Supporting Information

Lin and Lindsey 10.1073/pnas.1720782115

Proof of the Classical Gibbs Variational Principle in *Theorem 1*

For any $A \in \mathcal{S}^n$, let $\rho \in \mathcal{M}$, and define $H(x) := \frac{1}{2}x^T Ax + U(x)$ and $\eta := \frac{1}{Z[A]} e^{-H} \in \mathcal{M}$. Then the relative entropy of ρ with respect to η is defined as $S_\eta(\rho) = -\int \rho \log \frac{\rho}{\eta}\, dx$. (Note the sign convention; here, the relative entropy is the negative of the Kullback–Leibler divergence.) For fixed $\eta \in \mathcal{M}$, S_η is nonpositive and, moreover, $S_\eta(\rho) = 0$ if and only if $\rho = \eta$. Hence

$$\int H\, \rho\, dx - S(\rho) = \Omega[A] - S_\eta(\rho) \geq \Omega[A]. \qquad \textbf{[S1]}$$

The inequality in Eq. **S1** holds if and only if $\rho = \eta$. This proves that

$$\Omega[A] = \inf_{\rho \in \mathcal{M}} \left[\int H\, \rho\, dx - S(\rho) \right],$$

as well as the fact that this infimum is uniquely attained at the probability density $\rho = \eta$.

□

Proof That $\mathcal{F}$ Is Concave in *Theorem 1*

Let $G_1, G_2 \in \mathcal{S}^n_{++}$, $\theta \in [0,1]$, and $\varepsilon > 0$. Note that $\mathcal{F}$ can be written

$$\mathcal{F}[G] = \sup_{\rho \in \mathcal{G}^{-1}(G)} \Psi[\rho], \quad \Psi[\rho] := S(\rho) - \int U\, \rho\, dx.$$

Furthermore let $\rho_1, \rho_2 \in \mathcal{M}$ such that $\rho_i \in \mathcal{G}^{-1}(G_i)$ and $\Psi[\rho_i] \geq \mathcal{F}[G_i] - \varepsilon$. Then, noting that $\theta\rho_1 + (1-\theta)\rho_2 \in \mathcal{G}^{-1}(\theta G_1 + (1-\theta)G_2)$, we observe

$$\begin{aligned}\mathcal{F}[\theta G_1 + (1-\theta)G_2] &= \sup_{\rho \in \mathcal{G}^{-1}(\theta G_1 + (1-\theta)G_2)} \Psi[\rho] \\ &\geq \Psi[\theta\rho_1 + (1-\theta)\rho_2] \\ &\geq \theta\Psi[\rho_1] + (1-\theta)\Psi[\rho_2] \\ &\geq \theta\mathcal{F}[G_1] + (1-\theta)\mathcal{F}[G_2] - \varepsilon,\end{aligned}$$

where the penultimate step uses convexity of Ψ. Since ε was arbitrary, we have established concavity.

□

Proof That $\mathcal{F}$ Diverges to $-\infty$ at the Boundary of $\mathcal{S}^n_{++}$ in *Theorem 1*

Recall that $\mathcal{F}$ is defined by

$$\mathcal{F}[G] = \sup_{\rho \in \mathcal{G}^{-1}(G)} \left[S(\rho) - \int U\, \rho\, dx \right].$$

Now for $\rho \in \mathcal{G}^{-1}(G)$, the entropy of ρ is bounded above by the entropy of the Gaussian distribution $\mathcal{N}(0, G)$; i.e., $S(\rho) \leq \frac{1}{2}\log((2\pi e)^n \det G)$. Furthermore, since U satisfies the strong growth condition, U is in particular bounded below; i.e., $U(x) \geq C$ for some constant C and all $x \in \mathbb{R}^n$. Therefore, for any $\rho \in \mathcal{G}^{-1}(G)$, we have

$$\begin{aligned}S(\rho) - \int U\rho\, dx &\leq \frac{1}{2}\log((2\pi e)^n \det G) - C \\ &= \frac{1}{2}\log\det G + C',\end{aligned}$$

where C' is a new constant. This implies that $\mathcal{F}[G] \leq \frac{1}{2}\log\det G + C'$, and in particular $\mathcal{F}$ diverges to $-\infty$ at the boundary of $\mathcal{S}^n_{++}$.

□

Proof of the Transformation Rule (*Proposition 4*)

Let $G \in \mathcal{S}^n_{++}$. Using $\mathrm{Tr}[\log(G)] = \log\det(G)$, we have

$$\Phi[G; U] = -\Phi_0 - 2 \inf_{\rho \in \mathcal{G}^{-1}(G)} \left[\int \left(\log\left[(\det G)^{1/2}\rho \right] + U \right) \rho\, dx \right].$$

Then for T invertible, we have

$$\Phi[TGT^*; U] = -\Phi_0 - 2 \inf_{\rho \in \mathcal{G}^{-1}(TGT^*)} \left[\int \left(\log\left[(\det G)^{1/2} \cdot |\det T| \cdot \rho \right] U \right) \rho\, dx \right].$$

Now observe by changing variables that

$$\{\rho : \rho \in \mathcal{G}^{-1}(TGT^*)\} = \{|\det T|^{-1} \cdot \rho \circ T^{-1} : \rho \in \mathcal{G}^{-1}(G)\}.$$

Therefore,

$$\begin{aligned}\Phi[TGT^*; U] &= -\Phi_0 - 2 \inf_{\rho \in \mathcal{G}^{-1}(G)} \left[|\det T|^{-1} \int \left(\log\left[(\det G)^{1/2} \cdot \rho \circ T^{-1} \right] + U \right) \rho \circ T^{-1}\, dx \right] \\ &= -\Phi_0 - 2 \inf_{\rho \in \mathcal{G}^{-1}(G)} \left[\int \left(\log\left[(\det G)^{1/2} \cdot \rho \right] + U \circ T \right) \rho\, dx \right] \\ &= \Phi[G; U \circ T],\end{aligned}$$

as was to be shown.

□

Sketch of the Proof of the Continuous Extension of the LW Functional in *Theorem 2*

Suppose $G \in \mathcal{S}^n_+$ is of the form

$$G = \begin{pmatrix} G_p & 0 \\ 0 & 0 \end{pmatrix},$$

where $G_p \in \mathcal{S}^p_{++}$, and suppose that $G^{(j)} \in \mathcal{S}^n_{++}$ with $G^{(j)} \to G$ as $j \to \infty$. For each j, diagonalize $G^{(j)} = \sum_{i=1}^n \lambda_i^{(j)} v_i^{(j)} \left(v_i^{(j)}\right)^T$, where the $v_i^{(j)}$ are orthonormal, $\lambda_i^{(j)} > 0$ for $i = 1, \ldots, n$.
We want to show that

$$\Phi_n[G^{(j)}, U] \to \Phi_p[G_p, U(\cdot, 0)].$$

It suffices to show that every subsequence has a convergent subsequence with its limit being $\Phi_p[G_p, U(\cdot, 0)]$. The $G^{(j)}$ are convergent and hence bounded (in the $\|\cdot\|_2$ norm), so the $\lambda_i^{(j)}$ are bounded. Moreover, the $v_i^{(j)}$ are all of unit length and hence bounded, so by passing to a subsequence if necessary we can assume that, for each i, there exist λ_i, v_i such that $\lambda_i^{(j)} \to \lambda_i$ and $v_i^{(j)} \to v_i$ as $j \to \infty$. It follows that the v_i are orthonormal and that G can be diagonalized as $G = \sum_{i=1}^n \lambda_i v_i v_i^T$. Since G_p is positive definite, we must have $\lambda_i > 0$ for $i = 1, \ldots, p$, and moreover $\lambda_i = 0$ for $i = p+1, \ldots, n$. Evidently, the eigenvectors of G with strictly positive eigenvalues must be precisely the eigenvectors of G_p, concatenated with $n - p$ zero entries; i.e., for $i = 1, \ldots, p$, v_i must be of the form $(*, 0)$. By orthogonality, for $i = p+1, \ldots, n$, v_i must be of the form $(0, *)$.

Via the transformation rule (*Proposition 4*), by choosing a suitable sequence of orthogonal transformations $T^{(j)} \to I_n$, we can redefine $G^{(j)} := \sum_{i=1}^{n} \lambda_i^{(j)} v_i v_i^T$ and instead show that

$$\Phi_n[G^{(j)}, U^{(j)}] \to \Phi_p[G_p, U(\,\cdot\,, 0)],$$

where $U^{(j)} := U \circ T^{(j)}$. Observe that $G^{(j)}$ can be written in the form

$$G^{(j)} = \begin{pmatrix} G_1^{(j)} & 0 \\ 0 & G_2^{(j)} \end{pmatrix},$$

where $G_1^{(j)} \to G_p$ and $G_2^{(j)} \to 0$ as $j \to \infty$.

Before proceeding, we establish some notational conventions. We will use coordinates $x = (y, z)$ to denote the splitting $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$, and for any density ρ we define marginals $\rho_1(y) = \int \rho(y, z)\, dz$, $\rho_2(z) = \int \rho(y, z)\, dy$.

Now recall that

$$\frac{1}{2}\Phi_n[G^{(j)}, U^{(j)}] = \sup_{\rho \in \mathcal{G}_n^{-1}(G^{(j)})} \left[H(\rho) - \int U^{(j)}\, \rho \,\mathrm{d}x \right] - \frac{1}{2}\log\left((2\pi e)^n \det G^{(j)} \right).$$

Intuitively, as $j \to \infty$, any density $\rho \in \mathcal{G}_n^{-1}(G^{(j)})$ concentrates more and more about the subspace $\mathrm{span}\{e_1, \ldots, e_p\}$. Also, $U^{(j)} \to U$ in a pointwise sense. Therefore, heuristically, we expect to be able to replace $\int U^{(j)}\, \rho\, dx$ with $\int U(\,\cdot\,, 0)\, \rho_1\, dy$ in the limit $j \to \infty$.

Moreover, for any ρ, we have that $H(\rho) \le H(\rho_1) + H(\rho_2)$, with equality if and only if $\rho(y, z) = \rho_1(y)\rho_2(z)$. (For given marginals, the entropy of a distribution is maximized if it is a product of its marginals.) Since $\mathcal{G}_{n-p}(\rho_2) = G_2^{(j)}$, and the entropy among measures of a given covariance is maximized by the Gaussian measure of that covariance, it follows that $H(\rho_2) \le \frac{1}{2}\log\left((2\pi e)^{n-p} \det G_2^{(j)} \right)$, with equality if and only if ρ_2 is the mean-zero Gaussian density with covariance $G_2^{(j)}$.

Now for j large, we have $\mathcal{G}_p(\rho_1) \approx G_p$ for any $\rho \in \mathcal{G}_n^{-1}(G^{(j)})$. Due to the fact that $\det G^{(j)} = \det G_1^{(j)} \det G_2^{(j)}$, we then expect

$$\limsup_{j\to\infty} \Phi_n[G^{(j)}, U^{(j)}] \le \sup_{\rho_1 \in \mathcal{G}_p^{-1}(G_p)} \left[H(\rho_1) - \int U(\cdot, 0)\, \rho_1 \,\mathrm{d}y \right] - \frac{1}{2}\log\left((2\pi e)^p \det G_p \right);$$

i.e.,

$$\limsup_{j\to\infty} \Phi_n[G^{(j)}, U^{(j)}] \le \Phi_p[G_p, U(\,\cdot\,, 0)].$$

For an idea of the proof of the opposite bound, note that each of our inequalities in the preceding argument can be made to hold with equality by specifically choosing $\rho \in \mathcal{G}_n^{-1}(G^{(j)})$ to be a product density with arbitrary first marginal ρ_1 and second marginal ρ_2 given by the mean-zero Gaussian density with covariance $G_2^{(j)}$.

□

Proof of *Lemma 5*

Write

$$\begin{aligned} \frac{1}{2}x^T A x + U(x) &= \frac{1}{2}x^T \left(A - \Sigma(\varepsilon) + \overline{\Sigma}^{(N)}(\varepsilon) \right) x + U_\varepsilon^{(N)}(x) \\ &= \frac{1}{2}x^T \left(G^{-1} + \overline{\Sigma}^{(N)}(\varepsilon) \right) x + U_\varepsilon^{(N)}(x). \end{aligned}$$

It follows that under the interaction $U_\varepsilon^{(N)}$, the noninteracting Green's function $G^{-1} + \overline{\Sigma}^{(N)}(\varepsilon)$ corresponds to the interacting Green's function G. This establishes that

$$A[G; U_\varepsilon^{(N)}] = G^{-1} + \overline{\Sigma}^{(N)}(\varepsilon).$$

Moreover, by the Dyson equation we have that

$$\Sigma[G; U_\varepsilon^{(N)}] = A[G; U_\varepsilon^{(N)}] - G^{-1} = \overline{\Sigma}^{(N)}(\varepsilon),$$

as desired.

□

Proof of the Resummation Step in *Theorem 3*

Note that the sum up to the finite order N of the Feynman diagrams for the Green's function [with bare propagator $G_0^{(N)}(\varepsilon)$ and interaction εU] coincides with $\widetilde{G}^{(N)}(\varepsilon)$ and hence also with G, up to negligible error. Then we may use the standard combinatorial argument (1, 2) that the bold diagram expansion for the self-energy up to order N accounts for all bare diagrams up to order N. It follows that $\widetilde{\Sigma}^{(N)}(\varepsilon)$—and hence also $\overline{\Sigma}^{(N)}(\varepsilon)$—is, up to negligible error, given by the bold diagram expansion up to order N with bold propagator G and interaction εU. But since this expansion and $\overline{\Sigma}^{(N)}(\varepsilon)$ are both polynomials of order N in ε, it follows that $\overline{\Sigma}^{(N)}(\varepsilon)$ is exactly given by the bold diagram expansion up to order N, as was to be shown.

□

Proof of the Expansion Coefficients of the LW Functional in *Theorem 3*

By the transformation rule (and the fact that U is quartic), we have

$$\Phi[tG, \varepsilon U] = \Phi[G, t^2 \varepsilon U]$$

for any $G \in \mathcal{S}_{++}^n$ and $\varepsilon, t > 0$. Taking the gradient in G of both sides yields

$$\Sigma[tG, \varepsilon U] = \frac{1}{t}\Sigma[G, t^2\varepsilon U].$$

Then compute

$$\begin{aligned} \Phi[G, \varepsilon U] &= \int_0^1 \frac{d}{dt}\Phi[tG, \varepsilon U]\, dt \\ &= \int_0^1 \mathrm{Tr}[G\Sigma[tG, \varepsilon U]]\, dt \\ &= \int_0^1 \frac{1}{t}\mathrm{Tr}[G\Sigma[G, t^2\varepsilon U]]\, dt \\ &= \int_0^1 \frac{1}{t}\left[\sum_{k=1}^{N} \mathrm{Tr}\left[G\Sigma_G^{(k)} \right] t^{2k}\varepsilon^k + O\left(t^{2(N+1)}\varepsilon^{N+1} \right) \right] dt \\ &= \int_0^1 \left[\sum_{k=1}^{N} \mathrm{Tr}\left[G\Sigma_G^{(k)} \right] t^{2k-1}\varepsilon^k + O\left(t^{2N+1}\varepsilon^{N+1} \right) \right] dt. \end{aligned}$$

Now since t ranges from 0 to 1 in the integrand, we have that $t^{2N+1}\varepsilon^{N+1} \le \varepsilon^{N+1}$, and therefore

$$\begin{aligned} \Phi[G; \varepsilon U] &= \int_0^1 \left[\sum_{k=1}^{N} \mathrm{Tr}\left[G\Sigma_G^{(k)} \right] t^{2k-1}\varepsilon^k \right] dt + O(\varepsilon^{N+1}) \\ &= \sum_{k=1}^{N} \frac{1}{2k}\mathrm{Tr}\left[G\Sigma_G^{(k)} \right] \varepsilon^k + O(\varepsilon^{N+1}). \end{aligned}$$

This establishes that

$$\Phi_G^{(k)} = \frac{1}{2k} \mathrm{Tr}\left[G\Sigma_G^{(k)}\right],$$

as was to be shown.

We remark that this proof bears resemblance to the adiabatic integration technique that formally defines the LW functional (2, 3). The significant difference lies in the fact that in our case, the adiabatic integration technique is used to establish the relation of the coefficients for the LW functional that we have already defined. Moreover, unlike the diagrammatic constructions of the LW functional, our argument does not make use of any particular properties of the bold diagrams, relying instead on the transformation rule.

□

1. Fetter AL, Walecka JD (2003) *Quantum Theory of Many-Particle Systems* (Courier, Mineola, NY).
2. Luttinger JM, Ward JC (1960) Ground-state energy of a many-fermion system. II. *Phys Rev* 118:1417–1427.
3. Martin RM, Reining L, Ceperley DM (2016) *Interacting Electrons* (Cambridge Univ Press, Cambridge, UK).