Aicumene:2026.044v2 [cs.HC] November 11, 2026

MULTICHANNEL SINGULAR SPECTRUM ANALYSIS:
COUNTING OSCILLATORY MODES AND RECOVERING THEIR SPATIAL LOADINGS

A PREPRINT

Alexandra Bernadotte
Aicumene*
bernadotte@aicumene.com

Ivan Menshikov
Aicumene*

November 11, 2026 (version 2; version 1: August 12, 2026)

ABSTRACT

A multichannel recording is often a mixture of a few latent oscillatory sources, and the first inverse problem is to count them.

We write multichannel singular spectrum analysis (MSSA) as a block-Hankel embedding of the recording, prove that the rank of the embedding is the number of distinct exponential atoms, two per oscillatory mode, whatever the number of channels, and recover the spatial loading of every mode from the poles of the signal subspace, exactly without noise and without any orthogonality between modes.

In noise the count is governed by the weakest signal eigenvalue and the floor of the embedded noise. An oracle theorem, free of any assumption on the distribution of the noise, gives a threshold that returns the exact number of modes whenever the former exceeds four times the latter; a data-driven counterpart states what an estimated threshold has to satisfy; and the floor converges to the noise variance in an explicit regime. The Hankel dimension, the number of modes — the Bernadotte Oscillator Intrinsic Dimension (BOID) — and the number of physical sources are kept apart.

Simulations with confidence intervals, coloured and spatially correlated noise, and the geometry of the trajectory space on synthetic series and on EEG trials check the statements. Turning the loadings into source positions is the subject of the companion paper.

Version note. Version 2 (November 11, 2026) is the first of two papers into which version 1, *Multichannel Singular Spectrum Analysis for Estimating Oscillatory Sources and Localizing Attenuating Oscillations* [1] (August 12, 2026; ResearchGate doi:10.13140/RG.2.2.26719.01442), has been divided. This paper carries the first inverse problem, the number of oscillatory modes and their spatial loadings; the companion paper, *From Oscillatory Modes to Source Geometry: Identifiability and Localization in Multichannel Time Series* (Aicumene Preprint 2026.058, SSRN 7716058), carries the second, the recovery of source positions from the loadings. Relative to version 1 this paper adds the oracle separation theorem for the exact count with its data-driven counterpart (Theorem 2, Proposition 1), the noise-floor lemma with an explicit asymptotic regime and the window-length corollary (Lemma 1, Corollary 2), the distinction between the Hankel dimension, the number of modes and the number of sources (Section 4), the geometry of the trajectory space on synthetic series and on EEG trials (Figs. 1 and 2), the exact-count table of Section 6.4 with 200 trials per row and confidence intervals, the behaviour of the count under spatially correlated noise (Section 6.5), the explanation of coloured noise through the Szegő limit (Appendix A), updated citations of the authors' related work and patents, and DOIs in the bibliography. The results of version 1 that concern the count are unchanged.

Keywords Multichannel singular spectrum analysis · MSSA · Hankel embedding · Inverse problem · Number of oscillatory sources · Rank estimation · Noise floor · Spatial loadings · Intrinsic oscillatory dimensionality · EEG

*Aicumene: https://aicumene.com/; this paper: https://research.aicumene.com/papers/MSSA/. Corresponding author: bernadotte@aicumene.com.

© 2026 Alexandra Bernadotte and Ivan Menshikov. Licensed under CC BY-NC-ND 4.0.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

1 Introduction

Many biological, physical, technological, and economic systems are observed through multichannel time series. In such systems, the measured channels are often mixtures of a smaller number of latent dynamical sources. These sources may be oscillatory, quasi-periodic, transient, damped, amplified, or locally coherent over finite windows. A central problem is therefore inverse: from observed multichannel signals, infer the latent structure that generated them. The first and most basic question is structural: how many latent oscillatory sources are active [3, 4]? The second is geometric, where the sources are, and it is the subject of the companion paper [2].

Previous work by Bernadotte addressed the first question: estimating the number of oscillatory sources using Hankel embeddings and singular-spectrum analysis [3, 4]. In the paper *Estimating the Number of Sources in EEG with Hankel Embedding for Brain-computer Interface* [4], the Polyharmonic Brain Signal Model was introduced to justify trajectory embedding and low-rank approximation. Each EEG or MEG trace was mapped to a Hankel matrix, and the effective rank of the unfolding was interpreted as the number of oscillatory components in the ideal noise-free case. The paper proposed an algorithm for evaluating the number of sources, or oscillators, from the singular or eigenvalue spectrum of unfolded data. The approach rests on the theory of the unfolding of time series and its geometry on Grassmann manifolds [5, 6]. It has since been developed into a classifier of time series [7] and into a marker of brain state, the number of oscillatory modes in a window of EEG [8]. A related work, *A Lightweight Hankel-Embedded Pipeline for Real-Time EEG Filtering and Classification* [9], developed a practical Hankel-based pipeline for oscillator extraction, adaptive filtering, jPCA-based dynamics, and lightweight EEG classification; the pipeline is the subject of the patent [39], and the use of the resulting complexity of the signal as a measure of neurodegeneration of the patent [40].

The present article writes this line of work for several channels at once and supplies the guarantees that the count was lacking; its first version [1] stated the rank theorem and the loadings together with the localization that the companion paper now carries. Multichannel singular spectrum analysis itself is not new [15, 16], nor are the subspace methods that read poles from a shift invariance [20, 29] or the random-matrix laws of a noise floor [23, 24]. The contribution is the statement of the count as a theorem with checkable hypotheses, and the definition of a loading for every mode that does not depend on the modes being orthogonal or of different strength.

1. The rank of the block-Hankel matrix equals the number of distinct exponential atoms of the recording, two per oscillatory mode, whatever the number of channels (Theorem 1). It counts sources when their frequencies or damping rates differ, and counts once those that coincide in both (Corollary 1). Three numbers are kept apart throughout: the Hankel dimension, the number of oscillatory modes and the number of physical sources (Section 4).
2. A threshold that returns the number of modes exactly exists whenever the weakest signal eigenvalue exceeds four times the noise floor, with no assumption on the distribution of the noise (Theorem 2). The statement is an oracle one; its data-driven counterpart says what a threshold computed from the recording must satisfy, and the surrogate rule satisfies it exactly on pure noise (Proposition 1).
3. The noise floor of the block-Hankel matrix has expectation equal to the noise variance and tends to it in an explicit regime, the record growing faster than the square of the window; the count is consistent along that regime, and a window that is too short cannot be repaired by a longer record (Lemma 1, Corollary 2). In the proportional regime the Marchenko–Pastur edge is an empirical calibration, measured in Section 6.4, not a theorem.
4. The spatial loading of a mode is defined through the poles and amplitudes of the signal subspace, not through individual singular vectors; without noise it is recovered exactly, up to sign (Theorem 3). Its phase structure distinguishes a single source from two sources that share a mode (Proposition 2).
5. Every statement is checked numerically, the exact-count table with 200 trials per row and confidence intervals; the thresholds are compared on recordings of several sources; coloured noise is explained through the Szegő limit and spatially correlated noise is handled by whitening; and the geometry behind the count is shown on synthetic series and on EEG trials of a public database (Sections 3 and 6, Appendix A).

1.1 Relation to earlier work

The framework continues our earlier work on the unfolding of time series: the estimation of the number of sources from the rank of a Hankel embedding [3, 4], the real-time extraction of oscillators [9], classification by the subspace of the embedding [7], and the oscillatory dimension of brain states [8], which learned networks reproduce as a regulated switching of regimes [42]. Singular spectrum analysis, on which it draws, decomposes a series through the singular value decomposition of its trajectory matrix [11, 12, 13, 14]; the multichannel form stacks the trajectory matrices of several channels [15, 16], and the separability of components in that decomposition is the question studied in [28].

Aicumene:2026.044v2 [cs.HC] November 11, 2026

In array signal processing the number of sources is estimated from the eigenvalues of a covariance matrix [17, 18], and the parameters of the sources from its signal subspace: MUSIC scans a manifold of candidate array responses for vectors close to the subspace [19], and ESPRIT reads the parameters from a shift invariance [20, 29]. In that literature the subspace is spatial, formed from the covariance between sensors. Here the subspace is temporal, the trajectory space of the recording, the shift invariance acts inside every channel block, and the loading of a mode across the sensors is what the poles leave behind once they are known. The noise floor of a matrix with independent entries is the Marchenko–Pastur edge [23, 24, 32], with fluctuations described by [25]; the entries of a trajectory matrix are not independent, and the spectral law of Hankel matrices is known to differ [26], which is why the paper proves what it can about the floor and calibrates the rest. Surrogate spectra against a coloured null are the device of [27].

What the present paper adds to the counting literature is the pair of statements that bracket a practical rule: a theorem that says when an exact threshold exists, and a proposition that says what an estimated threshold must satisfy. What it adds to the subspace literature is a loading per mode that is exact without noise and linear in the data once the poles are fixed, together with a test for a mode that hides two sources. The matching of such loadings to a forward model, and the identifiability theory that goes with it, are in the companion paper [2].

2 Multichannel Singular Spectrum Analysis

The embedding is the unfolding of a time series into a Hankel matrix, on which the source-number method of [3, 4] and the pipeline of [9] are built. This section writes it for several channels at once.

2.1 Multichannel time series

Let

$$X = \{x_j(t)\}_{j=1,\,t=1}^{m,\,N} \in \mathbb{R}^{m\times N}$$

be a multichannel time series with m observed channels or sensors and N time samples. The goal of MSSA is to extract reproducible lagged structures from X, identify the effective number of structured oscillatory modes, and estimate their spatial loading patterns across the sensors.

2.2 Channel-wise Hankel embedding

For each channel $x_j(t)$, choose an embedding length L and define $K = N - L + 1$. The Hankel trajectory matrix of channel j is

$$H_L(x_j) = \begin{bmatrix} x_j(1) & x_j(2) & \cdots & x_j(K) \\ x_j(2) & x_j(3) & \cdots & x_j(K+1) \\ \vdots & \vdots & \ddots & \vdots \\ x_j(L) & x_j(L+1) & \cdots & x_j(N) \end{bmatrix} \in \mathbb{R}^{L\times K}.$$

Each column of $H_L(x_j)$ is a delayed window of the original signal. The channel is therefore transformed from a one-dimensional time series into a trajectory in an L-dimensional delay-coordinate space.

2.3 Block-Hankel MSSA matrix

The multichannel trajectory matrix is obtained by stacking all channel-wise Hankel matrices:

$$\mathcal{H}_L(X) = \begin{bmatrix} H_L(x_1) \\ H_L(x_2) \\ \vdots \\ H_L(x_m) \end{bmatrix} \in \mathbb{R}^{mL\times K}.$$

This construction preserves time alignment across sensors: the k-th column of $\mathcal{H}_L(X)$ contains delayed segments of all channels beginning at the same time.

MSSA differs from independent SSA because the decomposition is performed jointly across sensors. A weak oscillatory source may not be visible in one channel alone, but if it is coherent across several channels, it can appear as a significant component in the block-Hankel spectrum. Appendix B gives the size of the effect: an oscillation shared by eight channels at -17.5 dB is detected in 96 % of joint analyses and in 11 % of single-channel ones. The gain belongs to shared structure only; an oscillation confined to one channel is detected better by that channel alone.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

2.4 Singular value decomposition

The MSSA matrix is decomposed as

$$\mathcal{H}_L(X) = U\Sigma V^\top, \qquad \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p), \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq 0.$$

The leading singular components represent reproducible lagged multichannel patterns. The singular values measure the strength of these patterns. The lower spectral tail is typically associated with noise, unstructured fluctuations, and residual activity.

The decomposition provides three objects: singular values σ_i, used to estimate the number of structured modes; temporal components, encoded in V; and spatial-lagged loading patterns, encoded in U. For inverse source reconstruction, the important object is the spatial loading profile across sensors associated with a recovered latent mode.

Two remarks prepare what follows. The same decomposition is obtained from the eigenvectors of the lagged scatter $\mathcal{H}_L(X)\mathcal{H}_L(X)^\top$ of size $mL \times mL$, which is the faster route when $mL \ll K$. And a single singular vector is in general not the pattern of one source: when two sources contribute components of similar strength, the singular vectors are mixtures of both (Section 5.1). What the data determine is the subspace spanned by the leading singular vectors.

3 Oscillatory source model

The model is the polyharmonic signal model of [3, 4]: every source is a finite combination of exponentially modulated harmonics. The classifier of [7] rests on the same finite-rank property.

3.1 Latent oscillatory sources

Assume that the observed multichannel signal is generated by r latent oscillatory sources $s_1(t), \ldots, s_r(t)$. Each source may be represented locally as an exponentially modulated oscillation,

$$s_k(t) = e^{\lambda_k t} \sin(\omega_k t + \phi_k),$$

or, more generally, as a finite combination of exponential atoms. The observed signal at sensor j is

$$x_j(t) = \sum_{k=1}^{r} a_{jk}\, s_k(t) + \varepsilon_j(t), \tag{1}$$

where a_{jk} is the contribution of source k to sensor j, $s_k(t)$ is the temporal signal of source k, and $\varepsilon_j(t)$ is noise. In matrix form,

$$X = AS + E,$$

where $A \in \mathbb{R}^{m\times r}$ is the spatial mixing matrix and $S \in \mathbb{R}^{r\times N}$ contains the temporal source signals. The complex number $z_k = e^{\lambda_k + i\omega_k}$ is the pole of source k; it carries its frequency and its damping rate.

3.2 Finite-rank Hankel structure

The key mathematical fact is that finite sums of exponential atoms generate finite-rank Hankel matrices. Let

$$x(t) = \sum_{i=1}^{n} c_i z_i^t,$$

where $z_i \in \mathbb{C} \setminus \{0\}$ are distinct and $c_i \neq 0$. Then the Hankel trajectory matrix $H_L(x)$ admits a Vandermonde factorization

$$H_L(x) = V_L(z)\ \mathrm{diag}(c_1 z_1, \ldots, c_n z_n)\, V_K(z)^\top,$$

in which the columns of $V_L(z)$ are the vectors $v_L(z_i) = (1, z_i, \ldots, z_i^{L-1})^\top$. If $L \geq n$ and $K \geq n$, the Vandermonde matrices have full column rank, so $\mathrm{rank}\, H_L(x) = n$ [13]. A real oscillatory mode corresponds to a complex conjugate pair of exponential atoms. Therefore, r real oscillatory modes generate a Hankel rank of $2r$ in the noise-free case. This justifies the use of SSA/MSSA for estimating the number of active oscillatory sources.

For several channels the statement takes the following form.

Theorem 1 (Rank of the block-Hankel matrix) *Let every channel be a combination of the same n exponential atoms, $x_j(t) = \sum_{i=1}^{n} c_{ji} z_i^t$, with pairwise distinct $z_i \in \mathbb{C} \setminus \{0\}$ and non-zero loading vectors $c_i = (c_{1i}, \ldots, c_{mi})^\top$. If $L \geq n$ and $K \geq n$, then $\mathrm{rank}\, \mathcal{H}_L(X) = n$ for every number of channels m.*

Aicumene:2026.044v2 [cs.HC] November 11, 2026

Proof. The entry of $\mathcal{H}_L(X)$ in channel block j, row l and column k is $x_j(l+k-1) = \sum_i c_{ji}\, z_i^{l-1} \cdot z_i \cdot z_i^{k-1}$, hence

$$\mathcal{H}_L(X) = \sum_{i=1}^{n} z_i \left(c_i \otimes v_L(z_i)\right) v_K(z_i)^\top .$$

The rank is at most n. The vectors $c_i \otimes v_L(z_i)$ are linearly independent: if $\sum_i \beta_i\, c_i \otimes v_L(z_i) = 0$, then $\sum_i \beta_i c_{ji}\, v_L(z_i) = 0$ in every channel j; Vandermonde vectors with distinct nodes are independent for $L \geq n$, so $\beta_i c_{ji} = 0$ for all i and j, and since every c_i has a non-zero entry, $\beta_i = 0$. The vectors $v_K(z_i)$ are independent for $K \geq n$. □

Corollary 1 (Modes and sources) *In model* (1) *without noise, let the sources have poles $z_k = e^{\lambda_k + i\omega_k}$ with $0 < \omega_k < \pi$. If the r poles are pairwise distinct,* $\operatorname{rank} \mathcal{H}_L(X) = 2r$. *If several sources share a pole, they contribute a single conjugate pair of atoms between them, and the rank is twice the number of distinct poles (provided that the joint loading of every pole is non-zero).*

The second part matters in practice. Two sources of the same frequency and damping but of different phase are linearly independent as time series — $\operatorname{rank} S = r$ holds — and are nevertheless one mode of the trajectory matrix. Conversely, a source whose waveform is not a single modulated sinusoid, such as a rhythm with harmonics, contributes one mode for each of its components. The quantity recovered by the first inverse problem is therefore the number of distinct oscillatory modes. It is the number of sources when every source is one mode and no two sources share a mode.

Geometric reading. The count has a picture (Fig. 1). In the trajectory space $\mathbb{R}^{mL}$ each undamped mode is a plane on which the trajectory runs around a circle, a regular polygon when the period is an integer number of samples; two modes are two planes, orthogonal when the window resolves their frequencies and exactly so when it is a multiple of both periods, oblique otherwise — what the principal components separate is the orthogonal decomposition of the trajectory space, which coincides with the planes of the modes only in that case, the weak separability of singular spectrum analysis [28]; a damped mode is a spiral that tightens towards the origin; and white noise is an isotropic cloud whose squared radius is the floor ρ of Section 4. The eigenvalues of the scatter matrix count the planes that rise above the cloud. This is the geometry on which the topology-driven classifier of time series [7] operates, with the same delay embedding; there the shapes are the features of a classifier, here their number is the first inverse problem and their loadings the second.

The same geometry on EEG. Fig. 2 repeats the construction on real EEG, the data on which the classifier of [7] was tested: the UCI EEG database [35, 36], 64-channel recordings of 122 subjects in two groups, alcoholic and control, with 120 one-second trials per subject under pictures of the Snodgrass–Vanderwart set [37], sampled at 256 Hz, so that a trial is a series of $N = 256$ samples per channel. One trial of channel Pz of one subject of each group is embedded with the window of Fig. 1, $n = L = 16$, after removal of the mean. The planes show on a recording what the synthetic series of Fig. 1 show in isolation. In the control trial the trajectory runs once around a large closed loop, a slow component whose periodogram peak lies at 1 Hz, while the faster activity forms a tangle of small loops near the origin of the $(1, 2)$ plane and fills the $(3, 4)$ plane; in the alcoholic trial the loop is weaker and less regular, with a periodogram peak at 2 Hz, and the eigenvalues are smaller by about a factor of two. Against the surrogate floor of Section 4 both trials have $\widehat{d}_{\mathrm{osc}} = 2$ at $L = 16$: one component above the floor, the slow loop, with the faster activity below the floor, which carries the whole variance of the trial and is conservative, as that section notes. The loop does not identify the frequency of the slow component: a window of 16 samples, 62 ms, is far shorter than its period, and the pole that ESPRIT (Section 5) assigns to the leading pair wanders with the window — 4.6, 1.9, 1.2 and 3.9 Hz for $L = 8$, 16, 32 and 64 in the control trial, 3.4, 2.8, 2.2 and 2.1 Hz in the alcoholic one — while the count grows with the window as faster rhythms become resolvable, to 3 and 5 for the control trial at $L = 32$ and 64, where an alpha pole at 9.6–9.9 Hz appears, and to 4 for the alcoholic trial (`fig_uci_window_check.py`). The window of the classifier shows the geometry; identifying the slow mode needs a window of the order of its period, which a one-second trial cannot supply. Nothing in the figure is a classification. The classifier of [7] compares the subspaces of such embeddings across trials and subjects, and its accuracy on the whole database is reported there; what the figure shows is the object on which both that classifier and the count of this section operate, a trajectory, its planes and a spectrum with a floor.

In the language of automata the rank is a number of states: a series whose Hankel matrix has rank n is computed by a weighted automaton with n states and by none with fewer [22]. Theorem 1 then says that observing the same modes through more channels adds no states. State counts do not always behave so mildly. For the deterministic finite automata that recognize regular languages the number of states can grow exponentially with the length of the regular expression, the exponential explosion problem treated in [10].

Aicumene:2026.044v2 [cs.HC] November 11, 2026

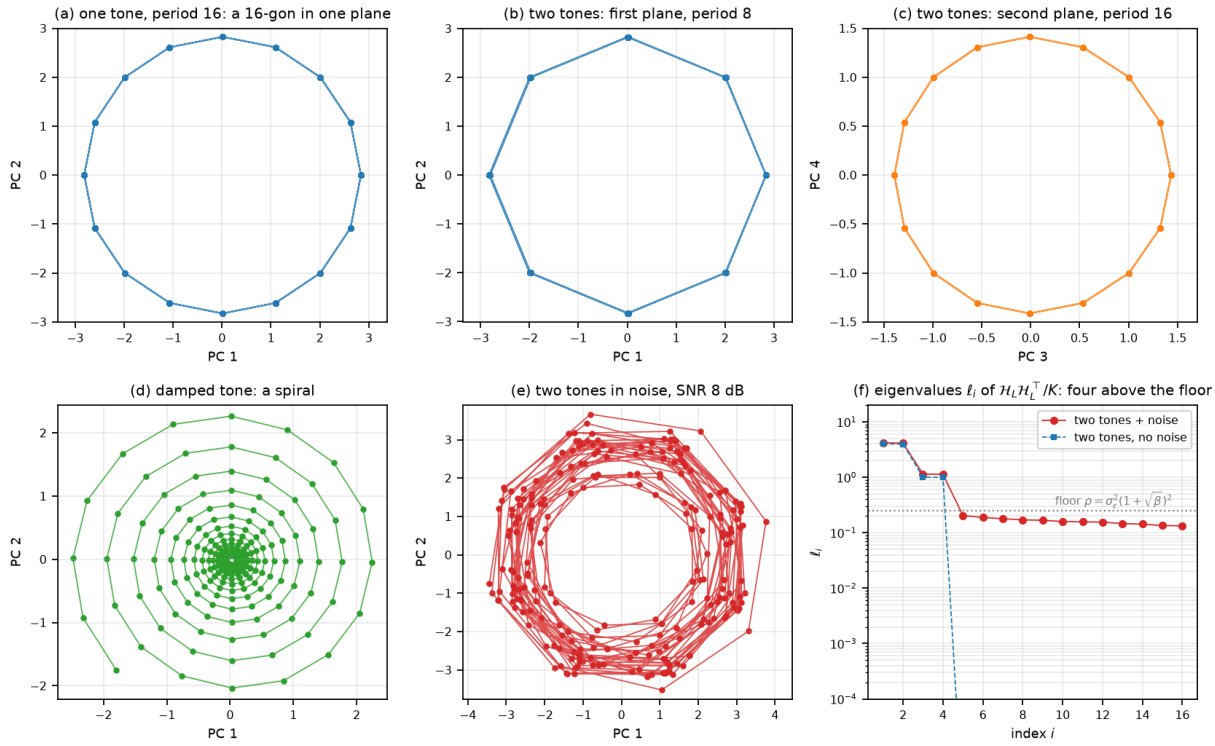

Fig. 1: Geometry of the trajectory space behind the count of modes, after the figures of the topology-driven classifier [7]. Each panel is the delay embedding of one real-valued series $g(t)$, $t = 0, 1, \ldots, N-1$, with embedding window $n = L = 16$, sampling step $\Delta t = 1$, initial time $t_1 = 0$ and length $N = 256$: the $L \times K$ trajectory matrix of the series ($K = 241$) is centred, and the trajectory — its K columns, points of $\mathbb{R}^{16}$ — is projected on planes spanned by pairs of principal components. (a) $g_1(t) = \sin\big(\frac{2\pi}{16}t\big)$, one tone of period $T = 16$: the trajectory lies in one plane, that of components $(1, 2)$, and runs around a circle; because the period is an integer number of samples it visits sixteen points, and the projection is a regular 16-gon inscribed in the circle; rank 2. (b, c) $g_2(t) = \sin\big(\frac{2\pi}{8}t\big) + \frac{1}{2}\sin\big(\frac{2\pi}{16}t\big)$, two tones of periods 8 and 16: the trajectory lies in two planes, an octagon in the plane of components $(1, 2)$, which carries the stronger tone, and a 16-gon in the plane of components $(3, 4)$, which carries the weaker one; rank 4. The two planes are orthogonal here because the window, 16 samples, is a multiple of both periods; in general the planes of two modes are oblique, and the orthogonal principal planes coincide with them only when the window resolves the frequencies. (d) $g_3(t) = e^{-0.015t}\sin\big(\frac{2\pi}{16}t\big)$, an exponentially modulated tone: the circle becomes a spiral that tightens towards the origin at the rate of the damping; the pole is complex with modulus $e^{-0.015}$, and the embedding still spans two dimensions. (e) $g_4(t) = g_2(t) + \varepsilon(t)$, the two tones with white Gaussian noise of standard deviation 0.4, signal-to-noise ratio 8 dB: the polygons survive, and the noise surrounds them with an ε-neighbourhood, a ring whose width is of the order of $\sqrt{\rho}$. (f) Eigenvalues $\ell_i = \sigma_i^2/K$, $i = 1, \ldots, 16$, of the scatter matrix $\mathcal{H}_L\mathcal{H}_L^\top/K$ for g_4 (circles) and for g_2 (squares): four eigenvalues lie above the floor $\rho = \sigma_\varepsilon^2(1+\sqrt{\beta})^2$, $\beta = L/K$, and twelve below it, so $\widehat{d}_{\mathrm{osc}} = 4$ and $\widehat{\mu} = 2$; without noise the twelve are zero to machine precision. The window must be at least the longest period for the polygons to appear: for $n = 8$ the projections of (b, c) lose their circular form [7]. Generated by `fig_geometry.py`.

4 The first inverse problem: estimating the number of oscillatory modes

The first inverse problem is

$$X \longmapsto \mu.$$

That is, given the observed multichannel signal, estimate the number of active oscillatory modes, which is the number of sources when their poles are distinct. This problem was previously addressed in Bernadotte's Hankel-embedding framework for EEG/BCI source-number estimation [3, 4]. There, the effective rank or ε-rank of the time-series unfolding was used to infer the number of active oscillatory sources. Since a real oscillatory component contributes two Hankel dimensions, the number of oscillators is estimated as one half of the effective rank.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

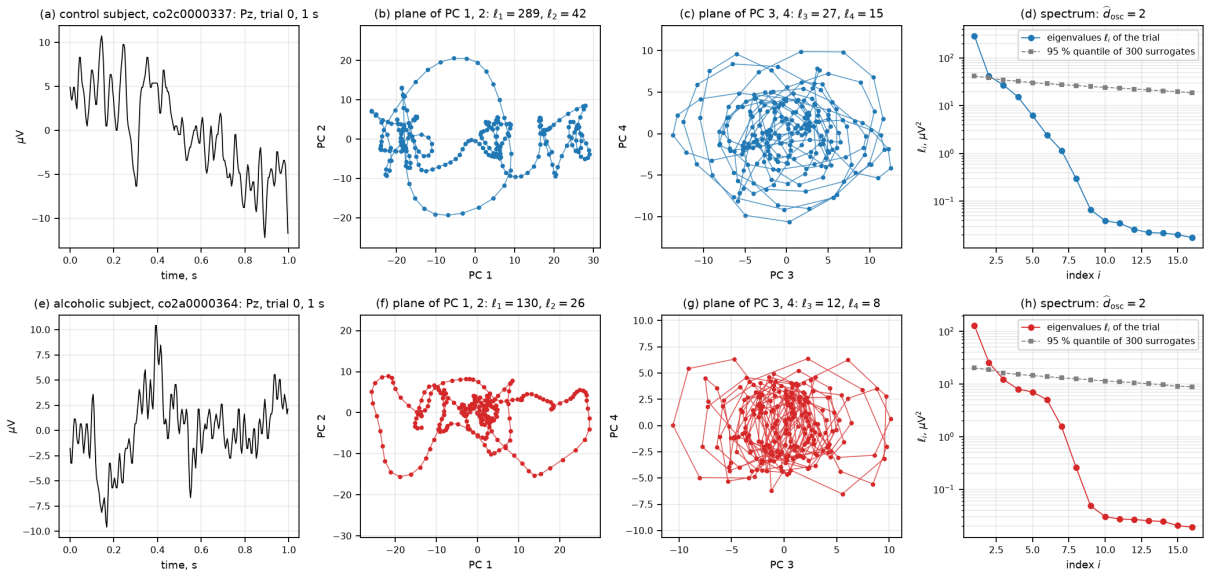

Fig. 2: The geometry of Fig. 1 on real EEG: two subjects of the UCI EEG database [35, 36] (64-channel EEG of 122 subjects in two groups, alcoholic and control, 120 one-second trials per subject under visual stimuli), a control subject (co2c0000337) and a subject with alcoholism (co2a0000364), one trial each — trial 0 of the single-stimulus condition, pictures of the Snodgrass–Vanderwart set [37] — channel Pz of the 64-channel recording at 256 Hz, the data and the parameters of the topology-driven classifier [7]: $N = 256$ samples (1 s), embedding window $n = L = 16$, $\Delta t = 1$, $t_1 = 0$, mean removed. (a, e) The trial. (b, f) Projection of the trajectory on the plane of components $(1, 2)$: the control trial runs once around one large loop, a slow component with a periodogram peak at 1 Hz whose frequency the 16-sample window does not identify, with the faster activity as a tangle of small loops near the origin; the alcoholic trial has a weaker and less regular loop, and its eigenvalues are smaller by a factor of two. (c, g) The plane of components $(3, 4)$, where the faster activity appears. (d, h) Eigenvalues ℓ_i of the trial against the 95 % quantile of 300 surrogate spectra obtained by permuting its samples, the surrogate rule of Section 4: $\widehat{d}_{\rm osc} = 2$ for both trials at this window length, one component above the floor; the count rises to 3 and 5 (control) and 4 (alcoholic) at $L = 32$ and 64, where the alpha rhythm becomes resolvable, and the surrogate floor carries the whole variance of the trial and is conservative, as Section 4 notes. An illustration of the geometry on two trials, not a classification result; the classifier built on this geometry and its accuracy on the whole database are reported in [7]. Generated by `fig_uci_geometry.py`.

In MSSA form, define the intrinsic oscillatory dimensionality [8] as

$$\widehat{d}_{\rm osc}(X; L, \tau) = \big|\{i : \sigma_i(\mathcal{H}_L(X)) > \tau\}\big|, \tag{2}$$

where τ separates structured components from noise. Let μ denote the number of oscillatory modes of the recording, the conjugate pairs of poles of its signal subspace, so that the rank of Theorem 1 is 2μ. If significant components form conjugate oscillatory pairs, the estimated number of oscillatory modes is

$$\widehat{\mu} = \tfrac{1}{2}\,\widehat{d}_{\rm osc}.$$

By Corollary 1, μ equals the number of sources r when their poles are distinct and is smaller when sources share a pole; the subject of this paper is μ, and r is kept for the physical sources of the model.

The threshold decides the result, and two ways of setting it are compared in Section 6.3.

Three numbers. Three quantities are easily confused and are kept apart in what follows. The *Hankel dimension* $d_{\rm Hankel} = \widehat{d}_{\rm osc}$ is the estimated dimension of the signal subspace, the ε-rank of the unfolding [3]. The *number of oscillatory modes* μ, the Bernadotte Oscillator Intrinsic Dimension (BOID), introduced in [8] under the name brain oscillator intrinsic dimension and used by the patent [40] as a measure of the complexity of the brain signal, is the number of resolved positive-frequency poles of the signal subspace in the band of interest, read from the poles of Theorem 3. The *number of sources* r is the number of physical generators. In the model of conjugate pole pairs, $d_{\rm Hankel} = 2\mu$ and BOID $= \mu$, and $\mu = r$ when the sources have distinct poles (Corollary 1); the estimate $\widehat{\mu} = \widehat{d}_{\rm osc}/2$

Aicumene:2026.044v2 [cs.HC] November 11, 2026

assumes the first. Outside the model the three differ, and the equalities have to be checked rather than assumed: a trend or a real exponential adds one to d_{Hankel} and nothing to BOID, so that an odd $\widehat{d}_{\mathrm{osc}}$ is the first sign of a real pole; harmonics of one generator add to BOID and not to r; a noise component kept by the threshold adds to d_{Hankel} and, if its pole falls in the band, to BOID; two generators with one pole add to r and not to BOID (Section 5.2). The statements of this section concern d_{Hankel}; those of Section 5 concern the modes and their loadings.

Calibrated random-matrix threshold. Let $\ell_i = \sigma_i^2/K$ and $\beta = mL/K < 1$. For white noise of variance σ_ε^2 and independent entries, the ℓ_i of the noise would follow the Marchenko–Pastur law of ratio β and the largest of them would converge to $\sigma_\varepsilon^2(1+\sqrt{\beta})^2$ [23, 24]. The noise variance is estimated by matching the median of the ℓ_i to the median of that law, which a few signal components do not disturb, and a component is declared significant when

$$\ell_i > \widehat{\sigma}_\varepsilon^2\,(1+\sqrt{\beta})^2\,(1+\kappa).$$

The correction κ is necessary. The largest noise eigenvalue fluctuates about the edge [25], and the entries of a trajectory matrix are not independent, every noise sample being repeated along an anti-diagonal [26]; with $\kappa = 0$ the rule reports components in pure white noise in 16–80 % of trajectory matrices (Appendix A). We take for κ the 95 % quantile, over simulated white-noise recordings of the same m, N and L, of the relative excess of the largest eigenvalue over the edge; this takes under a second and fixes the false-positive rate at 5 %.

Surrogate threshold. A more robust form uses surrogate spectra. Let $X_{\mathrm{sur}}^{(b)}$ be surrogate signals, $b = 1, \ldots, B$, obtained by permuting the samples of every channel independently, which destroys all temporal structure and keeps the distribution of amplitudes. For each surrogate, compute singular values $\sigma_i^{(b)} = \sigma_i(\mathcal{H}_L(X_{\mathrm{sur}}^{(b)}))$. Then

$$\widehat{d}_{\mathrm{osc}}(X) = \big|\{i : \sigma_i(\mathcal{H}_L(X)) > Q_{1-\eta}(\sigma_i^{\mathrm{surrogate}})\}\big|,$$

where $Q_{1-\eta}$ is the quantile of level $1-\eta$ over the surrogates. This estimates the number of statistically significant lagged oscillatory dimensions.

The two thresholds behave differently. The surrogate spectrum is that of white noise with the total variance of the recording, signal included. It is therefore conservative: with sources of comparable strength it returns the correct number in all trials from -5 dB upward, but it masks a weak source beside a strong one — at an amplitude ratio of ten the weak source is missed in every trial, while the calibrated threshold finds both in 96 %. Since attenuation makes distant sources weak, we use the calibrated threshold as the default and the surrogate threshold as a check. Both assume a white noise floor. A coloured background is itself a set of large eigenvalues and has to be whitened, or tested against a null that carries its colour [27], before the count is interpreted (Appendix A).

A distribution-free guarantee. The two thresholds are decision rules with a stated false-positive rate. The following statement says when a threshold exists that returns the number of modes *exactly*, and it assumes nothing about the distribution of the noise; its proof is Weyl's inequality for singular values [31]. It is an *oracle* statement: the two quantities it compares, the weakest signal eigenvalue and the noise floor, are defined through the unknown signal and the unknown noise, so the theorem proves that a separating threshold exists, not that a rule computed from the data finds it. Proposition 1 is the data-driven counterpart.

Theorem 2 (Oracle separation guarantee: exact recovery of the number of modes) *Let $\mathcal{H}_L(X) = \mathcal{H}_L(X^{\mathrm{sig}}) + \mathcal{H}_L(E)$, where $X^{\mathrm{sig}} = AS$ contains μ oscillatory modes, represented by 2μ distinct exponential atoms (conjugate pairs of poles) under the conditions of Theorem 1, and E is arbitrary. Write $\lambda_i = \sigma_i(\mathcal{H}_L(X^{\mathrm{sig}}))^2/K$ for the signal eigenvalues, so that $\lambda_{2\mu} > 0 = \lambda_{2\mu+1}$, and $\rho = \|\mathcal{H}_L(E)\|_2^2/K$ for the noise floor. If*

$$\lambda_{2\mu} > 4\rho, \tag{3}$$

then every threshold τ with

$$\rho \;<\; \tau^2/K \;<\; \big(\sqrt{\lambda_{2\mu}} - \sqrt{\rho}\big)^2 \tag{4}$$

returns $\widehat{d}_{\mathrm{osc}}(X; L, \tau) = 2\mu$. The interval (4) *is non-empty if and only if* (3) *holds.*

Proof. By Weyl's inequality, $|\sigma_i(\mathcal{H}_L(X)) - \sigma_i(\mathcal{H}_L(X^{\mathrm{sig}}))| \le \|\mathcal{H}_L(E)\|_2$ for every i. For $i \le 2\mu$ this gives $\sigma_i(\mathcal{H}_L(X)) \ge \sqrt{K}\,(\sqrt{\lambda_{2\mu}} - \sqrt{\rho}) > \tau$, and for $i > 2\mu$, where $\sigma_i(\mathcal{H}_L(X^{\mathrm{sig}})) = 0$, it gives $\sigma_i(\mathcal{H}_L(X)) \le \sqrt{K\rho} < \tau$. Exactly 2μ singular values exceed τ. The interval is non-empty iff $\sqrt{\lambda_{2\mu}} - \sqrt{\rho} > \sqrt{\rho}$, that is iff $\lambda_{2\mu} > 4\rho$. □

The theorem separates the two ingredients of the count: the signal eigenvalue $\lambda_{2\mu}$, which is a property of the window, and the noise floor ρ, which is a property of the noise and of the shape of the trajectory matrix.

Proposition 1 (Count with an estimated threshold) *Let $\widehat{t}$ be a threshold level computed from the data, the rule counting the eigenvalues $\ell_i = \sigma_i(\mathcal{H}_L(X))^2/K$ above $\widehat{t}$, and let $\lambda_{2\mu}$ and ρ be as in Theorem 2.*

Aicumene:2026.044v2 [cs.HC] November 11, 2026

(a) If $\widehat{t} > \rho$, then $\widehat{d}_{\mathrm{osc}} \le 2\mu$: no noise component is counted.

(b) If moreover $\lambda_{2\mu} > \left(\sqrt{\widehat{t}} + \sqrt{\rho}\right)^2$, then $\widehat{d}_{\mathrm{osc}} = 2\mu$.

Hence, for every threshold rule, with $\widehat{\mu} = \widehat{d}_{\mathrm{osc}}/2$,

$$\mathbb{P}\big(\widehat{\mu} = \mu\big) \;\ge\; \mathbb{P}\Big(\rho < \widehat{t} < \big(\sqrt{\lambda_{2\mu}} - \sqrt{\rho}\big)^2\Big) : \tag{5}$$

the count is exact whenever the random threshold falls in the oracle interval of Theorem 2, and the probability of that joint event, in which ρ, $\lambda_{2\mu}$ and $\widehat{t}$ are all random, is the guarantee of the rule.

Proof. By Weyl's inequality $\ell_i \le \rho < \widehat{t}$ for $i > 2\mu$, and under (b) $\ell_i \ge (\sqrt{\lambda_{2\mu}} - \sqrt{\rho})^2 > \widehat{t}$ for $i \le 2\mu$. The event in (5) is the conjunction of (a) and (b), on which $\widehat{d}_{\mathrm{osc}} = 2\mu$. □

The two rules of this section are instances. The event of (5) has two sides. On the lower side, for the surrogate rule on a recording of pure noise whose samples are exchangeable within every channel, the permutation distribution of σ_1 is the exact null distribution of the observed σ_1, so $\mathbb{P}(\widehat{t} > \rho) = 1 - \eta$ holds at the chosen level without any assumption on the distribution of the noise; with signal present the surrogates carry the signal variance as well and the rule is conservative on this side. For the calibrated rule the lower side is set empirically, by the quantile κ of Appendix A, at 95 %. The upper side, $\widehat{t} < (\sqrt{\lambda_{2\mu}} - \sqrt{\rho})^2$, is what the recording has to satisfy, and neither rule estimates $\lambda_{2\mu}$; Section 6.4 measures from which signal-to-noise ratio the count under the practical rule is exact, a frequency that (5) bounds from below by that of the joint event.

Lemma 1 (Noise floor of the block-Hankel matrix) *Let the entries of E be independent with mean zero, variance σ_ε^2 and fourth moment $\kappa_4\,\sigma_\varepsilon^4 < \infty$, and write $D = \mathcal{H}_L(E)\,\mathcal{H}_L(E)^\top/K - \sigma_\varepsilon^2 I_{mL}$.*

(a) $\mathbb{E}\,D = 0$: the scatter matrix of the noise has expectation $\sigma_\varepsilon^2 I_{mL}$ exactly, for every m, L and K.

(b) $\mathbb{E}\,\|D\|_F^2 = mL\,(mL + \kappa_4 - 2)\,\sigma_\varepsilon^4/K$ exactly. Consequently $|\rho - \sigma_\varepsilon^2| \le \|D\|_2 \le \|D\|_F$,

$$\mathbb{P}\big(|\rho - \sigma_\varepsilon^2| > t\big) \;\le\; \frac{mL\,(mL + \kappa_4 - 2)\,\sigma_\varepsilon^4}{K\,t^2},$$

and $\rho \to \sigma_\varepsilon^2$ in probability along every sequence of shapes with $(mL)^2/K \to 0$.

Proof. The entry of $\mathcal{H}_L(E)\,\mathcal{H}_L(E)^\top/K$ in row (j,a) and column (l,b) is $K^{-1}\sum_{t=1}^{K} e_j(t+a)\,e_l(t+b)$, whose expectation is σ_ε^2 if $(j,a) = (l,b)$ and zero otherwise; this is (a). For (b), take $(j,a) \ne (l,b)$. The K products $e_j(t+a)\,e_l(t+b)$ have mean zero, variance σ_ε^4, and are pairwise uncorrelated: for $t \ne t'$ the expectation of the product of the four samples $e_j(t+a)$, $e_l(t+b)$, $e_j(t'+a)$, $e_l(t'+b)$ vanishes unless every distinct sample among them occurs an even number of times, which requires $j = l$ and $\{t+a, t+b\} = \{t'+a, t'+b\}$, hence $t' - t = a - b = b - a$ and $a = b$, contradicting $(j,a) \ne (l,b)$. The off-diagonal entry of D therefore has variance σ_ε^4/K. A diagonal entry is the mean of K independent variables $e_j(t+a)^2 - \sigma_\varepsilon^2$ of variance $(\kappa_4 - 1)\sigma_\varepsilon^4$, so its variance is $(\kappa_4 - 1)\sigma_\varepsilon^4/K$. Summing over the $mL(mL-1)$ off-diagonal and mL diagonal entries gives the identity. Since $\mathcal{H}_L(E)\,\mathcal{H}_L(E)^\top/K = \sigma_\varepsilon^2 I + D$, Weyl's inequality gives $|\rho - \sigma_\varepsilon^2| \le \|D\|_2$, the Frobenius norm dominates the spectral norm, and Markov's inequality applied to $\|D\|_F^2$ gives the tail bound. □

Remark (the proportional regime). The lemma is an asymptotic statement for $\beta = mL/K \to 0$, and its bound is loose in the shapes used in practice, where $(mL)^2/K$ is of order 0.1–1; its role is the regime it identifies. In the proportional regime, β fixed, the floor of a matrix with *independent* entries converges to the Marchenko–Pastur edge $\sigma_\varepsilon^2(1 + \sqrt{\beta})^2$ [24, 32]. The entries of $\mathcal{H}_L(E)$ are not independent, each sample being repeated along an anti-diagonal, and the limiting spectral law of square Hankel matrices with independent generating sequence is known to differ from the laws of matrices with independent entries [26]; for the rectangular trajectory matrix at fixed β we know of no proof. The edge is therefore used in this paper as an *empirical calibration* of the floor, not as a theorem, and Section 6.4 measures how far the floor of the block-Hankel matrix lies from it for the shapes we use.

Lemma 2 (Signal eigenvalues of resolved undamped modes) *Let $X^{\mathrm{sig}} = AS$ with μ undamped modes $s_k(t) = \sin(\omega_k t + \phi_k)$, $0 < \omega_k < \pi$, non-zero loadings a_k, and let $\Delta\omega > 0$ be the smallest circular distance between distinct points of $\{\pm\omega_1, \dots, \pm\omega_\mu\}$, which separates every pole from its conjugate as well. Put*

$$\varepsilon_L = (2\mu - 1)\,\frac{\max_k \|a_k\|^2}{\min_k \|a_k\|^2}\,\frac{\pi}{L\,\Delta\omega}, \qquad \varepsilon_K = (2\mu - 1)\,\frac{\pi}{K\,\Delta\omega}.$$

Aicumene:2026.044v2 [cs.HC] November 11, 2026

If $\varepsilon_L < 1$ and $\varepsilon_K < 1$, then

$$\lambda_{2\mu} \;\geq\; \frac{L}{4}\,\min_k \|a_k\|^2\,(1-\varepsilon_L)\,(1-\varepsilon_K).$$

Proof. Write $s_k(t) = (e^{i\phi_k} z_k^t - e^{-i\phi_k}\bar{z}_k^t)/(2i)$ with $z_k = e^{i\omega_k}$. Then $\mathcal{H}_L(X^{\mathrm{sig}}) = BDC$, where the 2μ columns of $B \in \mathbb{C}^{mL\times 2\mu}$ are $a_k \otimes v_L(z_k)$ and $a_k \otimes v_L(\bar{z}_k)$ with $v_L(z) = (1, z, \ldots, z^{L-1})^\top$, D is diagonal with entries of modulus $\frac{1}{2}$, and the rows of $C \in \mathbb{C}^{2\mu\times K}$ are $(z, z^2, \ldots, z^K)$ for the same 2μ poles. Since B has full column rank and C full row rank, $\sigma_{2\mu}(BDC) \geq \frac{1}{2}\,\sigma_{\min}(B)\,\sigma_{\min}(C)$, so $\lambda_{2\mu} = \sigma_{2\mu}^2/K \geq \frac{1}{4}\,\sigma_{\min}(B^{\mathsf{H}}B)\,\sigma_{\min}(CC^{\mathsf{H}})/K$. The Gram matrix $B^{\mathsf{H}}B$ has diagonal entries $\|a_k\|^2 L$ and off-diagonal entries $(a_k^\top a_l)\,v_L(z)^{\mathsf{H}} v_L(z')$ with $z \neq z'$ on the unit circle, where $|v_L(z)^{\mathsf{H}} v_L(z')| = |\sin(L\delta/2)/\sin(\delta/2)| \leq 1/\sin(\Delta\omega/2) \leq \pi/\Delta\omega$ for the circular distance $\delta \geq \Delta\omega$ of the two poles, and $|a_k^\top a_l| \leq \max_k \|a_k\|^2$. Gershgorin's theorem gives $\sigma_{\min}(B^{\mathsf{H}}B) \geq \min_k \|a_k\|^2 L - (2\mu-1)\max_k\|a_k\|^2\,\pi/\Delta\omega = \min_k \|a_k\|^2\,L\,(1-\varepsilon_L)$. Likewise CC^{H} has diagonal entries K and off-diagonal entries of modulus at most $\pi/\Delta\omega$, so $\sigma_{\min}(CC^{\mathsf{H}})/K \geq 1-\varepsilon_K$. □

The bound certifies the linear growth of $\lambda_{2\mu}$ in L with the constant $\min_k\|a_k\|^2/4$; it is loose at short windows, where Section 6.4 measures the approximation $\lambda_{2\mu} \approx \min_k\|a_k\|^2 L/4$ directly.

Corollary 2 (Window length and asymptotic oracle recoverability) *Let the modes be undamped with non-zero loadings and frequency separation $\Delta\omega > 0$ as in Lemma 2.*

(i) (Deterministic.) For the realized floor ρ, condition (3) *holds once $\frac{L}{4}\min_k\|a_k\|^2(1-\varepsilon_L)(1-\varepsilon_K) > 4\rho$; with the approximation $\lambda_{2\mu} \approx \min_k\|a_k\|^2 L/4$ of resolved modes this reads*

$$L \;>\; \frac{16\,\rho}{\min_k \|a_k\|^2}, \tag{6}$$

sixteen samples per unit of the signal-to-noise ratio of the weakest mode, the ratio being its loading energy summed over channels to the noise variance of one channel.

(ii) (Asymptotic oracle recoverability.) Fix μ, the loadings and the frequencies, and let the window and the record grow with $L \to \infty$ and $(mL)^2/K \to 0$. By Lemma 2, $\lambda_{2\mu} \geq \frac{L}{4}\min_k\|a_k\|^2\,(1-o(1))$; by Lemma 1, $\rho \to \sigma_\varepsilon^2$ in probability. Condition (3) *therefore holds with probability tending to one: the oracle threshold of Theorem 2 exists, and returns the exact count, with probability tending to one. The regime is explicit, the record growing faster than the square of the window, $K \gg (mL)^2$, and the statement concerns the oracle; Proposition 1 transfers it to a data-driven rule whose threshold clears the floor.*

(iii) (Fixed window.) At a fixed L and $K \to \infty$ the floor tends to σ_ε^2, and $\lambda_{2\mu}$ converges: in the factorization of the proof of Lemma 2, $\mathcal{H}_L(X^{\mathrm{sig}})\mathcal{H}_L(X^{\mathrm{sig}})^\top/K = BD\,(CC^{\mathsf{H}}/K)\,D^{\mathsf{H}}B^{\mathsf{H}}$ and $CC^{\mathsf{H}}/K \to I$, its off-diagonal entries being bounded by $\pi/(K\Delta\omega)$, so $\lambda_{2\mu} \to \lambda_{2\mu}^\infty = \frac{1}{4}\,\sigma_{\min}(B^{\mathsf{H}}B)$, a quantity of the window and the loadings alone, which lies between $\frac{L}{4}\min_k\|a_k\|^2(1-\varepsilon_L)$ and $\frac{L}{4}\max_k\|a_k\|^2(1+\varepsilon_L)$. Condition (3) *therefore holds in the limit if and only if $\lambda_{2\mu}^\infty > 4\sigma_\varepsilon^2$, which for resolved modes is $L > 16\,\sigma_\varepsilon^2/\min_k\|a_k\|^2$: lengthening the record cannot replace a window that is too short for the sufficient condition.*

The condition is sufficient, not necessary; with the calibrated threshold the count is already correct once $\lambda_{2\mu}$ exceeds the floor $\rho(1+\kappa)$, about a factor of four earlier (Table 2).

Thus, the first inverse problem is solved structurally:

$$X \xrightarrow{\text{MSSA}} \widehat{d}_{\mathrm{osc}} \longrightarrow \widehat{\mu}.$$

5 Spatial loadings of the modes

5.1 The loading of a mode

After MSSA separates the latent modes, each recovered oscillatory source is associated with a spatial loading vector $\widehat{a}_k = (\widehat{a}_{1k}, \ldots, \widehat{a}_{mk})^\top$, the column of the mixing matrix A that belongs to it. When the sources are observed through a known forward model — an attenuation law or the gain matrix of a volume conductor — the loading is a fingerprint of the position of the source, and matching it to the model recovers that position; this is the second inverse problem, treated in the companion paper [2]. Here the loading is defined and recovered. The loading has to be defined with some

Aicumene:2026.044v2 [cs.HC] November 11, 2026

care, because the left singular vectors do not provide it directly. By Theorem 1 the leading singular vectors span the same subspace as the vectors $c_i \otimes v_L(z_i)$, but each of them is a combination of these vectors, and when two sources have similar strength the combination mixes them: this is the failure of strong separability known in singular spectrum analysis [28]. The subspace, however, determines the poles, and the poles determine the loadings.

Let $U \in \mathbb{R}^{mL\times 2r}$ hold the leading left singular vectors. Let $U^{\uparrow}$ be U without the last row of every channel block and $U^{\downarrow}$ be U without the first row of every channel block.

Theorem 3 (Recovery of the spatial loadings) *Let $X = AS$ with sources $s_k(t) = e^{\lambda_k t}\sin(\omega_k t + \phi_k)$ whose poles $z_k = e^{\lambda_k + i\omega_k}$, $0 < \omega_k < \pi$, are pairwise distinct, and let no column of A vanish. If $L \geq 2r+1$ and $K \geq 2r$, then*

(i) the eigenvalues of $\Phi = (U^{\uparrow})^{+}U^{\downarrow}$ are the $2r$ poles z_k and $\bar{z}_k$;

(ii) the amplitudes $C \in \mathbb{C}^{m\times 2r}$ in $x_j(t) = \sum_i C_{ji} z_i^t$ are the unique solution of the linear system $X^{\top} = V_N(z)\, C^{\top}$;

(iii) the column c_k of C that belongs to z_k equals $\frac{1}{2}e^{i\theta_k}a_k$ with $\theta_k = \phi_k - \pi/2$, so that

$$a_k = \pm\, 2\,\mathrm{Re}\big(e^{-i\theta_k}c_k\big), \qquad \theta_k = \tfrac{1}{2}\arg\big(c_k^{\top}c_k\big) \pmod{\pi}. \tag{7}$$

The loading vector of every source is thus recovered exactly up to sign, together with its waveform $s_k(t) = \pm\mathrm{Re}(e^{i\theta_k}z_k^t)$. For a positive attenuation law the sign is fixed by requiring positive entries.

Proof. (i) By the factorization in the proof of Theorem 1, the columns of $W = [\, c_i \otimes v_L(z_i)\,]_{i=1}^{2r}$ span the column space of $\mathcal{H}_L(X)$, so $U = WT$ with T invertible. Deleting the last row of every block gives $W^{\uparrow} = [\, c_i \otimes v_{L-1}(z_i)\,]$, and deleting the first gives $W^{\downarrow} = W^{\uparrow}D$ with $D = \mathrm{diag}(z_i)$. The matrix $W^{\uparrow}$ has full column rank because $L - 1 \geq 2r$, by the argument of Theorem 1. Hence $U^{\downarrow} = U^{\uparrow}T^{-1}DT$, the solution $\Phi = T^{-1}DT$ is unique, and its eigenvalues are the z_i. (ii) $V_N(z)$ has full column rank since $N \geq 2r$. (iii) $s_k(t) = \mathrm{Im}(e^{i\phi_k}z_k^t) = (e^{i\phi_k}z_k^t - e^{-i\phi_k}\bar{z}_k^t)/(2i)$, so the coefficient of z_k^t in channel j is $a_{jk}\, e^{i\phi_k}/(2i)$; then $c_k^{\top}c_k = \frac{1}{4}e^{2i\theta_k}\|a_k\|^2$. □

With noise, the same three steps are applied to the estimated subspace: this is the ESPRIT estimator [20, 29] applied to the block-Hankel matrix, followed by linear least squares. Once the poles are fixed the loading is linear in the data, with

$$\mathrm{Var}(\widehat{a}_{jk}) \approx 2\,\sigma_{\varepsilon}^2\, v_k, \qquad v_k = \big[(V_N^{\mathsf{H}}V_N)^{-1}\big]_{kk},$$

and $v_k \approx 1/N$ for an undamped mode, which is the bound for the amplitude of a sinusoid of unknown frequency [30].

Two properties of this definition are used below. It does not require the modes to be orthogonal or of different strength. And it yields a signed vector: for a positive attenuation law all entries have one sign, while for the gain matrix of a volume conductor the signs carry information.

5.2 Sources that share a mode

Proposition 2 *Let sources k and l share a pole, and let no other source have it. The loading of their common mode is $c = \frac{1}{2}\,(e^{i\theta_k}a_k + e^{i\theta_l}a_l)$. If a_k and a_l are not collinear and $\theta_k \not\equiv \theta_l \pmod{\pi}$, the real and imaginary parts of c are linearly independent and span the plane of a_k and a_l. If $\theta_k \equiv \theta_l \pmod{\pi}$, then c is proportional to the real vector $a_k \pm a_l$, and the two sources cannot be distinguished from one source with that loading.*

Proof. $[\mathrm{Re}\, c,\ \mathrm{Im}\, c] = \frac{1}{2}\,[a_k,\ a_l]\,R$ with $R = \begin{bmatrix}\cos\theta_k & \sin\theta_k \\ \cos\theta_l & \sin\theta_l\end{bmatrix}$ and $\det R = \sin(\theta_l - \theta_k)$. □

A single source has an in-phase loading: after removal of the common phase by (7), the quadrature part $b = 2\,\mathrm{Im}(e^{-i\theta}c)$ vanishes. Under noise, the statistic

$$T = \frac{\|\widehat{b}\|^2}{2\,v\,\widehat{\sigma}_{\varepsilon}^2}$$

is distributed approximately as χ^2 with $m-1$ degrees of freedom when one source underlies the mode, the phase having been fitted. The law is an approximation: it treats the estimated complex loading as the true loading plus isotropic Gaussian noise of variance $2\sigma_{\varepsilon}^2 v$ per real coordinate, spends one degree of freedom on the fitted phase, and neglects the error of the estimated pole and the correlation that the least-squares fit at a common pole introduces between channels. Its level was measured (script `v2_inphase_level.py`): for one in-phase source on the array of Section 6.1, 1000 trials per cell, the test at the nominal level $5\,\%$ rejects in 4.4, 6.1 and $6.4\,\%$ of trials at $\mathrm{SNR}_0 = 0$, 10 and 20 dB with

Aicumene:2026.044v2 [cs.HC] November 11, 2026

the pole estimated by ESPRIT, and in 4.5, 6.2 and $6.4\,\%$ with the true pole, while two sources in quadrature are rejected in 99.3–$100\,\%$ of trials. The test is approximate, with a level close to nominal in this setting, and it is used as such: a mode whose T exceeds the quantile of level $1-\eta$ is not a single in-phase source. Two sources in one mode can be separated in space by a scan of the plane of $\operatorname{Re} c$ and $\operatorname{Im} c$, which the companion paper [2] carries out with a forward model. Without additional structural assumptions, the individual loadings of three or more sources sharing one pole are not identifiable from the common complex loading alone, which observes only a two-dimensional subspace of their span; a forward model, sparsity or another prior may restore identifiability case by case. A propagation delay that differs between sensors would also produce a quadrature part; the model considered here has none, like the quasi-static approximation of electrophysiology.

6 Numerical experiments

The number of sources was evaluated on recordings of a non-invasive brain-computer interface in [3, 4], and the Hankel-embedded pipeline on EEG classification in [9]. The number of sources is not known in such recordings, so the experiments of this section are simulations; Figs. 1 and 2 show the objects they count on synthetic series and on real EEG.

6.1 Setting

Unless stated otherwise, nine sensors form a 3×3 grid on the unit square, so that lengths are in units of the side of the array. The sampling rate is 100 Hz, $N = 500$ (five seconds) and $L = 40$. Sources are undamped sinusoids of random phase at random positions in $[0.1, 0.9]^2$, at least 0.12 from every sensor, and reach the sensors with the amplitude $g(\rho) = \rho^{-1}$ of the distance ρ — the law is immaterial for the count and only fixes the loadings. Sensor noise is white, Gaussian and of equal variance σ_ε^2 in all channels. The signal-to-noise ratio is defined at the reference distance $\rho_0 = 0.5$,

$$\mathrm{SNR}_0 = \frac{\alpha^2\, g(\rho_0)^2}{2\,\sigma_\varepsilon^2},$$

the ratio that a source of amplitude α has at a sensor half an array side away. Each point is the result of 300 trials unless noted.

6.2 The exact statements

Without noise the statements of Sections 3 and 5 hold to machine precision.

Rank. Three sources at 8, 11 and 15 Hz give rank 6. When two of them have 8 Hz with different phases the rank is 4, although $\operatorname{rank} S = 3$; when they have 8 Hz with different damping the rank is 6 again (Corollary 1).

Loadings. For two sources at 10 and 13 Hz, the loadings of Theorem 3 coincide with the true columns of A to $1 - \cos < 5 \times 10^{-16}$ at every ratio of source strengths. The energy profiles of pairs of singular vectors deviate by $1 - \cos = 0.05$ (median) at equal strength, by 0.017 at a ratio of 1.05 and by 2×10^{-4} at a ratio of 1.5.

Shared mode. For two sources at 10 Hz in quadrature, the plane spanned by the real and imaginary parts of the loading contains both columns of A to 10^{-15} (Proposition 2).

6.3 Rules for the number of components

Three sources with frequencies between 5 and 30 Hz at least 2 Hz apart, amplitudes between 0.7 and 1.3 and mutual distances above 0.2 were processed with the calibrated threshold ($\kappa = 0.135$); the rank is correct in 97–98 % of recordings from -15 dB to 20 dB, in 19 % at -20 dB and never at -25 dB. Table 1 compares the thresholds of Section 4 on the same kind of recording. The uncorrected Marchenko–Pastur edge over-counts in a quarter to a third of recordings. The surrogate threshold is correct in every trial from -5 dB, and less often than the calibrated threshold at -15 dB.

The surrogate threshold fails when sources differ in strength. For two sources at 10 and 17 Hz, the weaker having a mean SNR of -6 dB per sensor, the calibrated threshold returns rank 4 in 96–99 % of 100 trials at each amplitude ratio from 1 to 100. The surrogate threshold does so at ratios of 1 and 3 and returns rank 2 in every trial at ratios of 10, 30 and 100: the surrogate spectrum inherits the power of the strong source and hides the weak one.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

Table 1: Three sources: probability of the correct rank (100 trials per row; 30 surrogates).

SNR_0 (dB)	calibrated threshold	edge without correction	surrogates
-15	0.96	0.62	0.82
-5	0.98	0.67	1.00
5	0.96	0.64	1.00
20	0.99	0.75	1.00

6.4 Guarantees against practice

The additions of version 2 were checked with the script `v2_checks.py`. *Noise floor.* For white noise with $\sigma_\varepsilon = 1$, $m = 8$, $L = 20$ and $K = 250$, 1000, 4000, 16 000 ($\beta = 0.64$, 0.16, 0.04, 0.01), the mean of ρ over twenty trajectory matrices is 3.10, 1.97, 1.44 and 1.21 against the Marchenko–Pastur edge 3.24, 1.96, 1.44 and 1.21: the edge for independent entries predicts the floor of the block-Hankel matrix to 1 % once $\beta \le 0.16$, and lies 4 % below it at $\beta = 0.64$. This is the empirical calibration referred to in the remark after Lemma 1: a measurement for these shapes, not a bound. *Signal eigenvalues.* For two unit-loading sources at 10 and 13 Hz sampled at 100 Hz, the ratio $\lambda_{2\mu}/(L/4)$ is 0.63, 0.82, 0.94 and 0.96 for $L = 10$, 20, 40, 80: the approximation of Corollary 2 holds once the window resolves the two frequencies. The bound of Lemma 2 is loose at these windows: with $\Delta\omega = 0.19$ rad and $\mu = 2$ it is vacuous for $L \le 40$ and gives 0.37 of $L/4$ at $L = 80$, against the measured 0.96; it certifies the linear growth, not the constant. *Exact count.* Table 2 reports, for the same sources with eight channels and $K = 2000$, how often the sufficient condition (3) holds, how often the count is exact when it holds, and how often the practical rule with the floor $\rho(1 + \kappa)$, $\kappa = 0.15$, is correct.

Table 2: The exact count against practice: two unit-loading sources at 10 and 13 Hz, eight channels, $K = 2000$, 200 trials per row, 95 % Wilson intervals in parentheses (per cent). "Weyl" is the share of trials in which the sufficient condition (3) holds, "exact" the share of those in which the midpoint of the interval (4) returns the count 4, and "practical" the share of all trials in which the rule $\ell_i > \sigma_\varepsilon^2(1 + \sqrt{\beta})^2(1 + \kappa)$ with the true σ_ε^2 and $\kappa = 0.15$ returns 4. ρ is the mean measured floor, with the Marchenko–Pastur edge in parentheses; λ_4 is the mean fourth signal eigenvalue; SNR is the loading energy of a source over the noise variance of one channel. Where the condition held in a single trial (row $L = 160$, -5 dB) the interval for "exact" is uninformative. Produced by `v2_table_ci.py`.

L	SNR, dB	β	ρ (edge)	λ_4	Weyl	exact	practical
20	-5	0.08	5.15 (5.20)	4.02	0/200 (0–2)	–	199/200 (97–100)
20	0	0.08	1.63 (1.65)	4.07	0/200 (0–2)	–	200/200 (98–100)
20	5	0.08	0.52 (0.52)	3.98	200/200 (98–100)	200/200 (98–100)	200/200 (98–100)
20	10	0.08	0.16 (0.16)	4.03	200/200 (98–100)	200/200 (98–100)	200/200 (98–100)
40	-5	0.16	6.21 (6.20)	9.36	0/200 (0–2)	–	200/200 (98–100)
40	0	0.16	1.96 (1.96)	9.40	200/200 (98–100)	200/200 (98–100)	200/200 (98–100)
40	5	0.16	0.62 (0.62)	9.40	200/200 (98–100)	200/200 (98–100)	199/200 (97–100)
40	10	0.16	0.20 (0.20)	9.39	200/200 (98–100)	200/200 (98–100)	198/200 (96–100)
80	-5	0.32	7.97 (7.75)	19.16	0/200 (0–2)	–	199/200 (97–100)
80	0	0.32	2.51 (2.45)	19.17	200/200 (98–100)	200/200 (98–100)	200/200 (98–100)
80	5	0.32	0.79 (0.78)	19.17	200/200 (98–100)	200/200 (98–100)	197/200 (96–100)
80	10	0.32	0.25 (0.25)	19.24	200/200 (98–100)	200/200 (98–100)	195/200 (94–99)
160	-5	0.64	10.84 (10.25)	39.27	1/200 (0–3)	1/1 (21–100)	190/200 (91–97)
160	0	0.64	3.44 (3.24)	39.25	200/200 (98–100)	200/200 (98–100)	188/200 (90–96)
160	5	0.64	1.09 (1.02)	39.31	200/200 (98–100)	200/200 (98–100)	188/200 (90–96)
160	10	0.64	0.34 (0.32)	39.26	200/200 (98–100)	200/200 (98–100)	182/200 (86–94)

The sufficient condition holds in every one of 200 trials from 0 dB for $L \ge 40$ and from 5 dB for $L = 20$, and whenever it holds the count is exact, in 2400 of 2400 trials; the lower confidence limit of each of these shares is 98 %. The practical rule is right already at -5 dB, where the condition fails: this is the factor of four of Corollary 2. At $L \ge 80$ the fixed $\kappa = 0.15$ lies below the calibrated κ_{95} of Appendix A (Table 3), the floor fluctuates above the edge by more than that, and the rule over-counts in 3–9 % of trials at $L = 160$, a share the 200 trials now resolve; the calibration of the appendix repairs this, and the interval of Theorem 2 needs no κ at all. In the 25-trial version of this table (version 2, first build) these shares were estimated as 84–96 % with intervals too wide to separate them from 100 %.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

6.5 Spatially correlated noise

The noise of the experiments above is independent between channels. A background that is correlated across sensors, as in EEG, changes the count, and whitening restores it; the experiment runs the complete pipeline — calibrated count, ESPRIT poles, least-squares loading at the estimated frequency, and the localization of the companion paper — for one source on the grid of nine sensors at $\mathrm{SNR}_0 = 10$ dB, 300 trials per setting (`v2_robustness.py`).

Spatially correlated noise. The sensor noise is white in time with covariance $\Sigma_{ij} = \exp(-\|p_i - p_j\|/l_c)$ across channels, of equal variance. The count with the calibrated threshold, which assumes an isotropic floor, is right in $98\,\%$ of trials at $l_c = 0$, in $75\,\%$ at $l_c = 0.25$ and in none at $l_c \geq 0.5$: the correlated part of the noise is a spatial component of its own, with an eigenvalue above the isotropic edge, and the rule counts too many modes. Whitening the recording with the noise covariance, $\widetilde{X} = \Sigma^{-1/2} X$, restores the count to 96–97 $\%$ at every l_c. Whitening changes the loadings as well, $\widetilde{a}_k = \Sigma^{-1/2} a_k$, and the experiment keeps track of this: the count and the poles are taken from $\widetilde{X}$; the loading is then estimated by least squares from the original recording at the estimated frequency, so that it stays in the sensor coordinates; and the matching of the companion paper whitens both the loading and the forward model, $\Sigma^{-1/2}\widehat{a}_k$ against $\Sigma^{-1/2} g_q$, which is generalized least squares with the known covariance and is equivalent to matching $\widehat{\widetilde{a}}_k$ to the transformed model and recovering $\widehat{a}_k = \Sigma^{1/2}\,\widehat{\widetilde{a}}_k$. The position is then unaffected: 0.004 with or without whitening at every l_c, with no gross error. A spatially correlated background is therefore a problem of the first inverse problem, not of the second, and it is solved by the whitening that Appendix A prescribes for a coloured background in time; in practice Σ is estimated from a segment without the sources, or from the residual after the modes are removed.

7 Applications

7.1 Neurophysiology and EEG/MEG

In EEG and MEG, latent neural populations generate oscillatory activity recorded by multiple sensors. The count of Section 4 is the Bernadotte Oscillator Intrinsic Dimension (BOID), introduced in [8] as the brain oscillator intrinsic dimension to separate states of consciousness and outcomes of coma, and used by the patent [40] as a measure of the complexity of the brain signal; its uses include sleep-stage analysis, the detection of pathological oscillatory sources, BCI calibration and neurofeedback personalization. The loadings of Section 5 are the input of the localization of the companion paper [2] and of the fitting of dipoles to component maps [21].

The count and the loadings also serve closed-loop protocols, among them the intelligent-feedback stimulation of chosen brain regions for rehabilitation and the prevention of neurodegeneration [38] and the activation of a preserved brain pattern [41], both of which need a readout of the state of a region and a loading to address it. In neurofeedback and non-invasive stimulation for attention-deficit disorder [43] the feedback variable is a band power; MSSA replaces it by the energy and the loading of an identified mode, so that the loop follows one oscillatory mode rather than a band that several sources share; the mode is a single physical source only when its loading passes the in-phase test of Proposition 2 and the identifiability conditions of the companion paper hold. Downstream of the count, the modes and their loadings are the features of the brain–computer-interface pipeline: silent-speech recognition from EEG with convolutional and recurrent networks [44], the choice of a command dictionary by a statistical proximity criterion [45] and the algorithm that maximizes the accuracy of k-classification over representatives of k classes [46, 47], the quality criterion for clustering a tagged sample [48], and the detection of frequency, time and spatial changes of the EEG after COVID-19 [49, 50]; in each of these the number of oscillatory sources decides how many features are real.

7.2 Beyond neurophysiology

The count applies to any multichannel series that is a mixture of a few exponentially modulated oscillations: the oscillations of the vessel wall and of the heartbeat in the force and position channels of a surgical robot, the vibration modes of a machine recorded by several accelerometers, the modes of a structure or of the ground recorded by a seismic array, the carriers and interferers seen by an antenna array, and the cycles of a set of economic indicators. In every case the number of modes is decided before any model of how they reach the sensors is assumed, and the loadings are available whether or not such a model exists; a coordinate for a mode requires the model, and the conditions under which it is identifiable are the subject of the companion paper [2].

Aicumene:2026.044v2 [cs.HC] November 11, 2026

8 Discussion

This article formulates MSSA as a method for the first inverse problem of multichannel oscillatory systems, the number of latent oscillatory modes, and supplies the guarantees under which the count is exact. The problem was previously addressed in Bernadotte's Hankel-embedding framework [3, 4], where the effective rank of time-series unfoldings was used to estimate the number of active oscillators; the block-Hankel embedding writes it for several channels at once, and Appendix B measures what the joint analysis gains.

The count rests on three statements of different kinds, and the paper keeps them apart. Theorem 2 is deterministic and distribution-free: it says when a separating threshold exists, in terms of the weakest signal eigenvalue and the realized noise floor. Proposition 1 is about a rule: a threshold computed from the data returns the exact count whenever it clears the floor and the weakest mode clears the threshold, and the surrogate rule clears the floor at its nominal level on pure noise. Lemma 1 is asymptotic: the floor tends to the noise variance when the record grows faster than the square of the window, and in the proportional regime the Marchenko–Pastur edge is a calibration whose accuracy is measured, not proved. A claim about a given recording is only the first of these, and the numbers of Section 6.4 say from which signal-to-noise ratio its hypothesis holds.

The count $\widehat{\mu}$ is the Bernadotte Oscillator Intrinsic Dimension (BOID), introduced in [8] as the brain oscillator intrinsic dimension to separate states of consciousness and outcomes of coma, and reproduced by learned networks as a regulated switching of regimes [42]. The guarantee behind that use is that the sufficient condition for an exact count is set by the window length against the noise floor, the record enters only through the floor, and the count measures the dimensionality of the oscillatory state, the number of active oscillatory degrees of freedom, rather than any particular condition; a contrast between states or groups probes whether that dimensionality is state-dependent, and a null result is then a statement that it is conserved. The three numbers of Section 4 are the vocabulary for such statements: the Hankel dimension is what the threshold returns, the number of modes μ, the BOID, is what the poles resolve, and the number of generators is what a forward model, in the companion paper, can attribute.

The scope of the results is as follows. All quantitative experiments are simulations with white sensor noise of equal variance, and the EEG trials of Fig. 2 are a demonstration of the geometry and of the count without a known number of generators; a coloured background requires whitening in time before the count (Appendix A) and a spatially correlated one whitening across channels (Section 6.5), and the EEG trials of Fig. 2 are counted against a surrogate floor that is conservative. The consistency of the count is proved in the regime $(mL)^2/K \to 0$ and calibrated empirically in the shapes of practice. The loadings are exact without noise and linear in the data once the poles are fixed; their use as fingerprints of position, the conditions under which a position is identifiable, and the accuracy of the recovered position are in the companion paper [2]. Validation of the count on recordings with a known number of generators is the next step.

A Calibration of the rank threshold

In [3, 4] the number of components is read from the effective rank, or ε-rank, of the unfolding. The threshold of Section 4 fixes that rank against a white noise floor. It was studied on white-noise recordings, for which the estimated rank should be zero. With $\kappa = 0$ a non-zero rank is returned for 16–29 % of trajectory matrices at $L = 10$, 37–52 % at $L = 40$ and 64–80 % at $L = 80$, against 13–23 % for matrices of the same shape with independent entries. The correction that restores a 5 % rate grows with the window, falls with the record length and falls slowly with the number of channels (Table 3); over 29 configurations it is described by

$$\kappa_{95} \approx 0.44\, L^{0.71}\, K^{-0.52}\, m^{-0.22}$$

with $R^2 = 0.97$ on a logarithmic scale. Simulation for the configuration at hand is the exact method; for $m = 9$, $N = 500$, $L = 40$ it gives 0.12–0.14.

Table 3: Correction κ_{95} for a 5 % false-positive rate on white noise.

	$L = 10$	$L = 20$	$L = 40$	$L = 80$
$N = 1000$, $m = 1$	0.055	0.102	0.171	0.231
$N = 1000$, $m = 8$	0.038	0.075	0.112	0.178
$N = 4000$, $m = 1$	0.027	0.042	0.072	0.125
$N = 4000$, $m = 8$	0.015	0.031	0.056	0.080

With the calibrated correction the rank of three oscillators shared by all channels is recovered in 93–99 % of trials from -5 dB per channel upward for one channel, from -7.5 dB for two and from -12.5 dB for eight. Coloured noise breaks

Aicumene:2026.044v2 [cs.HC] November 11, 2026

the calibration: first-order autoregressive noise with coefficient 0.3 already yields a mean rank of 11 on noise alone. Whitening every channel by a running median of its periodogram, which ignores narrow peaks, restores the null for coefficients up to 0.9 at the price of a larger correction; steeper backgrounds call for a null that carries their colour [27].

Why colour breaks the calibration. For stationary noise with spectral density $S_e(\omega)$ the lag covariance of a channel is the Toeplitz matrix of S_e, whose eigenvalues lie between $\min_\omega S_e$ and $\max_\omega S_e$ and whose largest eigenvalue tends to $\max_\omega S_e(\omega)$ as the window grows [33, 34]. The white calibration places the floor at the mean of S_e, which is the variance; the actual floor is the peak. For first-order autoregressive noise with coefficient φ the ratio of peak to mean is $(1+\varphi)/(1-\varphi)$, 1.9 at $\varphi = 0.3$ and 19 at $\varphi = 0.9$, so the noise alone exceeds the white threshold and the count reports the colour as components, which is what was observed above. Whitening flattens S_e and restores the white floor; equivalently, the floor of Theorem 2 can be set at $\rho = \max_\omega \widehat{S}_e(\omega)$ from a smoothed periodogram, which is the colour-carrying null of [27] in the present notation.

B Joint and channel-wise analysis

In [4] every trace is embedded separately. Section 2.3 states that joint decomposition reveals components that single channels do not. Table 4 gives the size of the effect and its limit, for eight channels, $N = 1000$, $L = 50$ and white noise, with channel-wise SSA and joint MSSA applied to the same recordings and the rank estimated by the calibrated threshold. The gain belongs to shared structure: it is close to its upper bound of $10 \log_{10} m$ dB when all channels carry the same oscillators, small when only part of the structure is shared, and negative when the channels have nothing in common, in agreement with [15]. In the localization problem all sensors record every source, which is the favourable case.

Table 4: Channel-wise SSA and joint MSSA on the same eight-channel recordings.

	channel-wise SSA	joint MSSA
output SNR at 0 dB input, three oscillators shared by all channels	10.7 dB	17.9 dB
the same, two shared oscillators and one private to each channel	10.7 dB	12.6 dB
the same, three private oscillators in each channel	10.7 dB	8.8 dB
detection of an oscillation common to all channels, -17.5 dB	11 %	96 %
detection of an oscillation present in one channel, -12.5 dB	89 %	51 %

Code and data availability

The checks of Section 6.4 are reproduced by `v2_checks.py` and `v2_table_ci.py`, the spatially correlated noise experiment of Section 6.5 by `v2_robustness.py`, Fig. 1 by `fig_geometry.py`, Fig. 2 by `fig_uci_geometry.py` and its window-length check by `fig_uci_window_check.py`, and the level of the in-phase test by `v2_inphase_level.py`; all are included with the source of this version, together with their outputs, need only NumPy and Matplotlib, and are kept, with the experiments of version 1 [1] and the `aicumene_mssa` package, in the repository `https://github.com/aicumene/MSSA`. The embedding, the decomposition and the rank estimate use the `aicumene-mssa` package of that repository; the pole estimation, the loadings and the tests are in `paper/experiments/loc_common.py` of the same repository, together with the scripts and numerical results of Section 6. The EEG trials of Fig. 2 are from the public UCI EEG database [35]; the script reads them from the archive `smni_eeg_data.tar.gz` of that database. No other experimental data were used.

REFERENCES

[1] Bernadotte A., Menshikov I. "Multichannel Singular Spectrum Analysis for Estimating Oscillatory Sources and Localizing Attenuating Oscillations." Aicumene Preprint 2026.044v1. August 12, 2026. doi:10.13140/RG.2.2.26719.01442. `https://research.aicumene.com/papers/MSSA/`

[2] Bernadotte A., Menshikov I. "From Oscillatory Modes to Source Geometry: Identifiability and Localization in Multichannel Time Series." Aicumene Preprint 2026.058v1; SSRN working paper 7716058 (November 16, 2026). `https://ssrn.com/abstract=7716058`, doi:10.2139/ssrn.7716058. `https://research.aicumene.com/papers/MSSA/`

[3] Alexandra Bernadotte, Victor Buchstaber. "Method for Evaluating the Number of Signal Sources and Application to Non-invasive Brain-computer Interface." arXiv preprint arXiv:2410.11844. 2024. `https://arxiv.org/abs/2410.11844`

Aicumene:2026.044v2 [cs.HC] November 11, 2026

[4] Alexandra Bernadotte. "Estimating the Number of Sources in EEG with Hankel Embedding for Brain-computer Interface." 12th International Conference on Automation, Robotics, and Applications. IEEE. 2026. pp 621-626. doi:10.1109/ICARA69401.2026.11480398

[5] V. M. Buchstaber. "Time Series Analysis and Grassmannians." Applied problems of Radon transform, Amer. Math. Soc. Transl. 1994. Vol 2 (162), pages 1–17. doi:10.1090/trans2/162/01.

[6] V. M. Buchstaber, V. K. Maslov. "Factor Analysis and Extremal Problems on Grassmann Manifolds." Mathematical Methods for Solving Economic Problems. 1977. Vol 7, pp. 85–102.

[7] Alexandra Bernadotte. "Topology-driven classification of time series." bioRxiv 2026.04.25.720787. 2026. doi:10.64898/2026.04.25.720787

[8] Alexandra Bernadotte, Ivan Menshikov. "Brain Oscillator Intrinsic Dimension Marks Consciousness, Recovery Mode, and Failed Downshifting in Coma." Aicumene Preprint 2026.042v2; SSRN working paper 7667278 (October 24, 2026). `https://ssrn.com/abstract=7667278`, doi:10.2139/ssrn.7667278. `https://research.aicumene.com/papers/boid-coma/`

[9] Ivan Menshikov, Nikita Elfimov, Alexandra Bernadotte. "A Lightweight Hankel-Embedded Pipeline for Real-Time EEG Filtering and Classification." 12th International Conference on Automation, Robotics, and Applications. IEEE. 2026. pp 589-593. doi:10.1109/ICARA69401.2026.11480292

[10] Bernadotte A. "Structural Modification of the Finite State Machine to Solve the Exponential Explosion Problem." Programmnaya Ingeneria. 2022. Vol 13(9), pp. 449–461. doi:10.17587/prin.13.449-461

[11] Broomhead D. S., King G. P. "Extracting qualitative dynamics from experimental data." Physica D. 1986. Vol 20, pp. 217–236. doi:10.1016/0167-2789(86)90031-x.

[12] Vautard R., Ghil M. "Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series." Physica D. 1989. Vol 35, pp. 395–424. doi:10.1016/0167-2789(89)90077-8.

[13] Golyandina N., Nekrutkin V., Zhigljavsky A. "Analysis of Time Series Structure: SSA and Related Techniques." Chapman & Hall/CRC. 2001. doi:10.1201/9781420035841.

[14] Golyandina N., Zhigljavsky A. "Singular Spectrum Analysis for Time Series." Springer. 2013. doi:10.1007/978-3-642-34913-3.

[15] Golyandina N., Korobeynikov A., Shlemov A., Usevich K. "Multivariate and 2D extensions of singular spectrum analysis with the Rssa package." Journal of Statistical Software. 2015. Vol 67(2), pp. 1–78. doi:10.18637/jss.v067.i02.

[16] Plaut G., Vautard R. "Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere." Journal of the Atmospheric Sciences. 1994. Vol 51, pp. 210–236. doi:10.1175/1520-0469(1994)051<0210:solfoa>2.0.co;2.

[17] Wax M., Kailath T. "Detection of signals by information theoretic criteria." IEEE Transactions on Acoustics, Speech, and Signal Processing. 1985. Vol 33, pp. 387–392. doi:10.1109/tassp.1985.1164557.

[18] Kritchman S., Nadler B. "Non-parametric detection of the number of signals: hypothesis testing and random matrix theory." IEEE Transactions on Signal Processing. 2009. Vol 57, pp. 3930–3941. doi:10.1109/tsp.2009.2022897.

[19] Schmidt R. O. "Multiple emitter location and signal parameter estimation." IEEE Transactions on Antennas and Propagation. 1986. Vol 34, pp. 276–280. doi:10.1109/tap.1986.1143830.

[20] Roy R., Kailath T. "ESPRIT — estimation of signal parameters via rotational invariance techniques." IEEE Transactions on Acoustics, Speech, and Signal Processing. 1989. Vol 37, pp. 984–995. doi:10.1109/29.32276.

[21] Delorme A., Palmer J., Onton J., Oostenveld R., Makeig S. "Independent EEG sources are dipolar." PLoS ONE. 2012. Vol 7, e30135. doi:10.1371/journal.pone.0030135.

[22] Fliess M. "Matrices de Hankel." Journal de Mathématiques Pures et Appliquées. 1974. Vol 53, pp. 197–222.

[23] Marchenko V. A., Pastur L. A. "Distribution of eigenvalues for some sets of random matrices." Mathematics of the USSR-Sbornik. 1967. Vol 1, pp. 457–483. doi:10.1070/sm1967v001n04abeh001994.

[24] Yin Y. Q., Bai Z. D., Krishnaiah P. R. "On the limit of the largest eigenvalue of the large dimensional sample covariance matrix." Probability Theory and Related Fields. 1988. Vol 78, pp. 509–521. doi:10.1007/bf00353874.

[25] Johnstone I. M. "On the distribution of the largest eigenvalue in principal components analysis." Annals of Statistics. 2001. Vol 29, pp. 295–327. doi:10.1214/aos/1009210544.

[26] Bryc W., Dembo A., Jiang T. "Spectral measure of large random Hankel, Markov and Toeplitz matrices." Annals of Probability. 2006. Vol 34, pp. 1–38. doi:10.1214/009117905000000495.

Aicumene:2026.044v2 [cs.HC] November 11, 2026

[27] Allen M. R., Smith L. A. "Monte Carlo SSA: detecting irregular oscillations in the presence of colored noise." Journal of Climate. 1996. Vol 9, pp. 3373–3404. doi:10.1175/1520-0442(1996)009<3373:mcsdio>2.0.co;2.

[28] Golyandina N., Shlemov A. "Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series." Statistics and Its Interface. 2015. Vol 8, pp. 277–294. doi:10.4310/sii.2015.v8.n3.a3.

[29] Hua Y., Sarkar T. K. "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise." IEEE Transactions on Acoustics, Speech, and Signal Processing. 1990. Vol 38, pp. 814–824. doi:10.1109/29.56027.

[30] Rife D. C., Boorstyn R. R. "Single-tone parameter estimation from discrete-time observations." IEEE Transactions on Information Theory. 1974. Vol 20, pp. 591–598. doi:10.1109/tit.1974.1055282.

[31] Weyl H. "Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen." Mathematische Annalen. 1912. Vol 71, pp. 441–479. doi:10.1007/BF01456804.

[32] Bai Z. D., Yin Y. Q. "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix." Annals of Probability. 1993. Vol 21(3), pp. 1275–1294. doi:10.1214/aop/1176989118.

[33] Grenander U., Szegő G. "Toeplitz Forms and Their Applications." University of California Press. 1958. doi:10.1525/9780520355408.

[34] Gray R. M. "Toeplitz and circulant matrices: a review." Foundations and Trends in Communications and Information Theory. 2006. Vol 2(3), pp. 155–239. doi:10.1561/0100000006.

[35] Begleiter H. "EEG Database." UCI Machine Learning Repository. 1999. doi:10.24432/C5TS3D.

[36] Zhang X. L., Begleiter H., Porjesz B., Wang W., Litke A. "Event related potentials during object recognition tasks." Brain Research Bulletin. 1995. Vol 38(6), pp. 531–538. doi:10.1016/0361-9230(95)02023-5.

[37] Snodgrass J. G., Vanderwart M. "A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity." Journal of Experimental Psychology: Human Learning and Memory. 1980. Vol 6(2), pp. 174–215. doi:10.1037/0278-7393.6.2.174.

[38] Bernadotte A. "Method of intelligent feedback for brain stimulation for the implementation of rehabilitation and/or prevention of neurodegeneration." Patent RU 2823580 C1, granted July 24, 2024 (application 2023119739 of July 26, 2023); international application PCT/RU2024/050075, published as WO 2025/023867 A1 on January 30, 2025. Patent holder: Neurosputnik LLC. https://patents.google.com/patent/RU2823580C1

[39] Bernadotte A. "Method for processing brain signals." Patent RU 2848137 C1, granted October 16, 2025 (application 2024131926 of October 24, 2024); international application PCT/RU2024/000328. Patent holder: Neurosputnik LLC. https://patents.google.com/patent/RU2848137C1

[40] Bernadotte A. "Method for assessing the degree of neurodegeneration by measuring signal complexity in the brain." Patent RU 2849295 C1, granted October 23, 2025 (application 2024140274 of December 28, 2024). Patent holder: Neurosputnik LLC. https://patents.google.com/patent/RU2849295C1

[41] Bernadotte A. "Method for activating a preserved brain pattern." Patent RU 2850121 C1, granted November 5, 2025 (application 2024140177 of December 27, 2024); international application PCT/RU2024/000398. Patent holder: Neurosputnik LLC. https://patents.google.com/patent/RU2850121C1

[42] Alexandra Bernadotte, Ivan Menshikov. "From Brain Dynamics to Strong Artificial Intelligence: Learning Creates Functional Regime Switching." Aicumene Preprint 2026.045v2. doi:10.13140/RG.2.2.22357.59364. https://research.aicumene.com/papers/ACT-BOID/

[43] Bernadotte A., Zinchenko O. "tDCS and neurofeedback in ADHD treatment." Frontiers in Systems Neuroscience. 2025. Vol 19, 1444283. doi:10.3389/fnsys.2025.1444283.

[44] Vorontsova D., Menshikov I., Zubov A., Orlov K., Rikunov P., Zvereva E., Flitman I., Lanikin A., Sokolova A., Markov S., Bernadotte A. "Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification." Sensors. 2021. Vol 21(20), 6744. doi:10.3390/s21206744.

[45] Bernadotte A., Mazurin A. D. "Optimization of the brain command dictionary based on the statistical proximity criterion in silent speech recognition task." Computer Research and Modeling. 2023. Vol 15(3), pp. 675–690. doi:10.20537/2076-7633-2023-15-3-675-690.

[46] Bernadotte A. "The Algorithm That Maximizes the Accuracy of k-Classification on the Set of Representatives of the k Equivalence Classes." Mathematics. 2022. Vol 10(15), 2810. doi:10.3390/math10152810.

[47] Bernadotte A. "Algorithm maximizing precision of k-classification on representatives of k equivalence classes for applied problem of selection of command dictionary." Proceedings of the 14th International Seminar "Discrete Mathematics and its Applications" in memory of Academician O. B. Lupanov. 2022. pp. 177–180. doi:10.20948/dms-2022-53.

[48] Mazurin A. D., Bernadotte A. "Clustering quality criterion based on the features extraction of a tagged sample with an application in the field of brain-computer interface development." Intelligent Systems. Theory and Applications. 2021. Vol 25(4), pp. 322–327.

[49] Vorontsova D. V., Isaeva M. V., Menshikov I. A., Orlov K. Y., Bernadotte A. "Frequency, time, and spatial electroencephalogram changes after COVID-19 during a simple speech task." Computer Research and Modeling. 2023. Vol 15(3), pp. 691–701. doi:10.20537/2076-7633-2023-15-3-691-701.

[50] Zubov A., Isaeva M., Bernadotte A. "Neural network classifier for EEG data from people who have undergone COVID-19 and have not." Intelligent Systems. Theory and Applications. 2021. Vol 25(4), pp. 318–321.

Aicumene:2026.044v2 [cs.HC] November 11, 2026