Contribution ID: **133** Type: **not specified**

k-Mixture Exponential Hopfield Network

Thursday, 28 August 2025 16:42 (3 minutes)

In this study, we propose the *k-Mixture Exponential Hopfield Network* (kMEHN) as a framework that bridges the Classical Hopfield Network (CHN) and the Modern Hopfield Network (MHN) by integrating their structural characteristics.
The CHN defines an energy function using a single symmetric weight matrix, where stored patterns correspond to stable energy minima \1.
In contrast, the MHN achieves high memory capacity and rapid convergence by employing a smooth, nonlinear energy function over a continuous vector space \2.

The proposed kMEHN inherits the Hebbian rule-based construction of weight matrices from CHN, and combines multiple such matrices via a sum of exponential terms, as used in the exponential-type MHN \[3], thereby forming a novel network that exhibits intermediate properties between the two models.
A key feature of kMEHN is its ability to construct an energy landscape that smoothly integrates contributions from multiple independent groups of memory patterns.
Each weight matrix $W^{(k)}$ is constructed from a specific set of memory patterns $\{\xi_\mu^{(k)}\}_{\mu=1}^{P_k}$ as follows:

(Equation 1)
$W^{(k)} = \frac{1}{N} \sum_{\mu=1}^{P_k} \xi_\mu^{(k)} (\xi_\mu^{(k)})^\top$

Here, each pattern $\xi_\mu^{(k)} \in \{-1, 1\}^N$ is a binary vector of length N, and P_k denotes the number of patterns in the k-th memory group.
The energy function of kMEHN for a state vector $s \in \{-1, 1\}^N$ is given by:

(Equation 2)
$E(s) = -\sum_{k=1}^{K} \exp(s^\top W^{(k)} s)$

This structure enables the integrated influence of multiple memory matrices, rather than relying on a single quadratic form as in CHN, resulting in an energy landscape formally similar to that of MHN.

To evaluate whether the proposed model functions effectively as an associative memory system, we conducted simulations using 4×4 black-and-white binary images.
The state space consisted of all possible image patterns (2^{16} in total), from which several patterns were selected as memory patterns.
The selected memory patterns were partitioned into K groups. Since there are multiple possible ways to partition the patterns, we exhaustively enumerated all possible group partition configurations. For each partition, we constructed a network based on the energy function defined above (Equation 2).

For every constructed network, we computed the energy of all possible states in the state space and identified local minima as those states whose energy could not be lowered by a single spin flip.
Among these local minima, those that matched memory patterns exactly were defined as **target states**, while all others were considered **spurious states**. This allowed us to quantitatively evaluate the number of target and spurious states for each partitioning.

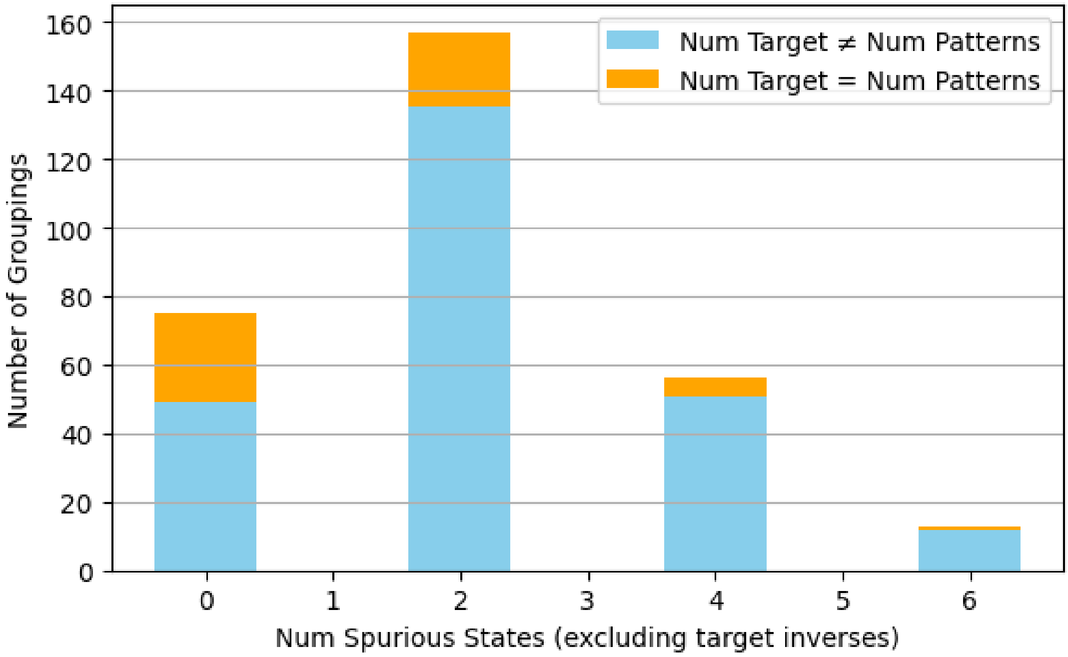

Figure 1: 7 memory patterns, K = 3 groups.

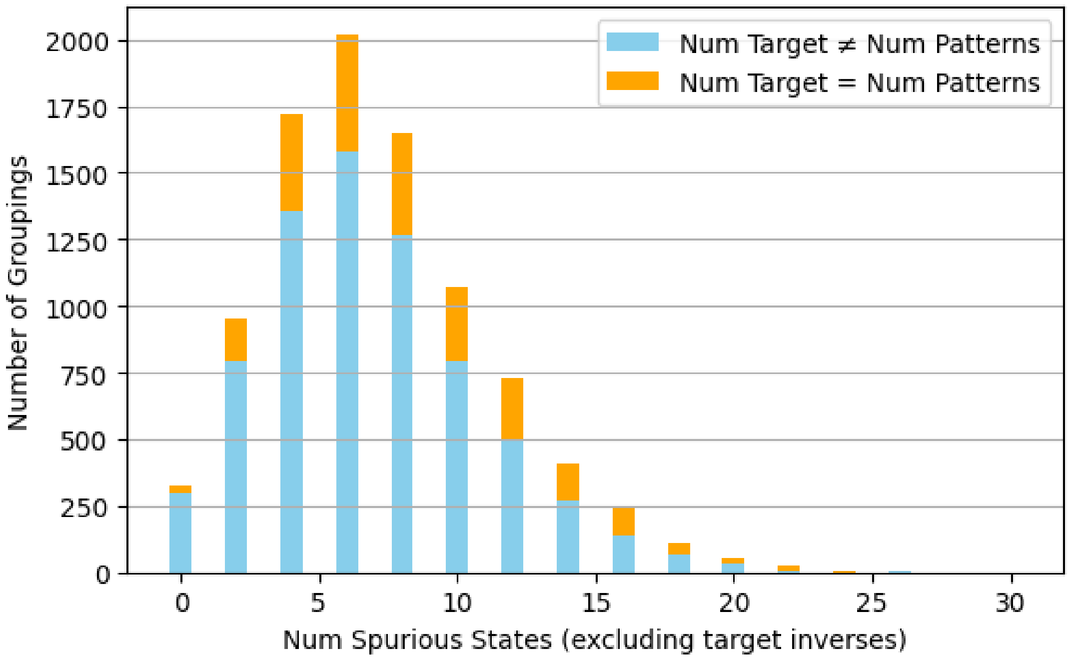

Figure 2: 10 memory patterns, K = 3 groups.

Figure 1: Spurious State Histogram **Description**:
Histograms showing the number of group partition patterns for each spurious state count (excluding bit-inverted target patterns).

- The horizontal axis indicates the number of spurious states.
- The vertical axis shows the number of partitioning patterns with that count.
- Bars are color-coded: **orange** for partitions in which the number of target states matches the number of memory patterns, and **sky blue** otherwise.

Due to the energy function depending on the quadratic forms involving the weight matrices, flipping all bits of a memory pattern does not change the energy, since these quadratic values remain the same. As a result, these inverted patterns consistently appear as spurious states.
These inverted patterns were excluded from the spurious state counts in the histograms.
Additionally, having the number of target states equal to the number of memory patterns indicates that all memory patterns are correctly stored, which is a desirable property for associative memory models.

The results demonstrate that there exist specific partitioning strategies in which spurious states are entirely eliminated.
This finding highlights a promising direction for controlling or reducing spurious states through the design of energy-based heterogeneous associative memory networks.

References

1. Hopfield, J. J. (1982). *Neural networks and physical systems with emergent collective computational abilities.* Proceedings of the National Academy of Sciences, **79**(8), 2554–2558.
2. Demircigil, M., Heusel, J., Löwe, M., Upgang, S., & Vermet, F. (2017). *On a model of associative memory with huge storage capacity.* Journal of Statistical Physics, **168**(2), 288–299.
3. Ramsauer, H., Schäfl, B., Lehner, J., Seidl, P., Widrich, M., Adler, T., … & Hochreiter, S. (2020). *Hopfield networks is all you need.* arXiv preprint arXiv:2008.02217.

Primary author: Mr HONDA, Yuki (Mie University)

Co-authors: Prof. SHIRAI, Nobu C. (Mie University); Prof. MORITA, Kento (Mie University); Prof. WAKABAYASHI, Tetsushi (Mie University)

Presenter: Mr HONDA, Yuki (Mie University)

Session Classification: Poster

Track Classification: Poster presentation