This is a review submitted to Mathematical Reviews/MathSciNet.

Reviewer Name: Konstantopoulos, Takis

Mathematical Reviews/MathSciNet Reviewer Number: 68397

Address:

Department of Mathematics
Uppsala University
PO Box 480
SE-75106 Uppsala
SWEDEN
takiskonst@gmail.com

Author: Hermon, Jonathan; Peres, Yuval

Title: The power of averaging at two consecutive time steps: proof of a mixing conjecture by Aldous and Fill.

MR Number: MR3729646

Primary classification:

Secondary classification(s):

Review text:

This paper answers a question posed by Aldous and Fill in their famous unpublished book https://www.stat.berkeley.edu/∼aldous/RWG/book.pdf on Markov chains. Consider a discrete-time, irreducible and time-reversible Markov chain $(X_0, X_1, \ldots)$ with values in a finite set with a (necessarily unique) stationary distribution π. Convergence of X_t to π as $t \to \infty$, $t \in \mathbb{N}$, may not be valid due to periodicity. There are many ways to tweak the chain so that convergence takes place. One is by considering the standard continuous-time chain $X^c_t := X_{N(t)}$, $t \geq 0$, where N is an independent Poisson process. Another is by replacing the transition probabilities $p(i, j)$ by $\frac{1}{2}(\delta(i, j) + p(i, j))$, where $\delta(i, j)$ is 1 or 0 according as $i = j$ or not; this gives the "lazy" chain X^L. A third one is by replacing the initial distribution μ by $\frac{1}{2}(\mu(i) + \sum_j \mu(j)p(j, i))$; this gives the "averaged" chain X^{ave}. These three variants have π as stationary distribution and converge to π. Let $d_c(t), d_L(t), d_{\text{ave}}(t)$ be the worst-case (with respect to the initial law) total-variation distance between π and X^c_t, X^L_t, X^{ave}_t, respectively. The Aldous-Fill conjecture is that there are functions ϕ, ψ such that $\phi(t)/t \to 1$ and $\psi(t) \to 0$, as $t \to \infty$, such that

$$d_{\text{ave}}(\phi(t)) \leq \psi(d_c(t)), \quad t \geq 0.$$

The paper answers this question affirmatively by explicitly prescribing the functions ϕ and ψ. It does so by providing explicit comparisons between the three total-variation distances. In addition, it is shown that if one of the chains X^c, X^L, X^{ave} exchibits cutoff then so do the others. One says that a chain X

converging to a stationary distribution exhibits cutoff if there is a natural parameter n of the chain (e.g., the size of the state space) such that if $t_n(\varepsilon)$ is inverse to the total-variation distance function $d_n(t)$ then $t_n(1-\varepsilon)/t_n(\varepsilon) \to 1$ as $n \to \infty$ for all $\varepsilon > 0$. The cutoff window is $t_n(\varepsilon) - t_n(1-\varepsilon)$. It is shown that X_L and X_{ave} have smaller cutoff window than X_c if cutoff takes place. The key ingredients in the analysis are natural couplings between the above chains and a maximal inequality due to Elias Stein (1961) stating that if P is a positive self-adjoint linear operator on $L^2(\Omega,\mu)$ into itself (where $L^2(\Omega,\mu)$ is the space of μ-square-integrable functions on the probability space (Ω,μ)) with spectrum contained in the unit interval $[0,1]$ then

$$\|\sup_{t\geq 0}(t+1)(P^{t+1}f - P^t f)\|_2 \leq C\|f\|_2,$$

for some universal positive constant C.