Physarum Inspired Dynamics to Solve Semi-Definite Programs

Yuan Gao*
yuanagao@mpi-inf.mpg.de

Hamidreza Kamkari†
hamidrezakamkari@gmail.com

Andreas Karrenbauer*
karrenba@mpi-inf.mpg.de

Kurt Mehlhorn*
mehlhorn@mpi-inf.mpg.de

Mohammadamin Sharifi†
sharifim689@gmail.com

Abstract

Physarum Polycephalum is a slime mold that can solve shortest path problems. A mathematical model based on Physarum's behavior, known as the Physarum Directed Dynamics, can solve positive linear programs. In this paper, we present a family of Physarum-based dynamics extending the previous work and introduce a new algorithm to solve *positive* Semi-Definite Programs (SDP). The Physarum dynamics are governed by orthogonal projections (w.r.t. time-dependent scalar products) on the affine subspace defined by the linear constraints. We present a natural generalization of the scalar products used in the LP case to the matrix space for SDPs, which boils down to the linear case when all matrices in the SDP are diagonal, thus, representing an LP. We investigate the behavior of the induced dynamics theoretically and experimentally, highlight challenges arising from the non-commutative nature of matrix products, and prove soundness and convergence under mild conditions. Moreover, we consider a more abstract view on the dynamics that suggests a slight variation to guarantee unconditional soundness and convergence-to-optimality. By simulating these dynamics using suitable discretizations, one obtains numerical algorithms for solving positive SDPs, which have applications in discrete optimization, e.g., for computing the Goemans-Williamson approximation for MaxCut or the Lovász theta number for determining the clique/chromatic number in perfect graphs.

*Max Planck Institute for Informatics
†Sharif University of Technology

1 Introduction

The Physarum computing model is an analog computing model motivated by the network dynamics of the slime mold Physarum Polycephalum. In wet-lab experiments, it was observed that the slime mold is able to solve shortest path problems [1]. A mathematical model for the dynamic behavior of the slime was proposed in [2]. Their slime network model is mathematically equivalent to an electrical network with time-varying resistors that react to the amount of electrical current flowing through them.

A variant of the Physarum dynamics, the directed Physarum dynamics, is known to solve positive linear programs in standard form [3, 4]. A positive linear program seeks to minimize a linear function $c^T x$ with a positive cost vector $c \in \mathbb{R}^n_{>0}$ subject to the constraints $Ax = b$ and $x \geq 0$. Here $A \in \mathbb{R}^{m \times n}$and $b \in \mathbb{R}^m$.

Since Semi-Definite Programming (SDP) problems generalize linear programs, it is natural to ask for what subclass of SDPs this dynamics can be adapted to. Before we turn to SDPs, we briefly review the dynamics for LPs.

The Physarum dynamics can be described by an autonomous dynamical system, i.e., its evolution over time is governed by a system of ordinary differential equations of the form

$$\dot{x}(t) = v(x(t))$$

where the *velocities* $v(\cdot)$ do not explicitly depend on t. For sake of of presentation, we may simply write x instead of $x(t)$. Moreover, we first consider the special case when $Ax = b$ is satisfied at any time. The Physarum dynamics for LPs can then be written as [3, 4]:

$$\dot{x} = G(A^T(AGA^T)^{\dagger}Ax - c), \tag{1}$$

where $G = C^{-1}X$ is the **Conductance** while C and X are the diagonal matrices defined by the objective coefficient c and the current point x, respectively. Equivalently, we also have:

$$\dot{x}(t) = -(I - GA^T(AGA^T)^{\dagger}A)w, \tag{2}$$

where $w \coloneqq Gc$. Note that $(I - GA^T(AGA^T)^{\dagger}A)$ is an orthogonal projection on the kernel of A under a weighted inner product. It has been shown in [3, 4] that the trajectory of the Physarum dynamic (2) stays within the feasible region and converges to an optimal solution of the linear program. Numerous results are driven by altering the dynamics or taking a more abstract view [5, 6, 7, 8]. These results motivate the research into generalizing the LP Physarum dynamic to other more powerful convex optimization problems. In this paper, we define and analyze a family of Physarum inspired dynamics for positive SDP.

Positive SDP In general, an SDP can be written as one of the primal-dual pair

$$\begin{aligned} &\min\{tr(C^T X) : tr(A_\ell^T X) = b_\ell \forall \ell \in [m],\, X \succeq 0\} \\ &\geq \max\{\sum_{\ell=1}^m b_\ell y_\ell : \sum_{\ell=1}^m y_\ell A_\ell + S = C,\, S \succeq 0\}, \end{aligned} \tag{3}$$

where we may assume without loss of generality that C and all A_ℓ are symmetric. This is due to the fact that X is symmetric and each formula of form $tr(MX)$ is equivalent to $tr(\frac{M+M^T}{2}X)$;

furthermore, we rewrite all of these conditions in this form where C will be replaced by $\frac{C+C^T}{2}$ and each A_ℓ matrices will be replaced by $\frac{A_\ell+A_\ell^T}{2}$. We will call an SDP **positive** if the cost matrix C is positive definite. Note that in this case, strong duality always holds because Slater's condition is trivially satisfied for the dual (maximization) problem as $y = 0$, $S = C \succ 0$ is a solution in the relative interior. Moreover, if C and all the matrices A_ℓ are diagonal, then it suffices to only consider diagonal solutions X and the SDP reduces to an LP. In this case, $X \succeq 0$ is equivalent to each of the diagonal elements being non-negative; which is the same as $diag(X) \geq 0$. The matrix trace function then reduces to a dot product of the diagonal vectors. Hence, the Physarum dynamics extends to positive SDPs restricted to diagonal matrices. But does it also extend to a more substantial subclass? Since the generalization of the conductances leaves some degrees of freedom, we will present a first ansatz and the subclass of positive SDPs for which we are able to prove soundness and convergence. Moreover, we report on computational experiments that suggest that this ansatz could actually work for all positive SDPs. However, to get there we make a second ansatz that allows us to prove soundness and convergence for all positive SDPs.

Positive SDP is an important subclass of SDP, e.g., they are general enough for the Goemans-Williamson algorithm for approximating MaxCut [9] or for computing the Lovász theta number to determine the clique/chromatic number in a perfect graph [10, 11]. This might not be obvious at a first glance but there are equivalent formulations with positive definite C in both cases because a suitable multiple of the identity can be added to the cost matrix, which only causes a shift in the objective value since the diagonal of X is fixed to all ones. In fact, whenever the dual SDP has a non-empty relative interior and one can find a witness with reasonable effort, one can replace the cost matrix C with the corresponding positive definite dual slack matrix, which turns the SDP into a positive one even when C was indefinite or negative definite (as for the two examples mentioned above).

Related Work The study of Physarum dynamics were started in the mathematical biology community, in the context of shortest path [1]. Its theoretical foundations in computing has been studied in a line of works [12, 13, 14]. Other works extended the Physarum dynamics to graph and network problem beyond shortest path, including the design of transportation networks [15, 16, 17], supply-chain networks [18], and flow problems [19, 5]. Going beyond the graph setting, [3, 4] showed that the directed Physarum dynamics can solve positive linear programs in standard form. Physarum dynamics are also studied in other more abstract optimization problems, including the basis pursuit problem [20, 8].

Many works benefit from modifying the Physarum dynamics or studying them in an abstract setting. In [6] the author revised the model of [2] by considering a different controlling variable for the adaptation mechanism, and gave a generalized Physarum dynamic which subsumes several Physarum dynamics as special cases, including [7] and the two-norm dynamics [5]. In [7] they analyzed the convergence of "non-uniform" variant of the Physarum dynamics. In [5] two different Physarum dynamics were proposed and analyzed for the multi-commodity flow problem. By studying a Meta-Algorithm, [8] established a connection between the Iteratively Reweighted Least Squares (IRLS) algorithm and the Physarum dynamics.

Our contribution In this paper, we generalize the LP Physarum dynamic to SDP, by presenting a general recipe for the SDP Physarum dynamic. We show that while being general enough, our framework already guarantees some very desirable properties. We propose a natural generalization of the LP Physarum conductance matrix and hence obtain our first ansatz. Key technical challenges are identified in the analysis of first ansatz's soundness and convergence to equilibrium, under mild conditions. We present strong empirical evidence for the dynamic's unconditional convergence to optimality. With our understanding of the technical challenges posed by the first ansatz, and following our general recipe, we propose a second conductance matrix and hence obtain our second ansatz. We prove unconditional guarantee on the second ansatz's soundness, and more importantly, convergence to optimality. We believe that this showcases the potential and power of our general framework for SDP Physarum dynamic. Although our implementation was geared towards the investigation of soundness and convergence instead of optimizing for performance, we are convinced that numerical algorithms obtained from carefully discretizing the Physarum dynamics motivate further research into practical SDP solvers based on nature-inspired algorithms.

Organization of the paper In section 2 we give an overview of our contributions and explain the general framework and some of the key results. In section 3 we investigate our general recipe of SDP Physarum dynamic and prove its desirable properties. Then in section 4 we propose a natural generalization of the LP Physarum dynamic's conductance matrix and hence give our first ansatz. In section 4.1 and 4.2 we analyse the soundness and covnergence properties of the first ansatz, under mild conditions. Section 5 proposes the second ansatz and proves unconditional soundness and convergence-to-optimality results. We also highlight key differences between the two Physarum dynamics, and the technical advantages of the second ansatz which enables us to prove the strong convergence results. Section 6 presents two numerical algorithms obtained from discretizing the SDP Physarum dynamic. In section 6.3 we discuss the experimental evaluations of the two numerical algorithms, providing empirical evidence for the unconditional soundness and convergence of the first ansatz. We provide practical implementations of the solvers and demonstrate their applications in various discrete optimization problems.

2 Overview

In this section we give an overview of the SDP Physarum dynamics that are studied in this paper. The SDP Physarum dynamics are naturally inspired by previous works in LP Physarum dynamic. These dynamics are charactarized by the choices of conductance matrices in the system of ordinary differential equations describing the evolution, and we give a concrete choice of the conductance, inspired by and closely related to the LP Physarum dynamic. However, the previous tehniques do not work out of the box, due to the noncommutative nature of matrix multiplication. We show that the SDP Physarum dynamic is sound and convergent, under mild conditions on the objective. We then move on to show that a slight variation on the choice of the conductance matrix enables us to prove the soundness and convergence-to-optimality of the SDP Physarum dynamic without these mild assumptions.

For an easier comparison with the LP case note that one can also write an SDP in a vectorized notation where $\underrightarrow{M}$ for an $n \times n$ matrix M is an n^2 vector whose $n(i-1)+j$th element is $M_{i,j}$, i.e., it is obtained by stacking up the columns of M on top of each other. Using this notation, we may write $tr(A^T B) = \sum_{i,j} a_{ij} b_{ij} = \underrightarrow{A}^T \underrightarrow{B}$. Consider $\mathcal{A}$ to be an $m \times n^2$ matrix defined as

$$\mathcal{A} = \begin{bmatrix} \underrightarrow{A_1}^T \\ \underrightarrow{A_2}^T \\ \dots \\ \underrightarrow{A_m}^T \end{bmatrix}. \tag{4}$$

Now we can also write SDP in the vectorized format, which resembles that of the LP case:

$$\min\{\underrightarrow{C}^T \underrightarrow{X} \ : \ \mathcal{A} \cdot \underrightarrow{X} = b, \ X \succeq 0\} = \max\{b^T y : \mathcal{A}^T y + \underrightarrow{S} = \underrightarrow{C}, \, S \succeq 0\} \text{ for } C \succ 0$$

where we use the convention $\min \emptyset = \infty$ in the case of primal infeasibility, which is equivalent to dual unboundedness because $C \succ 0$ implies dual feasibility.

A general recipe for SDP Physarum dynamics A natural way to extend the LP Physarum dynamic (2) to the SDP case is to lift the differential equations to matrices. For simplicity, here we assume that the SDP Physarum dynamic start with a feasible point $X(0)$. In particular, we say that a point X is **linearly feasible** if it satisfies all linear constraints. See section 4.3 for a detailed explanation on how to accommodate infeasible starting points. Our SDP Physarum dynamic is defined by the following differential equation:

$$\dot{\underrightarrow{X}} = -(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger \mathcal{A})G\underrightarrow{C}, \tag{5}$$

where, again, G is some **conductance** matrix that characterizes the SDP Physarum dynamics. While the dynamic itself is straightforward to generalize, the choice of conductance is not as obvious. A trivial generalization of the conductance of LP Physarum dynamic would be the $n^2 \times n^2$ matrix with the diagonal stacking all entries of $C^{-1}X$ while the other entries zero. This, however, will not work because this choice of conductance directly yields the LP Physarum dynamic for the problem

$$\min\left\{\underrightarrow{C}^T \underrightarrow{X} : \mathcal{A}\underrightarrow{X} = b, \underrightarrow{X} \geq 0\right\},$$

which does not agree with the SDP problem because a positive semi-definite matrix may have negative entries. Moreover, for $\underrightarrow{C}$, $\underrightarrow{X}(0)$ with non-negative entries, the dynamics with that conductance matrix will keep all entries of $\underrightarrow{X}$ non-negative, and for negative entries, convergence would not be guaranteed at all.

On the other hand, using the sole requirement $G \succeq 0$ without any further specification of its structure, our general recipe can already be shown to reduce the objective values monotonically over time. Writing the objective value along the trajectory of the dynamic as a function over time as

$$\mathcal{L}(t) := tr(CX(t)) = \underrightarrow{C}^T \underrightarrow{X}(t),$$

we obtain the following Theorem.

Theorem 2.1. *Using any symmetric positive semi-definite conductance G for the SDP Physarum dynamic* (5)*, then the following is true:*

1. $\frac{d}{dt}\mathcal{L}(t) \leq 0$

2. $\frac{d}{dt}\mathcal{L}(t)$ *becomes zero if and only if* $\dot{X}(t) = 0$

3. *If* $\mathcal{L}(t)$ *is bounded from below, the dynamic converges to equilibrium, i.e.,* $\lim_{t\to\infty} \left\|\dot{X}(t)\right\| = 0$

We restate and prove this theorem as Theorem 3.3. This theorem shows that while our general recipe is general enough to accomodate essentially any positive semi-definite conductance matrix, it still offers guarantees on how the objective values behave along the dynamics' trajectories.

First Ansatz Following the general recipe, we give a natural generalization of the conductance matrix employed in the LP Physarum dynamic (2), and hence get our first SDP Physarum dynamic, which we call the *first ansatz*. We also make a statement about having infeasible starting point, see section 4.3 for how we slightly modify the dynamic to cope with infeasible starting point. Under mild assumption on the objective C of the SDP, we obtain the following results on the first ansatz:

Theorem 2.2. *(Informal) For the class of positive SDPs where* C^{-1} *is linearly feasible, our first ansatz* $G := \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$ *for the conductance in* (5) *satisfies the following:*

1. *the dynamic stays linearly feasible,*

2. *the dynamic is sound, i.e.,* $X(T) \succ 0$ *for any finite time* $T \geq 0$.

3. *the dynamic reduces the objective and reaches equilibrium, in particular,* $\frac{d}{dt}tr(CX(t)) \leq 0$ *where equality holds if only if* $\dot{X}(t) = 0$. *Moreover,* $\lim_{t\to\infty} \left\|\dot{X}(t)\right\| = 0$.

If the equilibrium points are positive definite, then statement 3 implies that the first ansatz converges to optimality. In section 6.3 we also provide empirical evidences that the dynamic converges to optimality without such assumption. In section 6.1 we introduce the augmented SDP such that the linear feasibility of C^{-1} can be satisfied. The first statement of Theorem 2.2 is formally stated and proved as Theorem 3.2. The second statement is formally stated and proved as Theorem 4.2. The third statement is Corollary 4.4 implied by Theorem 2.1.

Second Ansatz The soundness proof of Theorem 2.2 relies on a mild assumption on the objective matrix C. Even though we can augment the problem so that the assumption is satisfied and that the conductance in the first ansatz is a natural extension of the conductance in (2), we study a different choice of conductance, namely $G := X \otimes X$ and obtain our *second ansatz*. With the second ansatz, we are able to remove the mild assumption in 2.2. We summarize it as the following:

Theorem 2.3. *(Informal) For the class of positive SDPs, our second ansatz* $G := X \otimes X$ *for the conductance in* (5) *satisfies the following:*

1. *the dynamic stays linearly feasible,*

2. *the dynamic is sound, i.e.,* $X(T) \succ 0$ *and for any finite time* $T \geq 0$,

3. *the dynamic reduces the objective and reaches equilibrium, in particular,* $\frac{d}{dt}tr(CX(t)) \leq 0$ *where equality holds if only if* $\dot{X}(t) = 0$. *Moreover,* $\lim_{t\to\infty} \left\|\dot{X}(t)\right\| = 0$.

Similar to before, if the equilibrium points are positive definite, then statement 3 implies that the second ansatz converges to optimality. The first statement of Theorem 2.3 is formally stated and proved as Theorem 3.2. The second statement is formally stated and proved as Theorem 5.2 and the third statement as Corollary 5.3. However, with the second ansatz, we are able to avoid the technical obstacles emerging from analyzing the first ansatz, and prove results that are much stronger: we show that the second ansatz coincide with a central path of SDP (3) when starting with a strictly feasible point, implying that the dynamic converges to optimality unconditionally. We believe that this shows the great power and potential in our general recipe: by cleverly designing the conductance, one can design Physarum dynamics that have strong theoretical guarantees. We state this result informally as the following:

Theorem 2.4. *(Informal) Fix any feasible point $F \succ 0$, there exists a central path of SDP* (3) *that coincides with the second ansatz starting from F. Consequently, the second ansatz converges to the optimum of SDP* (3) *when starting from any feasible point $X(0) \succ 0$.*

We discuss this result thoroughly in section 5.2 and formally state and prove Theorem 2.4 as Theorem 5.4.

3 A General Recipe for SDP Physarum Dynamics

In section 2 we briefly explained that our general recipe for SDP Physarum dynamic (5) is a natural extension of the LP Physarum dynamic to SDP. One might ask, why do we need a general framework of SDP Physarum dynamics that admit any choices of positive semidefinite conductance G? As it turns out, given any concrete choice of conductance matrix G that is positive semi-definite, the SDP Physarum dynamic already satisfies two important properties: starting with a linearly feasible $X(0)$,

1. the dynamic stays linearly feasible.
2. the objective value is reduced along the trajectory of the dynamic.

This means that once a positive semidefinite conductance matrix G is chosen such that the dynamic stays within the positive definite cone, an SDP Physarum dynamic that stays feasible and reduces objective is found.

In this section, we prove these two important properties of the general recipe of SDP Phyarum dynamic (5).

Linear feasibility Geometrically speaking, given any positive definite conductance G, the matrix $-(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A})$ is the projection into the kernel of $\mathcal{A}$, under the inner product:

$$\langle a, b\rangle_{G^{-1}} \coloneqq a^T G^{-1} b$$

We will expand on this in section 4.3 and obtain an equivalent formulation of the SDP Physarum dynamics. Now, what this geometric perspective entails for us is that in our general recipe of the SDP Phsarum dynamics, $\underrightarrow{\dot{X}}(t)$ is always in the kernel of $\mathcal{A}$.

Lemma 3.1. *Given any positive semi-definite conductance G, $-(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A})G\underset{\rightarrow}{C}$ is in the kernel of $\mathcal{A}$. In particular, given the SDP Physarum dynamic* (5)*, for any time t,*

$$\mathcal{A}\dot{\underset{\rightarrow}{X}}(t) = 0$$

Proof. Let $\tilde{\mathcal{A}} := \mathcal{A}G^{\frac{1}{2}}$, there exists v and d such that

$$G^{\frac{1}{2}}\underset{\rightarrow}{C} = \tilde{\mathcal{A}}^T v + d$$

where $\tilde{\mathcal{A}}d = 0$. Now we can verify that:

$$\begin{aligned}
\mathcal{A}\dot{\underset{\rightarrow}{X}}(t) &= -\mathcal{A}(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A})G\underset{\rightarrow}{C} \\
&= -\tilde{\mathcal{A}}G^{\frac{1}{2}}\underset{\rightarrow}{C} + \tilde{\mathcal{A}}\tilde{\mathcal{A}}^T(\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T)^\dagger\tilde{\mathcal{A}}G^{\frac{1}{2}}\underset{\rightarrow}{C} \\
&= -\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T v - \tilde{\mathcal{A}}d + \tilde{\mathcal{A}}\tilde{\mathcal{A}}^T(\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T)^\dagger\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T v + \tilde{\mathcal{A}}\tilde{\mathcal{A}}^T(\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T)^\dagger\tilde{\mathcal{A}}d \\
&= -\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T v + \tilde{\mathcal{A}}\tilde{\mathcal{A}}^T(\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T)^\dagger\tilde{\mathcal{A}}\tilde{\mathcal{A}}^T v \\
&= 0
\end{aligned}$$

So indeed $\dot{\underset{\rightarrow}{X}}(t)$ is in the kernel of $\mathcal{A}$. □

This leads to our next theorem that shows $X(t)$ stays linearly feasible, if the SDP Physarum dynamic starts with a linearly feasible $X(0)$.

Theorem 3.2. *Given a linearly feasible starting point for the SDP Physarum dynamics* (5)*, and a positive semi-definite conductance G, then at any time t,*

$$\mathcal{A}\underset{\rightarrow}{X}(t) = b$$

In other words, $X(t)$ stays feasible.

Proof. First note that:

$$\frac{d}{dt}\mathcal{A}\underset{\rightarrow}{X}(t) - b = \mathcal{A}\dot{\underset{\rightarrow}{X}}(t)$$

Since $X(0)$ is linearly feasible, $\mathcal{A}\underset{\rightarrow}{X}(0) - b = 0$. By fundamental theorem of calculus, we have:

$$\begin{aligned}
\mathcal{A}\underset{\rightarrow}{X}(t) - b &= \mathcal{A}\underset{\rightarrow}{X}(0) - b + \int_0^t \mathcal{A}\dot{\underset{\rightarrow}{X}}(s)ds \\
&= 0
\end{aligned}$$

where in the last equality we applied Lemma 3.1. Therefore, $X(t)$ stays feasible. □

Objective value along the dynamic's trajectory Now we move on to show that with any positive semi-definite conductance matrix G, our general recipe of SDP Physarum dynamic reduces the objective value. We restate Theorem 2.1 and give a formal proof:

Theorem 3.3. *Using any symmetric positive semi-definite conductance G for the SDP Physarum dynamic* (5)*, then the following is true:*

1. $\frac{d}{dt}\mathcal{L}(t) \leq 0$

2. $\frac{d}{dt}\mathcal{L}(t)$ *becomes zero if and only if* $\dot{X}(t) = 0$

3. If $\mathcal{L}(t)$ *is bounded from below, the dynamic converges to equilibrium* $\lim_{t\to\infty} \left\| \dot{X}(t) \right\| = 0$

Proof. It's easy to see that $\frac{d}{dt}\mathcal{L}(t) = tr(C\dot{X}(t))$. By the definition of the SDP Physarum dynamic, we have that:

$$\frac{d}{dt}\mathcal{L}(t) = -\underrightarrow{C}^T G \underrightarrow{C} + \underrightarrow{C}^T G \mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger \mathcal{A} G \underrightarrow{C}$$

1. Write $P_G := G^{\frac{1}{2}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A}G^{\frac{1}{2}}$. It's easy to see that P_G is symmetric. Moreover, $P_G^2 = G^{\frac{1}{2}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A}G^{\frac{1}{2}}G^{\frac{1}{2}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A}G^{\frac{1}{2}} = G^{\frac{1}{2}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A}G^{\frac{1}{2}}$, i.e. $P_G^2 = P_G$. Therefore, P_G is an orthogonal projection.

   In particular, for any v, $\|v\|_2 \geq \|P_G v\|_2$, where equality holds if and only if v is an eigenvector of P_G. Hence,

$$\begin{aligned}
\frac{d}{dt}\mathcal{L}(t) &= -\underrightarrow{C}^T G \underrightarrow{C} + \underrightarrow{C}^T G^{\frac{1}{2}} P_G G^{\frac{1}{2}} \underrightarrow{C} \\
&= -\underrightarrow{C}^T G \underrightarrow{C} + \underrightarrow{C}^T G^{\frac{1}{2}} P_G P_G G^{\frac{1}{2}} \underrightarrow{C} \\
&= -\underrightarrow{C}^T G \underrightarrow{C} + \left\| P_G G^{\frac{1}{2}} \underrightarrow{C} \right\|_2^2 \\
&\leq 0,
\end{aligned}$$

   where in the inequality we used the fact that P_G is an orthogonal projection.

2. Due to the proof above, $\frac{d}{dt}\mathcal{L}(t)$ becomes zero if only if the inequality is an equality in the previous proof. In particular, this implies that $G^{\frac{1}{2}}\underrightarrow{C}$ is an eigenvector of the projection P_G. Note that P_G only has two eigenvalues, 0 and 1. If $P_G G^{\frac{1}{2}}\underrightarrow{C} = 0$, then by definition $\underrightarrow{\dot{X}}(t) = -G\underrightarrow{C}$. However, in this case we also have that

$$\frac{d}{dt}\mathcal{L}(t) = -\underrightarrow{C}^T G \underrightarrow{C} = -\left\| G^{\frac{1}{2}} \underrightarrow{C} \right\|_2^2$$

   Therefore $G^{\frac{1}{2}}\underrightarrow{C}$ must be zero as well. This implies that $\underrightarrow{\dot{X}}(t) = 0$. On the other hand, if $P_G G^{\frac{1}{2}}\underrightarrow{C} = G^{\frac{1}{2}}\underrightarrow{C}$, then

$$\begin{aligned}
\underrightarrow{\dot{X}}(t) &= -G\underrightarrow{C} + G^{\frac{1}{2}} P_G G^{\frac{1}{2}} \underrightarrow{C} \\
&= -G\underrightarrow{C} + G^{\frac{1}{2}} G^{\frac{1}{2}} \underrightarrow{C} \\
&= 0
\end{aligned}$$

3. We know that $\mathcal{L}$ at time 0 is finite and its derivative is negative. The derivative remains negative as long as $\dot{X} \neq 0$. In addition, $\mathcal{L}(t) = tr(CX(t))$ is lower bounded from below. Therefore, $\dot{X}$ must eventually converge to 0.

□

4 First Ansatz

In section 2 we introduced a general recipe for SDP Physarum dynamic (5). In this section we study the soundness and convergence of the first ansatz. For most part of this section, we focus on the case where the dynamic has feasible starting point. We address the extension of the first ansatz to infeasible starting points in section 4.3. It remains unclear what is a natural generalization of the conductance matrix $C^{-1}X$ in the LP Physarum dynamic (2). We argue that by setting

$$G := \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1}),$$

the SDP Physarum dynamic reduces to the LP Physarum dynamic in the diagonal setting. Without going over all the details, we note that the key point here is to observe that, when C, X and A_ℓ's are diagonal, then the (k, ℓ)-th entry of $\mathcal{A}G\mathcal{A}^T$ is $tr(A_k X A_\ell C^{-1})$, which equals to $a_k^T C^{-1} X a_\ell$ where a_k and a_ℓ are the LP linear constraints. In other words, $\mathcal{A}G\mathcal{A}^T = ACX^{-1}A^T$ where A is the LP constraint matrix. Also note that $\mathcal{A}G\underrightarrow{C} = b$. Observe that $\frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$ is positive (semi-)definite when X is positive (semi-)definite.

4.1 Soundness of the First Ansatz

In this section we show that, under mild conditions, if $X(t)$ is a matrix funciton defined by (5) with conductance matrix $G = \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$, then in any finite time, $X(t)$ stays feasible. Two types of conditions comprise the feasibility of $X(t)$ in the SDP:

- Linear constraints of the form $tr(A_\ell X(t)) = b_\ell$ which can be summarized in $\mathcal{A} \cdot \underrightarrow{X} = b$.
- Positive semi-definiteness of $X(t)$, i.e., $X(t) \succeq 0$.

In section 3 we have shown that, as long as G stays positive semi-definite, our dynamic stays linearly feasible. Now it remains to show that $X(t)$ stays within the positive definite cone in any finte time.

Positive definiteness Proving that the Physarum dynamics stays within the positive definite cone is harder than proving linear feasibility. Typically one employs a **work function** type argument and shows that the work function does not explode to infinity. In [4] the work function for LP Physarum dynamic is defined to be $\mathcal{B}(t) := \sum_{i\in[n]} y_i c_i \ln x_i(t)$ where y is some feasible solution to the LP and c is the objective coefficient and $x(t)$ is the LP Physarum dynamic (2). This is an entropy type work function, and the c and y terms are introduced to counter the C^{-1} component in the conductance $C^{-1}X$. A crucial distinction between the SDP problem and the LP problem is that matrix multiplications are in general non-commutative. Consequently, differentiating a entropy type work function on matrices leads to very complicated expressions in terms of series. Handling these series would eventually require some commutativity assumption along the trajectory of the dynamic, which is unrealistic. Given the general recipe of SDP Physarum dynamics (5) and the choice of conductance $G = \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$, we define the following work function:

$$W(t) := \ln\det(X(t)) \tag{6}$$

This is also an adaptation of a well-known barrier function for the positive definite cone. We establish the following lemma connecting the positive definiteness of $X(t)$ with the derivative of the work function $W(t)$.

Lemma 4.1. *Consider any dynamics $X(t)$ with a starting point $X(0) \succ 0$. If there exists some constant μ (independent of the time t) such that*

$$\frac{d}{dt}W(t) \geq \mu,$$

then for any finite time $T \geq 0$, we have $X(T) \succ 0$

Proof. If at any time $X(t) \not\succ 0$, then by the assumption that $X(0) \succ 0$ and the continuity of the eigenvalues of $X(t)$, there must exists a finite time when $X(t)$ obtains a zero eigenvalue. Consider the first time $T > 0$ where $X(T)$ obtains a zero eigenvalue. Now we have

$$\begin{aligned}
\lim_{t \to T^-} W(t) &= W(0) + \lim_{t \to T^-} \int_{s=0}^{t} \frac{d}{ds}W(s)ds \\
&\geq W(0) + \mu T
\end{aligned}$$

However, by our assumption that $X(T)$ obtains a zero eigenvalue, $\lim_{t \to T^-} W(t)$ must be divergent and not lower bounded by any finite value. This is a contradoction and we get that for any finite time $T \geq 0, X(T) \succ 0$. $\square$

With Lemma 4.1, showing that our first ansatz, or indeed, any concrete instantiation of our general recipe of SDP Physarum dynamics, stays within the positive definite cone reduces to showing that the work function's derivative is lower bounded. This is where we have to introduce an additional condition on the objective of the SDP. We require that C^{-1} is linearly feasible, in other words, $\mathcal{A}\underrightarrow{C^{-1}} = b$. The reason is that in the LP Physarum dynamic's work function $\mathcal{B}(t)$, the entropy-like nature corrects and replaces the C^{-1} terms showing up in the derivative. Some may argue that requiring C^{-1} to be linearly feasible is too restrictive. However, in section 6.1, we present how to augment the problem so that C^{-1} is always linearly feasible. Moreover, in section 6.3, we show strong empirical evidences that the dynamic stays within the positive definite cone regardless of this mild assumption. In section 5 we show that a different choice of the conductance matrix will remove this condition on C^{-1}. Now, we lower bound the derivative of the work function for our first ansatz, and hence prove that the dynamic stays within the positive definite cone.

Theorem 4.2. *If C^{-1} and $X(0)$ are linearly feasible, then for the first ansatz where the conductance is $G = \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$, when $X \succ 0$, the derivative of the work function $W(t) = \ln\det(X(t))$ is lower bounded by a time-independent constant. Furthermore, by Lemma 4.1, the first ansatz stays in the positive definite cone at any finite time.*

Proof. First we compute the derivative of $W(t)$ [21]

$$\frac{d}{dt}W(t) = tr(X^{-1}(t)\dot{X}(t))$$

Note that $tr(X^{-1}\dot{X}(t)) = \underrightarrow{X^{-1}}^T \underrightarrow{\dot{X}}(t)$, so by the definition of the SDP Physarum dynamics (5), and that $G = \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$, we have

$$\begin{aligned}
\frac{d}{dt}W(t) &= \underrightarrow{X^{-1}}^T \underrightarrow{\dot{X}}(t) \\
&= -\underrightarrow{X^{-1}}^T (I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A})G\underrightarrow{C} \\
&= -\underrightarrow{X^{-1}}^T G\underrightarrow{C} + \underrightarrow{X^{-1}}^T G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}G\underrightarrow{C}
\end{aligned}$$

Note that [1] $G\underrightarrow{X^{-1}} = \frac{1}{2}((C^{-1} \otimes X)\underrightarrow{X^{-1}} + (X \otimes C^{-1})\underrightarrow{X^{-1}}) = \underrightarrow{C^{-1}}$, and similarly $G\underrightarrow{C} = \underrightarrow{X}$. Hence we have

$$\begin{aligned}
\frac{d}{dt}W(t) &= -\underrightarrow{C^{-1}}^T \underrightarrow{C} + \underrightarrow{C^{-1}}^T \mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}G\underrightarrow{C} \\
&= -n + \underrightarrow{C^{-1}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}\underrightarrow{X} \\
&= -n + b^T(\mathcal{A}G\mathcal{A}^T)^{-1}b \\
&\geq -n
\end{aligned}$$

where the third equality uses the condition that C^{-1} and $X(0)$ are linear feasible, and the inequality uses the fact that $\mathcal{A}G\mathcal{A}^T$, and therefore $(\mathcal{A}G\mathcal{A}^T)^{-1}$, is positive definite. $\square$

Remark 4.3. *In fact, based on the proof above, it suffices that any positive multiple of C^{-1} is linearly feasible. Note that scaling the objective does not change the solution of the problem.*

This concludes that, under the assumption that C^{-1} is linearly feasible, and with a feasible starting point $X(0) \succ 0$ the first ansatz is sound, in other word, the dynamic stays linearly feasible and within the positive definite cone.

4.2 Convergence of the First Ansatz

In this section we consider the convergence of the First Ansatz. In particular, we study how objective value behaves along the trajectory of the Physarum dynamic. Since we have established in section 4.1 that our first ansatz is sound, under the condition that C^{-1} is linearly feasible, we have the immediate corollary of Theorem 2.1:

Corollary 4.4. *Given a feasible starting point $X(0) \succ 0$, and that C^{-1} is linearly feasible, then first ansatz satisfies the following:*

1. $\frac{d}{dt}\mathcal{L}(t) \leq 0$
2. $\frac{d}{dt}\mathcal{L}(t)$ *becomes zero if and only if* $\dot{X}(t) = 0$
3. *The dynamic converges to equilibrium* $\lim_{t\to\infty} \left\|\dot{X}(t)\right\| = 0$

This corollary does not immediately imply the dynamic's convergence to optimality. The dynamic might not converge to a single point, but multiple equilibrium points. We define the following equilibrium set to charactarize them:

$$EQ_1 \coloneqq \{X : v_1(\underrightarrow{X}) = 0, \mathcal{A}\underrightarrow{X} = b, X \succeq 0\} \tag{7}$$

where we defined v_1 as the velocity corresponding to our autonomous SDP Physarum dynamic. In other words,

$$v_1(\underrightarrow{X}) = G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{\dagger}\mathcal{A}\underrightarrow{X} \text{ where } G = \frac{1}{2}\left(C^{-1} \otimes X + X \otimes C^{-1}\right).$$

[1]We frequently use the rule: $\underrightarrow{ABC} = (C^T \otimes A)\underrightarrow{B}$.

Moreover, we used the Moore-Penrose pseudo inverse to accomodate the possibility that the dynamic might approach singular points at the limit. Now we further charactarize the positive definite points in the equilibrium set EQ_1, and consequently deduce a conditional convergence-to-optimality result on the first ansatz.

Lemma 4.5. *If $X \in EQ_1$ and $X \succ 0$, then X is an optimal solution to SDP* (3)*.*

Proof. By definition (7), any point $X \in EQ_1$ is a feasible solution to SDP (3). If $X \succ 0$, we prove its optimality using week duality. Note that G is nonsingular if $X \succ 0$, therefore we have:

$$G^{-1}\underset{\rightarrow}{X} = \mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}\underset{\rightarrow}{X}$$

Write $y := (\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}\underset{\rightarrow}{X}$ and note that $G^{-1}\underset{\rightarrow}{X} = \underset{\rightarrow}{C}$, then we have that

$$\underset{\rightarrow}{C} = \mathcal{A}^T y$$

In particular, by the definition (3), $(y, 0)$ is a dual feasible solution. Moreover,

$$\begin{aligned} \underset{\rightarrow}{C}^T \underset{\rightarrow}{X} &= \underset{\rightarrow}{C}^T G\mathcal{A}^T p \\ &= \underset{\rightarrow}{X}^T \mathcal{A}^T p \\ &= b^T p \end{aligned}$$

where in the third equality we used the feasibily of X. This means that for the primal-dual pair $(X, y, 0)$, the primal and dual objective conincides. Therefore by weak duality we have that X is an optimal solution to the SDP. ☐

With Corollary 4.4 and Lemma 4.5, we can obtain the following conditional convergence-to-optimality result for our first ansatz.

Theorem 4.6. *Starting with a feasible $X(0) \succ 0$, assuming that C^{-1} is linearly feasible, if $\forall X \in EQ_1, X \succ 0$, then the first ansatz is sound and converges to the optimum of SDP* (3)*.*

We conjecture that the first ansatz converges to optimality regardless of the condition on C and EQ_1, and we summarize the conjecture as Conjecture 4.10 in section 4.3, which also captures the behavior of the dynamic with linearly infeasible starting point. We remark that in section 6.3 we present strong empirical evidence supporting our conjecture, motivating furthur research into this Physarum dynamic to confirm its soundness and convergence to optimality. In section 5 we show that choosing a different conductance for our general recipe of SDP Physarum dynamic suffices to guarantee both soundness and convergence to optimality, showcasing the power of our general framework.

4.3 Extension to Infeasible Starting Point

In the previous sections we discussed the soundness and convergence of our first ansatz with a feasible starting point. In this section, we show that it is possible to extend the dynamic to cope with infeasible starting points. We also introduce a different perspective on the definition of $\dot{X}(t)$ which will be the starting point of the extension of the dynamic.

Update problem Geometrically speaking, with positive definite G, $-(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A})$ is a projection into the kernel of $\mathcal{A}$, under the inner product induced by G^{-1}. In particular, given any w, we have the following:

$$-(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A})w = \text{argmin}_y \left\{ \|w - y\|^2_{G^{-1}} : \mathcal{A}y = 0 \right\}$$

We will not prove this claim, but instead, we prove the following charactarization of $\dot{X}(t)$ as an update problem:

Theorem 4.7. *Consider the SDP Physarum dynamic* (5)*, and positive semi-definite conductance* G*, then*

$$\dot{X} = \textit{argmin}_F \left\{ \underset{\rightarrow}{C}^T \underset{\rightarrow}{F} + \frac{1}{2}\underset{\rightarrow}{F}^T G^\dagger \underset{\rightarrow}{F} : \mathcal{A}\underset{\rightarrow}{F} = 0 \right\}, \tag{8}$$

Proof. We directly solve the minimization problem using the method of Lagrange multipliers. Using the Lagrange multiplier p, $\dot{X}$ is the solution to the following system:

$$\begin{aligned} \underset{\rightarrow}{C} + G^\dagger \underset{\rightarrow}{\dot{X}} &= \mathcal{A}^T p \\ \mathcal{A}\underset{\rightarrow}{\dot{X}} &= 0 \end{aligned} \tag{9}$$

Multiplying both sides of the first equation in (9) with $\mathcal{A}G$, we have that:

$$\mathcal{A}G\underset{\rightarrow}{C} = \mathcal{A}G\mathcal{A}^T p$$

Solving for p, we get

$$p = (\mathcal{A}G\mathcal{A}^T)^\dagger \mathcal{A}G\underset{\rightarrow}{C}$$

Plug it back into the first equation, we see that

$$\underset{\rightarrow}{\dot{X}} = -(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger\mathcal{A})G\underset{\rightarrow}{C},$$

which coincide with our original definition of the SDP Physarum dynamic. ☐

Introducing the update problem allows us to extend the SDP Physarum dynamic naturally to accommodate the infeasible starting point. In general, our first ansatz dynamic can be defined as the following:

$$\dot{X} = \text{argmin}_F \left\{ \underset{\rightarrow}{C}^T \underset{\rightarrow}{F} + \frac{1}{2}\underset{\rightarrow}{F}^T G^\dagger \underset{\rightarrow}{F} : \mathcal{A}\underset{\rightarrow}{F} = b - \mathcal{A}\underset{\rightarrow}{X}(t) \right\}, \tag{10}$$

where $G = \frac{1}{2}(C^{-1} \otimes X + X \otimes C^{-1})$ is still the conductance matrix. Similar to the proof of Theorem 4.7, we give without proof the solution to the general update problem, and write out the SDP Physarum dynamic explicitly:

$$\dot{X}(t) = G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^\dagger b - \underset{\rightarrow}{X}(t) \tag{11}$$

It should be easy to see that we can equivalently define $\dot{X}$ with an auxilary variable Q such that $\dot{X}(t) := Q(t) - X(t)$ where

$$Q = \text{argmin}_F \left\{ \underset{\rightarrow}{F}^T G^\dagger \underset{\rightarrow}{F} : \mathcal{A}\underset{\rightarrow}{F} = b \right\} \tag{12}$$

Approaching linear feasibility When the Physarum starts with a feasible point, Theorem 3.2 shows that the dynamic stays linearly feasible. We now show that with a linearly infeasible starting point, the dynamic (11) approaches linear feasibility exponentially.

Theorem 4.8. *Given any $X(0) \succ 0$, for any linear constraints ℓ, we have:*

$$tr(A_\ell X(t)) - b_\ell = e^{-t}(tr(A_\ell X(0)) - b_\ell)$$

In case $X(0)$ is linearly feasible, $X(t)$ remains linearly feasible and the dynamic (10) *reduces to the dynamic* (8)*.*

Proof. By definiton of the update problem (10), $\mathcal{A}\underset{\longrightarrow}{\dot{X}}(t) = b - \mathcal{A}\underset{\longrightarrow}{X}(t)$. We can write the following differential equation:

$$\begin{aligned} \frac{d}{dt}(tr(A_\ell X(t)) - b_\ell) &= tr(A_\ell \dot{X}(t)) \\ &= b_\ell - tr(A_\ell X(t)) \end{aligned}$$

Therefore, the function $\Delta_\ell(t) \coloneqq tr(A_\ell X(t)) - b_\ell$ can be characterized by

$$\dot{\Delta_\ell}(t) = -\Delta_\ell(t)$$

The solution to this differential equation is $\Delta_\ell(t) = e^{-t}\Delta_\ell(0)$, that is, $X(t)$ converges to linear feasibility exponentially fast.

In the case $X(0)$ is linearly feasible, $\Delta_\ell(0) = 0$ and $\Delta_\ell(t) = 0$, in other words, $X(t)$ stays linearly feasible. □

Positive definiteness Given Lemma 4.1, to show that the general first ansatz stays in the positive definite cone, we only need to show that the derivative of the work function $W(t) = \ln\det X(t)$ is lower bounded by some time-independent constant.

Theorem 4.9. *If C^{-1} is linearly feasible and $X(0) \succ 0$ (not necessarily linearly feasible), then for the general first ansatz, when $X(t) \succ 0$, the derivative of the work function $W(t) = \ln\det X(t)$ is lower bounded by a time-independent constant. Furthermore, by Lemma 4.1, the dynamic stays in the positive definite cone at any finite time.*

Proof. Following the proof of Theorem 4.2, we will lower bound $tr(X^{-1}\dot{X})$. By definition (11),

$$\begin{aligned} \underset{\longrightarrow}{X^{-1}}^T \underset{\longrightarrow}{\dot{X}} &= \underset{\longrightarrow}{X^{-1}}^T G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}b - n \\ &= \underset{\longrightarrow}{C^{-1}}^T \mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}b - n \end{aligned}$$

By our assumption that C^{-1} is linearly feasible, we have:

$$\begin{aligned} \underset{\longrightarrow}{X^{-1}}^T \underset{\longrightarrow}{\dot{X}} &= b^T(\mathcal{A}G\mathcal{A}^T)^{-1}b - n \\ &\geq -n, \end{aligned}$$

where the inequality follows from the positive definiteness of $(\mathcal{A}G\mathcal{A}^T)^{-1}$. The rest of the proof is identical to that of Theorem 4.2. □

Convergence The convergence results in section 4.2 is not so easily generalized. The main obstacle is that even though the errors Δ_ℓ approach zero exponentially fast, they still break the proofs where we use inequalities that are tight. Nevertheless, we present the following conjecture.

Conjecture 4.10. *If $X(0) \succ F$, for some feasible solution F, then the first ansatz stays in the positive definite cone and converges to the optimum of the SDP.*

In section 6.3 we present empirical evidence that this conjecture is indeed true.

5 Second Ansatz

Section 4 gives a natural generalization of the conductance matrix in the LP Physarum dynamic (2), but due to the intrinsic difficulties in matrix computations, we are only able to prove soundness for the subclass of positive SDPs where C^{-1} or any positive multiple thereof satisfies the linear constraints. In addition, the current techniques only give conditional convergence-to-optimality for first ansatz. In this section, we show that by choosing the conductance

$$G \coloneqq X \otimes X,$$

soundness is guaranteed for our general recipe of SDP Physarum dynamic (5). Moreover, we prove that, unconditionally, this dynamic converges to **optimality**. The resulting SDP Physarum dynamic is referred to as *second ansatz*. In this section we assume that the Physarum dynamic has a feasible starting point.

5.1 Soundness of the Second Ansatz

Recall that we state and prove Lemma 3.1 and Theorem 3.2 general enough to include different choices of the conductance G, in particular, our new choice $G = X \otimes X$. Therefore, to establish soundness of our second ansatz, we only need to show that $X(t)$ stays in the positive definite cone. In light of Lemma 4.1, we would like to give a lower bound on the derivative of the work function. Unfortunately, the lower bound that we can get in the following might not be time-independent, at least not at first glance.

Lemma 5.1. *Given $X(t) \succ 0$ at some point t, the derivative of the work function $W(t) = \ln\det(X(t))$ is lower bounded:*

$$\frac{d}{dt} \ln\det X(t) \geq -\sqrt{n} tr(CX(t)) \tag{13}$$

Proof. By definiton of $\dot{X}(t)$, we have:

$$\begin{aligned} \frac{d}{dt} \ln\det X(t) &= tr(X^{-1}(t)\dot{X}(t)) \\ &= -\underset{\longrightarrow}{X^{-1}}(I - G\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A})G\underset{\rightarrow}{C} \end{aligned}$$

Write $P_G \coloneqq G^{\frac{1}{2}}\mathcal{A}^T(\mathcal{A}G\mathcal{A}^T)^{-1}\mathcal{A}G^{\frac{1}{2}}$ and notice that P_G is an orthogonal projection. Plug it into the equation above:

$$\begin{aligned}\frac{d}{dt}\ln\det X(t) &= -\underrightarrow{X^{-1}}G^{\frac{1}{2}}(I-P_G)G^{-\frac{1}{2}}\underrightarrow{C} \\ &\geq -\sqrt{\underrightarrow{X^{-1}}^T G\underrightarrow{X^{-1}}}\left\|(I-P_G)G^{\frac{1}{2}}\underrightarrow{C}\right\|\end{aligned}$$

where the inequality follows from the Cauchy-Schwartz inequality. Now by the definition that $G = X \otimes X$, we get that $\underrightarrow{X^{-1}}^T G\underrightarrow{X^{-1}} = tr(I) = n$. On the other hand, since P_G is an orthogonal projection, it's easy to verify that $I - P_G$ is an orthogonal projection as well, therefore,

$$\left\|(I-P_G)G^{\frac{1}{2}}\underrightarrow{C}\right\| \leq \left\|G^{\frac{1}{2}}\underrightarrow{C}\right\|$$

Moreover, $\left\|G^{\frac{1}{2}}\underrightarrow{C}\right\|^2 = tr(XCXC)$. Now by the cyclic property of trace, we have:

$$\begin{aligned}tr(XCXC) &= tr\left(\left(X^{\frac{1}{2}}CX^{\frac{1}{2}}\right)\left(X^{\frac{1}{2}}CX^{\frac{1}{2}}\right)\right) \\ &= \sum_{i=1}^n \sigma_i^2, \text{ where } \sigma_i \text{ are the eigenvalues of } \left(X^{\frac{1}{2}}CX^{\frac{1}{2}}\right) \\ &\leq \left(\sum_{i=1}^n \sigma_i\right)^2 \\ &= tr^2\left(X^{\frac{1}{2}}CX^{\frac{1}{2}}\right)\end{aligned}$$

where in the first inequality we need the assumption that $C \succ 0$ and $X \succ 0$ which implies that all σ_i's are nonnegative.

Putting all these together,

$$\begin{aligned}\frac{d}{dt}\ln\det X &= tr(X^{-1}\dot{X}) \\ &\geq -\sqrt{n}\cdot tr\left(X^{\frac{1}{2}}CX^{\frac{1}{2}}\right) \\ &= -\sqrt{n}\cdot tr(CX)\end{aligned}$$

□

At first glance, the lower bound that we have in Lemma 5.1 is time dependent, but as it turns out, it is already enough for us to prove that $X(t)$ stays within the positive definite cone.

Theorem 5.2. *If $X(0) \succ 0$ is linearly feasible, then for the second ansatz where the conductance is $G = X \otimes X$, $X(t)$ stays in the positive definite cone at any finite time t.*

Proof. If at any time $X(t) \not\succ 0$, then by the assumption that $X(0) \succ 0$ and the continuity of the eigenvalues of $X(t)$, there must exists a finite time when $X(t)$ obtains a zero eigenvalue. Consider the first time $T > 0$ where $X(T)$ obtains a zero eigenvalue. In other words, for any $t < T, X(t) \succ 0$. Then by Theorem 2.1, for any $t \in [0,T)$, $\frac{d}{dt}\mathcal{L}(t) \leq 0$, that is,

$$\forall t \in [0,T), tr(CX(t)) \leq tr(CX(0))$$

Therefore, similar to the proof of Lemma 4.1, and by Lemma 5.1, we have

$$\begin{aligned}\lim_{t\to T^-} W(t) &= W(0) + \lim_{t\to T^-} \int_{s=0}^{t} \frac{d}{ds} W(s) ds \\ &\geq W(0) - \sqrt{n} \cdot tr(CX(0))T,\end{aligned}$$

which is a time-independet constant, given a finite starting point $X(0)$. However, by our assumption that $X(T)$ obtains a zero eigenvalue, $\lim_{t\to T^-} W(t)$ must be divergent and not lower bounded by any finite value. This yields a contradiction and we get that for any finite time $T \geq 0, X(T) \succ 0$. □

5.2 Convergence of the Second Ansatz

Having established the soundness of the second ansatz, we can directly conclude the following as a corollary of Theorem 2.1

Corollary 5.3. *Given a feasible starting point $X(0) \succ 0$, then the second ansatz satisfies the following:*

1. $\frac{d}{dt}\mathcal{L}(t) \leq 0$
2. $\frac{d}{dt}\mathcal{L}(t)$ *becomes zero if and only if* $\dot{X}(t) = 0$
3. *The dynamic converges to equilibrium* $\lim_{t\to\infty} \left\|\dot{X}(t)\right\| = 0$

Similar to before, this corollary does not immediately imply the dynamic's convergence to optimality. However, we show that our second ansatz indeed converge to the optimum of SDP (3) using an approach very different from the above. In fact, we prove that given any feasible starting point $X(0)$, the second ansatz coincides with a central path of the SDP. This argument is akin to the central path argument in [4], where the barrier in their work is an entropy type barrier. However, we use a more canonical logarithm type barrier. Before we proceed to the statement and proofs, we point out that the main reason that such arguments cannot be applied for the first ansatz again boils down to the noncommutative nature of matrix multiplication. Differentiating entropy-type barriers on matrices typically results in infinite series, and analyzing them in closed-form would require communtativity assumption. However, with a clever choice of conductance matrix, we manage to bypass this issue while still following our general recipe of SDP Physarum dynamics (5). We believe that it showcases the power and potential of our general framework and motivates research into designing other conductance matrices that leads to Physarum dynamics with better properties.

Central path and SDP Physarum dynamic Now we prove that the second ansatz starting from a feasible point $F \succ 0$ conincides with a central path of the SDP. Our barrier function is

$$f(X) := -\ln\det X + tr(F^{-1}X),$$

and with the parameter $t \geq 0$, points on our central path are

$$X(t) := \begin{array}{ll} \text{argmin} & t \cdot tr(CX) - \ln\det X + tr(F^{-1}X) \\ \text{s.t.} & \mathcal{A}\underset{\rightarrow}{X} = b \\ & X \succ 0 \end{array} \tag{14}$$

As $t \to \infty$, the central path (14) approaches to the optimal solution of SDP (3). Now we can state the main result of this section:

Theorem 5.4. *Fix any feasible point $F \succ 0$, the solution $X(t)$ to the central path problem* (14) *is the solution to the second ansatz with the starting point $X(0) = F$. Consequently, the Physarum dynamic converges to the optimum of SDP* (3) *when starting from any feasible point $X(0) \succ 0$.*

Proof. We first show that $X(0) = F$. With $t = 0$, we write down the Lagrangian equations of (14) (with the constraint $X \succ 0$ omitted):

$$-X^{-1} + F^{-1} = \mathcal{A}^T p$$
$$\mathcal{A}\underset{\rightarrow}{X} = b$$

Clearly $(F, 0)$ is a solution to the above system. This implies that F is the optimal solution to $\min \left\{ -\ln\det X + tr(F^{-1}X) : \mathcal{A}\underset{\rightarrow}{X} = b \right\}$. However, by our assumption that $F \succ 0$, we know that F is feasible for (14) when $t = 0$. Therefore, $X(0) = F$.

Now we prove that

$$\underset{\rightarrow}{\dot{X}}(t) = -(I - (X \otimes X)\mathcal{A}^T(\mathcal{A}(X \otimes X)\mathcal{A}^T)^{-1}\mathcal{A})(X \otimes X)\underset{\rightarrow}{C},$$

which, by the definition of our Physarum dynamic, implies that the central path is indeed the Physarum dynamic. Now we introduce the dual variable y for the linear constraints and consider the Lagrangian:

$$L(X, y) = t \cdot tr(CX) - \ln\det X + tr(F^{-1}X) + y^T(\mathcal{A}\underset{\rightarrow}{X} - b)$$

It is apriori not so clear why it is possible to drop the positive definiteness constraint from the central path problem. However, once we show that the "relaxed" central path problem coincide with the second ansatz, then due to our soundness statements in section 5.1 we see that we indeed did not lose anything by dropping the positive definiteness constraint. Now, we take the partial derivative of L with respect to X:

$$\frac{\partial}{\partial X} L(X, y) = tC - X^{-1} + F^{-1} + \sum_\ell y_\ell A_\ell$$

Setting it to zero, we get the relation between the optimal X and y:

$$X^{-1} = tC + F^{-1} + \sum_\ell y_\ell A_\ell$$

Now we differentiate with respect to t on both side and get

$$-X^{-1}\dot{X}X^{-1} = C + \sum_\ell \dot{y}_\ell A_\ell$$

In the vectorized notation, this is

$$\underset{\rightarrow}{\dot{X}} = (X \otimes X)\underset{\rightarrow}{C} + (X \otimes X)\mathcal{A}^T\dot{y}$$

However, since $\mathcal{A}\underset{\rightarrow}{X} = b$, we get $\mathcal{A}\underset{\rightarrow}{\dot{X}} = 0$. In terms of $\dot{y}$, this becomes

$$\mathcal{A}(X \otimes X)\underset{\rightarrow}{C} + \mathcal{A}(X \otimes X)\mathcal{A}^T\dot{y} = 0$$

Solving for $\dot{y}$ and plugging it back, we finally obtain

$$\underset{\rightarrow}{\dot{X}}(t) = -(I - (X \otimes X)\mathcal{A}^T(\mathcal{A}(X \otimes X)\mathcal{A}^T)^{-1}\mathcal{A})(X \otimes X)\underset{\rightarrow}{C},$$

in other words, the central path (14) coincides with the second ansatz. □

6 Algorithms and Experimental Results

In this section we introduce the augmented SDP and show that C^{-1} is feasible therein, explain our discretizations of the first ansatz and hence obtain numerical algorithms for solving SDP, and we present empirical evaluations of the numerical algorithms to complement where our current theoretical techniques are lacking. Beyond that, our empirical evaluations show the potential of nature-inspired algorithms in solving positive SDP. Our implementations and experiments are available at:

`https://github.com/HamidrezaKmK/PhysarumSDPSolver`

To accurately implement an algorithm for our Physarum solver framework, we define a set of update steps whereby we calculate $\dot{X}(t)$ according to the $X(t)$ value. Note that according to (5), $\dot{X}$ can be written as,

$$\underset{\rightarrow}{\dot{X}} = G\mathcal{A}^T(\underbrace{\mathcal{A}G\mathcal{A}^T}_{L})^{\dagger}\underbrace{\mathcal{A}G\underset{\rightarrow}{C}}_{b} - GC$$

Recall that according to Theorem 4.7's proof p is defined as the Lagrangian multipliers and the solution to the equation $Lp = b$. In both first ansatz and second ansatz, p can be considered as a dual candidate solution, and when the dynamic converges, it can produce a dual solution; therefore, we keep track of it to obtain a primal-dual solver.

That said, we split up the computation of the update step into three phases which repeat in all extensions of our SDP solvers:

1. We calculate matrix L element-by-element. $L_{i,j}$ is equal to $\underset{\rightarrow}{A_i}^T G \underset{\rightarrow}{A_j}$ in the general framework. For the first ansatz, it is equal to $tr(C^{-1}A_iXA_j)$ and for the second ansatz it is equal to $tr(XA_iXA_j)$.

2. We then calculate the solution to $Lp = b$ and present p as a dual candidate solution. In general, $p = L^{\dagger}b$.

3. Finally, we put it all together by calculating $\dot{X}$ using the following formula:

   $$\underset{\rightarrow}{\dot{X}} = G\mathcal{A}^T p - GC.$$

   With the first ansatz, the formula can be re-written in the following matrix notation without the need for vectorization:

   $$\dot{X} = \sum_{\ell=1}^{m} p_\ell \frac{C^{-1}A_\ell X + XA_\ell C^{-1}}{2} - X \tag{15}$$

   By analogy, one can also write down the following simplified formulation for the second ansatz in a similar fashion:

   $$\dot{X} = \sum_{\ell=1}^{m} p_\ell XA_\ell X - XCX \tag{16}$$

   We omit the steps of computations here.

6.1 Augmentation

Recall that in section 4, our theoretical guarantees relies on the assumption that C^{-1} is linearly feasible. While we conjecture that such assumption is unnecessary, in this section we show that SDP (3) can be augmented so that the inverse of the objective matrix is linearly feasible. To this end, we augment all matrices in (3) by one row and column. We define $\bar{C}$ and $\bar{A}_\ell$ according to the following scheme using a certain γ:

$$\bar{C} = \begin{pmatrix} \gamma C & 0 \\ 0 & 1 \end{pmatrix}, \bar{A}_i = \begin{pmatrix} A_i & 0 \\ 0 & \alpha_i \end{pmatrix} \quad , \quad \alpha_i = b_i - \frac{tr(A_i C^{-1})}{\gamma}.$$

We will also refer to an arbitrary matrix T as *augmented* if all entries of the right column and bottom row except bottom right entry are zero. With these, define the following SDP the **Augmented SDP**:

$$\min\{tr(\bar{C}\bar{X}) : tr(\bar{A}_\ell \bar{X}) = b_\ell \ \forall \ell \in [m],\ \bar{X} \succeq 0\}. \tag{17}$$

One can easily show $\bar{C}^{-1}$ is feasible in this case:

$$tr(\bar{A}_\ell \bar{C}^{-1}) = \frac{A_\ell C^{-1}}{\gamma} + \alpha_\ell = b_\ell$$

Therefore, by choosing $\bar{X}(0) = \bar{C}^{-1}$ we start with a feasible solution. By Theorem 3.2, $X(t)$ remains linearly feasible. Next, we present a lemma that states if $X(0)$ is augmented, then $X(t)$ stays augmented under certain conditions. We omit the proof here because it boils down to tedious verification and offers no insights. This lemma establishes a mapping between the dynamic in the augmented problem and the original problem.

Lemma 6.1. *For any given autonomous dynamic on $\bar{X}(t)$ with a symmetric and augmented starting point $\bar{X}(0)$ where*

$$\dot{\bar{X}}(t) = v(\bar{X}(t)),$$

$X(t)$ remains augmented throughout the dynamic if $v(X)$ is obtained by addition, multiplication, and inversion over symmetric augmented matrices including $\bar{X}$.

By taking into account the simplified formulas (15) and (16), both of these dynamic comply with the dynamics described in Lemma 6.1 in the augmented case. Therefore, we can re-write the dynamic in the augmented form as below:

$$\bar{X}(t) = \begin{pmatrix} \tilde{X}(t) & 0 \\ 0 & \beta(t) \end{pmatrix}, \dot{\bar{X}}(t) = \begin{pmatrix} \dot{\tilde{X}}(t) & 0 \\ 0 & b^T p - \beta(t) \end{pmatrix}$$

Each feasible solution of the original problem maps to a feasible solution $\bar{X}$ of the augmented problem by adding a row and a column with all zeros. The objective value of such an $\bar{X}$ is then scaled by γ, i.e., $tr(\bar{C}\bar{X}) = \gamma tr(CX)$. By an appropriate choice of γ, these objective values can be made smaller than the objective value of $\bar{C}^{-1}$, which is a constant. Hence we have the following:

$$tr(\bar{C}\bar{X}_{opt}) \leq \gamma tr(C X_{opt}) \tag{18}$$

If the dynamic converges to the optimum of the augmented SDP, and $\beta(t)$ converges to zero, then at sufficiently large time $T > 0$,

$$\begin{aligned}\gamma tr(CX_{opt}) &\geq tr(\bar{C}\bar{X}_{opt}) \\ &= tr(\bar{C}\bar{X}) - \epsilon\gamma \\ &= \gamma tr(\tilde{X}) + \beta(T) - \epsilon\gamma \\ &\geq \gamma tr(\tilde{X}) - \epsilon\gamma\end{aligned}$$

Therefore, the approximation $\tilde{X}$ has objective at most ϵ larger than the optimum objective value. Note that $\tilde{X}$ is almost feasible where the error is proportional to $\beta(t)$. So with the augmentation, we are approximating the original problem, both in terms of optimality and feasibility. We will also provide experimental evidence in Section 6.3 that shows $\beta(t)$ approaches zero for first ansatz where we need augmentation. Note that $\bar{X}(0) = \bar{C}^{-1}$ complies with the condition in Conjecture 4.10 as it is equal (therefore dominates) a feasible solution: $\bar{C}^{-1}$. Furthermore, this also provides evidence to back up the optimality convergence statement in Conjecture 4.10.

6.2 Discretization and Numerical Algorithms

In this section, we investigate the discretization of the SDP Physarum dynamic and introduce a framework to obtain a primal dual solver for SDP problems. The algorithms and experimental evaluations in this section complement our conjectures about the soundness and convergence to optimality of the first ansatz. Moreover, we believe that these nature-inspired numerical algorithms have strong potentials in its practicality.

To simulate the continuous dynamic in an algorithm, we will discretize it using the following [22]:

$$X(t+1) \leftarrow X(t) + h\dot{X}(t) \tag{19}$$

Intuitively, this equation simulates the SDP Physarum dynamic when $h \to 0$. As a consequence, for practical purpose, h should be chosen small enough.

Vanilla algorithm The vanilla algorithm is a straightforwad implementation of the discretization scheme (19). The algorithm consists of a sequence of iterations where in each iteration we solve for $\dot{X}(t)$ and update $X(t)$ with the discretized dynamic and step h to obtain $X(t+1)$. The process continues until $\dot{X}(t)$ becomes too small which indicates reaching an equilibrium point. We stop the iteration whenever $\left\|\dot{X}\right\|$ is smaller than some small constant ϵ. Thus we obtain the following algorithm:

Algorithm 1: Vanilla Physarum SDP solver

Input : $C, A_1, \ldots, A_m \in S_n$, $b \in \mathbb{R}^m$, $X(0)$.
Output: $(X^{eq} \succeq 0, p^{eq})$.

Let $t = 0$;
repeat
  $\dot{X}(t), p(t) \leftarrow$ SolveUpdateProblem$(C, A_1, ..., A_m, b, X(t))$ using Algorithm 2;
  Calculate small enough h;
  Update $X(t+1) \leftarrow X(t) + h\dot{X}(t)$;
  Increment t;
until $\left\|\dot{X}(t)\right\| \leq \epsilon$;
return $(X(t), p(t))$

Each iteration solves an update problem (or equivalently, a projection under weighted inner product). We give below the update problem solver in terms of $\dot{X}$ and p; moreover, we use the simplified formulas in (15) to implement the update problem.

Algorithm 2: Solve Update Problem

Input : $C, A_1, \ldots, A_m \in S_n$, $b \in \mathbb{R}^m$, X.
Output: $\dot{X}, p$.

Calculate the $m \times m$ matrix L and let $L_{i,j} \leftarrow \underrightarrow{A_i}^T G \underrightarrow{A_j}$;
Calculate $p \leftarrow L^{\dagger} b$;
Calculate $\dot{X}$ s.t. $\underrightarrow{\dot{X}} \leftarrow G\mathcal{A}^T p - GC$;
return $(\dot{X}, p)$

As mentioned before, h should be chosen small enough. One obvious upper-bound for h is that the update should preserve positive semi-definiteness of $X(t)$. There is a close form bound that guarantees positive semi-definiteness, or one can use exponential/binary search. We defer interested reader to our implementation.

In our experiments, we run Algorithm 1 on the augmented problems and map the augmented solution back to the original problem (recall that by Lemma 6.1 the iterates stays augmented). We provide empirical evidence that $\beta(t)$ converges to zero and that $\tilde{X}$ which was obtained from taking the upper-left $n \times n$ matrix of X yields a sound and optimal solution to our SDP.

Modified algorithm One of the problems encountered by Algorithm 1 is that $L = \mathcal{A}G\mathcal{A}^T$ becomes ill-conditioned as some eigenvalues of X, and in turn, G converge to zero while reaching an equilibrium point which is not full-rank. To circumvent this problem, we introduce the modified algorithm that facilitate numerical hacks to work with near-zero eigenvalues.

For the modified algorithm, we drop the condition on $X(0)$ being linearly feasible and we start with a large enough matrix $X(0) = \eta \times I$ where η is set to a large value such that $X(0)$ dominates an arbitrary feasible solution; this complies with the statement in Conjecture 4.10 where we consider non-feasible starting points. As a result, we do not augment the problem, and interestingly enough, the dynamic still converges to the optimum in our experiments. We introduce a set of "epochs" in our algorithm. Each epoch ends whenever either $\left\|\dot{X}(t)\right\|$ becomes less than a certain ϵ (i.e, the dynamic converges) or the minimum eigenvalue of $X(t)$ becomes smaller than ϵ. In the former case, we have reached equilibrium, but in the latter, we need to restart the dynamic and start another epoch.

We keep track of a set of linearly independent vectors, which we call a "diagonalization basis", that can simultaneously diagonalize $X(t)$ and C at any time t. The diagonalization basis will contain n vectors in the beginning since the initialization sees $X(0)$ as being full-rank. Assume that at a certain arbitrary epoch, $X(0)$ has rank k. Let us identify the diagonalization basis with $\{\tilde{u}_1, \tilde{u}_2, ..., \tilde{u}_k\}$ and obtain the $n \times k$ matrix $\tilde{U}$ by stacking these vectors in columns next to each other. $\tilde{U}$ is a time-invariant matrix that diagonalizes C and $X(t)$ at any time t; in other words, both $\tilde{U}^T C \tilde{U}$ and $\tilde{U} X(t) \tilde{U}$ are diagonal. The detail on why this matrix $\tilde{U}$ exists and how to obtain it is spared and we defer readers to take a look into our implementation which calculates these matrices using a simultaneous diagonalization technique on C and $X(0)$ based on generalized eigenvalues [23] at the beginning of each epoch.

With that in mind, when $X(t)$ obtains a near zero eigenvalue, the diagonal matrix $\tilde{U}^T X(t) \tilde{U}$ will obtain a near zero element λ_i on the diagonal. In that case, we pop $\tilde{u}_i$ from the diagonalization basis and start the new epoch. We define a projection of an arbitrary matrix M on the set of diagonalization basis vectors $\{\tilde{u}_1, ..., \tilde{u}_k\}$ as the solution to the following equation:

$$projection_{\tilde{U}}(M) = \sum_{i=1}^{k} c_i \tilde{u}_i \tilde{u}_i^T \text{ s.t } \{c_1, ..., c_m\} = \underset{\{z_1,...,z_m\}}{\arg\min} \left\| M - \sum_{i=1}^{k} z_i \tilde{u}_i \tilde{u}_i^T \right\|_2^2$$

Intuitively speaking, the *projection* function considers the best approximation of a matrix given a certain set of vectors. When $\tilde{u}_i^T X(t) \tilde{u}_i$ converges to zero, we may assume that $\tilde{u}_i$ is in the kernel of any optimum X; therefore, projecting all of the dynamic into a new basis by excluding $\tilde{u}_i$ will help avoid dealing with near zero eigenvalues in $X(t)$. As a result, we can work around numerical instability by calculating a projected G matrix in the update problem phase which will not be ill-conditioned.

The modified algorithm contains a lot of technical details embedded in our implementation; however, for simplicity, we do not expand on the theoretical foundations of these technical details. A simplified version of the algorithm using the projection intuition can be summarized in Algorithm 3. Additionally, we use the modified algorithm as strong evidence to back up Conjecture 4.10. Note that even though we do not start from a feasible solution, we converge to an optimum solution when we set $X(0)$ to be a large enough matrix.

6.3 Experimental Evaluation

In this section we discuss the dataset used to evaluate our algorithms and their implementations. These experimental results supports our Conjecture 4.10. For validation, our algorithms are compared with a standard SDPASolver[2] as ground truth.

We evaluate the objective value obtained from our dynamic using a metric identified as "gap" which calculates the absolute difference between the ground truth primal solution $tr(CX^*)$ and the objective value obtained from our dynamic: $tr(CX(\infty))$. We also evaluate the soundness of

[2] `http://sdpa.sourceforge.net/`

Algorithm 3: Modified Physarum SDP solver

Input : $C, A_1, \ldots, A_m \in S_n$, $b \in \mathbb{R}^m$, η.
Output: $(X^{eq} \succeq 0, p^{eq})$.

Set $X(0) = \eta \times I$;
Initialize the diagonalization basis in the $n \times n$ matrix $\tilde{U}$ s.t. $\tilde{U}^T X(t) \tilde{U}$ and $\tilde{U}^T C \tilde{U}$ remain diagonal throughout the dynamic;
repeat
  Pop any column in $\tilde{U}$ which corresponds to diagonal elements less than ϵ in $X(0)$;
  Set $X(0) \leftarrow \eta \times projection_{\tilde{U}}(X(0))$;
  Set $C \leftarrow projection_{\tilde{U}}(C)$;
  Let $t \leftarrow 0$;
  repeat
    Let $\dot{X}(t), p(t) \leftarrow$ SolveUpdateProblem$(C, A_1, ..., A_m, b, X(t))$ using Algorithm 2;
    Calculate small enough h;
    Update $X(t+1) \leftarrow X(t) + h\dot{X}(t)$;
    Increment t;
  until $\left\|\dot{X}(t)\right\| \leq \epsilon$ *or* $\lambda_{\min}(X(t)) \leq \epsilon$;
until $\left\|\dot{X}(t)\right\| \leq \epsilon$ *and* $\lambda_{\min}(projection(X(t))) \geq \epsilon$;
return $(X(t), p(t))$

our algorithm according to a real number identified as "infeasibility". The infeasibility metric is calculated as below:

$$infeasibility(X) = \max\left(\max_{\ell \in [m]} |b_\ell - tr(A_\ell X)|, \max(0, -\lambda_{\min}(X))\right)$$

For a sound dynamic, infeasibility should converge to zero and any deviation from zero impedes the soundness of the dynamic.

The dataset Table 1 shows a set of 180 positive SDP samples. The tests are generated using three different schemes on different matrix size initialization. Details as follows:

- **Random-Tests:** These tests are generated randomly. Given a certain n and m - the size of matrices and the number of conditions - a random positive definite matrix C and random symmetric matrices A_ℓ are generated. The datasets which are created using this scheme consist of "testset1" to "testset3", "large1", and "large2".

- **Vertex-Cover:** To generate these data we create a random graph with n vertices numbered from 1 to n and m edges. We will then create an SDP which is able to approximate the minimum vertex cover problem. For each random graph with n vertices, C is the identity matrix of size $(n+1) \times (n+1)$. Then the following linear condition matrices are generated:

  – For each vertex v in the graph an A_ℓ is created according to the following:

$$(A_\ell)_{i,j} = \begin{cases} -1 & \text{if } \min(i,j) = 1 \text{ and } \max(i,j) = v+1 \\ 2 & \text{if } i = j = v+1 \\ 0 & \text{otherwise} \end{cases}$$

and the respective b_ℓ is set to 0.

  - For each edge (v, w) in the graph an additional A_ℓ is created according to the following:

$$(A_\ell)_{i,j} = \begin{cases} -1 & \text{if } \min(i,j) = \min(v+1, w+1) \text{ and } \max(i,j) = \max(v+1, w+1) \\ 1 & \text{if } \min(i,j) = 1 \text{ and } \max(i,j) \in \{v+1, w+1\} \\ 0 & \text{otherwise} \end{cases}$$

    and the respective b_ℓ is set to 2.

  - An additional A_ℓ is created which is set to zero except one element on the upper-left corner and the respective b_ℓ is 1.

  These tests have $n + m + 1$ conditions and matrices of size $n + 1$.

- **Max-Cut:** We have created two sets of Max-Cut problems referred to as 'maxcut1' and 'maxcut2'. To generate them, we create a random graph with n vertices. An approximation of the maximum cut of the created graph can be computed using the following SDP problem

$$\arg\max\{tr(WX) \;:\; X \succeq 0, \text{Main diagonal entries of } X \text{ are all } 1\},$$

  where W is the Laplacian matrix of the graph. To satisfy the constraints of our algorithm, we transform the aforementioned to

$$\arg\min\{tr\left([\xi I - W]\, X\right) \;:\; X \succeq 0, \text{Main diagonal entries of } X \text{ are all } 1\},$$

  where ξ is big enough so that $[\xi I - W] \succ 0$. Note that as all main diagonal entries of X are required to be 1 which correspond to A_ℓ matrices that are set to be entirely zero except the entry on (i, i). That said, these tests have n linear constraints and matrices of size n.

Experimental results Now we explain our empirical results in details, thus providing evidence that first ansatz is sound and converges to optimum, in the augmented case and unconditionally (see Conjecture 4.10). We discuss the performances of the Algorithm 1 for the augmented case and Algorithm 3 for the general case. We extract the maximum *gap* and *infeasibility* for each of our test sets and will also provide empirical evidence that $\beta(t)$ converges to zero in the augmented case for Algorithm 1.

Performance of the vanilla algorithm Table 2 contains the results of Algorithm 1 on each of the datasets. Each row contains one of the datasets in Table 1. The solver provides accurate solutions on most of the datasets. In the "vertexcover1" and "vertexcover2" dataset however we have an increase in the number of linear conditions which makes the matrix L larger; therefore, making the update step more difficult. This can in turn make the algorithm both slower and more numerically unstable. In addition to that, the solver runs into numerical difficulties in these tests. We believe that there are two main causes:

- h is not chosen sufficiently small. We have only used a naive approach to choose h and by tuning it we might get more accurate results.
- L becomes ill-conditioned due to some of $\lambda_i(t)$ values vanishing.

TestSet Description	TestSet Name	No. of tests
Randomly generated definite SDP samples with $n = 5, m \in [3]$	testset1	20
Randomly generated definite SDP samples with $n = 10, m \in [5]$	testset2-1	14
	testset2-2	16
Randomly generated definite SDP samples with $n = 25, m \in [10]$	testset3-1	16
	testset3-2	4
Randomly generated vertex cover problems with $\|V(G)\| = 5, \|E(G)\| \in [10]$	vertexcover1-1	13
	vertexcover1-2	7
Randomly generated vertex cover problems with $\|V(G)\| = 20, \|E(G)\| \in [130]$	vertexcover2-1	4
	vertexcover2-2	16
Randomly generated definite SDP samples with $n = 50, m \in [5, 10]$	large1	10
Randomly generated definite SDP samples with $n = 100, m \in [5, 20]$	large2	20
Randomly generated vertex cover problems with $\|V(G)\| = 50, \|E(G)\| \in [10, 20]$	vertexcover3	10
Randomly generated max cut problems with $\|V(G)\| \in [20, 50]$	maxcut1	20
Randomly generated vertex cover problems with $\|V(G)\| = 100$	maxcut2	10

Table 1: Standard SDPLib dataset that is generated to validate the Physarum dynamics.

The latter is addressed by the modified algorithm 3.

The vanilla algorithm starts with a linearly feasible $X(0)$ by augmenting the problem. We have presented the maximum β values in the last iteration in the table. It is worthwhile mentioning that we did not augment the problem in "maxcut1" and "maxcut2" datasets as the identity matrix is a feasible solution for these tests; therefore, by setting $X(0) = I$ we obtain the linear feasibility condition which was the main motive for our augmentation technique. On the other hand, for the rest of the dataset, β values converged to zero closely, indicating that we manage to approach linear feasibility in the original problem.

Performance of the modified algorithm We test the performance of Algorithm 3 without augmentation. Table 3 shows the results for the modified dynamics on the datasets. This solver is able to solve all of the tests in the dataset with descent accuracy, thus providing empirical support for our conjecture on the unconditional convergence of first ansatz. However, it takes a longer time for the dynamic to converge accurately. We believe by a more appropriate choice of h one can also obtain a fast and more accurate Physarum solver.

Note that although we have claimed that the accuracy of this algorithm is better than Vanilla, Vanilla works better on the maxcut tests which make it a really good candidate to tackle these problems. In addition, "maxcut1" and "maxcut2" datasets contain relatively large matrices which is the reason why in some tests in order to get the desired accuracy the algorithm took approximately 38 minutes to become decently accurate.

TestSet	Ratio of tests with error below 10^{-2}	Average Time Spent	Maximum primal gap on tests with a lower than 10^{-2} primal gap	Maximum Infeasibility	Maximum β in last iteration for high accuracy tests
testset1	20/20	0.635 (sec)	2.7×10^{-8}	7.7×10^{-12}	7.4×10^{-54}
testset2-1	14/14	0.335 (sec)	9.5×10^{-7}	7.9×10^{-12}	9.5×10^{-7}
testset2-2	6/6	1 (sec)	1.3×10^{-6}	1.2×10^{-11}	3.8×10^{-114}
testset3-1	14/16	1.26 (sec)	7.9×10^{-7}	1.8×10^{-11}	9.3×10^{-30}
testset3-2	4/4	4.3 (sec)	4.7×10^{-7}	1.8×10^{-11}	7.3×10^{-189}
vertexcover1-1	13/13	1.4 (sec)	7×10^{-4}	0	7.4×10^{-5}
vertexcover1-2	5/7	1.8 (sec)	10^{-3}	7.2×10^{-5}	10^{-3}
large1	10/10	7.1 (sec)	1.2×10^{-8}	3.2×10^{-11}	7×10^{-54}
large2	20/20	25.8 (sec)	1.4×10^{-6}	6×10^{-11}	1.3×10^{-6}
vertexcover2-1	1/4	60 (sec)	2.6×10^{-9}	0	1.0×10^{-7}
vertexcover3	4/13	110 (sec)	4.8×10^{-5}	0	2.2×10^{-3}
vertexcover2-2	2/16	196 (sec)	3.3×10^{-9}	0	2.6×10^{-3}
maxcut1	20/20	13.4 (sec)	1.04×10^{-5}	0	–
maxcut2	10/10	155 (sec)	2.03×10^{-5}	0	–

Table 2: Results of the vanilla algorithm after augmenting the matrices with $\gamma = 0.01$. A test is considered accepted if the primal gap between the Physarum algorithm primal objective value and the ground truth is less than 10^{-2}.

Testset Name	Average Time Spent	Maximum Primal Gap	Maximum Infeasibility
testset1	7.5 (sec)	3.6×10^{-5}	5.9×10^{-12}
testset2-1	2.7 (sec)	2.9×10^{-6}	4.8×10^{-6}
testset2-2	65 (sec)	2.48×10^{-5}	1.1×10^{-13}
testset3-1	52 (sec)	1.9×10^{-7}	5.1×10^{-6}
testset3-2	115 (sec)	8×10^{-4}	5×10^{-7}
vertexcover1-1	9 (sec)	1.3×10^{-6}	2.4×10^{-7}
vertexcover1-2	41 (sec)	5.9×10^{-4}	2.6×10^{-4}
vertexcover2-1	8 (min)	4×10^{-6}	3.88×10^{-6}
vertexcover2-2	22 (min)	1.8×10^{-2}	3.3×10^{-5}
large1	32 (sec)	1.02×10^{-4}	5.9×10^{-7}
large2	170 (sec)	7.2×10^{-4}	5.5×10^{-11}
vertexcover3	7 (min)	1.7×10^{-4}	1.8×10^{-4}
maxcut1	3.5 (min)	9.4×10^{-3}	1.2×10^{-9}
maxcut2	38 (min)	9.1×10^{-3}	1.2×10^{-9}

Table 3: Results of the modified Algorithm 3 on the dataset. Maximum Primal Gap is the difference between the Physarum SDP solver and the SDPA baseline. Infeasibility measures how much X conflicts with the constraints. It is calculated by taking the maximum magnitude of the negative eigenvalue of X^{eq} and the maximum $|b_\ell - tr(A_\ell X)|$ for $1 \leq \ell \leq m$.

References

[1] T. Nakagaki, M. Iima, T. Ueda, Y. Nishiura, T. Saigusa, A. Tero, R. Kobayashi, and K. Showalter, "Minimum-risk path finding by an adaptive amoebal network," *Physical review letters*, vol. 99, no. 6, p. 068104, 2007.

[2] 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*, vol. 244, no. 4, pp. 553–564, 2007.

[3] A. Johannson and J. Zou, "A slime mold solver for linear programming problems," in *Conference on Computability in Europe*, pp. 344–354, Springer, 2012.

[4] D. Straszak and N. K. Vishnoi, "On a natural dynamics for linear programming," in *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, ITCS '16, (New York, NY, USA), p. 291, Association for Computing Machinery, 2016.

[5] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, "Physarum-inspired multi-commodity flow dynamics," *Theoretical Computer Science*, vol. 920, pp. 1–20, 2022.

[6] V. Bonifaci, "A revised model of fluid transport optimization in physarum polycephalum," *Journal of Mathematical Biology*, vol. 74, no. 3, pp. 567–581, 2017.

[7] A. Karrenbauer, P. Kolev, and K. Mehlhorn, "Convergence of the non-uniform physarum dynamics," *Theoretical Computer Science*, vol. 816, pp. 260–269, 2020.

[8] D. Straszak and N. K. Vishnoi, "Iteratively reweighted least squares and slime mold dynamics: connection and convergence," *Math. Program.*, vol. 194, no. 1, pp. 685–717, 2022.

[9] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *J. ACM*, vol. 42, p. 1115–1145, Nov. 1995.

[10] L. Lovász, "On the shannon capacity of a graph," *IEEE Transactions on Information theory*, vol. 25, no. 1, pp. 1–7, 1979.

[11] J. Kleinberg and M. X. Goemans, "The lovász theta function and a semidefinite programming relaxation of vertex cover," *SIAM Journal on Discrete Mathematics*, vol. 11, no. 2, pp. 196–204, 1998.

[12] V. Bonifaci, K. Mehlhorn, and G. Varma, "Physarum can compute shortest paths," in *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, (USA), p. 233–240, Society for Industrial and Applied Mathematics, 2012.

[13] L. Becchetti, V. Bonifaci, M. Dirnberger, A. Karrenbauer, and K. Mehlhorn, "Physarum can compute shortest paths: Convergence proofs and complexity bounds," in *Proceedings of the 40th International Conference on Automata, Languages, and Programming - Volume Part II*, ICALP'13, (Berlin, Heidelberg), p. 472–483, Springer-Verlag, 2013.

[14] T. Miyaji and I. Ohnishi, "Physarum can solve the shortest path problem on riemannian surface mathematically rigourously," *International Journal of Pure and Applied Mathematics*, vol. 47, no. 3, pp. 353–369, 2008.

[15] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, "Rules for biologically inspired adaptive network design," *Science*, vol. 327, no. 5964, pp. 439–442, 2010.

[16] S. Watanabe, A. Tero, A. Takamatsu, and T. Nakagaki, "Traffic optimization in railroad networks using an algorithm mimicking an amoeba-like organism, physarum plasmodium," *Biosystems*, vol. 105, no. 3, pp. 225–232, 2011.

[17] H. Yang, R. Mayne, and Y. Deng, "A bio-inspired network design method for intelligent transportation.," *International Journal of Unconventional Computing*, vol. 14, 2019.

[18] X. Zhang, A. Adamatzky, F. T. Chan, S. Mahadevan, and Y. Deng, "Physarum solver: a bio-inspired method for sustainable supply chain network design problem," *Annals of Operations Research*, vol. 254, no. 1, pp. 533–552, 2017.

[19] D. Straszak and N. K. Vishnoi, "Natural algorithms for flow problems," in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1868–1883, SIAM, 2016.

[20] E. Facca, F. Cardin, and M. Putti, "Physarum dynamics and optimal transport for basis pursuit," *arXiv preprint arXiv:1812.11782*, 2018.

[21] K. B. Petersen, M. S. Pedersen, *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, 2008.

[22] J. C. Butcher, *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[23] B. Ghojogh, F. Karray, and M. Crowley, "Eigenvalue and generalized eigenvalue problems: Tutorial," 2022.