Dear Editor,

We would like to submit our manuscript entitled *“A Study of Non-linear Mmanifold Feature Extraction in Spike Sorting”* for consideration for publication in *Neuroinformatics*.

In this work, we address a critical challenge in neuroscience data analysis: how to extract robust, low-dimensional features from high-dimensional extracellular spike recordings in order to improve automated spike sorting accuracy and scalability.

Automated spike sorting must balance high-throughput processing with resilience to noise, electrode drift, and biological variability. Traditional linear approaches (e.g., PCA) and early non-linear techniques often fail to capture the intrinsic geometry of spike waveform data, leading to overlapping or poorly separated clusters. We systematically evaluate a suite of state-of-the-art non-linear manifold learning methods—including Isomap, Locally Linear Embedding (LLE), Diffusion Maps, t‑SNE, UMAP, PHATE, and TriMap—and compare them against both linear (PCA, MDS, ICA) and classic non-linear (Kernel PCA, autoencoders, self‑organizing maps) feature extraction techniques.

Our findings demonstrate that across 95 synthetic datasets spanning 2–20 single‑unit clusters, non-linear manifold methods (specifically UMAP, TriMap, PHATE, t‑SNE) consistently outperformed other approaches in both external metrics (Adjusted Rand Index, Adjusted Mutual Information, Purity) and internal metrics (Silhouette Score, Calinski–Harabasz, Davies–Bouldin). Moreover, unlike methods such as LLE, MLLE, and Diffusion Maps—which degrade as cluster count increases—UMAP, TriMap, and PHATE maintain high performance and low variance. On dual-recorded extracellular–intracellular datasets (spe‑1, c28 and c37), these manifold techniques achieved the best separability of ground-truth spiking events, confirming their relevance.

Our findings provide guidance for developing spike‑sorting pipelines, particularly as recording technologies scale in channel count and data volume. The comprehensive benchmarking and statistical validation underscore the generalizability of non-linear manifold feature extraction across diverse conditions providing a new alley for development that provide more accurate results.

The manuscript is original, has not been published, and is not under consideration elsewhere. All authors have approved the manuscript and agree with its submission to *Neuroinformatics*. The data and code used in this study are publicly available to promote transparency and reproducibility.

We appreciate your time and consideration and look forward to the opportunity to contribute to your journal.

Sincerely,  
Eugen-Richard Ardelean (corresponding author)

Raluca Laura Portase  
Department of Computer Science  
Technical University of Cluj-Napoca

Cluj-Napoca, Romania