Full support and positive density in the zero-temperature SK model

Will Cook · An account of the quantitative proof of OpenAI (2026a) · 3 November 2026

AI assistance. Will Cook directed the work. AI agents based on large language models produced and revised this manuscript.

Abstract. For the pure, zero-field Sherrington–Kirkpatrick model at zero temperature, the Stieltjes measure of every integrable minimiser of the Parisi functional has full support on $[0,1)$ and a smooth density that is strictly positive on $(0,1)$. We reconstruct OpenAI's quantitative proof, which uses one spatial-moment inequality twice. On a positive constant interval, the inequality makes the derivative of the gradient second moment strictly convex, contradicting the endpoint conditions imposed by minimality, and together with the bounded gradient it also excludes terminal gaps. Once the support is full, moment identities give smoothness, and a local constant replacement transfers the inequality to the minimiser and forces its density to be positive. H.-B. Chen previously proved full support and smooth nonnegative density. We isolate the fixed-time spatial moment and the positive margin that survives the local replacement, and we import its shape comparisons and full polynomial identity. Along the way we derive the scalar value identity and explain its classical optimisation consequence. For the quantum approximate optimization algorithm (QAOA), we identify the rooted energy of a finite Gaussian observable and the separate circuit theorem that approximates it, and a further second-moment theorem then gives a measured near-optimal configuration on the same random instance, with the circuit word fixed before the number of spins grows. The construction gives no bound on the depth needed for a given accuracy.

1. The model and the result

For independent standard Gaussian couplings J_{ij}, the pure, zero-field Sherrington–Kirkpatrick (SK) Hamiltonian is

$$H_n(\sigma) = n^{-1/2} \sum_{i<j} J_{ij}\sigma_i\sigma_j, \qquad \sigma \in \{-1,1\}^n.$$

With the maximisation convention, its limiting ground-state energy per spin is

$$P_* := \lim_{n\to\infty} n^{-1}\mathbb{E}\max_{\sigma} H_n(\sigma).$$

The Parisi formula expresses P_* as a minimum over nondecreasing functions. We study the minimiser itself. Its increments charge every interval of overlap scales below one, and they have a strictly positive density at each interior scale. One quantitative bound gives both facts: it first excludes constant intervals and then, after an approximation, gives pointwise positivity.

The overlap of two configurations is $R(\sigma,\tau) = n^{-1}\sum_i \sigma_i\tau_i$, and it determines the energy covariance, $\mathbb{E}[H_n(\sigma)H_n(\tau)] = nR^2/2 - 1/2$. Hence the Gaussian law of the energies is encoded by the overlaps. We use the covariance convention $\xi(t) = t^2/2$, so $\xi'' = 1$. Adding one independent centred Gaussian of variance $1/2$ to every configuration's energy gives covariance $n\xi(R)$ and leaves the expected maximum unchanged, which is the normalisation of the ground-state Parisi formula of Auffinger and Chen (2017), Theorem 1. The model has no higher-order interactions and no external field.

Let $\mathcal{U}$ consist of the nonnegative, nondecreasing, right-continuous functions $\gamma : [0,1) \to [0,\infty)$ satisfying $\int_0^1 \gamma(t)\,dt < \infty$, and define the locally finite Stieltjes measure by

$$\mu_\gamma([0,t]) = \gamma(t), \qquad 0 \le t < 1.$$

Thus $\mu_\gamma((s,t]) = \gamma(t) - \gamma(s)$, so the measure charges each interval with the increment of γ across it. It may have a point mass (an atom) of size $\gamma(0)$ at zero, and its total mass need not be finite or equal to one. Support statements refer to the relative topology of $[0,1)$. A support gap is a relatively open interval of positive length and zero mass, and the cumulative coefficient is constant on it.

The Parisi functional $\mathcal{P}(\gamma)$ is defined through an auxiliary value function $\Phi_\gamma(t,x)$, in which $x \in \mathbb{R}$ is a spatial variable and the overlap parameter t plays the role of time. Write $\Phi = \Phi_\gamma$ and use subscripts for partial derivatives. The backward equation, terminal condition and functional are

$$\Phi_t = -\tfrac{1}{2}(\Phi_{xx} + \gamma(t)\Phi_x^2), \qquad \Phi(1,x) = |x|, \qquad \mathcal{P}(\gamma) = \Phi_\gamma(0,0) - \tfrac{1}{2}\int_0^1 t\gamma(t)\,dt.$$

The terminal value $|x|$ is the zero-temperature value of one spin in a field x. For an integrable coefficient, the solution of the PDE is defined by approximation from bounded monotone coefficients. OpenAI (2026a), §1 and Lemmas 2.1–2.2 give the exact control definition and the stability theorem.

To turn minimality of $\mathcal{P}$ into a condition on γ, we use a random process whose gradient moment measures the first-order cost of changing γ. Write $u = \Phi_x$ for the spatial gradient and $a = \Phi_{xx}$ for its spatial derivative. Let B be standard Brownian motion, and define the diffusion X and its self-consistency function q by

$$dX_t = \gamma(t)u(t,X_t)\,dt + dB_t, \qquad X_0 = 0, \qquad q(t) = \mathbb{E}[u(t,X_t)^2].$$

The discrepancy $q(t) - t$ enters the first variation of $\mathcal{P}$, and minimality will force $q(t) = t$ on the support of μ_γ. The SDE has a unique strong solution. Unless an endpoint limit is justified explicitly, all stochastic calculations below take place on compact time intervals ending before one.

The support of a minimiser is infinite by Auffinger, Chen and Zeng (2020), Theorem 1. Infinite support still allows gaps: the measure $\sum_{k\geq 1} 2^{-k}\delta_{1-2^{-k}}$ has infinitely many atoms and many gaps. Full support and a smooth nonnegative density were then proved in H.-B. Chen (2026), Theorem 1.3. We follow OpenAI's quantitative proof of the support conclusion and of its further strict-positivity result. The feature used twice is a lower bound that holds throughout every positive constant interval of an admissible coefficient, whether or not the coefficient minimises.

Theorem 1 (full support; quantitative proof following OpenAI (2026a), Theorem 1.1). If $\gamma \in \mathcal{U}$ minimises $\mathcal{P}$, then

$$\operatorname{supp}\mu_\gamma = [0,1), \qquad \gamma(t_2) > \gamma(t_1) \quad (0 \leq t_1 < t_2 < 1).$$

For the solution of the displayed SDE and every $0 \leq t < 1$,

$$q(t) = t, \qquad \mathbb{E}[a(t,X_t)^2] = 1.$$

Corollary 4 in §4 adds a smooth density, strictly positive at every interior point.

The existence of a minimiser is a separate input, from Auffinger and Chen (2017), Theorem 1 and the proof of Lemma 3. Integrability allows γ to grow near one. Zero field gives $u(0,0) = 0$, and the curvature calculation uses $\xi'' = 1$, so the theorem is stated only for the pure SK covariance at zero field. Nonzero field and general mixed-spin covariances lie outside its hypotheses.

Full support has a concrete meaning for the energy landscape. For each fixed $\rho \in [0,1)$ and $\delta, \eta > 0$, with probability tending to one there are configurations σ, τ with $|R(\sigma,\tau) - \rho| < \delta$ whose energies

are both at least $n(P_* - \eta)$. This follows from Auffinger, Chen and Zeng (2020), Proposition 1, equation (5) in the Gaussian normalisation above.

We isolate the spatial moment that the support and density arguments share, explain why it needs an integrated estimate, and identify what survives the local-replacement limit. The conclusions and the quantitative method are those of the cited sources.

The proof has two stages. A gap with two support endpoints faces two demands. Minimality forces q to agree with time at both endpoints, with endpoint slopes at most one. On a positive constant interval, a quantitative estimate makes q' strictly convex, so its integral over the gap is smaller than the gap's length, which contradicts the required change in q. Once every gap is excluded, the identity $q' = 1$ supplies the equation from which the same estimate forces $\gamma' > 0$. For a bounded gap, a qualitative crossing criterion already suffices. The density argument needs more: an estimate for a local constant replacement, which need not minimise, at a time where q'' need not vanish, with a positive lower bound that survives the limit.

Section 2 turns minimality into contact conditions and proves Theorem 1 from the curvature bound of Lemma 2. In §3 we derive the moment polynomial behind that bound, together with the fixed-time estimate of Proposition 3, which §4 applies to prove Corollary 4. Section 5 follows the support identities to the value formula, the classical algorithm and a finite Gaussian observable whose rooted energy a QAOA circuit approximates, and then identifies the additional concentration input for a measured configuration.

These optimisation consequences use full support only. Strict positivity of the density is a further analytic conclusion, and readers interested in optimisation may accept Theorem 1 and go directly to §5. Sections 2–4 assume Itô calculus and martingale convergence. The circuit construction is imported as a separate theorem. Omitted proofs are linked to *From Parisi full support to deterministic QAOA circuits: A long reasoning record*, the companion edition of 3 November 2026, and citations of the original sources stay beside the imported inputs.

2. Contact conditions and the exclusion of gaps

For an admissible coefficient, Φ is even, convex and one-Lipschitz in space. Its spatial derivatives of positive order are bounded and continuous on each strip $[0,T] \times \mathbb{R}$, $T < 1$, and the PDE holds for almost every time. The compact-strip approximation theorem of OpenAI (2026a), Lemmas 2.1–2.2 justifies the Itô calculations below even when γ has jumps. This control method has antecedents in Jagannath and Tobasco (2016) at positive temperature and in Auffinger and Chen (2017), Theorem 2 and Corollary 2 at zero temperature (see also Chen, Handschy and Lerman (2018), Proposition 2 and Theorem 5).

We fix a minimiser. Itô's formula cancels the drift of $u(t, X_t)$:

$$du(t, X_t) = a(t, X_t)\, dB_t, \qquad q(0) = 0, \qquad q'(t) = \mathbb{E}[a(t, X_t)^2].$$

In particular, q' is continuous across jumps of the coefficient. Minimality means that every allowed one-sided change has nonnegative first-order cost, and the one-sided variation formula computes that cost. When γ is replaced by $\gamma + \varepsilon\eta$ and $\varepsilon \downarrow 0$, the functional changes at the rate

$$D_+\mathcal{P}(\gamma)[\eta] = \tfrac{1}{2} \int_0^1 \eta(t)(q(t) - t)\, dt, \tag{1}$$

for every integrable direction η for which $\gamma + \varepsilon\eta \in \mathcal{U}$ for all sufficiently small $\varepsilon \geq 0$ (see OpenAI (2026a), Lemma 6.1). Earlier zero-temperature variation and support identities appear in Chen, Handschy and Lerman (2018), Propositions 3–4.

Adding a unit atom at s raises the cumulative function at every $t \geq s$, so its direction is $\eta = \mathbf{1}_{[s,1)}$. Its first-order cost is half of a tail integral, and minimality makes the integral nonnegative:

$$G(s) := \int_s^1 (q(t) - t)\, dt \geq 0.$$

Thus $G(s)/2$ is the first-order cost of the added atom.

Mass can also be removed. Let $\nu = \mu_\gamma|_K$ for a compact $K \subset [0,1)$. For $0 \leq \varepsilon \leq 1$, the measure $\mu_\gamma - \varepsilon\nu = (1 - \varepsilon \mathbf{1}_K)\mu_\gamma$ is nonnegative, and its cumulative function is $\gamma - \varepsilon\nu([0,\cdot]) \in \mathcal{U}$. Minimality gives

$$0 \leq D_+\mathcal{P}(\gamma)[-\nu([0,\cdot])] = -\tfrac{1}{2}\int_K G(s)\, d\mu_\gamma(s).$$

The bounds $|q - t| \leq 1$ and $\mu_\gamma(K) < \infty$ justify the change in the order of integration. Since $G \geq 0$, it vanishes almost everywhere for μ_γ, and by continuity it vanishes on $S = \operatorname{supp} \mu_\gamma$. The variation removes a fraction of all the mass in K, atoms and diffuse mass alike, and assumes no density or probability constraint.

At a support point $s > 0$, the function G has a minimum. Hence

$$G'(s) = s - q(s) = 0, \qquad G''(s) = 1 - q'(s) \geq 0, \quad \text{so} \quad q(s) = s, \quad q'(s) \leq 1. \tag{2}$$

If $0 \in S$, contact gives $G(0) = 0$, and zero field gives $q(0) = 0$ independently. Consequently $G(h) = \frac{1}{2}(1 - q'(0))h^2 + o(h^2)$, and nonnegativity gives the same slope bound at zero. These are the contact conditions of OpenAI (2026a), Lemma 6.2.

Lemma 2 (constant-interval curvature; OpenAI (2026a), Proposition 5.1). Let $\gamma \in \mathcal{U}$, with its associated diffusion and moment q. If $\gamma = c$ on an open interval $I \subset (0,1)$, called a plateau, then for $t \in I$

$$q'''(t) \geq \frac{c^2}{1000}\mathbb{E}[a(t, X_t)^4] > 0 \quad \text{if } c > 0. \tag{3}$$

If $c = 0$, then $q''(t) = \mathbb{E}[\Phi_{xxx}(t, X_t)^2] > 0$.

The lemma holds for every admissible coefficient, minimising or not. Section 3 derives its moment polynomial and explains the integrated estimate imported from the source.

Proof of Theorem 1, assuming Lemma 2. We treat the intervals of the complement of S by type. An interval with two support endpoints admits two contact conditions. Initial intervals without support at zero have coefficient zero. For a terminal gap, the bound $q \leq 1$ replaces the missing right contact condition.

Suppose first that (ℓ, r) is a support gap with endpoints in S and $r < 1$. The coefficient is a positive constant there, because support at ℓ and the absence of mass immediately to its right force $\mu_\gamma([0,\ell]) > 0$. Strict convexity of q' and (2) give

$$q'(t) < \frac{r - t}{r - \ell}q'(\ell) + \frac{t - \ell}{r - \ell}q'(r) \leq 1 \qquad (\ell < t < r).$$

Integrating contradicts $q(r) - q(\ell) = r - \ell$, so the geometric obstruction is an area deficit under q'.

An initial gap $[0, r)$ with $0 \notin S$ and $r \in S$ carries no mass, hence $\gamma = 0$ there. Its right endpoint gives $q(r) = r$ and $q'(r) \le 1$, and zero field gives $q(0) = 0$. The zero-plateau identity gives $q'' > 0$, hence $q'(t) < q'(r) \le 1$, and the same integral contradiction follows without assuming $G(0) = 0$. A gap $(0, r)$ with $0 \in S$ is already covered by the two-contact case.

A terminal gap $(r, 1)$ with nonempty support has $r \in S$ and a positive plateau height. Suppose that $q(t) \ge t$ for some $r < t < 1$. Strict convexity of q' and $q'(r) \le 1$ give

$$t - r \le \int_r^t q'(s)\, ds < \frac{t-r}{2}(q'(r) + q'(t)),$$

so $q'(t) > 1$. Convexity then gives $q'(s) > q'(t) > 1$ for $s > t$, and consequently $q(s) > t + (s - t)q'(t) > 1$ for $s < 1$ sufficiently close to one, contrary to $q \le 1$. Thus $q(t) < t$ on $(r, 1)$, and $G(r) < 0$ contradicts minimality. This case uses contact only at the left endpoint.

If S is empty, then $\gamma = 0$. Lemma 2 makes q strictly convex, and $q(0) = 0$ and $q \le 1$ give $q(t) < (t/b)q(b) \le t/b$ for $0 < t < b < 1$. Letting $b \uparrow 1$ gives $q(t) \le t$. Strict convexity applied once more with $t < b < 1$ now gives $q(t) < (t/b)q(b) \le t$, so $G(0) < 0$.

The terminal exclusion was already noted in Auffinger, Chen and Zeng (2020), Remark 7. Our argument uses only the curvature needed for the other gaps, and the full endpoint discussion is in long §5, "A shorter qualitative exclusion". A stronger boundary estimate $1 - q(t) \ge C\sqrt{1-t}$, for arbitrary coefficients constant near one, is proved in OpenAI (2026a), Lemma 6.3 and in long §5, terminal boundary estimate, and Chen (2026), Proposition 4.5, equations (4.27)–(4.38) gives the square-root asymptotic under the terminal-gap hypothesis. Antecedents of the zero-coefficient heat argument and the late-step variation appear in Chen, Handschy and Lerman (2018), Theorem 6.

These cases exhaust the complement of the relatively closed set S. Consequently $S = [0, 1)$, contact gives $q(t) = t$, and differentiation gives $\mathbb{E}a^2 = 1$, including the right-hand value at zero. Moreover, $\gamma(t_2) - \gamma(t_1) = \mu_\gamma((t_1, t_2]) > 0$.

This proves Theorem 1 from the stated analytic inputs.

3. The integrated estimate

Both arguments use one spatial moment. Recall $a = \Phi_{xx}$, write $b = \Phi_{xxx}$ and $d = \Phi_{xxxx}$, and for $0 < t < 1$ set

$$\mathcal{C}_\gamma(t) = \mathbb{E}\big[d^2 - 12\gamma(t)ab^2 + 6\gamma(t)^2a^4\big],$$

with the spatial factors evaluated at (t, X_t). The definition requires spatial regularity and no time derivative of γ, so the estimate we need can be stated at a single time.

Proposition 3 (fixed-time spatial-moment estimate). Let $\gamma \in \mathcal{U}$, let $0 < t < 1$ be a continuity point of γ, and suppose that $c = \gamma(t) > 0$. For its associated diffusion,

$$\mathcal{C}_\gamma(t) \ge \frac{c^2}{1000}\mathbb{E}[a(t, X_t)^4].$$

Neither minimality nor a constant neighbourhood of t is required. This is the consequence of OpenAI (2026a), Proposition 5.1 and Lemmas 2.1–2.2 stated in the companion's Proposition 4.1, §4.5. We first explain the estimate on a positive plateau and then obtain the fixed-time form by inserting a short plateau.

For a smooth coefficient, the two applications are connected by the identity

$$q''' + 2\gamma' \mathbb{E} a^3 = \mathcal{C}_\gamma.$$

On a plateau $\gamma' = 0$, while at a smooth minimiser full support gives $q = t$ and hence $q''' = 0$. Section 4 proves the required smoothness before this second use.

On a plateau with $\gamma = c > 0$, put $r = c\Phi_x = cu$, $v = r_x = ca$, $w = r_{xx}$ and $z = r_{xxx}$. This rescaling makes the derivative identities polynomial. The PDE becomes $r_t = -\frac{1}{2} r_{xx} - r r_x$, and the diffusion has spatial generator $\frac{1}{2}\partial_{xx} + r\partial_x$. Differentiating in space and applying Itô's formula gives the drift and diffusion pairs $(0, v)$, $(-v^2, w)$ and $(-3vw, z)$ for r, v, w, respectively. All expectations in this section evaluate spatial functions at (t, X_t). Starting from $c^2 q' = \mathbb{E} v^2$, Itô's formula gives

$$\begin{aligned} c^2 q'' &= \mathbb{E}[w^2 - 2v^3], \\ c^2 q''' &= \mathbb{E}[z^2 - 12vw^2 + 6v^4] = c^2 \mathcal{C}_\gamma(t). \end{aligned} \tag{4}$$

For example, the drifts of w^2 and v^3 are $z^2 - 6vw^2$ and $3vw^2 - 3v^4$, respectively.

The negative mixed term is the obstacle, and convexity, which gives $v \geq 0$, does not control it. Take, for instance, the smooth auxiliary terminal value $\log(2\cosh x)$ and the coefficient $\gamma \equiv 1$. The solution is $\Phi(t, x) = \log(2\cosh x) + (1 - t)/2$, so $r = \tanh x$, and at $r^2 = 1/2$ the polynomial in (4) equals $-19/8$. This example is a smoothed finite-step problem, and its coefficient is no minimiser of the absolute-value terminal problem. The calculation is in long §4.1, and that section also finds a negative pointwise value with the actual $|x|$ terminal datum for a small positive constant coefficient (long §4.1, pointwise obstruction).

The proof therefore subtracts a weighted divergence whose expectation under the law of X_t is zero, and what remains has a positive fourth-moment remainder. For this decomposition to have the required signs, the backward value function and the forward density must satisfy joint shape restrictions. Those comparisons and the exact polynomial identity are the substantial inputs from OpenAI (2026a), §§3–4. We explain their roles first for finite-step coefficients and then pass the integrated inequality to an integrable coefficient by approximation.

We begin with a positive nondecreasing finite-step coefficient γ_{step}, that is, a function constant on each interval of a finite partition, and we replace the terminal value $|x|$ by the smooth approximation $M^{-1}\log(2\cosh Mx)$, with $M \geq \|\gamma_{\text{step}}\|_\infty$. On a step of height c, set $h = e^{c\Phi}$. Then $h_t = -h_{xx}/2$ and $r = h_x/h$, and the same exponential removes the drift from the forward density equation. Indeed, writing $f = hg$ gives

$$\begin{aligned} f_t &= \tfrac{1}{2}(hg)_{xx} - (h_x g)_x \\ &= -\tfrac{1}{2} h_{xx} g + \tfrac{1}{2} h g_{xx} = h_t g + \tfrac{1}{2} h g_{xx}. \end{aligned}$$

Comparison with $f_t = h_t g + h g_t$ yields $g_t = g_{xx}/2$. One backward heat factor and one forward heat factor thus determine the law against which the moment polynomial is integrated.

For $s = -\partial_x \log g$, the source proves joint shape restrictions on $x > 0$: $r_x > 0$ and $s_x \geq 1/t$, both r and s are concave, and $H := w/(2r) + v \leq 0$ is nondecreasing. They also give

$$j := s_x - (s/r)v = r\,\partial_x(s/r) \geq 0.$$

This last inequality couples the backward gradient to the forward density, and separate monotonicity and concavity do not supply it. On the first step $s = x/t$, so concavity of r gives $j \geq 0$, and a

parabolic comparison propagates the sign within each step. At a forward jump with $k = c_+/c_- \geq 1$, continuity of Φ and f gives

$$r_+ = kr_-, \qquad s_+ = s_- + (k-1)r_-, \qquad j_+ = (s_x)_+ - \frac{s_+}{r_+}(r_x)_+ = j_-.$$

The joint sign therefore survives jumps of the coefficient as well. The quotients w/r and s/r extend smoothly through zero because r, s, w are odd and $r_x(0) > 0$. Both the boundary controls of the comparison and these extensions are analytic inputs from OpenAI (2026a), Lemmas 3.2–3.4, reconstructed in long §4.3. This ratio sign and its preservation at jumps also follow from Zhou (2026), Appendix E.3, (E.37) and (E.41), after rescaling the smoothed finite-step problem. Section 4.3 of the long record gives the change of variables and the resulting inequality.

The factorisation of the density gives $f_x/f = r - s$. For a suitable expression F in the shape quantities, define

$$\mathcal{W}(F) = r\partial_x F + (v + r^2 - rs)F.$$

Then $f\mathcal{W}(F) = \partial_x(rfF)$. The smooth extensions avoid an interior boundary term at zero, and the source's differentiated Gaussian tail estimates give $rfF \to 0$ at both spatial infinities for the polynomial correction used below. Hence $\mathbb{E}_f \mathcal{W}(F) = 0$, where $\mathbb{E}_f$ denotes integration against $f(x)\,dx$.

What remains is algebra under these analytic restrictions. The correction polynomial is chosen in two steps: first arrange a square and other nonnegative terms containing the highest derivative z, then cancel their terms linear in z with the weighted divergence. Only lower derivatives and the shape quantities then remain in the residual. OpenAI (2026a), Lemma 4.1 supplies an explicit polynomial F_2 and a decomposition

$$z^2 - 12vw^2 + 6v^4 = \mathcal{W}(F_2) + \kappa v^4 + \mathcal{R}, \qquad \mathcal{R} \geq 0, \qquad \kappa = \frac{1661507}{1521000000} > \frac{1}{1000}.$$

The coefficients of F_2 have both signs. Since the correction integrates to zero, the signs of F_2 and $\mathcal{W}(F_2)$ play no role, whereas the shape inequalities make every term of $\mathcal{R}$ nonnegative. Two of its actual terms are

$$\frac{337}{10000}\, r^2 j\,(-w/r - 7v)^2, \qquad (2rH_x)v\left(\frac{2331}{10000}r^2 + \frac{87}{2500}j\right).$$

The first needs the joint comparison $j \geq 0$, and the second also uses $H_x \geq 0$ on the positive half-line. Reflection preserves these signs on the negative half-line. The remaining terms use $H \leq 0$, forward concavity and the other shape bounds, so the positivity of the decomposition depends on the shapes the PDE supplies. Integration removes $\mathcal{W}(F_2)$ and keeps the fourth-moment margin, giving

$$\int_{\mathbb{R}} (z^2 - 12vw^2 + 6v^4) f\,dx \geq \frac{1}{1000}\int_{\mathbb{R}} v^4 f\,dx. \tag{5}$$

We use (5) as an imported analytic estimate. The full polynomial and its decomposition are in Lemma 4.1 and Table 1, and Proposition 4.2 integrates them using the signs and zero boundary flux supplied by Lemma 3.4. The definitions and the complete identity are reproduced in long §4.4, "The polynomial correction with zero expectation".

To obtain Lemma 2 for an integrable coefficient, choose positive monotone step coefficients γ_n that keep the plateau near a fixed interior time, with $\|\gamma_n - \gamma\|_1 \to 0$, and terminal parameters

$M_n \geq \|\gamma_n\|_\infty$ with $M_n \to \infty$. Spatial derivatives through order four then converge uniformly on a fixed strip below one, and the diffusions driven by the same Brownian motion converge (see OpenAI (2026a), Lemmas 2.1–2.2 and Proposition 5.1). Only v, w, z remain in (5), so the auxiliary logarithmic derivative s needs no convergence theorem. Dividing by c^2 gives (3), and strictness follows from $\Phi_{xx} > 0$ on the plateau (see OpenAI (2026a), Lemma 2.3). Long §4.5 specifies the plateau-preserving approximants and their common strip bounds.

The zero plateau has the separate heat identity stated in Lemma 2. Its strictness uses a positive diffusion density and the fact that Φ_{xxx} cannot vanish identically, since an even quadratic profile would contradict the one-Lipschitz bound and $|x| + O(1)$ growth.

Passage to a fixed time. Now let $\gamma \in \mathcal{U}$ be continuous at $t \in (0,1)$, with $c = \gamma(t) > 0$. For $0 < \delta < \min(t, 1-t)$ define

$$\gamma_\delta(s) = \begin{cases} c, & t-\delta \leq s < t+\delta, \\ \gamma(s), & \text{otherwise.} \end{cases}$$

The inequalities $\gamma(t-\delta) \leq c \leq \gamma(t+\delta)$ make both joins nondecreasing, so $\gamma_\delta \in \mathcal{U}$, although it need not minimise $\mathcal{P}$. Moreover,

$$\|\gamma_\delta - \gamma\|_1 \leq 2\delta\,[\gamma(t+\delta) - \gamma(t-\delta)] \longrightarrow 0.$$

We use a subscript δ for its value-function derivatives and diffusion. Lemma 2 and (4), already extended to integrable coefficients, give at the fixed time t

$$\mathcal{C}_{\gamma_\delta}(t) = \mathbb{E}[d_\delta^2 - 12c\,a_\delta b_\delta^2 + 6c^2 a_\delta^4] \geq \frac{c^2}{1000}\mathbb{E}[a_\delta^4]. \tag{6}$$

We now pass to the limit of shrinking plateaus and choose $t < T < T_1 < 1$. Monotonicity gives the common coefficient bound

$$\sup_{s \leq T_1} \gamma_\delta(s) \leq \frac{\|\gamma_\delta\|_1}{1 - T_1}.$$

This is the first uniform estimate, and the bounds on spatial derivatives require the smoothing theorem. Lemma 2.1 of OpenAI (2026a) gives uniform convergence in space and uniform bounds for $a_\delta, b_\delta, d_\delta$ on $[0,T]$, with the next spatial derivatives bounded as well. By Lemma 2.2, the diffusions coupled through the same Brownian motion satisfy $\sup_{s \leq T} |X_s^\delta - X_s| \to 0$. The derivative bounds control evaluation along these paths, so the integrands in (6) converge and are uniformly bounded, and dominated convergence passes their expectations to the limit. These estimates depend only on the common L^1 bound on the coefficient and the fixed distance from time one, whatever the width of the shrinking plateau. The result is $\mathcal{C}_\gamma(t) \geq \gamma(t)^2 \mathbb{E}a^4/1000$, which proves Proposition 3. Neither a time derivative of the coefficient nor minimality has entered the limit.

For a bounded gap, a qualitative crossing criterion already suffices: if $q'' = 0$ implies $q''' > 0$, then q' cannot attain an interior maximum, so the endpoint slope bounds give $q' < 1$ inside the gap and the same strict area deficit. Proposition 3 supplies more for the next argument. It applies at the chosen comparison time without assuming $q'' = 0$, and its fourth-moment lower bound survives the local replacement, where bare strict positivity could be lost in the limit.

4. Strict positivity of the density

Full support alone does not force a pointwise positive density. For instance, $\gamma(t) = \int_0^t (s - \frac{1}{2})^2 \, ds$ has full Stieltjes support but $\gamma'(\frac{1}{2}) = 0$: every interval of positive length has a positive integral, although the integrand vanishes at one point. The quantitative remainder in (5) rules this out for a minimiser. The identity $\mathbb{E}a^2 = 1$ first expresses γ as a quotient of continuous moments and then gives smoothness. Only afterwards do we apply the fixed-time spatial estimate to obtain $\gamma' > 0$.

Corollary 4 (smooth positive density; OpenAI (2026a), Corollary 6.4). Under the hypotheses of Theorem 1, $\gamma(0) = 0$ and $\gamma \in C^\infty([0,1))$, with derivatives at zero taken from the right. Its Stieltjes measure is $\mu_\gamma(dt) = \gamma'(t)\, dt$, and

$$2\gamma'(t)\mathbb{E}[a(t, X_t)^3] \geq \frac{\gamma(t)^2}{1000}\mathbb{E}[a(t, X_t)^4] > 0 \qquad (0 < t < 1). \tag{7}$$

Only nonnegativity of the density at zero is asserted.

Proof. All spatial factors in the expectations below are evaluated at (t, X_t). Since $da = -\gamma a^2\, dt + b\, dB$, Itô's formula applied to a^2 gives, for almost every t, $0 = \frac{d}{dt}\mathbb{E}a^2 = \mathbb{E}b^2 - 2\gamma\mathbb{E}a^3$, where the left side vanishes because full support gives $\mathbb{E}a^2 = 1$. Solving for the coefficient yields

$$\gamma(t) = \frac{\mathbb{E}b^2}{2\mathbb{E}a^3}. \tag{8}$$

At first (8) holds almost everywhere. Its denominator is at least two, because $a \geq 0$ and $\mathbb{E}a^3 \geq (\mathbb{E}a^2)^{3/2} = 1$. The moments are continuous, so right continuity of γ extends (8) to every time and proves that γ is continuous. Evenness gives $b(0,0) = 0$, hence $\gamma(0) = 0$.

For smoothness, fix $T < 1$ and consider the polynomial moments

$$M_P(t) = \mathbb{E}P(\Phi_x(t, X_t), \ldots, \partial_x^m \Phi(t, X_t)),$$

where P is a polynomial in any finite number m of variables. Spatial differentiation of the PDE and Itô's product rule give polynomials Q_P, R_P, involving finitely many further spatial derivatives, such that

$$M_P(t) - M_P(0) = \int_0^t [M_{Q_P}(s) + \gamma(s) M_{R_P}(s)]\, ds.$$

The identity involves γ and none of its derivatives. Bounded spatial derivatives on the fixed strip make every moment continuous. The integral identity then makes every moment C^1, and (8) makes γ C^1. Repeating the argument proves smoothness. This is the quotient-and-moment argument of Chen (2026), Proposition 4.6 and Appendix B.6, which uses the polynomial-moment method of Auffinger and Chen (2015), Lemma 2 and the proof of Theorem 2(ii). OpenAI (2026a), proof of Corollary 6.4 and long §9.1 give the explicit polynomials and the passage from initially nonsmooth coefficients.

The two differentiated moments are now

$$\frac{d}{dt}\mathbb{E}b^2 = \mathbb{E}[d^2 - 6\gamma a b^2], \qquad \frac{d}{dt}\mathbb{E}a^3 = \mathbb{E}[3ab^2 - 3\gamma a^4].$$

For a smooth time coefficient, differentiating $q'' = \mathbb{E}b^2 - 2\gamma\mathbb{E}a^3$ therefore gives

$$q^{(3)} + 2\gamma'\mathbb{E}a^3 = \mathcal{C}_\gamma(t). \tag{9}$$

This identity displays the two uses of the same spatial moment. On a plateau, $\gamma' = 0$ and $(v, w, z) = c(a, b, d)$ recover (4). For the now smooth minimiser $q = t$, so $q^{(3)} = 0$ and the left side is $2\gamma' \mathbb{E}a^3$, while Proposition 3 bounds the right side from below.

We now apply Proposition 3 to the smooth minimiser. Strict increase and $\gamma(0) = 0$ give $\gamma(t) > 0$ for $t > 0$, so (9) gives (7). Strictness follows from $\gamma(t) > 0$ and $\mathbb{E}a^4 \geq (\mathbb{E}a^2)^2 = 1$. In particular, Cauchy–Schwarz gives

$$\mathbb{E}a^3 \leq \sqrt{\mathbb{E}a^2\, \mathbb{E}a^4} = \sqrt{\mathbb{E}a^4} \leq \mathbb{E}a^4,$$

so the same inequality implies $\gamma'(t) \geq \gamma(t)^2/2000$ for $0 < t < 1$.

Proposition 3 passes spatial moments to the limit, and its proof does not identify q_δ''' with the third derivative of the limiting moment. Indeed, the minimiser has $q(t) = t$ and $q''' = 0$, and the coefficient derivative in (9) is exactly the term that the constant-interval formula omits. Long §§9.1–9.2 retain the smoothness argument and the fixed-time transfer, including the common approximation bounds.

Earlier and related estimates. The backward shape comparisons, the transformed density and the positive-temperature gap-crossing argument appear in Lopatto (2026), v3, Propositions 3.6, 7.1 and 9.1. H.-B. Chen (2026), Proposition 4.2 proves the zero-temperature crossing implication $q'' = 0 \Rightarrow q''' > 0$ under the minimiser convention of its §3.1. Lopatto's v4, Proposition 5.1 and Corollary 5.6 give qualitative crossing for arbitrary positive-temperature probability measures. In unit-noise coordinates, the cumulative coefficient there is $\gamma_\beta = \beta\alpha_\nu$, where $\alpha_\nu(t) = \nu([0, t])$, and the terminal value is $\beta^{-1} \log\cosh(\beta x)$. Every bounded $\gamma \in \mathcal{U}$ has this form: choose $\beta > 0$ with $\beta \geq \sup \gamma$ and put the remaining probability mass at one. The coefficient classes alone therefore do not distinguish the estimates. What distinguishes them is the terminal datum and the quantitative control through its zero-temperature limit.

Lopatto's external-field paper, Theorem 2.3 and §8 also uses a fixed-time spatial inequality for support and positive density. At a continuity point, monotonicity supplies its one-sided coefficient bounds, and on a smooth full-support interval of the minimiser its scalar conditions become equalities. Its Appendix B gives the positive-temperature, zero-field strict-density result. The bound used here instead applies at a nonstationary comparison time and keeps an explicit fourth-moment margin through the local replacement with terminal datum $|x|$. The rescaling and the exact comparison of hypotheses are in long §2.3, finite-temperature dictionary.

5. Consequences for optimisation

5.1 A martingale formula for the ground-state value

The support identities give a scalar representation of P_*. Set $M_t = u(t, X_t)$, the bounded gradient martingale, which satisfies $dM_t = a(t, X_t)\, dB_t$. Its quadratic variations distinguish the two Hessian moments:

$$d\langle M\rangle_t = a(t, X_t)^2\, dt, \qquad d\langle M, B\rangle_t = a(t, X_t)\, dt.$$

Full support gives $\mathbb{E}a(t, X_t)^2 = 1$, so each increment $M_t - M_s$ has unconditional variance $t - s$. This normalises the discrete increments below. The first moment of a gives instead the covariance with the driving noise. Let U be the almost-sure and L^2 limit of M_t. The value identity is

$$P_* = \mathbb{E}[UB_1] = \int_0^1 \mathbb{E}[a(t, X_t)]\, dt. \tag{10}$$

Boundedness gives $|U| \leq 1$, while $\mathbb{E}M_t^2 \to 1$ gives $\mathbb{E}U^2 = 1$, so $U^2 = 1$ almost surely.

The cancellation behind (10) runs as follows. Terminal convergence of the convex value functions gives $UX_1 = |X_1|$ (the equality is automatic on $X_1 = 0$), and Itô's formula gives $\mathcal{P}(\gamma) = \mathbb{E}|X_1| - \int_0^1 t\gamma(t)\,dt$. Multiplying $X_1 = B_1 + \int_0^1 \gamma(t)M_t\,dt$ by U cancels that integral, because $\mathbb{E}[UM_t] = t$, and the Parisi formula then gives (10). OpenAI (2026a), Corollary 7.1 and the long record, §6, "Recovering the value from a martingale", prove the endpoint limits and the stochastic-integral covariance. The argument uses only full support. Moreover, convexity and Cauchy–Schwarz give $0 \le \mathbb{E}a \le 1$, so

$$0 \le P_* - \int_0^T \mathbb{E}[a(t, X_t)]\,dt \le 1 - T.$$

We can therefore fix the terminal cutoff before discretising the diffusion.

The gradient-martingale and value methods have antecedents in Montanari, preprint v2, Lemmas 3.2–3.4, under its positive-temperature support assumption, and in the zero-temperature calculations of El Alaoui, Montanari and Sellke, preprint v1, Lemmas 6.5 and 6.13. Sellke, preprint v4, Proposition 3.10 and Lemma 3.11, p. 17 gives the closely related finite-cutoff cancellation by applying Itô's formula to Φ and $x\Phi_x$.

5.2 The classical algorithm

Full support and existence give an attained, strictly increasing minimiser in the monotone zero-temperature class, which is Assumption 2 of El Alaoui, Montanari and Sellke, preprint v1. Their Corollary 2.2 then gives, for every fixed $\varepsilon > 0$, a randomised algorithm that returns a configuration within εn of the finite SK maximum with probability tending to one over couplings and algorithmic randomness. It costs $C(\varepsilon)n^2$ real-arithmetic operations when the prescribed scalar functions and normalisers are available. This count includes no separate bound for numerical preprocessing or finite precision. OpenAI (2026a), Corollary 7.3 and long §9.4 work out the Gaussian convention and the operation count. Chen already notes the classical consequence in §1, p. 4.

The incremental approximate message-passing iterates of El Alaoui, Montanari and Sellke use the coupling matrix, and their scalar description uses the martingale multiplier and state drift

$$u_{\mathrm{EAMS}}(t, x) = a(t, x), \qquad v_{\mathrm{EAMS}}(t, x) = \gamma(t)u(t, x).$$

Their noise multiplier is thus the Hessian, whereas our u is the gradient. In our notation, the bounded martingale M_t and $\mathbb{E}a^2 = 1$ give the required variance normalisation. With the fixed-strip regularity and discretisation estimates, their Theorem 2 achieves energy arbitrarily close to the truncated first-moment integral in (10).

Strict positivity of the density is unnecessary for this algorithm.

5.3 A fixed QAOA circuit before the thermodynamic limit

The remaining consequence uses a separate circuit approximation theorem, stated as Proposition 5 below. We give its finite Gaussian input and identify the energy it approximates. The companion develops its operator and control construction, and the support and density proofs above need none of those additional prerequisites.

Let $\widehat{Z}_i, \widehat{X}_i$ be Pauli operators, and define the cost and mixer by

$$C_{n,J} = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij}\widehat{Z}_i\widehat{Z}_j, \qquad B_n = \sum_i \widehat{X}_i.$$

The eigenvalues of the cost are $H_n(\sigma)$. QAOA alternates cost evolution and mixing, starting from the uniform state $|+\rangle^{\otimes n}$ (Farhi, Goldstone and Gutmann (2014)). For a fixed sequence $W = (\theta_1, \beta_1, \dots, \theta_p, \beta_p)$, write

$$|\psi_{n,J}(W)\rangle = e^{-i\beta_p B_n} e^{-i\theta_p C_{n,J}} \cdots e^{-i\beta_1 B_n} e^{-i\theta_1 C_{n,J}} |+\rangle^{\otimes n},$$

$$v(W) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_J \langle \psi_{n,J}(W) | C_{n,J} | \psi_{n,J}(W) \rangle.$$

Farhi, Goldstone, Gutmann and Zhou (2022) established this fixed-angle limit. OpenAI (2026b), Theorem 1.1 proves that for every $\varepsilon > 0$ there is one finite W with $v(W) \geq P_* - \varepsilon$, whose depth and angles depend on ε but not on n or J. The instance still determines $C_{n,J}$. Here depth counts global cost–mixer layers, whose elementary gate counts depend on the implementation.

This order of limits differs from the fixed-instance depth limit in Farhi, Goldstone and Gutmann, §VI, which permits instance-dependent choices. Basso, Farhi, Marwaha, Villalonga and Zhou (2022), §7, equation (7.2) conjectured the common-circuit conclusion.

The finite Gaussian input. Fix the minimiser and its diffusion, a cutoff $0 < T < 1$, and a mesh $t_j = j\Delta$ with $\Delta = T/N$. The mesh count N is distinct from the number of spins n. From independent standard normals $Z_1, \dots, Z_{N+1}$, form

$$\begin{aligned} &Y_0 = 0, \\ &Y_j = Y_{j-1} + \sqrt{\Delta} Z_j + \Delta \gamma(t_{j-1}) u(t_{j-1}, Y_{j-1}), \\ &h_j = \frac{a(t_{j-1}, Y_{j-1})}{\sqrt{\mathbb{E} a(t_{j-1}, Y_{j-1})^2}}, \qquad g_j = h_j Z_j \quad (1 \leq j \leq N). \end{aligned}$$

OpenAI (2026a), Proposition 7.2 proves that the denominators are positive for all sufficiently large N. Since h_j depends only on $Z_1, \dots, Z_{j-1}$, the g_j are centred and orthonormal in L^2, although they need not be independent Gaussians. With Φ_{G} the standard normal distribution function, set

$$\begin{aligned} F_N &= \operatorname{sgn}\{u(T, Y_N) + 2\Phi_{\mathrm{G}}(Z_{N+1}) - 1\}, \\ \mathbb{E}[F_N \mid Z_1, \dots, Z_N] &= u(T, Y_N). \end{aligned}$$

The second equality holds because $\Phi_{\mathrm{G}}(Z_{N+1})$ is uniform on $[0, 1]$, so this random rounding preserves every correlation with a function of the preceding coordinates. The circuit construction needs one finite quantity, the adjacent pairing

$$\begin{aligned} S_{T,N} &:= \sum_{j=1}^{N} \mathbb{E}[F_N g_j]\, \mathbb{E}[F_N Z_{j+1}], \\ \lim_{N\to\infty} S_{T,N} &= I_T := \int_0^T \mathbb{E}[a(t, X_t)]\, dt. \end{aligned} \tag{11}$$

The same proposition proves this limit. Its proof starts from the Euler approximation and the gradient martingale, which give

$$\left\| u(T, Y_N) - \sum_{j=1}^{N} \sqrt{\Delta}\, a(t_{j-1}, Y_{j-1}) Z_j \right\|_2 \longrightarrow 0.$$

Conditional-mean preservation transfers these correlations to F_N for the first N coordinates. The rounding error itself stays of order one, since its squared L^2 error satisfies $\mathbb{E}(F_N - u(T, Y_N))^2 \to 1-T$. Bessel's inequality, applied separately to (g_j) and (Z_j), controls the errors of both correlation vectors in ℓ^2, and their bounded norms control the growing pairing by Cauchy–Schwarz. The terms with $j < N$ give a Riemann sum, and the last term, from the rounding coordinate, vanishes. All these limits keep T fixed. Only afterwards does $T \uparrow 1$ give P_*. The construction and the estimates are in long §7, "A finite Gaussian spin and its adjacent correlations".

Proposition 5 (circuit approximation; a consequence of OpenAI (2026b), §§2–5). For each fixed T, N in the preceding construction with finite positive normalisers, and every $\eta > 0$, there is a finite ordinary QAOA word W, with deterministic angles independent of n and J, such that

$$|v(W) - S_{T,N}| < \eta.$$

The adjacent pairing as a rooted energy. We use the rooted Gaussian representation of OpenAI (2026b), §2 and Lemma 3.1, which builds on the spin–boson description of Boulebnane et al. (2026). Its classical sector consists of square-integrable functions built recursively from a local Gaussian seed and child-generated Gaussian coordinates, with inner product given by expectation. For any finite list of fixed centred real child directions, their coordinates at the receiving root have a joint centred Gaussian law, and for two such directions f, g,

$$\mathbb{E}[\mathsf{G}(f)\mathsf{G}(g)] = \langle f, g\rangle.$$

These directions are fixed elements of the child space, even when their values depend on the child's Gaussian coordinates. Moreover, the seed of the receiving root is independent of these coordinates, and compatible embeddings preserve them as the recursive depth grows. The root shift L inserts a child direction into the first Gaussian chaos (the linear Gaussian part) with constant receiving-root state, and its adjoint projects onto that part and returns the child direction. A one-pair calculation at the start of long §8 compares the energy of this representation with the finite SK circuit before the general construction in §8.1.

We start with a standard normal receiving-root seed g_0 and put $Z_{j+1} = \mathsf{G}(g_j)$ recursively. The orthonormal child directions give independent standard Gaussian lifts, which reproduce the scalar law. Since the finite spin F_N depends only on those lifts, projection gives the exact rooted-energy identity

$$\begin{aligned} \mathsf{L}^* F_N &= \sum_{j=0}^{N} \mathbb{E}[F_N Z_{j+1}] g_j, \\ \mathcal{E}(F_N) := \mathrm{Re}\langle F_N, \mathsf{L}F_N\rangle &= \sum_{j=0}^{N} \mathbb{E}[F_N g_j]\, \mathbb{E}[F_N Z_{j+1}] = S_{T,N}. \end{aligned} \tag{12}$$

Independence of the receiving-root seed removes the $j = 0$ term. This identity concerns the finite Gaussian formula in a limiting rooted space and makes no statement about a finite-n circuit. An unshifted pairing can have the same scalar limit without identifying this energy. Long §8.1 proves the construction and the projection.

The shift carries the lift from a child direction g_j to the receiving coordinate Z_{j+1}, and a one-step example separates it from the unshifted pairing. For the nonminimising heat coefficient $\gamma = 0$, take $T = 4/5$ and $N = 1$. Then $g_1 = Z_1$ and $F_1 = \mathrm{sgn}(2Z_1 + Z_2)$, and Gaussian regression gives

$$S_{T,1} = \frac{4}{5\pi}, \qquad \mathbb{E}[F_1 g_1]\, \mathbb{E}[F_1 Z_1] = \frac{8}{5\pi}.$$

The example tests only the finite representation, since the limit in (11) depends on the minimiser. The calculation is in long §10.4, "An exact one-step calculation".

The construction behind Proposition 5. OpenAI (2026b), §§3–5 smooths the spin and approximates every transmitted function by a bounded smooth map that preserves oddness. Smoothing the final sign alone would be insufficient. The class exclusions of the construction prevent return through a traversed branch and reserve helper classes on which the output is zero, with positive probabilities fixed during synthesis. This finite formula is then approximated in energy using temporary Gaussian controls, with the same sitewise seeds throughout the word. Removing those controls produces an ordinary word, whose limiting energy is $v(W)$ by the fixed-word transfer. The approximation errors are absolute and can be chosen successively for each fixed $S_{T,N}$, before it is chosen close to P_*. Energy is the preserved quantity, and the conclusion about a measured configuration below uses a further concentration theorem. The complete error bound is in long §8.6, finite-formula energy approximation, and the constructions and their hypotheses are in long §§8.1–8.6.

Near-optimality follows from the decomposition

$$P_* - v(W) \le (P_* - I_T) + |I_T - S_{T,N}| + |S_{T,N} - v(W)|. \tag{13}$$

Choose T so that $1 - T < \varepsilon/3$, then N so that the second term is below $\varepsilon/3$, and finally apply Proposition 5 with $\eta < \varepsilon/3$. These finite choices give $v(W) > P_* - \varepsilon$ before n tends to infinity. Appendix A of the source supplies an alternative positive-temperature input to the same synthesis.

The construction gives no bound on the depth or on the cost-angle magnitudes in terms of the accuracy, and no efficient procedure for finding the angles. A separate estimate for the expected energy of a specified word appears in long §11.5. Its threshold depends on the depth and cost angles of the word, so without bounds on those parameters it gives no threshold in terms of optimisation accuracy alone. It also gives no rate for the concentration of measurements.

Corollary 6 (a measured configuration on the same instance). For every $\varepsilon > 0$ there is a finite ordinary QAOA word W_ε, chosen independently of n and J, such that one computational-basis measurement Σ satisfies

$$\Pr_{J,\Sigma}\Big\{\max_\sigma H_{n,J}(\sigma) - H_{n,J}(\Sigma) \le \varepsilon n\Big\} \longrightarrow 1 \qquad (n \to \infty). \tag{14}$$

Proof. For a fixed word, put $Y_n = H_{n,J}(\Sigma)/n$ and $K_n = \max_\sigma H_{n,J}(\sigma)/n$. Farhi, Goldstone, Gutmann and Zhou, §4, equations (13)–(22) prove

$$\mathbb{E}_{J,\Sigma} Y_n \to v(W), \qquad \mathbb{E}_{J,\Sigma} Y_n^2 \to v(W)^2.$$

The second moment is $\mathbb{E}_J\langle (C_{n,J}/n)^2\rangle$, so it includes measurement variance as well as disorder fluctuations. Thus $Y_n \to v(W)$ in joint L^2. Gaussian concentration of the maximum and its defining mean limit give $K_n \to P_*$ in probability. Choose W with $P_* - v(W) < \varepsilon/3$. With probability tending to one, both $K_n - P_* < \varepsilon/3$ and $v(W) - Y_n < \varepsilon/3$, which proves (14). The two energies use the same J, and the argument needs no independence between them. The concentration argument and its further probability consequences are in the companion's §11.2, "What is averaged, and a fixed-word probability consequence".

Formal verification

The proofs above are ordinary mathematical arguments that import three analytic inputs: the integrated estimate (5), the shape comparisons behind it and the compact-strip approximation. OpenAI's formalisation in Lean with mathlib contains two formal statements that bear on this paper. The formal statement of full support asserts the support and moment conclusions of Theorem 1, with an admissible minimiser as its hypothesis, and constructs the associated diffusion. Lean checks OpenAI's proof of the corresponding theorem.

A separate formal statement of the scalar consequences covers the value identity and the finite Gaussian approximation, with minimiser and diffusion premises for those clauses, and its third clause identifies the ground-state and Parisi values from a minimiser, constructing the diffusion internally. Its proof has not been checked, so those results rest here on the paper proofs cited in §5. The two formal statements use separate models of the order parameter, which have not been related to each other. Neither covers Corollary 4 or the convergence of QAOA. The long record compares each formal statement with the printed one in §12, "The formal statements and their scope", and gives versions and checking outcomes in §13, "Source checks and finite calculation".

References

- Auffinger, A. and Chen, W.-K. (2015). "On properties of Parisi measures." *Probability Theory and Related Fields* **161**, 817–850. Preprint v1: Lemma 2 and proof of Theorem 2(ii).
- Auffinger, A. and Chen, W.-K. (2017). "Parisi formula for the ground state energy in the mixed p-spin model." *Annals of Probability* **45**, 4617–4631. Preprint v2: Theorem 1 and proof of Lemma 3.
- Auffinger, A., Chen, W.-K. and Zeng, Q. (2020). "The SK model is infinite step replica symmetry breaking at zero temperature." *Communications on Pure and Applied Mathematics* **73**, 921–943. arXiv:1703.06872v2: Theorem 1, Proposition 1 and Remark 7.
- Basso, J., Farhi, E., Marwaha, K., Villalonga, B. and Zhou, L. (2022). "The Quantum Approximate Optimization Algorithm at High Depth for MaxCut on Large-Girth Regular Graphs and the Sherrington–Kirkpatrick Model." *TQC 2022*, 7:1–7:21. Full preprint v3: transfer theorem and §7, equation (7.2); the full preprint has separate numbering.
- Boulebnane, S., Khan, A., Liu, M., Larson, J., Herman, D., Shaydulin, R. and Pistoia, M. (2026). *Spin–Boson Mapping of the Quantum Approximate Optimization Algorithm*. Preprint v2: Theorem 2 (informal) and equations (10)–(11).
- Chen, H.-B. (2026). *FRSB in the SK spin glass: convergence to full-interval support at zero temperature*. arXiv:2607.18032v1, 20 July 2026. §1, p. 4; Theorem 1.3; Propositions 4.2, 4.5 and 4.6; Appendices B.6 and C.
- Chen, W.-K., Handschy, M. and Lerman, G. (2018). "On the energy landscape of the mixed even p-spin model." *Probability Theory and Related Fields* **171**, 53–95. Preprint v4: Proposition 2, Theorem 5, Lemma 3 and Propositions 3–4.
- de Moura, L. and Ullrich, S. (2021). "The Lean 4 Theorem Prover and Programming Language (System Description)." *CADE 28*. System description.
- El Alaoui, A., Montanari, A. and Sellke, M. (2021). "Optimization of mean-field spin glasses." *Annals of Probability* **49**, 2922–2960. arXiv:2001.00904v1: Theorem 2, p. 6; Remark 2.1 and

Theorem 3, p. 7; Assumption 2 and Corollary 2.2, p. 8; §§5–6.

- Farhi, E., Goldstone, J. and Gutmann, S. (2014). *A Quantum Approximate Optimization Algorithm*. Preprint v1. Algorithm definition; §VI concerns the fixed-instance depth limit.
- Farhi, E., Goldstone, J., Gutmann, S. and Zhou, L. (2022). "The Quantum Approximate Optimization Algorithm and the Sherrington–Kirkpatrick Model at Infinite Size." *Quantum* **6**, 759. arXiv:1910.08187v4.
- Jagannath, A. and Tobasco, I. (2016). "A dynamic programming approach to the Parisi functional." *Proceedings of the AMS* **144**, 3135–3150. Preprint v4. Positive-temperature control methods.
- Lopatto, P. (2026). *Full replica symmetry breaking in the Sherrington–Kirkpatrick model*. arXiv:2607.11756v3, revised 16 July 2026. Propositions 3.6, 7.1 and 9.1; related Cole–Hopf, density and crossing arguments.
- Lopatto, P. (2026). *Full replica symmetry breaking in the Sherrington–Kirkpatrick model*. arXiv:2607.11756v4, revised 5 October 2026. Proposition 5.1 and Corollary 5.6; qualitative crossing for arbitrary probability measures.
- Lopatto, P. (2026). *Parisi measure of the Sherrington–Kirkpatrick model with external field*. arXiv:2610.00127v1. Theorem 2.3 and Appendix B; related positive-temperature two-use argument.
- The mathlib Community (2020). "The Lean Mathematical Library." *Certified Programs and Proofs (CPP '20)*. Library description.
- Montanari, A. (2025). "Optimization of the Sherrington–Kirkpatrick Hamiltonian." *SIAM Journal on Computing* **54**, FOCS19-1–FOCS19-38. Preprint v2, 2019: §§2–3.
- OpenAI (2026a). *Full support of the zero-temperature Sherrington–Kirkpatrick order parameter*. 27 September 2026. Manuscript, fixed revision. Theorem 1.1; §§2–6; Corollaries 6.4 and 7.1; Proposition 7.2.
- OpenAI (2026b). *QAOA attains the SK ground-state energy in the thermodynamic-first limit*. 25 September 2026. Manuscript, fixed revision. Theorem 1.1; §§2–6 and Appendix A.
- Sellke, M. (2021). *Optimizing Mean Field Spin Glasses with External Field*. Preprint v4, 6 October 2021. Proposition 3.10 and Lemma 3.11, p. 17.
- Sherrington, D. and Kirkpatrick, S. (1975). "Solvable Model of a Spin-Glass." *Physical Review Letters* **35**, 1792–1796. Original model.
- Zhou, Y. (2026). "On the Gardner Transition in the Ising Pure p-Spin Glass II". arXiv:2608.06523v1, Appendix E.3, (E.37), (E.41).