Relative covariance in discrete Gibbs models

MANTIS DISCOVERY

MANTIS · MIT CSAIL

OCTOBER 2026

01

The question

Many learning methods need the covariance of a probabilistic model. An approximation can have a small error overall and still be badly wrong in a direction of small variance. When can we approximate the covariance of a discrete Gibbs model accurately in relative terms, small directions included?

02

Absolute and relative error

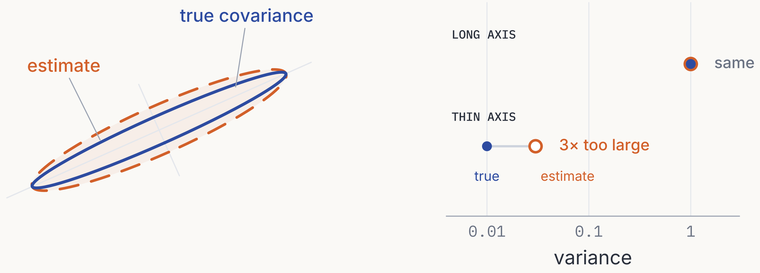


FIGURE 1. The dashed estimate is off by only 0.02 in absolute terms, but along the thin direction its variance is three times too large.

03

How we are approaching it

- Read the covariance scales off the affine spans of the model's energy sublevel sets.
- For distributions on matroid bases, compare with Laplacians through single-element exchanges.
- For hard-core models, use a cluster expansion in the activity.

04

A toy model

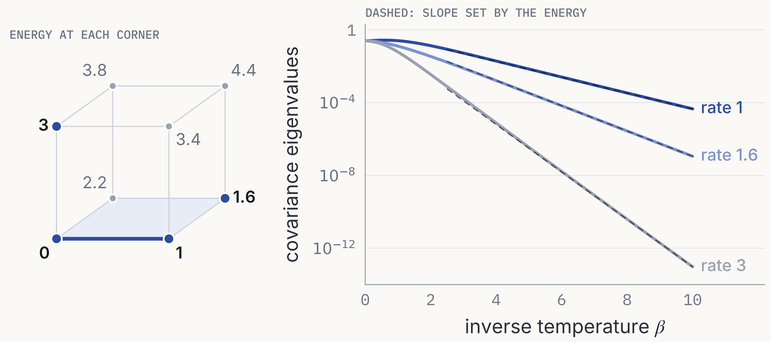


FIGURE 2. This toy model is not from the paper: its states are the eight corners of a cube, with the energies shown. As the inverse temperature β grows, the covariance eigenvalues fall at rates 1, 1.6 and 3 (dashed), the energies at which the low-energy corners first span a line, a plane and the whole space. The corner at 2.2 adds no new rate.

05

What we have so far

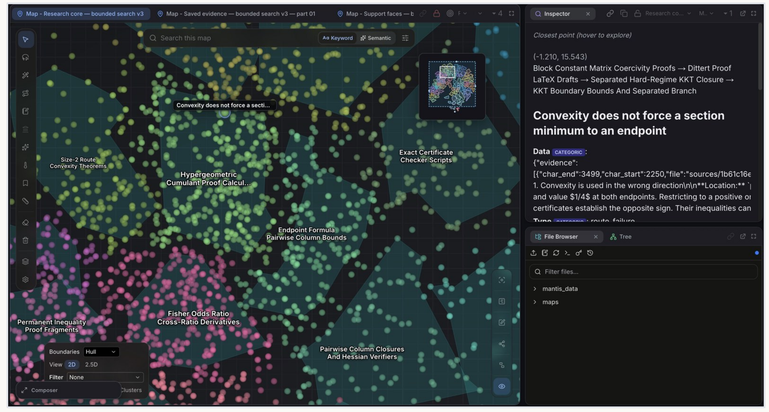
For a fixed finite Gibbs model, the asymptotic covariance scales come from the affine spans of energy sublevel sets, and the last energy at which the span grows sets the sampling threshold for relative recovery.

On matroid bases, single-element exchanges give a Laplacian comparison, and conditioning on bases within a fixed exchange distance gives a finite-temperature bound.

For hard-core models, the error is controlled by the local activity.

06

In Mantis



9,024 RECORDS · THE OPEN RECORD

Convexity does not force a section minimum to an endpoint.

The record is a rejected route: $f(x) = (x-1)^2$ on $[\frac{1}{2}, \frac{3}{2}]$ is convex, yet its minimum lies inside the interval.

ALSO ON RECORD

An explicit 3×3 matrix refutes a general tensor-norm multiplication rule.

07

Open questions

- In a small conditional subset-prediction problem, exact Newton was faster. Where does relative covariance control pay off in practice?
- The second-order hard-core approximation stayed accurate on growing strips at low activity. How high can the activity go?