Visual Spike Coding Using a Statistically Optimized Overcomplete Representation

Laurent Perrinet and Manuel Samuelides

`perrinet@cert.fr, samuelid@supaero.fr`

ONERA/DTIM, 2, av. Belin, 31055 Toulouse, France

Outline

Outline

1. Motivations:
   (a) Visual Spike Coding
   (b) Overcomplete representation of the image

Outline

1. Motivations:
   (a) Visual Spike Coding
   (b) Overcomplete representation of the image
2. Matching Pursuit

Outline

1. Motivations:
   (a) Visual Spike Coding
   (b) Overcomplete representation of the image
2. Matching Pursuit
3. Optimization through learning
   (a) Quantization
   (b) Whitening

Outline

1. Motivations:
   (a) Visual Spike Coding
   (b) Overcomplete representation of the image
2. Matching Pursuit
3. Optimization through learning
   (a) Quantization
   (b) Whitening

- Conclusion

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system
  - Based on neurophysiology of the retina
  - Using Spiking Neural Networks [Gerstner et al., 1999]

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system
  - $\Rightarrow$ Multiscale edge detection: Wavelet-like model [Van Rullen and Thorpe, 2001]

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system
  - $\Rightarrow$ Multiscale edge detection: Wavelet-like model [Van Rullen and Thorpe, 2001]
  - But: Needs a constrained basis of filters

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system
  - $\Rightarrow$ Multiscale edge detection: Wavelet-like model [Van Rullen and Thorpe, 2001]
  - But: Needs a constrained basis of filters
- Find a sparse and flexible representation

Motivation

- Build a Neuromorphic model of the neural coding in the primary visual system
  - $\Rightarrow$ Multiscale edge detection: Wavelet-like model [Van Rullen and Thorpe, 2001]
  - But: Needs a constrained basis of filters
- Find a sparse and flexible representation
  - Overcomplete representation of the image as a linear combination of some filters (set $\mathcal{D}_c$) from a large dictionary $\mathcal{D}$

$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

Overcomplete representation

$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

Overcomplete representation

$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

- Pros:
  - Potentially, larger dictionary = sparser code [Olshausen and Field, 1996]

Overcomplete representation

$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

- Pros:
  - Potentially, larger dictionary = sparser code
  - Similar image structure code with dictionary-compatible transforms

Overcomplete representation

$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

- Pros:
  - Potentially, larger dictionary = sparser code
  - Similar image structure code with dictionary-compatible transforms
- Cons:
  - Overlapping Basis: needs a specific algorithm to produce a sparse code

Overcomplete representation

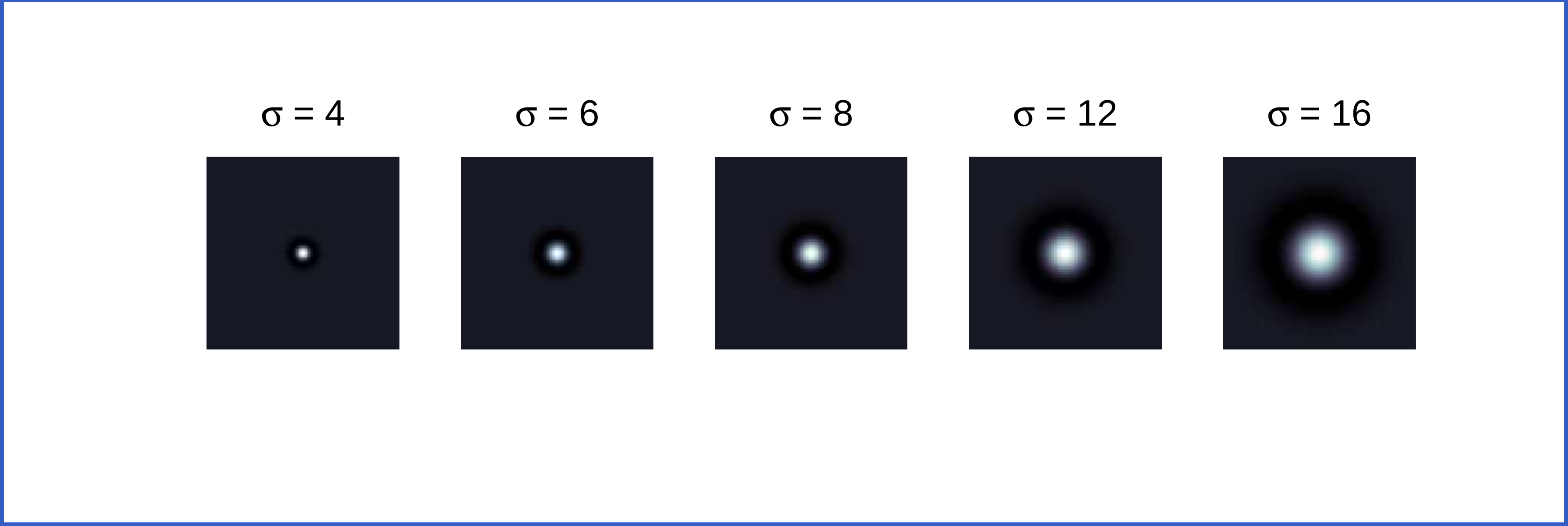
$$I = \sum_{i \in \mathcal{D}_c} a_i w_i$$

- Pros:
  - Potentially, larger dictionary = sparser code
  - Similar image structure code with dictionary-compatible transforms
- Cons:
  - Overlapping Basis: needs a specific algorithm to produce a sparse code
  - Optimal basis search: NP hard

Method: Greedy projection pursuit

Method: Greedy projection pursuit

- Define a dictionary of scaled and translated DOGs : $\mathcal{D} = \{w_i\}$



We define the norm $N^i = \|w_i\|$ and initialize the image $f^0 = f$

Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 =< f^0, w_i >$

Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 =< f^0, w_i >$
- Greedy projection pursuit

$$\begin{aligned} i^0 &= \text{ArgMax}_i(|C_i^0|) \\ f^1 &= f^0 - \frac{< f^0, w_{i^0} > w_{i^0}}{\|w_{i^0}\|^2} \\ &= f^0 - C_{i^0}^0 \frac{w_{i^0}}{{N_{i^0}}^2} \end{aligned}$$

Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 =< f^0, w_i >$
- Greedy projection pursuit

$$
\begin{aligned}
i^0 &= \text{ArgMax}_i(|C_i^0|) \\
C_i^1 &= C_i^0 - C_{i^0}^0 \frac{< w_{i^0}, w_i >}{{N_{i^0}}^2}
\end{aligned}
$$

Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 =< f^0, w_i >$
- Greedy projection pursuit

$$
\begin{aligned}
i^t &= \mathrm{ArgMax}_i(|C_i^t|) \\
C_i^{t+1} &= C_i^t - C_{i^t}^t \frac{< w_{i^t}, w_i >}{{N_{i^t}}^2}
\end{aligned}
$$

Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 = < f^0, w_i >$
- Greedy projection pursuit

$$
\begin{aligned}
i^t &= \text{ArgMax}_i(|C_i^t|) \\
C_i^{t+1} &= C_i^t - C_{i^t}^t \frac{< w_{i^t}, w_i >}{{N_{i^t}}^2}
\end{aligned}
$$

  - Matching Pursuit (MP) [Mallat and Zhang, 1993] for $N_i = 1$.

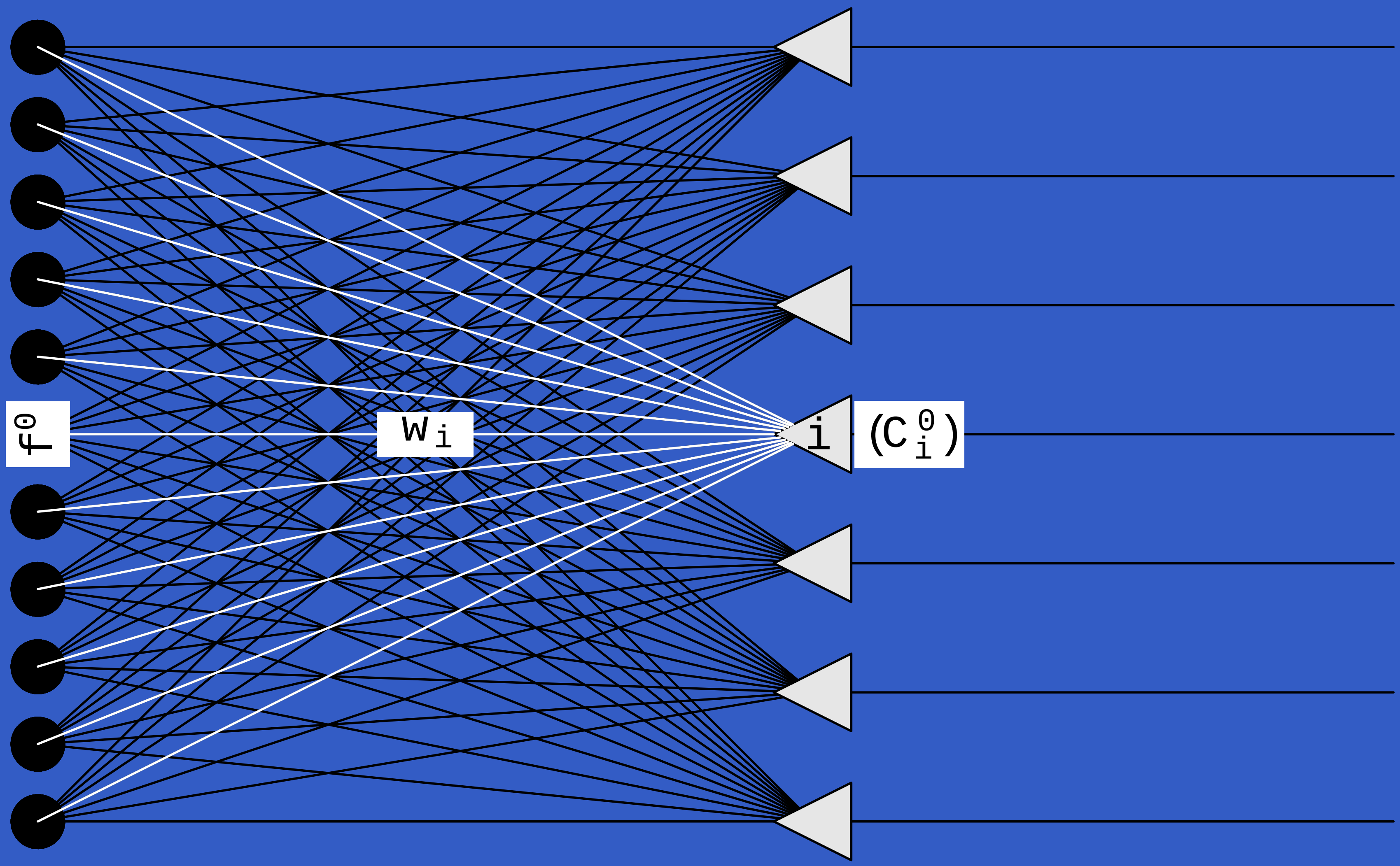
Method: Greedy projection pursuit

- Define a dictionary : $\mathcal{D} = \{w_i\}$
- Compute inner products once $C_i^0 =< f^0, w_i >$
- Greedy projection pursuit

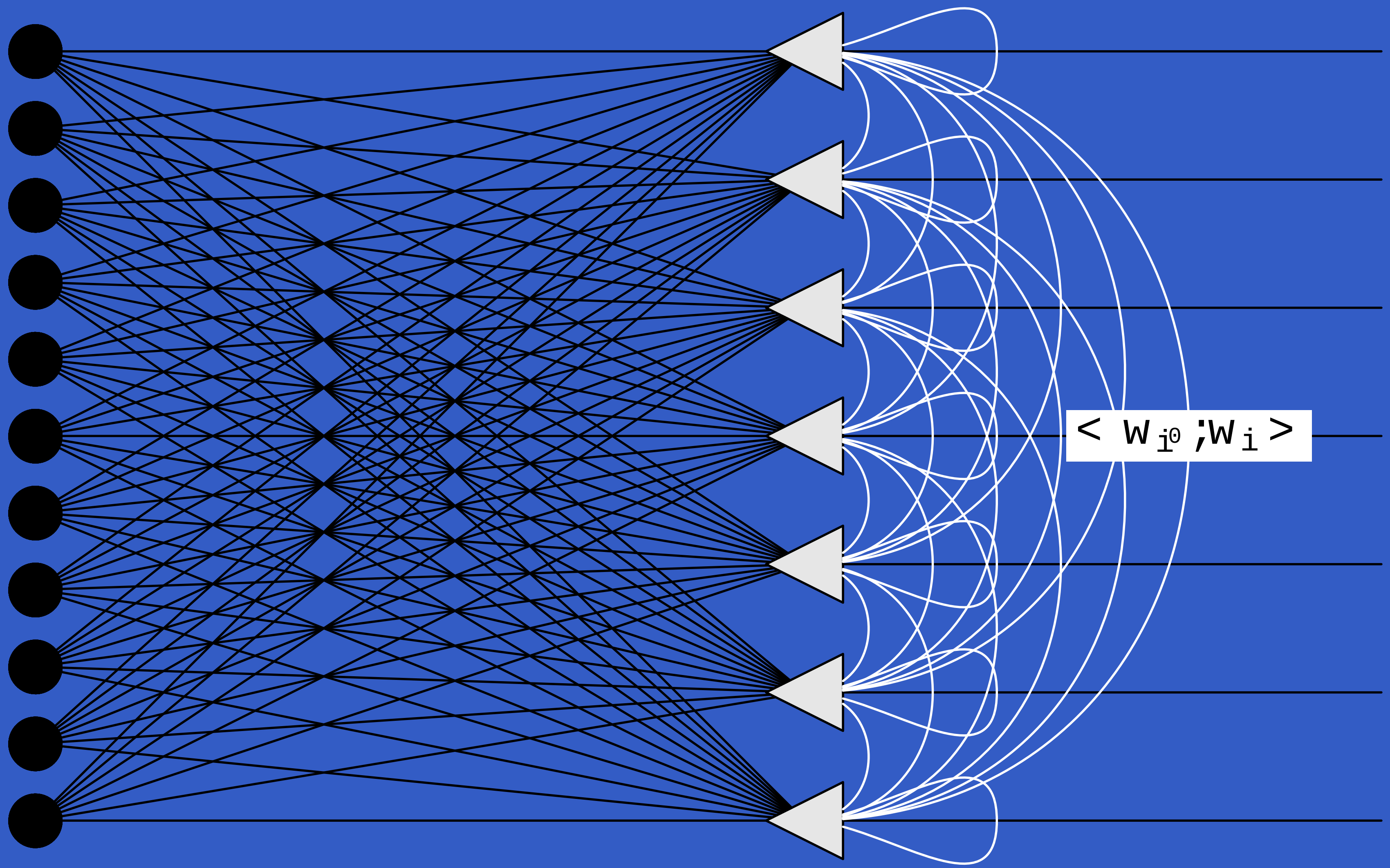
$$
\begin{aligned}
i^t &= \text{ArgMax}_i(|C_i^t|) \\
C_i^{t+1} &= C_i^t - C_{i^t}^t \frac{< w_{i^t}, w_i >}{{N_{i^t}}^2}
\end{aligned}
$$

  - Matching Pursuit (MP) [Mallat and Zhang, 1993] for $N_i = 1$.
  - Reconstruction: $\Rightarrow f = \sum_{t=1}^{T} C_{i^t}^t \frac{w_{i^t}}{{N_{i^t}}^2} + f^T$

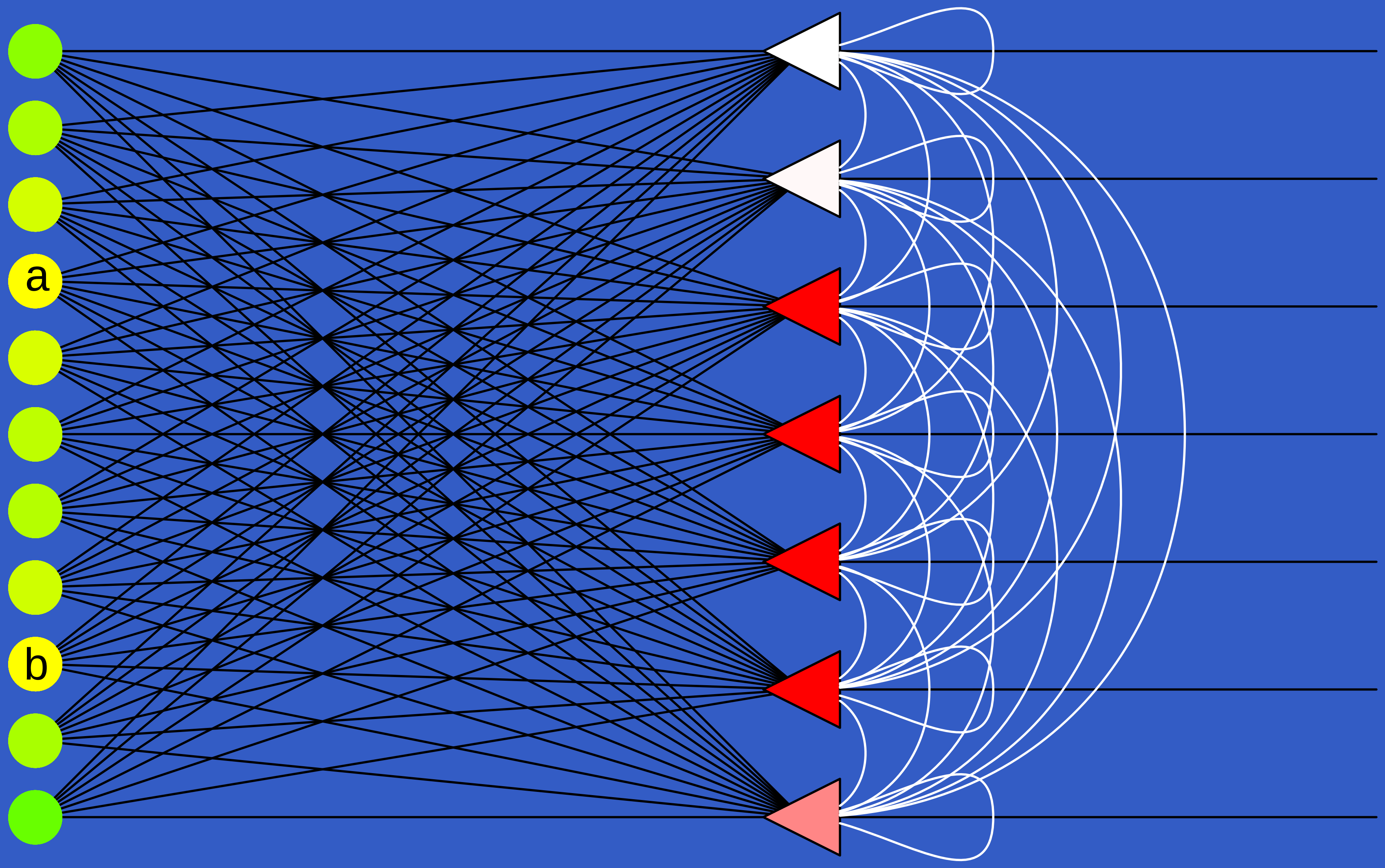
Spiking neural networks and MP



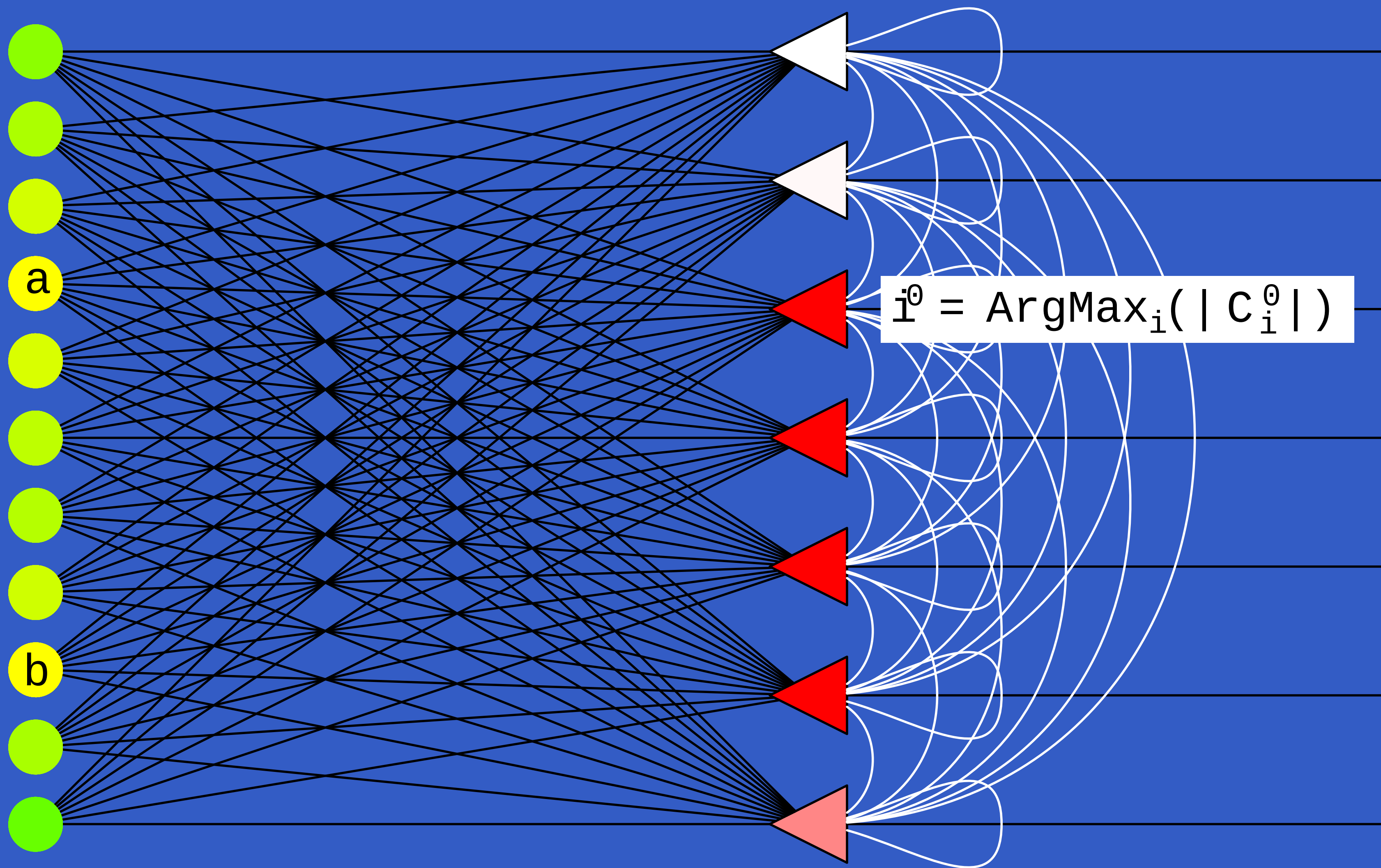
Spiking neural networks and MP



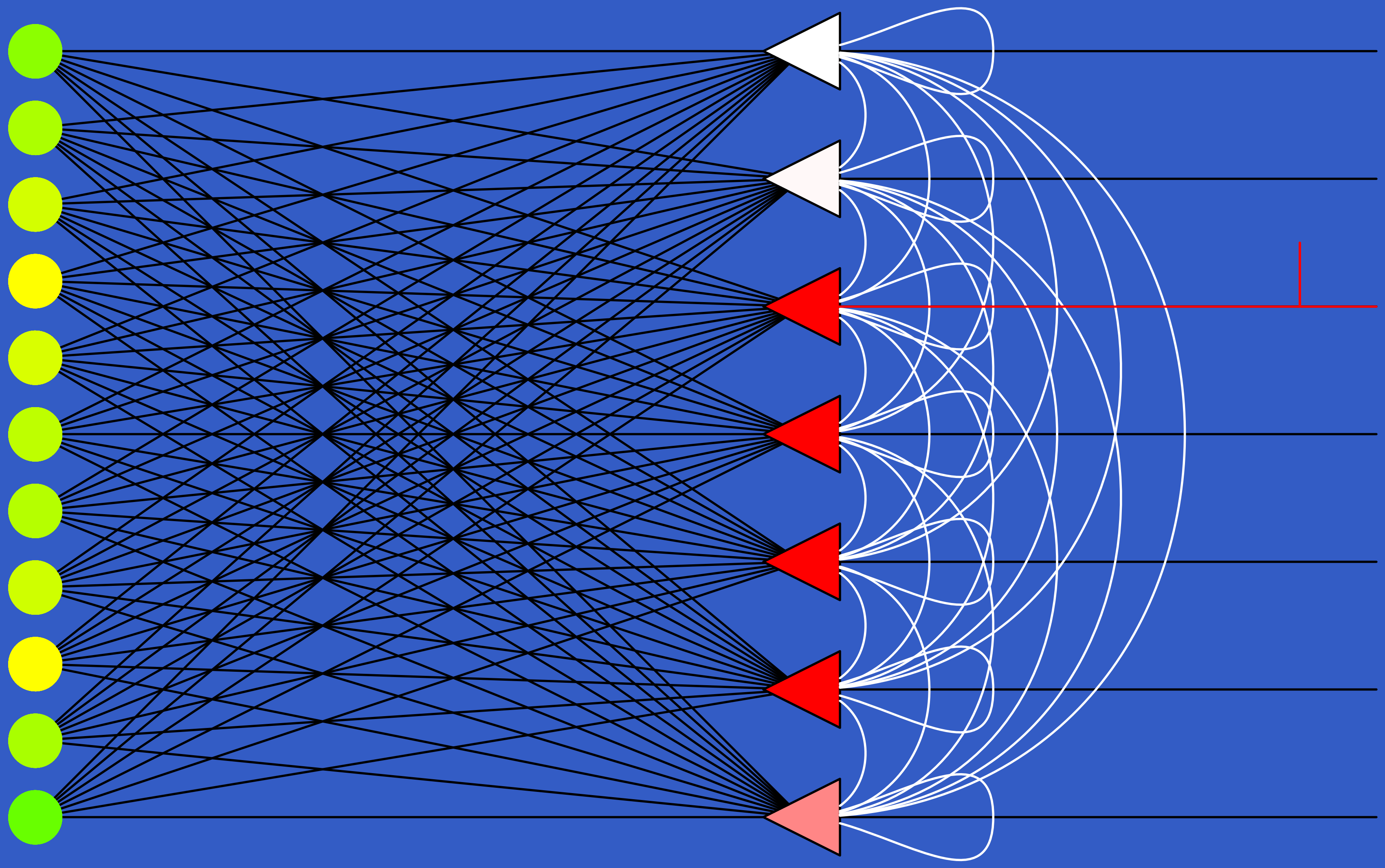
Spiking neural networks and MP



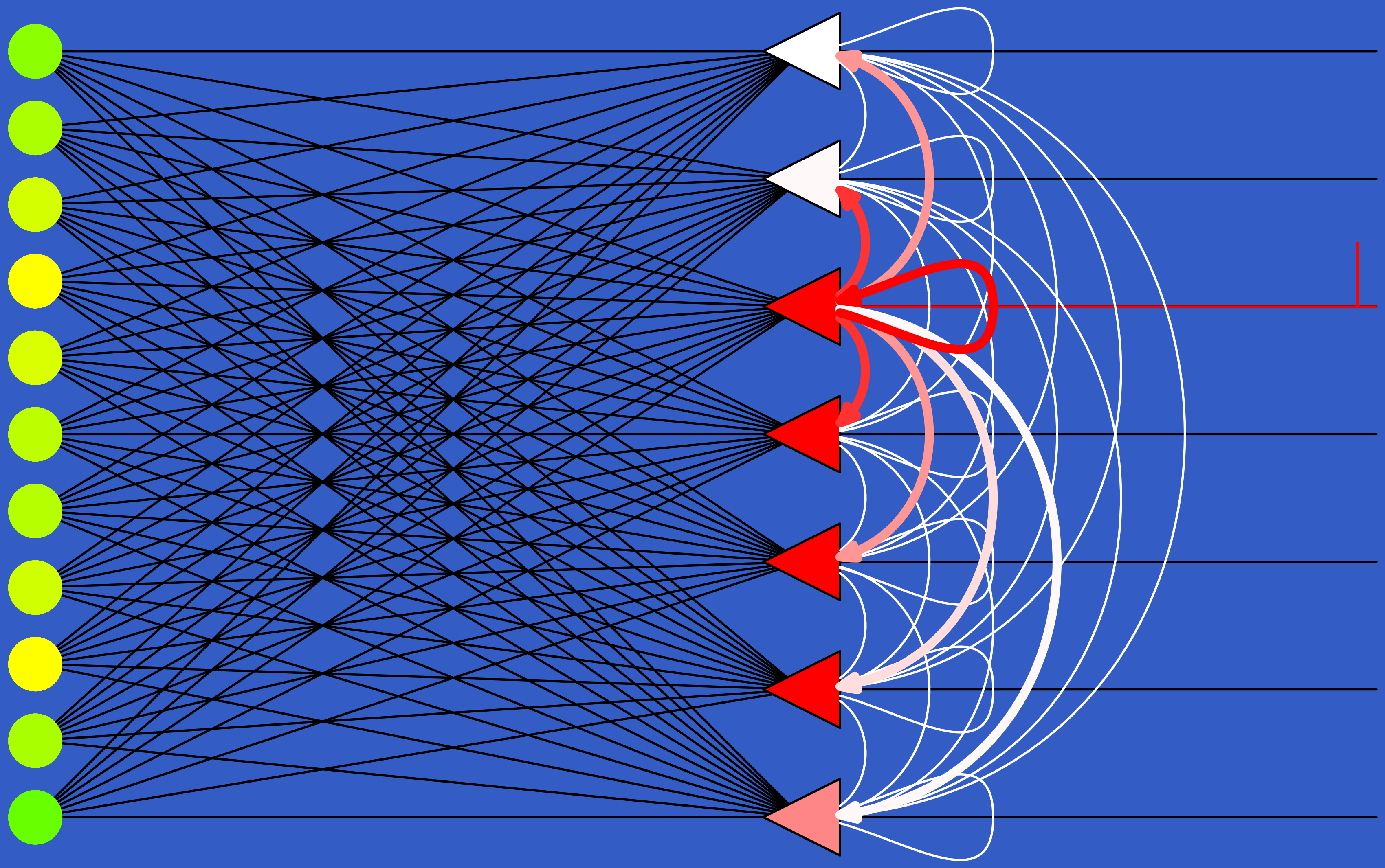
Spiking neural networks and MP



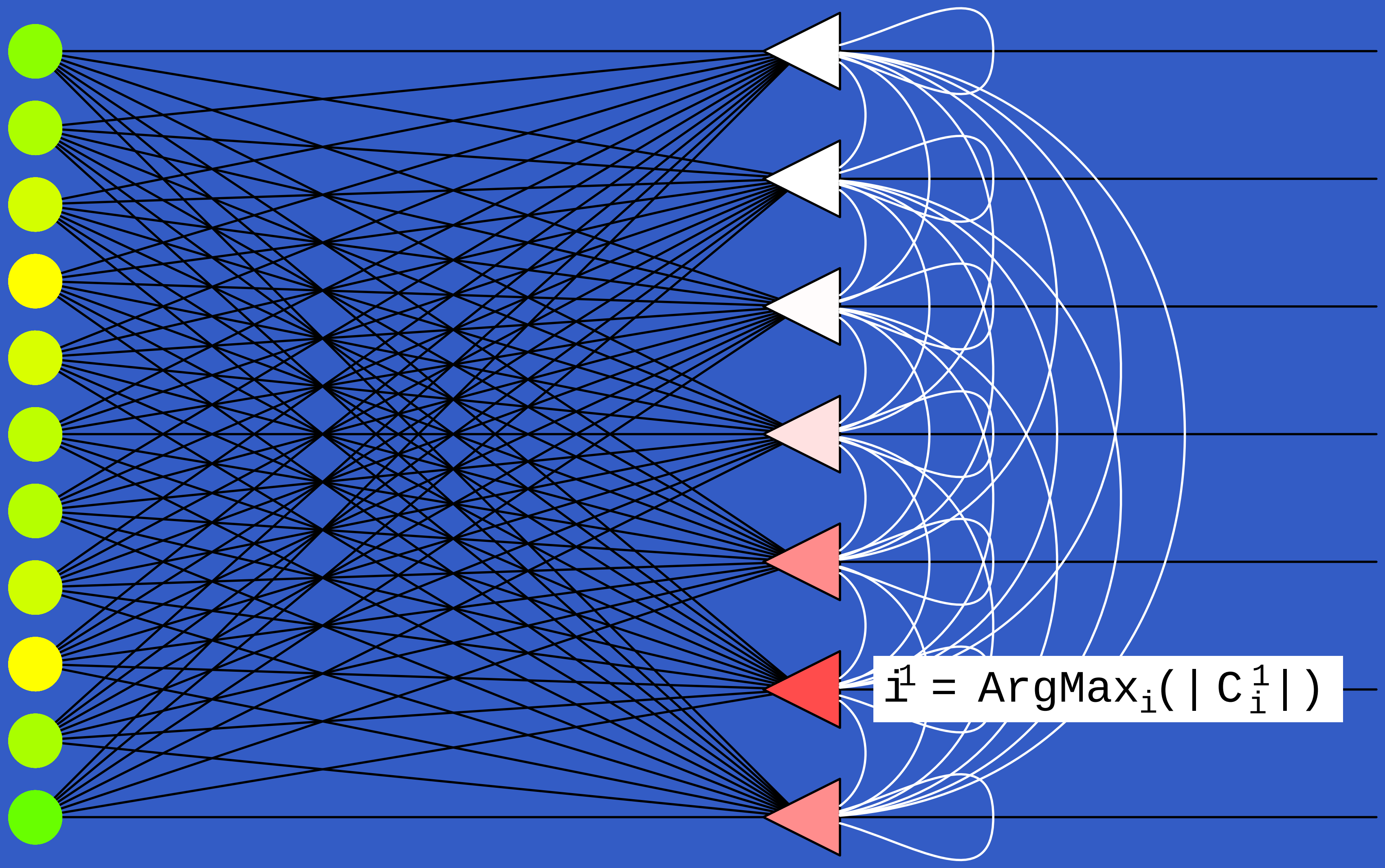
Spiking neural networks and MP



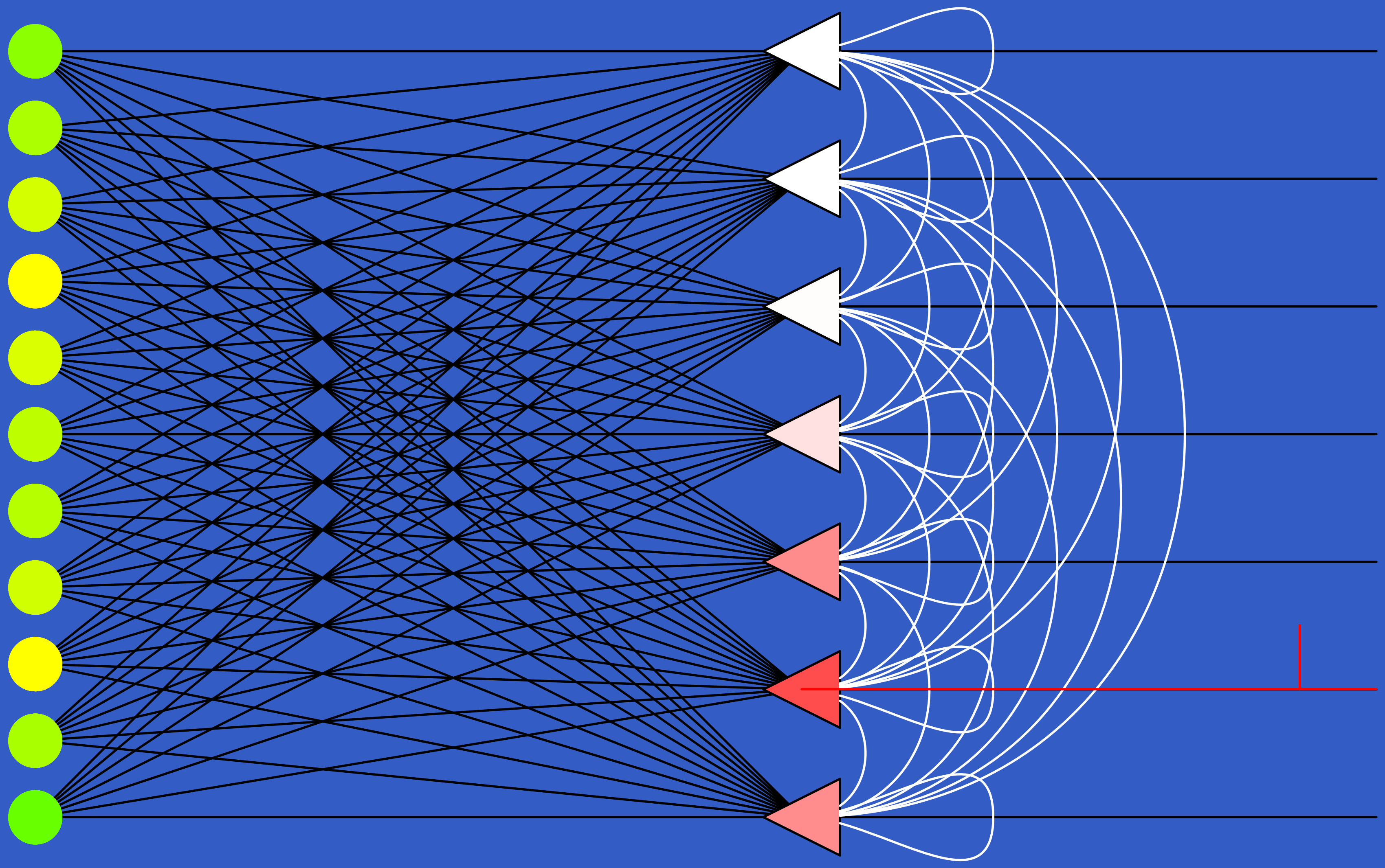
Spiking neural networks and MP



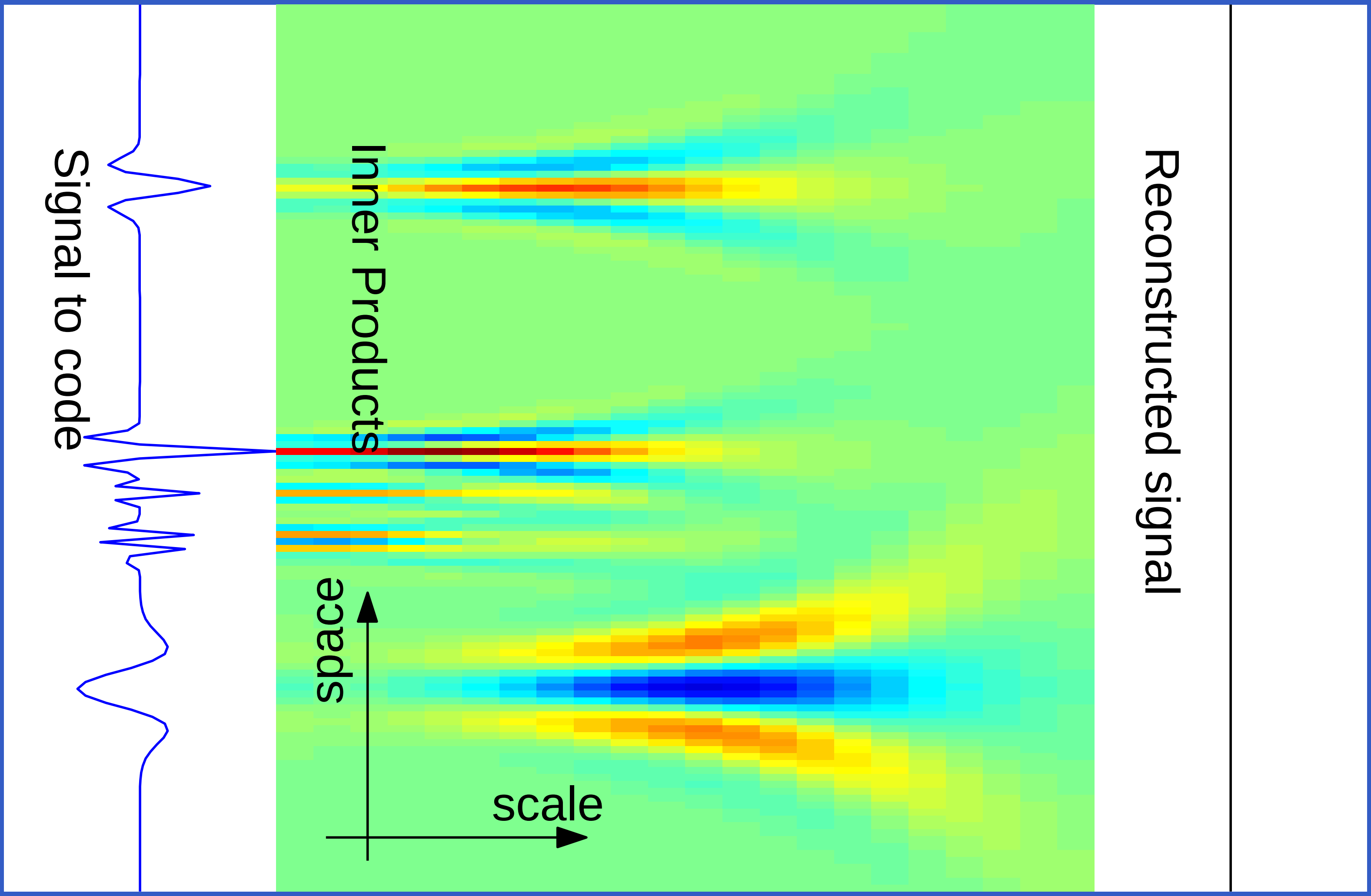
Spiking neural networks and MP



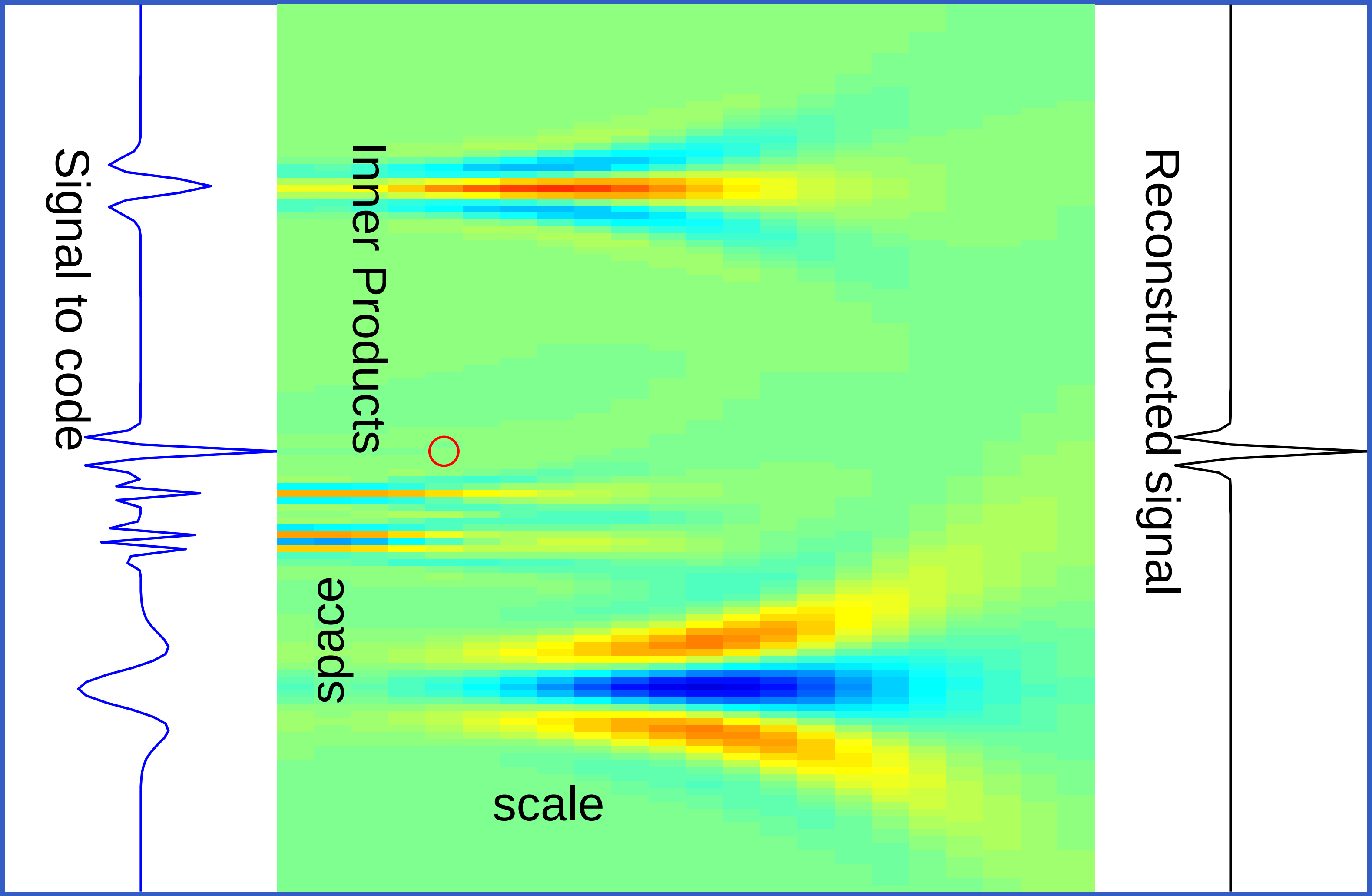
Spiking neural networks and MP



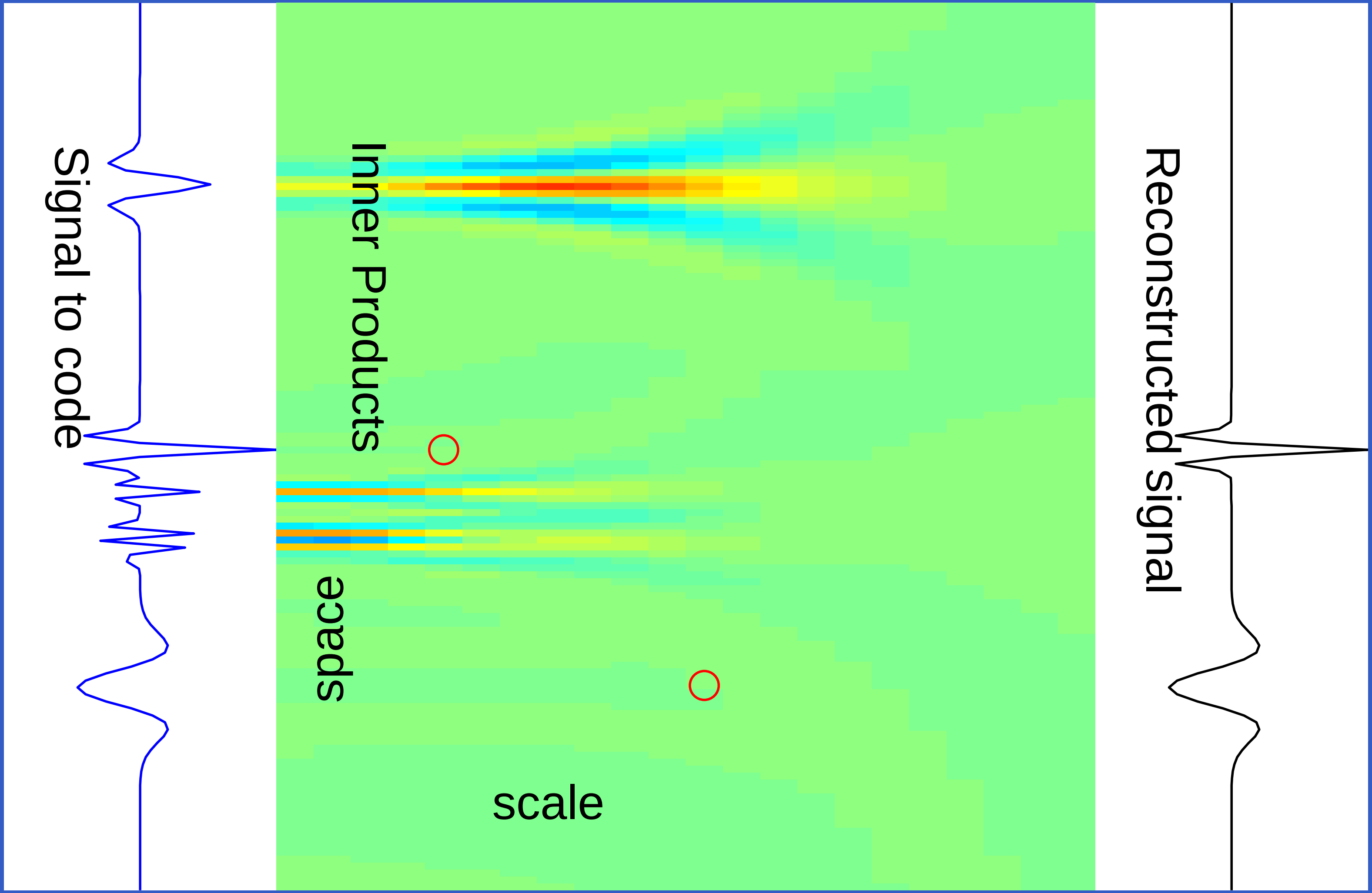
MP explained in plots



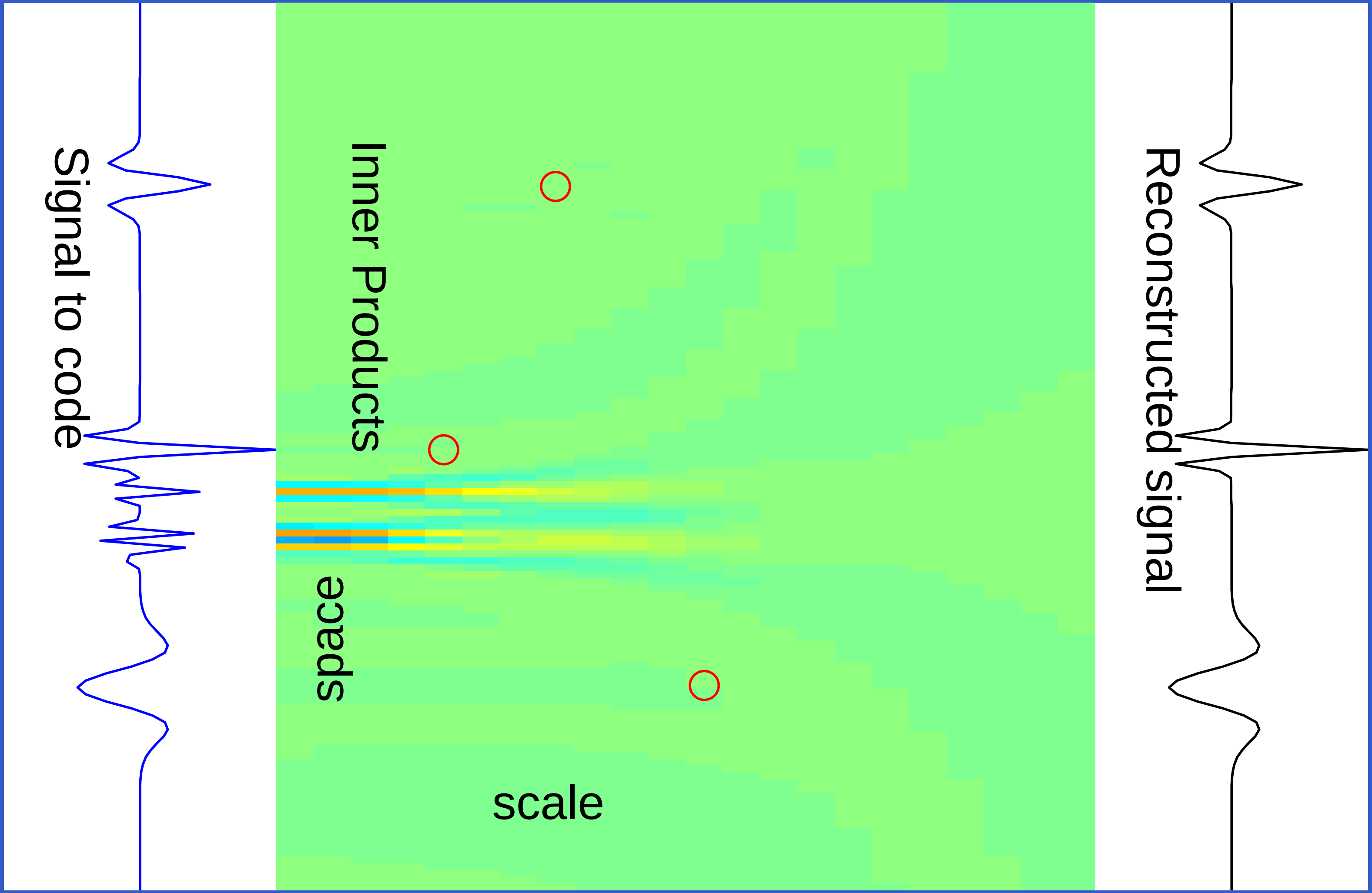
MP explained in plots



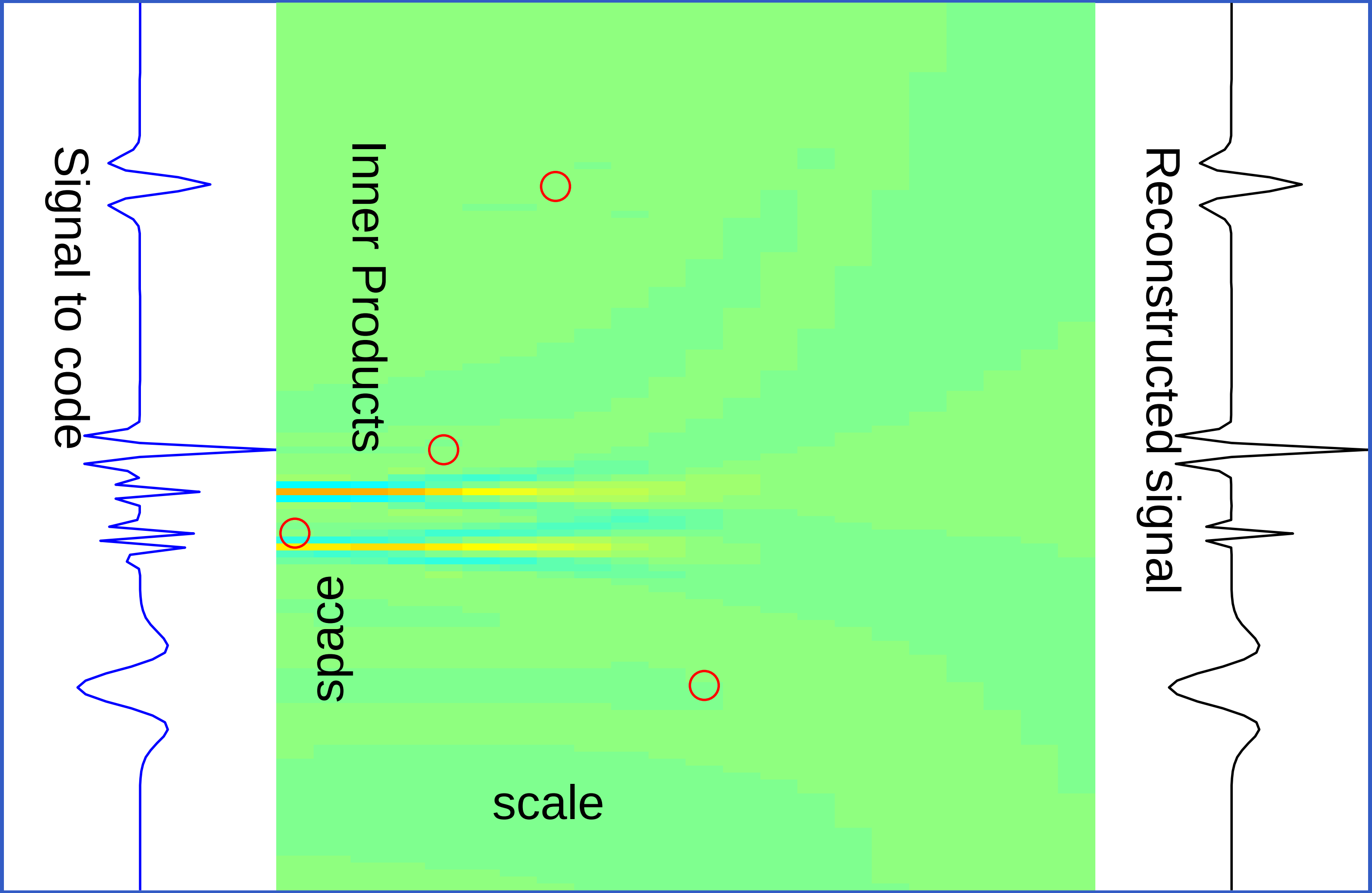
MP explained in plots



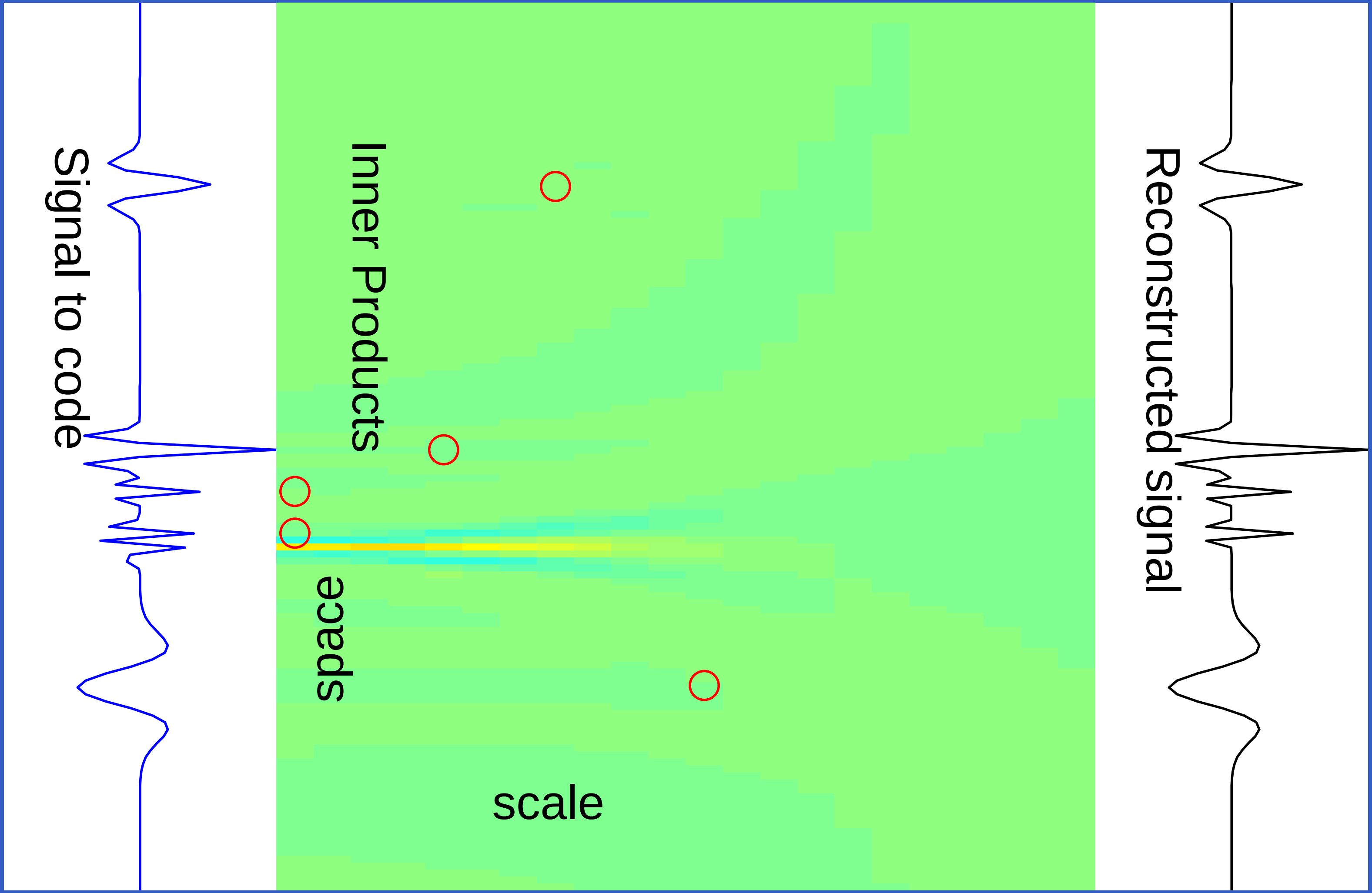
MP explained in plots



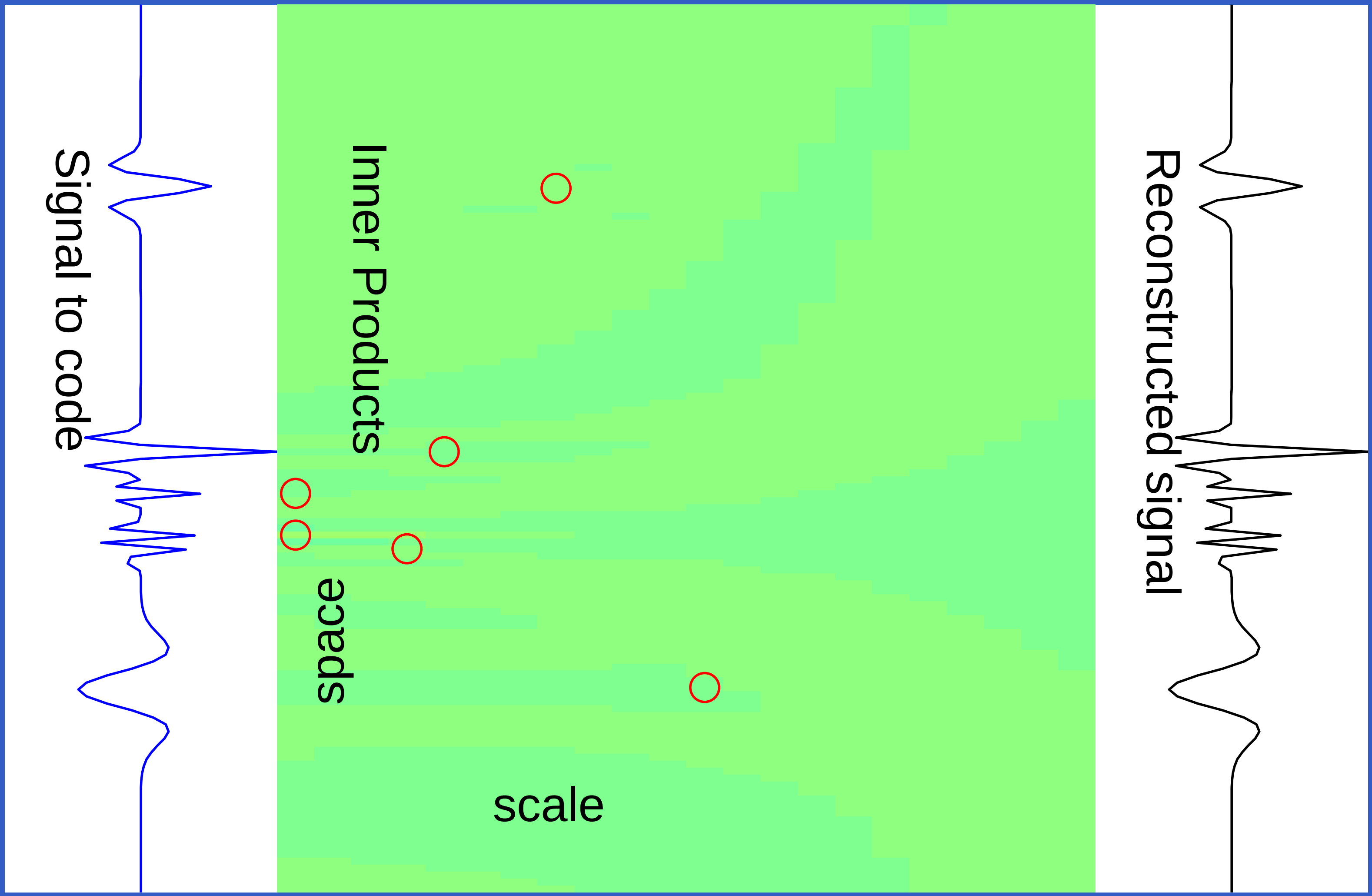
MP explained in plots



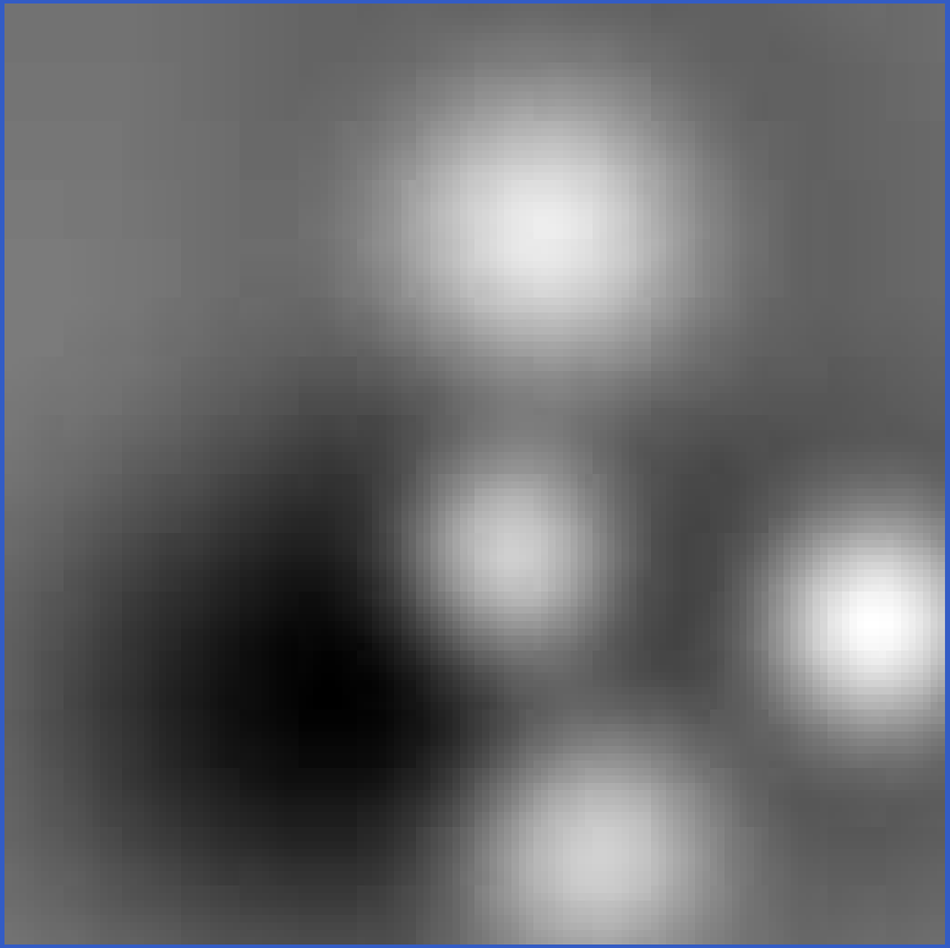
MP explained in plots



MP explained in plots

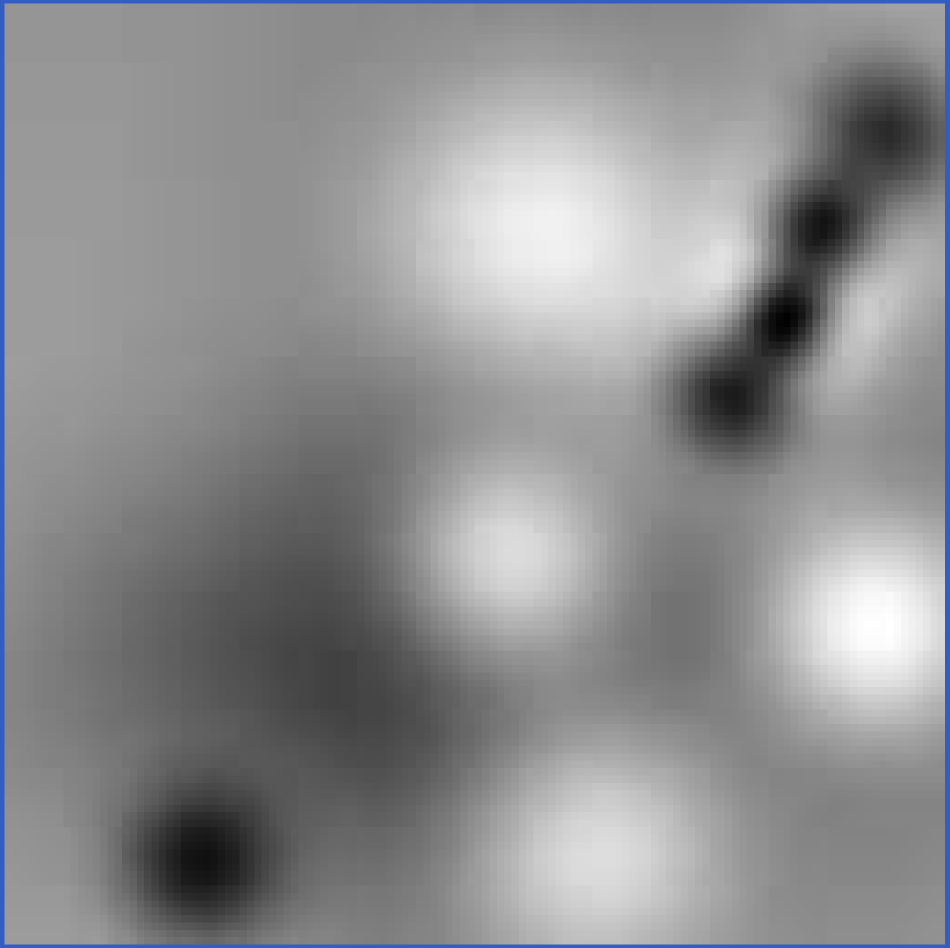


Results : image reconstruction (MP)



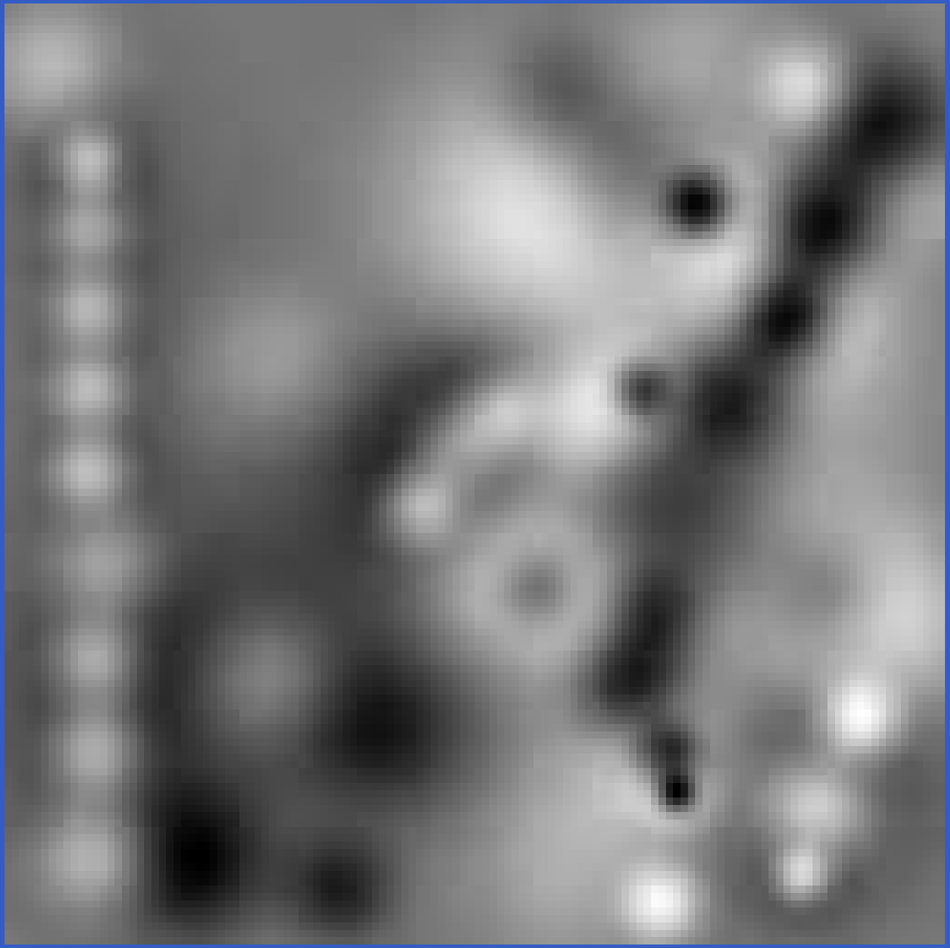
with 5 spikes

Results : image reconstruction (MP)



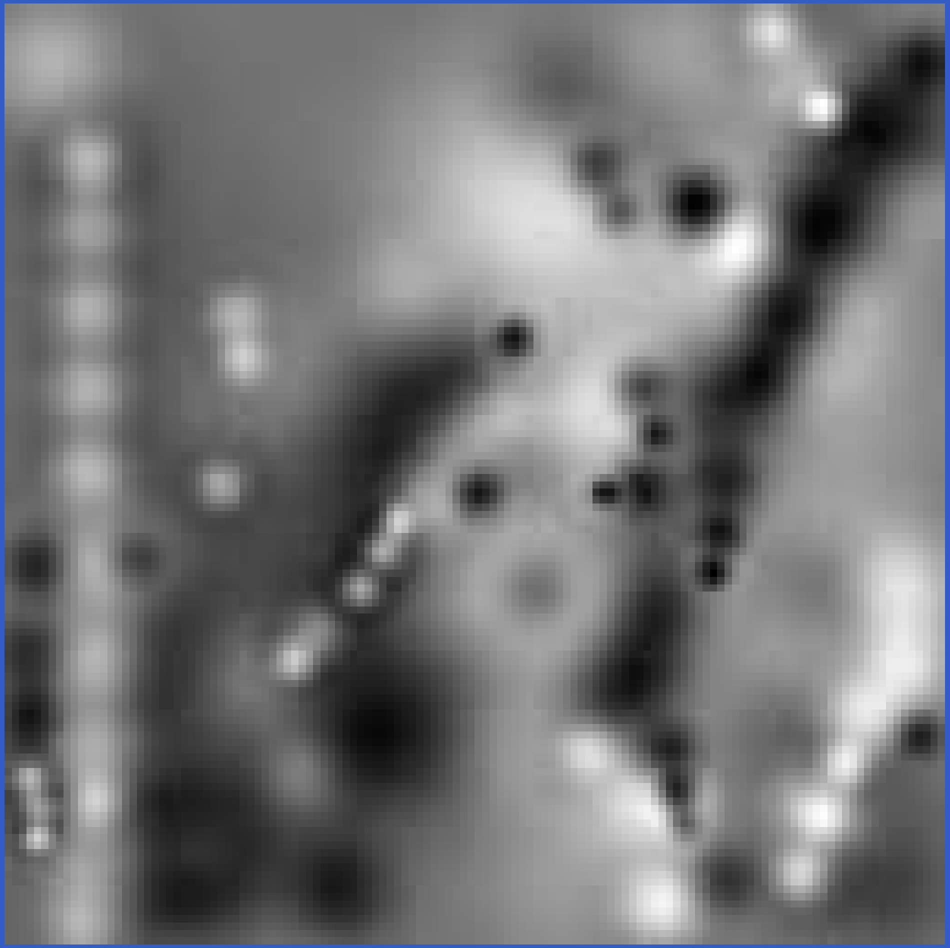
with 10 spikes

Results : image reconstruction (MP)



with 50 spikes

Results : image reconstruction (MP)



with 100 spikes

Results : image reconstruction (MP)

with 500 spikes

Results : image reconstruction (MP)

with 1000 spikes

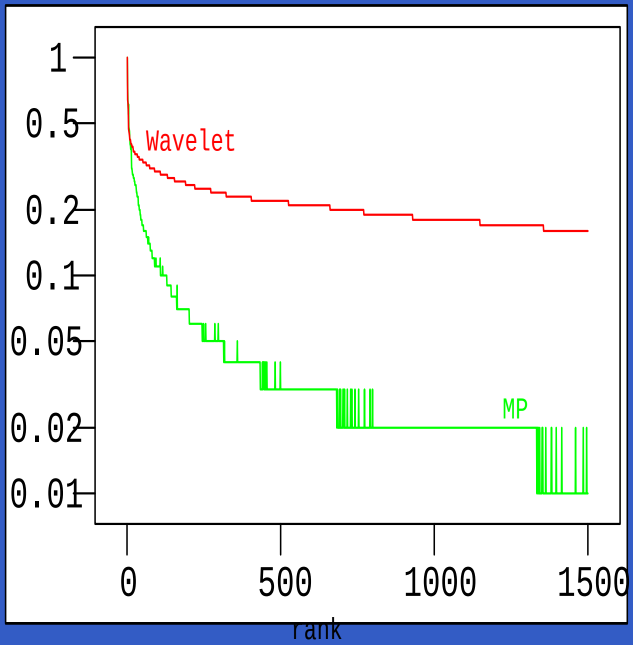
Matching Pursuit

Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.

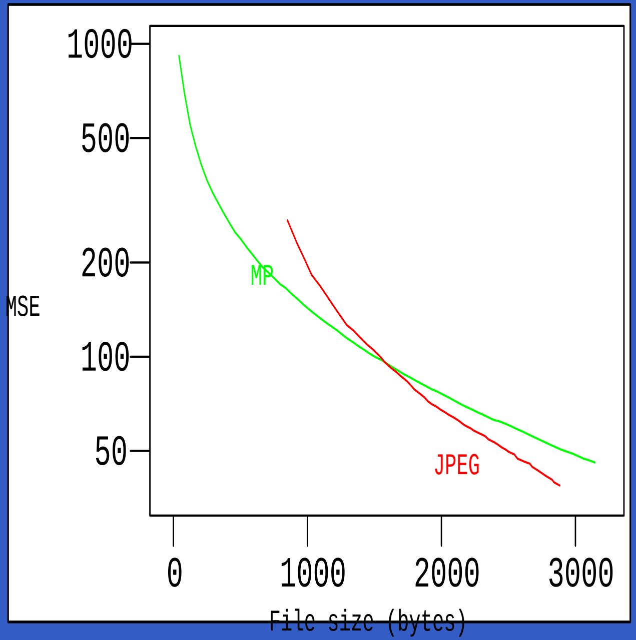
Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.
- Experimentally: Sparse Coding



Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.
- Experimentally: Mean Square Error



Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.
- But:
  - Compression performance just outperforms JPEG at low bit rates,

Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.
- But:
  - Compression performance just outperforms JPEG at low bit rates,
  - suboptimal (compared to Basis Pursuit),

Matching Pursuit

- Mallat and Zhang [1993] proved convergence of Matching Pursuit.
- But:
  - Compression performance just outperforms JPEG at low bit rates,
  - suboptimal (compared to Basis Pursuit),
  - may reveal non-features.

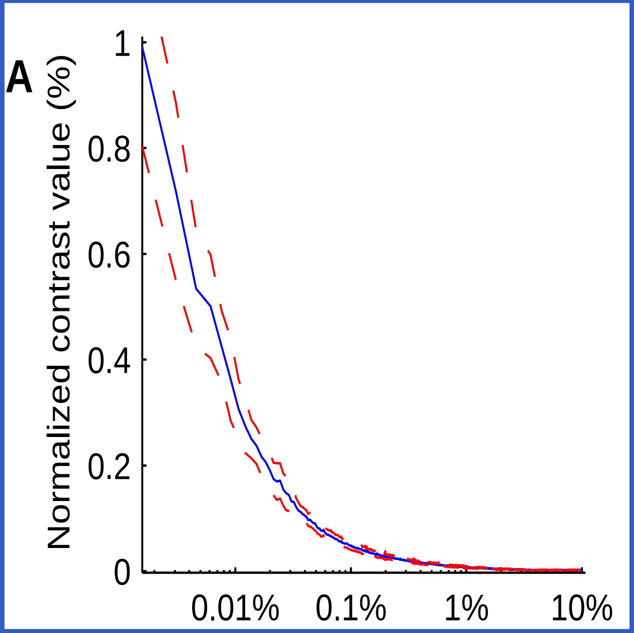
Learning with MP (1): quantization

Learning with MP (1): quantization

- Regularity across natural images of the coefficients' absolute value with their rank

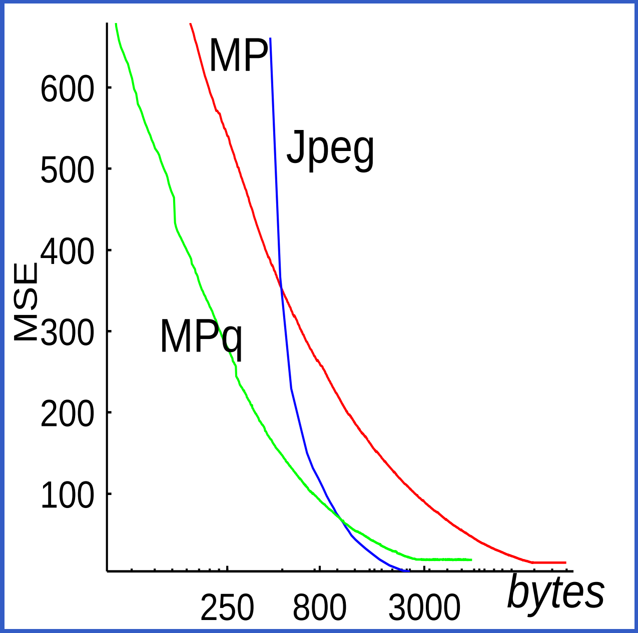
Learning with MP (1): quantization

- Regularity across natural images of the coefficients' absolute value with their rank
- Experimentally: Quantization (MPq)



Learning with MP (1): quantization

- Regularity across natural images of the coefficients' absolute value with their rank
- Experimentally: Mean Square Error



Learning with MP (1): quantization

- Regularity across natural images of the coefficients' absolute value with their rank
- Advantages:
  - Decreases bit weight of the coefficient to 1 (ON or OFF),

Learning with MP (1): quantization

- Regularity across natural images of the coefficients' absolute value with their rank
- Advantages:
  - Decreases bit weight of the coefficient to 1 (ON or OFF),
  - Versus classical wavelet coefficient quantization, MP may adapt to the quantization error.

Learning with MP (2): Whitening

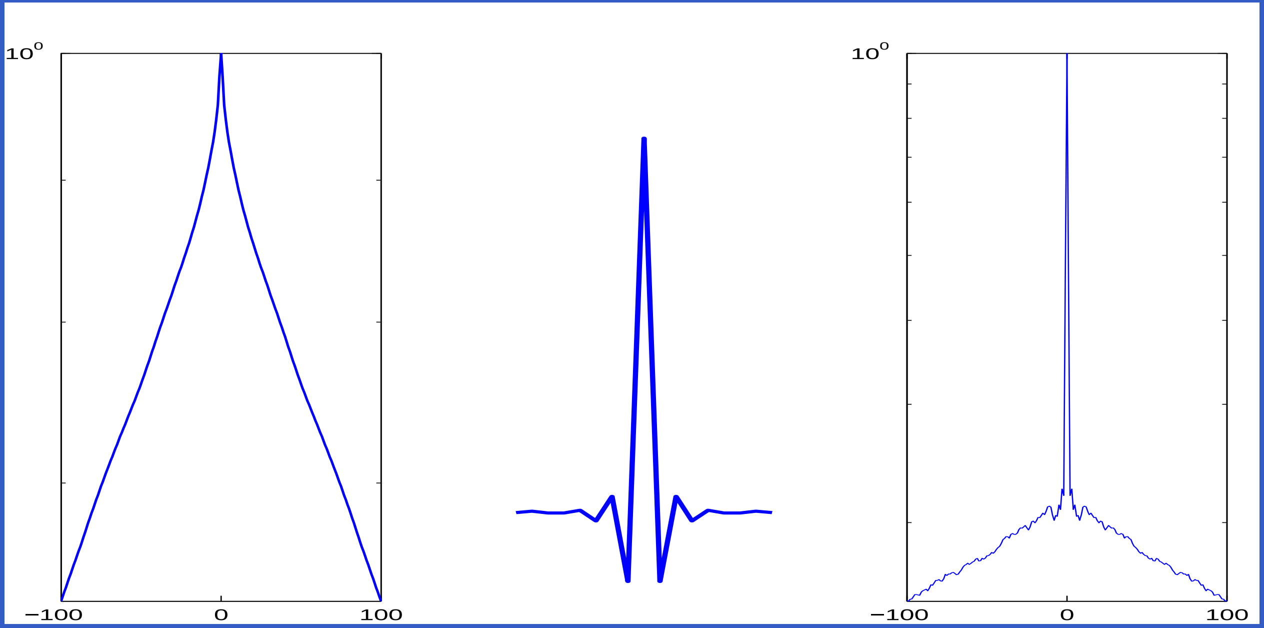
- MP: progressive and non-linear : importance of rank and of the competition between dictionary elements.

Learning with MP (2): Whitening

- MP: progressive and non-linear
- Statistics of natural images: decorrelating kernel K so that scales are whitened [Atick, 1992]

Learning with MP (2): Whitening

- MP: progressive and non-linear
- Statistics of natural images: decorrelating kernel K so that scales are whitened [Atick, 1992]

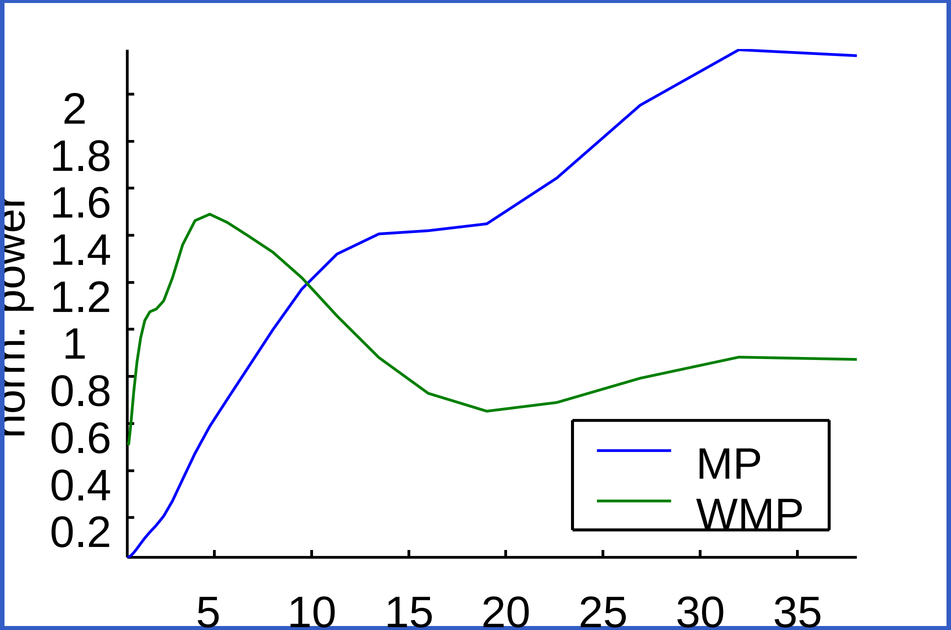


Learning with MP (2): Whitening

- MP: progressive and non-linear
- Statistics of natural images: decorrelating kernel K so that scales are whitened [Atick, 1992]
- Integrated in MP algorithm (Whitened MP, WMP):
  - $w_i^K = w_i * K$,
  - or use as a fast approximation $w_i^K = N_i . w_i$ with $N_i = \|\frac{w_i}{\|w_i\|} * K\|$,

Whitened MP : WMP

- The competition between scales is "fair": the same object at different scales is given the same importance. E.g., inner product power spectrum across scales for Lena:



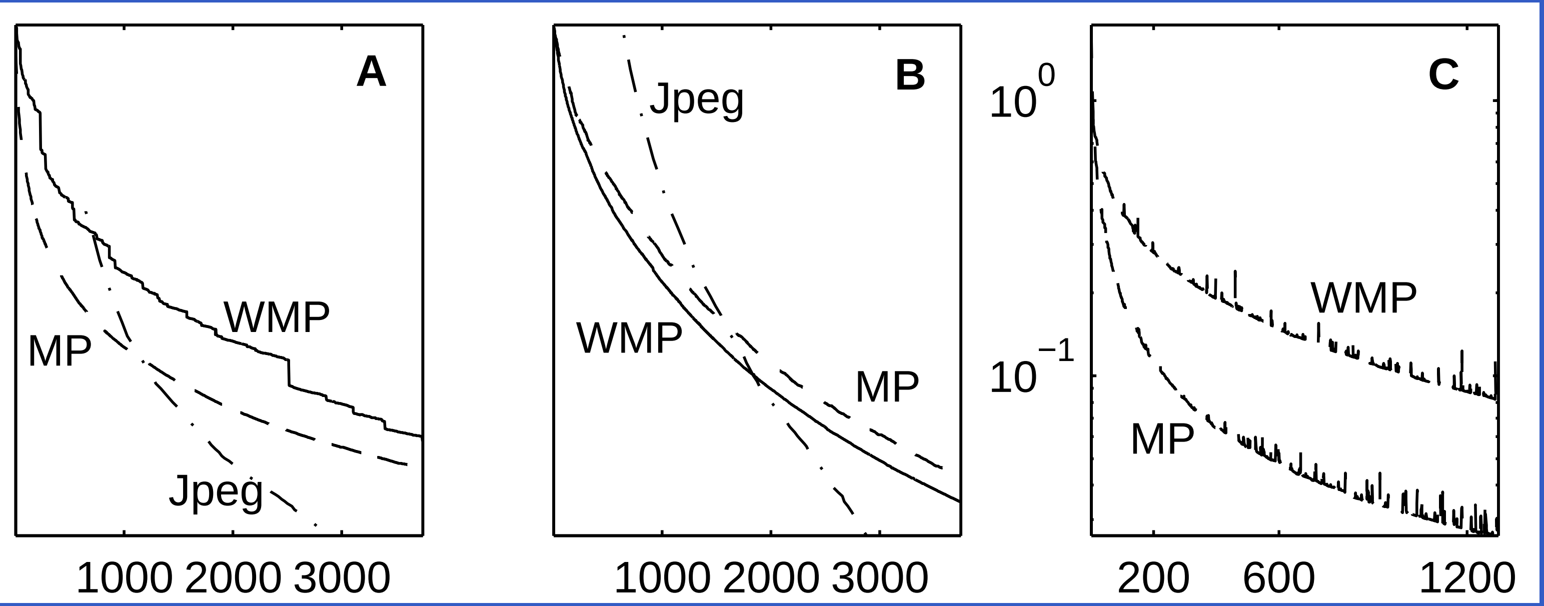
Whitened MP : WMP

- The competition between scales is "fair": the same object at different scales is given the same importance.
- Relates to the Mahalanobis distance,

Whitened MP : WMP

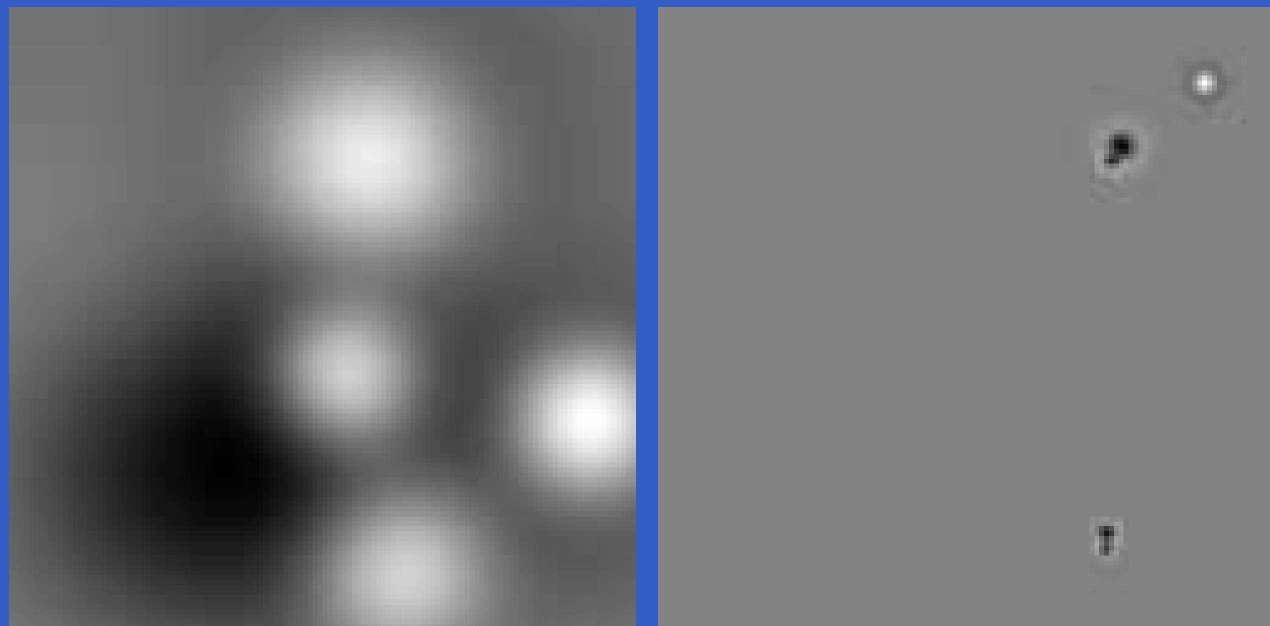
- The competition between scales is "fair": the same object at different scales is given the same importance.
- Relates to the Mahalanobis distance,
- Corresponds to a subjective distance between images.

Whitened MP: WMP(2)



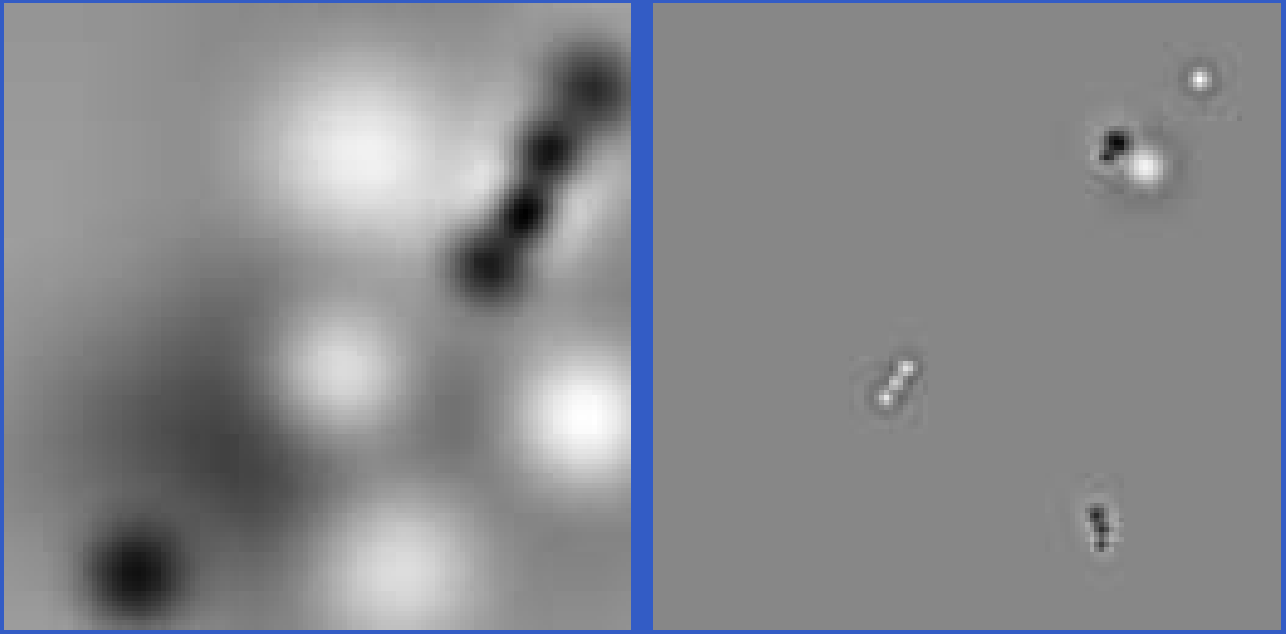
(A,B) Comparison of MSE in function of the structure book size (in $bytes$) and compared to JPEG at different qualities (dashed line: MP, plain line: WMP, dot-dashed line: JPEG) (A) Standard MSE, (B) Mahalanobis MSE; (C) Modulation function for MP and WMP in function of the spike's rank.

Results : image reconstruction (WMP)



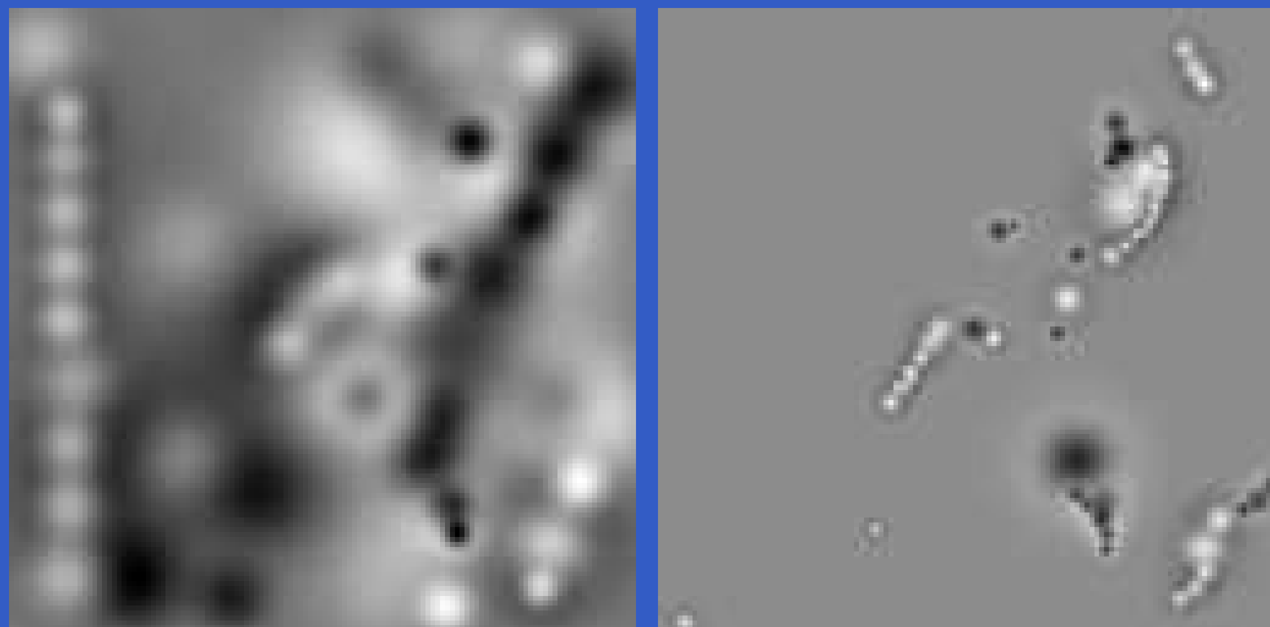
MP ——— WMP

Results : image reconstruction (WMP)



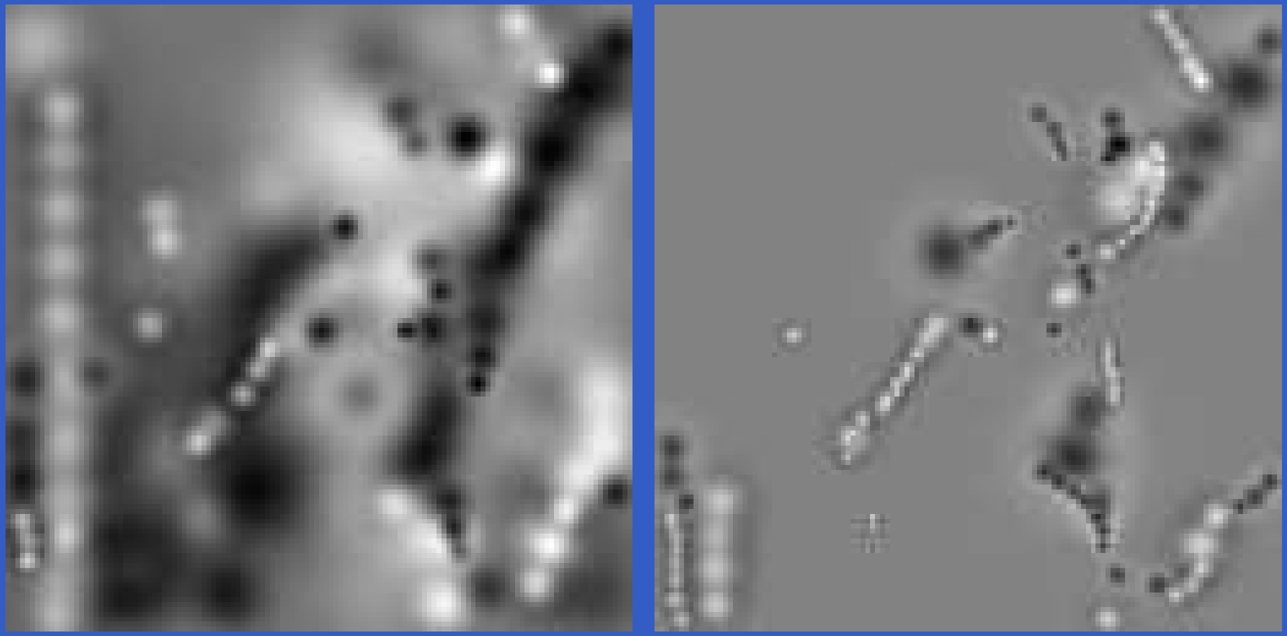
MP ——— WMP

Results : image reconstruction (WMP)



MP ——— WMP

Results : image reconstruction (WMP)



MP ——— WMP

Results : image reconstruction (WMP)

MP ———— WMP

Results : image reconstruction (WMP)

MP ——— WMP

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,
- Optimization through statistics of natural images,

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,
- Optimization through statistics of natural images,
- Future directions:

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,
- Optimization through statistics of natural images,
- Future directions:
  - Better results with direction filters,

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,
- Optimization through statistics of natural images,
- Future directions:
  - Better results with direction filters,
  - Multi-layered,

Conclusion

- Model of an overcomplete representation with spiking neurons with Matching Pursuit,
- Optimization through statistics of natural images,
- Future directions:
  - Better results with direction filters,
  - Multi-layered,
  - Link between statistical optimization and adaptive unsupervised learning: learning the dictionary

Beyond this talk

All source code and documents available on:

`http://laurent.perrinet.free.fr/code/retina.html`

References

J. J. Atick. Could information theory provide an ecological theory of sensory processing? *Neural Computation*, 3(2):213–252, 1992.

W. Gerstner, R. Kempter, J. L. V. Hemmen, and H. Wagner. *Spiking Neurons*, chapter 1, pages 3–54. The MIT Press, Cambridge, MA, 1999. URL `http://www.cis.tu-graz.ac.at/igi/maass/PNN.html`.

S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3414, 1993.

B. Olshausen and D. J. Field. Natural image statistics and effi cient coding. *Network*, 7: 333–339, 1996. URL `citeseer.nj.nec.com/olshausen96natural.html`.

R. Van Rullen and S. J. Thorpe. Rate coding versus temporal order coding: What the retina ganglion cells tell the visual cortex. *Neural Computation*, 13(6):1255–1283, 2001.