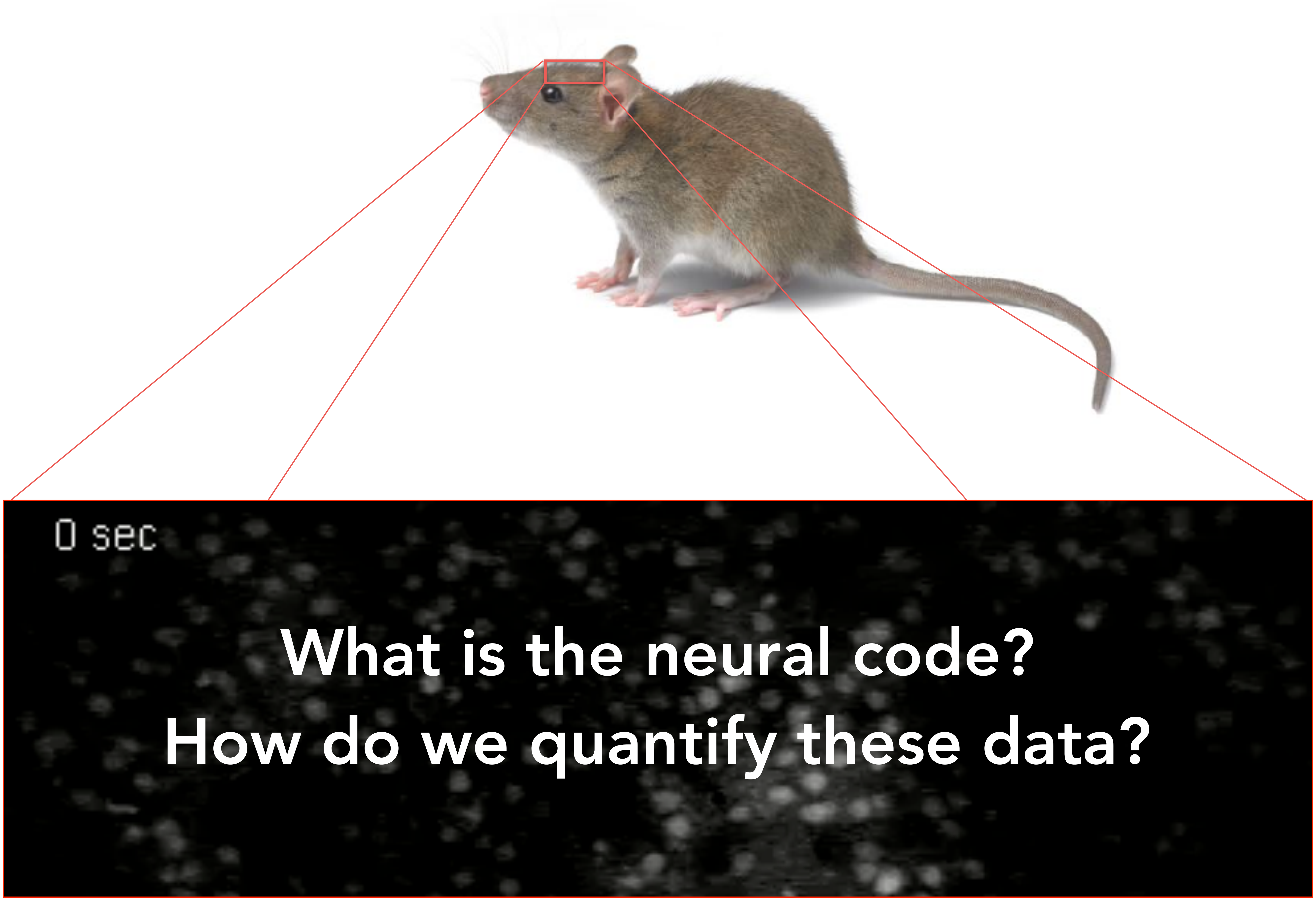
Information Processing and the Brain

Neural population data analysis

Dr. Cian O’Donnell
cian.odonnell@bristol.ac.uk



Golshani et al., J Neurosci (2009)

What we will cover

- Why analyse neural populations instead of single neurons?
- What properties should a good statistical model have?
- The independent-neuron model.
- Methods that capture spatial correlations:
    - Pairwise-maximum entropy
    - Population-count-based models
    - The dichotomised gaussian

Why neural populations?

The recent shift to neural populations

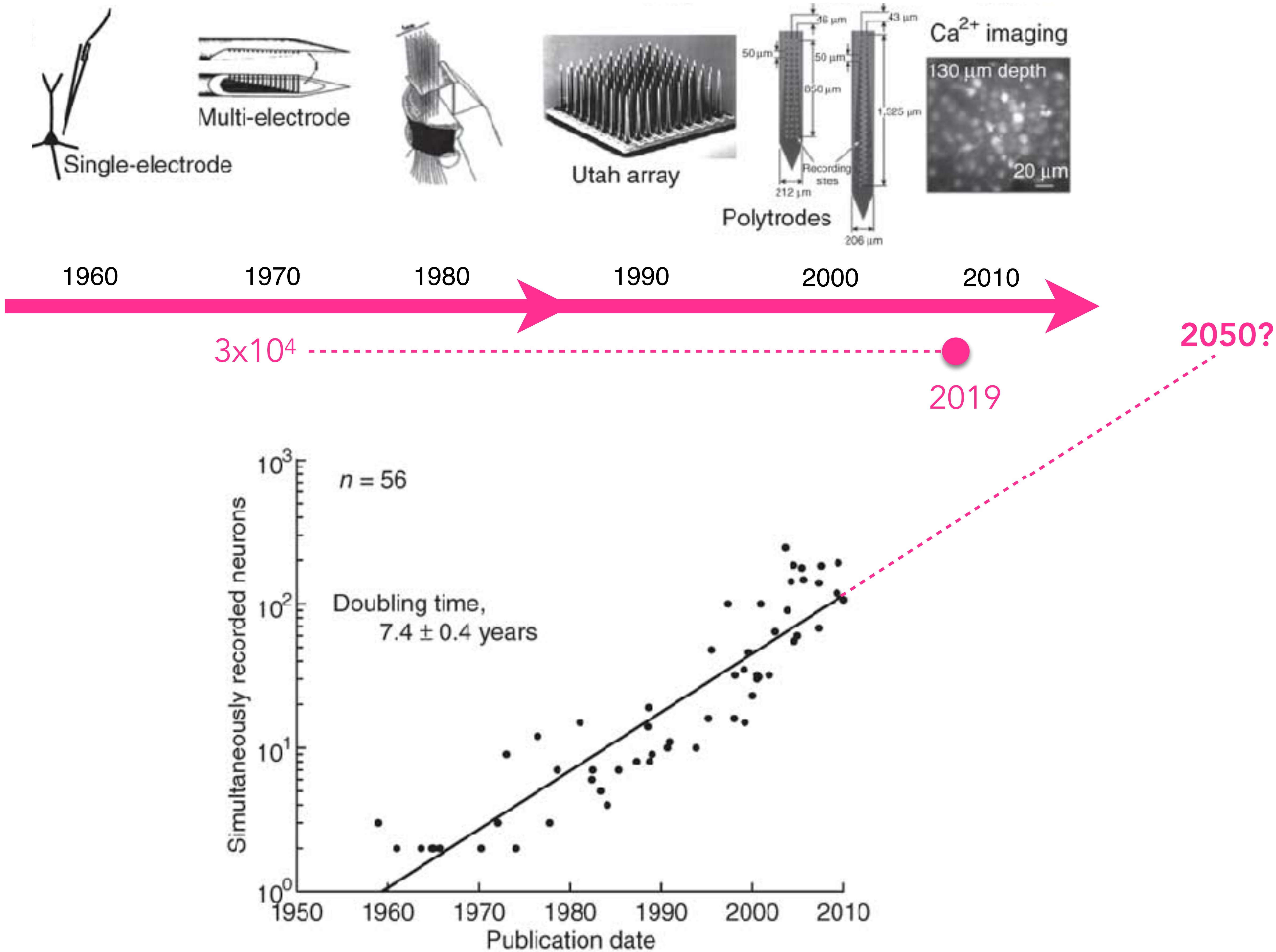
- Historically, neuroscientists focused on understanding the firing properties of single neurons (e.g. tuning curves, receptive fields, place cells).

- But the brain has a lot of neurons! (~100 billion in humans)
  What are they all for?

- This was partly due to limitations in recording techniques, but also because we didn't know what questions to ask.

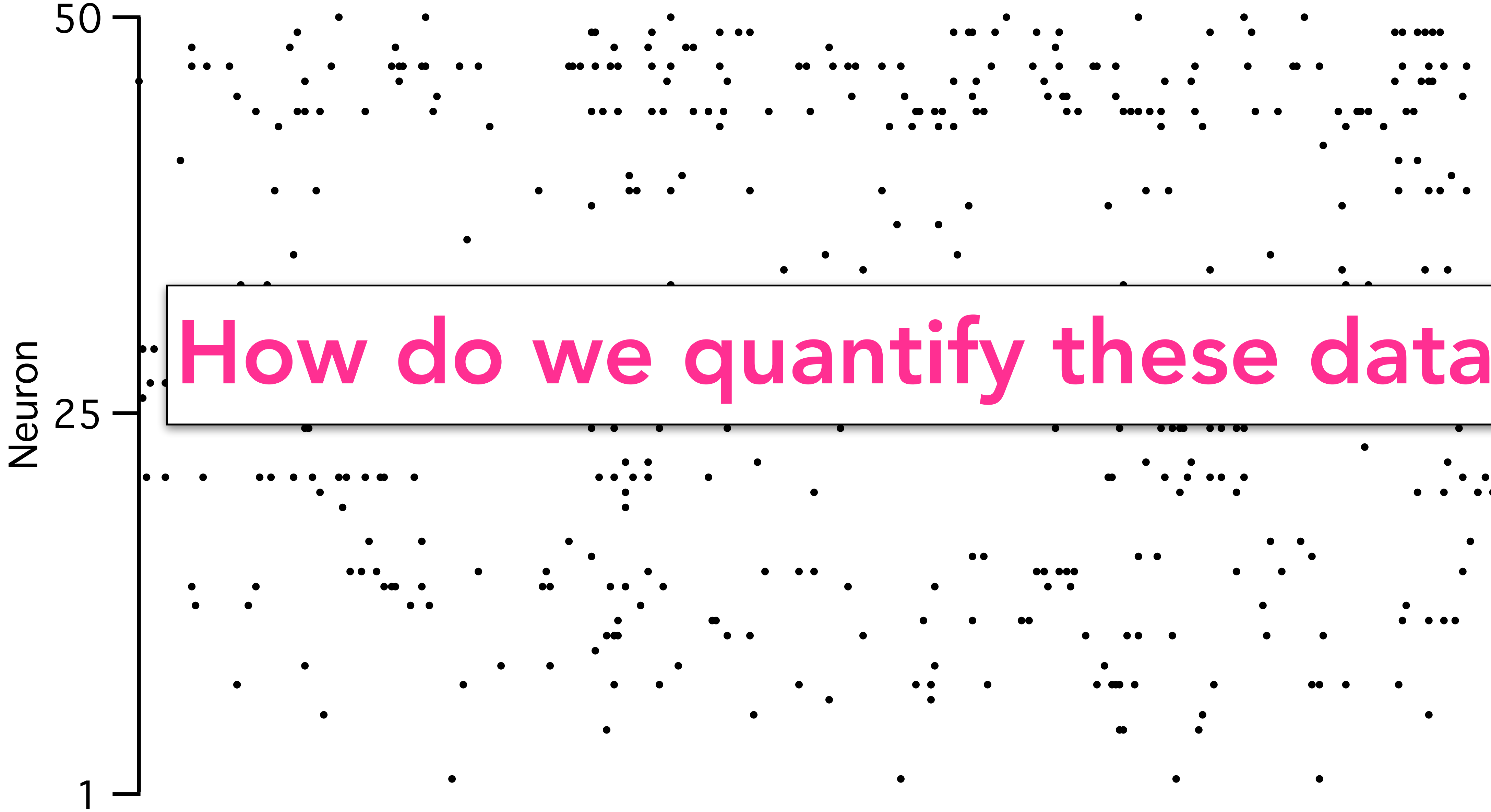
- But (as theorists have long pushed for) the trend is to now routinely record from 10s or 100s of neurons simultaneously.

- What we don't know is how to analyse or think about these data yet.



Stevenson & Kording, *Nat Neurosci* (2011)

Monkey primary visual cortex



How do we quantify these data?

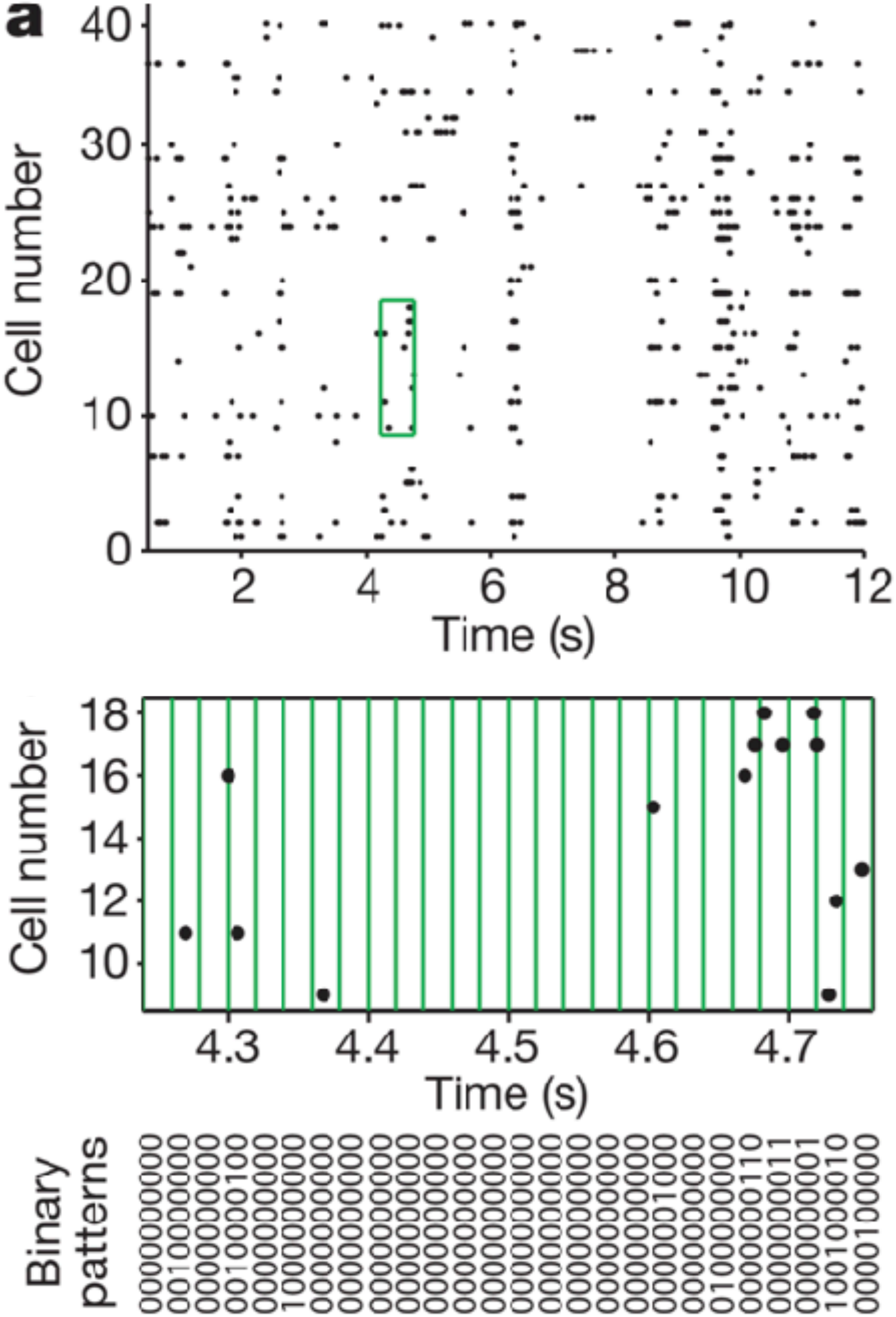
data from A. Kohn (Einstein College of Med)

What properties should a good statistical model for neural population data have?

Desirable properties of a statistical model for neural population data

- It captures the statistical structure in neural population data.
- Parameters can be fit efficiently for limited data.
- Parameter fitting can be done in reasonable computational time.
- We can generate samples from fitted model.
- Closed-form likelihood function.
- Model parameters are interpretable for humans.
- Model's latents variables are interpretable for humans.
- Can directly compute summary statistics from fitted parameters (entropy, distances between two parameter fits, etc)
- ~~Someone has coded it up in MATLAB or Python~~

How should we mathematically represent the neural activity?

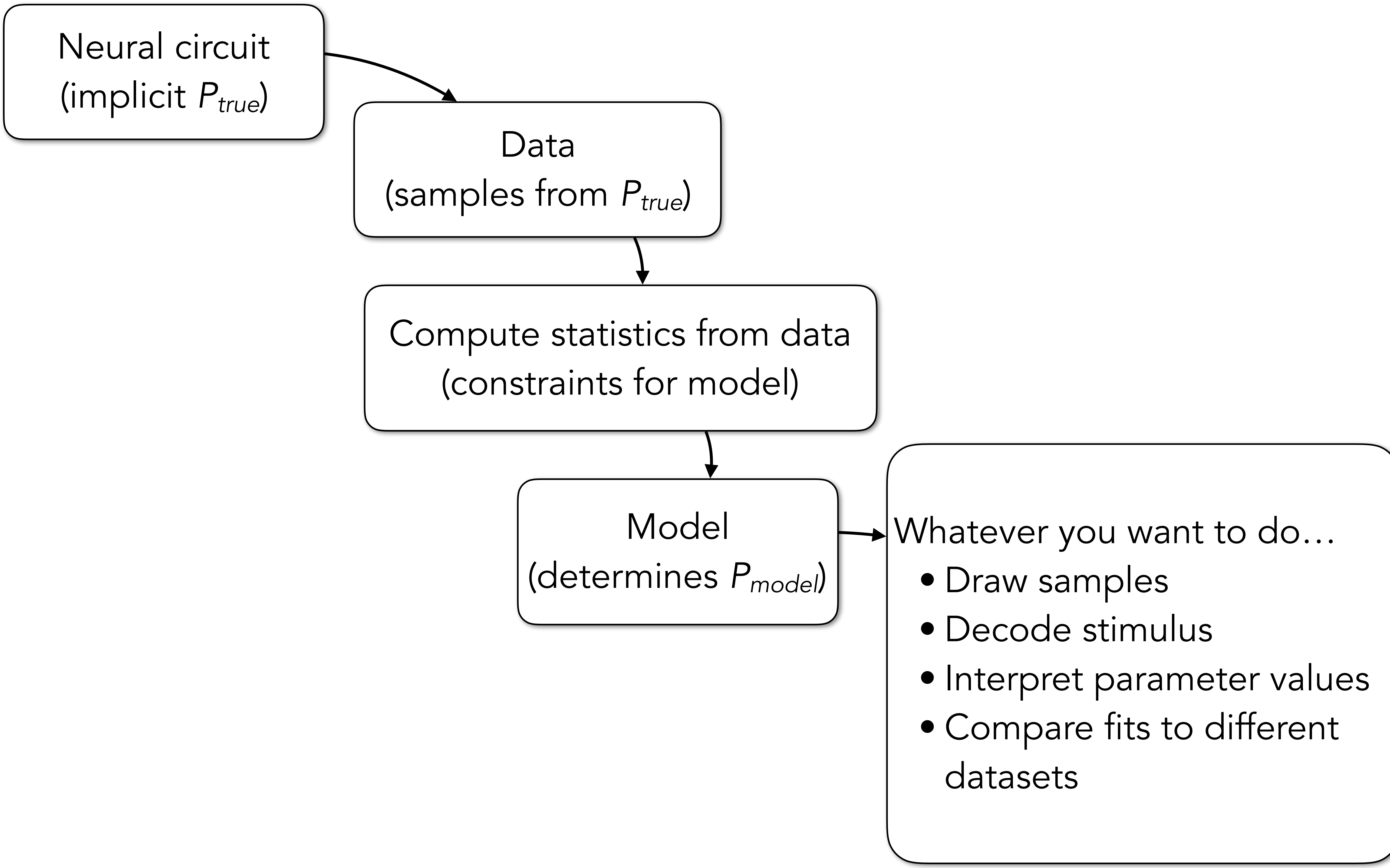


Schneidman et al, *Nature* (2006)

$$P(x_1, x_2, \ldots, x_N) = P(\mathbf{x}) = ?$$

The support of this distribution is the space of all possible binary patterns ~2^N

Modelling strategy



Given some data, how do we estimate $P(x)$?

The problem with histogramming

- If each pattern has a true probability of occurring p_i, then in the long run ($t \to \infty$) the maximum likelihood estimate $\hat{p}_i = n_1/(n_0 + n_1)$ converges towards p_i.

- However, since the number of possible patterns grows ~2^N, for reasonably-sized populations there may be many patterns with low p_i, which we observe only rarely. Maybe never.

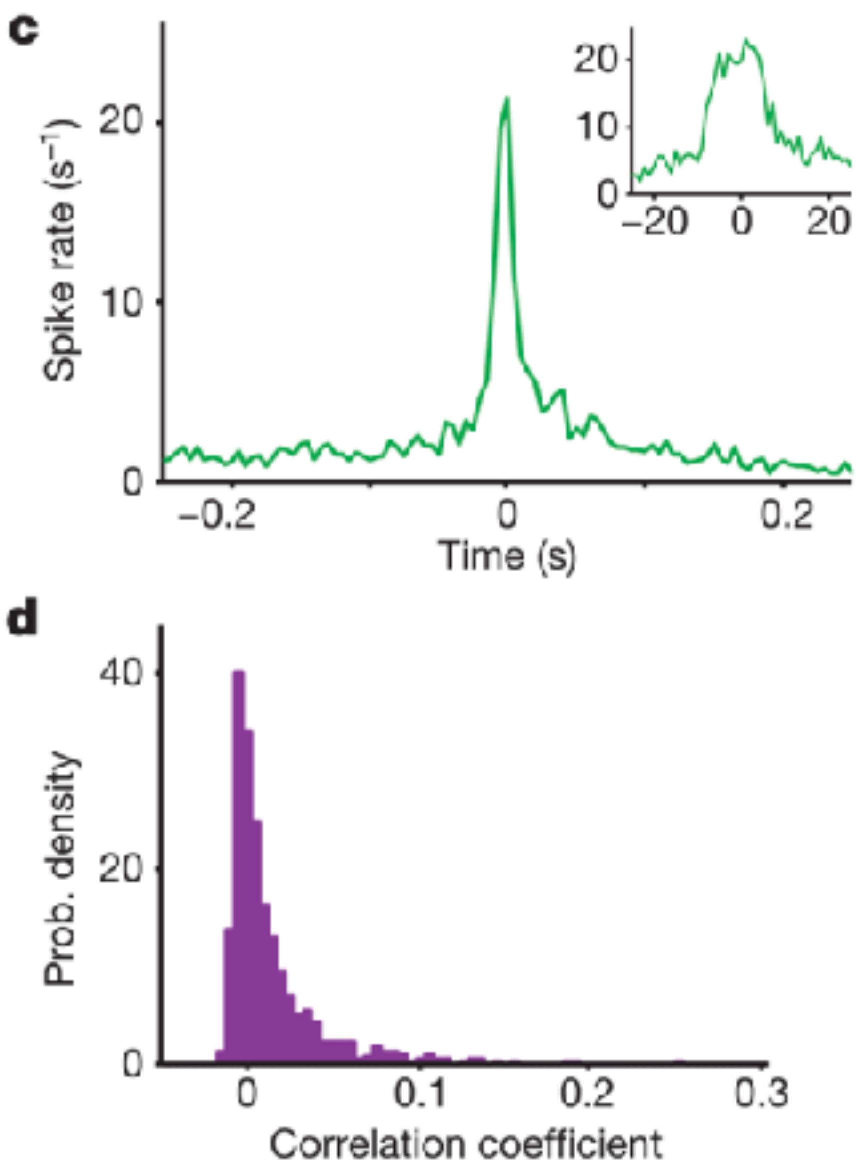
- For example, if we recorded some neural activity continually since the Big Bang (13.8B years ago = 4×10^{17} s), and binned our recording in 10 ms intervals, we could maximally observe the number of patterns corresponding to only $\log_2(4 \times 10^{19})$ ~ 65 neurons.

- These unobserved patterns are the brain's "dark matter"! They can exist (surely with non-zero probability) but need not ever occur.

The solution? Parametric models.

Independent neuron model

- The independent neuron model makes the assumption that neurons are statistically independent.
- This is wrong… but not *that* wrong.



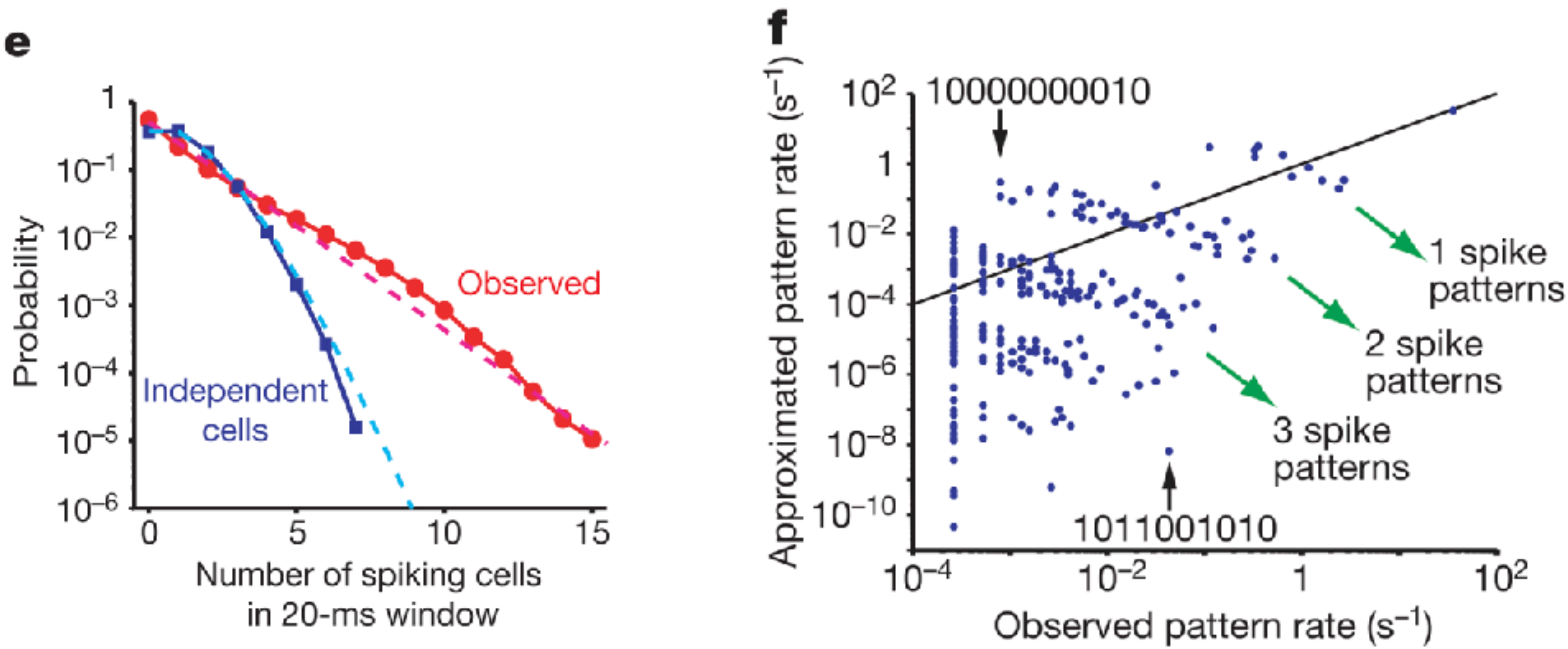
ON neurons

OFF neurons

$$p(\mathbf{x}) = \prod_i [x_i\sigma_i + (1 - x_i)(1 - \sigma_i)]$$

where σ_i is single neuron firing probability

Testing the independent model



Schneidman et al, *Nature* (2006)

We need a model that knows about correlations!

Three models with correlations

- Pairwise maximum entropy
- Population count-based models
- Dichotomised gaussian

Pairwise maximum entropy models

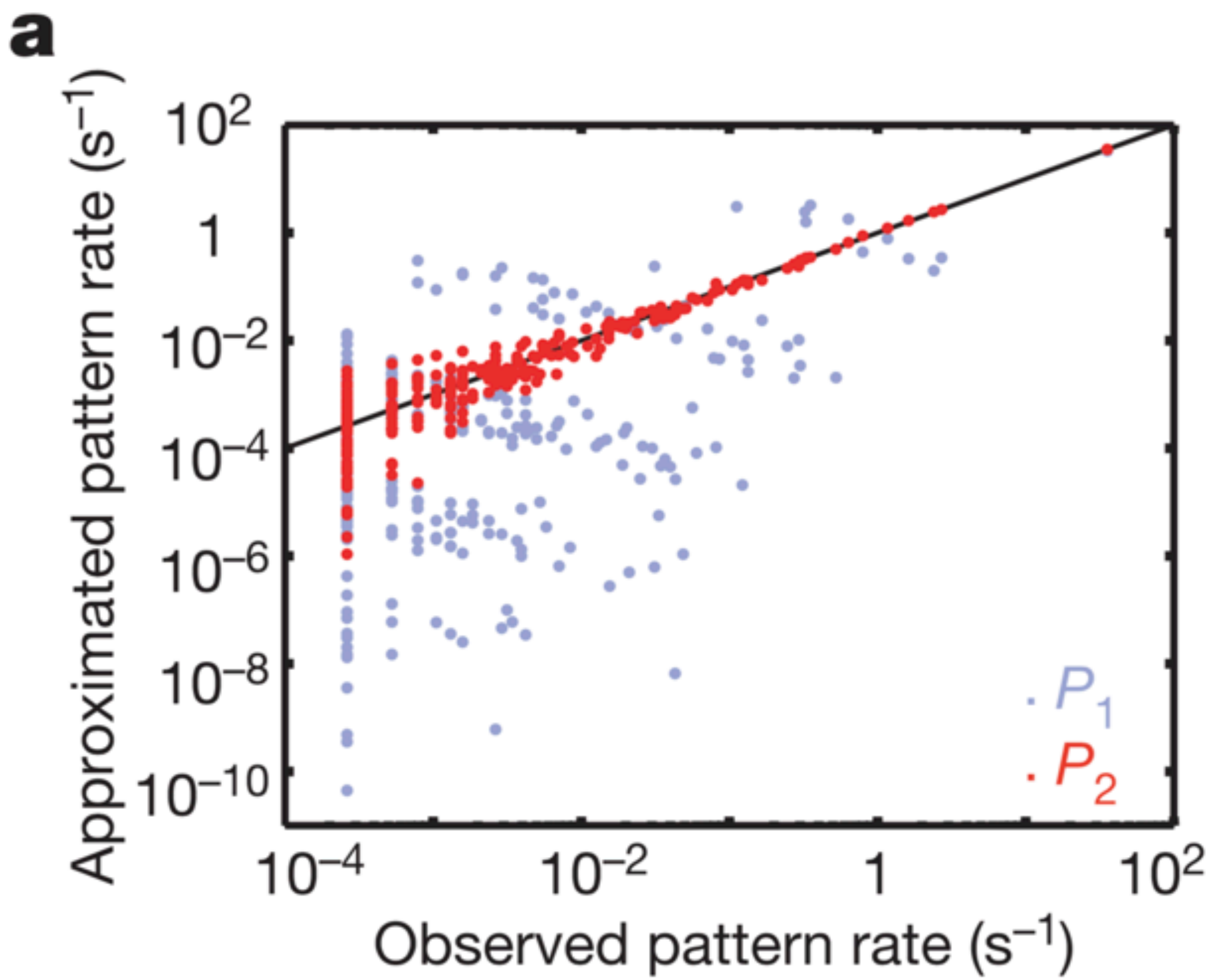
- Assume neural firing is as independent as possible, given the individual neuron firing rates and pairwise correlations.

$$p(x) = \frac{1}{Z} \exp[\sum_i h_i x_i + \sum_{i \neq j} J_{ij} x_i x_j]$$

where h_i are single neuron firing parameters and J_{ij} are pairwise coupling strengths.

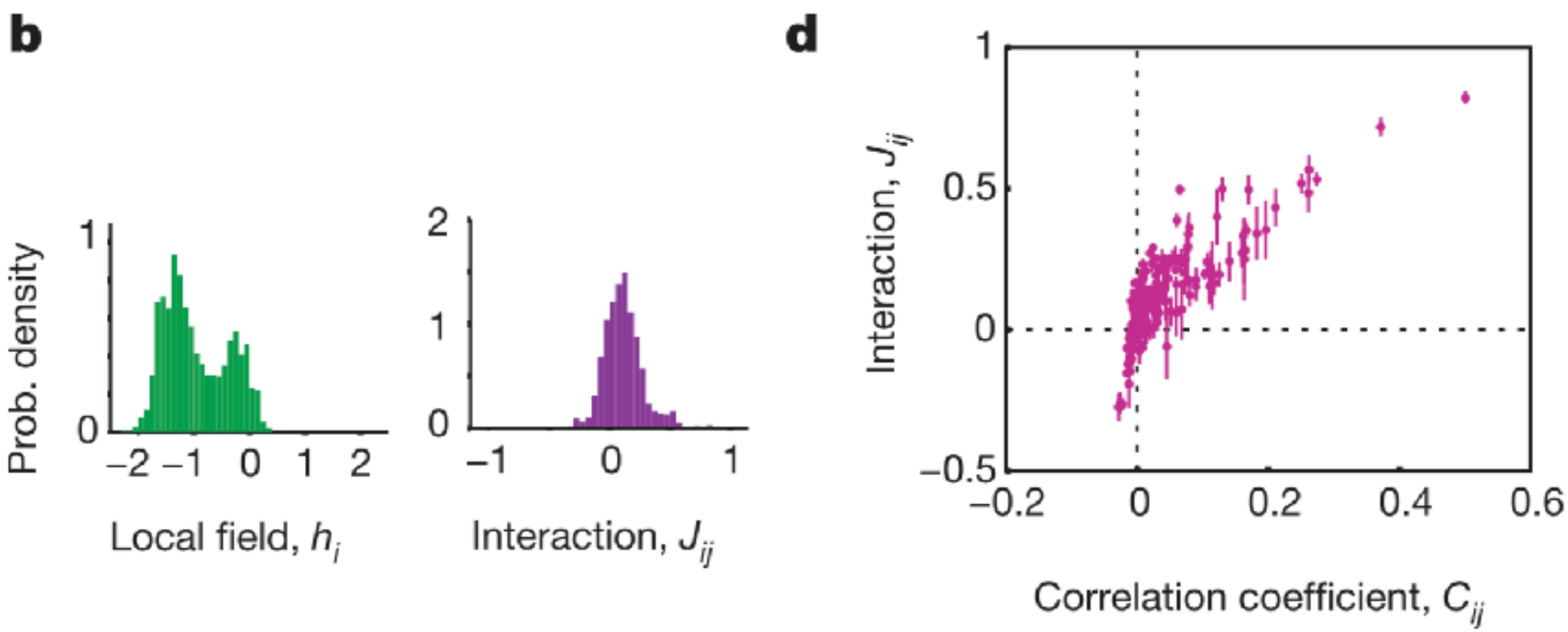
- h_i and J_{ij} are chosen so that $\mathbf{E}[x_i] = \hat{\sigma}_i$ and $\mathbf{E}[x_i x_j] = \hat{c}_{ij}$

Predictions of the pairwise maximum entropy model



Schneidman et al, *Nature* (2006)

Interpreting the parameter fits



Schneidman et al, *Nature* (2006)

Problems with pairwise maxent

- Data-hungry.
- Doesn't accurately match data from large numbers of neurons (even for infinite data).
- Computationally hard to fit parameters for large numbers of neurons.

Further reading

- Schneidman, E., Berry, M.J., Segev, R., Bialek, W., 2006. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440, 1007–1012

Part 2

Recap Tuesday's lecture

- The number of neurons we can simultaneously record from is increasing exponentially. Need analysis tools to make sense of these big data.
- Properties we would like in a good statistical model for neural population data.
- Independent neural model.
- Pairwise maximum entropy model.

Thursday’s lecture

- Elaborate on the pairwise maximum entropy model’s problems.
- Extending the model to include population count information.
- Dichotomised gaussian model as one example alternative to the “maximum entropy” framework.

Problems with pairwise maxent

- Data-hungry.
- Computationally hard to fit parameters for large numbers of neurons.
- Doesn’t accurately match data from large numbers of neurons (even for infinite data).

Problems with pairwise maxent

- Data-hungry.
- Computationally hard to fit parameters for large numbers of neurons.
- Doesn't accurately match data from large numbers of neurons (even for infinite data).

Problems with pairwise maxent

$$p(x) = \frac{1}{Z} \exp[\sum_i h_i x_i + \sum_{i \neq j} J_{ij} x_i x_j]$$

$$Z = \sum_x \left[\exp[\sum_i h_i x_i + \sum_{i \neq j} J_{ij} x_i x_j] \right]$$

- Evaluating the normalisation constant Z (sometimes called the "partition function") is computationally expensive.
- It involves enumerating the relative probabilities of all 2^N possible patterns of x.
- We need to know the value of Z for two reasons:
    - to do gradient ascent on the entropy with respect to the model parameters.
    - to compute absolute pattern probabilities after the model is fit.

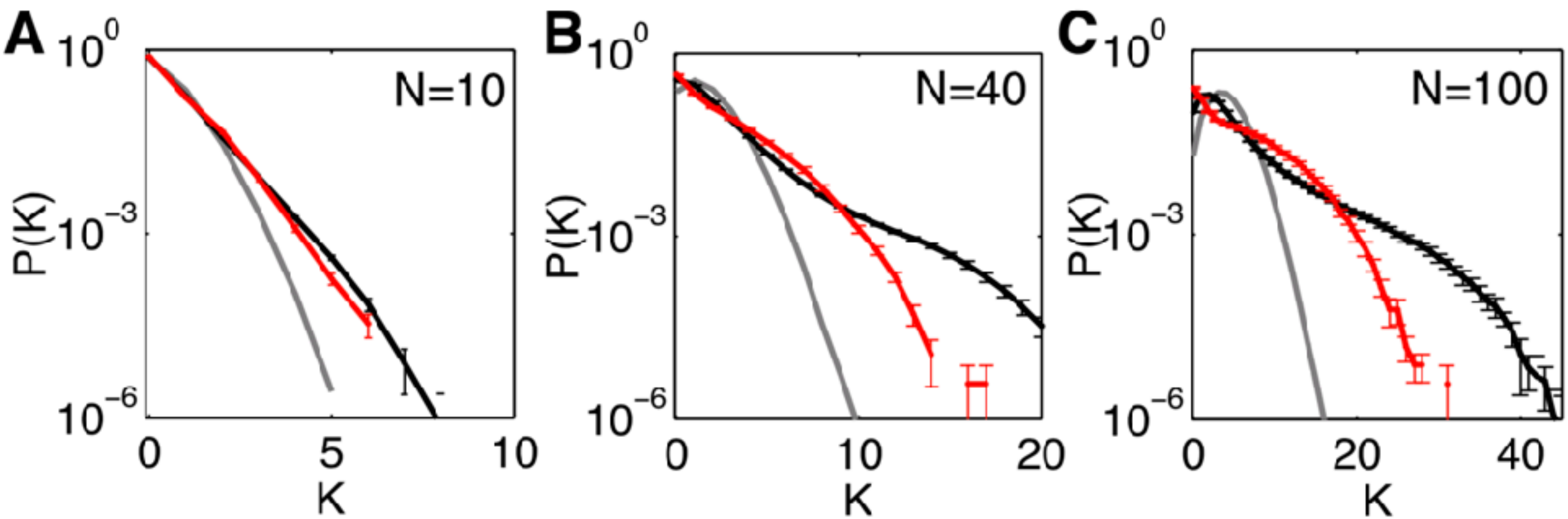
Problems with pairwise maxent

- Data-hungry.
- Computationally hard to fit parameters for large numbers of neurons.
- Doesn't accurately match data from large numbers of neurons (even for infinite data).

Problems with pairwise maxent

- Data-hungry.
- Computationally hard to fit parameters for large numbers of neurons.
- Doesn't accurately match data from large numbers of neurons (even for infinite data).

Problems with pairwise maxent

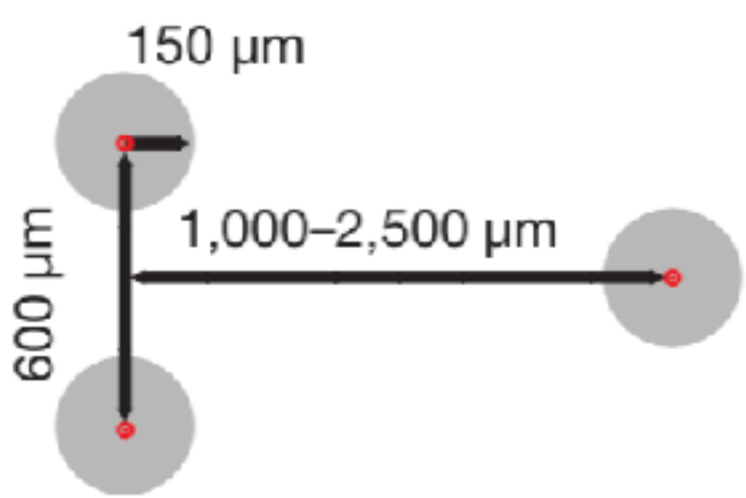


- Model does a good job of prediction synchrony distribution for N=10.
- But accuracy gets worse and worse for larger N.

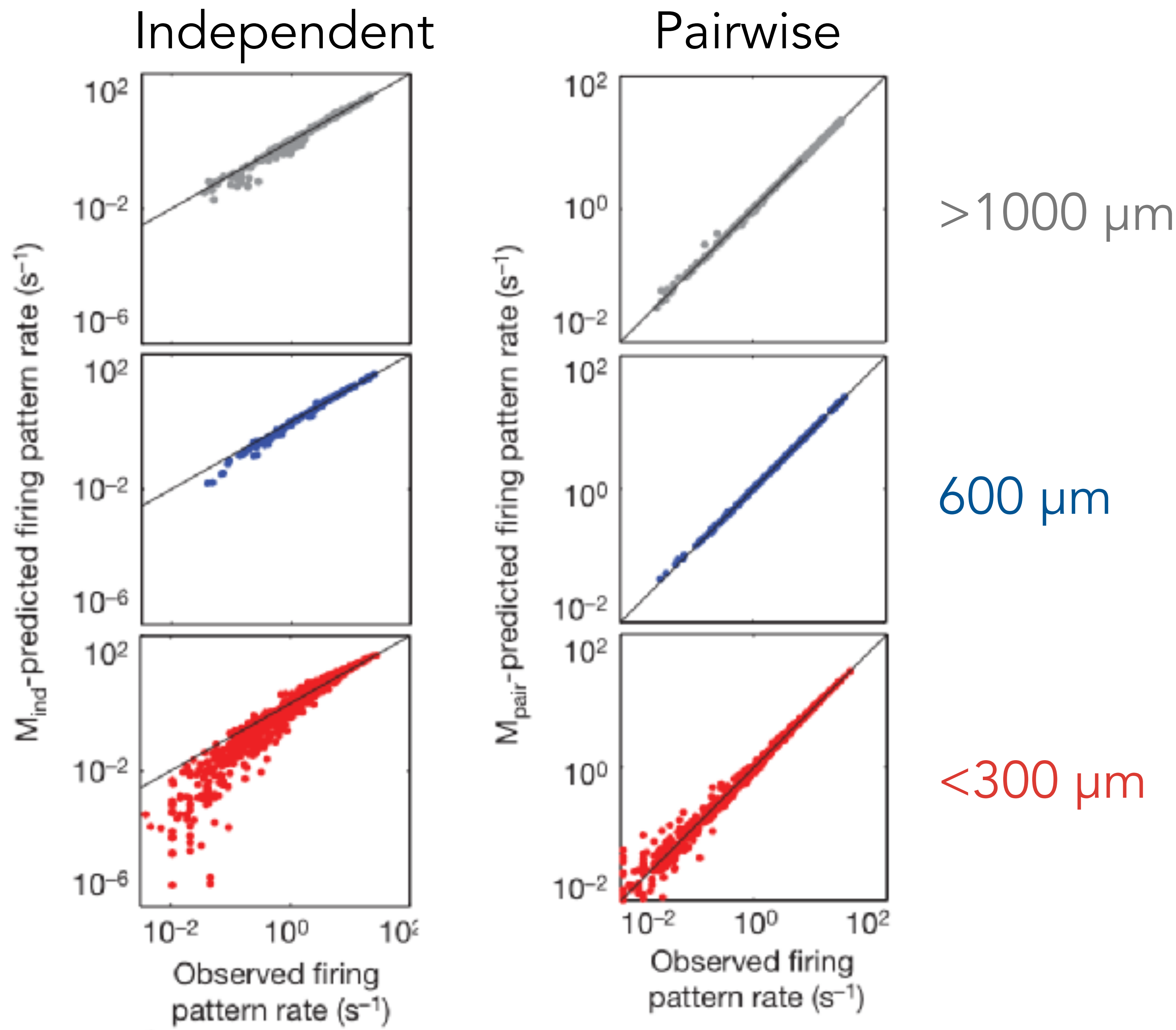
Tkacik et al, *PLoS Comp Biol* (2014)

Problems with pairwise maxent

Recording geometry



Model predicts patterns for monkey visual cortical neurons separated by >600 μm, but not for neighbours.

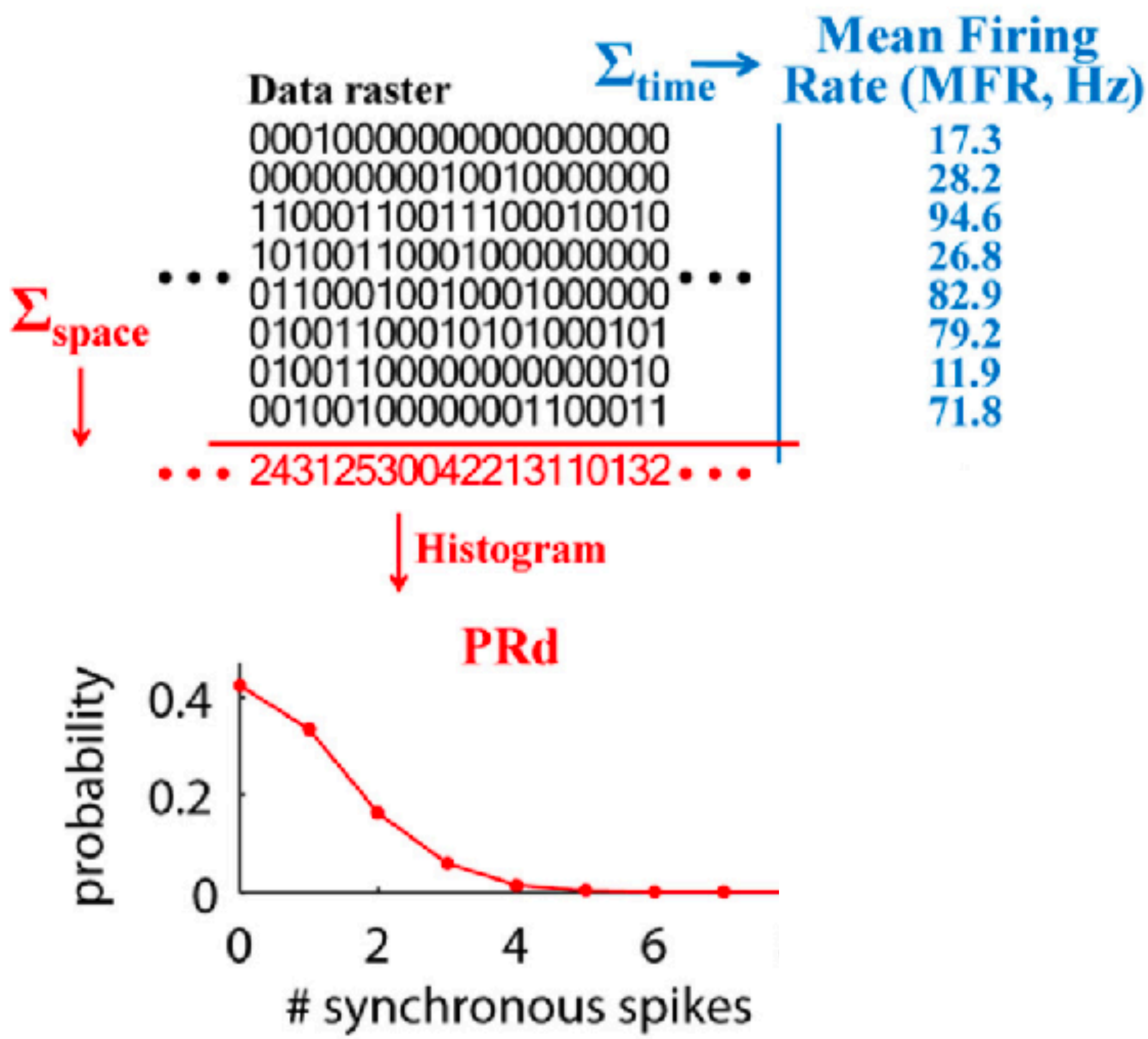


Ohiorhenuan et al, *Nature* (2010)

Why not just continue adding higher order correlations?

- You could go start adding triplet correlations, quadruplet correlations, and so on to the pairwise model.
- But the “curse of dimensionality” **will** beat you: the number of parameters grows quickly in N, and the statistical estimates become more data-hungry.
- A clever alternative statistic is the population count: the distribution for the number of neurons simultaneously active.
- The population count distribution has only N+1 parameters, and is easy to estimate accurately from limited data.

The population count



Okun et al, *J Neurosci* (2012)

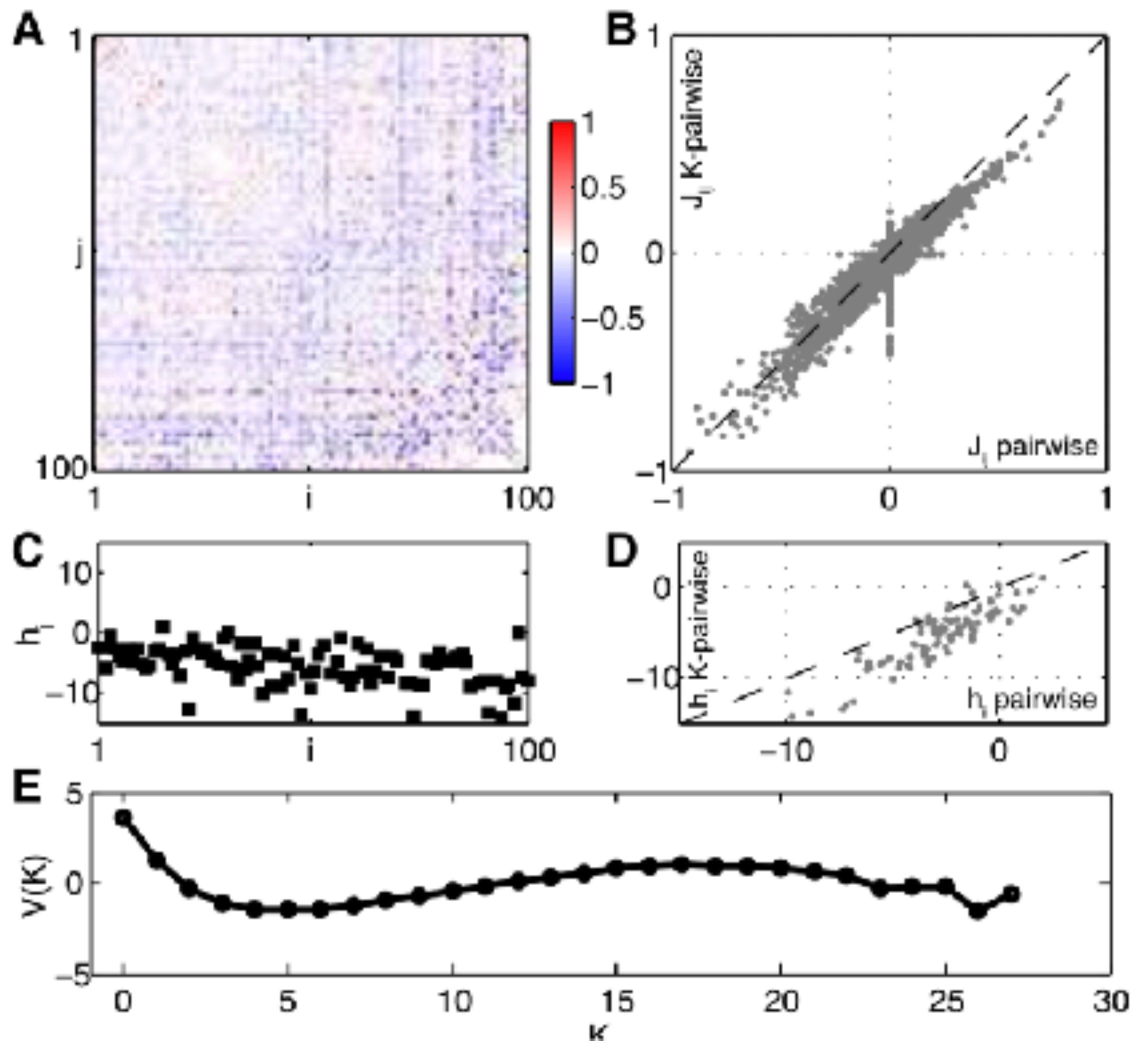
The K-pairwise model

Idea is to have a maximum entropy model that matches not only single neuron firing rates, and pairwise correlations, but also the population count distribution.

$$p(x) = \frac{1}{Z} \exp[\sum_i h_i x_i + \sum_{i \neq j} J_{ij} x_i x_j + V(\sum_i x_i)]$$

Tkacik et al, *PLoS Comp Biol* (2014)

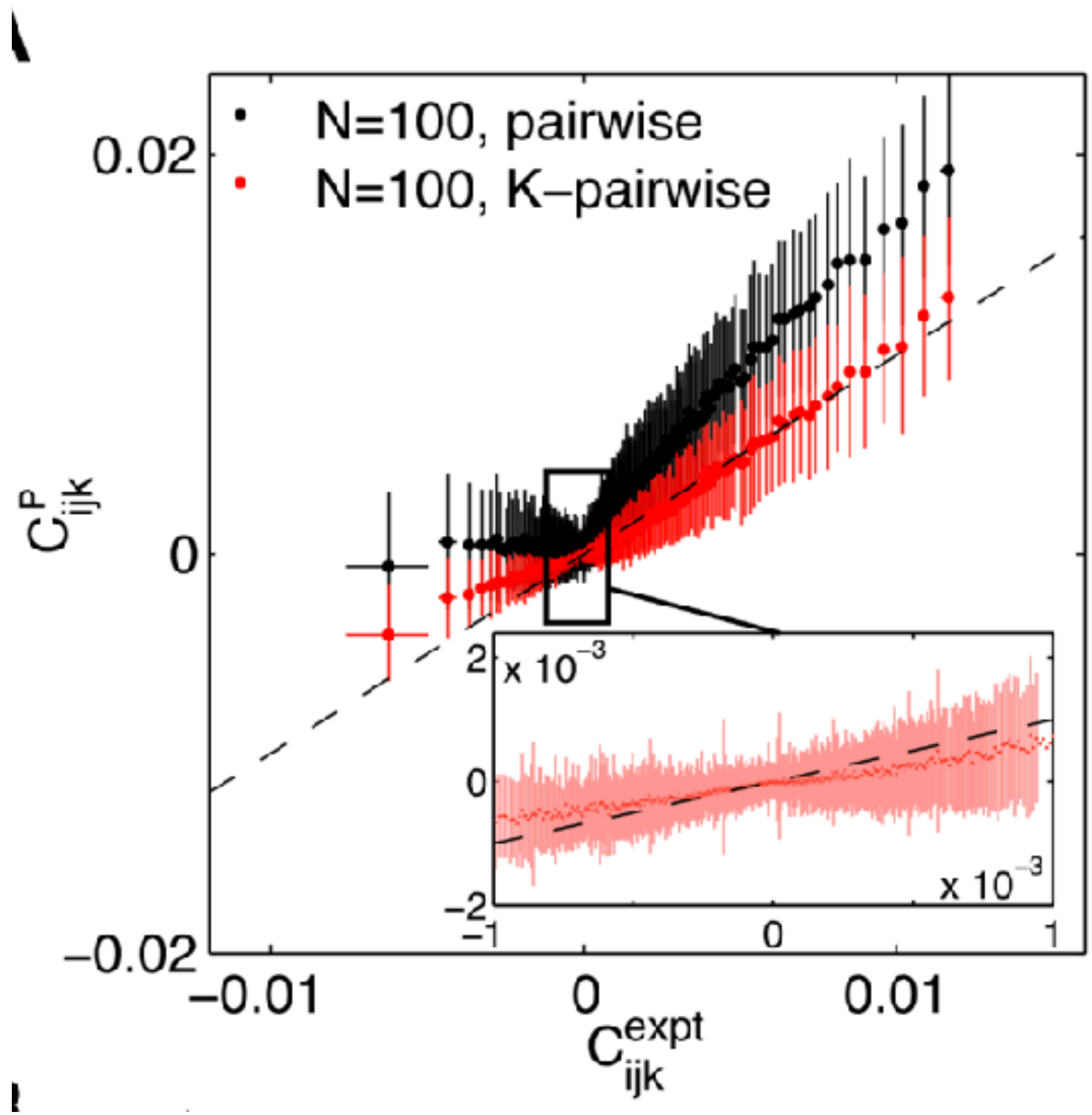
The K-pairwise model



Parameter values for K-pairwise are similar, but not the same, as for standard pairwise model.

Tkacik et al, *PLoS Comp Biol* (2014)

Accuracy of the K-pairwise maximum entropy model



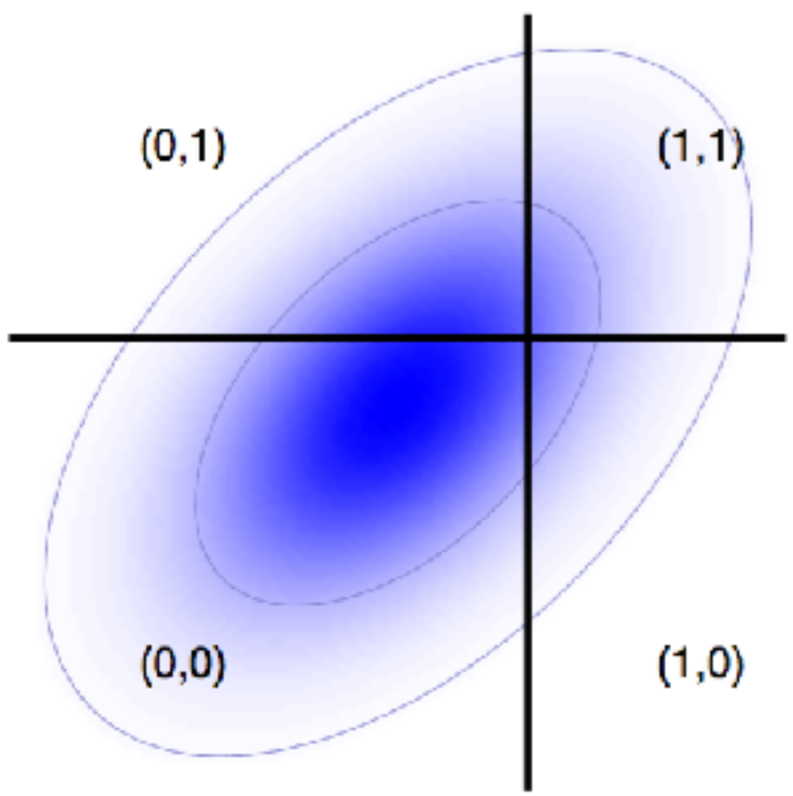
Tkacik et al, *PLoS Comp Biol* (2014)

So done deal?

- No. K-pairwise model still computationally still hard to fit for large *N*, limited correlation structure, doesn’t capture dynamics (temporal correlations), hard to include stimulus dependence.
- Plenty of alternative models out there (reliable interaction model, restricted Boltzmann machine, population tracking model*, cascaded logistic, etc).
- The search continues…

*O’Donnell et al, *Neural Comput* (2016)

The dichotomised gaussian model



- Basic idea is to imagine that our observed binary patterns were generated by thresholding some latent multivariate gaussian distribution.
- This model can match exactly the mean firing rates of individual neurons, and the pairwise correlations.
- Since a multivariate gaussian is entirely defined by its vector of N means and its NxN covariance matrix, there is some unique set of parameters to fit any observed set of neural firing rates and pairwise correlations.
- Interestingly — and in contrast to the pairwise maximum entropy model — this model generates high-order correlations (HOCs).
- Since the model's thresholding property mimics the spike threshold operation of real neurons, we might hope that these HOCs are similar to those in the brain.

Macke et al., *Neural Comput* (2009)

The dichotomised gaussian model

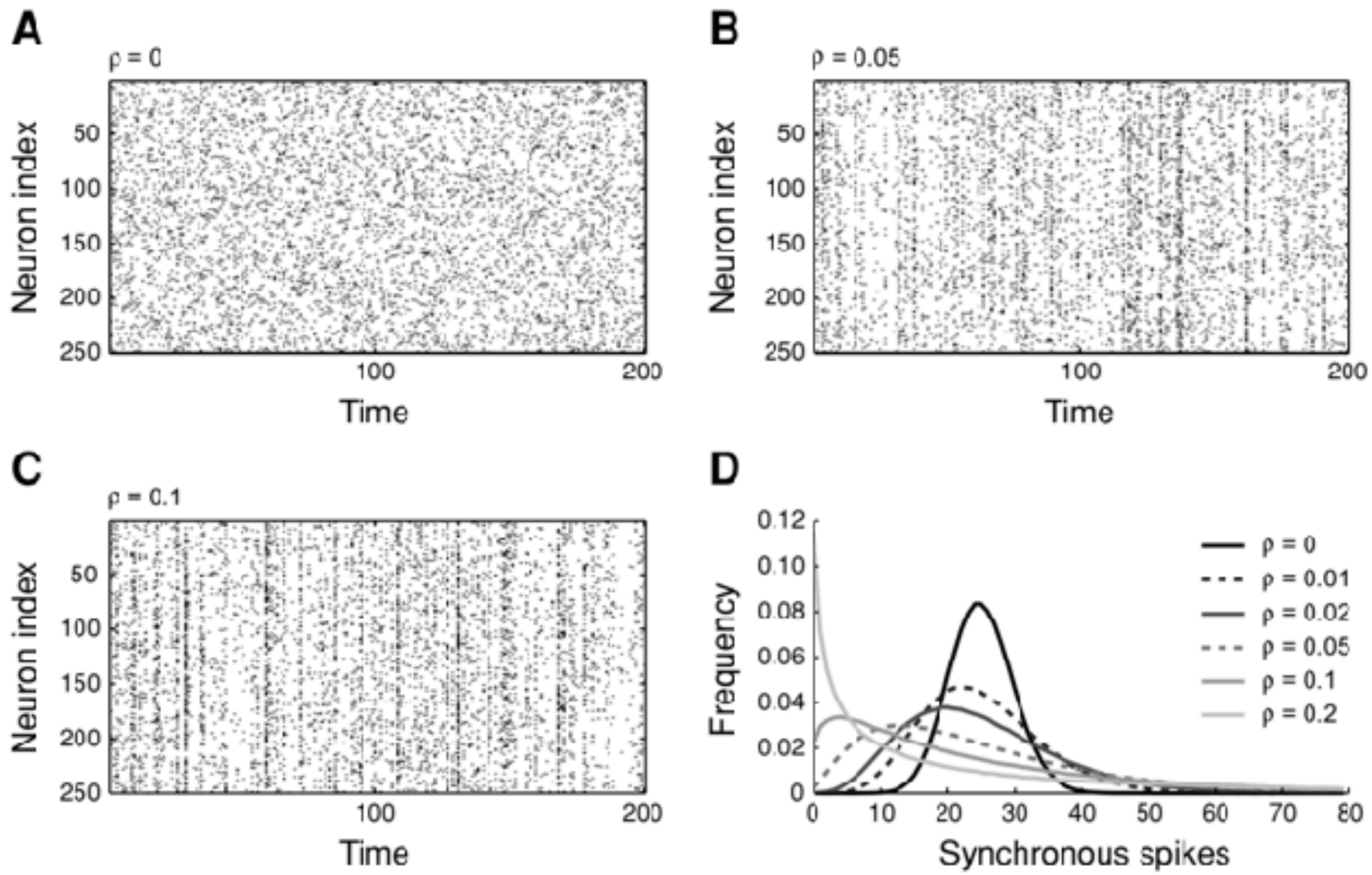


Figure 2: Raster plots of synthetically sampled multineuron firing patterns. The neurons depicted in panels A–C all have the same constant firing rate of 0.1 spike per bin and vary only in their correlation structure. (A) All neurons are independent. (B, C) The pairwise correlation between any pair of neurons is 0.05 and 0.1, respectively. Patterns in which many neurons fire simultaneously occur more frequently with increasing correlation strength. (D) Shows how the probability of observing k out of the 250 neurons to spike simultaneously varies with correlation.

Macke et al., *Neural Comput* (2009)

Summary

- Pairwise maximum entropy model is inaccurate for large N, and for some brain regions, and in certain activity regimes.
- Can improve model a lot by adding in population count information.
- Many limitations remain however.
- The dichotomised gaussian model is one different approach (among many) to the maximum entropy framework.
- The search continues for better statistical models for neural population data.

Further reading

- Tkacik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., and Berry, M.J. (2014). Searching for collective behavior in a large network of sensory neurons. PLoS Comput Biol 10, e1003408.

- Macke, J.H., Berens, P., Ecker, A.S., Tolias, A.S., and Bethge, M. (2009). Generating spike trains with specified correlation coefficients. Neural Comput 21, 397–423.