UNIVERSITY OF CALIFORNIA, SAN DIEGO

Distributed Resource Allocation and Optimization Algorithms Applied to Virus Spread Minimization

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Aerospace Engineering)

by

Eduardo Ramírez-Llanos

Committee in charge:

Professor Sonia Martínez, Chair
Professor Jorge Cortes
Professor Mauricio de Oliveira
Professor Jiawang Nie
Professor Alvaro Pelayo
Professor Sandip Roy

2017
The dissertation of Eduardo Ramírez-Llanos is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2017
DEDICATION

To Johi. I am truly thankful for having you in my life.
EPIGRAPH

The good thing about science is that it's true whether or not you believe in it.

—Neil deGrasse Tyson
# TABLE OF CONTENTS

Signature Page ......................................................... iii  
Dedication .............................................................. iv  
Epigraph ................................................................. v  
Table of Contents ...................................................... vi  
List of Figures .......................................................... ix  
Acknowledgements ..................................................... xi  
Vita .................................................................. xiv  
Abstract of the Dissertation ......................................... xvi  
Chapter 1  Introduction ............................................... 1  
Chapter 2  Preliminaries and notation ............................... 8  
    2.0.1 On the replicator dynamics ............................... 9  
    2.0.2 Partial stability for nonsmooth Lyapunov functions . 11  
    2.0.3 Set-valued analysis ...................................... 14  
    2.0.4 Lipschitz-type stability in nonsmooth convex programs 15  
    2.0.5 Stability for stochastic difference inclusions .......... 16  
    2.0.6 Convex analysis notions ................................ 19  
    2.1 Contributions of this thesis ............................... 20  
Chapter 3  A distributed dynamics for virus-spread control .... 22  
    3.1 Problem Formulation ....................................... 23  
        3.1.1 Problem statement .................................... 23  
        3.1.2 Solution characterization for balanced matrices . 26  
        3.1.3 Solution bound for unbalanced matrices .......... 30  
    3.2 The constrained Euler replicator algorithm ............... 32  
    3.3 Stability analysis ........................................... 34  
    3.4 Simulations .................................................. 41  
    3.5 Summary ..................................................... 44  
Chapter 4  Distributed and robust fair optimization for resource allocation .... 46  
    4.1 Problem statement, solution approach, and algorithms .. 47  
        4.1.1 Problem statement and solution approach .......... 47  
        4.1.2 Proposed algorithms ................................ 54  
        4.1.3 Stability analysis .................................... 55  

### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Evolution of $\delta(k)$ and the eigenvalues of $A(\delta(k))$ using (3.13).</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>Evolution of $\delta(k)$ and the eigenvalues of $A(\delta(k))$ using the constrained Euler replicator algorithm.</td>
<td>43</td>
</tr>
<tr>
<td>3.3</td>
<td>Trajectories of the constrained Euler replicator algorithm for the approximation to the optimum given by Lemma 3.4.</td>
<td>44</td>
</tr>
<tr>
<td>3.4</td>
<td>Trajectories of the constrained Euler replicator algorithm for the Enron network example.</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>Trajectories of $\delta(t)$ and eigenvalues $\lambda(A(\delta))$ of Example 4.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 20$.</td>
<td>77</td>
</tr>
<tr>
<td>4.2</td>
<td>Trajectories of the payoff function $\hat{h}(\delta)$ and the sum of resources of Example 4.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 20$.</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Trajectories of $\delta$ and $\hat{h}(\delta)$ of Example 4.1 for the $p$-robust box-gradient fairness algorithm. Agent 1 leaves the network at $t = 30$ and returns at $t = 50$.</td>
<td>78</td>
</tr>
<tr>
<td>4.4</td>
<td>Trajectories of $\kappa(t)$ and eigenvalue $\lambda_1(A(\kappa))$ of Example 4.2 for the $p$-robust box-gradient fairness algorithm. It is used $\epsilon = 26.16$, which satisfies (4.24).</td>
<td>79</td>
</tr>
<tr>
<td>4.5</td>
<td>Trajectories of $\kappa(t)$ and eigenvalue $\lambda_1(A(\kappa))$ of Example 4.2 for the $p$-robust box-gradient fairness algorithm. It is used $\epsilon = 10$, which does not satisfy (4.24).</td>
<td>80</td>
</tr>
<tr>
<td>4.6</td>
<td>Flight network for Example 4.3</td>
<td>81</td>
</tr>
<tr>
<td>4.7</td>
<td>Trajectories of $1_N^T \kappa(t)$ and eigenvalue $\lambda_1(A(\kappa))$ of Example 4.3.</td>
<td>81</td>
</tr>
<tr>
<td>4.8</td>
<td>Trajectories of $\kappa(t)$ for top 100 airports of Example 4.3.</td>
<td>82</td>
</tr>
<tr>
<td>5.1</td>
<td>Trajectories of $\delta(t)$ and $\lambda_1(A(\delta(t)))$ of Example 5.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 1500$.</td>
<td>119</td>
</tr>
<tr>
<td>6.1</td>
<td>Evolution of $\lambda_1(A(\delta))$ for Algorithm 3 when the distributed stopping condition for the estimation of the gradient is used in the Enron network example.</td>
<td>133</td>
</tr>
<tr>
<td>7.1</td>
<td>Evolution of the electricity market for Example 7.1. Negative trajectories correspond to the suppliers $z(k)$ and positive ones to the consumers $x(k)$. Dashed lines are the optimal consumption and production $x^<em>$ and $z^</em>$, respectively. In $k = 15 \times 10^3$ it is forced the states $x(k)$ and $z(k)$ to be zero during an interval of 100 iterations.</td>
<td>160</td>
</tr>
</tbody>
</table>
Figure 8.1: Evolution of the mobile robot for \( f = (p_1 - .9)^2 + |p_1 - .9| + (p_2 - 1)^2 + |p_2 - 1| \) with a box constraint \( p \in [0, 1]^2 \). The level sets of \( f \) are shown in colors and the trajectory of the robot is shown in black.

Figure 8.2: Evolution of the mobile robot for \( f = (p_1 - 5)^2 + |p_1 - 5| + (p_2 - 5)^2 + |p_2 - 5| \) with a box constraint \( p \in [0, 8]^2 \). The level sets of \( f \) are shown in colors and the trajectory of the robot is shown in red, the obstacles in blue.
ACKNOWLEDGEMENTS

I would especially like to thank my family for their unconditional support, love, and understanding over the years. To Johi, your incomparable love and sacrifice allow me to pursue my degree. Mom and Dad, you have been an exceptional inspiration to me. To my sister, for your encouragement that means a lot to me.

I would like to thank professor Sonia Martínez for her mentorship and technical support during the Ph.D. Sonia, I will forever be grateful for the knowledge and skills that I have gained working with you.

I thank my lab friends Ashish Cherukuri, Andres Cortés, Pedro Franco, Daniele Cavaglieri, Federico Bribiesca, Mike Ouimet, Robert Moroto, Shuxia Tang, Aman Khosa, Martin Sehr, Aamodh Suresh, and shahrouz Alimo for the academic and non-academic discussions.

This work could never have been accomplished without the generous support of Colciencias award 529-2011.

Finally, I would like to acknowledge my committee members, Professor Jorge Cortes, Professor Sandip Roy, Professor Mauricio de Oliveira, Professor Alvaro Pelayo, and Professor Jiawang Nie.

Coauthor acknowledgments: This thesis contains parts of the following publications with Professor Sonia Martínez as a coauthor:

tions of Automatic Control.


• E. Ramírez-Llanos and S. Martínez, “Distributed and robust fair resource allocation applied to virus spread minimization”, ACC, Chicago, IL, USA, 2015


At the end of each chapter, we specify the publication from where the content was taken.
VITA

2006 B. S. in Electronics Engineering *Honors degree*, Pontificia Universidad Javeriana, Cali, Colombia

2009 M. Sc. in Electronics Engineering, Universidad de los Andes, Colombia

2017 Ph. D. in Aerospace Engineering, University of California, San Diego

PUBLICATIONS


E. Ramírez-Llanos and S. Martínez, “Distributed and robust fair resource allocation applied to virus spread minimization”, ACC, Chicago, IL, USA, 2015


ABSTRACT OF THE DISSERTATION

Distributed Resource Allocation and Optimization Algorithms Applied to Virus Spread Minimization

by

Eduardo Ramírez-Llanos

Doctor of Philosophy in Engineering Sciences (Aerospace Engineering)

University of California, San Diego, 2017

Professor Sonia Martínez, Chair

The proliferation of large-scale networks like social networks, transportation networks, or smartgrids imposes new demands and challenges on the design of learning algorithms for optimal resource allocation. In a typical scenario, a group of agents decides how to coordinate the use of shared resources to solve a common goal while satisfying operational and communication constraints. The challenge is how to increase the network resilience given myopic agents with access to partial information. Under these settings, there is an emergence for the design of algorithms that are scalable, ro-
bust against adversarial or unknown environments, preserve privacy, and that allow the agents to take autonomous decisions on the resource utilization.

A real-world problem leading to such a scenarios arises in computer networks, epidemiology, and viral marketing, where a viral outbreak can be a threat to the security of interconnected infrastructure and the well-being of general population. The implementation of strategies to stop epidemics can be specially challenging when networks are managed by multiple operators who need to preserve the privacy and interests of their constituents.

Motivated by this situation, we consider a resource allocation problem for virus spread minimization. Based on a general contagion dynamics model, we characterize the optimal solution to the problem. We pose the problem objective as the minimization of the spectral radius of the contagion-dynamics matrix subject to operational constraints. We propose four algorithms to find the solution with provable convergence guarantees under different settings. The first algorithm, inspired by the Replicator Dynamics, implements the desired resource allocation for time-varying symmetric matrices. The second algorithm, designed in continuous-time, uses local and anonymous interactions, does not require knowledge of the total resource available to agents in order to converge to the solution, is robust to agents joining or departing the network, and to sporadic changes in the network topology, computation errors, and communication faults. The third algorithm, which is a discrete version of the second one, conserves the robustness properties of the previous one. Finally, we propose a stochastic algorithm, which extends the previous algorithms to scenarios where the closed-form expression of
the cost functions is unknown to the agents.
Chapter 1

Introduction

The operation of large-scale systems such as social networks, transportation networks, smartgrids, or robotic networks imposes new challenges on the design of learning algorithms for optimal resource allocation. Large-scale systems are formed by a set of interdependent and interconnected entities, which are semi-autonomous and interact in uncertain and adversarial scenarios. In a general setting, a group of agents decides how to allocate a set of scarce and distributed resources to solve a common objective while satisfying operational and communication constraints. The challenge is how to increase the network resilience. To achieve that goal, the algorithms should be scalable, robust against errors in communication or computation, preserve privacy, and that allow the agents in the network to take autonomous decisions on the resource utilization.

One prominent application of resource allocation can be found in the management of computer and public networks, where virus outbreaks threaten the security of critical infrastructure and the well-being of the general population. Interconnected and
dynamic networks make the design and implementation of strategies to stop epidemics even more challenging in scenarios where multiple operators are defined to solve the decision problem with only access to partial information. In this way, a collective decision must be made by the operators about which optimal responses should be implemented for the rapid vaccination and/or isolation of an infected population, possibly under scarce resources, restricted communications, and partial network knowledge. Thus, optimal policy implementations by multiple operators require distributed algorithms that can be used to learn the optimal strategies.

These scenarios can benefit from the development of distributed anonymous coordination algorithms that allow the implementation of best responses in a robust way, in the sense that it is not necessary to start from a specific initialization and the algorithm convergence is not affected by erroneous updates in the system state. Motivated by this, here we propose distributed and robust algorithms in continuous-time and discrete-time for resource allocation which have these properties.

There are many approaches to decentralized convex resource optimization for multiagent systems in the literature, for example some are based on dual decomposition methods, e.g., [1] for unconstrained or [2, 3] for constrained problems or based on alternating direction method of multipliers (ADMM), e.g., [4]. Other approaches are based on a combination of subgradients and consensus [5], a local version of the replicator equation [6, 7], gossip algorithms [8], saddle-point methods [9], or Laplacian gradient dynamics [10, 11]. However, these approaches are not proven to be robust since they assume no errors in communication or computations. In addition, for these approaches,
the total amount of resources available to all agents needs to be known in advance in order to initialize the algorithm. A different approach is shown in [12], where a distributed algorithm for resource allocation in power networks is proposed. However, the algorithm assumes separability of the cost function, which we do not assume.

Centralized algorithms that solve the problem of epidemic outbreak control in interconnected networks have been widely studied in the literature. There are two relevant approaches to handle this problem. One considers the isolation of the infected nodes by means of topology adaptation or quarantine, while others consider node vaccination and immunization. Along these lines, [13] proposes a convex optimization framework to find cost-optimal approaches to traffic control in epidemic outbreaks. However, the authors in [13] do not propose decentralized algorithms for the implementation of their solutions. In contrast, [14] proposes a decentralized algorithm to control the virus propagation by disconnecting nodes and applying an antivirus subject to resource constraints. This decentralized algorithm is based on the use of diagonal matrices in the control input, which are naturally distributed. Nonetheless, the algorithm that determines these diagonal matrices is not distributed itself. A recent formulation for designing optimal dynamic controllers to mitigate the infection through allocation of vaccines is given in [15]. However, the proposed strategy is centralized and the solution is suboptimal. More recently, [16] has proposed a distributed resource allocation strategy to control a virus outbreak in a network by building on the framework of [17]. The proposed algorithm is based on the ADMM algorithm. However, the communication cost of such algorithm is expensive as every operator in the network needs to interchange a local
estimate on the state of the full network. This approach does not scale as well as our
sparser algorithm, which only requires agents to keep estimates of local variables.

Here, we propose various distributed algorithms in continuous- and discrete-time
to solve a class of distributed resource allocation problems. Our approaches allows an
interconnected group of agents to collectively minimize a global cost function subject
to equality and inequality constraints. The first three approaches are posed in a deter-
ministic scenario, while the last approach is posed in a stochastic setting.

In Chapter 3, we study a virus spreading minimization problem based on a gen-
eral contagion dynamics model. We characterize the optimal allocation solution to the
virus problem by posing the problem objective as the minimization of the spectral ra-
dius of the contagion-dynamics matrix subject to operational constraints. By using the
Perron-Frobenius theorem and Lagrange multipliers theory, we obtain a novel charac-
terization of the critical points of the problem that applies to (not necessarily symmet-
ric) weight-balanced matrices. For other matrices, we give bounds for the solution in
terms of the associated symmetrized problem. After this, we propose a discrete-time
distributed algorithm that implements the desired resource allocation for symmetric ma-
trices. In contrast with previous work, our algorithm can be implemented under partial
information by the network nodes by means of local and anonymous interactions. More
precisely, our algorithm is based on a discretization of the local replicator dynamics that
is further adapted to ensure convergence of the solution to the virus mitigation problem,
while satisfying resource constraints. Using a novel discrete-time analysis, we are able
to provide a bound on the algorithm step size that guarantees convergence for agents
subject to time-varying interactions.

In Chapter 4, we propose three novel distributed continuous-time algorithms, the ROBUST MIN-MAX FAIRNESS, ROBUST GRADIENT FAIRNESS, and $p$-ROBUST BOX-GRADIENT FAIRNESS algorithms, which allows agents to converge to a solution that aims to improve their individual payoffs while subject to an equality resource constraint. These dynamics use local and anonymous interactions, do not require knowledge of the total resource available to agents in order to converge to the solution, are robust to agents joining or departing the network, and to sporadic changes in the network topology, computation errors, and communication faults. We analyze the ROBUST MIN-MAX FAIRNESS algorithm convergence over undirected and connected networks, while the ROBUST GRADIENT FAIRNESS, and $p$-ROBUST BOX-GRADIENT FAIRNESS algorithms over weight-balanced and strongly connected networks. We illustrate the applicability of our algorithms by a virus spreading allocation problem over computer and human networks.

Towards a more realistic implementation of the continuous-time dynamics, Chapter 5 presents and analyze two novel robust distributed discrete-time algorithms. The proposed algorithms allow agents to converge to a solution to a ball containing the solution as long as the chosen stepsize is sufficiently small. It is shown that the proposed algorithms are convergent to a neighborhood around the equilibrium even when there are temporary errors in communication or computation. Thus, agents do not require global knowledge of total resources in the network or any specific procedure for initialization. In addition, the proposed algorithms do not require separability of the cost function as long as the gradient is estimated in distributed way. We analyze the
algorithms over weight-balanced and strongly connected networks. We illustrate the applicability of our algorithms on a virus spreading problem over computer and human networks. In this application, we approximate the gradient of the cost function by means of the well-known distributed power iteration method. For that, we propose a distributed stopping criterion, shown in Chapter 6, for the well-known Power Iteration method for symmetric and Metzler matrices. We provide a bound on the accuracy of the approximations for the maximum eigenvalue of the matrix and its corresponding eigenvector. This result is applied to mitigate virus spread over a network. For that, we interconnect the Power Iteration algorithm together with the $p$-robust box-gradient fairness algorithm. This distributed algorithm allows an interconnected group of agents to collectively minimize a global cost function subject to equality and inequality constraints. The Power Iteration method and the distributed stopping criterion provides an approximation of the cost function’s gradient for each iteration. We show that the interconnection between the two methods is convergent and preserves the convergence properties of the $p$-robust box-gradient fairness algorithm. We illustrate the applicability of our algorithm by a virus spread allocation problem over computer dataset taken from the e-mail network of Enron Corporation.

Motivated by the clearing problem in electricity markets, Chapter 7 proposes and analyze a novel distributed discrete-time stochastic algorithm to solve a class of resource allocation problems. Our algorithm builds on the SP method. In particular, we extend our $p$-robust box-gradient fairness algorithm with an SP technique but, unlike previous works, we employ a constant step-size. In this way, our approach allows an
interconnected group of agents to collectively minimize a global cost function subject to both equality and inequality constraints, where the closed-form expression of the local cost functions is unknown to the agents. Under some technical conditions, we show that the algorithm converges in probability to a small neighborhood of the solution as long as the chosen step-size is sufficiently small. It is shown that the proposed algorithms are convergent to a neighborhood around the equilibrium even when there are temporary errors in communication or computation. Thus, agents do not require global knowledge of total resources in the network or employ any special procedure for initialization. Our algorithm is provable correct over weight-balanced and strongly connected networks. In the proofs, we employ Lyapunov theory together with tools from convex analysis and stochastic difference inclusions.

Finally, in Chapter 8, we propose a stochastic source seeking algorithm to drive a robot to an unknown source signal by only using measurements of the signal field. Our proposed algorithm builds on the SPSA algorithm. The novelty of our approach is that we consider nondifferentiable convex functions, fixed step-size, and the environment may have obstacles. We prove practical convergence to a ball and whose size depends on the step-size that contains the location of the source. For the proof, we use Lyapunov theory together with tools from convex analysis and stochastic difference inclusions. Our proof does not rely on stochastic approximation theory as is usually the case for algorithms in the literature based on SPSA. Finally, we show the applicability of the proposed algorithm in a 2D scenario for the source seeking problem.
Chapter 2

Preliminaries and notation

We denote by $\mathbb{R}^d_{\geq 0}$ the positive orthant of $\mathbb{R}^d$, for some $d \in \mathbb{N}$, $\text{diag}(a_1, \ldots, a_N)$ the $N \times N$ matrix with entries $a_i$ along the diagonal, $I_N$ the identity matrix of size $N \times N$, and $1_N \in \mathbb{R}^N$ the column vector whose elements are all equal to one. The spectrum of $A$ is denoted by $\text{spec}(A)$, an eigenvalue of $A$ is denoted by $\lambda_i(A) \in \text{spec}(A)$, its spectral radius by $\rho(A) = \max_i |\lambda_i(A)|$, and the 2-norm of $A$ is denoted by $\|A\|$. When we use inequalities for vectors, we refer to componentwise inequalities.

A real square matrix $A = [a_{ij}]$, $A \in \mathbb{R}^{N \times N}_{\geq 0}$, is called nonnegative, if its entries are nonnegative, i.e, $a_{ij} \geq 0$, for all $i, j \in \{1, \ldots, N\}$. A directed graph of order $N$ or digraph is a pair $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$, the vertex set, is a set with $N$ nodes, and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, the edge set, is a set of ordered pair of vertices called edges. We denote the graph at time $k$ as $G^{(k)} = (\mathcal{V}, \mathcal{E}(k))$ with edge set $\mathcal{E}(k) \subset \mathcal{V} \times \mathcal{V}$, $k \in \mathbb{N}$. Given a digraph $G$, we define the unweighted adjacency matrix of $G$ by $\mathcal{A}(G) \in \mathbb{R}^{N \times N}$ as $a_{ij} = 1$ if $(i, j) \in \mathcal{E}$, and $a_{ij} = 0$ otherwise. Given a nonnegative matrix $B \in \mathbb{R}^{N \times N}$, its associated weighted digraph $G(B)$
is the graph with $\mathcal{V} = \{1, \ldots, N\}$ and edge set defined by the following relationship: $(i, j) \in \mathcal{E}(B)$ if and only if $b_{ij} > 0$. The associated weight of the edge $(i, j)$ is given by the entry $b_{ij}$. The graph $\mathcal{G}(B)$ is said to be weight-balanced if $\sum_{j=1}^{N} b_{ij} = \sum_{j=1}^{N} b_{ji}$ for all $i \in \mathcal{V}$, in particular, $\mathcal{G}(B)$ is undirected if $b_{ij} = b_{ji}$ for all $(i, j) \in \mathcal{E}$. A pair of indices $i, j \in \mathcal{V}$ of an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ are called neighbors if $(i, j) \in \mathcal{E}$. We let $\mathcal{N}_i(\mathcal{G})$ denote the set of neighbors of $i$ in the digraph $\mathcal{G}$. A path in a graph is an ordered sequence of vertices such that any pair of consecutive vertices in the sequence is an edge of the graph. A graph is connected if there exists a path between any two vertices. If a graph is not connected, then it is composed of multiple connected components, that is, multiple connected subgraphs. In a connected graph $\mathcal{G}$, the distance from vertex $i$ to vertex $j$, denoted as $\text{dist}(i, j)$, is the length (number of edges) of a shortest $i$-$j$ path in $\mathcal{G}$.

### 2.0.1 On the replicator dynamics

Replicator dynamics [18, 19, 20] models the interaction of an homogeneous population, where fractions of individuals play a symmetric game. From the biological point of view, it can be seen as mechanism to model the behavior of a population whose individuals seek habitats with different conditions to feed or reproduce. This dynamics is represented by a first-order differential equation, which is composed by the replicator, its fitness, and the proportion in the population. The replicator represents one individual in the entire population. The fitness is the payoff that the individual gets during the game. Finally, the proportion in the population corresponds to the fraction of individuals in the population that changes as a result of their mutual interactions and fitnesses. A
particular choice of replicator dynamics is given by

\[ \dot{p}_i(t) = p_i(t)(f_i - \bar{f}), \]  

(2.1)

where \( p_i \) denotes the proportion of population that play one strategy \( i \in \{1, \ldots, N\} \), \( f_i : \mathbb{R} \to \mathbb{R} \) is the fitness, and \( \bar{f} \) is the average fitness described by \( \bar{f} = \sum_{j=1}^{N} p_j f_j \). The choice of \( \bar{f} \) in (2.1) imposes a useful restriction to the dynamics, as evolutions will belong to the simplex \( \Delta_p = \{ p \in \mathbb{R}_+^N \mid \sum_{i=1}^{N} p_i(t) = 1 \} \). When the equilibrium point \( p_i^* > 0 \) for all \( i \), then the steady state of (2.1) is achieved when \( f_i(p_i^*) = \bar{f}(p^*) \), where \( p^* = [p_1^*, \ldots, p_N^*] \).

The properties of (2.1) make it useful to solve distributed optimization problems subject to constraints like the virus problem we state in Section 3.1.1.

A local version of the original replicator dynamics in (2.1) is proposed in [6] to account for local interactions of fractions of the population over a graph \( G \). The local replicator dynamics is given by

\[ \dot{p}_i(t) = p_i(f_i \sum_{j \in N_i} p_j - \sum_{j \in N_i} p_j f_j), \]  

(2.2)

where \( N_i \) is the set of neighbors of \( i \) in the graph \( G \). If the choice of the fitness \( f_i \) only depends on information of the neighbors and itself, then the algorithm described in (2.2) is distributed. Moreover, since (2.2) does not require the exchange of identities, it is said that it accounts for anonymous interactions. The authors in [6] show that this algorithm conserves the most important characteristics of (2.1), i.e., i) the simplex is invariant, and
ii) the equilibrium point is asymptotically stable in $\Delta_p$.

### 2.0.2 Partial stability for nonsmooth Lyapunov functions

#### Continuous-time

The notions we introduce here follow [21]. Consider a nonlinear system of ODEs in $\mathbb{R}^d$

$$\dot{x} = X(x), \quad (2.3)$$

where $X : \mathbb{R}^d \to \mathbb{R}^d$. The state $x$ in system (2.3) is divided in two components: (1) the $y$-component used to study the stability of the equilibrium $x = 0$, and (2) other (non-controlled) $z$-component, so that $x = [y^T, z^T]^T$. We use $x(t) = x(t; t_0, x_0)$ to denote the solution of the system (2.3) given the initial condition $x_0$ at $t_0$.

**Definition 2.1.** Let $D_\varphi$ be a domain of initial conditions $x_0$ such that $\|y_0\|_2 < \varphi$ and $\|z_0\|_2 \leq L$, where $L \in \mathbb{R}_{>0}$. The origin of system (2.3) is said to be

i) y-stable if for any $\varepsilon > 0$, $t_0 \geq 0$, one can find $\varphi(\varepsilon, L) > 0$ such that $x_0 \in D_\varphi$ yields $\|y(t; t_0, x_0)\|_2 < \varepsilon$ for all $t \geq t_0$.

ii) globally asymptotically y-stable if it is y-stable, and an arbitrary solution $x(t)$ of the system (2.3) exists for all $t \geq 0$, $y_0 \in K_y$, where $K_y$ is an arbitrarily compact set in y-space, and $x(t)$ is y-bounded and satisfies $\lim\|y(t; t_0, x_0)\|_2 = 0$ as $t \to +\infty$.

The smooth version of the next lemma was introduced in [21], Theorem 2. We present here an adaptation to nonsmooth Lyapunov functions. The proof is straightforward when considering the analogous definitions from nonsmooth analysis and follows the same steps as in [21].
Lemma 2.1. Suppose that there exists a locally Lipschitz and regular scalar function $V$, and a continuous vector function $W : \mathbb{R}^d \rightarrow \mathbb{R}^d$ for the system (2.3) with $W(0) = 0$. Then, the origin is globally asymptotically $y$-stable if the following conditions are met

\begin{align}
V(0,z) &= 0, \\
\alpha_1(\|y\|_2) &\leq V(x) \leq \alpha_2(\|\vartheta\|_2), \\
\max_L \mathcal{L}XV &\leq -\alpha_3(\|\vartheta\|_2),
\end{align}

where $\vartheta = [y^\top, W(x)^\top]^\top$, $\alpha_1$ belongs to class $\mathcal{K}_\infty$, $\alpha_2, \alpha_3$ belong to class $\mathcal{K}$.

Discrete-time

The notions we introduce here follow [21, 22]. Given two sets $S$ and $T$, a set-valued map, denoted by $h : S \Rightarrow T$, associates to an element of $S$ a subset of $T$. Consider a discrete-time dynamical system given by the difference inclusion in $\mathbb{R}^n$

$$x(t + 1) \in \mathcal{H}(x(t), w(t)),$$

(2.5)

where $\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^m \Rightarrow \mathbb{R}^n$ is a set-valued map for some $n, m \in \mathbb{N}$, $x \in \mathbb{R}^n$ is the state, $w \in \mathbb{R}^m$ is the input, and $t \geq 0$. We assume that $\mathcal{H}$ assigns to each point $(x, w) \in \mathbb{R}^n \times \mathbb{R}^m$ a nonempty set $\mathcal{H}(x, w) \subset \mathbb{R}^n$. Consider the unforced system (2.5), i.e., $w = 0$. We divide the state $x$ in two components: (1) the $y$-component used to study the stability of the equilibrium $x^* = 0$, and (2) other (non-controlled) $z$-component, so that $x = [y^\top, z^\top]^\top$.

We use $x(t) = x(t; t_0, x_0)$ to denote the solution of the system (2.5) given the initial condition $x_0$ at $t_0$. When $w = 0$, we have the following definition.

Definition 2.2. Let $D_\varphi$ be a domain of initial conditions $x_0$ such that $\|y_0\|_2 < \varphi$ and $\|z_0\|_2 \leq L$, where $L \in \mathbb{R}_{>0}$. The origin of system (2.5) is said to be
1. **y-stable** if for any \( \varepsilon > 0, t_0 \geq 0 \), one can find \( \phi(\varepsilon, L) > 0 \) such that \( x_0 \in D_{\phi} \) yields \( \|y(t; t_0, x_0)\|_2 < \varepsilon \) for all \( t \geq t_0 \).

2. globally asymptotically **y-stable** if it is y-stable and an arbitrary solution \( x(t) \) of the system (2.5) exists for all \( t \geq 0 \), \( y_0 \in K_y \), where \( K_y \) is an arbitrarily compact set in y-space, and \( x(t) \) is y-bounded and satisfies \( \lim_{t \to +\infty} \|y(t; t_0, x_0)\|_2 = 0 \) as \( t \to +\infty \).

The smooth version of the next lemma was introduced in [21], Theorem 2. We present here an adaptation to nonsmooth Lyapunov functions. The proof is straightforward when considering the analogous definitions from nonsmooth analysis and follows the same steps as in [21].

**Lemma 2.2.** Suppose that there exists a continuous scalar function \( V : \mathbb{R}^n \to \mathbb{R} \) and a continuous vector function \( W : \mathbb{R}^n \to \mathbb{R}^n \) for the unforced system (2.5) in \( \mathbb{R}^n \) with \( W(0) = 0 \). Then, the origin is globally asymptotically y-stable if the following conditions are met

\[
\begin{align*}
V(0, z) & = 0, \\
\alpha_1(\|y\|_2) & \leq V(x) \leq \alpha_2(\|\theta\|_2), \\
\sup_{g \in H(x)} V(g) - V(x) & \leq -\alpha_3(\|\theta\|_2),
\end{align*}
\]

where \( \theta = [y^\top, W(x)^\top]^\top \), \( \alpha_1 \) belongs to class \( \mathcal{K}_\infty \), \( \alpha_2, \alpha_3 \) belong to class \( \mathcal{K} \).

Next definition is an adaptation of input-to-output stability (e.g., see [23]) when we consider the output of the system (2.5) as the controllable states \( y \).

**Definition 2.3.** (y-input-to-state stability): The system (2.5) is said to be y-input-to-state stable (for short y-ISS) if there exists a \( \mathcal{KL} \) function \( \psi \) and a \( \mathcal{K} \) function \( \phi \) such that for any initial state \( y(t_0) \) and any bounded input \( w(t) \), the solution \( y(t) \) for \( t \geq 0 \) satisfies

\[
\|y(t)\| \leq \psi(\|\theta_0\|, t - t_0) + \phi(\sup_{t_0 \leq \tau \leq t} w(\tau))
\]
Theorem 2.1. Suppose that there exists a continuous scalar function $V : \mathbb{R}^n \to \mathbb{R}$ such that

\begin{align}
V(0, z) &= 0, \\
\alpha_1(||y||) \leq V(x) \leq \alpha_2(||y||), \\
\sup_{g \in H(x)} V(g) - V(x) &\leq -\alpha_3(||y||), \quad \forall ||y|| \geq \rho(||w||)
\end{align}

$\forall (x, w) \in \mathbb{R}^n \times \mathbb{R}^m$, where $\theta = [y^T, W(x)^T]^T$, $\alpha_1$ belongs to class $\mathcal{K}_\infty$, $\alpha_2$, $\alpha_3$, $\rho$ belong to class $\mathcal{K}$, and a continuous vector function $W : \mathbb{R}^n \to \mathbb{R}^n$ for the system (2.5) with $W(0) = 0$. Then the system (2.5) is $y$-input-to-state stable with $\phi = \alpha_1^{-1}(\alpha_2(\rho))$.

2.0.3 Set-valued analysis

The notions we introduce here follow [24]. The following definitions assume $X$ and $Y$ to be metric spaces, in particular we assume $X$ and $Y$ to be subsets of $\mathbb{R}^N$ with the induced Euclidean topology.

Definition 2.4. (Graph of a set-valued map): Let $\psi : X \rightrightarrows Y$ be a set-valued map. The graph $Gr(\psi)$ of $\psi$ is the subset of $X \times Y$ associated to $\psi$, defined by

$$Gr(\psi) = \{(x, y) \in X \times Y \mid y \in \psi(x)\}.$$ 

Definition 2.5. (Upper hemicontinuity uhc): A set-valued map $\psi : X \rightrightarrows Y$ is upper hemicontinuous at $x \in X$ if for every open neighborhood $V$ of $\psi(x)$, the set $\{x' \in X \mid \psi(x') \subset V\}$ is a neighborhood of $x$. We say $\psi$ is upper hemicontinuous (for short uhc) on $X$, if it is upper hemicontinuous at every point of $X$.

Lemma 2.3. A singleton-valued, set-valued map is uhc if and only if it is a continuous function.
2.0.4 Lipschitz-type stability in nonsmooth convex programs

The notions we introduce here follow [25]. Consider the convex problem

$$\min_x f(x)$$

subject to \(Ax = y\),

where \(f : \mathbb{R}^N \to \mathbb{R}\) is assumed convex and finite, \(x \in \mathbb{R}^N\), and the operator \(A\) is linear from \(\mathbb{R}^N\) into \(\mathbb{R}^m\) for some \(m > 0\). The parameter is denoted by \(y\) and will also be called perturbation. We denote by \(x_0\) the optimal solution corresponding to the parameter \(y = 0\). We assume the so-called Mangasarian-Fromovitz regularity condition (MFC) holds at the optimal solution \(x_0\) for the parameter \(y = 0\):

- The rows \(A^i\) for \(i \in \{1, \ldots, r\}\) of the matrix \(A\) are linearly independent.
- There exists some vector \(z \in \mathbb{R}^N\) such that \(A^i z = 0\) for \(i \in \{1, \ldots, r\}\).

We assume moreover that the following local conditions are satisfied

- Superquadratic growth condition: There is \(\gamma > 0\) such that
  \[ f(x) \geq f(x_0) + Df(x_0; x - x_0) + \frac{\gamma}{2} \|x - x_0\|^2 \]

- Subquadratic growth condition: There is \(\Gamma \geq \gamma\) such that
  \[ f(x) \leq f(x_0) + Df(x_0; x - x_0) + \frac{\Gamma}{2} \|x - x_0\|^2 \]

for \(x \in x_0 + B\), where \(B\) is some compact convex neighborhood, \(Df(x_0; x - x_0)\) denotes the directional derivative of the function \(f\) at a point \(x_0\). The directional derivative and
the subdifferential of the function $f$ at a point $x$ are related by the relationship

$$Df(x; y) = \lim_{t \to 0^+} \frac{1}{t} (f(x + ty) - f(x)) \geq \xi^\top y \forall \xi \in \partial f(x),$$

and if we assume the function $f$ is continuous at the point $x$ then we have

$$Df(x; y) = \max_{\xi \in \partial f(x)} \xi^\top y.$$

For any full ranked matrix $A$ we denote $A^\# = A^\top (AA^\top)^{-1}$.

**Lemma 2.4.** ([25] Lipschitz-type stability of the optimal solution): Suppose the assumptions of superquadratic and subquadratic growth conditions holds and the MFC is satisfied. Then there exists $\epsilon > 0$ such that for all perturbations $y \in \mathbb{R}^m$, $\|y\| \leq \epsilon$, there exists $\sigma_0 \in \partial f(x_0)$ such that the following holds: if $\Sigma = \{\sigma \in \partial f(x_0) \mid \sigma^\top A^\# y > \lambda^\top y\}$ for any $\lambda$ satisfying $\sum_{i=1}^m \lambda_i A_i^\top \in \partial f(x_0)$ and $\lambda_i A_i x_0 = 0$, then for any optimal solution $x(y)$, the following alternative holds: either $\Sigma = \emptyset$, and then

$$\|x - x_0\| \leq \sqrt{\frac{\Gamma}{\gamma}} \|A^\#\| \|y\|, \quad (2.9)$$

or $\Sigma \neq \emptyset$, and then

$$\|x - x_0\| \leq \sqrt{\frac{\Gamma}{\gamma}} (\|A^\#\|^2 + \sup_{\sigma \in \Sigma} \frac{\|((\sigma - \sigma_0)^\top A^\#)^2\|}{\|1_{N \times N} - A^\# A(\sigma - \sigma_0)\|^2}) \|y\|. \quad (2.10)$$

### 2.0.5 Stability for stochastic difference inclusions

The notions we introduce here follow [26]. Consider a discrete-time, stochastic difference inclusion

$$x^+ \in \mathcal{H}_0(x, v^+), \quad v \sim \mu, \quad (2.11)$$
where \( x^+ \) is the state after an instantaneous change, \( \mathcal{H}_\alpha : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \) is a set-valued map for some \( n, m \in \mathbb{Z}_{>0} \) parameterized by \( \alpha \in \mathbb{R}_{>0} \) and which assigns non-empty set values, and \( x \in \mathbb{R}^n \) is the state. The notation \( v^+ \) and \( v \) refers to sequences of random input variables as explained next. Consider a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \( \Omega \) denotes the set of all possible outcomes, \( \mathcal{F} \) is the \( \sigma \)-field associated with \( \Omega \), and \( \mathbb{P} \) is the probability function that assigns a probability to events in \( \mathcal{F} \). In particular, we assume \( \mathcal{B}(\mathbb{R}^m) \subseteq \mathcal{F} \), where \( \mathcal{B}(\mathbb{R}^m) \) is the Borel field. In (2.11), we use \( v^+ \) and \( v \) as a place holder for a sequence of independent, identically distributed (i.i.d.) random variables \( v \triangleq \{ v_k \}_{k=0}^\infty \), that is, such that \( \mathbb{P}(v_k \in F) = \mathbb{P}(\{ w \in \Omega \mid v_k(w) \in F \}) \) is well defined and independent of \( k \) for each \( F \in \mathcal{B}(\mathbb{R}^m) \). We use \( \mathcal{F}_k \) to denote the collection of sets \( \{ w \in \Omega \mid (v_0(w), \ldots, v_k(w)) \in F \} \), \( F \in \mathcal{B}((\mathbb{R}^m)^{k+1}) \), which are the sub-\( \sigma \)-fields of \( \mathcal{F} \) that form the minimal filtration of the sequence \( v \). Due to the i.i.d property, each random variable has the same probability measure \( \mu : \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \) defined as \( \mu(F) = \mathbb{P}(v_k \in F) \) and, for almost all \( w \in \Omega \),

\[
E[f(v_0, \ldots, v_k, v_{k+1})|\mathcal{F}_k](w) = \int_{\mathbb{R}^m} f(v_0(w), \ldots, v_k(w), v)\mu(dv),
\]

for each \( k \in \mathbb{Z}_{\geq 0} \) and each measurable \( f : ((\mathbb{R}^m)^{k+2} \rightarrow \mathbb{R} \).

The sequence of random variables \( x \triangleq \{x_k\}_{k \geq 0} \), where \( x_k : \text{dom} x_k \subset \Omega \rightarrow \mathbb{R}^n \), \( k \in \mathbb{Z}_{\geq 0} \) with \( x_0 = x \) for all \( w \in \Omega \) and \( \text{dom} x_{k+1} \subset \text{dom} x_k \), is called a random process starting at \( x \in \mathbb{R}^n \). We say that \( x \) is adapted to the natural filtration of \( v \) if \( x_{k+1} \) is \( \mathcal{F}_k \) measurable for each \( k \in \mathbb{Z}_{\geq 0} \), i.e., \( x_{k+1}^{-1}(F) \in \mathcal{F}_k \) for each \( F \in \mathcal{B}(\mathbb{R}^m) \). A random process \( x \) starting from \( x \in \mathbb{R}^n \) that is adapted to the natural filtration of \( v \), together with a random
variable \( I : \Omega \to \mathbb{Z}_{\geq 0} \cup \{\infty\} \) (which denotes the number of elements in the sequence \( x \)) is a **random solution** of (2.11) starting at \( x \in \mathbb{R}^n \), denoted as \( x \in S(x) \), if \( x_0 = x \), \( x_{k+1}(w) \in \mathcal{H}_\alpha(x_k(w), v_{k+1}(w)) \) for all \( w \in \text{dom} \ x_{k+1} \triangleq \{ w \in \Omega \mid k + 1 \leq I \} \) and \( k \in \mathbb{Z}_{\geq 0} \).

We impose the following regularity condition on \( \mathcal{H} \).

**Assumption 2.1.** \( \mathcal{H} \) is locally bounded and \( v \mapsto \text{graph}(\mathcal{H}_\alpha(\cdot, v)) \triangleq \{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid y \in \mathcal{H}_\alpha(x, v) \} \) is measurable with closed values.

The compact set \( \mathcal{A} \subset \mathbb{R}^n \) is **stable in probability** for (2.11) if for each \( \epsilon > 0 \) and \( \varsigma > 0 \) there exists \( \omega > 0 \) such that, for each \( x \in \mathcal{A} + \omega \mathbb{B} \) and \( x \in S(x) \), \( \mathbb{P}(\text{graph}(x) \subset (\mathbb{Z}_{\geq 0} \times (\mathcal{A} + \epsilon \mathbb{B}))) \geq 1 - \varsigma \).

Next, we introduce the notion of stochastic stability called recurrence, wherein solutions return to a bounded set infinitely often. Roughly speaking, an open, bounded set is said to be recurrent if almost all solutions revisit the set infinitely often. A recurrent set is not necessarily stable in probability and recurrence does not imply that solutions stay bounded, but rather states that solutions reach a compact set with probability one.

**Definition 2.6.** *(Globally Recurrent Set):* An open, bounded set \( O \subset \mathbb{R}^n \) is said to be **globally recurrent** if \( E[\prod_{i \in \mathbb{Z}_{\geq 0}} \mathbb{1}_{\mathbb{R}^n \setminus O}(x_i)] = 0 \), for each \( x \in \mathbb{R}^n \) and each \( x \in S(x) \).

**Proposition 2.1.** *[26]* Consider the system (2.11) under Assumption 2.1. If there exists a radially unbounded, upper semicontinuous function \( V : \mathbb{R}^n \to \mathbb{R}_{\geq 0} \), and a continuous function \( \varrho : \mathbb{R}^n \to \mathbb{R}_{>0} \) such that

\[
\int_{\mathbb{R}^n} \max_{h \in \mathcal{H}_\alpha} V(h) \mu(dv) \leq V(x) - \varrho(x),
\]

for all \( x \in \mathbb{R}^n \setminus O \). Then, \( O \) is globally recurrent for (2.11).
In order to analyze the stability properties of systems of the form (2.11) with respect to compact sets $\mathcal{A} \subset \mathbb{R}^n$, we introduce the notion of input-to-state stability in probability. The system (2.11) is said to be input-to-state stable in probability (ISSp) relative to $\mathcal{A}$ if 1) $\mathcal{A}$ is stable in probability when $\alpha = 0$ and 2) there exists $\varphi \in \mathcal{K}_\infty$ such that, for each $\alpha > 0$, the open bounded set $\mathcal{A} + \varphi(\alpha)B^o$ is globally recurrent for (2.11).

**Definition 2.7.** (Mean-Square Practically Exponentially Stable Equilibrium): We say that the equilibrium point of (2.11) is mean-square practically exponentially stable (MSP-ES) if there exists $\alpha^* \in (0, 1)$, positive real numbers $\beta$, $\lambda < \frac{1}{\alpha^*}$, $\gamma$ and $\eta$, such that for all $\alpha \in (0, \alpha^*)$ we have

$$E[\|x_k\|^2] \leq \beta(1 - \alpha \lambda)^k\|x_0\|^2 + \gamma \alpha^n, \quad \forall k \in \mathbb{Z}_{\geq 0}.$$ 

**Proposition 2.2.** ([27]): Consider the system (2.11) under Assumption 2.1. If there exists an upper semicontinuous function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, positive constants $c_1, c_2, \lambda, K, \alpha^* \in (0, 1)$, and $\eta > 1$, such that for all $\alpha \in (0, \alpha^*)$

$$c_1\|x\|^2 \leq V(x) \leq c_2\|x\|^2,$$

$$\int_{\mathbb{R}^n} \max_{h \in \mathcal{H}_x} V(h)\mu(dv) \leq (1 - \alpha \lambda)V(x) + \alpha^\eta K,$$  

(2.13)

then, the equilibrium point is MSP-ES for (2.11).

2.0.6 **Convex analysis notions**

The notions we introduce here follow [28, 29]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a closed, proper, and convex function for some $n \in \mathbb{Z}_{>0}$. The subgradient of $f$ is the set-valued map $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by the subgradient set $\partial f(x) = \{\xi \in \mathbb{R}^n \mid f(x') \geq f(x) + \langle \xi, x' - x \rangle\}$. 

\(\xi^T(x' - x)\). We refer to \(df(x)\) as the *semi-derivative* function, which is the support function of the nonempty, compact, and convex set \(\partial f(x)\), i.e., \(df(x)(w) = \sup\{\xi^T w \mid \xi \in \partial f(x)\}\). The first order expansion of \(f\) for any point \(x\) is given by

\[
f(x + w) = f(x) + df(x)(w) + o(\|w\|).
\] (2.14)

We say that \(f\) satisfies the *superquadratic growth condition* if there exists \(\gamma > 0\) such that

\[
f(y) \geq f(x) + df(x)(y - x) + \frac{\gamma}{2}\|y - x\|^2,
\] (2.15)

for \(x, y \in \mathbb{R}^n\). In particular, a strongly convex function satisfies the superquadratic growth condition, and, if \(f\) is twice differentiable, this condition is equivalent to assuming \(\gamma I_n \leq \nabla^2 f(x)\) for \(x \in \mathbb{R}^n\). Similarly, we say that \(f\) satisfies the *subquadratic growth condition* if there is \(\Gamma\) such that

\[
f(y) \leq f(x) + df(x)(y - x) + \frac{\Gamma}{2}\|y - x\|^2,
\] (2.16)

for \(x, y \in \mathbb{R}^n\). *Note:* Some notations that are particular to each chapter, will be introduced at the beginning of the chapter itself.

### 2.1 Contributions of this thesis

The contents of this thesis can be summarized as follows: Chapter 3 studies the virus spreading minimization problem based on a general contagion dynamics model. We characterize the optimal allocation solution to the virus problem by posing the problem objective as the minimization of the spectral radius of the contagion-dynamics matrix subject to operational constraints. We propose an algorithm inspired by the Replica-
tor Dynamics that implements the desired resource allocation for time-varying symmetric matrices. Chapter 4 presents the second algorithm, which is designed in continuous-time. It uses local and anonymous interactions, does not require knowledge of the total resource available to agents in order to converge to the solution, is robust to agents joining or departing the network, and to sporadic changes in the network topology, computation errors, and communication faults. Chapter 5 shows the third proposed algorithm. The algorithm conserves the robustness properties of the previous algorithm. It is formulated in discrete-time and allows agents to converge to a solution to a ball containing the solution. Chapter 6 shows a distributed stopping criteria for the power iteration and its feedback interconnection with the algorithm shown in Chapter 5. Finally, Chapter 7, presents an algorithm for scenarios where the gradient is not explicitly given. The proposed algorithm is a discrete-time stochastic algorithm. It builds on the Simultaneous Perturbation (SP) method.
Chapter 3

A distributed dynamics for virus-spread control

In this chapter we study a virus spreading minimization problem based on a general contagion dynamics model. We characterize the optimal allocation solution to the virus problem by posing the problem objective as the minimization of the spectral radius of the contagion-dynamics matrix subject to operational constraints. By using the Perron-Frobenius theorem and Lagrange multipliers theory, we obtain a novel characterization of the critical points of the problem that applies to (not necessarily symmetric) weight-balanced matrices. For other matrices, we give bounds for the solution in terms of the associated symmetrized problem. After this, we propose a discrete-time distributed algorithm that implements the desired resource allocation for symmetric matrices. In contrast with previous work, our algorithm can be implemented under partial information by the network nodes by means of local and anonymous interactions. More
precisely, our algorithm is based on a discretization of the local replicator dynamics that is further adapted to ensure convergence of the solution to the virus mitigation problem, while satisfying resource constraints. Using a novel discrete-time analysis, we are able to provide a bound on the algorithm step size that guarantees convergence for agents subject to time-varying interactions.

### 3.1 Problem Formulation

This section introduces the contact network dynamics proposed in [30] and the problem statement given in [14]. Next, we extend a theorem in [14] for symmetric, irreducible matrices to weight-balanced and irreducible matrices. This extension is motivated by the possibility of having an asymmetric placement of edge isolation (e.g., quarantine or firewalls) making the interaction graph directed. Our proof relies on the Lagrange multiplier approach and the Perron-Frobenius theorem, instead of using a sensitivity formula. Finally, we propose a strategy for the minimization of the virus spread over a network such that minimizes the Perron eigenvalue of the symmetrized counterpart for any nonnegative matrix, and we characterize the goodness of this approximation.

#### 3.1.1 Problem statement

The virus dynamics over a network proposed in [30] is given by

\[
x^{(k+1)}_i = (1 - \prod_{j=1}^N (1 - a_{ji}x^{(k)}_j)),
\]  

(3.1)
where $x_i^{(k)} \in \mathbb{R}$ is the probability that node $i$ is infected at time $k$, $i \in \{1, \ldots, N\}$ and $a_{ji}$ is defined as

$$a_{ji} = \begin{cases} 
\beta_{ji}, & \text{for } j \neq i, \\
1 - \delta_i, & \text{for } j = i.
\end{cases}$$

Here, $\beta_{ji} \in [0, 1]$ is the probability that the virus from node $j$ infects node $i$, and $\delta_i \in [0, 1]$ is the probability of an infected node $i$ to be recovered. Using the Weierstrass product inequality, valid for $a_{ji}x_j^{(k)} \in [0, 1]$, we obtain the following upper bound

$$x_i^{(k+1)} \leq \sum_{j=1}^{N} a_{ji}x_j^{(k)}, \quad \forall i \in \{1, \ldots, N\}.$$ 

The previous inequality reads in vector notation as

$$x^{(k+1)} \leq A(\delta)x^{(k)}$$

where $x^{(k)} = [x_1^{(k)}, \ldots, x_N^{(k)}]^T$, $\delta = (\delta_1, \ldots, \delta_N) \in [0, 1]^N$, and $A(\delta) = [a_{ji}] \in \mathbb{R}^{N \times N}$. Let $G \equiv A(\delta) - I_N + D \equiv A(1_N)$, where $D = \text{diag}(\delta)$, and $\mathcal{G}(I_N + A(1_N)) = \mathcal{G}(I_N + G)$ be the graph associated with the contact dynamics matrix. We define the topology matrix of the network as the matrix $I_N + G$. When there is no confusion, we will denote $\mathcal{G}(I_N + G)$ by $\mathcal{G}$ and $\mathcal{A}$ for the associated unweighted adjacency matrix. In [30], authors prove next proposition that shows that the dominant eigenvalue, $\rho(A)$, governs the decay rate of infection.
**Proposition 3.1** ([30]). An epidemic described by (3.1) becomes extinct if and only if \( \rho(A(\delta)) < \alpha \) for some \( \alpha \in [0, 1] \) and for any initial infection \( x(0) \in [0, 1]^N \). Moreover, (3.1) will converge to zero exponentially fast with rate of convergence given by \( \alpha \).

We consider the following problems to minimize the effects of virus contagion. The \( \delta \)-virus mitigation problem is given by

\[
\min_{\delta \in [0,1]^N} \rho(A(\delta)), \\
\text{s.t. } \sum_{i=1}^{N} \delta_i = \Gamma, \\
\rho(A(\delta)) < 1.
\]

(3.3)

Depending on the value of \( \Gamma \), the \( \delta \)-virus mitigation problem is not feasible and we can only solve the relaxed virus spread minimization problem:

\[
\min_{\delta \in [0,1]^N} \rho(A(\delta)), \\
\text{s.t. } \sum_{i=1}^{N} \delta_i = \Gamma.
\]

(3.4)

Here, we only consider a partial vaccination strategy since we only consider as decision variable \( \delta \), while \( \beta \) is fixed. We further assume that there are enough resources in terms of isolation/quarantine capabilities (i.e., \( \beta \)), which make possible to balance the network interaction according to a finite-time distributed algorithm presented in [31].

**Remark 3.1.** When there are not enough resources to balance the network, we can use a similarity transformation \( B = XAX^{-1} \), where \( B \) is weight-balanced and \( X \) is a
square diagonal matrix that can be found by the algorithm given in [32]. The similarity transformation for balancing the network for irreducible matrices is proven to be unique in [33]. The algorithm in [32] requires information of the in- and out-neighbors. What prevents this algorithm to be computed in a distributed way, is the fact that at every step it is normalized using 1-norm, however, all results of this algorithm hold when it is used ∞-norm. By this modification in the algorithm, we can implement it in a distributed way. Since the algorithm converges asymptotically to the diagonal X, the result provided by this algorithm is only an approximation. The effect of such approximation is out of scope of this paper.

3.1.2 Solution characterization for balanced matrices

We present a characterization of the solution to a relaxed problem for weight-balanced matrices in Theorem 3.2 to provide sufficient conditions for feasibility for the δ-virus mitigation and virus spread minimization problems. Previous to this, the next theorem recalls that the function ρ is a convex function of δ and, hence, the problems introduced in Section 3.1.1 are convex. Proofs for all results can be found in the Appendix.

Theorem 3.1 ([34]). Let B be a nonnegative matrix and $D = \text{diag}(\delta_1, \ldots, \delta_N)$. Then, the Perron eigenvalue of $B + D$, $\rho(B + D)$, is a convex function of $D$.

The δ-virus mitigation problem is feasible when the set of $\delta$ satisfying $\delta \in [0, 1]^N$, $\rho(A(\delta)) < 1$ and $\sum_{i=1}^N \delta_i = \Gamma$ is non empty. On the other hand, the virus spread
MINIMIZATION is feasible when the set of $\delta$ satisfying $\delta \in [0, 1]^N$ and $\sum_{i=1}^N \delta_i = \Gamma$ is non empty. Denote by $r_i = 1 + \sum_{j=1, j \neq i}^N \beta_{ji}$ and $c_i = 1 + \sum_{j=1, j \neq i}^N \beta_{ij}$, the sum of row and column entries of $I_N + A(I_N)$, respectively.

**Theorem 3.2.** For a weight-balanced, nonnegative, and irreducible matrix $A(\delta)$ the solution to the virus spread minimization problem without the restrictions $\delta_i \in [0, 1], i \in \mathcal{V}$, is given by making the sums of each row of $A(\delta)$ equal to each other, i.e., $-\delta_i + r_i = -\delta_j + r_j$ for $i \neq j$. Precisely, the solution is characterized by

$$\rho^*(A(\delta^*)) = \frac{1}{N}(\sum_{j=1}^N r_j - \Gamma),$$

$$\delta^*_i = \frac{1}{N}(Nr_i - \sum_{j=1}^N r_j + \Gamma).$$

**Proof.** Since $A(\delta)$ is nonnegative and irreducible for all $\delta$, by the Perron-Frobenius theorem we have that $\rho(A) \leq \max_i (-\delta_i + r_i)$. In order to obtain an upper bound for $\rho^*(A(\delta^*))$, we minimize the maximum row sum of $A(\delta)$, i.e,

$$\min_{\delta} \max_{i} (-\delta_i + r_i)$$

s.t. $\sum_{i=1}^N \delta_i = \Gamma$.  

The problem stated in (3.7) can be reformulated as

$$\min_{z, \delta} z$$

s.t. $z \geq \max_i (-\delta_i + r_i)$,

$$\sum_{i=1}^N \delta_i = \Gamma,$$
where \( z \in \mathbb{R} \) is a new variable. It holds that \( z \geq \max_i (-\delta_i + r_i) \) if and only if \( z \geq -\bar{\delta}_i + r_i \) for all \( i \in \mathcal{V} \). Then, the Lagrangian of (3.8) is given by

\[
\mathcal{L}(z, \delta, \nu, \mu) = z + \nu \left( \sum_{i=1}^{N} \delta_i - \Gamma \right) + \sum_{i=1}^{N} \mu_i (-z - \delta_i + r_i),
\]

where \( \nu, \text{ and } \mu = [\mu_1, \ldots, \mu_N] \) are the KKT multipliers for the equality and inequality constraints, respectively. Suppose that \( \bar{\nu}^* = (\bar{\mu}^*, \bar{z}^*, \bar{\delta}^*, \bar{\nu}^*) \) is a solution to (3.8). From complementary slackness, we have \( \bar{\mu}^*_i (-\bar{z}^* - \bar{\delta}^*_i + r_i) = 0 \) for all \( i \in \{1, \ldots, N\} \). Moreover,

\[
\frac{\partial \mathcal{L}}{\partial \nu} \bigg|_{\bar{\nu}^*} = 1 - \sum_{i=1}^{N} \bar{\mu}^*_i = 0, \quad \frac{\partial \mathcal{L}}{\partial \delta} \bigg|_{\bar{\nu}^*} = \bar{\nu}^* - \bar{\mu}^*_i = 0, \text{ and } \bar{\nu}^* = \bar{\mu}^*_i = \frac{1}{N}. \]

A critical point is given when \( \bar{z}^* = -\bar{\delta}^*_i + r_i, \forall i \in \{1, \ldots, N\} \). Note that all \( \bar{\mu}^*_i \) are active (i.e, \( \bar{\mu}^*_i > 0 \)). From the fact that problem (3.8) gives an upper bound to the solution of the VIRUS SPREAD MINIMIZATION problem without the \( \delta \in [0, 1]^N \) restriction, it follows that \( \rho^*(A(\bar{\delta}^*)) \leq \bar{z}^* = \max_i (-\bar{\delta}^*_i + r_i) \).

Since \( \bar{z}^* = \max_i (-\bar{\delta}^*_i + r_i) = \min_i (-\bar{\delta}^*_i + r_i) = -\bar{\delta}^*_i + r_i, \forall i \), so \( \bar{z}^* = -\bar{\delta}^*_i + r_i, \forall i \). Similarly, an analogous condition holds when we consider \( \rho(A(\delta)^T) \geq \min_i (-\delta_i + c_i) \), which follows from the Perron-Frobenius theorem. The analogous problem to (3.7) with the alternative objective \( \max_i \delta \min_i (-\delta_i + c_i) \) leads to a solution \( \bar{z}^* = -\bar{\delta}^*_i + c_i, \forall i \), which satisfies \( \bar{z}^* \leq \rho^*(A(\bar{\delta}^*)^T) \).

From \( \rho(A(\delta)) = \rho(A(\delta)^T) \) for any \( \delta \), we obtain the relation \( \bar{z}^* \leq \rho^*(A(\bar{\delta}^*)) \leq \bar{z}^* \).

Using that \( \sum_i \delta_i = \Gamma \), and that \( \bar{z}^* = -\bar{\delta}^*_i + r_i \) (resp. \( \bar{z}^* = -\bar{\delta}^*_i + c_i \)) for all \( i \in \{1, \ldots, N\} \), we obtain (3.5) (resp. the analogous equation to (3.5) with \( c_i \) replacing \( r_i \)). From the weight-balanced property of \( I_N + G \), we have that \( r_i = c_i \) for all \( i \in \{1, \ldots, N\} \), and thus \( \bar{\delta}^*_i = \bar{\delta}^*_i \) for all \( i \in \{1, \ldots, N\} \). Thus, \( \bar{z}^* = \bar{z}^* = \rho^*(A(\bar{\delta}^*)) \).
Finally, (3.6) is obtained by replacing (3.5) into the expression \( \rho^*(A(\delta^*)) = -\delta_i^* + r_i. \)

\[ \square \]

**Corollary 3.1.** *(Sufficient conditions for problem feasibility): When \( A(\delta) \) is weight-balanced, nonnegative, and irreducible, then a feasible solution to the virus spread minimization problem is given by (3.6) if

\[
\max_i \left( \sum_{k,j=1,j\neq k}^N \beta_{jk} - N \sum_{j=1,j\neq i}^N \beta_{ji} \right) \leq \Gamma \leq \sum_{i,j=1,j\neq i}^N \beta_{ji} + N(1 - \max_i \sum_{j=1,j\neq i}^N \beta_{ji}). \quad (3.9)
\]

Moreover, a feasible solution to the \( \delta \)-virus mitigation problem is given by (3.6) if

\[
\sum_{i,j=1,j\neq i}^N \beta_{ji} < \Gamma \leq \sum_{i,j=1,j\neq i}^N \beta_{ji} + N(1 - \max_i \sum_{j=1,j\neq i}^N \beta_{ji}). \quad (3.10)
\]

**Proof.** For the left-hand side of the inequality (3.9), assume \( \delta_i^* \geq 0 \). It follows that

\[ Nr_i - \sum_{j=1}^N r_j + \Gamma \geq 0, \]

which is equivalent to \( \Gamma \geq \max_i (\sum_{k,j=1,j\neq k}^N \beta_{jk} - N \sum_{j=1,j\neq i}^N \beta_{ji}). \)

The right-hand side of inequalities (3.9) and inequality (3.10) follows from imposing \( 1 - \delta_i^* \geq 0, i \in V \). By replacing \( \delta_i^* \) in \( N - N\delta_i^* \geq 0 \) from (3.6), we have that

\[ \Gamma \leq \sum_{j=1}^N r_j - Nr_i + N \iff \Gamma \leq N + \sum_{i,j=1,j\neq i}^N \beta_{ji} - N(1 + \sum_{j=1,j\neq i}^N \beta_{ji}) + N. \]

It holds that \( \Gamma \leq N + \sum_{i,j=1,j\neq i}^N \beta_{ji} - N \sum_{j=1,j\neq i}^N \beta_{ji} \) if and only if \( \Gamma \leq \sum_{i,j=1,j\neq i}^N \beta_{ji} + N(1 - \max_i \sum_{j=1,j\neq i}^N \beta_{ji}) \). For the left-hand side of the inequality (3.10), assume \( \rho^*(A(\delta^*)) < 1 \). By replacing (3.5) in the previous inequality, we obtain \( \Gamma > \sum_{j=1}^N r_j - N \). The result
follows after replacing the definition of \( r_j, j \in \{1, \ldots, N\} \). Notice that \( \sum_{i,j=1,j\neq i}^{N} \beta_{ji} \geq \max_i(\sum_{k,j=1,j\neq k}^{N} \beta_{jk} - N \sum_{j=1}^{N} \beta_{ji}) \), then the left side of the inequality (3.10) implies that \( \delta_i^* \geq 0 \). □

3.1.3 Solution bound for unbalanced matrices

When the topology matrix \( I_N + A(1_N) \) is not weight-balanced and there are not enough resources to make it so as in [31], Theorem 3.2 is not applicable. Nonetheless, the virus spread minimization problem can be relaxed to minimizing \( \rho(\bar{A}(\delta)) \), where \( \bar{A}(\delta) = I_N - D + \frac{1}{2}(A(1_N) + A(1_N)^T) \). The next lemma shows that this upper bound to the solution \( \rho^*(A(\delta^*)) \), is at the same time upper bounded by \( \min_\delta \|A(\delta)\| \).

**Lemma 3.1.** Let \( \bar{A}(\delta) = I_N - D + \frac{1}{2}(A(1_N) + A(1_N)^T) \) be the symmetrization of \( A(\delta) \). Then \( \min_\delta \rho(\bar{A}(\delta)) \leq \min_\delta \|A(\delta)\| \) and \( \min_\delta \rho(\bar{A}(\delta)) \geq \rho^*(A(\delta^*)) \).

**Proof.** For the first inequality we use the fact that \( \rho(A) \leq \|A\| \) holds for any matrix \( A \), [35]. Notice that \( \bar{A}(\delta) \) can be expressed as \( \bar{A}(\delta) = \frac{1}{2}(A(\delta) + A^T(\delta)) \). Then, \( \|\bar{A}(\delta)\| = \frac{1}{2}\|A(\delta) + A^T(\delta)\| \leq \frac{1}{2}(\|A(\delta)\| + \|A^T(\delta)\|) = \|A(\delta)\| \), it follows that \( \min_\delta \rho(\bar{A}(\delta)) \leq \min_\delta \|\bar{A}(\delta)\| \leq \min_\delta \|A(\delta)\| \). For the second inequality, we refer the reader to [36], where it is proven that \( \rho(A) \leq \rho(\frac{1}{2}(A + A^T)) \) for any nonnegative matrix \( A \), so the result follows. □

Since it holds that \( \|A(\delta)\| \geq \rho(\bar{A}(\delta)) \geq \rho(A(\delta)) \) as shown in Lemma 3.1, an upper bound for a feasible solution of the \( \delta \)-virus mitigation problem is given by solv-
ing (3.11), which is based on the following lemma.

\[
\min_{\delta \in [0,1]^N} \rho(\tilde{A}(\delta)) \quad \text{s.t.} \quad \sum_{i=1}^N \delta_i = \Gamma, \quad (3.11)
\]

\[\rho(\tilde{A}(\delta)) < 1.\]

**Lemma 3.2.** *The $\delta$-virus mitigation problem is feasible if Problem (3.11) is feasible. In that case, an upper bound to the solution of the $\delta$-virus mitigation problem is given by a solution to Problem (3.11).*

**Proof.** We have that \(\min_{\delta} \|A(\delta)\| < 1 \implies \rho^*(A(\delta)) < 1\), then it holds that \(\min_{\delta} \rho(\tilde{A}(\delta)) < 1 \implies \rho^*(A(\delta)) < 1\). Thus, there exists $\delta$ satisfying the constraints of the $\delta$-virus mitigation problem if there exists $\delta$ satisfying the constraints of Problem (3.11).  

In the next lemma, we describe explicitly the upper bound given by solving Problem (3.11).

**Lemma 3.3.** *Consider a virus dynamics with associated nonnegative and irreducible $A(1_N)$. Let $\Gamma$ satisfy the sufficient condition (3.9) for the topology $I_N + \frac{1}{2}(A(1_N) + A(1_N)^T)$. Then, an upper bound for the solution of the virus spread minimization problem for $I_N + A(1_N)$ is given by $\rho^*(A(\delta^*)) \leq \frac{1}{2N} \left( \sum_{j=1}^{N} (r_j + c_j) - \Gamma \right)$.*

**Proof.** Notice that $\tilde{A}(\delta) = I_N - D + \frac{1}{2}(G + G^T)$, is symmetric, so we can apply Corollary 3.1 for $\rho^*(\tilde{A}(\delta^*))$ under the sufficient condition (3.9) for $\Gamma$. The upper bound follows from $\rho^*(A(\delta^*)) \leq \rho^*(\tilde{A}(\delta^*))$. 

Next, we characterize the distance of these bounds to the solution of the virus spread minimization problem.

**Lemma 3.4.** Consider a virus dynamics with associated nonnegative and irreducible $A(1_N)$. Let $\Gamma$ satisfy (3.9) for the topology $I_N + 1/2(A(1_N) + A(1_N)^T)$. Let $\delta^*, \delta^2$ and $\delta^3$ be the vector solutions given in (3.6) for the topology matrices $I_N + A(1_N)$, $I_N + A(1_N)^T$ and $I_N + 1/2(A(1_N) + A(1_N)^T)$, respectively. Let $\delta^*$ be the solution to the virus spread minimization problem and $e_i = |\delta^* - \delta^3|$, for $i \in V$, be the errors between the solution given by Lemma 3.3 and the optimal solution $\delta^*$. Then, $e_i \leq 1/2|\delta^1 - \delta^2|$, $i \in V$.

*Proof.* Note that the solution given by Lemma 3.3 is the middle point between $\delta^1$ and $\delta^2$. Also note that $\rho(A(\delta^1)) = \rho(A(\delta^2))$. By the convexity of $\rho(A(\delta))$ in Theorem 3.1, we know that $\min\{\delta^1, \delta^2\} \leq \delta^* \leq \max\{\delta^1, \delta^2\}$, so the result follows. $\square$

### 3.2 The constrained Euler replicator algorithm

This section describes the constrained Euler replicator algorithm proposed to solve the virus spread minimization problem. This algorithm is based on the replicator dynamics and a local version of it, see Section 2.0.1, which allows for the automatic satisfaction of the linear resource constraint.

Consider the probabilities of recovery $\delta \in [0, 1]^N$ and the network graph $G$, which are defined in the virus spread minimization problem. In what follows, we assume that the topology matrix $I_N + A(1_N)$ is symmetric, nonnegative, and the graph associated with it is connected. Recall that Theorem 3.2 shows that the solution to a relaxed version
of the virus spread minimization problem is given by a $\delta$ that makes the sum of the rows in matrix $A(\delta)$ to be equal. Motivated by the fact that at the equilibrium of the continuous-time replicator dynamics all fitnesses are equal and other constraints are also naturally satisfied, we want solve the virus spread minimization problem by employing a discretization of these dynamics and by defining local fitness as the $i$th row sum of matrix $A(\delta)$.

Using Euler first-order differences, we discretize the continuous-time local replicator dynamics (2.2),

$$p_i^{(k+1)} = p_i^{(k)} + \epsilon^{(k)}p_i^{(k)}(f_i^{(k)} \sum_{j \in N_i} p_j^{(k)} - \bar{f}_i^{(k)}),$$

(3.12)

where $k \in \mathbb{N}$, $\bar{f}_i^{(k)} = \sum_{j \in N_i} p_j^{(k)} f_j^{(k)}$, and $\epsilon^{(k)} > 0$. Define $p = [\delta_1, \ldots, \delta_N]^T \in [0, 1]^N$ and $f_i(p_i) = r_i - \Gamma p_i = r_i - \delta_i$, $i \in \{1, \ldots, N\}$, where recall that $r_i$ is the $i$th row sum of $I_N + A(1_N)$. Then $f_i(\Gamma p_i) \equiv f_i(\delta_i)$ is the $i$th row sum of $A(\delta)$. In compact form, the dynamics in (3.12) read as:

$$p^{(k+1)} = p^{(k)} + \epsilon^{(k)} P^{(k)} (F^{(k)} A p^{(k)} - \mathcal{A} \bar{f}^{(k)}),$$

(3.13)

where $\mathcal{A}$ is the unweighted adjacency matrix of $G$, $f^{(k)} = [f_1^{(k)}, \ldots, f_N^{(k)}]^T$, $P^{(k)} = \text{diag}(p^{(k)})$, $F^{(k)} = \text{diag}(f^{(k)})$, and $\bar{f}^{(k)} = [p_1^{(k)} f_1^{(k)}, \ldots, p_N^{(k)} f_N^{(k)}]^T$. To solve the problems of interest, we want to keep $\delta \in [0, 1]^N$. Notice that (3.12) does not constraint its states as $p_i \leq \frac{h_i}{\Gamma}$ for certain desired constraints $h_i > 0$ for $i \in \mathcal{V}$ (in our particular virus.
problem $h_i = 1$ for $i \in \mathcal{V}$). Because of this, we propose a variation of (3.12) called the **constrained Euler replicator** algorithm, whose convergence is analyzed in Section 3.3.

A short description of the constrained Euler replicator algorithm is given as follows. Each node computes its own state $\delta_i^{(k+1)} = \Gamma p_i^{(k+1)}$, $i \in \mathcal{V}$. If all trajectories are inside $\Omega \triangleq \{ \delta^{(k)} \in \mathbb{R}^N_+ | 1_h \delta^{(k)} = \Gamma, \delta^{(k)} \leq h_i \}$, then the algorithm reduces to (3.12). Otherwise, if node $i$ does not satisfy the constraint $\delta_i \leq h_i$, then it stores the difference in $\alpha_i \in \mathbb{R} \geq 0$ (i.e., $\alpha_i = \delta_i - h_i$) and puts $\delta_i^{(k+1)} = h_i$ (lines 4 to 6). Note that line 9 restores $\alpha_i/\Gamma$ to $p_i^{(k+1)}$.

### Algorithm 1: Constrained Euler Replicator

```
for $k > 0$ do
  2: Compute $p_i^{(k+1)}$ as in (3.12)
  3: $\delta_i^{(k+1)} = \Gamma p_i^{(k+1)}$
  4: if $\delta_i^{(k+1)} > h_i$ then
    5: $\alpha_i^{(k+1)} = \delta_i^{(k+1)} - h_i$
    6: $\delta_i^{(k+1)} = h_i$
  7: else $\alpha_i^{(k+1)} = 0$
  8: end if
  9: $p_i^{(k+1)} = \frac{\delta_i^{(k+1)} + \alpha_i^{(k+1)}}{\Gamma}$
  10: $k = k + 1$
end for
```

The computation of the step size for the constrained Euler replicator algorithm and the algorithm in (3.13) are discussed in the next section.

### 3.3 Stability analysis

In this section, we analyse the properties of the discrete-time algorithm (3.13) and provide a sufficient condition on $\epsilon^{(k)}$ that guarantees its stability. The algorithm can
be used to solve a relaxed version of the virus spread minimization problem, where the constraint $\delta_i \leq 1$ is omitted. Finally, we analyze the effects of the constrained Euler replicator algorithm when $h_i = 1$, for all $i \in \mathcal{V}$.

Next, we show that the algorithm in (3.13) conserves the most important characteristics of (2.1), i.e., i) the simplex is invariant for small enough step size as shown in Lemma 3.5, ii) all individuals get the same fitness at the equilibrium with the choice of an adequate fitness as shown in Lemma 3.6, and iii) the equilibrium point is asymptotically stable in $\Delta_p$ as shown in Theorem 3.3.

Lemma 3.5. (Invariance of $\Delta_p$ under (3.13)): The dynamics in (3.13) leaves $\Delta_p$ invariant for a sequence $\epsilon^{(k)} < (\max_i, j(f_i^{(k)} - f_j^{(k)}))^{-1}$, $k \geq 0$.

Proof. We have to show two properties to conclude that $\Delta_p$ is invariant under (3.13). For the first property, we multiply by $1^T_N$ on both sides of (3.13)

$$1^T_N p^{(k+1)} = 1^T_N p^{(k)} + \epsilon^{(k)} (\bar{f}^{(k)} A p^{(k)} - p^{(k)} A f^{(k)}),$$

where we use the fact $p^{(k)T} = 1^T_N \text{diag}(p^{(k)})$, and $\bar{f}^{(k)T} = p^{(k)T} \text{diag}(f^{(k)})$. Since the matrix $A$ is symmetric, then we have the property $p^{(k)T} A \bar{f}^{(k)} = f^{(k)T} A p^{(k)}$. Thus, we have that $1^T_N p^{(k+1)} = 1^T_N p^{(k)}$ for any $\epsilon^{(k)} > 0$. For the second property, we rewrite (3.12) as

$$p_i^{(k+1)} = p_i^{(k)} (1 - \Delta_i^{(k)}),$$

where $\Delta_i^{(k)} \equiv \epsilon^{(k)} (f_i^{(k)} - \sum_{j \in N_i} p_j^{(k)})$. Then,

$$\max_i \Delta_i^{(k)} = \epsilon^{(k)} \max_i \sum_{j \in N_i} p_j^{(k)} (f_j^{(k)} - f_j^{(k)}) \leq \epsilon^{(k)} \max_i (f_i^{(k)} - f_j^{(k)}).$$
Then, a sufficient condition for which $\max_i \Delta_i^{(k)} < 1$ is given when $\epsilon^{(k)} < \frac{1}{\max_j (f_i^{(k)} - f_j^{(k)})}$.

□

Lemma 3.6. (Equilibria of (3.13)): Assume that $\Gamma$ satisfies (3.9) and consider the dynamics (3.13) with initial condition $p^{(0)} \in \Delta_p$. Then the equilibrium points $p^*$ of (3.13) such that $p^* > 0$ coincide with those of the continuous-time replicator dynamics.

Proof. Clearly, an equilibrium point is $p^*_i = 0$ for all $i \in \mathcal{V}$, but by assumption we only consider $p^* > 0$. Now, suppose $p^* > 0$, then it must be that

$$\sum_{j \in N_i} p^*_j (f_i^* - f_j^*) = 0,$$  \hspace{1cm} (3.14)

for all $i \in \mathcal{V}$. By assumption we have that $\Gamma$ is large enough, or in other words, the size of $\Gamma$ allows $f_i^* = f_j^*$ for all $i, j \in \mathcal{V}$ and $p^* > 0$. Since $\mathcal{G}$ is connected, then (3.14) holds if $f_i^* = f_j^*$ for all $i, j \in \mathcal{V}$. Now we proceed to show the uniqueness of this solution.

For that, we want to show that the set $J = \{ p \in \mathbb{R}^N \mid f_i(p_i) = f_j(p_j) \forall i, j \in \mathcal{V}, 1_N^T p = 1 \}$ reduces to a single point $p^*$. By assumption $f_i$ is a strictly decreasing monotone function and its inverse $f_i^{-1}$ always exists. Notice that $f_i^{-1}$ is decreasing since $f_i \circ f_i^{-1} = 1$, and

$$\frac{\partial f_i^{-1}}{\partial p_i} = (\frac{\partial f_i}{\partial p_i})^{-1} < 0.$$  \hspace{1cm} (3.15)

Assume that there exits $\hat{\alpha} \neq \bar{\alpha}$ for $\hat{\alpha}, \bar{\alpha} \in \mathbb{R}$ such that $\hat{p} = f^{-1}(\hat{\alpha} 1_N)$, and $\bar{p} = f^{-1}(\bar{\alpha} 1_N)$, which satisfies $1_N^T \hat{p} = 1_N^T \bar{p} = 1$. Write $\hat{p} - \bar{p} = f^{-1}(\hat{\alpha} 1_N) - f^{-1}(\bar{\alpha} 1_N)$. Multiplying last expression by $1_N^T$, it follows

$$\sum_{j=1}^N (f_j^{-1}(\hat{\alpha}) - f_j^{-1}(\bar{\alpha})) = 1_N^T \hat{p} - 1_N^T \bar{p} = 0.$$  \hspace{1cm} (3.15)
Recall that \( f_i \) is strictly decreasing monotone function, i.e., for \( x, y \in \mathbb{R} \) we have that \( f_i(x) < f_i(y) \) iff \( x > y \). Using last fact we have that \( f_i^{-1}(x) < f_i^{-1}(y) \) iff \( x > y \). Without loss of generality assume \( \hat{\alpha} > \bar{\alpha} \) (if the opposite inequality is satisfied, simply switch the roles of \( \hat{\alpha} \) and \( \bar{\alpha} \) in what follows), then \( f_i^{-1}(\hat{\alpha}) < f_i^{-1}(\bar{\alpha}) \) (Recall \( \hat{\alpha} \neq \bar{\alpha} \) by assumption). We have that \( f_i^{-1}(\hat{\alpha}) - f_i^{-1}(\bar{\alpha}) < 0 \), which implies that \( f_i^{-1}(\hat{\alpha}) \) and \( f_i^{-1}(\bar{\alpha}) \) have the same sign. Therefore, the only solution to (3.15) is given when \( \hat{\alpha} = \bar{\alpha} \), and the set \( J \) reduces to a single point \( p^* \).

\[ \square \]

As a consequence of Lemma 3.6, each connected component of a disconnected graph arrives at a common equilibrium fitness. These equilibrium fitnesses can differ from one connected component to another.

**Lemma 3.7.**  
(Equilibrium point characterization under (3.10)): Let \( \mathcal{G} \) be a (not necessarily connected) graph, let \( \Gamma \) satisfy (3.9), and consider the dynamics (3.13) with initial condition \( p^{(0)} \in \Delta_p \). Then, the equilibrium point \( p^* > 0 \) of (3.13) is given by

\[ p_i^* = \frac{1}{|\mathcal{X}|\Gamma}(|\mathcal{X}|r_i + \Gamma \sum_{j \in \mathcal{X}} p_j^{(0)} - \sum_{j \in \mathcal{X}} r_j), \tag{3.16} \]

where \( i \in \mathcal{X} \) and \( (\mathcal{X}, \mathcal{E}_\mathcal{X}) \subset \mathcal{G} \) represents a connected component of \( \mathcal{G} \).

**Proof.** Note that \( \mathcal{G} \) is not necessary connected. In order to get (3.16), we use \( \sum_{j \in \mathcal{X}} p_j^{(k)} = \sum_{j \in \mathcal{X}} p_j^{(0)} \) for all \( k \geq 0 \), which is given by the conservativeness of the states in Lemma 3.5. We have that \( f_i^* = f_j^* \) for all \( i, j \in \mathcal{X} \) by Lemma 3.6. It follows \( \Gamma p_i^* \sum_{j \in \mathcal{X}} 1 = \sum_{j \in \mathcal{X}} (r_i - \)
Consider the Lyapunov function candidate $V$.

Theorem 3.3. (Sufficient conditions for the stability of (3.13)): Suppose $\Gamma$ satisfies (3.10), the set of neighbors in (3.12) are time-variant satisfying $\bigcup_{k \geq k_0} \mathcal{A}(1_N)(k)$ is connected for all $k_0 \in \mathbb{Z}_{\geq 0}$. Then, the algorithm (3.13) converges to the solution of the $\delta$-Virus Mitigation problem, and is asymptotically stable to this solution in $\Delta_\rho$ for a sequence $\epsilon(k) < \min\{(\max_{i,j}(f_i^{(k)} - f_j^{(k)}))^{-1}, (\Gamma \max_i p_i^{(k)})^{-1}\}, k \geq 0$.

Proof. Let $f(p) = [f_1(p_1), \ldots, f_N(p_N)]$ be the vector representation of the fitnesses. Notice that $V$ is a valid Lyapunov function because $f(p^*)$ is a global minimum by Theorem 3.2, and then $V(p^{(k)}) \geq V(p^*)$. Define $\Delta V^{(k)} = V(p^{(k+1)}) - V(p^{(k)})$, for all $k \geq 0$, where $V(p^{(k+1)}) = \max_i f_i^{(k+1)} = \max_i (r_i - \Gamma p_i^{(k+1)})$. Then,

$$\Delta V^{(k)} = \max_{i \in V} f_i^{(k+1)} - \max_{i \in V} f_i^{(k)}$$

$$= \max_{i \in V} (r_i - \Gamma p_i^{(k)}) - \epsilon^{(k)} \Gamma p_i^{(k)} f_i^{(k)} \sum_{j \in N_i} p_j^{(k)} + \epsilon^{(k)} \Gamma p_i^{(k)} p_i^{(k)} - \max_{i \in V} f_i^{(k)}$$

$$= \max_{i \in V} (f_i^{(k)} + \epsilon^{(k)} \Gamma p_i^{(k)} f_i^{(k)} - \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)}) - \max_{i \in V} f_i^{(k)}$$

Let $f_{\max}^{(k)} = \max_{i \in V} f_i^{(k)}$. Then,

$$\Delta V^{(k)} = \max_{i \in V} (f_i^{(k)} + \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} f_j^{(k)} - \epsilon^{(k)} \Gamma p_i^{(k)} f_i^{(k)} \sum_{j \in N_i} p_j^{(k)} - f_{\max}^{(k)})$$

$$= \max_{i \in V} (f_i^{(k)} - \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} f_j^{(k)} - f_{\max}^{(k)} + \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} (f_j^{(k)} + f_{\max}^{(k)} - f_{\max}^{(k)}))$$

$$= \max_{i \in V} (-\epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} (f_{\max}^{(k)} - f_j^{(k)})) - (1 - \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)}) (f_{\max}^{(k)} - f_i^{(k)})).$$

We require $\Delta V^{(k)} \leq 0$, so, it is sufficient to guarantee that $1 - \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} > 0$ since $(f_{\max}^{(k)} - f_i^{(k)}) \geq 0$. For that we choose $\epsilon^{(k)} < \min\left\{\frac{1}{\max_{i,j}(f_i^{(k)} - f_j^{(k)})}, \frac{1}{\Gamma \max_i p_i^{(k)}}\right\}$. Note...
that if \( \max_{i,j}(f_i^{(k)} - f_j^{(k)}) \geq \Gamma \max_i p_i^{(k)} \), then \( \epsilon^{(k)} \Gamma p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} < 1 \) since \( p_i^{(k)} \in \Delta_p \) by Lemma 3.5. On the other hand, if \( \max_{i,j}(f_i^{(k)} - f_j^{(k)}) < \Gamma \max_i p_i^{(k)} \), then \( \epsilon^{(k)} \frac{1}{\Gamma \max_i p_i^{(k)}} \leq \frac{1}{\max_{i,j}(f_i^{(k)} - f_j^{(k)})} \), and \( p_i^{(k)} \in \Delta_p \) by Lemma 3.5. Then, \( \Delta V^{(k)} \leq 0 \) in any case. We have shown that \( \Delta V^{(k)} \) is non-positive. We show next that there is no trajectory that can stay identically at points where \( \Delta V^{(k)} = 0 \) other than the equilibrium. There are three possible scenarios i) at the equilibrium, ii) when \( \max_i \) that if \( \max_i \) with strictly smaller fitness. Next, we characterize \( \gamma \) that is, \( \gamma \) defines the maximum distance between nodes with maximum fitness and nodes with strictly smaller fitness. Next, we characterize \( \gamma \) for time-invariant \( G \). Assume \( \Delta V^{(k)} = 0 \) but trajectories are not at the equilibrium. Define \( \Delta f_i^{(k)} = f_i^{(k+1)} - f_i^{(k)} \). Pick \( i \in V_{eq}^{(k)} \) such that there exists some neighbor \( j \) satisfying \( j \in N_i \cap (V \setminus V_{eq}^{(k)}) \). Note that node \( i \) always exists since the assumption that the graph is connected. Using an analogous procedure as the one to show that \( \Delta V^{(k)} \) is non-positive, and the fact that at least one node \( j \in N_i \) does not have maximum fitness, it follows that \( f_{\max}^{(k)} - f_j^{(k)} < 0 \) for some \( j \in N_i \), so that \( \Delta f_i^{(k)} = f_i^{(k+1)} - f_i^{(k)} = -p_i^{(k)} \sum_{j \in N_i} p_j^{(k)} (f_j^{(k)} - f_i^{(k)}) < 0 \). Then \( |V_{eq}^{(k+1)}| < |V_{eq}^{(k)}| \). Repeat this process \( \gamma \) times (note that \( \gamma \) is an upper bound for the time that it takes a node with maximum fitness to interact with a neighbor having strictly smaller fitness), then after \( m > k + \gamma \) at most the number of agents having \( f_{\max}^{(k)} \) is 1 or else \( f_i^{(m)} < f_{\max}^{(k)} \), which implies \( \Delta V^{(m+k)} < 0 \) if trajectories are not at equilibrium. Therefore, \( V(f^{(k+\gamma)}(p)) - V(f^{(k)}(p)) < 0 \). Note for the time-variant \( G^{(k)} \) the bound \( \gamma \).
can increase with respect to the one shown in the time-invariant $\mathcal{G}$ since there exists the possibility that at some instant a vertex in $\mathcal{V}_{\text{eq}}^{(k)}$ does not have a neighbor with fitness less than $f_{\text{max}}$, but $\gamma$ is finite since the union of graphs is connected as $k \to \infty$, which implies that trajectories do not get trapped in $\Delta V = 0$.

For case (iii), we have that at least one node $i \in \mathcal{V}_{\text{eq}}^{(k)}$ will have a neighbor $j \in \mathcal{V} \setminus \mathcal{V}_{\text{eq}}^{(k)}$ such that $f_{i}^{(k)} < f_{\text{max}}^{(k)}$ as $k \to \infty$, let say that this happens at instant $m + 1 \geq k$, then $\Delta f_{i}^{(m+1)} = f_{i}^{(m+1)} - f_{\text{max}}^{(m)} = -p_{i}^{(m)} \sum_{j \in \mathcal{N}_{i}} P_{j}^{(m)} (f_{\text{max}}^{(m)} - f_{j}^{(m)}) < 0$, then $|\mathcal{V}_{\text{eq}}^{(m+1)}| < |\mathcal{V}_{\text{eq}}^{(m)}|$. The same process occurs for all nodes in $\mathcal{V}_{\text{eq}}^{(k)}$ since the union of graphs is connected as $k \to \infty$. Finally, notice that the set $\{p \in \Delta_{p} \mid V(p) \leq c\}$ is bounded for any $c \in \mathbb{R}$. Therefore, by LaSalle’s invariance principle, every trajectory starting in $\Delta_{p}$ approaches $f^{*}$ as $k \to \infty$. □

**Remark 3.2.** To compute the $\epsilon^{(k)}$ given in Theorem 3.3 in a distributed way, agents can employ a min consensus algorithm. It means that every node takes the minimum of the messages of neighbors and their own. This algorithm has time complexity $\text{diam}(\mathcal{G})$ for fixed graphs $\mathcal{G}$. Therefore, to implement a new iteration of the dynamic equation (3.12), each node first implements a min consensus algorithm during $\text{diam}(\mathcal{G})$ rounds to obtain the new $\epsilon^{(k)}$.

**Remark 3.3.** To have a time-invariant $\epsilon$ given in Theorem 3.3, we can use the fact that $\max_{i,j}(f_{i}^{(k)} - f_{j}^{(k)}) \leq 2 \max_{i}|f_{i}^{(k)}| \leq 2 \max_{i} r_{i} + \Gamma$ provided $p^{(k)} \in (0,1)^{N}$ by Lemma 3.5. Then, $\epsilon$ can be chosen as $\epsilon < (2 \max_{i} r_{i} + \Gamma)^{-1}$. This time-invariant step size can be determined by using a min consensus algorithm before running the constrained Euler...
Repli9ator algorithm.

Remark 3.4. The evolution of \( p^{(k)} \) in the constrained Euler replicator algorithm is equivalent to the evolution of (3.13), then the equilibrium point, properties and the stability analysis already done for (3.13) hold for the constrained Euler replicator algorithm. However, notice when some \( \alpha_i^{(k)} > 0 \) in the constrained Euler replicator algorithm, then \( 1_N \delta^{(k)} \leq \Gamma \), but this can only happen for some finite time since \( p^{(k)} \to p^* \) asymptotically as shown in Theorem 3.3, and for \( k > K \) (i.e., when trajectories are close to the equilibrium) we know that \( 1_N \delta^{(k)} = \Gamma \).

3.4 Simulations

In this section, we illustrate the response of the discrete local replicator dynamics in (3.13) and the constrained Euler replicator algorithm for the symmetric matrix \( A(\delta) \), where \( A_{ii} = 1 - \delta_i \) for \( i \in \{1, \ldots, 5\} \), \( A_{21} = A_{12} = \frac{1}{2} \), \( A_{24} = A_{42} = \frac{1}{3} \), \( A_{34} = A_{43} = \frac{1}{4} \), \( A_{45} = A_{54} = \frac{1}{8} \), and all other entries are set to 0. The equality constraint is given by \( \Gamma = 3 \). In order to exemplify a switching topology, we use a pseudorandom number 0–1 given by a uniform distribution for every edge of the associated adjacency matrix of \( A(\delta) \) during the evolution \( \delta^{(k)} \). In Figure 3.1, we show the behavior of (3.13) for \( \epsilon = 1/2 \) and the initial condition \( \delta^{(0)} = \Gamma[2/768, 56/768, 2/6, 200/768, 254/768] \) and \( \lambda_1(A(\delta^{(0)})) = 1.424 \). Since \( \lambda_1(A(\delta^{(0)})) > 1 \), then the virus is spreading over the network. The optimal value is \( \delta^* = \Gamma[0.2056, 0.3167, 0.1222, 0.2750, 0.0806] \) and the minimizer is \( \lambda_1(A(\delta^*)) = .8833 \). Since \( \lambda_1(A(\delta^*)) < 1 \), we achieve the main objective that is to give
an optimal response to stop the epidemics by our dynamic algorithm.

Notice that, the discrete local replicator dynamics in (3.13) does not satisfy $\delta^{(k)} \leq 1$ for $k > 0$ in general. This fact is exemplified in Figure 3.1, where $\delta_3^{(0)} = 1$, $\delta_2^{(24)} = 1.0466$, and $\delta_4^{(6)} = 1.1077$. This issue is solved by the constrained Euler replicator algorithm, which performance is shown in Figure 3.2 for the same conditions as for Figure 3.1. Figure 3.2 shows that the constrained Euler replicator algorithm converges to the desired equilibrium point of (3.13). Also, note that $\delta_i^{(k)} \leq 1$ for $k \geq 0$ and $i \in \{1, \ldots, 5\}$, which shows that the algorithm constrains its states as expected.

Next example shows the performance of the constrained Euler replicator algorithm for initial conditions $\delta^{(0)} = \Gamma[1/256, 1/2, 127/256]$, $\Gamma = 2$, and the following unbalanced topology matrix $A(\delta)$ given by $A_{ii} = 1 - \delta_i$ for $i \in \{1, \ldots, 3\}$, $A_{21} = \frac{1}{16}$, $A_{12} = \frac{1}{4}$, $A_{32} = A_{23} = \frac{1}{16}$, $A_{31} = \frac{1}{8}$, and all other entries are set to 0. The trajectories for this
Figure 3.2: Evolution of $\delta(k)$ and the eigenvalues of $A(\delta(k))$ using the constrained Euler replicator algorithm.

example are shown in Figure 3.3, where we use $\bar{A}(\delta)$ as shown in Lemma 3.4 to approximate the solution given by the constrained Euler replicator algorithm to the optimal one. Using the same notation of the variables as defined in Lemma 3.4, we get $\delta^3 = [0.70415, 0.70415, 0.5917]$, $\rho(A(\delta^3)) = 0.5010$, $\delta^1 = [17/30, 187/240, 157/240]$, $\delta^2 = [101/120, 151/240, 127/240]$, $\rho(A(\delta^1)) = \rho(A(\delta^2)) = 0.5333$. The optimal value is $\rho^*(A(\delta^*)) = 0.5002$ for $\delta^* = [0.6884, 0.7199, 0.5917]$. This example shows that the expected error is achieved.

The following example is based on the e-mail communication network from the Enron corporation, constructed by taking the first 600 nodes from the dataset available in [37]. We fix the probabilities of transmission to be proportional to the in-degree of each node on the network and symmetrize the network. Figure 3.4 shows the performance of the constrained Euler replicator algorithm for this example. The initial
Figure 3.3: Trajectories of the constrained Euler replicator algorithm for the approximation to the optimum given by Lemma 3.4

conditions for this example are $\delta^{(0)} = \frac{\Gamma}{N} \mathbf{1}_N, \Gamma = 40$. Figure 3.4 shows that starting from $\rho(A(\delta^{(0)})) > 1$ we get $\rho(A(\delta^*)) = .9588$ by using the constrained Euler replicator algorithm. Since $\rho(A(\delta^*))$ is strictly less that 1, then by Lemma 3.2 the solution to the original problem (without the symmetrization) is feasible and thus, the epidemics dies out by Proposition 3.1.

3.5 Summary

We have studied a virus minimization problem for a general SIS model, characterizing an explicit solution to the problem for weight-balanced contagion-dynamics matrices. We have given an strategy that stabilizes the spread for general network topologies when there are enough network resources. Based on that characterization, we have
proposed a novel discrete-time distributed algorithm to stop the infection spread under time varying topologies. Our approach solves the optimization problem by allocating limited immunization resources under the system constraints.

**Publications associated with this chapter**

This chapter contains work previously published in:


Chapter 4

Distributed and robust fair optimization for resource allocation

In this chapter we present and analyze three novel distributed nonlinear algorithms, the robust min-max fairness, robust gradient fairness, and $p$-robust box-gradient fairness algorithms, which allows agents to converge to a solution that aims to improve their individual payoffs while subject to an equality resource constraint. These dynamics use local and anonymous interactions, do not require knowledge of the total resource available to agents in order to converge to the solution, are robust to agents joining or departing the network, and to sporadic changes in the network topology, computation errors, and communication faults. We analyze the robust min-max fairness algorithm convergence over undirected and connected networks, while the robust gradient fairness, and $p$-robust box-gradient fairness algorithms over weight-balanced and strongly connected networks. Finally, we illustrate the applicability of our algorithms
by a virus spreading allocation problem over computer and human networks.

### 4.1 Problem statement, solution approach, and algorithms

In this section, we introduce three related optimization problems, this is followed by the proposed **robust min-max fairness**, **robust gradient fairness**, and \(p\)-**robust box-gradient fairness** algorithms to converge to their solutions under complementary set of assumptions.

#### 4.1.1 Problem statement and solution approach

We consider a network of \(N\) agents connected over a digraph whose goal is to minimize the maximum payoff subject to limited resources. For the first problem we consider interactions given by an undirected graph, while for the last two problems we consider weight-balanced digraphs. The **min-max fairness** problem is given by

\[
\begin{align*}
\min_p & \quad g(p) \\
\text{subject to} & \quad 1_N^T p = 1_N^T u,
\end{align*}
\]

where \(g(p) \triangleq \max_{i \in \{1, \ldots, N\}} f_i(p_i), \ f_i : \mathbb{R} \to \mathbb{R}\) is the payoff, \(p = [p_1, \ldots, p_N]^T \in \mathbb{R}^N\) is the resource allocation, \(u_i \in \mathbb{R}\) is the input assumed to be constant that represents the
available quantity of resources for each agent, and \( u = [u_1, \ldots, u_N]^\top \). The problem will be solved under the following assumptions.

**Assumption 4.1.** (Min-max fairness problem assumptions): We assume that the set of optimal solutions for (4.1) is nonempty, and \( \frac{\partial f_i}{\partial p_i} \) is uniformly bounded. Moreover, it is assumed \( \frac{\partial f_i}{\partial p_i} < 0 \), \( \lim_{p_i \to -\infty} f_i(p_i) = +\infty \), \( \lim_{p_i \to +\infty} f_i(p_i) = -\infty \), and \( \frac{\partial f_k}{\partial p_i} = 0 \) for \( i \neq k \). ◊

**Remark 4.1.** An example payoff that satisfies Assumption 4.1 is the logistic per capita population growth rate \( f_i(p_i) = r_i(1 - \frac{p_i}{k_i}) \), where \( r_i \) and \( k_i \) are constants. This kind of function is commonly used in the literature for resource allocation in virus mitigation [38], engineering applications [39], game theory [40], and convex optimization [41].

An alternative formulation is given when considering a general \( g : \mathbb{R}^N \to \mathbb{R}_{\geq 0} \) in (4.1). We refer to this formulation as coupled fairness problem, which will be solved under the following assumptions.

**Assumption 4.2.** (Coupled fairness problem assumptions): We assume that the set of optimal solutions for the coupled fairness problem is nonempty, \( g \) is convex, continuously differentiable, \( g(p) \to +\infty \) as \( \|p\| \to +\infty \), \( \nabla_p g \) is uniformly bounded, and agent \( i \in V \) should be able to compute \( \frac{\partial g}{\partial p_i} \) using only information from \( N_i^{out} \). ◊

**Remark 4.2.** Some example payoff functions satisfying Assumption 4.2 are given when the objective function is separable and consists of a sum of convex functions. That is the case in power generation, e.g., [10] and [7]), in distributed optimization [11, 5, 42].

When \( p \) is constrained by lower and upper limits, we consider the following
The box-coupled fairness problem,

$$\min_p h(p)$$

subject to

$$\mathbf{1}_N^t p = \mathbf{1}_N^t u$$

$$\underline{p} \leq p \leq \overline{p},$$

(4.2)

where $h : \mathbb{R}^N \to \mathbb{R}_{\geq 0}$ is the payoff, $\underline{p}, \overline{p} \in \mathbb{R}^N$ are the lower and upper limits of the optimization variable. We refer to the last constraint in (4.2) as a box constraint. To solve this problem we state the following assumptions.

**Assumption 4.3.** (box-coupled fairness problem assumptions): We assume that the set of optimal solutions for (4.2) is nonempty, $h(p)$ is convex, continuously differentiable, and agent $i \in V$ should be able to compute $\frac{\partial h}{\partial p_i}$ using only local information from $N_i^{\text{out}}$.

**Remark 4.3.** Assumption 4.3 relaxes Assumption 4.2 since it does not suppose $\nabla_p h(p)$ to be uniformly bounded; nonetheless, this fact is implied since $h$ is locally Lipschitz over a compact set.

**Remark 4.4.** Similar assumptions to solve the box-coupled fairness problem have been made in [9]. However, the proposed algorithm to solve this problem is suboptimal. In Section 4.2, we apply the formulation of the box-coupled fairness problem to solve a virus mitigation problem. Some examples in the virus spreading minimization literature that use similar assumptions for the utility function are given in [14, 17].
Under the same assumptions as for the last problem and using the exact penalty method (see, e.g., [43]), we reformulate the box-coupled fairness problem as follows:

\[
\begin{align*}
\min_p & \; h(p) + J(p) \\
\text{subject to} & \; 1_N^T p = 1_N^T u, \\
\end{align*}
\]  

(4.3)

where \( J(p) \triangleq \epsilon \sum_{i=1}^N (\lceil p_i - \underline{p}_i \rceil_+ + \lfloor p_i - \overline{p}_i \rfloor_+) \) and \( \epsilon \in \mathbb{R}_{>0} \). In what follows we use the following definition.

**Definition 4.1.** We refer to \( F_{p,u} \leq \triangleq \{ p \in \mathbb{R}^N \mid 1_N^T p \leq 1_N^T u \} \), \( F_{p,u} \geq \triangleq \{ p \in \mathbb{R}^N \mid 1_N^T p \geq 1_N^T u \} \), \( F_{\text{box}} \triangleq \{ p \in \mathbb{R}^N \mid p \leq \underline{p} \} \), and \( F_{p,u}^p \triangleq F_{p,u} \cap F_{\text{box}} \).

We adopt a result from [10], where authors show that if

\[
\epsilon > 2 \max_{p \in F_{p,u}^p} \| \nabla_p h(p) \|_\infty,
\]

(4.4)

then the box-coupled fairness problem is equivalent to (4.3).

Under the assumptions we have laid out above, the next lemmas characterize the optimal solution to the min-max fairness, coupled fairness, and box-coupled fairness problems. Proofs for all results in this paper can be found in the Appendix.

**Lemma 4.1.** (Solution of the min-max fairness problem): Let Assumption 4.1, on the payoff characteristics for the min-max fairness problem, hold. Then, there exists a unique global minimizer \( p^* \) to the min-max fairness problem which satisfies

\[
\begin{align*}
1_N^T p^* &= 1_N^T u, \\
f_i(p^*) &= f_j(p^*), \; \forall i, j \in \{1, \ldots, N\}.
\end{align*}
\]

(4.5a)

(4.5b)
Proof. The min-max fairness problem can be reformulated as

\[
\begin{align*}
\min_{s, p} & \quad s \\
\text{subject to} & \quad s \geq \max_i f_i(p_i), \\
& \quad \mathbf{1}_N^T p = \mathbf{1}_N^T u,
\end{align*}
\]

(4.6)

where \( s \in \mathbb{R} \) is a new variable. It holds that \( s \geq \max_i f_i(p_i) \) if and only if \( s \geq f_i(p_i) \) for all \( i \in \{1, \ldots, N\} \). Then, the Lagrangian of (4.6) is given by

\[
L(s, p, \nu, \mu) = s + \nu \sum_{i=1}^N (p_i - u_i) + \sum_{i=1}^N \mu_i (-s + f_i(p_i)),
\]

where \( \nu \in \mathbb{R} \) and \( \mu = [\mu_1, \ldots, \mu_N] \) are the multipliers for the equality and inequality constraints, respectively. Suppose that \( y^* = (\mu^*, s^*, p^*, \nu^*) \) is a solution to (4.6). From complementary slackness, we have \( \mu_i^*(-s^* + f_i(p_i^*)) = 0 \), for all \( i \in \{1, \ldots, N\} \). Moreover,

\[
\left. \frac{\partial \varphi}{\partial s} \right|_{y^*} = 1 - \sum_i^N \mu_i^* = 0, \quad \left. \frac{\partial \varphi}{\partial p_i} \right|_{y^*} = \nu^* + \mu_i^* \left. \frac{\partial f_i}{\partial p_i} \right|_{p^*} = 0,
\]

which implies \( \nu^* \left( \left. \frac{\partial f_i}{\partial p_i} \right|_{p^*} \right)^{-1} = -\mu_i^* \) (notice that \( \left. \frac{\partial f_i}{\partial p_i} \right|_{p^*}^{-1} \) exists by Assumption 4.1, on the payoff characteristics for the min-max fairness problem.) Summing the \( \mu_i^* \), we get \( \nu^* \sum_{i=1}^N \left( \left. \frac{\partial f_i}{\partial p_i} \right|_{p^*} \right)^{-1} = -1 \). Hence, \( \nu^* \neq 0 \) and \( \mu_i^* \neq 0 \). It follows that the only critical point is given when \( s^* = f_i(p_i^*) \) for all \( i \in \{1, \ldots, N\} \), i.e., when \( f_i = f_j \) for all \( i, j \in \{1, \ldots, N\} \) satisfying \( \mathbf{1}_N^T p = \mathbf{1}_N^T u \).

Next, we show this critical point is a global minimum for (4.1). Recall that \( p \) must satisfy the equality constraint, i.e., \( p \in \{ p \in \mathbb{R}^N \mid \mathbf{1}_N^T p = \mathbf{1}_N^T u \} \). Notice that, by assumption, an increase (resp. decrease) in \( p_i \) results in a decrease (resp. increase) in
\( f_i(p_i) \). If \( p \neq p^* \) then, there is \( p_i \neq p_i^* \). If \( p_i < p_i^* \), then \( f_i(p_i) > f_i(p_i^*) \). If \( p_i > p_i^* \), then there exists \( l \in \{1, \ldots, N\} \) such that \( p_l < p_l^* \) because of the equality constraint. Then for any \( p \in \{ p \in \mathbb{R}^N | \mathbf{1}_N^\top p = \mathbf{1}_N^\top u \} \), \( \max_{i \in \{1, \ldots, N\}} f_i(p_i) \) increases with respect to \( \max_{i \in \{1, \ldots, N\}} f_i(p_i^*), \) when \( p \neq p^* \). Hence, \( f(p^*) \) is a global minimum for (4.1).

Next, we want to show that the set \( J = \{ p \in \mathbb{R}^N | f_i(p_i) = f_j(p_j), \forall i, j \in \{1, \ldots, N\}, \mathbf{1}_N^\top p = \mathbf{1}_N^\top u \} \) reduces to a single point \( p^* \). By assumption \( \frac{\partial f_i}{\partial p_i} < 0 \), then \( f_i \) is a strictly decreasing monotone function. Hence, \( f_i : \mathbb{R} \to \mathbb{R} \) is bijective, its inverse \( f_i^{-1} \) always exists, and it is decreasing since \( f_i \circ f_i^{-1} = 1 \) and \( \frac{\partial f_i^{-1}}{\partial p_i} = (\frac{\partial f_i}{\partial p_i})^{-1} < 0. \)

Next, we show that it reduces to a unique \( p^* \). Assume that there exits \( \hat{a} \neq \bar{a} \) for \( \hat{a}, \bar{a} \in \mathbb{R} \) such that \( \hat{p} = f_i^{-1}(\hat{a}\mathbf{1}_N) \) and \( \bar{p} = f_i^{-1}(\bar{a}\mathbf{1}_N) \), which satisfies \( \mathbf{1}_N^\top \hat{p} = \mathbf{1}_N^\top \bar{p} = \mathbf{1}_N^\top u. \)

Write \( \hat{p} - \bar{p} = f_i^{-1}(\hat{a}\mathbf{1}_N) - f_i^{-1}(\bar{a}\mathbf{1}_N) \). Multiplying the last expression by \( \mathbf{1}_N^\top \), it follows

\[
\sum_{j=1}^{N} (f_i^{-1}(\hat{a}) - f_i^{-1}(\bar{a})) = \mathbf{1}_N^\top \hat{p} - \mathbf{1}_N^\top \bar{p} = 0. \tag{4.7}
\]

Recall that \( f_i \) is strictly decreasing monotone function, i.e., for \( x, y \in \mathbb{R} \) we have that \( f_i(x) < f_i(y) \) if and only if \( x > y \). Using this fact we have that \( f_i^{-1}(x) < f_i^{-1}(y) \) if and only if \( x > y \). Without loss of generality assume \( \hat{a} > \bar{a} \) (if the opposite inequality is satisfied, simply switch the roles of \( \hat{a} \) and \( \bar{a} \) in what follows), then \( f_i^{-1}(\hat{a}) < f_i^{-1}(\bar{a}) \) (recall that \( \hat{a} \neq \bar{a} \) by assumption). We have that \( f_i^{-1}(\hat{a}) - f_i^{-1}(\bar{a}) < 0, \) which implies that \( f_i^{-1}(\hat{a}) \) and \( f_i^{-1}(\bar{a}) \) have the same sign for all \( i \in \{1, \ldots, N\} \). Therefore, the only solution to (4.7) is given when \( \hat{a} = \bar{a} \) and the set \( J \) reduces to a single point \( p^* \). \( \square \)

**Lemma 4.2.** (Solution of the coupled fairness problem): Let Assumption 4.2, on the payoff characteristics for the coupled fairness problem, hold. Then, there exists
a global minimizer $p^*$ to the coupled fairness problem which satisfies

\[
\begin{align*}
1_N^T p &= 1_N^T u, \quad (4.8a) \\
(\nabla_p g(p))_i &= (\nabla_p g(p))_j, \quad \forall i, j \in \{1, \ldots, N\}. \quad (4.8b)
\end{align*}
\]

\[\diamondsuit\]

Proof. Since $g$ is convex and the equality constraint is affine for the coupled fairness problem, then it is a convex problem. Moreover, since the set of optimal solutions is nonempty, then strong duality holds [41]. The Lagrangian is given by

\[
\mathcal{L}(p, \nu) = g(p) + \nu \sum_{i=1}^{N} (p_i - u_i),
\]

where $\nu \in \mathbb{R}$ is the multiplier for the equality constraint. Since the coupled fairness problem is convex and strong duality holds, then a point $p \in \mathbb{R}^N$ is a solution if and only if $p$ satisfies the first-order optimality conditions, and thus the result follows. \[\square\]

Next lemma is a result from applying exact penalty method and the characterization in [10].

**Lemma 4.3.** (Solution of the box-coupled fairness problem): Let Assumption 4.3, on the payoff characteristics for the box-coupled fairness problem, hold. Let $\epsilon \in \mathbb{R}_{>0}$.

Under the assumption

\[\epsilon > 2 \max_{p \in F \cap \alpha} \|\nabla_p h(p)\|_{\infty},\]

the solution to the box-coupled fairness problem satisfies

\[
\begin{align*}
\nu^* 1_N &\in \nabla_p h(p) + \partial J(p), \quad (4.9a) \\
1_N^T p^* &= 1_N^T u, \quad (4.9b)
\end{align*}
\]
where \( \nu \in \mathbb{R} \) is the Lagrange multiplier for the equality constraint of the box-coupled fairness problem.

In the following, we propose three related and distributed algorithms which successfully converge to the solutions of the problems introduced above under the corresponding assumptions. Our schemes take the idea from robust consensus, recently studied in [44], where the agent inputs are assigned not as initial states but as inputs into the dynamical system. We will refer to our method as the robust min-max fairness, the robust gradient fairness, and the \( p \)-robust box-gradient fairness algorithms.

### 4.1.2 Proposed algorithms

In order to solve the min-max fairness problem dynamically, we introduce the following robust min-max fairness algorithm,

\[
\begin{align*}
\dot{w} &= f - \hat{f}, \\
\dot{p} &= L(\dot{w} + w) - (p - u),
\end{align*}
\]

where \( w \in \mathbb{R}^N \) is an internal estimator state, \( L \) is the Laplacian matrix associated to an undirected graph \( G \), \( f = (f_i) \), and \( \hat{f} = (\hat{f}_i) \), where \( \hat{f}_i = \max_{k \in N(i)} f_k \), \( i \in \{1, \ldots, N\} \).

Exploiting the different properties of \( g \) in the coupled fairness problem, we alternatively consider

\[
\begin{align*}
\dot{w} &= -L \nabla_p g(p), \\
\dot{p} &= L(\dot{w} + w) - p + u,
\end{align*}
\]

where the other variables and constants have the same meaning as in (4.10).

Since the cost function of the box-coupled fairness problem is assumed non-smooth but convex, the previous algorithm can be adapted as the following \( p \)-robust
box-gradient fairness algorithm,

\begin{align}
\dot{w} & \in -L\hat{h}(p), \\
\dot{p} & = \dot{w} + w - p + u,
\end{align}

where \( \hat{h}(p) : \mathbb{R}^N \to \mathbb{R} \) is defined as \( \hat{h}(p) = h(p) + J(p) \). Notice that this function is convex, locally Lipschitz and regular [45], with generalized gradient \( \partial \hat{h}(p) : \mathbb{R}^N \Rightarrow \mathbb{R}^N \) given by \( \partial \hat{h}(p) = \nabla_p h(p) + \partial J(p) \), where

\[
(\partial J(p))_i = \begin{cases}
-\epsilon, & \text{if } p_i < \underline{p}_i \\
[-\epsilon, 0], & \text{if } p_i = \underline{p}_i \\
0, & \text{if } \underline{p}_i < p_i < \overline{p}_i \\
[0, \epsilon], & \text{if } p_i = \overline{p}_i \\
\epsilon, & \text{if } p_i > \overline{p}_i.
\end{cases}
\]

4.1.3 Stability analysis

One can show that the equilibrium points for the robust min-max fairness, the robust gradient fairness, and the \( p \)-robust box-gradient fairness dynamics coincide with the optimal solution given for the min-max fairness, coupled fairness, and box-coupled fairness problems under the stated assumptions and connectivity of graphs, respectively. In Theorems 4.1, 4.2, and 4.3 we present the stability analysis of such dynamics.

Lemma 4.4. (Equilibria of the robust min-max fairness algorithm): Let Assumption 4.1, on the payoff characteristics for the min-max fairness problem, hold. Assume that \( \mathcal{G} \) is
undirected and connected. Then, the robust min-max fairness algorithm has a unique solution \( p^* \) to
\[
\begin{align*}
\mathbf{1}_N^T p & = \mathbf{1}_N^T u, \\
f_i(p_i) & = f_j(p_j), \quad \forall i, j \in \{1, \ldots, N\},
\end{align*}
\]
(4.13a) (4.13b)

\begin{proof}
To obtain (4.13a), we have that at the equilibrium \( Lw^* - (p^* - u) = 0 \). It follows
\[
\mathbf{1}_N^T p^* = \mathbf{1}_N^T u + \mathbf{1}_N^T Lw^*.
\]
Since \( \mathcal{G} \) is undirected then \( \mathbf{1}_N^T L = 0 \), then (4.13a) follows.

In order to get (4.13b), notice that at the equilibrium \( \dot{w} = 0 \), which implies \( f = \hat{f} \).

Then, \( f_i = f_j \) for all \( i, j \in \{1, \ldots, N\} \) is a solution.

The uniqueness of the solution follows from Lemma 4.1.
\end{proof}

\textbf{Lemma 4.5.} (Equilibria of the robust gradient fairness algorithm): Let Assumption 4.2, on the payoff characteristics for the coupled fairness problem, hold. Assume that \( \mathcal{G} \) is weight-balanced and strongly connected. Then, the robust gradient fairness algorithm has a solution \( p^* \) to
\[
\begin{align*}
\mathbf{1}_N^T p & = \mathbf{1}_N^T u, \\
(\nabla_{p} g(p^*))_i & = (\nabla_{p} g(p^*))_j, \quad \forall i, j \in \{1, \ldots, N\},
\end{align*}
\]
(4.14a) (4.14b)

\begin{proof}
Similarly to the proof of Lemma 4.4, \( \mathbf{1}_N^T p^* = \mathbf{1}_N^T u \).

In order to get (4.14b), notice that at an equilibrium \( L^2 \nabla_{p} g(p^*) = 0 \) if and only if \( L \nabla_{p} g(p^*) = 0 \) as shown as follows. For the if part we have \( L v = 0 \), which implies \( v = \alpha \mathbf{1}_N \). Then \( L^2 v = 0 \). For the only if part we have \( L^2 v = 0 \). It implies \( L v = \alpha \mathbf{1}_N \).

Pre-multiplying last expression by \( \mathbf{1}_N^T \) we have \( \mathbf{1}_N^T L v = \alpha N = 0 \). Since \( N \geq 1 \), implies that \( \alpha = 0 \). Therefore \( \ker L = \ker L^2 \).
\end{proof}
It implies that $\frac{\partial g}{\partial p_i} = \frac{\partial g}{\partial p_j}$ for all $i, j \in \{1, \ldots, N\}$ is a solution to $L^2 \nabla_p g(p^*) = 0$. □

**Lemma 4.6.** *(Equilibria of the p-robust box-gradient fairness algorithm):* Let Assumption 4.3, on the payoff characteristics for the coupled fairness problem, hold. Assume that $w(0) = 0$. Let $G$ be weight-balanced and strongly connected. A point $p^*$ is a solution of the p-robust box-gradient fairness algorithm if and only if there exists $\alpha^* \in \mathbb{R}$ such that
\begin{align*}
1^T_N p^* &= 1^T_N u, \\ 
\alpha^* 1^T_N &= \partial \hat{h}(p). 
\end{align*}

*(4.15a)\hspace{1cm} (4.15b)\hspace{1cm} □

**Proof.** Notice that $1^T_N \dot{w} = 0$ since it is assumed that $G$ is weight-balanced. To get (4.15a), we have that at the equilibrium $1^T_N (w^* - p^* + u) = 0$. Since it is assumed $w(0) = 0$, it follows $1^T_N p^* = 1^T_N u$.

In order to get (4.15b), notice that at the equilibrium $L \eta^* = 0$, for $\eta^* \in \partial \hat{h}(p^*)$. From the fact $G$ is weight-balanced and strongly connected, we deduce that $L \eta^* = 0$ implies $\eta^* \in \text{span}\{1_N\}$. □

Before presenting our main results, the next lemma characterizes the invariance of $\mathcal{F}^{p,u}_{\leq}$ and $\mathcal{F}^{p,u}_{\geq}$ with respect to the robust min-max fairness dynamics. For the other proposed dynamics, the invariance of $\mathcal{F}^{p,u}_{\leq}$ and $\mathcal{F}^{p,u}_{\geq}$ holds under their respective assumptions.

**Lemma 4.7.** *(Invariance of the resource constraint under (4.10)):* Let Assumption 4.1, on the payoff characteristics for the min-max fairness problem, hold. Assume that $G$ is undirected. Then the sets $\mathcal{F}^{p,u}_{\leq}$ and $\mathcal{F}^{p,u}_{\geq}$ are strongly positively invariant under the robust min-max fairness dynamics. □
Proof. We analyze the case of the set $\mathcal{F}^{p,u}_\leq$ since for the case of $\mathcal{F}^{p,u}_\geq$, simply switch the $\leq$ inequality by $\geq$ in what follows. Pre-multiplying (4.10b) by $1_N^T$, it follows that $1_N^T \dot{p} = -1_N^T p + 1_N^T u$, which is a linear and first order system in the variable $1_N^T p$. Then $1_N^T p$ converges exponentially to the equilibrium $1_N^T u$, where there are not overshoots present since it is a first order system. Therefore, given $p(0) \in \mathcal{F}^{p,u}_\leq$, we have that $p(t) \in \mathcal{F}^{p,u}_\leq$. □

Remark 4.5. Notice that the condition to have $\mathcal{G}$ weight-balanced in Lemma 4.5 and Lemma 4.6 cannot be relaxed since this connectivity condition implies that the sets $\mathcal{F}^{p,u}_\leq$ and $\mathcal{F}^{p,u}_\geq$ are strongly positively invariant under the respective dynamics. This property will be used in the Lyapunov analysis in the following two theorems. On the other hand, when $\mathcal{G}$ is not weight-balanced, we refer to the reader to Remark 4.9, where it is explained one possibility to overcome this limitation.

Theorem 4.1. (Sufficient conditions for convergence of the robust min-max fairness algorithm): Let Assumption 4.1, on the payoff characteristics for the min-max fairness problem, hold. Assume that $\mathcal{G}$ is undirected and connected. Then for any constant input $u \in \mathbb{R}^N$, and any initial states $p(0), w(0) \in \mathbb{R}^N$ the solutions of the system (4.10) converge asymptotically to the solution of the min-max fairness problem. □

Proof. We want to show that the robust min-max fairness algorithm is $\gamma$-stable with respect to $p$ according to Definition 2.1 by using Lemma 2.1. When convenient, we denote the vector field in (4.10) by $\mathcal{K} : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N}$.

Recall that, by Lemma 4.7, $\mathcal{F}^{p,u}_\leq$ and $\mathcal{F}^{p,u}_\geq$ are strongly positively invariant under
Two possible cases can be distinguished: (i) when \( p(0) \in \mathcal{F}_{\leq} \), and (ii) \( p(0) \in \mathcal{F}_{\geq} \). For case (i) we consider the Lyapunov function candidate \( V : \mathcal{F}_{\leq} \times \mathbb{R}^N \rightarrow \mathbb{R}_{\geq 0} \),

\[
V(p, w) = \max_{i \in V} f_i(p_i) - f(p^*) + 2\|Lw - p + u\|_2,
\]

where \( f(p^*) \triangleq \max_{i \in \{1, \ldots, N\}} f_i(p^*_i) \) (recall that \( f_i(p^*_i) = f_j(p^*_j) \) by Lemma 4.4.) For case (ii) we consider the Lyapunov function

\[
V(p, w) = -\max_{i \in V} f_i(p_i) + f(p^*) + 2\|Lw - p + u\|_2.
\]

Since the analysis for (i) is analogous to that of (ii), we discuss only case (i). Next, we show that \( V(p, w) \) satisfies conditions (2.4a) and (2.4b) in Lemma 2.1. Notice that \( V(p, w) \) satisfies:

- \( V(p, w) > 0 \), if \( p \neq p^* \) and for all \( w \),

- \( V(p^*, w) = 0 \).

One can see the first condition holds by noticing that the solution to the problem:

\[
\min_p \max_{i \in V} f_i(p_i)
\]

subject to

\[
1_N^T p \leq 1_N^T u,
\]

is achieved when \( 1_N^T p = 1_N^T u \). Then, the minimizer of (4.16) is given by Lemma 4.1, where we have shown that \( f(p^*) \) is a global minimum (in the case (ii), we replace “\( \leq \)” by “\( \geq \)” in the inequality of (4.16) and show that \( f(p^*) \) is a global maximum.)

In order to verify the second condition, let define \( V_1(p) \triangleq \max_{i \in V} f_i(p_i) - f(p^*) \) and \( V_2(p, w) \triangleq 2\|Lw - p + u\|_2 \) which are each of the factors of \( V \). Notice that by the last analysis \( V_1(p) \geq 0 \). Moreover, if \( p = p^* \), implies \( \dot{p} = 0 \) in (4.10b), and then \( L(\dot{f}(p^*)) -
\( f(p^*) = L_w - p^* + u \). Thus, \( V(p^*, w) = \max_{i \in V} f_i(p^*) - f(p^*) + 2\|L(\hat{f}(p^*) - f(p^*))\| = 0 \)

if and only if \( p \in \{ p \in \mathbb{R}^N \mid f_i(p_i) = f_j(p_j), \forall i, j \in \{1, \ldots, N\}, L_w - p - u = 0 \} \) for fix \( w \).

Then the second condition is verified, and (2.4a) is satisfied.

By the last analysis together with the fact that \( V(p, w) \) is radially unbounded in \( p \) when we fix \( w \), then condition (2.4b) is satisfied, where \( a_1(\|y\|) \) and \( a_2(\|\theta\|) \) always can be found since \( V \) is positive definite with respect to \( y \triangleq p \) and \( \theta \triangleq [p, L_w - p + u] \) (see, e.g., [46], Lemma 4.3).

Notice that \( V_1(p) \) is differentiable almost everywhere, since it is locally Lipschitz. Let \( I(p) \) be the set of indices \( k \) for which \( f_k(p_k) = V_1(p) \). Since the trajectories of (4.10) are continuous and according to [45], \( \mathcal{L}_K V_1(p) \) is given by

\[
\mathcal{L}_K V_1(p) = \{ q \in \mathbb{R} \mid q = \left[ \sum_{i \in I(p)} \mu_i (\nabla_p f_i(p_i))^\top \right] \hat{p}, \\
\mu_i > 0 \text{ for } i \in I(p), \sum_{i \in I(p)} \mu_i = 1 \}
\]

\[
= \{ q \in \mathbb{R} \mid q = \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \hat{p}_i, \\
\mu_i > 0 \text{ for } i \in I(p), \sum_{i \in I(p)} \mu_i = 1 \}
\]

where we use the fact \( \frac{\partial f_i}{\partial p_i} = 0 \) for \( i \neq j \). It follows that for any \( q_1 \in \mathcal{L}_K V_1(p) \), we have

\[
q_1 = \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} ((L(f - \hat{f}))_i + (L_w - p + u)_i)
\]

\[
= \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) - (\hat{f}_i - \hat{f}_j) + \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} (L_w - p + u)_i
\]

\[
\leq \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) + \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} (L_w - p + u)_i
\]

\[
\leq \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) + \max_{j \in \{1, \ldots, N\}} \left| \max_{j \in \{1, \ldots, N\}} \left| \frac{\partial f_j}{\partial p_j} (L_w - p + u)_j \right| \sum_{i \in I(p)} \mu_i, \right.
\]
where we have used the facts \( f_i = f_j \) and \( f_j = f_i \) when \( i \in I(p) \). Next, we define \( K \triangleq -\text{diag}(\frac{\partial f_i}{\partial p_i}) \). By Assumption 4.1, \( \nabla_p f \) is uniformly bounded, then without loss of generality we can assume \( |\frac{\partial f_i}{\partial p_i}| \leq 0 \). Recall that \( \sum_{i \in I(p)} \mu_i = 1 \) and \( \|v\|_\infty = \max_{i \in \{1, \ldots, N\}} |v_i| \) for any vector \( v \in \mathbb{R}^N \). Then

\[
q_1 \leq \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) + \|Lw - p + u\|_\infty \leq \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) + \|Lw - p + u\|_2.
\]

where we have used \( \|v\|_\infty \leq \|v\|_2 \) and \( \lambda_{\max}(K) \leq 1 \). Next, we analyze \( V_2(p, w) \). Without loss of generality we assume \( Lw - p + u \neq 0 \) since set \( \{Lw - p + u = 0\} \) is invariant. In the case of \( Lw - p + u = 0 \), we omit \( V_2 \), i.e., \( V = V_1 \). Then,

\[
L_K V_2(p, w) = 2 \frac{(Lw - p + u)^\top (Lw - p)}{\|Lw - p + u\|_2} = -2 \frac{(Lw - p + u)^\top (Lw - p + u)}{\|Lw - p + u\|_2} = -2 \|Lw - p + u\|_2.
\]

It follows that for any \( q \in L_K V(p, w) \), we have

\[
q = \sum_{i \in I(p)} \mu_i \frac{\partial f_i}{\partial p_i} \sum_{j \neq i} a_{ij}(f_i - f_j) + \|Lw - p + u\|_2 \leq 2\|Lw - p + u\|_2 - 2\|Lw - p + u\|_2
\]

where we use \( \sum_{j \neq i} a_{ij} > 0 \) since the matrix associated to \( \mathcal{G} \) is irreducible, and hence \( N_i \neq \emptyset \) for any \( i \in \mathcal{V} \), and \( \frac{\partial f_i}{\partial p_i} < 0 \) by assumption.
Let
\[
z(\vartheta) \triangleq \max_{i \in V} \left( \frac{\partial f_i}{\partial p_i} \right) \min_{i \in V} \left( \sum_{j \neq i} a_{ij} (\max_{k \in V} (f_k) - f_j) \right) - \|Lw - p + u\|_2,
\]
which is a continuous function such that \( q \leq z(\vartheta) \leq 0 \), \( z(\vartheta^*) = 0 \), and \( z(\vartheta) < 0 \) when \( \vartheta \neq \vartheta^* \). Then \( \alpha_3(\vartheta) \) always can be found since the function \( z \) is negative definite with respect to \( \vartheta \) and it is a continuous function with respect to its arguments (see, [46], Lemma 4.3). Then (2.4c) is satisfied. Therefore, by Lemma 2.1, the equilibrium set is globally asymptotically \( \gamma \)-stable. \qed

**Theorem 4.2.** (Sufficient conditions for the convergence of the robust gradient fairness algorithm): Let Assumption 4.2, on the payoff characteristics for the coupled fairness problem, hold. Let \( G \) be weight-balanced and strongly connected. Then, for any constant input \( u \in \mathbb{R}^N \), and any initial states \( p(0) \), \( w(0) \in \mathbb{R}^N \) the solutions of the system (4.11) converge asymptotically to the solution of the coupled fairness problem. \end

**Proof.** Consider the Lyapunov function candidate \( V : \mathbb{R}^{2N} \to \mathbb{R}_{\geq 0}, V(p, w) = g(p) + 2\|Lw - p + u\|_2 \), where by assumption \( g(p) \geq 0 \). To show convergence to the set of equilibrium points \( p^*, w^* \), we use Lasalle’s invariant principle. When convenient, we denote the vector field in (4.11) by \( O : \mathbb{R}^{2N} \to \mathbb{R}^{2N} \). Since the trajectories of (4.11) are continuously differentiable and using \( Lw - p + u \neq 0 \) as explained in the proof of Theorem 4.1, then \( \mathcal{L}_O V(p) \) is given by
\[
\mathcal{L}_O V = -\nabla_p g(p)^\top L^2 \nabla_p g(p) + \nabla_p g(p)^\top (Lw - p + u) - 2\|Lw - p + u\|_2 \\
\leq -\nabla_p g(p)^\top L^2 \nabla_p g(p) + \|\nabla_p g(p)\|_2 \|Lw - p + u\|_2 - 2\|Lw - p + u\|_2.
\]
By assumption $\nabla p g(p)$ is uniformly bounded, then without loss of generality we can assume $\|\nabla p g(p)\| \leq 1$. Then,
\[
\mathcal{L}_O V \leq -\nabla p g(p)^T L^2 \nabla p g(p) + \|Lw - p + u\|_2 - 2\|Lw - p + u\|_2 \\
\leq -\nabla p g(p)^T L^2 \nabla p g(p) - \|Lw - p + u\|_2 \\
= -\frac{1}{2} \nabla p g(p)^T (L^2 + L^2) \nabla p g(p) - \|Lw - p + u\|_2 \\
\leq 0,
\]
where we use the facts $x^T L^2 x = x^T L^2^T x$ for any $x \in \mathbb{R}^N$, then $x^T L^2 x = \frac{1}{2} x^T (L^2 + L^2^T) x$. Also, we use the fact that $G$ is weight-balanced if and only if $1^T_N L = 0$, which implies that $G(L^2)$ is weight-balanced. Moreover, using the fact that $G$ is weight-balanced if and only if $L + L^T$ is positive semidefinite [47], then $L^2 + L^2^T$ is positive semidefinite. Since $G$ is strongly connected, it follows that $\mathcal{L}_O V(p) = 0$ if and only if $L \nabla p g(p) = 0$ and $Lw - p + u = 0$. Then, at the equilibrium $p^*$, (4.14) is satisfied.

Next, we show that $V(p, w)$ is radially unbounded when $w \in \mathcal{F}_c$, where $\mathcal{F}_c = \{w \in \mathbb{R}^N \mid 1^T_N w = c\}$ and $c$ is any given constant. We analyze two principal cases for $i, j \in V$: (i) when there exists $p_i \to \pm \infty$ and (ii) when there exists $w_i \to \pm \infty$ and $p$ is bounded. For the first case, recall that, by assumption, $g(p)$ is radially unbounded with respect to $p$, then the result holds. For the second case, we restrict the analysis over the set $w \in \mathcal{F}_c$. Then if there exist $w_i \to +\infty$, then there must exist $j \neq i$ such that $w_j \to -\infty$, and vice versa. This implies that $w = \alpha 1^T_N$ for $\alpha \to \pm \infty$ cannot happen, and thus $\|Lw - p + u\|_2 \to +\infty$ when there exists $i$ such that $w_i \to \pm \infty$. It follows that $V$ is radially unbounded when $w \in \mathcal{F}_c$.

Notice that $\mathcal{F}_c$ is strongly positively invariant under the coupled fairness dynamics by taking $c = 1^T_N w(0)$. Therefore, by LaSalle’s invariance principle, the equilibrium
set is globally asymptotically stable.

Next, in Lemma 4.8, it is shown the invariance of the set formed by the box constraints under the $p$-ROBUST BOX-GRADIENT FAIRNESS dynamics. This is important in applications for which box constraints are hard constraints (see, e.g., Section 4.2).

Next lemma follows a similar approach as in [10].

**Lemma 4.8.** (Invariance of the box constraints): Let Assumption 4.3, on the payoff characteristics for the box-coupled fairness problem, hold. Let $\mathcal{G}$ be weight-balanced and strongly connected. Then, the set $\mathcal{F}_p$ is strongly positively invariant under the $p$-ROBUST BOX-GRADIENT FAIRNESS dynamics provided that $\epsilon \in \mathbb{R}$ satisfies

$$
\epsilon > \frac{1}{\min_{(i,j) \in E} a_{ij}} (2d_{out,max} \max_p \|\nabla_p h(p)\|_\infty + \|w(0) - p(0) + u\|_\infty),
$$

where $d_{out,max} = \max_{i \in V} \sum_{j=1}^N a_{ij}$.

**Proof.** We begin by noting that, if $\epsilon$ satisfies (4.17), there exists $\alpha > 0$ such that

$$
\epsilon > \frac{1}{(\min_{(i,j) \in E} a_{ij})} (2d_{out,max} \max_{p \in \mathcal{F}_p} \|\nabla_p h(p)\|_\infty + \|w(0) - p(0) + u\|_\infty),
$$

where $\mathcal{F}_p \triangleq \{ p \in \mathbb{R}^N | p - \alpha 1_N \leq p \leq p + \alpha 1_N \}$ for $\alpha > 0$. For convenience, we use the notation $\psi : \mathbb{R}^N \to \mathbb{R}^N$ to refer to (4.12b), and $\mathcal{H} : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ to refer to (4.12).

Without loss of generality assume $p(t) \in \mathcal{F}^{p,u}_{\leq}$ (recall that by Lemma 4.7, the set $\mathcal{F}^{p,u}_{\leq}$ and $\mathcal{F}^{p,u}_{\geq}$ are invariant with respect to the $p$-ROBUST BOX-GRADIENT FAIRNESS dynamics). Then, trajectories can only leave the set $\mathcal{F}_p \cap \mathcal{F}^{p,u}_{\leq}$ by violating the box constraints.

We reason by contradiction. Assume that $\mathcal{F}_p \cap \mathcal{F}^{p,u}_{\leq}$ is not strongly positively invariant under the ROBUST GRADIENT FAIRNESS algorithm. This implies that there exists a
boundary point $p_{bd} \in \text{bd}(F^p_{\text{box}} \cap F^{p,a}_{\leq})$, a real number $\gamma > 0$, and a trajectory $t \to p(t)$ obeying (4.12) such that $p(0) = p_{bd}$ and $p(t) \notin F^p_{\text{box}} \cap F^{p,a}_{\leq}$ for all $t \in (0, \gamma)$.

Without loss of generality, assume that $p(t) \in F^{p,a}_{\text{box}} \cap F^{p,a}_{\leq}$ for all $t \in (0, \gamma)$.

Without loss of generality, there must exist a $i \in V$ such that $p_i(0) = \overline{p}_i$ and $p_i(t) > \overline{p}_i$ for all $t \in (0, \gamma)$. This means that there must exist $t \to \psi(t) \in -L\hat{h}(p(t)) + w(t) - p(t) + u$ and $\gamma_1 \in (0, \gamma)$ such that $\psi_i(t) \geq 0$ a.e. in $(0, \gamma_1)$.

Next, we show that this can only happen if $p_j(t) \geq \overline{p}_j$ for all $j \in N^{\text{out}}_i$. Since $p_i(t) > \overline{p}_i$, for all $t \in (0, \gamma_1)$, then $(\hat{\partial}h(p(t)))_i = \{(\nabla_p h(p(t)))_i + \epsilon\}$.

Therefore,

$$\psi_i(t) = -\sum_{j=1}^{N} a_{ij}(\nabla_p h(p(t)))_i + \epsilon - \eta_j(t) + w_i(t) - p_i(t) + u_i,$$

where $\eta_j(t) \in (\hat{\partial}h(p(t)))_j$, for all $j$.

Note that if $p_j(t) \geq \overline{p}_j(t)$, then $\eta_j(t) \leq (\nabla_p h(p(t)))_j + \epsilon$, whereas if $p_j(t) < \overline{p}_j(t)$, then $\eta_j(t) \leq (\nabla_p h(p(t)))_j$. For convenience, denote this latter set of neighbors by $N^C_i$.

Now, we upper bound $\psi_i(t)$ by

$$\psi_i(t) \leq -\sum_{j=1}^{N} a_{ij}(\nabla_p h(p(t)))_i - (\nabla_p h(p(t)))_j - \epsilon \sum_{j \in N^C_i} a_{ij} + w_i(t) - p_i(t) + u_i,$$

$$\leq 2d_{\text{out, max}} \max_p ||\nabla_p h(p)||_\infty + ||w(t) - p(t) + u||_\infty - \epsilon \sum_{j \in N^C_i} a_{ij},$$

$$< 0,$$

where the last inequality follows from (4.17) and the fact that $||w(t) - p(t) + u||_\infty \leq ||w(0) - p(0) + u||_\infty$. To see this last fact let $\phi_i \triangleq (w - p + u)_i$ and take a Lyapunov function $V(\phi) = \max_{i \in V}|\phi_i|$. Notice that $|\phi_i|$ is convex, then $|\phi_i|$ is regular. Since $|\phi_i|$ is locally Lipschitz, then $V(\phi)$ is locally Lipschitz and differentiable almost everywhere. Let $I(\phi)$
be the set of indices \( k \) for which \( |\phi_k| = V \). We follow a similar approach as in Theorem 4.1 for the next computation. For \( q \in \mathcal{L}_V V(\phi) \) we have that

\[
q \leq \sum_{i \in I(\phi)} \mu_i (\dot{w}_i - \dot{p}_i) \xi_i,
\]

where \( \xi_i \in \partial |\phi_i| \) and \( \mu_i \) has the same meaning as in Theorem 4.1. Then,

\[
q \leq - \min_{l \in V} |(w - p + u)_l|,
\]

which shows that \( ||w(t) - p(t) + u||_\infty \leq ||w(0) - p(0) + u||_\infty \).

Hence, \( \psi_i(t) \geq 0 \) only if \( p_j(t) \geq \bar{p}_j \) for all \( j \in \mathcal{N}_i^{\text{out}} \) and so the latter is true on \((0, \gamma_1)\) by continuity of the trajectories. Extending the argument to the neighbors of each \( j \in \mathcal{N}_i^{\text{out}} \), we obtain an interval \((0, \gamma_2) \subset (0, \gamma_1)\) over which all one- and two-hop neighbors of \( i \) have the states greater or equal to their respective maximum limits. Recursively, and since the graph is strongly connected and the number of nodes is finite, we get an interval \((0, \bar{\gamma})\) over which \( p(i) \geq \bar{p} \), which implies \( p(0) = \bar{p} \), contradicting the fact that \( \bar{p} \notin \mathcal{F}_p^{\text{box}} \cap \mathcal{F}_\leq^{p,\text{ut}} \). \( \square \)

The next theorem shows the convergence of the \( p \)-robust box-gradient fairness algorithm.

**Theorem 4.3.** *(Sufficient conditions for convergence of the \( p \)-robust box-gradient fairness algorithm):* Let Assumption 4.3, on the payoff characteristics for the box-coupled fairness problem, hold. Let \( \mathcal{G} \) be weight-balanced and strongly connected. Assume that \( w(0) = 0 \). For any constant input \( u \in \mathbb{R}^N \) and any initial state \( p(0) \), then the solutions of the system (4.12) converge asymptotically to the equilibrium point (4.15). \( \Diamond \)
Proof. For convenience, we use the notation \( \mathcal{H} : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N} \) to refer to (4.12). Since \( \hat{h}(p) \) is locally Lipschitz, then \( \partial \hat{h}(p) \) is nonempty, compact, and convex. Furthermore, the set-valued map \( \mathcal{H} \) is upper semicontinuous, locally bounded, and it takes nonempty, compact, and convex values \([45]\).

Consider the Lyapunov function candidate \( V : \mathbb{R}^{2N} \rightarrow \mathbb{R}_{\geq 0}, \ V(p, w) = \hat{h}(p) + 2\|w - p + u\|_2 \). The set-valued Lie derivative
\[
L_{\mathcal{H}}V = \{ -\psi^T L\psi + \psi^T (w - p + u) - 2\|w - p + u\|_2 \mid \psi \in \partial \hat{h}(p) \}.
\]

Following similar steps as in Theorem 4.2, we have that for any \( q \in L_{\mathcal{H}}V \), and \( \psi \in \partial \hat{h}(p) \),
\[
q \leq -\frac{1}{2} \psi^T (L + L^T)\psi - \|w - p + u\|_2 \leq 0,
\]
where we have used the fact that \( \partial \hat{h}(p) \) is uniformly bounded since \( \hat{h} \) is locally Lipschitz over the compact set \( \mathcal{F}_p^{\text{box}} \), which is strongly positively invariant by Lemma 4.8. Then, without loss of generality we can use \( \|\partial \hat{h}(p)\| \leq 1 \). Note that \( q = 0 \) only at the set of equilibria, otherwise \( q < 0 \). Therefore, we conclude by LaSalle’s invariant principle for nonsmooth functions (see \([45]\)) the equilibrium set is asymptotically stable. \( \square \)

Remark 4.6. Note that the \( p \)-robust box-gradient fairness algorithm requires a special initialization in the states \( w \) while it does not restrict it on \( p \). For this algorithm we are able to obtain a partial robust result that allows us to deal with the box constraints in the more general problem. Future work will be devoted to extend the robustness results to both variables of the algorithm.
4.2 Application to virus spread minimization

As an application, we employ the \textit{p-robust box-gradient fairness} algorithm on a virus spread minimization problem for computer and epidemic networks. Notice that if we remove the box constraints, we can similarly use the \textit{robust gradient fairness} algorithm to solve the same problem. In this section, first we present the problem statement and solution approach of Subsection 4.2.1. Then, we show how the proposed algorithm implements dynamically the solution to the virus spreading minimization problem.

4.2.1 Problem statement and solution approach

Inspired by previous work [48, 17, 49], we consider the following two problems to minimize the effects of virus contagion. The \textit{δ-virus mitigation} problem is defined by

\[
\min_{\delta} \lambda_1(A(\delta, \kappa)),
\]

subject to

\[
\begin{align*}
1^\top_N \delta &= 1^\top_N u_\delta, \\
\delta_i &\in [\underline{\delta}_i, \bar{\delta}_i], \forall i \in \{1, \ldots, N\},
\end{align*}
\]

(4.18)

where \(\kappa\) is fixed, \(1^\top_N u_\delta\) is the total amount of antivirus available, and the constants \(\underline{\delta}_i, \bar{\delta}_i \in [0, 1]\). The \textit{κ-virus mitigation} problem is given

\[
\min_{\delta} \lambda_1(A(\delta, \kappa)),
\]

subject to

\[
\begin{align*}
1^\top_N \delta &= 1^\top_N u_\delta, \\
\delta_i &\in \delta_i, \delta_i, \forall i \in \{1, \ldots, N\},
\end{align*}
\]
\[
\min_{\kappa} \lambda_1(A(\delta, \kappa)),
\]
subject to
\[
1_\mathbb{N}^\top \kappa = 1_\mathbb{N}^\top u_\kappa, \tag{4.19}
\]
\[\kappa_i \in [\underline{\kappa}_i, \overline{\kappa}_i], \forall i \in \{1, \ldots, N\},\]

where \(\delta\) is fixed, \(1_\mathbb{N}^\top u_\kappa\) is the total amount of isolation resources, and the constants \(\kappa_i, \kappa_i \in (0, 1]\).

**Remark 4.7.** We refer to \(\delta^*\) and \(\kappa^*\) as the solution of Problem (4.18) or Problem (4.19), respectively. These solutions maximize (minimize) the exponential decay (grow) rate of (3.1) subject to resource constraints. Moreover, if the solutions \(\delta^*\) or \(\kappa^*\) make the dominant eigenvalue \(\lambda_1(A(\delta^*, \kappa))\) or \(\lambda_1(A(\delta, \kappa^*))\) (for fixed \(\kappa\) or \(\delta\) depending on the problem) strictly negative, then this guarantees that the disease free equilibrium, \(q^* = 0\), is globally exponentially stable.

In order to analyze the solution of the \(\delta\)-virus mitigation and \(\kappa\)-virus mitigation problems, we show in the following two lemmas that \(\lambda_1(A(\delta, \kappa))\) is a convex function of \(\delta\), and \(\kappa\), respectively. Finally, in Lemma 4.11 we explicitly calculate the gradient of \(\lambda_1\) in order to use the \(p\)-robust box-gradient fairness algorithm using the exact penalty method presented in Section 4.1.

**Lemma 4.9 ([34]).** Let \(B\) be nonnegative, and \(C = \text{diag} (\delta_1, \ldots, \delta_N)\). Then, the maximum eigenvalue of \(B + C\), \(\lambda_1(B + C)\), is a convex function of \(C\). \(\diamond\)
Lemma 4.10. ([50] Convexity of $\lambda_1(KG)$): Let $G$ be positive semidefinite, and $K = \text{diag} (\kappa_1, \ldots, \kappa_N)$. Assume that $\kappa_i > 0$ for all $i \in \{1, \ldots, N\}$. Then, $\lambda_1$ of $KG$ is a convex function of $K$.

Remark 4.8. Notice that in general $\text{trace}(G) = 0$ for our virus application, so, in particular, $G$ cannot be positive semidefinite. This fact makes the $\kappa$-VIRUS MITIGATION problem to be non-convex, even when $G$ is symmetric. We refer to [50] for further discussion about the convexity of $\lambda_1(KG)$. To approximate the solution of the $\kappa$-VIRUS MITIGATION problem, we can use the inequality $\lambda_1(KG) \leq (\lambda_1(K^2G^2))^{\frac{1}{2}}$ [51]. When $G$ is symmetric, $G^2$ is positive semidefinite, then $\lambda_1(K^2G^2)$ is convex by Lemma 4.10. Therefore, we can obtain an upperbound of the solution of the $\kappa$-VIRUS MITIGATION problem. The analysis of the error by using this approach is out of scope of this paper.

Lemma 4.9 and Lemma 4.10 show that $\lambda_1(A(\delta, \kappa))$ is a convex function with respect to its arguments. Then we can aim to apply the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm on a fairness problem associated to this function. As shown in Lemma 4.3, information of the gradient of matrix $A(\delta, \kappa)$ is required in order to evaluate if a solution is optimal and in order to implement our algorithm. For that reason, in the following lemma we provide the analysis to obtain such a gradient.

Lemma 4.11. Let $v$, and $s$ be the left and right eigenvectors of the matrix $A(\delta, \kappa)$ as defined in (5.17). Then

$$\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \delta_i} = -c_i \frac{v_i s_i}{v^\top s}$$

(4.20)
and,

\[
\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \kappa_i} = \frac{v_i}{v^\top s} \sum_{j \neq i} \beta_{ij} s_j \tag{4.21}
\]

\[\diamond\]

Proof. First recall that \(\lambda_1(A(\delta, \kappa))\) is simple since \(A\) is irreducible. As a consequence of the implicit function theorem, \(\lambda_1(A(\delta, \kappa))\) is a differentiable function of \(\delta\) and \(\kappa\). Moreover, from [52] (Corollary 2.4, page 185),

\[
\frac{\partial \lambda_1(B)}{\partial b_{ij}} = \frac{v_i s_j}{v^\top s} \tag{4.22}
\]

for a given matrix \(B\) with simple eigenvalue \(\lambda_1(B)\). Then for the matrix \(A(\delta, \kappa)\) we apply the chain rule,

\[
\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \delta_i} = \frac{\partial \lambda_1(A(\delta, \kappa))}{\partial a_{ii}} \frac{\partial a_{ii}}{\partial \delta_i} = -\frac{c_i}{v^\top s} \frac{v_i s_i}{v^\top s}.
\]

And,

\[
\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \kappa_i} = \sum_{j \neq i} \frac{\partial \lambda_1(A(\delta, \kappa))}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial \kappa_i} = \sum_{j \neq i} \frac{v_i s_j}{v^\top s} \beta_{ij}.
\]

\[\diamond\]
4.2.2 Algorithm implementation and simulations

In many practical scenarios it is reasonable to seek for solutions which are decentralized (agents use only local information), scalable (each agent’s resource requirements like memory, computation, and communication are independent of the size of the network), asymptotically correct (if the inputs and the network are both constant, then each agents estimate of the optimal converges to the right value with zero steady-state error), and robust (it is not required a specific initialization and the convergence is not affected by a single erroneous update in the system state.) In particular, this robustness property allows for operators to be included or excluded from the control in online implementations. To dynamically solve the $\delta$-virus mitigation and the $\kappa$-virus mitigation problems given last requirements, we use the $p$-robust box-gradient fairness algorithm together with the exact penalty method as shown in Section 4.1. Since our distributed algorithm approach requires the local computation of the gradient, and the fact that the gradient for the $\delta$-virus mitigation and $\kappa$-virus mitigation problems is not naturally distributed, we approximate it by using a distributed algorithm in a faster time-scale approach. The distributed computation of the Perron-Frobenius eigenvector is out of scope of this paper, but it can be approximated by using the well-known power method algorithm or some variation for its distributed implementation (see, e.g., [53] or [54]).

Remark 4.9. We consider weight-balanced irreducible matrices, as opposed to symmetric irreducible matrices, to cover a broader spectrum of matrices. This consideration is
motivated by the possibility of having an asymmetric placement of edge isolation (e.g., quarantine or firewalls) making the interaction graph directed. If there are enough resources in terms of isolation/quarantine capabilities (i.e., \( \beta \)), then it is possible to balance network interactions by means of a finite-time distributed algorithm presented in [31].

**Theorem 4.4.** (Convergence of the \( p \)-robust box-gradient fairness algorithm to the solutions of the \( \delta \)-virus mitigation problem): Assume that \( G \) is weight-balanced and irreducible. Take \( \delta \triangleq p \), \( u_\delta \triangleq u \), \( \tilde{\delta} \triangleq p \), and \( \tilde{h}(\delta) \triangleq \lambda_1(A(\delta, \kappa)) + J(\delta) \), where \( \kappa \) is fixed for the \( p \)-robust box-gradient fairness algorithm and assume

\[
\epsilon > \frac{1}{\min_{(i,j) \in E} a_{ij}} \left( 2 \max_{i \in V} c_i + \|w(0) - \delta(0) + u\|_\infty \right), \tag{4.23}
\]

and let \( w(0) = 0 \). Then, for any constant input \( u \), and any initial state \( \delta(0) \), the solutions of the system (4.12) converge asymptotically to the solution of the \( \delta \)-virus mitigation problem.

\( \diamond \)

**Proof.** We show that all assumptions are satisfied in Theorem 4.3 to guarantee convergence of the algorithm for this specific virus problem to the equilibrium point:

i) We find \( \epsilon \) that satisfies (4.17). From (6.5) we deduce that \( \|\nabla_\delta \lambda_1(A(\delta, \kappa))\|_\infty \leq \max_{i \in V} c_i \) since the entries of the eigenvectors associated with \( \lambda_1 \) are positive.

ii) \( \tilde{h}_i(\delta) = \lambda_1(A(\delta, \kappa)) + J(\delta) \) is convex. From Lemma 4.9, \( \lambda_1(A(\delta, \kappa)) \) is a convex function of \( \delta \). Using the fact that function \( l \rightarrow [l]^+ \) is also convex, then \( J(\delta) \) is convex.

iii) \( \tilde{h}_i(\delta) \) is locally Lipschitz since \( \lambda_1(A(\delta, \kappa)) \) is continuously differentiable with bounded derivative and the function \( l \rightarrow [l]^+ \) is Lipschitz.
iv) Since other conditions of Assumption 4.3 hold, then by Lemma 4.8, if $\epsilon$ satisfies (4.23), then $\mathcal{F}_{\text{box}}^p$ is strongly positively invariant under the $p$-robust box-gradient fairness dynamics.

Using (i)-(iv), and Theorem 4.3, we conclude that the solutions of the $p$-robust box-gradient fairness algorithm converge asymptotically to the solutions of the $\delta$-virus mitigation problem.

Next, we show that there is a unique solution to the $\delta$-virus mitigation problem. Provided that $A(\delta, \kappa)$ for fix $\kappa$ is irreducible and symmetric, then $\lambda(\delta^*, \kappa)$ is given by a unique point since the normalized right eigenvector associated to $\lambda_1(A)$ is unique. □

Theorem 4.5. (Convergence of the $p$-robust box-gradient fairness to the solutions of the $\kappa$-virus mitigation problem): Assume that $G$ is positive semidefinite, weight-balanced, and irreducible. Take $\kappa \triangleq p$, $u_\kappa \triangleq u$, $\kappa \triangleq p$, $\bar{\kappa} \triangleq \bar{p}$, and $\hat{h}(\kappa) \triangleq \lambda_1(A(\delta, \kappa)) + J(\kappa)$, where $\delta$ is fixed for the $p$-robust box-gradient fairness algorithm and assume

$$
\epsilon > \frac{1}{(\min_{(i,j) \in \mathcal{E}} a_{ij})} (2 \max_{i \in \mathcal{V}} \sum_{j \neq i} b_{ij} + \|w(0) - \kappa(0) + u\|_\infty), \quad (4.24)
$$

and let $w(0) = 0$. Then for any constant input $u$, and any initial state $\kappa(0)$, the solutions of the system (4.12) converge asymptotically to the solution of the $\kappa$-virus mitigation problem. □

Proof. The proof follows similar steps as in the proof of Theorem 4.4. First, notice that all the assumptions in Theorem 4.3 are satisfied to guarantee convergence of the algorithm for this specific virus problem to the equilibrium point. To conclude, it should
be shown that the solution is unique. This step follows along the same steps of Theo-
rem 4.4. □

**Remark 4.10.** An alternative to the virus minimization problems of Section 4.2, is to
consider an infinite-horizon problem like \( \sum_{k=0}^{+\infty} \lambda_1(\delta(t + k))\varepsilon^k \), for fixed \( k, \varepsilon \in (0, 1) \), and
the same constraints as in the \( \delta \)-VIRUS MITIGATION problem for each \( k \). The solution to this
infinite-horizon problem is a sequence \( \delta = (\delta(t), \delta(t + 1), \ldots) \). Note that a minimizer of
this problem is given by \( \delta = (\delta^*, \delta^*, \ldots) \), where \( \delta^* \) is the solution to our original problem.
This is caused by the constant linear overapproximation to the problem. There are two
possible ways in which the algorithm to arrive at \( \delta^* \) can be implemented: implement it:

- **Operators could find \( \delta^* \) off-line and implement it afterwards.** In this case, they
  would actually apply \( \delta^* \) on the system from \( t = 0 \).

- **Operators could implement the algorithm online as time passes and apply a com-
  puted value through the algorithm for each \( t = k \).** In this case the first steps of the
  policy would not be optimal but they will approach the optimal value asymptoti-
  cally.

In both cases, the algorithm can adapt to operators who where available and drop out
or to intermittent interactions.

Next, we show three examples on virus spreading minimization. The first two
examples have the purpose of illustrating the behavior of the \( p \)-ROBUST BOX-GRADIENT
FAIRNESS algorithm over simple graphs. The last example tests the performance of our
algorithm for a real application over an air transportation network.
Example 4.1 (Optimizing in $\delta$). We illustrate the response of the $p$-robust box-gradient fairness algorithm for the weight-balanced graph associated with the following matrix,

$$A(\delta) = \begin{bmatrix} -c_1\delta_1 & 1/16 & 0 & 1/8 \\ 1/16 & -c_2\delta_2 & 5/16 & 0 \\ 0 & 1/8 & -c_3\delta_3 & 1/4 \\ 1/8 & 3/16 & 1/16 & -c_4\delta_4 \end{bmatrix},$$

where $c = [.7, .9, .89, 1]^T$, $u = [1, .2, .9, .6]^T$, $1_N^T u = 2.7$, $\bar{\delta} = 1_N$, and $\underline{\delta} = 0$. In Figure 4.1 and Figure 4.2, we show the behavior of the $p$-robust box-gradient fairness algorithm for initial conditions $\delta(0) = [.2, .1, .3, .1]^T$, $w(0) = 0$. The optimal value is $\delta^* = [0.6036, 0.7133, 0.7042, 0.6784]^T$, and $\lambda_1(A(\delta^*)) = -0.2628$. To illustrate the algorithm robustness, we introduce an erroneous update on the system state at time $t = 20$, where we force $p(20) = [1, .8, .9, .7]^T$ during half of a second. Figure 4.1 shows the trajectories for $\delta$ and $\lambda$, while Figure 4.2 shows the trajectories for $\hat{h}(\delta)$, and the sum of resources through time, i.e., $1_N^T \delta$. Note that the trajectories $\delta(t)$ converge to the desire equilibrium no matter the erroneous updates on the system state we have introduced.

Figure 4.3 shows the trajectories for $\delta$ and $\hat{h}(\delta)$ when agent 1 leaves the network at $t = 30$, and then it returns at $t = 50$. For $t \in [30, 50]$, we have $u_2 + u_3 + u_4 = 1.7$, otherwise $1_N^T u = 2.7$. This figure shows that a new equilibrium is created for $t \in [30, 50]$, which depends on the availability of the resources. Notice that trajectories converge to the right equilibrium before $t = 30$ and after $t = 50$.

Example 4.2 (Optimizing in $\kappa$). In this example we optimize in $\kappa$, while $\delta$ is fixed. For that, we use the symmetric matrix
Figure 4.1: Trajectories of $\delta(t)$ and eigenvalues $\lambda(A(\delta))$ of Example 4.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 20$.

Figure 4.2: Trajectories of the payoff function $\hat{h}(\delta)$ and the sum of resources of Example 4.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 20$. 
Figure 4.3: Trajectories of $\delta$ and $\hat{h}(\delta)$ of Example 4.1 for the $p$-robust box-gradient fairness algorithm. Agent 1 leaves the network at $t = 30$ and returns at $t = 50$.

$$A(\kappa) = -D + K \begin{bmatrix} 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/16 & 0 \\ 0 & 1/16 & 0 & 1/16 \\ 0 & 0 & 1/16 & 0 \end{bmatrix},$$

where we have fixed $D = .05I_N$. In Figure 4.4, we show the behavior of the $p$-robust box-gradient fairness algorithm for initial conditions $\kappa(0) = [.4, .5, .7, .6]^\top$, $w(0) = 0$, $u = [.3, .5, .2, .4]^\top$, $\epsilon = 26.16$, $1_N^\top u = 1.4$, $\bar{\kappa} = [.7, .7, 1, 1]^\top$, and $\kappa = 0.11 N$. The optimal value is $\kappa^* = [0.11, 0.1, 0.19, 1]^\top$, and $\lambda_1(A(\kappa^*)) = -0.0186$. This example illustrates that the set $F^*_\text{box}$ is invariant to the $p$-robust box-gradient fairness dynamics.

In Figure 4.5 we use same initial conditions as for Figure 4.4, but we set $\epsilon = 10$. In this case (4.24) is not satisfied, but the optimal value and the invariance of $F^*_\text{box}$ remain...
Figure 4.4: Trajectories of $\kappa(t)$ and eigenvalue $\lambda_1(A(\kappa))$ of Example 4.2 for the $p$-robust box-gradient fairness algorithm. It is used $\epsilon = 26.16$, which satisfies (4.24).

unchanged. This particular example illustrate that the conditions on the penalty are only sufficient.

Example 4.3 (Virus over an air transportation network). This example shows the performance of the $p$-robust box-gradient fairness algorithm over an air transportation network. This network is composed by the top 100 airports in the world in the sense of flow of passengers passing through them in 2011 according to http://www.theguardian.com/news/datablog/2012/may/04/world-top-100-airports. The idea is to show an hypothetical scenario where a virus is spread through the air transportation network, and the strategy to stop the epidemics is given by isolation or quarantine policies over the airports. This capacity of isolation is represented by the variable $\kappa$. Recall that a lower value in $\kappa$ represents that the airport is more isolated. This network is shown in Figure 4.6, which was constructed using information of location, and routes
Figure 4.5: Trajectories of $\kappa(t)$ and eigenvalue $\lambda_1(A(\kappa))$ of Example 4.2 for the $\rho$-robust box-gradient fairness algorithm. It is used $\epsilon = 10$, which does not satisfy (4.24).

taken from http://openflights.org/data.html. The links are calculated proportionally to the quantity of people of every airport. In Figure 4.7 and Figure 4.8, we show the performance of the algorithm using initial condition $\kappa(0) = 1_N$, $u = 0.21_N$ for all $i \in \{1, \ldots, N\}$, $1_N^\top u = 20$, $\kappa = 0.011_N$, $\bar{k} = 1_N$, $w(0) = 0$, and $\delta = 0.281_N$. In Figure 4.7 it is shown the sum of resources $1_N^\top \kappa(t)$, and $\lambda_1(A(\kappa(t)))$. Notice that $1_N^\top \kappa(t)$ converges exponentially to the value 20, while $\lambda_1(A(\kappa(t)))$ is decreasing over time. At time $t = 0$ we have $\lambda_1 = 5.3967$, and at time $t = 1000$, $\lambda_1 = -0.0297$. Figure 4.8 illustrates the trajectories of $\kappa(t)$, where we observe that all solutions reach steady state and preserve the box constraints.
Figure 4.6: Flight network for the top 100 airports in flow of passengers passing through them in 2011 for Example 4.3.

Figure 4.7: Trajectories of $1^T_N \kappa(t)$ and eigenvalue $\lambda_1(A(\kappa(t)))$ of Example 4.3.
Figure 4.8: Trajectories of $\kappa(t)$ for top 100 airports of Example 4.3.

4.3 Summary

Here, we have studied a class of fair optimization problems by characterizing the solution. Based on that characterization, we have proposed a robust and distributed algorithms which solves such optimization problems. We have proven the correctness of our algorithms by a refinement of partial stability theory for nonsmooth Lyapunov functions and standard stability theory. Finally, we have illustrated the performance of one of the proposed algorithms by a novel application for minimizing the virus spreading in an air transportation network. As future work we would like to study the proposed algorithms for a more general class of graphs and payoff functions.
Publications associated with this chapter

Parts of this chapter have been published in the following works:

- E. Ramírez-Llanos and S. Martínez, “Distributed and robust fair resource allocation applied to virus spread minimization”, ACC, Chicago, IL, USA, 2015

In this chapter, we present and analyze two novel distributed discrete-time non-linear algorithms to solve a class of distributed resource allocation problems. Our approach allows an interconnected group of agents to collectively minimize a global cost function subject to equality and inequality constraints. Under some technical conditions, we show that the algorithms converge to the solution in a practical way as long as the chosen stepsize is sufficiently small. It is shown that the proposed algorithms are convergent to a neighborhood around the equilibrium even when there are temporary errors in communication or computation. Thus, agents do not require global knowledge of total resources in the network or any specific procedure for initialization. In addition, the proposed algorithms do not require separability of the cost function as long as the gradient is estimated in distributed way. We analyze the algorithms over weight-balanced and
strongly connected networks. Finally, we illustrate the applicability of our algorithms on a virus spreading problem over computer and human networks. In this application we approximate the gradient of the cost function by means of the well-known distributed power iteration method.

5.1 Problem statement, solution approach, and algorithms

In this section, we introduce the optimization problem we are set out to solve, which is followed by the proposed robust gradient fairness and $p$-robust box-gradient fairness algorithms with guaranteed convergence to their corresponding optimizer under complementary sets of assumptions.

5.1.1 Problem statement and solution approach

We consider a network of $N$ agents connected over a digraph whose goal is to minimize a general $f : \mathbb{R}^N \rightarrow \mathbb{R}_{\geq 0}$ under resource constraints. The box-coupled fairness optimization problem is given by

\[
\begin{align*}
\min_{p} & \quad f(p) \\
\text{subject to} & \quad \mathbf{1}_N^T p = \mathbf{1}_N^T \bar{u}, \\
& \quad p \in [\underline{p}, \bar{p}]^N,
\end{align*}
\]
where \( f \) is the payoff, \( p = [p_1, \ldots, p_N]^T \in \mathbb{R}^N \) is the resource allocation, \( u_i \in \mathbb{R} \) is the input assumed to be constant that represents the available quantity of resources for each agent, \( \bar{u} = [\bar{u}_1, \ldots, \bar{u}_N]^T \), and \( \underline{p}, \bar{p} \in \mathbb{R}^N \) are the lower and upper limits of the optimization variable, respectively. We name the last constraint in (5.1) as the box constraint.

We simply refer to the problem with the box constraint omitted as the coupled fairness optimization problem. To solve both problems we state the following assumption.

**Assumption 5.1.** (Problem assumptions): We assume that there exist a unique solution to (5.1), the payoff \( f \) is twice continuously differentiable and bounded below. Also, an agent \( i \in \mathcal{V} \) should be able to compute \( \frac{\partial f}{\partial p_i} \) using only local information from \( N_i^{\text{out}} \) and \( \|\nabla_p f(p)\|_\infty \leq M \).

**Remark 5.1.** To solve the box-coupled fairness optimization problem, it is not required to have \( \nabla_p f(p) \) uniformly bounded, nonetheless, this fact is implied since \( h \) is locally Lipschitz over a compact set.

Under the same assumptions as for the last problem and using the exact penalty method (see, e.g., [43]), we reformulate the box-coupled fairness problem as follows:

\[
\begin{align*}
\min_p & \quad \hat{f}(p) \\
\text{subject to} & \quad 1^T_N p = 1^T_N \bar{u},
\end{align*}
\]

(5.2)

where \( \hat{f}(p) \doteq f(p) + J(p) \), \( J(p) \doteq \chi \sum_{i=1}^N ([p_i - p_i]_+ + [p_i - \bar{p}_i]_+) \), and \( \chi \in \mathbb{R}_{>0} \). In what follows we use the following notation. We refer to \( \mathcal{F}_{\leq}^{u,e} \doteq \{ p \in \mathbb{R}^N \mid 1^T_N p \leq 1^T_N \bar{u} + N \epsilon \} \), \( \mathcal{F}_{\geq}^{u,e} \doteq \{ p \in \mathbb{R}^N \mid 1^T_N p \geq 1^T_N \bar{u} - N \epsilon \} \), \( \mathcal{F}^{u,e} \doteq \mathcal{F}_{\leq}^{u,e} \cap \mathcal{F}_{\geq}^{u,e} \), and \( \mathcal{F}_{\text{box}}^{v} = \{ p \in \mathbb{R}^N \mid p - v 1_N \leq \}

\( p \leq \bar{p} + \nu 1_N \) for \( \epsilon \in \mathbb{R}_{\geq 0} \), and \( \nu \in \mathbb{R}_{>0} \).

Under the assumptions we have laid out above, the next lemma characterizes the optimal solution to the box-coupled fairness optimization problem. Next lemma is a result from applying the exact penalty method and the characterization in [10] to the above problem.

**Lemma 5.1.** *(Solution of the box-coupled fairness problem):* Let Assumption 5.1, on the payoff characteristics for the coupled fairness problem, hold. Let \( \chi \in \mathbb{R}_{>0} \) be such that

\[
\chi > 2 \max_{p \in F, u} \| \nabla_p f(p) \|_{\infty}.
\]

Then, the solution to the box-coupled fairness optimization problem satisfies

\[
\begin{align*}
\zeta^* 1_N & \in \nabla_p f(p) + \partial J(p), \\
1_N^T p^* & = 1_N^T \bar{u}, \\
\end{align*}
\]

where \( \zeta \in \mathbb{R} \) is the Lagrange multiplier for the equality constraint of the box-coupled fairness problem.

**Remark 5.2.** A more general equality constraint can be considered for the box-coupled fairness optimization problem. For example, if we consider the equality constraint of the form \( c^T p = 1_N^T \bar{u} \), where \( c_i \in \mathbb{R}_{>0} \). Then, we can use a new variable \( y_i = c_ip_i \) for \( i \in \{1, \ldots, N\} \). The box-coupled fairness optimization problem becomes

\[
\begin{align*}
\min_y & \quad f(\bar{y}) \\
\text{subject to} & \quad 1_N^T y = 1_N^T \bar{u}, \\
& \quad y_i \in [c_ip_i, c_ip_i], \quad \forall i,
\end{align*}
\]
where \( \bar{y} \triangleq [c_1^{-1}y_1, \ldots, c_n^{-1}y_N]^\top \). Notice that \( f \) is still convex with respect to \( y \) since it is a composition of an affine mapping.

Next, we propose two distributed discrete-time algorithms which successfully converge to the solutions of the box-coupled fairness and coupled fairness problems introduced above under the corresponding assumptions. We will refer to them as the robust gradient fairness and \( p \)-robust box-gradient fairness algorithms.

### 5.1.2 Proposed algorithms

In order to solve the coupled fairness problem dynamically, we introduce the following robust gradient fairness algorithm

\[
\begin{align*}
    w^+ &= w - \alpha L \nabla_p f(p) \quad (5.6a) \\
    p^+ &= p + \alpha(-L^2 \nabla_p f(p) + Lw - p + \bar{u}) \quad (5.6b)
\end{align*}
\]

where \( w \in \mathbb{R}^N \) is an internal estimator state, \( \alpha \in (0, 1) \) is the step size, and \( L \) is the Laplacian matrix associated to directed graph \( G \).

Since the cost function of the box-coupled fairness problem is assumed to be nonsmooth but convex, the previous algorithm can be adapted as the \( p \)-robust box-gradient fairness algorithm shown in Algorithm 2 to solve the box-coupled fairness problem, where

\[
\Sigma = \left\{ \begin{array}{l}
    w^+ \in w + \alpha(\xi_{\max} I_N - \xi) \\
    p^+ \in p + \alpha(-L\xi + Lw - p + u),
\end{array} \right. \quad (5.7a)
\]

\[
\Sigma = \left\{ \begin{array}{l}
    w^+ \in w + \alpha(\xi_{\max} I_N - \xi) \\
    p^+ \in p + \alpha(-L\xi + Lw - p + u),
\end{array} \right. \quad (5.7b)
\]
where \( w, \alpha, L, \) and \( u \) have the same meaning as in the previous algorithm, \( \hat{f} \) has the same meaning as in (5.2). \( \xi_{\text{max}} = \{ \xi_i \in (\partial \hat{f}(p))_i \mid i = \arg\max_{i \in V} \max \xi_i \} \), \( u \in \mathbb{R}^N \) is defined as \( u = \bar{u} + \bar{\epsilon} \), where \( \bar{u}_i \in \mathbb{R} \) is the input assumed to be constant that represents the available quantity of resources for each agent, \( \bar{u} = (\bar{u}_1, \ldots, \bar{u}_N)^\top \), and \( \bar{\epsilon} \in \mathbb{R}^N \) is defined as

\[
\bar{\epsilon}_i = \begin{cases} 
-\epsilon, & \text{if } p_i^+ = \overline{p}_i \\
\epsilon, & \text{if } p_i^+ = \underline{p}_i \\
0, & \text{otherwise,}
\end{cases}
\]

where \( \epsilon \in \mathbb{R}_{>0} \) is a given small constant satisfying \( \epsilon \leq \alpha \).

**Algorithm 2** One step of the box-coupled fairness algorithm for agent \( i \in V \)

1: \( \bar{\epsilon}_i = 0 \)
2: Compute \( \Sigma_i \) as in (5.7)
3: if \( p_i^+ = \overline{p}_i \) then
4: \( \bar{\epsilon}_i = -\epsilon \)
5: end if
6: if \( p_i^+ = \underline{p}_i \) then
7: \( \bar{\epsilon}_i = \epsilon \)
8: end if
9: Compute \( p_i^+ \) as in (5.7b)

Notice that \( \hat{f} \) is convex, locally Lipschitz, with generalized gradient \( \partial \hat{f} : \mathbb{R}^N \rightrightarrows \mathbb{R}^N \) given by \( \partial \hat{f}(p) = \nabla_p f(p) + \partial J(p) \). Then \( \xi \) above is an element in the generalized gradient of \( \hat{f} \).

**Remark 5.3.** Notice that the (5.7) does not allow to have \( p_i^+ = \overline{p}_i \) or \( p_i^+ = \underline{p}_i \) since it perturbs \( \Sigma_i \) using a small quantity \( \bar{\epsilon}_i \) at any time this happens.
5.2 Stability analysis

In this section, we show that the equilibrium points of the robust gradient fairness and \(p\)-robust box-gradient fairness dynamics coincide with the optimal solutions of the corresponding problems they solve, respectively, under the stated assumptions when \(\mathcal{G}\) is strongly connected and weight-balanced. Theorem 5.1 and Theorem 5.2 present the stability properties of both dynamics.

**Lemma 5.2. (Equilibria of the robust gradient fairness algorithm):** Let Assumption 5.1, on the payoff characteristics for the min-max fairness problem, hold. Let \(\mathcal{G}\) be a weight-balanced and strongly connected graph. Let the point \(p^*\) represents the solution of the robust gradient fairness algorithm. Then, the robust gradient fairness algorithm has a unique solution \(p^*\) to

\[
1^T_N p = 1^T_N \bar{u},
\]

\[
(\nabla_p f(p))_i = (\nabla_p f(p))_j, \quad \forall i, j \in \{1, \ldots, N\},
\]

\[
1^T_N p = 1^T_N \bar{u}.
\]

**Proof.** The proof follows similar steps as Lemma 5.3. \(\square\)

**Lemma 5.3. (Equilibria of the \(p\)-robust box-gradient fairness algorithm):** Let Assumption 5.1, on the payoff characteristics for the coupled fairness problem, hold. Let \(\mathcal{G}\) be a weight-balanced and strongly connected graph. Let \(\epsilon = 0\) in (5.7). The point \(p^*\) is the solution of the \(p\)-robust box-gradient fairness algorithm if and only if there exists \(\eta^* \in \mathbb{R}\) such that

\[
\eta^* 1_N \in \partial \hat{f}(p^*),
\]

\[
1^T_N p^* = 1^T_N \bar{u}.
\]

**Proof.** To obtain (5.9a), we have that at the equilibrium \(0 \in \xi_{\max}^* - \xi_i^*\) for all \(i \in \mathcal{V}\), where \(\xi \in \partial \hat{f}(p)\). Then \(0 \in \xi_{\max}^* 1_N - \xi^*\).
In order to get (5.9b), we have $0 \in \alpha(-L\xi^* + Lw^* - p^* + \bar{u})$. Pre-multiplying last expression by $1_N^T$, it follows $0 = -1_N^T p^* + 1_N^T \bar{u}$. □

Before presenting our main results, the next lemma characterizes the invariance of $\mathcal{F}_{\leq}^{u,0}$ and $\mathcal{F}_{\geq}^{u,0}$ with respect to the robust gradient fairness dynamics. The same is true for for the sets $\mathcal{F}_{\leq}^{u,\epsilon}$ and $\mathcal{F}_{\geq}^{u,\epsilon}$ with respect to the $p$-robust box-gradient fairness dynamics.

**Lemma 5.4.** (Invariance of the resource constraint under (5.6) and (5.7)): Let Assumption 5.1, on the payoff characteristics for the min-max fairness problem, hold. Let $G$ be a weight-balanced and strongly connected graph. Assume $\alpha \in (0,1)$ in (5.6) and (5.7). Then, the sets $\mathcal{F}_{\leq}^{u,0}$ and $\mathcal{F}_{\geq}^{u,0}$ are strongly positively invariant under the robust gradient fairness dynamics and the sets $\mathcal{F}_{\leq}^{u,\epsilon}$ and $\mathcal{F}_{\geq}^{u,\epsilon}$ are strongly positively invariant under the $p$-robust box-gradient fairness dynamics for given $\epsilon \in \mathbb{R}_{>0}$.

**Proof.** We only analyze the robust gradient fairness dynamics since the analysis the $p$-robust box-gradient fairness dynamics is analogous to the next analysis. Without loss of generality we analyze the case of the set $\mathcal{F}_{\leq}^{u,0}$ since for the case of $\mathcal{F}_{\geq}^{u,0}$ we simply switch the $\leq$ inequality by $\geq$ in what follows. Pre-multiplying (5.6b) by $1_N^T$, it follows that $1_N^T p^* = (1 - \alpha)1_N^T p + \alpha 1_N^T \bar{u}$, which is a linear and first order system in the variable $1_N^T p$. Since $\alpha \in (0,1)$, then $1_N^T p$ converges exponentially to $1_N^T \bar{u}$. Therefore, given $p(0) \in \mathcal{F}_{\leq}^{u,0}$, we have that $p(t) \in \mathcal{F}_{\leq}^{u,0}$. □

**Remark 5.4.** Notice that the condition to have $G$ weight-balanced cannot be relaxed in Lemma 5.2 and Lemma 5.3. Moreover, this connectivity property is a necessary
condition to show that the sets $F_{u,\varepsilon}^<$ and $F_{\geq u,\varepsilon}$ are strongly positively invariant under the respective dynamics. This important property will be used in the analysis of the following two theorems.

**Theorem 5.1.** (Sufficient conditions for convergence of the robust gradient fairness algorithm): Let Assumption 5.1, on the payoff characteristics for the coupled fairness problem, hold. Assume $f$ is radially unbounded and $\nabla^2 f(p) \leq \Gamma I$. Let $\mathcal{G}$ be a weight-balanced and strongly connected graph. For any constant input $\bar{u} \in \mathbb{R}^N$ and any initial state $p(0), w(0)$, the solutions of the system (5.6) converge asymptotically to the equilibrium point (5.8) if $\alpha \in (0, \min\{1, \frac{1}{2\Gamma \sigma_1(L^2)}\})$.

**Proof.** Consider the Lyapunov function $V : \mathbb{R}^{2N} \rightarrow \mathbb{R}_{\geq 0}$, $V(p, w) = V_1 + V_2$, where $V_1 = f(p) - f(p^*)$, $V_2 = (M + c)\|Lw - p + \bar{u}\|_2 + \Gamma \|Lw - p + \bar{u}\|_2^2$, and $c \in \mathbb{R}_{>0}$. We write (5.6b) as

$$p^+ = p + \alpha \phi,$$

where $\phi = -L^2 \nabla_p f(p) + Lw - p + \bar{u}$. Using a second-order Taylor expansion around $p$, there exists $\alpha' \in [0, 1]$ such that

$$f(p^+) = f(p) + \alpha' \nabla_p f(p) + \frac{1}{2} \alpha'^2 \nabla^2 f(p') \phi,$$

where $p' = p + \alpha' \phi$. Let $\Delta V_1 = V_1(p^+) - V_1(p)$, it follows that

$$\Delta V_1 \leq -\alpha \nabla_p f(p)^\top L^2 \nabla_p f(p) + \alpha \nabla_p f(p)^\top (Lw - p + \bar{u})$$

$$+ \alpha^2 \nabla_p f(p)^\top L^2 \nabla^2 f(p') L^2 \nabla f(p)$$

$$+ \alpha^2 (Lw - p + \bar{u})^\top \nabla^2 f(p')(Lw - p + \bar{u}),$$

where we have used the fact that $(x - y)^\top \nabla^2 f(p')(x - y) \geq 0$ for any $x, y \in \mathbb{R}^N$, which
implies \(x^T \nabla^2 f(p')x + y^T \nabla^2 f(p')y \geq x^T \nabla^2 f(p')y + y^T \nabla^2 f(p')x\), so that

\[
\frac{1}{2} \alpha^2 \phi^T \nabla^2 f(p') \phi \leq \alpha^2 \nabla_p f(p)^T L^2 \nabla^2 f(p') L^2 \nabla f(p) \\
+ \alpha^2 (Lw - p + \bar{u})^T \nabla^2 f(p')(Lw - p + \bar{u}).
\]

Consequently,

\[
\Delta V_1 \leq -\frac{1}{2} \alpha \nabla_p f(p)^T (L^2 + L^2) \nabla_p f(p) \\
+ \alpha \nabla_p f(p)^T (Lw - p + \bar{u}) \\
+ \alpha^2 \nabla_p f(p)^T L^2 \nabla^2 f(p') L^2 \nabla f(p) \\
+ \alpha^2 (Lw - p + \bar{u})^T \nabla^2 f(p')(Lw - p + \bar{u}) \\
\leq -\frac{1}{2} \alpha \lambda_2 (L^2 + L^2) \|\nabla_p f(p) - \text{Avg}(\nabla_p f(p))1_N\|_2^2 \\
+ \alpha^2 \Gamma \sigma^2_1 (L^2) \|\nabla_p f(p) - \text{Avg}(\nabla_p f(p))1_N\|_2^2 \\
+ \alpha \nabla_p f(p)^T (Lw - p + \bar{u}) + \alpha^2 \Gamma \|Lw - p + \bar{u}\|_2^2,
\]

where we have use the facts for any vector \(x \in \mathbb{R}^N\), and matrix \(A \in \mathbb{R}^N \times \mathbb{R}^N\), we have \(\|A\|_2 = \sigma_1(A)\) and \(\|Ax\|_2 \leq \|A\|_2\|x\|_2\) (e.g., see [35]). Also, we have used the assumption that \(\nabla^2 f(p) \leq \Gamma I\).

On the other hand,

\[
\Delta V_2 = (M + c)(\|Lw^+ - p^+ + u\|_2 - \|Lw - p + \bar{u}\|_2) \\
+ \Gamma (\|Lw^+ - p^+ + u\|_2^2 - \|Lw - p + \bar{u}\|_2^2) \\
= (M + c)((1 - \alpha)\|Lw - p + \bar{u}\|_2 - \|Lw - p + \bar{u}\|_2) \\
+ \Gamma ((1 - \alpha)^2 \|Lw - p + \bar{u}\|_2^2 - \|Lw - p + \bar{u}\|_2^2) \\
= -\alpha (M + c) \|Lw - p + \bar{u}\|_2 \\
+ \Gamma ((1 - \alpha)^2 - 1) \|Lw - p + \bar{u}\|_2^2.
\]

It follows,

\[
\Delta V \leq -\frac{1}{2} \alpha \lambda_2 (L^2 + L^2) \|\nabla_p f(p) - \text{Avg}(\nabla_p f(p))1_N\|_2^2 \\
+ \alpha^2 \Gamma \sigma^2_1 (L^2) \|\nabla_p f(p) - \text{Avg}(\nabla_p f(p))1_N\|_2^2 \\
+ \alpha \nabla_p f(p)^T (Lw - p + \bar{u}) - \alpha (M + c) \|Lw - p + \bar{u}\|_2 \\
+ \alpha^2 \Gamma \|Lw - p + \bar{u}\|_2^2 + \Gamma ((1 - \alpha)^2 - 1) \|Lw - p + \bar{u}\|_2^2
\]
where in the first inequality for the first term we have used the fact that $G$ is weight-balanced and strongly connected and then $x^T Lx \geq \lambda_2(L + L^T)\|x - \text{Avg}(x)1_N\|_2^2$ for $x \in \mathbb{R}^N$ (see [55]). Also, in the first inequality for the second term we have used that $x^T L^2 \nabla^2 f(p')Lx = (x - \text{Avg}(x)1_N)^T L^T \nabla^2 f(p')L(x - \text{Avg}(x)1_N)$ for $x \in \mathbb{R}^N$.

A sufficient condition to have $\Delta V \leq 0$ is given when $\alpha > 0$, $\alpha < 1$, and $\frac{1}{2} \lambda_2(L^2 + L^2) - \alpha \Gamma \sigma_1^2(L) \geq 0$. It follows that $\alpha \in (0, \min\{1, \frac{\lambda_2(L^2 + L^2)}{2\sigma_1^2(L)}\})$ guarantees $\Delta V \leq 0$.

Next, we show that $V(p, w)$ is radially unbounded over $w \in \mathcal{F}_b$, where $\mathcal{F}_b \doteq \{w \in \mathbb{R}^N \mid 1_N^T w = b\}$ and $b$ is any given constant. To analyze $V(p, w)$ when $\|(p, w)\|_2 \to +\infty$ consider two situations: (i) when there exists $p_i \to +\infty$ for some $i \in \mathcal{V}$ and (ii) when there exists $w_i \to +\infty$ for some $i \in \mathcal{V}$ and $p$ is bounded. If (i) holds, the result follows from $f(p)$ being radially unbounded with respect to $p$. Under (ii), if there exist $w_i \to +\infty$, then there must exist $j \neq i$ such that $w_j \to -\infty$, and vice versa with $w \in \mathcal{F}_b$. This implies that $w \neq d1_N$, for some $d \in \mathbb{R}$ as $\|w\|_2 \to +\infty$ and thus $\|Lw - p + \bar{u}\|_2 \to +\infty$ when there exists $i$ such that $w_i \to +\infty$. It follows that $V$ is radially unbounded over $\mathcal{F}_b$. Notice that $\mathcal{F}_b$ is strongly positively invariant under the robust gradient fairness algorithm by taking $b = 1_N^T w(0)$. Therefore, by the discrete-time Lyapunov Theorem, the equilibrium point is globally asymptotically stable. $\square$
Before presenting our main result in Theorem 5.2, we show five supporting lemmas.

**Lemma 5.5. (Bounds on the subgradients on \( F_{u,0} \)):** Let Assumption 5.1, on the payoff characteristics for the min-max fairness problem, hold for a function \( f : \mathbb{R}^n \to \mathbb{R} \) for some \( n \in \mathbb{Z}_{>0} \). Assume \( f \) satisfies the superquadratic growth condition for \( \gamma \in \mathbb{R}_{>0} \) and let \( p \in F_{u,0} \). Then,

\[
\|\xi - \text{Avg}(\xi)1_N\|_2 \geq \frac{1}{2}\gamma\|p - p^*\|_2,
\]

where \( p \in F_{u,0} \) and \( \xi \in \partial f(p) \).

**Proof.** Using the superquadratic growth condition for any point \( p, p' \in F_{u,0} \), one has

\[
f(p') \geq f(p) + df(p)(p' - p) + \frac{\gamma}{2}\|p' - p\|_2^2
\]

\[
= f(p) + \sup\{v^\top(p' - p) \mid v \in \partial f(p)\}
\]

\[
+ \frac{\gamma}{2}\|p' - p\|_2^2
\]

\[
= f(p) + \sup\{(v - \text{Avg}(v))1_N^\top(p' - p) \mid v \in \partial f(p)\}
\]

\[
+ \frac{\gamma}{2}\|p' - p\|_2^2
\]

where we have used the fact that \( \text{Avg}(v)1_N^\top(p' - p) = 0 \). It follows

\[
f(p') \geq f(p) + \xi^\top(p' - p) + \frac{\gamma}{2}\|p' - p\|_2^2,
\]

where \( \xi \in \partial f(p) \) and \( \bar{\xi} = \xi - \text{Avg}(\xi)1_N \). By taking \( p' = p^* \), we have

\[
f(p^*) \geq f(p) + \bar{\xi}^\top(p^* - p) + \frac{1}{2}\gamma\|p^* - p\|_2^2
\]

\[
\geq f(p) - \|\bar{\xi}\|_2\|p^* - p\|_2 + \frac{1}{2}\gamma\|p^* - p\|_2^2,
\]

where we have used \(-\|x\|_2\|y\|_2 \leq x^\top y \) for any \( x, y \in \mathbb{R}^N \). It follows,

\[
f(p^*) - f(p) \geq -\|\bar{\xi}\|_2\|p^* - p\|_2 + \frac{1}{2}\gamma\|p^* - p\|_2^2,
\]
By convexity recall that \( \hat{f}(p^*) \leq f(p) \), then
\[
0 \geq -\|\tilde{\xi}\|_2\|p^* - p\|_2 + \frac{1}{2}\gamma\|p^* - p\|_2^2,
\]
and thus, (5.10) follows. \( \square \)

Next lemma follows along the lines of the Maximum Theorem.

**Lemma 5.6.** (Continuity of the solutions of the perturbed box-coupled fairness problem): Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function. For \( p \in \mathbb{R}^n \) and \( v \in \mathbb{R} \), let
\[
\begin{align*}
  f^*(v) &= \min \{ f(p) \mid p \in C(v) \}, \\
  C^*(v) &= \operatorname{argmin} \{ f(p) \mid p \in C(v) \},
\end{align*}
\]
where \( C(v) = \{ p \in \mathbb{R}^n \mid \mathbf{1}_n^T p = v \} \). Assume that \( C^*(v) \) is single-valued for all \( v \) and non-empty. Then, \( f^* \) is convex at \( v \) and \( C^* \) is continuous at \( v \).

**Proof.** First we prove that \( C^* \) is uhc. Let \( p_n \in C^*(v_n) \) and \( v_n \in \mathbb{R} \) be sequences converging to \( p \in \mathbb{R}^N \) and \( v \in \mathbb{R} \), respectively. Take a sequence \( q_n \in C(v_n) \) with \( q_n \to \bar{p} \) for \( \bar{p} \in C^*(v) \). Then, \( f(\bar{p}) = \lim_{n \to +\infty} f(q_n) \geq \lim_{n \to +\infty} f(p_n) = f(p) \), which implies \( p \in C^*(v) \). If there is no subsequence of \( p_n \) converging to a finite value, then it can be assumed that \( p_n \) converges to infinity over the line \( \mathbf{1}_N^T p = v \). And so with the extended real number system, we have \( f(\infty) = f(\bar{p}) \). Since \( C^* \) is single-valued, then uhc coincides with continuity by Lemma 2.3.

Next, we show that \( f^* \) is convex. Let \( v, \bar{v} \in \mathbb{R} \), and let \( v' = \lambda v + (1 - \lambda)\bar{v} \) for \( \lambda \in (0, 1) \). Pick \( p \in C^*(v) \) and \( \bar{p} \in C^*(v) \). Let \( p' = \lambda p + (1 - \lambda)\bar{p} \). Notice that \( \text{Gr}(C) \) is a convex set. Since \( p \in C(v) \), \( \bar{p} \in C(\bar{v}) \), and \( C \) has a convex graph, then it must be
$p' \in C(v')$. Since $p'$ is feasible, but not necessarily optimal at $v'$, one has (by convexity of $f$)

$$f''(v') \leq f(p') = f(\lambda p + (1 - \lambda) \bar{p})$$

$$\leq \lambda f(p) + (1 - \lambda) f(\bar{p}) = \lambda f''(v) + (1 - \lambda) f''(\bar{v}),$$

which gives the desired result. □

**Lemma 5.7.** (On the Lipschitz continuity of $C^*$): Let Assumption 5.1 hold for a function $f : \mathbb{R}^n \to \mathbb{R}$ for some $n \in \mathbb{Z}_{>0}$. Assume $f$ satisfies the superquadratic growth condition for $\gamma \in \mathbb{R}_{>0}$ and the subquadratic growth condition for $\Gamma \in \mathbb{R}_{>0}$ with $\Gamma \geq \gamma$. Let $C^*$ be defined as in Lemma 5.6. Assume $C^*(y)$ is non-empty for every $y \in [a, b]$, for given $a, b \in \mathbb{R}$ such that $a \leq b$, $p^* = C^*(a)$. Then, $C^*$ is locally Lipschitz at $a$.

**Proof.** We want to apply Lemma 2.4. Notice that Assumption 5.1 contains the MFC assumption. Also, $f$ satisfies the superquadratic and subquadratic growth. Thus all assumptions on Lemma 2.4 are satisfied.

Using same terminology as in Section 2.0.4 we have $m = 1$, $A = 1_N^T$, $A^# = 1_N(1_N^T1_N)^{-1} = \frac{1}{N} 1_N$, $y, \lambda \in \mathbb{R}$, $p = x$, and $p^* = x_0$. For simplicity we use $a = 0$ (to see it use $\bar{x}_i = x_i - \frac{u_i}{N}$ and obtain $a = 0$.) Then, for $y \in [a, b]$ and $1_N^T x_0 = 0$, we have

$$\Sigma = \{ \sigma \in \partial f(x_0) | \frac{1}{N} 1_N^T \sigma > \lambda y, \lambda 1_N = \sigma_1, \sigma_1 \in \partial f(x_0) \}
$$

$$= \{ \sigma \in \partial f(x_0) | \frac{1}{N} 1_N^T \sigma > \lambda, \lambda = \frac{1}{N} 1_N^T \sigma_1, \sigma_1 \in \partial f(x_0) \}
$$

$$= \{ \sigma \in \partial f(x_0) | 1_N^T \sigma > 1_N^T \sigma_1, \sigma_1 \in \partial f(x_0) \} = \emptyset$$

where we have pre-multiplied $\lambda 1_N = \sigma_1$ by $1_N^T$, so that $N \lambda = 1_N^T \sigma_1$. Therefore, $C^*$ is locally Lipschitz at $a$ with Lipschitz constant given by (2.9). □
Lemma 5.8. (Bound on the subgradient of the box-coupled fairness problem): Let Assumption 5.1 hold for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ for some $n \in \mathbb{Z}_{>0}$. Assume $f$ satisfies the superquadratic growth condition for $\gamma \in \mathbb{R}_{>0}$. Let $C, f^*$ and $C^*$ be defined as in Lemma 5.6. Assume $C^*(v)$ is single-valued for every $v \in [a, b]$, for given $a, b \in \mathbb{R}$ such that $a \leq b$, $p^* = C^*(a)$, and $C^*$ locally Lipschitz at $a$. Then, for $p \in C(v)$, there exists $m \in \mathbb{R}_{>0}$ such that

$$\frac{\gamma^2}{8}\|p - p^*\|^2_2 \leq \|\xi - \text{Avg}(\xi)\|_2^2 + m(1_N^T(p - p^*))^2,$$

(5.11)

where $\xi \in \partial f(p)$.

Proof. Take $p_v \in C(v)$, then by the triangle inequality we have

$$\|p_v - p^*\|_2 \leq \|p_v - p_v^*\|_2 + \|p_v^* - p^*\|_2,$$

where $p_v^* \in C^*(v)$. The idea is to upper bound the right hand side of the last inequality. Using Lemma 5.5 for $\xi_v \in \partial f(p_v)$, one has $\|p_v - p_v^*\|_2 \leq \frac{2}{\gamma}\|\xi_v - \text{Avg}(\xi_v)1_N\|_2$.

It remains to bound the term $\|p_v^* - p^*\|_2$. Let $C_{[a,b]}^* \triangleq \{C^*(v') | v' \in [a, b]\}$. Since $C^*$ is continuous and $[a, b]$ is compact, then by the extreme value theorem, $\max_{v' \in [a,b]}\|p^* - p_v^*\|_2 < +\infty$. Exploiting the locally Lipschitz property of $C^*$ at $a$, there exists $\bar{m} \in \mathbb{R}_{>0}$ such that

$$\max_{\bar{v} \in [a,b]} \frac{\|p^* - p_{\bar{v}}^*\|_2}{1_N^T(p_{\bar{v}}^* - p^*)} \leq \bar{m} < +\infty.$$

Then, for $\xi_{\bar{v}} \in \partial f(p_{\bar{v}})$ we have that

$$\|p_{\bar{v}} - p^*\|_2 \leq \frac{2}{\gamma}\|\xi_{\bar{v}} - \text{Avg}(\xi_{\bar{v}})1_N\|_2 + \bar{m}1_N^T(p_{\bar{v}} - p^*).$$

Thus, for sufficiently large $m$ it follows that

$$\frac{\gamma^2}{8}\|p_{\bar{v}} - p^*\|^2_2 \leq \|\xi_{\bar{v}} - \text{Avg}(\xi_{\bar{v}})1_N\|^2_2 + m(1_N^T(p_{\bar{v}} - p^*))^2,$$
where we have used the fact that \((x - y)^2 \geq 0\) for \(x, y \in \mathbb{R}\), so that \(x^2 + y^2 \geq 2xy\). Therefore, (5.11) follows. \(\square\)

**Lemma 5.9.** (Positive semidefiniteness of class of functions): Let Assumption 5.1 hold for a function \(f : \mathbb{R}^n \to \mathbb{R}\) for some \(n \in \mathbb{Z}_{>0}\). Assume \(\gamma I \leq \nabla^2 f(p) \leq \Gamma I\) for some \(\gamma, \Gamma \in \mathbb{R}_{>0}\) and all \(p \in \mathbb{R}^n\). Let \(C, f^*\) and \(C^*\) be defined as in Lemma 5.6. Assume \(C^*(v)\) is single-valued for every \(v \in [a, b]\), for given \(a, b \in \mathbb{R}\) such that \(a \leq b\), \(p^* = C^*(a)\), and \(C^*\) locally Lipschitz at \(a\). Consider \(V' : \mathbb{R}^n \to \mathbb{R}_{\geq 0}\), \(V' = V_1 + V_4\), where \(V_1(p) = f(p) + J(p) - f(p^*) = \hat{f}(p) - f(p^*)\) and \(V_4(p) = m' \mathbf{1}_n^\top(p - p^*)\), for \(m' \in \mathbb{R}_{>0}\). Then, for \(p \in C(v)\), there exists \(m'\) such that \(V' \geq 0\).

**Proof.** Let \(\hat{f}^* : [a, b] \to \mathbb{R}\), \(\hat{f}^*(v) \triangleq \min\{\hat{f}(p) | p \in C(v)\}\), where \(C(v) = \{p \in \mathbb{R}^n | \mathbf{1}_n^\top p = v\}\). By Lemma 5.6, \(\hat{f}^*\) is convex. Using Lemma 5.10, the set \(\mathcal{F}_{\mathrm{box}}^{\nu}\) is strongly positively invariant under \(\Sigma_1\), so \(\hat{f}^*\) can be extended over an open interval containing \([a, b]\), and then \(\hat{f}^*\) is locally Lipschitz at \(a\).

Let \(\hat{f}^*_{\min} \triangleq \min\{\hat{f}^*(v) | v \in [a, b]\}\). If \(\hat{f}^*(a) = \hat{f}^*_{\min}\) then \(V_1 + V_4\) is positive semidefinite by the convexity of \(\hat{f}^*\) so that any positive \(m'\) works. Suppose that \(\hat{f}^*(a) > \hat{f}^*_{\min}\).

Since \(\hat{f}^*\) is continuous, then by the extreme value theorem (see, e.g., [56] theorem 4.16), \(\hat{f}^*\) attains a minimum in \([a, b]\), and then \(\hat{f}^*_{\min} > -\infty\). By convexity of \(\hat{f}\), we have that \(\hat{f}^*(v) \leq \hat{f}(p)\) for all \(p \in C(v)\) with \(v \in [a, b]\).

We are interested in those cases for which \(V_1 < 0\). For that reason, in the following analysis we assume that \(\hat{f}(p) < f^*(a)\) for \(p \in C(v)\) with \(v \in (a, b]\). Ex-
ploiting the locally Lipschitz property of $\hat{f}^*$ at $a$, then there exists $l \in \mathbb{R}_{>0}$ such that
\[
\frac{\hat{f}^*(a) - f(p)}{(1_N p - a)} \leq \sup_{v \in (a, b]} \frac{\hat{f}^*(a) - f^*(v)}{(v - a)} \leq l < +\infty \text{ for all } p \in C(v) \text{ with } v \in (a, b].
\]
It follows that $V_1 + V_4 \geq 0$ for $m' = l$. □

Next, in Lemma 5.10, it is shown the solutions of the $p$-ROBUST BOX-GRADIENT FAIRNESS dynamics are bounded.

**Lemma 5.10.** *(Boundedness of the p-ROBUST BOX-GRADIENT FAIRNESS dynamics):* Let Assumption 5.1, on the payoff characteristics for the coupled fairness problem, hold. Let $G$ be weight-balanced and strongly connected. Assume that
\[
\chi > \frac{1}{\min_{(i, j) \in E} a_{ij}} (2Md_{\text{out,max}} + \|Lw(0) - p(0) + \bar{u}\|_{\infty}),
\]
and
\[
\alpha < \frac{\min_{i \in V} \{p_i - p_j\}}{2d_{\text{out,max}} (M + \chi) + \|Lw(0) - p(0) + \bar{u}\|_{\infty}},
\]
where $d_{\text{out,max}} = \max_{i \in V} \sum_{j=1}^{N} a_{ij}$. Then, there exists $\nu$ such that
\[
\nu \leq \max\{|\nu_1|, |\nu_2|, \nu_3\}
\]
where $\nu_1 = \max\{1_N^T p(0), 1_N^T \bar{u} + N\epsilon\} - (N-1) \min_j p_j$, $\nu_2 = \min\{1_N^T p(0), 1_N^T \bar{u} + N\epsilon\} - (N-1) \max_j \bar{p}_j$, and $\nu_3 = \alpha(2d_{\text{out,max}} (M + \chi) + \|Lw(0) - p(0) + \bar{u}\|_{\infty})$, for which the set $\mathcal{F}^\nu_{\text{box}}$ is strongly positively invariant under the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm.

**Proof.** First we want to show that if $\chi$ satisfies (5.12), there exists $\nu > 0$ such that
\[
p(k) \in \mathcal{F}^\nu_{\text{box}} \cap \mathcal{F}^{u, \epsilon}_{\leq}, \text{ for } k \in \mathbb{Z}_{>0}.
\]

In the case that any agent $i \in \mathcal{V}$ has a $\bar{\epsilon}_i 
eq 0$, it means that $p_i^*$ has been perturbed by $\pm \epsilon$ and thus it is inside the box constraints. Then, without loss of generality we assume $\bar{\epsilon} = 0$ for all time in the following analysis.
Without loss of generality assume \( p(k) \in \mathcal{F}_{\leq}^{u,e} \), for \( k \in \mathbb{Z}_{>0} \) (recall that by Lemma 5.4, the sets \( \mathcal{F}_{\leq}^{u,e} \) and \( \mathcal{F}_{\geq}^{u,e} \) are invariant with respect to the \( p \)-robust box-gradient fairness dynamics). Then, the trajectories can only leave the set \( \mathcal{F}_{\text{box}}^\nu \cap \mathcal{F}_{\leq}^{u,e} \), for any \( \nu \), by violating the box constraints.

We reason this is not the case by contradiction. Assume that \( \mathcal{F}_{\text{box}}^\nu \) is not strongly positively invariant under the \( p \)-robust box-gradient fairness algorithm. This implies that there exists a boundary point \( p_{bd} \in \text{bd}(\mathcal{F}_{\text{box}}^\nu) \), a integer number \( \gamma > 0 \), and a trajectory \( k \mapsto p(k) \) obeying (5.7b) such that \( p(0) = p_{bd} \) and \( p(k) \notin \mathcal{F}_{\text{box}}^\nu \) for all \( k \in \{1, \ldots, \gamma\} \).

Without loss of generality, assume that \( p(k) \in \mathcal{F}_{\text{box}}^{\nu_1} \), for all \( k \in \{1, \ldots, \gamma\} \), for some \( \nu_1 \in \mathbb{R}_{>0} \) with \( \nu_1 > \nu \).

Without loss of generality, there must exist \( i \in \mathcal{V} \) such that \( p_i(0) = \bar{p}_i + \nu \) and \( p_i(k) > \bar{p}_i + \nu \) for all \( k \in \{1, \ldots, \gamma\} \). Define \( \Delta p = p(k+1) - p(k) \) and we assume that \( \alpha \) is small enough (we characterize the size of \( \alpha \) at the end of the proof) so that \( p(k+1) \) either satisfies the box constraint \( [p, \bar{p}] \) or it remains to be beyond the upper bound of the box constraint \( \bar{p} \). This means that there must exist \( k \) such that \( p(k+1) \in p(k) + \alpha(-L(\partial \hat{f}(p(k))) + Lw(k) - p(k) + \bar{u}) \) and \( \gamma_1 \in \{1, \ldots, \gamma\} \) such that \( \Delta p_i \geq 0 \) over \( \{1, \ldots, \gamma_1\} \). Without loss of generality assume \( \gamma_1 = \gamma \).

Next, we show that this can only happen if \( p_j(k) \geq \bar{p}_j + \nu \) for all \( j \in \mathcal{N}_i^{\text{out}} \) and all \( k \in \{1, \ldots, \gamma_1\} \).

Since \( p_j(k) > \bar{p}_j + \nu \), for all \( k \in \{1, \ldots, \gamma_1\} \), then \( (\partial \hat{f}(p))_i = \{(\nabla p f(p))_i + \chi\} \).
Therefore,

\[ p_i(k + 1) = p_i(k) - \alpha \sum_{j=1}^{N} a_{ij}((\nabla_p f(p))_i + \chi - \eta_j) + \alpha (Lw(k) - p(k) + \bar{u}), \]

where \( \eta_j \in (\partial \hat{f}(p))_j \), for all \( j \).

Note that if \( p_j \geq \bar{p}_j + \nu \), then \( \eta_j \leq (\nabla_p f(p))_j + \chi \), whereas if \( p_j < \bar{p}_j \), then \( \eta_j \leq (\nabla_p f(p))_j \). For convenience, denote this latter set of neighbors by \( N_i^c \). Now, we upper bound \( \Delta p_i \) as

\[
\Delta p_i \leq -\alpha \sum_{j=1}^{N} a_{ij}((\nabla_p f(p))_i - (\nabla_p f(p))_j) - \alpha \chi \sum_{j \in N_i^c} a_{ij} + \alpha (Lw - p + \bar{u})_i
\]
\[
\leq 2\alpha \max_{i \in V} |(\nabla_p f(p))_i| \sum_{j=1}^{N} a_{ij} + \alpha \|Lw - p + \bar{u}\|_{\infty} - \alpha \chi \sum_{j \in N_i^c} a_{ij}
\]
\[
\leq 2\alpha d_{out,max} \|\nabla_p f(p)\|_{\infty} + \alpha \|Lw - p + \bar{u}\|_{\infty} - \alpha \chi \sum_{j \in N_i^c} a_{ij}
\]
\[
\leq 2\alpha d_{out,max} M + \alpha \|Lw(0) - p(0) + \bar{u}\|_{\infty} - \alpha \chi \sum_{j \in N_i^c} a_{ij}
\]
\[
< 0,
\]

where the second last inequality follows fact that \( \|\nabla_p f(p)\|_{\infty} \leq M \) and the last inequality follows from (5.12). Also, we have used the fact that \( \|Lw(k) - p(k) + \bar{u}\|_{\infty} \leq \|Lw(0) - p(0) + \bar{u}\|_{\infty} \). To see this, take a Lyapunov function \( V = \|Lw - p + \bar{u}\|_{\infty} \), and check that \( \Delta V = -\alpha \|Lw - p + \bar{u}\|_{\infty} \) for \( \alpha \in (0, 1) \). In all, \( \Delta p_i \geq 0 \) is possible only if \( p_j \geq \bar{p}_j + \nu \) for all \( j \in N_i^{out} \) on \( t \in \{1, \ldots, \gamma_1\} \).

Extending the argument to the neighbors of each \( j \in N_i^{out} \), we have that \( \gamma_2 \leq \gamma_1 \) for some \( \gamma_2 \) over which all one- and two-hop neighbors of \( i \) have the states greater or equal to their respective maximum limits. Recursively, and since the graph is strongly connected and the number of nodes is finite, there is a \( \tilde{\gamma} \geq 1 \) over which \( p(k) \geq \bar{p} + \nu 1_N \) with \( p(0) \leq \bar{p} + \nu 1_N \), which contradicts the fact that \( p \in F_{\leq}^{u,\epsilon} \). The above argument shows that there exists a finite \( \nu \) such that \( p(k) \in F_{box}^{u,\epsilon} \cap F_{\leq}^{u,\epsilon} \) for all \( k > 0 \).
Next, we characterize $\nu$. We have two cases to analyze. First, assume that there exist $i \in V$ which in the next iteration is going to jump outside the box constraint. The worse-case scenario is given when $p_i(k) = p_{bd}$. Without loss of generality assume $p_i = \bar{p}_i$. Then $p_i(k + 1)$ is upper bounded by

$$p_i(k + 1) \leq p_i(k) + \alpha(\|\!-\!L\xi + Lw(k) - p(k) + \bar{u}\|_{\infty}) \leq \bar{p}_i + \alpha(\|\!-\!L\xi\|_{\infty} + \|Lw(0) - p(0) + \bar{u}\|_{\infty}) \leq \bar{p}_i + \alpha(2d_{out,max}(M + \chi) + \|Lw(0) - p(0) + \bar{u}\|_{\infty}),$$

Second, we follow a similar argument as the one we have shown for the existence of a finite $\nu$. Specifically, we have shown that if there exists $i \in V$ that is leaving from above (below) the box constraints, then all $j \in V \setminus \{i\}$ must leave the box constraints from above (below). Without loss of generality assume that $p \leq p(k)$ for all $k \geq 0$. Assume that there exists $i \in V$ such that $p_i(0) \geq \bar{p}_i$. By lemma 5.4, we have that $1^T_N p(k)$ has an exponential trend to $1^T_N u$, so $1^T_N p(k) \leq \max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\}$. As a worst case scenario assume $p_j = p_{-j}$ for all $j \in V \setminus \{i\}$. Then $p_i(k) \leq \max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\} - \sum_{j \neq i} p_{-j} \leq \max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\} - (N - 1) \min_j p_{-j}$.

By taking the above two cases we have

$$p_i(k) \leq \max\{\max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\} - (N - 1) \min_j p_{-j}, \bar{p}_i\} \leq \max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\} - (N - 1) \min_j p_{-j} \leq \max\{1^T_N p(0), 1^T_N \bar{u} + N\epsilon\} + \alpha(2d_{out,max}(M + \chi) + \|Lw(0) - p(0) + \bar{u}\|_{\infty}),$$

Therefore, (5.14) follows.

We have shown the invariance of the set $\mathcal{F}_\nu \cap \mathcal{F}_{\leq}^{u,\epsilon}$. This argument holds if a small enough $\alpha$ is assumed. The size of $\alpha$ matters to discard the case of a state jump from one side of the box constraint to the other, i.e., if $\|p_i(k) - \bar{p}_i\|_{\infty}$ for some $k > 0$ and $i \in V$, then it is not possible to have $p_i(k + 1) \leq p_{-j}$ or viceversa. Using (5.7b) with $\epsilon = 0$, we obtain the bound $|\Delta p_i| \leq \alpha(2d_{out,max}(M + \chi) + \|Lw(0) - p(0) + \bar{u}\|_{\infty})$. What we want is
to make the right hand side of the last inequality to be strictly less than \( \min_{i \in V} \{ \bar{p}_i - p_i \} \), which is satisfied when (5.13) holds.

\[ \square \]

**Theorem 5.2.** *(partial-ISS of the p-robust box-gradient fairness algorithm):* Let Assumption 5.1, on the payoff characteristics for the coupled fairness problem, hold. Assume \( G \) be a weight-balanced and strongly connected graph. Assume that the p-robust box-gradient fairness algorithm has access to an approximation of the gradient in the form \( \nabla_p f(p) + e \), where \( e \in \mathbb{R}^N \) is the error term for the approximation. Assume that \( e \) is uniformly bounded, i.e., \( \|e(t)\|_2 \leq \alpha K \) for some \( K \in \mathbb{R}_{>0} \). Then, for any constant input \( \bar{u} \in \mathbb{R}^N \) and any initial state \( p(0), w(0) \), the solutions of the system (5.7) converge asymptotically to a ball centered at the equilibrium point (5.9) with radius dependent on \( \alpha \).

**Proof.** Without loss of generality note that \( f \) can be assumed to be strongly convex under Assumption 5.1. To see this, if we define a new convex program by changing the cost function of (5.1) by \( f(p) + \|p - p^*\|_2^2 \), where \( p^* \) is the optimal solution to (5.1), then the solution of the original program and the new one is the same under Assumption 5.1. Furthermore, both problems satisfy the optimality condition (5.4) in Lemma 5.1. Therefore, it can be assumed that \( f \) is strongly convex and that \( \gamma I \leq \nabla^2_p f(p) \leq \Gamma I \) for some \( \gamma, \Gamma \in \mathbb{R}_{>0} \). Notice that the last upper bound on the Hessian matrix of \( f \) comes from the fact that the compact set \( F^\nu_{\text{box}} \) is strongly positively invariant under \( \Sigma \) by Lemma 5.10. Therefore the maximum eigenvalue of \( \nabla^2 f(p) \), which is a continuous function of \( p \) on
\( F_{\text{box}}^{\nu} \) is bounded above on \( F_{\text{box}}^{\nu} \), i.e., there exists a constant \( \Gamma \) such that \( \nabla^2 f(p) \leq \Gamma I \).

**Lyapunov function definition:**

With the help of the previous lemmas, we define a suitable Lyapunov function that is used to derive the stability result. Consider \( V : \mathbb{R}^{2N} \rightarrow \mathbb{R}_{\geq 0}, V = V_1 + V_2 + V_3 + V_4 \), where \( V_1(p, w) = f(p) + J(p) - f(p^*) \equiv \hat{f}(p) - f(p^*) \), \( V_2(p, w) = (M + \chi + \ell) \|Lw - p + u\|_2 + \Gamma \|Lw - p + u\|_2^2 \), for any \( \ell \in \mathbb{R}_{>0} \), \( V_3(p, w) = c_1 m(1_{N}^T(p - p^*))^2 \), \( c_1 = \frac{1}{2} \lambda_2 (L + L^T) \), for some \( m \in \mathbb{R}_{>0} \), and \( V_4(p, w) = m' 1_{N}^T(p - p^*) \), for some \( m' \in \mathbb{R}_{>0} \).

The constant \( m \) in \( V_3 \) is chosen from Lemma 5.8. By this lemma, there exists \( m \) such that \( \|\xi - \text{Avg}(\xi)1_N\|_2^2 + m(1_N^T(p - p^*))^2 \geq \frac{\gamma^2}{8} \|p - p^*\|_2^2 \) holds. To satisfy the assumptions of Lemma 5.8, take \( a = 1_N^T p^* \), \( b = 1_N^T p(0) \), and recall that, by Lemma 5.4, \( F_{\leq}^{u,e} \) and \( F_{\geq}^{u,e} \) are strongly positively invariant under \( \Sigma \), so without loss of generality we assume \( p(0) \in F_{\geq}^{u,e} \). Notice that the segment \([a - N\epsilon, b]\) is strongly positively invariant under \( \Sigma \) by Lemma 5.4.

Analogous to the choice of \( m \), the constant \( m' \) in \( V_4 \) is chosen from Lemma 5.9. By this lemma, there is \( m' \) such that \( V_1 + V_4 \geq 0 \). To satisfy the assumptions in Lemma 5.9 take \( a \) and \( b \) as defined for the case of \( m \). Recall that by assumption \( \chi \) satisfies (5.12) and \( \|\nabla_p f(p)\|_{\infty} \leq M \) for any \( p \in \mathbb{R}^N \).

First, we want to show that the box-coupled fairness algorithm is \( \gamma \)-ISS-stable with respect to \( p \) using Theorem 2.1. For that, we show that \( V(p, w) \) satisfies conditions (2.7a) and (2.7b) in Theorem 2.1, and then this holds for the dynamics \( \Sigma \). Notice that \( V \) satisfies:
\begin{itemize}
  \item $V(p, w) \geq 0$, for all $p, w$.
  \item $V(p, w) = 0$ if $p = p^*$ and for all $w$.
\end{itemize}

First condition is satisfied since $V_1 + V_4 \geq 0$, which implies $V \geq 0$. One can see the second condition holds by noticing that $V(p, w) = 0$ iff $p = p^*$ and $\|Lw - p + u\|_2 = 0$.

Next, we show that $V(p, w)$ satisfies (2.7c). Recall that the algorithm (5.7) has an approximation to the gradient in the form $\nabla_p f(p) + e$, so by including this error term in (5.7b) we write
\[
p^+ = p + \alpha \phi,
\]
where $\phi = -L \xi - Le + Lw - p + u$, and $\xi \in \partial \hat{f}(p)$. Since $\Sigma$ does not allow to have $p_i = \bar{p}_i$ or $p_i = \underline{p}_i$ for all $i \in V$, i.e., $p$ is never a point where $f$ is nonsmooth, then without loss of generality we assume $\xi = \nabla_p \hat{f}(p)$ in the following analysis. Also notice that the Hessian matrix $\nabla^2 \hat{f}(p) = \nabla^2 f(p)$ for all $p \in \mathbb{R}^N \setminus \{\bar{p}, \underline{p}\}$.

**One-step differences of $V$:**

The following analysis shows one-step differences of the summands of $V$, starting with $V_1$. Using the second-order Taylor expansion, we have
\[
\hat{f}(p^+) = \hat{f}(p) + \alpha \xi^T \phi + \frac{1}{2} \alpha^2 \phi^T \nabla^2 f(p) \phi + o(\alpha^2 \|\phi\|_2^2).
\]

It follows that
\[
\Delta V_1 = \hat{f}(p^+) - \hat{f}(p)
= \alpha \xi^T (-L \xi - Le + Lw - p + u)
+ \frac{1}{2} \alpha^2 \phi^T \nabla^2 f(p) \phi + o(\alpha^2 \|\phi\|_2^2).
\]
Continuing from the last equality, and using that \( \max\{\xi\} \leq M + \chi \)

\[
\Delta V_1 \leq -\alpha \xi^T L \xi + \alpha (M + \chi) \|L \xi\|_2 + \alpha (M + \chi) \|L w - p + u\|_2 + \frac{1}{2} \alpha^2 \phi^T \nabla^2 f(p) \phi \\
+ o(\alpha^2 \|\phi\|^2_2).
\]

Including now the expression for \( \phi = -L \xi - L e + L w - p + u \) in the quadratic Hessian term, and using the fact that \( x^T \nabla^2 f(p') x + y^T \nabla^2 f(p') y \geq x^T \nabla^2 f(p') y + y^T \nabla^2 f(p') x \)

for any \( x, y, p' \)

\[
\Delta V_1 \leq -\alpha \xi^T L \xi + \alpha (M + \chi) \|L \xi\|_2 \\
+ \alpha (M + \chi) \|L w - p + u\|_2 \\
+ \alpha^2 \xi^T L^T \nabla^2 f(p) L \xi \\
+ \alpha^2 e^T L^T \nabla^2 f(p) L e \\
+ \alpha^2 (L w - p + u)^T \nabla^2 f(p) (L w - p + u) + o(\alpha^2 \|\phi\|^2_2).
\]

In the following, we apply that, for any vector \( x \in \mathbb{R}^N \), and matrix \( A \in \mathbb{R}^{N \times N} \), we have \( \|A\|_2 = \sigma_1(A) \) and \( \|Ax\|_2^2 \leq \sigma_1(A)^2 \|x\|_2^2 \) (e.g., see [35]) together with the fact that \( \nabla^2 f(p) \leq \Gamma I \). Moreover, since \( G \) is weight-balanced and strongly connected, then

\( \xi^T L \xi \geq c_1 \|\hat{\xi}\|^2_2 \), for all \( \xi \in \mathbb{R}^N \), where \( \hat{\xi} = \xi - \text{Avg}(\xi) \mathbf{1}_N \) and \( c_1 = \lambda_2(L + L^T) \). Then,

\[
\Delta V_1 \leq -c_1 \alpha \|\xi - \text{Avg}(\xi) \mathbf{1}_N\|^2_2 \\
+ \alpha (M + \chi) \sigma_1(L) \|e\|_2 + \alpha (M + \chi) \|L w - p + u\|_2 \\
+ \alpha^2 \Gamma \sigma_1(L) (M + \chi)^2 \\
+ \alpha^2 \hat{\xi}^T L^T \nabla^2 f(p) L \xi \\
+ \alpha^2 e^T L^T \nabla^2 f(p) L e \\
+ \alpha^2 (L w - p + u)^T \nabla^2 f(p) (L w - p + u) + o(\alpha^2 \|\phi\|^2_2).
\]

Using the assumption \( \|e\|_2 \leq \alpha K \) and the fact that \( \nabla^2 f(p) \leq \Gamma I \) for some constant \( \Gamma \), we have

\[
\Delta V_1 \leq -c_1 \alpha \|\xi - \text{Avg}(\xi) \mathbf{1}_N\|^2_2 \\
+ \alpha^2 (M + \chi) \sigma_1(L) K + \alpha (M + \chi) \|L w - p + u\|_2 \\
+ \alpha^2 (M + \chi)^2 \sigma_1(L) K^2 \\
+ \alpha^2 \Gamma \sigma_1(L) K^2 \\
+ \alpha^2 K \|L w - p + u\|_2^2 + o(\alpha^2 \|\phi\|^2_2).
\]
Next, for $V_2(p, w) = (M + \chi + \ell)\|Lw - p + u\|_2 + \Gamma\|Lw - p + u\|_2^2$ and $\Delta V_2$, we have

$$
\Delta V_2 = (M + \chi + \ell)(\|Lw^+ - p^+ + u\|_2 - \|Lw - p + u\|_2) \\
+ \Gamma(\|Lw^+ - p^+ + u\|_2^2 - \|Lw - p + u\|_2^2) \\
= (M + \chi + \ell)((1 - \alpha)\|Lw - p + u\|_2 - \|Lw - p + u\|_2) \\
+ \Gamma((1 - \alpha)^2\|Lw - p + u\|_2^2 - \|Lw - p + u\|_2^2) \\
= -\alpha(M + \chi + \ell)\|Lw - p + u\|_2 + \Gamma((1 - \alpha)^2 - 1)\|Lw - p + u\|_2^2.
$$

In the following, to analyze $V_3$ and $V_4$ we assume $\bar{\epsilon} = 0$. We will show in the next steps that the solution $1_N^T p(k)$ for $k \in \mathbb{Z}_{\geq 0}$ goes exponentially to $1_N^T p^*$, which implies that it has the $\gamma$-ISS property. Therefore, when the perturbation $1_N^T \bar{\epsilon}$ is present, we already know that $\Delta V_3$ and $\Delta V_4$ are decreasing outside of an open ball of radius $1_N^T \bar{\epsilon}$, which can be made arbitrarily small by reducing $\epsilon$. For $V_3(p, w) = c_1 m(1_N^T(p - p^*))^2$ and $\Delta V_3$, we have

$$
\Delta V_3 = c_1 m((1_N^T(p^+ - p^*))^2 - (1_N^T(p - p^*))^2) \\
= c_1 m((1_N^T(p + \alpha(-L(\xi + e) + Lw - p + u) - p^*))^2 - (1_N^T(p - p^*))^2).
$$

Recall that $1_N^T L(\xi + e) = 0$ and $1_N^T Lw = 0$ since the $\mathcal{G}$ is weight-balanced. Using last two facts one has

$$
\Delta V_3 = c_1 m((1_N^T(p - p^* + \alpha(-p + u))|^2 - (1_N^T(p - p^*))^2) \\
= c_1 m((1_N^T(p - p^*))^2 + 2\alpha(1_N^T(p - p^*))(1_N^T(-p + u)) + \alpha^2(1_N^T(-p + u))^2 - (1_N^T(p - p^*))^2) \\
= c_1 m(2\alpha(1_N^T(p - p^*))^2 + \alpha^2(1_N^T(-p + u))^2) \\
= c_1 m(-\alpha(1_N^T(p - p^*))^2 - \alpha(1 - \alpha)(1_N^T(p - p^*))^2), \\
= -\alpha(2 - \alpha)V_3,
$$

where we have used the fact that $1_N^T p^* = 1_N^T u$, in obtaining the third and fourth equalities.

Finally, for $V_4(p, w) = m'1_N^T(p - p^*)$ and $\Delta V_4$ one has

$$
\Delta V_4 = m'(1_N^T(p^+ - p^*) - 1_N^T(p - p^*)) \\
= m'(1_N^T p + \alpha 1_N^T(-p + u) - 1_N^T p) \\
= -m'\alpha 1_N^T(p - u) = -m'\alpha V_3 \leq 0.
$$

Notice that $\Delta V_3 \leq 0$ and $\Delta V_4 \leq 0$ if $\alpha \in (0, 1)$. 

Define \( c_2 = \sigma_1^2(L)\Gamma(M + \chi)^2 \). Now we put all inequalities together to obtain an upper bound on \( \Delta V \)

\[
\Delta V \leq -\alpha c_1 \| \xi - \text{Avg}(\xi) \|_2^2 + \alpha^2(M + \chi)\sigma_1(L)K + \alpha^4 \Gamma \sigma_1^2(L)K^2 \\
+ \alpha^2 c_2 + o(\alpha^2 \| \phi \|_2^2) \\
- \alpha \ell \| Lw - p + u \|_2^2 - 2\alpha \Gamma(1 - \alpha) \| Lw - p + u \|_2^2 \\
- c_1 m\alpha(1 - \alpha)(1^T_N(p - p^*)^2 - c_4 m\alpha(1 - \alpha)(1^T_N(p - p^*))^2 \\
- m' \alpha V_4 \]

\[
= -\alpha c_1 \| \xi - \text{Avg}(\xi) \|_2^2 + m(1^T_N(p - p^*)^2) \\
+ \alpha^2(M + \chi)\sigma_1(L)K + \alpha^4 \Gamma \sigma_1^2(L)K^2 \\
+ \alpha^2 c_2 + o(\alpha^2 \| \phi \|_2^2) - \alpha \ell \| Lw - p + u \|_2^2 \\
- 2\alpha \Gamma(1 - \alpha) \| Lw - p + u \|_2^2 - c_4 m\alpha(1 - \alpha)(1^T_N(p - p^*))^2 \\
- m' \alpha V_4 .
\]

Using Lemma 5.8 together with the facts that \( F_\alpha^{u,\epsilon} \) is positively invariant and that \( \Delta V_3 \) and \( \Delta V_4 \) are decreasing outside of an open ball of radius \( 1^T_N\bar{\epsilon} \) which can be made arbitrarily small by reducing \( \epsilon \), we have

\[
\Delta V \leq -\alpha c_1 (1 - \theta)\gamma \| p - p^* \|_2^2 - \alpha c_4 \theta \gamma \| p - p^* \|_2^2 \\
+ \alpha^2 c_4 + \alpha^2 c_2 + \alpha^4 c_3 \\
+ o(\alpha^2 \| \phi \|_2^2) - \alpha \ell \| Lw - p + u \|_2^2 \\
- 2\alpha \Gamma(1 - \alpha) \| Lw - p + u \|_2^2 - c_4 m\alpha(1 - \alpha)(1^T_N(p - p^*))^2 \\
- m' \alpha V_4 ,
\]

where \( \theta \in (0, 1) \), \( c_3 = \Gamma \sigma_1^2(L)K^2 \), and \( c_4 = (M + \chi)\sigma_1(L)K \). In order to have \( \Delta V \leq 0 \), it must be \( \alpha \in (0, 1) \).

In order to have \( \Delta V \leq 0 \), it must be \( \alpha \in (0, 1) \). Moreover, in order to guarantee \( \Delta V \leq 0 \), consider first

\[
\| p - p^* \|_2^2 \geq \frac{\alpha c_4 + \alpha^3 c_3 + \alpha c_2}{c_1 \gamma \theta} \\
\frac{o(\alpha^2 \| L\xi \|_2^2) + o(\alpha^2 \| Lw - p + u \|_2^2)}{\alpha c_1 \gamma \theta},
\]
where $\xi \in \partial \tilde{f}(p)$. Recall that $\bar{u} \in \mathbb{R}^N$ is a given constant, $u = \bar{u} + \bar{\epsilon}$, and $\epsilon \leq \alpha$.

Notice that $o(||Lw - p + u||_2^2)$ converges exponentially fast to the segment $[-|1_N^T \bar{\epsilon}|, |1_N^T \bar{\epsilon}|] \subset [-N\epsilon, N\epsilon] \subset [-N\alpha, N\alpha]$ which can be made arbitrarily small by reducing $\alpha$. This fact can be shown if we define $W(p, w) = ||Lw - p + \bar{u}||_2^2$ and assume $\bar{\epsilon} = 0$. It follows that $\Delta W \leq ((1 - \alpha)^2 - 1)W$, which implies that $||Lw - p + \bar{u}||_2^2$ goes to zero exponentially and that has ISS property.

Also, notice that $o(||L\xi||_2^2) \leq l(M + \chi + \alpha K)^2$ for some $l \in \mathbb{R}_{>0}$. We define

$$\rho(\alpha) \triangleq \frac{\alpha c_4 + \alpha^3 c_3 + \alpha c_2}{c_1 \gamma \theta} + \frac{o(\alpha^2 ||L\xi||_2^2)}{\alpha c_1 \gamma \theta}.$$ 

Thus, $\Delta V \leq 0$ if $||p - p_i||_2^2 \geq \rho(\alpha)$, where $\rho$ is a class-$\mathcal{K}$.

Next, we show that $V(p, w)$ is radially unbounded over $w \in \mathcal{F}_b$, where $\mathcal{F}_b \triangleq \{w \in \mathbb{R}^N \mid 1_N^T w = b\}$ and $b$ is any given constant. The following analysis is analogous to the one in Theorem 5.1. Although $f(p)$ is not supposed to be radially unbounded with respect to $p$ as it is assumed in Theorem 5.1, the function $J(p)$ is radially unbounded. To analyze $\lim V(p, w)$ when $||(p, w)||_2 \to +\infty$ consider two situations: (i) when there exists $p_i \to \pm \infty$ $i \in \mathcal{V}$ and (ii) when there exists $w_i \to \pm \infty$ for some $i \in \mathcal{V}$ and $p$ is bounded.

If (i) holds, the result follows from $f(p) + J(p)$ being radially unbounded with respect to $p$. Under (ii), if there exist $w_i \to +\infty$, then there must exist $j \neq i$ such that $w_j \to -\infty$, and vice versa with $w \in \mathcal{F}_b$. This implies that $w \neq d 1_N$, for some $d \in \mathbb{R}$ as $||w||_2 \to +\infty$ and thus $||Lw - p + u||_2 \to +\infty$ when there exists $i$ such that $w_i \to \pm \infty$. It follows that $V$ is radially unbounded over $\mathcal{F}_b$. Notice that $\mathcal{F}_b$ is strongly positively invariant under the **box-coupled fairness algorithm** by taking $b = 1_N^T w(0)$. 
Since $V$ is radially unbounded and $\Delta V \leq 0$ if $\|p - p^*\|^2 \geq \rho(\alpha)$, then the sublevel sets of $V$ are compact and positively invariant. Since all assumptions in Theorem 2.1 are satisfied, then as $t \to +\infty$, solutions tend to a ball close to the equilibrium with radius depending on $\alpha$. \hfill \Box

**Remark 5.5.** Our motivation to include the error term $e$ in Theorem 5.2 is that in many applications the gradient of the payoff function only can be approximated. For example, in the following section, we show an application to the virus spread minimization, where the gradient is approximated by the well-known Power Iteration.

### 5.3 Application to virus spread minimization

As an application, we employ the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm on a virus spread minimization problem for computer and epidemic networks. Notice that if we remove the box constraints, we can similarly use the ROBUST GRADIENT FAIRNESS algorithm to solve the same problem. To dynamically solve the $\delta$-VIRUS MITIGATION and the $\kappa$-VIRUS MITIGATION problems using local information we use the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm together with the exact penalty method as shown in Section 5.1. Since our distributed algorithm approach requires the computation of the gradient, and the fact that the gradient for the $\delta$-VIRUS MITIGATION and $\kappa$-VIRUS MITIGATION problems is not naturally distributed, we approximate it by using the well-known Power Iteration.
5.3.1 Problem statement and solution approach

The SIS (susceptible-infected-susceptible) model for virus dynamics proposed in [30] is given by

\[ q_i(t + 1) = (1 - \prod_{j=1}^{N} (1 - a_{ij}q_j(t))), \quad (5.15) \]

where \( q_i(t) \in \mathbb{R} \) is the probability that node \( i \) is infected at time \( t \), \( i \in \{1, \ldots, N\} \) and \( a_{ji} \) is defined as

\[ a_{ij} = \begin{cases} \kappa_i \beta_{ij}, & \text{for } j \neq i, \\ 1 - c_i \delta_i, & \text{for } j = i. \end{cases} \]

Here, \( \kappa_i \in (0, 1] \) represents the scaling factor of the nominal weight \( \beta_{ij} \), \( \beta_{ij} \in [0, 1] \) is the probability that the virus from node \( j \) infects node \( i \), or in other words, it represents the isolation capability or firewall placed in the entering branches, \( c_i \in [0, 1] \) represents the district-specific scaling factor, and \( \delta_i \in [0, 1] \) is the probability of an infected node \( i \) to be recovered. Using the Weierstrass product inequality, valid for \( a_{ij}q_j(t) \in [0, 1] \), we obtain the following upper bound

\[ q_i(t + 1) \leq \sum_{j=1}^{N} a_{ij}q_j(t), \quad \forall i \in \{1, \ldots, N\}. \]
The previous inequality reads in vector notation as

\[ q(t + 1) \leq A(\delta, \kappa)q(t), \quad (5.16) \]

where \( q(t) = [q_1(t), \ldots, q_N(t)]^\top \) and \( A(\delta, \kappa) \in \mathbb{R}^{N \times N} \) is defined as

\[
A(\delta, \kappa) = \begin{bmatrix}
1 - c_1\delta_1 & \kappa_1\beta_{12} & \ldots & \kappa_1\beta_{1N} \\
\kappa_2\beta_{21} & 1 - c_2\delta_2 & \ldots & \kappa_2\beta_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\kappa_N\beta_{N1} & \kappa_N\beta_{N2} & \ldots & 1 - c_N\delta_N
\end{bmatrix}
= I_N - D + KG.
\]

Here, \( A(\delta, \kappa), D = \text{diag}(c)\text{diag}(\delta), K = \text{diag}(\kappa), \text{and } G = A(0, 1_N) - I_N \). Let \( \mathcal{G}(A(0, 1_N)) = \mathcal{G}(G) \) be the graph associated to the virus dynamics contact network. We define the \textit{topology matrix} of the network as the matrix \( G \). When there is no confusion, we will denote \( \mathcal{G}(G) \) by \( \mathcal{G} \). Next proposition shows that the dominant eigenvalue, \( \lambda_1(A(\delta, \kappa)) \), governs the growth/decay rate of infection.

**Proposition 5.1 ([30]).** \textit{An epidemic described by (5.15) becomes extinct if and only if }\( \rho(A(\delta, \kappa)) < 1 \). Moreover, when an epidemic is diminishing, the probability of infection \textit{decays at least exponentially over time.}
Inspired by [48, 17, 38, 14], we consider the following two problems to minimize the effects of virus contagion. The $\delta$-Virus mitigation problem is defined by

$$
\min_{\delta} \quad \lambda_1(A(\delta, \kappa))
$$

subject to

$$
1^T_N \delta = 1^T_N \bar{u}_\delta,
$$

$$
\delta \in [\delta, \bar{\delta}]^N,
$$

(5.18)

where $\kappa$ is fixed, $1^T_N \bar{u}_\delta$ is the total amount of antivirus available, and the constants $\delta_i, \bar{\delta}_i \in [0, 1]$. The $\kappa$-Virus mitigation problem is given

$$
\min_{\kappa} \quad \lambda_1(A(\delta, \kappa))
$$

subject to

$$
1^T_N \kappa = 1^T_N \bar{u}_\kappa,
$$

$$
\kappa \in [\kappa, \bar{\kappa}]^N,
$$

(5.19)

where $\delta$ is fixed, $1^T_N \bar{u}_\kappa$ is the total amount of isolation resources, and the constants $\kappa_i, \bar{\kappa}_i \in (0, 1]$.

**Remark 5.6.** We refer to $\delta^*$ and $\kappa^*$ to the solutions of Problem (5.18) or Problem (5.19), respectively. These solutions minimize the exponential decay/growth rate of (5.15) subject to resource constraints. Moreover, if the solutions $\delta^*$ or $\kappa^*$ make the dominant eigenvalue $\lambda_1(A(\delta^*, \kappa))$ or $\lambda_1(A(\delta, \kappa^*))$ (for $\kappa$ or $\delta$ depending on the problem) strictly less than one, then it is guaranteed that the disease free equilibrium, $q^* = 0$, is globally exponentially stable.
To analyze the solution of the $\delta$-virus mitigation and $\kappa$-virus mitigation problems, we show in the following two lemmas that $\lambda_1(A(\delta, \kappa))$ is a convex function of $\delta$, and $\kappa$, respectively. Finally, in Lemma 5.13 we explicitly calculate the gradient of $\lambda_1$ in order to use the $p$-robust box-gradient fairness algorithm using the exact penalty method presented in Section 5.1.

**Lemma 5.11** ([34]). Let $B$ be nonnegative, and $C = \text{diag}(\delta_1, \ldots, \delta_N)$. Then, the maximum eigenvalue of $B + C$, $\lambda_1(B + C)$, is a convex function of $C$.

**Lemma 5.12.** ([50] Convexity of $\lambda_1(KG)$): Let $G$ be positive semidefinite, and $K = \text{diag}(\kappa_1, \ldots, \kappa_N)$. Assume that $\kappa_i > 0$ for all $i \in \{1, \ldots, N\}$. Then, $\lambda_1$ of $KG$ is a convex function of $K$.

**Remark 5.7.** Notice that in general trace($G$) = 0 for our virus application, so, in particular, $G$ cannot be positive semidefinite. This fact makes the $\kappa$-virus mitigation problem to be non-convex, even when $G$ is symmetric. We refer to [50] for further discussion about the convexity of $\lambda_1(KG)$. To approximate the solution of the $\kappa$-virus mitigation problem, we can use the inequality $\lambda_1(KG) \leq (\lambda_1(K^2G^2))^{\frac{1}{2}}$ [51]. When $G$ is symmetric, $G^2$ is positive semidefinite, then $\lambda_1(K^2G^2)$ is convex by Lemma 5.12. Therefore, we can obtain an upperbound of the solution of the $\kappa$-virus mitigation problem. The analysis of the error by using this approach is out of scope of this paper.

Lemma 5.11 and Lemma 5.12 show that $\lambda_1(A(\delta, \kappa))$ is a convex function with respect to its arguments. Then we can aim to apply the $p$-robust box-gradient fairness algorithm on the resource allocation problem associated to this function. As shown in
Lemma 5.1, information of the gradient of matrix $A(\delta, \kappa)$ is required in order to evaluate if a solution is optimal and to implement our algorithm. For that reason, in the following lemma we provide the analysis to obtain such a gradient.

**Lemma 5.13.** Let $v,$ and $s$ be the left and right eigenvectors of the matrix $A(\delta, \kappa)$ as defined in (5.17). Then

\[
\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \delta_i} = -c_i \frac{v_i s_i}{v^T s}
\] (5.20)

and,

\[
\frac{\partial \lambda_1(A(\delta, \kappa))}{\partial \kappa_i} = \frac{v_i}{v^T s} \sum_{j \neq i} \beta_{ij} s_j
\] (5.21)

**Proof.** See Lemma 4.11. \qed

### 5.3.2 The Power Iteration

The power method is a well-known algorithm for approximating $\lambda_1(A)$ for $A \in \mathbb{R}^{N \times N}$. For a detailed description of this method the reader may consult [57, 58]. In this paper, we restrict our discussion for $A$ being primitive with $z(0) > 0$. In Remark 5.8, we explain how to relax the condition of $A$ being primitive to have $A$ Metzler and irreducible. The algorithm is given by

\[
z(t + 1) = \frac{Az(t)}{\|Az(t)\|_\infty},
\] (5.22)
where \( z(t) \in (0, 1] \) for \( t \in \mathbb{N} \). Under the assumptions listed above on \( A \) it can be proven that \( z(t) \to x \), as \( t \to +\infty \), where \( x \) is the right eigenvector associated to \( \lambda_1(A) \).

**Remark 5.8.** In general, the condition \( A \) to be primitive can be relaxed to have \( A \) non-negative and irreducible. For that, a shifted version of \( A \), \( A_c \triangleq A + cI \), where \( c > 0 \), can be used.

**Remark 5.9.** When \( A \) has negative elements on its diagonal, the algorithm can be used with a shifted version \( A_c \triangleq A - cI \), where \( c = \min_{i \in V} a_{ii} \).

### 5.3.3 Algorithm Implementation and Simulations

Next we show an example that illustrates the response of the \( p \)-robust box-gradient fairness algorithm to solve a particular \( \delta \)-virus mitigation problem. Before that, we summarize the required assumptions to solve the \( \delta \)-virus mitigation and the \( \kappa \)-virus mitigation problems by using Theorem 5.2. First, Lemma 5.12 show that \( \lambda_1(A) \) is a convex function respect \( \delta \) and \( \kappa \). For the case of \( \lambda_1(A) \), Example 5.1 shows that for a particular problem, the algorithm converges to the desired solution. Second, we require to have \( \lambda_1(A) \) bounded below and \( A(\delta, \kappa) \) has to be irreducible all time. These assumptions are satisfied since the set \( F^V_{\text{box}} \) is invariant to our dynamics as shown in Lemma 5.10. Finally, the computation of the gradient of \( \lambda_1(A) \) is required by using local information. To address this, we use the Power Iteration method, summarized in Section 5.3.2, to approximate the gradient as shown in Lemma (5.13). At each time step, the Power Iteration algorithm runs until a desired stopping condition is reached.
The reader may consult [59] for a discussion on the distributed stopping criterion.

**Example 5.1** (Optimizing in $\delta$). We illustrate the response of the $p$-robust box-gradient fairness algorithm for the directed topology matrix $G$ associated to $A(\delta, \kappa)$ for fix $\kappa$. We construct $G$ as a ring with $V \in \{1, \ldots, 10\}$, bidirectional edges given by $(i, i + 1) = 1/10$ for $i \in V$ (assume that if $i = 10$, then $i + 1 = 1$) and additional bidirectional edges given by $(1, 5) = (3, 9) = 1/8$. Moreover, we place directional edges $(i, i + 2) = 1/7$ for $i = \{1, \ldots, 8\}$ and $(i + 8, i) = 1/7$ for $i = \{1, 2\}$. We use $u_1 = 5.5$, $u_j = 0$ for $j \in V \setminus \{1\}$, $c_1 = c_3 = c_6 = 0.85$, $c_j = 1$ for $j \in V \setminus \{1, 3, 6\}$, $\overline{\delta} = .91N$, $\overline{\delta} = .21N$, $\chi = 10$, and $\alpha = 0.01$. In this example, we approximate the gradient of $\lambda_1(A)$ by the Power Iteration.

At each iteration of the $p$-robust box-gradient fairness algorithm, we run one iteration of the Power Iteration. In Figure 5.1, we show the behavior of the $p$-robust box-gradient fairness algorithm for a random initial condition with $\delta(0) \in [.2, .9]^N$, $w(0) \in [0, 1]^N$.

The optimal value is given by $\lambda_1(A(\delta^*)) = 0.8714$. We introduce an erroneous update on the system state at time $t = 1500$, where we force $p(1500)$ and $w(1500)$ to a random vector in $[0, 1]^N$. After $t = 1500$, the algorithm converges again to the optimal point.

Notice that $p(t) \in [0, 1]^N$ for $t \geq 0$ since $F_{box}^\nu$ is invariant to our dynamics.

### 5.4 Summary

We have considered a class of distributed resource allocation problems. We have proposed two novel discrete-time algorithms that converge in a practical way to the solution as long as the chosen stepsize is sufficiently small. In particular, the proposed
Figure 5.1: Trajectories of $\delta(t)$ and $\lambda_1(A(\delta(t)))$ of Example 5.1 for the $p$-robust box-gradient fairness algorithm. It is used an erroneous update on the system state at $t = 1500$.

algorithms are designed to be robust to temporary errors in communication or computations of agents. Our technical approach relies on results from algebraic graph theory, second-order convex analysis as well as nonsmooth partial stability. Simulations show that the algorithms converge for a wider set of problems. Motivated by applications to virus processes, we plan to extend available proofs that can help us relax the assumptions needed.

Publications associated with this chapter

This paper contains material that has been published in the following works:

Optimal Control Applications and Methods.

Chapter 6

Distributed stopping criteria for the

Power Iteration method

The well-known Power Iteration method is a simple algorithm for computing the maximum eigenvalue of a matrix and its corresponding eigenvector. When the matrix is Metzler and irreducible, it is known that this method is convergent under mild assumptions. Our particular interest in this approach is motivated by the design of a fully distributed algorithm that can be used in the minimization of the spectral radius of a nonnegative matrix under equality and inequality constraints. A specific real-world problem leading to such a setting arises in the mitigation of viruses over the Internet or the general population. We propose a distributed stopping criterion for the Power Iteration method. This algorithm is then used together with the $p$-robust box-gradient fairness algorithm proposed in [60] to mitigate virus spread over a network in a fully distributed manner.
Literature review. There are two main approaches in the literature to the virus mitigation problem. One consists of detecting and isolating infected nodes by topology adaptation or quarantine, while the other one aims to allocate antivirus on infected nodes. Centralized algorithms that fall under either one of the above categories have been proposed by many authors. For example, the authors in [13] propose a convex optimization framework to find cost-optimal solutions to control an epidemic outbreak by regulating the traffic between subpopulations. Another example is given in [17], where the authors propose a geometric programming framework to find the optimal allocation of resources under local constraints. However, centralized approaches are not implementable when networks are operated by multiple operators who need to preserve their local privacy and interest. These scenarios can be handled by means of local and anonymous interactions to stop the virus spread. Following this path, [14] and [61] propose decentralized controllers that mitigate virus epidemics. The work [14] proposes a decentralized algorithm based on the use of control matrices that have a diagonal structure, and thus, which are naturally distributed. However, the computation of these diagonal matrices is not distributed itself and can not be learned by operators. Along these lines, the authors in [61] propose a sparse allocation of limited resources on a subset of network nodes so that the dominant eigenvalue of a linear dynamical process associated with the network is minimized. More recently, [16] proposes a distributed resource allocation strategy to control a virus outbreak in a network by building on the framework of [17]. The proposed algorithm is based on Distributed Alternating Direction Method of Multipliers (D-ADMM) algorithm, however, the communication cost of such algorithm is
expensive as every operator in the network needs to interchange a local estimation of the states of the entire network nodes.

In our recent work [60], we propose the distributed $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm that allows an interconnected group of agents to collectively minimize a global cost function subject to equality and inequality constraints. It requires information of the gradient of the cost function, which in the particular problem of virus mitigation is not separable. Thus, we approximate the gradient of the cost function by the well-known Power Iteration method. The Power Iteration method was proposed originally in [62]. Its convergence and rate of convergence has been analyzed by many authors, see for instance [57, 58, 63]. Of particular interest in this paper is the criterion given in [58], where the authors provide a bound on the accuracy of the approximations for the maximum eigenvalue of a matrix and its corresponding eigenvector. However, to the best of the authors knowledge, a distributed stopping criterion with similar estimates of accuracy is not available.

Statement of contributions. We propose a distributed stopping criterion for the well-known Power Iteration method for symmetric and Metzler matrices. We provide a bound on the accuracy of the approximations for the maximum eigenvalue of the matrix and its corresponding eigenvector. This result is applied to mitigate virus spread over a network. For that, we interconnect the Power Iteration algorithm together with our recently developed $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm [60]. This distributed algorithm allows an interconnected group of agents to collectively minimize a global cost function subject to equality and inequality constraints. The Power Iteration method
and the distributed stopping criterion provides an approximation of the cost function’s
gradient for each iteration. We show that the interconnection between the two methods
is convergent and preserves the convergence properties of the \textit{p-robust box-gradient fairness} algorithm. Finally, we illustrate the applicability of our algorithm by a virus
spread allocation problem over computer dataset taken from the e-mail network of Enron Corporation.

6.1 The Power Iteration method

The Power Iteration is a well-known algorithm for approximating $\lambda_1(A)$ for $A \in \mathbb{R}^{N \times N}$. For a detailed description of this method the reader may consult [57, 58]. In this paper we restrict our discussion for $A$ being primitive with $z(0) > 0$. In Remark 6.1 we explain how to relax the condition of primitivity to Mezler and irreducible matrices. The algorithm is given by

$$z(t + 1) = \frac{A z(t)}{\|A z(t)\|_\infty}, \quad (6.1)$$

where $z(t) \in (0, 1]$ for $t \in \mathbb{N}$. Under the assumptions listed above on $A$, $z(t) \to x$, as $t \to +\infty$, where $x$ is the right eigenvector associated to $\lambda_1(A)$, see Lemma 6.4.

\textbf{Remark 6.1.} \textit{The condition $A$ to be primitive can be relaxed to have $A$ nonnegative and irreducible. For that, a shifted version of $A$, $A_c \doteq A + cI$, where $c > 0$, can be used. Moreover, when $A$ has negative elements on its diagonal, a shifted version of $A$, $A_c \doteq A + bI$, where $b > \min_{i \in V} a_{ii}$, can be used.}
6.1.1 Centralized stopping criterion

While running the Power Iteration one can compute \(\|r(t)\|_2 \leq \eta |\lambda(t)|\), where \(r(t) \triangleq \lambda(t)z(t) - Az(t)\), \(\lambda(t) \triangleq \frac{z(t)^\top Az(t)}{z(t)^\top z(t)}\), and \(\eta \in [0, 1]\). For a given \(\eta\), Theorem 6.1 shows how these can be used to provide a bound on the accuracy of the approximations of eigenvalues and eigenvectors. Leading to this result, we state two supporting lemmas.

**Lemma 6.1.** (Perturbed eigenvalues): Let \(A, B \in \mathbb{R}^{N \times N}\) be nonnegative and symmetric matrices. Then,
\[
|\lambda_1(A) - \lambda_1(B)| \leq \|A - B\|_2.
\]

*Proof.* Without loss of generality, assume \(\lambda_1(B) \geq \lambda_1(A)\). Notice that \(|\lambda_1(A) - \lambda_1(B)| = \min\{|\lambda' - \lambda_1(B)| \mid \lambda' \in \text{spec}(A)\}\) (in the case that \(\lambda_1(B) \leq \lambda_1(A)\), we have \(|\lambda_1(A) - \lambda_1(B)| = \min\{|\lambda' - \lambda_1(A)| \mid \lambda' \in \text{spec}(B)\}\)). From here, the proof follows the same steps as those in the proof of Proposition 3.11 in [58]. \(\square\)

**Lemma 6.2.** (Perturbed eigenvectors): Let \(A, B \in \mathbb{R}^{N \times N}\) be nonnegative and symmetric matrices. Let \(y\) be the eigenvector of \(B\) for the eigenvalue \(\lambda_1(B)\) and \(x\) be the eigenvector of \(A\) corresponding to \(\lambda_1(A)\). Then,
\[
\|x - y\|_2 \leq 2 \frac{\|A - B\|_2}{\lambda_1(A) - \lambda_2(A)} \|y\|_2.
\]

*Proof.* By Lemma 6.1, we have that \(|\lambda_1(A) - \lambda_1(B)| \leq \|A - B\|_2\). From here, the proof follows similar steps as those in the proof of Proposition 3.13 in [58]. \(\square\)

**Theorem 6.1.** Let \(A \in \mathbb{R}^{N \times N}\) be nonnegative, irreducible, and symmetric matrix. Let
\[ \eta \in [0, 1), \ |r(t)|_2 \leq \eta |\lambda(t)|, \text{ and } x \text{ be the eigenvector corresponding to } \lambda_1(A). \text{ Then,} \]
\[
|\lambda_1(A) - \lambda(t)| \leq \eta \lambda(t) \leq \frac{\eta}{1 - \eta} \lambda_1(A),
\]
\[
\|x - z(t)\|_2 \leq 2\eta(1 + \frac{\eta}{1 - \eta}) \frac{\lambda_1(A)}{\lambda_1(A) - \lambda_2(A)}.
\]

**Proof.** The proof follows along the same steps as in the proof of Corollary 4.10 in [58]. Whenever the proof employs Proposition 3.11 and 3.13, replace it by Lemma 6.1 and Lemma 6.2, respectively. \[\square\]

### 6.1.2 Distributed stopping criterion

In order to implement the stopping criterion presented in Section 6.1, each agent needs to compute \[\frac{z(t)^	op A z(t)}{z(t)^	op z(t)}\], which is not a naturally distributed computation. For that reason, we propose next a distributed stopping criterion for the Power Iteration while providing a bound on the accuracy of the approximations of eigenvalues and eigenvectors. While running the Power Iteration method, one can evaluate \[\sqrt{\bar{N}}|r_1(t)|_\infty \leq \eta \left( \min_{i \in V} h_i(t) \right)\], where \(\bar{N}\) can be the size of the network \(N\) or any upper bound of it, \(r_1(t) \triangleq (\max_{i \in V} h_i(t))z(t) - Az(t), h_i(t) \triangleq \frac{1}{z_i(t)} \sum_{j=1}^{\bar{N}} a_{ij} z_j(t)\), and \(\eta \in [0, 1)\). We rewrite (6.1)
\[
z_i(t + 1) = \frac{h_i(t)z_i}{\|Az(t)\|_\infty}.
\] (6.2)

Notice that (6.1) and (6.2) are equivalent. We show in Theorem 6.2 the estimates for the accuracy of the approximations of the eigenvector associated to \(\lambda_1(A)\). Before stating it, we present two supporting lemmas, which at the same time provide an alternative analysis for the convergence of \(h_i(t)\) for \(i \in \{1, \ldots, N\}\) to \(\lambda_1(A)\).
Lemma 6.3. Consider the algorithm given in (6.2), then it holds that \( h_i^* = h_j^* \) for all \( i, j \in V \), at equilibrium.

Proof. It holds that \( z_i^* \max_{l \in V} h_l^* z_l^* = h_i^* z_i^* \) for all \( i \in V \) at equilibrium of (6.2). It follows that \( h_i^* = \max_{l \in V} h_l^* z_l^* \) for all \( i \in V \). □

Lemma 6.4. Let \( A \) be a primitive matrix. Consider the dynamics in (6.2) with \( z(0) > 0 \). Then \( h_i(z(t)) \) converges asymptotically to the Perron eigenvalue of \( A \) and \( z(t) \) converges to the Perron eigenvector of \( A \).

Proof. Consider the Lyapunov function candidate \( V_2 : \mathbb{R}^N \to \mathbb{R}_{\geq 0}, V_2(z) = V_3(z) + V_4(z) \), where \( V_3(z) = \max_{i \in V} h_i(z) \) and \( V_4(z) = -\min_{i \in V} h_i(z) \), which is continuous for \( z \in (0,1]^N \) and positive definite. We define the equilibrium point of (6.2) as \( \Omega = \{ z \in \mathbb{R}^N_{>0} \mid h_i(z) = h_j(z) \forall i, j \in V \} \). To verify that \( \Omega \) is a point, notice that \( \|z(t)\|_{\infty} = 1 \) for all \( t > 0 \). Let \( \Delta V_1 = \max_{i \in V} h_i(z(t+1)) - \max_{i \in V} h_i(z(t)) \), it follows that

\[
\Delta V_3 = \max_{i \in V} \left( \frac{\|Az\|_{\infty}}{h_i(z)z_i} \sum_{j=1}^N a_{ij} \frac{h_j(z)z_j}{\|Az\|_{\infty}} \right) - \max_{i \in V} h_i(z) \\
= \max_{i \in V} \left( \frac{1}{h_i(z)z_i} \sum_{j=1}^N a_{ij} h_j(z)z_j \right) - \max_{i \in V} h_i(z) \\
\leq \max_{i \in V} \left( \frac{1}{h_i(z)z_i} \sum_{j=1}^N a_{ij} z_j \max_{i \in V} h_i(z) \right) - \max_{i \in V} h_i(z) \\
= 0
\]

From (6.3) and using the fact that \( A \) is assumed primitive, we can deduce that \( V_3 = 0 \) iff \( h_i = h_j \) for all \( i, j \in V \). Following similar steps as before for \( \Delta V_3 \) it can be shown that \( \Delta V_4 \leq 0 \) and \( V_4 = 0 \) iff \( h_i = h_j \) for all \( i, j \in V \).

It follows that \( \Delta V_2(z) = 0 \) iff \( z \in \Omega \), and \( \Delta V_2(z) < 0 \) otherwise. Therefore, by Lyapunov theorem, \( \Omega \) is asymptotically stable. Global convergence follows from the
Proof. We first prove that which implies that \( A \) is primitive and \( z(0) > 0 \), it follows that \( z(0) \notin \rho(A) \), which implies that \( \frac{A_t z(0)}{\|A_t z(0)\|_{\infty}} > 0 \) for all \( t > 0 \). By Theorem 1.1 in [57] and the facts that \( A \) is primitive and \( z(0) > 0 \), we know that \( \lim_{t \to +\infty} \frac{A_t z(0)}{\|A_t z(0)\|_{\infty}} = x > 0 \). Therefore, we can always find \( \bar{a} \in (0, 1) \) such that \( z(t) \in [\bar{a}, 1] \). \( \square \)

**Theorem 6.2.** (Accuracy): Let same assumptions as in Theorem 6.1, hold and \( z(0) > 0 \). Let the stopping criterion be given by \( \sqrt{N} |r_1(t)|_{\infty} \leq \eta |\lambda(t)| \). Then,

\[
\|x - z(t)\|_2 \leq 2\eta(1 + \frac{\eta}{1 - \eta}) \frac{\lambda_1(A)}{\lambda_1(A) - \lambda_2(A)}
\]

Proof. We first prove that \( \|r_1(t)\|_{\infty} \geq \|r(t)\|_{\infty} \). Since \( \text{proj}_z(Az) = \frac{z(t)^T A_t z(t)}{z(t)^T z(t)} z(t) = \lambda(t) z(t) \), then \( \|\lambda(t) z(t) - Az(t)\|_{\infty} = \|r(t)\|_{\infty} \) gives the minimum distance (best approximation) from \( Az(t) \) to the line spanned by \( z(t) \). Hence, \( \|r_1(t)\|_{\infty} \geq \|r(t)\|_{\infty} \).

Next, we show that the stop condition \( \sqrt{N} |r_1(t)|_{\infty} \leq \eta \min_{i \in V} h_i(t) \) implies that \( \sqrt{N} \|r(t)\|_{\infty} \leq \eta \lambda(t) \) for some \( T > 0 \) such that \( t \geq T \). By Lemma 6.4 we have that \( \max_{i \in V} h_i(z(t)) \to \lambda_1(A) \) and \( z(t) \to x \), which implies \( \|r_1(t)\|_{\infty} \to 0 \), and thus, \( T \) is finite. Using the fact that \( \|r_1(t)\|_{\infty} \geq \|r(t)\|_{\infty} \) and given the assumption \( \sqrt{N} |r_1(t)|_{\infty} \leq \eta(\min_{i \in V} h_i(z(t))) \) for \( t \geq T \), it follows that \( \|r(t)\|_2 \leq \sqrt{N} |r_1(t)|_{\infty} \leq \sqrt{N} |r_1(t)|_{\infty} \leq \eta \min_{i \in V} h_i(z(t)) \) for \( t \geq T \). Next we prove that \( \min_{i \in V} h_i(z(t)) \leq \lambda(t) \) for all \( t > 0 \).

Let \( h = [h_1, \ldots, h_N]^T \), \( h_{\min} = \min_{i \in V} h_i \), \( z \in \mathbb{R}_{>0}^N \). Recall that \( h_i = \frac{1}{z_i}(Az)_i \), so that \( h_i z_i = (Az)_i \). In vector form we have \( \text{diag}(h)z = Az \). Pre-multiplying last expression by \( z^T \) we have that \( z^T \text{diag}(h)z = z^T Az \). Multiplying last expression by \( (z^T z)^{-1} \), we have that

\[
\frac{z^T \text{diag}(h)z}{z^T z} = \frac{z^T Az}{z^T z} \geq h_{\min}.
\]
Therefore, \( \|r(t)\|_2 \leq \sqrt{N} \|r(t)\|_\infty \leq \sqrt{N} \|r_1(t)\|_\infty \leq \eta \min_{i \in V} h_i(z(t)) \leq \eta \lambda(t) \) for \( t \geq T \) and the conclusion follows by applying Theorem 6.1.

6.2 Application to virus spread minimization

Inspired by [38, 14], we consider the following problem to minimize the effects of virus contagion. The \( \delta \)-VIRUS MITIGATION problem is defined by

\[
\min_{\delta \in [\underline{\delta}, \overline{\delta}]^N} \lambda_1(A(\delta))
\]

s.t. \( \mathbf{1}_N^T \delta = \mathbf{1}_N^T u_\delta \),

where \( \mathbf{1}_N^T u_\delta \) is the total amount of antivirus available and the constants \( \underline{\delta}, \overline{\delta} \in [0, 1] \).

In Chapter 4 it is shown that the \( \delta \)-VIRUS MITIGATION problem is convex. In addition, we use the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm proposed in Chapter 4 to solve the BOX-COUPLED FAIRNESS problem.

6.2.1 Feedback interconnection

To dynamically solve the \( \delta \)-VIRUS MITIGATION problem using local information, we feedback interconnect the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm together with the Power Iteration method. To be able to use the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm, the computation of the gradient of \( \lambda_1(A) \) is required. This is given in the following lemma.
**Lemma 6.5.** Assume that \( A(\delta) \) is symmetric as defined in (5.17). Let \( s \) be the right eigenvector of \( A(\delta) \). Then,
\[
\frac{\partial \lambda_1(A(\delta))}{\partial \delta_i} = -c_i \frac{s_i^2}{\|s\|_2^2}.
\] (6.5)

Now, the computation of the gradient in (6.5) requires the computation of the right eigenvector associated with \( \lambda_1(A) \). We approximate this, by means of the Power Iteration method and the stopping criterion provided in Theorem 6.2. Algorithm 3 describes the feedback interconnection between the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm and the Power Iteration method. Briefly, lines 7-11 iterate the Power Iteration method until the stopping condition is achieved. After that, \( \nabla_{\delta} \lambda_1(A(\delta(t))) \) is approximated and used in the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm.

**Algorithm 3** Feedback interconnection

1: Inputs \( \alpha, \bar{N} \)
2: Agent \( i \in \mathcal{V} \) initializes \( s_i, w_i, p_i \)
3: for \( t > 0 \) do
4: \( \delta_i(t) = p_i(t) \)
5: \( z_i = s_i \)
6: \( j = 1 \)
7: while \( \sqrt{\bar{N}}\|r_1(t)\|_\infty \leq \alpha \min_{i \in \mathcal{V}} h_i(t) \) do
8: \( h_i = \frac{1}{s_i} \sum_{j=1}^{N} a_{ij} s_j \)
9: \( z_i(j + 1) = \frac{h_i(s_i(j))}{\|A(\delta(t))\|_\infty} \)
10: \( j = j + 1 \)
11: end while
12: \( s_i(t + 1) = z_i(j) \)
13: Compute \( \nabla_{\delta} \lambda_1(A(\delta(t))) \) as in (6.5)
14: Compute \( \partial J(\delta) \)
15: Compute \( w_i(t + 1) \) and \( p_i(t + 1) \) as in (3)
16: \( \delta_i(t + 1) = p_i(t + 1) \)
17: \( t = t + 1 \)
18: end for

**Remark 6.2.** In the Algorithm 3, line 13, we can replace the 2-norm by \( \infty \)-norm. This is
explained by the fact that the gradient in (6.5) represents a direction. The same direction is achieved by using different norms in the denominator of (6.5).

Next theorem states the convergence result for Algorithm 3.

**Theorem 6.3.** (Feedback interconnection): Consider the $\delta$-VIRUS MITIGATION problem, where $A$ is assumed to be primitive and symmetric. Take an $\alpha$ sufficiently small for the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm. Then, the solutions of Algorithm 3 converge asymptotically to a ball centered at a solution of the $\delta$-VIRUS MITIGATION problem with a radius dependent on $\alpha$.

**Proof.** First we show that all assumptions given in Assumption 4.1 for Algorithm 3 are satisfied.

i) $\lambda_1(A)$ is convex and infinitely continuous differentiable.

ii) Given $\nabla_\delta \lambda_1(A)$ in (6.5), then $\|\nabla_\delta \lambda_1(\delta)\|_2 \leq 1$.

iii) There is a unique solution to the $\delta$-VIRUS MITIGATION problem. Provided that $A(\delta)$ is irreducible and symmetric, then $\lambda(\delta^*)$ is given by a unique point since the normalized right eigenvector associated to $\lambda_1(A)$ is unique.

iv) Next, we show that the condition to have an approximation of the gradient in the form $\|e(t)\|_2 \leq \alpha K$ for some $K \in \mathbb{R}_{>0}$ is satisfied. Notice that lines 7-11 iterate the Power Iteration method until the stopping condition is achieved, and thus the desired error is achieved, i.e., at every step of the $p$-ROBUST BOX-GRADIENT FAIRNESS algorithm, the Power Iteration method runs as many steps as necessary to get the bound in the
estimation
\[ \|s - s(t)\|_2 \leq \alpha \frac{\lambda_1(A(\delta(t)))}{\lambda_1(A(\delta(t))) - \lambda_2(A(\delta(t)))}. \]
Since \( A(\delta(t)) \) is irreducible for all \( t \), then \( \lambda_1(A(\delta(t))) - \lambda_2(A(\delta(t))) > \text{constant} > 0 \) for all \( t > 0 \). It follows that we can find a finite constant \( R > 0 \) such that \( \alpha \frac{\lambda_1(A(\delta(t)))}{\lambda_1(A(\delta(t))) - \lambda_2(A(\delta(t)))} \leq \alpha R \) for all \( t \).

Using (i)-(iv), and Theorem 5.2, we conclude that the solutions of the \( p \)-ROBUST BOX-GRADIENT FAIRNESS algorithm converge asymptotically to the solutions of the \( \delta \)-VIRUS MITIGATION problem.

6.2.2 Simulations

The following example is based on an e-mail communication network from the Enron corporation, which is constructed by taking the first \( N = 1000 \) nodes from the dataset available in [37]. We take the parameter \( \eta = 0.03 \) and fix the probabilities of virus transmission to be proportional to the in-degree of each node on the network. Figure 6.1 shows the performance of Algorithm 3 for this example. The initial conditions are \( u(0) = .21_N, p(0) = .11_N, \) and \( w(0) = 1_N, \delta = 0.01, \bar{\delta} = .9, \) and \( \alpha = 0.001 \). Figure 6.1 shows the trajectory of \( \lambda_1(A(\delta)) \) over time. It starts from \( \lambda_1(A(\delta(0))) > 1 \) and we get \( \lambda_1(A(\delta^*)) = .8711 \). The maximum number of iterations of the Power Iteration at each iteration of Algorithm 3 algorithm is 9.
Figure 6.1: Evolution of $\lambda_1(A(\delta))$ for Algorithm 3 when the distributed stopping condition for the estimation of the gradient is used in the Enron network example.

6.3 Summary

We have proposed a distributed stopping criterion for the Power Iteration method. We provide bounds guarantees in the eigenvector estimates that allow us to interconnect this method with our recent developed $p$-Robust Box-Gradient Fairness algorithm with provable convergence guarantees. We have applied our result to minimize the spread of a virus over a complex network in a completely distributed fashion, which allows to preserve privacy to its constituents. As a future work, we want to extend our distributed criterion for the Power Iteration method together with the bounds on the estimates for more general classes of matrices.
Publications associated with this chapter

This paper contains material that has been published in the following works:

Chapter 7

Gradient-free distributed resource allocation via the SPSA method

In this chapter, we propose and analyze a novel distributed discrete-time stochastic algorithm to solve a class of distributed resource allocation problems. In particular, we extend our p-robust box-gradient fairness algorithm with an SP technique but, unlike previous works, we employ a constant step-size. In this way, our approach allows an interconnected group of agents to collectively minimize a global cost function subject to both equality and inequality constraints, where the closed-form expression of the local cost functions is unknown to the agents. Under some technical conditions, we show that the algorithm converges in probability to a small neighborhood of the solution as long as the chosen step-size is sufficiently small. It is shown that the proposed algorithms are convergent to a neighborhood around the equilibrium even when there are temporary errors in communication or computation. Thus, agents do not require global knowledge
of total resources in the network or employ any special procedure for initialization. Our
algorithm is provable correct over weight-balanced and strongly connected networks. In
the proofs, we employ Lyapunov theory together with tools from convex analysis and
stochastic difference inclusions.

7.1 Problem statement, solution approach, and algorithm

In this section, we introduce the optimization problem we are set out to solve,
which is followed by the proposed stochastic box-constrained gradient algorithm with
guaranteed convergence to their corresponding optimizer under complementary sets of
assumptions.

7.1.1 Problem statement and solution approach

We consider a network of \( N \) agents connected over a digraph whose goal is to
minimize the sum of local payoff functions \( f_i : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, i \in \{1, \ldots, N\} \), where no ex-

cplicit closed-form expression of the function \( f_i \) is available, under resource constraints.
The box-coupled fairness optimization problem is given by

\[
\min_{p \in [\underline{p}, \bar{p}]^N} \sum_{i=1}^{N} f_i(p_i) \\
\text{s.t. } \mathbf{1}_N^T p = \mathbf{1}_N^T \bar{u},
\]

(7.1)
where $f_i$ is the payoff, $p = [p_1, \ldots, p_N]^\top \in \mathbb{R}^N$ is the resource allocation, $\bar{u}_i \in \mathbb{R}$ is the input assumed to be constant that represents the available quantity of resources for each agent, $\bar{u} = [\bar{u}_1, \ldots, \bar{u}_N]^\top$, and $\underline{p}, \bar{p} \in \mathbb{R}^N$ are the lower and upper limits of the optimization variable, respectively. We name the last constraint in (7.1) as the box constraint. We simply refer to the problem with the box constraint omitted as the coupled fairness optimization problem. To solve both problems we state the following assumption.

**Assumption 7.1.** (Problem assumptions): *We assume that the box-coupled fairness has a unique solution and there is not explicit closed-form expression of the payoff function $f_i$, $i \in \{1, \ldots, N\}$. The only information available are measurements of $f_i$ at the parameter $p_i$. Furthermore, we assume $f_i$ is twice continuously differentiable and bounded below with uniform bounded gradients, i.e., there exist constant $M \in \mathbb{R}_{>0}$ such that $\max_{i \in \mathcal{V}} \left| \frac{\partial f_i(p)}{\partial p_i} \right| \leq M$ for $p \in \mathbb{R}^N$. An upper bound of $M$ is assumed to be known. An agent $i \in \mathcal{V}$ should be able to measure or obtain $f_j(p_j)$ for $j \in \mathcal{N}_i^\text{out}$. We assume that the box constraints are explicitly given.*

Under the same assumptions as for the last problem and using the exact penalty method (see, e.g., [43]), we reformulate the box-coupled fairness problem as follows:

$$\begin{align*}
\min_p \hat{f}(p) \\
s.t. \quad 1_N^\top p = 1_N^\top \bar{u},
\end{align*}$$

(7.2)

where $\hat{f}(p) \triangleq \sum_{i=1}^N f_i(p_i) + J(p)$, $J(p) \triangleq \chi \sum_{i=1}^N (\max([p_i - \bar{p}_i, 0]) + \max([\underline{p}_i - p_i, 0]))$, and $\chi \in \mathbb{R}_{>0}$. The next lemma characterizes the optimal solution to the box-coupled fairness optimization
Lemma 7.1. (Solution of the box-coupled fairness problem [10]): Let Assumption 7.1, on the payoff characteristics for the coupled fairness problem, hold. Let $\chi \in \mathbb{R}_{>0}$ be such that
\begin{equation}
\chi > 2 \max_{p \in \mathbb{R}^N} \| \nabla_p f(p) \|_{\infty}.
\end{equation}

Then, the solution $p^*$ to the box-coupled fairness optimization problem satisfies
\begin{align}
\zeta^* 1_N &\in \nabla_p f(p^*) + \partial J(p^*), \\
1_N^T p^* &= 1_N^T \bar{u},
\end{align}

where $\zeta \in \mathbb{R}$ is the Lagrange multiplier for the equality constraint of the box-coupled fairness problem.

Next, we propose a distributed discrete-time algorithm which successfully converges to the solutions of the box-coupled fairness problem introduced above under the corresponding assumptions. We will refer to them as the stochastic box-constrained gradient algorithm.

### 7.1.2 Proposed algorithm

In order to solve the box-coupled fairness problem dynamically, we introduce the stochastic box-constrained gradient algorithm shown in Algorithm 4, where
\begin{align}
\Sigma = \left\{ 
\begin{array}{l}
w^+ = w - \alpha L (g + \psi), \\
p^+ = p + \alpha (-L(g + \psi) + w - p + u),
\end{array}
\right.
\end{align}

where $w \in \mathbb{R}^N$ is an internal estimator state assumed $w(0) = 0$, $\alpha \in (0, 1)$ is the step-size, $L$ is the Laplacian matrix associated to directed graph $G$, $g(p, \delta, v) \triangleq$
\[[g_1(p_1, \delta, v_1), \ldots, g_N(p_N, \delta, v_N)]^T \in \mathbb{R}^N \text{ with } g_i : \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \text{ for } i \in \mathcal{V}, \text{ given as:} \]
\[
g_i(p, \delta, v) = \frac{f_i(p_i + \delta_1 v_i) - f_i(p_i - \delta_2 v_i)}{\delta_1 + \delta_2} v_i^{-1}. \]

Here, \( f_i : \mathbb{R} \to \mathbb{R} \) has the same meaning as in (7.1), \( \delta = (\delta_1, \delta_2)^T \in \mathbb{R}^2 \), the random variables \( \{v_k\}_{k \in \mathbb{Z}_{>0}} \) take values in \( \{-1, 1\}^N \), \( \psi \in \partial J(p) \), \( u \in \mathbb{R}^N \) is defined as \( u = \bar{u} + \bar{\epsilon} \), where \( \bar{u}_i \in \mathbb{R} \) is the input assumed to be constant that represents the available quantity of resources for each agent, \( \bar{u} = (\bar{u}_1, \ldots, \bar{u}_N)^T \), and \( \bar{\epsilon} \in \mathbb{R}^N \) is defined as

\[
\bar{\epsilon}_i = \begin{cases} 
-\epsilon, & \text{if } p_i^+ = \bar{p}_i \\
\epsilon, & \text{if } p_i^- = p_i \\
0, & \text{otherwise},
\end{cases}
\]

where \( \epsilon \in \mathbb{R}_{>0} \) is a given small constant satisfying \( \epsilon \leq \alpha \).

**Algorithm 4** One step of the **stochastic box-constrained gradient** algorithm for agent \( i \in \mathcal{V} \)

1: \( \tilde{\epsilon}_i = 0 \)
2: Compute \( \Sigma_i \) as in (7.5)
3: if \( p_i^+ = \bar{p}_i \) then
   4: \( \tilde{\epsilon}_i = -\epsilon \)
5: end if
6: if \( p_i^- = p_i \) then
   7: \( \tilde{\epsilon}_i = \epsilon \)
8: end if
9: Compute \( p_i^+ \) as in (7.5b)

**Remark 7.1.** Notice that the **stochastic box-constrained gradient** algorithm does not allow to have \( p_i^+ = \bar{p}_i \) or \( p_i^- = p_i \) since it perturbs \( \Sigma_i \) using a small quantity \( \tilde{\epsilon}_i \) at any time this happens.
Since the box constraints are explicitly given, the generalized gradient of the penalty function is directly used in the algorithm. We make the following assumption on the sequence of random variables $v$.

**Assumption 7.2.** (On the characteristics of the random input): The sequence of random variables $\{v_k\}_{k \in \mathbb{Z}_{\geq 0}}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $v_k : \Omega \rightarrow \{-1, 1\}^N$, is i.i.d. with $E[v_k] = 0$ for each $k \in \mathbb{Z}_{\geq 0}$.

In what follows we use the following notation. We refer to $\mathcal{F}_{\leq}^u \triangleq \{p \in \mathbb{R}^N \mid \mathbf{1}_N^T p \leq \mathbf{1}_N^T \bar{u} + N\epsilon\}$, $\mathcal{F}_{\geq}^u \triangleq \{p \in \mathbb{R}^N \mid \mathbf{1}_N^T p \geq \mathbf{1}_N^T \bar{u} - N\epsilon\}$, $\mathcal{F}^u \triangleq \mathcal{F}_{\leq}^u \cap \mathcal{F}_{\geq}^u$, and $\mathcal{F}_{\text{box}}^\nu = \{p \in \mathbb{R}^N \mid p - \nu \mathbf{1}_N \leq p \leq p + \nu \mathbf{1}_N\}$ for $\nu \in \mathbb{R}_{> 0}$.

**Remark 7.2.** For the easiness of presentation we neglect the presence of noise in the observations of $f_i$, $i \in \mathcal{V}$. However, from the analysis in the next section, practical convergence in expected value to the equilibrium point can be achieved under appropriate statistical properties on the noise.

### 7.2 Stability analysis

In this section, we show that the equilibrium point of the stochastic box-constrained gradient dynamics coincide with the optimal solutions of the corresponding problem under the stated assumptions when $\mathcal{G}$ is strongly connected and weight-balanced, and $g(p, \delta, v)$ is replaced by $\nabla_p f(p)$ in (7.5). Theorem 7.1 presents the stability properties of this dynamics.
Lemma 7.2. (Equilibria of the stochastic box-constrained gradient algorithm): Let Assumption 7.1, on the payoff characteristics for the coupled fairness problem, hold. Let $G$ be a weight-balanced and strongly connected graph. Let the point $p^*$ represent the solution of the stochastic box-constrained gradient algorithm. Replace $g(p, \delta, v)$ by $\nabla_p f(p)$ and let $\epsilon = 0$ in (7.5). Then, the point $p^*$ is the solution of the stochastic box-constrained gradient algorithm if and only if there exists $\eta^* \in \mathbb{R}$ such that

$$
\eta^* 1_N \in \partial \hat{f}(p^*),
$$

(7.6a)

$$
1_N^T p^* = 1_N^T \bar{u}.
$$

(7.6b)

Proof. To obtain (7.6a), we have that, at the equilibrium, $0 \in -\alpha L \xi^*$ for all $i \in \mathcal{V}$, where $\xi^* \in \partial \hat{f}(p^*)$. Then $\eta^* 1_N \in \partial \hat{f}(p^*)$. In order to get (7.6b), we have $0 \in \alpha (-L \xi^* + w^* - p^* + \bar{u})$. Pre-multiplying (7.6a) by $1_N^T$, we obtain $1_N^T w^* = 1_N^T \bar{w}$, which implies that $1_N^T w(k) = 1_N^T w(0) = 0$ for all $k \geq 0$. Using the last fact and pre-multiplying (7.6b) by $1_N^T$, it follows that $0 = -1_N^T p^* + 1_N^T \bar{u}$. \qed

The following lemma characterizes the invariance of $\mathcal{F}^u_\leq$ and $\mathcal{F}^u_\geq$ with respect to the stochastic box-constrained gradient dynamics.

Lemma 7.3. (Invariance of the resource constraint under (7.5)): Let Assumption 7.1, on the payoff characteristics for the min-max fairness problem, hold. Let $G$ be a weight-balanced and strongly connected graph. Assume $\alpha \in (0, 1)$ in (7.5). Then, the sets $\mathcal{F}^u_\leq$ and $\mathcal{F}^u_\geq$ are strongly positively invariant under the stochastic box-constrained gradient dynamics.

Proof. Without loss of generality we analyze the case of the set $\mathcal{F}^u_\leq$ since for the case of $\mathcal{F}^u_\geq$ we simply switch the $\leq$ inequality by $\geq$ in what follows. We define a suitable
Lyapunov function as follows. Consider $V : \mathbb{R}^{2N} \to \mathbb{R}_{\geq 0}$, $V(p, w) = 1_N^T(p - p^*)$. In the following, to analyze $V$ we assume $\bar{\epsilon} = 0$. We will show in that the solution $1_N^T p(k)$ for $k \in \mathbb{Z}_{>0}$ goes exponentially to $1_N^T p^*$, which implies that it has the Input-to-State Stability (ISS) property. Therefore, when the perturbation $1_N^T \bar{\epsilon}$ is present, we already know that $\Delta V$ is decreasing outside of an open ball of radius $1_N^T \bar{\epsilon}$, which can be made arbitrarily small by reducing $\epsilon$. Let $\Delta V = V(p^+, w^+) - V(p, w)$, it follows

\[
\Delta V = (1_N^T(p^* - p^*) - 1_N^T(p - p^*)) \\
= (1_N^T p + \alpha 1_N^T(-p + \bar{u}) - 1_N^T p) \\
= -\alpha 1_N^T(p - \bar{u}) = -\alpha V \leq 0,
\]

where we have used the facts that $1_N^T w^+ = 1_N^T w$, which implies that $1_N^T w(k) = 1_N^T w(0) = 0$ for $k \in \mathbb{Z}_{>0}$ and $1_N^T p^* = 1_N^T \bar{u}$, in obtaining the second equality. Note $\Delta V \leq 0$ if $\alpha \in (0, 1)$.

Now we analyze the most positive value of the perturbation $1_N^T \bar{\epsilon}$, which is given when all $i \in \mathcal{V}$ take the value $\epsilon$. Suppose that $p(0) \in \mathcal{F}_\le^u$, then, we have that $p(k) \in \mathcal{F}_\le^u$ due to the ISS property.

Before presenting our main results, the following lemma characterizes $g(p, \delta, \nu)$ in terms of the gradient of the cost function $f$. Lemma 7.5 shows that the trajectories of the stochastic box-constrained gradient dynamics are bounded.

**Lemma 7.4.** (SP approximation to the gradient): Let Assumption 7.2, on the characteristics of the random input, hold. Assume that $f$ is convex, finite, and twice differentiable
Then
\[ g_i(p, \delta, v) = \frac{\partial f(p)}{\partial p_i} + b_i + c_i, \quad (7.7) \]

where \( b_i = \sum_{j \neq i} \frac{\partial f(p)}{\partial p_j} \), for \( i \in \{1, \ldots, n\} \), and \( c = \frac{v^{-1}}{2(\delta_1 + \delta_2)} v^\top (\delta_1 \nabla^2 f(p^1) - \delta_2 \nabla^2 f(p^2)) v \), for \( p = p + \delta_j' v \) for some \( \delta_j' \in [0, 1] \) and \( j \in \{1, 2\} \).

**Proof.** Using a second-order Taylor expansion around \( p \), there exists \( \delta_1' \in [0, 1] \) such that
\[ f(p + \delta_1 v) = f(p) + \delta_1 v^\top \nabla_p f(p) + \frac{1}{2} \delta_1^2 v^\top \nabla^2 f(p^1) v, \quad (7.8) \]
where \( p^1 = p + \delta_1' v \). Similarly, there is \( \delta_2' \in [0, 1] \) such that
\[ f(p - \delta_2 v) = f(p) - \delta_2 v^\top \nabla_p f(p) + \frac{1}{2} \delta_2^2 v^\top \nabla^2 f(p^2) v, \quad (7.9) \]
where \( p^2 = p - \delta_2' v \). Subtracting (7.9) from (7.8) and dividing the result by \( \delta_1 + \delta_2 \), we have
\[
\frac{f(p + \delta_1 v) - f(p - \delta_2 v)}{\delta_1 + \delta_2} = v^\top \nabla_p f(p) + \frac{1}{2(\delta_1 + \delta_2)} v^\top \times (\delta_1^2 \nabla^2 f(p^1) - \delta_2^2 \nabla^2 f(p^2)) v.
\]

Multiplying the last equation by \( v^{-1} \) we have
\[ g_i(p, \delta, v) = v^\top \nabla_p f(p) v^{-1} + \frac{v^{-1}}{2(\delta_1 + \delta_2)} \times v^\top (\delta_1^2 \nabla^2 f(p^1) - \delta_2^2 \nabla^2 f(p^2)) v. \quad (7.10) \]

We analyze the \( i \)-th component of the first term of the right-hand side (RHS) of last equation:
\[(v^\top \nabla_p f(p) v^{-1})_i = \frac{1}{v_i} \sum_{j=1}^n v_j \frac{\partial f(p)}{\partial p_j} \]
\[= \frac{\partial f(p)}{\partial p_i} + b_i. \quad (7.11)\]

Replacing (7.11) in (7.10), (7.7) follows. \[\Box\]

**Lemma 7.5.** *(Boundedness of the stochastic box-constrained gradient dynamics):* Let Assumption 7.1, on the payoff characteristics for the coupled fairness problem, hold. Let \(G\) be weight-balanced and strongly connected. Assume that
\[\chi > \frac{1}{\min_{(i,j) \in E} a_{ij}} (2Md_{out,max} + \|w(0) - p(0) + \bar{u}\|_\infty), \quad (7.12)\]
and
\[\alpha < \frac{\min_{i \in V} \{\bar{p}_i - p_i\}}{2d_{out,max}(M + \chi) + \|w(0) - p(0) + \bar{u}\|_\infty}, \quad (7.13)\]
where \(d_{out,max} = \max_{i \in V} \sum_{j=1}^N a_{ij}\). Then, there exists \(\nu\) such that
\[\nu \leq \max\{|\nu_1|, |\nu_2|, \nu_3\} \quad (7.14)\]
where \(\nu_1 = \max\{1^\top_N p(0), 1^\top_N \bar{u} + N\epsilon\} - (N - 1) \min_j p_j\), \(\nu_2 = \min\{1^\top_N p(0), 1^\top_N \bar{u} + N\epsilon\} - (N - 1) \max_j \bar{p}_j\) and \(\nu_3 = \alpha(2d_{out,max} \max_{i \in V} \|\nabla_p f(p) + b + c\|_\infty + 2\chi d_{out,max} + \|w(0) - p(0) + u\|_\infty)\), for which the set \(\mathcal{F}^{\nu}_{box}\) is strongly positively invariant under the stochastic box-constrained gradient algorithm.

**Proof.** First we want to show that if \(\chi\) satisfies (7.12), there exists \(\nu > 0\) such that
\[p(k) \in \mathcal{F}^{\nu}_{box} \cap \mathcal{F}_{\leq 0}, \text{ for } k \in \mathbb{Z}_{\geq 0}.\]

In the case that any agent \(i \in V\) has a \(\bar{\epsilon}_i \neq 0\), it means that \(p_i^+\) has been perturbed by \(\pm \epsilon\) and thus it is inside the box constraints. Then, without loss of generality we assume \(\bar{\epsilon} = 0\) for all time in the following analysis.
Without loss of generality assume \( p(k) \in \mathcal{F}_u^\leq \), for \( k \in \mathbb{Z}_{>0} \) (recall that by Lemma 7.3, the sets \( \mathcal{F}_u^\leq \) and \( \mathcal{F}_u^\geq \) are invariant with respect to the stochastic box-constrained gradient dynamics). Then, the trajectories can only leave the set \( \mathcal{F}_{box}^\gamma \cap \mathcal{F}_u^\leq \), for any \( \nu \), by violating the box constraints.

We reason this is not the case by contradiction. Assume that \( \mathcal{F}_{box}^\gamma \) is not strongly positively invariant under the stochastic box-constrained gradient algorithm. This implies that there exists a boundary point \( p_{bd} \in \text{bd}(\mathcal{F}_{box}^\nu) \), an integer number \( \gamma > 0 \), and a trajectory \( k \mapsto p(k) \) obeying (7.5b) such that \( p(0) = p_{bd} \) and \( p(k) \notin \mathcal{F}_{box}^\nu \) for all \( k \in \{1, \ldots, \gamma\} \).

Without loss of generality, assume that \( p(k) \in \mathcal{F}_{box}^{\nu_1} \), for all \( k \in \{1, \ldots, \gamma\} \), for some \( \nu_1 \in \mathbb{R}_{>0} \) with \( \nu_1 > \nu \).

Without loss of generality, there must exist \( i \in \mathcal{V} \) such that \( p_i(0) = \overline{p_i} + \nu \) and \( p_i(k) > \overline{p_i} + \nu \) for all \( k \in \{1, \ldots, \gamma\} \). Define \( \Delta p \triangleq p(k+1) - p(k) \) and we assume that \( \alpha \) is small enough (we characterize the size of \( \alpha \) at the end of the proof) so that \( p(k+1) \) either satisfies the box constraint \([\underline{p}, \overline{p}]\) or it remains to be beyond the upper bound of the box constraint \( \overline{p} \). This means that there must exist \( k \) such that \( p(k+1) \in p(k) + \alpha(-L(\partial \hat{f}(p(k)) + b + c) + w(k) - p(k) + \bar{u}) \) and \( \gamma_1 \in \{1, \ldots, \gamma\} \) such that \( \Delta p_i \geq 0 \) over \( \{1, \ldots, \gamma_1\} \). Without loss of generality assume \( \gamma_1 = \gamma \).

Next, we show that this can only happen if \( p_j(k) \geq \overline{p_j} + \nu \) for all \( j \in \mathcal{N}_i^{\text{out}} \) and all \( k \in \{1, \ldots, \gamma_1\} \).

Since \( p_i(k) > \overline{p_i} + \nu \), for all \( k \in \{1, \ldots, \gamma_1\} \), then \( (\partial \hat{f}(p))_i = \{(\nabla_p f(p))_i + \chi\} \).

Therefore,
\[ p_i(k+1) = p_i(k) - \alpha \sum_{j=1}^{N} a_{ij}(\nabla p f(p))_i + b_i + c_i + \chi - \eta_j) + \alpha w(k) - p(k) + \bar{u}), \]

where \( \eta_j \in (\partial \hat{f}(p))_j + b_j + c_j \), for all \( j \).

Note that if \( p_j \geq \overline{p}_j + \nu \), then \( \eta_j \leq (\nabla p f(p))_j + b_j + c_j + \chi \), whereas if \( p_j < \overline{p}_j \), then \( \eta_j \leq (\nabla p f(p))_j + b_j + c_j \). For convenience, denote this latter set of neighbors by \( N_i^\leq \). Now, we upper bound \( \Delta p_i \) as

\[
\Delta p_i \leq -\alpha \sum_{j=1}^{N} a_{ij}(\nabla f(p))_i + b_i + c_i - (\nabla f(p))_j - b_j - c_j - \alpha \chi \sum_{j \in N_i^\leq} a_{ij} + \alpha w - p + \bar{u})_i
\]

\[
\leq 2\alpha \max_{i \in V}(\nabla f(p) + b + c) \sum_{i=1}^{N} a_{ij} + \alpha \|w - p + \bar{u}\|_{\infty} - \alpha \chi \sum_{j \in N_i^\leq} a_{ij}
\]

\[
\leq 2\alpha d_{\text{out,max}} \|\nabla f(p) + b + c\|_{\infty}
\]

\[
= 2\alpha d_{\text{out,max}} \|g(p, \delta, v)\|_{\infty}
\]

\[
\leq 2\alpha d_{\text{out,max}} M
\]

\[
+ \alpha \|w(0) - p(0) + \bar{u}\|_{\infty} - \alpha \chi \sum_{j \in N_i^\leq} a_{ij}
\]

\[
< 0,
\]

where the last inequality follows from (7.12) and the fact that \( \|g(p, \delta, v)\|_{\infty} \leq M \).

To see this last fact, notice that \( |g_i(p, \delta, v)| = \left| \frac{f((p_1 + \delta_1 v_1) - f((p_1 - \delta_2 v_2) v_1^{-1})}{\delta_1 + \delta_2} \right| \leq M v_i^2 \). Also, notice that \( \|w(k) - p(k) + \bar{u}\|_{\infty} \leq \|w(0) - p(0) + \bar{u}\|_{\infty} \). To see this, take a Lyapunov function

\[ V = \|w - p + \bar{u}\|_{\infty} \]

and check that \( \Delta V = -\alpha \|w - p + \bar{u}\|_{\infty} \) for \( \alpha \in (0, 1) \). In all, \( \Delta p_i \geq 0 \) is possible only if \( p_j \geq \overline{p}_j + \nu \) for all \( j \in N_i^\text{out} \) on \( t \in \{1, \ldots, \gamma_1\} \).

Extending the argument to the neighbors of each \( j \in N_i^\text{out} \), we have that \( \gamma_2 \leq \gamma_1 \) for some \( \gamma_2 \) over which all one- and two-hop neighbors of \( i \) have the states greater or equal to their respective maximum limits. Recursively, and since the graph is strongly
connected and the number of nodes is finite, there is a \( \tilde{\gamma} \geq 1 \) over which \( p(k) \geq \bar{p} + \nu \mathbf{1}_N \) with \( p(0) \leq \bar{p} + \nu \mathbf{1}_N \), which contradicts the fact that \( p \in \mathcal{F}^u \). The above argument shows that there exists a finite \( \nu \) such that \( p(k) \in \mathcal{F}^v_{\text{box}} \cap \mathcal{F}^u \) for all \( k > 0 \).

Next, we characterize \( \nu \). We have two cases to analyze. First, assume that there exist \( i \in \mathcal{V} \) which in the next iteration is going to jump outside the box constraint. The worse-case scenario is given when \( p_i(k) = p_{\text{bd}} \). Without loss of generality assume \( p_i = \bar{p}_i \). Then \( p_i(k + 1) \) is upper bounded by

\[
 p_i(k + 1) \leq p_i(k) + \alpha(\|L(\xi + b + c) + w(k) - p(k) + u\|_\infty) \\
 \leq \bar{p}_i + \alpha(\|L(\xi + b + c)\|_\infty + \|w(0) - p(0) + u\|_\infty) \\
 \leq \bar{p}_i + \alpha(2d_{\text{out,max}} \max_p \|\nabla_p f(p) + b + c\|_\infty \\
 + 2\chi d_{\text{out,max}} + \|w(0) - p(0) + u\|_\infty),
\]

Second, we follow a similar argument as the one we have shown for the existence of a finite \( \nu \). Specifically, we have shown that if there exists \( i \in \mathcal{V} \) that is leaving from above (below) the box constraints, then all \( j \in \mathcal{V} \setminus \{i\} \) must leave the box constraints from above (below). Without loss of generality assume that \( p \leq p(k) \) for all \( k \geq 0 \). Assume that there exists \( i \in \mathcal{V} \) such that \( p_i(0) \geq \bar{p}_i \). By lemma 7.3, we have that \( \mathbf{1}_N^T p(k) \) has an exponential trend to \( \mathbf{1}_N^T u \), so \( \mathbf{1}_N^T p(k) \leq \max\{\mathbf{1}_N^T p(0), \mathbf{1}_N^T \bar{u} + N\epsilon\} \). As a worst case scenario assume \( p_j = \bar{p}_j \) for all \( j \in \mathcal{V} \setminus \{i\} \). Then \( p_i(k) \leq \max\{\mathbf{1}_N^T p(0), \mathbf{1}_N^T \bar{u} + N\epsilon\} - \sum_{j \neq i} \bar{p}_j \leq \max\{\mathbf{1}_N^T p(0), \mathbf{1}_N^T \bar{u} + N\epsilon\} - (N - 1) \min_{j \neq i} \bar{p}_j \). By taking the above two cases we have

\[
 p_i(k) \leq \max\{\max\{\mathbf{1}_N^T p(0), \mathbf{1}_N^T \bar{u} + N\epsilon\} - (N - 1) \min_{j \neq i} \bar{p}_j, \\
 \bar{p}_i + \alpha(2d_{\text{out,max}} \max_p \|\nabla_p f(p) + b + c\|_\infty \\
 + 2\chi d_{\text{out,max}} + \|w(0) - p(0) + u\|_\infty).
\]

Therefore, (7.14) follows.
We have shown the invariance of the set $\mathcal{F}_\text{box}^\nu \cap \mathcal{F}_\leq^\mu$. This argument holds if a small enough $\alpha$ is assumed. The size of $\alpha$ matters to discard the case of a state jump from one side of the box constraint to the other, i.e., if $p_i(k) \geq \bar{p}_i$ for some $k > 0$ and $i \in \mathcal{V}$, then it is not possible to have $p_i(k+1) \leq \underline{p}_i$ or viceversa. Using (7.5b) with $\bar{\epsilon} = 0$, we obtain the bound $|\Delta p| \leq \alpha(2d_{\text{out},\max}(M + \chi) + ||w(0) - p(0) + \bar{u}||_\infty)$. What we want is to make the right hand side of the last inequality to be strictly less than $\min_{i \in \mathcal{V}}{p_i - p_i^*}$, which is satisfied when (7.13) holds.

**Lemma 7.6.** (Positive semidefiniteness of class of functions): Let Assumption 7.1 hold for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ for some $n \in \mathbb{Z}_{>0}$. Assume $\gamma I \leq \nabla^2 f(p) \leq \Gamma I$ for some $\gamma, \Gamma \in \mathbb{R}_{>0}$ and all $p \in \mathbb{R}^n$. Let $C$, $f^*$ and $C^*$ be defined as in Lemma 5.6. Assume $C^*(v)$ is single-valued for every $v \in [a, b]$, for given $a, b \in \mathbb{R}$ such that $a \leq b$, $p^* = C^*(a)$, and $C^*$ locally Lipschitz at $a$. Consider $V' : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq0}$, $V' = V_1 + V_4$, where $V_1(p) = f(p) + J(p) - f(p^*) = \hat{f}(p) - f(p^*)$ and $V_4(p) = m'1_n^T(p - p^*)$, for $m' \in \mathbb{R}_{>0}$. Then, for $p \in C(v)$, there exists $m'$ such that $V' \geq 0$.

**Proof.** Let $\hat{f}^* : [a, b] \rightarrow \mathbb{R}$, $\hat{f}^*(v) \triangleq \min\{\hat{f}(p) \mid p \in C(v)\}$, where $C(v) = \{p \in \mathbb{R}^n \mid 1_n^T p = v\}$. By Lemma 5.6, $\hat{f}^*$ is convex. Using Lemma 7.5, the set $\mathcal{F}_\text{box}^\nu$ is strongly positively invariant under $\Sigma_1$, so $\hat{f}^*$ can be extended over an open interval containing $[a, b]$, and then $\hat{f}^*$ is locally Lipschitz at $a$.

Let $\hat{f}^*_\min \triangleq \min\{\hat{f}^*(v) \mid v \in [a, b]\}$. If $\hat{f}^*(a) = \hat{f}^*_\min$, then $V_1 + V_4$ is positive semidefinite by the convexity of $\hat{f}^*$ so that any positive $m'$ works. Suppose that $\hat{f}^*(a) > \hat{f}^*_\min$. 

\[\hat{f}^* = \phi(f, p, \Sigma_1, \mathcal{V}, a, b)\]
Since $\hat{f}^*$ is continuous, then by the extreme value theorem (see, e.g., [56] theorem 4.16), $\hat{f}^*$ attains a minimum in $[a, b]$, and then $\hat{f}_{\min}^* > -\infty$. By convexity of $\hat{f}$, we have that $\hat{f}^*(v) \leq \hat{f}(p)$ for all $p \in C(v)$ with $v \in [a, b]$.

We are interested in those cases for which $V_1 < 0$. For that reason, in the following analysis we assume that $\hat{f}(p) < f^*(a)$ for $p \in C(v)$ with $v \in (a, b]$. Exploiting the locally Lipschitz property of $\hat{f}^*$ at $a$, then there exists $l \in \mathbb{R}_{>0}$ such that

$$\frac{\hat{f}^*(a) - f^*(p)}{(1,p-a)} \leq l < +\infty$$

for all $p \in C(v)$ with $v \in (a, b]$. It follows that $V_1 + V_4 \geq 0$ for $m' = l$.

**Theorem 7.1.** *(Stability of the stochastic box-constrained gradient algorithm):* Let Assumption 7.1, on the payoff characteristics for the coupled fairness problem, hold. Assume $G$ be a weight-balanced and strongly connected graph. Then, for any constant input $\bar{u} \in \mathbb{R}^N$ and any initial state $p(0)$, and $w(0) = 0$, the solution of the system (7.5) converges asymptotically in probability to an open ball of radius depending on $\alpha, \delta_1 + \delta_2$, and $\epsilon$ centered at the equilibrium point (7.6).

**Proof.** Without loss of generality note that $f$ can be assumed to be strongly convex under Assumption 7.1. To see this, if we define a new convex program by changing the cost function of (7.1) by $f(p) + \|p - p^*\|_2^2$, where $p^*$ is the optimal solution to (7.1), then the solution of the original program and the new one is the same under Assumption 7.1. Furthermore, both problems satisfy the optimality condition (7.4) in Lemma 7.1. Therefore, it can be assumed that $f$ is strongly convex and that $\gamma I \leq \nabla^2_p f(p) \leq \Gamma I$ for some $\gamma, \Gamma \in \mathbb{R}_{>0}$. Notice that the last upper bound on the Hessian matrix of $f$ comes from
the fact that the compact set \( F^\nu_{\text{box}} \) is strongly positively invariant under \( \Sigma \) by Lemma 7.5. Therefore the maximum eigenvalue of \( \nabla^2 f(p) \), which is a continuous function of \( p \) on \( F^\nu_{\text{box}} \), is bounded above on \( F^\nu_{\text{box}} \), i.e., there exists a constant \( \Gamma \) such that \( \nabla^2 f(p) \leq \Gamma I \).

**Lyapunov function definition**

With the help of the previous lemmas, we define a suitable Lyapunov function that is used to derive the stability result. Consider \( V : \mathbb{R}^{2N} \rightarrow \mathbb{R}_{\geq 0}, V = V_1 + V_2 + V_3 + V_4, \) where \( V_1(p,w) = f(p) + J(p) - f(p^*) \equiv \hat{f}(p) - f(p^*), V_2(p,w) = (M + \chi + \ell)||w - p + u||_2 + \Gamma||w - p + u||_2^2, \) for any \( \ell \in \mathbb{R}_{>0}, V_3(p,w) = c_1m(1_N^\top(p - p^*))^2, c_1 = \frac{1}{2}\lambda_2(L + L^\top), \) for some \( m \in \mathbb{R}_{>0}, \) and \( V_4(p,w) = m'(1_N^\top(p - p^*)), \) for some \( m' \in \mathbb{R}_{>0}. \)

The constant \( m \) in \( V_3 \) is chosen from Lemma 5.8. By this lemma, there exists \( m \) such that \( ||\xi - \text{Avg}(\xi)1_N||_2^2 + m(1_N^\top(p - p^*))^2 \geq \frac{\gamma^2}{8}||p - p^*||_2^2 \) holds. To satisfy the assumptions of Lemma 5.8, take \( a = 1_N^\top p^*, b = 1_N^\top p(0), \) and recall that, by Lemma 7.3, \( F_u^\nu \) and \( F_{\geq}^\nu \) are strongly positively invariant under \( \Sigma \), so without loss of generality we assume \( p(0) \in F_{\geq}^u \). Notice that the segment \([a - N\epsilon, b]\) is strongly positively invariant under \( \Sigma \) by Lemma 7.3.

Analogous to the choice of \( m \), the constant \( m' \) in \( V_4 \) is chosen from Lemma 7.6. By this lemma, there is \( m' \) such that \( V_1 + V_4 \geq 0. \) To satisfy the assumptions in Lemma 7.6 take \( a \) and \( b \) as defined for the case of \( m \). Recall that by assumption \( \chi \) satisfies (7.12) and \( ||\nabla_p f(p)||_\infty \leq M \) for any \( p \in \mathbb{R}^N \).

First, we want to show that there is a ball containing \( p^* \) with radius depending on \( \alpha \) which is globally recurrent. For that, we show that \( V(p,w) \) satisfies (2.12), and
then this holds for the dynamics $\Sigma$. We write

$$p^+ = p + \alpha \phi,$$

where $\phi = -L\xi - L(b + c) + w - p + u$, and $\xi \in \partial \hat{f}(p)$. Since $\Sigma$ does not allow to have $p_i = \bar{p}_i$ or $p_i = \underline{p}_i$ for all $i \in \mathcal{V}$, i.e., $p$ is never a point where $f$ is nonsmooth, then without loss of generality we assume $\xi = \nabla_p \hat{f}(p)$ in the following analysis. Also notice that the Hessian matrix $\nabla^2 \hat{f}(p) = \nabla^2 f(p)$ for all $p \in \mathbb{R}^N \setminus [\bar{p}, \underline{p}]$.

**One-step differences of $V$**

In the following the analyze the one-step differences of the summands of $V$, starting with $V_1$. Using the second-order Taylor expansion, we have

$$\hat{f}(p^+) = \hat{f}(p) + \alpha \xi^\top \phi + \frac{1}{2} \alpha^2 \phi^\top \nabla^2 f(p) \phi + o(\alpha^2 \|\phi\|^2).$$

It follows that

$$\Delta V_1 = \hat{f}(p^+) - \hat{f}(p)$$

$$= \alpha \xi^\top (-L\xi - L(b + c) + w - p + u)$$

$$+ \frac{1}{2} \alpha^2 \phi^\top \nabla^2 f(p) \phi + o(\alpha^2 \|\phi\|^2).$$

Continuing from the last equality, and using that $\max(|\xi|) \leq M + \chi$

$$\Delta V_1 \leq -\alpha \xi^\top L\xi - \alpha \xi^\top L(b + c)$$

$$+ \alpha(M + \chi)\|w - p + u\|_2 + \frac{1}{2} \alpha^2 \phi^\top \nabla^2 f(p) \phi$$

$$+ o(\alpha^2 \|\phi\|^2).$$

Including now the expression for $\phi = -L\xi - L(b + c) + w - p + u$ in the quadratic Hessian term, and using the fact that $x^\top \nabla^2 f(p') x + y^\top \nabla^2 f(p') y \geq x^\top \nabla^2 f(p') y + y^\top \nabla^2 f(p') x$ for any $x, y, p'$
\[
\Delta V_1 \leq -\alpha \xi^T L \xi + \alpha \xi^T L (b + c) \\
+ \alpha (M + \chi) ||w - p + u||_2 \\
+ \alpha^2 \xi^T L^T \nabla^2 f(p) L \xi \\
+ \alpha^2 (b + c)^T L^T \nabla^2 f(p) L (b + c) \\
+ \alpha^2 (w - p + u)^T \nabla^2 f(p) (w - p + u) + o(\alpha^2 ||\phi||_2^2).
\]

In the following, we apply that, for any vector \( x \in \mathbb{R}^N \), and matrix \( A \in \mathbb{R}^{N \times N} \), we have \( ||A||_2 = \sigma_1(A) \) and \( ||Ax||_2^2 \leq \sigma_1(A)^2 ||x||_2^2 \) (e.g., see [35]) together with the fact that \( \nabla^2 f(p) \leq \Gamma I \). Moreover, since \( G \) is weight-balanced and strongly connected, then \( \xi^T L \xi \geq c_1 ||\hat{\xi}||_2^2 \), for all \( \xi \in \mathbb{R}^N \), where \( \hat{\xi} = \xi - \text{Avg}(\xi) \mathbf{1}_N \) and \( c_1 = \lambda_2(L + L^T) \). Then,

\[
\Delta V_1 \leq -c_1 \alpha ||\xi - \text{Avg}(\xi) \mathbf{1}_N||_2^2 \\
+ \alpha \xi^T L (b + c) + \alpha (M + \chi) ||w - p + u||_2 \\
+ \alpha^2 \Gamma (\sigma_1^2(L)(M + \chi)^2) \\
+ \alpha^2 (b + c)^T L^T \nabla^2 f(p) L (b + c) \\
+ \alpha^2 \Gamma ||w - p + u||_2^2 + o(\alpha^2 ||\phi||_2^2).
\]

Next, for \( V_2(p, w) = (M + \chi + \ell)||w - p + u||_2 + \Gamma||w - p + u||_2^2 \) and \( \Delta V_2 \), we have

\[
\Delta V_2 = (M + \chi + \ell)(||w^* - p^* + u||_2 - ||w - p + u||_2) \\
+ \Gamma(||w^* - p^* + u||_2^2 - ||w - p + u||_2^2) \\
= (M + \chi + \ell)((1 - \alpha)||w - p + u||_2 - ||w - p + u||_2) \\
+ \Gamma((1 - \alpha)^2 ||w - p + u||_2^2 - ||w - p + u||_2^2) \\
= -\alpha (M + \chi + \ell)||w - p + u||_2 \\
+ \Gamma((1 - \alpha)^2 - 1)||w - p + u||_2^2.
\]

In the following, to analyze \( V_3 \) and \( V_4 \) we assume \( \epsilon = 0 \). We will show in the next steps that the solution \( \mathbf{1}_N^T p(k) \) for \( k \in \mathbb{Z}_{>0} \) goes exponentially to \( \mathbf{1}_N^T p^* \), which implies that it has the ISS property. Therefore, when the perturbation \( \mathbf{1}_N^T \bar{\epsilon} \) is present, we already know that \( \Delta V_3 \) and \( \Delta V_4 \) are decreasing outside of an open ball of radius \( \mathbf{1}_N^T \bar{\epsilon} \), which can be made arbitrarily small by reducing \( \epsilon \). For \( V_3(p, w) = c_4 m(\mathbf{1}_N^T (p - p^*))^2 \) and \( \Delta V_3 \), we have
\[ \Delta V_3 = c_1 M((1_n^T(p^+ - p^*)^2) - (1_n^T(p - p^*))^2) \\
= c_1 M((1_n^T(p + \alpha(-L(\xi + b + c) + w - p + u) - p^*))^2 - (1_n^T(p - p^*))^2). \]

Recall that \(1_n^T L(\xi + b + c) = 0\) since the \(G\) is weight-balanced. Moreover, \(1_n^T w = 0\) since \(1_n^T w^+ = 1_n^T w\). Using these two facts one has
\[
\Delta V_3 = c_1 M((1_n^T(p - p^* + \alpha(-p + u))^2 - (1_n^T(p - p^*))^2) \\
= c_1 M((1_n^T(p - p^*))^2 + 2\alpha(1_n^T(p - p^*))(1_n^T(-p + u)) + \alpha^2(1_n^T(-p + u))^2 - (1_n^T(p - p^*))^2) \\
= c_1 M(-2\alpha(1_n^T(p - p^*))^2 + \alpha^2(1_n^T(-p + u))^2) \\
= c_1 M(-\alpha(1_n^T(p - p^*))^2 - \alpha(1 - \alpha)(1_n^T(p - p^*))^2), \\
= -\alpha(2 - \alpha)V_3, \\
\]
where we have used the fact that \(1_n^T p^* = 1_n^T u\), in obtaining the third and fourth equalities. Finally, for \(V_4(p, w) = m'1_n^T(p - p^*)\) and \(\Delta V_4\) one has
\[
\Delta V_4 = m'(1_n^T(p^+ - p^*) - 1_n^T(p - p^*)) \\
= m'(1_n^T p + \alpha 1_n^T(-p + u) - 1_n^T p) \\
= -m'\alpha 1_n^T(p - u) = -m'\alpha V_4 \leq 0.
\]

Notice that \(\Delta V_3 \leq 0\) and \(\Delta V_4 \leq 0\) if \(\alpha \in (0, 1)\).

Define \(c_2 = \sigma_2^2(\Gamma(M + \chi)^2)\). Now we put all inequalities together to obtain an upper bound on \(\Delta V\)
\[
\Delta V \leq -\alpha c_1||\xi - \text{Avg}(\xi)1_n||_2^2 \\
+\alpha|\xi|L(b + c) \\
+\alpha^2 c_2 + \alpha^2(b + c)^T L^T \nabla^2 f(p)L(b + c) \\
+o(\alpha^2||\phi||_2^2) \\
-\alpha\ell||w - p + u||_2 - 2\alpha\Gamma(1 - \alpha)||w - p + u||_2^2 \\
-c_1 m\alpha(1_n^T(p - p^*))^2 - c_1 m\alpha(1 - \alpha)(1_n^T(p - p^*))^2 \\
- m'\alpha V_4 \\
= -\alpha c_1(||\xi - \text{Avg}(\xi)1_n||_2^2 + m(1_n^T(p - p^*))^2) \\
+\alpha|\xi|L(b + c) \\
+\alpha^2 c_2 + \alpha^2(b + c)^T L^T \nabla^2 f(p)L(b + c) \\
+o(\alpha^2||\phi||_2^2) - \alpha\ell||w - p + u||_2 \\
-2\alpha\Gamma(1 - \alpha)||w - p + u||_2^2 - c_1 m\alpha(1 - \alpha)(1_n^T(p - p^*))^2 \\
- m'\alpha V_4.
\]
Using Lemma 5.8 together with the facts that $\mathcal{F}_S^a$ is positively invariant and that $\Delta V_3$ and $\Delta V_4$ are decreasing outside of an open ball of radius $1_N^\epsilon$ which can be made arbitrarily small by reducing $\epsilon$, we have

\[
\Delta V \leq -\alpha c_1(1 - \theta)\gamma \|p - p^*\|_2^2 - \alpha c_1\theta \gamma \|p - p^*\|_2^2 + \alpha \xi^T L(b + c) + \alpha^2 c_2 + \alpha^2(b + c)^T L^T \nabla^2 f(p) L(b + c) + o(\alpha^2\|\phi\|_2^2) - \alpha \ell \|w - p + u\|_2 \\
-2\alpha \Gamma(1 - \alpha)\|w - p + u\|_2^2 - c_1 m \alpha (1 - \alpha)(1_N^\epsilon(p - p^*)^2 - m'\alpha V_4),
\]

where $\theta \in (0, 1)$. In order to have $\Delta V \leq 0$, it must be $\alpha \in (0, 1)$.

**Expected value of the difference of $V$**

At this point, we take expected values, to finally conclude the recurrence result.

By noticing that $E[b] = E[c] = 0$, it follows that

\[
E[\Delta V|\mathcal{F}_k] \leq -\alpha c_1(1 - \theta)\gamma \|p - p^*\|_2^2 - \alpha c_1\theta \gamma \|p - p^*\|_2^2 + \alpha \xi^T L(b + c) + \alpha^2 c_2 + o(\alpha^2\|\phi\|_2^2) - \alpha \ell \|w - p + u\|_2 \\
-2\alpha \Gamma(1 - \alpha)\|w - p + u\|_2^2 - c_1 m \alpha (1 - \alpha)(1_N^\epsilon(p - p^*)^2 - m'\alpha V_4).
\]

Now, we obtain upper bounds for the terms that have non-zero

\[
E[(b + c)^T L^T \nabla^2 f(p) L(b + c)|\mathcal{F}_k] \leq \sigma_1(L)\Gamma E[\|b\|_2^2 + \|c\|_2^2|\mathcal{F}_k].
\]

Recall from Lemma 7.4 that $g(p, \delta, \nu) = \nabla_p f(p) + b + c$, where $b_i = \sum_{j\in i} \nu_j \frac{\partial f(p)}{\partial p_j}$, for $i \in \{1, \ldots, N\}, c = \frac{\nu^2}{2(\delta_1 + \delta_2)} v^T (\delta_1^2 \nabla^2 f(p^1) - \delta_2^2 \nabla^2 f(p^2))v$, $p^1 = p + \delta_j' v$ for some $\delta_j' \in [0, 1]$ and $j \in \{1, 2\}$. It follows that

\[
\|b\|_2^2 \leq \|((\nabla_p f(p))^T v^{-1} - \nabla_p f(p))\|_2^2 \\
\leq \|((\nabla_p f(p))^T v^{-1})\|_2^2 + \|\nabla_p f(p)\|_2^2 \\
\leq NM^2(N + 1),
\]
and
\[
\|c\|_2 \leq \frac{1}{2(\delta_1 + \delta_2)}\|v\|_2\|v\|^2_2(\delta_1^2 + \delta_2^2)\Gamma \\
\leq \frac{1}{2(\delta_1 + \delta_2)}((\delta_1 + \delta_2)^2 - 2\delta_1\delta_2)\Gamma\|v\|_2^2 \\
\leq \frac{1}{2}\Gamma(\delta_1 + \delta_2)\mathcal{N}^2.
\]

To analyze when \(E[\Delta V|\mathcal{F}_k] \leq 0\), consider first
\[
\|p - p^*\|_2^2 \geq \frac{\alpha c_1(L)\Gamma((NM^2(N + 1)) + \frac{1}{4}\Gamma^2(\delta_1 + \delta_2)^2\mathcal{N}^3)}{\frac{\alpha c_1}{\Gamma}(\delta_1 + \delta_2)} \frac{c_1\gamma\theta}{\alpha c_1\gamma\theta} - \frac{o(\alpha^2\|L\xi\|_2^2) + o(\alpha^2\|b + c\|^2_2) + o(\alpha^2\|w - p + u\|_2^2)}{\alpha c_1\gamma\theta}.
\]

where \(\xi \in \partial \hat{f}(p)\). Recall that \(\bar{u} \in \mathbb{R}^N\) is a given constant, \(u = \bar{u} + \bar{\epsilon}\), and \(\epsilon \leq \alpha\).

Notice that \(o(\|w - p + u\|_2^2)\) converges exponentially fast to the segment \([-1^\top N\bar{\epsilon}, 1^\top N\bar{\epsilon}] \subset [-N\epsilon, N\epsilon] \subset [-Na, Na]\) which can be made arbitrarily small by reducing \(\alpha\). This fact can be shown if we define \(W(p, w) = \|w - p + \bar{u}\|_2^2\) and assume \(\bar{\epsilon} = 0\). It follows that \(\Delta W \leq ((1 - \alpha^2) - 1)W\), which implies that \(\|w - p + \bar{u}\|_2^2\) goes to zero exponentially and that has ISS property.

Also, notice that \(o(\|L\xi\|_2^2) \leq l(M + \chi)^2\) for some \(l \in \mathbb{R}_{>0}\) and that the term \(E[\|b + c\|^2_2|\mathcal{F}_k]\) is bounded. We define
\[
\varphi(\alpha) \equiv \frac{\alpha c_1(L)\Gamma((NM^2(N + 1)) + \frac{1}{4}\Gamma^2(\delta_1 + \delta_2)^2\mathcal{N}^3)}{\frac{\alpha c_1}{\Gamma}(\delta_1 + \delta_2)} \frac{c_1\gamma\theta}{\alpha c_1\gamma\theta} + \frac{o(\alpha^2)}{\alpha c_1\gamma\theta}.
\]

Thus, \(E[\Delta V|\mathcal{F}_k] \leq 0\) if \(\|p - p^*\|_2^2 \geq \varphi(\alpha)\), where \(\varphi\) is a class-'K' function.

Next, we show that \(V(p, w)\) is radially unbounded over \(w \in \mathcal{F}_b\), where \(\mathcal{F}_b \equiv \{w \in \mathbb{R}^N | 1_N^\top w = b\}\) and \(b\) is any given constant. To analyze \(\lim V(p, w)\) when \(\|(p, w)\|_2 \to +\infty\)
consider two situations: (i) when there exists \( p_i \rightarrow \pm \infty \ i \in V \) and (ii) when there exists \( w_i \rightarrow \pm \infty \) for some \( i \in V \) and \( p \) is bounded. If (i) holds, the result follows from \( f(p) + J(p) \) being radially unbounded with respect to \( p \). Under (ii), if there exist \( w_i \rightarrow +\infty \), then there must exist \( j \neq i \) such that \( w_j \rightarrow -\infty \), and vice versa with \( w \in F_b \).

This implies that \( w \neq d1_N \), for some \( d \in \mathbb{R} \) as \( \|w\|_2 \rightarrow +\infty \) and thus \( \|w - p + u\|_2 \rightarrow +\infty \) when there exists \( i \) such that \( w_i \rightarrow \pm \infty \). It follows that \( V \) is radially unbounded over \( F_b \).

Notice that \( F_b \) is strongly positively invariant under the box-coupled fairness algorithm by taking \( b = 1_N^T w(0) \).

Since \( V \) is radially unbounded and \( E[\Delta V|F_k] \leq 0 \) if \( \|p - p^*\|_2^2 \geq \varrho(\alpha) \), then the sublevel sets of \( V \) are compact and positively invariant. All assumptions in Proposition 2.1 are satisfied, then the set \( \{p \in \mathbb{R}^N \mid \|p - p^*\|_2^2 < \varrho(\alpha)\} \) is globally recurrent. Moreover, when \( \alpha = 0 \), \( \Sigma \) is stable in probability, which implies that the solution of the system (7.5) converges asymptotically in probability to an open ball of radius depending on \( \alpha \) and \( \delta_1 + \delta_2 \) centered at the equilibrium point (7.6).

\[ \square \]

### 7.3 Distributed economic dispatch with unknown utility functions

We consider an electricity market consisting of suppliers and consumers, where the profit and cost functions of power generation and consumption is unknown to the consumers and suppliers. We denote by \( x_i \) the electricity consumption of consumer \( i \in \{1, \ldots, n\} \). Each consumer is associated with a cost function \( f_i^c : \mathbb{R} \rightarrow \mathbb{R} \). We
denote by $z_j$ the electricity production by supplier $j \in \{1, \ldots, m\}$. We define $N$ as the dimension of the state space, i.e., $N = n + m$. The suppliers have an associated cost function $f_j^s : \mathbb{R} \to \mathbb{R}$. The market clearing procedure can be formulated as

$$
\begin{align*}
\min_{x,z} & \quad \sum_{i=1}^{n} f_c^i(x_i) + \sum_{i=1}^{m} f_s^i(z_i) \\
\text{s.t.} & \quad \sum_{i=1}^{n} x_i \leq \sum_{i=1}^{m} z_i, \\
& \quad x \in [x, \bar{x}], \ z \in [z, \bar{z}].
\end{align*}
$$

(7.15)

Notice the optimization problem above does not have the form of the box-coupled fairness optimization problem. However, we can transform the problem into

$$
\begin{align*}
\min_{x,z,s} & \quad \sum_{i=1}^{n} f_c^i(x_i) + \sum_{i=1}^{m} f_s^i(-z_i) \\
\text{s.t.} & \quad \sum_{i=1}^{n} x_i + \sum_{i=1}^{m} z_i + 1^\top_N s = 0, \\
& \quad s \geq 0, \ x \in [x, \bar{x}], \ z \in [-\bar{z}, -\bar{z}].
\end{align*}
$$

(7.16)

which allows us to use the stochastic box-constrained gradient algorithm, as explained next.

**Remark 7.3.** A more general equality constraint can be considered for the box-coupled fairness optimization problem. For example, if we consider the equality constraint of the form $c^\top p = 1^\top_N \bar{u}$, where $c_i \in \mathbb{R}_{>0}$. Then, we can use a new variable $y_i = c_i p_i$ for
\( i \in \{1, \ldots, N\}. \) The box-coupled fairness optimization problem becomes

\[
\min_{y_i \in [c_i p_i, c_i p_i]} f(\bar{y})
\]

\[
\text{s.t. } \mathbf{1}_N^T y = \mathbf{1}_N^T \bar{u},
\]

(7.17)

where \( \bar{y} \triangleq [c_1^{-1} y_1, \ldots, c_n^{-1} y_N]^T. \) Notice that \( f \) is still convex with respect to \( y \) since it is a composition of an affine mapping.

**Remark 7.4.** When an inequality constraint ‘\( \leq \)’ is considered instead of the equality constraint for the box-coupled fairness problem, we can add \( N \) slack variables \( s \in \mathbb{R}_+^N \) to this problem to convert the inequality constraint into an equality one. In this case, the box-coupled fairness optimization problem is equivalent to

\[
\min_{p \in [p, p], s \in \mathbb{R}_+^N} f(p)
\]

\[
\text{s.t. } \mathbf{1}_N^T p + \mathbf{1}_N^T s = \mathbf{1}_N^T \bar{u}.
\]

(7.18)

To see that the problem (7.18) is equivalent to the original one, first notice that if \( (p, s) \) is feasible for the problem (7.18), then \( p \) is feasible for the original problem, since

\[
\mathbf{1}_N^T s = \mathbf{1}_N^T \bar{u} - \mathbf{1}_N^T p \geq 0.
\]

Conversely, if \( p \) is feasible for the original problem, then \( (p, s) \) is feasible for the problem (7.18), where we take \( \mathbf{1}_N^T s = \mathbf{1}_N^T \bar{u} - \mathbf{1}_N^T p. \)

**Example 7.1** (Distributed electricity market). We consider an electricity market consisting in 10 consumers and 5 suppliers. This example is partially taken from [64]. Our example differs from the example in [64] since we consider a graph for the topology of the electricity market. We illustrate the response of the stochastic box-


7.4 Summary

Building on our previous algorithm \( p \)-robust box-gradient fairness and the simultaneous perturbation method, we have introduced a novel stochastic algorithm that allows a group of agents to find the minimizer of an unknown cost function while sat-
Figure 7.1: Evolution of the electricity market for Example 7.1. Negative trajectories correspond to the suppliers \( z(k) \) and positive ones to the consumers \( x(k) \). Dashed lines are the optimal consumption and production \( x^* \) and \( z^* \), respectively. In \( k = 15 \times 10^3 \) it is forced the states \( x(k) \) and \( z(k) \) to be zero during an interval of 100 iterations.

Isfying inequality and equality constraints. We have proven convergence in a practical way to the solution as long as the chosen step-size is sufficiently small. In particular, the proposed algorithm are designed to be robust to temporary errors in communication or computations of agents. Our technical approach relies on results nonsmooth Lyapunov theory, convex analysis, and stochastic difference inclusions. Motivated by applications to resource allocation and optimization, we plan to extend available proofs that can help us relax the assumptions needed.

**Publications associated with this chapter**

This paper contains material that has been published in the following works:
• E. Ramírez-Llanos and S. Martínez, “Gradient-free distributed resource allocation via simultaneous perturbation”. Communication, Control, and Computing (Allerton), 2016, USA.
Chapter 8

Stochastic source seeking for mobile robots in environments with obstacles via the SPSA method

Stochastic source seeking algorithms are used in mobile robotics to find a source of a radiation-like signal in GPS-denied environments. Applications range from biology, in understanding bacterial foraging, to security, for rescue operations and chemical detection. In a typical setting, the robot samples the signal emitted by the source by exploring the environment through a stochastic motion. The samples are used to steer and climb the gradient of the signal field, where the signal field might represent the spatial distribution of magnetic force, thermal signal, or chemical concentration.

Our approach is inspired by the simultaneous perturbation stochastic approximation (SPSA) algorithm. The SPSA method is a well-known algorithm usually applied to
estimate the gradient of a cost function from (noisy) measurements. It was proposed first in [65] and has been successfully applied in many optimization problems such parameter estimation, simulation optimization, resource allocation, and robotics.

The standard SPSA method uses a monotonically decreasing step-size to solve unconstrained optimization. In mobile robots, a decreasing step-size is not an option, since it is impossible to navigate with infinite precision, which is the case when the step-size is converging to zero. For this reason, we propose a modified version of the SPSA algorithm which uses a small, but constant step-size for environments that may contain obstacles.

**Literature review.** There are many approaches to stochastic source seeking for mobile robots in position-denied environments. An example is given in [66], where it is applied the extremum seeking framework to nonholonomic vehicles, or the application of the SPSA algorithm to mobile robots in [67, 68]. We follow the approach of [67], where the authors design a controller to drive a robot to the source by applying the SPSA algorithm without the use of the position information. The algorithm they propose uses standard assumptions found in SPSA in the literature such as thrice differentiability of the cost function and monotonically decreasing step-size. Those assumptions fit very well in many applications where direct measurements of the gradient are not available. However, in some applications, a decreasing step-size is not an option.

An alternative is to use a small, but constant step-size, which has been successfully applied in applications such as optimization of combustion control [69], mobile robots [68], and tracking and adaptive control [70].
In [69], a variation of the SPSA algorithm is proposed which decreases the oscillation against the constraints. The proposed algorithm is applied to an automotive combustion engine problem. Although [69] uses a constant step-size, no theoretical guarantees are given for fixed step-size.

In [68], a model-free algorithm is proposed based on stochastic approximation to find a source in environments with obstacles, which uses a constant step-size. A decreasing step-size is not desirable because the robot might get trapped in a location where the magnitude of the gradient is small. The convergence of the algorithm in [68] is shown through an experiment in a real world scenario. However, its convergence is not proven theoretically.

In [70], an algorithm inspired by SPSA is proposed for unconstrained optimization. The algorithm uses a constant step-size to minimize a cost function for three different tracking problems: a random walk, an optimization of UAV’s flight, and a load balancing. A drawback of their algorithm is that the cost function is assumed to be once differentiable and it solves an unconstrained optimization problem.

Following a different approach to the gradient-free algorithms, in [71] it is proposed an algorithm to guide a robot through an unknown obstacle environment using sensed information from a single intensity source. The proposed algorithm is similar in spirit to the well-known family of bug algorithms and its novelty is in the fact it uses less sensing information than any other algorithms in the literature. However, a drawback in their approach is the assumption on the availability of the gradient generated by the intensity source, which in general is not true.
Statement of contributions. We propose a stochastic source seeking algorithm to drive a robot to an unknown source signal by only using measurements of the signal field. Our proposed algorithm builds on the SPSA algorithm. The novelty of our approach is that we consider nondifferentiable convex functions, fixed step-size, and the environment may have obstacles. We prove practical convergence to a ball and whose size depends on the step-size that contains the location of the source. For the proof, we use Lyapunov theory together with tools from convex analysis and stochastic difference inclusions. Our proof does not rely on stochastic approximation theory as is usually the case for algorithms in the literature based on SPSA. Finally, we show the applicability of the proposed algorithm in a 2D scenario for the source seeking problem.

8.1 Problem statement

This section describes the stochastic source seeking problem for GPS-denied environments. This problem has been studied for obstacle-free environments in e.g. [66], [67], and [68]. In particular, we follow the approach of [67], except that we consider boundaries and obstacles in the environment. Suppose that a sufficiently small robot moves in $\mathbb{R}^n$ and its motion is described in the world coordinate frame by

$$
(p^\top, \theta^\top, \phi^\top)^\top = G(p(t), \theta(t), \phi(t))u(t),
$$

(8.1)

where $G : \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{(n+n_1+n_2)\times m}$ is a function describing the robot dynamics, $p(t) \in \mathbb{R}^n$ and $\theta(t) \in \mathbb{R}^{n_1}$ are the translational and orientational positions in the world of coordinate frame, $\phi(t) \in \mathbb{R}^{n_2}$ and $u(t) \in \mathbb{R}^m$ are the internal posture and the control input,
respectively.

Let $\mathcal{E}$ be the environment where the robot moves, which is assumed to be convex and compact. A tower broadcasts a signal, which is modeled by an intensity function $f$ over $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be the signal mapping, in which $f(p)$ yields the intensity at $p \in \mathbb{R}^n$, generated from a tower at $p^*$. The location of the tower $p^*$ can be or not in $\mathcal{E}$. The environment $\mathcal{E}$ and the signal mapping $f$ are unknown to the robot. The robot aims to solve the following optimization problem without knowledge of its own position and orientation

$$\min_{p \in \mathcal{E}} E[f(p)|p],$$

by only using measurements of $f(p(t))$. We consider two scenarios for $\mathcal{E}$: when $\mathcal{E}$ does not have obstacles and when it does. For the obstacle-free scenario, we assume $\mathcal{E}$ is a convex compact set. In the second scenario, we allow $\mathcal{E}$ to have obstacles. In both scenarios, the problem is to design an algorithm with guaranteed practical convergence to a small ball containing $p^*$. When $p^* \notin \mathcal{E}$ the robot should converge in practical way to a ball containing the closest point from $\mathcal{E}$ to $p^*$.

The robot is equipped with two sensors. First, we assume a contact sensor indicates whether the robot is touching the environment boundary or any obstacle inside the environment $\partial \mathcal{E}$

$$l_{\mathcal{E}}(p) = \begin{cases} 
1, & \text{if } p \in \partial \mathcal{E}, \\
0, & \text{otherwise}. 
\end{cases}$$

Second, the robot is equipped with an intensity sensor $l_I$, which indicates the strength of the signal from position $p$, i.e, $l_I(p) = f(p)$. Since the robot does not have position
information in the coordinate frame, it is necessary to adapt (8.1) to a body fixed frame.

The position of the robot in the body frame at time $t$ is given by

$$\begin{pmatrix}
z(t) \\
\psi(t) \\
\phi(t)
\end{pmatrix} = \begin{pmatrix}
R(-\theta(\tau))(p(t) - p(\tau)) \\
\theta(t) - \theta(\tau) \\
\phi(t)
\end{pmatrix},$$

where $t$ expresses a future time after $\tau$, $(z(t), \psi(t), \phi(t)) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ are the new coordinates, and $R(-\theta(\tau))$ is the rotation matrix of an angle $-\theta(\tau)$.

### 8.2 Proposed algorithm for the obstacle-free scenario

In this section we assume that there are no obstacles in $\mathcal{E}$, where $\mathcal{E}$ is described by a convex compact set. To find $p^*$, we propose the following algorithm, which is similar in spirit to the stochastic approximation algorithm for fixed step-size:

$$p_{k+1} = \Pi_{\mathcal{E}}[p_k - \alpha g(p_k, \delta(p_k, R_k v_k), R_k, v_k)],$$

where $k \in \mathbb{Z}_{\geq 0}$. To simplify the notation for aid in analysis, we write the above algorithm as a discrete-time dynamical system as follows $p^+ = \Pi_{\mathcal{E}}[p - \alpha g(p, \delta(p, R v), R, v)]$, where $p \in \mathbb{R}^n$ is the current state, $p^+ \in \mathbb{R}^n$ is the state at the next time step, $\Pi_{\mathcal{E}}$ is the projection on a convex compact set $\mathcal{E}$, and $g : \mathbb{R}^n \times \mathbb{R}^2 \times \text{SO}(n) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is given as:

$$g(p, \delta(p, R v), R, v) = \begin{cases}
R \frac{f(p+\delta_1 R v) - f(p-\delta_2 R v)}{\delta_1 + \delta_2} v^{-1}, & \text{if } \delta_1 + \delta_2 > 0, \\
0, & \text{otherwise.}
\end{cases}$$
Here, $f: \mathbb{R}^n \to \mathbb{R}$ is the function to be minimized, $\alpha \in \mathbb{R}_{>0}$ is the step-size, $R \in \text{SO}(n)$ is the uncertain time-varying rotation matrix, which by definition is an orthogonal matrix, and $\delta = (\delta_1, \delta_2)$, $\delta_i: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, for $i \in \{1, 2\}$ are defined as

$$
\delta_1(p, Rv) = \begin{cases} 
\bar{\delta}_1, & \text{if } p + \bar{\delta}_1 Rv \in \mathcal{E}, \\
\text{dist}_{Rv}(p, \partial \mathcal{E}), & \text{otherwise},
\end{cases}
$$

$$
\delta_2(p, Rv) = \begin{cases} 
\bar{\delta}_2, & \text{if } p - \bar{\delta}_2 Rv \in \mathcal{E} \\
\text{dist}_{-Rv}(p, \partial \mathcal{E}), & \text{otherwise},
\end{cases}
$$

where $\bar{\delta}_1, \bar{\delta}_2 \in \mathbb{R}_{\geq 0}$ are given constants satisfying $\bar{\delta}_1 + \bar{\delta}_2 > 0$, $\text{dist}_{Rv}(p, \partial \mathcal{E})$ is the distance between the point $p$ and the set $\partial \mathcal{E}$ along the direction $\pm Rv$, and the random variable $\{v_k\}_{k \in \mathbb{Z}_{\geq 0}}$ takes values in $\{-1, 1\}^n$. We assume that there is an algorithm which gives the distance from $p$ to the position where the robot found the obstacle. This routine can be designed using information of the acceleration and the contact sensor $l_E$ of the robot. We make the following assumption on the sequence of random variables $v$.

**Assumption 8.1.** (On the characteristics of the random input): The sequence of random variables $\{v_k\}_{k \in \mathbb{Z}_{\geq 0}}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $v_k: \Omega \to \{-1, 1\}^n$, is i.i.d. with $E[v_k] = 0$ for each $k \in \mathbb{Z}_{\geq 0}$.

**Remark 8.1.** For the easiness of presentation we neglect the presence of noise in the observations of $f$. However, from the analysis in the next section, practical convergence in expected value to the tower can still be achieved under appropriate statistical properties on the noise.
8.3 Convergence for the obstacle-free scenario

In this section, we derive the convergence results for the algorithm in (8.3). In particular, we show practical convergence in probability to a ball with fixed radius depending on $\alpha$ and $\tilde{\delta}_1 + \tilde{\delta}_2$ under different assumptions. We are able to characterize the size of this ball under the assumption of strong convexity of the cost function as shown in Theorem 8.1. However, when we do not have enough information on the cost function, like differentiability, we prove practical convergence in probability to a ball that can be made arbitrarily small by tuning $\alpha$ and $\tilde{\delta}$ as shown in Theorem 8.2. We begin by providing two supporting lemmas.

**Lemma 8.1.** (SPSA approximation to the gradient): Let Assumption 8.1, on the characteristics of the random input, hold. Assume that $f$ convex, finite, and twice differentiable. Then, if $\delta_1 + \delta_2 > 0$ we have

$$g_i(p, \delta, R, v) = \frac{\partial f(p)}{\partial p_i} + b_i + c_i,$$

(8.4)

where $b_i = \sum_{l,j,q,j \neq l} R_{il} R_{jv} \frac{\partial f(p)}{\partial p_q}$, for $i \in \{1, \ldots, n\}$, $c = \frac{Rv^\top R^\top}{2(\delta_1 + \delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) R v$, $p_j = p + \delta_j R v$ for some $\delta_j \in [0, 1]$ and $j \in \{1, 2\}$. Otherwise, if $\delta_1 + \delta_2 = 0$, we have $g_i(p, \delta, R, v) = 0$ for $i \in \{1, \ldots, n\}$.

**Proof.** For the case when $\delta_1 + \delta_2 = 0$, by definition it follows that $g_i(p, \delta, R, v) = 0$. Otherwise, when $\delta_1 + \delta_2 > 0$, by using a second-order Taylor expansion around $p$, there
exists \( \delta'_1 \in [0, 1] \) and \( p_1 = p + \delta'_1 Rv \) such that

\[
f(p + \delta_1 Rv) = f(p) + \delta_1 v^\top R^\top \nabla_p f(p) + \frac{1}{2} \delta_1^2 v^\top R^\top \nabla^2 f(p_1) Rv. \tag{8.5}
\]

Similarly, there is \( \delta'_2 \in [0, 1] \) and \( p_2 = p - \delta'_2 Rv \) such that

\[
f(p - \delta_2 Rv) = f(p) - \delta_2 v^\top R^\top \nabla_p f(p) + \frac{1}{2} \delta_2^2 v^\top R^\top \nabla^2 f(p_2) Rv. \tag{8.6}
\]

Subtracting (8.6) from (8.5) and dividing the result by \( \delta_1 + \delta_2 \), we have

\[
\frac{f(p + \delta_1 Rv) - f(p - \delta_2 Rv)}{\delta_1 + \delta_2} = v^\top R^\top \nabla_p f(p) + \frac{1}{2(\delta_1 + \delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) Rv.
\]

Multiplying last equation by \( Rv^{-1} \) we have

\[
R \frac{f(p + \delta_1 Rv) - f(p - \delta_2 Rv)}{\delta_1 + \delta_2} v^{-1} = v^\top R^\top \nabla_p f(p) Rv^{-1} + \frac{Rv^{-1}}{2(\delta_1 + \delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) Rv. \tag{8.7}
\]

We analyze next the \( i \)-th component of the first term of the right-hand side (RHS) of last equation:

\[
(v^\top R^\top \nabla_p f(p) Rv^{-1})_i = \sum_{l=1}^n \frac{R_{il}}{v_l} \sum_{j=1}^n \frac{v_j}{v_l} \sum_{q=1}^n \frac{R_{qj}}{v_l} \frac{\partial f(p)}{\partial p_q}
\]

\[
= \sum_{l=1}^n \frac{R_{il}}{v_l} \sum_{j=1}^n \frac{v_j}{v_l} \sum_{q=1}^n \frac{R_{qj}}{v_l} \frac{\partial f(p)}{\partial p_q} + \sum_{l=1}^n \frac{R_{il}}{v_l} \sum_{q=1}^n \frac{v_j}{v_l} \frac{\partial f(p)}{\partial p_q}
\]

\[
= \frac{\partial f(p)}{\partial p_i} + b_i, \tag{8.8}
\]

where we have used the fact that \( R \) is an orthogonal matrix. Replacing (8.8) in (8.7), (8.4) follows.
**Lemma 8.2.** *(Optimality bounds):* Assume $f$ is convex, finite, and satisfies the superquadratic growth condition in (2.15). Then, for all $\xi \in \partial f(p)$ and $p, p^* \in \mathbb{R}^n$ it holds

$$ (p - p^*)^\top \xi \geq \frac{\rho}{2} \|p^* - p\|^2, \tag{8.9} $$

and,

$$ \|\xi\| \geq \frac{\rho}{2} \|p^* - p\|. \tag{8.10} $$

**Proof.** We prove first inequality (8.9). By the assumption on the superquadratic growth condition (2.15) it holds

$$ f(p^*) \geq f(p) + (p^* - p)^\top \xi + \frac{\rho}{2} \|p^* - p\|^2, $$

for all $p, p^* \in \mathbb{R}^n$, and $\xi \in \partial f(p)$. Subtracting $f(p)$ on both sides, we have

$$ f(p^*) - f(p) \geq (p^* - p)^\top \xi + \frac{\rho}{2} \|p^* - p\|^2. $$

By noticing that $f(p^*) - f(p) \leq 0$, we have

$$ 0 \geq (p^* - p)^\top \xi + \frac{\rho}{2} \|p^* - p\|^2. $$

Then, (8.9) follows. Next, we prove (8.10). By noting that the RHS of (8.9) is bigger or equal than zero, it follows $\|(p - p^*)^\top \xi\| \geq \frac{\rho}{2} \|p^* - p\|^2$. By using the Cauchy-Schwarz inequality, it follows that $\|p - p^*\| \|\xi\| \geq \frac{\rho}{2} \|p^* - p\|^2$, which implies (8.10). $\square$

The next theorem shows algorithm convergence when $f$ is twice differentiable.

**Theorem 8.1.** *(Convergence when $f$ is twice differentiable):* Let Assumption 8.1, on the characteristics of the random input, hold. Assume that $f$ is convex, finite, twice differentiable, $\rho I_n \leq \nabla^2 f(p) \leq \Gamma I_n$, and $\|\nabla_p f(p)\| \leq M$. Assume $\alpha$ and $\bar{\delta}_1 + \bar{\delta}_2$ are
sufficiently small. Then, for any initial state $p_0$, the solution $p^*$ of the system (8.3) is MSP-ES with ultimate bound $O = \mathcal{E} \setminus Z$, where

$$Z = \{ p \in \mathcal{E} | \| p - p^* \|^2 \geq \frac{\alpha}{\rho}(M^2(n^2 + 2) + \frac{1}{4}(\bar{\delta}_1 + \bar{\delta}_2)^2 \Gamma^2 n^3) \}.$$  \hspace{1cm} (8.11)

**Proof.** Without loss of generality assume $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$. This is the case because, at any time $k > 0$ for which $\delta_1(p, Rv) + \delta_2(p, Rv) = 0$, with probability one, the dynamics in (8.3) will generate a feasible direction in finite time in $\mathcal{E}$ satisfying $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$.

Without loss of generality assume $p^* \in \mathcal{E}$ (the projection of $p^*$ on $\mathcal{E}$ is in $\mathcal{E}$ and unique.) By the non-expansive property of the projection operation, Algorithm (8.3), and the fact that $p^* \in \mathcal{E}$, we have

$$\| p^+ - p^* \|^2 = \| \Pi_\mathcal{E}[p + \alpha g(p, \delta(p, Rv), R, v)] - p^* \|^2$$

$$\leq \| p + \alpha g(p, \delta(p, Rv), R, v) - p^* \|^2$$

$$= \| p + \alpha(\nabla p f(p) + b + c) - p^* \|^2$$

$$= \| p - p^* \|^2 - 2\alpha(\nabla p f(p) + b + c)^\top (p - p^*) + \alpha^2\| \nabla p f(p) + b + c \|^2,$$

where $b_i = \sum_{i,j,q \neq i} R_{ij} R_{qj} \frac{\partial f(p)}{\partial p_i}$, for $i \in \{1, \ldots, n\}$, $c = \frac{Rv^{-1}}{2(\bar{\delta}_1 + \bar{\delta}_2)} v^\top R^\top (\bar{\delta}_1^2 \nabla^2 f(p_1) - \bar{\delta}_2^2 \nabla^2 f(p_2)) Rv$, $p_j = p + \delta'_j Rv$ for some $\delta'_j \in [0, 1]$ and $j \in \{1, 2\}$ (see Lemma 8.1 to learn how to get $b$ and $c$).

Let $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, $V = \| p - p^* \|^2$, and define $\Delta V = \| p^+ - p^* \|^2 - \| p - p^* \|^2$. We have

$$\Delta V \leq -2\alpha(\nabla p f(p) + b + c)^\top (p - p^*) + \alpha^2\| \nabla p f(p) + b + c \|^2.$$

By using (8.9), we have that $-(p - p^*)^\top \nabla p f(p) \leq -\frac{\rho}{2}\| p - p^* \|^2$. It follows that

$$\Delta V \leq -\alpha\rho\| p - p^* \|^2 - 2\alpha(b + c)^\top (p - p^*) + \alpha^2\| \nabla p f(p) + b + c \|^2.$$
By taking expectation operator $E[V(p^*)|\mathcal{F}_k]$, since $v_k$ is i.i.d with $E[v_k] = 0$ for each $k \in \mathbb{Z}_{\geq 0}$, and by noticing that $E[v_k^{-1}] = E[v_k]$, it follows that $E[b] = 0$. Next, we show that $E[c_i] = 0$ for $i \in \{1, \ldots, n\}$. We rewrite $c = m(v^T H v) R v$, where $m = \frac{1}{2(\delta_1 + \delta_2)}$, $H \triangleq R^T (\delta_1 \nabla^2 f(p_1) - \delta_2 \nabla^2 f(p_2)) R$, $H = (h_{ij})$, and we use the fact that $v = v^{-1}$. Then,

$$E[c_i] = m E \left[ \sum_{l=1}^n R_{il} v_i \sum_{k=1}^n v_k \sum_{j=1}^n h_{kj} v_j \right]$$

$$= m (q_i + z_i),$$

where $q_i = E[R_{ii} v_i \sum_{k=1}^n v_k \sum_{j=1}^n h_{kj} v_j]$ and $z_i = E[\sum_{l\neq i} R_{il} v_i \sum_{k=1}^n v_k \sum_{j=1}^n h_{kj} v_j]$. Expanding $q_i$,

$q_i = E[R_{ii} v_i (v_i \sum_{j=1}^n h_{ij} v_j + \sum_{k \neq i} v_k \sum_{j=1}^n h_{kj} v_j))]$

$$= R_{ii} E[h_{ii}] v_i^3 + v_i^2 \sum_{j \neq i} h_{ij} v_j + v_i \sum_{k \neq i} v_k^2 h_{kk} + v_i \sum_{j \neq k} v_k \sum_{j \neq k} h_{kj} v_j$

$$= 0,$$

where we have used the assumption that $v_k$ is i.i.d with $E[v_k] = 0$ and the fact that $v_i^3 = v_i$ for $i \in \{1, \ldots, n\}$. Analogous to last procedure, we expand $z_i$,

$z_i = E[\sum_{l\neq i} R_{il} v_i^3 \sum_{j=1}^n h_{ij} v_j + \sum_{l\neq i} R_{il} v_i \sum_{k \neq i} v_k \sum_{j=1}^n h_{kj} v_j]$

$$= E[\sum_{l\neq i} R_{il} v_i^3 h_{ii} + \sum_{l\neq i} R_{il} v_i^2 \sum_{j=1}^n h_{ij} v_j + \sum_{l\neq i} R_{il} v_i \sum_{k \neq i} v_k^2 h_{kk} + \sum_{l\neq i} R_{il} v_i \sum_{k \neq k} v_k \sum_{j \neq k} h_{kj} v_j]$

$$= 0.$$

Thus $E[c] = 0$. Therefore,

$$E[\Delta V | \mathcal{F}_k] \leq -\alpha \rho \| p - p^* \|^2$$

$$+ \alpha^2 (\| \nabla_p f(p) \|^2 + E[\| b \|^2 + \| c \|^2 | \mathcal{F}_k]).$$

(8.12)

Notice that $E[\| c \|^2] \leq \frac{1}{4} \Gamma^2 n^3 (\delta_1 + \delta_2)^2$ and from (8.8) we have

$$E[\| b \|^2 | \mathcal{F}_k] = E[\| v^T R^T \nabla_p f(p) R v^{-1} - \nabla_p f(p) \|^2 | \mathcal{F}_k]$$

$$\leq E[\| v^T R^T \|^2 \| \nabla_p f(p) \|^2 \| R v^{-1} \|^2 + \| \nabla_p f(p) \|^2 | \mathcal{F}_k]$$

$$\leq M^2 (n^2 + 1),$$
where we have used $\|\nabla_p f\| \leq M$. Using above upper bounds and replacing them in (8.12), it follows that

$$E[\Delta V|F_k] \leq -\alpha \rho V(p) + \frac{\alpha^2}{4} (\delta_1 + \delta_2)^2 \Gamma^2 n^3 + \alpha^2 M^2 (n^2 + 2).$$

It follows

$$E[\Delta V|F_k] \leq -\alpha \rho V(p) + \alpha^2 J,$$

where $J = \frac{1}{4} (\delta_1 + \delta_2)^2 \Gamma^2 n^3 + M^2 (n^2 + 2)$. Reorganizing these terms, we have

$$E[V(p^*)|F_k] \leq (1 - \alpha \rho) V(p) + \alpha^2 J.$$

Therefore, by Theorem 2.2 the equilibrium point is MSE-ES. Notice that the max inside the integral in (2.13) simplifies to a point because we do not have a differential inclusion.

The set $O$ given in (8.11) follows by noticing that $E[\Delta V|F_k] \leq 0$ if $\|p - p^*\|^2 \geq O$ and by noticing that $\delta_1 + \delta_2 \leq \bar{\delta}_1 + \bar{\delta}_2$. □

**Remark 8.2.** In the last theorem, knowledge of $\Gamma$ and $M$ are not required to prove convergence. Given that $\mathcal{E}$ is assumed to be compact, existence of $\Gamma$ is guaranteed. Furthermore, since $f$ is assumed locally Lipschitz, then there always exist a finite $M$ such that $\|\nabla_p f(p)\| \leq M$. However, we use those values to characterize the size of the ball where the trajectories converge to in expectation.

If $f$ is nondifferentiable, we are not able to characterize the size of the ball as in Theorem 8.1. However the next result shows practical convergence in probability to $p^*$ and that this ball can be made arbitrarily small by reducing $\alpha$ and $\bar{\delta}_1 + \bar{\delta}_2$ without the assumption on the superquadratic growth condition on $f$. 
Theorem 8.2. (Convergence when \( f \) is nonsmooth): Let Assumption 8.1, on the characteristics of the random input, hold. Assume that \( \alpha \) and \( \bar{\delta}_1 + \bar{\delta}_2 \) are sufficiently small. Moreover, assume that \( f \) is convex, and finite with a unique minimizer \( p^* \). Then, for any initial state \( p_0 \), the solution \( p^* \) of the system (8.3) is MSP-ES.

Proof. The proof follows along the lines of the proof of Theorem 8.1, except that we can not resort to the differentiability properties of \( f \) and we do not use the assumption on its superquadratic growth condition.

Since \( f \) is assumed to be convex and locally Lipschitz, then the set-valued map \( \partial f \) is locally bounded, upper semicontinuous, and takes non-empty, compact, and convex values \([45]\). Using the last fact, the sup in (2.14) can be replaced by a max, and then

\[
 f(p + \delta_1 Rv) = f(p) + \delta_1 v^T \bar{\xi} + o(\delta_1 ||Rv||),
\]

where \( \bar{\xi} = \arg\max_{\xi \in \partial f(p)} \{\xi^T Rv\} \). Similarly,

\[
 f(p - \delta_2 Rv) = f(p) - \delta_2 v^T \bar{\xi} + o(\delta_2 ||Rv||),
\]

where \( \bar{\xi} = \arg\min_{\xi \in \partial f(p)} \{\xi^T Rv\} \). Subtracting (8.14) from (8.13) and dividing the result by \( \delta_1 + \delta_2 \), we have

\[
 \frac{f(p + \delta_1 Rv) - f(p - \delta_2 Rv)}{\delta_1 + \delta_2} = \frac{1}{\delta_1 + \delta_2} (v^T R^T (\delta_1 \bar{\xi} + \delta_2 \bar{\xi}) + o(\delta_1 ||v||) - o(\delta_2 ||v||)),
\]

where we have used the assumption that \( R \) is an orthogonal matrix, then \( o(\delta_i ||Rv||) = o(\delta_i ||v||) \) for \( i \in \{1, 2\} \). Multiplying the last equation by \( Rv^{-1} \), we have

\[
 \frac{f(p + \delta_1 Rv) - f(p - \delta_2 Rv)}{\delta_1 + \delta_2} Rv^{-1} = v^T R^T (\delta_1 \bar{\xi} + \delta_2 \bar{\xi}) Rv^{-1} + \frac{Rv^{-1}}{\delta_1 + \delta_2} (o(\delta_1 ||v||) - o(\delta_2 ||v||)).
\]
We analyze the \(i\)-th component of the first term of the RHS of the last equation to obtain

\[
(v^\top R^\top (\delta_1 \vec{\xi} + \delta_2 \vec{\xi}) \frac{R v^{-1}}{\delta_1 + \delta_2})_i = \sum_{l=1}^n R_{il} \sum_{j=1}^n v_j \sum_{q=1}^n R_{qj} \frac{\delta_1 \vec{\xi}_q + \delta_2 \vec{\xi}_q}{\delta_1 + \delta_2} \\
= \frac{\delta_1 \vec{\xi}_i + \delta_2 \vec{\xi}_i}{\delta_1 + \delta_2} + b_i,
\]

where \(b_i = \sum_{l,j,q,l\neq l} R_{il} R_{qj} \frac{\delta_1 \vec{\xi}_q + \delta_2 \vec{\xi}_q}{\delta_1 + \delta_2}\). Replacing (8.16) in (8.15), it follows that

\[
\frac{f(p + \delta_1 R v) - f(p - \delta_2 R v)}{\delta_1 + \delta_2} R v^{-1} = \frac{\delta_1 \vec{\xi} + \delta_2 \vec{\xi}}{\delta_1 + \delta_2} + b + c,
\]

where \(c = \frac{R v^{-1}}{\delta_1 + \delta_2} (o(\delta_1||v||) - o(\delta_2||v||))\). It follows that

\[
g(p, \delta(p, R v), R, v) = \frac{\delta_1 \vec{\xi} + \delta_2 \vec{\xi}}{\delta_1 + \delta_2} + b + c.
\]

Without loss of generality assume \(\delta_1(p, R v) + \delta_2(p, R v) > 0\). This is the case because, at any time \(k > 0\) for which \(\delta_1(p, R v) + \delta_2(p, R v) = 0\), with probability one, the dynamics in (8.3) will generate a feasible direction in finite time in \(E\) satisfying \(\delta_1(p, R v) + \delta_2(p, R v) > 0\).

Further, without loss of generality assume \(p^* \in E\).

By the non-expansive property of the projection operation, the dynamics in (8.3), and the fact that \(p^* \in E\), we have

\[
||p^+ - p^*||^2 = ||\Pi_E [p - \alpha g(p, \delta(p, R v), R, v)] - p^*||^2 \\
\leq ||p - \alpha g(p, \delta(p, R v), R, v) - p^*||^2 \\
\leq ||p - \alpha (\frac{\delta_1 \vec{\xi} + \delta_2 \vec{\xi}}{\delta_1 + \delta_2} + b + c) - p^*||^2.
\]

It follows

\[
||p^+ - p^*||^2 \leq ||p - p^*||^2 + \alpha^2 ||\frac{\delta_1 \vec{\xi} + \delta_2 \vec{\xi}}{\delta_1 + \delta_2} + b + c||^2 - 2\alpha (\frac{\delta_1 \vec{\xi} + \delta_2 \vec{\xi}}{\delta_1 + \delta_2} + b + c)^\top (p - p^*).
\]
Let $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$, $V = \|p - p^*\|^2$, and define $\Delta V = \|p^+ - p^*\|^2 - \|p - p^*\|^2$. Then we have

$$\Delta V \leq -2\alpha \left( \frac{\delta_1 \xi + \delta_2 \xi}{\delta_1 + \delta_2} + b + c \right)^\top (p - p^*) + \alpha^2 \left[ \frac{\delta_1 \tilde{\xi} + \delta_2 \xi}{\delta_1 + \delta_2} + b + c \right]^2.$$

Let $f_s : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function satisfying the superquadratic growth condition for some $\rho \in \mathbb{R}_{> 0}$ such that $f_s(p^*) = f(p^*)$, $\xi^\top (p - p^*) \leq \xi^\top (p - p^*)$, $\xi \in \partial f(p)$, and $\xi_s \in \partial f_s(p)$ for all $p \in E$. Notice that $f_s$ always can be found since $p^*$ is assumed unique and $E$ is a compact set. Using the last fact, there exists $\rho > 0$ such that $-\xi^\top (p - p^*) \leq -\frac{\rho}{2} \|p - p^*\|^2$. It follows that

$$\Delta V \leq -\alpha \rho \frac{\delta_1 + \delta_2}{\delta_1 + \delta_2} \|p - p^*\|^2 - 2\alpha (b + c)^\top (p - p^*) + \alpha^2 \left[ \frac{\delta_1 \tilde{\xi} + \delta_2 \xi}{\delta_1 + \delta_2} + b + c \right]^2.$$

By noticing that $E[b] = E[c] = 0$, it follows that

$$E[\Delta V|\mathcal{F}_k] \leq -\alpha \rho \|p - p^*\|^2 + \alpha^2 \left[ \frac{\delta_1 \tilde{\xi} + \delta_2 \xi}{\delta_1 + \delta_2} + b + c \right]^2 E[\mathcal{F}_k].$$

From here, the proof follows similar steps as the proof of Theorem 8.1, where we use $\xi$ instead of $\nabla_p f$, and consider $O(1)$ terms instead of the upper bound of the Hessian. □

When the robot moves in a GPS-denied environment, the implementation of the projection operator $\Pi_E$ in (8.3) is challenging. However, if $p^* \in E$ and we use

$$p^+ = p - \alpha(p, \tilde{\delta}(p, Rv), R, v)g(p, \delta(p, Rv), R, v),$$

(8.17)

where $\alpha : \mathbb{R}^n \times \mathbb{R}^2 \times SO(n) \times \mathbb{R}^n \rightarrow \mathbb{R}$ is defined as

$$\alpha(p, \tilde{\delta}(p, Rv), R, v) = \begin{cases} \tilde{\alpha}, & \text{if } p - \tilde{\alpha}g(p, \delta(p, Rv), R, v) \in \mathcal{E}, \\ \text{dist}_{g}(p, \partial \mathcal{E}), & \text{otherwise}, \end{cases}$$

and all other variables defined as in (8.3) for given $\tilde{\alpha} \in \mathbb{R}_{> 0}$. Then, the results of Theorem 8.2 hold as is shown in the following corollary.
Corollary 8.1. Let Assumption 8.1, on the characteristics of the random input, hold. Assume \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex and finite with a unique minimizer \( p^* \in \mathcal{E} \). Assume that \( \bar{\alpha} \) and \( \bar{\delta}_1 + \bar{\delta}_2 \) are sufficiently small. Then, for any initial state \( p_0 \in \mathcal{E} \), the solution \( p^* \) of the system (8.17) is MSP-ES.

Proof. The proof follows by noticing that, whenever the robot detects a boundary, it stops until the next iteration of the dynamics in (8.17). Then, three considerations have to be taken into account along the lines of the proof of Theorem 8.2. First, notice that we can assume \( \delta_1(p, Rv) + \delta_2(p, Rv) > 0 \). This is the case because, at any time \( k > 0 \) for which \( \delta_1(p, Rv) + \delta_2(p, Rv) = 0 \), with probability one, the dynamics in (8.17) will generate a feasible direction in finite time in \( \mathcal{E} \) satisfying \( \delta_1(p, Rv) + \delta_2(p, Rv) > 0 \).

Second, if the robot stops at any time \( k > 0 \) when it is executing (8.17); it is because it finds a boundary, then the positive term in the RHS of the upper bound of \( E[\Delta V | F_k] \) given in the proof of Theorem 8.2 parameterized by \( \alpha \) is less or equal, with probability one, to the same positive term parameterized by \( \bar{\alpha} \). The last fact is explained by noticing that \( \alpha \leq \bar{\alpha} \). Third, when the robot is in \( \partial \mathcal{E} \), with probability one, it moves to a feasible direction in a finite time. \( \square \)
8.4 Proposed algorithm and analysis for environments with obstacles

In this section, we propose an algorithm to find the tower when there are obstacles in the environment and the intensity function $f$ is convex. We introduce some notation used in this section. Let $O$ be a disjoint set of obstacles. Each $O \in O$ is assumed to be closed with a connected piecewise-analytic boundary that is finite in length. Furthermore, the obstacles in $O$ are pairwise disjoint. There may be a countably infinite number of obstacles, but at most a finite number are contained in any fixed disc. The obstacle set $O$ may contain an outer obstacle $O_{\text{outer}}$ that is unbounded; all other obstacles are bounded. We consider the following motion primitives that are used in Algorithm 5:

$U_{\text{fwd}}$ The robot executes Algorithm (8.17), stopping only if it contacts an obstacle.

$U_{\text{fol}}$ The robot travels around the obstacle boundary clockwise, maintaining contact to the right at all times, with a fixed step-size $\bar{\alpha}$, and executing $g(p, \delta, Rv)$ at every time step. The robot stops implementing $U_{\text{fol}}$ only when $g(p, \delta, Rv)$ points to a feasible direction (i.e., a direction that gives a point in $E \setminus O$).

Remark 8.3. The proposed motion primitives are inspired by those in [71]. However, notice that $U_{\text{fol}}$ is different from the $U_{\text{fol}}$ proposed in [71] as the latter is defined to stop at a local minimum. Our motion primitive is defined to stop when the estimated gradient points towards a feasible direction.
Due to the fixed step-size, it could happen that the robot does not detect a point to leave an obstacle when traveling around it. If this happens, the robot will be trapped in the obstacle for all time. To avoid this case, we introduce a constant \( B \in \mathbb{R}_{>0} \). Assume that we know the Lipschitz constant \( L \), i.e., \(|f(p) - f(p')| \leq L\|p - p'\|\), then \( B \triangleq L \). We define \( C \in \mathbb{R}_{>0} \) to indicate an upper bound of \( E[\|b + c\|] \) for all \( p \in \mathcal{E} \setminus O \).

**Assumption 8.2.** (On the stepsize given \( B \)): We assume that \( \bar{\alpha} \) is small enough such that \( \bar{\alpha} \max\{B, C\} \) is strictly less than the minimum distance between two separated obstacles. Moreover, we assume that the robot is able to circumnavigate any obstacle \( O \in \mathcal{O} \) using the constant step-size \( \bar{\alpha} \).

**Algorithm 5** Algorithm for convex intensity functions

1: \( I_L = f(p) \)
2: \textbf{while} not hitting an obstacle \textbf{do}
3: \hspace{1em} execute \( U_{\text{fwd}} \)
4: \textbf{end while}
5: \( I_H = f(p) \)
6: \textbf{while} \( f(p) - \bar{\alpha}B \geq I_H \) \textbf{do}
7: \hspace{1em} execute \( U_{\text{fol}} \)
8: \textbf{end while}
9: go to Step 1

The following lemma is used in the main result of this section, given in Theorem 8.3.

**Lemma 8.3** ([71]). For every obstacle boundary \( \partial O \) and every possible tower location \( \mathbb{R}^2 \setminus O \), there exists at least one intensity local minimum \( p \in \partial O \) for which the topological disc centered at the tower \((0, 0)\) with radius \( \|p\| \) is disjoint from the interior of \( O \).
Theorem 8.3. (Convergence with obstacles in the environment): Let Assumption 8.1, on the characteristics of the random input, and Assumption 8.2, on the step-size given B and C, hold. Assume that \( \bar{\alpha} \) and \( \bar{\delta}_1 + \bar{\delta}_2 \) are sufficiently small. Assume that \( f \) is convex and finite with a unique minimizer \( p^* \). Then, for any initial \( p_0 \in \mathcal{E} \setminus \mathcal{O} \), Algorithm 5 causes the robot to reach a ball of radius \( \bar{\alpha} \) containing the tower with probability one.

Proof. Since \( f \) is assumed to be convex, when the robot moves using the estimated gradient in Step 3, then, in expectation, the distance to the tower is decreased as shown in Theorem 8.2. After the execution of Steps 2 to 4 for the first time, either the ball of size depending on \( \bar{\alpha} \) that contains the tower is located or the robot contacts the boundary of an obstacle \( \partial \mathcal{O} \). Assume the latter, by Step 5, \( IH \) stores the intensity at the boundary point where the robot hits an obstacle. On the other hand, notice that \( IL \) stores the intensity at the boundary point where the robot leaves an obstacle.

The main idea of the proof is that the intensity \( f \) decreases monotonically in expectation with every subsequent execution of Step 3. It might seem that an infinite loop is possible by failure to satisfy the condition of Step 6 or by the motion being blocked by an obstacle boundary. However, this is not the case because the following three facts: (i) let denote by \( P_m \) the set of points \( p \in \partial \mathcal{O} \) where there is a local minimum. By Lemma 8.3, \( P_m \) is not empty; (ii) for all \( p_m \in P_m \), the gradient \( \nabla_p f(p_m) \) (or an element of the generalized gradient, when \( f \) is non-differentiable) points to a feasible direction (i.e., a direction that gives a point in \( \mathcal{E} \setminus \mathcal{O} \)). Then, in expectation \( g(p_m, \delta, Rv) \) points to a feasible direction; and (iii) By the definition of \( B \) and by Step 6, the robot always
detects a ball of radius $\bar{a}B$ where there is local minimum in $\partial O$. This fact is explained as follows. First, note that by Step 6, the robot leaves $O$ whenever $f(p) - \bar{a}B < I_H$ or in other words, when the intensity $f(p)$ (plus some error term given by $aB$) is less than $I_H$ (recall that $I_H$ is the intensity recorded when the robot hit $O$). Second, since the robot moves in discrete steps, then by the definition of $B$, we know that the robot always hits a ball of radius $\bar{a}B$ centered at any local minimum in $\partial O$. Let denote the points $p \in \partial O$ where there is a feasible direction by $P_f$. Notice that $P_m \subset P_f$, which implies that the robot always can find a local minimum and then it always leaves the obstacle boundary and obtain a lower intensity value in expectation.

Notice that when the robot leaves an obstacle it is guaranteed that it will not contact a different obstacle in the next iteration, which is true since it is assumed that all different obstacles are separated by a distance bigger than $\bar{a}\max\{B, C\}$. By Step 6, the leaving point from an obstacle boundary $\partial O$ is closer to the goal than the hitting point. It can happen that an obstacle boundary $\partial O$ is contacted finite number of times, but in expectation, the robot will leave the obstacle, in finite time, and never contact it again. Therefore, when Step 3 is executed, the robot moves away from the obstacle and the intensity continues to decrease in expectation.

8.5 Simulations

Next we show an example that illustrates the response of our proposed Algorithm (8.3) to solve a particular source seeking problem. Figure 8.1 illustrates the evo-
olution of the mobile robot to a source $f = (p_1 - .9)^2 + |p_1 - .9| + (p_2 - 1)^2 + |p_2 - 1|$ with a box constraint $p \in [0, 1]^2$. Notice that the function $f$ is nondifferentiable and strongly convex, then it satisfies the conditions on Theorem 8.2. The tower is located at $p^* = [.9, 1]^T$. This simulation uses $\alpha = \bar{\delta}_1 = \bar{\delta}_2 = .02$ and $R_k = I_n$ for all $k \geq 0$. We have introduced additive gaussian noise in the measurements of the intensity signal of zero mean and variance .0001. The robot starts at $p_0 = [.6, .1]^T$. The robot converges to a ball containing the optimizer $p^*$, which in turn can be made arbitrarily small by decreasing the parameters $\alpha, \bar{\delta}_1, \text{and} \bar{\delta}_2$.

**Figure 8.1**: Evolution of the mobile robot for $f = (p_1 - .9)^2 + |p_1 - .9| + (p_2 - 1)^2 + |p_2 - 1|$ with a box constraint $p \in [0, 1]^2$. The level sets of $f$ are shown in colors and the trajectory of the robot is shown in black.

Next example shows the response of our proposed Algorithm (8.17) to solve a particular source seeking problem with obstacles. Figure 8.1 illustrates the evolution of the mobile robot to a source $f = (p_1 - 5)^2 + |p_1 - 5| + (p_2 - 5)^2 + |p_2 - 5|$ with a box constraint.
$p \in [0, 8]^2$. Notice that the function $f$ is nondifferentiable and strongly convex, then it satisfies the conditions on Theorem 8.2. The tower is located at $p^* = [5, 5]^T$. This simulation uses $\alpha = \delta_1 = \delta_2 = .1$ and $R_k = I_n$ for all $k \geq 0$. The robot starts at $p_0 = [0, 0]^T$. The robot converges to a ball containing the optimizer $p^*$, which in turn can be made arbitrarily small by decreasing the parameters $\alpha, \delta_1$, and $\delta_2$.

![Figure 8.2](image)  

**Figure 8.2:** Evolution of the mobile robot for $f = (p_1 - 5)^2 + |p_1 - 5| + (p_2 - 5)^2 + |p_2 - 5|$ with a box constraint $p \in [0, 8]^2$. The level sets of $f$ are shown in colors and the trajectory of the robot is shown in red, the obstacles in blue.

### 8.6 Summary

Building on the simultaneous perturbation stochastic approximation method, we have introduced a novel algorithm that allows a mobile robot to find the minimizer of an emitting signal. We are able to prove convergence to a ball around the optimizer of the emitting signal, even for non-differentiable signal case and restricting the motion of the
robot to a compact convex set. We have extended this scenario to include obstacles in the environment, which under mild assumptions, we have shown convergence in probability to a ball containing the minimizer.

**Publications associated with this chapter**

This paper contains material that has been published in the following works:

Chapter 9

Conclusions

In this thesis, we present multi-agent algorithms for optimal resource allocation and coordination in large-scale networks. In particular, we propose numerous algorithms to mitigate the spread of a virus in human and computer networks. In addition, we study the clearing problem for an electricity market scenario, and the stochastic source seeking for mobile robots.

Next, I describe two possible extensions to our proposed scenarios.

(i) Link removal: The problem of removing the most important nodes has been introduced [72]. The problem can be formulated as

\[
\begin{align*}
\min_{\kappa} & \quad \lambda_1(K \circ A) \\
\text{subject to} & \\
1_{\kappa}^\top \kappa & = \Gamma \\
\kappa_{ij} & \in \{0, 1\},
\end{align*}
\] (9.1)
where \( [K]_{ij} = 0 \) if \( [A]_{ij} = 0 \) and \( [K]_{ij} \neq \kappa_{ij} \) otherwise. Also, \( \kappa \) is the stacked vector of the nonzero entries \( [K]_{ij} \). We denote by \( \circ \) the entry-wise matrix product. Because the constraint \( \kappa_{ij} \in \{0, 1\} \) is Boolean, last problem is not easily solved by standard convex optimization methods. The problem can be relaxed as

\[
\min_{\kappa} \lambda_1(K \circ A) \\
\text{subject to} \\
1^\top \kappa = \Gamma \\
\kappa_{ij} \in [0, 1].
\]

\[(9.2)\]

In [72], the authors have shown the last problem is convex when \( K \) and \( A \) are symmetric matrices. However the major complication for a distributed implementation is in deciding which node takes care of a bidirectional link. The next idea should help to overcome the issue of dealing with bidirectional links. We already know that \( \lambda_1(\text{diag}(\kappa)A) \), for \( K \) diagonal and \( A \) positive semidefinite, is a convex function. What about if instead of \( A \) we use \( \mu(t) \max_{i,j \in E} a_{ij} I + A \), where \( \mu : \mathbb{R}_{>0} \to [0, 1] \) is a strictly decreasing sequence that starts from 1 and goes asymptotically to zero with rate of convergence slower than the \( p \)-robust box-gradient fairness algorithm. Then, when \( \mu(0) = 1 \), we solve a convex problem. By continuity of \( \lambda_1 \), we expect that the solution when \( \mu(t) \to 0 \) is not too far to the optimal solution when \( \mu \equiv 0 \). The challenge here is to show the error bounds for this approximation.

(ii) Discrete optimization: The idea is to use the stochastic box-constrained
gradient algorithm for discrete optimization. To start with, we can use the idea in [73], where SPSA is used to evaluate the gradient of the cost function only in integer points. This idea allows to have an algorithm in continuous space, which only evaluates the gradient in discrete points. However, to use the stochastic box-constrained gradient algorithm in this problem, we have to generalize the convergence result for non-smooth cost functions, which is left as a future work.
Bibliography


