Ising models and multiresolution quad-trees

Wilfrid S. Kendall
Department of Statistics
University of Warwick

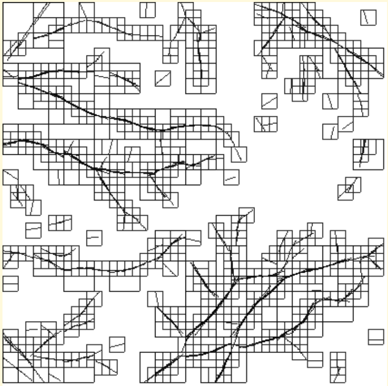
Roland G. Wilson
Department of Computer Science
University of Warwick

15 August 2002

Research supported by EPSRC research grant GR/M75785

ssa.tex 1.17 2002/10/22 10:35:15 Wilfrid W2k Wilfrid

Introduction



Multiresolution quad-trees are used in image analysis. For example: given a layer of pixels at finest resolution, successively aggregate blocks of pixels to produce coarser layers. A simple model then stipulates the black/white value of a pixel depends on its neighbours in the same layer and (with a different interaction strength) on its parent and daughters.

This talk describes initial steps in understanding the qualitative behaviour of this algorithm, addressing the following question:

> "What do phase transitions look like when the finest layer of pixels is unconstrained ("free boundary")?"

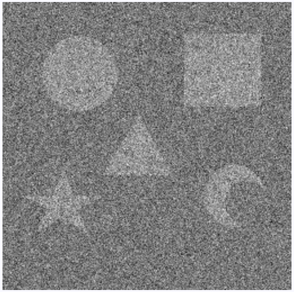
For details, see Kendall and Wilson (2002).

WARWICK
Image analysis
Generalized quad-trees
Random clusters
Percolation
Comparison
Simulation
Future work
References
Home Page
Title Page
Page 3 of 29
Go Back
Full Screen
Close
Quit

1. Image analysis

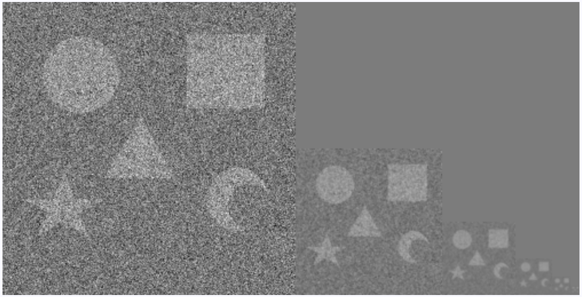
Wilson and Li (2002): Segmentation of noisy shapes.
Bhalerao et al. (2001), Thönnes et al. (2002): application to MCMC in medical imaging.

Multiresolution MAP algorithm, 1.3% misclassification:



$\Longleftrightarrow$

Quad-tree formed by successive averaging using "decimation":



Multiresolution: pros and cons

+ **FAST** since low resolution "steers" high resolution;

+ adapted to some kinds of **HIGH-LEVEL** objects;

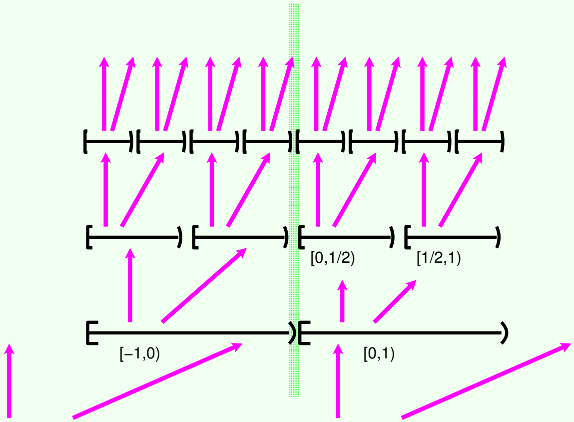
– can produce **"BLOCKY"** reconstructions:
  resolution hierarchy mediates all spatial interactions.

Possible solution

Add further explicitly *spatial* interactions?

2. Generalized quad-trees

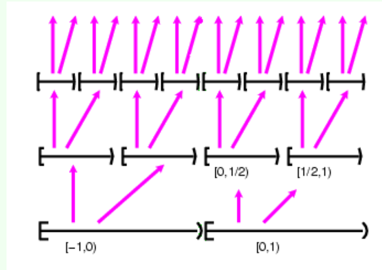
Define $\mathbb{Q}_d$ as graph whose vertices are cells of all dyadic tessellations of $\mathbb{R}^d$, with edges connecting each cell to its $2d$ neighbours, and also its parent (covering cell in tessellation at next resolution down) and its 2^d daughters (cells which it covers in next resolution up).



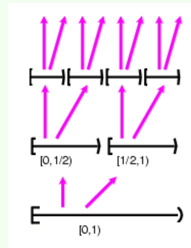
Case $d = 1$:
Neighbours at same level also are connected.

Remark: No spatial symmetry!

Further define:



- $\mathbb{Q}_{d;r}$ as subgraph of $\mathbb{Q}_d$ at resolution levels of r or higher;



- $\mathbb{Q}_d(\mathbf{o})$ as subgraph formed by $\mathbf{o}$ and all its descendants.

- **Remark:** there are many graph-isomorphisms between $\mathbb{Q}_{d;r}$ and $\mathbb{Q}_{d;s}$, with natural $\mathbb{Z}^d$-action;

- **Remark:** there are graph homomorphisms injecting $\mathbb{Q}(\mathbf{o})$ into itself, sending $\mathbf{o}$ to $\mathbf{x} \in \mathbb{Q}(\mathbf{o})$ *(semi-transitivity)*.

WARWICK
Image analysis
Generalized quad-trees
Random clusters
Percolation
Comparison
Simulation
Future work
References
Home Page
Title Page
◀◀ ▶▶
◀ ▶
Page 8 of 29
Go Back
Full Screen
Close
Quit

Simplistic analysis

Define J_λ to be strength of neighbour interaction, J_τ to be strength of parent interaction. If $S_x = \pm 1$ then probability of configuration is proportional to $\exp(-H)$ where

$$H = -\frac{1}{2} \sum_{\langle x,y \rangle \in \mathcal{E}(G)} J_{\langle x,y \rangle} (S_x S_y - 1)\,, \qquad (1)$$

for $J_{\langle x,y \rangle} = J_\lambda, J_\tau$ as appropriate.

If $J_\lambda = 0$ then the free Ising model on $\mathbb{Q}_d(\mathbf{o})$ is a *branching process* (Preston 1977; Spitzer 1975); if $J_\tau = 0$ then the Ising model on $\mathbb{Q}_d(\mathbf{o})$ decomposes into sequence of d-dimensional classical (finite) Ising models. So we *know* there is a phase change at $(J_\lambda, J_\tau) = (0, \ln(5/3))$ (branching processes), and *expect* one when $\lambda = 0+$, indeed at $(J_\lambda, J_\tau) = (\ln(1+\sqrt{2}), 0+)$ (2-dimensional Ising).

But is this all that there is to say?

3. Random clusters

A similar problem, concerning Ising models on products of trees with Euclidean lattices, is treated by Newman and Wu (1990). We follow them by exploiting the celebrated *Fortuin-Kasteleyn random cluster representation* (Fortuin and Kasteleyn 1972; Fortuin 1972a; Fortuin 1972b):

The Ising model is the marginal site process at $q = 2$ of a site/bond process derived from a dependent bond percolation model with configuration probability $\mathbb{P}_{q,p}$ proportional to

$$q^C \times \prod_{\langle x,y\rangle \in \mathcal{E}(G)} \left((p_{\langle x,y\rangle})^{b_{\langle x,y\rangle}} \times (1 - p_{\langle x,y\rangle})^{1-b_{\langle x,y\rangle}} \right) .$$

(where $b_{\langle x,y\rangle}$ indicates whether or not $\langle x, y\rangle$ is closed, and C is the number of connected clusters of vertices). Site spins are chosen to be the same in each cluster independently of other clusters with equal probabilities for ± 1.

Random cluster facts

- (Representation of Ising model.) The marginal bond process is Ising with

$$p_{\langle x,y\rangle} = 1 - \exp(-J_{\langle x,y\rangle})\,; \tag{2}$$

- (FK-comparison inequalities.) If $q \geq 1$ and A is an *increasing* event then

$$\mathbb{P}_{q,p}(A) \leq \mathbb{P}_{1,p}(A) \tag{3}$$

$$\mathbb{P}_{q,p}(A) \geq \mathbb{P}_{1,p'}(A) \tag{4}$$

where

$$p'_{\langle x,y\rangle} = \frac{p_{\langle x,y\rangle}}{p_{\langle x,y\rangle} + (1 - p_{\langle x,y\rangle})q} = \frac{p_{\langle x,y\rangle}}{q - (q-1)p_{\langle x,y\rangle}}\,.$$

Since $\mathbb{P}_{1,p}$ is bond percolation (bonds open or not independently of each other), we can find out about phase transitions by studying *independent* bond percolation.

4. Percolation

Independent bond percolation on products of trees with Euclidean lattices have been studied by Grimmett and Newman (1990), and these results were used in the Newman and Wu work on the Ising model. So we can make good progress by studying independent bond percolation on $\mathbb{Q}_d$, using p_τ for parental bonds, p_λ for neighbour bonds.

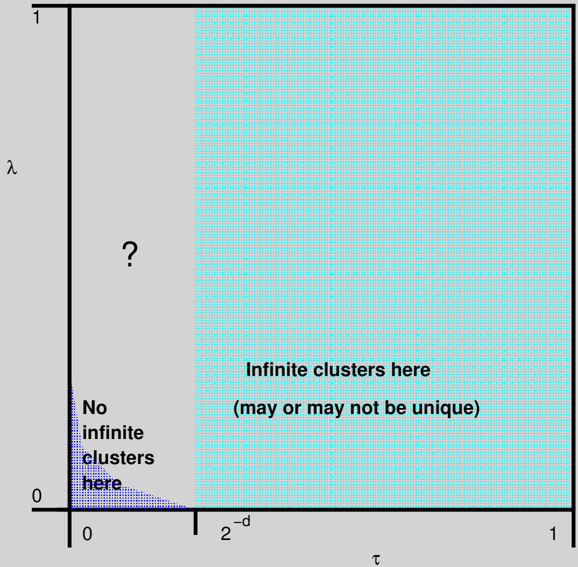
Theorem 1 *There is almost surely no infinite cluster in $\mathbb{Q}_{d;0}$ (and consequently in $\mathbb{Q}_d(\mathbf{o})$) if*

$$2^d \tau \mathcal{X}_\lambda \left(1 + \sqrt{1 - \mathcal{X}_\lambda^{-1}}\right) \quad < \quad 1\,,$$

where $\mathcal{X}_\lambda$ is the mean size of the percolation cluster at the origin for λ-percolation in $\mathbb{Z}^d$.

Modelled on Grimmett and Newman (1990, §3 and §5).

Get $\left(1 + \sqrt{1 - \mathcal{X}_\lambda^{-1}}\right)$ from matrix spectral asymptotics.



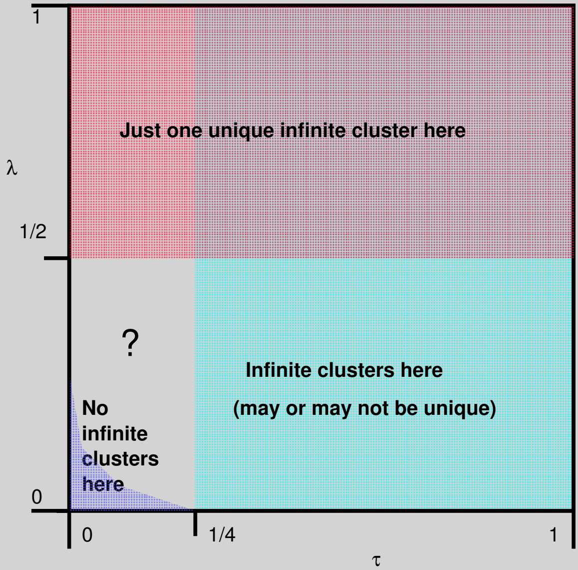
The story so far: small λ, small to moderate τ.

Case of small τ

Need $d = 2$ for mathematical convenience. Use Borel-Cantelli argument and planar duality to show, for supercritical $\lambda > 1/2$ (that is, supercritical with respect to planar bond percolation!), all but finitely many of the resolution layers $L_n = [1, 2^n] \times [1, 2^n]$ of $\mathbb{Q}_2(\mathbf{o})$ have just one large cluster each of diameter larger than constant $\times\, n$.

Hence ...

Theorem 2 *When* $\lambda > 1/2$ *and* τ *is* positive *there is one and only one infinite cluster in* $\mathbb{Q}_2(\mathbf{o})$.



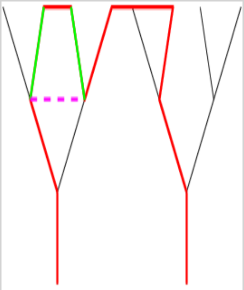
The story so far: adds small τ for case $d = 2$.

Uniqueness of infinite clusters

The Grimmett and Newman (1990) work was remarkable in pointing out that as τ increases so there is a *further* phase change, from many to just one infinite cluster for $\lambda > 0$. The work of Grimmett and Newman carries through for $\mathbb{Q}_d(\mathbf{o})$. However the relevant bound is *improved* by a factor of $\sqrt{2}$ if we take into account the hyperbolic structure of $\mathbb{Q}_d(\mathbf{o})$!

Theorem 3 *If $\tau < 2^{-(d-1)/2}$ and $\lambda > 0$ then there cannot be just one infinite cluster in $\mathbb{Q}_{d;0}$.*

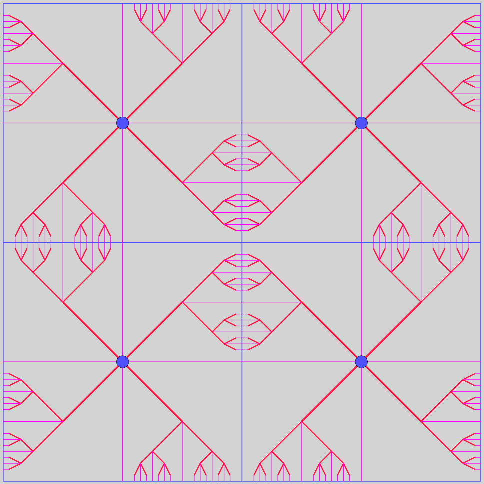
Method: sum weights of "up-paths" in $\mathbb{Q}_{d;0}$ starting, ending at level 0. For fixed s and start point there are infinitely many such up-paths containing s λ-bonds; but no more than $(1 + 2d + 2^d)^s$ which cannot be reduced by "shrinking" excursions. Hence control the mean number of open up-paths stretching more than a given distance at level 0.

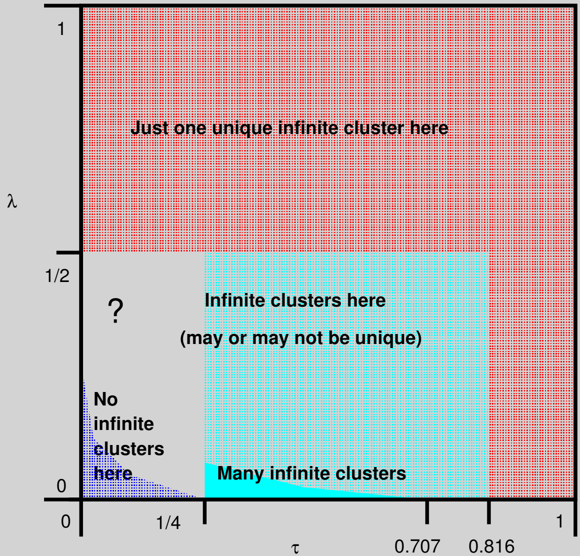


Contribution to upper bound on second phase transition:

Theorem 4 *If $\tau > \sqrt{2/3}$ then the infinite cluster of $\mathbb{Q}_{2:0}$ is almost surely unique for all positive λ.*

Method: prune bonds, branching processes, 2-dim comparison ...

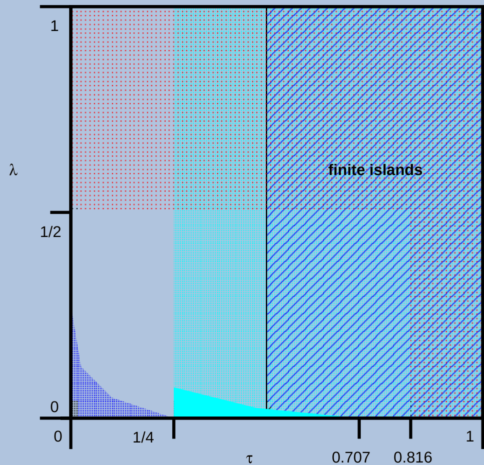




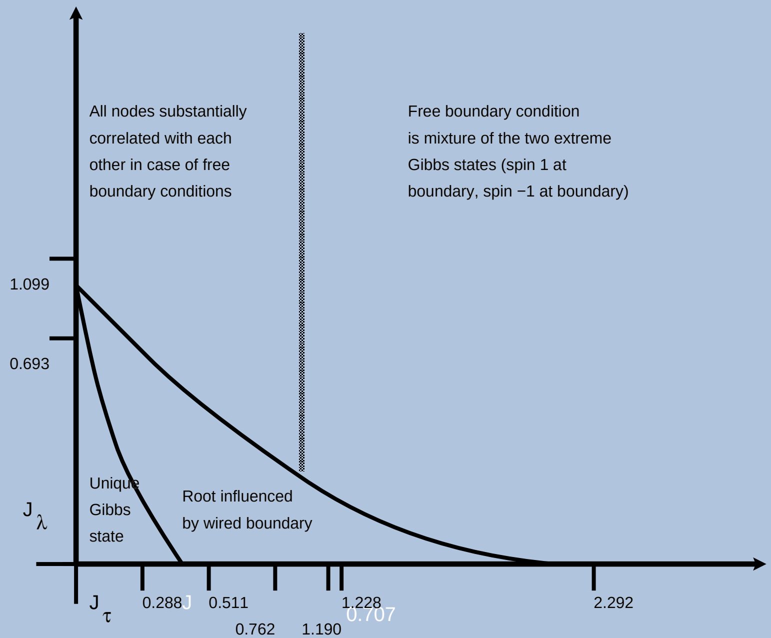
The story so far: includes uniqueness transition for case $d = 2$.

5. Comparison

We need to apply the Fortuin-Kasteleyn comparison inequalities (3) and (4). The event "just one unique infinite cluster" is **not** increasing, so we need more. Newman and Wu (1990) show it suffices to establish a *finite island property* for the site percolation derived under adjacency when all infinite clusters are removed. Thus:



Comparison arguments then show the following schematic phase diagram for the Ising model on $\mathbb{Q}_2(\mathbf{o})$:



6. Simulation

Approximate simulations confirm the general story:

http://www.dcs.warwick.ac.uk/~rgw/sira/sim.html

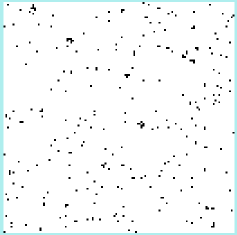
(1) Only 200 resolution levels;

(2) At each level, 1000 sweeps in scan order;

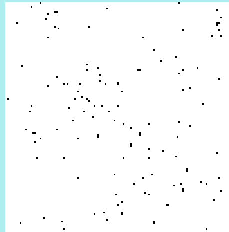
(3) At each level, simulate square sub-region of 128×128 pixels conditioned by mother 64×64 pixel region;

(4) Impose periodic boundary conditions on 128×128 square region;

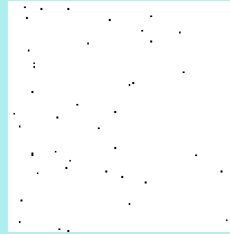
(5) At the coarsest resolution, all pixels set white. At subsequent resolutions, 'all black' initial state.



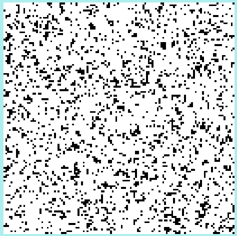
(a) $J_\lambda = 1, J_\tau = 0.5$



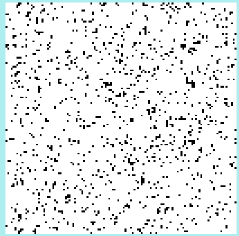
(b) $J_\lambda = 1, J_\tau = 1$



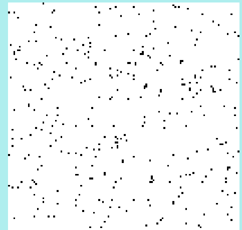
(c) $J_\lambda = 1, J_\tau = 2$



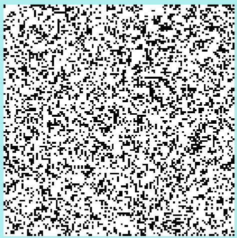
(d) $J_\lambda = 0.5, J_\tau = 0.5$



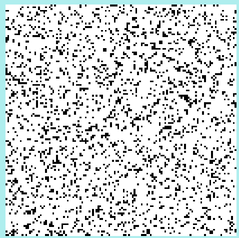
(e) $J_\lambda = 0.5, J_\tau = 1$



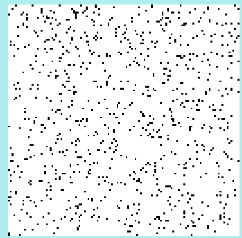
(f) $J_\lambda = 0.5, J_\tau = 2$



(g) $J_\lambda = 0.25, J_\tau = 0.5$



(h) $J_\lambda = 0.25, J_\tau = 1$



(i) $J_\lambda = 0.25, J_\tau = 2$

7. Future work

This is about the *free* Ising model on $\mathbb{Q}_2(\mathbf{o})$. Image analysis more naturally concerns the case of prescribed boundary conditions (say, image at finest resolution level . . .).

Question: will boundary conditions at "infinite fineness" propagate back to finite resolution?

Series and Sinaĭ (1990) show answer is yes for analogous problem on hyperbolic disk (2-dim, all bond probabilities the same).

Gielis and Grimmett (2001) point out (*eg*, in $\mathbb{Z}^3$ case) these boundary conditions translate to a *conditioning* for random cluster model, and investigate using large deviations.

Project: do same for $\mathbb{Q}_2(\mathbf{o})$ . . . and get quantitative bounds?

Home Page
Title Page
◀◀ ▶▶
◀ ▶

Go Back
Full Screen
Close
Quit

References[1]

Bhalerao, A., E. Thönnes, W. Kendall, and R. Wilson (2001). Inferring vascular structure from 2d and 3d imagery. In W. Niessen and M. Viergever (Eds.), *Medical Image Computing and Computer-Assisted Intervention, proceedings of MICCAI 2001*, Volume 2208 of *Springer Lecture Notes in Computer Science*, pp. 820–828. Springer-Verlag. Also: University of Warwick Department of Statistics Research Report 392.

Fortuin, C. (1972a). On the random-cluster model. II. The percolation model. *Physica 58*, 393–418.

Fortuin, C. (1972b). On the random-cluster model. III. The simple random-cluster model. *Physica 59*, 545–570.

[1] This is a rich hypertext bibliography. Journals are linked to their homepages. Stable URL links (as provided for example by JSTOR) have been added to entries where available. Access to such URLs may not be universal: in case of difficulty you should check whether you are registered (directly or indirectly) with the relevant provider.

Fortuin, C. and P. Kasteleyn (1972). On the random-cluster model. I. Introduction and relation to other models. *Physica 57*, 536–564.

Gielis, G. and G. Grimmett (2001, September). Rigidity of the interface for percolation and random-cluster models. Preprint, arXiv:math.PR/0109103. arXiv

Grimmett, G. and C. Newman (1990). Percolation in $\infty + 1$ dimensions. In *Disorder in physical systems*, pp. 167–190. New York: The Clarendon Press Oxford University Press.

Kendall, W. and R. Wilson (2002). Ising models and multiresolution quad-trees. Research report 402, Department of Statistics, University of Warwick.

Newman, C. and C. Wu (1990). Markov fields on branching planes. *Probability Theory and Related Fields 85*(4), 539–552.

Preston, C. (1977). Spatial birth-and-death processes. *Bull. Inst. Internat. Statist. 46*, 371–391.

Series, C. and Y. Sinaĭ (1990). Ising models on the Lobachevsky plane. *Communica-*

tions in Mathematical Physics 128(1), 63–76.

Spitzer, F. (1975). Markov random fields on an infinite tree. *The Annals of Probability 3*(3), 387–398.

Thönnes, E., A. Bhalerao, W. Kendall, and R. Wilson (2002). A Bayesian approach to inferring vascular tree structure from 2d imagery. In W. Niessen and M. Viergever (Eds.), *ICIP 2002*, Volume to appear. Also: University of Warwick Department of Statistics Research Report 391.

Wilson, R. and C. Li (2002). A class of discrete multiresolution random fields and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24(11)*, Accepted for publication.

A. Notes on proof of Theorem 1

Mean size of cluster at o bounded above by

$$\begin{aligned} & \sum_{n=0}^{\infty} \sum_{\underline{\mathbf{t}}:|\underline{\mathbf{t}}|=n} \mathcal{X}_\lambda \tau^n (\mathcal{X}_\lambda - 1)^{T(\underline{\mathbf{t}})} \mathcal{X}_\lambda^{n-T(\underline{\mathbf{t}})} \\ \leq\ & \sum_{n=0}^{\infty} \mathcal{X}_\lambda (\tau \mathcal{X}_\lambda)^n \sum_{\underline{\mathbf{t}}:|\underline{\mathbf{t}}|=n} (1 - \mathcal{X}_\lambda^{-1})^{T(\underline{\mathbf{t}})} \\ \leq\ & \sum_{n=0}^{\infty} \mathcal{X}_\lambda (2^d \tau \mathcal{X}_\lambda)^n \sum_{\underline{\mathbf{j}}:|\underline{\mathbf{j}}|=n} (1 - \mathcal{X}_\lambda^{-1})^{T(\underline{\mathbf{j}})} \\ \approx\ & \sum_{n=0}^{\infty} \mathcal{X}_\lambda (2^d \tau \mathcal{X}_\lambda)^n \left(1 + \sqrt{1 - \mathcal{X}_\lambda^{-1}}\right)^n . \end{aligned}$$

For last step, use spectral analysis of matrix representation

$$\sum_{\underline{\mathbf{j}}:|\underline{\mathbf{j}}|=n} (1 - \mathcal{X}_\lambda^{-1})^{T(\underline{\mathbf{j}})} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 - \mathcal{X}_\lambda^{-1} & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

B. Notes on proof of Theorem 2

Uniqueness: For negative exponent $\xi(1-\lambda)$ of dual connectivity function, set

$$\ell_n \quad = \quad (n \log 4 + (2+\epsilon) \log n)\, \xi(1-\lambda)\,.$$

More than one "ℓ_n-large" cluster in L_n forces existence of open path in dual lattice longer than ℓ_n. Now use Borel-Cantelli ....

On the other hand super-criticality will mean *some* distant points in L_n are inter-connected.

Existence: consider $4^{n-[n/2]}$ points in L_{n-1} and specified daughters in L_n. Study probability that

(a) parent percolates more than ℓ_{n-1},

(b) parent and child are connected,

(c) child percolates more than ℓ_n.

Now use Borel-Cantelli again ....

C. Notes on proof of Theorem 3

Two relevant lemmas:

Lemma 1 *Consider* $\boldsymbol{u} \in L_{s+1} \subset \mathbb{Q}_d$ *and* $\boldsymbol{v} = \mathcal{M}(\boldsymbol{u}) \in L_s \subset \mathbb{Q}_d$. *There are exactly* 2^d *solutions in* L_{s+1} *of*

$$\mathcal{M}(\boldsymbol{x}) \quad = \quad \mathcal{S}_{\boldsymbol{u};\boldsymbol{v}}(\boldsymbol{x})\,.$$

One is $\boldsymbol{x} = \boldsymbol{u}$. *The others are the remaining* $2^d - 1$ *vertices* $\boldsymbol{y}$ *such that the closure of the cell representing* $\boldsymbol{y}$ *intersects the vertex shared by the closures of the cells representing* $\boldsymbol{u}$ *and* $\mathcal{M}(\boldsymbol{u})$. *Finally, if* $\boldsymbol{x} \in L_{s+1}$ *does not solve* $\mathcal{M}(\boldsymbol{x}) = \mathcal{S}_{\boldsymbol{u};\boldsymbol{v}}(\boldsymbol{x})$ *then*

$$\|\mathcal{S}_{\boldsymbol{u};\boldsymbol{v}}(\boldsymbol{x}) - \mathcal{S}_{\boldsymbol{u};\boldsymbol{v}}(\boldsymbol{u})\|_{s,\infty} \quad > \quad \|\mathcal{M}(\boldsymbol{x}) - \mathcal{M}(\boldsymbol{u})\|_{s,\infty}\,. \qquad (5)$$

Lemma 2 *Given* distinct $\boldsymbol{v}$ *and* $\boldsymbol{y}$ *in the same resolution level. Count pairs of vertices* $\boldsymbol{u}$, $\boldsymbol{x}$ *in the resolution level one step higher, such that*

(a) $\mathcal{M}(\boldsymbol{u}) = \boldsymbol{v}$; (b) $\mathcal{M}(\boldsymbol{x}) = \boldsymbol{y}$; (c) $\mathcal{S}_{\boldsymbol{u};\boldsymbol{v}}(\boldsymbol{x}) = \boldsymbol{y}$.

There are at most 2^{d-1} *such vertices.*

D. Notes on proof of Theorem 4

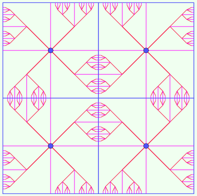
Prune! Then a direct connection is certainly established across the boundary between the cells corresponding to two neighbouring vertices **u**, **v** in L_0 if

(a) the τ-bond leading from **u** to the relevant boundary is open;

(b) a τ-branching process (formed by using τ-bonds mirrored across the boundary) survives indefinitely, where this branching process has family-size distribution Binomial$(2, \tau^2)$;

(c) the τ-bond leading from **v** to the relevant boundary is open.

Then there are infinitely many chances of making a connection across the cell boundary.



E. Notes on proof of infinite island property

Notion of "cone boundary" $\partial_c(S)$ of finite subset S of vertices: collection of daughters $\mathbf{v}$ of S such that $\mathbb{Q}_d(\mathbf{v}) \cap S = \emptyset$.

Use induction on S, building it layer L_n on layer L_{n-1} to obtain an isoperimetric bound: $\#(\partial_c(S)) \geq (2^d - 1)\#(S)$. Hence deduce

$$\mathbb{P}\left[S \text{ in island at } \mathbf{u}\right] \quad \leq \quad (1 - p_\tau(1-\eta))^{(2^d-1)n}$$

where $\#(S) = n$ and $\eta = \mathbb{P}\left[\mathbf{u} \text{ not in infinite cluster of } \mathbb{Q}_d(\mathbf{u})\right]$.

Upper bound on number $N(n)$ of self-avoiding paths S of length n beginning at $\mathbf{u}_0$:

$$N(n) \quad \leq \quad (1 + 2d + 2^d)(2d + 2^d)^n\,.$$

Hence upper bound on the mean size of the island:

$$\sum_{n=0}^{\infty} (1 + 2d + 2^d)(2d + 2^d)^n \eta_{\text{br}}^{n(1-2^{-d})}\,,$$

where η_{br} is extinction probability for branching process based on Binomial$(2^d, p_\tau)$ family distribution.