Electronic Journal of Probability

Vol. 9 (2004), Paper no. 24, pages 710-769.

Journal URL
http://www.math.washington.edu/~ejpecp/

Gaussian Scaling for the Critical Spread-out Contact Process above the Upper Critical Dimension

Remco van der Hofstad[1] and **Akira Sakai**[2]

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 behaviour 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 in such a way 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 their respective mean-field values. The results in this paper provide crucial ingredients to prove convergence of the finite-dimensional distributions for the contact process towards those for the canonical measure of super-Brownian motion, which we defer to a sequel of this paper.

The results in this paper also apply to oriented percolation, for which we reprove some of the results in [20] and extend the results to the local mean-field setting described above when $d \leq 4$.

Submitted to EJP on August 11, 2003. Final version accepted on August 30, 2004.

[1]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`

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

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., [22]).

Sakai [26, 27] 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 the main ingredient is the proof of 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. Since these Gaussian upper bounds imply the so-called triangle condition in [3], 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 [22, Chapter I.4] for an extensive account of the contact process on a tree.

Recently, van der Hofstad and Slade [20] 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 the r points, and then take $n \uparrow \infty$. That is, the finite-dimensional 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 [20] is based on the lace expansion and the inductive method of [19]. Important ingredients in [20] are detailed asymptotics and estimates of the oriented percolation two-point function. The proof for the higher-point functions then follows by deriving a lace expansion for the r-point functions together with an induction argument 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 uses the same strategy as applied to oriented percolation with $\varepsilon = 1$, i.e., an application of the lace expansion and the inductive method. However, to obtain the results for $\varepsilon \ll 1$, we use a different lace expansion from the two expansions used in [20, Sections 3.1–3.2], and modify the induction hypotheses of [19] to incorporate the ε-dependence. In order to extend the results from infrared bounds (as in [27]) to precise asymptotics (as in [20]), it is imperative to prove that the properly scaled lace expansion coefficients converge to a certain continuum limit. We can think of this continuum limit as giving rise to a lace expansion in continuous time, even though our proof is not based on the arising partial differential equation. In the proof that the continuum limit exists, we make heavy use of convergence results in [4] which show that the discretized contact process converges to the original continuous-time contact process.

In a sequel to this paper [18], we use the results proved here as a key ingredient in the proof that the finite-dimensional distributions of the critical contact process above four dimensions converge to those of the canonical measure of super-Brownian motion, as was proved in [20] 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 neighbours, at rate $\lambda \sum_{y \in \mathbf{C}_t} D(x-y)$, where $\lambda \geq 0$ is the infection rate and $D(x-y)$ represents the strength of the interaction between x and y. We denote by $\mathbb{P}^\lambda$ the associated probability measure.

The function D is a probability distribution over $\mathbb{Z}^d$ that is symmetric with respect to the lattice symmetries, and satisfies certain assumptions that 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 there are L-independent constants $C, C_1, C_2 \in (0, \infty)$ such that $D(o) = 0$, $\sup_{x\in\mathbb{Z}^d} D(x) \leq CL^{-d}$ and $C_1 L \leq \sigma \leq C_2 L$, where σ^2 is the variance of D:

$$\sigma^2 = \sum_{x\in\mathbb{Z}^d} |x|^2 D(x), \tag{1.1}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$. Moreover, we require that there is a $\Delta > 0$ such that

$$\sum_{x\in\mathbb{Z}^d} |x|^{2+2\Delta} D(x) \leq CL^{2+2\Delta}. \tag{1.2}$$

See Section 5.1.1 for the precise assumptions on D. A simple example of D is the uniform distribution over the cube of side length $2L$, excluding its center:

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

where $\|x\|_\infty = \sup_i |x_i|$ for $x = (x_1, \dots, x_d)$.

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 located at $o \in \mathbb{Z}^d$ at time 0.

By an extension of the results in [4, 10] to the spread-out contact process, there exists a unique critical value $\lambda_c \in (0, \infty)$ such that

$$\chi(\lambda) = \int_0^\infty dt\ \hat{\tau}_t^\lambda(0) \begin{cases} < \infty, & \text{if } \lambda < \lambda_c, \\ = \infty, & \text{if } \lambda \geq \lambda_c, \end{cases} \qquad \theta(\lambda) \equiv \lim_{t\uparrow\infty} \mathbb{P}^\lambda(\mathbf{C}_t \neq \varnothing) \begin{cases} = 0, & \text{if } \lambda \leq \lambda_c, \\ > 0, & \text{if } \lambda > \lambda_c, \end{cases} \tag{1.5}$$

where we denote the Fourier transform of a summable function $f : \mathbb{Z}^d \mapsto \mathbb{R}$ by

$$\hat{f}(k) = \sum_{x\in\mathbb{Z}^d} f(x)\, e^{ik\cdot x} \qquad (k \in [-\pi, \pi]^d). \tag{1.6}$$

We next describe our results for the sufficiently spread-out contact process at $\lambda = \lambda_c$ for $d > 4$.

1.2.1 Results above four dimensions

We now state the results for the two-point function. In the statements, σ and Δ are defined in (1.1)–(1.2), and we write $\|f\|_\infty = \sup_{x\in\mathbb{Z}^d} |f(x)|$ for a function f on $\mathbb{Z}^d$.

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

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

$$\frac{1}{\hat{\tau}_t^{\lambda_c}(0)} \sum_{x\in\mathbb{Z}^d} |x|^2 \tau_t^{\lambda_c}(x) = v\sigma^2 t \big[1 + O((1+t)^{-\delta})\big], \tag{1.8}$$

$$C_1 L^{-d}(1+t)^{-d/2} \leq \|\tau_t^{\lambda_c}\|_\infty \leq e^{-t} + C_2 L^{-d}(1+t)^{-d/2}, \tag{1.9}$$

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

The above results correspond to [20, Theorem 1.1], where the two-point function for sufficiently spread-out critical oriented percolation with $d > 4$ was proved to obey similar behaviour. The proof in [20] is based on the inductive method of [19]. We apply a modified version of this induction method to prove Theorem 1.1. The proof also reveals that

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

In a sequel to this paper [17], we will investigate the critical point in more detail and prove that

$$\lambda_c - 1 = \sum_{n=2}^{\infty} D^{*n}(o) + O(L^{-2d}), \tag{1.11}$$

holds for $d > 4$, where D^{*n} is the n-fold convolution of D in $\mathbb{Z}^d$. In particular, when D is defined by (1.3), we obtain (see [17, Theorem 1.2])

$$\lambda_c - 1 = L^{-d} \sum_{n=2}^{\infty} U^{\star n}(o) + O(L^{-d-1}), \tag{1.12}$$

where U is the uniform probability density over $[-1, 1]^d \subset \mathbb{R}^d$, and $U^{\star n}$ is the n-fold convolution of U in $\mathbb{R}^d$. The above expression was already obtained in [8], but with a weaker error estimate.

Let γ and β be the critical exponents for the quantities in (1.5), defined as

$$\chi(\lambda) \sim (\lambda_c - \lambda)^{-\gamma} \quad (\lambda < \lambda_c), \qquad \theta(\lambda) \sim (\lambda - \lambda_c)^{\beta} \quad (\lambda > \lambda_c), \tag{1.13}$$

where we use "$\sim$" in an appropriate sense. For example, the strongest form of $\chi(\lambda) \sim (\lambda_c - \lambda)^{-\gamma}$ is that there is a $C \in (0, \infty)$ such that

$$\chi(\lambda) = [C + o(1)]\,(\lambda_c - \lambda)^{-\gamma}, \tag{1.14}$$

where $o(1)$ tends to 0 as $\lambda \uparrow \lambda_c$. Other examples are the weaker form

$$\exists\, C_1, C_2 \in (0, \infty): \quad C_1(\lambda_c - \lambda)^{-\gamma} \le \chi(\lambda) \le C_2(\lambda_c - \lambda)^{-\gamma}, \tag{1.15}$$

and the even weaker form

$$\chi(\lambda) = (\lambda - \lambda_c)^{-\gamma + o(1)}. \tag{1.16}$$

See also [22, p.70] for various ways to define the critical exponents.

As discussed for oriented percolation in [20, Section 1.2.1], (1.7) and (1.9) imply finiteness at $\lambda = \lambda_c$ of the triangle function

$$\nabla(\lambda) = \int_0^\infty dt \int_0^t ds \sum_{x,y \in \mathbb{Z}^d} \tau_t^\lambda(y)\, \tau_{t-s}^\lambda(y - x)\, \tau_s^\lambda(x). \tag{1.17}$$

Extending the argument in [24] for oriented percolation to the continuous-time setting, we conclude that $\nabla(\lambda_c) < \infty$ implies the triangle condition of [1, 2, 3], under which γ and β are both equal to 1 in the form given in (1.15), independently of the value of d [3]. Since these d-independent values also arise on the tree [29, 34], we call them the mean-field values. The results (1.7)–(1.8) also show that the critical exponents ν and η, defined as

$$\frac{1}{\hat{\tau}_t^{\lambda_c}(0)} \sum_{x \in \mathbb{Z}^d} |x|^2 \tau_t^{\lambda_c}(x) \sim t^{2\nu}, \qquad \hat{\tau}_t^{\lambda_c}(0) \sim t^{\eta}, \tag{1.18}$$

take on the mean-filed values $\nu = 1/2$ and $\eta = 0$, in the stronger form given in (1.14). The result $\eta = 0$ proves that the statement in [22, Proposition 4.39] on the tree also holds for sufficiently spread-out contact process on $\mathbb{Z}^d$ for $d > 4$. See the remark below [22, Proposition 4.39]. Furthermore, following from bounds established in the course of the proof of Theorem 1.1, we can extend the aforementioned result of [3], i.e., $\gamma = 1$ in the form given in (1.15), to the precise asymptotics as in (1.14). We will prove this in Section 2.5.

So far, $d > 4$ is a sufficient condition for the mean-field behaviour for the spread-out contact process. It has been shown, using the hyperscaling inequalities in [28], that $d \geq 4$ is also a necessary condition for the mean-field behaviour. Therefore, the upper critical dimension for the spread-out contact process is 4, and one can expect log corrections in $d = 4$.

In [18], we will investigate the higher-point functions of the critical spread-out contact process for $d > 4$. These higher-point functions are defined for $\vec{t} \in [0, \infty)^{r-1}$ and $\vec{x} \in \mathbb{Z}^{d(r-1)}$ by

$$\tau_{\vec{t}}^{\lambda}(\vec{x}) = \mathbb{P}^{\lambda}(x_i \in \mathbf{C}_{t_i} \ \forall i = 1, \ldots, r-1). \tag{1.19}$$

The proof will be based on a lace expansion that expresses the r-point function in terms of s-point functions with $s < r$. On the arising equation, we will then perform induction in r, with the results for $r = 2$ given by Theorem 1.1. We discuss the extension to the higher point functions in somewhat more detail in Section 2.2, where we discuss the lace expansion. In order to bound the lace expansion coefficients for the higher point functions, the upper bounds in (1.7) for $k = 0$ and in (1.9) are crucial.

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 exhibit the mean-field behaviour as long as L remains finite, and Gaussian asymptotics are not expected to hold in this case. However, we can prove local Gaussian behaviour when the range grows in time as

$$L_T = L_1 T^b \qquad (T \geq 1), \tag{1.20}$$

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

$$\alpha = bd + \frac{d-4}{2} > 0. \tag{1.21}$$

Our main result is the following.

Theorem 1.2. *Let $d \leq 4$ and $\delta \in (0, 1 \wedge \Delta \wedge \alpha)$. Then, there is a $\lambda_T = 1 + O(T^{-\mu})$ for some $\mu \in (0, \alpha - \delta)$ such that, for sufficiently large L_1, there are positive and finite constants $C_1 = C_1(d)$ and $C_2 = C_2(d)$ such that, for every $0 < t \leq \log T$,*

$$\hat{\tau}_{Tt}^{\lambda_T}\Big(\frac{k}{\sqrt{\sigma_T^2 Tt}}\Big) = e^{-\frac{|k|^2}{2d}} \big[1 + O(T^{-\mu}) + O(|k|^2 (1+Tt)^{-\delta})\big], \tag{1.22}$$

$$\frac{1}{\hat{\tau}_{Tt}^{\lambda_T}(0)} \sum_{x \in \mathbb{Z}^d} |x|^2 \tau_{Tt}^{\lambda_T}(x) = \sigma_T^2 Tt \big[1 + O(T^{-\mu}) + O((1+Tt)^{-\delta})\big], \tag{1.23}$$

$$C_1 L_T^{-d} (1+Tt)^{-d/2} \leq \|\tau_{Tt}^{\lambda_T}\|_\infty \leq e^{-Tt} + C_2 L_T^{-d} (1+Tt)^{-d/2}, \tag{1.24}$$

with the error estimate in (1.22) uniform in $k \in \mathbb{R}^d$ with $|k|^2 / \log(2 + Tt)$ sufficiently small.

The upper bound on t in the statement can be replaced by any slowly varying function. However, we use $\log T$ to make the statement more concrete. The proof of Theorem 1.2 follows the same steps as the proof of Theorem 1.1.

First, we give a heuristic explanation of how (1.21) arises. Recall that, for $d > 4$, $\triangledown(\lambda_c) < \infty$ is a sufficient condition for the mean-field behaviour. For $d \le 4$, since $\triangledown(\lambda_T)$ cannot be defined in full space-time as in (1.17), we modify the triangle function as

$$\triangledown_{\rm ld}(\lambda_T) = \int_0^{T\log T} dt \int_0^t ds \sum_{x,y\in\mathbb{Z}^d} \tau_t^{\lambda_T}(y)\,\tau_{t-s}^{\lambda_T}(y-x)\,\tau_s^{\lambda_T}(x). \tag{1.25}$$

Using the upper bounds in (1.22) for $k = 0$ and in (1.24), we obtain

$$\triangledown_{\rm ld}(\lambda_T) \le C^2 \int_0^{T\log T} dt \int_0^t ds\ (e^{-tT} + C_2 L_T^{-d} T^{-d/2}) \le O(T^{-2}) + O(T^{2-bd-d/2}\log^2 T), \tag{1.26}$$

which is finite for all T whenever $bd > \frac{4-d}{2}$. We can find a similar argument in [33, Section 14].

Next, we compare the ranges needed in our results and in the results of Durrett and Perkins [8], in which the convergence of the rescaled contact process to super-Brownian motion was proved. As in (1.21) we need $bd > \frac{4-d}{2}$, while in [8] $bd = 1$ for all $d \ge 3$. For $d = 2$, which is a critical case in the setting of [8], the model with range $L_T^2 = T\log T$ was also investigated. In comparison, we are allowed to use ranges that grow to infinity slower than the ranges in [8] when $d \ge 3$, but the range for $d = 2$ in our results needs to be larger than that in [8]. It would be of interest to investigate whether Theorem 1.2 holds when $L_T^2 = T\log T$ (or even smaller) by adapting our proofs.

Finally, we give a conjecture on the asymptotics of λ_T as $T \uparrow \infty$. The role of λ_T is a sort of critical value for the contact process in the finite-time interval $[0, T\log T]$, and hence λ_T approximates the real critical value $\lambda_{c,T}$ that also converges to 1 in the mean-field limit $T \uparrow \infty$. We believe that the leading term of $\lambda_{c,T} - 1$, say c_T, is equal to that of $\lambda_T - 1$. As we will discuss below in Section 5.4, λ_T satisfies a type of recursion relation (5.41). We expect that, for $d \le 4$, we may employ the methods in [17] to obtain

$$\lambda_T = 1 + [1 + O(T^{-\mu})] \int_0^{T\log T} dt \int_{[-\pi,\pi]^d} \frac{d^dk}{(2\pi)^d}\ \hat{D}_T^2(k)\, e^{-[1-\hat{D}_T(k)]t}, \tag{1.27}$$

where D_T equals D with range L_T. (In fact, the exponent μ could be replaced by any positive number strictly smaller than α.) The integral with respect to $t \in \mathbb{R}_+$ converges when $d > 2$, and hence we may obtain for sufficiently large T that

$$\begin{aligned}\lambda_T &= 1 + [1 + O(T^{-\mu})]\bigg[\int_{[-\pi,\pi]^d} \frac{d^dk}{(2\pi)^d}\ \frac{\hat{D}_T^2(k)}{1-\hat{D}_T(k)} + O(T^{-bd-\frac{d-2}{2}})\bigg] \\ &= 1 + \sum_{n=2}^{\infty} D_T^{*n}(o) + O(L_T^{-d-\frac{\mu}{b}\wedge\frac{d-2}{2b}}),\end{aligned} \tag{1.28}$$

where we use (1.20) and the fact that the sum in (1.28) is $O(L_T^{-d})$. Based on our belief mentioned above, this would be a stronger result than the result in [8] when $d = 3, 4$, where $c_T = \sum_{n=2}^{\infty} D_T^{*n}(o)$. However, to prove this conjecture, we may require serious further work using block constructions used in [8].

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. These results apply in particular to oriented percolation, which is a special example of the discretized contact process. In Section 2.2, we briefly explain the lace expansion for the discretized contact process, and state the bounds on the lace expansion coefficients in Section 2.3. 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 1.1–1.2.

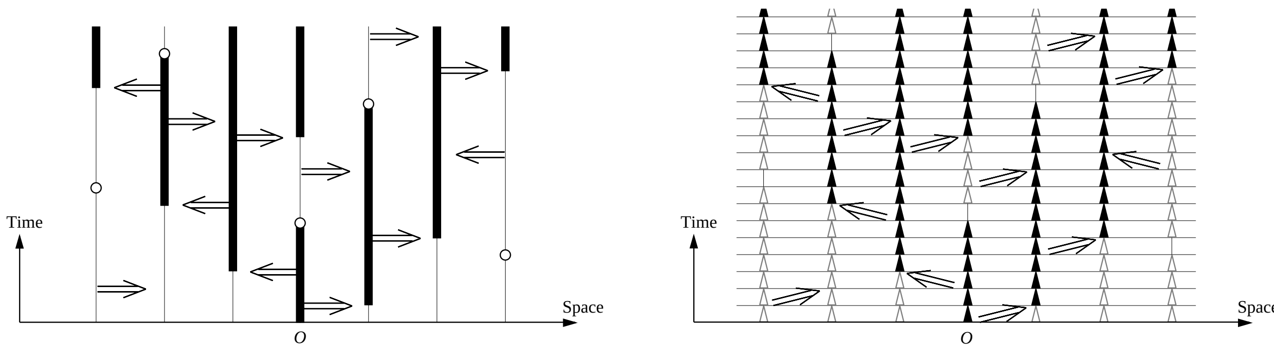

Figure 1: Graphical representation of the contact process and the discretized contact process.

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 $\{x \in \mathbb{Z}^d : (o,\, 0) \text{ is connected to } (x,\, t)\}$. See also [22, Section I.1].

Inspired by this percolation structure in space-time and following [27], we consider an oriented percolation approximation in $\mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$ to the contact process, where $\varepsilon \in (0,1]$ is a discretization parameter. We call this approximation the *discretized contact process*, and it is defined as follows. 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 a *temporal bond* if $x=y$, otherwise b is a *spatial bond*. 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

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

provided that $\|p_\varepsilon\|_\infty \leq 1$. 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 when $\varepsilon \downarrow 0$.

We also emphasize that the discretized contact process with $\varepsilon=1$ is equivalent to oriented percolation, for which $\lambda \in [0, \|D\|_\infty^{-1}]$ is the expected number of occupation bonds per site.

We denote by $(x,s) \longrightarrow (y,t)$ 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

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

Similarly to (1.5), the discretized contact process has a critical value $\lambda_c^{(\varepsilon)}$ satisfying

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

The main result for the discretized contact process with $\varepsilon\in(0,1]$ is the following theorem:

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

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

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

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

where all error terms are uniform in $\varepsilon\in(0,1]$. The error estimate in (2.4) is uniform in $k\in\mathbb{R}^d$ with $|k|^2/\log(2+t)$ sufficiently small.

Proposition 2.1 is the discrete analog of Theorem 1.1. The uniformity in ε 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. In particular, Proposition 2.1 applied to oriented percolation (i.e., $\varepsilon=1$) reproves [20, Theorem 1.1].

The discretized version of Theorem 1.2 is given in the following proposition:

Proposition 2.2 (Discretized results for $d\leq 4$). *Let $d\leq 4$ and $\delta\in(0,1\wedge\Delta\wedge\alpha)$. Then, there is a $\lambda_T=1+O(T^{-\mu})$ for some $\mu\in(0,\alpha-\delta)$ such that, for sufficiently large L_1, there are positive and finite constants $C_1=C_1(d)$ and $C_2=C_2(d)$ such that, for every $0<t\leq\log T$,*

$$\hat{\tau}_{Tt;\varepsilon}^{\lambda_T}\big(\tfrac{k}{\sqrt{\sigma_T^2 Tt}}\big) = e^{-\frac{|k|^2}{2d}} \big[1+O(T^{-\mu})+O(|k|^2(1+Tt)^{-\delta})\big], \tag{2.7}$$

$$\frac{1}{\hat{\tau}_{Tt;\varepsilon}^{\lambda_T}(0)} \sum_{x\in\mathbb{Z}^d} |x|^2 \tau_{Tt;\varepsilon}^{\lambda_T}(x) = \sigma_T^2 Tt\big[1+O(T^{-\mu})+O((1+Tt)^{-\delta})\big], \tag{2.8}$$

$$C_1 L_T^{-d}(1+Tt)^{-d/2} \leq \|\tau_{Tt;\varepsilon}^{\lambda_T}\|_\infty \leq (1-\varepsilon)^{Tt/\varepsilon} + C_2 L_T^{-d}(1+Tt)^{-d/2}, \tag{2.9}$$

where all error terms are uniform in $\varepsilon\in(0,1]$, and the error estimate in (2.7) is uniform in $k\in\mathbb{R}^d$ with $|k|^2/\log(2+Tt)$ sufficiently small.

Note that Proposition 2.2 applies also to oriented percolation, for which $\varepsilon=1$.

2.2 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 a brief motivation. We also motivate why we discretize time for the contact process.

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

$$(f*g)(x) = \sum_{y\in\mathbb{Z}^d} f(y)\,g(x-y). \tag{2.10}$$

We first motivate the basic idea underlying the expansion, similarly as in [20, 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 $\lambda D(y-x)$ with killing rate $1-\lambda$, we have the partial differential equation

$$\partial_t q_t^\lambda(x) = \lambda\,(D * q_t^\lambda)(x) - q_t^\lambda(x), \tag{2.11}$$

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

$$\partial_t \hat{q}_t^\lambda(k) = -[1 - \lambda \hat{D}(k)]\, \hat{q}_t^\lambda(k). \tag{2.12}$$

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

$$\hat{q}_t^\lambda(k) = e^{-[1-\lambda\hat{D}(k)]t}. \tag{2.13}$$

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

$$\hat{q}_t^1\big(\tfrac{k}{\sqrt{\sigma^2 t}}\big) = e^{-\frac{|k|^2}{2d}}\,[1 + o(1)], \tag{2.14}$$

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 (2.12), such as $-[1 - \lambda\hat{D}(k)]\,\hat{q}_{t-1}^\lambda(k)$, it would be much more involved 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 (recall (2.1))

$$q_{t;\varepsilon}^\lambda(x) = p_\varepsilon^{*t/\varepsilon}(x) \qquad (t \in \varepsilon\mathbb{N}). \tag{2.15}$$

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

$$q_{t;\varepsilon}^\lambda(x) = (p_\varepsilon * q_{t-\varepsilon;\varepsilon}^\lambda)(x) \qquad (t \in \varepsilon\mathbb{N}). \tag{2.16}$$

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 (2.12) 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 (2.16) can be solved using the Fourier transform to give

$$\hat{q}_{t;\varepsilon}^\lambda(k) = \hat{p}_\varepsilon(k)^{t/\varepsilon} = [1 - \varepsilon + \lambda\varepsilon\hat{D}(k)]^{t/\varepsilon} = e^{-[1-\lambda\hat{D}(k)]t + O(t\varepsilon[1-\hat{D}(k)]^2)}. \tag{2.17}$$

We note that the limit of $[\hat{q}_{t;\varepsilon}(k) - \hat{q}_{t-\varepsilon;\varepsilon}(k)]/\varepsilon$ exists and equals (2.12). In order to obtain the central limit theorem, we divide k by $\sqrt{\sigma^2 t}$. Then, uniformly in $\varepsilon > 0$, we have

$$\hat{q}_{t;\varepsilon}^1\big(\tfrac{k}{\sqrt{\sigma^2 t}}\big) = e^{-\frac{|k|^2}{2d} + O(|k|^{2+2\Delta}t^{-\Delta}) + O(\varepsilon|k|^4 t^{-1})}. \tag{2.18}$$

Therefore, the central limit theorem holds uniformly in $\varepsilon > 0$.

We follow Mark Kac's adagium: "Be wise, discretize!" for two reasons. Firstly, discretizing time allows us to obtain an expansion as in (2.11), and secondly, it allows us 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

$$\partial_t \hat{\tau}_t^\lambda(k) = -[1 - \lambda \hat{D}(k)]\, \hat{\tau}_t^\lambda(k) + \int_0^t ds\ \hat{\pi}_s^\lambda(k)\, \hat{\tau}_{t-s}^\lambda(k), \tag{2.19}$$

where $\hat{\pi}_s^\lambda$ are certain expansion coefficients. In order to derive the equation (2.19), we use that the discretized contact process is oriented percolation, for which lace expansions have been derived in the literature [20, 24, 25, 26, 27]. Clearly, the equation (2.19) is much more complicated than the corresponding equation for simple random walk in (2.11). Therefore, a simple solution to the equation as in (2.13) is impossible. We see no way to analyse the partial differential equation in (2.19) other than to discretize time combined with induction. It would be of interest to investigate whether (2.19) can be used directly.

We next explain the expansion for the discretized contact process in more detail, following the explanation in [20, 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). More precisely, a bond is pivotal for the connection from $(o, 0)$ to (x, t) when $(o, 0) \longrightarrow (x, t)$ in the possibly modified configuration in which the bond is made occupied, and $(o, 0)$ is not connected to (x, t) in the possibly modified configuration in which the bond is made vacant (see also Definition 3.1 below). In the strings-and-sausages picture, the strings are the pivotal bonds, and the sausages are the parts of the cluster from $(o, 0)$ in between the subsequent pivotal bonds. We expect that there are of the order t/ε 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, each sausage avoids the backbone from the endpoint of that sausage to (x, t), so that any connected path between the sausages is 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 (2.16) to the discretized contact process takes the form

$$\tau_{t;\varepsilon}^\lambda(x) = \sum_{s=0}^{t-\varepsilon}{}^{\bullet} (\pi_{s;\varepsilon}^\lambda * p_\varepsilon * \tau_{t-s-\varepsilon;\varepsilon}^\lambda)(x) + \pi_{t;\varepsilon}^\lambda(x) \qquad (t \in \varepsilon\mathbb{N}), \tag{2.20}$$

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 $\pi_{0;\varepsilon}^\lambda(x) = \delta_{o,x}$ and $\pi_{\varepsilon;\varepsilon}^\lambda(x) = 0$. Note that for $t = 0, \varepsilon$, we have $\tau_{0;\varepsilon}^\lambda(x) = \delta_{0,x}$ and $\tau_{\varepsilon;\varepsilon}^\lambda(x) = p_\varepsilon(x)$, which is consistent with (2.20).

Together with the initial values $\pi_{0;\varepsilon}^\lambda(x) = \delta_{o,x}$ and $\pi_{\varepsilon;\varepsilon}^\lambda(x) = 0$, the identity (2.20) gives an inductive definition of the sequence $\pi_{t;\varepsilon}^\lambda(x)$ for $t \geq 2\varepsilon$ with $t \in \varepsilon\mathbb{Z}_+$. However, to analyse the recursion relation (2.20), it will be crucial to have a useful representation for $\pi_{t;\varepsilon}^\lambda(x)$, and this is provided in Section 3. Note that (2.16) is of the form (2.20) with $\pi_{t;\varepsilon}^\lambda(x) = \delta_{o,x}\delta_{0,t}$, so that we can think of the coefficients $\pi_{t;\varepsilon}^\lambda(x)$ for $t \geq 2\varepsilon$ as quantifying the repulsive interaction between the sausages in the "string of sausages" picture.

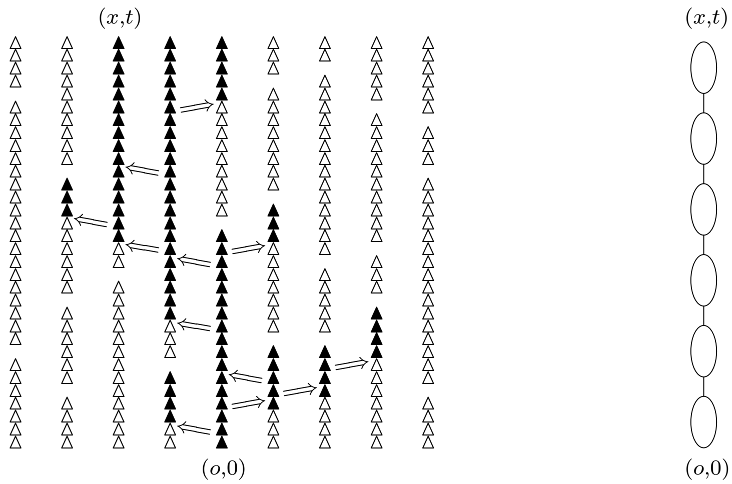

Figure 2: (a) A configuration for the discretized contact process. Open triangles △ denote occupied temporal bonds that are *not* connected from $(o,0)$, while closed triangles ▲ denote occupied temporal bonds that are connected from $(o,0)$. The arrows denote occupied spatial bonds, which represent the spread of the infection to neighbouring sites. (b) Schematic depiction of the configuration connecting $(o,0)$ and (x,t) as a "string of sausages."

Our proof will be based on showing that $\frac{1}{\varepsilon^2}\pi^\lambda_{t;\varepsilon}(x)$ for $t \geq 2\varepsilon$ is small at $\lambda = \lambda_c^{(\varepsilon)}$ if $d > 4$ and both t and L are large, uniformly in $\varepsilon > 0$. Based on this fact, we can rewrite the Fourier transform of (2.20) as

$$\frac{\hat{\tau}^\lambda_{t;\varepsilon}(k) - \hat{\tau}^\lambda_{t-\varepsilon;\varepsilon}(k)}{\varepsilon} = \frac{\hat{p}_\varepsilon(k) - 1}{\varepsilon}\,\hat{\tau}^\lambda_{t;\varepsilon}(k) + \varepsilon \sum_{s=\varepsilon}^{t-\varepsilon}{}^{\bullet}\, \frac{\hat{\pi}^\lambda_{s;\varepsilon}(k)}{\varepsilon^2}\,\hat{p}_\varepsilon(k)\,\hat{\tau}^\lambda_{t-s-\varepsilon;\varepsilon}(k) + \frac{\hat{\pi}^\lambda_{t;\varepsilon}(k)}{\varepsilon}. \tag{2.21}$$

Assuming convergence of $\frac{1}{\varepsilon^2}\hat{\pi}^\lambda_{s;\varepsilon}(k)$ to $\hat{\pi}^\lambda_s(k)$, which will be shown in Section 2.5, we obtain (2.19). Therefore, (2.19) is regarded as a small perturbation of (2.12) when $d > 4$ and $L \gg 1$, and this will imply the central limit theorem for the critical two-point function.

Now we briefly explain the expansion coefficients $\pi^\lambda_{t;\varepsilon}(x)$. In Section 3, we will obtain the expression

$$\pi^\lambda_{t;\varepsilon}(x) = \sum_{N=0}^{\infty} (-1)^N \pi^{(N)}_{t;\varepsilon}(x), \tag{2.22}$$

where we suppress the dependence of $\pi^{(N)}_{t;\varepsilon}(x)$ on λ. The idea behind the proof of (2.22) is the following. Let

$$\pi^{(0)}_{t;\varepsilon}(x) = \mathbb{P}^\lambda_\varepsilon((o,0) \Longrightarrow (x,t)) \tag{2.23}$$

denote the contribution to $\tau^\lambda_{t;\varepsilon}(x)$ from configurations in which there are no pivotal bonds, so that

$$\tau^\lambda_{t;\varepsilon}(x) = \pi^{(0)}_{t;\varepsilon}(x) + \sum_b \mathbb{P}^\lambda_\varepsilon\big(b \text{ first occupied and pivotal bond for } (o,0) \longrightarrow (x,t)\big), \tag{2.24}$$

where the sum over b is over bonds of the form $b = ((u,s),(v,s+\varepsilon))$. We write $\underline{b} = (u,s)$ for the starting point of the bond b and $\overline{b} = (v, s+\varepsilon)$ for its endpoint. Then, the probability on the right-hand side of (2.24) equals

$$\mathbb{P}^\lambda_\varepsilon\big((o,0) \Longrightarrow \underline{b},\ b \text{ occupied},\ \overline{b} \longrightarrow (x,t),\ b \text{ pivotal for } (o,0) \longrightarrow (x,t)\big). \tag{2.25}$$

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

$$\tau^{\lambda}_{t;\varepsilon}(x) = \pi^{(0)}_{t;\varepsilon}(x) + \sum_{s=0}^{t-\varepsilon}{}^{\bullet} \sum_{u,v\in\mathbb{Z}^d} \pi^{(0)}_{s;\varepsilon}(u)\, p_{\varepsilon}(v-u)\, \tau_{t-s-\varepsilon;\varepsilon}(x-v) - R^{(0)}_{t;\varepsilon}(x), \tag{2.26}$$

where

$$R^{(0)}_{t;\varepsilon}(x) = \sum_{b} \mathbb{P}^{\lambda}_{\varepsilon}\big((o,0) \Longrightarrow \underline{b},\ b \text{ occupied},\ \overline{b} \longrightarrow (x,t),\ b \text{ not pivotal for } (o,0) \longrightarrow (x,t)\big). \tag{2.27}$$

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 (2.22). 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 (2.22), 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 (2.22).

There are three ways to obtain the lace expansion in (2.20) for oriented percolation models. We use the expansion by Sakai [26, 27], as described in (2.23)–(2.27) above, based on inclusion-exclusion together with the Markov property for oriented percolation. For unoriented percolation, Hara and Slade [11] 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 [20] to investigate the oriented percolation r-point functions. The original expansion in [11] was for unoriented percolation, and does not make use of the Markov property. Nguyen and Yang [24, 25] derived an alternate expression for $\pi^{(N)}_{t;\varepsilon}(x)$ by adapting the lace expansion of Brydges and Spencer [7] 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.

It is not so hard to see that the Nguyen-Yang expansion is equivalent to the above expansion using inclusion-exclusion, just as for self-avoiding walks [23]. Since we find the Sakai expansion simpler, especially when dealing with the continuum limit, we prefer the Sakai expansion to the Nguyen-Yang expansion. In [20], the Hara-Slade expansion was used to obtain (2.22) with a *different* expression for $\pi^{(N)}_{t;\varepsilon}(x)$. 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 used in this paper produces the simplest diagrams.

In [20], the Nguyen-Yang expansion was also used to deal with the derivative of the lace expansion coefficients with respect to the percolation parameter p. In this paper, we use the inclusion-exclusion expansion also for the derivative of the expansion coefficients with respect to λ, rather than on two different expansions as in [20].

We now comment on the relative merits of the Sakai and the Hara-Slade expansion. Clearly, the Hara-Slade expansion is more general, as it also applies to unoriented percolation. On the other hand, the Sakai expansion is somewhat simpler to use, and the bounding diagrams on the arising Feynman diagrams are simpler. Finally, the resulting expressions for $\pi^{(N)}_{t;\varepsilon}(x)$ in the Sakai expansion allow for a continuum limit, where it is not clear to us how to perform this limit using the Hara-Slade expansion coefficients.

In [18], 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. The "decoupling" of disjoint paths is crucial in the derivation of the lace expansion for the higher point functions, and explains the importance of the Hara-Slade expansion for oriented percolation and the contact process.

To complete this discussion, we note that an alternative route to the contact process results is via (2.19). In [5], an approach using a Banach fixed point theorem was used to prove asymptotics of the two-point function for weakly self-avoiding walk. The crucial observation is that a lace expansion equation such as (2.19) can be viewed as a fixed point equation of a certain operator on sequence spaces. By proving properties of this operator, Bolthausen and Ritzman were able to deduce properties of the fixed point sequence, and thus of the weakly self-avoiding walk two-point function. It would be interesting to investigate whether such an approach may be used on (2.19) as well.

2.3 Bounds on the lace expansion

In order to prove the statements in Proposition 2.1, we will use induction in n, where $t = n\varepsilon \in \varepsilon\mathbb{Z}_+$. The lace expansion equation in (2.20) forms the main ingredient for this induction in time. We will explain the inductive method in more detail below. To advance the induction hypotheses, we clearly 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

$$\beta = L^{-d}. \tag{2.28}$$

We will use the following set of bounds:

$$|\hat\tau_{s;\varepsilon}(0)| \le K, \qquad |\nabla^2 \hat\tau_{s;\varepsilon}(0)| \le K\sigma^2 s, \qquad \|\hat D^2\, \hat\tau_{s;\varepsilon}\|_1 \le \frac{K\beta}{(1+s)^{d/2}}, \tag{2.29}$$

where we write $\|\hat f\|_1 = \int_{[-\pi,\pi]^d} \frac{d^dk}{(2\pi)^d}\, |\hat f(k)|$ for a function $\hat f : [-\pi,\pi]^d \mapsto \mathbb{C}$. The bounds on the lace expansion consist of the following estimates, which will be proved in Section 4.

Proposition 2.3 (Bounds on the lace expansion for $d > 4$). *Assume (2.29) for some λ_0 and all $s \le t$. Then, there are $\beta_0 = \beta_0(d,K) > 0$ and $C = C(d,K) < \infty$ (both independent of ε, L) such that, for $\lambda \le \lambda_0$, $\beta < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ with $2\varepsilon \le s \le t+\varepsilon$, $q = 0,2,4$ and $\Delta' \in [0, 1\wedge\Delta]$, and uniformly in $\varepsilon \in (0,1]$,*

$$\sum_{x\in\mathbb{Z}^d} |x|^q\, |\pi^\lambda_{s;\varepsilon}(x)| \le \frac{\varepsilon^2 C\sigma^q\beta}{(1+s)^{(d-q)/2}}, \tag{2.30}$$

$$\Big|\hat\pi^\lambda_{s;\varepsilon}(k) - \hat\pi^\lambda_{s;\varepsilon}(0) - \frac{a(k)}{\sigma^2}\, \nabla^2 \hat\pi^\lambda_{s;\varepsilon}(0)\Big| \le \frac{\varepsilon^2 C\beta\, a(k)^{1+\Delta'}}{(1+s)^{(d-2)/2-\Delta'}}, \tag{2.31}$$

$$|\partial_\lambda \hat\pi^\lambda_{s;\varepsilon}(0)| \le \frac{\varepsilon^2 C\beta}{(1+s)^{(d-2)/2}}. \tag{2.32}$$

The main content of Proposition 2.3 is that the bounds on $\hat\tau_{s;\varepsilon}$ for $s \le t$ in (2.29) imply bounds on $\hat\pi_{s;\varepsilon}$ for all $s \le t+\varepsilon$. This fact allows us to use the bounds on $\hat\pi_{s;\varepsilon}$ for all arising s in (2.20) 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 (2.29).

The proof of Proposition 2.3 is deferred to Section 4. Proposition 2.3 is probably false in dimensions $d \le 4$. 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 (1.20).

Proposition 2.4 (Bounds on the lace expansion for $d \le 4$). *Let $\alpha > 0$ in (1.21). Assume (2.29), with β replaced by $\beta_T = L_T^{-d}$ and σ^2 by σ_T^2, for some λ_0 and all $s \le t$. Then, there are $L_0 = L_0(d,K) < \infty$ (independent of ε) and $C = C(d,K) < \infty$ (independent of ε, L) such that, for $\lambda \le \lambda_0$, $L_1 \ge L_0$, $s \in \varepsilon\mathbb{Z}_+$ with $2\varepsilon \le s \le t+\varepsilon$, $q = 0,2,4$ and $\Delta' \in [0,1\wedge\Delta]$, the bounds in (2.30)–(2.32) hold for $t \le T\log T$, with β replaced by $\beta_T = L_T^{-d}$ and σ^2 by σ_T^2.*

The main point in Propositions 2.3–2.4 is the fact that we need to extract two factors of ε. One can see that such factors must be present by investigating, e.g., $\pi^{(0)}_{t;\varepsilon}(x)$, which is the probability that $(o,0)$ is doubly connected to (x,t). When $t>0$, 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 $N\geq 1$ then follows by induction in N.

2.4 Implementation of the inductive method

Our analysis of (2.20) begins by taking its Fourier transform, which gives the recursion relation

$$\hat\tau^{\lambda}_{t;\varepsilon}(k)=\sum_{s=0}^{t-\varepsilon}{}^{\bullet}\,\hat\pi^{\lambda}_{s;\varepsilon}(k)\,\hat p_{\varepsilon}(k)\,\hat\tau^{\lambda}_{t-s-\varepsilon;\varepsilon}(k)+\hat\pi^{\lambda}_{t;\varepsilon}(k)\qquad (t\in\varepsilon\mathbb{N}). \tag{2.33}$$

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 $\|\tau^{\lambda}_{s;\varepsilon}\|_1\equiv\sum_{x\in\mathbb{Z}^d}\tau^{\lambda}_{s;\varepsilon}(x)=\hat\tau^{\lambda}_{s;\varepsilon}(0)$ and $\|\tau^{\lambda}_{s;\varepsilon}\|_\infty\leq\|\hat\tau^{\lambda}_{s;\varepsilon}\|_1$ with $s\leq t-\varepsilon$. Therefore, the right-hand side of (2.33) explicitly involves $\hat\tau^{\lambda}_{s;\varepsilon}(k)$ only for $s\leq t-\varepsilon$. This opens up the possibility of an inductive analysis of (2.33). A general approach to this type of inductive analysis is given in [19]. 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

$$\lambda^{(\varepsilon)}_{\rm c}=1-\frac{1}{\varepsilon}\sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\,\hat\pi^{\lambda^{(\varepsilon)}_{\rm c}}_{s;\varepsilon}(0)\,\hat p^{\lambda^{(\varepsilon)}_{\rm c}}_{\varepsilon}(0), \tag{2.34}$$

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

$$A^{(\varepsilon)}=\frac{1+\sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\,\hat\pi^{\lambda^{(\varepsilon)}_{\rm c}}_{s;\varepsilon}(0)}{1+\frac{1}{\varepsilon}\sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\,s\,\hat\pi^{\lambda^{(\varepsilon)}_{\rm c}}_{s;\varepsilon}(0)\,\hat p^{\lambda^{(\varepsilon)}_{\rm c}}_{\varepsilon}(0)},\qquad v^{(\varepsilon)}=\frac{\lambda^{(\varepsilon)}_{\rm c}-\frac{1}{\sigma^2\varepsilon}\sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\,\nabla^2\big[\hat\pi^{\lambda^{(\varepsilon)}_{\rm c}}_{s;\varepsilon}(k)\,\hat p^{\lambda^{(\varepsilon)}_{\rm c}}_{\varepsilon}(k)\big]_{k=0}}{1+\frac{1}{\varepsilon}\sum_{s=2\varepsilon}^{\infty}{}^{\bullet}\,s\,\hat\pi^{\lambda^{(\varepsilon)}_{\rm c}}_{s;\varepsilon}(0)\,\hat p^{\lambda^{(\varepsilon)}_{\rm c}}_{\varepsilon}(0)}, \tag{2.35}$$

where we have added an argument $\lambda^{(\varepsilon)}_{\rm c}$ to emphasize that λ is critical for the evaluation of $\pi^{\lambda}_{t;\varepsilon}$ on the right-hand sides. Convergence of the series on the right-hand sides, for $d>4$, follows from Proposition 2.3. For oriented percolation, i.e., for $\varepsilon=1$, these equations agree with [20, (2.11-2.13)].

The result of induction is summarized in the following proposition:

Proposition 2.5 (Induction). *If Proposition 2.3 holds, then* (2.29) *holds for $s\leq t+\varepsilon$. Therefore,* (2.29) *holds for all $s\geq 0$ and* (2.30)–(2.32) *hold for all $s\geq 2\varepsilon$. Moreover, the statements in Proposition 2.1 follow, with the error terms uniform in $\varepsilon\in(0,1]$.*

There is also a low-dimensional version of Proposition 2.5, but we refrain from stating it.

2.5 Continuum limit

In this section we state the result necessary to complete the proof of Theorems 1.1–1.2 from Propositions 2.1–2.2. In particular, from now onwards, we specialize to the contact process.

Proposition 2.6 (Continuum limit). *Suppose that $\lambda^{(\varepsilon)}\to\lambda$ and $\lambda^{(\varepsilon)}\leq\lambda^{(\varepsilon)}_{\rm c}$ for ε sufficiently small. Then, for every $t>0$ and $x\in\mathbb{Z}^d$, there is a $\pi^{\lambda}_t(x)$ such that*

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^2}\pi^{\lambda^{(\varepsilon)}}_{t;\varepsilon}(x)=\pi^{\lambda}_t(x),\qquad\qquad \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon^2}[\partial_\ell\pi^{\ell}_{t;\varepsilon}(x)]_{\ell=\lambda^{(\varepsilon)}}=\partial_\lambda\pi^{\lambda}_t(x). \tag{2.36}$$

Consequently, for $\lambda \le \lambda_c$ and $q = 0, 2, 4$,

$$\sum_{x\in\mathbb{Z}^d} |x|^q \pi_t^\lambda(x) \le \frac{C\beta}{(1+t)^{(d-q)/2}}, \qquad \sum_{x\in\mathbb{Z}^d} \partial_\lambda \pi_t^\lambda(x) \le \frac{C\beta}{(1+t)^{(d-2)/2}}, \tag{2.37}$$

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

$$\lim_{\varepsilon\downarrow 0} A^{(\varepsilon)} = A, \qquad \lim_{\varepsilon\downarrow 0} v^{(\varepsilon)} = v. \tag{2.38}$$

Furthermore, $\partial_\lambda \pi_t^\lambda(x)$ is continuous in λ.

In Proposition 2.3, the right-hand sides of (2.30)–(2.32) are proportional to ε^2. The main point in the proof of Proposition 2.6 is that the lace expansion coefficients, scaled by ε^{-2}, converge as $\varepsilon \downarrow 0$, using the weak convergence of $\mathbb{P}_\varepsilon^\lambda$ to $\mathbb{P}^\lambda$ [4, Proposition 2.7].

In Section 6, 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 use

$$\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^\lambda(x) = \int_0^\lambda d\lambda' \, \frac{1}{\varepsilon^2}\partial_{\lambda'}\pi_{t;\varepsilon}^{\lambda'}(x). \tag{2.39}$$

where we use $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^0(x) = 0$ for $t > 0$. By the assumed pointwise convergence, the left-hand side converges to $\pi_t^\lambda(x)$, while the right-hand side converges to the integral of the limit of $\frac{1}{\varepsilon^2}\partial_{\lambda'}\pi_{t;\varepsilon}^{\lambda'}(x)$, denoted $f_t^{\lambda'}(x)$ for now, using the dominated convergence theorem. Therefore, for any $\lambda \le \lambda_c$,

$$\pi_t^\lambda(x) = \int_0^\lambda d\lambda' \, f_t^{\lambda'}(x), \tag{2.40}$$

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

Proof of Theorems 1.1–1.2 assuming Propositions 2.1–2.2 and 2.6. We 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 $\lambda > 0$,

$$\lim_{\varepsilon\downarrow 0} \tau_{t;\varepsilon}^\lambda(x) = \tau_t^\lambda(x). \tag{2.41}$$

Since $\tau_t^\lambda(x)$ is continuous in λ (see e.g., [22, pp.38–39]), we also obtain $\lim_{\varepsilon\downarrow 0} \tau_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x) = \tau_t^\lambda(x)$ for any $\lambda^{(\varepsilon)} \to \lambda$. Since $\lambda_c^{(\varepsilon)} \to \lambda_c$ [27, Section 2.1], $\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x)$ also converges to $\tau_t^{\lambda_c}(x)$. Using the uniformity in ε of the upper and lower bounds in (2.6), we obtain (1.9).

Next, we prove $\lim_{\varepsilon\downarrow 0} \hat{\tau}_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(k) = \hat{\tau}_t^{\lambda_c}(k)$ for every $k \in [-\pi, \pi]^d$ and $t \ge 0$. Note that the Fourier transform involves a sum over $\mathbb{Z}^d$, such as

$$\hat{\tau}_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(k) = \sum_{x\in\mathbb{Z}^d} \tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x) \, e^{ik\cdot x}. \tag{2.42}$$

To use the pointwise convergence of $\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x)$, we first show that the sum over $x \in \mathbb{Z}^d$ in (2.42) can be approximated by a finite sum. To see this, we note that

$$\tau_{t;\varepsilon}^\lambda(x) \le p_\varepsilon^{*t/\varepsilon}(x) = \sum_{n=0}^{t/\varepsilon} \binom{t/\varepsilon}{n} (1-\varepsilon)^{t/\varepsilon - n} (\lambda\varepsilon)^n \, D^{*n}(x). \tag{2.43}$$

For any fixed t, we can choose $\delta_R \geq 0$, which is ε-independent and decays to zero as $R \uparrow \infty$, such that

$$\sum_{x\in\mathbb{Z}^d:\|x\|_\infty>R} \tau_{t;\varepsilon}^{\lambda}(x) \leq \delta_R. \tag{2.44}$$

Therefore, the same holds for $\tau_t^\lambda(x)$, and hence we can approximate both $\hat\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(k)$ and $\hat\tau_t^{\lambda_c}(k)$ by sums over $x \in \mathbb{Z}^d$ with $\|x\|_\infty \leq R$, in which we use the pointwise convergence of $\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(x)$. Taking $R \uparrow \infty$, we obtain $\hat\tau_t^{\lambda_c}(k) = \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}(k)$.

Using the above, we obtain

$$\begin{aligned}
\hat\tau_t^{\lambda_c}\big(\tfrac{k}{\sqrt{v\sigma^2 t}}\big) &= \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}\big(\tfrac{k}{\sqrt{v\sigma^2 t}}\big) = \lim_{\varepsilon\downarrow 0} \hat\tau_{t;\varepsilon}^{\lambda_c^{(\varepsilon)}}\big(\tfrac{\sqrt{v^{(\varepsilon)}}}{\sqrt{v}} \tfrac{k}{\sqrt{v^{(\varepsilon)}\sigma^2 t}}\big)\\
&= \lim_{\varepsilon\downarrow 0} A^{(\varepsilon)}\, e^{-\frac{v^{(\varepsilon)}}{v}\frac{|k|^2}{2d}} \big[1 + O\big(\tfrac{v^{(\varepsilon)}}{v}|k|^2(1+t)^{-\delta}\big) + O((1+t)^{-(d-4)/2})\big]\\
&= A\, e^{-\frac{|k|^2}{2d}} \big[1 + O(|k|^2(1+t)^{-\delta}) + O((1+t)^{-(d-4)/2})\big],
\end{aligned} \tag{2.45}$$

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

Proof of (1.14) *assuming* (2.19) *and Proposition 2.6.* We now prove that, in the current setting, $\chi(\lambda) = \int_0^\infty dt\, \hat\tau_t^\lambda(0)$ satisfies the precise asymptotics in (1.14), assuming (2.19) and Proposition 2.6.

Let $\lambda < \lambda_c$. Since $\hat\tau_0^\lambda(0) = 1$ and $\hat\tau_\infty^\lambda(0) = 0$, using (2.19) we obtain

$$\begin{aligned}
-1 = \int_0^\infty dt\ \partial_t \hat\tau_t^\lambda(0) &= \int_0^\infty dt\ \Big[(\lambda-1)\,\hat\tau_t^\lambda(0) + \int_0^t ds\ \hat\pi_s^\lambda(0)\,\hat\tau_{t-s}^\lambda(0)\Big]\\
&= \Big[\lambda - 1 + \int_0^\infty ds\ \hat\pi_s^\lambda(0)\Big] \int_0^\infty dt\ \hat\tau_t^\lambda(0),
\end{aligned} \tag{2.46}$$

so that

$$\chi(\lambda) = \Big[1 - \lambda - \int_0^\infty ds\ \hat\pi_s^\lambda(0)\Big]^{-1}. \tag{2.47}$$

By (2.34) and Proposition 2.6, λ_c must satisfy

$$\lambda_c = 1 - \int_0^\infty ds\ \hat\pi_s^{\lambda_c}(0), \tag{2.48}$$

so that we can rewrite (2.47) as

$$\chi(\lambda) = [f(\lambda_c) - f(\lambda)]^{-1}, \tag{2.49}$$

where $f(\lambda) = \lambda + \int_0^\infty ds\, \hat\pi_s^\lambda(0)$, since, by (2.48), $f(\lambda_c) = 1$. By the mean-value theorem, together with the fact that $|\partial_\lambda \hat\pi_s^\lambda(0)|$ is integrable with respect to $s > 0$, there is a $\lambda_* \in (\lambda, \lambda_c)$ such that

$$\chi(\lambda) = [(\lambda_c - \lambda)\, f'(\lambda_*)]^{-1}. \tag{2.50}$$

By the continuity in λ of $\partial_\lambda \pi_t^\lambda(x)$ and its summability in $(x,t) \in \mathbb{Z}^d \times \mathbb{R}_+$ for $\lambda \leq \lambda_c$ due to (2.37), $f'(\lambda) = 1 + \int_0^\infty ds\ \partial_\lambda \hat\pi_t^\lambda(0)$ is also continuous in $\lambda \leq \lambda_c$. Therefore, we obtain (1.14) with $C = f'(\lambda_c)^{-1}$.

Finally, we note that the above proof, where the integral is replaced with a sum over $n \in \mathbb{Z}_+$, also shows that the stronger version of $\gamma = 1$ holds for oriented percolation. □

The proofs of Theorems 1.1–1.2 are now reduced to the proof of Propositions 2.1–2.2 and 2.6. Proposition 2.6 will be proved in Section 6. The proof of Propositions 2.1–2.2 is reduced to Propositions 2.3–2.5, which will be proved in Sections 4–5. The advancement of the induction hypotheses is deferred to Appendix A. We start in Section 3 by deriving the lace expansion (2.20).

3 Lace expansion

In this section, we derive the lace expansion in (2.20). The same type of recursion relation was used for *discrete* models, such as (weakly) self-avoiding walk in $\mathbb{Z}^d$ [7, 12, 15, 19, 21, 30, 31, 32] and oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$ [19, 20, 24, 25].

From now on, we will suppress the dependence on ε and λ when no confusion can arise, and write, e.g., $\pi_t(x) = \pi^{\lambda}_{t;\varepsilon}(x)$. In Section 3.1, we obtain (3.28), which is equivalent to the recursion relation in (2.20), and the expression (3.26) for $\pi_t(x)$. In Section 3.2, we obtain the expressions (3.34)–(3.35) for $\partial_\lambda \pi_t(x)$.

3.1 Expansion for the two-point function

In this section, we derive the expansion (3.28). We will also write $\Lambda = \mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$, and use bold letters $\boldsymbol{o}, \boldsymbol{x}, \dots$ to represent elements in Λ, such as $\boldsymbol{o} = (o, 0)$ and $\boldsymbol{x} = (x, t)$, and write $\tau(\boldsymbol{x}) = \tau_t(x)$, $\pi^{(N)}(\boldsymbol{x}) = \pi^{(N)}_t(x)$, and so on.

We recall that the two-point function is defined by

$$\tau(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{o} \longrightarrow \boldsymbol{x}). \tag{3.1}$$

Before starting with the expansion, we introduce some definition:

Definition 3.1. (i) For a bond $b = (\boldsymbol{u}, \boldsymbol{v})$, we write $\underline{b} = \boldsymbol{u}$ and $\overline{b} = \boldsymbol{v}$. We write $b \longrightarrow \boldsymbol{x}$ for the event that b is occupied and $\overline{b} \longrightarrow \boldsymbol{x}$.

(ii) Given a configuration, we say that $\boldsymbol{v}$ is *doubly connected to* $\boldsymbol{x}$, and we write $\boldsymbol{v} \Longrightarrow \boldsymbol{x}$, if there are at least two bond-disjoint paths from $\boldsymbol{v}$ to $\boldsymbol{x}$ consisting of occupied bonds. By convention, we say that $\boldsymbol{x} \Longrightarrow \boldsymbol{x}$ for all $\boldsymbol{x}$.

(iii) A bond is said to be *pivotal* for $\boldsymbol{v} \longrightarrow \boldsymbol{x}$ if $\boldsymbol{v} \longrightarrow \boldsymbol{x}$ in the possibly modified configuration in which that bond is made occupied, whereas $\boldsymbol{v}$ is not connected to $\boldsymbol{x}$ in the possibly modified configuration in which that bond is made vacant.

We split, depending on whether there is a pivotal bond for $\boldsymbol{o} \longrightarrow \boldsymbol{x}$, to obtain

$$\tau(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{o} \Longrightarrow \boldsymbol{x}) + \sum_b \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \text{ occupied \& pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}). \tag{3.2}$$

We denote

$$\pi^{(0)}(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{o} \Longrightarrow \boldsymbol{x}), \tag{3.3}$$

so that we can rewrite (3.2) as

$$\tau(\boldsymbol{x}) = \pi^{(0)}(\boldsymbol{x}) + \sum_b \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{x},\ b \text{ pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}). \tag{3.4}$$

Define

$$R^{(0)}(\boldsymbol{x}) = \sum_b \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{x},\ b \text{ not pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}), \tag{3.5}$$

then, by inclusion-exclusion on the event that b is pivotal for $\boldsymbol{o} \longrightarrow \boldsymbol{x}$, we arrive at

$$\tau(\boldsymbol{x}) = \pi^{(0)}(\boldsymbol{x}) + \sum_b \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{x}) - R^{(0)}(\boldsymbol{x}). \tag{3.6}$$

The event $\boldsymbol{o} \Longrightarrow \underline{b}$ only depends on bonds with time variables less than or equal to the one of $\underline{b}$, while the event $b \longrightarrow \boldsymbol{x}$ only depends on bonds with time variables larger than or equal to the one of $\underline{b}$. Therefore, by the Markov property, we obtain

$$\mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{x}) = \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b})\,\mathbb{P}(b \text{ occupied})\,\mathbb{P}(\overline{b} \longrightarrow \boldsymbol{x}) = \pi^{(0)}(\underline{b})\,p(b)\,\tau(\boldsymbol{x} - \overline{b}), \tag{3.7}$$

where we abuse notation to write

$$p(b) = p(\overline{b} - \underline{b}). \tag{3.8}$$

Therefore, we arrive at

$$\tau(\boldsymbol{x}) = \pi^{(0)}(\boldsymbol{x}) + (\pi^{(0)} \star p \star \tau)(\boldsymbol{x}) - R^{(0)}(\boldsymbol{x}), \tag{3.9}$$

where we use "$\star$" to denote convolution in Λ, i.e.,

$$(f \star g)(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \Lambda} f(\boldsymbol{y})\, g(\boldsymbol{x} - \boldsymbol{y}). \tag{3.10}$$

This completes the first step of the expansion, and we are left to investigate $R^{(0)}(\boldsymbol{x})$. For this, we need some further notation.

Definition 3.2. (i) Given a configuration and $\boldsymbol{x} \in \Lambda$, we define $\mathbf{C}(\boldsymbol{x})$ to be the set of sites to which $\boldsymbol{x}$ is connected, i.e., $\mathbf{C}(\boldsymbol{x}) = \{\boldsymbol{y} \in \Lambda : \boldsymbol{x} \longrightarrow \boldsymbol{y}\}$. Given a bond b, we also define $\tilde{\mathbf{C}}^b(\boldsymbol{x})$ to be the set of sites to which $\boldsymbol{x}$ is connected in the (possibly modified) configuration in which b is made vacant.

(ii) Given a site set $\mathbf{C}$, we say that $\boldsymbol{v}$ is connected to $\boldsymbol{x}$ *through* $\mathbf{C}$, if every occupied path connecting $\boldsymbol{v}$ to $\boldsymbol{x}$ has at least one bond with an endpoint in $\mathbf{C}$. This event is written as $\boldsymbol{v} \xrightarrow{\mathbf{C}} \boldsymbol{x}$. Similarly, we write $\{b \xrightarrow{\mathbf{C}} \boldsymbol{x}\} = \{b \text{ occupied}\} \cap \{\overline{b} \xrightarrow{\mathbf{C}} \boldsymbol{x}\}$.

We then note that

$$\{\boldsymbol{v} \longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{x},\ b \text{ not pivotal for } \boldsymbol{v} \longrightarrow \boldsymbol{x}\} = \big\{\boldsymbol{v} \longrightarrow \underline{b},\ b \xrightarrow{\tilde{\mathbf{C}}^b(\boldsymbol{v})} \boldsymbol{x}\big\}. \tag{3.11}$$

Therefore,

$$R^{(0)}(\boldsymbol{x}) = \sum_b \mathbb{P}\big(\boldsymbol{o} \Longrightarrow \underline{b},\ b \xrightarrow{\tilde{\mathbf{C}}^b(\boldsymbol{o})} \boldsymbol{x}\big). \tag{3.12}$$

The event $\{\boldsymbol{v} \xrightarrow{\mathbf{C}} \boldsymbol{x}\}$ can be decomposed into two cases depending on whether there is or is not a pivotal bond b for $\boldsymbol{v} \longrightarrow \boldsymbol{x}$ such that $\boldsymbol{v} \xrightarrow{\mathbf{C}} \underline{b}$. Let

$$E'(\boldsymbol{v}, \boldsymbol{y}; \mathbf{C}) = \{\boldsymbol{v} \xrightarrow{\mathbf{C}} \boldsymbol{y}\} \cap \big\{\nexists\, b \text{ pivotal for } \boldsymbol{v} \longrightarrow \boldsymbol{y} \text{ s.t. } \boldsymbol{v} \xrightarrow{\mathbf{C}} \underline{b}\big\}, \tag{3.13}$$

$$E(b, \boldsymbol{y}; \mathbf{C}) = \{b \text{ occupied}\} \cap E'(\overline{b}, \boldsymbol{y}; \mathbf{C}). \tag{3.14}$$

See Figure 3 for a schematic representation of the event $E(b, \boldsymbol{x}; \mathbf{C})$. If there are pivotal bonds for $\boldsymbol{v} \longrightarrow \boldsymbol{x}$, then we take the *first* such pivotal bond b for which $\boldsymbol{v} \xrightarrow{\mathbf{C}} \underline{b}$. Therefore, we have the partition

$$\{\boldsymbol{v} \xrightarrow{\mathbf{C}} \boldsymbol{x}\} = E'(\boldsymbol{v}, \boldsymbol{x}; \mathbf{C}) \,\dot{\cup}\, \dot{\bigcup_b} \big\{E'(\boldsymbol{v}, \underline{b}; \mathbf{C}) \cap \{b \text{ occupied \& pivotal for } \boldsymbol{v} \longrightarrow \boldsymbol{x}\}\big\}. \tag{3.15}$$

Defining

$$\pi^{(1)}(\boldsymbol{y}) = \sum_b \mathbb{P}\big(\{\boldsymbol{o} \Longrightarrow \underline{b}\} \cap E(b, \boldsymbol{x}; \tilde{\mathbf{C}}^b(\boldsymbol{o}))\big), \tag{3.16}$$

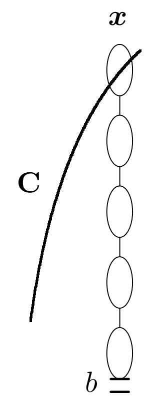

Figure 3: Schematic representation of the event $E(b,\boldsymbol{x};\mathbf{C})$.

we obtain

$$R^{(0)}(\boldsymbol{x})=\pi^{(1)}(\boldsymbol{x})+\sum_{b_1,b_2}\mathbb{P}\big(\{\boldsymbol{o}\Longrightarrow\underline{b}_1\}\cap E(b_1,\underline{b}_2;\tilde{\mathbf{C}}^{b_1}(\boldsymbol{o}))\cap\{b_2\text{ occupied \& pivotal for }\overline{b}_1\longrightarrow\boldsymbol{x}\}\big). \tag{3.17}$$

To the second term, we apply the inclusion-exclusion relation

$$\{b\text{ occupied \& pivotal for }\boldsymbol{v}\longrightarrow\boldsymbol{x}\}=\{\boldsymbol{v}\longrightarrow\underline{b},\ b\longrightarrow\boldsymbol{x}\}\setminus\big\{\boldsymbol{v}\longrightarrow\underline{b},\ b\xrightarrow{\tilde{\mathbf{C}}^{b}(\boldsymbol{v})}\boldsymbol{x}\big\}. \tag{3.18}$$

We define

$$R^{(1)}(\boldsymbol{x})=\sum_{b_1,b_2}\mathbb{P}\big(\{\boldsymbol{o}\Longrightarrow\underline{b}_1\}\cap E(b_1,\underline{b}_2;\tilde{\mathbf{C}}^{b_1}(\boldsymbol{o}))\cap\big\{b_2\xrightarrow{\tilde{\mathbf{C}}^{b_2}(\overline{b}_1)}\boldsymbol{x}\big\}\big), \tag{3.19}$$

so that we obtain

$$R^{(0)}(\boldsymbol{x})=\pi^{(1)}(\boldsymbol{x})+\sum_{b_1,b_2}\mathbb{P}\big(\{\boldsymbol{o}\Longrightarrow\underline{b}_1\}\cap E(b_1,\underline{b}_2;\tilde{\mathbf{C}}^{b_1}(\boldsymbol{o}))\cap\{b_2\longrightarrow\boldsymbol{x}\}\big)-R^{(1)}(\boldsymbol{x}), \tag{3.20}$$

where we use that

$$E'(\boldsymbol{v},\underline{b};\mathbf{C})\cap\{\boldsymbol{v}\longrightarrow\underline{b},\ b\longrightarrow\boldsymbol{x}\}=E'(\boldsymbol{v},\underline{b};\mathbf{C})\cap\{b\longrightarrow\boldsymbol{x}\}. \tag{3.21}$$

The event $\{\boldsymbol{o}\Longrightarrow\underline{b}_1\}\cap E(b_1,\underline{b}_2;\tilde{\mathbf{C}}^{b_1}(\boldsymbol{o}))$ depends only on bonds before $\underline{b}_2$, while $\{b_2\longrightarrow\boldsymbol{x}\}$ depends only on bonds after $\underline{b}_2$. By the Markov property, we end up with

$$\begin{aligned}R^{(0)}(\boldsymbol{x})&=\pi^{(1)}(\boldsymbol{x})+\sum_{b_2}\pi^{(1)}(\underline{b}_2)\,p(b_2)\,\tau(\boldsymbol{x}-\overline{b}_2)-R^{(1)}(\boldsymbol{x})\\&=\pi^{(1)}(\boldsymbol{x})+(\pi^{(1)}\star p\star\tau)(\boldsymbol{x})-R^{(1)}(\boldsymbol{x}),\end{aligned} \tag{3.22}$$

so that

$$\tau(\boldsymbol{x})=\pi^{(0)}(\boldsymbol{x})-\pi^{(1)}(\boldsymbol{x})+\big((\pi^{(0)}-\pi^{(1)})\star p\star\tau\big)(\boldsymbol{x})+R^{(1)}(\boldsymbol{x}). \tag{3.23}$$

This completes the second step of the expansion.

To complete the expansion for $\tau(\boldsymbol{x})$, we need to investigate $R^{(1)}(\boldsymbol{x})$ in more detail. Note that $R^{(1)}(\boldsymbol{x})$ involves the probability of a subset of $\big\{b_2\xrightarrow{\tilde{\mathbf{C}}^{b_2}(\underline{b}_1)}\boldsymbol{x}\big\}$.

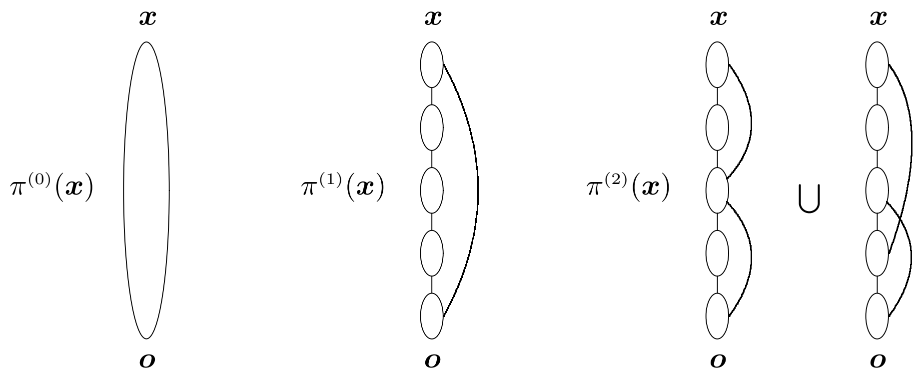

Figure 4: Schematic representations of $\pi^{(0)}(\boldsymbol{x})$, $\pi^{(1)}(\boldsymbol{x})$ and $\pi^{(2)}(\boldsymbol{x})$.

For this subset, we will use (3.15) and (3.18) again, and follow the steps of the above proof. The expansion is completed by repeating the above steps indefinitely. To facilitate the statement and the proof of the expansion, we make a few more definitions. For $\vec{b}_N = (b_1, \dots, b_N)$ with $N \geq 1$, we define

$$\tilde{E}^{(N)}_{\vec{b}_N}(\boldsymbol{x}) = \{\boldsymbol{o} \Longrightarrow \underline{b}_1\} \cap \bigcap_{i=1}^{N-1} E\big(b_i, \underline{b}_{i+1}; \tilde{\mathbf{C}}^{b_i}(\overline{b}_{i-1})\big) \cap E\big(b_N, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})\big), \tag{3.24}$$

where we use the convention that $\overline{b}_0 = \boldsymbol{o}$ and that the empty intersection, arising when $N = 1$, is the whole probability space. Also, we let

$$\tilde{E}^{(0)}_{\vec{b}_0}(\boldsymbol{x}) = \{\boldsymbol{o} \Longrightarrow \boldsymbol{x}\}. \tag{3.25}$$

Using this notation, we define

$$\pi^{(N)}(\boldsymbol{x}) = \sum_{\vec{b}_N} \mathbb{P}\big(\tilde{E}^{(N)}_{\vec{b}_N}(\boldsymbol{x})\big), \tag{3.26}$$

and denote the alternating sum by

$$\pi(\boldsymbol{x}) = \sum_{N=0}^{\infty} (-1)^N \pi^{(N)}(\boldsymbol{x}). \tag{3.27}$$

Note that the sum in (3.27) is a finite sum, as long as $t_{\boldsymbol{x}}$ is finite, where $t_{\boldsymbol{x}}$ denotes the time coordinate of $\boldsymbol{x}$, since each of the bonds $b_1, \dots, b_N$ eats up at least one time-unit ε, so that $\pi^{(N)}(\boldsymbol{x}) = 0$ for $N\varepsilon > t_{\boldsymbol{x}}$. The result of the expansion is summarized as follows.

Proposition 3.3 (The lace expansion). *For any $\lambda \geq 0$ and $\boldsymbol{x} \in \Lambda$,*

$$\tau(\boldsymbol{x}) = \pi(\boldsymbol{x}) + (\pi \star p \star \tau)(\boldsymbol{x}). \tag{3.28}$$

Proof. By (3.22), we are left to identify $R^{(1)}(\boldsymbol{x})$. For $N \geq 1$, we define

$$R^{(N)}(\boldsymbol{x}) = \sum_{\vec{b}_N} \mathbb{P}\big(\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap \big\{b_N \xrightarrow{\tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})} \boldsymbol{x}\big\}\big). \tag{3.29}$$

We prove below

$$R^{(N)}(\boldsymbol{x}) = \pi^{(N)}(\boldsymbol{x}) + (\pi^{(N)} \star p \star \tau)(\boldsymbol{x}) - R^{(N+1)}(\boldsymbol{x}). \tag{3.30}$$

The equation (3.28) follows by repeated use of (3.30) until the remainder $R^{(N+1)}(\boldsymbol{x})$ vanishes, which must happen at least when $N\varepsilon > t_{\boldsymbol{x}}$. To complete the proof of Proposition 3.3, we are left to prove (3.30), which is a generalization of (3.22).

First we rewrite $b_N \xrightarrow{\tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})} \boldsymbol{x}$ in (3.29). As in (3.15), this event can be decomposed into two cases, depending on whether there is or is not a pivotal bond b for $\overline{b}_N \longrightarrow \boldsymbol{x}$ such that $\overline{b}_N \xrightarrow{\tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})} \underline{b}$. The contribution where there is no such a bond equals $E(b_N, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1}))$. If there are such pivotal bonds, then we take the *first* bond b among these bonds and obtain (cf., (3.15))

$$\begin{aligned}\big\{b_N \xrightarrow{\tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})} \boldsymbol{x}\big\} &= E(b_N, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})) \\ &\quad \dot\cup \dot{\bigcup_b} \big\{E(b_N, \underline{b}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})) \cap \{b \text{ occupied \& pivotal for } \overline{b}_N \longrightarrow \boldsymbol{x}\}\big\}.\end{aligned} \tag{3.31}$$

By (3.26), the contribution from $E(b_N, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1}))$ in the right-hand side is $\pi^{(N)}(\boldsymbol{x})$, which is the first term in the right-hand side of (3.30). For the contribution from the remaining event in (3.31), we use (3.18) to arrive at

$$\sum_{\vec{b}_N, b} \mathbb{P}\big(\tilde{E}^{(N)}_{\vec{b}_N}(\underline{b}) \cap \{b \text{ occupied \& pivotal for } \overline{b}_N \longrightarrow \boldsymbol{x}\}\big) = \sum_{\vec{b}_N, b} \mathbb{P}\big(\tilde{E}^{(N)}_{\vec{b}_N}(\underline{b}) \cap \{b \longrightarrow \boldsymbol{x}\}\big) - R^{(N+1)}(\boldsymbol{x}). \tag{3.32}$$

The last term in the above expression is the last term in the right-hand side of (3.30). Again by the Markov property and (3.26), the first term in the right-hand side of (3.32) equals the second term in the right-hand side of (3.30). This completes the proof of (3.30) and thus the proof of Proposition 3.3. □

3.2 Representation for the derivative

In this section, we derive a formula for $\partial_\lambda \pi(\boldsymbol{x})$. To state the result below, we define

$$\texttt{piv}[\boldsymbol{v}, \boldsymbol{x}] = \{b : b \text{ pivotal for } \boldsymbol{v} \longrightarrow \boldsymbol{x}\}. \tag{3.33}$$

Proposition 3.4. *For $\lambda > 0$ and $\boldsymbol{x} \in \Lambda$,*

$$\partial_\lambda \pi(\boldsymbol{x}) = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \, \Pi^{(N)}(\boldsymbol{x}), \tag{3.34}$$

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

$$\Pi^{(N;n)}(\boldsymbol{x}) = \sum_{\substack{\vec{b}_N, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(N)}_{\vec{b}_N}(\boldsymbol{x}) \cap \big\{b \in \{b_n\} \,\dot\cup\, \texttt{piv}[\overline{b}_n, \underline{b}_{n+1}]\big\}\big), \tag{3.35}$$

and where $\underline{b}_{N+1}$ is defined to be $\boldsymbol{x}$.

By the same reason as in (3.27), the sum in (3.34) is a finite sum. We prove (3.34) by differentiating the expression (3.28) for $\tau(\boldsymbol{x})$ and comparing it with the expression for $\partial_\lambda \tau(\boldsymbol{x})$ obtained by using Russo's formula, rather than differentiating $\partial_\lambda \pi_t(x)$ directly. Possibly, one can also use direct differentiation of the expressions (3.26) for $\pi(\boldsymbol{x})$, but this is cumbersome because of the complex combination of *increasing* and *decreasing* events consisting of $\pi^{(N)}(\boldsymbol{x})$, where an increasing (respectively, decreasing) event is an event that is more (respectively, 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 $N = 0$. In particular, it is *not* true that $\partial_\lambda \pi^{(N)}(\boldsymbol{x}) = \frac{1}{\lambda}\Pi^{(N)}(\boldsymbol{x})$.

Proof. In the proof it will be convenient to split

$$\Pi^{(N,n)}(\boldsymbol{x}) = \Pi^{(N;n,1)}(\boldsymbol{x}) + \Pi^{(N;n,2)}(\boldsymbol{x}), \tag{3.36}$$

where $\Pi^{(N;n,1)}(\boldsymbol{x})$ is the contribution from $b = b_n$ in (3.35), whereas $\Pi^{(N;n,2)}(\boldsymbol{x})$ is the contribution from $b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]$.

To obtain an expression for $\partial_\lambda \tau(\boldsymbol{x})$, we use *Russo's formula* [4, 9]. Let E be an increasing event that depends only on finitely many spatial bonds. Then

$$\partial_\lambda \mathbb{P}(E) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \mathbb{P}(E \text{ occurs, } b \text{ is pivotal for } E), \tag{3.37}$$

where we use the fact that only spatial pivotal bonds for E are responsible to the differentiation with respect to λ. Let $\square_R = [-R, R]^d \cap \mathbb{Z}^d$. We apply (3.37) to $E = E_R(\boldsymbol{x}) \equiv \{\boldsymbol{o} \longrightarrow \boldsymbol{x} \text{ in } \square_R \times [0, t_{\boldsymbol{x}}]\}$, which is the set of bond configurations whose restriction on bonds $(\boldsymbol{u}, \boldsymbol{v}) \subset \square_R \times [0, t_{\boldsymbol{x}}]$ are in $\{\boldsymbol{o} \longrightarrow \boldsymbol{x}\}$. Note that $\lim_{R\to\infty} \mathbb{P}(E_R(\boldsymbol{x})) = \tau(\boldsymbol{x})$, and that, for any $\lambda_0 \in [0, \infty)$,

$$\lim_{R\to\infty} \partial_\lambda \mathbb{P}(E_R(\boldsymbol{x})) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \mathbb{P}(\boldsymbol{o} \longrightarrow \boldsymbol{x},\ b \text{ is pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}), \tag{3.38}$$

uniformly in $\lambda \in [0, \lambda_0]$, which we will show at the end of this section. Therefore, we can exchange the order of $\lim_{R\to\infty}$ and ∂_λ, and obtain

$$\partial_\lambda \tau(\boldsymbol{x}) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \mathbb{P}(\boldsymbol{o} \longrightarrow \boldsymbol{x},\ b \text{ is pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}). \tag{3.39}$$

We follow the same strategy as in Section 3.1 to obtain a recursion relation, now for $\partial_\lambda \tau(\boldsymbol{x})$ rather than for $\tau(\boldsymbol{x})$. Then, (3.39) equals

$$\begin{aligned} \partial_\lambda \tau(\boldsymbol{x}) = \frac{1}{\lambda} \sum_{b \text{ spatial}} \bigg[&\mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b},\ b \text{ occupied \& pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}) \\ &+ \sum_{b_1 < b} \mathbb{P}(\boldsymbol{o} \Longrightarrow \underline{b}_1,\ b_1 \text{ and } b \text{ occupied \& pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}) \bigg], \end{aligned} \tag{3.40}$$

where $\sum_{b_1<b}$ is the sum over bonds b_1 with $t_{\underline{b}_1} < t_{\underline{b}}$. The first and second terms in the brackets of the right-hand side correspond respectively to when b is or is not the first element of $\mathtt{piv}[\boldsymbol{o}, \boldsymbol{x}]$. The contribution from the first term is the same as (3.2), apart from the factor $\frac{1}{\lambda}$ and the restriction that b has to be a spatial bond. Thus, the first term equals

$$(\pi^{(0)} \star \varepsilon D \star \tau)(\boldsymbol{x}) + \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \big[\Pi^{(N;1,1)}(\boldsymbol{x}) + (\Pi^{(N;1,1)} \star p \star \tau)(\boldsymbol{x})\big], \tag{3.41}$$

where we abuse notation to write

$$D((y, s)) = D(y)\, \delta_{s,\varepsilon}. \tag{3.42}$$

For the second term in (3.40), we use

$$\begin{aligned} &\{b_1 \text{ and } b \text{ occupied \& pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}\} \\ &\quad = \{b_1 \text{ occupied \& pivotal for } \boldsymbol{o} \longrightarrow \boldsymbol{x}\} \cap \{b \in \mathtt{piv}[\overline{b}_1, \boldsymbol{x}]\}. \end{aligned} \tag{3.43}$$

We ignore the condition that b_1 is pivotal for $\boldsymbol{o} \longrightarrow \boldsymbol{x}$ and use inclusion-exclusion in the form (3.18) to make up for the arising error. Using the Markov property, the contribution from the second term in (3.40) is

$$\frac{1}{\lambda} \sum_{\substack{b_1, b: \\ b \text{ spatial}}} \mathbb{P}\big(\boldsymbol{o} \Longrightarrow \underline{b}_1,\ b_1 \longrightarrow \boldsymbol{x},\ b \in \texttt{piv}[\overline{b}_1, \boldsymbol{x}]\big) - Q^{(1)}(\boldsymbol{x}) = (\pi^{(0)} \star p \star \partial_\lambda \tau)(\boldsymbol{x}) - Q^{(1)}(\boldsymbol{x}), \tag{3.44}$$

where we define $Q^{(n)}(\boldsymbol{x})$ by

$$Q^{(n)}(\boldsymbol{x}) = \frac{1}{\lambda} \sum_{\substack{\vec{b}_n, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n-1)}_{\vec{b}_{n-1}}(\underline{b}_n) \cap \{b_n \xrightarrow{\tilde{\mathbf{C}}^{b_n}(\overline{b}_{n-1})} \boldsymbol{x}\} \cap \{b \in \texttt{piv}[\overline{b}_n, \boldsymbol{x}]\}\big), \tag{3.45}$$

and we recall that $\tilde{E}^{(0)}(\underline{b}_1) = \{\boldsymbol{o} \Longrightarrow \underline{b}_1\}$ and $\overline{b}_0 = \boldsymbol{o}$. Note that $Q^{(n)}(\boldsymbol{x})$ is similar to $R^{(n)}(\boldsymbol{x})$ in (3.29), except for the factor $\frac{1}{\lambda}$, the sum over spatial bonds b, and the extra condition $b \in \texttt{piv}[\overline{b}_n, \boldsymbol{x}]$. Therefore, by (3.40)–(3.41) and (3.44), we have

$$\begin{aligned} \partial_\lambda \tau(\boldsymbol{x}) &= (\pi^{(0)} \star \varepsilon D \star \tau)(\boldsymbol{x}) + (\pi^{(0)} \star p \star \partial_\lambda \tau)(\boldsymbol{x}) - Q^{(1)}(\boldsymbol{x}) \\ &\quad + \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \big[\Pi^{(N;1,1)}(\boldsymbol{x}) + (\Pi^{(N;1,1)} \star p \star \tau)(\boldsymbol{x})\big]. \end{aligned} \tag{3.46}$$

Below, we will use inclusion-exclusion to prove that, for $n \geq 1$,

$$\begin{aligned} Q^{(n)}(\boldsymbol{x}) &= (\pi^{(n)} \star \varepsilon D \star \tau)(\boldsymbol{x}) + (\pi^{(n)} \star p \star \partial_\lambda \tau)(\boldsymbol{x}) - Q^{(n+1)}(\boldsymbol{x}) \\ &\quad + \frac{1}{\lambda} \sum_{N=n+1}^{\infty} (-1)^{N-n} \big[\Pi^{(N;n+1,1)}(\boldsymbol{x}) + (\Pi^{(N;n+1,1)} \star p \star \tau)(\boldsymbol{x})\big] \\ &\quad + \frac{1}{\lambda} \sum_{N=n}^{\infty} (-1)^{N-n} \big[\Pi^{(N;n,2)}(\boldsymbol{x}) + (\Pi^{(N;n,2)} \star p \star \tau)(\boldsymbol{x})\big]. \end{aligned} \tag{3.47}$$

Before proving (3.47), we complete the proof of (3.34) assuming (3.47). By repeated applications of (3.47) to (3.46) until the remainder term $Q^{(n)}(\boldsymbol{x})$ vanishes, we obtain

$$\begin{aligned} \partial_\lambda \tau(\boldsymbol{x}) &- (\pi \star \varepsilon D \star \tau)(\boldsymbol{x}) + (\pi \star p \star \partial_\lambda \tau)(\boldsymbol{x}) \\ &= \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \big[\Pi^{(N;1,1)}(\boldsymbol{x}) + (\Pi^{(N;1,1)} \star p \star \tau)(\boldsymbol{x})\big] \\ &\quad + \frac{1}{\lambda} \sum_{n=1}^{\infty} (-1)^n \sum_{N=n+1}^{\infty} (-1)^{N-n} \big[\Pi^{(N;n+1,1)}(\boldsymbol{x}) + (\Pi^{(N;n+1,1)} \star p \star \tau)(\boldsymbol{x})\big] \\ &\quad + \frac{1}{\lambda} \sum_{n=1}^{\infty} (-1)^n \sum_{N=n}^{\infty} (-1)^{N-n} \big[\Pi^{(N;n,2)}(\boldsymbol{x}) + (\Pi^{(N;n,2)} \star p \star \tau)(\boldsymbol{x})\big] \\ &= \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \big[\Pi^{(N)}(\boldsymbol{x}) + (\Pi^{(N)} \star p \star \tau)(\boldsymbol{x})\big]. \end{aligned} \tag{3.48}$$

Differentiating both sides of (3.28) and comparing with the above expression, we obtain

$$\partial_\lambda \pi(\boldsymbol{x}) + (\partial_\lambda \pi \star p \star \tau)(\boldsymbol{x}) = \frac{1}{\lambda} \sum_{N=1}^{\infty} (-1)^N \big[\Pi^{(N)}(\boldsymbol{x}) + (\Pi^{(N)} \star p \star \tau)(\boldsymbol{x})\big]. \tag{3.49}$$

Using this identity, we prove (3.34) by induction on $t_{\boldsymbol{x}}/\varepsilon$. Since $\pi((x,0)) = \delta_{o,x}$ and $\Pi^{(N)}((x,0)) = 0$ for all $N \geq 1$, we obtain (3.34) for $t_{\boldsymbol{x}}/\varepsilon = 0$. Suppose that (3.34) holds for all $t_{\boldsymbol{x}}/\varepsilon \leq m$. Then the contribution from the second term in the brackets of (3.49) equals the second term on the left-hand side of (3.49), and thus (3.34) for $t_{\boldsymbol{x}}/\varepsilon = m+1$ holds. This completes the inductive proof of (3.34).

In order to complete the proof of Proposition 3.4, we prove (3.47). Because of the condition $b_n \xrightarrow{\tilde{\mathbf{C}}^{b_n}(\overline{b}_{n-1})} \boldsymbol{x}$ in (3.45), either the event $E(b_n, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_n}(\overline{b}_{n-1}))$ occurs or there is an occupied bond $b_{n+1} \in \mathtt{piv}[\overline{b}_n, \boldsymbol{x}]$ for which the event $E(b_n, \underline{b}_{n+1}; \tilde{\mathbf{C}}^{b_n}(\overline{b}_{n-1}))$ occurs. The contribution from the former case to $Q^{(n)}(\boldsymbol{x})$ is

$$\frac{1}{\lambda} \sum_{\substack{\vec{b}_n, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n-1)}_{\vec{b}_{n-1}}(\underline{b}_n) \cap E(b_n, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_n}(\overline{b}_{n-1})) \cap \{b \in \mathtt{piv}[\overline{b}_n, \boldsymbol{x}]\}\big) = \frac{1}{\lambda} \Pi^{(n;n,2)}(\boldsymbol{x}). \tag{3.50}$$

The contribution from the latter case is, as in (3.40),

$$\begin{aligned} \frac{1}{\lambda} \sum_{\substack{\vec{b}_n, b: \\ b \text{ spatial}}} \Bigg[& \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}) \cap \{b \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\big) \\ & + \sum_{b_{n+1} < b} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1}) \cap \{b_{n+1} \text{ and } b \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\}\big) \\ & + \sum_{b_{n+1} > b} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1}) \cap \{b \text{ and } b_{n+1} \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\}\big) \Bigg], \end{aligned} \tag{3.51}$$

where the first, second and third terms in the brackets 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 $\boldsymbol{x}$. The first term is similar to that in (3.40), and its contribution equals, as in (3.41),

$$(\pi^{(n)} \star \varepsilon D \star \tau)(\boldsymbol{x}) + \frac{1}{\lambda} \sum_{N=n+1}^{\infty} (-1)^{N-n} \big[\Pi^{(N;n+1,1)}(\boldsymbol{x}) + (\Pi^{(N;n+1,1)} \star p \star \tau)(\boldsymbol{x})\big]. \tag{3.52}$$

For the second term in (3.51), we apply (3.43), with b_1 and $\boldsymbol{o}$ being replaced respectively by b_{n+1} and $\overline{b}_n$, and use the inclusion-exclusion relation (3.18) and the Markov property. Then, the contribution from the second term equals, as in (3.44),

$$\begin{aligned} & \frac{1}{\lambda} \sum_{\substack{\vec{b}_{n+1}, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1}) \cap \{b_{n+1} \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\} \cap \{b \in \mathtt{piv}[\overline{b}_{n+1}, \boldsymbol{x}]\}\big) \\ & \quad = \frac{1}{\lambda} \sum_{\substack{\vec{b}_{n+1}, b: \\ b \text{ spatial}}} \mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1}) \cap \{b_{n+1} \longrightarrow \boldsymbol{x}\} \cap \{b \in \mathtt{piv}[\overline{b}_{n+1}, \boldsymbol{x}]\}\big) - Q^{(n+1)}(\boldsymbol{x}) \\ & \quad = (\pi^{(n)} \star p \star \partial_\lambda \tau)(\boldsymbol{x}) - Q^{(n+1)}(\boldsymbol{x}). \end{aligned} \tag{3.53}$$

For the third term in (3.51), we use

$$\begin{aligned} & \{b \text{ and } b_{n+1} \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\} \\ & \quad = \{b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]\} \cap \{b_{n+1} \text{ occupied \& pivotal for } \overline{b}_n \longrightarrow \boldsymbol{x}\}. \end{aligned} \tag{3.54}$$

By the inclusion-exclusion relation (3.18), the contribution from the third term equals

$$\begin{aligned}
&\frac{1}{\lambda}\sum_{\substack{\vec{b}_{n+1},b:\\ b\text{ spatial}}}\mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1})\cap\{b\in\texttt{piv}[\overline{b}_n,\underline{b}_{n+1}]\}\cap\{b_{n+1}\text{ occupied \& pivotal for }\overline{b}_n\longrightarrow\boldsymbol{x}\}\big)\\
&\quad=\frac{1}{\lambda}\sum_{\substack{\vec{b}_{n+1},b:\\ b\text{ spatial}}}\Big[\mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1})\cap\{b\in\texttt{piv}[\overline{b}_n,\underline{b}_{n+1}]\}\cap\{b_{n+1}\longrightarrow\boldsymbol{x}\}\big)\\
&\qquad\qquad-\mathbb{P}\big(\tilde{E}^{(n)}_{\vec{b}_n}(\underline{b}_{n+1})\cap\{b\in\texttt{piv}[\overline{b}_n,\underline{b}_{n+1}]\}\cap\{b_{n+1}\xrightarrow{\tilde{\mathbf{C}}^{b_{n+1}}(\overline{b}_n)}\boldsymbol{x}\}\big)\Big],
\end{aligned}\tag{3.55}$$

where the first term equals, by the Markov property,

$$\frac{1}{\lambda}(\Pi^{(n;n,2)}\star p\star\tau)(\boldsymbol{x}).\tag{3.56}$$

For the second term in (3.55), we use the same argument as above (3.31). Because of the condition $b_{n+1}\xrightarrow{\tilde{\mathbf{C}}^{b_{n+1}}(\overline{b}_n)}\boldsymbol{x}$, either the event $E(b_{n+1},\boldsymbol{x};\tilde{\mathbf{C}}^{b_{n+1}}(\overline{b}_n))$ occurs or there is an occupied bond $b_{n+2}\in\texttt{piv}[\overline{b}_{n+1},\boldsymbol{x}]$ such that $E(b_{n+1},\underline{b}_{n+2};\tilde{\mathbf{C}}^{b_{n+1}}(\overline{b}_n))$ occurs. By repeated use of inclusion-exclusion and the Markov property, as above (3.31), the contribution from the second term in (3.55) equals

$$\frac{1}{\lambda}\sum_{N=n+1}^{\infty}(-1)^{N-n}\big[\Pi^{(N;n,2)}(\boldsymbol{x})+(\Pi^{(N;n,2)}\star p\star\tau)(\boldsymbol{x})\big].\tag{3.57}$$

Combining (3.50), (3.52)–(3.53) and (3.56)–(3.57), we obtain (3.47). This completes the proof of Proposition 3.4, assuming the uniformity of (3.38). ☐

Proof of the uniformity of (3.38). Given $\lambda_0\in[0,\infty)$, we prove that $\partial_\lambda\mathbb{P}(E_R(\boldsymbol{x}))$ converges to the right-hand side of (3.38), uniformly in $\lambda\in[0,\lambda_0]$.

Recall that $E_R(\boldsymbol{x})=\{\boldsymbol{o}\longrightarrow\boldsymbol{x}\text{ in }\square_R\times[0,t_{\boldsymbol{x}}]\}$. The difference between $\partial_\lambda\mathbb{P}(E_R(\boldsymbol{x}))$ and the right-hand side of (3.38) is bounded by

$$\begin{aligned}
&\frac{1}{\lambda}\sum_{\substack{b\text{ spatial}\\ b\subset\square_R\times[0,t_{\boldsymbol{x}}]}}\mathbb{P}\big(b\text{ occupied \& pivotal for }E_R(\boldsymbol{x}),\text{ but not pivotal for }\boldsymbol{o}\longrightarrow\boldsymbol{x}\big)\\
&\qquad+\frac{1}{\lambda}\sum_{\substack{b\text{ spatial}\\ b\not\subset\square_R\times[0,t_{\boldsymbol{x}}]}}\mathbb{P}(b\text{ occupied \& pivotal for }\boldsymbol{o}\longrightarrow\boldsymbol{x}).
\end{aligned}\tag{3.58}$$

First, we bound the second term, using $\{b\text{ occupied \& pivotal for }\boldsymbol{o}\longrightarrow\boldsymbol{x}\}\subset\{\boldsymbol{o}\longrightarrow\underline{b}\}\cap\{b\longrightarrow\boldsymbol{x}\}$ as well as the Markov property and (2.43), by

$$\varepsilon\sum_{b\not\subset\square_R\times[0,t_{\boldsymbol{x}}]}p^{\star t_{\underline{b}}/\varepsilon}(\underline{b})\,D(\overline{b}-\underline{b})\,p^{\star(t_{\boldsymbol{x}}-t_{\overline{b}})/\varepsilon}(\boldsymbol{x}-\overline{b})\le\varepsilon\sum_{j=1}^{t_{\boldsymbol{x}}/\varepsilon}(1-\varepsilon+\lambda\varepsilon)^{t_{\boldsymbol{x}}/\varepsilon-j}\sum_{y\in\mathbb{Z}^d:\|y\|_\infty\ge R}p_\varepsilon^{*(j-1)}(y),\tag{3.59}$$

where we take the sum over the spatial component of $\boldsymbol{x}$ to obtain the bound. Similarly to (2.44), this is further bounded, uniformly in λ and ε, by $c\delta'_R$ where $c=c(\lambda_0,t_{\boldsymbol{x}})$ and $\delta'_R=\delta'_R(\lambda_0)$ are some finite constants satisfying $\lim_{R\to\infty}\delta'_R=0$.

Next, we consider the first term in (3.58). Note that, if b is pivotal for $E_R(\boldsymbol{x})$, but not pivotal for $\boldsymbol{o}\longrightarrow\boldsymbol{x}$, then there must be a detour from some $\boldsymbol{y}\in\square_R\times[0,t_{\underline{b}}]$ to another $\boldsymbol{z}\in\square_R\times[t_{\overline{b}},t_{\boldsymbol{x}}]$ that passes through $(\mathbb{Z}^d \setminus \Box_R) \times [0, t_{\boldsymbol{x}}]$ without traversing b. Therefore, the event in the first term of (3.58) is a subset of

$$\bigcup_{\substack{\boldsymbol{y},\boldsymbol{z},\boldsymbol{u}\in\Box_R\times[0,t_{\boldsymbol{x}}]\\ \boldsymbol{v}\notin\Box_R\times[0,t_{\boldsymbol{x}}]}} \{\boldsymbol{o} \longrightarrow \boldsymbol{y}\} \cap \big\{\{\boldsymbol{y} \longrightarrow \underline{b},\ b \longrightarrow \boldsymbol{z}\} \circ \{\boldsymbol{y} \longrightarrow \boldsymbol{u},\ (\boldsymbol{u},\boldsymbol{v}) \longrightarrow \boldsymbol{z}\}\big\} \cap \{\boldsymbol{z} \longrightarrow \boldsymbol{x}\}, \tag{3.60}$$

where $(\boldsymbol{u}, \boldsymbol{v})$ is the first bond along the detour that *crosses* the boundary of $\Box_R \times [0, t_{\boldsymbol{x}}]$, so that it is a spatial bond, and $E_1 \circ E_2$ is the event that E_1 and E_2 occur *disjointly*, i.e., there is a bond set B such that E_1 occurs on B and E_2 occurs on the complement of B. By the Markov property, the three events joined by $\cap$ are independent of each other. For the middle event, we use the *van den Berg-Kesten (BK) inequality* [9], which asserts that $\mathbb{P}(E_1 \circ E_2) \le \mathbb{P}(E_1)\,\mathbb{P}(E_2)$ when both E_1 and E_2 are *increasing* events, i.e., E_1 and E_2 are more likely to occur as λ increases, as in (3.60). Sometimes, we also make use of the *van den Berg-Kesten-Reimer (BKR) inequality* [6], which proves $\mathbb{P}(E_1 \circ E_2) \le \mathbb{P}(E_1)\,\mathbb{P}(E_2)$ for *any* events E_1 and E_2. Then, we use (2.43) as in (3.59) and obtain that the first term in (3.58) (even when we sum over the spatial component of $\boldsymbol{x}$) is bounded by $c'\delta''_R$, where $c' = c'(\lambda_0, t_{\boldsymbol{x}}, \varepsilon)$ and $\delta''_R = \delta''_R(\lambda_0)$ are some finite constants satisfying $\lim_{R\to\infty} \delta''_R = 0$. This completes the proof of the uniformity in λ of (3.38).

In the above proof, we did not care about the uniformity in ε, since it has been fixed and positive in this section. In fact, the above constant c' is of order $O(\varepsilon^{-2})$ and diverges as $\varepsilon \to 0$. This is because the contribution from (3.60) involves the sums over $t_{\overline{b}}(= t_{\underline{b}} + \varepsilon), t_{\boldsymbol{y}}, t_{\boldsymbol{z}}, t_{\boldsymbol{v}}(= t_{\boldsymbol{u}} + \varepsilon) \in \varepsilon\mathbb{Z}_+$ that give rise to the factor ε^{-4}. However, the factor ε^{-2} is cancelled by the bond occupation probabilities of the *spatial* bonds b and $(\boldsymbol{u}, \boldsymbol{v})$, and therefore $c' = O(\varepsilon^{-2})$. We could improve this to $c' = O(1)$ by using the ideas in Section 4.1, and hence obtain the uniformity in ε as well, though this is not necessary here. $\Box$

4 Bounds on the lace expansion

In this section, we prove Propositions 2.3–2.4. By (3.26)–(3.27) and (3.34)–(3.35), it suffices to prove the following bounds on $\pi_t^{(N)}(x)$ and $\Pi_t^{(N;n)}(x)$ in order to prove these propositions.

Lemma 4.1. *Suppose that* (2.29) *holds for some* λ_0 *and all* $s \le t$.

(i) Let $d > 4$. *Then, there are* $\beta_0 > 0$ *and* $C_K < \infty$ *such that, for* $\lambda \le \lambda_0$, $\beta < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ *with* $2\varepsilon \le s \le t + \varepsilon$, *and* $q = 0, 2, 4$,

$$\sum_x |x|^q\, \pi_s^{(N)}(x) \le \frac{\varepsilon^2 (C_K\beta)^{1\vee N} \sigma^q N^{q/2}}{(1+s)^{(d-q)/2}}, \qquad \text{for } N \ge 0, \tag{4.1}$$

$$\sum_x \Pi_s^{(N;n)}(x) \le \frac{\varepsilon^2 (C_K\beta)^N}{(1+s)^{(d-2)/2}}, \qquad \text{for } N \ge n \ge 1. \tag{4.2}$$

(ii) Let $d \le 4$ *with* $\alpha = bd - \frac{4-d}{2} > 0$, $\mu \in (0, \alpha)$ *and* $t \le T \log T$. *Then, there are* $\beta_0 > 0$ *and* $C_K < \infty$ *such that, for* $\lambda \le \lambda_0$, $\beta_1 < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ *with* $2\varepsilon \le s \le t + \varepsilon$, *and* $q = 0, 2, 4$,

$$\sum_x |x|^q\, \pi_s^{(N)}(x) \le \frac{\varepsilon^2 C_K \beta_T (C_K\hat{\beta}_T)^{0\vee(N-1)} \sigma_T^q N^{q/2}}{(1+s)^{(d-q)/2}}, \qquad \text{for } N \ge 0, \tag{4.3}$$

$$\sum_x \Pi_s^{(N;n)}(x) \le \frac{\varepsilon^2 C_K \beta_T (C_K\hat{\beta}_T)^{N-1}}{(1+s)^{(d-2)/2}}, \qquad \text{for } N \ge n \ge 1, \tag{4.4}$$

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

Proof of Propositions 2.3–2.4 assuming Lemma 4.1. The inequalities (2.30) and (2.32) follow from (3.26)–(3.27), (3.34)–(3.35) and (4.1)–(4.2), if β is sufficiently small. The proof of (2.31) is the same as that of Proposition 2.2(ii) in [20, Section 4.3], together with (2.30). This completes the proof of Proposition 2.3.

Proposition 2.4 is proved by using (4.3)–(4.4) instead of (4.1)–(4.2), if $\hat{\beta}_T$ is sufficiently small. $\Box$

Lemma 4.1 is proved in Sections 4.1–4.3. In Section 4.1, we first introduce certain diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$ that are defined in terms of two-point functions, and prove that these diagram functions are upper bounds on $\pi_t^{(N)}(x)$ and $\Pi_t^{(N;n)}(x)$, respectively. Then, we bound these diagram functions assuming the bounds in (2.29) on the two-point function, for $d > 4$ in Section 4.2 and for $d \le 4$ in Section 4.3. Finally, in Section 4.4, we use these diagram functions to obtain finite-volume approximations of $\pi_t^{(N)}(x)$ and $\Pi_t^{(N;n)}(x)$, which will be used in Section 6 to prove the continuum limit $\varepsilon \downarrow 0$.

4.1 Bounds in terms of the diagram functions

In this section, we prove that $\pi_t^{(N)}(x)$ and $\Pi_t^{(N;n)}(x)$ are bounded by certain diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$ that are defined below in terms of two-point functions.

The strategy in this section is similar to [20, Section 4.1] for oriented percolation in $\mathbb{Z}^d \times \mathbb{Z}_+$, where bounds on $\pi_t^{(N)}(x)$ were proved by using some diagram functions arising from the Hara-Slade lace expansion. Since the expansion used in this paper is somewhat simpler, we can use simpler diagram functions. However, to consider the case $\varepsilon \ll 1$ as in [27], extra care is needed to obtain the factor ε^2 in (4.36)–(4.54).

4.1.1 Preliminaries

Before defining the diagram functions, we start by some preliminaries. For $\boldsymbol{v} = (v, s) \in \Lambda$ and a bond b, we write $\boldsymbol{v}_+ = (v, s+\varepsilon)$ and $\{\boldsymbol{v} \longrightarrow b\} = \{\boldsymbol{v} \longrightarrow \underline{b}\} \cap \{b \text{ occupied}\}$ (cf., Definition 3.1(i)). For convenience, we will also use abbreviations, such as

$$\{\boldsymbol{v} \longrightarrow b \longrightarrow \boldsymbol{x}\} = \{\boldsymbol{v} \longrightarrow b\} \cap \{\overline{b} \longrightarrow \boldsymbol{x}\}. \tag{4.5}$$

Let $I'(\boldsymbol{v}, \boldsymbol{x}, \boldsymbol{x}) = \{\boldsymbol{v} \longrightarrow \boldsymbol{x}\}$, and define, for $\boldsymbol{y} \neq \boldsymbol{x}$,

$$I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x}) = \left\{ \bigcup_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{y}}} \{\boldsymbol{v} \longrightarrow b \longrightarrow \boldsymbol{x}\} \right\} \cup \left\{ \bigcup_{\substack{b \text{ spatial:} \\ \overline{b}=\boldsymbol{y}}} \{\boldsymbol{v} \longrightarrow (\underline{b}, \underline{b}_+) \longrightarrow \boldsymbol{x},\ b \text{ occupied}\} \right\}. \tag{4.6}$$

We note that $I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x})$ for $\boldsymbol{y} \neq \boldsymbol{x}$ equals $I'(\boldsymbol{v}, \boldsymbol{x}, \boldsymbol{x})$ with an extra spatial bond b being embedded (or added) along the connection from $\boldsymbol{v}$ to $\boldsymbol{x}$. Denoting

$$I(b, \boldsymbol{y}, \boldsymbol{x}) = \{b \text{ occupied}\} \cap I'(\overline{b}, \boldsymbol{y}, \boldsymbol{x}), \tag{4.7}$$

we define

$$M(b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y}) = \big\{ I(b, \boldsymbol{y}, \boldsymbol{x}) \circ \{\boldsymbol{v} \longrightarrow \boldsymbol{x}\} \big\} \cup \big\{ \{b \longrightarrow \boldsymbol{x}\} \circ I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x}) \big\}. \tag{4.8}$$

Note that, when neither $\overline{b}$ nor $\boldsymbol{v}$ is $\boldsymbol{x}$, the event $M(b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$ equals $\{b \longrightarrow \boldsymbol{x}\} \circ \{\boldsymbol{v} \longrightarrow \boldsymbol{x}\}$ with an extra spatial bond being embedded either between $\overline{b}$ and $\boldsymbol{x}$ due to $I(b, \boldsymbol{y}, \boldsymbol{x})$, or between $\boldsymbol{v}$ and $\boldsymbol{x}$ due to $I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x})$. In addition, we define $M^+(b, b', \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$ to be $M(b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$ with the connection from $\overline{b}$ to $\boldsymbol{x}$ being replaced by $\overline{b} \longrightarrow b' \longrightarrow \boldsymbol{x}$. For example, the second event $\{b \longrightarrow \boldsymbol{x}\} \circ I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x})$ in (4.8) is simply replaced by $\{\overline{b} \longrightarrow b' \longrightarrow \boldsymbol{x}\} \circ I'(\boldsymbol{v}, \boldsymbol{y}, \boldsymbol{x})$ in the definition of $M^+(b, b', \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$. Replacing the first event $I(b, \boldsymbol{y}, \boldsymbol{x}) \circ \{\boldsymbol{v} \longrightarrow \boldsymbol{x}\}$ in (4.8) is more complicated, due to the three possibilities of embedding b' into $\boldsymbol{v} \longrightarrow b \longrightarrow \boldsymbol{x}$ in (4.6) and the other three possibilities of embedding b' into $\boldsymbol{v} \longrightarrow (\underline{b}, \underline{b}_+) \longrightarrow \boldsymbol{x}$ in (4.6), and therefore we refrain from giving a formula for $M^+(b, b', \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$.

Recall (3.26) and (3.35) for the definitions of $\pi^{(N)}(\boldsymbol{x})$ and $\Pi^{(N;n)}(\boldsymbol{x})$ that involve the event $\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x})$. Our first claim is that $\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x})$ satisfies the following successive relations:

Lemma 4.2. *For $N \geq 1$,*

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

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

$$\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \cap \{b \in \mathtt{piv}[\overline{b}_N, \underline{b}_{N+1}]\} \subset \bigcup_{\boldsymbol{v}\in\Lambda} \Big\{ \big\{ \tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v}) \big\} \circ M^+(b_N, b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y}) \Big\}, \tag{4.11}$$

and for $N > n \geq 1$,

$$\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \cap \{b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]\} \subset \bigcup_{\boldsymbol{v}\in\Lambda} \Big\{ \big\{ \tilde{E}^{(N-2)}_{\vec{b}_{N-2}}(\underline{b}_{N-1}) \cap M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v}) \cap \{b \in \mathtt{piv}[\overline{b}_n, \underline{b}_{n+1}]\} \big\} \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y}) \Big\}, \tag{4.12}$$

where $\overline{b}_0 = \overline{b}_{-1} = \boldsymbol{o}$, $M(b_0, \overline{b}_{-1}; \underline{b}_1, \boldsymbol{v}) = \{\boldsymbol{o} \Longrightarrow \underline{b}_1\} \cap I'(\boldsymbol{o}, \boldsymbol{v}, \underline{b}_1)$, $\tilde{E}^{(0)}_{\vec{b}_0}(\underline{b}_1) = \{\boldsymbol{o} \Longrightarrow \underline{b}_1\}$ and $\tilde{E}^{(-1)}_{\vec{b}_{-1}}(\underline{b}_0)$ equals the whole probability space.

We note that the left-hand side of (4.12) is the same as that of (4.11), except that b is pivotal for $\overline{b}_n \longrightarrow \underline{b}_{n+1}$ with $n < N$.

Proof. The relation (4.9) follows immediately from (3.24) and

$$E(b_N, \boldsymbol{x}; \tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})) \subset \{b_N \longrightarrow \boldsymbol{x}\} \circ \{\overline{b}_{N-1} \longrightarrow \boldsymbol{x}\} = M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{x}). \tag{4.13}$$

We only prove (4.10), since (4.11)–(4.12) 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 $\underline{b}_{N-1}$, it suffices to prove

$$M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \subset \bigcup_{\boldsymbol{v}\in\Lambda} \big\{ M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \boldsymbol{v}) \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y}) \big\}. \tag{4.14}$$

Recall that $M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \underline{b}_N) = \{b_{N-1} \longrightarrow \underline{b}_N\} \circ \{\overline{b}_{N-2} \longrightarrow \underline{b}_N\}$. The event in the left-hand side of (4.14) implies existence of $\boldsymbol{v} \in \mathbf{C}(\overline{b}_{N-1})$ such that $\boldsymbol{v} \longrightarrow \underline{b}_N$ and $M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})$ occur disjointly. Therefore,

$$\begin{aligned} &M(b_{N-1}, \overline{b}_{N-2}; \underline{b}_N, \underline{b}_N) \cap M(b_N, \overline{b}_{N-1}; \boldsymbol{x}, \boldsymbol{y}) \\ &\quad \subset \bigcup_{\boldsymbol{v}\in\Lambda} \Big\{ \big\{ \{b_{N-1} \longrightarrow \boldsymbol{v} \longrightarrow \underline{b}_N\} \circ \{\overline{b}_{N-2} \longrightarrow \underline{b}_N\} \big\} \\ &\qquad\qquad \cup \big\{ \{b_{N-1} \longrightarrow \underline{b}_N\} \circ \{\overline{b}_{N-2} \longrightarrow \boldsymbol{v} \longrightarrow \underline{b}_N\} \big\} \Big\} \circ M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y}). \end{aligned} \tag{4.15}$$

We investigate the vicinity of $\boldsymbol{v} \in \Lambda$ in (4.15), where there are two disjoint connections, $\boldsymbol{v} \longrightarrow \underline{b}_N$ and $\boldsymbol{v} \longrightarrow \boldsymbol{x}$. Since there is at most one temporal bond growing out of each vertex in Λ, at least one of the two connections has to use a spatial bond at $\boldsymbol{v}$. Therefore,

$$\{\boldsymbol{v} \longrightarrow \underline{b}_N\} \circ \{\boldsymbol{v} \longrightarrow \boldsymbol{x}\} \subset \bigcup_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{v}}} \Big\{ \big\{ \{b \longrightarrow \underline{b}_N\} \circ \{\underline{b} \longrightarrow \boldsymbol{x}\} \big\} \cup \big\{ \{(\underline{b}, \underline{b}_+) \longrightarrow \underline{b}_N\} \circ \{b \longrightarrow \boldsymbol{x}\} \big\} \Big\}. \tag{4.16}$$

Substituting this relation into (4.15) and relabelling $\overline{b} = \boldsymbol{v}$ in the latter event $\{(\underline{b}, \underline{b}_+) \longrightarrow \underline{b}_N\} \circ \{b \longrightarrow \boldsymbol{x}\}$, we obtain (4.14), and thus (4.10). This completes the proof. ∎

4.1.2 Diagrammatic bounds

Inspired by the successive relations (4.9)–(4.12), we inductively construct the diagram functions $P_t^{(N)}(x)$ and $\tilde{P}_t^{(N;n)}(x)$ as follows. For $b=(\boldsymbol{u},\boldsymbol{v})$ with $\boldsymbol{u}=(u,s)$ and $\boldsymbol{v}=(v,s+\varepsilon)$, we abuse notation to 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

$$\varphi(\boldsymbol{x}-\boldsymbol{u})=\delta_{\boldsymbol{u},\boldsymbol{x}}+(p\star\tau)(\boldsymbol{x}-\boldsymbol{u}), \tag{4.17}$$

and

$$L(\boldsymbol{u},\boldsymbol{v};\boldsymbol{x})=\begin{cases}\varphi(\boldsymbol{x}-\boldsymbol{u})\ (\tau\star\lambda\varepsilon D)(\boldsymbol{x}-\boldsymbol{v})+(\varphi\star\lambda\varepsilon D)(\boldsymbol{x}-\boldsymbol{u})\ \tau(\boldsymbol{x}-\boldsymbol{v}), & \text{if } \boldsymbol{u}\neq\boldsymbol{v},\\ (D\star\tau)(\boldsymbol{x}-\boldsymbol{u})\ (\tau\star\lambda\varepsilon D)(\boldsymbol{x}-\boldsymbol{u})+(D\star\tau\star\lambda\varepsilon D)(\boldsymbol{x}-\boldsymbol{u})\ \tau(\boldsymbol{x}-\boldsymbol{u}), & \text{if } \boldsymbol{u}=\boldsymbol{v}.\end{cases} \tag{4.18}$$

We define

$$P^{(0)}(\boldsymbol{x})=\delta_{\boldsymbol{o},\boldsymbol{x}}+\lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};\boldsymbol{x}), \tag{4.19}$$

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 $(\lambda\varepsilon D\star\tau)(\boldsymbol{x})$ and $(\tau\star\lambda\varepsilon D)(\boldsymbol{x})$ for the contribution from the first term in (4.18) with $\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{o}$, and $(\lambda\varepsilon D\star\tau\star\lambda\varepsilon D)(\boldsymbol{x})$ and $\tau(\boldsymbol{x})$ for the contribution from the second term in (4.18) with $\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{o}$.

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

$$\tau(\boldsymbol{y}-\boldsymbol{v})\ (\lambda\varepsilon D\star\tau)(\boldsymbol{x}-\boldsymbol{y}), \tag{4.20}$$

and Construction $B^\ell_{\mathtt{temp}}(\boldsymbol{y})$ is defined to be the operation in which $\tau(\boldsymbol{x}-\boldsymbol{v})$ is replaced by

$$\sum_{b:\bar{b}=\boldsymbol{y}}\tau(\underline{b}-\boldsymbol{v})\ \lambda\varepsilon D(b)\ \mathbb{P}((\underline{b},\underline{b}_+)\longrightarrow\boldsymbol{x}). \tag{4.21}$$

We note that (4.20)–(4.21) are inspired by (4.6). The sum of the results of Construction $B^\ell_{\mathtt{spat}}(\boldsymbol{y})$ and Construction $B^\ell_{\mathtt{temp}}(\boldsymbol{y})$ is simply said to be the result of Construction $B^\ell(\boldsymbol{y})$. We define Construction $B^\ell(s)$ to be the operation in which Construction $B^\ell(y,s)$ is performed and then followed by summation over $y\in\mathbb{Z}^d$. Construction $B^\ell_{\mathtt{spat}}(s)$ and Construction $B^\ell_{\mathtt{temp}}(s)$ are defined similarly.

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

$$P^{(N)}(\boldsymbol{x})=2\lambda\varepsilon\sum_{\boldsymbol{v}\in\Lambda}P^{(N-1)}(\boldsymbol{v})\ L(\boldsymbol{v},\boldsymbol{v};\boldsymbol{x})+\sum_\ell\sum_{\substack{\boldsymbol{v},\boldsymbol{y}\in\Lambda\\ \boldsymbol{v}\neq\boldsymbol{y}}}P^{(N-1)}(\boldsymbol{v},B^\ell(\boldsymbol{y}))\ L(\boldsymbol{v},\boldsymbol{y};\boldsymbol{x}), \tag{4.22}$$

where $\sum_\ell$ is the sum over the $(N-1)^{\text{st}}$ admissible lines in each diagram. We define

$$\sum_\ell P^{(N-1)}(\boldsymbol{v},B^\ell(\boldsymbol{v}))=2\lambda\varepsilon P^{(N-1)}(\boldsymbol{v}). \tag{4.23}$$

Then, (4.22) equals

$$P^{(N)}(\boldsymbol{x})=\sum_\ell\sum_{\boldsymbol{v},\boldsymbol{y}\in\Lambda}P^{(N-1)}(\boldsymbol{v},B^\ell(\boldsymbol{y}))\ L(\boldsymbol{v},\boldsymbol{y};\boldsymbol{x}). \tag{4.24}$$

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

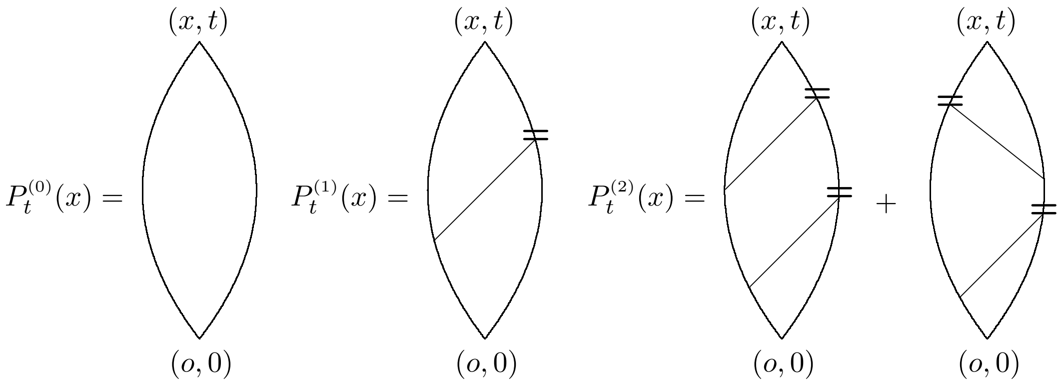

Figure 5: Graphical representations of $P_t^{(0)}(x), P_t^{(1)}(x)$ and $P_t^{(2)}(x)$. Lines indicate two-point functions, and small bars indicate a convolution with p_ε. Spatial bonds that are present at all vertices in the diagrams are left implicit.

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

$$\tilde{P}^{(N;n)}(\boldsymbol{x}) = \sum_{\ell} \sum_{\boldsymbol{y} \in \Lambda} P^{(N)}(\boldsymbol{x}, B^{\ell}_{\mathtt{spat}}(\boldsymbol{y})), \tag{4.25}$$

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

Thanks to the construction in terms of two-point functions, the diagram functions can be estimated by using (2.29), and this will be done in Sections 4.2–4.3. The following is the main statement of this section:

Lemma 4.3. *For $\lambda \geq 0$ and $N \geq n \geq 1$,*

$$\pi_t^{(N-1)}(x) \leq P_t^{(N-1)}(x), \qquad \Pi_t^{(N;n)}(x) \leq \tilde{P}_t^{(N;n)}(x). \tag{4.26}$$

Proof. We begin with proving $\pi^{(0)}(\boldsymbol{x}) \leq P^{(0)}(\boldsymbol{x})$. The first term in (4.19) 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 (4.15), at least one of two nonzero disjoint occupied paths from $\boldsymbol{o}$ has to use a spatial bond at $\boldsymbol{o}$. That is,

$$\pi^{(0)}(\boldsymbol{x}) \leq \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(\{b \longrightarrow \boldsymbol{x}\} \circ \{\boldsymbol{o} \longrightarrow \boldsymbol{x}\}) = \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})). \tag{4.27}$$

We use the same observation at $\boldsymbol{x}$: at least one of the two nonzero disjoint connections, $\overline{b} \longrightarrow \boldsymbol{x}$ and $\boldsymbol{o} \longrightarrow \boldsymbol{x}$, has to use another spatial bond at $\boldsymbol{x}$. Therefore, we can bound the right-hand side of (4.27) by $\lambda\varepsilon L(\boldsymbol{o}, \boldsymbol{o}; \boldsymbol{x})$ using the BK inequality. This completes the proof of $\pi^{(0)}(\boldsymbol{x}) \leq P^{(0)}(\boldsymbol{x})$.

Next, we consider $\pi^{(N)}(\boldsymbol{x})$ for $N \geq 1$. Let

$$\pi^{(n)}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\vec{b}_n} \mathbb{P}\big(\tilde{E}^{(n-1)}_{\vec{b}_{n-1}}(\underline{b}_n) \cap M(b_n, \overline{b}_{n-1}; \boldsymbol{x}, \boldsymbol{y})\big). \tag{4.28}$$

By the convention in Lemma 4.2, $\pi^{(0)}(\boldsymbol{x}, \boldsymbol{y}) = \mathbb{P}(\{\boldsymbol{o} \Longrightarrow \boldsymbol{x}\} \cap I'(\boldsymbol{o}, \boldsymbol{y}, \boldsymbol{x}))$. We prove below by induction that

$$\pi^{(n)}(\boldsymbol{x}, \boldsymbol{y}) \leq (2\lambda\varepsilon)^{-\delta_{\boldsymbol{x},\boldsymbol{y}}} \sum_{\ell} P^{(n)}(\boldsymbol{x}, B^{\ell}(\boldsymbol{y})) \tag{4.29}$$

holds for all $n \geq 0$, where $\sum_\ell$ is the sum over the n^{th} admissible lines. The inequality $\pi^{(N)}(\boldsymbol{x}) \leq P^{(N)}(\boldsymbol{x})$ for $N \geq 1$ follows from (4.9) and (4.29) for $\boldsymbol{y} = \boldsymbol{x}$, together with the convention in (4.24), i.e., $2\lambda\varepsilon P^{(N)}(\boldsymbol{x}) = \sum_\ell P^{(N)}(\boldsymbol{x}, B^\ell(\boldsymbol{x}))$.

For $n = 0$, we can assume $\boldsymbol{y} \neq \boldsymbol{x}$, since $\pi^{(0)}(\boldsymbol{x}) = \pi^{(0)}(\boldsymbol{x}, \boldsymbol{x}) \leq P^{(0)}(\boldsymbol{x})$ has already been proved. By the equivalence $\{\boldsymbol{o} \Longrightarrow \boldsymbol{x}\} \cap I'(\boldsymbol{o}, \boldsymbol{y}, \boldsymbol{x}) = \{\boldsymbol{o} \longrightarrow \boldsymbol{x}\} \circ I'(\boldsymbol{o}, \boldsymbol{y}, \boldsymbol{x})$ and by (4.6), we obtain

$$\begin{aligned}\pi^{(0)}(\boldsymbol{x}, \boldsymbol{y}) \leq & \sum_{\substack{b \text{ spatial:} \\ \underline{b} = \boldsymbol{y}}} \mathbb{P}(\{\boldsymbol{o} \longrightarrow \boldsymbol{x}\} \circ \{\boldsymbol{o} \longrightarrow b \longrightarrow \boldsymbol{x}\}) \\ & + \sum_{b : \overline{b} = \boldsymbol{y}} \lambda\varepsilon D(b)\ \mathbb{P}(\{\boldsymbol{o} \longrightarrow \boldsymbol{x}\} \circ \{\boldsymbol{o} \longrightarrow (\underline{b}, \underline{b}_+) \longrightarrow \boldsymbol{x}\}),\end{aligned} \tag{4.30}$$

where we use the BK inequality to derive $\lambda\varepsilon D(b)$ in the second sum.

We only prove that the first sum in (4.30) 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 $\pi^{(0)}(\boldsymbol{x}) \leq P^{(0)}(\boldsymbol{x})$. The second sum in (4.30) can be proved similarly to be bounded by $\sum_\ell P^{(0)}(\boldsymbol{x}, B^\ell_{\texttt{temp}}(\boldsymbol{y}))$. In the first sum in (4.30), there are three contributions: (i) $\boldsymbol{y} = \boldsymbol{o}$, (ii) $b = (\boldsymbol{y}, \boldsymbol{x})$ and (iii) $\boldsymbol{y} \neq \boldsymbol{o}$ and $\overline{b} \neq \boldsymbol{x}$. The contribution due to $\boldsymbol{y} = \boldsymbol{o}$ is bounded by

$$\lambda\varepsilon L(\boldsymbol{o}, \boldsymbol{o}; \boldsymbol{x}) = \sum_\ell P^{(0)}(\boldsymbol{x}, B^\ell_{\texttt{spat}}(\boldsymbol{o})) - 2(\lambda\varepsilon D \star \tau)(\boldsymbol{x})\ (\lambda\varepsilon D \star \tau \star \lambda\varepsilon D)(\boldsymbol{x}), \tag{4.31}$$

while the contribution due to $b = (\boldsymbol{y}, \boldsymbol{x})$ is bounded by

$$\begin{aligned}&\mathbb{P}(\{\boldsymbol{o} \longrightarrow \boldsymbol{x}\} \circ \{\boldsymbol{o} \longrightarrow \boldsymbol{y}\})\ \lambda\varepsilon D(\boldsymbol{x} - \boldsymbol{y}) \leq [(\lambda\varepsilon D \star \tau)(\boldsymbol{x})\ \tau(\boldsymbol{y}) + (\lambda\varepsilon D \star \tau)(\boldsymbol{y})\ \tau(\boldsymbol{x})]\ \lambda\varepsilon D(\boldsymbol{x} - \boldsymbol{y}) \\ &\quad = \sum_\ell P^{(0)}(\boldsymbol{x}, B^\ell_{\texttt{spat}}(\boldsymbol{y})) - [(\lambda\varepsilon D \star \tau)(\boldsymbol{y})\ (\tau \star \lambda\varepsilon D)(\boldsymbol{x}) + (\lambda\varepsilon D \star \tau \star \lambda\varepsilon D)(\boldsymbol{x})\ \tau(\boldsymbol{y})]\ \lambda\varepsilon D(\boldsymbol{x} - \boldsymbol{y}).\end{aligned} \tag{4.32}$$

We can estimate the case (iii) similarly, and obtain a bound which is $\lambda\varepsilon L(\boldsymbol{o}, \boldsymbol{y}; \boldsymbol{x})$ with one of the two τ's in each product $\tau \cdot \tau$ in (4.18) replaced by (4.20). Summarizing these bounds, we conclude that the first sum in (4.30) is bounded by $\sum_\ell P^{(0)}(\boldsymbol{x}, B^\ell_{\texttt{spat}}(\boldsymbol{y}))$. This completes the proof of (4.29) for $n = 0$, and initializes the inductive proof for $n \geq 1$.

To advance the induction hypotheses, we assume that (4.29) holds for $n = N - 1$, and that

$$\sum_{b : \underline{b} = \boldsymbol{u}} \mathbb{P}(M(b, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})) \leq (2\lambda\varepsilon)^{\delta_{\boldsymbol{u}, \boldsymbol{v}} - \delta_{\boldsymbol{x}, \boldsymbol{y}}} \sum_\ell L(\boldsymbol{u}, \boldsymbol{v}; \boldsymbol{x}, B^\ell(\boldsymbol{y})) \tag{4.33}$$

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 (4.24). We will prove (4.33) below. By (4.10) and the BKR inequality, together with (4.24), (4.29) and (4.33), we obtain

$$\begin{aligned}\pi^{(N)}(\boldsymbol{x}, \boldsymbol{y}) &\leq \sum_{\boldsymbol{u}, \boldsymbol{v}} \pi^{(N-1)}(\boldsymbol{u}, \boldsymbol{v}) \sum_{b_N : \underline{b}_N = \boldsymbol{u}} \mathbb{P}(M(b_N, \boldsymbol{v}; \boldsymbol{x}, \boldsymbol{y})) \\ &\leq (2\lambda\varepsilon)^{-\delta_{\boldsymbol{x}, \boldsymbol{y}}} \sum_{\ell, \ell'} \sum_{\boldsymbol{u}, \boldsymbol{v}} P^{(N-1)}(\boldsymbol{u}, B^{\ell'}(\boldsymbol{v}))\ L(\boldsymbol{u}, \boldsymbol{v}; \boldsymbol{x}, B^\ell(\boldsymbol{y})) \\ &= (2\lambda\varepsilon)^{-\delta_{\boldsymbol{x}, \boldsymbol{y}}} \sum_\ell P^{(N)}(\boldsymbol{x}, B^\ell(\boldsymbol{y})).\end{aligned} \tag{4.34}$$

This advances the induction hypotheses, and hence completes the proof of (4.29), assuming that (4.33) holds.

It thus remains to prove (4.33). 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 $\boldsymbol{x} \neq \boldsymbol{y}$ can be proved similarly to (4.29) for $n = 0$, and the proof of the remaining case is simpler because extracting the factors of ε at $\boldsymbol{u} \neq \boldsymbol{v}$ is unnecessary. Let $\boldsymbol{u} = \boldsymbol{o}$ in (4.33), by translation invariance. Then, the left-hand side of (4.33) equals the rightmost expression in (4.27), except for the condition that b is a spatial bond. By the same observation at $\boldsymbol{o}$ as in (4.27), we obtain

$$\begin{aligned}\sum_{b:\underline{b}=\boldsymbol{o}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})) &= \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})) + \mathbb{P}(M((\boldsymbol{o}, \boldsymbol{o}_+), \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})) \\ &\leq \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})) + \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(\{(\boldsymbol{o}, \boldsymbol{o}_+) \longrightarrow \boldsymbol{x}\} \circ \{b \longrightarrow \boldsymbol{x}\}) \\ &\leq 2 \sum_{\substack{b \text{ spatial:} \\ \underline{b}=\boldsymbol{o}}} \mathbb{P}(M(b, \boldsymbol{o}; \boldsymbol{x}, \boldsymbol{x})) \leq 2\lambda\varepsilon L(\boldsymbol{o}, \boldsymbol{o}; \boldsymbol{x}). \end{aligned} \tag{4.35}$$

This completes the proof of (4.33), and hence the proof of the first inequality in (4.26).

To prove the second inequality in (4.26), 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 (3.35) 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 $(p \star \tau)(\underline{b}_{n+1} - \underline{b}_n)$, replaced by $(\lambda\varepsilon D \star \tau)(\underline{b}_{n+1} - \underline{b}_n)$; 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)–(4.12) 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 bound on $\Pi^{(N;n,1)}(\boldsymbol{x})$, we obtain (4.25). This completes the proof the second inequality in (4.26), and hence the proof of Lemma 4.3. ☐

4.2 Estimate of the diagram functions above four dimensions

In this section, we bound the diagram functions for $d > 4$ as follows:

Lemma 4.4. *Let $d > 4$ and suppose that (2.29) holds for some λ_0 and all $s \leq t$. Then, there are $\beta_0 > 0$ and $C_K < \infty$ such that, for $\lambda \leq \lambda_0$, $\beta < \beta_0$, $s \in \varepsilon\mathbb{Z}_+$ with $2\varepsilon \leq s \leq t + \varepsilon$, and $q = 0, 2, 4$,*

$$\sum_x |x|^q \, P_s^{(N)}(x) \leq \frac{\varepsilon^2 (C_K\beta)^{1\vee N} \sigma^q N^{q/2}}{(1+s)^{(d-q)/2}}, \qquad \text{for } N \geq 0, \tag{4.36}$$

$$\sum_x \tilde{P}_s^{(N;n)}(x) \leq \frac{\varepsilon^2 (C_K\beta)^N}{(1+s)^{(d-2)/2}}, \qquad \text{for } N \geq n \geq 1. \tag{4.37}$$

Lemma 4.1(i) is an immediate consequence of Lemmas 4.3–4.4. To prove Lemma 4.4, we will use the following three lemmas.

Lemma 4.5. *Assume (2.29) for $s \leq t = n\varepsilon$ and $\lambda \in I_n$. Then, there is a $C_K = C_K(d, \lambda) < \infty$ such that the following bounds hold for $s \leq t$, $q = 0, 2$ and for that λ:*

$$\sum_x |x|^q \, (\tau_s * D)(x) \leq C_K \sigma^q (1+s)^{q/2}, \tag{4.38}$$

$$\sup_x |x|^q \, (\tau_s * D)(x) \leq \frac{C_K \sigma^q \beta}{(1+s)^{(d-q)/2}}, \tag{4.39}$$

$$\sup_x |x|^q \, \tau_s(x) \leq (1-\varepsilon)^{s/\varepsilon} \delta_{q,0} + \frac{C_K \sigma^q \beta}{(1+s)^{(d-q)/2}}. \tag{4.40}$$

Lemma 4.6. *Assume (2.29) for $s \leq t$. Let $f_t(x)$ be a diagram function that satisfies $\sum_x f_t(x) \leq F(t)$ by assigning l_1 or l_∞ norm to each diagram line and using (2.29) to estimate those norms. Then, there is a $C_K = C_K(d, \lambda) < \infty$ such that $\sum_x f_t(x, B^\ell(s)) \leq \varepsilon C_K F(t)$ for every $s \leq t$ and every admissible line ℓ.*

Lemma 4.7. *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 a $C = C(a, b, \kappa) < \infty$ such that*

$$\sum_{s_1=0}^{t}{}^{\bullet} \frac{\varepsilon}{(1+s_1)^a} \sum_{s_2=t-s_1}^{t}{}^{\bullet} \frac{\varepsilon}{(1+s_2)^b} \leq \frac{C}{(1+t)^{a\wedge b\wedge(a+b-2)-\kappa}}. \tag{4.41}$$

Lemmas 4.5 and 4.6 correspond respectively to Lemmas 4.3 and 4.6(a) in [20], and Lemma 4.7 corresponds to [19, Lemma 3.2]. The result of applying Lemma 4.7 is the same as [20, (4.26)] when $d > 4$ (see also the proof of Lemma 4.6(b) in [20]), but Lemma 4.7 can be applied to the lower dimensional case as well. We will use Lemmas 4.5–4.6 again in Sections 4.3–4.4, and Lemma 4.7 in Section 4.3 and Appendix A.

First, we prove Lemma 4.4 assuming Lemmas 4.5–4.7, and then prove these lemmas. We will use c to denote a finite positive constant whose exact value is unimportant and may change from line to line.

Proof of Lemma 4.4 assuming Lemmas 4.5–4.7. For (4.36), we only consider $q = 0$, since the other cases in (4.36) are proved along the same line of argument as in the last paragraph of [20, Section 4.2].

For $s' \leq s < t$, we use Lemma 4.5 to obtain

$$\sup_{u,v} \sum_x L((u,s),(v,s');(x,t)) \leq \frac{c\beta\varepsilon}{(1+t-s')^{d/2}}. \tag{4.42}$$

Since $P_t^{(0)}(x) = \lambda\varepsilon L(\boldsymbol{o},\boldsymbol{o};(x,t))$ for $t \geq 2\varepsilon$, this implies (4.36) with $q = N = 0$. By Lemma 4.6, we also obtain $\sum_\ell \sum_x P_t^{(0)}(x, B^\ell(s)) \leq c\varepsilon[\delta_{0,t} + \varepsilon^2\beta\,(1+t)^{-d/2}]$, where $\delta_{0,t}$ is the contribution from the first term in (4.19).

For $N \geq 1$, we note that, by (4.24) we have

$$\sum_x P_t^{(N)}(x) \leq \sum_{s,s'}{}^{\bullet} \bigg[\sum_\ell \sum_{u,v} P^{(N-1)}((u,s), B^\ell(v,s'))\bigg]\bigg[\sup_{u,v}\sum_x L((u,s),(v,s');(x,t))\bigg], \tag{4.43}$$

where $\sum_\ell$ is the sum over the $(N-1)^{\rm st}$ admissible lines in $P_s^{(N-1)}(u)$. Therefore, $\sum_x P_t^{(1)}(x)$ satisfies (4.36), and $\sum_\ell \sum_x P_t^{(1)}(x, B^\ell(s))$ is bounded by $c\varepsilon^3 C_K\beta\,(1+t)^{-d/2}$. This initializes the inductive proof of (4.36) for $N \geq 1$ with $q = 0$. Suppose that $\sum_x P_t^{(N-1)}(x, B^\ell(s))$ is bounded by $c\varepsilon^3(C_K\beta)^{N-1}(1+t)^{-d/2}$. Then, by (4.42) and Lemma 4.7, $\sum_x P_t^{(N)}(x)$ is bounded by $\varepsilon^2(C_K\beta)^N(1+t)^{-d/2}$ if C_K is sufficiently large. Note that the factor ε^2 is used in applying Lemma 4.7, to approximate $\varepsilon\sum^\bullet_{s\in\varepsilon\mathbb{Z}_+}$ by the Riemann sum. Using Lemma 4.6, we then obtain $\sum_x P_t^{(N)}(x, B^\ell(s)) \leq c\varepsilon^3(C_K\beta)^N(1+t)^{-d/2}$. This completes the proof of (4.36).

To prove (4.37), we first use Lemma 4.6 to obtain $\sum_x P_t^{(N)}(x, B^\ell_{\texttt{spat}}(s)) \leq c\varepsilon^3(C_K\beta)^N(1+t)^{-d/2}$ for every $s \leq t$, where ℓ is an $n^{\rm th}$ admissible line. Then, we sum the bound over $s \in [0,t]\cap\varepsilon\mathbb{Z}_+$ to obtain the desired bound in (4.37). Note that the factor ε is used in an approximation by the Riemann sum. This completes the proof. □

Proof of Lemma 4.5. The inequality (4.38) immediately follows from (2.29) and the properties of D. To prove (4.39)–(4.40), we use

$$(\tau_s * D)(x) \leq (1-\varepsilon)^{s/\varepsilon}D(x) + \lambda\varepsilon\sum_{j=1}^{s/\varepsilon}(1-\varepsilon)^{j-1}(D*\tau_{s-j\varepsilon}*D)(x), \tag{4.44}$$

and

$$\tau_s(x) \leq (1-\varepsilon)^{s/\varepsilon}\delta_{o,x} + \lambda s(1-\varepsilon)^{s/\varepsilon-1}D(x) + \lambda^2\varepsilon\sum_{j=2}^{s/\varepsilon}(j-1)\varepsilon\,(1-\varepsilon)^{j-2}(D*\tau_{s-j\varepsilon}*D)(x). \tag{4.45}$$

where $(1-\varepsilon)^n$ is the probability that $(o,0)$ is connected to $(o,n\varepsilon)$ along the temporal axis. Since the first term in (4.44) and the second term in (4.45), multiplied by $|x|^q$, are both bounded by $c\sigma^q\beta\,(1+s)^{-(d-q)/2}$ for any $x\in\mathbb{Z}^d$, we only need to consider the last terms in (4.44)–(4.45).

We fix $r\in(0,1)$ and use (2.29) to bound the second term in (4.44) by

$$\lambda\varepsilon\sum_{j=1}^{rs/\varepsilon}\frac{K\beta\,(1-\varepsilon)^{j-1}}{(1+s-j\varepsilon)^{d/2}}+\lambda\varepsilon\sum_{j=rs/\varepsilon}^{s/\varepsilon}\frac{K\beta\,(1-\varepsilon)^{j-1}}{(1+s-j\varepsilon)^{d/2}}\le\frac{c\beta}{(1+s)^{d/2}}. \tag{4.46}$$

By the same argument, the third term in (4.45) can be bounded by $c\beta\,(1+s)^{-d/2}$. This completes the proof of (4.39)–(4.40) for $q=0$.

For $|x|^2$ times the second term in (4.44) or the third term in (4.45), we have

$$\begin{aligned}|x|^2(D*\tau_s*D)(x)&\le 2\sum_y(|y|^2+|x-y|^2)\,(D*\tau_{s/2})(y)\,(\tau_{s/2}*D)(x-y)\\&\le 4\|D*\tau_{s/2}\|_\infty\sum_x|x|^2(\tau_{s/2}*D)(x).\end{aligned}\tag{4.47}$$

Applying (4.44) to $\|D*\tau_{s/2}\|_\infty$, using (2.29), and then separating the sum over j as in (4.46), we obtain (4.39)–(4.40) for $q=2$. This completes the proof of Lemma 4.5. □

Proof of Lemma 4.6. By the convention used in (4.24), for $s=t$ we have

$$\sum_{x,y}f_t(x,B^\ell(y,t))=\sum_x f_t(x,B^\ell(x,t))\le\sum_x\sum_\ell f_t(x,B^\ell(x,t))=\sum_x 2\lambda\varepsilon f_t(x)\le 2\lambda\varepsilon F(t), \tag{4.48}$$

where $\sum_\ell$ is the sum over the admissible lines arriving at (x,t). For $s<t$, Construction $B^\ell(s)$ replaces the diagram line ℓ, say $\tau_t(x)$, by $\lambda\varepsilon(\tau_s*D*\tau_{t-s-\varepsilon})(x)+\lambda\varepsilon(1-\varepsilon)(\tau_s*\tau_{t-s-\varepsilon})(x)$. By Lemma 4.5 and (2.29), we obtain

$$\lambda\varepsilon\|\tau_s*D*\tau_{t-s-\varepsilon}\|_1=\lambda\varepsilon\|\tau_s*D\|_1\,\|\tau_{t-s-\varepsilon}\|_1\le\lambda\varepsilon C_{4.5}K, \tag{4.49}$$

$$\lambda\varepsilon\|\tau_s*D*\tau_{t-s-\varepsilon}\|_\infty\le\lambda\varepsilon\|\tau_{s\vee(t-s-\varepsilon)}*D\|_\infty\,\|\tau_{s\wedge(t-s-\varepsilon)}\|_1\le\lambda\varepsilon\frac{2^{d/2}C_{4.5}K}{(1+t)^{d/2}}, \tag{4.50}$$

where $C_{4.5}$ is the constant in Lemma 4.5, and we use $s\vee(t-s-\varepsilon)\ge t/2$ to obtain the last inequality. The l_1 and l_∞ norms of $\lambda\varepsilon(1-\varepsilon)(\tau_s*\tau_{t-s-\varepsilon})(x)$ can be estimated in the same way. Therefore, the effect of Construction $B^\ell(s)$ is to obtain, at worst, an additional constant $C_K\varepsilon=2^{1+d/2}C_{4.5}\lambda\varepsilon$ in a bound. This completes the proof. □

Proof of Lemma 4.7. We prove (4.41) for $a\wedge b\ge 0$ and for $a\wedge b<0$ separately.

Let $a\wedge b\ge 0$. Separating 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 (4.41) by

$$\frac{c}{(1+t)^b}\sum_{s_1=0}^{t/2}{}^{\bullet}\frac{\varepsilon}{(1+s_1)^{a-1}}+\frac{c}{(1+t)^a}\sum_{s_1=t/2}^{t}{}^{\bullet}\sum_{s_2=t-s_1}^{t}{}^{\bullet}\frac{\varepsilon^2}{(1+s_2)^b}, \tag{4.51}$$

where the first term is bounded by $c'(1+t)^{-b-(a-2)\wedge 0+\kappa_a}$, where κ_a is an arbitrarily small but positive number if $a=2$, otherwise $\kappa_a=0$. Also, the double sum in (4.51) is

$$\sum_{s_2=0}^{t}{}^{\bullet}\frac{\varepsilon}{(1+s_2)^b}\sum_{s_1=t/2\vee(t-s_2)}^{t}{}^{\bullet}\varepsilon\;\le\;\sum_{s_2=0}^{t}\frac{\varepsilon}{(1+s_2)^{b-1}}\;\le\;\frac{c}{(1+t)^{(b-2)\wedge 0-\kappa_b}}, \tag{4.52}$$

where κ_b is an arbitrarily small but positive number if $b = 2$, otherwise $\kappa_b = 0$. This completes the proof of (4.41) for $a \wedge b \geq 0$.

Next, we consider the case $a \wedge b < 0$. Due to the symmetry of the left-hand side of (4.41) in terms of s_1, s_2, we suppose $b < a \wedge 0$. Then, we use the trivial inequality $(1 + s_2)^{-b} \leq (1 + t)^{-b}$. The remaining term equals $\varepsilon \sum_{0 \leq s_1 \leq t}^{\bullet} s_1 (1 + s_1)^{-a}$ and is bounded by $c\,(1 + t)^{-(a-2) \wedge 0 + \kappa_a}$. This completes the proof of (4.41) for $a \wedge b < 0$, and hence the proof of Lemma 4.7. ☐

4.3 Estimate of the diagram functions at and below four dimensions

In this section, we bound the diagram functions for $d \leq 4$ by using their inductive construction in (4.19) and (4.24)–(4.25) as well as Lemmas 4.5–4.7, as in the proof of Lemma 4.4 in Section 4.2, but we replace σ and β in Lemma 4.5 by σ_T and β_T, respectively.

Lemma 4.1(ii) is an immediate consequence of Lemma 4.3 and the following lemma:

Lemma 4.8. *Let $d \leq 4$ with $\alpha = bd - \frac{4-d}{2} > 0$, $\mu \in (0, \alpha)$ and $t \leq T \log T$, and suppose that* (2.29) *holds for some λ_0 and all $s \leq t$. Then, there are $\beta_0 > 0$ and $C_K < \infty$ 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 $q = 0, 2, 4$,*

$$\sum_x |x|^q \, P_s^{(N)}(x) \leq \frac{\varepsilon^2 C_K \beta_T (C_K \hat{\beta}_T)^{0 \vee (N-1)} \sigma_T^q N^{q/2}}{(1 + s)^{(d-q)/2}}, \qquad \textit{for } N \geq 0, \tag{4.53}$$

$$\sum_x \tilde{P}_s^{(N;n)}(x) \leq \frac{\varepsilon^2 C_K \beta_T (C_K \hat{\beta}_T)^{N-1}}{(1 + s)^{(d-2)/2}}, \qquad \textit{for } N \geq n \geq 1, \tag{4.54}$$

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

Proof. The proof is almost the same as that of Lemma 4.4. The only difference arises when we apply Lemma 4.7. Let $N \geq 1$ and suppose that the quantity in the first brackets in (4.43) is bounded by $c\varepsilon^3 C_K \beta_T (C\hat{\beta}_T)^{N-1}(1 + t)^{-d/2}$, where $\hat{\beta}_T = \beta_1 T^{-\mu}$ with $\mu \in (0, \alpha)$. Then, by Lemma 4.7 and (4.42) with β replaced by β_T, the right-hand side of (4.43) for $d \leq 4$ is bounded by

$$\frac{c\varepsilon^2 C_K \beta_T^2 (C_K \hat{\beta}_T)^{N-1}}{(1 + t)^{\frac{d}{2} \wedge (d-2) - \kappa}} = c\beta_1 T^{-bd}(1 + t)^{(4-d)/2 + \kappa} \frac{\varepsilon^2 C_K \beta_T (C_K \hat{\beta}_T)^{N-1}}{(1 + t)^{d/2}}, \tag{4.55}$$

where κ is an arbitrarily small but positive number if $d = 4$, otherwise $\kappa = 0$. Since $t \leq T \log T$ and $-bd + \frac{4-d}{2} = -\alpha < -\mu$, the factor in front of the fraction in the right-hand side is bounded by $C_K \hat{\beta}_T$ if C_K is sufficiently large, and thus we obtain (4.53) with $q = 0$.

For (4.53) with $q = 2$ and (4.54), we use Lemma 4.7 as in (4.55), with $(a, b) = (\frac{d}{2}, \frac{d-2}{2})$, to obtain the factor

$$\frac{\beta_T}{(1 + t)^{(d-2)/2 \wedge (d-3) - \kappa}} = \frac{\beta_1 T^{-bd}(1 + t)^{(4-d)/2 + \kappa}}{(1 + t)^{(d-2)/2}} \leq \frac{c\hat{\beta}_T}{(1 + t)^{(d-2)/2}}. \tag{4.56}$$

To prove (4.53) for $q = 4$, we apply Lemma 4.7 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

In this section, we prove that $\pi_t^{(N)}(x)$ can be approximated by

$$\pi_t^{(N)}(x \,|\, R) = \sum_{\vec{b}_N} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x}) \cap \{\mathbf{C}_{[0,t]} \subset \square_R\}\big), \tag{4.57}$$

where $\boldsymbol{x} = (x,t)$, $\square_R = [-R,R]^d \cap \mathbb{Z}^d$ and

$$\mathbf{C}_{[0,t]} = \bigcup_{s=0}^{t} \mathbf{C}_s(\boldsymbol{o}), \qquad\qquad \mathbf{C}_s(\boldsymbol{y}) = \{z \in \mathbb{Z}^d : \boldsymbol{y} \longrightarrow (z,s)\}. \tag{4.58}$$

We will also use the abbreviation $\mathbf{C}_s = \mathbf{C}_s(\boldsymbol{o})$. More precisely, we prove below that

$$\pi_t^{(N)}(x) = \pi_t^{(N)}(x \,|\, R) + o(1)\,\varepsilon^2, \tag{4.59}$$

where $o(1)$ is independent of ε and decays to zero as $R \to \infty$, by using the estimates for the diagram functions in Sections 4.1–4.3. This is a refined version of the finite containment argument used in proving the uniformity of (3.38), and will be useful in dealing with the continuum limit in Section 6.

Proof. First, we note that

$$\begin{aligned}0 \le \pi_t^{(N)}(x) - \pi_t^{(N)}(x \,|\, R) &\le \sum_{\vec{b}_N} \sum_{s=\varepsilon}^{t}{}^{\bullet}\, \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x}) \cap \{\mathbf{C}_{[0,s-\varepsilon]} \subset \square_R,\ \mathbf{C}_s \not\subset \square_R\}\big) \\ &\le \sum_{\vec{b}_N} \sum_{s=\varepsilon}^{t}{}^{\bullet} \sum_{\substack{u \in \square_R \\ v \notin \square_R}} \mathbb{P}\big(\tilde{E}_{\vec{b}_N}^{(N)}(\boldsymbol{x}) \cap \{\boldsymbol{o} \longrightarrow ((u,s-\varepsilon),(v,s))\}\big).\end{aligned} \tag{4.60}$$

Note that $((u,s-\varepsilon),(v,s))$ is a spatial bond, since $u \in \square_R$ and $v \notin \square_R$. The event in the rightmost expression is included in the union of two events: (i) $\boldsymbol{o} \longrightarrow ((u,s-\varepsilon),(v,s)) \longrightarrow \boldsymbol{x}$, or (ii) there is a vertex $(w,r) \in \square_R \times [0,s)$ such that $\boldsymbol{o} \longrightarrow (w,r) \longrightarrow \boldsymbol{x}$ and that $(w,r) \longrightarrow ((u,s-\varepsilon),(v,s))$ 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 $\sum_\ell \sum_s^\bullet \sum_{v \notin \square_R} P^{(N)}(\boldsymbol{x}, B^\ell(v,s))$, where $\sum_\ell$ is the sum over all admissible lines (i.e., the sum over $n = 1,\dots,N$ of the sum over the n^{th} admissible lines), and where we modified Construction $B^\ell_{\texttt{spat}}(v,s)$ by fixing the second endpoint (v,s), instead of fixing the first endpoint $(u,s-\varepsilon)$ as defined in (4.20). This bound, divided by ε^2, decays as $R \to \infty$ uniformly in ε, since the sum over $x \in \mathbb{Z}^d$ of the unrestricted sum $\sum_\ell \sum_s^\bullet P^{(N)}(\boldsymbol{x}, B^\ell(s))$ is bounded, by using Lemma 4.6, by $c\varepsilon^2 t\,(1+t)^{-d/2}$.

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

$$\sum_\ell \sum_{s=\varepsilon}^{t}{}^{\bullet} \sum_{r=0}^{s-\varepsilon}{}^{\bullet} \sum_{w \in \square_R} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \square_R} (\tau_{s-\varepsilon-r} * \lambda\varepsilon D)(v-w), \tag{4.61}$$

where we relabelled the second endpoint of the spatial bond in (4.21) as (w,r). For $w \in \square_{R/2}$, we use

$$\begin{aligned}&\sum_{w \in \square_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \square_R} (\tau_{s-\varepsilon-r} * D)(v-w) \\ &\quad \le P^{(N)}(\boldsymbol{x}, B^\ell(r)) \sup_{w \in \square_{R/2}} \sum_{v \notin \square_R} (\tau_{s-\varepsilon-r} * D)(v-w) \le P^{(N)}(\boldsymbol{x}, B^\ell(r)) \sum_{z \notin \square_{R/2}} (\tau_{s-\varepsilon-r} * D)(z),\end{aligned} \tag{4.62}$$

and for $w \in \square_R \setminus \square_{R/2}$, we use

$$\sum_{w \notin \square_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \sum_{v \notin \square_R} (\tau_{s-\varepsilon-r} * D)(v-w) \le \|\tau_{s-\varepsilon-r} * D\|_1 \sum_{w \notin \square_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)). \tag{4.63}$$

By Lemma 4.5, we have $\|\tau_{s-\varepsilon-r} * D\|_1 \le C_K$, so that $\sum_{z \notin \square_{R/2}} (\tau_{s-\varepsilon-r} * D)(z)$ decays to zero as $R \to \infty$, independently of ε. In addition, by Lemma 4.6, we have $\sum_\ell P^{(N)}(\boldsymbol{x}, B^\ell(r)) \le c\varepsilon^3(1+t)^{-d/2}$, so that $\sum_\ell \sum_{w \notin \square_{R/2}} P^{(N)}(\boldsymbol{x}, B^\ell(w,r)) \le o(1)\,\varepsilon^3(1+t)^{-d/2}$. Since there is another factor of ε in the summand of (4.61), while there are two summations over $\varepsilon\mathbb{Z}_+$ in (4.61), we conclude that (4.61) is $o(1)\,\varepsilon^2 t^2(1+t)^{-d/2}$. This completes the proof of (4.59). ∎

5 Inductive argument

In this section, we prove Proposition 2.1 by applying the inductive method of [19] 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 in (2.29). 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 a similar 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 (3.28) for the two-point function. Taking the Fourier transform with respect to the spatial component, we obtain (2.33), i.e.,

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

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

$$a(k) \equiv 1 - \hat{D}(k) \begin{cases} \asymp L^2|k|^2, & \text{if } \|k\|_\infty \leq L^{-1}, \\ > \eta, & \text{if } \|k\|_\infty > L^{-1}, \end{cases} \qquad a(k) < 2 - \eta \quad \forall k \in [-\pi,\pi]^d, \tag{5.2}$$

where $a \asymp b$ means that the ratio a/b is bounded away from zero and infinity. These assumptions correspond to Assumption D in [19].

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

$$f_n(k) = \hat{\tau}_{n\varepsilon}(k), \qquad e_n(k) = \hat{\pi}_{n\varepsilon}(k), \qquad g_{n+1}(k) = \hat{\pi}_{n\varepsilon}(k)\; \hat{p}_\varepsilon(k), \tag{5.3}$$

where the dependence on λ is left implicit, then (5.1) equals

$$f_{n+1}(k) = \sum_{m=0}^{n} g_{m+1}(k)\, f_{n-m}(k) + e_{n+1}(k), \tag{5.4}$$

with $f_0(k) = 1$. This is equivalent to the recursion relation (1.1) in [19]. The only difference is

$$f_1(k) = g_1(k) = \hat{p}_\varepsilon(k) = 1 - \varepsilon + \lambda\varepsilon\hat{D}(k), \tag{5.5}$$

whereas in [19], $f_1(k) = g_1(k) = \lambda\hat{D}(k)$. This change leads to a modification of the induction hypotheses in [19], the main reason being that we need to prove uniformity in ε. 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

$$-(2+\rho) < 0 < \tfrac{d}{2} - (2+\rho) < \gamma < \gamma + \delta < 1 \wedge \Delta \wedge \tfrac{d-4}{2}. \tag{5.6}$$

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

$$\lambda_n = 1 - \frac{1}{\varepsilon} \sum_{l=2}^{n} g_l(0; \lambda_{n-1}), \tag{5.7}$$

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

$$I_n = \lambda_n + \frac{K_1 \beta}{(1+n\varepsilon)^{(d-2)/2}} [-1, 1], \tag{5.8}$$

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

$$v_n = \frac{\lambda - \frac{1}{\sigma^2 \varepsilon} \sum_{l=2}^{n} \nabla^2 g_l(0)}{1 + \sum_{l=2}^{n} (l-1)\, g_l(0)}. \tag{5.9}$$

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

$$K_3 \gg K_1 \gg K_4 \gg 1, \qquad K_2, K_5 \gg K_4. \tag{5.10}$$

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

(H1)–(H2)

$$|\lambda_m - \lambda_{m-1}| \leq \frac{\varepsilon K_1 \beta}{(1+m\varepsilon)^{d/2}}, \qquad |v_m - v_{m-1}| \leq \frac{\varepsilon K_2 \beta}{(1+m\varepsilon)^{(d-2)/2}}, \tag{5.11}$$

(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

$$f_m(k) = \prod_{l=1}^{m} [1 - \varepsilon v_l\, a(k) + \varepsilon\, r_l(k)], \tag{5.12}$$

where $r_l(k)$ obeys the bounds

$$|r_l(0)| \leq \frac{K_3 \beta}{(1+l\varepsilon)^{(d-2)/2}}, \qquad |r_l(k) - r_l(0)| \leq \frac{K_3 \beta\, a(k)}{(1+l\varepsilon)^{\delta}}. \tag{5.13}$$

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

$$|f_m(k)| \leq \frac{K_4\, a(k)^{-2-\rho}}{(1+m\varepsilon)^{d/2}}, \qquad |f_m(k) - f_{m-1}(k)| \leq \frac{\varepsilon K_5\, a(k)^{-1-\rho}}{(1+m\varepsilon)^{d/2}}. \tag{5.14}$$

Instead of (5.12), we can alternatively write $f_m(k)$ as

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

where

$$f_m(0) = \prod_{l=1}^{m} [1 + \varepsilon\, r_l(0)], \qquad s_l(k) = \frac{\varepsilon v_l\, r_l(0)\, a(k) + [r_l(k) - r_l(0)]}{1 + \varepsilon\, r_l(0)}. \tag{5.16}$$

The induction hypothesis (H3) implies

$$|s_l(k)| \leq \frac{\varepsilon v_l |r_l(0)| a(k) + |r_l(k) - r_l(0)|}{1 - \varepsilon |r_l(0)|} \leq \frac{(1+\varepsilon v_l) K_3 \beta\, a(k)}{(1-\varepsilon K_3 \beta)(1+l\varepsilon)^{\delta}}. \tag{5.17}$$

In some cases, we will use (5.15)–(5.17), instead of (5.12)–(5.13). Moreover, by (5.15) and spatial symmetry, we obtain

$$\nabla^2 f_m(0) = f_m(0)\, \varepsilon \sum_{l=1}^{m} \big[-v_l \sigma^2 + \nabla^2 s_l(0)\big]. \tag{5.18}$$

The advancement of the induction hypotheses is a small modification of that in [19], 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 [19, 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 $m = 1, \ldots, n$. Then, $I_0 \supset I_1 \supset \cdots \supset I_n$.*

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

$$|f_m(k)| \le e^{cK_3\beta} e^{-m\varepsilon[1-c(K_1+K_2+K_3)\beta]\,a(k)}. \tag{5.19}$$

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

$$|\nabla^2 f_m(0)| \le [1 + c\,(K_1 + K_2 + K_3)\beta]\sigma^2 m\varepsilon. \tag{5.20}$$

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

$$\|\hat{D}^2 f_m\|_1 \le \frac{c\,(1 + K_4)\beta}{(1 + m\varepsilon)^{d/2}}. \tag{5.21}$$

The bounds (2.29) for $s \le n\varepsilon$ follow from Lemmas 5.2–5.4, if $K \gg K_4$. The proofs of Lemmas 5.1–5.4 are almost identical to those of [19, Lemmas 2.1–2.4], and are deferred to Appendix A.

By Lemma 5.1, if $\lambda \in I_m$ for some $m \ge 0$, then $\lambda \in I_0$ and hence, by (5.8),

$$|\lambda - 1| \le K_1\beta. \tag{5.22}$$

It also follows that $I_\infty = \bigcap_{m=0}^\infty I_m$ is a singleton λ_∞. As discussed in [19, Theorem 1.2], we obtain $\lambda_\infty = \lambda_c^{(\varepsilon)}$. Moreover, it follows from the second inequality of (5.11) that, for $\lambda \in I_m$,

$$|v_m - 1| \le \sum_{l=1}^m |v_l - v_{l-1}| + |v_0 - 1| \le \sum_{l=1}^m \frac{\varepsilon K_2\beta}{(1 + l\varepsilon)^{(d-2)/2}} + |\lambda - 1| \le (cK_2 + K_1)\beta. \tag{5.23}$$

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 (5.11)–(5.14) are incorporated for the same reason.

5.3 Proof of Proposition 2.1

Fix $\lambda = \lambda_c^{(\varepsilon)}$, 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 $\lambda_c = \lambda_c^{(\varepsilon)}$, $A = A^{(\varepsilon)}$ and $v = v^{(\varepsilon)}$.

Note that, by (5.11)–(5.13), we have that, for $n < m$,

$$|f_n(0) - f_m(0)| = \prod_{l=1}^n [1 + \varepsilon r_l(0)] \left|1 - \prod_{l=n+1}^m [1 + \varepsilon r_l(0)]\right| \le \frac{cK_3\beta}{(1 + n\varepsilon)^{(d-4)/2}}, \tag{5.24}$$

$$|v_n - v_m| \le \sum_{l=n+1}^m |v_l - v_{l-1}| \le \frac{cK_2\beta}{(1 + n\varepsilon)^{(d-4)/2}}, \tag{5.25}$$

so that $\{f_n(0)\}_{n=1}^\infty$ and $\{v_n\}_{n=1}^\infty$ are Cauchy sequences. Therefore, the limits $A = \lim_{n\to\infty} f_n(0)$ and $v = \lim_{n\to\infty} v_n$ exist, and satisfy

$$|f_n(0) - A| \le \frac{cK_3\beta}{(1 + n\varepsilon)^{(d-4)/2}}, \qquad |v_n - v| \le \frac{cK_2\beta}{(1 + n\varepsilon)^{(d-4)/2}}. \tag{5.26}$$

In particular, by Lemma 5.2 and (5.23), both A and v are equal to $1+O(\beta)$.

Let $t = n\varepsilon$ and $\tilde{k} = \frac{k}{\sqrt{v\sigma^2 t}} \in \mathcal{A}_n$. By (1.2), $a(\tilde{k}) = \frac{|k|^2}{2dvt} + O(|k|^{2+2\Delta} t^{-1-\Delta})$. Using (5.15), (5.17), (5.26) and $\delta < 1 \wedge \frac{d-4}{2}$, we obtain

$$
\begin{aligned}
f_n(\tilde{k}) &= \Big[A + \frac{O(\beta)}{(1+t)^{(d-4)/2}}\Big] \prod_{l=1}^{n} \Big[1 - \varepsilon\Big[v + \frac{O(\beta)}{(1+l\varepsilon)^{\delta}}\Big] a(\tilde{k})\Big] \\
&= \Big[A + \frac{O(\beta)}{(1+t)^{(d-4)/2}}\Big] \Big[1 - \frac{vt\, a(\tilde{k})}{n}\Big]^n \prod_{l=1}^{n} \Big[1 - \frac{\varepsilon O(\beta)\, a(\tilde{k})}{(1+l\varepsilon)^{\delta}}\Big] \\
&= \Big[A + \frac{O(\beta)}{(1+t)^{(d-4)/2}}\Big] e^{-\frac{|k|^2}{2d} + O(|k|^{2+2\Delta} t^{-\Delta}) + O(\varepsilon |k|^4 t^{-1})} \Big[1 + \varepsilon \sum_{l=1}^{t/\varepsilon} \frac{O(\beta |k|^2 t^{-1})}{(1+l\varepsilon)^{\delta}}\Big] \\
&= A e^{-\frac{|k|^2}{2d}} \Big[1 + \frac{O(\beta)}{(1+t)^{(d-4)/2}} + O(|k|^{2+2\Delta} t^{-\Delta}) + O(\varepsilon |k|^4 t^{-1}) + \frac{O(\beta |k|^2)}{(1+t)^{\delta}}\Big],
\end{aligned}
\tag{5.27}
$$

where the last error term follows from

$$
\varepsilon \sum_{l=1}^{t/\varepsilon} \frac{O(t^{-1})}{(1+l\varepsilon)^{\delta}} = O(t^{-1}) \big[(1+t)^{1-\delta} - 1\big] = O((1+t)^{-\delta}), \tag{5.28}
$$

for $\delta < 1$. Using $\frac{|k|^2}{t} \le c \frac{\log(2+t)}{1+t}$ for small $\tilde{k} \in \mathcal{A}_n$ and $\delta < 1 \wedge \Delta$, we have

$$
O(|k|^{2+2\Delta} t^{-\Delta}) \le O(|k|^2) \Big[\frac{\log(2+t)}{1+t}\Big]^{\Delta} \le \frac{O(|k|^2)}{(1+t)^{\delta}}, \tag{5.29}
$$

$$
O(\varepsilon |k|^4 t^{-1}) \le O(\varepsilon |k|^2) \frac{\log(2+t)}{1+t} \le \frac{O(\varepsilon |k|^2)}{(1+t)^{\delta}}. \tag{5.30}
$$

By (5.27)–(5.30), we obtain (2.4).

Let $e_1, \dots, e_d$ denote the standard basis vectors in $\mathbb{R}^d$. Then, by (5.17),

$$
|\nabla^2 s_l(0)| = \Big|\sum_{i=1}^{d} \lim_{h\to 0} \frac{s_l(he_i) - s_l(0)}{h^2}\Big| \le \frac{cK_3 \beta}{(1+l\varepsilon)^{\delta}} \Big|\sum_{i=1}^{d} \lim_{h\to 0} \frac{a(he_i)}{h^2}\Big| = \frac{cK_3 \sigma^2 \beta}{(1+l\varepsilon)^{\delta}}. \tag{5.31}
$$

Since $\delta < 1 \wedge \frac{d-4}{2}$, it follows from (5.18), (5.26), (5.28) and (5.31) that

$$
-\frac{\nabla^2 f_n(0)}{f_n(0)} = v\sigma^2 t \,[1 + O(\beta)\,(1+t)^{-\delta}], \tag{5.32}
$$

which is (2.5).

The upper bound in (2.6) is an immediate consequence of (4.40). For the lower bound, we consider the case of $t \ge 1$ and the case of $t < 1$, separately. When $t \ge 1$, we follow the proof of [19, Corollary 1.4] for oriented percolation. In this case, we use (2.4) and obtain the lower bound of the form $cL^{-d} t^{-d/2}$. When $t < 1$, we use the trivial inequality

$$
\|\tau_t\|_\infty \ge \|p_\varepsilon\|_\infty^{t/\varepsilon} \ge [(1-\varepsilon) \vee (\lambda \varepsilon \|D\|_\infty)]^{t/\varepsilon}, \tag{5.33}
$$

which can be bounded from below by an ε-independent multiple of $L^{-d}(1+t)^{-d/2}$. This completes the proof of Proposition 2.1. □

Finally, we derive the expressions (2.34)–(2.35) for λ_{c}, A and v. Recall (5.3). The expressions for λ_{c} and v immediately follow from (5.7), (5.9) and the fact that $\lambda_{\mathrm{c}} = \lambda_\infty$. To derive the expression for

$A = \lim_{n\to\infty} f_n(0)$, we follow the same strategy as in [19, 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 (5.4) with $k = 0$, multiplied by ε, over $n = 0, \dots, N-1$, 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

$$\begin{aligned} F_N &= g_1(0)\, F_{N-1} + \sum_{n=2}^N g_n(0)\, F_{N-n} + \varepsilon\sum_{n=2}^N e_n(0) + \varepsilon \\ &= F_{N-1} - \sum_{n=2}^\infty g_n(0)\, F_{N-1} + \sum_{n=2}^N g_n(0)\, F_{N-n} + \varepsilon\sum_{n=2}^N e_n(0) + \varepsilon. \end{aligned} \tag{5.34}$$

Taking the limit $N \to \infty$ of $f_N(0) = \frac{F_N - F_{N-1}}{\varepsilon}$, with the help of the bound (2.30), we obtain

$$A = -A\sum_{n=2}^\infty (n-1)g_n(0) + \sum_{n=2}^\infty e_n(0) + 1. \tag{5.35}$$

which, with the help of (5.3), gives the expression for A in (2.35). This completes the derivation of (2.34)–(2.35). □

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 $\alpha = bd - \frac{4-d}{2} > 0$, and let $\omega \in (\delta, 1 \wedge \alpha)$ and $\hat\beta_T = \beta_1 T^{-\mu}$ with $\mu \in (0, \alpha - \omega)$. We again define $\lambda_n = \lambda_n(T)$ and $v_n = v_n(\lambda)$ by (5.7) and (5.9), respectively, where we emphasize the dependence on T of λ_n. However, we replace (5.6), (5.8), (5.11) and (5.13), respectively, by

$$-(2+\rho) < 0 < \tfrac{d}{2} - (2+\rho) < \gamma < \gamma + \delta < \omega \wedge \Delta, \tag{5.36}$$

$$I_n = \lambda_n + \frac{K_1\hat\beta_T}{(1+n\varepsilon)^{1+\omega}}[-1, 1], \tag{5.37}$$

$$|\lambda_m - \lambda_{m-1}| \le \frac{\varepsilon K_1 \hat\beta_T}{(1+m\varepsilon)^{2+\omega}}, \qquad |v_m - v_{m-1}| \le \frac{\varepsilon K_2 \hat\beta_T}{(1+m\varepsilon)^{1+\omega}}, \tag{5.38}$$

$$|r_m(0)| \le \frac{K_3\hat\beta_T}{(1+m\varepsilon)^{1+\omega}}, \qquad |r_m(k) - r_m(0)| \le \frac{K_3\hat\beta_T}{(1+m\varepsilon)^{\delta}}\, a(k). \tag{5.39}$$

The induction hypotheses are that (H1)–(H4) hold for all $\lambda \in I_n$ and $m = 1, \dots, n$, 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 $n \le \frac{T}{\varepsilon}\log T$. This implies that we can use the bounds following from (H1)–(H4) in the advancement of the induction hypotheses. 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 (5.8) 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 $d \le 4$ (cf., the proof of Lemma 5.1 in Appendix A). To satisfy this requirement, we transfer some power exponent of β_T as

$$\frac{\beta_T}{(1+n\varepsilon)^{(d-2)/2}} = \frac{\beta_1 T^{-db}(1+n\varepsilon)^{(4-d)/2+\omega}}{(1+n\varepsilon)^{1+\omega}} \le \frac{c\beta_1 T^{-\mu}}{(1+n\varepsilon)^{1+\omega}}, \tag{5.40}$$

for $T \ge 1$, where we use $n\varepsilon \le T\log T$ and $-bd + \frac{4-d}{2} + \omega = -(\alpha - \omega) < -\mu$. This is the motivation of the changes in (5.37)–(5.39).

By the above changes, (5.22)–(5.23) 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 $m = 1, \dots, n$ 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 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 Proposition 2.2. Let $n \le \frac{T}{\varepsilon} \log T$ and $\lambda \in I_n$. In particular, the following results hold at $\lambda = \lambda_T$, which is defined as

$$\lambda_T = \lambda_{\frac{T}{\varepsilon}\log T}(T) = 1 - \frac{1}{\varepsilon} \sum_{l=2}^{\frac{T}{\varepsilon}\log T} g_l(0; \lambda_{\frac{T}{\varepsilon}\log T - 1}(T)). \tag{5.41}$$

By (5.39), we can bound $|f_n(0) - 1|$ by

$$\left| \prod_{m=1}^n [1 + \varepsilon r_m(0)] - 1 \right| \le \varepsilon \sum_{m=1}^n |r_m(0)| \prod_{l=m+1}^n [1 + \varepsilon |r_l(0)|] \le \varepsilon \sum_{m=1}^n \frac{K_3 \hat{\beta}_T \, e^{\varepsilon \sum_{l=m+1}^n |r_l(0)|}}{(1 + m\varepsilon)^{1+\omega}} \le c K_3 \hat{\beta}_T, \tag{5.42}$$

which proves that the asymptotic expected number of infected individuals is 1. Also, using (5.23) and (5.38), we have $v_n = 1 + O(\hat{\beta}_T)$, which means that the asymptotic diffusion constant is 1.

Let $n = Tt/\varepsilon$ with $t \le \log T$ and $\tilde{k} = \frac{k}{\sqrt{\sigma_T^2 Tt}} \in \mathcal{A}_n$. By (1.2), $a(\tilde{k}) = \frac{|k|^2}{2dTt} + O(|k|^{2+2\Delta}(Tt)^{-1-\Delta})$. Therefore,

$$\begin{aligned} f_n(\tilde{k}) &= [1 + O(\hat{\beta}_T)] \left[1 - [1 + O(\hat{\beta}_T)] \frac{Tt\, a(\tilde{k})}{n} \right]^n \\ &= e^{-\frac{|k|^2}{2d}} \left[1 + O(\hat{\beta}_T) + O(|k|^{2+2\Delta}(Tt)^{-\Delta}) + O(\varepsilon |k|^4 (Tt)^{-1}) \right]. \end{aligned} \tag{5.43}$$

By (5.29)–(5.30) for small $\tilde{k}$, the last two error terms can be replaced by $O(|k|^2(1+Tt)^{-\delta})$. This proves (2.7).

Using (5.18) and (5.31) as well as $v_n = 1 + O(\hat{\beta}_T)$, we obtain (2.8). The proof of (2.9) does not depend on d, and is the same as in Section 5.3. This completes the proof of Proposition 2.2. □

6 Continuum limit

In this section, we compute the limit of the lace expansion coefficients as $\varepsilon \downarrow 0$, and prove Proposition 2.6.

We prove below convergence of $\frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{\lambda}(x)$ for $t/\varepsilon \in [2, \infty) \cap \mathbb{Z}_+$ with a *fixed* $\lambda \le \lambda_c$, and then extend this to $\frac{1}{\varepsilon^2} \partial_\lambda \pi_{t;\varepsilon}^{\lambda}(x)$. The proof of the continuity in λ of $\partial_\lambda \pi_t^{\lambda}(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon^2} \partial_\lambda \pi_{t;\varepsilon}^{\lambda}(x)$ is more or less immediate from its finite containment property that is similar to the one for the discretized contact process in Section 4.4, and this will be discussed briefly at the end of this section. These statements imply convergence of $\frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x)$ whenever $\lambda^{(\varepsilon)} \to \lambda$ such that $\lambda^{(\varepsilon)} \le \lambda_c^{(\varepsilon)}$ for ε sufficiently small. Indeed, for any $\lambda_0 < \lambda \le \lambda_c$, we can write

$$\left| \frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x) - \pi_t^{\lambda_0}(x) \right| \le \frac{1}{\varepsilon^2} \left| \pi_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x) - \pi_{t;\varepsilon}^{\lambda_0}(x) \right| + \left| \frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{\lambda_0}(x) - \pi_t^{\lambda_0}(x) \right|. \tag{6.1}$$

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

$$\frac{1}{\varepsilon^2} \left| \pi_{t;\varepsilon}^{\lambda^{(\varepsilon)}}(x) - \pi_{t;\varepsilon}^{\lambda_0}(x) \right| \le \int_{\lambda_0}^{\lambda^{(\varepsilon)}} d\lambda' \; \frac{1}{\varepsilon^2} \left| \partial_{\lambda'} \pi_{t;\varepsilon}^{\lambda'}(x) \right|, \tag{6.2}$$

where we use $\lambda_0 \le \lambda_c^{(\varepsilon)}$ for sufficiently small ε, which is due to the fact that $\lambda_c^{(\varepsilon)}$ converges to $\lambda_c > \lambda_0$. Since the integrand is uniformly bounded (even when we sum over x) by $C\beta(1+t)^{-(d-2)/2}$, the limsup of the integral when $\varepsilon \downarrow 0$ is bounded by a multiple of $\lambda - \lambda_0$. By taking the limit $\lambda_0 \uparrow \lambda$ and using the fact that $\partial_\lambda \pi_t^\lambda(x)$ is continuous in λ, the first expression in (2.36) follows. Therefore, we are left to prove convergence of $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^\lambda(x)$ and $\frac{1}{\varepsilon^2}\partial_\lambda \pi_{t;\varepsilon}^\lambda(x)$ for every $\lambda \le \lambda_c$ and the continuity in λ of $\partial_\lambda \pi_t^\lambda(x)$.

We prove below that, for every $N \ge 0$, $\lambda \le \lambda_c$ and (x,t), there is a $\pi_t^{(N)}(x)$ such that

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

where we suppress the dependence on λ. That is, we will deal with *pointwise convergence*, rather than the uniform bounds in Section 4 which are valid for all (x,t) and $\varepsilon \le 1$, and hence all terms which are $o(1)$ as $\varepsilon \downarrow 0$ will be estimated away. By this pointwise convergence, together with the uniform bounds in Section 4 and the dominated convergence theorem, we have

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^\lambda(x) = \lim_{\varepsilon \downarrow 0} \sum_{N=0}^{\infty} (-1)^N \frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{(N)}(x) = \sum_{N=0}^{\infty} (-1)^N \pi_t^{(N)}(x) = \pi_t^\lambda(x). \tag{6.4}$$

This completes the proof of the pointwise convergence of $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^\lambda(x)$. The proof of the convergence of $\frac{1}{\varepsilon^2}\partial_\lambda \pi_{t;\varepsilon}^\lambda(x)$ is similar, and we will only discuss the necessary changes.

The proof of (6.3) is divided into several steps.

Statement of the induction hypothesis. Given a site set $\mathcal{C} \subset \mathbb{Z}^d$ (which may be an empty set), we define

$$\pi_{t;\varepsilon}^{(N)}(x;\mathcal{C}) = \sum_{\vec{b}_N} \mathbb{P}_\varepsilon^\lambda\big(\tilde{E}_{\vec{b}_N}^{(N)}(x,t) \cap \big\{\mathbf{C}_t(\overline{b}_N) \setminus \{x\} = \mathcal{C}\big\}\big), \tag{6.5}$$

where we recall $\tilde{E}_{\vec{b}_0}^{(0)}(x,t) = \{(o,0) \Longrightarrow (x,t)\}$, $\overline{b}_0 = (o,0)$ and the notation (4.58) for $\mathbf{C}_t(\overline{b}_N)$. We will use induction in N to prove that, for every $t > 0$, there is a $\pi_t^{(N)}(x;\mathcal{C})$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon^2} \pi_{t;\varepsilon}^{(N)}(x;\mathcal{C}) = \pi_t^{(N)}(x;\mathcal{C}). \tag{6.6}$$

The claim for $\pi_{t;\varepsilon}^{(N)}(x)$ in (6.3) then follows by summing over $\mathcal{C} \subset \mathbb{Z}^d$, together with the fact that the main contribution comes from $\mathcal{C} \subset \Box_R$ by the finite containment property in Section 4.4.

Initialization of the induction. First, we investigate $N = 0$. For $\mathbf{S}_1, \mathbf{S}_2, \mathbf{A} \subset \mathbb{Z}^d \times \varepsilon\mathbb{Z}_+$, we denote

$$\{\mathbf{S}_1 \longrightarrow \mathbf{S}_2\} = \bigcup_{\substack{\boldsymbol{s}_1 \in \mathbf{S}_1 \\ \boldsymbol{s}_2 \in \mathbf{S}_2}} \{\boldsymbol{s}_1 \longrightarrow \boldsymbol{s}_2\}, \qquad \{\mathbf{S}_1 \Longrightarrow \mathbf{S}_2\} = \bigcup_{\substack{\boldsymbol{s}_1, \boldsymbol{s}_1' \in \mathbf{S}_1 \\ \boldsymbol{s}_2, \boldsymbol{s}_2' \in \mathbf{S}_2}} \{\boldsymbol{s}_1 \longrightarrow \boldsymbol{s}_2\} \circ \{\boldsymbol{s}_1' \longrightarrow \boldsymbol{s}_2'\}, \tag{6.7}$$

and define

$$\mathbf{C}_t(\mathbf{A}) = \{x \in \mathbb{Z}^d : \mathbf{A} \longrightarrow (x,t)\} = \bigcup_{\boldsymbol{a} \in \mathbf{A}} \mathbf{C}_t(\boldsymbol{a}), \qquad \mathbf{C}(\mathbf{A}) = \bigcup_{t \ge 0} \mathbf{C}_t(\mathbf{A}). \tag{6.8}$$

Using the Markov property at time ε, we arrive at

$$\pi_{t;\varepsilon}^{(0)}(x;\mathcal{C}) = \sum_{\mathcal{A} \subset \mathbb{Z}^d : |\mathcal{A}| \ge 2} \bigg[\prod_{a \in \mathcal{A}} p_\varepsilon(a)\bigg]\bigg[\prod_{a \notin \mathcal{A}} [1 - p_\varepsilon(a)]\bigg] \times \mathbb{P}_\varepsilon^\lambda\left(\begin{matrix} \exists\, a, a'(\neq a) \in \mathcal{A} : \{(a,\varepsilon) \longrightarrow (x,t)\} \circ \{(a',\varepsilon) \longrightarrow (x,t)\} \\ \mathbf{C}_t(\mathcal{A} \times \{\varepsilon\}) \setminus \{x\} = \mathcal{C} \end{matrix}\right), \tag{6.9}$$

Since every $p_\varepsilon(a)$ for $a \neq o$ gives rise to a factor of ε, we immediately see that the main contribution comes from $\mathcal{A} = \{o, y\}$ for some $y \neq o$. Therefore, we obtain

$$\begin{aligned}
\frac{1}{\varepsilon}\pi^{(0)}_{t;\varepsilon}(x;\mathcal{C}) &= \sum_{y\in\mathbb{Z}^d} \lambda D(y)\, \mathbb{P}^\lambda_\varepsilon \begin{pmatrix} \{(o,\varepsilon) \longrightarrow (x,t)\} \circ \{(y,\varepsilon) \longrightarrow (x,t)\} \\ \mathbf{C}_t(\{(o,\varepsilon),(y,\varepsilon)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix} + o(1) \\
&= \sum_{y\in\mathbb{Z}^d} \lambda D(y)\, \mathbb{P}^\lambda_\varepsilon \begin{pmatrix} \{(o,0) \longrightarrow (x,t)\} \circ \{(y,0) \longrightarrow (x,t)\} \\ \mathbf{C}_t(\{(o,0),(y,0)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix} + o(1) \\
&= \sum_{y\in\mathbb{Z}^d} \lambda D(y)\, \mathbb{P}^\lambda_\varepsilon \begin{pmatrix} \{(o,0),(y,0)\} \Longrightarrow (x,t) \\ \mathbf{C}_t(\{(o,0),(y,0)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix} + o(1), \qquad (6.10)
\end{aligned}$$

where the second equality is due to the fact that $((o,0),(o,\varepsilon))$ or $((y,0),(y,\varepsilon))$ is vacant (with probability $(2-\varepsilon)\varepsilon$) in the symmetric difference between the events on both sides of the equality, and the third equality is due to the fact that the double connection from (o,ε) or (y,ε) gives rise to an extra factor of ε.

We repeat the same observation around (x,t) and obtain that, for every $y \in \mathbb{Z}^d$ and $\mathcal{C} \subset \mathbb{Z}^d$,

$$\begin{aligned}
&\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon \begin{pmatrix} \{(o,0),(y,0)\} \Longrightarrow (x,t) \\ \mathbf{C}_t(\{(o,0),(y,0)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix} \\
&\quad = \sum_{z\in\mathbb{Z}^d} \lambda D(x-z)\, \mathbb{P}^\lambda_\varepsilon \begin{pmatrix} \{(o,0),(y,0)\} \Longrightarrow \{(x,t),(z,t)\} \\ \mathbf{C}_t(\{(o,0),(y,0)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix} + o(1). \qquad (6.11)
\end{aligned}$$

Substituting (6.11) into (6.10) and using the weak convergence of $\mathbb{P}^\lambda_\varepsilon$ towards $\mathbb{P}^\lambda$ as formulated in [4, Proposition 2.7], we obtain

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon^2}\pi^{(0)}_{t;\varepsilon}(x;\mathcal{C}) = \sum_{y,z\in\mathbb{Z}^d} \lambda^2 D(y)\, D(x-z)\, \mathbb{P}^\lambda \begin{pmatrix} \{(o,0),(y,0)\} \Longrightarrow \{(x,t),(z,t)\} \\ \mathbf{C}_t(\{(o,0),(y,0)\}) \setminus \{x\} = \mathcal{C} \end{pmatrix}. \qquad (6.12)$$

Here and in the rest of this section, we use "$\longrightarrow$" and "$\Longrightarrow$" to denote connections in $\mathbb{Z}^d \times \mathbb{R}_+$, via the graphical representation in Section 2.1. This completes the proof of (6.6) for $N = 0$, with $\pi^{(0)}_t(x;\mathcal{C})$ for $t > 0$ defined to be the right-hand side of (6.12).

Preliminaries for the advancement. To advance the induction hypothesis, we first note that, by using the finite containment property of Section 4.4 and the Markov property at the time component of $\underline{b}_N$, we have

$$\begin{aligned}
\frac{1}{\varepsilon^2}\pi^{(N)}_{t;\varepsilon}(x;\mathcal{C}) &= \frac{1}{\varepsilon^2}\sum_{\vec{b}_N} \mathbb{P}^\lambda_\varepsilon\Big(\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(\underline{b}_N) \cap E(b_N,(x,t);\tilde{\mathbf{C}}^{b_N}(\overline{b}_{N-1})) \cap \big\{\mathbf{C}_t(\overline{b}_N) \setminus \{x\} = \mathcal{C}\big\}\Big) \\
&= \frac{1}{\varepsilon^2}\sum_{s=\varepsilon}^{t}{}^{\bullet} \sum_{y\in\mathbb{Z}^d} \sum_{\vec{b}_N:\underline{b}_N=(y,s-\varepsilon)} \sum_{\mathcal{C}'\subset\square_R:\mathcal{C}'\ni y} \mathbb{P}^\lambda_\varepsilon\big(\tilde{E}^{(N-1)}_{\vec{b}_{N-1}}(y,s-\varepsilon) \cap \{\mathbf{C}_{s-\varepsilon}(\overline{b}_{N-1}) = \mathcal{C}'\}\big) \\
&\qquad \times \mathbb{P}^\lambda_\varepsilon\Big(E\big(b_N,(x,t);\tilde{\mathbf{C}}^{b_N}(\mathcal{C}'\times\{s-\varepsilon\})\big) \cap \big\{\mathbf{C}_t(\overline{b}_N) \setminus \{x\} = \mathcal{C}\big\}\Big) + o(1), \qquad (6.13)
\end{aligned}$$

where, similarly to (6.8), we write $\tilde{\mathbf{C}}^{b_N}(\mathcal{C}'\times\{s-\varepsilon\}) = \bigcup_{v\in\mathcal{C}'}\tilde{\mathbf{C}}^{b_N}(v,s-\varepsilon)$, and $o(1)$ is independent of ε and decays to zero as $R \to \infty$. We now investigate the second probability in (6.13) when $\mathcal{C}' = \{y\}$ and when $\mathcal{C}' \supsetneq \{y\}$, separately.

When $\mathcal{C}' = \{y\}$, we recall the definitions (3.13)–(3.14) and use the Markov property at time s, similarly to the discussion around (6.9)–(6.10), to obtain that

$$\sum_{b_N : \underline{b}_N = (y, s-\varepsilon)} \frac{1}{\varepsilon} \mathbb{P}_\varepsilon^\lambda \Big(E(b_N, (x,t); \tilde{\mathbf{C}}^{b_N}(y, s-\varepsilon)) \cap \big\{ \mathbf{C}_t(\overline{b}_N) \setminus \{x\} = \mathcal{C} \big\} \Big)$$
$$= \sum_{y' \in \mathbb{Z}^d \setminus \{y\}} \lambda D(y'-y) \Big[\mathbb{P}_\varepsilon^\lambda \Big(E'\big((y,s),(x,t); \mathbf{C}(y',s)\big) \cap \big\{ \mathbf{C}_t(y,s) \setminus \{x\} = \mathcal{C} \big\} \Big)$$
$$+ \mathbb{P}_\varepsilon^\lambda \Big(E'\big((y',s),(x,t); \mathbf{C}(y,s)\big) \cap \big\{ \mathbf{C}_t(y',s) \setminus \{x\} = \mathcal{C} \big\} \Big) \Big] + o(1), \tag{6.14}$$

where $o(1)$ decays to zero as $\varepsilon \downarrow 0$, and the first probability in the brackets is the contribution from the case in which b_N is the temporal bond $((y, s-\varepsilon),(y,s))$, while the second probability is the contribution from the case in which b_N is the spatial bond $((y,s-\varepsilon),(y',s))$. In (6.14), we also use the fact that, with probability $1-o(1)$, $\tilde{\mathbf{C}}^{b_N}(y,s-\varepsilon) \cap (\mathbb{Z}^d \times [s,\infty))$ equals $\mathbf{C}(y',s)$ when $b_N = ((y,s-\varepsilon),(y,s))$, and equals $\mathbf{C}(y,s)$ when $b_N = ((y,s-\varepsilon),(y',s))$.

When $\mathcal{C}' \supsetneq \{y\}$, we again use the Markov property at time s, and then we use the fact that, with probability $1-o(1)$, every temporal bond growing from each site in $\mathcal{C}' \times \{s-\varepsilon\}$ is occupied, and all the spatial bonds growing from the sites in $\mathcal{C}' \times \{s-\varepsilon\}$ are vacant. Therefore, with probability $1-o(1)$, the subset of $\tilde{\mathbf{C}}^{b_N}(\mathcal{C}' \times \{s-\varepsilon\})$ after time s equals $\mathbf{C}\big((\mathcal{C}' \setminus \{y\}) \times \{s\}\big)$, and we have

$$\sum_{b_N : \underline{b}_N = (y, s-\varepsilon)} \mathbb{P}_\varepsilon^\lambda \Big(E\big(b_N, (x,t); \tilde{\mathbf{C}}^{b_N}(\mathcal{C}' \times \{s-\varepsilon\})\big) \cap \big\{ \mathbf{C}_t(\overline{b}_N) \setminus \{x\} = \mathcal{C} \big\} \Big)$$
$$= \mathbb{P}_\varepsilon^\lambda \Big(E'\big((y,s),(x,t); \mathbf{C}\big((\mathcal{C}' \setminus \{y\}) \times \{s\}\big)\big) \cap \big\{ \mathbf{C}_t(y,s) \setminus \{x\} = \mathcal{C} \big\} \Big) + o(1). \tag{6.15}$$

To deal with the event $E'((y,s),(x,t); \mathcal{A} \times \{s\})$ for $\mathcal{A} \subset \mathbb{Z}^d \setminus \{y\}$ in (6.14)–(6.15), we introduce some notation. We define the set of sites that are connected from (y,s) via a path which does not go through $\boldsymbol{v}$ by

$$\tilde{\mathbf{C}}^{\boldsymbol{v}}(y,s) = \bigcap_{b=(\,\cdot\,,\boldsymbol{v})} \tilde{\mathbf{C}}^{b}(y,s). \tag{6.16}$$

We also define

$$\mathcal{E}_{s,t}(y,x;\mathcal{A}) = \bigcup_{\boldsymbol{v}} \Big\{ \big\{ \boldsymbol{v} \notin \mathbf{C}(\mathcal{A} \times \{s\}) \big\} \cap \big\{ (y,s) \longrightarrow \boldsymbol{v} \Longrightarrow (x,t) \in \mathbf{C}(\mathcal{A} \times \{s\}) \setminus \tilde{\mathbf{C}}^{\boldsymbol{v}}(y,s) \big\} \Big\}, \tag{6.17}$$
$$\mathcal{R}_{s,t}(y,x;\mathcal{A}) = \bigcup_{\boldsymbol{v}} \Big\{ \big\{ \mathcal{A} \times \{s\} \longrightarrow \boldsymbol{v} \big\} \circ \big\{ (y,s) \longrightarrow \boldsymbol{v} \Longrightarrow (x,t) \big\} \Big\}. \tag{6.18}$$

By this notation, it is not hard to see that $E'((y,s),(x,t); \mathcal{A} \times \{s\})$ is rewritten as

$$E'((y,s),(x,t); \mathcal{A} \times \{s\}) = \mathcal{E}_{s,t}(y,x;\mathcal{A}) \;\dot\cup\; \mathcal{R}_{s,t}(y,x;\mathcal{A}). \tag{6.19}$$

The contribution from $\mathcal{R}_{s,t}(y,x;\mathcal{A})$ has an extra factor of ε, due to the fact that there are at least *two* spatial bonds at $\boldsymbol{v}$ (one before and one after $\boldsymbol{v}$), which leads to an error term as $\varepsilon \downarrow 0$. Therefore, we only need to focus on the contribution from $\mathcal{E}_{s,t}(y,x;\mathcal{A})$, i.e.,

$$\mathbb{P}_\varepsilon^\lambda \big(\mathcal{E}_{s,t}(y,x;\mathcal{A}) \cap \big\{ \mathbf{C}_t(y,s) \setminus \{x\} = \mathcal{C} \big\} \big). \tag{6.20}$$

Generalizing the definition (6.17) from a single end (x,t) to a pair $\{(x,t),(z,t)\}$ with $x \neq z$ as

$$\mathcal{E}_{s,t}(y,\{x,z\};\mathcal{A}) \tag{6.21}$$
$$= \bigcup_{\boldsymbol{v}} \Big\{ \big\{ \boldsymbol{v} \notin \mathbf{C}(\mathcal{A} \times \{s\}) \big\} \cap \big\{ (y,s) \longrightarrow \boldsymbol{v} \Longrightarrow \{(x,t),(z,t)\} \subset \mathbf{C}(\mathcal{A} \times \{s\}) \setminus \tilde{\mathbf{C}}^{\boldsymbol{v}}(y,s) \big\} \Big\},$$

and following the argument in (6.11) (see also the discussion around (6.9)–(6.10)), we obtain

$$\begin{aligned}&\frac{1}{\varepsilon}\mathbb{P}^\lambda_\varepsilon\big(\mathcal{E}_{s,t}(y,x;\mathcal{A})\cap\big\{\mathbf{C}_t(y,s)\setminus\{x\}=\mathcal{C}\big\}\big)\\&\qquad=\sum_{z\in\mathbb{Z}^d}\lambda D(x-z)\;\mathbb{P}^\lambda_\varepsilon\big(\mathcal{E}_{s,t}(y,\{x,z\};\mathcal{A})\cap\big\{\mathbf{C}_t(y,s)\setminus\{x\}=\mathcal{C}\big\}\big)+o(1).\end{aligned}\tag{6.22}$$

Advancement of the induction hypothesis. Now we advance the induction hypothesis in $N\geq 1$ by using (6.13)–(6.15) and (6.22).

First, we consider the contribution from $\mathcal{C}'=\{y\}$ in (6.13), which equals

$$\begin{aligned}&\sum_{s=\varepsilon}^{t}{}^{\bullet}\sum_{y\in\square_R}\sum_{b_N=(\,\cdot\,,(y,s-\varepsilon))}\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\varnothing)\;\frac{1}{\varepsilon^2}\mathbb{P}^\lambda_\varepsilon\big(E(b_N,(x,t);\tilde{\mathbf{C}}^{b_N}(y,s-\varepsilon))\cap\big\{\mathbf{C}_t(\overline{b}_N)\setminus\{x\}=\mathcal{C}\big\}\big)\\&\quad=\delta_{N,1}\sum_{b=(\boldsymbol{o},\,\cdot\,)}\frac{1}{\varepsilon^2}\mathbb{P}^\lambda_\varepsilon\big(E(b,(x,t);\tilde{\mathbf{C}}^{b}(\boldsymbol{o}))\cap\big\{\mathbf{C}_t(\overline{b})\setminus\{x\}=\mathcal{C}\big\}\big)\\&\qquad+\varepsilon^2\sum_{s=2\varepsilon}^{t}{}^{\bullet}\sum_{y\in\square_R}\frac{1}{\varepsilon^2}\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\varnothing)\sum_{b=((y,s-\varepsilon),\,\cdot\,)}\frac{1}{\varepsilon^2}\mathbb{P}^\lambda_\varepsilon\big(E(b,(x,t);\tilde{\mathbf{C}}^{b}(y,s-\varepsilon))\cap\big\{\mathbf{C}_t(\overline{b})\setminus\{x\}=\mathcal{C}\big\}\big),\end{aligned}\tag{6.23}$$

where we use $\pi^{(N-1)}_{0;\varepsilon}(y;\varnothing)=\delta_{N,1}\,\delta_{o,y}$ to obtain the first term in the right-hand side. We note that, by using the induction hypothesis, as well as (6.14) and (6.22), the second term is $O(\varepsilon)=o(1)$. Therefore, the first term is the main contribution. By using (6.14) and (6.22) again, as well as the weak convergence of $\mathbb{P}^\lambda_\varepsilon$, the limit in $R\uparrow\infty$ of the continuum limit of (6.23) equals

$$\begin{aligned}\delta_{N,1}\sum_{y,z\in\mathbb{Z}^d}\lambda^2D(y)\,D(x-z)\,\Big[&\mathbb{P}^\lambda\big(\mathcal{E}_{0,t}(o,\{x,z\};\{y\})\cap\big\{\mathbf{C}_t(\boldsymbol{o})\setminus\{x\}=\mathcal{C}\big\}\big)\\&+\mathbb{P}^\lambda\big(\mathcal{E}_{0,t}(y,\{x,z\};\{o\})\cap\big\{\mathbf{C}_t(y,0)\setminus\{x\}=\mathcal{C}\big\}\big)\Big].\end{aligned}\tag{6.24}$$

Next, we consider the contribution from $\mathcal{C}'\supsetneq\{y\}$ in (6.13), which equals

$$\begin{aligned}&\varepsilon\sum_{s=2\varepsilon}^{t}{}^{\bullet}\sum_{y\in\square_R}\sum_{\mathcal{A}\subset\square_R\setminus\{y\}:\mathcal{A}\neq\varnothing}\frac{1}{\varepsilon^2}\pi^{(N-1)}_{s-\varepsilon;\varepsilon}(y;\mathcal{A})\\&\qquad\times\sum_{z\in\mathbb{Z}^d}\lambda D(x-z)\;\mathbb{P}^\lambda_\varepsilon\big(\mathcal{E}_{s,t}(y,\{x,z\};\mathcal{A})\cap\big\{\mathbf{C}_t(y,s)\setminus\{x\}=\mathcal{C}\big\}\big)+o(1),\end{aligned}\tag{6.25}$$

where we use (6.15) and (6.22), as well as $\pi^{(N-1)}_{0;\varepsilon}(y;\mathcal{A})=0$ for $\mathcal{A}\neq\varnothing$ (so that the sum over s starts from $s=2\varepsilon$). By the dominated convergence theorem, as well as the induction hypothesis and the weak convergence of $\mathbb{P}^\lambda_\varepsilon$, the limit in $R\uparrow\infty$ of the continuum limit of (6.25) equals

$$\int_0^t ds\sum_{y,z\in\mathbb{Z}^d}\lambda D(x-z)\sum_{\substack{\mathcal{A}\subset\mathbb{Z}^d\setminus\{y\}\\\mathcal{A}\neq\varnothing}}\pi^{(N-1)}_{s}(y;\mathcal{A})\;\mathbb{P}^\lambda\big(\mathcal{E}_{s,t}(y,\{x,z\};\mathcal{A})\cap\big\{\mathbf{C}_t(y,s)\setminus\{x\}=\mathcal{C}\big\}\big).\tag{6.26}$$

Therefore, the limit $\pi^{(N)}_t(x;\mathcal{C})$ for $t>0$ exists and equals the sum of (6.24) and (6.26). This advances the induction hypothesis.

Bounds on π_t^λ in (2.37) and convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$. The bound on $\sum_{x\in\mathbb{Z}^d}|x|^q\pi_t^\lambda(x)$ follow immediately from the pointwise convergence of $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^\lambda(x)$, together with the uniform bounds in Proposition 2.3 and dominated convergence for the sum over x.

To prove convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$, we first note that by [27, Section 3.1], $\lambda_c^{(\varepsilon)}\to\lambda_c$. Convergence of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$ follows by dominated convergence, together with the identification of $A^{(\varepsilon)}$ and $v^{(\varepsilon)}$ in (2.35). Thus, we obtain that

$$A=\bigg[1+\int_0^\infty dt\; t\,\hat{\pi}_t^{\lambda_c}(0)\bigg]^{-1},\qquad\qquad v=A\bigg[\lambda_c-\frac{1}{\sigma^2}\int_0^\infty dt\;\hat{\nabla}^2\pi_t^{\lambda_c}(0)\bigg].\tag{6.27}$$

Convergence of $\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^\lambda(x)$ and the bound on $\partial_\lambda\pi_t^\lambda$ in (2.37). The only difference between $\frac{1}{\varepsilon^2}\partial_\lambda\pi_{t;\varepsilon}^\lambda(x)$ and $\frac{1}{\varepsilon^2}\pi_{t;\varepsilon}^\lambda(x)$ is the occurrence of the sum over spatial bonds b and the indicator of the event $b\in\{b_n\}\cup\mathtt{piv}[\overline{b}_n,\underline{b}_{n+1}]$. Clearly, the main term in the above comes from $b\in\mathtt{piv}[\overline{b}_n,\underline{b}_{n+1}]$. 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 [22, 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 (2.37) 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 2.3 and dominated convergence for the sum over $x\in\mathbb{Z}^d$.

Continuity in λ of $\partial_\lambda\pi_t^\lambda(x)$. Following the same strategy as above, we may obtain an explicit expression for $\partial_\lambda\pi_t^\lambda(x)$, similar to the expression obtained for $\pi_t^\lambda(x)$ from (6.12), (6.24) and (6.26). Let $\partial_\lambda\pi_t^\lambda(x\,|\,R)$ be equal to $\partial_\lambda\pi_t^\lambda(x)$ with the extra condition $\{\mathbf{C}_{[0,t]}\subset\square_R\}$ being imposed, as in (4.57) for the discretized contact process. Note that, as explained above, $\partial_\lambda\pi_t^\lambda(x)=\partial_\lambda\pi_t^\lambda(x\,|\,R)+o(1)$, where $o(1)$ decays to zero as $R\to\infty$, and that $\partial_\lambda\pi_t^\lambda(x\,|\,R)$ is continuous in λ since it depends only on events in the finite space-time box $\square_R\times[0,t]$. Therefore, $\partial_\lambda\pi_t^\lambda(x)$ is also continuous in λ. This completes the proof. $\square$

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 [19]. 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 (2.30)–(2.32), in order to obtain this uniformity. The argument for $d\le 4$ is almost identical, except for modifications due to the factors β_T and $\hat{\beta}_T$ in (4.53)–(4.54) and (5.37)–(5.39). We discuss the necessary changes for $d\le 4$ 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 (5.8) and (5.11),

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

where the last inequality is due to the fact that $f(\varepsilon)=(c+\varepsilon)(c-\varepsilon)^a$ is decreasing in $\varepsilon\ge0$ if $c>0$ and $a\ge1$, so that $f(\varepsilon)\le f(0)=c^{1+a}$ (in the above inequality, $c=1+m\varepsilon$ and $a=\frac{d-2}{2}$). This completes the proof of $I_m\subset I_{m-1}$. $\square$

Proof of Lemma 5.2. By (5.12)–(5.13) and the trivial inequality $1+x \le e^x$,

$$|f_m(0)| = \left|\prod_{l=1}^m [1+\varepsilon r_l(0)]\right| \le e^{\varepsilon\sum_{l=1}^m |r_l(0)|} \le e^{cK_3\beta}. \tag{A.2}$$

By (5.15)–(5.17) and (5.23), $|f_m(k)/f_m(0)|$ is bounded by

$$\left|\prod_{l=1}^m [1-\varepsilon v_l\, a(k)+\varepsilon\, s_l(k)]\right| \le e^{-\varepsilon\sum_{l=1}^m [v_l a(k)-|s_l(k)|]} \le e^{-m\varepsilon[1-c\,(K_1+K_2+K_3)\beta]\,a(k)}. \tag{A.3}$$

This completes the proof. □

Proof of Lemma 5.3. This is an immediate consequence of (5.18), (5.23), (5.31) and (A.2). □

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

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

where R_2 is empty if $m \gg 1$. Then,

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

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

$$\int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 |f_m(k)| \le c \int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 \le \frac{c\beta}{(1+m\varepsilon)^{d/2}}. \tag{A.5}$$

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

$$\int_{R_1} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 |f_m(k)| \le c \int_{R_1} \frac{d^dk}{(2\pi)^d}\, e^{-cm\varepsilon L^2|k|^2} \le \frac{c\beta}{(1+m\varepsilon)^{d/2}}. \tag{A.6}$$

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

$$\int_{R_2} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 |f_m(k)| \le c \int_{R_2} \frac{d^dk}{(2\pi)^d}\, \hat{D}(k)^2 r^{-m\varepsilon} \le c\beta\, r^{-m\varepsilon}. \tag{A.7}$$

Since $r^{-m\varepsilon} \le c\,(1+m\varepsilon)^{-d/2}$, 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

$$\frac{K_4}{(1+m\varepsilon)^{d/2}} \sum_{i=3}^4 \int_{R_i} \frac{d^dk}{(2\pi)^d}\, \frac{\hat{D}(k)^2}{a(k)^{2+\rho}}. \tag{A.8}$$

It thus suffices to bound the integral by $c\beta$. On R_3, we use the inequality $\hat{D}(k)^2 \le 1$ and the assumption $a(k) \asymp L^2|k|^2$ for $\|k\|_\infty \le L^{-1}$. Since $d > 2(2+\rho)$ (cf., (5.6)), we obtain

$$\int_{R_3} \frac{d^dk}{(2\pi)^d}\, \frac{\hat{D}(k)^2}{a(k)^{2+\rho}} \le \frac{c}{L^{4+2\rho}} \int_{\|k\|_\infty \le L^{-1}} \frac{d^dk}{|k|^{4+2\rho}} \le c\beta. \tag{A.9}$$

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

A.1.2 Initialization and advancement of the induction hypotheses

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

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

(H3) By (5.5) and (5.12), $r_1(k) \equiv \lambda - 1$ and thus $|r_1(k) - r_1(0)| \equiv 0$. Together with $\lambda \in I_1$, we obtain $|r_1(0)| \le K_1\beta/(1+\varepsilon)^{(d-2)/2}$. Therefore, (H3) holds, if $K_3 \ge K_1$.

(H4) By (5.5), $|f_1(k)| \le 1 + 3\varepsilon$ and $|f_1(k) - f_0(k)| \le 3\varepsilon$ for $\beta \ll 1$. Together with the trivial bound $a(k) \le 2$, (H4) is proved to hold, if $K_4 \ge (1+3\varepsilon)2^{2+\rho}(1+\varepsilon)^{d/2}$ and $K_5 \ge 3 \cdot 2^{1+\rho}(1+\varepsilon)^{d/2}$.

Next we advance the induction hypotheses for $\lambda \in I_{n+1}$ under the assumption that (H1)–(H4) hold for all $m \le n$. As mentioned below Lemma 5.4, this assumption implies (2.29) for all $s \le n\varepsilon$ if $K \gg K_4$, and thus implies (2.30)–(2.32) for all $s \le n\varepsilon + \varepsilon$. By (5.3), these bounds are translated into the following bounds for all $m \le n+1$: there is a $C_K < \infty$ such that

$$|e_m(k)| \le \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}}, \qquad |e_m(k) - e_m(0)| \le \frac{\varepsilon^2 C_K \beta\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}, \tag{A.10}$$

$$|g_m(k)| \le \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}}, \qquad |\nabla^2 g_m(0)| \le \frac{\varepsilon^2 C_K \sigma^2 \beta}{(1+m\varepsilon)^{(d-2)/2}}, \tag{A.11}$$

$$\Big| g_m(k) - g_m(0) - \frac{a(k)}{\sigma^2}\, \nabla^2 g_m(0) \Big| \le \frac{\varepsilon^2 C_K \beta\, a(k)^{1+\Delta'}}{(1+m\varepsilon)^{(d-2)/2-\Delta'}}, \tag{A.12}$$

$$|\partial_\lambda g_m(0)| \le \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{(d-2)/2}}. \tag{A.13}$$

We note that C_K depends on K and that, by Lemmas 5.2–5.4, K depends only on K_4 when $\beta \ll 1$. Therefore, we can choose C_K large depending only on K_4 when $\beta \ll 1$.

Advancement of (H1). By (5.7) and the mean-value theorem,

$$\begin{aligned}
\lambda_{n+1} - \lambda_n &= -\frac{1}{\varepsilon}\, g_{n+1}(0;\lambda_n) - \frac{1}{\varepsilon}\sum_{m=2}^{n} [g_m(0;\lambda_n) - g_m(0;\lambda_{n-1})] \\
&= -\frac{1}{\varepsilon}\, g_{n+1}(0;\lambda_n) - \frac{\lambda_n - \lambda_{n-1}}{\varepsilon} \sum_{m=2}^{n} \partial_\lambda g_m(0;\lambda_*),
\end{aligned} \tag{A.14}$$

for some λ_* between λ_n and λ_{n-1}. Since $\lambda_{n-1} \in I_n$ (cf., (5.8) and (5.11)), λ_* is also in I_n. By (A.11), (A.13) and (H1),

$$|\lambda_{n+1} - \lambda_n| \le \frac{\varepsilon C_K \beta}{[1+(n+1)\varepsilon]^{d/2}} + |\lambda_n - \lambda_{n-1}|\varepsilon \sum_{m=2}^{n} \frac{C_K \beta}{(1+m\varepsilon)^{(d-2)/2}} \le \frac{\varepsilon C_K (1 + cK_1\beta)\beta}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.15}$$

Therefore, (H1) holds for $n+1$, if $\beta \ll 1$ and $K_1 > C_K$. □

Advancement of (H2). Let $1 + M_n$ be the denominator of (5.9), and let N_n be the numerator of (5.9). Then,

$$v_{n+1} - v_n = \frac{\frac{-1}{\sigma^2\varepsilon}\nabla^2 g_{n+1}(0)}{1+M_{n+1}} - \frac{N_n\, n\, g_{n+1}(0)}{(1+M_{n+1})(1+M_n)}. \tag{A.16}$$

By (A.11), we obtain that, for $m \le n+1$,

$$|M_m| \le \varepsilon \sum_{l=2}^{m} \frac{(l-1)\varepsilon C_K \beta}{(1+l\varepsilon)^{d/2}} \le cC_K\beta, \qquad |N_m - \lambda| \le \varepsilon \sum_{l=2}^{m} \frac{C_K\beta}{(1+l\varepsilon)^{(d-2)/2}} \le cC_K\beta, \tag{A.17}$$

and

$$\Big|\frac{-1}{\sigma^2\varepsilon}\,\nabla^2 g_{n+1}(0)\Big| \le \frac{\varepsilon C_K\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}, \qquad |n\, g_{n+1}(0)| \le \frac{n\varepsilon^2 C_K\beta}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.18}$$

Therefore,

$$\begin{aligned} |v_{n+1} - v_n| &\le \frac{\varepsilon C_K\beta}{(1-cC_K\beta)[1+(n+1)\varepsilon]^{(d-2)/2}} + \frac{(\lambda + cC_K\beta)n\varepsilon^2 C_K\beta}{(1-cC_K\beta)^2[1+(n+1)\varepsilon]^{d/2}} \\ &= \frac{1-cC_K\beta + (\lambda+cC_K\beta)\frac{n\varepsilon}{1+(n+1)\varepsilon}}{(1-cC_K\beta)^2}\,\frac{\varepsilon C_K\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}} \le \frac{1+\lambda}{(1-cC_K\beta)^2}\,\frac{\varepsilon C_K\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}. \end{aligned} \tag{A.19}$$

Since $\lambda \in I_{n+1}$, (H2) holds for $n+1$, if $\beta \ll 1$ and $K_2 > 2C_K$. □

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 (5.4) by $f_n(k)$ and using $g_1(k) = 1 - \varepsilon + \lambda\varepsilon\hat{D}(k)$,

$$\begin{aligned} \frac{f_{n+1}(k)}{f_n(k)} &= g_1(k) + \sum_{m=1}^{n} g_{m+1}(k)\frac{f_{n-m}(k)}{f_n(k)} + \frac{e_{n+1}(k)}{f_n(k)} \\ &= 1 - \varepsilon v_{n+1}a(k) + \varepsilon\Bigg[v_{n+1}a(k) - 1 + \lambda\hat{D}(k) + \frac{1}{\varepsilon}\sum_{m=1}^{n} g_{m+1}(k)\frac{f_{n-m}(k)}{f_n(k)} + \frac{e_{n+1}(k)}{\varepsilon f_n(k)}\Bigg]. \end{aligned} \tag{A.20}$$

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

$$\begin{aligned} r_{n+1}(0) &= -1 + \lambda + \frac{1}{\varepsilon}\sum_{m=1}^{n} g_{m+1}(0)\frac{f_{n-m}(0)}{f_n(0)} + \frac{e_{n+1}(0)}{\varepsilon f_n(0)} \\ &= \Bigg[\lambda - 1 + \frac{1}{\varepsilon}\sum_{m=2}^{n+1} g_m(0)\Bigg] + \frac{1}{\varepsilon}\sum_{m=2}^{n+1} g_m(0)\Bigg[\frac{f_{n+1-m}(0)}{f_n(0)} - 1\Bigg] + \frac{e_{n+1}(0)}{\varepsilon f_n(0)} \\ &= r^{(1)}_{n+1}(0) + r^{(2)}_{n+1}(0) + r^{(3)}_{n+1}(0), \end{aligned} \tag{A.21}$$

where we denote the first, second and third terms in (A.21) by $r^{(1)}_{n+1}(0)$, $r^{(2)}_{n+1}(0)$ and $r^{(3)}_{n+1}(0)$, respectively. Similarly, we can obtain an expression for $r_{n+1}(k) - r_{n+1}(0)$. To do so, we note that, by (5.9),

$$v_{n+1} = \lambda - \frac{1}{\sigma^2\varepsilon}\sum_{m=2}^{n+1} \nabla^2 g_m(0) - v_{n+1}\sum_{m=2}^{n+1}(m-1)\, g_m(0). \tag{A.22}$$

Using this identity, we obtain

$$
\begin{aligned}
r_{n+1}(k) - r_{n+1}(0) &= (v_{n+1} - \lambda)\, a(k) + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} [g_m(k) - g_m(0)] \frac{f_{n+1-m}(k)}{f_n(k)} \\
&\quad + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} g_m(0) \left[\frac{f_{n+1-m}(k)}{f_n(k)} - \frac{f_{n+1-m}(0)}{f_n(0)} \right] + \frac{1}{\varepsilon} \left[\frac{e_{n+1}(k)}{f_n(k)} - \frac{e_{n+1}(0)}{f_n(0)} \right] \\
&= \frac{1}{\varepsilon} \sum_{m=2}^{n+1} \left[[g_m(k) - g_m(0)]\, \frac{f_{n+1-m}(k)}{f_n(k)} - \frac{a(k)}{\sigma^2}\, \nabla^2 g_m(0) \right] \\
&\quad + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} g_m(0) \left[\frac{f_{n+1-m}(k)}{f_n(k)} - \frac{f_{n+1-m}(0)}{f_n(0)} - \varepsilon v_{n+1}\, (m-1)\, a(k) \right] \\
&\quad + \frac{1}{\varepsilon} \left[\frac{e_{n+1}(k)}{f_n(k)} - \frac{e_{n+1}(0)}{f_n(0)} \right] \qquad \text{(A.23)} \\
&= \Delta r^{(1)}_{n+1}(k) + \Delta r^{(2)}_{n+1}(k) + \Delta r^{(3)}_{n+1}(k),
\end{aligned}
$$

where we denote the first, second and third terms in (A.23) by $\Delta r^{(1)}_{n+1}(k)$, $\Delta r^{(2)}_{n+1}(k)$ and $\Delta r^{(3)}_{n+1}(k)$, respectively.

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

Advancement of the first inequality in (5.13). We recall that $r_{n+1}(0)$ has been decomposed, as in (A.21), into $r^{(i)}_{n+1}(0)$ for $i = 1, 2, 3$. First, we investigate $r^{(1)}_{n+1}(0)$. By (5.7) and the mean-value theorem, we have

$$
\begin{aligned}
|r^{(1)}_{n+1}(0)| &\le |\lambda - \lambda_n| + |\lambda_n - \lambda_{n+1}| + \left| \lambda_{n+1} - 1 + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} g_m(0; \lambda) \right| \\
&= |\lambda_n - \lambda_{n+1}| + |\lambda - \lambda_n| + \left| \frac{1}{\varepsilon} \sum_{m=2}^{n+1} [g_m(0; \lambda) - g_m(0; \lambda_n)] \right| \\
&\le |\lambda_n - \lambda_{n+1}| + |\lambda - \lambda_n| \left[1 + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} |\partial_\lambda g_m(0; \lambda_*)| \right], \qquad \text{(A.24)}
\end{aligned}
$$

for some λ_* between λ and λ_n. Since $\lambda \in I_{n+1} \subset I_n$, λ_* is also in I_n. By (5.11), (5.8) and (A.13),

$$
|r^{(1)}_{n+1}(0)| \le \frac{\varepsilon K_1 \beta}{[1 + (n+1)\varepsilon]^{d/2}} + \frac{K_1 \beta}{(1 + n\varepsilon)^{(d-2)/2}} \left[1 + \sum_{m=2}^{n+1} \frac{\varepsilon C_K \beta}{(1 + m\varepsilon)^{(d-2)/2}} \right] \le \frac{c K_1 \beta}{[1 + (n+1)\varepsilon]^{(d-2)/2}}. \qquad \text{(A.25)}
$$

Therefore, we need $K_3 \gg K_1$.

Next we investigate $r^{(2)}_{n+1}(0)$. We will use the following results of Taylor's theorem applied to $h(t) = \prod_i (1 + c_i t)^{-1}$ with $|c_i| < 1$ for all i:

$$
|h(1) - h(0)| \le \sup_{t \in (0,1)} |h'(t)| \le \sum_i \frac{|c_i|}{1 - |c_i|}\, e^{\sum_j \frac{|c_j|}{1 - |c_j|}}, \qquad \text{(A.26)}
$$

$$
|h(1) - h(0) - h'(0)| \le \frac{1}{2} \sup_{t \in (0,1)} |h''(t)| \le \left(\sum_i \frac{|c_i|}{1 - |c_i|} \right)^2 e^{\sum_j \frac{|c_j|}{1 - |c_j|}}, \qquad \text{(A.27)}
$$

By (5.16), (A.11) and (A.26),

$$
|r^{(2)}_{n+1}(0)| \le \frac{1}{\varepsilon} \sum_{m=2}^{n+1} |g_m(0)| \left| \prod_{l=n+2-m}^{n} [1 + \varepsilon\, r_l(0)]^{-1} - 1 \right| \le \sum_{m=2}^{n+1} \frac{\varepsilon C_K \beta}{(1 + m\varepsilon)^{d/2}}\ \phi_m\, e^{\phi_m}, \qquad \text{(A.28)}
$$

where, by (5.13),

$$\phi_m = \sum_{l=n+2-m}^{n} \frac{\varepsilon|r_l(0)|}{1-\varepsilon|r_l(0)|} \leq \varepsilon \sum_{l=n+2-m}^{n} \frac{cK_3\beta}{(1+l\varepsilon)^{(d-2)/2}}, \tag{A.29}$$

and thus $e^{\phi_m} \leq e^{cK_3\beta}$ for all $m \leq n+1$. Substituting (A.29) into (A.28) and using Lemma 4.7 with $(a,b) = (\frac{d}{2}, \frac{d-2}{2})$, we obtain

$$|r_{n+1}^{(2)}(0)| \leq \frac{cC_K K_3 \beta^2}{[1+(n+1)\varepsilon]^{(d-2)/2}}. \tag{A.30}$$

Finally, we investigate $r_{n+1}^{(3)}(0)$. As in (A.28), $|f_n(0)^{-1}-1|$ is bounded by

$$\left|\prod_{l=1}^{n}[1+\varepsilon r_l(0)]^{-1} - 1\right| \leq \phi_{n+1}\, e^{\phi_{n+1}} \leq cK_3\beta. \tag{A.31}$$

Using (A.10), we obtain

$$|r_{n+1}^{(3)}(0)| \leq \frac{\varepsilon C_K(1+cK_3\beta)\beta}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.32}$$

The advancement of the first inequality in (5.13) is now completed by (A.21), (A.25), (A.30) and (A.32), if $\beta \ll 1$ and $K_3 \gg K_1 \vee C_K$. □

Advancement of the second inequality in (5.13). Recall that $k \in \mathcal{A}_{n+1}$, and that $r_{n+1}(k) - r_{n+1}(0)$ has been decomposed, as in (A.23), into $\Delta r_{n+1}^{(i)}(k)$ for $i = 1, 2, 3$.

First, we investigate $\Delta r_{n+1}^{(1)}(k)$, which is bounded as

$$\begin{aligned} |\Delta r_{n+1}^{(1)}(k)| \leq \frac{1}{\varepsilon}\sum_{m=2}^{n+1}\left|g_m(k) - g_m(0) - \frac{a(k)}{\sigma^2}\nabla^2 g_m(0)\right| + \frac{1}{\varepsilon}\sum_{m=2}^{n+1}|g_m(k)-g_m(0)|\left|\frac{f_{n+1-m}(0)}{f_n(0)} - 1\right| \\ + \frac{1}{\varepsilon}\sum_{m=2}^{n+1}|g_m(k)-g_m(0)|\left|\frac{f_{n+1-m}(k)}{f_n(k)} - \frac{f_{n+1-m}(0)}{f_n(0)}\right|. \end{aligned} \tag{A.33}$$

By (A.12) with $\delta < \Delta' < \frac{d-4}{2}$, the first sum is bounded by

$$\varepsilon\sum_{m=2}^{n+1}\frac{C_K\beta\, a(k)^{1+\Delta'}}{(1+m\varepsilon)^{(d-2)/2-\Delta'}} \leq cC_K\beta\, a(k)\left[\frac{\log[2+(n+1)\varepsilon]}{[1+(n+1)\varepsilon]}\right]^{\Delta'} \leq \frac{cC_K\beta\, a(k)}{[1+(n+1)\varepsilon]^{\delta}}, \tag{A.34}$$

while the second sum in (A.33) is first bounded similarly to (A.28), and then bounded, by using (A.12) with $\Delta' = 0$ and (A.29), as well as Lemma 4.7 with $a = b = \frac{d-2}{2}$, by

$$\sum_{m=2}^{n+1}\frac{2\varepsilon C_K\beta\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}\sum_{l=n+2-m}^{n}\frac{\varepsilon cK_3\beta}{(1+l\varepsilon)^{(d-2)/2}} \leq \frac{cC_K K_3\beta^2 a(k)}{[1+(n+1)\varepsilon]^{(d-2)/2\wedge(d-4)}} \leq \frac{cC_K K_3\beta^2 a(k)}{[1+(n+1)\varepsilon]^{2\delta}}, \tag{A.35}$$

where we use $\frac{d-2}{2} \wedge (d-4) = \frac{d-4}{2} + 1 \wedge \frac{d-4}{2} \geq 2\delta$. By using (A.12) with $\Delta' = 0$ again and (A.26), the third sum in (A.33) is bounded similarly to (A.28) by

$$\begin{aligned} \sum_{m=2}^{n+1}\frac{2\varepsilon C_K\beta\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}\left|\frac{f_{n+1-m}(0)}{f_n(0)}\right|\left|\prod_{l=n+2-m}^{n}[1-\varepsilon v_l\, a(k)+\varepsilon s_l(k)]^{-1} - 1\right| \\ \leq \sum_{m=2}^{n+1}\frac{2\varepsilon C_K\beta\, a(k)}{(1+m\varepsilon)^{(d-2)/2}}\left(1+\phi_m\, e^{\phi_m}\right)\psi_m(k)\, e^{\psi_m(k)}, \end{aligned} \tag{A.36}$$

where $\phi_m\, e^{\phi_m} \le cK_3\beta$ as discussed below (A.29), and

$$\psi_m(k) = \sum_{l=n+2-m}^{n} \frac{\varepsilon[v_l a(k) + |s_l(k)|]}{1-\varepsilon[v_l a(k) + |s_l(k)|]}. \tag{A.37}$$

By (5.17) and (5.23),

$$v_l a(k) + |s_l(k)| \le \left[v_l + \frac{(1+\varepsilon v_l)K_3\beta}{(1-\varepsilon K_3\beta)(1+l\varepsilon)^\delta}\right] a(k) \le [1 + c\,(K_1+K_2+K_3)\beta]\, a(k) \equiv q\, a(k). \tag{A.38}$$

Since $k \in \mathcal{A}_{n+1}$, $\psi_m(k)$ is bounded by

$$\psi_m(k) \le \frac{(m-1)\varepsilon q\, a(k)}{1-\varepsilon q\, a(k)} \le [1 + c\varepsilon\, a(k)](m-1)\varepsilon q\, a(k), \tag{A.39}$$

which is further 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 hence

$$e^{\psi_m(k)} \le c\, e^{\gamma q \log[2+(n+1)\varepsilon]} \le c\,[1+(n+1)\varepsilon]^{\gamma q}. \tag{A.40}$$

Substituting (A.39)–(A.40) into (A.36), and using $a(k) \le \gamma \frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon}$ and $\gamma q + \delta < 1 \wedge \frac{d-4}{2}$ for $\beta \ll 1$ (cf., (5.6) and (A.38)), we can bound (A.36) by

$$cC_K\beta\, a(k) \frac{\log[2+(n+1)\varepsilon]}{[1+(n+1)\varepsilon]^{1-\gamma q}}\ \varepsilon \sum_{m=2}^{n+1} \frac{(m-1)\varepsilon}{(1+m\varepsilon)^{(d-2)/2}} \le \frac{cC_K\beta\, a(k)}{[1+(n+1)\varepsilon]^\delta}. \tag{A.41}$$

By (A.33)–(A.35) and (A.41), if $\beta \ll 1$ and $K_3 \gg C_K$, we obtain

$$|\Delta r^{(1)}_{n+1}(k)| \le \frac{\frac{1}{3}K_3\beta a(k)}{[1+(n+1)\varepsilon]^\delta}. \tag{A.42}$$

Next, we investigate $|\Delta r^{(2)}_{n+1}(k)|$, which is bounded, by using (A.11) and (A.28), as

$$\begin{aligned}
|\Delta r^{(2)}_{n+1}(k)| &\le \frac{1}{\varepsilon} \sum_{m=2}^{n+1} |g_m(0)| \left|\frac{f_{n+1-m}(0)}{f_n(0)}\right| \left| \prod_{l=n+2-m}^{n} [1-\varepsilon v_l a(k) + \varepsilon s_l(k)]^{-1} - 1 - (m-1)\varepsilon v_{n+1} a(k)\right| \\
&\quad + \frac{1}{\varepsilon} \sum_{m=2}^{n+1} |g_m(0)| \left|\frac{f_{n+1-m}(0)}{f_n(0)} - 1\right| (m-1)\varepsilon v_{n+1} a(k) \\
&\le \varepsilon \sum_{m=2}^{n+1} \frac{C_K\beta(1+\phi_m e^{\phi_m})}{(1+m\varepsilon)^{d/2}} \left| \prod_{l=n+2-m}^{n} [1-\varepsilon v_l a(k) + \varepsilon s_l(k)]^{-1} - 1 - (m-1)\varepsilon v_{n+1} a(k)\right| \\
&\quad + \varepsilon \sum_{m=2}^{n+1} \frac{C_K\beta v_{n+1} a(k)\,(m-1)\varepsilon}{(1+m\varepsilon)^{d/2}}\ \phi_m e^{\phi_m}.
\end{aligned} \tag{A.43}$$

Using (5.23), (A.29), Lemma 4.7 with $a = b = \frac{d-2}{2}$ and $\delta < 1 \wedge \frac{d-4}{2}$, we can bound the second sum by $cC_K\beta a(k)[1+(n+1)\varepsilon]^{-2\delta}$. The first sum in (A.43) is bounded, by using (A.27), by

$$\begin{aligned}
&\varepsilon \sum_{m=2}^{n+1} \frac{cC_K\beta}{(1+m\varepsilon)^{d/2}} \left[\left| \prod_{l=n+2-m}^{n} [1-\varepsilon v_l a(k) + \varepsilon s_l(k)]^{-1} - 1 - \sum_{l=n+2-m}^{n} \varepsilon[v_l a(k) - s_l(k)]\right| \right.\\
&\qquad\qquad \left. + \left| \sum_{l=n+2-m}^{n} \varepsilon[(v_l - v_{n+1}) a(k) - s_l(k)]\right|\right] \\
&\le \varepsilon \sum_{m=2}^{n+1} \frac{cC_K\beta}{(1+m\varepsilon)^{d/2}} \left[\psi_m(k)^2 e^{\psi_m(k)} + \sum_{l=n+2-m}^{n} \varepsilon\left[\sum_{j=l+1}^{n+1} |v_j - v_{j-1}| a(k) + |s_l(k)|\right]\right].
\end{aligned} \tag{A.44}$$

Similarly to (A.41), the contribution from $\psi_m(k)^2 e^{\psi_m(k)}$ is bounded by $cC_K\beta a(k)[1+(n+1)]^{-\delta}$. By (5.11), (5.17) and Lemma 4.7 with $a = \frac{d}{2}$ and $b = \delta (< \frac{d-4}{2})$, the other contribution is bounded by

$$\varepsilon \sum_{m=2}^{n+1} \frac{cC_K\beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{n} \varepsilon \left[\frac{cK_2\beta a(k)}{(1+l\varepsilon)^{(d-4)/2}} + \frac{cK_3\beta a(k)}{(1+l\varepsilon)^{\delta}}\right] \leq \frac{cC_K(K_2+K_3)\beta^2 a(k)}{[1+(n+1)\varepsilon]^{\delta}}. \tag{A.45}$$

Therefore, if $\beta \ll 1$ and $K_3 \gg C_K$, we obtain

$$|\Delta r^{(2)}_{n+1}(k)| \leq \frac{\frac{1}{3}K_3\beta a(k)}{[1+(n+1)\varepsilon]^{\delta}}. \tag{A.46}$$

Finally, we investigate $|\Delta r^{(3)}_{n+1}(k)|$, which is bounded as

$$|\Delta r^{(3)}_{n+1}(k)| \leq \left|\frac{e_{n+1}(k)}{\varepsilon f_n(0)}\right| \left|\frac{f_n(0)}{f_n(k)} - 1\right| + \left|\frac{e_{n+1}(k) - e_{n+1}(0)}{\varepsilon f_n(0)}\right|. \tag{A.47}$$

By (A.31), $|f_n(0)| \geq 1 - cK_3\beta$. As in (A.36), $\left|\frac{f_n(0)}{f_n(k)} - 1\right|$ is bounded, by using (A.39)–(A.40), by $\psi_{n+1}(k)e^{\psi_{n+1}(k)} \leq c[1+(n+1)\varepsilon]^{1+\gamma q} a(k)$. Therefore, using (A.10) and taking β sufficiently small such that $\gamma q + \delta < \frac{d-4}{2}$ (cf., (5.6) and (A.38)), we obtain

$$|\Delta r^{(3)}_{n+1}(k)| \leq \frac{c\varepsilon C_K\beta a(k)}{[1+(n+1)\varepsilon]^{(d-2)/2-\gamma q}} \leq \frac{c\varepsilon C_K\beta a(k)}{[1+(n+1)\varepsilon]^{\delta+1}} \leq \frac{\frac{1}{3}K_3\beta a(k)}{[1+(n+1)\varepsilon]^{\delta}}, \tag{A.48}$$

if $K_3 \gg C_K$.

The advancement of the second inequality in (5.13) is now completed by (A.42), (A.46) and (A.48), if $\beta \ll 1$ and $K_3 \gg C_K$. □

Advancement of (H4). To advance (H4), we rewrite (5.4) as

$$f_{n+1}(k) = \left[g_1(k) + \sum_{m=2}^{n+1} g_m(k)\right] f_n(k) + W_{n+1}(k) + e_{n+1}(k), \tag{A.49}$$

where

$$W_{n+1}(k) = \sum_{m=2}^{n+1} g_m(k)[-f_n(k) + f_{n+1-m}(k)] = \sum_{m=2}^{n+1} g_m(k) \sum_{l=n+2-m}^{n} [f_{l-1}(k) - f_l(k)]. \tag{A.50}$$

Furthermore, using $g_1(k) = 1 - \varepsilon + \lambda\varepsilon\hat{D}(k) = 1 - \lambda\varepsilon a(k) + (\lambda - 1)\varepsilon$, we have

$$\begin{aligned} g_1(k) + \sum_{m=2}^{n+1} g_m(k) &= 1 - \left[\lambda - \frac{1}{\sigma^2\varepsilon}\sum_{m=2}^{n+1}\nabla^2 g_m(0)\right]\varepsilon a(k) + \varepsilon\left[\lambda - 1 + \frac{1}{\varepsilon}\sum_{m=2}^{n+1} g_m(0)\right] \\ &\quad + \sum_{m=2}^{n+1}\left[g_m(k) - g_m(0) - \frac{a(k)}{\sigma^2}\nabla^2 g_m(0)\right] \\ &= 1 - N_{n+1}\varepsilon a(k) + \varepsilon r^{(1)}_{n+1}(0) + X_{n+1}(k), \end{aligned} \tag{A.51}$$

where we recall N_n and $r^{(1)}_{n+1}(0)$ in (A.16) and (A.21), respectively, and denote the last sum in (A.51) by $X_{n+1}(k)$. Therefore,

$$f_{n+1}(k) = f_n(k)\left[1 - N_{n+1}\,\varepsilon\,a(k) + \varepsilon r^{(1)}_{n+1}(0) + X_{n+1}(k)\right] + W_{n+1}(k) + e_{n+1}(k). \tag{A.52}$$

We have already obtained $|N_{n+1} - \lambda| \le cC_K\beta$ in (A.17) and $|r^{(1)}_{n+1}(0)| \le cK_1\beta[1 + (n+1)\varepsilon]^{-(d-2)/2}$ in (A.25), while $X_{n+1}(k)$ equals ε times the first sum of (A.33) and is bounded, by using the leftmost expression of (A.34) with $\Delta' < \frac{d-4}{2}$, by $c\varepsilon C_K\beta\, a(k)^{1+\Delta'}$. We prove below that, for $k \notin \mathcal{A}_{n+1}$,

$$|W_{n+1}(k)| \le \varepsilon \frac{cC_K(1 + K_3\beta + K_5)\beta\, a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}}. \tag{A.53}$$

Assuming (A.53), we first advance the second inequality in (5.14), and then advance the first inequality in (5.14). To advance these inequalities, we will use the first inequality in (5.14) for $m = n$ in the extended region $\mathcal{A}^c_{n+1} = \mathcal{A}^c_n \,\dot\cup\, (\mathcal{A}_n \setminus \mathcal{A}_{n+1})$. We now verify the use of this inequality for $k \in \mathcal{A}_n \setminus \mathcal{A}_{n+1}$. When $n\varepsilon \le T$ for some large T, we can choose $K_4 \gg 1$ (depending on T) such that, for all $k \in [-\pi, \pi]^d$,

$$|f_n(k)| \le \|\tau_{n\varepsilon}\|_1 \le \|p^{*n}_\varepsilon\|_1 = (1 - \varepsilon + \lambda\varepsilon)^n \le e^{(\lambda-1)n\varepsilon} \le \frac{2^{-2-\rho}K_4}{(1 + n\varepsilon)^{d/2}} \le \frac{K_4 a(k)^{-2-\rho}}{(1 + n\varepsilon)^{d/2}}. \tag{A.54}$$

When $n\varepsilon > T$, we use Lemma 5.2 and $k \in \mathcal{A}_n \setminus \mathcal{A}_{n+1}$ (so that $\gamma \frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon} < a(k) \le \gamma \frac{\log(2+n\varepsilon)}{1+n\varepsilon}$) to obtain

$$\begin{aligned}|f_n(k)| &\le ce^{-n\varepsilon q a(k)} \le c(2 + n\varepsilon)^{-\frac{n\varepsilon}{1+(n+1)\varepsilon}\frac{\log[2+(n+1)\varepsilon]}{\log(2+n\varepsilon)} q\gamma} \\ &\le c(1 + n\varepsilon)^{-q'\gamma} = \frac{c}{(1 + n\varepsilon)^{d/2}} \frac{(1 + n\varepsilon)^{2+\rho}}{(1 + n\varepsilon)^{q'\gamma - [\frac{d}{2} - (2+\rho)]}} \le \frac{K_4 a(k)^{-2-\rho}}{(1 + n\varepsilon)^{d/2}},\end{aligned} \tag{A.55}$$

if $K_4 \gg 1$, where we use $q'\gamma > \frac{d}{2} - (2 + \rho)$ for $\beta \ll 1$ and $T \gg 1$ (cf., (5.6) and (A.38)).

Therefore, by using the first inequality in (5.14) with $m = n$ for $k \notin \mathcal{A}_{n+1}$, together with (5.22) and (A.52)–(A.53), we obtain

$$\begin{aligned}|f_{n+1}(k) - f_n(k)| &\le \varepsilon \frac{K_4 a(k)^{-2-\rho}}{(1 + n\varepsilon)^{d/2}} \left[(\lambda + cC_K\beta)a(k) + \frac{cK_1\beta}{[1 + (n+1)\varepsilon]^{(d-2)/2}} + cC_K\beta a(k)^{1+\Delta'}\right] \\ &\quad + \varepsilon \frac{cC_K(1 + K_3\beta + K_5)\beta a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}} + \frac{\varepsilon^2 C_K\beta}{[1 + (n+1)\varepsilon]^{d/2}} \\ &\le \varepsilon \frac{cK_4[1 + O(\beta)]\, a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}} + \varepsilon \frac{O(\beta)\, a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}} + \varepsilon^2 \frac{O(\beta)\, a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}},\end{aligned} \tag{A.56}$$

where we use $[1 + (n+1)\varepsilon]^{-(d-2)/2} \le a(k)^{(d-2)/2} \le 2^{(d-4)/2} a(k)$ to obtain the first term, and use $2^{-1-\rho} \le a(k)^{-1-\rho}$ for the third term. This completes the advancement of the second inequality in (5.14), if $\beta \ll 1$ and $K_5 \gg K_4$, under the hypotheses that (A.53) holds for $k \notin \mathcal{A}_{n+1}$.

Since (A.54) holds for $n \le T/\varepsilon$ independently of k, it remains to advance the first inequality of (H4) for $n > T/\varepsilon$. Similarly to (A.56), we have

$$\begin{aligned}|f_{n+1}(k)| \le \frac{K_4 a(k)^{-2-\rho}}{(1 + n\varepsilon)^{d/2}} &\left[|1 - N_{n+1}\varepsilon a(k)| + \frac{c\varepsilon K_1\beta}{[1 + (n+1)\varepsilon]^{(d-2)/2}} + c\varepsilon C_K\beta a(k)^{1+\Delta'}\right] \\ &+ \frac{c\varepsilon C_K(1 + K_3\beta + K_5)\beta a(k)^{-1-\rho}}{[1 + (n+1)\varepsilon]^{d/2}} + \frac{\varepsilon^2 C_K\beta}{[1 + (n+1)\varepsilon]^{d/2}}.\end{aligned} \tag{A.57}$$

Again, by $a(k) \le 2$, the sum of the last two terms is bounded by $\varepsilon O(\beta)\, a(k)^{-2-\rho}[1 + (n+1)\varepsilon]^{-d/2}$. To prove the first inequality in (5.14) with $m = n + 1$, it thus suffices to show that

$$\left[\frac{1 + (n+1)\varepsilon}{1 + n\varepsilon}\right]^{d/2} \left[|1 - N_{n+1}\varepsilon a(k)| + \frac{c\varepsilon K_1\beta}{[1 + (n+1)\varepsilon]^{(d-2)/2}} + c\varepsilon C_K\beta a(k)^{1+\Delta'}\right] < 1. \tag{A.58}$$

To achieve this inequality uniformly in $\varepsilon \le 1$, we consider the case in which $a(k) \le 1/2$ and the other case in which $a(k) > 1/2$ separately.

When $a(k) \le 1/2$, since $N_{n+1} = 1 + O(\beta)$ (cf., (5.22) and (A.17)), we have $|1 - N_{n+1}\varepsilon a(k)| = 1 - N_{n+1}\varepsilon a(k)$ for $\beta \ll 1$. Using $a(k)^{\Delta'} \le 2^{\Delta'}$ and then $a(k) > \gamma\frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon}$, we can bound (A.58) by

$$\begin{aligned}&\Big(1 + \frac{c\varepsilon}{1+n\varepsilon}\Big)\Big[1 - (1-c\beta)\varepsilon a(k) + \frac{c\varepsilon\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}\Big] \qquad \text{(A.59)}\\ &\quad \le 1 - \varepsilon\Big[(1-c\beta)\frac{\gamma\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon} - \frac{c}{1+n\varepsilon} - \Big(1+\frac{c\varepsilon}{1+n\varepsilon}\Big)\frac{c\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}\Big] < 1,\end{aligned}$$

if $\beta \ll 1$ and $T \gg 1$.

Since the above argument also applies to the case in which $1 - \varepsilon a(k) > 1 - (2-\eta)\varepsilon > 0$ (and $\beta \ll 1$, depending on η), it thus remains to consider the other case in which $1-(2-\eta)\varepsilon \le 0$ and $a(k) > 1/2$. In this case, since $\varepsilon \le 1$, we have

$$|1 - \varepsilon\, a(k)| \le [(2-\eta)\varepsilon - 1] \vee \Big(1 - \frac{\varepsilon}{2}\Big) \le 1 - \Big(\eta \wedge \frac{\varepsilon}{2}\Big). \qquad \text{(A.60)}$$

Since $N_{n+1} = 1 + O(\beta)$, (A.58) is bounded by

$$\begin{aligned}&\Big(1 + \frac{c\varepsilon}{1+n\varepsilon}\Big)\Big[1 - \Big(\eta \wedge \frac{\varepsilon}{2}\Big) + c\varepsilon\beta a(k) + \frac{c\varepsilon\beta}{[1+(n+1)\varepsilon]^{(d-2)/2}}\Big]\\ &\quad \le 1 - \Big[\Big(\eta\wedge\frac{\varepsilon}{2}\Big) - \frac{c\varepsilon}{1+n\varepsilon} - \Big(1+\frac{c\varepsilon}{1+n\varepsilon}\Big)c\varepsilon\beta\Big] < 1, \qquad \text{(A.61)}\end{aligned}$$

if $\beta \ll 1$ and $T \gg 1$, depending on η. This completes the proof of (A.58), and hence the advancement of the first inequality in (5.14), if $\beta \ll 1$, $T \gg 1$ and $K_4 \gg 1$, under the hypotheses that (A.53) holds for $k \notin \mathcal{A}_{n+1}$.

Proof of (A.53). 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 $\mu < l \le n$, f_l is in the domain of (H4). We separate the sum over l in (A.50) into two parts, corresponding respectively to $l \le \mu$ and $\mu < l \le n$, yielding $W_{n+1}(k) = W_{n+1}^{\le}(k) + W_{n+1}^{>}(k)$, where

$$|W_{n+1}^{\le}(k)| \le \sum_{m=n+2-\mu}^{n+1} \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{\mu} |f_{l-1}(k) - f_l(k)|, \qquad \text{(A.62)}$$

$$|W_{n+1}^{>}(k)| \le \sum_{m=2}^{n+1} \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=\mu\vee(n+1-m)+1}^{n} |f_{l-1}(k) - f_l(k)|. \qquad \text{(A.63)}$$

By (H4) and Lemma 4.7 with $a = b = \frac{d}{2}$, we easily obtain

$$|W_{n+1}^{>}(k)| \le \sum_{m=2}^{n+1} \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{n} \frac{\varepsilon K_5\, a(k)^{-1-\rho}}{(1+l\varepsilon)^{d/2}} \le \varepsilon\frac{cC_K K_5 \beta\, a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \qquad \text{(A.64)}$$

It remains to consider $|W_{n+1}^{\le}(k)|$. By (5.13), (5.23) and Lemma 5.2, we have

$$\begin{aligned}|f_{l-1}(k) - f_l(k)| &= |f_{l-1}(k)|\,\Big|1 - \big[1 - \varepsilon v_l a(k) + \varepsilon[r_l(k) - r_l(0)] + \varepsilon r_l(0)\big]\Big|\\ &\le ce^{-(l-1)\varepsilon q a(k)}\,\varepsilon\Big[a(k) + \frac{K_3\beta}{(1+l\varepsilon)^{(d-2)/2}}\Big], \qquad \text{(A.65)}\end{aligned}$$

where $q = 1 - O(\beta)$. We fix a small $r > 0$ and separate the sum over m in (A.62) into $\sum_{m>r(n+1)}$ and $\sum_{m\le r(n+1)}$ (the latter sum may be empty depending on μ). The contribution due to the former sum is bounded by

$$\frac{\varepsilon^2 C_K \beta}{[1+(n+1)\varepsilon]^{d/2}} \sum_{m=r(n+1)+1}^{n+1} \sum_{l=n+2-m}^{\mu} ce^{-(l-1)\varepsilon q a(k)}\, \varepsilon \left[a(k) + \frac{K_3\beta}{(1+l\varepsilon)^{(d-2)/2}}\right]$$
$$\leq \frac{\varepsilon^2 C_K \beta}{[1+(n+1)\varepsilon]^{d/2}} \sum_{m=r(n+1)+1}^{n+1} ce^{-(n+1-m)\varepsilon q a(k)}(1+K_3\beta)$$
$$\leq \varepsilon \frac{cC_K(1+K_3\beta)\beta a(k)^{-1}}{[1+(n+1)\varepsilon]^{d/2}} \leq \varepsilon \frac{cC_K(1+K_3\beta)\beta a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.66}$$

To investigate the contribution from $\sum_{m\leq r(n+1)}$, we use the inequality

$$e^{-(l-1)\varepsilon q a(k)} \left[a(k) + \frac{K_3\beta}{(1+l\varepsilon)^{(d-2)/2}}\right] \leq \frac{c(1+K_3\beta)a(k)^{-1-\rho}}{(1+l\varepsilon)^{d/2}}, \tag{A.67}$$

which we will prove below. Assuming this inequality and using Lemma 4.7 with $a = b = \frac{d}{2}$, we obtain that the contribution from $\sum_{m\leq r(n+1)}$ is bounded by

$$\sum_{m=n+2-\mu}^{r(n+1)} \frac{\varepsilon^2 C_K \beta}{(1+m\varepsilon)^{d/2}} \sum_{l=n+2-m}^{\mu} \varepsilon \frac{c(1+K_3\beta)a(k)^{-1-\rho}}{(1+l\varepsilon)^{d/2}} \leq \varepsilon \frac{cC_K(1+K_3\beta)\beta a(k)^{-1-\rho}}{[1+(n+1)\varepsilon]^{d/2}}. \tag{A.68}$$

This, together with (A.64) and (A.66), completes the proof of (A.53) for $k \notin \mathcal{A}_{n+1}$.

It remains to prove (A.67). First, we note that, by $m \leq r(n+1)$ and $n+2-m \leq l \leq \mu \leq n$, as well as $a(k) > \gamma \frac{\log[2+(\mu+1)\varepsilon]}{1+(\mu+1)\varepsilon}$, we have

$$(1+l\varepsilon)^{-(d-2)/2} \leq [1+(1-r)(n+1)\varepsilon]^{-(d-2)/2} \leq \frac{(1-r)^{-(d-2)/2}}{[1+(\mu+1)\varepsilon]^{1+(d-4)/2}} \leq c\, a(k). \tag{A.69}$$

Therefore, the left-hand side of (A.67) is bounded by $c(1+K_3\beta)a(k)\, e^{-(l-1)\varepsilon q a(k)}$. Similarly to (A.55), we have

$$e^{-(l-1)\varepsilon q a(k)} \leq (2+l\varepsilon)^{-\frac{(l-1)\varepsilon q\gamma}{1+(\mu+1)\varepsilon} \frac{\log[2+(\mu+1)\varepsilon]}{\log(2+l\varepsilon)} q\gamma} \leq (1+l\varepsilon)^{-q'\gamma}, \tag{A.70}$$

where $q' = \frac{(1-r)(n+1)\varepsilon}{1+(n+1)\varepsilon} q$. To bound (A.70), we fix $T \gg 1$ and consider the case in which $n\varepsilon \leq T$ and the other case separately. When $n\varepsilon \leq T$, since $l \leq n$, $a(k) \leq 2$ and $2+\rho > 0$, (A.70) is bounded as

$$\frac{(1+l\varepsilon)^{d/2-q'\gamma}}{(1+l\varepsilon)^{d/2}} \leq \frac{(1+T)^{d/2}}{(1+l\varepsilon)^{d/2}} \leq \frac{2^{2+\rho}(1+T)^{d/2}}{(1+l\varepsilon)^{d/2}}\, a(k)^{-2-\rho}. \tag{A.71}$$

When $n\varepsilon > T$, since $a(k) \leq \gamma \frac{\log(2+l\varepsilon)}{1+l\varepsilon}$, (A.70) is bounded as

$$\frac{1}{(1+l\varepsilon)^{d/2}} \frac{(1+l\varepsilon)^{2+\rho}}{(1+l\varepsilon)^{q'\gamma-[\frac{d}{2}-(2+\rho)]}} \leq \frac{c}{(1+l\varepsilon)^{d/2}}\, a(k)^{-2-\rho}, \tag{A.72}$$

where we use $q'\gamma > \frac{d}{2} - (2+\rho)$ for $\beta \ll 1$, $T \gg 1$ and $r \ll 1$. This completes the proof. □

Finally, we summarize the relations among the constants $K_1, \ldots, K_5$ that have been necessary in advancing the induction hypotheses. We have taken $\beta \ll 1$ and have chosen the constants $K_1, \ldots, K_5$ such that

$$K_1 > C_K, \qquad K_2 > 2C_K, \qquad K_3 \gg K_1, \qquad K_5 \gg K_4, \tag{A.73}$$

where, as stated below (A.13), C_K depends only on K_4 (when $\beta \ll 1$). This gives (5.10).

A.2 Advancement below and at four dimensions

The proofs of Lemmas 5.1–5.4 and the advancement of the induction hypotheses for $d \leq 4$ remain almost unchanged, except for the factors β_T and $\hat{\beta}_T$ in (4.53)–(4.54) and (5.37)–(5.39). 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 (5.38), we use (A.14) together with (A.11) and (A.13) with β replaced by β_T. Since $n\varepsilon \leq T \log T$, we obtain

$$|\lambda_{n+1} - \lambda_n| \leq \frac{\varepsilon C_K \beta_T}{[1+(n+1)\varepsilon]^{d/2}} + \frac{\varepsilon K_1 \hat{\beta}_T}{(1+n\varepsilon)^{2+\omega}} \sum_{m=2}^{n} \frac{\varepsilon C_K \beta_T}{(1+m\varepsilon)^{(d-2)/2}} \leq \varepsilon \frac{c C_K (1+K_1\hat{\beta}_T)\hat{\beta}_T}{[1+(n+1)\varepsilon]^{2+\omega}}, \tag{A.74}$$

where we use $\mu \in (0, \alpha-\omega)$. Similarly, we can advance the second inequality in (5.38) and the first inequality in (5.39).

We need a little more care in advancing the second inequality in (5.39) and the inequalities in (5.14). The second inequality in (5.39) is rewritten as in (A.23), and each term is bounded as in (A.42), (A.46) and (A.48) when $d > 4$. We can follow the same line when $d \leq 4$, except that, e.g., the factor β^2 in (A.35) is replaced by $\beta_T \hat{\beta}_T$, and that we use β_T to control the convolution in (A.35), where the power of $1 + l\varepsilon$ is replaced by $1+\omega$. If we have only one factor β as in (A.41), then we use $q = 1 + O(\hat{\beta}_T)$ with $\beta_1 = L_1^{-d} \ll 1$, as well as $\gamma + \delta < \omega$ and $\mu < \alpha - \omega$, to obtain

$$\begin{aligned} & cC_K \beta_T a(k) \frac{\log[2+(n+1)\varepsilon]}{[1+(n+1)\varepsilon]^{1-\gamma q}} [1+(n+1)\varepsilon]^{\frac{6-d}{2}} \\ & \leq \frac{cC_K \hat{\beta}_T a(k)}{[1+(n+1)\varepsilon]^{\delta}} T^{-bd+\mu} [1+(n+1)\varepsilon]^{\frac{4-d}{2}+\gamma q+\delta} \log[2+(n+1)\varepsilon] \\ & \leq \frac{cC_K \hat{\beta}_T a(k)}{[1+(n+1)\varepsilon]^{\delta}} T^{-(\alpha-\mu-\omega)} \leq \frac{cC_K \hat{\beta}_T a(k)}{[1+(n+1)\varepsilon]^{\delta}}. \end{aligned} \tag{A.75}$$

A similar argument applies to the advancement of the inequalities in (5.14). However, since $-\rho > \frac{4-d}{2} \geq 0$ (cf., (5.36)), we cannot use the trivial inequality $a(k) \leq 2$ to obtain, e.g., the low-dimensional version of (A.66). To overcome this difficulty, we use the factor β_T in the bound on $g_m(k)$ and $a(k) \leq \gamma \frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon}$ in (A.66), as well as $\mu < \alpha - \omega = bd + \frac{d-4}{2} - \omega$ and $\frac{d}{2} - (2+\rho) < \gamma < \omega$, to obtain

$$\beta_T \leq \beta_1 T^{-bd} \left[\frac{a(k)}{\gamma \frac{\log[2+(n+1)\varepsilon]}{1+(n+1)\varepsilon}} \right]^{-\rho} \leq c\hat{\beta}_T T^{\alpha-\omega-bd-\rho} a(k)^{-\rho} \leq c\hat{\beta}_T T^{-(\omega-\gamma)} a(k)^{-\rho} \leq c\hat{\beta}_T a(k)^{-\rho}. \tag{A.76}$$

This completes the advancement of the induction hypotheses for $d \leq 4$. □

Acknowledgements

The work of RvdH and AS was supported in part by the Netherlands Organisation for Scientific Research (NWO). The work of RvdH was carried out in part at Delft University of Technology, the Netherlands. The work of AS was supported in part by NSERC of Canada and was carried out in part at the University of British Columbia, Vancouver, Canada. 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] E. Bolthausen and C. Ritzmann. Strong pointwise estimates for the weakly self-avoiding walk. To appear in *Ann. Probab.*

[6] C. Borgs, J. T. Chayes and D. Randall. The van den Berg-Kesten-Reimer inequality. *Perplexing Problems in Probability: Festschrift in honor of Harry Kesten* (eds., M. Bramson and R. Durrett). Birkhäuser (1999): 159-173.

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

[8] 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.

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

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

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

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] R. van der Hofstad and A. Sakai. Critical points for spread-out self-avoiding walk, percolation and the contact process above the upper critical dimensions. Preprint (2004).

[18] 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.

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

[20] 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.

[21] 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.

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

[23] N. Madras and G. Slade. *The Self-Avoiding Walk*. Birkhäuser, Boston (1993).

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

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

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

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

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

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

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

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

[32] G. Slade. The lace expansion and the upper critical dimension for percolation. *Lectures in Applied Mathematics* **27** (1991):53–63.

[33] G. Slade. The lace expansion and its applications. Saint-Flour lecture notes. Preprint (2004).

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