Elsevier Editorial System(tm) for Theoretical Computer Science
Manuscript Draft

Manuscript Number:

Title: Two Results on Slime Mold Computations

Article Type: Regular Paper (10 - 40 pages)

Section/Category: C - Theory of natural computing

Keywords: physarum dynamics;
differential equations;
natural computing;
linear programming;
approximation algorithms

Corresponding Author: Mr. Pavel Kolev,

Corresponding Author's Institution: Max Planck Institute for Informatics

First Author: Pavel Kolev

Order of Authors: Pavel Kolev; Ruben Becker; Vincenzo Bonifaci; Andreas Karrenbauer; Kurt Mehlhorn

Two Results on Slime Mold Computations

Ruben Becker[1], Vincenzo Bonifaci[2], Andreas Karrenbauer[1], Pavel Kolev*[1], and Kurt Mehlhorn[1]

[1]Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany
{ruben,karrenba,pkolev,mehlhorn}@mpi-inf.mpg.de
[2]Institute for the Analysis of Systems and Informatics, National Research Council of Italy (IASI-CNR), Rome, Italy, vincenzo.bonifaci@iasi.cnr.it

Abstract

We present two results on slime mold computations. In wet-lab experiments (Nature'00) by Nakagaki et al. the slime mold Physarum polycephalum demonstrated its ability to solve shortest path problems. Biologists proposed a mathematical model, a system of differential equations, for the slime's adaption process (J. Theoretical Biology'07). It was shown that the process convergences to the shortest path (J. Theoretical Biology'12) for all graphs. We show that the dynamics actually converges for a much wider class of problems, namely undirected linear programs with a non-negative cost vector.

Combinatorial optimization researchers took the dynamics describing slime behavior as an inspiration for an optimization method and showed that its discretization can ε-approximately solve linear programs with positive cost vector (ITCS'16). Their analysis requires a feasible starting point, a step size depending linearly on ε, and a number of steps with quartic dependence on $\mathrm{opt}/(\varepsilon\Phi)$, where Φ is the difference between the smallest cost of a non-optimal basic feasible solution and the optimal cost (opt).

We give a refined analysis showing that the dynamics initialized with any strongly dominating point converges to the set of optimal solutions. Moreover, we strengthen the convergence rate bounds and prove that the step size is independent of ε, and the number of steps depends logarithmically on $1/\varepsilon$ and quadratically on opt/Φ.

*This work has been funded by the Cluster of Excellence "Multimodal Computing and Interaction" within the Excellence Initiative of the German Federal Government.

Contents

1 Introduction **3**
1.1 The Biologically-Grounded Model . . . 3
1.2 The Biologically-Inspired Model . . . 5

2 Convergence of the Continuous Undirected Dynamics: Simple Instances **8**
2.1 Preliminaries . . . 8
2.2 The Convergence Proof . . . 9

3 Convergence of the Continuous Undirected Dynamics: General Instances **11**
3.1 Existence of a Solution with Domain $[0, \infty)$ . . . 11
3.2 LP Duality . . . 15
3.3 Convergence to Dominance . . . 16
3.4 The Equilibrium Points . . . 17
3.5 Convergence . . . 18
3.6 Details of the Convergence Process . . . 20

4 Improved Convergence Results: Discrete Directed Dynamics **22**
4.1 Overview . . . 22
4.2 Useful Lemmas . . . 23
4.3 Strongly Dominating Capacity Vectors . . . 25
4.4 $x^{(k)}$ is Close to a Non-Negative Kernel-Free Vector . . . 26
4.5 $x^{(k)}$ is ε-Close to an Optimal Solution . . . 28
4.6 Proof of Theorem 4.2 . . . 30
4.7 Preconditioning . . . 31
4.8 A Simple Lower Bound . . . 32

5 Acknowledgements **32**

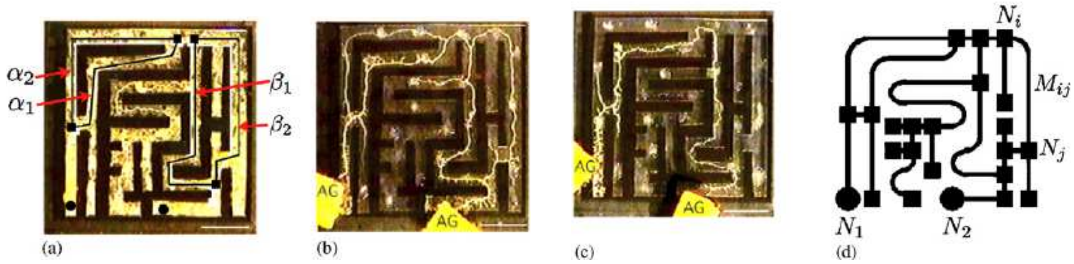

Figure 1: The experiment in [NYT00] (reprinted from there): (a) shows the maze uniformly covered by Physarum; yellow color indicates presence of Physarum. Food (oatmeal) is provided at the locations labeled AG. After a while the mold retracts to the shortest path connecting the food sources as shown in (b) and (c). (d) shows the underlying abstract graph. The video [Phy] shows the experiment.

1 Introduction

We present two results on slime mold computations, one on the biologically-grounded model and one on the biologically-inspired model. The first model was introduced by biologists to capture the slime's apparent ability to compute shortest paths. We show that the dynamics can actually do more. It can solve a wide class of linear programs with nonnegative cost vectors. The latter model was designed as an optimization technique inspired by the former model. We present an improved convergence result for its discretization. The two models are introduced and our results are stated in Sections 1.1 and 1.2 respectively. The results on the former model are shown in Sections 2 and 3, the results on the latter model are shown in Section 4.

1.1 The Biologically-Grounded Model

Physarum polycephalum is a slime mold that apparently is able to solve shortest path problems. Nakagaki, Yamada, and Tóth [NYT00] report about the following experiment; see Figure 1. They built a maze, covered it by pieces of Physarum (the slime can be cut into pieces which will reunite if brought into vicinity), and then fed the slime with oatmeal at two locations. After a few hours the slime retracted to a path that follows the shortest path in the maze connecting the food sources. The authors report that they repeated the experiment with different mazes; in all experiments, Physarum retracted to the shortest path.

The paper [TKN07] proposes a mathematical model for the behavior of the slime and argues extensively that the model is adequate. Physarum is modeled as an electrical network with time varying resistors. We have a simple *undirected* graph $G = (N, E)$ with distinguished nodes s_0 and s_1 modeling the food sources. Each edge $e \in E$ has a positive length c_e and a positive capacity $x_e(t)$; c_e is fixed, but $x_e(t)$ is a function of time. The resistance $r_e(t)$ of e is $r_e(t) = c_e/x_e(t)$. In the electrical network defined by these resistances, a current of value 1 is forced from s_0 to s_1. For an (arbitrarily oriented) edge $e = (u, v)$, let $q_e(t)$ be the resulting current over e. Then, the capacity of e evolves according to the differential equation

$$\dot{x}_e(t) = |q_e(t)| - x_e(t), \tag{1}$$

where $\dot{x}_e$ is the derivative of x_e with respect to time. In equilibrium ($\dot{x}_e = 0$ for all e), the flow through any edge is equal to its capacity. In non-equilibrium, the capacity grows (shrinks) if the absolute value of the flow is larger (smaller) than the capacity. In the sequel, we will mostly drop the argument t as is customary in the treatment of dynamical systems. It is well-known that the electrical flow q is the feasible flow minimizing energy dissipation $\sum_e r_e q_e^2$ (Thomson's principle).

We refer to the dynamics above as *biologically-grounded*, as it was introduced by biologists to model the behavior of a biological system. Miyaji and Ohnishi were the first to analyze convergence for special graphs (parallel links and planar graphs with source and sink on the same face) in [MO08]. In [BMV12] convergence was proven for *all* graphs. We state the result from [BMV12] for the special case that the shortest path is unique.

Theorem 1.1 ([BMV12]). *Assume $c > 0$ and that the undirected shortest path P^* from s_0 to s_1 w.r.t. the cost vector c is unique. Assume $x(0) > 0$. Then $x(t)$ in (1) converges to P^*. Namely, $x_e(t) \to 1$ for $e \in P^*$ and $x_e \to 0$ for $e \notin P^*$ as $t \to \infty$.*

[BMV12] also proves an analogous result for the undirected transportation problem; [Bon13] simplified the argument under additional assumptions. The paper [Bon15] studies a more general dynamics and proves convergence for parallel links.

In this paper, we extend this result to *non-negative undirected linear programs*

$$\min\{c^T x : Af = b,\ |f| \le x\}, \tag{2}$$

where $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$, $x \in \mathbb{R}^m$, $c \in \mathbb{R}^m_{\ge 0}$, and the absolute values are taken componentwise. Undirected LPs can model a wide range of problems, e.g., optimization problems such as shortest path and min-cost flow in undirected graphs, and the Basis Pursuit problem in signal processing [CDS98].

We use n for the number of rows of A and m for the number of columns, since this notation is appropriate when A is the node-edge-incidence matrix of a graph. A vector f is *feasible* if $Af = b$. We assume that the system $Af = b$ has a feasible solution and that there is no non-zero f in the kernel of A with $c_e f_e = 0$ for all e. A vector f lies in the kernel of A if $Af = 0$. The vector q in (1) is now the *minimum energy feasible solution*

$$q(t) = \operatorname*{argmin}_{f\in\mathbb{R}^m} \left\{ \sum_{e:x_e\neq 0} \frac{c_e}{x_e(t)} f_e^2 : Af = b \wedge f_e = 0 \text{ whenever } x_e = 0 \right\}. \tag{3}$$

We remark that q is unique; see Subsection 3.1.1. If A is the incidence matrix of a graph (the column corresponding to an edge e has one entry $+1$, one entry -1 and all other entries are equal to zero), (2) is a transshipment problem with flow sources and sinks encoded by a demand vector b. The condition that there is no solution in the kernel of A with $c_e f_e = 0$ for all e states that every cycle contains at least one edge of positive cost. In that setting, $q(t)$ as defined by (3) coincides with the electrical flow induced by resistors of value $c_e/x_e(t)$. We can now state our first main result.

Theorem 1.2. *Let $c \ge 0$ satisfy $c^T|f| > 0$ for every nonzero f in the kernel of A. Let x^* be an optimum solution of (2) and let $X_\star$ be the set of optimum solutions. Assume $x(0) > 0$. The following holds for the dynamics (1) with q as in (3):*

(i) The solution $x(t)$ exists for all $t \ge 0$.
(ii) The cost $c^T x(t)$ converges to $c^T x^*$ as t goes to infinity.
(iii) The vector $x(t)$ converges to $X_\star$.
(iv) For all e with $c_e > 0$, $x_e(t) - |q_e(t)|$ converges to zero as t goes to infinity.[1] If x^* is unique, $x(t)$ and $q(t)$ converge to x^* as t goes to infinity.

Item (i) was previously shown in [SV16a] for the case of a strictly positive cost vector. The result in [SV16a] is stated for the cost vector $c = \mathbf{1}$. The case of a general positive cost vector reduces to this special case by rescaling the solution vector x. We stress that the dynamics (1) is biologically-grounded. It was proposed to model a biological system and not as an optimization method. Nevertheless, it can solve a large class of non-negative LPs. Table 1 summarizes our first main result and puts it into context.

Sections 2 and 3 are devoted to the proof of our first main theorem. For ease of exposition, we present the proof in two steps. In Section 2, we give a proof under the following simplifying assumptions.

(A) $c > 0$,
(B) The basic feasible solutions of (2) have distinct cost,
(C) We start with a positive vector $x(0) \in X_{\text{dom}} \coloneqq \{\, x \in \mathbb{R}^n : \text{there is a feasible } f \text{ with } |f| \le x \,\}$.

Section 2 generalizes [Bon13]. For the undirected shortest path problem, condition (B) states that all simple undirected source-sink paths have distinct cost and condition (C) states that all source-sink cuts have a capacity of at least one at time zero (and hence at all times). The existence of a solution with domain $[0, \infty)$ was already shown in [SV16a]. We will show that X_{dom} is an invariant set, i.e., the solution stays in X_{dom}

[1] We conjecture that this also holds for the indices e with $c_e = 0$.

Reference	Problem	Existence of Solution	Convergence to OPT	Comments
[MO08]	Undirected Shortest Path	Yes	Yes	parallel edges, planar graphs
[BMV12]	Undirected Shortest Path	Yes	Yes	all graphs
[SV16a]	Undirected Positive LP	Yes	No	$c > 0$
Our Result	Undirected Nonnegative LP	Yes	Yes	1) $c \geq 0$ 2) $\forall v \in \ker(A) : c^T\|v\| > 0$

Table 1: Convergence results for the continuous undirected Physarum dynamics (1).

for all times, and that $E(x) = \sum_e r_e x_e^2 = \sum_e c_e x_e$ is a Lyapunov function for the dynamics (1), i.e., $\dot{E} \leq 0$ and $\dot{E} = 0$ if and only if $\dot{x} = 0$. It follows from general theorems about dynamical systems that the dynamics converges to a fixed point of (1). The fixed points are precisely the vectors $|f|$ where f is a feasible solution of (2). A final argument establishes that the dynamics converges to a fixed point of minimum cost.

In Section 3, we prove the general case of the first main theorem. We assume

(D) $c \geq 0$,

(E) $\text{cost}(z) = c^T|z| > 0$ for every nonzero vector z in the kernel of A,

(F) We start with a positive vector $x(0) > 0$.

Section 3 generalizes [BMV12] in two directions. First, we treat general undirected LPs and not just the undirected shortest path problem, respectively, the transshipment problem. Second, we replace the condition $c > 0$ by the requirement $c \geq 0$ and every nonzero vector in the kernel of A has positive cost. For the undirected shortest path problem, the latter condition states that the underlying undirected graph has no zero-cost cycle. Section 3 is technically considerably more difficult than Section 2. We first establish the existence of a solution with domain $[0, \infty)$. To this end, we derive a closed formula for the minimum energy feasible solution and prove that the mapping $x \mapsto q$ is locally-Lipschitz. Existence of a solution with domain $[0, \infty)$ follows by standard arguments. We then show that X_{dom} is an attractor, i.e., the solution $x(t)$ converges to X_{dom}. We next characterize equilibrium points and exhibit a Lyapunov function. The Lyapunov function is a normalized version of $E(x)$. The normalization factor is equal to the optimal value of the linear program $\max\{\, \alpha : Af = \alpha b,\ |f| \leq x \,\}$ in the variables f and α. Convergence to an equilibrium point follows from the existence of a Lyapunov function. A final argument establishes that the dynamics converges to a fixed point of minimum cost.

1.2 The Biologically-Inspired Model

Ito et al. [IJNT11] initiated the study of the dynamics

$$\dot{x}(t) = q(t) - x(t). \tag{4}$$

We refer to this dynamics as the directed dynamics in contrast to the undirected dynamics (1). The directed dynamics is *biologically-inspired* – the similarity to (1) is the inspiration. It was never claimed to model the behavior of a biological system. Rather, it was introduced as a biologically-inspired optimization method. The work in [IJNT11] shows convergence of this directed dynamics (4) for the directed shortest path problem and [JZ12, SV16c, Bon16] show convergence for general *positive linear programs*, i.e., linear programs with positive cost vector $c > 0$ of the form

$$\min\{c^T x : Ax = b,\ x \geq 0\}. \tag{5}$$

The *discrete versions* of both dynamics define sequences $x^{(t)}$, $t = 0, 1, 2, \ldots$ through

$$x^{(t+1)} = (1 - h^{(t)})x^{(t)} + h^{(t)}q^{(t)} \qquad \text{discrete directed dynamics;} \tag{6}$$

$$x^{(t+1)} = (1 - h^{(t)})x^{(t)} + h^{(t)}|q^{(t)}| \qquad \text{discrete undirected dynamics,} \tag{7}$$

where $h^{(t)}$ is the step size and $q^{(t)}$ is the minimum energy feasible solution as in (3). For the discrete dynamics, we can ask complexity questions. This is particularly relevant for the discrete directed dynamics as it was designed as an biologically-inspired optimization method.

For completeness, we review the state-of-the-art results for the discrete undirected dynamics. For the undirected shortest path problem, the convergence of the discrete undirected dynamics (7) was shown in [BBD+13]. The convergence proof gives an upper bound on the step size and on the number of steps required until an ε-approximation of the optimum is obtained. [SV16b] extends the result to the transshipment problem and [SV16a] further generalizes the result to the case of positive LPs. The paper [SV16b] is related to our first result. It shows convergence of the discretized undirected dynamics (7), we show convergence of the continuous undirected dynamics (1) for a more general cost vector.

We come to the discrete directed dynamics (6). Similarly to the undirected setting, Becchetti et al. [BBD+13] showed the convergence of (6) for the shortest path problem. Straszak and Vishnoi extended the analysis to the transshipment problem [SV16b] and positive LPs [SV16c].

Theorem 1.3. *[SV16c, Theorem 1.3] Let $A \in \mathbb{Z}^{n\times m}$ have full row rank ($n \le m$), $b \in \mathbb{Z}^n$, $c \in \mathbb{Z}^m_{>0}$, and let $D_S := \max\{|\det(M)| : M$ is a square sub-matrix of $A\}$.[2] Suppose the Physarum dynamics (6) is initialized with a feasible point $x^{(0)}$ of (5) such that $M^{-1} \le x^{(0)} \le M$ and $c^T x^{(0)} \le M \cdot \mathrm{opt}$, for some $M \ge 1$. Then, for any $\varepsilon > 0$ and step size $h \le \varepsilon/(\sqrt{6}\|c\|_1 D_S)^2$, after $k = O((\varepsilon h)^{-2}\ln M)$ steps, $x^{(k)}$ is a feasible solution with $c^T x^{(k)} \le (1+\varepsilon)\mathrm{opt}$.*

Theorem 1.3 gives an algorithm that computes a $(1+\varepsilon)$-approximation to the optimal cost of (5). In comparison to [BBD+13, SV16b], it has several shortcomings. First, it requires a feasible starting point. Second, the step size depends linearly on ε. Third, the number of steps required to reach an ε-approximation has a quartic dependence on $\mathrm{opt}/(\varepsilon\Phi)$. In contrast, the analysis in [BBD+13, SV16b] yields a step size independent of ε and a number of steps that depends only logarithmically on $1/\varepsilon$, see Table 2.

We overcome these shortcomings. Before we can state our result, we need some notation. Let $X_\star$ be the set of optimal solutions to (5). The distance of a capacity vector x to $X_\star$ is defined as $\mathrm{dist}(x, X_\star) := \inf\{\|x - x'\|_\infty : x' \in X_\star\}$. Let $\gamma_A := \gcd(\{A_{ij} : A_{ij} \ne 0\}) \in \mathbb{Z}_{>0}$ and

$$D := \max\{|\det(M)| : M \text{ is a square submatrix of } A/\gamma_A \text{ with dimension } n-1 \text{ or } n\}. \tag{8}$$

Let $\mathcal{N}$ be the set of non-optimal basic feasible solution of (5) and

$$\Phi := \min_{g\in\mathcal{N}} c^T g - \mathrm{opt} \ge 1/(D\gamma_A)^2, \tag{9}$$

where the inequality is well known [PS82, Lemma 8.6]. For completeness, we present a proof in Subsection 4.5. Informally, our second main result establishes the following properties of the Physarum dynamics (6):

(i) For any $\varepsilon > 0$ and any strongly dominating starting point[3] $x^{(0)}$, there is a fixed step size $h(x^{(0)})$ such that the Physarum dynamics (6) initialized with $x^{(0)}$ and $h(x^{(0)})$ converges to $X_\star$, i.e., $\mathrm{dist}(x^{(k)}, X_\star) < \varepsilon/(D\gamma_A)$ for large enough k.

(ii) The step size can be chosen *independently* of ε.

(iii) The number of steps k depends *logarithmically* on $1/\varepsilon$ and *quadratically* on opt/Φ.

(iv) The efficiency bounds depend on a *scale-invariant* determinant[4] D.

In Section 4.8, we establish a corresponding lower bound. We show that for the Physarum dynamics (6) to compute a point $x^{(k)}$ such that $\mathrm{dist}(x^{(k)}, X_\star) < \varepsilon$, the number of steps required for computing an ε-approximation has to grow linearly in $\mathrm{opt}/(h\Phi)$ and $\ln(1/\varepsilon)$, i.e. $k \ge \Omega(\mathrm{opt}\cdot(h\Phi)^{-1}\cdot\ln(1/\varepsilon))$. Table 2 puts our results into context.

---

[2] Using Lemma 3.1, the dependence on D_S can be improved to a scale-independent determinant D, defined in (8). For further details, we refer the reader to Subsection 4.2.

[3] We postpone the definition of strongly dominating capacity vector to Section 4.3. Every scaled feasible solution is strongly dominating. In the shortest path problem, a capacity vector x is strongly dominating if every source-sink cut $(S, \overline{S})$ has positive directed capacity, i.e., $\sum_{a\in E(S,\overline{S})} x_a - \sum_{a\in E(\overline{S},S)} x_a > 0$.

[4] Note that $(\gamma_A)^{n-1}D \le D_S \le (\gamma_A)^n D$, and thus D yields an exponential improvement over D_S, whenever $\gamma_A \ge 2$.

Reference	Problem	h step size	k number of steps	Guarantee
[BBD+13]	Shortest Path	indep. of ε	$\text{poly}(m, n, \|c\|_1, \|x^{(0)}\|_1) \cdot \ln(1/\varepsilon)$	$\text{dist}(x^{(k)}, X_\star) < \varepsilon$
[SV16b]	Transshipment	indep. of ε	$\text{poly}(m, n, \|c\|_1, \|b\|_1, \|x^{(0)}\|_1) \cdot \ln(1/\varepsilon)$	$\text{dist}(x^{(k)}, X_\star) < \varepsilon$
[SV16c]	Positive LP	depends on ε	$\text{poly}(\|c\|_1, D_S, \ln \|x^{(0)}\|_1) \cdot 1/(\Phi\varepsilon)^4$	$c^T x^{(k)} \leq (1+\varepsilon)\text{opt}$ $c^T x^{(k)} < \min_{g \in \mathcal{N}} c^T g$
Our Result	Positive LP	indep. of ε	$\text{poly}(\|c\|_1, \|b\|_1, D, \gamma_A, \ln \|x^{(0)}\|_1) \cdot \Phi^{-2} \ln(1/\varepsilon)$	$\text{dist}(x^{(k)}, X_\star) < \frac{\varepsilon}{D\gamma_A}$
Lower Bound	Positive LP	indep. of ε	$\Omega(\text{opt} \cdot (h\Phi)^{-1} \ln(1/\varepsilon))$	$\text{dist}(x^{(k)}, X_\star) < \varepsilon$

Table 2: Convergence results for the discrete directed Physarum dynamics (6).

We state now our second main result for the special case of a feasible starting point, and we provide the full version in Theorem 4.2 which applies for arbitrary strongly dominating starting point, see Section 4. We use the following constants in the statement of the bounds.

(i) $h_0 := c_{\min}/(4D\|c\|_1)$, where $c_{\min} := \min_i\{c_i\}$;

(ii) $\Psi^{(0)} := \max\{mD^2\|b/\gamma_A\|_1, \|x^{(0)}\|_\infty\}$;

(iii) $C_1 := D\|b/\gamma_A\|_1\|c\|_1,\ C_2 := 8^2 m^2 n D^5 \gamma_A^2 \|A\|_\infty \|b\|_1$ and $C_3 := D^3\gamma_A\|b\|_1\|c\|_1$.

Theorem 1.4. *Suppose $A \in \mathbb{Z}^{n\times m}$ has full row rank ($n \leq m$), $b \in \mathbb{Z}^n$, $c \in \mathbb{Z}^m_{>0}$ and $\varepsilon \in (0,1)$. Given a feasible starting point $x^{(0)} > 0$ the Physarum dynamics (6) with step size $h \leq (\Phi/\text{opt}) \cdot h_0^2/2$ outputs for any $k \geq 4C_1/(h\Phi) \cdot \ln(C_2\Psi^{(0)}/(\varepsilon \cdot \min\{1, x^{(0)}_{\min}\}))$ a feasible $x^{(k)} > 0$ such that $\text{dist}(x^{(k)}, X_\star) < \varepsilon/(D\gamma_A)$.*

We stated the bounds on h in terms of the unknown quantities Φ and opt. However, $\Phi/\text{opt} \geq 1/C_3$ by Lemma 3.1 and hence replacing Φ/opt by $1/C_3$ yields constructive bounds for h. Note that the upper bound on the step size does not depend on ε and that the bound on the number of iterations depends *logarithmically* on $1/\varepsilon$ and *quadratically* on opt/Φ.

What can be done if the initial point is not strongly dominating? For the transshipment problem it suffices to add an edge of high capacity and high cost from every source node to every sink node [BBD+13, SV16b]. This will make the instance strongly dominating and will not affect the optimal solution. We generalize this observation to positive linear programs. We add an additional column equal to b and give it sufficiently high capacity and cost. This guarantees that the resulting instance is strongly dominating and the optimal solution remains unaffected. Moreover, our approach generalizes and improves upon [SV16b, Theorem 1.2], see Section 4.7.

Proof Techniques: The crux of the analysis in [IJNT11, BBD+13, SV16b] is to show that for large enough k, $x^{(k)}$ is close to a *non-negative* flow $f^{(k)}$ and then to argue that $f^{(k)}$ is close to an optimal flow $f^\star$. This line of arguments yields a convergence of $x^{(k)}$ to $X_\star$ with a step size h chosen independently of ε.

In Section 4, we extend the preceding approach to positive linear programs, by generalizing the concept of non-negative cycle-free flows to non-negative *feasible kernel-free* vectors (Subsection 4.4). Although, we use the same high level ideas as in [BBD+13, SV16b], we stress that our analysis generalizes all relevant lemmas in [BBD+13, SV16b] and it uses arguments from linear algebra and linear programming duality, instead of combinatorial arguments. Further, our core efficiency bounds (Subsection 4.2) extend [SV16c] and yield a *scale-invariant* determinant dependence of the step size and are applicable for any strongly dominating point (Subsection 4.3).

2 Convergence of the Continuous Undirected Dynamics: Simple Instances

We prove Theorem 1.2 under the following simplifying assumptions:

(A) $c > 0$,

(B) The basic feasible solutions of (2) have distinct cost,

(C) We start with a positive vector $x(0) \in X_{\text{dom}} := \{\, x \in \mathbb{R}^n : \text{there is a feasible } f \text{ with } |f| \le x \,\}$.

This section generalizes [Bon13]. For the undirected shortest path problem, condition (B) states that all simple undirected source-sink paths have distinct cost and condition (C) states that all source-sink cuts have a capacity of at least one at time zero (and hence at all times). The existence of a solution with domain $[0, \infty)$ was already shown in [SV16a].

2.1 Preliminaries

Note that we may assume that A has full row-rank since any equation that is linearly dependent on other equations can be deleted without changing the feasible set. We continue to use n and m for the dimension of A. Thus, A has rank n. We continue by fixing some terms and notation. A *basic feasible solution* of (2) is a pair of vectors x and $f = (f_B, f_N)$, where $f_B = A_B^{-1} b$ and A_B is a square $n \times n$ non-singular sub-matrix of A and $f_N = 0$ is the vector indexed by the coordinates not in B, and $x = |f|$. Since f uniquely determines x, we may drop the latter for the sake of brevity and call f a basic feasible solution of (2). A feasible solution f is *kernel-free* or *non-circulatory* if it is contained in the convex hull of the basic feasible solutions.[5] We say that a vector f' is *sign-compatible* with a vector f (of the same dimension) or f-sign-compatible if $f'_e \neq 0$ implies $f'_e f_e > 0$. In particular, $\text{supp}(f') \subseteq \text{supp}(f)$. For a given capacity vector x and a vector $f \in \mathbb{R}^m$ with $\text{supp}(f) \subseteq \text{supp}(x)$, we use $E(f) = \sum_e (c_e/x_e) f_e^2$ to denote the *energy* of f. The energy of f is infinite, if $\text{supp}(f) \not\subseteq \text{supp}(x)$. We use $\text{cost}(f) = \sum_e c_e |f_e| = c^T |f|$ to denote the *cost* of f. Note that $E(x) = \sum_e (c_e/x_e) x_e^2 = \sum_e c_e x_e = \text{cost}(x)$. We define the constants $c_{\max} = \|c\|_\infty$ and $c_{\min} = \min_{e: c_e > 0} c_e$.

We use the following corollary of the finite basis theorem for polyhedra.

Lemma 2.1. *Let f be a feasible solution of (2). Then f is the sum of a convex combination of at most m basic feasible solutions plus a vector in the kernel of A. Moreover, all elements in this representation are sign-compatible with f.*

Proof. We may assume $f \ge 0$. Otherwise, we flip the sign of the appropriate columns of A. Thus, the system $Af = b,\ f \ge 0$ is feasible and f is the sum of a convex combination of at most m basic feasible solutions plus a vector in the kernel of A by the finite basis theorem [Sch99, Corollary 7.1b]. By definition, the elements in this representation are non-negative vectors and hence sign-compatible with f. □

Fact 2.2 (Grönwall's Lemma). *Let $A, B, \alpha, \beta \in \mathbb{R}$, $\alpha \neq 0$, $\beta \neq 0$, and let g be a continuous differentiable function on $[0, \infty)$. If $A + \alpha g(t) \le \dot{g}(t) \le B + \beta g(t)$ for all $t \ge 0$, then $-A/\alpha + (g(0) + A/\alpha) e^{\alpha t} \le g(t) \le -B/\beta + (g(0) + B/\beta) e^{\beta t}$ for all $t \ge 0$.*

Proof. We show the upper bound. Assume first that $B = 0$. Then

$$\frac{d}{dt} \frac{g}{e^{\beta t}} = \frac{\dot{g} e^{\beta t} - \beta g e^{\beta t}}{e^{2\beta t}} \le 0 \quad \text{implies} \quad \frac{g(t)}{e^{\beta t}} \le \frac{g(0)}{e^{\beta 0}} = g(0).$$

If $B \neq 0$, define $h(t) = g(t) + B/\beta$. Then

$$\dot{h} = \dot{g} \le B + \beta g = B + \beta(h - B/\beta) = \beta h$$

and hence $h(t) \le h(0) e^{\beta t}$. Therefore $g(t) \le -B/\beta + (g(0) + B/\beta) e^{\beta t}$. □

[5] For the undirected shortest path problem, we drop the equation corresponding to the sink. Then b becomes the negative indicator vector corresponding to the source node. Note that n is one less than the number of nodes of the graph. The basic feasible solutions are the simple undirected source-sink paths. A circulatory solution contains a cycle on which there is flow.

An immediate consequence of Grönwall's Lemma is that the undirected Physarum dynamics (1) initialized with any positive starting vector $x(0)$, generates a trajectory $\{x(t)\}_{t\geq 0}$ such that each time state $x(t)$ is a positive vector. Indeed, since $\dot{x}_e = |q_e| - x_e \geq -x_e$, we have $x_e(t) \geq x_e(0) \cdot \exp\{-t\}$ for every index e with $x_e(0) > 0$ and every time t. Further, by (1) and (3), it holds for indices e with $x_e(0) = 0$ that $x_e(t) = 0$ for every time t. Hence, the trajectory $\{x(t)\}_{t\geq 0}$ has a time-invariant support.

Fact 2.3 ([JZ12])**.** *Let $R = \mathrm{diag}(c_e/x_e)$. Then $q = R^{-1}A^T p$, where $p = (AR^{-1}A^T)^{-1}b$.*

Proof. q minimizes $\sum_e r_e q_e^2$ subject to $Aq = b$. The KKT conditions imply the existence of a vector p such that $Rq = A^T p$. Substituting into $Aq = b$ yields $p = (AR^{-1}A^T)^{-1}b$. □

Lemma 2.4. *X_{dom} is an invariant set, i.e., if $x(0) \in X_{\mathrm{dom}}$ then $x(t) \in X_{\mathrm{dom}}$ for all t.*

Proof. Let $q(t)$ be the minimum energy feasible solution with respect to $R(t) = \mathrm{diag}(c_e/x_e(t))$, and let $f(t)$ be such that $f(0)$ is feasible, $|f(0)| \leq x(0)$, and $\dot{f}(t) = q(t) - f(t)$. Then $\frac{d}{dt}(Af - b) = A(q - f) = b - Af$ and hence $Af(t) - b = (Af(0) - b)e^{-t} = 0$. Thus $f(t)$ is feasible for all t. Moreover,

$$\frac{d}{dt}(f - x) = \dot{f} - \dot{x} = q - f - (|q| - x) = q - |q| - (f - x) \leq -(f - x).$$

Thus $f(t) - x(t) \leq (f(0) - x(0))e^{-t} \leq 0$ by Gronwall's Lemma applied with $g(t) = f(t) - x(t)$ and $\beta = -1$, and hence $f(t) \leq x(t)$ for all t. Similarly,

$$\frac{d}{dt}(f + x) = \dot{f} + \dot{x} = q - f + (|q| - x) = q + |q| - (f + x) \geq -(f + x).$$

Thus $f(t) + x(t) \geq (f(0) + x(0))e^{-t} \geq 0$ by Grönwall's Lemma applied with $g(t) = f(t) + x(t)$ and $\alpha = -1$ and $A = 0$, and hence $f(t) \geq -x(t)$ for all t.

We conclude that $|f(t)| \leq x(t)$ for all t. Thus, $x(t) \in X_{\mathrm{dom}}$ for all t. □

2.2 The Convergence Proof

We will first characterize the equilibrium points. They are precisely the points $|f|$, where f is a basic feasible solution; the proof uses property (B). We then show that $E(x)$ is a Lyapunov function for (1), in particular, $\dot{E} \leq 0$ and $\dot{E} = 0$ if and only if x is an equilibrium point. For this argument, we need that the energy of q is at most the energy of x with equality if and only if x is an equilibrium point. This proof uses (A) and (C). It follows from the general theory of dynamical systems that $x(t)$ approaches an equilibrium point. Finally, we show that convergence to a non-optimal equilibrium is impossible.

Lemma 2.5 (Generalization of Lemma 2.3 in [Bon13])**.** *Assume (A) to (C). If f is a basic feasible solution of (2), then $x = |f|$ is an equilibrium point. Conversely, if x is an equilibrium point, then $x = |f|$ for some basic feasible solution f.*

Proof. Let f be a basic feasible solution, let $x = |f|$, and let q be the minimum energy feasible solution with respect to the resistances c_e/x_e. We have $Aq = b$ and $\mathrm{supp}(q) \subseteq \mathrm{supp}(x)$ by definition of q. Since f is a basic feasible solution there is a subset B of size n of the columns of A such that A_B is non-singular and $f = (A_B^{-1}b, 0)$. Since $\mathrm{supp}(q) \subseteq \mathrm{supp}(x) \subseteq B$, we have $q = (q_B, 0)$ for some vector q_B. Thus, $b = Aq = A_B q_B$ and hence $q_B = f_B$. Therefore $\dot{x} = |q| - x = 0$ and x is an equilibrium point.

Conversely, if x is an equilibrium point, $|q_e| = x_e$ for every e. By changing the signs of some columns of A, we may assume $q \geq 0$. Then $q = x$. Since $q_e = x_e/c_e A_e^T p$ where A_e is the e-th column of A by Lemma 2.3, we have $c_e = A_e^T p$, whenever $x_e > 0$. By Lemma 2.1, q is a convex combination of basic feasible solutions and a vector in the kernel of A that are sign-compatible with q. The vector in the kernel must be zero as q is a minimum energy feasible solution. For any basic feasible solution z contributing to q, we have $\mathrm{supp}(z) \subseteq \mathrm{supp}(x)$. Summing over the $e \in \mathrm{supp}(z)$, we obtain $\mathrm{cost}(z) = \sum_{e\in\mathrm{supp}(z)} c_e z_e = \sum_{e\in\mathrm{supp}(z)} z_e A_e^T p = b^T p$. Thus, the convex combination involves only a single basic feasible solution by assumption (B) and hence x is a basic feasible solution. □

The vector $x(t)$ dominates a feasible solution at all times. Since $q(t)$ is the minimum energy feasible solution at time t, this implies $E(q(t)) \leq E(x(t))$ at all times. A further argument shows that we have equality if and only if $x = |q|$.

Lemma 2.6 (Generalization of Lemma 3.1 in [Bon13])**.** *Assume* (A) *to* (C)*. At all times,* $E(q) \leq E(x)$*. If* $E(q) = E(x)$*, then* $x = |q|$*.*

Proof. Recall that $x(t) \in X_{\text{dom}}$ for all t. Thus, at all times, there is a feasible f such that $|f| \leq x$. Since q is a minimum energy feasible solution, we have

$$E(q) \leq E(f) \leq E(x).$$

If $E(q) = E(x)$ then $E(q) = E(f)$ and hence $q = f$ since the minimum energy feasible solution is unique. Also, $|f| = x$ since $|f| \leq x$ and $|f_e| < x_e$ for some e implies $E(f) < E(x)$. The last conclusion uses $c > 0$. □

Lyapunov functions are the main tool for proving convergence of dynamical systems. We show that $E(x)$ is a Lyapunov function for (1).

Lemma 2.7 (Generalization of Lemma 3.2 in [Bon13])**.** *Assume* (A) *to* (C)*.* $E(x)$ *is a Lyapunov function for* (1)*, i.e., it is continuous as a function of* x*,* $E(x) \geq 0$*,* $\dot{E}(x) \leq 0$ *and* $\dot{E}(x) = 0$ *if and only if* $\dot{x} = 0$*.*

Proof. E is clearly continuous and non-negative. Recall that $E(x) = \text{cost}(x)$. Let R be the diagonal matrix with entries c_e/x_e. Then

$$\begin{aligned}
\frac{d}{dt}\text{cost}(x) &= c^T(|q| - x) && \text{by (1)}\\
&= x^T R|q| - x^T Rx && \text{since } c = Rx\\
&= x^T R^{1/2} R^{1/2}|q| - x^T Rx &&\\
&\leq (q^T Rq)^{1/2}(x^T Rx)^{1/2} - x^T Rx && \text{by Cauchy-Schwarz}\\
&\leq (x^T Rx)^{1/2}(x^T Rx)^{1/2} - x^T Rx && \text{by Lemma 2.6}\\
&= 0.
\end{aligned}$$

Observe that $\frac{d}{dt}\text{cost}(x) = 0$ implies that both inequalities above are equalities. This is only possible if the vectors $|q|$ and x are parallel and $E(q) = E(x)$. Thus, $x = |q|$ by Lemma 2.6. □

It follows now from the general theory of dynamical systems that $x(t)$ converges to an equilibrium point.

Corollary 2.8 (Generalization of Corollary 3.3. in [Bon13].)**.** *Assume* (A) *to* (C)*. As* $t \to \infty$*,* $x(t)$ *approaches an equilibrium point and* $c^T x(t)$ *approaches the cost of the corresponding basic feasible solution.*

Proof. The proof in [Bon13] carries over. We include it for completeness. The existence of a Lyapunov function E implies by [LaS76, Corollary 2.6.5] that $x(t)$ approaches the set $\left\{ x \in \mathbb{R}^m_{\geq 0} : \dot{E} = 0 \right\}$, which by Lemma 2.7 is the same as the set $\left\{ x \in \mathbb{R}^m_{\geq 0} : \dot{x} = 0 \right\}$. Since this set consists of isolated points (Lemma 2.5), $x(t)$ must approach one of those points, say the point x_0. When $x = x_0$, one has $E(q) = E(x) = \text{cost}(x) = c^T x$. □

It remains to exclude that $x(t)$ converges to a nonoptimal equilibrium point.

Theorem 2.9 (Generalization of Theorem 3.4 in [Bon13])**.** *Assume* (A) *to* (C)*. As* $t \to \infty$*,* $c^T x(t)$ *converges to the cost of the optimal solution and* $x(t)$ *converges to the optimal solution.*

Proof. By the corollary, it suffices to prove the second part of the claim. For the second part, assume that $x(t)$ converges to a non-optimal solution z. Let x^* be the optimal solution and let $W = \sum_e x^*_e c_e \ln x_e$. Let $\delta = (\text{cost}(z) - \text{cost}(x^*))/2$. Note that for all sufficiently large t, we have $E(q(t)) \geq \text{cost}(z) - \delta \geq \text{cost}(x^*) + \delta$. Further, by definition $q_e = (x_e/c_e)A_e^T p$ and thus

$$\dot{W} = \sum_e x^*_e c_e \frac{|q_e| - x_e}{x_e} = \sum_e x^*_e |A_e^T p| - \text{cost}(x^*) \geq \sum_e x^*_e A_e^T p - \text{cost}(x^*) \geq \delta,$$

where the last inequality follows by $\sum_e x^*_e A_e^T p = b^T p = E(q) \geq \text{cost}(x^*) + \delta$. Hence $W \to \infty$, a contradiction to the fact that x is bounded. □

3 Convergence of the Continuous Undirected Dynamics: General Instances

We now prove the general case. We assume

(D) $c \geq 0$,

(E) $\text{cost}(z) > 0$ for every nonzero vector z in the kernel of A,

(F) We start with a positive vector $x(0) > 0$.

In this section, we generalize [BMV12] in two directions. First, we treat general undirected LPs and not just the undirected shortest path problem, respectively, the transshipment problem. Second, we substitute the condition $c > 0$ with $c \geq 0$ and every nonzero vector in the kernel of A has positive cost. For the undirected shortest path problem, the latter condition states that the underlying undirected graph has no zero-cost cycle.

3.1 Existence of a Solution with Domain $[0, \infty)$

In this section we show that a solution $x(t)$ to (1) has domain $[0, \infty)$. We first derive an explicit formula for the minimum energy feasible solution q and then show that the mapping $x \mapsto q$ is Lipschitz continuous; this implies existence of a solution with domain $[0, \infty)$ by standard arguments.

3.1.1 The Minimum Energy Solution

Recall that for $\gamma_A = \gcd(\{A_{ij} : A_{ij} \neq 0\}) \in \mathbb{Z}_{>0}$, we defined by

$$D = \max\left\{|\det(M)| \,:\, M \text{ is a square submatrix of } A/\gamma_A \text{ with dimension } n-1 \text{ or } n\right\}.$$

We derive now properties of the minimum energy solution. In particular, if every nonzero vector in the kernel of A has positive cost,

(i) the minimum energy feasible solution is kernel-free and unique (Lemma 3.2),

(ii) $|q_e| \leq D\|b/\gamma_A\|_1$ for every $e \in [m]$ (Lemma 3.3),

(iii) q is defined by (12) (Lemma 3.4), and

(iv) $E(q) = b^T p$, where p is defined by (12) (Lemma 3.5).

We note that for positive cost vector $c > 0$, these results are known.

We proceed by establishing some useful properties on basic feasible solutions.

Lemma 3.1. *Suppose $A \in \mathbb{Z}^{n \times m}$ is an integral matrix, and $b \in \mathbb{Z}^n$ is an integral vector. Then, for any basic feasible solutions f with $Af = b$ and $f \geq 0$, it holds that $\|f\|_\infty \leq D\|b/\gamma_A\|_1$ and $f_j \neq 0$ implies $|f_j| \geq 1/(D\gamma_A)$.*

Proof. Since f is a basic feasible solution, it has the form $f = (f_B, 0)$ such that $A_B \cdot f_B = b$ where $A_B \in \mathbb{Z}^{n \times n}$ is an invertible submatrix of A. We write $M_{-i,-j}$ to denote the matrix M with deleted i-th row and j-th column. Let Q_j be the matrix formed by replacing the j-th column of A_B by the column vector b. Then, using the fact that $\det(tA) = t^n \det(A)$ for every $A \in \mathbb{R}^{n \times n}$ and $t \in \mathbb{Z}$, the Cramer's rule yields

$$|f_B(j)| = \left|\frac{\det(Q_j)}{\det(A_B)}\right| = \frac{1}{\gamma_A}\left|\sum_{k=1}^{n} \frac{(-1)^{j+k} \cdot b_k \cdot \det\left(\gamma_A^{-1}[A_B]_{-k,-j}\right)}{\det\left(\gamma_A^{-1} A_B\right)}\right|$$

By the choice of γ_A, the values $\det(\gamma_A^{-1} A_B)$ and $\det(\gamma_A^{-1}[A_B]_{-k,-j})$ are integral for all k, it follows that

$$|f_B(j)| \leq D\|b/\gamma_A\|_1 \qquad \text{and} \qquad f_B(j) \neq 0 \quad \Longrightarrow \quad \frac{1}{D\gamma_A} \leq |f_B(j)|. \qquad \square$$

Lemma 3.2. *If every nonzero vector in the kernel of A has positive cost, the minimum energy feasible solution is kernel-free and unique.*

Proof. Let q be a minimum energy feasible solution. Since q is feasible, it can be written as $q_n + q_r$, where q_n is a convex combination of basic feasible solutions and q_r lies in the kernel of A. Moreover, all elements in this representation are sign-compatible with q by Lemma 2.1. If $q_r \neq 0$, the vector $q - q_r$ is feasible and has smaller energy, a contradiction. Thus $q_r = 0$.

We next prove uniqueness. Assume for the sake of a contradiction that there are two distinct minimum energy feasible solutions $q^{(1)}$ and $q^{(2)}$. We show that the solution $(q^{(1)} + q^{(2)})/2$ uses less energy than $q^{(1)}$ and $q^{(2)}$. Since $h \mapsto h^2$ is a strictly convex function from $\mathbb{R}$ to $\mathbb{R}$, the average of the two solutions will be better than either solution if there is an index e with $r_e > 0$ and $q_e^{(1)} \neq q_e^{(2)}$. The difference $z = q^{(1)} - q^{(2)}$ lies in the kernel of A and hence $\text{cost}(z) = \sum_e c_e|z_e| > 0$. Thus there is an e with $c_e > 0$ and $z_e \neq 0$. We have now shown uniqueness. □

Lemma 3.3. *Assume that every nonzero vector in the kernel of A has positive cost. Let q be the minimum energy feasible solution. Then $|q_e| \le D\|b/\gamma_A\|_1$ for every e.*

Proof. Since q is a convex combination of basic feasible solutions, $|q_e| \le \max_z |z_e|$ where z ranges over basic feasible solutions of the form $(z_B, 0)$, where $z_B = A_B^{-1}b$ and $A_B \in \mathbb{R}^{n\times n}$ is a non-singular submatrix of A. Thus, by Lemma 3.1 every component of z is bounded by $D\|b/\gamma_A\|_1$. □

In [SV16c], the bound $|q_e| \le D^2 m\|b\|_1$ was shown. We will now derive explicit formulae for the minimum energy solution q. We will express q in terms of a vector $p \in \mathbb{R}^n$, which we refer to as the *potential*, by analogy with the network setting, in which p can be interpreted as the electric potential of the nodes. The energy of the minimum energy solution is equal to $b^T p$. We also derive a local Lipschitz condition for the mapping from x to q. Note that for $c > 0$ these facts are well-known. Let us split the column indices $[m]$ of A into

$$P := \{\, e \in [m] : c_e > 0 \,\} \quad \text{and} \quad Z := \{\, e \in [m] : c_e = 0 \,\}. \tag{10}$$

Lemma 3.4. *Assume that every nonzero vector in the kernel of A has positive cost. Let $r_e = c_e/x_e$ and let R denote the corresponding diagonal matrix. Let us split A into A_P and A_Z, and q into q_P and q_Z. Since A_Z has linearly independent columns, we may assume that the first $|Z|$ rows of A_Z form a square non-singular matrix. We can thus write $A = \left[\begin{smallmatrix} A'_P & A'_Z \\ A''_P & A''_Z \end{smallmatrix}\right]$ with invertible A'_Z. Then the minimum energy solution satisfies*

$$\begin{bmatrix} A'_P & A'_Z \\ A''_P & A''_Z \end{bmatrix} \begin{bmatrix} q_P \\ q_Z \end{bmatrix} = \begin{bmatrix} b' \\ b'' \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} R_P & \\ & 0 \end{bmatrix} \begin{bmatrix} q_P \\ q_Z \end{bmatrix} = \begin{bmatrix} {A'_P}^T & {A''_P}^T \\ {A'_Z}^T & {A''_Z}^T \end{bmatrix} \begin{bmatrix} p' \\ p'' \end{bmatrix} \tag{11}$$

for some vector $p = \left[\begin{smallmatrix} p' \\ p'' \end{smallmatrix}\right]$; here p' has dimension $|Z|$. The equation system (11) has a unique solution given by

$$\begin{bmatrix} q_Z \\ q_P \end{bmatrix} = \begin{bmatrix} [A'_Z]^{-1}(b' - A'_P q_P) \\ R_P^{-1} A_P^T p \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} p' \\ p'' \end{bmatrix} = \begin{bmatrix} -[[A'_Z]^T]^{-1}[A''_Z]^T p'' \\ MR^{-1}M^T(b'' - A''_Z[A'_Z]^{-1}b') \end{bmatrix}, \tag{12}$$

where $M = A''_P - A''_Z[A'_Z]^{-1}A'_P$ is the Schur complement of the block A'_Z of the matrix A.

Proof. q minimizes $E(f) = f^T R f$ among all solutions of $Af = b$. The KKT conditions state that q must satisfy $Rq = A^T p$ for some p. Note that $2Rf$ is the gradient of the energy $E(f)$ with respect to f and that the $-A^T p$ is the gradient of $p^T(b - Af)$ with respect to f. We may absorb the factor -2 in p. Thus q satisfies (11).

We show next that the linear system (11) has a unique solution. The top $|Z|$ rows of the left system in (11) give

$$q_Z = [A'_Z]^{-1}(b' - A'_P q_P). \tag{13}$$

Substituting this expression for q_Z into the bottom $n - |Z|$ rows of the left system in (11) yields

$$Mq_P = b'' - A''_Z[A'_Z]^{-1}b'.$$

From the top $|P|$ rows of the right system in (11) we infer $q_P = R_P^{-1} A_P^T \cdot p$. Thus

$$M R_P^{-1} A_P^T \cdot p = b'' - A_Z'' [A_Z']^{-1} \cdot b'. \tag{14}$$

The bottom $n - |Z|$ rows of the right system in (11) yield $0 = A_Z^T p = [A_Z']^T p' + [A_Z'']^T p''$ and hence

$$p' = -[[A_Z']^T]^{-1} [A_Z'']^T p''. \tag{15}$$

Substituting (15) into (14) yields

$$\begin{aligned} b'' - A_Z''[A_Z']^{-1} b' &= M R_P^{-1} \left([A_P']^T p' + [A_P'']^T p'' \right) \\ &= M R_P^{-1} \left([A_P'']^T - [A_P']^T [[A_Z']^T]^{-1} [A_Z'']^T \right) p'' \\ &= M R_P^{-1} M^T p''. \end{aligned} \tag{16}$$

It remains to show that the matrix $M R_P^{-1} M^T$ is non-singular. We first observe that the rows of M are linearly independent. Consider the left system in (11). Multiplying the first $|Z|$ rows by $[A_Z']^{-1}$ and then subtracting A_Z'' times the resulting rows from the last $n - |Z|$ rows turns A into the matrix $Q = \left[\begin{smallmatrix} [A_Z']^{-1} A_P' & I \\ M & 0 \end{smallmatrix}\right]$. By assumption, A has independent rows. Moreover, the preceding operations guarantee that $\text{rank}(A) = \text{rank}(Q)$. Therefore, M has independent rows. Since R_P^{-1} is a positive diagonal matrix, $R_P^{-1/2}$ exists and is a positive diagonal matrix. Let z be an arbitrary nonzero vector of dimension $|P|$. Then $z^T M R_P^{-1} M^T z = (R_P^{-1/2} M^T z)^T (R_P^{-1/2} M^T z) > 0$ and hence $M R_P^{-1} M^T$ is non-singular. It is even positive semi-definite.

There is a shorter proof that the system (11) has a unique solution. However, the argument does not give an explicit expression for the solution. In the case of a convex objective function and affine constraints, the KKT conditions are sufficient for being a global minimum. Thus any solution to (11) is a global optimum. We have already shown in Lemma 3.2 that the global minimum is unique. □

We next observe that the energy of q can be expressed in terms of the potential.

Lemma 3.5. *Let q be the minimum energy feasible solution and let f be any feasible solution. Then $E(q) = b^T p = f^T A^T p$.*

Proof. As in the proof of Lemma 3.4, we split q into q_P and q_Z, R into R_P and R_Z, and A into A_P and A_Z. Then

$$\begin{aligned} E(q) &= q_P^T R_P q_P && \text{by the definition of } E(q) \text{ and since } R_Z = 0 \\ &= p^T A_P q_p && \text{by the right system in (11)} \\ &= p^T (b - A_Z q_Z) && \text{by the left system in (11)} \\ &= b^T p && \text{by the right system in (11).} \end{aligned}$$

For any feasible solution f, we have $f^T A^T p = b^T p$. □

3.1.2 The Mapping $x \mapsto q$ is Locally-Lipschitz

We show that the mapping $x \mapsto q$ is Lipschitz continuous; this implies existence of a solution $x(t)$ with domain $[0, \infty)$ by standard arguments. Our analysis builds upon Cramer's rule and the Cauchy-Binet formula. The Cauchy-Binet formula extends Kirchhoff's spanning tree theorem which was used in [BMV12] for the analysis of the undirected shortest path problem.

Lemma 3.6. *Assume $c \geq 0$, no non-zero vector in the kernel of A has cost zero, and that A, b, and c are integral. Let $\alpha, \beta > 0$. For any two vectors x and $\tilde{x}$ in $\mathbb{R}^m$ with $\alpha \mathbf{1} \leq x, \tilde{x} \leq \beta \mathbf{1}$, define $\gamma := 2m^n (\beta/\alpha)^n c_{\max}^n D^2 \|b/\gamma_A\|_1$. Then $\left| |q_e(x)| - |q_e(\tilde{x})| \right| \leq \gamma \|x - \tilde{x}\|_\infty$ for every $e \in [m]$.*

Proof. First assume that $c > 0$. By Cramer's rule

$$(AR^{-1}A^T)^{-1} = \frac{1}{\det(AR^{-1}A^T)}((-1)^{i+j}\det(M_{-j,-i}))_{ij},$$

where $M_{-i,-j}$ is obtained from $AR^{-1}A^T$ by deleting the i-th row and the j-th column. For a subset S of $[m]$ and an index $i \in [n]$, let A_S be the $n \times |S|$ matrix consisting of the columns selected by S and let $A_{-i,S}$ be the matrix obtained from A_S by deleting row i. If D is a diagonal matrix of size m, then $(AD)_S = A_S D_S$. The Cauchy-Binet theorem expresses the determinant of a product of two matrices (not necessarily square) as a sum of determinants of square matrices. It yields

$$\begin{aligned}\det(AR^{-1}A^T) &= \sum_{S\subseteq[m];\ |S|=n} (\det((AR^{-1/2})_S))^2 \\ &= \sum_{S\subseteq[m];\ |S|=n} (\prod_{e\in S} x_e/c_e)\cdot(\det A_S)^2.\end{aligned}$$

Similarly,

$$\det(AR^{-1}A^T)_{-i,-j} = \sum_{S\subseteq[m];\ |S|=n-1} (\prod_{e\in S} x_e/c_e)\cdot(\det A_{-i,S}\cdot\det A_{-j,S}).$$

Using $p = (AR^{-1}A^T)^{-1}b$, we obtain

$$p_i = \frac{\sum_{j\in[n]}(-1)^{i+j}\sum_{S\subseteq[m];\ |S|=n-1}(\prod_{e\in S} x_e/c_e)\cdot(\det A_{-i,S}\cdot\det A_{-j,S})b_j}{\sum_{S\subseteq[m];\ |S|=n}(\prod_{e\in S} x_e/c_e)\cdot(\det A_S)^2}. \tag{17}$$

Substituting into $q = R^{-1}A^Tp$ yields

$$\begin{aligned}q_e &= \frac{x_e}{c_e}A_e^Tp \\ &= \frac{x_e}{c_e}\sum_i A_{i,e}\cdot\frac{\sum_{j\in[n]}(-1)^{i+j+2n}\sum_{S\subseteq[m];\ |S|=n-1}(\prod_{e'\in S} x_{e'}/c_{e'})\cdot(\det A_{-i,S}\cdot\det A_{-j,S})b_j}{\sum_{S\subseteq[m];\ |S|=n}(\prod_{e'\in S} x_{e'}/c_{e'})\cdot(\det A_S)^2} \\ &= \frac{\sum_{S\subseteq[m];\ |S|=n-1}\left((\prod_{e'\in S\cup e} x_{e'}/c_{e'})\cdot\sum_{i\in[n]}(-1)^{i+n}A_{i,e}\det A_{-i,S}\cdot\sum_{j\in[n]}(-1)^{j+n}b_j\det A_{-j,S}\right)}{\sum_{S\subseteq[m];\ |S|=n}(\prod_{e'\in S} x_{e'}/c_{e'})\cdot(\det A_S)^2} \\ &= \frac{\sum_{S\subseteq[m];\ |S|=n-1}(\prod_{e'\in S\cup e} x_{e'}/c_{e'})\cdot\det(A_S|A_e)\cdot\det(A_S|b)}{\sum_{S\subseteq[m];\ |S|=n}(\prod_{e'\in S} x_{e'}/c_{e'})\cdot(\det A_S)^2},\end{aligned} \tag{18}$$

where $(A_S|A_e)$, respectively $(A_S|b)$, denotes the $n \times n$ matrix whose columns are selected from A by S and whose last column is equal to A_e, respectively b.

We are now ready to estimate the derivative $\partial q_e/\partial x_i$. Assume first that $e \neq i$. By the above, $q_e = \frac{x_e}{c_e}\frac{F+Gx_i/c_i}{H+Ix_i/c_i}$, where F, G, H and I are given implicitly by (18). Then

$$|\frac{\partial q_e}{\partial x_i}| = |\frac{x_e}{c_e}\frac{FI/c_i - GH/c_i}{(H+Ix_i/c_i)^2}| \le \frac{2\cdot\binom{m}{n-1}\beta^n D^2\|b/\gamma_A\|_1}{(\alpha/c_{\max})^n} \le \gamma.$$

For $e = i$, we have $q_e = \frac{Gx_e/c_e}{H+Ix_e/c_e}$, where G, H, and I are given implicitly by (18). Then

$$|\frac{\partial q_e}{\partial x_e}| = |\frac{GH/c_e}{(H+Ix_e/c_e)^2}| \le \frac{\binom{m}{n-1}\beta^n D^2\|b/\gamma_A\|_1}{(\alpha/c_{\max})^n} \le \gamma.$$

Finally, consider x and $\tilde{x}$ with $\alpha\mathbf{1} \le x, \tilde{x} \le \beta\mathbf{1}$. Let $\bar{x}_\ell = (\tilde{x}_1, \ldots, \tilde{x}_\ell, x_{\ell+1}, \ldots, x_m)$. Then

$$||q_e(x)| - |q_e(\tilde{x})|| \le |q_e(x) - q_e(\tilde{x})| \le \sum_{0\le\ell<m}|q_e(\bar{x}_\ell) - q_e(\bar{x}_{\ell+1})| \le \gamma\|x-\tilde{x}\|_1.$$

In the general case where $c \geq 0$, we first derive an expression for p'' similar to (17). Then the equations for p' in (12) yield p', the equations for q_P in (12) yield q_P, and finally the equations for q_Z in (12) yield q_Z. $\square$

We are now ready to establish the existence of a solution with domain $[0, \infty)$.

Lemma 3.7. *The solution to the undirected dynamics in* (1) *has domain* $[0, \infty)$*. Moreover,* $x(0)e^{-t} \leq x(t) \leq D\|b/\gamma_A\|_1 \cdot \mathbf{1} + \max(\mathbf{0}, x(0) - D\|b/\gamma_A\|_1 \cdot \mathbf{1})e^{-t}$ *for all* t*.*

Proof. Consider any $x_0 > 0$ and any $t_0 \geq 0$. We first show that there is a positive δ' (depending on x_0) such that a unique solution $x(t)$ with $x(t_0) = x_0$ exists for $t \in (t_0 - \delta', t_0 + \delta')$. By the Picard-Lindelöf Theorem, this holds true if the mapping $x \mapsto |q| - x$ is continuous and satisfies a Lipschitz condition in a neighborhood of x_0. Continuity clearly holds. Let $\varepsilon = \min_i (x_0)_i/2$ and let $U = \{\, x : \|x - x_0\|_\infty < \varepsilon \,\}$. Then for every $x, \tilde{x} \in U$ and every e

$$\big||q_e(x)| - |q_e(\tilde{x})|\big| \leq \gamma \|x - \tilde{x}\|_1,$$

where γ is as in Lemma 3.6. Local existence implies the existence of a solution which cannot be extended. Since q is bounded (Lemma 3.3), x is bounded at all finite times, and hence the solution exists for all t. The lower bound $x(t) \geq x(0)e^{-t} > 0$ holds by Fact 2.2 with $A = 0$ and $\alpha = -1$. Since $|q_e| \leq D\|b/\gamma_A\|_1$, $\dot{x} = |q| - x \leq D\|b/\gamma_A\|_1 \cdot \mathbf{1} - x$, we have $x(t) \leq D\|b/\gamma_A\|_1 \cdot \mathbf{1} + \max(\mathbf{0}, x(0) - D\|b/\gamma_A\|_1 \cdot \mathbf{1})e^{-t}$ by Fact 2.2 with $B = D\|b/\gamma_A\|_1 \cdot \mathbf{1}$ and $\beta = -1$. $\square$

3.2 LP Duality

The energy $E(x)$ is no longer a Lyapunov function, e.g., if $x(0) \approx \mathbf{0}$, $x(t)$ and hence $E(x(t))$ will grow initially. We will show that energy suitably scaled is a Lyapunov function. What is the appropriate scaling factor? In the case of the undirected shortest path problem, [BMV12] used the minimum capacity of any source-sink cut as a scaling factor. The proper generalization to our situation is to consider the linear program $\max\{\alpha : Af = \alpha b, |f| \leq x\}$, where x is a fixed positive vector. Linear programming duality yields the corresponding minimization problem which generalizes the minimum cut problem to our situation.

Lemma 3.8. *Let* $x \in \mathbb{R}^m_{>0}$ *and* $b \neq 0$*. The linear programs*

$$\max\{\alpha : Af = \alpha b, |f| \leq x\} \qquad \textit{and} \qquad \min\{|y^T A| x : b^T y = -1\} \tag{19}$$

are feasible and have the same objective value. Moreover, there is a finite set $Y_A = \{\, d^1, \ldots, d^K \,\}$ *of vectors* $d^i \in \mathbb{R}^m_{\geq 0}$ *that are independent of* x *such that the minimum above is equal to* $C_\star = \min_{d \in Y_A} d^T x$*. There is a feasible* f *with* $|f| \leq x/C_\star$*.* [6]

Proof. The pair $(\alpha, f) = (0, 0)$ is a feasible solution for the maximization problem. Since $b \neq 0$, there exists y with $b^T y = -1$ and thus both problems are feasible. The dual of $\max\{\alpha : Af - \alpha b = 0, f \leq x, -f \leq x\}$ has unconstrained variables $y \in \mathbb{R}^n$ and non-negative variables $z^+, z^- \in \mathbb{R}^m$ and reads

$$\min\{x^T(z^+ + z^-) : -b^T y = 1, A^T y + z^+ - z^- = 0,\ z^+, z^- \geq 0\}. \tag{20}$$

From $z^- = A^T y + z^+$, $z^+ \geq 0$, $z^- \geq 0$ and $x > 0$, we conclude $\min(z^+, z^-) = 0$ in an optimal solution. Thus $z^- = \max(0, A^T y)$ and $z^+ = \max(0, -A^T y)$ and hence $z^+ + z^- = |A^T y|$ in an optimal dual solution. Therefore, (20) and the right LP in (19) have the same objective value.

We next show that the dual attains its minimum at a vertex of the feasible set. For this it suffices to show that its feasible set contains no line. Assume it does. Then there are vectors $d = (y_1, z_1^+, z_1^-)$, d non-zero, and $p = (y_0, z_0^+, z_0^-)$ such that $(y, z^+, z^-) = p + \lambda d = (y_0 + \lambda y_1, z_0^+ + \lambda z_1^+, z_0^- + \lambda z_1^-)$ is feasible for all $\lambda \in \mathbb{R}$. Thus $z_1^+ = z_1^- = 0$. Note that if either z_1^+ or z_1^- would be non-zero then either $z_0^+ + \lambda z_1^+$ or $z_0^- + \lambda z_1^-$ would

[6] In the undirected shortest path problem, the d's are the incidence vectors of the undirected source-sink cuts. Let S be any set of vertices containing s_0 but not s_1, and let $\mathbf{1}^S$ be its associated indicator vector. The cut corresponding to S contains the edges having exactly one endpoint in S. Its indicator vector is $d^S = |A^T \mathbf{1}^S|$. Then $d^S_e = 1$ iff $|S \cap \{\, u, v \,\}| = 1$, where $e = (u, v)$ or $e = (v, u)$, and $d^S_e = 0$ otherwise. For a vector $x \geq 0$, $(d^S)^T x$ is the capacity of the source-sink cut $(S, V \setminus S)$. In this setting, $C_\star$ is the value of a minimum cut.

have a negative component for some lambda. Then $A^T y + z^+ + z^- = 0$ implies $A^T y_1 = 0$. Since A has full row rank, $y_1 = 0$. Thus the dual contains no line and the minimum is attained at a vertex of its feasible region. The feasible region of the dual does not depend on x.

Let (y^1, z_1^+, z_1^-) to (y^K, z_K^+, z_K^-) be the vertices of (20), and let $Y_A = \{\, |A^T y^1|, \ldots, |A^T y^K| \,\}$. Then

$$\begin{aligned}\min_{d \in Y_A} d^T x &= \min\{x^T(z^+ + z^-) : -b^T y = 1, A^T y + z^+ - z^- = 0,\ z^+, z^- \geq 0\} \\ &= \min\{|y^T A| x : b^T y = -1\}.\end{aligned}$$

We finally show that there is a feasible f with $|f| \leq x/C_\star$. Let $x' := x/C_\star$. Then $x' > 0$ and $\min_{d \in Y_A} d^T x' = \min_{d \in Y_A} d^T x / C_\star = C_\star / C_\star = 1$ and thus the right LP with $x = x'$ (19) has objective value 1. Hence, the left LP has objective value 1 and there is a feasible f with $|f| \leq x'$. □

3.3 Convergence to Dominance

In the network setting, an important role is played by the set of edge capacity vectors that support a feasible flow. In the LP setting, we generalize this notion to the set of *dominating states*, which is defined as

$$X_{\text{dom}} := \{x \in \mathbb{R}^m : \exists \text{ feasible } f : |f| \leq x\}.$$

An alternative characterization, using the set Y_A from Lemma 3.8, is

$$\mathcal{X}_1 := \{x \in \mathbb{R}^m_{\geq 0} : d^T x \geq 1 \text{ for all } d \in Y_A\}.$$

We prove that $X_{\text{dom}} = \mathcal{X}_1$ and that the set $\mathcal{X}_1$ is attracting in the sense that the distance between $x(t)$ and $\mathcal{X}_1$ goes to zero, as t increases.

Lemma 3.9. 1. *It holds that $X_{\text{dom}} = \mathcal{X}_1$. Moreover, $\lim_{t\to\infty} \text{dist}(x(t), \mathcal{X}_1) = 0$, where $\text{dist}(x, \mathcal{X}_1)$ is the Euclidean distance between x and $\mathcal{X}_1$.*

2. *If $x(t_0) \in \mathcal{X}_1$, then $x(t) \in \mathcal{X}_1$ for all $t \geq t_0$. For all sufficiently large t, $x(t) \in \mathcal{X}_{1/2} := \{x \in \mathbb{R}^n_{\geq 0} : d^T x \geq 1/2 \text{ for all } d \in Y_A\}$, and if $x \in \mathcal{X}_{1/2}$ then there is a feasible f with $|f| \leq 2x$.*

Proof. 1. If $x \in \mathcal{X}_1$, then $d^T x \geq 1$ for all $d \in Y_A$ and hence Lemma 3.8 implies the existence of a feasible solution f with $|f| \leq x$. Conversely, if $x \in X_{\text{dom}}$, then there is a feasible f with $|f| \leq x$. Thus $d^T x \geq 1$ for all $d \in Y_A$ and hence $x \in \mathcal{X}_1$. By the proof of Lemma 3.8, for any $d \in Y_A$, there is a y such that $d = |A^T y|$ and $b^T y = -1$. Let $Y(t) = d^T x$. Then

$$\dot{Y} = |y^T A| \dot{x} = |y^T A|(|q| - x) \geq |y^T A q| - Y = |y^T b| - Y = 1 - Y.$$

Thus for any t_0 and $t \geq t_0$, $Y(t) \geq 1 + (Y(t_0) - 1)e^{-(t-t_0)}$ by Lemma 2.2 applied with $A = 1$ and $\alpha = -1$. In particular, $\liminf_{t\to\infty} Y(t) \geq 1$. Thus $\liminf_{t\to\infty} \min_{d\in Y_A} d^T x \geq 1$ and hence $\lim_{t\to\infty} \text{dist}(x(t), \mathcal{X}_1) = 0$.

2. Moreover, if $Y(t_0) \geq 1$, then $Y(t) \geq 1$ for all $t \geq t_0$. Hence $x(t_0) \in \mathcal{X}_1$ implies $x(t) \in \mathcal{X}_1$ for all $t \geq t_0$. Since $x(t)$ converges to $\mathcal{X}_1$, $x(t) \in \mathcal{X}_{1/2}$ for all sufficiently large t. If $x \in \mathcal{X}_{1/2}$ there is f such that $Af = \frac{1}{2}b$ and $|f| \leq x$. Thus $2f$ is feasible and $|2f| \leq 2x$. □

The next lemma summarizes simple bounds on the values of resistors r, potentials p and states x that hold for sufficiently large t. Recall that $P = \{\, e \in [m] : c_e > 0 \,\}$ and $Z = \{\, e \in [m] : c_e = 0 \,\}$, see (10).

Lemma 3.10. 1. *For sufficiently large t, it holds that $r_e \geq c_e/(2D\|b/\gamma_A\|_1)$, $b^T p \leq 8D\|b/\gamma_A\|_1\|c\|_1$ and $|A_e^T p| \leq 8D^2\|b\|_1\|c\|_1$ for all e.*

2. *For all e, it holds that $\dot{x}_e/x_e \geq -1$ and for all $e \in P$, it holds that $\dot{x}_e/x_e \leq 8D^2\|b\|_1\|c\|_1/c_{\min}$.*

3. *There is a positive constant C such that for all $t \geq t_0$, there is a feasible f (depending on t) such that $x_e(t) \geq C$ for all indices e in the support of f.*

Proof. 1. By Lemma 3.7, $x_e(t) \le 2D\|b/\gamma_A\|_1$ for all sufficiently large t. It follows that $r_e = c_e/x_e \ge c_e/(2D\|b/\gamma_A\|_1)$. Due to Lemma 3.9, for large enough t, there is a feasible flow with $|f| \le 2x$. Together with $x_e(t) \le 2D\|b/\gamma_A\|_1$, it follows that

$$b^T p = E(q) \le E(2x) = 4c^T x \le 8D\|b/\gamma_A\|_1\|c\|_1.$$

Now, orient A according to q and consider any index e'. Recall that for all indices e, we have $A_e^T p = 0$ if $e \in Z$, and $q_e = (x_e/c_e) \cdot A_e^T p$ if $e \in P$. Thus $A_e^T p \ge 0$ for all e. If $e' \in Z$ or $e' \in P$ and $q_{e'} = 0$, the claim is obvious. So assume $e' \in P$ and $q_{e'} > 0$. Since q is a convex combination of q-sign-compatible basic feasible solutions, there is a basic feasible solution f with $f \ge 0$ and $f_{e'} > 0$. By Lemma 3.1, $f_{e'} \ge 1/(D\gamma_A)$. Therefore

$$f_{e'} A_{e'}^T p \le \sum_e f_e A_e^T p = b^T p \le 8D\|b/\gamma_A\|_1\|c\|_1$$

for all sufficiently large t. The inequality follows from $f_e \ge 0$ and $A_e^T p \ge 0$ for all e. Thus $A_{e'}^T p \le 8D^2\|b\|_1\|c\|_1$ for all sufficiently large t.

2. We have $\dot{x}_e/x_e = (|q_e| - x_e)/x_e \ge -1$ for all e. For e with $c_e > 0$

$$\frac{\dot{x}_e}{x_e} = \frac{|q_e| - x_e}{x_e} \le \frac{|A_e^T p|}{c_e} \le 8D^2\|b\|_1\|c\|_1/c_{\min}.$$

3. Let t_0 be such that $d^T x(t) \ge 1/2$ for all $d \in Y_A$ and $t \ge t_0$. Then for all $t \ge t_0$, there is f such that $Af = \frac{1}{2}b$ and $|f| \le x(t)$; f may depend on t. By Lemma 2.1, we can write $2f$ as convex combination of f-sign-compatible basic feasible solutions (at most m of them) and a f-sign-compatible solution in the kernel of A. Dropping the solution in the kernel of A leaves us with a solution which is still dominated by x.

   It holds that for every $e \in E$ with $f_e \ne 0$, there is a basic feasible solution g used in the convex decomposition such that $2|f_e| \ge |g_e| > 0$. By Lemma 3.1, every non-zero component of g is at least $1/(D\gamma_A)$. We conclude that $x_e \ge 1/(2D\gamma_A)$, for every e in the support of g. □

3.4 The Equilibrium Points

We next characterize the equilibrium points

$$F = \{\, x \in \mathbb{R}_{\ge 0} : |q| = x \,\}. \tag{21}$$

Let us first elaborate on the special case of the undirected shortest path problem. Here the equilibria are the flows of value one from source to sink in a network formed by undirected source-sink paths of the same length. This can be seen as follows. Consider any $x \ge 0$ and assume $\mathrm{supp}(x)$ is a network of undirected source-sink paths of the same length. Call this network $\mathcal{N}$. Assign to each node u, a potential p_u equal to the length of the shortest undirected path from the sink s_1 to u. These potentials are well-defined as all paths from s_1 to u in $\mathcal{N}$ must have the same length. For an edge $e = (u, v)$ in $\mathcal{N}$, we have $q_e = x_e/c_e(p_u - p_v) = x_e/c_e \cdot c_e = x_e$, i.e., $q = x$ is the electrical flow with respect to the resistances c_e/x_e. Conversely, if x is an equilibrium point and the network is oriented such that $q \ge 0$, we have $x_e = q_e = x_e/c_e(p_u - p_v)$ for all edges $e = (u, v) \in \mathrm{supp}(x)$. Thus $c_e = p_u - p_v$ and this is only possible if for every node u, all paths from u to the sink have the same length. Thus $\mathrm{supp}(x)$ must be a network of undirected source-sink paths of the same length. We next generalize this reasoning.

Theorem 3.11. *If $x = |q|$ is an equilibrium point and the columns of A are oriented such that $q \ge 0$, then all feasible solutions f with $\mathrm{supp}(f) \subseteq \mathrm{supp}(x)$ satisfy $c^T f = c^T x$. Conversely, if $x = |q|$ for a feasible q, A is oriented such that $q \ge 0$, and all feasible solutions f with $\mathrm{supp}(f) \subseteq \mathrm{supp}(x)$ satisfy $c^T f = c^T x$, then x is an equilibrium point.*

Proof. If x is an equilibrium point, $|q_e| = x_e$ for every e. By changing the signs of some columns of A, we may assume $q \geq 0$, i.e., $q = x$. Let p be the potential with respect to x. For every index $e \in P$ in the support of x, since $c_e > 0$ we have $q_e = \frac{x_e}{c_e} A_e^T p$ and hence $c_e = A_e^T p$. Further, for the indices $e \in Z$ in the support of x, we have $c_e = 0 = A_e^T p$ due to the second block of equations on the right hand side in (11). Let f be any feasible solution whose support is contained in the support of x. Then the first part follows by

$$\sum_{e \in \text{supp}(f)} c_e f_e = \sum_{e \in \text{supp}(f)} f_e A_e^T p = b^T p = E(q) = E(x) = \text{cost}(x).$$

For the second part, we misuse notation and use A to also denote the submatrix of the constraint matrix indexed by the columns in the support of x. We may assume that the rows of A are independent. Otherwise, we simply drop redundant constraints. We may assume $q \geq 0$; otherwise we simply change the sign of some columns of A. Then x is feasible. Let A_B be a square non-singular submatrix of A and let A_N consist of the remaining columns of A. The feasible solutions f with $\text{supp}(f) \subseteq \text{supp}(x)$ satisfy $A_B f_B + A_N f_N = b$ and hence $f_B = A_B^{-1}(b - A_N f_N)$. Then

$$c^T f = c_B^T f_B + c_N^T f_N = c_B A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) f_N.$$

Since, by assumption, $c^T f$ is constant for all feasible solutions whose support is contained in the support of x, we must have $c_N = A_N^T [A_B^{-1}]^T c_B$. Let $p = [A_B^{-1}]^T c_B$. Then it follows that $A^T p = \left[\begin{smallmatrix} A_B^T \\ A_N^T \end{smallmatrix}\right] [A_B^{-1}]^T c_B = \left[\begin{smallmatrix} c_B \\ c_N \end{smallmatrix}\right]$ and hence $Rx = A^T p$. Thus the pair (x, p) satisfies the right hand side of (11). Since x is feasible, it also satisfies the left hand side of (11). Therefore, x is the minimum energy solution with respect to x. ☐

Corollary 3.12. *Let g be a basic feasible solution. Then $|g|$ is an equilibrium point.*

Proof. Let g be a basic feasible solution. Orient A such that $g \geq 0$. Since g is basic, there is a $B \subseteq [m]$ such that $g = (g_B, g_N) = (A_B^{-1} b, 0)$. Consider any feasible solution f with $\text{supp}(f) \subseteq \text{supp}(g)$. Then $f = (f_B, 0)$ and hence $b = Af = A_B f_B$. Therefore, $f_B = g_B$ and hence $c^T f = c^T g$. Thus $x = |g|$ is an equilibrium point. ☐

This characterization of equilibria has an interesting consequence.

Lemma 3.13. *The set $L := \{c^T x : x \in F\}$ of costs of equilibria is finite.*

Proof. If x is an equilibrium, $x = |q|$, where q is the minimum energy solution with respect to x. Orient A such that $q \geq 0$. Then by Theorem 3.11, $c^T f = c^T x$ for all feasible solutions f with $\text{supp}(f) \subseteq \text{supp}(x)$. In particular, this holds true for all such basic feasible solutions f. Thus L is a subset of the set of costs of all basic feasible solutions, which is a finite set. ☐

We conclude this part by showing that the optimal solutions of the undirected linear program (2) are equilibria.

Theorem 3.14. *Let x be an optimal solution to (2). Then x is an equilibrium.*

Proof. By definition, there is a feasible f with $|f| = x$. Let us reorient the columns of A such that $f \geq 0$ and let us delete all columns e of A with $f_e = 0$. Consider any feasible g with $\text{supp}(g) \subseteq \text{supp}(x)$. We claim that $c^T x = c^T g$. Assume otherwise and consider the point $y = x + \lambda(g - x)$. If $|\lambda|$ is sufficiently small, $y \geq 0$. Furthermore, y is feasible and $c^T y = c^T x + \lambda(c^T g - c^T x)$. If $c^T g \neq c^T x$, x is not an optimal solution to (2). The claim now follows from Theorem 3.11. ☐

3.5 Convergence

In order to show convergence, we construct a Lyapunov function. The following functions play a crucial role in our analysis. Let $C_d = d^T x$ for $d \in Y_A$, and recall that $C_\star = \min_{d \in Y_A} d^T x$ denotes the optimum. Moreover, we define by

$$h(t) := \sum_e r_e |q_e| \frac{x_e}{C_\star} - E\left(\frac{x}{C_\star}\right) \quad \text{and} \quad V_d := \frac{c^T x}{C_d} \text{ for every } d \in Y_A.$$

Theorem 3.15. *(1) For every $d \in Y_A$, $\dot{C}_d \geq 1 - C_d$. Thus, if $C_d < 1$ then $\dot{C}_d > 0$.*

(2) If $x(t) \in \mathcal{X}_1$, then $\frac{d}{dt}\mathrm{cost}(x(t)) \leq 0$ with equality if and only if $x = |q|$.

(3) Let $d \in Y_A$ be such that $C_\star = d^T x$ at time t. Then it holds that $\dot{V}_d \leq h(t)$.

(4) It holds that $h(t) \leq 0$ with equality if and only if $|q| = \frac{x}{C_\star}$.

Proof. 1. Recall that for $d \in Y_A$, there is a y such that $b^T y = -1$ and $d = |A^T y|$. Thus $\dot{C}_d = d^T(|q| - x) \geq |y^T A q| - C_d = 1 - C_d$ and hence $\dot{C}_d > 0$, whenever $C_d < 1$.

2. Remember that $E(x) = \mathrm{cost}(x)$ and that $x(t) \in \mathcal{X}_1$ implies that there is a feasible f with $|f| = x$. Thus $E(q) \leq E(f) \leq E(x)$. Let R be the diagonal matrix of entries c_e/x_e. Then

$$
\begin{aligned}
\frac{d}{dt}\mathrm{cost}(x) &= c^T(|q| - x) && \text{by (1)}\\
&= x^T R^{1/2} R^{1/2} |q| - x^T R x && \text{since } c = Rx\\
&\leq (q^T R q)^{1/2} (x^T R x)^{1/2} - x^T R x && \text{by Cauchy-Schwarz}\\
&\leq 0 && \text{since } E(q) \leq E(x).
\end{aligned}
$$

If the derivative is zero, both inequalities above have to be equalities. This is only possible if the vectors $|q|$ and x are parallel and $E(q) = E(x)$. Let λ be such that $|q| = \lambda x$. Then $E(q) = \sum_e \frac{c_e}{x_e} q_e^2 = \lambda^2 \sum_e c_e x_e = \lambda^2 E(x)$. Since $E(x) > 0$, this implies $\lambda = 1$.

3. By definition of d, $C_\star = C_d$. By the first two items, we have $\dot{C}_\star = d^T|q| - C_\star$ and $\frac{d}{dt}\mathrm{cost}(x) = c^T|q| - \mathrm{cost}(x)$. Thus

$$
\begin{aligned}
\frac{d}{dt}\frac{\mathrm{cost}(x)}{C_\star} &= \frac{C_\star \frac{d}{dt}\mathrm{cost}(x) - \dot{C}_\star \mathrm{cost}(x)}{C_\star^2} = \frac{C_\star(c^T|q| - \mathrm{cost}(x)) - (d^T|q| - C_\star)\mathrm{cost}(x)}{C_\star^2}\\
&= \frac{C_\star \cdot c^T|q| - d^T|q| \cdot c^T x}{C_\star^2} \leq \sum_e r_e |q_e| \frac{x_e}{C_\star} - \sum_e r_e \Big(\frac{x_e}{C_\star}\Big)^2 = h(t),
\end{aligned}
$$

where we used $r_e = c_e/x_e$ and hence $c^T|q| = \sum_e r_e x_e |q_e|$, $c^T x = E(x)$, and $d^T|q| \geq |y^T A q| = 1$ since $d = |y^T A|$ for some y with $b^T y = -1$.

4. We have

$$
\sum_e r_e \tfrac{x_e}{C_\star} |q_e| = \sum_e r_e^{1/2} \tfrac{x_e}{C_\star} r_e^{1/2} |q_e| \leq \Big(\sum_e r_e \big(\tfrac{x_e}{C_\star}\big)^2\Big)^{1/2} \Big(\sum_e r_e q_e^2\Big)^{1/2} = E\big(\tfrac{x}{C_\star}\big)^{1/2} E(q)^{1/2}
$$

by Cauchy-Schwarz. Since $h(t) = \sum_e r_e |q_e| \frac{x_e}{C_\star} - E\big(\frac{x}{C_\star}\big)$ by definition, it follows that

$$
h(t) \leq E\big(\tfrac{x}{C_\star}\big)^{1/2} \cdot E(q)^{1/2} - E\big(\tfrac{x}{C_\star}\big) = E\big(\tfrac{x}{C_\star}\big)^{1/2} \cdot \Big(E(q)^{1/2} - E\big(\tfrac{x}{C_\star}\big)^{1/2}\Big) \leq 0
$$

since $x/C_\star$ dominates a feasible solution and hence $E(q) \leq E(x/C_\star)$. If $h(t) = 0$, we must have equality in the application of Cauchy-Schwarz, i.e., the vectors $x/C_\star$ and $|q|$ must be parallel, and we must have $E(q) = E(x/C_\star)$ as in the proof of part 2. □

We show now convergence against the set of equilibrium points. We need the following technical Lemma from [BMV12].

Lemma 3.16 (Lemma 9 in [BMV12])**.** *Let $f(t) = \max_{d \in Y_A} f_d(t)$, where each f_d is continuous and differentiable. If $\dot{f}(t)$ exists, then there is a $d \in Y_A$ such that $f(t) = f_d(t)$ and $\dot{f}(t) = \dot{f}_d(t)$.*

Theorem 3.17. *All trajectories converge to the set F of equilibrium points.*

Proof. We distinguish cases according to whether the trajectory ever enters $\mathcal{X}_1$ or not. If the trajectory enters $\mathcal{X}_1$, say $x(t_0) \in \mathcal{X}_1$, then $\frac{d}{dt}\mathrm{cost}(x) \leq 0$ for all $t \geq t_0$ with equality only if $x = |q|$. Thus the trajectory converges to the set of fix points. If the trajectory never enters $\mathcal{X}_1$, consider $V = \max_{d \in Y_A}(V_d + 1 - C_d)$.

We show that $\dot{V}$ exists for almost all t. Moreover, if $\dot{V}(t)$ exists, then $\dot{V}(t) \le 0$ with equality if and only if $|q_e| = x_e$ for all e. It holds that V is Lipschitz-continuous as the maximum of a finite number of continuously differentiable functions. Since V is Lipschitz-continuous, the set of t's where $\dot{V}(t)$ does not exist has zero Lebesgue measure (see for example [CLSW98, Ch. 3]). If $\dot{V}(t)$ exists, we have $\dot{V}(t) = \dot{V}_d(t) - \dot{C}_d(t)$ for some $d \in Y_A$ according to Lemma 3.16. Then, it holds that $\dot{V}(t) \le h(t) - (1 - C_d) \le 0$. Thus $x(t)$ converges to the set

$$\{\, x \in \mathbb{R}_{\ge 0} : \dot{V} = 0 \,\} = \{\, x \in \mathbb{R}_{\ge 0} : |q| = x/C \text{ and } C = 1 \,\} = \{\, x \in \mathbb{R}_{\ge 0} : |q| = x \,\}. \qquad \square$$

At this point, we know that all trajectories $x(t)$ converge to F. Our next goal is to show that $c^T x(t)$ converges to the cost of an optimum solution of (2) and that $|q| - x$ converges to zero. We are only able to show the latter for all indices $e \in P$, i.e. with $c_e > 0$.

3.6 Details of the Convergence Process

In the argument to follow, we will encounter the following situation several times. We have a non-negative function $f(t) \ge 0$ and we know that $\int_0^\infty f(t)dt$ is finite. We want to conclude that $f(t)$ converges to zero for $t \to \infty$. This holds true if f is Lipschitz continuous. Note that the proof of the following lemma is very similar to the proof in [BMV12, Lemma 11]. However, in our case we apply the Local Lipschitz condition that we showed in Lemma 3.6.

Lemma 3.18. *Let $f(t) \ge 0$ for all t. If $\int_0^\infty f(t)d(t)$ is finite and $f(t)$ is Lipschitz-continuous, i.e., for every $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(t') - f(t)| \le \varepsilon$ for all $t' \in [t, t+\delta]$, then $f(t)$ converges to zero as t goes to infinity. The functions $t \mapsto x^T R|q| - x^T Rx = c^T|q| - c^T x$ and $t \mapsto h(t)$ are Lipschitz continuous.*

Proof. If $f(t)$ does not converge to zero, there is $\varepsilon > 0$ and an infinite unbounded sequence t_1, t_2, …such that $f(t_i) \ge \varepsilon$ for all i. Since f is Lipschitz continuous there is $\delta > 0$ such that $f(t_i') \ge \varepsilon/2$ for $t_i' \in [t_i, t_i + \delta]$ and all i. Hence, the integral $\int_0^\infty f(t)dt$ is unbounded.

Since $\dot{x_e}$ is continuous and bounded (by Lemma 3.7), x_e is Lipschitz-continuous. Thus, it is enough to show that q_e is Lipschitz-continuous for all e. Since q_Z (recall that $Z = \{\, e : c_e = 0 \,\}$ and $P = [m] \setminus Z$) is an affine function of q_P, it suffices to establish the claim for $e \in P$. So let $e \in P$ be such that $c_e > 0$. First, we claim that $x_e(t+\varepsilon) \le (1 + 2K\varepsilon)x_e$ for all $\varepsilon \le K/4$, where $K = 8D^2\|b\|_1\|c\|_1/c_{\min}$. Assume that this is not the case. Let

$$\varepsilon = \inf\{\delta \le 1/4K : x_e(t+\delta) > (1 + 2K\delta)x_e(t)\},$$

then $\varepsilon > 0$ (since $\dot{x_e}(t) \le Kx_e(t)$ by Lemma 3.10) and, by continuity, $x_e(t+\varepsilon) \ge (1 + 2K\varepsilon)x_e(t)$. There must be $t' \in [t, t+\varepsilon]$ such that $\dot{x_e}(t') = 2Kx_e(t)$. On the other hand,

$$\dot{x_e}(t') \le Kx_e(t') \le K(1 + 2K\varepsilon)x_e(t) \le K(1 + 2K/4K)x_e(t) < 2Kx_e(t),$$

which is a contradiction. Thus, $x_e(t+\varepsilon) \le (1 + 2K\varepsilon)x_e$ for all $\varepsilon \le 1/4K$. Similarly, $x_e(t+\varepsilon) \ge (1 - 2K\varepsilon)x_e$. Now, let $\alpha = (1 - 2K\varepsilon)x_e$ and $\beta = (1 + 2K\varepsilon)x_e$. Then

$$||q_e(t+\delta)| - |q_e(t)|| \le M\|x(t+\delta) - x(t)\|_1 \le Mm(4K\varepsilon)x_e \le 8\varepsilon MmKD\|b/\gamma_A\|_1,$$

since $x_e \le 2D\|b/\gamma_A\|_1$ for sufficiently large t and where M is as in Lemma 3.6. Since $C_\star$ is at least 1/2 for all sufficiently large t, the division by $C_\star$ and $C_\star^2$ in the definition of $h(t)$ does not affect the claim. $\square$

Lemma 3.19. *For all $e \in [m]$ of positive cost, it holds that $|x_e - |q_e|| \to 0$ as t goes to infinity.*

Proof. For a trajectory ultimately running in $\mathcal{X}_1$, we showed $\frac{d}{dt}\text{cost}(x) \le x^T R|q| - x^T Rx \le 0$ with equality if and only if $x = |q|$. Also, $E(q) \le E(x)$, since x dominates a feasible solution. Furthermore, $x^T R|q| - x^T Rx$ goes to zero using Lemma 3.18. Thus

$$\sum_e r_e(x_e - |q_e|)^2 = \sum_e r_e x_e^2 + \sum_e r_e q_e^2 - 2\sum_e r_e x_e|q_e| \le 2\big(\sum_e r_e x_e^2 - \sum_e r_e x_e|q_e|\big)$$

goes to zero. Next observe that there is a constant C such that $x_e(t) \leq C$ for all e and t as a result of Lemma 3.7. Also $c_{\min} > 0$ and hence $r_e \geq c_{\min}/C$. Thus $\sum_e r_e(x_e - |q_e|)^2 \leq \frac{C}{c_{\min}} \cdot \sum_e (x_e - |q_e|)^2$ and hence $|x_e - |q_e|| \to 0$ for every e with positive cost. For trajectories outside $\mathcal{X}_1$, we argue about $||q_e| - \frac{x}{C_\star}|$ and use $C_\star \to 1$, namely

$$\sum_e r_e(\tfrac{x_e}{C_\star} - |q_e|)^2 \leq 2\big(\sum_e r_e(\tfrac{x_e}{C_\star})^2 - \sum_e r_e \tfrac{x_e}{C_\star}|q_e|\big) \to 0. \qquad \square$$

Note that the above does not say anything about the indices $e \in Z$ (with $c_e = 0$). Recall that $A_P q_P + A_Z q_Z = b$ and that the columns of A_Z are independent. Thus, q_Z is uniquely determined by q_P. For the undirected shortest path problem, the potential difference $p^T b$ between source and sink converges to the length of a shortest source-sink path. If an edge with positive cost is used by some shortest undirected path, then no shortest undirected path uses it with the opposite direction. We prove the natural generalizations.

Let OPT be the set of optimal solutions to (2) and let $E_{opt} = \cup_{x \in OPT} \operatorname{supp}(x)$ be the set of columns used in some optimal solution. The columns of positive cost in E_{opt} can be consistently oriented as the following Lemma shows.

Lemma 3.20. *Let x_1^* and x_2^* be optimal solutions to (2) and let f and g be feasible solutions with $|f| = x_1^*$ and $|g| = x_2^*$. Then there is no e such that $f_e g_e < 0$ and $c_e > 0$.*

Proof. Assume otherwise. Then $|g_e - f_e| = |g_e| + |f_e| > 0$. Consider $h = (g_e f - f_e g)/(g_e - f_e)$. Then $Ah = (g_e Af - f_e Ag)/(g_e - f_e) = b$ and h is feasible. Also, $h_e = \frac{g_e f_e - f_e g_e}{g_e - f_e} = 0$ and for every index e', it holds that $|h_{e'}| = \frac{|g_e f_{e'} - f_e g_{e'}|}{|g_e - f_e)|} \leq \frac{|g_e||f_{e'}| + |f_e||g_{e'}|}{|g_e| + |f_e|}$ and hence

$$\operatorname{cost}(h) < \operatorname{cost}(f) + \operatorname{cost}(g) = \frac{|g_e|}{|g_e| + |f_e|}\operatorname{cost}(x_1^*) + \frac{|f_e|}{|g_e| + |f_e|}\operatorname{cost}(x_2^*) = \operatorname{cost}(x_1^*),$$

a contradiction to the optimality of x_1^* and x_2^*. $\square$

By the preceding Lemma, we can orient A such that $f_e \geq 0$ whenever $|f|$ is an optimal solution to (2) and $c_e > 0$. We then call A *positively oriented.*

Lemma 3.21. *It holds that $p^T b$ converges to the cost of an optimum solution of (2). If A is positively oriented, then $\liminf_{t \to \infty} A_e^T p \geq 0$ for all e.*

Proof. Let x^* be an optimal solution of (2). We first show convergence to a point in L and then convergence to $c^T x^*$. Let $\varepsilon > 0$ be arbitrary. Consider any time $t \geq t_0$, where t_0 and C as in Lemma 3.10 and moreover $||q_e| - x_e| \leq \frac{C\varepsilon}{c_{\max}}$ for every $e \in P$. Then $x_e \geq C$ for all indices e in the support of some basic feasible solution f. For every $e \in P$, we have $q_e = \frac{x_e}{c_e} A_e^T p$. We also assume $q \geq 0$ by possibly reorienting columns of A. Hence

$$|c_e - A_e^T p| = \big|1 - \frac{q_e}{x_e}\big|c_e = \big|\frac{x_e - q_e}{x_e}\big|c_e \leq \frac{c_{\max}}{C}|q_e - x_e| \leq \varepsilon.$$

For indices $e \in Z$, we have $A_e^T p = 0 = c_e$. Since, $\|f\|_\infty \leq D\|b/\gamma_A\|_1$ (Lemma 3.1), we conclude

$$c^T f - p^T b = \sum_{e \in \operatorname{supp}(f)} (c_e - p^T A_e) f_e \leq \varepsilon \sum_e |f_e| \leq \varepsilon \cdot mD\|b/\gamma_A\|_1.$$

Since the set L is finite, we can let $\varepsilon > 0$ be smaller than half the minimal distance between elements in L. By the preceding paragraph, there is for all sufficiently large t, a basic feasible solution f such that $|c^T f - b^T p| \leq \varepsilon$. Since $b^T p$ is a continuous function of time, $c^T f$ must become constant. We have now shown that $b^T p$ converges to an element in L. We will next show that $b^T p$ converges to the optimum cost. Let x^* be an optimum solution to (2) and let $W = \sum_e x_e^* c_e \ln x_e$. Since $x(t)$ is bounded, W is bounded. We assume that A is positively oriented, thus there is a feasible f^* with $|f^*| = x^*$ and $f^*_e \geq 0$ whenever $c_e > 0$. By reorienting zero cost columns, we may assume $f^*_e \geq 0$ for all e. Then $Ax^* = b$. We have

$$\begin{aligned}
\dot{W} &= \sum_e x^*_e c_e \frac{|q_e| - x_e}{x_e} \\
&= \sum_{e;\ c_e > 0} x^*_e |A_e^T p| - \mathrm{cost}(x^*) && \text{since } q_e = \tfrac{x_e}{c_e} A_e^T p \text{ whenever } c_e > 0 \\
&= \sum_e x^*_e |A_e^T p| - \mathrm{cost}(x^*) && \text{since } A_e^T p = 0 \text{ whenever } c_e = 0 \\
&= \sum_e x^*_e (|A_e^T p| - A_e^T p) + b^T p - \mathrm{cost}(x^*)
\end{aligned}$$

and hence $b^T p - \mathrm{cost}(x^*)$ must converge to zero; note that $b^T p$ is Lipschitz continuous in t.

Similarly, $|A_e^T p| - A_e^T p$ must converge to zero whenever $x^*_e > 0$. This implies $\liminf A_e^T p \geq 0$. Assume otherwise, i.e., for every $\varepsilon > 0$, we have $A_e^T p < -\varepsilon$ for arbitrarily large t. Since p is Lipschitz-continuous in t, there is a $\delta > 0$ such that $A_e^T p < -\varepsilon/2$ for infinitely many disjoint intervals of length δ. In these intervals, $|A_e^T p| - A_e^T p \geq \varepsilon$ and hence W must grow beyond any bound, a contradiction. □

Corollary 3.22. *$E(x)$ and $\mathrm{cost}(x)$ converge to $c^T x^*$, whereas x and $|q|$ converge to OPT. If the optimum solution is unique, x and $|q|$ converge to it. Moreover, if $e \notin E_{opt}$, x_e and $|q_e|$ converge to zero.*

Proof. The first part follows from $E(x) = \mathrm{cost}(x) = b^T p$ and the preceding Lemma. Thus x and q converge to the set F of equilibrium points, see (21), that are optimum solutions to (2). Since every optimum solution is an equilibrium point by Theorem 3.14, x and q converge to OPT. For $e \notin E_{opt}$, $f_e = 0$ for every $f \in F \cap OPT$. Since x and $|q|$ converge to $F \cap OPT$, x_e and $|q_e|$ converge to zero for every $e \in E_{opt}$. □

4 Improved Convergence Results: Discrete Directed Dynamics

In this section, we present in its full generality our main technical result on the Physarum dynamics (6).

4.1 Overview

Inspired by the max-flow min-cut theorem, we consider the following primal-dual pair of linear programs: the primal LP is given by $\max\{\, t : Af = t \cdot b;\ 0 \leq f \leq x \,\}$ in variables $f \in \mathbb{R}^m$ and $t \in \mathbb{R}$, and its dual LP reads $\min\left\{\, x^T z : z \geq 0;\ z \geq A^T y;\ b^T y = 1 \,\right\}$ in variables $z \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. Since the dual feasible region does not contain a line and the minimum is bounded, the optimum is attained at a vertex, and in an optimum solution we have $z = \max\{0, A^T y\}$. Let V be the set of vertices of the dual feasible region, and let $Y := \{\, y : (z, y) \in V \,\}$ be the set of their projections on y-space. Then, the dual optimum is given by $\min\{\, \max\{0, y^T A\} \cdot x : y \in Y \,\}$. The set of *strongly dominating* capacity vectors x is defined as

$$X := \left\{\, x \in \mathbb{R}^m_{>0} : y^T A x > 0 \text{ for all } y \in Y \,\right\}.^7$$

Note that X contains the set of all scaled feasible solutions $\{x = tf \,:\, Af = b,\ f \geq 0,\ t > 0\}$.

We next discuss the choice of step size. For $y \in Y$ and capacity vector x, let $\alpha(y, x) := y^T A x$. Further, let $\alpha(x) := \min\{\, \alpha(y, x) : y \in Y \,\}$ and $\alpha^{(\ell)} := \alpha(x^{(\ell)})$. Then, for any $x \in X$ there is a feasible f such that $0 \leq f \leq x/\alpha(x)$, see Lemma 4.8. In particular, if x is feasible then $\alpha(x) = 1$, since $\alpha(y, x) = 1$ for all $y \in Y$. We partition the Physarum dynamics (6) into the following five regimes and define for each regime a fixed step size, see Subsection 4.3.

[7] In the shortest path problem (recall that $b = e_1 - e_n$) the set Y consists of all $y \in \{-1, +1\}^n$ such that $y_1 = 1 = -y_n$, i.e., y encodes a cut with $S = \{\, i : y_i = -1 \,\}$ and $\overline{S} = \{\, i : y_i = +1 \,\}$. The condition $y^T A x > 0$ translates into $\sum_{a \in E(S, \overline{S})} x_a - \sum_{a \in E(\overline{S}, S)} x_a > 0$, i.e., every source-sink cut must have positive directed capacity.

Corollary 4.1. *The Physarum dynamics* (6) *initialized with* $x^{(0)} \in X$ *and a step size* h *satisfies:*

- *If* $\alpha^{(0)} = 1$, *we work with* $h \leq h_0$ *and have* $\alpha^{(\ell)} = 1$ *for all* ℓ.
- *If* $1/2 \leq \alpha^{(0)} < 1$, *we work with* $h \leq h_0/2$ *and have* $1 - \delta \leq \alpha^{(\ell)} < 1$ *for* $\ell \geq h^{-1} \log(1/2\delta)$ *and* $\delta > 0$.
- *If* $1 < \alpha^{(0)} \leq 1/h_0$, *we work with* $h \leq h_0$ *and have* $1 < \alpha^{(\ell)} \leq 1 + \delta$ *for* $\ell \geq h^{-1} \cdot \log(1/\delta h_0)$ *and* $\delta > 0$.
- *If* $0 < \alpha^{(0)} < 1/2$, *we work with* $h \leq \alpha^{(0)} h_0$ *and have* $1/2 \leq \alpha^{(\ell)} < 1$ *for* $\ell \geq 1/h$.
- *If* $1/h_0 < \alpha^{(0)}$, *we work with* $h \leq 1/4$ *and have* $1 < \alpha^{(\ell)} \leq 1/h_0$ *for* $\ell = \lfloor \log_{1/(1-h)} h_0(\alpha^{(0)} - 1)/(1 - h_0) \rfloor$.

In each regime, we have $1 - \alpha^{(\ell+1)} = (1 - h)(1 - \alpha^{(\ell)})$.

We give now the full version of Theorem 1.4 which applies for any strongly dominating starting point.

Theorem 4.2. *Suppose* $A \in \mathbb{Z}^{n \times m}$ *has full row rank* ($n \leq m$), $b \in \mathbb{Z}^n$, $c \in \mathbb{Z}^m_{>0}$ *and* $\varepsilon \in (0, 1)$. *Given* $x^{(0)} \in X$ *and its corresponding* $\alpha^{(0)}$, *the Physarum dynamics* (6) *initialized with* $x^{(0)}$ *runs in two regimes:*

(i) The first regime is executed when $\alpha^{(0)} \notin [1/2, 1/h_0]$ *and it computes a point* $x^{(t)} \in X$ *such that* $\alpha^{(t)} \in [1/2, 1/h_0]$. *In particular, if* $\alpha^{(0)} < 1/2$ *then* $h \leq (\Phi/\text{opt}) \cdot (\alpha^{(0)} h_0)^2$ *and* $t = 1/h$. *Otherwise, if* $\alpha^{(0)} > 1/h_0$ *then* $h \leq \Phi/\text{opt}$ *and* $t = \lfloor \log_{1/(1-h)}[h_0(\alpha^{(0)} - 1)/(1 - h_0)] \rfloor$.

(ii) The second regime starts from a point $x^{(t)} \in X$ *with* $\alpha^{(t)} \in [1/2, 1/h_0]$, *it has a step size* $h \leq (\Phi/\text{opt}) \cdot h_0^2/2$ *and outputs for any* $k \geq 4C_1/(h\Phi) \cdot \ln(C_2 \Psi^{(0)}/(\varepsilon \cdot \min\{1, x^{(0)}_{\min}\}))$ *a vector* $x^{(t+k)} \in X$ *such that* $\text{dist}(x^{(t+k)}, X_\star) < \varepsilon/(D\gamma_A)$.

We stated the bounds on h in terms of the unknown quantities Φ and opt. However, $\Phi/\text{opt} \geq 1/C_3$ by Lemma 3.1 and hence replacing Φ/opt by $1/C_3$ yields constructive bounds for h.

Organization: This section is devoted to proving Theorem 4.2, and it is organized as follows: Subsection 4.2 establishes core efficiency bounds that extend [SV16c] and yield a *scale-invariant* determinant dependence of the step size and are applicable to strongly dominating points. Subsection 4.3 gives the definition of strongly dominating points and shows that the Physarum dynamics (6) initialized with such a point is well defined. Subsection 4.4 extends the analysis in [BBD+13, SV16b, SV16c] to positive linear programs, by generalizing the concept of non-negative flows to non-negative *feasible kernel-free* vectors. Subsection 4.5 shows that $x^{(\ell)}$ converges to $X_\star$ for large enough ℓ. Subsection 4.6 concludes the proof of Theorem 4.2.

4.2 Useful Lemmas

Recall that $R^{(\ell)} = \text{diag}(c) \cdot (X^{(\ell)})^{-1}$ is a positive diagonal matrix and $L^{(\ell)} \stackrel{\text{def}}{=} A(R^{(\ell)})^{-1} A^T$ is invertible. Let $p^{(\ell)}$ be the unique solution of $L^{(\ell)} p^{(\ell)} = b$. We improve the dependence on D_S in [SV16c, Lemma 5.2] to D.

Lemma 4.3. *[SV16c, extension of Lemma 5.2] Suppose* $x^{(\ell)} > 0$, $R^{(\ell)}$ *is a positive diagonal matrix and* $L = A(R^{(\ell)})^{-1} A^T$. *Then for every* $e \in [m]$, *it holds that* $\|A^T (L^{(\ell)})^{-1} A_e\|_\infty \leq D \cdot c_e / x^{(\ell)}_e$.

Proof. The statement follows by combining the proof in [SV16c, Lemma 5.2] with Lemma 3.1. □

We show next that [SV16b, Corollary 5.3] holds for x-capacitated vectors, which extends the class of feasible starting points, and further yields a bound in terms of D.

Lemma 4.4. *[SV16b, extension of Corollary 5.3] Let* $p^{(\ell)}$ *be the unique solution of* $L^{(\ell)} p^{(\ell)} = b$ *and assume* $x^{(\ell)}$ *is a positive vector with corresponding positive scalar* $\alpha^{(\ell)}$ *such that there is a vector* f *satisfying* $Af = \alpha^{(\ell)} \cdot b$ *and* $0 \leq f \leq x^{(\ell)}$. *Then* $\|A^T p^{(\ell)}\|_\infty \leq D\|c\|_1/\alpha^{(\ell)}$.

Proof. By assumption, f satisfies $\alpha^{(\ell)} b = Af = \sum_e f_e A_e$ and $0 \le f \le x^{(\ell)}$. This yields

$$\alpha^{(\ell)} \|A^T p^{(\ell)}\|_\infty = \|A^T (L^{(\ell)})^{-1} \cdot \alpha^{(\ell)} b\|_\infty = \|\sum_e f_e A^T (L^{(\ell)})^{-1} A_e\|_\infty$$
$$\le \sum_e f_e \|A^T (L^{(\ell)})^{-1} A_e\|_\infty \overset{\text{(Lem. 4.3)}}{\le} D \sum_e f_e \frac{c_e}{x_e^{(\ell)}} \le D\|c\|_1. \qquad \square$$

We note that applying Lemma 4.3 and Lemma 4.4 into the analysis of [SV16c, Theorem 1.3] yields an improved result that depends on the scale-invariant determinant D. Moreover, we show in the next Subsection 4.3 that the Physarum dynamics (6) can be initialized with any strongly dominating point.

We establish now an upper bound on q that does not depend on x. We then use this upper bound on q to establish a uniform upper bound on x.

Lemma 4.5. *For any $x^{(\ell)} > 0$, $\|q^{(\ell)}\|_\infty \le mD^2 \|b/\gamma_A\|_1$.*

Proof. Let f be a basic feasible solution of $Af = b$. By definition, $q_e^{(\ell)} = (x_e^{(\ell)}/c_e) A_e^T (L^{(\ell)})^{-1} b$ and thus

$$\left|q_e^{(\ell)}\right| = \left|\frac{x_e^{(\ell)}}{c_e} \sum_u A_e^T (L^{(\ell)})^{-1} A_u f_u\right| \le \frac{x_e^{(\ell)}}{c_e} \sum_u |f_u| \cdot \left|A_e^T (L^{(\ell)})^{-1} A_u\right| \le D\|f\|_1,$$

where the last inequality follows by

$$\left|A_e^T (L^{(\ell)})^{-1} A_u\right| = \left|A_u^T (L^{(\ell)})^{-1} A_e\right| \le \|A^T (L^{(\ell)})^{-1} A_e\|_\infty \overset{\text{(Lem. 4.3)}}{\le} D \cdot c_e / x_e^{(\ell)}.$$

By Cramer's rule and Lemma 3.1, we have $|q_e^{(\ell)}| \le D\|f\|_1 \le mD^2 \|b/\gamma_A\|_1$. $\square$

Let $k, t \in \mathbb{N}$. We denote by

$$\overline{q}^{(t,k)} = \sum_{i=t}^{t+k-1} \frac{h\,(1-h)^{t+k-1-i}}{1-(1-h)^k} q^{(i)} \quad \text{and} \quad \overline{p}^{(t,k)} = \sum_{i=t}^{t+k-1} p^{(i)}. \tag{22}$$

Straightforward checking shows that $A\overline{q}^{(t,k)} = b$. Further, for $C := \mathrm{diag}(c)$, $t \ge 0$ and $k \ge 1$, we have

$$x^{(t)} \prod_{i=t}^{t+k-1} [1 + h(C^{-1} A^T p^{(i)} - 1)] \;=\; x^{(t+k)} \;=\; (1-h)^k x^{(t)} + [1-(1-h)^k] \overline{q}^{(t,k)}.$$

We give next an upper bound on $x^{(k)}$ that is independent of k.

Lemma 4.6. *Let $\Psi^{(0)} = \max\{mD^2 \|b/\gamma_A\|_1, \|x^{(0)}\|_\infty\}$. Then $\|x^{(k)}\|_\infty \le \Psi^{(0)}$, $\forall k \in \mathbb{N}$.*

Proof. We prove the statement by induction. The base case $\|x^{(0)}\|_\infty \le \Psi^{(0)}$ is clear. Suppose the statement holds for some $k > 0$. Then, triangle inequality and Lemma 4.5 yield

$$\|x^{(k+1)}\|_\infty \le (1-h)\|x^{(k)}\|_\infty + h\|q^{(k)}\|_\infty \le (1-h)\Psi^{(0)} + h\Psi^{(0)} \le \Psi^{(0)}. \qquad \square$$

We show now convergence to feasibility.

Lemma 4.7. *Let $r^{(k)} = b - Ax^{(k)}$. Then $r^{(k+1)} = (1-h) r^{(k)}$ and hence $r^{(k)} = (1-h)^k (b - Ax^{(0)})$.*

Proof. By definition $x^{(k+1)} = (1-h)x^{(k)} + hq^{(k)}$, and thus the statement follows by

$$r^{(k+1)} = b - Ax^{(k+1)} = b - (1-h)Ax^{(k)} - hb = (1-h) r^{(k)}. \qquad \square$$

4.3 Strongly Dominating Capacity Vectors

For the shortest path problem, it is known that one can start from any capacity vector x for which the directed capacity of every source-sink cut is positive, where the directed capacity of a cut is the total capacity of the edges crossing the cut in source-sink direction minus the total capacity of the edges crossing the cut in the sink-source direction. We generalize this result. We start with the max-flow like LP

$$\max\{\, t : Af = t \cdot b;\; 0 \le f \le x \,\} \tag{23}$$

in variables $f \in \mathbb{R}^m$ and $t \in \mathbb{R}$ and its dual

$$\min\left\{\, x^T z : z \ge 0;\; z \ge A^T y;\; b^T y = 1 \,\right\} \tag{24}$$

in variables $z \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. The feasible region of the dual contains no line. Assume otherwise; say it contains $(z,y) = (z^{(0)}, y^{(0)}) + \lambda(z^{(1)}, y^{(1)})$ for all $\lambda \in \mathbb{R}$. Then, $z \ge 0$ implies $z^{(1)} = 0$ and further $z \ge A^T y$ implies $z^{(0)} \ge A^T y^{(0)} + \lambda A^T y^{(1)}$ and hence $A^T y^{(1)} = 0$. Since A has full row rank, we have $y^{(1)} = 0$. The optimum of the dual is therefore attained at a vertex. In an optimum solution, we have $z = \max\{0, A^T y\}$. Let V be the set of vertices of the feasible region of the dual (24), and let

$$Y := \{\, y : (z,y) \in V \,\}$$

be the set of their projections on y-space. Then, the optimum of the dual (24) is given by

$$\min_{y \in Y} \left\{ \max\{0, y^T A\} \cdot x \right\}. \tag{25}$$

The set of strongly dominating capacity vectors x is defined by

$$X := \left\{\, x \in \mathbb{R}^m_{>0} : y^T A x > 0 \text{ for all } y \in Y \,\right\}. \tag{26}$$

We next show that for all $x^{(0)} \in X$ and sufficiently small step size, the sequence $\{x^{(k)}\}_{k \in \mathbb{N}}$ stays in X. Moreover, $y^T A x^{(k)}$ converges to 1 for every $y \in Y$. We define by

$$\alpha(y,x) := y^T A x \quad \text{and} \quad \alpha(x) := \min\{\, \alpha(y,x) : y \in Y \,\}.$$

Let $\alpha^{(\ell)} := \alpha(x^{(\ell)})$. Then, $x^{(\ell)} \in X$ iff $\alpha^{(\ell)} > 0$. We summarize the discussion in the following Lemma.

Lemma 4.8. *Suppose $x^{(\ell)} \in X$. Then, there is a vector f such that $Af = \alpha^{(\ell)} \cdot b$ and $0 \le f \le x^{(\ell)}$.*

Proof. By the strong duality theorem applied on (23) and (24), it holds by (25) that

$$t = \min_{y \in Y} \left\{ \max\{0, y^T A\} \cdot x^{(\ell)} \right\} \ge \min_{y \in Y} y^T A x^{(\ell)} = \alpha^{(\ell)}.$$

The statement follows by the definition of (23). ☐

We demonstrate now that $\alpha^{(\ell)}$ converges to 1.

Lemma 4.9. *Assume $x^{(\ell)} \in X$. Then, for any $h^{(\ell)} \le \min\{1/4, \alpha^{(\ell)} h_0\}$ we have $x^{(\ell+1)} \in X$ and*

$$1 - \alpha^{(\ell+1)} = (1 - h^{(\ell)}) \cdot (1 - \alpha^{(\ell)}).$$

Proof. By applying Lemma 4.4 and Lemma 4.8 with $x^{(\ell)} \in X$, we have $\|A^T p^{(\ell)}\|_\infty \le D\|c\|_1/\alpha^{(\ell)}$ and hence for every index e it holds $-h^{(\ell)} \cdot c_e^{-1} x_e^{(\ell)} A_e^T p^{(\ell)} \ge -(h^{(\ell)} x_e^{(\ell)})/(2\alpha^{(\ell)} h_0) \ge -x_e^{(\ell)}/2$. Thus,

$$x_e^{(\ell+1)} = (1 - h^{(\ell)}) x_e^{(\ell)} + h^{(\ell)} \cdot [R_e^{(\ell)}]^{-1} A_e^T p^{(\ell)} \ge \frac{3}{4} x_e^{(\ell)} - \frac{1}{2} x_e^{(\ell)} = \frac{1}{4} x_e^{(\ell)} > 0.$$

Let $y \in Y$ be arbitrary. Then $y^T b = 1$ and hence $y^T r^{(\ell)} = y^T (b - A x^{(\ell)}) = 1 - y^T A x^{(\ell)} = 1 - \alpha(y, x^{(\ell)})$. The second claim now follows from Lemma 4.7. ☐

We note that the convergence speed crucially depends on the initial point $x^{(0)} \in X$, and in particular to its corresponding value $\alpha^{(0)}$. Further, this dependence naturally partitions the Physarum dynamics (6) into the five regimes given in Corollary 4.1.

4.4 $x^{(k)}$ is Close to a Non-Negative Kernel-Free Vector

In this subsection, we generalize [SV16b, Lemma 5.4] to positive linear programs. We achieve this in two steps. First, we generalize a result by Ito et al. [IJNT11, Lemma 2] to positive linear programs and then we substitute the notion of a non-negative cycle-free flow with a non-negative feasible kernel-free vector.

Throughout this and the consecutive subsection, we denote by $\rho_A := \max\left\{D\gamma_A,\, nD^2\|A\|_\infty\right\}$.

Lemma 4.10. *Suppose a matrix $A \in \mathbb{Z}^{n\times m}$ has full row rank and vector $b \in \mathbb{Z}^n$. Let g be a feasible solution to $Ag = b$ and $S \subseteq [n]$ be a subset of row indices of A such that $\sum_{i\in S}|g_i| < 1/\rho_A$. Then, there is a feasible solution f such that $g_i \cdot f_i \geq 0$ for all $i \in [n]$, $f_i = 0$ for all $i \in S$ and $\|f - g\|_\infty < 1/(D\gamma_A)$.*

Proof. W.l.o.g. we can assume that $g \geq 0$ as we could change the signs of the columns of A accordingly. Let $\mathbf{1}_S$ be the indicator vector of S. We consider the linear program

$$\min\{\mathbf{1}_S^T x \,:\, Ax = b,\ x \geq 0\}$$

and let *opt* be its optimum value. Notice that $0 \leq opt \leq \mathbf{1}_S^T g < 1/\rho_A$. Since the feasible region does not contain a line and the minimum is bounded, the optimum is attained at a basic feasible solution, say f. Suppose that there is an index $i \in S$ with $f_i > 0$. By Lemma 3.1, we have $f_i \geq 1/(D\gamma_A)$. This is a contradiction to the optimality of f and hence $f_i = 0$ for all $i \in S$.

Among the feasible solutions f such that $f_i g_i \geq 0$ for all i and $f_i = 0$ for all $i \in S$, we choose the one that minimizes $\|f - g\|_\infty$. For simplicity, we also denote it by f. Note that f satisfies $\mathrm{supp}(f) \subseteq \overline{S}$, where $\overline{S} = [m]\backslash S$. Further, since $f_S = 0$ and

$$A_S g_S + A_{\overline{S}} g_{\overline{S}} = Ag = b = Af = A_S f_S + A_{\overline{S}} f_{\overline{S}} = A_{\overline{S}} f_{\overline{S}}$$

we have $A_{\overline{S}}\left(f_{\overline{S}} - g_{\overline{S}}\right) = A_S g_S$. Let A_B be a linearly independent column subset of $A_{\overline{S}}$ with maximal cardinality, i.e. the column subset A_N, where $N = \overline{S} \setminus B$, is linearly dependent on A_B. Hence, there is an invertible square submatrix $A'_B \in \mathbb{Z}^{|B|\times|B|}$ of A_B and a vector $v = (v_B, 0_N)$ such that

$$\begin{pmatrix} A'_B \\ A''_B \end{pmatrix} v_B = A_B v_B = A_S g_S.$$

Let $r = (A_S g_S)_B$. Since A'_B is invertible, there is a unique vector v_B such that $A'_B v_B = r$. Observe that

$$|r_i| = \left|\sum_{j\in S} A_{i,j} g_j\right| \leq \|A\|_\infty \sum_{j\in S}|g_j| < \frac{\|A\|_\infty}{nD^2\|A\|_\infty} = \frac{1}{nD^2}.$$

By Cramer's rule $v_B(e)$ is quotient of two determinants. The denominator is $\det(A'_B)$ and hence at least one in absolute value. For the numerator, the e-th column is replaced by r. Expansion according to this column shows that the absolute value of the numerator is bounded by

$$\frac{D}{\gamma_A}\sum_{i\in B}|r_i| < \frac{D}{\gamma_A} \cdot \frac{|B|}{nD^2} \leq \frac{1}{D\gamma_A}.$$

Therefore, $\|f - g\|_\infty \leq 1/(D\gamma_A)$ and the statement follows. ☐

Lemma 4.11. *Let $q \in \mathbb{R}^m$, $p \in \mathbb{R}^n$ and $N = \{e \in [m] : q_e \leq 0 \text{ or } p^T A_e \leq 0\}$, where $Aq = b$ and $p = L^{-1}b$. Suppose $\sum_{e\in N}|q_e| < 1/\rho_A$. Then there is a non-negative feasible kernel-free vector f such that* $\mathrm{supp}(f) \subseteq E\backslash N$ *and $\|f - q\|_\infty < 1/(D\gamma_A)$.*

Proof. We apply Lemma 4.10 to q with $S = N$. Then, there is a non-negative feasible vector f such that $\mathrm{supp}(f) \subseteq E\backslash N$ and $\|f - q\|_\infty < 1/(D\gamma_A)$. By Lemma 2.1, f can be expressed as a sum of a convex combination of basic feasible solutions plus a vector w in the kernel of A. Moreover, all vectors in this representation are sign compatible with f, and in particular w is non-negative too.

Suppose for contradiction that $w \neq 0$. By definition, $0 = p^T A w = \sum_{e \in [m]} p^T A_e w_e$ and since $w \geq 0$ and $w \neq 0$, it follows that there is an index $e \in [m]$ satisfying $w_e > 0$ and $p^T A_e \leq 0$. Since f and w are sign compatible, $w_e > 0$ implies $f_e > 0$. On the other hand, as $p^T A_e \leq 0$ we have $e \in N$ and thus $f_e = 0$. This is a contradiction, hence $w = 0$. ☐

Using Corollary 4.1, for any point $x^{(0)} \in X$ there is a point $x^{(t)} \in X$ such that $\alpha^{(t)} \in [1/2, 1/h_0]$. Thus, we can assume that $\alpha^{(0)} \in [1/2, 1/h_0]$ and work with $h \leq h_0/2$, where $h_0 = c_{\min}/(2D\|c\|_1)$. We generalize next [SV16b, Lemma 5.4].

Lemma 4.12. *Suppose $x^{(t)} \in X$ such that $\alpha^{(t)} \in [1/2, 1/h_0]$, $h \leq h_0/2$ and $\varepsilon \in (0,1)$. Then, for any $k \geq h^{-1} \ln(8m\rho_A \Psi^{(0)}/\varepsilon)$ there is a non-negative feasible kernel-free vector f such that $\|x^{(t+k)} - f\|_\infty < \varepsilon/(D\gamma_A)$.*

Proof. Let $\beta^{(k)} \stackrel{\text{def}}{=} 1 - (1-h)^k$. By (22), vector $\overline{q}^{(t,k)}$ satisfies $A\overline{q}^{(t,k)} = b$ and thus Lemma 4.6 yields

$$\|x^{(t+k)} - \beta^{(k)}\overline{q}^{(t,k)}\|_\infty = (1-h)^k \cdot \|x^{(t)}\|_\infty \leq \exp\{-hk\} \cdot \Psi^{(0)} \leq \varepsilon/(8m\rho_A). \tag{27}$$

Using Corollary 4.1, we have $x^{(t+k)} \in X$ such that $\alpha^{(t+k)} \in (1/2, 1/h_0)$ for every $k \in \mathbb{N}_+$. Let $F_k = Q_k \cup P_k$, where $Q_k = \{e \in [m] : \overline{q}_e^{(t,k)} \leq 0\}$ and $P_k = \{e \in [m] : A_e^T \overline{p}^{(t,k)} \leq 0\}$. Then, for every $e \in Q_k$ it holds

$$|\overline{q}_e^{(t,k)}| \leq [\beta^{(k)}]^{-1} \cdot |x_e^{(t+k)} - \beta^{(k)}\overline{q}_e^{(t,k)}| \leq \varepsilon/(7m\rho_A). \tag{28}$$

By Lemma 4.6, $\|x^{(\cdot)}\| \leq \Psi^{(0)}$. Moreover, by (22) for every $e \in P_k$ we have

$$\begin{aligned} x_e^{(t+k)} &= x_e^{(t)} \prod_{i=t}^{k+t-1} \left[1 + h\left(c_e^{-1} A_e^T p^{(i)} - 1\right)\right] \\ &\leq x_e^{(t)} \cdot \exp\left\{-hk + (h/c_e) \cdot A_e^T \overline{p}^{(t,k)}\right\} \\ &\leq \exp\{-hk\} \cdot \Psi^{(0)} \\ &\leq \varepsilon/(8m\rho_A), \end{aligned}$$

and by combining the triangle inequality with (27), it follows for every $e \in P_k$ that

$$\begin{aligned} |\overline{q}_e^{(t,k)}| &\leq [\beta^{(k)}]^{-1} \cdot \left[|x_e^{(t+k)} - \beta^{(k)}\overline{q}_e^{(t,k)}| + |x_e^{(t+k)}|\right] \\ &\leq [\beta^{(k)}]^{-1} \cdot \varepsilon/(4m\rho_A) \\ &\leq \varepsilon/(3m\rho_A). \end{aligned} \tag{29}$$

Therefore, (28) and (29) yields that

$$\sum_{e \in F_k} |\overline{q}_e^{(t,k)}| \leq m \cdot \varepsilon/(3m\rho_A) \leq \varepsilon/(3\rho_A). \tag{30}$$

By Lemma 4.11 applied with $\overline{q}_e^{(t,k)}$ and $N = F_k$, it follows by (30) that there is a non-negative feasible kernel-free vector f such that $\text{supp}(f) \subseteq E \backslash N$ and

$$\|f - \overline{q}^{(t,k)}\|_\infty < \varepsilon/(3D\gamma_A).$$

By Lemma 4.5, we have $\|\overline{q}^{(t,k)}\|_\infty \leq mD^2\|b/\gamma_A\|_1$ and since $\Psi^{(0)} \geq mD^2\|b/\gamma_A\|_1$, it follows that

$$\begin{aligned} \|x^{(t+k)} - f\|_\infty &= \|x^{(t+k)} - \beta^{(k)}\overline{q}^{(t+k)} + \beta^{(k)}\overline{q}^{(t+k)} - f\|_\infty \\ &\leq \|x^{(t+k)} - \beta^{(k)}\overline{q}^{(t+k)}\|_\infty + \|f - \overline{q}^{(t+k)}\|_\infty + (1-h)^k \|\overline{q}^{(t+k)}\|_\infty \\ &\leq \frac{\varepsilon}{8m\rho_A} + \frac{\varepsilon}{3D\gamma_A} + \frac{\varepsilon \cdot mD^2\|b/\gamma_A\|_1}{8m\rho_A \cdot \Psi^{(0)}} \leq \frac{\varepsilon}{D\gamma_A}. \end{aligned}$$

☐

4.5 $x^{(k)}$ is ε-Close to an Optimal Solution

Recall that $\mathcal{N}$ denotes the set of non-optimal basic feasible solutions of (5) and $\Phi = \min_{g\in\mathcal{N}} c^T g - \text{opt}$. For completeness, we prove next a well known inequality [PS82, Lemma 8.6] that lower bounds the value of Φ.

Lemma 4.13. *Suppose $A \in \mathbb{R}^{n\times m}$ has full row rank, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}^m$ are integral. Then, $\Phi \geq 1/(D\gamma_A)^2$.*

Proof. Let $g = (g_B, 0)$ be an arbitrary basic feasible solution with basis matrix A_B, where $g_B(e) \neq 0$ and $|\text{supp}(g_B)| = n$. We write $M_{-i,-j}$ to denote the matrix M with deleted i-th row and j-th column. Let Q_e be the matrix formed by replacing the e-th column of A_B by the column vector b. Then, by Cramer's rule, we have

$$|g_B(e)| = \left|\frac{\det(Q_e)}{\det(A_B)}\right| = \frac{1}{\gamma_A}\left|\sum_{k=1}^{n}\frac{(-1)^{j+k}\cdot b_k\cdot\det\left(\gamma_A^{-1}[A_B]_{-k,-j}\right)}{\det\left(\gamma_A^{-1}A_B\right)}\right| \geq \frac{1}{D\gamma_A}.$$

Note that all components of vector g_B have denominator with equal value, i.e. $\det(A_B)$. Consider an arbitrary non-optimal basic feasible solution g and an optimal basic feasible solution $f^\star$. Then, $g_e = G_e/G$ and $f^\star_e = F_e/F$ are rationals such that $G_e, G, F_e, F \leq D\gamma_A$ for every e. Further, let $r_e = c_e\left(G_e F - F_e F\right) \in \mathbb{Z}$ for every $e \in [m]$, and observe that

$$c^T\left(g - f^\star\right) = \sum_e c_e\left(g_e - f^\star_e\right) = \frac{1}{GF}\sum_e r_e \geq 1/(D\gamma_A)^2,$$

where the last inequality follows by $c^T(g - f^\star) > 0$ implies $\sum_e r_e \geq 1$. □

Lemma 4.14. *Let f be a non-negative feasible kernel-free vector and $\varepsilon \in (0,1)$ a parameter. Suppose for every non-optimal basic feasible solution g, there exists an index $e \in [m]$ such that $g_e > 0$ and $f_e < \varepsilon/(2mD^3\gamma_A\|b\|_1)$. Then, $\|f - f^\star\|_\infty < \varepsilon/(D\gamma_A)$ for some optimal $f^\star$.*

Proof. Let $C = 2D^2\|b\|_1$. Since f is kernel-free, by Lemma 2.1 it can be expressed as a convex combination of sign-compatible basic feasible solutions $f = \sum_{i=1}^{\ell}\alpha_i f^{(i)} + \sum_{i=\ell+1}^{m}\alpha_i f^{(i)}$, where $f^{(1)},\ldots,f^{(\ell)}$ denote the optimal solutions. By Lemma 3.1, $f^{(i)}_e > 0$ implies $f^{(i)}_e \geq 1/(D\gamma_A)$. By the hypothesis, for every non-optimal $f^{(i)}$, i.e. $i \geq \ell+1$, there exists an index $e(i) \in [m]$ such that

$$1/(D\gamma_A) \leq f^{(i)}_{e(i)} \quad\text{and}\quad f_{e(i)} < \varepsilon/(mD\gamma_A\cdot C).$$

Thus,

$$\alpha_i/(D\gamma_A) \leq \alpha_i f^{(i)}_{e(i)} \leq \sum_{j=1}^{m}\alpha_j f^{(j)}_{e(i)} = f_{e(i)} < \varepsilon/(mD\gamma_A\cdot C),$$

and hence $\sum_{i=\ell+1}^{m}\alpha_i \leq \varepsilon/C$. By Lemma 3.1, we have $\|f^{(j)}\|_\infty \leq D\|b/\gamma_A\|_1 = C/(2D\gamma_A)$ for every j. Let $\beta \geq 0$ be an arbitrary vector satisfying $\sum_{i=1}^{\ell}\beta_i = \sum_{i=\ell+1}^{m}\alpha_i$. Let $\nu_i = \alpha_i + \beta_i$ for every $i \in [\ell]$ and let $f^\star = \sum_{i=1}^{\ell}\nu_i f^{(i)}$. Then, $f^\star$ is an optimal solution and we have

$$\begin{aligned}\|f^\star - f\|_\infty &= \left\|\sum_{i=1}^{\ell}\beta_i f^{(i)} - \sum_{i=\ell+1}^{m}\alpha_i\cdot f^{(i)}\right\|_\infty \\ &\leq \max_{i\in[1:m]}\left\|f^{(i)}\right\|_\infty\cdot\left(\sum_{i=1}^{\ell}\beta_i + \sum_{i=\ell+1}^{m}\alpha_i\right) \leq \frac{2\varepsilon}{C}\cdot\frac{C}{2D\gamma_A} = \frac{\varepsilon}{D\gamma_A}.\end{aligned}$$

□

In the following lemma, we extend the analysis in [SV16b, Lemma 5.6] from the transshipment problem to positive linear programs. Our result crucially relies on an argument that uses the parameter $\Phi = \min_{g\in\mathcal{N}} c^T g - \text{opt}$. It is here, where our analysis incurs the linear step size dependence on Φ/opt and the quadratic dependence on opt/Φ for the number of steps.

An important technical detail is that the first regime incurs an extra (Φ/opt)-factor dependence. At first glance, this might seem unnecessary due to Corollary 4.1, however a careful analysis shows its necessity (see (35) for the inductive argument). Further, we note that the undirected Physarum dynamics (7) satisfies $x_{\min}^{(t)} \geq (1-h)^t \cdot x_{\min}^{(0)}$, whereas the directed Physarum dynamics (6) might yield a value $x_{\min}^{(t)}$ which decreases with faster than exponential rate. As our analysis incurs a logarithmic dependence on $1/x_{\min}^{(0)}$, it is prohibitive to decouple the two regimes and give bounds in terms of $\log(1/x_{\min}^{(t)})$, which would be necessary as $x^{(t)}$ is the initial point of the second regime.

Lemma 4.15. *Let g be an arbitrary non-optimal basic feasible solution. Given $x^{(0)} \in X$ and its corresponding $\alpha^{(0)}$, the Physarum dynamics (6) initialized with $x^{(0)}$ runs in two regimes:*

(i) The first regime is executed when $\alpha^{(0)} \notin [1/2, 1/h_0]$ and computes a point $x^{(t)} \in X$ such that $\alpha^{(t)} \in [1/2, 1/h_0]$. In particular, if $\alpha^{(0)} < 1/2$ then $h \leq (\Phi/\mathrm{opt}) \cdot (\alpha^{(0)} h_0)^2$ and $t = 1/h$. Otherwise, if $\alpha^{(0)} > 1/h_0$ then $h \leq \Phi/\mathrm{opt}$ and $t = \lfloor \log_{1/(1-h)}[h_0(\alpha^{(0)}-1)/(1-h_0)] \rfloor$.

(ii) The second regime starts from a point $x^{(t)} \in X$ such that $\alpha^{(t)} \in [1/2, 1/h_0]$, it has step size $h \leq (\mathrm{opt}/\Phi) \cdot h_0^2/2$ and for any $k \geq 4 \cdot c^T g/(h\Phi) \cdot \ln(\Psi^{(0)}/\varepsilon x_{\min}^{(0)})$, guarantees the existence of an index $e \in [m]$ such that $g_e > 0$ and $x_e^{(t+k)} < \varepsilon$.

Proof. Similar to the work of [BBD+13, SV16b], we use a potential function that takes as input a basic feasible solution g and a step number ℓ, and is defined by

$$\mathcal{B}_g^{(\ell)} := \sum_{e \in [m]} g_e c_e \ln x_e^{(\ell)}.$$

Since $x_e^{(\ell+1)} = x_e^{(\ell)}(1 + h^{(\ell)}[c_e^{-1} \cdot A_e^T p^{(\ell)} - 1])$, we have

$$\begin{aligned} \mathcal{B}_g^{(\ell+1)} - \mathcal{B}_g^{(\ell)} &= \sum_e g_e c_e \ln \frac{x_e^{(\ell+1)}}{x_e^{(\ell)}} = \sum_e g_e c_e \ln \left(1 + h^{(\ell)} \left[\frac{A_e^T p^{(\ell)}}{c_e} - 1\right]\right) \\ &\leq h^{(\ell)} \sum_e g_e c_e \left[\frac{A_e^T p^{(\ell)}}{c_e} - 1\right] = h^{(\ell)} \left[-c^T g + [p^{(\ell)}]^T A g\right] = h^{(\ell)} \left[-c^T g + b^T p^{(\ell)}\right]. \end{aligned} \tag{31}$$

Let $f^\star$ be an optimal solution to (5). In order to lower bound $\mathcal{B}_{f^\star}^{(\ell+1)} - \mathcal{B}_{f^\star}^{(\ell)}$, we use the inequality $\ln(1+x) \geq x - x^2$, for all $x \in [-\frac{1}{2}, \frac{1}{2}]$. Then, we have

$$\begin{aligned} \mathcal{B}_{f^\star}^{(\ell+1)} - \mathcal{B}_{f^\star}^{(\ell)} &= \sum_e f_e^\star c_e \ln \frac{x_e^{(\ell+1)}}{x_e^{(\ell)}} = \sum_e f_e^\star c_e \ln \left(1 + h^{(\ell)} \left[\frac{A_e^T p^{(\ell)}}{c_e} - 1\right]\right) \\ &\geq \sum_e f_e^\star c_e \left(h^{(\ell)} \left[\frac{A_e^T p^{(\ell)}}{c_e} - 1\right] - [h^{(\ell)}]^2 \left[\frac{A_e^T p^{(\ell)}}{c_e} - 1\right]^2\right) \\ &\geq h^{(\ell)} \left(b^T p^{(\ell)} - \mathrm{opt} - h^{(\ell)} \cdot (1/2\alpha^{(\ell)} h_0)^2 \cdot \mathrm{opt}\right), \end{aligned} \tag{32}$$

where the last inequality follows by combining

$$\sum_e f_e^\star c_e [(c_e^{-1} A_e^T p^{(\ell)}) - 1] = [p^{(\ell)}]^T A f^\star - \mathrm{opt} = b^T p^{(\ell)} - \mathrm{opt},$$

$\|A^T p^{(\ell)}\|_\infty \leq D\|c\|_1/\alpha^{(\ell)}$ (by Lemma 4.4 and Lemma 4.8 applied with $x^{(\ell)} \in X$), $h_0 = c_{\min}/(4D\|c\|_1)$ and

$$h^{(\ell)} \sum_e f_e^\star c_e \cdot (c_e^{-1} A_e^T p^{(\ell)} - 1)^2 \leq h^{(\ell)} (2D\|c\|_1/\alpha^{(\ell)} c_{\min})^2 \mathrm{opt} = h^{(\ell)} (1/2\alpha^{(\ell)} h_0)^2 \mathrm{opt}.$$

Further, by combining (31), (32), $c^T g - \mathrm{opt} \geq \Phi$ for every non-optimal basic feasible solution g and provided that the inequality $h^{(\ell)}(1/2\alpha^{(\ell)} h_0)^2 \mathrm{opt} \leq \Phi/2$ holds, we obtain

$$\mathcal{B}_{f^\star}^{(\ell+1)} - \mathcal{B}_{f^\star}^{(\ell)} \geq h^{(\ell)} \left(b^T p^{(\ell)} - c^T g\right) + h^{(\ell)} \left(c^T g - \mathrm{opt} - \frac{\Phi}{2}\right) \geq \mathcal{B}_g^{(\ell+1)} - \mathcal{B}_g^{(\ell)} + \frac{h^{(\ell)} \Phi}{2}. \tag{33}$$

Using Corollary 4.1, we partition the Physarum dynamics (6) execution into three regimes, based on $\alpha^{(0)}$. For every $i \in \{1,2,3\}$, we show next that the i-th regime has a fixed step size $h^{(\ell)} = h_i$ such that $h^{(\ell)}(1/2\alpha^{(\ell)}h_0)^2\text{opt} \leq \Phi/2$, for every step ℓ in this regime.

By Lemma 4.9, for every $i \in \{1,2,3\}$ it holds for every step ℓ in the i-th regime that

$$\alpha^{(\ell)} = 1 - (1-h_i)^{\ell} \cdot (1 - \alpha^{(0)}). \tag{34}$$

Case 1: Suppose $\alpha^{(0)} > 1/h_0$. Notice that $h^{(\ell)} = \Phi/\text{opt}$ suffices, since $1/(2\alpha^{(\ell)}h_0) < 1/2$ for every $\alpha^{(\ell)} > 1/h_0$. Further, by applying (34) with $\alpha^{(t)} := 1/h_0$, we have $t = \lfloor \log_{1/(1-h^{(\ell)})}[h_0(\alpha^{(0)}-1)/(1-h_0)]\rfloor \leq (\text{opt}/\Phi) \cdot \log(\alpha^{(0)}h_0)$. Note that by (34) the sequence $\{\alpha^{(\ell)}\}_{\ell \leq t}$ is decreasing, and by Corollary 4.1 we have $1 < \alpha^{(t)} \leq 1/h_0$.

Case 2: Suppose $\alpha^{(0)} \in (0, 1/2)$. By (34) the sequence $\{\alpha^{(\ell)}\}_{\ell\in\mathbb{N}}$ is increasing and by Corollary 4.1 the regime is terminated once $\alpha^{(\ell)} \in [1/2, 1)$. Observe that $h^{(\ell)} = (\Phi/\text{opt}) \cdot (\alpha^{(0)}h_0)^2$ suffices, since $\alpha^{(0)} \leq \alpha^{(\ell)}$. Then, by (34) applied with $\alpha^{(t)} := 1/2$, this regime has at most $t = (\text{opt}/\Phi) \cdot (1/\alpha^{(0)}h_0)^2$ steps.

Case 3: Suppose $\alpha^{(0)} \in [1/2, 1/h_0]$. By (34) the sequence $\{\alpha^{(\ell)}\}_{\ell\in\mathbb{N}}$ converges to 1 (decreases if $\alpha^{(0)} \in (1, 1/h_0]$ and increases when $\alpha^{(0)} \in [1/2, 1)$. Notice that $h^{(\ell)} = (\Phi/\text{opt}) \cdot h_0^2/2$ suffices, since $1/2 \leq \alpha^{(\ell)} \leq 1/h_0$ for every $\ell \in \mathbb{N}$. We note that the number of steps in this regime is to be determined soon.

Hence, we conclude that inequality (33) holds. Further, using Case 1 and Case 2 there is an integer $t \in \mathbb{N}$ such that $\alpha^{(t)} \in [1/2, 1/h_0]$. Let $k \in \mathbb{N}$ be the number of steps in Case 3, and let $h := (\Phi/\text{opt}) \cdot h_0^2/2$. Then, for every $\ell \in \{t, \ldots, t+k-1\}$ it holds that $h^{(\ell)} = h$ and thus

$$\mathcal{B}_{f^\star}^{(t+k)} - \mathcal{B}_{f^\star}^{(0)} \geq \mathcal{B}_{g}^{(t+k)} - \mathcal{B}_{g}^{(0)} + \sum_{\ell=0}^{t+k-1} \frac{h^{(\ell)}\Phi}{2} \geq \mathcal{B}_{g}^{(t+k)} - \mathcal{B}_{g}^{(0)} + k \cdot \frac{h\Phi}{2}. \tag{35}$$

By Lemma 4.6, $\mathcal{B}_g^{(\ell)} \leq c^T g \cdot \ln \Psi^{(0)}$ for every basic feasible solution g and every $\ell \in \mathbb{N}$, and thus

$$\begin{aligned}
\mathcal{B}_g^{(t+k)} &\leq -k \cdot \frac{h\Phi}{2} + \mathcal{B}_g^{(0)} + \mathcal{B}_{f^\star}^{(t+k)} - \mathcal{B}_{f^\star}^{(0)} \\
&\leq -k \cdot \frac{h\Phi}{2} + c^T g \cdot \ln \Psi^{(0)} + \text{opt} \cdot \ln \Psi^{(0)} - \text{opt} \cdot \ln x_{\min}^{(0)} \\
&\leq -k \cdot \frac{h\Phi}{2} + 2c^T g \cdot \ln \frac{\Psi^{(0)}}{x_{\min}^{(0)}}.
\end{aligned}$$

Suppose for the sake of a contradiction that for every $e \in [m]$ with $g_e > 0$ it holds $x_e^{(t+k)} > \varepsilon$. Then, $\mathcal{B}_g^{(t+k)} > c^T g \cdot \ln \varepsilon$ yields $k < 4 \cdot c^T g/(h\Phi) \cdot \ln(\Psi^{(0)}/(\varepsilon x_{\min}^{(0)}))$, a contradiction to the choice of k. □

4.6 Proof of Theorem 4.2

By Corollary 4.1 and Lemma 4.15, if $x^{(0)} \in X$ such that $\alpha^{(0)} > 1/h_0$, we work with $h \leq \Phi/\text{opt}$ and after $t = \lfloor \log_{1/(1-h)}[h_0(\alpha^{(0)}-1)/(1-h_0)]\rfloor \leq (\text{opt}/\Phi) \cdot \log(\alpha^{(0)}h_0)$ steps, we obtain $x^{(t)} \in X$ such that $\alpha^{(t)} \in (1, 1/h_0]$. Otherwise, if $\alpha^{(0)} \in (0, 1/2)$ we work with $h \leq (\Phi/\text{opt}) \cdot (\alpha^{(0)}h_0)^2$ and after $t = 1/h$ steps, we obtain $x^{(t)} \in X$ such that $\alpha^{(t)} \in [1/2, 1)$. Hence, we can assume that $\alpha^{(t)} \in [1/2, 1/h_0]$ and set $h \leq (\Phi/\text{opt}) \cdot h_0^2$. Then, the Lemmas in Subsection 4.4 and 4.5 are applicable.

Let $E_1 := D\|b/\gamma_A\|_1\|c\|_1$, $E_2 := 8m\rho_A\Psi^{(0)}$, $E_3 := 2mD^3\gamma_A\|b\|_1$ and $E_4 := 8mD^2\|b\|_1$. Consider an arbitrary non-optimal basic feasible solution g.

By Lemma 3.1, we have $c^T g \leq E_1$ and thus both Lemma 4.12 and Lemma 4.15 are applicable with h, $\varepsilon^\star := \varepsilon/E_4$ and any $k \geq k_0 := 4E_1/(h\Phi) \cdot \ln[(E_2/\min\{1, x_{\min}^{(0)}\}) \cdot (D\gamma_A/\varepsilon^\star)]$. Hence, by Lemma 4.15, the Physarum dynamics (6) guarantees the existence of an index $e \in [m]$ such that $g_e > 0$ and $x_e^{(t+k)} < \varepsilon^\star/(D\gamma_A)$. Moreover, by Lemma 4.12 there is a non-negative feasible kernel-free vector f such that $\|x^{(t+k)} - f\|_\infty < \varepsilon^\star/(D\gamma_A)$. Thus, for the index e it follows that $g_e > 0$ and $f_e < 2\varepsilon^\star/D\gamma_A = (\varepsilon/2) \cdot (4/E_4 D\gamma_A) = \varepsilon/(2E_3)$. Then, Lemma 4.14, yields $\|f - f^\star\|_\infty < \varepsilon/(2D\gamma_A)$ and by triangle inequality we have $\|x^{(k)} - f^\star\|_\infty < \varepsilon/(D\gamma_A)$.

By construction, $\rho_A = \max\{D\gamma_A,\, nD^2\|A\|_\infty\} \le nD^2\gamma_A\|A\|_\infty$. Let $E_2' = 8mnD^2\gamma_A\|A\|_\infty\Psi^{(0)}$ and $E_5 = E_2'E_4 \cdot D\gamma_A = 8^2m^2nD^5\gamma_A^2\|A\|_\infty\|b\|_1$. Further, let $C_1 = E_1$ and $C_2 = E_5$. Then, the statement follows for any $k \ge k_1 := 4C_1/(h\Phi)\cdot\ln(C_2\Psi^{(0)}/(\varepsilon\cdot\min\{1, x_{\min}^{(0)}\}))$. □

4.7 Preconditioning

What can be done if the initial point is not strongly dominating? For the transshipment problem it suffices to add an edge of high capacity and high cost from every source node to every sink node [BBD+13, SV16b]. This will make the instance strongly dominating and will not affect the optimal solution.

In this section, we generalize this observation to positive linear programs. We add an additional column equal to b and give it sufficiently high capacity and cost. This guarantees that the resulting instance is strongly dominating and the optimal solution remains unaffected. We state now our algorithmic result.

Theorem 4.16. *Given an integral LP $(A,\, b,\, c>0)$, a positive $x^{(0)} \in \mathbb{R}^m$ and a parameter $\varepsilon \in (0,1)$. Let $([A\,|\,b],\, b,\, (c,\, c'))$ be an extended LP with $c' = 2C_1$ and $z^{(0)} := 1 + D_S\|x\|_\infty\|A\|_1\|b\|_1$.*[8] *Then, $(x^{(0)}; z^{(0)})$ is a strongly dominating starting point of the extended problem such that $y^T[A\,|\,b](x^{(0)},\, z^{(0)}) \ge 1$, for all $y \in Y$. In particular, the Physarum dynamics (6) initialized with $(x^{(0)},\, z^{(0)})$ and a step size $h \le h_0^2/C_3$, outputs for any $k \ge 4C_1\cdot(D\gamma_A)^2/h\cdot\ln(C_2\Upsilon^{(0)}/(\varepsilon\cdot\min\{1, x_{\min}^{(0)}\}))$ a vector $(x^{(k)}, z^{(k)}) > 0$ such that $\mathrm{dist}(x^{(k)}, X_\star) < \varepsilon/(D\gamma_A)$ and $z^{(k)} < \varepsilon/(D\gamma_A)$, where here $\Upsilon^{(0)} := \max\{\Psi^{(0)}, z^{(0)}\}$.*

Theorem 4.16 subsumes [SV16b, Theorem 1.2] for flow problems by giving a tighter asymptotic convergence rate, since for the transshipment problem A is a totally unimodular matrix and satisfies $D = D_S = 1$, $\gamma_A = 1$, $\|A\|_\infty = 1$ and $\Phi = 1$. We note that the scalar $z^{(0)}$ depends on the scaled determinant D_S, see Theorem 1.3.

4.7.1 Proof of Theorem 4.16

In the extended problem, we concatenate to matrix A a column equal to b such that the resulting constraint matrix becomes $[A\,|\,b]$. Let c' be the cost and let x' be the initial capacity of the newly inserted constraint column. We will determine c' and x' in the course of the discussion. Consider the dual of the max-flow like LP for the extended problem. It has an additional variable z' and reads

$$\min\left\{\, x^Tz + x'z' \,:\, z \ge 0;\ z' \ge 0;\ z \ge A^Ty;\ z' \ge b^Ty;\ b^Ty = 1 \,\right\}.$$

In any optimal solution, $z' = b^Ty = 1$ and hence the dual is equivalent to

$$\min\left\{\, x^Tz + x' \,:\, z \ge 0;\ z \ge A^Ty;\ b^Ty = 1 \,\right\}. \tag{36}$$

The strongly dominating set of the extended problem is therefore equal to

$$X = \left\{ \begin{pmatrix} x \\ x' \end{pmatrix} \in \mathbb{R}^{m+1}_{>0} \,:\, y^T[A\,|\,b]\begin{pmatrix} x \\ x' \end{pmatrix} > 0 \text{ for all } y \in Y \right\}. \tag{37}$$

The defining condition translates into $x' > -y^TAx$ for all $y \in Y$. We summarize the discussion in the following Lemma.

Lemma 4.17. *Given a positive $x \in \mathbb{R}^m$, let $\rho := \|b\|_1 D_S$ and $x' := 1 + \rho\|A\|_1\|x\|_\infty$, where $\|A\|_1 := \sum_{i,j}|A_{i,j}|$ and $D_S := \max\{|\det(A')| \,:\, A' \text{ is a square sub-matrix of } A\}$. Then, $(x;\ x')$ is a strongly dominating starting point of the extended problem such that $y^T[A\,|\,b](x;\ x') = y^TAx + x' \ge 1$, for all $y \in Y$.*

Proof. We show first that $\max_{y\in Y}\|y\|_\infty \le \rho$ implies the statement. Let $y \in Y$ be arbitrary. Since $|y^TAx| \le \|A\|_1\|x\|_\infty\|y\|_\infty$, we have $\max_{y\in Y}|y^TAx| \le \rho\|A\|_1\|x\|_\infty = x' - 1$ and hence $y^T[A\,|\,b](x;\ x') \ge 1$.

It remains to show that $\max_{y\in Y}\|y\|_\infty \le \rho$. The constraint polyhedron of the dual (36) is given in matrix notation as

$$P^{(ext)} := \left\{ \begin{pmatrix} z \\ y \end{pmatrix} \in \mathbb{R}^{m+n} \,:\, \begin{bmatrix} I_{m\times m} & -A^T \\ \mathbf{0}_m^T & b^T \\ I_{m\times m} & \mathbf{0}_{m\times n} \end{bmatrix} \begin{pmatrix} z \\ y \end{pmatrix} \begin{matrix} \ge \\ = \\ \ge \end{matrix} \begin{pmatrix} \mathbf{0}_m \\ 1 \\ \mathbf{0}_m \end{pmatrix} \right\}.$$

[8] We denote by $\|A\|_1 := \sum_{i,j}|A_{i,j}|$, i.e. we interpret matrix A as a vector and apply to it the standard ℓ_1 norm.

Let us denote the resulting constraint matrix and vector by $M \in \mathbb{R}^{2m+1 \times m+n}$ and $d \in \mathbb{R}^{2m+1}$, respectively.

Note that if $b = \mathbf{0}$ then the primal LP (23) is either unbounded or infeasible. Hence, we consider the non-trivial case when $b \neq \mathbf{0}$. Observe that the polyhedron $P^{(ext)}$ is not empty, since for any y such that $b^T y = 1$ there is $z = \max\{0, A^T y\}$ satisfying $(z; y) \in P^{(ext)}$. Further, $P^{(ext)}$ does not contain a line (see Subsection 4.3) and thus $P^{(ext)}$ has at least one extreme point $p' \in P^{(ext)}$. As the dual LP (24) has a bounded value (the target function is lower bounded by 0) and an extreme point exists ($p' \in P^{(ext)}$), the optimum is attained at an extreme point $p \in P^{(ext)}$. Moreover, as every extreme point is a basic feasible solution and matrix M has linearly independent columns (A has full row rank), it follows that p has $m + n$ tight linearly independent constraints.

Let $M_{B(p)} \in \mathbb{R}^{m+n \times m+n}$ be the basis submatrix of M satisfying $M_{B(p)} p = d_{B(p)}$. Since A, b are integral and $M_{B(p)}$ is invertible, using Laplace expansion we have $1 \leq |\det(M_{B(p)})| \leq \|b\|_1 D_S = \rho$. Let Q_i denotes the matrix formed by replacing the i-th column of $M_{B(p)}$ by the column vector $d_{B(p)}$. Then, by Cramer's rule, it follows that $|y_i| = |\det(Q_i)/\det(M_{B(p)})| \leq |\det(M_{B(p)})| \leq \rho$, for all $i \in [n]$. $\square$

It remains to fix the cost of the new column. Using Lemma 3.1, $\mathrm{opt} \leq c^T x^{(k)} \leq C_1$ for every $k \in \mathbb{N}$, and thus we set $c' := 2C_1$.

4.8 A Simple Lower Bound

Building upon [SV16b, Lemma B.1], we give a lower bound on the number of steps required for computing an ε-approximation to the optimum shortest path. In particular, we show that for the Physarum dynamics (6) to compute a point $x^{(k)}$ such that $\mathrm{dist}(x^{(k)}, X_\star) < \varepsilon$, the required number of steps k has to grow linearly in $\mathrm{opt}/(h\Phi)$ and $\ln(1/\varepsilon)$.

Theorem 4.18. *Let (A, b, c) be a positive LP instance such that $A = [1\ 1]$, $b = 1$ and $c = (\mathrm{opt}\ \mathrm{opt} + \Phi)^T$, where $\mathrm{opt} > 0$ and $\Phi > 0$. Then, for any $\varepsilon \in (0, 1)$ the discrete directed Physarum dynamics (6) initialized with $x^{(0)} = (1/2, 1/2)$ and any step size $h \in (0, 1/2]$, requires at least $k = (1/2h) \cdot \max\{\mathrm{opt}/\Phi, 1\} \cdot \ln(2/\varepsilon)$ steps to guarantee $x_1^{(k)} \geq 1 - \varepsilon$, $x_2^{(k)} \leq \varepsilon$. Moreover, if $\varepsilon \leq \Phi/(2\mathrm{opt})$ then $c^T x^{(k)} \geq (1 + \varepsilon)\mathrm{opt}$ as long as $k \leq (1/2h) \cdot \max\{\mathrm{opt}/\Phi, 1\} \cdot \ln(2\Phi/(\varepsilon \cdot \mathrm{opt}))$.*

Proof. Let $c_1 = \mathrm{opt}$ and $c_2 = \gamma\mathrm{opt}$, where $\gamma = 1 + \Phi/\mathrm{opt}$. We first derive closed-form expressions for $x_1^{(k)}$, $x_2^{(k)}$, and $x_1^{(k)} + x_2^{(k)}$. Let $s^{(k)} = \gamma x_1^{(k)} + x_2^{(k)}$. For any $k \in \mathbb{N}$, we have $q_1^{(k)} + q_2^{(k)} = 1$ and $q_1^{(k)}/q_2^{(k)} = (x_1^{(k)}/c_1)/(x_2^{(k)}/c_2) = \gamma x_1^{(k)}/x_2^{(k)}$. Therefore, $q_1^{(k)} = \gamma x_1^{(k)}/s^{(k)}$ and $q_2^{(k)} = x_2^{(k)}/s^{(k)}$, and hence

$$x_1^{(k)} = (1 + h(-1 + \gamma/s^{(k-1)}))x_1^{(k-1)} \quad \text{and} \quad x_2^{(k)} = (1 + h(-1 + 1/s^{(k-1)}))x_2^{(k-1)}. \tag{38}$$

Further, $x_1^{(k)} + x_2^{(k)} = (1 - h)(x_1^{(k-1)} + x_2^{(k-1)}) + h$, and thus by induction $x_1^{(k)} + x_2^{(k)} = 1$ for all $k \in \mathbb{N}$.

Therefore, $s^{(k)} \leq \gamma$ for all $k \in \mathbb{N}$ and hence $x_1^{(k)} \geq x_1^{(k-1)}$, i.e. the sequence $\{x_1^{(k)}\}_{k \in \mathbb{N}}$ is increasing and the sequence $\{x_2^{(k)}\}_{k \in \mathbb{N}}$ is decreasing. Moreover, since $h(-1 + 1/s^{(k-1)}) \geq h(1 - \gamma)/\gamma = -h\Phi/(\mathrm{opt} + \Phi)$ and using the inequality $1 - z \geq e^{-2z}$ for every $z \in [0, 1/2]$, it follows by (38) and induction on k that

$$x_2^{(k)} \geq \left(1 - \frac{h\Phi}{\mathrm{opt} + \Phi}\right)^k x_2^{(0)} \geq \frac{1}{2} \exp\left\{-k \cdot \frac{2h\Phi}{\mathrm{opt} + \Phi}\right\}.$$

Thus, $x_2^{(k)} \geq \varepsilon$ whenever $k \leq (1/2h) \cdot (\mathrm{opt}/\Phi + 1) \cdot \ln(2/\varepsilon)$. This proves the first claim.

For the second claim, observe that $c^T x^{(k)} = \mathrm{opt} \cdot x_1^{(k)} + \gamma\mathrm{opt} \cdot x_2^{(k)} = \mathrm{opt} \cdot (1 + (\gamma - 1)x_2^{(k)})$. This is greater than $(1+\varepsilon)\mathrm{opt}$ iff $x_2^{(k)} \geq \varepsilon \cdot \mathrm{opt}/\Phi$. Thus, $c^T x^{(k)} \geq (1+\varepsilon)\mathrm{opt}$ as long as $k \leq (1/2h) \cdot (\mathrm{opt}/\Phi + 1) \cdot \ln(2/(\varepsilon \cdot \mathrm{opt}/\Phi))$. $\square$

5 Acknowledgements

The author Pavel Kolev is funded by the Cluster of Excellence "Multimodal Computing and Interaction" within the Excellence Initiative of the German Federal Government.

References

[BBD+13] Luca Becchetti, Vincenzo Bonifaci, Michael Dirnberger, Andreas Karrenbauer, and Kurt Mehlhorn. Physarum can compute shortest paths: Convergence proofs and complexity bounds. In *ICALP*, volume 7966 of *LNCS*, pages 472–483, 2013.

[BMV12] Vincenzo Bonifaci, Kurt Mehlhorn, and Girish Varma. Physarum can compute shortest paths. *Journal of Theoretical Biology*, 309:121 – 133, 2012.

[Bon13] Vincenzo Bonifaci. Physarum can compute shortest paths: A short proof. *Inf. Process. Lett.*, 113(1-2):4–7, 2013.

[Bon15] Vincenzo Bonifaci. A revised model of network transport optimization in Physarum Polycephalum. November 2015.

[Bon16] Vincenzo Bonifaci. On the convergence time of a natural dynamics for linear programming. *CoRR*, abs/1611.06729, 2016.

[CDS98] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[CLSW98] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.

[IJNT11] Kentaro Ito, Anders Johansson, Toshiyuki Nakagaki, and Atsushi Tero. Convergence properties for the physarum solver. arXiv:1101.5249v1, January 2011.

[JZ12] A. Johannson and J. Zou. A slime mold solver for linear programming problems. In *CiE*, pages 344–354, 2012.

[LaS76] J. B. LaSalle. *The Stability of Dynamical Systems*. SIAM, 1976.

[MO08] T. Miyaji and Isamu Ohnishi. Physarum can solve the shortest path problem on Riemannian surface mathematically rigourously. *International Journal of Pure and Applied Mathematics*, 47:353–369, 2008.

[NYT00] T. Nakagaki, H. Yamada, and Á. Tóth. Maze-solving by an amoeboid organism. *Nature*, 407:470, 2000.

[Phy] http://people.mpi-inf.mpg.de/~mehlhorn/ftp/SlimeAusschnitt.webm.

[PS82] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1982.

[Sch99] Alexander Schrijver. *Theory of linear and integer programming*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 1999.

[SV16a] Damian Straszak and Nisheeth K. Vishnoi. IRLS and slime mold: Equivalence and convergence. *CoRR*, abs/1601.02712, 2016.

[SV16b] Damian Straszak and Nisheeth K. Vishnoi. Natural algorithms for flow problems. In *SODA*, pages 1868–1883, 2016.

[SV16c] Damian Straszak and Nisheeth K. Vishnoi. On a natural dynamics for linear programming. In *ITCS*, pages 291–291, New York, NY, USA, 2016. ACM.

[TKN07] A. Tero, R. Kobayashi, and T. Nakagaki. A mathematical model for adaptive transport network in path finding by true slime mold. *Journal of Theoretical Biology*, pages 553–564, 2007.