Comparison Inequalities and Fastest-mixing Markov Chains

James Allen FILL *The Johns Hopkins University, USA*, E-mail: jimfill@jhu.edu
Jonas KAHN *Université de Lille, France*, E-mail: jonas.kahn@math.univ-lille1.fr

KEY WORDS: Markov chains, comparison inequalities, fastest mixing, stochastic monotonicity, log-concave distributions, birth-and-death chains

MATHEMATICAL SUBJECT CLASSIFICATION: 60J10

Abstract: We introduce a new partial order on the class of stochastically monotone Markov kernels having a given stationary distribution π on a given finite partially ordered state space $\mathcal{X}$. When $K \preceq L$ in this partial order we say that K and L satisfy a *comparison inequality.* We establish that if $K_1, \dots, K_t$ and $L_1, \dots, L_t$ are reversible and $K_s \preceq L_s$ for $s = 1, \dots, t$, then $K_1 \cdots K_t \preceq L_1 \cdots L_t$. In particular, in the time-homogeneous case we have $K^t \preceq L^t$ for every t if K and L are reversible and $K \preceq L$, and using this we show that (for suitable common initial distributions) the Markov chain Y with kernel K mixes faster than the chain Z with kernel L, in the strong sense that *at every time* t the discrepancy—measured by total variation distance or separation or L^2-distance—between the law of Y_t and π is smaller than that between the law of Z_t and π.

Using comparison inequalities together with specialized arguments to remove the stochastic monotonicity restriction, we answer a question of Persi Diaconis by showing that, among all symmetric birth-and-death kernels on the path $\mathcal{X} = \{0, \dots, n\}$, the one (we call it the *uniform chain*) that produces fastest convergence from initial state 0 to the uniform distribution has transition probability 1/2 in each direction along each edge of the path, with holding probability 1/2 at each endpoint.

We also use comparison inequalities

(i) to identify, when π is a given log-concave distribution on the path, the fastest-mixing stochastically monotone birth-and-death chain started at 0, and

(ii) to recover and extend a Peres–Winkler result that extra updates do not delay mixing for monotone spin systems.

Among the fastest-mixing chains in (i), we show that the chain for uniform π is slowest in the sense of maximizing separation at every time.

References (selected)

[1] I. Benjamini, N. Berger, C. Hoffman, & E. Mossel (2005). Mixing times of the biased card shuffling and the asymmetric exclusion process, *Trans. Amer. Math. Soc.*, **357**(8), 3013–3029 (electronic).

[2] S. Boyd, P. Diaconis, P. Parrilo, & L. Xiao (2009). Fastest mixing Markov chain on graphs with symmetries, *SIAM J. Optim.*, **20**(2), 792–819.

[3] S. Boyd, P. Diaconis, J. Sun, & L. Xiao (2006). Fastest mixing Markov chain on a path, *Amer. Math. Monthly*, **113**(1), 70–74.

[4] S. Boyd, P. Diaconis, & L. Xiao (2004). Fastest mixing Markov chain on a graph, *SIAM Rev.*, **46**(4), 667–689 (electronic).

[5] P. Diaconis & J. A. Fill (1990). Strong stationary times via a new form of duality, *Ann. Probab.*, **18**(4), 1483–1522.

[6] J. A. Fill & J. Kahn (2012). Comparison inequalities and fastest-mixing Markov chains, *Annals of Applied Probability* (accepted subject to minor revision). Available from `http://www.ams.jhu.edu/~fill/`.

[7] Y. Peres & P. Winkler (2011). Can extra updates delay mixing?, Preprint, `arXiv:1112.0603v1 [math.PR]`.

[8] S. Roch (2005). Bounding fastest mixing, *Electron. Comm. Probab.*, **10**, 282–296 (electronic).

[9] J. Sun, S. Boyd, L. Xiao, & P. Diaconis (2006). The fastest mixing Markov process on a graph and a connection to a maximum variance unfolding problem, *SIAM Rev.*, **48**(4), 681–699 (electronic).