ (resp. $\eta_{t_{i-1}}(N-1) = 1$), we let $\eta_{t_i}(2) = 1$ (resp. $\eta_{t_i}(N-2) = 1$). Otherwise, we let $\eta_{t_i}(2) = \eta_{t_{i-1}}(2)$ (resp. $\eta_{t_i}(N-2) = \eta_{t_{i-1}}(N-2)$).

We leave it to the reader to check that this construction yields a Markov process η_t, with initial state η_0, and infinitesimal generator $N^2 \mathcal{L}_N$.

5.2. *Set of unknown sites.* Fix $x \in \Lambda_N$, and suppose that we want to determine the value of $\eta_t(x)$. To do so, we are going to explore the past of the process, and define a set $\hat{A}(s) \subset \Lambda_N$, $s < t$ of unknown sites at time s on which depends the value of $\eta_t(x)$. To do so, re-index the marks $t > t_{-1} > t_{-2} > \cdots > t_{-\kappa} > 0$ in the time interval $[0, t]$, we leave the set $\hat{A}(s)$ constant on each of the intervals $[t_{-1}, t]$, $[t_{-2}, t_{-1}), \ldots$ until time 0 is reached. For $k < 0$, assume that $\hat{A}(s)$ has been defined down to time t_k, i.e. on the whole segment $[t_k, t]$. To define $\hat{A}(s)$ down to time t_{k-1}, we define it at time t_k^- and leave it constant down to time t_{k-1}. Shorten $A := \hat{A}(t_k)$ and $A' := \hat{A}(t_k^-)$. We consider all the possible cases for the type of the mark t_k.

- If $t_k \in \mathcal{N}_{x,x+1}$, $x \in A$ and $x+1 \notin A$ (resp. $x+1 \in A$ and $x \notin A$), before the mark, the content of the unknown site x (resp. $x+1$), particle or hole, was in $x+1$ (resp. x), and we let $A' = A^{x,x+1} := A \cup \{x+1\} \setminus \{x\}$ (resp. $:= A \cup \{x\} \setminus \{x+1\}$). Otherwise, if both or neither of the two sites are in A, we just let $A' = A$.
- If $t_k \in \mathcal{N}_1^+$ (resp. $t_k \in \mathcal{N}_{N-1}^+$), then at the mark a particle was placed at site 1 (resp. $N-1$). Therefore, if site 1 (resp. $N-1$) was in the set A of unknowns, it can be removed, and we let $A' = A \setminus \{1\}$ (resp. $A' = A \setminus \{N-1\}$). Else, we let $A' = A$.
- If $t_k \in \mathcal{N}_1^-$ (resp. $t_k \in \mathcal{N}_{N-1}^-$), then at the mark, site 1 (resp. $N-1$) was emptied. Therefore, if site 1 (resp. $N-1$) was in the set A of unknowns, it is removed, and we let $A' = A \setminus \{1\}$ (resp. $A' = A \setminus \{N-1\}$). Else, we let $A' = A$.

- If $t_k \in \mathcal{N}^c$ (resp. $t_k \in \mathcal{N}^{c'}$), then at the mark, site 2 (resp. $N-2$) became a copy of site 1 (resp. $N-1$) so that if $2 \in A$ (resp. $N-2 \in A$), we let $A' = A \cup \{1\} \setminus \{2\}$ (resp. $A' = A \cup \{N-1\} \setminus \{N-2\}$). Else, we let $A' = A$.
- Finally, if $t_k \in \mathcal{N}^b$ (resp. $t_k \in \mathcal{N}^{b'}$), then at the mark t_k, something might have happened at site 2 (resp. $N-2$) if and only if site 1 (resp. $N-1$) was occupied. Else, nothing happened, so that we need to keep track of both of the sites 1 and 2 to know for sure the value at time t. More precisely, if $2 \in A$ (resp. $N-2 \in A$), we then let $A' = A \cup \{1\}$ (resp. $A' = A \cup \{N-1\}$). Else, we let $A' = A$.

We then iterate the process to build $\hat{\mathcal{A}}(s)$ backwards in time. Recall that elements of $\hat{\mathcal{A}}$ are *sites* whose value is unknown. To avoid ambiguity, we will refer to the elements of $\hat{\mathcal{A}}$ as *flags*. Once time 0 is reached, all the remaining flags in $\hat{\mathcal{A}}$ are determined by the initial configuration η_0, so that we let $\hat{\mathcal{A}}(0) = \emptyset$. We say that a mark t_k *affected* $\hat{\mathcal{A}}$ if at least one of the sites concerned by the mark was in $\hat{\mathcal{A}}(t_k)$. For example, a mark at time t_k of type $(x, x+1)$ affected $\hat{\mathcal{A}}$ iff. $\hat{\mathcal{A}}(t_k) \cap \{x, x+1\} \neq \emptyset$.

When a mark t_k is of type b (resp. b') and $2 \in \hat{\mathcal{A}}(t_k)$ (resp. $N-2 \in \hat{\mathcal{A}}(t_k)$), the process *branched* and a flag is *created* in $\hat{\mathcal{A}}$ (we recall that the process is constructed backwards in time). Note that this is the only way for the cardinal of $\hat{\mathcal{A}}$ to *increase*. When a mark t_k is of type $(\pm, 1)$ (resp. $(\pm, N-1)$) and $1 \in \hat{\mathcal{A}}(t_k)$ (resp. $N-1 \in \hat{\mathcal{A}}(t_k)$), a flag is *deleted* in $\hat{\mathcal{A}}$. Another way for the cardinal of $\hat{\mathcal{A}}$ to decrease is if the process meets a mark of type c or c' either at site 2 or $N-2$, while site 1 or $N-1$ was also in $\hat{\mathcal{A}}$, in which case site 2 (resp. $N-2$) is removed from $\hat{\mathcal{A}}$, while site 1 (resp. $N-1$) remains. The last way for the cardinal of $\hat{\mathcal{A}}$ to decrease is hitting time 0 in which case all remaining flags are removed.

We will call *boundary marks* the marks that make the cardinal of $\mathcal{A}$ change. Note, in particular, that not all marks of type c or c' are boundary marks, only those who occur while both sites of the boundary are flagged. In section 5.4, we explain how to recover from the process's value $\eta_x(t)$ from the process $\hat{\mathcal{A}}$ and the marks who affected it. For now, however, we study the process $\hat{\mathcal{A}}$ itself.

5.3. *Markov flag processes.* Recall that $\hat{\mathcal{A}}(s)$ is defined on $[0, t]$, observed backwards in time, and is right-continuous. For simplicity, we want to observe the process $\hat{\mathcal{A}}(s)$ forward in time, so that we define $(\mathcal{A}(s))_{s \in [0,t]}$ the right-continuous version of $\hat{\mathcal{A}}(-s)$: we set $\mathcal{A}(s) = \hat{\mathcal{A}}(-s)$ on $[0, t]$ except at the time of the marks t_k, at which $\mathcal{A}(t_k) = \mathcal{A}(t_k^+)$. Since the Poisson point processes forward and backward in time have the same distribution, one easily checks that up until time t (forward in time) $\mathcal{A}(s)$ is a Markov process on $\mathcal{P}(\Lambda_N)$ (the set of subsets of Λ_N) with generator $\mathcal{L}_N^\dagger = \mathcal{L}_N^{\dagger,1} + \mathcal{L}_N^{\dagger,2}$ acting on functions $f : \mathcal{P}(\Lambda_N) \to \mathbb{R}$. The generator $\mathcal{L}_N^{\dagger,1}$ represents the flag's motion, either due to the stirring dynamics or the copy mechanism,

$$\mathcal{L}_N^{\dagger,1} f(A) = N^2 \sum_{x=1}^{N-2} \left(f(A^{x,x+1}) - f(A) \right) + cN^{2-\theta} \mathbf{1}_{\{2 \in A\}} \left(f(A \cup \{2\} \setminus \{1\}) - f(A) \right)$$

$$+ c' N^{2-\theta} \mathbf{1}_{\{N-2 \in A\}} \left(f(A \cup \{N-2\} \setminus \{N-1\}) - f(A) \right).$$

whereas $\mathcal{L}_N^{\dagger,2}$ represents the branching (creation) and death (deletion) events at the boundary,

$$
\begin{aligned}
\mathcal{L}_N^{\dagger,2} f(A) =\,& r N^{2-\theta} \mathbf{1}_{\{1 \in A\}} \left(f(A \setminus \{1\}) - f(A) \right) \\
&+ r' N^{2-\theta} \mathbf{1}_{\{N-1 \in A\}} \left(f(A \setminus \{N-1\}) - f(A) \right) \\
&+ b N^{2-\theta} \mathbf{1}_{\{2 \in A\}} \left(f(A \cup \{1\}) - f(A) \right) \\
&+ b' N^{2-\theta} \mathbf{1}_{\{N-2 \in A\}} \left(f(A \cup \{N-1\}) - f(A) \right).
\end{aligned}
\tag{5.1}
$$

In the expression of $\mathcal{L}_N^{\dagger,1}$, we denoted

$$
A^{x,x+1} = \begin{cases}
A & \text{if } x, x+1 \notin A \text{ or } x, x+1 \in A \\
A \setminus \{x\} \cup \{x+1\} & \text{if } x \in A, \ x+1 \notin A \\
A \setminus \{x+1\} \cup \{x\} & \text{if } x+1 \in A, \ x \notin A.
\end{cases}
$$

the set A after exchange of sites x and $x+1$. We also define

$$
\widetilde{\mathcal{L}}_N^{\dagger,1} f(A) = N^2 \sum_{x=1}^{N-2} \left(f(A^{x,x+1}) - f(A) \right) + c N^{2-\theta} \mathbf{1}_{\{2 \in A\}} \left(f(A^{1,2}) - f(A) \right)
$$

$$
+ c' N^{2-\theta} \mathbf{1}_{\{N-2 \in A\}} \left(f(A^{N-1,N-2}) - f(A) \right),
$$

which is almost identical to $\mathcal{L}_N^{\dagger,1}$, except that when A contains either $\{1,2\}$ or $\{N-2, N-1\}$, and a mark of type c or c' occurs. In this case, instead of removing the flag at site 2 or $N-2$, $\widetilde{\mathcal{L}}_N^{\dagger,1}$ simply switches the two flags at the boundary.

Remark 5.1. We remark that the generator $\widetilde{\mathcal{L}}_N^{\dagger,1}$ above does *not* allow branchings, which are the only way for the flags to increase. Moreover, since in the particular case of the copies mentioned above we do not remove flags but exchange them, this generator *conserves* the number of flags.

In what follows, we forget the time horizon t, and consider two Markov processes $\mathcal{A}(s)$ (resp. $\mathcal{B}(s)$) started from a set $A \subset \Lambda_N$, and driven by the generator $\mathcal{L}_N^{\dagger}$ (resp. $\widetilde{\mathcal{L}}_N^{\dagger,1}$). Recall the construction using the Poisson point processes introduced above, since they are equivalent, we will alternatively use both descriptions (as Markov processes, or as functions of the Poisson point processes) for the two processes $\mathcal{A}$ and $\mathcal{B}$. Recall that we refer to any element of the processes $\mathcal{N}^b$ or $\mathcal{N}^{\pm}$ as *boundary marks* (For simplicity, we will often ignore the boundary marks of type c or c', we will prove later on that they occur with very small probability). The process $\mathcal{B}$ therefore evolves as $\mathcal{A}$, except that the marks that would branch or delete flags have no effect. Note in particular that until the first time $\tau > 0$ that $\mathcal{A}$ is affected by a boundary mark, we have $\mathcal{A}(s) = \mathcal{B}(s)$ for any $s < \tau$. We denote $\mathbb{P}_A^{\dagger}$ the joint distribution of those two processes started from the set A and $\mathbb{E}_A^{\dagger}$ the corresponding expectation.

Fix a starting set $A \subset \Lambda_N$, and give each element of A an arbitrary and unique label n, $n = 1, \ldots, |A|$. For $1 \leqslant n \leqslant |A|$, we follow the evolution of the flags in $\mathcal{A}$ and $\mathcal{B}$, and denote by $X_n(t)$ (resp. $Y_n(t)$) the position of the flag labeled n in $\mathcal{A}(t)$ (resp. $\mathcal{B}(t)$). Each time a new flag appears in $\mathcal{A}$, we label it by the smallest integer not used previously. Whenever a mark $(x, x+1)$ occurs while both x and $x+1$ are in $\mathcal{A}$ or $\mathcal{B}$, we switch the labels of the flags at x and $x+1$. In other words,

$$
\mathcal{A}(t) = \{ X_n(t), \ n \in K(t) \}, \quad \text{and} \quad \mathcal{B}(t) = \{ Y_n(t), \ n \leqslant |A| \}
$$

where $K(t)$ is the set of flag labels currently present at time t in $\mathcal{A}$. Further note that, fixed a label n, Y_n is a random walk on Λ_N reflected at the boundary, jumping at rate N^2 to each neighbor, and at an extra rate $cN^{2-\theta}$ (resp. $c'N^{2-\theta}$) from 2 to 1 (resp. $N-2$ to $N-1$). The processes X_n have the same distribution, except that they branch when they meet a branching mark (i.e. at rate $bN^{2-\theta}$, when $X_n = 2$, and at rate $b'N^{2-\theta}$ when $X_n = N - 2$), and die when they meet a mark $\pm$ (i.e. at rate $rN^{2-\theta}$ when $X_n = 1$ and at rate $r'N^{2-\theta}$ when $X_n = N - 1$). They can also die when a mark of type c, c' occurs while another flag is in the same boundary, but since we will prove the probability that this happens during $\mathcal{A}$'s lifespan is very small, we do not detail this possibility.

We first focus on the process $\mathcal{B}$. Denote by τ_n the first time at which the flag labeled n in $\mathcal{B}$ meets a boundary mark. Denote $\tau = \tau_1 \wedge \tau_2 \wedge \cdots \wedge \tau_m$ the first time a flag meets a boundary mark. For $n \leqslant |A|$, we define

$$C = \{\text{At time } \tau, \text{ there was only one flag at the left boundary}\}$$
$$= \{|\mathcal{B}(\tau) \cap \{1,2\}| = 1\}$$

and

$$D_\pm = \{\text{The mark encountered at time } \tau \text{ was of type } \pm\}.$$

We start with a technical Lemma, which is the main ingredient to prove Lemmas 3.2 and 4.2.

Lemma 5.2. *Define $\nu_\theta = \min(\theta, 1 - \theta)/4$. For any $A \subset \{1, \ldots, N/2\}$ containing site 3, and any $1 \leqslant n \leqslant |A|$*

$$\mathbb{P}_A^\dagger(C \cap D_\pm \cap \{\tau = \tau_n\}) = \frac{r\rho_\pm}{r+b}\mathbb{P}_A^\dagger(\tau = \tau_n) + |A|^2 O(N^{-\nu_\theta}), \tag{5.2}$$

where $\rho_+ = \overline{\rho}$ and $\rho_- = 1 - \overline{\rho}$. Furthermore, for any $A \subset \Lambda_N$

$$\mathbb{P}_A^\dagger(D_+ \cup D_-) \geqslant p + |A|^3 O(N^{-\nu_\theta}), \tag{5.3}$$

where

$$p := \min\left\{\frac{r}{r+b}, \frac{r'}{r'+b'}\right\} > \frac{1}{2}.$$

In the identities above, the $O(N^{-\nu_\theta})$ can be chosen uniform in the initial set A satisfying the relevant conditions.

Although this result involves significant technical difficulties, its content is fairly simple, so that before proving it, we briefly explain those two identities. Recall that we ultimately want to estimate the boundary densities (Lemma 3.2) and two-points correlations (Lemma 4.2), so that we need to examine the value of $\eta_t(3)$ and $\eta_t(3)\eta_t(x)$ for x of order δN. For the first identity, we concentrate on sets contained in the left half of the system and containing site 3 to avoid waiting for a long time for a flag to get close to a boundary, and to know with high probability that the next boundary mark is gonna be encountered at the left boundary. Equation (5.2) states that for any n, assuming that the next flag to encounter a boundary mark is labelled n, the probability that this boundary mark is of type $+$ (resp. $-$) converges to $\overline{\rho}r/(r + b)$ (resp. $(1 - \overline{\rho})r/(r + b)$). The second identity states that at the next boundary mark encountered, the probability that the affected flag branches rather than dies is less than $1/2$, which ultimately ensures that w.h.p. the branching process ultimately dies out after branching a finite number of times (cf. Corollary 5.3).

Proof of Lemma 5.2: This proof being a little intricate, we split it in several steps.

Step 1: a crude estimate on τ.   Recall (2.6). We first claim

$$\mathbb{P}^\dagger_A(\tau \geqslant N^{-\widehat{\theta}}) = O(N^{-\widehat{\theta}}) \tag{5.4}$$

for any set A containing site 3: Fix such a set, and without loss of generality label 1 the flag at site 3. The proof of this first claim follows the same steps as Lemma 5.3 of Erignoux et al. (2018), which we adapt to the slowed-down rates at the boundary. To prove (5.4), first note that

$$\mathbb{E}^\dagger_A(\tau) \leqslant \mathbb{E}^\dagger_A(\tau_1) = \mathbb{E}^\dagger_{\{3\}}(\tau).$$

Consider therefore a single flag initially at site 3, $Y_1(t)$ its position in $\mathcal{B}(t)$. This flag performs excursions away from site 3, either in the bulk $\{4, \ldots, N-1\}$, or in the left boundary $\{1, 2\}$, in which case it has a probability of order $N^{-\theta}$ of encountering a boundary mark, after which it ultimately gets back to site 3. By the Markov property, all the excursions of the flag are i.i.d. in $\mathcal{B}$. In what follows, we will neglect the possibility that the flag meets a boundary mark at the right boundary, which would only decrease τ and is therefore not an issue. Denote t_i the time of the i-th visit to site 3, i.e. $t_0 = 0$, and

$$t_i = \inf\{t > t_{i-1}, \quad Y_1(t) = 3, \quad Y_1(t^-) \neq 3\}.$$

For $i > 0$, denote by E_i the event

$$E_i = \{\text{The flag met a left boundary mark during the } i\text{-th excursion } [t_{i-1}, t_i)\}.$$

Since the excursions are i.i.d., so are the E_i's and $t_i - t_{i-1}$'s, and it is elementary to show (cf. Lemma 5.3 of Erignoux et al., 2018) that $\mathbb{E}^\dagger_{\{3\}}(t_i - t_{i-1}) = \mathbb{E}^\dagger_{\{3\}}(t_1) = O(N^{-1})$ (the excursions in the boundary last a time of order N^{-2}, whereas those in the bulk last a time of order N^{-1}). We can now write

$$\mathbb{E}^\dagger_{\{3\}}(\tau) = \sum_{i>0} \mathbb{E}^\dagger_{\{3\}}\left(\tau \mathbf{1}_{\{E_i \text{ and } E_j^c \text{ for } j<i\}}\right)$$

$$\leqslant \sum_{i>0} i\mathbb{E}^\dagger_{\{3\}}(t_1)\mathbb{P}^\dagger_{\{3\}}\left(E_i \text{ and } E_j^c \text{ for } j < i\right).$$

There exists a constant $c = c(r, b, c)$ such that each boundary excursion has a probability $\mathbb{P}^\dagger_{\{3\}}(E_1) \geqslant cN^{-\theta}$ of encountering a boundary mark. This finally yields as wanted

$$\mathbb{E}^\dagger_{\{3\}}(\tau) = O(N^{\theta-1}), \tag{5.5}$$

the factor N^θ being the expectation of the geometric number of excursions necessary before meeting a boundary mark. This together with Markov's inequality proves (5.4) uniformly for any set A containing site 3.

Step 2: inserting independent excursions at the left boundary. Because of the subtle correlations between the flag's motions and the marks they encounter, we detail this step, which is one of the main novelties w.r.t. Erignoux et al. (2018). We will call *excursion at the left boundary* a pair $(\tau_Z, Z(s), 0 \leqslant s \leqslant \tau_Z)$, satisfying

(1) $\tau_Z > 0$, $Z_0 = 2$, for any $s < \tau_Z$, $Z(s) \in \{1, 2\}$ and $Z_{\tau_Z} = 3$.
(2) $(Z(s))_{0 \leqslant s \leqslant \tau_Z}$ is right-continuous.

In order not to burden the notations, when referring to an excursion, we will simply denote it Z, the stopping time τ_Z will be implicitly associated with it.

Fix a realization of the marks $\boldsymbol{\mathcal{M}}$ on Λ_N, on which we build $\mathcal{A}$ and $\mathcal{B}$. On the same probability space, fix an i.i.d. family of boundary excursions $(Z^j)_{j\in\mathbb{Z}}$, each distributed as

$$\mathbb{P}(Z^j \in \cdot) = \mathbb{P}^\dagger_{\{2\}} \left((Y_1(s))_{0\leqslant s\leqslant H_3} \in \cdot \right),$$

where $H_3 = \inf\{s \geqslant 0, Y_1(s) = 3\}$. To each excursion Z^j, we also associate boundary marks distributed according to independent Poisson point processes $\mathcal{M}^j$ on $[0, \tau_{Z^j}]$ with the same intensity as the ones appearing in $\mathbb{P}^\dagger$. We take the family of excursions and their associated boundary marks independent of $\boldsymbol{\mathcal{M}}$. We are going to build a second process $\mathcal{B}'$, distributed as $\mathcal{B}$, using both the marks in $\boldsymbol{\mathcal{M}}$ and these independent excursions Z^j. Fix a set $A \subset \{1, \ldots, N/2\}$ containing site 3, and set $\mathcal{B}'(0) = A$. The two processes $\mathcal{B}, \mathcal{B}'$ then follow the marks in $\boldsymbol{\mathcal{M}}$ (recall that the boundary marks do not affect the evolution of either $\mathcal{B}$ or $\mathcal{B}'$), until the first time, t_1, that a flag is at site 2 while no other flag is at a distance less than $\ell_N := (\log N)^4$ from the left boundary,

$$t_1 = \inf\{s \geqslant 0, \ \mathcal{B}'(s) \cap \{1, \ldots, \ell_N\} = \{2\}\}.$$

After time t_1, we let all the flags in $\mathcal{B}'$ evolve according to the marks, except the flag that was at site 2 at time t_1, which then performs the excursion Z^1 until time

$$s_1 = \inf\{s \geqslant t_1, \ 3 \in \mathcal{B}'(s)\}.$$

Note that, most likely, we have $s_1 = t_1 + \tau_{Z^1}$, because otherwise another flag has reached site 3 before the excursion Z^1 finished. Since by construction all other flags were, when the excursion started, at a distance ℓ_N of the left boundary, this is very unlikely. If $s_1 \neq t_1 + \tau_{Z^1}$, we say that the construction *failed* after inserting the excursion Z^1. Regardless of whether or not the construction failed, we let $\mathcal{B}'$ evolve after time s_1 by following the marks in $\boldsymbol{\mathcal{M}}$ until the second time

$$t_2 = \inf\{s \geqslant s_1, \ \mathcal{B}'(s) \cap \{1, \ldots, \ell_N\} = \{2\}\}.$$

We then replace the trajectory of the flag at site 2 by the excursion Z^2, until time

$$s_2 = \inf\{s \geqslant t_2, \ 3 \in \mathcal{B}'(s)\}.$$

Once again, if $s_2 \neq t_2 + \tau_{Z^2}$ we say that the construction failed after inserting excursion Z^2, and then repeat the construction. More precisely, assume that $\mathcal{B}'$ has been built up until time s_{j-1}, it then follows the marks in $\boldsymbol{\mathcal{M}}$ until

$$t_j = \inf\{s \geqslant s_{j-1}, \ \mathcal{B}'(s) \cap \{1, \ldots, \ell_N\} = \{2\}\}.$$

After t_j, $\mathcal{B}'(s)$ follows the marks in $\boldsymbol{\mathcal{M}}$, except the flag present at site 2 at time t_j, which follows the trajectory Z^j until time $s_j = \inf\{s \geqslant t_j, \ 3 \in \mathcal{B}'(s)\}$. If $s_j \neq t_j + \tau_{Z^j}$, we say that the construction fails after inserting Z^j, and we then carry on with the same scheme. Recall that we built both $\mathcal{B}$ and $\mathcal{B}'$ on the same probability space, we denote $\mathbb{Q}_A$ their joint distribution starting from the set A. Further note that by Markov property, $\mathcal{B} \overset{(d)}{=} \mathcal{B}'$. We denote by τ' the first time a flag in $\mathcal{B}'$ meets a boundary mark.

Step 3: estimation of the probability that the construction failed before time τ'. Recall that the construction fails if, after inserting an excursion Z_j, a flag initially at a site $x > \ell_N$ gets to site 3 before the excursion ends. This means

either that the boundary excursions lasted more than $N^{-2}\ell_N$, or that another flag traveled a distance of order ℓ_N in a time $N^{-2}\ell_N$. Both of those probabilities are at most of order $|A|O(e^{-\sqrt{\ell_N}}) = O(N^{1-\log N})$, i.e. for any integer j,

$$\mathbb{Q}_A(\text{The construction failed after inserting } Z^j) = O(N^{-\log N/2}) \qquad (5.6)$$

We now only need a rough bound on the number of excursions inserted before time τ'. Define a discrete time random walk M_j as

$$M_{j+1} = M_j + N^{-3}\mathbf{1}_{\{s_j - t_j \geqslant N^{-3}\}},$$

which jumps at a distance N^{-3} to the right whenever an excursion is inserted that lasted longer than N^{-3}. Clearly, $\mathbb{P}(M_{j+1} = M_j) = O(N^{-1})$ for all steps j and independently from the other steps, because the jumps in $\mathcal{B}$ occur at rate $O(N^{-2})$. In particular, by a standard large deviation estimate, $\mathbb{P}(M_{N^4} \leqslant 1) = O(e^{-N})$. Furthermore, by construction, $M_j \geqslant 1 \Rightarrow s_j \geqslant 1$. These two remarks, together with (5.4), yield that

$$\mathbb{Q}_A(s_{N^4} \leqslant \tau') \leqslant \mathbb{Q}_A(\tau' \geqslant 1) + \mathbb{Q}_A(s_{N^4} \leqslant 1) = O(N^{-\widehat{\theta}}) + O(e^{-N}) = O(N^{-\widehat{\theta}}),$$

so that putting all those bounds together yield, by union bound,

$$\mathbb{Q}_A(\text{The construction failed before time } \tau') = O(N^{-\widehat{\theta}}) + O(N^4 N^{-\log N/2})$$
$$= O(N^{-\widehat{\theta}}). \qquad (5.7)$$

Step 4: estimation of the time spent with at least two flags close to the boundary. We introduce the time sets

$$I_t := \{s \leqslant t, \mathcal{B}'(s) \cap \{1,2\} \neq \emptyset\} := I_t^1 \cup I_t^2 \cup I_t^3,$$

where

$$I_t^1 = \bigcup_{j,\, s_j \leqslant t} [t_j, s_j)$$

$$I_t^2 = \{s \leqslant t, \mathcal{B}'(s) \cap \{1,2\} \neq \emptyset \text{ and } |\mathcal{B}'(s) \cap \{1, \ldots, \ell_N\}| \geqslant 2\}$$

and

$$I_t^3 = I_t \setminus (I_t^1 \cup I_t^2).$$

Note that for any $t > 0$, I_t, I_t^1 and I_t^2 can all be split into a disjoint union of a finite number of time segments of the form $[s, s')$, therefore so can I_t^3. We therefore write $I_3 = \sqcup_j [a_j, b_j)$. Since $I_t^3 \cap I_t^2 = \emptyset$, during each of the segments $[a_j, b_j)$, there is exactly one flag labeled $n = n(j)$ in $\{1, \ldots, \ell_N\}$ and this flag is in $\{1, 2\}$. Further note that we cannot have $Y_n(a_j^-) = 3$, because else, since there is no other flag in $\{1, \ldots, \ell_N\}$, a_j would have been the start of an excursion $Z^{j'}$, which is impossible since $I_t^3 \cap I_t^1 = \emptyset$. This means that at time a_j^-, another flag was at site ℓ_N and jumped to site $\ell_N + 1$ at time a_j. On the other hand, one can check that

$$b_j = \inf\{s \geqslant a_j,\, 2 \in \mathcal{B}(s) \text{ or } \ell_N \in \mathcal{B}(s)\}.$$

We claim that, letting

$$I_t^4 = \{s \leqslant t,\, |\mathcal{B}'(s) \cap \{1, \ldots, 2\ell_N\}| \geqslant 2\},$$

we have

$$\mathbb{Q}_A(I_{\tau'}^2 \cup I_{\tau'}^3 \not\subset I_{\tau'}^4) = O(N^{-\log N/4}). \qquad (5.8)$$

To prove this identity, first note that $I_{\tau'}^2 \subset I_{\tau'}^4$, therefore we only need to prove that with high probability $I_{\tau'}^3 \not\subset I_{\tau'}^4$. We already pointed out that if $[a_j, b_j)$ is one

of I_3's segments, at time a_j, a flag jumped from ℓ_N to $\ell_N + 1$, while another was
at the left boundary. Since the segment ends whenever the boundary flag reaches
site 3, in order to have $[a_j, b_j] \not\subset I_1^4$, the other flag must have reached site $2\ell_N + 1$
before time b_j. Once again, this means either that the boundary excursions lasted
more than $N^{-2}\ell_N$, or that the other flag traveled a distance of order ℓ_N in a time
$N^{-2}\ell_N$. Both of those probabilities are of order $O(e^{-\sqrt{\ell_N}}) = O(N^{-\log N})$, i.e. for
any integer j,

$$\mathbb{Q}_A([a_j, b_j] \not\subset I_{b_j}^4) = O(N^{-\log N}). \tag{5.9}$$

We now obtain a very rough estimate of the number of segments in $I_{\tau'}^2$, which we
bound by the total number of visits K to site ℓ_N by any flag occurring before time
τ',

$$K = |\{t \leqslant \tau', \ell_N \notin \mathcal{B}'(t^-) \text{ and } \ell_N \in \mathcal{B}'(t)\}|.$$

Not to burden the notation, simply denote $\mathbb{E}_A$ the expectation w.r.t. $\mathbb{Q}_A$. We
first write $\mathbb{E}_A(K) = \sum_{x \in A} \mathbb{E}_{\{x\}}(K)$. Each term of the sum is less than $\mathbb{E}_{\{\ell_N\}}(K)$.
Consider therefore a single flag initially at site ℓ_N. Each time it hits site ℓ_N, it has
a probability $1/2\ell_N$ of reaching the left boundary before getting back to site ℓ_N.
Once at the left boundary, it has a probability of order $O(N^{-\theta})$ of encountering a
boundary mark. In particular, we have $\mathbb{E}_{\{\ell_N\}}(K) = O(\ell_N N^\theta)$, so that

$$\mathbb{E}_A(|\{j \in \mathbb{N}, \ b_j \leqslant \tau'\}|) \leqslant \mathbb{E}_A(K) = |A|O(\ell_N N^\theta) = O(N^2).$$

Together with (5.9) and Markov's inequality, this bound proves (5.8).

We now estimate $|I_{\tau'}^4| := \int_0^{\tau'} \mathbf{1}_{|\mathcal{B}'(s) \cap \{1,\dots,2\ell_N\}| \geqslant 2} ds$. First, we write

$$|I_{\tau'}^4| \leqslant \sum_{n \neq n' \leqslant |A|} \int_0^{\tau'} \mathbf{1}_{\{Y_n(s), Y_{n'}(s) \in \{1,\dots,2\ell_N\}\}} ds.$$

In particular, according to (5.4) and Markov's inequality

$$\mathbb{Q}_A\left(|I_{\tau'}^4| \geqslant N^{\theta/2-2}\right) = \mathbb{Q}_A(\tau' \geqslant N^{-\widehat{\theta}}) + \mathbb{Q}_A(\tau' \leqslant N^{-\widehat{\theta}} \text{ and } |I_{\tau'}^4| \geqslant N^{\theta/2-2})$$

$$\leqslant O(N^{-\widehat{\theta}}) + |A|^2 N^{2-\theta/2} \max_{x,y \in A} \mathbb{E}_{\mathbb{Q}_{\{x,y\}}}\left(\int_0^{N^{-\widehat{\theta}}} \mathbf{1}_{\{Y_1(s), Y_2(s) \in \{1,\dots,2\ell_N\}\}} ds\right)$$

Assume that

$$\mathbb{E}_{\mathbb{Q}_{\{x,y\}}}\left(\int_0^{N^{-\widehat{\theta}}} \mathbf{1}_{\{Y_1(s), Y_2(s) \in \{1,\dots,2\ell_N\}\}} ds\right) = O(N^{-2}\ell_N^2 \log N) \tag{5.10}$$

uniformly in $x, y \in \Lambda_N$. Then, $\mathbb{Q}_A\left(|I_{\tau'}^4| \geqslant N^{\theta/2-2}\right) = |A|^2 O(N^{-\nu_\theta})$, where ν_θ was
defined in the statement of the Lemma. This bound together with (5.9) yields

$$\mathbb{Q}_A(|I_{\tau'}^2 \cup I_{\tau'}^3| \geqslant N^{\theta/2-2}) = |A|^2 O(N^{-\nu_\theta}),$$

and in particular, using (5.4), we finally obtain

$$\mathbb{Q}_A(|I_{N^{-\widehat{\theta}}}^2 \cup I_{N^{-\widehat{\theta}}}^3| \geqslant N^{\theta/2-2}) = |A|^2 O(N^{-\nu_\theta}), \tag{5.11}$$

where as before $|I_t^2 \cup I_t^3| = \int_{I_t^2 \cup I_t^3} ds$.

We now only need to prove (5.10). Since it is quite burdensome in terms of
notations, we will simply sketch the proof, it is rather elementary. See $\mathbf{Y}(s) =
(Y_1, Y_2)(s)$ as a two dimensional random walk on $\{(x, y) \in \Lambda_N^2, \ x \neq y\}$, reflected at

the boundaries $\Delta_1 = \{(1, y)$ or $(y, 1)\ y \in \Lambda_N \setminus \{1\}\}$ and $\Delta_2 = \{(N-1, y)$ or $(y, N-1),\ y \in \Lambda_N \setminus \{N-1\}\}$. Since we want an upper bound, we assume without loss of generality that $x, y \leqslant 2\ell_N$. Together, the four following claims, which we will not prove because they are elementary, prove (5.10):

(1) The random variable inside the expectation is bounded by $N^{-\widehat{\theta}} \leqslant 1$. Furthermore, since $x, y \leqslant 2\ell_N$, in a time $N^{-\widehat{\theta}}$ with probability $1 - O(e^{-CN^{\widehat{\theta}}})$, the random walk $\mathbf{Y}$ never hits the boundary Δ_2.

(2) The random walk $\mathbf{Y}$ performs excursions, either in the set $\Lambda = \{1, \ldots, 10\ell_N\}^2$ or in Λ^c. Each excursion in Λ lasts on average a time $O(\ell_N^2 N^{-2})$.

(3) Each time an excursion in Λ ends, the $\mathbf{Y}$ has a probability of order $\geqslant C/\log N$ of reaching Δ_2 before hitting $\{1, \ldots, 2\ell_N\}^2$, so that on average, $\mathbf{Y}$ performs $O(\log N)$ excursions in Λ before hitting Δ_2.

(4) This means that $\mathbf{Y}$ spends a time $O(\log N \ell_N^2 N^{-2})$ in $\{1, \ldots, 2\ell_N\}^2$ before hitting Δ_2. Together with the first claim, this proves (5.10).

Step 5: Proof of (5.2). We now have all the ingredients to prove (5.2). Give the flags an arbitrary label at time 0 (identical in $\mathcal{B}$ and $\mathcal{B}'$), and recall that we want to estimate $\mathbb{P}_A^{\dagger}(C \cap D_{\pm} \cap \{\tau = \tau_n\})$ for any $n \leqslant |A|$. Denote by τ_n' the first time the flag labeled n in $\mathcal{B}'$ meets a boundary mark. Finally, similarly to $D_{\pm}$, denote C' and $D_{\pm}'$ the events $C' = \{|\mathcal{B}'(\tau') \cap \{1, 2\}| = 1\}$ and $D_{\pm}' = \{$the mark met by $\mathcal{B}'$ a time τ' was of type $\pm\}$. Since both processes $\mathcal{B}$ and $\mathcal{B}'$ have the same distribution,

$$\mathbb{P}_A^{\dagger}(C \cap D_{\pm} \cap \{\tau = \tau_n\}) = \mathbb{Q}_A(C' \cap D_{\pm}' \cap \{\tau' = \tau_n'\}).$$

Denote by j^* the index of the first excursion Z^j to meet a boundary mark in $\mathcal{M}^j$, and let $D_{\pm}^*$ be the event $D_{\pm}^* = \{$The boundary mark met by Z^{j^*} was of type $\pm\}$. Further denote n^* the label of the flag performing the excursion Z^{j^*}. Recall that the flag labels go from 1 to $|A|$, we denote $n^* = 0$ if the construction failed before one of the excursions met a boundary marks. We claim that

$$\mathbb{Q}_A(C' \cap D' \cap \{\tau' = \tau_n'\}) = \mathbb{Q}_A(D_{\pm}^* \cap \{n^* = n\}) + |A|^2 O(N^{-\nu_\theta}). \qquad (5.12)$$

Denote τ_f the first time the construction fails, we proved in (5.7) that $\mathbb{Q}_A(\tau_f \leqslant \tau') = O(N^{-\widehat{\theta}})$. We now prove (5.12). We discard the possibility that the boundary mark was encountered at the right boundary, since with probability $O(N^{-\widehat{\theta}})$, as a consequence of (5.4), no flag made it past $N/2$ before time τ. The boundary mark encountered at time τ must therefore have appeared during I_t defined in (5.3). Furthermore, according to (5.11), and since boundary marks appear at rate $O(N^{2-\theta})$, the probability that a boundary mark appeared in $I_{N^{-\widehat{\theta}}}^2 \cup I_{N^{-\widehat{\theta}}}^3$ is $|A|^2 O(N^{-\theta/4})$. The first boundary mark encountered by $\mathcal{B}'$ must therefore have appeared with probability $1 - |A|^2 O(N^{-\nu_\theta})$ in I_t^1, i.e. during one of the inserted excursions, in which case, since we assumed the construction did not fail before it appeared, the mark was met by Z^{j^*}. This proves (5.12), because during the inserted excursions, the flag performing the excursion is alone at the boundary.

Note that all the excursions Z_j are independent from the process $(\mathcal{B}'(s))_{s \leqslant t_j}$ so that in particular, n^* being measurable w.r.t. $(\mathcal{B}'(s))_{s \leqslant t_{j^*}}$, we have

$$\mathbb{Q}_A(D_{\pm}^* \cap \{n^* = n\}) = \mathbb{Q}_A(D_{\pm}^*)\mathbb{Q}_A(n^* = n). \qquad (5.13)$$

Applying the same arguments as before, one obtains straightforwardly

$$\mathbb{Q}_A(\tau' = \tau'_n) = \mathbb{Q}_A(n^* = n) + |A|^2 O(N^{-\nu_\theta}). \tag{5.14}$$

Finally, given a flag at site 2 performing an excursion at the boundary conditioned to meeting a boundary mark (i.e. conditioned to not jumping to site 3 before meeting a boundary mark), letting $p_{1,\pm}$ (resp. p_2, p_3) the probability that the flag encountered a mark $\pm$ (resp. a branching mark, resp. no boundary mark) before coming back to site 2, one has the explicit formulas

$$p_{1,\pm} = \frac{(c + N^\theta) r \rho_\pm}{(b + c + N^\theta)(r + N^\theta)}, \; p_2 = \frac{b}{b + c + N^\theta} \; \text{and} \; p_3 = \frac{(c + N^\theta) N^\theta}{(b + c + N^\theta)(r + N^\theta)},$$

where $\rho_+ = \overline{\rho}$ and $\rho_- = 1 - \overline{\rho}$. In particular, assuming that an excursion met a boundary mark, the probability that the first boundary mark encountered was of type $\pm$, resp. branching, is

$$p_\pm^N = \frac{p_{1,\pm}}{p_{1,+} + p_{1,-} + p_2} = \frac{(c + N^\theta) r \rho_\pm}{(c + N^\theta) r + b(r + N^\theta)} = \frac{r \rho_\pm}{r + b} + O(N^{-\theta})$$

$$\text{resp.} \quad p_b^N = 1 - p_+^N - p_-^N = \frac{b}{r + b} + O(N^{-\theta}). \tag{5.15}$$

In particular $\mathbb{Q}_A(D_\pm^*) = p_\pm^N$, which, together with (5.12), (5.13) and (5.14), finally allows us to write

$$\mathbb{Q}_A(C' \cap D'_\pm \cap \{\tau' = \tau'_n\}) = \frac{r \rho_\pm}{r + b} \mathbb{Q}_A(\tau' = \tau'_n) + |A|^2 O(N^{-\nu_\theta}).$$

Since $\mathcal{B}'$ and $\mathcal{B}$ have the same distribution, this proves (5.2).

Step 6: Taking into account the right boundary. We now prove (5.3). By the Markov property, we first assume without loss of generality that $A \cap \{1, 2, N - 2, N - 1\} \neq \emptyset$, so that by an elementary adaptation of (5.4), $\mathbb{E}_A^\dagger(\tau) \leqslant CN^{\theta-1}$. Assuming $\tau \leqslant N^{-\widehat{\theta}}$, we can remove from A the flags initially at a distance more than $N^{1-\widehat{\theta}/4}$ of the boundaries, since one of those reaches the boundary before time τ with probability exponentially small in $N^{\widehat{\theta}/2}$. More precisely, let us denote

$$A_1 = A \cap \{1, \ldots, N^{1-\widehat{\theta}/4}\} \; \text{and} \; A_2 = A \cap \{N - N^{1-\widehat{\theta}/4}, \ldots, N - 1\},$$

according to (5.4), and shortening $D = D_+ \cup D_-$, we have

$$\mathbb{P}_A^\dagger(D) = \mathbb{P}_{A_1 \cup A_2}^\dagger(D) + O(N^{-\widehat{\theta}}).$$

Label initially the flags in increasing order from left to right. Then, according to (5.2) and its counterpart at the right boundary,

$$\mathbb{P}_{A_1 \cup A_2}^\dagger(D) = \sum_{n=1}^{|A_1 \cup A_2|} \mathbb{P}_{A_1 \cup A_2}^\dagger(D \text{ and } \tau = \tau_n)$$

$$= \frac{r}{r + b} \sum_{n=1}^{|A_1|} \mathbb{P}_{A_1 \cup A_2}^\dagger(\tau = \tau_n) + \frac{r'}{r' + b'} \sum_{n=|A_1|+1}^{|A_1 \cup A_2|} \mathbb{P}_{A_1 \cup A_2}^\dagger(\tau = \tau_n)$$

$$+ |A|^2 O(N^{-\nu_\theta}) \geqslant p + |A|^2 O(N^{-\nu_\theta})$$

where p was defined after (5.3). This concludes the proof. $\square$

Now that this Lemma is proved, we get back to the process $\mathcal{A}$ on which boundary marks have an effect. Start from $A = \{3\}$, and give the label 1 to the flag at site 3. Each time a flag is created by the generator (5.1), it is given the smallest label not already used up until this point. For any $t \geqslant 0$, we denote

$$\kappa(t) = 1 + \left| \left\{ s \leqslant t, |\mathcal{A}(s)| = |\mathcal{A}(s^-)| + 1 \right\} \right|. \tag{5.16}$$

the total number of labels used up to time s. Finally, we define

$$T = \inf\{s \geqslant 0, \ \mathcal{A}(s) = \emptyset\}. \tag{5.17}$$

Corollary 5.3. *Recall* (2.6). *For any $c > 0$, there exists $\varepsilon_1(c) > 0$ such that*

$$\mathbb{P}^\dagger_{\{3\}}(\kappa(T) > c \log N) = O(N^{-\varepsilon_1(c)}), \tag{5.18}$$

There also exists $\varepsilon_2 > 0$ such that

$$\mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2}) = O(N^{-\varepsilon_2}), \tag{5.19}$$

$$\mathbb{P}^\dagger_{\{3\}}(\exists t \in [0, T], \ \mathcal{A}(t) \ni N/2) = O(N^{-\varepsilon_2}), \tag{5.20}$$

$$\mathbb{P}^\dagger_{\{3\}}(\mathcal{A} \text{ encountered a boundary mark while it contained } \{1, 2\}) = O(N^{-\varepsilon_2}). \tag{5.21}$$

Proof of Corollary 5.3: Fix $c > 0$, We first prove (5.18), which is a consequence of equation (5.3), which states that regardless of the initial set, the probability that the first boundary mark encountered was of type $\pm$ (and therefore deleted a flag) is $p + o_N(1) > 1/2$. Consider the times $0 < t_1 < \cdots < t_m = T$ at which $|\mathcal{A}|$ changes. Set $t_0 = 0$, and consider the discrete time random walk $Y_k = |\mathcal{A}(t_k)| \in \mathbb{N}$ for $k \geqslant 0$. Note that Y_k is not a Markov process in itself since its jump rate depends on $\mathcal{A}(t_k)$. However, assuming $\kappa(T) > k$, one easily checks that $Y_k > 0$: else, there has been less than k updates of $|\mathcal{A}|$ before it became empty, so in particular the process must have branched less than k times. (Note that $Y_k = 0$ iff $\mathcal{A}(t_k) = \emptyset$). Furthermore, after j updates of $|\mathcal{A}|$, the latter cannot be more than $1 + j$. According to (5.3), there exists a constant C such that

$$\mathbb{P}^\dagger_{\{3\}}(Y_k > 0) \leqslant \mathbb{P}\left(1 + \sum_{j=1}^k B_j > 0\right),$$

where the B_j's are independent variables taking the value -1 (resp. 1) w.p. $q_j := p + C(j+1)^2 N^{-\nu_\theta}$ (resp. $1 - q_j$). We refer the reader to lemma 5.2 of Erignoux et al. (2018) for more details. In particular, by Markov inequality, for any positive λ

$$\mathbb{P}^\dagger_{\{3\}}(\kappa(T) > k) \leqslant \mathbb{P}^\dagger_{\{3\}}(Y_k > 0) \leqslant \mathbb{P}\left(\exp\left[\sum_{j=1}^k \lambda B_j\right] > e^{-\lambda}\right) \leqslant e^\lambda \prod_{j=1}^k \mathbb{E}\left(e^{\lambda B_j}\right).$$

Choose $p' = 1/4 + p/2 \in (1/2, p)$, for N large enough, and any $j \leqslant c \log N$, we have $q_j \geqslant p'$, so that for any N large enough and for any $\lambda > 0$

$$\mathbb{P}^\dagger_{\{3\}}(\kappa(T) > c \log N) \leqslant e^\lambda \left(p' e^{-\lambda} + (1 - p') e^\lambda\right)^{c \log N}.$$

Choose $\lambda = (p' - 1/2)^2$ small enough so that $p' e^{-\lambda} + (1 - p') e^\lambda < 1$ to obtain as wanted that there exists $\varepsilon_1 = \varepsilon_1(c)$ such that

$$\mathbb{P}^\dagger_{\{3\}}(\kappa(T) > c \log N) \leqslant O(N^{-\varepsilon_1(c)}),$$

which proves (5.18).

The second bound (5.19) is a direct consequence of the first. Thanks to the first bound, we first write for some $\varepsilon := \varepsilon_1(1) > 0$

$$\mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2}) \leqslant \mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2} \text{ and } \kappa(T) \leqslant \log N) + O(N^{-\varepsilon}).$$

We then bound from above T by the sum of the lifespans T_k (the difference between the time the label k encounters a boundary mark of type $\pm$, and the time the label k is introduced) of its flags, we obtain, since the lifespan only increases when a flag starts at site 3 instead of sites $1, 2$,

$$\mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2} \text{ and } \kappa(T) \leqslant \log N)$$

$$\leqslant \sum_{n=1}^{\log N} \mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2} \text{ and } \kappa(T) = n)$$

$$\leqslant \sum_{n=1}^{\log N} \mathbb{P}^\dagger_{\{3\}}\left(\sum_{k \leqslant n} T_k > N^{-\widehat{\theta}/2} \text{ and } \kappa(T) = n\right)$$

$$\leqslant \sum_{n=1}^{\log N} \mathbb{P}^\dagger_{\{3\}}\left(\max_{k \leqslant n} T_k > \frac{N^{-\widehat{\theta}/2}}{n} \text{ and } \kappa(T) = n\right)$$

$$\leqslant (\log N)^2 \mathbb{P}^\dagger_{\{3\}}\left(T_1 > \frac{N^{-\widehat{\theta}/2}}{\log N} \text{ and } \kappa(T) \leqslant \log N\right)$$

On the event $\kappa(T) \leqslant n$, the flag labeled 1 has branched at most n times. In particular, we can use (5.5) to obtain that

$$\mathbb{E}^\dagger_{\{3\}}\left(T_1 \mathbf{1}_{\{\kappa(T) \leqslant \log N\}}\right) \leqslant \log N \mathbb{E}^\dagger_{\{3\}}(\tau),$$

where as in (5.5), τ is the time the flag 1 waits before something happens to it at the left boundary. The right-hand side above, according to (5.5), is $O(N^{\theta-1} \log N)$, therefore by Markov inequality,

$$\mathbb{P}^\dagger_{\{3\}}(T > N^{-\widehat{\theta}/2}) \leqslant O(N^{-\widehat{\theta}}(\log N)^4) + O(N^{-\varepsilon}),$$

which proves (5.19) by choosing ε_2 strictly smaller than both $\widehat{\theta}$ and ε.

Identity (5.20) is an immediate consequence of the first two : assuming that $\kappa(T) \leqslant \log N$ and $T \leqslant N^{-\widehat{\theta}}$, by union bound, the probability that a flag reaches $x = N/2$ is less than $\log N \mathbb{P}_3\left(\sup_{0 \leqslant t \leqslant N^{-\widehat{\theta}}} X(t) \geqslant N/2\right)$, which is the probability that a random walker on $\{3, \ldots, N\}$, starting at site 3, jumping at rate N^2 and reflected at the boundaries, visits site $N/2$ before time $N^{-\widehat{\theta}}$, which is of order $O(e^{-N^{\widehat{\theta}/2}})$, thus proving (5.20).

Finally, (5.21) is a consequence of (5.18) and (5.11). The latter immediately yields, for any $A \subset \{1, \ldots, N/2\}$ and such that $A \cap \{1, 2\} \neq \emptyset$

$$\mathbb{P}^\dagger_A(\{1, 2\} \subset \mathcal{A}(\tau^-)) = |A|^2 O(N^{-\nu_\theta}). \tag{5.22}$$

Note we relaxed slightly the assumption on the set A, which can contain either site 1 or 2 and not just 2. This is not an issue, since with probability $1 - O(N^{-\theta})$, any flag starting from site 1 or 2 reaches site 3 before any boundary mark appeared.

Equation (5.22) is therefore a simple consequence of Markov's inequality and the fact that $\nu_\theta < \theta$. Assuming $\kappa(T) \leqslant \log N$, we have $|\mathcal{A}(s)| \leqslant 1 + \log N$ for any $s \leqslant T$, and at most $2 \log N + 2$ boundary marks were encountered by $\mathcal{A}$ before time T. On $\{\kappa(T) \leqslant \log N\} \cap \{\sup_{t \leqslant T} \mathcal{A}(t) < N/2\}$, we therefore use the Markov property together with (5.22) $2 \log N + 2$ times, to obtain that the probability in the left hand-side of (5.21) is less than

$$(2 \log N + 2)(\log N + 1)^2 O(N^{-\nu_\theta}) + \mathbb{P}_A^\dagger \left(\kappa(T) \geqslant \log N \ \text{ or } \ \sup_{t \leqslant T} \mathcal{A}(t) \geqslant N/2 \right).$$

This proves (5.21) and the Lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.4. *Determination tree.* Now that we studied the process $\mathcal{A}$, we can, with it, determine the nature of $\eta_t(x)$. Recall from Section 5.2 that the process $\mathcal{A}$ observed *forward in time* $0 \to t$, represents the evolution of the set of unknowns $\hat{\mathcal{A}}$ $t \to 0$ *backward in time*. To determine the value of $\eta_t(x)$, we start the process $\mathcal{A}$ from $\mathcal{A}(0) = \{x\}$, and, following the time evolution of $\mathcal{A}(s)$, build a labeled rooted tree $\mathcal{T}_s$, for $s \in [0, t]$. An example of this construction is represented in Figure 5.3. Define $\mathcal{T}_0$ as the trivial one-vertex rooted tree, and label 1 its only vertex corresponding to the label of the flag in $\mathcal{A}(0)$. Then, the tree remains unchanged until the first time τ at which the process $\mathcal{A}$ encounters a boundary mark. if the mark is of type $\pm$, we give the root a unique child, labeled $\pm$. If the mark is of branching type, we give the root two children, the first one labeled 1 as well, and the second one labeled 2.

We then carry on with the construction until time $T \wedge t$: each time one of the flags (labeled k) in $\mathcal{A}$ encounters a boundary mark at time τ, we build $\mathcal{T}_\tau$ from $\mathcal{T}_{\tau^-}$, with the following rules depending on the mark encountered by $\mathcal{A}$ at time τ. Denote k the label of the flag affected by the mark τ.

- If the mark is of type $\pm$, we give all leaves labeled k in $\mathcal{T}_{\tau^-}$ a unique child labeled $\pm$. This is the case where the value of site $X_k(\tau)$ was updated according to a reservoir (i.e. was filled if the mark was of type $+$, or emptied if the mark was of type $-$).
- If the mark is of type $(b, 2)$ (resp. $(b, N-2)$, we give all leaves labeled k in $\mathcal{T}_{\tau^-}$ two children. We label the first one k, and the second one k'. The integer k' is either the label of the flag at site 2 if $2 \in \mathcal{A}(\tau)$ (resp. at site $N - 2$ if $N - 2 \in \mathcal{A}(\tau)$), or $k' = \kappa(\tau) = \kappa(\tau^-) + 1$ which is the smallest unused label until now. This is the case where site 2 (resp. $N - 2$) was filled if site 1 was occupied, and nothing happened otherwise.

We carry on with this construction until $T_t := T \wedge t$. At time T_t, either the processes died ($\mathcal{A}(T_t) = \emptyset$), in which case the last leaf labeled k received a unique child labeled $\pm$, or it didn't, and there remains in $\mathcal{T}_{T_t}$ some leaves with labels k corresponding to the flag's labels in $\mathcal{A}(T_t)$. Up to this point we neglected the possibility that one of the boundary marks encountered was of type c or c' (recall that it requires for two flags to be at the same boundary at the time of the mark), in this case, we say that the construction failed. According to (5.21), the probability that the construction of the tree fails is $O(N^{-\varepsilon_2})$.

If $T_t := T$, we just let $\mathcal{T}_s = \mathcal{T}_T$ constant in $s \in [T, t]$. If however $T_t := t$, in order to build $\mathcal{T}_{T_t} = \mathcal{T}_t$ we give each leaf labeled k in $\mathcal{T}_{t^-}$ a unique child labeled $+$ (resp. $-$) if $\eta_0(X_k(t)) = 1$ (resp. $\eta_0(X_k(t)) = 1$) (where $X_k(t)$ is the position of

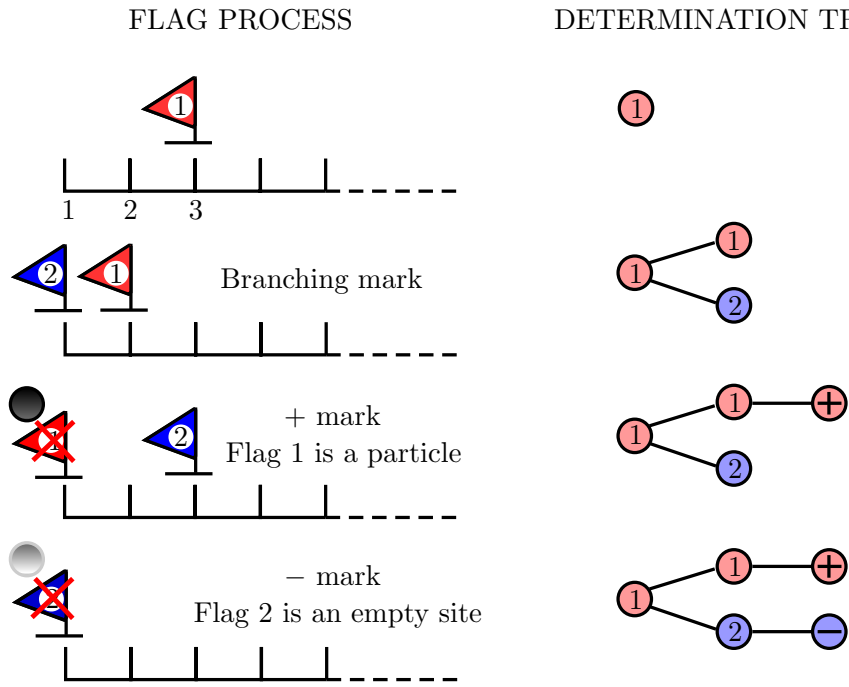FLAG PROCESS   DETERMINATION TREE

FIGURE 5.3. An example of construction of the determination tree. Each time a boundary mark is encountered by one of the flags in the process $\mathcal{A}$, the tree is updated accordingly, by adding two children (for a branching mark) to each node with the affected flag's label, or one single child labelled $\pm$ (for a $\pm$ mark).
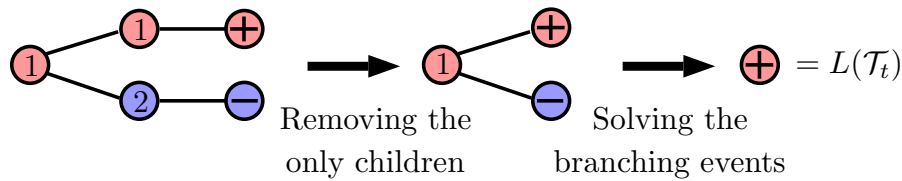

FIGURE 5.4. Solving the determination tree.

the flag labeled k in Λ_N at time t). For any rooted tree $\mathcal{T}$, and any vertex $v \in \mathcal{T}$, we denote $c(v)$ the number of its children, and $s(v)$ the number of its siblings (i.e. the number of children of its parent). We call *only children* the vertices such that $s(v) = 1$. One easily checks that, if the construction did not fail, the tree $\mathcal{T}_t$ has the following properties:

  (1) each leaf has a label $\pm$, and no other vertex has a label $\pm$.
  (2) Each vertex satisfies $c(v) \in \{0, 1, 2\}$.
  (3) The leaves are exactly the only children, i.e. $c(v) = 0$ iff $s(v) = 1$.

We denote by $\mathbb{T}$ the set of rooted trees satisfying these three properties.

Assuming the construction did not fail, we now recover, as represented in Figure 5.4 the value of $\eta_t(x)$ by "solving" the tree $\mathcal{T}_t$. We start by deleting the *only children*. For each only child deleted, we give their parent, which are now leaves, the same label $\pm$ its child had. Once this procedure is realized, there are no more only children, so that each vertex of the remaining tree has either two or zero children. Then, one vertex at a time, we choose an arbitrary vertex v with two children (v_1, v_2) which are both leaves. Then, we delete v_1 and v_2 from the tree, and give v:

- the label $+$ if v_2's label is also $+$, regardless of v_1's label. This is the case where there was a flag at site 1 or $N-1$, and a mark of branching type occurred. At the time of the mark, a particle was therefore placed at site 2 or $N-2$.
- the same label as v_1 if v_2's label is $-$. This is the case where at the time of the mark of branching type, there was not a particle at site 1 or $N-1$, so that the value of site 2 or $N-2$ did not change.

Ultimately, this procedure deletes the entire tree except the root, and gives the root a label $\pm := L(\mathcal{T}_t)$. This "solving" procedure can be defined for any tree in the set $\mathbb{T}$, so that we see L as a function $\mathbb{T} \to \{+, -\}$.

5.5. *Proof of Lemmas 3.2 and 4.2.* This construction is justified by the following result, which we will not prove because it is strictly analogous to Lemma 5.1, p.23 of Erignoux et al. (2018).

Lemma 5.4. *Assuming that the construction did not fail, we have*

$$\eta_t(x) = 1 \quad iff \quad L(\mathcal{T}_t) = +.$$

Recall the definition of the outcome's probabilities of an excursion $p_\pm^N$ and p_b^N introduced in (5.15), we define the distribution $m_\rho^N(\cdot)$ of a labeled random tree $\mathcal{T} \in \mathbb{T}$, built as follows :

- The root is labeled $*$.
- As long as there is still a leaf labeled $*$ in the tree, one such vertex chosen arbitrarily receives two children with probability p_b^N, each labeled $*$, and receives a unique child labeled $\pm$ w.p. $p_\pm^N$.
- The construction ends when there is no longer any leaf labeled $*$ (note that we assumed $r_2 > b$, so that a.s., for N large enough, this construction ends after a finite number of steps).

For any tree $\mathcal{T} \in \mathbb{T}$, we denote $|\mathcal{T}|$ its number of vertices. The main result of this section is the following.

Lemma 5.5. *There exists $\varepsilon_3 > 0$ such that for any $\mathcal{T} \in \mathbb{T}$,*

$$\sup_{t \geqslant N^{-\hat{\theta}}} \left| \mathbb{P}_{\{3\}}^\dagger(\mathcal{T}_t = \mathcal{T}) - m_\rho^N(\mathcal{T}) \right| = O(N^{-\varepsilon_3}).$$

where the identity $\mathcal{T}_t = \mathcal{T}$ means that the structure of the tree is the same, and that the labels $\pm$ of the leaves are identical as well.

Before proving the Lemma, we state the following result.

Corollary 5.6. *There exists $\varepsilon_4 > 0$ such that ,*

$$\sup_{t \geqslant N^{-\hat{\theta}}} \left| \rho_t^N(3) - \alpha \right| = \sup_{t \geqslant N^{-\hat{\theta}}} \left| \rho_t^N(3) - m_\rho^N(L(\mathcal{T}) = +) \right| + O(N^{-\theta}) = O(N^{-\varepsilon_4}),$$

where α is given by (2.3).

Proof of Lemma 5.5: Fix the initial set $A = \{3\}$. Given the process $\mathcal{A}$ built with the Poisson marks in $\mathcal{M}$, and given a family of independent left boundary excursions Z^j, we build in the same way that we built in Lemma 5.2 $\mathcal{B}'$ from $\mathcal{B}$ a process $\mathcal{A}'$ by inserting the independent excursions each time a flag reaches site 2 while no other flag is in $\{1, \ldots, \ell_N\}$, where $\ell_N = (\log N)^4$. However, instead of stopping the excursions Z at the time τ_Z when they reach site 3, we stop them at the first time τ_Z^m they either hit site 3 or meet a boundary mark. In particular, the time s_j at which we stop the inserted excursion is

$$s_j = (t_j + \tau_{Z^j}^m) \wedge \inf\{s \geqslant t_j, \, 3 \in \mathcal{A}'(s)\}.$$

We do not detail this construction here, it is exactly identical to the one performed in step 2 of the proof of Lemma 5.2, except that the process is affected by the boundary marks. As for $\mathcal{B}$, the two processes $\mathcal{A}$ and $\mathcal{A}'$ have the same distribution. We also denote $\mathbb{Q}_A$ their joint distribution, and T' and $\kappa'(t)$ the counterparts of T and $\kappa(t)$ (cf. (5.17) and (5.16)) for $\mathcal{A}'$. Once again, we say that this construction failed after inserting the j-th excursion if another flag reached site 3 before the j-th excursion Z^j reached site 3 or met a boundary mark, i.e. if $s_j \neq (t_j + \tau_{Z^j}^m)$. Define the events

$$F = \{\text{the construction failed before time } T'\},$$

$$M = \{\text{A boundary mark occurred before } T' \text{ outside of an excursion}\},$$

$$O = \{\exists s \in [0, T'], \, \mathcal{A}'(s) \ni N/2\} \quad \text{and} \quad S = \left\{T' \geqslant N^{-\widehat{\theta}}\right\}.$$

Using the same construction laid out at the beginning of Section 5.4, we build with the process $\mathcal{A}'$ a tree $\mathcal{T}'_t$ distributed as $\mathcal{T}_t$ because $\mathcal{A}$ and $\mathcal{A}'$ have the same distribution.

We further build a third process $\mathcal{A}''$, evolving exactly as $\mathcal{A}'$, except that:

- the boundary marks occurring outside of the inserted excursions are ignored.
- The time horizon t is ignored as well, and when reaching time t, $\mathcal{A}''$ keeps evolving by following the marks in $\mathcal{M}$ and inserting independent excursions until $T'' := \inf\{s \geqslant 0, \, \mathcal{A}''(s) = \emptyset\}$.

We finally build a third tree $\mathcal{T}''_t$ using $\mathcal{A}''$ following the same construction as before. On the event $F^c \cap M^c \cap O^c \cap S^c$, we have for any $t \geqslant N^{-\widehat{\theta}}$

$$\mathcal{A}'' = \mathcal{A}' \text{ on } [0, t],$$

so that in particular $\mathcal{T}'_t = \mathcal{T}''_t = \mathcal{T}''_{T''}$. Furthermore, by construction, $\mathcal{T}''_{T''}$ is distributed according to $m^N_{\frac{\rho}{\rho}}$. In particular, for any tree $\mathcal{T} \in \mathbb{T}$, and any $t \geqslant N^{-\widehat{\theta}}$

$$|\mathbb{Q}_{\{3\}}(\mathcal{T}'_t = \mathcal{T}) - m^N_{\frac{\rho}{\rho}}(\mathcal{T})| \leqslant 2\mathbb{Q}_{\{3\}}(F \cup M \cup O \cup S). \tag{5.23}$$

Since $\mathcal{A}$ and $\mathcal{A}'$ have the same distribution, according to (5.19) and (5.20) in Corollary 5.3,

$$\mathbb{Q}_{\{3\}}(O \cup S) = O(N^{-\varepsilon_2}).$$

Furthermore, using the same arguments used to prove (5.12) and the Markov property, on the event $\kappa'(T') \leqslant \log N$, one straightforwardly obtains

$$\mathbb{Q}_{\{3\}}(M \cap \{\kappa'(T') \leqslant \log N\}) = O((\log N)^3 N^{-\nu_\theta}). \tag{5.24}$$

We do not detail this step, it is enough to use (5.12) on each interval $[\tau_i, \tau_{i+1})$ of time between two consecutive boundary marks are met. With probability $1 - |\mathcal{A}(\tau_i)|^2 O(N^{-\nu_\theta})$, each of those boundary marks were met during an excursion, and on $\kappa'(T') \leqslant \log N$ there are at most $2\log N + 1$ such intervals, so that by union bound one obtains (5.24). In particular, using (5.18) and letting $\varepsilon = \min(\varepsilon_1(1), \nu_\theta/2)$ yields $\mathbb{Q}_{\{3\}}(M) = O(N^{-\varepsilon})$.

Similarly, using (5.7), a union bound and Markov's property, we obtain that $\mathbb{Q}_{\{3\}}(F \cap \{\kappa'(T') \leqslant \log N\}) \leqslant O(N^{-\widehat{\theta}} \log N)$, therefore using (5.18) and letting $\varepsilon' = \varepsilon_1(1) \wedge (\widehat{\theta}/2)$ yields $\mathbb{Q}_{\{3\}}(F) = O(N^{-\varepsilon'})$. Finally, letting $\varepsilon_3 = \min(\varepsilon_2, \varepsilon, \varepsilon')$, the right-hand side of (5.23) is $O(N^{-\varepsilon_3})$ for any $t \geqslant N^{\widehat{\theta}}$, which proves Lemma 5.5. $\quad\square$

Proof of Corollary 5.6: We start with the first identity. To prove it, one only needs to show that $m_{\bar{\rho}}^N(L(\mathcal{T}) = +) = \alpha + O(N^{-}\theta)$. Define the events

$$R_\pm = \{\text{the root has only one child, labeled } \pm\},$$

and

$$R_b = \{\text{the root has two children } v_1 \text{ and } v_2\}.$$

On R_b, for $i = 1, 2$, we denote $\mathcal{T}_i$ the sub-tree of $\mathcal{T}$ composed of v_i and its descendants. Note that conditionally to R_b, $\mathcal{T}_1$ and $\mathcal{T}_2$ are independent and distributed according to $m_{\bar{\rho}}^N$. Furthermore, by construction of the application L, we have the identity

$$\{L(\mathcal{T}) = +\} = R_+ \cup (R_b \cap \{L(\mathcal{T}_2) = +\}) \cup (R_b \cap \{L(\mathcal{T}_2) = -\} \cap \{L(\mathcal{T}_1) = +\}).$$

The three events in the union above are disjoint, so that taking the measure $m_{\bar{\rho}}^N$ of both sides of the identity above yields, shortening $\alpha_N = m_{\bar{\rho}}^N(L(\mathcal{T}) = +)$

$$\alpha_N = p_+^N + p_b^N \alpha_N + p_b^N (1 - \alpha_N) \alpha_N,$$

which rewrites using the definition (5.15) of $p_\pm^N$ and p_b^N,

$$r(\bar{\rho} - \alpha_N) + b\alpha_N(1 - \alpha_N) = O(N^{-\theta}),$$

which determines the boundary conditions for the equation, since it proves $\alpha_N = \alpha + O(N^{-\theta})$ defined in (2.3) as wanted.

We now prove that the supremum in the second term in the Corollary is $O(N^{-\varepsilon_4})$. Fix $\widehat{\theta} > 0$ and $t \geqslant N^{-\widehat{\theta}}$, and shorten $\mathbb{T}_k = \{\mathcal{T} \in \mathbb{T} \mid |\mathcal{T}| \leqslant k\}$

$$\rho_t^N(3) = \mathbb{P}_{\{3\}}^\dagger(L(\mathcal{T}_t) = +)$$
$$= \mathbb{P}_{\{3\}}^\dagger(L(\mathcal{T}_t) = + \text{ and } |\mathcal{T}_t| > k) + \sum_{\mathcal{T} \in \mathbb{T}_k} \mathbb{P}_{\{3\}}^\dagger(L(\mathcal{T}) = + \text{ and } \mathcal{T}_t = \mathcal{T}).$$

In particular, thanks to Lemma 5.5, we have for any k and any $t \geqslant N^{-\widehat{\theta}}$

$$\left| \rho_t^N(3) - m_{\bar{\rho}}^N(L(\mathcal{T}) = +) \right| \leqslant \mathbb{P}_{\{3\}}^\dagger(|\mathcal{T}_t| > k) + m_{\bar{\rho}}^N(|\mathcal{T}_t| > k) + O(|\mathbb{T}_k| N^{-\varepsilon_3}). \quad (5.25)$$

We now obtain a crude estimate on $|\mathbb{T}_k|$. Forgetting the leave's labels, to each tree $\mathcal{T} \in \mathbb{T}_k$, one can associate a unique full binary tree (whose vertices all have either 0 or 2 children) by removing all the leaves in $\mathcal{T}$, which are by assumption the only children. Since there are at most k leaves in a tree with k vertices, there are at most 2^k ways to associate to each leaf a label $\pm$. In particular, $|\mathbb{T}_k|$ is less than 2^k

times the number of full binary trees with less than k vertices. The number of full binary trees with k vertices is Catalan's number $C_{k-1} = O(4^k)$. In particular,

$$|\mathbb{T}_k| \leqslant 2^k \sum_{n=1}^{k} C_{n-1} = O(8^k).$$

We now choose $k = \frac{\varepsilon_3}{2 \log 8} \log N$, which yields that the last term in (5.25) is $O(N^{\varepsilon_3/2})$. According to (5.18), the first term in (5.25) is $O(N^{-\varepsilon_1(\varepsilon_3/2 \log 8)})$, whereas by construction of $m_{\bar{\rho}}^N$ the second term is $O(N^{-\varepsilon'})$ for some positive ε'. Choosing

$$\varepsilon_4 = \min(\varepsilon_3/2, \varepsilon_1 (\varepsilon_3/2 \log 8), \varepsilon')$$

proves Corollary 5.6. $\square$

Proof of Lemma 4.2: In order to estimate the correlations between site 3 and site $\ell \geqslant \delta N$, we start the process $\mathcal{A}$ from $\{3, \ell\}$, and denote $\mathcal{A}^{\{3\}}$, $\mathcal{A}^{\{\ell\}} \subset \mathcal{A}$ the sets of descendants of the flag initially at 3, ℓ. Let us denote τ the first time $\mathcal{A}^{\{3\}}$, $\mathcal{A}^{\{\ell\}}$ encounter,

$$\tau = \sup \left\{ t \leqslant 0, \ d\left(\mathcal{A}^{\{3\}}(t), \mathcal{A}^{\{\ell\}}\right)(t) \leqslant 1 \right\},$$

with the convention $d(\emptyset, A) = \infty$ for any set A. Also denote by $T^{\{3\}}$ the lifespan of $\mathcal{A}^{\{3\}}$,

$$T^{\{3\}} = \sup \left\{ t \leqslant 0, \ \mathcal{A}^{\{3\}}(t) = \emptyset \right\}.$$

Then, up until time τ, $\mathcal{A}^{\{3\}}$ and $\mathcal{A}^{\{\ell\}}$ can be coupled with two independent copies $\widetilde{\mathcal{A}}^{\{3\}}$ and $\widetilde{\mathcal{A}}^{\{\ell\}}$. In particular, we can write

$$|\varphi_t^N(3, \ell)| \leqslant \widehat{\mathbb{P}}_{\{3,l\}} \left(\tau > t \vee T^{\{3\}} \right) \leqslant \widehat{\mathbb{P}}_{\{3,\ell\}} \left(\tau > T^{\{3\}} \right).$$

To estimate the right-hand side, recall from (5.18) that with probability $1 - O(N^{-\varepsilon_1(1)})$, the total number of flags $C(T^{\{3\}})$ created by $\mathcal{A}^{\{3\}}$ is less than $\log N$. Further recall according to (5.19), $\widehat{\mathbb{P}}_{\{3,k\}}(T^{\{3\}} > N^{-\widehat{\theta}}) = O(N^{-\varepsilon_2})$. Finally, the probability that a flag travelled a distance at least $\delta N/4$ in a time $N^{-\widehat{\theta}}$ is of order $e^{-c\delta^2 N^{\widehat{\theta}}}$. In summary, in order to have $\tau > T^{\{3\}}$, one of three cases must have occurred.

- Either $\kappa(T) \geqslant \log N$, which occurs with probability $O(N^{-\varepsilon_1(1)})$.
- Or $T^{\{3\}} > N^{-\widehat{\theta}}$, which occurs with probability $O(N^{-\varepsilon_2})$.
- Or finally one of the (at most) $\log N + 2$ flags (either one of the two flags initially in the system or one of the flags created) must have travelled a distance $\delta N/4$ before a time $N^{-\varepsilon}$, which by union bound occurs with probability $O(\log N e^{-c\delta^2 N^{\widehat{\theta}}})$.

Letting $\varepsilon = \varepsilon_1 \wedge \varepsilon_2$, with probability $1 - O(N^{-\varepsilon})$, the two processes $\mathcal{A}^{\{3\}}$, $\mathcal{A}^{\{\ell\}}$ evolve independently up until the process $\mathcal{A}^{\{3\}}$ dies. All these bounds being independent of $\ell \geqslant \delta N$, this proves

$$\sup_{\delta N \leqslant \ell \leqslant N-1} |\varphi_t^N(3, \ell)| = O(N^{-\varepsilon}).$$

The second bound is identical. $\square$

References

De Masi, A., Presutti, E., Tsagkarogiannis, D., and Vares, M. E. Truncated correlations in the stirring process with births and deaths. *Electron. J. Probab.*, **17**, no. 6, 35 (2012). MR2878785.

Erignoux, C. Hydrodynamic limit of boundary driven exclusion processes with nonreversible boundary dynamics. *J. Stat. Phys.*, **172** (5), 1327–1357 (2018). MR3856946.

Erignoux, C., Gonçalves, P., and Nahum, G. Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part I, the critical regime and beyond (2020+). To appear in *J. Stat. Phys.*

Erignoux, C., Landim, C., and Xu, T. Stationary states of boundary driven exclusion processes with nonreversible boundary dynamics. *J. Stat. Phys.*, **171** (4), 599–631 (2018). MR3790154.