Coupled Neural Associative Memories

Amir Hesam Salavati, Amin Karbasi, Amin Shokrollahi

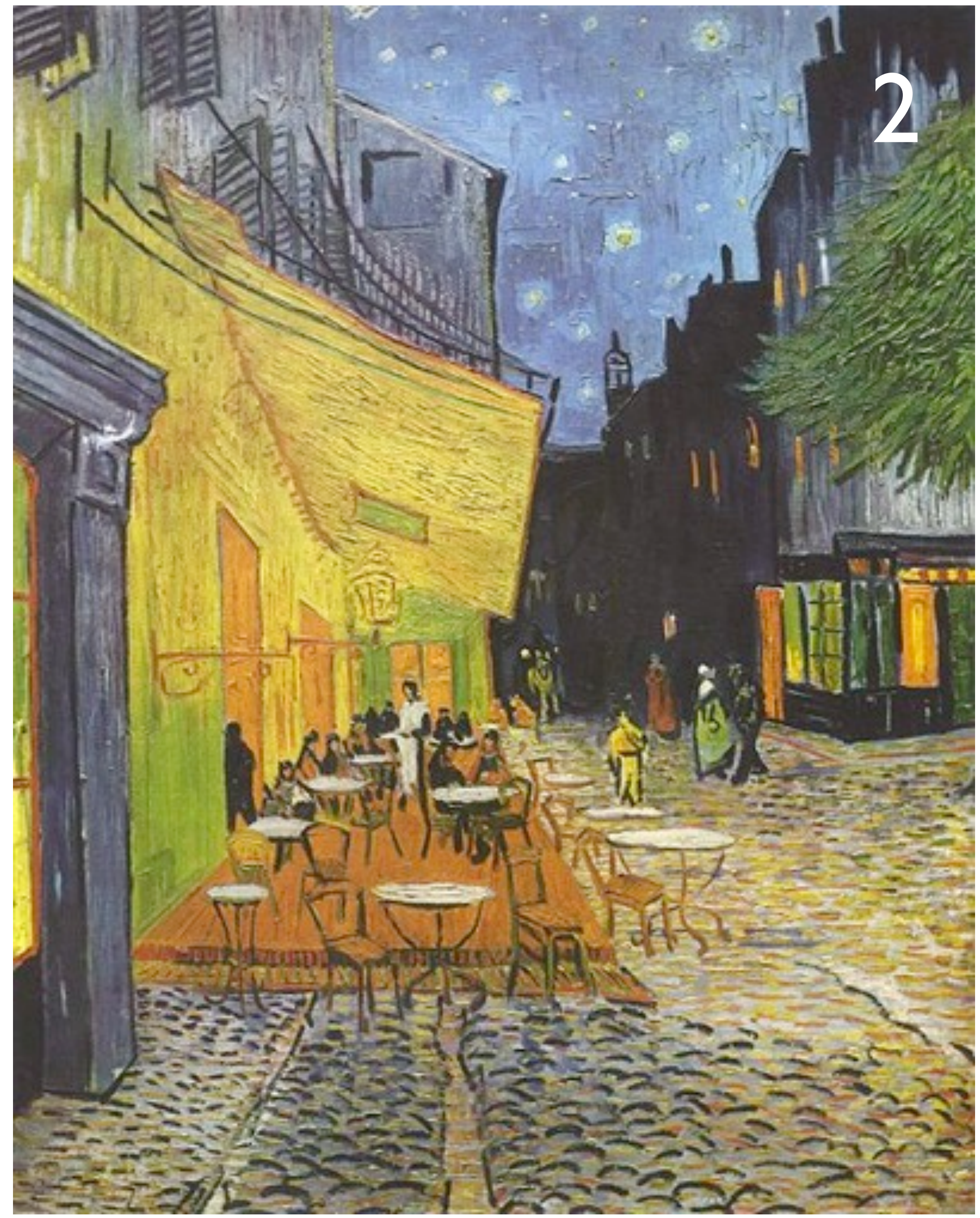
Puzzle!

Puzzle!

Memorize the following images

Puzzle!

Memorize the following images



Now answer!

Now answer!

What was the most similar painting to this one?

Now answer!

What was the most similar painting to this one?

Now answer!

What was the most similar painting to this one?

Neural Associative Memory

Neural Associative Memory

- Memorize
- Recall in presence of noise

Neural Associative Memory

- Memorize
- Recall in presence of noise

Learning

Neural Associative Memory

- Memorize
- Recall in presence of noise

Learning

Good noise tolerance

Neural Associative Memory

- Memorize
- Recall in presence of noise

Learning | Good noise tolerance | Large capacity

Neural Associative Memory

- Memorize
- Recall in presence of noise

Learning | Good noise tolerance | Large capacity

- Artificial neural networks to mimic brain:

  [Hopfield, 1982], [McEliece et al., 1987], [Venkatesh et al. 1989], [Komlos et al., 1993], [Lee, 2001], [Muezzinoglu et al. 2003], [Salavati et al. 2011], [Gripon et al., 2011], [Karbasi et al., 2012]

Neural Associative Memory

- Memorize
- Recall in presence of noise

Learning | Good noise tolerance | Large capacity

- Artificial neural networks to mimic brain:

  [Hopfield, 1982], [McEliece et al., 1987], [Venkatesh et al. 1989], [Komlos et al., 1993], [Lee, 2001], [Muezzinoglu et al. 2003], [Salavati et al. 2011], [Gripon et al., 2011], [Karbasi et al., 2012]

Traditional Approach

Traditional Approach

- [illegible]

Traditional Approach

- [illegible]

[Hopfield, 1982], [McEliece et al., 1987], [Venkatesh et al. 1989], [Komlos et al., 1993], [Lee, 2001], [Muezzinoglu et al. 2003]

Problem: versatility causes **low capacity**

Out of 2^n possible binary vectors of length n, only $O(n)$ can be memorized

Puzzle, Again!

Now memorize these images:

Puzzle, Again!

Now memorize these images:

Now Answer!

Now Answer!

What was the most similar painting to this one?

Structured Patterns

Structured Patterns

- Structured patterns increase the capacity
    - To $O(n^2)$ [Gripon & Berrou 2011]
    - To $O(a^n)$ with $a > 1$ [Kumar et al. 2011]

Structured Patterns

- Structured patterns increase the capacity
  - To $O(n^2)$ [Gripon & Berrou 2011]
  - To $O(a^n)$ with $a > 1$ [Kumar et al. 2011]

Learning

Good noise tolerance

Large capacity

In This Talk...

In This Talk...

- Improve the noise tolerance
  - Some history
  - New perspective: from *convolutional* to *coupled*
- Simulation results
- Ongoing work

The Model
&
Some History

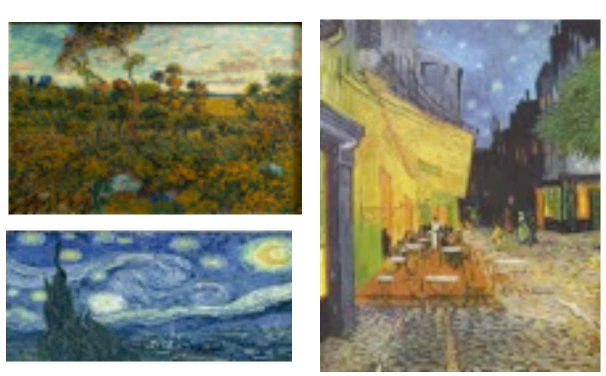
Neural Model

Neural Model

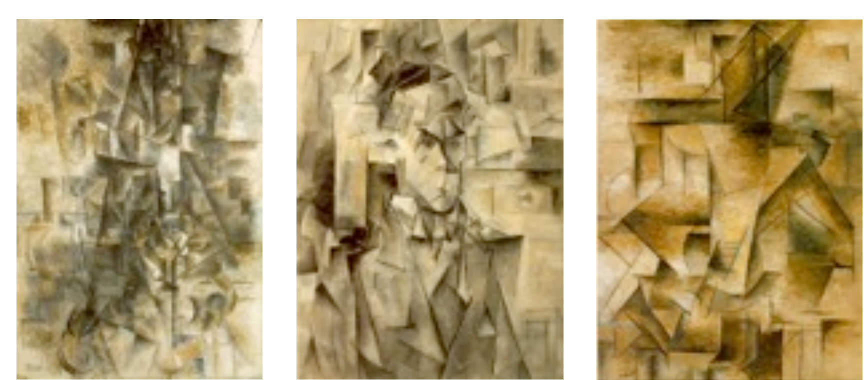
- Patterns
  - Vectors of length n
  - Integer values and non-negative (firing rate)
    - e.g. quantized grey level values

Neural Model

- Patterns
  - Vectors of length n
  - Integer values and non-negative (firing rate)
    - e.g. quantized grey level values
  - Strong local correlations: each sub-pattern comes from a subspace



vs.



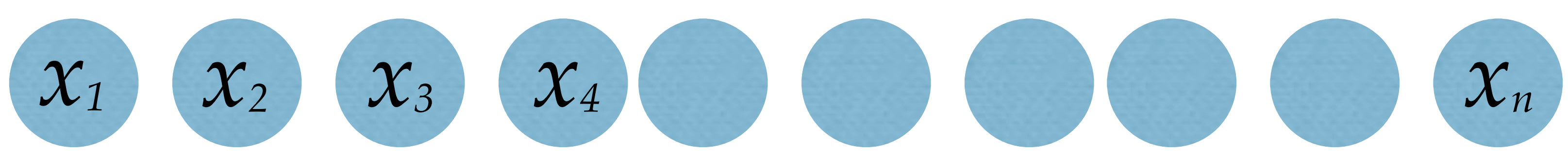
The Learning Process

The Learning Process

- Learn the dual vector(s) orthogonal to the sub-patterns
- Look for sparse orthogonal vectors

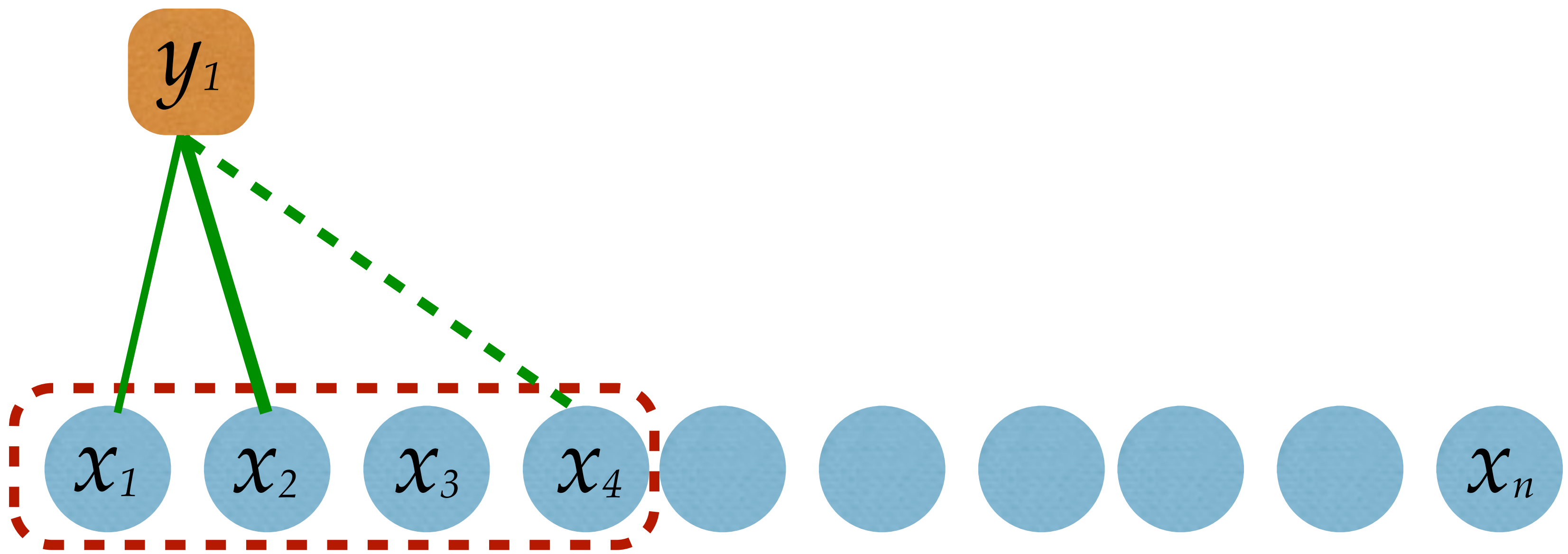
The Learning Process

- Learn the dual vector(s) orthogonal to the subspace patterns
- Look for sparse orthogonal vectors



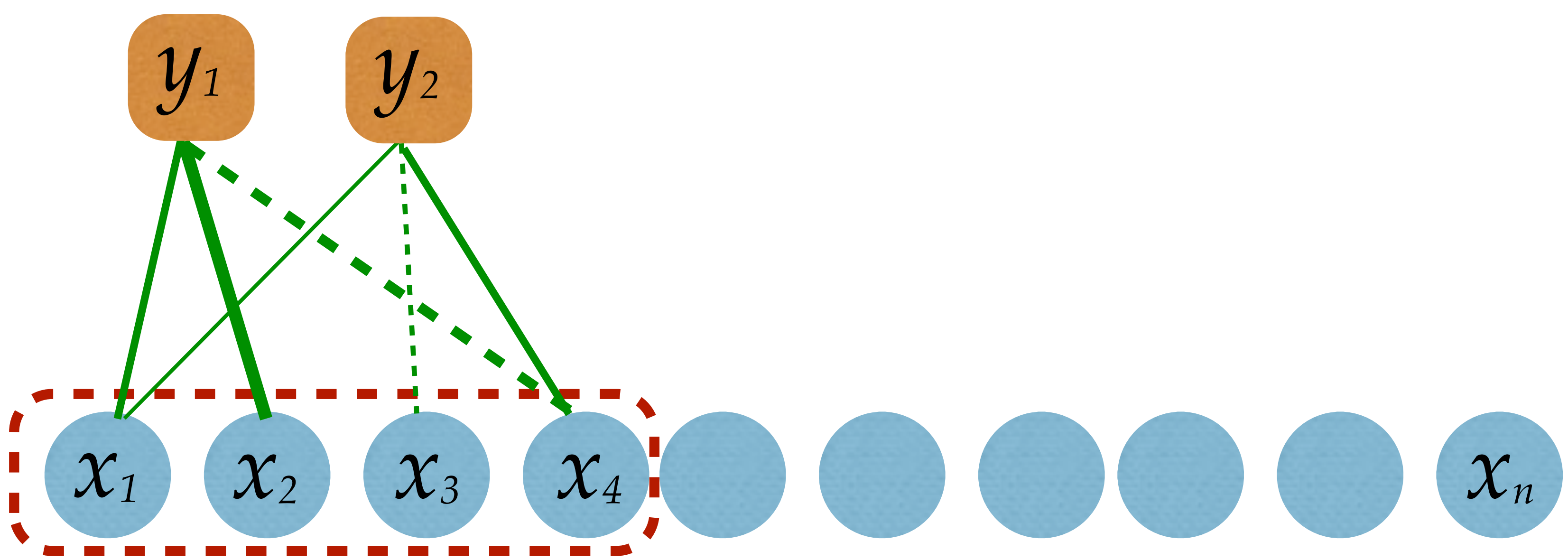
The Learning Process

- Learn the dual vector(s) orthogonal to the sub-patterns
- Look for sparse orthogonal vectors



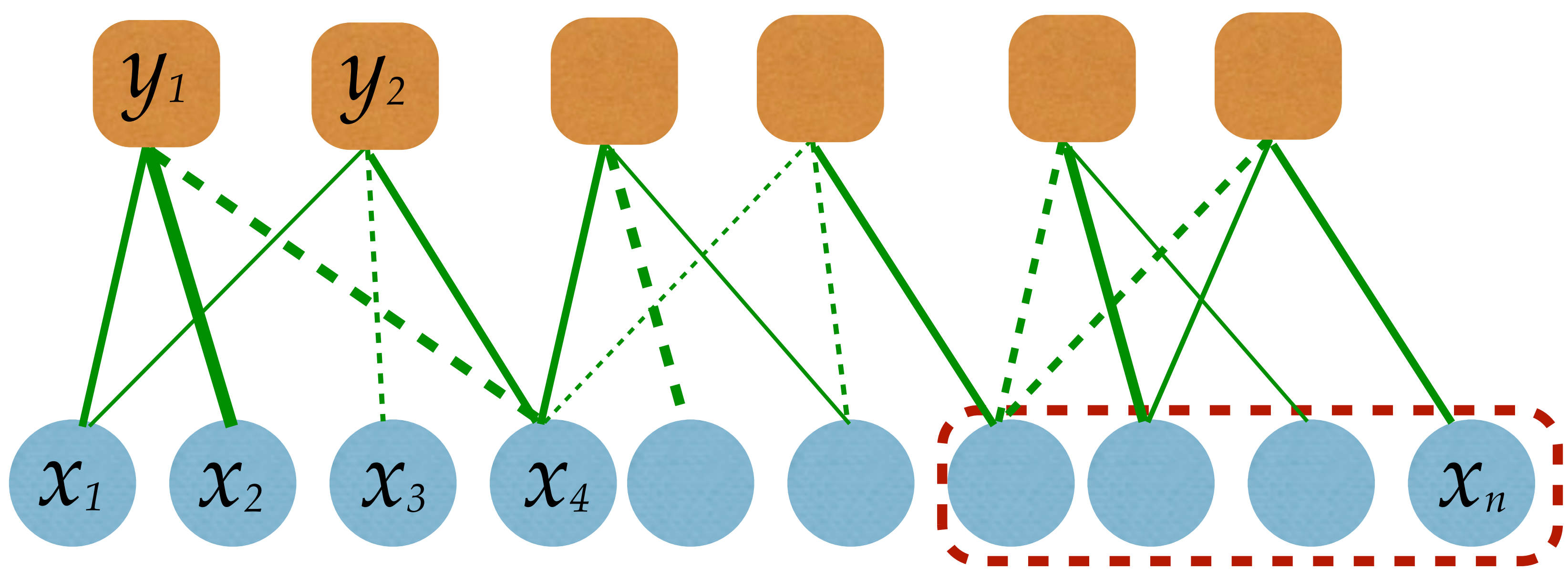
The Learning Process

- Learn the dual vector(s) orthogonal to the sub-patterns
- Look for sparse orthogonal vectors



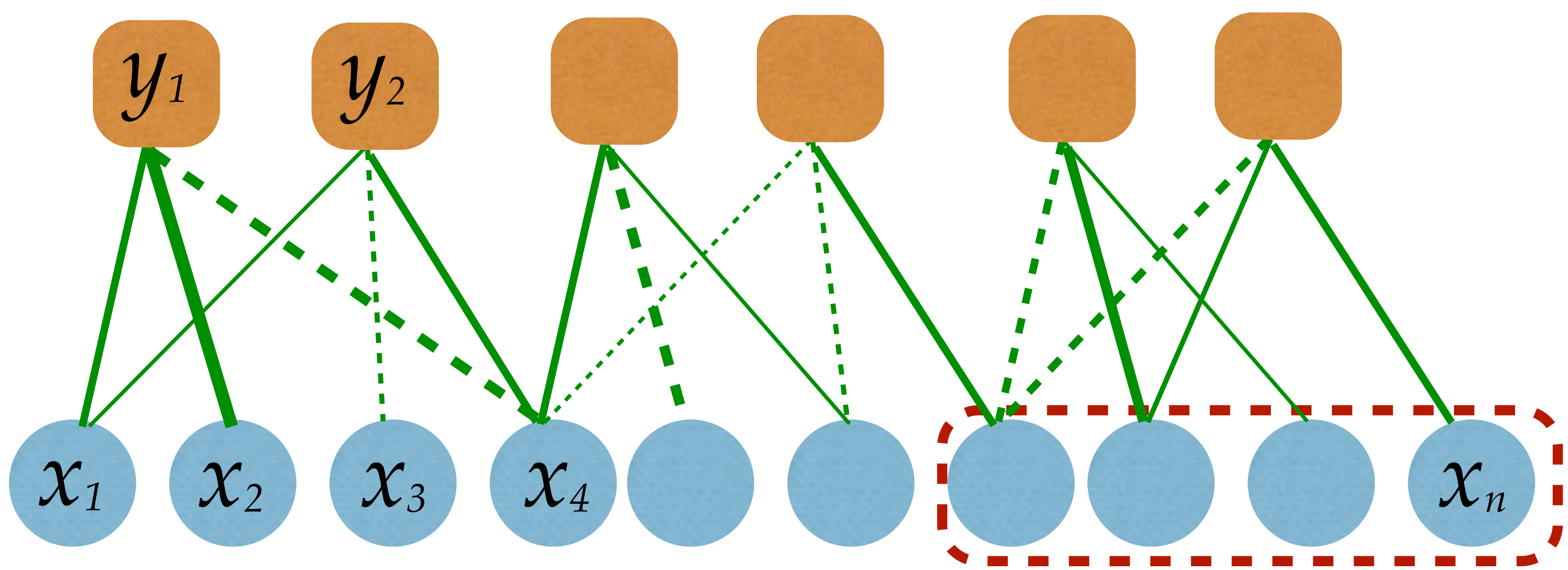
The Learning Process

- Learn the dual vector(s) orthogonal to the sub-patterns
- Look for sparse orthogonal vectors



The Learning Process

- Learn the dual vector(s) orthogonal to the subpatterns
- Look for sparse orthogonal vectors



All in all, we have a parity-check graph!

The Recall Phase

The Recall Phase

- **Theorem** [1]: Each block can correct *1* error reliably.

[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

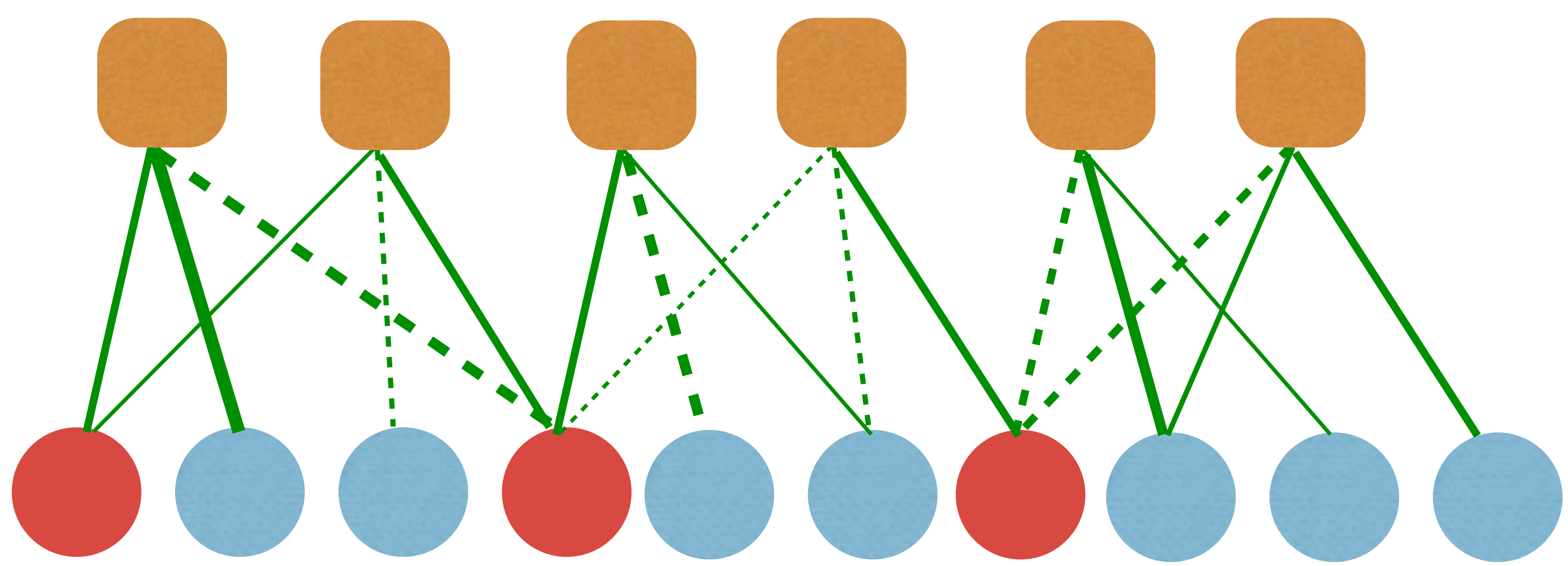
The Recall Phase

- **Theorem** [1]: Each block can correct *1* error reliably.
- However, the overlapping blocks help correct a linear fraction of errors [1].

[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

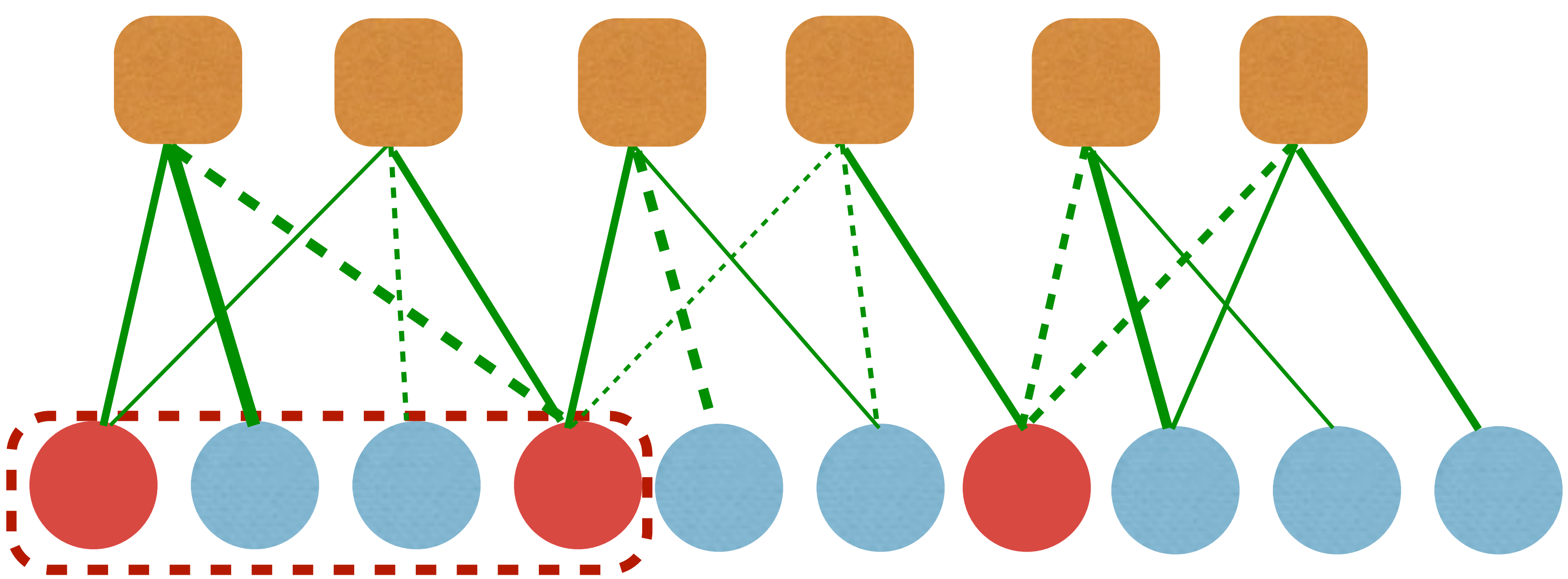
- Theorem [1]: [illegible]
- [illegible] [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

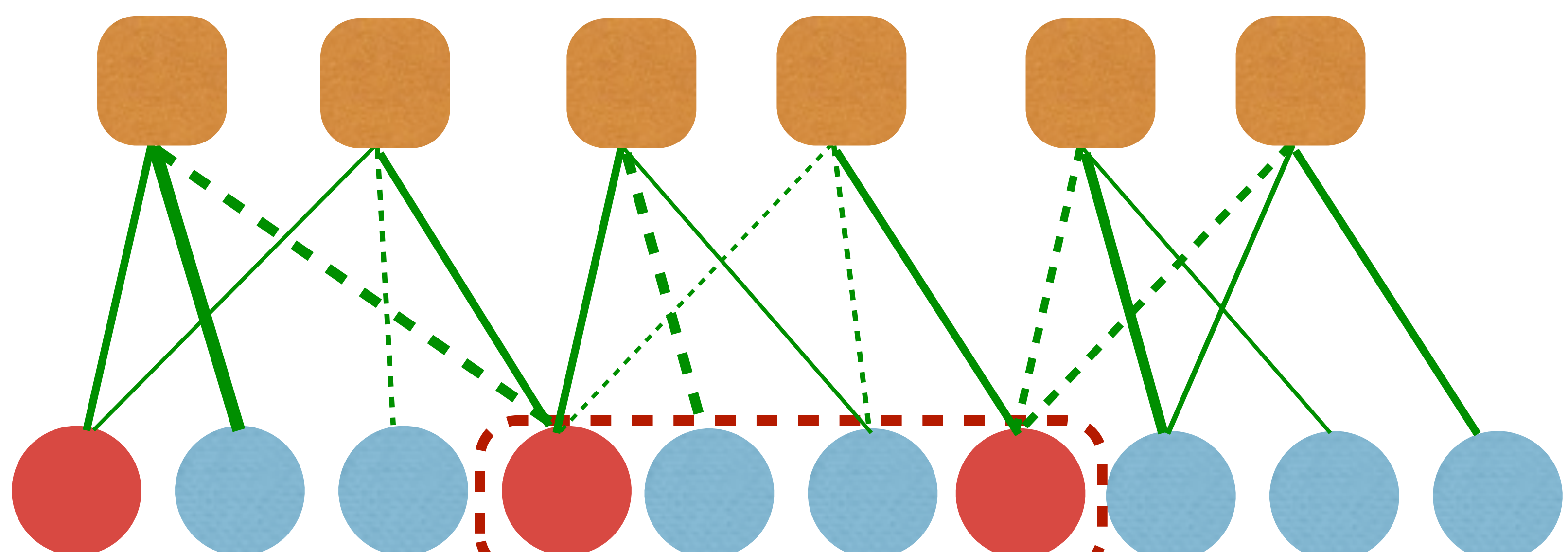
- **Theorem** [1]: Each block can correct 1 error reliably.
- However, the overlapping blocks help the correction of larger number of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

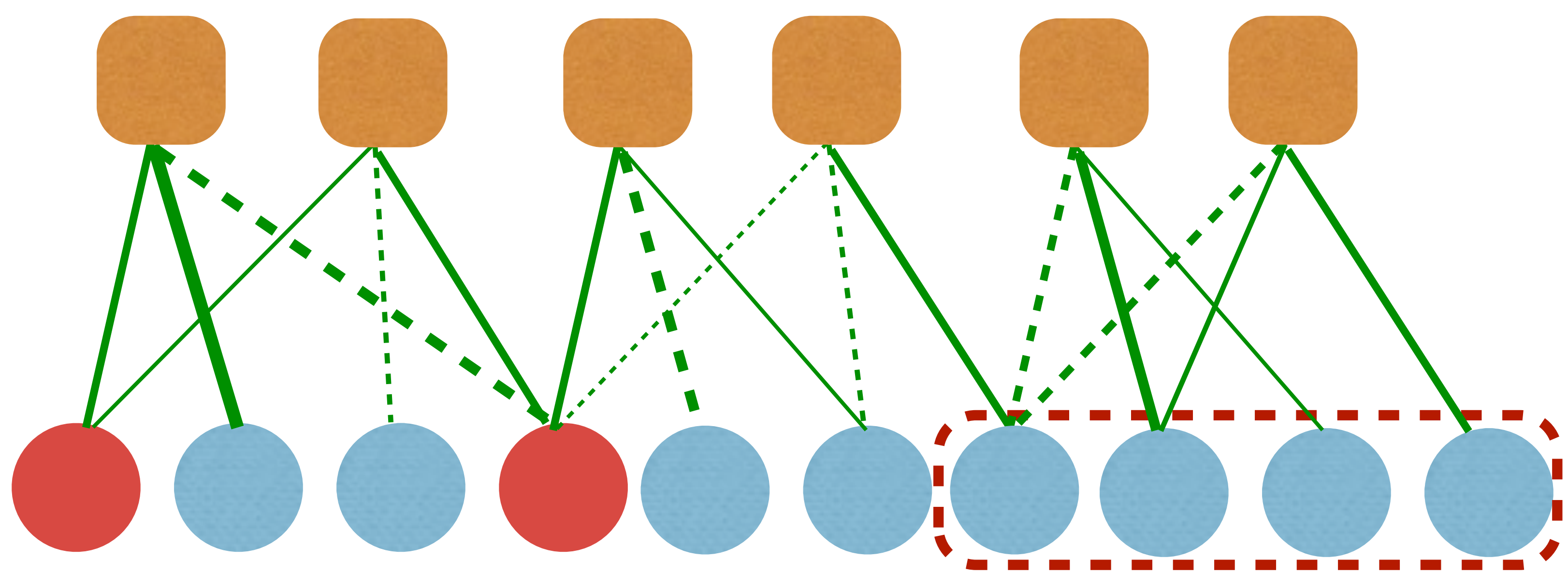
- **Theorem** [1]: Each block can correct 1 error reliably.
- However, the overlapping blocks help achieve correction of linear fraction of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

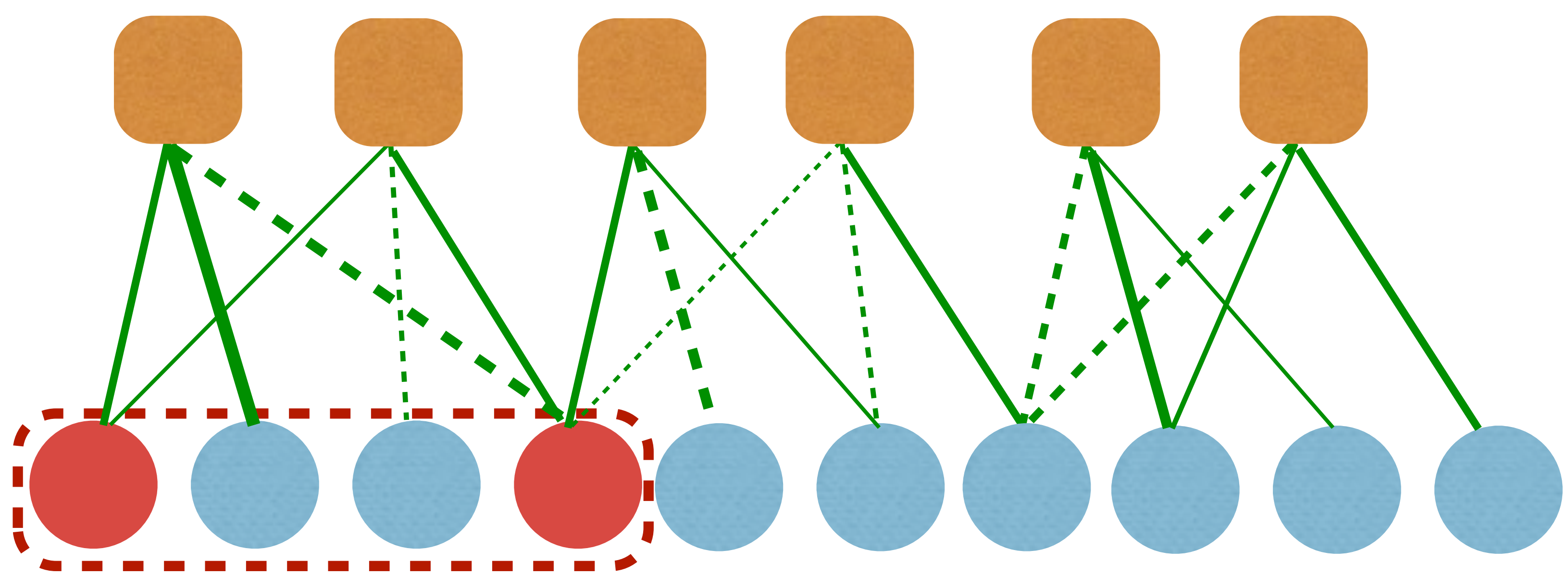
- **Theorem** [1]: Each block can correct 1 error reliably.
- However, the overlapping blocks help achieve correction of a linear fraction of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

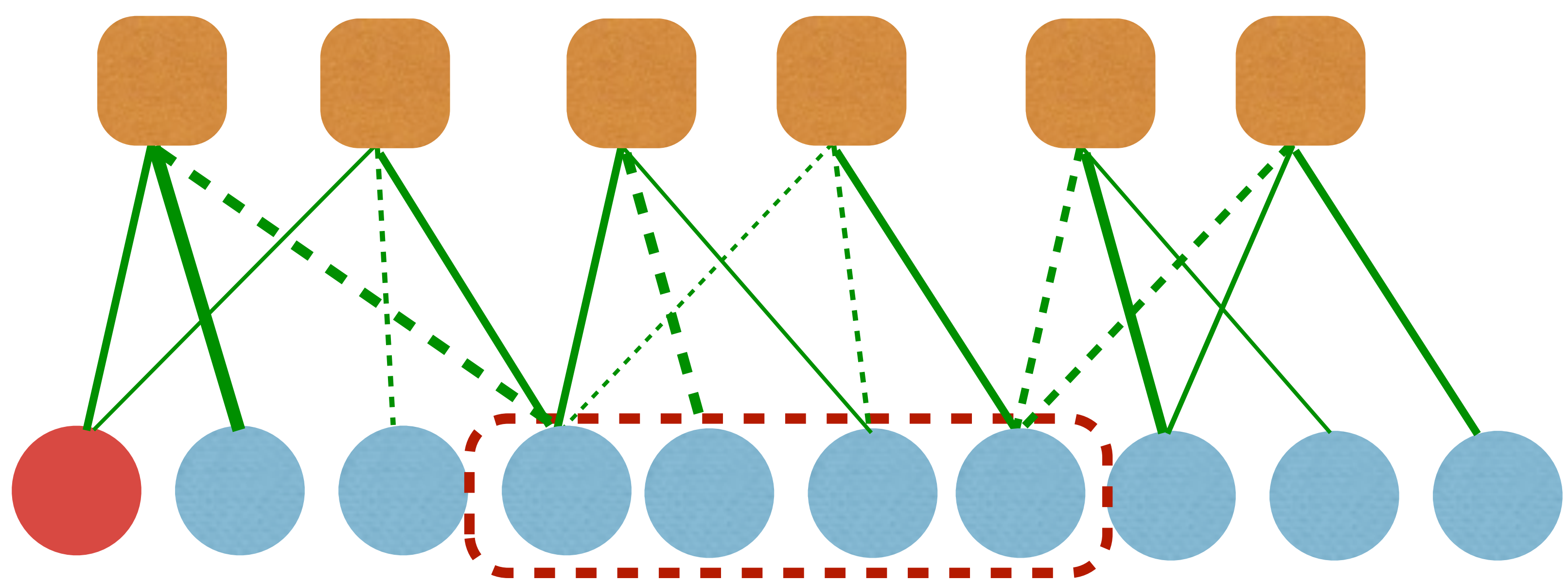
- **Theorem** [1]: Each block could correct *1* error reliably.
- However, the overlapping blocks help to achieve correction of a linear fraction of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

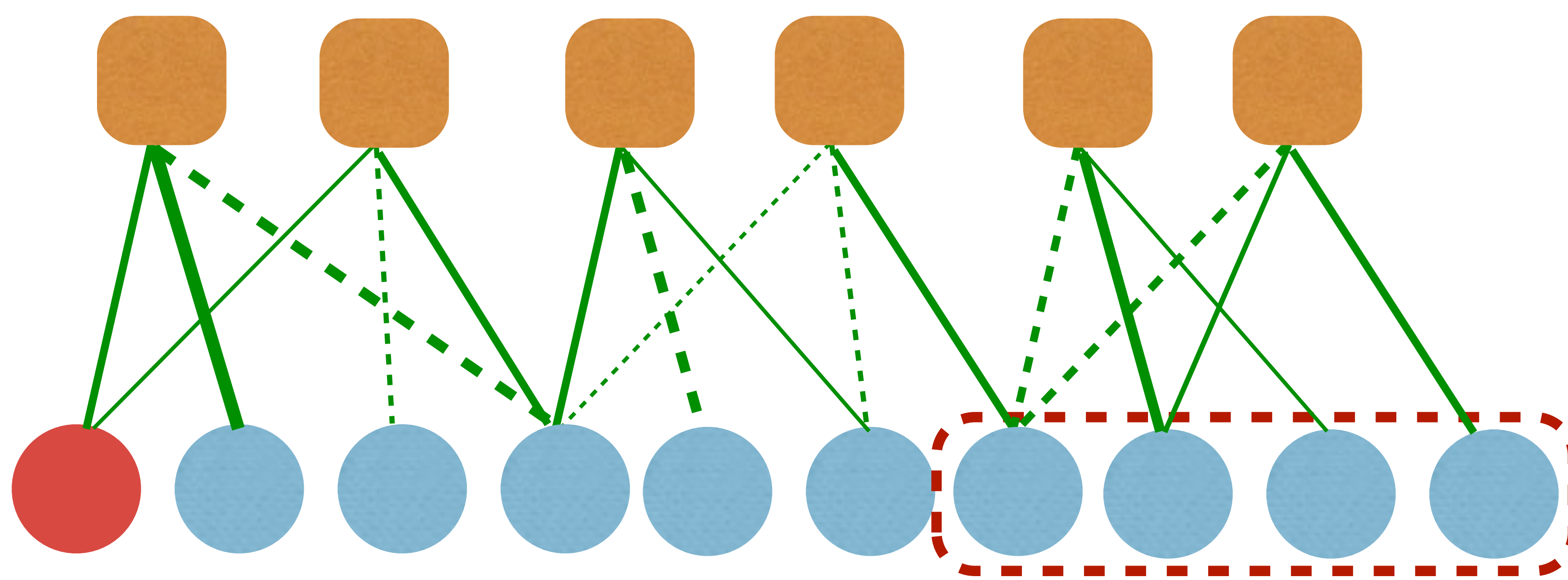
- **Theorem** [1]: Each block can correct *1* error reliably.
- However, the overlapping blocks help to achieve correction of a linear fraction of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

The Recall Phase

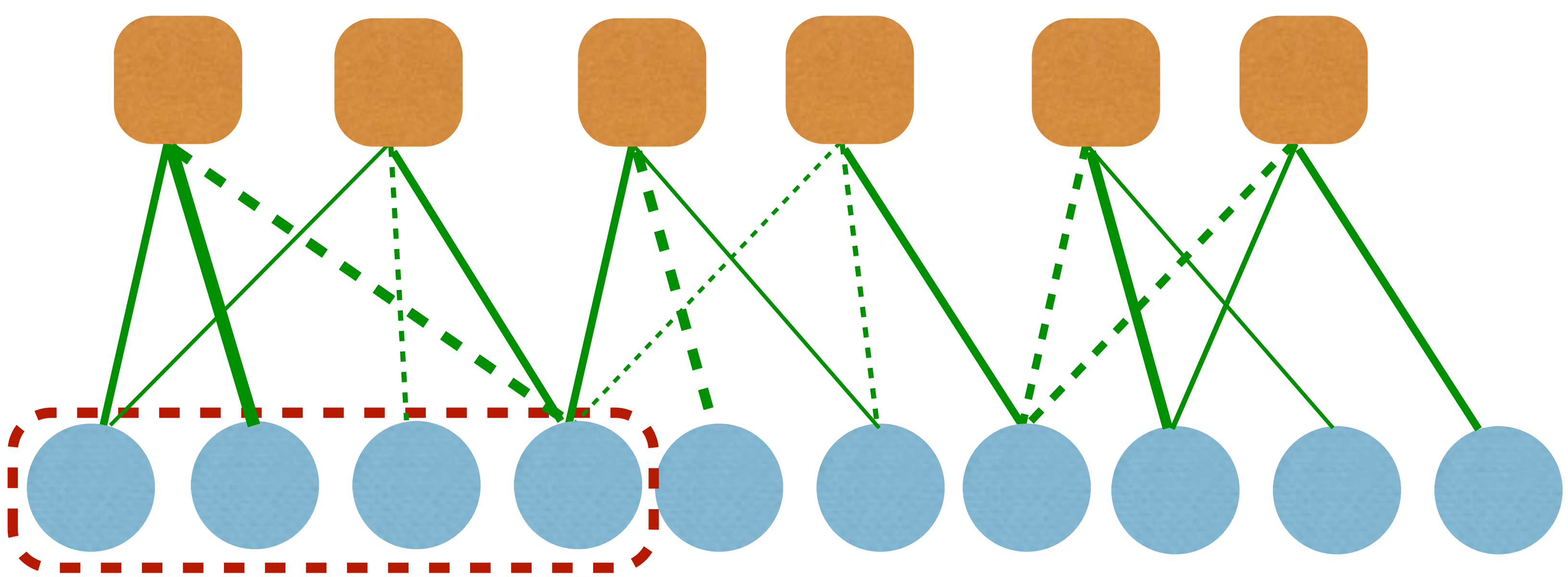
- **Theorem** [1]: Each block can correct 1 error reliably.
- However, the overlapping blocks help to achieve correction of linear fraction of errors [1].



[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

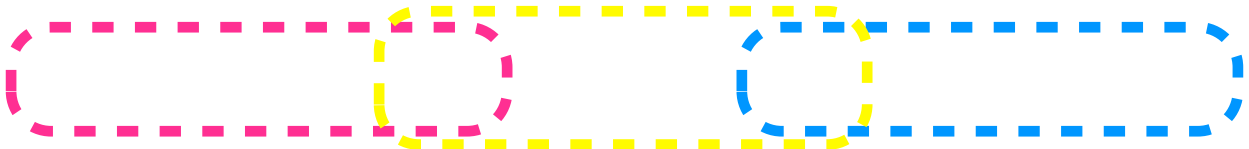
The Recall Phase

- **Theorem** [1]: Each block could correct 1 error reliably.
- However, the overlapping blocks help achieve correction of a linear fraction of errors [1].

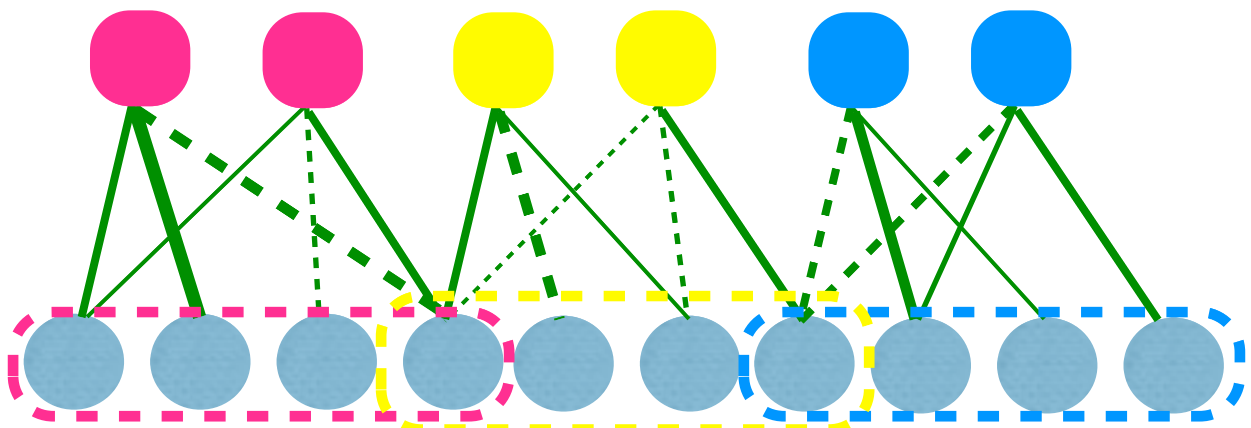


[1] *Iterative learning and denoising in convolutional neural associative memories*
A. Karbasi, A. H. Salavati, A. Shokrollahi, ICML 2013

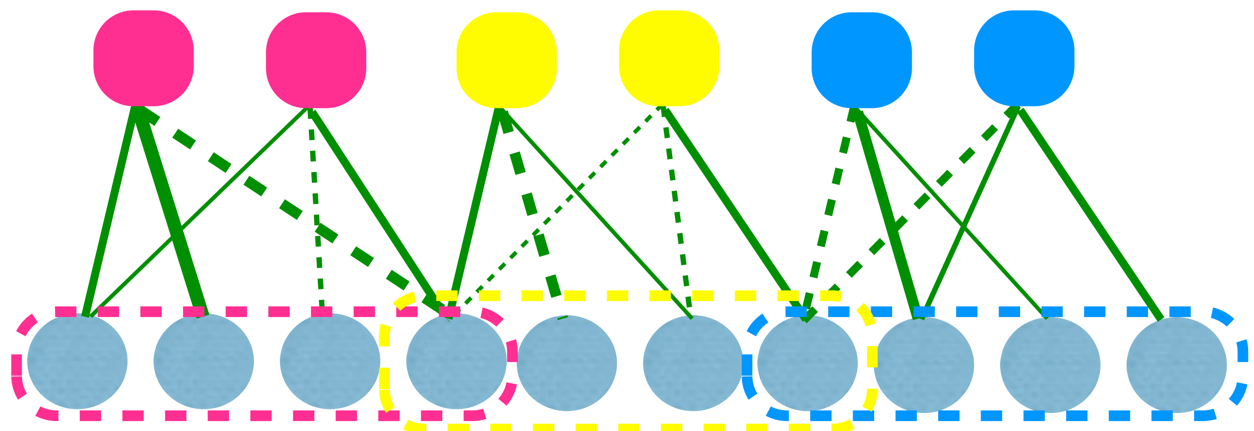
Relations to Peeling Decoder



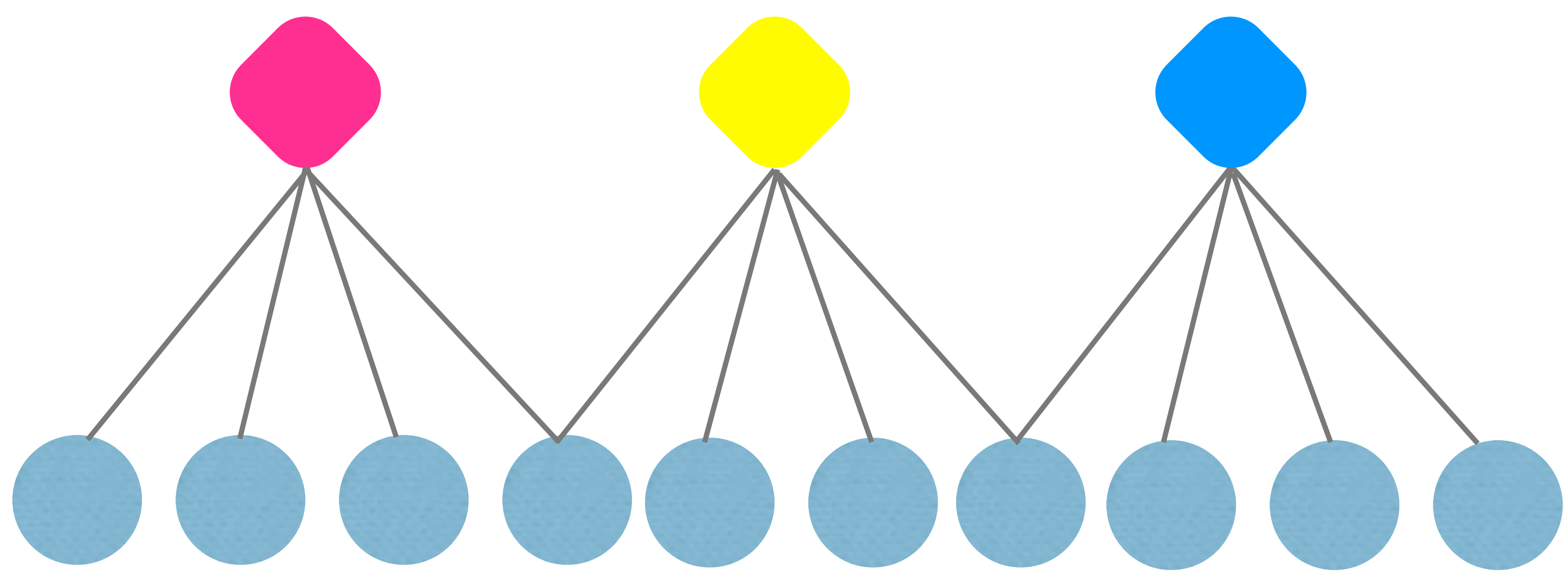
Relations to Peeling Decoder



Relations to Peeling Decoder



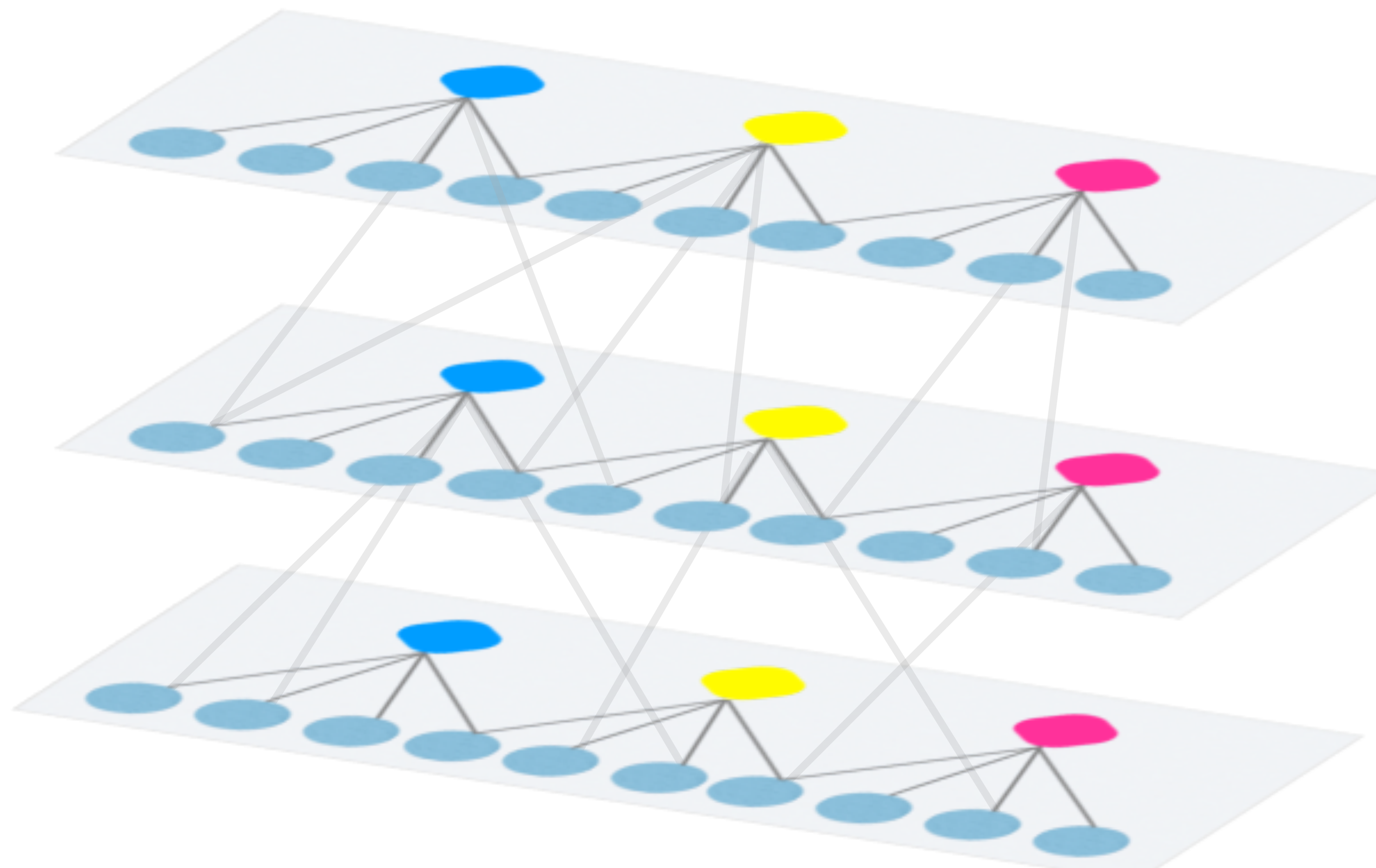
- Very similar to the *Peeling Decoder* on the following graph



Coupled Associative Memories

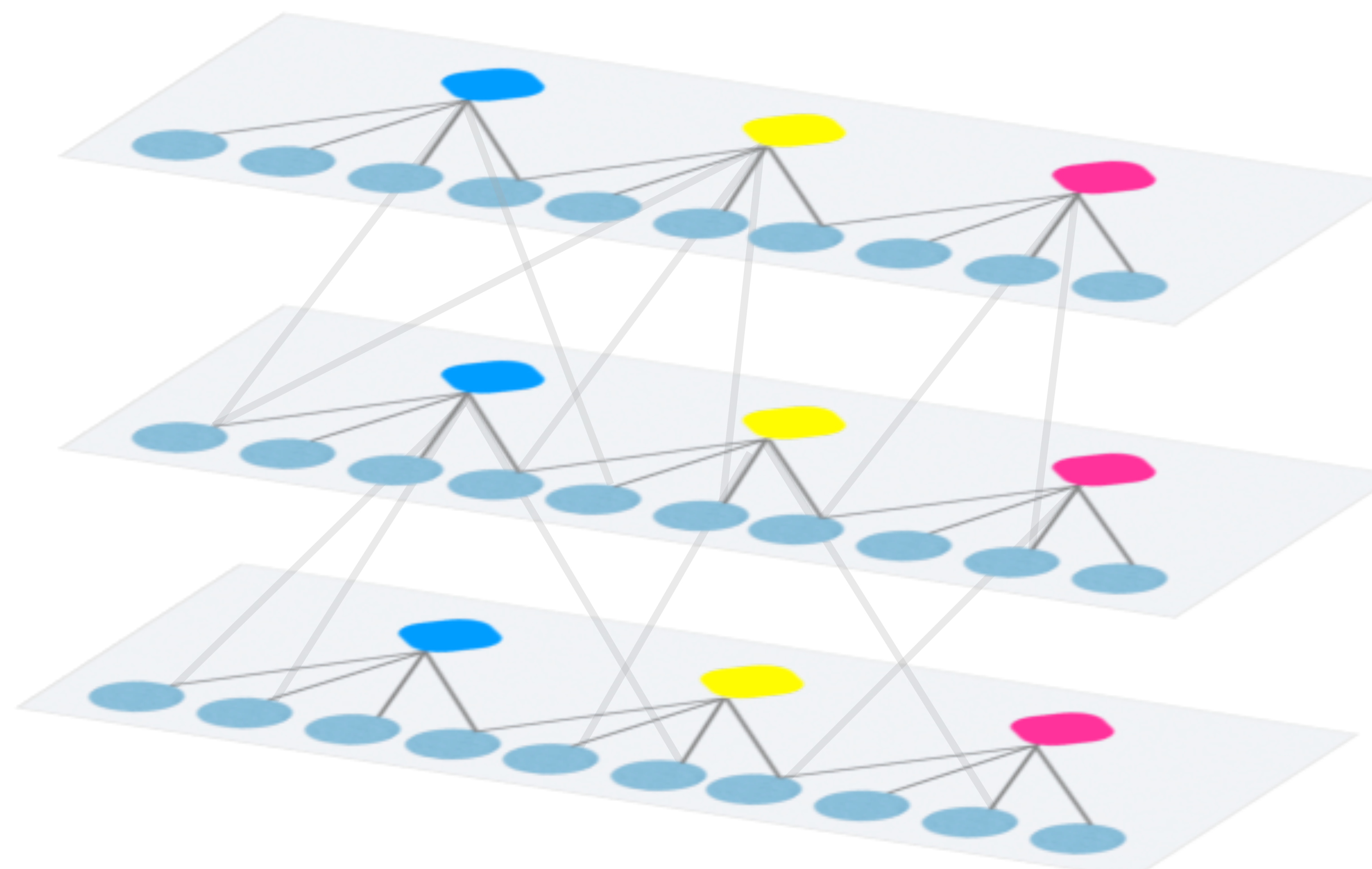
Coupling Neural Graphs

Coupling Neural Graphs



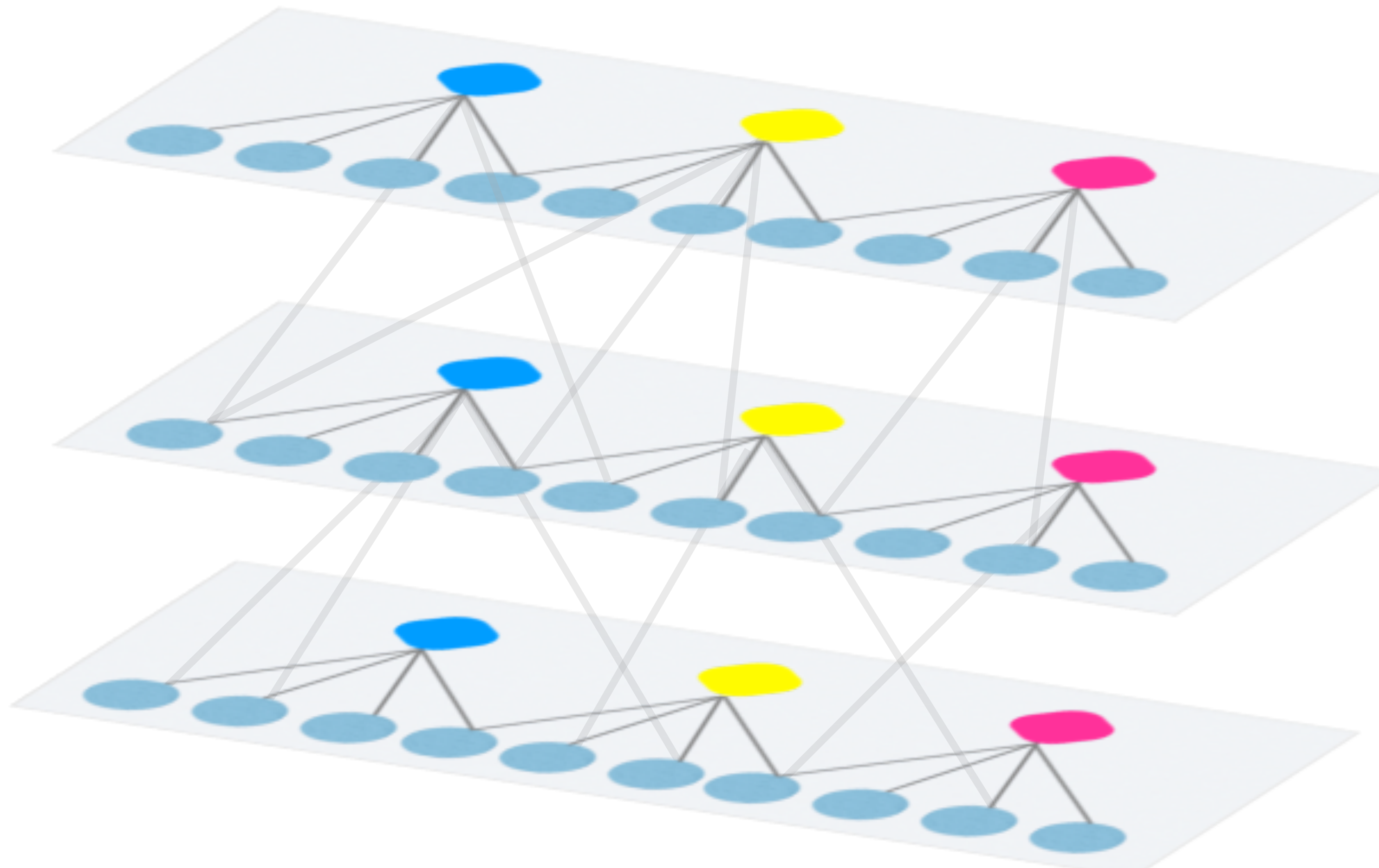
Coupling Neural Graphs

- Same decoding principle: sequential over *clusters* and then *planes*



Coupling Neural Graphs

- Same decoding principle: sequential over *clusters* and then *planes*
- Side information:
  freeze *borderline* neurons to the correct value

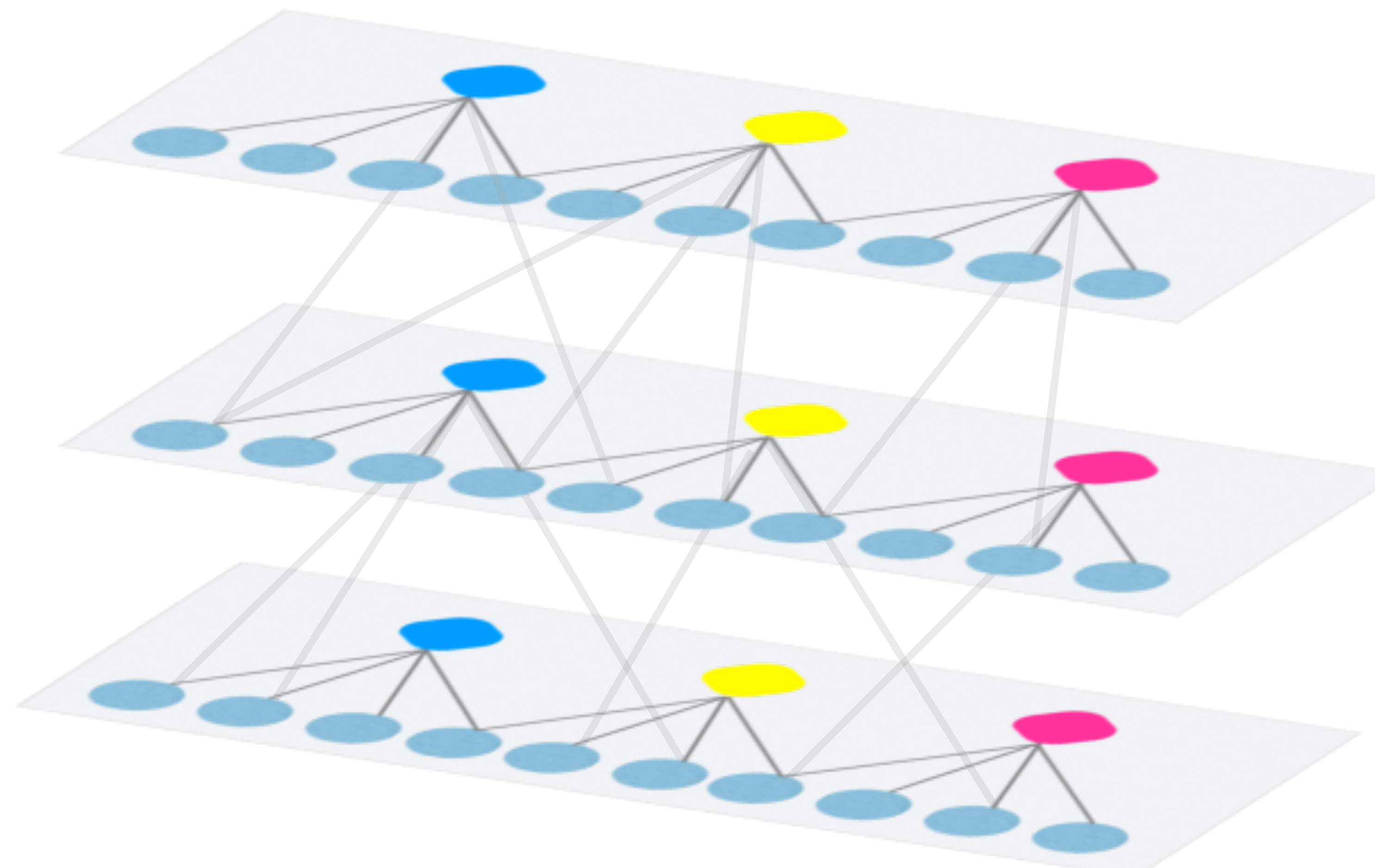


Coupling Neural Graphs

- Same decoding principle: sequential over *clusters* and then *planes*

- Side information:
  freeze *borderline* neurons to the correct value

- Parameters:

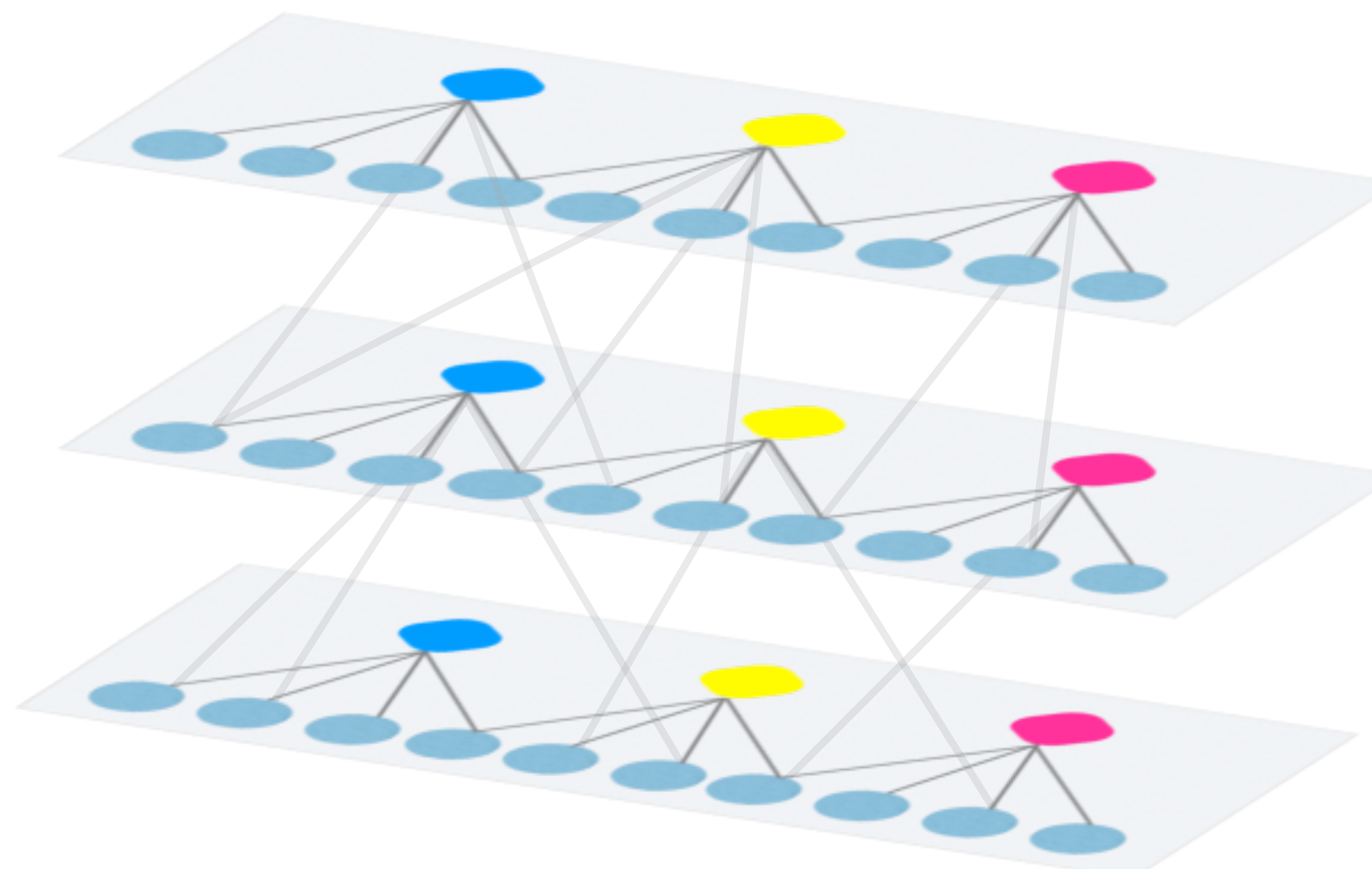


Coupling Neural Graphs

- Same decoding principles: sequentially over *clusters* and then *planes*

- Side information:
  freeze *borderline* neurons to the correct value

- Parameters:
  - D: number of planes

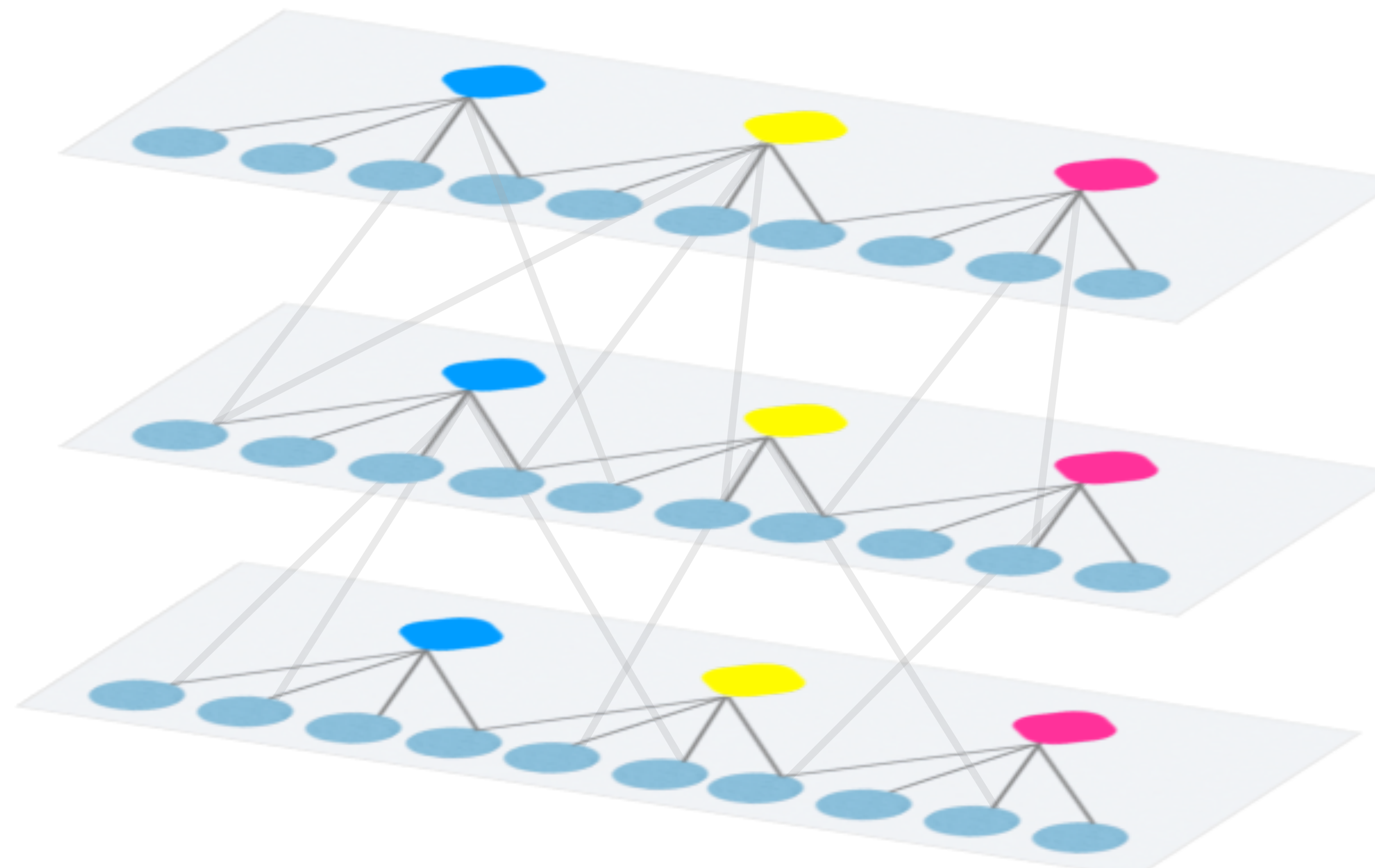


Coupling Neural Graphs

- Same decoding principle: sequential over *clusters* and then *planes*

- Side information:
  freeze *borderline* neurons to the correct value

- Parameters:
  - D: number of plane
  - L: number of clusters in each plane

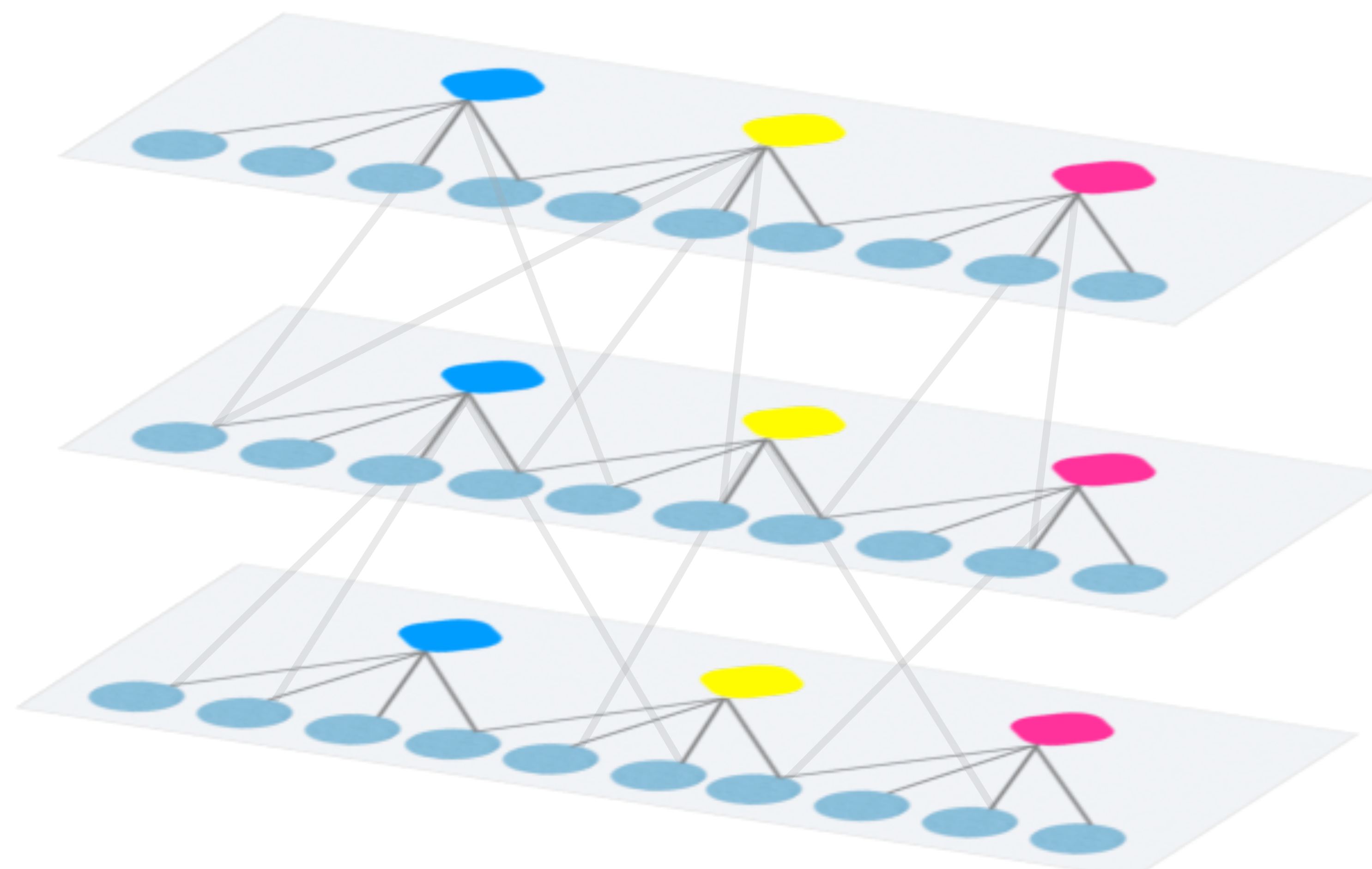


Coupling Neural Graphs

- Same decoding principle: sequential over *clusters* and then *planes*

- Side information:
  freeze *borderline* neurons to the correct value

- Parameters:
  - D: number of planes
  - L: number of clusters in each plane
  - Ω: coupling window



Biological Appeals

Biological Appeals

- Sdeirfondionfromdhecogilvdevels:

  ‘the_dfles’’(adoolot?)

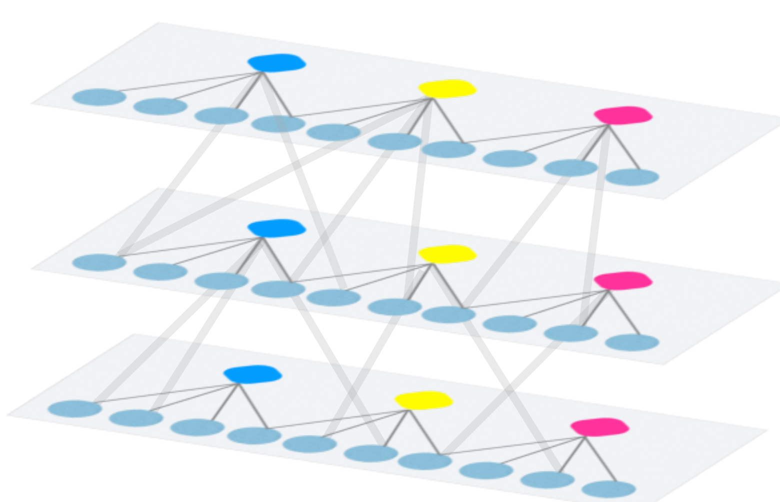
Biological Appeals

- Scelrfondiofromdhecogitivelevels:

  'the_drifles" (cdorlodt?)

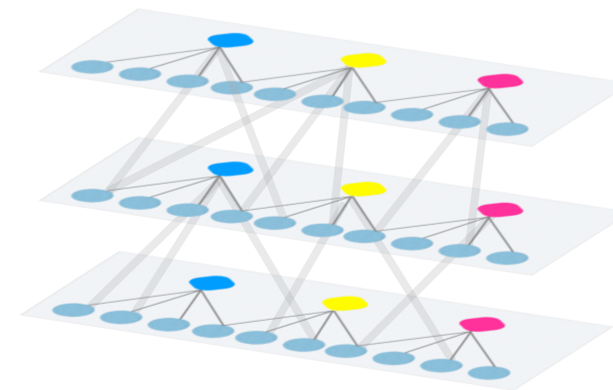
- Sinilar'spdidconectios" in mandiorbrain.

  Modha et al., *Cognitive computing*, Communications of the ACM, 2011.



Performance Analysis

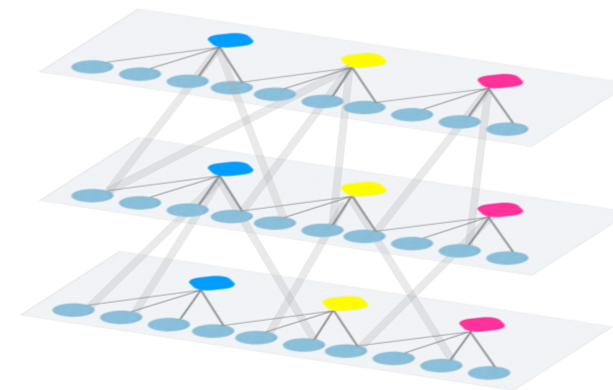
Performance Analysis



- Technical tools borrowed from [2]

[2] *A simple proof of threshold saturation for coupled scalar recursions*
A. Yedla, Y. Jian, P.S. Nguyen, H.D. Pfister, ISTC 2012.

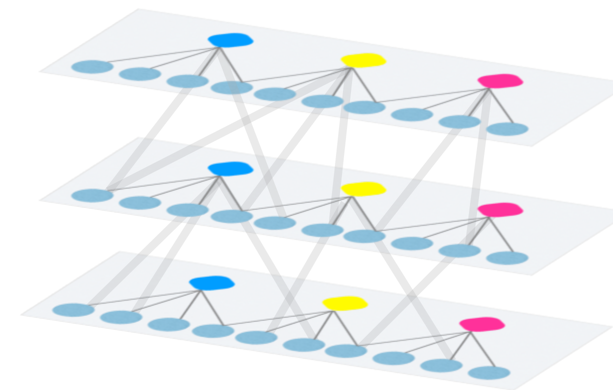
Performance Analysis



- Technical tools borrowed from [2]
- Define
  - p_e: "channel" error probability
  - $z(t)$: average probability of error in iteration t
  - $p_e^\dagger$: maximum p_e for which the uncoupled system is successful

[2] *A simple proof of threshold saturation for coupled scalar recursions*
A. Yedla, Y. Jian, P.S. Nguyen, H.D. Pfister, ISTC 2012.

Performance Analysis

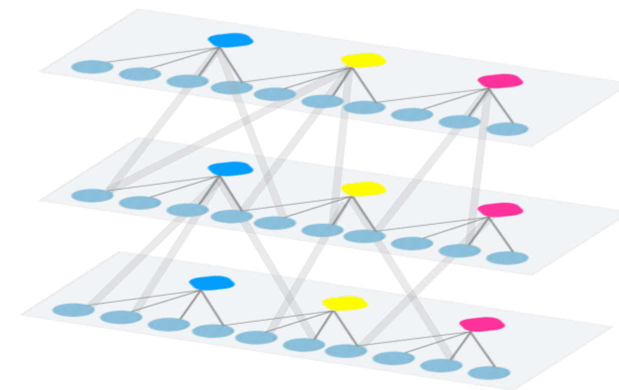


- Technical tools borrowed from [2]
- Define
  - p_e: "channel" error probability
  - $z(t)$: average probability of error in iteration t
  - $p_e^\dagger$: maximum p_e for which the *uncoupled* system is successful
- We define the potential $U(z; p_e)$ that has the property

$$U'(z; p_e) > 0 \text{ for } p_e < p_e^\dagger$$

[2] *A simple proof of threshold saturation for coupled scalar recursions*
A. Yedla, Y. Jian, P.S. Nguyen, H.D. Pfister, ISTC 2012.

Performance Analysis



- Technical tools borrowed from [2]
- Denote
  - p_e: "channel" error probability
  - $z(t)$: average probability of error in iteration t
  - $p_e^{\dagger}$: maximum p_e for which the *uncoupled* system is successful
- We define the potential $U(z; p_e)$ that has the property

$$U'(z; p_e) > 0 \text{ for } p_e < p_e^{\dagger}$$

- Define $p_e^{\dagger} < p_e^{*}$ to be the maximum p_e for which

$$\min_z U(z; p_e) > 0$$

[2] *A simple proof of threshold saturation for coupled scalar recursions*
A. Yedla, Y. Jian, P.S. Nguyen, H.D. Pfister, ISTC 2012.

Results

Error Correction Performance

Error Correction Performance

- [illegible] Ω [illegible] $p_e < p_e^*$.

Error Correction Performance

- [illegible] Ω [illegible] $p_e < p_e^*$.
    - Note that since $p_e^\dagger < p_e^*$ this means that the coupled system outperforms the uncoupled system.

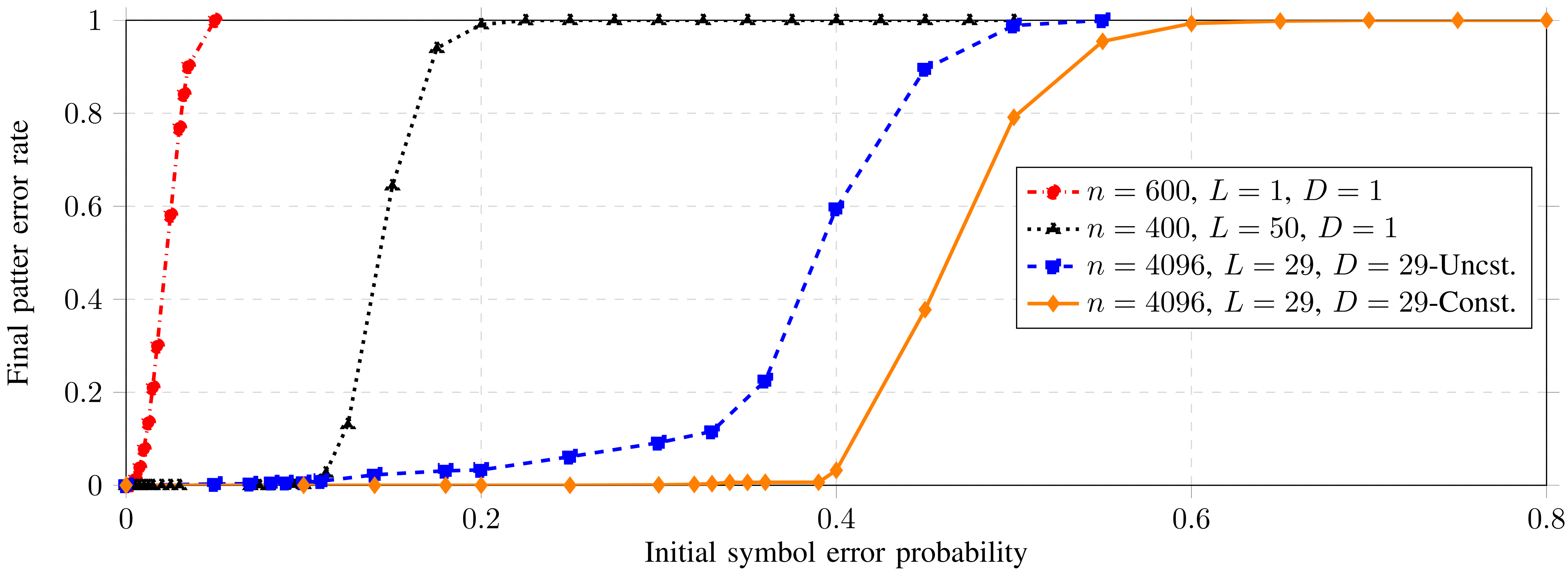
Error Correction Performance

- [illegible] Ω [illegible] $p_e < p_e^*$.
  - Note that since $p_e^\dagger < p_e^*$ this means that the coupled system outperforms the uncoupled system.
  - The lower bound for Ω provides a sufficient condition.

Simulations

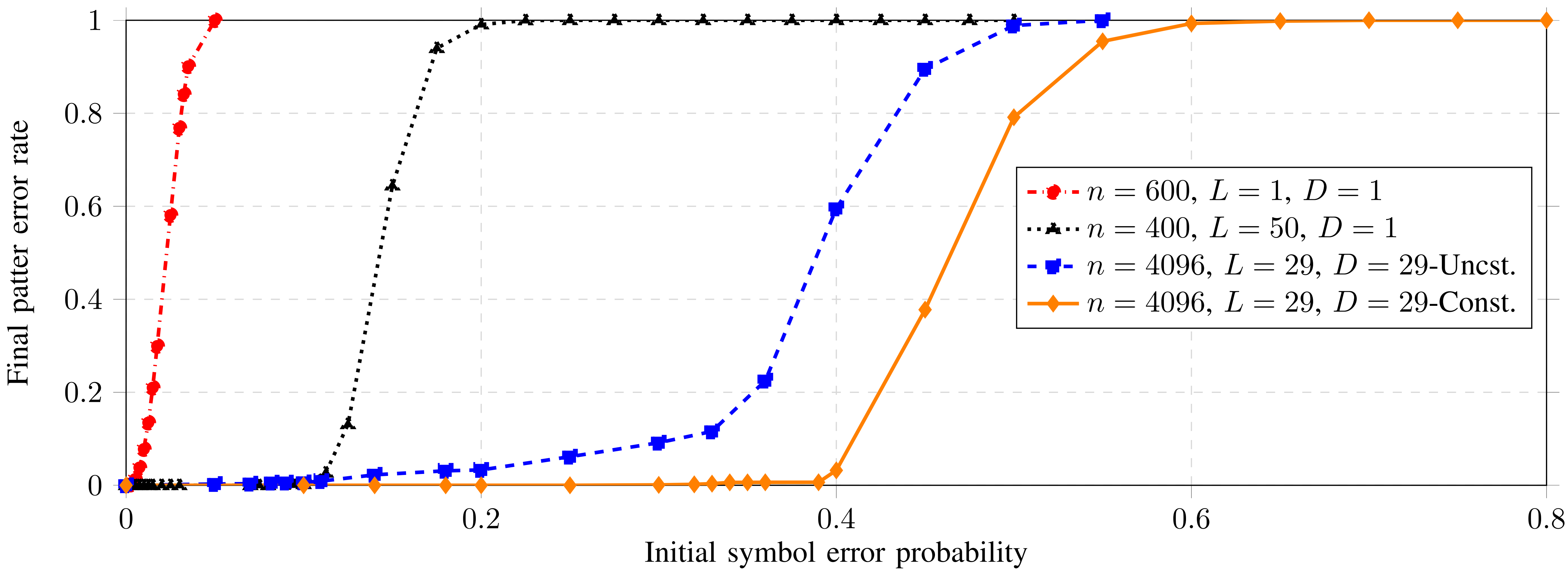
Simulations

- Pattern error probability vs initial errors



Simulations

- [illegible]



- [illegible]

	$p_e^\dagger$	p_e^*
$e = 1$	0.078	0.114
$e = 2$	0.197	0.394

Ongoing Work

Internal Noise Helps!

Internal Noise Helps!

- The neurons in our model were perfect, i.e. deterministic.

Internal Noise Helps!

- The neurons in our model were perfect, i.e. deterministic
- But real neurons are susceptible to *internal noise*
- So what happens if we introduce internal noise in our model?

Internal Noise Helps!

- The neurons in our model were perfect, i.e. deterministic.
- But real neurons are susceptible to *internal noise*.
- So what happens if we introduce internal noise in our model?

Counterintuitively, internal noise improves the recall performance!

the network achieves better thresholds in presence of internal noise.

Noise-Enhanced Associative Memories

A. Karbasi, A. H. Salavati, A. Shokrollahi, L. R. Varshney To appearn in NIPS 2013

Thank You!

Backup Slides

Pattern Retrieval Capacity

Pattern Retrieval Capacity

- **Theorem:** There exists a code X with C vectors of length n such that $C = a^k$, with $a \ge 2$, where $k = \mathrm{rank}(X) = O(n)$.