Bonding interactions between ligand-decorated colloidal particles
Online Supporting Information

Tine Curk*

*Institute of Physics, Chinese Academy of Sciences, 100190 Beijing, China and
Department of Chemistry, University of Maribor, Maribor, Slovenia*

Urban Bren

Department of Chemistry, University of Maribor, Maribor, Slovenia

Jure Dobnikar[†]

*Institute of Physics, Chinese Academy of Sciences, 100190 Beijing, China and
Department of Chemistry, University of Cambridge, Cambridge, UK*

I. DERIVATION OF THE ANALYTICAL MODEL

We will start from basic statistical mechanics and show that the expected value for the partition function due to bond formation of two ligand-coated particles, denoted by A and B, held at fixed distance d is given by the double exponential expression

$$\langle Q_b(d) \rangle = \sum_{i,j} e^{\tilde{n}_{\mathrm{A}} \tilde{n}_{\mathrm{B}} e^{-\beta \epsilon(d)}} \ . \tag{S1}$$

Here $\tilde{n}_{\mathrm{A}}$ and $\tilde{n}_{\mathrm{B}}$ are the mean number of ligands, respectively, on the particles A and B.

$$e^{-\beta \epsilon(d)} = \frac{1}{S_{\mathrm{A}} S_{\mathrm{B}}} \int_{S_{\mathrm{A}}} d\mathbf{r}_1 \int_{S_{\mathrm{B}}} d\mathbf{r}_2 \, e^{-\beta \Delta G(\mathbf{r}_1, \mathbf{r}_2)} \tag{S2}$$

is the effective bond formation free energy integrated over both particle surfaces S_{A} and S_{B}, keeping the inter-particle distance d fixed, and $\beta \equiv 1/(k_{\mathrm{B}}T)$ the inverse temperature. The bond formation free energy

$$\Delta G(\mathbf{r}_1, \mathbf{r}_2) = \Delta G_0 + \Delta G^{cnf}(\mathbf{r}_1, \mathbf{r}_2) \tag{S3}$$

is a sum of ΔG_0, the ligand-ligand hybridisation free energy in solution, and $\Delta G_{\mathrm{cnf}}(\mathbf{r}_1, \mathbf{r}_2)$, the configuration cost of forming the bond due to the two ligands being grafted to the particle's surface at positions $\mathbf{r}_1$ and $\mathbf{r}_2$ [1]. ΔG_0 is determined with respect to the standard concentration $\rho_0 = 1\mathrm{M}$.

In the main text we consider ligands being linked to particle's surface via an ideal polymeric linker resulting in harmonic stretching penalty with spring constant k describing the stiffness of the total bridge between the two particles. An individual linker connecting a ligand to a particle has a spring constant $2k$. The partition function of the two unbound ligands is simply a Gaussian integral over the two ligand positions:

$$q_u = \left(\frac{\pi}{\beta k} \right)^3 \tag{S4}$$

assuming that ligands are ideal and are not excluded by the particles. The bound partition function is a Gaussian integral over the ligand-ligand complex and it depends on the ligand grafting point positions

$$q_b(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\rho_0 N_{\mathrm{A}}} \left(\frac{\pi}{2\beta k} \right)^{3/2} e^{\beta \frac{k}{2} |\mathbf{r}_1 - \mathbf{r}_2|^2} \ , \tag{S5}$$

---

*Electronic address: tc387@cam.ac.uk
[†]Electronic address: jd489@cam.ac.uk

with $\rho_0 = 1$M, the standard molar concentration, and N_A the Avogadro's number. In this ideal harmonic case the configuration free energy can be expressed analytically

$$
\begin{aligned}
\beta \Delta G_{\mathrm{cnf}}(\mathbf{r}_1, \mathbf{r}_2) &= \ln \left(\frac{q_b(\mathbf{r}_1, \mathbf{r}_2)}{q_u} \right) \\
&= \frac{\beta k}{2} |\mathbf{r}_1 - \mathbf{r}_2|^2 + \ln \left[\rho_0 N_A \left(\frac{2\pi}{\beta k} \right)^{3/2} \right] ,
\end{aligned} \tag{S6}
$$

with the first term on the right the linker stretching penalty and the second term the unbound ligands free energy reference. We stress, however, that the following derivation is general and does not rely on the specific form of ΔG_{cnf}.

As we will demonstrate below, the double exponential expression (S1) for the bound partition function is exact when ligand positions on the particle's surface are uncorrelated and the number of ligands is Poisson distributed. The expression is also exact in the limit of weak bonds and holds for both mobile or immobile ligands. Furthermore, the formula generalises to multiple different ligand types

$$
\langle Q_b(d) \rangle = \sum_{i,j} e^{n_{A,i} n_{B,j} e^{-\beta \epsilon_{ij}(d)}} , \tag{S7}
$$

with $n_{A,i}$ and $n_{B,j}$ being the number of ligands of type i and j on the first and second particle and $\epsilon_{ij}(d)$ the interaction matrix between all ligand types i, j .

Moreover, as we shall see below, the same form of the expression also applies to multimeric complexes. For example trimeric complex contribution to the partition function, with two ligands from particle A and one from particle B, is

$$
\langle Q_{b,3}(d) \rangle = \sum_{i,j,k} \exp \left[n_{A,i} n_{A,j} n_{B,k} e^{-\beta \epsilon_{2,ij}^{1,k}(d)} \right] . \tag{S8}
$$

The sum on the right hand side takes into account all three-legged spiders with two "legs" on the first particle and one leg on the second. Tensor-like notation is employed where the subscript refers to ligands on the first particle and superscript to the ligands on the second particle. The trimeric complex formation free energy $\boldsymbol{\epsilon}_2^1(d)$ is a third order tensor with components $\epsilon_{2,ij}^{1,k}$.

A. Mobile vs. immobile ligands

Thermodynamic properties of a statistical ensemble do not depend on the value of the diffusion constant as long as it remains nonzero; $D > 0$. However, the ensemble average of the pair interaction remains the same for both mobile and immobile ligands if the ensemble properties of ligand distribution on isolated particles are the same for both cases [2]

$$
\langle Q_b^{\mathrm{mob}}(d) \rangle = \langle Q_b^{\mathrm{fixed}}(d) \rangle . \tag{S9}
$$

This argument assumes ergodicity, the time average yields the same result as the ensemble average. A particle with immobile ligands can be viewed as a particle with mobile ligands sampled at a specific moment in time. We therefore provide derivation of the mean interaction only for mobile binders, relying on the above argument stating that the result is equally applicable to quenched ligands. Below we provide derivation of bound partition function for mobile ligands, but Eq.(S9) ensures that the result is equally applicable to immobile ligands.

B. Poisson distributed ligands

We consider an ensemble of particle pairs separated by a distance d. Ligands of different types are grafted to the two particles, as described above the first particle has $n_{A,i}$ ligands of type A and the second particle $n_{B,j}$ ligands of type j. We use the notation where prime indices j' refer to specific ligands while

bare indices j refer to ligand types. The number of ligands per particle is Poisson distributed, i.e. the probability that the particle holds n_i ligands is

$$p(n_i) = e^{-\tilde{n}_i} \frac{\tilde{n}_i^{n_i}}{n_i!} \, , \tag{S10}$$

with $\tilde{n}_i$ the mean number of ligands of type i on the particle. The ligands are assumed to be mobile with no correlations in their spatial positions when the particles are isolated; we neglect ligand excluded volume and assume the ligands behave as an 2D ideal gas confined to the particle's surface.

1. Single ligand type

Let us first solve the problem in the case of a single ligand type. The number of possible distinct linking arrangements (the density of states) for a given number of links λ is

$$\Omega(\lambda) = \binom{n_A}{\lambda}\binom{n_B}{\lambda}\lambda! \, , \tag{S11}$$

because we need to choose λ bonds out of n_A ligands , λ bonds out of n_B ligands and there are $\lambda!$ ways of binding the chosen ligands together [3, 4].

The canonical partition function for a given number of ligands per particle is a sum over all possible number of bonds

$$Q_b(n_A, n_B, \epsilon | d) = \sum_{\lambda=0}^{n_A} \Omega(\lambda) e^{-\beta \epsilon(d)\lambda} \tag{S12}$$

with $\epsilon(d)$ given by Eq. (S2) and the distance between particles d is kept fixed. Note that the total number of terms in the sum is $\min[n_A, n_B]$. The density of states $\Omega(\lambda)$, however, is always zero when $\lambda > n_A$ or $\lambda > n_B$. Therefore, either n_A, n_B or ∞ can be used as the maximum number of terms in the above sum.

For en ensemble of particles with Poisson distributed ligands, Eq (S10), the average bound partition function

$$\langle Q_b \rangle^{\mathrm{P}}(\tilde{n}_A, \tilde{n}_B, \epsilon | d) = \langle Q_b(n_A, n_B, \epsilon | d) \rangle_{n_A, n_B} \tag{S13}$$

is a Poisson average over the ligand numbers, denoted by superscript $^{\mathrm{P}}$. The Poisson averaged partition function $\langle Q_b \rangle^{\mathrm{P}}(\tilde{n}_A, \tilde{n}_B, \epsilon | d)$ is essentially a grand partition function where ligand numbers are allowed to fluctuate and their mean values are determined by the chemical potential $\mu_A = k_B T \ln(n_A)$. Applying the Poisson distribution to the canonical partition (S12) we get

$$\langle Q_b \rangle^{\mathrm{P}}(\tilde{n}_A, \tilde{n}_B, \epsilon | d) = \sum_{n_A=0}^{\infty} e^{-\tilde{n}_A} \frac{\tilde{n}_A^{n_A}}{n_A!} \sum_{n_B=0}^{\infty} e^{-\tilde{n}_B} \frac{\tilde{n}_B^{n_B}}{n_B!} \sum_{\lambda=0}^{\infty} \Omega(\lambda) e^{-\beta \epsilon(d)\lambda} \, .$$

The sum over the λ bonds was extended to infinity such that all three sums are independent and can be rearranged. Inserting the density of states (S11) and altering the summation order

$$
\begin{aligned}
\langle Q_b \rangle^{\mathrm{P}}(\tilde{n}_A, \tilde{n}_B, \epsilon | d) &= \sum_{\lambda=0}^{\infty} \frac{1}{\lambda!} e^{-\beta \epsilon(d)\lambda} \sum_{n_A=0}^{\infty} e^{-\tilde{n}_A} \frac{\tilde{n}_A^{n_A}}{(n_A - \lambda)!} \sum_{n_B=0}^{\infty} e^{-\tilde{n}_B} \frac{\tilde{n}_B^{n_B}}{(n_B - \lambda)!} \\
&= \sum_{\lambda=0}^{\infty} \frac{1}{\lambda!} e^{-\beta \epsilon(d)\lambda} (\tilde{n}_A \tilde{n}_B)^{\lambda} \, ,
\end{aligned}
\tag{S14}
$$

where the last two sums in the first row are simply a Taylor expansion of an exponential function (by introducing a new variable: $x = n - \lambda$) and evaluate to $\tilde{n}_A^{\lambda}$ and $\tilde{n}_B^{\lambda}$, respectively. The final result follows directly since the above equation is again a Taylor expansion of en exponential function

$$\langle Q_b \rangle^{\mathrm{P}}(\tilde{n}_A, \tilde{n}_B, \epsilon | d) = \exp\left(\tilde{n}_A \tilde{n}_B e^{-\beta \epsilon(d)} \right) \, . \tag{S15}$$

3

The derivation involving an arbitrary number of distinct ligands is somewhat tedious, but otherwise follows exactly the same procedure as the above derivation for single ligand type. Let $\mathbf{n}_A$ and $\mathbf{n}_B$ denote vectors of ligand numbers on the A and B particles, e.g. $\mathbf{n}_A = [n_{A,1}, n_{A,2}...]$. The probability that particle A holds a vector $\mathbf{n}_A$ of ligands is a product of Poisson distributions

$$p(\mathbf{n}_A) = \prod_i p(n_{A,i}) \ , \tag{S16}$$

with $p(n_{A,i})$ given by Eq. (S10) applied to the particle A, and likewise for particle B.

The number of possible distinct linking arrangements between particles A and B is denoted by $\Omega(\boldsymbol{\lambda})$ and it depends on the given number of formed links between ij ligand types λ_{ij}. Note that $\boldsymbol{\lambda}$ is a matrix. The canonical partition function is a sum over all possible numbers of λ_{ij} bonds with binding free energy ϵ_{ij}

$$Q_b(\mathbf{n}_A, \mathbf{n}_B, \boldsymbol{\epsilon}|d) = \sum_{\boldsymbol{\lambda}} \Omega(\boldsymbol{\lambda}) e^{-\sum_{ij} \beta \epsilon_{ij} \lambda_{ij}} \ . \tag{S17}$$

where the sum represents a nested sum over all distinct receptor ligand pairs ij

$$\sum_{\boldsymbol{\lambda}} [\cdot] = \sum_{\lambda_{11}=0}^{\infty} \sum_{\lambda_{12}=0}^{\infty} \cdots \sum_{\lambda_{21}=0}^{\infty} \sum_{\lambda_{22}=0}^{\infty} \cdots [\cdot] \tag{S18}$$

to account for all possible states of distinct bonding arrangements. We note that the maximum term in each sum is set to inifinity, this choice was made for later convenience, as we will see below the density of states Ω is such that all terms where $\sum_j \lambda_{ij} > n_i$ are automatically zero.

The number of states $\Omega(\boldsymbol{\lambda})$ is given by a product of multinomial distributions because for each ligand type j on particle A we need to choose how many will bind to different ligand types i on particle B. Equivalently we need to choose among $n_{A,i}$ ligands how many will get attached to given ligand types $n_{B,j}$, and repeat for each ligand type. Finally, we need to bind ligands and receptors together and there are $\prod_{ij} \lambda_{ij}!$ ways of connecting them. Therefore, the density of states is

$$\Omega(\boldsymbol{\lambda}) = \prod_j \left(\frac{n_{B,j}!}{\prod_i (\lambda_{ij}!) \, (n_{B,j} - \sum_i \lambda_{ij})!} \right) \prod_i \left(\frac{n_{A,i}!}{\prod_j (\lambda_{ij}!) \, (n_{A,i} - \sum_j \lambda_{ij})!} \right) \prod_{ij} \lambda_{ij}! \ . \tag{S19}$$

The Poisson averaged partition function is a function of the mean ligand vectors $\tilde{\mathbf{n}}_A$ and $\tilde{\mathbf{n}}_B$

$$\langle Q_n \rangle^P (\tilde{\mathbf{n}}_A, \tilde{\mathbf{n}}_B, \boldsymbol{\epsilon}|d) = \sum_{\mathbf{n}_A=0}^{\infty} \sum_{\mathbf{n}_B=0}^{\infty} p(\mathbf{n}_A) p(\mathbf{n}_B) Q_b(\mathbf{n}_A, \mathbf{n}_B, \boldsymbol{\epsilon}|d) \tag{S20}$$

with the Poisson distributions $p(\mathbf{n}_A)$ and $p(\mathbf{n}_B)$ given by Eq. (S16) and the sum $\sum_{\mathbf{n}=0}^{\infty}[\cdot] = \prod_j \sum_{n_j=0}^{\infty}[\cdot]$ represents nested sum over all ligand types j. Inserting Eqs. (S16), (S17) and (S19) into the above equation (S20) we rearrange the summation order and obtain

$$\langle Q_b \rangle^P (\tilde{\mathbf{n}}_A, \tilde{\mathbf{n}}_B, \boldsymbol{\epsilon}|d) = \sum_{\boldsymbol{\lambda}=0}^{\infty} e^{-\sum_{ij} \beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\prod_{i,j} \lambda_{ij}!} \sum_{\mathbf{n}_A=0}^{\infty} \sum_{\mathbf{n}_B=0}^{\infty} \left[\prod_i \left(\frac{e^{-\tilde{n}_{A,i}} (\tilde{n}_{A,i})^{n_{A,i}}}{(n_{A,i} - \sum_j \lambda_{ij})!} \right) \prod_j \left(\frac{e^{-\tilde{n}_{B,j}} (\tilde{n}_{B,j})^{n_{B,j}}}{(n_{B,j} - \sum_i \lambda_{ij})!} \right) \right] \ .$$

The summation and product order of the two innermost sums can be rearranged $\sum_{\mathbf{n}} \prod_i [\cdot] = \prod_i \sum_{n_i} [\cdot]$, obtaining a product of individual sums, each sum being a Taylor expansion evaluating to $\tilde{n}_i^{\sum_j \lambda_{ij}}$. Hence, we obtain

$$\cdots = \sum_{\boldsymbol{\lambda}=0}^{\infty} e^{-\sum_{ij} \beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\prod_{i,j} \lambda_{ij}!} \prod_{ij} (\tilde{n}_{A,i} \tilde{n}_{B,j})^{\lambda_{ij}} \ .$$

Combining the products and swapping the summation and product order we get

$$\cdots = \prod_{ij} \sum_{\lambda_{ij}=0}^{\infty} e^{-\sum_{ij} \beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\lambda_{ij}!} (\tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j})^{\lambda_{ij}} \ , \tag{S21}$$

which is simply a Taylor expansion for exponential function and the final result follows

$$\langle Q_b \rangle^{\mathrm{P}} (\tilde{\mathbf{n}}_{\mathrm{A}}, \tilde{\mathbf{n}}_{\mathrm{B}}, \boldsymbol{\epsilon} | d) = \exp \left[\sum_{i,j} \tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j} e^{-\beta \epsilon_{ij}} \right] \ . \tag{S22}$$

The interaction free energy between the two particles due to bond formation is a sum over all ligand types

$$\begin{aligned} \beta F_b(d) &= -\ln(\langle Q_b \rangle^{\mathrm{P}} (\tilde{\mathbf{n}}_{\mathrm{A}}, \tilde{\mathbf{n}}_{\mathrm{B}}, \boldsymbol{\epsilon} | d)) \\ &= -\sum_{i,j} \tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j} e^{-\beta \epsilon_{ij}} \\ &= \tilde{\mathbf{n}}_{\mathrm{A}} e^{-\beta \boldsymbol{\epsilon}} \tilde{\mathbf{n}}_{\mathrm{B}}^T \ . \end{aligned} \tag{S23}$$

In the last row the result is written in standard matrix notation with the matrix $e^{-\beta \boldsymbol{\epsilon}}$ the element-wise exponential of $\boldsymbol{\epsilon}$ and $\tilde{\mathbf{n}}_{\mathrm{B}}^T$ is the transpose of the vector $\tilde{\mathbf{n}}_{\mathrm{B}}$.

Notice that each term in Eq. (S21) is proportional to the probability $p(\lambda_{ij})$ of forming λ_{ij} bonds, hence

$$p(\lambda_{ij}) = \frac{e^{-\beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\lambda_{ij}!} (\tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j})^{\lambda_{ij}}}{\sum_{\lambda_{ij}=0}^{\infty} e^{-\beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\lambda_{ij}!} (\tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j})^{\lambda_{ij}}} \ , \tag{S24}$$

independent of any other bonds among different i's and j's. The average number of formed bonds between ligand types i and j can, therefore, be obtained

$$\begin{aligned} \langle \lambda_{ij} \rangle &= \sum_{\lambda_{ij}} \lambda_{ij} p(\lambda_{ij}) \\ &= \frac{\sum_{\lambda_{ij}=0}^{\infty} e^{-\beta \epsilon_{ij} \lambda_{ij}} \frac{1}{(\lambda_{ij}-1)!} (\tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j})^{\lambda_{ij}}}{\sum_{\lambda_{ij}=0}^{\infty} e^{-\beta \epsilon_{ij} \lambda_{ij}} \frac{1}{\lambda_{ij}!} (\tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j})^{\lambda_{ij}}} \\ &= \tilde{n}_{\mathrm{A},i} \tilde{n}_{\mathrm{B},j} e^{-\beta \epsilon_{ij}} \ . \end{aligned} \tag{S25}$$

Hence, the average of the total number of bonds is exactly equal to the negative interaction free energy, Eq. (S23),

$$\langle \lambda^{\mathrm{tot}} \rangle = \sum_{ij} \langle \lambda_{ij} \rangle = -\beta F_b(d) \ . \tag{S26}$$

C. Weak interaction limit

In the weak interaction limit the partition function of two spherical particles separated by a distance d can also be derived analyticaly. As described above the first particle has $n_{\mathrm{A},i}$ ligands of type i and the second particle $n_{\mathrm{B},j}$ ligands of type j. We use the notation where prime indices j' refer to specific ligands while bare indices jr efer to ligand types. The probability that two specific ligands are bound is

$$p_{i'j'} = p_{i'} p_{j'} e^{-\beta \Delta G_{i'j'}(\mathbf{r}_{i'}, \mathbf{r}_{j'})} \ , \tag{S27}$$

with $p_{i'}$ and $p_{j'}$ the probability that, respectively, ligands i' and j' are free. $\Delta G_{i'j'}(\mathbf{r}_{i'}, \mathbf{r}_{j'})$ is the hybridisation free energy, Eq. (S3), which depends on the exact positions of the ligand grafting points, $\mathbf{r}_{i'}$ and $\mathbf{r}_{j'}$. This form implies that the probabilities of i' and j' to be free are not correlated.

The weak interaction limit is obtained when each and every ligand has a low probability of being bound: $p_{i'} \to 1$ for all $p_{i'}$. In this limit the valence limited nature of ligand-ligand interactions is not

important; any bond between two ligands can be formed irrespective of whether the same two ligands have some probability of already being bound. The partition function taking into account all possible states is a product over all links:

$$Q_b^{\mathrm{w}}(\mathbf{r}^{\mathbf{n_A}}, \mathbf{r}^{\mathbf{n_B}}) = \prod_{i',j'}(1 + e^{-\beta\Delta G(\mathbf{r}_{i'}, \mathbf{r}_{j'})}) \,, \tag{S28}$$

because each link between ligands i' and j' can be either present or not, irrespective of other links. Since all ligands are mostly unbound, $p_{i'} \to 1$, any correlations between ligand positions also become unimportant. For mobile ligands we must integrate Eq. (S28) over all possible independent links keeping the interparticle distance d fixed

$$Q_b^{\mathrm{w}}(d) = \frac{1}{S_1 S_2} \int_{S_1} d\mathbf{r}^{\mathbf{n_A}} \int_{S_2} d\mathbf{r}^{\mathbf{n_B}} \prod_{i',j'}(1 + e^{-\beta\Delta G(\mathbf{r}_{i'}, \mathbf{r}_{j'})}) = \prod_{i,j}\left(1 + e^{-\beta\epsilon_{ij}(d)}\right)^{n_i n_j} . \tag{S29}$$

We have used Eq. (S2) to obtain $\epsilon_{ij}(d)$ and the fact that upon integration any link between ligand types i and j will result in the same contribution. Finally, since all ligands are mostly unbound: $p_{ij} \sim 0$, the Boltzmann factor is small: $e^{-\beta\epsilon_{ij}(d)} = \frac{p_{ij}}{p_i p_j} < 1$ and the we can use the first order approximation to the exponential function obtaining

$$Q_b^{\mathrm{w}}(d) = \exp\left(\sum_{i,j} n_{\mathrm{A},i} n_{\mathrm{B},j} e^{-\beta\epsilon_{ij}(d)}\right) . \tag{S30}$$

Exactly the same expression is in the case of Poisson distributed ligands, Eq. (S22).

D.   Comparison with canonical theory

In Figures S1 and S2 we compare between valence limited interaction theory (VLIT) by Angioletti et. al [1], which self-consistently solves the canonical problem, Eqs. (S17,S19), and our double exponential solution (S22). For simplicity, we limit ourselves to a system of 2 spherical particles, each particle holding $n = n_{\mathrm{A}} = n_{\mathrm{B}}$ mobile binders with uncorrelated positions on the particle's surface. We use $n = \tilde{n}$ for comparison between canonical [1] and Poisson averaged free energies: $\beta F_b = \tilde{n}^2 e^{-\beta\epsilon}$. $e^{-\beta\epsilon} = \frac{p_{ij}}{p_i p_j}$ denotes the effective free energy of hybridising a single pair, p_{ij} is the probability that binders form a link (bond) and p_i, p_j are the probabilities that the two binders i and j are free. We observe that free energies converge when the number of ligands is large or when the individual bonds are weak.

In Figure S3 we shows that Poisson averaging the canonical solution (VLIT theory) gives us the same result as the double exponential.
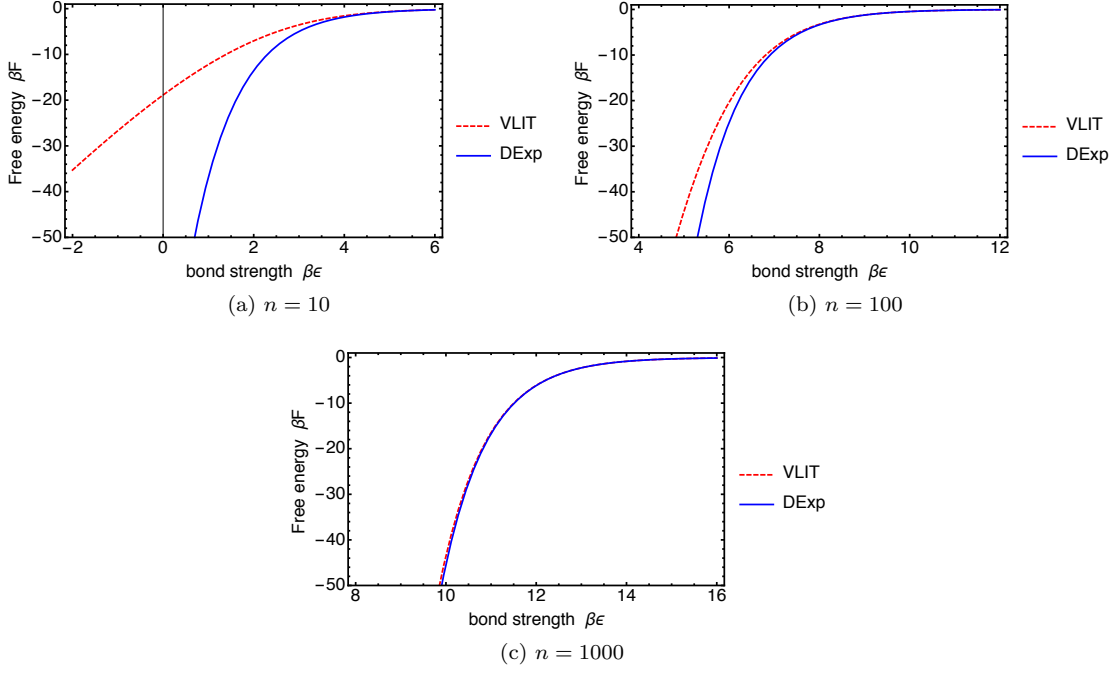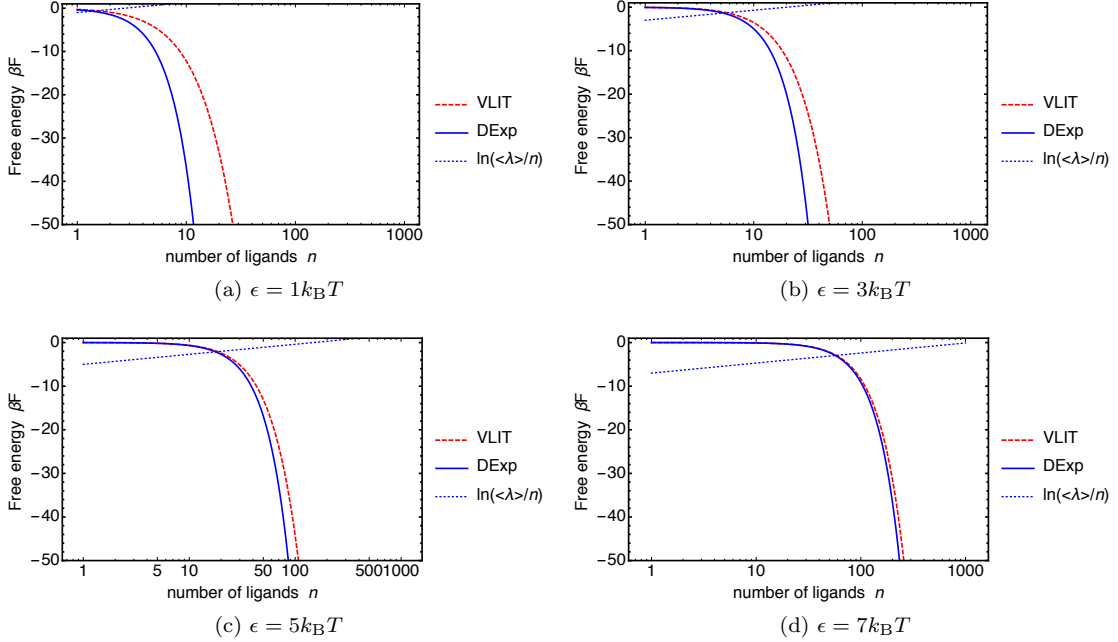
(a) $n = 10$

(b) $n = 100$

(c) $n = 1000$

FIG. S1: Comparison between canonical (VLIT) theory [1] and grand canonical (DExp) expression (S22) for the binding free energy between 2 particles each with n binders. As expected, at large n or weak bonds both expressions converge.

(a) $\epsilon = 1k_\mathrm{B}T$

(b) $\epsilon = 3k_\mathrm{B}T$

(c) $\epsilon = 5k_\mathrm{B}T$

(d) $\epsilon = 7k_\mathrm{B}T$

FIG. S2: Comparison between canonical (VLIT) theory [1] and grand canonical (DExp) expression (S22) for the binding free energy between 2 particles each holding n binders. We also show the logarithm of the expected fraction formed of bonds: $\ln(\langle\lambda\rangle/n)$, shown as dotted blue line. When the fraction of formed bonds is low, both the canonical and grand-canonical treatments converge.
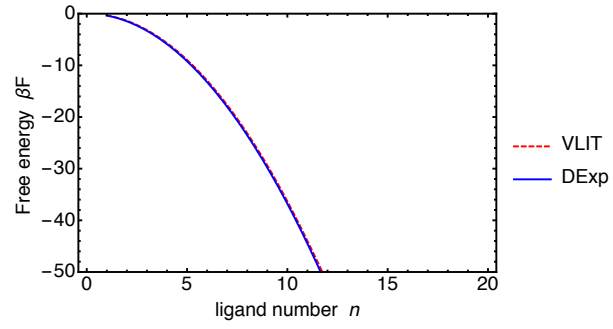
FIG. S3: Consistency check: Poisson averaging of the canonical (VLIT) theory [1] results in the same free energy as obtained the double exponential expression (DExp) (S22). The small deviation arises due to the approximations in the VLIT theory. Parameters: $\beta\epsilon = 1$

II. MULTIMERIC COMPLEXES GENERALIZATION

Recently, Mognetti et. al. [5, 6] extended the valence limited theory to multimeric complexes when more than two ligands can participate in forming a joint link. Our theory generalises to such multimeric complexes. For mathematical clarity we shall use tensor notation for dealing with multimeric complexes. The subscripts shall refer to the particle A and the superscripts to particle B. For dimeric links considered above a given configuration of links between the two colloids is described by a second order tensor $\boldsymbol{\lambda}_1^1$ where the sub- and super-scripts 1,1 mean that each bond is formed by linking 1 ligand from particle A and one ligand from particle B. En individual element of the tensor is $\lambda_{1,j}^{1,i}$. For a trimeric complex forming connection between the two particles there are two possible third order tensors: $\boldsymbol{\lambda}_2^1$ and $\boldsymbol{\lambda}_1^2$, because there can be two linkers from particle A and one from B, or vice versa. In this notation the topology of the complex, e.g. $\boldsymbol{\lambda}_2^0$ is a loop formed by two ligands on particle A, $\boldsymbol{\lambda}_1^1$ a bridge between the two particles, and $\boldsymbol{\lambda}_2^3$ captures all "spiders" complexes with two legs on the particle A and three legs on particle B. The same tensor notation will also be used for the ligand complexation free energy $\epsilon(d)$. For brevity we shall in the following omit the explicit notation of ϵ dependance on the inter-particle distance d.

A. Trimeric complexes

Let us first consider trimeric complexes where two ligands on particle A and one ligand on particle B form a trimeric complex. $\lambda_{2,ij}^{1,k}$ denotes the number of trimeric complexes between ligand types i and j on particle A and type k on particle B. We must chose $\lambda_{2,ij}^{1,k}$ linkers of type i on particle A, again $\lambda_{2,ij}^{1,k}$ linker of type j from particle A, and $\lambda_{2,ij}^{1,k}$ linkers of type k from particle B. Moreover there are $(\lambda_{2,ij}^{1,k}!)^2$ ways of connecting the chosen linkers together. Hence, the number of possible states of forming such a trimeric complex is

$$
\begin{aligned}
\Omega(\lambda_{2,ij}^{1,k}) &= \binom{n_{A,i}}{\lambda_{2,ij}^{1,k}} \binom{n_{A,j}}{\lambda_{2,ij}^{1,k}} \binom{n_{B,k}}{\lambda_{2,ij}^{1,k}} (\lambda_{2,ij}^{1,k}!)^2 \text{ , if } i \neq j \\
&= \binom{n_{i,A}}{\lambda_{2,ij}^{1,k}} \binom{n_{j,A} - \lambda_{2,ij}^{1,k}}{\lambda_{2,ij}^{1,k}} \binom{n_{B,k}}{\lambda_{2,ij}^{1,k}} (\lambda_{2,ij}^{1,k}!)^2 \text{ , if } i = j \text{ .}
\end{aligned}
\tag{S31}
$$

The cases where both ligands forming the complex are of the same type or not must be distinguished. $n_{i,A}$ and $n_{k,B}$ are the total number of ligands of type i and k, respectively, on particles A and B. The above number of states only considers $\boldsymbol{\lambda}_2^1$ trimeric complexes. Different topologies of complexes can, however, coexist.

Considering at most trimeric complexes and, to avoid tedious indexing, only one ligand type per particle, $i = j = k = 1$ (note this does not imply that the ligands on particle A and B are the same). The number of states when forming λ_1^1 dimeric bonds (bridges) and λ_2^1, λ_1^2 trimeric complexes (three legged spiders) is obtained by choosing λ_1^1, λ_2^1 and λ_1^2 ligands out of the total available n_A and n_B ligands

$$
\begin{aligned}
\Omega(\lambda_1^1, \lambda_2^1, \lambda_1^2) &= \binom{n_A}{\lambda_1^1} \binom{n_B}{\lambda_1^1} \lambda_1^1! \binom{n_A - \lambda_1^1}{\lambda_2^1} \binom{n_A - \lambda_1^1 - \lambda_2^1}{\lambda_2^1} \binom{n_B - \lambda_1^1}{\lambda_2^1} (\lambda_2^1!)^2 \times \\
&\quad \binom{n_A - \lambda_1^1 - 2\lambda_2^1}{\lambda_1^2} \binom{n_B - \lambda_1^1 - \lambda_2^1}{\lambda_1^2} \binom{n_B - \lambda_1^1 - \lambda_2^1 - \lambda_1^2}{\lambda_1^2} (\lambda_1^2!)^2 \\
&= \frac{n_A! n_B!}{\lambda_1^1! \lambda_2^1! \lambda_1^2! (n_A - \lambda_1^1 - 2\lambda_2^1 - \lambda_1^2)! (n_B - \lambda_1^1 - \lambda_2^1 - 2\lambda_1^2)!} \text{ .}
\end{aligned}
$$

(S32)

(S33)

The expression is rather complicated, and we haven't even considered different ligand types.

The canonical partition function is a sum over all possible linking states

$$
Q_b(n_A, n_B, \epsilon_1^1, \epsilon_2^1, \epsilon_1^2 | d) = \sum_{\lambda_1^1=0}^{\infty} \sum_{\lambda_2^1=0}^{\infty} \sum_{\lambda_1^2=0}^{\infty} \Omega(\lambda_1^1, \lambda_2^1, \lambda_1^2) e^{-\beta(\epsilon_1^1 \lambda_1^1 + \epsilon_2^1 \lambda_2^1 + \epsilon_1^2 \lambda_1^2)} \text{ ,}
\tag{S34}
$$

with the multimeric complex formation free energy employing the same tensor notation as we are using for the bonds λ. The dimerisation free energy ϵ_1^1 is equivalent to what we previously called $\epsilon(d)$ in

Eq. (S12). Analogous to the dimer free energy Eq. (S2), the trimeric complex formation free energy is defined as a three-ligand integral

$$e^{-\beta\epsilon_2^1} = \frac{1}{S_A^2 S_B} \int_{S_A} d\mathbf{r}_1 \int_{S_A} d\mathbf{r}_2 \int_{S_B} d\mathbf{r}_3 e^{-\beta\Delta G(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)} , \tag{S35}$$

for the case of two ligands on particle A and one on particle B. The trimer complexation free energy

$$\Delta G(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \Delta G_{0,\mathrm{tri}} + \Delta G_{\mathrm{cnf}}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \tag{S36}$$

depends on $\Delta G_{0,\mathrm{tri}}$, the trimer formation free energy in solution, and the configurational contribution, analogous to Eq. (S3).

Poisson averaging the canonical partition function, Eq. (S34), all correlations within the density of states, Eq. (S33), disappear

$$\begin{aligned}
\langle Q_b \rangle(d) &= \sum_{n_A=0}^{\infty} e^{-\tilde{n}_A} \frac{\tilde{n}_A^{n_A}}{n_A!} \sum_{n_B=0}^{\infty} e^{-\tilde{n}_B} \frac{\tilde{n}_B^{n_B}}{n_B!} Q_b(n_A, n_B, \epsilon_1^1, \epsilon_2^1, \epsilon_1^2 | d) \\
&= \langle Q_1^1 \rangle \langle Q_2^1 \rangle \langle Q_1^2 \rangle
\end{aligned} \tag{S37}$$

and the partition function can be written as a product of three independent contributions: dimeric links and two topologies of trimeric complexes. The dimeric partition function has already been evaluated above, Eq. (S15), we are only using a slightly different (tensor) notation

$$\langle Q_1^1 \rangle = \exp\left[\tilde{n}_A \tilde{n}_B e^{-\beta\epsilon_1^1}\right] . \tag{S38}$$

The trimeric contributions evaluate to

$$\begin{aligned}
\langle Q_2^1 \rangle &= \exp\left[\tilde{n}_A^2 \tilde{n}_B e^{-\beta\epsilon_2^1}\right] , \\
\langle Q_1^2 \rangle &= \exp\left[\tilde{n}_A \tilde{n}_B^2 e^{-\beta\epsilon_1^2}\right] .
\end{aligned} \tag{S39}$$

for a single ligand type.

Finally, the generalisation to different ligand types is tedious, but straightforward, and follows the same procedure as employed for dimeric links, Eqs. (S16-S22)

$$\begin{aligned}
\langle Q_2^1 \rangle &= \exp\left[\sum_{i,j,k} \tilde{n}_{A,i} \tilde{n}_{A,j} \tilde{n}_{B,k} e^{-\beta\epsilon_{2,ij}^{1,k}}\right] , \\
\langle Q_1^2 \rangle &= \exp\left[\sum_{i,j,k} \tilde{n}_{A,i} \tilde{n}_{B,j} \tilde{n}_{B,k} e^{-\beta\epsilon_{1,i}^{2,jk}}\right] .
\end{aligned} \tag{S40}$$

The free energy due to bond formation $\beta F_b(d) = -\ln(\langle Q_b \rangle(d))$ is, therefore, a sum of individual contributions due to dimeric and trimeric complexes

$$\beta F_b(d) = -\sum_{i,j} \tilde{n}_{A,i} \tilde{n}_{B,j} e^{-\beta\epsilon_{1,i}^{1,j}} - \sum_{i,j,k} \tilde{n}_{A,i} \tilde{n}_{A,j} \tilde{n}_{B,k} e^{-\beta\epsilon_{2,ij}^{1,k}} - \sum_{i,j,k} \tilde{n}_{A,i} \tilde{n}_{B,j} \tilde{n}_{B,k} e^{-\beta\epsilon_{1,i}^{2,jk}} . \tag{S41}$$

B. Higher order complexes

Generalisation to 4$^{\text{th}}$ and higher order complexes follows an analogous treatment to the trimeric complexes discussed above. The total partition function factorizes in topologies of the complexes

$$\langle Q_b \rangle(d) = \prod_{a=1} \prod_{b=1}^{a-1} \langle Q_a^b \rangle , \tag{S42}$$

with $\langle Q_a^b \rangle$ the Poisson averaged partition function of a $(a + b)$-meric complexes with a ligands from particle A and b ligands from particle B

$$\langle Q_a^b \rangle = \exp\left[\sum_{a\cdots}^{b\cdots} \prod_a (\tilde{n}_{\mathrm{A},\cdot}) \prod_b (\tilde{n}_{\mathrm{B},\cdot}) e^{-\beta \epsilon_{a,\cdots}^{b,\cdots}}\right] \ . \tag{S43}$$

The sum $\sum_{a\cdots}^{b\cdots}$ representing $(a + b)$ nested sums and $\epsilon_{a,\cdots}^{b,\cdots}$ is an element of a $(a + b)^{\mathrm{th}}$ order tensor specifying binding free energies for all combinations of ligand types within an $(a + b)$-meric complex.

Note that we do not need to consider loops, or complexes (spiders) with all 'legs' on a the same particle, e.g. Q_2^0. These complexes are also present for isolated particles, $d \to \infty$, and cancel out when calculating the interaction free energy as the ratio of bound-to-unbound partition functions $\beta F_b(d) = -\ln\left(\frac{Q_b(d)}{Q_b(\infty)}\right)$

---

[1] Stefano Angioletti-Uberti, Patrick Varilly, Bortolo M. Mognetti, and Daan Frenkel. Mobile linkers on dna-coated colloids: Valency without patches. *Phys. Rev. Lett.*, 113:128303, Sep 2014.

[2] Tine Curk, Jure Dobnikar, and Daan Frenkel. Optimal multivalent targeting of membranes with many distinct receptors. *Proc. Natl. Acad. Sci. U.S.A.*, pages 7210–7215, 2017.

[3] Pavel I. Kitov and David R. Bundle. On the nature of the multivalency effect: A thermodynamic model. *Journal of the American Chemical Society*, 125(52):16271–16284, 2003. PMID: 14692768.

[4] Francisco J Martinez-Veracoechea and Daan Frenkel. Designing super selectivity in multivalent nano-particle binding. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27):10963–10968, 2011.

[5] Lorenzo Di Michele, Stephan J. Bachmann, Lucia Parolini, and Bortolo M. Mognetti. Communication: Free energy of ligand-receptor systems forming multimeric complexes. *The Journal of Chemical Physics*, 144(16):161104, 2016.

[6] Omar A. Amjad, Bortolo M. Mognetti, Pietro Cicuta, and Lorenzo Di Michele. Membrane adhesion through bridging by multimeric ligands. *Langmuir*, 33(5):1139–1146, 2017. PMID: 28068766.