Cooperativity, absolute interaction, and algebraic optimization

Nidhi Kaihnsa, Yue Ren, Mohab Safey El Din, Johannes W R Martini

COOPERATIVITY, ABSOLUTE INTERACTION, AND ALGEBRAIC OPTIMIZATION

NIDHI KAIHNSA, YUE REN*, MOHAB SAFEY EL DIN, AND JOHANNES W. R. MARTINI

ABSTRACT. We consider a measure of cooperativity based on the minimal interaction required to generate an observed titration behavior. We describe the corresponding algebraic optimization problem and show how it can be solved using the nonlinear algebra tool SCIP. Moreover, we compute the minimal interactions and minimal molecules for several binding polynomials that describe the oxygen binding of various hemoglobins under different conditions. We compare their minimal interaction with the maximal slope of the Hill plot, and discuss similarities and discrepancies with a view towards the shapes of the binding curves.

1. INTRODUCTION

Interaction between components is a fundamental feature of biological systems. While a simple system of independent subunits is completely defined by its subunits, a complex system with interactions is more than the sum of its parts. A classical example of a small biological system with non-trivial interaction is hemoglobin with its four binding sites for oxygen [BHK04; Bar13; Hil13]. The *ligand* oxygen binds to the four binding sites of the *(target) molecule* hemoglobin and the interaction can be seen on the *overall (isotherm) binding curve*, which relates the average ligand saturation inside the system to the ligand concentration outside the system at constant temperature [Hil85]. While the binding curve of an independent system is linear, the binding curve of hemoglobin has a sigmoidal shape [Bar13; Hil13], see Figure 1. The sigmoidal shape implies that the extremal states of full and zero saturation are more stable than the intermediate states of partial saturation. This phenomenon is commonly referred to as *cooperativity*, named after the intuition that bound ligands affect, either positively or negatively, the chances of new ligands binding to the still open sites (see [SL13] for a review and interpretations). Cooperativity is ubiquitous in nature. It is an essential trait for transport molecules, and its importance in the formation of multi-protein complexes [RRS17], in general signal transduction processes [Sal+16; LFSR09; MSS16] and in the regulation of noise [GMD09; MMD13] has been discussed.

When investigating cooperativity, one interesting problem is to connect macroscopic behavior to microscopic behavior. Here, macroscopic refers to the measured average saturation at varying ligand concentration, and microscopic refers to which individual sites are saturated and how the different sites interact. An example is the
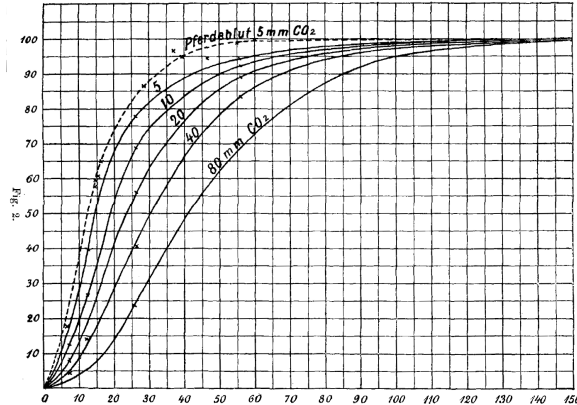
FIGURE 1. The oxygen concentration in dog and horse blood in relation to the ambient oxygen concentration [BHK04, Fig. 2].

aforementioned interpretation of occupied sites increasing the probability of ligands binding to unoccupied sites [SL13; OU04]. Ideally, such connections between microscopic and macroscopic behavior provide insight on the potential structure and the inner workings of the molecule.

However, the overall binding curve alone does not allow for much inference on the individual site binding curves, similar to how a large sum allows for little inference on its summands. Given a fixed overall binding curve, there generally is no unique molecule, defined by binding and interaction energies, with the prescribed macroscopic binding behavior, but infinitely many. Thus many works feature additional constraints which reduce the ambiguity, such as considering only small system with one or two binding sites [HA09] or symmetric systems with identical binding sites [Abe16; Ron+19]. These constraints have the added benefit of simplifying the mathematical problem, allowing for easier inference on interaction energy terms, conditional probabilities, correlations or other quantities. Ideally, constraints should be motivated by knowledge on the chemical structure of the molecule.

A biologically and mathematically interesting question is for instance whether there exist a molecule with independent sites for a fixed overall binding curve. Assuming evolutionary pressure on the transport properties of hemoglobin, it should not be possible to model the system's overall binding curve through an independent system, since otherwise the tetrameric structure of hemoglobin would hold no advantage over four monomers.

The decoupled sites representation [OCU01; MU13] is a central theorem in this context, which states that every overall binding curve can be generated by a unique *hypothetical molecule* with possibly non-real binding energies and independent binding sites. The binding energies are determined by the roots of the *binding polynomial* of the overall binding curve. In particular, the realness of the roots show whether

interaction is necessary to achieve the overall binding behavior and thus provide a qualitative definition of cooperativity.

In case interaction is required to generate the overall binding curve, a natural question is how to quantify the degree of cooperativity. This is classically done using the maximal slope of the Hill plot, which is the plot of binding curve on a logarithmic x-axis (see Figure 6). A new approach [Mar17] is to look at the minimal interaction required to generate the considered overall binding curve. The minimal interaction quantifies the minimal deviation from an independent system that is necessary to explain the observed overall binding behavior.

The manuscript in hand studies the problem of finding molecules generating a given overall binding curve with minimal interaction. We show that looking for the minimal interaction energy required to generate a given overall titration behavior, leads to an algebraic optimization problem which can be partially tackled using software such as SCIP [Gle+18]. We calculate the minimal molecules of various binding curves in existing datasets [CRBG86; Ike+83; Ima73]. These binding curves describe the oxygen binding of hemoglobins under different chemical treatments and under different temperatures. We compare their minimal interaction with the maximal slope of the Hill plot, and discuss similarities and discrepancies with a view towards the shapes of the binding curves.

2. Recapitulation of the mathematical framework

In this section, we briefly review the model for ligand binding based on the grand canonical ensemble of statistical mechanics and recall the notion of minimal absolute interaction from [Mar17]. We refer the reader to [Ben13; WG90; Hil85] for detailed expositions on the model.

Definition 2.1 A *hypothetical molecule* W with n sites is a complex point, whose coordinates are indexed by the subsets of $[n] := \{1, \ldots, n\}$ and with $w_\emptyset = 1$:

$$W = (w_I)_{I \subseteq [n]} \in \mathbb{C}^{2^n}.$$

A *(real) molecule* is a hypothetical molecule whose coordinates are real and positive.

We refer to w_I as *binding energies* if $|I| = 1$ and *interaction energies* if $|I| > 1$. For the sake of brevity, we abbreviate $w_{\{i_1,\ldots,i_r\}}$ to $w_{i_1 \ldots i_r}$ for $i_1 < \cdots < i_r$ (see Fig. 2).

Note that we refer to $w_{i_1 \ldots i_r}$ as *energy* even though it is actually representing the exponential of an energy difference [Ben13; WG90; Sch75]. To be precise:

- The binding energy w_i encodes the difference between the free energies of the completely unsaturated state and the state with a single ligand bound to site i.
- The interaction energy $w_{i_1 i_2}$ encodes the difference between the free energy when sites i_1, i_2 are occupied and the sum of the energies when either i_1 or i_2 is occupied.
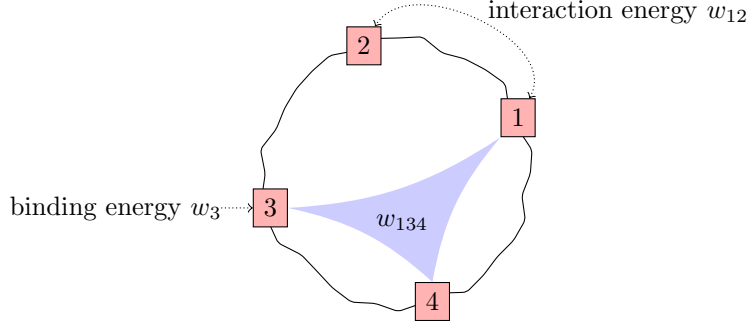
FIGURE 2. A molecule with 4 binding sites.

- Analogously, $w_{i_1 \dots i_r}$ encodes the difference between the free energy when sites $i_1, \dots, i_r$ are occupied and the sum of all energies when a proper subset of $i_1, \dots, i_r$ are occupied.

As such, the only "sensible" values for $w_{i_1 \dots i_r}$ are positive real numbers. However, mathematically it is easier to work with hypothetical molecules, and they allow for interesting phenomena such as the decoupled sites representation [OCU01; MU13]. In this article, we will mainly focus on real molecules.

Definition 2.2 The *overall binding curve* of a real molecule $W = (w_I)_{I \subseteq [n]} \in (\mathbb{R}_{>0})^{2^n}$ with n sites is defined as

$$\Psi(\Lambda) = \frac{n a_n \Lambda^n + (n-1) a_{n-1} \Lambda^{n-1} + \cdots + a_1}{a_n \Lambda^n + a_{n-1} \Lambda^{n-1} + \cdots + a_1 + a_0},$$

also commonly known as the Adair equation [ABJ25; SL13], with positive real coefficients

$$a_k := \sum_{|I|=k} \prod_{I' \subseteq I} w_{I'} \in \mathbb{R}_{>0}.$$

The *binding polynomial* $\Phi(W)$ is the univariate polynomial of degree n in its denominator,

$$\Phi(W) := a_n \Lambda^n + \cdots + a_1 \Lambda + a_0.$$

The binding polynomial Φ uniquely determines the overall binding curve Ψ.

Example 2.3 Let $W = (w_I)_{I \subseteq [3]} \in (\mathbb{R}_{>0})^{2^3}$ be a molecule with 3 sites. The binding polynomial $\Phi(W) = a_3 \Lambda^3 + a_2 \Lambda^2 + a_1 \Lambda + a_0$ is a real univariate polynomial of degree 3 whose coefficients are given by

$$\begin{aligned}
a_0 &= w_\emptyset = 1, \\
a_1 &= w_1 + w_2 + w_3, \\
a_2 &= w_1 w_2 w_{12} + w_1 w_3 w_{13} + w_2 w_3 w_{23}, \\
a_3 &= w_1 w_2 w_3 w_{12} w_{13} w_{23} w_{123}.
\end{aligned}$$

Note that two different molecules may have the same binding polynomial and thus the same binding curve, e.g.,

$$
\begin{array}{ccccccccc}
 & w_\emptyset & w_1 & w_2 & w_3 & w_{12} & w_{13} & w_{23} & w_{123} \\
\mathrm{W} = (& 1 & 3 & 1 & 1 & 1 & 1 & 2 & 2 &), \\
\mathrm{W}' = (& 1 & 2 & 2 & 1 & 1 & 1 & 1 & 3 &).
\end{array}
$$

Hence, the map Φ which maps a molecule with n sites to its binding polynomial is not injective.

Computing hypothetical or real molecules W with a fixed binding polynomial P can be a hard task, though there are well established algorithms and software for both. Let $\mathcal{M}_\mathbb{C}(P)$ resp. $\mathcal{M}_\mathbb{R}(P)$ denote the set of hypothetical resp. real molecules W with $\Phi(\mathrm{W}) = P$. Then $\mathcal{M}_\mathbb{C}(P)$ is a *algebraic variety*, i.e., a set of complex points given by polynomial equations, while $\mathcal{M}_\mathbb{R}(P)$ is a *semialgebraic set*, i.e., a set of real points given by polynomial equations and inequalities.

Computing a hypothetical molecule in $\mathcal{M}_\mathbb{C}(P)$ is generally easier than $\mathcal{M}_\mathbb{R}(P)$. It can be done purely numerically, e.g., using numerical algebraic geometry and software such as BERTINI, PHCPACK, HOM4PS or JULIAHOMOTOPYCONTINUATION. Computing a real molecule in $\mathcal{M}_\mathbb{R}(P)$ is a significantly more involved task. There are many approaches to it and most require some symbolic computation, such as RAGLIB which uses a mix of symbolic and numerical computations.

However, both $\mathcal{M}_\mathbb{C}(P)$ and $\mathcal{M}_\mathbb{R}(P)$ are of high dimension and degree [RMT19], so that it is unclear what information can be read off a randomly computed molecule. Rather, we need sensible constraints which gives us molecules of interest. Since the aim of this paper is to study cooperativity in terms of the interaction between the binding sites, we introduce the notions of absolute interaction and minimal interaction:

Definition 2.4 The *(absolute) interaction* of a molecule $\mathrm{W} = (w_I) \in (\mathbb{R}_{>0})^{2^n}$ is given by

$$
\|\mathrm{W}\| := \prod_{|I|>1} \max\left(w_I, w_I^{-1}\right).
$$

Since w_I represents the exponential of a difference of actual energies, which can be positive or negative, $\max(w_I, w_I^{-1})$ represents the exponential of the absolute value of that difference. Hence, the absolute interaction $\|\mathrm{W}\|$ represents the exponential of the sum of the absolute values of all energy differences. The minimal value $\|\mathrm{W}\|$ can adopt is 1 which corresponds to a minimal physical interaction energy of 0.

Definition 2.5 Given a binding polynomial P of degree n, the *minimal interaction* is

$$
\|P\| := \min\left\{ \|\mathrm{W}\| \ \Big| \ \mathrm{W} \in (\mathbb{R}_{>0})^{2^n} \text{ with } \Phi(\mathrm{W}) = P \right\}.
$$

We call a molecule W *minimal*, if $\|W\| = \|\Phi(W)\|$.

In other words, the minimal interaction $\|P\|$ is the minimal interaction required to generated the binding polynomial P. Thus it quantifies the minimal deviation from an independent system that is required to explain the observed overall binding behavior. The following proposition states that the minimum exists.

Proposition 2.6 ([Mar17, §4]) *For any univariate polynomial P of degree n with positive coefficients there exists a molecule $W \in (\mathbb{R}_{>0})^{2^n}$ with n sites such that $\Phi(W) = P$ and $\|W\| = \|P\|$.*

Since cooperativity is a property of the binding curve which emerges from the interaction between the sites, the minimal interaction of a binding polynomial is a natural candidate for quantifying cooperativity. In addition, [Mar17, §4] shows that it has several properties that are desirable of a quantifier.

Example 2.7 Consider the following binding polynomials:

$$
\begin{aligned}
P_1 &:= 4\Lambda^3 + 3\Lambda^2 + 2\Lambda + 1, \\
P_2 &:= 6\Lambda^3 + 7\Lambda^2 + 4\Lambda + 1 = (2\Lambda + 1)(3\Lambda^2 + 2\Lambda + 1), \\
P_3 &:= 6\Lambda^3 + 11\Lambda^2 + 6\Lambda + 1 = (\Lambda + 1)(2\Lambda + 1)(3\Lambda + 1).
\end{aligned}
$$

A brief computation reveals that the minimal absolute interactions are

$$\|P_1\| = 13.50, \qquad \|P_2\| = 3, \qquad \|P_3\| = 1.$$

The computation for $\|P_1\|$ is explicitly shown in Example 3.3. Figure 3 illustrates the minimal molecules for each binding polynomial. Since P_3 factorizes into three real linear factors, its minimal molecule has only trivial interactions.

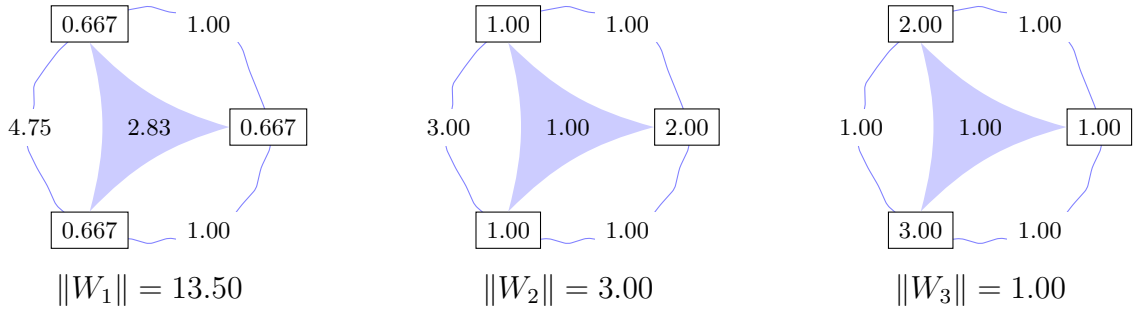$$\|W_1\| = 13.50 \qquad \|W_2\| = 3.00 \qquad \|W_3\| = 1.00$$

FIGURE 3. Minimal molecules for binding polynomial P_1, P_2, P_3.

3. COMPUTING THE MINIMAL INTERACTION

In this section, we consider the problem of computing the minimal absolute interaction for a given binding polynomial P:

$$
\begin{aligned}
&\underset{W}{\text{minimize}} && \|W\| \\
&\text{subject to} && \Phi(W) = P
\end{aligned}
\tag{1}
$$

For the sake of clarity, we will restrict ourselves to the case of molecules with 4 binding sites, though our arguments also apply to the general case. We will discuss what makes Problem (1) challenging, and what can or cannot be done towards solving it from a practical point of view.

3.1. **The initial problem.** Problem (1) is commonly called the *standard form* of a continuous optimization problem with $\|W\|$ being the *objective function* and $\Phi(\mathrm{W}) = P$ the *constraints*. The set of all molecules W which satisfy the constraints is the *feasible set*.

Generally speaking, there are many approaches for solving optimization problems depending on the type of the objective function and the constraints, which may also include inequalities. If both are linear, the problem is also known as a *linear program*. If the objective function is quadratic and the constraints are linear, the problem is also called a *quadratic program*. Both types of optimization problems are well-studied with plenty of literature and software solutions. If both objective function and constraints are polynomial, the problem is a *polynomial optimization problem* or an *algebraic optimization problem*. This is the type of problem that needs to be solved for Problem 1, and solving it generally involves using sums of squares approaches [Las01; Par03] or critical point methods [GS14].

For molecules with 4 binding sites specifically, Problem (1) is of the following form for a given binding polynomial $P = a_4\Lambda^4 + a_3\Lambda^3 + a_2\Lambda^2 + a_1\Lambda$:

$$
\begin{aligned}
& \underset{(w_I)_{I \subseteq [4]}}{\text{minimize}} && \prod_{|I|>1} \max\left(w_I, w_I^{-1}\right) \\
& \text{subject to} && a_1 = w_1 + w_2 + w_3 + w_4, \\
& && a_2 = w_1 w_2 w_{12} + w_1 w_3 w_{13} + w_1 w_4 w_{14} \\
& && \qquad + w_2 w_3 w_{23} + w_2 w_4 w_{24} + w_3 w_4 w_{34}, \\
& && a_3 = w_1 w_2 w_3 w_{12} w_{13} w_{23} w_{123} + w_1 w_2 w_4 w_{12} w_{14} w_{24} w_{124} \\
& && \qquad + w_1 w_3 w_4 w_{13} w_{14} w_{34} w_{134} + w_2 w_3 w_4 w_{23} w_{24} w_{34} w_{234}, \\
& && a_4 = w_1 w_2 w_3 w_4 w_{12} w_{13} w_{14} w_{23} w_{24} w_{34} w_{123} w_{124} w_{134} w_{234} w_{1234}.
\end{aligned}
\tag{2}
$$

Problem (2) is extremely hard to solve because of two reasons:

- the objective function contains 11 maxima,
- the constraints contain polynomials of degree up to 15.

In the following, we will briefly discuss the two challenges separately.

3.2. **Resolving the maxima in the objective function.** The maxima in the objective function of Problem (2) can be resolved in two ways.

One standard way is to reformulate the objective function as a maximum of polynomial functions and linearize the objective function through additional constraints: Abbreviating $\binom{[4]}{>1} := \{I \subseteq [4] \mid |I| > 1\}$, we define for each set of interactions

$\mathcal{I} \subseteq \binom{[4]}{>1}$

$$f_{\mathcal{I}}((w_I)_{I \subseteq [4]}) := \Big(\prod_{\substack{I \in \binom{[4]}{>1} \\ I \in \mathcal{I}}} w_I^{-1} \Big) \cdot \Big(\prod_{\substack{I \in \binom{[4]}{>1} \\ I \notin \mathcal{I}}} w_I \Big),$$

so that

$$\|(w_I)_{I \subseteq [4]}\| = \max_{\mathcal{I} \subseteq \binom{[4]}{>1}} f_{\mathcal{I}}((w_I)_{I \subseteq [4]}).$$

Then Problem (2) can be written as:

$$
\begin{aligned}
\underset{(w_I)_{I \subseteq [4]}}{\text{minimize}} \quad & r \\
\text{subject to} \quad & f_{\mathcal{I}}((w_I)_{I \subseteq [4]}) \leq r \quad \text{for all } \mathcal{I} \subseteq \binom{[4]}{>1}, \\
& a_1 = w_1 + w_2 + w_3 + w_4, \\
& a_2 = w_1 w_2 w_{12} + w_1 w_3 w_{13} + w_1 w_4 w_{14} \\
& \qquad + w_2 w_3 w_{23} + w_2 w_4 w_{24} + w_3 w_4 w_{34}, \\
& a_3 = w_1 w_2 w_3 w_{12} w_{13} w_{23} w_{123} + w_1 w_2 w_4 w_{12} w_{14} w_{24} w_{124} \\
& \qquad + w_1 w_3 w_4 w_{13} w_{14} w_{34} w_{134} + w_2 w_3 w_4 w_{23} w_{24} w_{34} w_{234}, \\
& a_4 = w_1 w_2 w_3 w_4 w_{12} w_{13} w_{14} w_{23} w_{24} w_{34} w_{123} w_{124} w_{134} w_{234} w_{1234}.
\end{aligned}
\tag{3}
$$

This resolves all maxima in the objective function at the cost of introducing 2^{11} additional constraints, one for each $\mathcal{I} \subseteq \binom{[4]}{>1}$.

Alternatively, one can decompose the space of molecules $(\mathbb{R}_{>0})^{2^4}$ into 2^{11} regions in which either $w_I \geq 1$ or $w_I \leq 1$ for all $\mathcal{I} \subseteq \binom{[4]}{>1}$. For each $\mathcal{I} \subseteq \binom{[4]}{>1}$, we define a region

$$O_{\mathcal{I}} := \Big\{ (w_I)_{I \subseteq [n]} \in (\mathbb{R}_{>0})^{2^4} \mid \text{for all } I \in \binom{[4]}{>1} : \ w_I \geq 1 \text{ if } I \notin \mathcal{I} \text{ and } w_I \leq 1 \text{ if } I \in \mathcal{I} \Big\}$$

so that

$$(\mathbb{R}_{>0})^{2^4} = \bigcup_{\mathcal{I} \subseteq \binom{[4]}{>1}} O_{\mathcal{I}} \quad \text{and} \quad \|W\| = f_{\mathcal{I}}(W) \text{ for } W \in \mathcal{O}_{\mathcal{I}}.$$

Finding the minimal absolute interaction in $O_{\mathcal{I}}$ hence becomes a straight-forward polynomial optimization problem with few constraints.

$$
\begin{aligned}
\underset{W \in O_{\mathcal{I}}}{\text{minimize}} \quad & f_{\mathcal{I}}(W) \\
\text{subject to} \quad & \Phi(W) = P
\end{aligned}
\tag{4}
$$

However, this approach requires solving 2^{11} polynomial optimization problems, one for each region $\mathcal{O}_{\mathcal{I}}$, which is not easily feasible.

One can reduce the number of regions that require consideration slightly by exploiting the intrinsic symmetry of the problem.
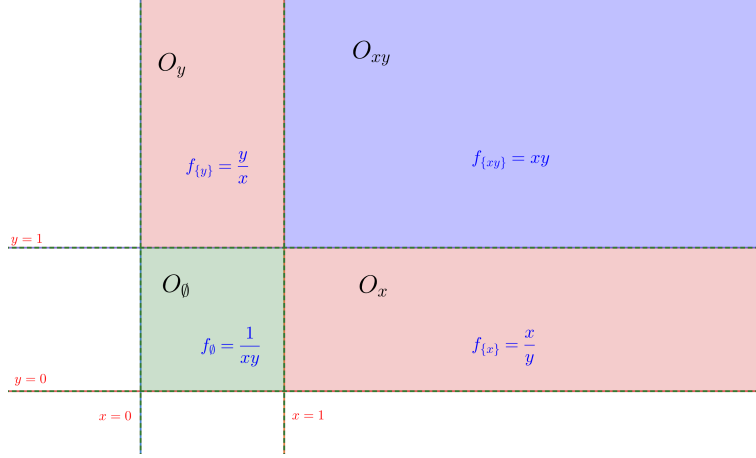
FIGURE 4. Partitions in plane

Example 3.1 Consider for example the following polynomial optimization problem:

$$\underset{(x,y)\in\mathbb{R}^2_{>0}}{\text{minimize}} \quad \max(x, x^{-1}) \cdot \max(y, y^{-1})$$
$$\text{subject to} \quad a_1 = x + y,$$
$$a_2 = x \cdot y.$$

The space $\mathbb{R}^2_{>0}$ can be decomposed into 4 regions on which the objective function is xy, $\frac{x}{y}$, $\frac{y}{x}$, $\frac{1}{xy}$ respectively, see Figure 4. As both objective function and constraints are symmetric with respect to permuting x and y, the minimum of $f_{\{y\}}$ in O_y is same as the minimum of $f_{\{x\}}$ in O_x. Therefore, it suffices to minimize over three regions instead of all four.

For Problem (2) specifically, this means:

Proposition 3.2 *There exists a natural action of the symmetric group S_4 on the space of molecules $(\mathbb{R}_{>0})^{2^4}$ by permuting the 4 binding sites:*

$$S_4 \times (\mathbb{R}_{>0})^{2^4} \longrightarrow (\mathbb{R}_{>0})^{2^4}, \quad (\sigma, (w_I)) \longmapsto \sigma \cdot (w_I) := (w_{\sigma(I)}).$$

This action permutes the regions $\mathcal{O}_{\mathcal{I}}$ and preserves binding polynomials as well as absolute interactions, i.e., $\Phi(\sigma \cdot W) = \Phi(W)$ and $\|\sigma \cdot W\| = \|W\|$ for all $W \in (\mathbb{R}_{>0})^{2^4}$ and all $\sigma \in S_4$.

Instead of computing in all 2^{11} regions, it therefore suffices to consider one region per S_4-orbit of regions. However, this still requires 249 regions to be checked, which is still too much for practical purposes.

3.3. **Reducing the degree of the constraints.** Given a binding polynomial P, the polynomial equations in Problem (3) arising from its coefficients $a_1, \ldots, a_4$ pose a serious computational problem. To sidestep this issue we now introduce new coordinates $s_I := \prod_{I' \subseteq I} w_{I'}$. Here, s_I represents the free energy difference of microstate

I to the fully unoccupied state. On the positive orthant $(\mathbb{R}_{>0})^{2^n}$, this defines a bijection

$$
\begin{array}{ccc}
W = (w_I) & \xrightarrow{\ \varphi\ } & \left(\displaystyle\prod_{I' \subseteq I} w_{I'}\right) \\[2mm]
\cap & & \cap \\[1mm]
(\mathbb{R}_{>0})^{2^n} & \longleftrightarrow & (\mathbb{R}_{>0})^{2^n} \\[1mm]
\cup & & \cup \\[1mm]
\left(\displaystyle\prod_{I' \subseteq I} s_{I'}^{(-1)^{|I \setminus I'|}}\right) & \xleftarrow{\ \varphi^{-1}\ } & (s_I) = S
\end{array}
$$

In the new coordinates, the formerly polynomial constraints $a_k = \sum_{|I|=k} \prod_{I' \subseteq I} w_{I'}$ become linear constraints $a_k = \sum_{|I|=k} s_I$ for $k = 1, \dots, n$. While the functions $f_{\mathcal{I}}$ become more complicated in the new coordinates, they still remain monomial and thus it has minimal impact on the runtime: For example, for $\mathcal{I} = \{\{1,2,3\}\}$

$$
f_{\mathcal{I}}(W) = w_{123}^{-1} \cdot \prod_{\substack{|I|>1, \\ I \neq \{1,2,3\}}} w_I \quad \text{becomes} \quad f_{\mathcal{I}} \circ \varphi^{-1}(S) = \frac{s_{1234} s_{12}^2 s_{13}^2 s_{23}^2}{s_{123}^2 s_1^3 s_2^3 s_3^3 s_4}.
$$

In the new coordinates, Problem (3) becomes:

$$
\begin{aligned}
\underset{(w_I)_{I \subseteq [4]}}{\text{minimize}} \quad & r \\
\text{subject to} \quad & f_{\mathcal{I}} \circ \varphi^{-1}(S) \leq r \quad \text{for all } \mathcal{I} \subseteq \{I \subseteq [n] \mid |I| > 1\}, \\
& a_1 = s_1 + s_2 + s_3 + s_4, \\
& a_2 = s_{12} + s_{13} + s_{14} + s_{23} + s_{24} + s_{34}, \\
& a_3 = s_{123} + s_{124} + s_{134} + s_{234}, \\
& a_4 = s_{1234}.
\end{aligned}
\tag{5}
$$

3.4. **Solving the polynomial optimization problem.** We will use SCIP [Gle+18] to solve polynomial optimization Problem (5), which is currently one of the fastest non-commercial solvers for non-linear programming. It uses a branch and bound method to solve the optimization problem with non-linear constraints. Example 3.3 shows a computation using SCIP for a molecule with 3 binding sites.

Example 3.3 Consider the polynomial P_1 of Example 2.7. Figure 5 shows the full input on the left and the partial output on the right. In the input, the first constraints c1, c2, c3 enforce $\sum_{|I|=k} s_I = a_k$ for $i = 1, 2, 3$. The remaining constraints c4 to c19 enforce $f_{\mathcal{I}} \circ \varphi^{-1}(S) \leq r$. For example, c19 states that $s_1 s_2 s_3 / s_{123} \leq r$ which is equivalent to $f_{\emptyset}(W) = (w_{123} w_{12} w_{13} w_{23})^{-1} \leq r$ in the coordinates w_I. The output states an approximate optimal value of $r = 13.5$ and lists all values of s_I of the minimal molecule, which gave rise to the numbers in Example 2.7.

```
Minimize                                SCIP version 6.0.0
obj: r                                  Copyright (C) 2002-2018 ZIB Berlin
Subject to
c1: s1+s2+s3=2                          SCIP> read input.pip
c2: s12+s13+s23=3
c3: s123=4                              SCIP> opt
c4: s123-r*s1*s2*s3<=0                   SCIP Status  : problem is solved
c5: s1*s2*s123-r*s3*s12^2<=0            Solving Time : 2.32
c6: s1*s3*s123-r*s2*s13^2<=0            Solving Nodes: 587
c7: s2*s3*s123-r*s1*s23^2<=0            Primal Bound : +1.349997004816e+01
c8: s12^2*s13^2*s23^2                   Dual Bound   : +1.349997004816e+01
    -r*s1^3*s2^3*s3^3*s123<=0           Gap          : 0.00 %
c9: s1^3*s2*s3*s123-r*s12^2*s13^2<=0
c10: s1*s2^3*s3*s123-r*s12^2*s23^2<=0   SCIP> display solution
c11: s13^2*s23^2-r*s1*s2*s3^3*s123<=0   objective value: 13.4999700481621
c12: s1*s2*s3^3*s123-r*s13^2*s23^2<=0   y               13.4999700481621
c13: s12^2*s23^2-r*s1*s2^3*s3*s123<=0   s1              0.666630230040994
c14: s12^2*s13^2-r*s1^3*s2*s3*s123<=0   s2              0.66617295525322
c15: s1^3*s2^3*s3^3*s123               s3              0.667196814705786
    -r*s12^2*s13^2*s23^2<=0            s12             0.444091987464541
c16: s1*s23^2-r*s2*s3*s123<=0           s13             0.444773633975962
c17: s2*s13^2-r*s1*s3*s123<=0           s23             2.1111343785595
c18: s3*s12^2-r*s1*s2*s123<=0           s123            4
c19: s1*s2*s3-r*s123<=0                 [...]

Bounds
0<s1<2
0<s2<2
0<s3<2
0<s12<3
0<s13<3
0<s23<3
4<=s123<=4
1<=r
End
```

FIGURE 5. Computing the minimal absolute interaction for a molecule with 3 sites using SCIP.

Unfortunately, the 2^{11} constraints in Problem (5) make it practically impossible to solve directly. Instead, we focus on computing an upper bound b_+ and a lower bound b_- for the minimal interaction. Both bounds are easier to compute, because they arise from problems with significantly less constraints. Our aim is to obtain bounds which either coincide or are sufficiently close to each other to allow us to compare the different minimal interactions with each other.

The upper bound b_+. Our upper bound b_+ for the solution to Problem (5) is the solution to Problem (4) for a region $O_{\mathcal{I}}$. In other words, instead of minimizing the interaction over all molecules in $(\mathbb{R}_{>0})^{2^4}$, we only minimize it over the molecules in $O_{\mathcal{I}}$. It is an upper bound as the global minimum need not be in the region that we have chosen. Computing the solution to Problem (4) still required changing to the new coordinates (s_I). Our region of choice was $O_{\emptyset}$, i.e., the region where $w_I \geq 1$ for all $I \in \binom{[4]}{>1}$, as it was the only region in which all computations terminated successfully during testing.

The lower bound b_-. Our lower bound b_- is the solution to a relaxed Problem (5) with some of the 2^{11} constraints of the form $f_{\mathcal{I}} \circ \varphi^{-1}(S) \leq r$ removed. It is a lower bound since removing constraints enlarges the feasible set over which the minimum is computed. Appendix A contains a list of constraints that remain in the final computation. We picked a region $O_{\mathcal{I}}$ in each S_4 orbit of regions and only considered constraints given by functions $f_{\mathcal{I}}$ that are dominant in one of the regions.

4. Experimental Results

In this section, we present the results of computational experiments on data from [Ima73; Ike+83] which is also summarized in [CRBG86]. Section 4.1 summarizes the origins of the data, Section 4.2 contains the bounds on the minimal interactions, and Section 4.3 contains molecules attaining the upper bound.

4.1. **Data.** The first data set consists of eight oxygen binding polynomials (P_1 to P_8) of human adult hemoglobin in its native form and chemically treated with iodoacetamide, N-ethylmaeimide, and carboxypeptidase A. The oxygen binding was observed both in absence and in presence of 2,3-diphosphoglycerate (DPG) [Ima73]. One of the goals was to determine how the chemical treatments and DPG affect the binding affinity of hemoglobin, the results are indicated by $\succ$ in Table 1.

The second data set contains five binding polynomials (P_9 to P_{13}) of native hemoglobin HbII of *Scapharca inaequivalvis* measured at different temperatures [Ike+83, Table 3], see Table 2. One of the goals was to investigate how temperature affects cooperativity in oxygen binding.

Table 3 lists the coefficients a_1, a_2 and a_3 of the binding polynomials P_1 to P_{13}, the coefficients a_0 and a_4 are normalized to 1. Furthermore, it lists the maximal slope to the Hill plot n_{max}, as reported by [Ima73] and [Ike+83], which relates the variance of the probability distribution on the macrostates $\{0, 1, 2, 3, 4\}$ at the respective ligand activity to the variance of a binomial distribution with the same mean [Abe05; MDH16]. A value larger than 1 indicates a variance higher than the maximal value an independent system can generate.

4.2. **Minimal absolute interactions.** Table 3 lists the upper (b_+) and lower (b_-) bounds for the minimal absolute interaction for the binding polynomials P_1 to P_{13}.

Human Hb	in the absence of 2mM DPG		in the presence of 2mM DPG
untreated	P_1	$\prec$	P_2
	$\curlyvee$		$\curlyvee$
treated with iodoacetamide	P_3	$\prec$	P_4
	$\curlyvee$		$\curlyvee$
treated with N-ethylmaeimide	P_5	$\prec$	P_6
	$\curlyvee$		$\curlyvee$
treated with carboxypeptidase A	P_7	$\prec$	P_8

TABLE 1. Conditions under which binding polynomials 1 to 8 were measured. $P_i \prec P_j$ means P_i that the Hill plot of P_i has a lower maximal slope compared to P_j.

Clam HbII	at 10°	at 15°	at 20°	at 25°	at 30°
	P_9	P_{10}	P_{11}	P_{12}	P_{13}

TABLE 2. Temperature in degree Celsius at which binding polynomials 9 to 13 were determined.

As explained in Section 3.4, solving Problem (5) for $n = 4$ is very hard, which is why we computed bounds instead. The upper bound is the minimal absolute interaction in the region $\mathcal{O}_\emptyset$, i.e., the minimal interaction in the region where all interactions are 1 or larger. The lower bound comes from a relaxation of Problem (5) which enlarges the set of molecules over which the minimum is computed.

Looking at the bounds b_+ and b_- in Table 3, we observe two things:

(1) Apart from P_2, P_3, and P_8, the bounds determine the minimal interaction up to a precision of 15%.

(2) The bounds are almost sufficient to rank the minimal interactions of P_1 to P_8 and P_9 to P_{13}:

$$\|P_6\| > \|P_4\| > \|P_1\| > \|P_3\| > \|P_5\| > \|P_7\| > \|P_8\|$$
$$\text{and} \qquad \|P_9\| = \|P_{10}\| = \|P_{12}\| > \|P_{11}\| > \|P_{13}\|$$

with $\|P_2\|$ possibly being larger than $\|P_6\|$ or between $\|P_4\|$ and $\|P_1\|$.

Comparing the ranking of the binding polynomials by their minimal interactions with the ranking by the maximal slope of their Hill plot $n_{\max}$ reveals the following similarities and differences:

P_9 to P_{13}. We begin by looking at polynomials P_9 to P_{13} as there are prominent differences between minimal interactions and maximal slopes of the Hill plots: While

	a_1	a_2	a_3	b_+	b_-	n_{max}
P_1	0.835	0.379	0.541	527	527	2.51
P_2	0.789	0.154	0.0648	3322	1600	3.09
P_3	1.42	2.42	0.752	111	63	1.63
P_4	0.647	0.568	0.0986	2002	1460	2.71
P_5	2.0	2.31	2.04	16	16	1.44
P_6	0.539	0.909	0.554	3033	3030	2.27
P_7	3.47	4.74	2.76	2.27	1.68	1.15
P_8	3.26	5.36	2.23	7.64	2.19	1.23
P_9	1.4	1.0	0.62	66	66	2.08
P_{10}	1.4	0.96	0.60	66	66	2.10
P_{11}	1.2	0.93	0.70	123	123	2.08
P_{12}	1.4	0.95	0.62	66	66	2.10
P_{13}	1.1	0.98	0.59	175	175	2.12

TABLE 3. Binding polynomials, bounds on the minimal absolute interaction and n_{max} as reported by [Ike+83; Ima73]. Coefficients a_0 and a_4 are normalized to 1 for all polynomials.

all maximal slopes are approximately equal, there are stark differences in the minimal interactions. This is because the minimal interaction depends on the entire binding curve, while the maximal slope does not. Figure 6 illustrates that while the binding curves of P_9 to P_{13} may have identical maximal slopes in logarithmic coordinates, there are subtle differences between both curves. The binding curve of P_{11} has a stronger sigmoidal shape, which is why more interaction is necessary to generate the binding behaviour.
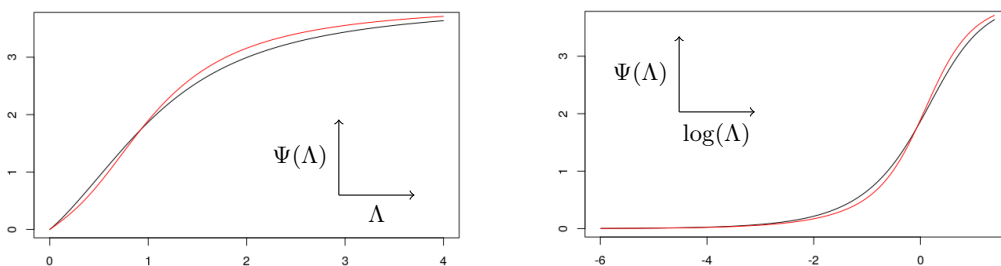
FIGURE 6. The binding curves of P_9 (black) and P_{11} (red) in normal (left) and logarithmic (right) scale.

P_1 to P_8. While the ranking by minimal interaction is different to the ranking by maximal slope, we retain nearly all important relations in Table 1, which indicate how the chemicals affect the binding behaviour. The only exceptions are P_4 and

P_6, where P_4 has a higher maximal slope and P_6 has a higher minimal interaction. This is not surprising, as Figure 7 illustrates that the binding curve P_4 has a higher maximal slope, while the binding curve P_6 has a stronger curvature.

Another explanation can be found in the coefficients of P_4 and P_6: In P_6, coefficient a_2 is larger than a_1 and a_3 and almost same size as $a_4 = 1$. This means that the distribution on the macrostates stabilizes around macrostate 2. This binding behaviour does not only requires more interaction between the binding sites, it also prevents large variances of the distribution on the macrostates and thus leads to a reduction of n_{max}. In P_4, coefficients a_1, a_2, a_3 are smaller than $a_4 = 1$. This means that the distribution on the macrostates leans toward macrostate 4, which leads to a higher value of n_{max}.
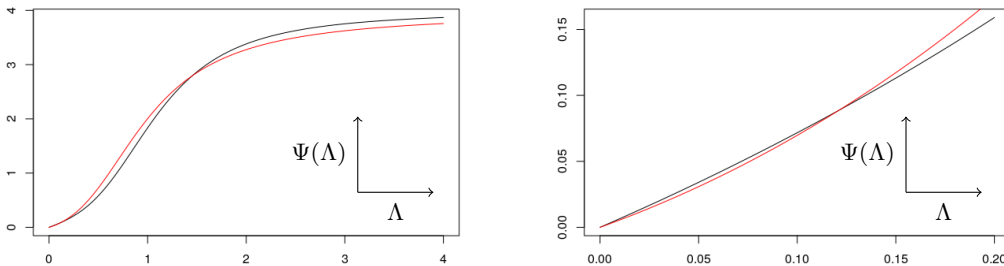
FIGURE 7. The binding curves of P_4 (black) and P_6 (red) on $[0, 4]$ (left) and zoomed in on $[0, 0.2]$ (right). Note how the red curve starts and ends lower as the black curve, but overtakes it in between.

4.3. **Molecules solving the optimization problem.** Table 4 lists all molecules realizing the upper bound b_+ of the minimal interaction. We see that the many interaction energies are close to 1. This is not surprising since we are minimizing the exponential of an L_1-norm and L_1 optimizations are known to provide sparse solutions [Tib96].

Looking at the structure of the solutions in more detail, we see that most solutions include a non-trivial weight for a pairwise interaction (w_{34}). The only polynomial whose molecule possesses trivial pairwise interaction is P_2 which has the highest n_{max}. This is plausible as a high n_{max} means a high variance of the macroscopic distribution, which in turn implies more weights towards macrostates 0 and 4. And any nontrivial pairwise interaction would put more weight on macrostate 2.

Moreover, we mentioned how the macroscopic distribution for P_6 stabilizes around 2 while the macroscopic distribution for P_4 leans towards 4. This can be seen in their minimal molecules in $O_\emptyset$, as the molecule for P_6 has large w_{34} and the molecule for P_4 has a large w_{1234}.

	w_1	w_2	w_3	w_4	$w_{12}\cdots w_{24}$	w_{34}	w_{123}	w_{124}	w_{134}	w_{234}	w_{1234}
P_1	0.209	0.209	0.209	0.209	1.00	3.70	1.00	1.00	14.5	1.00	9.79
P_2	0.0822	0.0822	0.0822	0.543	1.00	1.00	6.60	1.00	14.7	1.00	34.3
P_3	0.179	0.179	0.531	0.531	1.00	7.12	1.00	1.00	1.00	1.00	15.6
P_4	0.101	0.101	0.223	0.223	1.00	9.42	1.00	1.00	1.00	1.00	212
P_5	0.500	0.500	0.500	0.500	1.00	4.24	1.00	1.00	1.69	1.69	1.32
P_6	0.135	0.135	0.135	0.135	1.00	45.1	1.00	1.00	2.49	2.49	10.9
P_7	0.569	0.569	1.17	1.17	1.00	1.30	1.00	1.00	1.00	1.00	1.75
P_8	0.265	0.265	1.36	1.36	1.00	2.06	1.00	1.00	1.00	1.00	3.70
P_9	0.350	0.350	0.350	0.350	1.00	3.16	1.00	1.00	1.97	1.97	5.43
P_{10}	0.350	0.350	0.350	0.350	1.00	2.84	1.00	1.00	2.11	2.11	5.26
P_{11}	0.300	0.300	0.300	0.300	1.00	5.33	1.00	1.00	2.24	2.24	4.60
P_{12}	0.350	0.350	0.350	0.300	1.00	2.76	1.00	1.00	2.26	2.26	4.73
P_{13}	0.275	0.275	0.275	0.275	1.00	7.96	1.00	1.00	1.66	1.66	8.01

TABLE 4. Molecules of P_1 to P_{13} realizing the upper bound b_+. All interactions strictly larger than 1 are highlighted in blue.

5. DISCUSSION AND OUTLOOK

5.1. **Cooperativity and minimal absolute interaction.** While cooperativity has always been a property of the macroscopic binding behavior, it is often interpreted on a microscopic level, for instance that "binding of a ligand molecule increases the receptor's apparent affinity, and hence increases the chance of another ligand molecule binding" [SL13]. There is however no straightforward connection between the macroscopic and microscopic properties, as generally there are infinitely many different molecules with the same binding curve. Thus creating a bridge between macroscopic and microscopic properties usually relies on additional restrictions, such as having indistinguishable sites.

In this paper, we considered the minimal interaction necessary to generate a fixed overall binding curve. In other words, we study the minimal deviation from a molecule with no interaction that is needed to create a prescribed binding behavior. We showed that computing the minimal interaction gives rise to a challenging polynomial optimization problem, and proposed easier ways for obtaining upper and lower bounds. Applying our techniques to overall binding curves and comparing our results to the maximal slope of the Hill plot, we observe:

(1) The upper and lower bounds were sufficient to rank all but one of the curves.
(2) The complete ranking of the curves by minimal interaction may differ from the ranking by maximal Hill slope, but most of the important relations are the same.

(3) In the cases were minimal interaction and maximal Hill slope disagreed, the curve with the higher minimal interaction displayed a stronger curvature (see Figures 6 and 7).

Conceptually, the minimal interaction considers the entire curve and captures any deviation from an independent system, while the maximal Hill slope only indicates an abnormally high variance of the macroscopic distributions. This is also reflected in our computations, as our results based on the minimal interaction largely coincides with the results based on the maximal Hill slope and the small differences in the results could be explained by signs of interaction outside the region of maximal Hill slope. Additionally, the corresponding minimal molecules may give valuable insights into the microscopic dependencies of the molecule. While the minimal molecule may not be a physically correct representation of the actual chemical molecule that was measured, in a way it describes the simplest possible set of interactions required for the binding curve. Note that we chose to minimize the L_1-norm of the actual interaction energies because it was the most natural norm in our framework. Alternatively, other L_p-norms could be used, and it is unclear whether they yield better qualitative results. However, with current methods the minimal interaction is much harder to compute than the maximal Hill slope and thus new mathematical approaches are required.

5.2. **The mathematical problem.** Ligand binding raises many intriguing questions in applied algebra and geometry, and stands to benefit equally from further insights into the intrinsic geometry of the problems. While previous works have explored the application of polynomial system solving techniques to find decoupled molecules in a situation with different types of ligands [RMT19; MSU13], this work explored the use of algebraic optimization to calculate the minimal interaction required to generate a certain binding behavior.

We illustrated how searching for the minimal interaction energy required to generate a binding curve leads to a problem in algebraic optimization. Even though, computational tools are readily available and the structure of the problem seems relatively simple on first sight, it becomes complicated when more than 3 binding sites are considered. In particular the fact that we had to calculate upper and lower bounds by a relaxed problem and were not able to close their gap completely, shows that our approach is not yet ready for a simple ready-to-use tool to quantify cooperativity. Even though we were able to rank the polynomials in Section 4.2, we were not able to compute precise minimal energies for all polynomials. We thus believe further research into the computational problem is warranted, though it may need to come from a different direction.

Exploiting symmetries. One possibility would be to show that the minimal molecule have some symmetry, which would reduce the number of variables (interaction energies) in our problem. While both Problem 5 and the molecules in Table 4 show some

kind of symmetry, it is mathematically not clear whether that always needs to be the case. There are results using *representation theory* which relate symmetries in polynomial optimization problems to symmetries in their solutions [RS18], however the assumptions on the symmetry in the polynomial optimization problem are quite strong and not satisfied by Problem 5.

Tropical Geometry. Tropical geometry [MS15] revolves around functions on the max-plus semiring, such as the absolute interaction in logarithmic coordinates:

$$
\begin{array}{ccc}
(\mathbb{R}_{>0})^{2^n} & \longleftrightarrow & \mathbb{R}^{2^n} \\
\cup & & \cup \\
(w_I) & \longmapsto & (\log(w_I)) =: (x_I) \\
\\
\prod_{I \subseteq [n]} \max(w_I, w_I^{-1}) & \longmapsto & \sum_{I \subseteq [n]} \max(x_I, -x_I) =: \|(x_I)\|_{\mathrm{trop}}
\end{array}
$$

In tropical geometry, the function $\|\cdot\|_{\mathrm{trop}}$ is called a tropical rational function. It is a piecewise-linear function, linear on the image of each $\mathcal{O}_{\mathcal{I}}$, and the boundary separating its regions of linearity is called its tropical hypersurface T. Since the gradient of $\|\cdot\|_{\mathrm{trop}}$ is constant on each region of linearity, optimizing $\|\cdot\|_{\mathrm{trop}}$ on any set L boils down to understanding the intersection of L with T.

For our problem, L is the image of the feasible set given by the conditions $\Phi(W) = P$ which is an exponential variety of codimension n. Theoretically, there exist only a finite number of intersection patterns of L and T. Classifying these intersection patterns would allow us to compute minimal interactions and minimal molecules almost instantly.

Tensors. Tensors [Lan12] are multidimensional arrays of real numbers which generalize matrices. We can regard molecules as (positive) tensors in $\mathbb{R}^{2 \times \cdots \times 2} =: (\mathbb{R}^2)^{\otimes n}$, whose entries are indexed by $\{0, 1\}^n$ which is in bijection with the subsets of $[n]$.

From this point of view, many concepts in tensor theory have natural analogues in ligand binding, such as symmetric tensors and molecules with indistinguishable sites. Notably, a rank one tensors is up to scaling a tensor with entries

$$
(1, 0, \dots) \colon s_1, \quad (0, 1, 0, \dots) \colon s_2, \quad (1, 1, 0, \dots) \colon s_1 s_2, \quad \dots \quad \text{for some } s_i \in \mathbb{R}.
$$

Hence rank one tensors correspond to molecules with independent sites. Therefore, decomposing a given tensor into rank one tensors translates to decomposing a given molecule into different molecules with independent sites whose mix has the same binding curve. And the tensor rank translates to the minimal number of different molecules with independent sites needed to recreate the same binding curve. Thus the tensor rank of a molecule could serve as an interesting, discrete measure for cooperativity.

Dependency measures. From a stochastic or statistical point of view, the problem of cooperativity is a problem of quantifying (minimal) dependency, which is a classical research topic in these fields [Rén59; SW81; Koy+87]. In particular copulas [Sch91; DS10; Nel07] have become a common tool to describe dependencies between random variables. Copulas cannot be applied directly to our problem since they are defined on continuous, not discrete, random variables. Moreover, they aim at modeling the dependency of given random variables which in our scenario corresponds to the dependency of the sites of a known molecule. A copula then relates the one-dimensional marginal distributions to the multivariate minimization problem of finding the minimal dependency required to generate a given distribution of a sum of dependent variables is usually not treated. Nevertheless, the problem of cooperativity should also be treated from a more probabilistic point of view and the available concepts should be considered in more detail in the context of ligand binding.

References

[Abe05]   H. Abeliovich. "An empirical extremum principle for the hill coefficient in ligand-protein interactions showing negative cooperativity". In: *Biophysical Journal* 89.1 (2005), pp. 76–79.

[Abe16]   H. Abeliovich. "On Hill coefficients and subunit interaction energies". In: *Journal of Mathematical Biology* 73.6-7 (2016), pp. 1399–1411.

[ABJ25]   G. S. Adair, A. V. Bock, and H. F. Jr. "The hemoglobin system VI. The oxygen dissociation curve of hemoglobin". In: *Journal of Biological Chemistry* 63.2 (1925), pp. 529–545.

[Bar13]   J. Barcroft. "The combinations of haemoglobin with oxygen and with carbon monoxide. II". In: *Biochemical Journal* 7.5 (1913), p. 481.

[Ben13]   A. Y. Ben-Naim. *Cooperativity and regulation in biochemical processes.* Springer Science & Business Media, 2013.

[BHK04]   C. Bohr, K. Hasselbalch, and A. Krogh. "Ueber einen in biologischer Beziehung wichtigen Einfluss, den die Kohlensäurespannung des Blutes auf dessen Sauerstoffbindung übt 1". In: *Skandinavisches Archiv für Physiologie* 16.2 (1904), pp. 402–412.

[CRBG86]   P. Connelly, C. Robert, W. Briggs, and S. Gill. "Analysis of zeros of binding polynomials for tetrameric hemoglobins". In: *Biophysical Chemistry* 24.3 (1986), pp. 295–309.

[DS10]   F. Durante and C. Sempi. "Copula theory: an introduction". In: *Copula Theory and its Applications*. Springer, 2010, pp. 3–31.

[Gle+18]   A. Gleixner et al. *The SCIP Optimization Suite 6.0*. Technical Report. Optimization Online, July 2018. URL: http://www.optimization-online.org/DB_HTML/2018/07/6692.html.

[GS14]   A. Greuet and M. Safey El Din. "Probabilistic algorithm for polynomial optimization over a real algebraic set". In: *SIAM Journal on Optimization* 24.3 (2014), pp. 1313–1343.

[GMD09]   P. Gutierrez, D. Monteoliva, and L. Diambra. "Role of cooperative binding on noise expression". In: *Physical Review E* 80.1 (2009), p. 011914.

[Hil13]   A. V. Hill. "The combinations of hæmoglobin with oxygen and with carbon monoxide. I". In: *Biochemical Journal* 7.5 (1913), p. 471.

[Hil85]   T. L. Hill. *Cooperativity theory in biochemistry: steady-state and equilibrium systems*. Springer Series in Molecular and Cell Biology. Springer-Verlag New York, 1985.

[HA09]   C. A. Hunter and H. L. Anderson. "What is cooperativity?" In: *Angewandte Chemie International Edition* 48.41 (2009), pp. 7488–7499.

[Ike+83]   M. Ikeda-Saito et al. "Thermodynamic properties of oxygen equilibria of dimeric and tetrameric hemoglobins from Scapharca inaequivalvis". In: *Journal of Molecular Biology* 170.4 (1983), pp. 1009–1018.

[Ima73]   K. Imai. "Analyses of oxygen equilibriums of native and chemically modified human adult hemoglobins on the basis of Adir's stepwise oxygenation theory and the allosteric model of Monod, Wyman, and Changeux". In: *Biochemistry* 12.5 (1973), pp. 798–808.

[Koy+87]   R. A. Koyak et al. "On measuring internal dependence in a set of random variables". In: *The Annals of Statistics* 15.3 (1987), pp. 1215–1228.

[Lan12]   J. M. Landsberg. *Tensors: geometry and applications*. Vol. 128. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012, pp. xx+439.

[Las01]   J. Lasserre. "Global optimization with polynomials and the problem of moments". In: *SIAM Journal on Optimization* 11.3 (2001), pp. 796–817.

[LFSR09]   T. Lenaerts, J. Ferkinghoff-Borg, J. Schymkowitz, and F. Rousseau. "Information theoretical quantification of cooperativity in signalling complexes". In: *BMC Systems Biology* 3.1 (2009), p. 9.

[MS15]   D. Maclagan and B. Sturmfels. *Introduction to tropical geometry*. Vol. 161. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2015.

[Mar17]   J. W. R. Martini. "A measure to quantify the degree of cooperativity in overall titration curves". In: *Journal of Theoretical Biology* 432 (2017), pp. 33–37.

[MDH16]   J. W. R. Martini, L. Diambra, and M. Habeck. "Cooperative binding: a multiple personality". In: *Journal of Mathematical Biology* 72.7 (2016), pp. 1747–1774.

[MSS16]   J. W. R. Martini, M. Schlather, and S. Schütz. "A model for carrier-mediated biological signal transduction based on equilibrium ligand binding theory". In: *Bulletin of Mathematical Biology* 78.5 (2016), pp. 1039–1057.

[MSU13]   J. W. R. Martini, M. Schlather, and G. M. Ullmann. "On the interaction of different types of ligands binding to the same molecule Part II: systems with n to 2 and n to 3 binding sites". In: *Journal of Mathematical Chemistry* 51.2 (2013), pp. 696–714.

[MU13]   J. W. R. Martini and G. M. Ullmann. "A mathematical view on the decoupled sites representation". In: *Journal of Mathematical Biology* 66.3 (2013), pp. 477–503.

[MMD13]   D. Monteoliva, C. B. McCarthy, and L. Diambra. "Noise minimisation in gene expression switches". In: *PloS One* 8.12 (2013), e84020.

[Nel07]   R. B. Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.

[OCU01]   A. Onufriev, D. A. Case, and G. M. Ullmann. "A novel view of pH titration in biomolecules". In: *Biochemistry* 40.12 (2001), pp. 3413–3419.

[OU04]   A. Onufriev and G. M. Ullmann. "Decomposing complex cooperative ligand binding into simple components: connections between microscopic and macroscopic models". In: *The Journal of Physical Chemistry B* 108.30 (2004), pp. 11157–11169.

[Par03]   P. A. Parrilo. "Semidefinite programming relaxations for semialgebraic problems". In: *Mathematical Programming* 96.2 (2003), pp. 293–320.

[RMT19]   Y. Ren, J. W. R. Martini, and J. Torres. "Decoupled molecules with binding polynomials of bidegree (n,2)". In: *Journal of Mathematical Biology* 78.4 (Mar. 2019), pp. 879–898.

[Rén59]   A. Rényi. "On measures of dependence". In: *Acta Mathematica Hungarica* 10.3-4 (1959), pp. 441–451.

[RS18]   C. Riener and M. Safey El Din. "Real root finding for equivariant semialgebraic systems". In: *ISSAC'18—Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*. ACM, New York, 2018, pp. 335–342.

[Ron+19] C. Rong et al. "Homogeneous molecular systems are positively cooperative, but charged molecular systems are negatively cooperative". In: *The Journal of Physical Chemistry Letters* 10.8 (2019), pp. 1716–1721.

[RRS17]   R. D. Roy, C. Rosenmund, and M. I. Stefan. "Cooperative binding mitigates the high-dose hook effect". In: *BMC Systems Biology* 11.1 (2017), p. 74.

[Sal+16] D. V. Salakhieva et al. "Kinetic regulation of multi-ligand binding proteins". In: *BMC Systems Biology* 10.1 (2016), p. 32.

[Sch75]   J. A. Schellman. "Macromolecular binding". In: *Biopolymers* 14.5 (1975), pp. 999–1018.

[Sch91]   B. Schweizer. "Thirty years of copulas". In: *Advances in Probability Distributions with Given Marginals*. Springer, 1991, pp. 13–50.

[SW81]   B. Schweizer and E. F. Wolff. "On nonparametric measures of dependence for random variables". In: *The Annals of Statistics* 9.4 (1981), pp. 879–885.

[SL13]   M. I. Stefan and N. Le Novère. "Cooperative binding". In: *PLoS Computational Biology* 9.6 (2013), e1003106.

[Tib96]   R. Tibshirani. "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (1996), pp. 267–288.

[WG90]   J. Wyman and S. J. Gill. *Binding and linkage: functional chemistry of biological macromolecules*. University Science Books, 1990.

Appendix A. List of Constraints

Below is a compact representation of the list of 180 constraints we considered to compute the lower bounds b_- of minimal absolute interaction in Figure 3. There, $\{12, 34, 123\}$ represents the constraint $f_{\mathcal{I}} \leq r$ for $\mathcal{I} = \{12, 34, 123\}$, which is the absolute interaction $\|W\|$ for $W \in \mathcal{O}_{\mathcal{I}}$ where $w_{12}, w_{34}, w_{123} \leq 1$ and $w_I \geq 1$ otherwise.

$\emptyset$, {12}, {123}, {1234}, {12,13}, {12,34}, {12,123}, {12,134}, {123,124}, {12,1234}, {123,1234}, {12,13,14}, {12,13,23}, {12,13,24}, {12,13,123}, {12,13,124},{12,34,123}, {12,13,234}, {12,123,124}, {12,123,134}, {12,134,234}, {123,124,134}, {12,13,1234}, {12,34,1234}, {12,123,1234}, {12,134,1234}, {123,124,1234}, {12,13,14,23}, {12,13,24,34}, {12,13,14,123}, {12,13,23,123}, {12,13,24,123}, {12,13,23,124}, {12,13,24,134}, {12,13,14,234}, {12,13,123,124}, {12,13,124,134}, {12,13,123,234}, {12,34,123,134}, {12,34,123,124}, {12,13,124,234}, {12,123,124,134}, {12,123,134,234}, {123,124,134,234}, {12,13,14,1234}, {12,13,23,1234}, {12,13,24,1234}, {12,13,123,1234}, {12,13,124,1234}, {12,34,123,1234}, {12,13,234,1234}, {12,123,124,1234},{12,123,134,1234}, {12,134,234,1234}, {123,124,134,1234}, {12,13,14,23,24}, {12,13,14,23,123}, {12,13,14,23,124}, {12,13,24,34,123}, {12,13,14,23,234}, {12,13,14,123,124}, {12,13,23,123,124}, {12,13,24,123,124}, {12,13,24,123,234}, {12,13,24,123,134}, {12,13,23,124,134}, {12,13,14,123,234}, {12,13,24,134,234}, {12,13,123,124,134}, {12,13,123,124,234}, {12,34,123,124,134}, {12,13,124,134,234}, {12,123,124,134,234}, {12,13,14,23,1234}, {12,13,24,34,1234}, {12,13,14,123,1234}, {12,13,23,123,1234}, {12,13,24,123,1234}, {12,13,23,124,1234}, {12,13,24,134,1234}, {12,13,14,234,1234}, {12,13,123,124,1234}, {12,13,124,134,1234}, {12,13,123,234,1234}, {12,34,123,134,1234}, {12,34,123,124,1234}, {12,13,124,234,1234}, {12,123,124,134,1234}, {12,123,134,234,1234}, {123,124,134,234,1234}, {12,13,14,23,24,34}, {12,13,14,23,24,123}, {12,13,14,23,24,134}, {12,13,14,23,123,124}, {12,13,24,34,123,234}, {12,13,14,23,124,134}, {12,13,14,23,123,234}, {12,13,24,34,123,124}, {12,13,14,23,124,234}, {12,13,14,123,124,134}, {12,13,23,123,124,134}, {12,13,24,123,124,134}, {12,13,14,123,124,234}, {12,13,24,123,134,234}, {12,13,23,124,134,234}, {12,13,123,124,134,234}, {12,34,123,124,134,234}, {12,13,14,23,24,1234}, {12,13,14,23,123,1234}, {12,13,14,23,124,1234}, {12,13,24,34,123,1234}, {12,13,14,23,234,1234}, {12,13,14,123,124,1234}, {12,13,23,123,124,1234}, {12,13,24,123,124,1234}, {12,13,24,123,234,1234}, {12,13,24,123,134,1234}, {12,13,23,124,134,1234}, {12,13,14,123,234,1234}, {12,13,24,134,234,1234}, {12,13,123,124,134,1234}, {12,13,123,124,234,1234}, {12,34,123,124,134,234}, {12,13,124,134,234,1234}, {12,123,124,134,234,1234}, {12,13,14,23,24,34,123}, {12,13,14,23,24,123,124}, {12,13,14,23,24,

123,134}, {12,13,14,23,24,134,234}, {12,13,14,23,123,124,134}, {12,13,14,23,123,124,234}, {12,13,24,34,123, 124,134}, {12,13,14,23,124,134,234}, {12,13,14,123,124,134,234}, {12,13,23,123,124,134,234}, {12,13,24,123, 124,134,234}, {12,13,14,23,24,34,1234}, {12,13,14,23,24,123,1234}, {12,13,14,23,24,134,1234}, {12,13,14,23, 123,124,1234}, {12,13,24,34,123,234,1234}, {12,13,14,23,124,134,1234}, {12,13,14,23,123,234,1234}, {12,13, 24,34,123,124,1234}, {12,13,14,23,124,234,1234}, {12,13,14,123,124,134,1234}, {12,13,23,123,124,134,1234}, {12,13,24,123,124,134,1234}, {12,13,14,123,124,234,1234}, {12,13,24,123,134,234,1234}, {12,13,23,124,134, 234,1234}, {12,13,123,124,134,234,1234}, {12,34,123,124,134,234,1234}, {12,13,14,23,24,34,123,124}, {12, 13,14,23,24,123,124,134}, {12,13,14,23,24,123,134,234}, {12,13,14,23,123,124,134,234}, {12,13,24,34,123, 124,134,234}, {12,13,14,23,24,34,123,1234}, {12,13,14,23,24,123,124,1234}, {12,13,14,23,24,123,134,1234}, {12,13,14,23,24,134,234,1234}, {12,13,14,23,123,124,134,1234}, {12,13,14,23,123,124,234,1234}, {12,13,24, 34,123,124,134,1234}, {12,13,14,23,124,134,234,1234}, {12,13,14,123,124,134,234,1234}, {12,13,23,123,124, 134,234,1234}, {12,13,24,123,124,134,234,1234}, {12,13,14,23,24,34,123,124,134}, {12,13,14,23,24,123,124, 134,234}, {12,13,14,23,24,34,123,124,1234}, {12,13,14,23,24,123,124,134,1234}, {12,13,14,23,24,123,134,234, 1234}, {12,13,14,23,123,124,134,234,1234}, {12,13,24,34,123,124,134,234,1234}, {12,13,14,23,24,34,123,124, 134,234}, {12,13,14,23,24,34,123,124,134,1234}, {12,13,14,23,24,123,124,134,234,1234}, {12,13,14,23,24,34, 123,124,134,234,1234}

(Nidhi Kaihnsa) Division of Applied Mathematics, Brown University, 182 George Street, Providence, RI, 02912

Email address: nidhi_kaihnsa@brown.edu

URL: https://sites.google.com/view/kaihnsa/

(Yue Ren, corresponding author) Swansea University, Department of Mathematics, Swansea University Bay Campus, SA1 8JN Swansea, United Kingdom

Email address: yue.ren@swansea.ac.uk

URL: https://yueren.de

(Mohab Safey El Din) Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, Équipe PolSys

Email address: mohab.safey@lip6.fr

URL: https://www-polsys.lip6.fr/~safey/

(Johannes W. R. Martini) International Maize and Wheat Improvement Center, Texcoco, Mexico.

Email address: jwrmartini@gmail.com