Some Statistics from Lattice Statistical Mechanics

Ian G. Enting

The University of Melbourne

MASCOS

(ARC Center of Excellence for Mathematics and Statistics of Complex Systems)

Summary 2

An exploratory talk

Business-as-Usual statistical mechanics

Problems of statistical inference:

- calibration, reconstruction
  - 1-D ising
  - 2-D (Gibbs-Markov fields)
  - Pickard random fields

Gibbs distribution

Probability of microstate $x_1, x_2, \ldots x_N$

$$\Pr(x_1, x_2, \ldots x_N)$$

$$= \frac{1}{Z} \exp\left[-E(x_1, x_2, \ldots x_N)/k_{\mathrm{B}}T\right]$$

$E(x_1, x_2, \ldots x_N)$ is energy of micro-state

Z, the partition function, is the normalising factor for the distribution

k_{B} is Boltzmann's constant $=1.3806 \times 10^{-23}\mathrm{J/K}$

T is absolute temperature.

Thermodynamics

Thermodynamic quantities
= moments of Gibbs distribution
= derivatives of Z

$$\langle E \rangle = \sum E(x_1, x_2, \ldots x_N) \times \mathrm{Pr}(x_1, x_2, \ldots x_N)$$

Free energy $= -k_{\mathrm{B}} T \ln(Z)$

For example, Ising model ($x_j = \pm 1$) represents the physics of transitions in:

* binary alloys (beta-brass);
* liquid gas critical point
* anisotropic magnetism.

Solving lattice statistics problems

- Exact solutions (Z, low-order moments)
- Closed form aproximations (statistical closures)
- Exact series expansions: high-temperature — high-field (low-T)
- Renormalisation group: real space — reciprocal space
- Monte Carlo: Markov Chain Monte Carlo (MCMC) can be defined in terms of relative probabilities: — don't need Z

Random fields

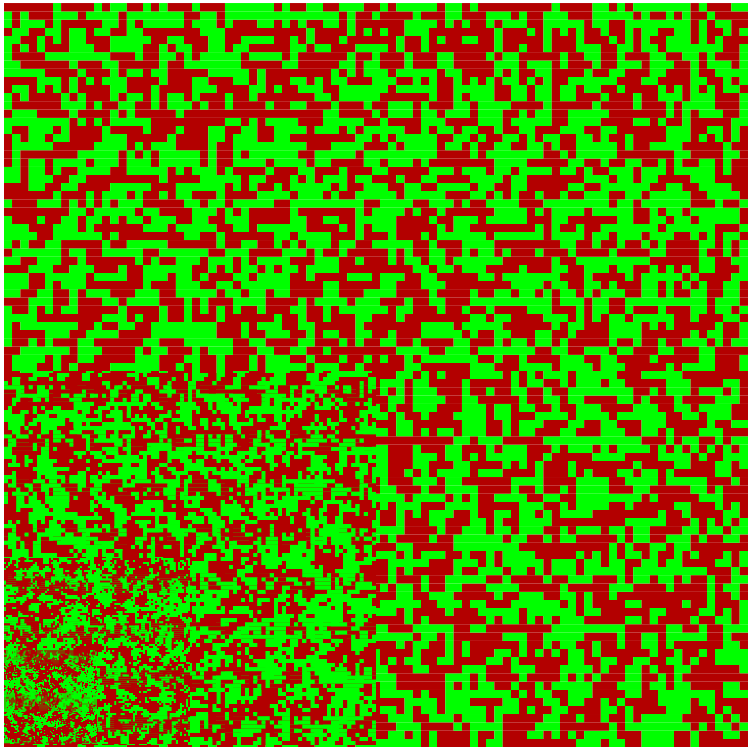
More general models where probabilities do not reflect energy or temperature.

Gibbs specification in terms of joint probability

Markov specification of probabilities, conditional on states of neighbours,

Hammersley-Clifford theorem shows equivalance of the Gibbs and Markov forms (as long as no probabilities are zero).

Critical points, scaling & fractals



- Special points — occurring as first-order transition disappears; *OR*
- Regime of critical behaviour — massless phases with Kosterlitz-Thouless transitions;

'*OR* "The ubiquity of the critical state may well be regarded as the first really solid discovery of complexity theory" — M Buchanan in *Ubiquity: The Science of History*, referring to self-organised criticality.

Physics, beyond thermodynamics

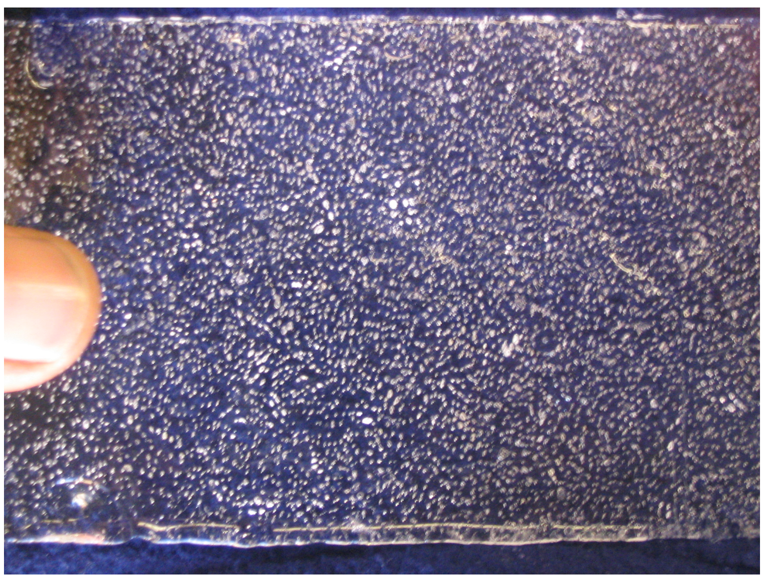
- Percolation:
  $q \to 1$ limit of q-state Potts model
  (Wu, Fortuin and Kasteleyn), but no longer
  related to temparture
  — numerous applications

- DLA: diffusion-limited aggregation

- geomorphology — percolation-related?

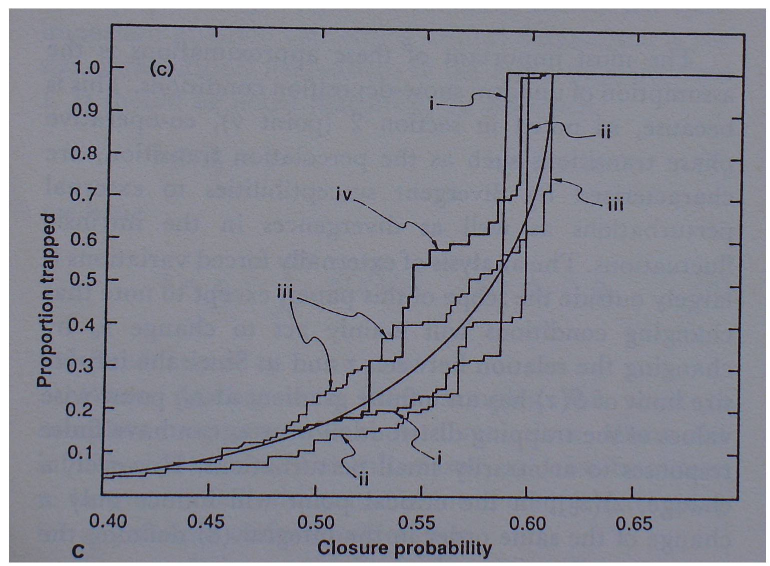
Percolation transition: Firn-to-ice

Past atmospheric composition can be measured in air bubbles trapped in polar ice.



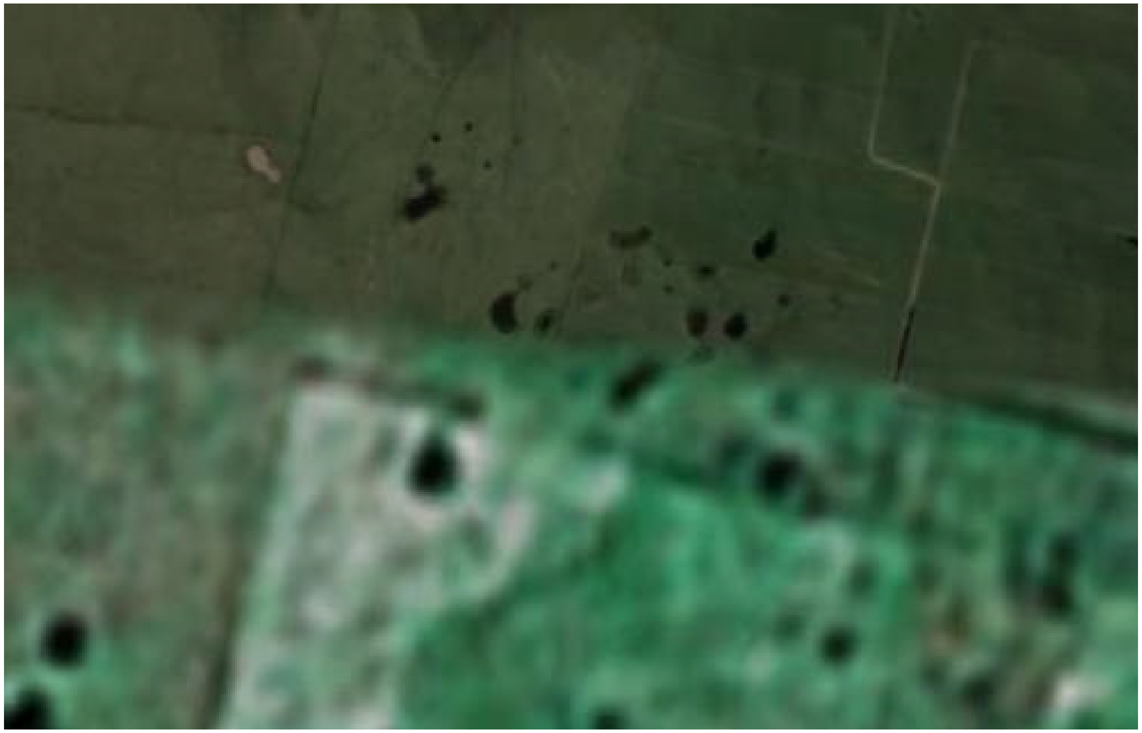
Photograph from CSIRO Atmospheric Research

Model bubble trapping as random closure of channels. ‘Critical fluctuations’ in trapped volume at close-off. (Simulations from Enting, J.Glaciol. **39** 133–142 (1993).)



Satellite data?

Statistical modelling to reconcile mutliple data streams.



Example from *Google Earth* (near Otway geosequestration test site)

Role in Earth System Science

- Use statistical models of heterogeneous systems:
  mechanistic models of behaviour of classes combined with statistical models of distribution, rather than spatially-explicit mechanistic modelling

- 'Null models' for testing significance of heterogeneity in observational data – especially aircraft/satellite remote sensing.

Aspects of Modelling

Simple model of response to forcing (Volterra eqn)

$$x(t) = \int_0^t R(t-t')\, f(t')\, dt' = \int_0^t R(t'')\, f(t-t'')\, dt''$$

Three forms of model application:

- calculate $x(t)$ given $R(t)$ and $f(t)$
  Forward model

- calculate $R(t)$ given $x(t)$ and $f(t)$
  calibration

- calculate $f(t)$ given $x(t)$ and $R(t)$
  deconvolution — data assimilation

Atmospheric CO_2 (century scale)

CO_2 concentration:

$$C(t) = C(t_0) + \int_{t_0}^{t} R(t-t')\, S_{\text{Fossil+land-use}}(t)\, dt'$$

For $t_0 \leq 1800$, $C(t_0) \approx 280$ ppm.

Formal equivalence between estimating $S(t)$ given $R(t)$, $C(t)$ and estimating $R((t)$ given $S(t)$, $C(t)$, but different forms of function expected.

(Actual model calibrations, i.e. estimates of $R(t)$, rely heavily on radio-carbon data).

Hidden Markov terminology

Hidden Markov Models, e.g. as applied to speech processing, have comparable three problems:

- forward problem:
  simulation?

- inverse problem:
  calibration = learning or training

- inverse problem:
  deconvolution = decoding

Lattice statistics modelling

Forward and inverse problems in Gibbs-Markov fields.

forward problem :

Usually hard to construct a realisation of the probability distribution (except by Markov Chain Monte Carlo).

calibration problem :

Maximum Likelihood is usually hard — partition function unknown.

Maximum conditional likelihood formally easier.

reconstruction :

Generally done as Markov Chain Monte Carlo.

Parameter estimation

Log likelihood is:

$\ell = -E(x_1, \ldots x_N)/kT - \ln Z$

(i.e. $-S/k$ in thermodynamic terms —

i.e. maximum likelihood = minimum entropy)

For $E(x_1, \ldots x_N)/kT = \sum_\alpha K_\alpha f_\alpha(x_1, \ldots x_N)$,

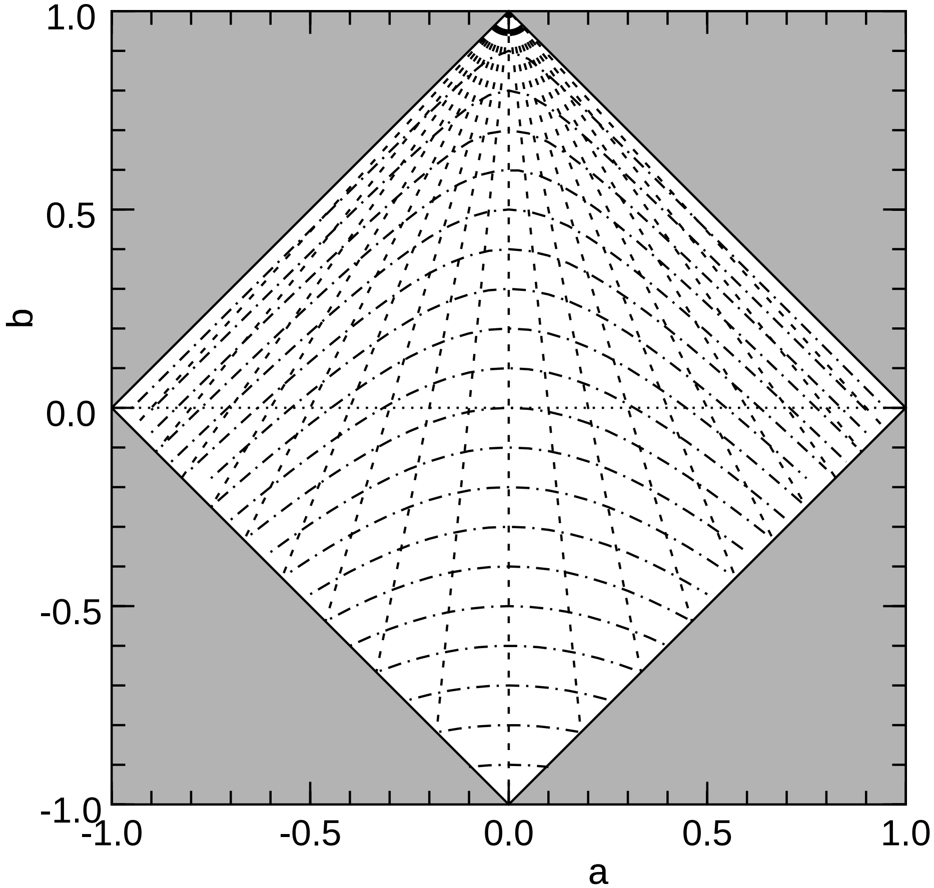
maximum likelihood estimates satisfy

$$\frac{\partial \ell}{\partial K_\alpha} = f_\alpha(x_1, \ldots x_N) - \langle f_\alpha \rangle$$

i.e. the maximum likelihood estimates of interaction strengths, K_α, are those that give expectations of spin products, f_α that equal observed averages.

Maximum likelihood is equivalent to fitting moments.

1-D Ising: Inference



Parameterise as Markov chain: $\Pr(\sigma_n|\sigma'_n, n' < n) = \frac{1}{2}[1 + \sigma_n(a + b\,\sigma_{n-1})]$ Plots shows contours of $\langle\sigma_n\rangle$ (straight lines) and $\langle\sigma_n\sigma_{n+1}\rangle$ in the a, b plane.

1-D Ising: reconstruction

$$\begin{array}{ccc} \sigma_1 & — & \nu_1 \\ | & & \\ \sigma_2 & — & \nu_2 \\ | & & \\ \sigma_3 & — & \nu_3 \\ | & & \\ \sigma_4 & — & \nu_4 \\ | & & \\ \sigma_5 & — & \nu_5 \\ | & & \\ \sigma_6 & — & \nu_6 \\ | & & \\ \sigma_7 & — & \nu_7 \\ | & & \\ \sigma_8 & — & \nu_8 \\ | & & \end{array}$$

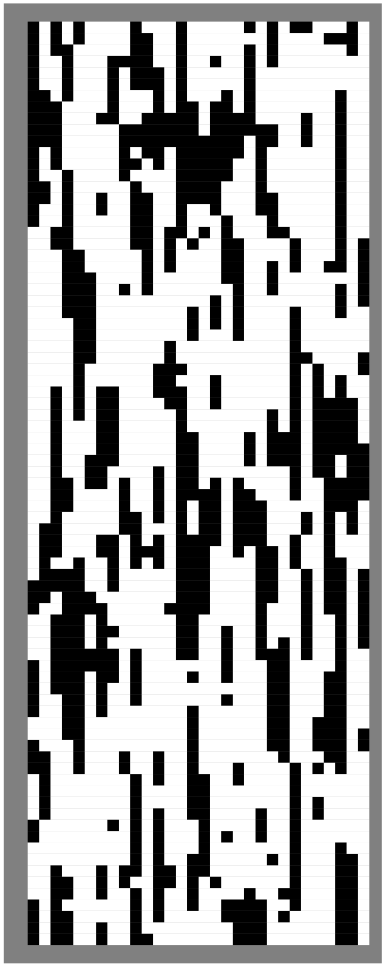
Represent reconstruction of Ising state, $\sigma_1 \ldots \sigma_N$, from noisy data, decorated model with $\nu_1, \ldots \nu_N$ as the degraded observations.

$$\Pr(\sigma_n | \nu_n, \sigma_{n-1}) = \frac{\Pr(\sigma_n | \sigma_{n-1}) \Pr(\nu_n | \sigma_n)}{\Pr(\nu_n | \sigma_{n-1})}$$

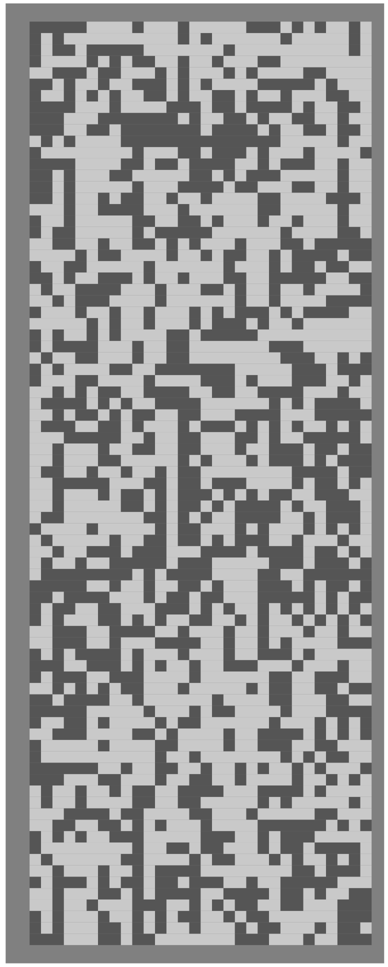
This gives a recursive one-sided procedure for state-estimation as $\Pr(\sigma_n | \nu_1 \ldots \nu_n)$

$$= \sum_{\sigma_{n-1} = \pm 1} \Pr(\sigma_n | \nu_n, \sigma_{n-1}) \Pr(\sigma_{n-1} | \nu_1, \ldots, \nu_{n-1})$$

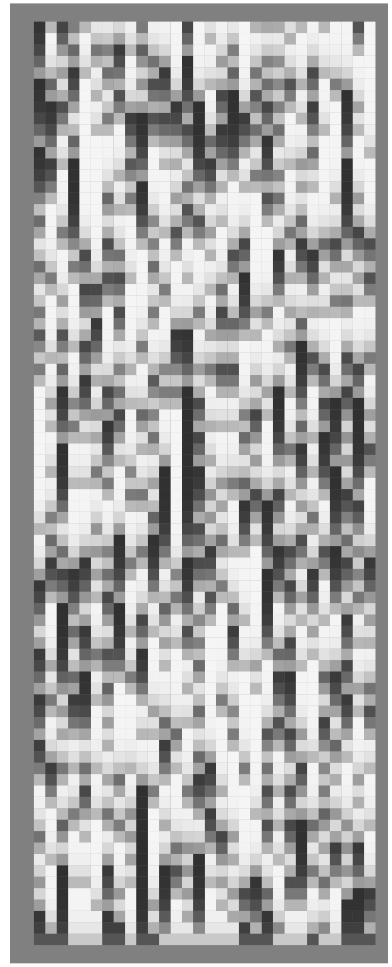
1-D Ising reconstruction



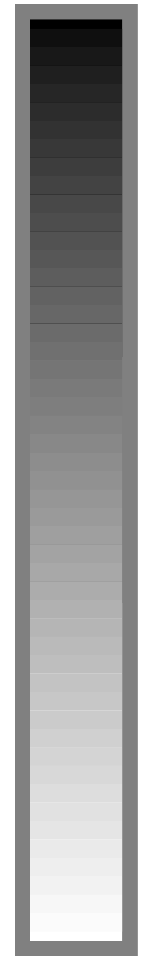
Time upwards, independent series



Noise on 50% of data.



Reconstruction as probabilities



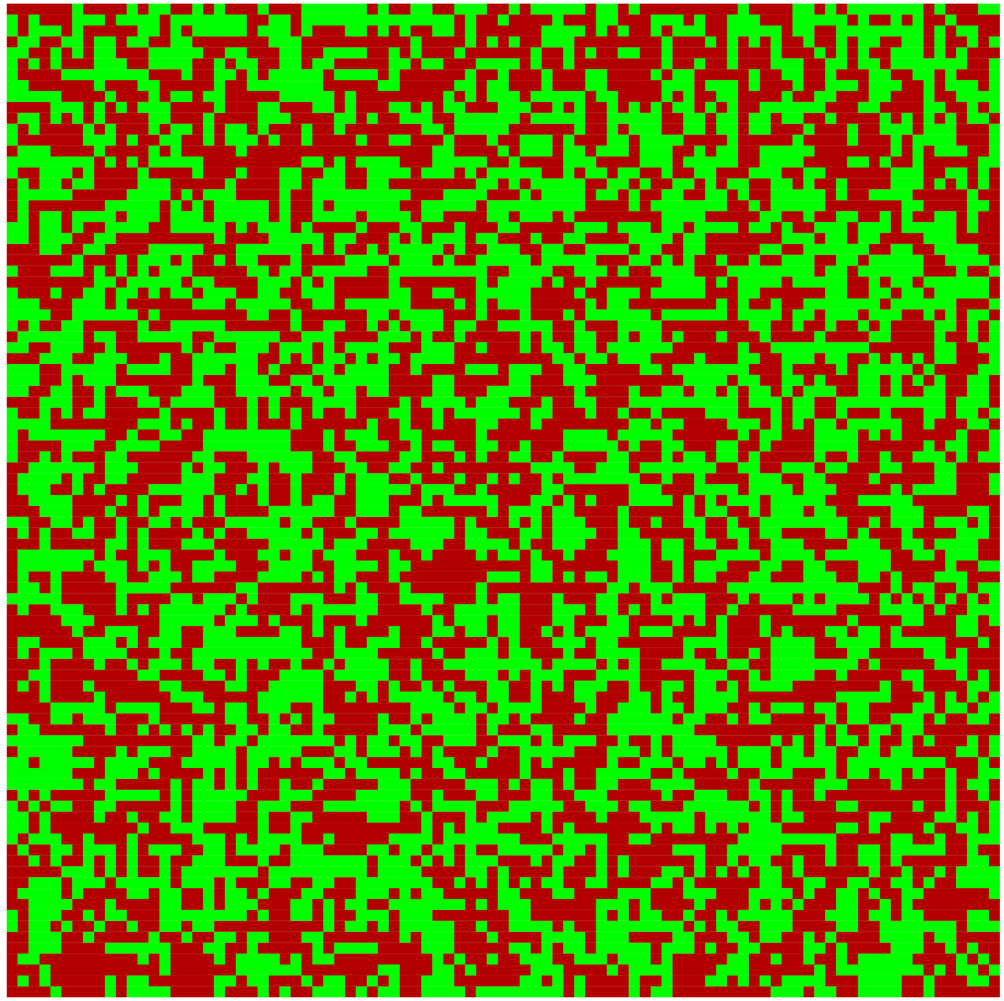
$\Pr(\sigma = 1)$ as greyscale

Inferring 2-D Ising interactions

Maximum likelihood: choose $\{K_\alpha\}$ so that for spin products $f_\alpha(x_1, \ldots x_N)$,

$$\langle f_\alpha \rangle = f_\alpha(x_1, \ldots x_N)_{\text{sample}}$$

The problems is that usually the $\langle f_\alpha \rangle$ are not known (as functions of the $\{K_\alpha\}$).

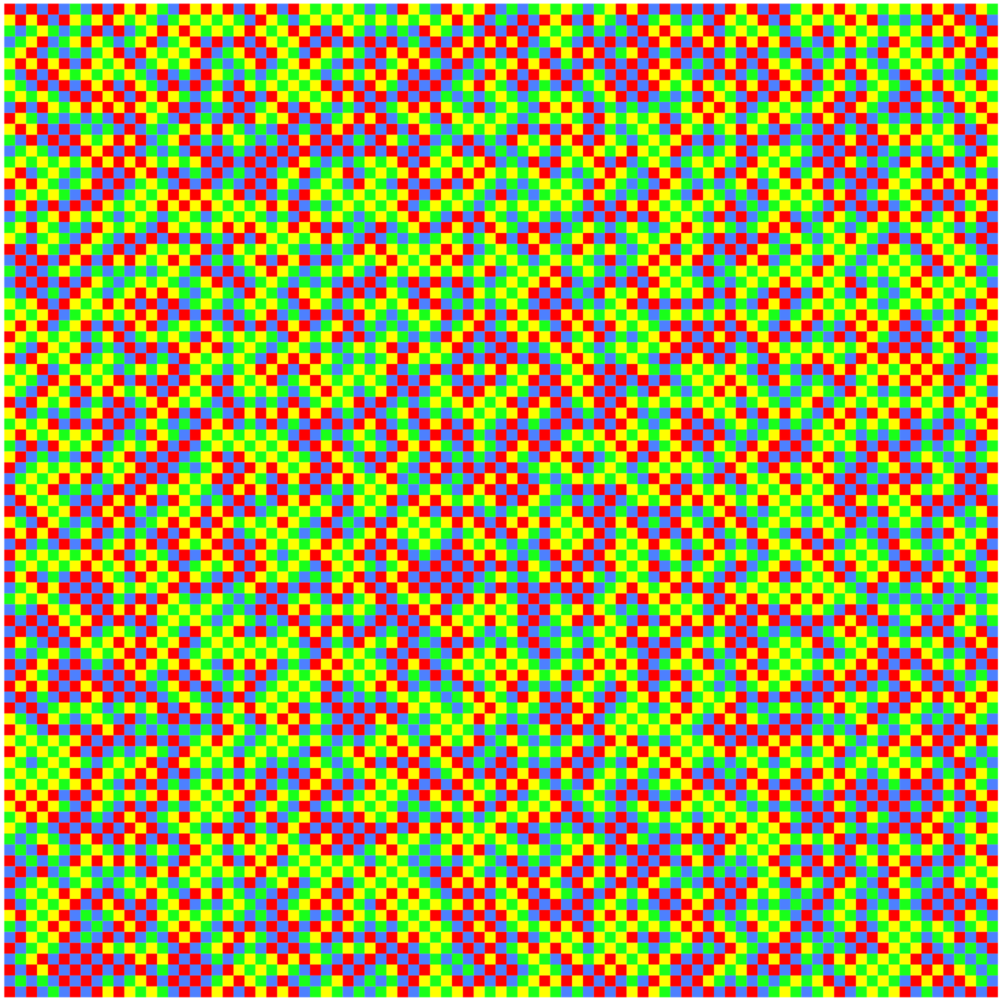


Conditional likelihoods

J. Besag: maximise the likelihood of one set of spins, conditional on others.

$$\Pr(x_j, j \in A | x_k, k \in B)$$

e.g. maximum (conditional) likelihood of green/red distribution, given blue/yellow sites). Loss of statistical efficiency and still potential for poor convergence, but computationally simple.



Beyond Besag

.	.	a1	.	.
.	a2	X1	a3	.
a4	X2	X3	X4	a5
.	a6	X5	a7	.
.	.	a8	.	.

Joint probability of X1 to X5 given a1 to a8 readily calculated.

Inference on 5/9 of sites (red or green), conditional on the other 4/9 (blue or yellow).

(Each site aj occurs as boundary of 2 sets of Xn.)

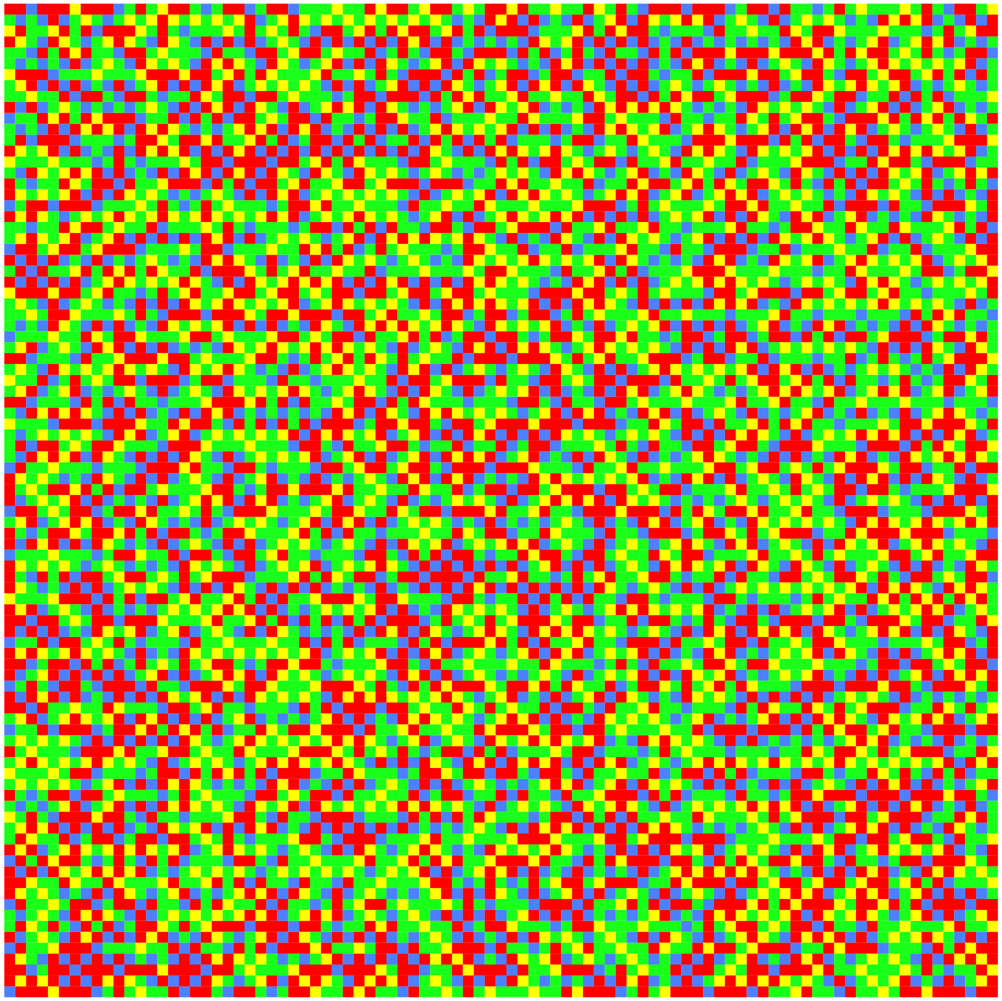


Image reconstruction

‘On the statistical analysis of dirty pictures’.
J. Besag, *J. Roy. Statist. Soc.*, **48**, 259–302.

Binary (i.e. Ising) models mainly relevant for classification, but also electron-micrography.

Markov Chain Monte Carlo techniques can sample posterior distributions either for reconstruction (posterior distribution as Gibbs-Markov field) or for ‘feature detection’ with Gibbs-Markov as ‘null model.

Directed problems

Crystal growth :

growing mixed disordered crystals

— a moving surface of a growing crystal

Stochastic cellular automata (SCA) :

— link to Wolfram (*A New Kind of Science*???)

Ising model disorder points :

— mapping back onto Gibbs Markov fields reveals 'hidden' symmetries (Enting 75).

Directed percolation :

— a special case, extensively studied

— in SCA terms a partly deterministic limit, since some probabilities are zero.

Stochastic cellular automata

Define a distribution in terms of conditional probabilities that can be used explicitly.

For $(0 \ldots M) \times (0 \ldots N)$ define a stochastic cellular automaton by:

$\Pr(\sigma_{m,n} | \sigma_{m-1,n}, \sigma_{m,n-1}, \sigma_{m-1,n-1}) =$
$\Pr(D|C, B, A)$.

Build up the random field, from boundaries, (0,0) to (0,N) and (0,0) to (M,0) by successive application of $\Pr(D|C, B, A)$.

Used to model disordered mixed crystals by Welberry. Various special solutions found.

Pickard fields

Reproducing a 1-D Markov chain across a planar lattice.

- D.J. Pickard (Statistics: R.S.Soc.Sci.): unilateral Markov fields
- R.J. Baxter (Theoretical Physics, R.S.Phys.S.): approximate eigenvectors of transfer matrix as products — variational approximations $\to$ corner transfer matrix
- T.R. Welberry (R.S.Chem): statistics of mixed crystals — 2-site correlations c.f. X-ray diffraction
- I.G. Enting (Theoretical Physics, R.S Phys.S.): Map Welberry conditional form onto generalised Ising models to reveal implicit symmetries.

ANU, Canberra 1976–77

Pickard specifications

- Applies to $(0 \ldots M) \times (0 \ldots N)$
- General stochastic cellular automaton defines: $\Pr(\sigma_{m,n} | \sigma_{m-1,n}, \sigma_{m,n-1}, \sigma_{m-1,n-1}) = \Pr(D|C,B,A)$
- Pickard defines boundaries $\sigma_{0,0}$ to $\sigma_{M,0}$ and $\sigma_{0,0}$ to $\sigma_{0,N}$ as Markov Chains.
- Using $\perp$ to denote independence, Pickard's constraints are $B \perp C|A$ along with:
  - $A \perp D|C$ and $A \perp D|B$ Case A
  - OR
  - $B \perp C|D$ Case B
- These conditions ensure stationarity over plane.
- Both lead to 'reversibility' – case A has rectangular symmetry.

Some results for Pickard fields

- Exponential decay of correlations:
$\langle \sigma_{n,m}\sigma_{0,0}\rangle - \langle \sigma_{n,m}\rangle\langle \sigma_{0,0}\rangle$
$= \alpha^{|m|}\beta^{|n|} \qquad \forall m, n$ Case A
$= \alpha^{|m|}\beta^{|n|} \qquad$ for $m \times n \geq 0$ case B

- Pair correlations don't uniquely characterise field (i.e. X-ray diffraction does not tie down structure). Even for spibn-symmetric cases, there are families whose members have identical pair correlations but differ in their 4-site correlations

Modelling with Pickard fields

- Modelling:
  — simple to construct explicit realisations by applying the defining rule

- Calibration:
  — match the defining probabilities to equal sample statistics??

- Reconstruction: — closed-form probabilities facilitate both Monte Carlo and maximum likelihood.

Symmetries

Mapping the SCA form onto the Gibbs-Markov form revealed ‘hidden symmetries (Enting, 1975) and led to additional closed form solutions of correlations (Welberry).

Special subcases have triangular lattice neighbourhood, $D \perp A|B, C$, and various symmetries lead to a set of cases with exact solutions, mostly identified in isolation as special Ising model cases.

Future directions?

- Applications to land-surface modelling
- Higher-order Pickard fields
- Use of Stochastic cellular automata (SCA) in the design of deterministic cellular automata (Rujan 1987).
  — looking in the Gibbs parameter space, 'interesting' cellular automata can occur when the SCA manifold approaches a critical manifold as probabilities $\rightarrow$ 0 or 1.

Beyond Pickard?

Pickard results not restricted to binary variables.

Can Pickard results produce a higher-order field by using variables with 2^{m^2} states?.

i.e. reproduce $m \times m$ squares of binary variables — overlapping to allow stationarity?

Is this possible?

Does it corrrespond to higher-order versions of corner-transfer-matrix approximations?

Is it useful? — bounds?

THANK YOU

Further information

- D.K. Pickard: Unilateral Markov Fields. *Adv. Appl. Prob,*, **12**, 655–671, 1980.
- T.R. Welberry. Diffuse X-ray scattering and models of disorder. *Rep. Prog. Phys.*, **48**, 1543–1593, 1985.
- N.A.C. Cressie *Statistics for Spatial Data*, Wiley, NY, 1993.
- P. Rujan. Cellular automata and statistical mechanics models. *J. Statistical Physics*, **49**, 139–222, 1987.

Comments on graphics

Hand-coded postscript.

Two main types:

1: Rendering a simulation from other software

— can render one data set in multiple ways

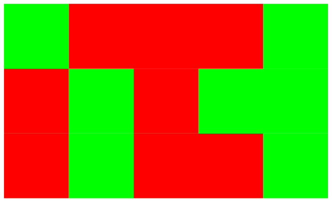
2: Simulation within the postscript

use internal random numbers in PS-interpreter

(in printer, viewer, pdf converter, etc)

Description in *Aust. M. S. Gazette*, **33**, 131 (2006). Enting's review of *Mathematical Illustrations: A Manual of Geometry and PostScript* by Bill Casselmann

Render external simulation



```
%!PS-Adobe-3.0 EPSF-3.0
%%BoundingBox: 90 90 565 560
% File monte.ps1 Plot Ising simulation
/dx 5 def /yc 100 def /x0 100 def /xc x0 def
dx setlinewidth
/A { 1 0 0 setrgbcolor block} def
/B {0 1 0 setrgbcolor block} def
/LF { /xc x0 def yc dx add /yc exch def} def
/block {newpath xc yc moveto dx 0 rlineto stroke xc dx
add /xc exch def} def
% End of file monte.ps1
% File monte0.psd: Dummy test data
A B A A B LF A B A B B LF B A A A B
showpage
```

Internal simulation

```
/unitrand {rand 65536 div 32768 div } def
/withinp {unitrand le } def
/setspin { withinp {-1} {1} ifelse}  def
.......
/pcase { 1 eq {A} {B} ifelse } def
/Print { {pcase} forall LF} def
...
/ycalc { bb exch dup dup aa xget exch 1 sub dup aa xget
exch bb xget func setspin put }  def
...
/Loop
{ bb 0 aa 0 get dim1 put 1 1 nsca {ycalc} for bb Print
aa 0 bb 0 get dim1 put 1 1 nsca {xcalc} for aa Print }  def
...
1 1 25 Loop for
```