Decomposing the modulation of interactions between neuronal populations

Marco Celotto[1,2,*] **J. Samuel Sooter**[3,*] **Sofie Ährlund-Richter**[1]
Kyle R. Jenks[1] **Mriganka Sur**[1,†] **Stefano Panzeri**[2,†]

[1]Picower Institute for Learning and Memory, Massachusetts Institute of Technology, USA
[2]Institute of Neural Information Processing, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany
[3]Department of Physics, University of Arkansas, USA

mcelotto@mit.edu, sooter@uark.edu, s.panzeri@uke.de

Abstract

Identifying, through analysis of large-scale neural recordings, subpopulations of neurons interacting across brain areas is key for understanding across-brain communication with cellular resolution. The communication subspace (CS) framework addresses this need by capturing additive interactions between subspaces of source and target neural populations. However, no current method identifies interpretable components of how a measured, multivariate modulator - such as behavioral state or the activity of another population - reshapes these interactions. Here we extend the CS framework by using a bilinear perturbation of reduced-rank regression to parameterize modulation as a low-rank tensor. We derive MICs - multiplicative interaction channels - a method to identify triplets made of a source and a target activity pattern and a modulator of their interaction. We develop a hierarchical fitting pipeline and provide a closed-form decomposition that quantifies whether modulation reshapes the modulator-averaged baseline interaction, recruits private dimensions of one population, or opens new interactions. In simulations, MICs reliably recovers the presence and geometry of ground-truth modulation even in the high-dimensional, low-sample regime. Applied to a benchmark macaque dataset, MICs recovers the geometry and functional form of the stimulus modulation of interactions between primary and secondary visual cortices. Applied to simultaneous calcium imaging of prefrontal axons and interneurons in mouse visual cortex, MICs reveals that behavioral state asymmetrically modulates top-down interactions, reconfiguring the prefrontal projection patterns that interact with a stable set of visual interneurons. MICs provides a data-efficient, interpretable account of how measured modulators reshape the interactions between neural populations. Code is available at `https://github.com/mcelotto/multiplicative_interaction_channels`.

1 Introduction

Brain computations supporting cognitive functions arise from interactions between multiple brain regions [1, 2]. Inter-area interactions are implemented in the brain with cellular-level resolution [3–5], as specific subpopulations of neurons dynamically interacting with other subpopulations. As modern recording technologies access hundreds to thousands of neurons across multiple regions [6–8], a fundamental computational challenge is to develop analysis tools that can map these interactions

*These authors contributed equally.
†These authors jointly supervised this work.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

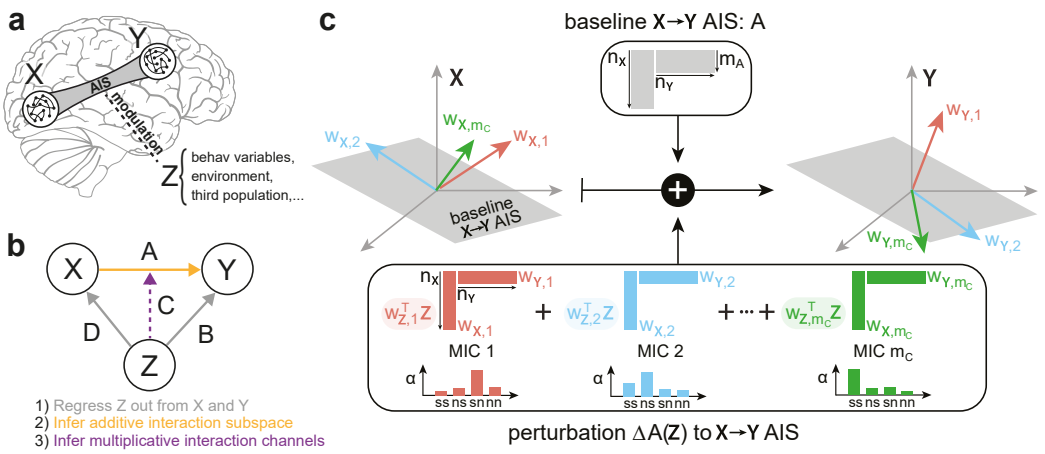

Figure 1: **Sketch of MICs**. (a) MICs quantify how the AIS between a multivariate source X and target Y changes as a function of a modulator Z. (b) Schematic of the modulation-channel fitting pipeline. (c) The perturbation $\Delta A(z)$ to the baseline $X \to Y$ mapping A is parameterized as a low-rank tensor: low-dimensional components of X and Z interact bilinearly to drive specific activity patterns in Y. Each channel's source and target directions need not lie within the baseline AIS; their geometry relative to it is summarized by four indices ($\alpha_{ss}, \alpha_{ns}, \alpha_{sn}, \alpha_{nn}$; see Sec. 3.4).

from simultaneous recordings of many neurons and work even in the under-sampled regime typical of brain data [9]. Classic pairwise approaches based on correlation [10] or the Wiener-Granger principle [11–14] were developed for scalar variables or small populations are impractical for high-dimensional recordings [15]. The *Communication Subspace* (CS; see Glossary A) [16] has emerged as a powerful framework to address this challenge. It uses Reduced-rank regression (RRR) to constrain the regression matrix relating two populations to low-rank to identify low-dimensional patterns of activity of one population that co-vary with low-dimensional activity patterns of the other, excluding variability private to each population [17]. The low-rank constraint dramatically reduces the number of free parameters, achieving data-efficiency and scalability. CS has become a popular tool to study inter-areal brain interactions with cellular resolution [18–22].

However, interactions between neural populations can change with cognitive state, movement parameters, task context, and activity of other brain regions. For example, attention *modulates* top-down interactions between higher cortical areas and sensory cortex [23, 24]. Yet the population-level structure of these modulations remains unknown. Modulation effects of a modulator on the relationship between other variables are usually described as multiplicative - rather than additive [25–28]. Moreover, modulatory variables can themselves be high-dimensional, such as multiple movement parameters [29] or the activity of other populations [30]. Despite the importance of modulatory interactions, CS analysis captures additive pairwise interactions between populations, but not the multiplicative modulation of their interaction by a third, potentially high-dimensional variable.

Here, we derive Multiplicative Interactions Channels (*MICs*), a method extending communication subspaces to capture how a high-dimensional modulator changes interactions between a source and a target population (1). We model the modulator-induced perturbation of the source-target mapping as a low-rank three-way tensor and parameterize it via a Canonical Polyadic (CP) decomposition [31–33]. Low-rank tensor decompositions can be robustly applied to neural population data [34–36] and to regression with multivariate covariates [37–39]. We thus parameterize the perturbation as a sum of rank-1 *multiplicative interaction channels*, each pairing one modulator pattern with the source-target interaction pattern it gates. We describe a hierarchical data-fitting procedure with cross-validated rank and regularization selection, and show that the geometry of each MIC relative to the modulator-averaged, baseline communication subspace can be decomposed into four interpretable components. We next validate MIC on simulated data. We then apply it to a public macaque V1-V2 benchmark dataset, recovering the expected stimulus dependence of the communication subspace with a single model and resolving its form and geometry, and to new simultaneous recordings of prefrontal axonal projections and visual cortical interneurons in mice.

2 Related work

Several methods model dynamically modulated interactions within or between neuronal populations, but none quantifies how a measured modulator reshapes the mapping between a source and a target population and describes the geometry of this modulation. They differ from MICs along at least one of two axes. First, *how the modulation enters*. In multi-region switching dynamical systems [40, 41], the interaction between regions changes with an inferred discrete state. In sequential autoencoders of inter-regional communication [42], messages between regions are inferred, and a third region can change the interaction only through the network's nonlinearity. In decomposed dynamical systems [43, 44], the dynamics are a combination of basis matrices with latent coefficients, from which behavior can be predicted [44]. These models can report that an interaction changed, but a measured variable never enters as an explicit modulator that the interaction depends on. Second, *what is modulated*. Low-rank tensor autoregressions [45, 46], behavior-constrained decomposed dynamics [44], and conditionally linear dynamical systems [47] modulate the recurrent dynamics of a single population. Several of these methods [43–47] share with MICs a decomposition into basis matrices. MICs differ in that the coefficients are functions of a measured modulator, the matrices map source to target, and each channel admits a geometric decomposition relative to the baseline interaction.

3 Mathematical derivation of multiplicative interaction channels

3.1 Background: the Communication Subspace (CS) framework

We consider simultaneous observations of three possibly high-dimensional variables: a *source* $x \in \mathbb{R}^{n_X}$, a *target* $y \in \mathbb{R}^{n_Y}$, and a *modulator* $z \in \mathbb{R}^{n_Z}$. For example, x and y could represent the activity of two neural populations, and z could be a set of behavioral variables (e.g., running speed, pupil diameter), or the activity of a third modulatory population. Given N simultaneous observations, such as time points in a multivariate time series, we build data matrices $\mathbf{X} \in \mathbb{R}^{N \times n_X}$, $\mathbf{Y} \in \mathbb{R}^{N \times n_Y}$, and $\mathbf{Z} \in \mathbb{R}^{N \times n_Z}$. The CS framework [16] identifies components of source activity linearly related to other specific components of target activity. It models the target as a low-rank linear function of the source via RRR:

$$\mathbf{Y} = \mathbf{X}A + \mathbf{E}, \qquad \operatorname{rank}(A) = m_A, \tag{1}$$

where $A \in \mathbb{R}^{n_X \times n_Y}$ and $\mathbf{E}$ denotes residual variability. The rank constraint implies that A can be factorized as $A = U_A V_A^\top$, with $U_A \in \mathbb{R}^{n_X \times m_A}$ and $V_A \in \mathbb{R}^{n_Y \times m_A}$. The columns of U_A and V_A define, respectively, the m_A source and target patterns that co-vary with each other; patterns outside $\operatorname{col}(U_A)$ and $\operatorname{col}(V_A)$ are private to each population [16]. The optimal A is obtained by ridge-RRR [48, 17] (Appendix B.1). We refer to $\operatorname{col}(U_A)$ and $\operatorname{col}(V_A)$ as the source and target *additive interaction subspaces* (AIS), rather than "communication subspaces" [16], as RRR at zero lag captures additive co-variation, not directed transmission [11, 12, 17], and source/target swapping does not, on its own, establish directionality [18]. Next, we will consider how to extend the CS framework beyond additive interactions.

3.2 Extending CS to multiplicative interactions beyond Additive Interaction Subspaces (AIS)

AIS capture *additive*, linear relationships between two variables. However, many interactions in neuroscience are hypothesized to be *modulated* by additional variables. We thus ask how the modulator Z changes the linear mapping from source to target (Fig. 1a). We model the instantaneous effective interaction operator as a z-dependent perturbation of the baseline mapping:

$$A_{\text{eff}}(z) = A + \Delta A(z), \tag{2}$$

where A is the baseline AIS and $\Delta A(z)$ captures the z-dependent modulation. We assume $\Delta A(z)$ to be linear in z, so that its action on the source is bilinear in x and z, with linearity in x at fixed z inherited from RRR. Bilinearity is deliberate: together with the low-rank parameterization below, it is what makes MICs sample-efficient (Section 4.1). We thus parameterize $\Delta A(z)$ as a sum of m_C rank-1 *multiplicative interaction channels*[1]:

$$\Delta A(z) = \sum_{i=1}^{m_C} (w_{Z,i}^\top z)\; w_{X,i}\, w_{Y,i}^\top, \tag{3}$$

[1] Unlike the RRR subspace - identifiable only as a column space - the CP components of the resulting tensor C satisfying Kruskal's condition [49] are individually identifiable (Appendix B.3); hence we call them *channels*, not subspaces.

where for each channel i, the weight vectors $w_{X,i} \in \mathbb{R}^{n_X}$, $w_{Y,i} \in \mathbb{R}^{n_Y}$, and $w_{Z,i} \in \mathbb{R}^{n_Z}$ define source, target, and modulator axes, respectively. Intuitively, MIC i captures the fact that when the modulator aligns with $w_{Z,i}$ (i.e., the $w_{Z,i}^\top z$ *score* is large), the effective mapping from source to target is perturbed by the rank-1 component $w_{X,i}\, w_{Y,i}^\top$, scaled proportionally to $w_{Z,i}^\top z$. This is a bilinear (multiplicative) interaction between source and modulator, analogous to an interaction term in a regression model, but operating in the high-dimensional setting where source, target, and modulator may each be high-dimensional (see Appendix B.2).

In matrix form, defining W_X, W_Z, W_Y as the factor matrices whose columns are $\{w_{X,i}\}$, $\{w_{Z,i}\}$, $\{w_{Y,i}\}$ respectively, and letting $\odot$ denote the element-wise (Hadamard) product, the modulation term across all N observations can be written compactly as:

$$\mathbf{Y}_{\text{mod}} = \begin{bmatrix} x_1^\top \Delta A(z_1) \\ \vdots \\ x_N^\top \Delta A(z_N) \end{bmatrix} = [(\mathbf{X}W_X) \odot (\mathbf{Z}W_Z)]\, W_Y^\top, \tag{4}$$

where $\mathbf{X}W_X \in \mathbb{R}^{N \times m_C}$ and $\mathbf{Z}W_Z \in \mathbb{R}^{N \times m_C}$ are the source and modulator scores (projections onto the respective axes), and their element-wise product captures the bilinear interaction. Mathematically, the collection $\{(w_{X,i}, w_{Z,i}, w_{Y,i})\}_{i=1}^{m_C}$ defines a rank-m_C CP decomposition [31–33] of a three-way tensor $\mathcal{C} \in \mathbb{R}^{n_Y \times n_X \times n_Z}$.

Beyond the baseline interaction A and the modulation tensor $\mathcal{C}$, the modulator may also exert direct additive effects on the target and on the source. We capture these with two additional low-rank linear terms: $B \in \mathbb{R}^{n_Z \times n_Y}$ (rank m_B) for the direct effect $Z \to Y$, and $D \in \mathbb{R}^{n_Z \times n_X}$ (rank m_D) for the effect $Z \to X$. Letting $\mathbf{X}' = \mathbf{X} - \mathbf{Z}D$ denote the *residualized source* (the component of source variability that is not linearly predictable from the modulator), the complete model is:

$$\mathbf{Y} = \underbrace{\mathbf{X}'A}_{\text{baseline AIS}} + \underbrace{\mathbf{Z}B}_{\substack{\text{additive} \\ Z \to Y}} + \underbrace{[(\mathbf{X}'W_X) \odot (\mathbf{Z}W_Z)]\, W_Y^\top}_{\text{modulation}} + \mathbf{E}. \tag{5}$$

The residualized source $\mathbf{X}'$ enters both the baseline interaction and the modulation terms. In Eq. 5, A and B capture the first-order effects of X and Z on Y, the modulation term captures their second-order $X \times Z$ interaction. The modulation term is a specific instance of CP-rank-constrained tensor-on-tensor regression [38], with the predictor restricted to the rank-1 outer product $x_n \otimes z_n$ of two distinct observed variables. It is this restriction, rather than the CP structure of $\mathcal{C}$ itself, that carries the bilinear-interaction form. The full MICs model (Eq. 5) embeds this term within an additive RRR framework with residualized source, a configuration not addressed in [38] (see Appendix B.2).

The baseline AIS operator A. We assume all variables X, Y, and Z are zero-centered at the stage of model fitting (see Section 3.3). The expected value of the modulatory perturbation is thus $\mathbb{E}[\Delta A(z)] = \sum_i \mathbb{E}[w_{Z,i}^\top z]\, w_{X,i} w_{Y,i}^\top = 0$. Thus, A represents the average effective AIS $\mathbb{E}[A_{\text{eff}}(z)] = A$. MICs capture how each observation deviates from this average, depending on z.

3.3 Hierarchical model fitting and modulatory explained variance

Fitting all parameters of Eq. 5 jointly is computationally challenging due to the interaction between the additive and multiplicative terms. Instead, we adopt a hierarchical procedure that sequentially isolates each component (Algorithm 1; Appendix B.6; Fig. 1b).

Step 1: Regressing out additive effects of Z. We first fit the additive effects of the modulator on both the target and the source via ridge-RRR:

$$B = \text{RRR}(\mathbf{Z}, \mathbf{Y};\, m_B, \lambda_B), \qquad D = \text{RRR}(\mathbf{Z}, \mathbf{X};\, m_D, \lambda_D), \tag{6}$$

where $\text{RRR}(\cdot, \cdot;\, m, \lambda)$ denotes ridge-RRR with rank m and penalty λ (see Appendix B.1), and compute residuals $\mathbf{X}' = \mathbf{X} - \mathbf{Z}D$, $\mathbf{Y}' = \mathbf{Y} - \mathbf{Z}B$. This removes variance in both X and Y that is linearly attributable to $\mathbf{Z}$ alone. Regressing Z out of X (via D) ensures that downstream terms capture X-Y relationship not confounded by Z independently driving both X and Y (see Sec. 4.1).

Step 2: Fitting baseline interaction. Given the residuals, we fit the baseline interaction matrix as:

$$A = \text{RRR}(\mathbf{X}', \mathbf{Y}';\, m_A, \lambda_A). \tag{7}$$

Step 3: Fitting MICs. We fit the modulation tensor to the residuals of the baseline interaction model, $\mathbf{Y}'' = \mathbf{Y}' - \mathbf{X}'A$, by minimizing the regularized loss:

$$\mathcal{L}(W_X, W_Z, W_Y) = \left\| \mathbf{Y}'' - [(\mathbf{X}'W_X) \odot (\mathbf{Z}W_Z)]\, W_Y^\top \right\|_F^2 + \lambda_C \left\| \mathcal{C} \right\|_F^2, \tag{8}$$

where $\|\mathcal{C}\|_F^2 = \mathrm{tr}\big(G_Y^\top (G_X \odot G_Z)\big)$ is the squared Frobenius norm of the tensor expressed in terms of the Gram matrices $G_X = W_X^\top W_X$, $G_Z = W_Z^\top W_Z$, $G_Y = W_Y^\top W_Y$ (see Appendix B.4). We minimize Eq. 8 via alternating least squares (ALS), cyclically updating each factor matrix while holding the other two fixed (Algorithm 2, [33, 38]). Each sub-problem reduces to a regularized linear system; closed-form update equations and computational details are provided in Appendix B.5. Eq. 8 is non-convex, so ALS is not guaranteed to reach the global optimum.

Rank and regularization selection. Each rank (m_A, m_B, m_C, m_D) and ridge penalty $(\lambda_A, \lambda_B, \lambda_C, \lambda_D)$ is selected via K-fold cross-validation. For a given rank, the ridge penalty is optimized using Bayesian optimization [50]. The rank is selected as the smallest value whose cross-validated performance is within one SE of the maximum [16, 51]. The same cross-validation folds are used across all steps preventing information leakage between hierarchical stages (Algorithm 1).

Modulatory explained variance Our primary measure of modulatory interaction strength is the *modulatory explained variance* ($\mathrm{EV}_{\mathrm{mod}}$): the cross-validated R^2 increment of the full model (Eq. 5) over the joint additive model $\mathbf{X}'A + \mathbf{Z}B$, averaged over folds (Appendix B.6). A positive $\mathrm{EV}_{\mathrm{mod}}$ indicates that the bilinear X-Z interaction explains target variance beyond the combination of additive terms. Analogously, $\mathrm{EV}_{\mathrm{AIS}}$ is the increment from adding $\mathbf{X}'A$ over the modulator-only prediction $\mathbf{Z}B$, yielding a variance partitioning of $\mathbf{Y}$ along the hierarchical fit.

3.4 Geometric decomposition of MICs

A MIC could modulate interaction geometry by rehaping baseline source-target interactions, engaging private source or target dimensions, or opening up new interaction pathways. We formalize these possibilities by decomposing each channel relative to the baseline AIS.

Let $P_X = \tilde{U}_A \tilde{U}_A^\top$ and $P_Y = \tilde{V}_A \tilde{V}_A^\top$ be orthogonal projectors onto the baseline source and target AIS, respectively, where $\tilde{U}_A$ and $\tilde{V}_A$ are orthonormal bases for $\mathrm{col}(U_A)$ and $\mathrm{col}(V_A)$. Using $I = P + (I - P)$ on both the source and target sides, we decompose each channel's rank-1 perturbation $\Delta A_i = w_{X,i} w_{Y,i}^\top$ into four orthogonal components (see Appendix B.7):

$$\Delta A_i = \underbrace{P_X \Delta A_i\, P_Y}_{\Delta A_i^{ss}} + \underbrace{(I - P_X)\Delta A_i\, P_Y}_{\Delta A_i^{ns}} + \underbrace{P_X \Delta A_i\, (I - P_Y)}_{\Delta A_i^{sn}} + \underbrace{(I - P_X)\Delta A_i\, (I - P_Y)}_{\Delta A_i^{nn}}, \tag{9}$$

where superscripts denote same (s) or new (n) dimensions relative to the AIS in the source/target respectively, and the score $w_{Z,i}^\top z$ is suppressed (it scales all components equally). Each component has a distinct interpretation (Fig.1c). ΔA_i^{ss}: modulation reshapes interactions within the baseline AIS. ΔA_i^{ns}: modulation makes source patterns outside the baseline AIS interact with target patterns within it. ΔA_i^{sn}: modulation makes target patterns outside the baseline AIS interact with source patterns within it. ΔA_i^{nn}: modulation creates new interaction pathways, outside the baseline AIS.

The four components are mutually orthogonal under the Frobenius inner product (Appendix B.7), yielding an additive decomposition of the channel strength $\|\Delta A_i\|_F^2$. We define fractional *geometry indices* as the fraction of channel strength carried by each component, $\alpha_k^{(i)} = \|\Delta A_i^k\|_F^2 \,/\, \|\Delta A_i\|_F^2$ for $k \in \{ss, ns, sn, nn\}$, thus having that $\sum_k \alpha_k^{(i)} = 1$. Importantly, the indices depend only on the orientation of $w_{X,i}$ and $w_{Y,i}$ relative to the baseline subspaces, not on the modulator, and factorize into independent source and target alignment terms - e.g., $\alpha_{ss}^{(i)} = \cos^2 \theta_X^{(i)} \cdot \cos^2 \theta_Y^{(i)}$, where $\theta_X^{(i)}$ ($\theta_Y^{(i)}$) is the angle between $w_{X,i}$ ($w_{Y,i}$) and the source (target) AIS - so that all four indices are determined by these two parameters alone (Appendix B.7.2). As a complement, we quantify each component's contribution to $\mathrm{EV}_{\mathrm{mod}}$ by fitting constrained MICs that restrict modulation to lie within or outside the baseline subspaces. This reduces to the unconstrained problem on the AIS-projected X' and Y'' (Appendix B.8), enabling direct cross-validated comparison of $\mathrm{EV}_{\mathrm{mod}}$ across geometric components.

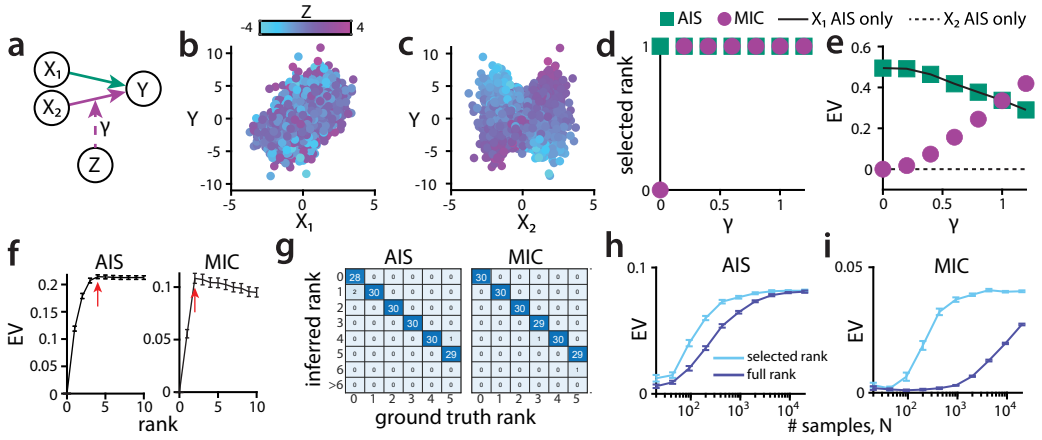

Figure 2: **Testing MICs on simulated data.** (a) Simulated scenario: a 2D source $X = (X_1, X_2)$ interacts with a 1D target Y; the $X_1 \to Y$ interaction is additive, whereas the $X_2 \to Y$ interaction is multiplied by γZ, where γ is the parameter setting the modulation strength. (b) Joint distribution of X_1 and Y; dots color indicates the value of Z. (c) Same as (b) for X_2 and Y. (d) Selected AIS and modulation ranks as function of γ. (e) Explained Variance (EV) by the AIS (green) and MICs (purple) vs γ. Solid and dashed black lines denote the AIS EV from X_1 and X_2 alone, respectively. (f) Model rank selection. AIS (left) and modulation (right) EV as function of the fitted rank (ground-truth ranks: $m_A = 4$, left; $m_C = 2$, right). Lines: mean $\pm$ SEM across cross-validation folds; red arrows: selected rank. (g) Heatmaps of ground-truth versus inferred ranks for the additive (left) and multiplicative (right) channels; $n_X = n_Y = n_Z = 20$, 5 simulations per (m_A,m_C) ground-truth rank combination (30 per marginal rank). (h) Variance explained by AIS as a function of sample size N, for the rank-selected low-rank model (cyan) and the full-rank model (blue); 30 simulations per scenario, $n_X = n_Y = n_Z = 20$; $m_A = 2$, $m_C = 1$. (i) Same as (h) for MICs.

4 Validating MICs on simulated data

We validated our pipeline on simulated data (see Appendix B.9 for simulation details).

4.1 Rank recovery, sample-efficiency, and benchmarks

To test whether MICs correctly quantify changes in source-target interactions as a function of a modulator Z, we first simulated a minimal scenario (Fig. 2a): a 2D source $X = (X_1, X_2)$, a 1D target Y, and a 1D modulator Z. X_1 drove Y additively, independently of Z (Fig. 2b), while the $X_2 \to Y$ interaction was gated by Z with strength γ (Fig. 2c). AIS correctly recovered a rank-1 subspace from X to Y, and MICs recovered a rank-1 modulation whenever $\gamma > 0$ (Fig. 2d). MIC EV increased monotonically with γ, while AIS EV decreased as a growing fraction of the variance of Y was absorbed by the Z-gated interaction - confirming that AIS is blind to modulatory contributions (Fig. 2e). Fitting AIS on X_1 or X_2 alone made this explicit: AIS EV from X_1 matched the full-X AIS EV, whereas AIS EV from X_2 remained at zero regardless of γ (Fig. 2e, dashed lines).

Next, we considered high-dimensional sources, targets, and modulators ($n_X = n_Y = n_Z = 20$). We varied the ground-truth AIS and MIC ranks from 0 to 5 drawing U_A, V_A, W_X, W_Y, W_Z at random for each simulation. For each fit, we selected the simplest model within 1 SEM of the maximum cross-validated EV (Fig. 2f, [16, 51]). Because this rule occasionally assigns positive ranks when the ground-truth rank is zero (Fig. S1), we retained a non-zero rank only if a residual-shuffle permutation test was significant ($p < 0.05$; Appendix B.13). The full pipeline recovered AIS and MIC ground-truth ranks in most simulations (Fig. 2g). Across all 16 on/off combinations of the four model terms, it yelded 4% and 0% false positives m_A and m_C (Fig. S1a-c). Skipping the significance test inflated them to 18% and 34%, skipping Step 1 inflated m_A whenever Z drove both X and Y (Fig. S1d-i).

We then assessed how AIS and MIC EV recovery scaled with sample size N and dimensionality $d = n_X = n_Y = n_Z$, against full-rank counterparts (no tensor-rank constraint on AIS or MICs).

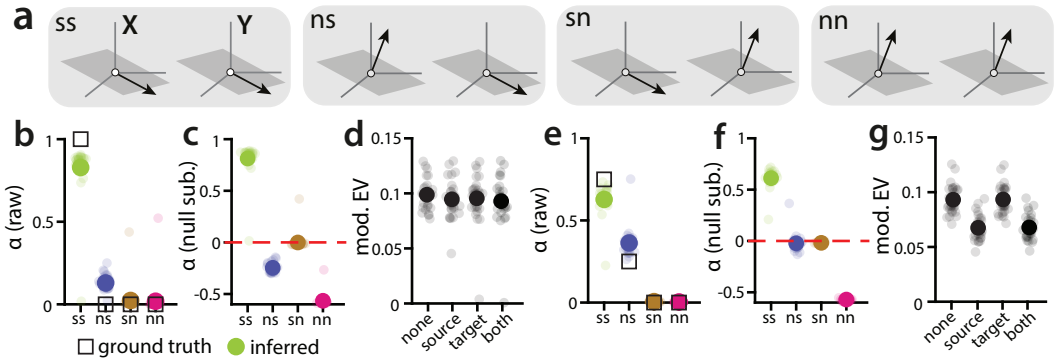

Figure 3: **Validating the geometric decomposition on simulations.** (a) The four geometric components (ss, ns, sn, nn) classifying how $w_{X,i}$ and $w_{Y,i}$ align with the baseline AIS (grey plane) on the source and target sides. (b) Raw geometric indices for a scenario with ground-truth (GT) $\alpha_{ss}^{\mathrm{GT}} = 1$ and all others zero. Squares: ground truth; filled circles: mean across simulations; light dots: individual simulations. (c) Same data after subtracting the per-simulation random-rotation null (dashed line: α equal to expected value under the null). (d) Constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ with the indicated MIC weights constrained to lie in the baseline AIS:"source": w_X; "target": w_Y; "both": both; "none": unconstrained fit. (e-g) Same as (b-d) for $\alpha_{ss}^{\mathrm{GT}} = 3/4$, $\alpha_{ns}^{\mathrm{GT}} = 1/4$. We ran 30 simulations in each scenario with 1000 samples each. $n_X = 50$, $n_Y = 5$, $n_Z = 1$ in all simulations.

Reduced-rank fits saturated the ground-truth EV with substantially fewer samples in both cases (Fig. 2h-i and S2a-c, f-h). Normalized recovery curves collapsed onto N/d for both reduced-rank models (Fig. S2d, i), versus N/d^2 for full-rank AIS and N/d^3 for full-rank MIC (Fig. S2e, j), consistent with $\mathcal{O}(d)$ per-rank parameter counts against $\mathcal{O}(d^2)$ and $\mathcal{O}(d^3)$ for unconstrained AIS and MICs (Appendix B.2). Rank restriction therefore reduces the sample requirement by a factor up to d for AIS and d^2 for MIC. Moreover, MICs also tolerated misspecification of the residual noise model: replacing Gaussian emissions on X and Y with Poisson distribution often used to describe spike counts [52] - or more extreme zero-inflated Poisson preserved $\mathrm{EV}_{\mathrm{mod}}$ scaling with overall modulation strength and recovery of ground truth MICs direction (Fig. S3).

Since prior work examined how AIS changes across discrete levels of a third variable [3, 19, 53, 54], we benchmarked MICs versus a discretization approach. We partitioned Z into N_{bins} quantile bins, refitted AIS within each bin, and computed the $\mathrm{EV}_{\mathrm{mod}}$ as the gap between these binned fits and a single global AIS fit ("Z-split"; same cross-validation folds). For 1D Z, $\mathrm{EV}_{\mathrm{mod}}$ was non-monotonic in N_{bins} - rising as finer bins resolved Z-dependent structure, then collapsing as per-bin fits became data-starved - but its peak stayed below MIC $\mathrm{EV}_{\mathrm{mod}}$ (Fig. S4a). For $\dim(Z) > 1$, where no natural discretization exists, Z-split applied to the leading principal component of Z degraded sharply as additional Z dimensions misaligned with that component, while MIC $\mathrm{EV}_{\mathrm{mod}}$ remained stable even in presence of distractor dimensions (Fig. S4b-d).

Finally, we tested MICs on generators that violate bilinearity (Appendix B.10; Fig. S5). In every case $\mathrm{EV}_{\mathrm{mod}}$ increased monotonically with the modulation strength: a nonlinear ground truth costs MICs sensitivity, not validity. This sensitivity loss is the price of bilinearity. A one-hidden-layer multilayer perceptron fit to the same residuals, with up to 100 times more parameters, captured less modulation than MICs at all sample sizes under bilinear or mildly nonlinear generators, and overtook MICs under strong nonlinearity only at large N, at the cost of interpretability (Appendix B.12; Fig. S6).

Together, these results show that MICs efficiently and reliably recover low-dimensional modulation structure in high-dimensional settings - a component of Y to which AIS is blind.

4.2 Numerical validation of the geometric decomposition

We next tested whether the geometry indices α and the constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ recover the correct MIC geometry with respect to AIS (Fig. 3a). All simulations used $n_X = 50$, $n_Y = 5$, and $n_Z = 1$, implementing a source-target dimensional imbalance which may be present in brain recordings.

We first simulated a MIC entirely confined to the baseline AIS ($\alpha_{ss}^{\mathrm{GT}} = 1$, all others zero). Raw indices returned a large α_{ss} but also a systematic $\alpha_{ns} > 0$ (Fig. 3b). This positive bias is a geometric artifact of the large difference between n_X and the baseline source rank m_A: because $n_X - m_A$ dimensions are available orthogonal to the AIS, small fit-driven deviations of $w_{X,i}$ out of the baseline subspace inflate α_{ns}. To correct for this, for each fitted MIC we drew $w_{X,i}$ and $w_{Y,i}$ independently and uniformly on their respective unit spheres, we recomputed the four indices, and subtracted the random-rotation null means from the raw α to obtain null-subtracted $\tilde{\alpha}$ (see Appendix B.7.3). After null subtraction only $\tilde{\alpha}_{ss}$ remained positive (Fig. 3c), consistent with ground truth. The constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ agreed: restricting the MIC to the ss component alone (both source and target MICs constrained to the respective AIS) recovered essentially all of the unconstrained $\mathrm{EV}_{\mathrm{mod}}$ (Fig. 3d). We next considered a mixed geometry ($\alpha_{ss}^{\mathrm{GT}} = 3/4$ and $\alpha_{ns}^{\mathrm{GT}} = 1/4$). Raw indices correctly individuated both components (Fig. 3e), but null subtraction pushed $\tilde{\alpha}_{ns}$ around zero (Fig. 3f): the true ns contribution was on the order of the random alignment the null was calibrated to remove, so the test dismissed a genuine component. Constrained EV resolved the ambiguity: the ss-only fit recovered $\approx 3/4$ of the unconstrained $\mathrm{EV}_{\mathrm{mod}}$. The gap was closed by allowing MICs to lie outside AIS in the source (target constraint; $\alpha_{ns} > 0$), but not by including sn (source constraint; Fig. 3g). A sweep over multi-MIC scenarios ($m_C > 1$) with randomly distributed ground-truth α further confirmed that geometry indices recover the correct geometry in more complex settings (Fig. S7). Geometry recovery also held under the nonlinear generators (Fig. S5). Since bilinearity may be violated in x or in z, we developed a held-out diagnostic that adds higher-order terms in either factor to a fitted channel (Appendix B.11). For sufficient modulation and nonlinearity strength, the diagnostics detected the leading-order nonlinearity of each generator and localized it to the correct factor, and did not detect any nonlinearity on bilinear data (Fig. S8).

Null-subtracted $\tilde{\alpha}$ and constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ are thus complementary: the former is closed-form and cheap, an efficient first-pass diagnostic that can provide false negatives; the latter is more expensive but directly and reliably quantifies each geometric term's contribution to $\mathrm{EV}_{\mathrm{mod}}$.

5 MICs quantifies stimulus and state modulation of cortical interactions

Stimulus modulation of V1-V2 interactions (benchmark dataset). We first applied MICs to the dataset [55] used to benchmark communication subspaces [16]. Anesthetized macaques (3 animals, 5 sessions) were shown drifting gratings with 8 different orientations θ while recording activity from V1 and V2 (Fig. 4a; Appendix B.14). Following [16], we analyzed 100 ms spike counts after subtracting the time-dependent trial-averaged activity to eliminate the 1st-order effect of θ on both populations ($B = D = 0$). Per-orientation communication subspace fits from previous work suggested that the linear map from V1 to V2 residuals depends on θ: predictive dimensions fit at one orientation generalize imperfectly to the others [16], and across-area subspaces vary approximately periodically with orientation difference [56]. Both observations rest on comparing separately fitted subspaces, without a statistical test of θ-dependence or a measure of its size. We used MICs to fit a single model across all orientations, test this dependence directly, and quantify it. Taking Z as a one-hot encoding of θ, which imposes no functional form, orientation significantly modulated the interaction in 5/5 sessions (shuffle test, Appendix B.13; $\mathrm{EV}_{\mathrm{mod}} = (1.3 \pm 0.1)\%$, mean $\pm$ SEM, i.e. $(16.6 \pm 2.5)\%$ of the θ-invariant $\mathrm{EV}_{\mathrm{AIS}}$; Fig. 4b, S9a-b; Table 1). Beyond detecting and quantifying the θ-dependence of interactions, MICs add two new results. First, it shows the form of the modulation: the modulator scores of an example MIC tracked the angular difference between orientations (Fig. 4c), and the periodic encoding $Z = (\sin 2\theta, \cos 2\theta)$ retained $(83.6 \pm 1.8)\%$ of the one-hot $\mathrm{EV}_{\mathrm{mod}}$ (Fig. S9d). Thus, the smooth first harmonic accounts for most of a modulation free to take any form, quantifying the periodicity previously observed qualitatively [56]. Second, it determines its geometry relative to the θ-invariant AIS, which per-condition refits cannot access: $\tilde{\alpha}_{ss}$, $\tilde{\alpha}_{ns}$ and $\tilde{\alpha}_{sn}$ exceeded the null (Fig. 4d, top). The V1- and V2-constrained fits, which add sn and ns to the both-in-AIS (ss-only) fit, respectively, both significantly exceeded it and did not differ significantly from each other, confirming the presence of ns and sn components (Fig. 4d, bottom). $\tilde{\alpha}_{nn}$ was negative, i.e. the nn component was weaker than expected for randomly oriented MICs, but not absent: an nn-only model still yielded significant $\mathrm{EV}_{\mathrm{mod}}$ in 5/5 sessions ($(0.24 \pm 0.05)\%$ Fig. S9c). The significant ns and sn components constrain the architecture of V1-V2 interactions. If orientation gated interaction between segregated channels each inside or outside the AIS, modulation would be confined to ss and nn (Fig. S9f, left). Instead, θ-invariant and θ-specific interactions are interleaved (Fig. S9f, right): V2 patterns in the AIS interact, at specific orientations, with V1 patterns outside it (ns), and vice versa (sn).

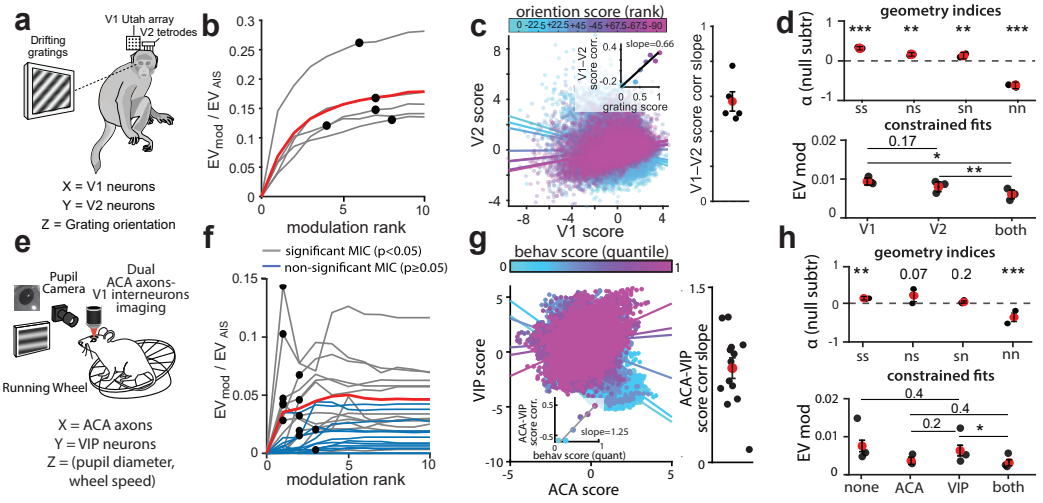

Figure 4: **Application of MICs to real data.** (a) Simultaneous V1-V2 recordings from macaques shown drifting oriented gratings. (b) Modulatory EV relative to AIS EV vs modulation rank; gray lines: individual sessions ($n = 5$, all with significant $\mathrm{EV}_{\mathrm{mod}}$); red: mean; black dots: selected rank. (c) Left: V1 vs V2 projection scores of one MIC from an example session. Color: orientation, ranked by its modulator score (one-hot encoding; labels: angular difference from the lowest-scoring orientation). Inset: V1 and V2 scores correlation within each orientation vs orientation score. Right: average correlation slope across MICs; dots: sessions. (d) Top: null-subtracted geometry indices. Bottom: modulatory EV under different geometry constraints (V1: source-in-AIS; V2: target-in-AIS; both: both constraints; n = 3 animals). (e) Simultaneous imaging of ACA axons and VIP neurons in V1 with pupil and wheel measures. (f) Same as (b) for ACA-VIP data; gray (blue) lines: sessions ($n = 28$) with significant (non-significant) $\mathrm{EV}_{\mathrm{mod}}$. (g) Same as (c) for ACA-VIP data; color: quantiles of behavioral state score. Correlations are computed within each quantile; right: sessions with significant $\mathrm{EV}_{\mathrm{mod}}$ ($n = 13$). (h) Same as (d) for ACA-VIP data (ACA: source-in-AIS; VIP: target-in-AIS; none: unconstrained; n = 3 animals). P-values in d,h: linear mixed-effects models with animal as random effect (Appendix B.14), testing each index against zero (top), and differences between constrained fits (bottom). *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Holm corrected across the comparisons shown, selected as in Appendix B.8. We plot mean ± SEM across sessions.

Behavioral state modulates prefrontal-visual top-down interactions (new dataset). Top-down projections from prefrontal cortex (PFC) to V1 gate visual processing in a state-dependent manner [57–59]. Anterior cingulate (ACA, part of PFC) axons modulate responses of Vasoactive Intestinal Peptide (VIP) interneurons in V1 [57], and the activity of both ACA axons and VIP V1 neurons depends on arousal and locomotion [58, 59]. However, whether and how behavioral state reshapes the ACA populations that interact top-down with V1 target populations is unknown. We thus used MICs to address this question. We recorded and publicly release a new dataset of simultaneous two-photon imaging of ACA axonal projections in V1 and V1 VIP interneurons in head-fixed mice during passive visual stimulation (drifting gratings or natural movies), while tracking pupil diameter and running speed (Fig. 4e; Appendix B.15; 3 animals, 28 sessions). We took ACA axons as source X and VIP neurons as target Y (following anatomical directionality), and defined the modulator $Z \in \mathbb{R}^2$ (hereafter *behavioral state*) as the joint vector of two well-established markers of behavioral state, pupil diameter and running speed [58, 60]. MICs identified significant $\mathrm{EV}_{\mathrm{mod}}$ in 13/28 sessions (residual circular shuffle test on held-out sequential data halves; Fig. S10a-c; Appendix B.13). In these sessions, MICs accounted for $(0.7 \pm 0.1)\%$ of total VIP population variance, i.e. $(4.7 \pm 1.1)\%$ of $\mathrm{EV}_{\mathrm{AIS}}$ (Table 2; mean selected rank $m_C = 1.69$). Thus, behavioral state modulates ACA-VIP interactions beyond driving each population alone (Fig. S10d-e), during both grating and natural-movie stimulation (Fig. S10f). The correlation between ACA and VIP MIC projection scores increased with the behavioral-state score $w_{Z,i}^{\top} z$ consistently across sessions (Fig. 4g). In contrast to V1-V2, the geometry of these MICs was asymmetric. $\tilde{\alpha}_{ss}$ was significantly above zero and $\tilde{\alpha}_{nn}$ significantly below it, whereas $\tilde{\alpha}_{ns}$ only trended above zero and $\tilde{\alpha}_{sn}$ did not (Fig. 4h, top). As

the null subtraction can miss a weak genuine component (Fig. 3f), we resolved the geometry with constrained fits (Fig. 4h, bottom). Adding ns to the both-in-AIS fit (VIP constraint) significantly increased $\mathrm{EV}_{\mathrm{mod}}$, whereas adding sn (ACA constraint) did not, nor did further adding sn and nn to the VIP-constrained fit (unconstrained model). Although the ACA- and VIP-constrained fits did not differ significantly, MIC-AIS alignment was significantly larger in VIP than in ACA (Fig. S10g), with MIC weights aligning strongly with the VIP AIS but weakly with the ACA AIS (Fig. S10i-j). Neither this asymmetry nor the geometry indices were explained by the ACA:VIP population-size imbalance (Fig. S10k-m). Both $\mathrm{EV}_{\mathrm{mod}}$ and the recovered MIC geometry were robust to random initialization (Fig. S11a-b). As in simulations, Z-split fits recovered less than half the $\mathrm{EV}_{\mathrm{mod}}$ of MICs, degrading further with the number of bins (Fig. S11c-e). Dropping pupil diameter from Z did not reduce $\mathrm{EV}_{\mathrm{mod}}$, whereas dropping running speed did, identifying running speed as the modulating factor (Fig. S11f). Moreover, in both datasets, the bilinearity diagnostic detected no significant nonlinearity in either x or z (Figs. S9e, S11g). Together, these analyses reveal that behavioral state primarily reconfigures ACA-axon activity patterns interacting with more stable VIP activity patterns.

6 Discussion

MICs extends the CS framework [16] to quantify how a multivariate modulator Z reshapes interactions between a source X and a target Y. Its low-rank structure ensures tractability in the under-sampled, high-dimensional regime of neural data [8, 9]. We provide MICS with a cross-validated hierarchical fitting pipeline a residual-shuffle significance test, a closed-form decomposition into orthogonal geometric components, enabling reliable detection and geometric characterization of whether modulation reshapes baseline interactions, recruits private dimensions of one population, or opens new interactions. This geometric quantification of how a measured modulator changes interactions complements existing models of dynamically modulated population interactions [40–47].

We validated MICs on both simulated and real neural data. On simulated data, MICs reliably recovered ground-truth modulation ranks and MICs geometry with respect to AIS, outperformed simpler approaches, and resolved modulatory EV using up to d^2 fewer samples than full-rank tensor regression. We tested MIC on two neural datasets. On a benchmark existing dataset, MICs recovered the expected stimulus dependence of V1-V2 communication [16, 56] with a single model, resolved its smooth periodic form and the interleaving of stimulus-invariant and stimulus-specific interaction patterns in both areas. On a new dataset that we publicly release (Appendix B.15), MICs revealed that behavioral state reconfigures the interactions between ACA-axon patterns and a stable set of V1 VIP patterns. This asymmetry could reflect a top-down control architecture in which state-dependent recruitment of distinct ACA projection patterns is routed through a conserved VIP interface [57], eliciting a consistent disinhibitory local influence onto V1 pyramidal cells [61].

Several limitations point to natural extensions. MICs capture linear effects of Z on the source-target mapping. The bilinear structure of MICs lends itself to extensions capturing also non-linear effects of Z on the source-target mapping [47], and to learning the modulator encoding function from data. Moreover, so far we restricted our analyses to interactions at zero time lag, with source and target assigned from anatomy. Similar to time-lagged CS analyses [56], using MICs with time-lagged-interactions, possibly combined with Wiener-Granger analyses [11, 12], would allow MICs to assess from data whether interactions are unidirectional or bidirectional, and to quantify how each direction is modulated. In addition, our significance test fits on one half of a session and tests generalization to the other, entirely held-out half. This makes it conservative and data-demanding. Thus more data-efficient tests will need to be developed. Finally, conclusions from our real datasets are restricted to the small populations of recorded animals. Further, we currently take the baseline as the session mean. Alternative schemes (e.g., anchoring the baseline at biologically interpretable values of the modulator) are possible. Finally, the subspaces jointly estimated by MICs for encoding (B, D) and interaction (AIS), possibly integrated with other tools to study the geometry of population coding ([62–64]), could address whether the population activity subspaces encoding information coincide or differ from those that transmit information about the modulating variable across areas, providing a unifying framework for information encoding and transmission.

By providing a compact, data-efficient, and theoretically grounded characterization of how a possibly high-dimensional modulator reshapes interactions between high-dimensional populations, MICs are well positioned to enable new discoveries about context-dependent interactions in neural systems.

Acknowledgments and Disclosure of Funding

We thank S.M. Lemke and members of the Sur lab for feedback and helpful discussions, and G. Sipe for the illustration in Fig. 4e. This work was supported by European Union's Horizon research and innovation programme grant agreement No. 101206609, Marie Sklodowska-Curie Global Fellowship - AstroError (MC); the German Federal Ministry of Education and Research (BMBF 01GQ2404 and 01DM26028; SP); NIH grants R01MH126351; R01MH133066; R01NS130361; R01MH085802; MURI Grant W911NF2110328; The Picower Institute Innovation Fund; and the Simons Foundation Autism Research Initiative through the Simons Center for the Social Brain (MS); NIH grant 1K99EY035752 (SÄR); NIH grant F32EY032756 (KRJ); the MIT MSRP Bio program (NSF STC award CCF-1231216; JSS).

References

[1] S.L. Bressler and V. Menon. Large-scale brain networks in cognition: emerging methods and principles. *Trends in cognitive sciences*, 14(6):277–290, 2010.

[2] F. Varela, J. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4):229–239, 2001.

[3] J. D. Semedo, A. I. Jasper, A. Zandvakili, A. Krishna, A. Aschner, C. K. Machens, A. Kohn, and B. M. Yu. Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, 13(1):1099, 2022.

[4] S. Panzeri, M. Moroni, H. Safaai, and C.D. Harvey. The structures and functions of correlations in neural population codes. *Nature Reviews Neuroscience*, 23(9):551–567, 2022.

[5] H. Safaai, A. Y. Wang, S. Kira, S. Blanco Malerba, S. Panzeri, and C. D. Harvey. Specialized structure of neural population codes in parietal cortex outputs. *Nature Neuroscience*, 28(12):2550–2560, 2025.

[6] J.J. Jun, N.A. Steinmetz, J.H. Siegle, D.J. Denman, M. Bauza, B. Barbarits, A.K. Lee, C.A. Anastassiou, A. Andrei, Ç. Aydın, M. Barbic, T.J. Blanche, V. Bonin, J. Couto, B. Dutta, Sergey L. Gratiy, D.A. Gutnisky, M. Häusser, B. Karsh, P. Ledochowitsch, C. Mora Lopez, C. Mitelut, S. Musa, M. Okun, M. Pachitariu, J. Putzeys, P.D. Rich, C. Rossant, W. Sun, K. Svoboda, M. Carandini, K.D. Harris, C. Koch, J. O'Keefe, and T.D. Harris. Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236, 2017.

[7] N. A. Steinmetz, C. Aydin, A. Lebedeva, M. Okun, M. Pachitariu, M. Bauza, M. Beau, J. Bhagat, C. Böhm, M. Broux, S. Chen, J. Colonell, R. J. Gardner, B. Karsh, F. Kloosterman, D. Kostadinov, C. Mora-Lopez, J. O'Callaghan, J. Park, J. Putzeys, B. Sauerbrei, R. J. J. van Daal, A. Z. Vollan, S. Wang, M. Welkenhuysen, Z. Ye, J. T. Dudman, B. Dutta, A. W. Hantman, K. D. Harris, A. K. Lee, E. I. Moser, J. O'Keefe, A. Renart, K. Svoboda, M. Häusser, S. Haesler, M. Carandini, and T. D. Harris. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*, 372(6539):eabf4588, 2021.

[8] C. Stringer, M. Pachitariu, N. Steinmetz, M. Carandini, and K. D. Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365, 2019.

[9] P. Gao and S. Ganguli. On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32:148–155, 2015.

[10] M.R. Cohen and A. Kohn. Measuring and interpreting neuronal correlations. *Nature Neuroscience*, 14(7):811–819, 2011.

[11] C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

[12] A.K. Seth, A.B. Barrett, and L. Barnett. Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297, 2015.

[13] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000.

[14] M. Celotto, J. Bím, A. Tlaie, V. De Feo, A. Toso, S. Lemke, D. Chicharro, H. Nili, M. Bieler, I. Hanganu-Opatz, T. Donner, A. Brovelli, and S. Panzeri. An information-theoretic quantification of the content of communication between brain regions. In *Advances in Neural Information Processing Systems*, volume 36, pages 64213–64265, 2023.

[15] E. N. Brown, R. E. Kass, and P. P. Mitra. Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nature Neuroscience*, 7:456–461, 2004.

[16] J.D. Semedo, A. Zandvakili, C.K. Machens, Byron M. Yu, and A. Kohn. Cortical areas interact through a communication subspace. *Neuron*, 102(1):249–259.e4, 2019.

[17] B. Wu and J. Pillow. Reduced rank regression for neural communication: a tutorial for neuroscientists. *arXiv:2512.12467*, 2025.

[18] C. J. MacDowell, A. Libby, C. I. Jahn, S. Tafazoli, A. Ardalan, and T. J. Buschman. Multiplexed subspaces route neural activity across brain-wide networks. *Nature Communications*, 16:3359, 2025.

[19] X. W. Chia, J. K. Tan, L. F. Ang, et al. Emergence of cortical network motifs for short-term memory during learning. *Nature Communications*, 14:6869, 2023.

[20] S. Borgognon, N. Macellari, A. M. Hickey, M. G. Perich, H. Javaheri, R. Ornelas-Kobayashi, M. Delacombaz, C. Hitz, F. Fallegger, S. P. Lacour, E. Bezard, E. M. Rouiller, J. Bloch, T. Milekovic, I. Seáñez, and G. Courtine. Regional specialization of movement encoding across the primate sensorimotor cortex. *Nature Communications*, 16(1):5729, 2025.

[21] N. H. Lam, A. Mukherjee, R. D. Wimmer, M. R. Nassar, Z. S. Chen, and M. M. Halassa. Prefrontal transthalamic uncertainty processing drives flexible switching. *Nature*, 637(8044):127–136, 2025.

[22] N. Binish, J. Terlau, J. Martini, J. J. Lin, R. T. Knight, and R. F. Helfrich. A communication subspace relays context-dependent actions from human prefrontal to motor cortex. *Nature Neuroscience*, 2026.

[23] J. Martínez-Trujillo and S. Treue. Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, 35(2):365–370, 2002.

[24] J.H. Reynolds and D.J. Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

[25] E. Salinas and P. Thier. Gain modulation: A major computational principle of the central nervous system. *Neuron*, 27(1):15–21, 2000.

[26] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13:51–62, 2012.

[27] K.A. Ferguson and J.A. Cardin. Mechanisms underlying gain modulation in the cortex. *Nature Reviews Neuroscience*, 21:80–92, 2020.

[28] S. M. Jayakumar, J. Menick, W. M. Czarnecki, J. Schwarz, J. W. Rae, S. Osindero, Y. W. Teh, T. Harley, and R. Pascanu. Multiplicative interactions and where to find them. In *International Conference on Learning Representations*, 2020.

[29] S. Musall, M. T. Kaufman, A. L. Juavinett, S. Gluf, and A. K. Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10):1677–1686, 2019.

[30] N. K. Totah, R. M. Neves, S. Panzeri, N. K. Logothetis, and O. Eschenko. The locus coeruleus is a complex and differentiated neuromodulatory system. *Neuron*, 99(5):1055–1068.e6, 2018.

[31] J. D. Carroll and J.-J. Chang. Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35(3):283–319, 1970.

[32] R. A. Harshman. Foundations of the parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, 1970.

[33] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[34] A. H. Williams, T. H. Kim, F. Wang, S. Vyas, S. I. Ryu, K. V. Shenoy, M. Schnitzer, T. G. Kolda, and S. Ganguli. Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron*, 98(6):1099–1115.e8, 2018.

[35] J. S. Seely, M. T. Kaufman, S. I. Ryu, K. V. Shenoy, J. P. Cunningham, and M. M. Churchland. Tensor analysis reveals distinct population structure that parallels the different computational roles of areas m1 and v1. *PLoS Computational Biology*, 12(11):e1005164, 2016.

[36] A. Onken, J. K. Liu, P. P. Karunasekara, I. Delis, T. Gollisch, and S. Panzeri. Using matrix and tensor factorizations for the single-trial analysis of population spike trains. *PLoS Computational Biology*, 12(11):e1005189, 2016.

[37] H. Zhou, L. Li, and H. Zhu. Tensor regression with applications in neuroimaging data analysis. *Journal of the American Statistical Association*, 108(502):540–552, 2013.

[38] E. F. Lock. Tensor-on-tensor regression. *Journal of Computational and Graphical Statistics*, 27(3):638–647, 2018.

[39] P. D. Hoff. Multilinear tensor regression for longitudinal relational data. *The Annals of Applied Statistics*, 9(3):1169–1193, 2015.

[40] J. I. Glaser, M. Whiteway, J. P. Cunningham, L. Paninski, and S. W. Linderman. Recurrent switching dynamical systems models for multiple interacting neural populations. In *Advances in Neural Information Processing Systems*, volume 33, pages 14867–14878, 2020.

[41] O. Karniol-Tambour, D. M. Zoltowski, E. M. Diamanti, L. Pinto, C. D. Brody, D. W. Tank, and J. W. Pillow. Modeling state-dependent communication between brain regions with switching nonlinear dynamical systems. In *International Conference on Learning Representations*, 2024.

[42] B. Liu, J. Sacks, and M. D. Golub. Accurate identification of communication between multiple interacting neural populations. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*. PMLR, 2025.

[43] N. Mudrik, R. Ly, O. Rübel, and A. S. Charles. CREIMBO: Cross-regional ensemble interactions in multi-view brain observations. In *International Conference on Learning Representations*, 2025.

[44] E. Yezerets, E. Yang, M. B. Ahrens, and A. S. Charles. Behavior-dLDS: A decomposed linear dynamical systems model for neural activity partially constrained by behavior. *Transactions on Machine Learning Research*, 2026.

[45] H. D. Lee, A. Warrington, J. I. Glaser, and S. W. Linderman. Switching autoregressive low-rank tensor models. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

[46] K. D. Harris, A. Aravkin, R. Rao, and B. W. Brunton. Time-varying autoregression with low-rank tensors. *SIAM Journal on Mathematics of Data Science*, 3(2):694–716, 2021.

[47] V. Geadah, A. Nejatbakhsh, D. Lipshutz, J. W. Pillow, and A. H. Williams. Modeling neural activity with conditionally linear dynamical systems. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

[48] A.J. Izenman. Reduced-rank regression for the multivariate linear model. *Journal of Multivariate Analysis*, 5(2):248–264, 1975.

[49] J.B. Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18(2):95–138, 1977.

[50] J. Snoek, H. Larochelle, and R. Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2012.

[51] J. Friedman, T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, Berlin, 2001.

[52] D. J. Tolhurst, J. A. Movshon, and A. F. Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, 23(8):775–785, 1983.

[53] Srinath R., D.A. Ruff, and M.R. Cohen. Attention improves information flow between neuronal populations without changing the communication subspace. *Current Biology*, 31(23):5299–5313.e4, 2021.

[54] O. Gozel and B. Doiron. Between-area communication through the lens of within-area neuronal dynamics. *Science Advances*, 10:eadl6120, 2024.

[55] Amin Zandvakili and Adam Kohn. Coordinated neuronal activity enhances corticocortical communication. *Neuron*, 87(4):827–839, 2015.

[56] E. Gokcen, A. I. Jasper, J. D. Semedo, A. Zandvakili, A. Kohn, C. K. Machens, and B. M. Yu. Disentangling the flow of signals between populations of neurons. *Nature Computational Science*, 2(8):512–525, 2022.

[57] S. Zhang, M. Xu, T. Kamigaki, J. P. Hoang Do, W. C. Chang, S. Jenvay, K. Miyamichi, L. Luo, and Y. Dan. Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345(6197):660–665, 2014.

[58] M. Vinck, R. Batista-Brito, U. Knoblich, and J. A. Cardin. Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, 86(3):740–754, 2015.

[59] S. Ährlund-Richter, Y. Osako, K. R. Jenks, E. Odom, H. Huang, Daniel B. Arnold, and M. Sur. Distinct roles of prefrontal subregion feedback to the primary visual cortex across behavioral states. *Neuron*, 114(3):492–506.e6, 2026.

[60] J. Reimer, E. Froudarakis, C. R. Cadwell, D. Yatsenko, G. H. Denfield, and A. S. Tolias. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, 84(2):355–362, 2014.

[61] H. J. Pi, B. Hangya, D. Kvitsiani, J. I. Sanders, Z. J. Huang, and A. Kepecs. Cortical interneurons that specialize in disinhibitory control. *Nature*, 503(7477):521–524, 2013.

[62] R.Q. Quiroga and S. Panzeri. Extracting information from neuronal populations: information theory and decoding approaches. *Nature Reviews Neuroscience*, 10(3):173–185, 2009.

[63] S. Bernardi, M. K. Benna, M. Rigotti, J. Munuera, S. Fusi, and C. D. Salzman. The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell*, 183(4):954–967.e21, 2020.

[64] D. Dupret, S. Fusi, and S. Panzeri. Neural population activity for memory: Properties, computations, and codes. *Neuron*, 114(3):390–407, 2026.

[65] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 2 edition, 2012.

[66] Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, 6(6):861–867, 1993.

[67] G. J. Broussard, Y. Liang, M. Fridman, E. K. Unger, G. Meng, X. Xiao, N. Ji, L. Petreanu, and L. Tian. In vivo measurement of afferent activity with axon-specific calcium imaging. *Nature Neuroscience*, 21(9):1272–1280, 2018.

[68] C. Meng, J. Zhou, A. Papaneri, T. Peddada, K. Xu, and G. Cui. Spectrally resolved fiber photometry for multi-component analysis of brain circuits. *Neuron*, 98(4):707–717.e4, 2018.

A Glossary

In this paper, we use the following abbreviations:

Abbreviation	Meaning
RRR	Reduced-rank regression.
CS	Communication subspace: the framework of [16], which uses RRR to identify covarying low-dimensional activity patterns of a source and a target population.
AIS	Additive interaction subspaces: the source and target subspaces estimated within the CS framework; renamed to emphasize that they capture additive covariation, not directed communication.
MIC	Multiplicative interaction channel: a triplet of source, target, and modulator patterns, in which the modulator pattern multiplicatively scales the interaction between the source and target patterns.
MICs	Our method, extending the CS framework to infer, within a single model, the subspaces encoding the modulator, the additive interaction subspaces, and the multiplicative interaction channels.
CP	Canonical Polyadic (tensor decomposition into a sum of rank-1 terms).
ALS	Alternating least squares.
EV	Explained variance.
GT	Ground truth.
IQR	Interquartile range.
MLP	Multilayer perceptron.
ReLU	Rectified linear unit.
SEM	Standard error of the mean.
ACA	Anterior cingulate area.
PFC	Prefrontal cortex.
V1, V2	Primary and secondary visual cortex.
VIP	Vasoactive intestinal peptide-expressing interneurons.
PSTH	Peri-stimulus time histogram (time-dependent firing rate averaged over all trials for a given stimulus).
ROI	Region of interest.

B Technical appendices

Throughout the appendix, we use the notation established in the main text: $\mathbf{X} \in \mathbb{R}^{N \times n_X}$ (source), $\mathbf{Y} \in \mathbb{R}^{N \times n_Y}$ (target), $\mathbf{Z} \in \mathbb{R}^{N \times n_Z}$ (modulator), all zero-centered column-wise. We use $\| \cdot \|_F$ for the Frobenius norm, $\langle K, L \rangle_F = \mathrm{tr}(K^\top L)$ for the Frobenius inner product, $\odot$ for the element-wise (Hadamard) product, and $\otimes$ for the outer (tensor) product.

B.1 Introduction to reduced-rank regression

We briefly summarize the ridge-regularized reduced-rank regression (ridge-RRR) estimator used throughout this work. For full derivations, see [48, 17].

Consider the linear model $\mathbf{Y} = \mathbf{X}A + \mathbf{E}$, where $A \in \mathbb{R}^{n_X \times n_Y}$ with $\mathrm{rank}(A) \leq m$. We seek to minimize the ridge-regularized loss subject to a rank constraint:

$$\mathcal{L}_{\mathrm{RRR}}(A) = \|\mathbf{Y} - \mathbf{X}A\|_F^2 + \lambda \|A\|_F^2, \qquad \text{subject to } \mathrm{rank}(A) \leq m. \tag{10}$$

Full-rank ridge estimate. Ignoring the rank constraint, setting $\nabla_A \mathcal{L} = 0$ yields:

$$A_{\mathrm{ridge}} = (\mathbf{X}^\top \mathbf{X} + \lambda I)^{-1} \mathbf{X}^\top \mathbf{Y}. \tag{11}$$

Rank-constrained projection. The key result (see [17], Appendices C-D, for a complete proof) is that incorporating the rank constraint is well approximated by projecting the ridge estimate onto the subspace spanned by the top m eigenvectors of a specific matrix. Concretely, the ridge-RRR solution is:

$$\hat{A} = A_{\text{ridge}}\, V_r V_r^\top, \tag{12}$$

where $V_r \in \mathbb{R}^{n_Y \times m}$ contains the top m eigenvectors of the symmetric positive semidefinite matrix $A_{\text{ridge}}^\top \mathbf{X}^\top \mathbf{X}\, A_{\text{ridge}}$. Equivalently, V_r is obtained by performing PCA on the $N \times n_Y$ matrix of ridge-regularized predictions $\mathbf{X}\, A_{\text{ridge}}$ and retaining the top m components.

Interpretation: source and target subspaces. The solution in Eq. 12 has the factorized form $\hat{A} = U_A V_A^\top$, with:

$$V_A = V_r \in \mathbb{R}^{n_Y \times m}, \qquad U_A = A_{\text{ridge}}\, V_r \in \mathbb{R}^{n_X \times m}. \tag{13}$$

These factors have direct interpretations in terms of the relationship between source and target variables:

- **Source AIS**, $\mathrm{col}(U_A) \subset \mathbb{R}^{n_X}$: the m-dimensional subspace of source patterns that co-vary with the target. Source patterns orthogonal to this subspace are "private" - they have no linear relationship with the target.
- **Target AIS**, $\mathrm{col}(V_A) \subset \mathbb{R}^{n_Y}$: the m-dimensional subspace of target patterns that can be linearly related to the source. Target patterns orthogonal to this subspace are "private" and unexplained by the source.

Note that while the subspaces $\mathrm{col}(U_A)$ and $\mathrm{col}(V_A)$ are uniquely determined by $\hat{A}$, the individual columns of U_A and V_A are not: for any invertible $L \in \mathbb{R}^{m \times m}$, the alternative factorization $\hat{A} = (U_A L)(V_A L^{-\top})^\top$ yields the same product. This rotational ambiguity means that individual "communication dimensions" within the subspace are not identifiable from the model alone.

B.2 From scalar interaction regression to the MICs modulation tensor

In this section, we position the MICs model within the broader landscape of interaction modeling: we start from scalar regression with an interaction term, generalize to the high-dimensional setting where the interaction is encoded by a three-way tensor $\mathcal{C}$, recover MICs as a CP-rank constraint on $\mathcal{C}$, and finally relate it to prior tensor regression. To build intuition, we briefly connect our model to the familiar setting of regression with interaction terms. Consider first a scalar source $x \in \mathbb{R}$, scalar modulator $z \in \mathbb{R}$, and scalar target $y \in \mathbb{R}$. A standard interaction model is:

$$y = \beta_x x + \beta_z z + \beta_{xz}\, xz + \varepsilon, \tag{14}$$

where $\beta_{xz}\, xz$ is the interaction term: the effect of x on y depends linearly on z, and vice versa.

Our model generalizes this to the high-dimensional setting where $x \in \mathbb{R}^{n_X}$, $y \in \mathbb{R}^{n_Y}$, $z \in \mathbb{R}^{n_Z}$. Without any dimensionality reduction, the target for observation n would be:

$$y_n = A^\top x_n + B^\top z_n + \sum_{j=1}^{n_X} \sum_{l=1}^{n_Z} (x_n)_j\, (z_n)_l\, c_{jl} + \varepsilon_n, \tag{15}$$

where $c_{jl} \in \mathbb{R}^{n_Y}$ is a vector of interaction coefficients for source dimension j and modulator dimension l. This full interaction model has $n_X \times n_Z \times n_Y$ free parameters in the interaction term alone, which is prohibitively large in typical neuroscience settings (e.g., $n_X = n_Y = n_Z = 50$ gives $125{,}000$ interaction parameters).

Equivalently, collecting the vectors $c_{jl} \in \mathbb{R}^{n_Y}$ into a single three-way array, i.e. *the full-rank modulation tensor*, $\mathcal{C} \in \mathbb{R}^{n_Y \times n_X \times n_Z}$ with entries $\mathcal{C}_{i,j,l} = (c_{jl})_i$, the modulation term in Eq. 15 is the contraction of $\mathcal{C}$ against the source and modulator vectors of observation n:

$$\sum_{j=1}^{n_X} \sum_{l=1}^{n_Z} c_{jl}\, (x_n)_j\, (z_n)_l \;=\; \mathcal{C} \times_2 x_n \times_3 z_n \;\in\; \mathbb{R}^{n_Y}. \tag{16}$$

Here $\times_k$ denotes the *mode-k vector product*, i.e., contraction of a tensor along its k-th index: for $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ and $u \in \mathbb{R}^{I_k}$, $(T \times_k u)_{i_1,\ldots,i_{k-1},i_{k+1},\ldots,i_N} = \sum_{i_k=1}^{I_k} T_{i_1,\ldots,i_N}\, u_{i_k}$ [33]. Mode 2 (size n_X) is thus contracted with the source x_n and mode 3 (size n_Z) with the modulator z_n, leaving mode 1 (size n_Y) as the output. Equivalently, $\mathcal{C}$ is evaluated on the rank-1 predictor $x_n \otimes z_n \in \mathbb{R}^{n_X \times n_Z}$ formed by source and modulator at observation n.

The MICs model (Eq. 5) greatly reduces the number of estimated parameters by imposing a low-rank structure on $\mathcal{C}$. Each MIC i contributes a rank-1 interaction:

$$y_n = A^\top x_n + B^\top z_n + \sum_{i=1}^{m_C} w_{Y,i}(w_{X,i}^\top x_n)(w_{Z,i}^\top z_n) + \varepsilon_n. \tag{17}$$

This is exactly Eq. 15 with the three-way array of interaction coefficients constrained to have the CP decomposition $c_{jl} = \sum_{i=1}^{m_C}(w_{X,i})_j (w_{Z,i})_l\, w_{Y,i}$, reducing the number of interaction parameters from $n_X \times n_Z \times n_Y$ to $m_C(n_X + n_Z + n_Y)$. For $m_C = 3$ and $n_X = n_Y = n_Z = 50$, this is a compression from $125{,}000$ to 450 parameters. Importantly, the low-rank structure is not merely a computational convenience: it provides interpretability. Each channel defines a specific source axis ($w_{X,i}$), modulator axis ($w_{Z,i}$), and target axis ($w_{Y,i}$), making it possible to ask which specific source patterns have their relationship with which specific target patterns modulated by which specific modulator patterns. In the scalar case (Eq. 14), there is only one such "channel" ($m_C = 1$, with $w_X = w_Z = w_Y = 1$, and the channel strength equals β_{xz}).

In isolation, the modulation term is a specific instance of the tensor-on-tensor regression framework of [38], with per-observation predictor tensor $x_n \otimes z_n \in \mathbb{R}^{n_X \times n_Z}$ and outcome $y_n \in \mathbb{R}^{n_Y}$. The rank-1 structure of the predictor - the outer product of two distinct observed variables rather than a single multiway array - is what endows the CP decomposition with its bilinear-interaction interpretation. The full MICs model (Eq. 5) embeds this term inside an additive RRR framework with residualized source $X' = X - ZD$, which falls outside [38]: Lock's framework fits a single coefficient tensor governing one predictor, and does not address the disentangling of additive from multiplicative effects of distinct predictors X and Z. Beyond the model itself, the hierarchical pipeline (Algorithm 1) handles the simultaneous presence of additive and multiplicative effects of Z under shared cross-validation folds, and the geometric decomposition (Section 3.4) exploits the existence of a baseline AIS operator A against which each MIC is projected - neither of which has a counterpart in [38].

B.3 Uniqueness of MICs: Kruskal's theorem

The use of *channel* rather than *subspace* to describe MICs reflects an algebraic distinction from AIS. Reduced-rank matrix factorizations identify column subspaces but not individual factor columns, which carry rotational indeterminacy. The three-way tensor structure of $\mathcal{C}$ breaks this indeterminacy: under a condition on the Kruskal ranks of the factor matrices, the individual rank-1 components of the CP decomposition are essentially unique, and identifiable up to permutation and per-channel rescaling (Kruskal's theorem). We state the theorem, verify the condition in our setting, and note its practical consequences for MIC analysis.

B.3.1 Kruskal rank and theorem statement

Definition B.1. The *Kruskal rank* of a matrix $M \in \mathbb{R}^{n \times r}$, denoted $k(M)$, is the largest integer j such that every subset of j columns of M is linearly independent.

Note that $k(M) \le \operatorname{rank}(M) \le \min(n, r)$, with equality $k(M) = \operatorname{rank}(M)$ when every subset of $\operatorname{rank}(M)$ columns is linearly independent. For a matrix whose columns are drawn independently from a continuous distribution on $\mathbb{R}^n$ with $n \ge r$, $k(M) = r$ with probability one.

Theorem B.2 (Kruskal, 1977 [49])**.** *Let $\mathcal{C} = \sum_{i=1}^{m_C} w_{Y,i} \otimes w_{X,i} \otimes w_{Z,i}$ be a rank-m_C CP decomposition with factor matrices $W_Y \in \mathbb{R}^{n_Y \times m_C}$, $W_X \in \mathbb{R}^{n_X \times m_C}$, $W_Z \in \mathbb{R}^{n_Z \times m_C}$. If*

$$k(W_Y) + k(W_X) + k(W_Z) \ge 2m_C + 2, \tag{18}$$

then the decomposition is essentially unique*: any other rank-m_C decomposition is related to it by a permutation of the channels and per-channel rescaling.*

In settings where, $n_X, n_Y, n_Z \gg m_C$ and the weight vectors are estimated from data with continuous noise, generically $k(W_Y) = k(W_X) = k(W_Z) = m_C$. Eq. 18 then reduces to $3m_C \ge 2m_C + 2$, i.e.,

$m_C \geq 2$ (uniqueness for $m_C = 1$ is trivial up to an overall scalar). This argument breaks when any factor dimension is small relative to m_C. For example, $k(W_Z) = 1$ for a scalar modulator ($n_Z = 1$), so Eq. 18 fails for all $m_C \geq 2$.

B.3.2 Practical implications for MIC analysis

Many natural analyses with AIS are subspace-level - for example, predicting the target response evoked by a given source pattern - and do not require identifiability of individual factor columns. Kruskal's theorem provides a stronger guarantee that becomes relevant when multiple MICs with distinct geometries coexist in the same population. Without column-level identifiability, two channels with different geometric profiles (e.g., one predominantly α_{ss}, another predominantly α_{ns}) could be rotated into mixtures, blurring their distinct alignment with the baseline AIS. Theorem B.2 ensures that this cannot happen: each channel retains its individual geometry. We do not fully exploit this property in the present manuscript - our current analyses summarize geometric structure at the level of MIC-averaged indices - but it opens the door to per-channel comparative analyses (e.g., contrasting the biological correlates of channels with different α profiles) that we view as a promising direction for future work.

The CP parametrization has an inherent per-channel scaling ambiguity: $(w_{X,i}, w_{Z,i}, w_{Y,i})$ and $(w_{X,i}/\beta,\ \beta\gamma\, w_{Z,i},\ w_{Y,i}/\gamma)$ define the same channel, since the outer product is multilinear. Under the identifiability conditions of Theorem B.2, this is the *only* remaining ambiguity. We resolve this by extracting unit-norm directions and absorbing the magnitude into a single *channel strength* $\sigma_i = \|w_{X,i}\| \cdot \|w_{Z,i}\| \cdot \|w_{Y,i}\|$, giving $\mathcal{C} = \sum_i \sigma_i\, \bar{w}_{Y,i} \otimes \bar{w}_{X,i} \otimes \bar{w}_{Z,i}$. This normalization is used throughout the paper when displaying or comparing MIC weights and scores (e.g., Fig. 4c, Fig. S10j-k). It does not affect the geometry indices α (which depend only on directions) or the channel's contribution to $\mathrm{EV}_{\mathrm{mod}}$.

B.4 The CP Frobenius penalty

The modulation tensor $\mathcal{C} \in \mathbb{R}^{n_Y \times n_X \times n_Z}$, with CP decomposition $\mathcal{C} = \sum_{i=1}^{m_C} w_{Y,i} \otimes w_{X,i} \otimes w_{Z,i}$, has squared Frobenius norm:

$$
\begin{aligned}
\|\mathcal{C}\|_F^2 &= \sum_{j,k,l} \mathcal{C}_{jkl}^2 = \sum_{j,k,l} \left(\sum_{i=1}^{m_C} (w_{Y,i})_j\, (w_{X,i})_k\, (w_{Z,i})_l \right)^2 \\
&= \sum_{j,k,l} \sum_{i=1}^{m_C} \sum_{r=1}^{m_C} (w_{Y,i})_j (w_{Y,r})_j \cdot (w_{X,i})_k (w_{X,r})_k \cdot (w_{Z,i})_l (w_{Z,r})_l \\
&= \sum_{i,r} (w_{Y,i}^\top w_{Y,r})(w_{X,i}^\top w_{X,r})(w_{Z,i}^\top w_{Z,r}) \\
&= \sum_{i,r} (G_Y)_{ir}\, (G_X)_{ir}\, (G_Z)_{ir} \\
&= \langle G_Y,\, G_X \odot G_Z \rangle_F = \mathrm{tr}\big(G_Y^\top (G_X \odot G_Z)\big)\,,
\end{aligned}
\tag{19}
$$

where $G_Y = W_Y^\top W_Y$, $G_X = W_X^\top W_X$, $G_Z = W_Z^\top W_Z \in \mathbb{R}^{m_C \times m_C}$ are the Gram matrices of the factor matrices, and the last line uses the identity $\sum_{ir} A_{ir} B_{ir} C_{ir} = \langle A, B \odot C \rangle_F$ for the element-wise product. This expression is efficient to compute (it involves only $m_C \times m_C$ matrices) and is used in the regularization term of the modulation loss (Eq. 8 in the main text).

B.5 ALS derivation for the modulation term

In this section, we derive ALS for the modulation loss of our model (Eq. 8). ALS for ridge-regularized CP-rank tensor regression is an established tool [33, 37, 38]. The closed-form updates derived below follow from specializing this framework - specifically the non-separable L_2 penalty $\lambda_C \|\mathcal{C}\|_F^2$ on the CP coefficient tensor of [38] - to the bilinear structure of Eq. 8. We provide the derivation for self-containedness and summarize it in algorithm 2.

The modulation loss of our model with the penalty term expressed as in Eq. B.4 is:

$$\mathcal{L}(W_X, W_Z, W_Y) = \|\mathbf{Y}'' - \Phi\, W_Y^\top\|_F^2 + \lambda_C \operatorname{tr}\big(G_Y^\top (G_X \odot G_Z)\big)\,, \tag{20}$$

where $\Phi = (\mathbf{X}' W_X) \odot (\mathbf{Z} W_Z) \in \mathbb{R}^{N \times m_C}$, and $\mathbf{Y}'' = \mathbf{Y}' - \mathbf{X}' A$ are the residuals from the additive model. Each ALS sweep updates W_Y, then W_X, then W_Z, while holding the other two fixed, similarly to the ALS procedures for CP-rank-constrained tensor regression developed in [37, 38].

Expanding the first terms in the RHS of Eq.20 yields:

$$\|\mathbf{Y}'' - \Phi W_Y^\top\|_F^2 = \operatorname{tr}(\mathbf{Y}''^\top \mathbf{Y}'') - 2\operatorname{tr}(W_Y \Phi^\top \mathbf{Y}'') + \operatorname{tr}(W_Y \Phi^\top \Phi W_Y^\top), \tag{21}$$

and, in the following, we refer to the two terms that depend on the factor matrices as the cross-term $-2\operatorname{tr}(W_Y \Phi^\top \mathbf{Y}'')$ and the quadratic term $\operatorname{tr}(W_Y \Phi^\top \Phi W_Y^\top)$.

B.5.1 Update for W_Y

With W_X and W_Z (and hence Φ, G_X, G_Z) fixed, the loss is quadratic in W_Y. Taking the derivative of Eq.20 with respect to $W_Y \in \mathbb{R}^{n_Y \times m_C}$ and setting it to zero yields:

$$\frac{\partial \mathcal{L}}{\partial W_Y} = -2\,\mathbf{Y}''^\top \Phi + 2\, W_Y \Phi^\top \Phi + 2\lambda_C\, W_Y (G_X \odot G_Z). \tag{22}$$

The derivative of the penalty term (i.e. the second term in the RHS of Eq.20) follows from $\frac{\partial}{\partial W_Y}\operatorname{tr}(W_Y^\top W_Y (G_X \odot G_Z)) = 2\, W_Y (G_X \odot G_Z)$, since $G_X \odot G_Z$ is symmetric. Setting the $\frac{\partial \mathcal{L}}{\partial W_Y}$ derivative to zero gives:

$$\boxed{W_Y^\top = \big(\Phi^\top \Phi + \lambda_C\,(G_X \odot G_Z)\big)^{-1}\, \Phi^\top \mathbf{Y}''.} \tag{23}$$

This is an $m_C \times m_C$ linear system (one right-hand side per target dimension), which is inexpensive to solve.

B.5.2 Update for W_X

With W_Z and W_Y fixed, the columns of W_X are coupled through both the Hadamard product in Φ and the CP penalty. We derive the gradient via total differentials.

Define $K_Z = \mathbf{Z} W_Z \in \mathbb{R}^{N \times m_C}$ and $K_Y = \mathbf{Y}'' W_Y \in \mathbb{R}^{N \times m_C}$. Considering a perturbation $W_X \to W_X + dW_X$, we compute the three contributions to $d\mathcal{L}$ coming from the cross-term, the quadratic term and the penalty term respectively:

$$d\mathcal{L} = d\mathcal{L}_{\text{cross}} + d\mathcal{L}_{\text{quad}} + d\mathcal{L}_{\text{pen}}. \tag{24}$$

(i) Cross term $-2\operatorname{tr}(W_Y \Phi^\top \mathbf{Y}'')$: Under $d\Phi = (\mathbf{X}'\, dW_X) \odot K_Z$ and using the Hadamard product identity $\langle A, B \odot C\rangle_F = \langle A \odot B, C\rangle_F$:

$$-2\operatorname{tr}(W_Y\, d\Phi^\top \mathbf{Y}'') = -2\operatorname{tr}\big((K_Y \odot K_Z)^\top \mathbf{X}'\, dW_X\big)\,. \tag{25}$$

(ii) Quadratic term $\operatorname{tr}(G_Y \Phi^\top \Phi)$:

$$2\operatorname{tr}(G_Y\, \Phi^\top d\Phi) = 2\operatorname{tr}\big((\Phi G_Y \odot K_Z)^\top \mathbf{X}'\, dW_X\big)\,. \tag{26}$$

(iii) Penalty term: Since $dG_X = dW_X^\top W_X + W_X^\top dW_X$,

$$\lambda_C\, d\big[\operatorname{tr}(G_Y^\top (G_X \odot G_Z))\big] = 2\lambda_C \operatorname{tr}\Big(\big[W_X (G_Y \odot G_Z)\big]^\top\, dW_X\Big)\,. \tag{27}$$

Therefore,

$$d\mathcal{L} = -2\operatorname{tr}\big((K_Y \odot K_Z)^\top \mathbf{X}'\, dW_X\big) + 2\operatorname{tr}\big((\Phi G_Y \odot K_Z)^\top \mathbf{X}'\, dW_X\big) + 2\lambda_C \operatorname{tr}\Big(\big[W_X (G_Y \odot G_Z)\big]^\top\, dW_X\Big)\,. \tag{28}$$

Combining and identifying the gradient of the scalar loss via $d\mathcal{L} = \operatorname{tr}(\nabla_{W_X} \mathcal{L}^\top\, dW_X)$:

$$\frac{\partial \mathcal{L}}{\partial W_X} = 2\,\mathbf{X}'^\top [(\Phi G_Y - K_Y) \odot K_Z] + 2\lambda_C\, W_X (G_Y \odot G_Z). \tag{29}$$

Block linear system. Setting Eq. 29 to zero does not yield a decoupled update for the columns of W_X, because the columns are coupled through ΦG_Y and through the penalty term. Let $L = G_Y \odot G_Z$ and $H = \mathbf{X'}^\top (K_Y \odot K_Z) \in \mathbb{R}^{n_X \times m_C}$. Writing $w_r = W_X^{(:,r)}$, $k_{Z,r} = K_Z^{(:,r)} \in \mathbb{R}^N$, and $h_j = H^{(:,j)}$, the stationarity condition for the j-th column is

$$\sum_{r=1}^{m_C} \left[(G_Y)_{rj}\, \mathbf{X'}^\top \mathrm{diag}(k_{Z,r} \odot k_{Z,j})\, \mathbf{X'} + \lambda_C\, L_{rj}\, I_{n_X} \right] w_r = h_j. \tag{30}$$

where I_{n_X} denotes the $n_X \times n_X$ identity matrix. Stacking the columns of W_X yields the linear system

$$\boxed{\mathbf{M}_X \,\mathrm{vec}(W_X) = \mathrm{vec}(H),} \tag{31}$$

where $\mathbf{M}_X \in \mathbb{R}^{n_X m_C \times n_X m_C}$ is a symmetric block matrix with (j, r)-th block ($n_X \times n_X$ each)

$$\mathbf{M}_X^{(jr)} = (G_Y)_{rj}\, \mathbf{X'}^\top \mathrm{diag}(k_{Z,r} \odot k_{Z,j})\, \mathbf{X'} + \lambda_C\, L_{rj}\, I_{n_X}. \tag{32}$$

The diagonal matrix need not be formed explicitly: for any $v \in \mathbb{R}^N$,

$$\mathbf{X'}^\top \mathrm{diag}(v)\, \mathbf{X'} = \mathbf{X'}^\top \big(\mathrm{diag}(v) \mathbf{X'}\big),$$

i.e. by scaling the rows of $\mathbf{X'}$ by v before left-multiplying by $\mathbf{X'}^\top$.

B.5.3 Update for W_Z

By symmetry between $\mathbf{X'}$ and $\mathbf{Z}$ in the Hadamard product $\Phi = (\mathbf{X'} W_X) \odot (\mathbf{Z} W_Z)$, the update for W_Z is obtained by swapping $(\mathbf{X'}, W_X, G_X)$ with $(\mathbf{Z}, W_Z, G_Z)$. Defining $K_X = \mathbf{X'} W_X$ (using the just-updated W_X), $L' = G_Y \odot G_X$, and $H' = \mathbf{Z}^\top (K_Y \odot K_X)$:

$$\boxed{\mathbf{M}_Z \,\mathrm{vec}(W_Z) = \mathrm{vec}(H'),} \tag{33}$$

with (j, r)-th block

$$\mathbf{M}_Z^{(jr)} = (G_Y)_{rj}\, \mathbf{Z}^\top \mathrm{diag}(k_{X,r} \odot k_{X,j})\, \mathbf{Z} + \lambda_C\, L'_{rj}\, I_{n_Z},$$

where $k_{X,r} = K_X^{(:,r)}$.

It is worth noticing that the ALS objective in Eq. 20 is non-convex. In principle, different initializations may converge to different local optima. Empirically, on the ACA-VIP data both $\mathrm{EV}_{\mathrm{mod}}$ and the recovered MIC geometry were stable across random ALS initializations (Fig. S11; Appendix B.15).

B.6 Hierarchical model fitting algorithm

For clarity, we summarize the full fitting pipeline in Algorithm 1. The algorithm mirrors the structure of the implementation described in the main text and in the code: first fit the additive effects of the modulator on source and target, then fit the baseline AIS on the residualized variables, and finally fit the MICs on the residuals of the baseline model.

In the pseudocode below, calligraphic symbols denote collections of fold-specific quantities across cross-validation splits. For example, $\mathcal{B} = \{B^{(k)}\}_{k=1}^K$ denotes the collection of $Z \to Y$ regression matrices obtained from each fold k, and $\mathcal{Z} = \{(Z_{\mathrm{train}}^{(k)}, Z_{\mathrm{test}}^{(k)})\}_{k=1}^K$ denotes the collection of train/test splits of Z. Algebraic expressions involving calligraphic variables indicate fold-wise operations. For instance,

$$\mathcal{Z}\mathcal{B} = \{(Z_{\mathrm{train}}^{(k)} B^{(k)},\ Z_{\mathrm{test}}^{(k)} B^{(k)})\}_{k=1}^K. \tag{34}$$

Algorithm 1 gives the global structure of the fitting pipeline. The final step, CFIT, requires solving the low-rank tensor regression problem for the modulation term. We summarize the complete ALS procedure used inside CFIT in Algorithm 2. The weight matrices are initialized randomly (from a standard normal distribution), and sweeps are repeated until the relative change in the loss falls below a tolerance ϵ or a maximum number of sweeps $maxIter$ is reached (in our implementation we set $\epsilon = 10^{-5}$ and $maxIter = 100$).

The computational cost per sweep is dominated by assembling and solving the $n_X m_C \times n_X m_C$ and $n_Z m_C \times n_Z m_C$ block systems for W_X and W_Z: $\mathcal{O}((n_X m_C)^2 N + (n_X m_C)^3)$ and analogously for W_Z. The W_Y update costs only $\mathcal{O}(N m_C n_Y + m_C^3)$.

Algorithm 1 Hierarchical fitting procedure for MICs

function FULLFIT(X, Y, Z)
  Split data into K cross-validation folds:

$$\{(X^{(k)}_{\text{train}}, X^{(k)}_{\text{test}}, Y^{(k)}_{\text{train}}, Y^{(k)}_{\text{test}}, Z^{(k)}_{\text{train}}, Z^{(k)}_{\text{test}})\}_{k=1}^{K}$$

  for $k = 1, \ldots, K$ **do**
    z-score the k-th train and test sets using the mean and standard deviation of the k-th training set
  end for
  $(\mathcal{B}, m_B) \leftarrow$ RRRFIT$(\mathcal{Z}, \mathcal{Y})$
  $(\mathcal{D}, m_D) \leftarrow$ RRRFIT$(\mathcal{Z}, \mathcal{X})$
  $\mathcal{X}' \leftarrow \mathcal{X} - \mathcal{Z}\mathcal{D}$
  $\mathcal{Y}' \leftarrow \mathcal{Y} - \mathcal{Z}\mathcal{B}$
  $(\mathcal{A}, m_A) \leftarrow$ RRRFIT$(\mathcal{X}', \mathcal{Y}')$
  $\mathcal{Y}'' \leftarrow \mathcal{Y}' - \mathcal{X}'\mathcal{A}$
  $(\mathcal{W}_X, \mathcal{W}_Z, \mathcal{W}_Y, m_C) \leftarrow$ CFIT$(\mathcal{X}', \mathcal{Y}'', \mathcal{Z})$
  return $(\mathcal{A}, \mathcal{B}, \mathcal{W}_X, \mathcal{W}_Z, \mathcal{W}_Y, \mathcal{D}, m_A, m_B, m_C, m_D)$
end function

function RRRFIT($\mathcal{X}, \mathcal{Y}$)
  for $m = 1, \ldots, \min(n_X, n_Y)$ **do**
    $\lambda^* \leftarrow$ ridge penalty λ that maximizes average fold-wise EV of rank-m $\mathcal{X} \to \mathcal{Y}$ regression matrix $\mathcal{A}(m, \lambda)$ (MATLAB `bayesopt`)
    $\mathrm{EV}(m) \leftarrow$ across-fold EV for $\lambda = \lambda^*$
    $\mathrm{sem}(m) \leftarrow$ across-fold standard error of mean EV for $\lambda = \lambda^*$
    $\mathcal{A}(m) \leftarrow \mathcal{A}(m, \lambda^*)$
  end for
  $m^* \leftarrow$ smallest m such that $\mathrm{EV}(m) + \mathrm{sem}(m_{\text{max}}) \geq \mathrm{EV}(m_{\text{max}})$, where m_{max} is the rank of max EV
  return $(\mathcal{A}(m^*), m^*)$
end function

function CFIT($\mathcal{X}', \mathcal{Y}'', \mathcal{Z}$)
  for $m = 1, \ldots,$ max rank **do**
    $\lambda^* \leftarrow$ Frobenius penalty λ that maximizes average fold-wise EV of rank-m MICs $(\mathcal{W}_X(m, \lambda), \mathcal{W}_Z(m, \lambda), \mathcal{W}_Y(m, \lambda))$ obtained using Algorithm 2 (MATLAB `bayesopt`)
    $\mathrm{EV}(m) \leftarrow$ across-fold EV for $\lambda = \lambda^*$
    $\mathrm{sem}(m) \leftarrow$ across-fold standard error of mean EV for $\lambda = \lambda^*$
    $(\mathcal{W}_X(m), \mathcal{W}_Z(m), \mathcal{W}_Y(m)) \leftarrow (\mathcal{W}_X(m, \lambda^*), \mathcal{W}_Z(m, \lambda^*), \mathcal{W}_Y(m, \lambda^*))$
  end for
  $m^* \leftarrow$ smallest m such that $\mathrm{EV}(m) + \mathrm{sem}(m_{\text{max}}) \geq \mathrm{EV}(m_{\text{max}})$, where m_{max} is the rank of max EV
  return $(\mathcal{W}_X(m^*), \mathcal{W}_Z(m^*), \mathcal{W}_Y(m^*), m^*)$
end function

Modulatory explained variance Denoting by $\hat{\mathbf{Y}}^{(k)}_{\text{add}} = {\mathbf{X}'_{\text{test}}}^{(k)} A^{(k)} + \mathbf{Z}^{(k)}_{\text{test}} B^{(k)}$ the joint additive model prediction on the test set of fold k, and by $\hat{\mathbf{Y}}^{(k)}_{\text{full}} = \hat{\mathbf{Y}}^{(k)}_{\text{add}} + \left[({\mathbf{X}'_{\text{test}}}^{(k)} W_X^{(k)}) \odot (\mathbf{Z}^{(k)}_{\text{test}} W_Z^{(k)}) \right] W_Y^{(k)\top}$ the full model prediction, the modulatory explained variance (Sec. 3.3) is

$$\mathrm{EV}_{\text{mod}} = \frac{1}{K} \sum_{k=1}^{K} \left[R^2_{\text{full}}(k) - R^2_{\text{add}}(k) \right], \tag{35}$$

where $R^2(k)$ is the coefficient of determination on the test set of fold k. Analogously, EV_{AIS} is the fold-averaged R^2 increment of $\hat{\mathbf{Y}}^{(k)}_{\text{add}}$ over the modulator-only prediction $\mathbf{Z}^{(k)}_{\text{test}} B^{(k)}$.

Algorithm 2 ALS for fitting MICs

Require: Residuals $\mathbf{Y}'' \in \mathbb{R}^{N \times n_Y}$, residualized source $\mathbf{X}' \in \mathbb{R}^{N \times n_X}$, modulator $\mathbf{Z} \in \mathbb{R}^{N \times n_Z}$, rank m_C, penalty λ_C, tolerance ϵ, max sweeps T
Ensure: Weight matrices $W_X \in \mathbb{R}^{n_X \times m_C}$, $W_Z \in \mathbb{R}^{n_Z \times m_C}$, $W_Y \in \mathbb{R}^{n_Y \times m_C}$

Initialize W_X, W_Z, W_Y randomly
Compute $K_X \leftarrow \mathbf{X}' W_X,\ K_Z \leftarrow \mathbf{Z} W_Z,\ \Phi \leftarrow K_X \odot K_Z$
Compute Gram matrices $G_X \leftarrow W_X^\top W_X,\ G_Z \leftarrow W_Z^\top W_Z$
for $t = 1, \ldots, T$ **do**
  Update W_Y: $W_Y^\top \leftarrow \left(\Phi^\top \Phi + \lambda_C \left(G_X \odot G_Z\right)\right)^{-1} \Phi^\top \mathbf{Y}''$ ▷ Eq. 23
  Update $G_Y \leftarrow W_Y^\top W_Y,\ K_Y \leftarrow \mathbf{Y}'' W_Y$
  Update W_X: Assemble $\mathbf{M}_X$ and H; solve $\mathbf{M}_X \operatorname{vec}(W_X) = \operatorname{vec}(H)$ ▷ Eq. 31
  Update $K_X \leftarrow \mathbf{X}' W_X,\ \Phi \leftarrow K_X \odot K_Z,\ G_X \leftarrow W_X^\top W_X$
  Update W_Z: Assemble $\mathbf{M}_Z$ and H'; solve $\mathbf{M}_Z \operatorname{vec}(W_Z) = \operatorname{vec}(H')$ ▷ Eq. 33
  Update $K_Z \leftarrow \mathbf{Z} W_Z,\ \Phi \leftarrow K_X \odot K_Z,\ G_Z \leftarrow W_Z^\top W_Z$
  Compute loss $\mathcal{L}^{(t)}$ via Eq. 20
  if $|\mathcal{L}^{(t)} - \mathcal{L}^{(t-1)}| / |\mathcal{L}^{(t-1)}| < \epsilon$ **then**
    break
  end if
end for

B.7 Geometric decomposition of MICs

In this section, we derive the four-component decomposition of each channel ΔA_i and the resulting form of the geometry indices:

$$\alpha_{ss}^{(i)} = \frac{\|\Delta A_i^{ss}\|_F^2}{\|\Delta A_i\|_F^2}, \qquad \alpha_{ns}^{(i)} = \frac{\|\Delta A_i^{ns}\|_F^2}{\|\Delta A_i\|_F^2}, \qquad \alpha_{sn}^{(i)} = \frac{\|\Delta A_i^{sn}\|_F^2}{\|\Delta A_i\|_F^2}, \qquad \alpha_{nn}^{(i)} = \frac{\|\Delta A_i^{nn}\|_F^2}{\|\Delta A_i\|_F^2}. \tag{36}$$

The results in this section follow from standard algebra of orthogonal projectors (see, e.g., [65]) and are reported here for clarity and completeness; we do not claim them as novel theoretical contributions.

B.7.1 Orthogonal decomposition

The decomposition $M = M^{ss} + M^{ns} + M^{sn} + M^{nn}$ (Eq. 9) follows by inserting $I = P + (I - P)$ on both sides of M:

$$M = \left[P_X + (I - P_X)\right] M \left[P_Y + (I - P_Y)\right]. \tag{37}$$

The pairwise orthogonality of the four components under the Frobenius inner product is a consequence of the idempotence of orthogonal projectors. For any two components that differ in their source-side projector (e.g., M^{ss} and M^{ns}):

$$\langle M^{ss}, M^{ns} \rangle_F = \operatorname{tr}\left(P_Y M^\top \underbrace{P_X (I - P_X)}_{=0} M P_Y \right) = 0, \tag{38}$$

and for components that differ in their target-side projector (e.g., M^{ss} and M^{sn}), the analogous identity using $(I - P_Y)P_Y = 0$ and the cyclic property of the trace applies. The Pythagorean norm decomposition $\|M\|_F^2 = \|M^{ss}\|_F^2 + \|M^{ns}\|_F^2 + \|M^{sn}\|_F^2 + \|M^{nn}\|_F^2$ is an immediate consequence.

B.7.2 Explicit form of α indices for rank-1 perturbations

For a rank-1 matrix $M = w_X w_Y^\top$, the four components remain rank-1 (or zero):

$$M^{ss} = (P_X w_X)(P_Y w_Y)^\top, \quad M^{ns} = ((I - P_X) w_X)(P_Y w_Y)^\top, \quad \text{etc.}, \tag{39}$$

since orthogonal projectors act linearly: $P_X (w_X w_Y^\top) P_Y = (P_X w_X)(P_Y w_Y)^\top$. Using the identity $\|ab^\top\|_F^2 = \|a\|^2 \|b\|^2$ for rank-1 matrices (which follows from $\|ab^\top\|_F^2 = \operatorname{tr}(ba^\top ab^\top) =$

$(a^\top a)(b^\top b)$), the squared Frobenius norms factorize:

$$\|M^{ss}\|_F^2 = \|P_X w_X\|^2 \|P_Y w_Y\|^2, \qquad \|M\|_F^2 = \|w_X\|^2 \|w_Y\|^2, \tag{40}$$

and analogously for the other components. The geometry indices (Eq. 36) therefore take the factorized form:

$$\alpha_{ss} = \cos^2\theta_X \cdot \cos^2\theta_Y, \qquad \alpha_{ns} = \sin^2\theta_X \cdot \cos^2\theta_Y, \tag{41}$$

$$\alpha_{sn} = \cos^2\theta_X \cdot \sin^2\theta_Y, \qquad \alpha_{nn} = \sin^2\theta_X \cdot \sin^2\theta_Y, \tag{42}$$

where $\cos^2\theta_X = \|P_X w_X\|^2/\|w_X\|^2$ is the squared cosine of the angle between w_X and the source AIS, and $\cos^2\theta_Y = \|P_Y w_Y\|^2/\|w_Y\|^2$ analogously for the target. The Pythagorean identity for projections ensures $\cos^2\theta + \sin^2\theta = 1$ on each side, so the four α values sum to one.

Remark B.3 (Independence from the modulator). Since each MIC i contributes $\Delta A_i(z) = (w_{Z,i}^\top z)\, w_{X,i} w_{Y,i}^\top$, the scalar $(w_{Z,i}^\top z)^2$ appears in both numerator and denominator of each α index and cancels. The geometry indices thus characterize a fixed geometric property of each channel, regardless of the instantaneous value of z or the modulator weight vector $w_{Z,i}$.

Remark B.4 (Two-parameter structure). The product form reveals that the full 2×2 table of α values is determined by just two free parameters: the source alignment $\cos^2\theta_X$ and the target alignment $\cos^2\theta_Y$. Defining $a_X = \cos^2\theta_X$ and $a_Y = \cos^2\theta_Y$:

$$\alpha_{ss} = a_X a_Y, \quad \alpha_{ns} = (1-a_X)a_Y, \quad \alpha_{sn} = a_X(1-a_Y), \quad \alpha_{nn} = (1-a_X)(1-a_Y). \tag{43}$$

B.7.3 Null randomized test of α indices

Raw geometry indices α_k are subject to a systematic bias: finite-sample estimation causes the fitted directions $\mathbf{w}_{X,i}$, $\mathbf{w}_{Y,i}$ to deviate from their true orientations, and the dimensional structure of the problem ($n_X \gg m_A$ in typical recordings) determines where those deviations land - predominantly outside the AIS, inflating components such as α_{ns} and α_{nn} even when the true geometry is confined to α_{ss} (Fig. 3b).

To correct for this, we estimate the expected α values under a *random-direction null*: for each fitted MIC i, we replace $\mathbf{w}_{X,i}$ and $\mathbf{w}_{Y,i}$ with independent draws from the uniform distribution on their respective unit spheres, recompute the four indices, and repeat for N_{shuf} shuffles. The null-subtracted index is

$$\tilde{\alpha}_k^{(i)} = \alpha_k^{(i)} - \langle \alpha_k^{\text{null}} \rangle, \qquad k \in \{ss,\, ns,\, sn,\, nn\}, \tag{44}$$

where $\langle\cdot\rangle$ denotes the mean over shuffles.

Values $\tilde{\alpha}_k > 0$ indicates that the fitted MIC aligns with geometric component k more than expected for a uniformly random direction in the same ambient space - that is, the channel exhibits geometric structure in that component beyond chance. The test is conservative: components whose true weight is small relative to the random-alignment baseline will be dismissed Fig. 3f), in which case constrained-fit EV_{mod} Appendix B.8) provides a more sensitive characterization of MICs geometry.

B.8 Constrained fitting: modulation within or outside baseline subspaces

As described in Section 4.2, we test whether modulatory variance is attributable to specific geometric components by fitting constrained models. The following result shows that constraining MIC factors to specified subspaces reduces, up to a constant, to an unconstrained problem on projected data, so that Algorithm 2 can be reused without modification.

B.8.1 General reduction

Proposition B.5 (Constrained fitting reduction)**.** *Let $Q_X \in \mathbb{R}^{n_X \times d_X}$ and $Q_Y \in \mathbb{R}^{n_Y \times d_Y}$ be semi-orthogonal matrices ($Q_X^\top Q_X = I_{d_X}$, $Q_Y^\top Q_Y = I_{d_Y}$), and consider the modulation loss (Eq. 20) constrained by the parametrization*

$$W_X = Q_X \widetilde{W}_X, \qquad W_Y = Q_Y \widetilde{W}_Y, \tag{45}$$

with $\widetilde{W}_X \in \mathbb{R}^{d_X \times m_C}$ and $\widetilde{W}_Y \in \mathbb{R}^{d_Y \times m_C}$ free. Then the constrained loss is equivalent, up to an additive constant independent of the free parameters, to the unconstrained modulation loss applied to the projected data $(\mathbf{X}'Q_X,\ \mathbf{Y}''Q_Y,\ \mathbf{Z})$.

Proof. Substituting Eq. 45 into the modulation loss and splitting the data-fit term using the fact that the prediction $\hat{\mathbf{Y}}$ lies entirely in $\mathrm{col}(Q_Y)$:

$$\|\mathbf{Y}'' - \hat{\mathbf{Y}}\|_F^2 = \|\mathbf{Y}''Q_Y Q_Y^\top - \hat{\mathbf{Y}}\|_F^2 + \|\mathbf{Y}''(I - Q_Y Q_Y^\top)\|_F^2. \tag{46}$$

The second term is constant in the free parameters. For the first, $Q_Y^\top Q_Y = I_{d_Y}$ gives $\|\mathbf{Y}''Q_Y - [(\mathbf{X}'Q_X\widetilde{W}_X) \odot (\mathbf{Z}W_Z)]\widetilde{W}_Y^\top\|_F^2$. The penalty simplifies analogously: $Q_X^\top Q_X = I_{d_X}$ and $Q_Y^\top Q_Y = I_{d_Y}$ give Gram matrices $\widetilde{G}_X = \widetilde{W}_X^\top\widetilde{W}_X$ and $\widetilde{G}_Y = \widetilde{W}_Y^\top\widetilde{W}_Y$, so $\|\mathcal{C}\|_F^2 = \mathrm{tr}(\widetilde{G}_Y^\top(\widetilde{G}_X \odot G_Z))$. Combining:

$$\mathcal{L}_{\mathrm{con.}}(\widetilde{W}_X, W_Z, \widetilde{W}_Y) = \left\|\mathbf{Y}''Q_Y - \left[(\mathbf{X}'Q_X\widetilde{W}_X) \odot (\mathbf{Z}W_Z)\right]\widetilde{W}_Y^\top\right\|_F^2 + \lambda_C\|\widetilde{\mathcal{C}}\|_F^2, \tag{47}$$

which is the unconstrained modulation loss in the projected variables. □ □

Eq. 47 can be minimized using the same ALS procedure (Algorithm 2). After optimization, the full-space factors are recovered via $W_X = Q_X\widetilde{W}_X$ and $W_Y = Q_Y\widetilde{W}_Y$.

B.8.2 Specific geometric constraints

The geometric components of the α decomposition correspond to specific choices of (Q_X, Q_Y). We list all eight possible constraints below for completeness. In the present paper we use four of these constraints, compared with each other and with the unconstrained fit (Sections 3.4 and 5): the both-in-AIS (α_{ss}-only) fit, the source-in-AIS fit (w_X in the source AIS, w_Y free; forces $\alpha_{ns} = \alpha_{nn} = 0$), the target-in-AIS fit (w_Y in the target AIS, w_X free; forces $\alpha_{sn} = \alpha_{nn} = 0$), and the α_{nn}-only fit (V1-V2 data; Fig. S9c). Comparing cross-validated $\mathrm{EV}_{\mathrm{mod}}$ across nested fits attributes specific geometric components. Rather than testing all pairs, we test only comparisons answering one geometric question each. (i) Contributions of α_{sn} and α_{ns}: the gap between the both-in-AIS and source-in-AIS (target-in-AIS) fits isolates α_{sn} (α_{ns}). (ii) Their difference: source-in-AIS vs. target-in-AIS fits. (iii) Contribution of α_{nn}: no pair of fits differs by α_{nn} alone, as the unconstrained fit adds $\alpha_{sn} + \alpha_{nn}$ to the target-in-AIS fit and $\alpha_{ns} + \alpha_{nn}$ to the source-in-AIS fit. If (i) detects no contribution from one of the two off-diagonal components, a non-significant gap for the comparison adding it together with α_{nn} indicates no detectable α_{nn} (ACA-VIP: unconstrained vs. target-in-AIS) - this reuses the unconstrained fit of the main analysis, requiring no additional model. If instead both off-diagonal components contribute (V1-V2), nested comparisons cannot isolate α_{nn}, and we resort to the more computationally expensive option of fitting an α_{nn}-only model and testing it against the residual-shuffle null (Appendix B.13). This yields four comparisons for ACA-VIP and three for V1-V2 data (Fig. 4d,h). Fig. S7e-h illustrates this for the four individual ground-truth geometries: when $\alpha_{ss} = 1$ the both-in-AIS constraint captures the unconstrained $\mathrm{EV}_{\mathrm{mod}}$; when $\alpha_{ns} = 1$ ($\alpha_{sn} = 1$) only the target-in-AIS (source-in-AIS) constrain captures the unconstrained $\mathrm{EV}_{\mathrm{mod}}$, while constraining the source (target) reduces $\mathrm{EV}_{\mathrm{mod}}$ to zero; when $\alpha_{nn} = 1$ any constraint reduces $\mathrm{EV}_{\mathrm{mod}}$ to zero. The α attribution from this approach is not univocal - nested gaps quantify the variance explained by adding a geometric component on top of a constrained baseline, rather than yielding a closed-form additive decomposition of $\mathrm{EV}_{\mathrm{mod}}$ into the four α terms - but comparison between these constrained models suffices to localize which components carry modulatory variance.

Constraint	Q_X	Q_Y
$\alpha_{ss} = 1$ (same source, same target)	$\tilde{U}_A$	$\tilde{V}_A$
$\alpha_{ns} = 1$ (new source, same target)	$\tilde{U}_A^\perp$	$\tilde{V}_A$
$\alpha_{sn} = 1$ (same source, new target)	$\tilde{U}_A$	$\tilde{V}_A^\perp$
$\alpha_{nn} = 1$ (new source, new target)	$\tilde{U}_A^\perp$	$\tilde{V}_A^\perp$
$\alpha_{nn} = \alpha_{sn} = 0$ (target in AIS)	—	$\tilde{V}_A$
$\alpha_{nn} = \alpha_{ns} = 0$ (source in AIS)	$\tilde{U}_A$	—
$\alpha_{ss} = \alpha_{ns} = 0$ (target out of AIS)	—	$\tilde{V}_A^\perp$
$\alpha_{ss} = \alpha_{sn} = 0$ (source out of AIS)	$\tilde{U}_A^\perp$	—

Here $\tilde{U}_A^\perp \in \mathbb{R}^{n_X \times (n_X - m_A)}$ and $\tilde{V}_A^\perp \in \mathbb{R}^{n_Y \times (n_Y - m_A)}$ are orthonormal bases for the orthogonal complements of the baseline source and target AIS, and "—" indicates that the corresponding factor

is left unconstrained. Constraining to the "same" subspaces projects the data onto m_A-dimensional subspaces, making the ALS block systems much cheaper (block size drops from $n_X \times n_X$ to $m_A \times m_A$).

Remark B.6 (Constrained fitting vs. post-hoc α indices). The constrained model directly optimizes modulation within a specified geometric subspace, whereas the post-hoc α indices decompose an unconstrained solution. These can give different results: the unconstrained fit might place a channel at an intermediate angle (e.g., $\alpha_{ss} \approx \alpha_{ns} \approx 0.5$), whereas the constrained α_{ss} model finds the best purely within-subspace solution. Comparing the cross-validated $\mathrm{EV}_{\mathrm{mod}}$ of constrained versus unconstrained models provides a direct test of whether modulatory variance is attributable to specific geometric components, complementing the descriptive α indices.

B.9 Simulation details

In this section, we provide additional implementation details on all simulations of the paper. Simulations in Figs. 2, 3, S1, and S2, S3, S4, and S7 all share the generative structure

$$\begin{aligned} \mathbf{Z} &= E_{\mathbf{Z}} \in \mathbb{R}^{N \times n_Z} \\ \mathbf{X} &= \mathbf{Z}D + E_{\mathbf{X}} \in \mathbb{R}^{N \times n_X} \\ \mathbf{Y} &= \mathbf{X}A + \mathbf{Z}B + \big[(\mathbf{X}W_X) \odot (\mathbf{Z}W_Z)\big]W_Y^\top + E_{\mathbf{Y}} \in \mathbb{R}^{N \times n_Y}, \end{aligned} \tag{48}$$

but differ in interaction structure (A, B, D, W_X, W_Y, W_Z), population dimensionalities (n_X, n_Y, n_Z), and sample size N. Noise terms $E_{\mathbf{X}}, E_{\mathbf{Y}}, E_{\mathbf{Z}}$ are i.i.d. Gaussian with unit variance unless stated otherwise. The low-rank factors are parametrized as

$$A = U_A V_A^\top, \quad B = U_B V_B^\top, \quad D = U_D V_D^\top,$$

where the columns of $U_A \in \mathbb{R}^{n_X \times m_A}$, $V_A \in \mathbb{R}^{n_Y \times m_A}$, $U_B \in \mathbb{R}^{n_Z \times m_B}$, $V_B \in \mathbb{R}^{n_Y \times m_B}$, $U_D \in \mathbb{R}^{n_Z \times m_D}$, $V_D \in \mathbb{R}^{n_X \times m_D}$, and the MIC weight vectors $\mathbf{w}_X, \mathbf{w}_Y, \mathbf{w}_Z$ are drawn independently and uniformly from the unit sphere, unless the simulation specifies a particular geometry. Throughout, rank selection uses the 1-SEM rule applied to 10-fold cross-validated EV curves.

Modulation strength sweep; Fig. 2(a-e). We first considered a minimal scenario with $n_X = 2$, $n_Y = 1$, $n_Z = 1$, and $N = 10^4$. The AIS matrix is fixed at $A = [1\ 0]^\top$, so that only the first source dimension drives Y additively (independent of Z), while the second drives Y multiplicatively through the MIC, with $\mathbf{w}_X = [0\ 1]^\top \cdot \gamma$, $\mathbf{w}_Y = \mathbf{w}_Z = 1$. $m_B = m_D = 0$ throughout. Modulation strength γ is swept from 0 to 1.2 in steps of 0.2 over 5 independent repeats. AIS EV is additionally evaluated fitting the model on X_1 and X_2 separately, to make explicit which source component is captured by the AIS.

High-dimensional rank recovery; Fig. 2(f, g). We then moved to a high-dimensional setting with $n_Z = n_X = n_Y = 20$ and $N = 2000$. Ground-truth ranks m_A and m_C are swept jointly from 0 to 5 (with $m_B = m_D = 0$ and Z, X drawn as independent Gaussians), with 5 independent simulations per rank pair, resulting in 30 simulations per marginal rank. Rank selection is followed by a permutation significance test on held-out predictions (1000 shuffles; $p < 0.05$ required to retain non-zero rank).

Data efficiency; Fig. 2h-i and Fig. S2. We assessed data efficiency by sweeping dimension $d := n_X = n_Y = n_Z \in \{10, 15, 20\}$ and sample size N over 10 log-spaced values in $[20,\ 20000]$, with $m_A = 2$, $m_C = 1$, $m_B = m_D = 0$, all factor matrices drawn uniformly on the unit sphere, and 30 simulations per (d, N). At each (d, N) the rank-selected fit (Algorithm 1) is compared against a full-rank counterpart in which the modulation tensor or AIS are fit without rank constraint. We report cross-validated AIS EV and $\mathrm{EV}_{\mathrm{mod}}$ (Fig. 2h-i for $d = 10$; Fig. S2a-c, f-h, all d). To probe scaling, both EVs are then normalized by the ground-truth EV computed on each test fold from the true A, W_X, W_Y, W_Z, and replotted against N rescaled by parameter count: N/d for low-rank fits, N/d^2 for full-rank AIS, N/d^3 for full-rank MICs (Fig. S2d-e, i-j).

Geometric decomposition, Fig. 3. We validated the geometric decomposition with $n_X = 50$, $n_Y = 5$, $n_Z = 1$, $m_A = 2$, $m_C = 1$, $N = 1000$, and 30 simulations per scenario. The source MIC weight $\mathbf{w}_X$ is drawn at a fixed angle θ_X to $\mathrm{span}(U_A)$, while the target weight $\mathbf{w}_Y$ is always

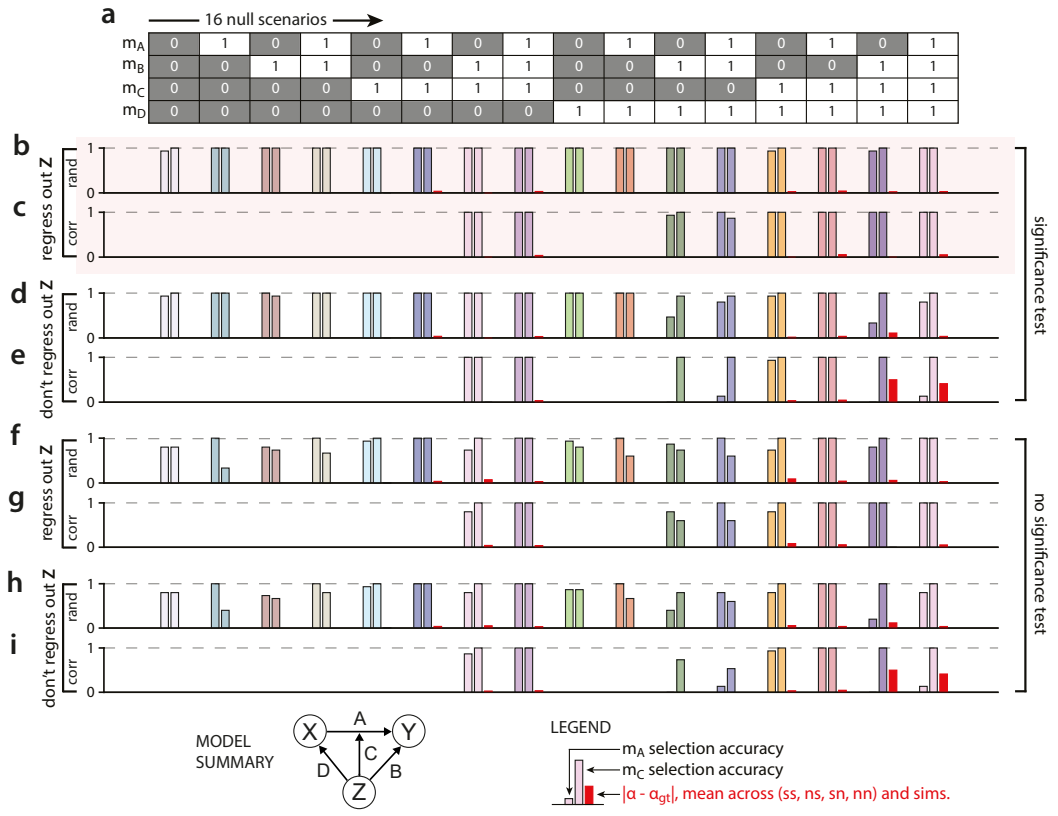

Figure S1: **Pipeline validation across all 16 rank combinations.** (a) Ground-truth rank indicator matrix: $m_A, m_B, m_C, m_D \in \{0, 1\}$ for each of the 16 simulated scenarios (columns). The model schematic (bottom left) is reproduced from Fig. 1b for reference. Panels (b-i) evaluate the MICs fitting pipeline across a $2\times2\times2$ factorial design: whether Z is regressed out of X and Y (Step 1 of Algorithm 1; panels b–c, f–g vs. d–e, h–i), whether the shuffle significance test defined in App. B.13 is applied to $\mathrm{EV}_{\mathrm{mod}}$ (panels b–e vs. f–i), and whether ground-truth interaction subspace vectors are drawn uniformly from the unit sphere (*rand*) or set to be perfectly co-aligned across model terms (*corr*; see below). In each sub-panel, bars for each scenario show m_A selection accuracy (left bar), m_C selection accuracy (center bar), and mean absolute geometry index error, averaged over the four geometric components (ss, ns, sn, nn) and simulations (right bar, red). (b) Pipeline with Step 1 and permutation significance test (*rand* condition): ground-truth weight vectors are drawn independently and uniformly from the unit sphere. (c) Same as (b), *corr* condition: for scenarios in which at least two of m_B, m_C, m_D are non-zero, the relevant low-rank components are set to be perfectly co-aligned. Specifically, when $m_B = m_C = 1$, the MIC and the $Z - Y$ interaction subspace are perfectly aligned in both Z and Y; when $m_D = m_C = 1$, the MIC and the $Z - X$ interaction subspace are perfectly aligned in both Z and X; when $m_B = m_D = 1$, the $Z \to Y$ and $Z \to X$ interaction subspaces are perfectly aligned in Z. (d-e) Same as (b-c) but without regressing Z out of X and Y (Step 1 omitted). (f-g) Same as (b-c) but without the permutation significance test. (h-i) Same as (b-c) but without Step 1 and without the permutation significance test. We ran 15 independent simulations per scenario, with $N = 1000$ samples per simulation and $n_X = n_Y = n_Z = 5$ in all simulations.

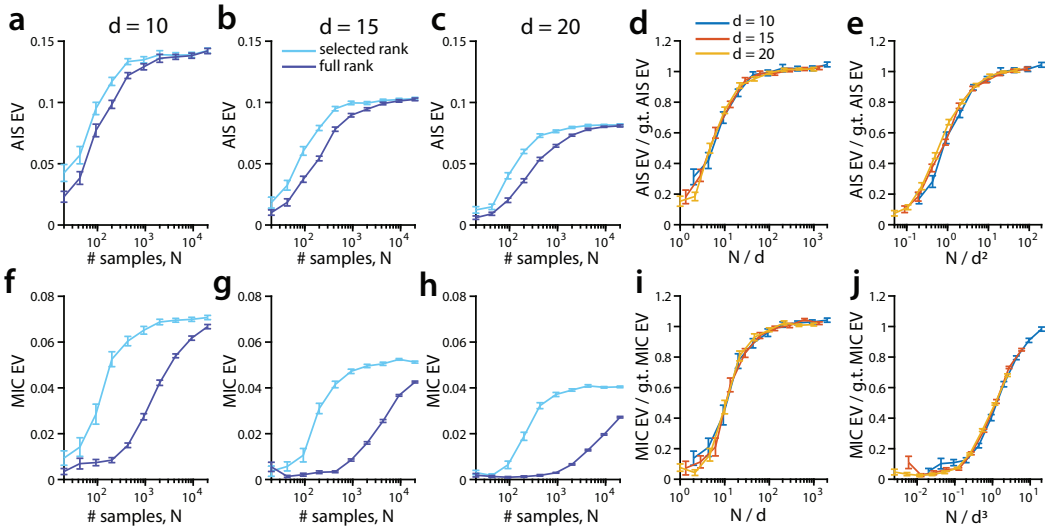

Figure S2: **Data efficiency of MICs relative to full-rank models.** (a) Variance explained by AIS as a function of sample size N, for the rank-selected low-rank model (cyan) and the full-rank model (blue), with $n_X = n_Y = n_Z = 10 =: d$. Lines show mean $\pm$ SEM across 30 independent simulations. (b-c) Same as (a), with $d = 15$ and $d = 20$. (d) Scaling collapse of low-rank AIS explained variance for the three values of d in (a-c). The number of samples is rescaled by a factor proportional to the number of parameters in the low-rank AIS model (specifically, $1/d$) and the explained variance is rescaled by the explained variance of the ground-truth AIS. (e) Same as (d), for the full-rank AIS model ($\propto d^2$ parameters). (f-h) Same as (a-c), for the low-rank and full rank MIC models. (i-j) Same as (d-e) for the low-rank (number of parameters $\propto d$) and full-rank (number of parameters $\propto d^3$) MIC models. Here the explained variance is rescaled by the explained variance of the ground-truth MIC. In all simulations, $m_A = 2$ and $m_C = 1$.

drawn within $\mathrm{span}(V_A)$ ($\theta_Y = 0$). Two ground-truth geometries are tested: $\theta_X = 0$, corresponding to $\alpha_{ss}^{\mathrm{GT}} = 1$ with all other components zero (Fig. 3b-d), and $\theta_X = \arccos\sqrt{3/4}$, corresponding to $\alpha_{ss}^{\mathrm{GT}} = 3/4$, $\alpha_{ns}^{\mathrm{GT}} = 1/4$ (Fig. 3e-g). The shuffle null for the α indices is constructed from 1000 random draws of $(\mathbf{w}_X, \mathbf{w}_Y)$ conditioned on the fitted $\hat{A}$, and constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ is evaluated for all combinations of source and target geometric constraints.

Models sweep with pipeline ablations; Fig. S1. We performed an exhaustive sweep over all $2^4 = 16$ binary rank combinations ($m_A, m_B, m_C, m_D \in \{0, 1\}$) with $n_Z = n_X = n_Y = 5$, $N = 1000$, and 15 simulations per scenario. For each scenario, two subspace alignment conditions are tested: *rand*, in which all weight vectors are drawn independently from the unit sphere; and *corr*, applied only when at least two of m_B, m_C, m_D are non-zero, in which the relevant weight vectors of all non-zero terms are set equal to a shared unit vector per space, making all interaction subspaces perfectly co-aligned. The full $2 \times 2 \times 2$ factorial of pipeline conditions - (regress out Z) $\times$ (permutation significance test) $\times$ (rand/corr) - is evaluated, yielding panels (b-i) of Fig. S1. The significance test uses 1000 permutation shuffles.

Non-Gaussian emission models, Fig. S3. This simulation departs from the others only in the observation step. To test robustness to emission-model misspecification, we used the generative process of Eq. 48 as a latent rate-level model and observed counts through a non-linear emission. Specifically, the latent variables $\mathbf{X}, \mathbf{Y}$ produced by Eq. 48 model, with W_Y rescaled by a modulation-strength scalar γ, are interpreted as log-rates after a constant baseline shift: $\boldsymbol{\eta}_X = h\,\mathbf{1}_{n_X}^\top + \mathbf{X}$ and $\boldsymbol{\eta}_Y = \mathbf{1}_{n_Y}^\top + \mathbf{Y}$, with $h = 2$ and entrywise clipping $|\eta| \leq 4$ for numerical stability. Three emission models were compared:

- *Gaussian benchmark*: $\mathbf{X}^{\mathrm{obs}} = \mathbf{X}$, $\mathbf{Y}^{\mathrm{obs}} = \mathbf{Y}$ (the latent rate-level variables, recovering the standard setup);

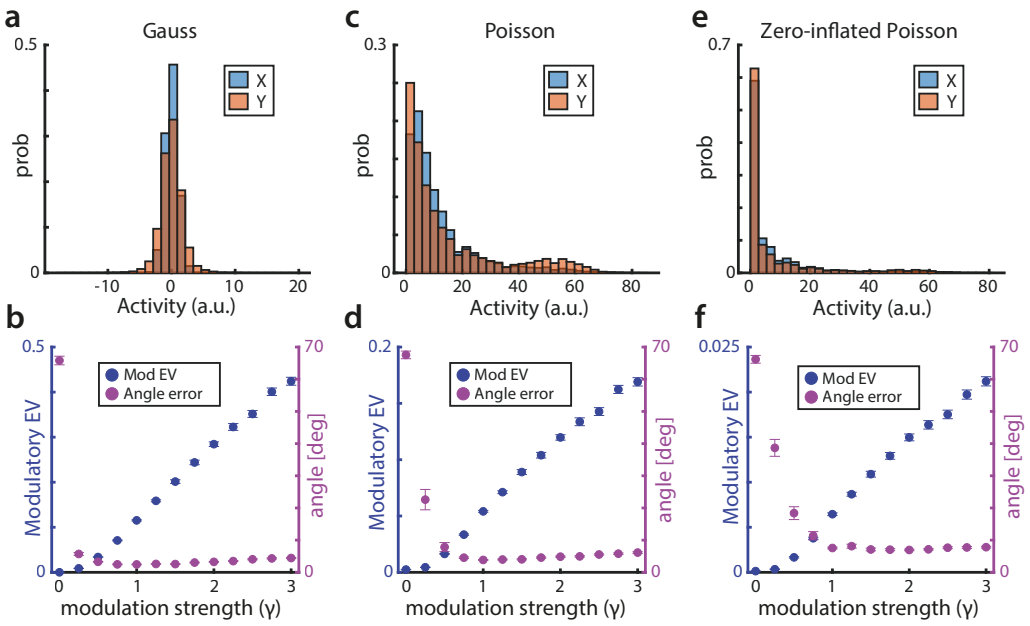

Figure S3: **MICs performance under non-Gaussian observation models.** All conditions share the same latent bilinear generative model, only the emission distribution of X and Y varies across scenarios. (a) Marginal distribution of X and Y for one example simulation in the Gaussian (linear benchmark) condition. (b) $\mathrm{EV}_{\mathrm{mod}}$ (blue) and mean angle between estimated and ground-truth MIC weight vectors (purple) as a function of modulation strength γ. (c-d) Same as (a-b) for Poisson-distributed X and Y. (e-f) Same as (a-b) for zero-inflated Poisson observations (zero-inflation probability $p = 0.5$). $n_X = n_Y = n_Z = 5$; ground-truth ranks $m_A = m_B = m_C = 1$, $m_D = 0$; $N = 10{,}000$ samples per simulation. Curves show mean $\pm$ SEM across 50 independent simulations.

- *Poisson*: $X_{ti}^{\mathrm{obs}} \sim \mathrm{Poisson}(e^{\eta_{X,ti}})$, $Y_{ti}^{\mathrm{obs}} \sim \mathrm{Poisson}(e^{\eta_{Y,ti}})$, independently across entries;
- *zero-inflated Poisson* (zero-inflation probability $p = 0.5$): $X_{ti}^{\mathrm{obs}} = D_{ti}^{X}\,\tilde{X}_{ti}$ with $\tilde{X}_{ti} \sim \mathrm{Poisson}(e^{\eta_{X,ti}})$ and $D_{ti}^{X} \sim \mathrm{Bernoulli}(1-p)$, analogously for Y.

We swept $\gamma \in [0, 3]$ in steps of 0.25, with 50 simulations per γ. Across conditions, $n_X = n_Y = n_Z = 5$, $N = 10^4$, $m_A = m_B = m_C = 1$, $m_D = 0$. The MICs pipeline – which assumes Gaussian observations – was applied without modification, fitting at the ground-truth ranks. We report $\mathrm{EV}_{\mathrm{mod}}$ and the mean principal angle between fitted and ground-truth $(\mathbf{w}_X, \mathbf{w}_Y, \mathbf{w}_Z)$ as a function of γ. $\mathrm{EV}_{\mathrm{mod}}$ increased monotonically with γ in all three conditions Fig. S3b,d,f), and the mean angle error decreased toward zero as γ grew, indicating reliable recovery of the ground-truth MIC geometry despite emission-model misspecification. As expected, absolute $\mathrm{EV}_{\mathrm{mod}}$ was attenuated under count noise relative to the Gaussian benchmark, with zero-inflated Poisson showing the largest attenuation.

Comparison with Z-splitting baseline; Fig. S4a-b. To benchmark MICs against the Z-splitting baseline, we simulated data with $n_X = n_Y = 5$, $n_Z \in \{1, 3, 5, 7, 9\}$, $N = 500$, $m_A = 2$, $m_C = 1$, and $m_B = m_D = 0$, running 100 independent simulations per n_X value. For the splitting baseline, $\mathbf{Z}$ is projected onto its first principal component and samples are partitioned into $N_{\mathrm{bins}} \in \{1, \ldots, 15\}$ equipopulated quantile bins; a separate rank-selected AIS is fit within each bin using the same cross-validation partition as MICs, and modulatory EV is defined as the difference between the mean within-bin AIS EV and the EV of a single AIS fit pooled across all samples. For $n_Z > 1$, this projection discards variance in $\mathbf{Z}$ beyond its leading principal component. MIC modulatory EV and the peak splitting EV (maximized over N_{bins}) are then compared as a function of n_Z.

Single-component and multi-MIC geometries, Fig. S7. We extended the geometric-decomposition validation in two directions. Across all panels, $n_X = n_Y = 20$, $n_Z = 10$, $m_A = 2$,

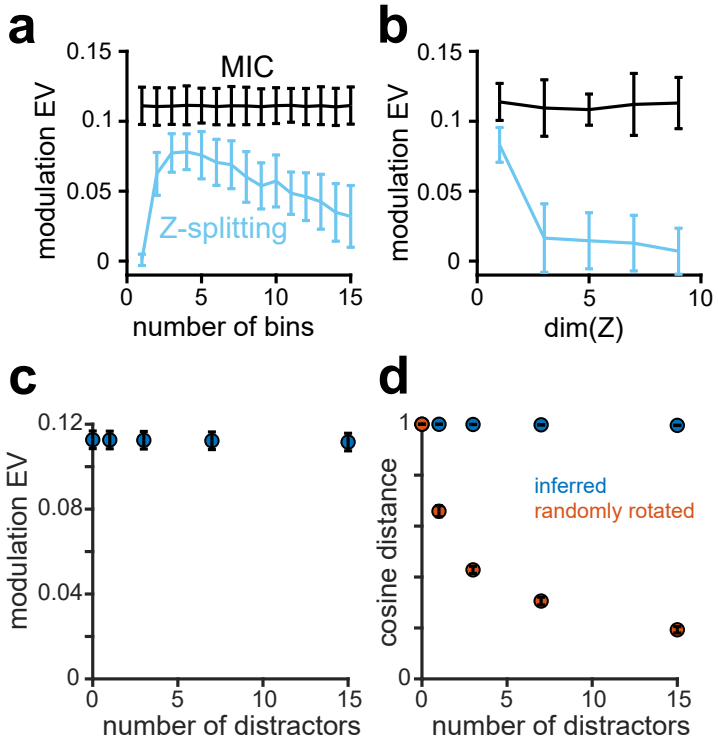

Figure S4: **Benchmarking MICs versus Z-splitting.** (a) Modulatory EV recovered by the MICs method (black) versus a benchmark that discretizes a 1D Z into quantile bins and refits the AIS within each bin (light blue), as a function of the number of bins. (b) Same comparison as a function of $\dim(Z)$, with the quantile-splitting benchmark applied to the first principal component of Z; 100 simulations per scenario, $N = 500$ samples per simulation. (c) $\mathrm{EV}_{\mathrm{mod}}$ recovered by MICs when one modulating dimension of Z is embedded among n non-modulatory distractor dimensions ($n_Z = n + 1$), as a function of n. (d) Cosine distance between ground-truth w_Z and the inferred w_Z (inferred) or a uniformly random unit vector in $\mathbb{R}^{n_Z}$ (randomly rotated). 20 simulations per n, $N = 2000$, $n_X = n_Y = 5$, $m_A = 2$, $m_C = 1$, $m_B = m_D = 0$.

$m_B = m_D = 0$, and $N = 1000$. Panels (a-h) cover the four pure single-component scenarios $\alpha_k^{\mathrm{GT}} = 1$, $k \in \{ss, ns, sn, nn\}$, with $m_C = 1$ and 10 simulations per scenario: $\mathbf{w}_X$ is drawn uniformly on the unit sphere of either $\mathrm{span}(U_A)$ (s on the source side) or its orthogonal complement (n), and analogously for $\mathbf{w}_Y$ with respect to $\mathrm{span}(V_A)$; $\mathbf{w}_Z$ is drawn uniformly on the unit sphere of $\mathbb{R}^{n_Z}$. For each fitted MIC, null-subtracted α indices use 1000 random-rotation shuffles, and constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ is evaluated under all four combinations of source and target geometric constraints (no/source/target/both). Panels (i-k) cover multi-MIC scenarios with $m_C \in \{1, 2, 3\}$ and 30 simulations per m_C: for each channel, angles θ_X, θ_Y are drawn independently and uniformly from $[0, \pi/2]$, and $\mathbf{w}_X$, $\mathbf{w}_Y$ are constructed at these angles to $\mathrm{span}(U_A)$ and $\mathrm{span}(V_A)$ respectively, yielding randomly distributed ground-truth indices $\alpha_{ss}^{\mathrm{GT}} = \cos^2\theta_X \cos^2\theta_Y$, $\alpha_{ns}^{\mathrm{GT}} = \sin^2\theta_X \cos^2\theta_Y$, $\alpha_{sn}^{\mathrm{GT}} = \cos^2\theta_X \sin^2\theta_Y$, $\alpha_{nn}^{\mathrm{GT}} = \sin^2\theta_X \sin^2\theta_Y$.

Identifying components of Z that drive modulation; Fig. S4c. We embedded one genuinely modulatory variable within a pool of n non-modulatory "distractors" and asked whether MICs recovers it. We ran 20 independent simulations per value of $n \in \{1, 3, 7, 15\}$, with $n_X = n_Y = 5$, $n_Z = n + 1$ (the n distractors plus one genuine), $N = 2000$, $m_A = 2$, $m_B = m_D = 0$, and $W_Z = (1, 0, ..., 0)^\top$. As the number of distractors increased, $\mathrm{EV}_{\mathrm{mod}}$ was essentially unchanged and the ground-truth and inferred W_Z remained closely aligned, indicating that MICs correctly recovered the genuine modulator.

In the ACA-VIP dataset, refitting MICs with wheel speed as the only modulator (i.e., removing pupil diameter from Z) did not reduce $\mathrm{EV}_{\mathrm{mod}}$ (0.0066 ± 0.0011 wheel only vs. 0.0071 ± 0.0012 with

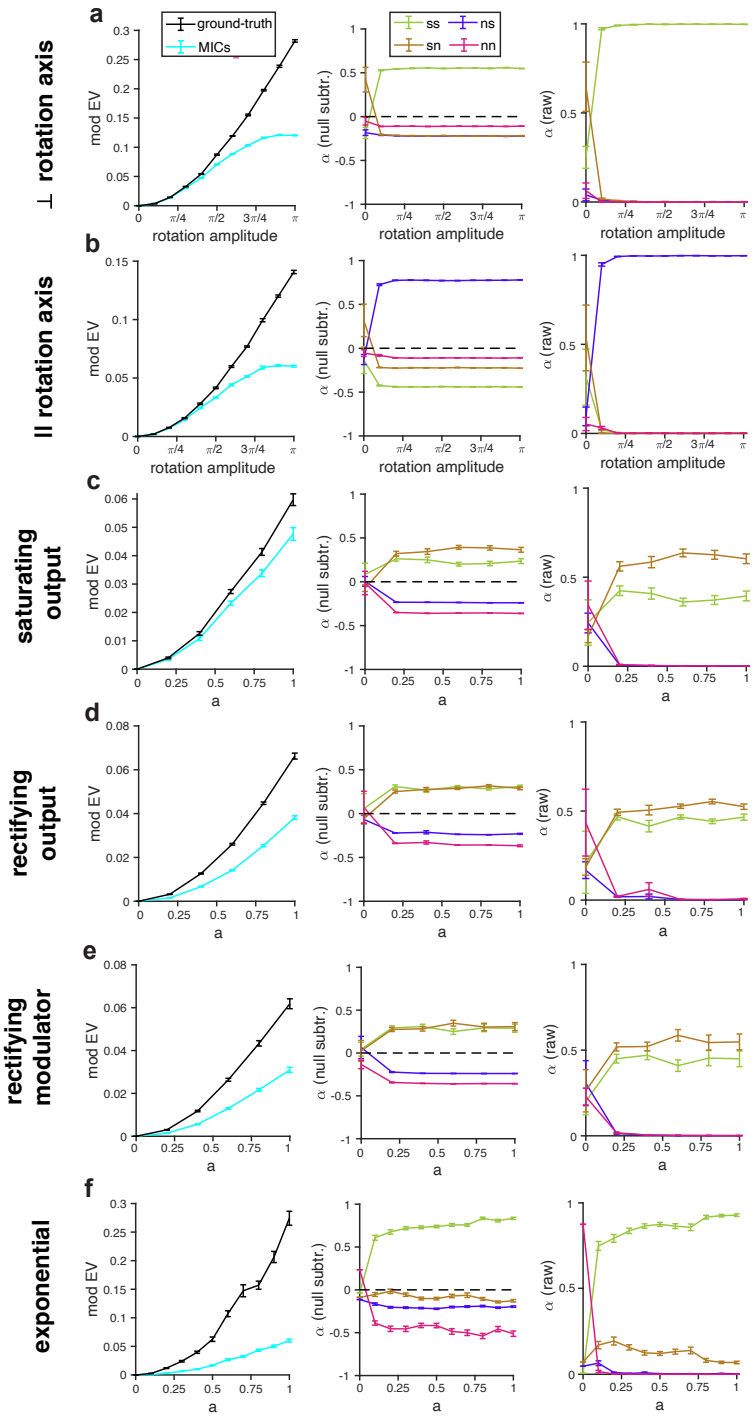

Figure S5: **MICs applied to nonlinear generators.** (a) Perpendicular-axis rotation generator (Appendidx B.10).Top: $\mathrm{EV}_{\mathrm{mod}}$ recovered by MICs (cyan) and by the ground-truth generator (black) as a function of the rotation amplitude. Middle: null-subtracted geometry indices of the inferred MICs (α_{ss}: green, α_{ns}: blue, α_{sn}: brown, α_{nn}: magenta). Bottom: raw geometry indices. (b) Same as (a) for the parallel-axis rotation generator. (c-f) Same as (a) for the saturating output (c), rectifying output (d), rectifying modulator (e), and exponential (f) generators, as a function of the modulation strength parameter a. Mean $\pm$ SEM across 20 simulations per condition (10 for rotations), $N = 10^4$ trials per simulation.

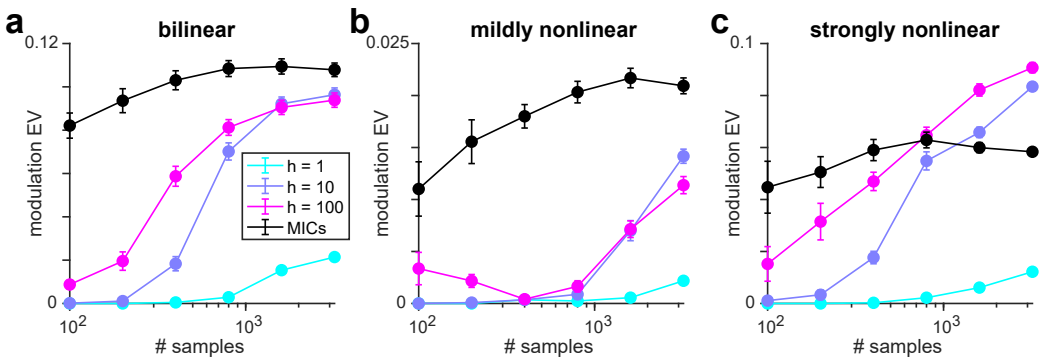

Figure S6: **Sample efficiency of MICs versus a multilayer perceptron.** (a) Cross-validated modulatory EV as a function of sample size N under a bilinear generator, for MICs and for one-hidden-layer MLPs with $h \cdot m_C$ hidden ReLU units. All models were fit to the residuals of the joint additive model using the same cross-validation folds (Appendix B.12). (b) Same as (a) for a mildly nonlinear generator (parallel-axis rotation, $a = 2\pi/5$; Appendix B.10). (c) Same as (a) for a strongly nonlinear generator (parallel-axis rotation, $a = \pi$). Mean $\pm$ SEM across 10 simulations per N.

both variables, $p = 0.64$; Fig. S11f), whereas refitting with pupil diameter as the only modulator significantly reduced it (0.0013 ± 0.0006; $p = 0.0007$ vs. both, $p = 0.003$ vs. wheel only). Thus, wheel speed is the main driver of the modulatory effect we observed.

B.10 Stress-testing MICs on nonlinear generators

The MICs model (Eq. 5) assumes that the source-target interaction is linear and that modulation is bilinear. We applied MICs to simulations of five different classes of nonlinear generators that break these assumptions, all of the form

$$y = f_a(x, z) + e$$

where e is zero-centered, unit-variance Gaussian noise and a is a scalar that controls the amount of modulation.

- *Rotations:* $f_a(x, z) = A^\top R(\theta = az, \mathbf{u})x$. Here, $n_X = n_Y = 3$, $n_Z = 1$, and $R(\theta, \mathbf{u})$ is the matrix that rotates vectors by angle θ about axis $\mathbf{u}$ in $\mathbb{R}^3$. For any fixed value of z, the source-target interaction is linear and characterized by an AIS (as assumed by MICs). Changes in z, however, elicit nonlinear changes in the AIS, breaking the assumptions of MICs. In particular, the AIS rotates on the source side. We set $A = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$ and either $\mathbf{u} = (0, 0, 1)^\top$ ("perpendicular"; the source-side AIS rotates within itself) or $\mathbf{u}$ uniformly drawn from $\mathrm{span}\{(1, 0, 0)^\top, (0, 1, 0)^\top\}$ ("parallel"; the source-side AIS rotates outside itself). Increasing a increases the rotation amplitude and hence the amount of modulation by z. We ran 10 independent simulations per value of a and type of rotation, with $N = 10^4$ trials per simulation and the components of x and z drawn independently from the zero-centered Gaussian with unit variance on each trial.
- *Saturating output:* $f_a(x, z) = \tanh(A^\top x + aW_Y((W_X^\top x) \odot (W_Z^\top z))$. If $\tanh$ were removed, the remaining expression would simply be the MICs model. We set $n_X = n_Y = 5$, $n_Z = 1$, $m_A = 2$, $m_C = 1$, $A = U_A V_A^\top$ with the columns of $U_A \in \mathbb{R}^{n_X \times m_A}$, $V_A \in \mathbb{R}^{n_Y \times m_A}$ drawn uniformly from the unit sphere, $W_X \in \mathbb{R}^{n_X}$ drawn uniformly from the column space of U_A, $W_Y \in \mathbb{R}^{n_Y}$ drawn uniformly from the subspace making an angle $\pi/4$ to the column space of V_A, and $W_Z = 1$. This setup ensures that the MICs model embedded within the output nonlinearity has geometry indices $\alpha_{ss} = \alpha_{sn} = 1/2$. As before, a controls the amount of modulation by z. We ran 20 simulations per value of a, again with $N = 10^4$ trials and $x_i, z_j \sim N(0, 1)$ on each trial.
- *Rectifying output:* $f_a(x, z) = \mathrm{ReLU}(A^\top x + aW_Y((W_X^\top x) \odot (W_Z^\top z))$. All other details are the same as in the *Saturating output* case.

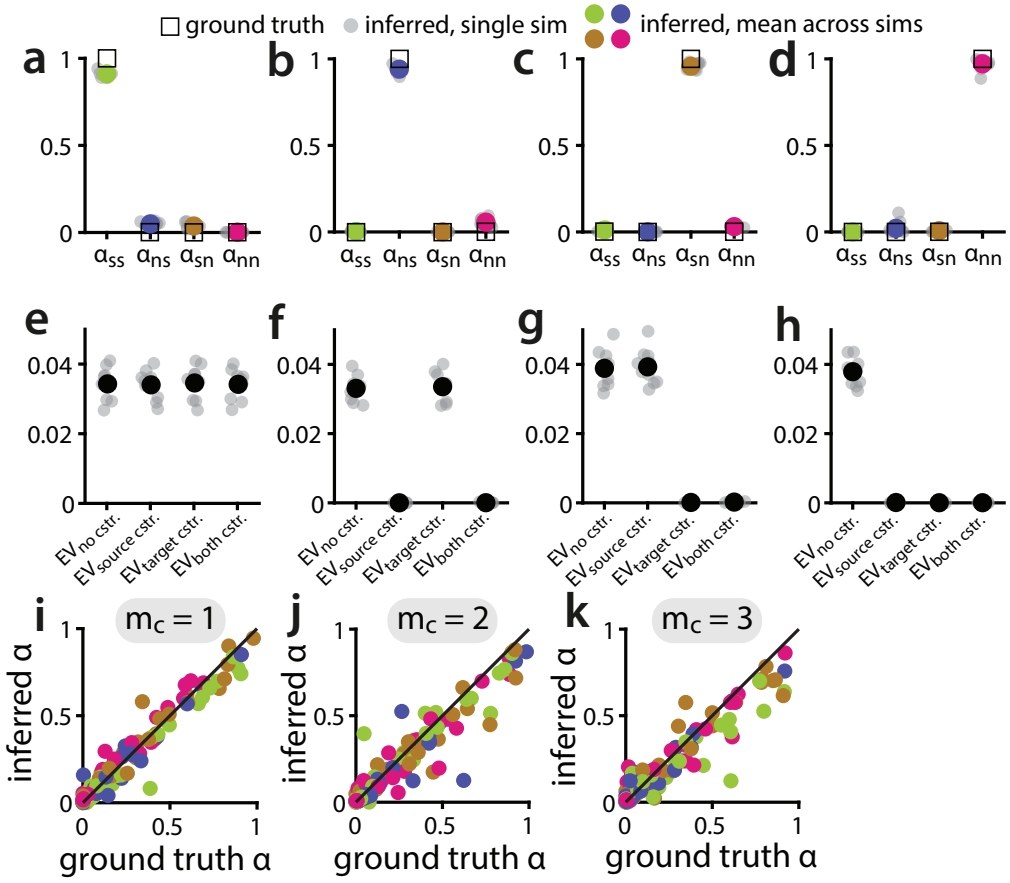

Figure S7: **Validating the geometric decomposition across single-component and multi-MIC scenarios.** (a) Raw geometry indices for a single-MIC scenario with ground-truth $\alpha_{ss}^{\mathrm{GT}} = 1$, all other components zero. Open squares: ground truth; filled colored circles: mean across simulations; light grey dots: individual simulations. (b-d) Same as (a) for ground-truth $\alpha_{ns}^{\mathrm{GT}} = 1$ (b), $\alpha_{sn}^{\mathrm{GT}} = 1$ (c), and $\alpha_{nn}^{\mathrm{GT}} = 1$ (d). (e) Constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ for the $\alpha_{ss}^{\mathrm{GT}} = 1$ scenario, evaluated under four constraints: *no cstr.*, unconstrained fit; *source cstr.*, source factor confined AIS, forcing $\alpha_{ns} = \alpha_{nn} = 0$; *target cstr.*, target confined to AIS, forcing $\alpha_{sn} = \alpha_{nn} = 0$; *both cstr.*, both source and target confined to AIS, leaving only α_{ss} admissible. Filled black circles: mean across simulations; light grey dots: individual simulations. (f-h) Same as (e) for the $\alpha_{ns}^{\mathrm{GT}} = 1$ (f), $\alpha_{sn}^{\mathrm{GT}} = 1$ (g), and $\alpha_{nn}^{\mathrm{GT}} = 1$ (h) scenarios. (i-k) Inferred against ground-truth α indices in multi-MIC scenarios with $m_C = 1$ (i), $m_C = 2$ (j), $m_C = 3$ (k); for each MIC, $\mathbf{w}_X$ is drawn to yield randomly-distributed ground-truth α profiles. Each dot is one channel from one simulation; colors indicate the four geometric components (α_{ss}: green, α_{ns}: blue, α_{sn}: brown, α_{nn}: magenta). Black line: identity. We ran 10 simulations per scenario in (a-h) and 30 simulations per m_C in (i-k). $n_X = n_Y = 20$, $n_Z = 10$, $m_A = 2$, $m_B = m_D = 0$, $N = 1000$ in all simulations.

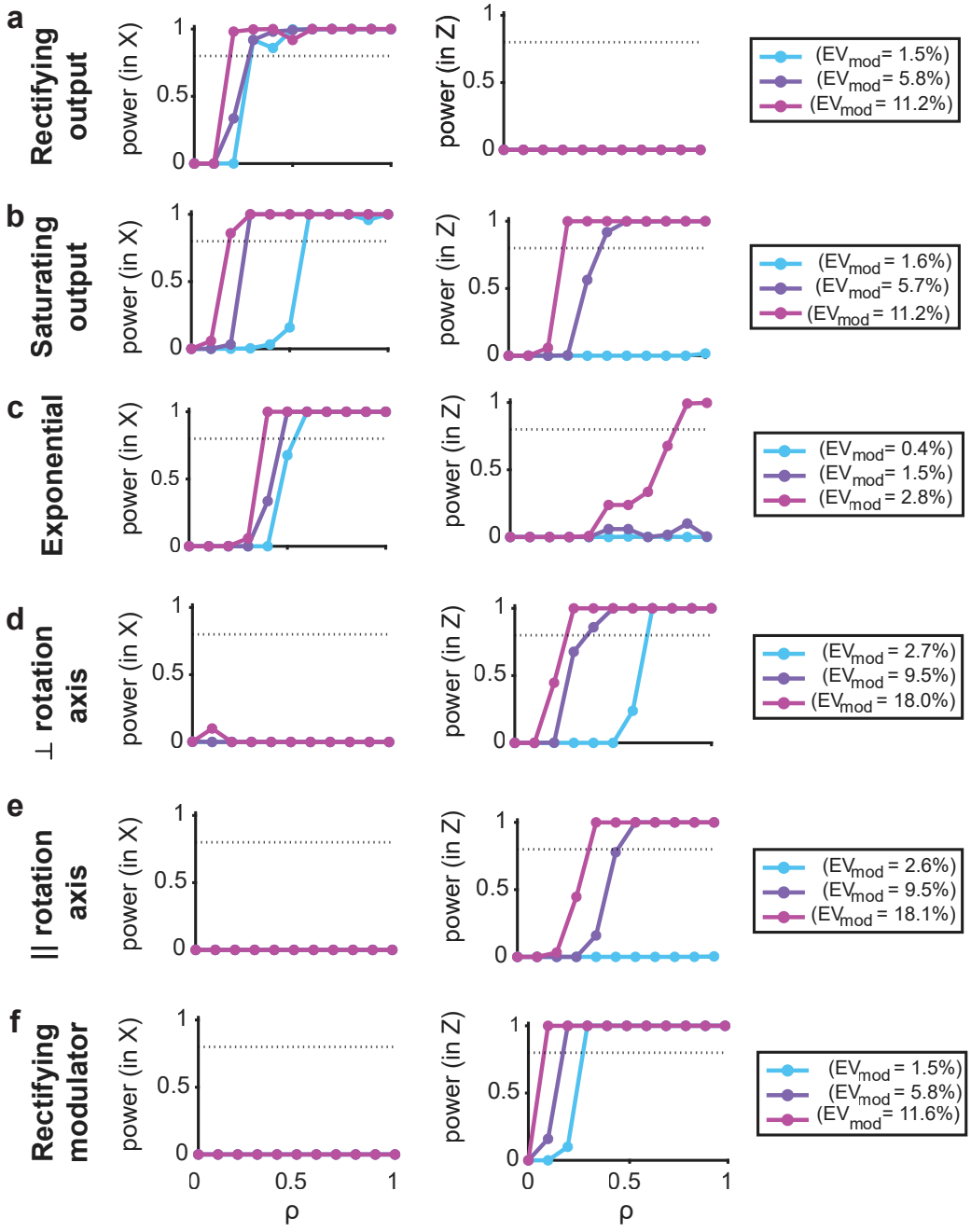

Figure S8: **Power of the bilinearity diagnostic.** Each nonlinear generator of Appendix B.10 was linearly interpolated with a bilinear MICs generator, so that $\rho = 0$ is purely bilinear and $\rho = 1$ is the original nonlinear generator. (a) Rectifying output generator. Left: power of detecting $\Delta\mathrm{EV}_{\mathrm{NL}} > 0$ with the nonlinear-in-x. Right: power of detecting $\Delta\mathrm{EV}_{\mathrm{NL}} > 0$ with nonlinear-in-z (sign test across 50 simulations, $\alpha = 0.05$; dotted line: power of 0.8), as a function of ρ. Colors: modulation strength $a = 1/3$ (light blue), $2/3$ (violet) and 1 (purple); legends report the $\mathrm{EV}_{\mathrm{mod}}$ recovered by MICs at $\rho = 0$ for each a. (b) Same as (a) for the saturating output generator. (c) Same as (a) for the exponential generator. (d-e) Same as (a) for the perpendicular-axis (d) and parallel-axis (e) rotation generators (rotation amplitude $a\pi$). (f) Same as (a) for the rectifying modulator generator. $N = 2 \times 10^4$ trials per simulation.

- *Rectifying modulator:* $f_a(x, z) = A^\top x + aW_Y((W_X^\top x) \odot (W_Z^\top \mathrm{ReLU}(z))$. All other details are the same as in the *Saturating output* case.
- *Exponential:* $f_a(x, z) = (A^\top x)^{1+a(2z-1)}$. The exponentiation acts componentwise. We set $n_X = n_Y = 5$, $n_Z = 1$, $m_A = 2$, and $A = U_A V_A^\top$ with the columns of $U_A \in \mathbb{R}^{n_X \times m_A}$, $V_A \in \mathbb{R}^{n_Y \times m_A}$ drawn uniformly from the portion of the unit sphere where all components are positive. If $a > 0$, then $z > 1/2$ results in a supralinear mapping from x to y, $z < 1/2$ in a sublinear mapping, and $z = 1/2$ in a linear mapping. In each trial (20 simulations per a, each with $N = 10^4$ trials), we drew the components of x and z uniformly from the interval $(0, 1)$. Thus, the generator is well-defined for $|a| \leq 1$ (for $|a| > 1$, the exponent can become negative).

We used two criteria to assess the performance of MICs on these nonlinear generators. First, we compared the $\mathrm{EV}_{\mathrm{mod}}$ returned by MICs to the "ground-truth" $\mathrm{EV}_{\mathrm{mod}}$, defined as the explained variance of the ground-truth generator's deterministic part, minus that of the same generator with z clamped to baseline ($z = 1/2$ for the *Exponential* generator, and $z = 0$ for all others). Across all generators, the $\mathrm{EV}_{\mathrm{mod}}$ of MICs monotonically increased with the ground-truth value as a increased, indicating that MICs could still track changes in the amount of modulation despite the generator mismatch (Fig. S5).

Second, we asked whether the modulation geometry indices inferred by MICs match the geometry built into the ground-truth generators. The *Rotations* generator alters which source (x) patterns interact with a fixed set of target (y) patterns, and hence should have $\alpha_{sn} \approx \alpha_{nn} \approx 0$. In the perpendicular axis case, the rotation acts within the AIS, so we expect $\alpha_{ss} \approx 1$. In the parallel axis case, on the other hand, the rotation acts outside the AIS, so we expect $\alpha_{ns} \approx 1$. MICs recovered all of these expectations, regardless of the rotation amplitude (Fig. (Fig. S5)a-b).

The *Saturating output*, *Rectifying ouput*, and *Rectifying modulator* generators were constructed so that, if the nonlinearity were removed, the ground-truth geometry indices would be $\alpha_{ss} = \alpha_{sn} = 1/2$ and $\alpha_{ns} = \alpha_{nn} = 0$. Despite the nonlinearities, MICs identified α_{ss} and α_{sn} as the indices exceeding the random rotation null (Fig. S5c-e). For the two output nonlinearities, x enters only through $W_X^\top x$ and $A^\top x$, both within the source AIS, so $\alpha_{ns} = \alpha_{nn} = 0$ is preserved, whereas the nonlinearity acts on the target and distorts the target MIC axis out of the target AIS, tilting the indices toward $\alpha_{sn} > \alpha_{ss}$.

Under the *Exponential* generator, z controls the exponent applied to the baseline output $A^\top x$. Thus, z primarily changes the overall gain of x-to-y interactions but can also distort which patterns in y are targeted by x. Accordingly, MICs returns α_{ss} as the largest index, with a smaller (but above random-rotation null) α_{sn} (Fig. (Fig. S5)f).

B.11 Diagnosing violations of bilinear modulation assumption

The MICs model assumes that modulation is bilinear, i.e. linear in the modulator z at fixed x and linear in the source x at fixed z. This assumption is deliberate: bilinearity, together with the low-rank CP parameterization, is what makes MICs data-efficient and geometrically interpretable. To diagnose violations of bilinearity, we added higher-order nonlinear terms to MICs and measured the change in performance compared to the bilinear baseline. Specifically, for a fitted channel (w_X, w_Y, w_Z), define the residualized source projection $s_X = \mathbf{X}' w_X$, the modulator projection $s_Z = \mathbf{Z} w_Z$, and the modulation-residual target projection $s_Y = \mathbf{Y}'' w_Y$. We tested the performance of a nested set of regressions of s_Y:

- *Bilinear:* $s_Y \sim s_X \odot s_Z$.
- *Nonlinear in x:* $s_Y \sim s_X \odot s_Z + s_X^2 \odot s_Z + s_X^3 \odot s_Z$.
- *Nonlinear in z:* $s_Y \sim s_X \odot s_Z + s_X \odot s_Z^2 + s_X \odot s_Z^3$.

We included terms up to third order to capture both rectification (second order) and saturation (third order) effects. We use the same cross-validation folds as were used to fit the MICs model and calculate the projections (s_X, s_Y, s_Z) for the train and test sets using the axes (w_X, w_Y, w_Z) from the train set. Each regression is ridge-regularized, with the penalty optimized by cross-validation over $[10^{-4}, 10^1]$. We report the difference $\Delta\mathrm{EV}_{\mathrm{NL}}(x)$ ($\Delta\mathrm{EV}_{\mathrm{NL}}(z)$) between the cross-validated explained variance of the *Nonlinear in x* (*Nonlinear in z*) regression and the bilinear regression. $\Delta\mathrm{EV}_{\mathrm{NL}}(x) > 0$ indicates a violation of linearity in x, and similarly $\Delta\mathrm{EV}_{\mathrm{NL}}(z) > 0$ indicates a

violation in z. Significance across simulation runs (or sessions, for real data analyses S9e and S11g) was assessed with a one-tailed sign (binomial) test on $\Delta\mathrm{EV}_{\mathrm{NL}} > 0$, one-tailed because only an increase in EV indicates a violation of bilinearity. In simulations, we report power: the probability that the sign test detects nonlinearity ($p < 0.05$) across $n = 50$ runs, computed from the fraction of runs with $\Delta\mathrm{EV}_{\mathrm{NL}} > 0$. For the bilinear generator ($\rho = 0$), this is the false-positive rate (Fig. S8).

We tested the diagnostic on simulations of the nonlinear generators in Appendix B.10 (Fig. S8). In the *Rotations* and *Rectifying modulator* generators, the nonlinearity acts directly on z, so the diagnostic should report a violation in z only. To determine how the diagnostic should behave on the remaining generators, we expanded them in power series around the baseline values of x and z:

- *Saturating output:* Besides the bilinear term, the lowest-order nonzero modulatory (i.e. containing both x and z) term is $a(A^\top x)^2 \odot W_Y\big((W_X^\top x) \odot (W_Z^\top z)\big)$, where the exponentiation acts componentwise. Thus, we expect the diagnostic to localize the violation primarily to x.
- *Rectifying output:* ReLU is not differentiable. However, it can be approximated by the smooth softplus function, $\mathrm{ReLU}(u) \approx \frac{1}{\beta}\ln(1 + e^{\beta u})$, with the error vanishing as $\beta \to \infty$. Under this approximation, the lowest-order modulatory term beyond the bilinear one is $\frac{a\beta}{4}(A^\top x) \odot W_Y\big((W_X^\top x) \odot (W_Z^\top z)\big)$, so we again expect the diagnostic to localize the violation to x.
- *Exponential:* Here, there are two lowest-order modulatory terms beyond bilinear, namely $-a(z - \frac{1}{2})\, b^{-1} \odot (A^\top(x - \frac{1}{2}\mathbf{1}))^2$ (nonlinear in x only) and $2a^2(z - \frac{1}{2})^2 \ln b \odot (2 + \ln b) \odot A^\top(x - \frac{1}{2}\mathbf{1})$ (nonlinear in z only), where $\mathbf{1} \in \mathbb{R}^{n_X}$ is the vector of 1s and $b = \frac{1}{2}A^\top\mathbf{1}$. Thus, the diagnostic should detect a violation in both x and z, with more power in x for $a < 1$, since the x term is first order in a and the z term second order.

Across all generators, the diagnostic detected a violation of bilinearity and localized it to the expected factor (x or z). Under a bilinear generator, the diagnostic correctly reported no violation.

B.12 Sample-efficiency of MICs compared to one-hidden-layer MLPs

In simulations of nonlinear generators (Appendix B.10), MICs correctly detected the geometry and presence of modulation, but with reduced sensitivity: the $\mathrm{EV}_{\mathrm{mod}}$ captured by MICs was typically smaller than the ground-truth, especially for strong modulation (Fig. S5). We therefore asked how much more $\mathrm{EV}_{\mathrm{mod}}$ a flexible nonlinear model could capture. Specifically, we compared MICs to a one-hidden-layer perceptron (MLP; ReLU activation function, h hidden units), a universal approximator of any continuous function [66]. We trained the MLP on exactly the same inputs and target output as the modulation term in MICs: it receives $[\mathbf{X}', \mathbf{Z}]$ and predicts $\mathbf{Y}''$, the residuals of the joint additive term in MICs. The MLP has $h(n_X + n_Y + n_Z) + h + n_Y$ parameters, which even at $h = m_C$ are more than that of MICs' modulation term ($m_C(n_X + n_Y + n_Z)$). We swept $h \in \{m_C, 10m_C, 100m_C\}$, up to 100 times more parameters than MICs. We used mean-squared-error loss, Adam (learning rate 10^{-3}, 200 epochs, minibatch size 32) and ridge regularization selected on held-out data (penalty grid 10^{-5} to 10^{-1}). We quantified the MLP's cross-validated explained variance ($\mathrm{EV}_{\mathrm{res}}^{\mathrm{MLP}}$) as the increment in explained variance over the joint additive model, exactly as we did for MICs $\mathrm{EV}_{\mathrm{mod}}$.

Note that while $\mathrm{EV}_{\mathrm{mod}}$ isolates the bilinear modulation term, $\mathrm{EV}_{\mathrm{res}}^{\mathrm{MLP}}$ captures any nonlinear structure in the residuals, whether from modulation or from nonlinearity in the marginal $x \to y$ and $z \to y$ mappings. Thus, the residual explained variance from the MLP is an upper bound on the variance attributable to modulation of any form: $\mathrm{EV}_{\mathrm{mod}} > \mathrm{EV}_{\mathrm{res}}^{\mathrm{MLP}}$ establishes that the flexible model does not extract more from the residuals than MICs does, whereas the reverse inequality is not specifically attributable to modulation.

Under a bilinear generator, we found $\mathrm{EV}_{\mathrm{mod}} > \mathrm{EV}_{\mathrm{res}}^{\mathrm{MLP}}$ regardless of sample size N (Fig. S6a). Even under a strongly nonlinear generator (parallel-axis rotation with amplitude $a = \pi$; Appendix B.10), MICs fares better than the MLP at small N, though the MLP eventually overtakes it for large enough N (Fig. S6c). Under a milder nonlinearity (parallel-axis rotation with $a = 2\pi/5$), MICs exceeded MLP at all N considered and degraded much more slowly as N decreased (Fig. S6b). Thus, while the MLP has higher flexibility to capture nonlinear modulations, this comes at the cost of reduced sample efficiency and interpretability.

B.13 Statistical testing of modulatory and AIS EV

Motivation. The 1-SEM rule used for rank selection (Algorithm 1; [16, 51]) is a heuristic that can yield spurious AIS or modulatory EV in a fraction of cases when the ground-truth m_A or m_C is zero. Across the $m_A^{\text{GT}} = 0$ or $m_C^{\text{GT}} = 0$ scenarios of Fig. S1 (*rand*, Step 1 on), the rule alone retained $m_A \geq 1$ and $m_C \geq 1$ in 18 % and 34 % of simulations, respectively. We therefore complement rank selection with a permutation test, applied hierarchically to AIS ($m_A > 0$) and to modulation ($m_C > 0$), that decides on held-out data whether each candidate term predicts Y better than expected under a null in which it is temporally unrelated to the test-set residuals it is meant to explain.

Test construction. We split the dataset into two non-overlapping halves (first and second half of samples). For each split, the MICs pipeline (rank selection, ridge regularization, and ALS) is fit on one half and evaluated on the other, so that all fitting and hyperparameter selection are confined to the training half. On the held-out half we form three predictions of Y of increasing complexity,

$$\hat{Y}_B = Z_{\text{test}} B, \qquad \hat{Y}_{AB} = \hat{Y}_B + X'_{\text{test}} A, \qquad \hat{Y}_{\text{full}} = \hat{Y}_{AB} + \left[(X'_{\text{test}} W_X) \odot (Z_{\text{test}} W_Z) \right] W_Y^\top,$$

where $X' = X - ZD$ uses the regression D fit in Step 1 of Algorithm 1. The observed held-out EV increments

$$\Delta\text{EV}_{AIS} = R^2(\hat{Y}_{AB}) - R^2(\hat{Y}_B), \qquad \Delta\text{EV}_{mod} = R^2(\hat{Y}_{\text{full}}) - R^2(\hat{Y}_{AB}),$$

quantify the variance in Y_{test} captured by AIS and by modulation, respectively. $R^2(\hat{Y})$ denotes the coefficient of determination of real Y from prediction $\hat{Y}$ compared against Y training mean.

To test significance of ΔEV_{mod}, we shuffle in time the predicted modulation component $\hat{\Delta}_{mod} = \hat{Y}_{\text{full}} - \hat{Y}_{AB}$ $n_{\text{shuf}} = 1000$ times. For time series, we apply random circular shifts uniformly drawn from $[\lfloor N_{\text{test}}/4 \rfloor, \lfloor 3N_{\text{test}}/4 \rfloor]$, where N_{test} denotes the number of test samples. For i.i.d. samples from simulations, we randomly permute $\hat{\Delta}_{mod}$ across all test samples. We then recompute $\Delta\text{EV}_{mod}^{shuff} = R^2(\hat{Y}_{AB} + \hat{\Delta}_{mod}^{shuff}) - R^2(\hat{Y}_{AB})$ on each shuffle. Circular shifting destroys the temporal alignment between $\hat{\Delta}_{mod}$ and the residual $Y_{\text{test}} - \hat{Y}_{AB}$ while preserving the marginal distribution and autocorrelation of $\hat{\Delta}_{mod}$, so that adding back a misaligned $\hat{\Delta}_{mod}^{shuff}$ to $\hat{Y}_{AB}$ generally inflates the sum of squared errors and yields a null statistic with negative mean. A genuine modulatory contribution should therefore exceed essentially all shuffles. Null EV are averaged across the two splits, and a one-tailed p-value is computed empirically, comparing ΔEV_{mod} to the $\Delta\text{EV}_{mod}^{shuff}$ distribution. ΔEV_{AIS} significance is tested identically with $\hat{\Delta}_{AIS} = X'_{\text{test}} A$ in place of $\hat{\Delta}_{mod}$ and $\hat{Y}_B$ in place of $\hat{Y}_{AB}$. A non-zero rank is retained only if $p < 0.05$; otherwise the corresponding term is fixed to zero in all downstream analyses (geometry indices, constrained-fit comparisons).

Testing m_A and m_C with this procedure reduces the average false-positive rate at $m_C^{\text{GT}} = 0$ from 34 % to 0 % across the 16 scenarios of Fig. S1 while preserving a true-positive rate of 100 % at $m_C^{\text{GT}} = 1$ (vs. 100 % without the test); analogous values for m_A: 18 %, 4 %, 100 %, 100 %.

B.14 Details and further analyses of V1-V2 data

Experimental procedures and preprocessing. We re-analyzed the publicly available V1-V2 dataset of Zandvakili and Kohn `https://crcns.org/data-sets/vc/v1v2-1`, a subset of the recordings described in [55] comprising the largest and best retinotopically aligned populations, with V2 electrodes targeted to the input layers, previously analyzed by Semedo et al. [16] and by Gokcen et al.[56], to which we refer for full experimental details. Briefly, V1 (Utah array) and V2 (tetrodes) populations with aligned receptive fields were recorded simultaneously in anesthetized macaques viewing drifting gratings of 8 orientations θ (22.5° steps; 300-400 trials per orientation, 1.28 s each). The dataset consists of 5 sessions from 3 animals (88-159 V1 and 24-37 V2 units per session). Following [16], we used 100 ms spike-count bins from 160 to 1160 ms after stimulus onset, retained units with firing rate ≥ 0.5 spikes/s, and subtracted for each unit, orientation and time bin the so called Peri-Stimulus-Time-Histogram, or PSTH, that is time-dependent firing rate averaged over trials with the same orientation, so that analyses target trial-to-trial residual co-fluctuations. Whereas [16] and [56] fit an independent model within each orientation, we fit a single model across all orientations, taking V1 units as source X, V2 units as target Y, and grating orientation as modulator Z. Because the PSTH is removed, orientation has no first-order effect on either population and the

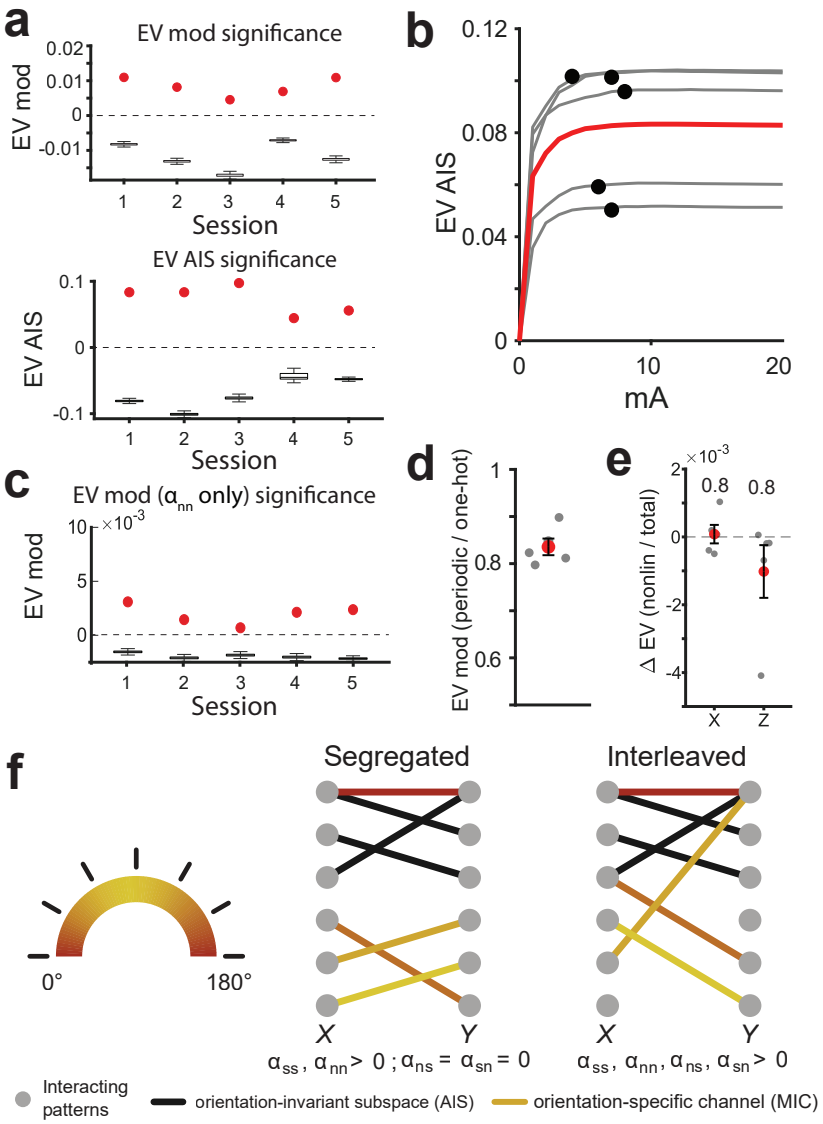

Figure S9: **Additional analyses of V1-V2 benchmark data.** All panels use the one-hot encoding of orientation unless stated. (a) Single-session significance of $\mathrm{EV}_{\mathrm{mod}}$ (top) and $\mathrm{EV}_{\mathrm{AIS}}$ (bottom) against the residual-shuffle null (Appendix B.13). Boxes: null median and IQR; whiskers: $1.5\times$IQR; red dots: observed held-out EV. Both terms were significant in 5/5 sessions. (b) $\mathrm{EV}_{\mathrm{AIS}}$ as a function of AIS rank m_A; gray lines: individual sessions; black dots: single-session selected rank; red: across-sessions mean. (c) As (a, top), for MICs constrained to the nn component only ($w_{X,i}$ orthogonal to the V1 AIS and $w_{Y,i}$ orthogonal to the V2 AIS; Appendix B.8). Despite $\tilde{\alpha}_{nn} < 0$ (Fig. 4d), nn-only modulation was significant in 5/5 sessions ($\mathrm{EV}_{\mathrm{mod}} = (0.24 \pm 0.05)\%$). (d) Ratio between $\mathrm{EV}_{\mathrm{mod}}$ obtained with the periodic encoding $Z = (\sin 2\theta, \cos 2\theta)$ and with the one-hot encoding; black dots: sessions; red: mean $\pm$ SEM ($(83.6 \pm 1.8)\%$). (e) Bilinearity diagnostic (Appendix B.11): change in cross-validated EV when higher-order terms in x or z are added to the fitted MICs ($\Delta\mathrm{EV}_{\mathrm{NL}}$). Gray dots: sessions; red: mean $\pm$ SEM; numbers: one-tailed binomial test p-values. (f) Schematic of segregated versus interleaved organization of orientation-specific interactions. Nodes: V1 (X) and V2 (Y) interacting activity patterns; black edges: θ-invariant interaction (AIS); colored edges: MICs, colored by the orientation at which they are most strongly engaged. Left) segregated: each MIC lies inside or outside the AIS at both ends (ss, nn), so $\alpha_{ns} = \alpha_{sn} = 0$. Right) interleaved: MICs also couple an AIS pattern on one side with a non-AIS pattern on the other (ns, sn), as observed in real data (Fig. 4d).

additive terms B and D explain zero variance; $\mathrm{EV}_{\mathrm{mod}}$ thus quantifies whether the V1-to-V2 residual mapping depends on θ. We used two encodings of θ: a one-hot encoding of the eight orientations ($Z \in \mathbb{R}^8$), which imposes no functional form on how the interaction depends on θ, and a periodic encoding $Z = (\sin 2\theta, \cos 2\theta) \in \mathbb{R}^2$. Unless stated otherwise, results refer to the one-hot encoding.

MICs fitting and main analyses. We fit the pipeline of Algorithm 1 separately to each session as for the ACA-VIP data (Appendix B.15): $K = 10$ temporally contiguous cross-validation folds, all variables z-scored at the fitting stage, and AIS and MICs ranks capped at $m_A^{\max} = 20$ and $m_C^{\max} = 10$ (mean selected ranks $m_A = 4.8$, $m_C = 6.4$). Single-session significance of $\mathrm{EV}_{\mathrm{mod}}$ and $\mathrm{EV}_{\mathrm{AIS}}$ was assessed via the residual-shuffle permutation test of Appendix B.13 with $n_{\mathrm{shuf}} = 1000$ shuffles per session (Fig. S9a). In Fig. 4b, $\mathrm{EV}_{\mathrm{mod}}$ is normalized by $\mathrm{EV}_{\mathrm{AIS}}$ to allow comparison across datasets. For the example MIC of Fig. 4c, orientations were ordered by the rank of their modulator score $Zw_{Z,i}$; the colorbar reports the angular difference of each orientation from the lowest-scoring one. The MIC correlation slope (Fig. 4c) was computed as for ACA-VIP data, except that the modulator score takes only 8 values: the Pearson correlation between $Xw_{X,i}$ and $Yw_{Y,i}$ was computed within each orientation and regressed against the orientation score rank-normalized to $[0, 1]$; slopes were averaged across MICs and then across sessions. Geometry indices (Fig. 4d, top) and constrained fits (Fig. 4d, bottom) were computed as for ACA-VIP data, with "V1" and "V2" denoting the constraint applied to the source and target factor, respectively; the nn-only model (Fig. S9c) constrains both factors to the orthogonal complement of the corresponding AIS (Appendix B.8). Because sessions from the same animal are not independent, statistics in Fig. 4d,h use linear mixed-effects models with a random intercept per animal. For each null-subtracted index, and for each selected constrained-fit comparison (Appendix B.8), we modeled the per-session value (the index, or the difference in $\mathrm{EV}_{\mathrm{mod}}$ between the two fits) as $d \sim 1 + (1 \mid \mathrm{animal})$ and tested the intercept against zero (two-tailed; Holm corrected across the four indices and across the three comparisons of Fig. 4d, respectively). For simplicity, tests in the supplementary figures take sessions as the unit of analysis. To test whether the modulation is captured by a smooth periodic function of θ, we refit MICs with the periodic encoding, using the same pipeline and folds, and report per session the ratio of its cross-validated $\mathrm{EV}_{\mathrm{mod}}$ to that of the one-hot encoding (Fig. S9d).

Table 1: Single-session explained variance for the V1-V2 dataset (one-hot encoding of θ; all values in %).

Session	$\mathrm{EV}_{\mathrm{mod}}$	$\mathrm{EV}_{\mathrm{AIS}}$	$\mathrm{EV}_{\mathrm{mod}}/\mathrm{EV}_{\mathrm{AIS}}$
1	1.49	10.06	14.81
2	1.25	9.58	13.05
3	1.23	10.17	12.09
4	0.84	5.03	16.70
5	1.53	5.85	26.15
Mean $\pm$ SEM	1.27 ± 0.12	8.14 ± 1.11	16.56 ± 2.52

B.15 Details and further analyses of ACA-VIP data

Experimental procedures and preprocessing. We collected simultaneous dual-color two-photon calcium imaging of ACA axonal projections in V1 and V1 VIP interneurons in head-fixed mice during viewing of visual stimuli, while continuously tracking pupil diameter and running speed. The dataset is publicly available at `https://osf.io/2ekj4/`. All procedures were approved by the Massachusetts Institute of Technology Committee on Animal Care and conformed to the National Institutes of Health guidelines for the care and use of laboratory animals. VIP-Cre mice (Jackson stock no: 031628, RRID:IMSR_JAX:031628) were anesthetized with isoflurane (1.5%) and given preemptive analgesia (extended release buprenex, 1mg/kg, and meloxicam, 5mg/kg, s.c.). After being placed in a stereotaxic frame and prepared for surgery, a small craniotomy was made above the ACA (AP: +1, ML: -0.3, DV: -0.9) and 0.3 ul of AAV1-hSyn-axon-GCaMP6s [67] was injected into the ACA with a micropipette. A larger craniotomy (3mm) was made above the V1 and 4 small injection of 0.15μl AAV9-DIO-jRGECO1a [68] was made into the V1 before placing the cranial window and securing it on the skull with dental cement. Visual stimuli were drifting gratings (8 directions $\times$ 3 contrasts, 10 trials per condition; 240 trials/session) or natural movies (5 movies $\times$ 3 contrasts, 10 trials per condition; 150 trials/session). Moreover, 25 of 28 sessions included 15 mild airpuffs (compressed air at 40 psi for 0.3 s) per session in alternating trial blocks to evoke arousal

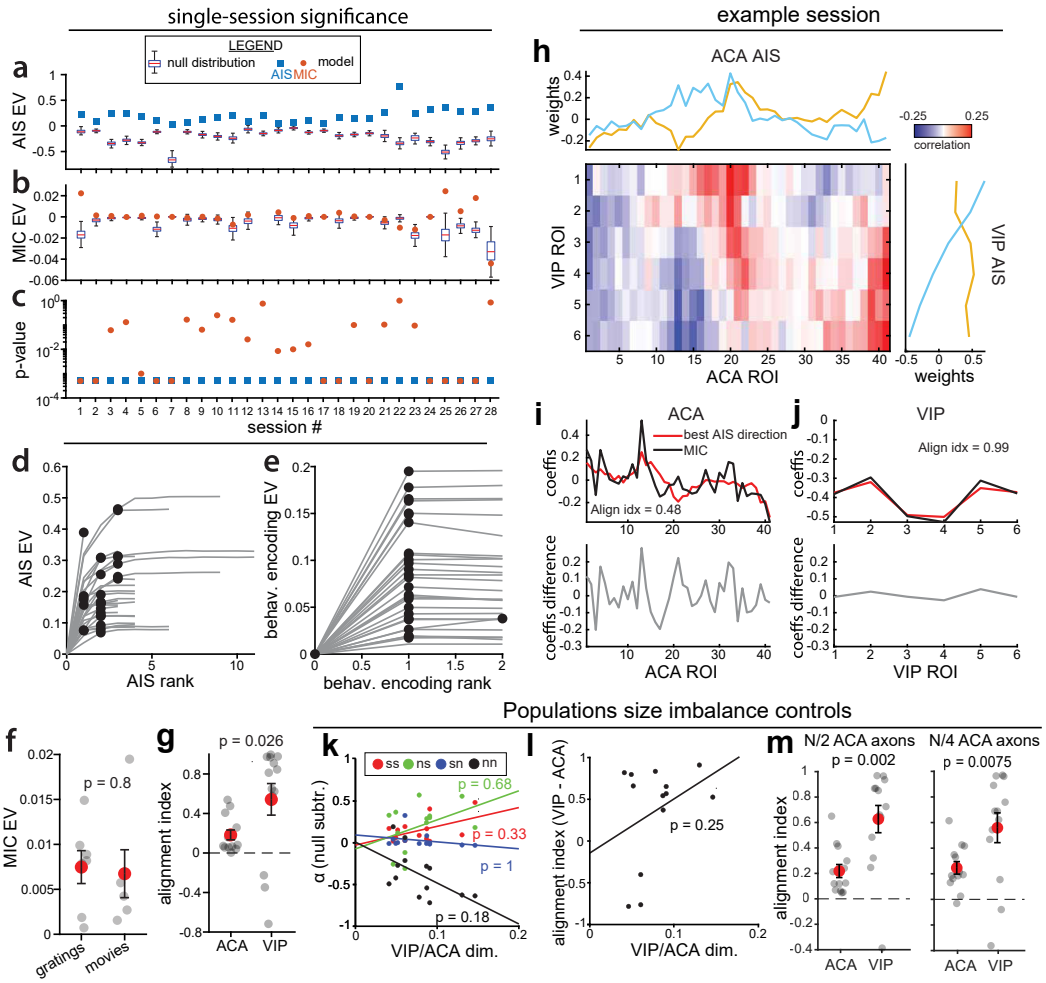

Figure S10: **Additional analyses of ACA-VIP recordings.** (a) Single-session AIS EV significance against a shuffle null (see App. B.13). Boxes: null distribution median and IQR; whiskers: $1.5\times$IQR; dots: mean held-out AIS EV. (b) Same as (a) but for EV MIC. (c) Single-session empirical AIS EV and MIC EV p-values. (d) AIS EV as a function of AIS rank; gray lines, individual session; black dots, single-session selected rank. (e) Same as (d) for behavioral state encoding EV (term B). (f) Comparison of MIC EV in sessions with different visual stimuli (drifting gratings vs. natural movies). Gray circles: individual sessions; red circles: mean $\pm$ SEM across sessions. Only sessions with significant (shuffle test $p < 0.05$, see (a)) are shown. (g) MIC-AIS alignment index for ACA weights (left) and VIP weights (right) across sessions. (h) AIS low-rank decomposition for one example session ($m_A = 2$). Heatmap: ACA-VIP cross-correlation matrix after regressing out behavioral state (Step 1, Alg. 1). Top: ACA AIS weights. Right: VIP AIS weights. For visualization purposes, ACA and VIP ROIs were sorted using optimal leafs order of hierarchical clustering (Euclidean distance, Ward linkage). (i) MIC alignment with AIS for ACA ($m_C = 1$). Top: MIC weights (black) and their best approximation within the AIS (red). Bottom: residual difference. Alignment index reported in panel. (j) Same as (i) for VIP. (k) Null-subtracted geometry indices α_{ss} (red), α_{ns} (green), α_{sn} (blue), and α_{nn} (black) as a function of the VIP:ACA population size ratio. Lines: best linear fits. (l) Same as (k), but for the difference between VIP and ACA MIC-AIS alignment index. (m) Same as (g), after refitting MICs with ACA axons randomly subsampled to half (left) or a quarter (right) of each session's population; per-session indices are averaged over 10 random subsamples. P-values show two-tailed t-test in (f), one-tailed paired t-test in (g,m), and Bonferroni-corrected Pearson's correlation tests in (k,l).

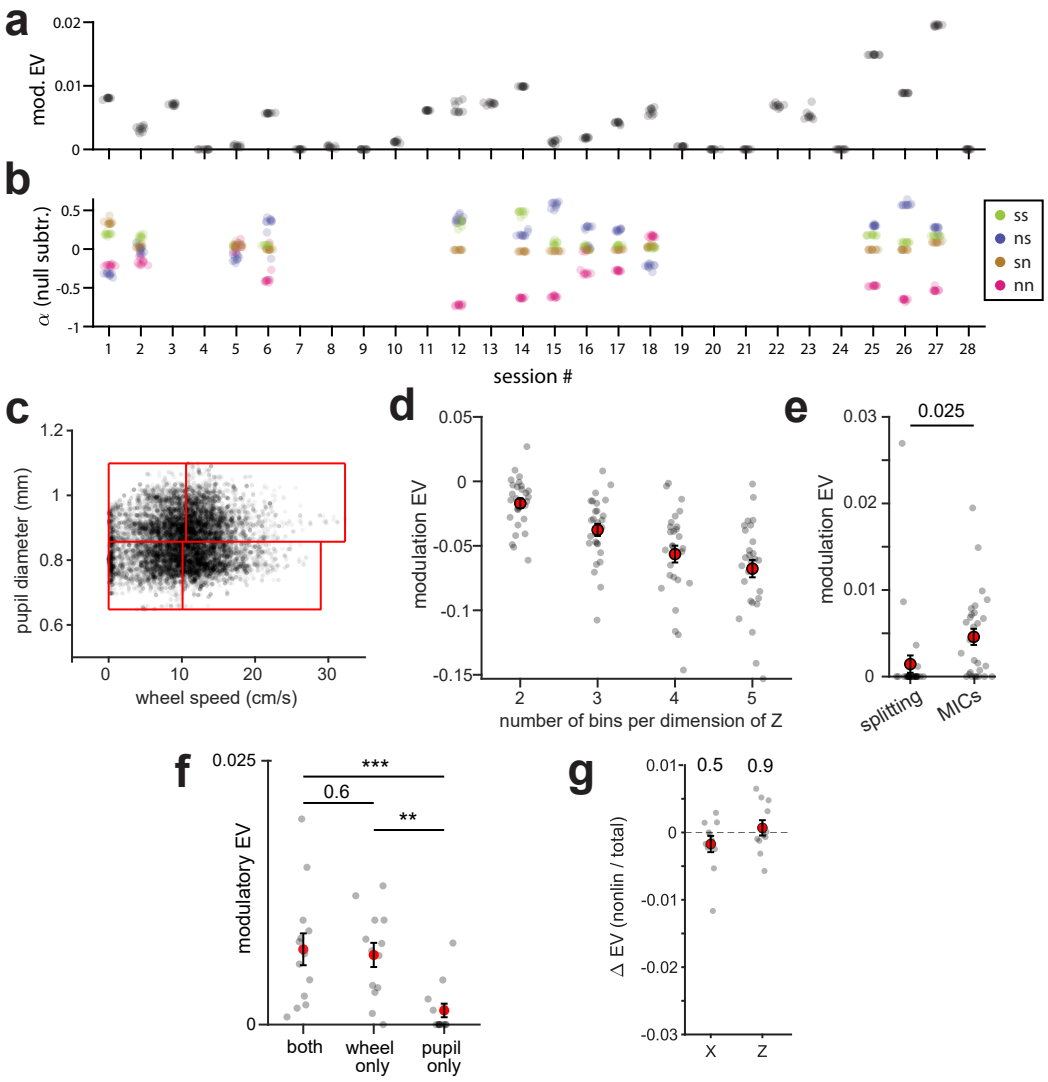

Figure S11: **Robustness and control analyses of ACA-VIP recordings.** For each of the 28 sessions, MICs were re-fitted 10 times with independently drawn random ALS initializations, holding all per-session ranks (m_A, m_B, m_C, m_D) fixed at the values selected in the main text. **(a)** $\mathrm{EV}_{\mathrm{mod}}$ for each session and ALS seed (one dot per seed). **(b)** Null-subtracted geometry indices $\tilde{\alpha}_{ss}$ (green), $\tilde{\alpha}_{ns}$ (blue), $\tilde{\alpha}_{sn}$ (brown), $\tilde{\alpha}_{nn}$ (pink) for the 13 sessions with significant $\mathrm{EV}_{\mathrm{mod}}$, shown for each seed; geometry was not interpreted for sessions without significant modulation. (c) Splitting-based alternative to MICs (Appendix B.15), example session: behavioral state is split into N_{split} wheel-speed quantile bins, each further split into N_{split} pupil-diameter quantile bins (red boxes; $N_{\mathrm{split}} = 2$), and a separate AIS is fit within each bin. (d) Splitting-based $\mathrm{EV}_{\mathrm{mod}}$ as a function of N_{split}. (e) Splitting-based $\mathrm{EV}_{\mathrm{mod}}$ at the best N_{split} per session versus MICs $\mathrm{EV}_{\mathrm{mod}}$. (f) $\mathrm{EV}_{\mathrm{mod}}$ with Z comprising both wheel speed and pupil diameter (both, as for all other analyses), wheel speed only (wheel only), or pupil diameter only (pupil only). (g) Bilinearity diagnostic, as in Fig. S9e. Gray dots: sessions (d, e: all 28; f, g: 13 with significant $\mathrm{EV}_{\mathrm{mod}}$); red: mean ± SEM. P-values: two-tailed t-test in (e,f), one-tailed binomial test in (g). **, $p < 0.01$; ***, $p < 0.001$.

events. Frames were acquired at 33 Hz and every other frame was averaged, yielding an effective sampling rate of ∼16.5 Hz. Imaging used a Prairie Ultima IV two-photon microscope with a MaiTai DeepSee laser (Spectra-Physics) tuned to 965 nm, resonant-galvo scanning (Prairie View 5.4) and a 20× objective (XLPlan N, 1.00 NA, Olympus) at 4× digital zoom. Behavioral apparatus, stimulus presentation and synchronization were as in [59]. Pupil diameter was extracted from 20 Hz face video by tracking 8 pupil-boundary markers with DeepLabCut. It was computed as the horizontal distance between two opposing markers, scaled by the median vertical-to-horizontal diameter ratio, and converted to mm (60.8 pixels/mm). Artifacts (frame-to-frame decreases below the 5th percentile of negative changes) were replaced by modified Akima interpolation, values ≥ 2.5 mm or missing were filled with the preceding value, and the trace was smoothed with a 1 s moving median. Running speed (cm/s) was computed from a rotary encoder (0.0158 cm/count), rectified, resampled to 20 Hz, and smoothed with a causal 0.5 s moving median. Calcium traces were neuropil-corrected as $F_{\text{ROI}} = F_{\text{ROI,raw}} - 0.7\, F_{\text{neuropil}}$. To obtain single-axon traces, ACA boutons were grouped into putative single-axon clusters by hierarchical clustering on pairwise temporal correlations, and within-cluster traces were averaged with weights proportional to ROI signal-to-noise ratio. All ACA-axon and VIP-neuron traces were linearly detrended within session prior to analysis. All variables were linearly interpolated onto a common 20 Hz time vector (∼24,600 samples/session, ∼20.5 min). We used 28 sessions (15 movies; 13 gratings) collected across 3 mice (1 male, 2 female; aged 3-5 months during imaging), taking as response variables the per-axon and per-neuron calcium traces. Included counts ranged 3-212 for ACA axons and 1-11 for VIP neurons. Wheel acquisition started slightly after and ended slightly before imaging; the time points without running speed (1–2% per session) were excluded.

MICs fitting and main analyses. We fit the pipeline of Algorithm 1 separately to each session, with $K = 10$ temporally contiguous cross-validation folds to avoid leakage between training and test sets due to temporal autocorrelation. We took ACA axons as source X and VIP neurons as target Y, consistent with anatomical directionality, and defined the modulator $Z \in \mathbb{R}^2$ as the joint, within-session-standardized vector of pupil diameter and wheel running speed. AIS and MICs ranks were capped at $m_A^{\max} = 20$ and $m_C^{\max} = 10$ for computational efficiency. The MIC correlation slope (Fig.4g) was computed per channel by projecting X, Y, Z onto $w_{X,i}$, $w_{Y,i}$, $w_{Z,i}$, partitioning $Zw_{Z,i}$ into 5 equipopulated quantiles, computing the Pearson correlation between $Xw_{X,i}$ and $Yw_{Y,i}$ within each quantile, and taking the ordinary least squares slope of these five correlations against quantile index; slopes were averaged across MICs and then across sessions. Geometry was characterized via null-subtracted indices $\tilde{\alpha}_{ss}$, $\tilde{\alpha}_{sn}$, $\tilde{\alpha}_{ns}$, $\tilde{\alpha}_{nn}$ (Fig. 4h, top), where $\tilde{\alpha}_{ns}$ and $\tilde{\alpha}_{sn}$ correspond to recruitment of private ACA (source) and VIP (target) dimensions, respectively. Each index was obtained by subtracting from the fitted α the mean over 1000 random rotations of $w_{X,i}$, $w_{Y,i}$ on their unit spheres (Appendix B.7.3). Constrained-fit $\mathrm{EV}_{\mathrm{mod}}$ (Fig. 4h, bottom) further restricts w_X and/or w_Y to lie within the corresponding AIS following Appendix B.8, with "ACA" and "VIP" denoting the constraint applied to the source and target factor, respectively. Statistics for main text analyses were computed as for V1-V2 data (Appendix B.14), with Holm correction across the four constrained-fit comparisons of Fig. 4h. For simplicity, tests in the supplementary figures take sessions as the unit of analysis. Single-session significance of $\mathrm{EV}_{\mathrm{mod}}$ and EV_A was assessed via the residual-shuffle permutation test of Appendix B.13 with $n_{\mathrm{shuf}} = 1000$ shuffles per session (Fig. S10a-c). Across all 28 sessions, $\mathrm{EV}_{\mathrm{mod}}$ correlated positively with the number of recorded VIP neurons (Pearson $r = 0.64$, $p = 1.9 \times 10^{-4}$), consistent with lower sensitivity in sessions with few target neurons.

Additional analyses. To visualize the MIC-AIS geometry in one example session (Fig. S10i), we computed the best approximation of each MIC factor within the corresponding AIS as $\hat{w}_X = P_X w_X$ and $\hat{w}_Y = P_Y w_Y$, with P_X, P_Y the orthogonal projectors onto the source and target AIS. To compare alignment across sessions (Fig. S10g), we summarized per-session MIC-AIS alignment by a normalized index that rescales the observed squared cosine by chance alignment,

$$\mathrm{align} \;=\; \frac{\cos^2\theta_{\mathrm{obs}} - \langle \cos^2\theta_{\mathrm{shuff}} \rangle}{1 - \langle \cos^2\theta_{\mathrm{shuff}} \rangle}, \tag{49}$$

where $\cos^2\theta_{\mathrm{obs}} = \|P\,w\|^2/\|w\|^2$ for the source ($w = w_X$, $P = P_X$) or target ($w = w_Y$, $P = P_Y$) factor (see Appendix B.7.2), and $\langle \cos^2\theta_{\mathrm{shuff}} \rangle$ is the same quantity averaged over uniform random samples of w on the unit sphere, matched in dimension and AIS rank. $\mathrm{align} = 0$ corresponds to chance alignment given the dimensionality; $\mathrm{align} = 1$ to a fitted MIC lying entirely within the AIS. We used a one-tailed test for this analysis under the hypothesis - motivated by geometry analyses in Fig 4 - that VIP MICs are more aligned to the AIS, compared to ACA. To rule out that the asymmetric

alignment between ACA and VIP factors was driven by the typical $n_{\mathrm{ACA}} : n_{\mathrm{VIP}} \approx 15$ population-size imbalance of sessions with significant $\mathrm{EV}_{\mathrm{mod}}$, we regressed the per-session ACA-vs-VIP alignment difference and the four null-subtracted α indices against the ratio $n_{\mathrm{VIP}}/n_{\mathrm{ACA}}$. As a direct control, we refit MICs after randomly subsampling ACA axons to $N/2$ and $N/4$ of each session's population (10 subsamples per session), and compared ACA and VIP alignment indices averaged across subsamples, as in Fig. S10g. VIP alignment remained significantly larger than ACA alignment, with a similar effect size, in both cases (Fig. S10m). Furthermore, to test whether airpuff-evoked arousal events drove the recovered modulation, we ran a control analysis recomputing $\mathrm{EV}_{\mathrm{mod}}$ after excluding the 10 s window following each airpuff onset (not shown); $\mathrm{EV}_{\mathrm{mod}}$ did not differ significantly from the full-session value, suggesting that modulation is driven by endogenous fluctuations in behavioral state rather than puff-evoked transients. Consistently, 2 of the 3 puff-free sessions were among the 13 sessions with significant $\mathrm{EV}_{\mathrm{mod}}$. Finally, to assess robustness of MIC fits to the non-convexity of the ALS objective (Appendix B.5), we re-fitted MICs on each session with 10 independent random ALS initializations, holding the per-session ranks (m_A, m_B, m_C, m_D) fixed at the values selected in the main text. For each seed we recomputed $\mathrm{EV}_{\mathrm{mod}}$ for all sessions and the null-subtracted α indices for the 13 sessions with significant $\mathrm{EV}_{\mathrm{mod}}$ (Fig. S11). Both $\mathrm{EV}_{\mathrm{mod}}$ and the null-subtracted α indices were tightly clustered across seeds, indicating that the ACA-VIP findings are stable under the non-convexity of the ALS objective.

As an alternative to fitting MICs to the ACA-VIP dataset, we split behavioral state into bins and fitted a separate ACA-VIP interaction model in each bin (Fig. S11c-e). Specifically, we first split into N_{split} bins by wheel speed quantile, and then further split into N_{split} sub-bins by pupil diameter quantile, for a total of N_{split}^2 bins. We then followed the same fitting procedure as in Fig. S4 to obtain a splitting-based $\mathrm{EV}_{\mathrm{mod}}$ value directly comparable to that outputted by MICs. Selecting N_{split} in each session to maximize $\mathrm{EV}_{\mathrm{mod}}$ resulted in (mean $\pm$ SEM across session) $\mathrm{EV}_{\mathrm{mod}} = 0.0018 \pm 0.0010$ (compared to $\mathrm{EV}_{\mathrm{mod}} = 0.0046 \pm 0.0009$ for MICs across all sessions; $p = 0.03$, two-tailed t-test). The same temporally contiguous cross-validation folds were used for fitting and assessing the performance of MICs and the splitting alternative. Repeating the same procedure with the order of wheel and pupil reversed when selecting the bins for Z yielded splitting $\mathrm{EV}_{\mathrm{mod}} = 0.0014 \pm 0.0010$ ($p = 0.02$, two-tailed t-test comparing to MICs $\mathrm{EV}_{\mathrm{mod}}$). In sum, increasing the number of splits systematically decreased the performance of the aggregate model, resulting in negative values of $\mathrm{EV}_{\mathrm{mod}}$, confirming that the binning approach is sample-inefficient.

To identify which behavioral variable drives the modulation, we refit MICs with Z comprising wheel speed only or pupil diameter only (Fig. S11f). Dropping pupil diameter did not reduce $\mathrm{EV}_{\mathrm{mod}}$ (0.0066 ± 0.0011 wheel only vs. 0.0071 ± 0.0012 both, $p = 0.64$), whereas dropping wheel speed did (0.0013 ± 0.0006 pupil only; $p = 0.0007$ vs. both, $p = 0.003$ vs. wheel only).

Table 2: Single-session explained variance for the ACA-VIP dataset, for the 13 sessions with significant $\mathrm{EV}_{\mathrm{mod}}$ (all values in %).

Session	$\mathrm{EV}_{\mathrm{mod}}$	$\mathrm{EV}_{\mathrm{AIS}}$	$\mathrm{EV}_{\mathrm{mod}}/\mathrm{EV}_{\mathrm{AIS}}$
1	0.82	19.02	4.31
2	0.27	13.20	2.05
3	0.07	24.10	0.29
4	0.57	12.13	4.70
5	0.79	7.65	10.33
6	0.99	6.88	14.39
7	0.16	7.98	2.01
8	0.19	12.24	1.55
9	0.42	9.21	4.56
10	0.67	16.06	4.17
11	1.49	46.01	3.24
12	0.89	31.35	2.84
13	1.95	28.88	6.75
Mean $\pm$ SEM	0.71 ± 0.15	18.05 ± 3.22	4.71 ± 1.07

B.16 Computational resources

All simulations and real data analyses ran on a workstation with a 24 cores Intel i9-14900K processor and 192 GB of RAM, running Windows 11 and MATLAB R2025b, with parallelization over 20 workers (MATLAB Parallel Computing Toolbox). The pipeline additionally requires the Statistics and Machine Learning Toolbox.

Simulations in Figure 2a-e ran in approximately 10 minutes. All other simulations in Figures 2, 3, and S1–S8 ran in less than one day. Real data analyses in Figures 4 and S9–S11 also completed in under one day on the same machine, except for the residual-permutation test and the constrained model fitting, which took approximately two days.

Code implementing the MICs pipeline, the simulations, and the real-data analyses is available at `https://github.com/mcelotto/multiplicative_interaction_channels`. The full ACA-VIP dual-color two-photon imaging dataset is available at `https://osf.io/2ekj4/`.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Yes, the abstract and the introduction reflect in an accurate manner the topics investigated in the paper and the results obtained.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Yes. The Discussion includes a limitations paragraph covering: (i) the bilinear form of the modulation, which captures only linear effects of Z on the source-target mapping (sensitivity to nonlinear modulation is characterized in Appendix B.10–B.12); (ii) the zero-lag, correlational nature of the method, with source and target assigned from anatomy, so no directional or causal claims are made; (iii) the conservative, data-demanding held-out significance test; (iv) the small number of recorded animals; and (v) the choice of the session mean as baseline. Computational cost and compute resources/runtimes are reported in Appendix B.6 and B.16.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best

judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The paper's sole novel theoretical result is Proposition B.5 (Appendix B.8.1), which shows that the geometry-constrained MIC loss reduces to the unconstrained loss on projected data. Assumptions and proof are provided in the proposition statement and in the proof immediately following it. Theorem B.2 (Appendix B.3) is Kruskal's classical uniqueness theorem [49], cited verbatim and stated with full assumptions for self-containedness. Other appendix material (ALS update equations in B.5 and orthogonal-projector decomposition in B.7) consists of computational derivations and standard algebraic derivations, not novel theoretical claims.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Yes. Reproducibility is ensured by: (i) public release of all code for the MICs pipeline, simulations, and real-data figures (`https://github.com/mcelotto/multiplicative_interaction_channels`); (ii) public release of the full ACA-VIP dataset (`https://osf.io/2ekj4/`); (iii) use of the publicly available V1–V2 dataset [55]; and (iv) full details of simulations (Appendix B.9) and real-data analyses (Appendix B.14–B.15).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.

- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: Code is publicly available at `https://github.com/mcelotto/multiplicative_interaction_channels` under the MIT License, including: (i) the MICs inference pipeline (Algorithm 1), (ii) all simulation-generation and analysis scripts (Figs. 2, 3, S1–S8), (iii) all real-data analysis scripts (Figs. 4, S9–S11), and (iv) running instructions. The full ACA-VIP dataset is publicly available at `https://osf.io/2ekj4/` (CC BY 4.0). The V1–V2 dataset [55] is publicly available on CRCNS (`https://crcns.org/data-sets/vc/v1v2-1`).

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The paper specifies all details. Hyperparameter selection and cross-validation procedures are described in Sec. 3.3 and summarized in Algorithm 1 (Appendix B.6).

Simulation- and real-data-specific parameters are provided in Appendix B.9 and Appendix B.13–B.15, respectively. Code is publicly available on GitHub.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Results are reported as mean ± SEM across simulation repetitions (simulated data) or across experimental sessions (real data), unless stated otherwise in figure captions. Main-text real-data tests (Fig. 4) use linear mixed-effects models with animal as random effect, Holm-corrected across comparisons. Supplementary analyses take sessions as the unit of analysis. The permutation test used to assess the significance of modulatory and AIS explained variance is described in Appendix B.13.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Appendix B.16 reports the computational resources (24 cores Intel i9-14900K processor, 192 GB RAM; Windows 11 / MATLAB R2025b), parallelization (20 cores), and per-figure runtimes. The full research project required no compute beyond what is reported.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.

- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: NeurIPS Codes of Ethics were reviewed, understood and followed.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: There is no societal impact of the work performed.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.

- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The V1–V2 dataset [55] was obtained from CRCNS (`https://crcns.org/data-sets/vc/v1v2-1`), used in accordance with the CRCNS terms of use, and credited to its original source [55] and prior analyses [16, 56] in Sec. 5 and Appendix B.14. The VIP-Cre mouse line (Jackson Laboratory stock no. 031628, RRID:IMSR_JAX:031628) and the AAV1-hSyn-axon-GCaMP6s [67] and AAV9-DIO-jRGECO1a [68] viral constructs are credited to their original sources via RRID and in-text citations in Appendix B.15.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: New code (the MICs inference pipeline and all simulation and real-data analysis scripts) is publicly released on GitHub (`https://github.com/mcelotto/multiplicative_interaction_channels`) under the MIT License, with running instructions. The full 28-session dataset of ACA-axon and V1 VIP-interneuron calcium imaging collected for this study is publicly released on OSF (`https://osf.io/2ekj4/`) under CC BY 4.0, with a README describing the file structure and preprocessing. Data collection procedures are described in Appendix B.15.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.

- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing and it does not involve research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve research with human subjects, and no IRB approval was needed.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs have not been used as an important, original, or non-standard component of the core methods in this research. Their use, if any, was limited to non-substantive tasks and did not impact the scientific rigor or originality of the methodology.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.

- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.