

ROOT STRUCTURE OF A FINITE ADSORPTION CELL

Root cascade, occupancy distribution, and spectral structure of the equilibrium response

Abstract

We consider the finite partition function of the simplified statistical isotherm (SSI) for the occupancy of a single zeolite cavity,

$$Z_m(x; \alpha) = \sum_{n=0}^m \left[\frac{(1 - n\alpha)^n}{n!} \right] x^n,$$

where m is the integer limiting capacity of the cell, $\alpha = \beta/V_c$ is the dimensionless free-volume parameter, and x is the dimensionless activity. For $0 < \alpha < 1/m$, all states $n = 0, \dots, m$ have positive weights, and $Z_m(x) > 0$ for any physical activity $x > 0$.

The coefficients Z_m specify the statistical weights of the states, while the normalized function

$$G_x(z) = \frac{Z_m(xz)}{Z_m(x)}$$

is the probability-generating function of the occupancy distribution. If all zeros of Z_m are real and negative,

$$Z_m(x) = \prod_{k=1}^m \left(1 + \frac{x}{r_k} \right),$$

then $G_x(z)$ factorizes into m Bernoulli factors, and the number of molecules in the cell has the distribution of a sum of m independent Bernoulli random variables with probabilities

$$\pi_k = \frac{x}{(x + r_k)}.$$

Consequently,

$$\langle N \rangle = \frac{\sum_{k=1}^m x}{(x + r_k)},$$

$$\text{Var}(N) = \frac{\sum_{k=1}^m x r_k}{(x + r_k)^2}.$$

In the variable $s = \ln x$, the function

$$\rho(s) = \frac{\text{Var}(N)}{m}$$

is nonnegative and normalized to unity, so it is convenient to regard it as the normalized density of the equilibrium response on the logarithmic-activity axis. In the fully real regime, ρ is the convolution of the discrete distribution $\ln r_k$ with the standard logistic kernel. The root spectrum therefore determines the mean position, variance, skewness, and higher cumulants of the response profile. The response maximum, half-occupancy point, and mean position in $\ln x$ correspond, respectively, to the mode, median, and mean of this normalized density.

For $m = 14$, the exact resultant and a complete count of real roots yield seven internal values of α at which a double root appears. The last of these boundaries is

$$\alpha_c = 0.0554547158587412, \\ \eta_c = 14 \alpha_c = 0.776366022022377.$$

At $\alpha = \alpha_c$, the colliding pair already forms a double real root; therefore, the fully real region begins at the boundary itself: $\alpha_c \leq \alpha < 1/14$. The CO_2 and CH_4 parameters for NaY lie within this region.

For $m \geq 2$, the quantities α and x are uniquely reconstructed from a single occupancy distribution using only P_0 , P_1 , and P_2 . Thereafter, the probabilities $P_3, \dots, P_m$ are completely determined by the same statistical weight and contain no new parameters.

In the near-boundary large-capacity regime, the reversed and normalized polynomial locally approaches a partial theta function. Uniform control of the coefficient tail makes it possible to carry out this passage together with the derivatives entering the double-root system. The first nondegenerate double point of the partial theta function determines a unique local asymptotic branch of double roots of the finite-dimensional problem:

$$\eta_m^{(*)} = 1 - \frac{\gamma_1}{\sqrt{m}} + O\left(\frac{1}{m}\right),$$

where

$$\gamma_1 = 0.92307837999 \dots$$

For the computed sequence $m=2, \dots, 30$, this branch coincides with the final boundary of the fully real spectrum. Preservation of this coincidence for arbitrarily large m requires separate global control of the remaining roots of the finite polynomial. These results connect the root structure of Z_m with the occupancy distribution and the shape of the equilibrium response of a finite cell.

Keywords: adsorption, finite partition function, free volume, polynomial zeros, root cascade, occupancy distribution, equilibrium response, partial theta function, factorial cumulants.

1. Problem formulation

A discrete statistical description of zeolite-cavity occupancy was proposed by Ruthven [1] and is used in the form of the simplified statistical isotherm (SSI). For NaY NIST RM 8850, the parameters of this description were determined from single-component CO_2 and CH_4 isotherms [2].

In the single-component case, the finite partition function has the form

$$Z_m(x; \alpha) = \sum_{n=0}^m \left[\frac{(1 - n\alpha)^n}{n!} \right] x^n,$$

where

$$\alpha = \frac{\beta}{V_c},$$

$$x = K(T)y,$$

and y denotes the physical variable of the chosen parameterization. For the original pressure parameterization of NaY, $y = p$ is used; for the fugacity parameterization, $y = f$ is used.

The single-cell formulation of the model, including the state distribution, limiting regimes, and the response identity, is given in [3]. In the notation used here,

$$P_n(x; \alpha) = \frac{a_n(\alpha) x^n}{Z_m(x; \alpha)},$$

$$a_n(\alpha) = \frac{(1 - n\alpha)^n}{n!},$$

and

$$\langle N \rangle = \frac{d \ln Z_m}{d \ln x},$$

$$\frac{d \langle N \rangle}{d \ln x} = \text{Var}(N).$$

Thus, the state probabilities, mean occupancy, and equilibrium response are determined by a single finite partition function.

The present work investigates the internal root structure of this sum: at which values of the free-volume parameter α a spectrum consisting entirely of negative real roots arises, how the number of real roots changes upon transition between different regimes, and what consequences this structure has for the occupancy distribution and equilibrium response. The connection between negative real zeros of a finite polynomial and the distribution of a sum of independent Bernoulli random variables follows from the theory of PF_∞ sequences [4]. For the family considered here, this factorization connects the roots with the complete occupancy distribution, its cumulants, and the shape of the equilibrium response.

For $m = 14$, the resultant with respect to x has seven roots in α within the physical interval $0 < \alpha < 1/14$. Between successive values, the number of real zeros of Z_{14} is successively 0, 2, 4, ..., 14. Thus, variation of α generates an ordered cascade of pairwise collisions terminating in a fully real spectrum.

In the fully real region, the factorization of Z_m converts the mean occupancy and its response into sums over the root scales r_k . In the variable $s = \ln x$, the normalized response

$$\rho(s) = \frac{\text{Var}(N)}{m}$$

is nonnegative and has unit integral normalization. Therefore, $\rho(s)$ can be regarded as a normalized response density on the $\ln x$ axis. In the fully real-spectrum regime, this density is the convolution of the discrete distribution $\ln r_k$ with a logistic kernel, which directly links the root spectrum to the position, width, skewness, and higher cumulants of the equilibrium response.

The complete distribution makes it possible to test the functional form of the statistical weight rather than only the mean response. For $m \geq 2$, the probabilities P_0 , P_1 , and P_2 uniquely determine α and x , after which $P_3, \dots, P_m$ are fixed without introducing new parameters and provide a direct

test of the dependence of the weight on occupancy. An independent estimate of α can be obtained from the first two coefficients of the low-activity expansion of the mean isotherm.

At large capacity, the upper edge of occupancy space $n = m, m-1, m-2, \dots$, after reversal of the polynomial, leads to a partial theta function. Below, it is shown that the transition to this function can be carried out locally uniformly together with the first two derivatives. This yields the asymptotics of the branch of double roots converging to the first spectral point of the partial theta function.

For NaY, the calculation is performed at fixed integer capacity $m=14$. The parameter α determines the relative statistical weights a_n and hence the shape of the occupancy distribution and the root structure of Z_m . The function $K(T)$ relates the dimensionless activity to the physical variable of the chosen parameterization,

$$x = K(T)y,$$

where $y=p$ in the original pressure representation and $y=f$ in the fugacity representation. The factor c_{ads} converts mean occupancy into absolute loading. This separation makes it possible to compare independently with the data the shape of the isotherm, its temperature position, and its absolute scale.

2. Finite cell with free volume

2.1. State weights

Let V_c be the volume of the selected adsorption cavity, and let β be the effective volumetric parameter of the SSI free-volume closure. For NaY, $V_c = 958.2 \text{ \AA}^3$ is used, corresponding to the adopted volume of a faujasite-framework supercage. The parameter β enters the statistical weight of a state and characterizes the reduction of the accessible configurational volume as occupancy increases. Its numerical value is a model parameter and need not coincide with the geometric volume of an isolated molecule.

For a state containing n molecules,

$$V_{free}(n) = V_c - n\beta.$$

After normalization of the configurational part, the state weight has the form

$$\begin{aligned} a_0 &= 1, \\ a_n(\alpha) &= \frac{(1 - n\alpha)^n}{n!}, n = 1, \dots, m, \\ \alpha &= \frac{\beta}{V_c}. \end{aligned}$$

The number m is the integer limiting capacity of the cell in the adopted SSI parameterization. It is fixed separately from the continuous parameter α . To preserve positive free volume in all states $n = 0, \dots, m$, it is necessary that

$$0 < \alpha < \frac{1}{m}.$$

Then $a_n > 0$ for all n , and for $x > 0$

$$Z_m(x; \alpha) > 0.$$

The roots considered below therefore lie outside the positive real axis of the activity variable, i.e. outside the physical domain.

2.2. Partition function, distribution, and observables

$$Z_m(x; \alpha) = \sum_{n=0}^m a_n(\alpha) x^n,$$

$$P_n(x; \alpha) = \frac{a_n(\alpha) x^n}{Z_m(x; \alpha)}.$$

The activity is defined as

$$x = K(T) y,$$

$$K(T) = K_0 \exp\left[\frac{E}{(RT)}\right].$$

Here E is the parameter governing the temperature dependence of the adsorption-equilibrium constant, R is the universal gas constant, and y is the physical variable of the chosen parameterization. In the original pressure treatment of NaY, $y = p$ is used. In the fugacity treatment, $y = f$ is used, and K_0 , E , and β must be estimated jointly in that same variable.

All occupancy moments are obtained from Z_m :

$$\langle N \rangle = \frac{x d(\ln Z_m)}{dx},$$

$$\text{Var}(N) = \frac{d\langle N \rangle}{d(\ln x)},$$

$$\kappa_q = \frac{d^q(\ln Z_m)}{d(\ln x)^q}.$$

The first of these equalities gives the mean loading of a single cell. Conversion to the amount of substance per mass of adsorbent requires a separate scale factor: the number of equivalent cells per gram divided by Avogadro's number. This factor does not affect the zeros of Z_m , α , or the normalized distribution P_n .

3. Factorization of the distribution for a real root spectrum

3.1. Statement

Let $Z(x) = \sum_{n=0}^m a_n x^n$, $a_n > 0$, and $a_0 = 1$. Consider the distribution $P_n(x) = \frac{a_n x^n}{Z(x)}$ at $x > 0$.

Theorem. For fixed $x > 0$, the distribution

$$P_n(x) = \frac{a_n x^n}{Z(x)}$$

can be represented as the distribution of a sum of m independent Bernoulli random variables if and only if all zeros of the polynomial Z are real and negative. Since this condition depends only on Z and not on the selected activity value, existence of such a factorization at one $x > 0$ is equivalent to its existence for all $x > 0$.

If the zeros have the form $\xi_k = -r_k$, $r_k > 0$, then

$$Z(x) = \prod_{k=1}^m \left(1 + \frac{x}{r_k} \right),$$

and the random variable N has the same distribution as

$$N = B_1 + \dots + B_m,$$

where B_k are independent and

$$Pr(B_k=1) = \pi_k(x) = \frac{x}{(x+r_k)},$$

$$Pr(B_k=0) = 1 - \pi_k(x) = \frac{r_k}{(x+r_k)}.$$

3.2. Proof

Introduce the normalized probability-generating function of the distribution:

$$G_x(z) = \sum_{n=0}^m P_n(x) z^n = \frac{Z(xz)}{Z(x)}.$$

For $\xi_k = -r_k$, factorization gives

$$G_x(z) = \prod_{k=1}^m \left[\frac{\left(1 + \frac{xz}{r_k} \right)}{\left(1 + \frac{x}{r_k} \right)} \right] = \prod_{k=1}^m [1 - \pi_k(x) + \pi_k(x)z].$$

Each factor is the probability-generating function of a Bernoulli random variable, while the product is the probability-generating function of their independent sum. This proves sufficiency.

Conversely, if N is a sum of independent Bernoulli random variables with probabilities $\pi_k(x)$, then

$$G_x(z) = \prod_{k=1}^m [1 - \pi_k(x) + \pi_k(x)z].$$

All zeros of this function with respect to z are negative and real. Since the zeros G_x are equal to $\frac{\xi_k}{x}$, the zeros of Z must likewise be negative and real. Necessity is proved.

3.3. Number of complex-conjugate root pairs

Define

$$C = \frac{(m - N_{real})}{2},$$

where N_{real} is the number of real zeros of Z counted with multiplicity. Since the coefficients are positive and $Z(0) = 1$, every real zero is negative. Thus, C is exactly equal to the number of complex-conjugate root pairs.

When

$$C=0$$

all m roots are real and negative, and the occupancy distribution factorizes into the distribution of a sum of m independent Bernoulli random variables. This factorization is an exact representation of the probability measure of a single finite cell. The Bernoulli variables correspond to the linear root factors Z_m and do not imply a partition of the physical cavity into m independent adsorption sites.

When

$$C>0$$

C complex-conjugate pairs remain, and therefore such a real Bernoulli representation of the distribution is absent.

If two complex roots collide on the negative real axis, a double real root arises at the collision point itself. Since N_{real} is counted with multiplicity, the value of C exactly at the boundary coincides with the value after the collision.

4. Consequences of root factorization

4.1. Spectral law of the response

Set

$$s = \ln x,$$

and introduce the normalized occupancy and normalized response

$$F(s) = \frac{\langle N \rangle}{m},$$

$$\rho(s) = \frac{\text{Var}(N)}{m} = \frac{dF}{ds}.$$

In the physical region $0 < \alpha < 1/m$, all statistical weights are positive. At $s \rightarrow -\infty$ we have $F \rightarrow 0$, while at $s \rightarrow +\infty$ we have $F \rightarrow 1$; for finite $x > 0$ the variance is positive. Therefore,

$$\rho(s) > 0,$$

$$\int_{-\infty}^{+\infty} \rho(s) ds = 1.$$

Consequently, $\rho(s)$ is a normalized density of the equilibrium response on the $s = \ln x$ axis, and $F(s)$ is the corresponding cumulative distribution function. These properties follow from the finite statistical measure and do not require all zeros to be real.

Additional spectral structure arises in the regime $C=0$. Denote

$$u_k = \ln r_k.$$

Then root factorization gives

$$F(s) = \frac{1}{m} \sum_{k=1}^m \frac{1}{1 + \exp(u_k - s)},$$

$$\rho(s) = \frac{1}{m} \sum_{k=1}^m \frac{1}{4 \cosh^2\left(\frac{s-u_k}{2}\right)}.$$

If

$$l(t) = \frac{1}{4 \cosh^2\left(\frac{t}{2}\right)}$$

is the standard logistic density and

$$\mu_m(u) = \frac{1}{m} \sum_{k=1}^m \delta(u - u_k),$$

then

$$\rho = \mu_m * l.$$

Thus, it is precisely in the fully real regime that the normalized response acquires an exact spectral representation as the convolution of a discrete measure of root scales with the universal logistic kernel.

This equality has a simple probabilistic interpretation. Let the random variable U take each value u_k with probability $1/m$, and let L have the standard logistic distribution and be independent of U. Then

$$S = U + L$$

has density $\rho(s)$.

Therefore, the cumulants of the response in the variable s separate into a contribution from the root spectrum and the known contribution of the logistic kernel. If

$$\bar{u} = \left(\frac{1}{m}\right) \sum_{k=1}^m u_k,$$

$$\sigma_u^2 = \left(\frac{1}{m}\right) \sum_{k=1}^m (u_k - \bar{u})^2,$$

then

$$K_1[\rho] = \bar{u},$$

$$K_2[\rho] = \sigma_u^2 + \frac{\pi^2}{3},$$

$$K_3[\rho] = \left(\frac{1}{m}\right) \sum_{k=1}^m (u_k - \bar{u})^3,$$

and

$$K_4[\rho] = \left(\frac{1}{m}\right) \sum_{k=1}^m (u_k - \bar{u})^4 - 3\sigma_u^4 + \frac{2\pi^4}{15}.$$

The first moments of the distribution of $\ln r_k$ can be determined from the shape of the equilibrium response without reconstructing the individual roots.

Since F is the distribution function corresponding to p , its median is determined by the condition

$$F\left(\frac{s_1}{2}\right) = \frac{1}{2},$$

which is equivalent to

$$\langle N \rangle = \frac{m}{2}.$$

Therefore, the half-occupancy activity is the median of the response distribution in $\ln x$. The maximum of p gives its mode, while $\bar{u}$ gives the mean. In general, these three locations differ.

From the leading coefficient Z_m it follows that

$$\prod_{k=1}^m r_k = \frac{1}{a_m}.$$

For the weight considered here,

$$a_m = \frac{(1 - m\alpha)^m}{m!},$$

therefore

$$\exp(\bar{u}) = \left(\prod_{k=1}^m r_k \right)^{1/m} = \frac{(m!)^{1/m}}{(1 - m\alpha)}.$$

The same quantity follows from successive switches of the most probable occupancy. If x_n is defined by the condition $P_n = P_{n+1}$, then

$$\begin{aligned} x_n &= \frac{a_n}{a_{n+1}}, \\ \prod_{n=0}^{m-1} x_n &= \frac{1}{a_m}, \\ \left[\prod_{n=0}^{m-1} x_n \right]^{1/m} &= \exp(\bar{u}). \end{aligned}$$

Thus, the geometric mean of the root scales, the geometric mean of all successive switching points, and the exponential of the mean value of $\ln x$ for the normalized response coincide.

Let

$$s = \ln x.$$

Since $\ln Z_m$ is the cumulant-generating function of the occupancy,

$$\begin{aligned} \frac{d \ln Z_m}{ds} &= \kappa_1 = \langle N \rangle, \\ \frac{d^2 \ln Z_m}{ds^2} &= \kappa_2 = \text{Var}(N), \\ \frac{d \text{Var}(N)}{ds} &= \kappa_3, \end{aligned}$$

$$\frac{d^2 \text{Var}(N)}{ds^2} = \kappa_4.$$

Consequently, any interior stationary point of the equilibrium response satisfies

$$\kappa_3 = 0,$$

and a strict nondegenerate local maximum additionally satisfies

$$\kappa_4 < 0.$$

In the fully real regime,

$$\begin{aligned}\kappa_3 &= \sum_{k=1}^m \pi_k (1 - \pi_k) (1 - 2\pi_k), \\ \kappa_4 &= \sum_{k=1}^m \pi_k (1 - \pi_k) [1 - 6\pi_k (1 - \pi_k)],\end{aligned}$$

where

$$\pi_k = \frac{x}{x + r_k}.$$

For the central CO₂ parameterization, the stationary point of the response is located at

$$x_{resp} = 14.6060570636,$$

where

$$\begin{aligned}\langle N \rangle &= 5.47855549124, \\ \text{Var}(N) &= 2.41896344326, \\ \kappa_3 &= 3.69 \cdot 10^{-15}, \\ \kappa_4 &= -0.69816620714.\end{aligned}$$

For this parameterization, uniqueness of the maximum is established algebraically. With an exact rational representation of the decimal value used,

$$\alpha = \frac{28771}{479100}$$

the condition

$$\kappa_3(x) = 0$$

after reduction to a common denominator has an integer numerator of degree 41. Exact isolation of its real roots shows that there is exactly one root on the positive half-axis.

For $x > 0$,

$$\text{Var}(N) > 0,$$

and at both ends of the physical half-axis,

$$\begin{aligned}\text{Var}(N) &\rightarrow 0 \text{ as } x \rightarrow 0, \\ \text{Var}(N) &\rightarrow 0 \text{ as } x \rightarrow \infty.\end{aligned}$$

Therefore, the unique positive stationary point is simultaneously the unique global maximum:

$$x_{resp} = 14.6060570636.$$

For the algebraic root count, a rational representation of the specified numerical value of α is used. The physical precision of β and V_c is unchanged by this procedure.

From the equality

$$\frac{\partial \ln P_n}{\partial \ln x} = n - \langle N \rangle$$

it follows that the Fisher information for the parameter $\ln x$ is

$$I_F(\ln x) = \text{Var}(\langle N \rangle).$$

4.2. Distribution invariants with respect to activity

For $1 \leq n \leq m-1$, define

$$I_n = \frac{P_{n+1} P_{n-1}}{P_n^2}.$$

Substitution of $P_n = a_n x^n / Z$ cancels the activity and normalization completely:

$$I_n = \frac{a_{n+1} a_{n-1}}{a_n^2}.$$

For the free-volume weight,

$$I_n = \frac{\left[\frac{n}{(n+1)} \right] \cdot \left[(1 - (n+1)\alpha)^{n+1} (1 - (n-1)\alpha)^{n-1} \right]}{\left[(1 - n\alpha)^{2n} \right]}.$$

Therefore, for a fixed cell each I_n is constant as x varies, while the entire sequence

$$I_1, \dots, I_{m-1}$$

is determined by the single parameter α . Constancy of I_n tests invariance of the coefficients of the finite partition function, while agreement of the entire sequence with the stated formula tests specifically the free-volume weight used here.

4.3. Unique reconstruction of α and x from a single distribution

Let $m \geq 2$ and suppose that at one unknown activity the probabilities P_0 , P_1 , and P_2 are measured, with all three probabilities positive. Define

$$R_1 = \frac{P_0 P_2}{P_1^2} = \frac{a_2}{a_1^2},$$

$$s_1 = \sqrt{2 R_1}.$$

For the weight considered here,

$$s_1 = \frac{(1 - 2\alpha)}{(1 - \alpha)},$$

from which, in the region $0 < \alpha < 1/2$, it follows uniquely that

$$\alpha = \frac{(1 - s_1)}{(2 - s_1)}.$$

After α has been reconstructed, the activity is obtained from the ratio of the first two probabilities:

$$\frac{P_1}{P_0} = a_1 x = (1 - \alpha) x,$$

$$x = \frac{P_1}{[(1 - \alpha) P_0]}.$$

Thus, for known capacity m , the probabilities P_0 , P_1 , and P_2 uniquely determine the two continuous model parameters α and x . Thereafter,

$$P_n^{\text{pred}} = \frac{\left[\frac{(1 - n\alpha)^n x^n}{n!} \right]}{\left[\frac{\sum_{j=0}^m (1 - j\alpha)^j x^j}{j!} \right]}.$$

The probabilities $P_3, \dots, P_m$ no longer contain new fitting parameters and become parameter-free predictions of the model. Their joint agreement with the observed distribution tests the functional form of the statistical weight $a_n(\alpha)$ over the entire sequence of occupancy states.

The parameter α can also be obtained from the initial isotherm curvature. For $x \rightarrow 0$,

$$\langle N \rangle = (1 - \alpha)x - \alpha(2 - 3\alpha)x^2 + O(x^3).$$

Define the dimensionless indicator

$$c_{\text{low}} = \frac{-\langle N \rangle''(0)}{[2\langle N \rangle'(0)]^2}.$$

Then

$$c_{\text{low}} = \frac{\alpha(2 - 3\alpha)}{(1 - \alpha)^2},$$

and on the physical branch,

$$\alpha = \frac{[1 + c_{\text{low}} - \sqrt{1 - c_{\text{low}}}]^2}{(3 + c_{\text{low}})}.$$

The activity scale $K(T)$ cancels in this ratio.

Let the absolute loading be related to the mean occupancy by

$$q = c_{\text{ads}} \langle N \rangle,$$

where c_{ads} is the amount of adsorbate per unit mass corresponding to an increase of the mean cell occupancy by one molecule. Then

$$c_{\text{low}} = \frac{-c_{\text{ads}} q''(0)}{[2q'(0)]^2},$$

where the derivatives are taken with respect to the same physical variable y used in the consistent parameterization

$$x = K(T)y.$$

If the saturation loading is known,

$$q_{sat} = mc_{ads},$$

then equivalently

$$c_{low} = \frac{-\left(\frac{q_{sat}}{m}\right)q''(0)}{[2q'(0)^2]}.$$

Consequently, α can be reconstructed from the first two coefficients of the low-activity expansion of the mean isotherm without prior knowledge of $K(T)$ and without measuring the full distribution P_n . In the fugacity parameterization $y = f$, whereas in the pressure parameterization $y = p$.

For experimental reconstruction of α , it is necessary to identify a low-activity interval in which the data permit stable determination of the linear and quadratic terms of the expansion

$$q(y) = h_1 y + h_2 y^2 + O(y^3).$$

Then

$$c_{low} = \frac{-c_{ads} h_2}{h_1^2}.$$

Direct estimation of $q''(0)$ from the first few noisy points is numerically unstable; therefore, h_1 and h_2 should be determined by a joint fit over the selected low-activity interval. The resulting α should remain stable under a reasonable variation of the upper boundary of this interval.

Such a reconstruction of α is determined by the local shape of the mean isotherm: $K(T)$ cancels, and the full-distribution data P_n do not enter the formula. For the 48-point NaY–CO₂ data set, this relation serves as a separate low-activity diagnostic, whereas the root diagram is constructed using the central value $\alpha = \beta/V_c$.

4.4. Renormalization of the distribution and temperature superposition

If, at x_1 , the distribution $P_n(x_1)$ is known, then for the same finite cell

$$P_n(x_2) = \frac{P_n(x_1) \left(\frac{x_2}{x_1}\right)^n}{\left[\sum_{j=0}^m P_j(x_1) \left(\frac{x_2}{x_1}\right)^j \right]}.$$

Since the activity dependence enters the statistical weight only through the factor x^n , this relation follows from the same set of coefficients a_n and contains no new parameters. Therefore, distributions measured at different activities provide a direct test of the invariance of the cell's statistical weights.

If

$$x = K_0 \exp \left[\frac{E}{(RT)} \right] y,$$

then, for unchanged m and α , two states with the same x correspond to the same distribution P_n . At different x , the distributions are related by the renormalization given above.

For the mean response this gives an even simpler law. Denote

$$s_y = \ln y.$$

Then

$$s = \ln x = s_y + \ln K(T),$$

and therefore

$$\rho_y(s_y; T) = \rho_x[s_y + \ln K(T)].$$

The temperature dependence $K(T)$ shifts the entire response along the $\ln y$ axis without changing its internal shape. In particular,

$$\text{mean}(\ln y) = \bar{u} - \ln K(T),$$

whereas all central cumulants of order two and higher are preserved.

Each root r_k defines its own scale of the chosen physical variable

$$y_k(T) = \frac{r_k}{K(T)}.$$

For

$$K(T) = K_0 \exp \left[\frac{E}{(RT)} \right]$$

we have

$$\ln y_k(T) = \ln \left(\frac{r_k}{K_0} \right) - \frac{E}{(RT)},$$

and therefore

$$\frac{d \ln y_k}{d \left(\frac{1}{T} \right)} = \frac{-E}{R}$$

for all k .

Different r_k define different characteristic occupancy scales. For unchanged m and α , all corresponding scales $y_k(T)$ have the same temperature slope $-E/R$. Thus, the multiscale root spectrum is compatible with a single energy parameter E governing the temperature dependence: the configurational occupancy scales differ, whereas their temperature dependence is common.

For constant m and α , the function $K(T)$ only shifts the normalized profile ρ along $\ln y$, without changing its width, skewness, or any other central cumulants.

4.5. Motion of the zeros of the generating function

From $G_x(z) = Z(xz)/Z(x)$ it follows that

$$z_k(x) = \frac{\xi_k}{x}.$$

As the activity changes, each zero of the generating function moves along a fixed ray toward the origin of the complex plane: its argument remains constant, while its modulus scales as $1/x$.

This property holds for any fixed set of coefficients a_n . Therefore, zeros of generating functions reconstructed at different activities must satisfy the same radial law

$$z_k(x) = \frac{\xi_k}{x}.$$

A change in their arguments or a deviation from the $1/x$ scaling means that the two distributions cannot be described by one invariant set of coefficients a_n .

4.6. Low- and high-activity moments of the root spectrum

In the regime $C = 0$,

$$\langle N \rangle = \frac{\sum_{k=1}^m x}{(x + r_k)}.$$

Let the roots be ordered so that

$$0 < r_1 \leq r_2 \leq \dots \leq r_m.$$

For each factor, when $|x| < r_k$, the convergent geometric series holds

$$\frac{x}{(x + r_k)} = \frac{x}{r_k} - \frac{x^2}{r_k^2} + \frac{x^3}{r_k^3} - \dots$$

Therefore, for

$$|x| < r_1$$

the entire sum has the convergent expansion

$$\langle N \rangle = S_{-1}x - S_{-2}x^2 + S_{-3}x^3 - \dots,$$

where

$$S_{-q} = \sum_{k=1}^m r_k^{-q}.$$

In particular,

$$S_{-1} = \sum_{k=1}^m \frac{1}{r_k} = a_1 = 1 - \alpha.$$

Similarly, when $|x| > r_k$

$$\frac{x}{(x + r_k)} = 1 - \frac{r_k}{x} + \frac{r_k^2}{x^2} - \frac{r_k^3}{x^3} + \dots$$

Consequently, for

$$|x| > r_m$$

the high-activity series converges,

$$m - \langle N \rangle = \frac{S_1}{x} - \frac{S_2}{x^2} + \frac{S_3}{x^3} - \dots,$$

where

$$S_q = \sum_{k=1}^m r_k^q.$$

Thus, the two edges of the isotherm are determined by two families of moments of the same root spectrum: inverse power sums at low activity and direct power sums as saturation is approached.

If

$$q = c_{ads} \langle N \rangle,$$

$$x = K(T)y,$$

then the low-activity expansion takes the form

$$q(y, T) = c_{ads} S_{-1} K(T) y - c_{ads} S_{-2} K(T)^2 y^2 + O(y^3).$$

Therefore, the initial slope with respect to the physical variable y is

$$H_y(T) = \left[\frac{dq}{dy} \right]_{y \rightarrow 0} = c_{ads} K(T) S_{-1}.$$

At large y,

$$q_{sat} - q(y, T) = \frac{c_{ads} S_1}{[K(T)y]} - \frac{c_{ads} S_2}{[K(T)^2 y^2]} + O(y^{-3}).$$

Consequently, the coefficient of the first high-activity term

$$A_{\infty}(T) = \lim_{y \rightarrow \infty} y [q_{sat} - q(y, T)]$$

is

$$A_{\infty}(T) = \frac{c_{ads} S_1}{K(T)} \cdot (B)$$

For the coefficients of the partition function,

$$S_{-1} = \frac{\sum_{k=1}^m 1}{r_k} = a_1 = 1 - \alpha,$$

whereas

$$S_1 = \sum_{k=1}^m r_k = \frac{a_{m-1}}{a_m} = \frac{m [1 - (m-1)\alpha]^{m-1}}{(1 - m\alpha)^m}.$$

These quantities are directly related to the extreme neighboring switches of the most probable occupancy. Since

$$x_n = \frac{a_n}{a_{n+1}},$$

we have

$$S_{-1} = \frac{1}{x_0},$$

$$S_1 = x_{m-1}.$$

Thus, the first coefficient at the low-activity edge is determined by the first switching point $P_0 = P_1$, while the first coefficient of the approach to saturation is determined by the final switching point $P_{m-1} = P_m$.

Equations (A) and (B) imply the temperature-invariant product

$$H_y(T) A_{\infty}(T) = c_{ads}^2 S_{-1} S_1, (C)$$

if m , α , and c_{ads} do not change with temperature. The function $K(T)$ shifts the low- and high-activity edges in opposite directions but cancels completely in their product.

Similarly, if

$$y_{resp}(T) = \frac{x_{resp}}{K(T)}$$

denotes the position of the response maximum in the same physical variable, then

$$H_y(T) y_{resp}(T) = c_{ads} S_{-1} x_{resp}. (D)$$

For the central CO₂ parameterization,

$$S_{-1} = 0.939947818827,$$

$$S_1 = 5627.71656187,$$

$$x_{resp} = 14.6060570636,$$

$$c_{ads} \approx 0.6204 \text{ mmol/g},$$

and therefore

$$H_y(T) y_{resp}(T) \approx 8.52 \text{ mmol/g}.$$

Relations (C) and (D) connect the two edges of the mean isotherm and the position of its maximum equilibrium response with the same root spectrum, without explicit dependence on the temperature-dependent factor $K(T)$.

The convergence domains are essential here. The low-activity series for the entire cell is limited by the nearest root r_1 , while the high-activity series is limited by the most distant root r_m . Saturation of only the first few root contributions does not yet imply that the series in $1/x$ converges for the full response.

5. Root cascade for capacity $m = 14$

5.1. Complete root cascade for $m = 14$

Introduce a polynomial with integer coefficients in x and α :

$$P_{14}(x, \alpha) = 14! Z_{14}(x; \alpha).$$

A multiple root in x is possible only when

$$D_{14}(\alpha) = \text{Res}_x \left(P_{14}, \frac{\partial P_{14}}{\partial x} \right) = 0.$$

Exact symbolic computation gives

$$D_{14}(\alpha) = A (14\alpha - 1)^{14} Q_{182}(\alpha),$$

where A is a nonzero integer constant, and Q_{182} is a primitive integer polynomial of degree 182. The factor $(14\alpha - 1)^{14}$ arises when the leading coefficient vanishes at the outer boundary $\alpha = 1/14$. Internal collisions are determined by the roots of Q_{182} .

The polynomial Q_{182} is square-free:

$$\gcd(Q_{182}, Q_{182}') = 1.$$

Exact rational isolation of all its real roots yields 14 disjoint intervals, each containing one simple root. Seven roots lie within

$$0 < \alpha < \frac{1}{14},$$

and seven more lie at $\alpha > 1/14$; for $\alpha \leq 0$, Q_{182} has no real roots.

Consequently, the physical interval contains exactly seven internal values of α at which Z_{14} has a multiple root. Each corresponding double root lies on the negative real axis.

Table 1. Complete cascade of boundaries for $m = 14$

N _o	α_c	$\eta_c = 14\alpha_c$	$y_c = -x_c$	C: before → after
1	0.0494260239663510	0.691964335528914	7.621923657224	7 → 6
2	0.0501166836198449	0.701633570677829	8.490256373172	6 → 5
3	0.0506824974743869	0.709554964641416	10.618975956400	5 → 4

N ₀	α_c	$\eta_c = 14\alpha_c$	$y_c = -x_c$	C: before → after
4	0.0513350543290141	0.718690760606197	15.294663125488	4 → 3
5	0.0521906921010993	0.730669689415391	26.935767220647	3 → 2
6	0.0534205930840320	0.747888303176448	67.194094619824	2 → 1
7	0.0554547158587412	0.776366022022377	392.673625281457	1 → 0

To determine the number of real zeros between successive values, it is sufficient to choose one rational test point in each interval. The exact count gives

$$\alpha = 0.0490 : 0 \text{ real roots, } C = 7;$$

$$\alpha = 0.0497 : 2 \text{ real roots, } C = 6;$$

$$\alpha = 0.0504 : 4 \text{ real roots, } C = 5;$$

$$\alpha = 0.0510 : 6 \text{ real roots, } C = 4;$$

$$\alpha = 0.0517 : 8 \text{ real roots, } C = 3;$$

$$\alpha = 0.0528 : 10 \text{ real roots, } C = 2;$$

$$\alpha = 0.0544 : 12 \text{ real roots, } C = 1;$$

$$\alpha = 0.0600 : 14 \text{ real roots, } C = 0.$$

Between neighboring roots Q_{182} , there are no multiple zeros Z_{14} ; therefore, these counts are constant throughout the corresponding interval.

At the collision point itself, the double real root is counted with multiplicity two. Thus, the complete classification is

$$0 < \alpha < 0.0494260239663510 : C = 7;$$

$$0.0494260239663510 \leq \alpha < 0.0501166836198449 : C = 6;$$

$$0.0501166836198449 \leq \alpha < 0.0506824974743869 : C = 5;$$

$$0.0506824974743869 \leq \alpha < 0.0513350543290141 : C = 4;$$

$$0.0513350543290141 \leq \alpha < 0.0521906921010993 : C = 3;$$

$$0.0521906921010993 \leq \alpha < 0.0534205930840320 : C = 2;$$

$$0.0534205930840320 \leq \alpha < 0.0554547158587412 : C = 1;$$

$$0.0554547158587412 \leq \alpha < \frac{1}{14} : C = 0.$$

Hence, the seven roots Q_{182} inside the physical interval exhaust all internal boundaries of the root structure of Z_{14} .

5.2. Local square-root law

Set $\delta\alpha = \alpha - \alpha_c$ and $\delta x = x - x_c$. For each of the seven boundaries, simplicity of the corresponding resultant root establishes that the collision is pairwise and transverse: Z_α and Z_{xx} are nonzero at the boundary point. Expansion of $Z = 0$ therefore gives

$$\delta x^2 = -2 \left(\frac{Z_\alpha}{Z_{xx}} \right) \delta\alpha + \text{higher-order terms},$$

that is,

$$\delta x \text{ proportional to } \pm \sqrt{\delta\alpha}$$

on the side with two real zeros, and

$$Imx \text{ proportional to } \sqrt{-\delta\alpha}$$

on the side with a complex pair, with the sign of the derivative ratio taken into account.

The square-root splitting characterizes the local rearrangement of the root geometry of Z_{14} as α varies. For physical activity $x > 0$,

$$Z_{14}(x; \alpha) > 0,$$

therefore, the mean occupancy and its derivatives with respect to x remain analytic when the identified α values are crossed. Thus, the root collision is a structural transition in the complex x plane but does not create a singularity of the equilibrium isotherm on the positive real axis. Below, these α values are referred to as boundaries of the root structure.

5.3. Boundary of the fully real spectrum

For fixed m and V_c , the parameter β must simultaneously satisfy the condition of positive free volume

$$0 < \beta < \frac{V_c}{m}.$$

For $m = 14$, the final root boundary is

$$\beta_c = V_c \alpha_c.$$

Because the colliding pair is already a double real root at the boundary, the region $C = 0$ is

$$\beta_c \leq \beta < \frac{V_c}{14}.$$

For $V_c = 958.2 \text{ \AA}^3$,

$$\beta_c \approx 53.14 \text{ \AA}^3,$$

$$\frac{V_c}{14} \approx 68.44 \text{ \AA}^3.$$

Thus, for $m = 14$, the fully real spectrum corresponds to

$$53.14 \text{ \AA}^3 \leq \beta < 68.44 \text{ \AA}^3$$

for the nominal value $V_c = 958.2 \text{ \AA}^3$.

The value α_c is determined by the specified polynomial and is reported to the numerical precision of the calculation. For the dimensional quantities β_c and $V_c/14$, the number of physically significant digits is set by the precision of the original V_c .

5.4. Near-boundary large-capacity limit and the partial theta function

5.4.1. Near-boundary scaling and reversed polynomial

The computed final boundaries for $m = 2, \dots, 30$ show that

$$\eta_c(m) = m \alpha_c(m)$$

approaches unity as m increases. To analyze this region, introduce

$$\eta = m\alpha,$$

$$\delta = 1 - \eta.$$

Taking into account

$$\alpha = \frac{\beta}{V_c}$$

we obtain

$$\delta = 1 - \frac{m\beta}{V_c} = \frac{[V_c - m\beta]}{V_c} = \frac{V_{free}(m)}{V_c}.$$

Consequently, δ has a direct physical meaning: it is the fraction of free volume remaining in the maximally occupied state $n = m$.

The condition that all statistical weights be positive

$$0 < \alpha < \frac{1}{m}$$

is equivalent to

$$\begin{aligned} 0 < \eta < 1, \\ \delta > 0. \end{aligned}$$

The limit $m \rightarrow \infty$ considered here refers to a sequence of finite systems approaching the upper boundary $\eta = 1$. It does not imply increasing the capacity of one fixed cell at unchanged α .

The numerical sequence of the final boundaries shows that

$$\delta_c(m) = 1 - \eta_c(m)$$

tends to zero and indicates a scale of order $m^{-1/2}$. Below, this scale is derived directly from the coefficients of the reversed polynomial, while the numerical sequence provides an independent indication of the same near-boundary regime.

To isolate the contribution of states close to maximum occupancy, set

$$n = m - k,$$

where k remains finite as $m \rightarrow \infty$, and reverse the polynomial with respect to its highest degree. Introduce the variable

$$z = \frac{\left(\frac{a_{m-1}}{a_m} \right)}{x}.$$

Then

$$\frac{Z_m(x; \alpha)}{(a_m x^m)} = \sum_{k=0}^m R_{m,k} z^k,$$

where

$$R_{m,k} = \frac{a_{m-k} a_m^{k-1}}{a_{m-1}^k}.$$

This normalization gives

$$R_{m,0} = R_{m,1} = 1,$$

and the index k directly measures the distance from the maximally occupied state.

Denote

$$t = \frac{(1-\delta)}{(m\delta)}.$$

For the free-volume coefficients,

$$a_n = \frac{(1-n\alpha)^n}{n!}$$

cancellation of powers of δ and factorial factors gives the exact expression

$$R_{m,k} = \left[\prod_{j=1}^{k-1} \left(1 - \frac{j}{m} \right) \right] \cdot \frac{(1+kt)^{m-k}}{(1+t)^{k(m-1)}}.$$

This formula itself determines the natural scale of the near-boundary limit. For $\delta \rightarrow 0$

$$t \sim \frac{1}{(m\delta)}.$$

For fixed k and $t \rightarrow 0$, the logarithm of $R_{m,k}$ has the structure

$$\ln R_{m,k} = - \left[\frac{k(k-1)}{2} \right] m t^2 + O(t) + O(m t^3) + O\left(\frac{1}{m}\right).$$

For the coefficients $R_{m,k}$ to have a nontrivial finite limit, the quantity $m t^2$ must remain of order unity. Since

$$m t^2 \sim \frac{1}{(m \delta^2)},$$

we obtain the required near-boundary scale

$$m \delta^2 = O(1),$$

that is,

$$\delta = \frac{\gamma}{\sqrt{m}},$$

where γ remains finite as $m \rightarrow \infty$.

Set

$$\varepsilon = \frac{1}{\sqrt{m}},$$

$$\delta = \gamma \varepsilon.$$

Then

$$t = \frac{\varepsilon}{\gamma} - \varepsilon^2.$$

5.4.2. Coefficient limit and the partial theta function

For each fixed k , the logarithmic expansion of the exact expression for $R_{m,k}$ gives

$$\ln R_{m,k} = \frac{-k(k-1)}{(2\gamma^2)} + \left[\frac{(k^3-k)}{(3\gamma^3)} \right] m^{-1/2} + O(m^{-1}).$$

Therefore,

$$R_{m,k} = \exp \left[\frac{-k(k-1)}{(2\gamma^2)} \right] \left[1 + O(m^{-1/2}) \right].$$

Introduce

$$q = \exp \left(\frac{-1}{\gamma^2} \right).$$

Then the coefficient limit has the form

$$\exp \left[\frac{-k(k-1)}{(2\gamma^2)} \right] = q^{k(k-1)/2},$$

which leads to the limiting series

$$\Theta(q, z) = \sum_{k=0}^{\infty} q^{k(k-1)/2} z^k.$$

This is the partial theta function in the normalization used here.

At this stage, coefficient convergence has been established for each fixed k . This is not sufficient for continuation of the double root: the entire coefficient tail must be controlled locally uniformly, together with the derivatives with respect to z and γ that enter the double-root system.

5.4.3. Uniform control of the coefficient tail

Denote

$$F_m(\gamma, z) = \sum_{k=0}^m R_{m,k} z^k.$$

For γ in any fixed compact interval around γ_1 and sufficiently large m , the quantity t is positive. From the exact formula for $R_{m,k}$ we obtain

$$\ln R_{m,k} = \sum_{j=1}^{k-1} \ln \left(1 - \frac{j}{m} \right) - m D_k(t) - k [\ln(1+kt) - \ln(1+t)],$$

where

$$D_k(t) = k \ln(1+t) - \ln(1+kt).$$

This quantity admits the exact integral representation

$$D_k(t) = k(k-1) \int_0^t \frac{u du}{(1+u)(1+ku)}.$$

Consequently,

$$D_k(t) > 0$$

for $k > 1$.

If $k \leq A\sqrt{m}$, then kt remains bounded, while mt^2 is separated from zero. Therefore, for some constant $c > 0$ independent of m and k ,

$$R_{m,k} \leq \exp[-ck(k-1)].$$

If $k > A\sqrt{m}$, then for sufficiently large m

$$\ln(1+t) \geq \frac{t}{2}$$

and

$$D_k(t) \geq \frac{kt}{2} - \ln(1+kt).$$

For any bounded set $|z| \leq R$, one can choose A sufficiently large so that the remaining part of the sum, after multiplication by $|z|^k$, decays exponentially in m . The same estimates remain valid after multiplication of the terms by k and $k(k-1)$, and therefore simultaneously control the first and second derivatives with respect to z .

These estimates can be strengthened to the uniform expansion required for the double-root system.

Fix a compact interval of γ values containing γ_1 and a bounded region

$$|z| \leq R.$$

Set

$$K_m = A\sqrt{\ln m},$$

where $A > 0$ is chosen sufficiently large.

For

$$0 \leq k \leq K_m$$

we have

$$\frac{k}{\sqrt{m}} \rightarrow 0$$

uniformly; therefore, expansion of the exact formula for $R_{m,k}$ gives

$$\ln R_{m,k} = \frac{-k(k-1)}{(2\gamma^2)} + \left[\frac{(k^3-k)}{(3\gamma^3)} \right] m^{-1/2} + O[m^{-1}P(k)],$$

where $P(k)$ is a polynomial of fixed degree and the remainder is uniform in γ over the chosen compact interval.

After exponentiation,

$$R_{m,k} = q^{k(k-1)/2} 1 + \left[\frac{(k^3-k)}{(3\gamma^3)} \right] m^{-1/2} + O[m^{-1}P_1(k)].$$

The Gaussian bound

$$R_{m,k} \leq \exp[-ck(k-1)]$$

majorizes the resulting series after multiplication by any fixed powers of k . It therefore controls not only F_m , $F_{m,z}$, and $F_{m,zz}$, but also the derivatives with respect to γ arising from differentiation of the coefficients $R_{m,k}$.

For

$$k > K_m$$

the same Gaussian tail, after an appropriate choice of A , is of order $o(m^{-1})$ on every bounded set in z ; the corresponding estimates remain valid for $F_{m,z}$, $F_{m,\gamma}$, and $F_{m,z\gamma}$.

Consequently,

$$F_m(\gamma, z) = \Theta(q(\gamma), z) + m^{-1/2} H(\gamma, z) + O(m^{-1})$$

locally uniformly together with the derivatives entering the system

$$F_m = 0,$$

$$F_{m,z} = 0,$$

where

$$H(\gamma, z) = \left[\frac{1}{(3\gamma^3)} \right] \sum_{k=0}^{\infty} (k^3 - k) q^{k(k-1)/2} z^k.$$

Thus, the passage to the partial theta function is carried out at the level of the double-root system itself, rather than merely pointwise in the coefficients.

5.4.4. Limiting double point and continuation of the finite-dimensional branch

For the partial theta function, the first positive spectral point is [5–7]

$$q_1 = 0.309249338600077 \dots,$$

and the corresponding double root in the variable z used here is

$$z_1 = -2.32037694428960 \dots$$

For $q < q_1$, the roots of the partial theta function are real and negative. At $q = q_1$, the two outermost real roots merge.

Since

$$q(\gamma) = \exp\left(\frac{-1}{\gamma^2}\right),$$

we obtain

$$\gamma_1 = \frac{1}{\sqrt{-\ln q_1}} = 0.923078379994697 \dots$$

At the double point,

$$\Theta(q_1, z_1) = 0,$$

$$\Theta_z(q_1, z_1) = 0,$$

$$\Theta_q(q_1, z_1) = 2.27449529399109 \dots,$$

$$\Theta_{zz}(q_1, z_1) = 0.261280341217975 \dots$$

In addition,

$$\frac{dq}{d\gamma} = \frac{2q}{\gamma^3} > 0.$$

Therefore, the Jacobian of the system

$$\begin{aligned}\Theta &= 0, \\ \Theta_z &= 0\end{aligned}$$

with respect to the variables (γ, z) is

$$J = \Theta_q \left(\frac{dq}{d\gamma} \right) \Theta_{zz} = 0.467320480630847 \dots \neq 0.$$

The double point is nondegenerate. Local C^1 convergence of the mappings

$$\Phi_m = \left(F_m, \frac{\partial F_m}{\partial z} \right)$$

to $\Phi = (\Theta, \Theta_z)$

and the nonzero Jacobian of the limiting system at the point (γ_1, z_1) allow application of the implicit function theorem. Hence, for sufficiently large m , there exists a unique branch in a neighborhood of the limiting double point,

$$\begin{aligned}\gamma_m^{(*)} &= \gamma_1 + O(m^{-1/2}), \\ z_m^{(*)} &= z_1 + O(m^{-1/2}).\end{aligned}$$

It follows that

$$\begin{aligned}\eta_m^{(*)} &= 1 - \frac{\gamma_1}{\sqrt{m}} + O\left(\frac{1}{m}\right), \\ \alpha_m^{(*)} &= \frac{1}{m} - \frac{\gamma_1}{m^{3/2}} + O\left(\frac{1}{m^2}\right).\end{aligned}$$

5.4.5. First correction term

The first correction term is determined from the linearization of the system

$$\begin{aligned}F_m(\gamma, z) &= 0, \\ F_{m,z}(\gamma, z) &= 0\end{aligned}$$

at the point (γ_1, z_1) .

Set

$$\begin{aligned}\gamma_m^{(*)} &= \gamma_1 + \frac{a_1}{\sqrt{m}} + O\left(\frac{1}{m}\right), \\ z_m^{(*)} &= z_1 + \frac{b_1}{\sqrt{m}} + O\left(\frac{1}{m}\right).\end{aligned}$$

At the limiting double point,

$$H = 1.09343809580966 \dots,$$

$$\Theta_y = \Theta_q \left(\frac{dq}{dy} \right) = 1.78857880563231 \dots,$$

therefore

$$a_1 = \frac{-H}{\Theta_y} = -0.611344656643799 \dots$$

For the second equation,

$$H_z = 0.156272638168742 \dots,$$

$$\Theta_{zy} = -0.642884978437576 \dots,$$

$$\Theta_{zz} = 0.261280341217975 \dots,$$

whence

$$b_1 = \frac{-[\Theta_{zy} a_1 + H_z]}{\Theta_{zz}} = -2.10232783688407 \dots$$

Consequently,

$$y_m^{(*)} = y_1 - \frac{0.611344656644}{\sqrt{m}} + O\left(\frac{1}{m}\right),$$

$$z_m^{(*)} = z_1 - \frac{2.10232783688}{\sqrt{m}} + O\left(\frac{1}{m}\right),$$

and

$$\eta_m^{(*)} = 1 - \frac{0.923078379995}{\sqrt{m}} + \frac{0.611344656644}{m} + O(m^{-3/2}).$$

Thus, the limiting double point determines the asymptotics of both coordinates: the boundary position and the corresponding normalized double root.

5.4.6. Relation to the standard normalization and comparison with finite m

The conventional form of the partial theta function

$$\theta(q, y) = \sum_{k=0}^{\infty} q^{k(k+1)/2} y^k$$

is related to the form used here by the substitution

$$z = qy.$$

Therefore,

$$y_1 = \frac{z_1}{q_1} \approx -7.503256.$$

The final boundaries computed to high precision for $m = 2, \dots, 30$ give

$$y_m = [1 - \eta_c(m)] \sqrt{m}$$

with the values

m	y_m
2	0.773459080339
5	0.812247294068

m	γ_m
10	0.829446949484
14	0.836761725634
20	0.844334152505
30	0.852825638266

and move toward γ_1 from the side and on the scale implied by the derived asymptotics.

The passage to the partial theta function and the asymptotics of the local branch of double roots are established analytically. For the computed sequence $m = 2, \dots, 30$, this branch coincides with the final boundary of the fully real spectrum. Preservation of this coincidence for arbitrarily large m requires separate global control of the remaining roots and does not follow from local convergence alone.

6. Numerical results for CO₂ and CH₄ in NaY

6.1. NaY parameters and robustness of the root classification

For the numerical analysis, the SSI parameters for single-component adsorption of CO₂ and CH₄ on NaY NIST RM 8850 reported by Wedler et al. [2] are used.

The volume of one faujasite-framework supercage is taken to be

$$V_c = 958.2 \text{ \AA}^3,$$

and the integer limiting capacity for both adsorbates is

$$m = 14.$$

At fixed capacity, the root spectrum is determined only by the pair

$$(m, \alpha),$$

where

$$\alpha = \frac{\beta}{V_c}.$$

The parameters K_0 and E do not enter the root equation. They determine the temperature dependence $K(T)$ and the mapping between the dimensionless activity and the physical variable of the chosen parameterization,

$$x = K(T)y,$$

where $y = p$ for the original pressure representation and $y = f$ for the fugacity representation.

To assess the robustness of the CO₂ root classification, the parameters K_0 , E , and β were re-estimated from the same 48 experimental points. Unweighted and weighted error functions were considered in the original pressure parameterization, as well as a consistent fugacity-based treatment using the Peng–Robinson equation of state [8] with joint re-estimation of all three parameters. Comparison of these variants shows how strongly the position of α relative to the final root boundary depends on the treatment of the same data set.

Table 2. Model parameters and position relative to the final root boundary

Quantity	CO₂	CH₄
Capacity m	14	14

Additional digits in the calculated dimensionless quantities are retained for reproducibility of the calculation. The number of physically significant digits is determined by the precision of the original V_c and β .

For the central CO₂ parameterization,

$$\alpha = 0.0600521811730,$$

whereas the final root boundary is

$$\alpha_c = 0.0554547158587.$$

Re-estimation of the parameters under different treatments of the original data set gave

$$0.0595675422 \leq \alpha \leq 0.0600536347.$$

Throughout this range,

$$C = 0,$$

and the minimum distance to the final boundary is

$$\alpha - \alpha_c = 0.00411282635.$$

The position of the response maximum varies only within

$$14.6057 \leq x_{resp} \leq 14.7173.$$

The parameters, residuals, and corresponding root spectra for these variants are given in the supplementary computational material.

The root classification is specified by the pair (m, α) , because changing m changes both the degree of the polynomial and the sequence of its coefficients. For the central value α_{CO_2} , the result for neighboring capacities is

m	$m\alpha_{CO_2}$	$\alpha_c(m)$	C
12	0.720626174076	0.063283805318	2
13	0.780678355249	0.059105008517	0
14	0.840730536422	0.055454715859	0
15	0.900782717595	0.052237900861	0
16	0.960834898769	0.049381101090	0

At $m = 17$, the condition $1 - m\alpha > 0$ is already violated, so positivity of all statistical weights a_n , $n = 0, \dots, 17$, is no longer preserved. For the central value α_{CO_2} , the $C = 0$ regime persists for $m = 13, \dots, 16$. Thus, the root classification is determined jointly by m and α .

For $m = 14$ and the nominal $V_c = 958.2 \text{ \AA}^3$, the final root boundary corresponds to

$$\beta_c = V_c \alpha_c \approx 53.14 \text{ \AA}^3.$$

For the central CO₂ parameterization,

$$\beta = 57.542 \text{ \AA}^3.$$

With V_c unchanged, reaching the final boundary corresponds to a decrease in β by approximately

$$4.41 \text{ \AA}^3,$$

or approximately 7.7%.

With β unchanged, the same boundary is obtained at

$$V_c \approx 1.038 \cdot 10^3 \text{ \AA}^3,$$

which is approximately 8.3% larger than the adopted value.

These values show how much β or V_c must change, with the remaining parameters fixed, for the central parameterization to reach the final root boundary. They therefore characterize the sensitivity of the root class within the adopted model, but they are not an estimate of the total physical uncertainty of the material. The latter includes uncertainty in V_c , the choice of integer capacity m , and the applicability of the free-volume description itself.

To convert the mean number of molecules in a cell to absolute loading, introduce

$$q = c_{ads} \langle N \rangle.$$

For NIST RM 8850, the micropore volume used is

$$v_{mi} = 0.358 \text{ cm}^3/\text{g}[9].$$

For $V_c = 958.2 \text{ \AA}^3$,

$$c_{ads} = \frac{1000 v_{mi}}{V_c N_A \cdot 10^{-24}} \approx 0.6204 \text{ mmol/g}.$$

per unit increase in mean cell occupancy.

The unrounded value of c_{ads} is used in the numerical calculations, whereas the number of reported significant digits is limited by the precision of the original v_{mi} and V_c .

The quantity c_{ads} sets only the vertical scale of the absolute isotherm. The roots Z_m , the parameter α , and the normalized distribution P_n do not depend on it.

6.2. Complete spectrum of zeros

Table 3. Negative real zeros $x = -r_k$

k	r_k for CO ₂	r_k for CH ₄
1	6.23187905420	6.15196905011
2	6.56317902327	6.48801962608
3	7.16445511061	7.09932805775
4	8.12360105248	8.07801014789
5	9.59907314524	9.59127716514
6	11.8767706735	11.9436538430
7	15.4933154916	15.7137229685
8	21.5232814993	22.0784555868
9	32.3179031367	33.6665366355
10	53.6654676788	57.1299309716
11	102.289044059	112.436675498
12	238.721971760	276.172307148
13	775.743308045	980.057036217
14	4338.40331214	6792.74997603

Both spectra contain a group of root scales r_k of order 6–20 and a long high-activity tail. Therefore, not all root contributions are saturated over the practical activity range. The high-activity expansion of the full response in $1/x$ converges only for

$$|x| > r_{14},$$

that is, beyond the scale of the most distant root.

6.3. Reconstruction of coefficients from the roots

For CO_2 , the coefficients computed directly and the coefficients of the product $\prod \left(1 + \frac{x}{r_k}\right)$ agree within numerical precision. Several values are listed in Table 4; the complete sequence contains $n = 0, \dots, 14$.

Table 4. Verification of the CO_2 coefficients

n	a_n from the weight	a_n from the product over roots	relative discrepancy
0	1	1	0
1	0.939947818826967	0.939947818826967	0
2	0.387108166581212	0.387108166581212	0
3	0.0918420467823145	0.0918420467823145	$3.1 \cdot 10^{-16}$
4	0.0138856421013694	0.0138856421013694	$2.5 \cdot 10^{-16}$
6	$9.51639145428830 \cdot 10^{-5}$	$9.51639145428830 \cdot 10^{-5}$	$2.8 \cdot 10^{-16}$
10	$2.85211927004226 \cdot 10^{-11}$	$2.85211927004226 \cdot 10^{-11}$	$4.5 \cdot 10^{-16}$
14	$7.75259132419446 \cdot 10^{-23}$	$7.75259132419443 \cdot 10^{-23}$	$3.6 \cdot 10^{-15}$

The maximum relative discrepancy over all coefficients is $3.64 \cdot 10^{-15}$. For the roots found, the normalized residual of the original polynomial $Z_{14}(\xi_k)$ was additionally monitored; its maximum value does not exceed $7.71 \cdot 10^{-72}$ for CO_2 and $8.38 \cdot 10^{-72}$ for CH_4 in high-precision calculations.

6.4. Numerical verification of the factorization of the full distribution

For CO_2 , the distribution P_n was computed in two ways:

$$\frac{\text{(a) directly from } a_n x^n}{Z};$$

(b) by successive convolution of fourteen Bernoulli distributions

$$\text{with probabilities } \pi_k = \frac{x}{x + r_k}.$$

On a grid of 2401 activity values and for all fifteen values of n , the maximum absolute discrepancy between the probabilities was $8.33 \cdot 10^{-15}$. The mean, variance, third cumulant, and fourth cumulant were compared simultaneously.

Table 5. Numerical verification of the CO_2 factorization

Quantity tested	maximum absolute discrepancy
Coefficients a_n	$2.78 \cdot 10^{-17}$
Complete distribution $P_0, \dots, P_{14}$	$8.33 \cdot 10^{-15}$

Quantity tested	maximum absolute discrepancy
Mean $\langle N \rangle$	$5.51 \cdot 10^{-14}$
Variance $\text{Var}(N)$	$8.44 \cdot 10^{-15}$
Third cumulant	$2.56 \cdot 10^{-14}$
Fourth cumulant	$3.49 \cdot 10^{-14}$

The numerical calculation confirms the root factorization at the level of the entire probability structure: the coefficients a_n , the distribution P_n , and the first four cumulants obtained from the linear root factors agree with the direct calculation from Z_{14} to machine precision.

7. Mean isotherm, residuals, and spectral scales for CO₂

For the 48 published NaY–CO₂ points at temperatures of 298.05–393.05 K and pressures of 0.027–3.018 MPa, the original temperature assigned to each experimental point was retained. The following measures of agreement between the calculated mean isotherm and the experimental data were obtained.

Table 6. Error in the mean amount of adsorbed CO₂

Range	N	MAE	RMSE	$\max \Delta q $	mean absolute relative error	mean z^2	$\max z $
All points	48	0.03582	0.05361	0.17854	1.562 %	33.47	19.84
$p \geq 0.25$ MPa	32	0.01552	0.02260	0.05081	0.267 %	4.56	4.62
$p < 0.15$ MPa	16	0.07641	0.08718	0.17854	4.152 %	91.28	19.84

The MAE, RMSE, and $\max|\Delta q|$ values are given in mmol/g. The mean absolute relative error is defined as

$$MARE = \frac{100\%}{N} \sum_{i=1}^N \frac{|\Delta q_i|}{q_{\text{exp},i}}.$$

For the normalized residual,

$$z_i = \frac{\Delta q_i}{u_i},$$

where, for the published expanded uncertainty U_i and coverage factor $k=2$

$$u_i = \frac{U_i}{2}.$$

The present calculation uses the temperatures of the individual rows of the original experimental data set. In Ref. [3], the same data were calculated using the nominal temperature branches 298.06, 333.06, 353.05, and 393.00 K. This explains the small difference between the summary metrics reported there and here; the statistical model and the set of experimental points remain the same.

For $p \geq 0.25$ MPa, the mean absolute error is 0.01552 mmol/g, whereas for $p < 0.15$ MPa it increases to 0.07641 mmol/g. In the low-pressure region, the following values are reached simultaneously:

$$\text{mean } z^2 = 91.28,$$

$$\max |z| = 19.84.$$

Consequently, the single-cell parameterization reproduces the main part of the measured range substantially more accurately, whereas at $p < 0.15$ MPa a systematic low-pressure mismatch remains that considerably exceeds the experimental uncertainty reported for those points. Therefore, the values in Table 6 characterize the quality of reproduction of the original data set used to parameterize the SSI. Independent testing of the predicted statistical structure requires additional observables, for example $\text{Var}(N)$, higher cumulants, or the complete distribution P_n .

The values in the table reproduce the published parameterization in which

$$x = K(T)p.$$

In the thermodynamic formulation using fugacity,

$$x = K(T)f.$$

The transition from p to f requires joint re-estimation of K_0 , E , and β because these parameters were obtained in the original pressure parameterization. At fixed α , such a transition does not change the root diagram Z_m , but it does change the correspondence between x and the experimental pressure.

If the function $\langle N \rangle(x)$ is known continuously and without error, then it uniquely determines $Z(x)$, because

$$\ln Z(x) = \int_0^x \left[\frac{\langle N \rangle(t)}{t} \right] dt,$$

$$Z(0) = 1.$$

For a finite noisy set of mean values, the inverse problem of reconstructing individual roots is much more poorly conditioned, especially for the distant high-activity tail. The complete distribution P_n , factorial cumulants, and coefficient criteria provide additional independent constraints on the coefficients of Z_m and on its root structure.

For CO_2 , the response distribution $\rho(s)$ has several distinct characteristic locations.

Table 7. Characteristic occupancy and response scales for CO_2

Quantity	x
First switching point $P_0 = P_1$	1.06388884571
Mode of $\rho(s)$: maximum of $\text{Var}(N)$	14.6060570636
Switching point $P_5 = P_6$	14.6901810343
Median of $\rho(s)$: $\langle N \rangle = 7$	27.7123337028
Exponential of the mean s for $\rho(s)$	37.9599141958
Last switching point $P_{13} = P_{14}$	5627.71656187

At the response mode,

$$\langle N \rangle = 5.47855549124,$$

$$\text{Var}(N) = 2.41896344326.$$

In logarithmic coordinates, the mode, median, and mean positions are

$$s_{mode} = \ln(14.6060570636) = 2.68143631005,$$

$$s_{med} = \ln(27.7123337028) = 3.32187757412,$$

$$\text{mean } s = 3.63653071336.$$

Thus,

$$s_{mode} < s_{med} < \text{mean } s.$$

This ordering is consistent with right-skewness of the response: the maximum is shifted toward smaller s relative to the median and the mean.

The integral characteristics are

$$\text{integral } \rho(s) ds = 1,$$

$$\text{mean } s = 3.63653071336,$$

$$\text{Var}_s(\rho) = 6.99962483116,$$

$$\text{standard deviation in } s = 2.64568040987.$$

Quantitatively, the sign and magnitude of the asymmetry are determined by the third cumulant

$$K_3[\rho] = 8.31668799672,$$

and by the normalized skewness

$$\frac{K_3[\rho]}{\text{Var}_s(\rho)^{3/2}} = 0.449095046025.$$

The positive value directly establishes right-skewness of the distribution ρ and links it to the long high-activity tail of the root spectrum.

The variance decomposes exactly:

$$\text{Var}_s(\rho) = 3.28986813370 + 3.70975669746.$$

The first term is $\pi^2/3$ and is generated by the logistic kernel. The second is the variance of $\ln r_k$. For CO_2 , the intrinsic spread of the root scales contributes about 53% of the total variance of the response in $\ln x$.

In the original pressure parameterization, each root corresponds to a pressure scale

$$p_k(T) = \frac{r_k}{K(T)},$$

at which $\pi_k = 1/2$. In the consistently constructed fugacity parameterization, analogously,

$$f_k(T) = \frac{r_k}{K(T)}.$$

In the general case, for the physical variable y used in the parameterization,

$$x = K(T)y,$$

$$y_k(T) = \frac{r_k}{K(T)}.$$

For

$$K(T) = K_0 \exp \left[\frac{E}{(RT)} \right]$$

we obtain

$$\ln y_k(T) = \ln \left(\frac{r_k}{K_0} \right) - \frac{E}{(RT)}.$$

Consequently, for unchanged m and α , all root scales have the common slope

$$\frac{-E}{R}$$

in coordinates $\ln y_k$ versus $1/T$, while their relative separations

$$\ln \left[\frac{y_k(T)}{y_j(T)} \right] = \ln \left(\frac{r_k}{r_j} \right)$$

are temperature-independent.

An individual logistic component has

$$\text{FWHM}_{\text{single}} = 3.52549435$$

along the $\ln x$ coordinate. All neighboring separations

$$\ln r_{k+1} - \ln r_k$$

for CO_2 are smaller than this width; therefore, the root components overlap and form a single broad response.

For the full CO_2 response,

$$\text{FWHM}_{\ln x} = 4.99223034,$$

which corresponds to

$$\frac{x_{\text{high}}}{x_{\text{low}}} = 147.2645.$$

For fixed m and α and an unchanged choice of the physical variable y ,

$$\ln x = \ln K(T) + \ln y.$$

Therefore, changing $K(T)$ only shifts the entire profile along $\ln y$. In the general thermodynamic representation $y = f$, whereas in the original pressure parameterization used above $y = p$. The width, skewness, and the remaining central cumulants of the normalized response are preserved under this shift.

Consequently, after accounting for the horizontal shift $K(T)$, a systematic temperature dependence of these characteristics cannot be explained by a change in affinity alone. Within the SSI with constant m and α , it implies a change in the distribution coefficients or a departure from such a single-cell description.

8. Homogeneous binomial control case and splitting of the root scales

For comparison, consider a system of m identical independent sites, for which

$$Z_L(x) = (1 + bx)^m.$$

All m roots of such a polynomial coincide:

$$r_1 = r_2 = \dots = r_m = \frac{1}{b}.$$

Therefore, the distribution of $\ln r_k$ collapses to a single point, and the normalized response is a single logistic kernel.

For the free-volume cell at $C = 0$, the factorization also contains m linear factors, but their scales r_k are different. The splitting of r_k creates, within a single finite cell, a distribution of characteristic activities and the corresponding width of the equilibrium response.

For $m = 1$, the partition function has the form

$$Z_1(x; \alpha) = 1 + (1 - \alpha)x,$$

therefore

$$\langle N \rangle = \frac{(1 - \alpha)x}{[1 + (1 - \alpha)x]}.$$

This is the Langmuir functional form with a renormalized affinity scale

$$K_{eff}(T) = (1 - \alpha) K(T).$$

The standard single-site form of the original finite single-cell model is obtained for

$$\begin{aligned} m &= 1, \\ \beta &= 0, \end{aligned}$$

that is, $\alpha = 0$ [3].

This case should be distinguished from the limit

$$\alpha \rightarrow 0$$

at fixed $m > 1$. Then

$$Z_m(x; 0) = \sum_{n=0}^m \frac{x^n}{n!},$$

and a truncated exponential sum arises rather than the binomial polynomial $(1 + bx)^m$.

The difference between the fully real SSI spectrum and the degenerate binomial spectrum can be measured directly from fluctuations. Set

$$\bar{p} = \frac{\langle N \rangle}{m}.$$

Then

$$\langle N \rangle \left[1 - \frac{\langle N \rangle}{m} \right] - \text{Var}(N) = \sum_{k=1}^m (\pi_k - \bar{p})^2.$$

Hence,

$$H_{spec} = 1 - \frac{\text{Var}(N)}{\langle N \rangle \left[1 - \frac{\langle N \rangle}{m} \right]}$$

can be written as

$$H_{spec} = \frac{\left[\sum_{k=1}^m (\pi_k - \bar{p})^2 \right]}{[m \bar{p} (1 - \bar{p})]}.$$

At finite activity, $0 < \bar{p} < 1$, and therefore

$$0 \leq H_{spec} \leq 1.$$

The value $H_{spec} = 0$ is attained if and only if all π_k are equal. In that case, the root spectrum degenerates to a single scale.

The same indicator has a direct interpretation for the shape of the mean isotherm. Since

$$F = \frac{\langle N \rangle}{m}$$

and

$$\frac{dF}{d \ln x} = \frac{\text{Var}(N)}{m},$$

the local logarithmic slope of the normalized occupancy is

$$n_H(x) = \frac{d \ln \left[\frac{F}{(1-F)} \right]}{d \ln x} = \frac{\text{Var}(N)}{\langle N \rangle \left[1 - \frac{\langle N \rangle}{m} \right]}.$$

Consequently,

$$n_H(x) = 1 - H_{spec}$$

and, in the fully real regime,

$$n_H(x) \leq 1.$$

Thus, H_{spec} simultaneously characterizes the splitting of the root scales and the reduction in the local isotherm slope relative to the degenerate binomial case. Violation of the condition $n_H \leq 1$ is incompatible with representing the occupancy distribution as a sum of independent Bernoulli random variables.

For CO₂ at the response maximum,

$$\langle N \rangle = 5.47855549124,$$

$$\text{Var}(N) = 2.41896344326,$$

whereas the binomial value at the same mean is

$$\langle N \rangle \left[1 - \frac{\langle N \rangle}{14} \right] = 3.33465761477.$$

Therefore,

$$H_{\text{spec}} = 0.27459915748,$$

$$n_H = 0.72540084252.$$

At the same mean, the variance of the SSI distribution is 27.5% lower than the corresponding binomial value. This difference is exactly equal to the sum of squared deviations of π_k from their mean and quantitatively characterizes the splitting of the root scales.

The difference is already visible in the first coefficients. For CO₂,

$$a_1 = 0.9399478188.$$

If b is chosen so that the homogeneous binomial model has the same linear coefficient,

$$b = \frac{a_1}{14},$$

then

$$a_2^L = 0.4101973117,$$

whereas the free-volume statistical weight gives

$$a_2 = 0.3871081666.$$

Consequently, deviation from the degenerate homogeneous binomial structure appears already in the second coefficient of the partition function and then persists throughout the entire sequence of states.

9. Newton inequalities as a necessary condition

If a polynomial of degree m has only negative real zeros, then the normalized coefficients

$$b_n = \frac{a_n}{C(m, n)}$$

satisfy the Newton inequalities

$$b_n^2 \geq b_{n-1} b_{n+1}.$$

Since $P_n = a_n x^n / Z$, the same condition can be checked from a single distribution:

$$\left[\frac{P_n}{C(m, n)} \right]^2 \geq \left[\frac{P_{n-1}}{C(m, n-1)} \right] \left[\frac{P_{n+1}}{C(m, n+1)} \right].$$

Violation of even one of these inequalities excludes $C = 0$. Satisfaction of all Newton inequalities is only a necessary condition.

For example, at $m = 14$ and $\alpha = 0.053$ the polynomial has two complex-conjugate pairs, $C = 2$, although all the stated inequalities are satisfied; the minimum ratio of the left-hand side to the right-hand side is 1.034779.

Therefore, the Newton inequalities are useful for preliminary testing of the coefficients or a measured P_n , but final classification requires either the full PF^∞ criterion or direct determination of the roots.

When working with small probabilities, the relative ratio of the two sides, rather than their absolute difference, must be compared: absolute values in the tail of the distribution can be very small even when the relative violation is substantial.

10. PF^∞ criterion and factorial cumulants

The Newton inequalities provide only a necessary local condition. A necessary and sufficient coefficient criterion follows from the Aissen–Edrei–Schoenberg–Whitney theorem [4].

For the finite sequence $a_0, \dots, a_m$, construct the two-sided Toeplitz matrix

$$T_{ij} = a_{j-i},$$

where

$$a_n = 0$$

for $n < 0$ and $n > m$. The sequence belongs to the PF^∞ class if all finite minors of this matrix are nonnegative.

For a polynomial with positive coefficients, the following conditions are equivalent:

1. all zeros Z are real and negative ;
2. sequence $a_0, \dots, a_m$ belongs to PF^∞ ;
3. $P_0, \dots, P_m$ is the distribution of a sum of m independent Bernoulli variables.

The transformation

$$a_n \rightarrow P_n = \frac{a_n x^n}{Z}$$

multiplies every Toeplitz minor by a positive factor. Therefore, its sign is independent of the activity x and of normalization. Total nonnegativity of the Toeplitz matrix is thus a coefficient criterion for $C = 0$ without explicitly computing the zeros.

In practice, violation of a Newton inequality or of a small Toeplitz minor immediately excludes $C = 0$. If all checked local conditions are satisfied, this is still insufficient: total nonnegativity of the corresponding matrix or exact root reconstruction is required.

In the regime $C = 0$,

$$G(1+t) = \prod_{k=1}^m (1 + \pi_k t),$$

where

$$\pi_k = \frac{x}{(x + r_k)}.$$

Therefore, the factorial cumulants are

$$\kappa_q^F = (-1)^{q-1} (q-1)! \sum_{k=1}^m \pi_k^q.$$

Their signs alternate:

$$\kappa_1^F \geq 0,$$

$$\kappa_2^F \leq 0,$$

$$\kappa_3^F \geq 0,$$

$$\kappa_4^F \leq 0,$$

...

Violation of any of these signs is incompatible with $C = 0$. Conversely, a finite number of correct signs does not replace the full root or PF_∞ criterion.

The Cauchy inequality simultaneously gives

$$\text{Var}(N) = \langle N \rangle - \sum_{k=1}^m \pi_k^2 \leq \langle N \rangle \left[1 - \frac{\langle N \rangle}{m} \right] \leq \frac{m}{4}.$$

The first bound is attained when the π_k are equal. The difference between this bound and the actual variance is exactly equal to the sum of squared deviations of π_k from their mean, which is what was used above in the definition of H_{spec} .

11. Observable consequences of the root representation

The mean isotherm specifies the first cumulant of the occupancy distribution. Fluctuation measurements provide higher cumulants, while the complete distribution P_n determines the entire finite sequence of probabilities. All these observables refer to the same finite statistical measure.

11.1. Mean isotherm

Let

$$q_{\text{sat}} = m c_{\text{ads}}$$

and

$$F(y, T) = \frac{q(y, T)}{q_{\text{sat}}} = \frac{\langle N \rangle}{m}.$$

At fixed temperature,

$$x = K(T)y,$$

therefore

$$\frac{dF}{d \ln y} = \frac{\text{Var}(N)}{m} = \rho.$$

Thus, the normalized profile ρ is determined directly by the derivative of the mean isotherm with respect to the logarithm of the chosen physical variable.

For a fixed cell,

$$\int \rho(\ln y) d \ln y = 1.$$

Through $K(T)$, temperature shifts this profile along $\ln y$ without changing its shape when m and α remain constant. The initial curvature of the isotherm determines α through the formulas of Section 4.3.

When processing experimental data, the derivative $dF/d \ln y$ should be obtained from a smooth fit to the isotherm with the data uncertainty taken into account.

In the fully real regime,

$$\rho = \mu_m * l,$$

where

$$l(t) = \frac{1}{\left[4 \cosh^2 \left(\frac{t}{2} \right) \right]}.$$

The Fourier transform of the logistic kernel is

$$\hat{l}(k) = \frac{\pi k}{\sinh(\pi k)},$$

with, by continuous extension,

$$\hat{l}(0) = 1.$$

Since $\hat{l}(k)$ has no real zeros, in the ideal noise-free case the function ρ uniquely determines the distribution of root scales μ_m .

On the experimental axis,

$$\ln x = \ln K(T) + \ln y,$$

and therefore the relative positions

$$\ln y_k = \ln r_k - \ln K(T).$$

are determined. The rapid decrease of $|\hat{l}(k)|$ at large $|k|$ makes reconstruction of individual roots poorly conditioned with respect to experimental noise. The mean, width, and skewness of ρ are determined more robustly.

11.2. Occupancy fluctuations

If the variance or higher occupancy cumulants are available independently of the mean isotherm, they make it possible to test directly the fluctuation consequences of the model:

$$\text{Var}(N) = \frac{d\langle N \rangle}{d \ln x},$$

the condition for a stationary point of the response,

$$\kappa_3 = 0,$$

the condition for a nondegenerate maximum,

$$\kappa_4 < 0,$$

and, in the regime $C = 0$, also the bound

$$\text{Var}(N) \leq \langle N \rangle \left[1 - \frac{\langle N \rangle}{m} \right].$$

and the alternating signs of the factorial cumulants.

11.3. Complete occupancy distribution

If $m \geq 2$ and the occupancy distribution of an individual cell is known, either directly or estimated from an ensemble of equivalent cells, then P_0 , P_1 , and P_2 determine α and x , after which the entire sequence $P_3, \dots, P_m$ is fixed.

As the activity changes, the distributions must be connected by the exact renormalization of Section 4.4, while the invariants I_n must remain unchanged.

The zeros of the generating function satisfy

$$z_k(x) = \frac{\xi_k}{x},$$

if the coefficients a_n remain unchanged.

The Newton inequalities provide a necessary condition for a fully real spectrum. The necessary and sufficient coefficient criterion is given by the PF^∞ property; equivalent information is provided by direct reconstruction of the roots.

For M identical independent cells, the generating function of the total occupancy is the M -th power of the single-cell generating function. The positions of its zeros are preserved, while their multiplicities are multiplied by M .

The contribution of a root ξ_k to the factorial cumulant of order q contains the factor

$$\left[\frac{x}{(x - \xi_k)} \right]^q.$$

Therefore, roots located far outside the activity range under study appear only weakly in the accessible lower-order cumulants.

The mean isotherm determines the integrated response profile, independent fluctuation data provide its cumulants, and the complete distribution P_n makes it possible to reconstruct the coefficients a_n and the corresponding root structure Z_m .

12. Reproducibility of the algebraic classification

Completeness of the root diagram for $m = 14$ is established by exact algebraic operations. To eliminate factorial denominators, the integer polynomial

$$P_{14}(x, \alpha) = 14! Z_{14}(x; \alpha).$$

is used. All internal values of α at which a multiple root in x is possible are contained among the zeros of the resultant

$$D_{14}(\alpha) = \text{Res}_x \left(P_{14}, \frac{\partial P_{14}}{\partial x} \right) = A(14\alpha - 1)^{14} Q_{182}(\alpha).$$

After separating the outer boundary $\alpha = 1/14$, the polynomial Q_{182} is tested for square-freeness. Complete rational isolation of its real roots determines all candidates for internal boundaries in the physical interval

$$0 < \alpha < \frac{1}{14}.$$

Between neighboring isolated values, a rational test point is selected at which the number of real zeros of P_{14} is counted exactly. Because there are no multiple roots inside such an interval, the number of real zeros is constant throughout the interval. It is precisely the combination of the resultant, complete isolation, and inter-boundary counting that provides a certificate of completeness for the root diagram.

High-precision numerical computation is used after the algebraic classification as an independent check. For the roots found, the residuals of P_{14} and its derivative, reconstruction of the original coefficients, factorization of the generating function, convolution of Bernoulli distributions, and agreement of the cumulants are verified. These checks characterize the numerical accuracy of the implementation but do not replace the algebraic proof of completeness.

The complete sequence of seven physical boundaries and the inter-boundary classification are given in Section 5.1. The algebraic verification of uniqueness of the CO_2 response maximum is presented together with the spectral response law itself in Section 4.1.

13. Discussion

The finite partition function

$$Z_m(x; \alpha) = \sum_{n=0}^m a_n(\alpha) x^n$$

unifies the coefficient, probabilistic, and root descriptions of a single cell. The coefficients a_n specify the statistical weights of the states, P_n gives the occupancy distribution, derivatives of $\ln Z_m$ give its cumulants, and the zeros of Z_m determine the factorization and root structure of the partition function.

The parameters α and x act on these objects in different ways. Changing x redistributes the probabilities P_n while leaving the coefficients a_n unchanged. Changing α alters the coefficients themselves and, together with them, the root spectrum Z_m .

For $m = 14$, this deformation passes through seven successive pairwise collisions. Beyond the final boundary, all fourteen roots are real and negative. The generating function of the distribution then factorizes, and the number of molecules has the distribution of a sum of fourteen independent Bernoulli random variables. The index k labels the root factors of the probability measure, while the physical cell is still described by a single volume V_c and a single parameter α .

Root factorization gives a direct representation of the equilibrium response. In the variable $s = \ln x$, the quantity

$$\rho(s) = \frac{\text{Var}(N)}{m}$$

is a normalized density and, for $C = 0$, has the form of an equally weighted mixture of m standard logistic densities shifted to the points

$$\ln r_k.$$

Therefore, the spread of root scales contributes directly to the width and skewness of the response already within a single finite cell with one value of α .

For CO_2 , this is seen directly from the locations of the mode, median, and mean of the response:

$$\begin{aligned} x_{\text{mode}} &= 14.6060570636, \\ x_{\text{median}} &= 27.7123337028, \\ \exp[\text{mean}(\ln x)] &= 37.9599141958. \end{aligned}$$

Their ordering

$$\text{mode} < \text{median} < \text{mean}$$

corresponds to the long high-activity tail. The normalized skewness of the distribution ρ is approximately 0.4491.

Within the SSI model considered here, splitting of the root scales inside a single free-volume cell generates an intrinsic contribution to the width and skewness of the equilibrium response. Therefore, a broad isotherm profile is compatible with a single-cell description if the corresponding finite polynomial has a sufficiently split root spectrum.

The parameters α , $K(T)$, and c_{ads} govern different properties of the isotherm. The parameter α determines the coefficients a_n and hence the shape of the normalized response and the relative positions of the roots. The function $K(T)$ sets the horizontal position of this profile on the logarithm of the physical variable, while c_{ads} sets the vertical scale of the absolute loading.

If m and α are constant, the temperature shift induced by $K(T)$ preserves all central cumulants of the normalized response. In the thermodynamic representation the physical variable is the fugacity f , while in the original pressure parameterization it is the pressure p .

Therefore, agreement of the entire temperature-superposed response shape simultaneously tests its width, skewness, and higher central characteristics and is a stronger condition than agreement only of the position of the maximum.

The sensitivity of individual roots to α differs substantially. For a simple root,

$$Z(-r_k, \alpha) = 0$$

we have

$$\frac{dr_k}{d\alpha} = \frac{Z_\alpha}{Z_x}$$

at $x = -r_k$. For the central CO₂ parameterization, the logarithmic sensitivity

$$\frac{d \ln r_k}{d \ln \alpha}$$

varies from -0.9663 for the first root to $+32.0285$ for the fourteenth. Thus, near this parameterization, a change in α affects the distant high-activity tail much more strongly than the lower part of the root spectrum.

The difference in the scales r_k directly affects their observability. The maximum of the CO₂ response is located at $x \approx 14.6$, that is, within the range of the first few r_k , whereas

$$r_{14} \approx 4.34 \cdot 10^3.$$

Therefore, data from the main occupancy range are substantially more sensitive to the lower and middle parts of the root spectrum than to the most distant roots. Consequently, the distant high-activity tail is much more weakly determined from the mean isotherm.

Comparison with the 48 NaY–CO₂ points shows good reproduction of the mean isotherm over the main part of the measured range; at low pressures, a systematic mismatch exceeding the tabulated uncertainty remains. The mean isotherm is used to estimate the parameters of the single-cell parameterization. Once they are fixed, the same polynomial Z_m , without introducing additional parameters, uniquely determines $\text{Var}(N)$, the higher cumulants, the complete distribution P_n , and the spectral characteristics of the response. Therefore, independent measurement of even some of these quantities provides a stricter test of the internal statistical structure of the model than repeated comparison only with the mean occupancy.

At large capacity, a near-boundary regime arises,

$$1 - m\alpha = O(m^{-1/2}),$$

in which the states $n = m, m-1, \dots$ form a boundary layer near the upper occupancy limit. After reversal and normalization of the polynomial, the coefficients of this layer converge to the coefficients of the partial theta function. For the branch of double roots converging to its first spectral point,

$$\eta_m^{(*)} = 1 - \frac{0.923078379995}{\sqrt{m}} + \frac{0.611344656644}{m} + O(m^{-3/2}).$$

The computed final boundaries for $m = 2, \dots, 30$ follow the same branch. The local limiting transition is established analytically. Global continuation of this identification to arbitrarily large m requires control of the remaining roots of the finite polynomial.

14. Conclusion

The finite partition function

$$Z_m(x; \alpha) = \sum_{n=0}^m \left[\frac{(1 - n\alpha)^n}{n!} \right] x^n$$

specifies the statistical weights of the states of a finite adsorption cell. The normalized function $Z_m(xz)/Z_m(x)$ determines the occupancy distribution, derivatives of $\ln Z_m$ give its cumulants, and the zeros of Z_m determine the factorization structure of the partition function.

For $m = 14$, the exact resultant has seven internal roots in α in the region

$$0 < \alpha < \frac{1}{14}.$$

Between them, the number of real zeros of Z_{14} changes successively as

$$0, 2, 4, 6, 8, 10, 12, 14.$$

The final boundary is

$$\alpha_c = 0.0554547158587412,$$

$$\eta_c = 0.776366022022377.$$

Since the double root at the boundary itself is already real, the fully real region is

$$\alpha_c \leq \alpha < \frac{1}{14}.$$

In this region,

$$Z_m(x) = \prod_{k=1}^m \left(1 + \frac{x}{r_k} \right),$$

and the occupancy distribution coincides with the distribution of a sum of m Bernoulli random variables with probabilities

$$\pi_k = \frac{x}{(x + r_k)}.$$

The same r_k determine the mean occupancy, variance, and the entire equilibrium response.

In the variable $s = \ln x$, the function

$$\rho(s) = \frac{\text{Var}(N)}{m}$$

is nonnegative and normalized to unity; that is, it defines a normalized density of the equilibrium response on the logarithmic-activity axis. For $C=0$, it is the convolution of a discrete measure concentrated at the points $\ln r_k$ with a logistic kernel. Therefore, the root spectrum directly determines the position, width, skewness, and higher cumulants of the response profile.

For CO_2 ,

$$x_{mode} = 14.6060570636,$$

$$x_{median} = 27.7123337028,$$

$$\exp[\text{mean}(\ln x)] = 37.9599141958,$$

and the normalized skewness of $\rho(s)$ is approximately

$$0.4491.$$

For the central CO_2 parameterization, the maximum of $\text{Var}(N)$ is unique: the algebraic equation

$$\kappa_3(x) = 0$$

has exactly one positive root, and at this point

$$\kappa_4 < 0.$$

The root spectrum simultaneously determines both edges of the isotherm. For $|x| < r_1$, the convergent series is determined by the inverse power sums r_k , while for $|x| > r_m$ the approach to saturation is determined by their direct power sums.

For the NaY–CO₂ and NaY–CH₄ parameters at $m = 14$, $C = 0$ is obtained. The root class is determined by the pair

$$(m, \alpha),$$

where

$$\alpha = \frac{\beta}{V_c}.$$

The parameters K_0 and E do not enter the root equation. If y denotes the physical variable used in the consistent parameterization,

$$x = K(T)y,$$

then each root defines the scale

$$y_k(T) = \frac{r_k}{K(T)},$$

and all root scales of a fixed cell have the same temperature slope

$$\frac{d \ln y_k}{d \left(\frac{1}{T} \right)} = \frac{-E}{R}.$$

In the thermodynamic representation $y = f$, whereas in the original pressure parameterization $y = p$. Consequently, within the SSI, a broad spectrum r_k defines an internal hierarchy of effective configurational occupancy scales that are shifted with temperature through the common affinity scale $K(T)$.

For $m \geq 2$, the complete distribution P_n determines α and x already from P_0 , P_1 , and P_2 , after which the remaining probabilities are fixed. The mean isotherm expresses the same structure through the normalized response profile and its temperature shift.

In the large-capacity limit, the reversed and normalized polynomial, together with the derivatives entering the double-root system, converges locally to a partial theta function. Nondegeneracy of its first double point determines a unique local asymptotic branch of finite-dimensional double roots:

$$\eta_m^{(*)} = 1 - \frac{0.923078379995}{\sqrt{m}} + \frac{0.611344656644}{m} + O(m^{-3/2}),$$

and simultaneously

$$z_m^{(*)} = z_1 + O(m^{-1/2}),$$

$$z_1 = -2.32037694428960 \dots$$

For the computed $m=2, \dots, 30$, this local branch coincides with the final boundary of the fully real spectrum. The derived asymptotic relation applies strictly to the double-root branch itself; preserving its identification with the final global root boundary for arbitrarily large m requires additional control of the remaining roots of the finite polynomial.

In the finite SSI model, the free-volume constraint produces a hierarchy of characteristic activity scales. Through the factorization of Z_m , this hierarchy is directly linked to the occupancy distribution and the shape of the equilibrium response.

References

1. Ruthven D. M. Simple Theoretical Adsorption Isotherm for Zeolites. *Nature Physical Science* 232 (1971), 70–71. DOI: 10.1038/physci232070a0.
2. Wedler C., Ferre A., Azzan H., Danaci D., Petit C., Pini R. Binary Adsorption Equilibria of Three $\text{CO}_2 + \text{CH}_4$ Mixtures on NIST Reference Zeolite Y (RM 8850) at Temperatures from 298 to 353 K and Pressures up to 3 MPa. *Journal of Chemical & Engineering Data* 69 (2024), 4216–4229. DOI: 10.1021/acs.jced.4c00358.
3. Rybakov P. I. Finite Single-Cell Statistical Model of Equilibrium CO_2 Adsorption on Zeolite NaY NIST RM 8850. Zenodo, 2026. DOI: 10.5281/zenodo.21035246.
4. Aissen M., Edrei A., Schoenberg I. J., Whitney A. M. On the Generating Functions of Totally Positive Sequences. *Proceedings of the National Academy of Sciences* 37 (1951), 303–307. DOI: 10.1073/pnas.37.5.303.
5. Kostov V. P. On the Double Zeros of a Partial Theta Function. *Bulletin des Sciences Mathématiques* 140 (2016), 98–111. DOI: 10.1016/j.bulsci.2015.04.006.
6. Kostov V. P., Shapiro B. Hardy–Petrovitch–Hutchinson’s Problem and Partial Theta Function. *Duke Mathematical Journal* 162 (2013), 825–861. DOI: 10.1215/00127094-2087264.
7. Sokal A. D. The Leading Root of the Partial Theta Function. *Advances in Mathematics* 229 (2012), 2603–2621. DOI: 10.1016/j.aim.2012.01.012.
8. Peng D.-Y., Robinson D. B. A New Two-Constant Equation of State. *Industrial & Engineering Chemistry Fundamentals* 15 (1976), 59–64. DOI: 10.1021/i160057a011.
9. Nguyen H. G. T. et al. Porosity, Powder X-Ray Diffraction Patterns, Skeletal Density, and Thermal Stability of NIST Zeolitic Reference Materials RM 8850, RM 8851, and RM 8852. *Journal of Research of NIST* 126 (2021), Article 126047. DOI: 10.6028/jres.126.047.

Appendix A. Brief summary of notation

Symbol	Meaning
V_c	volume of the selected adsorption cell; for NaY, the adopted volume of a faujasite-framework supercage
β	effective volumetric parameter of the SSI free-volume closure
$\alpha = \beta/V_c$	dimensionless free-volume parameter
m	integer limiting capacity of the cell
n	number of molecules in the cell

Symbol	Meaning
x	dimensionless activity; $x = K(T)y$
y	physical variable of the chosen parameterization: $y = p$ for the pressure parameterization and $y = f$ for the fugacity parameterization
p	gas pressure
f	gas fugacity
C_{ads}	factor converting mean occupancy to absolute loading; $q = c_{\text{ads}}\langle N \rangle$
a_n	weight of state n
Z_m	finite partition function
P_n	probability of occupancy n
ξ_k	k th zero of Z_m
r_k	$= -\xi_k$, positive scale of a real zero
C	number of complex-conjugate pairs of zeros
I_n	activity invariant $P_{n+1}P_{n-1}/P_n^2$
κ_q	occupancy cumulant of order q

Appendix B. Complete sequence of coefficients for CO₂

n	a_n
0	1
1	0.939947818826967
2	0.387108166581212
3	0.0918420467823145
4	0.0138856421013694
5	0.00139797513256981
6	$9.51639145428829 \cdot 10^{-5}$
7	$4.36164452391978 \cdot 10^{-6}$
8	$1.31739477555206 \cdot 10^{-7}$
9	$2.51805337766743 \cdot 10^{-9}$
10	$2.85211927004226 \cdot 10^{-11}$
11	$1.72599903968812 \cdot 10^{-13}$
12	$4.71945250722807 \cdot 10^{-16}$
13	$4.36293865925834 \cdot 10^{-19}$
14	$7.75259132419446 \cdot 10^{-23}$

Appendix C. Final boundary of the fully real spectrum for $m = 2, \dots, 30$

m	α_c	$\eta_c = m\alpha_c$	$y_c = -x_c$
2	0.226540919660986	0.453081839321973	2.58578643763
3	0.180479603994681	0.541438811984044	5.20372449567
4	0.149344562435611	0.597378249742446	9.63653085188
5	0.127350393436968	0.636751967184839	16.5574186841

m	α_c	$\eta_c = m\alpha_c$	$y_c = -x_c$
6	0.111058183028202	0.666349098169215	26.7940319846
7	0.0985177647823282	0.689624353476297	41.3403500988
8	0.0885676207216148	0.708540965772919	61.3764160661
9	0.0804784194271191	0.724305774844072	88.2911749672
10	0.0737705844135331	0.737705844135331	123.707759015
11	0.0681162516560575	0.749278768216632	169.511176878
12	0.0632838053179517	0.759405663815420	227.878511476
13	0.0591050085166534	0.768365110716494	301.311779050
14	0.0554547158587412	0.776366022022377	392.673625281
15	0.0522379008613756	0.783568512920634	505.226049575
16	0.0493811010902796	0.790097617444474	642.672361600
17	0.0468266221402738	0.796052576384655	809.202586628
18	0.0445285161432142	0.801513290577855	1009.54254875
19	0.0424497319504646	0.806544907058827	1249.00687395
20	0.0405600571963714	0.811201143927428	1533.55616843
21	0.0388346068546071	0.815526743946749	1869.85864168
22	0.0372526960989201	0.819559314176242	2265.35645827
23	0.0357969880500209	0.823330725150480	2728.33711783
24	0.0344528412057250	0.826868188937399	3268.01017860
25	0.0332078039849928	0.830195099624820	3894.58965663
26	0.0320512190648054	0.833331695684940	4619.38245049
27	0.0309739106351345	0.836295587148630	5454.88315921
28	0.0299679349648510	0.839102179015829	6414.87568084
29	0.02902637979982097	0.841765014194808	7514.54199855
30	0.0281432017811462	0.844296053434386	8770.57858243

For the computed values $m = 2, \dots, 30$, the final boundaries form a sequence consistent with the branch of double roots converging to the first spectral point of the partial theta function. For this branch,

$$\eta_m^{(*)} = 1 - \frac{\gamma_1}{\sqrt{m}} + \frac{0.611344656644}{m} + O(m^{-3/2}),$$

where

$$\gamma_1 = 0.923078379995 \dots$$

A convenient quantity for monitoring the leading limit remains

$$\gamma_m = [1 - \eta_c(m)]\sqrt{m}.$$

In the sequence shown, γ_m approaches γ_1 monotonically. For the computed $m = 2, \dots, 30$, the local limiting branch coincides with the final boundary of the fully real spectrum. Whether this coincidence persists for all sufficiently large m remains a separate global question and does not follow from the local limiting analysis alone.