# The Square Root Law of Steganography: Bringing Theory Closer to Practice

Andrew Ker
adk@cs.ox.ac.uk
University of Oxford
Department of Computer Science

5th ACM Workshop on Information Hiding & Multimedia Security
Philadelphia, 20 June 2017

# The moral of the story

***The Steganographer versus the Detector***

Suppose a payload of $m(n)$ bits hidden in cover of $n$ ‘pixels’.

If $m(n)$ is small enough, and the cover has no deterministic parts,
any Detector is unreliable.

If $m(n)$ is too large, and the embedding is imperfect,
some Detector will reliably detect it.

The ***critical rate*** $r(n)$ between these two cases is $O(\sqrt{n})$.

- In practice: observed robustly.
- In theory: ...

# The moral of the story

Given
- a probabilistic cover model,
- a probabilistic embedding model,

Suppose a payload of $m(n)$ bits hidden in cover of $n$ 'pixels'.

If $m(n)/r(n) \to 0$, and 'no determinism' in the cover model,
any Detector has $P_{fp} + P_{fn} \to 1$.

If $m(n)/r(n) \to \infty$, and 'no free bits' in the embedding process,
some Detector has $P_{fp} + P_{fn} \to 0$.

The ***critical rate*** $r(n)$ between these two cases is $O(\sqrt{n})$.

- In practice: observed robustly.
- In theory: ...

# The moral of the story

**Cover model**

- i.i.d. discrete pixels
  *[Ker, 2009] [Ker, 2010]*
- independent pixels
  *[Ker 2011]*
- i.i.d. continuous signals
  *[Bash, Goeckel, Towsley, 2012]*
- $1^{st}$-order Markov chain
  *[Filler, Ker, Fridrich, 2009]*

- In practice: observed robustly.
- In theory: only proved for very simple models.

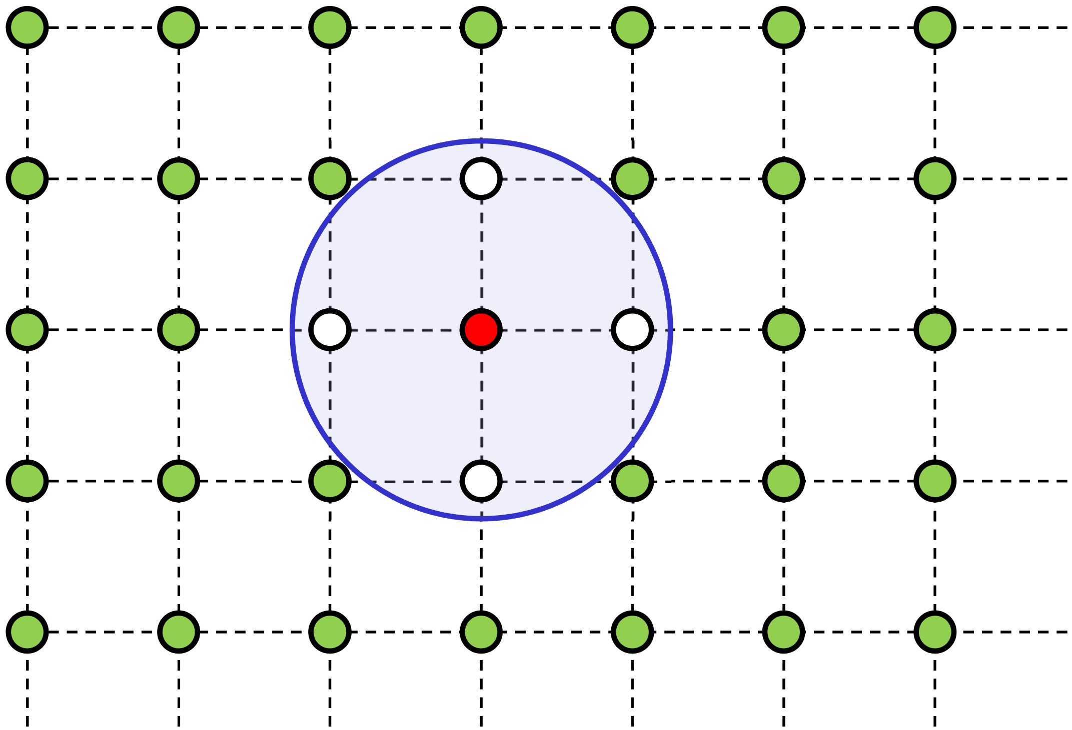
# Outline

- New result for wide class of covers models,
  - sketch proof, many details omitted.
- Examples.
- Conclusions.

# New square root law

| Cover model | Embedding model | Conditions | Critical Rate |
|---|---|---|---|
| **Markov random field** *(i)* of bounded degree, *(ii)* with exponential decay of local covariance<br><br>*- density of local neighbourhoods* $p_i(\mathbf{x})$ | use exactly *m* pixels<br><br>*- density of local neighbourhoods* $q_i(\mathbf{x})$ | 'No cover determinism': $p_i(\mathbf{x}) > \epsilon$<br><br>'No free bits': $\|p_i - q_i\|_1 > \epsilon \frac{m}{n}$ | $\sqrt{n}$ |

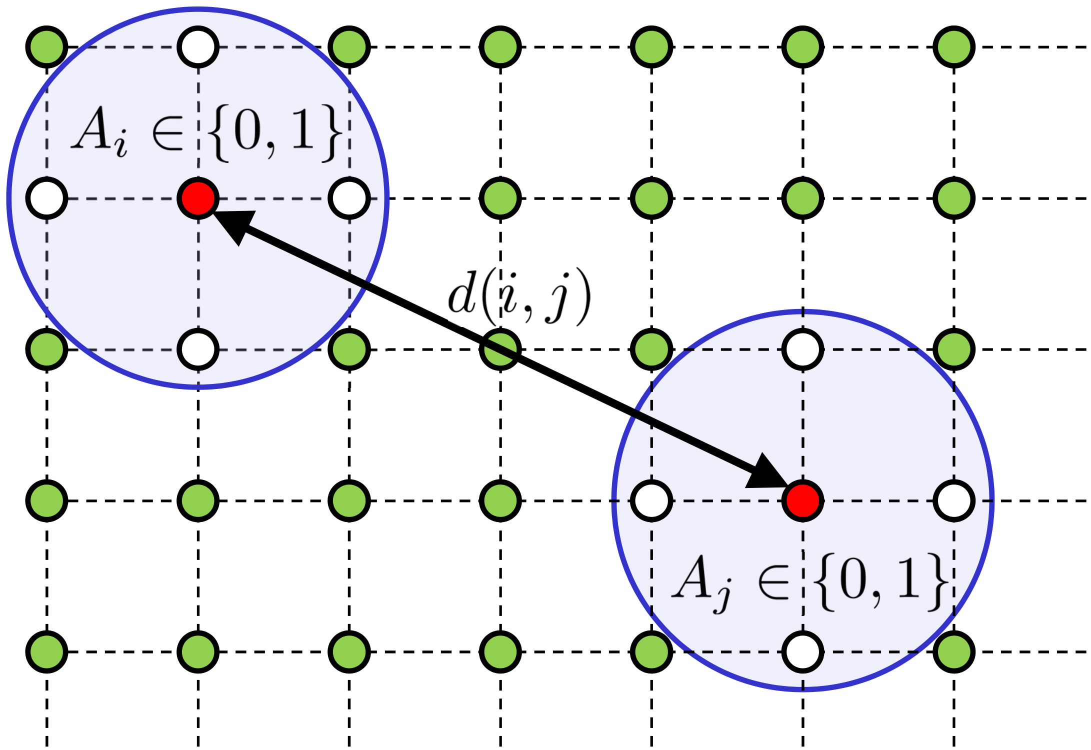
# Markov Random Field



Conditional on ○s, ● is independent of all ● s.

Bounded degree: maximum neighbourhood cardinality.

# Markov Random Field




Exponential decay of covariance: $\mathrm{Cov}[A_i, A_j] \leq Ce^{-cd(i,j)}$

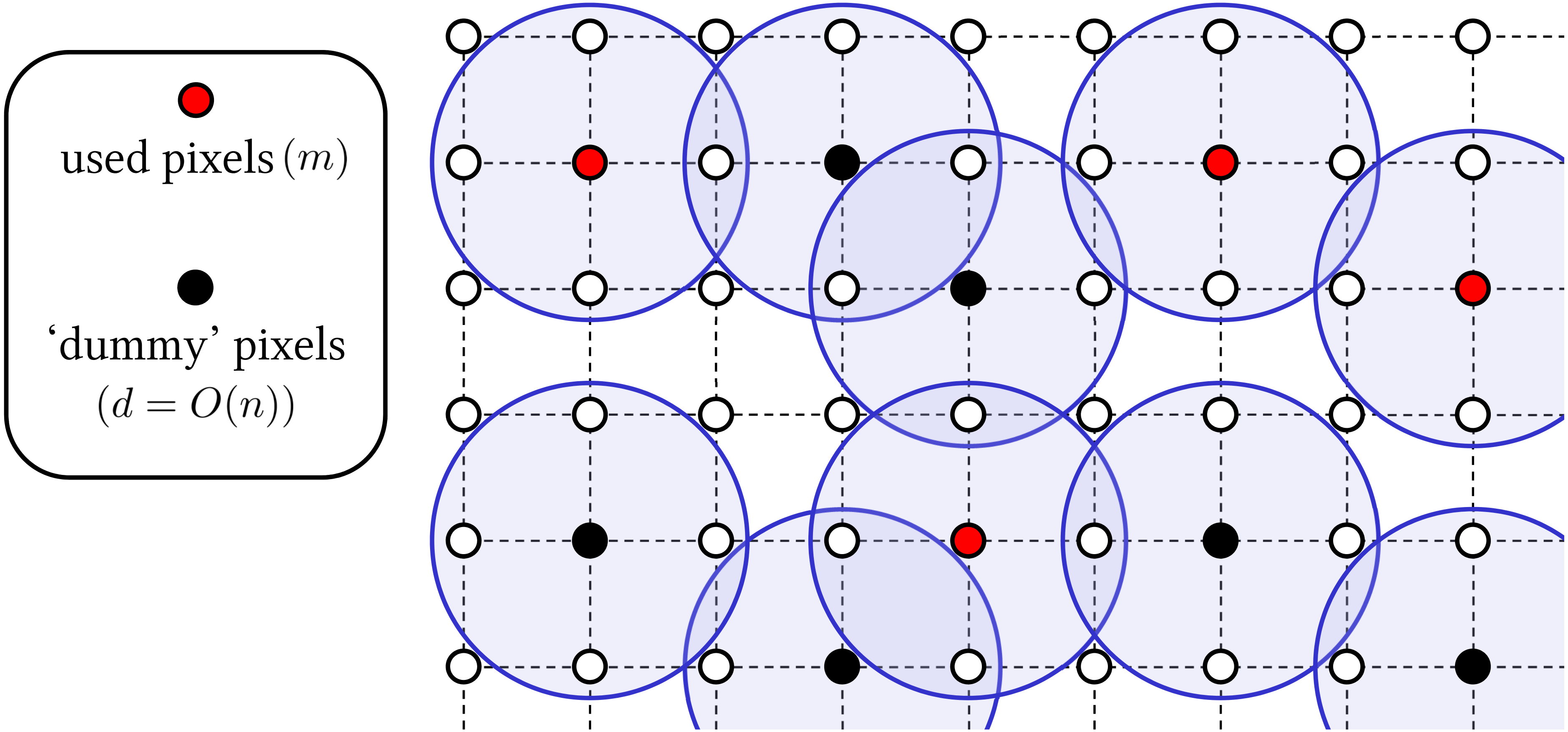
# New square root law

| Cover model | Embedding model | Conditions | Critical Rate |
|---|---|---|---|
| **Markov random field**<br>*(i)* of bounded degree,<br>*(ii)* with exponential decay of local covariance<br><br>- *density of local neighbourhoods* $p_i(\mathbf{x})$ | use exactly *m* pixels<br><br>- *density of local neighbourhoods* $q_i(\mathbf{x})$ | 'No cover determinism':<br>$p_i(\mathbf{x}) > \epsilon$<br><br>'No free bits':<br>$\lVert p_i - q_i \rVert_1 > \epsilon \frac{m}{n}$ | $\sqrt{n}$ |

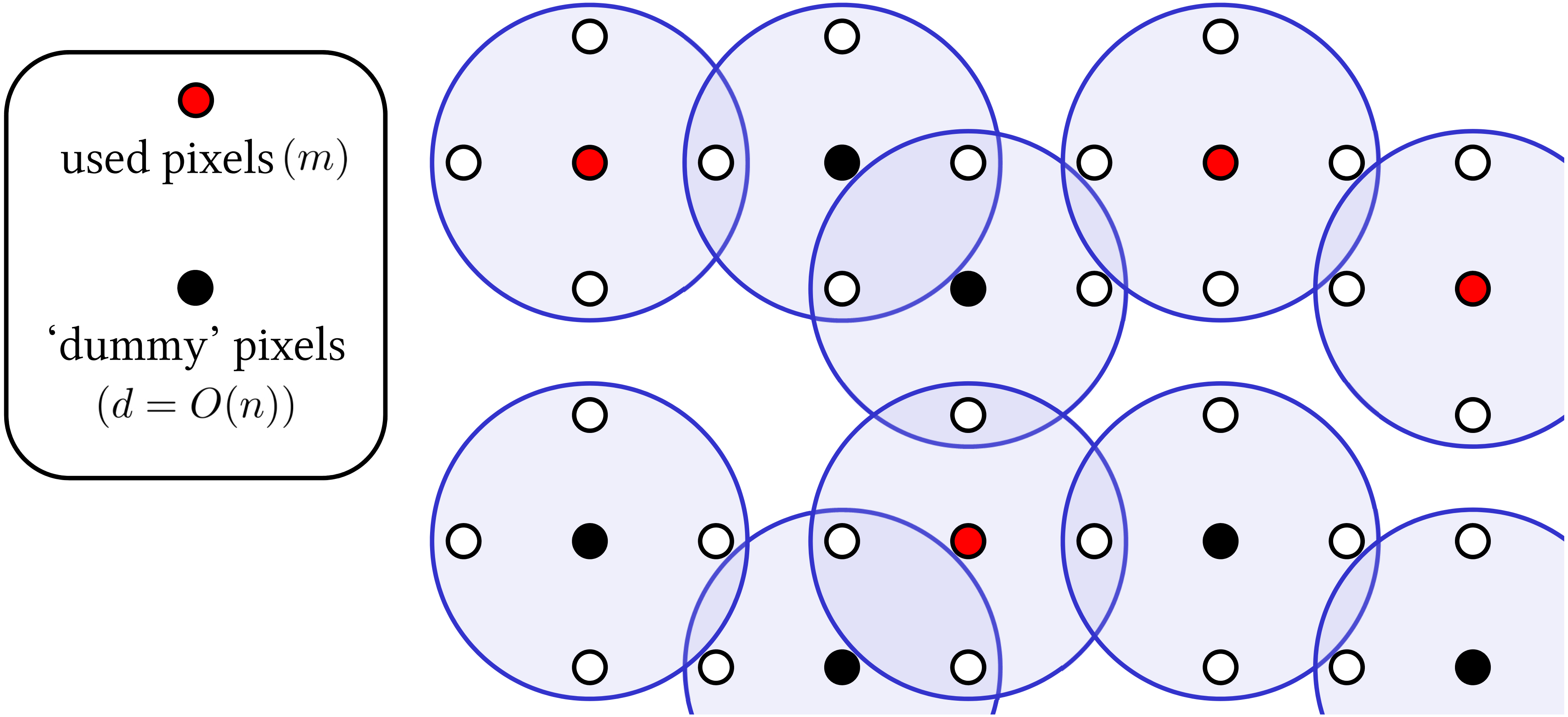
*'lower bound'* if $m/\sqrt{n} \to 0,$ then
all detectors are asymptotically random,

*'upper bound'* if $m/\sqrt{n} \to \infty,$ then
an asymptotically perfect detector exists.

# Sketch proof: lower bound



# Sketch proof: lower bound



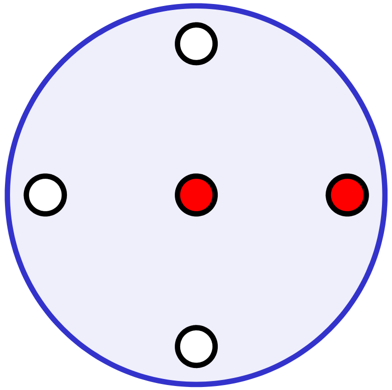

Conditional on ○s, ●s and ●s are mutually independent.

Detector is asymptotically random if $m/\sqrt{d+m} \to 0$.

# Sketch proof: lower bound

- Formalize side information argument (uses Total Variation).

- No item in the shortlist neighbours another. Probability of



tends to zero if $m/\sqrt{n} \to 0$.

- Apply correct Square Root Law for conditionally independent pixels.

# Sketch proof: upper bound

$$\mathrm{Var}_{\mathrm{cov}}[\textstyle\sum_i A_i] = \sum_{i,j} \mathrm{Cov}_{\mathrm{cov}}[A_i, A_j] \le \sum_{i,j} Ce^{-cd(i,j)} = O(n)$$

For each neighbourhood $i$ there is an indicator $A_i$ with

$$\mathrm{E}_{\mathrm{steg}}[A_i] - \mathrm{E}_{\mathrm{cov}}[A_i] > \epsilon\frac{m}{n}$$

$$\mathrm{E}_{\mathrm{steg}}[\textstyle\sum_i A_i] - \mathrm{E}_{\mathrm{cov}}[\sum_i A_i] > \epsilon m$$

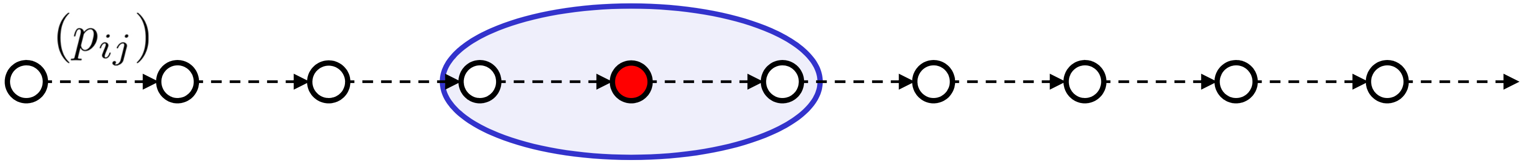
| | | | |
|---|---|---|---|
| *(i)* of bounded degree,<br>*(ii)* with exponential decay of local covariance<br><br>*- density of local neighbourhoods* $p_i(\mathbf{x})$ | *- density of local neighbourhoods* $q_i(\mathbf{x})$ | 'No free bits':<br>$\lVert p_i - q_i \rVert_1 > \epsilon\frac{m}{n}$ | |

$$\mathit{deflection}(\textstyle\sum_i A_i) \ge \dfrac{\Omega(m)}{O(\sqrt{n})}$$

so if $m/\sqrt{n} \to \infty$, $\sum_i A_i$ is an asymptotically perfect detector.

# Examples

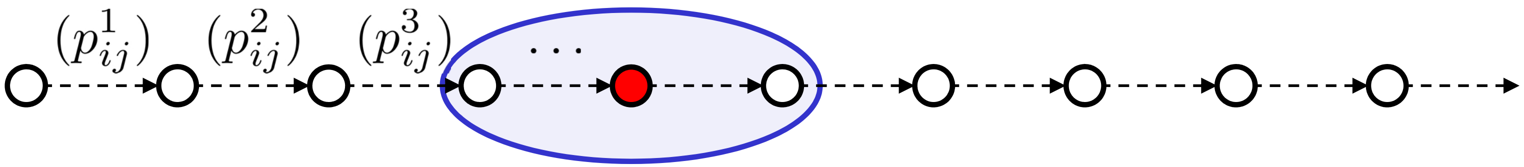
## Markov chain




| | | |
|---|---|---|
| Bounded maximum degree | ✓ | 2 |
| Exponential decay of covariance | ✓ | *follows from geometric ergodicity* |
| No determinism | ✓ | *if all* $p_{ij} > 0$ |
| No free bits | ✓ | *as long as embedding not perfect* |

# Examples

## Inhomogeneous Markov chain




| | | |
|---|---|---|
| Bounded maximum degree | ✓ | 2 |
| Exponential decay of covariance | ✓ | *follows from exponential forgetting* |
| No determinism | ✓ | *if all* $p_{ij}^k > \epsilon$ |
| No free bits | ✓ | *as long as embedding not perfect* |

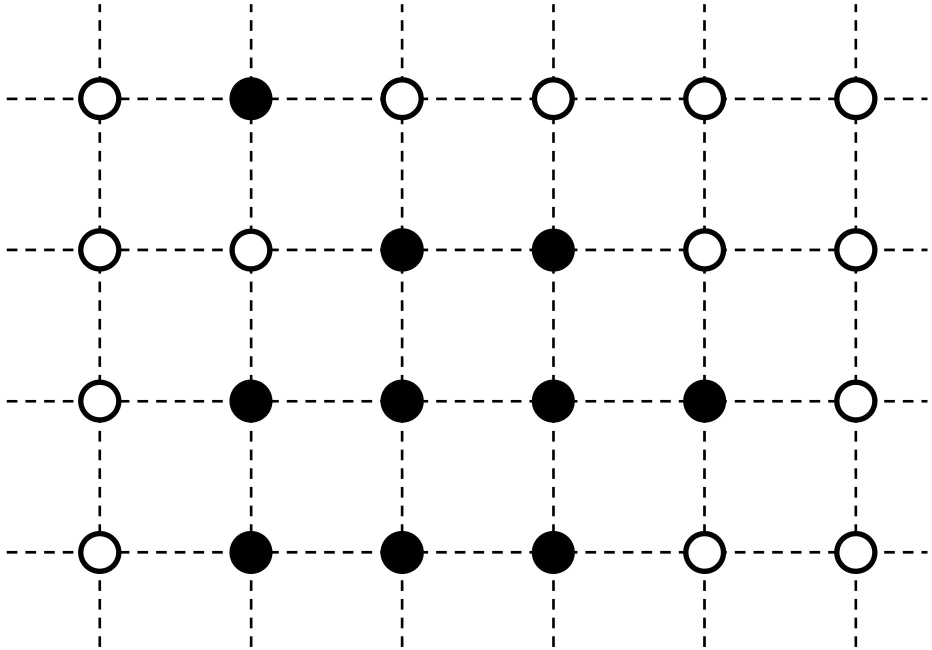
# Examples

## Ising model

$x_{ij} \in \{-1, +1\}$; density $\propto$

$$\exp\Big(\beta H \sum_{i,j} x_{ij} + \beta J \sum_{\substack{|i-i'|+\\|j-j'|=1}} \sum x_{ij} x_{i'j'}\Big)$$

higher $\beta H \to +1$ more likely
higher $\beta J \to$ neighbours more equal



# Examples

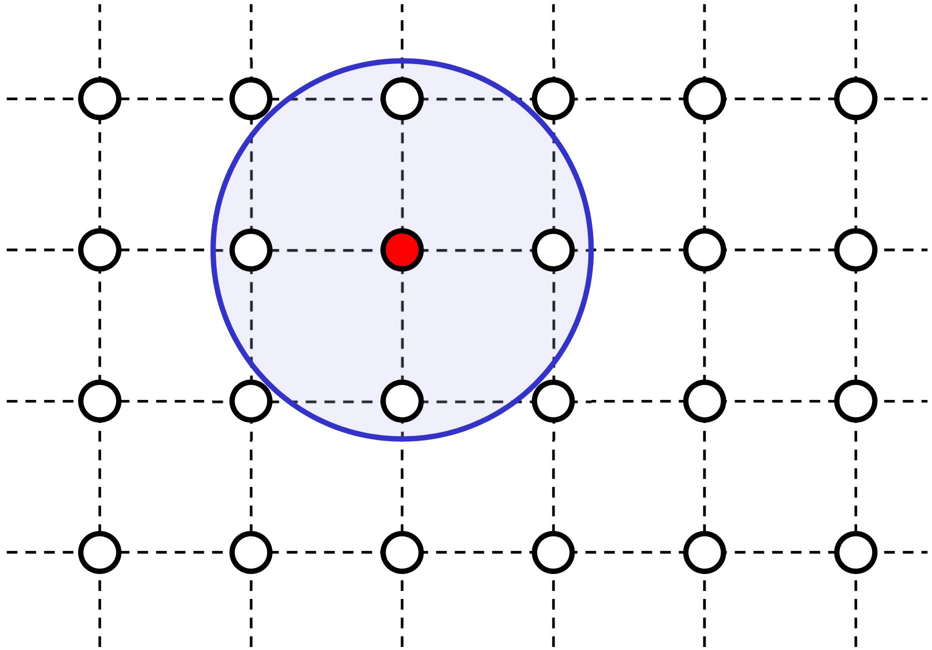
## Ising model

$x_{ij} \in \{-1,+1\}$; density $\propto$

$$\exp\Big(\beta H \sum_{i,j} x_{ij} + \beta J \sum_{\substack{|i-i'|+\\|j-j'|=1}} \sum x_{ij} x_{i'j'}\Big)$$

higher $\beta H \to +1$ more likely

higher $\beta J \to$ neighbours more equal



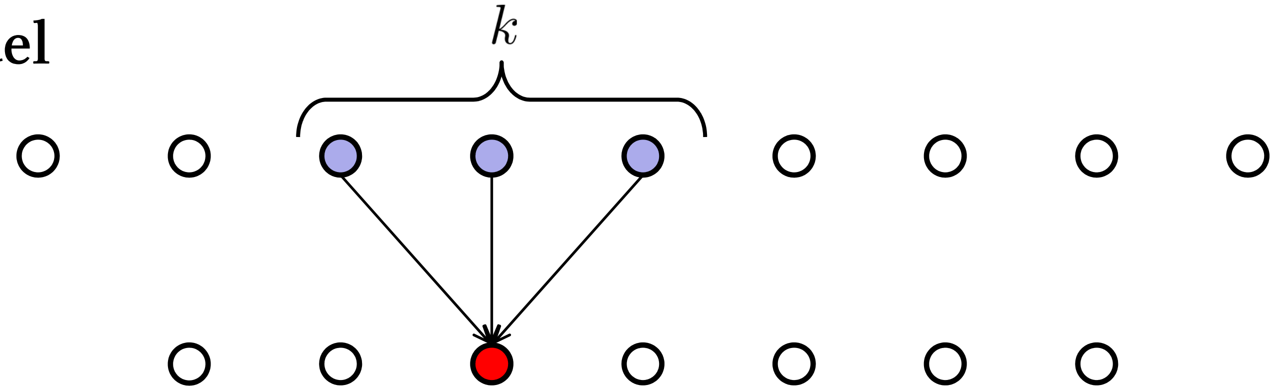
| | |
|---|---|
| Bounded maximum degree | ✓ 4 |
| Exponential decay of covariance | ✓ *from 'Dobrushin's condition' if* $\beta H = 0$ *and* $|\beta J|$ *small, or* $|\beta H|$ *large* |
| No determinism / no free bits | ✓ *unless* $\beta H = \infty$, $\beta J = \infty$, *or* $\beta H = \beta J = 0$ |

# Examples

**Hidden-layer model**

latent variables
(random)

observed pixels



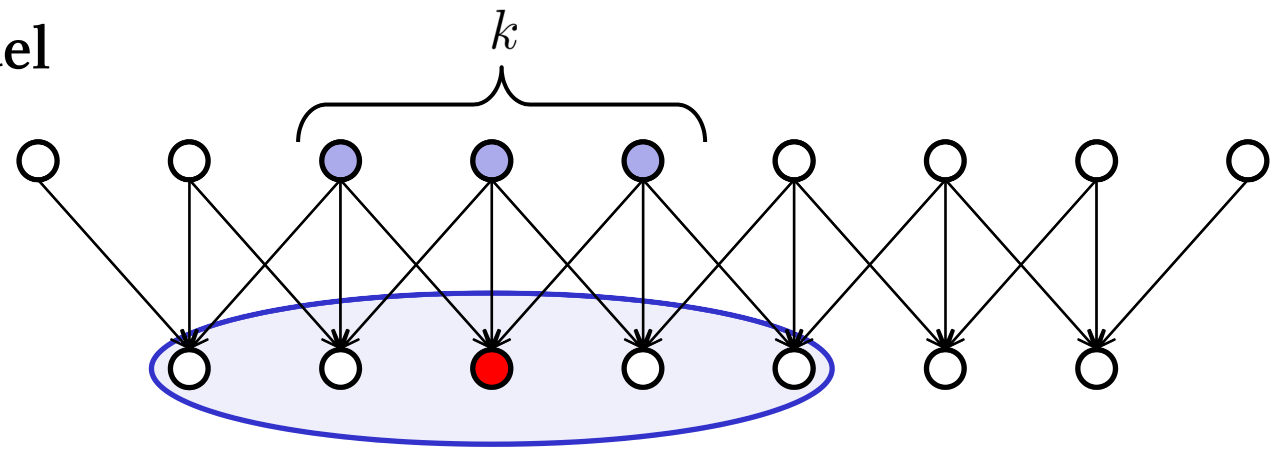

i.e. a deterministic function of random 'pre-cover'

# Examples

**Hidden-layer model**

latent variables
(random)

observed pixels




i.e. a deterministic function of random 'pre-cover'

| | | |
|---|---|---|
| Bounded maximum degree | ✓ | |
| Exponential decay of covariance | ✓ | $\mathrm{Cov}[x_i, x_j] = 0$ *for* $\lvert i - j \rvert > k$ |
| No determinism / no free bits | ✓ | *unless degenerate* |

# Conclusions

- There has been a gap between

  theory: *SRL for i.i.d. or 1$^{st}$-order Markov covers,*

  and practice: *SRL observed in real digital media objects.*

- New square root law for class of Markov Random Fields, including
  - (inhomogeneous / *k*-order) Markov chains,
  - some Ising models,
  - hidden-layer models.

  *'Local randomness, no long-range dependency'*

- Plenty of further directions...
  - adaptive embedding / source coding,
  - detector with imperfect information,
  - 'root rate'.

# Conclusions

- There has been a gap between

  theory: *SRL for i.i.d. or 1$^{st}$-order Markov covers,*

  and practice: *SRL observed in real digital media objects.*

- New square root law for class of Markov Random Fields, including
  - (inhomogeneous / $k$-order) Markov chains,
  - some Ising models,
  - hidden-layer models.

  *'Local randomness, no long-range dependency'*

- Plenty of further directions...

- There are also some interesting **non-examples** of the SRL.
  *to be continued...*