# CS294-180 Written Report: Kawasaki Dynamics Beyond the Uniqueness Threshold

Devon Ding and William Lin

December 2023

## 1 Introduction

Kawasaki dynamics on the canonical Ising model is a Markov Chain Monte Carlo algorithm that is shown in the main paper to mix quickly beyond the uniqueness threshold on random d-regular graphs, a result that was previously unknown and does not hold for Glauber dynamics on the Ising model.

The Ising model is a fundamental model in statistical physics modeling magnetism. The Ising model begins with $N$ particles, and an instance of this is represented by $\sigma \in \{\pm 1\}^N$. The probabilities of each $\sigma$ are then proportional to $\frac{1}{Z}e^{-\frac{\beta}{2}(\sigma, A\sigma)+(h,\sigma)}$. The normalizing factor is the partition function, but computing it is not the purpose of this paper. The canonical Ising model is the Ising model restricted to a fixed magnetization, which can be written to say the sample space is $\sigma \in \{\pm 1\}^N : \sum_{i=1}^N \sigma_i = Nm$, where $m$ is a magnetization constant such that $Nm \in Z$. This has the property of changing the normalizing constant, but the $\sigma$ are still distributed proportional to before.

On the random d-regular graph, a $\sigma \in \Omega_{N,m}$ is the assignment of vertices to $\{\pm 1\}$, the $A$ is often the Laplacian and $h$ denotes the magnetic field. In this framework for the canonical Ising model, the pluses can be thought of as particles and the minuses as holes, where there is a fixed number of particles. Kawasaki Dynamics on the random d-regular graph would describe the transitions between assignments as choosing a particle to randomly jump to an adjacent hole depending on the energy. Formally where $\sigma$ is an assignment where $i = 1, j = -1$ and $\sigma^{ij}$ is the assignment where $i = -1, j = 1$, we have that the jump probability is denoted $c(\sigma, i \to j) = c_{\pi_{N,m,h}}(\sigma, i \to j) = \mathbf{1}_{\sigma_i=1}\mathbf{1}_{\sigma_j=-1}\frac{\pi(\sigma^{ij})}{\sum_{k\notin(I(\sigma)\{i\}))}\pi(\sigma^{ik})}$ where $\pi$ is a measure and $I(\sigma)$ is the set of $i : \sigma_i = +1$.

The main result that is proved in the paper[1] is as follows:

**Theorem 1.1** (threshold for Kawasaki dynamics on random $d$-regular graphs)**.** For $d \geq 3$ and $\beta < \frac{1}{8\sqrt{d-1}}$, the Kawasaki dynamics of the canonical Ising model on a random d-regular graph on $N$ vertices mixes in $O_{d,\beta}(N \log^6 N)$ steps with high probability (where probability is over the randomness of the graph).

## 2 Thresholds and Intuitions

The uniqueness and reconstruction thresholds, $\beta_c \approx \frac{1}{d-1}, \beta_r \approx \frac{1}{\sqrt{d-1}}$ respectively, are thresholds for $\beta$ on random d-regular graphs. In particular, it is known that Glauber dynamics mixes fast with $\beta < \beta_c$, and slowly with $\beta > \beta_c$ on random d-regular graphs in the Ising model. However, the paper shows that beyond the uniqueness threshold $\beta_c$, it is still possible for Kawasaki dynamics to mix fast on random d-regular graphs on the canonical Ising model.

The *reconstruction problem* involves estimating the value of a specific vertex $r \in V$ when a boundary is fixed. For $i, j \in V, d(i,j)$, is their distance and $\bar{B}(i,t)$ are the set of vertices $j$ such that $d(i,j) \geq t$. Formally, the reconstruction problem is $(t,\epsilon)-$reconstructible for graph $G$ at $r$ if $||P_{r,\bar{B}(r,t)}\{\cdot,\cdot|G\} - P_r\{\cdot|G\}P_{\bar{B}(r,t)}\{\cdot|G\}||_{TV} \geq \epsilon$. This equation can be thought of as the effect of a root variable under the influence of an average boundary in a tree, as random d-regular graphs locally are tree-like as $n \to \infty$. Under the random d-regular graphs in the canonical Ising model, $\beta_r$ as a threshold means that $\beta < \beta_r$ is non-reconstructible. An intuition is that with lower $\beta$, the probabilities have less reliance on the structure of the graph induced by the $(\sigma, A\sigma)$ term, which would relate to the correlation between $r$ and its boundary. Thus, fast mixing would imply that it must be non-reconstructible.

For a high-level picture and intuition of what the results imply, we have that for the Ising model, $0 < \beta_c < \beta_r$. For the canonical Ising model, we are isolated to a fixed magnetization, while the Ising model contains all possible magnetizations. Because Glauber dynamics in the Ising model must mix through different magnetization levels, it intuitively seems more difficult to mix. For instance, we consider $\beta < \beta_c$. In this case, Glauber dynamics can mix quickly through the entire state space. Kawasaki dynamics constrained to a single magnetization slice would also mix quickly. In the case where $\beta_c < \beta < \beta_r$, Glauber dynamics mixing slow and Kawasaki mixing fast can be interpreted as connections within the same magnetization level remain strong, but between different magnetization levels are weaker. For instance, considering the state where all $\sigma_i = 1$, the Glauber dynamic change of changing the $\sigma_i$ to $-1$ (changing magnetization level) would incur a penalty from all neighbors scaled depending on $A$ and $\beta$ because of the $\frac{\beta}{2}(\sigma, A\sigma)$ term. However for mixing within a magnetization level, the number of $+1, -1$ are fixed, so switching the position would have around equal benefit. Suppose if $\sigma_i = 1, \sigma_j = -1$ were switched such that $\sigma_i = -1, \sigma_j = 1$. Changing $\sigma_i$ from 1 to -1 incurs benefit from its neighbors that were -1 and loss from those that were +1, and the same for $\sigma_j$. Intuitively this should cancel to be relatively neutral compared to the between magnetisation slices case.

## 3 Definitions

The paper proves and applies a modified Log-Sobolev inequality, so the following definitions are along a functional analysis line.

Let $A$ be a symmetric matrix with constant eigenvector $\vec{1} = (1,...,1)^N$, meaning $A\vec{1} = \lambda_1\vec{1}$, which is the matrix of interactions. Let $h$ be a vector of length $n$. The canonical Ising measure with coupling matrix $A_{ij}$ and external field $\vec{h} = (h_i)$ is defined by

$$E_\nu[F] \propto \sum_{\sigma \in \Omega_{N,m}} e^{\frac{-\beta}{2}(\sigma, A\sigma) + (h,\sigma)} F(\sigma), \Omega_{N,m} = \{\sigma \in \{\pm 1\}^N : \sum_{i=1}^{N} \sigma_i = Nm\},$$

where N is the number of particles and $m \in [-1,1]$ is the magnetisation level such that $Nm$ is an integer.

We then define the entropy and variance of a function for a measure $\nu$ as

$$\begin{aligned} \mathrm{Ent}_\nu(F) &= E_\nu[\Phi(F)] - \Phi(E_\nu[F]), \qquad \Phi(x) = x \log x; \\ \mathrm{Var}_\nu(F) &= E_\nu[\Phi(F)] - \Phi(E_\nu[F]), \qquad \Phi(x) = x^2. \end{aligned}$$

In the results, either the Spectral Condition (SC) or Covariance Condition (CC) are required.

- The Spectral Condition has that $\delta_\lambda = \lambda_N - \lambda_2$ where $\lambda_N$ is the largest eigenvalue and $\lambda_2$ is the second smallest of $A$. Then $\beta < 1/(2\delta_\lambda)$ and $C_\beta = 1/(1-2\beta\delta_\lambda)$.
- The Covariance Condition has $\bar{X}^0(\beta)$ as an upper bound on the largest eigenvalue of the covariance matrix $(\mathrm{Cov}_{\nu_{\beta,h}}(\sigma_i;\sigma_j))_{i,j}$ of the canonical Ising model on $\Omega_{N,m}$ and sets $C_\beta = \exp[\int_0^\infty \bar{X}^0(t)dt]$.

The general covariance condition is implied by the spectral condition. Note that the requirement $\beta < \frac{1}{8\sqrt{d-1}}$ for fast mixing comes from the spectral condition, as the spectrum for a random graphs adjacency matrix is contained in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ w.h.p., meaning $\delta_\lambda = 4\sqrt{d-1}$, and thus $\beta < \frac{1}{2\delta_\lambda} \implies \beta < \frac{1}{2\sqrt{d-1}}$

# 4 Main results

**Theorem 4.1** (modified log-Sobolev inequality for Kawasaki dynamics)**.** Assume either (SC) or (CC), and assume further that $\max_i \sum_{j\neq i} |A_{ij}| \leq \bar{A}$ and $\max_i |h_i| \leq \bar{h}$. Then uniformly in size $N$ and magnetisation $m$, the canonical Ising model satisfies

$$\mathrm{Ent}_\nu(F) \leq C_\beta C(\beta\bar{A}, \bar{h}) D_\nu(F, \log F),$$

where

$$D_\nu(F,G) = \frac{1}{2N} \sum_{i,j=1}^{N} E_\nu[(F(\sigma) - F(\sigma^{ij}))(G(\sigma) - G(\sigma^{ij}))].$$

From Theorem 4.1, a known technique shows that the modified log-Sobolev inequality implies the usual log-Sobolev inequality with additional factor $O_{\beta,h}(\log N)$, giving rise to Corollary 4.1.1.

**Corollary 4.1.1.** Under the same assumption as in Theorem 1.2 (and with the same Dirichlet form),

$$\mathrm{Ent}_\nu(F) \leq C_\beta C(\beta\bar{A}, \bar{h})(\log N) D_\nu(\sqrt{F}, \sqrt{F})$$

With this log-Sobolev inequality, we gain a log-Sobolev constant with application of the moving particle lemma to change the mean-field Dirichlet form to the Kawasaki Dirichlet form,

$$\mathrm{Ent}_\nu(F) \leq C(d, \beta\bar{A}, \bar{h}) \log^5 N \sum_{i\sim j} E_\nu[(\sqrt{F}(\sigma) - \sqrt{F}(\sigma^{ij}))^2]$$

This final bound on the entropy gives rise to our final result.

**Corollary 4.1.2** (mixing on random $d$-regular graph)**.** Let $d \geq 3$. For the canonical nearest-neighbor Ising model on a random d-regular graph on $N$ vertices with $\beta \leq \frac{1}{8\sqrt{d-1}}$, the inverse log-Sobolev constant of Kawasaki dynamics is bounded by $C(d, \beta, h) \log^5 N$. Hence Kawasaki dynamics mixes in at most $O_{d,\beta,h}(N \log^6 N)$ steps.

# 5 Proof Sketch

To prove Theorem 1.2, a dynamic decomposition of the original measure $\nu$ is critically used. On a high level, the proof consists of the following three parts.

1. Step 1. Compare $\mathrm{Ent}_\nu(F)$ with the "decomposed entropy" $\mathbf{E}_{\nu_0}(\mathrm{Ent}_\pi(F))$. This is accomplished via constructing a dynamic decomposition and checking the entropic stability conditions. Eventually, one will get for some constant $C_\beta$ that only depends on $\beta$,
$$\mathrm{Ent}_\nu(F) \leq C_\beta \mathbf{E}_{\nu_0}(\mathrm{Ent}_\pi(F)).$$
2. Step 2. Apply modified log-Sobolev inequality for the inner measure $\pi$. Since it is well-known that $\pi$ has a log-concave generating function, which implies entropic independence and modified log-Sobolev inequality[2], i.e.,
$$\mathrm{Ent}_\pi(F) \leq D_\pi(F, \log F).$$
Here $D_\pi$ is the Dirichlet form associated with density $\pi$.
3. Step 3. Compare Dirichlet forms. Let $D_\pi := D_{\pi_{N,m,\beta\phi}}$, then one can prove the following relationship between "decomposed Dirichlet form" and the original Dirichlet form,
$$\mathbf{E}_{\nu_0}(D_\pi(F, \log F)) \leq K(\beta\overline{A}) D_\nu(F, \log F).$$
In all, we have
$$\begin{aligned} \mathrm{Ent}_\nu(F) &\leq C_\beta \mathbf{E}_{\nu_0}(\mathrm{Ent}_\pi(F)) && \text{(compare entropy)} \\ &\leq C_\beta \mathbf{E}_{\nu_0}(D_\pi(F, \log F)) && \text{(modified LSI for } \pi) \\ &\leq C_\beta K(\beta\overline{A}) D_\nu(F, \log F) && \text{(compare Dirichlet forms).} \end{aligned}$$

## 5.1 The dynamic decomposition

At the core of the proof of Theorem 1.2 is the dynamic decomposition which we describe as follows. Consider a path between $\beta A$ and $\beta I$ as $C_t^{-1} = tA + (\beta - t)I$. Then $C_0^{-1} = \beta I$, and $C_\beta^{-1} = \beta A$. Specifically, let $X_{N,0} := \{\phi \in \mathbf{R}^n : \sum_i \phi_i = 0\}$, then we have
$$e^{-\frac{\beta}{2}(\sigma, A\sigma)} \propto e^{-\frac{1}{2}(\sigma, C_\beta^{-1}\sigma)} \propto \int_{X_{N,0}} e^{-\frac{1}{2}(\sigma-\phi, C_t^{-1}(\sigma-\phi))} e^{-\frac{1}{2}(\phi, (C_\beta - C_t)^{-1}\phi)} d\phi.$$

The second step is due to the convolution property of Gaussians: the sum of two independent Gaussians with covariance $C_t$ and $(C_\beta - C_t)$ is a Gaussian with covariance $C_\beta$.

Let's formally define the decomposition of $\nu$ into $\nu_t$ and $\mu_t^\phi$. We let
$$\begin{aligned} \frac{d\nu_t}{d\phi}(\phi) &\propto e^{-\frac{1}{2}(\phi, (C_\beta - C_t)^{-1}\phi) - V_t(\phi)}, && \phi \in X_{N,0} \\ \mu_t^\phi(\sigma) &\propto e^{-\frac{1}{2}(\sigma-\phi, C_t^{-1}(\sigma-\phi)) + (\sigma, h)} \propto e^{-\frac{1}{2}(\sigma, C_t^{-1}\sigma) + (C_t^{-1}\phi + h, \sigma)}, && \sigma \in \Omega_{N,m}. \end{aligned}$$

Here the renormalized potential is $V_t(\phi) := -\log \sum_{\sigma\in\Omega_{N,m}} e^{-\frac{1}{2}(\phi-\sigma, C_t^{-1}(\phi-\sigma))}$. The following two scenarios are of primal interest:

(a) When $t = 0$, we have

$$\mu_0^\phi = \pi_{N,m,\beta\phi+h}.$$

(b) When $t = \beta$, we have

$$\begin{aligned} \frac{d\nu_\beta}{d\phi}(\phi) &= \delta_0(\phi), && \phi \in X_{N,0} \\ \mu_\beta^\phi(\sigma) &\propto e^{-\frac{\beta}{2}(\sigma,A\sigma)+(\beta A\phi+h,\sigma)} && \propto e^{-\frac{\beta}{2}(\sigma,A\sigma)+(h,\sigma)}, \sigma \in \Omega_{N,m}. \end{aligned}$$

Notice that the outer measure is concentrated at $\phi = 0$, therefore the inner measure can be simplified as shown above. Clearly, by definition of $\nu$, we have $\mu_\beta^\phi = \nu$.

Therefore, when $t = 0$, it coincides with the base decomposition of $\nu$ into $\nu_0$ and $\pi$; when $t = \beta$, it exactly becomes the original measure $\nu$ (composed with Dirac measure).

## 5.2 Compare entropy

To compare the entropy, we recall the following result by Bauerschmidt et, al.[3] We give a proof sketch of this theorem.

**Theorem 5.1.** Assume there are $\alpha_t > 0$ such that

$$\forall \phi \in X_{N,0}: \quad \dot{C}_t C_t^{-1} \mathrm{Cov}(\mu_t^\phi) C_t^{-1} \dot{C}_t \preceq \alpha_t \dot{C}_t.$$

Then the measure $\mu_t^\phi$ satisfies the $\alpha_t$-entropic stability condition: $\forall \phi \in X_{N,0}$,

$$2\Big(\nabla\sqrt{\mathbf{E}_{\mu_t^\phi} F}\Big)^2_{\dot{C}_t}(\phi) \le \alpha_t \mathrm{Ent}_{\mu_t^\phi}(F).$$

This implies the entropic contraction: $\forall s > 0$,

$$\mathrm{Ent}_\nu(F) \le e^{\int_0^\beta \alpha_u du} \mathbf{E}_{\nu_0}(\mathrm{Ent}_\pi(F)).$$

*Proof sketch.* Let $\Phi(x) := x \log x$. Then we have

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{E}_{\nu_t}(\mathrm{Ent}_{\mu_t^\phi}(F)) &= \frac{\partial}{\partial t}\mathbf{E}_{\nu_t}(\mathbf{E}_{\mu_t^\phi}(\Phi(F)) - \Phi(\mathbf{E}_{\nu_t^\phi}(F))) \\ &= 2\mathbf{E}_{\nu_t}\Big(\Big(\nabla\sqrt{\mathbf{E}_{\mu_t^\phi} F}\Big)^2_{\dot{C}_t}\Big). \end{aligned}$$

One can think of it being analogous to the connections between the time derivative of the entropy along a Markov chain and the corresponding Dirichlet form. Now the $\alpha_t$-entropic stability condition followed by a standard Gronwall's inequality concludes the proof. □

We won't discuss how to derive entropic stability from the covariance condition cause it's lengthy. However, we give some high-level intuitions on what the covariance condition is about. Note that $\mathrm{Hess}V_t(\phi) = C_t^{-1} - C_t^{-1}\mathrm{Cov}(\mu_t^\phi)C_t^{-1}$, then the covariance condition is equivalent to

$$\dot{C}_t \mathrm{Hess}V_t(\phi)\dot{C}_t - \dot{C}_t C_t^{-1}\dot{C}_t \succeq -\alpha_t \dot{C}_t, \quad \forall t > 0.$$

Intuitively, this says that the smoothed potential function should satisfy some relaxed convexity conditions along the decomposition flow. Such relaxed convexity along the flow then implies entropic stability.

We then try to derive bounds on the covariance matrix of $\mu_t^\phi$ to apply Theorem 5.1 above. This motivates the following bound.

**Lemma 5.2** (covariance upper bound)**.** For $\|A\| \le 1$ and $\beta < \frac{1}{2}$, we have

$$\mathrm{Cov}(\mu_t^\phi) \preceq \left(\frac{1}{2} - t\right)^{-1} I.$$

The proof of this lemma involves using negative correlation of the inner measure $\pi$ as well as strict log-concavity of $\nu_0$ when $\beta < \frac{1}{2}$. Using this lemma, it's not hard to see that the main inequality in Theorem 5.1 is satisfied with $C_\beta = \int_0^\beta (\frac{1}{2} - t)^{-1} dt$.

We give some reasons why this decomposition is natural:

(a) It is known as Pochinski renormalization flow in the physics literature and is related to a reparametrized stochastic localization scheme;

(b) Using this decomposition, we can prove inequalities from a dynamic point of view, and we can use the properties of the basic decomposed measure $\pi$ and $\nu_0$ jointly.

## 5.3 Compare Dirichlet forms

To compare Dirichlet forms $\mathbf{E}_{\nu_0}(D_\pi(F, \log F))$ and $D_\nu(F, \log F)$, one notice that

$$\begin{aligned}
\mathbf{E}_{\nu_0}(D_\pi(F, \log F)) &= \mathbf{E}_{\nu_0}\Big(\sum_{\sigma\sim\sigma^{ij}} \pi(\sigma)c_\pi(\sigma,\sigma^{ij})(F(\sigma) - F(\sigma^{ij}))(\log(F(\sigma)) - \log(F(\sigma^{ij})))\Big) \\
&= \sum_{\sigma\sim\sigma^{ij}} \mathbf{E}_{\nu_0}\Big(\pi(\sigma)c_\pi(\sigma,\sigma^{ij})\Big)(F(\sigma) - F(\sigma^{ij}))(\log(F(\sigma)) - \log(F(\sigma^{ij}))), \\
D_\nu(F, \log F) &= \sum_{\sigma\sim\sigma^{ij}} \nu(\sigma)c_\nu(\sigma,\sigma^{ij})(F(\sigma) - F(\sigma^{ij}))(\log(F(\sigma)) - \log(F(\sigma^{ij}))).
\end{aligned}$$

So a pairwise comparison is sufficient for comparing the Dirichlet forms. That is, if $\mathbf{E}_{\nu_0}(\pi(\sigma)c_\pi(\sigma,\sigma^{ij})) \le K\nu(\sigma)c_\nu(\sigma,\sigma^{ij})$ for any $\sigma \sim \sigma^{ij}$, then we can deduce

$$\mathbf{E}_{\nu_0}(D_\pi(F, \log F)) \le K D_\nu(F, \log F).$$

To get a pairwise comparison bound, we are looking for some constant $K$ that upper bounds

$$\max_{\sigma\sim\sigma^{ij}} \frac{\mathbf{E}_{\nu_0}(\pi(\sigma)c_\pi(\sigma,\sigma^{ij}))}{\nu(\sigma)c_\nu(\sigma,\sigma^{ij})}.$$

However, the quantity $\frac{\mathbf{E}_{\nu_0}(\pi(\sigma)c_\pi(\sigma,\sigma^{ij}))}{\nu(\sigma)c_\nu(\sigma,\sigma^{ij})}$ can be interpreted as the expectation of some auxiliary random variable, which helps to derive a uniform bound on it. The exact derivation is computation-intensive and less intelligent and is left to the paper.

# References

[1] Bauerschmidt, R., Bodineau, T. and Dagallier, B., 2023. Kawasaki dynamics beyond the uniqueness threshold. arXiv preprint arXiv:2310.04609.

[2] Anari, N., Gharan, S.O. and Vinzant, C., 2018, October. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS) (pp. 35-46). IEEE.

[3] Bauerschmidt, R., Bodineau, T. and Dagallier, B., 2023. Stochastic dynamics and the Polchinski equation: an introduction. arXiv preprint arXiv:2307.07619.

[4] A. Gerschenfeld and A. Montanari. Reconstruction for models on random graphs. In 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), pages 194–204, 2007.