Entropy production in nonlinear recombination models.

Pietro Caputo

joint work with Alistair Sinclair

CIRM, Marseille
Apr. 19, 2017

Plan

- Nonlinear recombination models
- Convergence to equilibrium
- Entropy production estimates
- Sketch of proofs
- A general framework: Reversible Quadratic Systems
- Nonlinear stochastic Ising models
- Implementing Kac's program ?

Nonlinear recombination models

Classical model for genetic algorithms.

Consider sequences σ of length n from a finite alphabet S:

$$\sigma = (\sigma_1, \ldots, \sigma_n) \in \Omega = S^n$$

For any $A \subset [n] = \{1, \ldots, n\}$, write $\sigma = \sigma_A \sigma_{A^c}$.

Recombination at A (also "collision" or "mating"):

$$\Omega \times \Omega \ni (\sigma, \eta) = (\sigma_A \sigma_{A^c}, \eta_A \eta_{A^c}) \mapsto (\eta_A \sigma_{A^c}, \sigma_A \eta_{A^c}) = (\sigma', \eta')$$

Nonlinear recombination models

Classical model for genetic algorithms.

Consider sequences σ of length n from a finite alphabet S:

$$\sigma = (\sigma_1, \ldots, \sigma_n) \in \Omega = S^n$$

For any $A \subset [n] = \{1, \ldots, n\}$, write $\sigma = \sigma_A \sigma_{A^c}$.

Recombination at A (also "collision" or "mating"):

$$\Omega \times \Omega \ni (\sigma, \eta) = (\sigma_A \sigma_{A^c}, \eta_A \eta_{A^c}) \mapsto (\eta_A \sigma_{A^c}, \sigma_A \eta_{A^c}) = (\sigma', \eta')$$

If (σ, η) are sampled independently from a distribution p on Ω, and if A is sampled independently from a distribution ν on $[n]$, then (σ', η') has distribution

$$\sum_{A \subset [n]} \nu(A) \sum_{\sigma, \eta} p(\sigma) p(\eta) \, \mathbf{1}_{(\eta_A \sigma_{A^c}, \sigma_A \eta_{A^c}) = (\sigma', \eta')}$$

Nonlinear recombination II

If we only want the distribution of the first component σ':

$$\sum_{\eta'} \sum_{A \subset [n]} \nu(A) \sum_{\sigma,\eta} p(\sigma) p(\eta)\, \mathbf{1}_{(\eta_A \sigma_{A^c}, \sigma_A \eta_{A^c}) = (\sigma', \eta')}$$

Nonlinear recombination II

If we only want the distribution of the first component σ':

$$\sum_{\eta'} \sum_{A\subset[n]} \nu(A) \sum_{\sigma,\eta} p(\sigma)p(\eta)\,\mathbf{1}_{(\eta_A\sigma_{A^c},\sigma_A\eta_{A^c})=(\sigma',\eta')}$$

$$= \sum_{\sigma_A} \sum_{\eta_{A^c}} p(\sigma_A\sigma'_{A^c})p(\sigma'_A\eta_{A^c}) = \sum_{A\subset[n]} \nu(A)\, p_A(\sigma'_A)p_{A^c}(\sigma'_{A^c})$$

where p_A denotes the marginal on A.

Nonlinear recombination II

If we only want the distribution of the first component σ':

$$\sum_{\eta'} \sum_{A\subset[n]} \nu(A) \sum_{\sigma,\eta} p(\sigma)p(\eta)\, \mathbf{1}_{(\eta_A\sigma_{A^c},\sigma_A\eta_{A^c})=(\sigma',\eta')}$$

$$= \sum_{\sigma_A} \sum_{\eta_{A^c}} p(\sigma_A\sigma'_{A^c})p(\sigma'_A\eta_{A^c}) = \sum_{A\subset[n]} \nu(A)\, p_A(\sigma'_A)p_{A^c}(\sigma'_{A^c})$$

where p_A denotes the marginal on A.

In conclusion, one recombination maps the law p of the first sequence to

$$p \mapsto \Psi[p] := \sum_{A\subset[n]} \nu(A)\, p_A \otimes p_{A^c}\,,$$

Nonlinear, quadratic map.

Remark: the map Ψ preserves marginal at every single site $i \in [n]$:

$$\Psi[p]_i(\sigma_i) = p_i(\sigma_i)\,, \qquad i \in [n].$$

Evolution and convergence

Discrete time evolution:

$$p^{(k)} = \Psi[p^{(k-1)}], \quad k \in \mathbb{N}, \quad p^{(0)} = p.$$

Evolution and convergence

Discrete time evolution:

$$p^{(k)} = \Psi[p^{(k-1)}], \quad k \in \mathbb{N}, \quad p^{(0)} = p.$$

Continuous time evolution:

$$\begin{aligned} \partial_t p_t &= \Psi[p_t] - p_t \\ &= \sum_A \nu(A)(p_{t,A} \otimes p_{t,A^c} - p_t), \quad t \geqslant 0, \quad p_0 = p. \end{aligned}$$

Evolution and convergence

Discrete time evolution:

$$p^{(k)} = \Psi[p^{(k-1)}], \quad k \in \mathbb{N}, \quad p^{(0)} = p.$$

Continuous time evolution:

$$\begin{aligned}\partial_t p_t &= \Psi[p_t] - p_t \\ &= \textstyle\sum_A \nu(A)(p_{t,A} \otimes p_{t,A^c} - p_t), \quad t \geqslant 0, \quad p_0 = p.\end{aligned}$$

Convergence : If ν is nondegenerate
$[\forall i, j \in [n], \exists A \subset [n]$ s.t. $A \ni i, A^c \ni j$ and $\nu(A) > 0]$, then

$$p_t \to \pi := p_{0,1} \otimes \cdots \otimes p_{0,n}, \quad t \to \infty \quad \text{product of marginals.}$$

Evolution and convergence

Discrete time evolution:

$$p^{(k)} = \Psi[p^{(k-1)}], \quad k \in \mathbb{N}, \quad p^{(0)} = p.$$

Continuous time evolution:

$$\partial_t p_t = \Psi[p_t] - p_t$$
$$= \sum_A \nu(A)(p_{t,A} \otimes p_{t,A^c} - p_t), \quad t \geqslant 0, \quad p_0 = p.$$

Convergence : If ν is nondegenerate
$[\forall i, j \in [n], \exists A \subset [n]$ s.t. $A \ni i, A^c \ni j$ and $\nu(A) > 0]$, then

$$p_t \to \pi := p_{0,1} \otimes \cdots \otimes p_{0,n}, \quad t \to \infty \quad \text{product of marginals.}$$

[Geiringer 1944, ..., Rabani-Rabinovich-Sinclair 1998, ..., Baake-Baake-Salamat 2014, Martinez 2015]

Remark analogy with Boltzmann-like equations from kinetic theory: Ψ is a quadratic collision kernel as e.g. in the Kac model.

Later we discuss more general nonlinear evolutions which allow for convergence to non-product measures.

Trend to equilibrium: total variation distance

Consider the following examples of distribution ν:

1) Single site recombination: $\nu(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{A=\{i\}}$;
2) Single crossover: $\nu(A) = \frac{1}{n+1}\sum_{i=0}^{n} \mathbf{1}_{A=\{1,\ldots,i\}}$;
3) Uniform crossover: $\nu(A) = \frac{1}{2^n}$, for all $A \subset [n]$;
4) Bernoulli(q) model: $\nu(A) = q^{|A|}(1-q)^{n-|A|}$, some $q \in [0, \frac{1}{2}]$, .

Trend to equilibrium: total variation distance

Consider the following examples of distribution ν:

1) Single site recombination: $\nu(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{A=\{i\}}$;
2) Single crossover: $\nu(A) = \frac{1}{n+1}\sum_{i=0}^{n} \mathbf{1}_{A=\{1,\dots,i\}}$;
3) Uniform crossover: $\nu(A) = \frac{1}{2^n}$, for all $A \subset [n]$;
4) Bernoulli(q) model: $\nu(A) = q^{|A|}(1-q)^{n-|A|}$, some $q \in [0, \frac{1}{2}]$, .

Rabani-Rabinovich-Sinclair 1998 obtained rates of convergence of discrete-time evolution in total variation distance

$$\|\mu - \nu\|_{TV} = \frac{1}{2}\sum_{\sigma} |\mu(\sigma) - \nu(\sigma)|,$$

measured by the mixing time

$$T_{\mathrm{mix}}(\nu, n) = \max_{p^{(0)}} \min\{k \in \mathbb{N} : \ \|p^{(k)} - \pi\|_{TV} \ \leqslant \ \tfrac{1}{4}\}$$

Trend to equilibrium: total variation distance

Consider the following examples of distribution ν:

1) Single site recombination: $\nu(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{A=\{i\}}$;
2) Single crossover: $\nu(A) = \frac{1}{n+1}\sum_{i=0}^{n} \mathbf{1}_{A=\{1,\dots,i\}}$;
3) Uniform crossover: $\nu(A) = \frac{1}{2^n}$, for all $A \subset [n]$;
4) Bernoulli(q) model: $\nu(A) = q^{|A|}(1-q)^{n-|A|}$, some $q \in [0, \frac{1}{2}]$, .

Rabani-Rabinovich-Sinclair 1998 obtained rates of convergence of discrete-time evolution in total variation distance

$$\|\mu - \nu\|_{TV} = \frac{1}{2}\sum_{\sigma} |\mu(\sigma) - \nu(\sigma)|,$$

measured by the mixing time

$$T_{\mathrm{mix}}(\nu, n) = \max_{p^{(0)}} \min\{k \in \mathbb{N} : \ \|p^{(k)} - \pi\|_{TV} \ \leqslant \ \tfrac{1}{4}\}$$

For model 1 and 2, they prove $T_{\mathrm{mix}}(\nu, n) = O(n \log n)$;
for model 3: $T_{\mathrm{mix}}(\nu, n) = O(\log n)$;
for model 4: $T_{\mathrm{mix}}(\nu, n) = O(q^{-1} \log n)$ [coupling analysis].

Trend to equilibrium: entropy

Relative entropy: $$H(\mu|\nu) = \begin{cases} \sum_\sigma \mu(\sigma) \log\left(\frac{\mu(\sigma)}{\nu(\sigma)}\right) & \mu \ll \nu \\ +\infty & \mu \not\ll \nu \end{cases}$$

Pinsker's inequality: $\|\mu - \nu\|_{TV}^2 \leqslant \frac{1}{2} H(\mu|\nu)$.

Trend to equilibrium: entropy

Relative entropy: $H(\mu|\nu) = \begin{cases} \sum_\sigma \mu(\sigma) \log\left(\frac{\mu(\sigma)}{\nu(\sigma)}\right) & \mu \ll \nu \\ +\infty & \mu \not\ll \nu \end{cases}$

Pinsker's inequality: $\|\mu - \nu\|_{TV}^2 \leqslant \frac{1}{2} H(\mu|\nu)$.

Entropy decay rate: $\delta > 0$ such that for all $t \geqslant 0$:

$$\frac{\mathrm{d}}{\mathrm{dt}} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi). \qquad (1)$$

Trend to equilibrium: entropy

Relative entropy: $H(\mu|\nu) = \begin{cases} \sum_\sigma \mu(\sigma) \log\left(\frac{\mu(\sigma)}{\nu(\sigma)}\right) & \mu \ll \nu \\ +\infty & \mu \not\ll \nu \end{cases}$

Pinsker's inequality: $\|\mu - \nu\|_{TV}^2 \leqslant \frac{1}{2} H(\mu|\nu)$.

Entropy decay rate: $\delta > 0$ such that for all $t \geqslant 0$:

$$\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi). \tag{1}$$

Problem: Given product measure π, find $\delta = \delta(\pi, \nu, n) > 0$ s.t. this holds for all $t \geqslant 0$, for *all initial* p_0 with same marginals as π.

Trend to equilibrium: entropy

Relative entropy: $H(\mu|\nu) = \begin{cases} \sum_\sigma \mu(\sigma) \log\left(\frac{\mu(\sigma)}{\nu(\sigma)}\right) & \mu \ll \nu \\ +\infty & \mu \not\ll \nu \end{cases}$

Pinsker's inequality: $\|\mu - \nu\|_{TV}^2 \leqslant \frac{1}{2} H(\mu|\nu)$.

Entropy decay rate: $\delta > 0$ such that for all $t \geqslant 0$:

$$\frac{\mathrm{d}}{\mathrm{dt}} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi). \tag{1}$$

Problem: Given product measure π, find $\delta = \delta(\pi, \nu, n) > 0$ s.t. this holds for all $t \geqslant 0$, for *all initial* p_0 with same marginals as π.

In this case, decay to equilibrium: $H(p_t|\pi) \leqslant H(p_0|\pi)\, e^{-\delta t}$.

Trend to equilibrium: entropy

Relative entropy: $$H(\mu|\nu) = \begin{cases} \sum_\sigma \mu(\sigma) \log\left(\frac{\mu(\sigma)}{\nu(\sigma)}\right) & \mu \ll \nu \\ +\infty & \mu \not\ll \nu \end{cases}$$

Pinsker's inequality: $\|\mu - \nu\|_{TV}^2 \leqslant \frac{1}{2} H(\mu|\nu)$.

Entropy decay rate: $\delta > 0$ such that for all $t \geqslant 0$:

$$\frac{\mathrm{d}}{\mathrm{dt}} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi). \qquad (1)$$

Problem: Given product measure π, find $\delta = \delta(\pi, \nu, n) > 0$ s.t. this holds for all $t \geqslant 0$, for *all initial* p_0 with same marginals as π.

In this case, decay to equilibrium: $H(p_t|\pi) \leqslant H(p_0|\pi)\, e^{-\delta t}$.

Measuring convergence in terms of relative entropy is very natural in view of the analogy with kinetic theory; see e.g. Carlen-Carvalho 1992, or Desvillettes-Muhot-Villani 2011 review on Cercignani's conjecture. Unexplored thus far for nonlinear recombinations.

Functional inequalities

Fix π product measure. For any $f : \Omega \mapsto [0, \infty)$, define

$$f_A(\sigma) = \sum_{\eta_{A^c}} \pi_{A^c}(\eta_{A^c}) f(\sigma_A \eta_{A^c}) = \pi[f | \sigma_A].$$

Functional inequalities

Fix π product measure. For any $f : \Omega \mapsto [0, \infty)$, define

$$f_A(\sigma) = \sum_{\eta_{A^c}} \pi_{A^c}(\eta_{A^c}) f(\sigma_A \eta_{A^c}) = \pi[f | \sigma_A].$$

Suppose $f = \mu/\pi$ for some μ. Then $f_A = \frac{\mu_A}{\pi_A}$ and $f_A f_{A^c} = \frac{\mu_A \bigotimes \mu_{A^c}}{\pi}$. If $A = \{i\}$ write $f_i = f_{\{i\}}$. Note $f_i = 1$ means that μ and π have the same marginals: $\mu_i = \pi_i$.

Functional inequalities

Fix π product measure. For any $f : \Omega \mapsto [0, \infty)$, define

$$f_A(\sigma) = \sum_{\eta_{A^c}} \pi_{A^c}(\eta_{A^c}) f(\sigma_A \eta_{A^c}) = \pi[f | \sigma_A].$$

Suppose $f = \mu/\pi$ for some μ. Then $f_A = \frac{\mu_A}{\pi_A}$ and $f_A f_{A^c} = \frac{\mu_A \otimes \mu_{A^c}}{\pi}$. If $A = \{i\}$ write $f_i = f_{\{i\}}$. Note $f_i = 1$ means that μ and π have the same marginals: $\mu_i = \pi_i$. Define $\mathrm{Ent}(f) = \pi[f \log f] = H(\mu|\pi)$.

Setting $f = p_t/\pi$, the problem $\frac{\mathrm{d}}{\mathrm{dt}} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi)$ is equivalent to :

Problem: Given the measure π, find $\delta = \delta(\pi, \nu, n) > 0$ s.t.

$$\sum_A \nu(A)\, \pi \left[(f_A f_{A^c} - f) \log \frac{f_A f_{A^c}}{f} \right] \geqslant \delta\, \mathrm{Ent}(f)$$

for all $f : \Omega \mapsto [0, \infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

Functional inequalities

Fix π product measure. For any $f : \Omega \mapsto [0, \infty)$, define

$$f_A(\sigma) = \sum_{\eta_{A^c}} \pi_{A^c}(\eta_{A^c}) f(\sigma_A \eta_{A^c}) = \pi[f | \sigma_A].$$

Suppose $f = \mu/\pi$ for some μ. Then $f_A = \frac{\mu_A}{\pi_A}$ and $f_A f_{A^c} = \frac{\mu_A \otimes \mu_{A^c}}{\pi}$. If $A = \{i\}$ write $f_i = f_{\{i\}}$. Note $f_i = 1$ means that μ and π have the same marginals: $\mu_i = \pi_i$. Define $\mathrm{Ent}(f) = \pi[f \log f] = H(\mu|\pi)$.

Setting $f = p_t/\pi$, the problem $\frac{\mathrm{d}}{\mathrm{dt}} H(p_t|\pi) \leqslant -\delta\, H(p_t|\pi)$ is equivalent to :

Problem: Given the measure π, find $\delta = \delta(\pi, \nu, n) > 0$ s.t.

$$\sum_A \nu(A)\, \pi \left[(f_A f_{A^c} - f) \log \frac{f_A f_{A^c}}{f} \right] \geqslant \delta\, \mathrm{Ent}(f)$$

for all $f : \Omega \mapsto [0, \infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.
["nonlinear" Log-Sobolev inequality with constraints on marginals]

Main results I

Theorem (Entropy production estimates)

Let $\delta_* = \delta_*(\nu, n) := \inf_\pi \delta(\pi, \nu, n)$.

1) *Single site recombination:* $\frac{2}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n-1}$;
2) *Single crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n+1}$;
3) *Uniform crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-2^{-n+1}}{n-1}$;
4) *Be(q) model:* $\frac{4(1-(1-q/2)^n)}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-(1-q)^n-q^n}{n-1}$.

Main results I

Theorem (Entropy production estimates)

Let $\delta_* = \delta_*(\nu, n) := \inf_\pi \delta(\pi, \nu, n)$.

1) *Single site recombination:* $\frac{2}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n-1}$;

2) *Single crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n+1}$;

3) *Uniform crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-2^{-n+1}}{n-1}$;

4) *Be(q) model:* $\frac{4(1-(1-q/2)^n)}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-(1-q)^n-q^n}{n-1}$.

Upper bounds are obtained by calculating the two sides of the inequality for a suitable test function f [density of a mixture of Bernoulli] and a suitable product measure π [Be(w), $w \sim 2^{-n}$].

Main results I

Theorem (Entropy production estimates)

Let $\delta_* = \delta_*(\nu, n) := \inf_\pi \delta(\pi, \nu, n)$.

1) *Single site recombination:* $\frac{2}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n-1}$;
2) *Single crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1}{n+1}$;
3) *Uniform crossover:* $\frac{4}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-2^{-n+1}}{n-1}$;
4) *Be(q) model:* $\frac{4(1-(1-q/2)^n)}{n} + O(n^{-2}) \geqslant \delta_* \geqslant \frac{1-(1-q)^n-q^n}{n-1}$.

Upper bounds are obtained by calculating the two sides of the inequality for a suitable test function f [density of a mixture of Bernoulli] and a suitable product measure π [Be(w), $w \sim 2^{-n}$].

Lower bounds: By "linearization" we find $\delta(\pi, \nu, n) \geqslant \kappa(\pi, \nu, n)$, where $\kappa > 0$ is the optimal constant s.t.

$$\sum_A \nu(A)\left(\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})\right) \leqslant (1-\kappa)\,\mathrm{Ent}(f)\,, \qquad (2)$$

for all $f : \Omega \mapsto [0, \infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

Main results II

$$\sum_A \nu(A)\,\left(\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})\right) \;\leqslant\; (1-\kappa)\,\mathrm{Ent}(f)\,,$$

for all $f : \Omega \mapsto [0, \infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

Main results II

$$\sum_A \nu(A)\,(\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})) \leqslant (1-\kappa)\,\mathrm{Ent}(f)\,,$$

for all $f : \Omega \mapsto [0,\infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

With $\kappa = 0$ this is simply sub-additivity of relative entropy with respect to a product measure: $\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c}) \leqslant \mathrm{Ent}(f)$, for all $f \geqslant 0$ and all $A \subset [n]$.

Main results II

$$\sum_A \nu(A)\ (\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c}))\ \leqslant\ (1-\kappa)\,\mathrm{Ent}(f)\,,$$

for all $f : \Omega \mapsto [0, \infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

With $\kappa = 0$ this is simply sub-additivity of relative entropy with respect to a product measure: $\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})\ \leqslant\ \mathrm{Ent}(f)$, for all $f\ \geqslant\ 0$ and all $A \subset [n]$.

It cannot hold for $\kappa > 0$ without the marginal constraints on f: if f_i are non-trivial, then taking $f = \prod_i f_i$ yields $\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c}) = \mathrm{Ent}(f) = \sum_i \mathrm{Ent}(f_i)$ for all $A \subset [n]$.

Main results II

$$\sum_A \nu(A)\ (\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c}))\ \leqslant\ (1-\kappa)\,\mathrm{Ent}(f)\,,$$

for all $f : \Omega \mapsto [0,\infty)$ such that $\pi[f] = 1$ and $f_i = 1$ for all $i \in [n]$.

With $\kappa = 0$ this is simply sub-additivity of relative entropy with respect to a product measure: $\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})\ \leqslant\ \mathrm{Ent}(f)$, for all $f\ \geqslant\ 0$ and all $A \subset [n]$.

It cannot hold for $\kappa > 0$ without the marginal constraints on f: if f_i are non-trivial, then taking $f = \prod_i f_i$ yields $\mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c}) = \mathrm{Ent}(f) = \sum_i \mathrm{Ent}(f_i)$ for all $A \subset [n]$.

Theorem (Refined sub-additivity estimates)

Let $\kappa_* = \kappa_*(\nu, n) := \inf_\pi \kappa(\pi, \nu, n)$.

1) *Single site recombination:* $\kappa_* = \frac{1}{n-1}$;
2) *Single crossover:* $\kappa_* = \frac{1}{n+1}$;
3) *Uniform crossover:* $\kappa_* = \frac{1-2^{-n+1}}{n-1}$;
4) *Be(q) model:* $\kappa_* = \frac{1-(1-q)^n-q^n}{n-1}$.

Ideas of proof I (from δ to κ)

The lower bound $\delta(\pi, \nu, n) \geqslant \kappa(\pi, \nu, n)$, the "linearization", is a simple consequence of convexity:

$$\pi \left[(f_A f_{A^c} - f) \log \frac{f_A f_{A^c}}{f} \right] \geqslant \mathrm{Ent}(f) - \mathrm{Ent}(f_A) - \mathrm{Ent}(f_{A^c}).$$

Ideas of proof I (from δ to κ)

The lower bound $\delta(\pi,\nu,n) \geqslant \kappa(\pi,\nu,n)$, the "linearization", is a simple consequence of convexity:

$$\pi\left[(f_A f_{A^c} - f)\log\frac{f_A f_{A^c}}{f}\right] \geqslant \mathrm{Ent}(f) - \mathrm{Ent}(f_A) - \mathrm{Ent}(f_{A^c}).$$

Indeed: use $\pi\left[(f_A f_{A^c} - f)\log f_A f_{A^c}\right] = 0$, and

$$\begin{aligned}\pi\left[f_A f_{A^c}\log f\right] &\leqslant \pi\left[(f_A f_{A^c}\log f_A f_{A^c}\right]\\ &= \pi\left[f_A\log f_A\right] + \pi\left[f_{A^c}\log f_{A^c}\right]\\ &= \mathrm{Ent}(f_A) + \mathrm{Ent}(f_{A^c})\end{aligned}$$

Ideas of proof II (computation of optimal κ)

Fix $\mathcal{A}$ cover of $[n]$, i.e. a family of subsets covering $[n]$.

Lemma (Generalized sub-additivity inequalities)

For any $f \geqslant 0$:

$$\sum_A \mathrm{Ent}(f_A) \leqslant n_+(\mathcal{A}) \, \mathrm{Ent}(f)$$

where $n_+(\mathcal{A}) = \max_{k \in [n]} \#\{A \in \mathcal{A} : A \ni k\}$.

Ideas of proof II (computation of optimal κ)

Fix $\mathcal{A}$ cover of $[n]$, i.e. a family of subsets covering $[n]$.

Lemma (Generalized sub-additivity inequalities)

For any $f \geqslant 0$:

$$\sum_A \mathrm{Ent}(f_A) \leqslant n_+(\mathcal{A})\,\mathrm{Ent}(f)$$

where $n_+(\mathcal{A}) = \max_{k\in[n]} \#\{A \in \mathcal{A} : A \ni k\}$.

Proof. Use Shearer's inequality for Shannon's entropy. See [Shearer et al. 1986; Madiman-Tetali 2010; Balister-Bollobas 2012]. □

[See also (C, Menz, Tetali 2015) for some extensions of this to weakly-dependent non-product measures]

Ideas of proof II (computation of optimal κ)

Fix $\mathcal{A}$ cover of $[n]$, i.e. a family of subsets covering $[n]$.

Lemma (Generalized sub-additivity inequalities)

For any $f \geqslant 0$:

$$\sum_A \mathrm{Ent}(f_A) \leqslant n_+(\mathcal{A})\,\mathrm{Ent}(f)$$

where $n_+(\mathcal{A}) = \max_{k\in[n]} \#\{A \in \mathcal{A} : A \ni k\}$.

Proof. Use Shearer's inequality for Shannon's entropy. See [Shearer et al. 1986; Madiman-Tetali 2010; Balister-Bollobas 2012]. □

[See also (C, Menz, Tetali 2015) for some extensions of this to weakly-dependent non-product measures]

However, this lemma is not sufficient, it gives very poor bounds on κ (e.g. exp. small in n for uniform crossover). It is crucial to use:

Lemma (Sub-modularity, or strong sub-additivity)

For any $f \geqslant 0$, the function $A \mapsto h(A) := -\mathrm{Ent}(f_A)$ is sub-modular, i.e.

$$h(A) + h(B) \geqslant h(A \cap B) + h(A \cup B)\,, \qquad A, B \subset [n].$$

Reversible quadratic systems

Following [Rabinovich-Sinclair-Widgerson 1992] we define a general class of quadratic evolutions.

We do not require strict symmetry here but only reversibility.

Reversible quadratic systems

Following [Rabinovich-Sinclair-Widgerson 1992] we define a general class of quadratic evolutions.
We do not require strict symmetry here but only reversibility.

Idea: Take a Markov kernel Q on $\Omega \times \Omega$, that is $Q(\sigma, \sigma'; \tau, \tau') \geqslant 0$, $\sum_{\tau,\tau' \in \Omega} Q(\sigma, \sigma'; \tau, \tau') = 1$, and suppose that $\mu \in \mathrm{Prob}(\Omega)$ is such that $\mu \otimes \mu \in \mathrm{Prob}(\Omega \times \Omega)$ is reversible for Q.

Reversible quadratic systems

Following [Rabinovich-Sinclair-Widgerson 1992] we define a general class of quadratic evolutions.
We do not require strict symmetry here but only reversibility.

Idea: Take a Markov kernel Q on $\Omega \times \Omega$, that is $Q(\sigma, \sigma'; \tau, \tau') \geqslant 0$, $\sum_{\tau,\tau' \in \Omega} Q(\sigma, \sigma'; \tau, \tau') = 1$, and suppose that $\mu \in \mathrm{Prob}(\Omega)$ is such that $\mu \otimes \mu \in \mathrm{Prob}(\Omega \times \Omega)$ is reversible for Q.
Given an initial state $p \in \mathrm{Prob}(\Omega)$, construct the measure

$$[(p \otimes p)Q](\tau, \tau') = \sum_{\sigma,\sigma' \in \Omega} p(\sigma)p(\sigma')Q(\sigma, \sigma'; \tau, \tau').$$

Reversible quadratic systems

Following [Rabinovich-Sinclair-Widgerson 1992] we define a general class of quadratic evolutions.

We do not require strict symmetry here but only reversibility.

Idea: Take a Markov kernel Q on $\Omega \times \Omega$, that is $Q(\sigma, \sigma'; \tau, \tau') \geqslant 0$, $\sum_{\tau,\tau' \in \Omega} Q(\sigma, \sigma'; \tau, \tau') = 1$, and suppose that $\mu \in \mathrm{Prob}(\Omega)$ is such that $\mu \otimes \mu \in \mathrm{Prob}(\Omega \times \Omega)$ is reversible for Q.

Given an initial state $p \in \mathrm{Prob}(\Omega)$, construct the measure

$$[(p \otimes p)Q](\tau, \tau') = \sum_{\sigma,\sigma' \in \Omega} p(\sigma)p(\sigma')Q(\sigma, \sigma'; \tau, \tau').$$

Then take the projection on the first coordinate

$$\Psi[p](\tau) = \sum_{\tau' \in \Omega} [(p \otimes p)Q](\tau, \tau').$$

The nonlinear equation is then:

$$\partial_t p_t = \Psi[p_t] - p_t, \quad t \geqslant 0, \quad p_0 = p.$$

Reversible quadratic systems

Following [Rabinovich-Sinclair-Widgerson 1992] we define a general class of quadratic evolutions.

We do not require strict symmetry here but only reversibility.

Idea: Take a Markov kernel Q on $\Omega \times \Omega$, that is $Q(\sigma, \sigma'; \tau, \tau') \geqslant 0$, $\sum_{\tau,\tau' \in \Omega} Q(\sigma, \sigma'; \tau, \tau') = 1$, and suppose that $\mu \in \mathrm{Prob}(\Omega)$ is such that $\mu \otimes \mu \in \mathrm{Prob}(\Omega \times \Omega)$ is reversible for Q.
Given an initial state $p \in \mathrm{Prob}(\Omega)$, construct the measure

$$[(p \otimes p)Q](\tau, \tau') = \sum_{\sigma,\sigma' \in \Omega} p(\sigma)p(\sigma')Q(\sigma, \sigma'; \tau, \tau').$$

Then take the projection on the first coordinate

$$\Psi[p](\tau) = \sum_{\tau' \in \Omega}[(p \otimes p)Q](\tau, \tau').$$

The nonlinear equation is then:

$$\partial_t p_t = \Psi[p_t] - p_t, \quad t \geqslant 0, \quad p_0 = p.$$

Note: $p_t \to \mu$ or to some other μ' identified by the initial state p.
Important: Q is not assumed to be irreducible!

Nonlinear Stochastic Ising Model I

Fix a graph $G = ([n], E)$, external fields $\mathbf{h} = \{h_i\}$, and $\beta \in \mathbb{R}$:

$$\mu(\sigma) = \frac{1}{Z} \exp\left(\beta \sum_{ij \in E} \sigma_i \sigma_j + \sum_{i \in [n]} h_i \sigma_i\right)$$

$\mu = \mu_{G,\beta,\mathbf{h}}$ is a probability on $\Omega = \{-1, +1\}^n$.

Nonlinear Stochastic Ising Model I

Fix a graph $G = ([n], E)$, external fields $\mathbf{h} = \{h_i\}$, and $\beta \in \mathbb{R}$:

$$\mu(\sigma) = \frac{1}{Z} \exp\left(\beta \sum_{ij \in E} \sigma_i \sigma_j + \sum_{i \in [n]} h_i \sigma_i\right)$$

$\mu = \mu_{G,\beta,\mathbf{h}}$ is a probability on $\Omega = \{-1, +1\}^n$.
Consider the Reversible Markov Chain on $\Omega \times \Omega$ with kernel

$$Q(\sigma, \sigma'; \tau, \tau') = \sum_A \nu(A) Q_A(\sigma, \sigma'; \tau, \tau')$$

Nonlinear Stochastic Ising Model I

Fix a graph $G = ([n], E)$, external fields $\mathbf{h} = \{h_i\}$, and $\beta \in \mathbb{R}$:

$$\mu(\sigma) = \frac{1}{Z} \exp\left(\beta \sum_{ij \in E} \sigma_i \sigma_j + \sum_{i \in [n]} h_i \sigma_i\right)$$

$\mu = \mu_{G,\beta,\mathbf{h}}$ is a probability on $\Omega = \{-1, +1\}^n$.
Consider the Reversible Markov Chain on $\Omega \times \Omega$ with kernel

$$Q(\sigma, \sigma'; \tau, \tau') = \sum_A \nu(A) Q_A(\sigma, \sigma'; \tau, \tau')$$

$$Q_A(\sigma, \sigma'; \tau, \tau') = \alpha_A \, \mathbf{1}_{\tau = \sigma'_A \sigma_{A^c}, \, \tau' = \sigma_A \sigma'_{A^c}} + (1 - \alpha_A) \, \mathbf{1}_{\tau = \sigma, \, \tau' = \sigma'}$$

$$\alpha_A := \alpha_A(\sigma, \sigma') = \frac{\mu(\sigma'_A \sigma_{A^c}) \mu(\sigma_A \sigma'_{A^c})}{\mu(\sigma'_A \sigma_{A^c}) \mu(\sigma_A \sigma'_{A^c}) + \mu(\sigma) \mu(\sigma')}$$

Nonlinear Stochastic Ising Model I

Fix a graph $G = ([n], E)$, external fields $\mathbf{h} = \{h_i\}$, and $\beta \in \mathbb{R}$:

$$\mu(\sigma) = \frac{1}{Z} \exp\left(\beta \sum_{ij \in E} \sigma_i \sigma_j + \sum_{i \in [n]} h_i \sigma_i\right)$$

$\mu = \mu_{G,\beta,\mathbf{h}}$ is a probability on $\Omega = \{-1, +1\}^n$.
Consider the Reversible Markov Chain on $\Omega \times \Omega$ with kernel

$$Q(\sigma, \sigma'; \tau, \tau') = \sum_A \nu(A) Q_A(\sigma, \sigma'; \tau, \tau')$$

$$Q_A(\sigma, \sigma'; \tau, \tau') = \alpha_A \, \mathbf{1}_{\tau = \sigma'_A \sigma_{A^c}, \, \tau' = \sigma_A \sigma'_{A^c}} + (1 - \alpha_A) \, \mathbf{1}_{\tau = \sigma, \, \tau' = \sigma'}$$

$$\alpha_A := \alpha_A(\sigma, \sigma') = \frac{\mu(\sigma'_A \sigma_{A^c}) \mu(\sigma_A \sigma'_{A^c})}{\mu(\sigma'_A \sigma_{A^c}) \mu(\sigma_A \sigma'_{A^c}) + \mu(\sigma) \mu(\sigma')}$$

A kind of heat bath in $\Omega \times \Omega$ w.r.t. $\mu \otimes \mu$. Remarks:
1) $\beta = 0 \;\; \Rightarrow \;\; \alpha_A \equiv \frac{1}{2}$, "lazy" recombination model.
2) Kernel Q does not depend on external fields $\mathbf{h}$, but only on G, β.

Nonlinear Stochastic Ising Model II

The Markov chain yields a linear map on $\mathrm{Prob}(\Omega \times \Omega)$. The chain is not irreducible.

Nonlinear Stochastic Ising Model II

The Markov chain yields a linear map on $\mathrm{Prob}(\Omega \times \Omega)$. The chain is not irreducible. By projection one has the quadratic map on $\mathrm{Prob}(\Omega)$: $p \mapsto \Psi[p]$

$$\begin{aligned} \Psi[p](\tau) &= \sum_{\tau'} [(p \otimes p) Q](\tau, \tau') \\ &= \sum_{\sigma, \sigma', \tau'} p(\sigma) p(\sigma') Q(\sigma, \sigma'; \tau, \tau') \end{aligned}$$

Evolution equation:

$$\partial_t p_t = \Psi[p_t] - p_t, \quad t \geqslant 0, \quad p_0 = p.$$

Nonlinear Stochastic Ising Model II

The Markov chain yields a linear map on $\mathrm{Prob}(\Omega \times \Omega)$. The chain is not irreducible. By projection one has the quadratic map on $\mathrm{Prob}(\Omega)$: $p \mapsto \Psi[p]$

$$\Psi[p](\tau) = \sum_{\tau'} [(p \otimes p)Q](\tau, \tau') = \sum_{\sigma, \sigma', \tau'} p(\sigma)p(\sigma')Q(\sigma, \sigma'; \tau, \tau')$$

Evolution equation:

$$\partial_t p_t = \Psi[p_t] - p_t, \quad t \geqslant 0, \quad p_0 = p.$$

Fix $\nu(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{A=\{i\}}$ [as in Gibbs sampler].

Lemma (Convergence to equilibrium: H-theorem)

Fix a graph finite G, $\beta \in \mathbb{R}$, and $p \in \mathrm{Prob}(\Omega)$. Then $p_t \to \mu = \mu_{G,\beta,\mathbf{h}}$, as $t \to \infty$, where $\mathbf{h}$ is the unique set of external fields s.t. $\mu_i = p_i$ for all $i \in [n]$.

Nonlinear Stochastic Ising Model II

The Markov chain yields a linear map on $\mathrm{Prob}(\Omega \times \Omega)$. The chain is not irreducible. By projection one has the quadratic map on $\mathrm{Prob}(\Omega)$: $p \mapsto \Psi[p]$

$$\Psi[p](\tau) = \sum_{\tau'} [(p \otimes p)Q](\tau, \tau') = \sum_{\sigma, \sigma', \tau'} p(\sigma)p(\sigma')Q(\sigma, \sigma'; \tau, \tau')$$

Evolution equation:

$$\partial_t p_t = \Psi[p_t] - p_t, \quad t \geqslant 0, \quad p_0 = p.$$

Fix $\nu(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{A=\{i\}}$ [as in Gibbs sampler].

Lemma (Convergence to equilibrium: H-theorem)

Fix a graph finite G, $\beta \in \mathbb{R}$, and $p \in \mathrm{Prob}(\Omega)$. Then $p_t \to \mu = \mu_{G,\beta,\mathbf{h}}$, as $t \to \infty$, where $\mathbf{h}$ is the unique set of external fields s.t. $\mu_i = p_i$ for all $i \in [n]$.

[dynamics indep. of $\mathbf{h}$; if $\beta = 0$ we're back to recombination model]

Nonlinear Stochastic Ising Model III

In analogy with linear Gibbs sampler and nonlinear recombinations:

Conjecture There exists universal $c > 0$ s.t. for any graph G, any β: $|\beta| \leqslant c/\Delta$, where $\Delta = \max \deg(G)$,

$$H(p_t|\mu) \leqslant H(p|\mu)\, e^{-c\,t/n}, \quad t \geqslant 0. \tag{3}$$

for any initial $p \in \mathrm{Prob}(\Omega)$.

Nonlinear Stochastic Ising Model III

In analogy with linear Gibbs sampler and nonlinear recombinations:

Conjecture There exists universal $c > 0$ s.t. for any graph G, any β: $|\beta| \leqslant c/\Delta$, where $\Delta = \max \deg(G)$,

$$H(p_t|\mu) \leqslant H(p|\mu)\, e^{-c\,t/n}, \quad t \geqslant 0. \tag{3}$$

for any initial $p \in \mathrm{Prob}(\Omega)$.

A direct functional-analytic approach as for the $\beta = 0$ case seems to be very difficult. On the other hand, we hope to carry out some parts of Kac's program for the Boltzmann equation.

Nonlinear Stochastic Ising Model III

In analogy with linear Gibbs sampler and nonlinear recombinations:

Conjecture There exists universal $c > 0$ s.t. for any graph G, any β: $|\beta| \leqslant c/\Delta$, where $\Delta = \max \deg(G)$,

$$H(p_t|\mu) \leqslant H(p|\mu)\, e^{-c\,t/n}, \quad t \geqslant 0. \tag{3}$$

for any initial $p \in \mathrm{Prob}(\Omega)$.
A direct functional-analytic approach as for the $\beta = 0$ case seems to be very difficult. On the other hand, we hope to carry out some parts of Kac's program for the Boltzmann equation.

Remark: Adding a dissipative term (mutations, spin flip):

$$Q \to Q' = \frac{1}{2} Q + \frac{1}{2} Q_{\mathrm{Glauber}},$$

then conjecture holds true.

Implementing Kac' program ?

Work in progress with Arnaud Guillin:

1. Establish a particle system representation for which the nonlinear equation is a limiting object;
2. Prove entropy production estimates for the particle system that remain tight in the limit;
3. Establish propagation of chaos and so-called entropic chaos.
4. Deduce entropy decay for the nonlinear system.

See Carlen-Carvalho-Le Roux-Loss-Villani 2008 for a review of Kac's program for the Boltzmann-Kac equation.

Hope: the discrete combinatorial nature should allow us to avoid all the regularities issues that make the program very difficult in the kinetic theory setting.

Implementing Kac' program ?

Work in progress with Arnaud Guillin:

1. Establish a particle system representation for which the nonlinear equation is a limiting object;
2. Prove entropy production estimates for the particle system that remain tight in the limit;
3. Establish propagation of chaos and so-called entropic chaos.
4. Deduce entropy decay for the nonlinear system.

See Carlen-Carvalho-Le Roux-Loss-Villani 2008 for a review of Kac's program for the Boltzmann-Kac equation.

Hope: the discrete combinatorial nature should allow us to avoid all the regularities issues that make the program very difficult in the kinetic theory setting.

THANK YOU!