



## Cooperative binding: a multiple personality

Johannes W. R. Martini<sup>1,2,3</sup> · Luis Diambra<sup>4</sup> ·  
Michael Habeck<sup>2,5</sup>

Received: 30 September 2014 / Revised: 30 June 2015 / Published online: 29 August 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Cooperative binding has been described in many publications and has been related to or defined by several different properties of the binding behavior of the ligand to the target molecule. In addition to the commonly used Hill coefficient, other characteristics such as a sigmoidal shape of the overall titration curve in a linear plot, a change of ligand affinity of the other binding sites when a site of the target molecule becomes occupied, or complex roots of the binding polynomial have been used to define or to quantify cooperative binding. In this work, we analyze how the different properties are related in the most general model for binding curves based on the grand canonical partition function and present several examples which highlight differences between the cooperativity characterizing properties which are discussed. Our results mainly show that among the presented definitions there are not two which fully coincide. Moreover, this work poses the question whether it can make sense to distinguish between positive and negative cooperativity based on the macroscopic

---

✉ Johannes W. R. Martini  
jmartin2@gwdg.de

Luis Diambra  
ldiambra@gmail.com

Michael Habeck  
mhabeck@gwdg.de

- 1 Max Planck Institute for Developmental Biology, Tübingen, Germany
- 2 Felix Bernstein Institute for Mathematical Statistics in the Biosciences, Georg-August University Göttingen, Göttingen, Germany
- 3 Animal Breeding and Genetics Group, Department of Animal Sciences, Georg-August University Göttingen, Göttingen, Germany
- 4 Centro Regional de Estudios Geñomicos, Universidad Nacional de La Plata, La Plata, Argentina
- 5 Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

binding isotherm only. This article shall emphasize that scientists who investigate cooperative effects in biological systems could help avoiding misunderstandings by stating clearly which kind of cooperativity they discuss.

**Mathematics Subject Classification** 82B20 · 82B05 · 92C40 · 92C05 · 92E99

## 1 Introduction

The topic of cooperative binding has been addressed in many publications (e.g. [Ben-Naim 2001](#); [Gutierrez et al. 2012](#); [Hunter and Anderson 2009](#); [Ullmann and Ullmann 2011](#)) and different underlying equations, measures and definitions of “cooperativity” have been used since the discovery of cooperative binding of oxygen to hemoglobin ([Hill 1910, 1913](#)). Among the different equations which have been used to describe binding curves are the Hill equation, the Adair equation, the Klotz equation (for a review of the history of these equations see e.g. [Stefan and Le Novère 2013](#)) and several “measures” or “indicators” have been developed, sometimes based on the validity of the respective model. Even though already [Hill \(1985\)](#) provided a rigorous treatment of cooperative binding based on statistical mechanics nearly thirty years ago (which included results of even earlier work) and which in particular highlighted that different concepts of cooperativity do not necessarily coincide, these consistency problems are still not fully recognized by the scientific community investigating cooperative phenomena. For instance, [Hill \(1985\)](#) clearly highlighted that the significance of the A.V. Hill coefficient (in the form of how it was initially generalized from the Hill to the Adair equation) as an indicator of cooperativity depends on a certain kind of symmetry between the binding sites and should be substituted by another quantity if an asymmetric system is considered. In this work, we will follow Hill’s treatment of binding curves in the setup of the grand canonical ensemble, analyze under which circumstances different measures of cooperativity coincide and highlight difficulties which will arise if different measures are used. In particular, we present several hypothetical binding systems to illustrate differences between the concepts of cooperativity. We start with a recapitulation of the grand partition function in the next section which is followed by a section on definitions of cooperative binding related to the ligand activity (“effective concentration”) dependent behavior on the macrostates (the number of bound ligands) and a section on definitions of cooperative binding on the microstates (including the information which sites are occupied). Note that this work deals with cooperative ligand binding. The term “cooperativity” will not be discussed in other contexts.

## 2 The grand canonical ensemble

The grand canonical ensemble defines a family of parameterized probability distributions on the possible states of a small subsystem which is embedded in a much larger system with which it can exchange energy and particles. The chemical activity (“effective concentration”) of the particle in the whole system and the temperature of the whole system are the parameters. If we regard a molecule with  $n$  ligand binding

sites in solution as such a subsystem we obtain the following probability for a binding state  $k = (k_1, \dots, k_n) \in \{0, 1\}^n$ , where  $k_i = 0$  means that site  $i$  is unoccupied and  $k_i = 1$  that site  $i$  is occupied (Schellman 1975):

$$P(\{k\}) = \frac{g(k)\lambda^{|k|}}{\sum_k g(k)\lambda^{|k|}} \quad (1)$$

with

$$g(k) := \exp(-G(k)/RT) \quad (2)$$

the so-called Boltzmann factor,  $G(k)$  the free energy of the molecule in state  $k$  compared to a fixed, arbitrarily chosen reference state,  $R$  the Boltzmann constant,  $T$  the absolute temperature in °K,  $\lambda \geq 0$  the thermodynamic activity of the ligand in the environment and  $|k|$  the number of occupied sites in microstate  $k$  (note that the domain  $\lambda \geq 0$  will underly all statements in this work). The overall titration curve is given by the expectation of the number of bound ligands as a function of the activity of the ligand in the solution at fixed temperature (also called “isotherm”):

$$\Psi(\lambda) = \frac{n \cdot a_n \lambda^n + (n-1) \cdot a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda}{a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + 1}, \quad (3)$$

where the denominator is the grand partition function, which is a polynomial  $\Phi(\lambda) = a_n \lambda^n + \dots + 1$  of degree  $n$  in the case of a target molecule with  $n$  binding sites. Here, the coefficients  $a_i$  are the sum of all Boltzmann factors  $g(k)$  (Eq. 2) which belong to the respective macrostate  $i$ . In the context of ligand binding, Eq. (3) is called the Adair equation (Adair et al. 1925). The overall titration curve  $\Psi$  is what can be measured easily in experiments and we would like to deduce some characteristics of the family of measures described in Eq. (1) from  $\Psi$ . An example for a characteristic could be a change in “affinity” of the ligand to a certain site if another site changes its binding state. Note that even if Eq. (3) describes at first sight only the expectation of the distribution of  $|k|$  (macrostates) as a function of  $\lambda$ , it defines the whole family of distributions on the macrostates uniquely (see Martini 2014, p. 26). However, what is not uniquely determined is the original distribution on the microstates: an infinite number of families of distributions on the microstates produce the same overall binding curve.

In the following, we will introduce and discuss several quantities related to “cooperativity”, starting from definitions by properties of the distributions on the macrostates.

### 3 Cooperative binding as a property of the family of distributions on the macrostates

#### 3.1 The A. V. Hill equation, the Hill coefficient and its generalizations

##### 3.1.1 The A. V. Hill equation

The A. V. Hill equation was developed as a *phenomenological* description of the overall binding of oxygen to hemoglobin (Hill 1910, 1913) and it is not based on any

mechanistic principle of ligand binding. It is given by

$$\Psi(\lambda) = n \frac{K\lambda^\alpha}{1 + K\lambda^\alpha}. \quad (4)$$

with  $n$  denoting the maximal number of bound ligands,  $K$  the positive association constant and  $\alpha$  an appropriate positive constant. Even though Eq. (4) is “only” a phenomenological description and consequently not necessarily covered by Eq. (3), we will investigate under which circumstances Eq. (4) fits to the setup of Eq. (3) to generate an intuition of how to interpret  $\alpha$ . Firstly, we see that, e.g. for the case of  $\alpha = n = 2$  and  $K > 0$ , Eq. (4) fits the shape of Eq. (3) only if  $a_1 = 0$ . This is not possible, since the coefficients of the polynomial have to be positive, since they are sums of exponential terms. However, Eq. (4) can be interpreted in this example as a limit case with  $a_1$  much smaller than 1. To include these limit cases, we will from now on also use binding polynomials with coefficients equal to zero. Including these polynomials, the question arises whether there are other options than  $n = \alpha \in \mathbb{N}$  denoting the number of binding sites to obtain an equation of shape of Eq. (4) from Eq. (3). An answer is given by Proposition 1 and Example 1 (see the proof section for the derivation of the results of Proposition 1).

**Proposition 1** *Let Eq. (4) with  $K > 0$  be an overall binding curve of a molecule with  $n$  binding sites and  $a_n \neq 0$  (all sites can be occupied simultaneously). Then*

- (a)  $\alpha$  is a positive integer
- (b)  $\alpha$  divides  $n$
- (c) The coefficients  $a_{i\alpha}$  are nonzero for all  $i \in \{1, \dots, \frac{n}{\alpha}\}$ .

Note that Proposition 1 does not state that the denominator of Eq. (4) is necessarily the binding polynomial, which is true if  $\alpha = n$ . Indeed, binding polynomials can be constructed whose corresponding overall titration curve satisfies Eq. (4) and  $\alpha \neq n$ , which is illustrated by Example 1.

*Example 1* An example of a binding polynomial whose corresponding overall titration curve is a Hill equation with a denominator unequal to the binding polynomial is

$$\Phi(\lambda) = K^2\lambda^4 + 2K\lambda^2 + 1. \quad (5)$$

Its overall titration curve is given by

$$\Psi(\lambda) = \frac{4K^2\lambda^4 + 4K\lambda^2}{K^2\lambda^4 + 2K\lambda^2 + 1} = \frac{4K\lambda^2}{K\lambda^2 + 1} \quad (6)$$

Proposition 1 showed that within the framework of Eq. (3), the parameter  $\alpha$  can be regarded as a measure which tells us how many ligand particles bind simultaneously as a “package” to the target molecule.

### 3.1.2 The Hill coefficient

In Eq. (4), the parameter  $\alpha$  is called the Hill coefficient, but it is also a useful tool (based on the more general Definition 1) if the binding curve of interest is not of shape (4). As illustrated in Proposition 1, the parameter  $\alpha$  can be regarded as a measure which tells us how many ligand particles have to bind as a “package” to the target molecule. Even though, in the setup of the grand canonical ensemble,  $\alpha$  has to be a natural number (positive integer), we can generalize the domain of  $\alpha$  with different concepts. One possibility is to fit measured data to Eq. (4) with  $\alpha \in \mathbb{R}^+$ , thus “projecting” a curve of shape of Eq. (3) on the class of functions described by Eq. (4) and interpret  $\alpha$  in an analogous way, even if  $\alpha$  is not a natural number. A value of e.g. 1.5 for  $\alpha$  could be interpreted as packages of one and a half ligands which bind simultaneously. Consequently, a value of  $\alpha > 1$  in the best fitting function of shape of Eq. (4) can be interpreted as an indicator for cooperative binding. A second possibility to transfer the Hill coefficient  $\alpha$  to curves of shape of Eq. (3) is to generalize the definition such that it coincides with  $\alpha$  if we consider a curve of shape Eq. (4). This is done by Definition 1 which gives a more general definition of the Hill coefficient and which can be motivated by following idea: We start with the concept that (“positive”) cooperativity between different sites means that the binding of the ligand to one site increases the “affinity” of the ligand to the other sites. This view leads to the intuition that cooperativity should lead to a steep slope of the titration curve, since the binding enhances the binding to the other sites (similar to a positive feedback loop). Having the example of hemoglobin in mind, this means that only a “small” change in oxygen activity is required to transit from a distribution with mainly completely empty molecules to a distribution with mainly fully occupied molecules. Consequently, “positive” cooperativity should be related to the steepness of the slope of an overall titration curve.

**Definition 1** The Hill coefficient  $\eta$  of an overall titration curve  $\Psi$  is defined as the slope of

$$H_\Psi(\log(\lambda)) := \log\left(\frac{\Psi(\lambda)}{n - \Psi(\lambda)}\right) \quad (7)$$

as a function of the natural logarithm of the ligand activity  $\log(\lambda)$ , at the activity at which  $\Psi = \frac{n}{2}$  (Hill 1985, 1910, 1913).

Note that the Hill coefficient of Definition 1, coincides with  $\alpha$  of Eq. (4), if this equation describes the overall titration curve (see the proof section). Moreover note that the transformation of the overall titration curve is strictly monotonous, i.e.

$$\Psi(\lambda_0) > \Psi(\lambda_1) \Leftrightarrow H_\Psi(\log \lambda_0) > H_\Psi(\log \lambda_1),$$

which means that we can express the information that we find in the Hill plots  $H_\Psi$  also directly through  $\Psi$ .

We have summarized up to now that the Hill coefficient is motivated by the phenomenological Eq. (4) and that the usual way to generalize it to curves of shape of Eq. (3) is the use of Definition 1 which coincides with  $\alpha$  of Eq. (4) if the function is of this shape. Since the Hill coefficient is usually applied to curves of shape of Eq. (3),

the question arises which characteristics of the family of measures defined by Eq. (1) are really measured by the Hill coefficient. The answer to this question highlights a certain phenomenon of the grand partition function: Within this setup, characteristics of certain curves as a function of  $\lambda$  are related to characteristics of certain probability measures on  $\{0, \dots, n\}$  when  $\lambda$  is fixed. In particular, the Hill coefficient describes the variance of the distribution on the macrostates when  $\Psi = \frac{n}{2}$ .

**Proposition 2** *The Hill coefficient  $\nu$  of Definition 1 satisfies*

$$\eta = \frac{4}{n} \mathbb{V}(\lambda)_{|\Psi = \frac{n}{2}} \quad (8)$$

with  $\mathbb{V}(\lambda)$  denoting the variance of the distribution on the set of macrostates as a function of  $\lambda$  and  $\mathbb{V}(\lambda)_{|\Psi = \frac{n}{2}}$  its value at  $\Psi = \frac{n}{2}$ .

Proposition 2 states that the Hill coefficient, which measures the steepness of the slope of the expectation as a function of  $\log(\lambda)$  at  $\Psi = \frac{n}{2}$ , also gives the variance of the distribution on the macrostates at  $\Psi = \frac{n}{2}$ . This fact fits well to the idea of (“positive”) cooperativity increasing the “affinity” of the ligand to the other sites when a certain site is occupied, since this circumstance will lead to a distribution on the macrostates with more weight towards the extreme occupational states 0 and  $n$  than if the ligands bind independently and thus will increase the variance.

However, the question arises why the point  $\Psi = \frac{n}{2}$  is regarded. The answer is given by the fact that under a certain symmetry condition, the slope and consequently the variance of the distribution on the macrostates has a local extremum at this point., and that actually *abnormally high values* of variances of the distributions on the macrostates are quantities of interest. Before presenting the symmetry condition according to Hill (1985), we shortly describe why high values of variances of the distributions on the macrostates are of interest at all: High values of the variance of the distribution on the macrostates are important characteristics of the macroscopic system since the variances of independent systems define bounds. If a bound is exceeded, the observation can not result from an independent system. An “abnormally high” variance is an indicator for relevant interaction between the sites. To understand the threshold which is recognized by the Hill coefficient, let us regard an independent system. If the individual binding sites do not interact (in terms of interaction energy zero), the sites will bind the ligand stochastic independently for every value of  $\lambda$  (Martini et al. 2013b). Thus, the variance of the sum (the macrostates) equals the sum of the variances of the Bernoulli variables  $X_i$  describing whether site  $i$  is occupied or not. The variance of a Bernoulli variable is bounded by 0.25, which is reached if  $P(X_i = 1) = P(X_i = 0) = 0.5$ . Thus, a variance larger than  $0.25n$  (which at  $\Psi = \frac{n}{2}$  equals a Hill-coefficient of 1) indicates that the observed behavior cannot result from a system of stochastic independent binding sites [We will show later that this reference point of half-saturation which is used in Definition 1 and which can often be found in literature (e.g. Hill 1985; Onufriev and Ullmann 2004; Ge and Qian 2009; Hunter and Anderson 2009) is not in general meaningful].

Hill (1985) gives the following symmetry condition: a  $\lambda_0$  exists such that

$$\Psi(a\lambda_0) + \Psi(\lambda_0/a) = n \quad \forall a \neq 0 \quad (9)$$

and states that the “definition of the Hill coefficient [. . .] can be used for all systems with the above symmetry” (Hill 1985, p. 83). Since we did not find a proof for this statement in literature, we will prove Proposition 3 (proofs can be found in Sect. 6).

In the following, we will use the notation

$$f' := \frac{df}{d \log \lambda}$$

for the derivative of a function with respect to the natural logarithm of the activity of the ligand. Moreover, note that

$$\Psi' = \nabla(\lambda), \quad (10)$$

i.e. the derivative of the titration curve with respect to  $\log \lambda$  is the variance of the distribution on the macrostates  $|k| \in \{0, \dots, n\}$  (calculate the derivative or see e.g. Hill 1985).

**Proposition 3** *Let  $\Psi$  be an overall titration curve satisfying Eq. (9). Then  $\Psi'$  and thus the variance of the distribution on the macrostates  $|k|$  has a local extremum at  $\lambda_0$ .*

Hill (1985) also rewrote the condition of Eq. (9) to the equivalent statement that the coefficients of the binding polynomial  $\Phi(\lambda) = a_n \lambda^n + \dots + a_1 \lambda + 1$  fulfill

$$a_r \lambda_0^r = a_{n-r} \lambda_0^{n-r} \quad (11)$$

which means for the distribution at  $\lambda_0$  that

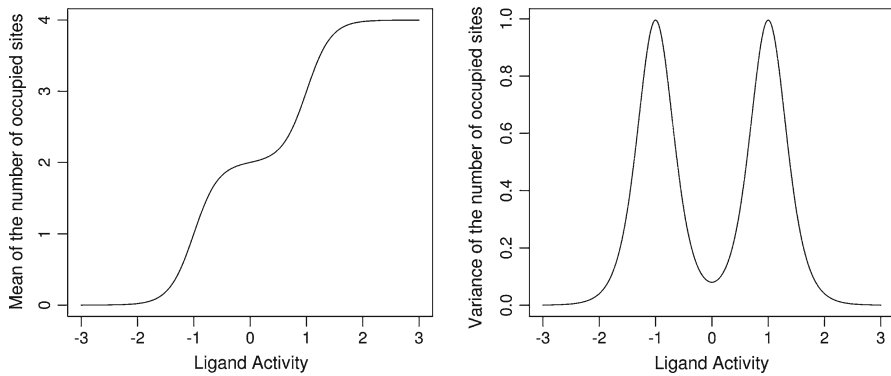
$$P_{\lambda_0}(|k| = r) = P_{\lambda_0}(|k| = n - r) \quad \forall r \in \{0, \dots, n\}.$$

In particular, this equation illustrates that if a  $\lambda_0$  exists which satisfies Eq. (9) [or equivalently Eq. (11)], this implies that  $\lambda_0$  also satisfies  $\Psi(\lambda_0) = \frac{n}{2}$ . Note that this kind of symmetry does not necessarily mean that the binding sites are identical, i.e. that they have the same energy levels. We will illustrate this point in Example 2.

*Example 2 (Homooligomers and symmetry)* Let  $M$  be a molecule with three binding sites. Moreover, let  $g_i$  be the Boltzmann factor of a microstate of macrostate  $i$ . Then the binding polynomial is given by

$$\Phi(\lambda) = g_3 \lambda^3 + 3g_2 \lambda^2 + 3g_1 \lambda + 1.$$

This special choice of the coefficients means that all the three sites are equal in terms of the free energy of binding or releasing a ligand from the site: the Boltzmann factor of a microstate only depends on the corresponding macrostate (this could be a valid assumption for homooligomers). In spite of the binding polynomial belonging to a symmetric system (in terms of being composed of energetically identical sites) it does not necessarily fulfill the symmetry condition of Eq. (9), which is illustrated by  $g_3 = 4, g_2 = 2, g_1 = 1$  and  $\lambda_0 = 0.603$  as the activity of half-saturation. Moreover, even any arbitrarily chosen binding polynomial with coefficients  $a_i$  can result from a



**Fig. 1** *Left side* the mean of the number of occupied sites as a function of the ligand activity (the “overall titration curve” or “isotherm”, a decadic logarithmic scale is used) of the polynomial of Example 3. *Right side* the variance as a function of the ligand activity (at constant temperature) of the system described by the polynomial of Example 3

molecule with identical sites with  $g_i = a_i \binom{n}{i}^{-1}$ , which illustrates that the discussed symmetry condition is not related to a chemical symmetry of the target molecule. Conversely, a binding polynomial fulfilling Eq. (9) does not necessarily come from a molecule in which all sites are identical. Eq. (9) describes a property of the distributions on the macrostates and not on the microstates. The symmetry condition (9) is not related to a chemical symmetry of the molecule. However, note here that the relation  $g_i = a_i \binom{n}{i}^{-1}$  allows us to deduce the free energies of the microstates from the binding polynomial (if we know that we are dealing with a chemically symmetric system).

We have seen that the Hill coefficient measures the variance of a certain distribution and that the variance is extreme at the ligand activity of half-saturation if condition Eq. (9) is satisfied. An important observation is that even if Eq. (9) is fulfilled, the Hill coefficient is only a locally extreme variance, but not necessarily a locally maximal variance which is illustrated by Example 3.

*Example 3* Let us regard the binding polynomial  $\Phi(\lambda) = \lambda^4 + 0.1\lambda^3 + 100\lambda^2 + 0.1\lambda + 1$  which satisfies Eq. (9). Half-saturation is reached at  $\lambda_0 = 1$ , where the corresponding variance reaches a local minimum of 0.08. The overall titration curve (isotherm) as well as the variance are illustrated in Fig. 1.

We have demonstrated up to now that—from a theoretical point of view—the use of the reference point  $\Psi = \frac{n}{2}$  is not fully convincing. Firstly, Example 3 showed that even if Eq. (9) is fulfilled by the system, we can measure a locally minimal variance. Secondly, if Eq. (9) is not fulfilled, the reference point  $\Psi = \frac{n}{2}$  does not in general give any special information. In particular, the variance at this point being smaller than  $0.25n$  does not necessarily mean that there is no other point at which the variance is “abnormally high”. This observations can motivate the following generalizations of the Hill coefficient.



### 3.1.3 An extension of the Hill coefficient definition: the maximal variance

An idea for the extension of the Hill coefficient to other systems which do not satisfy the symmetry condition of Eq. (9) can be found in literature (Hill 1985; Onufriev and Ullmann 2004). The concept is to substitute the point  $\Psi = \frac{n}{2}$  at which the variance (or the slope of the titration curve) is measured by the point at which the maximal variance is reached (see Hill 1985, p. 73). This idea is a direct consequence of the reference point  $\Psi = \frac{n}{2}$  not necessarily offering special information and of the initial question whether the threshold of  $0.25n$  for the variance is exceeded. The variance of a system of independent binding sites will always be smaller than or equal to this bound and thus, exceeding this threshold anywhere is an indicator for non-negligible interaction. Moreover, we have the obvious implication that, if the Hill coefficient is greater than 1 (which means the variance at  $\Psi = \frac{n}{2}$  exceeds  $0.25n$ ) this implies also that the maximal variance exceeds this threshold. Thus, the maximal variance exceeding this threshold is obviously a generalization of the Hill coefficient.

### 3.1.4 A further extension of the Hill coefficient: different binomial distributions as reference for the variance

Instead of using the variance of the binomial distribution  $\text{Bin}(n, p = 0.5)$  as constant reference and which defines the upper bound of  $0.25n$  for the variance of a system of independent variables, we can also compare the variance of the considered system point-wise to the variance of a binomial distribution  $\text{Bin}(n, p = \frac{\Psi}{n})$  (Abeliovich 2005):

**Proposition 4** (Binomial variance as reference) *Let us consider a target molecule with  $n$  binding sites. Moreover, let a  $\lambda_0$  exist such that*

$$\mathbb{V}(\lambda_0) > \Psi(\lambda_0) \left( 1 - \frac{\Psi(\lambda_0)}{n} \right). \quad (12)$$

*Then the observed overall titration curve cannot result from a system of  $n$  independent binding sites.*

The derivation of this result which we found in literature was based on the additional assumptions of *identical* binding sites. Note that this assumption is not made here (for a proof see Sect. 6).

Proposition 4 further generalizes the criterion of the maximal variance exceeding  $0.25n$ : In case that the maximal variance exceeds  $0.25n$ , the variance also exceeds the right-hand side of Eq. (12), since it is bounded by  $0.25n$ . Moreover, this criterion gives much sharper bounds for the variance that an independent system can exhibit and extends an elegant generalization procedure: The Hill coefficient of Definition 1 compares a variance at a specific point with a reference variance. The maximum variance criteria uses the same fixed threshold of  $0.25n$  but compares all variances of the system to this fixed boundary. At last, the criterion given by Eq. (12) defines saturation dependent bounds and checks whether this boundary function is exceeded at any point.

However, it is important to note here that both generalizations of the Hill coefficient are based on the idea of comparing the variance at fixed  $\lambda$  with the variance of a reference system. The following criterion is based on an alternative concept of investigating the *change* of the variance, when  $\lambda$  changes.

### 3.2 A sigmoidal shape in a linear plot

Even though [Stefan and Le Novère \(2013\)](#) use a definition of positive cooperativity based on microstates (a change in affinity of the ligand to a certain site if the ligand binds to another site), they state that if the normed titration curve “as a function of ligand concentration is sigmoidal in shape, as observed by Bohr for hemoglobin, this indicates positive cooperativity. If it is not, no statement can be made about cooperativity from looking at this plot alone”. We will investigate how this property of a sigmoidal binding curve is related to large variances which were the characteristics of interest in Sect. 3.1. We start with a definition of a sigmoidal shape.

**Definition 2** An overall titration curve  $\Psi$  is said to be of sigmoidal shape in a linear plot if a  $\lambda_0 > 0$  exists such that

$$\frac{d^2\Psi}{d\lambda^2} \Big|_{\lambda_0} > 0 \quad (13)$$

This condition can be rewritten to

#### Lemma 1

$$\lambda_0 \text{ exists such that } \frac{d^2\Psi}{d\lambda^2} \Big|_{\lambda_0} > 0 \Leftrightarrow \frac{\mathbb{V}(\lambda)}{\lambda} \text{ has a local maximum}$$

Proposition 5 states why a sigmoidal shape in a linear plot can be regarded as an indicator of cooperative binding.

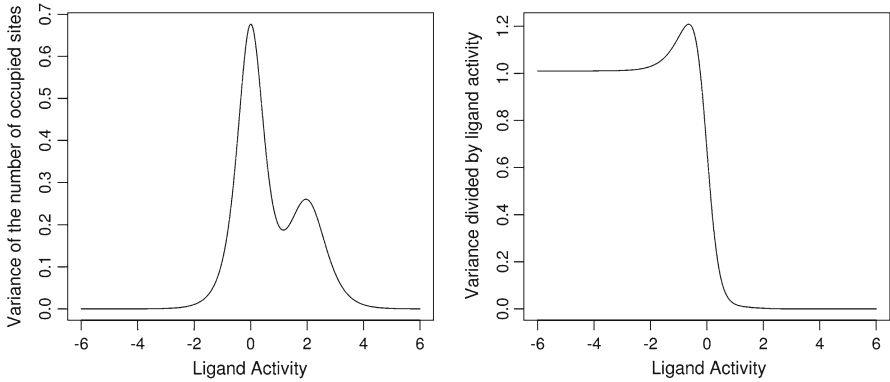
**Proposition 5** *Let  $\Psi(\lambda)$  be an overall titration curve of sigmoidal shape. Then  $\Psi(\lambda)$  cannot result from a system of independent binding sites.*

Proposition 5 gives another criterion for an “abnormal” ligand binding behavior of the system by comparing the change in variance to a linear function. The question arises how the indications of cooperativity by the different criteria are related. The answer is that there is no implication which is illustrated by the following examples (first we will compare the criterion of maximal variance exceeding  $0.25n$  to sigmoidal shape and then add an example comparing the latter to criterion (12)).

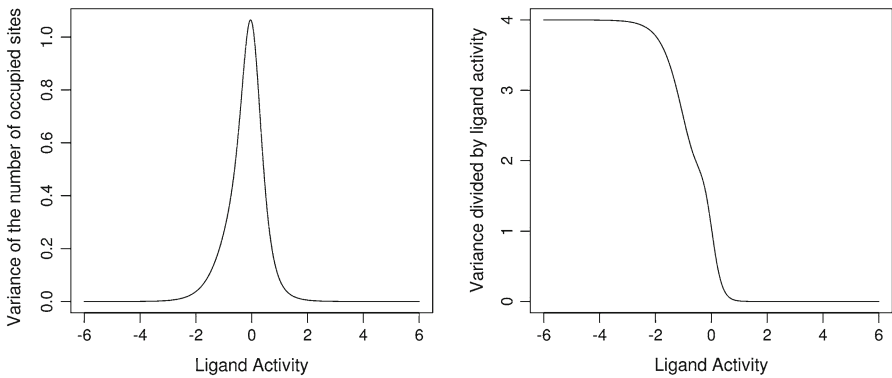
*Example 4* Let us consider the binding polynomial

$$0.01\lambda^3 + 1.01\lambda^2 + 1.01\lambda + 1.$$

Then the maximal variance of the system is smaller than 0.75 (for a formal proof see Sect. 6), which is the maximal value an independent system of three binding sites



**Fig. 2** *Left side* the variance of the number of occupied sites of the system of Example 4 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). *Right side* the variance divided by  $\lambda$  of the same binding system



**Fig. 3** *Left side* the variance of the number of occupied sites of the system of Example 5 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). *Right side* the variance divided by  $\lambda$  of the same binding system

could reach. However, the corresponding function  $\lambda^{-1}\nabla(\lambda)$  has a local maximum (criterion (12) is also satisfied). Both curves are illustrated in Fig. 2 (for earlier work on differences between the Hill coefficient and a sigmoidal shape, see also Ge and Qian 2009).

*Example 5* Let us consider the binding polynomial  $4\lambda^3 + 2\lambda^2 + 4\lambda + 1$ . Then the maximal variance of the system is approximately 1.06 which is greater than 0.75, which is the maximal value an independent system could reach. However, the corresponding function  $\lambda^{-1}\nabla_\lambda(|k|)$  does not have a local maximum (for a formal proof see Sect. 6). Since here, the maximal variance exceeds the threshold  $0.25n$ , criterion (12) is fulfilled, too. Both curves are illustrated in Fig. 3.

Example 5 shows that there are cases in which Eq. (12) is fulfilled but the curve is not sigmoidal. What remains to be shown is that the reverse case, which means having

sigmoidal shape but not fulfilling Eq. (12) exists as well. We did not find an example for this in the case of three binding sites. However, Example 6 describes a molecule with five binding sites and these properties.

*Example 6* The polynomial  $10\lambda^5 + 10^3\lambda^4 + (10^4 + 1)\lambda^3 + 110\lambda^2 + 10^3\lambda + 1$  defines a sigmoidal curve (to see this, plot the function  $\lambda^{-1}\nabla(\lambda)$  on the interval  $\lambda \in [10^{-2}, 1]$ ) but the variance does not exceed the bound given by Eq. (12) (see Sect. 6).

### 3.3 Non-real roots of the binding polynomial as a definition of cooperative ligand binding

The use of the existence of non-real roots of the binding polynomial as a definition of cooperative binding (Onufriev and Ullmann 2004; Martini 2014) has the big advantage that it unifies and generalizes both previously mentioned concepts in a way: A system with non-interacting, stochastically independent binding sites for every value of  $\lambda$  possesses a binding polynomial with real roots only (see e.g. Martini and Ullmann 2013). Consequently, the appearance of non-real roots indicates that the polynomial cannot belong to a system of independent binding sites. Conversely, if a binding polynomial has only real roots then none of the previously described criteria can be fulfilled, since it is not possible to distinguish the system from an independent system in which the binding energies of the binding sites are transformed roots of the polynomial (see e.g. Martini and Ullmann 2013; Martini et al. 2013b). Since both previously described types of criteria measure an abnormal deviation from the behavior of an independent system, both criteria imply that the binding polynomial has at least two roots which are non-real.

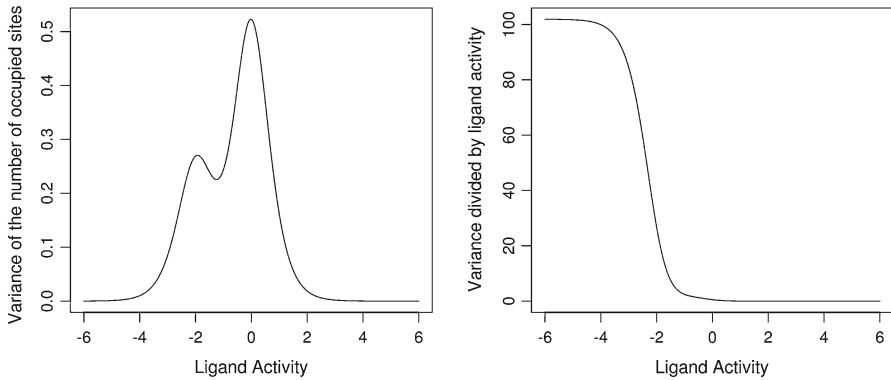
Moreover, there are cases in which the system exhibits “abnormal” behavior which can not result from an independent system, since the binding polynomial has complex roots, but both previously discussed criteria are not fulfilled, which is illustrated by Example 7. Still, the cooperativity of this system is captured by the property of the polynomial possessing non-real roots.

*Example 7* An example of a polynomial which shows a titration behavior which can not result from an independent system but whose variance does not exceed the boundary given by Eq. (12) and whose corresponding titration curve is not of sigmoidal shape is

$$100\lambda^3 + 191\lambda^2 + 101.9\lambda + 1.$$

Two of its roots are non-real. The fact that the two criteria of a maximal variance exceeding the threshold of 0.75 and being of sigmoidal shape are both not fulfilled is illustrated in Fig. 4. A theoretical treatment of  $\lambda^{-1}\nabla(\lambda)$  not having a local maximum and the bound (12) not being exceeded is given in Sect. 6.

However, compared to the criterion of measuring the maximal variance, the existence of non-real roots as a definition for cooperativity has the disadvantage that it is not directly obvious whether and how the cooperative behavior can be quantified. For a system with three binding sites, the number of non-real roots will either be zero or two and a quantification of the cooperative effects in the system might only



**Fig. 4** *Left side* the variance of the number of occupied sites of the system of Example 7 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). *Right side* the variance divided by  $\lambda$  of the same binding system

be possible when the imaginary parts of the roots are incorporated in an appropriate measure. Moreover, having the definition of the existence of non-real roots in mind, the question arises if it can make sense to distinguish between different types of cooperativity (positive and negative) for general systems based on macroscopic properties, especially since the connection between macroscopic and microscopic definitions of cooperative binding is not straightforward (We will show this later in Sect. 4). This question is important since a lot of research investigates effects related to positive and negative cooperativity (often with certain preconditions on the target molecule) (e.g. [Abeliovich 2005](#); [Alvarado et al. 2010](#); [Bush et al. 2012](#); [Erion and Strobel 2011](#); [Kaczanowska et al. 2014](#)).

### 3.4 Other measures of cooperativity based on the macroscopic behavior

Other additional approaches of capturing cooperative behavior as a real number exist. However, these concepts will have consistency problems if no additional conditions, typically such as a restriction to two binding sites [in this case several implications between the different types of cooperativity exist ([Onufriev and Ullmann 2004](#); [Martini 2014](#))] or the assumption of identical binding sites are imposed.

An example of another coefficient which captures cooperative effects is the  $\alpha$ -coefficient used in the review by [Hunter and Anderson \(2009\)](#). Assuming that the concentration of the free ligand is constant (this assumption is more or less fulfilled if the number of ligands is much higher than that of the target molecule and it is the basic assumption for the validity of the grand canonical ensemble), the coefficient  $\alpha$  of a system with binding polynomial  $a_2\lambda^2 + a_1\lambda + 1$  is given by

$$\alpha = \frac{4a_2}{a_1^2}.$$

This expression represents more or less the discriminant of the polynomial. Consequently, the polynomial will have non-real roots if  $\alpha > 1$  and thus the  $\alpha > 1$ -criterion coincides with the appearance of non-real roots ( $n = 2$ ). An open question is how to define  $\alpha$  in the case of more than two binding sites in the best way (as an  $\alpha_i$  for each binding step or an overall  $\alpha$ ) and how to decide then whether the target molecule binds its ligand cooperatively or not.

## 4 Cooperative binding as a property of the family of distributions on the microstates

### 4.1 Cooperative binding defined by interaction energies of binding sites

We consider a molecule with  $n = 2$  binding sites and the free energy differences of each state  $k$  with respect to the chosen reference state  $\{0\}^{n=2}$  (all sites unoccupied). As already previously mentioned,  $G(k)$  shall denote the difference between the free energies of the molecule in state  $k$  and the reference state. For a molecule with only two binding sites, the (relative) energies of the four states are

$$G((0, 0)) = 0, \quad G((1, 0)), \quad G((0, 1)), \quad G((1, 1))$$

We call  $G((1, 0))$  the binding energy of site 1,  $G((0, 1))$  the binding energy of site 2 and  $W_{1,2} := G((1, 1)) - G((1, 0)) - G((0, 1))$  the interaction energy. If the number of binding sites  $n$  is greater than 2, we can obtain also interaction terms of higher order, which will be illustrated later. For the moment, we consider a molecule with only pairwise interaction, which means that for any number of binding sites, the energies of every state are determined by the energies of states with one and two sites occupied. With  $k_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$  (only site  $i$  is occupied),  $g_i := \exp(-G(k_i)/RT)$  and  $w_{i,j} := \exp(-W_{i,j}/RT)$ , we can describe a molecule with  $n$  binding sites and only pairwise interaction by a tuple of  $n$  binding constants and  $\frac{n(n-1)}{2}$  interaction constants:

$$M = (g_1, \dots, g_n, w_{1,2}, \dots, w_{n-1,n}).$$

We can use the following definition of cooperative binding on microstate properties.

**Definition 3** Let  $M$  be a molecule with  $n$  binding sites and only pairwise interaction. Then sites  $i$  and  $j$  are said to bind the ligand

- positive cooperatively if  $w_{i,j} > 1$  (which corresponds to  $W_{i,j} < 0$ )
- non-cooperatively if  $w_{i,j} = 1$  (which corresponds to  $W_{i,j} = 0$ )
- negative cooperatively if  $w_{i,j} < 1$  (which corresponds to  $W_{i,j} > 0$ ).

At first sight, Definition 3 seems to be an appropriate way to define cooperativity. Advantages are that the definition is simple and unambiguous and that it captures the concept of an altered affinity of a ligand to a certain site if another site is occupied which often is the motivating idea (Berg et al. 2007; Stefan and Le Novère 2013): The binding of a ligand to the first binding site of a molecule with two binding sites leads

to a change of the energy level  $G(k_1)$  if the other site is unoccupied but the molecule gains  $G(k_1) + W_{1,2}$  if the second site is already occupied.

As a first observation, we see that there are molecules which exhibit different types of cooperativity according to Definition 3 but which are not distinguishable on the macroscopic level (Example 8 is taken from Martini and Ullmann 2013):

*Example 8* The following molecules  $(g_1, g_2, w_{1,2})$  are examples of positive, negative, or non cooperative ligand binding according to Definition 3, but share the same overall titration curve and thus all macroscopic properties.

$$(2, 1, 0.5), \left(1.5 + \sqrt{1.25}, 1.5 - \sqrt{1.25}, 1\right), \left(0.1, 2.9, \frac{100}{29}\right).$$

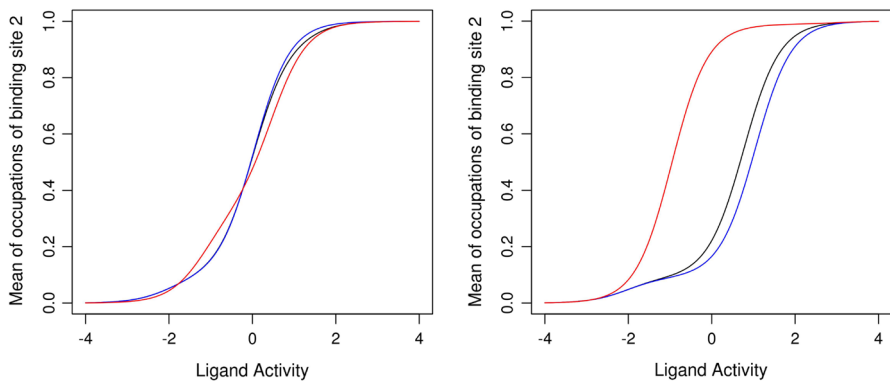
Since the roots of the binding polynomial are real, the overall titration curve is the sum of two standard Henderson–Hasselbalch (Henderson 1913; Hasselbalch 1916) curves describing the independent binding curves of the sites of the second molecule (Onufriev et al. 2001; Martini and Ullmann 2013; Martini et al. 2013b). Moreover, since the sites are energetically not identical, the point of maximal variance of both sites will not coincide, which means that the Hill coefficient of this system is smaller than 1. This fact will lead to a classification as undecidable or negative cooperative binding if the underlying assumption that the sites are equal is made. The curve is not sigmoidal [undecidable (Stefan and Le Novère 2013)], the roots are non complex (non cooperative binding), the coefficient  $\alpha$  (Hunter and Anderson 2009) is smaller than 1 (analogous to the Hill coefficient) and Definition 3 classifies the three systems as negative, non and positive cooperative molecules.

Example 8 illustrates that it is not in general possible to distinguish between the different types of cooperativity of Definition 3 based on the macroscopic binding isotherm. However, for two binding sites, we still have the implication that non-real roots of the binding polynomial imply a negative interaction energy (positive cooperative binding according to Definition 3 (Martini and Ullmann 2013, Corollary 2)). For more than four binding sites, it is not possible to deduce positive cooperativity as defined in Definition 3 from the existence of non-real roots. This point can be illustrated by a molecule with five binding sites (Example 3 of Martini and Ullmann 2013) showing negative cooperativity between all binding sites (according to Definition 3) and two non-real roots.

Another problem is that in larger systems ( $n > 2$ ), negative cooperativity in terms of interaction energy between different sites can also lead to microscopic effects which are usually expected for “positive” cooperative relations (depending on the definition of “affinity”): Example 9 illustrates that an occupation of a certain site can increase the probability of another site being occupied, in a molecule which is purely negative-cooperative, according to Definition 3.

*Example 9* Let us consider the molecules

$$M_1 = (g_1, g_2, g_3, w_{1,2}, w_{1,3}, w_{2,3}) = (100, 10, 1, 0.1, 0.1, 0.5)$$



**Fig. 5** *Left side* the (conditional) probability of site 2 being occupied in molecule  $M_1$  of Example 9 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used): *black line* unconditioned (or marginal) probability of site 2 being occupied. *Blue curve* conditional probability of site 2 being occupied, conditioned on site 3 being unoccupied. *Red curve* conditional probability of site 2 being occupied, conditioned on site 3 being occupied. *Right side* analogous color encoding for site 2 of molecule  $M_2$

and

$$M_2 = (100, 10, 1, 0.01, 0.001, 0.9).$$

All sites cooperate negatively according to Definition 3, in both molecules. In particular, for any fixed state of site 1, the ligand has a higher affinity to site 2, if site 3 is unoccupied than if site 3 is occupied: For instance for molecule  $M_1$ , the state transition from  $(0, 0, 0)$  to  $(0, 1, 0)$  gives the weight  $g_2$ , while the transition from  $(0, 0, 1)$  to  $(0, 1, 1)$  multiplies the factor  $g_2 w_{2,3} < g_2$  to the weight of the state. Analogously, if site 1 is occupied, the transition from  $(1, 0, 0)$  to  $(1, 1, 0)$  produces the factor  $g_2 w_{1,2}$  while the transition from  $(1, 0, 1)$  to  $(1, 1, 1)$  gives the factor  $g_2 w_{1,2} w_{2,3}$  for the weight of the state.

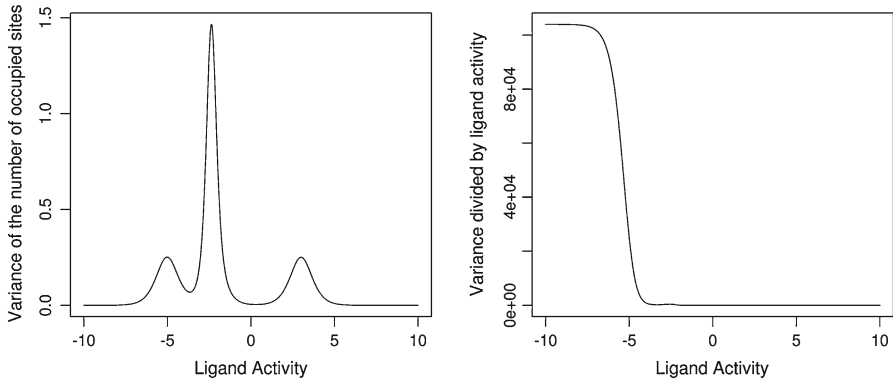
However, if we compare the marginal probability of site 2 being occupied to the conditional probabilities conditioned on site 3 being unoccupied and to the probability conditioned on site 3 being occupied (Fig. 5), we see that the probability of site 2 being occupied can be increased if site 3 is occupied. For molecule  $M_2$ , an occupation of site 3 does not decrease the probability of site 2 being occupied for any activity of the ligand. This effect is caused by the circumstance that the influence of an occupied site 1 on the binding to site 2 is more negative than that of site 3 and an occupation of site 3 reduces the probability of site 1 being occupied.

Example 9 shows that in a system which is fully negative cooperative according to Definition 3, the binding of a ligand to a certain site can increase the probability of another site being occupied. This is caused by the principle that “the enemy of my enemy is my friend” and this principle can also lead to a macroscopic behavior which will be classified as positive cooperativity, which is illustrated by Example 10.

*Example 10* Let us consider the molecule  $M$  with five binding sites, described by

$$M = (g_1, g_2, g_2, g_4, g_5, w_{1,2}, w_{1,3}, w_{1,4}, w_{1,5}, w_{2,3}, w_{2,4}, w_{2,5}, w_{3,4}, w_{3,5}, w_{4,5})$$





**Fig. 6** *Left side* the variance of the number of occupied sites of the system of Example 10 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). *Right side* the variance divided by  $\lambda$  of the same binding system

$$= (10^3, 10^3, 10^3, 10^3, 10^5, 1, 1, 1, 10^{-2}, 1, 1, 10^{-2}, 1, 10^{-2}, 10^{-2}).$$

According to Definition 3, all sites do either not cooperate or cooperate negatively. However, the polynomial

$$\Phi(\lambda) = 10^9\lambda^5 + 10004 \cdot 10^8\lambda^4 + 406 \cdot 10^7\lambda^3 + 10^7\lambda^2 + 104000\lambda + 1$$

has non-real roots and it is sigmoidal (based on Definition 2). Moreover, the variance exceeds the threshold of 1.25, which would be assessed as positive cooperative behavior (see Fig. 6).

Finally, another problem with Definition 3 appears when we deal with larger systems: The advantage of a simple discrimination between positive, non and negative cooperativity can vanish in larger systems, since additional interaction terms of higher order can contradict the pairwise interaction terms.

*Example 11* Let  $M$  be a system with three binding sites with microstate constants

$$\begin{aligned} g((1, 0, 0)) &= 1, & g((0, 1, 0)) &= 1, & g((0, 0, 1)) &= 1, \\ g((1, 1, 0)) &= 0.1, & g((1, 0, 1)) &= 0.1, & g((0, 1, 1)) &= 0.1, \\ g((1, 1, 1)) &= 1 \end{aligned}$$

Then the binding behavior can not be captured by a model with only pairwise interaction. Moreover, all pairwise terms  $w_{i,j}$  are smaller than 1 which would indicate negative cooperativity if Definition 3 would be extended to a model with interaction terms of higher order. Since the interaction term of order three  $w_{1,2,3}$  equals  $10^3$ , the question arises whether the sites should be called negatively cooperating or not.

## 4.2 Conditional probability functions

Another idea to distinguish between different types of cooperativity on microstates and which has already been mentioned in the context of Example 9 is the comparison of the conditional probabilities

$$P(k_i = 1|k_j = 0) \quad \text{vs.} \quad P(k_i = 1|k_j = 1)$$

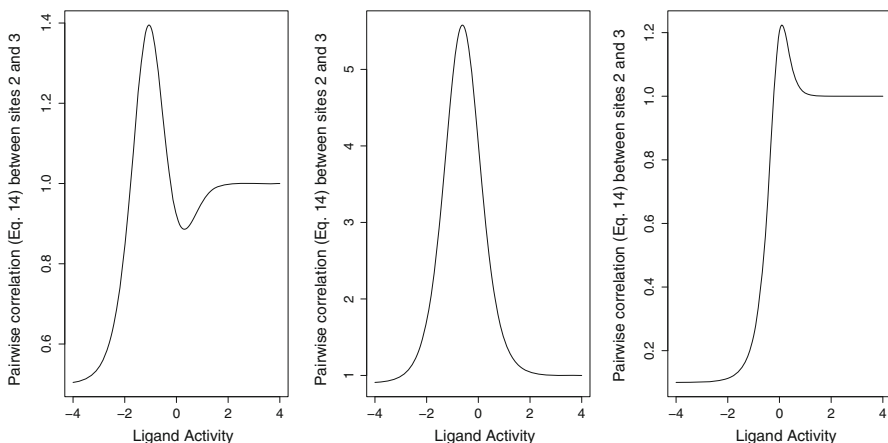
as functions of  $\lambda$  (for which the conditional probabilities exist, i.e.  $P(k_j = 0) \neq 0 \neq P(k_j = 1)$ ). Here the problem arises that a simple separation into positive, non and negative cooperativity may not be possible if the whole functions are considered, which has already been illustrated by molecule  $M_1$  of Example 9 (see Fig. 5), in which the conditional probability functions intersect. A variant to characterize a system's cooperative behavior by conditional probabilities was used by Ben-Naim (2001) who used the function

$$\frac{P(k_1 = 1, k_2 = 1)}{P(k_1 = 1)P(k_2 = 1)} = \frac{P(k_1 = 1|k_2 = 1)}{P(k_1 = 1)} = \frac{P(k_2 = 1|k_1 = 1)}{P(k_2 = 1)} \quad (14)$$

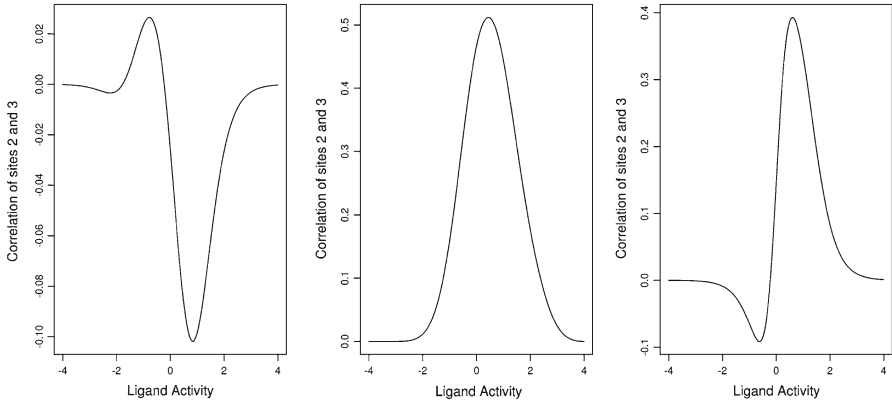
for systems with two binding sites and analogous generalizations for larger systems. Of special interest is here the limit

$$\lim_{\lambda \rightarrow 0} \frac{P(k_1 = 1|k_2 = 1)}{P(k_1 = 1)} \quad (15)$$

which can be used to distinguish between negative, positive and non-cooperativity by comparing the limit to the reference value 1 (Ben-Naim 2001). Obviously, Eq. (14)



**Fig. 7** Left side the pairwise “correlation” defined by Eq. (14) of sites 2 and 3 of system  $M_1$  of Example 9 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). Middle analogously for system  $M_2$  of Example 9. Right side the pairwise “correlation” of molecule  $M$  of Example 11 which includes interaction terms of higher order



**Fig. 8** *Left side* the pairwise correlation defined by Eq. (16) (rescaled by the products of the square roots of the variances) of sites 2 and 3 of system  $M_1$  of Example 9 as a function of the ligand activity (constant temperature, a decadic logarithmic scale is used). *Middle* analogously for system  $M_2$  of Example 9. *Right side* the pairwise correlation of molecule  $M$  of Example 11 which includes interaction terms of higher order

equals the constant function 1 if the sites are stochastically independent. Exemplarily, the pairwise “correlation” functions (Ben-Naim 2001) defined by Eq. (14) of sites 2 and 3 of molecules  $M_1$  and  $M_2$  of Example 9 are illustrated in Fig. 7.

### 4.3 The covariance function

Another possibility to characterize the cooperative behavior is to use the covariance function (or the stochastic correlation)

$$\text{Cov}(k_1, k_2) := \mathbb{E}k_1k_2 - \mathbb{E}k_1\mathbb{E}k_2. \tag{16}$$

The basic concept that underlies this approach is coinciding with Eq. (14): If the sites bind the ligand independently, the covariance will be zero and still a function instead of a single value has to be used to characterize the cooperativity between the sites fully. Compared to the approach of Eq. (14), we have the disadvantage here that the limit for  $\lambda \rightarrow 0$  does not give any information. The stochastic correlations (covariance of  $k_1, k_2$  divided by the square root of the product of both variances) of sites 2 and 3 of the molecule  $M$  of Example 11 and of molecules  $M_1$  and  $M_2$  of Example 9 are illustrated in Fig. 8. Both concepts of correlation highlight the fact that the degree of interaction between sites depends on the state of the environment. In particular, interacting binding sites can behave in an independent way at certain non-trivial ligand concentrations (when Eq. (14) equals one and Eq. (16) equals zero). The use of correlation functions instead of single values seem to be more appropriate to capture the interaction between sites. In particular, this concept also solves the problem of possibly contrary interaction energy terms of different order.

## 5 Summary and outlook

In this work, we discussed the relation between different definitions of cooperative binding and presented extreme examples to illustrate differences between the concepts of cooperativity. For cooperativity defined on the overall titration curve, i.e. on macroscopic properties, we highlighted that a Hill coefficient larger than 1 shows that the variance of the distribution on the macrostates of the considered system is at half-saturation higher than the maximal variance a system consisting of independent binding sites can reach. In this context, we underlined that the reference point of half-saturation does not in general offer special information and discussed generalizations of the Hill coefficient which also consider “abnormally high” variances. The most general criterion measuring “abnormally high” variances is the threshold given by Eq. (12) (Abeliovich 2005), which we proved also for systems consisting of non-identical binding sites. Moreover, we showed that a sigmoidal shape of an overall titration curve in a linear plot also indicates that the considered system cannot consist of independent binding sites and that—contrary to the Hill coefficient of Definition 1 and its generalizations—this criterion does not consider the value of the variance as important characteristic, but the behavior of the change in variance with respect to a change of  $\lambda$ . We presented different examples to show that these two concepts of measuring “abnormal” variance (by considering its value or its change) do not coincide. In this regard, we highlighted that non-real roots of the binding polynomial can be a unifying definition for cooperativity. This criterion offers the advantage that both previously mentioned concepts imply that the binding polynomial has at least two roots which are non-real and thus it is a generalization. Moreover, the appearance of non-real roots is actually also the only relevant observation to see that the observed titration curve cannot be a result of an independent system. Any binding polynomial which has real roots only can be regarded as an independent system of binding sites whose binding energies are given by transformation of the roots of the binding polynomial (Onufriev et al. 2001; Martini and Ullmann 2013; Martini et al. 2013b). Conversely, if the polynomial has non-real roots it can not belong to a system of independent binding sites, since the independence of the sites would lead to a binding polynomial which is the product of the polynomials (linear factors) of all individual sites and which consequently would have only real roots. However, based on this definition, it is not clear how the degree of cooperativity can be quantified and whether the discrimination between positive and negative cooperativity can be well defined. For cooperativity defined on the measures on microstates, we also showed that a single value alone as a description of the cooperative ligand binding behavior of a molecule might not be enough to characterize the system’s properties. Instead different correlation functions could be used. However, even based on the definitions on microstates it is not clear how to distinguish between positive and negative cooperativity well, since the qualitative behavior of the correlation can depend on the ligand activity, already if systems with three binding sites are regarded. Concerning the relation between the definitions of cooperativity on the macro- and microstates, we presented examples showing that “negative” cooperativity on the microscopic level can produce phenomena which are usually assigned to “positive” cooperative ligand binding. Considering additional information on the

macroscopic binding kinetics might allow the deduction of more microscopic properties from the macroscopic binding behavior (Martini et al. 2013a; Martini and Habeck 2015). The presented examples create the suspicion that a real ligand activity independent discrimination between “positive” and “negative” cooperativity, based on the overall binding isotherm only, is impossible if general binding systems are considered. Overall, this work showed that scientists investigating cooperative effects should clearly state what cooperativity means in the context of their work and check whether preconditions have to be satisfied by the target molecule for consistency of the used definition(s).

## 6 Proofs

*Proof of Proposition 1* (a) Let  $\Phi'$  denote the numerator of Eq. (3) (this is also the derivative of  $\Phi$  with respect to  $\log(\lambda)$ ). Then the constant  $n$  of Eq. (4) has to coincide with the number of binding sites, since if  $a_n \neq 0$  the overall binding curve will tend to the number of binding sites for  $\lambda \rightarrow \infty$ . By precondition, we know

$$\frac{\Phi'}{\Phi} = n \frac{K\lambda^\alpha}{1 + K\lambda^\alpha}$$

which states that a function  $\beta(\lambda)$  exists such that

$$\beta(\lambda)(1 + K\lambda^\alpha) = \Phi \quad \text{and} \quad \beta(\lambda)K\lambda^\alpha = \frac{1}{n}\Phi' \tag{17}$$

which implies

$$\beta(\lambda) = \Phi - \frac{1}{n}\Phi' \tag{18}$$

and thus  $\beta(\lambda)$  is a polynomial of degree smaller or equal to  $(n - 1)$ , since it is a difference of two polynomials with the same leading coefficient. Moreover, the constant term of  $\beta(\lambda)$  is 1, since the constant term of  $\Phi$  is 1. Thus,  $\alpha$  has to be a natural number (a positive integer), since otherwise  $\beta(\lambda)(1 + K\lambda^\alpha)$  would not be a polynomial.

(b) In the case of  $\alpha = 1$ , the statement is true. Let us consider now the case  $\alpha > 1$ . Let  $\beta(\lambda) = b_{n-1}\lambda^{n-1} + \dots + b_1\lambda + 1$ . Equation (17) implies that  $b_i = 0, \forall i > n - \alpha$ . Moreover, it shows that

$$b_i + b_{i-\alpha}K = a_i \quad \text{and} \quad b_{i-\alpha}K = \frac{i}{n}a_i \tag{19}$$

which implies

$$\frac{n-i}{i}b_{i-\alpha}K = b_i \quad \forall i \in \{\alpha, \dots, n\}. \tag{20}$$

For  $i = n$ , Eq. (20) is always fulfilled, since both sides are zero, independent of the value of  $b_{n-\alpha}$ . However, for  $i \neq n$ , this equation states that  $b_i = 0$  implies  $b_{i-\alpha} = 0$ , which gives

$$b_i = 0, \quad \forall i \in \{n - j\alpha + 1, \dots, n - (j - 1)\alpha - 1 \mid j \in \mathbb{N}\}$$

for any  $j$ , since  $b_i = 0, \forall i > n - \alpha$ . Thus,

$$\beta(\lambda) = b_{n-\alpha}\lambda^{n-\alpha} + b_{n-2\alpha}\lambda^{n-2\alpha} + \dots + \underbrace{b_{n-m\alpha}}_{=1} \tag{21}$$

and thus  $\alpha$  divides  $n$ .

(c)  $0 < a_n = b_{n-\alpha}K \Rightarrow b_{n-\alpha} > 0$ , since  $K > 0$ . Eq. (20) gives  $0 < b_{n-\alpha} = \frac{\alpha}{n-\alpha}b_{n-2\alpha}K$  which implies  $b_{n-2\alpha} > 0$ , as long as  $2\alpha \leq n$ . This implies that

$$a_{n-i\alpha} = b_{n-i\alpha} + b_{n-(i+1)\alpha}K > 0.$$

□

*Proof (Hill coefficient of a Hill equation)* Using Definition 1, the Hill coefficient of an equation of shape (4) is

$$\frac{d}{d \log(\lambda)} \left( \log \left( \frac{\frac{nK \exp(\alpha \log(\lambda))}{1+K \exp(\alpha \log(\lambda))}}{n - \frac{nK \exp(\alpha \log(\lambda))}{1+K \exp(\alpha \log(\lambda))}} \right) \right) = \frac{d}{d \log(\lambda)} (\log(n) + \alpha \log(\lambda)) = \alpha.$$

□

*Proof of Proposition 2* See Hill (1985), result in Eq. (11.41). □

*Proof of Proposition 3*  $\Psi = \frac{\Phi'}{\Phi}$  is a quotient of two polynomials in the variable  $\lambda$ . We introduced the notation  $f'$  for the derivative with respect to  $\log(\lambda)$ .

$$\Psi' = \frac{\Phi''\Phi - \Phi'\Phi'}{\Phi^2}$$

is also a rational function in the variable  $\lambda$ . To see that  $\Psi'$  has a local extremum we consider its derivative with respect to  $\lambda$ , which is a rational function in  $\lambda$ , since  $\Psi'$  is a rational function. Thus,  $\frac{\partial}{\partial \lambda} \Psi' = 0$  means that its numerator is equal to zero. Since this numerator is a polynomial, it equals zero on a finite number of points  $\lambda_1, \dots, \lambda_m$ . We assume the equality

$$\Psi'_{|\lambda=a\lambda_0} = \Psi'_{|\lambda=a^{-1}\lambda_0} \tag{22}$$

which we will deduce from Eq. (9) later.  $\Psi'$  is continuous (on  $\lambda \geq 0$ ) and it has only a finite number of local extremes, since  $\frac{\partial}{\partial \lambda} \Psi'$  has only a finite number of zeros. Let

$\epsilon > 0$  be a number smaller than the distance of  $\lambda_0$  to the nearest root of the enumerator of  $\frac{\partial}{\partial \lambda} \Psi'$ . Then the ball  $B_\epsilon(\lambda_0)$  does not include more than one extremum (which would have to be in  $\lambda_0$ ). Let  $(a_n)_{n \in \mathbb{N}} \rightarrow 1$  be a strictly monotonously increasing sequence with  $a_n \lambda_0 \in B_\epsilon(\lambda_0) \forall n$ . Then  $a_n \lambda_0 \rightarrow \lambda_0$  increases strictly monotonously, and  $\Psi'(a_n \lambda_0) \rightarrow \Psi'(\lambda_0)$  is strictly monotonously (either increasing or) decreasing, since otherwise another root of the denominator of  $\frac{\partial}{\partial \lambda} \Psi'$  within the ball  $B_\epsilon(\lambda_0)$  would exist. Then, the sequence  $\Psi'((a_n^{-1} \lambda_0)) \rightarrow \Psi'(\lambda_0)$  is strictly monotonously (either increasing or) decreasing, due to Eq. (22), but with  $a_n^{-1} \lambda_0$  strictly monotonously decreasing. Consequently,  $\Psi'$  has a local extreme in  $\lambda_0$ . To prove Eq. (22), we regard the following:

$$\Psi'(\lambda) = \frac{d}{d \log \lambda} (\Psi(\lambda) \circ e^{\log \lambda}) = \left[ \frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=e^{\log \lambda}} \cdot \lambda \right].$$

Thus, Eq. (22) is true if

$$\frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=e^{\log a \lambda_0}} \cdot a \lambda_0 = \frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=e^{\log a^{-1} \lambda_0}} \cdot a^{-1} \lambda_0 \tag{23}$$

which means (if  $\lambda_0 \neq 0$ )

$$\frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=a \lambda_0} = a^{-2} \frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=a^{-1} \lambda_0}. \tag{24}$$

To see that Eq. (24) is true, we consider the functions  $f(a) := a \lambda_0$  and  $g(a) = a^{-1} \lambda_0$  and rewrite Eq. (9)

$$\Psi \circ f(a) + \Psi \circ g(a) = n \tag{25}$$

Calculating the derivative with respect to  $a$  gives

$$\frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=a \lambda_0} \cdot \lambda_0 - \frac{d}{d \lambda} \Psi(\lambda) \Big|_{\lambda=a^{-1} \lambda_0} \lambda_0 a^{-2} = 0$$

which gives Eq. (24). □

*Proof of Proposition 4* Let  $X_i$  denote the Bernoulli variable describing whether site  $i$  is occupied or not. The variance of the distribution on the macrostates  $\mathbb{V}(\lambda)$  is given by the sum of the variances  $\mathbb{V}_\lambda(X_i)$  and covariances  $\text{Cov}_\lambda(X_i, X_j)$

$$\mathbb{V}(\lambda) = \sum_{i=1}^n \mathbb{V}_\lambda(X_i) + \sum_{i=1, j=1, i \neq j}^n \text{Cov}_\lambda(X_i, X_j)$$

We will show that  $\mathbb{V}(\lambda)$  exceeding the threshold of Eq. (12) implies that

$$\sum_{i=1, j=1, i \neq j}^n \text{Cov}_\lambda(X_i, X_j) > 0,$$

which in particular implies that the sum of the covariances is not zero and which again means that the considered system cannot consist of independent binding sites. For this purpose, we show that

$$\sum_{i=1}^n \mathbb{V}_\lambda(X_i) \leq \Psi \left(1 - \frac{\Psi}{n}\right).$$

With  $p_{i,\lambda}$  denoting the probability of site  $i$  being occupied at ligand activity  $\lambda$ , and thus with  $\Psi = \sum_{i=1}^n p_{i,\lambda}$ , the upper equation means

$$\sum_{i=1}^n \mathbb{V}_\lambda(X_i) = \sum_{i=1}^n p_{i,\lambda}(1 - p_{i,\lambda}) \leq \left(\sum_{i=1}^n p_{i,\lambda}\right) \left(1 - \frac{1}{n} \sum_{i=1}^n p_{i,\lambda}\right)$$

which is true if and only if

$$n \sum_{i=1}^n p_{i,\lambda}^2 \geq \left(\sum_{i=1}^n p_{i,\lambda}\right)^2.$$

This “sum of squares equation” is a result of Lagrange’s identity (see e.g. [Weisstein 2002](#), with one of the vectors of Lagrange’s identity with constant entry 1). This means that Eq. (12) implies  $\sum_{i=1, j=1, i \neq j}^n \text{Cov}_\lambda(X_i, X_j) > 0$  and consequently that the observed overall titration curve cannot result from a system of independent binding sites.  $\square$

*Proof of Lemma 1* Let a  $\lambda_0$  exist such that

$$\frac{d^2\Psi}{d\lambda^2} \Big|_{\lambda_0} > 0. \quad (26)$$

Since the second derivative of an overall titration curve will be negative if  $\lambda$  is sufficiently large, and since it is continuous, a  $\lambda_1$  exists such that  $\frac{d\Psi}{d\lambda}$  has a local maximum at  $\lambda_1$ . Moreover, since

$$\mathbb{V}(\lambda) = \frac{d\Psi}{d\lambda} \cdot \lambda \quad (27)$$

this implies that

$$\frac{\mathbb{V}(\lambda)}{\lambda}$$

has a local maximum at  $\lambda_1$  (analogously the converse).  $\square$

*Proof of Proposition 5* Let us assume that an independent system exists which produces the observed overall titration curve. We will prove the statement by contradiction.



We can write the variance of  $|k|$  as the sum the variance of the independent Bernoulli variables

$$\frac{\mathbb{V}(\lambda)}{\lambda} = \frac{\sum_{i=1}^n \mathbb{V}_\lambda(X_i)}{\lambda}.$$

The variance of a Bernoulli variable which is independent of the remaining system is given by

$$\frac{g_i \lambda}{(g_i \lambda + 1)^2}$$

with  $g_i$  characterizing the affinity of the ligand to the  $i$ th binding site and thus

$$\frac{\mathbb{V}(\lambda)}{\lambda} = \sum_{i=1}^n \frac{g_i}{(g_i \lambda + 1)^2}.$$

The right side of this equation is a sum of monotonously decreasing functions and thus cannot have a local maximum at  $\lambda_1$ . □

*Proof of Example 4* Example 4 shall demonstrate that a system in which the variance of the number of occupied sites is always below the bound  $0.25 \cdot n$  can be of sigmoidal shape in a linear plot, which means  $\frac{\mathbb{V}(\lambda)}{\lambda}$  has a local maximum.

The first property of  $\mathbb{V}(\lambda)$  being smaller than  $0.75$  can be seen in Fig. 2, however, we give a short proof to exclude the possibility that the local maximum is higher but is not recognized in the plot because of the choice of the lattice on which the function is evaluated. It is clear from the plot that values of  $\lambda$  exist such that  $\mathbb{V}(\lambda) < 0.75$ . What we have to show is (since  $\mathbb{V}(\lambda)$  is continuous) that  $\mathbb{V}(\lambda) - 0.75$  does not have a root on the positive real numbers. Writing

$$\mathbb{V}(\lambda) - 0.75 = \mathbb{E}|k|^2 - (\mathbb{E}|k|)^2 - 0.75 = \frac{Z_1}{\Phi} - \frac{Z_2}{\Phi^2} - 0.75$$

with  $\mathbb{E}$  denoting the expectation operator,  $\Phi$  the binding polynomial and  $Z_i$  the corresponding numerators gives that we are interested in the roots of the polynomial

$$Z_1 \cdot \Phi - Z_2 - 0.75\Phi^2 = 0.$$

The computer algebra program Maxima tells us that a positive root does not exist.

Moreover, to avoid discussions about the exact definition of a sigmoidal shape (exactly vs. at least one inflection point), we show that the function has exactly one inflection point, the following way: We use the computer algebra program Maxima, calculate the derivative of  $\frac{\mathbb{V}(\lambda)}{\lambda}$  with respect to  $\lambda$  (*diff(% , x, 1)*) and calculate the exact roots of the numerator, which is a polynomial (*allroots(..)*, *ratnumer(..)*). There is exactly one root in the positive numbers. □

*Proof of Example 5* The fact that  $\mathbb{V}(\lambda)$  is not bounded by 0.75 is obvious. What has to be shown is that the curve is not even slightly sigmoidal. Analogously, to the procedure described in the former proof, we calculate the derivative of  $\frac{\mathbb{V}(\lambda)}{\lambda}$  using the computer algebra program Maxima and determine the roots of the numerator. None of them is positive.  $\square$

*Proof of Example 6* The statement is that the binding isotherm defined by  $10\lambda^5 + 10^3\lambda^4 + (10^4 + 1)\lambda^3 + 110\lambda^2 + 10^3\lambda + 1$  is sigmoidal, but Eq. (12) is not fulfilled. To see that the binding curve is sigmoidal, plot the function  $\lambda^{-1}\mathbb{V}(\lambda)$  on the interval  $\lambda \in [10^{-2}, 1]$  to see that it has a local maximum. Equation (12) is not satisfied if

$$\mathbb{V}(\lambda) \leq \Psi(\lambda) \left(1 - \frac{\Psi(\lambda)}{n}\right) \quad \forall \lambda \geq 0.$$

To see that this is true, we rewrite this equation to

$$\frac{\Phi''}{\Phi} - \frac{\Phi'\Phi'}{\Phi^2} \leq \frac{\Phi'}{\Phi} \left(1 - \frac{\Phi'}{n\Phi}\right).$$

The binding polynomial  $\Phi$  has only real non-negative coefficients with constant term 1, and is consequently positive for all non-negative  $\lambda$ . We multiply both sides with  $\Phi^2$  to obtain

$$\Phi''\Phi - \Phi'\Phi' \leq \Phi'\Phi \left(1 - \frac{\Phi'}{n\Phi}\right)$$

and thus

$$\Phi''\Phi - (1 - 1/n)\Phi'\Phi' - \Phi'\Phi \leq 0$$

Plugging the polynomial of Example 6 into this equation, we have to show that

$$\begin{aligned} & -0.2 \cdot (2999900\lambda^8 + 59973000\lambda^7 + 598860006\lambda^6 - 19200120\lambda^5 \\ & - 59993400\lambda^4 + 359970\lambda^3 + 3998900\lambda^2) \leq 0 \quad \forall \lambda \geq 0 \end{aligned}$$

The left side represents a polynomial with negative leading coefficient with value 0 at  $\lambda = 0$ . Calculating its roots shows that it does not have any other root on the positive real numbers, which implies the statement.  $\square$

*Proof of Example 7* The corresponding titration curve is not sigmoidal: We used the computer algebra program maxima to calculate the derivative of  $\lambda^{-1}\mathbb{V}(\lambda)$ , which is a rational function. A local maximum has to be a root of its numerator. None of its roots is a positive real number.

To see that Eq. (12) is not fulfilled, calculate  $\Phi''\Phi - (1 - 1/n)\Phi'\Phi' - \Phi'\Phi$ , which is a polynomial of degree four with negative leading coefficient and no positive root.  $\square$

**Acknowledgments** L.D. is a researcher of CONICET (Argentina). J.W.R.M. would like to thank Mareike Busmann for helpful discussions. Moreover, we would like to thank our unknown reviewers for constructive suggestions.

## References

- Abeliovich H (2005) An empirical extremum principle for the Hill coefficient in ligand–protein interactions showing negative cooperativity. *Biophys J* 89(1):76–79
- Adair GS et al (1925) The hemoglobin system VI. The oxygen dissociation curve of hemoglobin. *J Biol Chem* 63(2):529–545
- Alvarado D, Klein DE, Lemmon MA (2010) Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell* 142(4):568–579
- Ben-Naim A (2001) *Cooperativity and regulation in biochemical processes*. Kluwer Academic/Plenum Publishers, Dordrecht
- Berg J, Tymoczko J, Stryer L (2007) *Biochemistry*. W. H. Freeman, New York
- Bush EC, Clark AE, DeBoever CM, Haynes LE, Hussain S, Ma S, McDermott MM, Novak AM, Wentworth JS (2012) Modeling the role of negative cooperativity in metabolic regulation and homeostasis. *PLoS One* 7(11):e48920
- Erion TV, Strobel SA (2011) Identification of a tertiary interaction important for cooperative ligand binding by the glycine riboswitch. *RNA* 17(1):74–84
- Ge H, Qian M (2009) Theoretical analysis of the relationship between positive/negative cooperativity and enzyme activation/inhibition. *Interdiscip Sci Comput Life Sci* 1(3):204–213
- Gutierrez PS, Monteoliva D, Diambra L (2012) Cooperative binding of transcription factors promotes bimodal gene expression response. *PLoS One* 7:e44812
- Hasselbalch K (1916) *Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensäure desselben, und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl*. Julius Springer, Berlin
- Henderson LJ (1913) *The fitness of the environment*. Macmillan, New York
- Hill AV (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J Physiol* 40:vi–vii
- Hill AV (1913) The combinations of haemoglobin with oxygen and with carbon monoxide. I. *Biochem J* 7(5):471
- Hill T (1985) *Cooperativity theory in biochemistry*
- Hunter CA, Anderson HL (2009) What is cooperativity? *Angew Chem Int Ed* 48(41):7488–7499
- Kaczanowska K, Harel M, Radić Z, Changeux J-P, Finn M, Taylor P (2014) Structural basis for cooperative interactions of substituted 2-aminopyrimidines with the acetylcholine binding protein. *Proc Natl Acad Sci* 111(29):10749–10754
- Martini JWR (2014) Excursions in the theory of ligand binding
- Martini JWR, Habeck M (2015) Comparison of the kinetics of different markov models for ligand binding under varying conditions. *J Chem Phys* 142(9):094104
- Martini JWR, Ullmann GM (2013) A mathematical view on the decoupled sites representation. *J Math Biol* 66(3):477–503
- Martini JWR, Habeck M, Schlather M (2013a) A derivation of the grand canonical partition function for systems with a finite number of binding sites using a Markov chain model for the dynamics of single molecules. *J Math Chem*
- Martini JWR, Schlather M, Ullmann GM (2013b) The meaning of the decoupled sites representation in terms of statistical mechanics and stochastics. *MATCH Commun Math Comput Chem* 70(3):829–850
- Onufriev A, Ullmann GM (2004) Decomposing complex cooperative ligand binding into simple components: connections between microscopic and macroscopic models. *J Phys Chem B* 108(30):11157–11169
- Onufriev A, Case DA, Ullmann GM (2001) A novel view of pH titration in biomolecules. *Biochemistry* 40(12):3413–3419

- Schellman JA (1975) Macromolecular binding. *Biopolymers* 14:999–1018
- Stefan M I, Le Novère N (2013, 06) Cooperative binding. *PLoS Computat Biol* 9(6):e1003106
- Ullmann RT, Ullmann GM (2011) Coupling of protonation, reduction and conformational change in azurin from *Pseudomonas aeruginosa* investigated with free energy measures of cooperativity. *J Phys Chem B* 115:10346–10359
- Weisstein EW (2002) CRC concise encyclopedia of mathematics. CRC Press, London