Thermodynamics of small systems with emergent structures

Jan Korbel,1,2‡ Simon David Lindner,1,2§ Rudolf Hanel,1,2‡ and Stefan Thurner1,2,3,4,5

1Section for the Science of Complex Systems, CeMSIIS,
Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria

2Complexity Science Hub Vienna, Josefstädterstrasse 39, A-1080 Vienna, Austria

3Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

4HASA, Schlossplatz 1, A-2361 Laxenburg, Austria

(Dated: April 15, 2020)

We derive the entropy for a closed system of particles that can form structures, molecules in the simplest case. The entropy differs from the Boltzmann-Gibbs entropy by a term that captures the molecule states. For large systems the approach is equivalent to the grand-canonical ensemble. For small systems large molecules start to play a dominant role. The number of molecules becomes a key quantity and appears explicitly in the second law of thermodynamics and in fluctuation theorems. We show that the fully connected Ising model with emergent structures exhibits a first-order phase transition.

Ludwig Boltzmann defined entropy as the logarithm of state multiplicity. The multiplicity of independent (but possibly interacting) systems is typically given by multinomial factors which lead to the well-known particular form of Boltzmann-Gibbs entropy and to the exponential growth of sample space as a function of the degrees of freedom. In recent decades much attention was given to systems with long-range and co-evolving interactions. These systems are sometimes called complex systems [1]. Many complex systems do not exhibit an exponential growth of sample space [2–5]. For correlated systems sample space typically grows sub-exponentially [6–13], Systems with super-exponential phase space growth were recently identified as those capable of forming emergent structures [5] [14]. A typical example of this kind are complex networks [1] [15], where complex behavior can lead to ensemble in-equivalence [16]. The most prominent example of systems forming emergent structures are chemical reaction networks [17–19]. The usual approach to model chemical reactions–where we have free particles and molecules that are composed of these– is through the grand-canonical ensemble with particle reservoirs such that the number of particles is conserved on average.

In this paper we present an alternative approach for closed systems where particles can interact and form structures. A toy example, the magnetic coin model, was recently introduced in [14], where n coins have two possible states (head and tail). In addition, since coins are magnetic, they can also form a third state, any two may create a bond state. The sample space of this model, $W(n)$, grows super-exponentially, $W(n) \sim n^{n/2} e^{2\sqrt{n}} \sim e^{n \log n}$. Here we generalize this model to arbitrary molecules sizes and to an arbitrary number of states. We then derive the entropy from the corresponding log-multiplicity and use it to compute thermodynamic quantities as the Helmholtz free energy. Compared to Boltzmann-Gibbs entropy there appears an extra term that captures the molecule states. With the help of stochastic thermodynamics we formulate the second law for systems with emergent structures or molecule states. We apply our findings to the well-known non-equilibrium Crooks' fluctuation theorem and the Jarzynski equality. Both theorems relate the probability distribution of the stochastic work done on a non-equilibrium system to thermodynamic state variables, such as the Helmholtz free energy and the temperature of the initial and final states. Unlike the grand canonical ensemble, where the number of molecules appears as stochastic chemical work in the first law, here the number of molecules becomes a state variable, naturally appearing in the second law. Finally, we discuss the fully connected Ising model where we allow the formation of molecules. Molecules feel neither the spin-spin interaction nor the external magnetic field. We show that the second-order transition in the usual fully connected Ising model changes to first-order when emergent structures are present.

Let us consider a system of n particles that can attain states from the set, $\{s_{1}^{(1)}, \ldots, s_{m_{1}}^{(1)}\}$. Here the superscript (1) means that the states belong to single particles, $m_{1}$ denotes the number of these states. A typical set of states can be the spin of the particle $\{\uparrow, \downarrow\}$, or a set of energy levels. Any two particles can create a two-particle state (molecule) with states $\{s_{i}^{(2)}, \ldots, s_{m_{2}}^{(2)}\}$. The particles can also form larger molecules up to a maximal size, $m$. Let us consider that $m$ is a fixed number, $m \leq n$. Generally, molecules of size $j$ have states $\{s_{i}^{(j)}, \ldots, s_{m_{j}}^{(j)}\}$. Let us consider a configuration, where we have $n_{i}^{(j)}$ molecules in state $s_{i}^{(j)}$. Altogether, we have $\sum_{ij} n_{i}^{(j)} = n$ particles. The entropy of this state is given by the Boltzmann formula

$$S\left(\left\{n_{i}^{(j)}\right\}\right) = \log W\left(\left\{n_{i}^{(j)}\right\}\right),$$

where $W$ is the multiplicity of the state. The multiplicity is the number of all distinct configurations of the system. The number of all configurations is simply $n!$, however,
many configurations represent the same micro-state. The number of configurations giving the same state can be expressed as the product of configurations with the same state for each \( s_i^{(j)} \). Let’s begin with the particles that do not form molecule states. The number of equivalent representations for one distinct state is \( \binom{n_i^{(j)}}{\binom{n_i^{(j)}}{j}} \), which corresponds to the number of permutations of all particles in the same state. For the molecule states, one can think about equivalent representations of one microstate in two steps: first permute all molecules, which gives \( \binom{n_i^{(j)}}{j} \) possibilities. Second, permute the particles in the molecule, which gives \( j! \) possibilities for every molecule, so we end up with \( (j!) \binom{n_i^{(j)}}{j} \) combinations. As an example, consider the case of four particles. The number of distinct micro-states for the case of two possible states is given by \( W = 4!/(2^2!2!) = 6 \), because we have states \( \{1, 1, 2, 2\}, \{1, 2, 1, 2\}, \{1, 2, 2, 1\}, \{2, 1, 2, 1\}, \{2, 1, 1, 2\} \), and \( \{2, 2, 1, 1\} \). Now imagine that these four particles form two molecules. The number of distinct micro-states is just \( W = 4!/(2!(2!)^2) = 3 \), i.e., states \( \{1, 1, 2, 2\}, \{1, 2, 1, 2\} \) and \( \{1, 2, 2, 1\} \). Note that compared to the previous example, we can identify states that can be achieved by re-labelling states \( 1 \leftrightarrow 2 \).

Thus, altogether the multiplicity of state \( s_i^{(j)} \) is \( (n_i^{(j)})!/(j!)^{n_i^{(j)}} \), and we can express the multiplicity as

\[
W(\{n_i^{(j)}\}) = \frac{n!}{\prod_{ij} \binom{n_i^{(j)}}{j} \binom{n_i^{(j)}}{j}^n}.
\]

Using Stirling formula for large \( n \) we get the entropy

\[
S(\{n_i^{(j)}\}) = n \log n - \sum_{ij} \left( n_i^{(j)} \log n_i^{(j)} + n_i^{(j)} \log j \right).
\]

Using the normalization condition, \( n = \sum_{ij} n_i^{(j)} \), and combining the first term with the remaining terms, we get the entropy in terms of ratios \( n_i^{(j)}/n \)

\[
\frac{S(\{n_i^{(j)}\})}{n} = -\sum_{ij} \frac{n_i^{(j)}}{n} \log \left( \frac{n_i^{(j)}}{n} \right) - \frac{n_i^{(j)}}{n} \log \left( \frac{j!}{n^{j-1}} \right).
\]

The \( n \)-particle entropy appears to be a multiple of a single-molecule entropy with ‘probabilities’, \( n_i^{(j)}/n = p_i^{(j)} \). Normalization is given by \( \sum_{ij} n_i^{(j)} = n \), and therefore, the quantity \( p_i^{(j)} = \sum_{ij} n_i^{(j)}/n \) can be interpreted as the probability that a particle is a part of a molecule in state \( s_i^{(j)} \). Finally, we express the entropy per particle as

\[
S(P) = -\sum_{ij} p_i^{(j)} \log p_i^{(j)} - \sum_{ij} p_i^{(j)} \log \left( \frac{j!}{n^{j-1}} \right),
\]

where the first term corresponds to the usual Boltzmann-Gibbs entropy. The second term arises from the presence of emergent structures. Note that the number of particles appears explicitly in the second term. There are two situations where the second term plays different roles: When all molecules are of the same type, say \( j \), then the second term is just an additive constant, \( \frac{1}{2} \log \left( \frac{j!}{n^{j-1}} \right) \), and the entropy is equivalent to Boltzmann-Gibbs entropy. When there are different types of molecule, which makes the sum over \( j \) non-trivial, the second term plays an important role: First, entropy is no longer maximized by the uniform distribution (from an axiomatic point of view, it does no longer fulfil the second Shannon-Khinchin axiom [21, 22]). Second, the explicit dependence on \( n \) changes the importance of the second term when re-scaling the system. Moreover, explicit dependence on system size, \( n \), destroys additivity (sometimes formulated as the fourth Shannon-Khinchin axiom) even in its generalized form [2] [23].

Up to here, we assumed an infinite range in the interactions between particles, which is unrealistic for chemical reactions where atoms only form molecules if they are within a short range. A simple correction is to divide the system into a fixed number of boxes; particles in the same box can form molecules, particles in different boxes can’t. We begin by calculating the multiplicity for two boxes. For simplicity, assume that they both contain \( n/2 \) particles. The multiplicity of a system with two boxes, \( W(\{n_i^{(j)}\}) \), is given by the sum of all possible divisions of \( \{n_i^{(j)}\} \) molecules with state \( s_i^{(j)} \) into the first box (containing \( \{1n_i^{(j)}\} \) molecules) and the second box (containing \( \{2n_i^{(j)}\} \) molecules), such that \( n_i^{(j)} = 1n_i^{(j)} + 2n_i^{(j)} \). The multiplicity is therefore given as

\[
\tilde{W}(\{n_i^{(j)}\}) = \sum_{1n_i^{(j)} + 2n_i^{(j)} = n_i^{(j)}} W(\{1n_i^{(j)}\}) W(\{2n_i^{(j)}\}),
\]

where \( W \) is the multiplicity in Eq. (2). The dominant contribution to the sum is given by the term, where \( 1n_i^{(j)} = 2n_i^{(j)} = n_i^{(j)}/2 \), so we can approximate the multiplicity as \( \tilde{W}(\{n_i^{(j)}\}) \approx W(\{n_i^{(j)}/2\})^2 \). Similarly, for \( b \) boxes we obtain that the multiplicity is

\[
\tilde{W}(\{n_i^{(j)}\}) = W(\{n_i^{(j)}/b\})^b = \frac{([n/b]!)^b}{\prod_{ij} \binom{[n_i^{(j)}/b]}{j} !} \binom{[n_i^{(j)}/b]}{j} ! \binom{n_i^{(j)}}{j} !,
\]

and the entropy becomes

\[
S(P) = -\sum_{ij} p_i^{(j)} \log p_i^{(j)} - \sum_{ij} p_i^{(j)} \log \left( \frac{j!}{c^{j-1}} \right),
\]

where \( c = n/b \) is the concentration of particles. A natural assumption is that the concentration is constant when approaching the thermodynamic limit, i.e.,
\( b(n) = \kappa n \), but results are still valid for any function \( b(n) \).

Let us now focus on equilibrium thermodynamics obtained, for example, by considering the maximum entropy principle. Let’s assume a Hamiltonian of the form

\[
H\left(\{n_{ij}\}\right) = \sum_{ij} n_{ij} \epsilon_{ij} = n \sum_{ij} p_{ij} \epsilon_{ij}.
\]  

Using Lagrange multipliers we maximize the functional

\[
S(P) - \alpha \left( \sum_{ij} j p_{ij} - 1 \right) - \beta \left( \sum_{ij} p_{ij} \epsilon_{ij} - U \right).
\]

Maximization leads to the set of equations,

\[
-\log \hat{p}_{ij} - 1 - \log \left( j^{\frac{1}{\beta-1}} \right) - \alpha j - \beta \epsilon_{ij} = 0,
\]

and the resulting distribution is

\[
\hat{p}_{ij} = \frac{1}{j!} \exp \left( -j \alpha - 1 - \beta \epsilon_{ij} \right).
\]

The Lagrange parameter, \( \alpha \), is obtained from the normalisation

\[
\sum_{j} \left[ \frac{e^{j-1}}{e(j-1)!} \sum_{i} e^{-\beta \epsilon_{ij}} \right] = 1,
\]

which is a polynomial equation of order maximally \( m \). The connection with thermodynamics follows through Eq. (11). By multiplying with \( \hat{p}_{ij} \) and summing over \( i, j \), we get

\[
S(P) - \sum_{ij} \hat{p}_{ij} - \alpha - \beta U = 0.
\]

Note that \( \sum_{ij} \hat{p}_{ij} = \sum_{ij} \hat{n}_{ij}/n = M/n \) is the number of molecules divided by the number of particles in the system. The number of molecules is

\[
M = n \sum_{j} \frac{e^{j-1}}{j!} \exp(-j \alpha) \sum_{i} \exp \left( -\beta \epsilon_{ij} \right).
\]

The Helmholtz free energy is thus obtained as

\[
F = U - \frac{1}{\beta} S = -\frac{\alpha}{\beta} - \frac{1}{\beta} M.
\]

The thermodynamic potential, \( \Omega = -\alpha/\beta \), plays the role of the grand potential (also Landau free energy), as known from the grand-canonical ensemble. The second term is the non-trivial contribution to free energy from the emergent structures of molecule states. The apparent difference between this approach and the usual grand-canonical ensemble is that while in the grand-canonical ensemble the number of molecules fluctuates around its average, here it is fixed and determined by the canonical ensemble, without the presence of a particle reservoir. Note that the division into boxes described above creates a situation that is similar to the grand-canonical ensemble, in the sense that subsystems can interchange particles with a reservoir. However, for the case of the exact calculation the total number of particles is strictly conserved.

To compare the presented exact approach with the grand-canonical ensemble, consider the simple chemical reaction \( 2X \rightleftharpoons X_{2} \). Without loss of generality, assume that free particles carry some energy \( \epsilon \). To make the comparison we introduce (similar to the case of the grand-canonical ensemble) the reduced sample space that is obtained by dividing the sample space volume by the factor \( n! \), i.e., \( W(n) = W(n)/n! \). Consequently, the reduced entropy, \( S(n) = S(n) - n \log n \), becomes extensive with respect to the number of particles. The factor \( 1/n! \) is typically introduced because of the indistinguishability of particles to avoid the Gibbs paradox. There are, however, other approaches for which the factor \( -n \log n \) appears naturally in the entropy for both distinguishable and indistinguishable particles [24].

Let us finally introduce the free energy using the reduced entropy as

\[
F = U - TS = F + T n \log n.
\]

\( F \) is equivalent to \( F \) in the sense that the response coefficients, as for example the specific heat or the compressibility remain the same, because they are expressed through second derivatives of the free energy and therefore the term \( T n \log n \) does not contribute. Note that
while $F$ decreases with temperature, $\mathcal{F}$ increases. In the Supplementary Material (SM) [24], we calculate the Helmholtz free energy for both approaches and show that the free energy becomes the same for the usual grand-canonical ensemble approach and the exact calculation with a strictly conserved number of particles (squares) and the grand-canonical approach (triangles) become the same. This is also demonstrated in Fig. 1 where we see that for large $n$, free energies for the exact computation with a conserved number of particles (squares) and the grand-canonical approach (triangles) become the same. The equivalence is even more evident from the inset, where we show the ratio of the free energies $\mathcal{F}_{\text{exact}}/\mathcal{F}_{\text{GC}} - 1$. Obviously, this term vanishes with an increasing number of particles regardless of temperature.

We now generalize the results to the case of an arbitrary non-equilibrium molecule state given by a distribution, $\tilde{p}_i^{(j)}(t) \equiv \tilde{p}_i^{(j)}_M(t)$. Assume that the evolution of the probability distribution is defined by a first-order Markovian linear master equation, as usually done in stochastic thermodynamics [33, 34]

\begin{equation}
\dot{p}_i^{(j)} = \sum_{kl} w_{ik}^{jl} \tilde{p}_k^{(l)} = \sum_{kl} \left( w_{ik}^{jl} - w_{kj}^{li} \right) \tilde{p}_i^{(j)},
\end{equation}

where $w_{ik}^{jl}$ are the transition rates. Note that normalization leads to $\sum_{ij} j \tilde{p}_i^{(j)} = 0$. Assuming detailed balance such that probability currents vanish, $w_{ik}^{jl} \tilde{p}_k^{(l)} = w_{ik} \tilde{p}_i^{(j)}$, implies that the distribution obtained from the maximum entropy principle is indeed given by the equilibrium distribution. From this we get detailed balance in the form

\begin{equation}
\frac{w_{ij}^{kl}}{w_{ki}^{ij}} = \frac{1}{l!} \frac{1}{i!} \frac{1}{j!} \exp \left[ \alpha(l-j) + \beta \left( \epsilon_l^{(i)} - \epsilon_i^{(j)} \right) \right].
\end{equation}

The time derivative of entropy is

\begin{equation}
\frac{dS}{dt} = - \sum_{ij} p_i^{(j)} \log p_i^{(j)} - \sum_{ij} \dot{p}_i^{(j)} \log \left( \frac{\dot{p}_i^{(j)}}{p_i^{(j)}} \right).
\end{equation}

Note that the second term does not vanish. It is equal to the change in the number of molecules per particle. Using the master equation and some straightforward calculations, we end up with the second law of thermodynamics for systems with molecules

\begin{equation}
\frac{dS}{dt} = - \frac{dM}{dt} + \beta \dot{Q} + \dot{S}_t,
\end{equation}

where $\dot{Q}$ is the heat flow and $\dot{S}_t$ is the non-negative entropy production, see SM [25]. Again, note the difference between the grand-canonical ensemble and the present approach: the number of molecules enters explicitly in the second law of thermodynamics, playing a similar role as the entropy. It is a state variable, independent of

![FIG. 2: Magnetization of the fully connected Ising model with molecule states for $n = 50$ and $n = 200$ for $J = 1$. Results for the Mean-field approximation (solid lines) are in very good agreement with Monte-Carlo simulations (symbols). The inset shows the well-known result for the fully connected Ising model without molecule states. For the system without molecules we observe the known second-order phase transition. For the system with molecules, the critical temperature decreases with number of molecules and the phase transition becomes first-order.](image)

the process. This has consequences for well-known non-equilibrium relations, such as the Crooks’ fluctuation theorem [35, 36]

\begin{equation}
\frac{P(W)}{P(-W)} = \exp (\beta(W - \Delta F) + \Delta M),
\end{equation}

and the Jarzynski equality [37]

\begin{equation}
\langle \exp(-\beta W) \rangle = \exp(-\beta \Delta F + \Delta M),
\end{equation}

for systems with emergent molecule states. Here, $P(W)$ is the probability of performing work along a path, starting at initial state, $A$, and ending in final state $B$; $P(W)$ is the probability of performing the work $W = -W$ from the initial state, $B$, ending in the final state, $A$; $\Delta F$ is the free energy difference between states $A$ and $B$, and $\Delta M$ is the difference between the number of molecules in the states $A$ and $B$. For a canonical ensemble, i.e., a constant number of molecules, $\Delta M = 0$, we recover the usual fluctuation theorems. For the case of the grand canonical ensemble, the stochastic work is composed of mechanical and chemical work, $W = W_{\text{mech}} + W_{\text{chem}}$. Both, mechanical and chemical work depend on the particular process of getting from initial state $A$ to the final $B$.

Finally, we discuss an example of a system with molecule states. Consider the fully connected Ising model [42–45] with a Hamiltonian that allows for pos-
sible molecule states
\[
H(\sigma_i) = -\frac{J}{n-1} \sum_{i \neq j, \text{free}} \sigma_i \sigma_j - h \sum_{j \text{free}} \sigma_j.
\]

(M24)

Molecule states do neither feel the spin-spin interaction nor the external magnetic field. Therefore, the sum only extends over free particles. In the mean-field approximation we use magnetization as
\[
m = \frac{1}{n} \sum_{i \neq j} \sigma_i
\]
and express the Hamiltonian as
\[
H^{\text{MF}}(\sigma_i) = -(Jm + h) \sum_{j, \text{free}} \sigma_j.
\]
The self-consistency equation \(m = -\frac{\partial \mathcal{F}}{\partial h}|_{h=0}\) leads to an equation for \(m\), which is calculated numerically (see SM) and shown in Fig. 2. We see that contrary to the mean-field approximation of the usual fully connected Ising model (without molecule states), the phase transition is no longer second-order, but becomes first-order. There is a bifurcation where the solutions for \(m = 0\) and \(m > 0\) are stable. The second-order transition is recovered for the \(c \to 0\) case, where no molecule states are allowed. The critical temperature is shifted towards zero for increasing \(c\). We performed Monte-Carlo simulations to check the result of the mean-field approximation, see SM. 25. In SM, we also discuss a number of further examples, such as the magnetic gas, or the size-dependent chemical potential.

Conclusions. We presented a way to establish thermodynamics of structure-forming systems (molecules in particular) based on the canonical ensemble with a modified entropy that is obtained by the proper counting of the system’s configurations. The approach is an alternative to the grand-canonical ensemble that yields identical results for large systems. However, there are also important differences that might have important consequences. The difference becomes significant, especially for small systems, i.e., systems where the interaction range becomes comparable with the system size. A typical example for such systems are chemical reactions at small scales. The presented results might be of direct use for chemical nano-motors 40.

In the presented approach the number of molecules serves as a fundamental macro-state variable. It naturally appears in the second law of thermodynamics. We showed that this fact has profound consequences, for example for fluctuation theorems, or for the order of phase-transitions. Further, the approach is fully consistent with the non-equilibrium description based on stochastic thermodynamics, which allows us to apply the results for arbitrary far-from-equilibrium systems at any scale. A natural question is now how the presented results are related to the well-known statistical physics of chemical reactions 38-41 where systems are composed of more than one type of atom.

We acknowledge support from the Austrian Science fund Project I 3073 and Austrian Research Promotion agency FFG Project 857136. We thank Tuam Pham for helpful discussions.

* jan.korbela@meduniwien.ac.at
† lindner.sd@gmail.com
‡ rudolf.hanel@meduniwien.ac.at
§ stefan.thurner@meduniwien.ac.at

[1] S. Thurner, P. Klimek, and R. Hanel. Introduction to the Theory of Complex Systems. Oxford University Press, Oxford (2018).
[2] R. Hanel and S. Thurner. A comprehensive classification of complex statistical systems and an axiomatic derivation of their entropy and distribution functions. Europhys. Lett. 93 (2011) 20006.
[3] R. Hanel and S. Thurner. When do generalized entropies apply? How phase space volume determines entropy. Europhys. Lett. 96 (2011) 50003.
[4] R. Hanel, S. Thurner, and M. Gell-Mann, How multiplicity determines entropy and the derivation of the maximum entropy principle for complex systems. Proc. Natl. Acad. Sci. U.S.A. 111, 6905 (2014).
[5] J. Korbel, R. Hanel and S. Thurner, Classification of complex systems by their sample-space scaling exponents. New J. Phys. 20 (2018) 093007.
[6] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. J. Stat. Phys. 52 (1988) 479.
[7] A.K. Rajagopal, Dynamic linear response theory for a nonextensive system based on the Tsallis prescription. Phys. Rev. Lett. 76 (1996) 3469.
[8] G. Kaniadakis, Statistical mechanics in the context of special relativity. Phys. Rev. E 66 (2002) 056125.
[9] P. Jizba and T. Arimitsu. The world according to Renyi: thermodynamics of multifractal systems. Ann. Phys. 312 (2004) 17.
[10] C. Anteneodo and A. R. Plastino, Maximum entropy approach to stretched exponential probability distributions. J. Phys. A 32(7) (1999) 1089.
[11] E. Lutz and F. Renzoni. Beyond Boltzmann-Gibbs statistical mechanics in optical lattices. Nat. Phys. 9 (2013) 615-619.
[12] A. Dechant, D.A. Kessler and E. Barkai, Deviations from Boltzmann-Gibbs statistics in confined optical lattices. Phys. Rev. Lett. 115 (2015) 173006.
[13] P. Jizba and J. Korbel, Maximum Entropy Principle in Statistical Inference: Case for Non-Shannonian Entropies. Phys. Rev. Lett. 122 (2019) 120601.
[14] H. J. Jensen, R. H. Pazuki, G. Pruessner and P. Tempesta, Statistical mechanics of exploding phase spaces: ontic open systems. J. Phys. A 51 (2018) 375002.
[15] V. Latora, V. Nicosia, and G. Russo, Complex networks: principles, methods and applications. Cambridge University Press, Cambridge, (2017).
[16] T. Squartini, J. de Mol, F. den Hoolander and D. Garlaschelli, Breaking of Ensemble Equivalence in Networks. Phys. Rev. Lett. 115 (2015) 268701.
[17] C. Berge, Graphs and Hypergraphs. North-Holland mathematical library (1973).
[18] O. N. Temkin, A. V. Zeigarnik, and D. G. Bonchev, Chemical reaction networks: a graph-theoretical approach. CRC Press (1996).
[19] C. Flamm, B. M. R. Stadler, and P. F. Stadler, Generalized topologies: hypergraphs, chemical reactions, and biological evolution. Advances in Mathematical Chemistry and Applications. Bentham Science Publishers, (2015) 300-328.

[20] L. Boltzmann, Über das Arbeitsquantum, welches bei chemischen Verbindungen gewonnen werden kann. Anzalen der Physik 258(5) (1884) 39-72.

[21] C. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. 27 (1948) 379.

[22] A.I. Khinchin, Mathematical Foundations of Information Theory. Dover Publications, New York (1957).

[23] P. Tempesta, Beyond the Shannon-Khinchin formulation: The composability axiom and the universal-group entropy. Ann. Phys. 365 (2016) 180.

[24] R.H. Swendsen, Statistical Mechanics of Classical Systems with Distinguishable Particles. J. of Stat. Phys. 107 (2002) 11431166.

[25] See Supplemental Material at *** for finer technical details, which include Refs. [26-32].

[26] A. Z. Panagiotopoulos, Monte Carlo methods for phase equilibria of fluids. J. Phys.: Condens. Matt. 12(3) (1999) 25-52.

[27] S. Thurner, B. Corominas-Murtra, and R. Hanel, Three faces of entropy for complex systems: Information, thermodynamics, and the maximum entropy principle. Phys. Rev. E 96 (2017) 032124.

[28] B. Smit and D. Frenkel, Explicit expression for finite size corrections to the chemical potential. J. Phys.: Condens. Matt. 1 (1989) 8659.

[29] J. I. Siepmann, I. R. McDonald, and D. Frenkel, Finite-size corrections to the chemical potential. J. Phys.: Condens. Matt. 4(3) (1992) 679.

[30] G.-B. Jo, Y.-R. Lee, J.-H. Choi, C. A. Christensen, T. H. Kim, J. H. Thywissen, D. E. Pritchard, and W. Ketterle, Itinerant Ferromagnetism in a Fermi Gas of Ultracold Atoms. Science 325 (2009) 5947.

[31] R. A. Duine and A. H. MacDonald, Itinerant Ferromagnetism in an Ultracold Atom Fermi Gas. Phys. Rev. Lett. 95 (2005) 230403.

[32] E. P. Verlinde, On the Origin of Gravity and the Laws of Newton. J. High Energy Phys. 2011(4) (2011) 29.

[33] U. Seifert, Stochastic thermodynamics: principles and perspectives. Eur. Phys. J. B 64 (2008) 423431.

[34] M. Esposito and C. Van den Broeck, The Three Faces of the Second Law: I. Master Equation Formulation. Phys. Rev. E 82 (2010) 011143.

[35] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys. Rev. E 60(3) (1999) 2721.

[36] M. Esposito and C. Van den Broeck, Three Detailed Fluctuation Theorems. Phys. Rev. Lett. 104 (2010) 090601.

[37] C. Jarzynski, Nonequilibrium equality for free energy differences. Phys. Rev. Lett. 78 (1997) 2690.

[38] D. Chandler and L. R. Pratt, Statistical mechanics of chemical equilibria and intramolecular structures of nonrigid molecules in condensed phases. J. Chem. Phys. 65(8) (1976) 2925-2940.

[39] H. J. Kreuzer, Nonequilibrium thermodynamics and its statistical foundations. Clarendon Press, Oxford and New York (1981).

[40] P. T. Cummings and G. Stell, Statistical mechanical models of chemical reactions: Analytic solution of models of $A + B \rightleftharpoons AB$ in the Percus-Yevick approximation.
SUPPLEMENTARY MATERIAL

This supplementary material to the paper *Thermodynamics of systems with emergent molecule structures* contains additional information, mainly on details of analytical and numerical computations. It also contains more examples that we mention in the main text.

Equivalence of the exact calculation of the sample space with the grand-canonical ensemble in the thermodynamic limit

Here we show the equivalence of the presented approach with the grand-canonical ensemble in the thermodynamic limit and the limit of low concentrations. Let us consider a chemical reaction, \(2X \rightleftharpoons X_2\), with \(n\) particles. Let us denote the number of particles \(X\) as \(n_X\) and the number of molecules \(X_2\) as \(n_{X_2}\). Let us consider, without loss of generality, that free particles have an energy \(\epsilon\) and molecules have zero energy. In both approaches we shall use the reduced sample space,

\[
W(n) = \frac{W(n)}{n!}
\]  

(25)
to make the entropy extensive and to avoid the Gibbs paradox when the numbering of particles is changed.

Exact calculation:

Let us start with the Boltzmann formula for entropy, which reads

\[
S(n_X, n_{X_2}) = \log W(n_X, n_{X_2}) = \log \left( \frac{n!}{n_X! n_{X_2}!} \frac{1}{n!} \right).
\]  

(26)

Thus, \(n!\) cancels out. By applying Stirling’s formula, we obtain the entropy in the following form

\[
S(n_X, n_{X_2}) = -n_X \log n_X - n_{X_2} \log n_{X_2} - n_{X_2} \log 2.
\]  

(27)

By dividing the system into \(b\) boxes, we obtain a similar formula as in the main text,

\[
S(n_X, n_{X_2}) = -n_X \log n_X - n_{X_2} \log n_{X_2} - n_{X_2} \log 2/c,
\]  

(28)

where \(c = 1/b\) is the reduced concentration. Let us now find the equilibrium distribution of the system, i.e., maximize the entropy under the normalization constraint, \(n_X + 2n_{X_2} = n\), and the energy constraint, \(c n_X = U\). From this we get that

\[
n_X = \exp(1 - \alpha - \beta \epsilon),
\]  

(29)

\[
n_{X_2} = \frac{c}{2} \exp(1 - 2\alpha).
\]  

(30)

The Lagrange multiplier \(\alpha\) can be calculated from the normalization constraint

\[
\exp(1 - \alpha - \beta \epsilon) + 2 \frac{c}{2} \exp(1 - 2\alpha) = n.
\]  

(31)

We obtain two solutions of the quadratic equation, from which only one has a physical meaning, i.e.,

\[
\alpha = \log \left( \frac{2ce^{\beta \epsilon}}{-1 + \sqrt{1 + 4cne^{1+2\beta \epsilon}}/2c} \right).
\]  

(32)

Helmholtz free energy can be obtained as

\[
\mathcal{F} = -\frac{\alpha n}{\beta} - \frac{n_X + n_{X_2}}{\beta} = \frac{n}{\beta} \log \left( \frac{-1 + \sqrt{1 + 4cne^{1+2\beta \epsilon}}}{2ce^{\beta \epsilon}} \right) - \frac{n}{\beta} 2 \left( \frac{cne^{3\beta \epsilon + 1} + \sqrt{4cne^{2\beta \epsilon + 1} + 1 + 1}}{\sqrt{4cne^{2\beta \epsilon + 1} + 1 + 1}} \right).
\]  

(33)

For large \(n\), the first term becomes dominant, so we have

\[
\mathcal{F}(\beta, \epsilon, n, c) \approx -\frac{\alpha n}{\beta} = \frac{n}{\beta} \log \left( \frac{-1 + \sqrt{1 + 4cne^{1+2\beta \epsilon}}}{2ce^{\beta \epsilon}} \right).
\]  

(34)
**Grand-canonical ensemble:** Let’s now compare the exact result with the usual approach using the grand-canonical ensemble. The partition function of the grand-canonical ensemble can be expressed as

\[
Z = \sum_{n_x, n_x^2 = 0}^{\infty} \frac{1}{n_x!} \exp(-\beta (\epsilon - \mu_X) n_x) \frac{1}{n_x^2!} \exp(\beta \mu_{X^2} n_x^2) = \exp \left( e^{\beta \mu_{X^2}} + e^{-\beta (\epsilon - \mu_X)} \right),
\]

where \( \mu_X \) and \( \mu_{X^2} \) are the chemical potentials. From the Gibbs-Duhem relation, we get that \( \mu_{X^2} = 2 \mu_X \). We denote the chemical potential by \( \mu \). The average number of particles can be calculated as

\[
\langle n \rangle = \frac{\partial \log Z}{\partial \mu} = 2e^{2\beta\mu} + e^{\beta(\mu - \epsilon)}.
\]

This relation serves as an equation for \( \mu \), which has the same form as Eq. (31), and the solution can be found as

\[
\mu = \frac{\log \left( \frac{-1 + \sqrt{1 + 8\langle n \rangle e^{2\beta\epsilon}}}{4e^{\beta\epsilon}} \right)}{\beta}.
\]

Helmholtz free energy can be expressed from the grand potential \( \Omega = -\beta \log Z \) as \( \mathcal{F} = \Omega + \mu \langle n \rangle \). By plugging in the chemical potential, we obtain that

\[
\mathcal{F} = \langle n \rangle \frac{\beta}{\beta} \log \left( \frac{-1 + \sqrt{1 + 8\langle n \rangle e^{2\beta\epsilon}}}{4e^{\beta\epsilon}} \right) - \frac{e^{-2\beta\epsilon} \left( 4\langle n \rangle e^{2\beta\epsilon} + \sqrt{8\langle n \rangle e^{2\beta\epsilon} + 1} - 1 \right)}{8\beta}.
\]

For large \( \langle n \rangle \), the fluctuations of particles diminish, so only the states with the average number of particles become relevant and we can set \( \langle n \rangle = n \). Moreover, the first term becomes dominant, so

\[
F(\beta, \epsilon, n) \approx \mu = \frac{\langle n \rangle}{\beta} \log \left( \frac{-1 + \sqrt{1 + 8\langle n \rangle e^{2\beta\epsilon}}}{4e^{\beta\epsilon}} \right),
\]

and we see that the free energies of both approaches coincide for \( \epsilon = 2 \). Thus, in thermodynamic limit, the exact calculation for low concentrations is equivalent to the grand-canonical ensemble.

**Derivation of the second law of thermodynamics for non-equilibrium molecule-forming systems**

The time derivative of entropy can be expressed as

\[
\frac{dS}{dt} = -\sum_{i,j} p_i^{(j)} \log p_i^{(j)} - \sum_{i,j} p_i^{(j)} - \sum_{i,j} p_i^{(j)} \log \left( \frac{j!}{(j-1)!} \right).
\]

By plugging in the master equation we can further obtain that

\[
\dot{S} = -\dot{\mathcal{M}} + \sum_{ijkl} w_{ij}^{kl} p_i^{(l)} \log p_i^{(j)} - \sum_{ijkl} w_{ij}^{kl} p_i^{(l)} \left( \frac{j!}{(j-1)!} \right)
\]

\[
= -\dot{\mathcal{M}} + \frac{1}{2} \sum_{ijkl} (w_{ij}^{kl} p_i^{(j)} - w_{ik}^{jl} p_i^{(l)}) \log \left( \frac{p_i^{(j)}}{p_k^{(l)}} \right) + \frac{1}{2} \sum_{ijkl} (w_{ij}^{kl} p_i^{(j)} - w_{ik}^{jl} p_i^{(l)}) \log \left( \frac{k!}{j!} \right) \cdot \frac{1}{e^{\beta}}
\]

\[
= -\dot{\mathcal{M}} + \frac{1}{2} \sum_{ijkl} (w_{ik}^{jl} p_i^{(j)} - w_{ik}^{j} p_i^{(l)}) \log \left( \frac{w_{ik}^{jl} p_i^{(j)}}{w_{ik}^{j} p_i^{(l)}} \right) + \frac{1}{2} \sum_{ijkl} (w_{ik}^{jl} p_i^{(j)} - w_{ik}^{j} p_i^{(l)}) \log \left( \frac{k!}{j!} \frac{e^{\beta}}{w_{i(k)}} \right)
\]

\[
= -\dot{\mathcal{M}} + \dot{S}_i + \frac{\beta}{2} \sum_{ijkl} (w_{ik}^{jl} p_i^{(j)} - w_{ik}^{j} p_i^{(l)}) (\epsilon_k^{(l)} - \epsilon_i^{(j)}) + \alpha \sum_{ijkl} (w_{ik}^{jl} p_i^{(j)} - w_{ik}^{j} p_i^{(l)}) (l - j).
\]

\[
\dot{S}_i = -\beta Q
\]
Let us note that from the first law of thermodynamics,

\[ \dot{U} = \sum_{ij} \dot{p}^{(j)}_i \epsilon^{(j)}_i + \sum_{ij} \dot{p}^{(j)}_i \epsilon^{(j)}_i = \dot{Q} + \dot{W}, \] (42)

the entropy flow is equal to the heat flow over the temperature. Let us focus on last term, which can be expressed as

\[ \frac{1}{2} \sum_{ijkl} (w^{ij}_{kl} p^{(j)}_i - w^{kl}_{ij} p^{(l)}_i) (l - j) = \frac{1}{2} \sum_{ijkl} w^{ij}_{kl} p^{(j)}_i l - \frac{1}{2} \sum_{ijkl} w^{kl}_{ij} p^{(l)}_i l + \frac{1}{2} \sum_{ijkl} w^{kl}_{ij} p^{(l)}_i j = \sum_{ij} \dot{p}^{(j)}_i j. \] (43)

This is nothing but the time derivative of the normalization condition, i.e., the number of particles in the system. But this remains constant. Thus, we can conclude that the whole term is identically zero. Therefore, the second law of thermodynamics can be expressed in the form

\[ \frac{dS}{dt} = -\dot{M} + \beta \dot{Q} + \dot{S}_i. \] (44)

The time derivative of the free energy can be calculated as

\[ \dot{F} = \dot{U} - T \dot{S} = \dot{W} + \dot{Q} + T \dot{M} - \dot{Q} - T \dot{S}_i. \] (45)

so we have

\[ \Delta S_i = \beta (W - \Delta F) + \Delta M. \] (46)

We can apply this expression directly to the Crooks fluctuation theorem and Jarzynski equality and obtain the formulas in the main text.

**Derivation of the self-consistency equation for magnetization in the fully connected Ising model**

The free energy for the fully connected Ising model is given by

\[ F = -\frac{\alpha}{\beta} - \frac{1 - p_{\parallel}}{\beta}, \] (47)

where \( \alpha \) and \( p_{\parallel} \) are the same as for the molecule gas in the magnetic field (see the following section on the molecule gas in the presence of the magnetic field), just with the effective field \( h_{\text{eff}} = (J_m + h) \). The self-consistency equation is obtained from the relation

\[ m = -\frac{\partial F}{\partial h}|_{h=0}, \] (48)

which leads to the following equation

\[ m = \frac{\sinh(\beta J_m) + \frac{\sinh(\beta J_m) \cosh(\beta J_m)}{\sqrt{\cosh^2(\beta J_m) + en}}}{\left( \cosh(\beta J_m) + \sqrt{\cosh^2(\beta J_m) + en} \right)^2} + \frac{n \left( \frac{\sinh(\beta J_m) + \frac{\sinh(\beta J_m) \cosh(\beta J_m)}{\sqrt{\cosh^2(\beta J_m) + en}}}{\left( \cosh(\beta J_m) + \sqrt{\cosh^2(\beta J_m) + en} \right)^2} \right)}{e^2 \left( \cosh(\beta J_m) + \sqrt{\cosh^2(\beta J_m) + en} \right)^3}. \] (49)

This equation has to be solved numerically, similarly to the case of the fully connected Ising model without molecule states. The solution is depicted in Fig. 2 in the main text.

**Monte Carlo simulation of the fully connected Ising model**

The Monte Carlo simulation for free particles with two states \( \{\uparrow, \downarrow\} \) and the two-particle molecule has one state \( \{\parallel\} \) with the Hamiltonian \( H(n_\uparrow, n_\downarrow, n_{\parallel}) = -h(n_\uparrow - n_\downarrow) \) is realised by defining two boxes similar to what Panagiotopoulos described in [1]. Box 1 contains the states \( \{\uparrow, \downarrow\} \) and Box 2 the states \( \{\parallel\} \). Two alternating kinds of move are tried. The first kind of move randomly chooses a particle in the \( \uparrow \) or \( \downarrow \) state and changes it to the other state. The move
is then accepted with the probability \( \min(1, \exp(-\beta \Delta H)) \). The second kind of move is either dividing a state \( \{\|\} \) particle in two state \( \{\uparrow, \downarrow\} \), or combining two random particles from Box 1 to a state \( \{\|\} \) particle. Which one of the two moves is tried is chosen randomly. The division move takes a Box 2 particle and deletes it. Then two new Box 1 particle are generated. The state of the new particle is either \( \uparrow \) or \( \downarrow \). Which of the two states the particles are created in are chosen randomly with the probability of the current distribution of the two states in Box 1. The move is then accepted with the probability \( \min(1, \frac{2n_\uparrow n_\downarrow}{n_\uparrow+n_\downarrow} \exp(-\beta \Delta H)) \), where \( n_\uparrow \) and \( n_\downarrow \) are the numbers for \( \uparrow \) and \( \downarrow \) particles after the move, respectively. The combination move takes two random particles in Box 1 deletes them and creates a new particle in Box 2. The move is then accepted with min \( \frac{\sqrt{1 + 4en - 1}}{2en} \frac{n_\uparrow n_\downarrow}{2n_\uparrow+n_\downarrow} \exp(-\beta \Delta H) \). Here, \( n_\uparrow \) and \( n_\downarrow \) are once again the number of \( \uparrow \) and \( \downarrow \) particles after the move, respectively. The simulation ends after the ensemble does not change significantly anymore and reaches an equilibrium. The final output is the mean over 1000 simulations at a certain temperature value. Each simulation consists of 2 million steps.

**System with free particles and two-particle molecules**

Let’s consider the simplest case of the system with molecule states. We have one free particle state and one two-particle molecule state. The equilibrium distribution without any constraint can be calculated as \( p^{(1)}(n) = \frac{\sqrt{1 + 4en - 1}}{2en} \) and \( p^{(2)}(n) = \frac{\sqrt{1 + 4en - 1}^2}{8en} \frac{n_\uparrow n_\downarrow}{2n_\uparrow+n_\downarrow} \). From this we see that the distribution, \( p^{(1)}, 2p^{(2)} \), is far from uniform distribution, but as \( n \) goes to infinity the former probability goes to zero, while the latter goes to one. The entropy can be calculated as

\[
S(n) = \frac{1}{2en} (\sqrt{1 + 4en - 1} - \log \left( \frac{2ne}{\sqrt{4en+1+1}} \right))
+ \frac{1}{8en} (\sqrt{1 + 4en - 1})^2 \log \left( \frac{8en}{\sqrt{4en+1+1}^2} \right)
+ \frac{1}{8en} (\sqrt{1 + 4en - 1} - 1)^2 \log \frac{n}{2} \approx \frac{1}{2} \log n .
\]

As a result, the main contribution to entropy is from \( p^{(2)} \). In the large \( n \) limit, the entropy is approximately half of the Boltzmann-Gibbs entropy of \( n \) particles. Interestingly, the same value is obtained for the case of sample space reducing process with \( n \) states [2].

**Molecule gas in the presence of a magnetic field**

Let us now consider a system where free particles can have two states \( \{\uparrow, \downarrow\} \) and two-particle molecules have one state \( \{\|\} \). This model is similar to the magnetic coin model discussed in the main text. The Hamiltonian corresponding to the magnetic field is

\[
H(n_\uparrow, n_\downarrow, n_\|) = -h(n_\uparrow - n_\downarrow) .
\]

The maximum entropy distribution is equal to \( p_\uparrow = \exp(-\alpha - 1 + \beta h), p_\downarrow = \exp(-\alpha - 1 - \beta h) \) and \( p_\| = \frac{2}{\beta} \exp(-1 - 2\alpha) \). The Lagrange parameter \( \alpha \) can be calculated from the normalization condition and we obtain that \( \alpha = \log \left( \sqrt{\cosh^2(\beta h) + en + \cosh(\beta h)} \right) \). Interestingly, we can recover the magnetic gas without any molecules by sending \( n \to 0 \), which gives us that \( \alpha = \log 2 \cosh(\beta h) \).

The specific heat \( c = T \frac{\partial S}{\partial T} \) with \( \beta = 1/T \) (let us set \( k_B = 1 \)) can also be calculated, and we see that there is a phase transition between the magnetic phase and molecule phase which grows with \( n \). In Fig. 3 we can observe the dependence on specific heat on the size of the system. Let us note that there is a dependence of critical temperature on \( N \), where \( \lim_{N \to \infty} T_c(N) = 0 \). However, as shown in the inset, even for large systems, the critical temperature is well separated from zero, since the convergence is very slow. Thus, for all real systems with a finite size, the critical temperature is positive, and we can observe the magnetic phase. Let’s mention that the magnetic gas has been observed for low temperatures experimentally [5, 6].

The equilibrium constant of the chemical reaction can be expressed as

\[
K_c = \frac{n_{X_2}}{n_X} = \frac{p_{X_2}}{(p_X)^2} = \exp \left( 2\beta h + \log \frac{ec}{2} \right) .
\]
FIG. 3: Top panel: Specific heat of the molecule model with magnetic field. We observe a phase transition for the magnetic model. The critical temperature decreases with \( n \) to zero very slowly (inset). Bottom panels: probabilities as the function of temperature for the model without molecules (left), and for the model with molecules (right).

Thus, we obtain the effective chemical potential, \( \Delta \mu^{eff} = 2 \epsilon + \frac{1}{3} \log \frac{\epsilon}{\beta} \), where the first term can be obtained from the ordinary grand-canonical ensemble of two-gas system and the second one is the correction which is obtained from the molecule entropy. This means that if the two gases are perfectly mixed in a small region so that every particle can interact with each other particle, the value of the chemical potential explicitly depends on the number of particles — with an increasing number of particles the chemical potential increases. Actually, the finite-size corrections to the chemical potential have been considered in several aspects, especially in case of interacting particles \[3, 4\]. In our case, the correction is simply because of the emergent structure of the molecule states.

Model with two and three-particle molecules

Similar to the previous examples, we now consider a model where we have two-particle molecules and three-particle molecules. The Hamiltonian is

\[
H(n^{(1)}, n^{(2)}, n^{(3)}) = \epsilon_1 n^{(1)} + \epsilon_2 n^{(2)} + \epsilon_3 n^{(3)}. \tag{53}
\]

The MaxEnt distribution is given by Eq. [12] \( p^{(1)} = \exp(-1 - \alpha - \beta \epsilon_1), p^{(2)} = \frac{1}{2} \exp(-1 - 2\alpha - \beta \epsilon_2) \) and \( p^{(3)} = \frac{e^2}{6} \exp(-1 - 3\alpha - \beta \epsilon_3) \). The Lagrange parameter \( \alpha \) can be deduced from the cubic equation \[13\] and its form is rather complicated. It is actually interesting to see which ‘phase’ dominates at what temperature. This can be actually
The second term can be interpreted as an entropic force (the term is also used in general relativity [7]). For our case, we have $c = n$, $\mathcal{E}^{(1)} = \epsilon_1$, $\mathcal{E}^{(2)} = \epsilon_2 + T \log \left( \frac{2!}{n^2} \right)$ and $\mathcal{E}^{(3)} = \epsilon_3 + T \log \left( \frac{3!}{n^2} \right)$. The system tends to remain in the lowest energy level. By comparing the effective energies we can obtain an illustrative picture of the system. This is illustrated in Fig. 4. For low temperatures most of the particles are in the free state, because $\mathcal{E}^{(1)}$ is the lowest energy level. Then, for higher temperatures, the two-particle molecules take over and finally, for high temperatures the three-particle molecule start to dominate. In the example we used the following values: $\epsilon_1 = 1$, $\epsilon_2 = 3$, $\epsilon_3 = 8$ and $n = 100$. Indeed, for other energy configurations, it may happen that some configurations will be suppressed because they will have unfavourable effective energy.