CHEMICAL DISTANCE IN PERCOLATION MODELS AND PHASE TRANSITIONS OF BALLISTIC ANNIHILATION

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Lily Wang Reeves
May 2023

CHEMICAL DISTANCE IN PERCOLATION MODELS AND PHASE TRANSITIONS OF BALLISTIC ANNIHILATION

Lily Wang Reeves, Ph.D.

Cornell University 2023

We consider problems in two areas at the intersection of discrete probability and statistical physics: the chemical distance in planar percolation models and phase transitions of ballistic annihilation. Although the two models are quite different, the techniques and results involved are often of combinatorial interest.

We first discuss the chemical distance in planar Bernoulli percolation and random cluster models in the critical regime. By extending Damron–Hanson–Sosoe's results, we give upper bounds on the expected number of edges to cross a square box from left to right, as well as from the origin to the boundary, in critical percolation clusters. Along the way, we also extend two classical results to new settings.

In the second part of the thesis, we consider a one-dimensional annihilating particle system called ballistic annihilation. We introduce two variants of the symmetric three-velocity ballistic annihilation, for which we prove the existence of phase transitions and compute the critical density using an exactly solvable approach pioneered by Haslegrave–Sidoravicius–Tournier. These variants have considerably more complicated dynamics and require tools that observe broader symmetry.

BIOGRAPHICAL SKETCH

Lily Reeves was born in Beijing, China. She obtained her Bachelor of Science with a major in mathematics from New York University in 2017. In the same year, she entered the Ph.D. program at the Center for Applied Mathematics at Cornell University. Besides math, she is a classically trained pianist.

ACKNOWLEDGMENTS

I would like to first thank my advisor Phil Sosoe for his generosity in sharing his knowledge and time and for showing me the ethics of scholarship. He has guided me through the highs and lows of my doctoral experience with wisdom and has always encouraged me to pursue my interests. I am grateful to have Phil as my advisor.

I would also like to thank my mentor and collaborator Matt Junge, who generously brought me along so many amazing journeys, including but not limited to the three summers when I was involved with the Baruch College discrete math REU. Along this line, I thank all my collaborators and others who provided helpful comments, including anonymous reviewers.

I thank my special committee members Lionel Levine and Pierre Patie for their support and many conversations over the years.

I thank Chuck Newman and Dan Stein for introducing me to the world of probability and statistical physics when I was an undergraduate student at NYU and Reza Gheissari for his continued mentorship over the years. Reza, Swee Hong Chan, and Ahmed Bou-Rabee have been so generous in sharing their experience with me while I was applying for postdoctoral positions, for which I am deeply grateful.

I would like to thank my peers and friends that I have met through CAM, Cornell, and/or the probability community who are great sources of inspiration: Dongping, Tianyi, Zach, Seraphina, Linus, Weichen, and many more with whom I have had the pleasure interacting with.

As a pianist, I have tremendously benefited from the lessons I took in the Cornell Department of Music from my teachers Shin Hwang, Xak Bjerken, and Andrew Zhou, who are also great friends. I thank my friends Cheryl, Alexis,

Mike, and Ji Young, as well as my other studiomates, for many memories of going to concerts together and more.

I thank my family—my parents, my sister, and my uncle and aunt—for always having my back. Finally, I thank Thomas, whom I admire both as a peer and as a fellow musician, for being my best friend and partner in life.

TABLE OF CONTENTS

Biographical Sketch . . . iii
Acknowledgments . . . iv
Table of Contents . . . vi
List of Figures . . . viii
Preface . . . ix

I Chemical Distance in Percolation Models 5

1 Introduction 6
- 1.1 Bernoulli percolation and the random cluster model . . . 6
- 1.2 Chemical distance . . . 10
- 1.3 Overview of Part I . . . 14
- 1.4 Notation . . . 14
- 1.5 Properties and techniques . . . 19
  - 1.5.1 Gluing construction . . . 23

2 Radial chemical distance in 2d critical percolation clusters 27
- 2.1 Introduction . . . 27
  - 2.1.1 Organization . . . 28
- 2.2 A three-arm path to $\partial B(n)$ . . . 29
  - 2.2.1 Definition of γ . . . 29
  - 2.2.2 Estimate on C_0^c. . . . 31
  - 2.2.3 Estimate on C_0. . . . 32
- 2.3 Proof of Theorem 2.1.1 . . . 38
- 2.4 Improving on γ: connecting shortcuts around three-arm points . 41
  - 2.4.1 Estimate on C_0^c . . . 42
  - 2.4.2 Estimate on C_0 . . . 52

3 Chemical distance on the planar random cluster model 66
- 3.1 Introduction . . . 66
  - 3.1.1 Organization . . . 67
  - 3.1.2 Lower bound for the random cluster model . . . 67
- 3.2 Proof outline . . . 68
- 3.3 Large deviation bound conditional on three arms . . . 74
- 3.4 Arm separation for the random cluster model . . . 85
- 3.5 Estimating without Reimer's inequality . . . 95
- 3.A Extremal distance and resistance . . . 99

4 Equivalence of polychromatic arm probabilities on the square lattice 100
4.1 Introduction 100
4.1.1 Notation 102
4.1.2 Organization 103
4.2 A shifting transformation 103
4.3 Proof of Proposition 4.1.1 108

II Ballistic Annihilation 120

5 Introduction 121
5.1 Background 121
5.2 Overview of Part II 124
5.3 Notation 125
5.4 HST's argument and the mass transport principle 126

6 Three-velocity coalescing ballistic annihilation 129
6.1 Introduction 129
6.1.1 Proof methods 132
6.1.2 Determining firstness 134
6.1.3 Organization 136
6.2 Recursion 136
6.3 Properties of g 144
6.4 Regularity conditions 146
6.4.1 Superadditivity 147
6.4.2 A necessary and sufficient condition for blockade survival 152
6.5 Proof of Theorem 6.1.1 155

7 Arrivals are universal in coalescing ballistic annihilation 156
7.1 Introduction 156
7.2 Proof of Theorem 7.1.1 158

8 Non-universality in clustered ballistic annihilation 172
8.1 Introduction 172
8.1.1 Notation 173
8.1.2 Results 175
8.1.3 Discussion 176
8.1.4 Proof overview 178
8.2 Recursion 179
8.3 Continuity 183
8.4 Proof of Theorem 8.1.1 188

Bibliography 191

LIST OF FIGURES

1.1 Constructions near the endpoints x_1, x_2. 26

2.1 The construction on scale M on the event C_0^c, Case A. 44
2.2 The construction inside the box $\mathscr{B}'$ to obtain a defect edge. 47
2.3 An illustration of the half-plane event $A^{hp}_{2,OO}(p(e); 5M/4, 3n/4)$. 48
2.4 The construction in $B(n)$ on the event C_0^c in Case B. 51
2.5 The construction for the lower bound (2.30). 54
2.6 The construction for the lower bound (2.31). 55
2.7 The construction for the lower bound (2.39). 61
2.8 The construction for the lower bound (2.44). 64

3.1 The construction used in the proof of Lemma 3.4.1. 89
3.2 Crossings in a family of annuli. 92
3.3 Two quads defined by the same four endpoints. 98

4.1 An illustration of "inheritance" in the shifting transformation. 106
4.2 An illustration of the arm construction in the proof of Lemma 4.3.1. 111
4.3 The geometric configuration in the proof of disjointness of Lemma 4.3.1. 113

5.1 An example of BA on a space-time diagram. 122
5.2 An illustration of a measure-preserving reversal. 127

6.1 Plots of the formula for $q(p)$ for different parameter choices (a, b, α). 132

7.1 An illustration of the reversal argument for (7.4). 163
7.2 An illustration of the reversal argument for (7.5). 164
7.3 Examples of configurations in $\vec{\bullet}_1 \wedge (x_1 \overset{1}{\leftharpoondown} \overleftarrow{\bullet}_n)_{(x_1,x_n]}$ for which $\vec{\bullet}_1 - \overleftarrow{\bullet}_n$ fails to occur. 169

PREFACE

- Chapter 2 is based on the paper *An estimate for the radial chemical distance in 2d critical percolation clusters* [SR22], written in collaboration with Philippe Sosoe. This paper was published in *Stochastic Processes and their Applications* in 2022.
- Chapter 3 is based on the manuscript *Nontrivial upper bound for chemical distance on the planar random cluster model*, of which I am the sole author.
- Chapter 4 is based on the preprint *Equivalence of polychromatic arm probabilities on the square lattice* [RS20b], written in collaboration with Philippe Sosoe. At the time of writing, this paper was under review for publication.
- Chapter 6 is based on the paper *Three-velocity coalescing ballistic annihilation* [BJL+23], written in collaboration with Luis Benitez, Matthew Junge, Hanbaek Lyu, and Maximus Redman. This paper was published in the *Electronic Journal of Probability* in 2023.
- Chapter 7 is based on the preprint *Arrivals are universal in coalescing ballistic annihilation* [CPJR22], written in collaboration with Darío Cruzado Padró and Matthew Junge. At the time of writing, this paper was under review for publication.
- Chapter 8 is based on the preprint *Non-universality in clustered ballistic annihilation* [JOSMRRS22], written in collaboration with Matthew Junge, Arturo Ortiz San Miguel, and Cynthia Rivera Sánchez. At the time of writing, this paper was under review for publication.
- Chapters 1 and 5, the two introductory chapters, contain material taken from all the aforementioned papers.

INTRODUCTION

The present thesis consists of two independent parts, each concerned with a discrete probabilistic model.

In the first part, we consider Bernoulli percolation and its dependent generalization, the random cluster model, on the two-dimensional square lattice $\mathbb{Z}^2$. A percolation configuration is an assignment of either open or closed status on each edge of the graph according to some probabilistic measure. The open edges form a random graph and the connected components of open edges are called percolation clusters. In Bernoulli *bond* percolation, each edge is independently labeled open with probability p and closed with probability $1 - p$. Let $\theta(p)$ denote the probability that the origin is contained in an infinite open cluster. Kesten famously showed that for Bernoulli percolation on $\mathbb{Z}^2$, $\theta(p)$ exhibits a phase transition at $p_c = 1/2$, meaning that $\theta(p)$ is zero for $p < p_c$ (the subcritical regime) and strictly positive for $p > p_c$ (the supercritical regime)[Kes80]. Moreover, he showed that $\theta(p_c) = 0$ at criticality, establishing the continuity of the phase transition.

Percolation clusters induce an intrinsic graph distance, which we call the *chemical distance*. We focus on estimating the dimension of the chemical distance for Bernoulli percolation at criticality, where it is not well understood even in the physics literature, nor is there a widely accepted conjecture for the value of the dimension.

Aizenman and Burchard [AB99] gave the best known lower bound: with high probability, $\mathrm{dist}_{\mathrm{chem}}(0, x) \geq |x|^{1+\eta}$. In [DHS21], Damron, Hanson, and Sosoe gave the best known upper bound. Let $B(n)$ denote a Euclidean ball (or box in

$\mathbb{Z}^2$) of radius n centered at the origin. By bootstrapping improvements upon the lowest crossing, they showed that the chemical distance between the two vertical boundaries of $B(n)$ has expected size of the order $O(n^{2-\delta}\pi_3(n))$, where $\pi_3(n)$ is the "three-arm" probability. In Chapter 2, we follow up on this work by showing the same upper bound for the radial chemical distance, the chemical distance between the origin and $\partial B(n)$ (Theorem 2.1.1). In the radial setting, even though there is no natural "lowest" crossing with which to compare the chemical distance, we construct a path that satisfies the $n^2\pi_3(n)$ size profile in Section 2.2 as a starting point for the bootstrapping argument.

The random cluster model is a dependent generalization of Bernoulli percolation with two parameters $p \in (0, 1)$, the edge weight, and $q > 0$, the cluster weight. Through the Edwards–Sokal coupling [ES88], it is coupled to the Ising model and Potts models. In Chapter 3, we extend the results in [DHS21] and Chapter 2 and derive the same non-trivial upper bound $n^{2-\delta}\pi_3(q; n)$ for both the horizontal and radial chemical distances to the random cluster model with $q \in [1, 4]$ (Theorems 3.1.1 and 3.1.2).

In Chapter 4, we turn our attention back to critical Bernoulli percolation on the square lattice. Configurations in a k-arm event contain k disjoint paths, each of a single color, open or dual-closed, crossing an annulus. A k-arm event is said to be *polychromatic* if not all paths in the configuration are of the same color. It is well-known that for critical *site* percolation on the triangular lattice, probabilities of different polychromatic k-arm events are comparable. In Chapter 4, we extend this result to critical bond percolation on the square lattice (Proposition 4.1.1).

The subject of the second part of the thesis is ballistic annihilation (BA). In it, infinitely many particles with velocities sampled from a probability measure move across the real line and mutually annihilate upon contact. In the canonical example, the *symmetric three-velocity ballistic annihilation*, particles are either static (which we call blockades) with probability p or move with velocity ±1 each with probability $\frac{1-p}{2}$. The question of interest is whether the model exhibits a nontrivial phase transition in $\theta(p)$, the probability that a prespecified blockade is never annihilated.

Although there have been conjectures and heuristic arguments in the 1990s that the critical density for blockade survival is $p_c = \inf\{p : \theta(p) > 0\} = 1/4$, it was not until 2018, first in a preprint and later published in [HST21], that Haslegrave, Sidoravicius, and Tournier introduced an exactly solvable combinatorial approach that confirmed $p_c = 1/4$. Our interest is in extending the methodology of [HST21] to other symmetric variants of BA.

In Chapters 6 and 7, we consider coalescing systems where instead of mutual annihilation, collisions may generate new particles. For a symmetric family of such systems, we locate the phase transition and compute $\theta(p)$ (now depending on more parameters) algebraically (Theorem 6.1.1). We further demonstrate universality by showing that the index of the first particle to arrive at the origin does not depend on the law of the initial particle spacings (Theorem 7.1.1).

Another form of universality is invariance with respect to the initial particle density. In Chapter 8, we introduce a clustered system in which each blockade (sampled with probability p) is replaced by a cluster of X blockades, where X is sampled independently from a nonnegative discrete distribution. We derive a closed formula for p_c:

$$p_c = \frac{1}{(1 + \mathbb{E}[X])^2 + \mathrm{Var}(X)},$$

which implies that BA lacks universality in initial particle densities (Theorem 8.1.1).

Part I

Chemical Distance in Percolation Models

CHAPTER 1

INTRODUCTION

1.1 Bernoulli percolation and the random cluster model

Percolation is modeled on a graph $G = (V, E)$, either on its edge set E, which we refer to as bond percolation, or on its vertex set V, which we refer to as site percolation. A bond percolation configuration ω is in $\{0, 1\}^E$, intuitively understood as assigning open or closed status to each edge where 1 symbolizes open and 0 closed. The connected components of open edges induced by ω are referred to as open clusters. Site percolation is defined analogously.

In **Bernoulli bond percolation**, each edge is independently labeled open with probability p and closed with probability $1 - p$. Let $\theta(p)$ denote the probability that the origin is contained in an infinite open cluster. (Note that $\theta(p)$ is clearly monotone in p.) One of the most fundamental results that makes Bernoulli percolation interesting is that, for many graphs, it undergoes a *nontrivial phase transition* in $\theta(p)$. That is, there is a critical probability $p_c(G)$ strictly between 0 and 1 such that $\theta(p) = 0$ for $p < p_c$ (the subcritical regime) and $\theta(p) > 0$ for $p > p_c$ (the supercritical regime).

On the hypercubic lattice $\mathbb{Z}^d$, the probability that the origin is contained in an open cluster of size n is bounded above by $p^n(2d)^n$, which follows from a simple first-moment argument where we consider the expected number of self-avoiding paths of length n starting from the origin that consist of only open edges. This immediately implies $p_c \geq 1/(2d)$. On the other hand, 0 is not in an infinite cluster if and only if there is a simple closed curve surrounding the origin consisting of only closed edges. The *Peierls argument* [Pei36] employs this idea and shows that, for $d \geq 2$, $\theta(p) > 0$ when p is sufficiently close to 1.

Most of the early efforts in the subject focused on locating the critical probability for Bernoulli bond percolation on the two-dimensional square lattice $\mathbb{Z}^2$. It took 23 years since the introduction of percolation by Broadbent and Hammersley in 1957 for Kesten to resolve this problem [Kes80]:

Theorem (Kesten 1980)**.** *The critical probability for Bernoulli bond percolation on* $\mathbb{Z}^2$ *is* $p_c(\mathbb{Z}^2) = 1/2$*.*

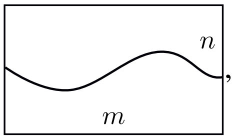
Winding back, in 1960, Harris provided the lower bound $p_c \geq 1/2$ by showing $\theta(\frac{1}{2}) = 0$ [Har60]. In 1978, Russo [Rus78] and independently Seymour–Welsh [SW78] developed what has become the Russo–Seymour–Welsh (RSW) theory, which we will discuss in more detail later in this section. RSW estimates turned out to be crucial building blocks for Kesten's final solution and one of the main tools for studying the critical regime. Moreover, Kesten and Harris's results together directly imply $\theta(p_c) = 0$ at criticality, establishing the *continuity of the phase transition* (also called a second-order phase transition in physics). In the subcritical regime, the probability that the origin is connected to a vertex at distance n through an open path decays exponentially fast in n [Kes80]. This is known as the sharpness of the phase transition.

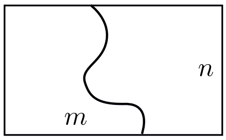
Establishing continuity of the phase transition, i.e., there is no infinite cluster at criticality, on $\mathbb{Z}^d$ for $d \geq 2$ became the next central question in the field. In high dimensions, via a powerful tool called lace expansion, it was shown by Hara and Slade in 1990 [HS90] for dimensions $d \geq 19$, and subsequently by Fitzner and van der Hofstad in 2017 [FvdH17a] for $d \geq 11$, that critical percolation on $\mathbb{Z}^d$ exhibits mean-field behavior, which would imply that $\theta(p_c(d)) = 0$. The mean field behavior is conjectured to hold for every $d \geq 6$ [HSvdH03, HvdH17]. For the intermediate dimensions $d = 3, 4$, and 5, continuity of the phase transition remains the largest open problem in the field.

Percolation on $\mathbb{Z}^2$ differs from higher dimensions in that it enjoys planar duality, which opens doors for combinatorial manipulations when studying crossing probabilities. For a percolation configuration ω, we define the dual configuration ω^* on the dual lattice $(\mathbb{Z}^2)^* = \mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$ by the relation

$$\omega^*(e^*) = \omega(e),$$

where e^* is the edge on the dual lattice that shares a midpoint with e. Then, the dual model to Bernoulli bond percolation with parameter p is another Bernoulli bond percolation with parameter $p^* = p$ but on the shifted square lattice $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$. An immediate consequence of duality is that an open path and a dual-closed path cannot intersect as the status of the edge (and the dual edge) where they intersect cannot be simultaneously open and closed. Let $\mathcal{H}(m, n)$ be the event that there is an open crossing between the left and right sides of the rectangle $[0, m] \times [0, n]$ and $\mathcal{V}(m, n)$ be the event that there is an open crossing between the top to bottom sides of the rectangle $[0, m] \times [0, n]$. Schematically,

$\mathcal{H}(m, n) =$  ,

$\mathcal{V}(m, n) =$  .

Let $\mathbb{P}_p$ be the Bernoulli percolation measure with parameter p on $\mathbb{Z}^2$. Then by duality, either there is an open horizontal crossing or there is a dual-closed vertical crossing:

$$\mathbb{P}_p(\mathcal{H}(m, n-1)) + \mathbb{P}_{1-p^*}(\mathcal{V}(m-1, n)) = 1$$

When $m = n$, we solve for $p = 1 - p^*$ such that $\mathbb{P}_p(\mathcal{H}(n, n-1)) = \mathbb{P}_{1-p^*}(\mathcal{V}(n-1, n))$. This is the intuitive reason why p_c should be 1/2. See Bollobás and Riordan's 2006 book [BR06] for a careful treatment of the Harris–Kesten theorem.

In the critical regime, Russo and independently Seymour–Welsh's results imply that the probability that there exists a horizontal crossing in a rectangle of a fixed aspect ratio is uniformly bounded away from 0 and 1:

Theorem (Russo–Seymour–Welsh 1978)**.** *Let* $0 < \alpha < \infty$. *Then there exists* $c_\alpha > 0$ *(independent of* n*) such that*

$$c_\alpha \leq \mathbb{P}_{p_c = \frac{1}{2}}(\mathcal{H}(\alpha n, n)) \leq 1 - c_\alpha.$$

The **random cluster model**, also called Fortuin–Kasteleyn (FK) percolation, is a *dependent* bond percolation model introduced by Fortuin and Kasteleyn [FK72] as a generalization of Bernoulli percolation. The random cluster measure on a percolation configuration is tuned by two parameters: $p \in (0, 1)$, the edge weight, and $q > 0$, the cluster weight. Then, the random cluster measure is proportional to

$$p^{\#\text{ open edges}}(1-p)^{\#\text{ closed edges}} q^{\#\text{ open clusters}}.$$

This measure yields Bernoulli percolation when $q = 1$. Through the Edwards–Sokal coupling [ES88], it corresponds to the Ising model when $q = 2$ and the q-state Potts model for integer $q \geq 3$. Sampling a random cluster configuration is equivalent to first sampling a Bernoulli percolation configuration with parameter p and then uniformly sampling a color out of q colors for each open cluster. On a finite graph, we obtain the q-state Potts model with inverse temperature $\beta = -\frac{1}{2}\log(1-p)$ as follows. For each cluster, we sample a state k uniformly from the q states and independently of other clusters. Then, we assign all spins in this cluster state k.

The duality relation is generalized as follows: if ω is a configuration sampled from a random cluster measure with parameters p and q, then its dual configuration ω^* is sampled from a random cluster measure with parameters p^* and q that satisfies

$$\frac{pp^*}{(1-p)(1-p^*)} = q.$$

This relation ultimately results in the critical edge weight being $p_c(q) = p_c(q)^* = \sqrt{q}/(1+\sqrt{q})$ for $q \geq 1$ [BDC12]. The random cluster model is positively associated when $q \geq 1$ and negatively associated when $q < 1$. The phase transition is continuous (i.e. at criticality, there is almost surely no infinite open cluster) for $1 \leq q \leq 4$ [DCST17] and discontinuous for $q > 4$ [DCGH⁺21, RS20a]. Duminil-Copin, Manolescu, and Tassion established RSW-type quad-crossing estimates in [DCMT21]. These results are crucial for applying classical Bernoulli percolation techniques despite the lack of independence.

1.2 Chemical distance

Although almost surely 2d critical percolation does not contain an infinite cluster, there are arbitrarily large clusters which, in some suitable limit, yield the incipient infinite cluster (IIC) [Kes86]. Kesten studied the simple random walk on the IIC in $\mathbb{Z}^2$ and showed that the displacement is *subdiffusive*, i.e., slower than a corresponding walk on all of $\mathbb{Z}^2$. Alexander and Orbach [AO82] famously conjectured that for general dimensions $\mathbb{Z}^d$, on the IIC, the $2n$-step return probability for a simple random walk is on the order of $n^{-4/3}$. This conjecture was confirmed for high dimensions by Kozma and Nachmias [KN09] by controlling the volume growth and effective resistance of the IIC. The *heat kernel estimates* used in their work come from understanding the intrinsic graph distance on percolation clusters, which we call the **chemical distance**. In high-dimensional critical percolation, the chemical distance between 0 and x is of order $|x|^2$ [FvdH17b]. All of the above results on high-dimensional critical percolation ultimately rely on the lace expansion technique to obtain the mean-field two-point function bound $|x-y|^{-d+2}$.

For two subsets A and B of V and a percolation configuration viewed as a subgraph of G, the chemical distance between A and B is the least number of edges in any path of open edges connecting A to B. We denote this distance by $\mathrm{dist}_{\mathrm{chem}}(A,B)$.

The chemical distance for both sub- and supercritical percolation is known to be linear, i.e., equivalent to the Euclidean distance, for general dimensions [GM90, AP96]. In the subcritical regime, linearity follows from the exponential decay of cluster volume, by Aizenman and Newman [AN84]. Roughly speaking, conditional on $\{x \leftrightarrow y\}$, that is, x and y are connected by a path of open edges, the probability $\mathrm{dist}_{\mathrm{chem}}(x,y)$ is superlinear is exponentially small. On the other hand, Antal and Pisztora [AP96] showed that when $p > p_c$, there is a constant ρ that depends on p and the dimension such that

$$\limsup_{|x|\to\infty} \frac{1}{|x|}\mathrm{dist}_{\mathrm{chem}}(0,x)\mathbf{1}\{0 \leftrightarrow x\} \le \rho(p,d)$$

almost surely. This implies that the chemical distance is at most linear.

It is known from work of Aizenman and Burchard [AB99] that, unlike in the sub- and supercritical cases, the chemical distance for critical percolation is super-linear:

Theorem (Aizenman–Burchard 1999)**.** *If A and B are at Euclidean distance greater than or equal to n, then there is $\eta > 0$ such that, with high probability*

$$\mathrm{dist}_{\mathrm{chem}}(A, B) \geq n^{1+\eta}. \tag{1.1}$$

In high dimensions, that is, rigorously for $d \geq 11$, the chemical distance is quadratic [FvdH17b].

The chemical distance in the critical phase is not well-understood. From the physics literature [EK85, Gra99, HN84, HTWBA85, HHS84, HS88, ZYDZ12], it is expected that there exists an exponent $s > 1$ such that if A and B are at Euclidean distance n, then:

$$\mathbb{E}[\mathrm{dist}_{\mathrm{chem}}(A, B) \mid A \leftrightarrow B] \approx n^s. \tag{1.2}$$

Here, $\mathbb{E}$ is the expectation with respect to the critical percolation measure $\mathbb{P}_{1/2}$. Despite remarkable progress on the derivation of other critical exponents, no approximation of the form (1.2) is known for any reasonable interpretation of $\approx$. See Schramm's survey [Sch11] for a list of problems in percolation, where the determination of the exponent s is listed as an important open problem. In particular, no clear connection has yet been discovered to the SLE process, which is used to derive other critical exponents in two-dimensional percolation. It is not at all clear how to relate the chemical distance, measured in lattice spacings, to any conformally invariant quantities.

Presently, no lower bound other than the one in (1.1) is known and the η in (1.1) is non-explicit. We remark that Aizenman and Burchard's result applies to a general family of random curves satisfying certain conditions. This includes shortest open connections in the random cluster model. We remark further on this lower bound in Section 3.1.2.

In [KZ93], Kesten and Zhang noted that if one restricts attention to paths inside a square box $B(n) = [-n, n]^2$ on $\mathbb{Z}^2$, and lets $A = \{-n\} \times [-n, n]$ and $B = \{n\} \times [-n, n]$ be the two vertical sides of $B(n)$, one obtains an upper bound by considering the *lowest crossing* ℓ_n of $B(n)$:

$$\mathbb{E}[\mathrm{dist}_{\mathrm{chem}}(A, B) \mid A \leftrightarrow B] \le \mathbb{E}[\#\ell_n \mid A \leftrightarrow B].$$

By now standard computations, the expected size of the lowest crossing can be expressed in terms of the *three-arm probability* to distance n:

$$\mathbb{E}[\#\ell_n \mid A \leftrightarrow B] \le Cn^2\pi_3(n), \tag{1.3}$$

see for example Morrow–Zhang [MZ05]. Here $\pi_3(n) = \mathbb{P}(A_3(n))$ is the probability that there are two open and one dual-closed connections from the origin to the boundary of the box $B(n)$. On the triangular lattice, the corresponding probability is computed in [SW01]:

$$\pi_3(n) = n^{-\frac{2}{3}+o(1)}.$$

In [DHS17, DHS21], Damron, Hanson, and Sosoe answered a question of Kesten–Zhang, showing that the upper bound (1.3) can be improved by a factor $n^{-\delta}$, for some $\delta > 0$:

Theorem 1.2.1 (Damron–Hanson–Sosoe 2021)**.**

$$\mathbb{E}[\mathrm{dist}_c(A, B) \mid A \leftrightarrow B] \le Cn^{-\delta}\mathbb{E}[\#\ell_n \mid A \leftrightarrow B]. \tag{1.4}$$

Simulation results in [ZYDZ12] suggest that the critical exponent s in (1.2) for the chemical distance is approximately 1.1308. The value of $\pi_3(n)$ on the square lattice is expected to be on the order of $n^{-2/3}$, as on the triangular lattice, although this has not yet been proved. Thus, we expect the optimal value of δ to be approximately 0.2. This is out of reach of current methods.

1.3 Overview of Part I

In Chapter 2, we derive an estimate for the chemical distance from the origin to the boundary of the box of side length $2n$, conditioned on the existence of an open connection. The estimate we obtain is the radial analogue of (1.4). In the present case, however, there is no lowest crossing in the box to compare to, so we construct a path γ from the origin to distance n that consists of "three-arm" points, and whose volume can thus be estimated by $O(n^2\pi_3(n))$. We then develop estimates for the existence of shortcuts around an edge e in the box, conditional on $\{e \in \gamma\}$, to obtain a bound of the form $O(n^{2-\delta}\pi_3(n))$ for some $\delta > 0$.

In Chapter 3, we extend the upper bounds derived for the horizontal and radial chemical distance for 2d Bernoulli percolation in [DHS21] and Chapter 2 to the planar random cluster model with cluster weight $1 \leq q \leq 4$. Along the way, we provide a complete proof of the strong arm separation lemma for the random cluster model.

In Chapter 4, we again consider 2d critical Bernoulli percolation on the square lattice. We prove an approximate color-switching lemma comparing k-arm probabilities for different polychromatic color sequences. This result is well-known for site percolation on the triangular lattice in [LSW02]. To handle the complications arising from the dual lattice, we introduce a shifting transformation to convert arms between the primal and dual lattices.

1.4 Notation

In this section, we summarize the notations we will use.

In Chapter 2 and 4, we consider Bernoulli percolation on the square lattice $\mathbb{Z}^2$ seen as a graph with the edge set $\mathcal{E}$ consisting of all pairs of nearest-neighbor vertices. We let $\mathbb{P}$ be the critical bond percolation measure

$$\mathbb{P} = \prod_{e \in \mathcal{E}} \frac{1}{2}(\delta_0 + \delta_1)$$

on the state space $\Omega = \{0, 1\}^{\mathcal{E}}$, with the product σ-algebra. An edge e is said to be *open* in the configuration $\omega \in \Omega$ if $\omega(e) = 1$ and *closed* otherwise. The dual lattice is written $((\mathbb{Z}^2)^*, \mathcal{E}^*)$, where

$$(\mathbb{Z}^2)^* = \mathbb{Z}^2 + \left(\frac{1}{2}, \frac{1}{2}\right)$$

and $\mathcal{E}^*$ its nearest-neighbor edges. Given $\omega \in \Omega$, we define $\omega^* \in \Omega^* = \{0, 1\}^{\mathcal{E}^*}$ by the relation $\omega^*(e^*) = \omega(e)$, where e^* is the dual edge that shares a midpoint with e. The edge e^* is said to be *dual-closed* if $\omega(e) = 0$. For any $V \subset \mathbb{R}^2$ we write

$$V^* = V + \left(\frac{1}{2}, \frac{1}{2}\right).$$

In Chapter 3, we work with the random cluster model on a finite subdomain of $(\mathbb{Z}^2, \mathcal{E})$. A finite subdomain $\mathcal{D} = (V, E)$ is defined by the (finite) edge set E and the vertex set V of all endpoints of the edges in E. Its boundary $\partial\mathcal{D}$ consists of the vertices in the topological boundary of $\mathcal{D}$.

A boundary condition ξ on $\mathcal{D}$ is a partition of $\partial\mathcal{D}$. All vertices in the same class of the partition are *wired together* and count towards the same connected component when defining the probability measure. In the *free* boundary condition, denoted by 0 in the superscript, no two vertices on the boundary are identified with each other.

Let $\mathcal{D} = (V, E)$ be a subdomain of $\mathbb{Z}^2$. For edge weight $p \in [0, 1]$ and cluster weight $q > 0$, the random cluster measure on $\mathcal{D}$ with boundary condition ξ is defined by

$$\phi^{\xi}_{p,q,\mathcal{D}}(\omega) = \frac{1}{Z^{\xi}_{p,q,\mathcal{D}}} p^{o(\omega)}(1-p)^{c(\omega)} q^{k(\omega^{\xi})}$$

where $o(\omega)$ is the number of open edges in ω, $c(\omega) = |E| - o(\omega)$ is the number of closed edges in ω, $k(\omega^{\xi})$ is the number of connected components of ω with consideration of the boundary condition ξ, and the partition function is defined by

$$Z^{\xi}_{p,q,\mathcal{D}} = \sum_{\omega \in \{0,1\}^E} p^{o(\omega)}(1-p)^{c(\omega)} q^{k(\omega^{\xi})}.$$

Dual configurations are defined analogously to those in Bernoulli percolation. The dual measure to the random cluster measure is of the form

$$\phi^{\xi}_{p^*,q,\mathcal{D}^*}(\omega^*) \propto (p^*)^{o(\omega^*)}(1-p^*)^{c(\omega^*)} q^{k((\omega^*)^{\xi})}$$

where the dual parameter p^* satisfies

$$\frac{pp^*}{(1-p)(1-p^*)} = q.$$

We use $\mathbb{E}$ to denote the expectation with respect to either the critical Bernoulli percolation measure $\mathbb{P}$ or the random cluster measure ϕ. The meaning of $\mathbb{E}$ is clear from the context.

A (lattice) *path* is a sequence $(v_0, e_1, v_1, \ldots, v_{N-1}, e_N, v_N)$ such that for all $k = 1, \ldots, N$, $\|v_{k-1} - v_k\|_1 = 1$ and $e_k = \{v_{k-1}, v_k\}$. A *circuit* is a path with $v_0 = v_N$. Given $\omega \in \Omega$, we say that $\gamma = (e_k)_{k=1,\ldots,N}$ is open in ω if $\omega(e_k) = 1$ for $k = 1, \ldots, N$.

We write $B(n)$ for the domain generated by the edges in $[-n,n]^2$ and $B(x,n)$ its translation by $x \in \mathbb{Z}^2$. When x is the origin $(0,0)$, we sometimes abbreviate $B((0,0),n)$ by $B(n)$. We sometimes abuse notation and write $B(e,n)$ for an edge e to mean the box $B(e_x,n)$ where e_x is the lower-left endpoint of e, defined as the first of the two endpoints of e in the lexicographic order on $\mathbb{Z}^2$. We denote by $\partial B(x, n)$ the set

$$\partial B(x, n) = \{(x_1, x_2) \in \mathcal{E} : x_1 \sim x_2, |x_1 - x|_\infty = n, |x_2 - x|_\infty = n\}.$$

$x \sim y$ means x and y are nearest neighbors on $\mathbb{Z}^2$.

For $n_1 \leq n_2$, we define an annulus centered at $x \in \mathbb{Z}^2$ as the difference between two boxes of different sizes centered at x:

$$\mathrm{Ann}(x; n_1, n_2) = B(x, n_2) \setminus B(x, n_1).$$

If the annulus is centered at the origin, we drop the x from the notation and instead write $\mathrm{Ann}(n_1, n_2)$.

Distances in Part I are measured in the ℓ_∞ norm

$$\|x - y\|_\infty = \max_{i=1,2} |x_i - y_i|, \quad \text{where } x, y \in \mathbb{Z}^2.$$

Arm events and arm exponents

A path of consecutive open or dual-closed edges is called an arm. We use O and C to represent open and resp. dual-closed arms. A *color sequence* σ of length k is a sequence $(\sigma_1, \ldots, \sigma_k) \in \{O, C\}^k$. Each σ_i indicates a "color", with O representing open and C representing dual-closed.

For $n_1 \leq n_2$ and a vertex x, we define a *k-arm event* with color sequence σ to be the event that there are k disjoint paths whose colors are specified by σ in the annulus $\mathrm{Ann}(x; n_1, n_2)$ connecting $\partial B(x, n_1)$ and $\partial B(x, n_2)$. Formally,

$$A_{k,\sigma}(x; n_1, n_2) := \left\{\partial B(x, n_1) \xleftrightarrow[\sigma]{\mathrm{Ann}(x; n_1, n_2)} \partial B(x, n_2)\right\}.$$

We write $A \overset{\mathcal{D}}{\underset{\sigma}{\leftrightarrow}} B$ to denote that vertex sets A and B are connected through a path of color σ in the domain $\mathcal{D}$.

We note a technical point: for $A_{k,\sigma}(x; n_1, n_2)$ to be defined, n_1 needs to be big enough for all k arms to be (vertex)-disjoint. We define $n_0(k)$ to be the smallest integer such that $|\partial B(n_0(k))| \geq k$. Color sequences that are equivalent up to cyclic order denote the same arm event.

In Section 2.4, we use *half-plane* versions of the arm events above. The half-plane event $A^{hp}_{k,\sigma}(n_1, n_2)$ is the event that $A_{k,\sigma}(n_1, n_2)$ occurs and all arms are contained in a half-plane $\{(x, y) \in \mathbb{R}^2 : v \cdot (x, y) \leq 0\}$ for some unit vector $v \in \{(0, 1), (1, 0)\}$.

We use special notations for the probabilities of certain arm events that occur frequently. We describe these probabilities in terms of the random cluster measure with some fixed cluster weight q, although some of these notations are only used in the context of Bernoulli percolation when $q = 1$. Let us fix n and the boundary condition ξ. Let $0 \leq n_1 < n_2 \leq n$.

Two-arm event We denote by $\pi_2(n_1, n_2)$ the two-arm probability for two arms, one open and one dual-closed:

$$\pi_2(n_1, n_2) := \phi^{\xi}_{B(n)}(A_{2,OC}(n_1, n_2));$$

Three-arm event We denote by $\pi_3(n_1, n_2)$ the probability for the three-arm event $A_3(n_1, n_2) = A_{3,OOC}(n_1, n_2)$ that there are two open arms and one dual-closed arm in the annulus $\mathrm{Ann}(n_1, n_2)$:

$$\pi_3(n_1, n_2) := \phi^{\xi}_{B(n)}(A_3(n_1, n_2)).$$

Alternating five-arm event There exists $c, C > 0$ such that

$$c(n_1/n_2)^2 \leq \phi^{\xi}_{B(n)}(A_{5,OCOCO}(n_1, n_2)) \leq C(n_1/n_2)^2.$$

Thus, the alternating five-arm event is said to have the universal arm exponent 2. For proof, see [DCMT21, Proposition 6.6].

Mono- and polychromatic arm events For $k \geq 3$, we denote by $\pi_k(n, N)$ the probability that there are $k-1$ disjoint open and one dual-closed arm joining the boundaries of the annulus. Monochromatic k-arm probabilities are denoted by π'_k:

$$\pi'_k(n_1, n_2) := \phi^{\xi}_{B(n)}(A_{k,O\cdots O}(n_1, n_2)) = \phi^{\xi}_{B(n)}(A_{k,C\cdots C}(n_1, n_2)).$$

We remark that the dependencies on the cluster weight q are implicit in the notations above.

Generally, we reserve the letter A for arm events, the letter B for boxes, and the capital letter C (with various fonts) for circuits and events related to circuits, for "closed" when used as a subscript, and at times for various constants, along with the small letter c. All other notations will be specified as needed.

1.5 Properties and techniques

Recall that the random cluster model undergoes a continuous phase transition when $q \in [1, 4]$. This is the scope of Chapter 3. Thus, unless stated otherwise, we state all the following properties and results in terms of the random cluster measure with $q \in [1, 4]$ and they automatically apply to Bernoulli bond percolation when $q = 1$. Elaboration on the following properties can be found in [Gri06] and [DC17] or as cited. Let us fix $1 \leq q < 4$, $p = p_c(q)$, a domain $\mathcal{D} = (V, E)$ of $\mathbb{Z}^2$, and some boundary condition ξ on $\partial\mathcal{D}$.

Domain Markov property

For any configuration $\omega' \in \{0,1\}^E$ and any subdomain $\mathcal{F} = (W, F)$ with $F \subset E$,

$$\phi^{\xi}_{\mathcal{D}}(\cdot_{|F} \mid \omega_e = \omega'_e, \forall e \notin F) = \phi^{\xi'}_{\mathcal{F}}(\cdot) \tag{DMP}$$

where the boundary conditions ξ' on $\mathcal{F}$ are defined as follows: x and y on $\partial\mathcal{F}$ are wired if they are connected in $\omega^{\xi}_{|E\setminus F}$.

Quad-crossing RSW

A discrete *quad* $(\mathcal{D}, a, b, c, d)$ is a discrete domain $\mathcal{D}$ with four vertices a, b, c, d on $\partial\mathcal{D}$ in counterclockwise order. The four vertices define four arcs $(ab), (bc), (cd)$, and (da), which make up the boundary $\partial\mathcal{D}$. The discrete domain $\mathcal{D}$ can be seen as a continuous domain of the plane.

The extremal distance $\ell_{\mathcal{D}}[(ab), (cd)]$ between (ab) and (cd) in $\mathcal{D}$ is is defined as the unique $\ell > 0$ such that there is a conformal map from $\mathcal{D}$ seen as a continuous domain to the rectangle $[0,1] \times [0,\ell]$, a, b, c, d being mapped to the four corners of $[0,1] \times [0,\ell]$ in counterclockwise order, starting from the lower-left corner. In Section 3.A, we provide an alternate definition for extremal distances in terms of the resistance of an electrical network.

Finally, $(\mathcal{D}, a, b, c, d)$ is said to be crossed from (ab) to (cd), which we denote by $(ab) \overset{\mathcal{D}}{\longleftrightarrow} (cd)$, if there is an open connection linking (ab) and (cd). Then, the crossing of a rectangle is clearly a special case of the crossing of a quad.

Theorem (Duminil-Copin–Manolescu–Tassion 2021). *For every $M > 0$, there exists $\eta = \eta(M) \in (0,1)$ such that for any discrete quad $(\mathcal{D}, a, b, c, d)$ and any boundary conditions ξ, if the extremal distance $\ell_{\mathcal{D}}[(ab),(cd)] \in [M^{-1}, M]$, then*

$$\eta \le \phi^{\xi}_{\mathcal{D}}[(ab) \overset{\mathcal{D}}{\longleftrightarrow} (cd)] \le 1 - \eta. \tag{RSW}$$

Fortuin–Kasteleyn–Ginibre (FKG) inequality

An event A is called increasing if for any $\omega \le \omega'$ (for the partial order on $\{0,1\}^E$), $\omega \in A$ implies that $\omega' \in A$. Similarly, A is decreasing if for any $\omega \le \omega'$, $\omega' \in A$ implies $\omega \in A$. For every pair of increasing events A and B,

$$\phi^{\xi}_{\mathcal{D}}(A \cap B) \ge \phi^{\xi}_{\mathcal{D}}(A)\phi^{\xi}_{\mathcal{D}}(B). \tag{FKG}$$

One technique that we will repeatedly use in Part I is the gluing construction. For Bernoulli percolation, this depends on the *generalized* FKG inequality combined with RSW estimates. It has become a standard technique since Kesten's work on critical percolation in the 80s [Kes86, Kes87]. We state the generalized FKG inequality as in [Nol08].

Theorem (Generalized FKG). *Consider two increasing events A^+, $\tilde{A}^+$, and two decreasing events A^-, $\tilde{A}^-$. Assume that there exist three disjoint finite sets of vertices $\mathcal{A}$, $\mathcal{A}^+$, and $\mathcal{A}^-$ such that A^+, A^-, $\tilde{A}^+$, and $\tilde{A}^-$ depend only on the sites in, respectively, $\mathcal{A} \cup \mathcal{A}^+$, $\mathcal{A} \cup \mathcal{A}^-$, $\mathcal{A}^+$, and $\mathcal{A}^-$. Then we have*

$$\hat{\mathbb{P}}(\tilde{A}^+ \cap \tilde{A}^- \mid A^+ \cap A^-) \ge \hat{\mathbb{P}}(\tilde{A}^+)\hat{\mathbb{P}}(\tilde{A}^-)$$

for any product measure $\hat{\mathbb{P}}$ on Ω.

There is no known proof for the equivalent of the generalized FKG inequality for the random cluster model. Therefore, we provide a proof for the gluing construction for the random cluster model without using generalized FKG inequality in Section 1.5.1.

Quasi-multiplicativity

Theorem (Duminil-Copin–Manolescu–Tassion 2021)**.** *Fix σ. There exist $c = c(\sigma, q) > 0$ and $C = C(\sigma, q) > 0$ such that for any boundary condition ξ and every $n_0(k) \leq n_1 \leq n_3 \leq n_2$,*

$$c\phi^{\xi}_{\mathcal{D}}(A_\sigma(n_1, n_2)) \leq \phi^{\xi}_{\mathcal{D}}(A_\sigma(n_1, n_3))\phi^{\xi}_{\mathcal{D}}(A_\sigma(n_3, n_2)) \leq C\phi^{\xi}_{\mathcal{D}}(A_\sigma(n_1, n_2)). \tag{QM}$$

As an immediate consequence of (QM) and (RSW), we have that there are constants $c, C > 0$ such that, uniformly in n,

$$c\phi^{\xi}_{\mathcal{D}}(A_\sigma(n/2, N)) \leq \phi^{\xi}_{\mathcal{D}}(A_\sigma(n, N)) \leq C\phi^{\xi}_{\mathcal{D}}(A_\sigma(n, 2N)). \tag{1.5}$$

Reimer's inequality

Let $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ be a probability space equipped with the product measure $\mathbf{P}$, where each Ω_i is a finite probability space. For two events A and B in Ω, their *disjoint occurrence*, denoted by $A \Box B$, is the event that their occurrences are observed on disjoint indices in $[n]$. Formally, if a configuration $\omega \in A \Box B$, then there exist subsets $I, J \subset [n]$ such that

- $I \cap J = \varnothing$,
- if $(\omega')_i = \omega_i$ for all $i \in I$, then $\omega' \in A$, and
- similarly, if $(\omega'')_j = \omega_j$ for all $j \in J$, then $\omega'' \in B$.

For Bernoulli percolation with parameter p, the van den Berg–Kesten/Reimer's inequality states: For any two events A and B,

$$\mathbf{P}(A \square B) \le \mathbf{P}(A)\mathbf{P}(B). \tag{Reimer}$$

In Bernoulli (bond) percolation, the Ω_i's are indexed by the edges. A classic application of Reimer's inequality is for arm events: The polychromatic two-arm event implies the disjoint occurrence of an open arm and a dual-closed arm. Thus,

$$\pi_2(n_1, n_2) \le \mathbb{P}(A_{1,O}(n_1, n_2) \square A_{1,C}(n_1, n_2)) \le \mathbb{P}(A_{1,O}(n_1, n_2))\mathbb{P}(A_{1,C}(n_1, n_2)).$$

Despite being a classical tool for Bernoulli percolation, Reimer's inequality is not known in the general form for the random cluster model, nor do we expect it to be true. A weak form of Reimer's inequality for the random cluster model is shown in [vdBG13]. We will further discuss this in Section 3.5.

1.5.1 Gluing construction

As discussed previously, the classic proof of the gluing construction for Bernoulli percolation uses the generalized FKG inequality, which is inaccessible for the random cluster model. See an exposition in [SR22, Section 1.4]. This section is dedicated to carefully examining the gluing construction for the random cluster model without using generalized FKG. The notations used in this section are independent of the rest of the document.

Fix $\delta > 0$, a positive integer k, and $n_1 < n_2 < n_3$ sufficiently large. Let $E(n_1, n_2)$ be the event such that:

1. there exist vertices $x_i \in \partial B(n_2)$ for $i = 1, \ldots, k$ and $\min_{i \neq j} |x_i - x_j| \geq 10\delta n_2$ such that x_i is connected to $\partial B(x_i, 2\delta n_2) \cap B(n_2)^c$ by a path of color σ_i;
2. $E(n_1, n_2)$ depends only on the status of the edges in $\mathrm{Ann}(n_1, n_2) \cup (\cup_{i=1}^k B(x_i, 2\delta n_2))$.

Similarly, let $F(2n_2, n_3)$ be the event such that:

1. there exist vertices $y_i \in \partial B(2n_2)$ for $i = 1, \ldots, k$ and $\min_{i \neq j} |y_i - y_j| \geq 20\delta n_2$ such that y_i is connected to $\partial B(y_i, 2\delta n_2) \cap B(2n_2)$ by a path of color σ_i;
2. $F(2n_2, n_3)$ depends only on the status of the edges in $\mathrm{Ann}(2n_2, n_3) \cup (\cup_{i=1}^k B(y_i, 2\delta n_2))$.

Proposition 1.5.1. *Let $E(n_1, n_2)$, $F(2n_2, n_3)$ be as the above. Then there exists $c > 0$ depending only on k, such that*

$$\begin{aligned} &\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(E(n_1, n_2) \cap F(2n_2, n_3) \cap \bigcap_{i=1}^{k} \{x_i \xleftrightarrow[\sigma_i]{\mathrm{Ann}(n_2,2n_2)} y_i\}\right) \\ &\qquad \geq c\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}(E(n_1, n_2) \cap F(2n_2, n_3)). \end{aligned} \tag{1.6}$$

Proof. Conditional on $E(n_1, n_2) \cap F(2n_2, n_3)$, we construct a set of k corridors, $T_1, \ldots, T_k$, each connecting $B(x_i, 2\delta n_2) \cap B(n_2)^c$ to $B(y_i, 2\delta n_2) \cap B(2n_2)$. Let $\gamma_1, \ldots, \gamma_k$ be a collection of (topological) paths that satisfy the following constraints:

- γ_i is a path in $\mathrm{Ann}(n_2, 2n_2)$ from x_i to y_i.
- The distance between any two γ_i, γ_j is at least $10\delta n_2$.
- The length of each γ_i is at most Cn_2 for some constant C.

Then, we let T_i be the δn_2 neighborhood of γ_i intersected with $\mathrm{Ann}(n_2, 2n_2)$. The T_i's are disjoint by construction. We show (1.6) by first dividing the right-hand side on both sides, converting the left-hand side into a conditional probability, and noting that

$$\begin{aligned}&\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(\cap_{i=1}^{k}\{x_i \xleftrightarrow[\sigma_i]{\mathrm{Ann}(n_2,2n_2)} y_i\} \,\middle|\, E(n_1,n_2)\cap F(2n_2,n_3)\right)\\ &\qquad\geq\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(\cap_{i=1}^{k}\{x_i \xleftrightarrow[\sigma_i]{T_i} y_i\} \,\middle|\, E(n_1,n_2)\cap F(2n_2,n_3)\right).\end{aligned}$$

It suffices to provide a constant lower bound for the right-hand side. We first use the tower rule for conditional expectations to isolate the occurrence of $\{x_1 \xleftrightarrow[\sigma_1]{T_1} y_1\}$.

$$\begin{aligned}&\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(\cap_{i=1}^{k}\{x_i \xleftrightarrow[\sigma_i]{T_i} y_i\} \,\middle|\, E(n_1,n_2)\cap F(2n_2,n_3)\right) &&(1.7)\\ &\quad=\mathbf{E}\left[\mathbf{E}\left[\mathbf{1}\left\{\cap_{i=1}^{k}\{x_i \xleftrightarrow[\sigma_i]{T_i} y_i\}\right\} \,\middle|\, \omega|_{T_1^c}, E(n_1,n_2)\cap F(2n_2,n_3)\right] \,\middle|\, E(n_1,n_2)\cap F(2n_2,n_3)\right].\end{aligned}$$

Here $\mathbf{E}$ denotes the expectation with respect to the measure $\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}$. Since $\cap_{i=2}^{k}\{x_i \xleftarrow[\sigma_i]{T_i} y_i\}$ is $\omega|_{T_1^c}$-measurable, the right-hand side can be rewritten as

$$\mathbf{E}\left[\mathbf{E}\left[\mathbf{1}\{x_1 \xleftrightarrow[\sigma_1]{T_1} y_1\} \,\middle|\, \omega|_{T_1^c}, E(n_1,n_2)\cap F(2n_2,n_3)\right]\cdot\mathbf{1}\left\{\cap_{i=2}^{k}\{x_i \xleftrightarrow[\sigma_i]{T_i} y_i\}\right\} \,\middle|\, E(n_1,n_2)\cap F(2n_2,n_3)\right]. \qquad (1.8)$$

We write the inner conditional expectation in (1.8) back in conditional probability form as $\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(x_1 \xleftrightarrow[\sigma_1]{T_1} y_1 \,\middle|\, \omega|_{T_1^c}, E(n_1,n_2)\cap F(2n_2,n_3)\right)$. Note that $\{x_1 \xleftrightarrow[\sigma_1]{T_1} y_1\}$ occurs if the following events occur simultaneously:

- $\{x_1 \xleftrightarrow[\sigma_1]{B(2\delta n_2)(x_1)\cap B(n_2)^c} \partial B(x_1, 2\delta n_2)\cap B(n_2)^c\}$;
- $\{y_1 \xleftrightarrow[\sigma_1]{B(2\delta n_2)(y_1)\cap B(2n_2)} \partial B(y_1, 2\delta n_2)\cap B(2n_2)\}$;
- there is a σ_1-path in T_1 connecting the two short sides of T_1;
- there is a half σ_1-circuit enclosing x_1 in the half annulus $\mathrm{Ann}(x_1;\delta n_2, 2\delta n_2)\cap B(n_2)^c$, the event of which we denote by C_1; and
- there is a half σ_1-circuit enclosing y_1 in the half annulus $\mathrm{Ann}(y_1;\delta n_2, 2\delta n_2)\cap B(2n_2)$, the event of which we denote by C_2.

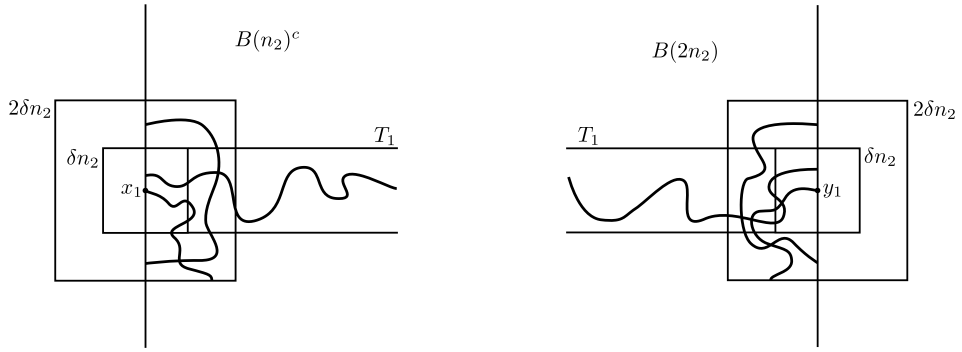


Figure 1.1: Constructions near the endpoints x_1, x_2.

See Figure 1.1. Since all these events are σ_1 connection events, (FKG) applies. To simplify notation, we denote by $\varphi(\cdot)$ the conditional measure $\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}(\cdot \mid \omega|_{T_1^c}, E(n_1,n_2) \cap F(2n_2,n_3))$. Then,

$$\begin{aligned}\varphi\left(x_1 \overset{T_1}{\underset{\sigma_1}{\longleftrightarrow}} y_1\right) \geq & \varphi\left(\partial B(n_2) \cap T_1 \overset{T_1}{\underset{\sigma_1}{\longleftrightarrow}} \partial B(2n_2) \cap T_1\right)\varphi(C_1)\varphi(C_2) \\ & \cdot \varphi\left(x_1 \xleftrightarrow[\sigma_1]{B(x_1,\delta n_2)\cap B(n_2)^c} \partial B(x_1,\delta n_2) \cap B(n_2)^c\right) \\ & \cdot \varphi\left(y_1 \xleftrightarrow[\sigma_1]{B(y_1,2\delta n_2)\cap B(2n_2)} \partial B(y_1,2\delta n_2) \cap B(2n_2)\right).\end{aligned}$$

The two probabilities on the last two lines are both 1 because the occurrence of the events is guaranteed by the conditioning on $E(n_1,n_2)$ and $F(2n_2,n_3)$. The cost of half circuits is constant by (RSW). Since the width and length of the corridor T_1 is of constant proportion, again by (RSW), the probability of having a σ_1-path connecting the two ends of the corridor is also constant. Therefore,

$$\phi^{\xi}_{\mathrm{Ann}(n_1,n_3)}\left(x_1 \overset{T_1}{\underset{\sigma_1}{\longleftrightarrow}} y_1 \,\middle|\, \omega|_{T_1^c}, E(n_1,n_2) \cap F(2n_2,n_3)\right) \geq c.$$

Plugging this back into (1.8), we have

$$\begin{aligned}(1.7) \geq & \mathbf{E}\left[c\mathbf{1}\left\{\cap_{i=2}^{k}\{x_i \overset{T_i}{\underset{\sigma_i}{\longleftrightarrow}} y_i\}\right\} \,\middle|\, E(n_1,n_2) \cap F(2n_2,n_3)\right] \\ = & c\mathbf{E}\left[\mathbf{1}\left\{\cap_{i=2}^{k}\{x_i \overset{T_i}{\underset{\sigma_i}{\longleftrightarrow}} y_i\}\right\} \,\middle|\, E(n_1,n_2) \cap F(2n_2,n_3)\right].\end{aligned}$$

Applying the same procedure sequentially to each $i = 2, \ldots, k$, we have a uniform lower bound. □

CHAPTER 2

RADIAL CHEMICAL DISTANCE IN 2D CRITICAL PERCOLATION CLUSTERS

2.1 Introduction

In this chapter, we consider Bernoulli percolation on the two-dimensional lattice $\mathbb{Z}^2$ at the critical density $p_c = \frac{1}{2}$. Rather than the shortest crossing across a box, we consider the expected distance from the origin to the boundary of the box $B(n)$, conditioned on the existence of an open connection. Such a "radial" estimate is more natural than the chemical distance across a box in many applications, notably the study of random walks on percolation clusters. Unlike in the case of a horizontal crossing, there is no natural crossing to compare to in this case. Nevertheless, we show that a bound of the form (1.4) also holds for the radial chemical distance.

Let $\{0 \leftrightarrow \partial B(n)\}$ be the event that there is an open connection from the origin $(0, 0)$ to $\partial B(n)$. On $\{0 \leftrightarrow \partial B(n)\}$, the random variable S_n is defined as the chemical distance between the origin and $\partial B(n)$.

Theorem 2.1.1. *There exist some $\delta > 0$ and constant $C > 0$ independent of n such that*

$$\mathbb{E}[S_n \mid 0 \leftrightarrow \partial B(n)] \leq C n^{2-\delta} \pi_3(n). \tag{2.1}$$

Note that, in addition to the absence of a lowest crossing to compare to, the conditioning in the estimate (2.1) is singular, in that the probability $\mathbb{P}(0 \leftrightarrow \partial B(n))$ tends to zero with $n \to \infty$. In the next subsection, we explain further why the arguments of the present chapter differ essentially from [DHS17, DHS21] despite the similarity in the results (c.f. in particular Sections 2.2 and 2.4).

2.1.1 Organization

The key estimate from [DHS21], (2.16) in Section 2.3, provides a method to build a path of expected length of $n^{2-\delta}\pi_3(n)$ by finding, with high probability, shortcuts around paths consisting of "three-arm" edges. Thus one gains the $n^{-\delta}$ factor in (2.1). Using this estimate requires two inputs, one of which, the existence of an open connection of volume $O(n^2\pi_3(n))$, is a given in the case of crossing paths in [DHS17, DHS16, DHS21] but is non-trivial in our case. The second, an estimate for the probability of shortcuts around an edge e, conditional on the edge belonging to a three-arm path, leads to considerably more involved constructions than in the aforementioned works.

Concerning the first point above, in the radial case, unlike the case considered in [DHS21], there is no canonical path of three-arm points connecting the origin to the boundary of $B(n)$. In Section 2.2, we present and complete a construction from the unpublished note [DHS16] to find a path γ of expected length $n^2\pi_3(n)$ conditional on the existence of a path from the origin to $\partial B(n)$. The idea is to consider successive innermost circuits around the origin, and build three-arm paths connecting these.

Secondly, we need an estimate for the conditional probability that there is no shortcut around an edge $e \in B(n)$, conditioned on $e \in \gamma$. We show that this probability is bounded up to a constant factor by the probability of the same shortcut event conditioned on a three-arm event centered at e. The requisite gluing constructions appear in Section 2.4. Given that the construction of γ is rather more complicated than that of the lowest crossing, the (RSW)/generalized (FKG) estimates here are more involved than the ones used to obtain the corresponding comparison in [DHS17, DHS16, DHS21]. In particular, forcing the occurrence of the event $e \in \gamma$ is a more delicate matter than in these works.

2.2 A three-arm path to $\partial B(n)$

The first step towards proving our main result is to find a replacement for the lowest path in [DHS17]. Here we use ideas in the (unpublished) note [DHS16]. Note however that, although the argument in [DHS16] is presented on the square lattice $\mathbb{Z}^2$, it relies on an inequality of Beffara and Nolin [BN11] comparing monochromatic and polychromatic arm events. This is expected to hold on $\mathbb{Z}^2$, but is so far known only on the triangular lattice. The modified argument we give here holds unconditionally for the square lattice.

2.2.1 Definition of γ

We start by defining the central object γ, a path from the origin to $\partial B(n)$. On the event $\{0 \leftrightarrow \partial B(n)\}$, let C_0 be the event that there is an open circuit around the origin in $B(n)$. The definition of γ and our estimate for S_n will depend on whether C_0 occurs.

Let us first fix a deterministic ordering on all edges in $B(n)$. On $\{0 \leftrightarrow \partial B(n)\} \cap C_0^c$, there exists a dual-closed arm from the origin to the boundary. Then, let $\mathfrak{c}$ be the first such path in the lexicographical order of paths viewed as sequences of edges. We let γ be the open arm from the origin to $\partial B(n)$ closest to the counterclockwise side of $\mathfrak{c}$. Here, "closest" is measured by the number of edges (or area) between γ and $\mathfrak{c}$.

On the event C_0, we construct an open path γ from 0 to $\partial B(n)$ based on an idea in [DHS16, Section 2.3]. We first need a number of definitions.

We begin by enumerating the successive innermost open circuits around the origin. We denote the number of such circuits by $\mathcal{K}$. For $1 \leq m \leq \mathcal{K}$, we denote by $\mathcal{C}_m$ the m-th innermost open circuit.

By definition of $\mathcal{C}_1$, there is a dual-closed path $\mathfrak{c}_1$ from a dual neighbor of the origin to the endpoint of the dual edge of an edge of the inner-most circuit $\mathcal{C}_1$. On $\{0 \leftrightarrow \partial B(n)\}$, there is also an open path from the origin to $\mathcal{C}_1$. We let σ_1 be the open path from the origin to $\mathcal{C}_1$ that is closest to the counterclockwise side of $\mathfrak{c}_1$. Similarly, for $m = 2, \ldots, \mathcal{K}$, there is a dual-closed path inside the region bounded by $\mathcal{C}_m$ but outside $\mathcal{C}_{m-1}$ joining the endpoint of a dual edge to $\mathcal{C}_{m-1}$ to the endpoint of a dual edge to $\mathcal{C}_m$. We let $\mathfrak{c}_m$ be the first such dual path, and σ_m be the open path connecting $\mathcal{C}_{m-1}$ to $\mathcal{C}_m$ that is closest to the counterclockwise side of $\mathfrak{c}_m$. Finally, by the definition of $\mathcal{K}$, there necessarily exists a dual path joining a dual edge to $\mathcal{C}_{\mathcal{K}}$ to $\partial B(n)$. We let $\mathfrak{c}_{\mathcal{K}+1}$ be the first such dual path. We define $\sigma_{\mathcal{K}+1}$ to be the open path joining $\mathcal{C}_{\mathcal{K}}$ to $\partial B(n)$ that is closest to the counterclockwise side of $\mathfrak{c}_{\mathcal{K}+1}$.

The definition of the path γ from 0 to $\partial B(n)$ on the event C_0 is now as follows. The initial segment of γ is σ_1. From the endpoint of σ_1 on $\mathcal{C}_1$, we follow the circuit $\mathcal{C}_1$ in the counterclockwise direction to the endpoint of σ_2 on $\mathcal{C}_1$. Then, we repeat this procedure for $m = 2, \ldots, \mathcal{K}$, concatenating σ_m with the open subpath $\tilde{\mathcal{C}}_m$ of $\mathcal{C}_m$ obtained by following this circuit in the counterclockwise direction, joining the endpoint of σ_m belonging to $\mathcal{C}_m$ to the endpoint σ_{m+1} on $\mathcal{C}_m$. Finally, we add $\sigma_{\mathcal{K}+1}$ to the path γ. The main result of this section is the following:

Lemma 2.2.1. *There is a constant $C > 0$ independent of n such that*

$$\mathbb{E}[\#\gamma \mid 0 \leftrightarrow \partial B(n)] \le Cn^2\pi_3(n).$$

Clearly, we then have

$$S_n \le \#\gamma, \tag{2.2}$$

from which it follows that the distance S_n is $O(n^2\pi_3(n))$ in expectation. In our construction, we frequently refer to the following quantity $M = M(e)$, defined for each edge e inside $B(n)$ by

$$M(e) := \min(\mathrm{dist}(e, 0), \mathrm{dist}(e, \partial B(n))). \tag{2.3}$$

The proof of Lemma 2.2.1 occupies the remainder of Section 2.2.

2.2.2 Estimate on C_0^c.

By duality, for each edge $e \in \gamma$, there is a dual-closed path from an endpoint of e^* to $\mathfrak{c}$. Thus, for each edge $e \in \gamma$, there are two open arms and a dual-closed arm from e to distance $M(e)$. The open arms are obtained by following γ from e in either direction, and the dual-closed arm is obtained by following the dual-closed path from e^* to $\mathfrak{c}$, and then following $\mathfrak{c}$ to the origin or $\partial B(n)$. Combining this fact with (2.2), we obtain:

$$\mathbb{E}[\#\gamma, C_0^c \mid 0 \leftrightarrow \partial B(n)] \le C \sum_{k=1}^{\lfloor n/2 \rfloor} \sum_{e:\ M(e)=k} \mathbb{P}(A_3(e,k) \mid 0 \leftrightarrow \partial B(n)). \tag{2.4}$$

Let $d = \mathrm{dist}(0, e)$. If in the inner sum $d < 2k$, $e \in B(2n/3)$. By independence, we have

$$\mathbb{P}(A_3(e,k) \mid 0 \leftrightarrow \partial B(n)) \le \frac{\mathbb{P}(0 \leftrightarrow \partial B(d/2))\mathbb{P}(A_3(e,k/2))\mathbb{P}(\partial B(3d/2) \leftrightarrow \partial B(n))}{\mathbb{P}(0 \leftrightarrow \partial B(n))}. \tag{2.5}$$

By (QM), we have for some constant $C > 0$,

$$\mathbb{P}(0 \leftrightarrow \partial B(n)) \geq C\mathbb{P}(0 \leftrightarrow \partial B(d/2))\mathbb{P}(\partial B(3d/2) \leftrightarrow \partial B(n)). \tag{2.6}$$

If $d \geq 2k$, then again by independence, we have

$$\mathbb{P}(A_3(e,k) \mid 0 \leftrightarrow \partial B(n)) \leq \frac{\mathbb{P}(0 \leftrightarrow \partial B(n/2))\mathbb{P}(A_3(e,k/2))}{\mathbb{P}(0 \leftrightarrow \partial B(n))}. \tag{2.7}$$

Similarly, by (1.5), we have for some constant $C > 0$,

$$\mathbb{P}(0 \leftrightarrow \partial B(n)) \geq C\mathbb{P}(0 \leftrightarrow \partial B(n/2)). \tag{2.8}$$

Using (2.5), (2.6), (2.7), and (2.8) to estimate (2.4), we have

$$\begin{aligned} \mathbb{E}[\#\gamma, C_0^c \mid 0 \leftrightarrow \partial B(n)] &\leq C \sum_{k=1}^{\lfloor n/2 \rfloor} \sum_{e:M(e)=k} \mathbb{P}(A_3(e,k/2)) \\ &\leq C \sum_{k=1}^{\lfloor n/2 \rfloor} k\pi_3(k). \end{aligned} \tag{2.9}$$

To bound the final quantity, we use the following result, proved for example in [Kes86, Eqn. (7)] and [DHS17, Proposition 16].

Proposition 2.2.2. *There exists $C > 0$ such that for all $L \in \mathbb{Z}^+$,*

$$\sum_{l=1}^{L} l\pi_3(l) \leq CL^2\pi_3(L).$$

Applying this to (2.9), we immediately obtain the desired bound

$$\mathbb{E}[\#\gamma, C_0^c \mid 0 \leftrightarrow \partial B(n)] \leq Cn^2\pi_3(n).$$

2.2.3 Estimate on C_0.

We estimate the expected volume $\mathbb{E}[\#\gamma \mid 0 \leftrightarrow \partial B(n)]$. Proceeding as in the previous case C_0^c (see (2.4)), it suffices to estimate the conditional probability that an edge $e \in B(n)$ belongs to any portion of γ.

The initial path σ_1

If $e \in \sigma_1$, then by duality there is a dual-closed path from a dual neighbor of e to the dual-closed path $\mathfrak{c}_1$, which then follows $\mathfrak{c}_1$ to a dual neighbor of the origin. There are also two open arms from e to the origin and to C_1 respectively. All three arms reach to distance at least $d = \text{dist}(e, 0)$. Recalling the notation (2.3), we have the estimate:

$$\mathbb{P}(e \in \sigma_1 \mid 0 \leftrightarrow \partial B(n)) \leq \mathbb{P}(A_3(e, d)) \leq C\pi_3(M(e)). \tag{2.10}$$

The final path $\sigma_{\mathcal{K}+1}$

For an edge $e \in \sigma_{\mathcal{K}+1}$, by duality there is a dual-closed path from a dual neighbor of e to the dual-closed path $\mathfrak{c}_{\mathcal{K}+1}$, which then follows $\mathfrak{c}_{\mathcal{K}+1}$ to $\partial B(n)$. There are also two open arms from e to $\partial B(n)$ and $C_{\mathcal{K}}$ respectively. As in the previous case, from e, there are three arms to distance at least $M(e)$ and we obtain the bound

$$\mathbb{P}(e \in \sigma_{\mathcal{K}+1} \mid 0 \leftrightarrow \partial B(n)) \leq C\pi_3(M(e)). \tag{2.11}$$

The intermediate paths σ_m for $m = 2, \ldots, \mathcal{K}$

In this case, we use a modified version of Lemma 11 from [DHS16].

Lemma 2.2.3. *Suppose $e \in \sigma_m$ for $2 \leq m \leq \mathcal{K}$. Fix $\epsilon > 0$ and an integer R such that for any $0 \leq n_1 < n_2$,*

$$\pi'_{2R+2}(n_1, n_2) \leq \pi_3(n_1, n_2)(n_1/n_2)^{\epsilon}. \tag{2.12}$$

There are $0 = \ell_0 < \ell_1 \leq \cdots \leq \ell_R \leq \lfloor \log M(e) \rfloor$ such that

1. *There are two disjoint open arms and one dual-closed arm from e to $\partial B(e, 2^{\ell_1 - 1})$.*

2. *For $i \geq 2$, if $\ell_{i-1} < \lfloor \log M(e) \rfloor$, there are $2i$ open arms and one dual-closed arm from $\partial B(e, 2^{\ell_{i-1}})$ to $\partial B(e, 2^{\ell_i - 1})$, all of which are disjoint.*

3. *If $\ell_R < \lfloor \log M(e) \rfloor$, there are $2R + 2$ disjoint open arms from $\partial B(e, 2^{\ell_R})$ to $\partial B(e, M(e))$.*

Proof. We first note that such an R that satisfies (2.12) exists by (Reimer) and because $\pi_3(n_1, n_2) \geq (n_1/n_2)^c$ for some constant c. See [Nol08, Proposition 14].

We observe that $e \in \sigma_m$ lies between the two circuits $\mathcal{C}_{m-1}$ and $\mathcal{C}_m$. Let x be the endpoint of σ_m intersecting $\mathcal{C}_{m-1}$, and let y be the other endpoint intersecting $\mathcal{C}_m$. Let ℓ'' be the smallest ℓ' such that $2^{\ell'} \geq M$ or $B(e, 2^{\ell'})$ contains both x and y.

Suppose first that $2^{\ell''} \geq M$. In this case, we let $\ell_1 = \lfloor \log M \rfloor$. Subsequently, we let all $\ell_i = \lfloor \log M \rfloor$ for $i > 1$. The two endpoints of σ_m inside $B(e, 2^{\ell_1})$ form two open arms from e to $\partial B(e, 2^{\ell_1})$. The edge of σ_m with endpoint y has a dual edge connected by a dual closed path to $\mathfrak{c}_m$. Since the same is true of the edge e^*, we obtain a closed arm from e^* which extends at least to distance $\mathrm{dist}(e, y) \geq 2^{\ell_1 - 1}$.

If $2^{\ell''} < M$, we let $\ell_1 = \ell''$. As in the previous case, we obtain two open arms and a dual closed arm from e to distance $2^{\ell_1 - 1}$. In addition, both x and y are contained in $B(e, 2^{\ell_1})$, and each lies on one of the open circuits $\mathcal{C}_{m-1}$ and $\mathcal{C}_m$ around 0, whereas $0 \notin B(e, 2^{\ell_1})$ since $2^{\ell_1} < M$.

If $m \leq 2$, we let $\ell_2 = \lfloor \log M \rfloor$. Otherwise, if $2^{\ell_1} < M$, we let ℓ'' be the least $\ell' \geq \ell_1$ such that $B(e, 2^{\ell'})$ intersects $\mathcal{C}_{m-2}$. If $2^{\ell''} \geq M$, we let $\ell_2 = \lfloor \log M \rfloor$, and if $2^{\ell''} < M$, we let $\ell_2 = l''$. In all three cases we have four open arms (following $\mathcal{C}_{m-1}$ and $\mathcal{C}_m$) and one dual-closed arm (the endpoint x has a dual edge connected by a dual closed path to a dual edge of $\mathcal{C}_{m-2}$) from $\partial B(e, 2^{\ell_1})$ to $\partial B(e, 2^{\ell_2 - 1})$.

Inductively, unless $\ell_{i-1} = \lfloor \log M \rfloor$, we define ℓ_i to be the least such that $B(e, 2^{\ell_i})$ intersects C_{m-i}. For each scale ℓ_i, if $\ell_i < \lfloor \log M \rfloor$, the box $B(e, 2^{\ell_i})$ crosses one more circuit than $B(e, 2^{\ell_{i-1}})$, thus resulting in two more open arms. At scale ℓ_R, one may not have a dual-closed arm from $\partial B(e, 2^{\ell_R})$ to $\partial B(e, M(e))$, but there are $2R+2$ disjoint open arms since the box $B(e, 2^{\ell_R})$ crosses $R + 1$ circuits, each resulting in two open arms. If for some i, $\ell_{i-1} = \lfloor \log M \rfloor$, we define all subsequent scales to be $\lfloor \log M \rfloor$. □

On the event $\left\{e \in (\cup_{m=2}^{\mathcal{K}} \sigma_m)\right\}$, by Lemma 2.2.3, computations similar to (2.5), (2.6), and a union bound over $0 = \ell_0 < \ell_1 \le \cdots \le \ell_R \le \lfloor \log M(e) \rfloor$, we have

$$\mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m \mid 0 \leftrightarrow \partial B(n)) \le C \sum_{0=\ell_0<\ell_1\le\cdots\le\ell_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R} \pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2R+2}(2^{\ell_R}, M). \quad (2.13)$$

Using (Reimer) in the form

$$\pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \le \pi_3(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2i-2}(2^{\ell_{i-1}}, 2^{\ell_i - 1}),$$

(1.5), and (2.12) we have

$$\begin{aligned} &\sum_{0=\ell_0<\ell_1\le\cdots\le l_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R} \pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2R+2}(2^{\ell_R}, M) \\ &\qquad \le C' \sum_{0=\ell_0<\ell_1\le\cdots\le l_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R} \pi_3(2^{\ell_{i-1}}, 2^{\ell_i}) \pi_3(2^{\ell_R}, M) \prod_{i=1}^{R} \pi'_{2i-2}(2^{\ell_{i-1}}, 2^{\ell_i}) \left(\frac{2^{\ell_R}}{M}\right)^{\epsilon} \end{aligned} \quad (2.14)$$

We are then able to use the standard gluing construction R times between each pair of consecutive scales for the polychromatic three-arm event:

$$\prod_{i=1}^{R} \pi_3(2^{\ell_{i-1}}, 2^{\ell_i}) \pi_3(2^{\ell_R}, M) \le C^R \pi_3(M(e)).$$

For the monochromatic probabilities in (2.14), we have the following estimate:

$$\pi'_j(n_1, n_2) \le (\pi_1(n_1, n_2))^j \le \left(\frac{n_1}{n_2}\right)^{\alpha j}, \text{ for some } \alpha > 0.$$

Plugging this back into (2.14), we have

$$(2.14) \le C' C^R \pi_3(M(e)) \sum_{0=\ell_0<\ell_1\le\cdots\le\ell_R}^{\lfloor \log M(e) \rfloor} \prod_{i=0}^{R-1} (2^\alpha)^{2i(\ell_i-\ell_{i+1})} \left(\frac{2^{\ell_R}}{M}\right)^\epsilon .$$

For the summation, we use the following estimate inductively: for $x = 2^\beta, \beta > 0$, we have

$$\sum_{i \le N} x^i \le c x^N .$$

for some constant $c = c(\beta)$. Therefore

$$\begin{aligned}
\sum_{0=\ell_0<\ell_1\le\cdots\le l_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R} \pi'_{2i-2}(2^{\ell_{i-1}}, 2^{\ell_i}) \left(\frac{2^{\ell_R}}{M}\right)^\epsilon &\le \sum_{0=\ell_0<\ell_1\le\cdots\le l_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R-1} (2^\alpha)^{2\ell_i} (2^\alpha)^{-2(R-1)\ell_R} \left(\frac{2^{\ell_R}}{M}\right)^\epsilon \\
&\le \sum_{\ell_R=1}^{\lfloor \log M(e) \rfloor} C (2^\alpha)^{2(R-1)\ell_R} (2^\alpha)^{-2(R-1)\ell_R} \left(\frac{2^{\ell_R}}{M}\right)^\epsilon \\
&\le C 2^{\epsilon \log M} M^{-\epsilon} \le C.
\end{aligned}$$

Inserting this estimate into (2.13), we obtain

$$\mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m \mid 0 \leftrightarrow \partial B(n)) \le C \pi_3(M(e)).$$

The circuits C_m for $m = 1, \ldots, \mathcal{K}$

For this estimate, we again use a modified lemma from [DHS16].

Lemma 2.2.4. *Suppose C_0 occurs and $e \in C_m$ for some $1 \le m \le \mathcal{K}$. Letting R be as in Lemma 2.2.3, there are $0 = \ell_0 < \ell_1 \le \cdots \le \ell_R \le \lfloor \log M(e) \rfloor$ such that:*

1. *e has two disjoint open arms and one dual-closed arm to $\partial B(e, 2^{\ell_1 - 1})$.*
2. *For $i \ge 2$, if $2^{\ell_{i-1}} < M$, there are $2i$ disjoint open arms and one dual-closed arm from $\partial B(e, 2^{\ell_{i-1}})$ to $\partial B(e, 2^{\ell_i - 1})$.*

3. *If* $\ell_R < \lfloor \log M(e) \rfloor$*, there are* $2R+2$ *disjoint open arms from* $\partial B(e, 2^{\ell_R})$ *to* $\partial B(e, M(e))$*.*

Proof. By duality, since $e \in C_m$, there is a dual-closed path in $\mathrm{int}(C_m)$ connecting e^* to a dual neighbor of 0 if $m = 1$ or to the dual of some edge $e' \in C_{m-1}$ if $m > 1$. Let ℓ_1 be the minimum ℓ' such that $2^{\ell'} \geq \mathrm{dist}(e, e')$ if $m > 1$ (or $2^{\ell} \geq \mathrm{dist}(e, 0)$ if $m = 1$). Then there are three arms, two open and one dual-closed, from e to $\partial B(e, 2^{\ell_1 - 1})$.

If $2^{\ell_1} < M(e)$, since $e' \in B(e, 2^{\ell_1})$, we can find four open arms from $\partial B(e, 2^{\ell_1})$ to $\partial B(e, M(e))$ by following the circuits C_m and C_{m-1} from e and e' in both directions. We also have a dual-closed arm from e' to a dual neighbor of some edge $e'' \in C_{m-2}$ if $m > 2$ or to the dual neighbor of 0 if $m = 2$. If $m > 2$, we define ℓ_2 to be the least $\ell_2 \geq \ell_1$ such that $B(e, 2^{\ell_2})$ intersects C_{m-2}. Then we have four disjoint open arms and one dual-closed arm from $\partial B(e, 2^{\ell_1})$ to $\partial B(e, 2^{\ell_2 - 1})$. If $m = 2$, we define $\ell_2 = \lfloor \log M(e) \rfloor$. There are four open arms from $\partial B(e, 2^{\ell_1})$ to $\partial B(e, M(e))$.

Inductively, if $m > R$ and $\ell_{i-1} < \lfloor \log M(e) \rfloor$, we let ℓ_i be the least $\ell_i \geq \ell_{i-1}$ such that $B(e, 2^{\ell_i})$ intersects C_{m-i}. Otherwise, if $\ell_{i-1} = \lfloor \log M(e) \rfloor$, we let $\ell_i = \cdots = \ell_R = \lfloor \log M(e) \rfloor$. If all R scales $\ell_1, \ldots, \ell_R$ are less than $\lfloor \log M(e) \rfloor$, at scale ℓ_R, $B(e, 2^{\ell_R})$ crosses $R+1$ circuits. So there are $2R+2$ disjoint open arms from $\partial B(e, 2^{\ell_R})$ to $\partial B(e, M(e))$. If $m \leq R$, then for all $i \geq m$, we define $\ell_i = \lfloor \log M(e) \rfloor$. $\square$

On the event $\{e \in \cup_{m=1}^{\mathcal{K}} C_m\}$, by the same arguments used to bound (2.13), we obtain

$$\begin{aligned}\mathbb{P}(e \in \cup_{m=1}^{\mathcal{K}} C_m \mid 0 \leftrightarrow \partial B(n)) &\leq C \sum_{0=\ell_0<\ell_1\leq\cdots\leq\ell_R}^{\lfloor \log M(e) \rfloor} \prod_{i=1}^{R} \pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2R+2}(2^{\ell_R}, M) \\ &\leq C\pi_3(M(e)). \end{aligned} \tag{2.15}$$

Summation on C_0

As in the case of C_0^c, we bound the length of the path γ defined at the beginning of this subsection.

$$\begin{aligned}\mathbb{E}[\#\gamma, C_0 \mid 0 \leftrightarrow \partial B(n)] \leq &\sum_{e \in B(n)} \mathbb{P}(e \in \{\sigma_1, \sigma_{\mathcal{K}+1}\} \mid 0 \leftrightarrow \partial B(n)) \\ &+ \sum_{e \in B(n)} \mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m \mid 0 \leftrightarrow \partial B(n)) \\ &+ \sum_{e \in B(n)} \mathbb{P}(e \in \cup_{m=1}^{\mathcal{K}} C_m \mid 0 \leftrightarrow \partial B(n)).\end{aligned}$$

Using (2.10), (2.11), (2.15), and summing over the values of e and M as in (2.4), we find

$$\mathbb{E}[\#\gamma, C_0 \mid 0 \leftrightarrow B(n)] \leq Cn^2 \pi_3(n).$$

2.3 Proof of Theorem 2.1.1

In this section, we combine the result in [DHS21], stated as Proposition 2.3.1, with the estimate of Lemma 2.2.1 to prove Theorem 2.1.1.

Let j be a sufficiently large integer, and $0 < \epsilon < 1$ be a small parameter. In [DHS21], the authors define a sequence of events $E_j(e, \epsilon, v)$ such that

- each event depends only on the status of edges in the annulus

  $$Ann_j = \text{Ann}(e_x; 2^j, 2^{j + \lfloor \log \frac{1}{\epsilon} \rfloor})$$

  around e. Recall the vertex $e_x \in \mathbb{Z}^2$ is the lower left endpoint of the edge e;
- If σ is any open path consisting of edges with three arms, two open and one dual closed extending to the outside of Ann_j which includes e, then

there exists an open path r inside Ann_j which is edge-disjoint from σ, but whose endpoints u and v lie on σ;

- Denoting by τ the portion of σ between u and v, we have $e \in \tau$;
- The number of edges in r is at most ν times the number of edges in τ. We say that r is a ν-shortcut around e:

$$\#r \leq \nu \cdot \#\tau.$$

See [DHS21, Section 5 and Appendix A] for proofs of these statements. The main result in [DHS21, Proposition 5.6] is that, for $0 < \delta < 1$ and for $\epsilon > 0$ sufficiently small, there exist constants $c, \hat{c} > 0$ such that

$$\mathbb{P}\Big(\bigcap_{j=\lceil \frac{\delta}{8} \log n \rceil}^{\lfloor \frac{\delta}{4} \log n \rfloor} E_j(e, \epsilon, n^{-c})^c \mid A_3(e, n^{\delta/2}) \Big) \leq 2^{-\hat{c} \frac{\delta \epsilon^4 \log n}{8 \log(1/\epsilon)}}. \tag{2.16}$$

The estimate (2.16) implies that, conditional on the existence of 3 arms to distance $n^{\delta/2}$, there is a shortcut around e which saves n^{-c} edges with probability at least $1 - n^{-\eta}$ for some small η. See (2.18). We will apply this to find shortcuts around the path γ constructed in Section 2.2.

Our main result follows from (2.16) and the following Proposition, the analogue of Proposition 8 in [DHS21], with the lowest crossing of a box replaced by our path γ.

Proposition 2.3.1. *Let $e \in B(n)$ and let $d \leq M/4 := \frac{1}{4} \min(\mathrm{dist}(e, 0), \mathrm{dist}(e, \partial B(n)))$. There are choices of $c, r > 0$ uniform in d such that for $k \leq d^{9/10}$ and any event E depending only on the status of edges in $B(e, k)$:*

$$\mathbb{P}(E \mid 0 \leftrightarrow \partial B(n), e \in \gamma) \leq c(d^{-r} + \mathbb{P}(E \mid A_3(e, d))).$$

Proposition 2.3.1 is proved in Section 2.4.

Proof of Theorem 2.1.1. Choose $\delta > 0$ small enough so that

$$n^{1+2\delta} \leq n^2 \pi_3(n), \tag{2.17}$$

and define the truncated box $\hat{B}(n) = B(n - n^\delta) \setminus B(n^\delta)$. This is possible because $\pi_3(n) \geq n^{-1+s}$ for some $s > 1$; see [DHS17, Lemma 3.1] for details. With our choice of $\delta > 0$ and $j \in \left(\frac{\delta}{8} \log n, \frac{\delta}{4} \log n\right)$, we have, if $e \in \hat{B}(n)$:

$$\mathbb{P}(\text{there is no } n^{-c}\text{-shortcut around } e \mid e \in \gamma) \leq \mathbb{P}\Big(\bigcap_{j=\lceil \frac{\delta}{8} \log n \rceil}^{\lfloor \frac{\delta}{4} \log n \rfloor} E_j(e, \epsilon, n^{-c})^c \mid e \in \gamma\Big). \tag{2.18}$$

Applying Proposition 2.3.1 and using (2.16), the quantity is bounded by

$$\begin{aligned} & cn^{-c'\delta r} + c\mathbb{P}\Big(\bigcap_{j=\lceil \frac{\delta}{8} \log n \rceil}^{\lfloor \frac{\delta}{4} \log n \rfloor} E_j(e, \epsilon, n^{-c})^c \mid A_3(e, n^{\delta/2})\Big) \\ & \leq cn^{-c'\delta r} + 2^{-\hat{c}\frac{\delta \epsilon^4 \log n}{8 \log(1/\epsilon)}} \end{aligned}$$

The last quantity is bounded by $n^{-\eta}$ for some $\eta > 0$.

Along the path γ, we now choose a collection of (vertex)-disjoint n^{-c}-shortcuts r_l such that the total length of the corresponding detoured paths τ_l is maximal. We define a path s from 0 to $\partial B(n)$ by taking the union of all the shortcuts r_l, together with all the edges of γ around which no shortcut exists.

Partitioning the edges in γ given $\{0 \leftrightarrow \partial B(n)\}$ into the truncated part of the box, the edges with n^{-c}-shortcuts, and the edges without shortcut, we estimate the expected size of s as follows:

$$\begin{aligned}
\mathbb{E}[\#s \mid 0 \leftrightarrow \partial B(n)] \leq Cn^{1+\delta} + n^{-c} \sum_{l} \mathbb{E}[\#\tau_l \mid 0 \leftrightarrow \partial B(n)] \\
+ \mathbb{E}[\#\{e \in \gamma \cap \hat{B}(n) : e \text{ has no } n^{-c}\text{-shortcut}\} \mid 0 \leftrightarrow \partial B(n)] \\
\leq Cn^{1+\delta} + n^{-c}\mathbb{E}[\#\gamma \cap \hat{B}(n) \mid 0 \leftrightarrow \partial B(n)] + n^{-\eta}\mathbb{E}[\#\gamma \cap \hat{B}(n) \mid 0 \leftrightarrow \partial B(n)] \\
\leq Cn^{1+\delta} + n^{-\min\{c,\eta\}}\mathbb{E}[\#\gamma \mid 0 \leftrightarrow \partial B(n)].
\end{aligned}$$

By (2.17) and Lemma 2.2.1, we now have

$$\mathbb{E}[S_n \mid 0 \leftrightarrow \partial B(n)] \leq \mathbb{E}[\#s \mid 0 \leftrightarrow \partial B(n)] \leq Cn^{2-\min\{c,\eta,\delta\}}\pi_3(n).$$

□

2.4 Improving on γ: connecting shortcuts around three-arm points

In this section, we derive the main estimate in Proposition 2.3.1 in Section 2.3: there are constants $c, r > 0$ uniform in d such that

$$\mathbb{P}(E(e,k) \mid 0 \leftrightarrow \partial B(n), e \in \gamma) \leq c(d^{-r} + \mathbb{P}(E(e,k) \mid A_3(e,d))) \tag{2.19}$$

where $E(e,k)$ is any event that depends only on the status of edges in the box of size k centered at e such that $k \leq d^{9/10}$ for $d \leq M/4$.

It suffices to show (2.19) for $d = M/4$ since for all $k^{10/9} \leq d \leq M/4$:

$$\mathbb{P}(E(e,k) \mid A_3(e,M/4)) \leq C\mathbb{P}(E(e,k) \mid A_3(e,d)). \tag{2.20}$$

To see (2.20), we note that $E(e,k) \cap A_3(e,M/4)$ implies $E(e,k) \cap A_3(e,d) \cap A_3(e;d,M/4)$. By independence, we have

$$\mathbb{P}(E(e,k), A_3(e,M/4)) \leq \mathbb{P}(E(e,k), A_3(e,d))\mathbb{P}(A_3(e;d,M/4)).$$

On the other hand, by the standard gluing constructions,

$$\mathbb{P}(A_3(e, M/4)) \geq c'\mathbb{P}(A_3(e, d))\mathbb{P}(A_3(e; d, M/4))$$

where c' holds uniformly in d. Then, (2.20) follows by the definition of conditional probability.

As in Section 2.2, we split the event $\{0 \leftrightarrow \partial B(n)\}$ into the event C_0 that there exists an open circuit around the origin in $B(n)$ and its complement C_0^c. In Sections 2.2.2 and 2.2.3 respectively, we defined a path γ from the origin to $\partial B(n)$ on each of these events. We estimate the conditional probability in (2.19) by splitting $\{0 \leftrightarrow \partial B(n)\} \cap \{e \in \gamma\}$ into a number of cases, depending on the location of e and which part of the path γ the edge lies on. By the decomposition

$$\{0 \leftrightarrow \partial B(n)\} \cap \{e \in \gamma\} = (\{e \in \gamma\} \cap C_0) \cup (\{e \in \gamma\} \cap C_0^c),$$

we have

$$\mathbb{P}(E(e, k) \mid 0 \leftrightarrow \partial B(n), e \in \gamma) \leq \mathbb{P}(E(e, k) \mid e \in \gamma, C_0^c) + \mathbb{P}(E(e, k) \mid e \in \gamma, C_0). \quad (2.21)$$

Hence it suffices to derive the estimate (2.19) with the left side replaced by $\mathbb{P}(E(e, k) \mid C, e \in \gamma)$, for $C = C_0$ or $C = C_0^c$. This involves intricate but standard gluing constructions using (RSW) and generalized (FKG) estimates. See Section 1.5.1 for a discussion of such constructions. In the interest of brevity, we do not spell out the full details of the applications of (FKG) and (RSW), but only indicate the relevant connections and provide figures for the reader's guidance.

2.4.1 Estimate on C_0^c

Recall from Section 2.2.2 that C_0^c is the event that there is no open circuit around the origin in $B(n)$. To minimize repetition, we treat this basic case carefully, and later indicate the necessary modifications to the argument for all other cases.

Recall also from Section 2.2.2 that $\mathfrak{c}$ is the first dual-closed path from the origin to $\partial B(n)$ in a fixed deterministic ordering of paths on the dual lattice. The path γ from the origin to $\partial B(n)$ was constructed by choosing the closest open path to $\mathfrak{c}$ on the counterclockwise side. Therefore, for each edge $e \in \gamma$, there must be a dual-closed path connecting e and $\mathfrak{c}$, resulting in an intersection point of the two dual-closed paths.

By (2.21), our task is to estimate

$$\mathbb{P}(E(e,k) \mid e \in \gamma, C_0^c) = \frac{\mathbb{P}(E(e,k), e \in \gamma, C_0^c)}{\mathbb{P}(e \in \gamma, C_0^c)}. \tag{2.22}$$

We find an upper bound for the numerator in (2.22) and, by a gluing construction, a lower bound for the denominator. We distinguish two cases, depending on the location of e.

Case A: e is in $B(n/2)$.

In this case, $M = \operatorname{dist}(e, 0)$. The event in the probability in the numerator in (2.22), $E(e,k) \cap C_0^c \cap \{e \in \gamma\}$, implies the occurrence of

- a two-arm event in $B(M/4)$,
- a three-arm event and $E(e,k)$ in $B(e, M/4)$,
- a two-arm event in the annulus $\operatorname{Ann}(2M, n)$.

By independence of disjoint regions, this gives the estimate

$$\mathbb{P}(E(e,k), e \in \gamma, C_0^c) \leq \mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(2M, n)) \tag{2.23}$$

The estimate for the denominator in (2.22) is somewhat more delicate, because we need to construct an event with probability of order

$$\mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(A_{3,OOC}(e,M/4))\mathbb{P}(A_{2,OC}(2M,n))$$

which ensures the occurrence of $\{e \in \gamma\}$. The relevant construction is illustrated in Figure 2.1.

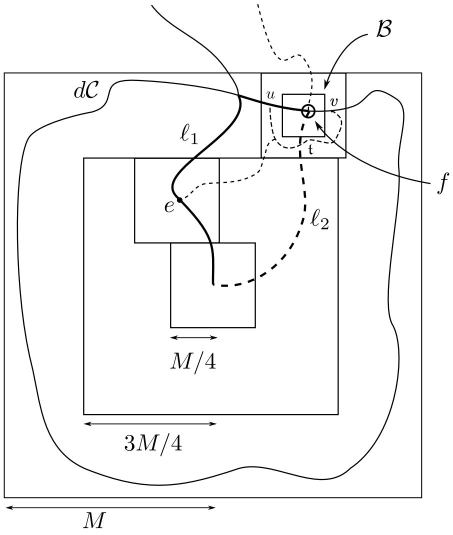


Figure 2.1: The construction on scale M around the origin 0 on the event C_0^c, Case A. ($M = \text{dist}(e, 0)$). Here ℓ_1 is an open path from the origin to $\partial B(n)$ and ℓ_2 is a dual-closed path from 0^* to $\partial B(n)^*$. Such paths necessarily exist on the event C_0^c. The circuit in black represents dC, a circuit with defect. The area bounded by the bold black and dotted curves is the domain J in the definition of the event D. The construction in the figure forces the edge e to belong to the open path γ. The figures are not to scale.

Definition 2.1. Let D be the event that

1. The edge e lies on an open arm ℓ_1 from the origin to $\partial B(n)$;
2. There is an open circuit dC with a defect in the annulus $\text{Ann}(3M/4, M)$. The defect edge f is contained in a box $\mathcal{B} \subset \text{Ann}(13M/16, 15M/16)$ of side length $M/8$;

3. There is a dual-closed arc $\mathfrak{f}$ inside dC, between the box $\mathcal{B}$ and a box $\mathcal{B}'$ with the same center as $\mathcal{B}$ and twice the side length of $\mathcal{B}$. This arc connects two edges that are dual to two open edges u and v on the circuit with defect dC. The edge u lies on the arc of dC between f and the endpoint of ℓ_1 on the circuit when the circuit is traversed in the counterclockwise direction. The edge v lies between f and the endpoint of ℓ_1 when the circuit is traversed in the clockwise direction. u and v are connected by an arc that consists only of open edges of dC inside $\mathcal{B}$ as well as the (closed) defect edge f;
4. A dual neighbor 0^* of the origin is connected to $\partial B(n)^*$ by a dual-closed path ℓ_2, which necessarily contains f^*, the dual of the defect edge in the second item. Consider the closed curve[1] obtained by joining the origin to 0^*, followed by ℓ_2, then following the circuit with defect dC in the *counterclockwise* direction from f to the endpoint of ℓ_1 on dC, and finally following ℓ_1 back to the origin. Denote the region bounded by this curve by J;
5. The dual edge e^* is connected to the arc $\mathfrak{f}$ by a dual-closed path lying inside J. In other words, $\mathfrak{f}$ is connected to e^* to the *clockwise* side of ℓ_1.

The significance of the event D is in the following:

Lemma 2.4.1. *The event D implies $C_0^c \cap \{e \in \gamma\}$.*

Proof. The dual-closed connections from the origin to $\partial B(n)^*$ imply the occurrence of C_0^c. Any dual-closed path from the origin to $\partial B(n)$ must contain the edge f^*, and thus cross the dual arc $\mathfrak{f}$. This includes the arc $\mathfrak{c}$ in the definition of γ in Section 2.2. Thus e^* is connected to $\mathfrak{c}$ by a dual-closed path starting at $\mathfrak{c}$ such that the other endpoint is connected to the clockwise side of ℓ_1. The edge e must thus be part of the counterclockwise closest open path to $\mathfrak{c}$ in $B(n)$, so $e \in \gamma$. □

---

[1] "Closed" here means that the curve is a continuous image of a circle.

Standard gluing constructions using generalized (FKG) as in Section **??** and [DHS17, Section 5] give the following.

Lemma 2.4.2. *There is a constant* $c > 0$ *such that, uniformly in the location of* $e \in B(n)$, *we have*

$$\mathbb{P}(D) \geq c\mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(2M, n)).$$

Proof. The proof involves a repeated application of the generalized (FKG) and (RSW) estimates, as well as (1.5), to construct the connections indicated in Figure 2.4 and force the occurrence of the event D, following two general principles:

- connections across boxes or annuli with aspect ratio on a fixed scale (either n or M) have probabilities lower bounded by constants independent of n, M, and
- open (resp. dual closed) connections between different scales $n_1 \ll n_2$ have probability costs comparable to arm events across the annulus $\mathrm{Ann}(n_1, n_2)$.

For the construction inside the box $\mathcal{B}$, which contains the defect edge f, we use the second moment method. See Figure 2.2. Denoting by $\mathcal{N}$ the number of dual-closed edges e' in the box of side length $M/16$ with the same center as $\mathcal{B}$, such that e' is connected to intervals I_1 and I_2 on the top, resp. bottom side of $\mathcal{B}'$ by two disjoint dual-closed arms, and connected to I_3 and I_4 on the left, resp. right side of $\mathcal{B}'$ by two disjoint open arms, we have

$$cM^2\pi_{4,OCOC}(M) \leq \mathbb{E}[\mathcal{N}] \leq CM^2\pi_{4,OCOC}(M).$$

A calculation analogous to that in [DHS17, Proposition 5.9] shows that

$$\mathbb{E}[\mathcal{N}^2] \leq C(\mathbb{E}[\mathcal{N}])^2.$$

By the second moment method, it follows that $\mathcal{N} > 0$ with positive probability, in which case there is at least one such edge e'. Using generalized (FKG), the dual-closed arms emanating from e' are extended into a dual-closed arm from the origin to $\partial B(n)$, while the open arms are extended into the circuit with defect dC, as shown in Figure 2.1. $\square$

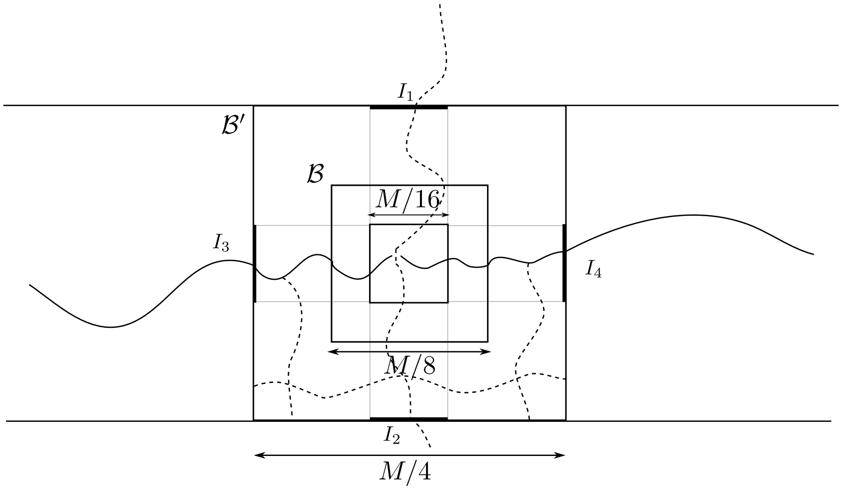


Figure 2.2: The construction inside the box $\mathcal{B}'$ to obtain a defect edge: we use the second moment method to show that, with probability bounded below uniformly in n, there is an edge inside a box of side length $M/16$ with four alternating arms (open, dual closed, open, dual closed) connected to prescribed sides of the box $\mathcal{B}'$. The open connections are then extended into an open circuit with defect. The dual-closed connections are extended into an arm from the origin to $\partial B(n)$. The overall cost of all connections indicated in this figure is bounded below by a constant.

Combining the previous two lemmas and using (1.5), we obtain that the denominator in (2.22) is bounded below:

$$\mathbb{P}(e \in \gamma, C_0^c) \geq c\mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(2M, n)),$$

where $c > 0$ is a positive constant. Together with the upper bound (2.23), the above implies the estimate

$$\mathbb{P}(E(e,k) \mid C_0^c, e \in \gamma) \leq C\mathbb{P}(E(e,k) \mid A_{3,OOC}(e, M/4)),$$

in Case A, $M = \text{dist}(e, 0)$.

Case B: e is closer to $\partial B(n)$.

In this case, $M = \text{dist}(e, \partial B(n))$.

To estimate the numerator in (2.22), we separate $B(n)$ into three regions: the inner box $B(n/4)$, the small box $B(e, M/4)$ around e, and the region $\text{Ann}(n/4, n) \setminus B(e, M/4)$.

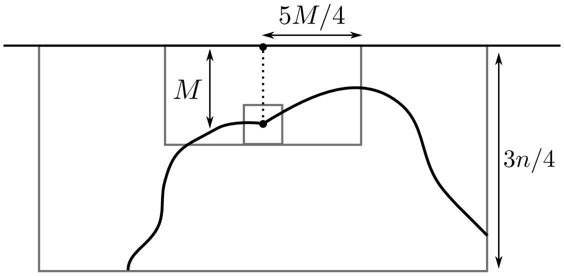


Figure 2.3: The existence of two open arms to distance n from an edge e at distance M from $\partial B(n)$ implies the two-arm event $A_{2,OO}(M/4)$ and the half-plane event $A^{hp}_{2,OO}(p(e); 5M/4, 3n/4)$.

Denote by $p_x(e)$ the projection onto $\partial B(n)$ of the lower left endpoint of e (the first of the two endpoints in the lexicographic order on $\mathbb{Z}^2$) along the x-axis, and by $p_y(e)$ the projection of the lower left endpoint onto $\partial B(n)$ along the y-axis. Finally, let $p(e)$ be the ℓ^2 projection of the lower left endpoint onto $\partial B(n)$.

We make extensive use of the *half-plane* events $A^{hp}_{k,\sigma}$. Their relevance here is illustrated in Figure 2.3. For example, if $e \in B(n)$ is at distance $M \ll n$ from the boundary and has two open arms to distance of order n, then this implies the simultaneous occurrence of the two-arm event $A_{2,OO}(e, M/4)$ *and* the occurrence inside $B(n) \cap \text{Ann}(p(e); 5M/4, 3n/4)$ of two open arms. This last event is the half-plane arm event $A^{hp}_{2,OO}(e; 5M/4, 3n/4)$.

To simplify the proofs of the estimates below, we will assume that

$$M = \mathrm{dist}(e, \partial B(n)) = \mathrm{dist}(e, p(e)) \ll \max\{\mathrm{dist}(e, p_x(e)), \mathrm{dist}(e, p_y(e))\} \geq cn.$$

This amounts to assuming that the edge e is not close to a corner of the box $\partial B(n)$. The case when e is close to a corner is dealt with by similar constructions to those in this section, involving the use of quarter-plane arm events instead of half-plane events.

The event $\{E(e,k), e \in \gamma, C_0^c\}$ implies

- the existence of two arms, one open and one dual-closed arm in $B(n/4)$; the open arm connects the origin to $\partial B(n/4)$, the dual-closed arm connects a dual neighbor of the origin to $\partial B(n/4)^*$,
- the joint occurrence, inside $B(e, M/4)$, of $E(e,k)$ and a three-arm event (two open arms, one dual-closed) from e to distance $M/4$, and
- the existence, inside the region $B(n) \cap (B(p(e), 3n/4) \setminus B(e, 5M/4))$ of two arms, one open and one dual-closed, from $\partial B(e, 5M/4)$ to $\partial B(p(e), 3n/4)$. Here $p(e)$ is the orthogonal projection of the lower left endpoint of the edge e onto the boundary $\partial B(n)$.

By independence, this gives the upper bound

$$\begin{aligned}\mathbb{P}(E(e,k), e \in \gamma, C_0^c) &\leq \mathbb{P}(E(e,k), A_{3,OOC}(e, M/4), A_{2,OC}^{hp}(p(e); 5M/4, 3n/4), A_{2,OC}(n/4)) \\ &= \mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}^{hp}(p(e); 5M/4, 3n/4))\mathbb{P}(A_{2,OC}(n/4)).\end{aligned} \tag{2.24}$$

For the lower bound, we introduce an event F that implies the event $\{e \in \gamma\} \cap C_0^c$ when $M = \mathrm{dist}(e, \partial B(n))$. See Figure 2.4 for an illustration.

Definition 2.2. The event F is defined by the simultaneous occurrence of the following:

1. the edge e lies on an open arm ℓ_1 from the origin to $\partial B(n)$;
2. there is an open circuit dC with a defect in the annulus $B(n/8, n/4)$. The defect edge f is contained in a box $\mathcal{B} \subset \mathrm{Ann}(5n/32, 7n/32)$ of side length $n/16$;
3. there is a dual-closed arc $\mathfrak{f}$ outside dC in the annulus between the box $\mathcal{B}$ and a box $\mathcal{B}'$ with the same center as $\mathcal{B}$ and twice the side length of $\mathcal{B}$. This arc connects two edges that are dual to two open edges u and v on the circuit with defect dC. The edge u lies on the arc of dC between f and the endpoint of ℓ_1 on the circuit when the circuit is traversed in the counterclockwise direction. The edge v lies between f and the endpoint of ℓ_1 when the circuit is traversed in the clockwise direction. u and v are connected by an arc that consists only of open edges of dC inside $\mathcal{B}$ as well as the (closed) defect edge f;
4. a dual neighbor 0^* of the origin is connected to $\partial B(n)^*$ by a dual-closed path ℓ_2, which necessarily contains the dual edge f^*, the dual to the defect edge in the first item; Consider the closed curve formed by concatenating f^* and the portion of ℓ_2 from f^* to $\partial B(n)$, then following $\partial B(n)$ in the *counterclockwise* direction from the endpoint of ℓ_2 on $\partial B(n)$ to the endpoint of ℓ_1 on dC, and finally following dC in the clockwise direction until one reaches f. Denote by J the region bounded by this curve;
5. The dual edge e^* is connected to the arc $\mathfrak{f}$ by a dual-closed path lying inside J. In other words, $\mathfrak{f}$ is connected at e^* to the *clockwise* side of ℓ_1.

By the same argument as for Lemma 2.4.2, we obtain the following.

Lemma 2.4.3. *Let $e \in B(n)$ and suppose that $M = \operatorname{dist}(e, \partial B(n))$. The event F implies $\{e \in \gamma\} \cap C_0^c$.*

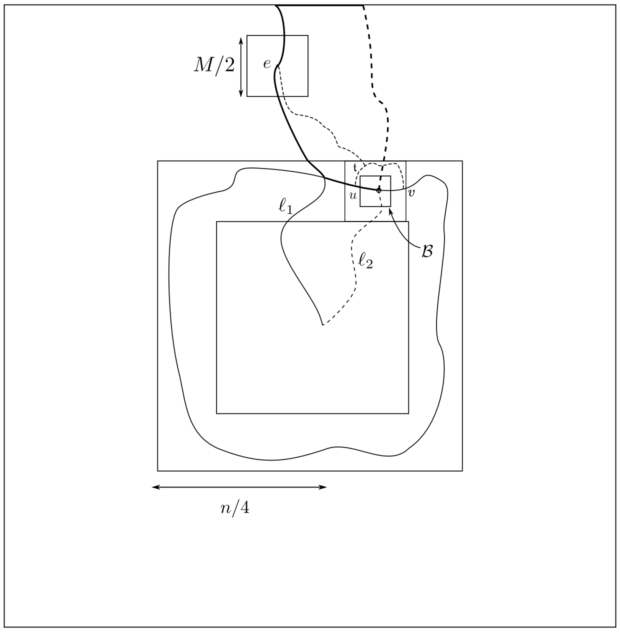


Figure 2.4: The construction in $B(n)$ on the event C_0^c in Case B ($M = \operatorname{dist}(e, \partial B(n))$). As in Figure 2.1, ℓ_1 and ℓ_2 are open, respectively dual-closed paths from the origin to distance n and the circuit in black represents a dual-closed circuit with defect. The construction forces $e \in \gamma$.

By standard gluing constructions illustrated in Figure 2.4, one has the following:

Lemma 2.4.4. *There is a constant $c > 0$ such that, if $M = \operatorname{dist}(e, \partial B(n))$ then, uniformly in the location of e, we have*

$$\mathbb{P}(F) \geq c\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4))\mathbb{P}(A_{2,OC}(n/4)).$$

Taken together, the last two lemmas give a lower bound for the denominator in (2.22) in case $M = \operatorname{dist}(e, \partial B(n))$:

$$\mathbb{P}(e \in \gamma, C_0^c) \geq c\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4))\mathbb{P}(A_{2,OC}(n/4)). \quad (2.25)$$

Combining the upper bound (2.24) with the lower bound (2.25), we find

$$\mathbb{P}(E(e,k) \mid e \in \gamma, C_0^c) \leq C\mathbb{P}(E(e,k) \mid A_3(e, M/4)).$$

2.4.2 Estimate on C_0

When there exists at least one open circuit around the origin (C_0 occurs), the path γ, defined in Section 2.2.3, consists of portions of the successive innermost open circuits $C_1, \ldots, C_{\mathcal{K}}$ around the origin, as well as open paths $\sigma_1, \ldots, \sigma_{\mathcal{K}+1}$ joining the origin to C_1, the successive circuits to each other, and finally $C_{\mathcal{K}}$ to $\partial B(n)$.

We write

$$\mathbb{P}(E(e,k) \mid e \in \gamma, C_0) = \frac{\mathbb{P}(E(e,k), e \in \gamma, C_0)}{\mathbb{P}(e \in \gamma, C_0)}.$$

The upper bound for the numerator will be obtained by a union bound along the decomposition

$$\gamma \subset \left(\cup_{m=1}^{\mathcal{K}} C_m \right) \cup \left(\cup_{m=1}^{\mathcal{K}+1} \sigma_m \right),$$

using estimates close to those obtained in Section 2.2 for the volume $\#\gamma$. For the denominator, we use

$$\mathbb{P}(e \in \gamma, C_0) \geq \mathbb{P}(e \in \sigma_1, C_0) + \mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0) + \mathbb{P}(e \in \sigma_{\mathcal{K}+1}, C_0). \tag{2.26}$$

We then obtain lower bounds on the terms on the right side of (2.26) by (RSW)/(FKG) constructions which force a given edge to belong to one of the portions of γ. As in the previous case, the constructions used depend on the location of the edge e in $B(n)$.

Case A: the edge e is closer to the origin than to $\partial B(n)$

In this case, we have

$$M = \mathrm{dist}(e, 0).$$

We write

$$\frac{\mathbb{P}(E(e,k), e \in \gamma, C_0)}{\mathbb{P}(e \in \gamma, C_0)} \leq \mathbb{P}(E(e,k) \mid e \in \sigma_1, C_0) + \mathbb{P}(E(e,k) \mid e \in \sigma_{\mathcal{K}+1}, C_0) + \mathbb{P}(E(e,k) \mid e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0) + \frac{\mathbb{P}(E(e,k), e \in \cup_{m=1}^{\mathcal{K}} C_m, C_0)}{\mathbb{P}(e \in \gamma, C_0)}. \tag{2.27}$$

Estimate for $e \in \sigma_1$. We estimate the conditional probability

$$\mathbb{P}(E(e,k) \mid e \in \sigma_1, C_0). \tag{2.28}$$

When $M = \mathrm{dist}(e, 0)$ (e is closer to the origin), the event $\{E(e,k), e \in \sigma_1, C_0\}$ implies

- a two-arm event in $B(M/4)$: an open arm from the origin to $\partial B(M/4)$, and a dual-closed arm from a dual neighbor of the origin to $\partial B(M/4)^*$;
- the occurrence of $E(e,k)$ and a three-arm event (two open, one dual-closed) in $B(e, M/4)$ and
- an open arm in the annulus $\mathrm{Ann}(2M, n)$ from $\partial B(2M)$ to $\partial B(n)$.

By independence of the regions involved, we obtain the upper bound

$$\mathbb{P}(e \in \sigma_1, C_0, E(e,k)) \leq \mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))\mathbb{P}(A_{1,O}(2M, n)). \tag{2.29}$$

We bound the denominator in (2.28), $\mathbb{P}(e \in \sigma_1, C_0)$, below using a construction analogous to that in Section 2.4.1, Case A, but inside an open circuit on scale M. See Figure 2.5. This forces $e \in \sigma_1$. Thus:

$$\mathbb{P}(e \in \sigma_1, C_0) \geq c\mathbb{P}(A_{2,OC}(M/4))\mathbb{P}(A_{3,OOC}(e, M/4))\, \mathbb{P}(A_{1,O}(2M, n)). \tag{2.30}$$

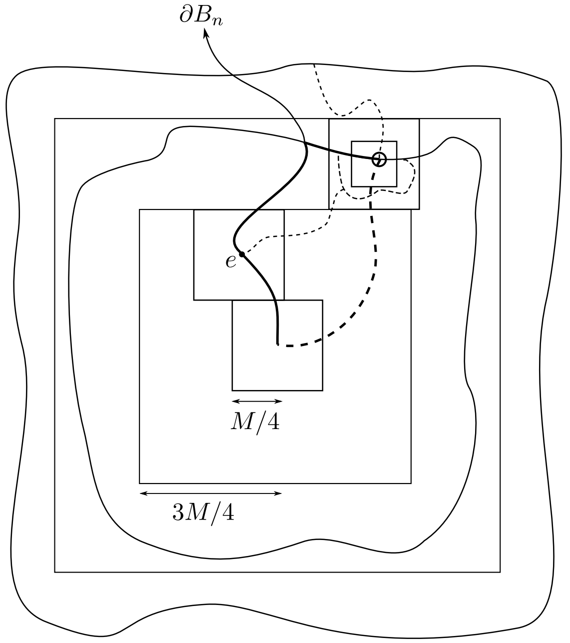


Figure 2.5: The construction for the lower bound (2.30).

Combining (2.29) and (2.30), we have

$$\mathbb{P}(E(e,k) \mid e \in \sigma_1, C_0) \leq C\mathbb{P}(E(e,k) \mid A_3(e, M/4)).$$

Estimate for $\sigma_{\mathcal{K}+1}$. We now estimate the conditional probability

$$\mathbb{P}(E(e,k) \mid e \in \sigma_{\mathcal{K}+1}, C_0).$$

When e is closer to the origin, the event $\{E(e,k), e \in \sigma_{\mathcal{K}+1}, C_0\}$ implies

- the joint occurrence, inside $B(e, M/4)$, of the event $E(e,k)$, and a three-arm event $A_{3,OOC}(e, M/4)$;
- a two-arm event (one open arm and one dual-closed arm) in the annular region $\mathrm{Ann}(2M, n)$; and
- the existence of an open arm connection in $B(M/4)$.

This gives the upper bound:

$$\mathbb{P}(E(e,k), e \in \sigma_{\mathcal{K}+1}, C_0) \leq \mathbb{P}(A_{1,O}(M/4))\mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(2M,n)).$$

By the construction illustrated in Figure 2.6, we also have the lower bound

$$\mathbb{P}(e \in \sigma_{\mathcal{K}+1}, C_0) \geq c\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(2M,n)). \tag{2.31}$$

Combining the previous two estimates, we obtain

$$\mathbb{P}(E(e,k) \mid e \in \sigma_{\mathcal{K}+1}, C_0) \leq C\mathbb{P}(E(e,k) \mid A_3(e, M/4)).$$

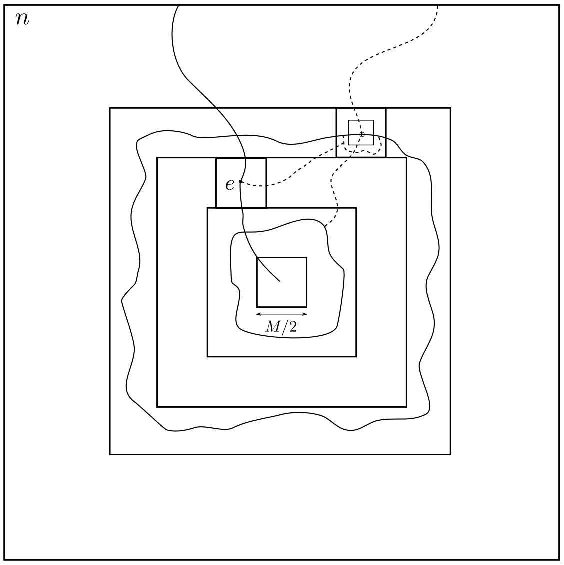


Figure 2.6: The construction for the lower bound (2.31).

The intermediate paths σ_m for $m = 2, \ldots, \mathcal{K}$. We estimate the conditional probability

$$\mathbb{P}(E(e,k) \mid e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0). \tag{2.32}$$

Here we use a modification of Lemma 2.2.3 from Section 2.2 to obtain an upper bound. Let $k' = k^{12/11}$, so that $k' \leq (M/4)^{54/55} \ll M$.

Lemma 2.4.5. *Suppose* $e \in \sigma_m$ *for some* $m = 2, \dots, \mathcal{K}$*, and both* $A_{3,OOC}(e,k)$ *and* $E(e,k)$ *occur. Then either (i) there exists* $\log(k') < \ell_1 \le \cdots \le \ell_R \le \lfloor \log(M/4) \rfloor$ *such that*

1. $E(e,k)$ *and* $A_{3,OOC}(e,k')$ *occur;*
2. *there are two open arms and one dual-closed arm, inside* $\mathrm{Ann}(e, k', 2^{\ell_1 - 1})$*, from* $\partial B(e,k')$ *to* $\partial B(e, 2^{\ell_1 - 1})$*;*
3. *for* $i \ge 2$*, if* $\ell_{i-1} < \lfloor \log(M/4) \rfloor$*, there are* $2i$ *open arms and one dual-closed arm from* $\partial B(e, 2^{\ell_{i-1}})$ *to* $\partial B(e, 2^{\ell_i - 1})$*;*
4. *if* $\ell_R < \lfloor \log(M/4) \rfloor$*, there are* $2R + 2$ *disjoint open arms from* $\partial B(e, 2^{\ell_R})$ *to* $\partial B(e, M(e))$*; and*
5. *there is an open arm from the origin to distance* $M/4$ *and one open arm from* $\partial B(2M)$ *to* $\partial B(n)$*.*

or (ii) there are $0 = \ell_0 \le \ell_1 \le \cdots \le \ell_R \le \lfloor \log(M/4) \rfloor$ *and* $\ell_1 \le \log(k')$ *such that*

1. $A_{3,OOC}(e, 2^{\ell_1 - 1})$ *occurs;*
2. *for* $i \ge 2$*, if* $\ell_{i-1} < \lfloor \log(M/4) \rfloor$*, there are* $2i$ *disjoint open arms and one dual-closed arm from* $\partial B(e, 2^{\ell_{i-1}})$ *to* $\partial B(e, 2^{\ell_i - 1})$*;*
3. *if* $\ell_R < \lfloor \log(M/4) \rfloor$*, there are* $2R + 2$ *disjoint open arms from* $\partial B(e, 2^{\ell_R})$ *to* $\partial B(e, M/4)$*; and*
4. *there is an open arm from the origin to distance* $M/4$ *and one open arm from* $\partial B(2M)$ *to* $\partial B(n)$*.*

The second case occurs when there is no space between C_m and C_{m+1} to fit the box $B(e,k')$.

The estimate in case (*ii*) uses an adaptation of the idea in Lemma 2.2.3:

$$
\begin{aligned}
&\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0, \text{case } (ii)) \\
&\qquad \le C\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)) \\
&\qquad\quad \times \sum_{0=l_0 \le \ell_1 \le \log(k')} \pi_3(2^{\ell_1}) \sum_{\ell_1 \le \ell_2 \le \ldots \le \ell_R} \prod_{i=2}^{R} \pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2R+2}(2^{\ell_R}, M/4).
\end{aligned}
$$

For $i = 3, \ldots, R$, we use (Reimer) in the form

$$\pi_{2i+1}(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \le \pi_5(2^{\ell_{i-1}}, 2^{\ell_i - 1}) \pi'_{2i-4}(2^{\ell_{i-1}}, 2^{\ell_i - 1}).$$

By similar computations to those in Section 2.2.3, we obtain a bound for the inner sum over $\ell_1, \ldots, \ell_R$ of $c^{R-2}\pi_5(2^{\ell_1}, 2^{\ell_R})$. Using (Reimer), (QM), and (2.12), we are left with

$$
\begin{aligned}
&\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0, \text{case } (ii)) \\
\le\ & C\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)) \sum_{0 \le \ell_1 \le \log(k')} \pi_3(2^{\ell_1}) \sum_{\ell_R \le \log(M/4)} \pi_5(2^{\ell_1}, 2^{\ell_R}) \pi_3(2^{\ell_R}, M/4) \left(\frac{2^{\ell_R}}{M/4}\right)^{\epsilon} \\
\le\ & C\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n))\pi_3(M/4) \sum_{\ell_R \le \log(M/4)} \left(\frac{2^{\ell_R}}{M/4}\right)^{\epsilon} \sum_{\ell_1 \le \log(k')} 2^{-2\alpha(\ell_R - \ell_1)} \\
\le\ & C\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n))\pi_3(M/4) \left(\frac{4k'}{M}\right)^{2\alpha}.
\end{aligned}
$$

Since k is chosen to satisfy $k \le (M/4)^{9/10}$, the right side of the final quantity above is then bounded by

$$C\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n))\pi_3(M/4)(M/4)^{-r}$$

for some $r > 0$.

For the first case (*i*) of Lemma 2.4.5, similarly to Lemma 2.2.3 and in the previous case, the probability of items (2) – (4) is bounded by

$$\sum_{\lceil \log(k') \rceil < \ell_1 \le \cdots \le \ell_R}^{\lfloor \log(M/4) \rfloor} \mathbb{P}(A_{3,OOC}(e, k', 2^{\ell_1 - 1}))) \prod_{i=1}^{R-1} \pi_{2i+1}(2^{\ell_i}, 2^{\ell_{i+1}-1})) \pi'_{2R+2}(2^{\ell_R}, M/4) \le C\pi_3(k', M/4).$$

We now claim:

$$\begin{aligned} &\mathbb{P}(E(e,k), A_{3,OOC}(e,k'))\mathbb{P}(A_{3,OOC}(k',M/4)) \\ \leq\ & C\mathbb{P}(E(e,k), A_{3,OOC}(e,M/4)) + C(k/k')^{\delta}\pi_3(M/4), \end{aligned} \tag{2.33}$$

for some $\delta > 0$. Note that if we remove the event $E(e,k)$ from the probability on both sides, (2.33) follows immediately from arm separation and gluing, as summarized in Section 1.5.1. However, the presence of the event $E(e,k)$ could bias the probability of $A_{3,OOC}(e,k')$ and thus prevent arm separation on the boundary of $B(e,k')$. This can be circumvented by another classical technique: decoupling at interfaces. The interfaces we consider here are dual-closed circuits around the origin with two open edge-defects. We will be brief in our presentation. We refer the reader to [DS11, Section 6], [DHS16, Proposition 6.4], and [DHS21, Lemma 4.4] for related estimates proved using this technique.

We introduce the event $S_{k,k'}$ that there is a dual-closed circuit with two open defects in some annulus $\mathrm{Ann}(e;2^j,2^{j+1})$, $j = \lfloor \log(2k) \rfloor, \ldots, \lfloor \log(k') \rfloor$. We begin by excluding the event $S_{k,k'}$. Using the argument in [DHS21, Theorem 4.1], there is a $\delta > 0$ such that

$$\mathbb{P}(S_{k,k'}^c \mid A_{3,OOC}(e,k')) \leq C(k/k')^{\delta}.$$

Thus,

$$\mathbb{P}(E(e,k), A_{3,OOC}(e,k')) \leq \mathbb{P}(E(e,k), A_{3,OOC}(e,k'), S_{k,k'}) + C(k/k')^{\delta}\pi_3(k'). \tag{2.34}$$

For any dual-closed circuit D with two open defects, the two defects partition the circuit into two arcs. For the deterministic ordering fixed at the beginning of Section 2.2.1, let the arc with the smallest edge be denoted $\mathrm{Arc}_1(D)$, and the other $\mathrm{Arc}_2(D)$. We define $X_-(D,i)$ be the event that e is connected to D by three disjoint arms: two open arms from e to the two defects, and a dual-closed arm to $\mathrm{Arc}_i(D)$. Similarly, we let $X_+(D,k',i)$ be the event that D is connected to the boundary of the box $B(e,k')$ by three disjoint arms: two open arms emanating from the two defects, and a dual-closed arm from $\mathrm{Arc}_i(D)$.

We then partition the event $\{E(e,k), A_{3,OOC}(e,k'), S_{k,k'}\}$ at the innermost circuit with defects $\mathcal{D}$ in the region $B(e,2k,k')$:

$$\mathbb{P}(E(e,k), A_{3,OOC}(e,k'), S_{k,k'}) = \sum_{D} \sum_{i=1,2} \mathbb{P}(E(e,k), \mathcal{D} = D, X_-(D,i))\mathbb{P}(X_+(D,k',i)). \tag{2.35}$$

By the 3-arm analogue of external arm separation as in [DS11, Lemma 6.2], we have

$$\mathbb{P}(X_+(D,k',i)) \le C\mathbb{P}(\tilde{X}_+(D,k',i)), \tag{2.36}$$

where in the $\tilde{X}_+(D,k',i)$ the arms from the circuit with defects to $\partial B(e,k')$ are well-separated and have given landing sites. See [Nol08, Definitions 7 & 9] for the relevant definitions. Inserting (2.36) into (2.35), we have now showed that

$$\mathbb{P}(E(e,k), A_{3,OOC}(e,k'), S_{k,k'}) \le C\mathbb{P}(E(e,k), \tilde{A}_{3,OOC}(e,k')),$$

where the arms on $\partial B(e,k')$ in the event $\tilde{A}_{3,OOC}(e,k')$ are well-separated and have their endpoints in specified intervals on the boundary. We can now use standard gluing to obtain that

$$\mathbb{P}(E(e,k), \tilde{A}_{3,+,OOC}(e,k'))\pi_3(k',M) \le C\mathbb{P}(E(e,k), A_{3,OOC}(e,M)).$$

The claim (2.33) now follows from this and (2.34), using $\pi_3(k')\pi_3(k',M) \le C\pi_3(e,M)$.

Together with items (1) and (5), we have the estimate

$$\begin{aligned}
&\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}}\sigma_m, C_0, \text{case } (i)) \\
&\quad \le \mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n))\mathbb{P}(E(e,k), A_{3,OOC}(e,k'))C\pi_3(k',M/4) \\
&\quad \le C'\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n))\left(\mathbb{P}(E(e,k), A_{3,OOC}(e,M/4)) + (k/k')^{\delta}\pi_3(M/4)\right).
\end{aligned} \tag{2.37}$$

Thus, we obtain the upper bound

$$\begin{aligned}&\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0)\\&\leq C\left(\mathbb{P}(E(e,k), A_{3,OOC}(M/4)) + M^{-r}\pi_3(M/4)\right)\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)).\end{aligned} \tag{2.38}$$

For the denominator in (2.32), we have the following lower bound, a consequence of the construction illustrated in Figure 2.7:

$$\mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0) \geq c\mathbb{P}(A_{3,OOC}(M/4))\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)). \tag{2.39}$$

Finally, by (2.38) and (2.39), we have

$$\frac{\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0)}{\mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0)} \leq C(\mathbb{P}(E(e,k) \mid A_3(e,M/4)) + M^{-r}).$$

Definition 2.3. For later use, we denote the event defined in Lemma 2.4.5, case (ii) by $\mathcal{A}_R(e, M/4)$, where it is always understood that R is chosen such that (2.12) holds. The previous computations show that

$$\mathbb{P}(\mathcal{A}_R(e, M/4)) \leq cM^{-r}\pi_3(M/4),$$

for some $r > 0$.

The circuits C_m, $m = 1, \ldots, \mathcal{K}$. This case is dealt with similarly to the probability of the event $\{e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0\}$. The upper bound in this case is of the same form as (2.38):

$$\begin{aligned}&\mathbb{P}(E(e,k), e \in \cup_{m=1}^{\mathcal{K}} C_m, C_0)\\&\quad\leq C(\mathbb{P}(E(e,k), A_{3,OOC}(e,M/4)) + M^{-r}\pi_3(e,M/4))\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)).\end{aligned}$$

To control the denominator in the last term of (2.27), we use the lower bound (2.39)

$$\mathbb{P}(e \in \gamma, C_0) \geq \mathbb{P}(e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0) \geq c\mathbb{P}(A_{3,OOC}(e,M/4))\mathbb{P}(A_{1,O}(M/4))\mathbb{P}(A_{1,O}(2M,n)).$$

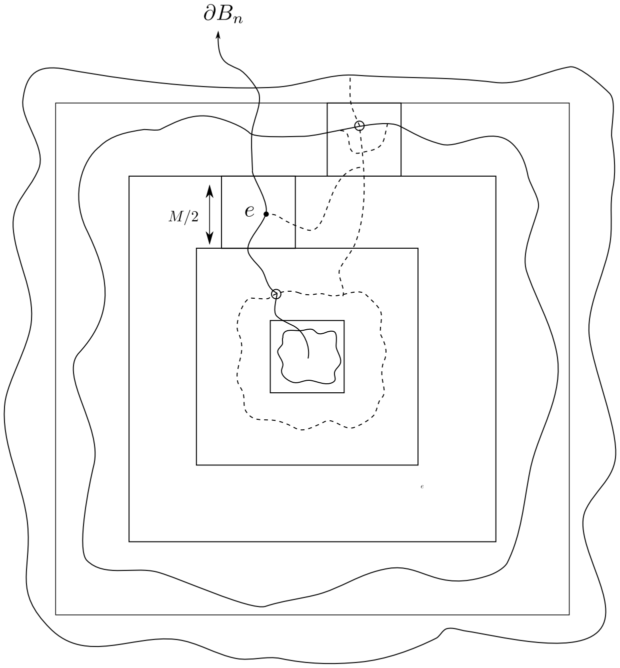


Figure 2.7: The construction for the lower bound (2.39).

Case B: e is closer to $\partial B(n)$ than the origin.

In this case, we estimate the ratio

$$\frac{\mathbb{P}(E(e,k), e \in \gamma, C_0)}{\mathbb{P}(e \in \gamma, C_0)}$$

using the lower bound (2.26) instead of (2.27).

Estimate for $e \in \sigma_1$. We estimate

$$\frac{\mathbb{P}(E(e,k), e \in \sigma_1, C_0)}{\mathbb{P}(e \in \gamma, C_0)}.$$

When the edge e is closer to the boundary ($M = \operatorname{dist}(e, \partial B(n))$), the following occur:

1. three arms, two open and one dual-closed, inside $B(e, M/4)$ from e to $\partial B(e, M/4)$;

2. four arms, three open and one dual-closed, in the semi-annular region $B(n) \cap \mathrm{Ann}(p(e); 5M/4, 3n/4)$ and
3. two arms, one open and one dual-closed, from the origin to distance $n/4$.

By independence, this results in the bound

$$\mathbb{P}(e \in \sigma_1, C_0) \leq \mathbb{P}(A_{3,OOC}(e, M/4), E(e,k))\mathbb{P}(A_{2,OC}(n/4))\mathbb{P}(A^{hp}_{2,OC}(5M/4, 3n/4)). \quad (2.40)$$

For the lower bound, we use the construction illustrated in Figure 2.8 to obtain the estimate

$$\begin{aligned}\mathbb{P}(e \in \gamma, C_0) &\geq \mathbb{P}(e \in \sigma_{\mathcal{K}+1}) \\ &\geq c\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A_{1,O}(n/4))\mathbb{P}(A^{hp}_{2,OC}(5M/4, 3n/4)) \qquad (2.41) \\ &\geq c\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A_{2,OC}(n/4))\mathbb{P}(A^{hp}_{2,OC}(5M/4, 3n/4)).\end{aligned}$$

Using (2.40) and (2.41), we have

$$\frac{\mathbb{P}(E(e,k), e \in \sigma_1, C_0)}{\mathbb{P}(e \in \gamma, C_0)} \leq C\mathbb{P}(E(e,k) \mid A_3(e, M/4)).$$

Estimate for $e \in \sigma_{\mathcal{K}+1}$. When e is closer to the boundary ($M = \mathrm{dist}(e, \partial B(n))$), we use the following lemma to obtain an upper bound for the probability $\mathbb{P}(E(e,k), e \in \sigma_{\mathcal{K}+1}, C_0)$.

Lemma 2.4.6. *Suppose $e \in \sigma_{\mathcal{K}+1}$ and $M = \mathrm{dist}(e, \partial B(n))$. If $E(e,k) \cap C_0$ occurs, then there exists $\log(2k) < j \leq l \leq \log(3n/4)$ such that*

1. *$E(e,k)$ and the three-arm event $A_{3,OOC}(e, M/4)$ occur;*
2. *there are two arms in $B(n) \cap \mathrm{Ann}(p(e); 5M/4, 2^{j-1})$, one open and one dual-closed, from $\partial B(p(e), 5M/4)$ to $\partial B(p(e), 2^{j-1})$;*
3. *there are three arms, two open and one dual-closed, in $B(n) \cap \mathrm{Ann}(p(e); 2^j, 2^{l-1})$ from $\partial B(p(e), 2^j)$ to $\partial B(p(e), 2^{l-1})$;*

4. *there are four open arms inside* $B(n) \cap \mathrm{Ann}(p(e); 2^l, 3n/4)$ *and*

5. *there is an open arm in* $B(n/4)$ *from the origin to* $\partial B(n/4)$.

Proof. On C_0, enumerate the successive innermost circuits as $\mathcal{C}_1, \ldots, \mathcal{C}_{\mathcal{K}}$. If $e \in \sigma_{\mathcal{K}+1}$, then e and $p(e)$ lie outside $\mathcal{C}_{\mathcal{K}}$. We let j be the least integer such that $B(p(e), 2^j) \cap \mathcal{C}_{\mathcal{K}} \neq \varnothing$, and $l \geq j$ be the least l such that $B(p(e), 2^l) \cap \mathcal{C}_{\mathcal{K}-1} \neq \varnothing$ if $\mathcal{K} \geq 2$. Otherwise, we set $l = \lceil \log(3n/4) \rceil$. It is now easy to check that the claims regarding the arm events hold. □

Decomposing according to the distances $2^j, 2^l$, we estimate the ratio

$$\frac{\mathbb{P}(E(e,k), e \in \sigma_{\mathcal{K}+1}, C_0)}{\mathbb{P}(A_{1,O}(n/4))\mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))}$$

by the sum

$$\begin{aligned}
&\sum_{\log(2k)<j\leq l\leq \log(n/4)} \mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 2^{j-1}))\mathbb{P}(A^{hp}_{3,OOC}(p(e); 2^j, 2^{l-1}))\mathbb{P}(A^{hp}_{4,OOOO}(2^l, 3n/4)) \\
&\leq C \sum_{\log(2k)<j\leq l\leq \log(n/4)} \mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 2^{j}))\mathbb{P}(A^{hp}_{2,OC}(p(e); 2^j, 2^{l}))\frac{2^{\epsilon j}}{2^{\epsilon l}}\mathbb{P}(A^{hp}_{2,OO}(2^l, 3n/4))\frac{2^{2\epsilon l}}{n^{2\epsilon}} \\
&\leq C' \mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4)).
\end{aligned} \tag{2.42}$$

This gives the upper bound

$$\mathbb{P}(E(e,k), e \in \sigma_{\mathcal{K}+1}, C_0) \leq C' \mathbb{P}(A_{1,O}(n/4))\mathbb{P}(E(e,k), A_{3,OOC}(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4)). \tag{2.43}$$

The main estimate for the denominator $\mathbb{P}(e \in \gamma, C_0)$ in this case is

$$\mathbb{P}(e \in \gamma, C_0) \geq \mathbb{P}(e \in \sigma_{\mathcal{K}+1}, C_0) \geq c\mathbb{P}(A_{1,O}(n/4))\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4)). \tag{2.44}$$

This is obtained by the construction illustrated in Figure 2.8. Combining (2.43)

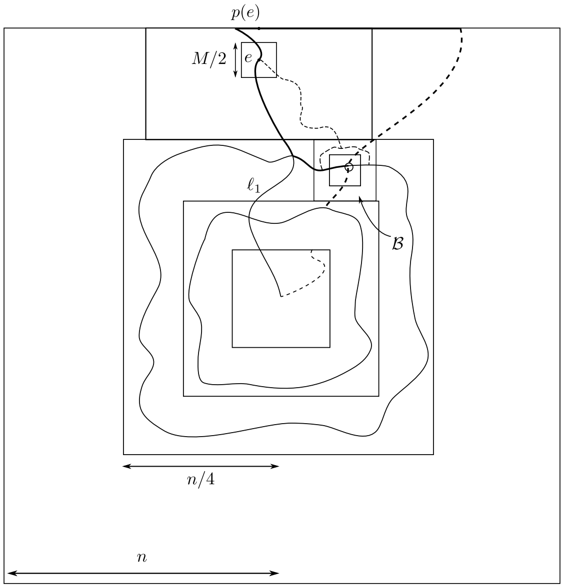


Figure 2.8: The construction near the boundary used to obtain the lower bound (2.44).

and (2.44), we obtain the desired estimate:

$$\mathbb{P}(E(e,k) \mid e \in \sigma_{\mathcal{K}+1}, C_0) \leq C\mathbb{P}(E(e,k) \mid A_3(e, M/4)), \quad \text{if } M = \operatorname{dist}(e, \partial B(n)).$$

Estimate for $e \in \cup_{m=2}^{\mathcal{K}} \sigma_m$. When $M = \operatorname{dist}(e, \partial B(n))$, we have the upper bound

$$\begin{aligned} \mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0) &\leq (C\mathbb{P}(A_{3,OOC}(e, M/4), E(e,k)) + M^{-r}\pi_3(e, M/4))) \\ &\quad \cdot \mathbb{P}(A_{1,O}(e, n/4))\mathbb{P}(A_{2,OC}^{hp}(p(e); 5M/4, 3n/4)). \end{aligned} \tag{2.45}$$

This estimate follows from the following lemma, which is proved along the same lines as Lemma 2.2.3 in Section 2.2 and Lemma 2.4.6 in Section 2.4.2.

Lemma 2.4.7. *Suppose $M = \operatorname{dist}(e, \partial B(n))$. Then,*

1. *(a) either the event $E(e,k)$ and the three-arm event $A_{3,OOC}(e, M/4)$ occur;*

   (b) or the event $\mathcal{A}_R(e, M/4)$ in Definition (2.3) *occurs;*

2. *there are two open arms in $B(n) \cap \operatorname{Ann}(p(e); 5M/4, 3n/4)$ and;*

3. *there is an open arm in* $B(e, n/4)$ *from the origin to* $\partial B(e, n/4)$.

Proof sketch. The proof for item (1) is very similar to the proof for Lemma 2.2.3. If there is no dual-closed arm to distance $M/4$, then C_m or C_{m+1} intersects $B(e, M/4)$. We can then find a sequence of scales $\ell_1 \leq \ell_2 \leq \cdots \leq \ell_R$ such that the annulus on each scale contains two more open arms than the preceding one. For item (2), there are at least two arms consisting of portions of C_{m+1} in $B(n) \cap \mathrm{Ann}(p(e); 5M/4, 3n/4)$. □

Applying the lemma and decomposing according to the distance j as in (2.37), (2.42), we obtain the estimate (2.45).

For the denominator, we use the lower bound (2.44) established for the event $\{e \in \sigma^{\mathcal{K}+1}\}$.

$$\mathbb{P}(e \in \sigma_{\mathcal{K}+1}, C_0) \geq c\mathbb{P}(A_{3,OOC}(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4))\mathbb{P}(A_{1,O}(n/4)). \quad (2.46)$$

Combining (2.45) and (2.46), we find

$$\frac{\mathbb{P}(E(e,k), e \in \cup_{m=2}^{\mathcal{K}} \sigma_m, C_0)}{\mathbb{P}(e \in \gamma, C_0)} \leq C(\mathbb{P}(E(e,k) \mid A_{3,OOC}(e, M/4)) + M^{-r}).$$

Estimate for $e \in \cup_{m=1}^{\mathcal{K}+1} C_m$. This case can be handled very similarly to that in the previous section. By a variant of Lemma 2.4.7, we obtain the upper bound:

$$\begin{aligned}\mathbb{P}(E(e,k), e \in \cup_{m=1}^{\mathcal{K}+1} C_m, C_0) \leq C(&\mathbb{P}(E(e,k), A_3(e, M/4)) \\ &+ M^{-r}\pi_3(e, M/4))\mathbb{P}(A^{hp}_{2,OC}(p(e); 5M/4, 3n/4))\mathbb{P}(A_{1,O}(n/4)).\end{aligned} \quad (2.47)$$

We use the lower bound established for $\mathbb{P}(e \in \gamma, C_0)$ in (2.44). Combined with the upper bound (2.47), we obtain

$$\frac{\mathbb{P}(E(e,k), e \in \cup_{m=1}^{\mathcal{K}+1} C_m, C_0)}{\mathbb{P}(e \in \gamma, C_0)} \leq C\mathbb{P}(E(e,k) \mid A_{3,OOC}(M/4)).$$

CHAPTER 3

CHEMICAL DISTANCE ON THE PLANAR RANDOM CLUSTER MODEL

3.1 Introduction

In this chapter, we consider the random cluster model on the square lattice $\mathbb{Z}^2$ with fixed cluster weight $q \in [1,4]$ and critical edge weight $p = p_c(q)$.

From the introduction (Section 1.1), it is known that when $q \in [1,4]$, the random cluster model exhibits a continuous phase transition [DCST17] as well as enjoys positive association ((FKG) inequality). These facts combined with recent developments in RSW-type quad-crossing probabilities [DCMT21] allow us to pursue an upper bound for the horizontal chemical distance in the form of (1.4). Let $\mathcal{H}_n$ be the event that there exists a horizontal open crossing between the left and right sides of the box $B(n)$ and H_n the length of the shortest such crossing. Let $\mathbb{E}$ denote the expectation with respect to the random cluster measure $\phi^{\xi}_{p_c,q,B(n)}$.

Theorem 3.1.1. *For any boundary condition ξ, there is a $\delta > 0$ and a constant $C > 0$ independent of n such that*

$$\mathbb{E}[H_n \mid \mathcal{H}_n] \leq Cn^{2-\delta}\phi^{\xi}_{p_c,q,B(n)}(A_3(n)).$$

Furthermore, we also obtain an upper bound for the radial chemical distance of the same form as that derived in Chapter 2. Recall that S_n denotes the length of the shortest crossing connecting the origin to $\partial B(n)$.

Theorem 3.1.2. *There exist some $\delta > 0$ and constant $C > 0$ independent of n such that*

$$\mathbb{E}[S_n \mid 0 \leftrightarrow \partial B(n)] \leq Cn^{2-\delta}\phi^{\xi}_{p_c,q,B(n)}(A_3(n)).$$

3.1.1 Organization

The general strategy to prove Theorem 3.1.1 aligns with [DHS21], which we outline in Section 3.2 to provide context. We aim to point to the similarities and highlight the differences between the proofs for the two models to ensure readability while minimizing the amount of repetition. In Section 3.3 and 3.4, we provide the proofs of a large deviation bound conditional on a three-arm event and the random-cluster analogue of the strong arm separation lemma. Both proofs involve strategic applications of the domain Markov property (DMP) to circumvent the lack of independence.

We can extend the result of the main theorem to the radial chemical distance, following the approach in [SR22] for Bernoulli percolation. Since most of the arguments in [SR22] rely solely on independence and gluing constructions, they extend to the random cluster model when substituted with (DMP) and gluing constructions detailed in Section 1.5.1. The remaining challenge is to find a way to bound the probability of a specific event without the use of Reimer's inequality. Such a method will be detailed in Section 3.5.

3.1.2 Lower bound for the random cluster model

The Aizenman-Burchard lower bound (1.1) applies when the following criterion on the probability of simultaneous traversal of separated rectangles is satisfied: A collection of rectangles R_j is called *well-separated* when the distance between any two rectangles is at least as large as the diameter of the larger. The following criterion is formulated for the random cluster measure.

Hypothesis ([AB99]). *Fix $\delta > 0$. There exist $\sigma > 0$ and some $\rho < 1$ with which for every collection of k well-separated rectangles, $A_1, \ldots, A_k$, of aspect ratio σ and lengths $\ell_1, \ldots, \ell_k \geq \delta n$,*

$$\phi^{\xi}_{B(n)} \begin{pmatrix} A_1, \ldots, A_k \text{ are traversed (in} \\ \text{the long direction) by segments} \\ \text{of an open crossing} \end{pmatrix} \leq K\rho^k.$$

This hypothesis is satisfied as a consequence of the weak (polynomial) mixing property [DC13]: There exists $\alpha > 0$ such that for any $2k < m < n$ and any event A depending only on edges in $B(k)$ and event B depending only on the edges in $\mathrm{Ann}(m, n)$,

$$|\phi^{\xi}_{B(n)}(A \cap B) - \phi^{\xi}_{B(n)}(A)\phi^{\xi}_{B(n)}(B)| \leq \left(\frac{k}{m}\right)^{\alpha} \phi^{\xi}_{B(n)}(A)\phi^{\xi}_{B(n)}(B),$$

uniform in the boundary condition ξ.

3.2 Proof outline

As an extension to the results derived in [DHS21], the proof of the main result follows the same strategy with modifications in certain arguments. For completeness, we outline the proof with an emphasis on the present application and point to the main differences. An alternate, more detailed outline is offered in [DHS21, Section 2].

The proof is essentially divided into three steps: In the first step, we construct shortcuts around edges on the lowest crossing and show that the existence of such shortcuts has a "good" probability. The second step uses an iterative scheme to improve upon shortcuts. Finally, we find the maximal collection of disjoint shortcuts and sum up the total savings.

Step 0: The lowest crossing ℓ_n

The estimate on the length of ℓ_n relies on the observation that ℓ_n consists only of three-arm points: since ℓ_n is the lowest crossing, by duality, from every edge e in ℓ_n there are two disjoint open arms and a dual-closed arm to distance $\mathrm{dist}(e, B(n))$. In conjunction with some smoothness control, we have

$$\mathbb{E}[\#\ell_n \mid \mathcal{H}_n] \leq Cn^2\pi_3(n).$$

Step 1: Construction of shortcuts

For any $\epsilon > 0$, an edge e on the lowest crossing ℓ_n, we look for an arc r over e that saves at least $(1/\epsilon - 1)\#r$ edges. The event $\hat{E}_k(e) = \hat{E}_k(e, \epsilon, \delta)$ describes such an arc circumventing e on scale k. The exact definition of $\hat{E}_k$ is quite involved, see [DHS21, Section 5]. We only state the properties and results relevant to the argument:

1. $\hat{E}_k(e)$ depends only on $\mathrm{Ann}(e; 2^k, 2^K)$ where $K = k + \lfloor \log(1/\epsilon) \rfloor$;
2. For each $e \in \ell_n$, $\hat{E}_k(e, \epsilon, \delta)$ implies the existence of an δ-shortcut around e. That is, there is an open arc $r \subset B(e, 3 \cdot 2^k)$ such that r only intersects with ℓ_n at its two endpoints $u(e)$ and $v(e)$ and

$$\frac{\#r}{\#\tau} \leq \delta;$$

where τ denotes the portion of ℓ_n between $u(e)$ and $v(e)$. See [DHS21, Proposition 5.4].

3. $\hat{E}'_k$ is a similar event to $\hat{E}_k$ on scale k that relates to a "U-shaped region" and $\mathfrak{s}_k$ is the shortest path in the U-shaped region. If for some $\epsilon \in (0, \frac{1}{2})$, $\delta > 0$, and $k \geq 1$,

   $$\mathbb{E}[\#\mathfrak{s}_k \mid \hat{E}'_k] \leq \delta 2^{2k} \pi_3(2^k)$$

   holds, then for all $L \geq 1$,

   $$\phi^{\xi}_{B(n)}(\hat{E}_k(e, \epsilon, \delta) \mid A_3(e, 2^L)) \geq c\epsilon^4 \quad \text{for all } L \geq 1. \tag{3.1}$$

   See [DHS21, Equation (5.29)].

Property (2) relies mostly on topological considerations and therefore applies to the random cluster model. For property (3), we refrain from elaborating further despite the original proof using, at times, independence, generalized (FKG), and gluing constructions. This is because we feel the techniques to convert these arguments for the random cluster model are sufficiently represented in the proofs that we do include, especially in that of the following proposition; and completely reproducing all necessary parts of the proof of property (3) would require lots of notation and mostly verbatim steps that translate directly for the random cluster model.

The key result in this step is that the probability that no shortcut exists for any scale k is small:

Proposition 3.2.1. *There is a constant $\hat{c}$ such that if $\delta_j > 0$, $j = 1, \ldots, L$, is a sequence of parameters such that for some $\epsilon \in (0, \frac{1}{4})$,*

$$\mathbb{E}[\#\mathfrak{s}_j \mid \hat{E}'_j] \leq \delta_j 2^{2j} \pi_3(2^j), \tag{3.2}$$

then for any $L' < L$,

$$\phi^{\xi}_{B(n)}\left(\cap_{j=L'}^{L}\hat{E}_j(e,\epsilon,\delta_j)^c \mid A_3(e,2^L)\right) \le 2^{-\hat{c}\frac{\epsilon^4}{\log(1/\epsilon)}(L-L')}.$$

This estimate follows from a large deviation bound conditioned on a three-arm event. The argument uses an application of the strong separation lemmas from [DS11] to decorrelate $E_k(e)$ and $E_{k'}(e)$ if $|k-k'| \ge C\log\frac{1}{\epsilon}$. In this chapter, we recreate all the above arguments as their extensions to the random cluster model are nontrivial due to the lack of independence.

Let $\hat{\mathfrak{C}}_k$ denote the existence of a large stack of circuits in $\mathrm{Ann}(2^{10kN}, 2^{10(k+1)N})$, see Definition 3.1.

Theorem 3.2.2 ([DHS21], Theorem 4.1)**.** *There exist universal $c > 0$ and $N_0 > 0$ such that for any $N \ge N_0$, any $L', L \ge 0$ satisfying $L - L' \ge 40$, and any event E_k satisfying*

- *(A) E_k depends on the status of the edges in $\mathrm{Ann}(2^{kN}, 2^{(k+1)N})$ and*
- *(B) there exists a uniform constant $c_0 > 0$ such that $\phi^{\xi}_{B(n)}(E_{10k+5} \cap \hat{\mathfrak{C}}_k \mid A_3(2^L)) \ge c_0$ for all $n \ge 0$ and $0 \le k \le \frac{L}{10N} - 1$.*

Then,

$$\phi^{\xi}_{B(n)}\left(\sum_{k=\lceil \frac{L'}{10N}\rceil}^{\lfloor \frac{L}{10N}\rfloor - 1} \mathbf{1}\{E_{10k+5}, \hat{\mathfrak{C}}_k\} \le cc_0\frac{L-L'}{N} \;\middle|\; A_3(2^L)\right) \le \exp(-cc_0\frac{L-L'}{N}).$$

Proof of Proposition 3.2.1 subject to Theorem 3.2.2. Let $E_k = \hat{E}_{kN}$. By property (3), the combination of (3.2) with the observation that the occurrence of a circuit in $\mathrm{Ann}(2^{(10k+i)N}, 2^{(10k+i+1)N})$ conditional on a three-arm event has constant probability due to (RSW) and gluing constructions (see Proposition 1.5.1) implies that

$$\phi^{\xi}_{B(n)}(E_{10k+5} \cap \hat{\mathfrak{C}}_k \mid A_3(2^L)) \ge c'\epsilon^4$$

for $0 \leq k \leq \frac{L}{10N} - 1$. Note that c' is now uniform in k. We observe the following chain of set inclusions with changes of indices including $j = \ell N$ in the equality and $\ell = 10k$ in the first inclusion:

$$\bigcap_{j=L'}^{L} \hat{E}_j^c = \bigcap_{\ell=\lceil \frac{L'}{N} \rceil}^{\lfloor \frac{L}{N} \rfloor} E_\ell^c \subset \bigcap_{k=\lceil \frac{L'}{10N} \rceil}^{\lfloor \frac{L}{10N} \rfloor - 1} (E_{10k+5})^c \subset \left\{ \sum_{k=\lceil \frac{L'}{10N} \rceil}^{\lfloor \frac{L}{10N} \rfloor - 1} \mathbf{1}\{E_{10k+5}, \hat{\mathfrak{C}}_k\} = 0 \right\}.$$

Thus, applying Theorem 3.2.2 by choosing $N = \lfloor \log(1/\epsilon) \rfloor$ and $L - L' \geq 40$, we obtain

$$\begin{aligned} \phi_{B(n)}^{\xi}\left(\cap_{k=L'}^{L} \hat{E}_k(e, \epsilon, \delta_j)^c \,\middle|\, A_3(e, 2^L) \right) &\leq \phi_{B(n)}^{\xi}\left(\sum_{k=\lceil \frac{L'}{10N} \rceil}^{\lfloor \frac{L}{10N} \rfloor - 1} \mathbf{1}\{E_{10k+5}, \hat{\mathfrak{C}}_k\} = 0 \,\middle|\, A_3(2^L) \right) \\ &\leq \exp(-\frac{cc'\epsilon^4}{\log(1/\epsilon)}(L - L')). \end{aligned}$$

□

Step 2: Iteration in the “U-shaped region”

In this step, we inductively improve the length of the “best possible” shortcuts for a fixed scale. The function of this step is to ensure that (3.2) is satisfied.

Proposition 3.2.3 ([DHS21], Proposition 7.1)**.** *There exist constants C, C' such that for any $\epsilon > 0$ sufficiently small, $L \geq 1$, and $2^k \geq (C\epsilon^{-4}(\log(1/\epsilon)^2))^L$, we have*

$$\mathbb{E}[\#\mathfrak{s}_k \mid \hat{E}_k'] \leq (C'\epsilon^{1/2})^L 2^{2k} \pi_3(2^k).$$

The constructions are detailed in [DHS21, Section 6 & 7] and we only give the high level heuristics. In step 1, shortcuts are constructed in a “U-shaped” region. Conditional on an event $\hat{E}_k'$ which is a superset of $\hat{E}_k$ for the “U-shaped” region at scale 2^k, the starting estimate of a piece of shortcut is

$$\mathbb{E}[\#\mathfrak{s}_k \mid E_k'] \leq C_1 2^{2k} \pi_3(2^k).$$

The factor 2^{2k} comes from the five-arm points in the construction. Suppose, at stage i, one can construct a shortcut of the order at most

$$\mathbb{E}[\#\mathfrak{s}_k \mid E'_k] \leq \delta_k(i) 2^{2k} \pi_3(2^k).$$

Through constructions detailed in [DHS21, Section 7], we get an additional gain of $\sim \epsilon^{1/2}$ as long as there is enough space, i.e., when $2^k \geq C(\epsilon)^i$ for some $C(\epsilon) \sim \epsilon^{-4}(\log \frac{1}{\epsilon})^2$. We then iterate this procedure.

Since the proof of this proposition relies mostly on intricate algebraic manipulations, we simply cite the conclusion and refer the reader to the original paper for more explanation.

Step 3: Compilation

The final estimate accounts for edges too close to the origin or the boundary, edges on ℓ_n that don't have shortcuts in Step 1, and a maximal collection of disjoint shortcuts that are optimized in Step 2. For the reader's benefit, we recreate the compilation here.

We first define a truncated box $\hat{B}(n) = B(n - n^\delta) \setminus B(n^\delta)$ for $\delta > 0$ small enough such that $n^{1+2\delta} \leq n^2 \pi_3(n)$. For each $e \in \hat{B}(n)$, we let $L' = \lceil \frac{\delta}{8} \log n \rceil$ and $L = \lfloor \frac{\delta}{4} \log n \rfloor$.

We apply Proposition 3.2.1 for $L' = \lceil \frac{\delta}{8} \log n \rceil$ and $L = \lfloor \frac{\delta}{4} \log n \rfloor$ and obtain

$$\begin{aligned} &\phi^{\xi}_{B(n)}(\text{there is no } n^{-c}\text{-shortcut around } e \mid e \in \ell_n) \\ &\leq \phi^{\xi}_{B(n)}\Big(\bigcap_{j=\lceil \frac{\delta}{8} \log n \rceil}^{\lfloor \frac{\delta}{4} \log n \rfloor} \hat{E}_j(e, \epsilon, n^{-c})^c \mid A_3(e, n^{\delta/2}) \Big) \\ &\leq 2^{-\hat{c}\epsilon^4 \frac{\delta}{8} \log n} \\ &\leq n^{-\eta} \end{aligned}$$

for some $\eta > 0$.

We choose a collection of n^{-c}-shortcuts around edges of ℓ_n such that the shortcuts are disjoint and the number of edges circumvented is maximal. Conditional on the existence of a horizontal crossing, any edge e in $B(n)$ falls into one of three categories: in the margin of the box, with no n^{-c}-shortcut, or with a n^{-c}-shortcut. Thus, H_n has the following estimate:

$$\begin{aligned} \mathbb{E}_{\mathbb{Z}^2}[H_n \mid \mathcal{H}_n] &\leq Cn^{1+\delta} + n^{-\eta}\mathbb{E}[\#\ell_n \mid \mathcal{H}_n] + n^{-c}\mathbb{E}[\#\ell_n \mid \mathcal{H}_n] \\ &\leq Cn^{-\min\{\delta,\eta,c\}} n^2 \pi_3(n). \end{aligned}$$

3.3 Large deviation bound conditional on three arms

In this section, we prove Theorem 3.2.2. We first introduce some notations and definitions. Fix some integer $N > 0$ and for any integer $k \geq 1$, we define $\mathfrak{C}_k$ to be the event that there is a dual-closed circuit in $\mathrm{Ann}(2^{kN}, 2^{(k+1)N})$ with two defect dual-open edges. Similarly, let $\mathfrak{D}_k$ be the event that there is an open circuit in $\mathrm{Ann}(2^{kN}, 2^{(k+1)N})$ with one defect closed edge.

Definition 3.1. For any $k \geq 1$, we define $\hat{\mathfrak{C}}_k$, the compound circuit event in $\mathrm{Ann}(2^{10kN}, 2^{10(k+1)N})$, as the simultaneous occurrence of the following events:

1. for $i = 1, 4, 6, 9$, $\mathfrak{C}_{10k+i}$ occurs in $\mathrm{Ann}(2^{(10k+i)N}, 2^{(10k+i+1)N})$ and
2. $\mathfrak{D}_{10k}$ occurs in $\mathrm{Ann}(2^{10kN}, 2^{(10k+1)N})$.

For simplicity of notation, we define $I_{L',L}$ and $J_{L',L}$ to be (random) collections of indices of scales:

$$I_{L',L} := \left\{ k = \lceil \frac{L'}{10N} \rceil, \dots, \lfloor \frac{L}{10N} \rfloor - 1 : E_{10k+5} \cap \hat{\mathfrak{C}}_k \text{ occurs} \right\}.$$
$$J_{L',L} := \left\{ k = \lceil \frac{L'}{10N} \rceil, \dots, \lfloor \frac{L}{10N} \rfloor - 1 : \hat{\mathfrak{C}}_k \text{ occurs} \right\}.$$

The estimate in Theorem 3.2.2 relies on a two-step strategy: first estimating $\# I_{L',L}$ using the standard Chernoff bound, then expanding the resulting expectation by conditioning on nested filtrations. Since these two steps themselves apply to general random variables, we provide a high-level summary of the proof and only recreate the parts that are sensitive to the model.

To start, we want to condition on there existing sufficiently many decoupling circuits $\hat{\mathfrak{C}}_k$. We quantify this probability using the next Proposition which is proved later in this section.

Proposition 3.3.1 ([DHS21], Proposition 4.2). *There exist $c_1 > 0$ and $N_0 \geq 1$ such that for all $N \geq N_0$ and $L, L' \geq 0$ with $L - L' \geq 40$,*

$$\phi^{\xi}_{B(n)}\left(\# J_{L',L} \leq c_1 \frac{L - L'}{N} \,\middle|\, A_3(2^L) \right) \leq \exp(-c_1(L - L')).$$

Combining Proposition 3.3.1 and the Chernoff bound, we have

$$\begin{aligned} &\phi^{\xi}_{B(n)}\left(\# I_{L',L} \leq c \frac{L - L'}{N} \,\middle|\, A_3(2^L) \right) \\ &\qquad \leq \exp(-c_1(L - L')) + \exp\left(c \frac{L - L'}{N} \right) \mathbb{E}\left[e^{-\# I_{L',L}} \mathbf{1}\{\# J_{L',L} \geq c_1 \frac{L - L'}{N}\} \,\middle|\, A_3(2^L) \right]. \end{aligned}$$

We decompose the expectation over all possible sets of $J_{L',L}$.

$$\sum_{\mathcal{J}:\#\mathcal{J}\geq c_1\frac{L-L'}{N}} \mathbb{E}\left[e^{-\#I_{L',L}} \mid J_{L',L}=\mathcal{J}, A_3(2^L)\right] \phi^{\xi}_{B(n)}(J_{L',L}=\mathcal{J} \mid A_3(2^L)). \tag{3.3}$$

We enumerate $\mathcal{J}=\{k_1,\dots,k_R\}$. Then, conditional on $J_{L',L}=\mathcal{J}$, we have $\#I_{L',L} = \sum_{r=1}^R \mathbf{1}\{E_{10k_r+5}\}$. Define the filtration $(\mathcal{F}_r)$ by

$$\mathcal{F}_r = \sigma\{E_{10k_1+5},\dots,E_{10k_{r-1}+5}\} \cap \{J_{L',L}=\mathcal{J}\} \cap A_3(2^L) \quad \text{for } r=1,\dots,R.$$

Thus, the expectation in (3.3) can be expanded as

$$\mathbb{E}[e^{-\mathbf{1}\{E_{10k_1+5}\}}\cdots\mathbb{E}[e^{-\mathbf{1}\{E_{10k_{R-1}+5}\}}\mathbb{E}[e^{-\mathbf{1}\{E_{10k_R+5}\}} \mid \mathcal{F}_R] \mid \mathcal{F}_{R-1}]\cdots \mid \mathcal{F}_1]$$

where for each $r=1,\dots,R$, we have

$$\mathbb{E}[e^{-\mathbf{1}\{E_{10k_r+5}\}} \mid \mathcal{F}_r] = 1-(1-e^{-1})\phi^{\xi}_{B(n)}(E_{10k_r+5} \mid \mathcal{F}_r). \tag{3.4}$$

Thus, we introduce the following lemma to give a uniform bound on the conditional probability above and decouple E_{10k_r+5} from $\sigma\{E_{10k_1+5},\dots,E_{10k_{r-1}+5}\}$ while conditional on $A_3(2^L)$.

Lemma 3.3.2. *There exists a universal constant $c>0$ such that the following holds. For any $k,L\geq 0$ and $N\geq 1$ satisfying $k\leq \left\lfloor \frac{L}{10N}\right\rfloor - 1$ and any events F and G depending on the status of edges in $B(2^{10kN})$ and $B(2^{10(k+1)N})^c$ respectively, one has*

$$\phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, A_3(2^L), F, G) \geq c\phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, A_3(2^L)).$$

Let $k=\lceil \frac{L'}{10N}\rceil,\dots,\lfloor\frac{L}{10N}\rfloor-1$. For any F depending on edges in $B(2^{10kN})$ and any $\mathcal{J}$ containing k, we have as a result of Lemma 3.3.2

$$\phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, F, J_{L',L}=\mathcal{J}, A_3(2^L)) \geq c\phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, A_3(2^L)) \geq c',$$

with the second inequality owing to (3.1) and gluing constructions (see Section 1.5.1) for mitigating $\hat{\mathfrak{C}}_k$. Inserting back into (3.4), we have

$$\mathbb{E}[e^{-\mathbf{1}\{E_{10k_r+5}\}} \mid \mathcal{F}_r] \leq 1 - c'(1 - e^{-1})$$

Putting everything together, we have

$$\begin{aligned}
&\phi^{\xi}_{B(n)}\left(\#I_{n',n} \leq c\frac{L - L'}{N} \mid A_3(2^L)\right) \\
&\qquad \leq \exp(-c_1(L - L')) + \exp\left(c\frac{L - L'}{N}\right)(1 - c'(1 - e^{-1}))^{c_1 \frac{L-L'}{N}} \\
&\qquad \leq \exp\left(-c''\frac{L - L'}{N}\right),
\end{aligned}$$

which proves Theorem 3.2.2.

Proof of Proposition 3.3.1. We first note the set relation

$$\begin{aligned}
\left\{\#J_{n',n} \leq c_1\frac{L - L'}{N}, A_3(2^L)\right\} \subset\ & A_3\left(2^{10N\lceil \frac{L'}{10N} \rceil}\right) \\
& \cap \left\{\bigcap_{m=10\lceil \frac{L'}{10N} \rceil}^{10\lfloor \frac{L}{10N} \rfloor - 1} A_3(2^{mN}, 2^{(m+1)N}), \#J_{L',L} \leq c_1\frac{L - L'}{N}\right\} \\
& \cap A_3\left(2^{10N\lfloor \frac{L}{10N} \rfloor}, 2^L\right).
\end{aligned}$$

Since the three events on the right-hand side are disjoint, we apply (DMP) twice and obtain

$$\begin{aligned}
&\phi^{\xi}_{B(n)}\left(\#J_{n',n} \leq c_1\frac{L - L'}{N}, A_3(2^L)\right) \\
&\qquad \leq \phi^{\xi''}_{B(2^{10N\lceil \frac{L'}{10N} \rceil})}\left(A_3\left(2^{10N\lceil \frac{L'}{10N} \rceil}\right)\right) \\
&\qquad \quad \times \phi^{\xi'}_{B(2^{10N\lfloor \frac{L}{10N} \rfloor})}\left(\bigcap_{m=10\lceil \frac{L'}{10N} \rceil}^{10\lfloor \frac{L}{10N} \rfloor - 1} A_3(2^{mN}, 2^{(m+1)N}), \#J_{L',L} \leq c_1\frac{L - L'}{N}\right) \qquad (3.5) \\
&\qquad \quad \times \phi^{\xi}_{B(n)}\left(A_3(2^{10N\lfloor \frac{L}{10N} \rfloor}, 2^L)\right).
\end{aligned}$$

Since three-arm probabilities are of the same order uniform over boundary conditions, it suffices to show that (3.5) can be bounded by

$$O\left(\exp(-c''(L-L'))\phi^{\xi'}_{B(2^{10N\lfloor \frac{L}{10N}\rfloor})}\left(A_3(2^{10N\lceil \frac{L'}{10N}\rceil}, 2^{10N\lfloor \frac{L}{10N}\rfloor})\right)\right). \tag{3.6}$$

For each scale m, Let X_m be the indicator function on the event $A_3(2^{10mN}, 2^{10(m+1)N})$ occurs but $\hat{\mathfrak{C}}_m$ does not. Then,

$$\left\{\bigcap_{m=10\lceil \frac{L'}{10N}\rceil}^{10\lfloor \frac{L}{10N}\rfloor-1} A_3(2^{mN}, 2^{(m+1)N}), \#J_{L',L} \le c_1\frac{L-L'}{N}\right\} \subset \left\{\sum_{m=\lceil \frac{L'}{10N}\rceil}^{\lfloor \frac{L}{10N}\rfloor-1} X_m \ge \lfloor \frac{L}{10N}\rfloor - \lceil \frac{L'}{10N}\rceil - c_1\frac{L-L'}{N}\right\}.$$

Although the X_m's are not independent, when applying (DMP), the dependence between events is only reflected on the boundary condition. We will establish an upper bound uniform over boundary conditions for each X_m, thus allowing us access to a set of independent Y_m's that stochastically dominate the X_m's.

For now, the Y_m's are independent Bernoulli random variables with parameters $p_m \in (2^{-\beta N}, 1)$ to be chosen later. We use an elementary lemma from [DHS21] on the concentration of independent Bernoulli random variables.

Lemma 3.3.3 ([DHS21], Lemma 4.3)**.** *Given $\epsilon_1 \in (0,1)$ and $M \ge 1$, if $Y_1, \dots, Y_M$ are any independent Bernoulli random variables with parameters $p_1, \dots, p_M$, respectively, satisfying $p_i \in [\epsilon_1, 1]$ for all i, then for all $r \in (0,1)$,*

$$\mathbf{P}\left(\sum_{m=1}^{M} Y_m \ge rM\right) \le (1/\epsilon_1)^{M(1-r)} 2^M \prod_{m=1}^{M} p_m.$$

Applying Lemma 3.3.3 by taking $M = \lfloor \frac{L}{10N} \rfloor - \lceil \frac{L'}{10N} \rceil$ and $r = 1 - 20c_1$, we have

$$\begin{aligned} &\mathbf{P}\left(\sum_{m=\lceil \frac{L'}{10N}\rceil}^{\lfloor \frac{L}{10N}\rfloor - 1} X_m \geq \lfloor \frac{L}{10N} \rfloor - \lceil \frac{L'}{10N} \rceil - c_1 \frac{L-L'}{N}\right) \\ &\qquad \leq \mathbf{P}\left(\sum_{m=\lceil \frac{L'}{10N}\rceil}^{\lfloor \frac{L}{10N}\rfloor - 1} Y_m \geq \lfloor \frac{L}{10N} \rfloor - \lceil \frac{L'}{10N} \rceil - c_1 \frac{L-L'}{N}\right) \\ &\qquad \leq \left(2^{20c_1\beta N+1}\right)^{\lfloor \frac{L}{10N} \rfloor - \lceil \frac{L'}{10N} \rceil} \prod_{m=\lceil \frac{L'}{10N}\rceil}^{\lfloor \frac{L}{10N}\rfloor - 1} p_m. \end{aligned} \tag{3.7}$$

We now construct the Y_m's by computing $\mathbf{P}(X_m = 1) = \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \hat{\mathfrak{C}}^c_m\right)$. By a union bound,

$$\begin{aligned} &\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \hat{\mathfrak{C}}^c_m\right) \\ &\qquad = \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{D}^c_{10m} \cup \bigcup_{i=1,4,6,9} \mathfrak{C}^c_{10m+i}\right) \\ &\qquad \leq \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{D}^c_{10m}\right) \\ &\qquad\quad + \sum_{i=1,4,6,9} \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{C}^c_{10m+i}\right). \end{aligned} \tag{3.8}$$

Using (DMP) twice, for each i and similarly the probability with $\mathfrak{D}^c_{10m}$,

$$\begin{aligned} \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{C}^c_{10m+i}\right) &\leq \phi^{\eta_2}_{B(2^{(10m+i)N})}\left(A_3(2^{10mN}, 2^{(10m+i)N})\right) \\ &\quad \times \phi^{\eta_1}_{B(2^{(10m+i+1)N})}\left(A_3(2^{(10m+i)N}, 2^{(10m+i+1)N}), \mathfrak{C}^c_{10m+i}\right) \\ &\quad \times \phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{(10m+i+1)N}, 2^{10(m+1)N})\right), \end{aligned} \tag{3.9}$$

where η_1, η_2 are randomly induced boundary conditions.

Claim 3.1. *There exists $\alpha \in (0,1)$ such that for any boundary condition η, we have*

$$\phi^{\eta}_{B(2^{(10m+i+1)N})}\left(A_3(2^{(10m+i)N}, 2^{(10m+i+1)N}), \mathfrak{C}^c_{10m+i}\right) \leq c2^{-\alpha N} \phi^{\eta}_{B(2^{(10m+i+1)N})}\left(A_3(2^{(10m+i)N}, 2^{(10m+i+1)N})\right),$$

$$\phi^{\eta}_{B(2^{(10m+1)N})}\left(A_3(2^{10mN}, 2^{(10m+1)N}), \mathfrak{D}^c_{10m}\right) \leq c2^{-\alpha N} \phi^{\eta}_{B(2^{(10m+1)N})}\left(A_3(2^{10mN}, 2^{(10m+1)N})\right).$$

Plugging this back into (3.9) and using gluing constructions similar to that demonstrated in Section 1.5.1, we have

$$\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{C}^c_{10m+i}\right) \leq c2^{-\alpha N}\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N})\right). \tag{3.10}$$

Similarly,

$$\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \mathfrak{D}^c_{10m}\right) \leq c2^{-\alpha N}\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N})\right). \tag{3.11}$$

Plugging (3.10) and (3.11) into (3.8) and using (QM), we have

$$\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N}), \hat{\mathfrak{C}}^c_m\right) \leq 5c2^{-\alpha N}\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N})\right).$$

By choosing $p_m = 5c2^{-\alpha N}\phi^{\xi'}_{B(2^{10(m+1)N})}\left(A_3(2^{10mN}, 2^{10(m+1)N})\right)$, we have

$$(3.7) \leq \left(2^{20c_1\beta N+1}\right)^{\lfloor \frac{L}{10N}\rfloor - \lceil \frac{L'}{10N}\rceil} C2^{-\alpha N}\phi^{\xi'}_{B(2^{10N\lfloor \frac{L}{10N}\rfloor})}\left(A_3(2^{10N\lceil \frac{L'}{10N}\rceil}, 2^{10N\lfloor \frac{L}{10N}\rfloor})\right),$$

which demonstrates (3.6). □

Proof of Claim 3.1. By duality/Menger's theorem, $A_3(2^{kN}, 2^{(k+1)N}) \cap \mathfrak{C}^c_k$ (resp. $A_3(2^{kN}, 2^{(k+1)N}) \cap \mathfrak{D}^c_k$) is equivalent to the disjoint occurrence of $A_3(2^{kN}, 2^{(k+1)N})$ and $A_{1,O}(2^{kN}, 2^{(k+1)N})$ (resp. $A_{1,C^*}(2^{kN}, 2^{(k+1)N})$). Let $\mathcal{U}$ denote an explorable (random) region that contains two open and one dual-closed arm. Then, conditional on $\mathcal{U}$, the complement region is independent up to some boundary condition. This type of independence argument is standard in Bernoulli percolation and a detailed treatment can be found in Section 4.3, for example.

$$\begin{aligned}
&\phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N}), \mathfrak{C}^c_k\right) \\
&\qquad = \phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N}) \square A_{1,O}(2^{kN}, 2^{(k+1)N})\right) \\
&\qquad \leq \sum_{\text{admissible } U} \phi^{\iota}_{U^c}\left(A_{1,O}(2^{kN}, 2^{(k+1)N})\right)\phi^{\eta}_{B(2^{(k+1)N})}(\mathcal{U} = U).
\end{aligned}$$

The one-arm probability decays at $(2^N)^{-\alpha}$ for some $\alpha \in (0,1)$ due to quad-crossing (RSW) estimates. Thus,

$$\phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N}), \mathfrak{C}_k^c\right) \leq 2^{-\alpha N}\phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N})\right).$$

By the same reasoning,

$$\phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N}), \mathfrak{D}_k^c\right) \leq 2^{-\alpha N}\phi^{\eta}_{B(2^{(k+1)N})}\left(A_3(2^{kN}, 2^{(k+1)N})\right).$$

□

Now we prove Lemma 3.3.2.

Proof of Lemma 3.3.2. Recall that F and G depend on the status of edges in $B(2^{10kN})$ and $B(2^{10(k+1)N})^c$ respectively. We focus on demonstrating how we remove the conditioning on F as the other side is similar.

$$\phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, A_3(2^L), F, G) \geq c_2 \phi^{\xi}_{B(n)}(E_{10k+5} \mid \hat{\mathfrak{C}}_k, A_3(2^L), G). \tag{3.12}$$

Recall from Definition 3.1 that $\hat{\mathfrak{C}}_k$ is the simultaneous occurrence of a stack of circuits. We use these circuits to separate F and E_{10k+5} while conditioning on $A_3(2^L)$.

A dual-closed circuit C with two defect edges is naturally divided into two arcs between these defects consisting of open connections. Fix some deterministic ordering of arcs and label the two arcs $\mathrm{Arc}_1(C)$ and $\mathrm{Arc}_2(C)$ in this ordering. Let $X_-(C,i)$ be the event such that:

1. $\mathfrak{D}_{10k}$ occurs;
2. C is the *innermost* dual-closed circuit with two defect edges in $\mathrm{Ann}(2^{(10k+1)N}, 2^{(10k+2)N})$;

3. the origin is connected to the two defects in C through two disjoint open paths;
4. $(\frac{1}{2}, -\frac{1}{2})$ is connected to $\mathrm{Arc}_i(C)$ through a dual-closed path.

Note that the occurrence of $\mathfrak{D}_{10k}$, together with the two open paths through the origin, guarantees that item (4) only occurs for one of the two arcs. Hence $X_-(C, i)$ occurs for exactly one choice of C and i. Similarly, let $X_+(\mathcal{D}, j)$ be the event such that:

1. $\mathfrak{C}_{10k+i}$ occurs for $i = 6, 9$;
2. $\mathcal{D}$ is the *innermost* dual-closed circuit with two defect edges in $\mathrm{Ann}(2^{(10k+4)N}, 2^{(10k+5)N})$;
3. the two defects are connected to $\partial B(2^L)$ through disjoint open paths;
4. $\mathrm{Arc}_j(\mathcal{D})$ is connected to $\partial B(2^L)$ through a dual-closed path.

We also need a three-arm event between C and $\mathcal{D}$. The dual-closed arm connects an arc of C and an arc of $\mathcal{D}$ and the indices of the arcs are important. Thus, let $X(C, \mathcal{D}, i, j)$ be the event such that:

1. there is a dual-closed path connecting $\mathrm{Arc}_i(C)$ and $\mathrm{Arc}_j(\mathcal{D})$ in the region between C and $\mathcal{D}$;
2. there is a pair of disjoint open paths in the region between C and $\mathcal{D}$ connecting a defect of C and a defect of $\mathcal{D}$.

Note that for each pair of i and j, topologically there is only one possible way to connect the defects of C and $\mathcal{D}$ with open paths.

Then, the occurrence of $\hat{\mathfrak{C}}_k$ allows for a decomposition in admissible $\mathcal{C}$, $\mathcal{D}$, and $i, j = 1, 2$ for the following event.

$$\begin{aligned}&\phi^{\xi}_{B(n)}(E_{10k+5}, \hat{\mathfrak{C}}_k, A_3(2^L), F, G)\\ &\qquad = \sum_{\mathcal{C},\mathcal{D},i,j} \phi^{\xi}_{B(n)}(F, X_-(\mathcal{C}, i), X(\mathcal{C}, \mathcal{D}, i, j), X_+(\mathcal{D}, j), E_{10k+5}, G).\end{aligned}$$

Since $\mathcal{C}$ is the innermost circuit, its position can be determined via exploration in the interior of $\mathcal{C}$. Using (DMP) first on $\mathrm{Ext}(\mathcal{C}) := B(n) \setminus \mathrm{Int}(\mathcal{C})$, the boundary condition induced by the configuration in $\mathrm{Int}(\mathcal{C})$ identifies only the two defect edges in $\mathcal{C}$, which we denote by $0^*(\mathcal{C})$. Thus, we have

$$\begin{aligned}&\phi^{\xi}_{B(n)}(F, X_-(\mathcal{C}, i), X(\mathcal{C}, \mathcal{D}, i, j), X_+(\mathcal{D}, j), E_{10k+5}, G)\\ &\qquad = \phi^{\xi}_{B(n)}(F, X_-(\mathcal{C}, i))\phi^{0^*(\mathcal{C})}_{\mathrm{Ext}(\mathcal{C})}(X(\mathcal{C}, \mathcal{D}, i, j), X_+(\mathcal{D}, j), E_{10k+5}, G).\end{aligned}$$

Using (DMP) again on $\mathrm{Ext}(\mathcal{D})$, the configuration in $\mathrm{Int}(\mathcal{D}) \setminus \mathrm{Ext}(\mathcal{C})$ induces a free boundary condition on $\mathcal{D}$. Thus, we have

$$\begin{aligned}&\phi^{\xi}_{B(n)}(E_{10k+5}, \hat{\mathfrak{C}}_k, A_3(2^L), F, G)\\ &\qquad = \sum_{\mathcal{C},\mathcal{D},i,j} \phi^{\xi}_{B(n)}(F, X_-(\mathcal{C}, i))\phi^{0^*(\mathcal{C})}_{\mathrm{Ext}(\mathcal{C})}(X(\mathcal{C}, \mathcal{D}, i, j))\phi^{0}_{\mathrm{Ext}(\mathcal{D})}(X_+(\mathcal{D}, j), E_{10k+5}, G)\end{aligned} \tag{3.13}$$

A similar decomposition gives,

$$\begin{aligned}&\phi^{\xi}_{B(n)}(\hat{\mathfrak{C}}_k, A_3(2^L), G)\\ &\qquad = \sum_{\mathcal{C}',\mathcal{D}',i',j'} \phi^{\xi}_{B(n)}(X_-(\mathcal{C}', i'))\phi^{0^*(\mathcal{C}')}_{\mathrm{Ext}(\mathcal{C}')}(X(\mathcal{C}', \mathcal{D}', i', j'))\phi^{0}_{\mathrm{Ext}(\mathcal{D}')}(X_+(\mathcal{D}', j'), G)\end{aligned} \tag{3.14}$$

Multiplying (3.13) and (3.14) gives

$$\begin{aligned}&\phi^{\xi}_{B(n)}(E_{10k+5}, \hat{\mathfrak{C}}_k, A_3(2^L), F, G)\phi^{\xi}_{B(n)}(\hat{\mathfrak{C}}_k, A_3(2^L), G)\\ &\quad = \sum_{\substack{\mathcal{C},\mathcal{D},i,j\\ \mathcal{C}',\mathcal{D}',i',j'}} \Big[\phi^{\xi}_{B(n)}(F, X_-(\mathcal{C}, i))\phi^{0^*(\mathcal{C})}_{\mathrm{Ext}(\mathcal{C})}(X(\mathcal{C}, \mathcal{D}, i, j))\phi^{0}_{\mathrm{Ext}(\mathcal{D})}(X_+(\mathcal{D}, j), E_{10k+5}, G)\\ &\qquad \times \phi^{\xi}_{B(n)}(X_-(\mathcal{C}', i'))\phi^{0^*(\mathcal{C}')}_{\mathrm{Ext}(\mathcal{C}')}(X(\mathcal{C}', \mathcal{D}', i', j'))\phi^{0}_{\mathrm{Ext}(\mathcal{D}')}(X_+(\mathcal{D}', j'), G)\Big].\end{aligned} \tag{3.15}$$

We use the following estimate which is the random cluster analogue of [DS11, Lemma 6.1]. It is essentially a corollary of the so-called strong separation lemmas, the random cluster version of which we prove in Section 3.4. To get from the strong separation lemmas to the following estimate, a proof sketch of no essential difference can be found in [DS11].

There exists a uniform constant c such that the following holds for all choices of circuits $C, C', \mathcal{D}, \mathcal{D}'$ and arc indices i, i', j, j':

$$\frac{\phi^{0^*(C)}_{\mathrm{Ext}(C)}(X(C,\mathcal{D}',i,j'))\phi^{0^*(C')}_{\mathrm{Ext}(C')}(X(C',\mathcal{D},i',j))}{\phi^{0^*(C)}_{\mathrm{Ext}(C)}(X(C,\mathcal{D},i,j))\phi^{0^*(C')}_{\mathrm{Ext}(C')}(X(C',\mathcal{D}',i',j'))} < c. \tag{3.16}$$

Applying (3.16) to the summand of (3.15), we have

$$\begin{aligned}
&\phi^{\xi}_{B(n)}(F, X_-(C,i))\phi^{0^*(C)}_{\mathrm{Ext}(C)}(X(C,\mathcal{D},i,j))\phi^{0}_{\mathrm{Ext}(\mathcal{D})}(X_+(\mathcal{D},j), E_{10k+5}, G)\\
&\qquad\qquad \times \phi^{\xi}_{B(n)}(X_-(C',i'))\phi^{0^*(C')}_{\mathrm{Ext}(C')}(X(C',\mathcal{D}',i',j'))\phi^{0}_{\mathrm{Ext}(\mathcal{D}')}(X_+(\mathcal{D}',j'), G)\\
&\quad > c^{-1}\phi^{\xi}_{B(n)}(F, X_-(C,i))\phi^{0^*(C)}_{\mathrm{Ext}(C)}(X(C,\mathcal{D}',i,j'))\phi^{0}_{\mathrm{Ext}(\mathcal{D}')}(X_+(\mathcal{D}',j'), G)\\
&\qquad\qquad \times \phi^{\xi}_{B(n)}(X_-(C',i'))\phi^{0^*(C')}_{\mathrm{Ext}(C')}(X(C',\mathcal{D},i',j))\phi^{0}_{\mathrm{Ext}(\mathcal{D})}(X_+(\mathcal{D},j), E_{10k+5}, G).
\end{aligned}$$

Summing over $C, \mathcal{D}, i, j, C', \mathcal{D}', i', j'$, by (DMP),

$$\begin{aligned}
(3.15) > &c^{-1} \sum_{\substack{C,\mathcal{D},i,j\\ C',\mathcal{D}',i',j'}} \Big[\phi^{\xi}_{B(n)}(F, X_-(C,i))\phi^{0}_{\mathrm{Ext}(\mathcal{D}')}(X_+(\mathcal{D}',j'), G)\phi^{0^*(C)}_{\mathrm{Ext}(C)}(X(C,\mathcal{D}',i,j'))\\
&\qquad\qquad \times \phi^{\xi}_{B(n)}(X_-(C',i'))\phi^{0}_{\mathrm{Ext}(\mathcal{D})}(X_+(\mathcal{D},j), E_{10k+5}, G)\phi^{0^*(C')}_{\mathrm{Ext}(C')}(X(C',\mathcal{D},i',j))\Big]\\
=&c^{-1}\phi^{\xi}_{B(n)}(\hat{\mathfrak{C}}_k, A_3(2^L), F, G)\phi^{\xi}_{B(n)}(E_{10k+5}, \hat{\mathfrak{C}}_k, A_3(2^L), G).
\end{aligned}$$

Dividing both sides by $\phi^{\xi}_{B(n)}(\hat{\mathfrak{C}}_k, A_3(2^L), G)\phi^{\xi}_{B(n)}(\hat{\mathfrak{C}}_k, A_3(2^L), F, G)$, we have (3.12) with $c_2 = c^{-1}$.

From (3.12), we can remove the conditioning on G using a nearly identical argument. □

Although the above proof is formally similar to the proof of Lemma 4.4 in [DHS21], it heavily relies on the domain Markov property (DMP), so the choice of the domain and the order of application are crucial.

3.4 Arm separation for the random cluster model

As indicated in the proof of Lemma 3.3.2, (3.16) depends on the following two strong arm separation lemmas in combination with the gluing constructions explained in Section 1.5.1.

Lemma 3.4.1 (External Arm Separation)**.** *Fix an integer $m \geq 2$ and let $n_1 \leq n_2 - 3$. Consider an open circuit C in $B(2^{n_1})$ with m defects $e_1, \ldots, e_m$. Let $\mathcal{A}(C, 2^{n_2})$ be the event that*

1. *there are $2m$ alternating disjoint open arms and dual-closed arms from C to $\partial B(2^{n_2})$ in $B(2^{n_2}) \setminus \mathrm{Int}(C)$;*
2. *the m dual-closed paths emanate from e_j^* to $\partial B(2^{n_2})^*$, respectively.*

We note that the dependence on the defects $e_1, \ldots, e_m$ is implicit in the notation of C. Let $\tilde{\mathcal{A}}(C, 2^{n_2})$ be the event that $\mathcal{A}(C, 2^{n_2})$ occurs with $2m$ arms $\gamma_1, \ldots, \gamma_{2m}$ (open, dual-closed alternatingly) whose endpoints in $\partial B(2^{n_2})$ or $\partial B(2^{n_2})^*$, $f_1, \ldots, f_{2m}$, satisfy

$$2^{-n_2} \min_{k \neq l} |f_k - f_l| \geq \frac{1}{2m}.$$

Then, there is a constant $c(m) > 0$ independent of n_1, n_2, C, and the boundary condition ξ such that

$$\phi^{\xi}_{B(2^{n_2})}(\mathcal{A}(C, 2^{n_2})) \leq c(m)\phi^{\xi'}_{B(2^{n_2})}(\tilde{\mathcal{A}}(C, 2^{n_2}))$$

for some boundary condition ξ' on $B(2^{n_2})$.

Lemma 3.4.2 (Internal Arm Separation)**.** *Fix an integer $m \geq 2$ and let $n_3 + 3 \leq n_4$. Consider an open circuit $\mathcal{D}$ in $B(2^{n_4})^c$ with m defects $g_1, \ldots, g_m$. Let $\mathcal{B}(2^{n_3}, \mathcal{D})$ be the event that*

1. *there are $2m$ alternating disjoint open arms and dual-closed arms from $\partial B(2^{n_3})$ to $\mathcal{D}$ in $\mathrm{Int}(\mathcal{D}) \setminus B(2^{n_3})$;*
2. *the m dual-closed paths emanate from g_j^* to $\partial B(2^{n_3})^*$, respectively.*

Let $\tilde{\mathcal{B}}(2^{n_3}, \mathcal{D})$ be the event that $\mathcal{B}(2^{n_3}, \mathcal{D})$ occurs with $2m$ arms $\gamma_1, \ldots, \gamma_{2m}$ (open, dual-closed alternatingly) whose endpoints in $\partial B(2^{n_3})$ or $\partial B(2^{n_3})^*$, $h_1, \ldots, h_{2m}$, satisfy

$$2^{-n_3} \min_{k \neq l} |h_k - h_l| \geq \frac{1}{2m}.$$

Then, there is a constant $c'(m) > 0$ independent of $n_3, n_4, \mathcal{D}$, and the boundary condition ξ such that

$$\phi^{\xi}_{\mathrm{Int}(\mathcal{D})}(\mathcal{B}(2^{n_3}, \mathcal{D})) \leq c'(m) \phi^{\xi}_{\mathrm{Int}(\mathcal{D})}(\tilde{\mathcal{B}}(2^{n_3}, \mathcal{D}))$$

for some boundary condition ξ' on $\mathcal{D}$.

Remark 3.1. *ξ' arises due to a technical challenge in the proof. However, for the purpose of* (3.16)*, any boundary condition suffices as the* (RSW) *estimates we have are uniform in boundary conditions.*

The proofs of Lemma 3.4.1 and 3.4.2 are similar, and we only provide the proof for the former.

Arm separation techniques are classical techniques that date back to Kesten [Kes87, Nol08]. They were first developed to show well-separatedness for arms crossing square annuli. In our case, the annulus consists of one square boundary and one circuitous boundary. The main obstacle for directly applying the classical arm separation arguments is that the geometry of the circuit may generate bottlenecks that prevent arms from being separated on certain scales. In the first part of the proof, we address this through a construction that "leads" the interfaces to the boundary of $B(2^{n_1})$. We note that this part of the proof for the random cluster model is identical to that of [DS11, Lemma 6.2] as the constructions are purely topological. However, we include here for the reader's convenience. The second step is to define a family of disjoint annuli in levels, which groups the arms based on their relative distances. In the following proof, the details for this step is provided last. The final part of the proof depends on an arm separation statement in each annuli defined in the previous step, for which we provide the details in Lemma 3.4.3.

We note that our proof is stated in full generality compared to the proof in [DS11] which is stated for $m = 2$, and therefore slightly deviates from it in notation.

Proof of Lemma 3.4.1. Given the circuit C with defects $e_1, \ldots, e_m$, we assume the occurrence of $\mathcal{A}(C, 2^{n_2})$. The first step is to "extend" the circuit C to $\partial B(2^{n_1})$ so that the arms will not be tangled due to the geometry of C.

For $i = 1, \ldots, m$, α_i^l be the counterclockwise-most dual-closed path emanating from e_i^* to $\partial B(2^{n_1} + 1/2)$ in $B(2^{n_1} + 1/2) \setminus C$ and α_i^r the clockwise-most dual-closed path emanating from e_i^* to $\partial B(2^{n_1} + 1/2)$ in $B(2^{n_1} + 1/2) \setminus C$. We denote by a_i^l the first vertex on $\partial B(2^{n_1})$ to the counterclockwise side of α_i^l and a_i^r the first vertex on $\partial B(2^{n_1})$ to the clockwise side of α_i^r. Let β_i^l be the counterclockwise-most open path from the lower right end-vertex of e_i to a_i^r in $B(2^{n_1}) \setminus C$ and β_i^r be the clockwise-most open path from the top left end-vertex of e_{i+1} to a_{i+1}^l in $B(2^{n_1}) \setminus C$. Here, the indices are cyclic, meaning that $i = i \bmod m$.

Note that it is necessary that $a_i^l \neq a_i^r$; it is possible that $a_i^r = a_{i+1}^l$, but by the assumption that $\mathcal{A}(C, n_2)$ occurs, a_{i+1}^l must be on the clockwise side of a_i^r on $\partial B(2^{n_1})$.

We identify the last intersection of α_i^l and α_i^r, which can possibly be e_i. Let α_i be the union of the piece of α_i^l from the last intersection to $\partial B(2^{n_1} + 1/2)$ and the piece of α_i^r from the last intersection to $\partial B(2^{n_1} + 1/2)$. Let R_i be the domain bounded by α_i and the piece of $\partial B(2^{n_1})$ between a_i^l and a_i^r on a_i^l's clockwise side. We now define a path β_i. If β_i^l and β_i^r intersect, we define β_i analogously to α_i. Otherwise, we define β_i to be the union of the piece of β_i^l from its last intersection with C to $\partial B(2^{n_1})$, the piece of β_i^r from its last intersection with C to $\partial B(2^{n_1})$, and the piece of C that connects the aforementioned two pieces. Let S_i be the domain bounded by β_i and the piece of $\partial B(2^{n_1})$ between a_i^r and a_{i+1}^l on a_i^r's clockwise side. Note that in the case $a_i^r = a_{i+1}^l$, β_i and S_i consist of only the vertex a_i^r.

Let $R := (B(2^{n_2}) \setminus B(2^{n_1})) \cup (\cup_{i=1}^m R_i) \cup (\cup_{i=1}^m S_i)$. Note that once $\{\alpha_i, \beta_i\}_i$ is fixed, the conditional distribution of the cluster configuration inside R is (uniquely) determined by the status of α_i and β_i. Let $\mathcal{A}(R)$ denote the event that

1. there is a dual-closed arm connecting α_i to $\partial B(2^{n_2})^*$ in R for $i = 1, \ldots, m$;
2. there is an open arm connecting β_i to $\partial B(2^{n_2})$ in R for $i = 1, \ldots, m$.

Let $\tilde{\mathcal{A}}(R)$ be the event that $\mathcal{A}(R)$ occurs with $2m$ arms $\gamma_1, \ldots, \gamma_{2m}$ (dual-closed, open alternatingly) whose endpoints in $\partial B(2^{n_2})$ or $\partial B(2^{n_2})^*$, $f_1, \ldots, f_{2m}$, satisfy $2^{-n_2} \min_{k \neq l} |f_k - f_l| \geq 1/(2m)$. Lemma 3.4.1 is then equivalent to

$$\phi_{B(2^{n_2})}^{\xi}(\mathcal{A}(R)) \leq c(m) \phi_{B(2^{n_2})}^{\xi'}(\tilde{\mathcal{A}}(R))$$

for some constant $c(m) > 0$ that only depends on m and some boundary condition ξ'.

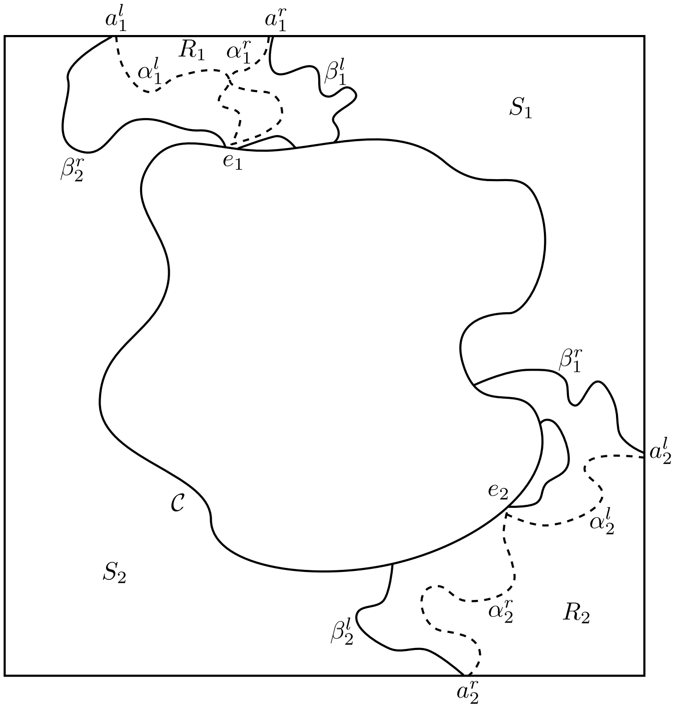


Figure 3.1: A representation of the construction in the first step with $m = 2$. This figure is topologically equivalent to [DS11, Fig. 4], with a relabeling.

Let us first relabel the vertices $a_1^l, a_1^r, \dots, a_m^l, a_m^r$ by $x_1, \dots, x_{2m}$ where $x_{2i-1} = a_i^l$ and $x_{2i} = a_i^r$ for $i = 1, \dots, m$. The next step is to identify *critical scales*, scales of neighborhoods of these vertices comparable to the distance beween them. We now informally introduce the notion of *level-j annuli* so we can finish the proof before returning to formally defining them at the end of the proof.

For $j = 1, \dots, 2m-1$, $\mathcal{I}_j$ is a collection of indices. $\mathcal{I}_j$ keeps track of groups of $j+1$ vertices among $x_1, \dots, x_{2m}$ on level j and the index indicates the first vertex in a group in clockwise order. For each level j, the difference of two indices in $\mathcal{I}_j$ is at least $j+1$.

Let $\mathcal{L}_j$ be the collection of level-j annuli: $\mathcal{L}_j := \{\mathrm{Ann}_j(i) : i \in \mathcal{I}_j\}$. The level-$j$ annuli $\mathrm{Ann}_j(i)$ satisfy:

- If $i \in \mathcal{I}_j$, that is, $\mathrm{Ann}_j(i)$ is nonempty, then $\mathrm{Ann}_j(i)$ is centered on $\partial B(2^{n_1})$ and both its inner box and outer box enclose exactly $j+1$ vertices $x_i, x_{i+1}, \ldots, x_{i+j}$. That is, $\mathrm{Ann}_j(i)$ is crossed by j arms.
- Level-j annuli are mutually disjoint and disjoint from annuli of other levels.
- There is exactly one level-$2m$ annulus and at most $\lfloor \frac{2m}{j+1} \rfloor$ (and possibly zero) level-j annuli.
- All level-j annuli are contained in $B(2^{n_1+1})$, for $j = 1, \ldots, 2m-1$. The level-$2m$ annulus is $\mathrm{Ann}_{2m}(1) = \mathrm{Ann}(2^{n_1+1}, 2^{n_2})$.

$\mathcal{A}(R)$ implies the simultaneous occurrence of crossings in each of the annuli defined above intersected with the domain R, which can cause the annuli to have irregular boundaries. However, since all annuli (excluding the level-$2m$ annulus) are centered on $\partial B(2^{n_1})$ and the boundary of R (the α_i and β_i) are in the interior of $B(2^{n_1})$, each annulus $\mathrm{Ann}_j(i)$ intersected with R necessarily contains one of the top-, bottom-, left-, or right-half of $\mathrm{Ann}_j(i)$. We call the half annulus $\mathrm{Ann}_j^{\mathrm{h}}(i)$. If there are two choices, choose the top or bottom over the left or right.

For $j = 1, \ldots, 2m-1$, let $E_j(i)$ be the event that there exist j disjoint crossings in $\mathrm{Ann}_j^{\mathrm{h}}(i)$ such that the color of each crossing is determined by the vertices the annulus encloses. In particular, let E_{2m} be the event that Ann_{2m} is crossed by $2m$ disjoint alternating open and dual-closed crossings. Then $\mathcal{A}(R)$ implies the occurence of $\cap_{j=1}^{2m} \cap_{i\in\mathcal{I}_j} E_j(i)$. By repeatedly applying (DMP), we have

$$\phi^{\xi}_{B(2^{n_2})}(\mathcal{A}(R)) \leq \phi^{\xi}_{B(2^{n_2})}(\cap_{j=1}^{2m} \cap_{i\in\mathcal{I}_j} E_j(i)) = \prod_{j=1}^{2m} \prod_{i\in\mathcal{I}_j} \phi^{\xi_{j,i}}_{D_{j,i}}(E_j(i)), \tag{3.17}$$

where $D_{j,i} = \cup_{k=1}^{j-1}(\cup_{D\in\mathcal{L}_k} D) \cup (\cup_{k=1}^{i} \mathrm{Ann}_j(k))$, that is, D_j is the union of all annuli up to level $j-1$ union the union of all annuli in on level-j up to index i. The exception is $D_{2m} = B(2^{n_2})$. And $\xi_{j,i}$ is some implicit boundary condition on $\partial D_{j,i}$ induced by conditioning on the outside and $\xi_{2m} = \xi$. Note that $D_{j_1,i_1} \subset D_{j_2,i_2}$ if $j_1 < j_2$ or if $j_1 = j_2$ and $i_1 < i_2$.

Let $\tilde{E}_j(i)$ be the event that $E_j(i)$ occurs and the exit points of the crossings are separated, that is the distance between any two exit points are at least $\delta/2m$ times the length of the boundary of the box they are on, for some $\delta > 1/8$. The following lemma is an arm separation statement that compares the separated event to the regular arm event.

Lemma 3.4.3. *For any i, j, ξ, there is a $C = C(m) > 0$ such that for any $j = 1, \dots, 2m$,*

$$\phi^{\xi}_{D_{j,i}}(E_j(i)) \leq C\phi^{\xi}_{D_{j,i}}(\tilde{E}_j(i)).$$

Proof. This a classical result using (RSW) and (FKG) estimates except on half-annuli. Nonetheless, all parts of the classical argument apply. We refer the reader to the proof of [CDCH16, Proposition 5.6]. □

Applying Lemma 3.4.3 to each probability in the RHS of (3.17) and then (DMP) and we have

$$\begin{aligned}
\prod_{j=1}^{2m}\prod_{i\in\mathcal{I}_j}\phi^{\xi_{j,i}}_{D_{j,i}}(E_j(i)) &\leq c\prod_{j=1}^{2m}\prod_{i\in\mathcal{I}_j}\phi^{\xi_{j,i}}_{D_{j,i}}(\tilde{E}_j(i)) \\
&= c\phi^{\xi_{1,i_1}}_{D_{1,i_1}}(\tilde{E}_1(i_1))\phi^{\xi_{1,i_2}}_{D_{1,i_2}}(\tilde{E}_1(i_2))\prod_{k=3}^{|\mathcal{I}_1|}\phi^{\xi_{1,i_k}}_{D_{1,i_k}}(\tilde{E}_1(i_k))\prod_{j=2}^{2m}\prod_{i\in\mathcal{I}_j}\phi^{\xi_{j,i}}_{D_{j,i}}(\tilde{E}_j(i)) \\
&= c\phi^{\xi'_{1,i_2}}_{D_{1,i_2}}(\tilde{E}_1(i_1)\cap\tilde{E}_1(i_2))\prod_{k=3}^{|\mathcal{I}_1|}\phi^{\xi_{1,i_k}}_{D_{1,i_k}}(\tilde{E}_1(i_k))\prod_{j=2}^{2m}\prod_{i\in\mathcal{I}_j}\phi^{\xi_{j,i}}_{D_{j,i}}(\tilde{E}_j(i)) \\
&\dots \\
&= c\phi^{\xi'}_{B(2^{n_2})}(\cap_{j=1}^{2m}\cap_{i\in\mathcal{I}_j}\tilde{E}_j(i)),
\end{aligned}$$

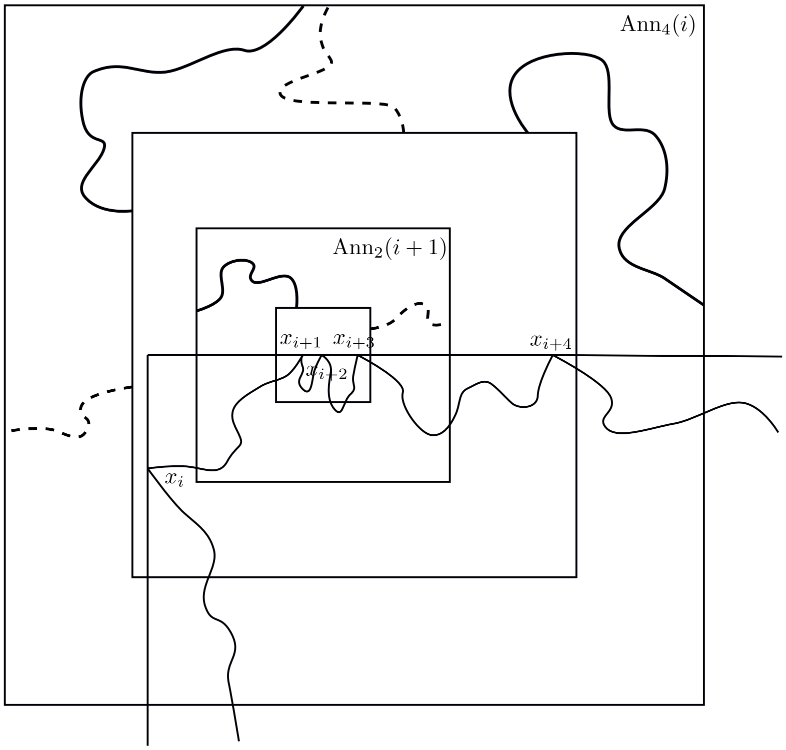


Figure 3.2: A representation of the crossings in a family of annuli: $x_i, x_{i+1}, x_{i+2}, x_{i+3}$, and x_{i+4} are five vertices on $\partial B(2^{n_1})$. There are two disjoint crossings, one open and one dual-closed, in the annulus $\mathrm{Ann}_2(i+1)$. There are four disjoint crossings, alternatingly dual-closed and open, in the annulus $\mathrm{Ann}_4(i)$.

where ξ' is some boundary condition on $B(2^{n_2})$ induced by reversely applying (DMP).

It remains to "glue" the crossings that occur in the $\tilde{E}_j$ events together so that $\tilde{\mathscr{A}}(R)$ occurs, which we refer to Section 1.5.1 for details on the gluing constructions. We make a special note that connecting a crossing in the inner-most annulus inward to the boundary of R (α_i or β_i) has a constant cost due to (RSW). Then, there exists c' that depends only on m such that

$$\phi^{\xi'}_{B(2^{n_2})}(\cap_{j=1}^{2m} \tilde{E}_j) \leq c' \phi^{\xi'}_{B(2^{n_2})}(\tilde{\mathscr{A}}(R)),$$

as desired.

We now formally define $\mathcal{I}_j$ and $\mathcal{L}_j$. Recall that $x_1, \ldots, x_{2m}$ are $2m$ vertices on

$\partial B(2^{n_1})$. Let us again use cyclic indexing, i.e. $i = i \bmod (2m)$. Recall further that each level-j annulus is crossed by j arms and encloses $j+1$ vertices. The purpose of defining these annuli (and groupings of vertices) is to identify which arms are close relative to the scale and which ones are far away.

We start with level 1. Let $\mathcal{I}_1$ be the collection of indices such that the index i is I_1 if the distance between x_i and x_{i+1} is logarithmically smaller than the distance between them and any other adjacent vertices, that is, $i \in \mathcal{I}_1$ if

$$|x_i - x_{i+1}| < 2^{-5} \cdot \min\{|x_{i-1} - x_i|, |x_{i+1} - x_{i+2}|\}. \tag{3.18}$$

For any $i \in \mathcal{I}_1$, we define

$$\ell_1(i) := \min\{\ell : \exists x \in 2^\ell \mathbb{Z}^2 \cap \partial B(2^{n_1}) \text{ such that } B(x, 2^\ell) \supset \{x_i, x_{i+1}\}\}.$$

Let $x_1(i)$ be the center for such a box $B(x_1(i), 2^{\ell_1(i)})$. If there are several choices for $x_1(i)$, we choose the first in lexicographical order. Next, we define

$$\ell_1'(i) := \min\{\ell \geq \ell_1(i) : B(x_1(i), 2^\ell) \ni x_{i-1} \text{ or } B(x_1(i), 2^\ell) \ni x_{i+2}\} - 3.$$

Condition (3.18) guarantees the existence of $x_1(i)$ and ensures that $\ell_1(i) < \ell_1'(i) \leq n_1$. Let $\mathrm{Ann}_1(i) := \mathrm{Ann}(x_1(i); 2^{\ell_1(i)}, 2^{\ell_1'(i)})$. Finally, we let $\mathcal{L}_1 := \{\mathrm{Ann}_1(i) : i \in \mathcal{I}_1\}$.

For $j = 2, \ldots, 2m-2$, we define $\mathcal{I}_j$ and $\mathcal{L}_j$ inductively. Let $\mathcal{I}_j$ again be a collection of indices. An index i is in $\mathcal{I}_j$ if

$$\max_{k,\ell \in \{i,\ldots,i+j\}} |x_k - x_\ell| < 2^{-5} \cdot \min\{|x_{i-1} - x_i|, |x_{i+j} - x_{i+j+1}|\}. \tag{3.19}$$

For any $i \in \mathcal{I}_j$, we define

$$\begin{aligned}\ell_j(i) := \min\{\ell : \exists x \in 2^\ell \mathbb{Z}^2 \cap \partial B(2^{n_1}) &\text{ such that} \\ &B(x, 2^\ell) \supset \{x_i, \ldots, x_{i+j}\} \cup (\cup_{h=1}^{j-1} \cup_{k=i}^{i+j} \mathrm{Ann}_h(k))\}\end{aligned}$$

where $\mathrm{Ann}_h(k) = \emptyset$ if $(k, k+1, \ldots, k+h) \notin \mathcal{I}_h$. Let $x_j(i)$ be the center for such a box $B(x_j(i), 2^{\ell_j(i)})$. If there are several choices for $x_j(i)$, we choose the first in lexicographical order. Next, we define

$$\ell'_j(i) := \min\{\ell \geq \ell_j(i) : B(x_j(i), 2^{\ell}) \ni x_{i-1} \text{ or } B(x_j(i), 2^{\ell}) \ni x_{i+j+1}\} - 3.$$

Again, condition (3.19) guarantees the existence of $x_j(i)$ and ensures that $\ell_j(i) < \ell'_j(i) \leq n_1$. We let $\mathrm{Ann}_j(i) := \mathrm{Ann}(x_j(i); 2^{\ell_j(i)}, 2^{\ell'_j(i)})$ and $\mathcal{L}_j := \{\mathrm{Ann}_j(i) : i \in \mathcal{I}_j\}$. Note that the definition of $\ell_j(i)$ ensures that level-j annuli are disjoint from annuli of lower levels.

The case $j = 2m-1$ is different from the previous cases: for there to be a level-$(2m-1)$ annulus, all $2m$ vertices must be concentrated relative to the scale of $\partial B(2^{n_1})$. We say $1 \in \mathcal{I}_{2m-1}$ if

$$\max_{k,\ell \in \{1,\ldots,2m\}} |x_k - x_\ell| < 2^{n_1-5}.$$

If $\mathcal{I}_{2m-1}$ is nonempty, we define ℓ_{2m-1} similar to before:

$$\begin{aligned} \ell_{2m-1} := \min\{\ell : \exists x \in 2^{\ell}\mathbb{Z}^2 \cap \partial B(2^{n_1}) \text{ such that} \\ B(x, 2^{\ell}) \supset \{x_1, \ldots, x_{2m}\} \cup (\cup_{h=1}^{2m-2} \cup_{k=1}^{2m} \mathrm{Ann}_h(k))\} \end{aligned}$$

Let x_{2m-1} be the center for such a box $B(x_{2m-1}, 2^{\ell_{2m-1}})$. If there are several choices for x_{2m-1}, we choose the first in lexicographical order. For the second-to-last level, we define

$$\ell'_{2m-1} := n_1 - 2.$$

Similarly to before, we define $\mathrm{Ann}_{2m-1}(1) := \mathrm{Ann}(x_{2m-1}; 2^{\ell_{2m-1}}, 2^{\ell'_{2m-1}})$. Let $\mathcal{L}_{2m-1} = \{\mathrm{Ann}_{2m-1}(1)\}$ if $\mathcal{I}_{2m-1}$ is nonempty and empty otherwise.

Finally, we define $\mathcal{L}_{2m} := \{\mathrm{Ann}_{2m}(1)\} = \{\mathrm{Ann}(2^{n_1+1}, 2^{n_2})\}$.

□

Remark 3.2. *Although Lemmas 3.4.1 and 3.4.2 are stated for* $2m$ *alternating arms, the proof can be adapted to accommodate any color sequence such that the dual-closed arms and defect edges are matched, thus including the three-arm case. For consecutive open arms, the constructions in step one and two remain the same except there are multiple open arms emanating from* β_i *in the definition of* $\mathcal{A}(R)$*. This subsequently changes the definitions of* $E_j(i)$*, but the argument carries through as the arm separation statement still holds. Consecutive closed arms can be considered as having zero open arm between them, the argument for which follows the consecutive open arms case essentially.*

3.5 Estimating without Reimer's inequality

In the radial case in Chapter 2, there is no natural crossing like the lowest crossing to compare to. Instead, we consider "lowest-like" paths between successive circuits around the origin. One nuisance in this construction occurs when two circuits are close and there is not enough space for there to be three arms to a large distance. However, if this happens, closeby circuits form a bottleneck which implies an arm event with more than three arms. This ensures that the three-arm probability is an upper bound. The details of the construction are encapsulated in Lemma 2.2.3, and similarly in Lemmas 2.4.5 and 2.4.7.

Recall that $A_3(n_1, n_2)$ denotes the three-arm event in the annulus $\mathrm{Ann}(n_1, n_2)$ and $\pi_3(n_1, n_2)$ its probability with domain $B(n)$ and boundary condition ξ. For any $j \geq 3$, let $A_j(n_1, n_2)$ ($\pi_j(n_1, n_2)$, resp.) denote the polychromatic j-arm event (probability, resp.) with exactly $j - 1$ disjoint open arms and one dual-closed arm. Let $\pi'_j(n_1, n_2)$ denote the monochromatic j-arm probability. In an abuse of notation, for a box "centered at an edge", we write $B(e, n)$ in place of $B(e_x, n)$

where e_x denotes the first endpoint of the edge e in lexicographical order.

Lemma (Restatement of Lemma 2.2.3)**.** *Fix $\epsilon > 0$ and an integer R such that for any $0 \leq n_1 < n_2$, $\pi'_{2R+2}(n_1, n_2) \leq \pi_3(n_1, n_2)(n_1/n_2)^{\epsilon}$. Let $\mathcal{H}_R(e, M)$ be the event that there exist $0 = \ell_0 \leq \ell_1 \leq \cdots \leq \ell_R = \lfloor \log(M) \rfloor$:*

1. *$A_{3,OOC}(e, 2^{\ell_1 - 1})$ occurs;*
2. *for $i \geq 2$, if $\ell_{i-1} < \lfloor \log(M) \rfloor$, there are $2i$ disjoint open arms and one closed dual arm from $\partial B(e, 2^{\ell_{i-1}})$ to $\partial B(e, 2^{\ell_i - 1})$; and*
3. *if $\ell_R < \lfloor \log(M) \rfloor$, there are $2R + 2$ disjoint open arms from $\partial B(e, 2^{\ell_R})$ to $\partial B(e, M)$.*

Then,

$$\phi^{\xi}_{B(n)}(\mathcal{H}_R(e, M)) \leq cM^{-r}\pi_3(M).$$

The above estimate relies essentially on the following proposition.

Proposition 3.5.1. *Let $0 \leq n_1 < n_2$, $i \geq 1$,*

$$\pi_{2i+1}(n_1, n_2) \leq \left(\frac{n_2}{n_1}\right)^{(2i-2)\alpha_1} \pi_3(n_1, n_2).$$

where α_1 is the one-arm exponent.

In Bernoulli percolation, this is done by applying Reimer's inequality. A weak form of Reimer's inequality for the random cluster model can be found in [vdBG13]. However, it requires the events to not only have disjoint occurrences but also occur on disjoint clusters. The arms in the arm event π_{2i+1} that [SR22] concerns belong to the same cluster since they are portions of consecutive circuits chained together by a radial arm. Therefore, the weak estimate is not applicable to our problem. We provide a proof using conditional probability and quad-crossing (RSW).

Proof. It suffices to show that

$$\pi_{2i+1}(n_1, n_2) \le \left(\frac{n_2}{n_1}\right)^{\alpha_1} \pi_{2i}(n_1, n_2)$$

for $2i \ge 3$.

Since there is at least one dual-closed arm in any configuration of $A_{2i+1}(n_1, n_2)$, we condition on a dual-closed arm and the first open arm on its clockwise side and the (consecutive) first $2i-2$ disjoint open arms on its counterclockwise side and apply (DMP). As in the proof of Claim 3.1, let $\mathcal{U}$ denote the explorable (random) region that contains the $2i$ arms and whose boundaries consist of a dual-closed arm, an open arm, and portions of $\partial B(n_1)$ and $\partial B(n_2)$. Then, conditional on $\mathcal{U}$, the complement region is independent up to some boundary condition.

$$\begin{aligned}\phi^{\xi}_{B(n)}(A_{2i+1,CO\cdots O}(n_1, n_2)) &= \sum_{\text{admissible } U} \phi^{\xi}_{B(n)}(A_{1,O}(n_1, n_2, U^c) \mid \mathcal{U} = U)\phi^{\xi}_{B(n)}(\mathcal{U} = U) \\ &\le \sum_{\text{admissible } U} \phi^{\eta}_{B(n)\setminus U}(A_1(n_1, n_2, U^c))\phi^{\xi}_{B(n)}(\mathcal{U} = U) \qquad (3.20)\end{aligned}$$

where $A_1(n_1, n_2, U^c)$ is the one arm event in $\mathrm{Ann}(n_1, n_2)$ restricted to U^c and η is uniquely determined by ξ and U. The one-arm event in $\mathrm{Ann}(n_1, n_2) \setminus U$ can be estimated using quad-crossing (RSW) and an analogue of (QM) $\log(n_2/n_1)$ times as follows:

$$\phi^{\xi}_{B(n)\setminus U}(A_1(n_1, n_2, U^c)) \le c \prod_{\ell=\lfloor \log(n_1)\rfloor}^{\lceil \log(n_2)\rceil - 1} \phi^{\eta}_{B(n)\setminus U}(A_1(2^{\ell}, 2^{\ell+1}, U^c)) \le c c_4^{\log(n_2/n_1)+1}.$$

Applying the above estimate into (3.20) and we have

$$(3.20) \le \left(\frac{n_2}{n_1}\right)^{c'} \sum_{\text{admissible } U} \phi^{0}_{B(n)}(\mathcal{U} = U) = \left(\frac{n_2}{n_1}\right)^{c'} \phi^{0}_{B(n)}(A_{2i,CO\cdots O}(n_1, n_2)).$$

We note that to apply quad-crossing (RSW), the extremal distance for each quad $\mathrm{Ann}(2^{\ell}, 2^{\ell+1}) \setminus U$, which for convenience we call $\mathcal{D}$ here, needs to be uniformly lower bounded over all admissible U. To our advantage, bottlenecks in $\mathcal{D}$ make

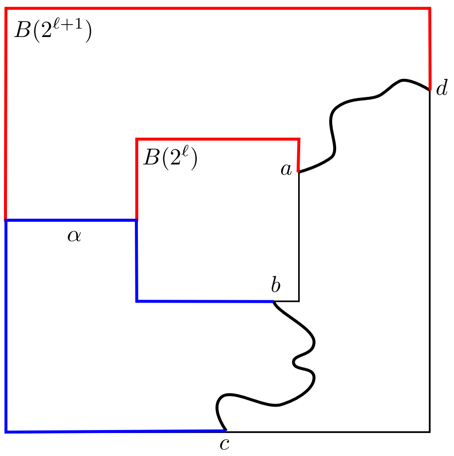


Figure 3.3: Two quads defined by the same four endpoints. The blue arc is $(bc)'$ and the red arc is $(da)'$. They both traverse through α.

the extremal distance larger. The boundary of $\mathcal{D}$ defines four arcs: (ab) on $\partial B(2^{\ell})$, (cd) on $\partial B(2^{\ell+1})$, and (bc) and (da) in the interior of $\mathrm{Ann}(2^{\ell}, 2^{\ell+1})$. Indeed, if $\mathcal{D}$ is contained in another quad $\mathcal{D}'$ with the same landing arcs as $\mathcal{D}$, then

$$\ell_{\mathcal{D}}[(ab),(cd)] \geq \ell_{\mathcal{D}'}[(ab),(cd)]. \tag{3.21}$$

We verify (3.21) in Appendix 3.A. Let α be a topological path in U, disjoint from (bc) and (da), and $\mathcal{D}'$ be all of $\mathrm{Ann}(2^{\ell}, 2^{\ell+1})$ with arcs (ab), (cd), and $(bc)'$, $(da)'$, where $(bc)'$ consists of a portion of $\partial B(2^{\ell})$, α, and a portion of $\partial B(2^{\ell+1})$ and similarly, $(da)'$ consists of another portion of $\partial B(2^{\ell})$, α, and another portion of $\partial B(2^{\ell+1})$. Clearly, $\mathcal{D}$ is contained in $\mathcal{D}'$, see Figure 3.3. Then,

$$\ell_{\mathcal{D}'}[(ab),(cd)] \geq \frac{2^{\ell}}{\min\{\#(ab), \#(cd)\}} \geq \frac{1}{16}.$$

$\square$

APPENDIX

3.A Extremal distance and resistance

In this section, we verify (3.21) through the definition of extremal distance by the resistance of an electrical network.

Definition 3.2 ([CDCH16]).

$$\ell_\Omega[(ab),(cd)] := \sup_{g:\mathcal{E}(\Omega)\to\mathbb{R}_+} \frac{\left[\inf_{\gamma:(ab)\overset{\Omega}{\leftrightarrow}(cd)} \sum_{e\in\gamma} g_e\right]^2}{\sum_{e\in\mathcal{E}(\Omega)} g_e^2}.$$

Let Ω_2 be a rectangle with vertices a, b, c, d, labeled in counterclockwise order. Then, the arcs (ab), (bc), (cd), and (da) are the four sides of the rectangle. Let $(bc)'$ be an arc from b to c contained in Ω_2, and $(da)'$ be an arc from d to a contained in Ω_2. Then, Ω_1 bounded by (ab), $(bc)'$, (cd), and $(da)'$ is a subdomain of Ω_2. We want to show

$$\ell_{\Omega_1}[(ab),(cd)] \geq \ell_{\Omega_2}[(ab),(cd)]. \tag{3.22}$$

For any fixed $g : \mathcal{E}(\Omega_2) \to \mathbb{R}_+$, since $\mathcal{E}(\Omega_1) \subset \mathcal{E}(\Omega_2)$, we have $\{\gamma : (ab) \overset{\Omega_1}{\leftrightarrow} (cd)\} \subset \{\gamma : (ab) \overset{\Omega_2}{\leftrightarrow} (cd)\}$. Then,

$$\inf_{\gamma:(ab)\overset{\Omega_1}{\leftrightarrow}(cd)} \sum_{e\in\gamma} g_e \geq \inf_{\gamma:(ab)\overset{\Omega_2}{\leftrightarrow}(cd)} \sum_{e\in\gamma} g_e.$$

For the denominator, we use $\mathcal{E}(\Omega_1) \subset \mathcal{E}(\Omega_2)$ again:

$$\sum_{e\in\mathcal{E}(\Omega_1)} g_e^2 \leq \sum_{e\in\mathcal{E}(\Omega_2)} g_e^2.$$

Therefore,

$$\frac{\left[\inf_{\gamma:(ab)\overset{\Omega_1}{\leftrightarrow}(cd)} \sum_{e\in\gamma} g_e\right]^2}{\sum_{e\in\mathcal{E}(\Omega_1)} g_e^2} \geq \frac{\left[\inf_{\gamma:(ab)\overset{\Omega_2}{\leftrightarrow}(cd)} \sum_{e\in\gamma} g_e\right]^2}{\sum_{e\in\mathcal{E}(\Omega_2)} g_e^2}.$$

(3.22) follows from taking supremum over all $g : \mathcal{E}(\Omega_2) \to \mathbb{R}_+$.

CHAPTER 4

EQUIVALENCE OF POLYCHROMATIC ARM PROBABILITIES ON THE SQUARE LATTICE

4.1 Introduction

We consider critical Bernoulli percolation on the square lattice $\mathbb{Z}^2$ where $p = p_c = \frac{1}{2}$. *Arm events*, defined by the simultaneous occurrence of several long-range connections across annuli, have been extensively studied in this setting, especially since the pioneering works by Kesten in the 1980s, see [Kes80, Kes86, Kes87, SW01]. See [Nol08] for a survey.

Configurations in a k-arm event have k disjoint paths, each of a single color, open or dual-closed, connecting the two components of the boundary of an annulus. A k-arm event is said to be *polychromatic* if not all paths in the configuration are of the same color. It is well-known that for critical *site* percolation on the triangular lattice, probabilities of different polychromatic arm events are comparable up to constants that are independent of the size of the annulus. In this chapter, we prove that the same holds on the square lattice.

Some of the key results accessible for site percolation on the triangular lattice remain out of reach for the square lattice. While many of the techniques involving the use of correlation inequalities, like the generalized (FKG) inequality and van den Berg–Kesten (Reimer) inequality, can be applied to very general lattices, intrinsic differences in the duality relations satisfied by the two models have so far impeded the extension of key results for critical and near-critical percolation on the triangular lattice to bond percolation on the square lattice.

For example, the arm exponents for critical site percolation on the triangular lattice are known [LSW02, SW01]: $\alpha_1 = 5/48$ and $\alpha_k = (k^2 - 1)/12$ for $k \geq 2$. Since this result relies on approximation by the Schramm–Loewner Evolution (SLE) process, the values of arm exponents on the square lattice remain conjectural.

The proof of the equivalence of different polychromatic arm probabilities on the triangular lattice uses *color-switching*, a combinatorial trick with a venerable history in critical percolation, see for example [ADA99]. Most notably, an *exact* version of color-switching is used in Smirnov's celebrated proof of conformal invariance for critical site percolation [Smi01].

We adapt to the square lattice an approximate color-switching argument appearing in [Nol08]. It shows that arm probabilities for different color sequences of the same length are asymptotically equivalent, as long as they contain arms of both colors:

Proposition 4.1.1. *Consider percolation on the square lattice and let $k \geq 3$, $n_0(k) < n < N$. Let σ, σ' be two polychromatic color sequences of length k. Denote by $A_{k,\sigma}(n, N)$ the event that there exist k disjoint arms, color-coded by σ, from scale n to N. (See Sections 1.4 and 4.1.1 for precise definitions.) Then, there exists a constant $C = C(k)$ independent of σ, σ', n, and N such that*

$$\mathbb{P}(A_{k,\sigma}(n, N)) \leq C\mathbb{P}(A_{k,\sigma'}(n, N)).$$

This proposition is not unexpected, but we could not locate it in the literature. Our method, using a measure-preserving shift, has also not appeared in this context. We therefore found it worthwhile to provide a complete proof.

4.1.1 Notation

In this chapter, we often work with the "flipped" configuration. Therefore it is important to not only consider (primal)-open and dual-closed connections but also the dual-open and primal-closed ones. For this reason, we give altered notations and definitions for some events defined in Section 1.4.

A *color sequence* σ of length k is a sequence $(\sigma_1, \ldots, \sigma_k) \in \{O, O^*, C, C^*\}^k$. Each σ_i indicates a color. The colors are encoded O for open, O^* for dual-open, C for closed, and C^* for dual-closed. In percolation theory, the most important color sequences consist of primal-open and/or dual-closed connections since primal-open paths and dual-closed paths must be mutually disjoint. However, in the proof of the main proposition, it will be useful to consider primal-closed and dual-open arms as well.

We use the convention $(O^*)^* = O$, $(C^*)^* = C$, and $\sigma^* = (\sigma_1, \ldots, \sigma_k)^* = (\sigma_1^*, \ldots, \sigma_k^*)$. Similarly, $\overline{\sigma}$ denotes the *flipped* color sequence, with the conventions $\overline{O} = C$, $\overline{C} = O$, $\overline{O^*} = C^*$, $\overline{C^*} = O^*$, and $\overline{(\sigma_1, \ldots, \sigma_k)} = (\overline{\sigma_1}, \ldots, \overline{\sigma_k})$.

A primal-open (-closed, resp.) arm in $\mathrm{Ann}(n, N)$ connecting $\partial B(n)$ and $\partial B(N)$ is a path of open (closed, resp.) edges in $\mathrm{Ann}(n, N)$ with one endpoint lying in $\partial B(n)$ and another endpoint in $\partial B(N)$.

We say a *dual*-closed (-open, resp.) arm in $\mathrm{Ann}(n, N)^*$ *connects* $\partial B(n)$ and $\partial B(N)$ if the (primal) path obtained by shifting by $(-1/2, -1/2)$ connects $\partial B(n)$ and $\partial B(N)$.

Definition 4.1. For $n \le N$, we define the k-arm event with color sequence σ to be the event that there are k disjoint paths whose colors are specified by σ in clockwise order in the annulus $\mathrm{Ann}(n, N)$ connecting $\partial B(n)$ and $\partial B(N)$.

Color sequences that are equivalent up to cyclic order denote the same arm event. As in Section 1.4, we define $n_0(k)$ to be the smallest integer such that $|\partial B(n_0(k))| \geq k$.

4.1.2 Organization

We use the basic idea of color-switching. Unlike in site percolation on the triangular lattice, when flipping the status of edges in a region, a primal-open arm becomes primal-closed instead of dual-closed; a dual-closed arm becomes dual-open instead of primal-open. To address this problem, we introduce a shifting transformation in Section 4.2 to convert between the two lattices. To effectively apply the transformation to the region bounded by two arms in an annulus, the arms cannot come too close to the boundaries. In Section 4.3, we show that, at the cost of a constant factor, we may assume that the arms remain at least a fixed distance away from each other. In the final part of Section 4.3, we prove the main result.

4.2 A shifting transformation

Our first lemma is the main ingredient of the proof of our result. We introduce a transformation that shifts a configuration in a region by $(1/2, 1/2)$ to convert arms between the primal and the dual lattices.

Let us begin by defining what is meant by a *region* of an annulus bounded by two curves. Let $n < N$ and let γ_1 and γ_2 be two (primal or dual) paths connecting $\partial B(n)$ and $\partial B(N)$. In particular, if γ_1 or γ_2 is a dual path, we use "connecting" as in the definition of the arm events above, which is distinct from the usual topological sense. Consider γ_1, γ_2, as well as $\partial B(n)$, $\partial B(N)$ as curves in $\mathbb{R}^2$. We define $\varphi(\gamma_1, \gamma_2) \subset \mathrm{Ann}(n, N)$ to be the Jordan curve[1] obtained by concatenating $\gamma_1 \cap \mathrm{Ann}(n, N)$, a portion of $\partial B(N)$, $\gamma_2 \cap \mathrm{Ann}(n, N)$, and a portion of $\partial B(n)$ in counterclockwise orientation. If γ_1 or γ_2 lies on the dual lattice and does not (topologically) connect $\partial B(n)$ and $\partial B(N)$, we add the shortest line segment connecting the endpoints of γ_1 to $\partial B(n)$ or $\partial B(N)$ to ensure we obtain a closed curve.

Definition 4.2. A *region* is a connected set of edges. Given $\varphi(\gamma_1, \gamma_2)$ as above, the *region in* $\mathrm{Ann}(n, N)$ *with boundary* φ is the set S of edges that lie in the interior of φ, together with all edges of $(\varphi \cap \partial B(n)) \cup (\varphi \cap \partial B(N))$.

Definition 4.3. For a region $S \subset \mathcal{E}$, we define the event $A_{k,\sigma}(S)$ by the occurrence of the event $A_{k,\sigma}(n, N)$, with the additional condition that the k arms consist of edges of S or edges dual to S. The arms in $A_{k,\sigma}(S)$ are necessarily disjoint from $\partial S \setminus (\partial \mathrm{Ann}(n, N))$.

We introduce a variant of the arm events $A_{k,\sigma}$ with an additional separation condition. Fix some integer constant $\ell \geq 5$, the separation between the arms.

Definition 4.4. Let $\tilde{A}_{k,\sigma}(n, N)$ be the event that there are k disjoint arms from distance n to N, color-coded by σ and any two of the arms are at distance at least ℓ in $\mathrm{Ann}(2n, N/2)$.

For a region S with boundary φ, we define $\tilde{A}_{k,\sigma}(S)$ to be the event the event that there are k disjoint arms connecting $\partial B(n)$ and $\partial B(N)$ in S, color-coded by σ, and the k arms are at distance at least ℓ from each other and the curve φ.

---

[1] A Jordan curve is a simple closed curve in $\mathbb{R}^2$.

Lemma 4.2.1. *Let $k \geq 2$ and let σ be some color sequence of length k. For $n_0(k) < n < N$, let $\varphi = \varphi(\gamma_1, \gamma_2)$ be a Jordan curve given by two disjoint arms γ_1, γ_2, and $S \subset \mathcal{E}$ be the region with boundary φ, excluding any edges in $\gamma_1 \cap \gamma_2$. Then,*

$$\mathbb{P}(\tilde{A}_{k,\sigma}(S)) \leq \mathbb{P}(A_{k,\sigma^*}(S \cap \operatorname{Ann}(2n, N/2))).$$

Proof. Our goal is to define an invertible and thus measure preserving transformation T on configurations of the edges in S such that

$$T(\tilde{A}_{k,\sigma}(S)) \subset A_{k,\sigma^*}(S \cap \operatorname{Ann}(2n, N/2)). \tag{4.1}$$

Since all configurations have equal probability, we have

$$\mathbb{P}(\tilde{A}_{k,\sigma}(S)) = \sum_{\omega \in \tilde{A}_{k,\sigma}(S)} \mathbb{P}(\omega) = \sum_{\omega \in \tilde{A}_{k,\sigma}(S)} \mathbb{P}(T(\omega))$$

Using the bijectivity of T and (4.1) respectively, the above is bounded by

$$\begin{aligned}\mathbb{P}(\tilde{A}_{k,\sigma}(S)) &\leq \sum_{\omega' \in T(\tilde{A}_{k,\sigma}(S))} \mathbb{P}(\omega') \\ &\leq \mathbb{P}(A_{k,\sigma^*}(S \cap \operatorname{Ann}(2n, N/2))).\end{aligned}$$

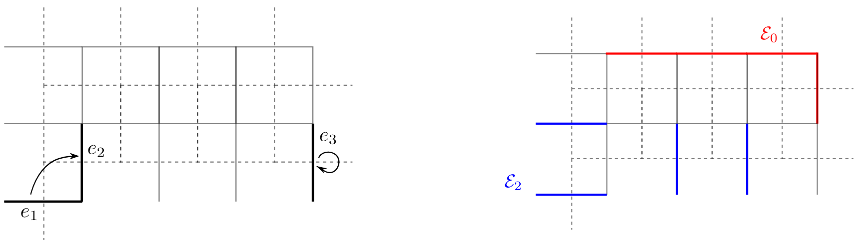
This completes the proof provided the transformation T exists.

To define T, we first choose some deterministic ordering of all edges in $B(N)$. This induces an ordering of the edges in S, which we enumerate as $e_1, \ldots, e_m$.

Given an initial configuration $(\omega(e_1), \ldots, \omega(e_m)) \in \{0, 1\}^S$, we determine an image configuration

$$(\omega'(e_1), \ldots, \omega'(e_m))$$

as an intermediate step to defining T by the following correspondence:



(a) e_2 inherits its status in ω' from e_1 (in ω). e_3 inherits its status in ω' from itself.

(b) Under transformation T, the red edges belong to $\mathcal{E}_0$; the blue edges belong to $\mathcal{E}_2$; and the black edges belong to $\mathcal{E}_1$.

Figure 4.1: All solid lines are edges in S and all dotted lines are dual edges of the edges in S.

- If $(e_i)^* - (1/2, 1/2) \in S$, then we let

  $$\omega'^*((e_i)^*) = \omega'(e_i) := \omega\left((e_i)^* - \left(\frac{1}{2}, \frac{1}{2}\right)\right).$$

  (Note that $(e_i)^* - (1/2, 1/2)$ is an edge on the primal lattice $\mathcal{E}$.) In this case, we say that e_i *inherits* its status in ω' from (the status of) $(e_i)^* - (1/2, 1/2)$ (in ω).

- If $(e_i)^* - (1/2, 1/2) \notin S$, then the status of e_i in ω' remains the same as in ω:

  $$\omega'(e_i) := \omega(e_i)$$

  In this case, we say e_i *inherits* its status in ω' from itself.

See Figure 4.1a for an illustration.

We classify the edges of S into three sets:

1. An edge e is in $\mathcal{E}_0(T)$ if no edge inherits its status in ω' from e in ω.
2. An edge e is in $\mathcal{E}_1(T)$ if exactly one edge, including possibly e itself, inherits its status in ω' from e in ω.

3. An edge e is in $\mathcal{E}_2(T)$ if two edges, including possibly e itself, inherit their status in ω' from e in ω.

Notice that the sets $\mathcal{E}_i$, $i = 0, 1, 2$ do not depend on ω: $e \in \mathcal{E}_0$ if $e^* - (1/2, 1/2) \in S$ but $e^* + (1/2, 1/2) \notin S$, $e \in \mathcal{E}_2$ if $e^* + (1/2, 1/2) \in S$ but $e^* - (1/2, 1/2) \notin S$, and an edge e is in $\mathcal{E}_1$ if either $e^* \pm (1/2, 1/2) \in S$ or $e^* \pm (1/2, 1/2) \notin S$.

By counting the number of edges inheriting their status from each of the sets $\mathcal{E}_i$, $i = 0, 1, 2$, we have:

$$|S| = |\mathcal{E}_0| + |\mathcal{E}_1| + |\mathcal{E}_2| = 0 \cdot |\mathcal{E}_0| + |\mathcal{E}_1| + 2|\mathcal{E}_2|$$

Therefore, $|\mathcal{E}_0| = |\mathcal{E}_2|$. See Figure 4.1b for an illustration.

We now assign new statuses to the edges in $\mathcal{E}_2$. Enumerate the edges in $\mathcal{E}_0$ and $\mathcal{E}_2$ according to the deterministic order fixed in the beginning so that

$$\mathcal{E}_0 = (e_{\alpha_1}, \ldots, e_{\alpha_K}),$$
$$\mathcal{E}_2 = (e_{\beta_1}, \ldots, e_{\beta_K})$$

where $K = |\mathcal{E}_0| = |\mathcal{E}_2|$. Set

$$T(\omega)(e_{\beta_i}) := \omega(e_{\alpha_i}), \quad i = 1, \ldots, K.$$

For the remaining edges $e \notin \mathcal{E}_2$, we let

$$T(\omega)(e) := \omega'(e).$$

This definition guarantees that T is invertible. It is easily checked that its inverse is the following map T': given an initial configuration $(\omega'(e_1), \ldots, \omega'(e_m))$, first assign the status of each edge e_i as follows.

$$\omega(e_i) = \begin{cases} \omega'((e_i)^* + (\frac{1}{2}, \frac{1}{2})) & \text{if } (e_i)^* + (\frac{1}{2}, \frac{1}{2}) \in S, \\ \omega(e_i) & \text{otherwise.} \end{cases} \tag{4.2}$$

Next, define $\mathcal{E}'_0 = \mathcal{E}_0(T')$ to be the set of edges e such that no edge inherits its status in ω from e, and $\mathcal{E}'_2 = \mathcal{E}_2(T')$ to be the set of edges such that two edges inherit their status in ω from e. Here, *inheritance* is defined as it was used in the definition of the transformation T. Then,

$$\mathcal{E}'_2 = \mathcal{E}_0(T),$$

$$\mathcal{E}'_0 = \mathcal{E}_2(T).$$

For $e \notin \mathcal{E}'_2$, we define

$$T'(\omega')(e) = \omega(e),$$

with ω as in (4.2). For $e \in \mathcal{E}'_2 = \mathcal{E}_0(T) = (e_{\alpha_1}, \ldots, e_{\alpha_K})$, we let

$$T'(\omega'(e_{\alpha_i})) = \omega'(e_{\beta_i}).$$

We now show (4.1). For any arm γ in a configuration $\omega \in \tilde{A}_{k,\sigma}(S)$, the translated path γ^* is contained in S at least from $\partial B(2n)$ to $\partial B(N/2)$, since the original arms are ℓ-separated in the sub-annulus $\mathrm{Ann}(2n, N/2)$. Finally, $\gamma^* \subset \mathrm{Ann}(n+1, N-1)$ in the configuration $T(\omega)$ receives the same color as γ in the configuration ω since all but two edges (the two edges with endpoints lying on $\partial B(n)$ or $\partial B(N)$) in γ^* lie in the set $\mathcal{E}_1$, and thus receive their status from the edges in ω.

□

4.3 Proof of Proposition 4.1.1

To apply Lemma 4.2.1 in the proof of the main result, we need the arms in the shifted region to be at least distance ℓ apart. This is well known to happen with high probability. Indeed, if any two arms come close, they form a bottleneck that generates a 6-arm event. Since we could not locate a proof that applies in our exact situation, we provide one for completeness. The point of the following lemma 4.3.1 is that, to show that a six-arm event is unlikely to appear anywhere in a box using existing results, we must ensure that we witness the arms in the right order, that is, they must contain an *alternating* 5-arm subsequence.

For $\delta > 0$, let $m = m(\delta)$ be the least integer such that $(\pi_1(n))^{2m+2} > n^{-2-\delta}$, and let $L = 2^{\lfloor \log(d(e)/2) \rfloor}$.[2]

Lemma 4.3.1. *Let* $d(e) = \mathrm{dist}(e, \partial B(n))$, $k \geq 4$ *and* $8\ell n_0(k) \leq 8n \leq N$. *Suppose* $A_{k,\sigma} \setminus \tilde{A}_{k,\sigma}$ *occurs, where* $\sigma = (\sigma_1, \ldots, \sigma_k)$. *There are two arms* γ_{-1} *and* γ_1 *from* $\partial B(n)$ *and* $\partial B(N)$ *such that the shortest distance between edges of* γ_1 *and* γ_{-1} *in* $\mathrm{Ann}(2n, N/2)$ *is less than* ℓ *and* γ_{-1} *is of color* σ_i, γ_1 *is of color* σ_{i+1} *for some* $1 \leq i \leq k$. *We relabel the colors* σ_i *and* σ_{i+1} *as* β_{-1} *and* β_1, *respectively.*

Then there exists an edge $e \in \mathrm{Ann}(2n, N/2)$ *and integer scales* $\lceil \log \ell \rceil = \ell_0 \leq \ell_1 \leq \cdots \leq \ell_{m-1} \leq \log(d(e)/2)$, *such that the following event* $E(e)$ *occurs:*

1. *For* $0 \leq i \leq m-2$, *if* $\ell_{i+1} \geq \ell_i + 2$, *there are* $6 + 2i$ *disjoint arms from* $\partial B(e, 2^{1+\ell_i})$ *to* $\partial B(e, 2^{\ell_{i+1}})$. *There exists an integer* $0 \leq j \leq i$ *such that the* $6+2i$ *arms appear with the color sequence* Σ_i *given by:*

$$\beta_{-1},\ \beta_{-2}, \ldots, \beta_{-i+j-1}, \overline{\beta_{-i+j-1}},\ \beta_{-i+j-1}, \ldots$$
$$\beta_{-2},\ \beta_{-1},\ \beta_1,\ \beta_2, \ldots,\ \beta_{j+1},\ \overline{\beta_{j+1}},\ \beta_{j+1}, \ldots,\ \beta_2,\ \beta_1.$$

2. *If* $\ell_{m-1} + 1 < \log(d(e)/2)$, *there are* $2m+2$ *disjoint arms from* $\partial B(e, 2^{1+\ell_{m-1}})$ *to* $\partial B(e, L)$. *There exists an integer* $0 \leq j \leq m-1$ *such that the* $2m+2$ *arms appear*

---

[2] log denotes the logarithm with base 2.

in the sequence Σ_{m-1}*:*

$$\beta_{-1},\ \beta_{-2}, \dots,\ \beta_{-m+j},\ \beta_{-m+j}, \dots,\ \beta_{-2},\ \beta_{-1},\ \beta_1,\ \beta_2, \dots,\ \beta_{j+1},\ \beta_{j+1}, \dots,\ \beta_2,\ \beta_1.$$

Proof. We first discuss how to determine γ_{-1} and γ_1. On the event that $A_{k,\sigma}$ occurs but $\tilde{A}_{k,\sigma}$ does not, there are k disjoint arms with the color sequence σ and there exists a pair of two arms such that the shortest distance between them in $\mathrm{Ann}(2n, N/2)$ is less than ℓ. These arms are of colors σ_i and σ_{i+1} for some i between 1 and k. If several choices are available, choose the first pair of arms in some fixed deterministic order on all finite paths in the annulus. We denote the arm of color σ_i by $\tilde{\gamma}_{-1}$ and the arm of color σ_{i+1} by $\tilde{\gamma}_1$. Since the two arms come within ℓ, the set $\mathcal{E} \subset \mathrm{Ann}(2n, N/2)$ of edges e such that $B(e, \ell)$ intersects both $\tilde{\gamma}_{-1}$ and $\tilde{\gamma}_1$ is non-empty. We choose $e \in \mathcal{E}$ such that $\mathrm{dist}(e, \tilde{\gamma}_{-1}) + \mathrm{dist}(e, \tilde{\gamma}_1)$ is minimized. We then choose γ_{-1} and γ_1, two arms of colors σ_i and σ_{i+1} such that both arms intersect $B(e, \ell)$ and the area enclosed by $\gamma_{-1}, \gamma_1, \partial B(n)$ and $\partial B(N)$ containing e is maximized among all pairs $(\tilde{\gamma}_{-1}, \tilde{\gamma}_1)$. We assume that this region lies clockwise of γ_{-1} inside $\mathrm{Ann}(n, N)$ and counterclockwise of γ_1. Otherwise, exchange their labels for the rest of the proof.

From now on, we relabel the colors σ_i and σ_{i+1} as β_{-1} and β_1. There are four disjoint arms, two of color β_{-1} and two of color β_1, from $B(e, \ell)$ to $B(e, d(e)/2)$, obtained by following γ_{-1} and γ_1 in either direction from inside $B(e, \ell)$ outward. See Figure 4.2.

Let γ_{-2} be the open or dual-closed arm disjoint from γ_{-1} such that the area enclosed by γ_{-1}, γ_{-2}, $\partial B(n)$, and $\partial B(N)$ on γ_{-1}'s counterclockwise side is minimized. That is, γ_{-2} is the "closest" disjoint arm on γ_{-1}'s counterclockwise side. We denote the color of γ_{-2} by β_{-2}. Similarly, let γ_2 denote the open or dual-closed arm

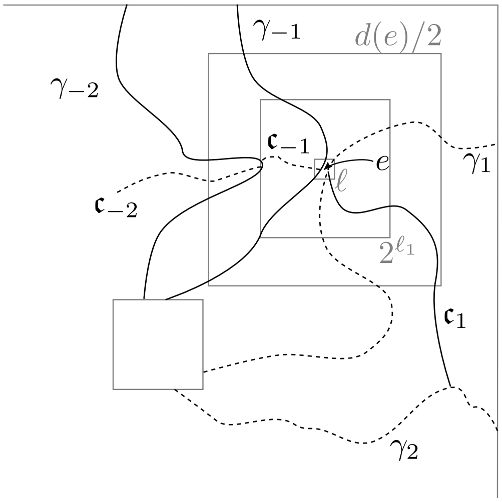


Figure 4.2: A representation of the arm construction in the proof of Lemma 4.3.1. The outer grey boundary represents a portion of $\partial B(N)$; the small grey box in the bottom left represents $\partial B(n)$.

disjoint from γ_1 such that the area enclosed by $\gamma_1, \gamma_2, \partial B(n)$, and $\partial B(N)$ on γ_1's clockwise side is minimized. We denote the color of γ_2 by β_2. By duality and the minimality of γ_{-2}, there is a $\overline{\beta_{-1}}$ arm, say $\mathfrak{c}_{-1}$, from an edge of γ_{-1} in $B(e, \ell)$ to γ_{-2}. Similarly, there is a $\overline{\beta_1}$ arm, denoted $\mathfrak{c}_1$, from an edge of γ_1 in $B(e, \ell)$ to γ_2. See Figure 4.2.

Let $\ell' \geq \log \ell$ be the largest integer such that neither γ_{-2} nor γ_2 intersects $B(e, 2^{\ell'})$. If $2^{\ell'} \geq L$, we let $\ell_1 = \lfloor \log(d(e)/2) \rfloor$. There are six arms in the annulus $\mathrm{Ann}(e; \ell, d(e)/2)$: a portion of γ_1, a portion of $\mathfrak{c}_1$, another portion of γ_1, a portion of γ_{-1}, a portion of $\mathfrak{c}_{-1}$, and another portion of γ_{-1}. Otherwise if $2^{\ell'} < L$, we let $\ell_1 = \ell'$ and we have the same arms as in the previous case crossing $\mathrm{Ann}(e; \ell, 2^{\ell_1})$ instead. In this case, without loss of generality, we assume γ_{-2} intersects $B(e, 2^{\ell_1+1})$. We associate to the scale ℓ_1 the collection

$$C_1 = \{\gamma_{-2}, \gamma_{-1}, \gamma_1\}$$

of three arms intersecting $B(e, 2^{\ell_1+1})$.

For any $1 \leq i \leq m-2$, we inductively define the scale ℓ_{i+1}. If $\ell_i = \lfloor \log(d(e)/2) \rfloor$, we let $\ell_r = \ell_i$ for $r \geq i+1$. Otherwise, $\ell_i < \lfloor \log(d(e)/2) \rfloor$ and there is a collection

$$C_i = \{\gamma_{-i+j-1}, \ldots, \gamma_{-1}, \gamma_1, \ldots, \gamma_{j+1}\}$$

of $i+2$ arms associated with the scale ℓ_i, with j ($j \leq i$) arms on the clockwise side of γ_1. These arms all intersect $B(e, 2 \cdot 2^{\ell_i})$. Their colors are labeled $\beta_{-i+j-1}, \ldots, \beta_{-1}, \beta_1, \ldots, \beta_{j+1}$. Let γ_{-i+j-2} be the open or dual-closed arm disjoint from γ_{-i+j-1} such that the area of the region enclosed by $\gamma_{-i+j-1}, \gamma_{-i+j-2}, \partial B(n)$, and $\partial B(N)$ on γ_{-i+j-1}'s counterclockwise side is minimized. Its color is labeled β_{-i+j-2}. Similarly, we define γ_{j+2} the closest arm to the clockwise side of γ_{j+1} and label its color β_{j+2}.

By duality and the minimality of γ_{-i+j-2}, there is a $\overline{\beta_{-i+j-1}}$ arm, $\mathfrak{c}_{-i+j-1}$, from an edge of γ_{-i+j-1} in $B(e, 2^{1+\ell_{i-1}})$ to γ_{-i+j-2}. Similarly, there is a $\overline{\beta_{j+1}}$ arm, $\mathfrak{c}_{j+1}$, from an edge of γ_{j+1} in $B(e, 2^{1+\ell_i})$ to γ_{j+2}.

Let $\ell'' \geq \ell_i$ be the largest integer such that neither γ_{-i+j-2} nor γ_{j+2} intersects $B(e, 2^{\ell''})$. If $2^{\ell''} \geq L$, we let $\ell_{i+1} = \lfloor \log(d(e)/2) \rfloor$. Otherwise, we let $\ell_{i+1} = \ell''$. As in the case $i = 1$, we form C_{i+1} from C_i by adding the arm, either γ_{-i+j-2} or γ_{j+2}, that intersects $B(e, 2^{1+\ell_{i+1}})$ for this scale. Suppose, without loss of generality, that the newly recorded arm is γ_{j+2}:

$$C_{i+1} := C_i \cup \{\gamma_{j+2}\}.$$

The $i+3$ arms in C_{i+1} cross the box $B(e, 2^{1+\ell_{i+1}})$. We associate the collection C_{i+1} to the scale ℓ_{i+1}. Each of the arms in C_i crosses the annulus $\mathrm{Ann}(e; 2^{1+\ell_i}, 2^{\ell_{i+1}})$ twice. In addition, both $\mathfrak{c}_{-i+j-1}$ and $\mathfrak{c}_j$ cross $\mathrm{Ann}(e; 2^{1+\ell_i}, 2^{\ell_{i+1}})$. Thus, the annulus is crossed by $6 + 2i$ arms.

For the m-th annulus, if $\ell_{m-1} + 1 < \lfloor \log(d(e)/2) \rfloor$, then each of the $m+1$ arms

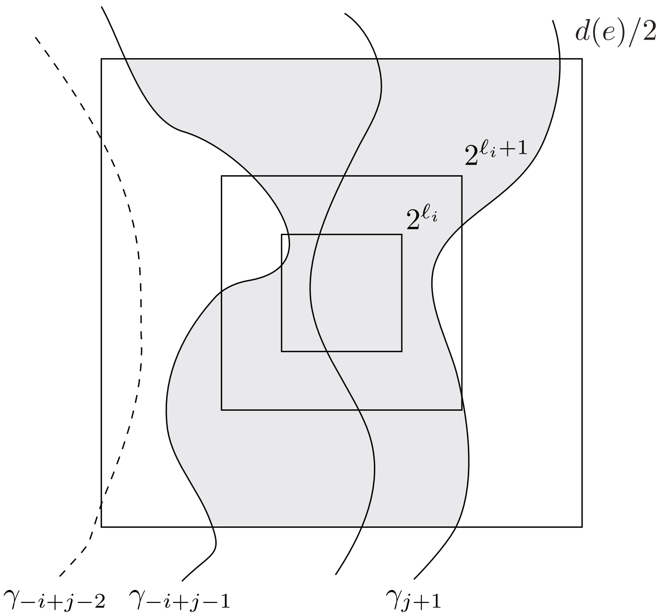


Figure 4.3: The geometric configuration in the proof of disjointness of Lemma 4.3.1. The shaded region represents R_2.

in C_{m-1} crosses $\mathrm{Ann}(e; 2^{1+\ell_{m-1}}, L)$ twice, proving item 2.

Finally, we address the disjointness of arms in each annulus. For the first annulus $\mathrm{Ann}(e; \ell, 2^{\ell_1})$, γ_{-1}, γ_1, $\mathfrak{c}_{-1}$, and $\mathfrak{c}_1$ are disjoint between $\partial B(e, \ell)$ and $\partial B(e, 2^{\ell_1})$ by definition. Inductively, for any non-empty annulus $\mathrm{Ann}(e; 2^{1+\ell_i}, 2^{\ell_{i+1}})$, $0 \leq i \leq m-2$, we have associated a collection C_i of $i+2$ arms to the scale ℓ_i. Suppose all of them are disjoint. The "outermost" arms, labeled γ_{-i+j-1} and γ_{j+1}, partition the box $B(e, 2^{\ell_{i+1}})$ into three regions: R_1, the region on the counterclockwise side of γ_{-i+j-1}; R_2, the region between γ_{-i+j-1} and γ_j through which the other i arms pass; and R_3, the region on the clockwise side of γ_{j+1}. See Figure 4.3 for an illustration. The collection C_{i+1} is formed by adding either γ_{-i+j-2} or γ_{j+2} to C_i. Neither passes through R_2 and both are necessarily disjoint from γ_{-i+j-1} and γ_{j+1}. Thus they must be disjoint from all $i + 2$ arms in C_i. It follows that all $i + 3$ arms in C_{i+1} are disjoint. □

In the previous proof, $\mathrm{Ann}(e; \ell, L)$ was divided into m annuli containing six arms in $\mathrm{Ann}(e; \ell, 2^{\ell_1})$ and two additional arms for every successive scale larger than ℓ_1. The scales ℓ_i, $i \leq r \leq m$ might coincide for some $r < m$ if we reach distance L before the m-th step of induction. In this case, the corresponding annuli $\mathrm{Ann}(e; 2^{1+\ell_i}, 2^{\ell_{i+1}})$ are empty.

Corollary 4.3.1.1. *For $0 \leq i \leq m-2$, the color sequence Σ_i in Lemma 4.3.1 contains at least two non-consecutive occurrences of open colors and resp. dual-closed colors.*

Proof. We recall that for $0 \leq i \leq m-2$, the $(6+2i)$-arm color sequence is of the form

$$\beta_{-1},\ \beta_{-2}, \ldots, \beta_{-i+j-1}, \overline{\beta_{-i+j-1}},\ \beta_{-i+j-1}, \ldots$$
$$\beta_{-2},\ \beta_{-1},\ \beta_1,\ \beta_2, \ldots,\ \beta_{j+1},\ \overline{\beta_{j+1}},\ \beta_{j+1}, \ldots,\ \beta_2,\ \beta_1.$$

If $\beta_{j+1} = \beta_{-i+j-1}$, then the two occurrences of β_{j+1} are non-consecutive and $\overline{\beta_{j+1}}$ and $\overline{\beta_{-i+j-1}}$ are two non-consecutive occurrences of the $\overline{\beta_{j+1}}$ color.

If $\beta_{j+1} = \overline{\beta_{-i+j-1}}$, then the two occurrences of β_{j+1} are non-consecutive and $\overline{\beta_{j+1}}$ and β_{-i+j-1} are two non-consecutive occurrences of the $\overline{\beta_{j+1}}$ color. ☐

A landing sequence $I = \{I_i\}_{1 \leq i \leq k}$ on $\partial B(n)$ is a sequence of disjoint subintervals of $\partial B(n)$ in clockwise order with $|I_i| \geq \delta n$ for some $\delta > 0$ and all i. A landing sequence I' on $\partial B(N)$ is defined analogously. We use a result in [Nol08], where it is proved that the probability of a k-arm event is comparable to the probability of the same event with extra landing conditions.

$$\mathbb{P}(A_{k,\sigma}(n, N)) \asymp \mathbb{P}(A_{k,\sigma}^{I/I'}(n, N)), \tag{4.3}$$

see [Nol08, Theorem 11].

Although the setting of Nolin's paper is site percolation on the triangular lattice, the proof applies to the square lattice, since the main techniques used in the proof are Russo-Seymour-Welsh estimates and generalized FKG inequality, both of which apply to the square lattice. Using (4.3), we may work with the events with prescribed landing zones for all arms.

The next lemma shows that enforcing a separation condition in the arm events does not essentially change the order of the probabilities.

Lemma 4.3.2. *Let integer $k \geq 2$, and σ be some color sequence. Let n be sufficently large and $N \geq 8n$, and specify two landing sequences I, I' on $\partial B(n)$ and $\partial B(N)$, and $\epsilon > 0$. Then:*

$$\mathbb{P}(\tilde{A}^{I,I'}_{k,\sigma}(n,N)) \geq (1-\epsilon)\mathbb{P}(A^{I/I'}_{k,\sigma}(n,N)).$$

Proof. We show the following:

$$\mathbb{P}(A^{\times,I,I'}_{k,\sigma}(n,N)) \leq \epsilon\mathbb{P}(A^{I,I'}_{k,\sigma}(n,N)),$$

where $A^{\times,I,I'}_{k,\sigma}$ denotes the event $A^{I,I'}_{k,\sigma} \setminus \tilde{A}^{I,I'}_{k,\sigma}$. That is, $A^{I,I'}_{k,\sigma}$ occurs but there are two arms that come closer than ℓ in $\mathrm{Ann}(2n, N/2)$.

Let the event $E(e)$ be as in Lemma 4.3.1, then

$$\mathbb{P}(A^{\times,I,I'}_{k,\sigma}(n,N)) \leq \sum_{e \in B(2n,N/2)} \mathbb{P}(A^{\times,I,I'}_{k,\sigma}(n,N), E(e)).$$

Using independence, Lemma 4.3.1, and carving out a sub-annulus to distance $d(e)/2$ around e, the terms on the RHS can be bounded as

$$\begin{aligned}
&\mathbb{P}(A^{\times,I,I'}_{k,\sigma}(n,N), E(e)) \\
&\quad \leq \mathbb{P}(A^{I}_{k,\sigma}(n, d(e)/2))\mathbb{P}(A^{I'}_{k,\sigma}(2d(e), N)) \\
&\quad \times \sum_{\ell_0 \leq \ell_1 \leq \cdots \leq \ell_{m-1}}^{L} \prod_{i=0}^{m-2} \mathbb{P}(A_{6+2i,\Sigma_i}(e; 2^{\ell_i+1}, 2^{\ell_{i+1}}))\mathbb{P}(A_{2m+4,\Sigma_{m-1}}(e; 2^{\ell_{m-1}+1}, L). \qquad (4.4)
\end{aligned}$$

Here, we have denoted $\ell_0 = \lceil \log \ell \rceil$ and $L = \lfloor \log(d(e)/2) \rfloor$.

By the generalized FKG inequality and quasi-multiplicativity [Nol08, Proposition 12], we have

$$\mathbb{P}(A^{I}_{k,\sigma}(n, d(e)/2))\mathbb{P}(A^{I'}_{k,\sigma}(2d(e), N)) \leq C\mathbb{P}(A^{I,I'}_{k,\sigma}(n, N))$$

where the constant C depends only on k and σ.

On the other hand, it is known the alternating five-arm probability has a universal exponent 2:

$$\mathbb{P}(A_{5,OC^*OOC^*}(n, N)) \leq C\left(\frac{n}{N}\right)^{-2}$$

for some constant $C > 0$. See [KSZ98, Lemma 5].

By Corollary 4.3.1.1, for any $0 \leq i \leq m-2$, the color sequence Σ_i can be split into an alternating five-color sequence ((O, C^*, O, O, C^*) or its flipped sequence) and a color sequence of length $2i+1$, while maintaining the relative order within each subsequence. We can now apply Reimer's inequality [Rei00] to obtain

$$\begin{aligned}\mathbb{P}(A_{6+2i,\Sigma_i}(e; 2^{\ell_i+1}, 2^{\ell_{i+1}})) &\leq \pi_5(2^{\ell_i+1}, 2^{\ell_{i+1}})\mathbb{P}(A_{2i+1}(e; 2^{\ell_i+1}, 2^{\ell_{i+1}})) \\ &\leq \pi_5(2^{\ell_i+1}, 2^{\ell_{i+1}})(\pi_1(2^{\ell_i+1}, 2^{\ell_{i+1}}))^{2i+1}.\end{aligned}$$

Here we denote $\pi_5(n, N) = \mathbb{P}(A_{5,OC^*OOC^*}(n, N))$ and $\pi_1(n, N) = \mathbb{P}(A_1(n, N))$. For π_1, we have the estimate

$$\pi_1(n, N) \leq \left(\frac{n}{N}\right)^{\delta}, \quad \text{for some } \delta > 0.$$

Similarly, for the final annulus, we have

$$\mathbb{P}(A_{2m+4,\Sigma_{m-1}}(e; 2^{\ell_{m-1}+1}, L)) \leq \left(\frac{2^{\ell_{m-1}}}{L}\right)^{2+(2m-1)\delta}.$$

Plugging the above estimates in the product in line (4.4), we have

$$\begin{aligned}\text{product in (4.4)} &\leq \prod_{i=0}^{m-2} (2^{\ell_i - \ell_{i+1}})^{2+(2i+1)\delta} \left(\frac{2^{\ell_{m-1}}}{L}\right)^{2+(2m-1)\delta} \\ &= \left(\frac{\ell}{L}\right)^{2+\delta} \prod_{i=0}^{m-2} (2^{2\delta})^{\ell_i} (2^{2\delta})^{-(m-2)\ell_{m-1}} \left(\frac{2^{\ell_{m-1}}}{L}\right)^{2(m-1)\delta}.\end{aligned}$$

Since we have the estimate $\sum_{i \leq N} x^i \leq c x^N$ for any $x = 2^y$, $y > 0$, and some constant $c = c(y)$, we have

$$\sum_{\ell_0 \leq \ell_i \leq \ell_{i+1}} (2^{2i\delta})^{\ell_i} \leq c(2^{2i\delta})^{\ell_{i+1}}.$$

Summing over all possible values of $\ell_1, \ldots, \ell_{m-1}$, we have

$$\begin{aligned}(4.4) &\leq \left(\frac{\ell}{L}\right)^{2+\delta} c^{m-2} \sum_{\ell_{m-1}=\ell_0}^{\lfloor \log(d(e)/2) \rfloor} (2^{2\delta})^{(m-2)\ell_{m-1}} (2^{2\delta})^{-(m-2)\ell_{m-1}} \left(\frac{2^{\ell_{m-1}}}{L}\right)^{2(m-1)\delta} \\ &\leq \left(\frac{\ell}{L}\right)^{2+\delta} c^{m-2} (2^{\lfloor \log(d(e)/2) \rfloor})^{-2(m-1)\delta} c (2^{2(m-1)\delta})^{\lfloor \log(d(e)/2) \rfloor} \\ &\leq C \left(\frac{\ell}{L}\right)^{2+\delta}.\end{aligned}$$

Summing over the location of e, we have

$$\begin{aligned}\mathbb{P}(A_{k,\sigma}^{\times,I,I'}(n,N)) &\leq C\mathbb{P}(A_{k,\sigma}^{I,I'}(n,N)) \sum_{e \in B(2n,N/2)} (d(e)/2)^{-2-\delta} \\ &\leq C\mathbb{P}(A_{k,\sigma}^{I,I'}(n,N)) \sum_{k=1}^{\lceil \log(N/4n) \rceil} \left(n 2^{k-1}\right)^{-2-\delta} n^2 2^{2k} \\ &\leq C' N^{-\delta} \mathbb{P}(A_{k,\sigma}^{I,I'}(n,N)),\end{aligned}$$

where the factor $n^2 2^{2k}$ estimates the number of edges in the k-th annulus of the sum. Choosing N such that $C' N^{-\delta} \leq \epsilon$, we obtain the stated result. □

Proof of Proposition 4.1.1. By Lemma 4.3.2 and (4.3), for any $\epsilon > 0$ and some choice of landing sequences I, I' which we will specify later on, there exists a $C > 0$ such that

$$\mathbb{P}(A_{k,\sigma}(n,N)) \leq C\mathbb{P}(A_{k,\sigma}^{I,I'}(n,N)) \leq \frac{C}{1-\epsilon} \mathbb{P}(\tilde{A}_{k,\sigma}^{I,I'}(n,N)),$$

where $\tilde{A}_{k,\sigma}$ is the ℓ-separated k-arm event, see Definition 4.4. It suffices to bound the probability on the right in the last display.

It suffices to consider the case when σ and σ' differ by one entry. For general polychromatic color sequences σ and σ', we consider a sequence interpolating between σ and σ' with at most k steps, such that any two consecutive color sequences differ in a single entry.

Without loss of generality, we assume that σ and σ' differ only in the k-th entry and moreover we assume that $\sigma_1 = O$ and $\sigma_2 = C^*$. Fix two consecutive landing zones on $\partial B(N)$ for an open and a dual-closed arm, say I_1, I_2, corresponding to the first two entries. Let γ_1, γ_2 be the pair of open and dual-closed arms closest to each other on γ_1's clockwise side, such that γ_1 lands on I_1 and γ_2 lands on I_2. Finally, we inductively let $\gamma_3, \ldots, \gamma_{k-1}$ be such that γ_i is the arm with color σ_i landing on I_i such that the region enclosed by γ_{i-1} and γ_i on γ_{i-1}'s clockwise side is minimal and the γ_j's are all disjoint.

We then denote by U^c the region enclosed by γ_1, γ_{k-1}, $\partial B(n)$, and $\partial B(N)$ that excludes γ_k. By minimality, the event $\{U^c = R\}$ depends only on the status of edges in R. In particular, the configuration in the complement region $U = \mathrm{Ann}(n, N) \setminus U^c$ is independent of U^c. Moreover, there is an arm γ_k with color σ_k in U such that γ_k is at distance at least ℓ from γ_1 and γ_{k-1}, two parts of the boundary of U.

$$\begin{aligned}\mathbb{P}(\tilde{A}^{I_1,\ldots,I_k}_{k,O,C^*,\sigma_3,\ldots,\sigma_k}(n,N)) &\le \sum_{\text{admissible } S} \mathbb{P}(\tilde{A}^{I_k}_{1,\sigma_k}(U), U = S) \\ &= \sum_{\text{admissible } S} \mathbb{P}(U = S)\mathbb{P}(\tilde{A}^{I_k}_{1,\sigma_k}(S))\end{aligned}$$

The last two sums are over the possible values of S of the region U.

With fixed $\gamma_1, \ldots, \gamma_{k-1}$, we flip the percolation configuration in the region S.

We have

$$\mathbb{P}(\tilde{A}^{I_k}_{1,\sigma_k}(S)) = \mathbb{P}(\tilde{A}^{I_k}_{1,\overline{\sigma}_k}(S)). \tag{4.5}$$

Recall that $\overline{\sigma}_k$ denotes the flipped color sequence.

Having flipped the configuration in S, we use the transformation T defined in the proof of Lemma 4.2.1 to shift the configuration in the region S to the dual lattice. By Lemma 4.2.1, we have

$$\mathbb{P}(\tilde{A}_{1,\overline{\sigma}_k}(S)) \leq \mathbb{P}(A_{1,\overline{\sigma}^*_k}(S \cap \mathrm{Ann}(2n, N/2))).$$

Inserting this inequality into (4.5) we have

$$\begin{aligned}
\mathbb{P}(A^{I_1,\ldots,I_k}_{k,O,C^*,\sigma_3,\ldots,\sigma_k}(n,N)) &\leq \sum_{\text{admissible } S} \mathbb{P}(U = S)\mathbb{P}(A_{1,\overline{\sigma}^*_k}(S \cap \mathrm{Ann}(2n, N/2))) \\
&\leq C\mathbb{P}(A_{k,O,C^*,\sigma_3,\ldots,\sigma_{k-1},\overline{\sigma}^*_k}(2n, N/2)) \\
&\leq C\mathbb{P}(A_{k,O,C^*,\sigma_3,\ldots,\sigma_{k-1},\overline{\sigma}^*_k}(n, N)).
\end{aligned}$$

The final inequality follows from a standard argument using RSW estimates. See [Nol08, Proposition 16]. □

Part II

Ballistic Annihilation

CHAPTER 5

INTRODUCTION

5.1 Background

Annihilating particle systems were introduced in the 1980s by Toussaint and Nobel physics laureate Wilczek to study the kinetics of chemical reactions [TW83]. Most of the early effort was in studying diffusive motions, where particles follow random walk or Brownian motion trajectories, see [BL91] on two-type annihilation $A+B \to \varnothing$ on $\mathbb{Z}^d$ and [Arr83] on one-type annihilation $A+A \to \varnothing$ (also known as annihilating random walk) on $\mathbb{Z}^d$. This class of systems falls under the name of diffusion-limited annihilating systems (DLAS), which remains a subject of active research today. Interest in annihilating dynamics with *ballistic* particle trajectories, i.e. where particles have constant velocities, arose as an extremal case of DLAS.

In **ballistic annihilation** (BA), infinitely many particles with velocities sampled from a probability measure ν move across the real line and mutually annihilate upon contact. The canonical example, the *symmetric three-velocity ballistic annihilation*, samples velocities independently from the probability measure

$$\nu = \frac{1-p}{2}\delta_{-1} + p\delta_0 + \frac{1-p}{2}\delta_1. \tag{5.1}$$

For convenience, we call velocity-0 particles blockades, and velocity-±1 particles arrows.

The main statistic of interest is $\theta(p)$, the probability that a prespecified blockade is never annihilated, see (5.3) for the formal definition. By ergodicity, the

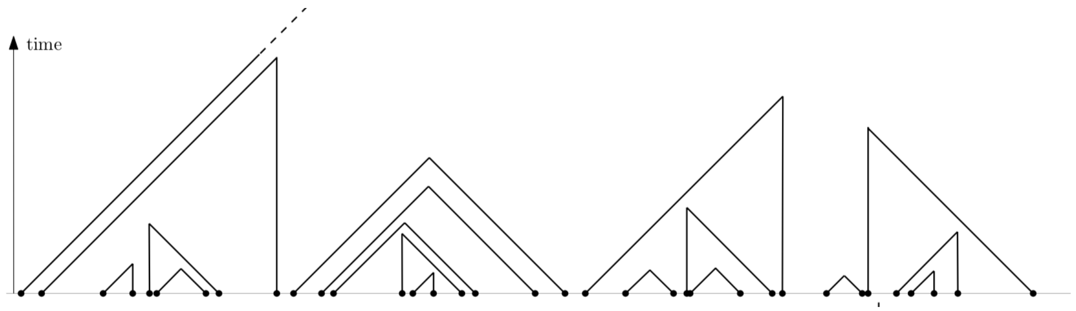


Figure 5.1: An example of BA on a space-time diagram. Figure courtesy of [HST21].

limiting proportion of surviving blockades converges to $p\theta(p)$. Thus,

$$p_c = \inf\{p \colon \theta(p) > 0\} \tag{5.2}$$

represents the critical initial blockade density for species survival.

Droz, Rey, Frachebourg, and Piasecki [DRFP95] and, later in more detail, Krapivsky, Redner and Leyvraz [KRL95] inferred in the 1990s that there is a critical density for blockade survival and $p_c = 1/4$. A rough intuition for why (given in [DRFP95]) is that arrow–arrow collisions should occur, on average, at twice the rate of arrow–blockade collisions. Under this assumption, on average, six arrows are removed for every two blockades. This suggests that the critical density starts with a 1:3 ratio of blockades to arrows, hence p_c should be 1/4. This heuristic assumes uniformity of particle types and ignore the most challenging feature of BA—that it exhibits long-range dependence. BA is also sensitive to perturbation. Changing an arrow into a blockade may increase the lifespans of other arrows. Thus, it is not obvious how to establish monotonicity for $\theta(p)$, among other related quantities, in p without quantitatively computing them.

The revival of interest in BA from the mathematics community is linked to a puzzle called the "bullet problem," which first circulated through the IBM research challenge [KW]. In this problem, a bullet of a random velocity uniformly sampled from $[0, 1]$ is fired every second from the origin and annihilated upon collision with other bullets. It is conjectured that there exists a critical velocity s_c such that the first bullet survives with positive probability if and only if its velocity is greater than s_c. This problem proved incredibly difficult and there is little rigorous literature so far besides [DKJ$^+$19] and [BM20]. In particular, Dygert et al. [DKJ$^+$19] considered the bullet problem with discrete speed distributions and showed that the first bullet survives with positive probability if it has the second-fastest speed. This result yields a lower bound for the survival probability of blockades in BA through a coupling that interchanges time and space, see [DKJ$^+$19, ST17] for elaborations on the coupling.

Despite some progress towards upper bounds on p_c [DKJ$^+$19, ST17, BGJ19], showing that $p_c = 1/4$ remained an open problem. Even proving the much weaker statement that $p_c > 0$ was a problem widely advertised by Sidoravicius in the mid-2010s. In a breakthrough paper first posted in 2018, Haslegrave, Sidoravicius, and Tournier [HST21] introduced an exactly solvable combinatorial approach that confirmed $p_c = 1/4$. The authors also established finer details such as the asymptotics of the density of particles by type and the “skyline process” of collision types, many of which exhibit universal properties with respect to the distribution of initial particle spacings.

The proof in [HST21] relies heavily on the symmetry of the velocity measure ν in (5.1). Indeed, one of their breakthrough contributions is a measure-preserving reversal argument, which fails for any asymmetrical ν. In [JL22], Junge and Lyu considered an asymmetrical system with velocity distribution

$$\nu(p, \lambda, w) = (1 - \lambda)(1 - p)\delta_{-w} + p\delta_0 + \lambda(1 - p)\delta_1,$$

with $p, \lambda \in [0,1]$ and $w > 0$. Despite being unable to locate the phase transition, Junge and Lyu derived non-trivial bounds for the thresholds in p and λ. Their main contribution in [JL22] is introducing the more robust *mass transport principle* (MTP), which only relies on translation invariance, as a replacement of [HST21]'s reversal argument. This circumvents the lack of symmetry in the generalized system. Note that Haslegrave and Tournier [HT21] also found an application of the mass transport principle for computing a moment generating function associated to BA.

5.2 Overview of Part II

In Chapter 6, we introduce a coalescing variant of the three-velocity ballistic annihilation in which collisions may generate new particles. For a symmetric three-parameter family of such systems, we compute the survival probability of stationary particles at a given initial density. This allows us to describe a phase transition for stationary particle survival.

In Chapter 7, we continue to consider coalescing ballistic annihilation processes. For a family of symmetric coalescing rules, we prove that the law of the index of the first particle to arrive at the origin does not depend on the law for spacings between particles.

In Chapter 8, we introduce a variant of ballistic annihilation with superimposed clusters of stationary particles. We provide a simple formula for the critical initial density in terms of the mean and variance of the cluster size.

5.3 Notation

We let $(x_k)_{k\in\mathbb{Z}}$ be an ordered sequence of starting locations for particles. To standardize placements, set $x_0 = 0$ and assume that $x_k - x_{k-1}$ are sampled independently according to a continuous distribution with support contained in $(0, \infty)$. At each x_k, we place a particle $\bullet_k$ where the particle type is independently sampled from ν in (5.1). We denote the different types of particles by

$$\dot{\bullet}_k = \{\bullet_k \text{ is a blockade}\}, \vec{\bullet}_k = \{\bullet_k \text{ is a right arrow}\}, \text{ and } \overleftarrow{\bullet}_k = \{\bullet_k \text{ is a left arrow}\}.$$

Blockades are stationary while left and right arrows move with velocities -1 and 1, respectively. Collision events and visits to a location $u \in \mathbb{R}$ are specified by

$$\begin{aligned}
\vec{\bullet}_j \longleftrightarrow \overleftarrow{\bullet}_k &= \{\vec{\bullet}_j \text{ and } \overleftarrow{\bullet}_k \text{ mutually annihilate}\}\\
\dot{\bullet}_j \leftarrow \overleftarrow{\bullet}_k &= \{\overleftarrow{\bullet}_k \text{ mutually annihilates with a blockade at } x_j\}\\
\bullet_j \;\text{---}\; \bullet &= \cup_{k>j}\{\bullet_j \;\text{---}\; \bullet_k\}\\
\bullet \;\text{---}\; \bullet_k &= \cup_{j<k}\{\bullet_j \;\text{---}\; \bullet_k\}\\
u \leftarrow \overleftarrow{\bullet} &= \{u \text{ is visited by a particle from the right}\}.
\end{aligned}$$

We use --- to denote any collision type. The complements of collision events are denoted by $\not\longleftrightarrow$, $\not\leftarrow$, and $\not\rightarrow$. The events $\vec{\bullet}_j \rightarrow \dot{\bullet}_k$ and $\vec{\bullet} \rightarrow u$ are defined similarly. Note that we count an arrow destroying a blockade as also visiting the site housing the blockade, e.g. $\{\dot{\bullet}_k \leftarrow \bullet\} \subseteq \{x_k \leftarrow \bullet\}$.

It is often advantageous to restrict to the system with only the particles started in a specified interval $I \subseteq \mathbb{R}$. We notate this restriction by including I as a subscript on the event. For example, $(\dot{\bullet}_j \longleftrightarrow \overleftarrow{\bullet}_k)_{[x_j,x_k]}$ is the event that $\bullet_j$ is a blockade that mutually annihilates with a left arrow started at x_k when restricted to only the particles in $[x_j, x_k]$.

A fundamental statistic associated to BA is the probability $\theta(p)$ that the origin is not visited conditional on $\dot{\bullet}_0$. Let

$$q = q(p) := \mathbb{P}_p((0 \leftarrow \overleftarrow{\bullet})_{(0,\infty)}).$$

Then formally, we define

$$\theta(p) := \mathbb{P}_p((\vec{\bullet} \not\rightarrow 0)_{(-\infty,0)} \wedge (0 \not\leftarrow \overleftarrow{\bullet})_{(0,\infty)}). \tag{5.3}$$

By independence and symmetry, $\theta(p) = (1 - q)^2$. The Birkhoff Ergodic Theorem ensures that the limiting density of surviving blockades is $p\theta(p)$.

Following conventions, we hereafter consider the one-sided process on $(0, \infty)$ unless stated otherwise. Accordingly, we drop the subscript $(0, \infty)$ from our event notation.

5.4 HST's argument and the mass transport principle

The basic idea of Haslegrave–Sidoravicius–Tournier's argument is to partition the event associated to q based on the velocity of $\bullet_1$. Recall that all events are restricted to the one-sided process on $(0, \infty)$:

$$q = \mathbb{P}((0 \leftarrow \bullet) \wedge \overleftarrow{\bullet}_1) + \mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1) + \mathbb{P}((0 \leftarrow \bullet) \wedge \vec{\bullet}_1).$$

For the first term, observe that $(0 \leftarrow \bullet) \wedge \overleftarrow{\bullet}_1 = \overleftarrow{\bullet}_1$. Then the probability is $\frac{1-p}{2}$. For the second term, conditional on $\dot{\bullet}_1$, the origin is visited if and only if $\dot{\bullet}_1$ annihilates with a left arrow and another left arrow visits the starting location of said left arrow. By independence and translation invariance, the probability in the second term is pq^2. The third term is the most tricky. A key probability associated to this term is

$$s := \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \rightarrow \dot{\bullet})).$$

In [HST21], it was observed that $s = \frac{1}{2}pq^2$. This comes from the insightful observation through a measure-preserving reversal argument that the event measured by s has the same probability as the event in which two left particles arrive at 0 with the time of the first arrival strictly smaller then the time between the first and second arrivals, see Figure 5.2. Two arrivals occur with probability q^2 and, even with no explicit knowledge of the arrival time distribution, symmetry and independence give that the first arrival takes less time with probability 1/2.

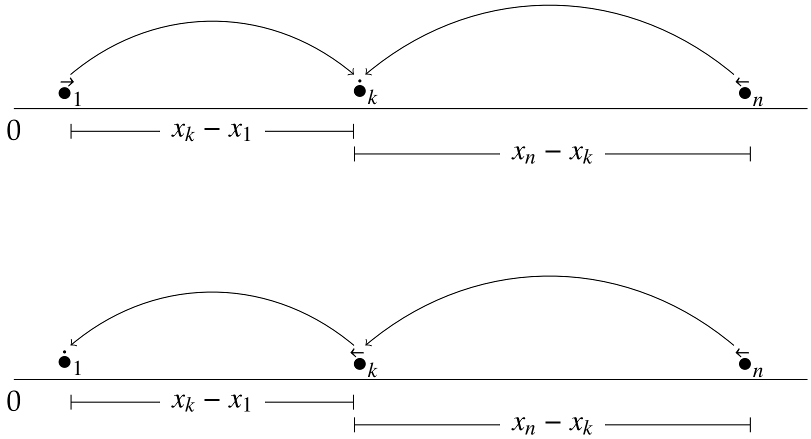


Figure 5.2: A configuration in $\{(0 \leftarrow \overleftarrow{\bullet}_n) \wedge (\vec{\bullet}_1 \rightarrow \dot{\bullet}_k)\}$ (top) and its reversal on $[x_1, x_k]$ (bottom). Particles between $\bullet_1$ and $\bullet_k$ and $\bullet_k$ and $\bullet_n$ are not shown. The arced arrows indicate that the particle is the first to reach that site among all particles started under the arc.

The recursive equation in q gives the following proposition.

Proposition (Haslegrave–Sidoravicius–Tournier 2021)**.**

$$0 = 1 - q - p - pq + pq^2 + pq^3 = (1 - q)(1 - p(1 + q)^2).$$

This implies that either $q = 1$ or $q = p^{-1/2} - 1$. Since q is a probability, $q = 1$ when $p \leq 1/4$. This gives the first nontrivial lower bound for p_c.

However, it is not a priori that q takes root 1 when $p \in (0, 1/4)$ and root $p^{-1/2} - 1$ when $p > 1/4$. HST showed that both

$$\{p \in (1/4, 1) : \theta(p) = 0\} \text{ and}$$
$$\{p \in (1/4, 1) : \theta(p) > 0\}$$

are open intervals. Thus, the first set must be both open and closed in the subspace topology on $(1/4, 1)$ and, combined with other reasons, must be empty. This gives the upper bound $p_c \leq 1/4$ and settles the problem.

Before ending this chapter, we state an important tool for obtaining recursive formulas for q in the coming chapters, which depends only on translation invariance.

Proposition 5.4.1 (Mass transport principle)**.** *Consider a family of non-negative random variables Z(m,n) for integers $m, n \in \mathbb{Z}$ such that its distribution is diagonally invariant under translation, i.e., for any integer ℓ, $Z(m + \ell, n + \ell)$ has the same distribution as $Z(m, n)$. Then for each $m \in \mathbb{Z}$:*

$$\mathbb{E} \sum_{n \in \mathbb{Z}} Z(m, n) = \mathbb{E} \sum_{n \in \mathbb{Z}} Z(n, m). \tag{MTP}$$

Proof. Using Fubini's theorem and translation invariance of $\mathbb{E}[Z(m, n)]$, we obtain

$$\mathbb{E} \sum_{n \in \mathbb{Z}} Z(m, n) = \sum_{n \in \mathbb{Z}} \mathbb{E}[Z(m, n)] = \sum_{n \in \mathbb{Z}} \mathbb{E}[Z(2m - n, m)] = \sum_{n \in \mathbb{Z}} \mathbb{E}[Z(n, m)] = \mathbb{E} \sum_{n \in \mathbb{Z}} Z(n, m).$$

□

CHAPTER 6

THREE-VELOCITY COALESCING BALLISTIC ANNIHILATION

6.1 Introduction

We consider ballistic motion in which collisions sometimes result in the generation of new particles. This is inspired by earlier work from physicists [CPY90, TEW98, BEK00]. However, none of these works considered the three-velocity setting. We remark that [HST21] allowed for a coalescence rule in which a particle is selected uniformly at random to survive a triple collision. The primary reason for considering this case was to resolve technical difficulties that arise in the presence of triple collisions, rather than investigate coalescence dynamics. Note that when the spacings between particles are sampled from an atomless distribution, there are almost surely no triple collisions.

We require more notation to describe coalescing systems. The initial conditions with particles $\bullet_k$ at x_k for $k \in \mathbb{Z}$ assigned velocities from $\{-1, 0, 1\}$ remain unchanged. We denote two particles meeting at the same location by $\bullet_j — \bullet_k$. Upon meeting, a reaction takes place. Either the particles coalesce and form a new particle with an independently sampled new velocity, or the particles mutually annihilate.

In general, there are three types of collisions: $\vec{\bullet} — \dot{\bullet}$, $\dot{\bullet} — \overleftarrow{\bullet}$, and $\vec{\bullet} — \overleftarrow{\bullet}$. Each collision may result in one of four reactions: generating a left arrow, right arrow, blockade, or mutual annihilation (denoted by $\varnothing$).

We define the *three-parameter coalescing ballistic annihilation* (TCBA) covered by our results. Fix parameters $0 \le a, b, \alpha < 1$ with $a + b \le 1$. Using the notation

$$\bullet \text{—} \bullet \implies \Theta, \quad x$$

to denote a collision resulting in an outcome $\Theta \in \{\dot{\bullet}, \vec{\bullet}, \overleftarrow{\bullet}, \varnothing\}$ with probability x, we have the following collision rules:

$$\vec{\bullet} \text{—} \overleftarrow{\bullet} \implies \begin{cases} \overleftarrow{\bullet}, & a/2 \\ \vec{\bullet}, & a/2 \\ \dot{\bullet}, & b \\ \varnothing, & 1-(a+b) \end{cases} \qquad \begin{array}{l} \dot{\bullet} \text{—} \overleftarrow{\bullet} \implies \begin{cases} \overleftarrow{\bullet}, & \alpha \\ \varnothing, & 1-\alpha \end{cases} \\ \vec{\bullet} \text{—} \dot{\bullet} \implies \begin{cases} \vec{\bullet}, & \alpha \\ \varnothing, & 1-\alpha \end{cases} \end{array} . \tag{6.1}$$

So TCBA allows for arrows to survive collisions with blockades and other arrows or to generate a blockade after colliding with an arrow. Note that BA is the special case $a = b = \alpha = 0$.

For all but the $\vec{\bullet} \text{—} \overleftarrow{\bullet} \implies \dot{\bullet}$ reaction, it is mathematically equivalent to view coalescence as one of the particles surviving the collision. Taking this perspective, we have the head of the arrow point to the particle that is destroyed. For example, $\vec{\bullet}_j \to \overleftarrow{\bullet}_k$ denotes the event that the left arrow started at x_k is destroyed by the right arrow started at x_j. The survived particle is still denoted by $\vec{\bullet}_j$. If mutual annihilation occurs, then we continue to write $\bullet \longleftrightarrow \bullet$. We denote the case in which two arrows collide and generate a blockade by $\vec{\bullet}_m \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_n$. We denote blockades generated from such collisions by $\hat{\bullet}_{m,n}$, and denote a generic blockade generated from such a reaction by $\hat{\bullet}$.

Recall that $(0 \leftarrow \bullet)$ is implicitly restricted to the positive real line. For TCBA, we define $q = q(a, b, \alpha, p) := \mathbb{P}(0 \leftarrow \bullet)$ and θ as in (5.3):

$$\theta = \theta(a, b, \alpha, p) := \mathbb{P}_p((\vec{\bullet} \nrightarrow 0)_{(-\infty,0)} \wedge (0 \nleftarrow \overleftarrow{\bullet})_{(0,\infty)}) = (1-q)^2.$$

We define $p_c = p_c(a,b,\alpha)$ as the value of $p_c^- = p_c^-(a,b,\alpha) := \inf\{p\colon q(a,b,\alpha,p) < 1\}$ and $p_c^+ = p_c^+(a,b,\alpha) := \sup\{p\colon q(a,b,\alpha,p) = 1\}$ when they coincide. Our main result gives formulas for q and p_c.

Theorem 6.1.1. *For any TCBA it holds that*

$$p_c = p_c(a,b,\alpha) = \frac{1 - b(1-\alpha)}{4 - 3\alpha - (a+b)(1-\alpha)} \tag{6.2}$$

with $q(p) = 1$ *for* $p \le p_c$ *and*

$$q(p) = \frac{\sqrt{(1-\alpha)\left(b(1-p)^2 - p(a(1-p) + p\alpha - 1)\right)} - p(1-\alpha)}{(1-\alpha)((1-a)p + b(1-p))} \tag{6.3}$$

for $p > p_c$.

This generalizes [HST21, Theorem 1] in which the formula $q(p) = p^{-1/2} - 1$ for $p \ge 1/4$ and otherwise $q(p) = 1$ is established. Given the notorious sensitivity of BA to perturbation, it is noteworthy that we can describe systems whose local behavior is markedly different from BA. The form of (6.2) illustrates how the location of the phase transition depends in a subtle way on the coalescence rules. This suggests that it would be difficult to infer p_c from heuristic arguments such as those given in the introductions of [DRFP95, BGJ19] for BA. Note that there are parameter choices that result in arbitrarily small and large values of p_c. See Figure 6.1 for a depiction of the function q for various parameter choices.

Although Theorem 6.1.1 extends the main result from [HST21] to ballistic systems with coalescence, it is still restricted to systems with reflection symmetry. For asymmetric three-velocity ballistic annihilation without coalescence—for example, left and right particles have different speeds or probabilities of occurring—universal bounds for p_c were obtained in [JL22]. However, it seems to be difficult to establish a sharp phase transition and/or formulas for p_c and

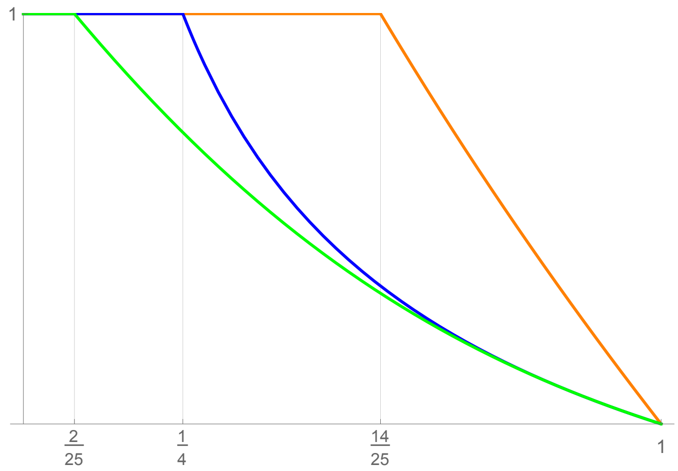


Figure 6.1: Plots of the formula for $q(p)$ from Theorem 6.1.1 for different parameter choices (a, b, α). The horizontal axis is p, and the vertical axis is the probability q. The green curve is for $(1/8, 3/4, 0)$, which gives $p_c = 2/25$. The blue curve is for $(0, 0, 0)$, so the usual BA, which has $p_c = 1/4$. The orange curve is $(1/4, 1/2, 3/4)$, which gives $p_c = 14/25$.

$q(p)$ in asymmetric cases. The lack of reflection symmetry introduces a number of additional quantities to be solved explicitly. Another interesting extension is to allow blockades to survive collisions (e.g., $\dot{\bullet} \,—\, \vec{\bullet} \implies \dot{\bullet}$) or particles moving in the opposite direction of the reactant to be generated (e.g., $\dot{\bullet} \,—\, \vec{\bullet} \implies \overleftarrow{\bullet}$). Such cases do not appear to offer the same sort of renewal as the cases we cover in Theorem 6.1.1. This is discussed more in Remark 6.2. In Chapter 8, a solution to the process in which blockades survive multiple collisions was found, see Corollary 8.1.1.3.

6.1.1 Proof methods

The high-level idea for establishing the phase transition is that first we derive an equation involving only q, a, b, α, and p. Then, we analyze the roots (in p) of this equation to find p_c and show that there is indeed a phase transition.

Our first step is deriving the identity

$$0 = (1 - q)g(p, q) \tag{6.4}$$

in Proposition 6.2.4 for an explicit function g. To obtain the identity we partition q in terms of the velocity assigned to $\bullet_1$. Recall that all events below are restricted to the one-sided process on $(0, \infty)$:

$$q = \mathbb{P}((0 \leftarrow \bullet) \wedge \overleftarrow{\bullet}_1) + \mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1) + \mathbb{P}((0 \leftarrow \bullet) \wedge \vec{\bullet}_1). \tag{6.5}$$

In Lemma 6.2.2, each of these three terms is expanded into a formula involving the parameters. These formulas are derived by further partitioning on the type of particle that destroys $\bullet_1$ and then observing some form of renewal. For example, the interdistances and types of particles started in (x_j, ∞) are conditionally independent of the event $(\dot{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}_j)$. Thus, the probability the origin is visited at least once is again equal to q.

While in the same spirit as the proof of [HST21, Proposition 5], where the identity $q = 1 - p(1 + q - q^2 - q^3)$ is derived, the coalescing case is more complicated. An important tool is the mass transport principle (MTP) derived from translation invariance. The use of the mass transport principle is inspired by what was done in [JL22].

The identity at (6.4) implies that q is either equal to 1 or a root of g. Let p_* be the claimed value of p_c in Theorem 6.1.1. It is straightforward to prove that the roots of g fall outside of $[0, 1]$ for $p \in [0, p_*)$, and so $q = 1$ on that interval. Moreover, at $p = p_*$ the only positive root of g is equal to 1. So we additionally have $q(p_*) = 1$. Recall that there is no known direct proof that q is continuous. On the interval $p \in (p_*, 1)$, we cannot easily rule out the possibility that q jumps between 1 and a root of g. To rule out this pathology, we generalize an approach from [HST21] and prove that the set $I = \{p \in (p_*, 1) \colon q(p) < 1\}$ is both open and closed in the subspace topology on $(p_*, 1)$. These observations along with the fact that I is nonempty (Lemma 6.4.6) imply that $I = (p_*, 1)$ and thus q must be equal to the unique positive root of g on this interval.

The proof that I^c is open follows what was done in [HST21]; we approximate q from below with continuous functions $q_k = \mathbb{P}((0 \leftarrow \bullet)_{(0,x_k]})$ to write I^c as a union of sets which are open because of an explicit comparison to the nonnegative root of g. The proof that I is open relies on a necessary and sufficient condition for blockades to survive with positive probability that is similar to what was done in [HST21]. Coalescence again introduces new challenges. The result [HST21, Lemma 12] is devoted to showing that the net number of surviving blockades versus left arrows is superadditive when combining processes on adjacent intervals. This no longer holds with TCBA since arrows can destroy multiple blockades. We introduce a weighting scheme to account for this in Lemma 6.4.3. The argument concludes by showing that $q(p) < 1$ if and only if there are on average more surviving blockades than arrows in our weighting scheme. This condition is different and a little more efficient than what was used for the analogue in [HST21]. See Remark 6.3. We use this condition to prove that I is open in Lemma 6.4.5 and that I is nonempty in Lemma 6.4.6.

6.1.2 Determining firstness

Since arrows may survive multiple collisions, we give more detail concerning how reactions are decided. Arrow-arrow collisions are decided in some generic way that is compatible with (6.1) and consistent when combining the process restricted to different intervals. For example, left arrows carry a queue of instructions for what reaction occurs for each right arrow they meet. An upcoming lemma (Lemma 6.4.3) requires particular care with how arrow-blockade reactions are decided. To facilitate the presentation of the proof of Lemma 6.4.3, we associate to each arrow $\bullet_i$ a *quiver* of σ_i *sharp arrows*. Each sharp arrow represents one blockade that $\bullet_i$ is able to destroy. Accordingly, $\sigma_i \sim \text{Geometric}_1(1-\alpha)$ is distributed as a geometric random variable supported on $1, 2, \dots$ with success parameter $1 - \alpha$.

Each arrow in the quiver has the same direction and position as $\bullet_i$. If $\bullet_i$ is destroyed by another arrow, then all arrows in the quiver are destroyed as well. Whenever $\bullet_i$ meets a blockade, one of the sharp arrows from its quiver mutually annihilates with the blockade. When the last sharp arrow in the quiver of $\bullet_i$ is destroyed, so is $\bullet_i$. It is straightforward to verify that the quiver formulation for deciding collisions is equivalent to TCBA according to (6.1).

Lastly, it will be necessary to distinguish whether the first arrow that visits a location will destroy and survive, or mutually annihilate with, the next blockade it meets. Accordingly, for $u \in \mathbb{R}$, we introduce the events

$$\{u \overset{1}{\Leftarrow} \bullet\} := \left\{ \begin{array}{c} u \text{ is first visited by an arrow whose quiver} \\ \text{contains at least two sharp arrows} \end{array} \right\}$$

$$\{u \overset{1}{\longleftrightarrow} \bullet\} := \left\{ \begin{array}{c} u \text{ is first visited by an arrow whose quiver} \\ \text{contains exactly one sharp arrow} \end{array} \right\}.$$

6.1.3 Organization

In Section 6.2, we use the mass transport principle (MTP) to derive a recursive formula for q which leads to the identity at (6.4) in Proposition 6.2.4. Section 6.3 contains additional information about g and its roots. Section 6.4 proves regularity of the set $I = \{p \in (p_*, 1) \colon q(p) < 1\}$. Finally, Section 6.5 combines the other sections to give a short proof of Theorem 6.1.1.

6.2 Recursion

We begin by stating and proving the main tool beside (MTP) for obtaining a recursive expression for q.

Proposition 6.2.1. *Let $c = 1 - (a + b)$. The following equations hold so long as the parameters in the denominators are nonzero:*

$$\begin{aligned}\frac{1}{a/2}\mathbb{P}((\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet})_{(0,\infty)}) &= \frac{1}{b}\mathbb{P}((\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet})_{(0,\infty)}) = \frac{1}{c}\mathbb{P}((\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})_{(0,\infty)}), \\ \mathbb{P}((u \leftarrow \bullet)_{(u,\infty)}) &= \frac{1}{\alpha}\mathbb{P}((u \overset{1}{\Leftarrow} \bullet)_{(u,\infty)}) = \frac{1}{1-\alpha}\mathbb{P}((u \overset{1}{\longleftrightarrow} \bullet)_{(u,\infty)}). \qquad (6.6)\end{aligned}$$

Proof. These are simple derivations from the probabilities of the various outcomes that occur when two particles meet. □

We define a few auxiliary probabilities that will be useful when expanding (6.5):

$$s := P((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k \text{ for some } k > 1)) + \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet}))$$

$$r := P((0 \not\leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k \text{ for some } k > 1)) + \mathbb{P}((0 \not\leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})).$$

Lemma 6.2.2. *Suppose that* $c = 1 - (a + b) > 0$*. For TCBA restricted to* $(0, \infty)$ *it holds that*

$$\mathbb{P}((0 \leftarrow \bullet) \wedge \overleftarrow{\bullet}_1) = \frac{1-p}{2} \tag{6.7}$$

$$\mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1) = \alpha pq + (1-\alpha)pq^2 \tag{6.8}$$

$$\mathbb{P}((0 \leftarrow \bullet) \wedge \vec{\bullet}_1) = \left(q + \frac{a/2}{c} + \frac{b}{c}\frac{\mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1)}{p}\right)\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) + s \tag{6.9}$$

$$\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) = \frac{\frac{1-p}{2} - s - r}{1 + \frac{a/2}{c} + \frac{b}{c}} = \frac{c\left(pq^2(1-\alpha) - 2pq(1-\alpha) - p + 1\right)}{-a + b(2-q)q(1-\alpha) + 2} \tag{6.10}$$

Proof of (6.7). This is simply the observation that $\mathbb{P}((0 \leftarrow \bullet) \wedge \overleftarrow{\bullet}_1) = \mathbb{P}(\overleftarrow{\bullet}_1)$. ☐

Proof of (6.8). We condition on $\{\dot{\bullet}_1\}$ and consider the two ways that $\dot{\bullet}_1$ can be annihilated:

$$\mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1) = p\mathbb{P}(\dot{\bullet}_1 \leftarrow \overleftarrow{\bullet} \mid \dot{\bullet}_1) + p\mathbb{P}((0 \leftarrow \bullet) \wedge (\dot{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) \mid \dot{\bullet}_1).$$

By (6.6), $\mathbb{P}(\dot{\bullet}_1 \leftarrow \overleftarrow{\bullet} \mid \dot{\bullet}_1) = \mathbb{P}(x_1 \overset{1}{\Leftarrow} \overleftarrow{\bullet}) = \alpha q$. For the second term,

$$\mathbb{P}((0 \leftarrow \bullet) \wedge (\dot{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) \mid \dot{\bullet}_1) = \sum_{k>1} \mathbb{P}((x_1 \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_k)_{(x_1,x_k]} \wedge (x_k \leftarrow \bullet)_{(x_k,\infty)}).$$

Since the events on $(x_1, x_k]$ and (x_k, ∞) are independent, we have

$$\begin{aligned}\mathbb{P}((0 \leftarrow \bullet) \wedge (\dot{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) \mid \dot{\bullet}_1) &= \sum_{k>1} \mathbb{P}((x_1 \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_k)_{(x_1,x_k]})q \\ &= \mathbb{P}(x_1 \overset{1}{\longleftrightarrow} \overleftarrow{\bullet})q = (1-\alpha)q^2.\end{aligned}$$

The last equality uses (6.6) again. Putting both terms together, we obtain the claimed formula. ☐

Proof of (6.9). We partition on the various ways that $\vec{\bullet}_1$ is destroyed to write $\mathbb{P}((0 \leftarrow \bullet) \wedge \vec{\bullet}_1)$ as

$$\mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})) + \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet})) + \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet})) \tag{6.11}$$

$$+ \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})) + \mathbb{P}((0 \leftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})). \tag{6.12}$$

The first term of (6.11) is equal to $\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})q$ via the renewal that occurs after $\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}$. For the second term in (6.11), the left arrow that destroys $\vec{\bullet}_1$ will reach 0. The change of measure in Proposition 6.2.1 gives $\mathbb{P}(\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet}) = ((a/2)/c)\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})$. For the third term in (6.11), we note that Proposition 6.2.1 ensures that $\mathbb{P}(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \bullet) = (b/c)\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})$. Next, after a blockade is generated by $\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \bullet$, having 0 visited is a translation of the event $\{(0 \leftarrow \bullet) \wedge \dot{\bullet}_1\}$ conditional on $\dot{\bullet}_1$ already being present (since a blockade was generated). Thus, we write the probability as $\mathbb{P}((0 \leftarrow \bullet) \wedge \dot{\bullet}_1)/p$.

Lastly, observe that the two terms at (6.12) have sum equal to s. $\square$

Proof of (6.10). The proof uses the fact that right arrows in $(0,\infty)$ is destroyed with probability 1. This is proven in [ST17, Lemma 3.4] and the same observation holds for TCBA. The basic reason right arrows cannot survive in the right half-line is that, if they did, symmetry ensures left arrows also survive in the left half-line. Ergodicity gives a positive density of left and right surviving arrows in the full-line, which is a contradiction. This lets us form the partition

$$\mathbb{P}(\vec{\bullet}_1) = \frac{1-p}{2} = \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}) + \mathbb{P}(\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet}) + \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) + \mathbb{P}(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}).$$

The first term on the right side is equal to $s + r$. The last three terms can be transformed using Proposition 6.2.1 which gives

$$\frac{1-p}{2} = s + r + \frac{a/2}{c}\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) + \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) + \frac{b}{c}\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}). \tag{6.13}$$

Solving for $\mathbb{P}(\vec{\bullet} \longleftrightarrow \overleftarrow{\bullet})$ gives the first claimed equality in (6.10).

To obtain the final expression that involves only parameters and q, we use formulas for s and r which we will derive in Lemma 6.2.3. Recall that $a+b+c = 1$ and notice that the formulas for s and r depend only on parameters, q, and

$\hat{p} = \mathbb{P}(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet})$, which by Proposition 6.2.1 can be written as $(b/c)\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})$. Substitute this into the formulas for s and r at (6.14) and (6.15). Then, substitute these expressions into the first equality at (6.10). This results in a linear equation in $\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})$. Solving and simplifying this linear equation, which we used mathematical software to carry out, gives the claimed equality. □

Remark 6.1. *If $c = 0$, similar formulas as in Lemma 6.2.2 could be derived using whichever parameter of a and b is nonzero. For example, if $c = 0$ and $a > 0$, then we would write the terms involving arrow-arrow collision that partition $\mathbb{P}((0 \longleftarrow \bullet) \wedge \vec{\bullet}_1)$ in* (6.11)

$$\mathbb{P}((0 \longleftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})) + \mathbb{P}((0 \longleftarrow \bullet) \wedge (\vec{\bullet}_1 \longleftarrow \overleftarrow{\bullet})) + \mathbb{P}((0 \longleftarrow \bullet) \wedge (\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}))$$

in terms of $\mathbb{P}(\vec{\bullet}_1 \longleftarrow \overleftarrow{\bullet})$ using Proposition 6.2.1, rather than in terms of $\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})$ as we did under the assumption $c > 0$. The application of Proposition 6.2.1 to rewrite (6.13) *could be handled similarly.*

In the following derivations, we consider events that occur on systems restricted to various intervals. So moving forward we include all subscripts for clarity.

Lemma 6.2.3. *Recall that we denote a blockade generated from a $\vec{\bullet}_m \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_n$ collision by $\hat{\bullet}_{m,n}$. Let $\hat{p} = \mathbb{P}((\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet})_{(0,\infty)})$. The following identities hold for TCBA:*

$$\begin{aligned} s &= \mathbb{P}((0 \leftarrow \bullet)_{(0,\infty)} \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})_{(0,\infty)}) + \mathbb{P}((0 \leftarrow \bullet)_{(0,\infty)} \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})_{(0,\infty)}) \\ &= \frac{1}{2}(p + \hat{p})(1-\alpha)q^2 \qquad (6.14) \\ r &= \mathbb{P}((0 \not\leftarrow \bullet)_{(0,\infty)} \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})_{(0,\infty)}) + \mathbb{P}((0 \not\leftarrow \bullet)_{(0,\infty)} \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})_{(0,\infty)}) \\ &= (p + \hat{p})(1-\alpha)q(1-q) \qquad (6.15) \end{aligned}$$

Proof of (6.14). For $m, n \in \mathbb{Z}$ define the indicator random variable

$$\begin{aligned} Z(m,n) = \mathbf{1}\{(\vec{\bullet}_m \longleftrightarrow \dot{\bullet}_n)_{[x_m,\infty)} \wedge (x_n \leftarrow \bullet)_{(x_n,\infty)}\} \\ + \sum_{m<k<n} \mathbf{1}\{(\vec{\bullet}_m \longleftrightarrow \hat{\bullet}_{k,n})_{[x_m,\infty)} \wedge (x_n \leftarrow \bullet)_{(x_n,\infty)}\}. \end{aligned}$$

Note that $Z(m,n) = 0$ when $m \geq n$. $Z(m,n)$ describes the ways in which $\vec{\bullet}_m$ is annihilated by either $\dot{\bullet}_n$ or a $\hat{\bullet}$-particle generated by $\overleftarrow{\bullet}_n$, and subsequently x_n is visited by a left arrow. When the event in the indicator $Z(m,n)$ occurs, we also have $x_m \leftarrow \bullet$. The events in the indicators in the sum $\sum_{n>1} Z(1,n)$ form a partition of the event $\{(0 \leftarrow \bullet)_{(0,\infty)} \wedge ((\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})_{(0,\infty)} \vee (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})_{(0,\infty)})\}$. This gives

$$s = \mathbb{E} \sum_{n>1} Z(1,n). \tag{6.16}$$

Next, let us define $\overrightarrow{\tau}^{(x_n)}$ and $\overleftarrow{\tau}^{(x_n)}$ as the times at which x_n is first visited in TCBA restricted to $(-\infty, x_n)$ and (x_n, ∞), respectively. For example, if 0 is first visited by $\overleftarrow{\bullet}_k$, then $\overleftarrow{\tau}^{(0)} = x_k$. By symmetry $\overrightarrow{\tau}^{(x_n)}$ and $\overleftarrow{\tau}^{(x_n)}$ are i.i.d. Moreover by ergodicity, $\overrightarrow{\tau}^{(x_{n_1})}$ and $\overrightarrow{\tau}^{(x_{n_2})}$ have the same distribution for $n_1 \leq n_2$. Since the interdistances between particles are according to a continuous distribution, we have

$$\mathbb{P}(\overrightarrow{\tau}^{(x_{n_1})} < \overleftarrow{\tau}^{(x_{n_2})} \mid \overrightarrow{\tau}^{(x_{n_1})}, \overleftarrow{\tau}^{(x_{n_2})} < \infty) = \frac{1}{2}. \tag{6.17}$$

By (MTP) and the fact that $Z(m,n) = 0$ for $m \geq n$, (6.16) is equal to

$$\begin{aligned} \mathbb{E} \sum_{m<1} Z(m,1) &= \sum_{m<1} \mathbb{P}\left((\vec{\bullet}_m \longleftrightarrow \dot{\bullet}_1)_{[x_m,\infty)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\ &\quad + \sum_{m<k<1} \mathbb{P}\left((\vec{\bullet}_m \longleftrightarrow \hat{\bullet}_{k,1})_{[x_m,\infty)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\ &= \mathbb{P}\Big((\dot{\bullet}_1) \wedge (\vec{\bullet} \overset{1}{\longleftrightarrow} x_1)_{(-\infty,x_1)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)} \wedge (\overrightarrow{\tau}^{(x_1)} < \overleftarrow{\tau}^{(x_1)})\Big) \\ &\quad + \mathbb{P}\left(\exists\, k < 1 \,:\, (\vec{\bullet} \longleftrightarrow \hat{\bullet}_{k,1})_{(-\infty,\infty)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\ &= \frac{1}{2} p(1-\alpha)q^2 + \frac{1}{2}\hat{p}(1-\alpha)q^2. \end{aligned} \tag{6.18}$$

We will conclude by justifying the equality at (6.18), which will complete the proof of (6.14).

For the first term in (6.18), note that (6.6) ensures it is equal to

$$(1-\alpha)\mathbb{P}\Big((\dot{\bullet}_1) \wedge (\vec{\bullet} \rightarrow x_1)_{(-\infty,x_1)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)} \wedge (\vec{\tau}^{(x_1)} < \overleftarrow{\tau}^{(x_1)})\Big).$$

Notice that the events $(\vec{\bullet} \overset{1}{\longleftrightarrow} x_1)_{(-\infty,x_1)}$ and $(x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}$ imply that $\vec{\tau}^{(x_1)}, \overleftarrow{\tau}^{(x_2)} < \infty$ but provide no further observation regarding the values of these arrival times, that is, they are independent of the event $(\vec{\tau}^{(x_1)} < \overleftarrow{\tau}^{(x_1)})$. Applying (6.17) and independence of the one-sided processes gives

$$\mathbb{P}\Big((\dot{\bullet}_1) \wedge (\vec{\bullet} \overset{1}{\longleftrightarrow} x_1)_{(-\infty,x_1)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)} \wedge (\vec{\tau}^{(x_1)} < \overleftarrow{\tau}^{(x_1)})\Big) = \frac{1}{2}(1-\alpha)pq^2.$$

For the second term in (6.18), it suffices to show that

$$\begin{aligned}
&\mathbb{P}\left(\exists\, k < 1 \,:\, (\vec{\bullet} \longleftrightarrow \hat{\bullet}_{k,1})_{(-\infty,\infty)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\
&\quad = \mathbb{P}\begin{pmatrix} \exists\, k < 1 \,:\, (\vec{\bullet} \overset{1}{\longleftrightarrow} x_k)_{(-\infty,x_k)} \wedge (\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]} \\ \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)} \wedge (\vec{\tau}^{(x_k)} < \overleftarrow{\tau}^{(x_1)}) \end{pmatrix} = \frac{1}{2}(1-\alpha)\hat{p}q^2. \quad (6.19)
\end{aligned}$$

The first equality in (6.19) is justified by observing that, for each $k < 1$, the event

$$\left\{(\vec{\bullet} \longleftrightarrow \hat{\bullet}_{k,1})_{(-\infty,\infty)} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right\}$$

is equal to

$$\Big\{(\vec{\bullet} \overset{1}{\longleftrightarrow} x_k)_{(-\infty,x_k)} \wedge (\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]} \wedge (x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)} \wedge (\vec{\tau}^{(x_k)} < \overleftarrow{\tau}^{(x_1)})\Big\}.$$

Indeed, observe that the right and left arrows have the same constant speed and the blockade $\hat{\bullet}_{k,1}$ is at the midpoint $x_{k,1} = (x_k + x_1)/2$ of the interval $[x_k, x_1]$. In order for the blockade created by the event $(\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]}$ at location $x_{k,1}$ to be annihilated by a right-arrow, the visit corresponding to $(\vec{\bullet} \overset{1}{\longleftrightarrow} x_k)_{(-\infty,x_k)}$ must occur before the first visit to x_1 by left-arrow.

For the second equality in (6.19), note that the event $(\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]}$ can occur only for one $k < 1$, and the restricted processes on $(-\infty, x_k)$ and $[x_k, x_1]$ are independent conditional on $(\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]}$. It follows that, conditional on $(\vec{\bullet} \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{(-\infty,x_1]}$, defining k to be the unique (random) index such that $(\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_1)_{[x_k,x_1]}$ occurs, $(\vec{\bullet} \to x_k)_{(-\infty,x_k)}$ and $(x_1 \leftarrow \overleftarrow{\bullet})_{(x_1,\infty)}$ are independent and symmetric in the sense that the random times $\vec{\tau}^{(x_k)}$ and $\overleftarrow{\tau}^{(x_1)}$ are independent with the same distribution. This yields the second equality in (6.19), as desired. □

Remark 6.2. *It seems that the major obstacle to extending TCBA to a four-parameter family that includes the reactions* $[\vec{\bullet} - \dot{\bullet} \implies \dot{\bullet}]$ *and* $[\dot{\bullet} - \overleftarrow{\bullet} \implies \dot{\bullet}]$ *consists in an asymmetry that arises when comparing certain arrival times of arrows in the one-sided processes. Call a visit to the origin "strong" if the arrow would destroy a blockade there. Call it "weak" if a blockade at* 0 *would survive the interaction. Let* $\vec{S}$ *and* $\vec{W}$ *be the times of the first strong and weak visits to* 0 *from the left in the one-sided process on* $(-\infty, 0)$*, and similarly for* $\overleftarrow{S}$ *and* $\overleftarrow{W}$*. Extending* (6.14) *to these reactions would then require us to compute*

$$\mathbb{P}((\overleftarrow{S} > \vec{S}) \wedge (\vec{S} > \vec{W}) \mid \overleftarrow{S}, \vec{S} < \infty). \tag{6.20}$$

Unlike when comparing $\vec{\tau}$ *and* $\overleftarrow{\tau}$*, there is no obvious symmetry that allows us to compute the value of* (6.20)*. Hence one needs to carry* (6.20)*, as well as an additional term to account for* $\vec{\bullet}_1$ *being destroyed by weakly visiting a blockade, into the equation for* s *at* (6.14)*. And, eventually solving the main recursion for* q *becomes intractable without a value for* (6.20)*. A similar difficulty has been observed in the asymmetric BA considered in [JL22]. There,* $\vec{\tau}$ *and* $\overleftarrow{\tau}$ *may have different distribution so* $\mathbb{P}(\vec{\tau} \le \vec{\tau})$ *cannot be inferred via symmetry. It may depend on the system parameters.*

Proof of (6.15). The proof is similar to (6.14), but uses the modified indicators

$$Z'(m,n) = \mathbf{1}\{(\vec{\bullet}_m \longleftrightarrow \dot{\bullet}_n)_{[x_m,\infty)} \wedge (x_n \nleftarrow \bullet)_{(x_n,\infty)}\}$$
$$+ \sum_{m<k<n} \mathbf{1}\{(\vec{\bullet}_m \longleftrightarrow \hat{\bullet}_{k,n})_{[x_m,\infty)} \wedge (x_n \nleftarrow \bullet)_{(x_n,\infty)}\}$$

The difference from the indicators used in (6.14) is that we require that x_n *not* be visited by a left arrow.

By the definition, $Z'(m,n) = 0$ for $m \geq n$. Hence

$$r = \mathbb{E}\sum_{n>1} Z'(1,n) = \mathbb{E}\sum_{n\in\mathbb{Z}} Z'(1,n) = \mathbb{E}\sum_{m\in\mathbb{Z}} Z'(m,1) = \mathbb{E}\sum_{m<1} Z'(m,1),$$

where the second equality uses the mass transport principle. Using similar reasoning as in the proof of (6.14), this expands to be the claimed formula for r. Here is the derivation, but without the similar explanations given in the proof of (6.14):

$$\begin{aligned}
r &= \sum_{m<1} \mathbb{P}((\vec{\bullet}_m \longleftrightarrow \dot{\bullet}_1)_{[x_m,\infty)} \wedge (x_1 \nleftarrow \bullet)_{(x_1,\infty)}) \\
&\qquad + \sum_{m<k<1} \mathbb{P}\left((\vec{\bullet}_m \longleftrightarrow \hat{\bullet}_{k,1})_{[x_m,\infty)} \wedge (x_1 \nleftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\
&= \mathbb{P}\left((\dot{\bullet}_1) \wedge (\vec{\bullet} \overset{1}{\longleftrightarrow} x_1)_{(-\infty,x_1)} \wedge (x_1 \nleftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\
&\qquad + \mathbb{P}\left(\exists\, k < 1 \ : \ (\vec{\bullet} \longleftrightarrow \hat{\bullet}_{k,1})_{(-\infty,\infty)} \wedge (x_1 \nleftarrow \overleftarrow{\bullet})_{(x_1,\infty)}\right) \\
&= (1-\alpha)pq(1-q) + (1-\alpha)\hat{p}q(1-q).
\end{aligned}$$

□

Proposition 6.2.4. *For TCBA, it holds for all $p \in [0,1)$ that $0 = (1-q)g(p,q)$ with g defined as*

$$g(u,v) = \frac{1 - u - 2(1-\alpha)uv - b(1-\alpha)v^2 - (1-(a+b))(1-\alpha)uv^2}{2 - a + b(1-\alpha)(2-v)v}. \tag{6.21}$$

Proof. Lemma 6.2.2 gives a formula for each term in the partition for q at (6.3) that depends only on p, q, a, b, and α. This yields $q = G(p, q, a, b, \alpha)$ for an explicit function G. Some algebra gives that $0 = -q + G(p, q, a, b, \alpha) = (1 - q)g(p, q)$ as claimed. □

6.3 Properties of g

We will derive some elementary, but useful properties of the function g from (6.21). It will be helpful for presenting our arguments to distinguish p_c from its claimed formula in Theorem 6.1.1. So, we introduce the term

$$p_* = \frac{1 - b(1 - \alpha)}{4 - 3\alpha - (a + b)(1 - \alpha)}. \tag{6.22}$$

Note that the denominator is positive as it can be rearranged as $4 - (a + b) - (3 - (a + b))\alpha$. With the notation of p_*, the first part of Theorem 6.1.1 could be restated as $p_c = p_*$ for all TCBA. Note that p_* is derived by solving the equation $g(u, 1) = 0$. In some sense this corresponds to finding the transition point for q being a root of $1 - q$ to being a root of $g(p, q(p))$ seen in Figure 6.1. We record the fact that p_* is the unique such solution in the following lemma.

Lemma 6.3.1. *$u = p_*$ is the unique solution to $g(u, 1) = 0$.*

Proof. It is easily checked that $g(p_*, 1) = 0$. Moreover, since g is linear in u, it follows that this solution is unique. □

Lemma 6.3.2. *$\partial_u g(u, v), \partial_v g(u, v) < 0$ for all TCBA and $u, v \in [0, 1]$.*

Proof. The partial derivatives of g are

$$\begin{aligned}
\partial_u g(u,v) &= -\frac{v^2(1-\alpha)(1-(a+b)) + 2v(1-\alpha) + 1}{2-a+b(2-v)v(1-\alpha)}, \\
\partial_v g(u,v) &= -\frac{2(1-\alpha)((2-a)u + b(1-u))\Big((1-a)v + bv^2(1-\alpha) + 1\Big)}{(2-a+b(2-v)v(1-\alpha))^2}.
\end{aligned}$$

As all parameters lie in $[0,1)$, it is easy to check that both partial derivatives are negative. □

Lemma 6.3.3.

$$\begin{gathered}
g(u,v) > 0 \textit{ for } (u,v) \in (0,p_*) \times [0,1] \\
g(u,v) < 0 \textit{ for } (u,v) \in (p_*,1] \times [1,\infty].
\end{gathered}$$

Proof. Lemma 6.3.1 states that $u = p_*$ is the unique solution to $g(u,1) = 0$. Combining this observation with Lemma 6.3.2 immediately implies the claimed inequalities. □

Lemma 6.3.4. *For all TCBA, the equation* $g(p, q(p)) = 0$ *has two distinct solutions:*

$$q_\pm(p) := \frac{-p(1-\alpha) \pm \sqrt{(1-\alpha)\left(b(1-p)^2 - p(a(1-p) + p\alpha - 1)\right)}}{(1-\alpha)((1-a)p + b(1-p))}. \qquad (6.23)$$

Moreover, it holds that:

(i) $q_-(p) < 0$ *for* $p \in [0,1)$.

(ii) $q_+(p) > 1$ *for* $p \in [0,p_*)$.

(iii) $q_+(p) < 1$ *for* $p \in (p_*,1)$.

Proof. Since the numerator of g is quadratic in v, the formula for $q_\pm(p)$ follows from the quadratic formula. Notice that the discriminant D of (6.23) can be rewritten as

$$\frac{D}{1-\alpha} = p(1-(a+b)) + b(1-p) + p^2(a+b) - p^2\alpha$$

After multiplying the $p(1-(a+b))$ term by p we have for $p \in [0,1)$ that

$$\begin{aligned} \frac{D}{1-\alpha} &> p^2(1-(a+b)) + b(1-p) + p^2(a+b) - p^2\alpha \\ &= b(1-p) + p^2(1-\alpha) \\ &> 0. \end{aligned}$$

Thus, q_- and q_+ are distinct.

Since the numerator of q_- is negative and the denominator is positive, we immediately observe (i). To deduce (ii) and (iii), first notice that the uniqueness observation in Lemma 6.3.1 along with the fact that $g(p_*, q_+(p_*)) = 0$ imply that $q_+(p_*) = 1$. This observation and Lemma 6.3.2 together imply (ii) and (iii). □

6.4 Regularity conditions

The goal of this section is to prove the following proposition.

Proposition 6.4.1. $I := \{p \in (p_*, 1) \colon q(p) < 1\} = (p_*, 1)$.

Proof. It follows from Lemma 6.4.2 and Lemma 6.4.5 that $\{p \in (p_*, 1) \colon q(p) < 1\}$ and its complement are both open in the subspace topology on $(p_*, 1)$. Thus, $I = \varnothing$ or $I = (p_*, 1)$. By Lemma 6.4.6, I is nonempty so the latter holds. □

Lemma 6.4.2. $I^c := \{p \in (p_*, 1) \colon q(p) = 1\}$ *is open in the subspace topology on* $(p_*, 1)$.

Proof. Let g be as in (6.21). Lemma 6.3.3 gives that there are two distinct solutions $v \in \{q_-, q_+\}$ to $g(p, v) = 0$. Lemma 6.3.3 (i) states that $q_- < 0$ for all p. As $q(p)$ is a probability, we cannot have $q(p) = q_-(p)$. The remaining possibilities are that $q = 1$ or $q = q_+$.

Let $q_k = \mathbb{P}((0 \leftarrow \bullet)_{[0,x_k]})$. The quantities $q_k \uparrow q$ only involve the initial configurations concerning finitely many particles. Conditioning on the velocities of particles in the configuration and integrating to account for interdistances between particles gives a polynomial in p. It follows from Lemma 6.3.3 (iii) that

$$q(p) = 1 \text{ if and only if } q(p) > q_+(p) \text{ for } p > p_*. \tag{6.24}$$

We use (6.24) to give the following characterization of the set on which $q(p) = 1$:

$$\{p \in (p_*, 1) \colon q(p) = 1\} = \{p \in U \colon q(p) > q_+(p)\} = \bigcup_{k=1}^{\infty} \{p \in U \colon q_k(p) > q_+(p)\}.$$

Continuity of the q_k and $q_+(p)$ ensures that the sets $\{p \in U \colon q_k(p) > q_+(p)\}$ are open. Thus, so is the union which is equal to I^c. □

6.4.1 Superadditivity

The idea underlying the upcoming Lemma 6.4.5, that I is open, was developed in [ST17], and extended in [HST21, Lemma 10]. A more general version tailored to the asymmetric setting is proven in [JL22]. Coalescence makes for new challenges. Before getting to the lemma, we introduce some additional notation.

Let

$$B(j,k) = \sum_{i=j}^{k} \mathbf{1}\{(\dot{\bullet}_i \text{ survives})_{[x_j,x_k]}\}$$

be the number of blockades that survive in the process restricted to $[x_j, x_k]$. Note that $B(j,k)$ only counts surviving blockades from the initial configuration of particles; blockades generated from $\vec{\bullet} - \overleftarrow{\bullet} \implies \dot{\bullet}$ reactions do not contribute. We use the quiver interpretation of TCBA described in Section 6.1.2.

To briefly summarize, each arrow in the process, call such particles *original arrows*, carries a *quiver* of Geometric$(1-\alpha)$ many *sharp arrows*. When the arrow meets a blockade one of the sharp arrows mutually annihilates with the blockade, but the original arrow along with its quiver of any remaining sharp arrows continue moving. The original arrow is destroyed at the moment its last sharp arrow is destroyed. When two original arrows meet, some arbitrary rule is used to decide the reaction. If an original arrow is destroyed in a $\vec{\bullet} - \overleftarrow{\bullet}$ collision, its quiver of sharp arrows is also destroyed.

In the quiver formulation, we view sharp arrows as distinct particles that follow the same trajectory as the original arrow. These arrows track how many blockades the original arrow could potentially destroy. Let $A(j,k)$ be the number of left and right sharp arrows that survive in $[x_j, x_k]$. Define the count

$$N(j,k) = B(j,k) - A(j,k). \tag{6.25}$$

This is in some sense a worst-case weighting of surviving blockades; every surviving sharp arrow is treated like it will ultimately destroy a blockade.

We now show that N is superadditive after merging intervals. In the analogue [HST21, Lemma 12], it is required that the right interval have no surviving right arrows. We show that, with our more general weighting scheme, superadditivy holds with no hypotheses about surviving arrows. This extra level of generality ends up being crucial for proving the upcoming Lemma 6.4.4. See Remark 6.3.

Lemma 6.4.3. *Let $k < \ell$ be positive integers. For any initial assignment of particle types and spacings to $(\bullet_i)_{i\in\mathbb{Z}}$ we have*

$$N(1,\ell) \geq N(1,k) + N(k+1,\ell). \tag{6.26}$$

Proof. Let $I = [x_1, x_k]$ and $J = [x_{k+1}, x_\ell]$. We show that when sharp arrows that survived in the process on I or J are destroyed, this has at worst a net-neutral effect on the difference between surviving blockades and sharp arrows. We track these changes in real time.

Using any index system that uniquely identifies blockades and sharp arrows, define $\mathcal{B}^I, \mathcal{B}^J, \mathcal{A}^I$, and $\mathcal{A}^J$ to be the sets of blockades and sharp arrows that survive in the processes restricted to I and J. The sets $\mathcal{B}^I$ and $\mathcal{B}^J$ only include blockades from the original configuration; blockades generated from $\vec{\bullet}\!-\!\overleftarrow{\bullet} \implies \dot{\bullet}$ reactions are not counted. We will define a pair of set-valued processes $(\mathcal{A}_t, \mathcal{B}_t)$ with the following properties:

- (i) $\mathcal{B}_t$ is the set of blockades from $\mathcal{B}^I \cup \mathcal{B}^J$ that are still surviving at time t in the combined process on $I \cup J$.
- (ii) $\mathcal{A}_0 = \mathcal{A}^I \cup \mathcal{A}^J$.
- (iii) $\mathcal{A}_t$ is non-increasing and $\mathcal{B}_t$ may decrease only when a decrease of the same magnitude occurs in $\mathcal{A}_t$.
- (iv) For $T := x_\ell - x_1$, we have $|\mathcal{B}_T| = B(1, \ell)$ and $|\mathcal{A}_T| = A(1, \ell)$.

Using the characterization of $N(1, \ell)$ at (6.25), these properties give (6.26) since

$$N(1,k) + N(k+1,\ell) \overset{(i),(ii)}{=} |\mathcal{B}_0| - |\mathcal{A}_0| \overset{(iii)}{\leq} |\mathcal{B}_T| - |\mathcal{A}_T| \overset{(iv)}{=} N(1,\ell).$$

It remains to define $(\mathcal{A}_t, \mathcal{B}_t)$ and prove that it satisfies (i), (ii), (iii), and (iv).

We take property (i) as the definition of $\mathcal{B}_t$ and (ii) as the definition of $\mathcal{A}_0$. Let $t_0 = 0$ and $t_1 > t_0$ be the first time in the combined process on $I \cup J$ that a sharp arrow, say μ, from $\mathcal{A}_{t_0}$ is annihilated. If the collision involves two sharp arrows from $\mathcal{A}_{t_0}$, let μ be the left arrow. There are three possible ways that μ is destroyed:

(I) If μ is annihilated by a blockade $\dot{\bullet}_i$ counted by $\mathcal{B}_{t_0}$, then set $\mathcal{A}_{t_1} = \mathcal{A}_{t_0} \setminus \{\mu\}$ and $\mathcal{B}_{t_1} = \mathcal{B}_{t_0} \setminus \{\dot{\bullet}_i\}$.

(II) If μ is annihilated from hitting a blockade $\dot{\bullet}$ that does not belong to $\mathcal{B}_{t_0}$, then $\mathcal{B}_{t_1} = \mathcal{B}_{t_0}$.

  (a) Let μ' be the first sharp arrow that: reaches the location of $\dot{\bullet}$ in the process on $I \cup J$, has the opposite direction of μ, and is not counted by $\mathcal{A}_{t_0}$. Set $\mathcal{A}_{t_1} = (\mathcal{A}_{t_0} \setminus \{\mu\}) \cup \{\mu'\}$.

  (b) If there is no such μ', then set $\mathcal{A}_{t_1} = \mathcal{A}_{t_0} \setminus \{\mu\}$.

In this case, the blockade $\dot{\bullet}$ is generated by an arrow-arrow collision between times t_0 and t_1. In (II)(a), we account for the case that μ' would have been annihilated by $\dot{\bullet}$ in the combined process if not for μ. Note that μ' may or may not have already been in $\mathcal{A}_{t_0}$.

(III) If μ is destroyed by another arrow μ'', then define $\mathcal{A}_{t_1}$ to be $\mathcal{A}_{t_0}$ minus all of the sharp arrows in the same quiver as μ as well as—if the reaction is mutual annihilation and $\mu'' \in \mathcal{A}_{t_0}$—all arrows in the same quiver as μ''. Set $\mathcal{B}_{t_1} = \mathcal{B}_{t_0}$.

Iterate this procedure by considering the next sharp arrow from $\mathcal{A}_{t_j}$ to be destroyed at time $t_{j+1} > t_j$. This gives new values of $\mathcal{A}_{t_{j+1}}$ and $\mathcal{B}_{t_{j+1}}$ according to whichever of (I), (II), or (III) occurs. We make the process right continuous for $t \geq 0$ by setting $\mathcal{A}_t = \mathcal{A}_{t_j}$ for all $t \in [t_j, t_{j+1})$ and $j \geq 0$. By construction, we have (iii).

To show (iv), first notice that, since arrows have unit speed, $\mathcal{A}_t$ and $\mathcal{B}_t$ must fixate after the length of the interval $T = x_\ell - x_1$ has elapsed. $\mathcal{A}_t$ tracks all potentially surviving sharp arrows. By construction, $\mathcal{A}_T$ does not contain non-surviving sharp arrows; when a sharp arrow is annihilated, it is removed from the process. The only time a new sharp arrow is added to $\mathcal{A}_t$ is when (II.a) occurs, in which case the added sharp arrow has survived thus far.

Next, we show that *all* surviving sharp arrows in the combined process are contained in $\mathcal{A}_T$. First observe that $\mathcal{A}_T$ contains all sharp arrows from $\mathcal{A}_0$ that survive in the combined process, as an arrow is removed only if it is annihilated. If a sharp arrow survives in the combined process but is not in $\mathcal{A}_0$, then it must be because the particle that it was going to be annihilated by in the separate process is annihilated by an arrow from the other interval in the combined process. Recall that although blockades can be generated through collisions, arrows cannot. In fact, the "unleashing" of an arrow is only possible when it was going to be annihilated with a blockade and this is captured in case (II.a) when the unleashed arrow is added to $\mathcal{A}_t$. This gives the first part of (iv), that $\mathcal{A}_T = A(1, \ell)$.

To show the second half of (iv), that $|\mathcal{B}_T| = B(1, l)$, we first note that $\mathcal{B}_0$ contains all possible surviving blockades in the combined process – non-surviving blockades in the separate processes would be destroyed by the same sharp arrows that destroyed them in the separate processes if not by other arrows from opposite intervals prior. Since the dynamic construction of $\mathcal{B}_t$ only removes annihilated blockades, $\mathcal{B}_T$ consists of all surviving blockades by the end of the combined process. □

6.4.2 A necessary and sufficient condition for blockade survival

Lemma 6.4.4. *Let* $N_k = N(1,k)$ *with* N *defined at* (6.25). *For all* $p \in (0,1)$ *it holds that*

$$\theta(p) > 0 \iff \text{there exists } k \geq 1 \text{ with } \mathbb{E}N_k > 0. \tag{6.27}$$

Proof. The forward implication is analogous to [HST21, Proposition 11]. The key observation is that for any x_i in an interval I, $P((\dot{\bullet}_i \text{ survives})_I)$ is decreasing in I. This continues to hold with the coalescence rules in TCBA. To see why, suppose that we have a configuration of particles in I with $(\dot{\bullet}_i \text{ does not survive})_I$. Keeping this configuration fixed, there is no manner in which one could add particles outside of I to intercept the particle that destroys $\dot{\bullet}_i$ in time. This monotonicity ensures that

$$\mathbb{E}B(1,k) = \sum_{i=1}^{k} \mathbb{P}((\dot{\bullet}_i \text{ survives})_{[x_1,x_k]}) \geq \sum_{i=1}^{k} \mathbb{P}((\dot{\bullet}_i \text{ survives})_{\mathbb{R}}) = kp\,\theta(p) \to \infty$$

with $B(1,k)$ from (6.25).

Let $\overleftarrow{A}_k$ be the sharp left arrows counted by $A(1,k)$ from (6.25) and $\vec{A}_k$ the sharp right arrows so that $A(1,k) = \overleftarrow{A}_k + \vec{A}_k$. We will next show that $\mathbb{E}\overleftarrow{A}_k$ and $\mathbb{E}\vec{A}_k$ are bounded. Symmetry ensures that $\mathbb{E}\overleftarrow{A}_k = \mathbb{E}\vec{A}_k$, so we only provide the argument for $\overleftarrow{A}_k$. For each $i \geq 1$, the event that $\overleftarrow{\bullet}_i$ survives from the restriction to $[x_1, x_k]$ is non-decreasing in k, since such arrows do not interact with any particles to their right. This monotonicity implies that $\overleftarrow{A}_k \uparrow \overleftarrow{A}_\infty$. Let T denote the number of surviving left arrows counted from $\overleftarrow{A}_\infty$ so that $\overleftarrow{\bullet}_{i_1}, \dots, \overleftarrow{\bullet}_{i_T}$ reach 0 for some integers $i_1, \dots, i_T \geq 1$. When $\theta(p) > 0$, by using a renewal property of TCBA, T is a geometric random variable supported on $0, 1, \dots$ with parameter $1 - q$. Hence we have

$$\overleftarrow{A}_k \uparrow \overleftarrow{A}_\infty := \textstyle\sum_{j=1}^{T} G_j,$$

where each G_j an independent geometric random variable supported on $1, 2, \dots$ with parameter $1 - \alpha$. The G_j count how many additional sharp arrows remain in the quiver of the jth arrow to arrive to 0. G_j has a geometric distribution because of the memoryless property. Hence,

$$\mathbb{E}A(1,k) = 2\mathbb{E}\overleftarrow{A}_k \le \frac{2}{1-q}\frac{1}{1-\alpha} < \infty.$$

Since $\mathbb{E}B(1,k) \to \infty$ as $k \to \infty$, we have $\mathbb{E}N_k = \mathbb{E}B(1,k) - \mathbb{E}A(1,k) > 0$ for some large k. This shows the forward implication of (6.27).

Towards proving the reverse implication of (6.27), suppose that $k \ge 1$ is such that $\mathbb{E}N_k > 0$. We reveal the configuration on k consecutive particles at a time. To this end, let $\tilde{N}^{(i)} = N((i-1)k + 1, ik)$ and denote $\tilde{S}_n = \sum_{i=1}^{n} \tilde{N}^{(i)}$, $n \ge 1$. Notice that the $\tilde{N}^{(i)}$ are i.i.d. with common distribution N_k and $\mathbb{E}N_k = \mathbb{E}\tilde{N}^{(1)} > 0$. Hence $(\tilde{S}_n - \tilde{S}_1)_{\ge 1}$ is a random walk with positive drift, so it stays strictly above 0 with positive probability. Independently from this event, with positive probability the first k particles right of 0 are blockades, i.e., $\tilde{N}^{(1)} = k$. Applying Lemma 6.4.3 gives that

$$N(1,nk) \ge N^{(1)} + (\widetilde{S}_n - \widetilde{S}_1) > k, \qquad \text{for all } n \ge 2 \tag{6.28}$$

with positive probability.

On the event that (6.28) holds, we claim that 0 is never visited by a left-moving particle. Indeed, consider the $(n+1)$st reveal of the k particles from the interval $J_{n+1} := [x_{1+nk}, x_{(n+1)k}]$. Since $N(1,nk) > k$, there are least $k+1$ surviving blockades in $I_n := [x_1, x_{nk}]$. If 0 is reached by one of the particles from J_{n+1}, then no blockades survive in I_n. However, at most k new blockades are introduced by J_{n+1}. So, if this occurs, we would have $N(1,(n+1)k) \le k$, which contradicts (6.28). It follows that $\theta(p) > 0$ since 0 is never visited with at least the probability that (6.28) occurs. □

Remark 6.3. *When we apply Lemma 6.4.3 at* (6.28) *it is important that we have a formulation of superadditivity that accounts for the effects of surviving right arrows. Unlike the approach used in [HST21], we cannot extend each interval of size k a random amount to eliminate any surviving right-arrows without changing the distribution of* $\tilde{N}^{(i)}$*. Indeed, expanding the interval until the surviving right arrows are removed may bring in left arrows that not only destroy the surviving right arrows, but also some of the blockades in the initial segment. Our formulation of subadditivity could be applied to the argument in [HST21] and would slightly simplify their proof of the analogue of Lemma 6.4.4.*

One might further wonder why we can get away with what appears to be a stronger necessary condition ($\theta > 0 \implies \exists k \colon \mathbb{E}[B_k - (\vec{A}_k + \overleftarrow{A}_k)] > 0$) *than that observed in [HST21]* ($\theta > 0 \implies \exists k \colon \mathbb{E}[B_k - \overleftarrow{A}_k] > 0$)*. As the proof of Lemma 6.4.4 reveals, there are* $O(1)$ *blockades and* $\Omega(n)$ *surviving blockades. Thus, it does not matter if we include the worst-case remaining impact from the "dust" of surviving arrows. For similar reasons, we do not need to include the blockades resulting from* $\vec{\bullet} - \overleftarrow{\bullet} \implies \dot{\bullet}$ *reactions in* B_k*.*

Lemma 6.4.5. $\{p \in (0,1) \colon q(p) < 1\}$ *is open in the subset topology on* $(0,1)$*.*

Proof. Using Lemma 6.4.4, we have

$$\{p \in (p_*, 1) \colon \theta(p) > 0\} = \bigcup_{k \geq 1} \{p \in (p_*, 1) \colon \mathbb{E}N_k > 0\}.$$

The function $p \mapsto \mathbb{E}N_k$ is easily seen to be a finite polynomial in p, and so the sets in the union are open. □

Lemma 6.4.6. *For TCBA, it holds that*

$$p_c^+ \leq \frac{1}{2 - \alpha} < 1.$$

Proof. By Lemma 6.4.4, it suffices to prove that $\mathbb{E}N_1 > 0$ for p large enough. We can easily compute

$$\mathbb{E}N_1 = p - (1 - p)\frac{1}{1 - \alpha}.$$

This is strictly greater than 0 for $p > 1/(2 - \alpha) < 1$ as claimed. □

6.5 Proof of Theorem 6.1.1

Proof. Let p_* be as defined at (6.22). We will write q in place of $q(p)$ and similarly for the functions q_- and q_+ from Proposition 6.2.4. Recall that Proposition 6.2.4 gives that $v = q$ is a solution to the equation $(1 - v)g(p, v) = 0$ which has solutions $\{1, q_-, q_+\}$. Lemma 6.3.3 (i) states that $q_- < 0$. Consequently, q must equal either 1 or q_+. For $p \in [0, p_*)$, Lemma 6.3.3 (ii) states that $q_+ > 1$. We immediately deduce that $q = 1$ for $p \in [0, p_*)$. Next, Proposition 6.4.1 states that $q < 1$ for $p \in (p_*, 1)$. As $q_- < 0$, this leaves $q = q_+$ as the only possible value for q for p in the interval $(p_*, 1)$. Lastly, Lemma 6.3.1 implies that $q_+(p_*) = 1$, and thus $q(p_*)$, which must equal either 1 or $q_+(p_*)$, also equals 1. This gives the claimed value of p_c and claimed formula for $q(p)$.

□

CHAPTER 7

ARRIVALS ARE UNIVERSAL IN COALESCING BALLISTIC ANNIHILATION

7.1 Introduction

We restrict our attention to the *three-parameter coalescing ballistic annihilation* (TCBA) systems considered in Chapter 6. TCBA allows for moving particles to spontaneously survive collisions (equivalently, to generate a new moving particle), or to generate a $\dot{\bullet}$-particle. Fix parameters $0 \le a, b, c \le 1$ with $a + b \le 1$. Let $[\bullet \mathbin{—} \bullet \implies \Theta,\ \theta]$ denote a collision that generates $\Theta \in \{\dot{\bullet}, \vec{\bullet}, \overleftarrow{\bullet}, \varnothing\}$ independently with probability θ. The reaction rules are:

$$\vec{\bullet} \mathbin{—} \overleftarrow{\bullet} \implies \begin{cases} \overleftarrow{\bullet}, & a/2 \\ \vec{\bullet}, & a/2 \\ \dot{\bullet}, & b \\ \varnothing, & 1-(a+b) \end{cases} \qquad \begin{aligned} \dot{\bullet} \mathbin{—} \overleftarrow{\bullet} &\implies \begin{cases} \overleftarrow{\bullet}, & c \\ \varnothing, & 1-c \end{cases} \\ \vec{\bullet} \mathbin{—} \dot{\bullet} &\implies \begin{cases} \vec{\bullet}, & c \\ \varnothing, & 1-c \end{cases} \end{aligned}. \tag{7.1}$$

We will refer to the special case $a = b = c = 0$ with only mutual annihilation as *simple ballistic annihilation*.

Denote the event that the site x is visited by the particle started at x_k by $x \mathbin{—} \overleftarrow{\bullet}_k$. Let $A = \min\{k \colon 0 \mathbin{—} \overleftarrow{\bullet}_k\}$ be the index of the first particle to reach the origin. It was proven in [HST21, Theorem 2] that the law of A does not depend on the spacing distribution μ for simple ballistic annihilation. We generalize this to TCBA.

Theorem 7.1.1. *The law of A does not depend on μ for TCBA and $\mathbb{E}[t^A]$ satisfies the* *recursion at* (7.20).

One of the main quantities of interest in ballistic annihilation is $q := \mathbb{P}(A < \infty)$ and the phase transition $p_c = \sup\{p \colon q = 1\}$. It is proven in [HST21] for simple ballistic annihilation and in Chapter 6 for TCBA that q and p_c do not depend on μ. In [HST21], the authors further discovered that the *skyline* of collision shapes for $p > p_c$ does not depend on μ. In [HT21], additional spacing universality properties were observed for simple ballistic annihilation as well as for an asymmetric version introduced in [JL22]. Broutin and Marckert proved that the related bullet process with finitely many particles has a universal law governing the number of surviving particles that does not depend on the velocity or spacing laws [BM20].

Ballistic annihilation dynamics are notoriously sensitive to perturbation; changing the velocity of a single particle can have cascading effects. This feature makes the coalescing version significantly more complex. Our interest in establishing Theorem 7.1.1 comes from a desire to understand the limits of techniques successfully applied to simple ballistic annihilation. Theorem 7.1.1 marks a step in this direction and suggests that TCBA may share other universality properties with simple ballistic annihilation. An additional feature of Theorem 7.1.1 is the implicit recursion of the generating function of A. A special case of this recursion was utilized in [HST21, Theorem 3] to describe the rate of decay of $\mathbb{P}(A > n)$. Our more general formula at (7.20) is a first step towards describing the right tail of the distribution of A in TCBA.

The method of proving Theorem 7.1.1 is similar to what was done in [HST21]. The idea is to prove by induction that the coefficients of the generating function $\mathbb{E}[t^A]$ do not depend on μ. Coalescence makes the details more involved and requires additional considerations. For example, we must distinguish between strong and weak visits. This brings in a second generating function related to the index of the first strong visit to 0.

7.2 Proof of Theorem 7.1.1

We will write $\hat{\bullet}$ to denote a stationary particle generated from a $\vec{\bullet} — \overleftarrow{\bullet}$ collision. If the collision involved $\vec{\bullet}_j$ and $\overleftarrow{\bullet}_k$, then we write $\hat{\bullet}_{j,k}$ for the stationary particle now inhabiting $(x_j + x_k)/2$.

For positive integers j and k, we define the collision events

$$\begin{aligned}
\dot{\bullet}_j \longleftrightarrow \overleftarrow{\bullet}_k &:= (\dot{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\dot{\bullet}_j — \overleftarrow{\bullet}_k \implies \varnothing\} \\
\dot{\bullet}_j \leftarrow \overleftarrow{\bullet}_k &:= (\dot{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\dot{\bullet}_j — \overleftarrow{\bullet}_k \implies \overleftarrow{\bullet}_k\} \\
\vec{\bullet}_j \longleftrightarrow \overleftarrow{\bullet}_k &:= (\vec{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\vec{\bullet}_j — \overleftarrow{\bullet}_k \implies \varnothing\} \\
\vec{\bullet}_j \leftarrow \overleftarrow{\bullet}_k &:= (\vec{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\vec{\bullet}_j — \overleftarrow{\bullet}_k \implies \overleftarrow{\bullet}_k\} \\
\vec{\bullet}_j \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_k &:= (\vec{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\vec{\bullet}_j — \overleftarrow{\bullet}_k \implies \hat{\bullet}_{j,k}\} \\
\vec{\bullet}_j \leftharpoonup \overleftarrow{\bullet}_k &:= (\vec{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\vec{\bullet}_j — \overleftarrow{\bullet}_j \implies \text{all but } \vec{\bullet}_j\} \\
\vec{\bullet}_j \leftharpoondown \overleftarrow{\bullet}_k &:= (\vec{\bullet}_j — \overleftarrow{\bullet}_k) \wedge \{\vec{\bullet}_j — \overleftarrow{\bullet}_k \implies \vec{\bullet}_j\}.
\end{aligned}$$

Specify generic collision events by:

$$\begin{aligned}
\bullet_j — \overleftarrow{\bullet} &:= \{\text{there exists } k \text{ with } \{\bullet_j — \overleftarrow{\bullet}_k\}\} \\
\vec{\bullet}_j — \dot{\bullet} &:= \{\text{there exists } k \text{ with } \{\vec{\bullet}_j — \dot{\bullet}_k\}\} \\
\vec{\bullet}_j — \hat{\bullet} &:= \{\text{there exist } k \text{ and } \ell \text{ with } \{\vec{\bullet}_j — \hat{\bullet}_{k,\ell}\}\},
\end{aligned}$$

where — can be replaced by any admissible collision event.

To determine how reactions occur, we assign to each $\vec{\bullet}$-particle a stack of independent instructions for $\vec{\bullet} \text{---} \dot{\bullet}$ reaction types with probabilities as at (7.1). When $\vec{\bullet}_j$ collides with a blockade the smallest index unused instruction is used to determine the reaction type. We assign to each $\overleftarrow{\bullet}$-particle two independent stacks of reaction instructions distributed as at (7.1) to determine the outcomes of $\vec{\bullet} \text{---} \overleftarrow{\bullet}$ and $\dot{\bullet} \text{---} \overleftarrow{\bullet}$ collisions. This construction ensures that the reaction type of the next collision can be read off from the instructions before it occurs. Thus, the following visiting events are well-defined.

$$x_j \overset{\cdot}{\longleftrightarrow} \bullet_k := (x_j \text{---} \overleftarrow{\bullet}_k) \wedge \{\overleftarrow{\bullet}_k \text{ mutually annihilate in its next } \dot{\bullet} \text{ collision}\}$$

$$x_j \leftarrow \overleftarrow{\bullet}_k := (x_j \text{---} \overleftarrow{\bullet}_k) \wedge \{\overleftarrow{\bullet}_k \text{ survives its next } \dot{\bullet} \text{ collision}\}$$

$$x_j \leftharpoonup \overleftarrow{\bullet}_k := (x_j \text{---} \overleftarrow{\bullet}_k) \wedge \{\text{the next } \vec{\bullet}\text{-particle } \overleftarrow{\bullet}_k \text{ meets is annihilated}\}$$

$$x_j \leftharpoondown \overleftarrow{\bullet}_k := (x_j \text{---} \overleftarrow{\bullet}_k) \wedge \{\text{the next } \vec{\bullet}\text{-particle } \overleftarrow{\bullet}_k \text{ meets survives}\}.$$

Given an interval I and an event B, we write B_I for the event in TCBA restricted to only the particles initially in I. We will use various forms of renewal that occur in TCBA. These come from the fact that the particles behind a moving particle cannot influence events involving the moving particle. For example, $\mathbb{P}((x_\ell \text{---} \overleftarrow{\bullet}) \mid (\dot{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}_\ell)) = \mathbb{P}(x_\ell \text{---} \overleftarrow{\bullet}) = \mathbb{P}(0 \text{---} \overleftarrow{\bullet})$.

We will call a visit to x_j by a left-moving particle where the left-moving particle will destroy the next right-moving particle it meets (as in $\{x_j \leftharpoonup \overleftarrow{\bullet}_j\}$) a *strong* visit. On the other hand, we refer to *weak* visits to x_j as those by a left-moving particle whose next collision with a right-moving particle will result in the right-moving particle surviving. We designate first visits to a given site with the following notation:

$$x_j \overset{1}{\text{---}} \overleftarrow{\bullet}_k := \{\overleftarrow{\bullet}_k \text{ is the first left-moving particle to reach } x_j\}.$$

It will be necessary to refine the notion of a first visit into the following events:

$$x_j \overset{1}{\leftharpoonup} \overset{\leftarrow}{\bullet}_k := \{\overset{\leftarrow}{\bullet}_k \text{ is the first left-moving particle to strong visit } x_j\}$$

$$x_j \overset{1}{\leftharpoondown} \overset{\leftarrow}{\bullet}_k := \{\overset{\leftarrow}{\bullet}_k \text{ is the first left-moving particle to weak visit } x_j\}$$

$$x_j \overset{1}{\longleftrightarrow} \overset{\leftarrow}{\bullet}_k := (x_j \overset{1}{\relbar\relbar} \overset{\leftarrow}{\bullet}_k) \wedge (x_j \overset{\cdot}{\longleftrightarrow} \overset{\leftarrow}{\bullet}_k)$$

$$x_j \overset{1}{\leftarrow} \overset{\leftarrow}{\bullet}_k := (x_j \overset{1}{\relbar\relbar} \overset{\leftarrow}{\bullet}_k) \wedge (x_j \leftarrow \overset{\leftarrow}{\bullet}_k).$$

We define the index of the first particle to strongly visit 0 by

$$A^* := \min\{k\colon 0 \leftharpoonup \overset{\leftarrow}{\bullet}_j\}.$$

Let $p_n := \mathbb{P}(A = n)$ and $p_n^* := \mathbb{P}(A^* = n)$. We define the generating functions for $0 \le t \le 1$

$$f(t) = \mathbb{E}[t^A] = \sum_{n=0}^{\infty} p_n t^n \qquad \text{and} \qquad f^*(t) = \mathbb{E}[t^{A^*}] = \sum_{n=0}^{\infty} p_n^* t^n.$$

These are related by the following formula.

Lemma 7.2.1. *For all* $0 \le t \le 1$ *it holds that*

$$f^*(t) = \frac{(1 - \frac{a}{2}) f(t)}{1 - \frac{a}{2} f(t)}.$$

Proof. We can decompose $\{A^* = n\}$ in terms of the number of weak visits w to 0 that precede the first strong visit from $\overset{\leftarrow}{\bullet}_n$:

$$p_n^* = \sum_{w=0}^{n-1} \sum_{0<\ell_1<\cdots<\ell_w<n} \mathbb{P}\left((0 \leftharpoondown \overset{\leftarrow}{\bullet}_{\ell_1}) \wedge \cdots (0 \leftharpoondown \overset{\leftarrow}{\bullet}_{\ell_w}) \wedge (0 \leftharpoonup \overset{\leftarrow}{\bullet}_n)\right)$$

This is the same as having $w + 1$ visits to 0 where the first w visits are weak and the last is strong. Formally,

$$p_n^* = \sum_{w=0}^{n-1} \left(\frac{a}{2}\right)^w \left(1 - \frac{a}{2}\right) \sum_{0<\ell_1<\cdots<\ell_w<n} p_{\ell_1} p_{\ell_2 - \ell_1} \cdots p_{\ell_w - \ell_{w-1}} p_{n-\ell_w}.$$

Exchanging the order of summation via Fubini's theorem gives

$$\begin{aligned}
f^*(t) &= \sum_{n=0}^{\infty} p_n^* t^n \\
&= \sum_{w=0}^{\infty} \left(\frac{a}{2}\right)^w \left(1 - \frac{a}{2}\right) \sum_{n=0}^{\infty} \sum_{0<\ell_1<\cdots<\ell_w<n} p_{\ell_1} t^{\ell_1} \cdots p_{\ell_w - \ell_{w-1}} t^{\ell_2 - \ell_{w-1}} p_{n-\ell_w} t^{n-\ell_w} \\
&= \sum_{w=0}^{\infty} \left(\frac{a}{2}\right)^w \left(1 - \frac{a}{2}\right) f(t)^{w+1}.
\end{aligned}$$

This is a geometric series whose closed form is the claimed formula for $f^*(t)$. □

Our main tool is the following decomposition result for $\mathbb{P}(A = n)$.

Proposition 7.2.2. *Let* $p_n := \mathbb{P}(A = n)$. *For* $n \geq 2$, *it holds that*

$$p_n = \alpha_n + \dot{\beta}_n + \hat{\beta}_n + \gamma_n + \hat{\gamma}_n + \overleftarrow{\gamma}_n \tag{7.2}$$

with

$$\alpha_n := \mathbb{P}[(A = n) \wedge \dot{\bullet}_1] = c p p_{n-1} + (1 - c) p \sum_{1<k<n} p_{k-1} p_{n-k} \tag{7.3}$$

$$\dot{\beta}_n := \mathbb{P}[(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})] = \frac{1 - c}{2} p \sum_{1<k<n} p_{k-1} p_{n-k} \tag{7.4}$$

$$\hat{\beta}_n := \mathbb{P}[(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})] = \frac{1 - c}{2} \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1} p_{k-1} p_{n-\ell} \tag{7.5}$$

$$\gamma_n := \mathbb{P}[(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})] = \sum_{1<k<n} \delta_k p_{n-k} \tag{7.6}$$

$$\begin{aligned}
\hat{\gamma}_n &:= \mathbb{P}[(A = n) \wedge (\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet})] \\
&= c \sum_{1<k<n} \hat{\delta}_k p_{n-k} + (1 - c) \sum_{1<k<\ell<n} \hat{\delta}_k p_{\ell-k} p_{n-\ell}
\end{aligned} \tag{7.7}$$

$$\overleftarrow{\gamma}_n := \mathbb{P}[(A = n) \wedge (\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet})] = \frac{a}{2} \bar{\delta}_n, \tag{7.8}$$

and

$$\begin{aligned}\delta_n^* := \mathbb{P}(\vec{\bullet}_1 \leftharpoonup \overleftarrow{\bullet}_n) &= \frac{1-p}{2} p_{n-1}^* - \sum_{1<k<n} (\dot{\beta}_k + \hat{\beta}_k) p_{n-k}^* \\ &\quad - \frac{1}{2}\left(1 - \frac{a}{2}\right)(1-c)c \sum_{1<k<n} p_{k-1} p_{n-k} \\ &\quad - \frac{1}{2}\left(1 - \frac{a}{2}\right)(1-c)c \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1} p_{k-1} p_{n-\ell}. \end{aligned} \tag{7.9}$$

Also let

$$\bar{\delta}_n := \mathbb{P}(\vec{\bullet}_1 \mathbin{—} \overleftarrow{\bullet}_n) = \frac{\delta_n^*}{1 - \frac{a}{2}} \tag{7.10}$$

$$\delta_n := \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}_n) = (1 - (a+b))\bar{\delta}_n \tag{7.11}$$

$$\hat{\delta}_n := \mathbb{P}(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_n) = b\bar{\delta}_n. \tag{7.12}$$

Proof of (7.2). This is a partitioning of the event $\{A = n\}$ based on the velocity of $\bullet_1$. We use the fact (ensured by symmetry) that $\vec{\bullet}_1$ is almost surely annihilated as observed in the proof of (6.10). □

Proof of (7.3). Conditional on $\dot{\bullet}_1$, there are precisely two manners in which $A = n$. One is that $\overleftarrow{\bullet}_n$ is the first left-moving particle to reach x_1 and the reaction $\dot{\bullet} - \overleftarrow{\bullet} \implies \overleftarrow{\bullet}$ occurs. This occurs with probability

$$\mathbb{P}(\dot{\bullet}_1)\mathbb{P}(\dot{\bullet} \mathbin{—} \overleftarrow{\bullet} \implies \overleftarrow{\bullet}) p_{n-1} = pcp_{n-1}. \tag{7.13}$$

The other manner in which $A = n$ may occur conditional on $\dot{\bullet}_1$, is if there is some $1 < k < n$ such that $\bullet_k$ is the first particle to reach x_1 from the right and a $[\dot{\bullet} \mathbin{—} \overleftarrow{\bullet} \implies \varnothing]$ reaction occurs. Then $\overleftarrow{\bullet}_n$ is the first to reach x_k from the right. This second part happens with probability $\mathbb{P}(A = k-1)\mathbb{P}(A = n-k) = p_{k-1}p_{n-k}$. So, for each k we acquire the probability

$$\mathbb{P}(\dot{\bullet}_1)\mathbb{P}(\dot{\bullet} \mathbin{—} \overleftarrow{\bullet} \implies \varnothing) p_{k-1} p_{n-k} = p(1-c) p_{k-1} p_{n-k}.$$

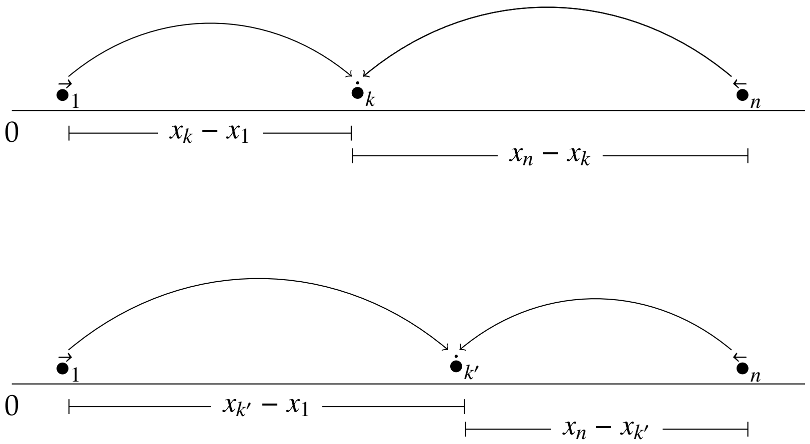


Figure 7.1: A configuration $\omega \in \{(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k)\}$ (top) and its reversal (bottom). Particles between $\bullet_1$ and $\bullet_k$ and $\bullet_k$ and $\bullet_n$ are not shown. The arced arrows indicate that the particle is the first to reach that site among all particles started under the arc. Summing over all k and k' gives complementary events. This allows us to bypass any computations involving the interdistances.

Summing over k and combining with (7.13) gives (7.3). □

Proof of (7.4). The event $\{(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k)\}$ occurs if and only if the following hold:

- $\dot{\bullet}_k$ occurs.
- The first particle to reach x_k from the left is $\vec{\bullet}_1$, which mutually annihilates with $\dot{\bullet}_k$.
- The first particle to reach x_k from the right is $\overleftarrow{\bullet}_n$.
- And, $x_k - x_1 < x_n - x_k$.

We can then write

$$\dot{\beta}_n = \sum_{1<k<n} \mathbb{P}((\dot{\bullet}_k) \wedge (\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{(0,x_k)} \wedge (x_k \overset{1}{-\!\!-} \overleftarrow{\bullet}_n)_{(x_k,\infty)} \wedge (x_k - x_1 < x_n - x_k)).$$

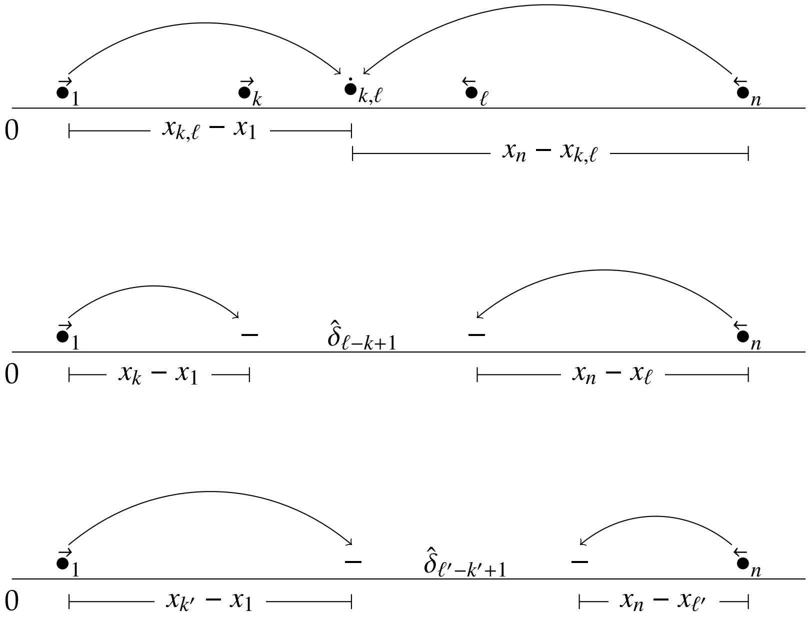


Figure 7.2: A configuration $\omega \in \{(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_{k,\ell})\}$ (top). The middle figure shows an equivalent formulation conditional on $\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_\ell$. The bottom figure shows the configuration after reversing the particles in $[x_1, x_k)$.

Given a configuration $\omega \in \{(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k)\}$ of particle locations and velocities, define $\mathrm{rev}_n(\omega)$ to be the reversed configuration. The particle at $x \in [x_1, x_n]$ corresponds to the particle in $\mathrm{rev}_n(\omega)$ with position $x_1 + (x_n - x)$ and is moving in the opposite direction. The symmetry of the parameters ensures that reversing the configuration preserves the probability: $\mathbb{P}(\omega) = \mathbb{P}(\mathrm{rev}_n(\omega))$. Since reversing maps the index k to $k' = n + 1 - k$, we may also write

$$\dot{\beta}_n := \textstyle\sum_{1<k'<n} \mathbb{P}\big((\dot{\bullet}_{k'}) \wedge (\vec{\bullet}_1 \overset{1}{\text{---}} x_{k'})_{(0,x_{k'})} \wedge (x_{k'} \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_n)_{(x_{k'},\infty)} \wedge (x_{k'} - x_1 > x_n - x_{k'})\big). \quad (7.14)$$

Since reactions are determined independently, we can swap the reaction types in (7.14) so that the probabilities

$$\mathbb{P}\big((\dot{\bullet}_{k'}) \wedge (\vec{\bullet}_1 \overset{1}{\text{---}} x_{k'})_{(0,x_{k'})} \wedge (x_{k'} \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_n)_{(x_{k'},\infty)} \wedge (x_{k'} - x_1 > x_n - x_{k'})\big)$$

and

$$\mathbb{P}\Big((\dot{\bullet}_{k'}) \wedge (\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_{k'})_{(0,x_{k'})} \wedge (x_{k'} \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_{k'},\infty)} \wedge (x_{k'} - x_1 > x_n - x_{k'})\Big)$$

are equal.

Summing the two formulas for $\dot{\beta}_n$ and combining terms with the same index partitions the comparison between $x_k - x_1$ and $x_n - x_k$. See Figure 7.1. Note that this depends crucially on the continuity of μ, which ensures that we need not worry about events like $\{x_k - x_1 = x_n - x_k\}$. Thus,

$$\begin{aligned} 2\dot{\beta}_n &= \sum_{1<k<n} \mathbb{P}((\dot{\bullet}_k) \wedge (\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{(0,x_k)} \wedge (x_k \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_k,\infty)}) \\ &= \sum_{1<k<n} p(1-c)p_{k-1}p_{n-k}. \end{aligned}$$

At the second step we apply independence. Dividing by 2 gives the claimed formula. □

Proof of (7.5). The event $\{(A = n) \wedge (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet}_{k,\ell})\}$ occurs if and only if the following hold:

- $\hat{\bullet}_{k,\ell}$ is generated from $\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_\ell$ at $x_{k,\ell} = (x_k + x_\ell)/2$ for some $1 < k < \ell < n$.
- The first particle to reach $x_{k,\ell}$ from the left of x_k is $\vec{\bullet}_1$, which mutually annihilates with $\dot{\bullet}_{k,\ell}$.
- The first particle to reach $x_{k,\ell}$ from the right of x_ℓ is $\overleftarrow{\bullet}_n$.
- And, $x_{k,\ell} - x_1 < x_n - x_{k,\ell}$.

Thus,

$$\begin{aligned} \hat{\beta}_n = \sum_{1<k<\ell<n} \mathbb{P}\big((\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{(0,x_k)} &\wedge (\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_\ell)_{[x_k,x_\ell]} \\ &\wedge (x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)} \wedge (x_{k,\ell} - x_1 < x_n - x_{k,\ell})\big). \end{aligned}$$

Let $G_{k,\ell} = (\vec{\bullet}_k \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_\ell)_{[x_k,x_\ell]}$ so that $\mathbb{P}(G_{k,\ell}) = \hat{\delta}_{\ell-k+1}$. Conditioning gives

$$\hat{\beta}_n = \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1}\mathbb{P}((\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{(0,x_k)} \wedge (x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)}$$
$$\wedge\, (x_{k,\ell} - x_1 < x_n - x_{k,\ell}) \mid G_{k,\ell}).$$

Since moving particles have unit speed, we have $x_{k,\ell} - x_1 < x_n - x_{k,\ell}$ if and only if $x_k - x_1 < x_n - x_\ell$. Using this observation and the fact that the events $(\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{(0,x_k)}$ and $(x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)}$ are independent of $G_{k,\ell}$ yields

$$\hat{\beta}_n = \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1}\mathbb{P}((\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{[x_1,x_k)} \wedge (x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)}$$
$$\wedge\, (x_k - x_1 < x_n - x_\ell)\big). \tag{7.15}$$

By reversing the configuration of particles in $[x_1, x_n]$ as with the proof of (7.4) (illustrated in Figure 7.2), we may also write

$$\hat{\beta}_n = \sum_{1<k'<\ell'<n} \hat{\delta}_{\ell'-k'+1}\mathbb{P}((\vec{\bullet}_1 \overset{1}{\relbar\!\relbar} x_{k'})_{[x_1,x_{k'})} \wedge (x_{\ell'} \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_n)_{(x_{\ell'},\infty)}$$
$$\wedge\, (x_{k'} - x_1 > x_n - x_{\ell'})\big). \tag{7.16}$$

Since reactions are determined independently, we can swap the reaction types in (7.16) as we did in the proof of (7.4). We may then rewrite (7.16) as

$$\hat{\beta}_n = \sum_{1<k'<\ell'<n} \hat{\delta}_{\ell'-k'+1}\mathbb{P}((\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_{k'})_{[x_1,x_{k'})} \wedge (x_{\ell'} \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_{\ell'},\infty)}$$
$$\wedge\, (x_{k'} - x_1 > x_n - x_{\ell'})\big). \tag{7.17}$$

Summing the two formulations of $\hat{\beta}$ at (7.15) and (7.17) removes the interval comparisons. Thus,

$$\begin{aligned} 2\hat{\beta}_n &= \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1}\mathbb{P}\big((\vec{\bullet}_1 \overset{1}{\longleftrightarrow} x_k)_{[x_1,x_k)} \wedge (x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)}\big) \\ &= (1-c) \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1} p_{k-1} p_{n-\ell}. \end{aligned}$$

Dividing by 2 gives (7.5). □

Proof of (7.6). Taking the definition of δ_n for granted, it is straightforward to see that

$$\mathbb{P}(\gamma_n) = \sum_{1<k<n} \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}_k)\mathbb{P}(x_k \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n) = \sum_{1<k<n} \delta_k p_{n-k}$$

which gives (7.6). □

Proof of (7.7). The event in the probability at (7.7) may occur in two ways. One, there exists a $1 < k < n$ such that: $(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_k) \wedge (\hat{\bullet}_{1,k} \leftarrow \overleftarrow{\bullet}_n)$ occurs. Each such event is equivalent to $(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_k) \wedge (x_k \leftarrow \overleftarrow{\bullet}_n)$, which has probability $\hat{\delta}_k c p_{n-k}$. The other manner in which the event in the probability at (7.7) may occur is if for $1 < k < \ell < n$ we have

$$(\vec{\bullet}_1 \overset{\hat{\bullet}}{\longleftrightarrow} \overleftarrow{\bullet}_k) \wedge (x_{1,k} \overset{1}{\longleftrightarrow} \overleftarrow{\bullet}_\ell) \wedge (x_\ell \overset{1}{\relbar\!\relbar} \overleftarrow{\bullet}_n)).$$

Conditional independence ensures that this event has probability $\hat{\delta}_k(1 - c)p_{\ell-k+1}p_{n-\ell}$ as claimed in the second part of (7.7). □

Proof of (7.8). The formula for $\overleftharpoon{\gamma}_n$ is the simple observation that the event in question occurs if and only if $\{\vec{\bullet}_1 \leftarrow \overleftarrow{\bullet}_n\}$, which has the claimed probability. □

Proof of (7.9), (7.10), (7.11), *and* (7.12). The main work is proving (7.9). The other three formulas follow immediately by specifying the reaction. Towards (7.9), let

$$G = \vec{\bullet}_1 \wedge (x_1 \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n)_{(x_1,\infty)}.$$

We can easily compute $\mathbb{P}(G) = \frac{1-p}{2} p^*_{n-1}$. We further claim that

$$(\vec{\bullet}_1 \leftharpoonup \overleftarrow{\bullet}_n) = G \setminus [\dot{B}_1 \cup \dot{B}_2 \cup \hat{B}_1 \cup \hat{B}_2] \tag{7.18}$$

with

$$\dot{B}_1 = \bigcup_{1<k<n} (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet})_{[x_1,\infty)} \wedge (x_1 \xleftrightarrow{1} \overleftarrow{\bullet}_k)_{[x_1,x_k]} \wedge (x_k \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n)_{(x_k,\infty)},$$

$$\dot{B}_2 = \bigcup_{1<k<n} (\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}_k)_{[x_1,\infty)} \wedge (x_k \overset{1}{\leftarrow} \overleftarrow{\bullet}_n)_{(x_k,\infty)} \wedge (x_k \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n)_{(x_k,\infty)}$$

$$\wedge\, (x_k - x_1 < x_n - x_k),$$

$$\hat{B}_1 = \bigcup_{1<k<n} (\vec{\bullet}_1 \longleftrightarrow \hat{\bullet})_{[x_1,\infty)} \wedge (x_1 \xleftrightarrow{1} \overleftarrow{\bullet}_k)_{[x_1,x_k]} \wedge (x_k \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n)_{(x_k,\infty)},$$

and

$$\hat{B}_2 = \bigcup_{1<k<\ell<n} (\vec{\bullet}_1 \xleftrightarrow{1} x_k)_{[x_1,x_k)} \wedge (\vec{\bullet}_k \xleftrightarrow{\hat{\bullet}} \overleftarrow{\bullet}_\ell)_{[x_k,x_\ell]} \wedge (x_\ell \overset{1}{\leftarrow} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)}$$

$$\wedge\, (x_\ell \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n)_{(x_\ell,\infty)} \wedge (x_k - x_1 < x_n - x_\ell).$$

To see why (7.18) holds, first note that G is necessary for $\vec{\bullet}_1 \leftharpoonup \overleftarrow{\bullet}_n$. Next, we claim that $\dot{B}_1 \cup \dot{B}_2 \cup \hat{B}_1 \cup \hat{B}_2$ contains precisely the configurations in G for which $\vec{\bullet}_1$ does not collide with $\overleftarrow{\bullet}_n$. Indeed, $\vec{\bullet}_1$ cannot collide and be destroyed by a smaller index $\overleftarrow{\bullet}$-particle, since otherwise, that smaller index $\overleftarrow{\bullet}$-particle would strongly visit x_1 before $\overleftarrow{\bullet}_n$ in the process restricted to $(x_1, x_n]$. So, the configurations from G for which $\vec{\bullet}_1$ does not collide with $\overleftarrow{\bullet}_n$ must have $\vec{\bullet}_1$ mutually annihilating with a blockade. The events in $\dot{B}_1 \cup \dot{B}_2$ describe the configurations for which $\vec{\bullet}_1 \longleftrightarrow \dot{\bullet}$ and $x_1 \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n$. See Figure 7.3. The configurations in $\hat{B}_1 \cup \hat{B}_2$ describe the configurations for which $\vec{\bullet}_1 \longleftrightarrow \hat{\bullet}$ and $x_1 \overset{1}{\leftharpoonup} \overleftarrow{\bullet}_n$.

Using independence and the definition of $\dot{\beta}_k$, it is easily seen that

$$\mathbb{P}(\dot{B}_1) = \sum_{1<k<n} \dot{\beta}_k p^*_{n-k}.$$

A similar reversal argument as in the proof of (7.4) yields

$$\mathbb{P}(\dot{B}_2) = \left(1 - \frac{a}{2}\right) \sum_{1<k<n} \frac{1}{2}(1-c)c p_{k-1} p_{n-k}.$$

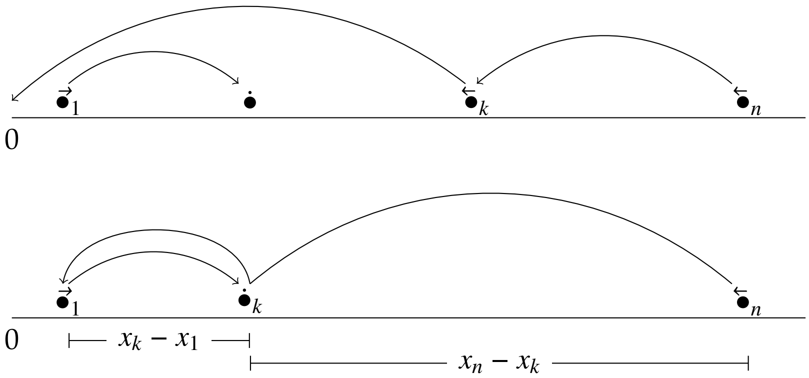


Figure 7.3: The top diagram shows a configuration in $\vec{\bullet}_1 \wedge (x_1 \overset{1}{\leftharpoondown} \overleftarrow{\bullet}_n)_{(x_1,x_n]}$ for which $\vec{\bullet}_1 — \overleftarrow{\bullet}_n$ fails to occur. Arrows indicate that the particle from the tail of the arrow is the first to visit the location at the head of the arrow. The bottom diagram shows another type of configuration in which this may occur. Note that $\overleftarrow{\bullet}_n$ survives the indicated $\dot{\bullet} — \overleftarrow{\bullet}_n$ collision.

Similarly,

$$\mathbb{P}(\hat{B}_1) = \sum_{1<k<n} \hat{\beta}_k p^*_{n-k},$$

and a reversal argument like the one used to obtain (7.5) yields

$$\mathbb{P}(\hat{B}_2) = \left(1 - \frac{a}{2}\right) \sum_{1<k<\ell<n} \frac{1}{2}(1-c)c\hat{\delta}_{\ell-k+1} p_{k-1} p_{n-\ell}.$$

Since all the individual events in $\dot{B}_1 \cup \dot{B}_2 \cup \hat{B}_1 \cup \hat{B}_2$ are disjoint, we subtract these equations from (7.18) to obtain (7.10).

□

Proof of Theorem 7.1.1. For $n = 1$, it is immediate that $p_1 = \mathbb{P}(\overleftarrow{\bullet}_1) = (1-p)/2$ and the quantities in (7.3)–(7.12) are all equal to 0. Hence, none depend on μ. Given $n \geq 2$, it follows from Proposition 7.2.2 that the quantities in (7.3)–(7.12) can be expressed solely in terms of quantities with index strictly less than n. Thus, we may proceed by induction to infer that these quantities do not depend on μ for all n. It follows then from (7.2) that p_n does not depend on μ. Since

$f(t) = \mathbb{E}[t^A] = \sum_{n\geq 1} p_n t^n$ uniquely determines the distribution of A, we obtain the first part of Theorem 7.1.1.

The implicit recursion for f is obtained by summing the generating functions corresponding to both sides of (7.2) and then applying the equations (7.3)–(7.12). This gives

$$\begin{aligned}
A(t) &:= \sum_{n\geq 0} \alpha_n t^n = cptf(t) + (1-c)ptf(t)^2 \\
B(t) &:= \sum_{n\geq 0} (\dot{\beta}_n + \hat{\beta}_n)t^n = \frac{1-c}{2} ptf(t)^2 + \frac{1-c}{2} \hat{D}(t)f(t)^2 \\
C(t) &:= \sum_{n\geq 0} (\gamma_n + \hat{\gamma}_n + \overleftarrow{\gamma}_n)t^n \\
&= D(t)f(t) + c\hat{D}(t)f(t) + (1-c)\hat{D}(t)f(t)^2 + \frac{a}{2}\bar{D}(t) \\
D^*(t) &:= \sum_{n\geq 0} \delta_n^* t^n \\
&= \frac{1-p}{2} tf^*(t) - B(t)f^*(t) - \frac{1}{2}\left(1 - \frac{a}{2}\right)(1-c)c(t + \hat{D}(t))f(t)^2 \\
\bar{D}(t) &:= \sum_{n\geq 0} \bar{\delta}_n t^n = \frac{1}{1 - \frac{a}{2}} D^*(t) \\
D(t) &:= \sum_{n\geq 0} \delta_n t^n = (1 - (a+b))\bar{D}(t) \\
\hat{D}(t) &:= \sum_{n\geq 0} \hat{\delta}_n t^n = b\bar{D}(t).
\end{aligned}$$

We may apply Lemma 7.2.1 to write the $f^*(t)$ terms in $D^*(t)$ in terms of $f(t)$. As an example of the calculations that lead to the formulas for A, B, C, D, and $\hat{D}$, we provide the derivation for summing the $\hat{\beta}_n$. First, we apply the formula at (7.5) to write

$$\sum_{n=0}^{\infty} \hat{\beta}_n t^n = \frac{1-c}{2} \sum_{n=0}^{\infty} \sum_{1<k<\ell<n} \hat{\delta}_{\ell-k+1} p_{k-1} p_{n-\ell} t^n. \tag{7.19}$$

Expanding and rearranging the sums, then applying Fubini's theorem gives

(7.19) is equal to

$$\frac{1-c}{2}\sum_{k=0}^{\infty} p_k t^k \sum_{\ell=0}^{\infty} \hat{\delta}_\ell t^\ell \sum_{n=0}^{\infty} p_n t^n = \frac{1-c}{2}\hat{D}(t) f(t)^2.$$

The other derivations are similar.

We have thus established that

$$f(t) = p_0 + p_1 t + A(t) + B(t) + C(t) \tag{7.20}$$

with $p_0 = 0$ and $p_1 = (1-p)/2$. The characterization is implicit since the formula for $\bar{D}(t)$ is also a recursive equation and may not necessarily have a solution. For the purpose of illustration, here is the recursion for the special case (proven in [HST21, Theorem 2]) of simple ballistic annihilation $a = b = c = 0$,

$$f(t) = -\frac{1}{2}ptf(t)^4 + ptf(t)^2 + \frac{1}{2}tf(t)^2 + \frac{1-p}{2}t.$$

$\square$

CHAPTER 8

NON-UNIVERSALITY IN CLUSTERED BALLISTIC ANNIHILATION

8.1 Introduction

Many of the findings in [HST21] are *universal* in the sense that the results hold for any continuous law of particle spacings. For example, $p_c = 1/4$ so long as triple collisions almost surely do not occur [BGJ19, HST21]. Note that [HST21] also proved results concerning universality when a random arrow survives a triple collision. Additional universality properties with respect to particle spacings were observed in the followup work by Haslegrave and Tournier [HT21] as well as Cruzado-Padro, Junge, and the author [CPJR22]. Broutin and Marckert discovered that a closely related bullet process with finitely many particles has a universal law governing the number of surviving particles that does not depend on velocity or spacing laws [BM20].

A canonical form of universality is invariance with respect to the average particle density. It is physically and mathematically natural to allow for clusters of superimposed particles, as is standard in other diffusion-limited annihilating systems [BL91]. To test the robustness of BA dynamics to the initial particle density, we introduce a variant of BA with random clusters of multiple blockades. We prove that the analogue of the critical value (5.2) depends on more than simply the average initial density of particles. Thus, three-velocity BA lacks this type of universality. To our knowledge, this is a new discovery that was not previously conjectured.

8.1.1 Notation

The notation used in this chapter is similar to Section 5.3 but includes some new elements.

The initial sampling of particle spacings remains unchanged from Section 5.3. Let X be a nonnegative integer-valued random variable with probability distribution $\mu = (\mu_k)_{k \geq 0}$, and let $f(t) = \mathbb{E}[t^X] = \sum_{k=0}^{\infty} \mu_k t^k$ be the probability generating function. In an abuse of notation, we will write $\mathbb{E}[\mu]$ and $\mathrm{Var}(\mu)$ for the mean and variance of X. Adopt the convention that $\mathrm{Var}(\mu) = \infty$ whenever $\mathbb{E}[\mu] = \infty$. Take $(X_n)_{n \in \mathbb{Z}}$ to be independent and μ-distributed. In clustered ballistic annihilation, each site x_n either independently starts with a *cluster* of X_n-blockades with probability $p \in [0, 1]$, or otherwise contains a single arrow whose velocity is sampled uniformly from ±1. We will sometimes refer to the starting number of blockades in a cluster as the *size* and write *k-cluster* to refer to a cluster of size k. Blockades are stationary. Left and right arrows move with velocities -1 and $+1$, respectively.

Define *μ-clustered ballistic annihilation* to have the just-described starting configuration at time 0. As time evolves, particles move at their assigned velocities. When two arrows collide, both vanish from the system. When an arrow collides with a cluster containing $k \geq 1$ remaining blockades, the arrow vanishes and one blockade is removed from the cluster (so $k - 1$ blockades remain). A more formal construction of BA that easily generalizes to include clusters can be found in [HST21].

We denote the events that a cluster starts at x_n by $\dot{\bullet}_n$, or that a left or right arrow starts at x_n by $\overleftarrow{\bullet}_n$ and $\vec{\bullet}_n$, respectively. When x_n contains a cluster, we denote the starting size with a superscript $\dot{\bullet}_n^{X_n}$. We will frequently refer to $\dot{\bullet}, \overleftarrow{\bullet}, \vec{\bullet}$ as particles. Accordingly, collision events and visits to a location $u \in \mathbb{R}$ are specified by

$$\begin{aligned}
\vec{\bullet}_m \longleftrightarrow \overleftarrow{\bullet}_n &= \{\vec{\bullet}_m \text{ and } \overleftarrow{\bullet}_n \text{ mutually annihilate}\}\\
\vec{\bullet}_m \longleftrightarrow \overleftarrow{\bullet} &= \{\vec{\bullet}_m \text{ mutually annihilates with an arrow}\}\\
\dot{\bullet}_n \leftarrow \overleftarrow{\bullet}_m &= \{\overleftarrow{\bullet}_m \text{ mutually annihilates with a blockade at } x_n\}\\
\dot{\bullet}_n^k \leftarrow \overleftarrow{\bullet}_m &= \{\overleftarrow{\bullet}_m \text{ mutually annihilates with a blockade at } x_n, X_n = k\}\\
\dot{\bullet} \leftarrow \overleftarrow{\bullet}_n &= \{\overleftarrow{\bullet}_n \text{ mutually annihilates with a blockade}\}\\
u \leftarrow \overleftarrow{\bullet} &= \{u \text{ is visited by a } \overleftarrow{\bullet}\}\\
u \overset{j}{\leftarrow} \overleftarrow{\bullet}_m &= \{\overleftarrow{\bullet}_m \text{ is the } j\text{th } \overleftarrow{\bullet} \text{ to arrive to } u\}.
\end{aligned}$$

The events $\vec{\bullet}_m \to \dot{\bullet}_n$, $\vec{\bullet}_m \to \dot{\bullet}$, $\vec{\bullet}_m \to \dot{\bullet}^k$, $\vec{\bullet} \to u$, and $\vec{\bullet} \overset{j}{\to} u$ are defined similarly.

As before, unless indicated otherwise, the default is that events are one-sided, i.e. restricted to $(0, \infty)$. So, $\mathbb{P}(0 \leftarrow \overleftarrow{\bullet}) = \mathbb{P}((0 \leftarrow \overleftarrow{\bullet})_{(0,\infty)})$.

We now define the generalization of θ from Section 5.3 for μ-clustered BA:

$$\theta = \theta(p, \mu) := \mathbb{P}((\vec{\bullet} \not\to 0)_{(-\infty,0)} \wedge (0 \not\leftarrow \overleftarrow{\bullet})_{(0,\infty)}).$$

As before, we work with the one-sided complement

$$q = q(p, \mu) := \mathbb{P}(0 \leftarrow \overleftarrow{\bullet}),$$

so that $\theta = (1 - q)^2$. Define the critical value

$$p_c = p_c(\mu) := \inf\{p \colon \theta(p, \mu) > 0\}.$$

8.1.2 Results

Our main result is a simple formula for p_c that depends on both the mean and variance of μ. We also provide an implicit formula for q.

Theorem 8.1.1. *For μ-clustered BA it holds that*

$$p_c = \frac{1}{(\mathbb{E}[\mu] + 1)^2 + \mathrm{Var}(\mu)}. \tag{8.1}$$

Moreover, q is continuous, strictly decreasing on $[p_c, 1]$, and solves

$$\frac{(1-q)^2}{(1-q^2)\, q^2 f'(q) - 2qf(q) + q^2 + 1} = p \tag{8.2}$$

with $f(q) = \sum_{k=0}^{\infty} \mu_k q^k$ the probability generating function of μ.

A surprising consequence of Theorem 8.1.1 is that there is no phase transition whenever μ has infinite variance.

Corollary 8.1.1.1. *If* $\mathrm{Var}(\mu) = \infty$*, then* $p_c = 0$.

Another corollary is that the value $p_c = 1/4$ in BA from [HST21] is maximal among all systems with $\mathbb{E}[\mu] = 1$.

Corollary 8.1.1.2. *If $\mu_1 = 1$, then $p_c = 1/4$. For all other μ with $\mathbb{E}[\mu] = 1$, we have $p_c < 1/4$.*

Lastly, in Chapter 6, we studied a coalescing version of ballistic annihilation in which particles sometimes survive collisions. The primary interest was determining the analogue of p_c for these systems. However, we were unable to analyze the case in which blockades survive each collision with some fixed probability, see Remark 6.2. This is equivalent to μ-clustered ballistic annihilation with μ a geometric distribution. Thus, Theorem 8.1.1 gives the value of p_c in this unsolved case.

Corollary 8.1.1.3. *Let* $\beta \in (0,1)$ *and* μ *be a geometric distribution with parameter* β*, i.e.,* $\mu_k = (1-\beta)^{k-1}\beta$ *for* $k \geq 1$*. For* μ*-clustered ballistic annihilation it holds that*

$$p_c = \frac{\beta^2}{\beta^2 + \beta + 2}$$

and, by solving (8.2) *for* q*, we have for* $p \geq p_c$

$$q(p) = \frac{\sqrt{p^2\beta - p^2 - p\beta + 2p} - p\beta + \beta - 1}{p\beta^2 - p\beta + p - \beta^2 + 2\beta - 1}.$$

8.1.3 Discussion

There is no robust general theory that tells us whether or not a given interacting particle system will have a universal phase transition. On $\mathbb{Z}^d$, branching processes, diffusion-limited-annihilating systems, and activated random walk are processes known to have phase transitions that do not depend on the initial particle density [AN04, CRS18, RSZ19]. The frog model and directed parking processes on d-ary trees have phase transitions that depend on more than the average density [CH19, BBJ21, JR19]. Furthermore, the number of visits to a distinguished site varies monotonically with the concentration of the initial particle placements for these processes on general families of graphs [JJ18, BBJJ22].

Our result is of a similar character as the main result of Curien and Hénard from [CH19]. They studied the parking process on critical Galton-Watson trees whose offspring distribution has variance Σ^2. A parking spot is placed at each vertex of such a tree and a random number of cars with mean m and variance σ^2 arrive independently at the vertices. Cars drive towards the root and park in the first spot they encounter. Only one car is allowed per spot, so this reaction can be viewed as ballistic particles (cars) mutually annihilating with stationary blockades (spots). [CH19, Theorem 1] established that the expected number of cars to reach the root is infinite if and only if $(1-m)^2 - \Sigma^2(\sigma^2 + m^2 - 1) < 0$. Like our result, the phase behavior depends (in a surprisingly simple way) on both the mean and variance of the initial configuration. Note that a more complicated formula that depends on the entire distribution of X was observed for parking on binary trees in [ACCH23].

It is a priori unclear whether or not p_c depends on more than $\mathbb{E}[\mu]$. On one hand, BA has dynamics similar to the systems considered in [BBJJ22]. So it is reasonable to expect some sensitivity to the initial density of particles. On the other hand, the mean-field heuristic presented in [DRFP95] and further clarified in [KRL95, Section (b)] suggests that p_c might be universal. The explanation in [KRL95] assumes that arrow–arrow collisions, on average, occur at twice the rate of blockade–left arrow collisions. This is, in their words, "based on the expectation that the relative number of annihilation events is proportional to the relative velocities of the collision partners." If this "expectation", which seems to only depend on the relative velocities of particle types, still holds in μ-clustered BA, then the same heuristic would predict universality.

Theorem 8.1.1 settles the question. Put concisely, the more volatile μ becomes, the more space for arrow-arrow collisions, which enhances blockade survival. In a loose sense, our theorem says that the order particles are placed plays a role in determining p_c. A more detailed heuristic for why the variance plays a role in the formula for p_c in Theorem 8.1.1 comes from considering the extreme case in which $\mu_0 = (k-1)/k$ and $\mu_k = 1/k$ with k a large integer. As $\mathrm{Var}(\mu) = k-1$ and $\mathbb{E}[\mu] = 1$, Theorem 8.1.1 implies that $p_c = 1/(k+3)$. To see intuitively why this is the correct order, suppose that 0 contains a k-cluster. Let

x_N be the next site to the right of 0 that contains a k-cluster. We have that N is a geometric random variable with parameter p/k. Thus, we expect on the order of $(1-p)k/p$ arrows in $(0, x_N)$ along with some 0-clusters. The number of arrows that reach the boundary of $(0, x_N)$ should be comparable to the magnitude of the discrepancy between left and right arrows started in $(0, x_N)$. By the central limit theorem, the discrepancy is on the order of $\sqrt{k/p}$, and so this order of left arrows from $(0, x_N)$ will reach 0 [EF85]. For these arrows to eliminate a significant portion of the k-cluster at 0, we would need $k \approx \sqrt{k/p}$, or equivalently, $p \approx 1/k \sim 1/\operatorname{Var}(\mu)$.

8.1.4 Proof overview

Our proof has three main parts. Section 8.2 is devoted to proving the recursive equation for q in Proposition 8.2.1. This is inspired by what was done in [HST21], but instead uses a version of the mass transport principle first observed in [JL22] and refined in [BJL+23] (Chapter 6). The basic idea is to partition the event associated to q based on the velocity of $\bullet_1$.

An important probability for deriving this recursion is $s_k = \mathbb{P}((0 \leftarrow \overset{\hookleftarrow}{\bullet}) \wedge (\vec{\bullet}_1 \to \dot{\bullet}^k))$. In [HST21], it was observed that $s_1 = (1/2)pq^2$. Computing s_k for $k > 1$ in the proof of (8.4) is more involved. After applying the mass transport principle, this event partitions into various events in which $k+1$ left arrows arrive to 0 while satisfying non-symmetric spacing requirements. Remarkably, a broader symmetry than what was used in [HST21] (see (8.13)) makes this case tractable and yields the simple formula $s_k = (1/2)p\mu_k kq^{k+1}$. In the proof of (8.5), we use similar methods to give a relatively simple formula for the companion probability $r_k = \mathbb{P}((0 \not\leftarrow \overset{\curvearrowleft}{\bullet}) \wedge (\vec{\bullet}_1 \rightarrow \dot{\bullet}^k))$. With these quantities in hand, it is straightforward to obtain (8.3). The second part is proving that q is continuous in p. The proof closely follows the argument that θ is continuous in asymmetric three-velocity ballistic annihilation from [JL22].

The last step, in Section 8.4, involves analyzing the recursion from Proposition 8.2.1. The recursion implies that $0 = (1-q)h(p,q)$ for an explicit function h. This tells us that either $q = 1$ or solves $h(p,q) = 0$. We prove that $h(u,1)$ has the unique solution $u = p_c$ from Theorem 8.1.1. The goal is then to show that q, for any μ-clustered BA, continuously switches from being identically 1 for $p \le p_c$ to the unique curve determined by (8.2). A priori, it is not obvious how to prove that the roots of h are well-behaved and that q faithfully follows them. The continuity of q observed in Theorem 8.3.6 is crucial for ruling out the pathology that q jumps between 1 and solutions to $h = 0$. This is more robust than past approaches [HST21, BJL+23].

8.2 Recursion

The goal of this section is to prove the following recursive formula.

Proposition 8.2.1.

$$q = \frac{1-p}{2} + pqf(q) + s + q\left(\frac{1-p}{2} - s - r\right) \tag{8.3}$$

with

$$s := \mathbb{P}((0 \leftarrow \overset{\curvearrowleft}{\bullet}) \wedge (\vec{\bullet}_1 \rightarrow \dot{\bullet})) = \frac{pq^2}{2} f'(q) \tag{8.4}$$

$$r := \mathbb{P}((0 \leftarrow \overset{\curvearrowleft}{\bullet}) \wedge (\vec{\bullet}_1 \rightarrow \dot{\bullet})) = \frac{pq\left(q^2 f'(q) - qf'(q) - f(q) + 1\right)}{1-q}. \tag{8.5}$$

Proof of (8.3). We partition q in terms of the velocity of the first particle

$$q = \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \overleftarrow{\bullet}_1) + \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \dot{\bullet}_1) + \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \vec{\bullet}_1) \tag{8.6}$$

and will provide a formula for each summand. It is immediate that

$$\mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \overleftarrow{\bullet}_1) = \frac{1-p}{2}. \tag{8.7}$$

For the second summand, we further partition on the size of $\dot{\bullet}_1$ to write

$$\mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \dot{\bullet}) = \sum_{k=0}^{\infty} \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \dot{\bullet}_1 \wedge (X_1 = k)).$$

If $X_1 = k$, then $k+1$ left arrows must arrive at x_1 in order for 0 to be visited. This happens if and only if the jth left arrow to arrive reaches the starting location of the $(j-1)$th left arrow to arrive for $j = 2, \dots, k+1$. By similar reasoning as [HST21, Lemma 7], each of these arrivals is conditionally independent and has probability q. Thus, for $k \geq 0$ we have

$$\mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \dot{\bullet}_1 \wedge (X_1 = k)) = p \cdot \mu_k q^{k+1}.$$

Summing over k gives

$$\mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \dot{\bullet}_1) = pqf(q). \tag{8.8}$$

A similar argument as [ST17, Lemma 3.3] implies that all arrows are eventually annihilated. Since $\mathbb{P}(\vec{\bullet}_1 \to \dot{\bullet}) = s + r$, we may write

$$\begin{aligned} \mathbb{P}(\vec{\bullet}_1) = \frac{1-p}{2} &= \mathbb{P}(\vec{\bullet}_1 \to \dot{\bullet}) + \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) \\ &= s + r + \mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}). \end{aligned} \tag{8.9}$$

For 0 to be visited on the event $\{\vec{\bullet}_1\}$, the particle $\vec{\bullet}_1$ must first be annihilated. We partition on the collision type:

$$\begin{aligned}\mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge \vec{\bullet}_1) &= \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge (\vec{\bullet}_1 \to \dot{\bullet})) + \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge (\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet})) \\ &= s + q\mathbb{P}(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) \qquad (8.10) \\ &= s + q\left(\frac{1-p}{2} - s - r\right). \qquad (8.11)\end{aligned}$$

The equality at (8.10) follows from the definition of s and the fact that $\mathbb{P}(0 \leftarrow \overleftarrow{\bullet} \mid \vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}) = q$. This fact follows from the observation that conditional on $(\vec{\bullet}_1 \longleftrightarrow \overleftarrow{\bullet}_j)$ for some $j > 1$, $(0 \leftarrow \overleftarrow{\bullet})$ occurs if and only if $(x_j \leftarrow \overleftarrow{\bullet})_{(x_j,\infty)}$, which has probability q. The move to (8.11) then uses (8.9). Combining (8.7), (8.8), and (8.11) in (8.6) gives (8.3). □

Next, we will prove the formulas for s and r at (8.4) and (8.5), respectively.

Proof of (8.4). Let $s_k = \mathbb{P}((0 \leftarrow \overleftarrow{\bullet}) \wedge (\vec{\bullet}_1 \to \dot{\bullet}^k))$ so that $s = \sum_{k=0}^{\infty} s_k$. We will use the mass transport principle to relate the event associated to s_k to one that involves $k+1$ arrows arriving to the site containing a k-cluster. To this end, define

$$Z_k^j(a,b) = \sum_{c \in \mathbb{Z}} \Big[\mathbf{1}\{\dot{\bullet}_b^k \wedge (\vec{\bullet}_a \xrightarrow{j} x_b)_{[x_a,x_b)} \wedge (x_b \xleftarrow{k+1-j} \overleftarrow{\bullet}_c)_{(x_b,x_c]} \wedge (x_b - x_a < x_c - x_b)\}\Big]$$

for $a, b, j, k \in \mathbb{Z}$.

Observe that

$$s_k^j := \mathbb{P}((\vec{\bullet}_1 \xrightarrow{j} \dot{\bullet}^k) \wedge (0 \leftarrow \overleftarrow{\bullet})) = \mathbb{E} \sum_{b \in \mathbb{Z}} Z_k^j(1,b).$$

Define $\vec{D}_j$ to be the starting distance from x_1 of the jth particle to arrive to x_1 in the process restricted to particles in $(-\infty, x_1)$. We set $\vec{D}_j = \infty$ whenever fewer than j particles ever visit x_1. Define $\overleftarrow{D}_j$ similarly, but on (x_1, ∞). By MTP and independence, s_k^j is equal to

$$\begin{aligned}\mathbb{E}\textstyle\sum_{a\in\mathbb{Z}} Z_k^j(a,1) &= \mathbb{P}(\dot{\bullet}_1^k)\mathbb{P}((\vec{\bullet} \xrightarrow{j} x_1)_{(-\infty,x_1)})\mathbb{P}((x_1 \xleftarrow{k+1-j} \overleftarrow{\bullet})_{(x_1,\infty)})\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}) \\ &= p\cdot\mu_k q^j q^{k+1-j}\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}). \qquad (8.12)\end{aligned}$$

Since $s_k = \sum_{j=1}^k s_k^j$, (8.12) gives

$$s_k = p\cdot\mu_k q^{k+1}\sum_{j=1}^{k}\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}).$$

If k is even, then grouping summands gives

$$\sum_{j=1}^{k}\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}) = \sum_{j=1}^{k/2}\Big[\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}) + \mathbb{P}(\vec{D}_{k+1-j} < \overleftarrow{D}_j)\Big] = \frac{k}{2}. \qquad (8.13)$$

We have $\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}) + \mathbb{P}(\vec{D}_{k+1-j} < \overleftarrow{D}_j) = 1$. This is because $\vec{D}_j$, $\overleftarrow{D}_{k+1-j}$, $\vec{D}_{k+1-j}$ and $\overleftarrow{D}_j$ are continuous and identically distributed random variables. Moreover, the $\vec{D}_j$ and $\overleftarrow{D}_{k+1-j}$ as well as $\vec{D}_{k+1-j}$ and $\overleftarrow{D}_j$ are pairwise independent. Using similar reasoning, if $k = 2m+1$ is odd, then we can write $\sum_{j=1}^k \mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j})$ as

$$\sum_{j=1}^{m}\Big[\mathbb{P}(\vec{D}_j < \overleftarrow{D}_{k+1-j}) + \mathbb{P}(\vec{D}_{k+1-j} < \overleftarrow{D}_j)\Big] + \mathbb{P}(\vec{D}_{m+1} < \overleftarrow{D}_{m+1}),$$

which equals $m + (1/2) = k/2$. Hence, $s_k = p\cdot\mu_k q^{k+1}(k/2)$. Summing gives

$$s = \sum_{k=1}^{\infty} s_k = \frac{pq^2}{2}\sum_{k=0}\mu_k k q^{k-1} = \frac{pq^2}{2}f'(q).$$

□

Proof of (8.5). Let $r_k = \mathbb{P}((0 \not\leftarrow \overleftarrow{\bullet}) \wedge (\vec{\bullet}_1 \to \dot{\bullet}^k))$ so that $r = \sum_{k=0}^{\infty} r_k$. As we did for the proof of (8.4), we apply (MTP) with new indicators

$$W_k^{i,j}(a,b) = \textstyle\sum_{c\in\mathbb{Z}} \mathbf{1}\{\dot{\bullet}_b^k \wedge (\vec{\bullet}_a \xrightarrow{i} x_b)_{[x_a,x_b)} \wedge (x_b \xleftarrow{j} \overleftarrow{\bullet}_c)_{(x_b,x_c]} \wedge (x_c \not\leftarrow \overleftarrow{\bullet})_{(x_c,\infty)}\}$$

for $i, j, k, a, b \in \mathbb{Z}$. Let $(\vec{\bullet}_1 \xrightarrow{i} \dot{\bullet}^k \xleftarrow{j^*} \overleftarrow{\bullet})$ denote the event that $\vec{\bullet}_1$ is the ith right arrow to annihilate with a k-cluster and *exactly* j left arrows visit that same k-cluster. Observe that for $i + j \leq k$ with $i \neq 0$, we have

$$r_k^{i,j} := \mathbb{P}((0 \not\leftarrow \overleftarrow{\bullet}) \wedge (\vec{\bullet}_1 \xrightarrow{i} \dot{\bullet}^k \xleftarrow{j^*} \overleftarrow{\bullet})) = \mathbb{E} \sum_{b \in \mathbb{Z}} W_k^{i,j}(1, b).$$

By (MTP) and independence,

$$\begin{aligned} r_k^{i,j} &= \mathbb{E} \sum_{a \in \mathbb{Z}} W_k^{i,j}(a, 1) \\ &= \mathbb{P}(\dot{\bullet}_1^k)\mathbb{P}((\vec{\bullet} \xrightarrow{i} x_1)_{(-\infty, x_1)})\mathbb{P}((x_1 \xleftarrow{j} \overleftarrow{\bullet})_{(x_1, \infty)}) \times \mathbb{P}(\overleftarrow{D}_{j+1} = \infty \mid \overleftarrow{D}_j < \infty) \\ &= p \cdot \mu_k q^i q^j (1 - q). \end{aligned}$$

We then have

$$r_k = \sum_{i=1}^{k} \sum_{j=0}^{k-i} r_k^{i,j} = p \cdot \mu_k \sum_{i=1}^{k} \sum_{j=0}^{k-i} q^{i+j}(1 - q).$$

Applying the formula $\sum_{i=0}^{m} a^i = (1 - a^{m+1})/(1 - a)$ twice, gives

$$r_k = p \cdot \mu_k \frac{q\left(kq^{k+1} - kq^k - q^k + 1\right)}{1 - q}.$$

Hence,

$$r = \sum_{k=1}^{\infty} r_k = \frac{pq\big(q^2 f'(q) - qf'(q) - f(q) + 1\big)}{1 - q}.$$

□

8.3 Continuity

The goal of this section is to prove that q is continuous in p by proving that it is both upper and lower semi-continuous. We begin by recalling these definitions and stating a few classical facts. A function φ is *upper semi-continuous* (USC) at each $p_0 \in [0, 1]$ if and only if $\limsup_{p \to p_0} \varphi(p) \leq \varphi(p_0)$. It is *lower semi-continuous*

(LSC) at each $p_0 \in [0,1]$ if and only if it holds that $\liminf_{p \to p_0} \varphi(p) \geq \varphi(p_0)$. Rather than working directly with the definition, we will apply the following properties. See [JL22] for proofs.

Fact 8.3.1. The following hold.

(a) φ is continuous if and only if φ is USC and LSC.

(b) If there exists a sequence of LSC functions φ_n with $\varphi_n \uparrow \varphi$, then φ is LSC.

(c) If $\varphi(p) = \sup_n(\varphi_n(p))$ with φ_n LSC, then φ is LSC.

(d) If φ_1 and φ_2 are LSC, then $\max(\varphi_1, \varphi_2)$ is LSC.

(e) φ is LSC if and only if $-\varphi$ is USC.

(f) If ψ is continuous and φ is LSC, then $\psi \circ \varphi$ is LSC. Similarly, if φ is USC, then $\psi \circ \varphi$ is USC.

(g) If φ and ψ are both LSC or USC, then so is $\varphi + \psi$.

That q is LSC follows almost immediately from its definition.

Proposition 8.3.2. *q is LSC for $p \in [0,1]$.*

Proof. The events $Q_n = \{(0 \leftarrow \overleftarrow{\bullet})_{(0,x_n)}\}$ involve finitely many particles. After conditioning on the velocities of these particles and integrating over all possible spacings, $\mathbb{P}(Q_n)$ is a finite degree polynomial in p, and thus continuous. Moreover, $Q_n \subseteq Q_{n+1}$, thus the $\mathbb{P}(Q_n)$ are increasing in n. Since $q = \lim_{n \to \infty} \mathbb{P}(Q_n)$, it follows from Fact 8.3.1 (b) that q is LSC. $\square$

We next aim to prove that q is USC. This is more difficult and involves an indirect characterization of $\theta = (1-q)^2$ that takes a supremum over functionals of configurations with only finitely many particles. Let $\dot{N}(j,k)$ be the number of blockades that survive in ballistic annihilation restricted to the particles in $[x_j, x_k]$. Similarly, let $\overleftarrow{N}(j,k)$ and $\overrightarrow{N}(j,k)$ count the number of surviving left and right arrows. Define the random variables that track the difference between the number of surviving blockades and arrows in the process restricted to only the particles in $[x_j, x_k]$:

$$W(j,k) = \dot{N}(j,k) - \overleftarrow{N}(j,k) - \overrightarrow{N}(j,k).$$

Lemma 8.3.3. *$n^{-1}\mathbb{E}_p[W(1,n)]$ is continuous in p for all $n \geq 1$.*

Proof. The random variables $W(1,n)$ involve only finitely many particles. Thus, $\mathbb{E}_p[W(1,n)]$ is a finite degree polynomial in p and is continuous. □

Lemma 8.3.4. *$\theta = \max\left(0, \sup_{n\geq 1} n^{-1}\mathbb{E}_p[W(1,n)]\right)$ for all $p \in [0,1]$.*

Proof. The proof has four steps. Fortunately, it requires little modification from the blueprint developed in [JL22]. We explain the basic idea of each step and refer the reader to the appropriate reference.

Step 1. For all integers $j < k < \ell$ it holds that $W(j,\ell) \geq W(j,k) + W(k+1,\ell)$.

Proof. This superadditivity property is proven in Lemma 6.4.3 for a more general variant of ballistic annihilation in which particles sometime survive collisions. The basic idea is that surviving arrows from the restrictions to $[x_j, x_k]$ and $[x_{k+1}, x_\ell]$ have a non-decreasing effect on $W(j,\ell)$. Surviving arrows either destroy other surviving arrows, which augments $W(j,\ell)$. Or, surviving arrows destroy blockades, which may cause a chain reaction, but, regardless, the effect is worst-case neutral on $W(j,\ell)$. The argument does not change if multiple blockades are present at a site. □

Step 2. $\lim_{k\to\infty} k^{-1}\overleftarrow{N}(1,k) = 0 = \lim_{k\to\infty} k^{-1}\overrightarrow{N}(1,k)$.

Proof. This is proven in [JL22, Proposition 12] for asymmetric ballistic annihilation. It is much simpler to deduce for symmetric systems. Using translation invariance of the velocity configuration, Birkhoff's Ergodic Theorem gives that the limits equal the probability an arrow is never annihilated. [ST17, Lemma 3.3] observes that this quantity must be zero, as otherwise, Birkhoff's Ergodic Theorem gives the contradiction that there is a positive densities of surviving left and right arrows. This reasoning still applies with the possibility of multiple blockades at a single site. □

Step 3. Let $N_{\mathbb{R}}(1,k)$ denote the number of blockades that survive in $[x_1, x_k]$ in ballistic annihilation with all particles in $\mathbb{R}$ present. If $\theta > 0$, then

$$\lim_{k\to\infty} k^{-1}\dot{N}(1,k) = \theta = \lim_{k\to\infty} k^{-1}N_{\mathbb{R}}(1,k).$$

Proof. This is proven in [JL22, Proposition 12]. It follows from the definition of θ and the strong law of large numbers that $\lim_{k\to\infty} k^{-1}N_{\mathbb{R}}(1,k) = \theta$. So, it suffices to prove that

$$\lim_{k\to\infty} k^{-1}[\dot{N}(1,k) - \dot{N}_{\mathbb{R}}(1,k)] = 0.$$

First, observe that blockade survival is a decreasing event as the interval of restriction is expanded. So, $\dot{N}(1,k) - \dot{N}_{\mathbb{R}}(1,k) \geq 0$. From there, the main idea is that at most a geometric random variable with parameter q, call it R_k, of the surviving blockades in ballistic annihilation restricted to $[x_1, x_k]$ are removed from right arrows entering at x_1, and the same for an independent and identically distributed geometric random variable of left arrows entering at x_k, call it L_k. So, $\dot{N}(1,k) - N_{\mathbb{R}}(1,k) \leq L_k + R_k$ Since these random variables have exponential tails and constant parameter q, it is easy to infer from the Borel-Cantelli lemma that $\lim_{k\to\infty} k^{-1}[R_k + L_k] = 0$ almost surely. ☐

Step 4. Let $\theta_0 := \max\left(0, \sup_{k\geq 1} k^{-1}\mathbb{E}_p[W(1,k)]\right)$. It holds that $\theta = \theta_0$.

Proof. The proof is similar to [JL22, Lemma 10]. First, we prove that $\theta \leq \theta_0$. Combining Steps 2 and 3 and Fatou's lemma gives

$$\theta = \lim_{k\to\infty} k^{-1}W(1,k) = \mathbb{E}_p\left[\liminf_{k\to\infty} k^{-1}W(1,k)\right] \leq \liminf_{k\to\infty} k^{-1}\mathbb{E}_p[W(1,k)] \leq \theta_0.$$

Next, we show that $\theta \geq \theta_0$. This is immediate when $\theta_0 = 0$, so suppose that $\theta_0 > 0$. Then, there is an integer k with $\mathbb{E}_p[W(1,k)] > 0$. Letting $K_m = km$ for $m \geq 0$, we see that $S_n := \sum_{m=0}^{n-1} W(K_m + 1, K_{m+1})$ is a random walk with positive drift. The law of large numbers gives that $S_n > 0$ for all $n \geq 1$ with positive probability. Step 1 implies that

$$W(1, K_n) \geq S_n \qquad \forall n \geq 1. \tag{8.14}$$

This is enough to deduce that 0 is never visited with positive probability, which gives $\theta > 0$. See the proof of [JL22, Lemma 10] for more details.

We will use this framework to prove that $\theta > \delta$ for arbitrary $\delta \in (0,1)$ with $\delta < \theta_0$. Let $k \geq 1$ be such that $k^{-1}\mathbb{E}_p[W(1,k)] > \delta$. Steps 2 and 3 imply that $\theta = \lim_{n\to\infty} \frac{1}{n}W(1,n)$. Multiplying by n/n, applying (8.14) and then the strong law of large numbers gives

$$\theta = \liminf_{n\to\infty} \frac{n}{K_n}\frac{1}{n}W(1,K_n) \geq \liminf_{n\to\infty} \frac{n}{K_n}\frac{1}{n}S_n = k^{-1}\mathbb{E}_p[W(1,k)] > \delta$$

as desired. ☐

☐

Proposition 8.3.5. *q is USC for $p \in [0, 1]$.*

Proof. It follows that θ is LSC from Lemmas 8.3.3 and 8.3.4 along with Facts 8.3.1 (c) and (d). Since $\theta = (1-q)^2$, we have $q = 1 - \sqrt{\theta}$. Facts 8.3.1 (e) and (f) imply that $-\sqrt{\theta}$ is USC. Since 1 is USC, q can be expressed as the sum of two USC functions and by Fact 8.3.1 (g) is USC. □

Theorem 8.3.6. *q is continuous for $p \in [0, 1]$.*

Proof. This follows immediately from Propositions 8.3.2 and 8.3.5 along with Fact 8.3.1 (a). □

8.4 Proof of Theorem 8.1.1

Proof of Theorem 8.1.1. Subtracting q from both sides of (8.3) in Proposition 8.2.1 gives $0 = g(p, q)$ with $g\colon [0, 1]^2 \to \mathbb{R}$ defined as

$$g(u, v) := \frac{u\left(1 - v^2\right) v^2 f'(v) + 2uvf(v) - uv^2 + u + v^2 - 2v + 1}{2(1 - v)}. \tag{8.15}$$

Let $h(u, v) = g(u, v)/(1 - v)$ so that Proposition 8.2.1 implies

$$0 = (1 - q)h(p, q). \tag{8.16}$$

The goal is to show that (p, q) solves $1 - v = 0$ for $p \le p_c$ and transitions to solving $h(u, v) = 0$ for $p \ge p_c$.

Inspecting (8.15), we see that $h(u, v)$ is linear in u. Solving $h(u, v) = 0$ yields

$$u = \frac{(1 - v)^2}{(1 - v^2)\, v^2 f'(v) - 2vf(v) + v^2 + 1} =: F(v).$$

Thus,

Fact 8.4.1. If $h(u, v) = 0$, then $u = F(v)$.

Using L'Hospital's rule twice and basic generating function properties

$$\lim_{v\to 1} F(v) = \frac{1}{f(1) + 3f'(1) + f''(1)} = \frac{1}{(1 + \mathbb{E}[X])^2 + \mathrm{Var}(X)} =: p_*.$$

By Fact 8.4.1,

Fact 8.4.2. $(u, v) = (p_*, 1)$ is the unique solution to $1 - v = 0 = h(u, v)$.

Since q is continuous (Theorem 8.3.6) with $q(1) = 0$, it follows from (8.16) and Fact 8.4.2 that $(p_*, 1)$ is the only point at which (p, q) can continuously transition from solving $1 - v = 0$ to solving $h(u, v) = 0$. So,

Fact 8.4.3. If $p \geq p_*$, then $h(p, q(p)) = 0$.

Combining Facts 8.4.1 and 8.4.3 gives

Fact 8.4.4. $p = F(q(p))$ for $p \geq p_*$.

Fact 8.4.4 says that F is a left inverse of q on the domain $p \geq p_*$ i.e., if $q(p) = y$ for $p \geq p_*$, then $F(y) = p$. It is an elementary exercise in analysis that this and the continuity of q imply that

Fact 8.4.5. q is continuous and strictly decreasing for $p \geq p_*$.

Facts 8.4.4 and 8.4.5 (along with Theorem 8.3.6) imply (8.2) in Theorem 8.1.1.

It remains to prove that $p_c = p_*$ as claimed at (8.1). Suppose that $p > p_*$. Fact 8.4.3 implies that $h(p, q(p)) = 0$. Fact 8.4.2 ensures that $q(p_*) = 1$. Fact 8.4.1 requires that $q(p) \neq 1$. Since $q(p)$ is a probability, we then have $q(p) < 1$. So, $p_c \leq p_*$.

To see the reverse inequality, suppose that there exists $p_0 < p_*$ with $q(p_0) = v < 1$. Fact 8.4.5 and $q(1) = 0$ imply that $q\colon [p_*, 1] \to [0, 1]$ is a continuous bijection. Thus, there is $p_1 > p_*$ with $q(p_1) = v$. As $v < 1$, (8.16) implies that $h(p_0, v) = 0 = h(p_1, v)$. This contradicts Fact 8.4.1, which requires that $p_0 = p_1 = F(v)$. So, $q = 1$ for all $p \geq p_*$. Thus, $p_c \geq p_*$. $\square$

BIBLIOGRAPHY

[AB99] M. Aizenman and A. Burchard, *Hölder regularity and dimension bounds for random curves*, Duke Math. J. **99** (1999), 419–453.

[ACCH23] D. Aldous, A. Contat, N. Curien, and O. Hénard, *Parking on the infinite binary tree*, Probab. Theory Relat. Fields (2023), 1–24.

[ADA99] M. Aizenman, B. Duplantier, and A. Aharony, *Path-crossing exponents and the external perimeter in 2d percolation*, Phys. Rev. Lett. **83** (1999), 1359–1362.

[AN84] M. Aizenman and C. M. Newman, *Tree graph inequalities and critical behavior in percolation models*, J. Stat. Phys. **36** (1984), no. 1-2, 107–143.

[AN04] K. B. Athreya and P. E. Ney, *Branching processes*, Courier Corporation, 2004.

[AO82] S. Alexander and R. Orbach, *Density of states on fractals:fractons*, J. Physique Lett. **43** (1982), no. 17, 625–631.

[AP96] P. Antal and A. Pisztora, *On the chemical distance for supercritical Bernoulli percolation*, Ann. Probab. **24** (1996), 1036–1048.

[Arr83] R. Arratia, *Site recurrence for annihilating random walks on Z_d*, Ann. Probab. (1983), 706–713.

[BBJ21] R. Bahl, P. Barnet, and M. Junge, *Parking on supercritical Galton-Watson trees*, ALEA **18** (2021).

[BBJJ22] R. Bahl, P. Barnet, T. Johnson, and M. Junge, *Diffusion-limited annihilating systems and the increasing convex order*, Electron. J. Probab. **27** (2022), 1 – 19.

[BDC12] V. Beffara and H. Duminil-Copin, *The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$*, Probab. Theory Relat. Fields **153** (2012), no. 3, 511–542.

[BEK00] R. A. Blythe, M. R. Evans, and Y. Kafri, *Stochastic ballistic annihilation and coalescence*, Phys. Rev. Lett. **85** (2000), no. 18, 3750.

[BGJ19] D. Burdinski, S. Gupta, and M. Junge, *The upper threshold in ballistic annihilation*, ALEA **16** (2019), 1077–1087.

[BJL$^+$23] L. Benitez, M. Junge, H. Lyu, M. Redman, and L. Reeves, *Three-velocity coalescing ballistic annihilation*, Electron. J. Probab. **28** (2023), no. 56, 1–18.

[BL91] M. Bramson and J. L. Lebowitz, *Asymptotic behavior of densities for two-particle annihilating random walks*, J. Stat. Phys. **62** (1991), no. 1-2, 297–372.

[BM20] N. Broutin and J.-F. Marckert, *The combinatorics of the colliding bullets*, Random Struct. Algorithms **56** (2020), no. 2, 401–431.

[BN11] V. Beffara and P. Nolin, *On monochromatic arm exponents for 2d critical percolation*, Ann. Probab. **39** (2011), no. 4, 1286–1304.

[BR06] B. Bollobás and O. Riordan, *Percolation*, Cambridge University Press, 2006.

[CDCH16] D. Chelkak, H. Duminil-Copin, and C. Hongler, *Crossing probabilities in topological rectangles for the critical planar fk-ising model*, Electron. J. Probab. **21** (2016), 28 pp.

[CH19] N. Curien and O. Hénard, *The phase transition for parking on Galton–Watson trees*, arXiv e-prints (2019).

[CPJR22] D. Cruzado Padró, M. Junge, and L. Reeves, *Arrivals are universal in coalescing ballistic annihilation*, arXiv e-prints (2022).

[CPY90] G. F. Carnevale, Y. Pomeau, and W. R. Young, *Statistics of ballistic agglomeration*, Phys. Rev. Lett. **64** (1990), no. 24, 2913–2916.

[CRS18] M. Cabezas, L. T. Rolla, and V. Sidoravicius, *Recurrence and density decay for diffusion-limited annihilating systems*, Probability Theory and Related Fields **170** (2018), no. 3, 587–615.

[DC13] H. Duminil-Copin, *Parafermionic observables and their applications to planar statistical physics models*, Ensaios Matematicos **25** (2013), 1–371.

[DC17] ———, *Lectures on the Ising and Potts models on the hypercubic lattice*, arXiv e-prints (2017).

[DCGH$^+$21] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion, *Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$*, Ann. Sci. Éc. Norm. Supér. **54** (2021), no. 6, 1363–1413.

[DCMT21] H. Duminil-Copin, I. Manolescu, and V. Tassion, *Planar random-cluster model: fractal properties of the critical phase*, Probab. Theory Relat. Fields **181** (2021), no. 1-3, 401–449.

[DCST17] H. Duminil-Copin, V. Sidoravicius, and V. Tassion, *Continuity of the phase transition for planar random-cluster and Potts models with $1 \le q \le 4$*, Comm. Math. Phys. **349** (2017), 47–107.

[DHS16] M. Damron, J. Hanson, and P. Sosoe, *On the chemical distance in critical percolation II*, arXiv e-prints (2016).

[DHS17] ———, *On the chemical distance in critical percolation*, Electron. J. Probab. **22** (2017), 1–43.

[DHS21] ———, *Strict inequality for the chemical distance exponent in two-dimensional critical percolation*, Commun. Pure Appl. Math **74** (2021), no. 4, 679–743.

[DKJ$^+$19] B. Dygert, C. Kinzel, M. Junge, A. Raymond, E. Slivken, J. Zhu, et al., *The bullet problem with discrete speeds*, Electron. Commun. Probab. **24** (2019).

[DRFP95] M. Droz, P.-A. Rey, L. Frachebourg, and J. Piasecki, *Ballistic-annihilation kinetics for a multivelocity one-dimensional ideal gas*, Phys. Rev. E **51** (1995), no. 6, 5541–5548 (eng).

[DS11] M. Damron and A. Sapozhnikov, *Outlets of 2d invasion percolation and multiple-armed incipient infinite clusters*, Probab. Theory Relat. Fields **150** (2011), 257–294.

[EF85] Y. Elskens and H. L. Frisch, *Annihilation kinetics in the one-dimensional ideal gas*, Phys. Rev. A **31** (1985), no. 6, 3812.

[EK85] B. F. Edwards and A. R. Kerstein, *Is there a lower critical dimension for chemical distance?*, J. Phys. A Math. Gen. **18** (1985), no. 17, L1081–L1086.

[ES88] R. G. Edwards and A. D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988), 2009–2012.

[FK72] C. M. Fortuin and P. W. Kasteleyn, *On the random-cluster model: I. introduction and relation to other models*, Physica **57** (1972), no. 4, 536–564.

[FvdH17a] R. Fitzner and R. van der Hofstad, *Generalized approach to the non-backtracking lace expansion*, Probab. Theory Relat. Fields **169** (2017), no. 3, 1041–1119.

[FvdH17b] ———, *Mean-field behavior for nearest-neighbor percolation in* $d >$ 10, Electron. J. Probab. **22** (2017), 1–65.

[GM90] G. R. Grimmett and J. M. Marstrand, *The supercritical phase of percolation is well behaved*, Proc. Roy. Soc. London Ser. A **430** (1990), 439–457.

[Gra99] P. Grassberger, *Pair connectedness and shortest-path scaling in critical percolation*, J. Phys. A Math. Gen. **32** (1999), no. 35, 6233–6238.

[Gri06] G. Grimmett, *The random-cluster model*, Grundlehren der mathematischen Wissenschaften, vol. 333, Springer-Verlag, 2006.

[Har60] T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Math. Proc. Cambridge Philos. Soc. **56** (1960), no. 1, 13–20.

[HHS84] H. J. Herrmann, D. C. Hong, and H. E. Stanley, *Backbone and elastic backbone of percolation clusters obtained by the new method of 'burning'*, J. Phys. A Math. Gen. **17** (1984), no. 5, L261–L266.

[HN84] S. Havlin and R. Nossal, *Topological properties of percolation clusters*, J. Phys. A Math. Gen. **17** (1984), no. 8, L427–L432.

[HS88] H. J. Herrmann and H. E. Stanley, *The fractal dimension of the mini-*

mum path in two- and three-dimensional percolation, J. Phys. A Math. Gen. **21** (1988), no. 17, L829–L833.

[HS90] T. Hara and G. Slade, *Mean-field critical behaviour for percolation in high dimensions*, Commun. Math. Phys. **128** (1990), 333–391.

[HST21] J. Haslegrave, V. Sidoravicius, and L. Tournier, *Three-speed ballistic annihilation: phase transition and universality*, Selecta Mathematica **27** (2021), no. 84.

[HSvdH03] T. Hara, G. Slade, and R. van der Hofstad, *Critical two-point functions and the lace expansion for spread-out high-dimensional percolation and related models*, Ann. Probab. **31** (2003), no. 1, 349–408.

[HT21] J. Haslegrave and L. Tournier, *Combinatorial universality in three-speed ballistic annihilation*, In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius, vol. 77, Springer International Publishing, 2021, pp. 487–517.

[HTWBA85] S. Havlin, B. Trus, G. H. Weiss, and D. Ben-Avraham, *The chemical distance distribution in percolation clusters*, J. Phys. A Math. Gen. **18** (1985), no. 5, L247–L249.

[HvdH17] M. Heydenreich and R. van der Hofstad, *Progress in high-dimensional percolation and random graphs*, Springer, 2017.

[JJ18] T. Johnson and M. Junge, *Stochastic orders and the frog model*, Ann. Inst. H. Poincaré Probab. Statist. **54** (2018), no. 2, 1013–1030.

[JL22] M. Junge and H. Lyu, *The phase structure of asymmetric ballistic annihilation*, Ann. Appl. Probab. **32** (2022), no. 5, 3797–3816.

[JOSMRRS22] M. Junge, A. Ortiz San Miguel, L. Reeves, and C. Rivera Sánchez, *Non-universality in clustered ballistic annihilation*, arXiv e-prints (2022).

[JR19] T. Johnson and L. T. Rolla, *Sensitivity of the frog model to initial conditions*, Electron. Commun. Probab. **24** (2019), 1–9.

[Kes80] H. Kesten, *The critical probability of bond percolation on the square lattice equals* $1/2$, Commun. Math. Phys. **74** (1980), no. 1, 41–59.

[Kes86] ______, *The incipient infinite cluster in two-dimensional percolation*, Probab. Theory Relat. Fields **73** (1986), 369–394.

[Kes87] ______, *Scaling relations for 2d-percolation*, Comm. Math. Phys. **109** (1987), no. 1, 109–156.

[KN09] G. Kozma and A. Nachmias, *The Alexander-Orbach conjecture holds in high dimensions*, Inventiones mathematicae **178** (2009), no. 3, 635–654.

[KRL95] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Ballistic annihilation kinetics: The case of discrete velocity distributions*, Phys. Rev. E **51** (1995), no. 5, 3977.

[KSZ98] H. Kesten, V. Sidoravicius, and Y. Zhang, *Almost All Words Are Seen In Critical Site Percolation On The Triangular Lattice*, Electron. J. Probab. **3** (1998), 1 – 75.

[KW] M. Kleber and D. Wilson, *"Ponder This" IBM research challenge*, `https://research.ibm.com/haifa/ponderthis/challenges/May2014.html`, May 2014.

[KZ93] H. Kesten and Y. Zhang, *The tortuosity of occupied crossings of a box in critical percolation*, J. Stat. Phys. **70** (1993), no. 3, 599–611.

[LSW02] G. Lawler, O. Schramm, and W. Werner, *One-arm exponent for critical 2d percolation*, Electron. J. Probab. **7** (2002), 1–13.

[MZ05] G. J. Morrow and Y. Zhang, *The sizes of the pioneering, lowest crossing and pivotal sites in critical percolation on the triangular lattice*, Ann. Appl. Probab. **15** (2005), no. 3, 1832–1886.

[Nol08] P. Nolin, *Near-critical percolation in two dimensions*, Electron. J. Probab. **13** (2008), 1562–1623.

[Pei36] R. Peierls, *On ising's model of ferromagnetism*, Math. Proc. Cambridge Philos. Soc. **32** (1936), no. 3, 477–481.

[Rei00] D. Reimer, *Proof of the van den berg-kesten conjecture*, Comb. Probab. Comput. **9** (2000), no. 1, 27–32.

[RS20a] G. Ray and Y. Spinka, *A short proof of the discontinuity of phase transition in the planar random-cluster model with $q > 4$*, Comm. Math. Phys. **378** (2020), 1977–1988.

[RS20b] L. Reeves and P. Sosoe, *Equivalence of polychromatic arm probabilities on the square lattice*, arXiv e-prints (2020).

[RSZ19] L. T. Rolla, V. Sidoravicius, and O. Zindy, *Universality and sharpness in activated random walks*, Ann. Henri Poincaré **20** (2019), no. 6, 1823–1835.

[Rus78] L. Russo, *A note on percolation*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **43** (1978), no. 1, 39–48.

[Sch11] O. Schramm, *Conformally invariant scaling limits: an overview and a collection of problems*, pp. 1161–1191, Springer New York, New York, NY, 2011.

[Smi01] S. Smirnov, *Critical percolation in the plane: Conformal invariance, Cardy's formula, scaling limits*, C. R. Acad. Sci. Paris Sér. I Math. **333** (2001), 239–244.

[SR22] P. Sosoe and L. Reeves, *An estimate for the radial chemical distance in 2d critical percolation clusters*, Stoch. Process. Their Appl. **147** (2022), 145–174.

[ST17] V. Sidoravicius and L. Tournier, *Note on a one-dimensional system of annihilating particles*, Electron. Commun. Probab. **22** (2017), 9 pp.

[SW78] P. D. Seymour and D. J. A. Welsh, *Percolation probabilities on the square lattice*, Ann. Discrete Math. **3** (1978), 227–245.

[SW01] S. Smirnov and W. Werner, *Critical exponents for two-dimensional percolation*, Math. Res. Lett. **8** (2001), 729–744.

[TEW98] B. Toth, A. Ermakov, and W. Werner, *On some annihilating and coalescing systems*, J. Stat. Phys. **91** (1998), no. 5-6, 845–870.

[TW83] Doug Toussaint and Frank Wilczek, *Particle–antiparticle annihilation in diffusive motion*, Chem. Phys. **78** (1983), no. 5, 2642–2647.

[vdBG13] J. van den Berg and A. Gandolfi, *Bk-type inequalities and generalized random-cluster representations*, Probab. Theory Relat. Fields **157** (2013), 157–181.

[ZYDZ12] Z. Zhou, J. Yang, Y. Deng, and R. M. Ziff, *Shortest-path fractal dimension for percolation in two and three dimensions*, Phys. Rev. E **86** (2012), 061101.