Reconnecting statistical physics and combinatorics beyond ensemble equivalence

Tiziano Squartini

IMT School for Advanced Studies, Lucca, Italy

Diego Garlaschelli

Lorentz Institute for Theoretical Physics, University of Leiden, The Netherlands

(Dated: July 8, 2018)

In statistical physics, the challenging combinatorial enumeration of the configurations of a system subject to hard constraints (microcanonical ensemble) is mapped to a mathematically easier calculation where the constraints are softened (canonical ensemble). However, the mapping is exact only when the size of the system is infinite and if the property of ensemble equivalence (EE), i.e. the asymptotic identity of canonical and microcanonical large deviations, holds. For finite systems, or when EE breaks down, statistical physics is currently believed to provide no answer to the combinatorial problem. In contrast with this expectation, here we establish exact relationships connecting conjugate ensembles in full generality, even for finite system size and when EE does not hold. We also show that in the thermodynamic limit the ensembles are directly related through the matrix of canonical (co)variances of the constraints, plus a correction term that survives only if this matrix has an infinite number of finite eigenvalues. These new relationships restore the possibility of enumerating microcanonical configurations via canonical probabilities, thus reconnecting statistical physics and combinatorics in realms where they were believed to be no longer in mutual correspondence.

For virtually any system consisting of a large number n of interacting elements, an exact microscopic description is unfeasible and has to be replaced by a statistical one involving a probability distribution $P(\mathbf{G})$ over the microstates $\mathbf{G} \in \mathcal{G}$ of the system consistent with a certain number K of macroscopic properties $\vec{C}^* = (C_1^*, \ldots, C_K^*)$ that are empirically accessible. Here $\vec{C}^* \equiv \vec{C}(\mathbf{G}^*)$ is the particular, known value of the K-dimensional vector $\vec{C}(\mathbf{G})$ of macroscopic constraints, $\mathbf{G}^* \in \mathcal{G}$ is the (unknown) microstate, and $\mathcal{G}$ is the set of allowed microstates. For concreteness, throughout the paper we assume that $\mathbf{G}$ only takes discrete values, so that $P(\mathbf{G})$ is a probability mass function. The principle of maximum entropy states that the optimal $P(\mathbf{G})$ is the one that maximizes the entropy functional

$$S[P] \equiv - \sum_{\mathbf{G} \in \mathcal{G}} P(\mathbf{G}) \ln P(\mathbf{G}), \qquad (1)$$

as such maximum-entropy probability is *"maximally non-committal with respect to missing information"* [1]. The resulting theory is statistical physics, which describes large systems subject to a few macroscopic constraints such as total energy and total number of particles [2, 3].

The usefulness of statistical physics crucially depends on whether one is able to calculate $P(\mathbf{G})$ explicitly. Imposing *hard* constraints, i.e. restricting oneself to the configurations $\mathbf{G}$ such that $\vec{C}(\mathbf{G}) = \vec{C}^*$, leads to the so-called *microcanonical* maximum-entropy probability

$$P_{\text{mic}}(\mathbf{G}|\vec{C}^*) = \Omega_{\vec{C}^*}^{-1} \delta_{\vec{C}^*, \vec{C}(\mathbf{G})}, \qquad (2)$$

where $\Omega_{\vec{C}^*}$ is the number of microstates compatible with the constraint $\vec{C}^*$.

The calculation of $\Omega_{\vec{C}^*}$ requires complicated and often unknown combinatorial enumeration formulae, a math-

ematical challenge that makes the microcanonical ensemble generally intractable. A cornerstone of statistical physics, dating back to Gibbs [3] and known as 'ensemble equivalence' (EE), is the assumption that, in the thermodynamic limit $n \to \infty$, the hard values $\vec{C}^*$ of the constraints can be safely replaced by *soft* (i.e. suitably fluctuating) values, since for large systems the fluctuations around the hard values are expected to vanish. The result is the mathematically tractable *canonical* ensemble [1] where the constraints are enforced only as expected values, i.e. $\langle \vec{C} \rangle = \vec{C}^*$. The maximization of the entropy yields in this case [1]

$$P_{\text{can}}(\mathbf{G}|\vec{\theta}^*) = \frac{e^{-\vec{\theta}^* \cdot \vec{C}(\mathbf{G})}}{\sum_{\mathbf{G}' \in \mathcal{G}} e^{-\vec{\theta}^* \cdot \vec{C}(\mathbf{G}')}} = \frac{e^{-\mathcal{H}(\mathbf{G}, \vec{\theta}^*)}}{\mathcal{Z}(\vec{\theta}^*)}, \qquad (3)$$

where $\vec{\theta}$ is a K-dimensional vector of Lagrange multipliers needed to control the expected value $\langle \vec{C} | \vec{\theta} \rangle \equiv \sum_{\mathbf{G} \in \mathcal{G}} P_{\text{can}}(\mathbf{G}|\vec{\theta}) \vec{C}(\mathbf{G})$, $\mathcal{H}(\mathbf{G}, \vec{\theta}) \equiv \vec{\theta} \cdot \vec{C}(\mathbf{G}) = \sum_{k=1}^{K} \theta_k C_k(\mathbf{G})$ is (somewhat improperly) called the 'Hamiltonian' function (the dot indicating the scalar product), $\mathcal{Z}(\vec{\theta}) = \sum_{\mathbf{G} \in \mathcal{G}} e^{-\vec{\theta} \cdot \vec{C}(\mathbf{G})}$ is the partition function, and $\vec{\theta}^*$ is the specific value such that

$$\langle \vec{C} | \vec{\theta}^* \rangle = \vec{C}^*, \qquad (4)$$

---

where we have defined the canonical ensemble average of a generic function $X(\mathbf{G})$ as

$$\langle X|\vec{\theta}\rangle \equiv \sum_{\mathbf{G}\in\mathcal{G}} P_{\text{can}}(\mathbf{G}|\vec{\theta})X(\mathbf{G}). \qquad (5)$$

Equivalently, the value $\vec{\theta}^*$ is the one that maximizes the log-likelihood function $\mathcal{L}^*(\vec{\theta}) \equiv \ln P_{\text{can}}(\mathbf{G}^*|\vec{\theta})$ [4, 5] (see SI).

Note that eq. (4) fixes the relationship between the dual quantities $\vec{C}^*$ and $\vec{\theta}^*$. We can therefore simplify the notation and set $P_{\text{mic}}^*(\mathbf{G}) \equiv P_{\text{mic}}(\mathbf{G}|\vec{C}^*)$ and $P_{\text{can}}^*(\mathbf{G}) \equiv P_{\text{can}}(\mathbf{G}|\vec{\theta}^*)$, the asterisk indicating that both probabilities are ultimately specified by the unique value $\vec{C}^*$. Similarly, after inserting eqs. (2) and (3) into eq. (1), we can denote the resulting entropies of the two ensembles as

$$S_{\text{mic}}^* \equiv S[P_{\text{mic}}^*] = \ln\Omega_{\vec{C}^*} \qquad (6)$$

$$S_{\text{can}}^* \equiv S[P_{\text{can}}^*] = -\sum_{\mathbf{G}\in\mathcal{G}} P_{\text{can}}(\mathbf{G}|\vec{\theta}^*)\ln P_{\text{can}}(\mathbf{G}|\vec{\theta}^*). \ (7)$$

An important equality, proven in SI and crucial for our later results, is the one between the canonical entropy and minus the *maximized* log-likelihood $\mathcal{L}^*(\vec{\theta}^*)$:

$$S_{\text{can}}^* = -\ln P_{\text{can}}(\mathbf{G}^*|\vec{\theta}^*) = -\mathcal{L}^*(\vec{\theta}^*). \qquad (8)$$

Throughout the paper, we assume that in the thermodynamic limit $\vec{\theta}^*$ is not a critical point of a phase transition, i.e. it is in the interior of a thermodynamic phase.

Obviously, for finite n the microcanonical and canonical ensembles are necessarily different. A mathematical formalization of this difference is discussed later. An informal example is the different physical interpretation of the two ensembles, the microcanonical one being viewed as a model for systems in energetic isolation (if the constraint is the total energy) and the canonical one as a model for systems in thermal equilibrium with a heat bath (if the Lagrange multiplier is the inverse temperature). So it only makes sense to talk about EE in the thermodynamic limit, if relative energy fluctuations vanish. This already implies that the combinatorial problem cannot be solved via statistical physics when n is finite.

In the thermodynamic limit, provided EE holds, the generally unfeasible calculations in the microcanonical ensemble can be replaced by easier calculations in the canonical ensemble. For instance, most textbooks treat the microcanonical ensemble by relying on the approximation $\Omega_{\vec{C}^*} \approx e^{S_{\text{can}}^*} = 1/P_{\text{can}}^*(\mathbf{G}^*)$ where the symbol "$\approx$" means "equal up to a factor assumed to be subexponential in n" and the last equality follows from eq. (8). To make a concrete example, for systems with given total energy E^* one often assumes $\Omega_{E^*} \approx e^{+\beta E^*}\mathcal{Z}(\beta)$ where $\beta = (k_B T)^{-1}$ is the inverse temperature, k_B is Boltzmann's constant and T is the temperature. Stated more rigorously, this assumption is

$$\Omega_{\vec{C}^*} = e^{S_{\text{can}}^* - o(n)} = \frac{e^{-o(n)}}{P_{\text{can}}^*(\mathbf{G}^*)} \qquad (9)$$

(where $o(n)$ denotes a quantity that, if divided by n, vanishes as n diverges) or equivalently

$$\lim_{n\to\infty}\frac{\ln\Omega_{\vec{C}^*}}{n} = -\lim_{n\to\infty}\frac{\ln P_{\text{can}}^*(\mathbf{G}^*)}{n} = \lim_{n\to\infty}\frac{S_{\text{can}}^*}{n}. \ (10)$$

Therefore, in presence of EE, the enumeration problem required to calculate the 'hard' microcanonical quantity $\Omega_{\vec{C}^*}$ can be mapped exactly, in the thermodynamic limit, to a 'softened' canonical problem requiring only the calculation of $P_{\text{can}}^*(\mathbf{G}^*)$ or S_{can}^*. This establishes a tight connection between statistical physics and combinatorics. An equivalent condition is

$$P_{\text{mic}}^*(\mathbf{G}^*) = \Omega_{\vec{C}^*}^{-1} = e^{o(n)-S_{\text{can}}^*} = P_{\text{can}}^*(\mathbf{G}^*)e^{o(n)}, \quad (11)$$

or ultimately

$$\lim_{n\to\infty}\frac{1}{n}\ln\frac{P_{\text{mic}}^*(\mathbf{G}^*)}{P_{\text{can}}^*(\mathbf{G}^*)} = 0, \qquad (12)$$

which essentially states that EE corresponds to the identity of the large deviation properties of the two conjugate ensembles at rate n, i.e. the two probability measures are *exponentially equivalent* at rate n.

When the size n of the system is finite, or when the assumption of EE as stated by eq. (12) is violated, the two ensembles are different and the enumeration problem cannot be directly solved via a canonical calculation. Indeed, while most statistical physics textbooks still convey the idea that EE is expected to hold in very general circumstances, several examples of the breakdown of EE in a diverse range of systems have actually been provided over the last decades [6–13]. A recently investigated class of such systems is random networks with topological constraints [14–21], to which we will briefly refer later in this paper. For systems with nonequivalent ensembles, it has been shown that microcanonical and canonical calculations of various properties no longer agree. Therefore, for these systems, the connection between statistical physics and combinatorics appears essentially lost, just like for systems with finite size.

At this point, an open problem of practical relevance emerges. Even accepting that the useful mapping to the canonical ensemble is only justified in the regime of EE, how to rigorously determine whether EE holds, without having to calculate the microcanonical ensemble? In fact, if EE can only be verified *after* having calculated *both* canonical and microcanonical properties and compared them, then the practical usefulness of the canonical ensemble evaporates, because one will have to trust the latter only when the microcanonical one is also tractable.

A second open problem, related to the previous one but of more fundamental theoretical importance and generality, should also be highlighted. Is it possible, in absence of EE and/or when the size of the system is finite, to restore some (possibly modified) rigorous mathematical correspondence between the canonical and microcanonical ensembles, so that the intractability of the latter can still be circumvented by performing feasible calculations

in the former? In other words, is there a general and stronger mathematical relationship connecting conjugate statistical ensembles, irrespective of whether EE holds and possibly for finite system size?

We now provide definite solutions to these two problems. One of our main results, proven in SI, is the following exact formula, which is valid in general and reduces the calculation of $\Omega_{\vec{C}^*}$ to purely canonical quantities:

$$\Omega_{\vec{C}^*} = P_{\text{mic}}^{-1}(\mathbf{G}^*|\vec{C}^*) = \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K}\, P_{\text{can}}^{-1}(\mathbf{G}^*|\vec{\theta}^* + i\vec{\psi}) \quad (13)$$

where $\vec{\pi} \equiv (\pi, \ldots, \pi)$ is the K-dimensional vector with all components equal to π and the argument $\vec{\theta}$ of $P_{\text{can}}(\mathbf{G}^*|\vec{\theta})$ has been extended to complex values. The above equation establishes a seemingly esoteric, but exact relationship connecting the canonical and microcanonical probabilities. When the integral in eq. (13) can be carried out explicitly, the microcanonical ensemble can be calculated *exactly* from an extension of the canonical ensemble to the complex domain. Remarkably, the formula is valid *even for finite n and even when EE breaks down*, thus extending the usefulness of the canonical ensemble significantly and unveiling a tighter and more general connection between 'hard' combinatorial enumeration and 'soft' statistical physics.

Due to the importance of the above result, we may wonder whether eq. (13), which in principle needs to be evaluated case by case and is therefore not very transparent, takes some more explicit and general form, independently of the model specification. As we show in SI, this is indeed possible. Before showing the result, we need some precautions. Note that, for certain choices of $\vec{\theta}^*$ or $\vec{C}^*$, some of the K components of $\vec{C}^*$ may be dependent on the other ones (see SI for details). In such a case, there is only an effective number $\tilde{K} \leq K$ of independent constraints. Upon ordering the entries of $\vec{C}$ in such a way that the first $\tilde{K}$ ones are independent, we define the reduced $\tilde{K} \times \tilde{K}$ matrix $\tilde{\boldsymbol{\Sigma}}^*$ of the canonical (co)variances, evaluated at $\vec{\theta}^*$, of the $\tilde{K}$ independent constraints. The entries of $\tilde{\boldsymbol{\Sigma}}^*$ are

$$\tilde{\Sigma}_{k,l}^* = \text{Cov}[C_k, C_l]_{\vec{\theta}^*} \qquad k, l = 1, \ldots, \tilde{K} \quad (14)$$

and its eigenvalues, which are all strictly positive, are denoted as $\{\lambda_k^*\}_{k=1}^{\tilde{K}}$. The strict positivity of these eigenvalues is a consequence of the semi-positive definiteness of covariance matrices, together with the fact that all the $\tilde{K}$ rows of $\tilde{\boldsymbol{\Sigma}}^*$ are independent, as follows from the independence of the corresponding $\tilde{K}$ constraints. Armed with the above notation, we can carry out a saddle-point [22] calculation (see SI) to obtain the following asymptotic expansion for $\Omega_{\vec{C}^*}$:

$$\Omega_{\vec{C}^*} = \frac{e^{S_{\text{can}}^*} T^*}{\sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}} \quad (15)$$

where

$$T^* \equiv \prod_{k=1}^{\tilde{K}} \left[1 + O\left(1/\lambda_k^*\right)\right]. \quad (16)$$

Equation (15) makes the leading term of eq.(13) manifest and renders the connection between the hard enumeration problem and its softened dual more explicit. In particular, upon comparison with eq.(9), it shows that an important role is played not only by the canonical entropy S_{can}^*, but also by the canonical covariances $\tilde{\boldsymbol{\Sigma}}^*$. We will discuss the asymptotic behaviour of T^* later. For the moment, we note that if no n-dependence is assumed in the problem, i.e. question is counting how many configurations are compatible with a fixed constraint $\vec{C}^*$ for a given finite value of n, then eqs. (15) and (16) already provide the corresponding answer. However, the usual problem in combinatorics is that of assuming a certain n-dependence of the configuration space and of the enforced constraint. In this case, the counting question needs an asymptotic answer for large n. This is precisely what we are going to investigate at the end, when moving to the thermodynamic limit.

We now show how the above results relate to the property of ensemble (non)equivalence and how they can be exploited to generalize eqs.(9), (10), (11) and (12) to the case of finite system sizes and/or nonequivalent ensembles. We start from the definition of the relative entropy (or Kullback-Leibler divergence) between the canonical and microcanonical probabilities [20, 23]:

$$\Delta^* \equiv S[P_{\text{mic}}^* || P_{\text{can}}^*] = \sum_{\mathbf{G}} P_{\text{mic}}^*(\mathbf{G}) \ln \frac{P_{\text{mic}}^*(\mathbf{G})}{P_{\text{can}}^*(\mathbf{G})}. \quad (17)$$

The relative entropy rigorously quantifies the additional uncertainty contained in $P_{\text{can}}^*(\mathbf{G})$ as compared to $P_{\text{mic}}^*(\mathbf{G})$. This additional uncertainty is due to the fluctuating nature of the constraints in the canonical ensemble. An important result, proven in ref. [20], is that Δ^* is only determined by the *local* values $P_{\text{can}}^*(\mathbf{G}^*)$ and $P_{\text{mic}}^*(\mathbf{G}^*)$ achieved by any of the configurations $\mathbf{G}^*$ that realize the sharp value $\vec{C}^*$ of the constraints:

$$\Delta^* = \ln \frac{P_{\text{mic}}^*(\mathbf{G}^*)}{P_{\text{can}}^*(\mathbf{G}^*)} = \ln \frac{1}{\Omega_{\vec{C}^*} P_{\text{can}}^*(\mathbf{G}^*)}. \quad (18)$$

For finite n, P_{mic}^* and P_{can}^* are different and $\Delta^* > 0$, in line with our previous discussion of the distinction between thermal equilibrium and energetic isolation. However, we can now quantify Δ^* exactly by inserting eq.(13) into eq.(18):

$$\begin{aligned}
\Delta^* &= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} \frac{P_{\text{can}}(\mathbf{G}^*|\vec{\theta}^*)}{P_{\text{can}}(\mathbf{G}^*|\vec{\theta}^* + i\vec{\psi})} \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} \langle e^{i\vec{\psi}\cdot(\vec{C}^* - \vec{C})} | \vec{\theta}^* \rangle, \quad (19)
\end{aligned}$$

where we have used eq. (5) with $X(\mathbf{G}) = e^{i\vec{\psi}\cdot[\vec{C}^* - \vec{C}(\mathbf{G})]}$ (see SI for details). Equation (19) is valid in general

(again, even for finite n and when EE breaks down) and provides a way to calculate the relative entropy exactly *as a function of only the canonical probability* (extended to the complex domain).

We now consider the thermodynamic limit. Inserting the asymptotic result (15) into eq. (18) leads to

$$\Delta^* = \ln \frac{\sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}}{T^*}, \tag{20}$$

where we have used eq. (8). It is now useful to re-express the above result in terms of a suitable relative entropy *density*. Indeed, the definition of EE in the measure sense [23] is the vanishing of the relative entropy n-density as $n \to \infty$, i.e.

$$\lim_{n\to\infty} \frac{\Delta^*}{n} = 0. \tag{21}$$

Using eq. (18), we indeed see that the above condition is the same as eq. (12). When EE breaks down, i.e. when the above limit is strictly positive, then eqs. (9), (10), (11) and (12) no longer hold, and the nexus between combinatorics and statistical physics is apparently severed. However, this turns out not to be the case, because a nontrivially modified version of all those equations is still in place, as we now show.

Let us slightly change our perspective and look for an increasing positive sequence α_n such that, irrespective of whether EE holds, the following limit is finite:

$$\delta^*_{\alpha_\infty} \equiv \lim_{n\to\infty} \frac{\Delta^*}{\alpha_n}, \qquad \delta^*_{\alpha_\infty} \in (0, +\infty). \tag{22}$$

We call $\delta^*_{\alpha_\infty}$ the *limiting α_n-density* of Δ^*. We can then restate the definition of EE in the measure sense by saying that the ensembles are equivalent iff $\alpha_n = o(n)$. We will return on this point later. An important result, following from eq.(20) and proven in SI, is

$$\delta^*_{\alpha_\infty} = \lim_{n\to\infty} \frac{\ln \sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}}{\alpha_n} + \tau^*_{\alpha_\infty}. \tag{23}$$

where

$$\tau^*_{\alpha_\infty} = -\lim_{n\to\infty} \frac{\ln T^*}{\alpha_n} \tag{24}$$

is zero, unless there is an infinite (growing at least like α_n) number of eigenvalues of $\tilde{\boldsymbol{\Sigma}}^*$ that have a finite limit as $n \to \infty$ (see SI). This key result finally allows us to discuss the asymptotic behaviour of T^* in terms of its effects on $\tau^*_{\alpha_\infty}$.

Whenever the number of constraints is finite, then $\tau^*_{\alpha_\infty} = 0$ irrespective of the value of the eigenvalues of $\tilde{\boldsymbol{\Sigma}}^*$. This applies to most traditional situations in statistical physics, where there are only a handful of constraints including properties like total energy, total number of particles (possibly of different chemical species),

etc. In the context of network ensembles, this also applies to e.g. random graphs with a finite number of global constraints, like total number of links and/or triangles and/or wedges [14–17].

Also, whenever the canonical fluctuations of all constraints grow as the size of the system grows, then $\tau^*_{\alpha_\infty} = 0$ irrespective of the number of constraints. Note that this is also a very natural situation, since the constraints are typically chosen to be extensive quantities whose mean and variance grows with n. Besides the traditional examples mentioned above, this includes systems with *local* (i.e. particle-specific) constraints, such as dense random graphs with given degrees, for which $\alpha_n = n \ln n$, $K = \tilde{K} = n$, and $\tilde{\boldsymbol{\Sigma}}^*$ has diagonal entries equal to the (diverging) variances of the (diverging) degrees. A rigorous mathematical proof that this class of graphs obeys eq. (23) with $\tau^*_{\alpha_\infty} = 0$ is left for a separate paper [24].

If the number of constraints is infinite and, roughly speaking, the canonical fluctuations of all but a finite number of constraints grow as the size of the system grows, then all but a finite number of eigenvalues will diverge, and again $\tau^*_{\alpha_\infty} = 0$. This is a useful criterion in hybrid situations where $\tilde{\boldsymbol{\Sigma}}^*$ has both finite and infinite eigenvalues. So, in order to have a finite correction $\tau^*_{\alpha_\infty} > 0$ surviving in eq. (23), we need an infinite number of finite constraints. We expect this to occur in sparse random graphs with given degrees, for which $\alpha_n = n$, $K = \tilde{K} = n$, and $\tilde{\boldsymbol{\Sigma}}^*$ is nearly diagonal with diagonal entries equal to the (finite) variances of the (finite) degrees [24]. When $\tau^*_{\alpha_\infty} > 0$, the value of $\delta^*_{\alpha_\infty}$ has to be calculated independently from eq. (23) for the specific model, using either the explicit form of $\Omega_{\vec{C}^*}$ (if known) or carrying out the integral in eq. (19) explicitly and then taking the limit (22).

Note that, in our approach, whether $\tau^*_{\alpha_\infty}$ is zero has nothing to do with EE or its breakdown. By definition of α_n, $\delta^*_{\alpha_\infty}$ is always strictly positive, irrespective of the value of $\tau^*_{\alpha_\infty}$. Whether EE holds depends on a different property, namely whether $\alpha_n = o(n)$. Our results above enable the calculation of $\delta^*_{\alpha_\infty}$, and hence α_n, in terms of purely canonical quantities. This allows checking whether the ensembles are equivalent, without having to calculate microcanonical quantities.

We can now exploit our results in order to provide exact and more general replacements for eqs. (9), (10), (11) and (12), valid even when EE is violated. Note that eq. (22) implies

$$\Delta^* = \alpha_n \delta^*_{\alpha_\infty} + o(\alpha_n), \tag{25}$$

where $\delta^*_{\alpha_\infty}$ is given by eq. (23). This implies that eq. (18) can be rearranged as follows:

$$\begin{aligned} \Omega_{\vec{C}^*} &= e^{S^*_{\mathrm{can}} - \Delta^*} \\ &= e^{S^*_{\mathrm{can}} - \alpha_n \delta^*_{\alpha_\infty} - o(\alpha_n)} \\ &= \frac{e^{-\alpha_n \delta^*_{\alpha_\infty} - o(\alpha_n)}}{P^*_{\mathrm{can}}(\mathbf{G}^*)}. \end{aligned} \tag{26}$$

The above expression is the correct extension of eq. (9). We see the appearance of the extra term $\delta^*_{\alpha_\infty}$ of finite order. As a consequence, eq. (10) should be generalized to

$$
\begin{aligned}
\lim_{n\to\infty} \frac{\ln \Omega_{\vec{C}^*}}{\alpha_n} &= \lim_{n\to\infty} \frac{S^*_{\mathrm{can}}}{\alpha_n} - \delta^*_{\alpha_\infty}, \\
&= \lim_{n\to\infty} \frac{S^*_{\mathrm{can}} - \ln\sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}}{\alpha_n} - \tau^*_{\alpha_\infty}
\end{aligned}
\tag{27}
$$

which represents a novel, corrected identify mapping the enumeration problem to the calculation of the canonical entropy *and* the relative entropy or the covariance matrix of the constraints. Correspondingly, eqs. (11) and (12) should be replaced by

$$
P^*_{\mathrm{mic}}(\mathbf{G}^*) = P^*_{\mathrm{can}}(\mathbf{G}^*)\, e^{\Delta^*} = P^*_{\mathrm{can}}(\mathbf{G}^*)\, e^{\alpha_n \delta^*_{\alpha_\infty} + o(\alpha_n)}
\tag{28}
$$

and

$$
\lim_{n\to\infty} \frac{1}{\alpha_n} \ln \frac{P^*_{\mathrm{mic}}(\mathbf{G}^*)}{P^*_{\mathrm{can}}(\mathbf{G}^*)} = \delta^*_{\alpha_\infty}
\tag{29}
$$

respectively.

Equations (26), (27), (28) and (29) restore an exact, nontrivially modified connection between statistical ensembles, valid irrespective of whether they are equivalent. In presence of EE, i.e. if $\alpha_n = o(n)$, then $\alpha_n \delta^*_{\alpha_\infty} + o(\alpha_n) = o(n)$ and these expressions reduce to eqs. (9), (10), (11) and (12) respectively. By contrast, if EE does not hold, then the factor neglected in eq. (9) is actually (at least) exponential in n, i.e. the volume of microcanonical configurations is exponentially smaller than estimated in that equation. Remarkably, our results restore, in a properly modified way, the possibility of conveniently expressing microcanonical quantities in terms of purely canonical ones, even beyond the regime of EE. As a byproduct, this shows that it is indeed possible to check for EE (i.e. whether $\alpha_n = o(n)$) without having to calculate the microcanonical ensemble.

We finally note that all choices of α_n with the same leading order are equivalent. This can be exploited to introduce a sort of 'natural' $\tilde{\alpha}_n$ defined as

$$
\tilde{\alpha}_n = \ln\sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}
\tag{30}
$$

(note that the term on the right is n-dependent, despite the absence of an explicit symbol "n" in our notation). The above choice is admissible, because it always ensures $\delta^*_{\tilde{\alpha}_\infty} \in (0, +\infty)$. Moreover, it allows to rewrite a number of expressions in more compact form. In particular, checking for EE simply reduces to checking whether $\tilde{\alpha}_n = o(n)$, i.e.

$$
\ln\sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)} = o(n) \iff \mathrm{EE}.
\tag{31}
$$

This explicit and simple criterion is an important byproduct of the results presented in this paper. Moreover, if

there is at most a finite number of eigenvalues of $\tilde{\boldsymbol{\Sigma}}^*$ that have a finite limit as $n \to \infty$, then $\tau^*_{\tilde{\alpha}_\infty} = 0$ and we can rewrite eqs. (23) and (25) as

$$
\begin{aligned}
\delta^*_{\tilde{\alpha}_\infty} &= 1, \tag{32} \\
\Delta^* &= \tilde{\alpha}_n + o(\tilde{\alpha}_n). \tag{33}
\end{aligned}
$$

Similarly, eqs. (26) and (27) become

$$
\Omega_{\vec{C}^*} = e^{S^*_{\mathrm{can}} - \tilde{\alpha}_n - o(\tilde{\alpha}_n)},
\tag{34}
$$

$$
\lim_{n\to\infty} \frac{\ln \Omega_{\vec{C}^*}}{\tilde{\alpha}_n} = \lim_{n\to\infty} \frac{S^*_{\mathrm{can}}}{\tilde{\alpha}_n} - 1,
\tag{35}
$$

while eqs. (28) and (29) become

$$
P^*_{\mathrm{mic}}(\mathbf{G}^*) = P^*_{\mathrm{can}}(\mathbf{G}^*)\, e^{\tilde{\alpha}_n + o(\tilde{\alpha}_n)},
\tag{36}
$$

$$
\lim_{n\to\infty} \frac{1}{\tilde{\alpha}_n} \ln \frac{P^*_{\mathrm{mic}}(\mathbf{G}^*)}{P^*_{\mathrm{can}}(\mathbf{G}^*)} = 1
\tag{37}
$$

respectively. Equations (32)-(37) are an elegant and compact summary of our results in the case $\tau^*_{\tilde{\alpha}_\infty} = 0$.

To summarize, in this paper we have established new general connections between conjugate ensembles. So far, the microcanonical and canonical ensembles have been regarded as dual representations of the same system that are valid under different physical conditions (such as energetic isolation and thermal equilibrium) and only coincide under certain assumption (thermodynamic limit and ensemble equivalence). By contrast, we have found that the relationship between the two ensembles is much stronger than mere conjugacy or duality, as it can be formulated in terms of mathematical identities that allow the two ensembles can be calculated from each other in full generality, even when neither the thermodynamic limit nor the assumption of ensemble equivalence are in order. Our results significantly expand the usefulness of the canonical ensemble as a tool to carry out *exact* physical and combinatorial calculations that are unfeasible in the microcanonical one even when EE breaks down and/or for finite system size. In particular, we found new and very general enumeration formulae (both exact and asymptotic) that can be applied to a variety of (possibly unsolved) combinatorial problems. What is exciting from a physical point of view is that the enumeration of all the configurations with a given hard value of the constraints can be re-expressed in terms of the conjugate canonical distribution with softened constraints, extended to the complex domain. In the thermodynamic limit, the two ensembles are connected via the relative entropy density, which we found to be proportional to the logarithm of the determinant of the matrix of canonical covariances of the constraints, plus a possible correction term that is nonzero only if that matrix has an infinite number of finite eigenvalues. These results offer new insight into the foundations of statistical physics and its connections to combinatorial enumeration. Moreover, they significantly expand the toolkit for studying systems with nonequivalent ensembles.

Acknowledgments

The authors acknowledge fruitful discussions with Frank den Hollander and Andrea Roccaverde.

[1] E.T. Jaynes, *Phys. Rev.* **106**, 4 (1957).
[2] L. Boltzmann, *Wiener Berichte* **2**, 373-435 (1877).
[3] J.W. Gibbs, *Elementary Principles of Statistical Mechanics*, Yale University Press, New Haven, Connecticut (1902).
[4] D. Garlaschelli and M.I. Loffredo, *Phys. Rev. E* **78**, 015101(R) (2008).
[5] T. Squartini and D. Garlaschelli, *New J. Phys.* **13**, 083001 (2011).
[6] R.S. Ellis, K. Haven, and B. Turkington, *J. Stat. Phys.* **101**, 999 (2000).
[7] M. Blume, V.J. Emery, and R.B. Griffiths, *Phys. Rev. A* **4**, 1071 (1971).
[8] J. Barré, D. Mukamel, and S. Ruffo, *Phys. Rev. Lett.* **87**, 030601 (2001).
[9] R.S. Ellis, H. Touchette, and B. Turkington, *Physica A* **335**, 518 (2004).
[10] D. Lynden-Bell, *Physica A* **263**, 293 (1999).
[11] P.-H. Chavanis, *Astron. & Astrophys.* **401**, 15 (2003).
[12] M. D'Agostino, *Phys. Lett. B* **473**, 219 (2000).
[13] J. Barré and B. Goncalves, *Physica A* **386**, 212-218 (2007).
[14] C. Radin and L. Sadun, *J. Phys. A: Math. Theor.* **46**, 305002 (2013).
[15] C. Radin and M. Yin, *The Annals of Applied Probability* **23**, 2458-2471 (2013).
[16] S. Chatterjee and P. Diaconis, *The Annals of Statistics* **41**, 2428-2461 (2013).
[17] F. den Hollander, M. Mandjes, A. Roccaverde, N.J. Starreveld, `https://arxiv.org/abs/1703.08058`.
[18] K. Anand and G. Bianconi, *Phys. Rev. E* **80**, 045102 (2009).
[19] T. Squartini, R. Mastrandrea, and D. Garlaschelli, *New J. Phys.* **17**, 023052 (2015).
[20] T. Squartini, J. de Mol, F. den Hollander, D. Garlaschelli, *Physical review letters*, **115**(26), 268701 (2015).
[21] D. Garlaschelli, F. den Hollander, A. Roccaverde, *Journal of Physics A: Mathematical and Theoretical* **50**(1), 015001 (2016).
[22] M.V. Fedoryuk, "Asymptotic Methods in Analysis" in: Gamkrelidze R.V. (eds) *Analysis I. Encyclopaedia of Mathematical Sciences*, vol 13. Springer, Berlin, Heidelberg (1989).
[23] H. Touchette, *J. Stat. Phys.* **159**(5), 987-1016 (2015).
[24] D. Garlaschelli, F. den Hollander, A. Roccaverde. *in preparation* (2017).

SUPPLEMENTARY INFORMATION

accompanying the paper
*"Reconnecting statistical physics and combinatorics
beyond ensemble equivalence"*
by T. Squartini and D. Garlaschelli

The likelihood function

We first recall the result, shown e.g. in [4, 5], that the value $\vec{\theta}^*$ that realizes $\langle \vec{C}|\vec{\theta}^*\rangle = \vec{C}^*$ in eq. (4) can be equivalently defined as the value that maximizes the *log-likelihood function*

$$
\begin{aligned}
\mathcal{L}^*(\vec{\theta}) &\equiv \ln P_{\text{can}}(\mathbf{G}^*|\vec{\theta}) \\
&= -\vec{\theta}\cdot\vec{C}^* - \ln\mathcal{Z}(\vec{\theta})
\end{aligned}
\tag{S1}
$$

(where the asterisk now indicates dependence of $\mathcal{L}$ on $\mathbf{G}^*$ and hence on $\vec{C}^*$, but *not* on $\vec{\theta}^*$). In other words,

$$
\vec{\theta}^* = \text{argmax}_{\vec{\theta}}\left\{\mathcal{L}^*(\vec{\theta})\right\}.
\tag{S2}
$$

To prove the above result, we calculate the first partial derivatives of $\mathcal{L}^*$

$$
\begin{aligned}
\frac{\partial}{\partial\theta_k}\mathcal{L}^*(\vec{\theta}) &= -C_k^* - \frac{1}{\mathcal{Z}(\vec{\theta})}\frac{\partial}{\partial\theta_k}\mathcal{Z}(\vec{\theta}) \\
&= -C_k^* + \sum_{\mathbf{G}\in\mathcal{G}} C_k(\mathbf{G})\frac{e^{-\vec{\theta}\cdot\vec{C}(\mathbf{G})}}{\mathcal{Z}(\vec{\theta})} \\
&= \langle C_k|\vec{\theta}\rangle - C_k^*
\end{aligned}
\tag{S3}
$$

and note that eq. (4) implies that $\vec{\theta}^*$ is such that

$$
\frac{\partial}{\partial\theta_k}\mathcal{L}^*(\vec{\theta})\bigg|_{\vec{\theta}=\vec{\theta}^*} = \langle C_k|\vec{\theta}^*\rangle - C_k^* = 0 \quad \forall k.
\tag{S4}
$$

In other words, $\vec{\theta}^*$ is a stationary point for $\mathcal{L}^*$. It is also the *only* such point.

To prove that $\vec{\theta}^*$ is also a global maximum, we consider the second derivatives

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta_k\partial\theta_l}\mathcal{L}^*(\vec{\theta}) &= \sum_{\mathbf{G}\in\mathcal{G}} C_k(\mathbf{G})\frac{e^{-\vec{\theta}\cdot\vec{C}(\mathbf{G})}}{\mathcal{Z}(\vec{\theta})}\sum_{\mathbf{G}\in\mathcal{G}} C_l(\mathbf{G})\frac{e^{-\vec{\theta}\cdot\vec{C}(\mathbf{G})}}{\mathcal{Z}(\vec{\theta})} \\
&\quad - \sum_{\mathbf{G}\in\mathcal{G}} C_k(\mathbf{G})C_l(\mathbf{G})\frac{e^{-\vec{\theta}\cdot\vec{C}(\mathbf{G})}}{\mathcal{Z}(\vec{\theta})} \\
&= \langle C_k|\vec{\theta}\rangle\langle C_l|\vec{\theta}\rangle - \langle C_kC_l|\vec{\theta}\rangle \\
&= -\text{Cov}[C_k, C_l]_{\vec{\theta}}
\end{aligned}
\tag{S5}
$$

and evaluate them at the point $\vec{\theta} = \vec{\theta}^*$, to get

$$
\frac{\partial^2}{\partial\theta_k\partial\theta_l}\mathcal{L}^*(\vec{\theta})\bigg|_{\vec{\theta}=\vec{\theta}^*} = -\text{Cov}[C_k, C_l]_{\vec{\theta}^*} = -\Sigma_{k,l}^*
\tag{S6}
$$

where $\Sigma_{k,l}^*$, as in eq. (14), denotes the covariance of the constraints C_k and C_l under the canonical probability P_{can}^*. The above result implies that, at the particular point $\vec{\theta}^*$, the Hessian matrix $\mathbf{\Lambda}^*$ of second derivatives of $\mathcal{L}^*$ is equal to minus the covariance matrix $\mathbf{\Sigma}^*$ of the constraints.

Since covariance matrices are positive semi-definite (i.e. their eigenvalues are all non-negative), $\mathbf{\Lambda}^*$ is negative semi-definite (i.e. its eigenvalues are all non-positive). The only case when some eigenvalues of $\mathbf{\Sigma}^*$ are zero is when some of the K constraints are not linearly independent of the other constraints. This means that there are some redundant constraints, i.e. some of the imposed properties are linear combinations of other imposed properties. While this situation is generally not encountered, it may happen in certain phases of some models. If we exclude this circumstance, $\mathbf{\Sigma}^*$ is positive definite and $\mathbf{\Lambda}^*$ is negative definite, which implies that $\vec{\theta}^*$ is a maximum of $\mathcal{L}^*(\vec{\theta})$. Since it is the only stationary point, it is necessarily a global maximum for $\mathcal{L}^*(\vec{\theta})$.

By contrast, in the case when only an effective number $\tilde{K} < K$ if constraints are linearly independent of each other, there will be degenerate maxima for $\mathcal{L}^*(\vec{\theta})$. If we order the constraints in such a way that the first $\tilde{K}$ ones are all linearly independent, while the following $K - \tilde{K}$ ones are dependent on the first $\tilde{K}$ ones, then $\vec{\theta}^*$ will be a global maximum of $\mathcal{L}^*(\vec{\theta})$ in the subspace spanned by the first $\tilde{K}$ Lagrange multipliers, while along the other $K - \tilde{K}$ directions $\mathcal{L}^*(\vec{\theta})$ will be constant. Correspondingly, there will be a $(K - \tilde{K})$-dimensional family of degenerate solutions $\vec{\theta}^*$ to eq. (4).

In either case, the above discussion proves eq. (S2), which is a crucial result allowing us to prove a variety of other identities below.

Relation between canonical entropy and likelihood

We here prove the important result that the maximized log-likelihood $\mathcal{L}^*(\vec{\theta}^*)$ coincides with $-S_{\text{can}}^*$. This is easily demonstrated as follows:

$$
\begin{aligned}
\mathcal{L}^*(\vec{\theta}^*) &= \ln P_{\text{can}}^*(\mathbf{G}^*) \\
&= -\vec{\theta}^*\cdot\vec{C}(\mathbf{G}^*) - \ln\mathcal{Z}(\vec{\theta}^*) \\
&= -\vec{\theta}^*\cdot\vec{C}^* - \ln\mathcal{Z}(\vec{\theta}^*) \\
&= -\vec{\theta}^*\cdot\langle\vec{C}|\vec{\theta}^*\rangle - \ln\mathcal{Z}(\vec{\theta}^*) \\
&= \langle\ln P_{\text{can}}^*(\mathbf{G})|\vec{\theta}^*\rangle \\
&= -\sum_{\mathbf{G}} P_{\text{can}}^*(\mathbf{G})\ln P_{\text{can}}^*(\mathbf{G}) \\
&= -S_{\text{can}}^*.
\end{aligned}
\tag{S7}
$$

The above result proves eq. (8) and will be useful multiple times in the following.

Exact relation between $\Omega_{\vec{C}^*}$ and P_{can}^*

Here we prove one of our main results, i.e. the exact formula (13) connecting $\Omega_{\vec{C}^*}$, P_{mic}^* and P_{can}^*. Recall that we assumed that $C_k(\mathbf{G})$ can only take integer values, e.g. because it counts certain properties of $\mathbf{G}$. Importantly, we also assume that the parameter value $\vec{\theta}^*$ conjugate to $\vec{C}^*$ is not a critical point of any phase transition. The possibility that $\vec{\theta}^*$ approaches a phase boundary is discussed separately elsewhere in the paper.

Using the integral representation $\delta_{x,y} = \int_{-\pi}^{+\pi} \frac{d\psi}{2\pi} e^{i\psi(x-y)}$ of the Kronecker symbol (where $\psi \in [-\pi, +\pi] \subset \mathbb{R}$), and extending it to the K-dimensional case, we can write

$$
\begin{aligned}
\Omega_{\vec{C}^*} &= \sum_{\mathbf{G} \in \mathcal{G}} \delta_{\vec{C}^*, \vec{C}(\mathbf{G})} \\
&= \sum_{\mathbf{G} \in \mathcal{G}} \prod_{k=1}^{K} \int_{-\pi}^{+\pi} \frac{d\psi_k}{2\pi} e^{i\psi_k[C_k^* - C_k(\mathbf{G})]} \\
&= \sum_{\mathbf{G} \in \mathcal{G}} \prod_{k=1}^{K} \int_{\alpha_k} \frac{dz_k}{2\pi i} e^{z_k[C_k^* - C_k(\mathbf{G})]}, \quad \text{(S8)}
\end{aligned}
$$

where in the last equality we have rewritten the integral in terms of the complex variable $z_k = \theta_k + i\psi_k$ (with θ_k and ψ_k real) and reinterpreted the real domain of integration $[-\pi, +\pi] \subset \mathbb{R}$ for ψ_k as a segment $\alpha_k \subset \mathbb{C}$ for z_k, going from the point of (real, imaginary) coordinates $(0, -\pi)$ to the point $(0, +\pi)$ along the imaginary axis. This trick, although apparently pointless, allows us to perform a convenient operation, as we now explain.

Note that the function $f_k(z_k) = e^{z_k[C_k^* - C_k(\mathbf{G})]}$ being integrated is analytical over the entire complex plane (hence it is an entire function) and its integral along any closed contour is therefore zero:

$$
\oint dz_k f_k(z_k) = 0. \quad \text{(S9)}
$$

For convenience, let us consider the closed rectangular contour formed by the following four segments in the complex plane: the aforementioned upward vertical segment α_k going from $(0, -\pi)$ to $(0, +\pi)$, the horizontal segment β_k going from $(0, +\pi)$ to $(\theta_k^*, +\pi)$, the downward vertical segment γ_k going from $(\theta_k^*, +\pi)$ to $(\theta_k^*, -\pi)$, and the horizontal segment δ_k going from $(\theta_k^*, -\pi)$ back to the starting point $(0, -\pi)$. We can rewrite eq. (S9) for this particular contour as

$$
\left[\int_{\alpha_k} dz_k + \int_{\beta_k} dz_k + \int_{\gamma_k} dz_k + \int_{\delta_k} dz_k \right] f_k(z_k) = 0. \quad \text{(S10)}
$$

Now, since $C_k(\mathbf{G})$ (and hence C_k^*) is assumed to take only integer values, we have the following periodicity:

$$
f_k(z_k + 2\pi i) = e^{(z_k + 2\pi i)[C_k^* - C_k(\mathbf{G})]} = f_k(z_k). \quad \text{(S11)}
$$

This implies that, for any value of θ_k, f_k takes the same value in each pair of points of the type $(\theta_k, -\pi)$, $(\theta_k, +\pi)$.

Since the segments β_k and δ_k go over such pairs of points in opposite direction, we have

$$
\left[\int_{\beta_k} dz_k + \int_{\delta_k} dz_k \right] f_k(z_k) = 0 \quad \text{(S12)}
$$

which, together with eq. (S10), implies

$$
\int_{\alpha_k} dz_k f_k(z_k) = -\int_{\gamma_k} dz_k f_k(z_k) = \int_{\omega_k} dz_k f_k(z_k), \quad \text{(S13)}
$$

where ω_k is the upward vertical segment going from $(\theta_k^*, -\pi)$ to $(\theta_k^*, +\pi)$, i.e. the same as γ_k but going in opposite direction. Using eq. (S13), and denoting with $\vec{\omega}$ the collection of all segments $\{\omega_k\}_{k=1}^{K}$, eq. (S8) becomes

$$
\begin{aligned}
\Omega_{\vec{C}^*} &= \sum_{\mathbf{G} \in \mathcal{G}} \prod_{k=1}^{K} \int_{\omega_k} \frac{dz_k}{2\pi i} e^{z_k[C_k^* - C_k(\mathbf{G})]} \\
&= \sum_{\mathbf{G} \in \mathcal{G}} \int_{\vec{\omega}} \frac{d\vec{z}}{(2\pi i)^K} e^{\vec{z} \cdot [\vec{C}^* - \vec{C}(\mathbf{G})]} \\
&= \int_{\vec{\omega}} \frac{d\vec{z}}{(2\pi i)^K} e^{\vec{z} \cdot \vec{C}(\mathbf{G}^*)} \sum_{\mathbf{G} \in \mathcal{G}} e^{-\vec{z} \cdot \vec{C}(\mathbf{G})} \\
&= \int_{\vec{\omega}} \frac{d\vec{z}}{(2\pi i)^K} e^{\mathcal{H}(\mathbf{G}^*, \vec{z})} \mathcal{Z}(\vec{z}) \\
&= \int_{\vec{\omega}} \frac{d\vec{z}}{(2\pi i)^K} P_{\mathrm{can}}^{-1}(\mathbf{G}^*|\vec{z}), \\
&= \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} P_{\mathrm{can}}^{-1}(\mathbf{G}^*|\vec{\theta}^* + i\vec{\psi}), \quad \text{(S14)}
\end{aligned}
$$

where we have extended the domain of $\mathcal{H}(\mathbf{G}^*, \vec{\theta})$, $\mathcal{Z}(\vec{\theta})$ and $P_{\mathrm{can}}(\mathbf{G}^*|\vec{\theta})$, all viewed as functions of $\vec{\theta}$, from $\mathbb{R}^K$ to $\mathbb{C}^K$. In so doing, they have become complex functions of complex vectors. To produce the last equality, we have rewritten the integral for z_k along the complex segment $\omega_k \subset \mathbb{C}$ back as an integral for ψ_k along the real interval $[-\pi, +\pi] \subset \mathbb{R}$. Correspondingly, we have denoted $\pm\vec{\pi} = \pm\pi\vec{1}$, where $\vec{1}$ is the K-dimensional vector with all unit entries.

Note that, to be able to interchange the order of sum and integral in the derivation of eq. (S14) even in the limit $n \to \infty$ (which we will take later on), we have to assume uniform convergence of the partial sums of $\sum_{\mathbf{G} \in \mathcal{G}} e^{-\vec{z} \cdot \vec{C}(\mathbf{G})}$ (which becomes an infinite series in such limit) to $\mathcal{Z}(\vec{z})$. Note that our assumptions also ensure that $\mathcal{Z}(\vec{z})$ is a holomorphic (or analytic) function in a domain containing $\vec{\omega}$. For finite n, this is automatically ensured by the fact that $\mathcal{Z}(\vec{z})$ is a finite sum of exponentials (which are all holomorphic). In the thermodynamic limit $n \to \infty$, the analyticity of $\mathcal{Z}(\vec{z})$ is ensured by our assumption that $\vec{\omega}$ does not cross any boundary between different thermodynamic phases, i.e. that the point $\vec{\theta}^*$, which is the only real point crossed by $\vec{\omega}$, is not a critical point of any phase transition.

Recalling from eq. (2) that $P_{\mathrm{mic}}(\mathbf{G}^*|\vec{C}^*) = \Omega_{\vec{C}^*}^{-1}$, eq. (S14) immediately proves eq. (13) in the main text,

which provides an exact relationship between the micro-canonical and the canonical probability extended to the complex domain. Remarkably, this relationship is valid even for finite n and even if EE does not hold.

Asymptotic relation between Ω^* and the canonical (co)variances

Here we prove another main result, i.e. we start from eq. (S14) and, using a saddle point calculation, arrive at its asymptotic expression given in eq. (15). We keep assuming that $\vec{\theta}^*$ is not a critical point of any phase transition.

First, extending the log-likelihood function to the complex domain via the definition

$$\mathcal{L}^*(\vec{z}) \equiv \ln P_{\text{can}}(\mathbf{G}^*|\vec{z}), \qquad \text{(S15)}$$

we rewrite eq. (S14) as

$$\Omega_{\vec{C}^*} = \int_{\vec{\omega}} \frac{d\vec{z}}{(2\pi i)^K} \, e^{-\mathcal{L}^*(\vec{z})}. \qquad \text{(S16)}$$

Second, we note that $\mathcal{L}^*(\vec{z})$ is a holomorphic function, which follows from the analyticity of $P_{\text{can}}(\mathbf{G}^*|\vec{z})$ away from phase transitions and from the fact that, when $P_{\text{can}}(\mathbf{G}^*|\vec{z})$ is real-valued, it is positive and hence its logarithm is analytic.

Third, we notice that the point $\vec{z}^* \in \vec{\omega}$ of coordinates $(\vec{\theta}^*, \vec{0})$ is a saddle point for $\mathcal{L}^*(\vec{z})$. To see this, we extend the calculation in eq. (S3) to the complex domain to get

$$\frac{\partial}{\partial z_k} \mathcal{L}^*(\vec{z}) = \langle C_k|\vec{z}\rangle - C_k^* \qquad \text{(S17)}$$

and note that eq. (S4) implies

$$\left. \frac{\partial}{\partial z_k} \mathcal{L}^*(\vec{z}) \right|_{\vec{z}=\vec{z}^*} = \langle C_k|\vec{\theta}^*\rangle - C_k^* = 0 \quad \forall k. \qquad \text{(S18)}$$

The above result shows that $\vec{z}^*$ is a stationary point for $\mathcal{L}^*(\vec{z})$. As always happens in the complex domain, such a point is necessarily a saddle point. This can be confirmed here by noting that, as we move across $\vec{z}^*$ along the real direction, the second derivatives are given by eq. (S6) and the corresponding Hessian matrix is, as already mentioned, negative semi-definite. On the contrary, as we show later, as we move across $\vec{z}^*$ along the imaginary direction, all second derivatives have the opposite sign and the Hessian matrix is therefore positive semi-definite. The sign difference of the second derivatives calculated along the real and imaginary directions implies that $\vec{z}^*$ is a saddle point.

Fourth, we notice that $\vec{z}^*$ is the only relevant saddle point for evaluating eq. (S16). Due to the periodicity $\mathcal{L}^*(\vec{z}) = \mathcal{L}^*(\vec{z}+2\pi i\vec{u})$ where $\vec{u}$ is any vector in $\mathbb{Z}^K$, all the points of the type $\vec{z}^*+2\pi i\vec{u}$, having coordinates $(\vec{\theta}^*, 2\pi\vec{u})$, are also saddle points for $\mathcal{L}^*(\vec{z})$. However, among all

these saddle points, $\vec{z}^*$ is the only one encountered along the path $\vec{\omega}$. It is then reasonable to assume that there are no other saddle points along $\vec{\omega}$, i.e. no other points $(\vec{\theta}^*, \vec{\psi})$ with $\vec{\psi} \neq 0$ such that $\langle \vec{C}|\vec{\theta}^* + i\vec{\psi}\rangle = \vec{C}^*$ (the existence of such points is unlikely, given that $\vec{C}^* \in \mathbb{R}^K$). At the very least, we assume that those points, if present, can be avoided via a suitable deformation of $\vec{\omega}$ (since $\mathcal{L}^*(\vec{z})$ is analytical, such deformation is possible, unless those points form an infinite curve separating $\vec{z}^*$ from the end-points $\pm\vec{\pi}$ of integration, or a closed curve surrounding $\vec{z}^*$ and/or $\pm\vec{\pi}$).

The above considerations imply that, in the thermodynamic limit, eq. (S16) naturally lends itself to a saddle-point calculation [22] leading to eq. (15). We now show this in detail. We define the complex-valued function

$$\mathcal{U}^*(\vec{\psi}) = \mathcal{L}^*(\vec{\theta}^* + i\vec{\psi}), \quad \mathcal{U}^* : \mathbb{R}^K \to \mathbb{C} \qquad \text{(S19)}$$

and Taylor-expand it around $\vec{\psi} = \vec{0}$:

$$\begin{aligned}
\mathcal{U}^*(\vec{\psi}) &= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{\sum_k a_k = m} \frac{\partial^m}{\partial \psi_1^{a_1} \dots \partial \psi_K^{a_K}} \mathcal{U}(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} \prod_{k=1}^{K} \psi_k^{a_k} \\
&= \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} + \sum_{k=1}^{K} \psi_k \frac{\partial}{\partial \psi_k} \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} \\
&\quad + \frac{1}{2} \sum_{k,l} \psi_k \psi_l \frac{\partial^2}{\partial \psi_k \partial \psi_l} \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} + \mathcal{V}^*(\vec{\psi}), \text{(S20)}
\end{aligned}$$

where $\mathcal{V}^*(\vec{\psi}) = O(|\vec{\psi}|^3)$ contains all the higher-order terms and is such that $\mathcal{V}^*(\vec{0}) = 0$. Note that the function $\mathcal{U}^*$ has the periodicity $\mathcal{U}^*(\vec{\psi}) = \mathcal{U}^*(\vec{\psi} + 2\pi\vec{u})$ for any $\vec{u} \in \mathbb{Z}^K$, therefore the above expansion only makes sense within one such period. This is consistent with the fact that, in eq. (S14), the domain of integration for each variable ψ_k is the single period $[-\pi, +\pi]$. Now, from the identity

$$\frac{\partial^m}{\partial \psi_1^{a_1} \dots \partial \psi_K^{a_K}} \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} = i^m \frac{\partial^m}{\partial \theta_1^{a_1} \dots \partial \theta_k^{a_k}} \mathcal{L}^*(\vec{\theta}) \bigg|_{\vec{\theta}=\vec{\theta}^*}$$

we realize that all the m-th order derivatives of $\mathcal{U}^*(\vec{\psi})$ at $\vec{\psi} = \vec{0}$ are real if m is even and purely imaginary if m is odd, because all the derivatives of $\mathcal{L}^*(\vec{\theta})$, when evaluated at $\vec{\theta} = \vec{\theta}^*$, are real. Moreover, we can use eqs. (S4), (S6) and (S7) to obtain

$$\mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} = \mathcal{L}^*(\vec{\theta}) \bigg|_{\vec{\theta}=\vec{\theta}^*} = -S_{\text{can}}^*, \qquad \text{(S21)}$$

$$\frac{\partial}{\partial \psi_k} \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} = i \frac{\partial}{\partial \theta_k} \mathcal{L}^*(\vec{\theta}) \bigg|_{\vec{\theta}=\vec{\theta}^*} = 0, \qquad \text{(S22)}$$

$$\frac{\partial^2}{\partial \psi_k \partial \psi_l} \mathcal{U}^*(\vec{\psi}) \bigg|_{\vec{\psi}=\vec{0}} = -\frac{\partial^2}{\partial \theta_k \partial \theta_l} \mathcal{L}^*(\vec{\theta}) \bigg|_{\vec{\theta}=\vec{\theta}^*} = \Sigma_{k,l}^*. \text{(S23)}$$

Equation (S23) confirms that $\vec{z}^*$ is a stationary point for $\mathcal{L}^*(\vec{z})$, as anticipated.

The expansion in eq. (S20) can then be rewritten as

$$\mathcal{U}^*(\vec{\psi}) = -S^*_{\text{can}} + \frac{1}{2}\sum_{k,l}\psi_k\psi_l\Sigma^*_{k,l} + \mathcal{V}^*(\vec{\psi}) \quad (S24)$$

and, via eq. (S19), we can rewrite eq. (S16) as

$$
\begin{aligned}
\Omega_{\vec{C}^*} &= \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{-\mathcal{L}^*(\vec{\theta}^*+i\vec{\psi})} \\
&= \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{-\mathcal{U}^*(\vec{\psi})}, \\
&= e^{S^*_{\text{can}}} \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{-\frac{1}{2}\vec{\psi}\,\mathbf{\Sigma}^*\vec{\psi}} e^{-\mathcal{V}^*(\vec{\psi})}. \quad (S25)
\end{aligned}
$$

We are now ready for a saddle-point calculation [22]. Rather than using the multidimensional version of the method, which requires the existence of a common large factor at the exponent for all the coordinates of $\vec{\psi}$, we prefer iterating the one-dimensional version upon diagonalizing $\mathbf{\Sigma}^*$, so that we can keep track of the asymptotic behaviour of each of its eigenvalues separately. Let us therefore carry out a volume-preserving (i.e. with unit Jacobian determinant) linear change of variables $\vec{\psi} \mapsto \vec{\xi}(\vec{\psi})$ that diagonalizes $\mathbf{\Sigma}^*$ while keeping $\vec{\xi}(\vec{0}) = \vec{0}$ and rewrite

$$\Omega_{\vec{C}^*} = e^{S^*_{\text{can}}} \int_{\vec{\xi}^-}^{\vec{\xi}^+} \frac{d\vec{\xi}}{(2\pi)^K} \frac{e^{-\frac{1}{2}\sum_{k=1}^K \lambda_k^* \xi_k^2}}{e^{\mathcal{W}^*(\vec{\xi})}}, \quad (S26)$$

where λ_k^* is the k-th eigenvalue of $\mathbf{\Sigma}^*$, $\mathcal{W}^*(\vec{\xi}) = \mathcal{V}^*[\vec{\psi}(\vec{\xi})]$ (so that $\mathcal{W}^*(\vec{0}) = \mathcal{V}^*(\vec{0}) = 0$), and $\vec{\xi}^{\pm} = \vec{\xi}(\pm\vec{\pi})$.

As discussed previously, the K eigenvalues of $\mathbf{\Sigma}^*$ cannot be negative. We already mentioned that if only $\tilde{K} \leq K$ constraints are mutually independent, then $K - \tilde{K}$ eigenvalues will be zero. Let us now carry out the integration in eq. (S26) over a single variable ξ_l for which the corresponding eigenvalue λ_l^* is zero. In such a case, eq. (S24) implies that, along the direction ξ_l, the function $\mathcal{U}^*[\vec{\psi}(\vec{\xi})]$ is constant and equal to the value $\mathcal{U}^*[\vec{\psi}(\vec{\xi})]\big|_{\xi_l=0}$. This means that, along ξ_l, $\mathcal{W}^*(\vec{\xi})$ is also constant and equal to $\mathcal{W}^*(\vec{\xi})\big|_{\xi_l=0}$. Therefore, if $\lambda_l^* = 0$ then in eq. (S26) the integral over ξ_l is the definite integral of a constant function, which generates a contribution $\xi_l^+ - \xi_l^- = 2\pi$. This follows from the fact that the trasformation $\vec{\psi} \mapsto \vec{\xi}$ is volume-preserving, so it should keep the area along each redundant direction equal to the original area 2π (because each such direction can be added arbitrarily without changing the volume spanned by the non-redundant directions). This contribution effectively replaces the factor $(2\pi)^K$ at the denominator with $(2\pi)^{K-1}$.

$$\Omega_{\vec{C}^*} = e^{S^*_{\text{can}}} \int_{\vec{\xi}_{-l}^-}^{\vec{\xi}_{-l}^+} \frac{d\vec{\xi}_{-l}}{(2\pi)^{K-1}} \frac{e^{-\frac{1}{2}\sum_{k\neq l}\lambda_k^*\xi_k^2}}{e^{\mathcal{W}^*(\vec{\xi})}\big|_{\xi_l=0}}, \quad (S27)$$

where $\vec{\xi}_{-l}$ denotes the $(K-1)$-dimensional vector including all variables except ξ_l. Repeating the procedure for all the zero eigenvalues, the factor $(2\pi)^{K-1}$ becomes $(2\pi)^{\tilde{K}}$.

Let us now start again from eq. (S26) and consider an eigenvalue λ_l^* that is strictly positive. The corresponding term $e^{-\frac{1}{2}\lambda_l^*\xi_l^2}$ in the integral is now an unnormalized Gaussian function with zero mean and variance $1/\lambda_l^*$. If λ_l^* grows with n, the Gaussian function $e^{-\frac{1}{2}\lambda_l^*\xi_l^2}$ will become more and more peaked aroud zero as $n \to \infty$ and its limit becomes the Dirac delta function $\delta(\xi_l)$ times the normalization factor $\sqrt{2\pi/\lambda_l^*}$. Integrating over the variable ξ_l, for $n \to \infty$ the limiting $\delta(\xi_l)$ will select the unique value $e^{-\mathcal{W}^*(\vec{\xi})}\big|_{\xi_l=0}$ in eq. (S26). If λ_l^* remains finite (and positive) as n grows, there is an additional correction factor that can be estimated [22] as $[1 + O(1/\lambda_l^*)]$. So in general we get

$$\Omega_{\vec{C}^*} = \frac{e^{S^*_{\text{can}}}}{\sqrt{\lambda_l^*}} \int_{\vec{\xi}_{-l}^-}^{\vec{\xi}_{-l}^+} \frac{d\vec{\xi}_{-l}}{(2\pi)^{K-1/2}} \frac{e^{-\frac{1}{2}\sum_{k\neq l}\lambda_k^*\xi_k^2}}{e^{\mathcal{W}^*(\vec{\xi})}\big|_{\xi_l=0}} [1 + O(1/\lambda_l^*)] \quad (S28)$$

where $\vec{\xi}_{-l}$ denotes again the $(K-1)$-dimensional vector including all variables except ξ_l.

Iterating the above calculations for all eigenvalues, and using either eq.(S27) (whenever $\lambda_l^* = 0$) or eq. (S28) (whenever $\lambda_l^* > 0$), we arrive at the asymptotic expansion

$$
\begin{aligned}
\Omega_{\vec{C}^*} &= \frac{e^{S^*_{\text{can}}}}{\sqrt{\lambda_1^* \cdots \lambda_{\tilde{K}}^*}} \frac{\prod_{k=1}^{\tilde{K}}[1 + O(1/\lambda_k^*)]}{(2\pi)^{\tilde{K}-\tilde{K}/2}\, e^{\mathcal{W}^*(\vec{0})}} \\
&= \frac{e^{S^*_{\text{can}}}}{\sqrt{(2\pi)^{\tilde{K}}\det\tilde{\mathbf{\Sigma}}^*}} \prod_{k=1}^{\tilde{K}}[1 + O(1/\lambda_k^*)] \\
&= \frac{e^{S^*_{\text{can}}} T^*}{\sqrt{\det(2\pi\tilde{\mathbf{\Sigma}}^*)}}, \quad (S29)
\end{aligned}
$$

where $\tilde{\mathbf{\Sigma}}^*$ is the reduced $\tilde{K} \times \tilde{K}$ matrix of the canonical (co)variances of the $\tilde{K} \leq K$ independent constraints, and we have reordered the original K constraints in such a way that the first $\tilde{K}$ ones are independent (so the first $\tilde{K}$ eigenvalues are non-zero). The above derivation proves eq. (15). The asymptotic behaviour of T^* as a function of that of the eigenvalues of $\tilde{\mathbf{\Sigma}}^*$ is discussed later.

Exact calculation of Δ^*

We now use the previous results in order to calculate the relative entropy Δ^* explicitly. Inserting eq. (S14)

into eq. (18) implies

$$
\begin{aligned}
\Delta^* &= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} \frac{P_{\mathrm{can}}(\mathbf{G}^*|\vec{\theta}^*)}{P_{\mathrm{can}}(\mathbf{G}^*|\vec{\theta}^* + i\vec{\psi})} \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} \frac{e^{-\mathcal{H}(\mathbf{G}^*,\vec{\theta}^*)}}{\mathcal{Z}(\vec{\theta}^*)} \frac{\mathcal{Z}(\vec{\theta}^* + i\vec{\psi})}{e^{-\mathcal{H}(\mathbf{G}^*,\vec{\theta}^* + i\vec{\psi})}} \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{-\vec{\theta}^* \cdot \vec{C}^* + \vec{\theta}^* \cdot \vec{C}^* + i\vec{\psi} \cdot \vec{C}^*} \frac{\mathcal{Z}(\vec{\theta}^* + i\vec{\psi})}{\mathcal{Z}(\vec{\theta}^*)} \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{i\vec{\psi} \cdot \vec{C}^*} \frac{\mathcal{Z}(\vec{\theta}^* + i\vec{\psi})}{\mathcal{Z}(\vec{\theta}^*)} \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} e^{i\vec{\psi} \cdot \vec{C}^*} \langle e^{-i\vec{\psi} \cdot \vec{C}} | \vec{\theta}^* \rangle \\
&= -\ln \int_{-\vec{\pi}}^{+\vec{\pi}} \frac{d\vec{\psi}}{(2\pi)^K} \langle e^{i\vec{\psi} \cdot (\vec{C}^* - \vec{C})} | \vec{\theta}^* \rangle \quad \text{(S30)}
\end{aligned}
$$

where we have used the definition (5) of ensemble average of the (complex) function $X(\mathbf{G}) = e^{i\vec{\psi} \cdot [\vec{C}^* - \vec{C}(\mathbf{G})]}$ and exploited the fact that

$$
\begin{aligned}
\frac{\mathcal{Z}(\vec{\theta}^* + i\vec{\psi})}{\mathcal{Z}(\vec{\theta}^*)} &= \frac{\sum_{\mathbf{G}} e^{-(\vec{\theta}^* + i\vec{\psi}) \cdot \vec{C}(\mathbf{G})}}{\sum_{\mathbf{G}} e^{-\vec{\theta}^* \cdot \vec{C}(\mathbf{G})}} \\
&= \sum_{\mathbf{G}} P_{\mathrm{can}}(\mathbf{G}|\vec{\theta}^*) e^{-i\vec{\psi} \cdot \vec{C}} \\
&= \langle e^{-i\vec{\psi} \cdot \vec{C}} | \vec{\theta}^* \rangle \quad \text{(S31)}
\end{aligned}
$$

Equation (S30) proves eq. (19) in the main text. If the integral can be carried out explicitly, eq. (S30) allows an exact calculation of Δ^* in terms of only the canonical probability P_{can}^*, thus avoiding any microcanonical enumeration.

Asymptotic behaviour of T^*

We now study the asymptotic behaviour of the correction term T^* appearing in eqs. (15) and (20).

We already noted that, in general, some of the nonzero eigenvalues of $\tilde{\boldsymbol{\Sigma}}^*$ may diverge, while some may remain finite, as $n \to \infty$. If there are $K_\infty \le \tilde{K}$ diverging eigenvalues, and if we choose an ordering such that these eigenvalues are the first ones, then their contribution to T^* is

$$
\prod_{k=1}^{K_\infty} [1 + O(1/\lambda_k^*)] = 1 + O\left(\sum_{k=1}^{K_\infty} 1/\lambda_k^*\right) = 1 + O(K_\infty/\lambda_\infty^*)
$$

$$(S32)$$

where λ_∞^* is the (diverging) harmonic mean of the K_∞ diverging eigenvalues:

$$
\frac{1}{\lambda_\infty^*} = \frac{1}{K_\infty} \sum_{k=1}^{K_\infty} \frac{1}{\lambda_k^*}. \quad \text{(S33)}
$$

The remaining $\tilde{K} - K_\infty$ finite eigenvalues give a contribution to T^* equal to

$$
\begin{aligned}
\prod_{k=K_\infty+1}^{\tilde{K}} [1 + O(1/\lambda_k^*)] &= \prod_{k=K_\infty+1}^{\tilde{K}} O(1/\lambda_k^*) \\
&= O\left((\bar{\lambda}^*)^{-(\tilde{K} - K_\infty)}\right), \text{(S34)}
\end{aligned}
$$

where

$$
\bar{\lambda}^* \equiv \left(\prod_{k=K_\infty+1}^{\tilde{K}} \lambda_k^*\right)^{1/(\tilde{K} - K_\infty)} \quad \text{(S35)}
$$

is the (finite) geometric mean of the $\tilde{K} - K_\infty$ finite eigenvalues.

Combining eqs. (S32) and (S34), we arrive at

$$
T^* = [1 + O(K_\infty/\lambda_\infty^*)] O[(\bar{\lambda}^*)^{-(\tilde{K} - K_\infty)}] \quad \text{(S36)}
$$

and, taking the logarithm, we get

$$
\ln T^* = \ln[1 + O(K_\infty/\lambda_\infty^*)] + O(\tilde{K} - K_\infty). \quad \text{(S37)}
$$

Correspondingly, the quantity $\tau_{\alpha_\infty}^*$ defined in eq. (24) becomes

$$
\tau_{\alpha_\infty}^* = -\lim_{n \to \infty} \frac{\ln[1 + O(K_\infty/\lambda_\infty^*)] + O(\tilde{K} - K_\infty)}{\alpha_n}
$$

$$(S38)$$

The above equation is very important as it informs us about the limiting behaviour of $\delta_{\alpha_\infty}^*$, as we now discuss.

Let us first consider the case when all the nonzero eigenvalues are diverging, i.e. $K_\infty = \tilde{K}$. Note that this is the typical situation, since the constraints are generally macroscopic quantities whose values, as well as those of their fluctuations (hence the eigenvalues of the covariance matrix), diverge at least like n in the thermodynamic limit. If the number of constraints remains finite as $n \to \infty$ (i.e. $K_\infty < \infty$), then $\ln T^* = O(1/\lambda_\infty^*) \to 0$ and $\Delta^* \to \ln \sqrt{\det(2\pi\tilde{\boldsymbol{\Sigma}}^*)}$. Stated more rigorously, for any increasing α_n we get $\tau_{\alpha_\infty}^* = 0$. If all eigenvalues are diverging, and the number K_∞ of constraints diverges like λ_∞^*, $\ln T^*$ has now a finite limit but still vanishes upon dividing by α_n. We therefore still obtain $\tau_{\alpha_\infty}^* = 0$. The same holds true as long as K_∞ diverges slower than the astronomically big number $\lambda_\infty^* e^{\alpha_n}$. Therefore $\tau_{\alpha_\infty}^* = 0$ whenever all eigenvalues diverge.

In the opposite case when all eigenvalues have a finite limit as $n \to \infty$ (i.e. $K_\infty = 0$), eq. (S37) implies $\ln T^* = O(\tilde{K})$. Upon dividing by α_n and taking the limit, this term vanishes, so $\tau_{\alpha_\infty}^* = 0$ still holds, as long as the number $\tilde{K}$ of constraints is bounded or diverges slower than α_n. By contrast, if the number of constraints grows like α_n or faster, $\tau_{\alpha_\infty}^* > 0$ and its value has to be evaluated case by case.

Combining all the results discussed so far, we conclude that eq. (23) with $\tau_{\alpha_\infty}^* = 0$ always holds as an exact result, unless $\boldsymbol{\Sigma}^*$ has an infinite number (growing at least like α_n) of finite eigenvalues, in which case $\tau_{\alpha_\infty}^* > 0$ and has to be calculated for the specific model under scrutiny.