Andrés Ortiz-Muñoz 11/12/2021 3 PM [Zoom]

Combinatorics and Stochasticity for Chemical Reaction Networks

$$\mathcal{M}f = \sum_{x\in\mathbb{N}^S} \frac{z^x}{x!}\langle z'^x e^{z'}, f\rangle = \sum_{x\in\mathbb{N}^S} (z+1)^x f_x$$

Combinatorics and Stochasticity for Chemical Reaction Networks

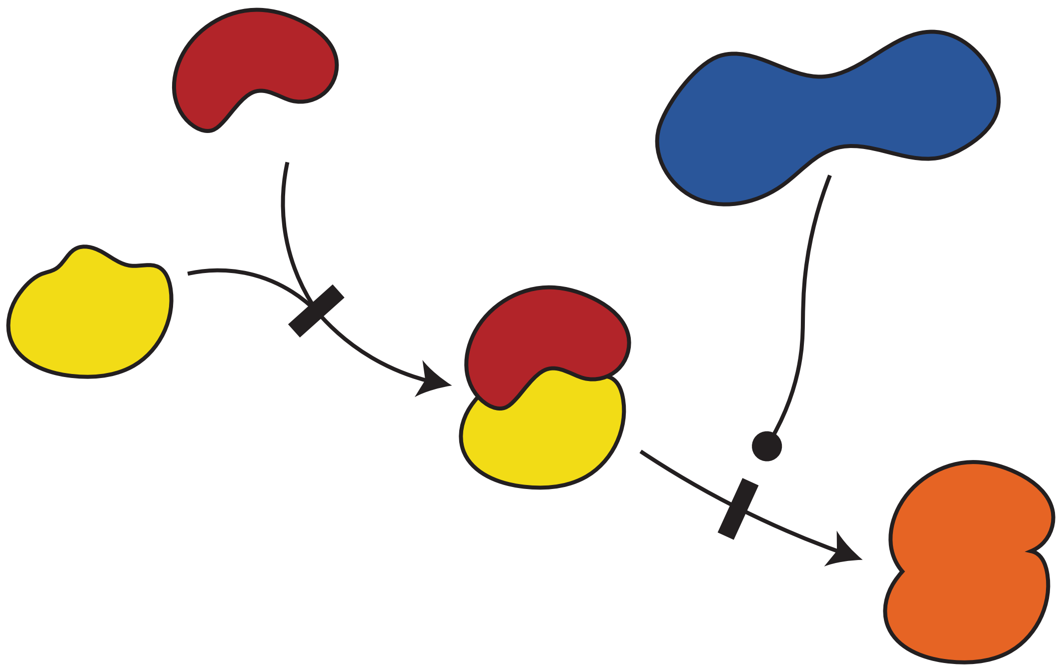
Andrés Ortiz-Muñoz
PhD Advisor: Erik Winfree

Thesis Defense
November 12, 2021

Overview

- Chemical reaction networks
- Stochasticity
- Combinatorics
- Formal Semantics

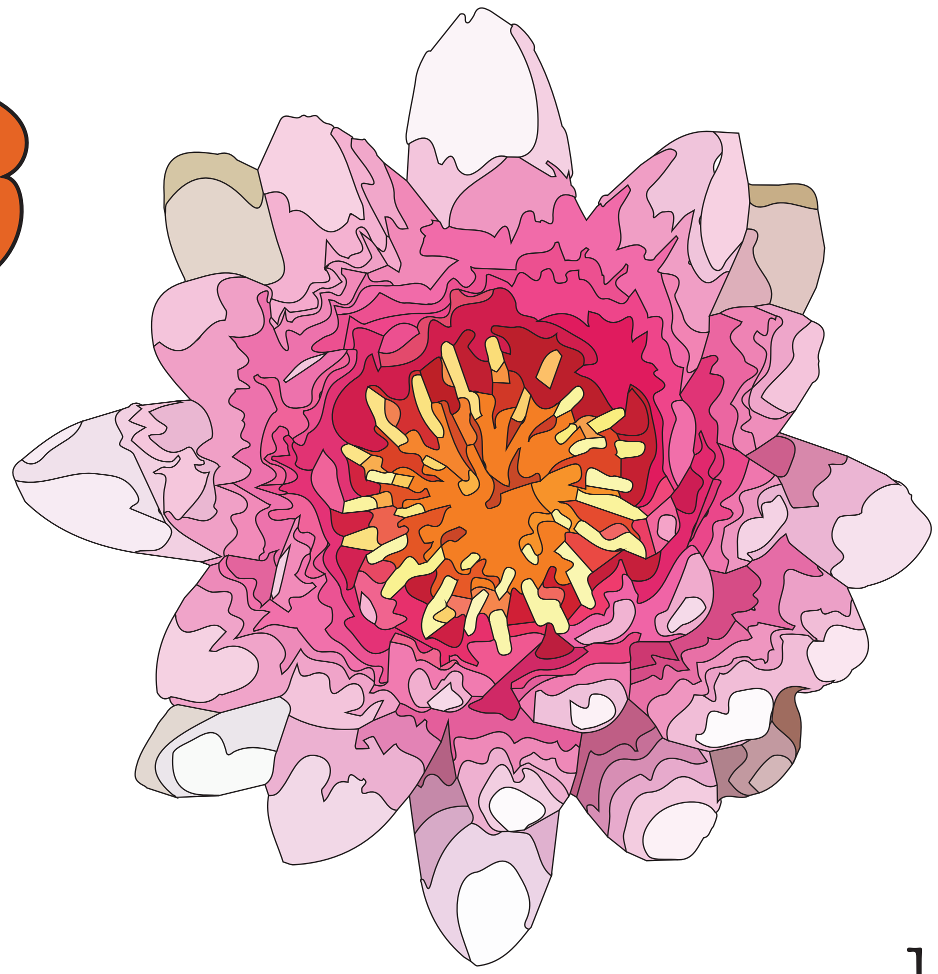
Flowers are theorems



$$\dfrac{\dfrac{A \quad B}{A \wedge B} \qquad (A \wedge B) \to C}{C}$$

$$\forall n \forall x \forall y \forall z.(n \geq 3 \to x^n + y^n \neq z^n)$$

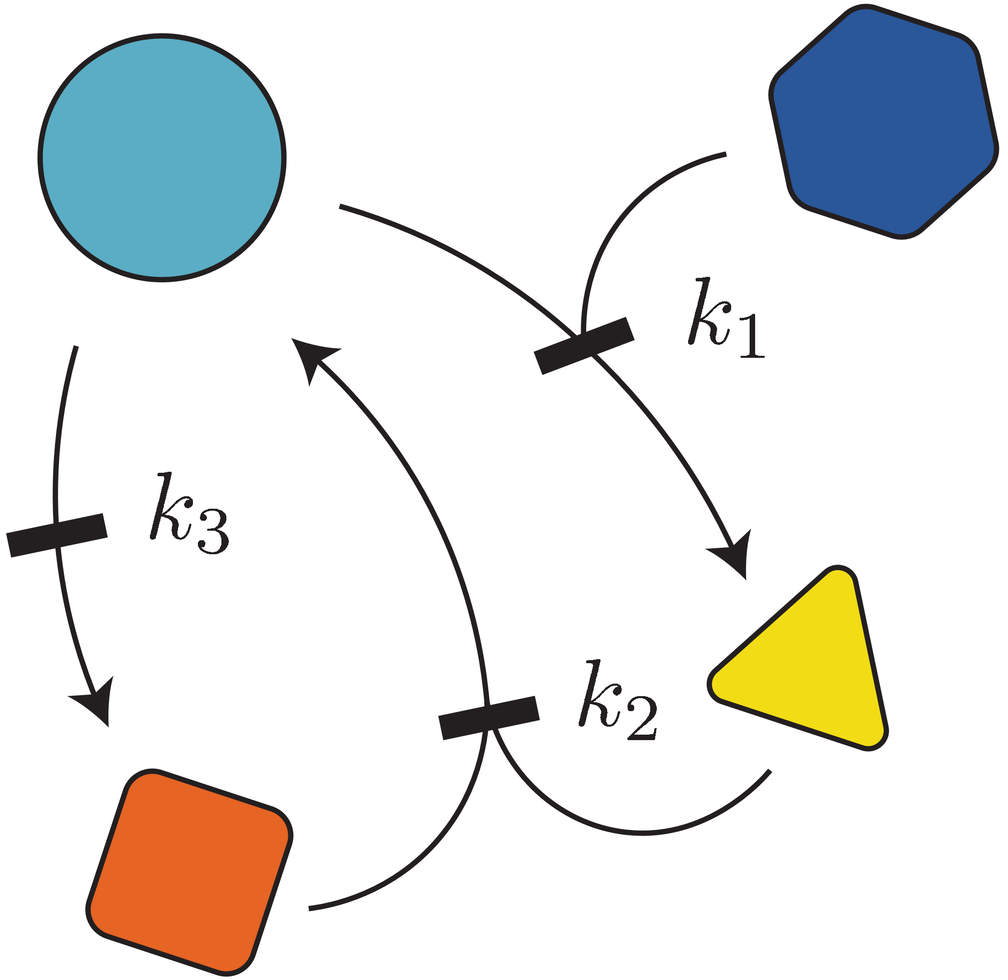
Theorems are flowers



Chemical Reaction Networks

How can we model biomolecular systems mathematically?

Chemical reaction networks are abstract models of systems with interacting species



S species

R reactions

$$\rho : R \times S \to \mathbb{N}$$

reactants

$$\pi : R \times S \to \mathbb{N}$$

products

$$k : R \to \mathbb{R}^+$$

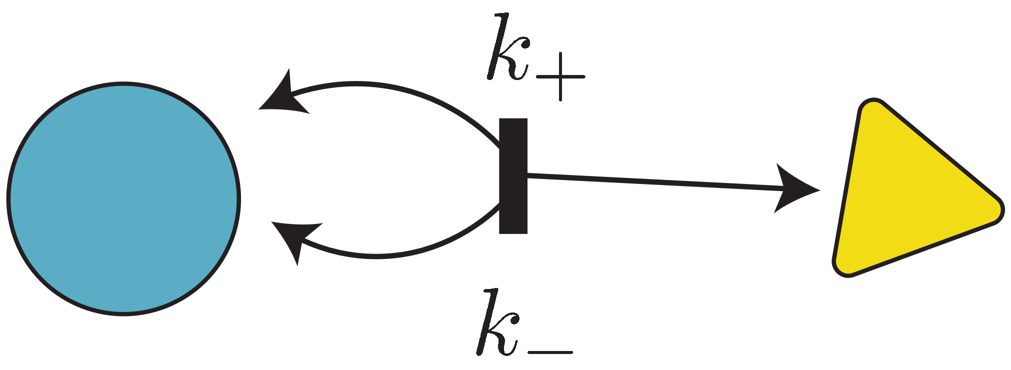
rate constants

Horn, F. and Jackson, R., 1972. General mass action kinetics. Archive for rational mechanics and analysis, 47(2), pp.81-116.

Continuous models are deterministic

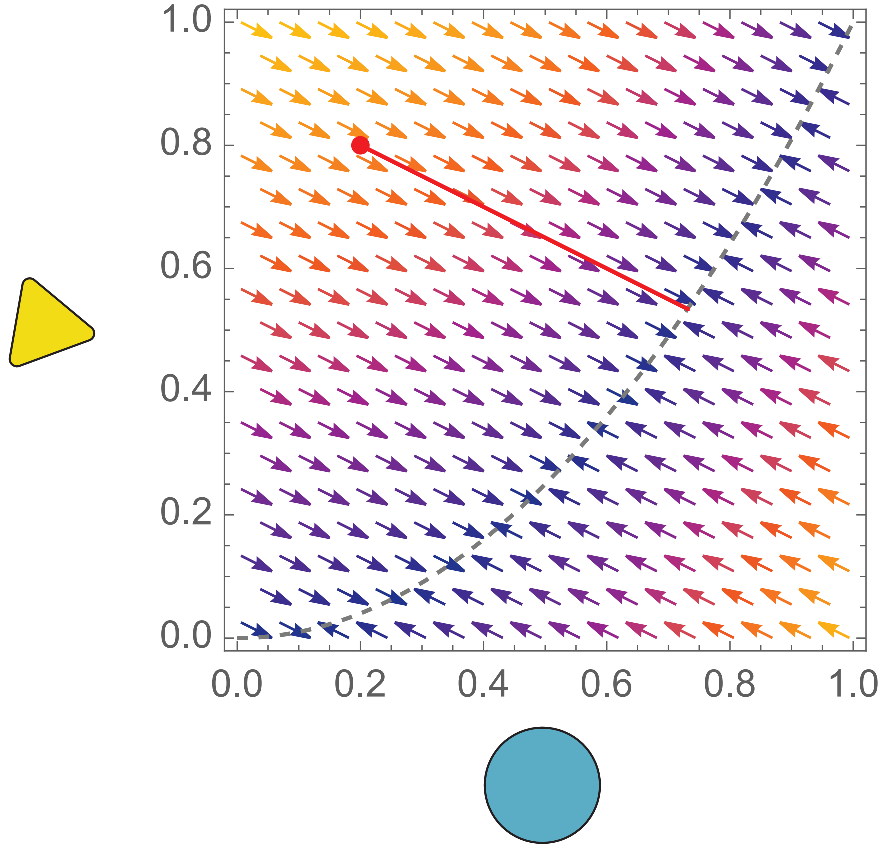
$$\frac{d[A]}{dt} = 2k_-[B] - 2k_+[A]^2$$

$$\frac{d[B]}{dt} = k_+[A]^2 - k_-[B]$$



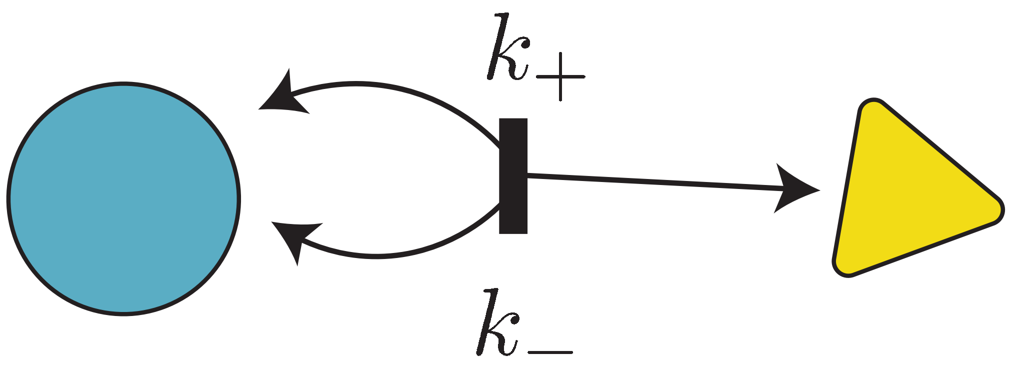
$$2A \xrightarrow{k_+} B$$

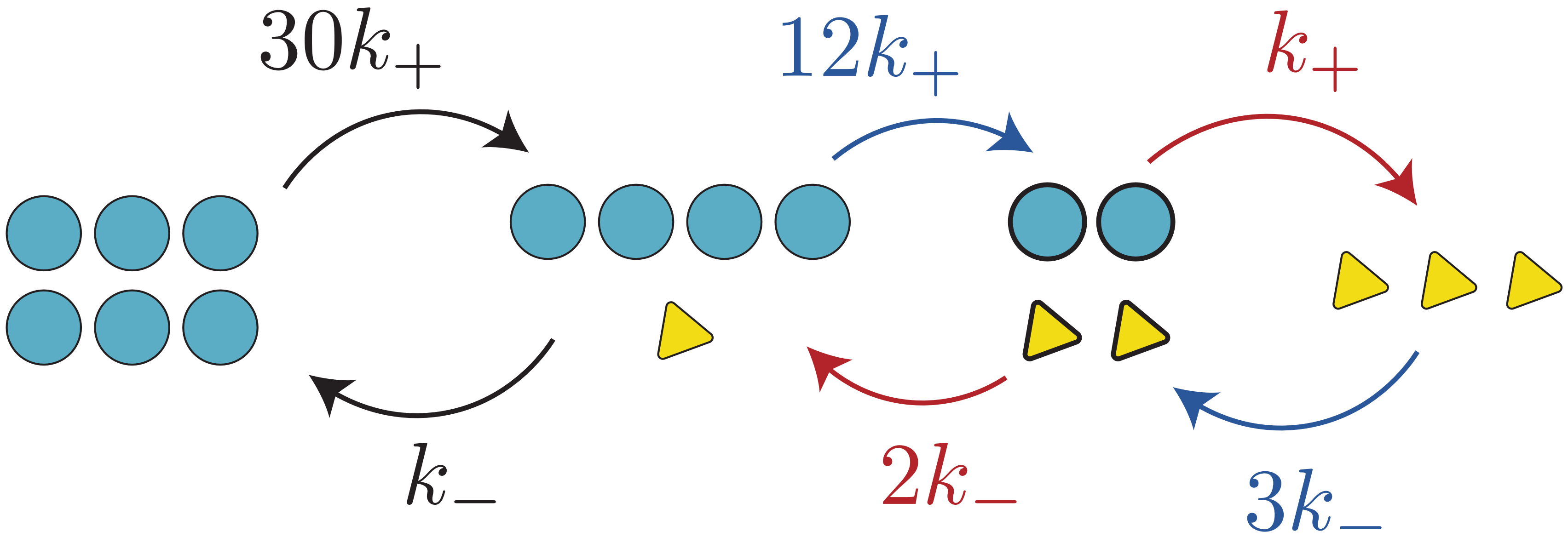
$$B \xrightarrow{k_-} 2A$$

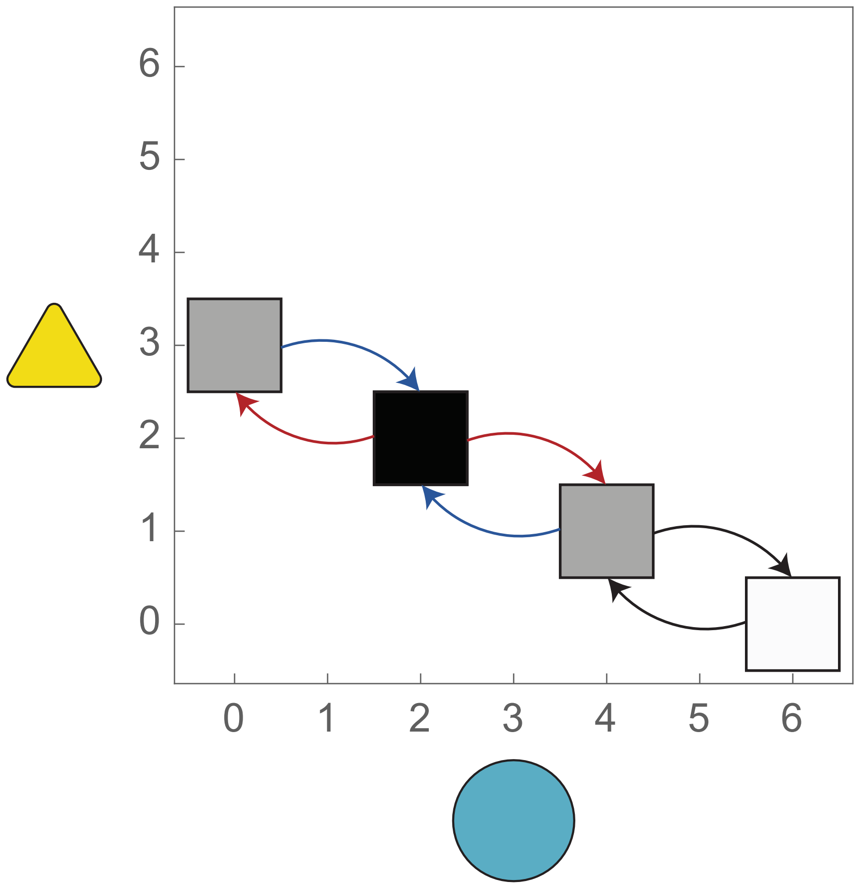


Discrete models are stochastic

$$\frac{dp_{2,2}}{dt} = 12k_+ p_{4,1} + 3k_- p_{0,3}$$
$$- k_+ p_{2,2} - 2k_- p_{2,2}$$





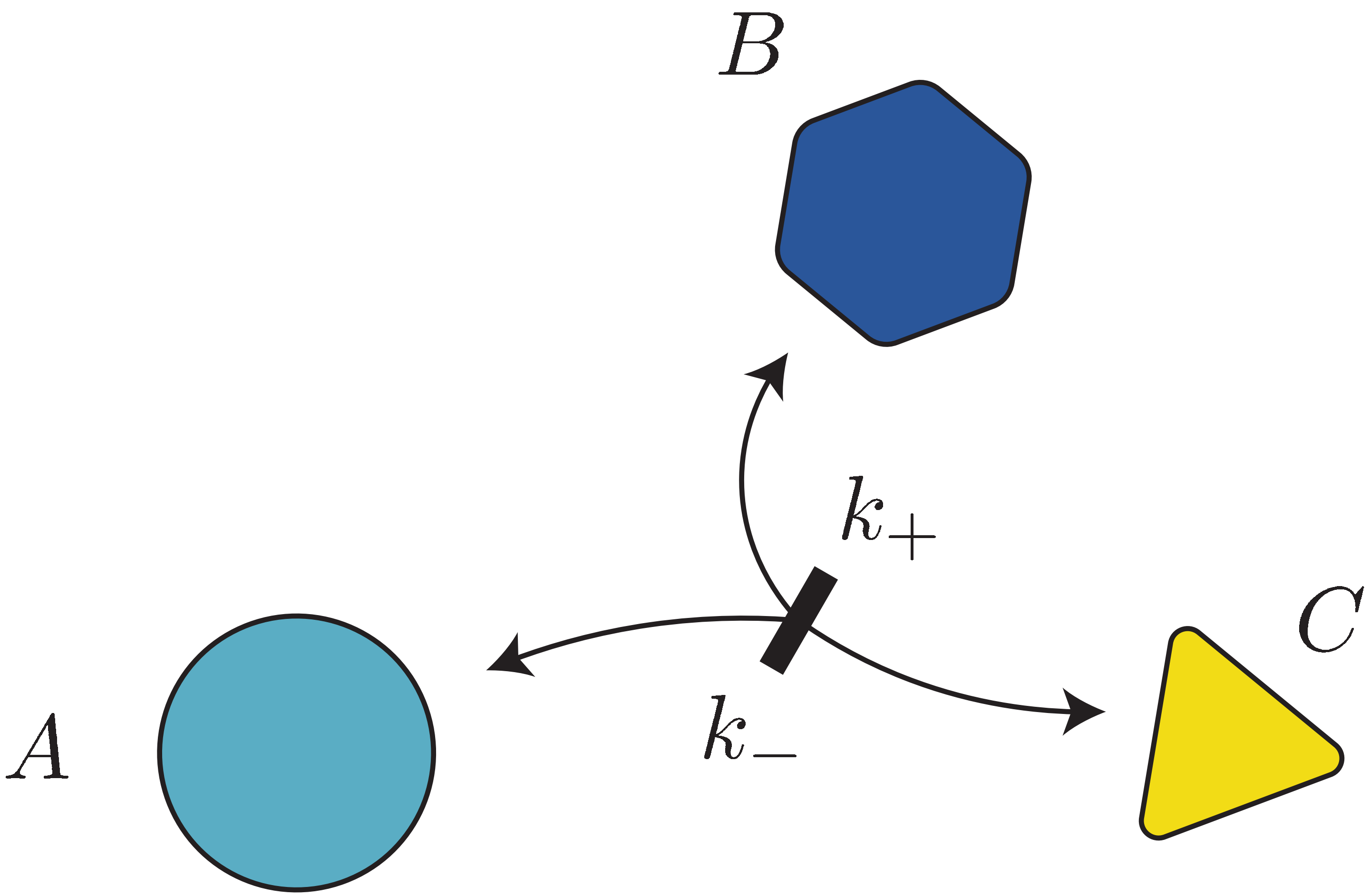


Van Kampen, N.G., 1992. Stochastic processes in physics and chemistry (Vol. 1). Elsevier.

Stochasticity

What are discrete, stochastic chemical systems capable of?

If each reaction is in equilibrium with its reverse reaction then the system is in detailed balance



$$G_C - G_A - G_B = \ln \frac{k_-}{k_+}$$

$$\varepsilon_A = e^{-G_A}$$

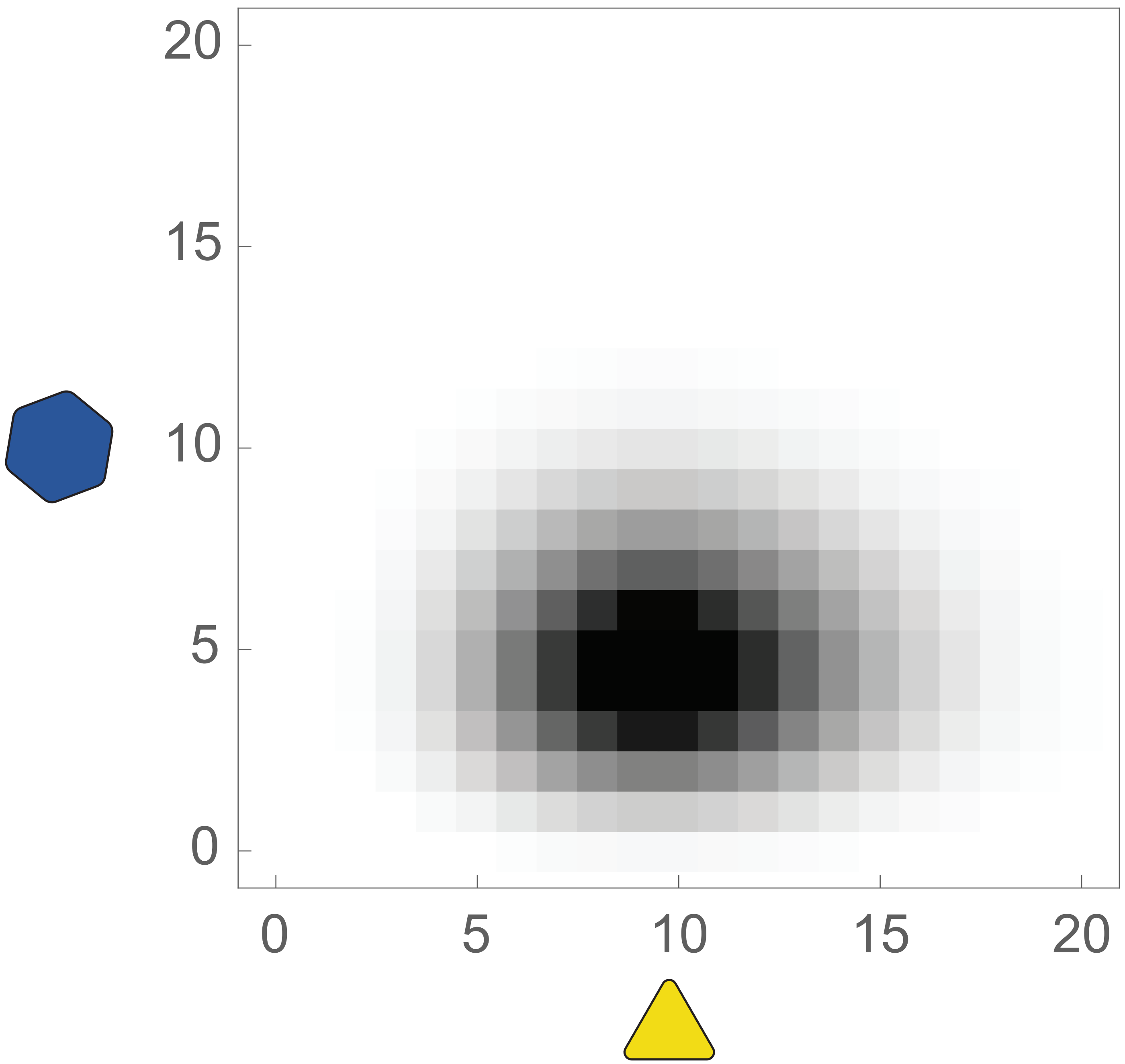
$$\varepsilon_B = e^{-G_B}$$

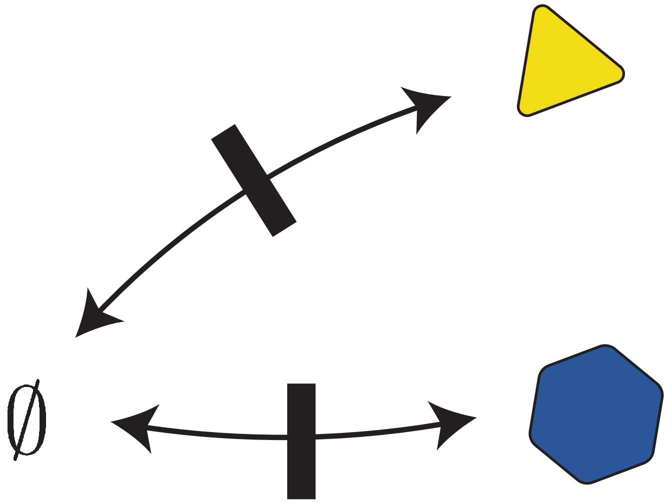
$$\varepsilon_C = e^{-G_C}$$

$$k_+ \varepsilon_A \varepsilon_B = k_- \varepsilon_C$$

Horn, F. and Jackson, R., 1972. General mass action kinetics. Archive for rational mechanics and analysis, 47(2), pp.81-116.

Detailed balanced systems have stationary multivariate Poisson distributions



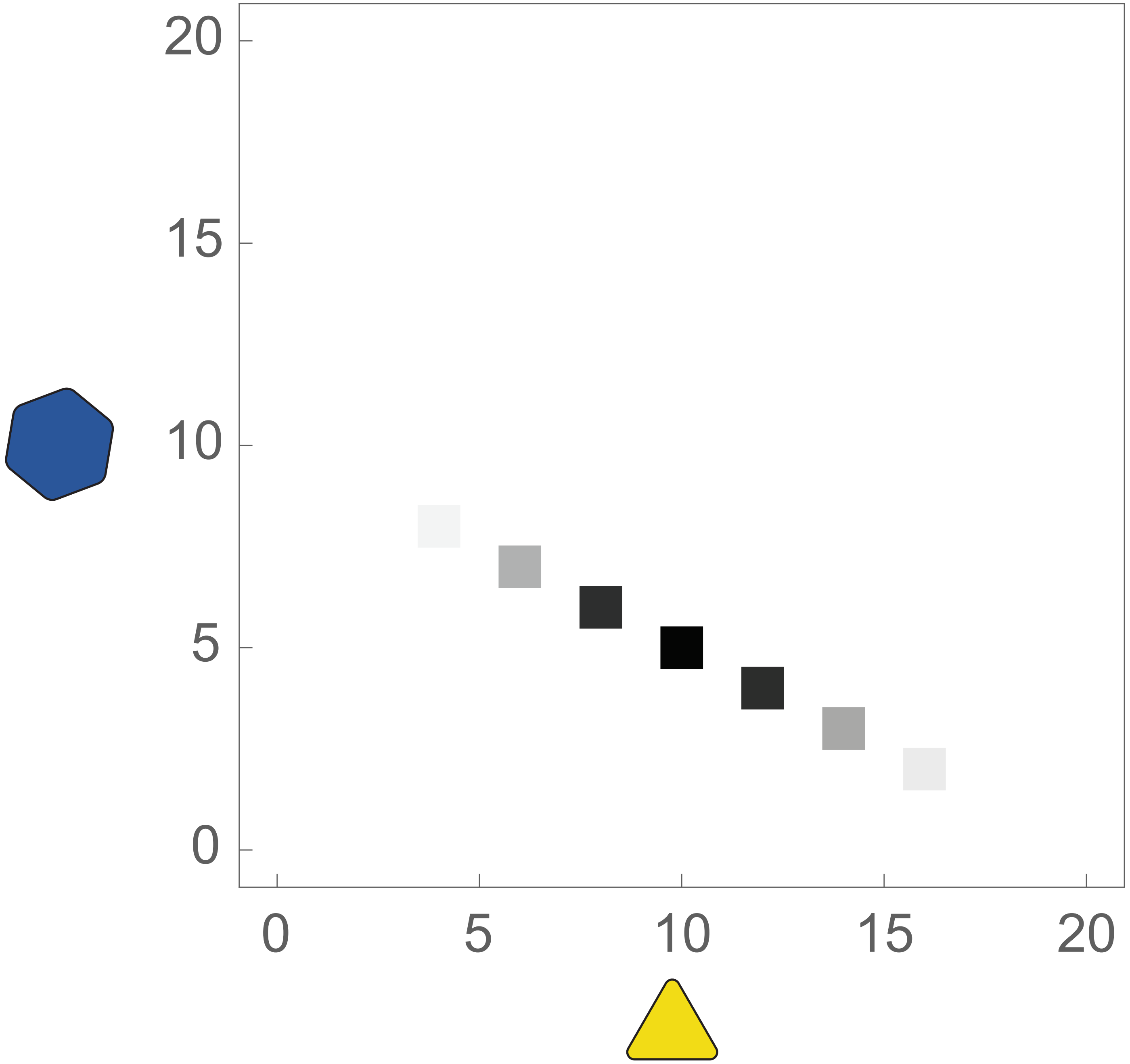


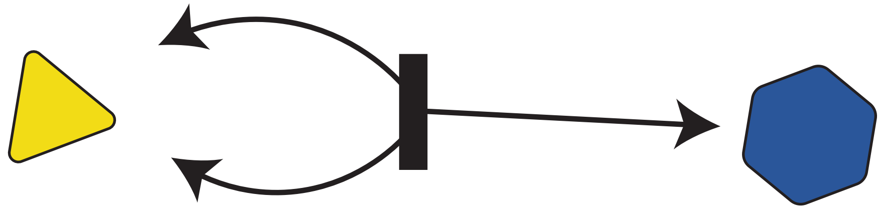
$$p_{m,n} \propto \frac{\varepsilon_A^m \varepsilon_B^n}{m!n!}$$

multivariate Poisson distribution

Van Kampen, N.G., 1992. Stochastic processes in physics and chemistry (Vol. 1). Elsevier.

Detailed balanced systems have stationary multivariate Poisson distributions



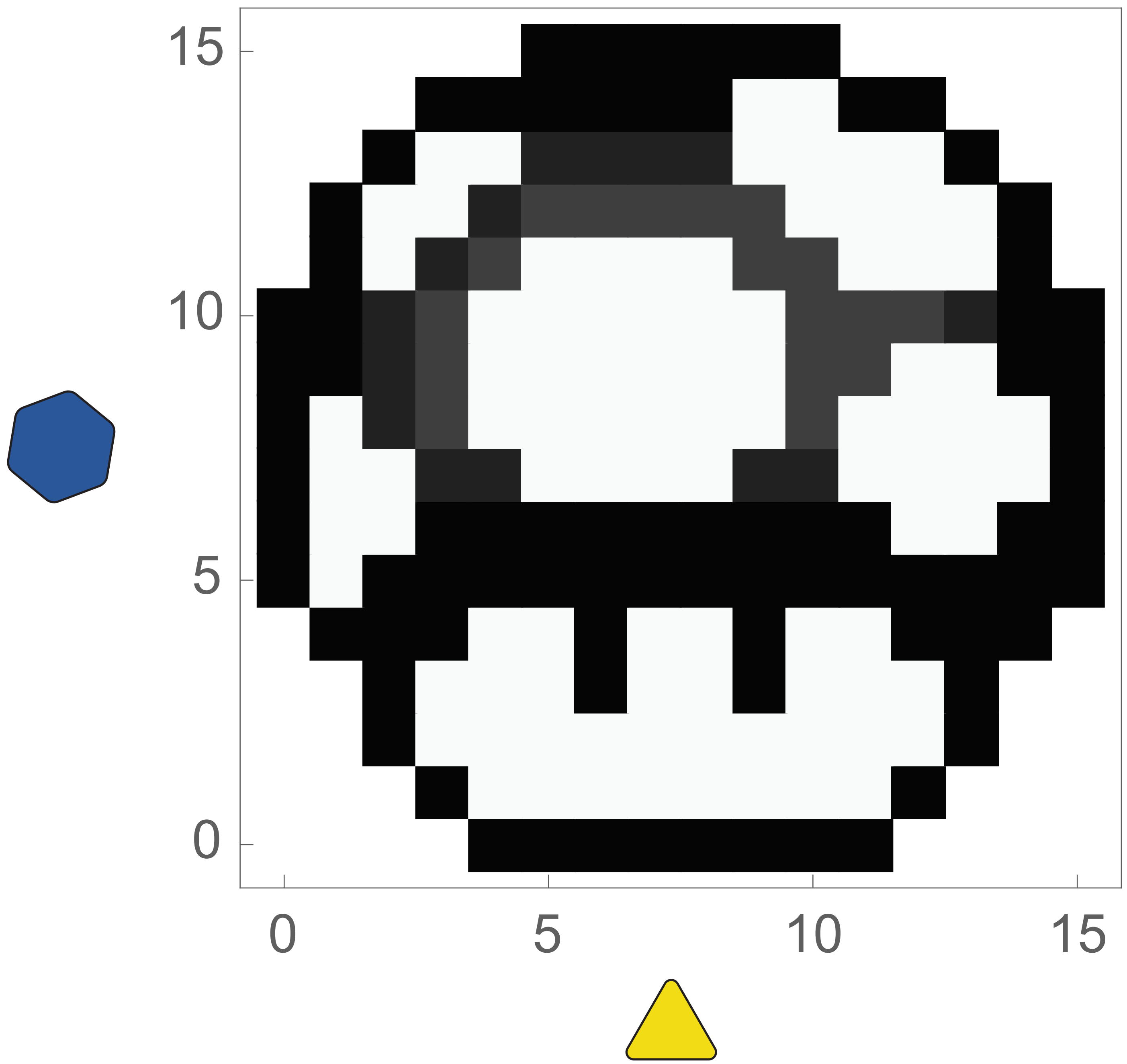


$$p_{m,n} \propto \frac{\varepsilon_A^m \varepsilon_B^n}{m!n!}$$

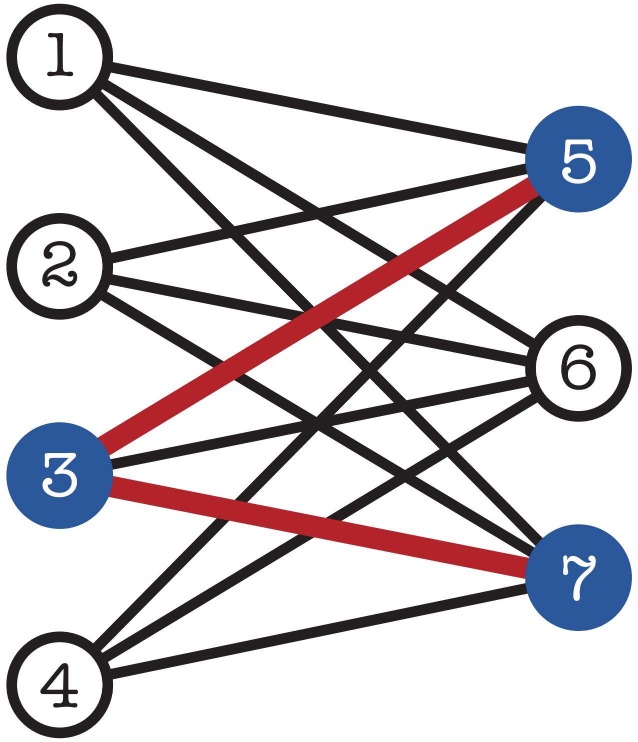
restricted multivariate
Poisson distribution

Van Kampen, N.G., 1992. Stochastic processes in physics and chemistry (Vol. 1). Elsevier.

Can detailed balanced systems produce complex distributions?



Botlzmann machines are stochastic models of neural networks



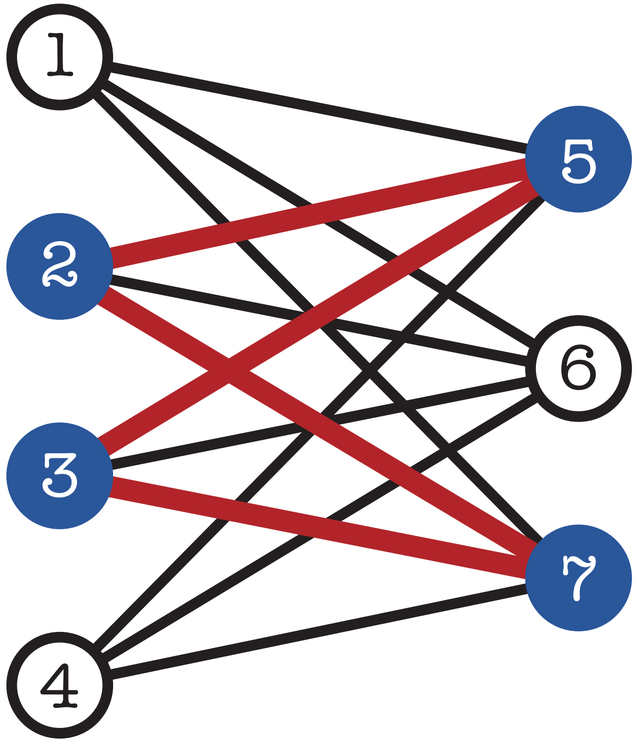
$$E(0,0,1,0;1,0,1) =$$

$$-w_{3,5} - w_{3,7} + \theta_3 + \theta_5 + \theta_7$$

$$p(x) = \frac{e^{-E(x)}}{Z}$$

Hinton, G.E. and Sejnowski, T.J., 1983, June. Optimal perceptual inference. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (Vol. 448).

Botlzmann machines are stochastic models of neural networks

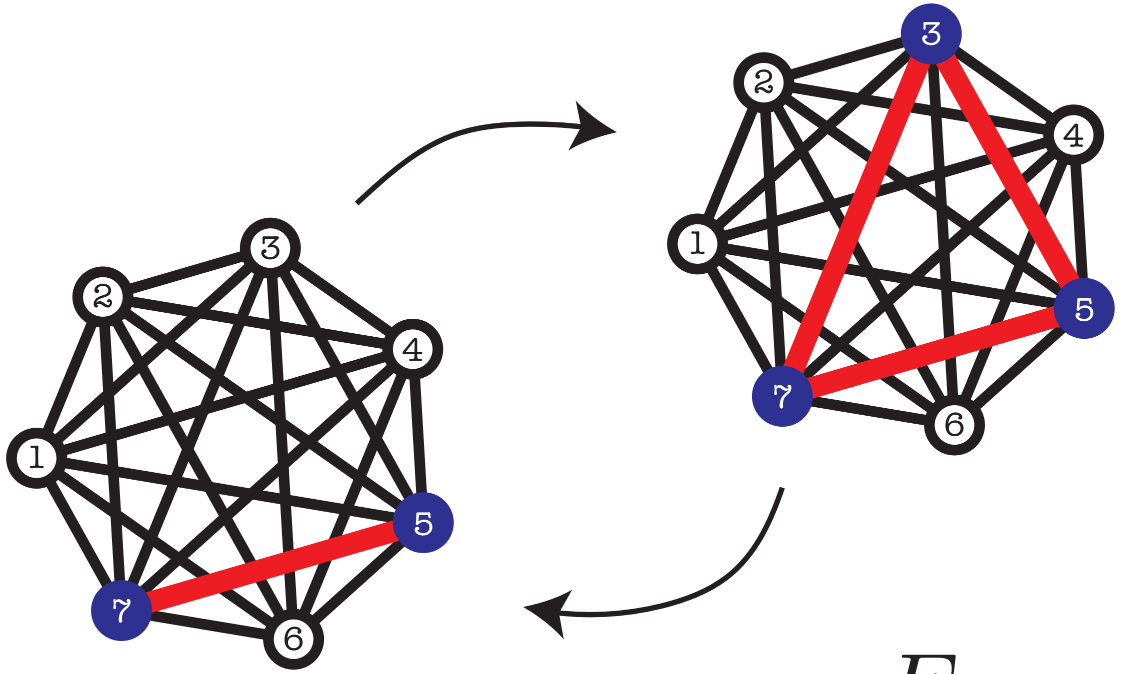


$$E(0,0,1,0;1,0,1) =$$

$$-w_{2,5} - w_{2,7} - w_{3,5} - w_{3,7} + \theta_3 + \theta_5 + \theta_7 + \theta_2$$

$$p(x) = \frac{e^{-E(x)}}{Z}$$

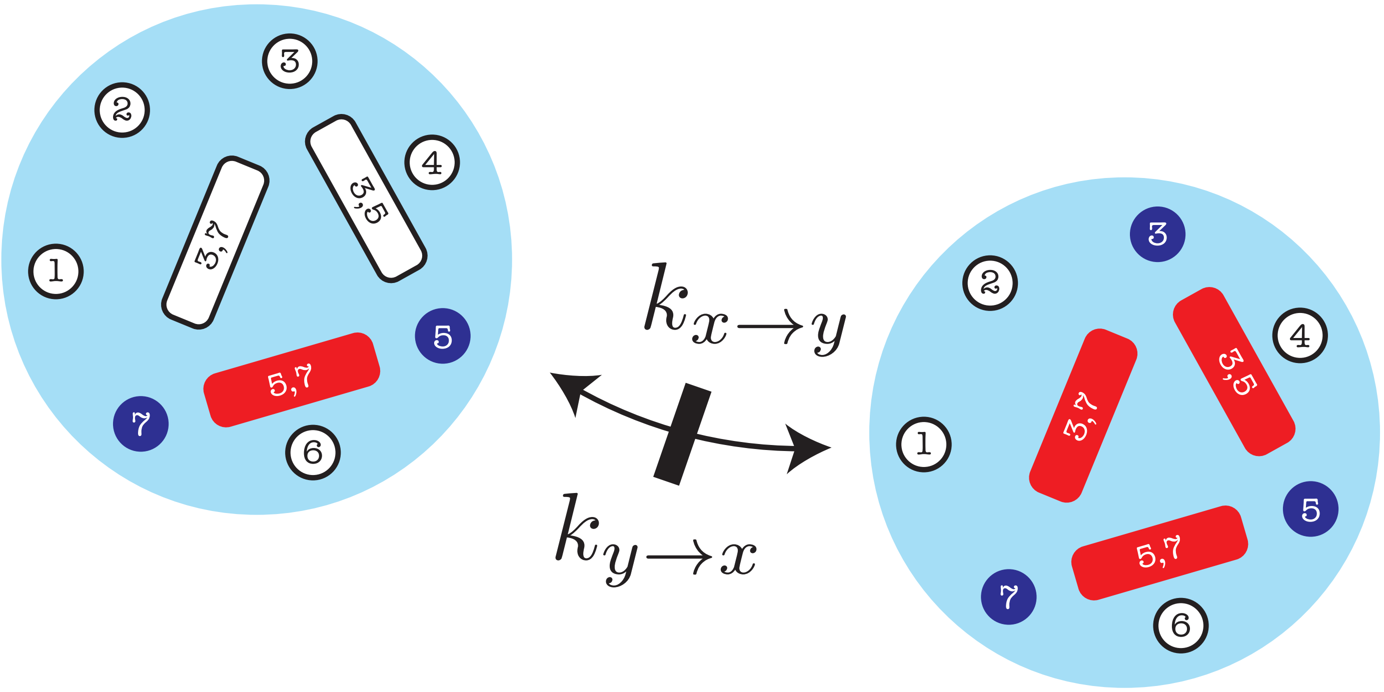
Hinton, G.E. and Sejnowski, T.J., 1983, June. Optimal perceptual inference. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (Vol. 448).



Chemical systems can behave like Boltzmann machines

$$E_x = -w_{5,7} + \theta_5 + \theta_7$$

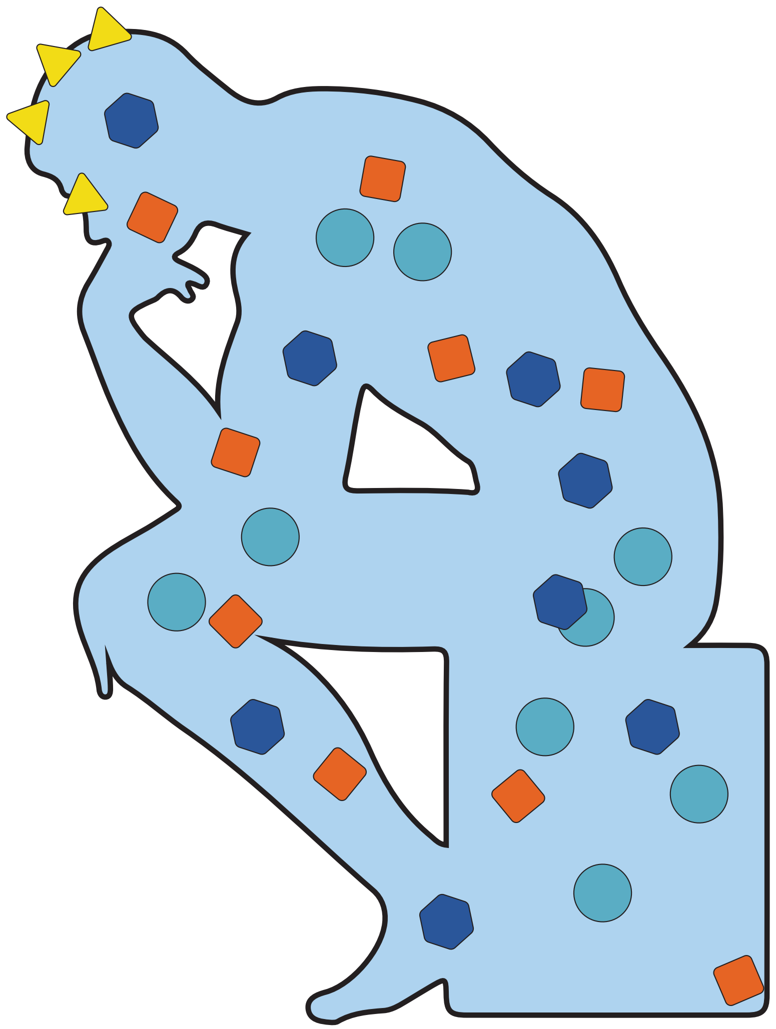
$$E_y = -w_{3,5} - w_{3,7} - w_{5,7} + \theta_3 + \theta_5 + \theta_7$$



$$\ln \frac{k_{y \to x}}{k_{x \to y}} = E_y - E_x$$

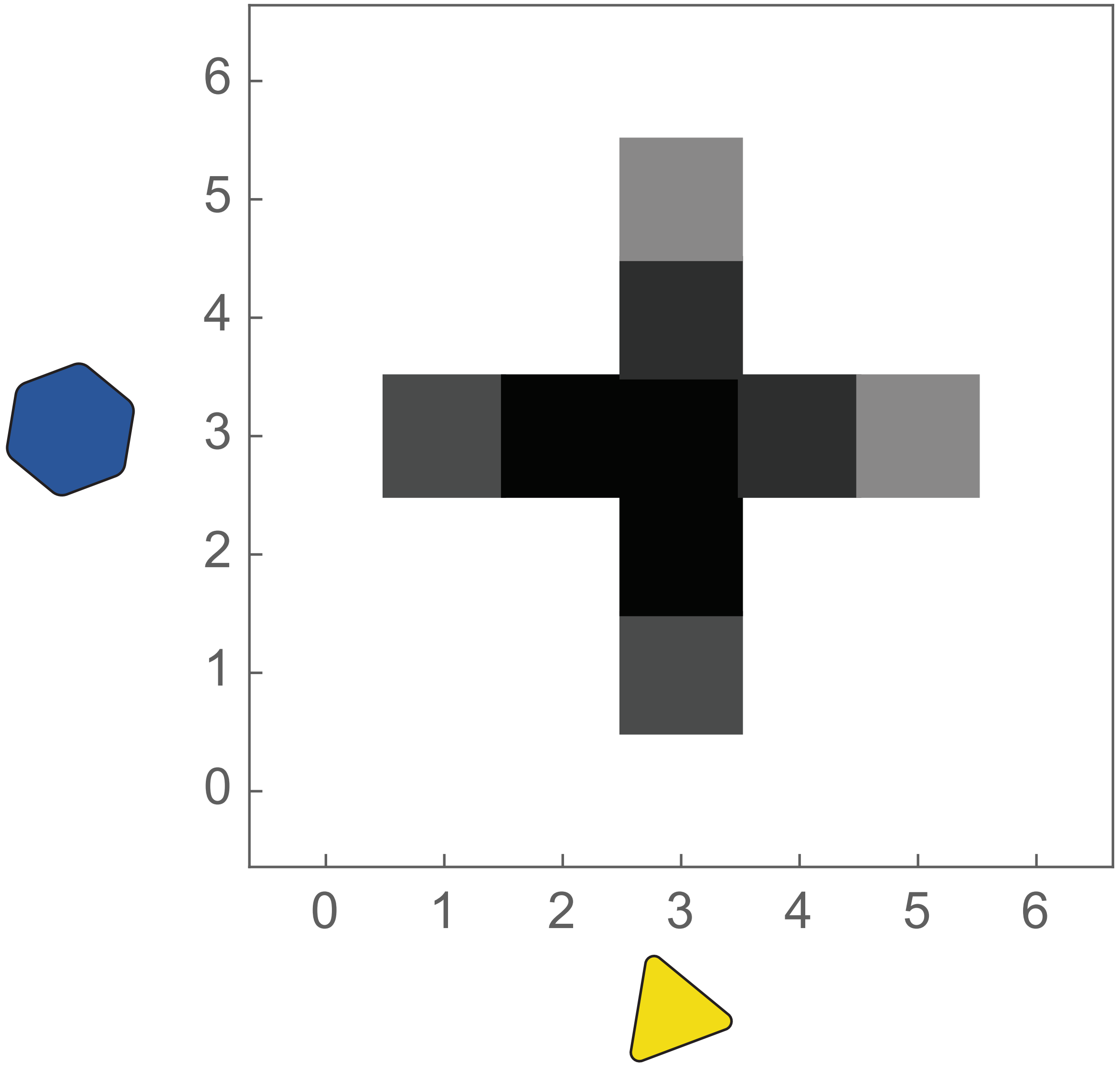
$$= -w_{3,5} - w_{3,7} + \theta_3$$

Poole W., Ortiz-Muñoz, A., et al. (2017) Chemical Boltzmann Machines. In: Brijder R., Qian L. (eds) DNA Computing and Molecular Programming. DNA 2017. Lecture Notes in Computer Science, vol 10467. Springer, Cham.



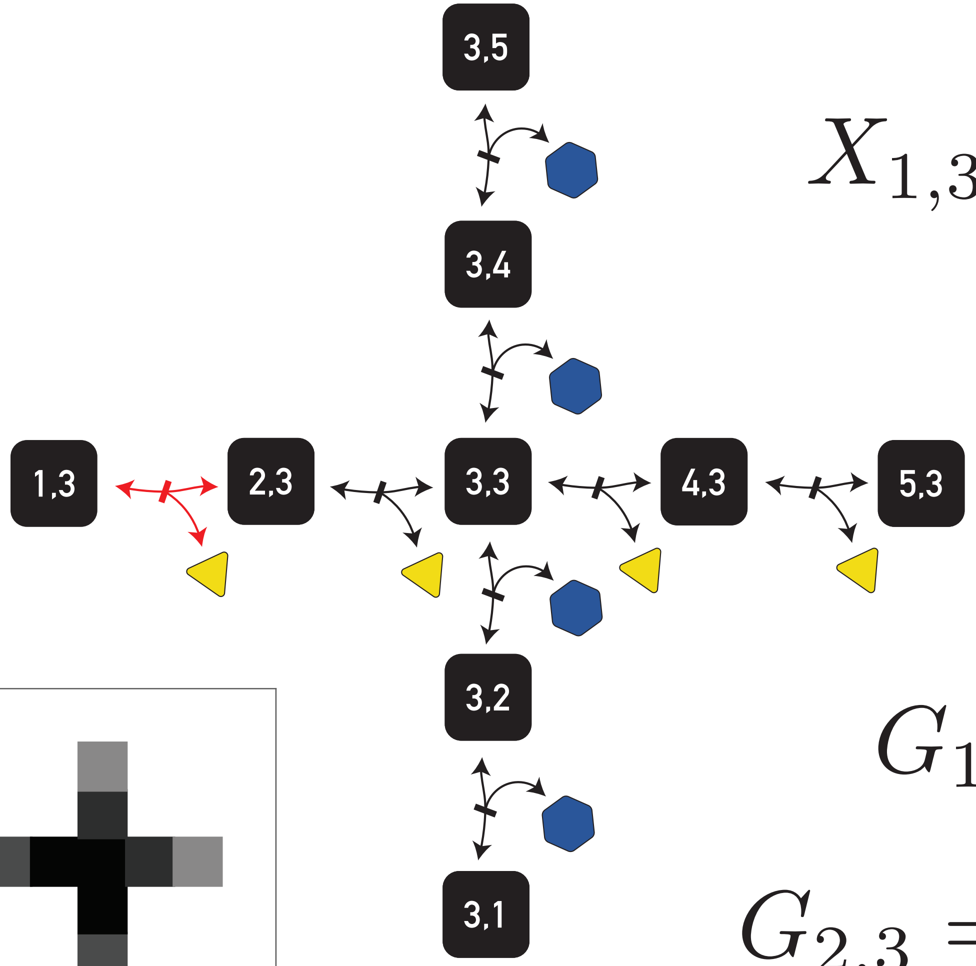
Chemical systems can produce
complex distributions,
sample them,
and perform inference

Detailed-balanced systems with hidden species can produce any distribution



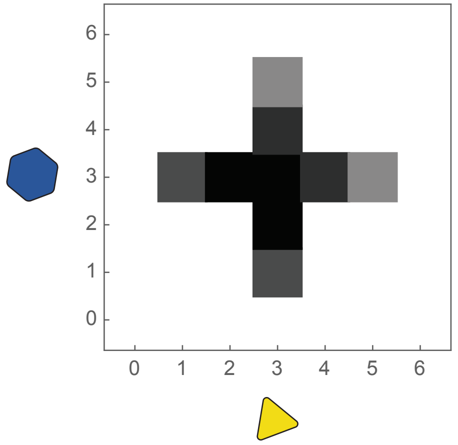
Cappelletti, D., Ortiz-Muñoz, A., Anderson, D.F. and Winfree, E., 2020. Stochastic chemical reaction networks for robustly approximating arbitrary probability distributions. Theoretical Computer Science, 801, pp.64-95.

Detailed-balanced systems with hidden species can produce any distribution



$$X_{1,3} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} X_{2,3} + A$$

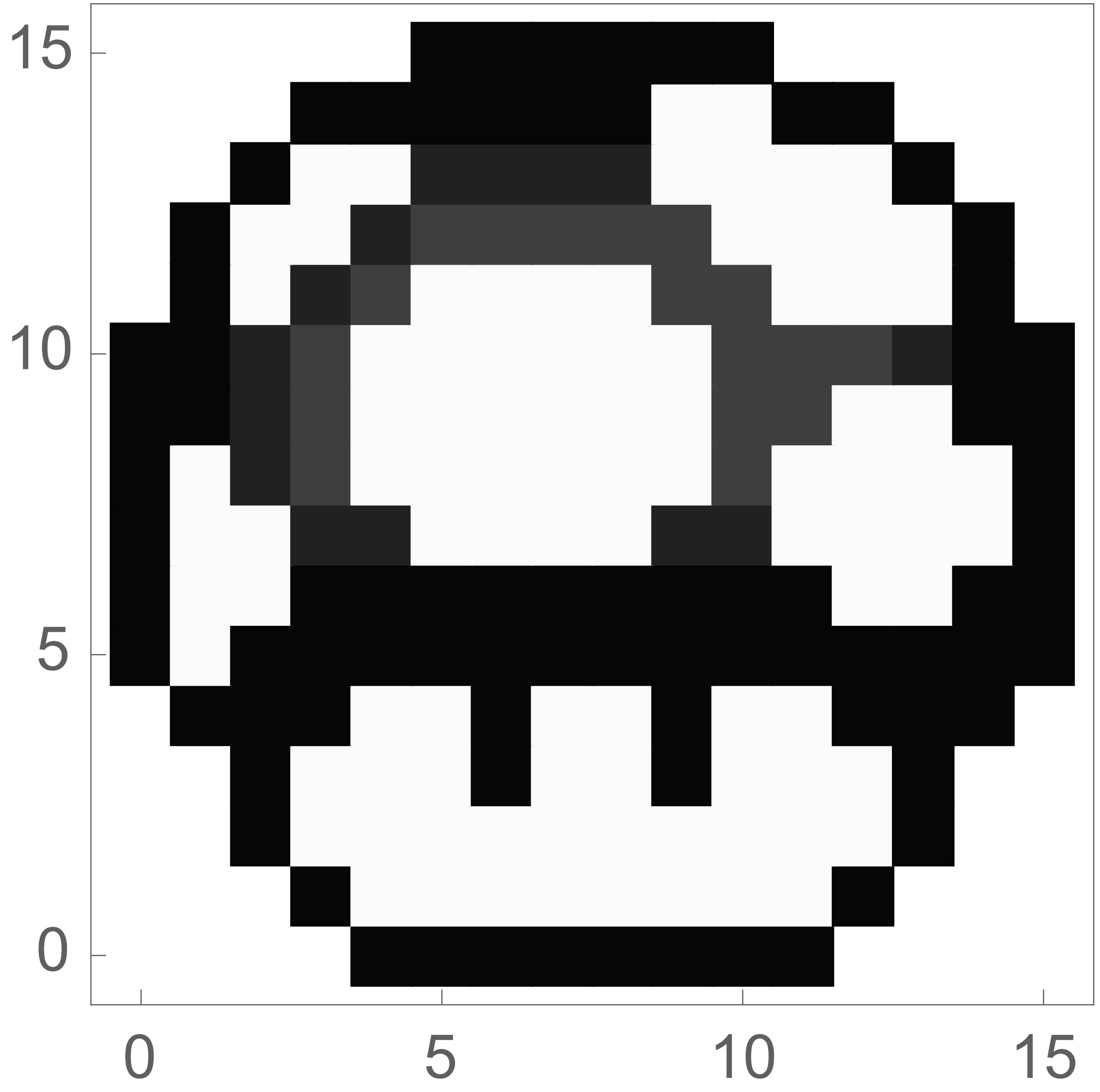
$$\ln \frac{k_-}{k_+} = G_{2,3} + G_A - G_{1,3}$$



$$G_{1,3} = -\ln(1!3!q_{2,3}) - G_A - 3G_B$$

$$G_{2,3} = -\ln(2!3!q_{2,3}) - 2G_A - 3G_B$$

Cappelletti, D., Ortiz-Muñoz, A., Anderson, D.F. and Winfree, E., 2020. Stochastic chemical reaction networks for robustly approximating arbitrary probability distributions. Theoretical Computer Science, 801, pp.64-95.

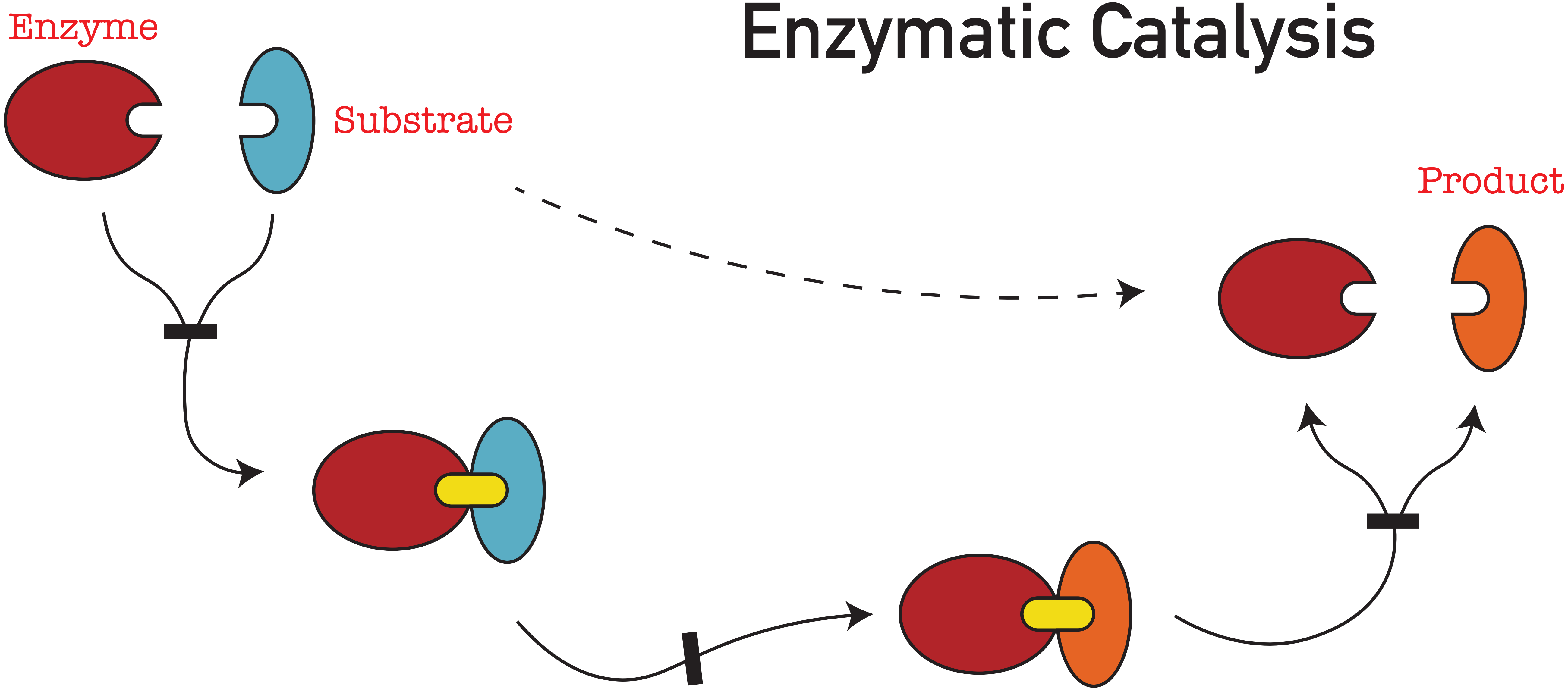


Chemical reaction networks can approximate arbitrary probability distributions!

Combinatorics

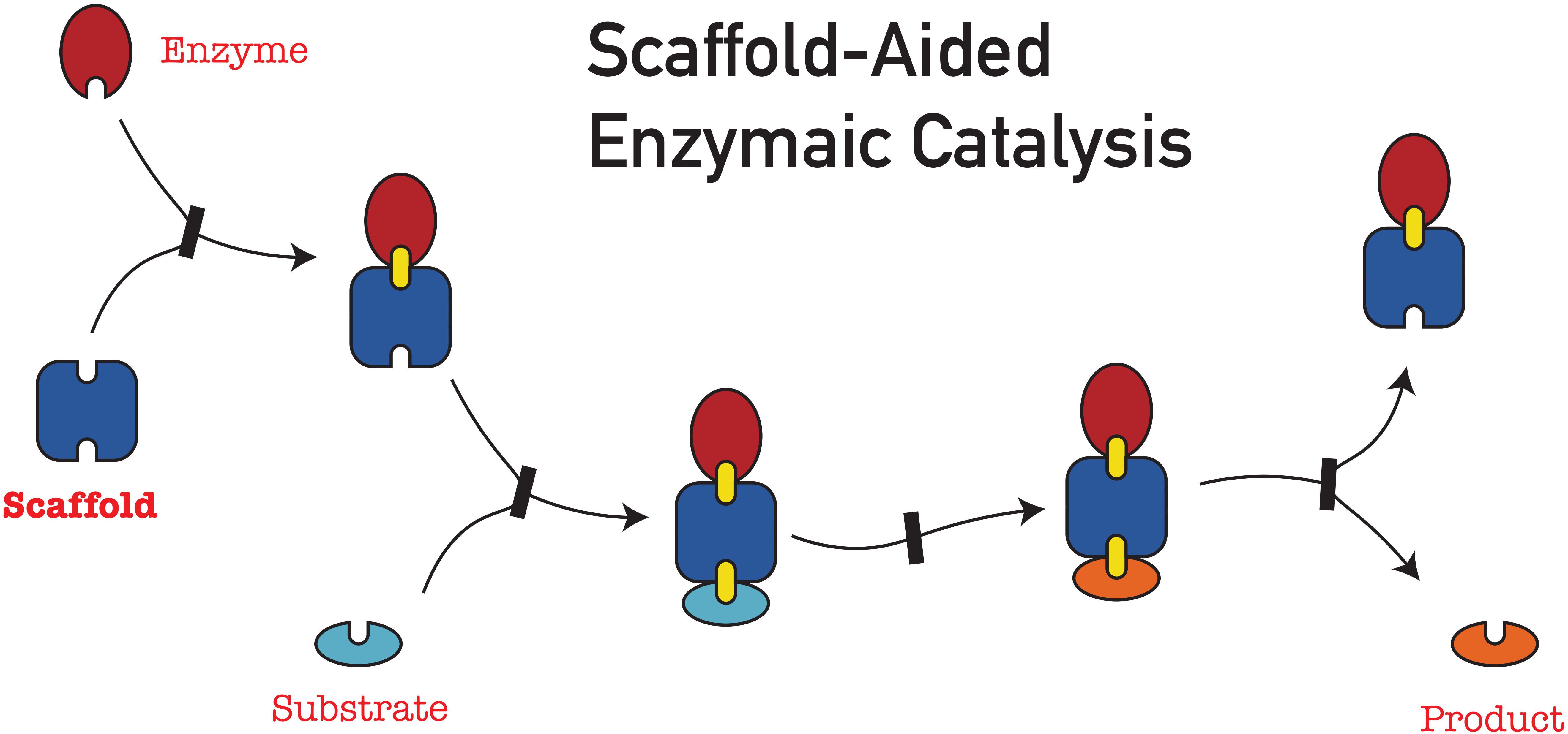
What role does combinatorics play in biology?

Enzymatic Catalysis



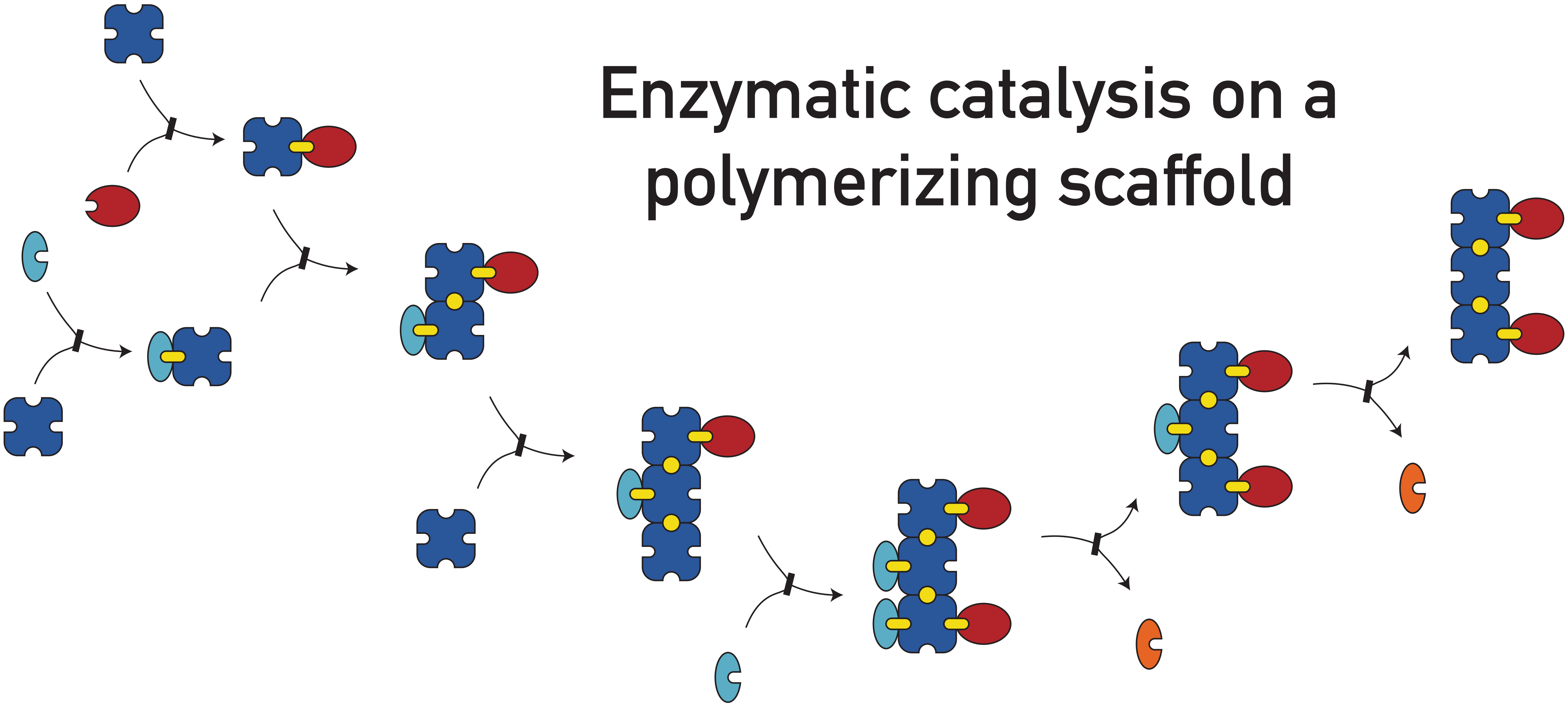
Cornish-Bowden, A., 2013. Fundamentals of enzyme kinetics. John Wiley & Sons.

Scaffold-Aided Enzymaic Catalysis



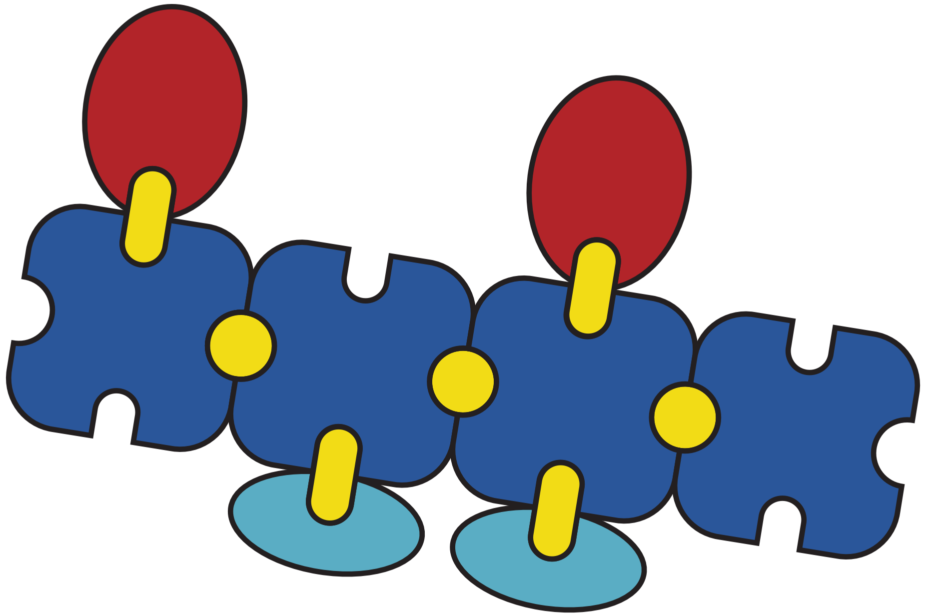
Ikeda, S., et. al. 1998. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3β and β-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. The EMBO journal, 17(5), pp.1371-1384.

Enzymatic catalysis on a polymerizing scaffold

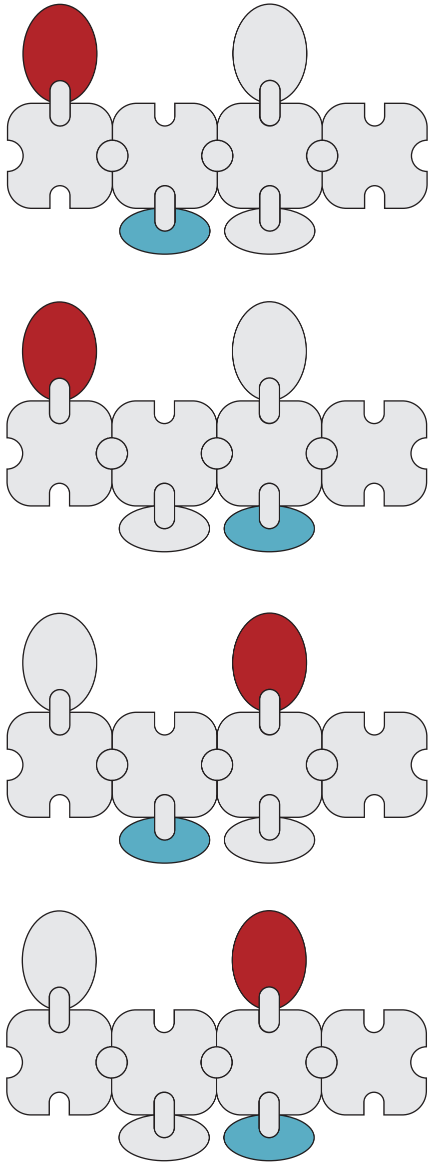


Fiedler, M., et.al. 2011. Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating β-catenin. Proceedings of the National Academy of Sciences, 108(5), pp.1937-1942.

The catalytic potential of a complex is its effective concentration of potential enzyme-substrate interactions

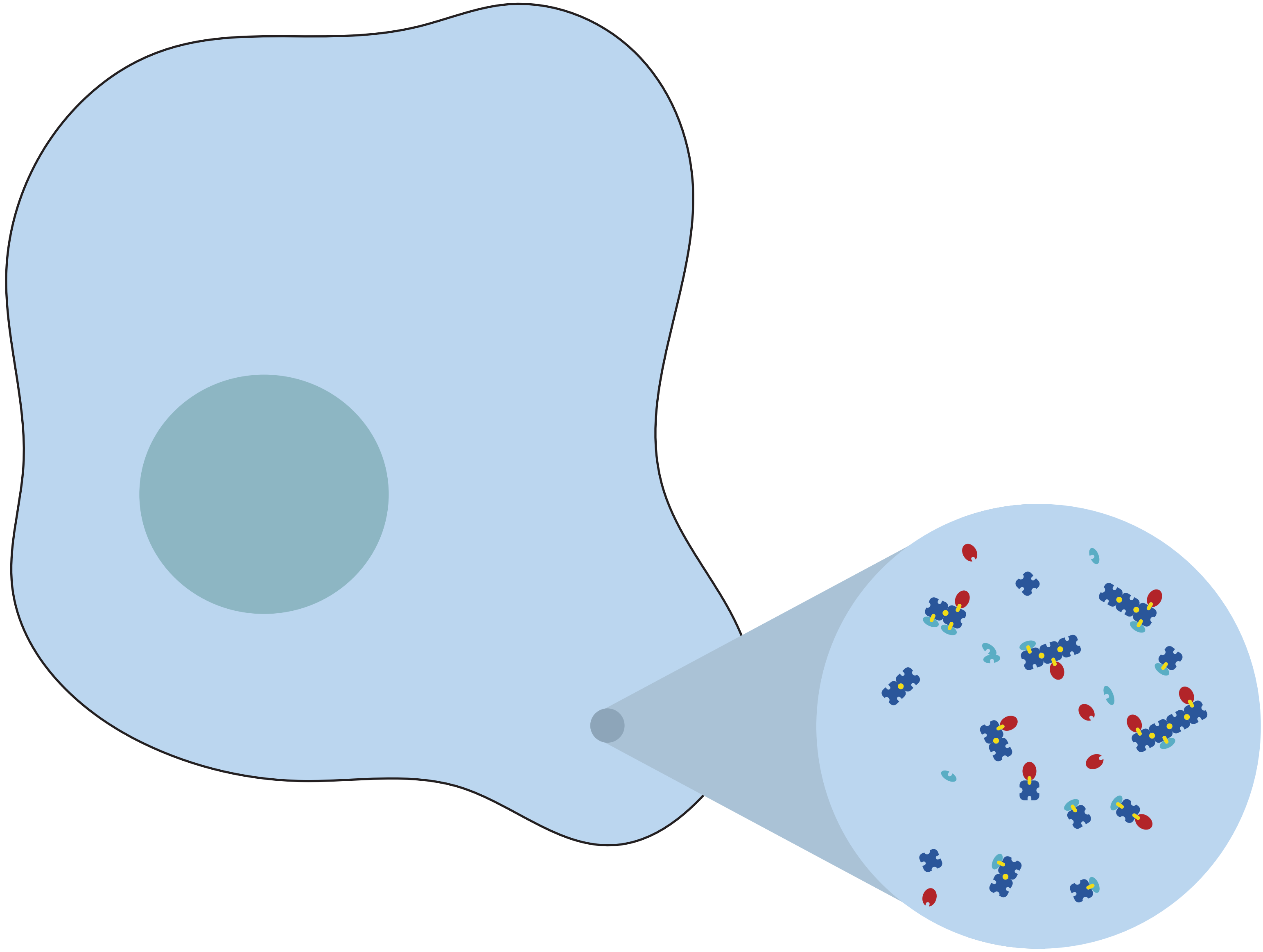


complex



potential interactions

Ortiz-Muñoz, A., Medina-Abarca, H.F. and Fontana, W., 2020. Combinatorial protein-protein interactions on a polymerizing scaffold. Proceedings of the National Academy of Sciences, 117(6), pp.2930-2937.



What is the catalytic potential of a cell?

Generating functions encode complex combinatorial patterns as algebraic expressions

$$x^4 = \blacksquare\blacksquare\blacksquare\blacksquare$$

$$\frac{d}{dx}x^4 = 4x^3$$

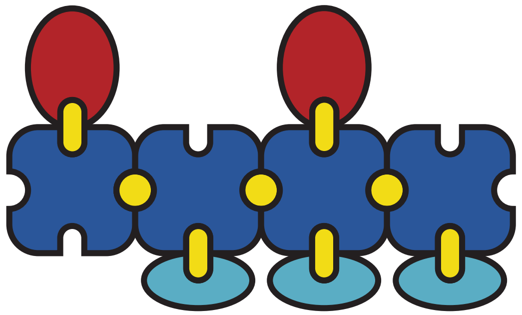
$$= \square\blacksquare\blacksquare\blacksquare + \blacksquare\square\blacksquare\blacksquare + \blacksquare\blacksquare\square\blacksquare + \blacksquare\blacksquare\blacksquare\square$$

$$f = 1 + x + x^2 + x^3 + \cdots = 1 + \blacksquare + \blacksquare\blacksquare + \blacksquare\blacksquare\blacksquare + \cdots$$

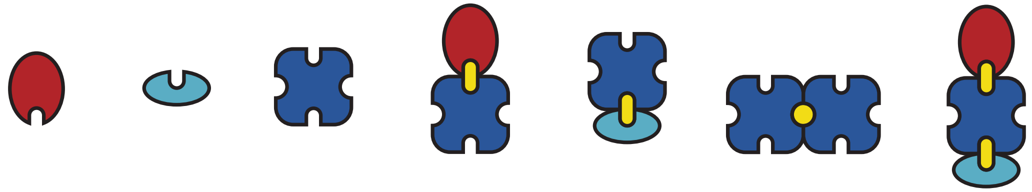
$$= 1 + \blacksquare f = 1 + xf = \boxed{\frac{1}{1-x}}$$

$$\frac{df}{dx} = \frac{1}{(1-x)^2} = f^2 = (1 + \blacksquare + \blacksquare\blacksquare + \cdots)\,\square\,(1 + \blacksquare + \blacksquare\blacksquare + \cdots)$$

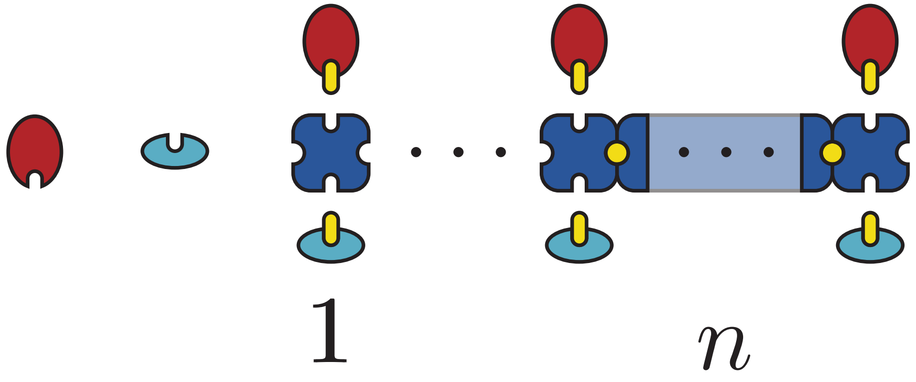
Flajolet, P. and Sedgewick, R., 2009. Analytic combinatorics. cambridge University press.

 $= \alpha^2\beta^3\sigma^3a^2b^3s^4$

We can use generating functions to encode the class of all possible complexes

$W =$  $\cdots$

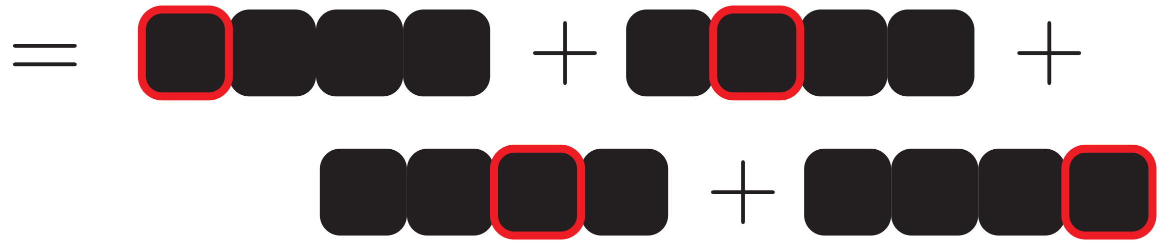
$$= a + b + s + \alpha a s + \beta b s + \sigma s^2 + \alpha\beta a b s + \cdots$$

$=$  $\cdots$

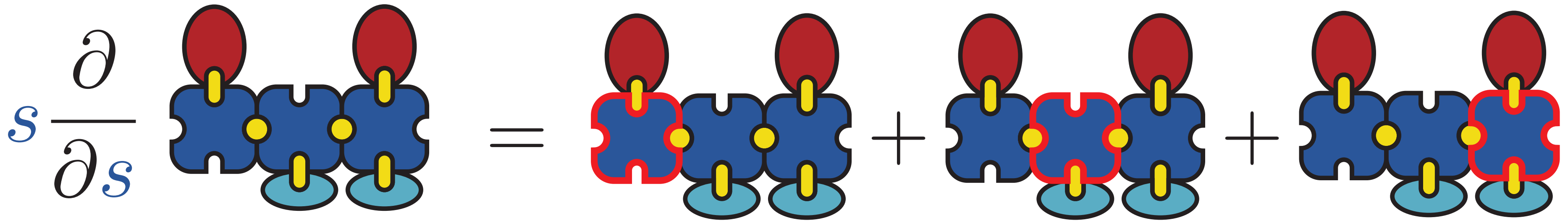
$$= a + b + \sum_{n=1}^{\infty} \sigma^{n-1} s^n (1+\alpha a)^n (1+\beta b)^n$$

$$= a + b + \frac{s(1+\alpha a)(1+\beta b)}{1-\sigma s(1+\alpha a)(1+\beta b)}$$

$$x\frac{d}{dx}x^4 = 4x^4$$

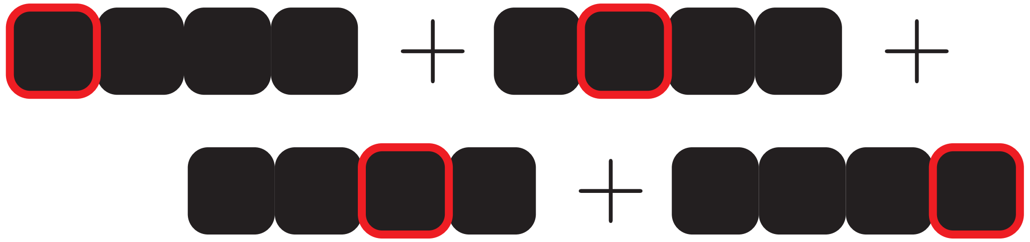


We can take derivatives of the complex generating function to express **conservation laws**

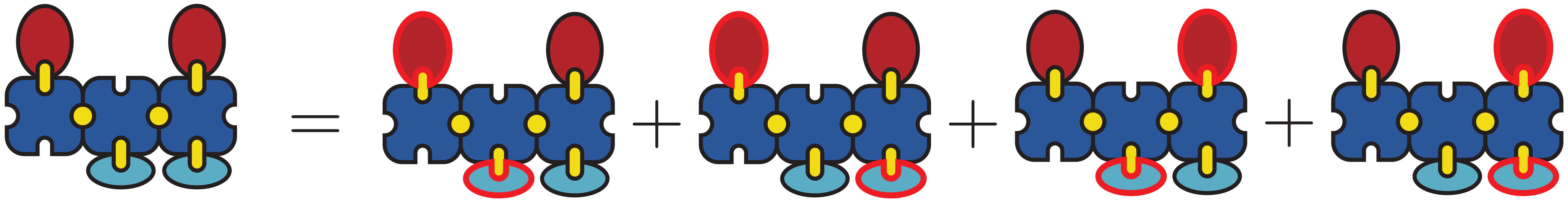


$$t_A = a\frac{\partial W}{\partial a} \qquad t_B = b\frac{\partial W}{\partial b} \qquad t_S = s\frac{\partial W}{\partial s}$$

$$x\frac{d}{dx}x^4 = 4x^4$$

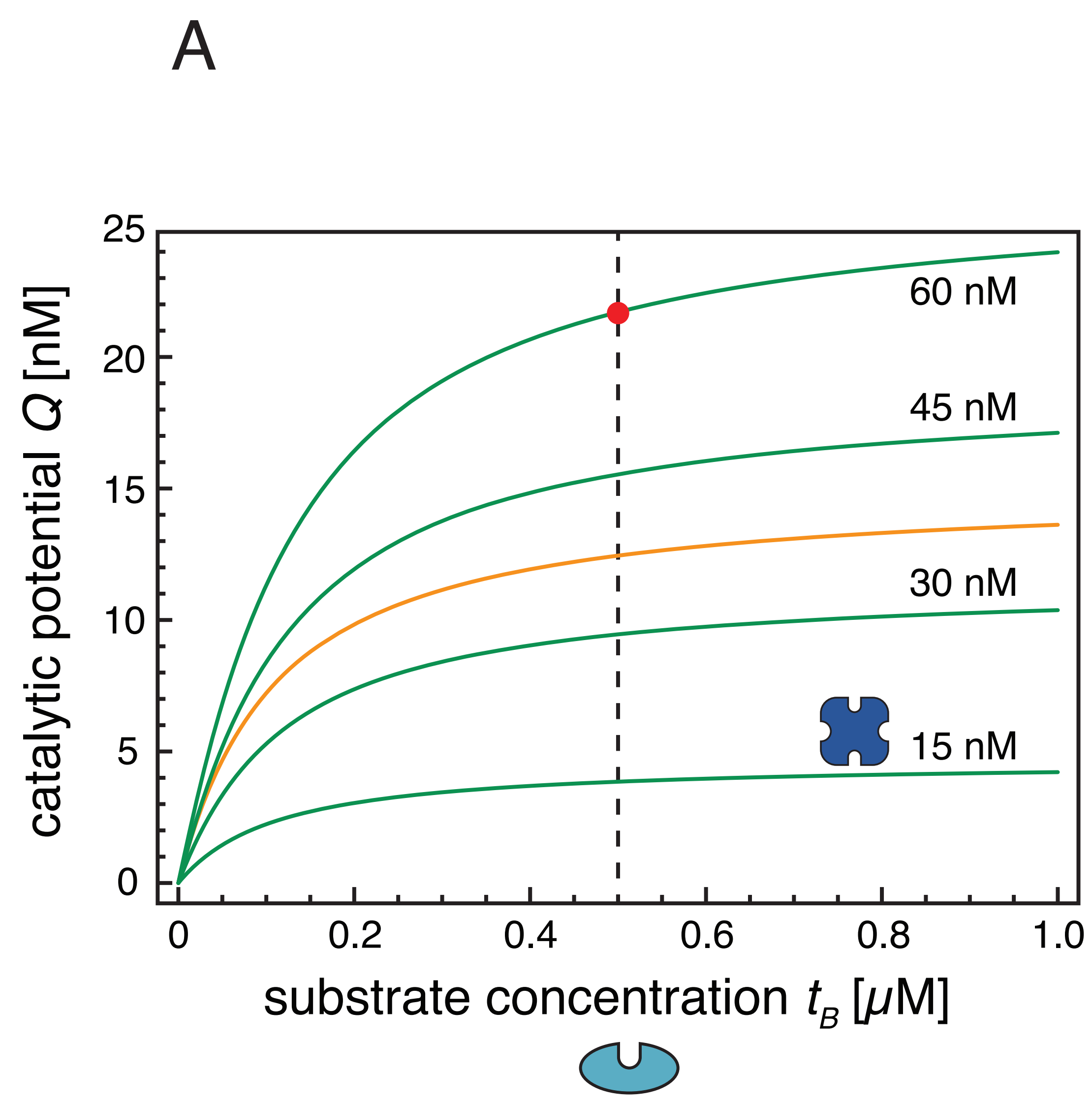
$=$ 

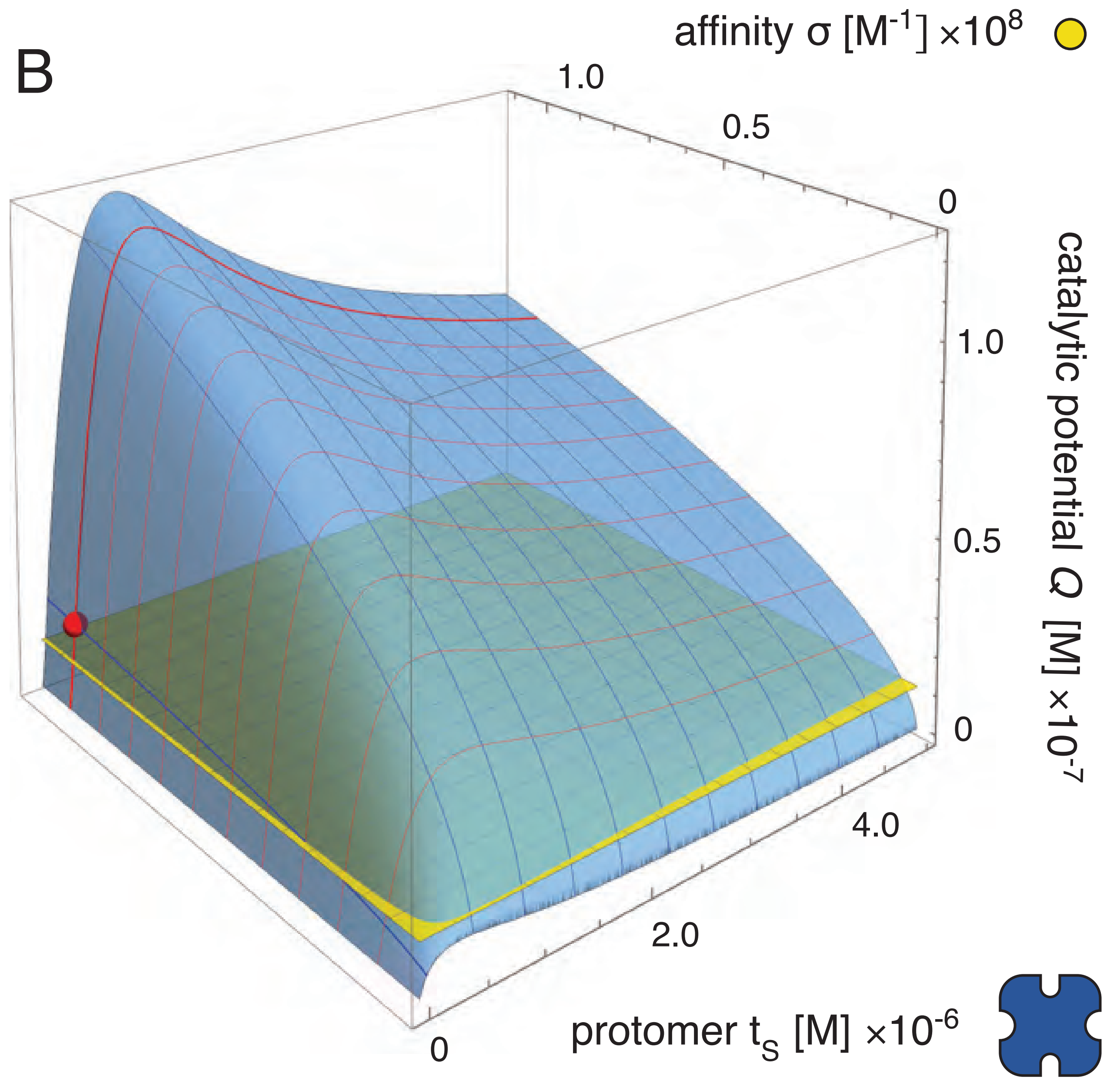
We can take derivatives of the complex generating function to extract the catalytic potential

$$ab\frac{\partial^2}{\partial a \partial b}$$ 

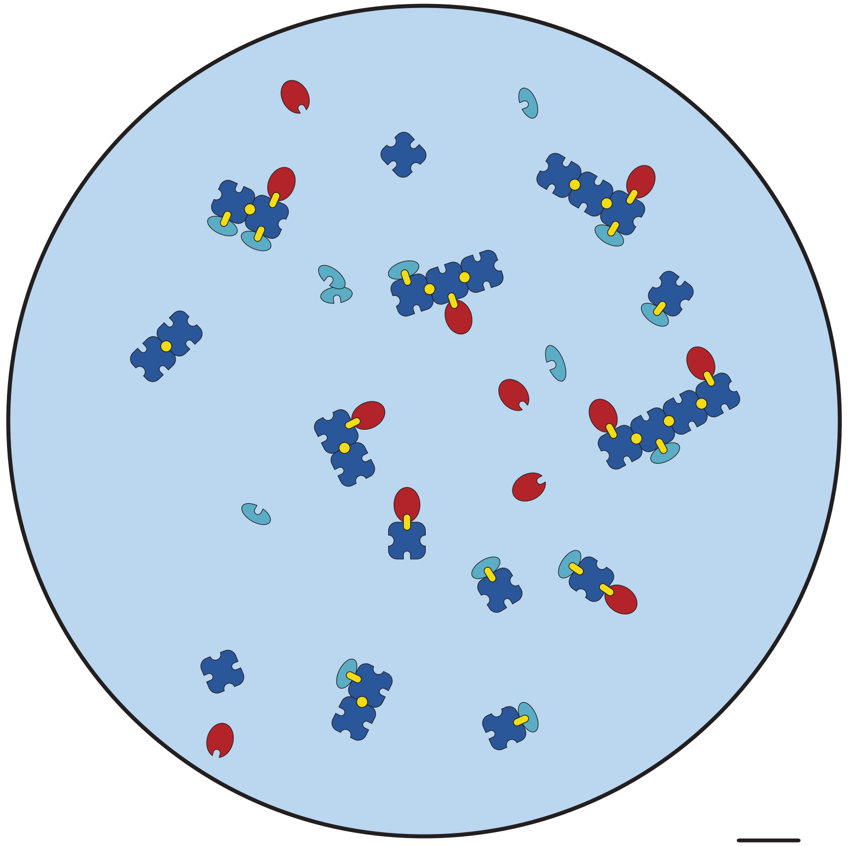
$$Q = ab\frac{\partial^2 W}{\partial a \partial b} = \alpha\beta abs\frac{1+\sigma s(1+\alpha a)(1+\beta b)}{(1-\sigma s(1+\alpha a)(1+\beta b))^3}$$

Catalytic potential is sensitive to polymerization parameters



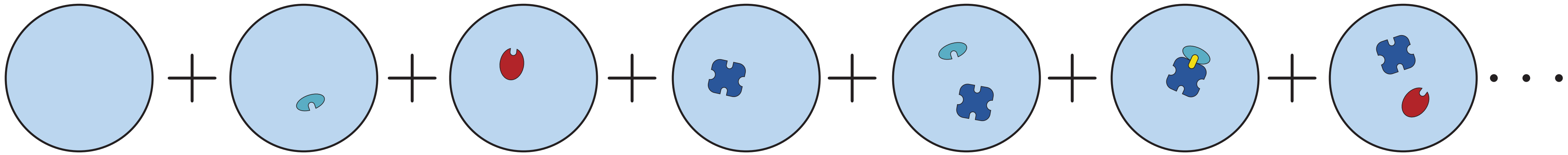


Ortiz-Muñoz, A., Medina-Abarca, H.F. and Fontana, W., 2020. Combinatorial protein–protein interactions on a polymerizing scaffold. Proceedings of the National Academy of Sciences, 117(6), pp.2930-2937.

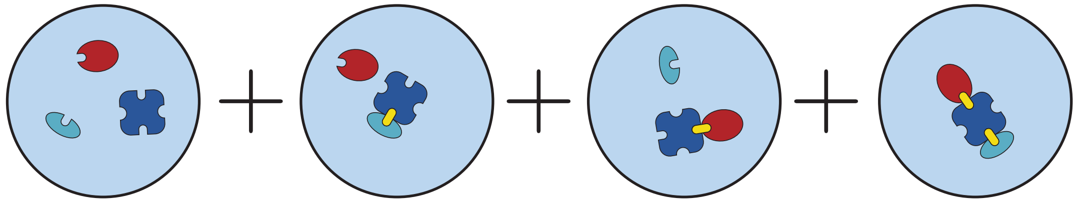


The stochasticity of the system can also be analyzed using generating functions

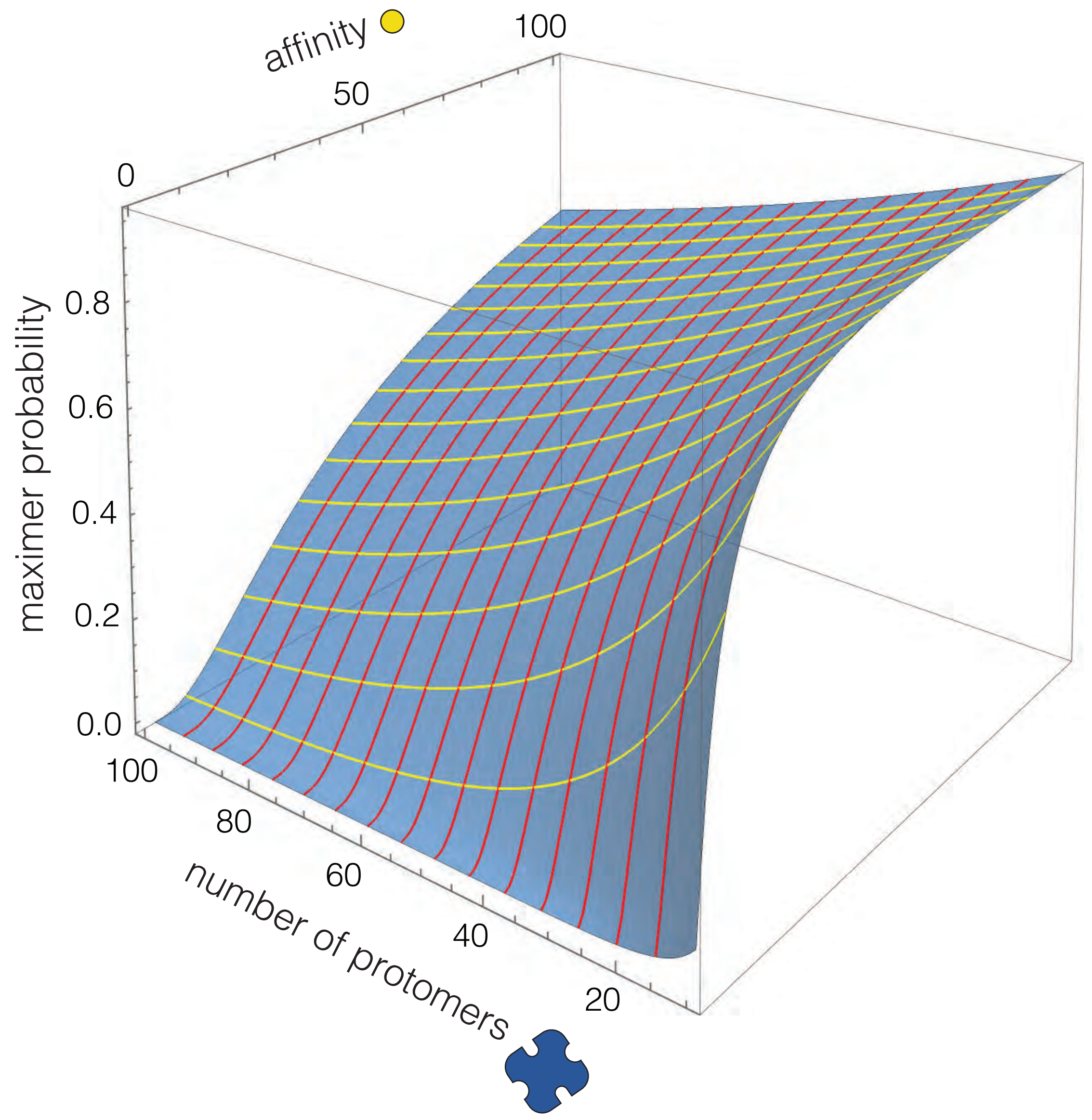
$$Z = e^W$$ ensemble generating function

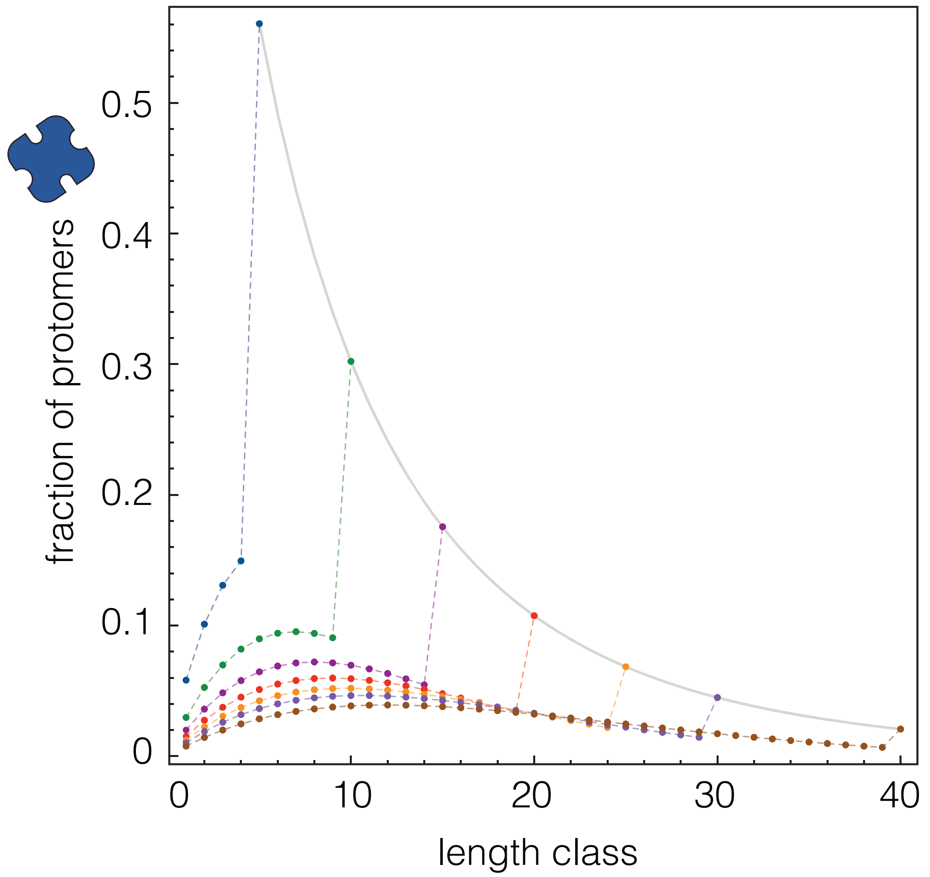
$=$ 

$$= \sum_{(t_A,t_B,t_C)\in\mathbb{N}^3} Z_{t_A,t_B,t_C}\, a^{t_A} b^{t_B} c^{t_C}$$

$Z_{1,1,1} =$ 

The discrete system has a polymer of maximum length





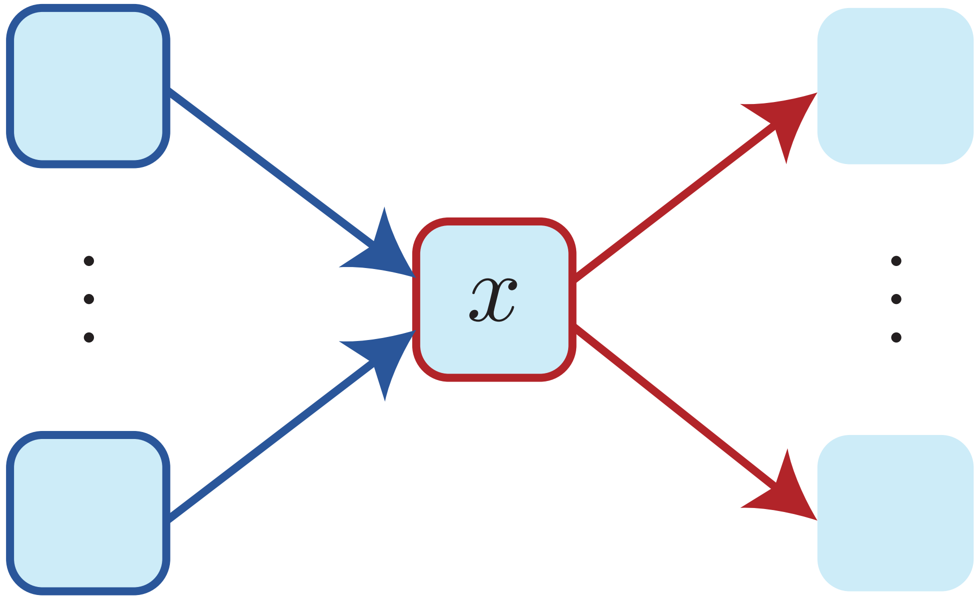
Ortiz-Muñoz, A., Medina-Abarca, H.F. and Fontana, W., 2020. Combinatorial protein–protein interactions on a polymerizing scaffold. Proceedings of the National Academy of Sciences, 117(6), pp.2930-2937.

Formal Semantics

Can we interpret each variable
formally rather than numerically?

The chemical master equation is a probability flux equation

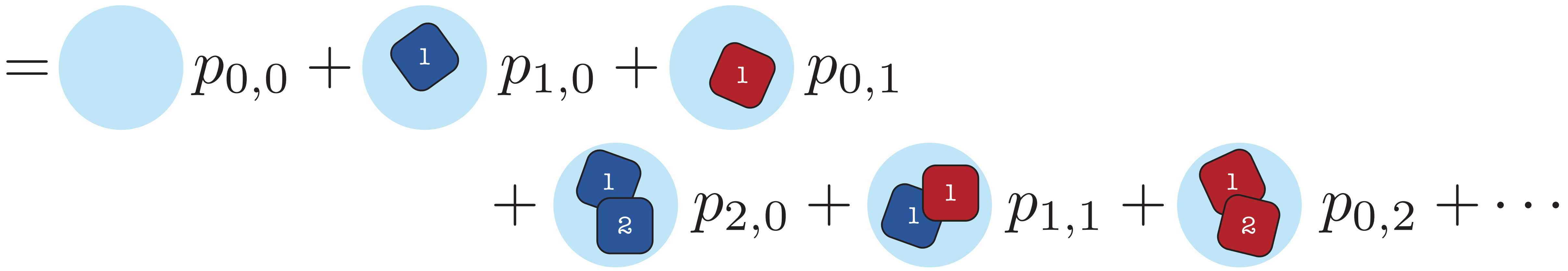
$$\frac{dp_x}{dt} = \sum_{r \in R} k_r \frac{(x + \rho_r - \pi_r)!}{(x - \pi_r)!} p_{x+\rho_r-\pi_r} - k_r \frac{x!}{(x - \rho_r)!} p_x$$



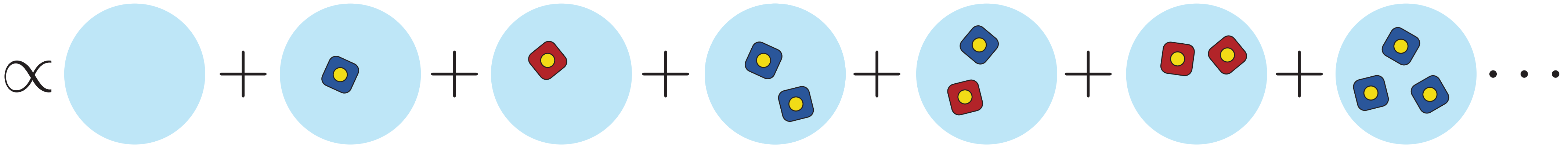
Van Kampen, N.G., 1992. Stochastic processes in physics and chemistry (Vol. 1). Elsevier.

Probability generating functions can encode complex probability distributions in algebraic expressions

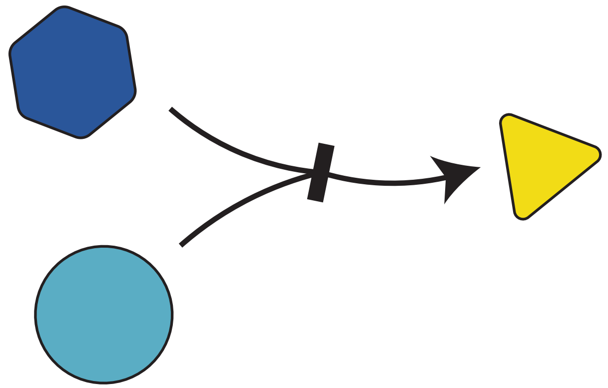
$$f = \sum_{(m,n)\in\mathbb{N}^2} p_{m,n} z_A^m z_B^n$$



$$f = e^{\lambda_A (z_A - 1) + \lambda_B (z_B - 1)} = \sum_{(m,n)\in\mathbb{N}} \left(e^{-\lambda_A} \frac{\lambda_A^m}{m!} \right) \left(e^{-\lambda_B} \frac{\lambda_B^n}{n!} \right) z_A^m z_B^n$$



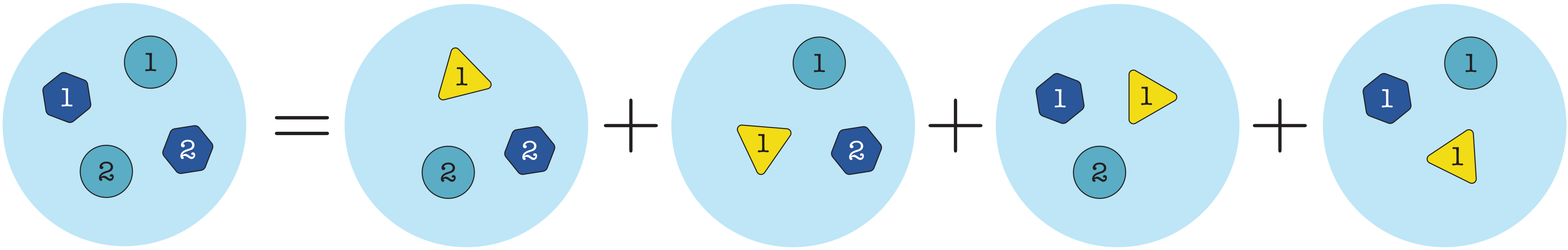
Baez, J.C. and Biamonte, J.D., 2018. Quantum techniques in stochastic mechanics. World Scientific.



Reaction operators can be encoded as partial differential operators

$$z_C \partial_A \partial_B (z_A^2 z_B^2) = 4 z_A z_B z_C$$

$z_C \partial_A \partial_B$



$$\boxed{\mathcal{A}_r}\, f = k_r (z_C - z_A z_B) \partial_A \partial_B f$$

Baez, J.C. and Biamonte, J.D., 2018. Quantum techniques in stochastic mechanics. World Scientific.

$$\sum_{r \in R} \mathcal{A}_r = \mathcal{A} \qquad \frac{df}{dt} = \mathcal{A} f$$

The stochastic dynamics of a system can be obtained by applying reaction operators recursively

$$f = f_0 + \int_0^t \mathcal{A} f \, dt$$

$$\int_0^t \mathcal{A} \, dt = \mathcal{R} \qquad f = f_0 + \mathcal{R} f$$

$$= f_0 + \mathcal{R} f_0 + \mathcal{R}\mathcal{R} f_0 + \mathcal{R}\mathcal{R}\mathcal{R} f_0 + \cdots$$

$$= f_0 + t \mathcal{A} f_0 + \frac{t^2}{2} \mathcal{A}\mathcal{A} f_0 + \frac{t^3}{6} \mathcal{A}\mathcal{A}\mathcal{A} f_0 + \cdots$$

$$= \boxed{e^{t\mathcal{A}} f_0} = \mathcal{E} f_0$$

Baez, J.C. and Biamonte, J.D., 2018. Quantum techniques in stochastic mechanics. World Scientific.

$$\boxed{\mathcal{R}} = \boxed{\rightarrow} - \boxed{\downarrow}$$

$$f = f_0 + \boxed{\mathcal{R}}\, f$$

$$f + \boxed{\downarrow}\, f = f_0 + \boxed{\rightarrow}\, f$$

The stochastic dynamics can be expressed as a path integral in terms of waiting operators

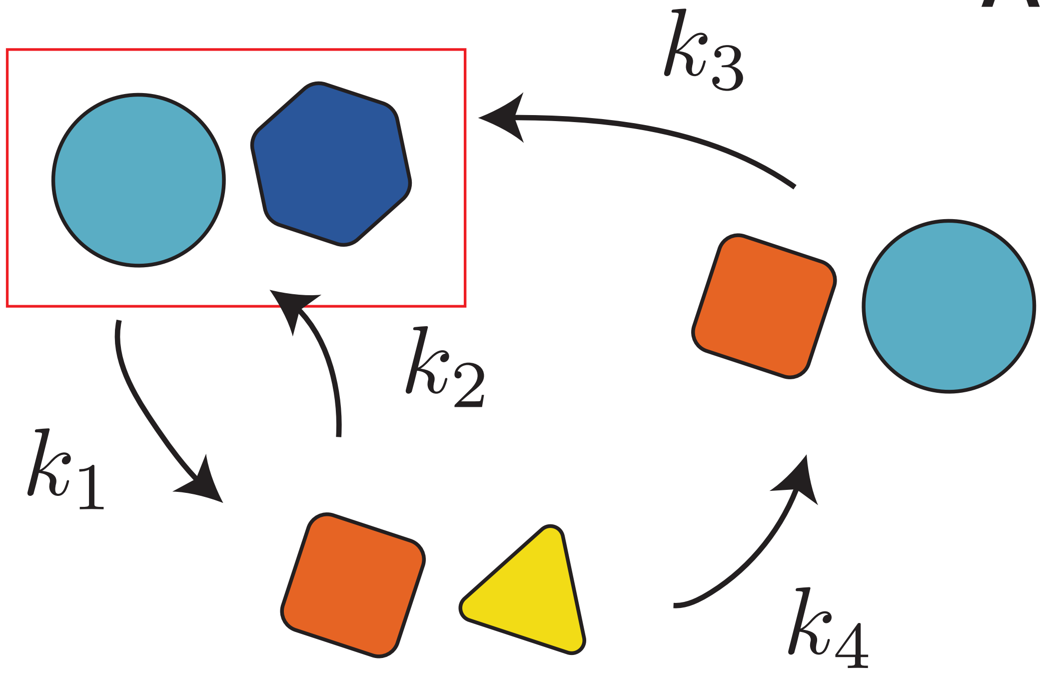
$$\boxed{\mathcal{W}} = 1 - \boxed{\downarrow} + \boxed{\downarrow}\boxed{\downarrow} - \boxed{\downarrow}\boxed{\downarrow}\boxed{\downarrow} + \cdots = \frac{1}{1 + \boxed{\downarrow}}$$

$$f = \boxed{\mathcal{W}}\, f_0 + \boxed{\mathcal{W}}\boxed{\rightarrow}\, f$$

$$= \boxed{\mathcal{W}}\, f_0 + \boxed{\mathcal{W}}\boxed{\rightarrow}\boxed{\mathcal{W}}\, f_0 + \boxed{\mathcal{W}}\boxed{\rightarrow}\boxed{\mathcal{W}}\boxed{\rightarrow}\boxed{\mathcal{W}}\, f_0 + \cdots$$

$$\boxed{\mathcal{W}}\, f_0 = e^{-t\boxed{\mathcal{A}^-}} f_0$$

Orit-Muñoz, A. 2021. Combinatorics and stochasticity for chemical reaction networks. PhD Thesis.



A system has a stationary Poisson distribution if, and only if, it is complex-balanced

$$k_1 \varepsilon_A \varepsilon_B = k_2 \varepsilon_C \varepsilon_D + k_3 \varepsilon_A \varepsilon_C$$

$$\mathcal{A}\, e^{\varepsilon_A z_A + \varepsilon_B z_B + \varepsilon_C z_C + \varepsilon_D z_D} = 0$$

$$p_{i,j,k,l} \propto \frac{\varepsilon_A^i}{i!} \frac{\varepsilon_B^j}{j!} \frac{\varepsilon_C^k}{k!} \frac{\varepsilon_D^l}{l!}$$

$$0 = \mathcal{A} e^{\varepsilon z} = e^{\varepsilon z} \sum_{c \in C} z^c \left(\sum_{r:\pi_r = c} k_r \varepsilon^{\rho_r} - \sum_{r:\rho_r = c} k_r \varepsilon^c \right)$$

$$\sum_{r:\pi_r = c} k_r \varepsilon^{\rho_r} = \sum_{r:\rho_r = c} k_r \varepsilon^c$$

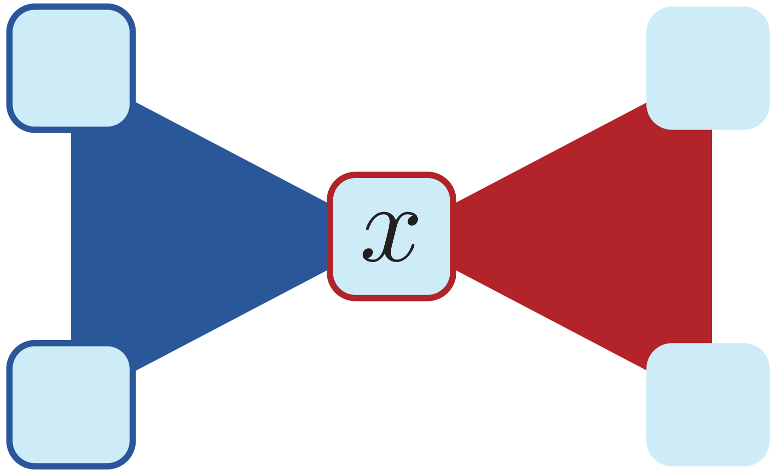
Anderson, D.F., Craciun, G. and Kurtz, T.G., 2010. Product-form stationary distributions for deficiency zero chemical reaction networks. Bulletin of mathematical biology, 72(8), pp.1947-1970.
Cappelletti, D. and Wiuf, C., 2016. Product-form poisson-like distributions and complex balanced reaction systems. SIAM Journal on Applied Mathematics, 76(1), pp.411-432.
Orit-Muñoz, A. 2021. Combinatorics and stochasticity for chemical reaction networks. PhD Thesis.

Factorial moments are expectation values of permutations

$$\mu_3 = (3 \cdot 2 \cdot 1)p_3 + (4 \cdot 3 \cdot 2)p_4 + (5 \cdot 4 \cdot 3)p_5 + \cdots = \sum_{n \geq 3} \frac{n!}{(n-3)!} p_n$$

$$p_n = e^{-\lambda} \frac{\lambda^n}{n!} \qquad \mu_n = \lambda^n$$

$$\frac{d\mu_x}{dt} = \sum_{r \in R} k_r \left(\sum_{y \leq \pi_r} \binom{\pi_r}{y} \frac{x!}{(x-y)!} \mu_{x-y+\pi_r} - \sum_{y \leq \rho_r} \binom{\rho_r}{y} \frac{x!}{(x-y)!} \mu_{x-y+\rho_r} \right)$$



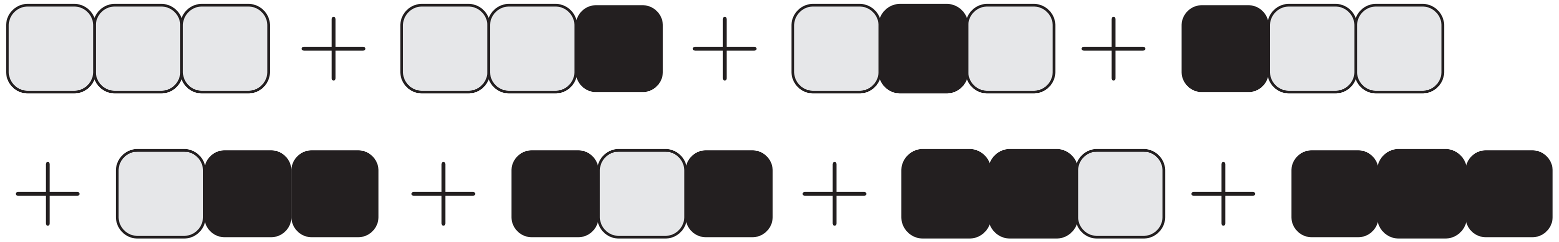
Smith, E. and Krishnamurthy, S., 2017. Flows, scaling, and the control of moment hierarchies for stochastic chemical reaction networks. Physical Review E, 96(6), p.062102.

$$m = \mu_0 + \mu_1 z + \frac{\mu_2}{2} z^2 + \frac{\mu_3}{6} z^3 + \cdots = \sum_{n \in \mathbb{N}} \frac{\mu_n}{n!} z^n$$

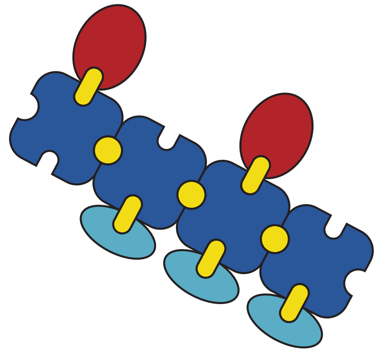
$$\boxed{m(z) = f(z+1)} \qquad f = e^{\lambda(z-1)} \quad m = e^{\lambda z}$$

The generating functions of factorial moments and probability have a simple relationship

$$(z+1)^3 = 1 + 3z + 3z^2 + z^3$$



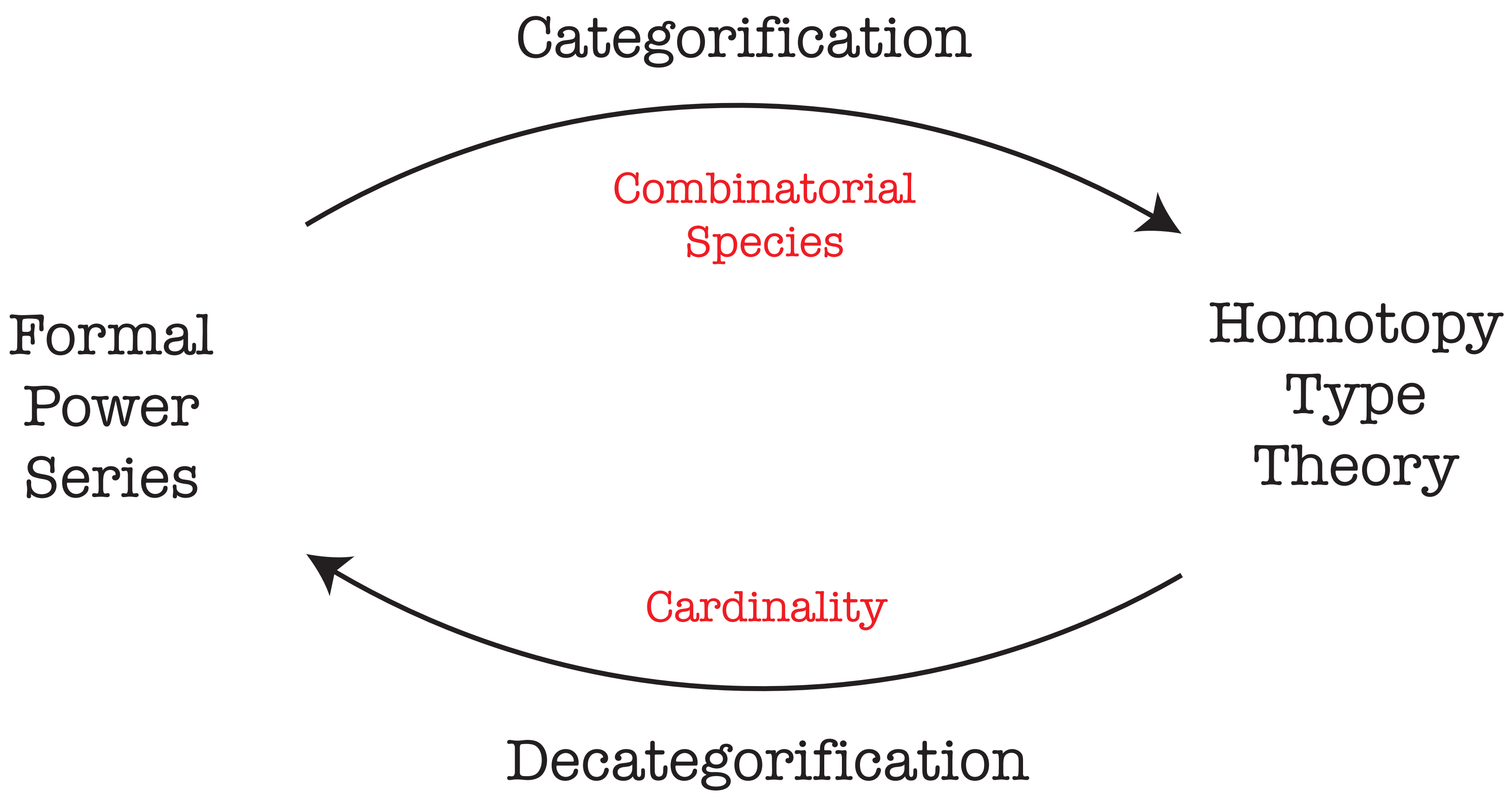
$$\frac{dm}{dt} = \sum_{r \in R} k_r \left((z+1)^{\varrho_r^+} - (z+1)^{\varrho_r^-} \right) \partial_z^{\varrho_r^-} m$$



Behr, N., Duchamp, G.H. and Penson, K.A., 2017. Combinatorics of chemical reaction systems. arXiv preprint arXiv:1712.06575.
Orit-Muñoz, A. 2021. Combinatorics and stochasticity for chemical reaction networks. PhD Thesis.

The Barrier of Objects

"In Nature, interaction involves objects directly and never by a numerical value describing them. Stepping outside of conventional dynamical systems requires taking this observation seriously."
W. Fontana & L. Buss, 1992



W. Fontana and L. W. Buss. 1996. "The barrier of objects: From dynamical systems to bounded organizations."
In Boundaries and Barriers. J. Casti and A. Karlqvist (eds.), Pp. 56–116. Reading MA: Addison-Wesley.

Acknowledgements

Erik Winfree

Walter Fontana

Acknowledgements

William Poole
Manoj Gopalkrishnan
Thomas Ouldridge
David Anderson
Daniele Cappelletti
Robert Johnson
Fangzhou Xiao
Damien Woods
Chris Thachuck
Dave Doty
Stephan Badelt
Frits Dannenberg
Héctor Medina

Caltech

Acknowledgements

Manuel Razo
Porfirio Quintero
Enrique Amaya
Manolo Flores
David Larios
Alejandro Granados
Jorge Castellanos

Santa Fe Institute

Acknowledgements

Acknowledgements

No Soy Poeta

Dicen que no soy poeta
tiene razón quien lo diga
Tan solo escribo versos
para bendecir a Dios
o elogiar a una hormiga

Pascual Ortiz Saucedo