Gaussian scaling for the critical spread-out contact process above the upper critical dimension

Remco van der Hofstad[*]
Akira Sakai[†]

August 7, 2003

Abstract

We consider the critical spread-out contact process in $\mathbb{Z}^d$ with $d \geq 1$, whose infection range is denoted by $L \geq 1$. The two-point function $\tau_t(x)$ is the probability that $x \in \mathbb{Z}^d$ is infected at time t by the infected individual located at the origin $o \in \mathbb{Z}^d$ at time 0. We prove Gaussian behavior for the two-point function with $L \geq L_0$ for some finite $L_0 = L_0(d)$ for $d > 4$. When $d \leq 4$, we also perform a local mean-field limit to obtain Gaussian behaviour for $\tau_{tT}(x)$ with $t > 0$ fixed and $T \to \infty$ when the infection range depends on T such that $L_T = LT^b$ for any $b > (4-d)/2d$.

The proof is based on the lace expansion and an adaptation of the inductive approach applied to the discretized contact process. We prove the existence of several critical exponents and show that they take on mean-field values. The results in this paper provide crucial ingredients to prove convergence of the finite-dimensional distributions for the contact process towards the canonical measure of super-Brownian motion, which we defer to a sequel of this paper.

1 Introduction and results

1.1 Introduction

The contact process is a model for the spread of an infection among individuals in the d-dimensional integer lattice $\mathbb{Z}^d$. We suppose that the origin $o \in \mathbb{Z}^d$ is the only infected individual at time 0, and that every infected individual may infect a healthy individual at a distance less than $L \geq 1$. We refer to this model as the *spread-out* contact process. The rate of infection is denoted by λ, and it is well known that there is a phase transition in λ (see e.g., [20]).

Sakai [24, 25] has proved that when $d > 4$, the sufficiently spread-out contact process has several critical exponents which are equal to those of branching random walk. The proof by Sakai uses the lace expansion for the time-discretized contact process, and by proving the so-called infrared bound uniformly in the time discretization. Thus, we can think of his results as proving Gaussian upper bounds for the two-point function of the critical contact process. Together with the results by Barsky and Wu [3], and the fact that these Gaussian upper bounds imply the so-called triangle condition, it follows that certain critical exponents take on their mean-field values, i.e., the values for branching random walk. These values also agree with the critical exponents appearing on the tree. See [20, Section 2.4] for an extensive account of the contact process on a tree.

Recently, van der Hofstad and Slade [18] proved that for all $r \geq 2$, the r-point functions for sufficiently spread-out critical oriented percolation with spatial dimension $d > 4$ converge to those of the canonical measure of super-Brownian motion when we scale space by $n^{1/2}$, where n is the largest temporal component among those r points, and then take $n \uparrow \infty$. Their result implies that the finite-dimensional

---

[*] Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. `r.w.v.d.hofstad@TUE.nl`

[†] EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. `sakai@eurandom.tue.nl`

distributions of the critical oriented percolation cluster when it survives up to time n converge to those of the canonical measure of super-Brownian motion. The proof in [18] is based upon the lace expansion and the inductive method of [17].

Important ingredients in [18] are detailed asymptotics and estimates of the oriented percolation two-point function. The proof for the higher-point functions then follows by a lace expansion on the r-point function together with induction in r.

In this paper, we prove the two-point function results for the contact process via a time discretization. The discretized contact process is oriented percolation in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ with $\varepsilon \in (0,1]$, and the proof is based on the strategy for oriented percolation with $\varepsilon = 1$, i.e., on application of the lace expansion and the inductive method. To obtain the results for small ε, we have to extract the correct powers of ε as in [24] from intersection points in the Feynman-type diagrams appearing in the lace expansion. This leads to a modification of the induction hypotheses of [17]. In order to extend the results from infrared bounds (as in [24]) to precise asymptotics (as in [18]), it is imperative to prove that the properly scaled lace expansion coefficients converge to some continuum limit. We can think of this continuum limit as giving rise to a the lace expansion in continuous time, even though our proof is not based on the arising partial differential equation.

In a sequel to this paper [15], the results obtained in this paper form the key ingredient in the proof that the finite-dimensional distributions of the critical contact process above 4 dimensions converge to those of the canonical measure of super-Brownian motion, as was proved in [18] for oriented percolation.

1.2 The spread-out contact process and main results

We define the spread-out contact process as follows. Let $\mathbf{C}_t \subset \mathbb{Z}^d$ be the set of infected individuals at time $t \in \mathbb{R}_+$, and let $\mathbf{C}_0 = \{o\}$. An infected site x recovers in a small time interval $[t, t+\varepsilon]$ with probability $\varepsilon + o(\varepsilon)$ independently of t, where $o(\varepsilon)$ is a function that satisfies $\lim_{\varepsilon\to 0} o(\varepsilon)/\varepsilon = 0$. In other words, $x \in \mathbf{C}_t$ recovers at rate 1. A healthy site x gets infected, depending on the status of its neighboring sites, at rate $\lambda \sum_{y\in\mathbf{C}_t} D(x-y)$, where $\lambda \geq 0$ is the infection rate. We denote by $\mathbb{P}^\lambda$ the associated probability measure. The function $D : \mathbb{Z}^d \mapsto [0,1]$ is a probability distribution which is symmetric with respect to the lattice symmetries and satisfying certain assumptions. The assumptions on D involve a parameter $L \geq 1$ which serves to spread out the infections, and will be taken to be large. In particular, we require that $D(o) = 0$, $\sup_{x\in\mathbb{Z}^d} D(x) \leq CL^{-d}$. Moreover, with σ defined as

$$\sigma^2 = \sum_x |x|^2 D(x), \tag{1.1}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$, we require that $C_1 L \leq \sigma \leq C_2 L$ and that there exists a $\Delta > 0$ such that

$$\sum_x |x|^{2+2\Delta} D(x) \leq CL^{2+2\Delta}. \tag{1.2}$$

See Section 5.1.1 for the precise assumptions on D. An example of D is

$$D(x) = \frac{1}{(2L+1)^d - 1} \mathbb{1}_{\{0<\|x\|_\infty\leq L\}}, \tag{1.3}$$

which is the uniform distribution on the cube of width L.

The *two-point function* is defined as

$$\tau_t^\lambda(x) = \mathbb{P}^\lambda(x \in \mathbf{C}_t), \qquad x \in \mathbb{Z}^d,\ t \in \mathbb{R}_+. \tag{1.4}$$

In words, $\tau_t^\lambda(x)$ is the probability that at time t, the individual located at $x \in \mathbb{Z}^d$ is infected due to the infection at time 0 located at $o \in \mathbb{Z}^d$. For a summable function $f : \mathbb{Z}^d \mapsto \mathbb{R}$, we define its Fourier transform for $k \in [-\pi,\pi]^d$ by

$$\hat{f}(k) = \sum_{x\in\mathbb{Z}^d} e^{ik\cdot x} f(x), \tag{1.5}$$

By an extension of the results in [8, 4] to the spread-out contact process, there exists a unique critical value $\lambda_{\rm c} \in (0,\infty)$ such that

$$\int_0^\infty dt\, \hat\tau_t^\lambda(0) \begin{cases} <\infty, & \text{if } \lambda<\lambda_{\rm c},\\ =\infty, & \text{otherwise,}\end{cases} \qquad \lim_{t\uparrow\infty} \mathbb{P}^\lambda(\mathbf{C}_t\neq\varnothing) \begin{cases} =0, & \text{if } \lambda\leq\lambda_{\rm c},\\ >0, & \text{otherwise.}\end{cases} \tag{1.6}$$

We will next investigate the sufficiently spread-out contact process at the critical value $\lambda_{\rm c}$ above 4 spatial dimensions.

1.2.1 Results above four dimensions

We now state the results for the two-point function. In the statements, Δ is defined in (1.2) and σ in (1.1).

Theorem 1.1. *Let $d>4$ and $\delta\in(0,1\wedge\Delta\wedge\frac{d-4}{2})$. There exists $L_0=L_0(d)$ such that, for $L\geq L_0$, there exist finite and positive constants v, A (depending on d, L), and C_1, C_2 (depending only on d), such that*

$$\hat\tau_t^{\lambda_{\rm c}}\big(\tfrac{k}{\sqrt{v\sigma^2 t}}\big) = A\, e^{-\frac{|k|^2}{2d}}\, \big[1+O(|k|^2(1+t)^{-\delta})+O((1+t)^{-(d-4)/2})\big], \tag{1.7}$$

$$\frac{1}{\hat\tau_t^{\lambda_{\rm c}}(0)}\sum_x |x|^2\tau_t^{\lambda_{\rm c}}(x) = v\sigma^2 t\,\big[1+O((1+t)^{-\delta})\big], \tag{1.8}$$

$$C_1L^{-d}(1+t)^{-d/2}\leq \sup_x \tau_t^{\lambda_{\rm c}}(x)\leq e^{-t}+C_2L^{-d}(1+t)^{-d/2}, \tag{1.9}$$

with the error estimate in (1.7) *uniform in $k\in\mathbb{R}^d$ with $|k|^2(\log(2+t))^{-1}$ sufficiently small.*

The above results correspond to [18, Theorem 1.1], where the two-point function for sufficiently spread-out critical oriented percolation with $d>4$ was proved to obey the same behavior. The proof in [18] is based on the inductive method of [17]. We apply a modified version of this induction method to prove Theorem 1.1, and obtain

$$\lambda_{\rm c}=1+O(L^{-d}), \qquad A=1+O(L^{-d}), \qquad v=1+O(L^{-d}). \tag{1.10}$$

As discussed in [18, Section 1.2.1], (1.7) and (1.9) imply the triangle condition of [1, 2, 3]. Under that condition, it has been proved [3] that the critical exponents γ and β for the quantities in (1.6) are both equal to 1, independently of the value of d. The fact that these exponents also arise on the tree is given in [3] in combination with [26] and [30]. The result (1.8) also shows that ν, which is defined by

$$\frac{1}{\hat\tau_t^{\lambda_{\rm c}}(0)}\sum_x |x|^2\tau_t^{\lambda_{\rm c}}(x)\sim t^{2\nu}, \tag{1.11}$$

equals $\nu=\frac12$. Here we use $\sim$ to denote any way in which the left- and right-hand side can be equal to each other (see also [20, p. 70]). Moreover, with the critical exponent η defined by

$$\hat\tau_t^{\lambda_{\rm c}}(0)\sim t^{\eta}, \tag{1.12}$$

(1.7) shows that $\eta=0$. This result proves that the statement in [20, Proposition 4.39] on the tree also holds for sufficiently spread-out contact process on $\mathbb{Z}^d$ and $d>4$. See the remark below [20, Proposition 4.39]. Finally, we can extend the proof of $\gamma=1$ to precise asymptotics by showing that there exists a constant C such that

$$\int_0^\infty dt\,\hat\tau_t^\lambda(0)=C(\lambda_{\rm c}-\lambda)^{-1}[1+o(1)]. \tag{1.13}$$

The form of $\gamma=1$ in (1.13) is stronger than the one proved by Barsky and Wu [3], as they only show that the left-hand side multiplied by $\lambda_{\rm c}-\lambda$ is bounded above and below.

In [25], a pair of hyperscaling inequalities for the contact process and oriented percolation critical exponents were derived, which implies that d has to be greater than or equal to 4 in order to have the mean-field values. Therefore, the upper critical dimension for the spread-out contact process is 4.

In [15], we will investigate the higher-point functions of the critical spread-out contact process. These higher-point functions are defined for $\vec{t} \in [0,\infty)^{r-1}$ and $\vec{x} \in \mathbb{Z}^{(r-1)d}$ by

$$\tau_{\vec{t}}^{\lambda}(\vec{x}) = \mathbb{P}^{\lambda}(x_i \in \mathbf{C}_{t_i} \forall i = 1,\dots,r-1). \tag{1.14}$$

The proof will be based on a lace expansion, which expresses the r-point function in terms of s-point functions for all $s < r$. On the arising equation, we will then perform induction in r, with the results for $r = 2$ given by Theorem 1.1. In order to bound the coefficients in the lace expansion for the higher point functions, equations (1.9) and (1.7) for $k = 0$ will be crucial.

In a sequel to this paper [16], we will investigate the critical point in more detail and prove that for $d > 4$, with $G(x) = \sum_{n\geq 0} D^{*n}(x)$ and D^{*n} the n-fold convolution of D, we have

$$\lambda_{\rm c} - 1 = (D * D * G)(o) + O(L^{-2d}). \tag{1.15}$$

Thus, if D is defined by (1.3), then

$$\lambda_{\rm c} - 1 = \frac{1}{(2L+1)^d - 1} \sum_{x:0<\|x\|\leq L} G(x) + O(L^{-2d}) = \frac{1}{(2L)^d} \int_{[-1,1]^d} d^d x \sum_{n\geq 1} U^{*n}(x) + O(L^{-(d+1)}), \tag{1.16}$$

where $U(x)$ is the uniform probability density on $[-1,1]^d$. The above expression was already obtained in [6], but with a weaker error estimate.

1.2.2 Results below and at four dimensions

We also consider the low-dimensional case, i.e., $d \leq 4$. In this case, the contact process is believed *not* to be mean-field, and Gaussian asymptotics are not expected to hold as long as L remains finite. However, following the rescaling of Durrett and Perkins in [6], we can prove Gaussian asymptotics when the range and time grow simultaneously. We suppose that the infection range grows as

$$L_T = L_1 T^b, \tag{1.17}$$

where $L_1 \geq 1$ is the initial infection range. We denote by σ_T^2 the variance of D in this situation. We will assume that

$$\alpha = bd + \frac{d-4}{2} > 0. \tag{1.18}$$

Theorem 1.2. *Let $d \leq 4$ and $\delta \in (0, 1 \wedge \Delta \wedge \alpha)$. There exist $\mu \in (0, \alpha - \delta)$ and $\lambda_T = 1 + O(T^{-\mu})$ such that, for sufficiently large L_1, there exist C_1, C_2 (depending only on d), such that for all $0 < t \leq \log T$, as $T \uparrow \infty$*

$$\hat{\tau}_{Tt}^{\lambda_T}\Big(\frac{k}{\sqrt{\sigma_T^2 Tt}}\Big) = e^{-\frac{|k|^2}{2d}} \big[1 + O(T^{-\mu}) + O(|k|^2 T^{-\delta})\big], \tag{1.19}$$

$$\frac{1}{\hat{\tau}_{Tt}^{\lambda_T}(0)} \sum_x |x|^2 \tau_{Tt}^{\lambda_T}(x) = \sigma_T^2 Tt \big[1 + O(T^{-\mu})\big], \tag{1.20}$$

$$C_1 L_T^{-d} T^{-d/2} \leq \sup_x \tau_{Tt}^{\lambda_T}(x) \leq e^{-tT} + C_2 L_T^{-d} T^{-d/2}, \tag{1.21}$$

with the error estimate in (1.19) uniform in $k \in \mathbb{R}^d$ with $|k|^2(\log(2+Tt))^{-1}$ sufficiently small.

The proof of Theorem 1.2 follows the same steps as the proof of Theorem 1.1.

In [6], a similar situation was investigated, and Durrett and Perkins have proved convergence of the rescaled contact process to super-Brownian motion. We now compare the ranges needed in our results and in [6]. We need that $\alpha = bd + \frac{d-4}{2} > 0$, so that $bd > \frac{4-d}{2}$. In [6], $bd = 1$ for all $d \geq 3$. For $d = 2$, which is a critical case in the setting of [6], Durrett and Perkins investigate the model with range $L_T^2 = T \log T$. In comparison, we are allowed to use ranges that grow to infinity slower than the ranges in [6] when $d \geq 3$, but the range for $d = 2$ in our results needs to be larger than that in [6]. It would be of interest to investigate whether a range $L_T^2 = T \log T$ or even smaller is possible by adapting our proofs.

2 Outline of the proof

In this section, we provide an outline of the proof of our main results. This section is organized as follows. In Section 2.1, we explain what the discretized contact process is, and state the results for the discretized contact process. In Section 2.2, we briefly explain the lace expansion for the discretized contact process. In Section 2.4, we explain how to use induction to prove the asymptotics for the discretized contact process. In Section 2.5, we state the results concerning the continuum limit, and show that the results for the discretized contact process together with the continuum limit imply the main results in Theorems 1.1 and 1.2.

2.1 Discretization

By the graphical representation, the contact process can be constructed as follows. We consider $\mathbb{Z}^d \times \mathbb{R}_+$ as space-time. Along each time line $\{x\} \times \mathbb{R}_+$, we place points according to a Poisson process with intensity 1, independently of the other time lines. For each ordered pair of distinct time lines from $\{x\} \times \mathbb{R}_+$ to $\{y\} \times \mathbb{R}_+$, we place directed bonds $((x,\, t), (y,\, t))$, $t \geq 0$, according to a Poisson process with intensity $\lambda\, D(y-x)$, independently of the other Poisson processes. A site (x, s) is said to be *connected to* (y, t) if either $(x, s) = (y, t)$ or there is a non-zero path in $\mathbb{Z}^d \times \mathbb{R}_+$ from (x, s) to (y, t) using the Poisson bonds and time line segments traversed in the increasing time direction without traversing the Poisson points. The law of $\mathbf{C}_t$ defined in Section 1.2 is equivalent to that of $\{x \in \mathbb{Z}^d : (o,\, 0)$ is connected to $(x,\, t)\}$. See also [20, Section I.1].

Inspired by this percolation structure in space-time and following [24], we consider the following oriented percolation process in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ with $\varepsilon \in (0, 1]$ being a discretization parameter. A directed pair $b = ((x, t), (y, t+\varepsilon))$ of sites in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ is called a *bond*. In particular, b is said to be *temporal* if $x = y$, otherwise *spatial*. Each bond is either *occupied* or *vacant* independently of the other bonds, and a bond $b = ((x, t), (y, t+\varepsilon))$ is occupied with probability

$$p_\varepsilon(y-x) = \begin{cases} 1-\varepsilon, & \text{if } x = y, \\ \lambda\varepsilon\, D(y-x), & \text{otherwise,} \end{cases} \tag{2.1}$$

provided that $\sup_x p_\varepsilon(x) \leq 1$. We denote the associated probability measure by $\mathbb{P}_\varepsilon^\lambda$. It is proved in [4] that $\mathbb{P}_\varepsilon^\lambda$ weakly converges to $\mathbb{P}^\lambda$ as $\varepsilon \downarrow 0$. See Figure 2.1 for a graphical representation of the contact process and the discretized contact process. As explained in more detail in Section 2.2, we prove our main results by proving the results first for the discretized contact process, and then taking the continuum limit.

We denote by $(x,\, s) \rightarrow (y,\, t)$ the event that (x, s) is *connected to* (y, t), i.e., either $(x, s) = (y, t)$ or there is a non-zero path in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ from (x, s) to (y, t) consisting of occupied bonds. The *two-point function* is defined as

$$\tau_{t;\varepsilon}^\lambda(x) = \mathbb{P}_\varepsilon^\lambda((o, 0) \rightarrow (x, t)). \tag{2.2}$$

Similarly to (1.6), the discretized contact process has a critical value $\lambda_{\mathrm{c}}^{(\varepsilon)}$ satisfying

$$\varepsilon \sum_{t \in \varepsilon\mathbb{Z}_+} \hat{\tau}_{t;\varepsilon}^\lambda(0) \begin{cases} < \infty, & \text{if } \lambda < \lambda_{\mathrm{c}}^{(\varepsilon)}, \\ = \infty, & \text{otherwise,} \end{cases} \qquad \lim_{t \uparrow \infty} \mathbb{P}_\varepsilon^\lambda(\mathbf{C}_t \neq \varnothing) \begin{cases} = 0, & \text{if } \lambda \leq \lambda_{\mathrm{c}}^{(\varepsilon)}, \\ > 0, & \text{otherwise.} \end{cases} \tag{2.3}$$

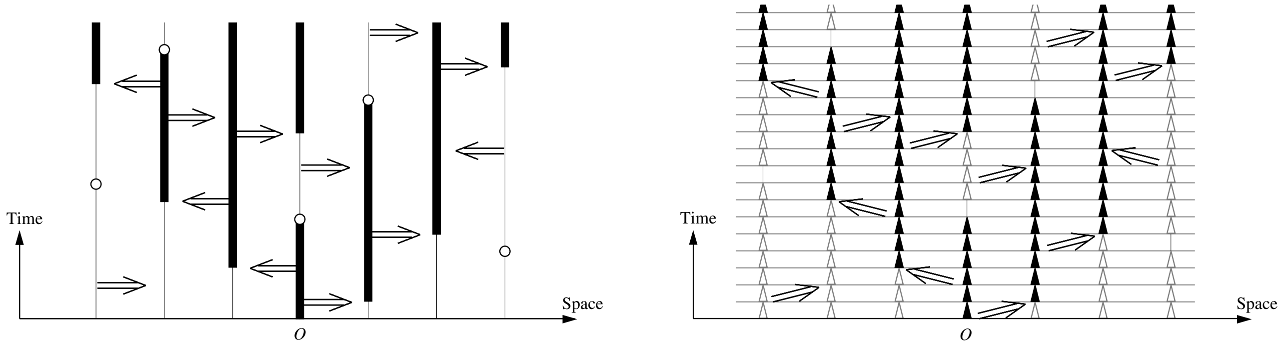

Figure 1: Graphical representation of the contact process and the discretized contact process.

The main result for the discretized contact process is the following theorem:

Proposition 2.1 (Discretized results). *Let $d > 4$ and $\delta \in (0, 1 \wedge \Delta \wedge \frac{d-4}{2})$. There exists $L_0 = L_0(d)$ such that, for $L \geq L_0$, there exist finite and positive constants $v^{(\varepsilon)}$, $A^{(\varepsilon)}$ (depending on d, L), and C_1, C_2 (depending only on d), such that*

$$\hat{\tau}_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}\Big(\frac{k}{\sqrt{v^{(\varepsilon)}\sigma^2 t}}\Big) = A^{(\varepsilon)} e^{-\frac{|k|^2}{2d}} \big[1 + O(|k|^2(1+t)^{-\delta}) + O((1+t)^{-(d-4)/2})\big], \tag{2.4}$$

$$\frac{1}{\hat{\tau}_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(0)} \sum_x |x|^2 \tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x) = v^{(\varepsilon)}\sigma^2 t \big[1 + O((1+t)^{-\delta})\big], \tag{2.5}$$

$$C_1 L^{-d}(1+t)^{-d/2} \leq \sup_x \tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x) \leq e^{-t} + C_2 L^{-d}(1+t)^{-d/2}, \tag{2.6}$$

where all error terms are uniform in $\varepsilon \in (0, 1]$. The error estimate in (1.7) *is uniform in $k \in \mathbb{R}^d$ with $|k|^2(\log(2+t))^{-1}$ sufficiently small.*

Proposition 2.1 is the discrete analog of Theorem 1.1. The uniformity of the error terms is crucial, as this will allow us to take the limit $\varepsilon \downarrow 0$ and to conclude the results in Theorem 1.1 from the corresponding statements in Proposition 2.1. There is also a discretized version of Theorem 1.2, but we refrain from stating it. Note that Proposition 2.1 applies also to oriented percolation, for which $\varepsilon = 1$.

2.2 The expansion

The proof of Proposition 2.1 makes use of the lace expansion, which is an expansion for the two-point function. We postpone the derivation of the expansion to Section 3, and here we provide only brief motivation. We will also motivate why we discretize time for the contact process.

We will make use of the convolution of functions, which is defined for absolutely summable functions f, g on $\mathbb{Z}^d$ by

$$(f * g)(x) = \sum_y f(y)g(x-y). \tag{2.7}$$

We first motivate the basic idea underlying the expansion, similarly as in [18, Section 2.1.1], by considering the much simpler corresponding expansion for continuous-time random walk. For continuous-time random walk making jumps from x to y at rate $\lambda D(y-x)$, and being killed at rate $1-\lambda$, we have the partial differential equation

$$\partial_t q_t^\lambda(x) = \lambda(D * q_t^\lambda)(x) - q_t^\lambda(x), \tag{2.8}$$

where $q_t^\lambda(x)$ is the probability that continuous-time random walk with step distribution D started at $o \in \mathbb{Z}^d$ is at time t in $x \in \mathbb{Z}^d$. By taking the Fourier transform, we obtain

$$\partial_t \hat{q}_t^\lambda(k) = -[1 - \lambda \hat{D}(k)]\hat{q}_t^\lambda(k). \tag{2.9}$$

In this simple case, the above equation is readily solved to yield that

$$\hat{q}_t^\lambda(k) = e^{-[1-\lambda\hat{D}(k)]t}. \tag{2.10}$$

We see that $\lambda = 1$ is the critical value, and the central limit theorem at $\lambda = \lambda_{\mathrm{c}} = 1$ follows by a Taylor expansion of $1 - \hat{D}(k)$ for small k, yielding

$$\hat{q}_t^1\big(\tfrac{k}{\sqrt{\sigma^2 t}}\big) = e^{-\frac{|k|^2}{2d}}[1 + o(1)], \tag{2.11}$$

where $|k|^2 = \sum_{j=1}^d k_i^2$ (recall also (1.1)).

The above solution is quite specific to continuous-time random walk. When we would have a more difficult function on the right-hand side of (2.9), such as $-[1 - \lambda\hat{D}(k)]\hat{q}_{t-1}^\lambda(k)$, it would be much more difficult to solve the above equation, even though one would expect that the central limit theorem at the critical value still holds.

A more robust proof of central limit behaviour uses induction in time t. Since time is continuous, we first discretize time. The two-point function for discretized continuous-time random walk is defined by setting $q_{0;\varepsilon}^\lambda(x) = \delta_{0,x}$ and

$$q_{t;\varepsilon}^\lambda(x) = p_\varepsilon^{*t/\varepsilon}(x) \qquad (t \in \varepsilon\mathbb{N}). \tag{2.12}$$

To obtain a recursion relation for $q_{t;\varepsilon}^\lambda(x)$, we simply observe that by independence of the underlying random walk

$$q_{t;\varepsilon}^\lambda(x) = (p_\varepsilon * q_{t-\varepsilon;\varepsilon}^\lambda)(x) \qquad (t \in \varepsilon\mathbb{N}). \tag{2.13}$$

We can think of this as a simple version of the lace expansion, applied to random walk, which has no interaction.

For the discretized continuous-time random walk, we can use induction in n for all $t = n\varepsilon$. If we can further show that the arising error terms are uniform in ε, then we can take the continuum limit $\varepsilon \downarrow 0$ afterwards, and obtain the result for the continuous-time model. The above proof is more robust, and can for instance be used to deal with the situation where the right-hand side of (2.9) equals $-[1-\lambda\hat{D}(k)]\hat{q}_{t-1}^\lambda(k)$. This robustness of the proof is quite valuable when we wish to apply it to the contact process.

The identity (2.13) can be solved using the Fourier transform to give

$$\hat{q}_{t;\varepsilon}^\lambda(k) = \hat{p}_\varepsilon(k)^{t/\varepsilon} = [1 - \varepsilon + \lambda\varepsilon\hat{D}(k)]^{t/\varepsilon} = e^{-t[1-\lambda\hat{D}(k)]+O(t\varepsilon[1-\hat{D}(k)]^2)}.$$

In order to obtain the central limit theorem, we divide k by $\sqrt{\sigma^2 t}$. Then, uniformly in $\varepsilon > 0$, we have that

$$\hat{q}_{t;\varepsilon}^1\big(\tfrac{k}{\sqrt{\sigma^2 t}}\big) = e^{-\frac{|k|^2}{2d}+O(k^{2+2\Delta}t^{-\Delta}+k^4\varepsilon t^{-1})}. \tag{2.14}$$

Therefore, the central limit theorem holds uniformly in $\varepsilon > 0$. We stress here that we must have that $\hat{q}_{t;\varepsilon}(k) - \hat{q}_{t-\varepsilon;\varepsilon}(k)$ is of order ε in order for the limit over ε to exist.

We follow Mark Kac's adagium: “Be wise, discretize!” for the contact process for two reasons. Discretizing time allows us to obtain an expansion as in (2.8), and to analyse the arising equation. The lace expansion, which is explained in more detail below, can be used for the contact process to produce an equation of the form

$$\partial_t \tau_t^\lambda(x) = \lambda(D * \tau_t^\lambda)(x) - \tau_t^\lambda(x) + \int_0^t (\pi_s^\lambda * \tau_{t-s}^\lambda)(x)ds, \tag{2.15}$$

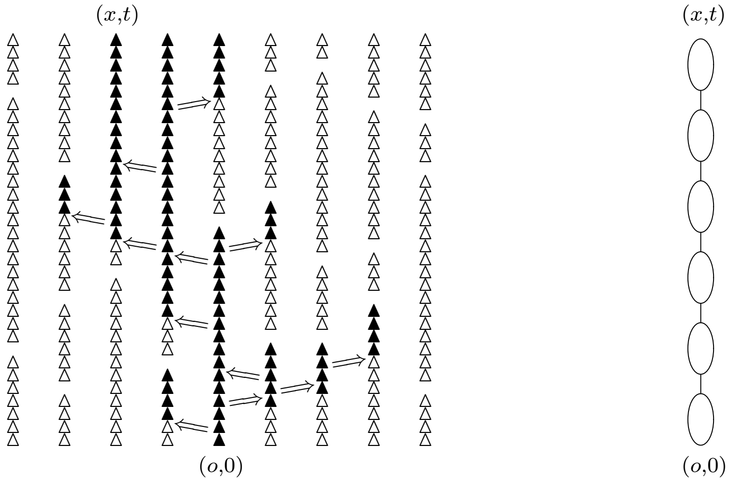

Figure 2: (a) A configuration for the discretized contact process. Open triangles △ denote temporal bonds, closed triangles ▲ denote temporal bonds that are infected, and the arrows denote spread of the infection to neighbours. (b) Schematic depiction of the configuration connecting $(o, 0)$ and (x, t) as a "string of sausages."

where π_s^λ are certain expansion coefficients. In order to derive the equation (2.15), we use that the discretized contact process is oriented percolation, and lace expansions have been derived in the literature for oriented percolation [18, 21, 22]. Clearly, the equation (2.15) is much more complicated than the corresponding equation for simple random walk (2.8). Therefore, a simple solution to the equation as in (2.10) is impossible. We see no way to analyse the partial differential equation in (2.15) other than to discretize time combined with induction. It would be of interest to investigate whether (2.15) can be used directly.

We next explain the expansion for the discretized contact process in more detail, following the explanation in [18, Section 2.1.1]. For the discretized contact process, we will regard the part of the oriented percolation cluster connecting $(o, 0)$ to (x, t) as a "string of sausages." An example of such a cluster is shown in Figure 2. The difference between oriented percolation and random walk resides in the fact that for oriented percolation, there can be multiple paths of occupied bonds connecting $(o, 0)$ to (x, t). However, for $d > 4$, each of those paths passes through the same *pivotal bonds*, which are the essential bonds for the connection from $(o, 0)$ to (x, t) (see also Definition 3.1). We expect that there are of the order t/ε such pivotal bonds. For instance, the first black triangle indicates that $(o, 0)$ is connected to (o, ε), and this bond is pivotal for the connection from $(o, 0)$ to (x, t).

Using this picture, we can think of the oriented percolation two-point function as a kind of random walk two-point function with a distribution describing the statistics of the sausages, taking steps in both space and time. Due to the nature of the pivotal bonds, there can be no connection from one sausage to a later sausage other than the connection via the pivotal bonds between these sausages. Therefore, there is a kind of repulsive interaction between the sausages. The main part of our proof shows that this interaction is weak for $d > 4$.

Fix $\lambda \geq 0$. As we will prove in Section 3 below, the generalisation of (2.13) to the discretized contact process takes the form

$$\tau_{t;\varepsilon}^\lambda(x) = \sum_{s=0}^{t-\varepsilon}{}^{\bullet} (\pi_{s;\varepsilon}^\lambda * p_\varepsilon * \tau_{t-s-\varepsilon;\varepsilon}^\lambda)(x) + \pi_{t;\varepsilon}^\lambda(x) \qquad (t \in \varepsilon\mathbb{N}), \tag{2.16}$$

where we use the notation $\sum^\bullet$ to denote sums over $\varepsilon\mathbb{Z}_+$ and the coefficients $\pi_{t;\varepsilon}^\lambda(x)$ will be defined in Section 3. In particular, $\pi_{t;\varepsilon}^\lambda(x)$ depends on λ, is invariant under the lattice symmetries, and $\pi_{0;\varepsilon}^\lambda(x) = \delta_{o,x}$,

$\pi^{\lambda}_{\varepsilon;\varepsilon}(x)=0$. Note that for $t=0,\varepsilon$, we have $\tau^{\lambda}_{0;\varepsilon}(x)=\delta_{0,x}$ and $\tau^{\lambda}_{\varepsilon;\varepsilon}(x)=p_{\varepsilon}(x)$, which is consistent with (2.16).

Together with the initial values $\pi^{\lambda}_{0;\varepsilon}(x)=\delta_{o,x}$, $\pi^{\lambda}_{\varepsilon;\varepsilon}(x)=0$, the identity (2.16) gives an inductive definition of the sequence $\pi^{\lambda}_{t;\varepsilon}(x)$ for $t\geq 2\varepsilon$ with $t\in\varepsilon\mathbb{Z}_{+}$. However, to analyse the recursion relation (2.16) it will be crucial to have a useful representation for $\pi^{\lambda}_{t;\varepsilon}(x)$, and this is provided in Section 3. Note that (2.13) is of the form (2.16) with $\pi^{\lambda}_{t;\varepsilon}(x)=\delta_{o,x}\delta_{0,t}$, so that we can think of the coefficients $\pi^{\lambda}_{t;\varepsilon}(x)$ for $t\geq 2\varepsilon$ as quantifying the repulsive interaction between the sausages in the "string of sausages" picture.

Our proof will be based on showing that $\frac{1}{\varepsilon^2}\pi^{\lambda}_{t;\varepsilon}(x)$ is small for $\lambda=\lambda_{\rm c}^{(\varepsilon)}$ if $d>4$ and t and L are both large, uniformly in $\varepsilon>0$. In fact, we prove that $\frac{1}{\varepsilon^2}\pi^{\lambda}_{t;\varepsilon}(x)$ converges to a certain $\pi^{\lambda}_{t}(x)$, which is the coefficient of the continuous-time lace expansion in (2.15). The reason that we need to divide by ε^2 is that the sum over s in (2.16) times ε is a Riemann sum approximation to an integral. Thus, the sum over s eats up one factor of ε. The second factor is needed since, similarly to the proof for time-discretized continuous-time random walk in (2.14), we need that $\hat{\tau}^{\lambda}_{t;\varepsilon}(k)-\hat{\tau}^{\lambda}_{t-\varepsilon;\varepsilon}(k)$ is of order ε. Therefore, for $d>4$, (2.16) can be regarded as a small perturbation of (2.13), and this will imply the central limit theorem for the critical two-point function.

In Section 3, we will express $\pi^{\lambda}_{t;\varepsilon}(x)$ as

$$\pi^{\lambda}_{t;\varepsilon}(x)=\sum_{N=0}^{\infty}(-1)^{N}\pi^{(N)}_{t;\varepsilon}(x), \tag{2.17}$$

where we suppress the dependence of $\pi^{(N)}_{t;\varepsilon}(x)$ on λ. We now briefly explain the idea behind the proof of (2.17). Let

$$\pi^{(0)}_{t;\varepsilon}(x)=\mathbb{P}^{\lambda}_{\varepsilon}((o,0)\rightrightarrows(x,t)) \tag{2.18}$$

denote the contribution to $\tau^{\lambda}_{t;\varepsilon}(x)$ from configurations where there is no pivotal bond, so that

$$\tau^{\lambda}_{t;\varepsilon}(x)=\pi^{(0)}_{t;\varepsilon}(x)+\sum_{b}\mathbb{P}^{\lambda}_{\varepsilon}(b\text{ first occupied and pivotal bond for }(o,0)\to(x,t)), \tag{2.19}$$

where the sum over b is over bonds of the form $b=((u,s),(v,s+\varepsilon))$. Write $\overline{b}=(v,s+\varepsilon)$ for the endpoint of the bond b and $\underline{b}=(u,s)$ for its starting point. Then we have that the probability on the right-hand side of (2.19) equals

$$\mathbb{P}^{\lambda}_{\varepsilon}\big((o,0)\rightrightarrows\underline{b},b\text{ occupied},\overline{b}\to(x,t),b\text{ pivotal for }(o,0)\to(x,t)\big). \tag{2.20}$$

We ignore the intersection with the event that b is pivotal for $(o,0)\to(x,t)$, and obtain using the Markov property that

$$\tau^{\lambda}_{t;\varepsilon}(x)=\pi^{(0)}_{t;\varepsilon}(x)+\sum_{s=0}^{t-\varepsilon}{}^{\bullet}\sum_{u,v}\pi^{(0)}_{s;\varepsilon}(u)p_{\varepsilon}(v-u)\tau_{t-s-\varepsilon;\varepsilon}(x-v)-R^{(0)}_{t;\varepsilon}(x), \tag{2.21}$$

where

$$R^{(0)}_{t;\varepsilon}(x)=\sum_{b}\mathbb{P}^{\lambda}_{\varepsilon}\big((o,0)\rightrightarrows\underline{b},b\text{ occupied},\overline{b}\to(x,t),b\text{ not pivotal for }(o,0)\to(x,t)\big). \tag{2.22}$$

We will investigate the error term $R^{(0)}_{t;\varepsilon}(x)$ further, again using inclusion-exclusion, by investigating the first pivotal bond after $\overline{b}$ to arrive at (2.17). The term $\pi^{(1)}_{t;\varepsilon}(x)$ is the contribution to $R^{(0)}_{t;\varepsilon}(x)$ where such a pivotal does not exist. Thus, in $\pi^{(0)}_{t;\varepsilon}(x)$ for $t\geq\varepsilon$ and in $\pi^{(1)}_{t;\varepsilon}(x)$ for all $t\geq 0$, there is at least one loop, which, for L large, should yield a small correction only. In (2.17), the contributions from $N\geq 2$ have at least two loops and are thus again smaller, even though all $N\geq 0$ give essential contributions to $\pi^{\lambda}_{t;\varepsilon}(x)$ in (2.17).

There are three ways to obtain the lace expansion in (2.16) for oriented percolation models. We use the expansion by Sakai [23, 24], as described in (2.18)–(2.22) above, based on inclusion-exclusion together with the Markov property for oriented percolation. For unoriented percolation, Hara and Slade [9] developed an expression for $\pi^{\lambda}_{t;\varepsilon}(x)$ in terms of sums of nested expectations, by repeated use of inclusion-exclusion and using the independence of percolation. This expansion, and its generalizations to the higher-point functions, was used in [18] to investigate the oriented percolation r-point functions. The original expansion in [9] was for unoriented percolation, and does not make use of the Markov property. Nguyen and Yang [21, 22] derived an alternate expression for $\pi^{(N)}_{t;\varepsilon}(x)$ by adapting the lace expansion of Brydges and Spencer [5] for weakly-self avoiding walk. In the graphical representation of the Brydges-Spencer expansion, laces arise which give the "lace expansion" its name. Even though in many of the lace expansions for percolation type models, such as oriented and unoriented percolation, no laces appear, the name has stuck for historical reasons.

We believe that the Nguyen-Yang expansion is equivalent to the above expansion using inclusion-exclusion, just as for self-avoiding walks, the lace expansion can be derived using inclusion-exclusion [29]. The formulas for $\pi^{(N)}_{t;\varepsilon}(x)$ in the inclusion-exclusion expansion are somewhat simpler than the ones in the Nguyen-Yang expansion. The Nguyen-Yang expansion was also used in [18, Section 3.2] to deal with the derivative of the lace expansion coefficients with respect to the percolation parameter p. The Hara-Slade expansion, on the other hand, yields (2.17) with a *different* expression for the $\pi^{(N)}_{t;\varepsilon}(x)$, even though, in either expansion, $\pi^{(N)}_{t;\varepsilon}(x)$ is nonnegative for all t, x, N, and can be represented in terms of Feynman-type diagrams. The Feynman diagrams are similar for the three expansions, and obey similar estimates, even though the expansion using inclusion-exclusion produces the simplest diagrams. We will use the inclusion-exclusion expansion in [23, 24], as well as an adaptation to deal with derivatives with respect to λ, since this expansion most easily deals with the discretized contact process, in particular with the uniformity in the small parameter ε.

In [15], we will adapt the expansion in Section 3 to deal with the discretized contact process and oriented percolation higher-point functions. For this, we will need ingredients from the Hara-Slade expansion to compare occupied paths living on a common time interval, with independent paths. This independence does not follow from the Markov property, and therefore the Hara-Slade expansion, which does not require the Markov property, will be crucial.

2.3 Bounds on the lace expansion

In what follows, we will use the notation $\|f\|_\infty = \sup_{x\in\mathbb{Z}^d} |f(x)|$ for a function $f : \mathbb{Z}^d \mapsto \mathbb{C}$, and $\|\hat{f}\|_1 = (2\pi)^{-d}\int_{[-\pi,\pi]^d} |\hat{f}(k)| d^dk$ for a function $\hat{f} : [-\pi,\pi]^d \mapsto \mathbb{C}$.

In order to prove the statements in Proposition 2.1, we will use induction in n, where $t = n\varepsilon$. The lace expansion equation in (2.16) forms the main ingredient for this induction in time. We will explain the inductive method in more detail below. It is clear that to advance the induction, we need to have certain bounds on the lace expansion coefficients. The form of those bounds will be explained now. The statement of the bounds involve the small parameter

$$\beta = L^{-d}. \tag{2.23}$$

We will use the following set of bounds:

$$|\hat{\tau}_{s;\varepsilon}(0)| \leq K, \qquad |\nabla^2 \hat{\tau}_{s;\varepsilon}(0)| \leq K\sigma^2 s, \qquad \|\hat{D}^2\hat{\tau}_{s;\varepsilon}\|_1 \leq \frac{K\beta}{(1+s)^{d/2}}. \tag{2.24}$$

The bounds on the lace expansion consist of the following estimates, which will be proved in Section 4.

Proposition 2.2 (Bounds on the lace expansion above four dimensions). *Assume* (2.24) *for some λ_0 and all $s \leq t$. Then there exist $\beta_0 > 0$ and $C < \infty$ (both depending on d and K, but not on ε, L*

and Δ'*) such that for* $\lambda \leq \lambda_0$*,* $\beta < \beta_0$*,* $s \in \varepsilon\mathbb{Z}_+$ *with* $2\varepsilon \leq s \leq t+\varepsilon$*,* $q = 0, 2, 4$ *and* $\Delta' \in [0, 1 \wedge \Delta]$*, and uniformly in* $\varepsilon \in (0, 1]$*,*

$$\sum_x |x|^q \, |\pi^{\lambda}_{s;\varepsilon}(x)| \leq \frac{\varepsilon^2 C \sigma^q \beta}{(1+s)^{(d-q)/2}}, \tag{2.25}$$

$$\Big|\hat{\pi}^{\lambda}_{s;\varepsilon}(k) - \hat{\pi}^{\lambda}_{s;\varepsilon}(0) - \frac{a(k)}{\sigma^2}\, \nabla^2 \hat{\pi}^{\lambda}_{s;\varepsilon}(0)\Big| \leq \frac{\varepsilon^2 C \beta \, a(k)^{1+\Delta'}}{(1+s)^{(d-2)/2-\Delta'}}, \tag{2.26}$$

$$|\partial_\lambda \hat{\pi}^{\lambda}_{s;\varepsilon}(0)| \leq \frac{\varepsilon^2 C \beta}{(1+s)^{(d-2)/2}}. \tag{2.27}$$

The main content of Proposition 2.2 is that the bounds on $\hat{\tau}_{s;\varepsilon}$ for $s \leq t$ in (2.24) imply bounds on $\hat{\pi}_{s;\varepsilon}$ for all $s \leq t+\varepsilon$. This fact allows us to use the bounds on $\hat{\pi}_{s;\varepsilon}$ for all arising s in (2.16) in order to advance the appropriate induction hypotheses. Of course, in order to complete the inductive argument, we need that the induction statements imply the bounds in (2.24).

The proof of Proposition 2.2 is deferred to Section 4. Proposition 2.2 is probably false in dimensions $d \leq 4$. However, when the range increases with T as in Theorem 1.2, we are still able to obtain the necessary bounds. In the statement of the bounds, we recall that L_T is given in (1.17).

Proposition 2.3 (Bounds on the lace expansion below and at four dimensions). *Let* $\alpha > 0$ *in* (1.18)*. Assume* (2.24) *for some* λ_0 *and all* $s \leq t$*. Then there exist* $L_0 < \infty$ *(depending on* d *and* K*, but not on* ε*, and* Δ'*) and* $C < \infty$ *(depending on* d *and* K*, but not on* ε*,* L *and* Δ'*) such that for* $\lambda \leq \lambda_0$*,* $L_1 \geq L_0$*,* $s \in \varepsilon\mathbb{Z}_+$ *with* $2\varepsilon \leq s \leq t+\varepsilon$*,* $q = 0, 2, 4$ *and* $\Delta' \in [0, 1 \wedge \Delta]$*, the bounds in* (2.25)*–*(2.27) *hold for* $t \leq T \log T$ *with* β *replaced by* $\beta_T = L_T^{-d}$ *and* σ^2 *by* σ_T^2*.*

The main point in Propositions 2.2-2.3 is the fact that we need to extract two factors of ε. One can see that such factors must be present by investigating the case $N = 0$, when $\pi^{(0)}_{t;\varepsilon}(x)$ is the probability that $(o, 0)$ is doubly connected to (x, t). Thus, there must be at least two spatial bonds, one emanating from $(o, 0)$ and one pointing into (x, t). By (2.1), these two spatial bonds give rise to two powers of ε. The proof for $N \geq 1$ follows then by induction in N.

2.4 Implementation of the inductive method

Our analysis of (2.16) begins by taking its Fourier transform, which gives the recursion relation

$$\hat{\tau}^{\lambda}_{t;\varepsilon}(k) = \sum_{s=0}^{t-\varepsilon}{}^{\bullet}\, \hat{\pi}^{\lambda}_{s;\varepsilon}(k) \hat{p}_{\varepsilon}(k) \hat{\tau}^{\lambda}_{t-s-\varepsilon;\varepsilon}(k) + \hat{\pi}^{\lambda}_{t;\varepsilon}(k) \qquad (t \in \varepsilon\mathbb{N}). \tag{2.28}$$

The right-hand side of (2.28) explicitly involves $\hat{\tau}^{\lambda}_{s;\varepsilon}(k)$ only for $s \leq t-\varepsilon$. As already explained in Section 2.3, it is possible to estimate $\hat{\pi}^{\lambda}_{s;\varepsilon}(k)$, for all $s \leq t$, in terms of $\hat{\tau}^{\lambda}_{s;\varepsilon}(0)$ and $\sup_x \tau^{\lambda}_{s;\varepsilon}(x)$ with $s \leq t-\varepsilon$. This opens up the possibility of an inductive analysis of (2.28). A general approach to this type of inductive analysis is given in [17]. However, here we will need the uniformity in the variable ε, and therefore we will state a version of the induction in Section 5 that is adapted to the uniformity in ε and thus the continuum limit. The advancement of the induction hypotheses is deferred to Appendix A.

Moreover, we will show that the critical point is given implicitly by the equation

$$\lambda_{\mathrm{c}}^{(\varepsilon)} = 1 - \frac{1}{\varepsilon} \sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\, \hat{p}_{\varepsilon}^{\lambda_{\mathrm{c}}^{(\varepsilon)}}(0) \hat{\pi}^{\lambda_{\mathrm{c}}^{(\varepsilon)}}_{s;\varepsilon}(0), \tag{2.29}$$

and that the constants $A^{(\varepsilon)}$, $v^{(\varepsilon)}$ of Proposition 2.1 are given by

$$A^{(\varepsilon)}=\frac{1+\sum\limits_{s=2\varepsilon}^{\infty}{}^{\bullet}\,\hat{\pi}_{s;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(0)}{1+\frac{1}{\varepsilon}\sum\limits_{s=2\varepsilon}^{\infty}{}^{\bullet}\,s\hat{p}_{\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(0)\hat{\pi}_{s;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(0)},\qquad v^{(\varepsilon)}=\frac{\lambda_{\rm c}^{(\varepsilon)}-\frac{1}{\varepsilon\sigma^2}\sum\limits_{s=2\varepsilon}^{\infty}{}^{\bullet}\,\nabla^2[\hat{p}_{\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}\hat{\pi}_{s;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}](0)}{1+\frac{1}{\varepsilon}\sum\limits_{s=2\varepsilon}^{\infty}{}^{\bullet}\,s\hat{p}_{\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(0)\hat{\pi}_{s;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(0)},\tag{2.30}$$

where we have added an argument $\lambda_{\rm c}^{(\varepsilon)}$ to emphasize that λ is critical for the evaluation of $\pi_{t;\varepsilon}^{\lambda}$ on the right sides. Convergence of the series on the right sides, for $d>4$, follows from Proposition 2.2.

The result of induction is summarized in the following proposition:

Proposition 2.4 (Induction). *Assume that Proposition 2.2 holds. Then there exist $\beta_0>0$ and $C<\infty$ (both depending on d and K, but not on ε and Δ') such that for $\beta<\beta_0$, $q=0,2,4$ and $\Delta'\in[0,1\wedge\Delta]$, the statements in Proposition 2.1 follow, where the error terms are uniform in $\varepsilon\in(0,1]$.*

There is also a low-dimensional version of Proposition 2.4, but we refrain from stating it.

2.5 The continuum limit

In this section we state the result necessary to complete the proof of Theorem 1.1 from Proposition 2.1.

Proposition 2.5 (Continuum limit). *Suppose that $\lambda^{(\varepsilon)}\to\lambda$ and $\lambda^{(\varepsilon)}\le\lambda_{\rm c}^{(\varepsilon)}$ for ε sufficiently small. Then, for every $t>0$ and $x\in\mathbb{Z}^d$, there exists $\pi_t(x)$ such that*

$$\lim_{\varepsilon\downarrow0}\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x)=\pi_t^{\lambda}(x),\qquad\lim_{\varepsilon\downarrow0}\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^{\lambda}(x)=\partial_\lambda\pi_t^{\lambda}(x).\tag{2.31}$$

Consequently, for $q\in\{0,2,4\}$,

$$\sum_{x\in\mathbb{Z}^d}|x|^q\pi_t^{\lambda}(x)\le\frac{C\beta}{(1+t)^{(d-q)/2}},\qquad\sum_{x\in\mathbb{Z}^d}\partial_\lambda\pi_t^{\lambda}(x)\le\frac{C\beta}{(1+t)^{(d-2)/2}},\tag{2.32}$$

and there exist finite and positive constants $A=1+O(L^{-d})$ and $v=1+O(L^{-d})$ such that

$$\lim_{\varepsilon\downarrow0}A^{(\varepsilon)}=A,\qquad\lim_{\varepsilon\downarrow0}v^{(\varepsilon)}=v.\tag{2.33}$$

In the proof of Proposition 2.5, we will make essential use of the fact that $\mathbb{P}_\varepsilon^{\lambda}$ weakly converges to $\mathbb{P}^{\lambda}$ as $\varepsilon\downarrow0$ as proved in [4, Proposition 2.7].

From Proposition 2.2, we see that the two powers of ε are present in $\pi_{t;\varepsilon}^{\lambda}(x)$. The main point in the proof of Proposition 2.5 is that the properly rescaled lace expansion coefficients converge.

In Section 6 below we will show that $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^{\lambda}(x)$ and $\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^{\lambda}(x)$ both converge pointwise. We now show that this implies that the limit of $\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^{\lambda}(x)$ equals $\partial_\lambda\pi_t^{\lambda}(x)$. To see this, we note that since $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^{0}(x)=0$ for $t>0$, we have

$$\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^{\lambda}(x)=\int_0^{\lambda}\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^{\lambda'}(x)d\lambda'.\tag{2.34}$$

The integral on the right-hand side converges to the integral of the limit $f_t^{\lambda'}(x)$ of $\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^{\lambda}(x)$, by the assumed pointwise convergence, and dominated convergence. Also, the left-hand side converges by assumption to $\pi_t^{\lambda}(x)$. Thus, we obtain that

$$\pi_t^{\lambda}(x)=\int_0^{\lambda}f_t^{\lambda'}(x)d\lambda',\tag{2.35}$$

which indeed implies that $f_t^\lambda(x) = \partial_\lambda \pi_t^\lambda(x)$.

We next prove the main results in Theorems 1.1- 1.2.

Conclusion of the proof of Theorems 1.1 and 1.2 assuming Propositions 2.1–2.5.
We will only prove Theorem 1.1, since the proof of Theorem 1.2 is identical. By [4, Proposition 2.7], we have that for every (x,t) and $\lambda > 0$,

$$\lim_{\varepsilon\downarrow 0} \tau_{t;\varepsilon}^\lambda(x) = \tau_t^\lambda(x). \tag{2.36}$$

Since $\tau_{t;\varepsilon}^\lambda(x)$ is monotonically increasing in λ, and since $\lambda \mapsto \tau_t^\lambda(x)$ is continuous (see e.g., [20, pp. 38-39]), we also obtain that for any $\lambda^{(\varepsilon)} \to \lambda$, we have that $\lim_{\varepsilon\downarrow 0} \tau_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x) = \tau_t^\lambda(x)$.

Sakai [24, Section 2.1] proved that $\lambda_{\rm c}^{(\varepsilon)} \to \lambda_{\rm c}$, so that by the above argument $\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(x)$ converges to $\tau_t^{\lambda_{\rm c}}(x)$. Using this, we first prove the upper bound in (1.9). We have that for every x

$$\tau_t^{\lambda_{\rm c}}(x) = \lim_{\varepsilon\downarrow 0} \tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(x) \leq \lim_{\varepsilon\downarrow 0} \left[e^{-t} + C_2 L^{-d}(1+t)^{-d/2}\right] = e^{-t} + C_2 L^{-d}(1+t)^{-d/2}, \tag{2.37}$$

where we use that the bound is uniform in $\varepsilon \in (0,1]$. The lower bound in (1.9) follows similarly.

We next prove that for any fixed $t \geq 0$ and $k \in [-\pi,\pi]^d$, we have that $\hat\tau_t^{\lambda_{\rm c}}(k) = \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(k)$. Note that by definition,

$$\hat\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(k) = \sum_{x\in\mathbb{Z}^d} e^{ik\cdot x} \tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(x). \tag{2.38}$$

Thus, using the convergence of $\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(x)$, we must to show that the sum over $x \in \mathbb{Z}^d$ in (2.38) can be approximated by a finite sum. To see this, we note that

$$\tau_{t;\varepsilon}^\lambda(x) \leq p_\varepsilon^{*t/\varepsilon}(x), \tag{2.39}$$

so that

$$\tau_{t;\varepsilon}^\lambda(x) \leq \sum_{n=0}^{t/\varepsilon} \binom{t/\varepsilon}{n} (1-\varepsilon)^{t/\varepsilon - n} (\lambda\varepsilon)^n D^{*n}(x). \tag{2.40}$$

Clearly, for any fixed t and R large enough, we have that

$$\sum_{\|x\|_\infty > R} \tau_{t;\varepsilon}^\lambda(x) \leq \delta_R, \tag{2.41}$$

where $\delta_R \downarrow 0$ as $R \to \infty$ uniformly in $\varepsilon > 0$. Therefore, the same holds for the contact process two-point function $\tau_t^\lambda(x)$, and thus we can approximate both $\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(k)$ and $\hat\tau_t^{\lambda_{\rm c}}(k)$ by sums over x with $\|x\|_\infty \leq R$. This proves that $\hat\tau_t^{\lambda_{\rm c}}(k) = \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}(k)$.

Using the above, we obtain that

$$\begin{aligned}
\hat\tau_t^{\lambda_{\rm c}}\big(\tfrac{k}{\sqrt{v\sigma^2 t}}\big) &= \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}\big(\tfrac{k}{\sqrt{v\sigma^2 t}}\big) = \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_{\rm c}^{(\varepsilon)}}\big(\tfrac{k}{\sqrt{v^{(\varepsilon)}\sigma^2 t}} \tfrac{\sqrt{v^{(\varepsilon)}}}{\sqrt{v}}\big) \\
&= \lim_{\varepsilon\downarrow 0} A^{(\varepsilon)}\, e^{-\frac{|k|^2}{2d}\frac{v^{(\varepsilon)}}{v}} \left[1 + O(|k|^2(1+t)^{-\delta}) + O((1+t)^{-(d-4)/2})\right] \\
&= A\, e^{-\frac{|k|^2}{2d}} \left[1 + O(|k|^2(1+t)^{-\delta}) + O((1+t)^{-(d-4)/2})\right],
\end{aligned} \tag{2.42}$$

which proves (1.7). Similar arguments can be used for (1.8). □

We complete the proof by proving (2.15) and (1.13). For (2.15), we note that all terms in the sum are of order ε^2, except for the contribution from $s = 0$, which equals $(p_\varepsilon * \tau_{t-\varepsilon;\varepsilon}^\lambda)(x) = \tau_{t-\varepsilon;\varepsilon}^\lambda(x) + \lambda\varepsilon(D * \tau_{t-\varepsilon;\varepsilon})(x) - \varepsilon\tau_{t-\varepsilon;\varepsilon}^\lambda(x)$. Thus, we obtain that

$$\frac{1}{\varepsilon}[\tau_{t;\varepsilon}^\lambda(x) - \tau_{t-\varepsilon;\varepsilon}^\lambda(x)] = \lambda(D * \tau_{t-\varepsilon;\varepsilon}^\lambda)(x) - \tau_{t-\varepsilon;\varepsilon}^\lambda(x) + \varepsilon \sum_{s=2\varepsilon}^{t}{}^{\bullet} \frac{1}{\varepsilon^2}(\pi_{s;\varepsilon}^\lambda * p_\varepsilon * \tau_{t-s-\varepsilon;\varepsilon}^\lambda)(x). \tag{2.43}$$

Using the pointwise convergence of all terms and dominated convergence for the sum over s, we obtain (2.15). To prove (1.13), we use (2.15) together with $\hat{\tau}_0^\lambda(0) = 1$, which implies that for all $\lambda < \lambda_c$

$$\begin{aligned} 1 &= \int_0^\infty \partial_t \hat{\tau}_t^\lambda(0) dt = \int_0^\infty \Big[\lambda \hat{\tau}_t^\lambda(0) - \hat{\tau}_t^\lambda(0) + \int_0^t \hat{\pi}_s^\lambda(0) \hat{\tau}_{t-s}^\lambda(0) ds\Big] dt \\ &= \int_0^\infty \hat{\tau}_t^\lambda(0) dt \Big[\lambda - 1 + \int_0^\infty \hat{\pi}_s^\lambda(0) ds\Big], \end{aligned} \tag{2.44}$$

so that

$$\int_0^\infty \hat{\tau}_t^\lambda(0) dt = \Big[\lambda - 1 + \int_0^\infty \hat{\pi}_s^\lambda(0) ds\Big]^{-1}. \tag{2.45}$$

By (2.29), we must have that λ_c satisfies

$$\lambda_c = 1 - \int_0^\infty \hat{\pi}_s^{\lambda_c}(0) ds, \tag{2.46}$$

so that we can rewrite (2.45) as

$$\int_0^\infty \hat{\tau}_t^\lambda(0) dt = [f(\lambda) - f(\lambda_c)]^{-1}, \tag{2.47}$$

where $f(\lambda) = \lambda + \int_0^\infty \hat{\pi}_s^\lambda(0) ds$. By the mean-value theorem, together with the fact that $|\partial_\lambda \hat{\pi}_s^\lambda(0)|$ is integrable, there exists a $\lambda_* \in (\lambda, \lambda_c)$ such that

$$\int_0^\infty \hat{\tau}_t^\lambda(0) dt = -(\lambda_c - \lambda)^{-1} f'(\lambda_*)^{-1}. \tag{2.48}$$

By continuity of $\lambda \mapsto f'(\lambda)$ as $\lambda \uparrow \lambda_c$, we obtain (1.13) with $C = -1/f'(\lambda_c)$. ☐

The proofs of Theorems 1.1 and 1.2 are now reduced to the proof of the Propositions 2.1 and 2.5. Proposition 2.5 will be proved in Section 6. The proof of Proposition 2.1 is reduced to Propositions 2.2–2.3, 2.4 and 2.5, which will be proved in Sections 4 and 5, respectively. The advancement of the induction hypotheses is deferred to Appendix A. We start in Section 3 by deriving the lace expansion (2.16).

3 The lace expansion

In this section, we derive the lace expansion in (2.16). The same type of recursion relation was used for *discrete* models, such as (weakly) self-avoiding walk in $\mathbb{Z}^d$ [5, 10, 13, 17, 19, 27, 28, 29] and oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$ [17, 18, 21, 22].

From now on, we will suppress the dependence on ε and λ when no confusion can arise, and write, e.g., $\pi_t(x) = \pi_{t;\varepsilon}^\lambda(x)$. In Section 3.1, we obtain (3.4), which is equivalent to the recursion relation in (2.16), and the expression (3.5) for $\pi_t(x)$. In Section 3.2, we obtain the expressions (3.12)–(3.13) for $\partial_\lambda \pi_t(x)$.

3.1 Expansion for the two-point function

Before starting the expansion for $\tau_t(x)$, we introduce some notation.

Definition 3.1. (i) We write $b \to (y, s)$ for the event that b is occupied and $\overline{b} \to (y, s)$, where we recall that $\overline{b}$ is the top of the bond b.

(ii) We write $(y, s) \to (x, t)$ *off* b for the event that (y, s) is connected to (x, t) without using b, and define $\mathbf{C}^b(y, s) = \{(x, t) \in \mathbb{Z}^d \times \varepsilon\mathbb{Z}_+ : (y, s) \to (x, t) \text{ off } b\}$.

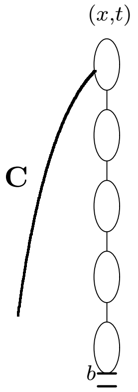

Figure 3: A graphical representation of the event $E[b,(x,t);\mathbf{C}]$. Only the last sausage is allowed to share sites with $\mathbf{C}$.

(iii) If $(y,s) \to \underline{b}$ and $\overline{b} \to (x,t) \notin \mathbf{C}^b(y,s)$, then b is said to be *pivotal* for $(y,s) \to (x,t)$. We define $\mathtt{piv}[(y,s),(x,t)]$ to be the set of pivotal bonds for $(y,s) \to (x,t)$.

(iv) If $(y,s) \to (x,t)$ and $\mathtt{piv}[(y,s),(x,t)] = \varnothing$ (i.e., either $(y,s) = (x,t)$ or there are at least two non-zero bond-disjoint occupied paths from (y,s) to (x,t)), then we write $(y,s) \rightrightarrows (x,t)$ and say that (y,s) is *doubly-connected* to (x,t).

For $\mathbf{C} \subset \mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ and a bond b with $\overline{b} \notin \mathbf{C}$ or $\underline{b} \notin \mathbf{C}$, we define

$$F[b,(x,t);\mathbf{C}] = \bigcup_{(y,s)\in\mathbf{C}} \{b \to (y,s) \to (x,t)\} \tag{3.1}$$

to be the event that b is occupied and that there is at least one occupied path from $\overline{b}$ to (x,t) passing through $\mathbf{C}$. Let

$$E[b,(x,t);\mathbf{C}] = F[b,(x,t);\mathbf{C}] \cap \Big(\{\overline{b} \rightrightarrows (x,t)\} \mathbin{\dot\cup} \big(\{\mathtt{piv}[\overline{b},(x,t)] \neq \varnothing\} \cap \bigcap_{b' \in \mathtt{piv}[\overline{b},(x,t)]} F[b,\underline{b}';\mathbf{C}]^c\big)\Big) \tag{3.2}$$

denote the event that b is occupied, that there is at least one occupied path from $\overline{b}$ to (x,t) passing through $\mathbf{C}$, and that either $\overline{b} \rightrightarrows (x,t)$, or that for any $b' \in \mathtt{piv}[\overline{b},(x,t)] \neq \varnothing$, there are no occupied paths from $\overline{b}$ to $\underline{b}'$ passing through $\mathbf{C}$. If we regard the connection from $\overline{b}$ to (x,t) as a string of sausages as in Figure 2, the event $E[b,(x,t);\mathbf{C}]$ implies that the last sausage is the first one among all sausages which intersects $\mathbf{C}$. See Figure 3 for a graphical representation of the event $E[b,(x,t);\mathbf{C}]$.

For $n \geq 1$, we will write $\vec{b}_n$ for the ordered set of n bonds $(b_1,\dots,b_n)$. For $N \geq 0$, we define

$$\tilde{E}^{(N)}_{\vec{b}_N}(x,t) = \{(o,0) \rightrightarrows \underline{b}_1\} \cap \bigcap_{i=1}^{N} E[b_i,\underline{b}_{i+1};\mathbf{C}^{b_i}(\overline{b}_{i-1})], \tag{3.3}$$

where we use the convention that $\overline{b}_0 = (o,0)$ and $\underline{b}_{N+1} = (x,t)$, and write $\tilde{E}^{(0)}(x,t) = \{(o,0) \rightrightarrows (x,t)\}$, where the subscript $\vec{b}_0$ has been dropped.

Proposition 3.2 (The lace expansion). *For* $\lambda \geq 0$ *and* $(x,t) \in \mathbb{Z}^d \times \varepsilon\mathbb{N}$,

$$\tau_t(x) = \pi_t(x) + \sum_{s=0}^{t-\varepsilon}{}^{\bullet} (\pi_s * p_\varepsilon * \tau_{t-s-\varepsilon})(x), \tag{3.4}$$

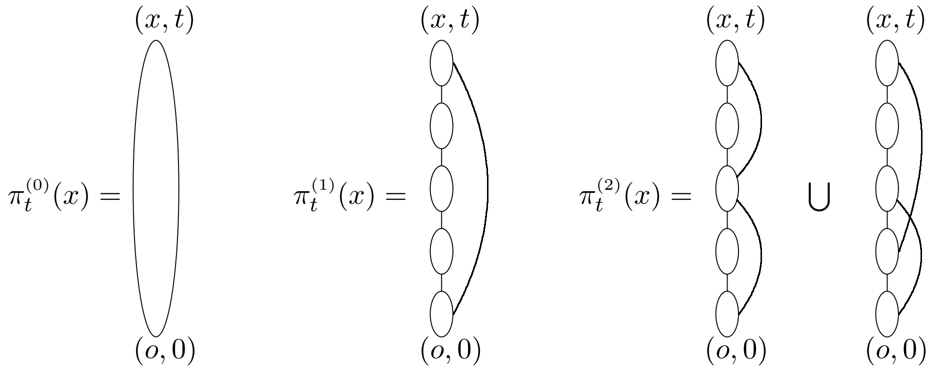

Figure 4: Graphical representations of $\pi_t^{(0)}, \pi_t^{(1)}$ and $\pi_t^{(2)}$.

where $\pi_t(x) = \sum_{N=0}^{\infty}(-1)^N \pi_t^{(N)}(x)$ *with*

$$\pi_t^{(N)}(x) = \sum_{\vec{b}_N} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(x,t)\big). \tag{3.5}$$

Note that the sum over N in (3.5) is a finite sum, as long as t/ε is finite, since each of the bonds $b_1, \dots, b_N$ eats up at least a time displacement ε, so that $\pi_t^{(N)}(x) = 0$ for $N > t/\varepsilon$. The above proposition is equivalent to [24, Proposition 4.1], in which the expansion was performed a finite number of times, and the remainder was kept unexpanded. Although the proof is the same, we prove Proposition 3.2 to make this paper self-contained, as well as to apply the same idea to the proof of Proposition 3.3 below.

Proof. The two-point function $\tau_t(x)$ is the probability of the event $(o,0) \to (x,t)$. This event can be decomposed into two cases depending on whether $\texttt{piv}[(o,0),(x,t)]$ is empty or not. The contribution due to $\texttt{piv}[(o,0),(x,t)] = \varnothing$ is $\pi_t^{(0)}(x)$. When $\texttt{piv}[(o,0),(x,t)] \neq \varnothing$, we fix the first element b of $\texttt{piv}[(o,0),(x,t)]$ and then sum over b to obtain

$$\tau_t(x) = \pi_t^{(0)}(x) + \sum_b \mathbb{P}((o,0) \rightrightarrows \underline{b};\ b \to (x,t) \notin \mathbf{C}^b(o,0)). \tag{3.6}$$

We ignore the condition $(x,t) \notin \mathbf{C}^b(o,0)$ in (3.6) and use inclusion-exclusion to make up for the arising error. Thus, (3.6) equals

$$\pi_t^{(0)}(x) + \sum_b \mathbb{P}((o,0) \rightrightarrows \underline{b};\ b \to (x,t)) - Q_t^{(1)}(x),$$

where we define $Q_t^{(n)}(x)$ for $n \geq 1$ by

$$Q_t^{(n)}(x) = \sum_{\vec{b}_n} \mathbb{P}\big(\tilde{E}_{\vec{b}_{n-1}}^{(n-1)}(\underline{b}_n);\ b_n \to (x,t) \in \mathbf{C}^{b_n}(\overline{b}_{n-1})\big), \tag{3.7}$$

and where we recall that $\tilde{E}^{(0)}(\underline{b}_1) = \{(o,0) \rightrightarrows \underline{b}_1\}$ and $\overline{b}_0 = (o,0)$. We next note that we can write

$$\{(o,0) \rightrightarrows \underline{b};\ b \to (x,t)\} = \{(o,0) \rightrightarrows \underline{b}\} \cap \{b \text{ occupied}\} \cap \{\overline{b} \to (x,t)\}.$$

The first depends on bonds before the time variable of $\underline{b}$, the second on bonds between the time variables of $\underline{b}$ and $\overline{b}$, and the third on bonds after the time variable of $\overline{b}$. Thus, the Markov property, which holds for general oriented percolation models, yields that

$$\sum_b \mathbb{P}((o,0) \rightrightarrows \underline{b};\ b \to (x,t)) = \sum_{s=0}^{t-\varepsilon}{}^{\bullet} (\pi_s^{(0)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x). \tag{3.8}$$

This completes the first step of the inclusion-exclusion procedure, and we are left to identify $Q_t^{(1)}(x)$. In order to do so, we prove below, again using inclusion-exclusion, that, for $n \geq 1$,

$$Q_t^{(n)}(x) = -Q_t^{(n+1)}(x) + \pi_t^{(n)}(x) + \sum_s\nolimits^{\bullet} (\pi_s^{(n)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x). \tag{3.9}$$

The equations (3.4)–(3.5) follow by repeated use of (3.9) until the remainder $Q_t^{(n)}(x)$ vanishes, which must happen when $n > t/\varepsilon$.

Thus, to complete the derivation of the lace expansion, we are left to prove (3.9). Because of the condition $(x,t) \in \mathbf{C}^{b_n}(\overline{b}_{n-1})$ in (3.7), either the event $E[b_n,(x,t);\mathbf{C}^{b_n}(\overline{b}_{n-1})]$ occurs or there is a unique occupied bond $b_{n+1} \in \texttt{piv}[\overline{b}_n,(x,t)]$ such that the event $E[b_n,\underline{b}_{n+1};\mathbf{C}^{b_n}(\overline{b}_{n-1})]$ occurs, i.e.,

$$\begin{aligned}\{b_n \to (x,t) \in \mathbf{C}^{b_n}(\overline{b}_{n-1})\} &= E[b_n,(x,t);\mathbf{C}^{b_n}(\overline{b}_{n-1})]\\ &\quad \dot\cup \dot{\bigcup_{b_{n+1}}} \Big(E[b_n,\underline{b}_{n+1};\mathbf{C}^{b_n}(\overline{b}_{n-1})] \cap \{b_{n+1} \to (x,t) \notin \mathbf{C}^{b_{n+1}}(\overline{b}_n)\}\Big). \end{aligned} \tag{3.10}$$

By (3.5), the contribution from the former case is $\pi_t^{(n)}(x)$, which is the second term in (3.9). In the contribution from the latter case we again ignore the condition $(x,t) \notin \mathbf{C}^{b_{n+1}}(\overline{b}_n)$ and use inclusion-exclusion to make up for it, as well as (3.7), to arrive at

$$\sum_{\vec{b}_{n+1}} \mathbb{P}\big(\tilde{E}_{\vec{b}_n}^{(n)}(\underline{b}_{n+1});\, b_{n+1} \to (x,t) \notin \mathbf{C}^{b_{n+1}}(\overline{b}_n)\big) = \sum_{\vec{b}_{n+1}} \mathbb{P}\big(\tilde{E}_{\vec{b}_n}^{(n)}(\underline{b}_{n+1});\, b_{n+1} \to (x,t)\big) - Q_t^{(n+1)}(x). \tag{3.11}$$

Again by the Markov property, the first term in the right-hand side of (3.11) equals the last term in (3.9). This completes the proof of (3.9) and thus the proof of Proposition 3.2. ☐

3.2 Representation for the derivative

We obtain the expression for $\partial_\lambda \pi_t(x)$ in the following proposition.

Proposition 3.3. *For $\lambda > 0$ and $(x,t) \in \mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$,*

$$\partial_\lambda \pi_t(x) = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \Pi_t^{(N)}(x), \tag{3.12}$$

where $\Pi_t^{(N)}(x) = \sum_{n=1}^N \Pi_t^{(N;n)}(x)$ with

$$\Pi_t^{(N;n)}(x) = \sum_{\substack{\vec{b}_N, b:\\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(x,t);\, b \in \{b_n\} \cup \texttt{piv}[\overline{b}_n, \underline{b}_{n+1}]\big), \tag{3.13}$$

and where $\underline{b}_{N+1}$ is defined to be (x,t).

By the same reason as in (3.5), the sum in (3.12) is a finite sum. We prove (3.12) by differentiating the expression (3.4) for $\tau_t(x)$ and comparing it with the expression for $\partial_\lambda \tau_t(x)$ obtained by using Russo's formula, rather than differentiation of $\partial_\lambda \pi_t(x)$ directly. Direct differentiation of the expressions (3.5) for $\pi_t(x)$ may be possible, but is cumbersome because of the complex combination of *increasing* and *decreasing* events consisting of $\pi_t^{(N)}(x)$, where an increasing (resp., decreasing) event is an event that is more (resp., less) likely to occur as λ increases. We note that instead of a difference of two terms due to the pivotals for the increasing and decreasing events, we only obtain a single sum over pivotals. Thus, an intricate cancellation takes place. This is further demonstrated by the fact that there is no contribution from $N = 0$. Thus, it is *not* true that $\partial_\lambda \pi_t^{(N)}(x) = \frac{1}{\lambda}\Pi_t^{(N)}(x)$.

Proof. In the proof it will be convenient to split $\Pi_t^{(N,n)}(x) = \Pi_t^{(N;n,1)}(x) + \Pi_t^{(N;n,2)}(x)$, where the first is the contribution from $b = b_n$, whereas the second is the contribution from $b \in \texttt{piv}[\overline{b}_n, \underline{b}_{n+1}]$.

To obtain an expression for $\partial_\lambda \tau_t(x)$, we use *Russo's formula* [4, 7]. Let E be an increasing event that depends only on finitely many *spatial* bonds. Then

$$\partial_\lambda \mathbb{P}(E) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \mathbb{P}(E \text{ occurs; } b \text{ is pivotal for } E),$$

where we used the fact that only spatial pivotal bonds for E are responsible to the differentiation with respect to λ. We apply this formula to $E = \{(o,0) \to (x,t)\}$, which is possible as long as t/ε is finite[1], and obtain

$$\partial_\lambda \tau_t(x) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \mathbb{P}((o,0) \to (x,t);\, b \in \texttt{piv}[(o,0),(x,t)]). \tag{3.14}$$

We follow the same strategy as in Section 3.1 to obtain a recursion relation, now for $\partial_\lambda \tau_t(x)$ rather than for $\tau_t(x)$.

Since $b \in \texttt{piv}[(o,0),(x,t)]$ is equivalent to the event that $(o,0) \to \underline{b}$ and $\overline{b} \to (x,t) \notin \mathbf{C}^b(o,0)$, the right hand side of (3.14) equals

$$\frac{1}{\lambda} \sum_{b \text{ spatial}} \Bigg[\mathbb{P}((o,0) \rightrightarrows \underline{b};\, b \to (x,t) \notin \mathbf{C}^b(o,0)) \\ + \sum_{b_1} \mathbb{P}((o,0) \rightrightarrows \underline{b}_1;\, b_1 \to b \to (x,t) \notin \mathbf{C}^{b_1}(o,0) \cup \mathbf{C}^b(o,0)) \Bigg], \tag{3.15}$$

where the first and second terms correspond respectively to when b is or is not the first element of $\texttt{piv}[(o,0),(x,t)]$. The contribution from the first term is the same as (3.6), apart from the factor $\frac{1}{\lambda}$ and the restriction that b has to be a spatial bond. Thus, the first term equals

$$\sum_s{}^{\bullet} (\pi_s^{(0)} * \varepsilon D * \tau_{t-s-\varepsilon})(x) + \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \Bigg[\Pi_t^{(N;1,1)}(x) + \sum_s{}^{\bullet} (\Pi_s^{(N;1,1)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x) \Bigg]. \tag{3.16}$$

For the second term in (3.15), we use the relation $\mathbf{C}^b(o,0) = \mathbf{C}^{b_1}(o,0) \cup \mathbf{C}^b(\overline{b}_1)$, which holds on the event $\{(o,0) \rightrightarrows \underline{b}_1\} \cap \{b_1 \text{ is occupied}\}$. We ignore the condition that $(x,t) \notin \mathbf{C}^{b_1}(o,0)$ and use inclusion-exclusion to make up for the arising error. Using the Markov property, the contribution from the second term in (3.15) is

$$\frac{1}{\lambda} \sum_{\substack{b_1, b: \\ b \text{ spatial}}} \mathbb{P}((o,0) \rightrightarrows \underline{b}_1;\, b_1 \to b \to (x,t) \notin \mathbf{C}^b(\overline{b}_1)) - R_t^{(1)}(x) = \sum_s{}^{\bullet} (\pi_s^{(0)} * p_\varepsilon * \partial_\lambda \tau_{t-s-\varepsilon})(x) - R_t^{(1)}(x), \tag{3.17}$$

where we define $R_t^{(n)}(x)$ by

$$R_t^{(n)}(x) = \frac{1}{\lambda} \sum_{\substack{\vec{b}_n, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}_{\vec{b}_{n-1}}^{(n-1)}(\underline{b}_n);\, b_n \to (x,t) \in \mathbf{C}^{b_n}(\overline{b}_{n-1});\, b \in \texttt{piv}[\overline{b}_n,(x,t)]\big), \tag{3.18}$$

---

[1] First we apply Russo's formula to the restricted two-point function $\tau_{t;R}(x)$ inside the space-time box $[-R,R]^d \times [0,t]$ with free boundary condition. Then (3.14) follows from the fact that $\tau_{t;R}(x) \to \tau_t(x)$ and that $\partial_\lambda \tau_{t;R}(x)$ converges to the right-hand side of (3.14), uniformly in $\lambda \in [\lambda_0, \lambda_1]$ for any $0 < \lambda_0 \le \lambda_1 < \infty$. This is because the coupling function is summable, i.e., $\sum_x D(x) = 1 < \infty$.

and we recall that $\tilde{E}^{(0)}(\underline{b}_1) = \{(o,0) \rightrightarrows \underline{b}_1\}$ and $\overline{b}_0 = (o,0)$. Note that $R_t^{(n)}(x)$ is similar to $Q_t^{(n)}(x)$ in (3.7), except for the factor $\frac{1}{\lambda}$, the sum over spatial bonds b, and the extra condition $b \in \texttt{piv}[\overline{b}_n,(x,t)]$. We will use inclusion-exclusion below to prove that, for $n \geq 1$,

$$\begin{aligned}
R_t^{(n)}(x) = &- R_t^{(n+1)}(x) + \sum_s{}^{\bullet} \big[(\pi_s^{(n)} * \varepsilon D * \tau_{t-s-\varepsilon})(x) + (\pi_s^{(n)} * p_\varepsilon * \partial_\lambda \tau_{t-s-\varepsilon})(x)\big] \\
&+ \frac{1}{\lambda} \sum_{N=n+1}^{\infty} (-1)^{N-n} \bigg[\Pi_t^{(N;n+1,1)}(x) + \sum_s{}^{\bullet} (\Pi_t^{(N;n+1,1)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg] \\
&+ \frac{1}{\lambda} \sum_{N=n}^{\infty} (-1)^{N-n} \bigg[\Pi_t^{(N;n,2)}(x) + \sum_s{}^{\bullet} (\Pi_t^{(N;n,2)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg]. \qquad (3.19)
\end{aligned}$$

Before proving (3.19), we complete the proof of Proposition 3.3 assuming (3.19). By (3.15)–(3.17) and repeated use of (3.19) until the remainder term $R_t^{(n)}(x)$ vanishes,

$$\begin{aligned}
&\partial_\lambda \tau_t(x) - \sum_s{}^{\bullet} \big[(\pi_s * \varepsilon D * \tau_{t-s-\varepsilon})(x) + (\pi_s * p_\varepsilon * \partial_\lambda \tau_{t-s-\varepsilon})(x)\big] \qquad (3.20) \\
&\quad = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \bigg[\Pi_t^{(N;1,1)}(x) + \sum_s{}^{\bullet} (\Pi_t^{(N;1,1)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg] \\
&\qquad + \frac{1}{\lambda} \sum_{n=1}^{\infty} (-1)^n \sum_{N=n+1}^{\infty} (-1)^{N-n} \bigg[\Pi_t^{(N;n+1,1)}(x) + \sum_s{}^{\bullet} (\Pi_t^{(N;n+1,1)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg] \\
&\qquad + \frac{1}{\lambda} \sum_{n=1}^{\infty} (-1)^n \sum_{N=n}^{\infty} (-1)^{N-n} \bigg[\Pi_t^{(N;n,2)}(x) + \sum_s{}^{\bullet} (\Pi_t^{(N;n,2)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg] \\
&\quad = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \bigg[\Pi_t^{(N)}(x) + \sum_s{}^{\bullet} (\Pi_s^{(N)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg].
\end{aligned}$$

Differentiating both sides of (3.4) and comparing with the above expression, we obtain that

$$\partial_\lambda \pi_t(x) + \sum_s{}^{\bullet} (\partial_\lambda \pi_s * p_\varepsilon * \tau_{t-s-\varepsilon})(x) = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \bigg[\Pi_t^{(N)}(x) + \sum_s{}^{\bullet} (\Pi_s^{(N)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg]. \qquad (3.21)$$

We prove (3.12) by induction on t/ε. Since $\pi_0(x) = \delta_{o,x}$ and $\Pi_0^{(N)}(x) = 0$ for all $N \geq 1$, we obtain (3.12) for $t/\varepsilon = 0$. Suppose that (3.12) holds for all $t/\varepsilon \leq m$. Then the contribution from the second term in the brackets of (3.21) equals the second term on the left-hand side of (3.21), and thus (3.12) for $t/\varepsilon = m+1$ holds. This completes the inductive proof of (3.12).

In order to complete the proof of Proposition 3.3, we prove (3.19). Because of the condition $(x,t) \in \mathbf{C}^{b_n}(\overline{b}_{n-1})$ in (3.18), either the event $E[b_n,(x,t);\mathbf{C}^{b_n}(\overline{b}_{n-1})]$ occurs or there is an occupied bond $b_{n+1} \in \texttt{piv}[\overline{b}_n,(x,t)]$ such that the event $E[b_n,\underline{b}_{n+1};\mathbf{C}^{b_n}(\overline{b}_{n-1})]$ occurs. The contribution from the former case to $R_t^{(n)}(x)$ is

$$\frac{1}{\lambda} \sum_{\substack{\vec{b}_n,b:\\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n-1)}_{\vec{b}_{n-1}}(\underline{b}_n) \cap E[b_n,(x,t);\mathbf{C}^{b_n}(\overline{b}_{n-1})];\ b \in \texttt{piv}[\overline{b}_n,(x,t)]\big) = \frac{1}{\lambda} \Pi_t^{(n;n,2)}(x). \qquad (3.22)$$

The contribution from the latter case is, as in (3.15),

$$\frac{1}{\lambda} \sum_{\substack{\vec{b}_n,b:\\ b \text{ spatial}}} \Big[\mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b});\ b \to (x,t) \notin \mathbf{C}^{b}(\overline{b}_n)\big)$$

$$
\begin{aligned}
&+\sum_{b_{n+1}} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1});\ b_n \to b_{n+1} \to b \to (x,t) \notin \mathbf{C}^b(\overline{b}_n) \cup \mathbf{C}^{b_{n+1}}(\overline{b}_n)\big)\\
&+\sum_{b_{n+1}} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1});\ b_n \to b \to b_{n+1} \to (x,t) \notin \mathbf{C}^b(\overline{b}_n) \cup \mathbf{C}^{b_{n+1}}(\overline{b}_n)\big)\bigg],
\end{aligned}
\tag{3.23}
$$

where the first, second and third terms correspond respectively to when $b_{n+1} = b$, when b_{n+1} is between $\overline{b}_n$ and $\underline{b}$, and when b_{n+1} is between $\overline{b}$ and (x,t). The first term is similar to that in (3.15), and its contribution equals, as in (3.16),

$$
\sum_s{}^{\bullet}(\pi_s^{(n)} * \varepsilon D * \tau_{t-s-\varepsilon})(x) + \frac{1}{\lambda}\sum_{N=n+1}^{\infty}(-1)^{N-n}\bigg[\Pi_t^{(N;n+1,1)}(x) + \sum_s{}^{\bullet}(\Pi_s^{(N;n+1,1)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg]. \tag{3.24}
$$

For the second term in (3.23), we use the relation $\mathbf{C}^b(\overline{b}_n) = \mathbf{C}^{b_{n+1}}(\overline{b}_n) \cup \mathbf{C}^b(\overline{b}_{n+1})$, which holds on the event $\overline{b}_n \to b_{n+1}$, ignore the condition $(x,t) \notin \mathbf{C}^{b_{n+1}}(\overline{b}_n)$, and make up for the arising error using inclusion-exclusion. By the Markov property, the contribution from the second term in (3.23) is, as in (3.17),

$$
\begin{aligned}
&\frac{1}{\lambda}\sum_{\substack{\vec{b}_{n+1},b:\\ b\text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1});\ b_{n+1} \to b \to (x,t) \notin \mathbf{C}^b(\overline{b}_{n+1})\big) - R_t^{(n+1)}(x)\\
&\qquad\qquad = \sum_s{}^{\bullet}(\pi_s^{(n)} * p_\varepsilon * \partial_\lambda \tau_{t-s-\varepsilon})(x) - R_t^{(n+1)}(x).
\end{aligned}
\tag{3.25}
$$

For the third term in (3.23), the condition $(x,t) \notin \mathbf{C}^b(\overline{b}_n)$ can be replaced by $\underline{b}_{n+1} \notin \mathbf{C}^b(\overline{b}_n)$, because $\{\underline{b}_{n+1} \notin \mathbf{C}^b(\overline{b}_n)\} \setminus \{(x,t) \notin \mathbf{C}^b(\overline{b}_n)\}$ implies the existence of an occupied path from $\overline{b}_n$ to (x,t) without using b and b_{n+1}, and this violates $(x,t) \notin \mathbf{C}^{b_{n+1}}(\overline{b}_n)$. By inclusion-exclusion, the third term in (3.23) equals

$$
\begin{aligned}
\frac{1}{\lambda}\sum_{\substack{\vec{b}_{n+1},b:\\ b\text{ spatial}}} &\Big[\mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1});\ b_n \to b \to \underline{b}_{n+1} \notin \mathbf{C}^b(\overline{b}_n);\ b_{n+1} \to (x,t)\big)\\
&\quad - \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1});\ b_n \to b \to \underline{b}_{n+1} \notin \mathbf{C}^b(\overline{b}_n);\ b_{n+1} \to (x,t) \in \mathbf{C}^{b_{n+1}}(\overline{b}_n)\big)\Big],
\end{aligned}
\tag{3.26}
$$

where the first term equals, by the Markov property,

$$
\frac{1}{\lambda}\sum_s{}^{\bullet}(\Pi_s^{(n;n,2)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x). \tag{3.27}
$$

For the second term in (3.26), we use the same argument as above (3.10). Because of the condition $(x,t) \in \mathbf{C}^{b_{n+1}}(\overline{b}_n)$, either the event $E[b_{n+1}, (x,t); \mathbf{C}^{b_{n+1}}(\overline{b}_n)]$ occurs or there is an occupied bond b_{n+2} such that $E[b_{n+1}, \underline{b}_{n+2}; \mathbf{C}^{b_{n+1}}(\overline{b}_n)]$ occurs. By repeated use of inclusion-exclusion and the Markov property, as above (3.10), the second term in (3.26) equals

$$
\frac{1}{\lambda}\sum_{N=n+1}^{\infty}(-1)^{N-n}\bigg[\Pi_t^{(N;n,2)}(x) + \sum_s{}^{\bullet}(\Pi_s^{(N;n,2)} * p_\varepsilon * \tau_{t-s-\varepsilon})(x)\bigg]. \tag{3.28}
$$

Equation (3.19) is obtained by combining (3.22), (3.24)–(3.25) and (3.27)–(3.28). This completes the proof of Proposition 3.3. ☐

4 Bounds on the lace expansion

In this section, we prove Propositions 2.2–2.3 following the same strategy as in [18, Section 4] for oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$.

To prove Propositions 2.2–2.3, we first bound $\pi_t^{(N)}(x)$ and $\Pi_t^{(N;n)}(x)$ by certain diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$, respectively, which are constructed in Section 4.1.2 in terms of two-point functions. Then, using the bounds in (2.24) on the two-point function, we estimate these diagram functions. This completes the proof of Propositions 2.2–2.3. These main two steps are summarized in the following lemmas.

Lemma 4.1. *For $\lambda \geq 0$ and $N \geq n \geq 1$,*

$$\pi_t^{(N-1)}(x) \leq P_t^{(N-1)}(x), \qquad \qquad \Pi_t^{(N;n)}(x) \leq \tilde{P}_t^{(N;n)}(x). \tag{4.1}$$

Lemma 4.2. *Suppose that (2.24) holds for some λ_0 and all $s \leq t$.*

(i) Let $d > 4$. There exist $\beta_0 > 0$ and $C < \infty$ such that, for $\lambda \leq \lambda_0$, $\beta < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ with $2\varepsilon \leq s \leq t+\varepsilon$, and $q = 0, 2, 4$,

$$\sum_x |x|^q \, P_t^{(N)}(x) \leq \frac{C\beta\varepsilon^2\sigma^q N^{q/2}}{(1+t)^{(d-q)/2}} (C\beta)^{0\vee(N-1)}, \quad \text{for } N \geq 0, \tag{4.2}$$

$$\sum_x \tilde{P}_t^{(N;n)}(x) \leq \frac{C\beta\varepsilon^2}{(1+t)^{(d-2)/2}} (C\beta)^{N-1}, \quad \text{for } 1 \leq n \leq N. \tag{4.3}$$

(ii) Let $d \leq 4$ with $\alpha \equiv bd - \frac{4-d}{2} > 0$ and $t \leq T \log T$. There exist $\beta_0 > 0$, $\mu \in (0, \alpha)$ and $C < \infty$ such that, for $\lambda \leq \lambda_0$, $\beta_1 < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ with $2\varepsilon \leq s \leq t + \varepsilon$, and $q = 0, 2, 4$,

$$\sum_x |x|^q \, P_t^{(N)}(x) \leq \frac{C\beta_T\varepsilon^2\sigma_T^q N^{q/2}}{(1+t)^{(d-q)/2}} (C\hat{\beta}_T)^{0\vee(N-1)}, \quad \text{for } N \geq 0, \tag{4.4}$$

$$\sum_x \tilde{P}_t^{(N;n)}(x) \leq \frac{C\beta_T\varepsilon^2}{(1+t)^{(d-2)/2}} (C\hat{\beta}_T)^{N-1}, \quad \text{for } 1 \leq n \leq N, \tag{4.5}$$

where $\beta_T = \beta_1 T^{-bd}$ and $\hat{\beta}_T = \beta_1 T^{-\mu}$.

We prove Lemma 4.1 in Section 4.1 and Lemma 4.2 in Sections 4.2–4.3.

Proof of Propositions 2.2–2.3 assuming Lemmas 4.1–4.2. The inequalities (2.25) and (2.27) follow from (3.5), (3.12)–(3.13) and (4.1)–(4.3), if β is sufficiently small. The proof of (2.26) is the same as that of Proposition 2.2(ii) in [18, Section 4.3], together with (2.25). This completes the proof of Proposition 2.2.

Proposition 2.3 is proved by using (4.4)–(4.5) instead of (4.2)–(4.3), if $\hat{\beta}_T$ is sufficiently small. $\square$

4.1 Bounds in terms of the diagram functions

In this section, we define the diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$ and prove Lemma 4.1.

This section is similar to [18, Section 4.1] for oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$, where bounds on $\pi_t^{(N)}(x)$ are proved by using diagram functions that arise from the Hara-Slade lace expansion. Since the expansion used in this paper is somewhat simpler, we use simpler diagram functions. However, to consider the case $\varepsilon \ll 1$, as in [24], extra care is needed to obtain the factor ε^2 in (4.2)–(4.5). The diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$, with $t > 0$, are defined, in (4.16) and (4.19)–(4.20) below, to be of order ε^2 for this purpose.

4.1.1 Preliminaries

From now on, we will frequently use bold letters $\boldsymbol{o}, \boldsymbol{x}, \dots$ to represent elements in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$, such as $\boldsymbol{o} = (o, 0)$ and $\boldsymbol{x} = (x, t)$, and write $\tau(\boldsymbol{x}) = \tau_t(x)$, $P^{(N)}(\boldsymbol{x}) = P_t^{(N)}(x)$, and so on. For $\boldsymbol{v} = (v, s)$, let $\boldsymbol{v}_+ = (v, s + \varepsilon)$.

Recall the definitions (3.5) and (3.13) for $\pi^{(N)}(\boldsymbol{x})$ and $\Pi^{(N)}(\boldsymbol{x})$, which involve the event $\tilde{E}^{(N)}_{\vec{b}_N}(\boldsymbol{x})$. Let $I(\overline{b}_N, \boldsymbol{y}, \boldsymbol{x})$ equal $\{\overline{b}_N \to \boldsymbol{x}\}$ if $\boldsymbol{y} = \boldsymbol{x}$, otherwise

$$I(\overline{b}_N, \boldsymbol{y}, \boldsymbol{x}) = \bigcup_{\substack{b \text{ spatial:} \\ \underline{b} = \boldsymbol{y}}} \{\overline{b}_N \to b \to \boldsymbol{x}\} \cup \bigcup_{\substack{b \text{ spatial:} \\ \overline{b} = \boldsymbol{y}}} \big\{\{\overline{b}_N \to b\} \circ \{(\underline{b}, \underline{b}_+) \to \boldsymbol{x}\}\big\}. \tag{4.6}$$

Here $E_1 \circ E_2$ is the event that E_1 and E_2 occur *disjointly*, i.e., there is a set B of bonds such that E_1 occurs on B and E_2 occurs on the complement of B. We also define

$$M(b, \boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v}) = \big\{\{b \to \boldsymbol{u}\} \circ \{\boldsymbol{w} \to \boldsymbol{u}\}\big\} \cap \{I(\overline{b}, \boldsymbol{v}, \boldsymbol{u}) \cup I(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u})\}, \tag{4.7}$$

$$M'(b, b', \boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v}) = M(b, \boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v}) \cap \{\overline{b} \to b' \to \boldsymbol{u}\}. \tag{4.8}$$

The main observation to bound the diagrams by induction in N is the following lemma.

Lemma 4.3. *For $N \geq n \geq 1$,*

$$\tilde{E}^{(N)}_{\vec{b}_N}(\boldsymbol{x}) \subset \tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{x}), \tag{4.9}$$

$$\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \subset \bigcup_{\boldsymbol{v}} \Big\{\{\tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v})\} \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})\Big\}, \tag{4.10}$$

$$\begin{aligned} &\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \cap \{b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]\} \\ &\quad \subset \begin{cases} \bigcup_{\boldsymbol{v}} \Big\{\{\tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v})\} \circ M'(b_N, b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})\Big\}, & \text{if } n = N, \\ \bigcup_{\boldsymbol{v}} \Big\{\{\tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v}) \cap \{b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]\}\} \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})\Big\}, & \text{if } n < N, \end{cases} \end{aligned} \tag{4.11}$$

where $\tilde{E}^{(m)}_{\vec{b}_m}(\underline{b}_{m+1})$ equals $\{\boldsymbol{o} \rightrightarrows \underline{b}_1\}$ if $m = 0$, and equals the whole probability space if $m = -1$, and we write $M(b_0, \overline{b}_{-1}; \underline{b}_1, \boldsymbol{v}) = \{\boldsymbol{o} \rightrightarrows \underline{b}_1\} \cap I(\boldsymbol{o}, \boldsymbol{v}, \underline{b}_1)$.

We will use Lemma 4.3 in Section 4.1.3 to prove Lemma 4.1 by induction on N.

Proof of Lemma 4.3. The inclusion in (4.9) immediately follows from the definition (3.3) and

$$E[b_N, \boldsymbol{x}; \mathbf{C}^{b_N}(\overline{b}_{N-1})] \subset \{b_N \to \boldsymbol{x}\} \circ \{\overline{b}_{N-1} \to \boldsymbol{x}\} = M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{x}).$$

We only prove (4.10), since (4.11) can be proved similarly. First, we use (4.9) to obtain $\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \subset \tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \underline{b}_N)$. Since $\tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1})$ depends only on bonds before the time variable of $\underline{b}_{N-1}$, it suffices to prove that

$$M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \subset \bigcup_{\boldsymbol{v}} \{M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v}) \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})\}. \tag{4.12}$$

Recall that $M(b_{N-1},\overline{b}_{N-2};\underline{b}_N,\underline{b}_N)=\{b_{N-1}\to\underline{b}_N\}\circ\{\overline{b}_{N-2}\to\underline{b}_N\}$. The event in the left-hand side of (4.12) implies existence of $\boldsymbol{v}\in\mathbf{C}(\overline{b}_{N-1})$ such that $\boldsymbol{v}\to\underline{b}_N$ and $M(b_N,\boldsymbol{v};\boldsymbol{x},\boldsymbol{y})$ occur disjointly. Therefore,

$$\begin{aligned}&M(b_{N-1},\overline{b}_{N-2};\underline{b}_N,\underline{b}_N)\cap M(b_N,\overline{b}_{N-1};\boldsymbol{x},\boldsymbol{y})\\&\subset\bigcup_{\boldsymbol{v}}\Big\{\big\{\{b_{N-1}\to\boldsymbol{v}\to\underline{b}_N\}\circ\{\overline{b}_{N-2}\to\underline{b}_N\}\big\}\cup\big\{\{b_{N-1}\to\underline{b}_N\}\circ\{\overline{b}_{N-2}\to\boldsymbol{v}\to\underline{b}_N\}\big\}\Big\}\circ M(b_N,\boldsymbol{v};\boldsymbol{x},\boldsymbol{y}).\end{aligned}\tag{4.13}$$

We investigate the vicinity of $\boldsymbol{v}$ in (4.13), where there are two disjoint connections, $\boldsymbol{v}\to\underline{b}_N$ and $\boldsymbol{v}\to\boldsymbol{x}$. Since there is at most one temporal bond growing out of each vertex in $\mathbb{Z}^d\times\varepsilon\mathbb{Z}_+$, at least one of the two connections has to use a spatial bond at $\boldsymbol{v}$. Therefore,

$$\{\boldsymbol{v}\to\underline{b}_N\}\circ\{\boldsymbol{v}\to\boldsymbol{x}\}\subset\bigcup_{\substack{b\text{ spatial:}\\ \underline{b}=\boldsymbol{v}}}\Big\{\big\{\{b\to\underline{b}_N\}\circ\{\underline{b}\to\boldsymbol{x}\}\big\}\cup\big\{\{(\underline{b},\underline{b}_+)\to\underline{b}_N\}\circ\{b\to\boldsymbol{x}\}\big\}\Big\}.$$

Substituting this relation into (4.13) and relabelling $\overline{b}=\boldsymbol{v}$ in the latter event $\{(\underline{b},\underline{b}_+)\to\underline{b}_N\}\circ\{b\to\boldsymbol{x}\}$, we obtain (4.12), and thus (4.10). This completes the proof. ☐

4.1.2 Construction of the diagram functions

We inductively construct the diagram functions $P^{(N)}(\boldsymbol{x})$ and $\tilde{P}^{(N;n)}(\boldsymbol{x})$ that will be proved in Section 4.1.3 to bound $\pi^{(N)}(\boldsymbol{x})$ and $\Pi^{(N;n)}(\boldsymbol{x})$ respectively. This inductive construction will also be useful to prove Lemma 4.2 in Sections 4.2–4.3.

For $b=(\boldsymbol{u},\boldsymbol{v})$ with $\boldsymbol{u}=(u,s)$ and $\boldsymbol{v}=(v,s+\varepsilon)$, we abuse notation and write $p(b)$ or $p(\boldsymbol{v}-\boldsymbol{u})$ for $p_\varepsilon(v-u)$, and $D(b)$ or $D(\boldsymbol{v}-\boldsymbol{u})$ for $D(v-u)$. Let

$$\varphi(\boldsymbol{x}-\boldsymbol{u})=\delta_{\boldsymbol{u},\boldsymbol{x}}+\sum_{b:\underline{b}=\boldsymbol{u}}p(b)\,\tau(\boldsymbol{x}-\overline{b}),\tag{4.14}$$

and

$$L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x})=\begin{cases}\displaystyle\sum_{b:\overline{b}=\boldsymbol{x}}\lambda\varepsilon D(b)\,[\varphi(\boldsymbol{x}-\boldsymbol{u})\,\tau(\underline{b}-\boldsymbol{v})+\varphi(\underline{b}-\boldsymbol{u})\,\tau(\boldsymbol{x}-\boldsymbol{v})], & \text{if } \boldsymbol{u}\neq\boldsymbol{v},\\ \displaystyle\sum_{b:\overline{b}=\boldsymbol{x}}\lambda\varepsilon D(b)\sum_{b':\underline{b}'=\boldsymbol{u}}D(b')\,[\tau(\boldsymbol{x}-\overline{b}')\,\tau(\underline{b}-\boldsymbol{u})+\tau(\underline{b}-\overline{b}')\,\tau(\boldsymbol{x}-\boldsymbol{u})], & \text{if } \boldsymbol{u}=\boldsymbol{v}.\end{cases}\tag{4.15}$$

We define

$$P^{(0)}(\boldsymbol{x})=\delta_{\boldsymbol{o},\boldsymbol{x}}+\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x}),\tag{4.16}$$

and define the *zeroth admissible lines* to be the two lines from $\boldsymbol{o}$ to $\boldsymbol{x}$ in each diagram of $\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x})$. With lines, we mean here $(\lambda\varepsilon D*\tau)(\boldsymbol{x})$ and $(\tau*\lambda\varepsilon D)(\boldsymbol{x})$ for the first term in the brackets in (4.15) with $\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{o}$, and $(\lambda\varepsilon D*\tau*\lambda\varepsilon D)(\boldsymbol{x})$ and $\tau(\boldsymbol{x})$ for the second term in the brackets in (4.15) with $\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{o}$.

Given an admissible line ℓ from $\boldsymbol{y}$ to $\boldsymbol{z}$ of a diagram function, say $\tau(\boldsymbol{z}-\boldsymbol{y})$ for simplicity, and given $\boldsymbol{v}\neq\boldsymbol{z}$, Construction $B^{\ell}_{\texttt{spat}}(\boldsymbol{v})$ is defined to be the operation in which $\tau(\boldsymbol{z}-\boldsymbol{y})$ is replaced by

$$\tau(\boldsymbol{v}-\boldsymbol{y})\sum_{b:\underline{b}=\boldsymbol{v}}\lambda\varepsilon D(b)\;\tau(\boldsymbol{z}-\overline{b}),\tag{4.17}$$

and Construction $B^{\ell}_{\texttt{temp}}(\boldsymbol{v})$ is defined to be the operation in which $\tau(\boldsymbol{z}-\boldsymbol{y})$ is replaced by

$$\sum_{b:\overline{b}=\boldsymbol{v}}\tau(\underline{b}-\boldsymbol{y})\;\lambda\varepsilon D(b)\;\mathbb{P}((\underline{b},\underline{b}_+)\to\boldsymbol{z}),\tag{4.18}$$

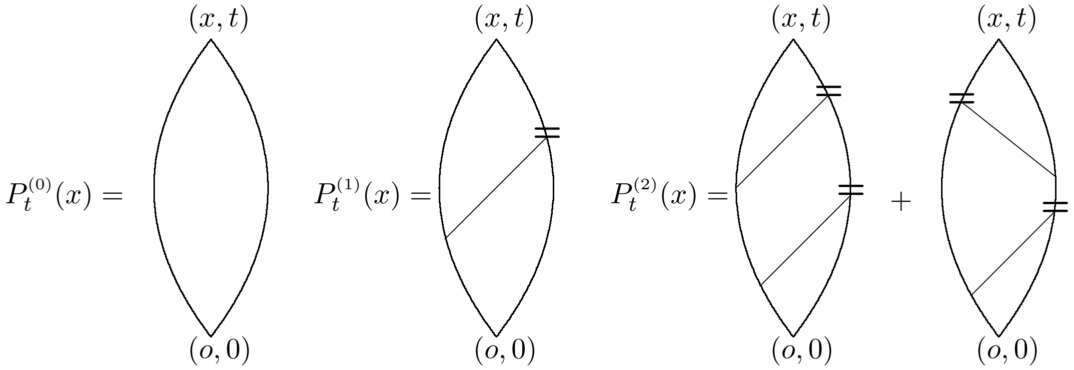

Figure 5: Graphical representations of $P_t^{(0)}(x), P_t^{(1)}(x)$ and $P_t^{(2)}(x)$. Lines indicate two-point functions, and small bars indicate a convolution with p_ε. Spatial bonds that are present at all vertices in the diagrams are left implicit.

where we recall that $(\underline{b}, \underline{b}_+)$ is a temporal bond. The sum of the results of Construction $B^\ell_{\texttt{spat}}(\boldsymbol{v})$ and Construction $B^\ell_{\texttt{temp}}(\boldsymbol{v})$ is simply said to be the result of Construction $B^\ell(\boldsymbol{v})$. We define Construction $B^\ell(s)$ to be the operation in which Construction $B^\ell(v,s)$ is performed and then followed by summation over $v \in \mathbb{Z}^d$. Construction $B^\ell_{\texttt{spat}}(s)$ and Construction $B^\ell_{\texttt{temp}}(s)$ are defined similarly.

We denote the result of applying Construction $B^\ell(\boldsymbol{v})$ to a diagram function $f(\boldsymbol{x})$ by $f(\boldsymbol{x}, B^\ell(\boldsymbol{v}))$, and define $f(\boldsymbol{x}, B^\ell_{\texttt{spat}}(\boldsymbol{v}))$ and $f(\boldsymbol{x}, B^\ell_{\texttt{temp}}(\boldsymbol{v}))$ similarly. We construct $P^{(N)}(\boldsymbol{x})$ from $P^{(N-1)}(\boldsymbol{x})$ by

$$P^{(N)}(\boldsymbol{x}) = 2\lambda\varepsilon \sum_{\boldsymbol{u}} P^{(N-1)}(\boldsymbol{u})\ L(\boldsymbol{u}, \boldsymbol{u}; \boldsymbol{x}) + \sum_\ell \sum_{\boldsymbol{u},\boldsymbol{v}:\boldsymbol{u}\neq\boldsymbol{v}} P^{(N-1)}(\boldsymbol{u}, B^\ell(\boldsymbol{v}))\ L(\boldsymbol{u}, \boldsymbol{v}; \boldsymbol{x}),$$

where $\sum_\ell$ is the sum over the $(N-1)^{\text{st}}$ admissible lines in each diagram. We write $\sum_\ell P^{(N-1)}(\boldsymbol{u}, B^\ell(\boldsymbol{u})) = 2\lambda\varepsilon P^{(N-1)}(\boldsymbol{u})$ to summarize the above expression as

$$P^{(N)}(\boldsymbol{x}) = \sum_\ell \sum_{\boldsymbol{u},\boldsymbol{v}} P^{(N-1)}(\boldsymbol{u}, B^\ell(\boldsymbol{v}))\ L(\boldsymbol{u}, \boldsymbol{v}; \boldsymbol{x}). \tag{4.19}$$

We call the newly added lines, contained in $L(\boldsymbol{u}, \boldsymbol{v}; \boldsymbol{x})$, the N^{th} *admissible lines.*

Finally, for $N \geq n \geq 1$, we define

$$\tilde{P}^{(N;n)}(\boldsymbol{x}) = \sum_\ell \sum_{\boldsymbol{v}} P^{(N)}(\boldsymbol{x}, B^\ell_{\texttt{spat}}(\boldsymbol{v})), \tag{4.20}$$

where $\sum_\ell$ is the sum over the n^{th} admissible lines.

4.1.3 Proof of Lemma 4.1

First, we prove $\pi^{(0)}(\boldsymbol{x}) \leq P^{(0)}(\boldsymbol{x})$. The first term in (4.16) is the contribution from the case of $\boldsymbol{o} = \boldsymbol{x}$. If $\boldsymbol{o} \neq \boldsymbol{x}$, there are at least two nonzero disjoint occupied paths from $\boldsymbol{o}$ to $\boldsymbol{x}$. As explained below (4.13), at least one of two nonzero disjoint occupied paths from $\boldsymbol{o}$ has to use a spatial bond at $\boldsymbol{o}$. That is,

$$\pi^{(0)}(\boldsymbol{x}) \leq \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(\{b \to \boldsymbol{x}\} \circ \{\boldsymbol{o} \to \boldsymbol{x}\}) = \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})). \tag{4.21}$$

We use the same observation at $\boldsymbol{x}$: at least one of the two nonzero disjoint connections, $\overline{b} \to \boldsymbol{x}$ and $\boldsymbol{o} \to \boldsymbol{x}$, has to use another spatial bond at $\boldsymbol{x}$. Therefore, we can bound the right-hand side of (4.21) by

$\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x})$ using the *BK inequality* [4, 7], which asserts that $\mathbb{P}(E_1\circ E_2)\le\mathbb{P}(E_1)\,\mathbb{P}(E_2)$ for increasing events E_1 and E_2. This completes the proof of $\pi^{(0)}(\boldsymbol{x})\le P^{(0)}(\boldsymbol{x})$.

Next, we consider $\pi^{(N)}(\boldsymbol{x})$ for $N\ge1$. Let

$$\pi^{(n)}(\boldsymbol{x},\boldsymbol{y})=\sum_{\vec{b}_n}\mathbb{P}\big(\tilde{E}^{(n-1)}_{\vec{b}_{n-1}}(\underline{b}_n)\cap M(b_n,\overline{b}_{n-1};\boldsymbol{x},\boldsymbol{y})\big). \tag{4.22}$$

By the convention in Lemma 4.3, $\pi^{(0)}(\boldsymbol{x},\boldsymbol{y})=\mathbb{P}(\{\boldsymbol{o}\rightrightarrows\boldsymbol{x}\}\cap I(\boldsymbol{o},\boldsymbol{y},\boldsymbol{x}))$. We prove below by induction that

$$\pi^{(n)}(\boldsymbol{x},\boldsymbol{y})\le(2\lambda\varepsilon)^{-\delta_{\boldsymbol{x},\boldsymbol{y}}}\sum_{\ell}P^{(n)}(\boldsymbol{x},B^{\ell}(\boldsymbol{y})) \tag{4.23}$$

holds for all $n\ge0$, where $\sum_\ell$ is the sum over the n^{th} admissible lines. The inequality $\pi^{(N)}(\boldsymbol{x})\le P^{(N)}(\boldsymbol{x})$, $N\ge1$, follows from (4.9) and (4.23) for $\boldsymbol{y}=\boldsymbol{x}$, together with the convention $\sum_\ell P^{(N)}(\boldsymbol{x},B^{\ell}(\boldsymbol{x}))=2\lambda\varepsilon P^{(N)}(\boldsymbol{x})$.

For $n=0$, we can assume $\boldsymbol{y}\ne\boldsymbol{x}$ because $\pi^{(0)}(\boldsymbol{x})=\pi^{(0)}(\boldsymbol{x},\boldsymbol{x})\le P^{(0)}(\boldsymbol{x})$ has already been proved. By the equivalence $\{\boldsymbol{o}\rightrightarrows\boldsymbol{x}\}\cap I(\boldsymbol{o},\boldsymbol{y},\boldsymbol{x})=\{\boldsymbol{o}\to\boldsymbol{x}\}\circ I(\boldsymbol{o},\boldsymbol{y},\boldsymbol{x})$ and the definition (4.6), we obtain

$$\pi^{(0)}(\boldsymbol{x},\boldsymbol{y})\le\sum_{\substack{b\text{ spatial:}\\ \underline{b}=\boldsymbol{y}}}\mathbb{P}(\{\boldsymbol{o}\to\boldsymbol{x}\}\circ\{\boldsymbol{o}\to b\to\boldsymbol{x}\})+\sum_{b:\overline{b}=\boldsymbol{y}}\lambda\varepsilon D(b)\,\mathbb{P}(\{\boldsymbol{o}\to\boldsymbol{x}\}\circ\{\boldsymbol{o}\to(\underline{b},\underline{b}_+)\to\boldsymbol{x}\}). \tag{4.24}$$

We prove that the first sum in the right-hand side is bounded by $\sum_\ell P^{(0)}(\boldsymbol{x},B^{\ell}_{\texttt{spat}}(\boldsymbol{y}))$, by investigating the vicinity of $\boldsymbol{o}$ and $\boldsymbol{x}$ in the diagram functions, as in the proof of $\pi^{(0)}(\boldsymbol{x})\le P^{(0)}(\boldsymbol{x})$. There are three contributions: (i) $\boldsymbol{y}=\boldsymbol{o}$, (ii) $b=(\boldsymbol{y},\boldsymbol{x})$, and (iii) $\boldsymbol{y}\ne\boldsymbol{o}$ and $\overline{b}\ne\boldsymbol{x}$. The contribution due to $\boldsymbol{y}=\boldsymbol{o}$ is bounded by $\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x})\,\delta_{\boldsymbol{o},\boldsymbol{y}}$, while the contribution due to $b=(\boldsymbol{y},\boldsymbol{x})$ is bounded by

$$\lambda\varepsilon D(\boldsymbol{x}-\boldsymbol{y})\sum_{\boldsymbol{w}}\lambda\varepsilon D(\boldsymbol{w})\ [\tau(\boldsymbol{x}-\boldsymbol{w})\,\tau(\boldsymbol{y})+\tau(\boldsymbol{y}-\boldsymbol{w})\,\tau(\boldsymbol{x})]. \tag{4.25}$$

We can similarly estimate the remaining case (iii), and obtain a bound which is $\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{y};\boldsymbol{x})$ with one of two τ's in each product $\tau\cdot\tau$ in (4.15) replaced by (4.17). Reorganizing the obtained bounds, we conclude that the first sum in (4.24) is bounded by $\sum_\ell P^{(0)}(\boldsymbol{x},B^{\ell}_{\texttt{spat}}(\boldsymbol{y}))$. In the same way, the second sum in (4.24) can be bounded by $\sum_\ell P^{(0)}(\boldsymbol{x},B^{\ell}_{\texttt{temp}}(\boldsymbol{y}))$. This completes the proof of (4.23) for $n=0$, and initializes the inductive proof for $n\ge1$.

To advance the induction, we assume that (4.23) holds for $n=N-1$, and that

$$\sum_{b:\underline{b}=\boldsymbol{u}}\mathbb{P}(M(b,\boldsymbol{v};\boldsymbol{x},\boldsymbol{y}))\le(2\lambda\varepsilon)^{\delta_{\boldsymbol{u},\boldsymbol{v}}-\delta_{\boldsymbol{x},\boldsymbol{y}}}\sum_{\ell}L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x},B^{\ell}(\boldsymbol{y})) \tag{4.26}$$

holds, where we write $\sum_\ell L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x},B^{\ell}(\boldsymbol{x}))=2\lambda\varepsilon L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x})$, similarly to the convention used in (4.19). We will prove (4.26) below. By Lemma 4.3, the BK inequality, (4.23) and (4.26), we obtain

$$\begin{aligned}\pi^{(N)}(\boldsymbol{x},\boldsymbol{y})&\le\sum_{\boldsymbol{u},\boldsymbol{v}}\pi^{(N-1)}(\boldsymbol{u},\boldsymbol{v})\sum_{b_N:\underline{b}_N=\boldsymbol{u}}\mathbb{P}(M(b_N,\boldsymbol{v};\boldsymbol{x},\boldsymbol{y}))\\&\le(2\lambda\varepsilon)^{-\delta_{\boldsymbol{x},\boldsymbol{y}}}\sum_{\ell,\ell'}\sum_{\boldsymbol{u},\boldsymbol{v}}P^{(N-1)}(\boldsymbol{u},B^{\ell'}(\boldsymbol{v}))\ L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x},B^{\ell}(\boldsymbol{y}))=(2\lambda\varepsilon)^{-\delta_{\boldsymbol{x},\boldsymbol{y}}}\sum_{\ell}P^{(N)}(\boldsymbol{x},B^{\ell}(\boldsymbol{y})).\end{aligned}$$

This advances the induction, and hence completes the proof of (4.23), assuming that (4.26) holds.

It thus remains to prove (4.26). We only consider the case of $\boldsymbol{u}=\boldsymbol{v}$ and $\boldsymbol{x}=\boldsymbol{y}$, since it explains why the factor $2\lambda\varepsilon$ is in the definition of the diagram functions. The case of $\boldsymbol{u}=\boldsymbol{v}$ and $\boldsymbol{x}\ne\boldsymbol{y}$ can be proved similarly to (4.23) for $n=0$, and the proof of the remaining case is simpler because extracting the factors of ε at $\boldsymbol{u}\ne\boldsymbol{v}$ is unnecessary. Let $\boldsymbol{u}=\boldsymbol{o}$ in (4.26), by translation invariance. Then the left-hand side of

(4.26) equals the rightmost expression in (4.21), except for the condition that b has to be a spatial bond. By the same observation at $\boldsymbol{o}$ as in (4.21), we obtain

$$\begin{aligned}\sum_{b:\underline{b}=\boldsymbol{o}} \mathbb{P}(M(b,\boldsymbol{o};\boldsymbol{x},\boldsymbol{x})) &= \sum_{\substack{b \text{ spatial:}\\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b,\boldsymbol{o};\boldsymbol{x},\boldsymbol{x})) + \mathbb{P}(M((\boldsymbol{o},\boldsymbol{o}_+),\boldsymbol{o};\boldsymbol{x},\boldsymbol{x}))\\ &\le \sum_{\substack{b \text{ spatial:}\\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b,\boldsymbol{o};\boldsymbol{x},\boldsymbol{x})) + \sum_{\substack{b \text{ spatial:}\\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(\{(\boldsymbol{o},\boldsymbol{o}_+)\to\boldsymbol{x}\}\circ\{b\to\boldsymbol{x}\})\\ &\le 2\sum_{\substack{b \text{ spatial:}\\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b,\boldsymbol{o};\boldsymbol{x},\boldsymbol{x})) \le 2\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x}). \end{aligned} \tag{4.27}$$

This completes the proof of (4.26), and hence the proof of the first inequality in (4.1).

To prove the second inequality in (4.1), we recall that $\Pi^{(N,n)}(\boldsymbol{x}) = \Pi^{(N;n,1)}(\boldsymbol{x}) + \Pi^{(N;n,2)}(\boldsymbol{x})$, where the first and second terms are the contributions to (3.13) from $b = b_n$ and from $b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]$, respectively. There are two n^{th} admissible lines terminating at $\underline{b}_{n+1}$, one from $\underline{b}_n$ and the other from some vertex $\boldsymbol{w}$. We can bound $\Pi^{(N;n,1)}(\boldsymbol{x})$ by $P^{(N)}(\boldsymbol{x})$ with the nonzero admissible line from $\underline{b}_n$, say $(p * \tau)(\underline{b}_{n+1} - \underline{b}_n)$, replaced by $(\lambda\varepsilon D * \tau)(\underline{b}_{n+1} - \underline{b}_n)$; if $\boldsymbol{w} = \underline{b}_n$, we replace the factor p in one of the two admissible lines by $\lambda\varepsilon D$ as above, and add both contributions. For $\Pi^{(N;n,2)}(\boldsymbol{x})$, we use (4.11) to obtain the bound $\sum_\ell \sum_{b\neq b_n} P^{(N)}(\boldsymbol{x}, B^\ell_{\mathtt{spat}}(\underline{b}))$, where $\sum_\ell$ is the sum over the n^{th} admissible lines. Together with the above bound on $\Pi^{(N;n,1)}(\boldsymbol{x})$, we obtain (4.20). This completes the proof the second inequality in (4.1), and hence the proof of Lemma 4.1. ☐

4.2 Estimate of the diagram functions above four dimensions

In this section, we prove Lemma 4.2(i) by using the inductive construction of $P^{(N)}(\boldsymbol{x})$ and $\tilde{P}^{(N;n)}(\boldsymbol{x})$ in (4.19)–(4.20) and the following three lemmas.

Lemma 4.4. *Assume* (2.24) *for $s \le t$ and $\lambda \in I_n$. Then there exists $C < \infty$ depending on K, but not on ε, such that the following bounds hold for $s \le t$, $q = 0, 2$ and for that λ:*

$$\sum_x |x|^q\,(\tau_s * D)(x) \le C\sigma^q(1+s)^{q/2}, \tag{4.28}$$

$$\sup_x |x|^q\,(\tau_s * D)(x) \le \frac{C\sigma^q\beta}{(1+s)^{(d-q)/2}}, \tag{4.29}$$

$$\sup_x |x|^q\,\tau_s(x) \le e^{-s}\delta_{q,0} + \frac{C\sigma^q\beta}{(1+s)^{(d-q)/2}}. \tag{4.30}$$

Lemma 4.5. *Assume* (2.24) *for $s \le t$. Let $f_t(x)$ be a diagram function that satisfies $\sum_x f_t(x) \le F(t)$ by assigning l_1 or l_∞ norm to each diagram line and using* (2.24) *to estimate those norms. Then there exists $c < \infty$ depending only on d and K in* (2.24) *such that $\sum_x f_t(x, B^\ell(s)) \le c\varepsilon F(t)$ holds for every $s \le t$ and every admissible line ℓ.*

Lemma 4.6. *Let $a, b \in \mathbb{R}$, and let κ be a positive number if a or b is 2, and zero otherwise. Then there exists $c = c(a, b, \kappa) < \infty$ such that*

$$\sum_{s_1=0}^{t}{}^{\bullet} \frac{\varepsilon}{(1+s_1)^a} \sum_{s_2=t-s_1}^{t}{}^{\bullet} \frac{\varepsilon}{(1+s_2)^b} \le \frac{c}{(1+t)^{a\wedge b\wedge(a+b-2)-\kappa}}. \tag{4.31}$$

Lemmas 4.4 and 4.5 correspond respectively to Lemmas 4.3 and 4.6(a) in [18], and Lemma 4.6 corresponds to [17, Lemma 3.2]. The result of applying Lemma 4.6 is similar to [18, Lemma 4.6(b)]. We will use Lemmas 4.4–4.5 again in Sections 4.3–4.4, and Lemma 4.6 in Section 4.3 and Appendix A.

Proof of Lemma 4.2(i) assuming Lemmas 4.4–4.6. For (4.2), we only consider $q = 0$; the other cases in (4.2) are proved along the same argument in the last paragraph of [18, Section 4.2]. For $s' \le s < t$, we use Lemma 4.4 to obtain

$$\sup_{u,v} \sum_x L((u,s),(v,s');(x,t)) \le \frac{c\beta\varepsilon}{(1+t-s')^{d/2}}. \tag{4.32}$$

Since $P^{(0)}_t(x) = \lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};(x,t))$ for $t \ge 2\varepsilon$, this implies (4.2) with $q = N = 0$.

By Lemma 4.5, we obtain $\sum_\ell \sum_x P^{(0)}_t(x, B^\ell(s)) \le c\varepsilon[\delta_{0,t} + \varepsilon^2\beta/(1+t)^{d/2}]$, where $\delta_{0,t}$ is the contribution from the first term in (4.16). Therefore, $\sum_x P^{(1)}_t(x)$ satisfies (4.2), and $\sum_\ell \sum_x P^{(1)}_t(x, B^\ell(s))$ is bounded by $c\varepsilon^3 C\beta/(1+t)^{d/2}$. This initializes the inductive proof of (4.2) for $q = 0$. Suppose that $\sum_x P^{(N-1)}_t(x, B^\ell(s))$ is bounded by $c\varepsilon^3(C\beta)^{N-1}/(1+t)^{d/2}$. Then by (4.19),

$$\sum_x P^{(N)}_t(x) \le \sum_{s,s'}{}^{\bullet} \Bigg[\sum_\ell \sum_{u,v} P^{(N-1)}((u,s), B^\ell(v,s'))\Bigg]\Bigg[\sup_{u,v} \sum_x L((u,s),(v,s');(x,t))\Bigg], \tag{4.33}$$

where $\sum_\ell$ is the sum over the $(N-1)^{\text{st}}$ admissible lines in $P^{(N-1)}_s(u)$. By (4.32) and Lemma 4.6, the right-hand side of (4.33) is bounded by $\varepsilon^2(C\beta)^N/(1+t)^{d/2}$ if C is sufficiently large. We note that ε^2 is used in applying Lemma 4.6, in order to approximate $\varepsilon\sum^{\bullet}_{s\in\varepsilon\mathbb{Z}_+}$ by the Riemann sum. This completes the proof of (4.2).

To prove (4.3), we first use Lemma 4.5 to obtain $\sum_x P^{(N)}_t(x, B^\ell_{\mathtt{spat}}(s)) \le c\varepsilon^3(C\beta)^N/(1+t)^{d/2}$, where ℓ is an n^{th} admissible line. Then we sum the bound over $s \in \varepsilon\mathbb{Z}_+$ with $0 \le s \le t$ to obtain the desired bound in (4.3). Note that ε is used in an approximation by the Riemann sum. This completes the proof. ☐

Proof of Lemma 4.4. The inequality (4.28) easily follows from (2.24) and the properties of D. To prove (4.29)–(4.30), we use

$$(\tau_s * D)(x) \le (1-\varepsilon)^{s/\varepsilon} D(x) + \lambda\varepsilon \sum_{j=1}^{s/\varepsilon} (1-\varepsilon)^{j-1} (D * \tau_{s-j\varepsilon} * D)(x), \tag{4.34}$$

and

$$\tau_s(x) \le (1-\varepsilon)^{s/\varepsilon}\delta_{o,x} + \lambda s(1-\varepsilon)^{s/\varepsilon-1} D(x) + \lambda^2\varepsilon \sum_{j=2}^{s/\varepsilon} (j-1)\varepsilon\,(1-\varepsilon)^{j-2}(D * \tau_{s-j\varepsilon} * D)(x). \tag{4.35}$$

where $(1-\varepsilon)^k$ is the probability that $(o,0)$ is connected to $(o,k\varepsilon)$ along the temporal axis. By the trivial inequality $(1-\varepsilon)^{s/\varepsilon} \le e^{-s}$ and the properties of D, both the first term in (4.34) and the second term in (4.35), multiplied by $|x|^q$, are bounded by $c\sigma^q\beta/(1+s)^{(d-q)/2}$ for any $x \in \mathbb{Z}^d$. We fix $r \in (0,1)$ and use (2.24) to bound the second term in (4.34) by

$$\varepsilon \sum_{j=1}^{rs/\varepsilon} \frac{c\beta e^{-j\varepsilon}}{(1+s-j\varepsilon)^{d/2}} + \varepsilon \sum_{j=rs/\varepsilon}^{s/\varepsilon} \frac{c\beta e^{-j\varepsilon}}{(1+s-j\varepsilon)^{d/2}} \le \frac{c'\beta}{(1+s)^{d/2}}.$$

By the same argument, the third term in (4.35) can be bounded by $c\beta/(1+s)^{d/2}$. This completes the proof of (4.29)–(4.30) for $q = 0$.

For $|x|^2$ times the second term in (4.34) or the third term in (4.35), we have

$$\begin{aligned} |x|^2(D * \tau_s * D)(x) &\le 2\sum_y (|y|^2 + |x-y|^2)\,(D * \tau_{s/2})(y)\,(\tau_{s/2} * D)(x-y) \\ &\le 4\Big[\sup_x (\tau_{s/2} * D)(x)\Big] \sum_x |x|^2 (\tau_{s/2} * D)(x). \end{aligned}$$

We obtain (4.29)–(4.30) for $q = 2$ by applying (4.34) to $(\tau_{s/2} * D)(x)$ and using (2.24), as above. This completes the proof of Lemma 4.4. ☐

Proof of Lemma 4.5. By definition, the case of $s = t$ is trivial. For $s < t$, Construction $B^{\ell}(s)$ replaces the diagram line ℓ, say $\tau_t(x)$, by $\lambda\varepsilon(\tau_s * D * \tau_{t-s-\varepsilon} + \tau_s * \tau_{t-s})(x)$. By Lemma 4.4, we obtain

$$\lambda\varepsilon\|\tau_s * D * \tau_{t-s-\varepsilon}\|_1 = \lambda\varepsilon\|\tau_s\|_1\,\|\tau_{t-s-\varepsilon}\|_1 \le c\varepsilon K^2,$$

$$\lambda\varepsilon\|\tau_s * D * \tau_{t-s-\varepsilon}\|_\infty \le \lambda\varepsilon\|\tau_{s\vee(t-s-\varepsilon)}\|_\infty\,\|\tau_{s\wedge(t-s-\varepsilon)}\|_1 \le \frac{c\varepsilon K^2}{(1+t)^{d/2}},$$

where we used $s \vee (t-s-\varepsilon) \ge t/2$. The contribution from the other term can be estimated in the same way. Therefore, the effect of Construction $B^{\ell}(s)$ is to obtain, at worst, an additional constant $c\varepsilon K$ in a bound. This completes the proof. □

Proof of Lemma 4.6. We prove Lemma 4.6 for the case of $a, b \ge 0$ and for the other case that a or b is negative separately.

First, we consider the case of $a, b \ge 0$. Decomposing the sum over s_1 into $\sum^{\bullet}_{0\le s_1\le t/2}$ and $\sum^{\bullet}_{t/2<s_1\le t}$, and using $s_2 \ge t - s_1 \ge t/2$ in the former sum, we can bound the left-hand side of (4.31) by

$$\frac{c}{(1+t)^b}\sum_{s_1=0}^{t/2}{}^{\bullet}\,\frac{\varepsilon}{(1+s_1)^{a-1}} + \frac{c}{(1+t)^a}\sum_{s_1=t/2}^{t}{}^{\bullet}\,\sum_{s_2=t-s_1}^{t}{}^{\bullet}\,\frac{\varepsilon^2}{(1+s_2)^b}. \tag{4.36}$$

The double sum is

$$\sum_{s_2=0}^{t}{}^{\bullet}\,\frac{\varepsilon}{(1+s_2)^b}\sum_{s_1=t/2\vee(t-s_2)}^{t}{}^{\bullet}\,\varepsilon \;\le\; \sum_{s_2=0}^{t}\frac{\varepsilon}{(1+s_2)^{b-1}} \;\le\; \frac{c}{(1+t)^{(b-2)\wedge 0-\kappa}},$$

where $\kappa > 0$ if $b = 2$, otherwise $\kappa = 0$. Also, the sum in the first term of (4.36) is bounded by $c/(1+t)^{(a-2)\wedge 0-\kappa}$, where $\kappa > 0$ if $a = 2$, and zero otherwise.

Next, we consider the case that a or b is negative. Due to the symmetry of the left-hand side of (4.31) in terms of s_1, s_2, we suppose $b < a \wedge 0$. Then we use the trivial inequality $(1+s_2)^{-b} \le (1+t)^{-b}$. The remaining quantity $\varepsilon\sum^{\bullet}_{0\le s_1\le t} s_1/(1+s_1)^a$ is bounded by $c/(1+t)^{(a-2)\wedge 0-\kappa}$, where $\kappa > 0$ if $a = 2$, and zero otherwise. This completes the proof. □

4.3 Estimate of the diagram functions at and below four dimensions

In this section, we prove the bounds in Lemma 4.2(ii) on the diagram functions for $d \le 4$. As in the proof of Lemma 4.2(i) in Section 4.2, we will use the inductive construction of $P^{(N)}(\boldsymbol{x})$ and $\tilde{P}^{(N;n)}(\boldsymbol{x})$ in (4.19)–(4.20) and Lemmas 4.4–4.6, with σ and β in Lemma 4.4 replaced by σ_T and β_T, respectively.

Proof of Lemma 4.2(ii). The proof is almost the same as that of Lemma 4.2(i). The only difference arises when we apply Lemma 4.6. Let $N \ge 1$ and suppose that the quantity in the first brackets in (4.33) is bounded by $c\varepsilon^3 C\beta_T(C\hat{\beta}_T)^{N-1}/(1+t)^{d/2}$, where $\hat{\beta}_T = \beta_1 T^{-\mu}$ with $\mu \in (0, \alpha)$. Then by Lemma 4.6 and (4.32) with β replaced by β_T, the right-hand side of (4.33) for $d \le 4$ is bounded by

$$\frac{cC\beta_T^2(C\hat{\beta}_T)^{N-1}\varepsilon^2}{(1+t)^{\frac{d}{2}\wedge(d-2)-\kappa}} = \frac{C\beta_T(C\hat{\beta}_T)^{N-1}\varepsilon^2}{(1+t)^{d/2}}\,c\beta_1 T^{-bd}(1+t)^{(4-d)/2+\kappa}, \tag{4.37}$$

where κ is positive (but arbitrarily small) if $d = 4$, and zero otherwise. Since $t \le T\log T$ and $\mu < \alpha$, the last factor is bounded by $C\hat{\beta}_T$ if C is sufficiently large, and thus we obtain (4.4) with $q = 0$.

For (4.4) with $q = 2$ and (4.5), we use Lemma 4.6 as in (4.37), with $(a, b) = (\frac{d}{2}, \frac{d-2}{2})$, to obtain

$$\frac{\beta_T}{(1+t)^{(d-2)/2\wedge(d-3)-\kappa}} = \frac{\beta_1 T^{-bd}(1+t)^{(4-d)/2+\kappa}}{(1+t)^{(d-2)/2}} \le \frac{c\hat{\beta}_T}{(1+t)^{(d-2)/2}}.$$

To prove (4.4) for $q = 4$, we apply Lemma 4.6 as above, with $(a, b) = (\frac{d}{2}, \frac{d-4}{2})$ and $(\frac{d-2}{2}, \frac{d-2}{2})$. This completes the proof. □

4.4 Finite containment

We define $\mathbf{C}_r(\boldsymbol{z}) = \{w \in \mathbb{Z}^d : \boldsymbol{z} \to (w,r)\}$ and $\mathbf{C}_{[s,t]}(\boldsymbol{z}) = \bigcup_{r=s}^t \mathbf{C}_r(\boldsymbol{z})$, and simply write $\mathbf{C}_r = \mathbf{C}_r(\boldsymbol{o})$ and $\mathbf{C}_{[s,t]} = \mathbf{C}_{[s,t]}(\boldsymbol{o})$. In this section, we prove that $\pi_t^{(N)}(x)$ can be approximated by

$$\pi_t^{(N)}(x \,|\, R) = \sum_{\vec{b}_N} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x});\ \mathbf{C}_{[0,t]} \subset \Box_R\big),$$

where $\Box_R = [-R,R]^d \cap \mathbb{Z}^d$. More precisely, we prove below

$$\pi_t^{(N)}(x) = \pi_t^{(N)}(x \,|\, R) + o(1)\,\varepsilon^2, \tag{4.38}$$

where $o(1)$ is independent of ε and decays as $R \to \infty$, by assuming (2.24) and using estimates for the diagram functions in Sections 4.1–4.3. This result will be used to perform the continuum limit in Section 6.

Proof. First, we note that

$$\begin{aligned} 0 \le \pi_t^{(N)}(x) - \pi_t^{(N)}(x \,|\, R) &\le \sum_{\vec{b}_N} \sideset{}{^\bullet}\sum_{s=\varepsilon}^{t} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x});\ \mathbf{C}_{[0,s-\varepsilon]} \subset \Box_R;\ \mathbf{C}_s \not\subset \Box_R\big) \\ &\le \sum_{\vec{b}_N} \sideset{}{^\bullet}\sum_{s=\varepsilon}^{t} \sum_{\substack{u \in \Box_R \\ v \notin \Box_R}} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x});\ \boldsymbol{o} \to ((u,s-\varepsilon),(v,s))\big). \end{aligned}$$

The event in the rightmost expression is included in the union of two events: (i) $\boldsymbol{o} \to ((u,s-\varepsilon),(v,s)) \to \boldsymbol{x}$, or (ii) there is a vertex $(w,r) \in \Box_R \times [0,s)$ such that $\boldsymbol{o} \to (w,r) \to \boldsymbol{x}$ and that $(w,r) \to ((u,s-\varepsilon),(v,s))$ disjointly from $\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x})$. The contribution from the case (i) and from the case (ii) with $(w,r) = (u,s-\varepsilon)$ can be bounded by $\sum_\ell \sideset{}{^\bullet}\sum_s \sum_{v \notin \Box_R} P^{(N)}(\boldsymbol{x}, B^\ell(v,s))$, where $\sum_\ell$ is the sum over $n \in \{1,\dots,N\}$ of the sum over the n^{th} admissible lines, and where we modified Construction $B^\ell_{\mathtt{spat}}(v,s)$ by fixing the second endpoint (v,s), instead of fixing the first endpoint $(u,s-\varepsilon)$ as defined in (4.17). This bound, divided by ε^2, decays as $R \to \infty$ uniformly in ε, because, by Lemma 4.5, the sum over $x \in \mathbb{Z}^d$ of the unrestricted sum $\sum_\ell \sideset{}{^\bullet}\sum_s P^{(N)}(\boldsymbol{x}, B^\ell(s))$ is bounded by $c\varepsilon^2 t/(1+t)^{d/2}$, where c is a finite constant depending only on β and N (and T if $d \le 4$).

The contribution from the case (ii) with $(w,r) \ne (u,s-\varepsilon)$ can be bounded by

$$\sum_\ell \sideset{}{^\bullet}\sum_{s=\varepsilon}^{t} \sideset{}{^\bullet}\sum_{r=0}^{s-\varepsilon} \sum_{w \in \Box_R} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \Box_R} (\tau_{s-\varepsilon-r} * \lambda\varepsilon D)(v-w), \tag{4.39}$$

where we relabelled the second endpoint of the spatial bond in (4.18) as (w,r). For $w \in \Box_{R/2}$, we use

$$\begin{aligned} \sum_{w \in \Box_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \Box_R} (\tau_{s-\varepsilon-r} * D)(v-w) &\le P^{(N)}(\boldsymbol{x}, B^\ell(r)) \sup_{w \in \Box_{R/2}} \sum_{v \notin \Box_R} (\tau_{s-\varepsilon-r} * D)(v-w) \\ &\le P^{(N)}(\boldsymbol{x}, B^\ell(r)) \sum_{z \notin \Box_{R/2}} (\tau_{s-\varepsilon-r} * D)(z), \end{aligned}$$

and for $w \in \Box_R \setminus \Box_{R/2}$, we use

$$\sum_{w \notin \Box_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \Box_R} (\tau_{s-\varepsilon-r} * D)(v-w) \le \|\tau_{s-\varepsilon-r} * D\|_1 \sum_{w \notin \Box_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)).$$

Since $\|\tau_{s-\varepsilon-r} * D\|_1 = \|\tau_{s-\varepsilon-r}\|_1 \le K$ and, by Lemma 4.5, $P^{(N)}(\boldsymbol{x}, B^\ell(r)) \le c\varepsilon^3$, where c depends only on β, N and t (and T if $d \le 4$), the expression (4.39) is $o(1)\,\varepsilon^2$, similarly to the case (i). This completes the proof of (4.38). $\Box$

5 Inductive argument

In this section, we prove Proposition 2.1 by applying the inductive method of [17] for self-avoiding walk in $\mathbb{Z}^d$ and for oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$, to the recursion equation (5.1) for oriented percolation in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$. To consider the case of $\varepsilon \ll 1$, we will modify the induction hypotheses to incorporate the dependence on ε. We expect that a similar method could be used for *continuous-time* weakly self-avoiding walk above its upper critical dimension.

First we consider $d > 4$ in Sections 5.1–5.3. In Section 5.1, we give the modified version of the induction hypotheses. In Section 5.2, we show several consequences of the induction hypotheses, mainly the bounds (2.24). In Section 5.3, we prove Proposition 2.1. We complete this section by proving the results for $d \leq 4$ in Section 5.4. Since the same strategy applies to the lower-dimensional case, we only discuss the necessary changes.

5.1 Induction hypotheses

5.1.1 General assumptions

In Section 3.1, we derived the recursion equation (3.4) for the two-point function. Taking the Fourier transform with respect to the spatial component, we obtain (2.28), i.e.,

$$\hat{\tau}_{t+\varepsilon}(k) = \sum_{s=0}^{t}{}^{\bullet}\, \hat{\pi}_s(k)\, \hat{p}_\varepsilon(k)\, \hat{\tau}_{t-s}(k) + \hat{\pi}_{t+\varepsilon}(k), \tag{5.1}$$

where $\hat{p}_\varepsilon(k) = 1 - \varepsilon + \lambda\varepsilon\hat{D}(k)$ with $\hat{D}(k) = \sum_x D(x)\, e^{ik\cdot x}$. The probability distribution $D : \mathbb{Z}^d \mapsto [0,1]$ satisfies the assumptions in Section 1.2. In addition, we assume that there exists $\eta > 0$ such that

$$a(k) \equiv 1 - \hat{D}(k) \begin{cases} \asymp L^2|k|^2, & \text{if } \|k\|_\infty \leq L^{-1}, \\ > \eta, & \text{if } \|k\|_\infty > L^{-1}, \end{cases} \qquad a(k) < 2 - \eta \quad \forall k \in [-\pi, \pi]^d,$$

where the notation $\asymp$ means that the left-hand side is bounded from above and below by a multiple of the right-hand side. These assumptions correspond to Assumption D in [17].

If we replace t by $n\varepsilon$, and write

$$f_n(k) = \hat{\tau}_t(k), \qquad e_n(k) = \hat{\pi}_t(k), \qquad g_{n+1}(k) = \hat{\pi}_t(k)\, \hat{p}_\varepsilon(k), \tag{5.2}$$

then (5.1) equals

$$f_{n+1}(k) = \sum_{m=0}^{n} g_{m+1}(k)\, f_{n-m}(k) + e_{n+1}(k), \tag{5.3}$$

with $f_0(k) = 1$. This equals the recursion relation (1.1) investigated in [17], where the dependence on λ is left implicit. The only difference resides in the fact that

$$f_1(k) = g_1(k) = 1 - \varepsilon + \lambda\varepsilon\hat{D}(k), \tag{5.4}$$

whereas in (1.1) of [17], it was assumed that $f_1(k) = g_1(k) = \lambda\hat{D}(k)$. This change leads to a minor modification of the induction hypotheses in [17]. Further technical changes are explained in Section 5.1.2.

5.1.2 Statement of the induction hypotheses

Fix γ, δ and ρ according to

$$-(2+\rho) < 0 < \tfrac{d}{2} - (2+\rho) < \gamma < \gamma + \delta < 1 \wedge \Delta \wedge \tfrac{d-4}{2}. \tag{5.5}$$

We define λ_n recursively by $\lambda_0 = \lambda_1 = 1$ and, for $n \geq 2$,

$$\lambda_n = 1 - \frac{1}{\varepsilon} \sum_{l=2}^{n} g_l(0; \lambda_{n-1}), \tag{5.6}$$

where we explicitly write the dependence of $g_l(0)$ on λ_{n-1}. Let

$$I_n = \lambda_n + \frac{K_1 \beta}{(1 + n\varepsilon)^{(d-2)/2}} [-1, 1], \tag{5.7}$$

and define $v_n = v_n(\lambda)$ by $v_0 = v_1 = \lambda$ and, for $n \geq 2$,

$$v_n = \frac{\lambda - \frac{1}{\sigma^2 \varepsilon} \sum_{l=2}^{n} \nabla^2 g_l(0)}{1 + \sum_{l=2}^{n} (l-1)\, g_l(0)}. \tag{5.8}$$

Let $K_1, \ldots, K_5$ be some positive and finite constants that are independent of β and ε, and are related by

$$K_3 \gg K_1 > K_4 \gg 1, \qquad K_2 \gg K_1, K_4, \qquad K_5 \gg K_4. \tag{5.9}$$

The induction hypotheses are that the following (H1)–(H4) hold for all $\lambda \in I_n$ and $m = 1, \ldots, n$.

(H1)–(H2)

$$|\lambda_m - \lambda_{m-1}| \leq \frac{\varepsilon K_1 \beta}{(1 + m\varepsilon)^{d/2}}, \qquad |v_m - v_{m-1}| \leq \frac{\varepsilon K_2 \beta}{(1 + m\varepsilon)^{(d-2)/2}}, \tag{5.10}$$

(H3) For $k \in \mathcal{A}_m \equiv \{k : a(k) \leq \gamma \frac{\log(2+m\varepsilon)}{1+m\varepsilon}\}$, $f_m(k)$ can be written in the form

$$f_m(k) = \prod_{l=1}^{m} [1 - \varepsilon v_l\, a(k) + \varepsilon\, r_l(k)], \tag{5.11}$$

where $r_l(k)$ obeys the bounds

$$|r_l(0)| \leq \frac{K_3 \beta}{(1 + l\varepsilon)^{(d-2)/2}}, \qquad |r_l(k) - r_l(0)| \leq \frac{K_3 \beta\, a(k)}{(1 + l\varepsilon)^{\delta}}. \tag{5.12}$$

(H4) For $k \notin \mathcal{A}_m$, $f_m(k)$ obeys the bounds

$$|f_m(k)| \leq \frac{K_4\, a(k)^{-2-\rho}}{(1 + m\varepsilon)^{d/2}}, \qquad |f_m(k) - f_{m-1}(k)| \leq \frac{\varepsilon K_5\, a(k)^{-1-\rho}}{(1 + m\varepsilon)^{d/2}}. \tag{5.13}$$

Instead of (5.11), we can write $f_m(k)$ as

$$f_m(k) = f_m(0) \prod_{l=1}^{m} [1 - \varepsilon v_l\, a(k) + \varepsilon\, s_l(k)], \tag{5.14}$$

where

$$f_m(0) = \prod_{l=1}^{m} [1 + \varepsilon\, r_l(0)], \qquad s_l(k) = \frac{\varepsilon v_l\, r_l(0)\, a(k) + [r_l(k) - r_l(0)]}{1 + \varepsilon\, r_l(0)}. \tag{5.15}$$

The induction hypothesis (H3) implies

$$|s_l(k)| \leq \frac{\varepsilon v_l |r_l(0)| a(k) + |r_l(k) - r_l(0)|}{1 - \varepsilon |r_l(0)|} \leq \frac{(1 + \varepsilon v_l) K_3 \beta\, a(k)}{(1 - \varepsilon K_3 \beta)(1 + l\varepsilon)^{\delta}}. \tag{5.16}$$

In some cases, we will use (5.14)–(5.16), instead of (5.11)–(5.12). Moreover, by (5.14) and spatial symmetry, we obtain

$$\nabla^2 f_m(0) = f_m(0)\, \varepsilon \sum_{l=1}^{m} \big[-v_l \sigma^2 + \nabla^2 s_l(0)\big]. \tag{5.17}$$

The advancement of the induction hypotheses is a small modification of that in [17], which we add to keep this paper self-contained. The advancement is deferred to Appendix A.

5.2 Consequences of the induction hypotheses

We assume $\beta \ll 1$ and use c to denote a positive and finite constant that may depend on d, γ, δ, ρ, but not on $K_i, k, n, \beta, \varepsilon$. The value of c may change from line to line.

The following four lemmas, corresponding respectively to [17, Lemmas 2.1, 2.2, 2.4, 2.3], are consequences of the induction hypotheses (H1)–(H4) for $d > 4$.

Lemma 5.1. *Assume (H1) for $m = 1, \dots, n$. Then $I_0 \supset I_1 \supset \cdots \supset I_n$.*

Lemma 5.2. *Let $\lambda \in I_n$ and assume (H2)–(H3) for $m = 1, \dots, n$. For $k \in \mathcal{A}_m$,*

$$|f_m(k)| \le e^{cK_3\beta}\, e^{-m\varepsilon[1-c\,(K_1+K_2+K_3)\beta]\,a(k)}.$$

Lemma 5.3. *Let $\lambda \in I_n$ and assume (H2)–(H3) for $m = 1, \dots, n$. Then*

$$|\nabla^2 f_m(0)| \le [1 + c\,(K_1 + K_2 + K_3)\beta]\sigma^2 m\varepsilon.$$

Lemma 5.4. *Let $\lambda \in I_n$ and assume (H2)–(H4) for $m = 1, \dots, n$. Then*

$$\|\hat{D}^2 f_m\|_1 \le \frac{c\,(1 + K_4)\beta}{(1 + m\varepsilon)^{d/2}}.$$

The bounds (2.24) for $m = 1, \dots, n$ follow from Lemmas 5.2–5.4, with $K \gg K_4$.

By Lemma 5.1, if $\lambda \in I_m$ for some $m \ge 0$, then $\lambda \in I_0$ and hence, by (5.7),

$$|\lambda - 1| \le K_1\beta. \tag{5.18}$$

It also follows that $I_\infty = \bigcap_{m=0}^\infty I_m$ is a singleton λ_∞. As discussed in [17, Theorem 1.2], we obtain $\lambda_\infty = \lambda_{\mathrm{c}}^{(\varepsilon)}$. Moreover, it follows from the second inequality of (5.10) that, for $\lambda \in I_m$,

$$|v_m - 1| \le \sum_{l=1}^m |v_l - v_{l-1}| + |v_0 - 1| \le \sum_{l=1}^m \frac{\varepsilon K_2\beta}{(1 + l\varepsilon)^{(d-2)/2}} + |\lambda - 1| \le (cK_2 + K_1)\beta. \tag{5.19}$$

We note that the factor ε in the numerator is necessary to approximate the sum by the Riemann sum when ε is small. The factors of ε in (5.10)–(5.13) are incorporated for the same reason. The proofs of Lemmas 5.1–5.4 are almost identical to those of [17, Lemmas 2.1–2.4], and are deferred to Appendix A.

5.3 Proof of Proposition 2.1

Fix $\lambda = \lambda_{\mathrm{c}}^{(\varepsilon)}$, so that the induction hypotheses (H1)–(H4) and Lemmas 5.1–5.4 hold for all $m \in \mathbb{N}$. From now on, we suppress the dependence on ε and write $\lambda_{\mathrm{c}} = \lambda_{\mathrm{c}}^{(\varepsilon)}$, $A = A^{(\varepsilon)}$ and $v = v^{(\varepsilon)}$. We also write $O(\,\cdot\,)$ for an ε-independent term, which may depend on L.

By Lemma 5.2, $A \equiv \lim_{n\to\infty} f_n(0)$ exists and is $1 + O(\beta)$. Moreover, by (5.11)–(5.12),

$$|f_n(0) - A| = \prod_{l=1}^n [1 + \varepsilon r_l(0)] \left|1 - \prod_{l=n+1}^\infty [1 + \varepsilon r_l(0)]\right| \le \frac{cK_3\beta}{(1 + n\varepsilon)^{(d-4)/2}}. \tag{5.20}$$

Also, by (5.19), $v \equiv \lim_{n\to\infty} v_n$ exists and is $1 + O(\beta)$. Moreover, by (5.10),

$$|v_n - v| \le \sum_{l=n+1}^\infty |v_l - v_{l-1}| \le \frac{cK_2\beta}{(1 + n\varepsilon)^{(d-4)/2}}. \tag{5.21}$$

Let $t = n\varepsilon$ and $\frac{k}{\sqrt{v\sigma^2 t}} \in \mathcal{A}_n$. By (1.2), we have $a(k) = \frac{\sigma^2}{2d}|k|^2 + O(|k|^{2+2\Delta})$, where constants in the error term may depend on L. Together with (5.14), (5.16), (5.20)–(5.21) and $\delta < 1 \wedge \frac{d-4}{2}$, we obtain

$$\begin{aligned} f_n(\tfrac{k}{\sqrt{v\sigma^2 t}}) &= [A + c\beta(1+t)^{-(d-4)/2}] \prod_{l=1}^{t/\varepsilon} \Big[1 - \varepsilon\big[v + c\beta(1+l\varepsilon)^{-\delta}\big]\, a(\tfrac{k}{\sqrt{v\sigma^2 t}})\Big] \\ &= Ae^{-\frac{|k|^2}{2d}}\big[1 + O(\beta(1+t)^{-(d-4)/2}) + O(|k|^{2+2\Delta}t^{-\Delta}) + O(|k|^2(1+t)^{-\delta})\big]. \end{aligned} \tag{5.22}$$

Since $\big(\frac{|k|^2}{t}\big)^{\Delta} \le c\big[\frac{\log(2+t)}{1+t}\big]^{\Delta} \le c\,(1+t)^{-\delta}$ for $\Delta > \delta$, this completes the proof of (2.4).

Let $e_1, \ldots, e_d$ denote the standard basis vectors in $\mathbb{R}^d$. Then, by (5.16),

$$|\nabla^2 s_l(0)| = \left|\sum_{i=1}^{d} \lim_{h\to 0} \frac{s_l(he_i) - s_l(0)}{h^2}\right| \le \frac{cK_3\beta}{(1+l\varepsilon)^{\delta}} \left|\sum_{i=1}^{d} \lim_{h\to 0} \frac{a(he_i)}{h^2}\right| = \frac{cK_3\sigma^2\beta}{(1+l\varepsilon)^{\delta}}. \tag{5.23}$$

Since $\delta < 1$, it follows from (5.17), (5.20)–(5.21) and (5.23) that

$$-\frac{\nabla^2 f_n(0)}{f_n(0)} = v\sigma^2 t[1 + O(\beta(1+t)^{-\delta})],$$

which is (2.5).

The upper bound in (2.6) is an immediate consequence of (4.30). The lower bound is proved as in [17, Corollary 1.4], with the help of (5.22). This completes the proof of Proposition 2.1. ☐

Finally, we derive the expressions (2.29)–(2.30) for λ_c, A and v. Recall (5.2). The expressions for λ_c and v immediately follow from (5.6), (5.8) and the fact that $\lambda_c = \lambda_\infty$. To derive the expression for $A = \lim_{n\to\infty} f_n(0)$, we follow the same strategy as in [17, p.424]. Let $F_N = \varepsilon\sum_{n=0}^{N} f_n(0)$, which can be approximated by $AN\varepsilon$ as $N \to \infty$. Summing the recursion equation (5.3) with $k = 0$, multiplied by ε, over $n = 0, \ldots, N-1$, and using $f_0(0) = 1$, $e_1(0) = 0$, $g_1(0) = 1 + (\lambda_c - 1)\varepsilon$ and the expression for λ_c, we have

$$\begin{aligned} F_N &= g_1(0)\,F_{N-1} + \sum_{n=2}^{N} g_n(0)\,F_{N-n} + \varepsilon\sum_{n=2}^{N} e_n(0) + \varepsilon \\ &= F_{N-1} - \sum_{n=2}^{\infty} g_n(0)\,F_{N-1} + \sum_{n=2}^{N} g_n(0)\,F_{N-n} + \varepsilon\sum_{n=2}^{N} e_n(0) + \varepsilon. \end{aligned}$$

Taking the limit $N \to \infty$ of $f_N(0) = \frac{F_N - F_{N-1}}{\varepsilon}$, with the help of the bound (2.25), we obtain

$$A = -A\sum_{n=2}^{\infty}(n-1)g_n(0) + \sum_{n=2}^{\infty} e_n(0) + 1.$$

which gives the expression for A in (2.30). This completes the derivation of (2.29)–(2.30). ☐

5.4 Discussion on changes below and at four dimensions

In dimensions $d \le 4$, the induction analysis in Sections 5.1–5.3 no longer works as long as the infection range is fixed, and we need to incorporate the factor $L_T = L_1 T^b$ into the induction hypotheses.

Recall $\alpha = bd - \frac{4-d}{2} > 0$. Letting $\omega \in (0, 1 \wedge \alpha)$ and $\hat{\beta}'_T = \beta_1 T^{-\mu'}$ with $\mu' \in (0, \alpha - \omega)$, we replace (5.5), (5.7), (5.10) and (5.12), respectively, by

$$-(2+\rho) < 0 < \tfrac{d}{2} - (2+\rho) < \gamma < \gamma + \delta < \omega \wedge \Delta, \tag{5.24}$$

$$I_n = \lambda_n + \frac{K_1 \hat{\beta}'_T}{(1+n\varepsilon)^{1+\omega}}[-1,1], \tag{5.25}$$

$$|\lambda_m - \lambda_{m-1}| \le \frac{\varepsilon K_1 \hat{\beta}'_T}{(1+m\varepsilon)^{2+\omega}}, \qquad |v_m - v_{m-1}| \le \frac{\varepsilon K_2 \hat{\beta}'_T}{(1+m\varepsilon)^{1+\omega}}, \tag{5.26}$$

$$|r_m(0)| \le \frac{K_3 \hat{\beta}'_T}{(1+m\varepsilon)^{1+\omega}}, \qquad |r_m(k) - r_m(0)| \le \frac{K_3 \hat{\beta}'_T}{(1+m\varepsilon)^{\delta}}\, a(k). \tag{5.27}$$

The induction hypotheses are that (H1)–(H4) hold for all $\lambda \in I_n$ and $m = 1, \ldots, n$, where we assume that $n\varepsilon \le T \log T$. It suffices to prove the main statement for sufficiently small $\beta_1 > 0$, i.e., for sufficiently large initial infection ranges L_1.

We now discuss the induction hypotheses. One of the key ingredients in the induction is the fact that the intervals I_n are decreasing in $n \le (T \log T)/\varepsilon$. This implies that we can use the bounds following from (H1)–(H4) in the advancement of the induction. One would expect that one could choose $I_n = \lambda_n + K_1 \beta_T (1+n\varepsilon)^{-(d-2)/2}[-1,1]$, i.e., by simply replacing β in (5.7) by β_T. However, to obtain a decreasing sequence of I_n, it is required for the power exponent $(d-2)/2$ in the width of I_n to be greater than 1, and it is not the case when $d < 4$ (for more details, see the proof of Lemma 5.1 in Appendix A); for the marginal case $d = 4$, the power exponent is also required to be greater than 1 to obtain summability. To satisfy this requirement, we transfer some power exponent of β_T as

$$\frac{\beta_T}{(1+n\varepsilon)^{(d-2)/2}} = \frac{\beta_1 T^{-db}(1+n\varepsilon)^{(4-d)/2+\omega}}{(1+n\varepsilon)^{1+\omega}} \le \frac{\beta_1 T^{-\mu'}}{(1+n\varepsilon)^{1+\omega}},$$

for sufficiently large T, where we used $n\varepsilon \le T \log T$ and $\mu' \in (0, \alpha - \omega)$. This is the motivation of the changes in (5.25)–(5.27).

By the above changes, (5.18)–(5.19) are modified by replacing β with $\hat{\beta}'_T$. We have that λ and $v_m(\lambda)$ are both $1 + O(\hat{\beta}'_T)$ for $\lambda \in I_m$ and $m = 1, \ldots, n$ with $n\varepsilon \le T \log T$. Similarly, we replace β in Lemmas 5.2–5.3 and Lemma 5.4 by $\hat{\beta}'_T$ and β_T, respectively, although the proofs of these lemmas will remain unchanged. However, the proof of the main result does change, due to the fact that the constants A and v for $d > 4$ are replaced by 1, and the fact that there is no unique limit of $\bigcap_{m=1}^n I_m$.

Proof of the main result for $d \le 4$. Let $n\varepsilon \le T \log T$ and $\lambda \in I_n$. By (5.27), we can bound $|f_n(0) - 1|$ by

$$\left| \prod_{m=1}^n [1 + \varepsilon r_m(0)] - 1 \right| \le \varepsilon \sum_{m=1}^n |r_m(0)| \prod_{l=m+1}^n [1 + \varepsilon |r_l(0)|] \le \varepsilon \sum_{m=1}^n \frac{K_3 \hat{\beta}'_T \, e^{\varepsilon \sum_{l=m+1}^n |r_l(0)|}}{(1+m\varepsilon)^{1+\omega}} \le c K_3 \hat{\beta}'_T, \tag{5.28}$$

which proves that the asymptotic expected number of infected individuals is 1. Also, as discussed before, $v_n = 1 + O(\hat{\beta}'_T)$, which means that the asymptotic diffusion constant is 1.

Let $Tt = n\varepsilon$ and $k' = \frac{k}{\sigma_T \sqrt{Tt}} \in \mathcal{A}_n$. By (1.2), we have $a(k') = \frac{|k|^2}{2dTt} + \frac{O(|k|^{2+2\Delta})}{(Tt)^{1+\Delta}}$. Therefore,

$$f_n(k') = [1 + O(\hat{\beta}'_T)]\big[1 - \varepsilon[1 + O(\hat{\beta}'_T)]\, a(k')\big]^{Tt/\varepsilon} = e^{-\frac{|k|^2}{2d}} \big[1 + O(\hat{\beta}'_T) + O(|k|^{2+2\Delta}(Tt)^{-\Delta})\big].$$

For small t, we use $\big(\frac{|k|^2}{Tt}\big)^{\Delta} \le c\big[\frac{\log(2+Tt)}{1+Tt}\big]^{\Delta} \le c\,(1+Tt)^{-\delta}$ for $\Delta > \delta$, to replace $O(|k|^{2+2\Delta}(Tt)^{-\Delta})$ by $O(|k|^2(1+Tt)^{-\delta})$. This proves the discrete version of (1.19).

The discrete version of (1.20) follows from (5.17) and (5.23), with the help of the fact that both $f_n(0)$ and v_n are $1 + O(\hat{\beta}'_T)$. The proof of the discrete version of (1.21) does not depend on d, and is the same as in Section 5.3. This completes the proof of the main result for $d \le 4$. □

6 Continuum limit

In this section, we compute the limit as $\varepsilon \downarrow 0$, and prove Proposition 2.5.

We will first prove convergence of $\frac{1}{\varepsilon^2}\pi^{\lambda}_{t;\varepsilon}(x)$ for $\lambda < \lambda_{\mathrm{c}}$ fixed, and then extend this to $\frac{1}{\varepsilon^2}\partial_\lambda \pi^{\lambda}_{t;\varepsilon}(x)$. These two statements will imply convergence of $\frac{1}{\varepsilon^2}\pi^{\lambda^{(\varepsilon)}}_{t;\varepsilon}(x)$ whenever $\lambda^{(\varepsilon)} \to \lambda$ such that $\lambda^{(\varepsilon)} \leq \lambda^{(\varepsilon)}_{\mathrm{c}}$ for ε sufficiently small. Indeed, for any $\lambda' < \lambda \leq \lambda_{\mathrm{c}}$, we can write

$$\Big|\frac{1}{\varepsilon^2}\pi^{\lambda^{(\varepsilon)}}_{t;\varepsilon}(x) - \pi^{\lambda'}_t(x)\Big| \leq \frac{1}{\varepsilon^2}\Big|\pi^{\lambda^{(\varepsilon)}}_{t;\varepsilon}(x) - \pi^{\lambda'}_{t;\varepsilon}(x)\Big| + \Big|\frac{1}{\varepsilon^2}\pi^{\lambda'}_{t;\varepsilon}(x) - \pi^{\lambda'}_t(x)\Big|. \tag{6.1}$$

The second term in (6.1) converges to zero by assumption, while we estimate the first term by

$$\frac{1}{\varepsilon^2}\Big|\pi^{\lambda^{(\varepsilon)}}_{t;\varepsilon}(x) - \pi^{\lambda'}_{t;\varepsilon}(x)\Big| \leq \int_{\lambda'}^{\lambda^{(\varepsilon)}} \frac{1}{\varepsilon^2}\Big|\partial_\lambda \pi^{\mu}_{t;\varepsilon}(x)\Big| d\mu. \tag{6.2}$$

Since $\lambda^{(\varepsilon)}_{\mathrm{c}} \to \lambda_{\mathrm{c}}$ and $\lambda' < \lambda_{\mathrm{c}}$, we must have that $\lambda' \leq \lambda^{(\varepsilon)}_{\mathrm{c}}$ for ε sufficiently small. Therefore, the function in the integral is uniformly bounded (even when we sum over x) by $C\beta(1+t)^{-(d-2)/2}$. Thus, the limsup of the integral when $\varepsilon \downarrow 0$ is bounded by a multiple of $\lambda - \lambda'$. By taking the limit of $\lambda' \uparrow \lambda$ and using continuity of $\pi^{\lambda}_t(x)$ as proved in Section 2.5, (2.31) follows. Therefore, we are left to prove convergence of $\frac{1}{\varepsilon^2}\pi^{\lambda}_{t;\varepsilon}(x)$ and $\frac{1}{\varepsilon^2}\partial_\lambda \pi^{\lambda}_{t;\varepsilon}(x)$ for fixed $\lambda < \lambda_{\mathrm{c}}$.

We will prove that for every $N \geq 0$, every $\lambda < \lambda_{\mathrm{c}}$ and (x,t) fixed, there exists $\pi^{(N)}_t(x)$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x) = \pi^{(N)}_t(x), \tag{6.3}$$

where we have suppressed the dependence on λ. Thus, we will deal with *pointwise convergence,* rather than the uniform bounds in Section 4 above that provide bounds in (x,t) valid for all ε. Therefore, all terms that are $o(1)$ as $\varepsilon \downarrow 0$ will be estimated away. The proof is divided into several steps.

Double connections. Recall that in Section 4.4, we have shown that we may restrict the cluster $\mathbf{C}_{[0,t]}$ to a finite box. Thus, in the sequel of the proof, we will always assume that $\mathbf{C}_{[0,t]} \subset \square_R$, and omit the dependence on R. The second step in the proof is to determine the limit of probabilities involving double connections.

For sets of sites $\mathbf{S}_1, \mathbf{S}_2 \subset \mathbb{Z}^d \times \mathbb{R}_+$, we introduce the notation

$$\{\mathbf{S}_1 \to \mathbf{S}_2\} = \bigcup_{\boldsymbol{s}_1 \in \mathbf{S}_1, \boldsymbol{s}_2 \in \mathbf{S}_2} \{\boldsymbol{s}_1 \to \boldsymbol{s}_2\}, \qquad \{\mathbf{S}_1 \rightrightarrows \mathbf{S}_2\} = \bigcup_{\boldsymbol{s}_1, \boldsymbol{s}_1' \in \mathbf{S}_1, \boldsymbol{s}_2, \boldsymbol{s}_2' \in \mathbf{S}_2} \{\boldsymbol{s}_1 \to \boldsymbol{s}_2\} \circ \{\boldsymbol{s}_1' \to \boldsymbol{s}_2'\}. \tag{6.4}$$

We will first show that for any $\mathcal{C} \subset \mathbb{Z}^d$ containing at least two elements,

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\mathbb{P}^{\lambda}_{\varepsilon}\big(\mathcal{C} \times \{0\} \rightrightarrows (x,t)\big) = \lambda \sum_{z \in \mathbb{Z}^d} D(x-z)\mathbb{P}^{\lambda}\big(\mathcal{C} \times \{0\} \rightrightarrows \{(x,t),(z,t)\}\big). \tag{6.5}$$

In order to prove (6.5), we note that $\{y \in \mathbb{Z}^d : ((y,t-\varepsilon),(x,t))$ occupied$\}$ should contain at least two elements, and that for every x, y, z distinct,

$$\frac{1}{\varepsilon}\mathbb{P}^{\lambda}_{\varepsilon}\big((y,t-\varepsilon) \to (x,t) \text{ and } (z,t-\varepsilon) \to (x,t)\big) = o(1), \tag{6.6}$$

so that the main contribution in (6.5) comes from the case where $\{y \in \mathbb{Z}^d : ((y,t-\varepsilon),(x,t))$ occupied$\}$ contains x and a single other site in $\mathbb{Z}^d$.

We first use the Markov property at time $t-\varepsilon$ to obtain

$$\begin{aligned}\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon(\mathcal{C}\times\{0\}\rightrightarrows(x,t)) &= \lambda\sum_{z\in\mathbb{Z}^d}D(x-z)\mathbb{P}^\lambda_\varepsilon\Big(\{\mathcal{C}\times\{0\}\to(x,t-\varepsilon)\}\circ\{\mathcal{C}\times\{0\}\to(z,t-\varepsilon)\}\Big)+o(1)\\ &= \lambda\sum_{z\in\mathbb{Z}^d}D(x-z)\mathbb{P}^\lambda_\varepsilon\big(\mathcal{C}\times\{0\}\rightrightarrows\{(x,t-\varepsilon),(z,t-\varepsilon)\}\big)+o(1), \qquad (6.7)\end{aligned}$$

where in the latter equality, we use the fact that the probability that $\mathcal{C}\times\{0\}$ is doubly connected to $(x,t-\varepsilon)$ or $(z,t-\varepsilon)$ is $o(1)$.

We can also assume that no spatial bond leading to (x,t) and (z,t) is occupied, since the probability of the complement is $o(1)$. Then, we arrive at

$$\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon\big(\mathcal{C}\times\{0\}\rightrightarrows(x,t)\big)=\lambda\sum_{z\in\mathbb{Z}^d}D(x-z)\mathbb{P}^\lambda_\varepsilon\big(\mathcal{C}\times\{0\}\rightrightarrows\{(x,t),(z,t)\}\big)+o(1). \qquad (6.8)$$

Using the weak convergence of $\mathbb{P}^\lambda_\varepsilon$ towards $\mathbb{P}^\lambda$ as formulated in [4, Proposition 2.7] we obtain (6.5) for any $\mathcal{C}$ finite. Here we note that for $\mathcal{C}$ finite and $F=\big\{\mathcal{C}\times\{0\}\rightrightarrows\{(x,t),(z,t)\}\big\}$, the boundary event ∂F of F satisfies $\mathbb{P}^\lambda(\partial F)=0$.

Statement and initialization of the induction in N. We write $\pi^{(0)}_{t;\varepsilon}(x;\mathcal{C})=\mathbb{P}^\lambda_\varepsilon((o,0)\rightrightarrows(x,t),\mathbf{C}_t\backslash\{x\}=\mathcal{C})$ and, for $N\geq1$,

$$\pi^{(N)}_{t;\varepsilon}(x;\mathcal{C})=\sum_{\vec{b}_N}\mathbb{P}^\lambda_\varepsilon\big(\tilde{E}^{(N)}_{\vec{b}_N}(x,t),\mathbf{C}_t(\overline{b}_{N-1})\backslash\{x\}=\mathcal{C}\big). \qquad (6.9)$$

We will use induction in N to prove that for every $\mathcal{C}\subset\mathbb{Z}^d$ and every $t>0$, there exists $\pi^{(N)}_t(x;\mathcal{C})$ such that

$$\lim_{\varepsilon\downarrow0}\frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x;\mathcal{C})=\pi^{(N)}_t(x;\mathcal{C}). \qquad (6.10)$$

The claim for $\pi^{(N)}_{t;\varepsilon}(x)$ then follows by summation over $\mathcal{C}\subset\mathbb{Z}^d$, together with the fact that the main contribution comes from finite sets $\mathcal{C}$ by the finite containment property.

We first investigate $N=0$. We use the Markov property at time ε to arrive at

$$\pi^{(0)}_{t;\varepsilon}(x;\mathcal{C})=\sum_{\mathcal{A}\subset\mathbb{Z}^d}\prod_{a\in\mathcal{A}}p_\varepsilon(a)\Big[\prod_{a\notin\mathcal{A}}(1-p_\varepsilon(a))\Big]\mathbb{P}^\lambda_\varepsilon\Big(\mathcal{A}\times\{\varepsilon\}\rightrightarrows(x,t),\mathbf{C}_t(\mathcal{A}\times\{\varepsilon\})\backslash\{x\}=\mathcal{C}\Big), \qquad (6.11)$$

where, for any $\mathbf{A}\subset\mathbb{Z}^d\times\mathbb{R}_+$, we write

$$\mathbf{C}_t(\mathbf{A})=\bigcup_{\boldsymbol{a}\in\mathbf{A}}\mathbf{C}_t(\boldsymbol{a}). \qquad (6.12)$$

Since every $a\neq o$ gives rise to a factor ε, we immediately see that the main contribution comes from $\mathcal{A}=\{o,y\}$ for a unique $y\in\mathbb{Z}^d\backslash\{o\}$. Thus, using (6.5),

$$\lim_{\varepsilon\downarrow0}\frac{1}{\varepsilon^2}\pi^{(0)}_{t;\varepsilon}(x;\mathcal{C})=\lambda^2\sum_{y,z\in\mathbb{Z}^d}D(y)D(x-z)\mathbb{P}^\lambda\Big(\{(o,0),(y,0)\}\rightrightarrows\{(x,t),(z,t)\},(\mathbf{C}_t(o,0)\cup\mathbf{C}_t(y,0))\backslash\{x\}=\mathcal{C}\Big). \qquad (6.13)$$

This completes the proof for $N=0$ with $\pi^{(0)}_t(x;\mathcal{C})$, for $t>0$, defined to be the right-hand side of (6.13).

The contribution due to $\underline{b}_N=(o,0)$. We will now advance the induction. We first extract the term for $\underline{b}_N=(o,0)$ in (6.9), which is special and only arises when $N=1$. This term equals, where we recall that $\overline{b}_0=(o,0)$,

$$\sum_y\mathbb{P}^\lambda_\varepsilon\Big(E[((o,0),(y,\varepsilon)),(x,t);\mathbf{C}^{((o,0),(y,\varepsilon))}(o,0)]\cap\{\mathbf{C}_t\backslash\{x\}=\mathcal{C}\}\Big). \qquad (6.14)$$

We introduce some notation that will be helpful to deal with the sets $E[b_N,(x,t);C^{b_N}(\overline{b}_{N-1})]$ present in (6.9). For a set of sites $\mathcal{C}\subset\mathbb{Z}^d$ and $y\notin\mathcal{C}$, we let

$$\mathcal{E}_{s,t}(y,x;\mathcal{C}) = \bigcup_{(u,r)}\Big\{\big\{(y,s)\to(u,r)\rightrightarrows(x,t)\big\}\cap\big\{(u,r)\notin\mathbf{C}(\mathcal{C}\times\{s\})\big\}\cap\big\{(x,t)\in\mathbf{C}(\mathcal{C}\times\{s\})\backslash\mathbf{C}^{(u,r)}(y,s)\big\}\Big\},$$

where we write $\mathbf{C}^{(u,r)}(y,s)=\{(v,t):(y,s)\to(v,t)$ via a path that does not use $(u,r)\}$. In terms of this definition, and writing $\overline{b}_N=(y,s)$, we can rewrite

$$\begin{aligned}E[b_N,(x,t);\mathbf{C}^{b_N}(\overline{b}_{N-1})] &= \mathcal{E}_{s,t}(y,x;\mathbf{C}_s(\overline{b}_{N-1})\backslash\{y\})\cap\{b_N\text{ occ.}\}\\ &\qquad \dot\cup\bigcup_{(u,r)}\Big\{\big\{(y,s)\to(u,r)\rightrightarrows(x,t)\big\}\circ\big\{\mathcal{C}\times\{s\}\to(u,r)\big\}\cap\{b_N\text{ occ.}\}\Big\}.\end{aligned}\tag{6.15}$$

For the second line of this event, we obtain an extra factor of ε due to the fact that there must be *two* spatial bonds present at (u,r) (one before and one after (u,r)), which leads to an error term as $\varepsilon\downarrow 0$. Thus, we will focus on the first line of this event.

When $\underline{b}_N=(o,0)$, we must have that there are two occupied bonds emanating from $(o,0)$, since $(o,0)\rightrightarrows(x,t)$ on $E[b_N,(x,t);\mathbf{C}^{b_N}(o,0)]$. Similarly to the case for $N=0$, the main contribution comes from configurations where there is a single spatial occupied bond from the origin, and the other occupied bond is a temporal bond. Using the Markov property at time ε, we obtain

$$\begin{aligned}&\frac{1}{\varepsilon^2}\sum_y\mathbb{P}_\varepsilon^\lambda\big(E[((o,0),(y,\varepsilon)),(x,t);\mathbf{C}^{((o,0),(y,\varepsilon))}(o,0)]\cap\{\mathbf{C}_t\backslash\{x\}=\mathcal{C}\}\big)\\ &\quad=\frac{1}{\varepsilon}\sum_w\lambda D(w)\Big[\mathbb{P}_\varepsilon^\lambda\Big(\mathcal{E}_{0,t}\big(o,x;\{o,w\}\big)\cap\big\{\mathbf{C}_t\big(\{(o,0),(w,0)\}\big)\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &\qquad+\mathbb{P}_\varepsilon^\lambda\Big(\mathcal{E}_{0,t}\big(w,x;\{o,w\}\big)\cap\big\{\mathbf{C}_t\big(\{(o,0),(w,0)\}\big)\backslash\{x\}=\mathcal{C}\big\}\Big)\Big]+o(1).\end{aligned}\tag{6.16}$$

Here the first term corresponds to $y=o$ and the spatial bond $((o,0),(w,\varepsilon))$ being occupied, while the second term corresponds to $y=w\neq o$. We have now extracted one of the necessary factors of ε. However, the problem with the event on the right-hand side of (6.16) is that its probability still tends to zero when $\varepsilon\downarrow 0$ due to the fact that there must be at least two occupied bonds leading to (x,t). To define its limit, for $z\neq x$, let $\mathcal{E}_{s,t}(y,x,z;\mathcal{C})$ be the event

$$\begin{aligned}\mathcal{E}_{s,t}(y,x,z;\mathcal{C}) = \bigcup_{(u,r)}\Big\{&\big\{(y,s)\to(u,r)\rightrightarrows\{(x,t),(z,t)\}\big\}\cap\big\{(u,r)\notin\mathbf{C}(\mathcal{C}\times\{s\})\big\}\\ &\cap\big\{\{(x,t),(z,t)\}\subset\mathbf{C}(\mathcal{C}\times\{s\})\backslash\mathbf{C}^{(u,r)}(y,s)\big\}\Big\}.\end{aligned}\tag{6.17}$$

Thus, $\mathcal{E}_{s,t}(y,x,z;\mathcal{C})$ is identical to $\mathcal{E}_{s,t}(y,x;\mathcal{C})$, apart form the fact that the role of the single site (x,t) is taken over by the two sites (x,t) and (z,t).

Note that by an argument similar to (6.5), we have that for all $\mathcal{C}\subset\mathbb{Z}^d$ and any $y\in\mathbb{Z}^d$,

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{P}_\varepsilon^\lambda(\mathcal{E}_{s,t}(y,x;\mathcal{C}))=\lambda\sum_{z\in\mathbb{Z}^d}D(x-z)\mathbb{P}^\lambda\big(\mathcal{E}_{s,t}(y,x,z;\mathcal{C})\big).\tag{6.18}$$

In particular, for the contribution due to $\underline{b}_N=(o,0)$, we obtain that

$$\begin{aligned}&\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^2}\sum_y\mathbb{P}_\varepsilon^\lambda\Big(E\big[((o,0),(y,\varepsilon)),(x,t);\mathbf{C}^{((o,0),(y,\varepsilon))}(o,0)\big]\cap\big\{\mathbf{C}_t\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &\quad=\sum_{w,z}\lambda^2D(w)D(x-z)\Big[\mathbb{P}^\lambda\Big(\mathcal{E}_{0,t}\big(o,x,z;\{o,w\}\big)\cap\big\{\mathbf{C}_t\big(\{(o,0),(w,0)\}\big)\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &\qquad+\mathbb{P}^\lambda\Big(\mathcal{E}_{0,t}\big(w,x,z;\{o,w\}\big)\cap\big\{\mathbf{C}_t\big(\{(o,0),(w,0)\}\big)\backslash\{x\}=\mathcal{C}\big\}\Big)\Big].\end{aligned}\tag{6.19}$$

Advancement of the induction. We now continue the advancement of the induction hypothesis. We note that when $\underline{b}_N \neq (o,0)$, the main contribution in $\pi^{(N)}_{t;\varepsilon}(x;\mathcal{C})$ comes from b_N that are temporal bonds. See also (6.15) and note that for temporal bonds b_N, we have that $\mathbb{P}^\lambda_\varepsilon(b_N \text{ occupied}) = 1-\varepsilon = 1-o(1)$, so that we can ignore the intersection with $\{b_N \text{ occupied}\}$ in (6.15). Moreover, in a similar fashion, we may assume that no spatial bonds leading into $\mathbf{C}_s(\overline{b}_{N-1})\times\{s\}$ are occupied, so that $\mathbf{C}_{s-\varepsilon}(\overline{b}_{N-1}) = \mathbf{C}_s(\overline{b}_{N-1})$. Conditioning on $\mathbf{C}_{s-\varepsilon}(\overline{b}_{N-1}) = \mathcal{C}'$ and using the Markov property at time $s-\varepsilon$, we obtain that for $N \geq 1$

$$\pi^{(N)}_{t;\varepsilon}(x;\mathcal{C}) = \sum_{s=\varepsilon}^{t}{}^{\bullet} \sum_{y\in\Box_R}\sum_{\mathcal{C}'\subset\Box_R} \pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\mathcal{C}')\mathbb{P}^\lambda_\varepsilon\Big(\mathcal{E}_{s,t}(y,x;\mathcal{C}')\cap\big\{\mathbf{C}_t(y,s)\backslash\{x\}=\mathcal{C}\big\}\Big)+o(\varepsilon^2). \tag{6.20}$$

We have already proved that the contribution due to $\underline{b}_N = (o,0)$, which only arises when $N=1$, converges. Thus, we may now restrict attention to $\underline{b}_N \neq (o,0)$. In this case, $s-\varepsilon > 0$, and the induction hypothesis applies to $\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\mathcal{C}')$. In particular, $\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\mathcal{C}')$ is order ε^2. Since the sums over y and $\mathcal{C}'$ are finite, it suffices to prove that for any $\mathcal{C},\mathcal{C}'\subset\mathbb{Z}^d$ finite and any $y\in\mathbb{Z}^d$,

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^2}\sum_{s=2\varepsilon}^{t}{}^{\bullet}\,&\pi^{(N-1)}_{s;\varepsilon}(y;\mathcal{C}')\mathbb{P}^\lambda_\varepsilon\Big(\mathcal{E}_{s,t}(y,x;\mathcal{C}')\cap\big\{\mathbf{C}_t(y,s)\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &=\lambda\sum_{z\in\mathcal{C}\backslash\{x\}}D(x-z)\int_0^t\pi^{(N-1)}_s(y;\mathcal{C}')\mathbb{P}^\lambda\big(\mathcal{E}_{s,t}(y,x,z;\mathcal{C}')\cap\{\mathbf{C}_t(y,s)\backslash\{x\}=\mathcal{C}\}\big)ds.\end{aligned} \tag{6.21}$$

In order to do so, we will use dominated convergence.

We write the contribution to $s\geq 2\varepsilon$ as

$$\begin{aligned}\frac{1}{\varepsilon^2}\sum_{s=2\varepsilon}^{t}{}^{\bullet}\,&\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\mathcal{C}')\mathbb{P}^\lambda_\varepsilon\Big(\mathcal{E}_{s,t}(y,x;\mathcal{C}')\cap\big\{\mathbf{C}_t(y,s)\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &=\int_\varepsilon^t\frac{1}{\varepsilon^2}\pi^{(N-1)}_{s_\varepsilon-\varepsilon;\varepsilon}(y;\mathcal{C}')\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon\Big(\mathcal{E}_{s_\varepsilon,t}(y,x;\mathcal{C}')\cap\big\{\mathbf{C}_t(y,s_\varepsilon)\backslash\{x\}=\mathcal{C}\big\}\Big)ds,\end{aligned} \tag{6.22}$$

where $s_\varepsilon = \varepsilon\lceil s/\varepsilon\rceil$. By the induction hypothesis, we have that the first term in the integral converges pointwise to $\pi^{(N-1)}_s(y;\mathcal{C}')$. Moreover, by an argument similar to (6.18), the second term converges pointwise to

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon\Big(&\mathcal{E}_{s_\varepsilon,t}(y,x;\mathcal{C}')\cap\big\{\mathbf{C}_t(y,s_\varepsilon)\backslash\{x\}=\mathcal{C}\big\}\Big)\\ &=\lambda\sum_{z\in\mathbb{Z}^d}D(z-x)\mathbb{P}^\lambda\Big(\mathcal{E}_{s,t}(y,x,z;\mathcal{C}')\cap\{\mathbf{C}_t(y,s)\backslash\{x\}=\mathcal{C}\}\Big).\end{aligned} \tag{6.23}$$

Therefore, by dominated convergence, we obtain (6.21).

This advances the induction hypotheses, and proves the claim. In particular, we have that for every N, $\frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x)$ converges to $\pi^{(N)}_t(x)$. Moreover, by the uniform bound in Section 4, we have that for L large and all $\lambda<\lambda_{\rm c}$, $\sum_x\frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x)\leq C\beta^{N\vee 1}(1+t)^{-d/2}$. Therefore, again by dominated convergence, now for the sum over $N\geq 0$,

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^2}\pi^\lambda_{t;\varepsilon}(x)=\lim_{\varepsilon\downarrow 0}\sum_{N=0}^\infty(-1)^N\frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x)=\sum_{N=0}^\infty(-1)^N\pi^{(N)}_t(x)=\pi^\lambda_t(x). \tag{6.24}$$

This completes the proof of pointwise convergence of $\frac{1}{\varepsilon^2}\pi^\lambda_{t;\varepsilon}(x)$.

Bounds on π^λ_t in (2.32) and convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$. The bound on $\sum_{x\in\mathbb{Z}^d}|x|^q\pi^\lambda_t(x)$ follow immediately from the pointwise convergence of $\frac{1}{\varepsilon^2}\pi^\lambda_{t;\varepsilon}(x)$, together with the uniform bounds in Proposition 2.2 and dominated convergence for the sum over x.

To prove convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$, we first note that by [24, Section 3.1], $\lambda_{\rm c}^{(\varepsilon)} \to \lambda_{\rm c}$. Convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$ follows by dominated convergence, together with the identification of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$ in (2.30). Thus, we obtain that

$$A = \big[1 + \int_0^\infty t\hat{\pi}_t^{\lambda_{\rm c}}(0)dt\big]^{-1}, \qquad\qquad v = A\big[\lambda_{\rm c} - \frac{1}{\sigma^2}\int_0^\infty \hat{\nabla}^2\pi_t^{\lambda_{\rm c}}(0)dt\big]. \tag{6.25}$$

Convergence of $\frac{1}{\varepsilon^2}\partial_\lambda \pi_{t;\varepsilon}^{\lambda}(x)$ and bound on $\partial_\lambda \pi_t^{\lambda}$. The only difference of $\frac{1}{\varepsilon^2}\partial_\lambda \pi_{t;\varepsilon}^{\lambda}(x)$ compared to $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^{\lambda}(x)$ is the occurrence of the sum over spatial bonds b and the indicator of the event $b \in \texttt{piv}[\overline{b}_n, \underline{b}_{n+1}] \cup \{b_n\}$. Clearly, the main term in the above comes from $b \in \texttt{piv}[\overline{b}_n, \underline{b}_{n+1}]$. The extra inclusion of this event gives rise to an extra integral over the time variable r' and an indicator that the arrow $((w, r'), (w', r'))$ is pivotal for the connection from $\overline{b}_n$ to $\underline{b}_{n+1}$ (see [20, p. 61] for the definition of a pivotal arrow). Apart from this minor modification, the proof remains unchanged. The bound on $\sum_{x\in\mathbb{Z}^d} |\partial_\lambda \pi_t^{\lambda}(x)|$ in (2.32) follows immediately from the pointwise convergence of $\frac{1}{\varepsilon^2}\partial_\lambda \pi_{t;\varepsilon}^{\lambda}(x)$, together with the uniform bounds in Proposition 2.2 and dominated convergence for the sum over x. $\square$

Acknowledgements

The work of AS was supported in part by NSERC of Canada. The work of RvdH was carried out in part at Delft University of Technology. This project was initiated during an extensive visit of RvdH to the University of British Columbia, Vancouver, Canada. We thank Gordon Slade and Ed Perkins for stimulating discussions during various stages of the project.

References

[1] M. Aizenman and C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.* **36** (1984): 107–143.

[2] D. J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.* **19** (1991): 1520–1536.

[3] D. J. Barsky and C. C. Wu. Critical exponents for the contact process under the triangle condition. *J. Statist. Phys.* **91** (1998): 95–124.

[4] C. Bezuidenhout and G. Grimmett. Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* **19** (1991): 984–1009.

[5] D.C. Brydges and T. Spencer. Self-avoiding walk in 5 or more dimensions. *Commun. Math. Phys.*, **97** (1985):125–148.

[6] R. Durrett and E. Perkins. Rescaled contact processes converge to super-Brownian motion in two or more dimensions. *Probab. Th. Rel. Fields* **114** (1999): 309–399.

[7] G. Grimmett. *Percolation.* Springer, Berlin (1999).

[8] G. Grimmett and P. Hiemer. Directed percolation and random walk. *In and Out of Equilibrium* (ed. V. Sidoravicius). Birkhäuser (2002): 273-297.

[9] T. Hara and G. Slade. Mean-field critical behaviour for percolation in high dimensions. *Commun. Math. Phys.*, **128** (1990):333–391.

[10] T. Hara and G. Slade. Self-avoiding walk in five or more dimensions. I. The critical behaviour. *Commun. Math. Phys.*, **147** (1992):101–136.

[11] T. Hara and G. Slade. The scaling limit of the incipient infinite cluster in high-dimensional percolation. I. Critical exponents. *J. Statist. Phys.* **99** (2000): 1075–1168.

[12] T. Hara and G. Slade. The scaling limit of the incipient infinite cluster in high-dimensional percolation. II. Integrated super-Brownian excursion. *J. Math. Phys.* **41** (2000): 1244–1293.

[13] R. van der Hofstad, F. den Hollander and G. Slade. A new inductive approach to the lace expansion for self-avoiding walks. *Probab. Th. Rel. Fields* **111** (1998): 253–286.

[14] R. van der Hofstad, F. den Hollander and G. Slade. Construction of the incipient infinite cluster for the spread-out oriented percolation above 4+1 dimensions. *Commun. Math. Phys.* **231** (2002): 435–461.

[15] R. van der Hofstad and A. Sakai. Convergence of the critical finite-range contact process to super-Brownian motion above 4 spatial dimensions. In preparation.

[16] R. van der Hofstad and A. Sakai. Critical points for spread-out contact processes and oriented percolation. In preparation.

[17] R. van der Hofstad and G. Slade. A generalised inductive approach to the lace expansion. *Probab. Th. Rel. Fields* **122** (2002): 389–430.

[18] R. van der Hofstad and G. Slade. Convergence of critical oriented percolation to super-Brownian motion above 4+1 dimensions. *Ann. Inst. H. Poincaré Probab. Statist.* **39** (2003): 413–485.

[19] R. van der Hofstad and G. Slade. The lace expansion on a tree with application to networks of self-avoiding walks. *Adv. Appl. Math.* **30** (2003): 471–528.

[20] T. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes.* Springer, Berlin (1999).

[21] B. G. Nguyen and W.-S. Yang. Triangle condition for oriented percolation in high dimensions. *Ann. Probab.* **21** (1993): 1809–1844.

[22] B. G. Nguyen and W.-S. Yang. Gaussian limit for critical oriented percolation in high dimensions. *J. Statist. Phys.* **78** (1995): 841–876.

[23] A. Sakai. Analyses of the critical behavior for the contact process based on a percolation structure. *Ph.D. thesis* (2000).

[24] A. Sakai. Mean-field critical behavior for the contact process. *J. Statist. Phys.* **104** (2001): 111–143.

[25] A. Sakai. Hyperscaling inequalities for the contact process and oriented percolation. *J. Statist. Phys.* **106** (2002): 201–211.

[26] R. Schonmann. The triangle condition for contact processes on homogeneous trees. *J. Statist. Phys.* **90** (1998): 1429–1440.

[27] G. Slade. The diffusion of self-avoiding random walk in high dimensions. *Commun. Math. Phys.*, **110** (1987): 661-683.

[28] G. Slade. Convergence of self-avoiding random walk to Brownian motion in high dimensions. *J. Phys. A: Math. Gen.*, **21** (1988):L417-L420.

[29] G. Slade, G. The lace expansion and the upper critical dimension for percolation. Lectures in Applied Mathematics, **27** (1991):53–63, (Mathematics of Random Media, eds. W.E. Kohler and B.S. White, A.M.S., Providence).

[30] C. C. Wu. The contact process on a tree: behavior near the first phase transition. *Stochastic Process. Appl.* **57** (1995): 99–112.

A Advancement of the induction hypotheses

In this appendix, we prove Lemmas 5.1–5.4 and we advance the induction hypotheses. We discuss the case of $d > 4$ in Appendix A.1, which is quite similar to the argument in [17]. The main difference is due to the required uniformity in ε. We will explain in detail how to use the factors of ε contained in the induction hypotheses and in the bounds (2.25)–(2.27), in order to obtain this uniformity. The argument for $d \leq 4$ is almost identical, except for modifications due to the factors β_T and $\hat{\beta}'_T$ in (4.4)–(4.5) and (5.25)–(5.27). We discuss the necessary changes for $d \leq 4$ in Appendix A.2.

A.1 Advancement above four dimensions

A.1.1 Proofs of Lemmas 5.1–5.4

Recall the induction hypotheses (H1)–(H4) and the definitions of λ_n, I_n and v_n in Section 5.1.2. We now prove Lemmas 5.1–5.4 using the induction hypotheses.

Proof of Lemma 5.1. We prove $\lambda \in I_{m-1}$ assuming $\lambda \in I_m$. By (5.7) and (5.10),

$$|\lambda - \lambda_{m-1}| \leq |\lambda - \lambda_m| + |\lambda_m - \lambda_{m-1}| \leq K_1 \beta \frac{1+(m+1)\varepsilon}{(1+m\varepsilon)^{d/2}} \leq \frac{K_1\beta}{[1+(m-1)\varepsilon]^{(d-2)/2}},$$

where the last inequality is the result of the fact that $f(\varepsilon) = (c+\varepsilon)(c-\varepsilon)^a$ is decreasing in $\varepsilon > 0$, if $c > 0$ and $a \geq 1$ (in the above inequality, $c = 1 + m\varepsilon$ and $a = \frac{d-2}{2}$). This completes the proof that $I_m \subset I_{m-1}$. □

Proof of Lemma 5.2. By (5.11)–(5.12) and the trivial inequality $1 + x \leq e^x$,

$$|f_m(0)| = \left|\prod_{l=1}^{m}[1+\varepsilon r_l(0)]\right| \leq e^{\varepsilon \sum_{l=1}^{m} |r_l(0)|} \leq e^{cK_3\beta}. \tag{A.1}$$

By (5.14)–(5.16) and (5.19), $|f_m(k)/f_m(0)|$ is bounded by

$$\left|\prod_{l=1}^{m}[1-\varepsilon v_l\, a(k) + \varepsilon\, s_l(k)]\right| \leq e^{-\varepsilon \sum_{l=1}^{m}[v_l a(k) - |s_l(k)|]} \leq e^{-m\varepsilon[1-c\,(K_1+K_2+K_3)\beta]\, a(k)}. \tag{A.2}$$

This completes the proof. □

Proof of Lemma 5.3. This is an immediate consequence of (5.17), (5.19), (5.23) and (A.1). □

Proof of Lemma 5.4. Recalling $\mathcal{A}_m \equiv \{k : a(k) \leq \gamma \frac{\log(2+m\varepsilon)}{1+m\varepsilon}\}$, we define

$$R_1 = \{k \in \mathcal{A}_m : \|k\|_\infty \leq L^{-1}\}, \qquad R_2 = \{k \in \mathcal{A}_m : \|k\|_\infty > L^{-1}\},$$
$$R_3 = \{k \notin \mathcal{A}_m : \|k\|_\infty \leq L^{-1}\}, \qquad R_4 = \{k \notin \mathcal{A}_m : \|k\|_\infty > L^{-1}\},$$

where R_2 is empty if m is sufficiently large. Then

$$\|\hat{D}^2 f_m\|_1 = \sum_{i=1}^{4} \int_{R_i} \frac{d^dk}{(2\pi)^d}\; \hat{D}(k)^2 |f_m(k)|.$$

On R_1, we consider the cases of $m\varepsilon < 1$ and $m\varepsilon \geq 1$ separately. If $m\varepsilon < 1$, we use Lemma 5.2 to obtain

$$\int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2|f_m(k)| \leq c \int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 \leq c\beta.$$

If $m\varepsilon \geq 1$, we use the inequality $\hat{D}^2(k) \leq 1$, Lemma 5.2, and then the assumption $a(k) \asymp L^2|k|^2$ for $\|k\|_\infty \leq L^{-1}$, to obtain

$$\int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2|f_m(k)| \leq c \int_{R_1} \frac{d^dk}{(2\pi)^d}\, e^{-cm\varepsilon L^2|k|^2} \leq c\beta\,(m\varepsilon)^{-d/2}.$$

Summarizing both cases, we obtain the desired bound on the contribution from R_1.

On R_2, we use Lemma 5.2 and the assumption $a(k) > \eta$ for $\|k\|_\infty > L^{-1}$ to conclude that there exists an $r > 1$ independently of β such that

$$\int_{R_2} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2|f_m(k)| \leq c \int_{R_2} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 r^{-m\varepsilon} \leq c\beta\, r^{-m\varepsilon}.$$

Since $r^{-m\varepsilon} \leq c\,(1+m\varepsilon)^{-d/2}$, we obtain the desired bound on the contribution from R_2.

On R_3 and R_4, we use (H4). Then the contribution from these two regions is bounded by

$$\frac{K_4}{(1+m\varepsilon)^{d/2}} \sum_{i=3}^{4} \int_{R_i} \frac{d^dk}{(2\pi)^d}\, \frac{\hat{D}(k)^2}{a(k)^{2+\rho}}.$$

It thus suffices to bound the integral by $c\beta$. On R_3, we use the inequality $\hat{D}(k)^2 \leq 1$ and the assumption $a(k) \asymp L^2|k|^2$ for $\|k\|_\infty \leq L^{-1}$. Since $\frac{d-4}{2} - \rho > 0$,

$$\int_{R_3} \frac{d^dk}{(2\pi)^d}\, \frac{\hat{D}(k)^2}{a(k)^{2+\rho}} \leq \frac{c}{L^{4+2\rho}} \int_{\|k\|_\infty \leq L^{-1}} \frac{d^dk}{|k|^{4+2\rho}} \leq c\beta.$$

On R_4, we use the assumption $a(k) > \eta$ for $\|k\|_\infty > L^{-1}$ and the fact that $\int \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 \leq \beta$, to obtain the desired bound $c\beta$ on the integral over R_4. This completes the proof. □

A.1.2 Advancement of the induction hypotheses

First we verify that the induction hypotheses hold for $n = 1$.

(H1)–(H2) By (5.6) and (5.8), $|\lambda_1 - \lambda_0| = |v_1 - v_0| = 0$.

(H3) By (5.4) and (5.11), $r_1(k) \equiv \lambda - 1$ and thus $|r_1(k) - r_1(0)| \equiv 0$. Together with $\lambda \in I_1$, we obtain $|r_1(0)| \leq K_1\beta/(1+\varepsilon)^{(d-2)/2}$. Therefore (H3) holds if $K_3 \geq K_1$.

(H4) By (5.4), $|f_1(k)| \leq 1+3\varepsilon$ and $|f_1(k) - f_0(k)| \leq 3\varepsilon$ for sufficiently small β. Together with the trivial bound $a(k) \leq 2$, (H4) is proved to hold if $K_4 \geq (1+3\varepsilon)2^{2+\rho}(1+\varepsilon)^{d/2}$ and $K_5 \geq 3{\cdot}2^{1+\rho}(1+\varepsilon)^{d/2}$.

Next we advance the induction hypotheses for $\lambda \in I_{n+1}$ under the assumption that (H1)–(H4) hold for all $m \leq n$ and that (2.25)–(2.27) hold for all $s/\varepsilon \leq n+1$. Recalling (5.2), we immediately obtain the following bounds on $e_m(k)$ and $g_m(k)$: there exists $C_g < \infty$ such that

$$|e_m(k)| \leq \frac{\varepsilon^2 C_g\beta}{(1+m\varepsilon)^{d/2}}, \qquad |e_m(k) - e_m(0)| \leq \frac{\varepsilon^2 C_g\beta\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}, \tag{A.3}$$

$$|g_m(k)| \leq \frac{\varepsilon^2 C_g\beta}{(1+m\varepsilon)^{d/2}}, \qquad |\nabla^2 g_m(0)| \leq \frac{\varepsilon^2 C_g\sigma^2\beta}{(1+m\varepsilon)^{(d-2)/2}}, \tag{A.4}$$

$$\left|g_m(k) - g_m(0) - \frac{a(k)}{\sigma^2}\,\nabla^2 g_m(0)\right| \leq \frac{\varepsilon^2 C_g\beta\, a(k)^{1+\Delta'}}{(1+m\varepsilon)^{(d-2)/2-\Delta'}}, \tag{A.5}$$

$$|\partial_\lambda g_m(0)| \leq \frac{\varepsilon^2 C_g\beta}{(1+m\varepsilon)^{(d-2)/2}}. \tag{A.6}$$

Advancement of (H1). By (5.6) and the mean value theorem,

$$\begin{aligned}\lambda_{n+1}-\lambda_n &= -\frac{1}{\varepsilon}\,g_{n+1}(0;\lambda_n)-\frac{1}{\varepsilon}\sum_{m=2}^{n}[g_m(0;\lambda_n)-g_m(0;\lambda_{n-1})]\\ &= -\frac{1}{\varepsilon}\,g_{n+1}(0;\lambda_n)-\frac{\lambda_n-\lambda_{n-1}}{\varepsilon}\sum_{m=2}^{n}\partial_\lambda g_m(0;\lambda_*),\end{aligned} \tag{A.7}$$

for some λ_* between λ_n and λ_{n-1}. By (5.7) and (H1), $\lambda_* \in I_n$. By (A.4), (A.6) and (H1),

$$|\lambda_{n+1}-\lambda_n| \le \frac{\varepsilon C_g\beta}{[1+(n+1)\varepsilon]^{d/2}}+|\lambda_n-\lambda_{n-1}|\varepsilon\sum_{m=2}^{n}\frac{C_g\beta}{(1+m\varepsilon)^{(d-2)/2}} \le \frac{\varepsilon C_g(1+cK_1\beta)\beta}{[1+(n+1)\varepsilon]^{d/2}}.$$

Therefore (H1) holds for $n+1$, if β is sufficiently small and $K_1 > C_g$. $\square$

Advancement of (H2). Let $1+M_n$ and N_n be the denominator and the numerator of (5.8) for $n \ge 2$ respectively. Then

$$v_{n+1}-v_n = \frac{\frac{-1}{\sigma^2\varepsilon}\nabla^2 g_{n+1}(0)}{1+M_{n+1}}-\frac{N_n\,n\,g_{n+1}(0)}{(1+M_{n+1})(1+M_n)}. \tag{A.8}$$

By (A.4), we obtain that, for $m \le n+1$,

$$|M_m| \le \varepsilon\sum_{l=2}^{m}\frac{(l-1)\varepsilon C_g\beta}{(1+l\varepsilon)^{d/2}} \le cC_g\beta, \qquad |N_m-\lambda| \le \varepsilon\sum_{l=2}^{m}\frac{C_g\beta}{(1+l\varepsilon)^{(d-2)/2}} \le cC_g\beta, \tag{A.9}$$

and

$$\Big|\frac{-1}{\sigma^2\varepsilon}\,\nabla^2 g_{n+1}(0)\Big| \le \frac{\varepsilon C_g\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}, \qquad |n\,g_{n+1}(0)| \le \frac{n\varepsilon^2 C_g\beta}{[1+(n+1)\varepsilon]^{d/2}}.$$

Therefore,

$$|v_{n+1}-v_n| \le \frac{1+\lambda}{(1-cC_g\beta)^2}\,\frac{\varepsilon C_g\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}.$$

Since $\lambda \in I_{n+1}$, (H2) holds if β is sufficiently small and $K_2 > 2C_g$. $\square$

Advancement of (H3). First we derive expressions for $r_{n+1}(0)$ and $r_{n+1}(k)-r_{n+1}(0)$. By dividing both sides of (5.3) by $f_n(k)$ and using $g_1(k) = 1-\varepsilon+\lambda\varepsilon\hat{D}(k)$,

$$\begin{aligned}\frac{f_{n+1}(k)}{f_n(k)} &= g_1(k)+\sum_{m=1}^{n}g_{m+1}(k)\frac{f_{n-m}(k)}{f_n(k)}+\frac{e_{n+1}(k)}{f_n(k)}\\ &= 1-\varepsilon v_{n+1}a(k)+\varepsilon\Big[v_{n+1}a(k)-1+\lambda\hat{D}(k)+\frac{1}{\varepsilon}\sum_{m=1}^{n}g_{m+1}(k)\frac{f_{n-m}(k)}{f_n(k)}+\frac{e_{n+1}(k)}{\varepsilon f_n(k)}\Big].\end{aligned}$$

Therefore $r_{n+1}(k)$ equals the quantity in the brackets. In particular,

$$\begin{aligned}r_{n+1}(0) &= -1+\lambda+\frac{1}{\varepsilon}\sum_{m=1}^{n}g_{m+1}(0)\frac{f_{n-m}(0)}{f_n(0)}+\frac{e_{n+1}(0)}{\varepsilon f_n(0)}\\ &= \Big[\lambda-1+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}g_m(0)\Big]+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}g_m(0)\Big[\frac{f_{n+1-m}(0)}{f_n(0)}-1\Big]+\frac{e_{n+1}(0)}{\varepsilon f_n(0)}.\end{aligned} \tag{A.10}$$

Let $r^{(1)}_{n+1}(0)$, $r^{(2)}_{n+1}(0)$ and $r^{(3)}_{n+1}(0)$ denote the first, second and third terms, respectively. Similarly,

$$\begin{aligned} r_{n+1}(k)-r_{n+1}(0)=&(v_{n+1}-\lambda)\,a(k)+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}[g_m(k)-g_m(0)]\frac{f_{n+1-m}(k)}{f_n(k)}\\ &+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}g_m(0)\bigg[\frac{f_{n+1-m}(k)}{f_n(k)}-\frac{f_{n+1-m}(0)}{f_n(0)}\bigg]+\frac{1}{\varepsilon}\bigg[\frac{e_{n+1}(k)}{f_n(k)}-\frac{e_{n+1}(0)}{f_n(0)}\bigg]. \end{aligned} \tag{A.11}$$

By (5.8), we can rewrite v_{n+1} as

$$v_{n+1}=\lambda-\frac{1}{\sigma^2\varepsilon}\sum_{m=2}^{n+1}\nabla^2 g_m(0)-v_{n+1}\sum_{m=2}^{n+1}(m-1)\,g_m(0).$$

Substituting the above identity into (A.11), we obtain

$$\begin{aligned} r_{n+1}(k)-r_{n+1}(0)=&\frac{1}{\varepsilon}\sum_{m=2}^{n+1}\bigg[[g_m(k)-g_m(0)]\,\frac{f_{n+1-m}(k)}{f_n(k)}-\frac{a(k)}{\sigma^2}\,\nabla^2 g_m(0)\bigg]\\ &+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}g_m(0)\bigg[\frac{f_{n+1-m}(k)}{f_n(k)}-\frac{f_{n+1-m}(0)}{f_n(0)}-\varepsilon v_{n+1}\,(m-1)\,a(k)\bigg]\\ &+\frac{1}{\varepsilon}\bigg[\frac{e_{n+1}(k)}{f_n(k)}-\frac{e_{n+1}(0)}{f_n(0)}\bigg]. \end{aligned} \tag{A.12}$$

Let $\Delta r^{(1)}_{n+1}(k)$, $\Delta r^{(2)}_{n+1}(k)$ and $\Delta r^{(3)}_{n+1}(k)$ denote the first, second and third terms, respectively.

Thus, to advance (H3), we are left to bound $r^{(i)}_{n+1}(0)$ and $\Delta r^{(i)}_{n+1}(k)$ for $i=1,2,3$.

Advancement of the first inequality in (5.12)*.* Note that $r^{(1)}_{n+1}(0;\lambda_n)=\lambda_n-\lambda_{n+1}$. By the mean value theorem,

$$\begin{aligned} \big|r^{(1)}_{n+1}(0)\big| &\le |\lambda-\lambda_n|+|\lambda_n-\lambda_{n+1}|+\bigg|\lambda_{n+1}-1+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}g_m(0;\lambda)\bigg|\\ &=|\lambda-\lambda_n|+\big|r^{(1)}_{n+1}(0;\lambda_n)\big|+\bigg|\frac{1}{\varepsilon}\sum_{m=2}^{n+1}\big[g_m(0;\lambda)-g_m(0;\lambda_n)\big]\bigg|\\ &\le \big|r^{(1)}_{n+1}(0;\lambda_n)\big|+|\lambda-\lambda_n|\bigg[1+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}\big|\partial_\lambda g_m(0;\lambda_*)\big|\bigg], \end{aligned}$$

for some $\lambda_*\in I_n$ between λ and λ_n, because $\lambda\in I_{n+1}$ and $\lambda_n\in I_n$. By using (A.6), (5.7) and (5.10), the right-hand side is bounded by

$$\frac{\varepsilon K_1\beta}{[1+(n+1)\varepsilon]^{d/2}}+\frac{K_1\beta}{(1+n\varepsilon)^{(d-2)/2}}\bigg[1+\varepsilon\sum_{m=2}^{n+1}\frac{C_g\beta}{(1+m\varepsilon)^{(d-2)/2}}\bigg]\le\frac{cK_1\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}. \tag{A.13}$$

Next we investigate $r^{(2)}_{n+1}(0)$. By (A.4) and (5.15), together with Taylor's theorem applied to $h(t)=\prod_i(1+c_it)^{-1}$:

$$|h(1)-h(0)|\le\sup_{t\in(0,1)}|h'(t)|\le\sum_i\frac{|c_i|}{1-|c_i|}\,e^{\sum_j\frac{|c_j|}{1-|c_j|}}, \tag{A.14}$$

$$|h(1)-h(0)-h'(0)|\le\frac{1}{2}\sup_{t\in(0,1)}|h''(t)|\le\bigg(\sum_i\frac{|c_i|}{1-|c_i|}\bigg)^2e^{\sum_j\frac{|c_j|}{1-|c_j|}}, \tag{A.15}$$

we can bound $\big|r^{(2)}_{n+1}(0)\big|$ as

$$\big|r^{(2)}_{n+1}(0)\big| \le \frac{1}{\varepsilon}\sum_{m=2}^{n+1}|g_m(0)|\bigg|\prod_{l=n+2-m}^{n}[1+\varepsilon\, r_l(0)]^{-1}-1\bigg| \le \sum_{m=2}^{n+1}\frac{\varepsilon C_g\beta}{(1+m\varepsilon)^{d/2}}\,\phi_{m,n}\,e^{\phi_{m,n}}, \tag{A.16}$$

where, by (5.12),

$$\phi_{m,n}=\sum_{l=n+2-m}^{n}\frac{\varepsilon|r_l(0)|}{1-\varepsilon|r_l(0)|}\le\varepsilon\sum_{l=n+2-m}^{n}\frac{cK_3\beta}{(1+l\varepsilon)^{(d-2)/2}}, \tag{A.17}$$

and $e^{\phi_{m,n}}\le e^{cK_3\beta}$ for all $m\le n+1$. Moreover, by Lemma 4.6 with $(a,b)=(\frac{d}{2},\frac{d-2}{2})$, the right-hand side of (A.16) is bounded by $cC_gK_3\beta^2[1+(n+1)\varepsilon]^{-(d-2)/2}$.

Finally, we investigate $r^{(3)}_{n+1}(0)$. As in (A.16), $|f_n(0)^{-1}-1|$ is bounded by

$$\bigg|\prod_{l=1}^{n}[1+\varepsilon r_l(0)]^{-1}-1\bigg|\le\phi_{n+1,n}\,e^{\phi_{n+1,n}}\le cK_3\beta. \tag{A.18}$$

Together with (A.3),

$$\big|r^{(3)}_{n+1}(0)\big|\le\frac{\varepsilon C_g(1+cK_3\beta)\beta}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.19}$$

Assembling (A.13), the statement below (A.17), and (A.19), we obtain the first inequality in (5.12) if β is sufficiently small and $K_3\gg K_1\vee C_g$. □

Advancement of the second inequality in (5.12). Let $k\in\mathcal{A}_{n+1}$. First we bound $\big|\Delta r^{(1)}_{n+1}(k)\big|$ by

$$\frac{1}{\varepsilon}\sum_{m=2}^{n+1}\bigg|g_m(k)-g_m(0)-\frac{a(k)}{\sigma^2}\nabla^2 g_m(0)\bigg|+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}|g_m(k)-g_m(0)|\bigg|\frac{f_{n+1-m}(0)}{f_n(0)}-1\bigg|$$
$$+\frac{1}{\varepsilon}\sum_{m=2}^{n+1}|g_m(k)-g_m(0)|\bigg|\frac{f_{n+1-m}(k)}{f_n(k)}-\frac{f_{n+1-m}(0)}{f_n(0)}\bigg|. \tag{A.20}$$

By (A.5) with $\delta<\Delta'<\frac{d-4}{2}$, the first sum is bounded by

$$\varepsilon\sum_{m=2}^{n+1}\frac{C_g\beta\,a(k)^{1+\Delta'}}{(1+m\varepsilon)^{(d-2)/2-\Delta'}}\le cC_g\beta\,a(k)\bigg[\frac{\log[2+(n+1)\varepsilon]}{[1+(n+1)\varepsilon]}\bigg]^{\Delta'}\le\frac{cC_g\beta\,a(k)}{[1+(n+1)\varepsilon]^{\delta}}. \tag{A.21}$$

Using Lemma 4.6 with $a=b=\frac{d-2}{2}$, (5.12), (A.5) with $\Delta'=0$, (A.6) and (A.16)–(A.17), we bound the second sum of (A.20) by

$$\sum_{m=2}^{n+1}\frac{2\varepsilon C_g\beta\,a(k)}{(1+m\varepsilon)^{(d-2)/2}}\sum_{l=n+2-m}^{n}\frac{\varepsilon cK_3\beta}{(1+l\varepsilon)^{(d-2)/2}}\le\frac{cC_gK_3\beta^2a(k)}{[1+(n+1)\varepsilon]^{(d-2)/2\wedge(d-4)}}, \tag{A.22}$$

where $\frac{d-2}{2}\wedge(d-4)=\frac{d-4}{2}+1\wedge\frac{d-4}{2}\ge2\delta$. By (A.14) and the same argument as in (A.16), the third sum of (A.20) is bounded by

$$\sum_{m=2}^{n+1}\frac{2\varepsilon C_g\beta\,a(k)}{(1+m\varepsilon)^{(d-2)/2}}\bigg|\frac{f_{n+1-m}(0)}{f_n(0)}\bigg|\bigg|\prod_{l=n+2-m}^{n}[1-\varepsilon v_l\,a(k)+\varepsilon s_l(k)]^{-1}-1\bigg|$$
$$\le\sum_{m=2}^{n+1}\frac{2\varepsilon C_g\beta\,a(k)}{(1+m\varepsilon)^{(d-2)/2}}\big(1+\phi_{m,n}\,e^{\phi_{m,n}}\big)\,\psi_{m,n}(k)\,e^{\psi_{m,n}(k)}, \tag{A.23}$$

where

$$\psi_{m,n}(k) = \sum_{l=n+2-m}^{n} \frac{\varepsilon v_l a(k) + \varepsilon |s_l(k)|}{1 - \varepsilon v_l a(k) - \varepsilon |s_l(k)|}. \tag{A.24}$$

As discussed below (A.17), $\phi_{m,n}\, e^{\phi_{m,n}} \le cK_3\beta$. By (5.19) and (5.16),

$$v_l\, a(k) + |s_l(k)| \le \left[v_l + (1 + \varepsilon c K_3\beta) \frac{(1+\varepsilon v_l) K_3 \beta}{(1+l\varepsilon)^\delta} \right] a(k) \le [1 + c\,(K_1 + K_2 + K_3)\beta]\, a(k).$$

Let $q = 1 + c\,(K_1 + K_2 + K_3)\beta$. Then

$$\psi_{m,n}(k) \le \frac{(m-1)\varepsilon q\, a(k)}{1 - \varepsilon q\, a(k)} \le [1 + c\varepsilon\, a(k)](m-1)\varepsilon q\, a(k), \tag{A.25}$$

which is bounded by $\gamma q \big[1 + c\varepsilon \frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon}\big] \log[2 + (n+1)\varepsilon]$, and thus

$$e^{\psi_{m,n}(k)} \le c\, e^{\gamma q \log[2+(n+1)\varepsilon]} \le c\,[1 + (n+1)\varepsilon]^{\gamma q}. \tag{A.26}$$

Since $\delta < 1 \wedge \frac{d-4}{2} - \gamma$, we take q to be close to 1, by letting β be small, to bound (A.23) by

$$cC_g\beta\, a(k) \frac{\log[2 + (n+1)\varepsilon]}{[1 + (n+1)\varepsilon]^{1-\gamma q}} \sum_{m=2}^{n+1} \frac{\varepsilon}{(1+m\varepsilon)^{(d-4)/2}} \le \frac{cC_g\beta\, a(k)}{[1+(n+1)\varepsilon]^\delta}. \tag{A.27}$$

By (A.20)–(A.22) and (A.27), $\big|\Delta r^{(1)}_{n+1}(k)\big|$ is bounded by the second bound of (5.12) if β is sufficiently small and $K_3 \gg C_g$.

Next we use (A.4) and (A.16)–(A.17) to bound $\big|\Delta r^{(2)}_{n+1}(k)\big|$ by

$$\begin{aligned}
&\frac{1}{\varepsilon} \sum_{m=2}^{n+1} |g_m(0)| \left| \frac{f_{n+1-m}(0)}{f_n(0)} \right| \left| \prod_{l=n+2-m}^{n} [1 - \varepsilon v_l\, a(k) + \varepsilon\, s_l(k)]^{-1} - 1 - (m-1)\varepsilon v_{n+1}\, a(k) \right| \\
&\qquad + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} |g_m(0)| \left| \frac{f_{n+1-m}(0)}{f_n(0)} - 1 \right| (m-1)\varepsilon v_{n+1}\, a(k) \\
&\le \sum_{m=2}^{n+1} \frac{\varepsilon C_g \beta\,(1 + \phi_{m,n}\, e^{\phi_{m,n}})}{(1+m\varepsilon)^{d/2}} \left| \prod_{l=n+2-m}^{n} [1 - \varepsilon v_l\, a(k) + \varepsilon\, s_l(k)]^{-1} - 1 - (m-1)\varepsilon v_{n+1}\, a(k) \right| \\
&\qquad + \sum_{m=2}^{n+1} \frac{\varepsilon C_g \beta v_{n+1}\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}\, \phi_{m,n}\, e^{\phi_{m,n}}.
\end{aligned} \tag{A.28}$$

By (5.12), (5.19) and Lemma 4.6, the second sum in (A.28) is bounded by the same bound as in (A.22). We use (A.15) to bound the first sum in (A.28) by

$$\sum_{m=2}^{n+1} \frac{c\varepsilon C_g \beta}{(1+m\varepsilon)^{d/2}} \left[\psi_{m,n}(k)^2\, e^{\psi_{m,n}(k)} + \sum_{l=n+2-m}^{n} \big[\varepsilon |v_l - v_{n+1}|\, a(k) + \varepsilon |s_l(k)| \big] \right].$$

By (A.25)–(A.26), the contribution from $\psi_{m,n}(k)^2\, e^{\psi_{m,n}(k)}$ is bounded by the same bound in (A.27). By (5.10), (5.16) and Lemma 4.6, the other contribution is bounded by

$$\sum_{m=2}^{n+1} \frac{c\varepsilon C_g \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{n} \left[\frac{c\varepsilon K_2 \beta\, a(k)}{(1+l\varepsilon)^{(d-4)/2}} + \frac{c\varepsilon K_3 \beta\, a(k)}{(1+l\varepsilon)^\delta} \right] \le \frac{cC_g\,(K_2+K_3)\beta^2 a(k)}{[1+(n+1)\varepsilon]^\delta}.$$

Therefore $\big|\Delta r^{(2)}_{n+1}(k)\big|$ is also bounded by the second bound of (5.12) if β is sufficiently small.

Finally, we bound $\big|\Delta r^{(3)}_{n+1}(k)\big|$ by

$$\left|\frac{e_{n+1}(k)}{\varepsilon f_n(0)}\right|\left|\frac{f_n(0)}{f_n(k)}-1\right|+\left|\frac{e_{n+1}(k)-e_{n+1}(0)}{\varepsilon f_n(0)}\right|. \tag{A.29}$$

By (A.18), $|f_n(0)|\geq 1-cK_3\beta$. By (A.25)–(A.26), $\big|\frac{f_n(0)}{f_n(k)}-1\big|$ is bounded by $\psi_{n+1,n}(k)\,e^{\psi_{n+1,n}(k)}\leq c\,[1+(n+1)\varepsilon]^{1+\gamma q}\,a(k)$. Together with (A.3), (A.29) is bounded by

$$\frac{c\varepsilon C_g\beta\,a(k)}{[1+(n+1)\varepsilon]^{(d-2)/2-\gamma q}}. \tag{A.30}$$

Suppose that q is sufficiently close to 1. Since $\gamma+\delta<1\wedge\frac{d-4}{2}$, this implies the desired bound if $K_3\gg C_g$. This completes the advancement of the second inequality in (5.12), and thus the advancement of (H3). $\square$

Advancement of (H4). Let $k\notin\mathcal{A}_{n+1}$. To advance (H4), we rewrite (5.3) as

$$f_{n+1}(k)=\left[g_1(k)+\sum_{m=2}^{n+1}g_m(k)\right]f_n(k)+W_{n+1}(k)+e_{n+1}(k),$$

where

$$W_{n+1}(k)=\sum_{m=2}^{n+1}g_m(k)\,[-f_n(k)+f_{n+1-m}(k)]=\sum_{m=2}^{n+1}g_m(k)\sum_{l=n+2-m}^{n}[f_{l-1}(k)-f_l(k)]. \tag{A.31}$$

Furthermore,

$$\begin{aligned}g_1(k)+\sum_{m=2}^{n+1}g_m(k)=&1-\left[\lambda-\frac{1}{\sigma^2\varepsilon}\sum_{m=2}^{n+1}\nabla^2 g_m(0)\right]\varepsilon\,a(k)+\left[(\lambda-1)\varepsilon+\sum_{m=2}^{n+1}g_m(0)\right]\\&+\sum_{m=2}^{n+1}\left[g_m(k)-g_m(0)-\frac{a(k)}{\sigma^2}\nabla^2 g_m(0)\right]\\=&1-N_{n+1}\,\varepsilon\,a(k)+\varepsilon r^{(1)}_{n+1}(0)+X_{n+1}(k),\end{aligned}$$

where $X_{n+1}(k)=\sum_{m=2}^{n+1}\big[g_m(k)-g_m(0)-\frac{a(k)}{\sigma^2}\nabla^2 g_m(0)\big]$, and we recall that N_n is the numerator of (5.8). Therefore,

$$f_{n+1}(k)=f_n(k)\,\big[1-N_{n+1}\,\varepsilon\,a(k)+\varepsilon r^{(1)}_{n+1}(0)+X_{n+1}(k)\big]+W_{n+1}(k)+e_{n+1}(k). \tag{A.32}$$

By (A.9), $|N_{n+1}-\lambda|\leq cC_g\beta$. By (A.13), $|r^{(1)}_{n+1}(0)|\leq cK_1\beta[1+(n+1)\varepsilon]^{-(d-2)/2}$. The term $X_{n+1}(k)$ equals the first term of (A.20) multiplied by ε, and is bounded by $c\varepsilon C_g\beta\,a(k)^{1+\Delta'}$ if $\Delta'<\frac{d-4}{2}$ (see the left-hand side of (A.21)). We prove below

$$|W_{n+1}(k)|\leq\frac{c\varepsilon C_g(1+K_3\beta+K_5)\beta\,a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.33}$$

We advance (H4) assuming (A.33). By the same reason as in [17, Section 3.4], it suffices to consider only $n\varepsilon>T$ for some large T.

First we claim that the first inequality in (5.13) with $m=n$, i.e.,

$$|f_n(k)|\leq\frac{K_4\,a(k)^{-2-\rho}}{(1+n\varepsilon)^{d/2}} \tag{A.34}$$

holds not only for $k \notin \mathcal{A}_n$ but also for $k \in \mathcal{A}_n \setminus \mathcal{A}_{n+1}$: (A.34) holds for all $k \notin \mathcal{A}_{n+1}$. To prove (A.34) when $k \in \mathcal{A}_n \setminus \mathcal{A}_{n+1}$ (so that $\gamma\frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon} < a(k) \le \gamma\frac{\log(2+n\varepsilon)}{1+n\varepsilon}$), we use Lemma 5.2 to obtain

$$\begin{aligned} |f_n(k)| &\le ce^{-n\varepsilon q\, a(k)} \le c\,[1+(n+1)\varepsilon]^{-\frac{n\varepsilon q\gamma}{1+(n+1)\varepsilon}} \le c\,(1+n\varepsilon)^{-q'\gamma} \\ &\le \frac{c}{(1+n\varepsilon)^{d/2}} \frac{(1+n\varepsilon)^{2+\rho}}{(1+n\varepsilon)^{q'\gamma-(\frac{d-4}{2}-\rho)}} \le \frac{c\,a(k)^{-2-\rho}}{(1+n\varepsilon)^{d/2}}, \end{aligned} \tag{A.35}$$

where q' was taken to be close to 1, by letting T/ε sufficiently large and letting $q = 1 - O(\beta)$ sufficiently close to 1, such that $q'\gamma > \frac{d-4}{2} - \rho$ holds. It thus suffices to take $K_4 \gg 1$.

Thanks to the above extension of the first inequality in (5.13), together with (A.32)–(A.33),

$$\begin{aligned} |f_{n+1}(k) - f_n(k)| \le &\frac{\varepsilon K_4\, a(k)^{-2-\rho}}{(1+n\varepsilon)^{d/2}} \bigg[(\lambda + cC_g\beta)\, a(k) + \frac{cK_1\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}} + cC_g\beta\, a(k)^{1+\Delta'}\bigg] \\ &+ \frac{c\varepsilon C_g(1+K_3\beta+K_5)\beta\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}} + \frac{\varepsilon^2 C_g\beta}{[1+(n+1)\varepsilon]^{d/2}}. \end{aligned} \tag{A.36}$$

By (5.18) and the inequality $[1+(n+1)\varepsilon]^{-(d-2)/2} \le [\gamma\log[2+(n+1)\varepsilon]/a(k)]^{-(d-2)/2} \le c\,a(k)$, the second bound in (5.13) with $m = n+1$ holds if β is sufficiently small and $K_5 \gg K_4$.

As in (A.36),

$$\begin{aligned} |f_{n+1}(k)| \le &\frac{K_4\, a(k)^{-2-\rho}}{(1+n\varepsilon)^{d/2}} \bigg[|1 - N_{n+1}\,\varepsilon\, a(k)| + \frac{c\varepsilon K_1\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}} + c\varepsilon C_g\beta\, a(k)^{1+\Delta'}\bigg] \\ &+ \frac{c\varepsilon C_g(1+K_3\beta+K_5)\beta\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}} + \frac{\varepsilon^2 C_g\beta}{[1+(n+1)\varepsilon]^{d/2}}. \end{aligned} \tag{A.37}$$

To prove the first bound in (5.13) with $m = n+1$, it suffices to show that

$$\bigg[|1 - N_{n+1}\,\varepsilon\, a(k)| + \frac{c\varepsilon K_1\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}} + c\varepsilon C_g\beta\, a(k)^{1+\Delta'}\bigg] \bigg[\frac{1+(n+1)\varepsilon}{1+n\varepsilon}\bigg]^{d/2} \tag{A.38}$$

is strictly less than 1. To achieve this uniformly in $\varepsilon \le 1$, we will use separate arguments for $a(k) \le 1/2$ and $a(k) > 1/2$. These arguments will be valid only when T is large enough.

Suppose that $a(k) \le 1/2$. Since $N_{n+1} = 1 + O(\beta)$ by (5.18) and (A.9), we have $|1 - N_{n+1}\,\varepsilon\, a(k)| = 1 - N_{n+1}\,\varepsilon\, a(k)$ for sufficiently small β. Therefore (A.38) is bounded by

$$\begin{aligned} &\bigg[1 - (1-c\beta)\varepsilon\, a(k) + \frac{c\beta\varepsilon}{[1+(n+1)\varepsilon]^{(d-2)/2}}\bigg]\Big(1 + \frac{c\varepsilon}{1+n\varepsilon}\Big) \\ &\le 1 - \varepsilon\bigg[(1-c\beta)\frac{\gamma\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon} - \frac{c\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}} - \frac{c}{1+n\varepsilon}\bigg] < 1, \end{aligned}$$

if T is sufficiently large.

The above argument also applies when $1 - \varepsilon\, a(k) > 1 - (2-\eta)\varepsilon > 0$ and $\beta \ll 1$ (depending on η). Therefore we assume $1 - (2-\eta)\varepsilon \le 0$ and $a(k) > 1/2$. In this case, since $\varepsilon \le 1$, we have

$$|1 - \varepsilon\, a(k)| \le [(2-\eta)\varepsilon - 1] \vee \Big(1 - \frac{\varepsilon}{2}\Big) \le 1 - \Big(\eta \wedge \frac{\varepsilon}{2}\Big).$$

Since $N_{n+1} = 1 + O(\beta)$, (A.38) is bounded by

$$\begin{aligned} &\bigg[1 - \Big(\eta \wedge \frac{\varepsilon}{2}\Big) + c\beta\varepsilon + \frac{c\beta\varepsilon}{[1+(n+1)\varepsilon]^{(d-2)/2}}\bigg]\Big(1 + \frac{c\varepsilon}{1+n\varepsilon}\Big) \\ &\le 1 - \bigg[\Big(\eta \wedge \frac{\varepsilon}{2}\Big) - c\beta\varepsilon - \frac{c\beta\varepsilon}{[1+(n+1)\varepsilon]^{(d-2)/2}} - \frac{c\varepsilon}{1+n\varepsilon}\bigg] < 1, \end{aligned}$$

if T is large enough. This completes the advancement of (H4) under the assumption that (A.33) holds.

Proof of (A.33). Given $k \notin \mathcal{A}_{n+1}$, let $\mu = \mu(k) = \max\{l \in \mathbb{N} : k \in \mathcal{A}_l\}$. For $l \le \mu$, f_l is in the domain of (H3), while for $l > \mu$, f_l is in the domain of (H4). We divide the sum over l in (A.31) into two parts, corresponding respectively to $l \le \mu$ and $l > \mu$, yielding $W_{n+1}(k) = W^{\le}_{n+1}(k) + W^{>}_{n+1}(k)$, where

$$|W^{\le}_{n+1}(k)| \le \sum_{m=n+2-\mu}^{n+1} \frac{\varepsilon^2 C_g \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{\mu} |f_{l-1}(k) - f_l(k)|, \tag{A.39}$$

$$|W^{>}_{n+1}(k)| \le \sum_{m=2}^{n+1} \frac{\varepsilon^2 C_g \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=\mu\vee(n+1-m)+1}^{n} |f_{l-1}(k) - f_l(k)|. \tag{A.40}$$

By (H4) and Lemma 4.6,

$$|W^{>}_{n+1}(k)| \le \sum_{m=2}^{n+1} \frac{\varepsilon^2 C_g \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{n} \frac{\varepsilon K_5\, a(k)^{-1-\rho}}{(1+l\varepsilon)^{d/2}} \le \frac{c\varepsilon C_g K_5 \beta\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.41}$$

For $W^{\le}_{n+1}(k)$, we apply (H3) to obtain

$$\begin{aligned} |f_{l-1}(k) - f_l(k)| &= |f_l(k)| \left| \left[1 - \varepsilon v_l\, a(k) + \varepsilon[r_l(k) - r_l(0)] + \varepsilon\, r_l(0)\right]^{-1} - 1 \right| \\ &\le c\varepsilon e^{-l\varepsilon q\, a(k)} \left[a(k) + \frac{K_3 \beta}{(1+l\varepsilon)^{(d-2)/2}} \right], \end{aligned} \tag{A.42}$$

with $q = 1 - O(\beta)$. We fix a small $r > 0$ and separate $\sum_m$ in (A.39) into $\sum_{m>r(n+1)}$ and $\sum_{m\le r(n+1)}$ (the latter range may be empty depending on μ). The contribution due to $\sum_{m>r(n+1)}$ is bounded by

$$\begin{aligned} &\frac{c\varepsilon^2 C_g \beta}{[1+(n+1)\varepsilon]^{d/2}} \sum_{m=r(n+1)+1}^{n+1} e^{-(n+2-m)\varepsilon q\, a(k)}\, (1+K_3\beta) \\ &\qquad \le \frac{c\varepsilon C_g \beta (1+K_3\beta)\, a(k)^{-1}}{[1+(n+1)\varepsilon]^{d/2}} \le \frac{c\varepsilon C_g \beta (1+K_3\beta)\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \end{aligned} \tag{A.43}$$

For the contribution from $\sum_{m\le r(n+1)}$, we use the inequalities $m \le r\,(n+1)$, $n+2-m \le l \le \mu$ and $\mu \le n$ to obtain, as in (A.35),

$$\begin{aligned} e^{-l\varepsilon q\, a(k)} &\le [2+(\mu+1)\varepsilon]^{-\frac{l\varepsilon q\gamma}{1+(\mu+1)\varepsilon}} \le (1+l\varepsilon)^{-\frac{l\varepsilon q\gamma}{1+(\mu+1)\varepsilon}} \le (1+l\varepsilon)^{-\frac{(1-r)(n+1)\varepsilon q\gamma}{1+(n+1)\varepsilon}} = (1+l\varepsilon)^{-q'\gamma} \\ &\le \frac{c}{(1+l\varepsilon)^{d/2}} \frac{[1+(\mu+1)\varepsilon]^{2+\rho}}{[1+(\mu+1)\varepsilon]^{q'\gamma-(\frac{d-4}{2}-\rho)}} \le \frac{c\, a(k)^{-2-\rho}}{(1+l\varepsilon)^{d/2}} \end{aligned}$$

where $q' = \frac{(1-r)(n+1)\varepsilon q}{1+(n+1)\varepsilon}$ was taken to be close to 1 such that$q'\gamma > \frac{d-4}{2} - \rho$ holds. Therefore, by Lemma 4.6, the contribution from $\sum_{m\le r(n+1)}$ is bounded by

$$\sum_{m=n+2-\mu}^{r(n+1)} \frac{\varepsilon^2 C_g \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{\mu} \frac{c\varepsilon(1+K_3\beta)\, a(k)^{-1-\rho}}{(1+l\varepsilon)^{d/2}} \le \frac{c\varepsilon C_g \beta(1+K_3\beta)\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.44}$$

By (A.41) and (A.43)–(A.44), we obtain the inequality (A.33). □

A.2 Advancement below and at four dimensions

The proofs of Lemmas 5.1–5.4 and the advancement of the induction hypotheses for $d \le 4$ remain almost unchanged, except for the factors β_T and $\hat{\beta}'_T$ in (4.4)–(4.5) and (5.25)–(5.27). We simply replace β by $\hat{\beta}'_T$

in the proofs of Lemmas 5.1–5.3, and by β_T in the proof of Lemma 5.4 (we also replace $d/2$ by $2+\omega$ in the proof of Lemma 5.1). In the advancement of the first inequality in (5.26), we use (A.7) together with (A.4) and (A.6) with β replaced by β_T. Since $n\varepsilon \le T\log T$, we obtain

$$|\lambda_{n+1}-\lambda_n| \le \frac{\varepsilon C_g \beta_T}{[1+(n+1)\varepsilon]^{d/2}} + \frac{\varepsilon K_1 \hat{\beta}'_T}{(1+n\varepsilon)^{2+\omega}} \sum_{m=2}^{n} \frac{\varepsilon C_g \beta_T}{(1+m\varepsilon)^{(d-2)/2}} \le \frac{\varepsilon C_g (1+cK_1\beta_1 T^{-\mu})\hat{\beta}'_T}{[1+(n+1)\varepsilon]^{2+\omega}},$$

where $\mu \in (0,\alpha)$. Similarly, we can advance the second inequality in (5.26) and the first inequality in (5.27).

We need a little more care in advancing the second inequality in (5.27) and the inequalities in (5.13). The second inequality in (5.27) is rewritten as in (A.12), and each term of that expression is bounded as in (A.20)–(A.30) for $d > 4$. There is no difficulty to extend those bounds to the lower-dimensional case when we have β^2 as in (A.22); that term is replaced by $\beta_T\hat{\beta}'_T$ for $d \le 4$, and we use β_T to control the convolution in (A.22), where the power exponent of $1+l\varepsilon$ is $1+\omega$ for $d \le 4$. When we have only one β as, for example, in (A.27), we use $\gamma+\delta < \omega$ and $\mu' < \alpha-\omega$ to obtain the desired bound

$$cC_g\beta_T\, a(k)[1+(n+1)\varepsilon]^{\frac{4-d}{2}+\gamma q}\log[2+(n+1)\varepsilon] \le \frac{cC_g\hat{\beta}'_T T^{-(\omega-\gamma-\delta)}a(k)}{[1+(n+1)\varepsilon]^{\delta}},$$

where we assumed that q is close to 1.

A similar argument applies to the advancement of the inequalities in (5.13). However, since $\rho < -\frac{4-d}{2}$, we cannot use the trivial inequality $a(k) \le 2$ to obtain, for example, the lower-dimensional version of (A.43). To overcome this difficulty, we use the factor β_T in the bound on $g_m(k)$ to obtain

$$\beta_T \le \frac{\beta_1 T^{-bd}}{a(k)^{\rho}}\left[\frac{1+(n+1)\varepsilon}{\gamma\log[2+(n+1)\varepsilon]}\right]^{-\rho} \le \frac{c\hat{\beta}'_T T^{-\omega+\frac{d-4}{2}-\rho}}{a(k)^{\rho}} \le \frac{c\hat{\beta}'_T T^{-(\omega-\gamma)}}{a(k)^{\rho}},$$

where we used $\mu' < \alpha-\omega$ and $\frac{d-4}{2}-\rho < \gamma < \omega$. This completes the advancement of the induction hypotheses for $d \le 4$. □