On the collective behaviors of bio-inspired distributed systems

Francesco d'Amore

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Sur les Comportements Collectifs de Systèmes Distribués Bio-Inspirés

Francesco D'AMORE

Université Côte d'Azur, Inria, CNRS, I3S, France

Présentée en vue de l'obtention du grade de docteur en Informatique **d'**Université Côte d'Azur

Dirigée par : Nicolas NISSE, Chargé de Recherche, Inria, France
Co-encadrée par : Emanuele NATALE, Chargé de Recherche, CNRS, France
Soutenue le : 17 Octobre 2022

Devant le jury, composé de :
Pierre FRAIGNIAUD, Directeur de Recherche, CNRS, France
Robert ELSÄSSER, Full Professor, University of Salzburg, Austria
Frederik MALLMANN-TRENN, Assistant Professor, King's College London, United Kingdom
Alain JEAN-MARIE, Directeur de Recherche, Inria, France

SUR LES COMPORTEMENTS COLLECTIFS DE SYSTÈMES DISTRIBUÉS BIO-INSPIRÉS

On the Collective Behaviors of Bio-Inspired Distributed Systems

Francesco D'AMORE

Jury :

Rapporteurs

Pierre FRAIGNIAUD, Directeur de Recherche, CNRS, France
Robert ELSÄSSER, Full Professor, University of Salzburg, Austria
Frederik MALLMANN-TRENN, Assistant Professor, King's College London, United Kingdom

Examinateurs

Alain JEAN-MARIE, Directeur de Recherche, Inria, France

Directeur de thèse

Nicolas NISSE, Chargé de Recherche, Inria, France

Co-encadrant de thèse

Emanuele NATALE, Chargé de Recherche, CNRS, France

Université Côte d'Azur

Francesco D'AMORE

Sur les Comportements Collectifs de Systèmes Distribués Bio-Inspirés

xv+152 p.

Ce document a été préparé avec LaTeX2e et la classe `these-ISSS` version v. 2.10.

Impression : sommaire.tex – 5/12/2022 – 16:46

Révision pour la classe : these-ISSS.cls,v 2.10 2020/06/24 14:16:37 mpelleau

To the people who believed in me and taught me to believe.

Sur les Comportements Collectifs de Systèmes Distribués Bio-Inspirés

Résumé

Récemment, la communauté algorithmique a manifesté un intérêt croissant pour l'utilisation de ses outils théoriques à la compréhension des systèmes complexes, notamment biologiques, tels que les colonies d'insectes, les volées d'oiseaux et les réseaux de neurones. Nous contribuons à l'étude de ces systèmes dans trois directions différentes. Premièrement, nous analysons des dynamiques computationnelles pour les tâches de coordination stochastique dans les systèmes multi-agents. En particulier, nous nous focalisons sur le problème du consensus dans des environnements où la communication est bruyante : nous analysons deux dynamiques d'opinion, les dynamiques UNDECIDED-STATE et 3-MAJORITY, et nous prouvons qu'elles présentent une transition de phase à des seuils de bruit différents. En dessous du seuil, ces dynamiques atteignent rapidement une phase métastable de quasi-consensus ; au-dessus, aucune forme de consensus n'est possible. Deuxièmement, nous étudions les Lévy walks, des marches aléatoires qui modélisent des schémas de mouvement trouvés dans la nature, dont la distribution de la longueur de pas suit une loi de puissance. Nous analysons leur temps d'arrêt (hitting time) parallèle et les utilisons pour concevoir un algorithme optimal pour l'ANTS problem, un problème de recherche distribuée sur $\mathbb{Z}^2$ qui capture certains aspects de la théorie du butinage. Troisièmement, nous considérons l'Assembly Calculus, un modèle distribué du cerveau récemment proposé, qui consiste en des neurones et des synapses stylisés, et nous testons expérimentalement ses capacités, largement inexplorées, en mettant en œuvre des heuristiques connues pour la tâche de planification du monde des blocs. Nous montrons empiriquement que des programmes grands et complexes dans ce modèle s'exécutent correctement et de manière fiable.

Mots-clés : Calcul distribué, Algorithmes naturels, Systèmes biologiques.

On the Collective Behaviors of Bio-Inspired Distributed Systems

Abstract

In recent years there has been a surge of interest on behalf of the algorithmic community in applying its theoretical tools to the understanding of complex systems, in particular biological ones, such as insect colonies, flocks of birds, and networks of neurons. We contribute to the investigation of such systems in three different directions. First, we analyze computational dynamics for stochastic coordination tasks in multi-agent systems: in particular, we focus on the consensus problem in environments where communication is affected by some form of noise. In this setting, we analyze two known opinion dynamics, the UNDECIDED-STATE and the 3-MAJORITY dynamics, and prove that they exhibit a phase-transition at different noise thresholds. Below the threshold, the two dynamics quickly reach an almost-consensus metastable phase; above, no form of consensus is possible. Second, we study Lévy walks, i.e., special random walks known to model many movement patterns found in nature, characterized by a step-length density distribution proportional to a power-law. We analyze their parallel hitting time and show how to use them to design an almost optimal algorithm for the ANTS problem, a distributed search problem on $\mathbb{Z}^2$ which captures some aspects of animal foraging theory. Third, we consider the Assembly Calculus, a recently proposed distributed computational model of the brain which consists of stylized spiking neurons and synapses, and we test experimentally its capabilities, largely unexplored. In particular, we implement known heuristics for the Blocks World planning task; we empirically prove that reasonably large and complex programs in the Assembly Calculus run correctly and reliably.

Keywords: Distributed computing, Natural algorithms, Biological systems.

Acknowledgments

My doctoral journey owes its development and maturation to many people who have accompanied me over the past three years. Quoting the acknowledgments of a doctoral colleague of mine, thanking all the people that have helped me in these years is a very difficult task, and I apologize in advance if I forget anyone.

First, I want to thank my supervisors: Emanuele and Nicolas. Emanuele was a source of constant stimuli that challenged my intellect not only in mathematical research per se (which is already a lot), but also in making me ask questions about what research is worth pursuing. I cannot say that I have achieved the goal of only investigating what is worthy; however, I have developed an inner voice that plays the role of devil's advocate, repeatedly questioning me about the depth and the foresight of my work. I hope this voice can guide me through my future research.

Although I did not collaborate scientifically with Nicolas, I am truly grateful to him: starting with the willingness to be my supervisor, to the care with which he followed all formal aspects of the progress of my Ph.D. journey, he was always available to help me.

Second, I want to thank Alain Jean-Marie, Pierre Fraigniaud, Robert Elsässer, and Frederik Mallmann-Trenn, for agreeing to serve on the jury for my doctoral defense. This important step in my professional career was possible thanks to their availability. In addition, I thank Pierre, Robert, and Frederik for taking the time to review my dissertation manuscript and the valuable comments they sent me. Since the reviews of submitted papers are circumstantial to specific works, these are the first real evaluations of my contribution as a researcher.

I also want to thank the COATI team, which welcomed me three years ago when I first joined it as a newbie, and provided me with a confortable and attractive working environment. I cannot mention each of the COATIs because there are too many who are either part of the team now or have been part of it in the past, but I want to mention David and Patricia, the team leader and the team assistant, who have built (along with everybody else) this "team spirit." I also want to mention those who have become friends, with whom I shared experiences, thoughts, concerns, and fun outside the workplace: Foivos, Hicham, Gosia, Arthur, Igor, Aurora, Amadeus, Paulo, Damien, Ali.

Special thanks goes to all the coauthors of my papers, who helped me improve my research skills in a broad sense: from mathematical insights into difficult problems, to expository clarity of scientific writing. Special mention goes to Andrea Clementi, George Giakkoupis, and Pierluigi Crescenzi who, among the others, have been a source of academic inspiration for me.

Facing this path with serenity was also made possible by all the friendships built within the Inria and I3S institutes. I want to mention Sara, Giulia R., Sardor, Edoardo, Vladimir, Igor, Giulia M., Gianluigi, Riccardo, Lorenzo. A special thanks goes to Michelangelo who, besides being a doctoral colleague and a friend, is also my housemate, and helped bring joy and amusement to the home environment. I won't forget the present and past members of ADSTIC, with whom we shared the beautiful experience of organizing events for other (unfortunate and lost) young Ph.D. students: Jules, Nina, Romain, Amélie, Florian, Laetitia L., Laetitia G., Rémy. Last but not least, I also mention Emilie, Felix, and Yann, with whom I shared some outside activities such as climbing in Arkose, or ADSTIC social events.

The next part of these acknowledgments is written in Italian, as it concerns only Italian people. Non posso non ringraziare il gruppo di amici (più o meno indipendente dal lavoro) costruito qui a Nizza in questi anni, con cui ho condiviso la vita quotidiana durante e dopo i vari confinamenti: Pupi, Isa, Matto Ferranti, Matteo Frigo, Stefano, Paolo, Alessandra. Grazie per tutte le esperienze vissute assieme, dalle passeggiate agli aperitivi. Un ringraziamento particolare va a Massimo, Franco e Giovanni, con i quali forse ho condiviso ancor di più la vita e le passioni, tra la montagna, il cinema, l'arrampicata e il semplice chiacchierare.

Un paragrafo a sé va dedicato a Costantino, mio amico di infanzia, che ha vissuto con me a Nizza quest'ultimo anno come un vero fratello, tra cene, serie TV, escursioni, sport e chiacchiere: è stato un vero sostegno.

Seppure più a distanza e meno quotidianamente, vivendo luoghi ed esperienze differenti, i miei amici degli anni universitari hanno sempre fatto sentire la loro presenza e il loro affetto: Andrea, Maria Sofia, Federico, Primula, Guglielmo, Davide, Benedetta, Alessandro C. Non da meno, i miei amici di una vita, Alessandro L. e Angelo. Termino questo paragrafo ringraziando due "nuove" amiche, che sono entrate nella mia vita in quanto coinquiline della mia ragazza, Stephanie e Sabrina: in questi anni, ci siamo tenuti compagnia e ci siamo stati di sostegno a lungo durante le varie zone verdi, arancioni e rosse, condividendo anche viaggi e avventure.

Un ruolo speciale ricoprono, ormai da anni, i Gesuiti che sono amici miei e della mia famiglia: Pietro, Andrea, Diego, Claudio, Renato, Giovanni. Con sfumature e ruoli differenti, ognuno mi aiuta e ha aiutato a crescere.

È più difficile invece ringraziare la mia famiglia, mamma, papà e Valeria, per l'enorme sforzo che tutti noi siamo stati chiamati a vivere in questi anni. Siamo rimasti vicini, unendoci ancor di più in modo inaspettato di fronte alle avversità di questi anni. La mia famiglia è ed è stata di enorme sostegno in tutte le mie scelte. Ringrazio anche i miei zii, zio Mino, zia Lella, zio Paolo, per la vicinanza che sempre mi hanno fatto sentire, e mia nonna, Consolina, per la sua instancabile presenza e operosità. Non da meno sono stati gli zii più lontani, zia Teta, Federica, e Fabio, sempre presenti nel bisogno e non. Voglio menzionare anche i miei cugini "allargati": Monica, Alessandro, il nuovo arrivato Danielino, Michele, Linda, e i piccoli Davide e Daniele.

Come anche feci per i ringraziamenti della tesi di laurea magistrale, dedico l'ultimo paragrafo alla mia dolce metà, Sabrina. Ha dipinto e colorato questi anni come solo lei sa fare. Prendendomi per mano, mi ha accompagnato passo per passo in tutte le situazioni che siamo stati chiamati a vivere, alcune semplici, altre meno; ogni giorno ricevo il dono del suo sostegno e del suo amore. Grazie.

Preface

This thesis presents a major part of the research I conducted during my doctoral studies. During these years, I mainly studied bio-inspired distributed systems and their collective behaviors, both theoretically and experimentally. The purpose of this preface is to briefly list the main topics covered in this thesis and the work that has not been included.

— Chapter 1 presents an overview of the interplay between biology and distributed computing, and then introduces the topics I investigate in the thesis and discusses related work.

— Chapter 2 provides the analysis of two opinion dynamics, the UNDECIDED-STATE and the 3-MAJORITY dynamics, in the presence of a communication noise feature. The results are based on the papers [D'Amore et al., 2020], [D'Amore and Ziccardi, 2022], and [D'Amore et al., 2022a].

— Chapter 3 is dedicated to the analysis of the hitting time of the so-called Lévy walks, special random walks whose step-length density distribution follows a power law, and shows how to use them to obtain an optimal solution for a distributed search problem. These results are presented in the paper [Clementi et al., 2021].

— Chapter 4 investigates the Assembly Calculus, a distributed model of the brain, and presents the results of simulations testing its computational capabilities. Such experiments have been presented in [D'Amore et al., 2022b].

In addition to this, during my doctoral studies, I also investigated other topics that are outside the scope of this thesis. The most important one, which led to two currently submitted works, concerns a randomized version of the subset sum problem in its one-dimensional [Carvalho Walraven da Cunha et al., 2022] and multidimensional [Becchetti et al., 2022] formulations.

Contents

CHAPTER 1

Introduction

Algorithmic research has classically focused on the design of efficient algorithms to solve well defined problems in computer science, often arising from engineering and technology or, more in general, from the theory of computation. However, in recent years, there has been a surge of interest on behalf of the algorithmic community in applying its theoretical tools — the "computational lens" [Karp, 2011] — to the understanding of complex systems. The study of complex systems lies in an interdisciplinary field of research and aims at explaining how systems of relatively simple entities organize themselves, into a collective whole that creates patterns, uses information, evolves and learns in some cases, and achieves a common task without relying on any central controller. The word complex comes from the Latin root plectere: to weave, entwine [Mitchell and Toroczkai, 2010]. Complex systems are composed of many simple parts that are entwined and interact giving rise to complex, emergent behaviors, the latter being the objects of investigation of disciplines such as physics, biology, chemistry, and, more recently, computer science.

Different complex systems in nature (understood in a broad sense), such as insect colonies [Musco et al., 2017], flocks of birds [Chazelle, 2009], immune systems [Harris et al., 2012], networks of neurons [Afek et al., 2011,Dasgupta et al., 2017,Papadimitriou et al., 2020], and social networks [Mossel and Tamuz, 2017], benefit from and propose many challenges to computational science [Feinerman and Korman, 2013]. The algorithmic processes that arise from the modeling of complex systems have been called *natural algorithms* [Chazelle, 2012], and have been an hot topic from which both computer science and biology have profited. On the one hand, the algorithmic perspective — or computational lens — has helped catching behavioral properties of biological systems (e.g., estimation of the bird flocking convergence time [Chazelle, 2009], proof that the slime mold is able to compute shortest paths [Bonifaci et al., 2012]); on the other hand, biological systems have helped designing new algorithms for well-known problems (e.g., design of a distributed maximal independent set algorithm from the fly's nervous system [Afek et al., 2011]).

Biological systems are often distributed in nature, which is why the main field of computer science for which this line of research is particularly fruitful is distributed computing: its advance in the understanding of topics such as dynamic networks, mobile agents, population protocols, and network computing in general indicates that this field has reached the maturity level of being useful also for the context of understanding large biological systems [Feinerman and Korman, 2013].

With respect to this, many important results related to natural algorithms have already been established other than the already mentioned ones: a distributed model of how ants could estimate their density in a bounded space has been proposed and studied [Musco et al., 2017]; some toy models on how biological systems can spread information and reach agreement despite the presence of communication noise have been analyzed [Feinerman et al., 2017]; a problem on multi-agent systems that have to locate nearby treasures as fast as possible (the ANTS problem) has been for-

malized and investigated, with links to foraging theory [Feinerman et al., 2012, Feinerman and Korman, 2017]; it has been discovered that the fly's olfactory circuit solves the similarity search problem using a novel variant of a traditional computer science algorithm (called locality-sensitive hashing) [Dasgupta et al., 2017]; a distributed computational model of the brain based on assemblies of neurons has been designed [Papadimitriou et al., 2020]; etc.

In this thesis we investigate and contribute to three main lines of research which fall under the broad natural algorithms field. The first one (Section 1.1 and Chapter 2) is about opinion dynamics for quickly reaching agreement in multi-agent systems under a communication noise feature inspired by the model proposed in [Feinerman et al., 2017]. Opinion dynamics are a class of randomized protocols executed by systems of agents with limited power (bounded memory and little communication capacities) which rely on simple, lightweight, elementary rules; they're often used to achieve tasks such as broadcast [Fraigniaud and Natale, 2019] or consensus [Becchetti et al., 2020a]. The second one (Section 1.2 and Chapter 3) studies a specific type of random walks, the Lévy walks, which have been established to model biological organisms movement patterns [Reynolds, 2018]. We uses the properties of such walks to give an optimal solution to the Ants Nearby Treasure Search (ANTS) problem [Feinerman and Korman, 2017]. The third and final one (Section 1.3 and Chapter 4) investigates experimentally the computational capabilities of the Assembly Calculus [Papadimitriou et al., 2020], a computational distributed model of the brain which consists only in stylized spiking neurons and synapses, by implementing known heuristics for the Blocks World planning task [Gupta and Nau, 1992].

Notation For us to continue, we need to introduce the notation used in the thesis: we make use of the conventional Bachmann–Landau notation for asymptotic behaviors of function, which we now recall. Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be any two functions. We write $f(x) = \mathcal{O}(g(x))$ if a constant $M > 0$ and a value $x_0 \in \mathbb{R}$ exist such that $|f(x)| \leq M|g(x)|$ for any $x > x_0$. Similarly, we write $f(x) = \Omega(g(x))$ if a constant $m > 0$ and a value $x_0 \in \mathbb{R}$ exist such that $|f(x)| \geq m|g(x)|$ for any $x > x_0$; finally, we write $f(x) = \Theta(g(x))$ if two constants $0 < m < M$ and a value $x_0 \in \mathbb{R}$ exist such that $m|g(x)| \leq |f(x)| \leq M|g(x)|$ for all $x > x_0$. Moreover, we write $f(x) = o(g(x))$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$, and $f(x) = \omega(g(x))$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| = \infty$. We also mention the polylog function. By writing $f(x) = \operatorname{polylog} x$, we mean that a constant $m > 0$ exists such that $f(x) = \Theta(\log^m x)$. We complete this paragraph by introducing the notion of high probability. An event E depending on a parameter n holds with high probability (in short, w.h.p.) with respect to n if a constant $c > 0$ exists such that $\Pr[E] \geq 1 - n^{-c}$. In general, we are interested in stating theoretical results that hold with high probability with respect to some input parameter.

1.1 Consensus dynamics with uniform communication noise

The consensus problem. The consensus problem is a fundamental problem in distributed computing in which we have a system of agents supporting some opinions that interact between each other by exchanging messages over some communication graph, with the goal of reaching an agreement on some *valid* opinion (i.e., an opinion initially present in the system) [Becchetti et al., 2020a]. In particular, many research papers focus on the majority consensus problem where the goal is to converge towards the initial majority opinion. The numerous theoretical studies in this area are justified by many different application scenarios, ranging from social networks [Mossel et al., 2014, Acemoglu et al., 2013], swarm robotics [Bayindir, 2016], cloud computing, communi-

cation networks [Ruan and Mostofi, 2008], and distributed databases [Dietzfelbinger et al., 2010], to biological systems [Feinerman et al., 2017, Fraigniaud and Natale, 2019]. As for the latter, the goal of the majority consensus problem is to model some real-world scenarios where biological entities need to communicate and agree in order to pursue some collective task. Many biological entities in different real situations perform this type of process, e.g., molecules [Carroll, 2004], bacteria [Bassler, 2002], flock of birds [Ben-Shahar et al., 2014], school of fish [Sumpter et al., 2008], or social insects [Franks et al., 2002], such as honeybees [Reina et al., 2017].

In the aforementioned applicative scenarios, communication among agents is often affected by some form of noise: that is why one of the main goal in network information theory is to guarantee reliable communications in noisy networks [Gamal and Kim, 2011]. In this context, error-correcting codes are very effective methods to reduce communication errors in computer systems [Moon, 2005, Koetter and Kschischang, 2008], and this is why many theoretical studies of the (majority) consensus problem assume that communication between entities occurs without error, and instead consider some adversarial behavior (e.g., byzantine fault [Becchetti et al., 2016]). Despite their effectiveness in computer applications, error-correcting codes are quite useless if we want to model consensus in biological systems. Indeed, they involve sending complicated codes through communication links, and it is reasonable to assume that biological type entities communicate between each other in a simpler way. For this reason, in recent years many works have been focusing on the study of the (majority) consensus problem where the communication between entities is unreliable and subjected to some form noise [Feinerman et al., 2017, Fraigniaud and Natale, 2019, Cruciani et al., 2021].

Opinion dynamics. Opinion dynamics are mathematical models to investigate the way a decentralized multi-agent system is able to reach some form of consensus. Their study is a hot topic touching several research areas such as distributed computing [Ghaffari and Lengler, 2018, Cruciani et al., 2019, Becchetti et al., 2020a], social networks [Acemoglu et al., 2013, Mossel and Tamuz, 2017], and system biology [Cardelli and Csikàsz-Nagy, 2012, Boczkowski et al., 2018b]. Typical examples of opinion dynamics are the VOTER model, the averaging rules, and the majority rules. Some of such dynamics share a surprising efficiency and resiliency that exploit common computational principles, as they rely on simple, lightweight, local, elementary rules. A definition of dynamics, as given in [Becchetti et al., 2020a], follows.

Definition 1.1.1 (Dynamics)**.** A dynamics is a synchronous distributed algorithm characterized by a very simple structure. In particular, the state of an agent at round $t \in \mathbb{N}$ only depends on its state and a symmetric function [1] of the vector of states of its neighbors, while the update rule is the same for every communication graph and for every agent, and it does not change over time.

The first consensus dynamics that have been studied in the presence of noise communication are linear opinion dynamics, i.e., where the state-update function is linear in the neighbors states, such as the VOTER dynamics and the AVERAGING dynamics: in the VOTER dynamics, an agent samples a neighbor u.a.r. and simply adopts its opinion; in the AVERAGING dynamics, an agent computes and adopts the *average* opinion of its neighbors (the set of opinions must be ordered). In particular, they have been studied in the presence of uniform noise communication [Lin et al., 2007] or in the presence of some communities of *stubborn agents* (i.e., agents that never change

1. A function f is symmetric if its value is the same no matter the order of its arguments.

opinion) [Mobilia, 2003, Mobilia et al., 2007, Yildiz et al., 2013]. Due to the communication unreliability, only *metastable* forms of consensus can be achieved, where a large subset of the agents agree on some opinion while other opinions remain supported by smaller subsets of agents, and this setting lasts for a relatively-long time. However, the VOTER model has a slow convergence time even in fully connected networks and a large initial bias towards some majority opinion [Hassin and Peleg, 2001], and the AVERAGING dynamics requires agents to perform non-trivial computation and, more importantly, to have large local memory. For these reasons, linear opinion dynamics struggle explaining the observed metastable consensus in multi-agent systems [Boczkowski et al., 2019, Condon et al., 2019, Emanuele Natale, 2017], and many research papers have begun to investigate new, more plausible, non-linear opinion dynamics.

In this thesis, we consider a network of n agents supporting opinions from a finite set of opinions Σ, and we study the processes arising from the synchronous version of the UNDECIDED-STATE and the 3-MAJORITY dynamics with uniform communication noise. In the UNDECIDED-STATE dynamics there is an extra opinion, the undecided state, other than those in Σ, and works as follows: at any round, every agent samples a neighbor uniformly at random (u.a.r.) and pulls its opinion. If the opinion is different from the agent's one, the agent becomes undecided. An undecided agent, instead, just applies the VOTER rule and adopts any pulled opinion; we remark that for an agent to change its opinion in Σ, the UNDECIDED-STATE dynamics needs at least two rounds, as it must pass through the undecided state. Instead, in the 3-MAJORITY dynamics, at any round, every agent samples u.a.r. three neighbors (with replacement) and pulls their opinions; then, it adopts the majority opinion among the pulled ones, if there is any, while ties are broken u.a.r. The noise model we consider is inspired from [Feinerman et al., 2017] and is defined as follows: let $p \in [0, 1]$ be any real value. For each pulled opinion, with probability $1 - p$ the opinion keeps intact; instead, with probability p it is sampled u.a.r. among all possible ones.

In the following section, we state our main contributions, which are based on the works [D'Amore et al., 2020], [D'Amore et al., 2022a], and [D'Amore and Ziccardi, 2022].

1.1.1 Our contribution

We consider the synchronous version of the dynamics in the binary opinion case over the fully connected network of n agents. Even though the complete graph is a strong assumption for such communication networks, we remark that, at every round, an agent pulls an opinion from a constant number of neighbors: therefore, the round-by-round communication pattern results is a dynamic graph with $\mathcal{O}(n)$ edges. Furthermore, such a model can be used to capture the behavior of bio-inspired multi-agent systems in which mobile agents meet randomly at a relatively high rate.

In the aforementioned setting, we prove that the processes induced by the UNDECIDED-STATE and the 3-MAJORITY dynamics exhibits a phase-transition. At any time $t \geq 0$, let s_t be the bias of the system, i.e., the difference between the majority opinion community size and the minority opinion one at time t. Our results are summarized in the following theorems.

Theorem 1.1.1 (Phase-transition of the U-dyn)**.** *Let $\{s_t\}_{t\geq 0}$ be the bias of the process induced by the* UNDECIDED-STATE *dynamics with uniform noise probability p. We prove the following.*

— *If $p < 1/2$, let $s_0 = \Omega(\sqrt{n \log n})$ be the bias at the beginning of the process. Then, there exists a time $\tau_1 = \mathcal{O}(\log n)$ such that, w.h.p., the process at time τ_1 reaches a metastable almost-consensus phase characterized by a bias $s_{\tau_1} = \Theta(n)$ towards the initial majority opinion. Moreover, the bias keeps of magnitude $\Theta(n)$ for $\exp(\Theta(n))$ rounds w.h.p.*

— *If $p < 1/2$, let $s_0 = \mathcal{O}(\sqrt{n \log n})$ be the bias at the beginning of the process. Then, there exists a time $\tau_2 = \mathcal{O}(\log n)$ such that, w.h.p. , the system becomes unbalanced towards an opinion, i.e.,*

$$|s_{\tau_2}| = \Omega(\sqrt{n \log n}).$$

— *If $p > 1/2$, let $s_0 = \Omega(\sqrt{n \log n})$ be the bias at the beginning of the process. Then, there exists a time $\tau_3 = \mathcal{O}(\log n)$ such that, w.h.p. , at time τ_3 the majority opinion is lost, i.e., $s_{\tau_3} = \mathcal{O}(\sqrt{n})$. In addition, for $n^{\Theta(1)}$ additional rounds the absolute value of the bias is bounded by $\mathcal{O}(\sqrt{n \log n})$ w.h.p.*

Theorem 1.1.2 (Phase transition of the 3Maj-dyn)**.** *Let $\{s_t\}_{t \geq 0}$ be the bias of the process induced by the* 3-MAJORITY *dynamics with uniform noise probability p. We prove the following.*

— *If $p < 1/3$, let $s_0 = \Omega(\sqrt{n \log n})$ be the bias at the beginning of the process, $s_{eq} = \frac{n}{1-p}\sqrt{\frac{1-3p}{1-p}}$, and let $\varepsilon > 0$ be any sufficiently small constant. Then, there exists a time $\tau_1 = \mathcal{O}(\log n)$ such that, w.h.p., the process at time τ_1 reaches a metastable almost-consensus phase characterized by the equilibrium point s_{eq}, i.e.,*

$$s_{\tau_1} \in [(1-\varepsilon)s_{eq}, (1+\varepsilon)s_{eq}].$$

Moreover, the bias oscillates in such interval for $n^{\Theta(1)}$ rounds w.h.p.

— *If $p < 1/3$, let $s_0 = \mathcal{O}(\sqrt{n \log n})$ be the bias at the beginning of the process. Then, there exists a time $\tau_2 = \mathcal{O}(\log n)$ such that, w.h.p. , the system becomes unbalanced towards an opinion, i.e.,*

$$|s_{\tau_2}| = \Omega(\sqrt{n \log n}).$$

— *If $p > 1/3$, let $s_0 = \Omega(\sqrt{n \log n})$ be the bias at the beginning of the process. Then, there exists a time $\tau_3 = \mathcal{O}(\log n)$ such that, w.h.p., at time τ_3 the majority opinion is lost, i.e., $s_{\tau_3} = \mathcal{O}(\sqrt{n})$. In addition, with constant probability, at time $\tau_3 + 1$ the majority opinion changes. Moreover, for $n^{\Theta(1)}$ additional rounds the absolute value of the bias is $\mathcal{O}(\sqrt{n \log n})$ w.h.p.*

A more exhaustive statement of our results and their analysis is given in Chapter 2. Theorems 1.1.1 and 1.1.2 show that the 3-MAJORITY dynamics is less resilient to noise than the UNDECIDED-STATE dynamics, despite in the 3-MAJORITY dynamics more communication per-round are allowed. Indeed, the phase transition of the two dynamics are, respectively, at the thresholds $p = 1/3$ and $p = 1/2$; the two-phases update-rule of the UNDECIDED-STATE dynamics turns out to be more robust to noise and, hence, a swarm of agents would benefit from it. With respect to this, in Chapter 2 we show that the UNDECIDED-STATE dynamics can be derived by a discretization of the differential equations describing a best-of-N nest-site selection processes in honeybees [Reina et al., 2017]. We remark that the obtained phase transitions separate qualitatively the behaviors of the UNDECIDED-STATE and the 3-MAJORITY dynamics from that of the VOTER model which is, to the best of our knowledge, the only linear opinion dynamics (with a finite opinion set) which has been rigorously analyzed in the presence of uniform communication noise or stubborn agents [Mobilia et al., 2007, Yildiz et al., 2013]. For this dynamics, it has been shown that no form of consensus can be reached in the presence of uniform noise: this hints at a more general phenomenon for non-linear dynamics with fast convergence to some metastable consensus. We remind that in the (a)synchronous version of the VOTER model, any agent pulls just a neighbor u.a.r. and adopts its opinion.

Furthermore, in Chapter 2 we also show that the noise model we consider in the complete network (for the general case of k opinions) is equivalent to a model without any communication noise and where $\frac{pn}{(1-p)}$ stubborn agents (that is, they never change opinion), organized in k communities where each the community holds a different opinion, are added to the network.

We remark that our result on the 3-MAJORITY dynamics is more complete than that on the UNDECIDED-STATE dynamics. A first difference lies in the fact that in the 3-MAJORITY dynamics we find a precise equilibrium value s_{eq} that is attractive for the bias when the noise is below the phase-transition threshold. Secondly, we characterize in detail what happens in the metastable almost-consensus phase: for every arbitrary small value $\varepsilon > 0$, we prove that the bias oscillates in the interval $[(1-\varepsilon)s_{\text{eq}}, (1+\varepsilon)s_{\text{eq}}]$ for polynomial time w.h.p. Instead, in the UNDECIDED-STATE dynamics we could only prove the bias converges to a larger interval of width $\Theta(n)$, without arbitrarily approaching an equilibrium state. On the other hand, when the noise probability is above the threshold, we also show that in the 3-MAJORITY dynamics the majority opinion switches every $\mathcal{O}(\log n)$ rounds with constant probability. In order to prove this, some drift analysis results with super-martingale arguments are used [Lehre and Witt, 2014]. The reason for the lack of such refinements in the results for the UNDECIDED-STATE dynamics lies in the fact that this dynamics is more difficult to analyze, as there is a third opinion to consider and the round-by-round update-rule depends on the opinion of the agent. Nevertheless, we remark that we think the UNDECIDED-STATE process should behave in such a way.

As future directions that interest us, sparser topologies are worth to be investigated theoretically. We believe that, as long as the communication graph shows strong connection properties, similar phase transitions will be exhibited. Indeed, in Chapter 2 we also show some performed simulations corroborating this hypothesis. Furthermore, it would be interesting to see whether the UNDECIDED-STATE and the 3-MAJORITY dynamics with an arbitrary number of possible opinions, with the same noise model, have the exact same phase transitions: in general, $p = 1/2$ for the UNDECIDED-STATE dynamics means that half of the round-by-round communications are non-noisy, on average, while $p = 1/3$ in the 3-MAJORITY dynamics corresponds to the fact that, for each node and at each round, exactly one communication among the three ones is noisy in expectation.

We believe this thesis contributes to the research endeavor of exploring the interplay between communication noise and stochastic interaction pattern in multi-agent systems.

1.1.2 Related works

On the UNDECIDED-STATE dynamics. The UNDECIDED-STATE dynamics has been originally studied as an efficient majority-consensus population protocol, that is, where the process is asynchronous and at each round two agents sampled u.a.r. interact, by [Angluin et al., 2008] and independently by [Benezit et al., 2009] for the binary opinion case. They prove that, w.h.p., within a logarithmic number of parallel rounds, all agents support the initial majority opinion. Some works have then extended the analysis of the UNDECIDED-STATE dynamics to non-complete topologies. In the Poisson-clock model (formally equivalent to the population protocol model), the authors of [Draief and Vojnovic, 2012] derive an upper bound on the expected convergence time of the dynamics that holds for arbitrary connected graphs, which is based on the location of eigenvalues of some contact rate matrices. They also instantiate their bound for particular network topologies. Successively, going back to the population protocol model, [Mertzios et al., 2016] provides an analysis when the initial states of agents are assigned independently at random, and they

also derive "bad" initial configurations on certain graph topologies such that the initial minority opinion eventually becomes the majority one. As for the use of UNDECIDED-STATE dynamics as a generic synchronous consensus protocol for the binary opinion case, where all agents update their state in parallel, [Clementi et al., 2018] shows that the convergence time of the UNDECIDED-STATE dynamics is, w.h.p., logarithmic. In particular, whenever the initial bias towards some majority opinion is of order $\Omega\left(\sqrt{n \log n}\right)$, the process achieves majority consensus. Interestingly enough, our results show that the behavior of the UNDECIDED-STATE dynamics in presence of noise is qualitatively equivalent to the non-noisy case, except for the complete consensus as the final configuration, when the noise is under the phase-transition threshold.

As for the many opinion case, the authors of [Becchetti et al., 2015] show that the UNDECIDED-STATE dynamics in the complete graph is able to reach plurality consensus provided that the ratio between the initial plurality opinion community size and the second largest one is at least some positive constant. More precisely, if the number of opinions is k and the number of agents supporting opinion i is c_i in the initial configuration, they define the monochromatic distance of the initial configuration to be $\mathrm{md} = \sum_{i \leq k} \frac{c_i}{c_1}$, where it is assumed wlog that $c_1 \geq c_2 \geq \cdots \geq c_k$; if $c_1 \geq (1+\alpha)c_2$, with $\alpha > 0$ being a constant, the convergence time to plurality consensus is upper bounded by $\mathcal{O}(\mathrm{md} \log n)$ and lower bounded by $\Omega(\mathrm{md})$, w.h.p. Thus, the convergence time is linear in the monochromatic distance of the initial configuration. We remark that $1 \leq \mathrm{md} \leq k$, so the more general upper bound on the consensus time is $\mathcal{O}(k \log n)$. Moreover, in [Becchetti et al., 2015] it is also shown how to adapt the UNDECIDED-STATE dynamics to solve the plurality consensus on expander graphs.

More in general, the UNDECIDED-STATE dynamics has also been considered as a model of some mechanism occurring in the biology of a cell [Cardelli and Csikàsz-Nagy, 2012] and it has been employed as a sub-routine of efficient majority consensus protocols: indeed, [Ghaffari and Parter, 2016], [Berenbrink et al., 2016], and [Elsässer et al., 2017] consider majority consensus in the synchronous model, and design protocols (based on the UNDECIDED-STATE dynamics) which w.h.p. converge in polylogarithmic time even if the number of initial opinions is very large. It is worth mentioning the more recent work [Bankhamer et al., 2022], which analyzes a variant of the UNDECIDED-STATE dynamics both in its asynchronous and synchronous versions, which solve the plurality consensus in parallel time $\mathcal{O}(\log^2 n)$ for an arbitrary number of opinions, starting from any initial configuration, at the cost of a per-agent memory capable of storing $k \cdot \mathcal{O}(\log n)$ states. We remark that, in the former UNDECIDED-STATE dynamics, the agent needs a memory capable of storing only $k+1$ states, namely, the k opinions and the undecided state.

On of the 3-MAJORITY dynamics. The study of the 3-MAJORITY dynamics arises on the ground of the results obtained for the MEDIAN dynamics in [Doerr et al., 2011]. The MEDIAN dynamics considers a totally ordered opinion set, in which each agent pulls two neighbor opinions i, j u.a.r. and then updates its opinion k to the median between i, j, and k. The dynamics turns out to be a fault-taulerant, efficient dynamics for the majority consensus problem. However, as pointed out in [Becchetti et al., 2020a], the MEDIAN dynamics may not guarantee with high probability convergence to a valid opinion in case of the presence of an adversary, which is needed for the consensus problem. Moreover, the opinion set must have an ordering, property that might not be met by applicative scenarios such as biological systems [Becchetti et al., 2020a]. These facts naturally lead researchers to look for efficient dynamics that satisfy the above requirements.

To the best of our knowledge, [Abdullah and Draief, 2015] is the first work analyzing the h-MAJORITY dynamics. In detail, in the h-MAJORITY dynamics we have n nodes and, at every

round, every node pulls the opinion from h random neighbors and sets his new opinion to the majority one (ties are broken arbitrarily). More extensive characterizations of the 3-MAJORITY dynamics over the complete graph are given in [Becchetti et al., 2017, Berenbrink et al., 2017, Becchetti et al., 2016, Ghaffari and Lengler, 2018].

In [Becchetti et al., 2017] it is shown that the 3-MAJORITY dynamics is a fast, fault-tolerant protocol for (valid) majority consensus in the case of $k \geq 2$ opinions, provided that there is an initial bias towards some majority opinion. Furthermore, [Becchetti et al., 2017] shows an exponential time-gap between the 3-MAJORITY consensus process and the median process in [Doerr et al., 2011], thus establishing its efficiency. In [Becchetti et al., 2016], the analysis is extended to any (even balanced) initial configuration in the many-opinion case, in the presence of a different kind of bounded adversaries. The authors of [Becchetti et al., 2016] emphasize how the absence of an initial majority opinion considerably complicates the analysis, as it must be proved that the process breaks the initial symmetry despite the presence of the adversary. Indeed, before the symmetry breaking, the adversary is more likely to cause undesired behaviors. The strongest result about the convergence of the 3-MAJORITY is that in [Ghaffari and Lengler, 2018]. The authors show that in the case of k opinions, the process converges in time $O(k \log n)$ rounds, and it is tight when $k = O(\sqrt{n})$. The 3-MAJORITY dynamics is also studied in different topologies: [Kang and Rivera, 2019] analyzes the 3-MAJORITY process in graphs of minimum degree n^{α}, with $\alpha = \Omega\left((\log \log n)^{-1}\right)$, starting from random biased binary configurations.

On the 2-CHOICES dynamics. Another important and efficient opinion dynamics for the majority consensus problem is the 2-CHOICES dynamics. A quick description of its update-rule follows: at each round, each agent samples two neighbors u.a.r. and updates its opinion to the majority opinion among its former opinion and the two sampled neighbor opinions, if there is any. Otherwise, it keeps its opinion. We just want to remark that the expected round-by-round behavior of the 2-CHOICES dynamics and that of the 3-MAJORITY are the same, while the actual behaviors differ substantially in high probability [Berenbrink et al., 2017]. This is why mean-field arguments are sometimes not sufficient to analyze such processes. For example, we have ran simple experiments that suggest that our uniform noise model on the 2-CHOICES dynamics yields a threshold noise value $p = 1/2$, just like the UNDECIDED-STATE dynamics. For an overview on the state of the art about opinion dynamics in general, we defer the reader to [Becchetti et al., 2020a].

Consensus dynamics in the presence of noise or stubborn agents. The authors of [Wang and Liu, 2009] initiate the study of the consensus problem in the presence of communication noise. They consider the Vicsek model [Vicsek et al., 1995], in which they introduce a noise feature and a notion of robust consensus. Subsequently, dynamics for the consensus problem with noisy communications have received considerable attention. In particular, as mentioned in the introduction, this direction is motivated, among many reasons, by the desire to find models for the consensus problem in natural phenomena [Feinerman et al., 2017].

The communication noise studied in this type of problem can be devided in two types: uniform (or unbiased) and non-uniform (or biased). The uniform case wants to capture errors in communications between agents in real-world scenarios. The non-uniform communication noise instead describes the case in which agents have a preferred opinion. The authors of [Feinerman et al., 2017] are the first to explicitly focus on the uniform noise model. In detail, they study the broadcast and the majority consensus problem when the opinion set is binary. In their model of noise,

every bit in every exchanged message is flipped independently with some probability smaller than $1/2$. As a result, the authors give natural protocols that solve the aforementioned problems efficiently. The work [Fraigniaud and Natale, 2019] generalizes the above study to opinion sets of any cardinality.

As for the non-uniform communication noise case, in [Cruciani et al., 2021] it is considered the h-MAJORITY dynamics (where the majority update-rule is performed over a sample of h neighbors) with a binary opinion set $\{Alpha, Beta\}$, where they have a probability p that any received message is flipped towards a *fixed* preferred opinion, say *Beta*, while with probability $1 - p$ the former message keeps intact. The authors suppose there is an initial majority agreeing on *Alpha*, and they analyze the *time of disruption*, that is the time the initial majority is subverted. They prove there exists a threshold value $p^\star$ (which depends on h), such that 1) if $p < p^\star$, the time of disruption is at least polynomial, w.h.p., and 2) if $p > p^\star$, the time of disruption is constant, w.h.p. Their result holds for any sufficiently dense graph. We remark that our work on the 3-MAJORITY dynamics differs from the one in [Cruciani et al., 2021] as there is no preferred opinion, and the noise affecting communications may result in any possible opinion.

As we show in Chapter 2, the noise feature we consider is equivalent to a model in which communities of stubborn agents (i.e., they never change opinion) are added to the network. Hence, we discuss some previous works that consider such a model. In [Yildiz et al., 2013], the authors focus on the VOTER model, and show that the presence stubborn agents with opposite opinions precludes the convergence to consensus. The work [Mukhopadhyay et al., 2020] studies the asynchronous VOTER rule and the asynchronous majority rule dynamics with Poisson clocks, when the opinion set is binary using mean-field techniques, in presence of agents that either have a probability (which depends on their current opinion) not to update when the clock ticks, or are stubborn. In the second case, which directly relates with this thesis, they show that for the 3-MAJORITY dynamics there are either one or two possible stable equilibria, depending on the sizes of the stubborn communities, which are reached in logarithmic time. If the two sizes are close between each other and not too large, then agreement on both opinions is possible in the steady state. Otherwise, either no agreement is possible, or the process converges to an agreement towards a single opinion (that of the largest stubborn community). This thesis includes the case in which the two stubborn communities have equal size, which corresponds to the uniform communication noise model. Nevertheless, we have some crucial differences: first of all, this thesis consider the synchronous version of the 3-MAJORITY dynamics, which cannot be analyzed with the same tools. Indeed, in each update round, the synchronous model has non-zero probability to reach any of the two monochromatic configurations. This feature is absent in the asynchronous version, since at each update, with probability equal to 1, at most one agent can change opinion. Furthermore, we want to remark that mean-field arguments do not capture important aspects of the process, such as metastability, which in [Mukhopadhyay et al., 2020] is shown only through simulations. For example, the actual behavior of the process in the long-term is oscillation in a very small interval around the equilibrium values, spending long times in those intervals, and eventually switching between the two. We characterize the width of the oscillation interval and show there is high probability of convergence, providing also a lower bound on the time the process spends in the equilibrium interval.

1.2 Lévy walks

A *Lévy walk* is a random walk process in which jump lengths are drawn from a power-law distribution. Thus, the walk consists of a mix of long trajectories and short, random movements. Over the last two decades, Lévy walks have attracted significant attention, as a result of increasing empirical evidence that the movement patterns of various animal species resemble Lévy walks. Examples of such species range from snails [Reynolds et al., 2017], bees [Reynolds et al., 2007], and albatross birds [Viswanathan et al., 1996], to sharks [Humphries et al., 2010, Sims et al., 2008], deers [Focardi et al., 2009], and humans [Boyer et al., 2006, Raichlen et al., 2014], among others [Reynolds, 2018]. Nowadays, Lévy walks are the most prominent movement model in biology [Reynolds, 2018], at least among models with comparable mathematical simplicity and elegance [Viswanathan et al., 2011].

The *Lévy flight foraging hypothesis*, put forward by Viswanathan et al. [Viswanathan et al., 1999, Viswanathan et al., 2008], stipulates that the observed Lévy walk movement patterns in animals must have been induced by natural selection, due to the optimality of Lévy walks in searching for food. Indeed, it has been shown that Lévy walks achieve (near) optimal search time in certain settings. In particular, Levy walks with exponent parameter $\alpha = 2$ are optimal for searching sparse randomly distributed revisitable targets [Viswanathan et al., 1999]. However, these results were formally shown just for one-dimensional spaces [Buldyrev et al., 2001], and do not carry over to higher-dimensions [Levernier et al., 2020]. Very recently, a new argument was provided in [Guinard and Korman, 2021] supporting the optimality of Lévy walks with $\alpha = 2$. In the considered setting, the space is a two-dimensional square torus of area n, and the walk must find a single, randomly selected target. Two critical model assumptions are that the target may have an arbitrary diameter D, and that the Lévy walk is "intermittent," i.e., cannot detect the target during a jump, only at the end of the jump. Under these assumptions, the Cauchy walk (that is, a Lévy walk with power-law exponent $\alpha = 2$) was shown to achieve a (near) optimal search time of $\tilde{O}(n/D)$, whereas exponents $\alpha \neq 2$ are suboptimal. We remark that if the target has a fixed size $D = 1$ or the walk is not intermittent, then all exponents $\alpha \geq 2$ or $\alpha \leq 2$, respectively, are optimal as well. Currently, the latter is the only work (other than ours) that shows evidence supporting the optimality of Lévy walks in d-dimensional spaces for $d > 1$.

In this thesis we focus on Lévy walks on two-dimensional spaces. Concretely, we assume the infinite lattice $\mathbb{Z}^2$, and consider the Manhattan distance as the underlying metric. To determine each jump of the Lévy walk, an integer distance d is chosen independently at random such that the probability of $d = j$ is inversely proportional to j^α, where $\alpha \in (1, \infty)$ is the exponent parameter of the walk. A destination v is then chosen uniformly at random among all nodes at distance d from the current node u, and in the next d steps, the process moves from u to v along a shortest lattice path approximating the straight line segment $\overline{uv}$.

We evaluate the search efficiency of Lévy walks on $\mathbb{Z}^2$, by analysing the parallel hitting time of multiple walks originating at the same node. Precisely, we assume that $k \geq 1$ independent Lévy walks start simultaneously from the origin $(0, 0)$ of the lattice. Then the parallel hitting time for any given target node u^* is the first step when some walk visits u^*. This very basic setting can be viewed as a model of natural cooperative foraging behavior, such as the behavior of ants around their nest. In fact, our setting is as a special instance of the more general ANTS problem introduced by Feinerman and Korman [Feinerman and Korman, 2017]. The ANTS problem asks for a search strategy for k independent agents that minimizes the parallel hitting time for an unknown target, subject to limited communication before the search starts. The related works are further discussed

in Section 1.2.3. In the following section, we discuss our main results which are based on the work [Clementi et al., 2021], while the detailed statements and their mathematical analyses are given in Chapter 3.

1.2.1 Our results

Hitting time bounds for a single Lévy walk. A main technical contribution of this thesis is an analysis of the hitting time $\tau_\alpha(u^*)$ of a single Lévy walk with exponent $\alpha \in (1,\infty)$, for an arbitrary target node u^*. We show the following bounds on $\tau_\alpha(u^*)$, assuming the Lévy walk starts at the origin $(0,0)$, and u^*'s distance to the origin is $\ell = \|u^*\|_1$.

Consider first the *super-diffusive* regime, where $\alpha \in (2,3)$. In this regime, jump lengths have bounded mean and unbounded variance. Roughly speaking, we show that in the first $t_\ell = \Theta(\ell^{\alpha-1})$ steps, [2] the walk stays inside a ball of radius $(t_\ell \cdot \operatorname{polylog} \ell)^{\frac{1}{\alpha-1}}$ with significant probability, while only a constant fraction of those steps are inside the smaller ball of radius ℓ. We also show a monotonicity property, which roughly implies that the probability of visiting a node decreases as the node's distance from the origin increases. Therefore, a constant fraction of the t_ℓ steps visits nodes at distances between ℓ and $\ell \cdot \operatorname{polylog} \ell$, and the visit probability of each of these nodes is upper bounded by that of node u^*. We thus obtain that the probability of visiting u^* within t_ℓ steps is $\Omega\left(t_\ell/\ell^2 \operatorname{polylog} \ell\right)$.

If we consider a smaller number of steps, $t = \mathcal{O}\left(t_\ell/\operatorname{polylog} \ell\right)$, then it is very likely that the walk stays in a ball of radius smaller than ℓ, and we show a simple bound of $\mathcal{O}\left((t/t_\ell)^2 \cdot t_\ell/\ell^2\right)$ for the probability of $\tau_\alpha(u^*) \leq t$, i.e., ignoring $\operatorname{polylog} \ell$ factors, the probability decreases by a factor of $\mathcal{O}\left((t/t_\ell)^2\right)$.

On the other hand, if we consider a larger number of steps (even if $t \to \infty$), the probability that u^* is hit does not increase significantly, just by at most a $\operatorname{polylog} \ell$ factor.

Therefore, in regime $\alpha \in (2,3)$, $\Theta(\ell^{\alpha-1})$ steps suffice to maximize the hitting probability (within $\operatorname{polylog} \ell$ factors), while reducing this time reduces the probability super-linearly.

The *diffusive* regime, $\alpha \in (3,\infty)$, is similar to the case of a simple random walk, as jump lengths have bounded mean and bounded variance. We show that $\mathcal{O}\left(\ell^2 \operatorname{polylog} \ell\right)$ steps suffice to hit the target with probability $\Omega\left(1/\operatorname{polylog} \ell\right)$, while for a smaller number of steps t, the probability decreases by a factor of $\mathcal{O}\left((t/\ell^2)^2\right)$. The behavior is similar also in the threshold case of $\alpha = 3$, even though the variance of the jump length is unbounded in this case.

Finally, in the *ballistic* regime, $\alpha \in (1,2]$, where jump lengths have unbounded mean and unbounded variance, the behavior is similar to that of a straight walk along a random direction. We show that the target is hit with probability $\Omega\left(1/\ell \operatorname{polylog} \ell\right)$ in the first $\Theta\left(\ell\right)$ steps, while increasing the number of steps does not increase this probability significantly.

Below we give formal statements of these results, for the case where α is independent of ℓ, as $\ell \to \infty$. More refined statements and their proofs are given in Chapter 3. (Sections 3.3 to 3.5).

Theorem 1.2.1. *Let α be any real constant in $(2,3)$ and u^* any node in $\mathbb{Z}^2$ with $\ell = \|u^*\|_1$. Then:*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell^{\alpha-1}\right)\right] = \Omega\left(1/\ell^{3-\alpha} \log^2 \ell\right)$;

(b) $Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(t^2/\ell^{\alpha+1}\right)$, *for any step* $\ell \leq t = \mathcal{O}\left(\ell^{\alpha-1}\right)$;

(c) $Pr\left[\tau_\alpha(u^*) < \infty\right] = \mathcal{O}\left(\log \ell/\ell^{3-\alpha}\right)$.

2. Note that t_ℓ is of the same order as the expected number of steps before the first jump of length greater than ℓ.

Theorem 1.2.2. *Let α be any real constant in $[3,\infty)$ and u^* any node in $\mathbb{Z}^2$ with $\ell = \|u^*\|_1$. Then:*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = \Omega\left(1/\log^4 \ell\right)$;

(b) $Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(t^2 \log \ell/\ell^4\right)$, *for any step t with $\ell \leq t = \mathcal{O}\left(\ell^2\right)$.*

Theorem 1.2.3. *Let α be any real constant in $(1,2]$ and u^* any node in $\mathbb{Z}^2$ with $\ell = \|u^*\|_1$. Then:*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell\right)\right] = \Omega\left(1/\ell \log \ell\right)$;

(b) $Pr\left[\tau_\alpha(u^*) < \infty\right] = \mathcal{O}\left(\log^2 \ell/\ell\right)$.

Parallel Lévy walks with common exponent. Consider $k \geq 1$ independent identical Lévy walks with exponent $\alpha \in (1,\infty)$, that start simultaneously at the origin. Let $\tau_a^k(u^*)$ denote the parallel hitting time for node u^*, i.e., the first step when some walk visits u^*. It is straightforward to derive upper and lower bounds on $\tau_a^k(u^*)$ from the corresponding bounds on the hitting time of a single Lévy walk. For example, the next statement is a direct corollary of Theorem 1.2.1Part a.

Corollary 1.2.4. *Let α be any real constant in $(2,3)$ and u^* any node in $\mathbb{Z}^2$ with $\ell = \|u^*\|_1$. Then*

$$Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\ell^{\alpha-1}\right)\right] = 1 - e^{-\Omega\left(k/\ell^{3-\alpha}\log^2 \ell\right)}.$$

From the bounds we obtain for τ_α^k, it follows [3] that, for each pair of k and $\ell = \|u^*\|_1$ with $\text{polylog}\,\ell \leq k \leq \ell\,\text{polylog}\,\ell$, there is a unique optimal exponent $\alpha = 3 - \frac{\log k}{\log \ell} + \mathcal{O}\left(\frac{\log\log \ell}{\log \ell}\right)$, which minimizes $\tau_\alpha^k(u^*)$, w.h.p. Moreover, increasing or decreasing this exponent by an arbitrarily small constant term, respectively increases the hitting time by a $\text{poly}(\ell)$ factor, or the walks never hit u^* with probability $1 - o(1)$. For the case of $k \leq \text{polylog}\,\ell$ or $k \geq \ell\,\text{polylog}\,\ell$, all exponents $\alpha \in [3,\infty)$ or $\alpha \in (1,2]$, respectively, achieve the same optimal value of $\tau_\alpha^k(u^*)$ (within $\text{polylog}\,\ell$ factors). Formal statements of these results are given in Chapter 3 (Sections 3.3 to 3.5). The theorem below bounds the parallel hitting time for (near) optimal choices of α.

Theorem 1.2.5. *Let u^* be any node in $\mathbb{Z}^2$, and $\ell = \|u^*\|_1$.*

(a) If $\log^6 \ell \leq k \leq \ell \log^4 \ell$, then for $\alpha = 3 - \frac{\log k}{\log \ell} + 5\frac{\log\log \ell}{\log \ell}$, $Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\frac{\ell^2 \log^6 \ell}{k}\right)\right] = 1 - e^{-\omega(\log \ell)}$;

(b) If $k = \omega(\log^5 \ell)$, then $Pr\left[\tau_3^k(u^*) = \mathcal{O}\left(\ell^2\right)\right] = 1 - e^{-\omega(\log \ell)}$;

(c) If $k = \omega(\ell \log^2 \ell)$, then $Pr\left[\tau_2^k(u^*) = \mathcal{O}\left(\ell\right)\right] = 1 - e^{-\omega(\log \ell)}$.

Observe that for any given k, ℓ with $k = \omega(\log^5 \ell)$, if we choose the exponent α as in Theorem 1.2.5, then

$$\Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left((\ell^2/k)\log^6 \ell + \ell\right)\right] = 1 - e^{-\omega(\log \ell)}. \tag{1.1}$$

3. In fact, we use more refined versions of Theorems 1.2.1 to 1.2.3, to obtain bounds on τ_α^k which allow α to be a function of ℓ and k.

Parallel Lévy walks with random exponents The right choice of α, according to Theorem 1.2.5, requires knowledge of the values of k and ℓ (at least within polylogarithmic factors). We propose a very simple randomized strategy for choosing the exponents of the k Lévy walks, which almost matches the parallel hitting time bounds of Theorem 1.2.5, for all distances ℓ *simultaneously*. The strategy does not require knowledge of ℓ, and works as long as $k \geq \operatorname{polylog} \ell$. Interestingly, it does not require knowledge of k either. The strategy is the following:

> *The exponent of each walk is sampled independently and uniformly at random from the real interval* $(2, 3)$.

The next theorem bounds the resulting parallel hitting time $\tau^k_{rand}(u^*)$, for an arbitrary node u^*. Its proof is given in Chapter 3 (Section 3.6).

Theorem 1.2.6. *Let u^* by any node in $\in \mathbb{Z}^2$, $\ell = \|u^*\|_1$, and $k \geq \log^8 \ell$. Then*

$$Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left((\ell^2/k)\log^7 \ell + \ell \log^3 \ell\right)\right] = 1 - e^{-\omega(\log \ell)}. \tag{1.2}$$

By comparing Eq. (1.1) and Eq. (1.2), we observe that indeed the hitting time of the randomized strategy is only by a $\operatorname{polylog} \ell$ factor worse than that of the deterministic strategy based on Theorem 1.2.5, which knows ℓ and k. Moreover, this hitting time is optimal within a $\operatorname{polylog} \ell$ factor among *all* possible search strategies (deterministic or randomized) that do not know ℓ within a constant factor, since a universal lower bound of $\Omega\left(\ell^2/k + \ell\right)$ with constant probability applies to all such strategies, as observed in [Feinerman and Korman, 2017].

1.2.2 Implications on Lévy Hypothesis and Distributed Search

As already mentioned, our setting of k independent walks starting from the same location, aiming to hit an unknown target, can be viewed as a basic model of animals' foraging behavior around a central location, such as a nest, a food storage area, or a sheltered environment. The assumption that walks are independent is approximately true for certain animal species such as ants Cataglyphis, which lack pheromone-based marking mechanisms [Razin et al., 2013]. Our results suggest that if the typical or maximum distance ℓ of the food (target) from the nest (source) is fixed, then a group of animals executing parallel Lévy walks with the same exponent can optimize search efficiency by tuning the exponent value and/or the number k of animals participating in the foraging. In that setting, no universally optimal exponent value exists, as the optimal exponent depends on k and ℓ. An alternative, novel approach suggested by our last result is that each animal performs a Lévy walk with a randomly chosen exponent. This strategy, which surprisingly achieves near optimal search efficiency for all distance scales, implies that different members of the same group follow different search patterns. The existence of such variation in the search patterns among individuals of the same species requires empirical validation.

In the context of the related ANTS problem [Feinerman and Korman, 2017], our result on parallel Lévy walks with randomly selected exponents directly implies a uniform solution to the problem (i.e., independent of k and ℓ), which is extremely simple and natural, and is optimal within $\operatorname{polylog} \ell$ factors, w.h.p.

1.2.3 Related Work

On the Lévy walks. Lévy walks (also referred to as Lévy flights) have been studied mostly by physicists, and mainly in continuous spaces [Zaburdaev et al., 2015, Reynolds, 2018]. The idea

that biological organisms could perform Lévy walks was first suggested in the mid 80s [Shlesinger and Klafter, 1986], as a potentially more efficient search strategy compared to Brownian motion. Lévy walks attracted significant attention after experimental work in the mid 90s showed that albatross birds follow Lévy walk-like trajectories [Viswanathan et al., 1996], a pattern that was subsequently observed for various other organisms as well [Reynolds et al., 2017, Reynolds et al., 2007, Humphries et al., 2010, Sims et al., 2008, Focardi et al., 2009, Boyer et al., 2006, Raichlen et al., 2014]. Even though statistical and methodological flaws were later pointed out in several of these works [Edwards et al., 2007], there is currently ample evidence that many animals do exhibit Lévy walk movements [Viswanathan et al., 2011, Humphries et al., 2012].

The Lévy flight foraging hypothesis. A possible explanation for this phenomenon is the *Lévy foraging hypothesis* [Viswanathan et al., 1999, Viswanathan et al., 2008]: "According to the optimal foraging theory [Werner and Hall, 1974], natural selection drives species to adopt the most economically advantageous foraging pattern. Thus species must have adapted to follow Lévy walks because Lévy walks optimize search efficiency." The main theoretical argument in support of this hypothesis was provided in [Viswanathan et al., 1999], stating that a Lévy walk with exponent $\alpha = 2$ (known as Cauchy walk) maximizes the number of visits to targets, when targets are sparse and uniformly distributed. This result has been formally shown for one-dimensional spaces [Buldyrev et al., 2001], but is not true for higher-dimensional spaces [Levernier et al., 2020], at least not without additional assumptions [Buldyrev et al., 2021, Levernier et al., 2021].

Our results add a new perspective to the Lévy foraging hypothesis. Unlike [Viswanathan et al., 1999] and [Guinard and Korman, 2021], we consider a *collective* search setting, where k individuals start from the same source and move independently. The space is two-dimensional as in [Guinard and Korman, 2021] (but discrete and unbounded), and there is a single target (of unit size). If rough information about the target's distance ℓ to the source is known then letting all individuals execute identical Lévy walks with a specific exponent, which depends on k and ℓ, achieves (near) optimal search time. If no information on ℓ is available, then using a random exponent for each walk, sampled independently from the super-diffusive range $(2, 3)$, still achieves near optimal search time, for all distances ℓ.

The ANTS problem. In the Ants-Nearby-Treasure-Search (ANTS) problem [Feinerman and Korman, 2017], k identical (probabilistic) agents starting from the same location, search for an unknown target on $\mathbb{Z}^2$. Agents do not know k, and cannot communicate (or see each other). However, before the search begins, each agent receives a b-bit *advice* from an oracle. In [Feinerman and Korman, 2017], matching upper and lower bounds are shown for the trade-off between the expected time until the target is found, and the size b of the advice. The proposed optimal algorithms repeatedly execute the following steps: walk to a random location in a ball of a certain radius (chosen according to the algorithm specifics), perform a spiral movement of the same radius as the ball's, then return to the origin. We shall compare our results only with those for the setting in which no advice is given. In [Feinerman and Korman, 2017], it is shown that no uniform algorithm (where each agent starts at the same state) provides expected hitting time $\mathcal{O}\left((\ell^2/k + \ell)\log k\right)$, and an algorithm achieving expected hitting time $\mathcal{O}\left((\ell^2/k + \ell)\log^{1+\varepsilon} k\right)$ is given, for any constant $\varepsilon > 0$, which has the same structure described above. However, as argued by the authors, such algorithms are only relatively simple and, thus, a simpler one, the *Harmonic Search* Algorithm (HSA), is provided. The HSA works as follows: if δ is any positive constant, each agent moves to

a node u sampled with a power-law distribution proportional to $d(u)^{-2-\delta}$, where $d(u)$ is the distance of the agent from u. Then it performs a spiral search around u for $d(u)^{2+\delta}$ steps; [4] afterwards it goes back to the origin and repeats the procedure. With probability $1-\varepsilon$, for any constant $\varepsilon > 0$, this algorithm achieves hitting time $\mathcal{O}\left(\ell^{2+\delta}/k + \ell\right)$. Even if the algorithm we propose doesn't match the best hitting time shown in [Feinerman and Korman, 2017] for the case of no advice, it does beat the HSA and it is arguably simpler: each agent performs a Lévy walk with a uniformly random exponent sampled from $(2, 3)$. Agents do not need to return to the origin periodically and, more importantly, the search processes do not consist of different stages. Our algorithm is Monte Carlo, and finds the target w.h.p. in time that is larger than the optimal by at most a polylogarithmic factor.

Variants of the ANTS problem have been studied, where agents are (synchronous or asynchronous) finite state machines, which can communicate during the execution whenever they meet [Emek et al., 2014,Emek et al., 2015,Cohen et al., 2017,Lenzen et al., 2017]. Another variant, involving parallel search on the line by k non-communicating agents, is considered in [Fraigniaud et al., 2016].

Random walks in general and optimality results. In our analysis, we derive upper and lower bounds on the hitting time of a Lévy walk on $\mathbb{Z}^2$. Bounds on the hitting time and related quantities for Lévy walks on the (one-dimensional) real line are given in [Palyulin et al., 2019]. Bounds for general random walks on $\mathbb{Z}^d$, for $d \geq 1$, in the case where the walk has bounded second (or higher) moments can be found in [Uchiyama, 2011]. Recall that Lévy walks have unbounded second moment when $\alpha \leq 3$.

In [Boczkowski et al., 2018a, Guinard and Korman, 2020], tight bounds were shown for the cover time on the cycle of a random walk with k different jump lengths. The optimal walk in this case is one that approximates (using k levels) a Lévy walk with exponent $\alpha = 2$.

When $\alpha \in (3, \infty)$, a Lévy walk on $\mathbb{Z}^d$ behaves similarly to a simple random walk, as the variance of the jump length is bounded. In particular, as $\alpha \to \infty$, the Lévy walk jump converges in distribution to that of a simple random walk. Parallel independent simple random walks have been studied extensively on finite graphs, under various assumptions for their starting positions [Alon et al., 2011, Efremenko and Reingold, 2009, Elsässer and Sauerwald, 2011, Ivaskovic et al., 2017, Kanade et al., 2019]. A main objective of that line of work has been to quantify the "speedup" achieved by k parallel walks on the cover time, hitting times, and other related quantities, compared to a single walk.

Small-world networks. The following basic network model has been proposed by Kleinberg to study the small world phenomenon [Kleinberg, 2000]. A square (or, more generally, d-dimensional) finite lattice is augmented by adding one "long-range" edge from each node u, to a uniformly random node v among all nodes at lattice distance k, where distance k is chosen independently for each u, from a power-law distribution with exponent α. That is, the distribution of long-range edges is the same as the jump distribution of a Lévy walk with the same exponent. It was shown that (distributed) greedy routing is optimized when $\alpha = 1$, whereas for $\alpha \neq 1$ the

4. We remark that even if this step may seem complex, a simple random walk for the same amount of steps would be substantially equivalent.

expected routing time is slowed down by polynomial factors [Kleinberg, 2000]. [5] This result is of similar nature as our result for the hitting time of k identical Lévy walks, where exactly one exponent is optimal. However, in our case, this exponent depends on the target distance. In Kleinberg's network, exponent $\alpha = 1$ ensures that the lengths of long-range links are uniformly distributed over all *distance scales*, which facilitates fast routing. In our randomized strategy, availability of a sufficient number of walks with the right exponent is achieved by choosing the exponents uniformly at random over the interval $(2, 3)$.

1.3 The Assembly Calculus

Intelligence in all its forms is one of the most astounding phenomena of the brain whose functioning, however, remains a mystery. How reasoning, problem-solving, decision-making, planning, empathy, language, art are achieved through the activity of neurons and synapses is undoubtedly an interesting question and is the subject of active research. Nevertheless, despite tremendous advances over the past decades in our understanding of neural mechanisms, increasingly assisted and propelled by machine learning, we are still very far from understanding how the brain begets the mind. The difficulty lies in the huge gap of scale and methodology between Experimental Neuroscience and Cognitive Science [Papadimitriou et al., 2020]. This frustration was articulated in a most eloquent way by Nobel laureate Richard Axel, who declared in a 2018 interview [Axel, 2018]: *"We do not have a logic for the transformation of neural activity to thought and action. I consider discerning [this logic] as the most important future direction in Neuroscience"*.

The *Assembly Calculus* (AC) is a recently proposed formal computational distributed system [Papadimitriou et al., 2020]. As far as we know, it is the only computational system in the literature whose explicit purpose is to bridge through computation the gap between neurons and intelligence, that is to say, to function as Axel's logic. The basic data item of the AC is the *assembly of neurons,* a large stable set of neurons believed to represent an idea, object, word, etc., while its operations (project, associate, merge, etc.) create and manipulate assemblies in response to stimuli and other brain events. Importantly, these operations can be provably simulated through the activity of stylized neurons and synapses. Moreover, the AC is a Turing complete computational distributed system that tries to found its basics on the principles of neuroscience. In the next section, we provide a comprehensive introduction to the AC; however, the interested reader may want to read [Papadimitriou et al., 2020].

Is the AC the bridging "logic" sought by Axel? One avenue for pursuing this important question is to demonstrate empirically that reasonably complex cognitive phenomena can be formulated and implemented in the AC framework. Indeed, in the original paper [Papadimitriou et al., 2020] it was argued that aspects of language generation can be handled by the operations of the AC, while in a subsequent work [Mitropolsky et al., 2021], a parser implemented in the AC was demonstrated to analyze syntactically reasonably complex sentences of English, and it was argued that it can be generalized to more complex features as well as other natural languages. Another avenue is to show that assemblies can mediate *learning*: in a very recent paper [Dabagia et al., 2022], it is shown both theoretically and experimentally that in the AC it is possible to develop

5. In the paper, the exponent considered is that of choosing the endpoint of u's long-range link to be a given node v at distance k, which is proportional to $1/k^{\beta}$, where $\beta = \alpha + d - 1$. Thus the optimal exponent is $\beta = 2$ for the square lattice, and $\beta = d$ for the d-dimensional lattice.

mechanisms for learning to classify well-separated classes of stimuli, including clustered distributions and linear threshold functions with margin, under a very mild supervision.

Our contribution in this direction is to demonstrate that a program in the AC is capable of implementing reasonably sophisticated stylized planning strategies – in particular, heuristics for solving tasks in the blocks world [Gupta and Nau, 1991, Slaney and Thiébaux, 2001]. A *blocks-world configuration* is defined by a set of *stacks,* where a stack is a sequence of unique *blocks,* each sitting on top of the previous one. A stack of size one is just a *block sitting on the table* (see, e.g., Fig. 4.1 in Chapter 4). A configuration can be manipulated by moving a block from the top of a stack (or from the table) to the top of another stack (or to the table). A *task* in the blocks world is the following: given a starting configuration C_{init} and a goal configuration C_{goal}, find a sequence of actions which transforms C_{init} into C_{goal}. It was shown in [Gupta and Nau, 1992] that solving a task in the blocks world with the smallest number of actions is NP-Complete, and it was observed that the following provides a simple 2-approximation strategy: *move to the table all blocks that are not in their final positions, and then move these blocks one by one to their final positions.*

Here we implement this strategy in the AC. From the exposition of this implementation and demonstration, which happens to employ representations and structures of a different style from those needed for language tasks [Papadimitriou et al., 2020, Mitropolsky et al., 2021], we believe that it will become clear that more complicated heuristics for solving related tasks can be similarly implemented in the AC.

In fact, the kind of representations needed for planning, involving long "chains" of assemblies linked through strong synaptic connections, reveals a limitation of the AC which was not apparent before: we find empirically that there are limits, depending on the parameters of the execution model, such as the number of excitatory neurons per brain area, synaptic density, synaptic plasticity, and assembly size, on the length of such chains that can be implemented reliably. As chaining is also used in the Turing machine simulation demonstrating the completeness of the AC [Papadimitriou et al., 2020], such limitations are significant because they bound from above the space complexity, and therefore the parallel time complexity, of AC computations. The model is formally introduced in Chapter 4, where we also describe our implementation and our results; we also discuss and quantify the aforementioned issue about chaining in the experimental validation section. These results are based on our work [D'Amore et al., 2022b].

1.3.1 Related Work

Terry Winograd introduced the blocks world half a century ago as the context for his language understanding system SHRDLU [Winograd, 1971], but since then blocks-world planning has been widely investigated, primarily because such tasks appear to capture several of the difficulties posed to planning systems [Gupta and Nau, 1991, Gupta and Nau, 1992]. There has been extensive work in AI on blocks world problems, including recently on leveraging artificial neural networks (ANNs) for solving them, and learning to solve them from examples (e.g., the Neural Logic Machines of [Dong et al., 2019], or Neural Turing Machines, which are used for related problem-solving tasks [Graves et al., 2014]).

Bridging the gap between low-level models of neural activity in the brain and high-level symbolic systems modelling cognitive processes is a fundamental open problem in artificial intelligence and neuroscience at large [Doursat, 2013, Chady, 1999]. Several computational cognitive-science papers address the problem of solving (or learning to solve) block-worlds tasks in higher-level computational models of cognition, such as ACT-R or SOAR (see for instance [Kennedy and

Trafton, 2006, Kurup, 2008, Panov, 2017]). In contrast to this thesis, however, these papers utilize high-level languages and data structures for the programming of these systems, without providing a link, as we do, to the behavior of stylized neurons and synapses, in an effort to remain as faithful as possible to the ways animal brains would solve these tasks. Furthermore, in this thesis the brain does not *learn* how to solve the planning task: we simply implement known heuristics for the problem.

Less related to our problem is the literature on block stacking (see, for example, [Hayashi, 2007, Tian et al., 2020]). These papers focus on the ability of humans and chimpanzees to place a block on top of an existing tower without toppling it.

Finally, it is worth mentioning some previous works on solving planning tasks through spiking neural networks, such as [Rueckert et al., 2016, Basanisi et al., 2020], in which the attention is more focused on learning world models.

A spiking neural network framework not unlike ours is Nengo [Bekolay et al., 2014]. The crucial difference is that our framework is based on the known behavior called assemblies which enable higher levels of abstraction such as the AC, and carrying out far more advanced tasks such as in [Mitropolsky et al., 2021] and the present thesis, while Nengo does not deal with assemblies.

1.3.2 An informal description of the AC

The Assembly Calculus (AC) [Papadimitriou et al., 2020] is a computational distributed system for modeling a dynamical system of firing neurons. In this system, there is a finite number of *areas*, each containing n neurons. The neurons of an area form a random Erdös-Rényi directed graph $G_{n,p}$, where p is the probability that two neurons of the area are connected. Moreover, certain ordered pairs of areas are connected one to another through an Erdös-Rényi directed bipartite graph $G_{n,p}$. The directed connections in and between areas are called *fibers*.

In the AC, neurons in an area A fire in discrete time steps, and are subject to stylized forms of *inhibition* and *plasticity*. For what concerns inhibition, at any time step, we assume only k_A of the n neurons fire, that is, the ones that previously received the highest total input from all other areas; these k_A neurons are sometimes called the *winners*, as a result of some *winner-take-all* (WTA) competition. Plasticity is modelled by assuming that, if, at a given time step, neuron x fires and, at the next time step, an out-neighbor neuron y of x fires, then the weight of the synapse from x to y (which is 1 at the beginning) is multiplied by $(1 + \beta_A)$, where $\beta_A > 0$. In the original definition of the AC, a process of *homeostasis* was also modelled through a periodic renormalization, at a different time scale, of the synaptic weights, in order to avoid the generation of huge weights. Such process is of course part of any realistic brain system, also providing a mechanism for *forgetting*.

Lastly, yet importantly, the AC allows *inhibiting* and *disinhibiting* areas and fibers at different time steps. The exact mechanism through which areas and fibers are (dis)-inhibited may vary; in a recent paper modeling syntactic processing using the AC, [Mitropolsky et al., 2021] model specific neurons as having (dis)-inhibitory effects on areas or fibers. In this thesis, (dis)-inhibition is always determined by which areas and fibers fired at the previous time step.

The most important emergent object in the AC is the *assembly*, that is, a stable set of k_A highly interconnected neurons in an area A. It is emergent in the sense that assemblies are not a primitive of the model; instead, they are formed through its more basic operations. Assemblies are by now well known and widely studied in neuroscience, and are thought to represent concepts, ideas, objects, words, etc., and are increasingly believed in recent years to play a central role in cognitive processes [Buzsáki, 2010], often called "the alphabet of the brain" [Buzsáki, 2021].

In terms of classical thinking in AI, one could think of assemblies as the boundary in the brain between sub-symbolic and symbolic computation.

The AC makes possible to perform certain *operations* with assemblies, described in Chapter 4: in fact, it is through these operations that assemblies are created, in a way that guarantees high connectivity. In [Papadimitriou et al., 2020], the authors demonstrate, both mathematically and through simulation, that these operations are "possible" in the sense that they can be stably performed with high probability in the dynamical system of neurons outlined in the previous paragraphs. In this thesis, we mostly make use of one of these operations: *projection* of an assembly in an area into another assembly in another area, which works as follows. Let us assume that an assembly x of k_A neurons of the area A has just fired into an area B (presumably through a disinhibited fiber going from A to B), and assume that B was quiescent at that time (no neurons were firing). This will result in a set w_1 of k_B neurons (the winners) firing at the next time step. Next, the neurons in B will receive inputs not only from the k_A neurons of the assembly in A, which will continue to fire, but also from the neurons in w_1 through recurrent connections within B: this will result in a set w_2 of k_B neurons (the new winners) firing at the next time step, and so on. It has been proved that, under appropriate values of the parameters n, k_A, k_B, β, and p, this process converges with high probability to an assembly y of k_B neurons in B, which is called the projection of x into B and can be thought as a copy of x in B such that, from now on, y will fire every time x fires.

In theory, the projection and other operations are emergent behavior of the distributed system as a direct or indirect response to external stimuli; in practice and for simplicity, it is a function we call while coding.

For a complete description of the AC the reader is referred to [Papadimitriou et al., 2020], where in addition to stability of various assembly operations, it is also proved that, under certain assumptions, this computational system is capable of performing arbitrary computations as long as the space required does not exceed $\frac{n}{k_A}$ (under much milder assumptions, $\sqrt{\frac{n}{k_A}}$).

In this thesis, similarly to the parser of [Mitropolsky et al., 2021], the AC programs work by projecting between all pairs of disinhibited areas along disinhibited fibers at each time step. For brevity, this operation, i.e. a simultaneous set of projections between multiple areas, is called *strong projection.* For a detailed description about the AC operations and our programs, we defer the reader to Chapter 4 .

CHAPTER 2

Opinion Dynamics with Uniform Communication Noise

In this chapter we present the analysis of the UNDECIDED-STATE *dynamics and the* 3-MAJORITY *dynamics over complete networks, in the binary opinion case, where a feature of uniform communication noise is introduced in the system.*

2.1 Introduction

In this chapter we analyze two popular opinion dynamics in, the UNDECIDED-STATE dynamics and the 3-MAJORITY dynamics, in presence of uniform communication noise and in the synchronous setting, when the underlying communication network is the complete graph, with two possible opinions.

We prove that both dynamics exhibit very similar phase-transitions: when the noise is below some threshold, the dynamics are able to quickly break any initial symmetry and reach in logarithmic time in the size of the network a metastable phase of almost-consensus which lasts for at least a polynomial number of rounds, with high probability. When, instead, the noise exceeds the threshold, no form of consensus is possible and any information about the initial majority opinion is lost in logarithmic time, with high probability. The UNDECIDED-STATE dynamics turns out to be more resilient to noise as its noise threshold is higher than that of the 3-MAJORITY: this aspect will be further discussed in the final section Section 2.8. Such results are based on our two works [D'Amore et al., 2020] and [D'Amore and Ziccardi, 2022].

These similar behaviors hint at more general characterizations of simple non-linear opinion dynamics that have a strong drift towards the agreement in noiseless settings.

Roadmap. In Section 2.2 we give the main definitions and characterizations of the processes we study, as well as the definition of the noise model. We then state our results in Section 2.3; Section 2.4 is devoted to the analysis of the UNDECIDED-STATE dynamics. In Section 2.5 we show that the UNDECIDED-STATE dynamics turns out to be a specific case of a best-of-N nest site selection process in honeybees [Reina et al., 2017]. The analysis of the 3-MAJORITY dynamics is given in Section 2.6. The chapter concludes with some simulations (Section 2.7) and a discussion on our results and potential research directions (Section 2.8).

2.2 Preliminaries

Opinion dynamics. Let $G = (V, E)$ be a finite graph of n nodes (the agents), where each node is labelled uniquely with labels in $[n] \coloneqq \{1, \dots, n\}$. Furthermore, each node supports an opinion from a set of opinions Σ. An opinion dynamics defines a stochastic process $\{M_t\}_{t\in\mathbb{N}}$ which is described by the opinion of the nodes at each time step, i.e., $M_t = (i_1(t), \dots, i_n(t)) \in \Sigma^n$ for every $t \geq 0$, where $i_j(t)$ is the opinion of node j at time t. For any given node and any given time $t \geq 0$, its state at time $t+1$ is given by some update rule that is a symmetric function of the vector of states of the node neighbors at time t, which is the same for all nodes of the graph and for any time [Emanuele Natale, 2017]. Since M_t depends only on M_{t-1}, it follows that the process is a Markov chain. In the following, we will call the state of the process also by *configuration of the graph*. We remark that the latter definition is a tentative attempt, as it is still possible to come up with very complex dynamics even respecting the above properties; instead, opinion dynamics are characterized by simple update-rules.

The UNDECIDED-STATE dynamics. In the UNDECIDED-STATE dynamics there is an additional state/opinion, i.e., the undecided state, besides the possible opinions a node can support, and the updating rule works as follows: at every round $t \in \mathbb{N}$, each agent u samples a neighbor v

independently and uniformly at random and, at the next round, it gets a new opinion according to the rule given in Table 2.1.

$u \setminus v$	undecided	opinion i	opinion j
undecided	undecided	i	j
opinion i	i	i	undecided
opinion j	j	undecided	j

TABLE 2.1 – The update rule of the USD.

The 3-MAJORITY dynamics. The 3-MAJORITY update rule as follows: given any $t \geq 0$, at time $t+1$ each node $u \in V$ samples three neighbors in G independently uniformly at random (with repetition) and updates its opinion to the majority one among the sampled neighbor opinions at time t, if there is any. Otherwise, it adopts a random opinion among the sampled ones. For the sake of clarity, we remark that when u samples a neighbor node twice, the corresponding opinion counts twice.

The communication noise. We introduce an uniform communication noise feature in the dynamics as follows: let $0 < p < 1$ be a constant. When a node pulls a neighbor opinion, there is probability p that the received opinion is sampled u.a.r. in Σ; instead, with probability $1-p$, the former opinion keeps intact and is received.

Opinion dynamics on the complete graph in the binary setting. The communication network we focus on is the complete graph $G = K_n$ with self loops in the binary opinion case, i.e., $\Sigma = \{\textit{Alpha}, \textit{Beta}\}$. For the symmetry of the network, the state of the process is fully characterized by the number of nodes supporting each of the opinions, which implies that the nodes do not require unique IDs. Hence, we can write $M_t = (a_t, b_t)$, where a_t is the number of the nodes supporting opinion *Alpha* at time t, and b_t is the analogous for opinion *Beta*. Moreover, since at each time t, $a_t + b_t = n$, it suffices to know $\{b_t\}_{t \geq 0}$ to fully describe the process. [1]

We define the bias of the process at time t by

$$s_t = a_t - b_t, \tag{2.1}$$

which takes value in $\{-n, \ldots, n\}$. We remark that $s_t > 0$ if the majority opinion at time t is *Alpha* and $s_t < 0$ if it is *Beta*. We say that configurations having bias $s_t \in \{n, -n\}$ are monochromatic, meaning that every node supports the same opinion, while a configuration with $s_t = 0$ is symmetric. In the introduction, we took the bias to be $|s_t|$ but, for the sake of the analysis, we consider its *signed* version here. We finally remark that, conditional on any configuration at time $t-1$, the random variable a_t (and, analogously, b_t and q_t) is the sum of i.i.d. Bernoulli r.v.s, which allows us to make use of the popular Chernoff bounds (Lemmas B.1 and B.2 and Lemma C.1 in Appendix B). In detail, if $X_i^{(t)}$ is the r.v. yielding 1 if node i adopts opinion *Alpha* at round $t+1$, and 0 otherwise, then $a_t = \sum_{i \in [n]} X_i^{(t)}$. Similarly, we can write $b_t = \sum_{i \in [n]} Y_i^{(t)}$, where $Y_i^{(t)}$ is

1. For the UNDECIDED-STATE dynamics, $M_t = (a_t, b_t, q_t)$, where q_t is the number of nodes that are undecided at time t, and $a_t + b_t + q_t = n$. Thus, the knowledge of at least two variables between $\{a_t, b_t, q_t\}$ is required to describe the configuration of the process.

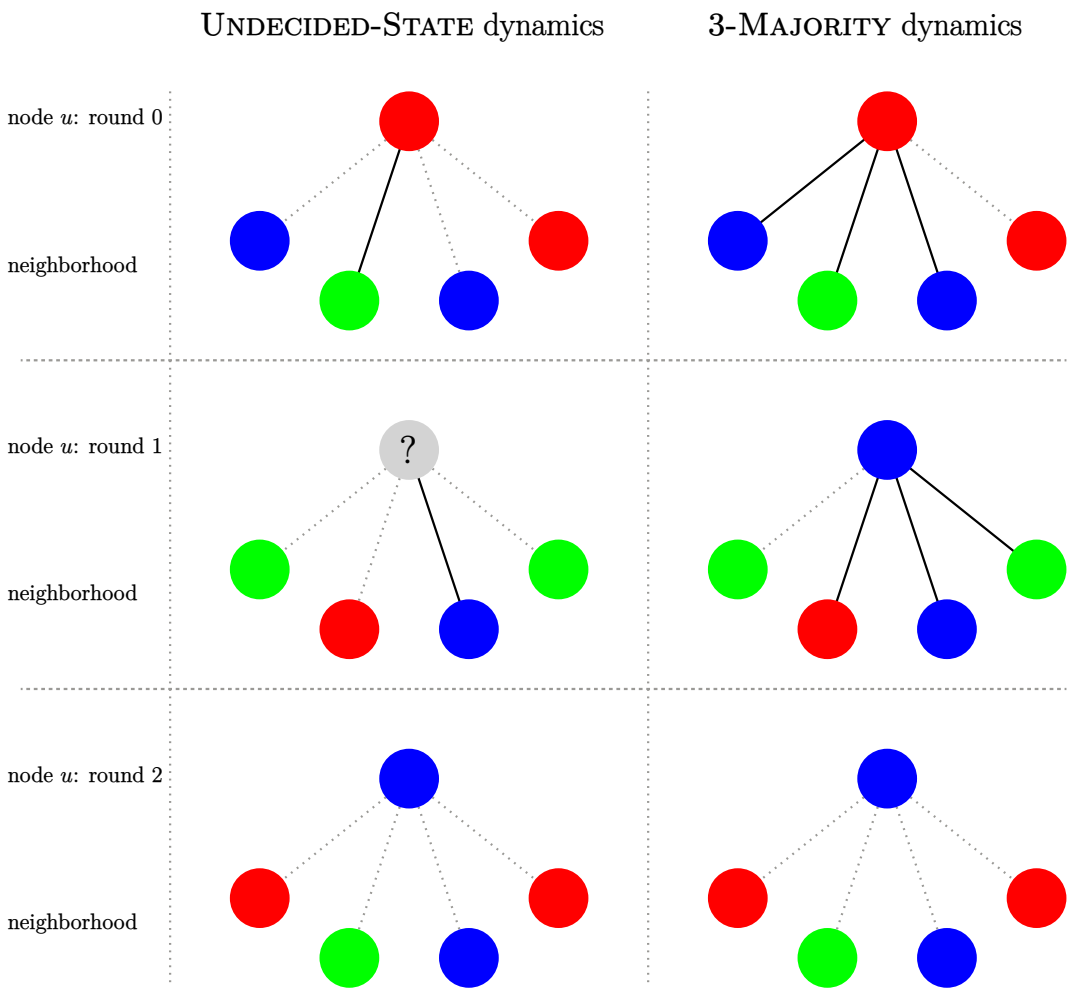


Figure 2.1 – Update rules of the two dynamics shown for a node u in two consequent rounds in which we suppose that the neighbor opinions do not change. The full edges represent the sampled neighbors at each round.

the r.v. yielding 1 if node i adopts opinion *Beta* at round $t+1$, and 0 otherwise. Therefore, for (2.1),

$$s_t = \sum_{i \in [n]} X_i^{(t)} - Y_i^{(t)}, \tag{2.2}$$

where the $(X_i^{(t)} - Y_i^{(t)})_i$ variables are i.i.d. taking values in $\{-1, 1\}$. For this reason, we can apply the Hoeffding bound (Lemma B.3) to the bias. Notice that these dynamics requires no labeling of the agents, i.e., the network can be anonymous.

2.2.1 Oblivious Noise and Stubborn Agents

In this section we show that our noise model is equivalent to a noiseless model in which stubborn agents are added to the graph. We can now consider the following more general message-oblivious model of noise.

Definition 2.2.1 (Oblivious noise model)**.** We say that the communication is affected by oblivious noise if the value of any sent message changes according to the following scheme:

(i) with probability $1 - p_{\text{noise}}$, independently of the value of the sent message, the message remains unchanged;

(ii) otherwise, the noise acts on the message and it changes its value according to a fixed distribution $\mathbf{p} = p_1, ..., p_m$ over the possible message values $1, ..., m$.

In other words, the probability that the noise changes any message to message i is $p_{\text{noise}} \cdot p_i$. It is immediate to verify that the definition of noise adopted in our model corresponds to the aforementioned model of oblivious noise in the special case $m = 3$, $p_{\text{noise}} = p$, and $p_{Alpha} = p_{Beta} = p_{undecided} = \frac{1}{3}$ for the UNDECIDED-STATE dynamics, and the special case $m = 2$, $p_{\text{noise}} = p$, and $p_{Alpha} = p_{Beta} = \frac{1}{2}$ for the 3-MAJORITY dynamics.

Recalling that an agent is said to be stubborn if it never updates its state [Yildiz et al., 2013], we now observe that the process under the above noise model is in fact equivalent to the same process in a noiseless setting with stubborn agents.

Lemma 2.2.1. *Consider any opinion dynamics on the complete graph of n nodes with self-loops, and opinion set $\Sigma = \{1, ..., m\}$. The following two processes are equivalent, i.e., their restrictions on the former graph of n nodes have the same transition probabilities.*

(a) the dynamics in the presence of oblivious noise with parameters p_{noise} and $\mathbf{p} = p_1, ..., p_m$;

(b) the dynamics with $n_{stub} = \frac{p_{noise}}{1-p_{noise}} n$ additional stubborn agents present in the system, of which: $n_{stub} \cdot p_1$ are stubborn agents supporting opinion 1, $n_{stub} \cdot p_2$ are stubborn agents supporting opinion 2, and so on.

Proof. Consider the complete graph of n nodes, K_n, over which the former process runs. Consider also the complete graph $K_{n+n_{stub}}$, which contains a sub-graph isomorphic to K_n which we denote by $\tilde{K}_n$. Let $H = K_{n+n_{stub}} \setminus \tilde{K}_n$ be the subgraph of stubborn nodes. In the former model (a), the probability an agent pulls opinion $j \in \{1, \ldots, m\}$ at any given round is

$$(1 - p_{\text{noise}})\frac{c_j}{n} + p_{\text{noise}} \cdot p_j \ ,$$

where c_j is the size of the community of agents supporting opinion j; in the model defined in (b), the probability a non-stubborn agent pulls opinion j at any given round is

$$\frac{c_j + n_{stub} \cdot p_j}{n + n_{stub}} = \frac{c_j + \frac{p_{\text{noise}}}{1-p_{\text{noise}}} n \cdot p_j}{n + \frac{p_{\text{noise}}}{1-p_{\text{noise}}} n} = (1 - p_{\text{noise}}) \cdot \frac{c_j}{n} + p_{\text{noise}} \cdot p_j \ .$$

Let C and $\tilde{C}$ be the set of all possible configurations of, respectively, K_n and $\tilde{K}_n$. The latter result implies there exists a bijection $\phi : C \to \tilde{C}$ such that the probability to go from configuration M_t to $\mathbf{y}$ in C is the same as that of going from configuration $\phi(M_t)$ to $\phi(\mathbf{y})$ in $\tilde{C}$. □

Basically, this equivalence implies that any result we state for the process defined in (a) has an analogous statement for the process defined in (b).

2.3 Results

In this section we state our main theorems about the phase-transitions of the UNDECIDED-STATE and the 3-MAJORITY dynamics. We remark that, for any two positive functions $g(n)$, $f(n)$,

and for any positive constant c, by $g(n) = \mathcal{O}_c(f(n))$ we mean that there exists a constant $h(c) > 0$ depending on c such that $g(n) \leq h(c)f(n)$ for n large enough. We will also write $\Omega_c(f(n))$ with an analogous meaning.

Phase-transition of the UNDECIDED-STATE dynamics. The first theorem shows how the dynamics solves the majority consensus problem when the noise is below some threshold, even if in a "weak" form (since only an almost-consensus is reached). Section 2.4.3 is devoted to the proof of this theorem.

Theorem 2.3.1 (Victory of the majority)**.** *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* UNDECIDED-STATE *dynamics with uniform noise probability $p < 1/2$. For any $\varepsilon > 0$, let $\check{s}_\varepsilon = \frac{n}{1-p}\sqrt{(1-2p)(1-2p/3) - \varepsilon}$ and $\hat{s}_\varepsilon = \frac{n}{1-p}\sqrt{(1-2p)(1-2p/3) + \varepsilon}$. Furthermore, let $\varepsilon_1 = (1-2p)^2/2$, $\varepsilon_2 = 2p(3-p)/9$, and let $\gamma > 0$ be any constant. For any starting configuration M_0 having bias $s_0 \geq \gamma\sqrt{n \log n}$, the following holds w.h.p.:*

(i) there exists a time $\tau_1 = \mathcal{O}_{p,\gamma}(\log n)$ such that $\check{s}_{\varepsilon_1}(1-2p/3)^2 \leq s_{\tau_1} \leq \hat{s}_{\varepsilon_2}$;

(ii) there exists a value $c = \Theta_p(1)$ such that, for all $k \leq exp(cn)$, $\check{s}_{\varepsilon_1}(1-2p/3)^2 \leq s_{\tau_1+k} \leq \hat{s}_{\varepsilon_2}$.

A symmetric result for the case $s_0 \leq -\gamma\sqrt{n \log n}$ holds in the same way. Next theorem characterizes the symmetry-breaking phase and is proved in Section 2.4.4.

Theorem 2.3.2 (Symmetry breaking)**.** *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* UNDECIDED-STATE *dynamics with uniform noise probability $p < 1/2$, and let $\gamma > 0$ be any positive constant. Then, for any starting configuration M_0 such that $|s_0| \leq \gamma\sqrt{n \log n}$, w.h.p. there exists a time $\tau_2 = \mathcal{O}_{\gamma,p}(\log n)$ such that $|s_{\tau_2}| \geq \gamma\sqrt{n \log n}$.*

The next and final theorem shows the victory of noise when the noise parameter is above the threshold, even starting from a monochromatic configuration. Section 2.4.5 is devoted to the proof of this theorem.

Theorem 2.3.3 (Victory of noise)**.** *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* UNDECIDED-STATE *dynamics with uniform noise probability $p > 1/2$. Let $\gamma > 0$ be any positive constant. Then, for any starting configuration M_0 such that $|s_0| \geq \gamma\sqrt{n \log n}$, the following holds w.h.p. :*

(i) there exists a time $\tau_3 = \mathcal{O}_{\gamma,p}(\log n)$ such that $|s_{\tau_3}| \leq \gamma\sqrt{n \log n}$;

(ii) there exists a value $c = \Theta_{\gamma,p}(1)$ such that, for all $k \leq n^c$, it holds that $|s_{\tau_3+k}| \leq \gamma\sqrt{n \log n}$.

Informally, although the analysis is technically complex, it can be appreciated from it that the phase transition phenomenon in question is ultimately based on the exponential drift of the dynamics toward the majority opinion in the absence of noise: as long as the noise is kept under a certain threshold, the dynamics manages to quickly amplify and sustain the bias towards the majority opinion; as soon as the noise level reaches the threshold, the expected increase of the majority bias abruptly decreases below the standard deviation of the process and the ability of the dynamics to preserves a signal towards the initial majority rapidly vanishes.

Phase-transition of the 3-MAJORITY dynamics. The 3-MAJORITY dynamics has a very similar phase-transition but turns out to be less resilient to noise than the UNDECIDED-STATE dynamics. Next theorem shows the victory of the majority whenever the noise is below some threshold, and it is proved in Section 2.6.1.

Theorem 2.3.4 (Victory of the majority). *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* 3-MAJORITY *dynamics with uniform noise probability $p < 1/3$. Let $\varepsilon > 0$ be any arbitrarily small constant (such that $\varepsilon < 1/3$ and $\varepsilon^2 \leq (1-3p)/2$) and let $\gamma > 0$ be any constant. Let $s_{eq} = \frac{n}{(1-p)}\sqrt{\frac{1-3p}{1-p}}$. Then, for any starting configuration M_0 such that $s_0 \geq \gamma\sqrt{n\log n}$ the following holds w.h.p. :*

(i) there exists a time $\tau_1 = \mathcal{O}_{\gamma,\varepsilon,p}(\log n)$ such that $(1-\varepsilon)s_{eq} \leq s_{\tau_1} \leq (1+\varepsilon)s_{eq}$;

(ii) there exists a value $c = \Theta_{\gamma,\varepsilon,p}(1)$ such that, for all $k \leq n^c$, $(1-\varepsilon)s_{eq} \leq s_{\tau_1+k} \leq (1+\varepsilon)s_{eq}$.

Our second theorem shows how the dynamics is capable of quickly breaking the initial symmetry. By applying also Theorem 2.3.4, it shows that the consensus problem is solved. The proof of the theorem is shown in Section 2.6.2.

Theorem 2.3.5 (Symmetry breaking). *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* 3-MAJORITY *dynamics with uniform noise probability $p < 1/3$, and let $\gamma > 0$ be any positive constant. Then, for any starting configuration M_0 such that $|s_0| \leq \gamma\sqrt{n\log n}$ and for any sufficiently large n, w.h.p. there exists a time $\tau_2 = \mathcal{O}_{\gamma,p}(\log n)$ such that $|s_{\tau_2}| \geq \gamma\sqrt{n\log n}$.*

Our last theorem shows that no form of consensus is possible when $p > 1/3$, and it is proved in Section 2.6.3.

Theorem 2.3.6 (Victory of noise). *Let $\{M_t\}_{t\geq 0}$ be the process induced by the* 3-MAJORITY *dynamics with uniform noise probability $p > 1/3$. Let $\varepsilon > 0$ be any arbitrarily small constant (such that $\varepsilon < \min\{1/4, (1-p), (3p-1)/2\}$) and let $\gamma > 0$ be any positive constant. Then, for any starting configuration M_0 such that $|s_0| \geq \gamma\sqrt{n\log n}$ and for any sufficiently large n, the following holds w.h.p. :*

(i) there exists a time $\tau_3 = \mathcal{O}_{\varepsilon,p}(\log n)$ such that $s_{\tau_3} = \mathcal{O}_{\varepsilon}(\sqrt{n})$ and, moreover, the majority opinion switches at the next round with probability $\Theta_{\varepsilon}(1)$;

(ii) there exists a value $c = \Theta_{\gamma,\varepsilon}(1)$ such that, for all $k \leq n^c$, it holds that $|s_{\tau_3+k}| \leq \gamma\sqrt{n\log n}$.

Stubborn process. All previous theorems have analogous statements in the stubborn model according to Lemma 2.2.1. In particular, the UNDECIDED-STATE dynamics exhibits a phase transition at $p = 1/2$ while the 3-MAJORITY dynamics at $p = 1/3$.

2.4 Analysis of the UNDECIDED-STATE dynamics

In this section we analyze the UNDECIDED-STATE dynamics. In Section 2.4.1 we define some notation and give some preliminary results for our analysis. Then, in Section 2.4.3 we prove Theorem 2.3.1; in Section 2.4.4 we prove Theorem 2.3.2; finally, Section 2.4.5 is devoted to the proof of Theorem 2.3.3.

2.4.1 Notation, Characterization, and Mean-Field Behavior.

In this subsection, we characterize the mean-field behavior of the UNDECIDED-STATE dynamics. To do so, we introduce some notation and describe useful properties of the process. Let us call UNDECIDED-STATE process the process induces by the UNDECIDED-STATE dynamics with uniform communication noise starting from any given configuration. The UNDECIDED-STATE

process at time t, M_t, is uniquely determined by the number of *Alpha* nodes, a_t and the number of *Beta* nodes, b_t. Accordingly to this notation, we call q_t the number of undecided nodes at time t, and $s_t = a_t - b_t$ the *bias* of the configuration at time t. M_t is thus a finite-state non-reversible Markov chain with no absorbing states. Once a configuration at time $t-1 \geq 0$ is fixed, we use lower case letters a, b, q, s to refer to random variables a_{t-1}, b_{t-1}, q_{t-1}, and s_{t-1}. Notice that we consider the bias as $a-b$ instead of $|a-b|$ since the expectation of $|a_t - b_t|$ is much more difficult to evaluate than that of $a_t - b_t$.

The round-by-round expected behaviors of the above key random variables is described by the following equations:

$$\mathbb{E}\left[a_t \mid M_{t-1}\right] = \frac{a}{n}(1-p)(a+2q) + \frac{p}{3}(2a+q); \tag{2.3}$$

$$\mathbb{E}\left[b_t \mid M_{t-1}\right] = \frac{b}{n}(1-p)(b+2q) + \frac{p}{3}(2b+q); \tag{2.4}$$

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] = s\left(1 - \frac{p}{3} + \frac{q}{n}(1-p)\right); \tag{2.5}$$

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] = \frac{pn}{3} + \frac{1-p}{2n}\left[2q^2 + (n-q)^2 - s^2\right]. \tag{2.6}$$

Proof of Equations (2.3) *to* (2.6). Conditional on any configuration at time $t-1$, the probability an agent pulls opinion *Alpha* at the next round is

$$(1-p)\frac{a}{n} + \frac{p}{3},$$

and symmetrical expressions hold for opinion *Beta* and the undecided state. An agent updates its opinion to *Alpha* at time t if it is undecided at time $t-1$ and pulls opinion *Alpha*, or if it supports opinion *Alpha* at time t and pulls either opinion *Alpha* or the undecided state. If V_q and V_a denote the sets of agents supporting the undecided state and opinion *Alpha*, respectively, at time t, we have

$$\begin{aligned}
\mathbb{E}\left[a_t \mid M_{t-1}\right] &= \sum_{u \in V_a}\left((1-p)\frac{a}{n} + \frac{p}{3} + (1-p)\frac{q}{n} + \frac{p}{3}\right) + \sum_{u \in V_q}\left((1-p)\frac{a}{n} + \frac{p}{3}\right) \\
&= a\left((1-p)\frac{a}{n} + \frac{p}{3} + (1-p)\frac{q}{n} + p\right) + q\left((1-p)\frac{a}{n} + \frac{p}{3}\right) \\
&= (1-p)\frac{a}{n}(a+2q) + \frac{p}{3}(2a+q).
\end{aligned}$$

Similarly, we get the conditional expectation of B. Then

$$\begin{aligned}
\mathbb{E}\left[s_t \mid M_{t-1}\right] &= \mathbb{E}\left[a_t \mid M_{t-1}\right] - \mathbb{E}\left[b_t \mid M_{t-1}\right] \\
&= \frac{1-p}{n}\left[(a-b)(a+b) + 2q(a-b)\right] + \frac{2p}{3}(a-b) \\
&= \frac{1-p}{n}\left[s(n-q) + 2qs\right] + \frac{2ps}{3} = s\left(1 - \frac{p}{3} + (1-p)\frac{q}{n}\right) \\
&= s\left(1 - \frac{p}{3} + (1-p)\frac{q}{n}\right), \\
\mathbb{E}\left[q_t \mid M_{t-1}\right] &= + \frac{a}{n}\left[\frac{p}{3}(a+q) + \left(1 - \frac{2p}{3}\right)b\right] + \frac{b}{n}\left[\frac{p}{3}(b+q) + \left(1 - \frac{2p}{3}\right)a\right] \\
&\quad + \frac{q}{n}\left[\frac{p}{3}(a+b) + (1 - \frac{2p}{3})q\right] \\
&= \frac{p}{3n}\left[a^2 + b^2 + 2q(a+b)\right] + \frac{1 - \frac{2p}{3}}{n}\left[2ab + q^2\right] \\
&= \frac{p}{3}n + \frac{1-p}{n}\left[2ab + q^2\right] = \frac{p}{3}n + \frac{1-p}{2n}\left[2q^2 + (n-q)^2 - s^2\right].
\end{aligned}$$

□

2.4.2 Equilibria in expectation.

Notice that, from Eq. (2.5) the bias keeps constant in expectation iff $s = 0$ or

$$\begin{aligned}
& 1 - \frac{p}{3} + \frac{q}{n}(1-p) = 1 \\
\iff & q = \frac{pn}{3(1-p)}.
\end{aligned}$$

If we substitute the latter value in Eq. (2.6) we get that q keeps constant, in expectation, iff

$$\begin{aligned}
& \frac{pn}{3} + \frac{1-p}{2n}\left[\frac{2p^2n^2}{9(1-3p)^2} + \frac{(3-4p)^2n^2}{9(1-3p)^2} - s^2\right] = \frac{pn}{3(1-p)} \\
\iff & s^2 \cdot \frac{1-p}{2n} = \frac{2p^2n + (3-4p)^2n - 6p^2}{18(1-3p)} \\
\iff & s^2 = \frac{n^2}{3(1-p)^2} \cdot (1-2p)(3-2p) \\
\iff & s = \pm\frac{n}{1-p}\sqrt{\frac{(1-2p)(3-2p)}{3}}.
\end{aligned}$$

Let

$$s_{\text{eq}} = \frac{n}{1-p}\sqrt{\frac{(1-2p)(3-2p)}{3}} \quad \text{and} \quad q_{\text{eq}} = \frac{pn}{3(1-p)}.$$

The configuration with $s = \pm s_{\text{eq}}$ and $q = q_{\text{eq}}$ is in equilibrium, in expectation. Theorems 2.3.1 and 2.3.2 show that, if $p < 1/2$, the latter is an "attractive" configuration, i.e., the process converges and oscillates in an interval of width $\Theta_p n$ around the above configuration w.h.p., no

matter what the initial parameters are. Instead, Theorem 2.3.3 shows taht if $p > 1/2$, the process converges to a configuration with $s = \mathcal{O}\left(\sqrt{n \log n}\right)$ w.h.p.

For the sake of the analysis, we define the following quantities for any $\varepsilon > 0$:

$$\begin{aligned}
\check{s}_\varepsilon &= \frac{n}{1-p}\sqrt{\frac{(1-2p)(3-2p)}{3} - \varepsilon}; \\
\hat{s}_\varepsilon &= \frac{n}{1-p}\sqrt{\frac{(1-2p)(3-2p)}{3} + \varepsilon}; \\
\check{q}_\varepsilon &= \frac{\left(\frac{p}{3} - \varepsilon\right) n}{1-p}; \\
\hat{q}_\varepsilon &= \frac{\left(\frac{p}{3} + \varepsilon\right) n}{1-p}.
\end{aligned}$$

2.4.3 UNDECIDED-STATE dynamics victory of the majority

In this section we prove Theorem 2.3.1. Wlog, in the remainder, for a given configuration M_t, we will assume $a \geq b$. Indeed, as it will be clear from the results, if $s = a - b \geq \gamma\sqrt{n \log n}$, then the plurality opinion does not change for $\exp\left(\Omega_p(n)\right)$ rounds, w.h.p., and the argument for the case $b > a$ is symmetric. Let $p < 1/3$. The key point to prove the first claim of the theorem is to show that, if the bias of the configuration is less than $\check{s}_\varepsilon$, and the number of undecided nodes is above $\hat{q}_\varepsilon$, then the bias at the next round increases by a constant factor and the number of undecided nodes remains greater than $\hat{q}_\varepsilon$, w.h.p.

Lemma 2.4.1. *Let M_{t-1} be a configuration such that $q \geq \hat{q}_\varepsilon$ and $s \geq 0$ for any arbitrarily small constant $\varepsilon > 0$. Then, in the next round, $s_t \geq s\,(1 + \varepsilon/2)$, with probability $1 - \exp\left(-\varepsilon^2 s^2/(2^3 n)\right)$.*

Remark 2.4.1 – If $s \geq \gamma\sqrt{n \log n}$ for any constant $\gamma > 0$, then the statement holds with probability $1 - n^{-\varepsilon^2\gamma^2/2^3}$.

Proof of Lemma 2.4.1. We first notice that $q \geq \hat{q}_\varepsilon$ in eq. (2.5) implies

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq s\,(1 + \varepsilon)\,.$$

Let $\lambda = (\varepsilon/2) \cdot (s/n)$. Then, for the Hoeffding bound (Lemma B.3).

$$\begin{aligned}
\Pr\left[s_t \geq s\left(1 + \frac{\varepsilon}{2}\right) \,\middle|\, M_{t-1}\right] &= \Pr\left[s_t \geq s\,(1+\varepsilon) - \lambda n \mid M_{t-1}\right] \geq 1 - \exp\left(-\frac{2\lambda^2 n}{4}\right) \\
&= 1 - \exp\left(-\frac{\varepsilon^2 s^2}{2^3 n}\right).
\end{aligned}$$

□

Lemma 2.4.2. *Let ε be any constant with $(1-2p)^2/2 \leq \varepsilon < (1-2p)(1-2p/3)$. Let M_{t-1} be a configuration such that $s \leq \check{s}_\varepsilon$. Then, in the next round, $q_t \geq \hat{q}_{\varepsilon/12}$ with probability $1 - \exp\left(-\varepsilon^2 n/\left(2^3 3^2 (1-p)^2\right)\right)$*

Proof of Lemma 2.4.2. Notice that eq. (2.6) reaches its minimum in $q = n/3$. Combined with $s \leq \check{s}_\varepsilon$, we get

$$\begin{aligned}
\mathbb{E}\left[q_t \mid M_{t-1}\right] &\geq \frac{pn}{3} + \frac{1-p}{2n}\left[\frac{2n^2}{3} - \frac{(1-2p)(3-2p)n^2 - 3\varepsilon n^2}{3(1-p)^2}\right] \\
&= \frac{pn}{3} + \frac{2(1-p)^2 n - (1-2p)(3-2p)n + 3\varepsilon n}{6(1-p)} \\
&\geq \frac{2pn - 2p^2 n + 2n - 4pn + 2p^2 n - 3n + 6pn - 4p^2 n + 3\varepsilon n}{6(1-p)} \\
&= \frac{-4p^2 n + 6pn - n + 3\varepsilon n}{6(1-p)} = \frac{-(1-2p)^2 n + 2pn + 3\varepsilon n}{6(1-p)}.
\end{aligned}$$

Now, for $(1-2p)^2/2 \leq \varepsilon < (1-2p)(1-2p/3)$, we get

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \geq \frac{-(1-2p)^2 n + 2pn + 3\varepsilon n}{6(1-p)} \geq \frac{(2p+\varepsilon)n}{6(1-p)} = \frac{p + \varepsilon/6}{1-p} n = \hat{q}_{\frac{\varepsilon}{6}}.$$

By the additive form of Chernoff bound (Lemma B.2), we get

$$\Pr\left[q_t \leq \hat{q}_{\frac{\varepsilon}{12}} \;\middle|\; M_{t-1}\right] = \Pr\left[q_t \leq \hat{q}_{\frac{\varepsilon}{6}} - \frac{\varepsilon}{12(1-p)} n \;\middle|\; M_{t-1}\right] \leq \exp\left(-\frac{\varepsilon^2 n}{2^3 3^2 (1-p)^2}\right).$$

□

As we will see at the end of this section, the two lemmas above ensure that the system eventually reaches a configuration with bias $s > \check{s}_\varepsilon$ within $\mathcal{O}(\log n)$ rounds, w.h.p. We now consider configurations in which $s > \check{s}_\varepsilon$ and derive a useful bound on the possible decrease of s.

Lemma 2.4.3. *Let M_{t-1} be any configuration such that $s \geq \gamma\sqrt{n \log n}$ for any constant $\gamma > 0$. Then, in the next round, it holds that $s_t \geq s\,(1 - 2p/3)$ with probability $1 - \exp\left(-p^2 s^2/(2 \cdot 3^2 n)\right)$.*

Remark 2.4.2 – If $s \geq \gamma\sqrt{n \log n}$ for any constant $\gamma > 0$, then the statement holds with probability $1 - n^{-p^2\gamma^2/(2\cdot 3^2)}$.

Proof of Lemma 2.4.3. From eq. (2.5), we get

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq s(1 - p/3).$$

Let $\lambda = p \cdot (s/3n)$. Then, for the Hoeffding bound (Lemma B.3),

$$\Pr\left[s_t \geq s\,(1-2p/3) \mid M_{t-1}\right] = \Pr\left[s_t \geq s(1-p/3) - \lambda n \mid M_{t-1}\right] \geq 1 - \exp\left(-\frac{p^2 s^2}{2 \cdot 3^2 n}\right).$$

□

Lemma 2.4.3 is used to control the decrease of the bias, showing that it keeps of magnitude $\Theta\left(n\right)$ for at least a polynomial number of rounds. The next two lemmas provide an upper bound on the bias during this phase.

Lemma 2.4.4. *Let M_{t-1} be any configuration, and let $\varepsilon = (p/6)\cdot(1+p)$. Then, in the next round, $q_t \geq \check{q}_\varepsilon$, with probability $1 - \exp\left(-p^2 n/2\cdot 3^2\right)$.*

Proof of Lemma 2.4.4. From Eq. (2.6), we get

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \geq \frac{pn}{3} + \frac{1-p}{2n}\left[2q^2 + (n-q)^2 - (n-q)^2\right] = \frac{pn}{3} + \frac{(1-p)q^2}{n} \geq \frac{pn}{3}.$$

where we used that $s \leq n - q$. Let $\lambda = pn/6$. For the additive form of Chernoff bound (Lemma B.2), we get

$$\Pr\left[q_t \geq \frac{pn}{6} \,\middle|\, M_{t-1}\right] = \Pr\left[q_t \geq pn/3 - \lambda \mid M_{t-1}\right] \geq 1 - \exp\left(-\frac{2\lambda^2}{n}\right) = 1 - \exp\left(-\frac{p^2 n}{18}\right).$$

We conclude the proof by observing that

$$\frac{(p/3-\varepsilon)n}{1-p} = \frac{p(1-p)n}{6(1-p)} = \frac{pn}{6}.$$

□

Lemma 2.4.5. *Let M_{t-1} be a configuration with $q \geq \check{q}_{\varepsilon_1}$, for any constant $0 < \varepsilon_1 < p/3$. Let $\varepsilon_2 = 2p(3-p)/9$. Then, in the next round, $s_t \leq \hat{s}_{\varepsilon_2}$ with probability $1 - \exp\left(-(p^2/2\cdot 3^2)\cdot(p/3-\varepsilon_1)^2 n/(1-p)^2\right)$.*

Proof of Lemma 2.4.5. From Eq. (2.4), we get that

$$\mathbb{E}\left[b_t \mid M_{t-1}\right] \geq \frac{pq}{3} \geq \frac{p}{3}\check{q}_{\varepsilon_1} = \frac{p(p/3-\varepsilon_1)n}{3(1-p)}.$$

Let $\lambda = (p/12)\cdot(p/3-\varepsilon_1)n/(1-p)$. Then, Lemma B.2 implies that

$$\begin{aligned}\Pr\left[b_t \geq \frac{p(p/3-\varepsilon_1)n}{6(1-p)} \,\middle|\, M_{t-1}\right] &= \Pr\left[b_t \geq \frac{p(p/3-\varepsilon_1)n}{3(1-p)} - \lambda \,\middle|\, M_{t-1}\right] \geq 1 - \exp\left(-\frac{2\lambda^2}{n}\right) \\ &= 1 - \exp\left(-\frac{p^2(p/3-\varepsilon_1)^2 n}{2\cdot 3^2(1-p)^2}\right).\end{aligned}$$

Notice that $s_t \leq n - 2b_t$. Since

$$n - \frac{p(p/3-\varepsilon_1)n}{3(1-p)} \leq \check{s}_{\varepsilon_2},$$

for $\varepsilon_2 = 2p(3-p)/9$, then $s_t \leq \hat{s}_{\varepsilon_2}$ with probability $1-\exp\left(-(p^2/2\cdot 3^2)\cdot(p/3-\varepsilon_1)^2 n/(1-p)^2\right)$.

□

Proof of Theorem 2.3.1. Suppose the initial configuration bias is outside the interval $I = [\check{s}_{\varepsilon_1}, \hat{s}_{\varepsilon_2}]$. By Lemmas 2.4.4 and 2.4.5, we have that in two rounds the bias is less than or equal to $\hat{s}_{\varepsilon_2}$ with probability $1 - 2\exp\left(-p^4 n/2^3 3^4\right)$ for the union bound, and keeps bounded by the same value for $T = \exp\left(-p^4 n/2^4 3^4\right)$ rounds with probability $1 - \exp\left(-\Omega_p(n)\right)$ for Lemmas A.1 and A.2.

Lemma 2.4.3 and Remark 2.4.2 imply that the bias at the second round is no less than $\gamma'\sqrt{n\log n}$ with probability $1-n^{-p^2\gamma^2/2\cdot 3^2}$, with $\gamma'=\gamma(1-2p/3)$. Define the event

$$\begin{aligned}
\mathcal{B}_k &= \left\{ s_k \geq s_{k-1}\left(1+\frac{\varepsilon_1}{2\cdot 12}\right)\right\};\\
\mathcal{C}_k &= \left\{ s_k \geq \gamma'\sqrt{n\log n}\right\} \cap \left\{ q_{k-1} \geq \hat{q}_{\varepsilon_1/12}\right\};\\
\mathcal{D}_k &= \{ s_k \geq \check{s}_{\varepsilon_1}\}\\
\mathcal{E}_k &= \mathcal{B}_k \cap \mathcal{C}_k.
\end{aligned}$$

Notice that, conditional on $\mathcal{E}_{k-1}\cap\mathcal{D}_{k-1}^C$, the probability of $\mathcal{E}_{k+1}$ is $1-3n^{-(1/2^3)(\varepsilon_1 ps)^2/3^2 n}$ for Lemmas 2.4.1 to 2.4.3, Remarks 2.4.1 and 2.4.2, and the union bound, where ε_1 depends only on p. Let $\tau^*=\min\{k\geq 0 : s_k \geq \check{s}_{\varepsilon_1}\}$. Notice that $\cap_{k=1}^{T_2}(\mathcal{F}_k\cup\mathcal{D}_k)\subseteq\{\tau^*\leq T_2\}$, with $T_2=\log n/\log(1+\varepsilon_1/24)$. Then, by Lemmas A.1 and A.2, we have that $\tau^*\leq T_2$ with probability $1-n^{\Omega_{p,\gamma}(1)}$. Let τ be the first time s_t lies in the interval $I=[\check{s}_{\varepsilon_1},\hat{s}_{\varepsilon_2}]$. For what we showed above, $\tau=\max\{2,\tau^*\}$, which is at most T_2 with high probability for the union bound.

Suppose now we start in a configuration inside the interval I. Then, the bias keeps bounded by $\hat{s}_{\varepsilon_2}$ for time T_1-2 with probability $1-\exp(-\Omega_p(n))$. Furthermore, it can decreases to, at most, $\check{s}_{\varepsilon_1}(1-2p/3)^2$ for Lemma 2.4.3 with probability at least $1-\exp(-\Omega_p(n))$ before starting to increase again towards interval I with probability at least $1-\exp(-\Omega_p(n))$. Thus, the bias keeps inside the interval Δ for $\exp(\Omega_p(n))$ rounds, w.h.p. for the chain rule. □

2.4.4 UNDECIDED-STATE dynamics: symmetry-breaking

In this section, we consider the UNDECIDED-STATE process starting from arbitrary initial configurations: in particular, from configurations having no bias. Informally, Theorem 2.3.2 states that when $p<1/3$, the UNDECIDED-STATE process is able to break the symmetry of any perfectly-balanced initial configuration; then, Theorem 2.3.1 applies and the system computes almost consensus within $O(\log n)$ rounds, w.h.p.

Interestingly enough, Theorem 2.3.3 (proved in 2.4.5) implies that the same threshold $p=1/3$ for computing almost-consensus is sharp as higher values of noise prevent any strong bias drift. What follows is an outline of the proof of the theorem, while more details are given in the next section.

Outline of Proof of Theorem 2.3.2. If the initial configuration M_0 has bias $s=\Omega(\sqrt{n\log n})$ then the claim of the theorem is equivalent to that of Theorem 2.3.1, so we are done. Hence, we next assume the initial bias s be $o(\sqrt{n\log n})$: for this case, our proof proceeds along the following main steps.

Step I. Whenever the bias s is small, i.e., $o(n)$, we prove that, within the next $\mathcal{O}(\log n)$ rounds, the number of undecided nodes turns out to keep always in a suitable linear range: roughly speaking, we get that this number lies in $(n/3,n/2]$, w.h.p.

Step II. Whenever s is very small, i.e., $s=o(\sqrt{n})$, there is no effective drift towards any opinion. However, we can prove that, thanks to Step I, the random variable S, representing the bias in the next round, has *high variance*, i.e., $\Theta(n)$. The latter holds since S can be written as a suitable sum whose addends include some random variables having binomial distribution of expectation 0: so, we can apply the Berry-Esseen result (Lemma B.6) to get a lower bound on the variance of S. Then, thanks to this large variance, classical arguments for the standard deviation imply

that, in this parameter range, there is a positive constant probability that S will get some value of magnitude $\Omega(\sqrt{n})$ (see Claim 1 of Lemma 2.4.8). Not surprisingly, in this phase, we find out that the variance of S is not decreased by the communication noise. We can thus claim that the process, at every round, has positive constant probability to reach a configuration having bias $s = \omega(\sqrt{n})$ and $q \in (n/3, n/2]$. Then, after $O(\log n)$ rounds, this event will happen w.h.p.
Step III. Once the process reaches a configuration with $s = \omega(\sqrt{n})$ and $q \in (n/3, n/2]$, we then prove that the expected bias increases by a constant factor (which depends on p). Observe that we cannot use here the same round-by-round concentration argument that works for bias over $\sqrt{n \log n}$ (this is in fact the minimal magnitude required to apply the Chernoff's bounds [Dubhashi and Panconesi, 2009]). We instead exploit a useful general tool [Clementi et al., 2018] that bounds the stopping time of some class of Markov chains having rather mild conditions on the drift towards their absorbing states (see Lemma 2.4.6). This tool in fact allows us to consider the two phases described, respectively, in Step II and Step III as a unique *symmetry-breaking* phase of the process. Our final technical contribution here is to show that the conditions required to apply this tool hold whenever the communication noise parameter is such that $p \in (0, 1/3)$. This allows us to prove that, within $O(\log n)$ rounds, the process reaches a configuration with bias $s = \Omega(\sqrt{n} \log n)$, w.h.p. □

2.4.4.1 Proofs: symmetry breaking

The proof of Theorem 2.3.2 essentially relies on the following lemma which has been proved in [Clementi et al., 2018], for which we report a proof that corrects some minor mistakes.

Lemma 2.4.6. *Let $\{X_t\}_{t\in\mathbb{N}}$ be a Markov Chain with finite-state space Ω and let $f : \Omega \mapsto [0, n]$ be a function that maps states to integer values. Let c_3 be any positive constant and let $m = c_3\sqrt{n}\log n$ be a target value. Assume the following properties hold:*

(1) for any positive constant h, a positive constant $c_1 < 1$ exists such that for any $x \in \Omega$: $f(x) < m$,

$$\Pr\left[f(X_{t+1}) < h\sqrt{n} \mid X_t = x\right] < c_1;$$

(2) there exist two positive constants δ and c_2 such that for any $x \in \Omega : h\sqrt{n} \le f(x) < m$,

$$\Pr[f(X_{t+1}) < (1+\delta)f(X_t) \mid X_t = x] < e^{-c_2 f(x)^2/n}.$$

Then the process reaches a state x such that $f(x) \ge m$ within $\mathcal{O}_{c_1,c_3}(\log n)$ rounds with probability at least $1 - 2/n$.

Proof. Define a set of hitting times $T := \{\tau(i)\}_{i\in\mathbb{N}}$, where

$$\tau(i) = \inf_{i\in\mathbb{N}} \{t : t > \tau(i-1), f(X_t) \ge h\sqrt{n}\},$$

setting $\tau(0) = 0$. By the first hypothesis, for every $i \in \mathbb{N}$, the expectation of $\tau(i)$ is finite, which implies $\Pr[\tau(i) < +\infty] = 1$. Hence, $\tau(i)$ is well defined for any $i \in \mathbb{N}$ and is a stopping time. Now, define the following stochastic process which is a subsequence of $\{X_t\}_{t\in\mathbb{N}}$:

$$\{R_i\}_{i\in\mathbb{N}} = \{X_{\tau(i)}\}_{i\in\mathbb{N}}.$$

Observe that $\{R_i\}_{i\in\mathbb{N}}$ is still a Markov chain as the Markov property holds. Indeed, if $\{x_1,\dots,X_{i-1}\}$ is a set of states in Ω, then

$$\begin{aligned}
&\Pr\left[R_i = x \mid R_{i-1} = x_{i-1},\dots,R_1 = x_1\right]\\
&= \Pr\left[X_{\tau(i)} = x \;\middle|\; X_{\tau(i-1)} = x_{i-1},\dots,X_{\tau(1)} = x_1\right]\\
&= \sum_{t(i)>\cdots>t(1)\in\mathbb{N}} \Pr\left[X_{t(i)} = x \;\middle|\; X_{t(i-1)} = x_{i-1},\dots,X_{t(1)} = x_1\right]\\
&\quad\cdot \Pr\left[\tau(i) = t(i),\dots,\tau(1) = t(1)\right]\\
&\overset{(a)}{=} \Pr\left[X_{\tau(i)} = x \;\middle|\; X_{\tau(i-1)=x_{i-1}}\right]\\
&= \Pr\left[R_i = x \mid R_{i-1} = x_{i-1}\right],
\end{aligned}$$

where (a) holds for the Strong Markov property. By definition, the state space of R is $\{x \in \Omega : f(x) \geq h\sqrt{n}\}$. Moreover, the second hypothesis still holds for this new Markov chain. Indeed:

$$\begin{aligned}
&\Pr\left[f(R_{i+1} < (1+\epsilon)f(R_i) \mid R_i = x\right]\\
&= 1 - \Pr\left[f(R_{i+1} \geq (1+\epsilon)f(R_i) \mid R_i = x\right]\\
&= 1 - \Pr\left[f(X_{\tau(i+1)} \geq (1+\epsilon)f(X_{\tau(i)}) \;\middle|\; X_{\tau(i)} = x\right]\\
&\leq 1 - \Pr\left[f(X_{\tau(i+1)} \geq (1+\epsilon)f(X_{\tau(i)}), \tau(i+1) = \tau(i)+1 \;\middle|\; X_{\tau(i)} = x\right]\\
&= 1 - \Pr\left[f(X_{\tau(i)+1} \geq (1+\epsilon)f(X_{\tau(i)}) \;\middle|\; X_{\tau(i)} = x\right]\\
&= 1 - \Pr\left[f(X_{t+1} \geq (1+\epsilon)f(X_t) \mid X_t = x\right]\\
&< e^{-c_2 f(x)^2/n}.
\end{aligned}$$

These two properties are sufficient to study the number of rounds required by the new Markov chain $\{R_i\}_{i\in\mathbb{N}}$ to reach the target value m. Indeed, by defining the random variable $Z_i = \frac{f(R_i)}{\sqrt{n}}$, and considering the following "potential" function, $Y_i = exp\left(\frac{m}{\sqrt{n}} - Z_i\right)$, we can compute its expectation at the next round as follows. Let us fix any state $x \in \Omega$ such that $h\sqrt{n} \leq f(x) < m$, and define $z = \frac{f(x)}{\sqrt{n}}$, $y = exp\left(\frac{m}{\sqrt{n}} - z\right)$. We have

$$\begin{aligned}
\mathbb{E}\left[Y_{i+1} | R_i = x\right] &\leq \Pr\left[f(R_{i+1}) < (1+\epsilon)f(x)\right] e^{m/\sqrt{n}}\\
&\quad + \Pr\left[f(R_{i+1}) \geq (1+\epsilon)f(x)\right] e^{m/\sqrt{n}-(1+\epsilon)z}\\
\text{(from Hypothesis (2))} &\leq e^{-c_2 z^2}\cdot e^{m/\sqrt{n}} + 1\cdot e^{m/\sqrt{n}-(1+\epsilon)z}\\
&= e^{m/\sqrt{n}-c_2 z^2} + e^{m/\sqrt{n}-z-\epsilon z}\\
&= e^{m/\sqrt{n}-z}(e^{z-c_2 z^2} + e^{-\epsilon z})\\
&\leq e^{m/\sqrt{n}-z}(e^{-2} + e^{-2}) \qquad (2.7)\\
&< \frac{e^{m/\sqrt{n}-z}}{e}\\
&= \frac{y}{e},
\end{aligned}$$

where in (2.7) we used that z is always at least h and thanks to Hypothesis (1) we can choose a sufficiently large h.

By applying the Markov inequality and iterating the above bound, we get

$$\Pr[Y_i > 1] \leq \frac{\mathbb{E}[Y_i]}{1} \leq \frac{\mathbb{E}[Y_{i-1}]}{e} \leq \cdots \leq \frac{\mathbb{E}[Y_0]}{e^{\tau_R}} \leq \frac{e^{m/\sqrt{n}}}{e^i}.$$

We observe that if $Y_i \leq 1$ then $R_i \geq m$, thus by setting $i = m/\sqrt{n} + \log n = (c_3+1)\log n$, we get:

$$\Pr\left[R_{(c_3+1)\log n} < m\right] = \Pr\left[Y_{(c_3+1)\log n} > 1\right] < \frac{1}{n}. \tag{2.8}$$

Our next goal is to give an upper bound on the hitting time $\tau_{(c_3+1)\log n}$. Note that the event "$\tau_{(c_3+1)\log n} > c_4 \log n$" holds if and only if the number of rounds such that $f(X_t) \geq h\sqrt{n}$ (before round $c_4 \log n$) is less than $(c_3+1)\log n$. Thanks to Hypothesis (1), at each round t there is at least probability $1 - c_1$ that $f(X_t) \geq h\sqrt{n}$. This implies that, for any positive constant c_4, the probability $\Pr\left[\tau_{(c_3+1)\log n} > c_4 \log n\right]$ is bounded by the probability that, within $c_4 \log n$ independent Bernoulli trials, we get less then $(c_3+1)\log n$ successes, where the success probability is at least $1 - c_1$. We can thus choose a sufficiently large $c_4 = c_4(c_1, c_3)$ and apply the multiplicative form of the Chernoff bound (Lemma B.1), obtaining

$$\Pr\left[\tau_{(c_3+1)\log n} > c_4 \log n\right] < \frac{1}{n}. \tag{2.9}$$

We are now ready to prove the Lemma using (2.8) and (2.9), indeed

$$\begin{aligned}
\Pr[X_{c_4 \log n} \geq m] &> \Pr\left[R_{(c_3+1)\log n} \geq m \wedge \tau_{(c_3+1)\log n} \leq c_4 \log n\right] \\
&= 1 - \Pr\left[R_{(c_3+1)\log n} < m \vee \tau_{(c_3+1)\log n} > c_4 \log n\right] \\
&\geq 1 - \Pr\left[R_{(c_3+1)\log n} < m\right] + \Pr\left[\tau_{(c_3+1)\log n} > c_4 \log n\right] \\
&> 1 - \frac{2}{n}.
\end{aligned}$$

Hence, choosing a suitable large c_4, we have shown that in $c_4 \log n$ rounds the process reaches the target value m, w.h.p. □

Our goal is to apply the above lemma to the UNDECIDED-STATE process (which defines a finite-state Markov chain) starting with bias of size $o(\sqrt{n \log n})$ where we set $f(M_t) = s_t$, $c_3 = \gamma > 0$ for some constant $\gamma > 0$, and $m = \gamma\sqrt{n}\log n$: this would imply the upper bound $\mathcal{O}(\log n)$ on the number of rounds needed to reach a configuration having bias $\Omega(\sqrt{n \log n})$, w.h.p., breaking the symmetry because Theorem 2.3.1 then holds. To this aim, with the next two lemmas we show that the UNDECIDED-STATE process satisfies the hypotheses of Lemma 2.4.6 in this setting, w.h.p.

Lemma 2.4.7. *Let ϵ be any constant such that $(1-2p)^2/2 \leq \epsilon < (1-2p)(1-2p/3)$. Let M_{t-1} be any configuration in which $s \leq \check{s}_\epsilon$, and $q \leq \frac{n}{2}$. Then, in the next round, it holds that $\hat{q}_{\epsilon/12} \leq q_t \leq \frac{n}{2}$ with probability $1 - \exp\left(-\epsilon^2 p^2 n/(2^4 3^4)\right)$.*

Proof of Lemma 2.4.7. Lemma 2.4.2 implies that $q_t \geq \hat{q}_{\epsilon/12}$ with probability $1 - \exp\left(-\epsilon^2 n / \left(2^3 3^2 (1-p)^2\right)\right)$. At the same time, by Eq. (2.6) it holds that

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \leq \frac{pn}{3} + \frac{1-p}{2n}\left[2q^2 + (n-q)^2\right] = f(q).$$

For $q \leq \frac{n}{2}$, the maximum of f is obtained either at $q_1 = 0$ or at $q_2 = \frac{n}{2}$. Then,

$$\begin{aligned} f(q_1) &= \left(\frac{p}{3} + \frac{1-p}{2}\right) n = \frac{(3-p)n}{6}; \\ f(q_2) &= pn + \frac{1-p}{2n}\left[\frac{n^2}{2} + n^2 - n^2 + \frac{n^2}{4}\right] = \left(\frac{p}{3} + \frac{3(1-p)}{8}\right) n \\ &= \frac{(9-p)n}{24}. \end{aligned}$$

Therefore, we have $f(q) \leq (3-p)n/6$ for $q \leq \frac{n}{2}$, since $(3-p)/6 > (9-p)/24$ for $p < 1$. By the additive form of Chernoff bound (Lemma B.2),

$$\begin{aligned} \Pr\left[q_t \geq \frac{n}{2} \,\middle|\, M_{t-1}\right] &= \Pr\left[q_t \geq \frac{(3-p)n}{6} + \frac{pn}{6} \,\middle|\, M_{t-1}\right] \leq \Pr\left[q_t \geq \mathbb{E}\left[q_t \mid M_{t-1}\right] + \frac{pn}{6} \,\middle|\, M_{t-1}\right] \\ &\leq \exp\left(-p^2 n / 2 \cdot 3^2\right). \end{aligned}$$

Hence, the joint probability that $q_t \leq n/2$ and $q_t \geq \hat{q}_{\epsilon/12}$ is at least $1 - \exp\left(-\epsilon^2 p^2 n/(2^4 3^4)\right)$, since $\epsilon < 1$ and $p < 1$. □

Lemma 2.4.8. *Let ϵ be any constant such that $(1-2p)^2/2 \leq \epsilon < (1-2p)(1-2p/3)$. Let M_{t-1} be any configuration such that $q \in \left[\hat{q}_{\epsilon/12}, n/2\right]$. Then, it holds that*

(1) for any constant $h > 0$ there exists a constant $c_1 > 0$ (depending only on p) such that

$$\Pr\left[|s_t| < h\sqrt{n}) \mid M_{t-1}\right] < c_1;$$

(2) there exist two positive constants δ and c_2 (depending only on p) such that

$$\Pr\left[|s_t| \geq (1+\delta)s \mid M_{t-1}\right] \geq 1 - e^{-c_2 \frac{s^2}{n}}.$$

Proof. As for the first item, let $\mathbf{x} = (a, b, q)$ and $\mathbf{x}' = (a', b', q')$ be two states such that $|s| = |a - b| < h\sqrt{n}\log n$, $|s'| = |a' - b'| = 0$, $q = q'$. A simple domination argument implies that

$$\Pr\left[|s_t| < h\sqrt{n} \mid M_{t-1} = \mathbf{x}\right] \leq \Pr\left[|s_t| < h\sqrt{n} \mid M_{t-1} = \mathbf{x}'\right].$$

Thus, we can bound just the second probability, where the initial bias is zero, which implies that $a = b$.

Define a_t^q, b_t^q, q_t^q the random variables counting the nodes that were undecided in the configuration $\mathbf{x}'$ and that, in the next round, get the opinion *Alpha*, *Beta*, and undecided, respectively. Similarly, a_t^a (b_t^b) counts the nodes that support opinion *Alpha* (*Beta*) in the configuration $\mathbf{x}'$ and that, in the next round, still support the same opinion. Trivially, $a_t = a_t^q + a_t^a$ and $b_t = b_t^q + b_t^b$. Moreover, observe that, among these random variables, only a_t^q and b_t^q are mutually dependent.

Thus, conditional on the event $\{M_{t-1} = \mathbf{x}'\}$, if $\alpha = \mathbb{E}\left[a_t^a \mid M_{t-1} = \mathbf{x}'\right] = \mathbb{E}\left[b_t^b \,\middle|\, M_{t-1} = \mathbf{x}'\right]$, it holds that

$$\Pr\left[|s_t| \geq h\sqrt{n}\right] \geq \Pr\left[a_t \geq b_t + h\sqrt{n}\right] \geq \Pr\left[a_t^q \geq b_t^q + h\sqrt{n}\right] \Pr\left[a_t^a \geq \alpha\right] \Pr\left[b_t^b \leq \alpha\right].$$

The random variables $a_t^a - \alpha$ and $b_t^b - \alpha$ follow binomial distributions with expectation 0 (recall that $a = b$), and finite second and third moment. Thus, the Berry-Esseen inequality (Lemma B.6 in Appendix B) allows us to approximate up to an arbitrary-small constant $\epsilon_1 > 0$ (as long as n is large enough) both the random variables with a normal distribution that has expectation 0. Thus,

$$\Pr\left[a_t^a \geq \alpha\right] = \Pr\left[b_t^b \leq \alpha\right] \geq \left(\frac{1}{2} - \epsilon_1\right).$$

As for the random variable $a_t^q - b_t^q$, notice that conditional on the event $\{q - q_t^q = k\}$, it is the sum of k Rademacher random variables. The hypothesis $q \leq \frac{n}{2}$ allows us to use the Chernoff bound on q_t^q and show that $q_t^q \leq \frac{3}{4}q$ w.h.p. Thus, since $q \geq \hat{q}_{\epsilon/12}$, it holds that $q - q_t^q = \Theta(n)$ w.h.p. It follows that the conditional variance of $a_t^q - b_t^q$ given $q - q_t^q$ yields $\Theta(n)$ w.h.p., and $a_t^q - b_t^q$ conditional on the event $E = \{q - q_t^q = \Theta(n)\}$ can be approximated by a normal distribution up to an arbitrary-small constant $\epsilon_2 > 0$. Then, we have that

$$\Pr\left[a_t^q \geq b_t^q + h\sqrt{n}\right] \geq \Pr\left[a_t^q \geq b_t^q + h\sqrt{n} \mid E\right] \Pr\left[E\right] \geq \epsilon_2.$$

Setting $c_1 = \epsilon_1 \cdot \epsilon_2$, we get property (1). We can choose ϵ_1, ϵ_2 to be equal to p, so that c_1 depends only on p. As for property (2), by Eq. (2.5) and the hypothesis on q we have that

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq s\left(1 + \epsilon/12\right).$$

We can get the property applying the Hoeffding bound (Lemma B.3), getting that

$$\Pr\left[s_t \leq s\left(1 + \frac{\epsilon}{24}\right) \,\middle|\, M_{t-1}\right] = \Pr\left[s_t \leq s\left(1 + \frac{\epsilon}{12}\right) - \frac{\epsilon s}{24} \,\middle|\, M_{t-1}\right] \leq \exp\left(-\frac{\epsilon^2 s^2}{2^7 3^2 n}\right),$$

which is the thesis. □

The reader may notice that Lemma 2.4.7 requires the number of undecided nodes to be inside the interval $\left[\hat{q}_{\epsilon/12}, n/2\right]$. We will later take care of this issue with Lemma 2.4.11, showing that whenever this number of undecided agents is not within the above interval, in at most $\mathcal{O}(\log n)$ rounds it will. Furthermore, Lemma 2.4.7 guarantees that the condition on the undecided nodes holds "only" w.h.p., while Lemma 2.4.6 requires this condition to hold with probability 1. We show this issue can be solved using a coupling argument similar to that in [Clementi et al., 2018]. The key point is that, starting from any configuration M_t with $q_t \in \left[\hat{q}_{\epsilon/12}, n/2\right]$, the probability that the process goes in one of those "bad" configurations with q outside the above interval is negligible. Intuitively speaking, the configurations actually visited by the process before breaking symmetry do satisfy the hypothesis of Lemma 2.4.6. In order to make this argument rigorous, we define a *pruned* process, by removing all the unwanted transitions.

Let $\bar{s} \in \{0, 1, \ldots, n\}$, and $\mathbf{z}_{\bar{s}} = (\bar{a}, \bar{b}, \bar{q})$ the configuration with bias $\bar{s} = \bar{a} - \bar{b}$, and undecided nodes $\bar{q} = n/2$. Let $p_{\mathbf{x},\mathbf{y}}$ be the probability of a transition from the configuration $\mathbf{x} = (a_\mathbf{x}, b_\mathbf{x}, q_\mathbf{x})$ to the configuration $\mathbf{y} = (a_\mathbf{y}, b_\mathbf{y}, q_\mathbf{y})$ in the UNDECIDED-STATE process. The PRUNED process

behaves exactly as the original process but every transition from a configuration $\mathbf{x}$ such that $q_{\mathbf{x}} \in \left[\hat{q}_{\epsilon/12}, n/2\right]$ and $s_{\mathbf{x}} = a_{\mathbf{x}} - b_{\mathbf{x}} = \mathcal{O}(\sqrt{n \log n})$ to a configuration $\mathbf{y}$ such that $q_{\mathbf{y}} < \hat{q}_{\epsilon/12}$ or $q_{\mathbf{y}} > n/2$ has probability $p_{\mathbf{x},\mathbf{y}} = 0$. Moreover, for any $\bar{s} \in [n]$, starting from the configuration $\mathbf{x}$, the probability of reaching the configuration $\mathbf{z}_{\bar{s}}$ is

$$p'_{\mathbf{x},\mathbf{z}_{\bar{s}}} = p_{\mathbf{x},\mathbf{z}_{\bar{s}}} + \sum_{\substack{\mathbf{y}:\ s_{\mathbf{y}} = \bar{s} \text{ and} \\ q_{\mathbf{y}} \notin \left[\hat{q}_{\epsilon/12}, \frac{n}{2}\right]}} p_{\mathbf{x},\mathbf{y}}.$$

All the other transition probabilities remain the same. Observe that the PRUNED process is defined in such a way that it has exactly the same marginal probability of the original process with respect to (in short, w.r.t.) the random variable s_t; thus, Lemma 2.4.8 holds for the PRUNED process as well and we can apply Lemma 2.4.6. Then, the PRUNED process reaches a configuration having bias $\Omega(\sqrt{n \log n})$ within $\mathcal{O}(\log n)$ rounds, w.h.p., as shown in the following lemma.

Lemma 2.4.9. *Let ϵ be any constant such that $(1-2p)^2/2 \leq \epsilon < (1-2p)(1-2p/3)$. Furthermore, let $\gamma > 0$ be any constant. Starting from any configuration M_t such that $q_t \in \left[\hat{q}_{\epsilon/12}, n/2\right]$ and $s_t \leq \gamma\sqrt{n \log n}$, the* PRUNED *process reaches a configuration having bias $\Omega(\sqrt{n \log n})$ within $\mathcal{O}_{p,\gamma}(\log n)$ rounds with probability $1 - 2/n$.*

Proof of Lemma 2.4.9. Let $\gamma > 0$ be a constant and $m = \gamma\sqrt{n} \log n$ be the target value of the bias in Lemma 2.4.6. Since $q_t \in \left[\hat{q}_{\epsilon/12}, n/2\right]$ and $s_t = \mathcal{O}(\sqrt{n \log n})$, the PRUNED process always satisfies Lemma 2.4.8, and thus we can apply Lemma 2.4.6 (setting the function $f(M_t) = s_t$), which gives us that the PRUNED process process reaches a configuration $\mathbf{y}$ having bias $s \geq m = \Omega(\sqrt{n \log n})$ within $\mathcal{O}_{p,\gamma}(\log n)$ rounds with probability $1 - 2/n$. □

We now want to go back to the original process. The definition of the PRUNED process suggests a natural coupling between it and the original one. If the two process are in different states, then they act independently, while, if they are in the same state M_t, they move together unless the UNDECIDED-STATE process goes in a configuration $\mathbf{y}$ such that $q \notin \left[\hat{q}_{\epsilon/12}, n/2\right]$. In that case, the PRUNED process goes in $\mathbf{z}_{s_{\mathbf{y}}}$, where $s_{\mathbf{y}}$ is the bias of configuration $\mathbf{y}$. In the proof of the next lemma, we show that the time the PRUNED process takes to reach bias $\Omega(\sqrt{n \log n})$ stochastically dominates the one of the original process, giving the result.

Lemma 2.4.10. *Let γ, ϵ be any two positive constants, with $(1-2p)^2/2 \leq \epsilon < (1-2p)(1-2p/3)$. Starting from any configuration M_0 such that $q_0 \in \left[\hat{q}_{\epsilon/12}, n/2\right]$ and $s_0 \leq \gamma\sqrt{n \log n}$, the* UNDECIDED-STATE *process reaches a configuration having bias $s \geq \gamma\sqrt{n \log n}$ within $\mathcal{O}_{p,\gamma}(\log n)$ rounds with probability at least $1 - 3/n$.*

Proof of Lemma 2.4.10. Let $\{M_t\}$ and $\{\mathbf{Y}_t\}$ be the original process and the pruned one, respectively. Denote the set of possible initial configuration according to the hypothesis by H. Note that if $M_t = \mathbf{Y}_t$, then

$$\mathbf{Y}_{t+1} = \begin{cases} M_{t+1} & \textbf{if } M_{t+1} \in H \\ \mathbf{z}_{s_{t+1}} & \text{otherwise} \end{cases}.$$

Let $\tau = \inf\{t : \mathbb{N} : |s_t| \geq \sqrt{n \log n}\}$, and let $\tau* = \inf\{t \in \mathbb{N} : |s_{\mathbf{Y}_t}| \geq \sqrt{n \log n}\}$. For any configuration $\mathbf{x} \in H$, define $\rho^t_{\mathbf{x}}$ the event that the two processes $\{M_t\}$ and $\{\mathbf{Y}_t\}$ have separated

at round $t+1$, i.e., $\rho_{\mathbf{x}}^t = \{M_t = \mathbf{Y}_t = \mathbf{x}\} \cap \{M_{t+1} \neq \mathbf{Y}_{t+1}\}$. Observe that, if the two couple processes in the same configuration $\mathbf{x} \in H$ and $\tau > c \log n$, then either $\tau* > c \log n$ or there exists a round $t \leq c \log n$ such that for some $\mathbf{x} \in H$ the event $\rho_{\mathbf{x}}^t$ has occurred. Hence, if $\mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}[\cdot]$ is the joint probability for the couple $(M_t, \mathbf{Y_t})$ which both start at $\mathbf{x}_0$, we have

$$\begin{aligned}
&\mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}[\tau > c \log n] \\
&\leq \mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}\left[\{\tau* > c \log n\} \cup \{\exists t \leq c \log n, \exists \mathbf{x} \in H : \rho_{\mathbf{x}}^t\}\right] \\
&\leq \mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}[\tau* > c \log n] + \mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}\left[\exists t \leq c \log n, \exists \mathbf{x} \in H : \rho_{\mathbf{x}}^t\right].
\end{aligned}$$

We choose a suitable constant c which depends on p and γ, in order to apply Lemmas 2.4.6 and 2.4.7. As for the first item, since Lemma 2.4.6 holds for the PRUNED process, we have that it is upper bounded by $2/n$. As for the second term, we get that

$$\begin{aligned}
&\mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}\left[\exists t \leq c \log n, \exists \mathbf{x} \in H : \rho_{\mathbf{x}}^t\right] \leq \sum_{t=1}^{c \log n} \mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}\left[\exists \mathbf{x} \in H : \rho_{\mathbf{x}}^t\right] \\
&= \sum_{t=1}^{c \log n} \sum_{\mathbf{x} \in H} \mathbb{P}'_{\mathbf{x}_0,\mathbf{x}_0}\left[\rho_{\mathbf{x}}^t\right] \overset{(a)}{\leq} \sum_{t=1}^{c \log n} \frac{n^2}{e^{-\frac{\epsilon^2 p^2 n}{2^4 3^2}}} \leq \frac{1}{n},
\end{aligned}$$

where in (a) the second inequality we used Lemma 2.4.7, and the fact that $|H|$ is at most all the combinations of parameters q and s. □

Now, we take care of those cases in which the starting configuration M_{t-1} is such that $q \notin \left[\hat{q}_{\epsilon/12}, n/2\right]$. If $q < \hat{q}_{\epsilon/12}$ then, for Lemma 2.4.7, $q_t \in \left[\hat{q}_{\epsilon/12}, n/2\right]$ w.h.p. Next lemma takes care of the case $q > n/2$.

Lemma 2.4.11. *Let M_{t-1} be any starting configuration such that $q > \frac{n}{2}$. Then, at the next round, it holds that $q_t \leq q\,(1 - p/6)$ with probability $1 - \exp\left(-p^2 n / 2^3 \cdot 3^2\right)$.*

Proof of Lemma 2.4.11. From Eq. (2.6), we have that

$$\begin{aligned}
\mathbb{E}\left[q_t \mid M_{t-1}\right] - \frac{q(3-2p)}{3} &\leq \frac{pn}{3} + \frac{1-p}{2n}\left[2q^2 + (n-q)^2\right] - \frac{q(3-p)}{3} \\
&= 3q^2\left(\frac{1-p}{2n}\right) - \frac{q(6-4p)}{3} + n\left(\frac{3-p}{6}\right) = f(q).
\end{aligned}$$

Now, $f(q)$ takes its maximum either in $q_1 = n/2$, or in $q_2 = n$. Then,

$$\begin{aligned}
f(q_1) &= \frac{3n}{4}\left(\frac{1-p}{2}\right) - \frac{n(6-4p)}{6} + n\left(\frac{3-p}{6}\right) = n \cdot \frac{9 - 9p - 24 + 16p + 12 - 4p}{24} \\
&= n \cdot \frac{-1+p}{8}; \\
f(q_2) &= 3n\left(\frac{1-p}{2}\right) - \frac{n(6-4p)}{3} + n\left(\frac{3-p}{6}\right) = n \cdot \frac{9 - 9p - 12 + 8p + 3 - p}{6} \\
&= -\frac{pn}{3}.
\end{aligned}$$

Since $p < 1/2$, we have that $f(q) < 0$ for $n/2 < q \leq n$. Thus,

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \leq q\left(1 - \frac{p}{3}\right),$$

and we can use the Chernoff bound (Lemma B.2).

$$\begin{aligned}
\Pr\left[q_t \geq q\left(1 - \frac{p}{6}\right) \,\middle|\, M_{t-1}\right] &= \Pr\left[q_t \geq q\left(1 - \frac{p}{3}\right) + \frac{pq}{6} \,\middle|\, M_{t-1}\right] \\
&\leq \Pr\left[q_t \geq \mathbb{E}\left[q_t \mid M_{t-1}\right] + \frac{pn}{12} \,\middle|\, M_{t-1}\right] \\
&\leq \exp\left(-\frac{p^2 n}{2^3 \cdot 3^2}\right).
\end{aligned}$$

□

Finally, we are ready to prove Theorem 2.3.2.

Proof of Theorem 2.3.2: Wrap-Up. Let $\gamma = p/3$, and let $m = \gamma\sqrt{n \log n}$ be the target value of the bias in Lemma 2.4.6. Let M_{t-1} be any initial configuration having bias $|s| < m$. Let ϵ be any constant such that $(1-2p)^2/2 \leq \epsilon < (1-2p)(1-2p/3)$. We have two cases.

(i) If the number of undecided nodes is such that $q_t \in \left[\hat{q}_{\epsilon/12}, n/2\right]$, then Lemma 2.4.10 implies that the UNDECIDED-STATE process reaches a configuration having bias $s \geq \gamma\sqrt{n \log n}$ in $\mathcal{O}_p(\log n)$ rounds with probability $1 - 3/n$;

(ii) else, if the starting configuration is such that $q_t \notin \left[\hat{q}_{\epsilon/12}, n/2\right]$, then, for Lemmas 2.4.7 and 2.4.11, the UNDECIDED-STATE process reaches within $\mathcal{O}_p(\log n)$ rounds a configuration having the number of undecided nodes $q \in \left[\hat{q}_{\epsilon/12}, n/2\right]$, with probability $1 - \exp\left(-p^2 n/2^4 \cdot 3^2\right)$ for Lemmas A.1 and A.2. Then, either the bias is $s \geq \gamma\sqrt{n \log n}$, or we are in case (i). As Lemmas A.1 and A.2 in the preliminaries imply, the probability that case (ii) and then case (i) take place is a high probability (in this case, with probability at least $1 - 4/n$).

Then, Theorem 2.3.1 gives the desired result. □

2.4.5 UNDECIDED-STATE dynamics: victory of noise

This section provide all technical lemmas to prove Theorem 2.3.3. We assume the starting configuration M_0 to have bias $s = a - b \geq \gamma\sqrt{n \log n}$ for some constant $\gamma > 0$; the case in which $b > a$ is analogous. Let $1/2 < p < 1$ be the probability of noise, with p being a constant.

We first show that if the bias is positive, it can become negative but its absolute value will be bounded by $\mathcal{O}\left(n \log n\right)$. The same holds if it is negative, by symmetry.

Lemma 2.4.12. *Let M_{t-1} be any configuration such that $s \geq 0$. Then, for any $\gamma > 0$, $s_t \geq -\gamma\sqrt{n \log n}$ with probability $1 - n^{-\gamma^2/2}$.*

Proof of Lemma 2.4.12. From Eq. (2.5), we have that $\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq 0$. If, Since $p \leq 1$, we have $\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq s\left(1 - p/3\right) \geq 0$. Let $\lambda = \gamma\sqrt{\log n / n}$. Then, by the Hoeffding bound (Lemma B.3), we have that

$$\Pr\left[s_t \geq -\gamma\sqrt{n \log n} \,\middle|\, M_{t-1}\right] = \Pr\left[s_t \geq -\lambda n \mid M_{t-1}\right] \leq \exp\left(-\frac{\gamma^2 \log n}{2}\right) = n^{-\frac{\gamma^2}{2}}.$$

□

We distinguish two cases: $p > 3/4$ and $1/2 < p \leq 3/4$.

2.4.5.1 First case: $p > \frac{3}{4}$ large noise.

Let $\varepsilon > 0$ be a constant such that $3/4 + \varepsilon < p \leq 1$. We first show a bound on the decrease of the bias.

Lemma 2.4.13. *Let M_{t-1} be a configuration such that $s \geq \gamma\sqrt{n \log n}$ for any constant $\gamma > 0$. Then, in the next round, it holds that $s_t \leq s\,(1-2\varepsilon)$ with probability $1 - n^{-2\varepsilon^2\gamma^2/9}$.*

Proof of Lemma 2.4.13. Equation (2.5) implies that

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \leq s_t\left[1 - \frac{1}{4} - \frac{\varepsilon}{3} + 1 - \frac{3}{4} - \varepsilon\right] = s_t\left(1 - \frac{4\varepsilon}{3}\right),$$

since $q/n \leq 1$. Let $\lambda = 2\varepsilon(s/3n)$. From the Hoeffding bound (Lemma B.3), we have that

$$\Pr\left[s_t \leq s\left(1 - \frac{2\varepsilon}{3}\right) \,\middle|\, M_{t-1}\right] = \Pr\left[s_t \leq s\left(1 - \frac{4\varepsilon}{3}\right) + \lambda n \,\middle|\, M_{t-1}\right] \leq \exp\left(-\frac{2\varepsilon^2 s^2}{9n}\right) \leq n^{-\frac{2\varepsilon^2\gamma^2}{9}}.$$

□

2.4.5.2 Second case: $1/2 < p \leq 3/4$ small noise.

Let $\varepsilon > 0$ be a constant such that $1/2 + 3\varepsilon = p \leq 3/4$,[2] which implies that $\varepsilon \leq 1/12$. We remark that, in contrast to the previous case, here the parameter p for the noise is *exactly* $1/2 + 3\varepsilon$ for the sake of the analysis.

The following lemma states that, if $s \geq \frac{2}{3}n$, the bias decreases exponentially at the next round, w.h.p. On the other hand, if the bias is at most $\frac{2}{3}n$, it cannot grow over $\frac{2}{3}n$, w.h.p.

Lemma 2.4.14. *Let M_{t-1} be any configuration. The following holds:*
(1) if $s \geq \frac{2}{3}n$, then $s_t \leq s(1-\varepsilon)$ with probability $1 - \exp\left(-2\varepsilon^2 n/9\right)$;
(2) if $s \leq \frac{2}{3}n$, then $s_t \leq \frac{2}{3}n$ with probability $1 - \exp\left(-16\varepsilon^2 n/9\right)$.

Proof of Lemma 2.4.14. We first address claim (1). Consider the first statement. If $s \geq \frac{2}{3}n$, then $q \leq \frac{1}{3}n$. Thus, from Eq. (2.5), we get

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \leq s\left[1 - \frac{1}{6} - \varepsilon + \frac{1}{3} - \frac{1}{6} - \varepsilon\right] = s\,(1-2\varepsilon).$$

Let $\lambda = \varepsilon(s/n)$. By the Hoeffding bound (Lemma B.3), we conclude that

$$\Pr\left[s_t \leq s(1-\varepsilon) \mid M_{t-1}\right] = \Pr\left[s_t \leq s(1-2\varepsilon) + \lambda n \mid M_{t-1}\right] \leq \exp\left(-\frac{\varepsilon^2 s^2}{2n}\right) \leq \exp\left(-\frac{2\varepsilon^2 n}{9}\right),$$

since $s \geq 2n/3$.

2. The factor 3 in 3ε is useful for calculations.

Now we prove claim (2). Since $q \leq n - s$, from Eq. (2.5) we get that

$$\begin{aligned}\mathbb{E}\left[s_t \mid M_{t-1}\right] &\leq s \cdot \left[1 - \frac{1}{6} - \varepsilon + \frac{n-s}{n} \cdot \left(1 - \frac{1}{2} - 3\varepsilon\right)\right] \\ &\leq \frac{2n}{3} \cdot \left[1 - \frac{1}{6} - \varepsilon + 1 - \frac{1}{2} - 3\varepsilon - \frac{2}{3} + \frac{1}{3} + 2\right] \\ &= \frac{2n}{3} \cdot \left[1 - 4\varepsilon\right].\end{aligned}$$

Let $\lambda = 4\varepsilon(2/3)$. For the Hoeffding bound (Lemma B.3), it holds that

$$\Pr\left[s_t \leq \frac{2n}{3} \,\middle|\, M_{t-1}\right] = \Pr\left[s_t \leq \frac{2n}{3}\left(1 - 4\varepsilon\right) + \lambda n \,\middle|\, M_{t-1}\right] \leq \exp\left(-\frac{16\varepsilon^2 n}{9}\right).$$

□

Thus, we just have to take care of cases in which the bias is no more than $\frac{2}{3}n$. The key-point to show the decrease of the bias, as long as it is $\Omega\left(\sqrt{n \log n}\right)$, it is the condition $q \leq \breve{q}_{\varepsilon/2} = (1 + 3\varepsilon)(n/3)/(1 - 6\varepsilon)$, as shown in the next lemma.

Lemma 2.4.15. *Let M_{t-1} be a configuration such that $s \geq \gamma\sqrt{n \log n}$ for some constant $\gamma > 0$. If $q \leq \breve{q}_{\varepsilon/2}$, then in the next round it holds that $s_t \leq s\left(1 - \frac{\varepsilon}{4}\right)$ with probability $1 - n^{-\varepsilon^2\gamma^2/2^6}$.*

Proof of Lemma 2.4.15. From Eq. (2.5) it follows that

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \leq s \cdot \left[1 - \frac{1}{6} - \varepsilon + \frac{1}{6} + \frac{\varepsilon}{2}\right] = s\left(1 - \frac{\varepsilon}{2}\right).$$

Let $\lambda = (\varepsilon/4)(s/n)$. By the Hoeffding bound (Lemma B.3), we get

$$\Pr\left[s_t \leq s\left(1 - \frac{\varepsilon}{4}\right) \,\middle|\, M_{t-1}\right] = \Pr\left[s_t \leq s\left(1 - \frac{\varepsilon}{2}\right) + \lambda n \,\middle|\, M_{t-1}\right] \leq \exp\left(-\frac{\varepsilon^2 s^2}{2^6 n}\right) \leq n^{-\frac{\varepsilon^2\gamma^2}{2^6}}.$$

□

We now analyze the dynamics by partitioning the interval $\left(0, \frac{2}{3}n\right]$ and seeing what happens to the bias in each element of the partition. Let $\beta = \frac{2\sqrt{2\varepsilon}}{\sqrt{(1+6\varepsilon)(1-6\varepsilon)}}$ and define $S_{-1} \coloneqq (0, \beta n]$, S_i the sequence of intervals

$$S_i \coloneqq \left(\left(\frac{3}{2}\right)^i \beta n, \left(\frac{3}{2}\right)^{i+1} \beta n\right]$$

for $i = 0, 1, \ldots, k-2$ where $k = \left\lceil \log_{\frac{2}{3}}(\beta) - 1 \right\rceil$, and $S_{k-1} \coloneqq \left(\left(\frac{3}{2}\right)^{k-1} \beta n, \frac{2}{3}n\right]$. Furthermore, just for completeness, we define $S_k \coloneqq \left(\frac{2}{3}n, n\right]$. In the next lemmas, we show that as long as $s \in S_i$ for $i = -1, \ldots, k-1$, $s = \Omega(\sqrt{n \log n})$, and $\bar{q}_{i+1} \leq q \leq \breve{q}_{\varepsilon/2}$ for some decreasing sequence $\bar{q}_{-1}, \ldots, \bar{q}_k$ accurately chosen, then, at the next round, the bias decreases exponentially w.h.p. and the number of undecided nodes moves to the interval $\left[\bar{q}_i, \breve{q}_{\varepsilon/2}\right]$ w.h.p. Note that since ε is a constant, so is k. The following lemma determines the sequence $\bar{q}_i$.

Lemma 2.4.16. *Let M_{t-1} be any configuration.*

1. *If* $-1 \leq i \leq k-1$ *and* $s \leq \left(\frac{3}{2}\right)^{i+1} \beta n$, *it holds that* $q_t \geq \frac{n}{3} - \frac{2n\varepsilon}{1+6\varepsilon}\left(\frac{3}{2}\right)^{2i+3}$ *with probability* $1 - \exp\left(-2\varepsilon^2(3/2)^{4i+4}n/(1+6\varepsilon)^2\right)$;
2. *If* $s \leq \frac{2}{3}n$, *it holds that* $q_t \geq \frac{2}{9}n + \frac{\varepsilon}{3}n$ *with probability* $1 - \exp\left(-2\varepsilon^2 n/9\right)$;
3. *Without any condition on* s, *it holds that* $q_t \geq \frac{n}{12} + \varepsilon n$ *with probability* $1 - \exp\left(-\varepsilon^2 n/2\right)$.

Proof of Lemma 2.4.16. We start proving claim (1). From Equation (2.6), we have that

$$\begin{aligned}\mathbb{E}\left[q_t \mid M_{t-1}\right] &\geq \frac{n}{3} - \frac{1-6\varepsilon}{n}\left(\frac{s}{2}\right)^2 \geq \frac{n}{3} - \frac{1-6\varepsilon}{4}\left(\frac{3}{2}\right)^{2i+2}\beta^2 n \\ &= n\left[\frac{1}{3} - \frac{2\varepsilon}{1+6\varepsilon}\left(\frac{3}{2}\right)^{2i+2}\right].\end{aligned}$$

Thus, using the additive form of Chernoff bound (Lemma B.2) with $\lambda = \frac{\varepsilon}{1+6\varepsilon}\left(\frac{3}{2}\right)^{2i+2} n$, we have that $q_t \geq n\left(\frac{1}{3} - \frac{2\varepsilon}{1+6\varepsilon}\left(\frac{3}{2}\right)^{2i+3}\right)$ with probability $1 - \exp\left(-2\varepsilon^2(3/2)^{4i+4}n/(1+6\varepsilon)^2\right)$.

As for claim (2), we have that

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \geq \frac{n}{3} - \frac{1-6\varepsilon}{n}\left(\frac{s}{2}\right)^2 \geq \frac{n}{3} - n\frac{1-6\varepsilon}{9} = \frac{2+6\varepsilon}{9}n$$

and we conclude by using the additive Chernoff bound with $\lambda = \varepsilon n/3$, getting that $q_t \geq \frac{2}{9}n + \frac{\varepsilon}{3}n$ with probability $1 - \exp\left(-2\varepsilon^2 n/9\right)$.

To prove claim (3), we use that $s \leq n$ and observe that

$$\mathbb{E}\left[q_t \mid M_{t-1}\right] \geq \frac{n}{3} - \frac{1-6\varepsilon}{n}\left(\frac{s}{2}\right)^2 \geq \frac{n}{3} - n\frac{1-6\varepsilon}{4} = \frac{1+18\varepsilon}{12}n.$$

We conclude with the additive Chernoff bound (Lemma B.2) with $\lambda = \varepsilon n/2$, getting $q_t \geq \frac{n}{12} + \varepsilon n$ with probability $1 - \exp\left(-\varepsilon^2 n/2\right)$. □

Define $\bar{q}_i \coloneqq n\left(\frac{1}{3} - \frac{2\varepsilon}{1+6\varepsilon}\left(\frac{3}{2}\right)^{2i+3}\right)$ for $i = -1, ..., k-2$, $\bar{q}_{k-1} \coloneqq \frac{2}{9}n + \frac{\varepsilon}{3}n$, and $\bar{q}_k \coloneqq \frac{n}{12} + \varepsilon n$, and notice that they form a decreasing sequence. With the next lemmas, we take care of controlling the behaviour of the number of undecided nodes when $s > \inf(S_i)$ for $-1 \leq i \leq k-1$.

Lemma 2.4.17. *Let* $-1 \leq i \leq k-1$ *and let* M_{t-1} *be a configuration such that* $\bar{q}_{i+1} \leq q \leq \check{q}_{\varepsilon/2}$ *and* $s > \inf(S_i)$. *Then, at the next round,* $\bar{q}_i \leq q_t \leq \check{q}_{\varepsilon/2}$ *with probability* $1 - \exp\left(-2\varepsilon^2 n/(1-6\varepsilon)^2\right)$.

Proof of Lemma 2.4.17. Define $f(q)$ equal to $\mathbb{E}\left[q_t \mid M_{t-1}\right] = \frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)q^2 - \frac{1-6\epsilon}{2}q + \frac{5-6\epsilon}{12}n - \frac{1-6\epsilon}{n}\left(\frac{s}{2}\right)^2$. We are going to evaluate $f(q)$ in $\bar{q}_{i+1}$ and in $\bar{q} = \frac{n(1+3\epsilon)}{3(1-6\epsilon)}$. We take care of different cases: first, we assume $i = -1$, with the condition that $s > 0$. Thus

$$\begin{aligned}f(\bar{q}_0) \leq{}& \frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)n^2\left[\frac{1}{9} + \frac{4\epsilon^2}{(1+6\epsilon)^2}\left(\frac{3}{2}\right)^6 - \frac{4\epsilon}{3(1+6\epsilon)}\left(\frac{3}{2}\right)^3\right] \\ &- \frac{1-6\epsilon}{2}n\left[\frac{1}{3} - \frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^3\right] + \frac{5-6\epsilon}{12}n \\ ={}& \frac{n}{3} + n\frac{3\epsilon^2(1-6\epsilon)}{(1+6\epsilon)^2}\left(\frac{3}{2}\right)^6 = \frac{n}{3}\left[1 + \frac{3^8\epsilon^2(1-6\epsilon)}{2^6(1+6\epsilon)^2}\right];\end{aligned}$$

now, we observe that

$$\begin{aligned}
&1+\frac{3^8\epsilon^2(1-6\epsilon)}{2^6(1+6\epsilon)^2}-\frac{1}{1-6\epsilon}\\
&<\frac{-6\epsilon}{1-6\epsilon}+\frac{3^8\epsilon^2}{2^6(1+6\epsilon)^2}\\
&=\frac{-2^6\cdot 6\epsilon(1+6\epsilon)^2+3^8\epsilon^2(1-6\epsilon)}{2^6(1-6\epsilon)(1+6\epsilon)^2}\\
&=\frac{3\epsilon\left(-2^7-3\cdot 2^9\epsilon-3^2\cdot 2^9\epsilon^2+3^7\epsilon-2\cdot 3^8\epsilon^2\right)}{2^6(1-6\epsilon)(1+6\epsilon)^2}\\
&=\frac{3\epsilon\left(-2^7+3\epsilon(3^6-2^9)-2\cdot 3^2\epsilon^2(2^8+3^6)\right)}{2^6(1-6\epsilon)(1+6\epsilon)^2}<0
\end{aligned}$$

where in the last inequality we have used that $-2^7+3\epsilon(3^6-2^9)<0$ for $\epsilon\leq\frac{1}{12}$. Thus, $f(\bar{q}_0)<\frac{n}{3(1-6\epsilon)}$.

Second, we assume $0\leq i\leq k-3$, with the condition that $s>\left(\frac{3}{2}\right)^i\beta n$.

$$\begin{aligned}
f(\bar{q}_{i+1})&\leq\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)n^2\left[\frac{1}{9}+\frac{4\epsilon^2}{(1+6\epsilon)^2}\left(\frac{3}{2}\right)^{4i+10}-\frac{4\epsilon}{3(1+6\epsilon)}\left(\frac{3}{2}\right)^{2i+5}\right]\\
&\quad-\frac{1-6\epsilon}{2}n\left[\frac{1}{3}-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i+5}\right]+\frac{5-6\epsilon}{12}n-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i}n\\
&=\frac{n}{3}+n\frac{3\epsilon^2(1-6\epsilon)}{(1+6\epsilon)^2}\left(\frac{3}{2}\right)^{4i+10}-n\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i}\\
&=\frac{n}{3}\left\{1+\frac{3\epsilon}{(1+6\epsilon^2)}\left(\frac{3}{2}\right)^{2i}\left[3\epsilon(1-6\epsilon)\left(\frac{3}{2}\right)^{2(i+5)}-2(1+6\epsilon)\right]\right\};
\end{aligned}$$

for the evaluation of $f(\bar{q}_{i+1})$ we observe that β is a constant in $(0,1)$ and that

$$3\epsilon(1-6\epsilon)\left(\frac{3}{2}\right)^{2(i+5)}-2(1+6\epsilon)<3\epsilon\left(\frac{2}{3}\frac{1}{\beta}\left(\frac{3}{2}\right)^3\right)^2-2<\frac{3^5}{2^7}-2<0$$

because $\left(\frac{3}{2}\right)^{i+2}<\frac{2}{3\beta}$ for $i+2\leq k$; thus $f(\bar{q}_{i+1})\leq\frac{n}{3}$.

Let now $i=k-2$; thus $s>\left(\frac{3}{2}\right)^{k-2}\beta n$. We now evaluate $f(\bar{q}_{k-1})$.

$$\begin{aligned}
f(\bar{q}_{k-1})&\leq\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)n^2\left[\frac{4}{81}+\frac{\epsilon^2}{9}+\frac{4\epsilon}{27}\right]-\frac{1-6\epsilon}{2}n\left[\frac{2}{9}+\frac{\epsilon}{3}\right]+\frac{5-6\epsilon}{12}n\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-4}n\\
&=\frac{-8(1-6\epsilon)+45-54\epsilon}{108}n-\frac{\epsilon(1-6\epsilon)}{18}n-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-4}n\\
&=\frac{37-12\epsilon+36\epsilon^2}{108}n-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-4}n\\
&=\frac{n}{3}\left[1+\frac{(1-6\epsilon)^2}{36}-\frac{6\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-4}\right]\\
&=\frac{n}{3}\left[1+\left(\frac{1-6\epsilon}{6}-\sqrt{\frac{6\epsilon}{1+6\epsilon}}\cdot\frac{3^{k-2}}{2^{k-2}}\right)\cdot\left(\frac{1-6\epsilon}{6}+\sqrt{\frac{6\epsilon}{1+6\epsilon}}\cdot\frac{3^{k-2}}{2^{k-2}}\right)\right].
\end{aligned}$$

Observe that, by definition of k, we have

$$\begin{aligned}&\frac{1-6\epsilon}{6}-\sqrt{\frac{6\epsilon}{1+6\epsilon}}\cdot\frac{3^{k-2}}{2^{k-2}}\leq\frac{1-6\epsilon}{6}-\sqrt{\frac{6\epsilon}{1+6\epsilon}}\cdot\frac{2^3}{3^3\beta}\\&=\frac{1-6\epsilon}{6}-\frac{\sqrt{3(1-6\epsilon)}\cdot 2^2}{3^3}=\frac{9-54\epsilon-8\sqrt{3(1-6\epsilon)}}{54}\\&<\frac{9-8\sqrt{3(1-6\epsilon)}}{54}<0\end{aligned}$$

for $\epsilon\leq\frac{1}{12}$. Thus, $f(\bar{q}_{k-1})\leq\frac{n}{3}$.

Let now $i=k-1$, which implies that $s>\left(\frac{3}{2}\right)^{k-1}\beta n$. We evaluate $f(\bar{q}_k)$:

$$\begin{aligned}f(\bar{q}_k)&\leq\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)n^2\left[\frac{1}{144}+\epsilon^2+\frac{\epsilon}{6}\right]-\frac{1-6\epsilon}{2}n\left[\frac{1}{12}+\epsilon\right]+\frac{5-6\epsilon}{12}n-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-2}n\\&=\frac{-7(1-6\epsilon)+80-96\epsilon}{192}n-\frac{3\epsilon(1-6\epsilon)}{8}n-\frac{2\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-2}n\\&=\frac{n}{3}\left[\frac{219-162\epsilon-216\epsilon+1296\epsilon^2}{192}-\frac{6\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-2}\right]\\&=\frac{n}{3}\left[1+\frac{27-378\epsilon+1296\epsilon^2}{192}-\frac{6\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-2}\right].\end{aligned}$$

By definition of k, we have that

$$\frac{27-378\epsilon+1296\epsilon^2}{192}-\frac{6\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2k-2}\leq\frac{27-378\epsilon+1296\epsilon^2}{192}-\frac{1-6\epsilon}{2}\left(\frac{2}{3}\right)^3<0,$$

where the first and the second inequalities hold for $\epsilon\leq\frac{1}{12}$. Thus, $f(\bar{q}_k)\leq\frac{n}{3}$.

We finally evaluate $f(\bar{q})$:

$$\begin{aligned}f(\bar{q})&\leq\frac{n}{12}\left(\frac{(1+3\epsilon)^2}{(1-6\epsilon)}\right)-\frac{n(1+3\epsilon)}{6}+\frac{n(5-6\epsilon)}{12}\\&=\frac{n}{12}\left(\frac{1-6\epsilon+9\epsilon^2}{1-6\epsilon}-2-6\epsilon+5-6\epsilon\right)\\&=\frac{n}{12(1-6\epsilon)}\left(1-6\epsilon+9\epsilon^2-2+6\epsilon+36\epsilon^2+5-36\epsilon+36\epsilon^2\right)\\&=\frac{n}{12}\left(\frac{4-36\epsilon+81\epsilon^2}{1-6\epsilon}\right)=\frac{n}{12}\left(\frac{(2-9\epsilon)^2}{1-6\epsilon}\right).\end{aligned}$$

It holds that $f(\bar{q})\leq\frac{n}{3(1-6\epsilon)}$ (remember that $\epsilon\leq 1/12$); in all cases, f is no more than $\frac{n}{3(1-6\epsilon)}$, and, from an immediate application of the additive Chernoff bound (Lemma B.2) with $\lambda=\frac{\epsilon n}{1-6\epsilon}$, and by observing that $\bar{q}_i\geq\bar{q}_{i+1}$, we get that

$$\bar{q}_i\leq q_t\leq\check{q}_{\epsilon/2}=\frac{n(1+3\epsilon)}{3(1-6\epsilon)}$$

with probability $1-\exp\left(-2\epsilon^2 n/(1-6\epsilon)^2\right)$. □

At the same time, the following lemma implies that the possible decrease of the bias cannot move it from S_i beyond S_{i-1}.

Lemma 2.4.18. *Let M_{t-1} be a configuration such that $s \geq \gamma\sqrt{n \log n}$ for some constant $\gamma > 0$. If $s > \left(\frac{3}{2}\right)^i \beta n$ for some $0 \leq i \leq k$, then $s_t > \inf(S_{i-1})$ with probability $1 - \exp\left(-\varepsilon^2 (3/2)^{2i} \beta^2 n/2\right)$.*

Proof of Lemma 2.4.18. We have

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \geq s\left(\frac{5}{6} - \varepsilon\right).$$

The Hoeffding bound (Lemma B.3) implies that $s_t \geq s\left(\frac{5}{6} - 2\varepsilon\right)$ with probability $1 - \exp\left(-\varepsilon^2 (3/2)^{2i} \beta^2 n/2\right)$. Observe that

$$s\left(\frac{5}{6} - 2\varepsilon\right) > \left(\frac{3}{2}\right)^i \beta n \left(\frac{5}{6} - 2\varepsilon\right) \geq \left(\frac{3}{2}\right)^{i-1} \beta n = S_{i-1}$$

since $\frac{3}{2}\left(\frac{5}{6} - 2\varepsilon\right) \geq 1$, which concludes the proof. □

We still need to control the process when $q > \breve{q}_{\varepsilon/2}$. The next lemma addresses this issue by showing that there is a decrease of the number of undecided nodes when they are more than $\breve{q}_\varepsilon = (n/3)/(1 - 6\varepsilon)$, and provides a lower bound on the decrease, depending on the bias.

Lemma 2.4.19. *Let M_{t-1} be a configuration such that $q \geq \breve{q}_\varepsilon$. Then, it holds that*

(1) $q_t \leq q\,(1 - \varepsilon)$ with probability $1 - \exp\left(-2\varepsilon^2 (n/3)/(1 - 6\varepsilon)^2\right)$;

(2) if $s \leq \sup(S_i)$ for any $-1 \leq i \leq k - 1$, then $q_t \geq \bar{q}_{i+1}$ with probability $1 - \exp\left(-\varepsilon^2 n/18\right)$.

Proof of Lemma 2.4.19. Consider the first claim. We define $f(q) = \frac{3}{4}\left(\frac{1-6\epsilon}{n}\right) q^2 - \frac{1-6\epsilon}{2} q + \frac{5-6\epsilon}{12} n$. We now show that $f(q) \leq q\,(1 - 2\epsilon)$. Indeed, $f(q) - q\,(1 - 2\epsilon)$ is equal to

$$\begin{aligned} &\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right) q^2 - \left(\frac{1-6\epsilon}{2} + 1 - 2\epsilon\right) q + \frac{5-6\epsilon}{12} n \\ =\ &\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right) q^2 - \left(\frac{3-10\epsilon}{2}\right) q + \frac{5-6\epsilon}{12} n. \end{aligned}$$

This expression is a convex parabola which has its maximum in either $\bar{q}_1 = \frac{n}{3(1-6\epsilon)}$ or $\bar{q}_2 = n$. We calculate $f(q) - q\,(1 - 2\epsilon)$ in these two points

$$\begin{aligned} &\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right) q_1^2 - \left(\frac{3-10\epsilon}{2}\right) q_1 + \frac{5-6\epsilon}{12} n \\ =\ &\frac{n^2}{12(1-6\epsilon)}\left(1 - 6 + 20\epsilon + 5 - 36\epsilon + 36\epsilon^2\right) \\ =\ &\frac{n^2}{12(1-6\epsilon)}\left[-4\epsilon(4 - 9\epsilon)\right] < 0 \end{aligned}$$

for all $0 < \epsilon \leq \frac{1}{12}$. At the same time it holds that

$$\begin{aligned}
&\frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)q_2^2 - \left(\frac{3-10\epsilon}{2}\right)q_2 + \frac{5-6\epsilon}{12}n \\
&= \frac{n}{12}\left[9 - 54\epsilon - 18 + 60\epsilon + 5 - 6\epsilon\right] \\
&= -\frac{-n}{3} < 0.
\end{aligned}$$

Thus, $\mathbb{E}\left[q_t \mid M_{t-1}\right] \leq q(1-2\epsilon)$. The additive form of Chernoff bound (Lemma B.2) with $\lambda = \frac{\epsilon n}{3(1-6\epsilon)}$ implies that $q_t < q(1-\epsilon)$ with probability $1 - \exp\left(-2\epsilon^2(n/3)/(1-6\epsilon)^2\right)$.

As for the second claim, we consider two cases. First, assume $i < k-1$, thus $s \leq \left(\frac{3}{2}\right)^{i+1}\beta n$. From Eq. (2.6), we observe that

$$\begin{aligned}
\mathbb{E}\left[q_t \mid M_{t-1}\right] &\geq \frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)\frac{n^2}{9(1-6\epsilon)^2} + \frac{5-6\epsilon}{12}n - \left(\frac{1-6\epsilon}{2}\right)\frac{n}{3(1-6\epsilon)} - \frac{1-6\epsilon}{n}\left(\frac{s}{2}\right)^2 \\
&\geq \frac{n}{12}\left(\frac{4-24\epsilon+36\epsilon^2}{1-6\epsilon}\right) - \frac{2n\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i+2} \\
&\geq \frac{n}{3(1-6\epsilon)}\left(1-3\epsilon\right)^2 - \frac{2n\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i+2} \\
&\geq \frac{n}{3} - \frac{2n\epsilon}{1+6\epsilon}\left(\frac{3}{2}\right)^{2i+5} + \frac{19n\epsilon}{4}\left(\frac{3}{2}\right)^{2i+2} \\
&= \bar{q}_{i+1} + \frac{19n\epsilon}{4}\left(\frac{3}{2}\right)^{2i+2}.
\end{aligned}$$

We conclude applying the additive Chernoff bound with $\lambda = \frac{19\epsilon n}{4}\left(\frac{3}{2}\right)^{2i+2}$, obtaining $q_t \geq \bar{q}_{i+1}$ with probability $1 - \exp\left(-19^2\epsilon^2(n/8)(3/2)^{4i+4}\right)$.

Second, let $i = k-1$; then $s \leq \frac{2}{3}n$. As before

$$\begin{aligned}
\mathbb{E}\left[q_t \mid M_{t-1}\right] &\geq \frac{3}{4}\left(\frac{1-6\epsilon}{n}\right)\frac{n^2}{9(1-6\epsilon)^2} + \frac{5-6\epsilon}{12}n - \left(\frac{1-6\epsilon}{2}\right)\frac{n}{3(1-6\epsilon)} - \frac{1-6\epsilon}{n}\left(\frac{s}{2}\right)^2 \\
&\geq \frac{n}{12(1-6\epsilon)}\left(4-24\epsilon+36\epsilon^2\right) - \frac{n(1-6\epsilon)}{9} \geq \frac{n\left(1-3\epsilon\right)^2}{3(1-6\epsilon)} - \frac{n(1-6\epsilon)}{9} \\
&= \frac{n}{9(1-6\epsilon)}\left(2-6\epsilon-9\epsilon^2\right) = \frac{n}{12}\left(\frac{4(2-6\epsilon-9\epsilon^2)}{3(1-6\epsilon)}\right) \\
&\geq \frac{n}{12}(1+14\epsilon) = \bar{q}_k + \frac{\epsilon n}{6}
\end{aligned}$$

since

$$\frac{4(2-6\epsilon-9\epsilon^2)}{3(1-6\epsilon)} - (1+14\epsilon) > \frac{5-48\epsilon}{3(1-6\epsilon)} > 0$$

for $\epsilon \leq \frac{1}{12}$. Thus, we conclude applying the additive form of Chernoff bound with $\lambda = \frac{\epsilon}{6}n$, obtaining $q_t \geq \bar{q}_k$ with probability $1 - \exp\left(-\epsilon^2 n/18\right)$. Hence, the second claim holds with probability at least $1 - \exp\left(-\epsilon^2 n/18\right)$, which is less than $1 - \exp\left(-19^2\epsilon^2(n/8)(3/2)^{4i+4}\right)$. □

The next and last lemma guarantees that once the process reaches a configuration having bias $\mathcal{O}(\sqrt{n \log n})$, then it "enters" a metastable phase that lasts $\Omega(n^{\lambda'})$ rounds w.h.p. in which the absolute value of the bias remains $\mathcal{O}(\sqrt{n \log n})$, since it can be used symmetrically when the bias is negative.

Lemma 2.4.20. *Let M_{t-1} be any configuration. If $s \leq \gamma\sqrt{n \log n}$ for some constant $\gamma > 0$ and $q \leq \frac{n}{3}\left(\frac{1+3\varepsilon}{1-6\varepsilon}\right)$, then $s_t \leq 2\gamma\sqrt{n \log n}$ with probability $1 - n^{-\gamma^2/2}$.*

Proof of Lemma 2.4.20. From Eq. (2.5), we have

$$\mathbb{E}\left[s_t \mid M_{t-1}\right] \leq \gamma\sqrt{n \log n}\left(\frac{5}{6} - \varepsilon + \frac{1}{6} + \frac{\varepsilon}{2}\right) \leq \gamma\sqrt{n \log n}.$$

We conclude applying the Hoeffding bound with $\lambda = \gamma\sqrt{\log n / n}$, obtaining that $s_t \leq 2\gamma\sqrt{n \log n}$ with probability $1 - n^{-\gamma^2/2}$. □

2.4.5.3 Proof of Theorem 2.3.3

If $p > 3/4$, then there is an $\varepsilon > 0$ such that $p > 3/4 + \varepsilon$. Lemmas 2.4.12 and 2.4.13, and Lemmas A.1 and A.2 imply that the process reaches a configuration $\mathbf{y}$ with bias $|s(\mathbf{y})| \leq (\gamma/2) \cdot \sqrt{n \log n}$ within time $\mathcal{O}_{p,\gamma}(\log n)$ w.h.p., since ε can be chosen as $(4p-1)/8$. Furthermore, the bias remains bounded by the same value for $n^{\Omega_{p,\gamma}(1)}$ rounds, w.h.p.

Let us now assume $1/2 + 3\varepsilon = p \leq 3/4$ for any $\varepsilon > 0$. The proof is divided into different cases. Let $\gamma > 0$ be any constant. We remark that ε depends only on p.

(1) $s \in S_i$ for some $-1 \leq i \leq k-1$ and $s \geq (\gamma/2) \cdot \sqrt{n \log n}$,
  (1.1) $\bar{q}_{i+1} \leq q \leq \frac{n(1+3\varepsilon)}{3(1-6\varepsilon)}$: the bias decreases exponentially fast each round, w.h.p., until $|s| \leq (\gamma/2) \cdot \sqrt{n \log n}$ due to the combination of Lemmas A.1 and A.2, Lemma 2.4.12, and Lemmas 2.4.15 to 2.4.18. This phase lasts $\mathcal{O}_{p,\gamma}(\log n)$ rounds;
  (1.2) $\frac{n(1+3\varepsilon)}{3(1-6\varepsilon)} < q$: Lemmas A.1 and A.2, and Lemmas 2.4.16 and 2.4.19 imply that in $\mathcal{O}_p(\log n)$ rounds the number of undecided nodes reaches the interval $\left[\bar{q}_{i+1}, \frac{n}{3(1-6\varepsilon)}\right]$ where i is such that the round before the undecided nodes become less than $\frac{n}{3(1-6\varepsilon)}$, the bias is in S_i: recall that during this whole process (which lasts $\mathcal{O}_p(\log n)$ rounds) the bias never goes over $\frac{2}{3}n$ thanks to Lemma 2.4.14. At the same time, the value of the bias will belong to one set between $S_{i-1}, \ldots, S_{k-1}$ due to Lemma 2.4.18. Since $\bar{q}_i$ is a decreasing sequence, we are in Case 1.1, and we conclude.
  (1.3) $q < \bar{q}_{i+1}$: in this case, Lemma 2.4.16 implies $q_t \geq \bar{q}_{i+1}$ in the next round, w.h.p. Since Lemma 2.4.14 guarantees that the bias remains under the value $\frac{2}{3}n$ w.h.p., either we are in Case 1.1 or in Case 1.2, and we conclude.
(2) $s > \frac{2}{3}n$: Lemma 2.4.14 implies that the bias gets less than or equal to $\frac{2}{3}n$ in $\mathcal{O}_p(\log n)$ rounds, w.h.p.; then we are in Case 1 and we conclude.

Now, we can suppose the process starts from a configuration $\mathbf{y}$ having bias $0 \leq s(\mathbf{y}) \leq (\gamma/2) \cdot \sqrt{n \log n}$, and such that $\bar{q}_0 \leq q(\mathbf{y}) \leq \frac{n}{3}\left(\frac{1+3\varepsilon}{1-6\varepsilon}\right)$, as Case 1.1 or 1.3 leaves it. In the next round, it holds that the number of undecided nodes is $\bar{q}_{-1} \leq q_t \leq \frac{n}{3}\left(\frac{1+3\varepsilon}{1-6\varepsilon}\right)$, w.h.p., due to Lemma 2.4.17; at the same time, w.h.p., $|s_t| \leq \gamma\sqrt{n \log n}$ for Lemmas 2.4.12 and 2.4.20 (which can be used symmetrically on $A - B$ and $B - A$). Thus, the absolute value of the bias is either still less than $(\gamma/2) \cdot \sqrt{n \log n}$ or has become greater than or equal to $|s_t| \geq (\gamma/2) \cdot \sqrt{n \log n}$, in which case it

starts decreasing exponentially fast each round, w.h.p. for Lemma 2.4.15, until becoming smaller than $(\gamma/2) \cdot \sqrt{n \log n}$ (as explained in Case 1.1, which works analogously if the bias is negative, because of symmetry). This phase lasts $n^{\Omega_{p,\gamma}(1)}$ rounds w.h.p. for Lemmas A.1 and A.2.

2.5 Best-of-N nest-site selection process in honeybees.

In both its non-noisy and noisy versions, the UNDECIDED-STATE dynamics can be seen as a specific case of a best-of-N selection process in honeybees [Reina et al., 2017]. In this section, we show how to derive it. We first describe an equation system that models the decision process of a bee swarm to select a new nest among N possible ones (the *options*). This mathematical model has been introduced in [Seeley et al., 2012] and also analyzed in [Reina et al., 2017] (see Eq.s (1) in the latter work). In the next subsection, we show how to derive the UNDECIDED-STATE dynamics.

In detail, consider any population of bees. Let x_i denote the proportion of bees agreeing on (which are *committed* to) option i, and x_u be the proportion of bees which are undecided (or *uncommitted*). The dynamics is described by the following system

$$\begin{cases} \frac{\mathrm{d}x_i}{\mathrm{d}t} = \gamma_i x_u - \alpha_i x_i + \rho_i x_u x_i - x_i \sum_{1 \le i \le N} \beta_{ji} x_j; \\ x_u = 1 - \sum_{1 \le i \le N} x_i. \end{cases} \tag{2.10}$$

We now provide an intuitive explanation of the equation system, and the description of the parameters. Uncommitted bees explore the environment and discover possible options. While encountering option i, the uncommitted bee estimates its quality ν_i and may commit to that option at a rate γ_i, where $\gamma_i \propto \nu_i$ (more frequent commitments to better quality nests). Committed bees may spontaneously revert, through *abandonment*, to an uncommitted state at rate α_i. The abandonment rate is assumed to be inversely proportional to the option's quality α_i ($\propto \nu_i^{-1}$). This abandonment feature allows bees to quickly discard bad options, and endows the swarm with a degree of flexibility since bees are not locked into their commitment state [Reina et al., 2017]. In addition to these two spontaneous individual transitions, bees interact with each other to achieve agreement in two forms: *recruitment* and *cross-inhibition*. The rate at which uncommitted bees are recruited to option i depends on the number of bees committed to i and on the strength of the recruitment process for i, that is ρ_i ($\propto \nu_i$) . Thus, the recruitment process is a form of positive feedback. On the other hand, cross-inhibition is a form of negative feedback: when a bee committed to option i encounters another bee committed to another option j, (with $j \neq i$), the first may deliver stop signals to the second which reverts to an uncommitted state at a rate β_{ji} ($\propto \nu_i$).

We discretize Eq. (2.10) by using the forward Euler method [Iserles, 2008] with $\mathrm{d}t = 1$, obtaining

$$\begin{cases} x_i^{(t+1)} = \gamma_i x_u^{(t)} - (\alpha_i + 1) x_i^{(t)} + \rho_i x_u^{(t)} x_i^{(t)} - x_i^{(t)} \sum_{1 \le i \le N} \beta_{ji} x_j^{(t)}; \\ x_u^{(t+1)} = 1 - \sum_{1 \le i \le N} x_i^{(t+1)}. \end{cases} \tag{2.11}$$

As we show in the next paragraph, a specific parameter choice yields exactly the mean-field behavior of the UNDECIDED-STATE dynamics.

Derivation of the UNDECIDED-STATE dynamics. Consider the equation system in Eq. (2.11), and the following parameter choice [3] (for $i, j = 1, ..., N$)

$$\begin{aligned}
&\rho_i = 1 - p;\\
&\alpha_i = \frac{p}{N+1};\\
&\gamma_i = \frac{p}{N+1};\\
&\beta_{i,i} = 0;\\
&\beta_{i,j} = 1 - p; \text{ if } i \neq j,
\end{aligned}$$

which yields the following equations:

$$x_i^{(t+1)} = x_i^{(t)} + \frac{p}{N+1}x_u^{(t)} - \frac{p}{N+1}x_i^{(t)} + (1-p)\,x_u^{(t)}x_i^{(t)} - \sum_{\substack{1\leq j\leq N\\ j\neq i}} (1-p)\,x_i^{(t)}x_j^{(t)}; \quad (2.12)$$

$$x_u^{(t+1)} = 1 - \sum_{1\leq i\leq N} x_i^{(t+1)}. \quad (2.13)$$

We can now observe that Eqs. (2.12) and (2.13) are the mean-field equations of the UNDECIDED-STATE process we consider in our work. Indeed, if y_i represents the number of agents supporting opinion i at some given configuration M_t, Y_i is the r.v. yielding the number of agents supporting opinion i at the next round (y_u, Y_u have analogous meaning for the undecided state), by substituting $x_i^{(t+1)} = \mathbb{E}\left[Y_i \mid M_t\right]$, $x_i^{(t)} = y_i$ (the same for y_u, Y_u) we obtain the expected round-by-round behavior of the UNDECIDED-STATE process. If $N = 2$, we get exactly Eqs. (2.3), (2.4) and (2.6).

Significance of the parameters. We assume that the two opinion/options (nest-sites) *Alpha* and *Beta* have the same quality. The rate of the recruitment process for option i is, in our setting, $\rho_i = 1 - p$, which implies that only noise is responsible for the non recruitment of uncommitted bees by committed bees. Since all nests have the same qualities, we choose the rate at which a bee committed to option i reverts to an uncommitted state when encountering a bee committed to option $j \neq i$ to be $B_{j,i} = 1 - p = \rho_i$. The abandonment rate is the spontaneous reversion to an uncommitted state from a committed state (say, committed to option i), and here it is assumed to be governed only by noise at a rate $\alpha_i = p/3$. Such choices are reasonable in the setting described in Section 2.5. The only difference lies in the choice of γ_i, the spontaneous commission rate to option i for an uncommitted bee. Usually, such parameter is proportional to the nest quality. In our setting, the nest quality plays no role in such a spontaneous phenomenon; instead, only noise is responsible, which corresponds to the choice $\gamma_i = \alpha_i = p/3$. Even though, our setting is a strong restriction of the more general framework described by Eq. (2.11) ([Reina et al., 2017]), we believe it effectively describes a realistic scenario in which there is a somewhat equivalence between the option quality and the evolution of the system is dominated by another important factor (i.e., noise). In particular, based on our analysis on the convergence time, this setting represents a sort of worst-case scenario for the swarm system that, in general, requires a relatively longer symmetry-breaking phase but still achieves an almost-agreement.

3. We remark that these parameter choice is not the common one indicated in [Reina et al., 2017].

2.6 Analysis of the 3-MAJORITY dynamics

In this section we analyze the 3-MAJORITY dynamics: first, we give some preliminary results. Afterwards, in Section 2.6.1 we prove Theorem 2.3.4, in Section 2.6.2 we prove Theorem 2.3.5, while Section 2.6.3 is devoted to the proof Theorem 2.3.6.

Notice that in the 3-MAJORITY dynamics, a configuration of the system at time t is completely determined by the value of the bias at time t, i.e., $s_t = a_t - b_t$, since $a_t = (n + s_t)/2$ and $b_t = (n - s_t)/2$). Thus, in this section, we refer to the induced process by s_t instead of M_t.

We now give the expectation of the bias at time t, conditional on its value at time $t - 1$.

Lemma 2.6.1. *Let $\{s_t\}_{t\geq 0}$ be the process induced by the* 3-MAJORITY *dynamics with uniform noise probability $p \in (0, 1)$. The conditional expectation of the bias is*

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] = \frac{s(1-p)}{2}\left(3 - \frac{s^2}{n^2}(1-p)^2\right). \tag{2.14}$$

Proof. Let $b = b_t$ and $a = a_t$. Then $s = b - a$ and $n = a + b$, which implies $b = (n + s)/2$ and $a = (n - s)/2$. The probability that, when a node samples a neighbor, it receives opinion *Beta* is $b' = (b/n) \cdot (1 - p) + p/2$, where $(b/n) \cdot (1 - p)$ is the probability to receive a non-noisy message which contains opinion *Beta*, and $p/2$ is the contribution of the noise. Analogously, the probability that it receives opinion *Alpha* is $a' = (a/n) \cdot (1-p)+p/2$. Then, the probability the node updates its opinion to *Beta* is $(b')^3 + 3a'(b')^2$. So, for Eq. (2.1), we have that

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] = 2n\left((b')^3 + 3a'(b')^2\right) - n = \frac{s(1-p)}{2}\left(3 - \frac{s^2}{n^2}(1-p)^2\right),$$

where the last equation follows from simple calculations. □

By the lemma above, we deduce that there are up to three equilibrium configurations in expectation. The first one corresponds to $s = 0$, and the other (possible) equilibrium correspond to the condition

$$\frac{1-p}{2}\left(3 - \frac{s^2}{n^2}(1-p)^2\right) = 1$$

The latter condition results in

$$s = \pm\frac{n}{(1-p)} \cdot \sqrt{\frac{3(1-p)-2}{(1-p)}} = \pm\frac{n}{(1-p)} \cdot \sqrt{\frac{1-3p}{1-p}},$$

which is well defined if only if $p \leq 1/3$. We will denote the absolute value of the latter two values by s_{eq}.

2.6.1 3-MAJORITY dynamics: victory of the majority

The aim of this subsection is to prove Theorem 2.3.4: so, in each statement we assume that $\{s_t\}_{t\geq 0}$ is the process induced by the 3-MAJORITY dynamics with uniform noise probability $p < 1/3$.

We first show a lemma which states that, for any small constant $\varepsilon > 0$, whenever $s_{t-1} \notin [(1 - \varepsilon)s_{\text{eq}}, (1 + \varepsilon)s_{\text{eq}}]$, then s_t gets closer to the interval.

Lemma 2.6.2. *For any constant $\varepsilon > 0$ such that $\varepsilon^2 < (1-3p)/2$ and for any $\gamma > 0$, if $s \geq \gamma\sqrt{n \log n}$, the following hold*

(i) if $s \leq (1-\varepsilon)s_{eq}$, then Pr $[s_t \geq (1+3\varepsilon^2/4)s \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4/32}}$;

(ii) if, $s \geq (1+\varepsilon)s_{eq}$, then Pr $[s_t \leq (1-3\varepsilon^2/4)s \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4/32}}$.

Proof. We first notice that

$$(1-\varepsilon)s_{\text{eq}} \leq \frac{n}{1-p}\sqrt{\frac{1-3p-2\varepsilon^2}{1-p}}, \tag{2.15}$$

which holds since $\varepsilon^2 \leq (1-3p)/2$ and can be proved with simple calculations.

For Lemma 2.6.1, if each $s \leq (1-\varepsilon)s_{\text{eq}}$, then

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] = \frac{s(1-p)}{2}\left(3 - \frac{s^2}{n^2}(1-p)^2\right) \geq s\left(\frac{3-3p}{2} - \frac{1-3p-2\varepsilon^2}{2}\right) = s(1+\varepsilon^2).$$

where the inequality follows from (2.15). Since (2.2), for the Hoeffding bound (Lemma B.3), it holds that

$$\Pr\left[s_t \leq s(1+\varepsilon^2) - s\varepsilon^2/4 \mid s_{t-1} = s\right] \leq e^{-s^2\varepsilon^4/(32n)} \leq e^{-\gamma^2\varepsilon^4 \log n/32} \leq \frac{1}{n^{\gamma^2\varepsilon^4/32}}.$$

The second inequality in the lemma follows by a symmetric argument, observing that

$$(1+\varepsilon)s_{\text{eq}} \geq \frac{n}{1-p}\sqrt{\frac{1-3p+2\varepsilon^2}{1-p}},$$

for ε such that $\varepsilon^2 < (1-3p)/2$. □

The following lemma serves to bound how far the bias can get from the interval $[(1+\varepsilon)s_{\text{eq}}, (1-\varepsilon)s_{\text{eq}}]$.

Lemma 2.6.3. *For any constants $\varepsilon > 0$ and $\gamma > 0$, if $s \geq \gamma\sqrt{n \log n}$, the following hold*

(i) if $s \leq (1+\varepsilon)s_{eq}$, then Pr $[s_t \geq (1-\varepsilon-\varepsilon^2)s \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^2/16}}$;

(ii) if $s \geq (1-\varepsilon)s_{eq}$ with $\varepsilon < 1$, then Pr $[s_t \leq (1+\varepsilon)s \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^2p^2}}$.

Proof. The proof is similar to that of the previous lemma. From Lemma 2.6.1, we get that

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \geq s\left(1 - \varepsilon - \frac{\varepsilon^2}{2}\right),$$

which follows since $s \leq (1+\varepsilon)s_{\text{eq}}$ by simple calculations. For the Hoeffding bound (Lemma B.3) we get

$$\Pr\left[s_t \leq s\left(1-\varepsilon-\frac{\varepsilon^2}{2}\right) - \frac{\varepsilon^2 \cdot s}{2} \,\middle|\, s_{t-1} = s\right] \leq e^{-\frac{\gamma^2\varepsilon^4}{16}} = \frac{1}{n^{\frac{\gamma^2\varepsilon^2}{16}}}.$$

The second claim comes symmetrically from Lemma 2.6.1 by observing that, since $s \geq (1-\varepsilon)s_{\text{eq}}$

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \leq s\left(1 + (1-3p)\varepsilon\right).$$

The Hoeffding bound implies

$$\Pr\left[s_t \geq s\left(1+\varepsilon\right) \mid s_{t-1} = s\right] \leq \Pr\left[s_t \geq s\left(1+(1-3p)\varepsilon\right) + 2p\varepsilon \cdot s \mid s_{t-1} = s\right] \leq e^{-\gamma^2\varepsilon^2p^2} = \frac{1}{n^{\gamma^2\varepsilon^2p^2}}.$$

□

We provide another lemma to control the behavior of the bias. The proof consists in the application of simple concentration bounds.

Lemma 2.6.4. *For any constant $k > 0$, the following hold:*

(i) if $s \geq s_{eq}$, then $Pr[s_t \geq 2s_{eq}/3 \mid s_{t-1} = s] \geq 1 - 1/n^k$.

(ii) if $0 \leq s \leq 2s_{eq}/3$, then $Pr[s_t \leq s_{eq} \mid s_{t-1} = s] \geq 1 - 1/n^k$.

Proof. Let k be any arbitrarily large constant. As for (i), Lemma 2.6.1 gives that

$$\mathbb{E}[s_t \mid s_{t-1} = s] \geq s(1-p) \geq s_{\text{eq}}(1-p),$$

since $s_{\text{eq}} \leq s \leq n$. Then, let $\delta = (1-3p)/3 > 0$. By using the Hoeffding bound, it holds that

$$\Pr[s_t \leq s_{\text{eq}}(1-p) - \delta \cdot s_{\text{eq}} \mid s_{t-1} = s] \leq e^{-\frac{\delta^2 s_{\text{eq}}^2}{4}} \leq \frac{1}{n^k},$$

where the latter inequality holds since $s_{\text{eq}} = \Theta(n)$ and $s_{\text{eq}} > (2k/\delta)\log n$ for a sufficiently large n. As for (ii), Lemma 2.6.1 implies that

$$\mathbb{E}[s_t \mid s_{t-1} = s] \leq \frac{3s(1-p)}{2} \leq s_{\text{eq}}(1-p),$$

which is true since $0 \leq s \leq 2s_{\text{eq}}/3$. The Hoeffding bound then gives

$$\Pr[s_t \geq s_{\text{eq}}(1-p) + ps_{\text{eq}} \mid s_{t-1} = s] \leq e^{-\frac{p^2 s_{\text{eq}}^2}{4}} \leq \frac{1}{n^k},$$

where the latter inequality holds since $s_{\text{eq}} = \Theta(n)$ and so $s_{\text{eq}} > (2k/p)\log n$ for a sufficiently large n.

□

We can piece together the above lemmas, which imply the following corollary.

Corollary 2.6.5. *For any constant $\varepsilon > 0$ such that $\varepsilon < 1/3$ and $\varepsilon^2 < (1-3p)/2$, the following hold:*

(i) if $|s_{eq} - s| \leq (\varepsilon/4)s_{eq}$, then

$$Pr[|s_{eq} - s_t| \leq \varepsilon s_{eq} \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^2p^2/2^5}};$$

(ii) if $(\varepsilon/4)s_{eq} \leq |s_{eq} - s| \leq s_{eq}/3$, then

$$Pr\left[|s_{eq} - s_t| \leq |s_{eq} - s| \cdot \left(1 - \frac{3\varepsilon^2}{2^5}\right) \,\middle|\, s_{t-1} = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4p^2/(2^{18}3^2)}}.$$

Proof. First, we prove (i). From Lemma 2.6.3 and the union bound, we have that

$$\Pr\left[\left(1 - \frac{\varepsilon}{4} - \frac{\varepsilon^2}{16}\right) \cdot \left(1 - \frac{\varepsilon}{4}\right) s_{\text{eq}} \leq s_t \leq \left(1 + \frac{\varepsilon}{4}\right) \cdot \left(1 + \frac{\varepsilon}{4}\right) s_{\text{eq}} \,\middle|\, s_{t-1} = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^2p^2/2^5}}.$$

The claim follows by osberving that

$$\left[\left(1-\frac{\varepsilon}{4}-\frac{\varepsilon^2}{16}\right)\cdot\left(1-\frac{\varepsilon}{4}\right)s_{\text{eq}},\left(1+\frac{\varepsilon}{4}\right)\cdot\left(1+\frac{\varepsilon}{4}\right)s_{\text{eq}}\right]\subseteq\left[(1-\varepsilon)s_{\text{eq}},(1+\varepsilon)s_{\text{eq}}\right].$$

As for claim (ii), we divide the proof in two different cases. Suppose, first, that $2s_{\text{eq}}/3 \leq s \leq (1-\varepsilon/4)s_{\text{eq}}$. A constant $\varepsilon/4 \leq \delta \leq 1/3$ exists such that $s=(1-\delta)s_{\text{eq}}$. Then, from Lemmas 2.6.2 and 2.6.3

$$\Pr\left[(1-\delta)\left(1+\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}\leq s_t\leq s_{\text{eq}}\,\middle|\,s_{t-1}=s\right]\geq 1-\frac{1}{n^{\gamma^2\varepsilon^4p^2/2^{14}}}.$$

Notice that

$$\begin{aligned}&\left|s_{\text{eq}}-(1-\delta)\left(1+\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}\right| = s_{\text{eq}}-(1-\delta)\left(1+\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}\\ &=\left(s_{\text{eq}}-(1-\delta)s_{\text{eq}}\right)\cdot\left[1-\frac{(1-\delta)\cdot\frac{3\varepsilon^2}{2^6}\cdot s_{\text{eq}}}{\delta\cdot s_{\text{eq}}}\right]\leq(s_{\text{eq}}-s)\cdot\left[1-\frac{3\varepsilon^2}{2^5}\right],\end{aligned}$$

where in the last inequality we used that $\delta<1/3$. Hence,

$$\Pr\left[\left|s_{\text{eq}}-s_t\right|\leq\left|s_{\text{eq}}-s\right|\cdot\left[1-\frac{3\varepsilon^2}{2^5}\right]\,\middle|\,s_{t-1}=s\right]\geq 1-\frac{1}{n^{\gamma^2p^2\varepsilon^4/2^{14}}}. \tag{2.16}$$

Second, suppose $(1+\varepsilon/4)s_{\text{eq}}\leq s\leq 3s_{\text{eq}}/2$. Again, a constant $\varepsilon/4\leq\delta\leq 1/3$ exists such that $s=(1+\delta)s_{\text{eq}}$. From Lemmas 2.6.2 and 2.6.3, it holds that

$$\Pr\left[(1+\delta)\left(1-\delta-\delta^2\right)s_{\text{eq}}\leq s_t\leq(1+\delta)\left(1-\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}\,\middle|\,s_{t-1}=s\right]\geq 1-\frac{1}{n^{\gamma^2\varepsilon^4p^2/(2^{18}3^2)}},$$

for the union bound, since $\delta\leq 1/3$. Notice that

$$\begin{aligned}&\left|s_{\text{eq}}-(1+\delta)\left(1-\delta-\delta^2\right)s_{\text{eq}}\right| = s_{\text{eq}}-(1+\delta)\left(1-\delta-\delta^2\right)s_{\text{eq}}\\ &=\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[\frac{(1+\delta)(\delta+\delta^2)s_{\text{eq}}}{\delta s_{\text{eq}}}-1\right]\leq\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[\frac{16}{9}-1\right]\\ &=\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[1-\frac{2}{9}\right],\end{aligned}$$

where the inequality holds since $\delta\leq 1/3$. By simple calculations, it can be seen that $(1+\delta)\left(1-\frac{3\varepsilon^2}{2^6}\right)\geq 1$. Then, we have also that

$$\begin{aligned}&\left|s_{\text{eq}}-(1+\delta)\left(1-\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}\right| = (1+\delta)\left(1-\frac{3\varepsilon^2}{2^6}\right)s_{\text{eq}}-s_{\text{eq}}\\ &=\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[1-\frac{(1+\delta)\cdot\frac{3\varepsilon^2}{2^6}s_{\text{eq}}}{\delta s_{\text{eq}}}\right]\overset{(a)}{\leq}\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[1-3\left(1+\frac{\varepsilon}{4}\right)\cdot\frac{3\varepsilon^2}{2^6}\right]\\ &\overset{(b)}{\leq}\left((1+\delta)s_{\text{eq}}-s_{\text{eq}}\right)\cdot\left[1-\frac{9\varepsilon^2}{2^6}\right],\end{aligned}$$

where (a) holds since $\varepsilon/4 \leq \delta \leq 1/3$, and (b) holds since $\varepsilon > 0$. Thus,

$$\Pr\left[\left|s_{\text{eq}} - s_t\right| \leq \left|s_{\text{eq}} - s\right| \cdot \left[1 - \frac{9\varepsilon^2}{2^6}\right] \,\middle|\, s_{t-1} = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4 p^2/(2^{18}3^2)}}\,. \tag{2.17}$$

Combining Eqs. (2.16) and (2.17), we get that, whenever $(\varepsilon/4)s_{\text{eq}} \leq \left|s_{\text{eq}} - s\right| \leq s_{\text{eq}}/3$, then

$$\Pr\left[\left|s_{\text{eq}} - s_t\right| \leq \left|s_{\text{eq}} - s\right| \cdot \left[1 - \frac{3\varepsilon^2}{2^5}\right] \,\middle|\, s_{t-1} = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4 p^2/(2^{18}3^2)}}\,.$$

□

We are finally ready to prove the theorem.

Proof of Theorem 2.3.4. We divide the proof in different cases. First, suppose that $(\varepsilon/4)s_{\text{eq}} \leq \left|s_{\text{eq}} - s\right| \leq \varepsilon s_{\text{eq}}$. Let $T_1 = n^{\gamma^2\varepsilon^4 p^2/(2^{19}3^2)}$. Then, from Corollary 2.6.5.(i) and (ii), for the chain rule, we have that

$$\Pr\left[\bigcap_{k=1}^{T}\left\{\left|s_{\text{eq}} - s_{t+k}\right| \leq \varepsilon s_{\text{eq}}\right\} \,\middle|\, s_t = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4 p^2/(2^{20}3^2)}}.$$

Second, suppose that $\varepsilon s_{\text{eq}} \leq \left|s_{\text{eq}} - s\right| \leq s_{\text{eq}}/3$. Then, from Corollary 2.6.5.(ii), for the chain rule, a time T_2 exists, with

$$T_2 = \mathcal{O}\left(-\frac{\log n}{\log\left(1 - \frac{3\varepsilon^2}{2^5}\right)}\right) = \mathcal{O}\left(\log n/\varepsilon^2\right)$$

such that

$$\Pr\left[\left|s_{\text{eq}} - s_{t+T_2}\right| \leq \varepsilon s_{\text{eq}} \mid s_t = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4 p^2/(2^{20}3^2)}}.$$

Third, suppose that $s \leq 2s_{\text{eq}}/3$. From Lemma 2.6.2.(i) and Lemma 2.6.4.(ii), for the chain rule and the union bound, there is a time

$$T_3 = \mathcal{O}\left(\frac{\log n}{\log\left(1 + \frac{3\varepsilon^2}{4}\right)}\right) = \mathcal{O}\left(\log n/\varepsilon^2\right)$$

such that

$$\Pr\left[2s_{\text{eq}}/3 \leq s_{t+T_3} \leq s_{\text{eq}} \mid s_t = s\right] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^4/2^6}}\,.$$

Then, we are in one of the first two cases, and we conclude for the chain rule.

Fourth, suppose that $s \geq (1+\frac{1}{3})s_{\text{eq}}$. From Lemma 2.6.2.(ii) and Lemma 2.6.4.(i), for the chain rule, a time T_4 exists, with $T_4 = \mathcal{O}(\log n)$, such that

$$\Pr\left[\left|s_{\text{eq}} - s_{T_4}\right| \leq s_{\text{eq}}/3 \mid s_t = s\right] \geq 1 - \frac{1}{n^{\gamma^2 3^4/2^6}}\,.$$

The theorem follows with $\tau_1 = \mathcal{O}(T_2 + T_3 + T_4)$ □

2.6.2 3-MAJORITY dynamics: symmetry-breaking

The aim of this section is to prove Theorem 2.3.5: so, in each statement we assume that $\{s_t\}_{t\geq 0}$ is the process induced by the 3-MAJORITY dynamics with uniform noise probability $p < 1/3$. The symmetry breaking analysis relies on the same lemma we exploited for the UNDECIDED-STATE dynamics, Lemma 2.4.6 in Section 2.4.4, and the proof follows the same outline.

Again, the goal is to apply the above lemma to the 3-MAJORITY process, which defines a Markov chain. In particular, we claim the hypothesis of Lemma 2.4.6 are satisfied when the bias of the system is $o\left(\sqrt{n\log n}\right)$, with $f(M_t) = s(M_t)$, $m = \gamma\sqrt{n}\log n$ for any constant $\gamma > 0$. Then, Lemma 2.4.6 implies the process reaches a configuration with bias greater than $\Omega\left(\sqrt{n\log n}\right)$ within time $\mathcal{O}\left(\log n\right)$, w.h.p. We need to prove that the two hypotheses hold.

Lemma 2.6.6. *For any constant $c_3 > 0$, let s be a value such that $|s| < c_3\sqrt{n}\log n$. Then,*

(i) for any positive constant $h > 0$, there exists a positive constant $c_1 < 1$ (which depends only on h), such that

$$Pr\left[s_t < h\sqrt{n} \mid s_{t-1} = s\right] < c_1;$$

(ii) two positive constants δ, c_2 exist (depending only on p), such that if $|s| \geq h\sqrt{n}$, then

$$Pr\left[s_t < (1+\delta)s \mid s_{t-1} = s\right] < e^{-\frac{c_2 s^2}{n}}.$$

Proof. As for the first claim, a simple domination argument implies that

$$\Pr\left[|s_t| < h\sqrt{n} \mid s_{t-1} = s\right] \leq \Pr\left[|s_t| < h\sqrt{n} \mid s_{t-1} = 0\right]. \tag{2.18}$$

Thus, we can bound just the second probability, where the initial bias is zero. As shown in **??**, s_t is a sum of n i.i.d. Rademacher r.v.s with zero mean and unitary variance. We can hence make use of the Lemma B.6 (Berry-Esseen inequality). In particular, let $\Phi(x)$ be the cumulative function of a standard normal distribution. A constant $C > 0$ exists such that

$$\left|\Pr\left[s_t \leq h\sqrt{n} \mid s_{t-1} = 0\right] - \Phi(h)\right| \leq \frac{C}{\sqrt{n}}.$$

Since $\Phi(h) = c$ for some constant $c > 0$ which depends only on h, we have that

$$c - \frac{C}{\sqrt{n}} \leq \Pr\left[s_t \leq h\sqrt{n} \mid s_{t-1} = 0\right] \leq c + \frac{C}{\sqrt{n}}.$$

Since $\Pr\left[|s_t| < h\sqrt{n} \mid s_{t-1} = 0\right] \leq \Pr\left[s_t \leq h\sqrt{n} \mid s_{t-1} = 0\right]$, for n large enough we get

$$\Pr\left[|s_t| < h\sqrt{n} \mid s_{t-1} = 0\right] < 2c.$$

By setting $c_1 = c/2$ and from Eq. (2.18) we get claim (i).

As for the second claim, assume $s > 0$ and $h\sqrt{n} \leq s \leq h\sqrt{n}\log n$. By Lemma 2.6.1 and the fact that $h\sqrt{n} \leq s \leq h\sqrt{n}\log n \leq (1-\sqrt{\varepsilon})s_{eq}$, we have (as in Lemma 2.6.2)

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] = \frac{s(1-2p)}{2}\left(3 - \frac{s^2}{n^2}(1-2p)^2\right) \geq s\left(\frac{3}{2} - 3p - \frac{1-6p-2\varepsilon}{2}\right) = s(1+\varepsilon).$$

From the Hoeffding bound (Lemma B.3), we get that

$$\Pr\left[s_t \leq s\left(1+\varepsilon\right) - s\varepsilon/4 \mid s_{t-1} = s\right] \leq e^{-s^2\varepsilon^2/(32n)}.$$

Observe that $\Pr\left[|s_t| \leq s\left(1+3\varepsilon/4\right) \mid s_{t-1} = s\right] \leq \Pr\left[s_t \leq s\left(1+3\varepsilon/4\right) \mid s_{t-1} = s\right]$. Thus, we have the claim by setting $\delta = 3\varepsilon/4$ and $c_2 = \varepsilon^2/32$. □

The symmetry breaking is then a simple consequence of the above Lemma.

Proof of Theorem 2.3.5. Apply Lemmas 2.6.6 and 2.4.6 with $h = c_3 = \gamma$. □

2.6.3 3-MAJORITY dynamics: victory of noise

In this subsection, we prove Theorem 2.3.6: so, in each statement, we assume that $\{s_t\}_{t\geq 0}$ is the process induced by the 3-MAJORITY dynamics with uniform noise probability $p > 1/3$.

We make use of tools from drift analysis (Lemma D.1) to the absolute value of the bias of the process, showing that it reaches magnitude $\mathcal{O}(\sqrt{n})$ quickly. Then, since the standard deviation of the bias is $\Theta(\sqrt{n})$, we have constant probability that the majority opinion switches Lemma 2.6.9. Finally, with Lemma 2.6.10, we show that the bias keeps bounded in absolute value by $\mathcal{O}(\sqrt{n \log n})$.

Lemma 2.6.7. *For any constant $\varepsilon > 0$ such that $\varepsilon < (1-p)$, if $s \geq 2\sqrt{n}/(\varepsilon^2)$, the following holds*

$$\mathbb{E}\left[|s_t| \mid s_{t-1} = s\right] \leq \mathbb{E}\left[s_t \mid s_{t-1} = s\right] \cdot \left(1 + \frac{\varepsilon}{2}\right).$$

Proof. Trivially, it holds that

$$|s_t| \leq |s_t - \mathbb{E}\left[s_t \mid s_{t-1} = s\right]| + |\mathbb{E}\left[s_t \mid s_{t-1} = s\right]|.$$

Furthermore, from Lemma 2.6.1, we have that $\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \geq 0$ as long as $s \geq 0$. By writing

$$|s_t - \mathbb{E}\left[s_t \mid s_{t-1} = s\right]| = \sqrt{\left(s_t - \mathbb{E}\left[s_t \mid s_{t-1} = s\right]\right)^2},$$

and by using the Jensen's inequality for a concave function (i.e., the square root), it follows that

$$\begin{aligned}\mathbb{E}\left[|s_t| \mid s_{t-1} = s\right] &\leq \sqrt{\mathbb{E}\left[\left(s_t - \mathbb{E}\left[s_t \mid s_{t-1} = s\right]\right)^2 \,\middle|\, s_{t-1} = s\right]} + \mathbb{E}\left[s_t \mid s_{t-1} = s\right] \\ &= \sigma\left(|s_t| \mid s_{t-1} = s\right) + \mathbb{E}\left[s_t \mid s_{t-1} = s\right], \end{aligned} \tag{2.19}$$

where $\sigma(x)$ represents the standard deviation of a r.v. x. As pointed out in the preliminaries (**??**), the bias can be written as the sum of i.i.d. random variables $Y_i^{(t)}$ taking values in $\{-1, +1\}$. For such sum of variables, the variance is linear:

$$\sigma\left(|s_t| \mid s_{t-1} = s\right)^2 = \sum_{i=1}^{n} \sigma\left(Y_i^{(t)} \,\middle|\, s_{t-1} = s\right)^2 \leq n,$$

where the latter inequality holds since $\sigma\left(Y_i^{(t)} \,\middle|\, s_{t-1} = s\right)^2 \leq 1$ for every i. Furthermore, from Lemma 2.6.1, we deduce that

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \geq \frac{s(1-p)(3-(1-p)^2)}{2} \geq s(1-p).$$

Since $s \geq \frac{2\sqrt{n}}{\varepsilon^2} \geq \frac{2\sqrt{n}}{\varepsilon(1-p)}$, we get that $\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \geq \frac{2\sqrt{n}}{\varepsilon}$. By using the latter facts in Eq. (2.19), we obtain

$$\begin{aligned}\mathbb{E}\left[|s_t| \mid s_{t-1}\right] &\leq \mathbb{E}\left[s_t \mid s_{t-1} = s\right] \cdot \left(1 + \frac{\sigma\left(|s_t| \mid s_{t-1} = s\right)}{\mathbb{E}\left[s_t \mid s_{t-1} = s\right]}\right) \leq \mathbb{E}\left[s_t \mid s_{t-1} = s\right] \cdot \left(1 + \frac{\sqrt{n}}{\frac{2\sqrt{n}}{\varepsilon}}\right) \\ &\leq \mathbb{E}\left[s_t \mid s_{t-1} = s\right] \cdot \left(1 + \frac{\varepsilon}{2}\right).\end{aligned}$$

□

With next lemma, we show that the absolute value of the process quickly becomes of magnitude $\mathcal{O}(\sqrt{n})$.

Lemma 2.6.8. *For any constant $\varepsilon > 0$ such that $\varepsilon < \min\{(1-p),(3p-1)/2\}$ we define $s_{min} = \sqrt{n}/\varepsilon^2$. Then, for any starting configuration s_0 such that $s_0 \geq s_{min}$, with probability at least $1 - 1/n$ there exists a time $\tau = \mathcal{O}_\varepsilon(\log n)$ such that $|s_\tau| \leq s_{min}$.*

Proof. Let $h(x) = \frac{\varepsilon \cdot x}{2}$ be a function. Let $X_t = |s_t|$ if $s_t \geq s_{\min}$, otherwise $X_t = 0$. We now estimate $\mathbb{E}[X_t - X_{t-1} \mid X_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}]$, where $\mathcal{F}_t$ is the natural filtration of the process X_t. We have that

$$\begin{aligned}
&\mathbb{E}[X_t - X_{t-1} \mid X_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}] = \mathbb{E}[X_t \mid X_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}] - X_{t-1} \\
&\overset{(a)}{\leq} \mathbb{E}[|s_t| \mid s_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}] - s_{t-1} \overset{(b)}{\leq} \mathbb{E}[s_t \mid s_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}] \cdot \left(1 + \frac{\varepsilon}{2}\right) - s_{t-1} \\
&\overset{(c)}{\leq} s_{t-1}(1-\varepsilon)\left(1 + \frac{\varepsilon}{2}\right) - s_{t-1} \leq -\frac{\varepsilon \cdot s_{t-1}}{2},
\end{aligned}$$

where (a) holds because $X_t \leq |s_t|$, (b) holds for Lemma 2.6.7, and (c) holds for Lemma 2.6.1. Thus,

$$\mathbb{E}[X_{t-1} - X_t \mid X_{t-1} \geq s_{\min}, \mathcal{F}_{t-1}] \geq h(X_{t-1}) .$$

Since $h'(x) = \varepsilon/2 > 0$, we can apply Lemma D.1.(iii) (B), which is a drift analysis result for super-martingale-like processes. Let τ be the first time $X_t = 0$ or, equivalently, $|s_t| < s_{\min}$. Then

$$\begin{aligned}
\Pr[\tau > t \mid s_0] &< \exp\left[-\frac{\varepsilon}{2} \cdot \left(t - \frac{2}{\varepsilon} - \int_{s_{\min}}^{s_0} \frac{2}{\varepsilon \cdot y}\, \mathrm{d}y\right)\right] \leq \exp\left[-\frac{\varepsilon}{2} \cdot \left(t - \frac{2}{\varepsilon} - \int_{s_{\min}}^{n} \frac{2}{\varepsilon \cdot y}\, \mathrm{d}y\right)\right] \\
&= \exp\left[-\frac{\varepsilon}{2} \cdot \left(t - \frac{2}{\varepsilon} - \frac{2}{\varepsilon}(\log n - \log s_{\min})\right)\right] \\
&= \exp\left[-\frac{\varepsilon}{2} \cdot \left(t - \frac{2}{\varepsilon} - \frac{2}{\varepsilon}((\log n)/2 + 2\log \varepsilon)\right)\right] \\
&\leq \exp\left[-\frac{\varepsilon \cdot t}{2} + 1 + \frac{\log n}{2}\right] .
\end{aligned}$$

If $t = 4(\log n)/\varepsilon$, then we get that $\Pr[\tau > t \mid s_0] < e^{-3(\log n)/2+1} < 1/n$. □

Next lemma states that, whenever the absolute value of the bias is of order of $\mathcal{O}(\sqrt{n})$, then the majority opinion switches at the next round with constant probability.

Lemma 2.6.9. *For any constant $\varepsilon > 0$ such that $\varepsilon < 1/4$, and let s_{t-1} be a configuration such that $|s_{t-1}| = s \leq \sqrt{n}/\varepsilon$. Then, the majority opinion switches at the next round with constant probability.*

Proof. Wlog we assume $s_{t-1} > 0$. Now, $s_{t-1} = b_{t-1} - a_{t-1}$, with $n/2 < b_{t-1} \leq n/2 + \sqrt{n}/(2\varepsilon)$ and $n/2 - \sqrt{n}/(2\varepsilon) \leq a_{t-1} < n/2$. Both b_{t-1} and a_{t-1} can be expressed as the sum of i.i.d. Bernoulli r.v.s. Since $\mathbb{E}[a_t \mid n/2 - \sqrt{n}/(2\varepsilon) \leq a_{t-1} < n/2] \leq n/2$, we have

$$\Pr\left[a_t \geq \frac{n}{2} + \frac{\sqrt{n}}{2\varepsilon} \,\middle|\, s_{t-1} = s\right] = \Pr\left[a_t \geq \frac{n}{2} \cdot \left(1 + \frac{1}{\varepsilon\sqrt{n}}\right) \,\middle|\, s_{t-1} = s\right] \geq e^{-\frac{9}{2\varepsilon^2}},$$

where the latter inequality holds for the reverse Chernoff bound (Lemma C.1), whose hypothesis is satisfied since $\varepsilon < 1/4$. Thus, there is at least constant probability that the majority opinion switches. □

Next lemma shows that the signed bias decreases each round.

Lemma 2.6.10. *For any constant $\varepsilon > 0$ such that $\varepsilon \leq (3p-1)/2$, the following hold*

(i) if $s \geq \frac{\gamma}{2}\sqrt{n \log n}$, then Pr $[s_t \leq (1 - 3\varepsilon/4)s \mid s_{t-1} = s] \geq 1 - \frac{1}{n^{\gamma^2\varepsilon^2/2^7}}$;

(ii) if $s \geq 0$, then Pr $\left[-\frac{\gamma}{2}\sqrt{n \log n} \leq s_t \leq s + \frac{\gamma}{2}\sqrt{n \log n} \mid s_{t-1} = s\right] \geq 1 - \frac{2}{n^{\gamma^2/8}}$.

Proof. From Lemma 2.6.1, for each $s \geq 0$ it holds that

$$\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \leq \frac{3s(1-p)}{2} \leq (1-\varepsilon)s, \tag{2.20}$$

where the second inequality is true since $\varepsilon \leq (3p-1)/2$. We now apply the Hoeffding bound (Lemma B.3) to s_t:

$$\Pr\left[s_t \geq (1-\varepsilon)s + \varepsilon \cdot s/4\right] \leq e^{-s^2\varepsilon^2/(32n)} \leq e^{-\gamma^2\varepsilon^2 \log n/2^7} \leq \frac{1}{n^{\frac{\gamma^2\varepsilon^2}{2^7}}}.$$

As for the second claim, we notice that, from Eq. (2.20), $\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \leq s$. The Hoeffding bound (Lemma B.3) now implies that

$$\Pr\left[s_t \geq s + \frac{\gamma}{2}\sqrt{n \log n}\right] \leq e^{-\gamma^2 \log n/8} \leq \frac{1}{n^{\gamma^2/8}}.$$

Moreover, from Lemma 2.6.1, for any $0 \leq s \leq n$, $\mathbb{E}\left[s_t \mid s_{t-1} = s\right] \geq 0$. Applying again the Hoeffding bound, we get that

$$\Pr\left[s_t \geq -\frac{\gamma}{2}\sqrt{n \log n} \mid s_{t-1} = s\right] \leq e^{-\gamma^2 \log n/8} \leq \frac{1}{n^{\gamma^2/8}},$$

For the union bound, we get the second claim. □

We are ready to prove Theorem 2.3.6.

Proof of Theorem 2.3.6. Claim (i) follows directly from Lemmas 2.6.8 and 2.6.9. As for claim (ii), whenever the bias at some round $t = \tau + k$ becomes $|s_t| \geq (\gamma/2)\sqrt{n \log n}$, from Lemma 2.6.10.(ii) (and its symmetric statement), we have that $|s_t| \leq \gamma\sqrt{n \log n}$ with probability $1 - 2/n^{\frac{\gamma^2}{8}}$. Then, from Lemma 2.6.10.(i) it follows that the bias starts decreasing each round with probability $1 - 1/n^{\gamma^2\varepsilon^2/2^7}$ until reaching $(\gamma/2)\sqrt{\log n}$. This phase in which the absolute value of the bias keeps bounded by $\left|\gamma\sqrt{n \log n}\right|$ lasts for at least $n^{\gamma^2\varepsilon^2/2^8}$ with probability at least $1 - 1/(2n^{\gamma^2\varepsilon^2/2^8})$ for the chain rule. □

2.7 Experiments

In this section, we describe the experiments we conducted on the two dynamics to show, in practice, the behavior of the bias for different input sizes and noise parameters. We also tested our results on different topologies.

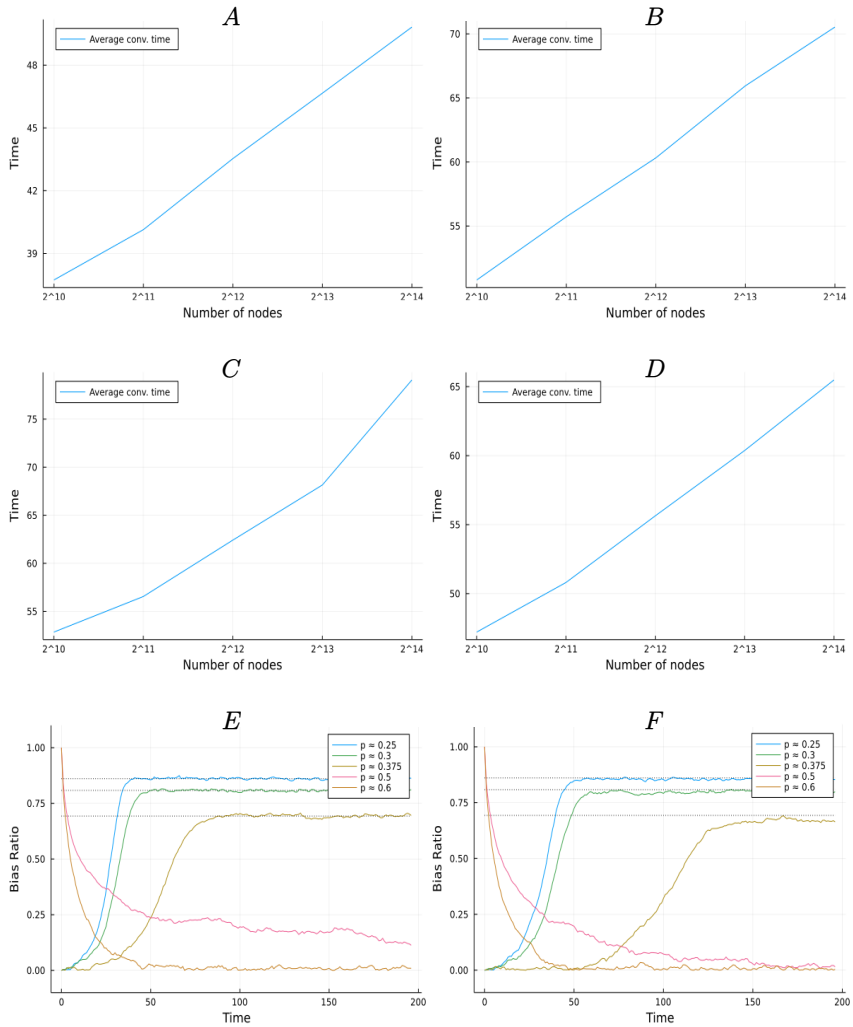


Figure 2.2 – (A) Clique, victory of majority: average convergence time to almost consensus for a clique with noise parameter $p = 1/4$. (B) Clique, victory of noise: average convergence time to a symmetric configuration for a clique with noise parameter $p = 3/5$. (C) Erdös-Rényi graph, victory of majority: same as in (A) but for an Erdös-Rényi graph. (D) Erdös-Rényi graph, victory of noise: same as in (B) but for an Erdös-Rényi graph. (E) Clique, bias behavior: evolution of the bias for different values of the noise parameter $p \in \{1/4, 3/8, 3/7, 1/2, 3/5\}$ in the clique of size 2^{14}. (F) Erdös-Rényi graph, bias behavior: same as in (E) but for an Erdös-Rényi graph of size 2^{14}. In (A), (B), (C), and (D) the x-axis shows the size of the graph in logarithmic scale and the y-axis the average time of convergence over 1000 trials. In (E) and (F), the x-axis shows the time instant of the process, while the y-axis indicates the ratio of the bias, i.e., $|s_t|/n$ for a given configuration M_t. The dotted lines represent the expected equilibrium values for the clique. In all figures, the starting configuration is symmetric and random for $p < 1/2$, and the monochromatic one for $p \geq 1/2$.

2.7.1 UNDECIDED-STATE dynamics: simulations

We simulated the UNDECIDED-STATE dynamics with values of the input size in the range $n \in \{2^{10}, 2^{11}, \ldots, 2^{14}\}$, and for noise probabilities in the range $p \in \{1/4, 3/8, 3/7, 1/2, 3/5\}$. Besides confirming the phase transition predicted by our theoretical analysis, the outcomes show this behavior emerges even for reasonable sizes of the system. Fig. 2.2.(A) shows a convergence time to almost-consensus [4] in the complete graph (equivalently, clique) for $p = 1/4$ which is perfectly approximated by a logarithmic function. The convergence time has been calculated as the average over 1000 trials. In Fig. 2.2.(B), the same process is shown for a "victory of noise" case, i.e., for $p = 3/5$. The behavior of the bias for the different values of p over the clique with 2^{14} nodes is shown in Fig. 2.2.(E). Our intuition suggests that our results carry on for sparser topologies that exhibit good expansion properties. To establish experimentally this, we ran the same simulations on the Erdös-Rényi graph model of size n, with edge probability $\log_2(n)/n$: it is well-known that this setting generate, w.h.p., a connected graph with logarithmic node degree. We briefly recall what an undirected Erdös-Rényi graph is.

Definition 2.7.1 (Undirected Erdös-Rényi graph)**.** Let $n \in \mathbb{N}$ and $p \in (0, 1)$. An undirected graph $G = (V, E)$ is constructed according to the Erdös-Rényi model if $|V| = n$ and for any $x, y \in V$, $\{x, y\} \in E$ with probability p. Such a graph is denoted by $G_{n,p}$.

When $p \sim \log_2(n)/n$, each node has logarithmic degree w.h.p. The results in Fig. 2.2.(C) and Fig. 2.2.(D) seem to show that even such a topology yields logarithmic convergence time; moreover, Fig. 2.2.(F) suggests that the phase-transition threshold either is very close or coincides with the same value $p = 1/2$.

2.7.2 3-MAJORITY dynamics: simulations

We simulated the 3-MAJORITY dynamics with values of the input size in the range $n \in \{2^{10}, 2^{11}, \ldots, 2^{14}\}$, and for noise probabilities in the range $p \in \{1/6, 1/5, 1/4, 1/3, 1/2\}$ as we did for the UNDECIDED-STATE dynamics in Section 2.7.1, The same observations of Section 2.7.1 carry on, as one can observe in Fig. 2.3.

2.8 Discussion and future work

While our mathematical analysis for the UNDECIDED-STATE and the 3-MAJORITY dynamics does not directly apply to other opinion dynamics (even if for some of them, such as the 2-CHOICHES, we believe it does), it might suggest the emergence of a general phase-transition phenomenon for a large class of dynamics characterized by an exponential drift towards consensus configurations. Our works thus naturally pose the general question of whether it is possible to provide a characterization of opinion dynamics with stochastic interactions, in terms of their critical behavior with respect to uniform communication noise.

As for the specific mathematical questions that follow from our results, our assumption of a complete topology as underlying graph is, for several real MAS, a rather strong condition. However, two remarks on this issue follow. On one hand, we observe that, according to the adopted communication model, at every round, every agent can pull information from a constant number of

4. Both for the UNDECIDED-STATE dynamics and the 3-MAJORITY dynamics, we consider as convergence time the first time the bias of the system reaches its expected equilibrium value.

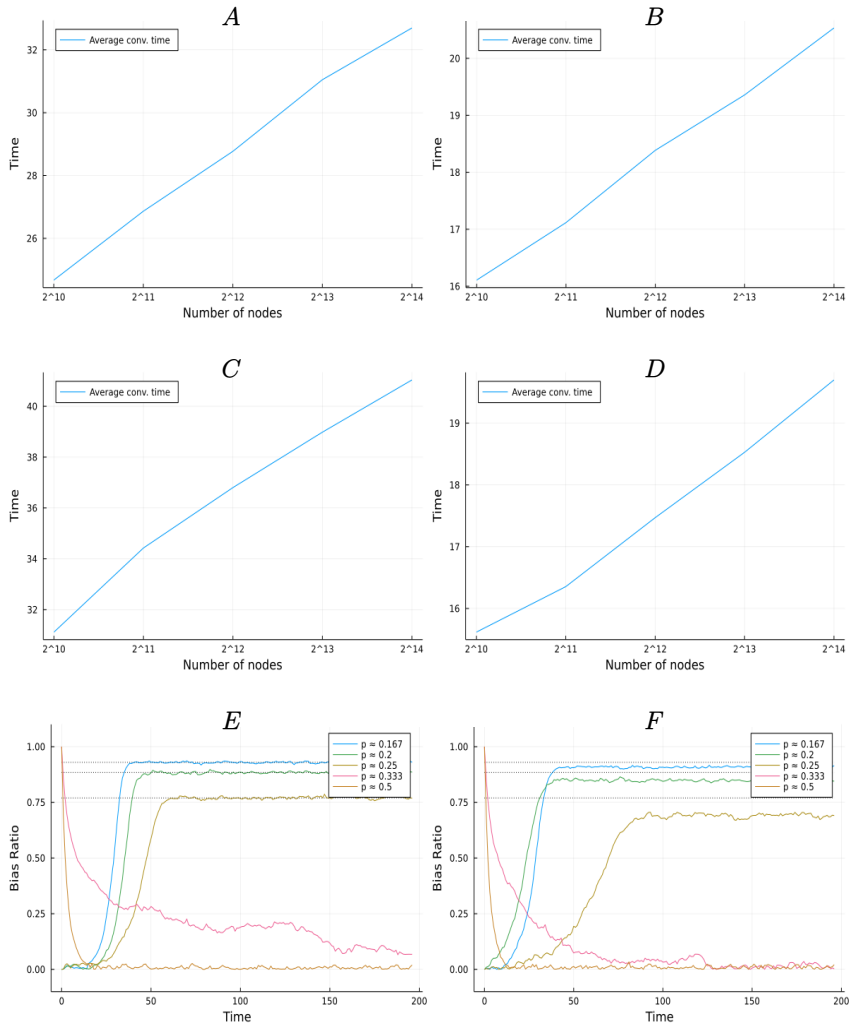


Figure 2.3 – (A) Clique, victory of majority: average convergence time to almost consensus for a clique with noise parameter $p = 1/6$. (B) Clique, victory of noise: average convergence time to a symmetric configuration for a clique with noise parameter $p = 1/2$. (C) Erdös-Rényi graph, victory of majority: same as in (A) but for an Erdös-Rényi graph. (D) Erdös-Rényi graph, victory of noise: same as in (B) but for an Erdös-Rényi graph. (E) Clique, bias behavior: evolution of the bias for different values of the noise parameter $p \in \{1/6, 1/5, 1/4, 1/3, 1/2\}$ in the clique of size 2^{14}. (F) Erdös-Rényi graph, bias behavior: same as in (E) but for an Erdös-Rényi graph of size 2^{14}. In (A), (B), (C), and (D) the x-axis shows the size of the graph in logarithmic scale and the y-axis the average time of convergence over 1000 trials. In (E) and (F), the x-axis shows the time instant of the process, while the y-axis indicates the ratio of the bias, i.e., $|s_t|/n$ for a given configuration M_t. The dotted lines represent the expected equilibrium values for the clique. In all figures, the starting configuration is symmetric and random for $p < 1/3$, and the monochromatic one for $p \geq 1/3$.

other agents: the dynamic communication pattern is thus random and sparse. This setting may model opportunistic MAS where mobile agents use to meet randomly, at a relatively-high rate. On the other hand, we believe that a similar transition phase does hold even for sparse topologies having good expansion/conductance [Hoory et al., 2006]. Our conjecture is supported by the set of experimental tests we performed on the classic Erdös-Rényi model: the results confirm the presence of a threshold behavior of the UNDECIDED-STATE and the 3-MAJORITY dynamics equivalent to those we proved for the complete graph. We instead believe that the behavior of the dynamics over non-expanding graphs cannot be directly exploited to get fast and reliable almost-consensus since it strongly depends on the specific "geometric" shape of the initial configuration: hence, this study requires to introduce further, more refined parameters. Furthermore, it would be interesting to investigate and characterize the same model in presence of many opinions. The phase-transitions we have shown hint at phenomena that are independent from the number of opinions: in the UNDECIDED-STATE dynamics, the threshold $p = 1/2$ corresponds to the setting where half of the round-by-round communication are non-noisy, on average; as for the 3-MAJORITY dynamics, the threshold $p = 1/3$ represents the case where exactly 2 out of 3 communications for each node at each round are non-noisy, on average.

The UNDECIDED-STATE dynamics turns out to have a higher noise threshold than the 3-MAJORITY dynamics: the feature of the UNDECIDED-STATE dynamics of changing opinion in at least two rounds (first become undecided, then copy) may result in a higher resilience to uniform communication disturbances, while in the 3-MAJORITY dynamics nodes change opinion more easily and are, thus, more affected by noise.

Finally, the properties we show for the UNDECIDED-STATE and the 3-MAJORITY dynamics in the case of the complete graphs might also result useful to perform another fundamental task in MAS (which is not the subject of our work): that of distributed community detection [Cruciani et al., 2019, Becchetti et al., 2020b]. Informally, if the MAS has a hidden community structure, the evolution of the processes might have a long metastable phase where the opinion/state of each agent is consistent with the hidden community it belongs to.

CHAPTER 3

Search via Parallel Lévy Walks on $\mathbb{Z}^2$

In this chapter we study the hitting time of parallel Lévy walks over the infinite two-dimensional lattice and employ them to design an optimal algorithm for a distributed search problem.

3.1 Introduction

In this chapter we study the parallel hitting time of Lévy walks in the infinite two-dimensional lattice and we employ the obtained bounds to design an almost-optimal search algorithm for the *Ants Nearby Treasure Search* (ANTS) problem. Informally, a Lévy walk is just a random walk whose step-length density distribution is proportional to a power-law, that is, $f(d) \sim \frac{1}{d^\alpha}$ for any $\alpha > 1$, which moves at constant speed. In the ANTS problem, one is given k agents in $\mathbb{Z}^2$ initially placed at the origin, and a special node, the target, placed by an adversary at Manhattan distance ℓ from the origin; the agents move independently and in parallel, crossing at most one edge each round. The task is to design a search algorithm which finds the target as fast as possible. In [Feinerman and Korman, 2017], a simple lower bound for the problem is shown: in particular, any search algorithm that is oblivious to ℓ requires time $\Omega\left(\ell^2/k + \ell\right)$ both in expectation and with constant probability.

Roadmap. This chapter is organized as follows. In Section 3.2, we give some basic definitions and facts, and investigate a monotonicity property that plays a key role in our analysis. In Section 3.3, we provide the analysis of the regime $\alpha \in (2, 3]$ while in Sections 3.4 and 3.5, we investigate the regimes $\alpha \in (1, 2]$ and $\alpha \in (3, \infty)$, respectively. Finally, in Section 3.6, we use the results from Section 3.3 to analyze the efficiency of our simple distributed search algorithm. We conclude with a discussion on the consequence of our result and future research directions in Section 3.7.

3.2 Preliminaries

3.2.1 Main definitions and properties

For each point $s = (x, y) \in \mathbb{R}^2$, we write $\|s\|_p$ to denote its p-norm $(|x|^p + |y|^p)^{1/p}$. The p-norm distance between points $s = (x, y)$ and $s' = (x', y')$ is $\|s - s'\|_p = (|x - x'|^p + |y - y'|^p)^{1/p}$. We consider the infinite grid graph $G = (\mathbb{Z}^2, E)$, where $E = \{\{u, v\} \colon \|u - v\|_1 = 1\}$. The shortest-path distance between two nodes $u, v \in \mathbb{Z}^2$ in G equals $\|u - v\|_1$. In the following, we will say just *distance* to refer to the shortest-path distance. We denote by $R_d(u)$ the set of all nodes $v \in \mathbb{Z}^2$ that are at distance exactly d from u, i.e., $R_d(u) = \{v \in \mathbb{Z}^2 \colon \|u - v\|_1 = d\}$. We also define $B_d(u) = \{v \in \mathbb{Z}^2 \colon \|u - v\|_1 \leq d\}$ and $Q_d(u) = \{v \in \mathbb{Z}^2 \colon \|u - v\|_\infty \leq d\}$. See Fig. 3.1 for an illustration.

By $\overline{uv}$ we denote the straight-line segment on the real plane $\mathbb{R}^2$ between nodes u and v. A *direct-path* between u and v in G is a shortest path that "closely follows" the real segment $\overline{uv}$. See Fig. 3.2 for an illustration.

Definition 3.2.1 (Direct-path)**.** A *direct-path* from node u to v is a shortest path $u, u_1, \ldots, u_k = v$, where $k = \|u - v\|_1$, and for each $1 \leq i < k$, $u_i \in R_i(u)$ and $\|u_i - w_i\|_2 = \min_{v' \in R_i(u)} \|v' - w_i\|_2$, where w_i is the (unique) point w in the real segment $\overline{uv}$ with $\|u - w\|_1 = i$.

It is not hard to verify that $u, u_1, \ldots, u_k$ is indeed a path of G. Also, unlike point w_i, node u_i is not necessarily unique, since there may be two different nodes in $R_i(u)$ that are closest to w_i. We prove the next simple fact about direct-paths.

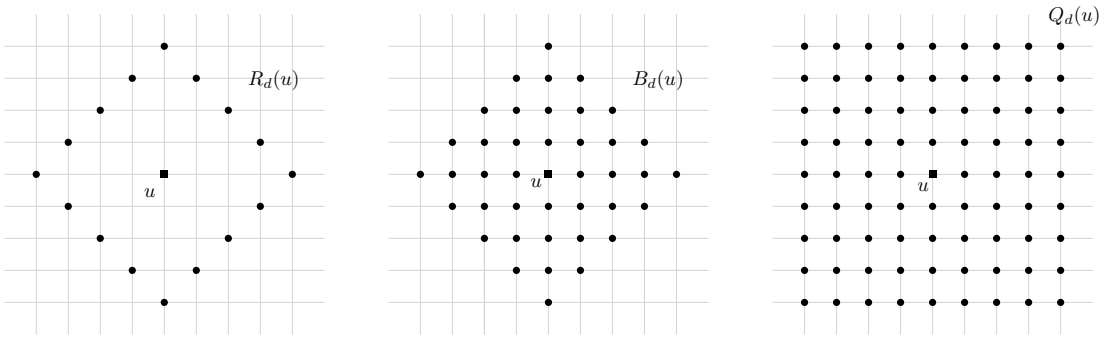


Figure 3.1 – Illustrations of $R_d(u)$, $B_d(u)$, and $Q_d(u)$, for $d = 4$.

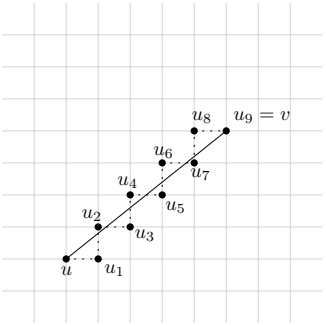


Figure 3.2 – Example of a line segment $\overline{uv}$ and the direct-path between u and v.

Lemma 3.2.1. *Let $u \in \mathbb{Z}^2$ and $d \geq 1$ be an integer. Suppose we sample a node v uniformly at random from set $R_d(u)$, and then sample a direct-path $u, u_1, \ldots, u_d = v$ from u to v uniformly at random among all such paths. Then, for every $1 \leq i < d$ and $w \in R_i(u)$,*

$$\frac{(i/d) \cdot \lfloor d/i \rfloor}{4i} \leq \boldsymbol{Pr}[u_i = w] \leq \frac{(i/d) \cdot \lceil d/i \rceil}{4i}.$$

Proof. For $1 \leq i < d$, consider the real rhombus $\tilde{R}_i(u)$ which is the set $\{v \in \mathbb{R}^2 : \|u - v\|_1 = i\}$. Project each element of $R_d(u)$ on $\tilde{R}_i(u)$ as in Fig. 3.3. We obtain $4d$ equidistant points $\bar{v}_1, \ldots, \bar{v}_{4d}$ in Euclidean distance. Then, for each $w \in R_i(u)$, consider the set $C_w \subseteq \{\bar{v}_1, \ldots, \bar{v}_{4d}\}$ of points that are closest to w than to any other node of $R_i(u)$ in Euclidean distance. Note that each $\bar{v}_j$ belongs to either one or two sets C_w; in the latter case we say that $\bar{v}_j$ is *shared*. For the cardinality of set C_w, we have the following cases: (i) if $d \equiv 0 \pmod{i}$ then $|C_w| = d/i + 1$, and two of the elements of C_w are shared; (ii) if $d \not\equiv 0 \pmod{i}$ then either $|C_w| = \lfloor d/i \rfloor$ and no elements of C_w are shared, or $|C_w| = \lceil d/i \rceil$ and at most one element of C_w is shared.

Choose a node $v \in R_d(u)$ u.a.r. and look at a node $w \in R_i(u)$. The probability that w is on a direct-path chosen u.a.r. is exactly the probability that the projection $\bar{v}$ of v on $\tilde{R}_i(u)$ belongs to C_w, where shared points contribute by $1/2$. In all cases, this probability is between $\lfloor d/i \rfloor / (4d)$ and $\lceil d/i \rceil / (4d)$. □

A (discrete-time) *jump process* on $\mathbb{Z}^2$ is just an infinite sequence of random variables $(J_t)_{t \geq 0}$, where $J_t \in \mathbb{Z}^2$ for each integer $t \geq 0$. We say that the process *visits* node $v \in \mathbb{Z}^2$ at step $t \geq 0$ if $J_t = u$. Our analysis will focus on the following two jump processes.

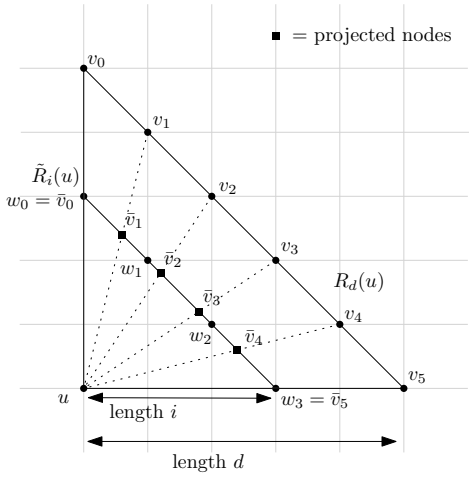


Figure 3.3 – Projection from $R_d(u)$ to $\tilde{R}_i(u)$, with $d = 5$, $i = 3$.

Definition 3.2.2 (Lévy flight)**.** A jump process $L^f = (L^f_t)_{t\geq 0}$ on $\mathbb{Z}^2$ is a Lévy flight with exponent parameter $\alpha \in (1, \infty)$, and start node $s \in \mathbb{Z}^2$, if $L^f_0 = s$, and for each $t \geq 0$, if $L^f_t = u$ then $L^f_{t+1} = v$, where node $v \in \mathbb{Z}^2$ is selected as follows: First a jump distance d is chosen independently at random such that

$$\Pr[d = 0] = 1/2, \text{ and } \Pr[d = i] = c_\alpha / i^\alpha \text{ for } i \geq 1, \tag{3.1}$$

where c_α is a normalizing factor. Then, node v is chosen independently and uniformly at random among all nodes in $R_d(u)$ (i.e., all nodes at distance d from u).

Definition 3.2.3 (Lévy walk)**.** A jump process $L^w = (L^w_t)_{t\geq 0}$ on $\mathbb{Z}^2$ is a Lévy walk with exponent parameter $\alpha \in (1, \infty)$, and start node $s \in \mathbb{Z}^2$, if $L^w_0 = s$, and the process consists of a infinite sequence of *jump-phases*, where each jump-phase is defined as follows: Suppose that the jump-phase begins at step $t + 1$ (the first jump-phase begins at the first step), and suppose also that $L^w_t = u$. First a distance d and a node v at distance d from u are chosen, in exactly the same way as in the Lévy flight. If $d = 0$ then the jump-phase has length 1, and $L^w_{t+1} = u$, i.e., the process stays put for one step. If $d \neq 0$ then the jump-phase has length d, and in the next d steps the process follows a path $u, u_1, \ldots, u_d = v$ chosen uniformly at random among all direct-paths from u to v, i.e., $L^w_{t+i} = u_i$, for all $1 \leq i \leq d$. [1]

We observe that a Lévy flight is a Markov chain, while a Lévy walk is not.

Remark 3.2.1 – Throughout the analysis, we assume α to be a (not necessarily constant) real value greater than $1 + \epsilon$ for some arbitrarily small constant $\epsilon > 0$.

We will often use the following bound on the tail distribution of the jump length d chosen according to (3.1):

$$\Pr[d \geq i] = \Theta\left(1/i^{\alpha-1}\right) \tag{3.2}$$

The next statement follows immediately from Lemma 3.2.1.

1. Our analysis works also if an arbitrary direct-path between u and v is selected, instead of a random one.

Corollary 3.2.2. *Let $u, v \in \mathbb{Z}^2$, and $d = \|u-v\|_1 > 0$. If a Lévy walk is at node u at the beginning of a jump-phase, then the probability it visits v during the jump-phase is $\Theta\left(1/d^{\alpha}\right)$.*

Definition 3.2.4 (Hitting Time)**.** The *hitting time* for node $u^* \in \mathbb{Z}^2$ of a jump process is the first step $t \geq 0$ when the process visits u. For a set of k independent jump processes that run in parallel, their *parallel hitting time* for u^* is the first step in which some (at least one) of the k processes visits u^*.

We will denote by $\tau_\alpha(u^*)$ the hitting time for u^* of a single Lévy walk processes with exponent α starting from the origin $\mathbf{0} = (0,0)$; and by $\tau_\alpha^k(u^*)$ the parallel hitting time for u^* of k independent copies of the above Lévy walk. Unless stated otherwise, we will always assume that the starting node of a jump processes is the origin $\mathbf{0} = (0,0)$.

For a Lévy flight $L^f = (L_t^f)_{t\geq 0}$, we denote by $Z_u^f(t)$ the number of times the process visits node $u \in \mathbb{Z}^2$ until step t, i.e., $Z_u^f(t) = |\{i\colon L_i^f = u\} \cap \{1, \ldots, t\}|$. We define $Z_u^w(t)$ similarly for a Lévy walk.

3.2.2 Bounds via monotonicity

We will now describe an intuitive monotonicity property that applies to a family of jump processes that includes Lévy flights (but not Lévy walks). We then use this property, and the similarity between Lévy flights and walks, to show upper bounds on the probability that a Lévy walk visits a given target. We start by defining the family of *monotone radial* jump processes.

Definition 3.2.5 (Monotone radial process)**.** A jump process $(J_t)_{t\geq 0}$ is *monotone radial* if, for any pair of nodes $u, v \in \mathbb{Z}^2$, and any $t \geq 0$, $\Pr\left[J_{t+1} = v \mid J_t = u\right] = \rho(\|u - v\|_1)$, for some non-increasing function ρ.

Clearly, Lévy flights are monotone radial processes. For all such processes, we use geometric arguments to prove the following property. The proof is deferred to Appendix F.1.

Lemma 3.2.3 (Monotonicity property)**.** *Let $(J_i)_{i\geq 0}$ be any monotone radial jump process. For every pair $u, v \in \mathbb{Z}^2$ and any $t \geq 0$, if $\|v\|_\infty \geq \|u\|_1$ then $\Pr\left[J_t = u\right] \geq \Pr\left[J_t = v\right]$.*

Next, we use Lemma 3.2.3 to upper bound the probability that a target node is visited during a given jump-phase of a Lévy walk, and then bound the probability that the target is visited during *at least one* jump-phase.

Lemma 3.2.4. *Let u^* be an arbitrary node with $\ell = \|u^*\|_1$. Let $\epsilon > 0$ be an arbitrarily small constant. For any $\alpha \geq 1 + \epsilon$, the probability that a Lévy walk visits u^* during its i-th jump-phase is $\mathcal{O}\left(\mu \cdot (\ell^{-2} + \ell^{-\alpha})\right)$ if $\alpha \neq 2$, where $\mu = \min\{\log \ell, \left|\frac{1}{2-\alpha}\right|\}$, and $\mathcal{O}\left(\log \ell/\ell^2\right)$ if $\alpha = 2$.*

Proof. For any $v \in B_{\ell/4}(u^*)$, the probability that the i-th jump starts in v is at most $\mathcal{O}(1/\ell^2)$ due to Lemma 3.2.3, since the process restricted only to the jumps endpoints is a Lévy flight. Moreover, for any $1 \leq d \leq \ell/4$, there are at most $4d$ nodes in $B_{\ell/4}(u^*)$ located at distance d from u^*. Then, from the chain rule and Corollary 3.2.2, the probability that the i-th jump starts from $B_{\ell/4}(u^*)$ and the agent visits the target during the jump-phase is bounded by

$$\mathcal{O}\left(\frac{1}{\ell^2}\right) \sum_{d=1}^{\ell/4} 4d \cdot \mathcal{O}\left(\frac{1}{d^\alpha}\right) + \mathcal{O}\left(\frac{1}{\ell^2}\right),$$

the term $\mathcal{O}(1/\ell^2)$ being the contribution of u^* itself. The above expression equals $\mathcal{O}\left(\mu\ell^{2-\alpha}/\ell^2\right)$ if $\alpha \neq 2$, and $\mathcal{O}\left(\log \ell/\ell^2\right)$ if $\alpha = 2$. For any fixed node v, denote by F_i event that that, during a jump-phase starting in v, the agent visits the target. We now prove that $\Pr[F_i] = \mathcal{O}\left(\mu/\ell^\alpha\right)$.

Let V_i be the event that the starting point of the i-th jump is in $B_{\ell/4}(u^*)$. Notice that $\Pr\left[F_i \mid \overline{V_i}\right]$ is at most $\mathcal{O}\left(1/\ell^\alpha\right)$ for Corollary 3.2.2. Then,

$$\Pr[F_i] \leq \Pr[F_i \mid V_i]\Pr[V_i] + \Pr\left[F_i \mid \overline{V_i}\right] \leq \Pr[F_1 \mid V_i] + \mathcal{O}\left(\frac{1}{\ell^\alpha}\right).$$

Then, if $\alpha > 2$,

$$\Pr[F_i] = \mathcal{O}\left(\frac{\mu}{\ell^2} + \frac{1}{\ell^\alpha}\right) = \mathcal{O}\left(\frac{\mu}{\ell^2}\right).$$

If $\alpha = 2$, $\Pr[F_i] = \mathcal{O}\left(\log \ell/\ell^2\right)$. And if $1 < \alpha < 2$,

$$\Pr[F_i] = \mathcal{O}\left(\frac{\mu}{\ell^\alpha} + \frac{1}{\ell^\alpha}\right) = \mathcal{O}\left(\frac{\mu}{\ell^\alpha}\right).$$

□

Lemma 3.2.5. *Let u^* be an arbitrary node with $\|u^*\|_1 = \ell$. Let $\epsilon > 0$ be an arbitrarily small constant. The probability that a Lévy walk with exponent $1+\epsilon \leq \alpha < 3$ visits u^* at least once (at any step t) is $\mathcal{O}\left(\mu \log \ell(\ell^{-1} + \ell^{3-\alpha})\right)$ if $\alpha \neq 2$, where $\mu = \min\{\log \ell, \left|\frac{1}{2-\alpha}\right|\}$, and $\mathcal{O}\left(\frac{\log^2 \ell}{\ell}\right)$ if $\alpha = 2$.*

Proof. For each $i \geq 0$, consider the first time t_i the agent is at distance at least $\lambda_i = 2^i\ell$ from the origin. From Eq. (3.2), the probability any jump has length no less than $2\lambda_i$ is at least $c/\lambda_i^{\alpha-1}$, for some constant $c > 0$. Define, for $i \geq 1$, the values $\tau_i = 2c^{-1}\lambda_i^{\alpha-1}\log\lambda_i$. The probability that $t_i \geq n\tau_i$ is bounded from above by the probability that no jump between the first $n\tau_i$ jumps has length at least $2\lambda_i$. Since the jump lengths are mutually independent, from Eq. (3.2) we get

$$\Pr[t_i \geq n\tau_i] \leq \left[1 - \frac{c}{\lambda_i^{\alpha-1}}\right]^{\frac{2}{c}n\lambda_i^{\alpha-1}\log\lambda_i} \leq \exp\left(-2n\log\lambda_i\right) = \frac{1}{\lambda_i^{2n}} = \frac{1}{2^{2ni}\ell^{2n}},$$

where we have used the well known inequality $1 - x \leq e^{-x}$ that holds for every real x. We next bound the expected number of visits to the target node from time t_i to time t_{i+1} as follows:

$$\begin{aligned}\mathbb{E}\left[Z^w_{u^*}(t_{i+1}) - Z^w_{u^*}(t_i)\right] \leq{}& \mathbb{E}\left[Z^w_{u^*}(t_{i+1}) - Z^w_{u^*}(t_i) \mid t_{i+1} \leq \tau_{i+1}\right]\Pr[t_{i+1} \leq \tau_{i+1}] \\ &+ \sum_{n\geq 1}\mathbb{E}\left[Z^w_{u^*}(t_{i+1}) - Z^w_{u^*}(t_i) \mid n\tau_{i+1} < t_{i+1} \leq (n+1)\tau_{i+1}\right]\cdot\Pr[t_{i+1} \geq n\tau_i].\end{aligned} \tag{3.3}$$

We now proceed by analysing three different ranges for the exponent α and use Lemma 3.2.4. Notice that the agent starts at distance $\Omega(\lambda_i)$ from the target.

If $\alpha > 2$, the expression in Eq. (3.3) is equal to

$$\mathcal{O}\left(\frac{\mu\tau_{i+1}}{\lambda_i^2}\right) + \sum_{n\geq 1}\mathcal{O}\left(\frac{\mu(n+1)\tau_{i+1}}{2^{2ni}\ell^{2n}\lambda_i^2}\right) = \mathcal{O}\left(\frac{\mu\tau_i}{\lambda_i^2}\right)$$

since the sum $\sum_{n\geq 1}(n+1)\left(2^i\ell\right)^{-2n}$ is less than a constant. If $\alpha = 2$, the expression in Eq. (3.3) is

$$\mathcal{O}\left(\frac{\tau_{i+1}\log\lambda_i}{\lambda_i^2}\right) + \sum_{n\geq 1}\mathcal{O}\left(\frac{(n+1)\tau_{i+1}\log\lambda_i}{2^{2ni}\ell^{2n}\lambda_i^2}\right) = \mathcal{O}\left(\frac{\tau_i\log\lambda_i}{\lambda_i^2}\right).$$

And if $1 < \alpha < 2$, the expression is

$$\mathcal{O}\left(\frac{\mu\tau_{i+1}}{(2-\alpha)\lambda_i^\alpha}\right) + \sum_{n\geq 1}\mathcal{O}\left(\frac{\mu(n+1)\tau_{i+1}}{2^{2ni}\ell^{2n}\lambda_i^\alpha}\right) = \mathcal{O}\left(\frac{\mu\tau_i}{\lambda_i^\alpha}\right).$$

The same bounds with $i = 1$ hold for the expected number of visits to the target until time t_1. From the facts above, for $\alpha > 2$, the expected total number of visits to the target is bounded by

$$\mathcal{O}\left(\frac{\mu\tau_1}{\lambda_1^2}\right) + \sum_{i\geq 1}\mathcal{O}\left(\frac{\mu\tau_i}{\lambda_i^2}\right) = \mathcal{O}\left(\frac{\mu\log\ell}{\ell^{3-\alpha}}\right) + \sum_{i\geq 1}\mathcal{O}\left(\mu\cdot\frac{\log(2^i)+\log\ell}{(2^{2i(3-\alpha)}\ell^{3-\alpha})}\right) = \mathcal{O}\left(\frac{\mu\log\ell}{\ell^{3-\alpha}}\right).$$

Similarly, for $\alpha = 2$, we obtain the bound $\mathcal{O}\left(\log^2\ell/\ell^{3-\alpha}\right)$, while, for $1 < \alpha < 2$, we get

$$\mathcal{O}\left(\frac{\mu\log\ell}{\ell}\right) + \sum_{i\geq 1}\mathcal{O}\left(\mu\cdot\frac{\log(2^i)+\log\ell}{(2^{2i(3-\alpha)}\ell)}\right) = \mathcal{O}\left(\frac{\mu\log\ell}{\ell}\right).$$

Finally, we bound the probability the agent visits the target at least once using the Markov's inequality. □

3.3 The case $\alpha \in (2, 3]$

In this section, we analyze the hitting time of Lévy walks when the exponent parameter α belongs to the range $(2, 3]$. In this case, the jump length has bounded mean and unbounded variance.

Recall that $\tau_\alpha(u^*)$ is the hitting time for target u^* of a single Lévy walk with exponent α, and $\tau_\alpha^k(u^*)$ is the parallel hitting time for u^* of k independent copies of the above Lévy walk. All walks start from the origin.

We will prove the following bounds on the hitting time τ_α.

Theorem 3.3.1. *Let $\alpha \in (2,3)$, $u^* \in \mathbb{Z}^2$, and $\ell = \|u^*\|_1$. Let $\mu = \min\{\log\ell, \frac{1}{\alpha-2}\}$, $\nu = \min\{\log\ell, \frac{1}{3-\alpha}\}$, and $\gamma = \frac{(\log\ell)^{\frac{2}{\alpha-1}}}{(3-\alpha)^2}$. Then:*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\mu\cdot\ell^{\alpha-1}\right)\right] = \Omega\left(1/(\gamma\cdot\ell^{3-\alpha})\right)$, if $3-\alpha = \omega(1/\log\ell)$;

(b) $Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(\mu\nu\cdot t^2/\ell^{\alpha+1}\right)$, for any step $\ell \leq t = \mathcal{O}\left(\ell^{\alpha-1}/\nu\right)$;

(c) $Pr\left[\tau_\alpha(u^*) < \infty\right] = \mathcal{O}\left(\mu\cdot\log\ell/\ell^{3-\alpha}\right)$.

Using Theorem 3.3.1, we can easily obtain the following bounds on the parallel hitting time τ_α^k. The proof is given in Section 3.3.5.

Corollary 3.3.2. *Let $u^* \in \mathbb{Z}^2$ and $\ell = \|u^*\|_1$, and let k be any integer such that $\log^6\ell \leq k \leq \ell\log^4\ell$. Let $\alpha^* = 3 - \frac{\log k}{\log\ell}$, and $\overline{\alpha} = \max\{2, \alpha^* - 4\frac{\log\log\ell}{\log\ell}\}$. Then:*

(a) For $\alpha = \alpha^* + 5\frac{\log\log\ell}{\log\ell}$, *Pr* $\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\ell^2\log^6\ell/k\right)\right] = 1 - e^{-\omega(\log\ell)}$;

(b) For $\overline{\alpha} < \alpha < 3$, *Pr* $\left[\tau_\alpha^k(u^*) \le (\ell^2/k)\cdot\ell^{(\alpha-\alpha^*)/2}/\log^4\ell\right] = o(1)$;

(c) For $2 < \alpha \le \alpha^*$, *Pr* $\left[\tau_\alpha^k(u^*) < \infty\right] = \mathcal{O}\left(\log^2\ell/\ell^{\alpha^*-\alpha}\right)$.

Corollary 3.3.2Part a states that $\tau_\alpha^k(u^*) \le (\ell^2/k)\cdot\operatorname{polylog}\ell$, w.h.p., when $\alpha \approx \alpha^* = 3 - \log k/\log\ell$. One the other hand, Corollary 3.3.2Part b says that the lower bound $\tau_\alpha^k(u^*) \ge (\ell^2/k)\cdot\ell^{(\alpha-\alpha^*)/2}/\operatorname{polylog}\ell$ holds with probability $1 - o(1)$ for any $\alpha \ge \alpha^* - \mathcal{O}(\log\log\ell/\log\ell)$, and Corollary 3.3.2Part c says that the target is never hit with probability at least $1 - \operatorname{polylog}\ell/\ell^{\alpha^*-\alpha}$, if $\alpha \le \alpha^*$. Therefore, the optimal hitting time of ℓ^2/k is achieved (modulo $\operatorname{polylog}\ell$ factors) only for values of α very close to α^*, precisely only if $|\alpha - \alpha^*| = \mathcal{O}(\log\log\ell/\log\ell)$.

For the threshold case $\alpha = 3$, similar bounds to those of Theorem 3.3.1 apply, modulo some $\operatorname{polylog}\ell$ factors, as stated in the next theorem.

Theorem 3.3.3. *Let* $u^* \in \mathbb{Z}^2$ *and* $\ell = \|u^*\|_1$. *Then:*

(a) Pr $[\tau_3(u^*) = \mathcal{O}(\ell^2)] = \Omega\left(1/\log^4\ell\right)$;

(b) Pr $[\tau_3(u^*) \le t] = \mathcal{O}(t^2\log\ell/\ell^4)$, *for any step* t *with* $\ell \le t = \mathcal{O}(\ell^2/\log\ell)$.

The next corollary gives bounds on the parallel hitting time τ_3^k. The proof is similar to that of Corollary 3.3.2.

Corollary 3.3.4. *Let* $u^* \in \mathbb{Z}^*$ *and* $\ell = \|u^*\|_1 = \ell$, *and let* k *be any integer such that* $\omega\left(\log^5\ell\right) \le k \le \ell^2/\log^2\ell$. *Then:*

(a) Pr $\left[\tau_3^k(u^*) = \mathcal{O}(\ell^2)\right] = 1 - e^{-\omega(\log\ell)}$.

(b) Pr $\left[\tau_3^k(u^*) \le \ell^2/\sqrt{k}\right] = o(1)$.

Corollary 3.3.4.a says that $\tau_3^k(u^*) = \mathcal{O}(\ell^2)$, w.h.p., for any $k \ge \operatorname{polylog}\ell$, and Corollary 3.3.4.b provides a very crude lower bound indicating that increasing k beyond $\operatorname{polylog}\ell$, can only result in sublinear improvement.

3.3.1 Proof of Theorems 3.3.1 and 3.3.3

We will use the following key lemma, which is shown in Section 3.3.2. The lemma provides an upper bound on the hitting time of a Lévy flight, assuming the maximum jump length is capped to some appropriate value.

Lemma 3.3.5 (Lévy flight with $\alpha \in (2,3]$)**.** *Let* h_f *be the hitting time of a Lévy flight for target* u^* *with* $\|u^*\|_1 = \ell$. *Let* $\mathcal{E}_t$ *be the event that each of the first* t *jumps has length less than* $(t\log t)^{1/(\alpha-1)}$. *Then, there is a* $t = \Theta\left(\ell^{\alpha-1}\right)$ *such that:*

(a) Pr $[h_f \le t \mid \mathcal{E}_t] = \Omega\left(1/(\gamma\ell^{3-\alpha})\right)$, *if* $2 < \alpha \le 3 - \omega(1/\log\ell)$, *where* $\gamma = \frac{(\log\ell)^{\frac{2}{\alpha-1}}}{(3-\alpha)^2}$;

(b) Pr $[h_f \le t \mid \mathcal{E}_t] = \Omega\left(1/\log^4\ell\right)$, *if* $\alpha = 3$.

The second lemma we need is an upper bound on the hitting time of a Lévy walk in terms of the hitting time of the capped Lévy flight considered above. The proof proceeds by coupling the two processes, and is given in Section 3.3.3.

Lemma 3.3.6. *Let h_f and $\mathcal{E}_t$ be defined as in Lemma 3.3.5, and let $\tau_\alpha(u^*)$ be the hitting time of a Lévy walk with the same exponent α, for the same target u^*. Then, for any $2 < \alpha \leq 3$ and step t, and for $\mu = \min\{\log \ell, \frac{1}{\alpha-2}\}$,*

$$Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\mu t\right)\right] \geq \left(1 - \mathcal{O}\left(1/\log t\right)\right) \cdot \left(Pr\left[h_f \leq t \mid \mathcal{E}_t\right] - e^{-t^{\Theta(1)}}\right).$$

The last lemma we need is the following lower bound on the hitting time of a Lévy walk, proved in Section 3.3.4.

Lemma 3.3.7. *Let $u^* \in \mathbb{Z}^2$ and $\ell = \|u^*\|_1$. For any step $t \geq \ell$,*

(a) $Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(\frac{\nu \mu t^2}{\ell^{\alpha+1}}\right)$ if $\alpha \neq 3$ and $t = \mathcal{O}\left(\ell^{\alpha-1}/\nu\right)$, where $\nu = \min\{\log \ell, \frac{1}{3-\alpha}\}$ and $\mu = \min\{\log \ell, \frac{1}{\alpha-2}\}$;

(b) $Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(\frac{t^2 \log \ell}{\ell^{\alpha+1}}\right)$ if $\alpha = 3$ and $t = \mathcal{O}\left(\ell^2/\log \ell\right)$.

Proof of Theorem 3.3.1. Lemmas 3.3.5 and 3.3.6 imply that, for some $t = \Theta\left(\ell^{\alpha-1}\right)$,

$$\Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\mu t\right)\right] = \Omega\left(\frac{(3-\alpha)^2}{\ell^{3-\alpha} \log^{\frac{2}{\alpha-1}} \ell}\right),$$

obtaining (a). Parts (b) and (c) follow from Lemma 3.3.7 and Lemma 3.2.5, respectively. □

Proof of Theorem 3.3.3. The proof proceeds in exactly the same way as the proof of Theorem 3.3.1, using Lemmas 3.3.5 and 3.3.6 to show Part a, and Lemma 3.2.5 to show Part b. □

3.3.2 Proof of Lemma 3.3.5

We first define some notation. Then we give an overview of the analysis, before we provide the detailed proof.

Recall that $(L_i^f)_{i \geq 0}$ denotes the Lévy flight process, and $Z_u^f(i) = |\{j \colon L_j^f = u\} \cap \{1, \ldots, i\}|$ is the number of visits to node u in the first i steps. Let $t = \Theta\left(\ell^{\alpha-1}\right)$ be a step to be fixed later. For each $i \geq 1$, let S_i be the length of the i-th jump of the Lévy flight, i.e.,

$$S_i = \|L_i^f - L_{i-1}^f\|_1.$$

Define also the events

$$E_i = \{S_i \leq (t \log t)^{1/(\alpha-1)}\},$$

and let $\mathcal{E}_i = \bigcap_{j=1}^{i} E_j$. For each node u and $i \geq 0$, let

$$p_{u,i} = \Pr\left[L_i^f = u \,\middle|\, \mathcal{E}_i\right],$$

and note that $\mathbb{E}\left[Z_u^f(i) \,\middle|\, \mathcal{E}_i\right] = \sum_{j=0}^{i} p_{u,j}$. We partition the set of nodes into disjoint sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ defined as follows:

$$\mathcal{A}_1 = \{v \colon \|v\|_\infty \leq \ell\}$$

$$\mathcal{A}_2 = \begin{cases} \{v \colon \|v\|_1 \leq 2(t \log t)^{1/(\alpha-1)}\} \setminus \mathcal{A}_1 & \text{if } \alpha \in (2,3) \\ \{v \colon \|v\|_1 \leq 2\sqrt{t} \log t\} \setminus \mathcal{A}_1 & \text{if } \alpha = 3 \end{cases}$$

$$\mathcal{A}_3 = \mathbb{Z}^2 \setminus (\mathcal{A}_1 \cup \mathcal{A}_2).$$

3.3.2.1 Proof overview

We discuss just the case of $\alpha \in (2,3)$; the case of $\alpha = 3$ is similar. We assume all probability and expectation quantities below are conditional on the event $\mathcal{E}_t$, and we omit writing this conditioning explicitly.

First, we show a simple upper bound on the mean number of visits to $\mathcal{A}_1$ until step $t = \Theta\left(\ell^{\alpha-1}\right)$, namely,

$$\sum_{v\in\mathcal{A}_1} \mathbb{E}\left[Z_v^f(t)\right] \le ct,$$

for a constant $c < 1$: with constant probability the walk visits a node outside $\mathcal{A}_1$ in the first $t/2$ steps, and after that at most a constant fraction of steps visit nodes in $\mathcal{A}_1$, by symmetry.

To bound the mean number of visits to $\mathcal{A}_2$, we use the monotonicity property from Section 3.2.2 and the fact that $\|v\|_\infty \ge \ell = \|u^*\|_1$ for all $v \in \mathcal{A}_2$, to obtain

$$\sum_{v\in\mathcal{A}_2} \mathbb{E}\left[Z_v^f(t)\right] \le |\mathcal{A}_2| \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right] \le 4(t\log t)^{1/(\alpha-1)} \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right].$$

For the number of visits to $\mathcal{A}_3$ we obtain the following bound using Chebyshev's inequality, for a constant c',

$$\sum_{v\in\mathcal{A}_3} \mathbb{E}\left[Z_v^f(t)\right] \le c't/((3-\alpha)\log t).$$

From the above results, and the fact that the total number of visits to all three sets is t, we get

$$ct + 4(t\log t)^{1/(\alpha-1)} \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right] + c't/[(3-\alpha)\log t] \ge t,$$

which implies

$$\mathbb{E}\left[Z_{u^*}^f(t)\right] = \Omega\left(t^{\frac{\alpha-3}{\alpha-1}} \cdot (\log t)^{-\frac{2}{\alpha-1}}\right)$$

if $3-\alpha = \omega(1/\log t)$. We can express the probability of $h_f \le t$ in terms of the above mean as

$$\begin{aligned}\Pr[h_f \le t] &= \Pr\left[Z_{u^*}^f(t) > 0\right] \\ &= \mathbb{E}\left[Z_{u^*}^f(t)\right] / \mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, Z_{u^*}^f(t) > 0\right].\end{aligned}$$

We have

$$\mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, Z_{u^*}^f(t) > 0\right] \le \mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] + 1.$$

We also compute

$$\mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] = \mathcal{O}(1/(3-\alpha)^2).$$

Combining the last four equations yields

$$\Pr[h_f \le t] = \Omega\left(t^{\frac{\alpha-3}{\alpha-1}} \cdot (\log t)^{-\frac{2}{\alpha-1}} \cdot (3-\alpha)^2\right),$$

and substituting $t = \Theta\left(\ell^{\alpha-1}\right)$ completes the proof.

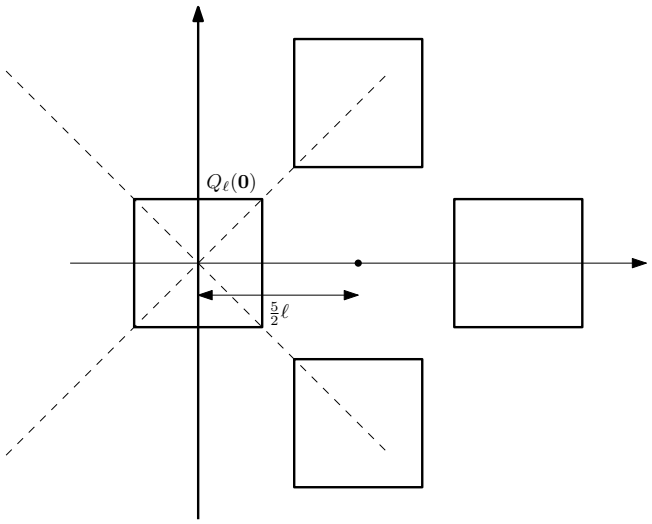


Figure 3.4 – The disjoint zones at least as equally likely as $Q_\ell(\mathbf{0})$ to be visited.

3.3.2.2 Detailed proof

We give now the details of the analysis. Throughout the section $2 < a \leq 3 - \omega(1/\log \ell)$.

Lemma 3.3.8 (Bound on the visits to $\mathcal{A}_1$)**.** *For any $t = \Theta(\ell^{\alpha-1})$ large enough, there is a constant $c \in (0,1)$ such that*

$$\sum_{v \in Q_\ell(\mathbf{0})} \mathbb{E}\left[Z_v^f(t) \,\middle|\, \mathcal{E}_t \right] \leq ct.$$

Proof. We bound the probability the walk has moved to distance $\frac{5}{2}\ell$ at least once, within time $t = \Theta\left(\ell^{\alpha-1}\right)$, by the probability that at least one of the performed jumps is no less than 5ℓ (we denote this latter event by H). Indeed, if there is a jump of length at least 5λ, the walk moves necessarily to distance no less than $\frac{5}{2}\ell$. Then,

$$\Pr\left[S_j \geq 5\ell \mid S_j \leq (t\log t)^{\frac{1}{\alpha-1}} \right] = \sum_{k=5\ell}^{(t\log t)^{\frac{1}{\alpha-1}}} \frac{c_\alpha}{k^\alpha} \overset{(*)}{\geq} \frac{c_\alpha}{\alpha-1}\left(\frac{1}{(5\ell)^{\alpha-1}} - \frac{1}{t\log t} \right) \overset{(\star)}{\geq} \frac{c_\alpha}{2(\alpha-1)(5\ell)^{\alpha-1}},$$

where $(*)$ follows for the integral test (Lemma E.1), while $(\star)$ easily holds for a large enough ℓ since $t = \Theta(\ell^{\alpha-1})$. Thanks to the mutual independence among the random destinations chosen by the agent, the probability of the event "the desired jump takes place within time $c' \cdot 2(\alpha-1)(5\ell)^{\alpha-1}/c_\alpha$" is bounded by

$$1 - \left[1 - \frac{c_\alpha}{2(\alpha-1)(5\ell)^{\alpha-1}} \right]^{c' \frac{2(\alpha-1)(5\ell)^{\alpha-1}}{c_\alpha}} \geq \frac{3}{4},$$

for some constant $c' > 0$ and for ℓ large enough. Hence, by choosing $t \geq 4c' \cdot 2(\alpha-1)(5\ell)^{\alpha-1}/c_\alpha$, the desired jump takes place with probability $\frac{3}{4}$, within time $\frac{t}{4}$. Once reached such a distance (conditional on the previous event), Fig. 3.4 shows there are at least other 3 mutually disjoint regions which are at least as equally likely as $Q_\ell(\mathbf{0})$ to be visited at any future time.

Thus, the probability to visit $Q_\ell(\mathbf{0})$ at any future time step is at most $\frac{1}{4}$. Observe that

$$\begin{aligned}\mathbb{E}\left[\sum_{v\in Q_\ell(\mathbf{0})} Z_v^f(t) \mid \mathcal{E}_t\right] &= \mathbb{E}\left[\sum_{v\in Q_\ell(\mathbf{0})} Z_v^f(t) \mid H, \mathcal{E}_t\right] \Pr[H \mid \mathcal{E}_t] \\ &\quad + \mathbb{E}\left[\sum_{v\in Q_\ell(\mathbf{0})} Z_v^f(t) \mid H^C, \mathcal{E}_t\right] \Pr\left[H^C \mid \mathcal{E}_t\right] \\ &\le \left(\frac{1}{4}t + \frac{1}{4}\cdot\frac{3}{4}t\right)\frac{3}{4} + t\cdot\frac{1}{4} \\ &= \frac{t}{4}\left(1+\frac{3}{4}+\frac{9}{16}\right) = \frac{37}{64}t,\end{aligned}$$

and the proof is completed. □

For the rest of Section 3.3.2, let $t = \Theta\left(\ell^{\alpha-1}\right)$ as in Lemma 3.3.12

Remark 3.3.1 – The monotonicity property (Lemma 3.2.3) holds despite the conditional event $\mathcal{E}_t$. The proof is exactly the same.

Notice that, from $\mathbb{E}\left[Z_v^f(t) \mid \mathcal{E}_t\right] = \sum_{i=0}^t p_{v,i}$ and the monotonicity property, we easily get the following bound.

Corollary 3.3.9. $\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right] \ge \mathbb{E}\left[Z_v^f(t) \mid \mathcal{E}_t\right]$ *for all* $v \notin Q_{d_u}(\mathbf{0})$.

Namely, the more the node is "far" (according to the sequence of squares $\{Q_d(\mathbf{0})\}_{d\in\mathbb{N}}$) from the origin, the less it is visited on average. Thus, each node is visited at most as many times as the origin, on average. This easily gives an upper bound on the total number of visits to $\mathcal{A}_2$ until time t, namely, by taking $u = u^*$ and by observing that each $v \in \mathcal{A}_2$ lies outside $Q_\ell(\mathbf{0})$, we get that the average number of visits to $\mathcal{A}_2$ is at most the expected number of visits to the target u^* (i.e., $\mathbb{E}\left[Z_{u^*}^f(t) \mid \mathcal{E}_t\right]$) times (any upper bound of) the size of $\mathcal{A}_2$: in formula, it is upper bounded by $\mathbb{E}\left[Z_{u^*}^f(t) \mid \mathcal{E}_t\right] \cdot 4(t\log t)^{\frac{2}{\alpha-1}}$ if $\alpha \in (2,3)$, and by $\mathbb{E}\left[Z_{u^*}^f(t) \mid \mathcal{E}_t\right] \cdot 4t\log^2 t$ if $\alpha = 3$.

The next lemma considers $\mathcal{A}_3$.

Lemma 3.3.10 (Bound on visits to $\mathcal{A}_3$)**.** *It holds that*

$$\sum_{\substack{v=(x,y):\\ |x|+|y|\ge 2(t\log t)^{\frac{1}{\alpha-1}}}} \mathbb{E}\left[Z_v^f(t) \mid \mathcal{E}_t\right] = \mathcal{O}\left(\frac{t}{(3-\alpha)\log t}\right) \quad \textit{if } \alpha \in (2,3); \tag{3.4}$$

$$\sum_{\substack{v=(x,y):\\ |x|+|y|\ge 2\sqrt{t}\log t}} \mathbb{E}\left[Z_v^f(t) \mid \mathcal{E}_t\right] = \mathcal{O}\left(\frac{t}{\log t}\right) \quad \textit{if } \alpha = 3. \tag{3.5}$$

Proof. Let $L_{t'}^f$ be the two dimensional random variable representing the coordinates of the node the agent performing the Lévy flight is located in at time t'. Consider the projection of the Lévy flight on the x-axis, namely the random variable $X_{t'}$ such that $L_{t'}^f = (X_{t'}, Y_{t'})$. The random variable $X_{t'}$ can be expressed as the sum of t' random variables S_j^x, $j = 1,\dots,t'$, representing the jumps (with sign) that the projection of the walk takes at each of the t' rounds. The partial

distribution of the jumps along the x-axis, conditional on the event $\mathcal{E}_t$, can be derived as follows. [2] For any $0 \leq d \leq (t \log t)^{\frac{1}{\alpha-1}}$,

$$\begin{aligned}
&\Pr\left[S_j^x = \pm d \mid S_j \leq (t\log t)^{\frac{1}{\alpha-1}}\right] \\
&= \left[\frac{1}{2} + \sum_{k=1}^{(t\log t)^{\frac{1}{\alpha-1}}} \frac{c_\alpha}{2k^{\alpha+1}}\right] \mathbb{1}_{d=0} + \left[\frac{c_\alpha}{2d^{\alpha+1}} + \sum_{k=1+d}^{(t\log t)^{\frac{1}{\alpha-1}}} \frac{c_\alpha}{k^{\alpha+1}}\right] \mathbb{1}_{d\neq 0}, \qquad (3.6)
\end{aligned}$$

where: $\mathbb{1}_{d\in A}$ returns 1 if $d \in A$ and 0 otherwise, the term

$$\frac{\mathbb{1}_{d=0}}{2} + \frac{c_\alpha}{2d^{\alpha+1}}\mathbb{1}_{d\neq 0}$$

is the probability that the original jump lies along the horizontal axis and has "length" exactly d (there are two such jumps if $d > 0$), and, for $k \geq 1 + d$, the terms

$$\frac{c_\alpha}{2k^{\alpha+1}}\mathbb{1}_{d=0} + \frac{c_\alpha}{k^{\alpha+1}}\mathbb{1}_{d\neq 0}$$

are the probability that the original jump has "length" exactly k and its projection on the horizontal axis has "length" d (there are two such jumps if $d = 0$, and four such jumps if $d > 0$). Observe that (3.6) is of the order of

$$\Theta\left(\frac{1}{d^{\alpha+1}} + \sum_{k=1+d}^{(t\log t)^{\frac{1}{\alpha-1}}} \frac{1}{k^{\alpha+1}}\right).$$

By the integral test (Lemma E.1 in E), we know that this probability is

$$\Pr\left[S_j^x = \pm d \mid \mathcal{E}_j\right] = \Theta\left(\frac{1}{d^\alpha}\right).$$

Due to symmetry, it is easy to see that $\mathbb{E}\left[X_{t'} \mid \mathcal{E}_t\right] = 0$ for each time t', while

$$\mathrm{Var}\left(X_{t'} \mid \mathcal{E}_t\right) = \sum_{i=1}^{t'} \mathrm{Var}\left(S_j^x \mid \mathcal{E}_j\right) = t'\mathrm{Var}\left(S_1^x \mid E_1\right)$$

since $S_1^x, \ldots, S_{t'}^x$ are i.i.d.

As for the case $\alpha \in (2,3)$, the variance of S_1^x conditioned to the event $E_1 = \left\{S_1 \leq (t\log t)^{\frac{1}{\alpha-1}}\right\}$, can be bounded as follows

$$\mathrm{Var}\left(S_1^x \mid E_1\right) \leq \sum_{k=1}^{(t\log t)^{\frac{1}{\alpha-1}}} \mathcal{O}\left(\frac{k^2}{k^\alpha}\right) \overset{(*)}{=} \mathcal{O}\left(\frac{1}{3-\alpha}\left[(t\log t)^{\frac{3-\alpha}{\alpha-1}} - 1\right]\right) = \mathcal{O}\left(\frac{(t\log t)^{\frac{3-\alpha}{\alpha-1}}}{3-\alpha}\right),$$

where, in $(*)$, we used the integral test (Lemma E.1). Observe that the event $\mathcal{E}_t = \bigcap_{i=1}^t E_i$ has probability

$$\Pr\left[\mathcal{E}_t\right] = 1 - \mathcal{O}\left(\frac{1}{\log t}\right).$$

2. We remark that in Appendix F.2 we estimate the unconditional distribution of the jump projection length on the x-axis (Lemma F.1) for any $\alpha > 1$. Nevertheless, in this case we are conditioning on the event the original two dimensional jump is bounded, and thus we cannot make use of Lemma F.1.

Then, for each $t' \leq t$, from the Chebyshev's inequality and the fact that $\mathbb{E}\left[X_{t'} \mid \mathcal{E}_t\right] = 0$,

$$\Pr\left[|X_{t'}| \geq (t\log t)^{\frac{1}{\alpha-1}} \mid \mathcal{E}_t\right] \leq \frac{t'\mathrm{Var}\left(S_1^x \mid E_1\right)}{(t\log t)^{\frac{2}{\alpha-1}}} \leq \frac{t\mathrm{Var}\left(S_1^x \mid E_1\right)}{(t\log t)^{\frac{2}{\alpha-1}}} = \mathcal{O}\left(\frac{1}{(3-\alpha)\log t}\right),$$

which implies that

$$\Pr\left[|X_{t'}| \geq (t\log t)^{\frac{1}{\alpha-1}}\right] \leq \Pr\left[|X_{t'}| \geq (t\log t)^{\frac{1}{\alpha-1}} \mid \mathcal{E}_t\right] + \Pr\left[\mathcal{E}_t^C\right] = \mathcal{O}\left(\frac{1}{(3-\alpha)\log t}\right).$$

Then, the probability that both $X_{t'}$ and $Y_{t'}$ are less than $(t\log t)^{\frac{1}{\alpha-1}}$ (call the corresponding events $A_{x,t'}$ and $A_{y,t'}$, respectively) is

$$\Pr\left[A_{x,t'} \cap A_{y,t'}\right] = \Pr\left[A_{x,t'}\right] + \Pr\left[A_{y,t'}\right] - \Pr\left[A_{x,t'} \cup A_{y,t'}\right] \geq 1 - \mathcal{O}\left(\frac{1}{(3-\alpha)\log t}\right),$$

for any $t' \leq t$. Then, let $Z'(t)$ be the random variable indicating the number of times the Lévy flight visits the set of nodes whose coordinate absolute values are both no less than $(t\log t)^{\frac{1}{\alpha-1}}$, until time t. Then,

$$\mathbb{E}\left[Z'(t) \mid \mathcal{E}_t\right] = \sum_{\substack{v=(x,y) \\ |x|+|y| \geq 2(t\log t)^{\frac{1}{\alpha-1}}}} \mathbb{E}\left[Z_v^f(t) \mid \mathcal{E}_t\right],$$

and

$$\mathbb{E}\left[Z'(t) \mid \mathcal{E}_t\right] = \sum_{i=1}^{t} \Pr\left[\left(A_{x,i} \cap A_{y,i}\right)^C \mid \mathcal{E}_t\right] = t \cdot \mathcal{O}\left(\frac{1}{(3-\alpha)\log t}\right) = \mathcal{O}\left(\frac{t}{(3-\alpha)\log t}\right),$$

which proves Eq. (3.4).

As for the case $\alpha = 3$, the variance of S_1^x conditional on E_1 is $\mathcal{O}\left(\log(t\log t)\right)$. Then, we look at the probability that $|X_{t'}|$ is at least $\sqrt{t} \cdot \log t$ conditional on $\mathcal{E}_t$, which is, again, $\mathcal{O}\left(1/\log t\right)$. Finally, the proof proceeds in exactly the same way of the previous case, obtaining Eq. (3.5). □

Lemma 3.3.11. *For $t = \Theta(\ell^{\alpha-1})$,*

$$ct + \mathbb{E}\left[Z_{u^*}^f(t) \mid \mathcal{E}_t\right] \cdot 4(t\log t)^{\frac{2}{\alpha-1}} + \mathcal{O}\left(\frac{t}{(3-\alpha)\log t}\right) \geq t \;\; \textit{if } \alpha \in (2,3); \tag{3.7}$$

$$ct + \mathbb{E}\left[Z_{u^*}^f(t) \mid \mathcal{E}_t\right] \cdot 4t\log^2 t + \mathcal{O}\left(\frac{t}{\log t}\right) \geq t \;\; \textit{if } \alpha = 3. \tag{3.8}$$

Proof. Suppose the agent has made t jumps for some $t = \Theta(\ell^{\alpha-1})$ (the same t of Lemma 3.3.8), thus visiting exactly t nodes. Then,

$$\mathbb{E}\left[\sum_{v\in\mathbb{Z}^2} Z_v^f(t) \mid \mathcal{E}_t\right] = t.$$

As for Eq. (3.7), we observe that, from Lemma 3.3.8, the number of visits to $\mathcal{A}_1 = Q_\ell(\mathbf{0})$ until time t is at most ct, for some constant $c \in (0,1)$. From Lemma 3.3.10, the number of visits to $\mathcal{A}_3$

is at most $\mathcal{O}\left(t/\left((3-\alpha)\log t\right)\right)$. Thanks to Corollary 3.3.9, each of the remaining nodes, i.e., the nodes in $\mathcal{A}_2$ (whose size is at most $4(t\log t)^{\frac{2}{\alpha-1}}$), is visited by the agent at most $\mathbb{E}\left[Z^f_{u^*}(t) \mid \mathcal{E}_t\right]$ times. It follows that

$$ct + \mathbb{E}\left[Z^f_{u^*}(t) \mid \mathcal{E}_t\right] \cdot 4(t\log t)^{\frac{2}{\alpha-1}} + \mathcal{O}\left(\frac{t}{(3-\alpha)\log t}\right) \geq t.$$

As for Eq. (3.8), we proceed as for the first case above, by noticing that the number of visits to $\mathcal{A}_2$ is at most $\mathbb{E}\left[Z^f_{u^*}(t) \mid \mathcal{E}_t\right] \cdot (4t\log^2 t)$. This gives Eq. (3.8). □

The next two lemmas provide a clean relationship between the probability to hit a node u within time t to the average number of visits to the origin and to the average number of visits to u itself. In particular, the first lemma estimate the average number of visits to the origin. For any $t \geq 0$ and $\alpha \in (2,3]$, let $\mathbb{E}\left[Z^f_{\mathbf{0}}(t) \mid \mathcal{E}_t\right] = a_t(\alpha)$.

Lemma 3.3.12 (Visits to the origin)**.**
(a) If $\alpha \in (2,3)$, then $a_t(\alpha) = \mathcal{O}\left(1/(3-\alpha)^2\right)$.
(b) If $\alpha = 3$, then $a_t(3) = \mathcal{O}\left(\log^2 t\right)$.

Proof. For the case $\alpha \in (2,3)$, we proceed as follows. Since $\mathbb{E}\left[Z^f_{\mathbf{0}}(t) \mid \mathcal{E}_t\right] = \sum_{k=1}^{t} p_{\mathbf{0},k}$, it suffices to accurately bound the probability $p_{\mathbf{0},k}$ for each $k = 1,\ldots,t$. Let us make a partition of the natural numbers in the following way

$$\mathbb{N} = \bigcup_{t'=1}^{\infty} \left[\mathbb{N} \cap \left[2t'\log t', 2(t'+1)\log(t'+1)\right)\right].$$

For each $k \in \mathbb{N}$, there exists t' such that $k \in [2t'\log t', 2(t'+1)\log(t'+1))$. Then, within $2t'\log t'$ steps, we claim that the walk has moved to distance $\lambda = \frac{(t')^{\frac{1}{\alpha-1}}}{2}$ at least once, with probability $\Omega\left(\frac{1}{(t')^2}\right)$. Indeed, if there is one jump of length at least 2λ, then the walk has necessarily moved to a distance at least λ from the origin. We now bound the probability that one jump is at least 2λ. For the integral test and for $\lambda > 0$, we get

$$\begin{aligned}
\Pr\left[S_j \geq 2\lambda \mid S_j \leq (t\log t)^{\frac{1}{\alpha-1}}\right] &\geq \frac{1}{\Pr\left[S_j \leq (t\log t)^{\frac{1}{\alpha-1}}\right]} \left[\int_{2\lambda}^{(t\log t)^{\frac{1}{\alpha-1}}} \frac{c_\alpha}{s^\alpha} ds\right] \\
&\geq \frac{c_\alpha}{\alpha-1}\left(\frac{1}{t'} - \frac{1}{t\log t}\right) \geq \frac{c_\alpha}{\alpha-1}\left(\frac{1-\frac{t'}{t\log t}}{t'}\right) \\
&\geq \frac{c_\alpha}{\alpha-1}\left(\frac{1-\frac{1}{2\log(t')\log t}}{t'}\right) = \Omega\left(\frac{1}{t'}\right),
\end{aligned}$$

where the last inequality holds since $2t'\log t' \leq t$. Thus, the probability that the first $2t'\log t'$ jumps are less than 2λ is

$$\begin{aligned}
\Pr\left[\cap_{j=1}^{2t'\log t'}\{S_j < 2\lambda\} \mid \mathcal{E}_t\right] &\overset{(*)}{=} \left[1 - \Pr\left[S_1 < 2\lambda \mid S_1 \leq (t\log t)^{\frac{1}{\alpha-1}}\right]\right]^{2t'\log t'} \\
&\geq \left[1 - \Omega\left(\frac{1}{t'}\right)\right]^{2t'\log t'} = \mathcal{O}\left(\frac{1}{(t')^2}\right),
\end{aligned}$$

where in $(*)$ we used the independence among the agent's jumps. Once the agent reaches such a distance, Lemma 3.2.3 implies that there are at least $\lambda^2 = \Omega\left((t')^{\frac{2}{\alpha-1}}\right)$ different nodes that are at least as equally likely as $\mathbf{0}$ to be visited at any given future time. Thus, the probability to reach the origin at any future time is at most $\mathcal{O}\left(\frac{1}{(t')^{\frac{2}{\alpha-1}}}\right) = \mathcal{O}\left(\frac{1}{(t')^{1+\epsilon}}\right)$ with $\epsilon = (3-\alpha)/(\alpha-1) > 0$: in particular the bound holds for $p_{\mathbf{0},k}$. Observe that in an interval $[2t'\log t', 2(t'+1)\log(t'+1))$ there are

$$2(t'+1)\log(t'+1) - 2t'\log t' = 2t'\left[\log\left(1+\frac{1}{t'}\right)\right] + 2\log(t'+1) = \mathcal{O}\left(\log t'\right)$$

integers. Let L_t^f be the two-dimensional random variable denoting the node visited at time t by an agent which started from the origin, and let $H_{t'}$ be the event $\cup_{j=1}^{2t'\log t'}\{S_j \geq 2\lambda\}$. Observe that, by the law of total probability,

$$p_{\mathbf{0},k} = \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}, \mathcal{E}_t\right]\Pr\left[H_{t'} \mid \mathcal{E}_t\right] + \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}^C, \mathcal{E}_t\right]\Pr\left[H_{t'}^C \mid \mathcal{E}_t\right].$$

Thus, if $I_{t'} = [2t'\log t', 2(t'+1)\log(t'+1))$, we get

$$\begin{aligned}
\sum_{k=1}^{t} p_{\mathbf{0},k} &\leq \sum_{t'=1}^{t}\sum_{k\in I_{t'}} p_{\mathbf{0},k} \\
&\leq \sum_{t'=1}^{t}\left[\Pr\left[L_t^f = \mathbf{0} \mid H_{t'}, \mathcal{E}_t\right]\Pr\left[H_{t'} \mid \mathcal{E}_t\right] + \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}^C, \mathcal{E}_t\right]\Pr\left[H_{t'}^C \mid \mathcal{E}_t\right]\right]\mathcal{O}(\log t') \\
&\leq \sum_{t'=1}^{t}\left[\mathcal{O}\left(\frac{1}{(t')^{1+\epsilon}}\right) + \mathcal{O}\left(\frac{1}{(t')^2}\right)\right]\mathcal{O}(\log t') \\
&= \sum_{t'=1}^{t}\mathcal{O}\left(\frac{\log t'}{(t')^{1+\epsilon}}\right) \overset{(\star)}{=} \mathcal{O}\left(\frac{1}{\epsilon^2}\right) = \mathcal{O}\left(\frac{1}{(3-\alpha)^2}\right),
\end{aligned}$$

where for $\star$ we used the integral test and partial integration. In particular, it holds that

$$\int_1^t \frac{\log(x)}{x^{1+\epsilon}}\,\mathrm{d}x = \frac{1}{\epsilon^2}\left[\frac{\epsilon(1-\log x)}{x^\epsilon} - \frac{\epsilon+1}{x^\epsilon}\right]_1^t.$$

For the case $\alpha = 3$, we can consider the same argument above for the previous case where we fix $\lambda = \sqrt{t'}$. Then the proof proceeds as in the previous case by observing that the average number of visits until time t is, now, of magnitude $\mathcal{O}\left(\log^2 t\right)$. □

Lemma 3.3.13. *Let $u \in \mathbb{Z}^2$ be any node. Then,*

(i) $\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right] \leq a_t(\alpha)$,

(ii) $1 \leq \mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right] \leq a_t(\alpha)$,

(iii) $\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right]/a_t(\alpha) \leq Pr\left[Z_u^f(t) > 0 \mid \mathcal{E}_t\right] \leq \mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right]$.

Proof. Claim (i) is a direct consequence of (ii), since $\mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right] \geq \mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right]$. As for Claim (ii), let τ be the first time the agent visits u. Then, conditional

on $Z_u^f(t) > 0$, τ is at most t, and

$$\mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right] = \mathbb{E}\left[Z_{\mathbf{0}}^f(t-\tau) \mid \tau \le t, \mathcal{E}_t\right] \le \mathbb{E}\left[Z_{\mathbf{0}}^f(t) \mid \mathcal{E}_t\right] = a_t(\alpha).$$

Notice that this expectation is at least 1 since we have the conditional event. As for Claim (iii), let us explicitly write the term $\mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right] \cdot \Pr\left[Z_u^f(t) > 0 \mid \mathcal{E}_t\right]$:

$$\begin{aligned}
&\sum_{i=1}^{t} i \Pr\left[Z_u^f(t) = i \mid Z_u^f(t) > 0, \mathcal{E}_t\right] \cdot \Pr\left[Z_u^f(t) > 0 \mid \mathcal{E}_t\right] \\
&= \sum_{i=1}^{t} i \frac{\Pr\left[Z_u^f(t) = i, Z_u^f(t) > 0, \mathcal{E}_t\right]}{\Pr\left[Z_u^f(t) > 0, \mathcal{E}_t\right]} \cdot \frac{\Pr\left[Z_u^f(t) > 0, \mathcal{E}_t\right]}{\Pr\left[\mathcal{E}_t\right]} \\
&= \sum_{i=1}^{t} i \frac{\Pr\left[Z_u^f(t) = i, Z_u^f(t) > 0, \mathcal{E}_t\right]}{\Pr\left[\mathcal{E}_t\right]} = \sum_{i=1}^{t} i \Pr\left[Z_u^f(t) = i \mid \mathcal{E}_t\right] \\
&= \mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right].
\end{aligned}$$

Then,

$$\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right] \ge \Pr\left[Z_u^f(t) > 0 \mid \mathcal{E}_t\right] = \frac{\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right]}{\mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right]} \ge \frac{\mathbb{E}\left[Z_u^f(t) \mid \mathcal{E}_t\right]}{a_t(\alpha)},$$

since, from Claim (ii), $\mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0, \mathcal{E}_t\right] \le a_t(\alpha)$. □

We can now complete the proof of Lemma 3.3.5, as follows. From Lemma 3.3.11 we have that $\mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, \mathcal{E}_t\right] = \Omega\left((3-\alpha)^2 / \left(t^{(3-\alpha)/(\alpha-1)} (\log t)^{2/(\alpha-1)}\right)\right)$ if $\alpha \in (2,3)$ and $3 - \alpha = \omega(1/\log t)$ and $\mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, \mathcal{E}_t\right] = \Omega\left(1/(\log t)^2\right)$ if $\alpha = 3$. Then, Lemma 3.3.12 and claim (iii) of Lemma 3.3.13 give the results by substituting $t = \Theta\left(\ell^{\alpha-1}\right)$.

3.3.3 Proof of Lemma 3.3.6

Let S_j be the random variable denoting the j-th jump-length. From Eq. (3.2), we get

$$\Pr\left[S_j > (t \log t)^{\frac{1}{\alpha-1}}\right] = \Theta\left(\frac{1}{t \log t}\right).$$

Let E_j be the event $\left\{S_j \le (t \log t)^{\frac{1}{\alpha-1}}\right\}$, and let $\mathcal{E}_t$ be the intersection of E_j for $j = 1, \ldots, t$. Notice that, by the union bound, the probability of $\mathcal{E}_t$ is $1 - \mathcal{O}(1/\log t)$. We next apply the multiplicative form of the Chernoff bound to the sum of S_j, conditional on the event $\mathcal{E}_t$. This is possible since the variable $S_j/(t \log t)^{\frac{1}{\alpha-1}}$ takes values in $[0,1]$. To this aim, we first bound the expectation of the sum of the random variables S_j, for $j = 1, \ldots, t$ conditional on $\mathcal{E}_t$.

$$\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{t} S_j \,\middle|\, \mathcal{E}_t\right] &= \sum_{j=1}^{t} \mathbb{E}\left[S_j \mid \mathcal{E}_t\right] = \Theta(t) + t \frac{c_\alpha}{\Pr\left[\mathcal{E}_t\right]} \sum_{d=1}^{(t \log t)^{\frac{1}{\alpha-1}}} \frac{d}{d^\alpha} \\
&\le \Theta(t) + 2 c_\alpha t \sum_{d=1}^{(t \log t)^{\frac{1}{\alpha-1}}} \frac{1}{d^{\alpha-1}} \overset{(a)}{\le} \mathcal{O}(\mu t),
\end{aligned}$$

where in (a) we have $\mu = \min\{\log \ell, \frac{1}{\alpha-2}\}$ for the integral test (Lemma E.1). We now use the Chernoff bound (Lemma B.1) on the normalized sum of all jumps, to show that such a sum is at most linear in $\mathcal{O}(\mu t)$ with probability $1 - \exp\left(-t^{\Theta(1)}\right)$, conditional on $\mathcal{E}_t$. In formula,

$$\begin{aligned}\Pr\left[\sum_{j=1}^{t} S_j \geq \Theta(\mu t) \;\middle|\; \mathcal{E}_t\right] &= \Pr\left[\frac{\sum_{j=1}^{t} S_j}{(t\log t)^{\frac{1}{\alpha-1}}} \geq \frac{2\Theta(\mu t)}{(t\log t)^{\frac{1}{\alpha-1}}} \;\middle|\; \mathcal{E}_t\right] \\ &\leq \exp\left(-\frac{2\Theta(\mu t)}{3\left((t\log t)^{\frac{1}{\alpha-1}}\right)}\right) \\ &\leq \exp\left(-\Theta\left(\frac{\mu t^{\frac{\alpha-2}{\alpha-1}}}{(\log t)^{\frac{1}{\alpha-1}}}\right)\right) \leq \exp\left(-\Theta\left(t^{\frac{\alpha-2}{2(\alpha-1)}}\right)\right).\end{aligned}$$

Then, define

$$\begin{aligned}A &= \{\text{the Lévy walk finds the target within time } \Theta(\mu t)\} = \{\tau_\alpha(u^*) \leq \Theta(\mu t)\}, \\ A_1 &= \left\{\sum_{j=1}^{t} S_j = \Theta(\mu t)\right\}, \text{ and} \\ A_2 &= \{\text{the Lévy flight finds the target within } t \text{ jumps}\} = \{h_f \leq t\}.\end{aligned}$$

Observe that the event $A_1 \cap A_2$ implies that the Lévy walk finds the target within t jumps, which, in turn, implies the event A. Indeed, $A_1 \cap A_2$ implies the target is found in one of the t jump endpoints, and the overall amount of steps is $\Theta(\mu t)$. Let $p(t) = \Pr[h_f \leq t \mid \mathcal{E}_t]$. Then

$$\begin{aligned}\Pr[A] \geq \Pr[A_1, A_2] &\geq \Pr[A_1, A_2, \mathcal{E}_t] \\ &= \Pr[\mathcal{E}_t]\left[\Pr[A_1 \mid \mathcal{E}_t] + \Pr[A_2 \mid \mathcal{E}_t] - \Pr[A_1 \cup A_2 \mid \mathcal{E}_t]\right] \\ &\geq \left(1 - \mathcal{O}\left(\frac{1}{\log t}\right)\right)\left[1 - exp(-t^{\Theta(1)}) + p(t) - 1\right] \\ &= \left(1 - \mathcal{O}\left(\frac{1}{\log t}\right)\right)\left(p(t) - exp\left(-t^{\Theta(1)}\right)\right),\end{aligned}$$

where in the second line we used the definition of conditional probability and the inclusion-exclusion principle, and in the third line we used that $\Pr[\mathcal{E}_t] = (1 - \mathcal{O}(1/\log t))$, $\Pr[A_1 \mid \mathcal{E}_t] \geq 1 - \exp\left(-t^{\Theta}(1)\right)$, and $\Pr[A_1 \cup A_2 \mid \mathcal{E}_t] \leq 1$.

3.3.4 Proof of Lemma 3.3.7

Let X_i be the x-coordinate of the agent at the end of the i-th jump. For any $i \leq t$, we bound the probability that $X_i > \ell/4$. The probability that there is a jump whose length is at least ℓ among the first i jumps is $\Theta\left(i/\ell^{\alpha-1}\right)$. We first consider the case $\alpha \in (2,3)$. Conditional on the event that the first i jump-lengths are all smaller than ℓ (event C_i), the expectation of X_i is zero and its variance is

$$\Theta(1) + i \cdot \sum_{d=1}^{\ell/4} \Theta\left(\frac{d^2}{d^\alpha}\right) = \Theta\left(i \cdot \nu \ell^{3-\alpha}\right),$$

for the integral test (Lemma E.1), where $\nu = \min\{\log \ell, \frac{1}{3-\alpha}\}$. Chebyshev's inequality implies that

$$\Pr\left[|X_i| \geq \ell/4 \mid C_i\right] \leq \frac{\Theta\left(i \cdot \nu \ell^{3-\alpha}\right)}{\Theta(\ell^2)} = \Theta\left(\frac{i\nu}{\ell^{\alpha-1}}\right).$$

Since the conditional event has probability $1 - \Theta\left(i/\ell^{\alpha-1}\right)$, then the "unconditional" probability of the event $|X_i| \leq \ell/4$ is

$$\left[1 - \Theta\left(\frac{i}{\ell^{\alpha-1}}\right)\right]\left[1 - \mathcal{O}\left(\frac{i\nu}{\ell^{\alpha-1}}\right)\right] = 1 - \mathcal{O}\left(\frac{\nu t}{\ell^{\alpha-1}}\right),$$

since $i \leq t$, for t which is some $\mathcal{O}\left(\ell^{\alpha-1}/\nu\right)$. The same result holds analogously for Y_i (the y-coordinate of the agent after the i-th jump), thus obtaining $|X_i| + |Y_i| \leq \ell/2$, with probability $1 - \mathcal{O}\left(\nu t/\ell^{\alpha-1}\right)$ by the union bound.

During the first jump-phase, thanks to Corollary 3.2.2, the probability the agents visits the target is $\mathcal{O}\left(1/\ell^{\alpha}\right)$. Let $2 \leq i \leq t$. We want to estimate the probability that during i-th jump-phase the agents visits the target, having the additional information that $t = \mathcal{O}\left(\ell^{\alpha-1}\right)$. As in the proof of Lemma 3.2.4, we consider the node u^* where the target is located on, and the rhombus centered in u^* that contains the nodes within distance $\ell/4$ from u^*, namely $B_{\ell/4}(u^*)$. Let F_i be the event that during the i-th jump-phase the agent visits the target; let V_{i-1} be the event that the $(i-1)$-th jump ends in $B_{\ell/4}(u^*)$, and let W_{i-1} be the event that the $(i-1)$-th jump ends at distance farther than $\ell/2$ from the origin. Finally, let L_i^f be the two-dimensional random variable denoting the coordinates of the node the agent is located on at the end of the i-th jump-phase. Then,

$$\begin{aligned}\Pr\left[F_i \mid V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] &= \sum_{v \in B_{\ell/4}(u^*)} \Pr\left[F_i \mid L_i^f = v\right]\Pr\left[L_i^f = v \mid W_{i-1}\right] \\ &\leq \mathcal{O}\left(\frac{1}{\ell^2}\right) \sum_{v \in B_{\ell/4}(u^*)} \Pr\left[F_i \mid L_i^f = v\right],\end{aligned}$$

where in the above inequalities we used the monotonicity property (Lemma 3.2.3, which holds since the process restricted to the jump endpoints is a Lévy flight), and the fact that, for each $v \in B_{\ell/4}(u^*)$, there are at least $\Theta\left(\ell^2\right)$ nodes at distance at least $\ell/2$ from the origin which are more likely to be the destination of the i-th jump than v. Then, we proceed as in the proof of Lemma 3.2.4 and obtain

$$\Pr\left[F_i \mid V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] = \mathcal{O}\left(\frac{\mu}{\ell^2}\right), \tag{3.9}$$

where $\mu = \min\{\log \ell, \frac{1}{\alpha-2}\}$. By the law of total probabilities, we get

$$\begin{aligned}\Pr\left[F_i\right] &= \Pr\left[F_i \mid W_{i-1}\right]\Pr\left[W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\ &= \left[\Pr\left[F_i \mid W_{i-1}, V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right]\Pr\left[V_{i-1}^C \mid W_{i-1}\right]\right]\Pr\left[W_{i-1}\right] \\ &+ \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\ &\leq \left[\Pr\left[F_i \mid V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right]\right]\Pr\left[W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\ &\leq \left[\mathcal{O}\left(\frac{\mu}{\ell^2}\right) + \mathcal{O}\left(\frac{1}{\ell^{\alpha}}\right)\right]\mathcal{O}\left(\frac{\nu t}{\ell^{\alpha-1}}\right) + \mathcal{O}\left(\frac{1}{\ell^{\alpha}}\right) = \mathcal{O}\left(\frac{\nu\mu t}{\ell^{\alpha+1}}\right),\end{aligned} \tag{3.10}$$

where for second-last inequality we used that $V_{i-1} \subset W_{i-1}$ and that $\Pr\left[V_{i-1}^C \mid W_{i-1}\right] \leq 1$, while for the last inequality we used Eq. (3.9), and that $\Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right] = \mathcal{O}\left(1/\ell^\alpha\right)$, which is true because the jump starts in a node whose distance form the target is $\Omega(\ell)$, and that $\Pr\left[F_i \mid W_{i-1}^C\right] = \mathcal{O}\left(1/\ell^\alpha\right)$, which is true for the same reason.

Thus, by the union bound and by Eq. (3.10), the probability that during at least one between the t jump-phases, the agent finds the target is

$$\mathcal{O}\left(\frac{1}{\ell^\alpha}\right) + (t-1)\mathcal{O}\left(\frac{\nu\mu t}{\ell^{\alpha+1}}\right) = \mathcal{O}\left(\frac{\nu\mu t^2}{\ell^{\alpha+1}}\right)$$

since $t \geq \ell$, which gives the first claim of the lemma by observing that within time t at most t jumps can be performed.

Consider now the case $\alpha = 3$. The proof proceeds exactly as in the first case, with the only key difference that the variance of X_i is $\Theta\left(i \log \ell\right)$. This means that the probability that $|X_i|$ is at least $\ell/4$ conditional to C_i is $\mathcal{O}\left(\log \ell/\ell^2\right)$, and the "unconditional" probability that $|X_i|$ is less than $\ell/4$ is $1 - \mathcal{O}\left(t \log \ell/\ell^2\right)$. It thus follows that

$$\Pr\left[F_i\right] = \mathcal{O}\left(t \log \ell/\ell^4\right).$$

Then we get the second claimed bound of the lemma: $\mathcal{O}\left(t^2 \log \ell/\ell^4\right)$.

3.3.5 Proof of Corollary 3.3.2

From Theorem 3.3.1Part a and the independence among the agents, we get that

$$\Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\mu\ell^{\alpha-1}\right)\right] = 1 - \left[1 - \Omega\left(\frac{1}{\gamma\ell^{3-\alpha}}\right)\right]^k \geq 1 - e^{\Omega\left(\frac{k}{\gamma\ell^{3-\alpha}}\right)},$$

where we have used the inequality $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$. Then, if $\alpha = \alpha^* + 5\frac{\log\log\ell}{\log\ell}$,

$$\Pr\left[\tau_\alpha^k(u^*) = o\left(\frac{\ell^2 \log^6 \ell}{k}\right)\right] = 1 - e^{-\omega(\log \ell)},$$

since $\mu \leq \frac{1}{\alpha-2} \leq \frac{\log\ell}{\log\log\ell}$, $\ell^{\alpha-1} = \frac{\ell^2 \log^5 \ell}{k}$, $\gamma = o\left(\log^4 \ell\right)$ and $\ell^{3-\alpha} = \frac{k}{\log^5 \ell}$, thus giving Claim (a). From Theorem 3.3.1.Part b and the independence among the agents, we get

$$\Pr\left[\tau_\alpha^k(u^*) > t\right] = \left[1 - \mathcal{O}\left(\frac{\mu\nu t^2}{\ell^{\alpha+1}}\right)\right]^k,$$

for $\ell \leq t = o\left(\ell^{\alpha-1}/\nu\right)$. Let $t = \frac{\ell^2 \cdot \ell^{\frac{\alpha-\alpha^*}{2}}}{k \log^4 \ell}$ which is a function in $o\left(\ell^{\alpha-1}/\nu\right)$ since $\alpha > \overline{\alpha}$. If $t \geq \ell$, we get

$$\Pr\left[\tau_\alpha^k(u^*) > t\right] = \left[1 - \mathcal{O}\left(\frac{\mu\nu\ell^4}{k^2\ell^{\alpha^*+1}\log^8 \ell}\right)\right]^k \geq e^{-\mathcal{O}\left(\frac{1}{\log^6 \ell}\right)} = 1 - \mathcal{O}\left(\frac{1}{\log^6 \ell}\right),$$

since $\mu\nu \leq \log^2 \ell$, $\ell^{\alpha^*+1} = \ell^4/k$. Notice that, in the inequality we have used that $1 - x \geq e^{-\frac{x}{1-x}}$ if $x < 1$, and in the last equality we have used the Taylor's expansion of the exponential function. If $t < \ell$, we get $\Pr\left[\tau_\alpha^k(u^*) > t\right] = 1$ (at least ℓ steps are needed to reach the target). Therefore, Claim (b) follows. Finally, from Theorem 3.3.1.Part c and the independence among the agents, we get

$$\Pr\left[\tau_\alpha^k(u^*) = \infty\right] = \left[1 - \mathcal{O}\left(\frac{\mu \log \ell}{\ell^{3-\alpha}}\right)\right]^k = \left[1 - \mathcal{O}\left(\frac{\log^2 \ell}{k\ell^{\alpha^*-\alpha}}\right)\right]^k \geq \exp\left(-\mathcal{O}\left(\frac{\log^2 \ell}{\ell^{\alpha^*-\alpha}}\right)\right),$$

since $\mu \leq \log^\ell$, $\ell^{3-\alpha} = \ell^{3-\alpha^*} \cdot \ell^{\alpha^*-\alpha} = k\ell^{\alpha^*-\alpha}$. Notice that in the last inequality we have used that $1 - x \geq e^{-\frac{x}{1-x}}$ for $x < 1$, which is our case since $k \geq \log^6 \ell$. Hence, we have Claim (c).

3.4 The case $\alpha \in (1, 2]$

We now analyze the hitting time of Lévy walks with parameter $\alpha \in (1, 2]$, which is the exponent range for which the jump length has unbounded mean and unbounded variance. We show the following theorems.

Theorem 3.4.1. *Let $\alpha \in [1 + \epsilon, 2)$, where $\epsilon > 0$ is an arbitrarily small constant. Let $u^* \in \mathbb{Z}^2$, and $\ell = \|u^*\|_1$. Let $\mu = \min\{\log \ell, \frac{1}{2-\alpha}\}$. Then:*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell\right)\right] = \Omega\left(1/\mu\ell\right)$;

(b) $Pr\left[\tau_\alpha(u^*) < \infty\right] = \mathcal{O}\left(\mu \log \ell/\ell\right)$.

Theorem 3.4.2. *Let $u^* \in \mathbb{Z}^2$ and $\ell = \|u^*\|_1$. Then:*

(a) $Pr\left[\tau_2(u^*) = \mathcal{O}\left(\ell\right)\right] = \Omega\left(1/\ell \log \ell\right)$;

(b) $Pr\left[\tau_2(u^*) < \infty\right] = \mathcal{O}\left(\log^2 \ell/\ell\right)$.

The above theorems imply the following bounds on the parallel hitting time.

Corollary 3.4.3. *Let $\alpha \in [1 + \epsilon, 2]$, where $\epsilon > 0$ is an arbitrarily small constant. Let $u^* \in \mathbb{Z}^2$ and $\ell = \|u^*\|_1$. Let $\mu = \min\{\log \ell, \frac{1}{2-\alpha}\}$. Then:*

(a) $Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\ell\right)\right] = 1 - e^{-\omega(\log \ell)}$, if $k = \omega\left(\ell \log^2 \ell\right)$;

(b) $Pr\left[\tau_\alpha^k(u^*) < \infty\right] = o(1)$, if $k = o\left(\ell/\log^2 \ell\right)$.

Recall also that the trivial lower bound $\tau_\alpha^k(u^*) \geq \ell$ holds.

3.4.1 Proof of Theorems 3.4.1 and 3.4.2

We first show the following lemma, which bounds the hitting time for a single Lévy walk.

Lemma 3.4.4. *Let $\alpha \in (1, 2]$ and $u^* \in \mathbb{Z}^2$ with $\|u^*\|_1 = \ell$. Then,*

(a) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}(\ell)\right] = \Omega\left(\frac{1}{\mu\ell}\right)$, if $\alpha \in [1 + \epsilon, 2)$ for an arbitrarily small constant $\epsilon > 0$, where $\mu = \min\{\frac{1}{\alpha-2}, \log \ell\}$.

(b) $Pr\left[\tau_\alpha(u^*) = \mathcal{O}(\ell)\right] = \Omega\left(\frac{1}{\ell \log \ell}\right)$, if $\alpha = 2$.

Proof. Consider a single agent moving according the Lévy walk with parameter $\alpha \in (1,2]$. By Equation (3.2) in Section 3.2, the probability the agent chooses a jump of length at least ℓ is of the order of $\Theta\left(1/d^{\alpha-1}\right)$. Let c be some constant to be fixed later, and let μ_α be equal to $\min\{\frac{1}{\alpha-2}, \log \ell\}$ if $\alpha < 2$, and to $\log \ell$ if $\alpha = 2$. Then, the probability that all the first $c\ell^{\alpha-1}/\mu_\alpha$ jumps have length less than ℓ is

$$\left(1 - \Theta\left(\frac{1}{\ell^{\alpha-1}}\right)\right)^{\frac{c\ell^{\alpha-1}}{\mu_\alpha}}$$

which is greater than positive constant strictly less than 1 thanks to the inequality $\exp(-x/(1-x)) < (1-x)$ for $x < 1$. Let E_i be the event that the i-th jump-length is less than ℓ and $\mathcal{E}_i = \cap_{1\le j\le i} E_j$. By what has been said before, we have

$$\Pr\left[\mathcal{E}_i\right] \ge \Theta\left(1\right) \quad \text{for all } i \le c\ell^{\alpha-1}/\mu_\alpha.$$

Conditional on $\mathcal{E}_i$, the sum of the first i jumps is at most $3\ell/4$ with constant probability. Indeed, if $j < i$, the expected value of S_j is, for the integral test (Lemma E.1)

$$\mathbb{E}\left[S_j \mid \mathcal{E}_i\right] = \mathcal{O}\left(1\right) + \sum_{d=1}^{\ell-1} \frac{c_\alpha d}{d^\alpha} = \mathcal{O}\left(\mu_\alpha \ell^{2-\alpha}\right).$$

Thus,

$$\mathbb{E}\left[\sum_{j=1}^{i} S_j \mid \mathcal{E}_i\right] \le \sum_{j=1}^{c\ell^{\alpha-1}/\mu_\alpha} \mathbb{E}\left[S_j \mid \mathcal{E}_i\right] = \mathcal{O}\left(c\ell\right).$$

We choose c small enough so that this expression is less than $\ell/2$. Conditional on $\mathcal{E}_i$, the $\{S_j\}_{j\le i}$ random variables are non negative and we can use the Markov's inequality to get that their sum is bounded by $3\ell/4$ with constant probability. Indeed

$$\Pr\left[\sum_{j=1}^{i} S_j \ge \frac{3\ell}{4} \,\middle|\, \mathcal{E}_i\right] \le \Pr\left[\sum_{j=1}^{i} S_j \ge \frac{3\mathbb{E}\left[\sum_{j=1}^{i} S_j \,\middle|\, \mathcal{E}_i\right]}{2} \,\middle|\, \mathcal{E}_i\right] \le \frac{2}{3}.$$

The latter implies there is at least constant probability the agent has displacement at most $3\ell/4$ from the origin during the first $c\ell^{\alpha-1}/\mu_\alpha$ jumps and in time $\mathcal{O}\left(\ell\right)$ (since the sum of all jumps is at most linear), without any conditional event:

$$\begin{aligned}\Pr\left[\sum_{j\le i} S_j \le 3\ell/4\right] &\ge \Pr\left[\sum_{j\le i} S_j \le 3\ell/4 \,\middle|\, \mathcal{E}_i\right] \Pr\left[\mathcal{E}_i\right] \\ &\ge \Theta(1)\end{aligned}$$

for each $i \le c\ell^{\alpha-1}/\mu_\alpha$.

Define the event $W_i = \{\sum_{j\le i} S_j \le 3\ell/4\}$. We now compute the probability that, given $i \le c\ell^{\alpha-1}/\mu_\alpha$, in the first $i-1$ jumps the displacement has been at most $3\ell/4$ and during the i-th jump-phase the agent finds the target. Let F_i be such the latter event. Since

$$\Pr\left[F_i, W_{i-1}\right] = \Pr\left[F_i \mid W_{i-1}\right] \Pr\left[W_{i-1}\right],$$

we estimate $\Pr[F_i \mid W_{i-1}]$. Let L_t^w be the two-dimensional random variable representing the coordinates of the nodes the Lévy walk visits at time t. If t_i is the time the agent ends the i-th jump-phase, we have

$$\Pr[F_i \mid W_{i-1}] \geq \sum_{v \in Q_{3\ell/4}(\mathbf{0})} \Pr\left[F_i \;\middle|\; L^w_{t_{i-1}} = v, W_{i-1}\right] \Pr\left[L^w_{t_{i-1}} = v \;\middle|\; W_{i-1}\right].$$

By Corollary 3.2.2, the term $\Pr\left[F_i \;\middle|\; L^w_{t_{i-1}} = v, W_{i-1}\right]$ is $\Theta\left(1/\ell^\alpha\right)$, and, since $L^w_{t_{i-1}} \in Q_{3\ell/4}(\mathbf{0})$ is implied by W_{i-1}, we have

$$\begin{aligned}
&\sum_{v \in Q_{3\ell/4}(\mathbf{0})} \Pr\left[F_i \;\middle|\; L^w_{t_{i-1}} = v, W_{i-1}\right] \Pr\left[L^w_{t_{i-1}} = v \;\middle|\; W_{i-1}\right] \\
&\geq \Theta\left(\frac{1}{\ell^\alpha}\right) \cdot \Pr\left[L^w_{t_{i-1}} \in Q_{3\ell/4}(\mathbf{0}) \;\middle|\; W_{i-1}\right] = \Theta\left(\frac{1}{\ell^\alpha}\right),
\end{aligned}$$

implying $\Pr[F_i, W_{i-1}] = \Omega\left(\frac{1}{\ell^\alpha}\right)$ for all $i \leq c\ell^{\alpha-1}/\left(1 + \log \ell \cdot \mathbb{1}_{[\alpha=2]}\right)$. Then, for the chain rule, the probability that none of the events $F_i \cap W_{i-1}$ holds for each $i \leq \ell^{\alpha-1}/\log(c\ell)$ is

$$\begin{aligned}
\Pr\left[\bigcup_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} (F_i \cap W_{i-1})\right] &= 1 - \Pr\left[\bigcap_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} (F_i^C \cup W_{i-1}^C)\right] \\
&= 1 - \prod_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} \Pr\left[F_i^C \cup W_{i-1}^C \mid \bigcap_{j \leq i-1} (J_j^C \cup W_{j-1}^C)\right] \\
&= 1 - \prod_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} \left(1 - \Pr\left[F_i \cap W_{i-1} \mid \bigcap_{j \leq i-1} (J_j^C \cup W_{j-1}^C)\right]\right) \\
&\overset{(*)}{\geq} 1 - \prod_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} \left(1 - \Pr\left[F_i \cap W_{i-1}, \bigcap_{j \leq i-1} (J_j^C \cup W_{j-1}^C)\right]\right) \\
&\overset{(\star)}{=} 1 - \prod_{i \leq \frac{c\ell^{\alpha-1}}{\mu_\alpha}} \left(1 - \Pr[F_i \cap W_{i-1}]\right) \\
&= 1 - \left(1 - \Omega\left(\frac{1}{\ell^\alpha}\right)\right)^{\frac{c\ell^{\alpha-1}}{\mu_\alpha}} \\
&\geq 1 - e^{-\Omega\left(\frac{c}{\mu_\alpha \ell}\right)} = \Omega\left(\frac{c}{\mu_\alpha \ell}\right),
\end{aligned}$$

where, $(*)$ holds since $\Pr[A \mid B] \geq \Pr[A, B]$, $(\star)$ holds since $W_{i-1} \subseteq (W_{j-1}^C \cup J_j^C)$ for $j \leq i-1$, and the last equality holds by the inequality $e^{-x} \geq 1 - x$ for all x, and by the Taylor's expansion of $f(x) = e^x$. Then, there is probability at least $\Omega\left(c/\left(\mu_\alpha \ell\right)\right)$ to find the target within time $\mathcal{O}(\ell)$. □

Lemma 3.4.4 gives Part a of Theorems 3.4.1 and 3.4.2, while Part b comes from Lemma 3.2.5.

3.4.2 Proof of Corollary 3.4.3

First, suppose $\alpha \in [1+\epsilon, 2)$. From Corollary 3.4.3.Part a and the independence between agents, we get that

$$\Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}(\ell)\right] = 1 - \left[\Omega\left(\frac{1}{\mu\ell}\right)\right]^k \geq 1 - e^{-\Omega\left(\frac{k}{\ell \log \ell}\right)},$$

where we used the inequality $1 - x \leq e^{-x}$ for every real x and the bound $\mu \leq \log \ell$. From Corollary 3.4.3.Part b and the independence between agents, we get that

$$\Pr\left[\tau_\alpha^k(u^*) = \infty\right] = \left[1 - \mathcal{O}\left(\frac{\mu \log \ell}{\ell}\right)\right]^k \geq \exp\left(-\mathcal{O}\left(\frac{k \log^2 \ell}{\ell}\right)\right),$$

where we used again $\mu \leq \log \ell$ and the inequality $1 - x \geq e^{-\frac{x}{1-x}}$ for every real x. Thus, if $k = o\left(\ell / \log^2 \ell\right)$, for the Taylor's expansion of the exponential function, we get hitting time ∞ with probability $1 - o(1)$. For $\alpha = 2$ the proof proceeds exactly in the same way.

3.5 The case $\alpha \in (3, \infty)$

We analyze now the hitting time of Lévy walks with parameter $\alpha \in (3, \infty)$, which is the exponent range for which the jump length has bounded mean and bounded variance.

Theorem 3.5.1. *Let $\alpha \in (3, \infty)$ and $u^* \in \mathbb{Z}^2$ with $\ell = \|u^*\|_1$. Let $\nu = \min\{\log \ell, \frac{1}{\alpha - 3}\}$, and $\gamma = \frac{\alpha^2}{(\alpha-3)^2}$. Then:*

(a) $\Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = \Omega\left(1/(\gamma \log^4 \ell)\right)$, if $\alpha \geq 3 + \omega(\log\log \ell / \log \ell)$;

(b) $\Pr[\tau_\alpha(u^*) \leq t] = \mathcal{O}\left(\nu \cdot t^2 / \ell^{\alpha+1}\right)$, for any step $\ell \leq t = \mathcal{O}\left(\ell^2/\nu\right)$.

From the above result, we easily obtain the following bounds on the parallel hitting time.

Corollary 3.5.2. *Let $\alpha \in (3, \infty)$, $u^* \in \mathbb{Z}^2$, $\ell = \|u^*\|_1$, and $1 \leq k \leq o\left(\ell^2\right)$. Let $\nu = \min\{\log \ell, \frac{1}{\alpha-3}\}$ and $\gamma = \frac{\alpha^2}{(\alpha-3)^2}$. Then:*

(a) $\Pr\left[\tau_\alpha^k(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = 1 - e^{-\omega(\log \ell)}$, if $k \geq \Omega\left(\log^6 \ell\right)$ and $\alpha \geq 3 + \omega\left(\frac{\log\log \ell}{\log \ell}\right)$;

(b) $\Pr\left[\tau_\alpha^k(u^*) \leq \ell^2/\sqrt{k}\right] = 1 - o(1)$.

Corollary 3.5.2.a says that $\tau_\alpha^k(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)$, w.h.p., for $k \geq \operatorname{polylog} \ell$ and $\alpha \geq 3 + \omega\left(\frac{\log\log \ell}{\log \ell}\right)$, and Corollary 3.5.2.b provides a crude lower bound indicating that increasing k beyond $\operatorname{polylog} \ell$, can only result in sublinear improvement.

3.5.1 Proof of Theorem 3.5.1

The structure of the proof is similar to that for Theorems 3.3.1 and 3.3.3. We will use the next three lemmas, which are analogous to Lemmas 3.3.5 to 3.3.7, respectively

Lemma 3.5.3 (Lévy flight with $\alpha \in (3,\infty)$)**.** *Let h_f be the hitting time of a Lévy flight for target $u^* \in \mathbb{Z}^2$, and let $\ell = \|u^*\|_1$. If $\alpha - 3 = \omega\left(\log\log \ell / \log \ell\right)$, then*

$$Pr\left[h_f = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = \Omega\left(\frac{(\alpha-3)^3}{\alpha^2 \log^4 \ell}\right).$$

Lemma 3.5.4. *Let h_f be defined as in Lemma 3.5.3, and let $\tau_\alpha(u^*)$ be the hitting time of a Lévy walk with the same $\alpha \in (3,\infty)$, for the same target. Then, for every step t,*

$$Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(t\right)\right] \geq Pr\left[h_f \leq t\right] - \mathcal{O}\left(\frac{\alpha}{(\alpha-3)t}\right).$$

Lemma 3.5.5. *Let $\alpha \in (3,\infty)$, $u^* \in \mathbb{Z}^2$, and $\ell = \|u^*\|_1$. For any step t such that $\ell \leq t = \mathcal{O}\left(\ell^2/\nu\right)$,*

$$Pr\left[\tau_\alpha(u^*) \leq t\right] = \mathcal{O}\left(\nu t^2/\ell^{\alpha+1}\right),$$

where $\nu = \min\{\log \ell, \frac{1}{\alpha-3}\}$.

The proofs of the above lemmas are given in Sections 3.5.2 to 3.5.4, respectively. Using these lemmas can now prove our main result as follows. From Lemmas 3.5.3 and 3.5.4, by substituting $t = \Theta\left(\ell^2 \log^2 \ell\right)$, we get

$$\Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = \Omega\left(\frac{(\alpha-3)^2}{\alpha^2 \log^4 \ell}\right) - \mathcal{O}\left(\frac{\alpha}{(\alpha-3)\ell^2 \log^2 \ell}\right) = \Omega\left(\frac{(\alpha-3)^2}{\alpha^2 \log^4 \ell}\right)$$

if $3 + \omega\left(\log\log \ell / \log \ell\right) \leq \alpha$, which is Part a of Theorem 3.5.1. Also, by applying Lemma 3.5.5, we get Part b of Theorem 3.5.1.

3.5.2 Proof of Lemma 3.5.3

The proof is similar to that of Lemma 3.3.5, in Section 3.3.2. We reuse some of the notation defined there. Namely, $(L_i^f)_{i\geq 0}$ denotes the Lévy flight process, S_i is the length of the i-th jump, and $Z_u^f\left(i\right)$ is the number of visits to u in the first i steps. We also use the following modified definitions: For each node u and $i \geq 0$, we let

$$p_{u,i} = \Pr\left[L_i^f = u\right],$$

thus $\mathbb{E}\left[Z_u^f\left(i\right)\right] = \sum_{j=0}^{i} p_{u,j}$. We partition $\mathbb{Z}^2$ into sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ as follows. Let $\delta > 0$ be some value to be fixed later. Then,

$$\mathcal{A}_1 = \{v\colon \|v\|_\infty \leq \ell\}$$
$$\mathcal{A}_2 = \{v\colon \|v\|_1 \leq 4\sqrt{2(1+\delta)t\log t}\} \setminus \mathcal{A}_1$$
$$\mathcal{A}_3 = \mathbb{Z}^2 \setminus (\mathcal{A}_1 \cup \mathcal{A}_2).$$

3.5.2.1 Proof overview

First, we bound the mean number of visits to $\mathcal{A}_1$ until a given step t. We show that

$$\mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] = \mathcal{O}\left(\log^2 t\right).$$

The monotonicity property from Section 3.2.2 implies that

$$\sum_{v\in\mathcal{A}_1} \mathbb{E}\left[Z_v^f(t)\right] \leq |\mathcal{A}_1| \cdot \mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] \leq c(3-\alpha)(\ell \log t)^2/\alpha,$$

where c is a constant. To bound the mean number of visits to $\mathcal{A}_2$, as before, we use the monotonicity property again to obtain

$$\sum_{v\in\mathcal{A}_2} \mathbb{E}\left[Z_v^f(t)\right] \leq |\mathcal{A}_2| \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right] \leq 32(1+\delta)t\log^2 t \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right].$$

For the number of visits to $\mathcal{A}_3$, using a Chernoff-Hoeffding bound we show that

$$\sum_{v\in\mathcal{A}_3} \mathbb{E}\left[Z_v^f(t)\right] \leq c'\left(t^{1-(\alpha-3)/2}+1\right),$$

for some constant c' and for $\delta = \Theta\left(1/(\alpha-3)^2\right)$ large enough. Combining the above we obtain

$$c\ell^2\log^2 t + 32(1+\delta)t\log^2 t \cdot \mathbb{E}\left[Z_{u^*}^f(t)\right] + c'\left(t^{1-(\alpha-3)/2}+1\right) \geq t.$$

By choosing $t = \Theta\left(\ell^2\log^2\ell\right)$ and $\alpha - 3 = \omega\left(\log\log\ell/\log\ell\right)$, the above inequality implies

$$\mathbb{E}\left[Z_{u^*}^f(t)\right] = \Omega\left(\frac{1}{(1+\delta)\log^2\ell}\right).$$

Since

$$\Pr\left[h_f \leq t\right] = \Pr\left[Z_{u^*}^f(t) > 0\right] = \mathbb{E}\left[Z_{u^*}^f(t)\right] / \mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, Z_{u^*}^f(t) > 0\right],$$

and $\mathbb{E}\left[Z_{u^*}^f(t) \,\middle|\, Z_{u^*}^f(t) > 0\right] \leq \mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] + 1 = \mathcal{O}\left(\log^2 t\right)$, we obtain $\Pr\left[h_f \leq t\right] = \Omega\left((\alpha-3)^2/(\alpha^2\log^4\ell)\right)$.

3.5.2.2 Detailed proof

Lemma 3.5.6. *For any $t \geq 0$, $\mathbb{E}\left[Z_{\mathbf{0}}^f(t)\right] = b_t = \mathcal{O}\left(\log^2 t\right)$.*

Proof. First, we show the following. Let $L_{t'}^f$ be the two dimensional random variable representing the coordinates of the agent performing the Lévy flight at time t'. Let $L_{t'}^f = (X_{t'}, Y_{t'})$ and consider the projection of the Lévy flight on the x-axis $X_{t'}$: it can be expressed as the sum of t' random variables S_j^x, $j = 1,\ldots,t'$, representing the projection of the jumps (with sign) of the agent on the x-axis at times $j = 1,\ldots,t'$. The partial distribution of the jumps along the x-axis is given by Lemma F.1 in Appendix F.2, and states that, for any given $d \geq 1$, we have

$$\Pr\left[S_j^x = \pm d\right] = \Theta\left(\frac{1}{d^\alpha}\right).$$

Since $\mathbb{E}\left[Z_{\mathbf{0}}^{f}(t)\right] = \sum_{k=1}^{t} p_{\mathbf{0},k}$, it suffices to accurately bound the probability $p_{\mathbf{0},k}$ for each $k = 1,\ldots,t$. Let us partition the natural numbers in the following way

$$\mathbb{N} = \bigcup_{t'=1}^{\infty}\left[\mathbb{N}\cap\left[2t'\log t', 2(t'+1)\log(t'+1)\right)\right].$$

For each $k \in \mathbb{N}$, there exists t' such that $k \in [2t'\log t', 2(t'+1)\log(t'+1))$. Then, within $2t'\log t'$ steps the walk has moved to distance $\Theta\left(\sqrt{t'}\right)$ at least once, with probability $\Omega\left(\frac{1}{(t')^2}\right)$. Indeed, the sequence $\{S_j^x\}_{1\leq j\leq t'}$ consists of i.i.d. r.v.s with zero mean and finite variance $\Omega(1)$. Thus, the central limit theorem (Lemma B.5 in Appendix B) implies that the variable

$$\frac{S_1^x+\cdots+S_t^x}{\sigma\sqrt{t}}$$

converges in distribution to a standard normal random variable Z, with $\sigma = \Omega(1)$. Let $\epsilon > 0$ be a small enough constant, then there exists a $\overline{t'}$ large enough, such that for all $t' \geq \overline{t'}$ it holds that

$$\Pr\left[S_1^x+\cdots+S_{t'}^x \geq \sigma\sqrt{t'}\right] \geq \Pr[Z\geq 1] - \epsilon = \frac{c}{2} > 0,$$

for some suitable constant $c \in (0,1)$. The symmetrical results in which the normalized sum is less than $-\sigma\sqrt{t'}$ holds analogously. Thus, for all $t' \geq \overline{t'}$, we have that $\left|\sum_{j=1}^{t'} S_j^x\right| \geq \sigma\sqrt{t'}$ with constant probability $c > 0$. In $2t'\log t'$ jumps, we have $2\log t'$ sets of t' consequent i.i.d. jumps. For independence, the probability that at least in one round before round $2t'\log t'$ the Lévy flight has displacement $\Theta\left(\sqrt{t'}\right)$ from the origin is at least

$$1-(1-c)^{2\log t'} = 1-\mathcal{O}\left(\frac{1}{(t')^2}\right).$$

Once reached such a distance, there are at least $\lambda^2 = \Theta(t')$ different nodes that are at least as equally likely as $\mathbf{0}$ to be visited at any given future time for the monotonicity property (Lemma 3.2.3). Thus, the probability to reach the origin at any future time is at most $\mathcal{O}(1/t')$. Let $H_{t'}$ be the event that in any instant before time $2t'\log t'$ the Lévy flight has displacement at least $\Theta\left(\sqrt{t'}\right)$. Observe that

$$p_{\mathbf{0},k} = \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}\right]\Pr[H_{t'}] + \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}^C\right]\Pr\left[H_{t'}^C\right],$$

by the law of total probability. We remark that in an interval $[2t'\log t', 2(t'+1)\log(t'+1))$ there are

$$2(t'+1)\log(t'+1) - 2t'\log t' = 2t'\left[\log\left(1+\frac{1}{t'}\right)\right] + 2\log(t'+1) = \mathcal{O}(\log t')$$

integers. Thus, if $I_{t'} = [2t'\log t', 2(t'+1)\log(t'+1))$, we have

$$\begin{aligned}
\sum_{k=1}^{t} p_{\mathbf{0},k} &\leq \sum_{t'=1}^{t}\sum_{k\in I_{t'}} p_{\mathbf{0},k} \\
&\leq \sum_{t'=1}^{t}\left[\Pr\left[L_t^f = \mathbf{0} \mid H_{t'}\right]\Pr[H_{t'}] + \Pr\left[L_t^f = \mathbf{0} \mid H_{t'}^C\right]\Pr\left[H_{t'}^C\right]\right]\mathcal{O}(\log t') \\
&\leq \overline{t'} + \sum_{t'=\overline{t'}}^{t}\left[\mathcal{O}\left(\frac{1}{t'}\right) + \mathcal{O}\left(\frac{1}{(t')^2}\right)\right]\mathcal{O}(\log t') = \mathcal{O}\left(\log^2 t\right),
\end{aligned}$$

since $\overline{t'}$ is a constant. □

We have also the following.

Lemma 3.5.7. *For any node $u \in \mathbb{Z}^2$, it holds that*
(i) $\mathbb{E}\left[Z_u^f(t)\right] \leq b_t$;
(ii) $1 \leq \mathbb{E}\left[Z_u^f(t) \mid Z_u^f(t) > 0\right] \leq b_t$;
(iii) $\mathbb{E}\left[Z_u^f(t)\right]/b_t \leq Pr\left[Z_u^f(t) > 0\right] \leq \mathbb{E}\left[Z_u^f(t)\right]$.

Proof. The proof is exactly as that of Lemma 3.3.13. □

Thus, the total number of visits to $\mathcal{A}_1$ is upper bounded by $m_{u^*} b_t$, where $m_{u^*} = |Q_\ell(\mathbf{0})|$. Furthermore, from the monotonicity propert (Lemma 3.2.3), the following holds.

Corollary 3.5.8. *For any node u in $\mathbb{Z}^2$, we have $\mathbb{E}\left[Z_u^f(t)\right] \geq \mathbb{E}\left[Z_v^f(t)\right]$ for all $v \notin Q_{d_u}(\mathbf{0})$.*

Namely, almost all the nodes that are "further" than u from the origin are less likely to be visited at any given future time. This easily gives an upper bound on the total number of visits to $\mathcal{A}_2$ until time t, namely, by taking $u = u^*$ and by observing that each $v \in \mathcal{A}_2$ lies outside $Q_\ell(\mathbf{0})$, we get that the average number of visits to $\mathcal{A}_2$ is at most the expected number of visits to the target u^* (i.e., $\mathbb{E}\left[Z_{u^*}^f(t)\right]$) times (any upper bound of) the size of $\mathcal{A}_2$: in formula, it is upper bounded by $\mathbb{E}\left[Z_{u^*}^f(t)\right] \cdot 32(1+\delta)t\log^2 t$.

We also give a bound to the average number of visits to nodes that are further roughly $\sqrt{t}\cdot\log t$ from the origin.

Lemma 3.5.9. *A sufficiently large positive real δ exists such that $\delta = \Theta\left(1/(\alpha-3)^2\right)$ and*

$$\sum_{\substack{v\in\mathbb{Z}^2 : \\ \|v\|_1 \geq 4\sqrt{2(1+\delta)t\log t}}} \mathbb{E}\left[Z_v^f(t)\right] = \mathcal{O}\left(t^{1-\frac{\alpha-3}{2}} + 1\right).$$

Proof. Since $\alpha > 3$, the expectation and the variance of a single jump-length are finite. By Equation (3.2) in the preliminaries (Section 3.2), the probability a jump length is at least $\sqrt{t}$ is $\Theta\left(1/t^{\frac{\alpha-1}{2}}\right)$. Let us call A_j the event that the j-th jump-length is less than $\sqrt{t}$. Let us recall that L_j^f is the random variable denoting the coordinates of the nodes the corresponding Lévy flight visits at the end of the j-th jump. We can write $L_j^f = (X_j, Y_j)$, where X_j is x-coordinate of the Lévy walk after the j-th jump, and Y_j is the y-coordinate. Then, X_j can be seen as the sum $\sum_{i=1}^j S_i^x$ of j random variables representing the projections of the jumps along the x-axis. For symmetry, $\mathbb{E}[X_j] = 0$ for each j, while $\mathrm{Var}(X_j) = j\mathrm{Var}(S_1^x) = \mathcal{O}(j/(\alpha-3)+j) = \mathcal{O}(\alpha j/(\alpha-3))$ since S_1^x has finite variance $\mathcal{O}(1+1/(\alpha-3))$. This comes by observing that $S_1^x \leq S_1$. Then, conditional on $A = \cap_{i=1}^t A_i$, we can apply the Chernoff bound (Lemma B.4) on the sum of the

first j jumps, for $j \leq t$. We have

$$\begin{aligned}\Pr\left[|X_t| \geq 2\sqrt{2(1+\delta)t}\log t \mid A\right] &\leq 2\exp\left(-\frac{8(1+\delta)t\cdot\log^2 t}{\mathcal{O}\left(\frac{\alpha t}{\alpha-3}\right)+\Theta\left(\sqrt{(1+\delta)t}\cdot\log t\right)\sqrt{t}}\right)\\ &\leq 2\exp\left(-\Theta\left(\frac{\alpha-3}{\alpha}\sqrt{1+\delta}\cdot\log t\right)\right)\\ &\leq \frac{2}{t^{\Theta\left(\frac{\alpha-3}{\alpha}\sqrt{1+\delta}\right)}},\end{aligned}$$

which is less than $1/t^2$ if we choose $\delta = \Theta\left(1/(\alpha-3)^2\right)$ large enough. The same result holds for the random variable X_j for each $j < t$, since the variance of X_j is smaller than the variance of X_t. Notice that

$$\begin{aligned}\Pr\left[\cap_{j=1}^{t}\{|X_j| < 2\sqrt{2(1+\delta)t}\log t\} \mid A\right] &= 1-\Pr\left[\cup_{j=1}^{t}\{|X_t| \geq 2\sqrt{2(1+\delta)t}\log t\} \mid A\right]\\ &\geq 1-\frac{t}{t^2} = 1-\frac{1}{t},\end{aligned}$$

and that

$$\Pr[A] = 1-\Pr\left[A^C\right] = 1-\Pr\left[\cup_{j=1}^{t}A_j^C\right] \geq 1-\mathcal{O}\left(\frac{t}{t^{\frac{\alpha-1}{2}}}\right) = 1-\mathcal{O}\left(\frac{1}{t^{\frac{\alpha-3}{2}}}\right).$$

An analogous argument holds for the random variable Y_t conditioned to the event A. Then,

$$\begin{aligned}&\Pr\left[\cap_{j=1}^{t}\{\|X_j\|_1 < 2\sqrt{2(1+\delta)t}\cdot\log t\}, \cap_{j=1}^{t}\{\|Y_j\|_1 < 2\sqrt{2(1+\delta)t}\cdot\log t\}\right]\\ &\geq \Pr\left[\cap_{j=1}^{t}\{\|X_j\|_1 < 2\sqrt{2(1+\delta)t}\cdot\log t\}, \cap_{j=1}^{t}\{\|Y_j\|_1 < 2\sqrt{2(1+\delta)t}\cdot\log t\} \mid A\right]\Pr[A]\\ &\geq \left(2\Pr\left[\cap_{j=1}^{t}\{\|X_j\|_1 < 2\sqrt{2(1+\delta)t}\cdot\log t\} \mid A\right]-1\right)\Pr[A]\\ &\overset{(*)}{\geq} \left[2\left(1-\frac{1}{t}\right)-1\right]\left(1-\mathcal{O}\left(\frac{1}{t^{\frac{\alpha-3}{2}}}\right)\right)\\ &\geq 1-\mathcal{O}\left(\frac{1}{t^{\frac{\alpha-3}{2}}}+\frac{1}{t}\right),\end{aligned}$$

where $(*)$ holds for symmetry (the distribution of Y_t is the same as the one of X_t) and for the union bound. Thus, in t jumps (which take at least time t), the walk has never reached distance $4\sqrt{2(1+\delta)t}\cdot\log t$, w.h.p. The average number of visits until time t to nodes at distance at least $4\sqrt{2(1+\delta)t}\cdot\log t$ is then less than $t\cdot\mathcal{O}\left(1/t^{\frac{\alpha-3}{2}}+1/t\right) = \mathcal{O}\left(t^{1-\frac{\alpha-3}{2}}+1\right)$. □

The following puts together the previous estimations in order to get a lower bound on the average number of visits the target u^*. Let $\delta > 0$ be as given in Lemma 3.5.9 for the rest of the section.

Lemma 3.5.10. *For every node $u^* \in \mathbb{Z}^2$ and every time $t \geq 1$,*

$$m_{u^*}b_t + \mathbb{E}\left[Z_{u^*}^f(t)\right]\cdot 32(1+\delta)(t\log^2 t) + \mathcal{O}\left(t^{1-\frac{\alpha-3}{2}}+1\right) \geq t.$$

Proof. Suppose the agent has made t jumps, thus visiting t nodes. Then,

$$\mathbb{E}\left[\sum_{v\in\mathbb{Z}^2} Z_v^f(t)\right] = t.$$

We divide the plane in different zones, and we bound the number of visits over each zone in expectation. From Lemma 3.5.7, the number of visits inside $\mathcal{A}_1 = Q_\ell(\mathbf{0})$ until time t is at most $m_{u^*}b_t$, where $m_{u^*} = |Q_\ell(\mathbf{0})| = 4\ell^2$. From Lemma 3.5.9, the number of visits $\mathcal{A}_3$ is at most $\mathcal{O}\left(t^{1-\frac{\alpha-3}{2}}\right)$. Each of the remaining nodes, i.e., the nodes in $\mathcal{A}_2$, which are at most $32(1+\delta)(t\log^2 t)$ in total, is visited by the agent at most $\mathbb{E}\left[Z_u^f(t)\right]$ times, for Corollary 3.5.8. Then, we have that

$$m_{u^*}b_t + \mathbb{E}\left[Z_{u^*}^f(t)\right]\cdot 32(1+\delta)(t\log^2 t) + \mathcal{O}\left(t^{1-\frac{\alpha-3}{2}}+1\right) \geq t.$$

□

We can now complete the proof of Lemma 3.5.3 as follows. Lemma 3.5.10 implies that

$$\mathbb{E}\left[Z_{u^*}^f(t)\right] = \Omega\left(\frac{t - t^{1-\frac{\alpha-3}{2}} - 1 - m_{u^*}b_t}{(1+\delta)t\log^2 t}\right),$$

while Lemma 3.5.7 implies that

$$\Pr[h_f \leq t] = \Omega\left(\frac{t - t^{1-\frac{\alpha-3}{2}} - 1 - m_{u^*}b_t}{(1+\delta)t\log^2 t\cdot b_t}\right).$$

Lemma 3.5.6 gives $b_t = \mathcal{O}\left(\log^2 t\right)$, while *Lemma* 3.5.9 gives $\delta = \Theta\left(1/(\alpha-3)^2\right)$. If $t = \Theta\left(\ell^2\log^2\ell\right)$ is large enough and $\alpha - 3 = \omega\left(\log\log\ell/\log\ell\right)$, so that $t - t^{1-\frac{\alpha-3}{2}} - m_{u^*}b_t = \Theta(t)$, we get the result.

3.5.3 Proof of Lemma 3.5.4

If S_i is the random variable yielding the i-th jump length, then it has expectation $\Theta(1)$ and variance. This means that the sum $\bar{S}_t = \sum_{i=1}^t S_i$ has expectation $\Theta(t)$ and variance $\mathcal{O}\left(t + t/(\alpha-3)\right) = \mathcal{O}\left(\alpha t/\alpha - 3\right)$. Then, from Chebyshev's inequality,

$$\Pr\left[\bar{S}_t \geq \Theta(t) + t\right] \leq \frac{\mathrm{Var}\left(\bar{S}_t\right)}{t^2} = \mathcal{O}\left(\frac{\alpha}{(\alpha-3)t}\right).$$

Hence,

$$\Pr[h_w = \mathcal{O}(t)] \geq \Pr\left[h_f \leq t, \bar{S}_t \leq \Theta(t) + t\right] = \Pr[h_f \leq t] - \mathcal{O}\left(\frac{\alpha}{(\alpha-3)t}\right),$$

where the latter equality is obtained using the union bound.

3.5.4 Proof of Lemma 3.5.5

Let X_i be the x-coordinate of the agent at the end of the i-th jump-phase. For any $i \leq t$, we bound the probability that $X_i > \ell/4$. The probability that there is a jump whose length is at least ℓ among the first i jumps is $\mathcal{O}(i/\ell^{\alpha-1})$ for the union bound. Conditional on the event that the first i jump lengths are all smaller than ℓ (event $\mathcal{E}_i$), the expectation of X_i is zero and its variance is

$$i \cdot \sum_{d=1}^{\ell/4} \Theta\left(d^2/d^\alpha\right) = \mathcal{O}\left(i\nu\ell^{3-\alpha}\right),$$

for the integral test (Lemma E.1), where $\nu = \min\{\log \ell, \frac{1}{\alpha-3}\}$. Chebyshev's inequality implies that

$$\Pr\left[|X_i| \geq \ell/4 \mid \mathcal{E}_i\right] \leq \frac{\mathcal{O}\left(i\nu\ell^{3-\alpha}\right)}{\Theta(\ell^2)} = \mathcal{O}\left(\frac{i\nu}{\ell^{\alpha-1}}\right),$$

Since the conditional event has probability $1-\mathcal{O}(i/\ell^{\alpha-1})$, then the "unconditional" probability that of the event $|X_i| \leq \ell/4$ is

$$\left[1-\mathcal{O}\left(\frac{i}{\ell^{\alpha-1}}\right)\right]\cdot\left[1-\mathcal{O}\left(\frac{i\nu}{\ell^{\alpha-1}}\right)\right] = 1-\mathcal{O}\left(\frac{\nu t}{\ell^{\alpha-1}}\right),$$

since $i \leq t$, with t which is some function in $\mathcal{O}\left(\ell^{\alpha-1}/\nu\right)$. The same result holds analogously for Y_i (the y-coordinate of the agent after the i-th jump), obtaining that $|X_i| + |Y_i| \leq \ell/2$ with probability $1 - \mathcal{O}\left(\nu t/\ell^{\alpha-1}\right)$ by the union bound.

Consider the first jump-phase. The probability the agents visits the target during it is $\mathcal{O}(1/\ell^\alpha)$ for Corollary 3.2.2 (Section 3.2). Now, let $2 \leq i \leq t$. We want to estimate the probability the agent visits the target during the i-th jump-phase. We recall that $B_{\ell/4}(u^*)$ is the rhombus centered in u^* that contains the nodes at distance at most $\frac{\ell}{4}$ from u^*. We denote the event that the agent visits the target during the i-th jump-phase by F_i. Furthermore, let V_{i-1} be the event that the $(i-1)$-th jump ends in $B_{\ell/4}(u^*)$, and W_{i-1} the event that $(i-1)$-th jump ends at distance farther than $\ell/2$ from the origin. Then, by the law of total probabilities, we have

$$\begin{aligned}
\Pr\left[F_i\right] &= \Pr\left[F_i \mid W_{i-1}\right]\Pr\left[W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\
&= \left[\Pr\left[F_i \mid W_{i-1}, V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right]\Pr\left[V_{i-1}^C \mid W_{i-1}\right]\right]\Pr\left[W_{i-1}\right] \\
&\quad + \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\
&\overset{(*)}{\leq} \left[\Pr\left[F_i \mid V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right]\right]\Pr\left[W_{i-1}\right] + \Pr\left[F_i \mid W_{i-1}^C\right]\Pr\left[W_{i-1}^C\right] \\
&\overset{(\star)}{\leq} \left[\mathcal{O}\left(\frac{1}{\ell^2}\right) + \mathcal{O}\left(\frac{1}{\ell^\alpha}\right)\right]\mathcal{O}\left(\frac{\nu t}{\ell^{\alpha-1}}\right) + \mathcal{O}\left(\frac{1}{\ell^\alpha}\right) = \mathcal{O}\left(\frac{\nu t}{\ell^{\alpha+1}}\right) \qquad (3.11)
\end{aligned}$$

where in $(*)$ we used that $V_{i-1} \subset W_{i-1}$ and that $\Pr\left[V_{i-1}^C \mid W_{i-1}\right] \leq 1$, while in $(\star)$ we used that

$$\Pr\left[F_i \mid V_{i-1}\right]\Pr\left[V_{i-1} \mid W_{i-1}\right] = \mathcal{O}\left(\frac{1}{\ell^2}\right), \text{ (the proof is below)}$$

that $\Pr\left[F_i \mid W_{i-1}, V_{i-1}^C\right] = \mathcal{O}\left(1/\ell^\alpha\right)$ because the jump starts in a node whose distance form the target is $\Omega(\ell)$, and that $\Pr\left[F_i \mid W_{i-1}^C\right] = \mathcal{O}\left(1/\ell^\alpha\right)$ for the same reason. As for the term

$\Pr[F_i \mid V_{i-1}] \cdot \Pr[V_{i-1} \mid W_{i-1}]$ we observe the following. Let t_i be the time at the end of the i-th jump phase. Then

$$\begin{aligned}\Pr[F_i \mid V_{i-1}] \Pr[V_{i-1} \mid W_{i-1}] &= \sum_{v \in B_{\ell/4}(u^*)} \Pr\left[F_i \mid L^w_{t_i} = v\right] \Pr\left[L^w_{t_i} = v \mid W_{i-1}\right] \\ &\leq \mathcal{O}\left(\frac{1}{\ell^2}\right) \sum_{v \in B_{\ell/4}(u^*)} \Pr\left[F_i \mid L^w_{t_i} = v\right],\end{aligned}$$

since Lemma 3.2.3 holds in a consequent way conditional on W_{i-1}, and since, for each $v \in B_{\ell/4}(u^*)$, there are at least $\Theta(\ell^2)$ nodes at distance at least $\ell/2$ from the origin which are more probable to be visited than v. Then, we proceed similarly to the proof of Lemma 3.2.4 to show that $\sum_{v \in B_{\ell/4}(u^*)} \Pr\left[F_i \mid L^w_{t_i} = v\right] = \mathcal{O}(1)$, and we obtain $\Pr[F_i \mid V_{i-1}] \Pr[V_{i-1} \mid W_{i-1}] = \mathcal{O}(1/\ell^2)$.

Thus, by the union bound and by the inequality (3.11), the probability that at least during one of the t jump-phases the agent finds the target is, for some $t = \mathcal{O}(\ell^2/\nu)$,

$$\frac{1}{\ell^\alpha} + \mathcal{O}\left(\frac{\nu t^2}{\ell^{\alpha+1}}\right) = \mathcal{O}\left(\frac{\nu t^2}{\ell^{\alpha+1}}\right),$$

since $t \geq \ell$.

3.5.5 Proof of Corollary 3.5.2

From Theorem 3.5.1.Part a and the independence between agents we ge that

$$\Pr\left[\tau^k_\alpha(u^*) = \mathcal{O}\left(\ell^2 \log^2 \ell\right)\right] = 1 - \left[1 - \Omega\left(\frac{1}{\gamma \log^4 \ell}\right)\right]^k \geq 1 - e^{-\Omega\left(\frac{k}{\gamma \log^4 \ell}\right)},$$

where we have used the inequality $1 - x \leq e^{-x}$ for all x. Then, part (a) follows. Let $t = \ell^2/k \cdot \sqrt{k}/\log^2 \ell$; hence, we have $t \leq o(\ell^{\alpha-1}/\nu)$ since $k = o(\ell^2)$. If $t < \ell$, then $\Pr\left[\tau^k_\alpha(u^*) > t\right] = 1$, since ℓ steps are needed to reach distance ℓ. If $t \geq \ell$, from Theorem 3.5.1.Part b and the independence between agents, we get that

$$\Pr\left[\tau^k_\alpha(u^*) > t\right] = \left[1 - \mathcal{O}\left(\frac{\nu t^2}{\ell^{\alpha+1}}\right)\right]^k \geq \exp\left(-\mathcal{O}\left(\frac{k t^2 \log \ell}{\ell^{\alpha+1}}\right)\right),$$

where we have used the inequality $1 - x \geq e^{-\frac{x}{1-x}}$ for $x < 1$, that $\nu t^2/\ell^{\alpha+1} = o(1)$, and that $\nu \leq \log \ell$. Then, by substituting $t = \ell^2/k \cdot \sqrt{k}/\log^2 \ell$, and by the Taylor's expansion of the exponential function, we get

$$\exp\left(-\mathcal{O}\left(\frac{k t^2 \log \ell}{\ell^{\alpha+1}}\right)\right) = 1 - \mathcal{O}\left(\frac{1}{\log^3 \ell}\right).$$

Corollary 3.5.2.Part b follows by observing that $\ell^2/k \cdot \sqrt{k}/\log^2 \ell \leq \ell^2/\sqrt{k}$.

3.6 Distributed search algorithm

In this section, we prove the following theorem, which provides a simple distributed search algorithm, which allows k agents to find an arbitrary, unknown target on $\mathbb{Z}^2$ in optimal time (modulo polylogarithmic factors).

Theorem 3.6.1. *Consider k independent Lévy walks that start simultaneously from the origin, and the exponent of each walk is sampled independently and uniformly at random from the real interval $(2,3)$. Let $\tau^k_{rand}(u^*)$ be the parallel hitting time for a given target u^*. If $k \geq \log^8 \ell$, and $\ell = \|u^*\|_1$ is large enough, then*

$$Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left((\ell^2/k)\cdot \log^7 \ell + \ell \log^3 \ell\right)\right] = 1 - e^{-\omega(\log \ell)}.$$

We need the next lemma, which is a slight generalization of Corollary 3.3.2.a that bounds the hitting time of a collection of Lévy walks with different exponent values.

Lemma 3.6.2. *Consider k independent Lévy walks that start simultaneously from the origin, and the exponent of each walk is in $[\alpha_1, \alpha_2]$. Let h_{diff} be the parallel hitting time for a target u^* with $\|u^*\|_1 = \ell$. If $2 < \alpha_1 \leq \alpha_2 \leq 3 - \epsilon$ and $\epsilon = \omega(1/\log \ell)$, then*

$$Pr\left[h_{diff} = \mathcal{O}\left(\frac{\ell^{\alpha_2 - 1}}{\alpha_1 - 2}\right)\right] = 1 - e^{-\Omega\left(\frac{(3-\alpha_2)^2 k}{\ell^{3-\alpha_1}\log^2 \ell}\right)}.$$

Proof. First, we recall that the k agents move independently from each other. Let $\tau_\alpha(u^*)$ be the hitting time of a single walk. If $\alpha \in [\alpha_1, \alpha_2]$, then from Lemmas 3.3.5 and 3.3.6,

$$\Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha-1}}{\alpha - 2}\right)\right] = \Omega\left(\frac{(3-\alpha)^2}{\ell^{3-\alpha}\log^{\frac{2}{\alpha-1}} \ell}\right) = \Omega\left(\frac{(3-\alpha_2)^2}{\ell^{3-\alpha_1}\log^2 \ell}\right),$$

provided that $3 - \alpha_2 = \omega(1/\log \ell)$. Observe that $\Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha_2-1}}{\alpha_1-2}\right)\right] \geq \Pr\left[\tau_\alpha(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha-1}}{\alpha-2}\right)\right]$. Then,

$$\Pr\left[h_{diff} = \mathcal{O}\left(\frac{\ell^{\alpha_2-1}}{\alpha_1-2}\right)\right] = 1 - \left(1 - \Omega\left(\frac{(3-\alpha_2)^2}{\ell^{3-\alpha_1}\log^2 \ell}\right)\right)^k \leq 1 - e^{-\Omega\left(\frac{(3-\alpha_2)^2 k}{\ell^{3-\alpha_1}\log^2 \ell}\right)}.$$

□

We can now prove our main result.

Proof of Theorem 3.6.1. Fix k, ℓ such that $k \geq \log^8 \ell$, and let $\epsilon = \log\log \ell / \log \ell$. Let $\alpha \in [2+\epsilon, 3-2\epsilon]$, and let k_α be the number of Lévy walks whose exponent is in the interval $[\alpha, \alpha+\epsilon]$. Then $\mathbb{E}[k_\alpha] = \epsilon k$, and by the Chernoff bound (Lemma B.1),

$$\Pr[k_\alpha \geq \epsilon k/2] = 1 - e^{-\Omega(\epsilon k)}.$$

Clearly, the parallel hitting time of the k Lévy walks is upper bounded by the parallel hitting time of the k_α Lévy walks whose exponent is in $[\alpha, \alpha+\epsilon]$. Then, from Lemma 3.6.2, it follows that

$$\Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha+\epsilon-1}}{\alpha-2}\right) \,\middle|\, k_\alpha \geq \epsilon k/2\right] = 1 - e^{-\Omega\left(\frac{(3-\alpha-\epsilon)^2 \epsilon k}{\ell^{3-\alpha}\log^2 \ell}\right)}.$$

Combining the last two equations, we obtain

$$\Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha+\epsilon-1}}{\alpha-2}\right)\right] = 1 - e^{-\Omega\left(\frac{(3-\alpha-\epsilon)^2\epsilon k}{\ell^{3-\alpha}\log^2\ell}\right)}. \tag{3.12}$$

We distinguish the following two cases.

Case $\log^7\ell \le k \le \ell\log^3\ell$**.** Choose some $\alpha \in [2+\epsilon, 3-2\epsilon]$ such that

$$k = \frac{\ell^{3-\alpha}\cdot\log^2\ell}{(3-\alpha-\epsilon)^2\cdot\epsilon}\cdot\log\ell\cdot\log\log\ell.$$

Such an α exists because the values of function $f(\alpha) = \frac{\ell^{3-\alpha}\cdot\log^3\ell\cdot\log\log\ell}{(3-\alpha-\epsilon)^2\cdot\epsilon}$ at the extreme points of α are $f(2+\epsilon) \ge \ell\log^3\ell \ge k$ and $f(3-2\epsilon) \le \log^8\ell \le k$. Substituting the above value of α to Eq. (3.12), we obtain

$$\Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left(\frac{\ell^{\alpha+\epsilon-1}}{\alpha-2}\right)\right] = 1 - e^{-\omega(\log\ell)}.$$

Thus, with probability $1 - e^{-\omega(\log\ell)}$,

$$k\cdot\tau^k_{rand}(u^*) = k\cdot\mathcal{O}\left(\frac{\ell^{\alpha+\epsilon-1}}{\alpha-2}\right) = \mathcal{O}\left(\frac{\ell^2\cdot\log^5\ell}{(3-\alpha-\epsilon)^2(\alpha-2)}\right) = o\left(\ell^2\cdot\log^7\ell\right).$$

Case $k \ge \ell\log^3\ell$**.** In this case, we set $\alpha = 2+\epsilon$ and substitute this value of α to Eq. (3.12) to obtain $\Pr\left[\tau^k_{rand}(u^*) = \mathcal{O}\left(\ell^{1+2\epsilon}/\epsilon\right)\right] = 1 - e^{-\Omega\left(\frac{k\log\log\ell}{\ell\log^2\ell}\right)} = 1 - e^{-\omega(\log\ell)}$. Thus, with probability $1 - e^{-\omega(\log\ell)}$,

$$\tau^k_{rand}(u^*) = \mathcal{O}\left(\ell^{1+2\epsilon}/\epsilon\right) = o\left(\ell\cdot\log^3\ell\right).$$

Combining the two cases completes the proof. □

3.7 Discussion and future work

Our setting, that is also a specific instance of the ANTS problem, serves as an abstract and simplified model of cooperative foraging, where a group of k walkers (that move approximately independently between one another) is searching for some food location around a nest site, a food storage area, or a sheltered environment. If the typical or maximum distance ℓ of the food (target) from the nest (source) is fixed and known by the walkers executing parallel Lévy walks with the same exponent, the latters can optimize search efficiency by tuning the exponent value and/or the number k of animals participating in the foraging. Naturally, this would imply (partial) knowledge of ℓ. In that setting, no universally optimal exponent value exists, as the optimal exponent depends on k and ℓ. As an alternative, novel approach suggested by our last result, that does not require any knowledge of ℓ, k, nor communication, is that each animal performs a Lévy walk with a randomly chosen exponent. As discussed in Section 1.2.2, this strategy, which surprisingly achieves near optimal search efficiency for all distance scales, implies that different members of the same group

follow different search patterns, and offers an almost-optimal uniform solution (within a $\operatorname{polylog} \ell$ factor) to the ANTS problem that is extremely simple and natural. Our work mathematically corroborates the Lévy flight foraging hypothesis by showing that Lévy walks are surprisingly efficient movement models even in the parallel search.

What other studies that may further establish the optimal performances of Lévy walks deserve to be addressed? What follows is an extremely partial answer. An obvious generalization of our work to the n-dimensional space arises: we expect a generally different behavior. With respect to this, notice that the simple random walk also undergoes changes in behavior as it becomes transient when $n \geq 3$. Moreover, we have the intuition that the properties of the super-diffusive regime, i.e., when $\alpha \in (2, 3)$ in our formalization, can be further exploited in other distributed problems: information dissemination in a thorus is an important task in distributed computing and it has been studied extensively. While the behavior of simple random walks that spread information is well known both in the two-dimensional case [Pettarin et al., 2011] and in the three-dimensional one [Lam et al., 2012], that of power-law tail random walks is still unexplored. As for the latter, we believe that for some density regimes (number of agents / size of the thorus), Lévy walks can be of strong improvements.

CHAPTER 4

Planning with Biological Neurons and Synapses

In this chapter we consider the Assembly Calculus and we test experimentally its capabilities. In particular, we empirically demonstrate that complex programs such as heuristics for the planning task in the Blocks World run correctly and reliably.

4.1 Introduction

In this chapter we investigate the capabilities of the Assembly Calculus (AC) [Papadimitriou et al., 2020], a recently proposed computational system for modeling a dynamical, distributed system of firing neurons following biological principles. Our contribution in this thesis is to demonstrate that a program in the AC is capable of implementing reasonably sophisticated stylized planning strategies – in particular, heuristics for solving tasks in the blocks world [Slaney and Thiébaux, 2001]; in particular, we code the planning heuristics in the AC framework in the Julia programming language [Bezanson et al., 2017]. The exposed results are based on the work [D'Amore et al., 2022b].

Roadmap. In the following, we briefly describe the content of this chapter. In Section 4.2 we define the Assembly Calculus (Section 4.2.1) and the Blocks World and the planning task (Section 4.2.2). Afterwards, in Sections 4.3 and 4.4 we describe the AC programs we designed to implement the planning heuristics, and discuss the experiments we ran, respectively. In Section 4.5 we summarize our results and open to further research directions that can strengthen the AC theory and widen its application domain. Finally, in Section 4.6 we give some more technical details about our experiments and their reproducibility, and further discuss the limits of the Assembly calculus.

4.2 Preliminaries

4.2.1 Crash course in the Assembly Calculus

In this section we formally introduce the Assembly Calculus framework, as it is described in [Papadimitriou et al., 2020]. We start recalling what an Erdös-Rényi graph is.

Definition 4.2.1 (Directed Erdös-Rényi graph)**.** Let $n \in \mathbb{N}$ and $p \in (0,1)$. A directed graph $G = (V, E)$ is constructed according to the Erdös-Rényi model if $|V| = n$ and for any $x, y \in V$, $(x, y) \in E$ with probability p. Such a graph is denoted by $G_{n,p}$.

We will equivalently denote any edge (x, y) of the graph by xy. We remark that we are making an abuse of notation for the sake of simplicity as $G_{n,p}$ represented the undirected Erdös-Rényi graph in Definition 2.7.1 (Chapter 2). The following is the definition of brain area.

Definition 4.2.2 (Brain area)**.** Given any $n_A \in \mathbb{N}$ and $p_A \in (0,1)$, a brain area A with n_A neurons and connection probability p_A is a directed Erdös-Rényi G_{n_A,p_A}.

Here, p_A is the probability that two nodes/neurons of area A are connected. We are now ready to define the brain.

Definition 4.2.3 (Brain)**.** Let $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$ be a finite set of disjoint brain areas, and let $\mathcal{C} \subseteq \mathcal{A} \times \mathcal{A}$ be any set of different pairs of brain areas. Furthermore, for each $(A_i, A_j) \in \mathcal{C}$, let $p_{A_i,A_j} \in (0,1)$. Finally, let $f : V \times \mathbb{N} \to \{0,1\}$ and $w : E \times \mathbb{N} \to \mathbb{R}_+$ be two functions. The brain is a graph $\mathcal{B} = (V, E, w, f)$ where $V = \sqcup_{i=1}^{m} V(A_i)$ and [1] $E = (\sqcup_{i=1}^{m} E(A_i)) \sqcup E^\star$, where $E^\star$ is characterized as follows:

$$\forall (A_i, A_j) \in \mathcal{C}, \forall x \in A_i, \forall y \in A_j, (x, y) \in E^\star$$

1. $\sqcup$ is the disjoint union.

with probability p_{A_i,A_j}. The functions w and f are called, respectively, the edge weight function and the neuron activation function.

We refer to brain edges as *fibers* and to brain nodes as *neurons*. In order to describe the AC dynamics, we must specify how the neuron activation function and the edge weight function are defined: we accomplish this recursively. We first remark that for each neuron x and any time t, $f(x,t) = 1$ means that neuron x is firing (or, equivalently, active) at time t, and that for each edge $e = xy$, $w(e,t)$ yields the weight of e at time t.

For each area $A \in \mathcal{A}$, let $k_A \in \mathbb{N}$ be an integer called the $k - CAP$. At time $t = 0$, all weights are initialized to be 1. At any time t, each neuron x receives an input $I(x,t)$ that can either be provided by the external environment, or depend on the neurons that are firing (active) at time $t-1$. In the latter case, the input for a neuron is the quantity

$$I(x,t) = \sum_{y \in \mathcal{N}(x)} w(yx, t-1) \cdot f(y, t-1).$$

A neuron receiving an input is said to be *excited*. In any area A, a neuron x fires according to the k-Winners Take All rule (k-WTA), that is, if and only if it belongs to the set of the k_A neurons with the highest input. This phenomenon tries to capture what happens in the brain though populations of inhibitory neurons [Papadimitriou et al., 2020], but WTA dynamics that tackle this problem in details are still missing; a further discussion on works related to this dynamics is given in Section 4.5. Hence, we only postulate this aspect by assuming exclusively its final effect. Formally, if F_A^t is the set of the k_A neurons with the highest input at time t in area A,[2]

$$f(x,t) = \begin{cases} 1 & \text{if } x \in F_A^t, \\ 0 & \text{otherwise.} \end{cases}$$

The behavior of the weights at time $t \geq 1$ depends on the neurons that fired at time $t-1$, as the weight update rule follows Hebbian learning: for each area $A \in \mathcal{A}$, upon the firing of both endpoints in consequent rounds, the weight of an edge xy is multiplied by a factor $(1+\beta)$ where $\beta = \beta_A > 0$ is a constant which is specific for area A. If, instead, the edge connects different areas, say A_i and A_j, then the Hebbian constant depends on the ordered pair, i.e. $\beta = \beta_{A_i,A_j}$. Formally, for each edge $(x,y) \in E$,

$$w(xy,t) = [w(xy,t-1)(1+\beta)] \cdot [f(y,t)f(x,t-1)] + w(xy,t-1) \cdot [1 - f(y,t)f(x,t-1)]\,.$$

As already mentioned in the introduction (Section 1.3), in the original definition of the AC, a process of *homeostasis* was also modelled through a periodic renormalization, at a different time scale, of the synaptic weights, in order to avoid the generation of huge weights. Such process is of course part of any realistic brain system, also providing a mechanism for *forgetting*. We will not implement here this feature of the model for simplicity, but its presence would not affect our results.

Furthermore, an area of excitatory neurons can be *inhibited*, that is, its neurons cannot fire. This aspect can also be accomplished trough populations of inhibitory neurons that send inhibitory inputs to other neurons preventing them from firing. Similarly, an inhibited area can be *disinhibited*, i.e. its neurons can now fire again; disinhibition can be realized by a second population of inhibitory neurons that inhibits the first one. We also assume the same operations

2. Ties are broken uniformly at random.

can be carried out on fibers, in accordance with [Mitropolsky et al., 2021]. For a set of areas $\mathcal{A}$, we write $\texttt{inhibitArea}\,(\mathcal{A})$ to denote the operation of inhibiting all areas in $\mathcal{A}$; similarly, by $\texttt{disinhibitArea}\,(\mathcal{A})$ we denote the operation f disinhibiting all areas in $\mathcal{A}$. If, instead, $\mathcal{F}$ is a set of pairs of areas, $\texttt{inhibitFiber}\,(\mathcal{F})$ inhibits all fibers between any area pair in $\mathcal{F}$, while $\texttt{disinhibitFiber}\,(\mathcal{F})$ disinhibits all fibers between the pairs in $\mathcal{F}$. Here, (dis)inhibition is always determined by which areas and fibers fired at the previous time step. Such operators are only stylized functions that are postulated in the model which capture some aspects of real neuronal (dis)inhibition mechanisms; obviously, the use we make of them is a strong simplification, even though it is justified in other works [Dabagia et al., 2022]. As we already discussed for the projection operation in the introduction (Section 1.3), the (dis)inhibition operation is an emergent behavior of the distributed system as a direct or indirect response to external stimuli, induced by specific neurons firing; in practice and for simplicity, it is a function we call while coding.

In this framework, we are ready to give the definition of an assembly.

Definition 4.2.4 (Assembly)**.** Given any disinhibited area A of a brain $\mathcal{B}$, an assembly F_A is a stable set of k_A neurons in A, where stable means that the firing of these k_A neurons results deterministically in the firing of the same set of neurons at the next round when no external inputs are provided.

It is easy to verify that when a set of neurons is stable for two consecutive rounds according to Definition 4.2.4, it will be stable in any other future round unless the network state is modified. In the AC framework, we improperly call the stability of an assembly the ratio measuring the overlap of two consecutives sets of neurons firing: while a set of neurons that is highly stable is not formally an assembly w.r.t. Definition 4.2.4, the AC programs keep running realiably as long as the overlap is high, and it is also realistic to suppose that a real brain assembly isn't actually a fixed set of neurons, but a highly stable set of neurons.

Usually, we denote the set of k_A neurons an assembly consists of by $\mathbf{x}$. An *active* assembly is an assembly that has just fired, i.e. the set of neurons with the highest inputs in the area. We remark that an active assembly is unique in a given area. Consequently, the *activation* of an assembly corresponds to its firing. This operation is denoted by $\texttt{activate}\,(\mathbf{x})$. We remark that we assume the neurons keep track of their firing state when they're inhibited, which implies that an assembly that was active before inhibition is still active after disinhibition.

Intuitively, an assembly represents a particular concept or idea, and can be manipulated by strengthening the connections between existing assemblies (association), or by creating new assemblies as a result of a projection of existing assemblies. For the purposes of our work, we consider only the *projection* operation. By projection of an assembly $\mathbf{x}$, we mean the creation of an assembly $\mathbf{y}$ in a downstream area that can be thought of as a "copy" of $\mathbf{x}$, and such that $\mathbf{y}$ will henceforth fire every time x fires. We report the description of this operation as it is given in [Papadimitriou et al., 2020].

Projection operation. Consider two distinct disinhibited brain areas A and B such that A is connected to B. Let $\mathbf{x}$ be an active assembly in A. The operation $\texttt{project}\,(A, \mathbf{x}, B)$ entails activating, repeatedly, assembly $\mathbf{x}$ while B is disinhibited. Such repeated activation creates, in the disinhibited area B, a sequence of sets of k_B cells, let us call them $\mathbf{y_1}, \mathbf{y}_2, \ldots, \mathbf{y}_t, \ldots$. The mathematical details are quite involved, but the intuition is the following: cells in B can be thought of as competing for synaptic input. At the first step, only $\mathbf{x}$ provides synaptic input, and thus $\mathbf{y_1}$ consists of the k_B cells in B which happen to have the highest sum of synaptic weights originating

in $\mathbf{x}$; note that these weights are subsequently increased by a factor of $(1+\beta)$ due to plasticity. At the second step, neurons in both $\mathbf{x}$ and $\mathbf{y}_1$ spike, and, as a result, a new set $\mathbf{y}_2$ of "winners" from among cells of B is selected; for typical parameters, $\mathbf{y}_1$ overlaps heavily with $\mathbf{y}_2$. This continues as long as $\mathbf{x}$ keeps firing, with certain cells in $\mathbf{y}_t$ replaced by either "new winners" cells that never participated in a $\mathbf{y}_{t'}$ with $t' < t$, or by "old winners" cells that did participate in some $\mathbf{y}_{\mathbf{t}'}$ with $t' < t$. We say that the process has converged when there are no new winners. For a large range of parameters and for high enough plasticity, this process can be proved to converge exponentially fast, with high probability, to create an assembly $\mathbf{y}$, the result of the projection [Legenstein et al., 2018].

Remark 4.2.1 – Upon further firing of $\mathbf{x}$, $\mathbf{y}_t$ may evolve further slowly, or cycle periodically, with past winners coming in and out of $\mathbf{y}_t$; in fact, this mode of assembly firing (cells of the assembly alternating in firing) is very much in accordance with how assemblies have been observed to fire in Ca+ imaging experiments in mice; see, for example, [Carrillo-Reid et al., 2019].

It is theoretically possible that new winner cells may come up after convergence; but it can be proved that this is an unlikely event [Papadimitriou et al., 2020], and in simulations the overlap of sets of neurons firing consecutively in the projection operation is usually very large. The number of steps required for convergence depends on the parameters, but, most crucially, on the plasticity coefficient β. From now on, y will fire every time x fires.

An extension of the projection operation we consider is the *strong projection* operation, as it is introduced in [Mitropolsky et al., 2021].

Strong projection operation. Consider all disinhibited areas of the dynamical system, and all disinhibited fibers in and between them. This defines an undirected graph (which in our usage will always be a tree). Call a disinhibited area active if it contains an active assembly. Now, suppose that all these assemblies fire simultaneously, into every other disinhibited adjacent area through every disinhibited fiber, and these areas fire in turn, possibly creating new assemblies and firing further down the graph, until the process stabilizes (that is, the same neurons keep firing from one step to the next). We denote this systemwide operation by `strongProject()`. Note that `strongProject()` is almost syntactic sugar, as it simply abbreviates a sequence of projections (which can be done in the AC model); however, the notion of an active area is a small addition to the AC. Though this modification is minor, it simplifies our implementation, but it could be removed at the expense of more AC brain areas and perhaps time steps.

During the strong projection operation, it is possible that two assemblies in adjacent areas send synaptic inputs between the two areas. As a result, we may have three situations: the two assemblies can just strengthen the connections between them without being modified; the two assemblies are actually slightly modified by this process but stabilize quickly; new assemblies arise. Experimentally, under a wide range of parameters, we lie in the first two cases (mostly, the first one). This results in the *association* of assemblies, where the firing of one results in the firing of the other.

Finally, we describe the readout mechanism that we postulate in the AC, following [Papadimitriou et al., 2020].

Readout mechanism. In [Buzsáki, 2010], the author proposes that, for assemblies to be functionally useful, readout mechanisms must exist that sense the current state of the assembly system

and trigger appropriate further action. We thus introduce another operation; for any given area A, `isAssembly` (A) returns `true` if A contains an active assembly, and `false` otherwise. In the code, this is accomplished making the set of k_A neurons with the highest input fire and checking whether, at the next round, the same set (or a set with a very large overlap with the former one) is firing in turn.

On the convergence of the AC operations. As mentioned in 1.3.2, the AC makes possible to perform such *operations* with assemblies: assemblies are created, in a way that guarantees high connectivity, through such operations. In [Legenstein et al., 2018, Papadimitriou and Vempala, 2019, Papadimitriou et al., 2020], the authors give theoretical guarantees that some operations are "possible" in the sense that they can be stably performed with high probability in the dynamical system of neurons outlined in the previous paragraphs, at least under some restricted range of parameters, and they also demonstrate the same through simulations. We here state the theoretical guarantees of convergence of the projection operation, which are (arguably) the most interesting.

Theorem 4.2.1 (Projection convergence)**.** *Let A and B two brain areas with the same number of neurons n, the same connection probability p, the same k-CAP, and the same plasticity β. Furthermore, suppose A has outgoing fibers towards B like a one-way bipartite Erdös-Rényi graph with probability p. Consider any assembly in A that fires repeatedly in B due to some continuous external stimuli. Let $\beta^\star = \sqrt{\frac{\sqrt{2}-1}{1+\sqrt{\frac{pk}{\ln n}}}}$. Then, with high probability, the process converges*[3] *and the total number of neurons in B activated during the entire process is bounded by*

(i) $k + o(k)$ *if* $\beta > \beta^\star$*;*

(ii) $k \cdot \left(\frac{n}{k}\right)^{\frac{1}{4\beta}}$ *if* $\beta < \beta^\star$*.*

Moreover, for $\beta > \beta^\star$, any future activation of the upstream assembly leads to the activation of all but $o(k)$ of the previously activated neurons in B with high probability.

Experimentally, the convergence speed is very fast. E.g., when $n = 10^6, k = 50, \beta = 0.1, p = 0.1$, we find that an assembly is formed with a projection operation after ~ 35 spikes (firings) (Section 4.4). We remark that in the more recent work [Dabagia et al., 2022], generalizations of this theorem are provided in that the projection convergence is analyzed and proved for non-fixed upstream assemblies, i.e., the source that is projecting comes from a distribution over the neurons and changes over time.

Wrap-up. Our AC programs are described with the operations in defined in the previous paragraphs. Inhibition and disinhibition are primitives of the AC system, whereas strong projection (tantamount to a set of simultaneous projections) is an emergent property of the AC's dynamical system. We use other such "emergent" operations, i.e., the readout mechanism `isAssembly` (), that are not primitives of the AC system, but can be stably implemented with its basic operations. In Table 4.1, we summarize the operations (primitive and non primitive) of the AC system, that we will use in our work. Note that the assembly activation operation is a special operation, which causes an assembly in a special area fire. Such an area is said to be *explicit*, and is responsible for the interaction with the external environment: e.g., when the subject is exposed to some idea, concept, or sight, a unique corresponding assembly in an explicit area fires.

3. We remind that a projection is said to converge if there is a time t such that the sets of neurons firing in response in the downstream area at time t and $t + 1$ are the same.

Operation	Input	Semantics
`activate` $(\mathbf{x})$	Assembly $\mathbf{x}$	Makes the assembly $\mathbf{x}$ fire in an explicit area
`disinhibitArea` $(\mathcal{A})$	Set $\mathcal{A}$ of areas	Disinhibit all the areas in $\mathcal{A}$
`disinhibitFiber` $(\mathcal{F})$	Set $\mathcal{F}$ of pairs of areas	Disinhibit the fibers between any pair of areas in $\mathcal{F}$
`inhibitArea` $(\mathcal{A})$	Set $\mathcal{A}$ of areas	Inhibit all the areas in $\mathcal{A}$
`inhibitFiber` $(\mathcal{F})$	Set $\mathcal{F}$ of pairs of areas	Inhibit the fibers between any pair of areas in $\mathcal{F}$
`isAssembly` (A)	Area A	Verify whether there is an active assembly in the area A
`project` $(A, \mathbf{x}, B)$	Areas A and B	Executes a projection of (the active assembly in) the area A to the area B
`strongProject`$()$		Executes a strong projection involving all the disinhibited areas and fibers

TABLE 4.1 – The AC operations (primitive and non primitive) used in our work.

For a complete description of the AC the reader is referred to [Papadimitriou et al., 2020], where in addition to stability of various assembly operations, it is also proved that, under certain assumptions, this computational system is capable of performing arbitrary computations as long as the space required does not exceed $\frac{n}{k}$ (under much milder assumptions, $\sqrt{\frac{n}{k}}$), assuming k to be the same k-CAP for all brain areas.

4.2.2 The Blocks World

A blocks world (BW) configuration C is a set of *stacks* $C = \{S_1, \ldots, S_m\}$, where each stack S_i is a sequence of *blocks* $S_i = (b_1^{(i)}, \ldots, b_{k_i}^{(i)})$, from top to bottom [Slaney and Thiébaux, 2001]. Each block $b_j^{(i)}$ is assumed to be a unique integer between 1 and s (see Fig. 4.1).

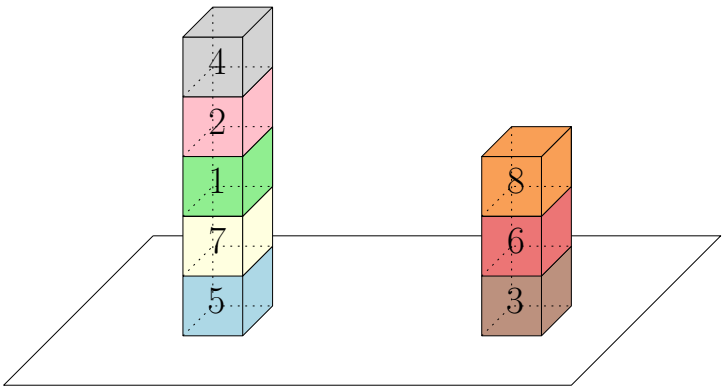


Figure 4.1 – Example of a BW configuration with 8 blocks. Such a configuration is denoted as $\{(4, 3, 1, 7, 5), (8, 6, 3)\}$.

Planning task. In the planning task [Gupta and Nau, 1991, Gupta and Nau, 1992], we are given two blocks world configurations, an initial one C_{init} and a target one C_{target} constituting in exactly the same blocks (Fig. 4.2). Three possible actions are allowed: moving a block from the top of a stack to the table (creating a new stack); moving a block from the top of a stack to the top of another stack; moving a block from the table to the top of a stack. The goal is to output the sequence of moves that transforms the C_{init} into C_{target}. Among the many possible strategies that solve this problem, the simplest one is to "demolish" all stacks by putting all blocks on the table, and then build the final configuration from scratch.

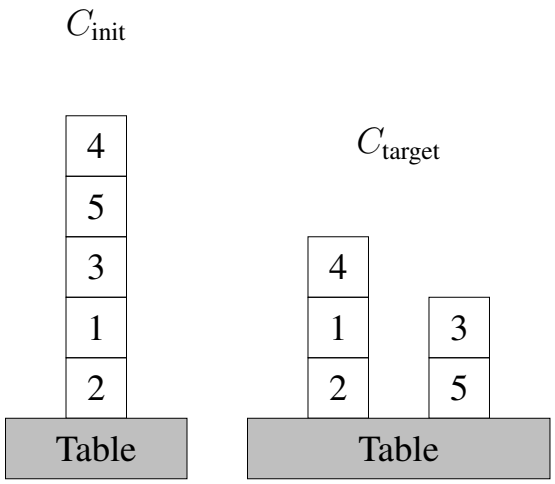


Figure 4.2 – Planning task input example.

In general, the problem of computing the minimum number of moves is NP-hard, but a very simple 2-approximation algorithm exists [Gupta and Nau, 1992] and can be summarized as follows: *move all blocks that are not in their final position to the table, and then place them on their proper stack one by one.*

The blocks world AC framework. All programs we write for the planning task work on a common set of brain areas connected with bi-directional fibers; equivalently, we assume all Erdös-Rényi graphs to be undirected. This assumption simplifies the implementation and it is also carried on by previous works [Papadimitriou et al., 2020, Mitropolsky et al., 2021]. We work with one single explicit area, BLOCKS, which already contains a fixed assembly for every possible block. As we said, this type of area is special, as an assembly can be activated explicitly with the presentation of the corresponding block (a number) in input. For a block b and its corresponding assembly $\mathbf{x}_b$, we write $\texttt{activateBlock}\,(b) = \texttt{activate}\,(\mathbf{x}_b)$. There are four other areas used in our AC programs: HEAD, NODE_0, NODE_1, and NODE_2. HEAD is connected to the NODE_0 area via fibers, while each NODE area is connected to BLOCKS, and to each other in the shape of a triangle: NODE_0 is connected with NODE_1, which is connected with NODE_2, which is connected with NODE_0 (see Fig. 4.3). All of these areas are standard brain areas of the AC system, containing the same number of neurons n, of which at most k fire at any time, where k is the same among all areas. Also, we make the assumptions that the plasticity β and the connection probability p is the same for the whole brain to make the simulation coding easier. Such simplifications are assumed also in the related works, such as [Papadimitriou et al., 2020, Mitropolsky et al., 2021, Dabagia et al., 2022]; we remark that it is possible to work without these assumptions, but it would be more challenging from a coding point of view and would burden the notation.

4.2.3 Our contribution

Our contribution in this work is to demonstrate that a program in the AC framework is capable of implementing reasonably sophisticated stylized planning strategies — in particular, the two aforementioned heuristics for solving tasks in the blocks world — and of running in reasonable time. Furthermore, we empirically establish the functioning of *chaining*, that is, the encoding of

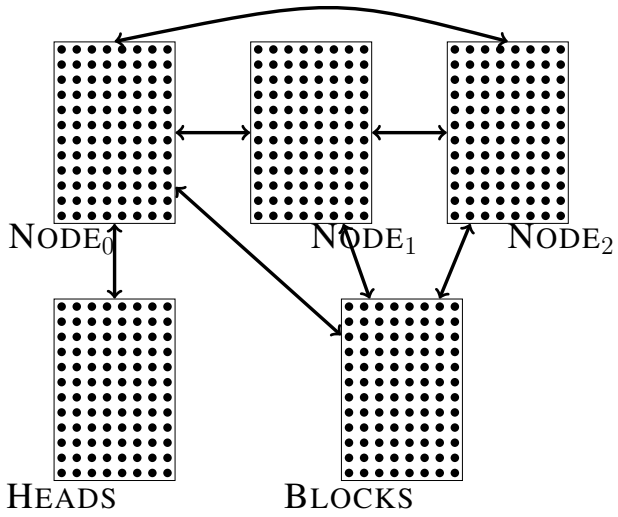


Figure 4.3 – The five main areas used by our programs, along with the connections through fibers.

an arbitrarily large number of non-overlapping assemblies in a constant number of areas; the latter phenomenon was only hypothesized to work in [Mitropolsky et al., 2021], but left unexplored. Further details are given in Section 4.3.1.

Roadmap. We shall at first concentrate on *configurations with a single stack* — already a meaningful problem — and we shall eventually graduate to multiple stacks (see Section 4.3.5). We next describe four AC programs: (a) a program that takes the input — a sequence of integers representing a stack — and creates a list-like structure, in a set of brain areas and fibers, for representing the stack (Section 4.3.1); (b) a program that removes the top block of a stack thus represented (Section 4.3.2); (c) a program that adds a new block to the represented stack (Section 4.3.3); and (d) a program for computing the *intersection* of two stacks represented this way, that is, the longest common suffix of the two sequences, read from bottom to top (Section 4.3.4).

4.3 The blocks world AC programs

In the following sections, we describe our programs and show example images for the initial configuration in 4.1.

4.3.1 The Parser

The parser (see Algorithm 1) processes each block in a stack sequentially, starting from the top. When it analyses the first block (see lines 2-4), the three areas BLOCKS, HEAD, and NODE$_0$, and the fibers between HEAD and NODE$_0$ and between NODE$_0$ and BLOCKS are disinhibited. The block assembly is then activated and a strong projection is performed, thus creating a connection between the assembly in BLOCKS corresponding to the block and an assembly in NODE$_0$, and between this latter assembly and an assembly in HEAD (see the red dashed lines in Fig. 4.4-C1). Successively, the HEAD area and the fibers between HEAD and NODE$_0$ and between NODE$_0$ and BLOCKS are inhibited. For each other block in the stack (see lines 6-11), the NODE area next

Algorithm 1: PARSER (S)

input: a stack S of blocks $b_1, b_2, \ldots, b_s$.

```
1  disinhibitArea({BLOCKS, HEAD, NODE_0});
2  disinhibitFiber({(HEAD, NODE_0), (NODE_0, BLOCKS)});
3  activateBlock(b_1); strongProject();
4  inhibitArea({HEAD}); inhibitFiber({(HEAD, NODE_0), (NODE_0, BLOCKS)});
5  foreach i with 2 ≤ i ≤ s do
6      p = (i − 2) mod 3; c = (i − 1) mod 3;
7      disinhibitArea({NODE_c});
8      disinhibitFiber({(NODE_p, NODE_c), (NODE_c, BLOCKS)});
9      activateBlock(b_i); strongProject();
10     inhibitArea({NODE_p});
11     inhibitFiber({(NODE_p, NODE_c), (NODE_c, BLOCKS)});
12 end
13 inhibitArea({BLOCKS, NODE_((s−1) mod 3)});
```

to the one (i.e., $\text{NODE}_{i \bmod 3}$) currently disinhibited (i.e., $\text{NODE}_{i+1 \bmod 3}$) is disinhibited, and the fibers between this NODE area and the BLOCKS area and between the two NODE areas are disinhibited. The next block assembly is then activated and a strong projection is performed, creating a connection between the assembly in BLOCKS and an assembly in the NODE area just disinhibited, and between this latter assembly and the assembly previously activated in the previous NODE area (see the red dashed lines in the Fig. 4.4-C2,C3,C4). After this and before the next block, this latter NODE area and the fibers between it and the NODE area after it, and those between the NODE area after it and the BLOCKS area, are inhibited.

The final data structure is a *chain* of assemblies starting from an assembly in HEAD and passing through assemblies in the NODE areas (see Fig. 4.4-C6). Note that this chain can contain more than one assembly in the same NODE area: for instance, in Fig. 4.4-C6, the chain contains two assemblies in NODE_0 and NODE_1. Each assembly in the chain is also connected to the assembly in BLOCKS corresponding to a block in the stack. For instance, the sequence of such assemblies in Figure 4.4-C6 corresponds to the sequence of blocks $4, 5, 3, 1, 2$, which is exactly the sequence of blocks in the stack from top to bottom (see the left part of Fig. 4.2). Note that Algorithm 1 uses a constant number of brain areas (that is, five), independently of the number of blocks in the stack. The number of blocks that can be robustly stored with the parser depends on the relation between the number of neurons n and the k for the k-WTA rule. Further details on the limits of chaining is given in Section 4.4.

4.3.2 Removing the Top Block

In order to implement in AC the algorithm which transforms an input stack of blocks into a target stack of blocks, we start by describing an AC program to remove a block from the top of a stack. This program uses the same areas and fibers of the parser described in the previous section (see Figure 4.3-B), with the addition of fibers between HEAD with NODE_1, and HEAD with NODE_2. Intuitively, these fibers are needed to allow changing the head of the chain representing the current stack, without having to shift all the assemblies one position to the left.

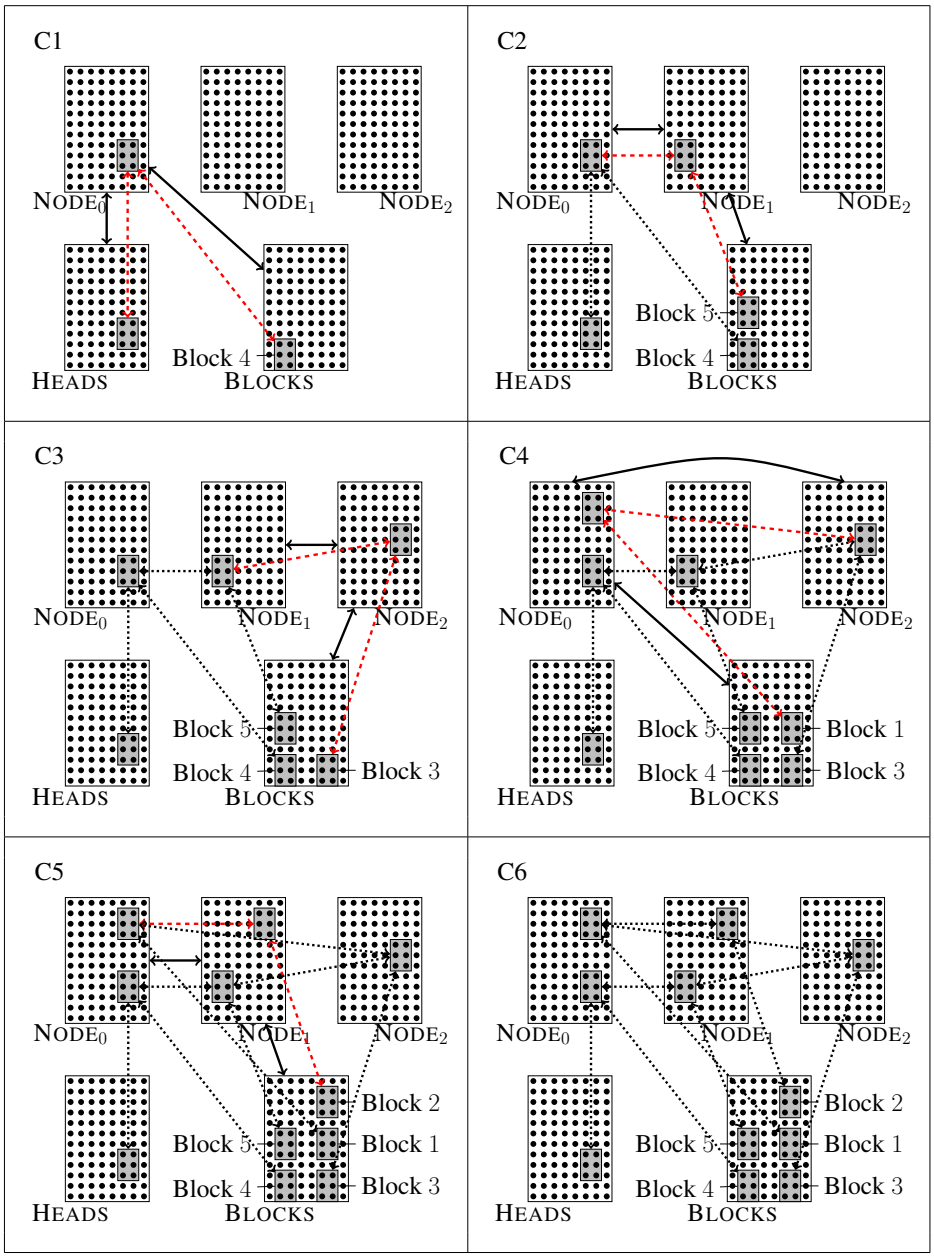


Figure 4.4 – **C1-6.** The behavior of the parser AC program for the initial configuration in 4.1. The black solid lines denote the fibers of Fig. 4.3 which are disinhibited. The red dashed lines denote the newly created connections between assemblies in different areas, while the black dotted lines denote the connections previously created.

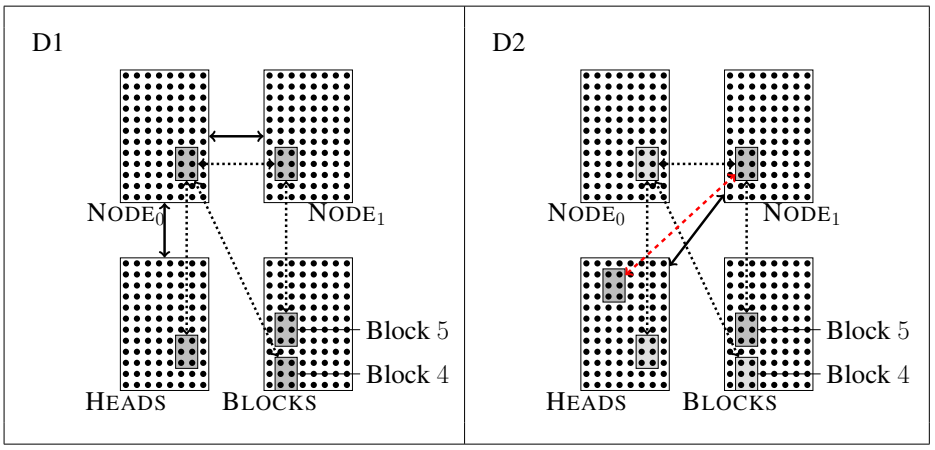


Figure 4.5 – **D1-2.** The behavior of the AC program which removes the block from the top of a stack, with input the data structure resulting from the parser execution (only the areas involved in the remove operation are shown). The black solid lines denote the fibers which are disinhibited. The red dashed lines denote the newly created connections between assemblies in different areas, while the black dotted lines denote the already existing connections.

The AC program, which "removes" the block from the top of the stack, uses the connections created by the parser in order to activate the assembly in the NODE$_1$, which is connected to the block just below the top block (that is block 5 in Figure 4.5-D1,D2). This is done by projecting from the HEAD into NODE$_0$, and projecting from NODE$_0$ into the NODE$_1$ (see Figure 4.5-D1). Through strong projection, the program successively creates a new connection from the active assembly in the NODE$_1$ area to a new assembly in the HEAD area (see the red dashed line in Figure 4.5-D2).

Note that the connections between the light gray assemblies in Figure 4.5-D2 are still active, but they will not be used in the future since the last active assembly in the HEAD area is now connected to the assembly in the NODE$_1$ area. These connections, indeed, might later disappear because of a process of homeostasis, which can be modeled in the AC system through a sort of "renormalization" (as described in [Papadimitriou et al., 2020]). In a certain sense, the system will slowly "forget" which block was on the top of the stack, before a removal operation.

The removal of the top block can be repeated as many times as the number of blocks in the stack. The only difference is that the activation of the assembly in NODE corresponding to the block below the top one is done by projecting HEAD into the NODE area corresponding to the top block, and then projecting from this NODE area to the one following it (in modular arithmetic).

In order to maintain an updated representation of the blocks world configuration, we use four additional brain areas to store the chain of blocks which have been removed and that, hence, are currently on the table. This chain can be implemented in the AC system exactly the same way we did when parsing a stack of blocks. Then, when we want to read the current data structure stored in the AC system, we examine the stack of blocks represented in HEAD and the NODE areas, as well as the chain of blocks on the table in the additional areas.

4.3.3 Putting a Block on Top of the Stack

The second operation we need in order to implement a minimal planning algorithm for the blocks world problem is *putting* a block on top of the stack. The AC program, for this operation first projects the block from in BLOCKS into the NODE area preceding (in modular arithmetic) the NODE area currently connected to HEAD, and then projects the newly created assembly into HEAD (see Figure 4.6-E1). Successively, the program executes a strong projection between the four areas in order to correctly connect them (see Figure 4.6-E2). Once again, an active connection between the HEAD area and a NODE area will still exist after the execution of the AC program, but this connection will not be used in the future.

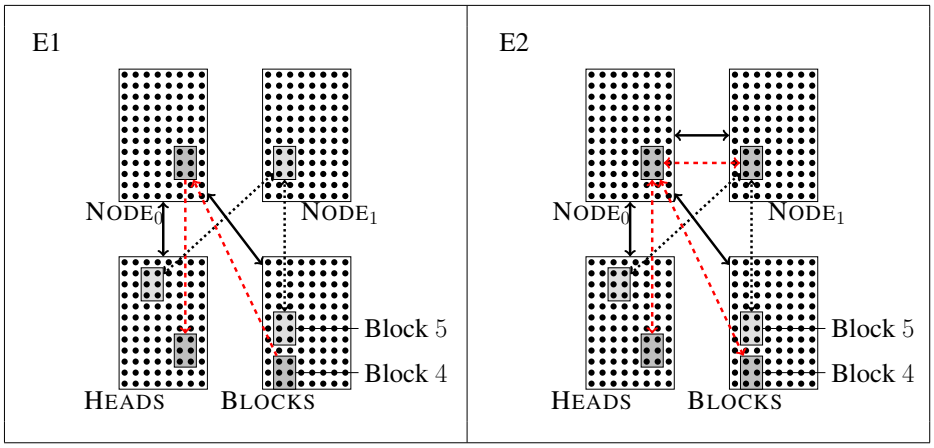


Figure 4.6 – **E1-2.** The behavior of the AC program which put the block 4 on top of the stack, above the block 5. The black solid lines denote the fibers which are disinhibited. The red dashed lines denote the newly created connections (unidirectional and bidirectional) between assemblies in different areas, while the black dotted lines denote the already existing connections.

4.3.4 Computing the Intersection of Two Stacks

The pop and put operations described in the previous two sections are sufficient to implement a simple planning algorithm, which consists in moving all the blocks on the table (by using pop), and by then moving the blocks on the table on top of the stack (by using put) according to the target stack. We remark that the target stack is also stored through a parsing in an another, reserved set of four areas connected to BLOCKS.

In order to improve this algorithm and execute the two-approximation algorithm described in the Section 4.2.2, we need an AC program which implements a third operation, that is, finding the *intersection* of two stacks. This operation looks for the common sub-stack of the two stacks (starting from the bottom) and return the highest block in this sub-stack. Then only the blocks above this block have to be moved on the table and reassembled in the right order.

In a nutshell, this can be achieved in AC by first reaching the bottom of the two stacks which have to be compared, and then proceeding upwards until we find two different blocks, or the end of one of the two stacks.

4.3.5 Multiple Stacks

So far in this exposition we have concerned ourselves with configurations consisting of one stack. In our experiments (see the next section) we have implemented up to five stacks by employing a different set of four areas for each stack. This is a bit unsatisfactory, because it implies that the maximum number of stacks that can be handled by the brain is encoded in the brain architecture.

With multiple stacks one has to solve the *matching problem:* identifying pairs of stacks in the input and output that must be transformed one to the other. Naively, this can be done by comparing all pairs of stacks, but this entails effort that is quadratic in the number of stacks. This latter strategy is the one currently employed in our experiments.

4.4 Experiments

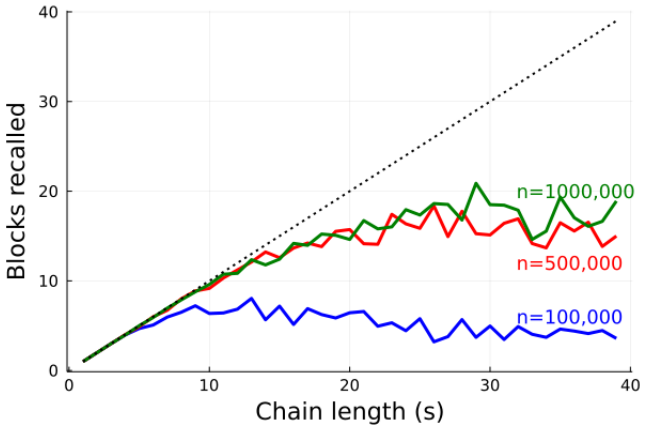
A software system for programming in the AC, as well as implementations of the algorithms described in this work, have been written in Julia [Bezanson et al., 2017]. We make use of the Java generator for BW configurations available at [Koeman, 2020], based on [Slaney and Thiébaux, 2001]. We ran experiments on over 100 blocks-world configurations, with up to five stacks and 10, 20, and 30 blocks. The algorithm worked correctly in every instance. We have used various settings of the parameters n, k, p, β – a particularly good set of parameters is $n = 10^6, k = 50, p = 0.1, \beta = 0.1$. Interestingly, the algorithms do not work in all parameter settings, because of limits on the chaining operation (see the next discussion). The Julia source code can be found at [jBrain, 2021].

In general, the amount of rounds of strong project (parallel spikings of neurons) needed to carry out the BW tasks seems to be around 35 spikes per block processed (parse, popped, or pushed), which, assuming roughly 50 Hz spikes for excitatory neurons in the brain, is around 1.4 seconds per operation.

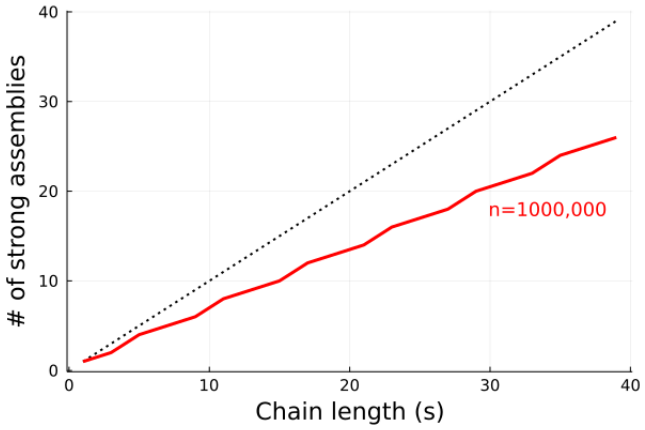
Limits of the AC. An unexpected finding of our simulations is that they are stable only under very specific parameter settings. The bottleneck of the planning algorithms is in parsing the chain of blocks, that is, memorizing the sequence of blocks so they can be read out reliably. In isolation we call this operation "chaining".

The results in this section, which describe some properties and limits of chaining, can be viewed as theoretical properties of the AC. First, we find it is only possible to chain a rather limited number of blocks. For instance, even though with $n = 10^6$ and $k = 50$ there is, at least in theory, space for $10^6/50 = 200000$ non-overlapping assemblies, even with strong p and β, we can only reliably chain up to 20 blocks. This is illustrated in Figure 4.7a, which shows how many of s blocks were successfully read out after chaining. Generally, for higher values of n (and a higher $n : k$ ratio), longer portions of the chain tend to be correctly stored, but the operation is highly noisy: in some trials it will fail and then succeed for a *longer* chain. Indeed, unlike the assembly operations described in [Papadimitriou et al., 2020] (Project, Merge, and so on) which are either stable with overwhelming probability under appropriate parameters, or do not succeed if the parameters are not appropriately strong, chaining appears to push the computational power of the AC to its limits, and often succeeds or fails between repeated trials with the same parameters.

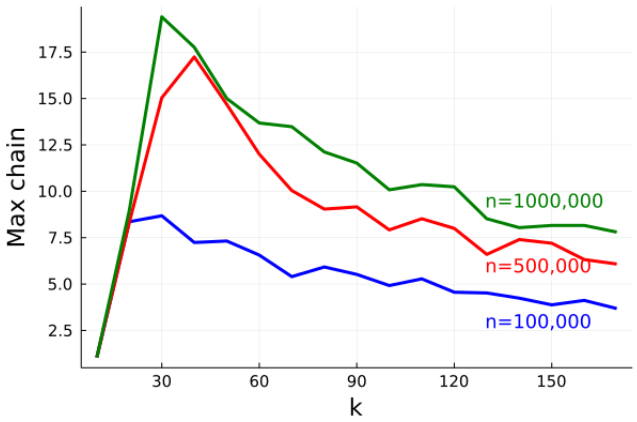
One can also look at a related property: after chaining, how many of the assemblies in the NODE_i areas during readout are "strong" in the sense that they pass the `isAssembly` () test with



(a)



(b)



(c)

Figure 4.7 – Experiments on the “chaining" operation, the bottleneck of the AC planning algorithm. (a) shows number of blocks correctly chained for various chain length; (b) shows number of “strong" assemblies formed in chaining; (c) shows maximal chain length that is correctly parsed for varying k. (b) and (c) show averages over 50 trials per parameter setting (exact numbers, including sample standard deviation, are provided in the appendix). In these charts, $p = \beta = 0.1$ was used, in (a) and (b) $k = 50$.

a high threshold (i.e., firing those k neurons recursively results in the same set of k winners)? Interestingly, this proportion, which is significantly less than the maximum of s, does not change significantly when we vary n, p, β– there appears to be a natural proportion of strong assemblies formed during chaining (Figure 4.7b).

Finally, in Figure 4.7c we varied k and found the maximally long chain that succeeded completely. These experiments again showed that for higher $n : k$ ratio, longer chains are possible, and that for each setting of n there is a narrow window of optimal k that allows for the longest chains– above of this range, as we increase k the maximum chain does not change, i.e. it appears to settle to some natural lower bound. A more thorough analysis of chaining is an important direction in AC theory, since such maneuvers could be subroutines in various cognitive processes (for instance, [Mitropolsky et al., 2021] suggest using it for processing chains of identical parts of speech, such as multiple adjectives in a noun phrase).

4.5 Discussion

The aim of this work is not so much to produce a performing system, but to demonstrate experimentally that reasonably large and complex programs in the AC can execute correctly and reliably, and in particular can implement in a natural manner planning strategies for solving instances of the blocks world problem. In fact, the implementation of these strategies is based on the realization of a *list-like data structure* which makes use of a *constant* number of brain regions. Confirming theoretical insights, we have experimentally found that the structure's reliability depends on the ratio between the number of neurons and the size of the assemblies in each region — even though the dependency was a bit more constraining than we had expected. The reasons and extent of this shortcoming must be the object of further investigation.

We have also shown how simple manipulations of the data structure (such as the top, pop, and append operations) can be realized by making use of a constant number of brain regions. These manipulations allowed us to implement planning strategies based on two basic kinds of moves, that is, moving the block from the top of a stack to the table, and putting a block from the table to the top of a stack. All our programs work for an *arbitrary* number of blocks and a *bounded* number of stacks, while a possible improvement involves implementing a version with an arbitrary number of stacks.

After syntactic analysis in language and blocks world planning, what other investigations of the AC can be carried out? On the one hand, one can ask what other stylized cognitive functions can be described by and implemented in the assembly calculus. There is recent work dealing with *learning* though assemblies of neurons [Dabagia et al., 2022]. Two further realms of cognition come to mind, and they happen to be closely related: *reasoning,* as well as *planning and problem solving* in less specialized domains than BW. More interestingly for the distributed computing community, one can strengthen the AC theory by proposing reasonable dynamics for the k-Winners Take All rule. At this moment, no dynamics is known to run efficiently in such a network (even though the k-WTA rule is well-established in neuroscience and there have been some attempts to design k-WTA dynamics in different networks [Majani et al., 1988,Lynch et al., 2017]), and a theoretical study on plausible solutions is necessary. In [Lynch et al., 2017], input neurons receive a signal and a layer of output neurons must output the winner among all signals; in our case, we would like to have a network where input and output neurons coincide. In [Dabagia et al., 2022] another work [Binas et al., 2014] is cited, which proposes an implementation k-WTA circuits based on the

interaction between excitatory and inhibitory neurons; nevertheless, the authors of the latter work model a 1-WTA dynamics, where the nodes of the network are units representing the average behaviors of clusters of many neurons. Hence, there is just one winning unit which results in many neurons firing together, a phenomenon which is far from the AC framework. In [Majani et al., 1988], a continuous dynamics in the Hopfield model is proposed, which converges to a configuration such that only the k neurons with the highest initial inputs carry positive weights: a possible starting research direction may be discretizing this latter dynamics. Similarly, the (dis)inhibition operations need a distributed, microscopic implementation.

4.6 More details on the experiments

The operation of parsing works in every possible instance we tried provided that the constraints shown in the subsection on the limits of AC are met. We tried parsing randomly generated BW configurations with 10, 20, and 30 blocks divided in multiple stacks, with the following parameters: $n = 4 \times 10^6$ neurons, $p = 0.1, \beta = 0.1, k = 50$. The intersect operation needs two parsed stacks as input, and runs correctly each time the parsing operation of the stacks works correctly. Removing the top block (that is, the pop operation) and putting on top of a stack (that is, the put operation) run as well in the aforementioned settings. The used machine is a DELL laptop with an Intel(R) Core(TM) i7-8665U CPU @ 1.90GHz processor, 32GB Ram, running Fedora 33. The input BW configurations are specified in the `bw_instances` folder at [jBrain, 2021].

Stricter constraints are needed for the whole planning operation. Since the previous operations have to be run many times and one after the other other, the graph representing the brain and its connectivity grows quite quickly. On a machine like the one described above, only BW configurations with at most 10 blocks are well handled (good parameters to test this are the same as for parsing). Otherwise, the program requires too much time in order to be completed, even with the same set of parameters. For the planning with 10 blocks, we used a machine Dell R940 quad-Xeon SP Gold 6148 @ 2.40GHz (80 cores) with 1024 GB of RAM.

In the case of BW configurations with 10 blocks, we have verified the correctness of the AC programs implementing the two planning algorithms (the one without and the one with the intersect operation) on 100 BW configurations randomly chosen, such that each stack has at most 7 blocks, to avoid chaining issues (these configurations are specified in the file `planning_inputs.txt`). The execution on all these instances correctly run and finished in reasonable time.

Due to the discussion on the limits of AC (i.e., the limits on the maximum chain lengths), in the case of BW configurations between 20 and 30 blocks, we have limited ourselves to verify the correctness of the implementation of the basic operations used by the two algorithms, that is, the parser, the pop, the push, and the intersect operations (roughly 30 simulations). All runs were completed in reasonable time (few minutes - up to 20 in the case of more demanding operations) without errors. Also in this case, we used $n = 4 \times 10^6$ neurons for each brain area. We remark that the 20 and 30 blocks must be split among several stacks of maximum length up to 7, otherwise the parsing procedure may fail with the above number of neurons (these BW configurations are specified in the file `operation_inputs.txt`).

The verification of the entire planning algorithms in the case of 20 and 30 blocks (even split among several stacks of up to 7 blocks each) requires more memory and time, due to the large computations needed to represent "bigger" brains. After running the planning algorithms in these

cases, nevertheless, we observed that the initial actions performed by the brain were correct, which makes us believe that the algorithms would also correctly work in its entirety.

Limits of the AC. Table 4.2 shows the outcome of the experiments on chaining. In particular, it shows the mean number of blocks (over 50 runs) we can chain with $n = 10^5, 5 \times 10^5, 10^6$ neurons for each brain area, $k = 50$ (the number of neurons an assembly is composed of), and the standard deviation. With 10^6 neurons, we can reliably chain up to $10, 11$ blocks, but it's better if the number of blocks is less than 8. If we lower the number of neurons, less blocks can be reliably parsed.

Table 4.3 shows the outcome of the experiments on the ratio between n and k which can chain the higher number of blocks. The best reliability is obtained when k increases together with n. For $n = 10^5$, $20 \leq k \leq 30$ seems to be best. For $n = 5 \times 10^5, 10^6$, $30 \leq k \leq 40$ works better. The means and the standard deviations are obtained over 50 runs of the experiment.

In order to execute these experiments, the reader can execute the following terminal command:

```
julia experiments/chaining_experiments.jl
```

Neurons	Blocks	Mean	Std	Neurons	Blocks	Mean	Std	Neurons	Blocks	Mean	Std
10^5	1	1	0	$5 \cdot 10^5$	1	1	0	10^6	1	1	0
10^5	2	2	0	$5 \cdot 10^5$	2	2	0	10^6	2	2	0
10^5	3	3	0	$5 \cdot 10^5$	3	3	0	10^6	3	3	0
10^5	4	3,96	0,28	$5 \cdot 10^5$	4	4	0	10^6	4	3,96	0,28
10^5	5	4,68	0,94	$5 \cdot 10^5$	5	4,88	0,63	10^6	5	5	0
10^5	6	5,08	1,86	$5 \cdot 10^5$	6	6	0	10^6	6	5,92	0,57
10^5	7	5,96	2,13	$5 \cdot 10^5$	7	6,72	1,21	10^6	7	6,96	0,28
10^5	8	6,5	2,64	$5 \cdot 10^5$	8	8	0	10^6	8	7,88	0,63
10^5	9	7,22	3,22	$5 \cdot 10^5$	9	8,88	0,85	10^6	9	8,78	1,3
10^5	10	6,36	3,86	$5 \cdot 10^5$	10	9,16	2,22	10^6	10	9,6	1,81
10^5	11	6,42	4,53	$5 \cdot 10^5$	11	10,32	1,91	10^6	11	10,74	1,38
10^5	12	6,84	4,4	$5 \cdot 10^5$	12	11,22	2,47	10^6	12	10,84	2,57
10^5	13	8,04	4,69	$5 \cdot 10^5$	13	12,1	2,77	10^6	13	12,36	2,45
10^5	14	5,66	4,98	$5 \cdot 10^5$	14	13,22	2,74	10^6	14	11,78	4,26
10^5	15	7,18	5,46	$5 \cdot 10^5$	15	12,58	4,52	10^6	15	12,42	4,51
10^5	16	5,14	4,89	$5 \cdot 10^5$	16	13,66	4,35	10^6	16	14,18	3,77
10^5	17	6,9	5,82	$5 \cdot 10^5$	17	14,22	5,41	10^6	17	13,96	5,03
10^5	18	6,24	5,25	$5 \cdot 10^5$	18	13,82	6,73	10^6	18	15,22	4,99
10^5	19	5,86	4,78	$5 \cdot 10^5$	19	15,52	5,87	10^6	19	15,08	6,19
10^5	20	6,44	5,91	$5 \cdot 10^5$	20	15,72	6,25	10^6	20	14,62	6,66
10^5	21	6,58	4,99	$5 \cdot 10^5$	21	14,14	8,16	10^6	21	16,72	6,87
10^5	22	4,94	4,55	$5 \cdot 10^5$	22	14,08	7,85	10^6	22	15,8	7,69
10^5	23	5,32	5,73	$5 \cdot 10^5$	23	17,42	7,85	10^6	23	16,02	8,04
10^5	24	4,44	3,97	$5 \cdot 10^5$	24	16,36	7,51	10^6	24	17,94	7,88
10^5	25	5,78	5,07	$5 \cdot 10^5$	25	15,86	9,67	10^6	25	17,34	9,2
10^5	26	3,2	2,85	$5 \cdot 10^5$	26	18,36	9	10^6	26	18,62	9,24
10^5	27	3,78	4,11	$5 \cdot 10^5$	27	14,94	10,15	10^6	27	18,52	9,36
10^5	28	5,68	4,76	$5 \cdot 10^5$	28	17,76	10,88	10^6	28	16,78	9,76
10^5	29	3,7	3,42	$5 \cdot 10^5$	29	15,26	9,08	10^6	29	20,88	10,21
10^5	30	4,96	4,18	$5 \cdot 10^5$	30	15,12	10,13	10^6	30	18,5	11,14
10^5	31	3,46	3,72	$5 \cdot 10^5$	31	16,42	12	10^6	31	18,44	11,79
10^5	32	4,9	3,87	$5 \cdot 10^5$	32	16,92	10,97	10^6	32	17,88	11,87
10^5	33	4,04	3,81	$5 \cdot 10^5$	33	14,16	11,63	10^6	33	14,62	11,38
10^5	34	3,7	3,59	$5 \cdot 10^5$	34	13,68	10,82	10^6	34	15,54	12,04
10^5	35	4,62	4,7	$5 \cdot 10^5$	35	16,48	11,78	10^6	35	19,34	11,9
10^5	36	4,4	4,16	$5 \cdot 10^5$	36	15,56	12,73	10^6	36	17,04	12,71
10^5	37	4,12	3,49	$5 \cdot 10^5$	37	16,54	12,67	10^6	37	16,08	12,29
10^5	38	4,46	3,88	$5 \cdot 10^5$	38	13,84	12,86	10^6	38	16,64	13,46
10^5	39	3,58	3,24	$5 \cdot 10^5$	39	15	12,68	10^6	39	18,9	13,67

TABLE 4.2 – Table for the chain length experiment. We check if $n \in \{10^5, 5 \cdot 10^5, 10^6\}$ neurons for brain area are sufficient to parse a stack of a given number of blocks. Here, the mean number of correctly parsed blocks and its standard deviation are shown over 50 runs of the experiment.

Neurons	k	**Mean**	**Std**	**Neurons**	k	**Mean**	**Std**	**Neurons**	k	**Mean**	**Std**
10^5	10	1,04	0,28	$5 \cdot 10^5$	10	1,04	0,28	10^6	10	1,04	0,28
10^5	20	8,36	4,25	$5 \cdot 10^5$	20	8,32	5,22	10^6	20	8,96	4,47
10^5	30	8,68	3,03	$5 \cdot 10^5$	30	15,04	6,38	10^6	30	19,4	6,93
10^5	40	7,24	2,93	$5 \cdot 10^5$	40	17,24	5,68	10^6	40	17,76	6,13
10^5	50	7,32	2,47	$5 \cdot 10^5$	50	14,68	5,53	10^6	50	15	5,51
10^5	60	6,56	2,43	$5 \cdot 10^5$	60	12	4,15	10^6	60	13,68	5,01
10^5	70	5,4	2,06	$5 \cdot 10^5$	70	10,04	2,98	10^6	70	13,48	4,92
10^5	80	5,92	1,99	$5 \cdot 10^5$	80	9,04	3,86	10^6	80	12,12	4,34
10^5	90	5,52	1,97	$5 \cdot 10^5$	90	9,16	3,1	10^6	90	11,52	3,87
10^5	100	4,92	1,98	$5 \cdot 10^5$	100	7,92	2,83	10^6	100	10,08	3,96
10^5	110	5,28	1,57	$5 \cdot 10^5$	110	8,52	2,82	10^6	110	10,36	3,73
10^5	120	4,56	1,58	$5 \cdot 10^5$	120	8	2,92	10^6	120	10,24	3,38
10^5	130	4,52	1,64	$5 \cdot 10^5$	130	6,6	2,56	10^6	130	8,52	4,05
10^5	140	4,24	1,33	$5 \cdot 10^5$	140	7,4	2,71	10^6	140	8,04	2,78
10^5	150	3,88	1,15	$5 \cdot 10^5$	150	7,2	2,43	10^6	150	8,16	2,97
10^5	160	4,12	1,08	$5 \cdot 10^5$	160	6,32	2,12	10^6	160	8,16	3,23
10^5	170	3,68	1,19	$5 \cdot 10^5$	170	6,08	2,18	10^6	170	7,8	3,21

TABLE 4.3 – Table for the max chain length experiment. We check, for a given number of k (the number of neurons an assembly is composed of), the maximum chain length the brain can correctly parse, with $n \in \{10^5, 5 \cdot 10^5, 10^6\}$ neurons in each area of the brain. Here, the mean number of the maximum chain length correctly parsed blocks and its standard deviation are shown over 50 runs of the experiment.

CHAPTER 5

Conclusion et Perspectives

The myriad of dynamics and algorithms found in nature attract great interest from interdisciplinary fields. In this thesis, we have studied them from a computational, mostly theoretical, perspective, and tried to capture some properties with the hope of hinting at more general phenomena and aspects.

Opinion dynamics are common models to describe how a system of agents achieves collective tasks, such as consensus or community detection. Our results are in line with other recent works showing how such systems are characterized by phase-transitions that depend on the parameters of the considered settings and reflect characteristics of the surrounding environment [Cruciani et al., 2018, Mukhopadhyay et al., 2020, Cruciani et al., 2021]. This phenomenon suggests that in noisy-like environments biological organisms should be less likely to gather since collective tasks are more difficult to carry out. In general, the study of opinion dynamics is still far from exhausted: a research direction which we find very interesting consists in dynamics settings in which the topology of the network changes over time. This can be achieved in various ways: one of the simpler one, is imagining the agents as random walks over some graph, such as a thorus, and the interaction/message exchange is concurrent with meeting at a node. Such temporal structure of the underlying communication graph surely gives more plausibility to the investigated model and can provide a more complete understanding of real dynamical systems.

As for movement patterns, in nature many cost functions are optimized for different reasons: it can be the overall energy spent by a biological system, or the hitting time to find some targets. Often, variations in such cost functions can make the difference between life and death. In the introduction, we already discussed that it is no news that Lévy walks are excellent search strategies in many different settings [Viswanathan et al., 1999, Reynolds, 2018], as they seem to optimize hitting times of sparse targets. Nevertheless, to the best of our knowledge, we are the first to introduce their parallel study and to show their optimal performance in a distributed search problem on the infinite two-dimensional lattice [Feinerman and Korman, 2017]. Inspired by the Lévy flight foraging hypothesis, we began to study Lévy walks and ended up corroborating it. We believe Lévy walks must be further investigated as there's still much to discover: e.g., how well do they perform in other distributed tasks? We think, for example, about information dissemination, epidemic processes, etc.

Moving instead to neuronal computing, we observe that in the last decades there has been a lot of interest in the study of neuronal networks in computer science. Understanding how the brain works — at least, some aspects of it — has been beneficial to computer science: consider, e.g., the fly's nervous system that led to the design of a distributed maximal independent set algorithm

[Afek et al., 2011]), or its olfactory circuit that solves the similarity search problem using a novel locality-sensitive hashing [Dasgupta et al., 2017]. Modeling cognitive phenomena is, of course, a much more difficult task. The Assembly Calculus takes a step towards this direction and we think it is able to capture some important functions of the brain that serve to cognition, such as the projection or the association of assemblies. Our contribution here has been to empirically prove that with the AC framework we are able to implement large and complex heuristics that run relatively quickly and reliably. Nevertheless, we remark that for higher cognitive functions (even the planning itself) the model needs to be strengthen at a microscopic level; with respect to this, we think that interesting and open research directions are finding and modeling reasonable k-Winners-Take-All dynamics, not only restricted to the AC framework but also more in general, as the k-WTA dynamics seems to be a fundamental process in neuroscience. Furthermore, the AC would benefit also from an accurate modeling of (dis)inhibition processes.

Bibliography

[Abdullah and Draief, 2015] Abdullah, M. A. and Draief, M. (2015). Global majority consensus by local majority polling on graphs of a given degree sequence. *Discret. Appl. Math.*, 180:1–10.

[Acemoglu et al., 2013] Acemoglu, D., Como, G., Fagnani, F., and Ozdaglar, A. E. (2013). Opinion fluctuations and disagreement in social networks. *Math. Oper. Res.*, 38(1):1–27.

[Afek et al., 2011] Afek, Y., Alon, N., Barad, O., Hornstein, E., Barkai, N., and Bar-Joseph, Z. (2011). A biological solution to a fundamental distributed computing problem. *Science*, 331:183 – 185.

[Alon et al., 2011] Alon, N., Avin, C., Koucký, M., Kozma, G., Lotker, Z., and Tuttle, M. R. (2011). Many random walks are gaster than one. *Combinatorics, Probability and Computing*, 20(4):481–502.

[Angluin et al., 2008] Angluin, D., Aspnes, J., and Eisenstat, D. (2008). A simple population protocol for fast robust approximate majority. *Distributed Computing*, 21(2):87–102.

[Axel, 2018] Axel, R. (2018). Q&A. *Neuron*, 99(6):1110–1112. doi: 10.1016/j.neuron.2018.09.003.

[Bankhamer et al., 2022] Bankhamer, G., Berenbrink, P., Biermeier, F., Elsässer, R., Hosseinpour, H., Kaaser, D., and Kling, P. (2022). Fast consensus via the unconstrained undecided state dynamics. In Naor, J. S. and Buchbinder, N., editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022*. SIAM.

[Basanisi et al., 2020] Basanisi, R., Brovelli, A., Cartoni, E., and Baldassarre, G. (2020). A generative spiking neural-network model of goal-directed behaviour and one-step planning. *PLOS Computational Biology*, 16(12):1–32.

[Bassler, 2002] Bassler, B. L. (2002). Small talk: Cell-to-cell communication in bacteria. *Cell*, 109(4).

[Bayindir, 2016] Bayindir, L. (2016). A review of swarm robotics tasks. *Neurocomputing*, 172:292–321.

[Becchetti et al., 2022] Becchetti, L., Carvalho Walraven da Cunha, A., Clementi, A., d'Amore, F., Lesfari, H., Natale, E., and Trevisan, L. (2022). On the Multidimensional Random Subset Sum Problem. Research report, Inria & Université Cote d'Azur, CNRS, I3S, Sophia Antipolis, France ; Sapienza Università di Roma, Rome, Italy ; Università Bocconi, Milan, Italy ; Università di Roma Tor Vergata, Rome, Italy.

[Becchetti et al., 2020a] Becchetti, L., Clementi, A. E. F., and Natale, E. (2020a). Consensus dynamics: An overview. *SIGACT News*, 51(1).

[Becchetti et al., 2015] Becchetti, L., Clementi, A. E. F., Natale, E., Pasquale, F., and Silvestri, R. (2015). Plurality consensus in the gossip model. In Indyk, P., editor, *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 371–390. SIAM.

[Becchetti et al., 2017] Becchetti, L., Clementi, A. E. F., Natale, E., Pasquale, F., Silvestri, R., and Trevisan, L. (2017). Simple dynamics for plurality consensus. *Distributed Comput.*, 30(4):293–306.

[Becchetti et al., 2016] Becchetti, L., Clementi, A. E. F., Natale, E., Pasquale, F., and Trevisan, L. (2016). Stabilizing consensus with many opinions. In Krauthgamer, R., editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 620–635. SIAM.

[Becchetti et al., 2020b] Becchetti, L., Clementi, A. E. F., Natale, E., Pasquale, F., and Trevisan, L. (2020b). Find your place: Simple distributed algorithms for community detection. *SIAM J. Comput.*, 49(4):821–864.

[Bekolay et al., 2014] Bekolay, T., Bergstra, J., Hunsberger, E., DeWolf, T., Stewart, T., Rasmussen, D., Choo, X., Voelker, A., and Eliasmith, C. (2014). Nengo: a Python tool for building large-scale functional brain models. *Frontiers in Neuroinformatics*, 7(48):1–13.

[Ben-Shahar et al., 2014] Ben-Shahar, O., Dolev, S., Dolgin, A., and Segal, M. (2014). Direction election in flocking swarms. *Ad Hoc Networks*, 12:250–258.

[Benezit et al., 2009] Benezit, F., Thiran, P., and Vetterli, M. (2009). Interval consensus: From quantized gossip to voting. In *ICASSP 2009*, pages 3661–3664.

[Berenbrink et al., 2017] Berenbrink, P., Clementi, A. E. F., Elsässer, R., Kling, P., Mallmann-Trenn, F., and Natale, E. (2017). Ignore or comply?: On breaking symmetry in consensus. In Schiller, E. M. and Schwarzmann, A. A., editors, *Proceedings of the ACM Symposium on Principles of Distributed Computing, PODC*. ACM.

[Berenbrink et al., 2016] Berenbrink, P., Friedetzky, T., Giakkoupis, G., and Kling, P. (2016). Efficient Plurality Consensus, Or: the Benefits of Cleaning up from Time to Time. In *ICALP 2016*, volume 55 of *LIPIcs*, pages 136:1–136:14.

[Bezanson et al., 2017] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.

[Binas et al., 2014] Binas, J., Rutishauser, U., Indiveri, G., and Pfeiffer, M. (2014). Learning and stabilization of winner-take-all dynamics through interacting excitatory and inhibitory plasticity. *Frontiers in Computational Neuroscience*, 8.

[Boczkowski et al., 2018a] Boczkowski, L., Guinard, B., Korman, A., Lotker, Z., and Renault, M. (2018a). Random walks with multiple step lengths. In *LATIN 2018: Theoretical Informatics*, pages 174–186.

[Boczkowski et al., 2019] Boczkowski, L., Korman, A., and Natale, E. (2019). Minimizing message size in stochastic communication patterns: fast self-stabilizing protocols with 3 bits. *Distributed Computing*.

[Boczkowski et al., 2018b] Boczkowski, L., Natale, E., Feinerman, O., and Korman, A. (2018b). Limits on reliable information flows through stochastic populations. *PLOS Computational Biology*, 14(6):e1006195.

[Bonifaci et al., 2012] Bonifaci, V., Mehlhorn, K., and Varma, G. (2012). Physarum can compute shortest paths. *Journal of theoretical biology*, 309:121–33.

[Boyer et al., 2006] Boyer, D., Ramos-Fernández, G., Miramontes, O., Mateos, J. L., Cocho, G., Larralde, H., Ramos, H., and Rojas, F. (2006). Scale-free foraging by primates emerges from

their interaction with a complex environment. *Proceedings of the Royal Society B: Biological Sciences*, 273(1595):1743–1750.

[Buldyrev et al., 2001] Buldyrev, S. V., Havlin, S., Kazakov, A. Y., da Luz, M. G. E., Raposo, E. P., Stanley, H. E., and Viswanathan, G. M. (2001). Average time spent by Lévy flights and walks on an interval with absorbing boundaries. *Phys. Rev. E*, 64:041108.

[Buldyrev et al., 2021] Buldyrev, S. V., Raposo, E. P., Bartumeus, F., Havlin, S., Rusch, F. R., da Luz, M. G. E., and Viswanathan, G. M. (2021). Comment on "inverse square Lévy walks are not optimal search strategies for $d \geq 2$". *Phys. Rev. Lett.*, 126:048901.

[Buzsáki, 2010] Buzsáki, G. (2010). Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron*, 68(3):362–385.

[Buzsáki, 2021] Buzsáki, G. (2021). *The Brain from Inside Out*. Oxford University Press, Oxford, reprint edition edition.

[Cardelli and Csikàsz-Nagy, 2012] Cardelli, L. and Csikàsz-Nagy, A. (2012). The Cell Cycle Switch Computes Approximate Majority. *Scientific Reports*, 2:656.

[Carrillo-Reid et al., 2019] Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A., and Yuste, R. (2019). Controlling visually guided behavior by holographic recalling of cortical ensembles. *Cell*, 178:447–457.e5.

[Carroll, 2004] Carroll, M. C. (2004). The complement system in regulation of adaptive immunity. *Nature Immunology*, 5:981–986.

[Carvalho Walraven da Cunha et al., 2022] Carvalho Walraven da Cunha, A., d'Amore, F., Giroire, F., Lesfari, H., Natale, E., and Viennot, L. (2022). Revisiting the Random Subset Sum problem. Research report, Inria Sophia Antipolis - Méditerranée, Université Côte d'Azur ; Inria Paris.

[Chady, 1999] Chady, M. (1999). Modelling higher cognitive functions with hebbian cell assemblies. In Hendler, J. and Subramanian, D., editors, *Proceedings of AAAI/IAAI 1999, July 18-22, 1999, Orlando, Florida, USA*, page 943. AAAI Press.

[Chazelle, 2009] Chazelle, B. (2009). Natural algorithms. In Mathieu, C., editor, *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, pages 422–431. SIAM.

[Chazelle, 2012] Chazelle, B. (2012). Natural algorithms and influence systems. *Commun. ACM*, 55(12):101–110.

[Chung and Lu, 2006] Chung, F. and Lu, L. (2006). Concentration inequalities and martingale inequalities: A survey. *Internet Mathematics*, 3.

[Clementi et al., 2021] Clementi, A., d'Amore, F., Giakkoupis, G., and Natale, E. (2021). Search via parallel lévy walks on z2. In *Proceedings of the 2021 ACM Symposium on Principles of Distributed Computing*, PODC'21, page 81–91, New York, NY, USA. Association for Computing Machinery.

[Clementi et al., 2018] Clementi, A. E. F., Ghaffari, M., Gualà, L., Natale, E., Pasquale, F., and Scornavacca, G. (2018). A tight analysis of the parallel undecided-state dynamics with two colors. In Potapov, I., Spirakis, P. G., and Worrell, J., editors, *43rd International Symposium on Mathematical Foundations of Computer Science, MFCS 2018, August 27-31, 2018, Liverpool, UK*, volume 117 of *LIPIcs*, pages 28:1–28:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Cohen et al., 2017] Cohen, L., Emek, Y., Louidor, O., and Uitto, J. (2017). Exploring an infinite space with finite memory scouts. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 207–224.

[Condon et al., 2019] Condon, A., Hajiaghayi, M., Kirkpatrick, D., and Manuch, J. (2019). Approximate majority analyses using tri-molecular chemical reaction networks. *Natural Computing*.

[Cruciani et al., 2021] Cruciani, E., Mimun, H. A., Quattropani, M., and Rizzo, S. (2021). Phase transitions of the k-majority dynamics in a biased communication model. In *ICDCN '21: International Conference on Distributed Computing and Networking*. ACM.

[Cruciani et al., 2018] Cruciani, E., Natale, E., Nusser, A., and Scornavacca, G. (2018). Phase Transition of the 2-Choices Dynamics on Core-Periphery Networks. In *AAMAS '18*, pages 777–785.

[Cruciani et al., 2019] Cruciani, E., Natale, E., and Scornavacca, G. (2019). Distributed Community Detection via Metastability of the 2-Choices Dynamics. In *AAAI 2019*, Honolulu, Hawaii, United States.

[Dabagia et al., 2022] Dabagia, M., Vempala, S. S., and Papadimitriou, C. (2022). Assemblies of neurons learn to classify well-separated distributions. In Loh, P.-L. and Raginsky, M., editors, *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 3685–3717. PMLR.

[D'Amore et al., 2020] D'Amore, F., Clementi, A. E. F., and Natale, E. (2020). Phase transition of a non-linear opinion dynamics with noisy interactions - (extended abstract). In Richa, A. W. and Scheideler, C., editors, *Structural Information and Communication Complexity - 27th International Colloquium, SIROCCO 2020, Paderborn, Proceedings*, volume 12156 of *Lecture Notes in Computer Science*, pages 255–272. Springer.

[D'Amore et al., 2022a] D'Amore, F., Clementi, A. E. F., and Natale, E. (2022a). Phase transition of a nonlinear opinion dynamics with noisy interactions. *Swarm Intelligence*, 16(4):261–304.

[D'Amore et al., 2022b] D'Amore, F., Mitropolsky, D., Crescenzi, P., Natale, E., and Papadimitriou, C. H. (2022b). Planning with biological neurons and synapses. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 21–28. AAAI Press.

[D'Amore and Ziccardi, 2022] D'Amore, F. and Ziccardi, I. (2022). Phase transition of the 3-majority dynamics with uniform communication noise. In Parter, M., editor, *Structural Information and Communication Complexity - 29th International Colloquium, SIROCCO 2022, Paderborn, Germany, June 27-29, 2022, Proceedings*, volume 13298 of *Lecture Notes in Computer Science*, pages 98–115. Springer.

[Dasgupta et al., 2017] Dasgupta, S., Stevens, C. F., and Navlakha, S. (2017). A neural algorithm for a fundamental computing problem. *Science*, 358(6364):793–796.

[Dietzfelbinger et al., 2010] Dietzfelbinger, M., Goerdt, A., Mitzenmacher, M., Montanari, A., Pagh, R., and Rink, M. (2010). Tight thresholds for cuckoo hashing via XORSAT. In Abramsky, S., Gavoille, C., Kirchner, C., auf der Heide, F. M., and Spirakis, P. G., editors, *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010*, Lecture Notes in Computer Science. Springer.

[Doerr, 2018] Doerr, B. (2018). Probabilistic tools for the analysis of randomized optimization heuristics. *CoRR*, abs/1801.06733.

[Doerr et al., 2011] Doerr, B., Goldberg, L. A., Minder, L., Sauerwald, T., and Scheideler, C. (2011). Stabilizing consensus with the power of two choices. In Rajaraman, R. and auf der Heide, F. M., editors, *SPAA 2011: Proceedings of the 23rd Annual ACM Symposium on Parallelism in Algorithms and Architectures*. ACM.

[Dong et al., 2019] Dong, H., Mao, J., Lin, T., Wang, C., Li, L., and Zhou, D. (2019). Neural logic machines. In *International Conference on Learning Representations*.

[Doursat, 2013] Doursat, R. (2013). Bridging the Mind-Brain Gap by Morphogenetic "Neuron Flocking": The Dynamic Self-Organization of Neural Activity into Mental Shapes. In *AAAI Fall Symposia*.

[Draief and Vojnovic, 2012] Draief, M. and Vojnovic, M. (2012). Convergence speed of binary interval consensus. *SIAM J. on Control and Optimization*, 50(3):1087–1109.

[Dubhashi and Panconesi, 2009] Dubhashi, D. P. and Panconesi, A. (2009). *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press.

[Edwards et al., 2007] Edwards, A. M., Phillips, R. A., Watkins, N. W., Freeman, M. P., Murphy, E. J., Afanasyev, V., Buldyrev, S. V., Luz, M. G. E. d., Raposo, E. P., Stanley, H. E., and Viswanathan, G. M. (2007). Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449(7165):1044–1048.

[Efremenko and Reingold, 2009] Efremenko, K. and Reingold, O. (2009). How well do random walks parallelize? In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX*, pages 476–489.

[Elsässer et al., 2017] Elsässer, R., Friedetzky, T., Kaaser, D., Mallmann-Trenn, F., and Trinker, H. (2017). Brief Announcement: Rapid Asynchronous Plurality Consensus. In *PODC '17*, pages 363–365.

[Elsässer and Sauerwald, 2011] Elsässer, R. and Sauerwald, T. (2011). Tight bounds for the cover time of multiple random walks. *Theoretical Computer Science*, 412(24):2623 – 2641.

[Emanuele Natale, 2017] Emanuele Natale (2017). *On the Computational Power of Simple Dynamics*. PhD Thesis, Sapienza University of Rome.

[Emek et al., 2015] Emek, Y., Langner, T., Stolz, D., Uitto, J., and Wattenhofer, R. (2015). How many ants does it take to find the food? *Theoretical Computer Science*, 608:255 – 267.

[Emek et al., 2014] Emek, Y., Langner, T., Uitto, J., and Wattenhofer, R. (2014). Solving the ANTS problem with asynchronous finite state machines. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014*, pages 471–482.

[Feinerman et al., 2017] Feinerman, O., Haeupler, B., and Korman, A. (2017). Breathe before speaking: efficient information dissemination despite noisy, limited and anonymous communication. *Distributed Comput.*

[Feinerman and Korman, 2013] Feinerman, O. and Korman, A. (2013). Theoretical distributed computing meets biology: A review. In Hota, C. and Srimani, P. K., editors, *Distributed Computing and Internet Technology, 9th International Conference, ICDCIT 2013, Bhubaneswar, India, February 5-8, 2013. Proceedings*, volume 7753 of *Lecture Notes in Computer Science*, pages 1–18. Springer.

[Feinerman and Korman, 2017] Feinerman, O. and Korman, A. (2017). The ANTS problem. *Distributed Comput.*, 30(3):149–168.

[Feinerman et al., 2012] Feinerman, O., Korman, A., Lotker, Z., and Sereni, J.-S. (2012). Collaborative search on the plane without communication. In *Proceedings of the 2012 ACM Symposium on Principles of Distributed Computing*, PODC '12, page 77–86, New York, NY, USA. Association for Computing Machinery.

[Feller, 1968] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley.

[Focardi et al., 2009] Focardi, S., Montanaro, P., and Pecchioli, E. (2009). Adaptive lévy walks in foraging fallow deer. *PLoS One*, 4(8):e6587.

[Fraigniaud et al., 2016] Fraigniaud, P., Korman, A., and Rodeh, Y. (2016). Parallel exhaustive search without coordination. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing, STOC*, pages 312–323.

[Fraigniaud and Natale, 2019] Fraigniaud, P. and Natale, E. (2019). Noisy rumor spreading and plurality consensus. *Distributed Comput.*

[Franks et al., 2002] Franks, N., Pratt, S., Mallon, E., Britton, N., and Sumpter, D. (2002). Information flow, opinion polling and collective intelligence in house-hunting social insects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 357:1567–83.

[Gamal and Kim, 2011] Gamal, A. E. and Kim, Y. (2011). *Network Information Theory*. Cambridge University Press.

[Ghaffari and Lengler, 2018] Ghaffari, M. and Lengler, J. (2018). Nearly-tight analysis for 2-choice and 3-majority consensus dynamics. In *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*, PODC '18, page 305–313, New York, NY, USA. Association for Computing Machinery.

[Ghaffari and Parter, 2016] Ghaffari, M. and Parter, M. (2016). A Polylogarithmic Gossip Algorithm for Plurality Consensus. In *PODC '16*, pages 117–126.

[Graves et al., 2014] Graves, A., Wayne, G., and Danihelka, I. (2014). Neural turing machines.

[Guinard and Korman, 2020] Guinard, B. and Korman, A. (2020). Tight bounds for the cover times of random walks with heterogeneous step lengths. In Paul, C. and Bläser, M., editors, *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France*, volume 154 of *LIPIcs*, pages 28:1–28:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Guinard and Korman, 2021] Guinard, B. and Korman, A. (2021). Intermittent inverse-square lévy walks are optimal for finding targets of all sizes. *Science Advances*, 7:eabe8211.

[Gupta and Nau, 1991] Gupta, N. and Nau, D. S. (1991). Complexity results for blocks-world planning. In *Proceedings of the Ninth National Conference on Artificial Intelligence - Volume 2*, AAAI'91, page 629–633. AAAI Press.

[Gupta and Nau, 1992] Gupta, N. and Nau, D. S. (1992). On the complexity of blocks-world planning. *Artificial Intelligence*, 56(2):223–254.

[Harris et al., 2012] Harris, T. H., Banigan, E. J., Christian, D. A., Konradt, C., Wojno, E. D. T., Norose, K., Wilson, E. H., John, B., Weninger, W., Luster, A. D., Liu, A. J., and Hunter, C. A.

(2012). Generalized lévy walks and the role of chemokines in migration of effector cd8+ t cells. *Nature*, 486:545 – 548.

[Hassin and Peleg, 2001] Hassin, Y. and Peleg, D. (2001). Distributed Probabilistic Polling and Applications to Proportionate Agreement. *Information and Computation*, 171(2):248–268.

[Hayashi, 2007] Hayashi, M. (2007). Stacking of blocks by chimpanzees: developmental processes and physical understanding. *Animal Cognition*, 10:89–103.

[Hoory et al., 2006] Hoory, S., Linial, N., and Wigderson, W. (2006). Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S)*, 43:439–561.

[Humphries et al., 2010] Humphries, N. E., Queiroz, N., Dyer, J. R., Pade, N. G., Musyl, M. K., Schaefer, K. M., Fuller, D. W., Brunnschweiler, J. M., Doyle, T. K., Houghton, J. D., et al. (2010). Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature*, 465(7301):1066–1069.

[Humphries et al., 2012] Humphries, N. E., Weimerskirch, H., Queiroz, N., Southall, E. J., and Sims, D. W. (2012). Foraging success of biological Lévy flights recorded in situ. *Proceedings of the National Academy of Sciences*, 109(19):7169–7174.

[Iserles, 2008] Iserles, A. (2008). *Euler's method and beyond*, page 3–18. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2 edition.

[Ivaskovic et al., 2017] Ivaskovic, A., Kosowski, A., Pajak, D., and Sauerwald, T. (2017). Multiple random walks on paths and grids. In Vollmer, H. and Vallée, B., editors, *34th Symposium on Theoretical Aspects of Computer Science, STACS*, pages 44:1–44:14.

[jBrain, 2021] jBrain (2021). https://github.com/piluc/jBrain.

[Kanade et al., 2019] Kanade, V., Mallmann-Trenn, F., and Sauerwald, T. (2019). On coalescence time in graphs: When is coalescing as fast as meeting?: Extended abstract. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 956–965.

[Kang and Rivera, 2019] Kang, N. and Rivera, N. (2019). Best-of-three voting on dense graphs. In Scheideler, C. and Berenbrink, P., editors, *The 31st ACM on Symposium on Parallelism in Algorithms and Architectures, SPAA*. ACM.

[Karp, 2011] Karp, R. M. (2011). Understanding science through the computational lens. *J. Comput. Sci. Technol.*, 26(4):569–577.

[Kennedy and Trafton, 2006] Kennedy, W. G. and Trafton, J. G. (2006). Long-term symbolic learning in soar and act-r. In *Proceedings of the Seventh International Conference on Cognitive Modeling*, page 166–171.

[Klein and Young, 2015] Klein, P. and Young, N. E. (2015). On the number of iterations for dantzig–wolfe optimization and packing-covering approximation algorithms. *SIAM Journal on Computing*, 44(4):1154–1172.

[Kleinberg, 2000] Kleinberg, J. (2000). The small-world phenomenon: An algorithmic perspective. In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing, STOC*, pages 163–170.

[Koeman, 2020] Koeman, V. (2020). The blocks world. https://github.com/eishub/blocksworld#readme. [Online; last access 08-September-2021].

[Koetter and Kschischang, 2008] Koetter, R. and Kschischang, F. R. (2008). Coding for errors and erasures in random network coding. *IEEE Trans. Inf. Theory*.

[Korolev and Shevtsova, 2010] Korolev, V. and Shevtsova, I. (2010). On the upper bound for the absolute constant in the berry–esseen inequality. *Theory of Probability and its Applications*, 54.

[Kurup, 2008] Kurup, U. (2008). *Design and use of a bimodal cognitive architecture for diagrammatic reasoning and cognitive modeling*. Ph.D. diss., Graduate School of the Ohio State University.

[Lam et al., 2012] Lam, H., Liu, Z., Mitzenmacher, M., Sun, X., and Wang, Y. (2012). Information dissemination via random walks in d-dimensional space. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, page 1612–1622, USA. Society for Industrial and Applied Mathematics.

[Legenstein et al., 2018] Legenstein, R., Maass, W., Papadimitriou, C. H., and Vempala, S. S. (2018). Long term memory and the densest k-subgraph problem. In Karlin, A. R., editor, *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, volume 94 of *LIPIcs*, pages 57:1–57:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Lehre and Witt, 2014] Lehre, P. K. and Witt, C. (2014). Concentrated hitting times of randomized search heuristics with variable drift. In Ahn, H. and Shin, C., editors, *Algorithms and Computation - 25th International Symposium, ISAAC 2014, Jeonju, Korea, Proceedings*, volume 8889 of *Lecture Notes in Computer Science*, pages 686–697. Springer.

[Lenzen et al., 2017] Lenzen, C., Lynch, N. A., Newport, C., and Radeva, T. (2017). Searching without communicating: Tradeoffs between performance and selection complexity. *Distributed Comput.*, 30(3):169–191.

[Levernier et al., 2020] Levernier, N., Textor, J., Bénichou, O., and Voituriez, R. (2020). Inverse square Lévy walks are not optimal search strategies for $d \geq 2$. *Phys. Rev. Lett.*, 124:080601.

[Levernier et al., 2021] Levernier, N., Textor, J., Bénichou, O., and Voituriez, R. (2021). Reply to "comment on 'inverse square lévy walks are not optimal search strategies for $d \geq 2$"'. *Phys. Rev. Lett.*, 126:048902.

[Lin et al., 2007] Lin, W., Zhixin, L., and Lei, G. (2007). Robust consensus of multi-agent systems with noise. In *2007 Chinese Control Conference*.

[Lynch et al., 2017] Lynch, N. A., Musco, C., and Parter, M. (2017). Computational tradeoffs in biological neural networks: Self-stabilizing winner-take-all networks. In Papadimitriou, C. H., editor, *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, volume 67 of *LIPIcs*, pages 15:1–15:44. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Majani et al., 1988] Majani, E., Erlanson, R., and Abu-Mostafa, Y. (1988). On the k-winners-take-all network. In Touretzky, D., editor, *Advances in Neural Information Processing Systems*, volume 1. Morgan-Kaufmann.

[Mertzios et al., 2016] Mertzios, G. B., Nikoletseas, S. E., Raptopoulos, C. L., and Spirakis, P. G. (2016). Determining majority in networks with local interactions and very small local memory. *Distributed Computing*, 30(1):1–16.

[Mitchell and Toroczkai, 2010] Mitchell, M. and Toroczkai, Z. (2010). Complexity: A guided tour. *Physics Today*, 63:47–.

[Mitropolsky et al., 2021] Mitropolsky, D., Collins, M. J., and Papadimitriou, C. H. (2021). A Biologically Plausible Parser. In *Transactions of the Association for Computational Linguistics*. arXiv: 2108.02189.

[Mitzenmacher and Upfal, 2005] Mitzenmacher, M. and Upfal, E. (2005). *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press.

[Mobilia, 2003] Mobilia, M. (2003). Does a single zealot affect an infinite group of voters? *Phys. Rev. Lett.*, 91:028701.

[Mobilia et al., 2007] Mobilia, M., Petersen, A., and Redner, S. (2007). On the role of zealotry in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08029.

[Moon, 2005] Moon, T. K. (2005). *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley-Interscience, USA.

[Mossel et al., 2014] Mossel, E., Neeman, J., and Tamuz, O. (2014). Majority dynamics and aggregation of information in social networks. *Auton. Agents Multi Agent Syst.*, 28(3):408–429.

[Mossel and Tamuz, 2017] Mossel, E. and Tamuz, O. (2017). Opinion exchange dynamics. *Probability Surveys*, 14:155–204.

[Mukhopadhyay et al., 2020] Mukhopadhyay, A., Mazumdar, R. R., and Roy, R. (2020). Voter and majority dynamics with biased and stubborn agents. *Journal of Statistical Physics*.

[Musco et al., 2017] Musco, C., Su, H., and Lynch, N. A. (2017). Ant-inspired density estimation via random walks. *Proc. Natl. Acad. Sci. USA*, 114(40):10534–10541.

[Palyulin et al., 2019] Palyulin, V. V., Blackburn, G., Lomholt, M. A., Watkins, N. W., Metzler, R., Klages, R., and Chechkin, A. V. (2019). First passage and first hitting times of Lévy flights and Lévy walks. *New Journal of Physics*, 21(10):103028.

[Panov, 2017] Panov, A. I. (2017). Behavior planning of intelligent agent with sign world model. *Biologically Inspired Cognitive Architectures*, 19:21–31.

[Papadimitriou and Vempala, 2019] Papadimitriou, C. H. and Vempala, S. S. (2019). Random projection in the brain and computation with assemblies of neurons. In Blum, A., editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 57:1–57:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Papadimitriou et al., 2020] Papadimitriou, C. H., Vempala, S. S., Mitropolsky, D., Collins, M., and Maass, W. (2020). Brain computation by assemblies of neurons. *Proceedings of the National Academy of Sciences*, 117(25):14464–14472.

[Pettarin et al., 2011] Pettarin, A., Pietracaprina, A., Pucci, G., and Upfal, E. (2011). Tight bounds on information dissemination in sparse mobile networks. In *Proceedings of the 30th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing*, PODC '11, page 355–362, New York, NY, USA. Association for Computing Machinery.

[Raichlen et al., 2014] Raichlen, D. A., Wood, B. M., Gordon, A. D., Mabulla, A. Z., Marlowe, F. W., and Pontzer, H. (2014). Evidence of Lévy walk foraging patterns in human hunter–gatherers. *Proceedings of the National Academy of Sciences*, 111(2):728–733.

[Razin et al., 2013] Razin, N., Eckmann, J.-P., and Feinerman, O. (2013). Desert ants achieve reliable recruitment across noisy interactions. *Journal of the Royal Society Interface*, 10(82):20130079.

[Reina et al., 2017] Reina, A., Marshall, J. A. R., Trianni, V., and Bose, T. (2017). Model of the best-of-n nest-site selection process in honeybees. *Phys. Rev. E*, 95:052411.

[Reynolds et al., 2017] Reynolds, A., Santini, G., Chelazzi, G., and Focardi, S. (2017). The Weierstrassian movement patterns of snails. *Royal Society open science*, 4(6):160941.

[Reynolds, 2018] Reynolds, A. M. (2018). Current status and future directions of Lévy walk research. *Biology Open*, 7(1):bio030106.

[Reynolds et al., 2007] Reynolds, A. M., Smith, A. D., Reynolds, D. R., Carreck, N. L., and Osborne, J. L. (2007). Honeybees perform optimal scale-free searching flights when attempting to locate a food source. *Journal of Experimental Biology*, 210(21):3763–3770.

[Ruan and Mostofi, 2008] Ruan, Y. and Mostofi, Y. (2008). Binary consensus with soft information processing in cooperative networks. In *Proceedings of the 47th IEEE Conference on Decision and Control, CDC 2008*. IEEE.

[Rueckert et al., 2016] Rueckert, E., Kappel, D., Tanneberg, D., Pecevski, D., and Peters, J. (2016). Recurrent spiking networks solve planning tasks. *Scientific Reports*, 6.

[Seeley et al., 2012] Seeley, T. D., Visscher, P. K., Schlegel, T., Hogan, P. M., Franks, N. R., and Marshall, J. A. R. (2012). Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science*, 335(6064):108–111.

[Shlesinger and Klafter, 1986] Shlesinger, M. F. and Klafter, J. (1986). Lévy walks versus Lévy flights. In *On Growth and Form: Fractal and Non-Fractal Patterns in Physics*, pages 279–283. Springer Netherlands, Dordrecht.

[Sims et al., 2008] Sims, D. W., Southall, E. J., Humphries, N. E., Hays, G. C., Bradshaw, C. J., Pitchford, J. W., James, A., Ahmed, M. Z., Brierley, A. S., Hindell, M. A., et al. (2008). Scaling laws of marine predator search behaviour. *Nature*, 451(7182):1098–1102.

[Slaney and Thiébaux, 2001] Slaney, J. and Thiébaux, S. (2001). Blocks world revisited. *Artificial Intelligence*, 125(1):119–153.

[Sumpter et al., 2008] Sumpter, D. J., Krause, J., James, R., Couzin, I. D., and Ward, A. J. (2008). Consensus decision making by fish. *Current Biology*, 18(22):1773–1777.

[Tian et al., 2020] Tian, M., Luo, T., and Cheung, H. (2020). The development and measurement of block construction in early childhood: A review. *Journal of Psychoeducational Assessment*, 38(6):767–782.

[Uchiyama, 2011] Uchiyama, K. (2011). The first hitting time of a single point for random walks. *Electron. J. Probab.*, 16:1960–2000.

[Vicsek et al., 1995] Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75:1226–1229.

[Viswanathan et al., 1996] Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., and Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature*, 381(6581):413–415.

[Viswanathan et al., 1999] Viswanathan, G. M., Buldyrev, S. V., Havlin, S., da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (1999). Optimizing the success of random searches. *Nature*, 401(6756):911–914.

[Viswanathan et al., 2011] Viswanathan, G. M., Luz, M. G. E. d., Raposo, E. P., and Stanley, H. E. (2011). *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters*. Cambridge University Press, Cambridge ; New York.

[Viswanathan et al., 2008] Viswanathan, G. M., Raposo, E. P., and da Luz, M. G. E. (2008). Lévy flights and superdiffusion in the context of biological encounters and random searches. *Physics of Life Reviews*, 5(3):133–150.

[Wang and Liu, 2009] Wang, L. and Liu, Z. (2009). Robust consensus of multi-agent systems with noise. *Sci. China Ser. F Inf. Sci.*

[Werner and Hall, 1974] Werner, E. E. and Hall, D. J. (1974). Optimal foraging and the size selection of prey by the bluegill sunfish (lepomis macrochirus). *Ecology*, 55(5):1042–1052.

[Winograd, 1971] Winograd, T. (1971). Procedures as a representation for data in a computer program for understanding natural language.

[Yildiz et al., 2013] Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A., and Scaglione, A. (2013). Binary opinion dynamics with stubborn agents. *ACM Trans. Econ. Comput.*, 1(4).

[Zaburdaev et al., 2015] Zaburdaev, V., Denisov, S., and Klafter, J. (2015). Lévy walks. *Rev. Mod. Phys.*, 87:483–530.

List of figures

List of definitions

Appendix

A Chain rule and union bound

Lemma A.1. *Let $\eta > \lambda > 0$ be two constants, and $0 < p \leq 1$ be a probability. Consider any family of events $\{\xi_i\}_{i \leq M}$ with $M > 1$ being some integer. Suppose $Pr[\xi_1] \geq p$, and, for $i \geq 2$, that $Pr[\xi_i \mid \xi_1, \ldots, \xi_{i-1}] = Pr[\xi_i \mid \xi_{i-1}] \geq p$. The following holds.*

(i) If $p = 1 - n^{-\eta}$ and $M \leq n^{\lambda}$, then $\cap_{i \leq M} \xi_i$ holds with probability $1 - \mathcal{O}\left(n^{-(\eta-\lambda)}\right)$.

(ii) If $p = 1 - \exp(-\eta n)$ and $M \leq e^{\lambda}$. Then $\cap_{i \leq M} \xi_i$ holds with probability $1 - \mathcal{O}\left(e^{-(\eta-\lambda)n}\right)$.

Proof. We have that

$$\begin{aligned}
\Pr[\cap_{i \leq M} \xi_i] &= \Pr[\xi_M \mid \cap_{i \leq M-1} \xi_i] \cdot \Pr[\cap_{i \leq M-1} \xi_i] \\
&= \Pr[\xi_M \mid \xi_{M-1}] \cdot \Pr[\cap_{i \leq M-1} \xi_i] \\
&= \prod_{i=2}^{M} \Pr[\xi_i \mid \xi_{i-1}] \cdot \Pr[\xi_1] \\
&\geq p^M
\end{aligned}$$

Now, let $f(n) = 1 - p = o(1)$. Notice that $1 - x \geq e^{-x/(1-x)}$ for $|x| < 1$. Then,

$$\begin{aligned}
p^M &= [1 - f(n)]^M \\
&\geq \exp\left(-\frac{Mf(n)}{1 - f(n)}\right) \\
&\geq \exp(-2Mf(n)) \\
&\overset{(a)}{\geq} 1 - 4Mf(n),
\end{aligned}$$

where (a) holds of the exponential function Taylor's expansion, since $Mf(n) = o(1)$ by the hypotheses. As for (i), we get

$$p^M \geq 1 - \mathcal{O}\left(\frac{n^{\lambda}}{n^{\eta}}\right) = 1 - \mathcal{O}\left(\frac{1}{n^{\eta-\lambda}}\right).$$

As for (ii), we get

$$p^M \geq 1 - \mathcal{O}\left(\frac{\exp(\lambda n)}{\exp(\eta n)}\right) = 1 - \mathcal{O}\left(e^{-(\eta-\lambda)n}\right).$$

□

More easily, the union bound implies the following.

Lemma A.2. *Let $\eta > \lambda > 0$ be two constants, and $0 < p \leq 1$ be a probability. Consider any family of events $\{\xi_i\}_{i \leq M}$ with $M > 1$ being some integer. Suppose $Pr[\xi_i] \geq p$ for all i. The following holds.*

(i) If $p = 1 - n^{-\eta}$ and $M \leq n^{\lambda}$, then $\cap_{i \leq M} \xi_i$ holds with probability $1 - n^{-(\eta-\lambda)}$.

(ii) If $p = 1 - \exp(-\eta n)$ and $M \leq e^{\lambda}$. Then $\cap_{i \leq M} \xi_i$ holds with probability $1 - e^{-(\eta-\lambda)n}$.

B Concentration bounds

We recall the concentration results that we use in the analysis. For an overview on the forms of Chernoff bounds see [Dubhashi and Panconesi, 2009] or [Doerr, 2018].

Lemma B.1 (Multiplicative forms of Chernoff bounds)**.** *Let $X_1, X_2, \dots, X_n$ be independent $\{0,1\}$ random variables. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}[X]$. Then:*

(i) for any $\delta \in (0,1)$ and $\mu \leq \mu_+ \leq n$, it holds that

$$Pr\left[X \geq (1+\delta)\mu_+\right] \leq e^{-\frac{1}{3}\delta^2\mu_+}, \tag{B.1}$$

(ii) for any $\delta \in (0,1)$ and $0 \leq \mu_- \leq \mu$, it holds that

$$Pr\left[X \leq (1-\delta)\mu_-\right] \leq e^{-\frac{1}{2}\delta^2\mu_-}. \tag{B.2}$$

Lemma B.2 (Additive forms of Chernoff bounds)**.** *Let $X_1, X_2, \dots, X_n$ be independent $\{0,1\}$ random variables. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}[X]$. Then:*

(i) for any $0 < \lambda < n$ and $\mu \leq \mu_+ \leq n$, it holds that

$$Pr\left[X \geq \mu_+ + \lambda\right] \leq e^{-\frac{2}{n}\lambda^2}, \tag{B.3}$$

(ii) for any $0 < \lambda < \mu_-$ and $0 \leq \mu_- \leq \mu$, it holds that

$$Pr\left[X \leq \mu_- - \lambda\right] \leq e^{-\frac{2}{n}\lambda^2}. \tag{B.4}$$

We also make use of the Hoeffding bounds [Mitzenmacher and Upfal, 2005].

Lemma B.3 (Hoeffding bounds)**.** *Let $0 < a < b$ be two constants. Let $X_1, X_2, \dots, X_n$ be independent random variables such that $Pr\left[a \leq X_i \leq b\right] = 1$ and $\mathbb{E}\left[X_i\right] = \mu/n$ for all $i \leq n$, and let $X = \sum_{i=1}^{n} X_i$. Then:*

(i) for any $\lambda > 0$ and $\mu \leq \mu_+$, it holds that

$$Pr\left[X \geq \mu_+ + \lambda n\right] \leq \exp\left(-\frac{2\lambda^2 n}{(b-a)^2}\right); \tag{B.5}$$

(ii) for any $\lambda > 0$ and $0 \leq \mu_- \leq \mu$, it holds that

$$Pr\left[X \leq \mu_- - \lambda n\right] \leq \exp\left(-\frac{2\lambda^2 n}{(b-a)^2}\right). \tag{B.6}$$

Additive Chernoff bound using variance [Chung and Lu, 2006, Theorem 3.4].

Lemma B.4. *Let $X_1, \dots, X_n$ be independent random variables satisfying $X_i \leq \mathbb{E}\left[X_i\right] + M$ for some $M \geq 0$, for all $i = 1, \dots, n$. Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}\left[X\right]$, and $\sigma^2 = \mathrm{Var}\left(X\right)$. Then, for any $\lambda > 0$,*

$$Pr\left[X \geq \mu + \lambda\right] \leq \exp\left(-\frac{\lambda^2}{\sigma^2 + \frac{M\lambda}{3}}\right). \tag{B.7}$$

Central limit theorem [Feller, 1968, Chapter X].

Lemma B.5. *Let $\{X_k\}_{k\geq 1}$ be a sequence of i.i.d. real random variables. Let $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \mathrm{Var}(X_1)$, and $S_n = \sum_{k=1}^{n} X_k$ for any $n \geq 1$. Let $\Phi : \mathbb{R} \to [0,1]$ be the cumulative distribution function of a standard normal distribution. Then, for any $\beta \in \mathbb{R}$,*

$$\lim_{n\to\infty} Pr\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \beta\right] = \Phi(\beta).$$

The Berry-Esseen theorem is well treated in [Korolev and Shevtsova, 2010], and it gives an estimation on "how far" is the distribution of the normalized sum of i.i.d. random variables to the standard normal distribution.

Lemma B.6 (Berry-Esseen)**.** *Let $\{X_k\}_{k\geq 1}$ be a sequence of i.i.d. real random variables. Let $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \mathrm{Var}(X_1)$, and $S_n = \sum_{k=1}^{n} X_k$ for any $n \geq 1$. Let $\Phi : \mathbb{R} \to [0,1]$ be the cumulative distribution function of a standard normal distribution. Then, there exists a positive constant $C > 0$ such that*

$$\sup_{x\in\mathbb{R}} \left| Pr\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] - \Phi(x)\right| \leq \frac{C}{\sqrt{n}}$$

for all $n \geq 1$.

C Anti-concentration inequalities

Finally, we use some anti-concentration inequalities known as reverse Chernoff bounds. The proof can be found in the appendix of [Klein and Young, 2015].

Lemma C.1 (Reverse Chernoff bounds)**.** *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\{0,1\}$ random variables. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}[X]$, with $\mu \leq n/2$. Furthermore, let $\delta \in (0, 1/2]$ be a constant. If $\delta^2\mu \geq 3$, then:*

(i) for any $\mu \leq \mu_+ \leq n$, it holds that

$$Pr[X \geq (1+\delta)\mu_+] \geq e^{-9\delta^2\mu_+}; \tag{C.8}$$

(ii) for any $0 \leq \mu_- \leq \mu$, it holds that

$$Pr[X \leq (1-\delta)\mu_-] \geq e^{-9\delta^2\mu_-}. \tag{C.9}$$

D Hitting time for (sub/super)-martingales

We make use of the following general result on (super/sub)-martingales, which can be found in [Lehre and Witt, 2014].

Lemma D.1. *Let $\{X_t\}_{t\in\mathbb{N}}$ be a stochastic process adapted to a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{N}}$, over some state space $S \subseteq \{0\} \cup [x_{min}, x_{max}]$, where $x_{min} \geq 0$. Let $h : [x_{min}, x_{max}] \to \mathbb{R}^+$ be a function such that $1/h(x)$ is integrable and $h(x)$ differentiable on $[x_{min}, x_{max}]$. Define $T := \min\{t \in \mathbb{N} \mid X_t = 0\}$. Then, the following hold.*

(i) *If* $\mathbb{E}[X_t - X_{t+1} \mid X_t \geq x_{min}, \mathcal{F}_t] \geq h(X_t)$ *for all* $t \in \mathbb{N}$ *and* $\frac{\mathrm{d}h(x)}{\mathrm{d}x} \geq 0$ *for all* $x \in [x_{min}, x_{max}]$, *then*

$$\mathbb{E}[T \mid X_0] \leq \frac{x_{min}}{h(x_{min})} + \int_{x_{min}}^{X_0} \frac{1}{h(y)} \, \mathrm{d}y.$$

(ii) *If* $\mathbb{E}[X_t - X_{t+1} \mid X_t \geq x_{min}, \mathcal{F}_t] \leq h(X_t)$ *for all* $t \in \mathbb{N}$ *and* $\frac{\mathrm{d}h(x)}{\mathrm{d}x} \leq 0$ *for all* $x \in [x_{min}, x_{max}]$, *then*

$$\mathbb{E}[T \mid X_0] \geq \frac{x_{min}}{h(x_{min})} + \int_{x_{min}}^{X_0} \frac{1}{h(y)} \, \mathrm{d}y.$$

(iii) *If* $\mathbb{E}[X_t - X_{t+1} \mid X_t \geq x_{min}, \mathcal{F}_t] \geq h(X_t)$ *for all* $t \in \mathbb{N}$ *and* $\frac{\mathrm{d}h(x)}{\mathrm{d}x} \geq \lambda$ *for some* $\lambda > 0$ *and all* $x \in [x_{min}, x_{max}]$, *then*

$$Pr[T > t \mid X_0] < \exp\left(-\lambda\left(t - \frac{x_{min}}{h(x_{min})} - \int_{x_{min}}^{X_0} \frac{1}{h(y)} \, \mathrm{d}y\right)\right).$$

(iv) *If* $\mathbb{E}[X_t - X_{t+1} \mid X_t \geq x_{min}, \mathcal{F}_t] \leq h(X_t)$ *for all* $t \in \mathbb{N}$ *and* $\frac{\mathrm{d}h(x)}{\mathrm{d}x} \leq -\lambda$ *for some* $\lambda > 0$ *and all* $x \in [x_{min}, x_{max}]$, *then*

$$Pr[T < t \mid X_0] < \frac{e^{\lambda t} - e^{\lambda}}{e^{\lambda} - 1} \exp\left(-\lambda\left(\frac{x_{min}}{h(x_{min})} + \int_{x_{min}}^{X_0} \frac{1}{h(y)} \, \mathrm{d}y\right)\right).$$

E Integral test

Lemma E.1. *Let* $0 < d < d_{max}$ *be any integers. For any* $\alpha > 1$,

$$\frac{1}{(\alpha-1)(d)^{\alpha-1}} \leq \sum_{k \geq d} \frac{1}{k^{\alpha}} \leq \frac{1}{(\alpha-1)(d)^{\alpha-1}} + \frac{1}{d^{\alpha}}, \quad \textit{and} \tag{E.10}$$

$$\frac{1}{(\alpha-1)}\left(\frac{1}{d^{\alpha-1}} - \frac{1}{d_{max}^{\alpha-1}}\right) \leq \sum_{k=d}^{d_{max}} \frac{1}{k^{\alpha}} \leq \frac{1}{(\alpha-1)}\left(\frac{1}{d^{\alpha-1}} - \frac{1}{d_{max}^{\alpha-1}}\right) + \frac{1}{d^{\alpha}}. \tag{E.11}$$

Also,

$$\log\left(\frac{d_{max}}{d}\right) \leq \sum_{k=d}^{d_{max}} \frac{1}{k} \leq \log\left(\frac{d_{max}}{d}\right) + \frac{1}{d}, \tag{E.12}$$

and for any $0 < \alpha < 1$,

$$\frac{(d_{max})^{1-\alpha} - d^{1-\alpha}}{1-\alpha} \leq \sum_{k=d}^{d_{max}} \frac{1}{k^{\alpha}} \leq \frac{(d_{max})^{1-\alpha} - d^{1-\alpha}}{1-\alpha} + \frac{1}{d^{\alpha}}. \tag{E.13}$$

Proof. By the integral test, it holds that

$$\int_{d}^{d_{\max}} \frac{1}{k^{\alpha}} dk \leq \sum_{k=d}^{d_{\max}} \frac{1}{k^{\alpha}} \leq \int_{d}^{d_{\max}} \frac{1}{k^{\alpha}} dk + \frac{1}{d^{\alpha}}.$$

Straightforward calculations give the result for Eqs. (E.11) to (E.13). As for Eq. (E.10), it comes from the integral test letting $d_{\max} \to \infty$. □

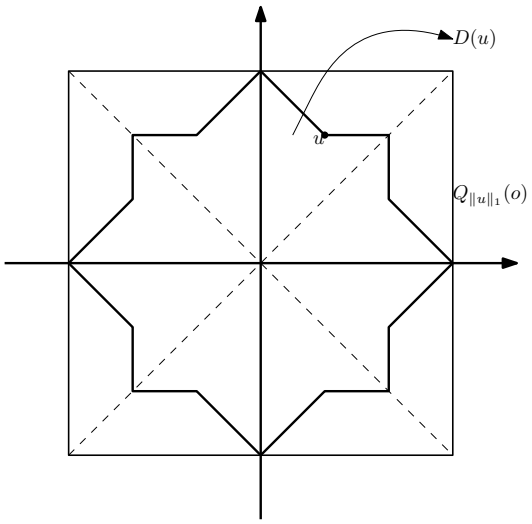


Figure F.1 – The set $D(u)$, consisting in all inner nodes of the "star", and the square $Q_{d_u}(\mathbf{0})$.

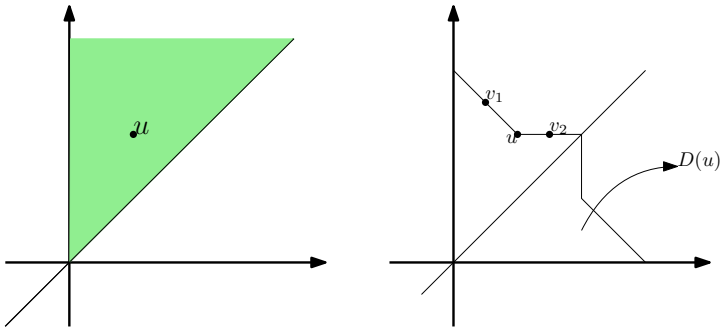


Figure F.2 – The "area" in which we take u, and the possible choices of v.

F Omitted proofs from Chapter 3

F.1 Proof of Lemma 3.2.3

For any node w, define $D(w)$ as the set $B_{\|w\|_1}(\mathbf{0}) \cup Q_{\|w\|_\infty}(\mathbf{0})$. Notice that $D(u) \subseteq Q_{\|u\|_1}(\mathbf{0})$. Then it suffice to prove that for all nodes $v \notin D(u)$ we have

$$\Pr[J_t = u] \geq \Pr[J_t = v].$$

Let $u = (x_u, y_u)$ and, without loss of generality, suppose u is in the first quadrant and not below the main bisector, i.e., in the set $\{(x, y) \in \mathbb{Z}^2 : y \geq 0, x \geq y\}$ (Fig. F.2). If we show that, for any v in $\{v_1 = (x_u - 1, y_u + 1), v_2 = (x_u + 1, y_u)\}$ (Fig. F.2)., we have $\Pr[J_t = u] \geq \Pr[J_t = v]$, than we have the statement. Indeed, for any $v \notin D(u)$ that "lives" in the highlighted area in Fig. F.2, there exists a sequence of nodes $u = w_0, w_1, \ldots, w_k = v$ from u to v such that w_{i+1} belongs to the set

$$\{(x_{w_i} - 1, y_{w_i} + 1), (x_{w_i} + 1, y_{w_i})\},$$

where $w_i = (x_{w_i}, y_{w_i})$, as Fig. F.3 shows. Thus, if the thesis is true for $v \in \{v_1, v_2\}$, then it is true also for all $v \notin D(u)$ in the highlighted area in Fig. F.2. At the same time, for any other $v \notin D(u)$,

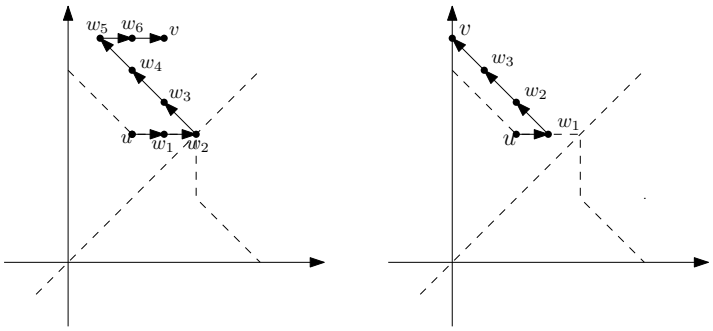


Figure F.3 – Two "path" examples.

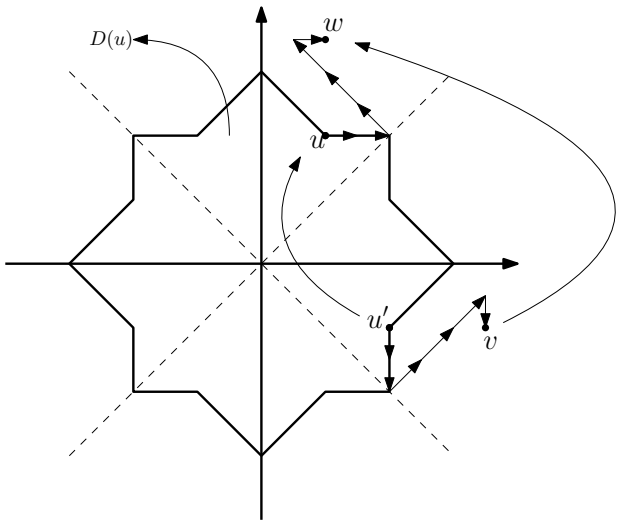


Figure F.4 – Symmetrical argument.

outside the highlighted area in Fig. F.2, there exists a symmetrical argument explained in Fig. F.4. Thus, if the thesis is true for all $v \notin D(u)$ in the highlighted area in Fig. F.2, then it is also true for any $v \notin D(u)$. We now consider some geometric constructions which will be used in the proof, one for each choice of v. The following description is showed in Fig. F.5 .

(i) $v = (x_u - 1, y_u + 1)$: consider the strict line defined by $r : y = x + (y_u - x_u) + 1$ (i.e., the line in $\mathbb{R}^2$ which is the set of points that are equidistant from u and v in Euclidean distance). Call $V \subset \mathbb{Z}^2$ the set of nodes that are "above" this line, namely the ones that are closer to v than u. Define $U = \mathbb{Z}^2 \backslash (V \cup r)$ the complementary set without line r. Consider the injective function $f : V \to U$ such that $f(x, y) = (y - (y_u - x_u) - 1, x + (y_u - x_u) + 1)$, which is the symmetry with respect to r. It trivially holds that for any $w \in V$, $\|w - v\|_1 = \|f(w) - u\|_1$ and $\|w - u\|_1 = \|f(w) - v\|_1$. Furthermore, it holds that for each $w \in V$, either $w \notin D(f(w))$, or w lies on the "border" of $D(f(w))$. All these properties are well-shown in Fig. F.6.

(ii) $v = (x_u + 1, y_u)$: the same construction can be done in this case. Indeed, the strict line will be $x = x_u + \frac{1}{2}$, and the injective function $f(x, y) = (2x_u + 1 - x, y)$. The same properties we have seen in the previous case hold here too.

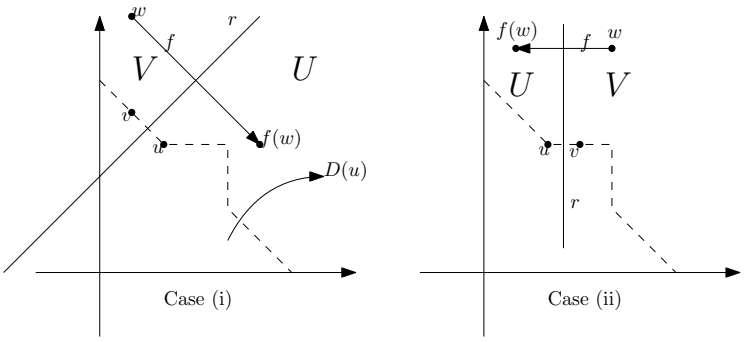


Figure F.5 – Geometric constructions in the two cases.

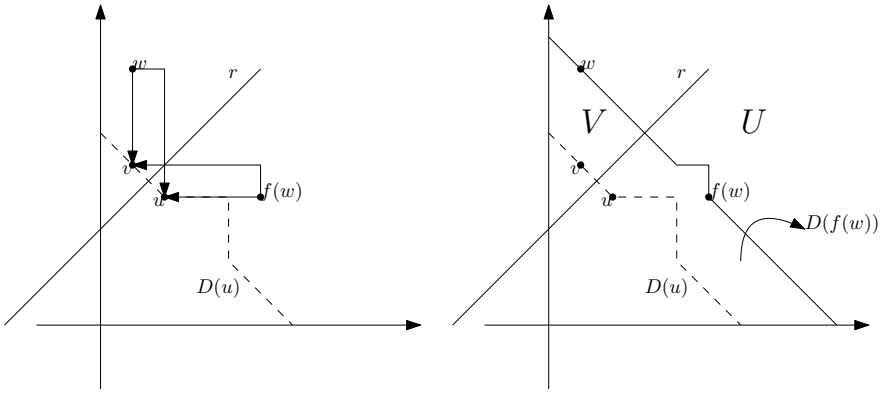


Figure F.6 – Left: equivalence between distances. Right: $w \notin D(f(w))$.

Now we go for the proof. For any time i, and any two nodes $u', v' \in \mathbb{Z}^2$, define

$$p^i(u', v') = \Pr\left[\mathbf{X}_i = v' \mid \mathbf{X}_0 = u'\right].$$

Let $v \in \{v_1, v_2\}$. We show that $p^t(\mathbf{0}, u) \geq p^t(\mathbf{0}, v)$ by induction on t. The base case is $t = 1$. From the monotonicity, we know that

$$p^1(\mathbf{0}, u) - p^1(\mathbf{0}, v) \geq 0$$

for any u and v in $\mathbb{Z}^2$ such that $\|u\|_1 \leq \|v\|_1$. We now suppose $t \geq 2$ and the thesis true for $t-1$. Fix u and v as in Fig. F.2; then, for the geometric construction we made above, it holds that

$$\begin{aligned} p^t(\mathbf{0}, u) - p^t(\mathbf{0}, v) &= \sum_{w \in \mathbb{Z}^2} p^{t-1}(\mathbf{0}, w) \left(p^1(w, u) - p^1(w, v)\right) \\ &\geq \sum_{w \in U} p^{t-1}(\mathbf{0}, w) \left(p^1(w, u) - p^1(w, v)\right) + \sum_{w \in V} p^{t-1}(\mathbf{0}, w) \left(p^1(w, u) - p^1(w, v)\right) \end{aligned}$$

where last inequality is immediate for case (ii), indeed the line r does not contain elements of $\mathbb{Z}^2$, while in case (i) the sum over nodes in line r is zero. Then, the previous value is equal to

$$\sum_{w \in V} p^{t-1}(\mathbf{0}, f(w)) \left(p^1(f(w), u) - p^1(f(w), v)\right) + \sum_{w \in V} p^{t-1}(\mathbf{0}, w) \left(p^1(w, u) - p^1(w, v)\right)$$

because of the definition of $f : V \to U$, and, changing the sign of the second sum, we obtain

$$\sum_{w \in V} p^{t-1}(\mathbf{0}, f(w)) \left(p^1(f(w), u) - p^1(f(w), v) \right) - \sum_{w \in V} p^{t-1}(\mathbf{0}, w) \left(p^1(w, v) - p^1(w, u) \right).$$

Now, observe that the definition of f implies that for each $w \in V$, $\|w - v\|_1 = \|f(w) - u\|_1$ and $\|f(w) - v\|_1 = \|w - u\|_1$ (Fig. F.6). Thus we can group out the term $p^1(f(w), u) - p^1(f(w), v) = p^1(w, v) - p^1(w, u)$, and we have

$$\sum_{w \in V} \left(p^{t-1}(\mathbf{0}, f(w)) - p^{t-1}(\mathbf{0}, w) \right) \left(p^1(f(w), u) - p^1(f(w), v) \right). \tag{F.14}$$

We observe that $p^{t-1}(\mathbf{0}, f(w)) - p^{t-1}(\mathbf{0}, w) \geq 0$ by the inductive hypothesis, since either $w \notin D(f(w))$ or w lies on the "border" of $D(f(w))$ (Fig. F.6), and $p^1(f(w), u) - p^1(f(w), v) \geq 0$ by definition of f, since the distance between $f(w)$ and u is no more than the distance between $f(w)$ and v. It follows that (F.14) is non-negative, and, thus, the thesis.

F.2 Projection of a Lévy Flight Jump

Let L_t^f be the two dimensional random variable representing the coordinates of an agent performing an α-Lévy flight at time t, for any $\alpha > 1$. Consider the projection of the Lévy flight on the x-axis, namely the random variable $X_{t'}$ such that $L_t^f = (X_t, Y_t)$. The random variable X_t can be expressed as the sum of t random variables S_j^x, $j = 1, \ldots, t$, representing the projection of the jumps (with sign) of the agent on the x-axis at times $j = 1, \ldots, t$. With the next lemma, we prove that the jump projection length has the same tail distribution as the original jump length.

Lemma F.1. *The probability that a jump S_j^x has length equal to d is $\Theta(1/d^\alpha)$.*

Proof. The partial distribution of the jumps along the x-axis is given by the following. For any $d \geq 0$,

$$\Pr\left[S_j^x = \pm d\right] = \left[\frac{1}{2} + \sum_{k=1}^{\infty} \frac{c_\alpha}{2k^{\alpha+1}}\right] \mathbb{1}_{d=0} + \left[\frac{c_\alpha}{2d^{\alpha+1}} + \sum_{k=1+d}^{\infty} \frac{c_\alpha}{k^{\alpha+1}}\right] \mathbb{1}_{d \neq 0}, \tag{F.15}$$

where $\mathbb{1}_{d \in A}$ returns 1 if $d \in A$ and 0 otherwise, the term

$$\frac{\mathbb{1}_{d=0}}{2} + \frac{c_\alpha}{2d^{\alpha+1}} \mathbb{1}_{d \neq 0}$$

is the probability that the original jump lies along the horizontal axis and has "length" exactly d (there are two such jumps if $d > 0$), and, for $k \geq 1 + d$, the terms

$$\frac{c_\alpha}{2k^{\alpha+1}} \mathbb{1}_{d=0} + \frac{c_\alpha}{k^{\alpha+1}} \mathbb{1}_{d \neq 0}$$

are the probability that the original jump has "length" exactly k and its projection on the horizontal axis has "length" d (there are two such jumps if $d = 0$, and four such jumps if $d > 0$). By the integral test (Lemma E.1) we know that quantity (F.15) is

$$\Pr\left[S_j^x = \pm d\right] = \Theta\left(\frac{1}{d^{\alpha+1}}\right).$$

□

Sur les Comportements Collectifs de Systèmes Distribués Bio-Inspirés

Francesco D'AMORE

Résumé

Récemment, la communauté algorithmique a manifesté un intérêt croissant pour l'utilisation de ses outils théoriques à la compréhension des systèmes complexes, notamment biologiques, tels que les colonies d'insectes, les volées d'oiseaux et les réseaux de neurones. Nous contribuons à l'étude de ces systèmes dans trois directions différentes. Premièrement, nous analysons des dynamiques computationnelles pour les tâches de coordination stochastique dans les systèmes multi-agents. En particulier, nous nous focalisons sur le problème du consensus dans des environnements où la communication est bruyante : nous analysons deux dynamiques d'opinion, les dynamiques UNDECIDED-STATE et 3-MAJORITY, et nous prouvons qu'elles présentent une transition de phase à des seuils de bruit différents. En dessous du seuil, ces dynamiques atteignent rapidement une phase métastable de quasi-consensus ; au-dessus, aucune forme de consensus n'est possible. Deuxièmement, nous étudions les Lévy walks, des marches aléatoires qui modélisent des schémas de mouvement trouvés dans la nature, dont la distribution de la longueur de pas suit une loi de puissance. Nous analysons leur temps d'arrêt (hitting time) parallèle et les utilisons pour concevoir un algorithme optimal pour l'ANTS problem, un problème de recherche distribuée sur $\mathbb{Z}^2$ qui capture certains aspects de la théorie du butinage. Troisièmement, nous considérons l'Assembly Calculus, un modèle distribué du cerveau récemment proposé, qui consiste en des neurones et des synapses stylisés, et nous testons expérimentalement ses capacités, largement inexplorées, en mettant en œuvre des heuristiques connues pour la tâche de planification du monde des blocs. Nous montrons empiriquement que des programmes grands et complexes dans ce modèle s'exécutent correctement et de manière fiable.

Mots-clés : Calcul distribué, Algorithmes naturels, Systèmes biologiques.

Abstract

In recent years there has been a surge of interest on behalf of the algorithmic community in applying its theoretical tools to the understanding of complex systems, in particular biological ones, such as insect colonies, flocks of birds, and networks of neurons. We contribute to the investigation of such systems in three different directions. First, we analyze computational dynamics for stochastic coordination tasks in multi-agent systems: in particular, we focus on the consensus problem in environments where communication is affected by some form of noise. In this setting, we analyze two known opinion dynamics, the UNDECIDED-STATE and the 3-MAJORITY dynamics, and prove that they exhibit a phase-transition at different noise thresholds. Below the threshold, the two dynamics quickly reach an almost-consensus metastable phase; above, no form of consensus is possible. Second, we study Lévy walks, i.e., special random walks known to model many movement patterns found in nature, characterized by a step-length density distribution proportional to a power-law. We analyze their parallel hitting time and show how to use them to design an almost optimal algorithm for the ANTS problem, a distributed search problem on $\mathbb{Z}^2$ which captures some aspects of animal foraging theory. Third, we consider the Assembly Calculus, a recently proposed distributed computational model of the brain which consists of stylized spiking neurons and synapses, and we test experimentally its capabilities, largely unexplored. In particular, we implement known heuristics for the Blocks World planning task; we empirically prove that reasonably large and complex programs in the Assembly Calculus run correctly and reliably.

Keywords: Distributed computing, Natural algorithms, Biological systems.