Neural noise improves path representation in a simulated network of grid, place, and time cells

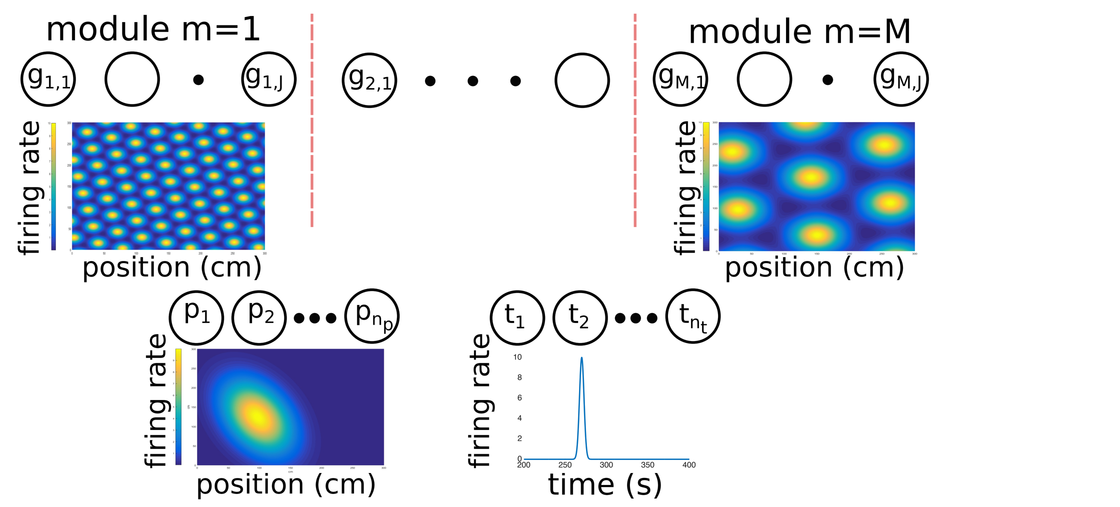
David M. Schwartz[1] and O. Ozan Koyluoglu[1,2,3]

1. Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, AZ
2. Dept. of Electrical Engineering and Computer Science, U.C. Berkeley, CA
3. Huawei R&D, Santa Clara, CA

Abstract

- The joint activity of grid, place, and time cell populations forms a neural code for paths.
- We measure the performance of a network of these populations, as well as interneurons, which implement biologically realizable de-noising algorithms.
- Simulations demonstrate that representation improves when activity of a small fraction of the population is corrupted by noise.

A code for paths in space and time

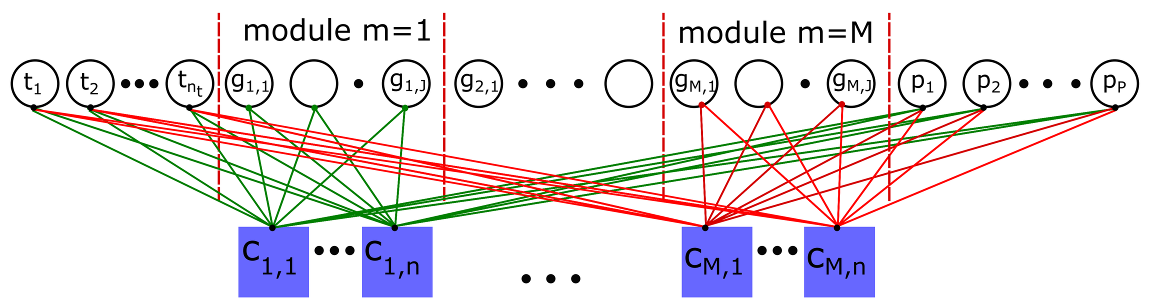


- **Components:** N neurons, M grid modules (m), with J neurons, and T time cells
- **Grid cell tuning curves:**

$$g_{m,j}(\mathbf{s}) = \frac{f_{\max}}{Z}\exp\left[\sum_{k=1}^{3}\cos\left(\frac{4}{\lambda_m\sqrt{3}}\mathbf{u}(\theta_k-\theta_{m,j})\cdot(\mathbf{s}-\mathbf{c}_{m,j})+\frac{3}{2}\right)-1\right]$$

  - $\mathbf{u}(\theta_k-\theta_{m,j})$ is a unit vector in the direction of $\theta_k-\theta_{m,j}$.
  - $\mathbf{s}\in[0,L]\times[0,L]$ is the position stimulus.
  - $\mathbf{c}_{m,j}$, $\theta_{m,j}$, and λ_m are spatial phase offset, orientation offset, and scaling ratio.
  - Orientations: $\theta_{m,j}\in\{-60°,0°,60°\}$.
  - Z is a normalizing constant (≈ 2.857399).
  - $f_{\max}$ is the grid cell's maximum firing rate.
- **Place cells** have bivariate Gaussian tuning curves with mean $\boldsymbol{\xi}\in[0,L]\times[0,L]$, $\rho\in[-\frac{1}{2},\frac{1}{2}]$, and covariance $\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$.
- **Time cells** have univariate Gaussian tuning curves with mean $v_t\in[0,\tau]$ and variance $\sigma_t^2\in[0.5,8]$ seconds.
- **Codewords** (i.e., rows of $\underline{C}$) are formed by concatenating activities of these cells evoked by positions and times from paths recorded from a rat engaging in a spatial navigation task.

De-noising network



- **This network** is a bipartite graph consisting of N pattern neurons and N_I interneurons.
- **Clustering:**
  - Interneurons are split into M distinct clusters of n interneurons per cluster, each connected to a distinct grid module.
  - Interneurons are initially connected randomly to any grid cell in the corresponding module, and any place and time cell.

Subspace learning

- **Before denoising** is possible, this network must learn (i.e., adapt its weights for) the hybrid code.
- **Code subspace learning** is complete when the interneurons may be read to determine if the states of the pattern neurons map to a valid codeword, i.e., when the network has developed a connectivity matrix, W, whose rows are approximately perpendicular to the code space.
- **(anti)Hebbian learning update rule:**

$$\mathbf{w}\leftarrow\mathbf{w}-\alpha_t\left(y\left(\mathbf{x}-\frac{y\mathbf{w}}{\|\mathbf{w}\|^2}\right)+\eta\Gamma(\mathbf{w},\theta)\right),$$

  - α_t is the learning rate at iteration t.
  - $y=\mathbf{x}'\mathbf{w}$ is the scalar projection of $\mathbf{x}$ onto $\mathbf{w}$.
  - θ is a sparsity threshold.
  - η is a penalty coefficient.
  - Γ is a sparsity enforcing function, approximating the gradient of a penalty function, $g(\mathbf{w})=\sum_{k=1}^{m}\tanh(\sigma\mathbf{w}_k^2)$, which, for appropriate choices of σ, penalizes non-sparse solutions early in the learning procedure.

De-noising algorithms

- **Goal:** Recover the correct pattern of activity, $\mathbf{x}$ from the noisy state, $\mathbf{x}_n=\mathbf{x}+\mathbf{n}$, where $\mathbf{n}$ is this noise pattern.
- **$\mathbf{x}_nW'$ reveals inconsistencies in $\mathbf{x}_n$** that the de-noising algorithm seeks to correct in the feedback stage. To see this, consider that $\mathbf{x}_nW'=(\mathbf{x}+\mathbf{n})W'=\mathbf{x}W'+\mathbf{n}W'\approx 0+\mathbf{n}W'$.

Algorithm 1 Sequential de-noising

Require: local weights, W_i, for each cluster, $i\in\{1,...,M\}$, noisy pattern, $\mathbf{x}_n$, stopping threshold, ϵ
Ensure: denoised pattern, $\mathbf{x}_d$

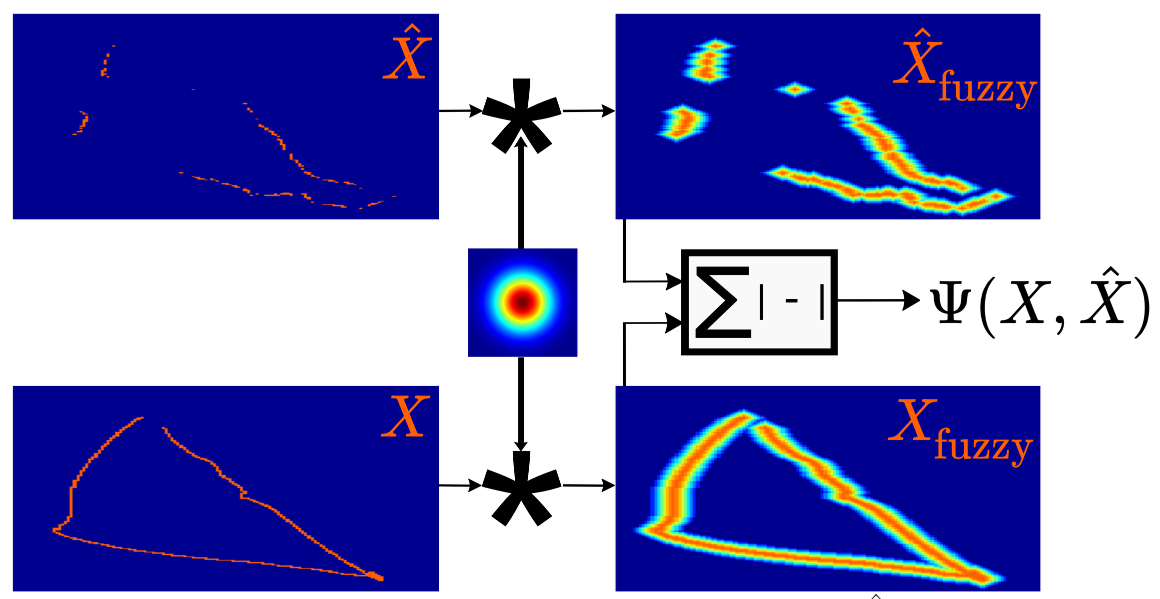
```
x_d ← x_n
while t < T_max or a cluster has an unsatisfied constraint do
    for each cluster, i ∈ {1,...,M} do
        x ← subpattern corresponding to cluster i
        d ← Modular_Recall(x, W_i)
        if |dW_i| ≤ ε then
            x_d(cluster i's subpattern indices) ← d
        end if
    end for
    t ← t + 1
end while
```

Algorithm 2 Modular Recall

Require: local weights for this cluster, W, maximum number of iterations, $T_{\max}$, noisy subpattern, $\mathbf{x}$, feedback threshold, ϕ
Ensure: denoised subpattern, $\mathbf{d}$

```
d ← p
while t < T_max do
    y ← xW'
    r ← y'W
    if ||y|| < ε then
        break;
    end if
    f ← (|y'|·|W|) / (Σ_{i=1}^{m} |W|)
    for each pattern neuron, j do
        if f_j ≥ φ then f_j = sign(x_j)
        else f_j = 0
        end if
    end for
    d ← d + f
end while
```

Fuzzy mean SAD

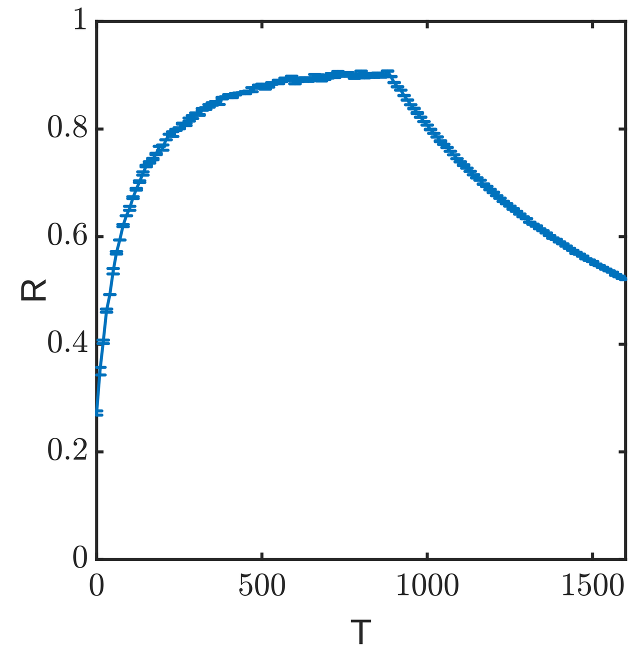


- **Denote** true and estimated path images as X and $\hat{X}$.
- **Construct** X_{fuzzy} and $\hat{X}_{\text{fuzzy}}$, by convolving X and $\hat{X}$ with a bivariate Gaussian distribution of standard deviation, σ_{fuzzy}
- **Fuzzy mean SAD** (sum of absolute differences) is

$$\Psi(X,\hat{X})=\frac{1}{\sum_{i=1}^{N_p}(X_{\text{fuzzy}}(i)+\hat{X}_{\text{fuzzy}}(i))}\sum_{i=1}^{N_p}|\hat{X}_{\text{fuzzy}}(i)-X_{\text{fuzzy}}(i)|,$$

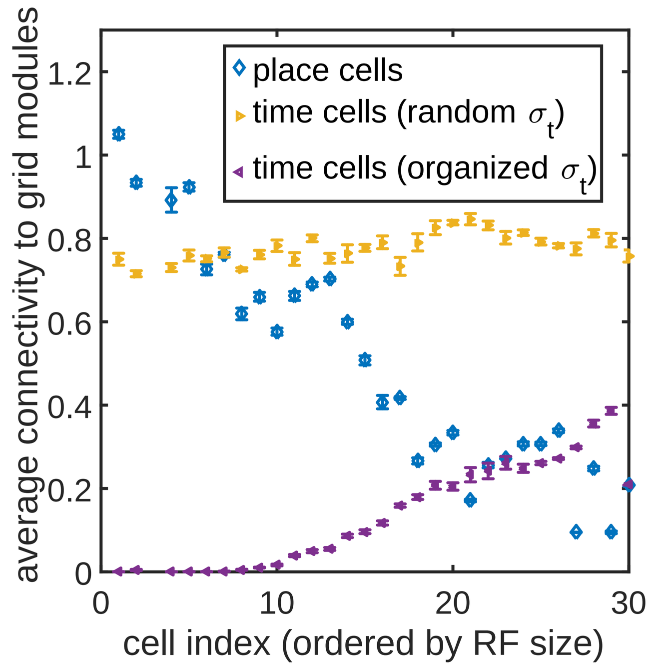
where N_p is the number of pixels in the image.

Coding theoretic results



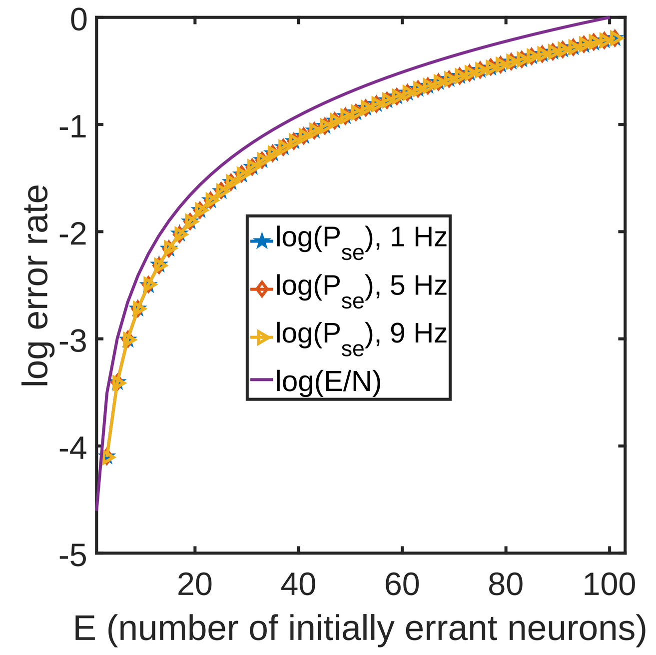
- **Define** $R=\frac{\text{rank}(\underline{C})}{N}$, normalized rank of the code.
- **R increases** with increasing T until additional time cells contribute only redundant information, at which point their inclusion reduces rank.

Subspace learning results



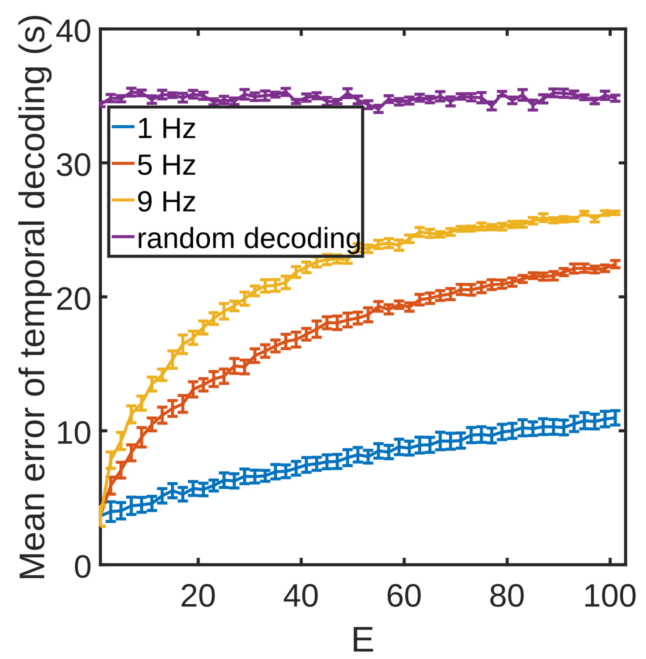
- **Define** connection strength from place cells to grid modules by $\frac{1}{N_I}(\sum_{(i,j)}|w_{i,j}w_{i,p}|)$ (for interneurons, i, grid cells, j, in module m).
- **Average connectivity** of place cells to grid modules decreases with increasing place field width.
- **Surprisingly** time cells exhibit the opposite trend when organized as in [5], where cells firing later in a sequence had wider receptive fields.

De-noising results

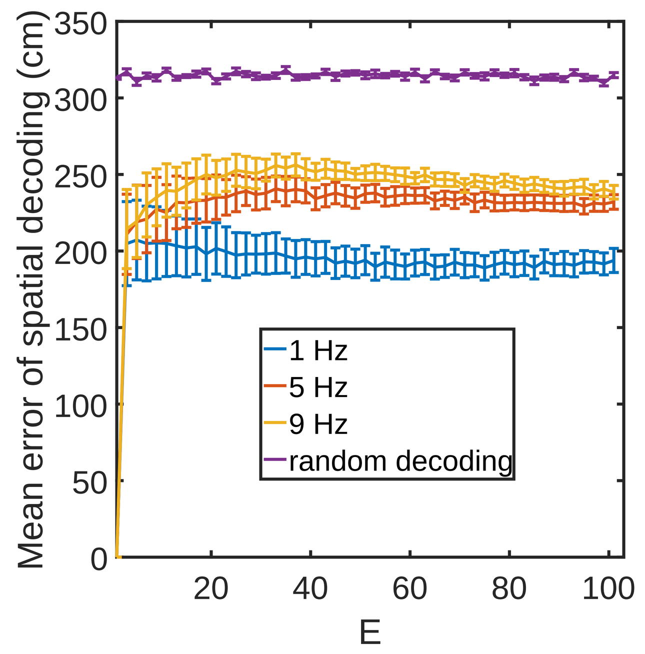


- **Error rate** (fraction of incorrect neurons after de-noising) for E errant neurons before de-noising for different noise frequencies
- **De-noising** reduces errors: $\log_{10}(\frac{E}{N})>\log_{10}(P_{\text{se}})\iff\frac{E}{N}>P_{\text{se}}\iff E>NP_{\text{se}}$.

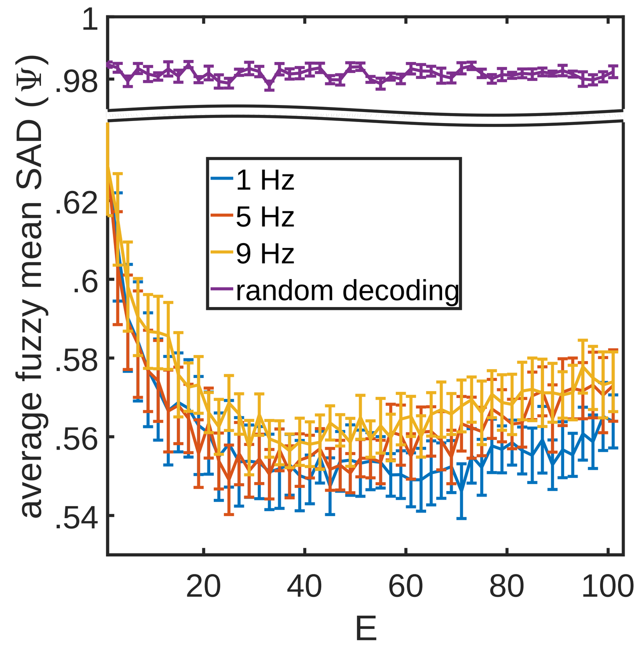
Decoding results



- **Temporal decoding error** increases with increasing magnitude and frequency of noise.



- **Spatial decoding error** also increases with increasing magnitude and frequency of noise.



- **Fuzzy mean SAD** is minimized for intermediate values of E and frequency of noise.
- **Surprisingly** fuzzy mean SAD drops off quickly for small and increasing E and increases slowly when E is near N.

Discussion

- **Readily de-noisable codes** including all cell types may be constructed by proper choice of population parameters.
- **Specific accuracy of decoding** position or time alone decreases with increasing frequency and ubiquity of noise.
- **Average strength of connection** from place cells to grid modules decreases with increasing width of place field.
- **A population of time cells** in which σ_t is positively correlated with v_t exhibits the opposite trend. Surprisingly, without this correlation, the trend disappears.
- **Accuracy of path representation** is maximized when a small number of participating cells are subject to noise with intermediate intensity.

References

[1] H. Stensola, T. Stensola, T. Solstad, K. Frland, M.B. Moser, and E. Moser. "The entorhinal grid map is discretized," Nature 492, no. 7427 (2012): 72-78.
[2] C. MacDonald, K. Lepage, U. Eden, H. Eichenbaum. "Hippocampal 'time cells' bridge the gap in memory for discontiguous events." Neuron 71, no. 4 (2011): 737-749.
[3] E. Oja, J. Karhunen. "On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix." Journal of Mathematical Analysis and Applications 106, no. 1 (1985): 69-84.
[4] A. Karbasi, A. Salavati, A. Shokrollahi. "Iterative learning and denoising in convolutional neural associative memories." In ICML (1), pp. 445-453. 2013.
[5] D. M. Salz., Z. Tiganj, S. Khasnabish, A. Kohley, D. Sheehan, M. W. Howard, and H. Eichenbaum. "Time cells in hippocampal area CA3." Journal of Neuroscience 36, no. 28 (2016): 7476-7484.

Acknowledgments

We thank Sam Gianelli, Blaine Harper, and Jean-Marc Fellous of CENL (Dept. of Psychology University of Arizona) for curating position and time data.
This work was supported by NSF grant IIS 1464349.