BEYOND GAUSSIAN ASSUMPTIONS: DISTRIBUTION-AWARE CHANNEL CAPACITY FOR EFFECTIVE CONNECTIVITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Effective-connectivity estimation from brain signals often relies on Gaussian residual modeling, which enables tractable estimation but can discard informative distributional structure and distort inferred directed interactions when empirical residuals are non-Gaussian. We show across multiple modalities, species, and experimental conditions that both brain signals and fitted channel residuals frequently deviate from Gaussianity. We therefore introduce a distribution-aware, information-theoretic measure of effective connectivity based on channel capacity under general residual distributions. To estimate the resulting capacity from empirical, potentially non-Gaussian residuals, we develop a dual-flow min–max estimator based on normalizing flows, in which a generator searches over admissible input distributions under a power constraint while an observer estimates output entropy. We provide a theoretical characterization of the estimator, showing that the observer objective recovers differential entropy up to a KL approximation term, that the formulation reduces to classical Gaussian capacity as a special case, and that residual entropy can alter achievable information rates beyond variance; game-gap and error analyses further characterize optimization and approximation sources. In brain-like simulations with known directed connectivity, Dual-flow achieves the highest AUROC and AUPRC across ten conditions spanning diverse network topologies, hidden drivers, feedback, and heterogeneous hemodynamics, compared with Gaussian capacity, Granger causality, VAR-LiNGAM, and GIMME. Applied to multimodal brain signals, the method reveals time- and condition-resolved directed interactions consistent with known neurobiological circuitry. Together, these results establish a principled distribution-aware framework for effective-connectivity estimation beyond Gaussian residual modeling.

1 INTRODUCTION

Estimating time-resolved directed interactions among brain regions is a central problem in computational neuroscience and neuroimaging (Friston, 2011). Effective-connectivity (EC) methods aim to move beyond undirected association by quantifying how activity in one region predicts, drives, or transmits information to another (Friston, 2011; Valdes-Sosa et al., 2011; Bielczyk et al., 2019). Such directed measures are particularly important for understanding dynamic brain states, task-evoked processing, and multimodal brain-signal measurements, including BOLD fMRI, EEG, LFP, and optical neuroimaging (Seth et al., 2015; Pan et al., 2011; Lake et al., 2020; Vafaii et al., 2024).

EC estimation is model-based: different approaches define method-specific measures of directed interaction under different assumptions (Valdes-Sosa et al., 2011; Deshpande & Hu, 2012; Bielczyk et al., 2019). A common strategy for making EC estimation tractable is to impose Gaussian, linear, or second-order statistical assumptions (Granger, 1969; Lütkepohl, 2005; Barnett & Seth, 2014; Friston et al., 2003; Stephan et al., 2010). These assumptions underlie many practical estimators and enable closed-form analysis or stable finite-sample estimation (Barnett & Seth, 2014; Friston et al., 2003; Cover & Thomas, 2006). However, whether such Gaussian assumptions are empirically justified across diverse brain-signal modalities remains an open question. Brain signals may exhibit heavy-tailed, skewed, burst-like, or state-dependent structure arising from latent brain states, physiological variability, nonlinear measurement processes, or artifacts (Zhang et al., 2020; Mitra et al., 2018; Mandino et al., 2025; Vafaii et al., 2024). If such distributional structure is present, Gaussian

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

estimators may obscure or distort directed interactions (Seth et al., 2015; Friston et al., 2003; Stephan et al., 2010). Prior work has often adopted modality-specific distributional assumptions for tractable analysis—for example, Rician models for MRI magnitude noise (Gudbjartsson & Patz, 1995; Noh & Solo, 2011; Wegmann et al., 2017) and approximate Gaussian models in many resting-state fMRI and electrophysiology analyses (Granger, 1969; Lütkepohl, 2005; Barnett & Seth, 2014; Pan et al., 2011). Non-Gaussian approaches such as LiNGAM exploit non-Gaussian disturbances to identify causal structure (Shimizu et al., 2006; Hyvärinen & Smith, 2013). However, a systematic cross-modal diagnosis of brain-signal and fitted channel-residual distributions, especially in the context of EC estimation, has received relatively little attention.

Channel capacity quantifies the maximum information rate supported by a channel under specified input constraints (Shannon, 1948; Cover & Thomas, 2006; Gallager, 1968). In EC analysis, a directional interaction between brain regions can be modeled as an empirical communication channel, whose capacity provides an information-theoretic measure reflecting both channel dynamics and the residual distribution (Cover & Thomas, 2006; Jian et al., 2026). This allows EC estimates to capture distributional structure beyond second-order statistics. For linear additive Gaussian channels, capacity can be computed analytically or semi-analytically (Cover & Thomas, 2006; Gallager, 1968). However, Gaussian formulas characterize residual uncertainty through covariance alone and may miss capacity differences arising from non-Gaussian residual structure.

Estimating channel capacity beyond Gaussian settings is challenging because it requires both optimizing over admissible input distributions and estimating output entropy under unknown, potentially non-Gaussian noise (Cover & Thomas, 2006; Gallager, 1968). Classical nonparametric entropy estimators can be unstable in finite samples and are difficult to embed inside this optimization (Kozachenko, 1987), while neural mutual-information estimators typically target a fixed joint distribution rather than the capacity-achieving input distribution (Belghazi et al., 2018; Poole et al., 2019). Thus, a practical estimator must support flexible distribution modeling while remaining compatible with constrained capacity maximization (Rezende & Mohamed, 2015; Dinh et al., 2016).

In this work, we challenge the Gaussian assumption in brain-signal EC analysis. We assess brain-signal Gaussianity across multiple modalities and species and examine the distributions of fitted channel residuals. Motivated by the observed deviations, we introduce a distribution-aware, information-theoretic EC measure based on channel capacity. We model each directed pair of regions of interest (ROIs) as a finite impulse response (FIR) channel, using fitted residuals as samples from an unknown continuous noise distribution. To estimate this measure, we develop a dual-flow min–max formulation that couples power-constrained input-distribution optimization with output-entropy estimation. The formulation recovers Gaussian capacity as a special case. Its assumptions of local linear time-invariance and additive white noise may limit applicability to strongly nonlinear dynamics or temporally correlated residuals. This proposed method is summarized in Fig. 1.

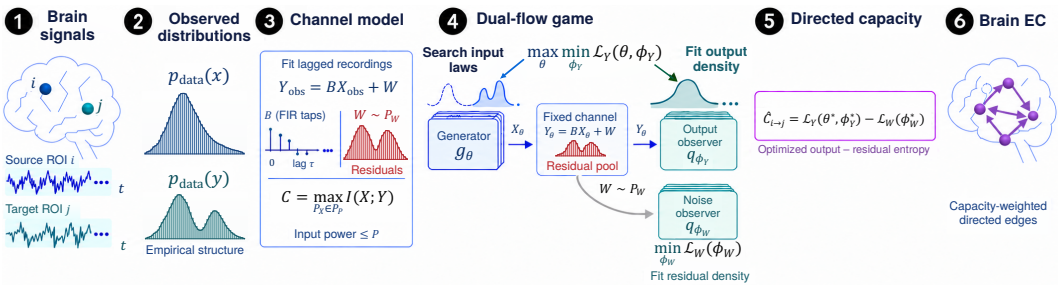


Figure 1: **Proposed distribution-aware channel-capacity framework for EC.** Steps 1 & 2: Select the source and target ROIs and retrieve their signals. Step 3: Perform a linear regression to obtain the linear channel model and a residual pool. Steps 4 & 5: Compute the channel capacity using a dual-flow minimax game. Step 6: Repeat for all directed ROI pairs to obtain the EC graph.

Contributions. We make four contributions. First, we assess Gaussianity across brain-signal modalities and conditions, documenting departures in observed signals and fitted channel residuals. Second, we formulate distribution-aware channel-capacity EC using empirical residual resampling under an unknown continuous noise distribution. Third, we develop a dual-flow adversarial normalizing-flow estimator for capacity beyond Gaussian settings. Fourth, we demonstrate superior directed-edge recovery on simulated brain-like signals with ground-truth EC spanning diverse network topologies

and hemodynamic confounds. Applications to brain signals reveal time- and condition-resolved directed interactions consistent with known neurobiological circuitry.

2 Background and related work

EC methods under Gaussian and non-Gaussian assumptions. Widely used EC methods include multivariate autoregressive and Granger-causal models (Granger, 1969; Lütkepohl, 2005; Barnett & Seth, 2014), as well as information-theoretic measures such as transfer entropy and directed information (Schreiber, 2000; Massey, 1990). Many tractable estimators in these families rely on Gaussian residual assumptions or second-order statistics, which may inadequately capture heavy-tailed, skewed, or state-dependent distributions in empirical brain signals and fitted residuals. Although modality-specific distributional models have been studied, systematic cross-modal assessment of these assumptions in EC estimation remains limited.

Non-Gaussian EC methods also exist. LiNGAM-family methods exploit non-Gaussian disturbances to identify causal directions under structural and independence assumptions (Shimizu et al., 2006; Hyvärinen & Smith, 2013). Our framework instead uses the empirical residual distribution to estimate channel-capacity EC while optimizing the input distribution under a power constraint. Gaussian capacity (GCap) assumes Gaussian residuals, whereas our Dual-flow estimator uses empirical residual samples without this assumption. Further background and interpretation of channel-capacity EC are provided in Appendix A.1, with the evaluated baseline methods detailed in Appendix A.2.

Normalizing flows for distribution-aware entropy estimation. Normalizing flows are invertible density models that transform a simple base distribution into a flexible target distribution while retaining exact likelihood evaluation through the change-of-variables formula (Rezende & Mohamed, 2015; Dinh et al., 2016). Because differential entropy can be estimated from log densities, flows provide a practical tool for entropy estimation when the target distribution is unknown and potentially non-Gaussian. In the present work, we build on this property with an adversarial generator–observer flow architecture: the generator searches over admissible input distributions under a power constraint, while the observer estimates output entropy from samples generated by the empirical channel. Unlike Gaussian-capacity estimators, our framework uses empirically resampled residual noise, allowing channel-capacity estimates to reflect non-Gaussian residual structure in brain-signal measurements.

3 Distribution-aware channel capacity for EC

3.1 Empirical FIR Channels with Residual Resampling

For each temporal window or condition segment, let $\{x_t\}_{t=1}^{T}$ and $\{y_t\}_{t=1}^{T}$ denote brain-signal time series from a source and target ROI, and approximate the directed interaction $x \to y$ by a locally stationary finite impulse response channel capturing finite-lag effects,

$$y_t = \sum_{\ell=0}^{k-1} b_\ell x_{t-\ell} + w_t, \qquad t = k, \dots, T, \tag{1}$$

where $b = (b_0, \dots, b_{k-1})$ is the channel filter, w_t is additive residual noise, and T denotes the number of time points in the analyzed segment. For each ROI pair and segment, we estimate $\widehat{b}$ by least squares and compute residuals

$$\widehat{w}_t = y_t - \sum_{\ell=0}^{k-1} \widehat{b}_\ell x_{t-\ell}, \qquad t = k, \dots, T. \tag{2}$$

Unlike Gaussian-noise channel models, we do not impose a parametric form on the residual distribution. We assume that the additive noise W is drawn from an unknown continuous distribution P_W with finite differential entropy, and use the fitted residuals as empirical samples from this distribution. During training, expectations over P_W are approximated by uniformly resampling from the residual pool $\{\widehat{w}_t\}_{t=k}^{T}$. The residual pool is not treated as a discrete noise distribution; rather, the residuals are finite samples from an unknown continuous residual-noise distribution P_W. Thus the differential-entropy objective is defined with respect to the underlying continuous distribution, while empirical residual resampling provides Monte Carlo minibatches for the flow-based estimator. The fidelity of

this approximation depends on residual-pool size and representativeness, the adequacy of the fitted FIR model, and approximate within-segment stationarity (Appendix C.1).

Stacking Eq. 1 over valid samples gives the vector channel

$$Y = BX + W, \qquad W \sim P_W, \tag{3}$$

where B is the convolution operator induced by the FIR filter. Within each segment, we assume that the additive FIR approximation is locally valid and that W is independent of admissible inputs X. Under an average input-power budget $P > 0$, the admissible distributions form the set $\mathcal{P}_P = \{P_X : \mathbb{E}_{P_X}\|X\|_2^2 \le dP\}$, where d is the input dimension and P is the allowed power per sample. Under this empirical channel formulation,

$$C = \sup_{P_X \in \mathcal{P}_P} I(X;Y) = \sup_{P_X \in \mathcal{P}_P} h(BX + W) - h(W), \qquad W \sim P_W. \tag{4}$$

Thus capacity estimation reduces to maximizing output entropy while subtracting residual entropy; a derivation of Eq. 4 is given in Appendix B.1. This identity also clarifies why residual shape matters: Gaussian capacity depends only on second-order residual statistics, whereas the general additive-channel formulation depends on the full residual entropy.

3.2 Dual-Flow Capacity Estimator

Eq. 4 requires optimizing over admissible input distributions while estimating differential entropy under unknown residual noise. We introduce a Dual-flow estimator to address this by combining three complementary ingredients: an information-theoretic channel-capacity objective, a game-theoretic generator–observer interaction, and normalizing flows for flexible non-Gaussian distribution modeling. The generator searches over power-constrained inputs, while the observer estimates the induced output entropy (Fig. 2).

The generator g_θ maps Gaussian latent samples $Z_1 \sim \mathcal{N}(0, I)$ to candidate inputs and enforces the power constraint by RMS normalization,

$$X_\theta = \mathrm{Norm}_P(g_\theta(Z_1)), \qquad Y_\theta = BX_\theta + W, \quad W \sim P_W. \tag{5}$$

The observer q_ϕ is an invertible flow that induces the density

$$p_\phi(y) = p_U(q_\phi(y))\left|\det J_{q_\phi}(y)\right|, \qquad U \sim \mathcal{N}(0, I). \tag{6}$$

Its negative log-likelihood is

$$L_\phi(y) = -\log p_\phi(y) = -\log\left|\det J_{q_\phi}(y)\right| + \frac{1}{2}\|q_\phi(y)\|_2^2 + \frac{d}{2}\log(2\pi). \tag{7}$$

Lemma 1 (Observer-flow entropy identity)**.** *Let P_Y be the output distribution and let p_ϕ be the density induced by the observer flow. Then*

$$\mathbb{E}_{Y\sim P_Y}[-\log p_\phi(Y)] = h(Y) + \mathrm{KL}(P_Y\|P_\phi). \tag{8}$$

Consequently, for a sufficiently expressive observer family, minimizing the expected negative log-likelihood estimates $h(Y)$, with equality when $P_\phi = P_Y$.

For the signal-plus-noise output and residual-noise terms, define

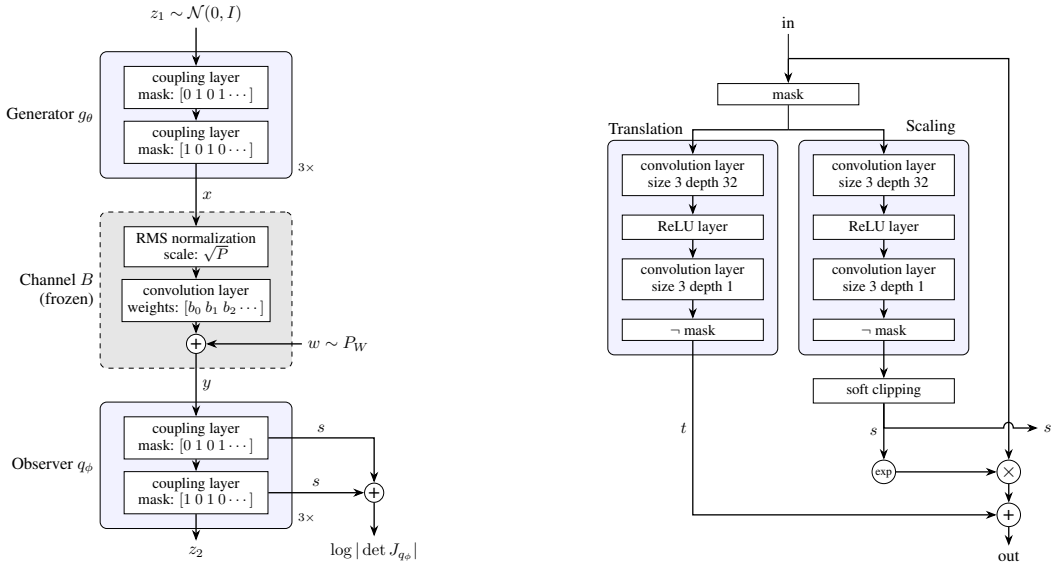
$$\mathcal{L}_Y(\theta, \phi_Y) = \mathbb{E}\left[-\log p_{\phi_Y}\left(B\mathrm{Norm}_P(g_\theta(Z_1)) + W\right)\right], \tag{9}$$

$$\mathcal{L}_W(\phi_W) = \mathbb{E}[-\log p_{\phi_W}(W)]. \tag{10}$$

The resulting flow-restricted capacity estimator is

$$\widehat{C}_{\Theta,\Phi} = \max_{\theta\in\Theta}\min_{\phi_Y\in\Phi}\mathcal{L}_Y(\theta,\phi_Y) - \min_{\phi_W\in\Phi}\mathcal{L}_W(\phi_W). \tag{11}$$

The generator searches for an admissible input distribution that maximizes the observer-estimated output entropy, while the observer estimates the induced output density. We use the adversarial terminology to describe this optimization structure, but do not claim strong duality for the nonconvex neural parameterization.



(a) Dual-flow estimator.

(b) 1D Real NVP layer.

Figure 2: **Dual-flow estimator and flow architecture.** (a) The generator searches over admissible input distributions under a power constraint, while the observer estimates channel-output entropy. (b) A 1D Real NVP coupling layer used in the generator and observer.

Both the generator and observer use 1D Real NVP flows with three coupling-layer pairs (Fig. 2b; architecture details in Appendix C.1). In each training iteration, we generate a training sequence of length L by sampling a latent sequence $Z_1 \sim \mathcal{N}(0, I_L)$ and drawing L residual-noise samples by uniformly resampling from the fitted residual pool. These samples are propagated through the generator, empirical FIR channel, and observer to obtain Monte Carlo estimates of the entropy objectives. We alternate observer updates minimizing negative log-likelihood with generator updates maximizing the signal-plus-noise entropy estimate; gradients are computed by automatic differentiation and parameters are optimized with *Adam*. The full training procedure is given in Algorithm 1 (Appendix C.1). Capacity estimates showed limited sensitivity to the generated training-sequence length L over the tested range (Appendix C.2).

To estimate capacity, we train signal-plus-noise and noise-only objectives in parallel using matched residual samples. The noise-only objective estimates $h(W)$ by applying an observer flow directly to resampled residual noise. This paired estimation reduces Monte Carlo variance in the entropy difference. The final estimate is

$$\widehat{C} = \widehat{h}(Y) - \widehat{h}(W). \tag{12}$$

Convergence is declared when an exponential moving average of this paired entropy difference reaches stationarity.

3.3 Key Theoretical Properties

We next characterize the key theoretical properties of the Dual-flow formulation. We first establish the observer-flow entropy identity, then show that the formulation recovers classical Gaussian capacity as a special case and that residual distributional structure can affect capacity beyond variance. We further characterize optimization and approximation errors of the generator–observer game. Corresponding controlled checks are reported in Appendix C.3.

Entropy estimation by the observer flow. Lemma 1 shows that the observer negative log-likelihood equals the target entropy plus a KL approximation error, justifying the minimized observer loss as an entropy estimator inside the capacity objective (Appendix B.2). In controlled settings with analytically known noise entropy, the observer accurately recovered the target entropy across Gaussian and non-Gaussian residual laws (Fig. 6A-i; Appendix C.3).

Gaussian recovery. When $W \sim \mathcal{N}(0, \Sigma_W)$, the population capacity formulation in Eq. 4 reduces to the classical Gaussian-channel capacity:

$$C_{\mathrm{G}} = \max_{\Sigma_X \succeq 0,\ \mathrm{tr}(\Sigma_X) \le dP} \frac{1}{2} \log\det\left(I + \Sigma_W^{-1/2} B \Sigma_X B^\top \Sigma_W^{-1/2}\right). \tag{13}$$

Thus the generalized formulation includes classical Gaussian capacity as a special case while allowing the residual uncertainty term $h(W)$ to reflect empirical non-Gaussian structure; the formal derivation is given in Appendix B.3. In controlled Gaussian FIR channels with known analytic capacity, the full Dual-flow estimator recovered the corresponding capacity across representative channel structures (Fig. 6A-ii; Appendix C.3).

Residual entropy beyond residual variance. For a scalar additive channel $Y = X + W$ with $\mathbb{E}[X^2] \le P$, $\mathrm{Var}(W) = \sigma_W^2$, and entropy power $N(W) = (2\pi e)^{-1} \exp(2h(W))$, the entropy power inequality gives the achievable-rate lower bound

$$C(W) \ge \frac{1}{2} \log\left(1 + \frac{P}{N(W)}\right) \ge \frac{1}{2} \log\left(1 + \frac{P}{\sigma_W^2}\right), \tag{14}$$

with equality in the second inequality only when W is Gaussian. Thus residual distributions with matched variance but different entropy power can support different achievable rates under the same input-power constraint. A matched-Gaussian estimator cannot distinguish these cases because it replaces the residual distribution by a Gaussian approximation with the same second-order statistics. The theoretical result and its FIR extension are derived in Appendix B.4; the corresponding controlled matched-variance experiment is shown in Fig. 6B (Appendix C.3).

Remark 1 (Approximation sources)**.** *The finite-sample estimate differs from the population capacity because of residual/channel estimation, generator restriction, observer density approximation, nonconvex optimization, and Monte Carlo sampling. We summarize these sources as*

$$|\widehat{C}_{\Theta,\Phi} - C| \lesssim \varepsilon_{\mathrm{resid}} + \varepsilon_{\mathrm{gen}} + \varepsilon_{\mathrm{obs}} + \varepsilon_{\mathrm{opt}} + \varepsilon_{\mathrm{MC}}. \tag{15}$$

This decomposition is meant as a practical guide to estimator limitations rather than a finite-sample guarantee. Additional details, including the game-gap diagnostic for the flow-restricted optimization problem and the corresponding error decomposition, are provided in Appendix B.5 and Appendix B.6.

3.4 Time- and Condition-Resolved EC Estimation

The same empirical residual-aware estimator can be applied either to sliding temporal windows or to pre-defined experimental conditions. For continuously recorded data, such as resting-state LFP–BOLD recordings, we estimate dynamic directed connectivity in sliding windows. For block-structured task data, such as motor fMRI, we estimate condition-resolved directed connectivity across independently defined task and rest windows rather than interpreting the estimates as a continuous temporal trajectory. For each window or condition segment τ and directed ROI pair $i \to j$, we estimate the FIR filter, resample from the fitted residual pool, and compute $\widehat{C}_{i\to j}(\tau)$.

The FIR order k is selected separately for each ROI pair and segment. We use BIC for sufficiently long windows and corrected AIC for shorter windows with low sample-to-parameter ratios (Hurvich & Tsai, 1989). Given the selected k, FIR coefficients are estimated by least squares, and the resulting residuals are retained as empirical samples from the unknown residual-noise distribution rather than replaced by a parametric Gaussian model.

4 Results

The experiments serve three complementary roles: distributional diagnosis, ground-truth validation, and real-data application. We first diagnose departures from Gaussianity in brain signals across modalities and in fitted rat BOLD channel residuals (Section 4.1; Appendix D.1). We then validate directed-edge recovery in brain-like networks with known ground-truth EC across diverse network and hemodynamic conditions (Section 4.2; Appendices E). Finally, real-data analyses assess task sensitivity and biological plausibility in tongue-motion fMRI (Section 4.3; Appendix D.2), with complementary neurophysiological evidence from cross-modal correspondence (Appendix D.3). Computational requirements and runtime are reported in Appendix F.

4.1 Empirical Brain Signals and Channel Residuals Exhibit non-Gaussian Structure

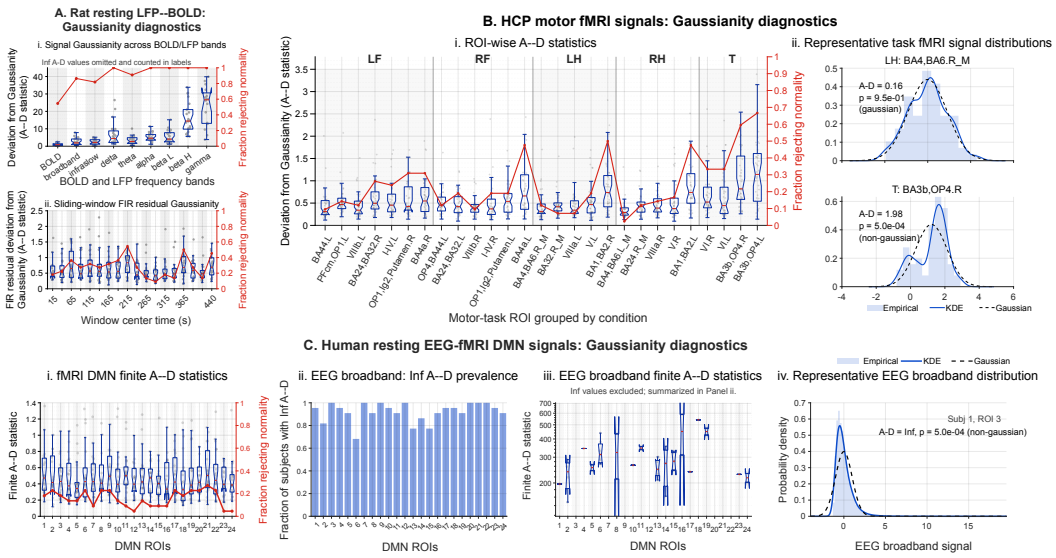


Figure 3: **Gaussianity diagnostics across neural signal modalities and datasets.** Anderson–Darling (A–D) statistics quantify deviations from matched Gaussian distributions. **A**, Rat resting LFP–BOLD signals show frequency-dependent signal non-Gaussianity (*i*); sliding-window FIR residuals show time-varying Gaussianity, with the red curve indicating the fraction of subjects rejecting residual normality (*ii*). **B**, HCP motor fMRI shows ROI- and task-dependent Gaussianity variation (*i*) with representative task-ROI distributions and matched Gaussian fits (*ii*). **C**, Concurrent EEG–fMRI DMN signals show modest finite A–D statistics for fMRI BOLD (*i*) but widespread infinite A–D statistics for broadband EEG; infinite values are summarized by prevalence (*ii*), excluded from finite-value boxplots (*iii*), and illustrated by a representative distribution (*iv*). Dataset descriptions, preprocessing, and Gaussianity-analysis details are provided in Appendix D.1.

Many EC models assume, either explicitly or implicitly, that model residuals, prediction errors, or observation noise can be adequately characterized by second-order statistics. This assumption is central to linear autoregressive and Granger-causal frameworks, where directed interactions are inferred from reductions in prediction-error variance or covariance under VAR/MVAR models (Granger, 1969; Bressler & Seth, 2011; Seth et al., 2015). Related Gaussian noise assumptions also appear in generative EC frameworks such as dynamic causal modeling, where observation and state noise are commonly modeled within Gaussian state-space or variational inference formulations (Friston et al., 2003; Stephan et al., 2010).

Across rat resting LFP–BOLD recordings (Fig. 3A), HCP motor-task fMRI (Fig. 3B), and human resting simultaneous EEG–fMRI datasets (Fig. 3C), empirical signal distributions were not uniformly Gaussian. Rat LFP–BOLD recordings showed frequency-dependent deviations from Gaussianity, with stronger departures in higher-frequency LFP bands. HCP motor fMRI showed ROI- and task-dependent variation in Anderson–Darling (A–D) statistics, and representative task-ROI distributions showed visible mismatch from matched Gaussian fits. In simultaneous EEG–fMRI, BOLD signals from default mode network (DMN) regions of interest (ROIs) showed relatively modest finite A–D statistics, whereas broadband EEG-derived DMN signals frequently produced numerically infinite A–D statistics, indicating severe departures from Gaussianity.

These signal-level diagnostics are directly relevant to EC estimation because Gaussian residual or prediction-error models reduce uncertainty to variance or covariance. Under the additive FIR channel model (Eq. 3), the residual distribution P_W directly determines the uncertainty term entering capacity estimation. We therefore fit local FIR channels in sliding windows and examined the resulting 100-sample residual time series in each window. In the rat resting BOLD channel, residual A–D statistics varied across windows, and the fraction of subjects rejecting residual normality also changed over time (Fig. 3A-ii). Thus, the issue is not only that neural signals can be non-Gaussian, but also that the channel residual distribution can be time-varying and only locally Gaussian-like (see

Fig. 7a in the Appendix for a sample time-varying noise histogram in the LFP-BOLD dataset). Together, these observations motivate estimating P_W from empirical residual samples rather than approximating residual uncertainty solely through Gaussian covariance structure. Fig. 7b shows that larger signal-level Anderson–Darling statistics were associated with greater differences between GCap and Dual-flow capacity estimates.

4.2 Brain-Like Simulations with Ground-Truth EC

Because ground-truth EC is unavailable in neuroimaging, we validated Dual-flow's EC-recovery capability across ten brain-like simulation conditions against Gaussian capacity (GCap), its within-framework matched-Gaussian counterpart, as well as GC, VAR-LiNGAM, and GIMME (see Appendix A.2 for baseline methods). The conditions capture network-level and hemodynamic variation challenges relevant to brain EC estimation, including indirect pathways, hidden drivers, feedback, modular organization, heterogeneous HRFs, their coexistence, and increasing network scale (3–28 ROIs; 50 realizations per condition). Simulation design and ground-truth EC topologies are detailed in Appendix E.1, with the evaluation metrics in Appendix E.2.

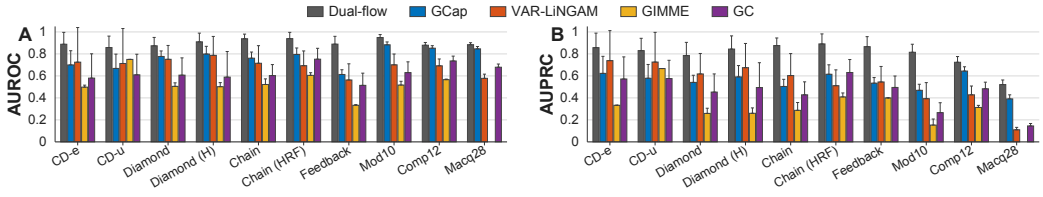


Figure 4: **Directed-edge recovery across brain-like simulations.** Bars show mean AUROC (**A**) and AUPRC (**B**) across 50 realizations per condition; error bars denote $+1$ SD (upper half shown for clarity). Dual-flow achieved the highest mean AUROC and AUPRC in every condition. Full results are reported in Appendix E.2.

Dual-flow achieved the highest AUROC and AUPRC in all ten conditions (AUROC .858–.949; AUPRC .521–.891; Fig. 4). In the common-driver conditions, edge ranking remained strong despite larger variability in threshold-dependent metrics; because each realization contains only two true and four null edges, sensitivity and FPR change in coarse increments of .50 and .25, respectively, making these metrics intrinsically discrete. The Diamond, hidden-Diamond, and Chain conditions tested robustness to convergent and indirect pathways and to an unobserved common driver, while Chain (HRF) isolated regional hemodynamic heterogeneity. The Feedback and Mod10 conditions extended the evaluation to reciprocal/cyclic interactions and modular organization. Performance remained strong when these challenges coexisted in Comp12 (AUROC/AUPRC .879/.724) and in the sparse 28-ROI recurrent Macq28 network (.884/.521), where only 52 of 756 candidate directed edges were true (random AUPRC .069). Together, these results show that Dual-flow preserves directed-edge ranking across increasingly challenging network and hemodynamic conditions, including their coexistence and larger network scale. Full results with additional metrics are reported in Appendix E.2.

4.3 Distribution-Aware Capacity Detects Directed Interactions in Tongue-Motion fMRI

We applied the proposed estimator to the tongue-motion condition of the motion-evoked fMRI dataset, as this condition showed the strongest departure from Gaussianity among the movement conditions in Fig. 3B-i (see Appendix D.2 for implementation details). For each subject, condition, and directed ROI pair, we fit a local FIR channel and computed four directed-connectivity estimates: Dual-flow, GCap, GC, and VAR-LiNGAM. This analysis evaluates whether Dual-flow can detect task-related directed interactions beyond GCap, GC, and VAR-LiNGAM in a condition whose signals showed relatively strong departure from Gaussianity, rather than focusing on within-run temporal variability. Following established hypothesis-driven EC practice that restricts analysis to task-relevant ROIs and prespecified circuits (Smith et al., 2012; Deshpande & Hu, 2012), we predefined four tongue-motion ROIs from the HCP motor dataset's task-activation maps, following its established activation-mapping framework (Barch et al., 2013); details are provided in Appendix D.1.

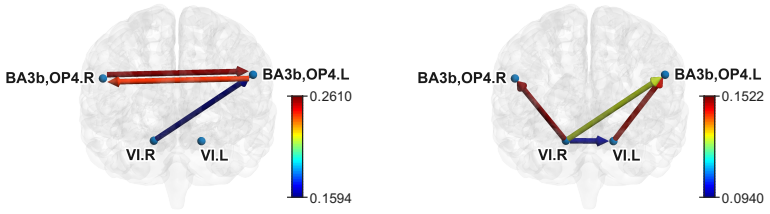


Figure 5: **Task-related directed connections of Dual-flow (left) and GCap (right)** visualized in BrainNet Viewer (Xia et al., 2013). Result with all four EC metrics, as well as their respective connectivity matrices in the task state and rest states, are reported in Fig. 8.

For the tongue-motion task, prior literature on contralateral cerebello-cortical organization and bilateral orofacial sensorimotor control supports four physiologically plausible candidate interactions: (i) VI.R $\longrightarrow$ BA3b, OP4.L; (ii) VI.L $\longrightarrow$ BA3b, OP4.R; (iii) BA3b, OP4.L $\longrightarrow$ BA3b, OP4.R; and (iv) BA3b, OP4.R $\longrightarrow$ BA3b, OP4.L (Manto et al., 2012; Sasegbon & Hamdy, 2023; Rizzolatti & Luppino, 2001; Todorov, 2004). These literature-supported interactions provide a physiologically plausible reference but not ground-truth labels for the exact directed edges in the present fMRI data. Task-sensitive edges that were estimated by each EC method were required to meet both FDR-corrected tongue-versus-rest significance ($q < 0.1$) and a prespecified cross-subject stability criterion (CV below the 30th percentile); see Appendix D.2 for implementation details and Table 3 for numerical results. As shown in Fig. 5, Dual-flow identified three of the four candidates—(i), (iii), and (iv)—while the missing mirrored cerebello-cortical interaction (ii) may reflect asymmetry in the sampled data or limited statistical power. GCap did not identify interaction (iii) and instead identified connections between the left cerebellum and left sensorimotor cortex; this difference is consistent with the stronger departure from Gaussianity observed in motor-cortical signals (Fig. 3B-i), although the real-data analysis does not isolate non-Gaussianity as its sole cause. GC and VAR-LiNGAM did not identify any significant tongue-evoked directed connections (Table. 3).

5 Conclusion

We introduced a distribution-aware channel-capacity framework that incorporates the distributional structure of brain signals and fitted residuals into effective-connectivity (EC) estimation. Rather than reducing residual uncertainty to Gaussian second-order statistics, Dual-flow uses empirical residual samples and optimizes the admissible input distribution to estimate the information-carrying capacity of a fitted directional interaction under general residual distributions. Across modalities and conditions, brain-signal and fitted-residual distributions frequently departed substantially from Gaussianity, providing empirical support for explicitly incorporating distributional structure into EC estimation. We theoretically characterize the framework by establishing Gaussian-capacity recovery, the effect of residual structure beyond variance on achievable information rates, and the main approximation and optimization errors. In brain-like simulations with known ground-truth EC, Dual-flow showed strong directed-edge recovery across diverse network and hemodynamic challenges and consistently outperformed alternative baselines. Empirical applications further demonstrate its potential as a practical tool for time- and condition-resolved EC analysis, with potential value for both basic neuroscience and translational studies of task-evoked or disrupted effective connectivity.

Limitations. The current formulation assumes a linear FIR channel with additive noise within each temporal window or condition segment, which may be inadequate for strongly nonlinear dynamics. The current implementation is designed for hypothesis-driven EC analysis of prespecified circuits and moderate-sized subnetworks, an established use case in cognitive and clinical neuroimaging (Smith et al., 2012; Deshpande & Hu, 2012; Bielczyk et al., 2019). A remaining limitation is that it does not yet scale efficiently to unrestricted whole-brain exploratory analysis, which represents a distinct use case aimed at broader network-discovery questions (Deshpande & Hu, 2012; Bielczyk et al., 2019).

AI USE STATEMENT

We used generative AI tools to help refine hypotheses and diagnostics, and improve the organization and wording of the manuscript. The authors checked the resulting derivations, references, figures, and claims against the cited sources and experiments. The authors take responsibility for the final content of the paper and all AI-assisted artifacts.

ETHICS STATEMENT

This work uses three previously collected datasets. The human motor-task fMRI data are from the Human Connectome Project (Barch et al., 2013), collected under protocols approved by the Washington University in St. Louis institutional review board with written informed consent from all participants. The simultaneous EEG–fMRI data are publicly available through OpenNeuro (Cha et al., 2025) and were collected under the originating institution's approved protocol. The rat LFP–fMRI data were acquired under a protocol approved by the originating institution's animal care and use committee (Pan et al., 2011; Zhang et al., 2020). No new human or animal data were collected for this study, and all analyses were performed on de-identified data. The proposed estimator is a research tool for scientific analysis.

REPRODUCIBILITY STATEMENT

Implementation and training details are provided in Appendix C.1, simulation design and evaluation procedures in Appendix E, and empirical-data preprocessing and analysis in Appendix D. Theoretical derivations and computational benchmarks are provided in Appendices B and F, respectively. Code will be made publicly available upon paper acceptance.

REFERENCES

Fahimeh Arab, AmirEmad Ghassami, Hamidreza Jamalabadi, Megan AK Peters, and Erfan Nozari. Whole-brain causal discovery using fmri. *Network Neuroscience*, 9(1):392–420, 2025.

Deanna M Barch, Gregory C Burgess, Michael P Harms, Steven E Petersen, Bradley L Schlaggar, Maurizio Corbetta, Matthew F Glasser, Sandra Curtiss, Sachin Dixit, Cindy Feldt, et al. Function in the human connectome: task-fmri and individual differences in behavior. *Neuroimage*, 80: 169–189, 2013.

Lionel Barnett and Anil K. Seth. The MVGC multivariate granger causality toolbox: A new approach to granger-causal inference. *Journal of Neuroscience Methods*, 223:50–68, 2014. doi: 10.1016/j.jneumeth.2013.10.018.

Jeremie Barral, Xiao-Jing Wang, and Alex D Reyes. Propagation of temporal and rate signals in cultured multilayer networks. *Nature communications*, 10(1):3969, 2019.

Mohamed Ishmael Belghazi, Aristide Baratin, Sai Rajeshwar, Sherjil Ozair, Yoshua Bengio, Aaron Courville, and R. Devon Hjelm. Mutual information neural estimation. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 531–540, 2018.

Adriene M Beltz and Kathleen M Gates. Network mapping with gimme. *Multivariate behavioral research*, 52(6):789–804, 2017.

Natalia Z Bielczyk, Sebo Uithol, Tim van Mourik, Paul Anderson, Jeffrey C Glennon, and Jan K Buitelaar. Disentangling causal webs in the brain using functional magnetic resonance imaging: A review of current approaches. *Network Neuroscience*, 3(2):237–273, 2019.

Steven L. Bressler and Anil K. Seth. Wiener–granger causality: A well established methodology. *NeuroImage*, 58(2):323–329, 2011. doi: 10.1016/j.neuroimage.2010.02.059.

Younghwa Cha, Yeji Lee, Eunhee Ji, SoHyun Han, Sunhyun Min, Hyoungkyu Kim, Minseo Cho, Haesung Lee, Youngjai Park, and Joon-Young Moon. "sustained attention task (gradcpt) dataset using simultaneous eeg-fmri and dti", 2025.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ, 2 edition, 2006. ISBN 978-0-471-24195-9.

Gopikrishna Deshpande and Xiaoping Hu. Investigating effective brain connectivity from fmri data: past findings and current issues with reference to granger causality analysis. *Brain connectivity*, 2 (5):235–245, 2012.

Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803*, 2016.

Karl J Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

Karl J Friston, Lee Harrison, and Will Penny. Dynamic causal modelling. *Neuroimage*, 19(4): 1273–1302, 2003. doi: 10.1016/S1053-8119(03)00202-7.

Robert G Gallager. *Information theory and reliable communication*, volume 588. Springer, 1968.

Kathleen M Gates and Peter CM Molenaar. Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, 63(1):310–319, 2012.

Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, et al. Meg and eeg data analysis with mne-python. *Frontiers in Neuroscience*, 7:267, 2013.

Clive W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969. doi: 10.2307/1912791.

Hákon Gudbjartsson and Samuel Patz. The rician distribution of noisy mri data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995.

Daniel A Handwerker, John M Ollinger, and Mark D'Esposito. Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*, 21(4):1639–1651, 2004.

Clifford M Hurvich and Chih-Ling Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

Aapo Hyvärinen and Stephen M Smith. Pairwise likelihood ratios for estimation of non-gaussian structural equation models. *The Journal of Machine Learning Research*, 14(1):111–152, 2013.

Aapo Hyvärinen, Kun Zhang, Shohei Shimizu, and Patrik O Hoyer. Estimation of a structural vector autoregression model using non-gaussianity. *Journal of Machine Learning Research*, 11(5), 2010.

Shiro Ikeda and Jonathan H Manton. Capacity of a single spiking neuron channel. *Neural Computation*, 21(6):1714–1748, 2009.

Jianan Jian, Benjamin Li, Nurahmed Multezem, Francesca Mandino, Evelyn MR Lake, and Nan Xu. Channel capacity for time-resolved effective connectivity in functional neuroimaging. *bioRxiv*, pp. 2026–03, 2026.

Leonenko Kozachenko. Sample estimate of the entropy of a random vector. *Probl. Pered. Inform.*, 23:9, 1987.

Evelyn MR Lake, Xinxin Ge, Xilin Shen, Peter Herman, Fahmeed Hyder, Jessica A Cardin, Michael J Higley, Dustin Scheinost, Xenophon Papademetris, Michael C Crair, et al. Simultaneous cortex-wide fluorescence ca2+ imaging and whole-brain fmri. *Nature methods*, 17(12):1262–1271, 2020.

Helmut Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, Heidelberg, 2005. ISBN 978-3-540-26239-8. doi: 10.1007/978-3-540-27752-1.

Donald M MacKay and Warren S McCulloch. The limiting information capacity of a neuronal link. *The bulletin of mathematical biophysics*, 14(2):127–135, 1952.

Francesca Mandino, Corey Horien, Xilin Shen, Gabriel Desrosiers-Grégoire, Wendy Luo, Marija Markicevic, R Todd Constable, Xenophon Papademetris, Mallar M Chakravarty, Richard F Betzel, et al. Multimodal identification of the mouse brain using simultaneous ca2+ imaging and fmri. *Communications biology*, 8(1):665, 2025.

Mario Manto, James M Bower, Adriana Bastos Conforto, José M Delgado-García, Suzete Nascimento Farias Da Guarda, Marcus Gerwig, Christophe Habas, Nobuhiro Hagura, Richard B Ivry, Peter Mariën, et al. Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. *The Cerebellum*, 11(2):457–487, 2012. doi: 10.1007/S12311-011-0331-9.

Nikola T Markov, Maria M Ercsey-Ravasz, AR Ribeiro Gomes, Camille Lamy, Loic Magrou, Julien Vezoli, Pierre Misery, Arnaud Falchier, Rene Quilodran, Marie-Alice Gariel, et al. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral cortex*, 24(1):17–36, 2014.

James L. Massey. Causality, feedback and directed information. *Proceedings of the 1990 International Symposium on Information Theory and Its Applications*, pp. 303–305, 1990.

Cooper J Mellema and Albert A Montillo. Novel machine learning approaches for improving the reproducibility and reliability of functional and effective connectivity from functional mri. *Journal of neural engineering*, 20(6):066023, 2023.

Anish Mitra, Andrew Kraft, Patrick Wright, Benjamin Acland, Abraham Z Snyder, Zachary Rosenthal, Leah Czerniewski, Adam Bauer, Lawrence Snyder, Joseph Culver, et al. Spontaneous infra-slow brain activity has unique spatiotemporal dynamics and laminar structure. *Neuron*, 98(2):297–305, 2018.

Laura Murray, Blaise B Frederick, and Amy C Janes. Data-driven connectivity profiles relate to smoking cessation outcomes. *Neuropsychopharmacology*, 49(6):1007–1013, 2024.

Jae Kwang Noh and Victor Solo. Rician distributed fmri: asymptotic power analysis and cramér–rao lower bounds. *IEEE Transactions on Signal Processing*, 59(3):1322–1328, 2011.

Wen-Ju Pan, Garth Thompson, Matthew Magnuson, Waqas Majeed, Dieter Jaeger, and Shella Keilholz. Broadband local field potentials correlate with spontaneous fluctuations in functional magnetic resonance imaging signals in the rat somatosensory cortex under isoflurane anesthesia. *Brain connectivity*, 1(2):119–131, 2011.

Roberto D Pascual-Marqui. Discrete, 3d distributed, linear imaging methods of electric neuronal activity. part 1: exact, zero error localization. *arXiv preprint arXiv:0710.3341*, 2007.

Ben Poole, Sherjil Ozair, Aaron van den Oord, Alexander A. Alemi, and George Tucker. On variational bounds of mutual information. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 5171–5180, 2019.

Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows. *Proceedings of the 32nd International Conference on Machine Learning*, 37:1530–1538, 2015.

Giacomo Rizzolatti and Giuseppe Luppino. The cortical motor system. *Neuron*, 31(6):889–901, 2001.

Alard Roebroeck, Elia Formisano, and Rainer Goebel. Mapping directed influence over the brain using granger causality and fmri. *Neuroimage*, 25(1):230–242, 2005.

Ruben Sanchez-Romero, Joseph D Ramsey, Kun Zhang, Madelyn RK Glymour, Biwei Huang, and Clark Glymour. Estimating feedforward and feedback effective connections from fmri time series: Assessments of statistical methods. *Network Neuroscience*, 3(2):274–306, 2019.

Ayodele Sasegbon and Shaheen Hamdy. The role of the cerebellum in swallowing. *Dysphagia*, 38(2): 497–509, 2023. doi: 10.1007/s00455-021-10271-x.

Alexander Schaefer, Ru Kong, Evan M Gordon, Timothy O Laumann, Xi-Nian Zuo, Avram J Holmes, Simon B Eickhoff, and BT Thomas Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral cortex*, 28(9):3095–3114, 2018.

Thomas Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000. doi: 10.1103/PhysRevLett.85.461.

Anil K Seth, Adam B Barrett, and Lionel Barnett. Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297, 2015. doi: 10.1523/JNEUROSCI.4399-14.2015.

Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

Woodrow L Shew, Hongdian Yang, Shan Yu, Rajarshi Roy, and Dietmar Plenz. Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *The Journal of neuroscience*, 31(1):55–63, 2011.

Shohei Shimizu, Patrik O. Hoyer, Aapo Hyvärinen, and Antti Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(10), 2006.

Jason F Smith, Ajay Pillai, Kewei Chen, and Barry Horwitz. Effective connectivity modeling for fmri: six issues and possible solutions using linear dynamic systems. *Frontiers in systems neuroscience*, 5:104, 2012.

Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.

Klaas E. Stephan, Will D. Penny, Rosalyn J. Moran, Hanneke E. M. den Ouden, Jean Daunizeau, and Karl J. Friston. Ten simple rules for dynamic causal modeling. *NeuroImage*, 49(4):3099–3109, 2010. doi: 10.1016/j.neuroimage.2009.11.015.

Garth John Thompson, Michael Donelyn Merritt, Wen-Ju Pan, Matthew Evan Magnuson, Joshua Koehler Grooms, Dieter Jaeger, and Shella Dawn Keilholz. Neural correlates of time-varying functional connectivity in the rat. *Neuroimage*, 83:826–836, 2013.

Emanuel Todorov. Optimality principles in sensorimotor control. *Nature neuroscience*, 7(9):907–915, 2004.

Hadi Vafaii, Francesca Mandino, Gabriel Desrosiers-Grégoire, David O'Connor, Marija Markicevic, Xilin Shen, Xinxin Ge, Peter Herman, Fahmeed Hyder, Xenophon Papademetris, et al. Multimodal measures of spontaneous brain activity reveal both common and divergent patterns of cortical functional organization. *Nature Communications*, 15(1):229, 2024.

Pedro A Valdes-Sosa, Alard Roebroeck, Jean Daunizeau, and Karl Friston. Effective connectivity: influence, causality and biophysical modeling. *Neuroimage*, 58(2):339–361, 2011.

Bruno Wegmann, Anders Eklund, and Mattias Villani. Bayesian rician regression for neuroimaging. *Frontiers in Neuroscience*, 11:586, 2017.

Mingrui Xia, Jinhui Wang, and Yong He. Brainnet viewer: a network visualization tool for human brain connectomics. *PloS one*, 8(7):e68910, 2013.

Nan Xu, R Nathan Spreng, and Peter C Doerschuk. Initial validation for the estimation of resting-state fmri effective connectivity by a generalization of the correlation approach. *Frontiers in neuroscience*, 11:271, 2017.

Xiaodi Zhang, Wen-Ju Pan, and Shella Dawn Keilholz. The relationship between bold and neural activity arises from temporally sparse events. *Neuroimage*, 207:116390, 2020.

Yao Zhu, Tianming Huang, Ruolin Li, Qianrong Yang, Chaoyue Zhao, Ming Yang, Bin Lin, DIRECT Consortium, and Xuzhou Li. Distinct resting-state effective connectivity of large-scale networks in first-episode and recurrent major depression disorder: evidence from the rest-meta-mdd consortium. *Frontiers in Neuroscience*, 17:1308551, 2023.

APPENDIX CONTENTS

A Background and Related Work **15**

A.1 Channel Capacity for Effective Connectivity: Motivation and Interpretation 15

A.2 Baseline EC Methods 15

B Proofs and Justifications for Theoretical Properties **16**

B.1 Additive-Channel Capacity Identity 17

B.2 Proof of Lemma 1: Observer-Flow Entropy Identity 17

B.3 Gaussian Capacity as a Special Case 17

B.4 Entropy-Power Bounds for Matched-Variance Non-Gaussian Residuals 18

B.5 Game-Gap Diagnostic 19

B.6 Detailed Error Decomposition 20

C Algorithm, Training Sensitivity, and Controlled Validation of the Dual-Flow Estimator **22**

C.1 Algorithm and Implementation Details 22

C.2 Sensitivity to Training Sequence Length 23

C.3 Controlled Validation of Estimator Properties 23

D Implementation Details/Empirical Data Analysis **24**

D.1 Brain Datasets and Gaussianity Assessment 24

D.2 Task-Versus-Rest Sensitivity Analysis in Tongue-Motion fMRI 26

D.3 Cross-Modal Correspondence of Time-Varying EC Estimates 27

E Ground-Truth Effective-Connectivity Validation in Brain-Like Simulations **28**

E.1 Simulation Design 28

E.2 Evaluation Procedure and Performance 29

F Computational Requirements **31**

A Background and Related Work

A.1 Channel Capacity for Effective Connectivity: Motivation and Interpretation

Effective connectivity (EC) is inherently model-based: different approaches define directed interaction through different models, assumptions, and statistical objectives (Roebroeck et al., 2005; Valdes-Sosa et al., 2011; Deshpande & Hu, 2012; Bielczyk et al., 2019). For example, GC measures directed predictability, whereas structural equation modeling and LiNGAM-based methods infer directed relationships under different assumptions about network structure and residual distributions. Accordingly, their outputs are not directly interchangeable measures of connection strength. Smith et al. (2011) compared such methods by their ability to recover known edges and directions from simulated fMRI. Following this practice, we introduce channel capacity not as a universal measure of EC strength, but as a distinct information-theoretic measure evaluated through directed-edge recovery.

Brain-signal and residual distributions are not nuisance details; they directly shape the information-carrying limit of a fitted directional interaction. Our framework incorporates this distributional structure directly into a channel-capacity EC measure. For each ordered ROI pair $X \to Y$, we fit a directional finite impulse response (FIR) model and interpret its fitted dynamics and residuals as a noisy communication channel. Under an input-power constraint, capacity quantifies the maximum information rate supported by the fitted channel, reflecting its directional dynamics and residual distribution while optimizing over admissible input distributions. GCap evaluates this quantity after replacing the fitted residuals with matched-variance Gaussian noise, whereas Dual-flow approximates the empirical residual distribution through residual resampling and optimizes the input distribution. Their comparison therefore assesses the contribution of distribution-aware capacity estimation while holding the fitted FIR channel fixed.

Like other pairwise measures, channel capacity is conditional on the fitted model and observed variables. For example, if X drives both Y and Z, an analysis of Y and Z may detect directional dependence induced by the omitted common driver X, despite the absence of a direct $Y \leftrightarrow Z$ connection. Indirect pathways, feedback, and heterogeneous hemodynamic responses may similarly affect the fitted channel. We evaluate robustness to these factors using simulations with prespecified ground-truth connectivity (Section 4.2; Appendix E).

Channel capacity has previously been used to characterize information limits in single neurons and neuronal circuits (MacKay & McCulloch, 1952; Ikeda & Manton, 2009; Shew et al., 2011; Barral et al., 2019). Here, our novelty is to formulate the capacity of a fitted directional interaction between brain regions as a model-based EC measure under general residual distributions. This measure characterizes distribution-dependent communication potential beyond coupling magnitude alone. In this study, the channel is instantiated by a pairwise FIR–residual model within short windows; extension to multivariate autoregressive or other dynamical channel models is a future direction. Because the resulting capacity is conditional on the selected model and observed variables, it should not be interpreted as proof of a direct anatomical connection, established neural causation, or the amount of information actually transmitted in vivo. Before this metric can be translated into a clinically actionable biomarker, rigorous disease-specific validation in independent cohorts will be needed.

A.2 Baseline EC Methods

We selected four representative EC baselines spanning complementary frameworks that remain in current neuroscience and neuroimaging use, supporting their inclusion as contemporary baselines (Zhu et al., 2023; Mellema & Montillo, 2023; Arab et al., 2025; Murray et al., 2024). These methods can be applied to the same ROI timeseries and evaluated against the same directed ground truth without auxiliary experimental inputs or anatomical priors. Table 1 summarizes their main differences.

Gaussian Capacity (GCap). GCap serves as the within-framework ablation for Dual-flow. Both methods use the same fitted directional FIR channel, input-power constraint, and edge-selection rule. GCap replaces the empirical residual distribution with a matched-variance Gaussian model and computes the corresponding Gaussian-channel capacity, whereas Dual-flow uses empirical residual

Table 1: Distributional structure and EC estimation principles of the proposed method and evaluated baselines.

Method	Distributional structure	EC estimation principle
Dual-flow	Residual distribution approximated by empirical samples.	Incorporates empirical residual shape into channel-capacity estimation while optimizing the input distribution.
GCap	Gaussian residuals; variance estimated from empirical data.	Computes channel capacity using a matched-variance Gaussian residual model.
GC	Gaussian residuals for conventional parametric inference.	Compares residual variances of reduced and full autoregressive models to quantify directed predictability.
VAR-LiNGAM	Independent non-Gaussian structural errors; distributional form unspecified.	Uses non-Gaussianity and independence to identify directed connections.
GIMME	Gaussian signal model under conventional SEM estimation.	Uses signal covariance structure to select group- and individual-level directed paths.

samples and learns an admissible input distribution. Their comparison therefore tests whether residual distributional shape beyond variance affects EC estimates.

Granger Causality (GC). Several GC variants for EC estimation were evaluated by Smith et al. (2011). Here, we implement the pairwise Granger Bn approach, the best-performing GC variant in that benchmark, to provide a direct comparison with our pairwise channel model. This approach uses BIC to select the model order up to a maximum lag of n. Throughout this paper, GC refers to this pairwise implementation. For each ordered ROI pair $X \to Y$, GC compares a reduced autoregressive model that predicts Y from its own past with a full model that additionally includes the past of X. The reduction in residual variance quantifies the additional predictive contribution of X (Granger, 1969; Barnett & Seth, 2014). The conventional parametric F-test assesses this improvement under Gaussian-error assumptions (Seth et al., 2015).

VAR-LiNGAM. LiNGAM-family methods identify linear causal structure under assumptions of independent non-Gaussian structural errors and acyclicity (Shimizu et al., 2006). VAR-LiNGAM extends this framework to timeseries data, estimating contemporaneous and lagged directed effects while retaining an acyclic contemporaneous structure (Hyvärinen et al., 2010). The specific forms of the structural error distributions are unspecified; their non-Gaussianity and independence enable identification of the directed structure and connection coefficients. We include VAR-LiNGAM as a representative non-Gaussian EC baseline. Our framework uses distributional information differently: it incorporates the empirical residual distribution directly into a channel-capacity EC measure while optimizing the admissible input distribution under a power constraint.

GIMME. Group Iterative Multiple Model Estimation (GIMME) is a network-level method based on unified structural equation modeling (Gates & Molenaar, 2012; Sanchez-Romero et al., 2019). It identifies contemporaneous and lagged connections supported across participants and then adds participant-specific paths. Thus, GIMME estimates group- and individual-level networks, whereas Dual-flow assigns a continuous capacity score to each ordered ROI pair. Published guidance indicates that GIMME generally works well with 3–20 ROIs, with the required time-series length depending on network size and model complexity (Beltz & Gates, 2017). The software documentation specifies at least 30 time points per evaluated timeseries, with 60 or more recommended.[1]

B Proofs and Justifications for Theoretical Properties

This appendix gives the formal derivations supporting the identities and theoretical properties used in the main text. Throughout, differential entropies are assumed to exist and be finite, the additive noise is independent of admissible inputs, and all distributions are absolutely continuous with respect to Lebesgue measure unless otherwise stated. The fitted residual pool is treated as a finite sample from

[1] https://github.com/GatesLab/gimme

an unknown continuous residual-noise distribution; empirical resampling is used for Monte Carlo approximation of expectations, not as the definition of the continuous entropy itself.

B.1 Additive-Channel Capacity Identity

Consider the additive FIR vector channel

$$Y = BX + W, \tag{16}$$

where B is fixed within a temporal window or condition segment and X is independent of W. For any admissible input distribution $P_X \in \mathcal{P}_P$, the mutual information satisfies

$$I(X;Y) = h(Y) - h(Y|X). \tag{17}$$

Conditioning on $X = x$ gives

$$Y|X = x = Bx + W. \tag{18}$$

Differential entropy is invariant to deterministic translations, so

$$h(Y|X = x) = h(Bx + W) = h(W). \tag{19}$$

Averaging over X yields

$$h(Y|X) = \mathbb{E}_X[h(Y|X = x)] = h(W). \tag{20}$$

Therefore,

$$I(X;Y) = h(BX + W) - h(W). \tag{21}$$

Taking the supremum over all admissible input distributions gives Eq. 4.

B.2 Proof of Lemma 1: Observer-Flow Entropy Identity

Let $q_\phi : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible differentiable flow with Jacobian $J_{q_\phi}(y)$, and let $U \sim \mathcal{N}(0, I)$ have density p_U. The density induced on Y by the observer flow is

$$p_\phi(y) = p_U(q_\phi(y)) \left|\det J_{q_\phi}(y)\right|. \tag{22}$$

Let P_Y denote the true output distribution with density p_Y. The expected observer negative log-likelihood is the cross-entropy from P_Y to P_ϕ:

$$\mathbb{E}_{Y\sim P_Y}[-\log p_\phi(Y)] = \int p_Y(y) \log \frac{1}{p_\phi(y)}\, dy. \tag{23}$$

Add and subtract $\log p_Y(y)$:

$$\mathbb{E}_{P_Y}[-\log p_\phi(Y)] = \int p_Y(y) \log \frac{1}{p_Y(y)}\, dy + \int p_Y(y) \log \frac{p_Y(y)}{p_\phi(y)}\, dy \tag{24}$$

$$= h(Y) + \mathrm{KL}(P_Y \| P_\phi). \tag{25}$$

Thus, if the observer family contains the true density or can approximate it arbitrarily well in KL divergence, the minimized observer loss equals $h(Y)$.

B.3 Gaussian Capacity as a Special Case

Assume $W \sim \mathcal{N}(0, \Sigma_W)$ with $\Sigma_W \succ 0$, and let $K_X = \mathrm{Cov}(X)$ for an arbitrary admissible input X satisfying $\mathrm{tr}(K_X) \le dP$. Then

$$\mathrm{Cov}(Y) = BK_XB^\top + \Sigma_W. \tag{26}$$

Among all distributions with a fixed covariance, the Gaussian distribution maximizes differential entropy. Therefore,

$$h(Y) \le \frac{1}{2} \log\left((2\pi e)^d \det(BK_XB^\top + \Sigma_W)\right). \tag{27}$$

Since W is Gaussian,

$$h(W) = \frac{1}{2} \log\left((2\pi e)^d \det \Sigma_W\right). \tag{28}$$

Using the additive-channel identity,

$$I(X;Y) \le \frac{1}{2} \log\det\left(I + \Sigma_W^{-1/2} BK_XB^\top \Sigma_W^{-1/2}\right). \tag{29}$$

The upper bound is achieved by Gaussian X with covariance Σ_X solving the constrained optimization problem in Eq. 13, yielding the classical Gaussian-channel capacity.

B.4 Entropy-Power Bounds for Matched-Variance Non-Gaussian Residuals

For the scalar additive channel $Y = X + W$, choose $X \sim \mathcal{N}(0, P)$ independent of W. The entropy power inequality gives

$$N(X + W) \geq N(X) + N(W) = P + N(W). \tag{30}$$

Therefore,

$$I(X;Y) = h(X + W) - h(W) \tag{31}$$

$$= \frac{1}{2} \log \frac{N(X + W)}{N(W)} \tag{32}$$

$$\geq \frac{1}{2} \log \left(1 + \frac{P}{N(W)} \right). \tag{33}$$

Because capacity is the supremum over admissible input distributions, the same quantity provides an achievable-rate lower bound on $C(W)$. Moreover, Gaussian distributions maximize entropy at fixed variance, so $N(W) \leq \sigma_W^2$, with equality only when W is Gaussian. This yields Eq. 14 and shows that matched residual variance does not, in general, imply matched achievable information rate.

Extension to the FIR channel. To obtain the bounds used in the controlled matched-variance experiment, we consider the finite-dimensional additive FIR channel in its vector form

$$Y = BX + W, \tag{34}$$

where $B \in \mathbb{R}^{d \times d}$ is the finite-dimensional FIR convolution operator induced by the fixed filter b, the admissible input satisfies

$$\mathbb{E}\|X\|_2^2 \leq dP, \tag{35}$$

and the residual noise W has matched per-sample variance σ_W^2. In the Gaussian-surrogate model, the residual covariance is taken to be $\Sigma_W = \sigma_W^2 I_d$.

Gaussian surrogate capacity. By the Gaussian special case in Appendix B.3, replacing the true residual law by a Gaussian distribution with the same variance gives

$$C_{\mathrm{G}}^{\mathrm{FIR}} = \frac{1}{2d} \max_{\Sigma_X \succeq 0,\ \mathrm{tr}(\Sigma_X) \leq dP} \log \det \left(I_d + \frac{1}{\sigma_W^2} B \Sigma_X B^\top \right). \tag{36}$$

Equivalently, if λ_i are the squared singular values of B and $p_i^\star$ are the corresponding water-filling allocations under the power constraint $\sum_i p_i^\star \leq dP$, then

$$C_{\mathrm{G}}^{\mathrm{FIR}} = \frac{1}{2d} \sum_i \log \left(1 + \frac{p_i^\star \lambda_i}{\sigma_W^2} \right). \tag{37}$$

Because B, P, and σ_W^2 are fixed across matched-variance residual conditions, $C_{\mathrm{G}}^{\mathrm{FIR}}$ is fixed across residual distributions.

Lower bound. Let $\Sigma_X^\star$ denote the Gaussian water-filling input covariance attaining Eq. 36. A lower bound for the same FIR model follows from the vector entropy power inequality applied to the admissible Gaussian input $X_G \sim \mathcal{N}(0, \Sigma_X^\star)$:

$$C^{\mathrm{FIR}}(W) \geq \frac{1}{2} \log \left(1 + \frac{\det(B \Sigma_X^\star B^\top)^{1/d}}{N(W)} \right), \tag{38}$$

where the vector entropy power of W is

$$N(W) = \frac{1}{2\pi e} \exp \left(\frac{2}{d} h(W) \right). \tag{39}$$

In practice, $B \Sigma_X^\star B^\top$ may be rank-deficient because only a subset of singular modes receive nonzero water-filling power. Let

$$\mathcal{I}_\star = \{ i : \lambda_i > 0,\ p_i^\star > 0 \}, \qquad r = |\mathcal{I}_\star|. \tag{40}$$

Restricting the entropy-power argument to the active output subspace yields the implementable lower bound used in the simulation:

$$C_{\mathrm{LB}}^{\mathrm{FIR}}(W) = \frac{r}{2d} \log\left(1 + \frac{\left(\prod_{i\in\mathcal{I}_\star} \lambda_i p_i^\star\right)^{1/r}}{N(W)}\right). \tag{41}$$

For i.i.d. residuals with per-sample differential entropy $h(w)$, Eq. 39 reduces to

$$N(W) = \frac{1}{2\pi e} \exp(2h(w)). \tag{42}$$

Upper bound. An upper bound for the same FIR model follows from the fact that, for fixed covariance, the Gaussian distribution maximizes differential entropy. Let h_G denote the per-sample entropy of a matched Gaussian residual with variance σ_W^2:

$$h_G = \frac{1}{2} \log(2\pi e\, \sigma_W^2). \tag{43}$$

Then

$$C^{\mathrm{FIR}}(W) \le C_{\mathrm{G}}^{\mathrm{FIR}} + h_G - h(w), \tag{44}$$

where $h(w)$ is the per-sample differential entropy of the true residual law. Equivalently, in vector form,

$$C^{\mathrm{FIR}}(W) \le C_{\mathrm{G}}^{\mathrm{FIR}} + \frac{1}{d}\left[\frac{1}{2}\log\big((2\pi e)^d \det(\sigma_W^2 I_d)\big) - h(W)\right]. \tag{45}$$

Implications for matched-variance residuals. For Gaussian residuals, $N(W) = \sigma_W^2$ and $h(w) = h_G$, so the upper bound in Eq. 44 collapses exactly to the Gaussian-surrogate capacity:

$$C_{\mathrm{UB}}^{\mathrm{FIR}}(W_{\mathrm{Gaussian}}) = C_{\mathrm{G}}^{\mathrm{FIR}}. \tag{46}$$

By contrast, the lower bound in Eq. 41 is generally not tight even in the Gaussian case, so typically

$$C_{\mathrm{LB}}^{\mathrm{FIR}}(W_{\mathrm{Gaussian}}) \le C_{\mathrm{G}}^{\mathrm{FIR}} = C_{\mathrm{UB}}^{\mathrm{FIR}}(W_{\mathrm{Gaussian}}). \tag{47}$$

For matched-variance non-Gaussian residuals, $C_{\mathrm{G}}^{\mathrm{FIR}}$ remains fixed, whereas the lower and upper bounds depend on residual entropy through $N(W)$ and $h(w)$. Thus residual distributional structure can alter the range of achievable information rates even when the FIR channel, input-power constraint, and residual variance are held fixed.

B.5 Game-Gap Diagnostic

Although the main text does not rely on a theory-heavy saddle-point claim, the dual-flow objective can be monitored with an empirical game-gap diagnostic.

Proposition 1 (Game gap for the dual-flow objective)**.** *Let*

$$V_{\Theta,\Phi} = \sup_{\theta\in\Theta} \inf_{\phi_Y\in\Phi} \mathcal{L}_Y(\theta, \phi_Y) \tag{48}$$

denote the signal-plus-noise value of the flow-restricted game. For a trained pair $(\widehat{\theta}, \widehat{\phi}_Y)$, define the generator and observer regrets

$$r_{\mathrm{gen}} = \sup_{\theta\in\Theta} \mathcal{L}_Y(\theta, \widehat{\phi}_Y) - \mathcal{L}_Y(\widehat{\theta}, \widehat{\phi}_Y), \tag{49}$$

$$r_{\mathrm{obs}} = \mathcal{L}_Y(\widehat{\theta}, \widehat{\phi}_Y) - \inf_{\phi_Y\in\Phi} \mathcal{L}_Y(\widehat{\theta}, \phi_Y). \tag{50}$$

Define the game gap as

$$\mathrm{Gap}_Y = r_{\mathrm{gen}} + r_{\mathrm{obs}}. \tag{51}$$

Then

$$\left|\mathcal{L}_Y(\widehat{\theta}, \widehat{\phi}_Y) - V_{\Theta,\Phi}\right| \le \max\{r_{\mathrm{gen}}, r_{\mathrm{obs}}\} \le \mathrm{Gap}_Y. \tag{52}$$

Thus a small game gap indicates that neither player can substantially improve the objective by unilateral re-optimization, and bounds the remaining optimization error relative to the flow-restricted saddle-point value.

Proof. Let

$$V_{\Theta,\Phi} = \sup_{\theta\in\Theta} \inf_{\phi\in\Phi} \mathcal{L}_Y(\theta,\phi) \tag{53}$$

be the max–min value of the signal-plus-noise flow game. For any fixed observer $\widehat{\phi}_Y$, weak duality gives

$$V_{\Theta,\Phi} \le \sup_{\theta\in\Theta} \mathcal{L}_Y(\theta,\widehat{\phi}_Y). \tag{54}$$

By the definition of the generator regret,

$$\sup_{\theta\in\Theta} \mathcal{L}_Y(\theta,\widehat{\phi}_Y) = \mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) + r_{\mathrm{gen}}. \tag{55}$$

Therefore

$$V_{\Theta,\Phi} \le \mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) + r_{\mathrm{gen}}. \tag{56}$$

Similarly, for any fixed generator $\widehat{\theta}$,

$$\inf_{\phi\in\Phi} \mathcal{L}_Y(\widehat{\theta},\phi) \le V_{\Theta,\Phi}. \tag{57}$$

By the definition of the observer regret,

$$\inf_{\phi\in\Phi} \mathcal{L}_Y(\widehat{\theta},\phi) = \mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) - r_{\mathrm{obs}}. \tag{58}$$

Thus

$$\mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) - r_{\mathrm{obs}} \le V_{\Theta,\Phi}. \tag{59}$$

Combining Eqs. 56 and 59 gives

$$\mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) - r_{\mathrm{obs}} \le V_{\Theta,\Phi} \le \mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) + r_{\mathrm{gen}}. \tag{60}$$

Therefore

$$\left|\mathcal{L}_Y(\widehat{\theta},\widehat{\phi}_Y) - V_{\Theta,\Phi}\right| \le \max\{r_{\mathrm{gen}}, r_{\mathrm{obs}}\} \le r_{\mathrm{gen}} + r_{\mathrm{obs}}. \tag{61}$$

This proves Eq. 52. The same reasoning applies to the noise-only observer objective after removing the generator player, where the relevant optimization error is the suboptimality gap

$$\mathcal{L}_W(\widehat{\phi}_W) - \inf_{\phi_W\in\Phi} \mathcal{L}_W(\phi_W). \tag{62}$$

In practice, this diagnostic can be estimated by restarting or continuing optimization of one player while holding the other fixed.

B.6 Detailed Error Decomposition

The error decomposition in Eq. 15 separates channel and residual estimation, generator restriction, observer approximation, optimization, and Monte Carlo error. It is an accounting of approximation sources rather than a finite-sample guarantee. Throughout this section, differential entropy is defined for the underlying continuous distributions. Empirical residual resampling is used to approximate expectations and is not assigned a differential entropy.

Let C denote the population capacity of the additive FIR channel with operator B and continuous residual-noise law P_W. For $X_\theta = \mathrm{Norm}_P(g_\theta(Z_1))$, define the generator-restricted capacity

$$C_\Theta = \sup_{\theta\in\Theta} \{h(BX_\theta + W) - h(W)\}, \qquad W \sim P_W. \tag{63}$$

The generator-restriction error is

$$\varepsilon_{\mathrm{gen}} = C - C_\Theta. \tag{64}$$

This term measures the loss from restricting the admissible input distributions to the generator family.

Using the population losses $\mathcal{L}_Y$ and $\mathcal{L}_W$, define the flow-restricted entropy-difference objective

$$V_{\Theta,\Phi} = \sup_{\theta\in\Theta} \inf_{\phi_Y\in\Phi} \mathcal{L}_Y(\theta,\phi_Y) - \inf_{\phi_W\in\Phi} \mathcal{L}_W(\phi_W). \tag{65}$$

For $Y_\theta = BX_\theta + W$, Lemma 1 gives

$$\inf_{\phi_Y\in\Phi} \mathcal{L}_Y(\theta,\phi_Y) = h(Y_\theta) + a_\Phi(\theta), \tag{66}$$

$$a_\Phi(\theta) = \inf_{\phi_Y\in\Phi} \mathrm{KL}(P_{Y_\theta}\|P_{\phi_Y}), \tag{67}$$

$$\inf_{\phi_W\in\Phi} \mathcal{L}_W(\phi_W) = h(W) + b_\Phi, \tag{68}$$

$$b_\Phi = \inf_{\phi_W\in\Phi} \mathrm{KL}(P_W\|P_{\phi_W}). \tag{69}$$

Consequently, the observer-approximation error satisfies

$$\varepsilon_{\mathrm{obs}} := |V_{\Theta,\Phi} - C_\Theta| \le \sup_{\theta\in\Theta} a_\Phi(\theta) + b_\Phi. \tag{70}$$

Let $\widehat{B}$ be the fitted FIR operator and let $\widehat{P}_W$ denote the distribution induced by resampling the fitted residual pool. Conditional on the observed data, define the empirical-resampling losses

$$\widetilde{\mathcal{L}}_Y(\theta,\phi_Y) = \mathbb{E}_{Z_1,\widehat{W}}\left[-\log p_{\phi_Y}\left(\widehat{B}X_\theta + \widehat{W}\right)\right], \tag{71}$$

$$\widetilde{\mathcal{L}}_W(\phi_W) = \mathbb{E}_{\widehat{W}}\left[-\log p_{\phi_W}(\widehat{W})\right], \qquad \widehat{W}\sim\widehat{P}_W. \tag{72}$$

These are expected log-density losses evaluated on resampled data, not differential entropies of the empirical distribution. Define their optimized difference as

$$\widetilde{V}_{\Theta,\Phi} = \sup_{\theta\in\Theta} \inf_{\phi_Y\in\Phi} \widetilde{\mathcal{L}}_Y(\theta,\phi_Y) - \inf_{\phi_W\in\Phi} \widetilde{\mathcal{L}}_W(\phi_W). \tag{73}$$

The channel and residual approximation term is

$$\varepsilon_{\mathrm{resid}} = |\widetilde{V}_{\Theta,\Phi} - V_{\Theta,\Phi}|. \tag{74}$$

This term captures the effect of replacing the population channel and residual expectations with the fitted FIR operator and finite residual pool. Increasing the number of resampled training points does not, by itself, eliminate this discrepancy.

For trained parameters $(\widehat{\theta},\widehat{\phi}_Y,\widehat{\phi}_W)$, let

$$\widetilde{C}_{\mathrm{train}} = \widetilde{\mathcal{L}}_Y(\widehat{\theta},\widehat{\phi}_Y) - \widetilde{\mathcal{L}}_W(\widehat{\phi}_W). \tag{75}$$

Applying the argument of Proposition 1 to the empirical-resampling losses gives

$$|\widetilde{C}_{\mathrm{train}} - \widetilde{V}_{\Theta,\Phi}| \le \mathrm{Gap}_{\mathrm{game}}, \tag{76}$$

where $\mathrm{Gap}_{\mathrm{game}}$ is the sum of the generator regret, signal-plus-noise observer regret, and noise-only observer suboptimality for these losses.

Finally, define

$$\varepsilon_{\mathrm{MC}} = |\widehat{C}_{\Theta,\Phi} - \widetilde{C}_{\mathrm{train}}| \tag{77}$$

as the error from evaluating the trained empirical-resampling objectives using finite Monte Carlo samples. Adding and subtracting the intermediate quantities yields

$$|\widehat{C}_{\Theta,\Phi} - C| \le \varepsilon_{\mathrm{resid}} + \varepsilon_{\mathrm{gen}} + \varepsilon_{\mathrm{obs}} + \mathrm{Gap}_{\mathrm{game}} + \varepsilon_{\mathrm{MC}}. \tag{78}$$

This decomposition assumes that the displayed quantities are finite. It does not establish that the terms are small or provide convergence rates; it identifies the distinct sources of discrepancy between the finite-sample estimate and population capacity.

C Algorithm, Training Sensitivity, and Controlled Validation of the Dual-Flow Estimator

C.1 Algorithm and Implementation Details

Algorithm 1 Dual-flow empirical channel-capacity estimation

Require: Source signal $x_{1:T}$, target signal $y_{1:T}$, FIR order k, power constraint P, training-sequence length L, generator g_θ, observer flows q_{ϕ_Y} and q_{ϕ_W}

Fit FIR coefficients $\widehat{b}$ from x to y by least squares.
Compute residuals $\widehat{w}_t = y_t - \sum_{\ell=0}^{k-1} \widehat{b}_\ell x_{t-\ell}$ for $t = k, \dots, T$.
repeat
    Sample a latent sequence $Z_1 \sim \mathcal{N}(0, I_L)$ of length L.
    Sample a residual-noise sequence W of length L by uniformly resampling with replacement from $\{\widehat{w}_t\}_{t=k}^{T}$.
    Generate admissible inputs $X_\theta = \mathrm{Norm}_P(g_\theta(Z_1))$.
    Compute signal-plus-noise outputs $Y_\theta = BX_\theta + W$.
    Update observer q_{ϕ_Y} to minimize $-\log p_{\phi_Y}(Y_\theta)$.
    Update noise observer q_{ϕ_W} to minimize $-\log p_{\phi_W}(W)$.
    Update generator g_θ to maximize the signal-plus-noise objective $\mathcal{L}_Y(\theta, \phi_Y)$.
    Estimate capacity $\widehat{C} = \widehat{h}(Y_\theta) - \widehat{h}(W)$.
until the exponential moving average of $\widehat{C}$ reaches stationarity
return $\widehat{C}$

Flow architecture. Both the generator and observer are implemented using three pairs of 1D Real NVP coupling layers (Fig. 2b). Each coupling layer partitions the input into frozen and transformed components using a binary mask, and paired layers use complementary masks so that all dimensions are transformed. The transformed component is updated by scale and translation networks, each implemented with two one-dimensional convolutional layers and an intermediate ReLU nonlinearity. In each observer coupling layer, the sum of the scale-network outputs gives the log-determinant of the layer Jacobian; summing these terms across all observer layers yields $\log|\det J_{q_\phi}(y)|$ in Eq. 7. To stabilize training, scale outputs are softly clipped as $s \mapsto 2\tanh(s/2)$.

Training and optimization. At each training iteration, a latent sequence $Z_1 \sim \mathcal{N}(0, I_L)$ of length L and L residual-noise samples drawn by uniform resampling from the fitted residual pool are propagated through the generator, FIR channel, and observer. We alternate observer updates that minimize negative log-likelihood with generator updates that maximize the signal-plus-noise entropy estimate. Gradients are computed by automatic differentiation and parameters are optimized with *Adam*. Training stops when the exponential moving average of the paired entropy-difference estimate reaches stationarity. Sensitivity to the generated training-sequence length is reported in Appendix C.2.

Residual-resampling approximation. In empirical data, the true residual law P_W is unknown, so we approximate expectations under this law by uniformly resampling with replacement from the fitted residual pool. The quality of this approximation depends on the size and representativeness of the pool and the adequacy of the fitted FIR model. Under our indexing, an observed segment of length T contributes $T - k + 1$ fitted residuals, from which L residual samples are drawn with replacement at each training iteration. Thus, T determines the amount of empirical residual information available, whereas L determines the generated training-sequence length used for Monte Carlo estimation; increasing L does not add information beyond the observed residual pool. Independent resampling captures the empirical marginal residual distribution but does not preserve temporal dependence, and therefore relies on treating residuals as approximately independent draws from a stable within-segment distribution. Strong temporal dependence or within-segment changes in the residual distribution can compromise this approximation.

C.2 Sensitivity to Training Sequence Length

We evaluated sensitivity to the generated training-sequence length L, using $L = 1024$ as the default. We tested $L \in \{256, 512, 1024, 2048, 4096\}$ using 50 random initializations for each of four residual distributions (Gaussian, Uniform, Student's t, and Exponential), yielding 1,000 runs in total. Across all conditions, mean capacity estimates differed from their corresponding $L = 1024$ values by at most 4.30%, and all mean shifts were no larger than the corresponding across-initialization SD at $L = 1024$. All 1,000 runs completed without numerical failure and produced finite estimates, indicating limited sensitivity to L within the tested range (Table 2). The Real-NVP architecture with three coupling-layer pairs and the EMA-based convergence criterion were fixed across all experiments rather than tuned to individual datasets. A broader architecture-depth ablation was not performed and is acknowledged as a limitation.

Table 2: Capacity estimates across training sequence lengths. Values are mean ± SD across 50 random initializations. Max. Δ is the largest relative deviation from the $L = 1024$ value.

Distribution	L=256	L=512	L=1024	L=2048	L=4096	Max. Δ
Gaussian	.128±.006	.129±.008	.131±.007	.132±.011	.134±.011	2.25%
Uniform	.236±.011	.238±.010	.237±.012	.239±.012	.238±.011	1.11%
Student's t	.184±.011	.181±.010	.184±.006	.185±.013	.190±.017	3.23%
Exponential	.337±.030	.346±.034	.332±.038	.344±.029	.341±.037	4.30%

C.3 Controlled Validation of Estimator Properties

Entropy estimation and Gaussian-capacity recovery. We first evaluated the estimator in analytically tractable settings where the target entropy or channel capacity is known. For entropy validation, we set the FIR coefficients to zero so that the channel output consisted only of sampled noise. Noise samples were drawn from Gaussian, Uniform, Exponential, and Student's t distributions, with $\nu = 3$ for Student's t. For each distribution, we performed $N = 60$ independent training runs and compared the observer-flow entropy estimate with the ground-truth differential entropy. Across all four distributions, the observer accurately recovered the known entropy, with small estimation bias across independent runs (Fig. 6A-i). For Gaussian-capacity validation, we then evaluated the full generator–observer game in FIR channels with known analytic capacity. We tested four representative channels: a bandpass filter, a bandstop filter, a maximum-phase version of the bandpass filter obtained by time reversal, and a maximum-phase sparse multipath filter with coefficients $b = (0, 0, 0, 0, 0, 0, 0.1, 0, 0.3, 0, 0, 0, 0, 0, 0, 0.4)$. The bandpass and bandstop filters were constructed using second-order Butterworth filters with passband or stopband between 0.3 and 0.8 times the Nyquist frequency and truncated to the first 16 impulse-response coefficients. Across these representative channel structures, the full Dual-flow estimator recovered the corresponding analytic Gaussian FIR capacities with small estimation bias over 60 independent runs (Fig. 6A-ii). Together, Fig. 6A-i–ii validates observer-based entropy estimation and generator-based capacity maximization against known ground truth before application to empirical non-Gaussian residuals.

Matched-variance FIR validation. We next evaluated the theoretical prediction that residual distributions with identical variance can nevertheless support different achievable information rates because of distributional structure beyond second-order statistics. We considered additive FIR channels with fixed filter B, fixed input-power constraint P, and matched residual variance σ_W^2, while varying only the residual distribution; the corresponding Gaussian-surrogate capacity and entropy-dependent theoretical bounds are derived in Appendix B.4. Because B, P, and σ_W^2 are fixed, the Gaussian-surrogate capacity is identical across residual laws, whereas the entropy-power lower bound and entropy-based upper bound vary with residual entropy, as shown in Fig. 6B. Dual-flow estimates likewise varied across residual distributions and tracked these entropy-dependent differences rather than remaining fixed at the Gaussian-surrogate value (Fig. 6B). In particular, for several non-Gaussian residual laws, including Uniform, Student's t, and skewed-Exponential residuals, the Gaussian surrogate fell below the theoretical lower bound, demonstrating that a matched-variance Gaussian approximation can underestimate achievable information rates (Fig. 6B). In contrast, Dual-flow remained within the theoretical lower and upper bounds for all non-Gaussian

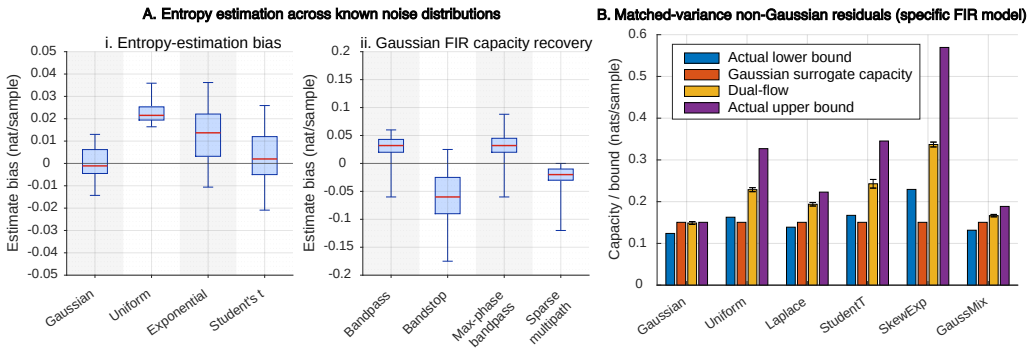


Figure 6: **Controlled validation of Dual-flow. A**, Validation in analytically tractable settings. **(i)** Entropy-estimation bias relative to ground truth across known noise distributions over 60 runs. **(ii)** Capacity-estimation bias relative to analytic Gaussian FIR capacity across representative channels over 60 runs. **B**, Matched-variance non-Gaussian residuals in a fixed FIR channel. Bars show the actual lower bound, Gaussian surrogate capacity, Dual-flow estimate, and actual upper bound for residual distributions with matched variance. Error bars denote mean $\pm$ s.e.m. where applicable.

residual cases and agreed with the Gaussian-surrogate capacity in the Gaussian case (Fig. 6B). These controlled matched-variance simulations therefore show that residual distribution shape alone can alter achievable information rates even when the channel filter, input power, and residual variance are fixed, and that Dual-flow captures these entropy-dependent differences beyond variance-based Gaussian capacity.

D Implementation Details/Empirical Data Analysis

D.1 Brain Datasets and Gaussianity Assessment

We evaluated Gaussianity in three datasets spanning human task-fMRI, human simultaneous EEG–fMRI, and rodent LFP–fMRI, chosen to sample distinct recording modalities, temporal scales, and experimental contexts. For the human task-fMRI and EEG–fMRI datasets, the analysis focused on observed signal distributions. For the rat LFP–fMRI dataset, we additionally examined fitted local channel residuals to test whether residual distributions can deviate from Gaussianity even when the corresponding BOLD signals appear approximately Gaussian at the whole-scan level. Across datasets, departures from Gaussianity were summarized primarily using Anderson–Darling (A–D) statistics relative to matched Gaussian references, with larger values indicating stronger deviations from Gaussianity.

Human HCP motor-task fMRI. To evaluate Gaussianity in task-evoked human fMRI, we used the motor-task dataset from the Human Connectome Project test–retest cohort (Barch et al., 2013). The minimally preprocessed 3T fMRI data ($TR = 0.72$ s) included 45 participants, each completing two motor-task runs. The task paradigm comprised visually cued movements of the left foot, right foot, left hand, right hand, and tongue, together with interleaved resting blocks (Barch et al., 2013). Task-specific activation maps were estimated using FSL FEAT following the HCP motor-task design and standard preprocessing choices commonly used for this dataset (Barch et al., 2013). ROI time series were extracted from validated task-activated regions in motor, somatosensory, and cerebellar areas, whose spatial patterns were consistent with canonical motor-task activation reported previously (Barch et al., 2013; Rizzolatti & Luppino, 2001; Manto et al., 2012). For the Gaussianity analysis, we evaluated ROI-level fMRI signals from the task conditions and matched rest blocks.

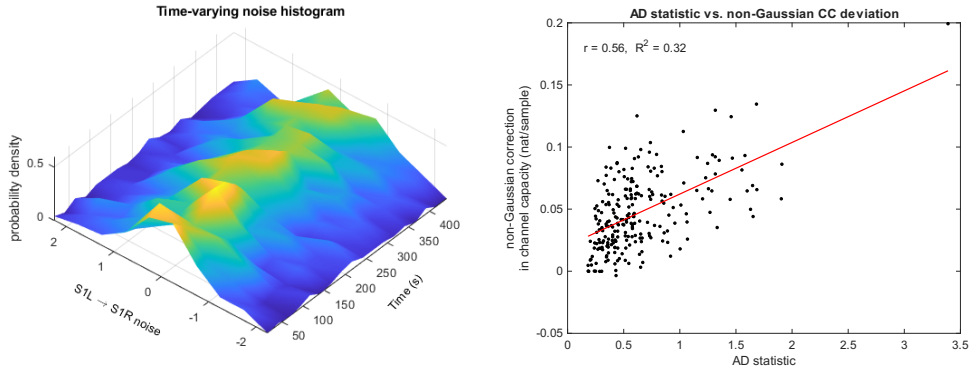
Human simultaneous EEG–fMRI. We further analyzed the eyes-open resting-state runs of a simultaneous EEG–fMRI resting-state dataset (Cha et al., 2025). The dataset originally contained 28 participants and included concurrent EEG, fMRI, and diffusion-weighted imaging; after excluding subjects with interrupted EEG acquisition or substantial motion artifacts, the final sample consisted of 22 subjects (Cha et al., 2025). fMRI data were acquired on a 3T Siemens Prisma scanner with

$TR = 2.0$ s, and EEG was recorded concurrently using a 64-channel MR-compatible Brain Products system at 5000 Hz (Cha et al., 2025). For EEG processing, we used the derivative data with MR-related cardiac artifacts already removed by the dataset authors, followed by standard preprocessing including rereferencing, ICA-based artifact removal, detrending, and source reconstruction with eLORETA on the fsaverage template (Cha et al., 2025; Pascual-Marqui, 2007; Gramfort et al., 2013). Source-space signals were extracted from Schaefer-atlas 100 ROIs (Schaefer et al., 2018), and subsequent analysis focused on default mode network regions to match the resting-state fMRI analysis. In contrast to frequency-specific analyses, here we used broadband source-space EEG signals for the Gaussianity diagnosis. We then evaluated Gaussianity for both the fMRI ROI signals and the broadband EEG source signals.

Rat concurrent LFP–fMRI. We used a rat dataset with concurrent 9.4T fMRI and LFP recordings from the left and right primary somatosensory cortex (S1), which served as the two ROIs, to assess Gaussianity in both hemodynamic and electrophysiological signals. The data were acquired under dexmedetomidine (DMED) anesthesia and consisted of repeated resting-state scans with fMRI repetition time $TR = 0.5$ s. fMRI preprocessing followed standard procedures in SPM12, including motion correction, spatial smoothing, global signal regression, and temporal band-pass filtering (0.01–0.1 Hz). For the electrophysiology data, gradient-related artifacts were removed, recordings were aligned to the fMRI frame structure, and broadband LFP power was computed after denoising and low-pass filtering. We also computed band-limited power (BLP) in six frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), low beta (14–25 Hz), high beta (25–40 Hz), and gamma (40–100 Hz) (Pan et al., 2011; Thompson et al., 2013). Quality control and ROI definition for the bilateral S1 fMRI signals followed prior work on this dataset (Zhang et al., 2020; Pan et al., 2011; Thompson et al., 2013). For the Gaussianity analysis, we examined the resting BOLD signals and corresponding LFP broadband power and BLP timeseries.

To further assess whether approximate scan-level Gaussianity can mask local distributional deviations relevant to directed-connectivity estimation, we also fit local FIR channels to the rat resting-state BOLD signals in sliding windows and computed residual samples as Eq. 2. We then quantified Gaussianity of these window-specific residuals. This residual analysis was performed only for the rat BOLD data and was intended to test whether local channel residuals can exhibit stronger non-Gaussianity than the corresponding whole-scan BOLD signals.

Fig. 7a illustrates temporal variation in the fitted residual distribution of a representative BOLD recording. Separately, across windows from BOLD and LFP band-limited power time series, larger signal-level A–D statistics were associated with larger differences between GCap and Dual-flow capacity estimates (Pearson $r = 0.56$; Fig. 7b). This association links stronger signal-level departures from Gaussianity to greater disagreement between Gaussian and distribution-aware EC estimates.



(a) (b)

Figure 7: **Sliding-window analysis of the LFP–BOLD dataset.** (a) Time-resolved noise histograms in a representative BOLD recording reveal time-variability of the distribution and the deviations from Gaussianity. (b) The difference between GCap and Dual-flow estimates correlates with the Anderson–Darling (AD) statistic of the windowed time series (Pearson $r = 0.56$). Each dot represents the results computed for a window from a BOLD or an LFP band-limited power timeseries.

D.2 Task-Versus-Rest Sensitivity Analysis in Tongue-Motion fMRI

We analyzed the first tongue-motion trial and matched resting-fixation period from acquisition run 1 for each of the 44 participants with usable data from the 45-participant HCP test–retest cohort. Each segment contained 16 fMRI time points (approximately 12 s) and was analyzed separately without additional sliding windows. GIMME was not evaluated because these segments were shorter than its documented minimum time-series length (Appendix A.2). The optimal FIR model order was estimated independently for each directed ROI pair and trial using the corrected Akaike information criterion (AICc) (Hurvich & Tsai, 1989).

For each directed ROI pair, we fit the local FIR model (Eq. 1) within each task or rest trial, and estimated four directed-connectivity measures from the same ROI time series: (i) Dual-flow, which treats the fitted residuals as empirical samples from an unknown continuous residual-noise distribution; (ii) Gaussian capacity (GCap), which uses the same fitted FIR filter but replaces residual uncertainty by a matched Gaussian model based on residual second-order statistics; (iii) pairwise Granger causality as a directed-prediction baseline; and (iv) VAR-LiNGAM, which uses non-Gaussian innovations for structural identification in a vector-autoregressive model. This parallel construction allowed us to compare whether explicitly modeling residual distribution shape improves sensitivity to task-evoked directed interactions beyond Gaussian capacity and Granger-causal baselines.

To quantify task sensitivity, we compared tongue-task and rest EC estimates for each directed ROI pair across the 44 participants, treating participants as the unit of inference. Using the same criterion across estimators, we defined task-sensitive directed interactions as ROI pairs with low between-subject variability (CV below the 30th percentile) and a significant task-related increase relative to rest, assessed using right-tailed paired tests followed by false-discovery-rate correction ($q < 0.1$). Per-pair results for all four estimators are reported in Table 3. Three directed edges met these criteria for Dual-flow and four for GCap; none met the criteria for GC or VAR-LiNGAM. Fig. 8 displays the identified Dual-flow and GCap edges on the brain, alongside the four nonsignificant edges with the lowest FDR-adjusted p-values for each of GC and VAR-LiNGAM. These 16-point task segments provide a practical stress test under limited observations, in which Dual-flow detected task-sensitive directed interactions consistent with the literature-supported circuitry described in Section 4.3.

Table 3: Student's t-test positive false discovery rate (pFDR) and coefficient of variation (CV) comparing Dual-flow, GCap, GC, and VAR-LiNGAM between task and rest conditions for each directed ROI pair in the HCP tongue task dataset. A directed connection is considered task-related if pFDR < 0.1 and CV $< 30\%$ percentile; values meeting each threshold are shown in bold.

Sender ROI	Receiver ROI	Dual-flow		GCap		GC		VAR-LiNGAM	
		pFDR	CV	pFDR	CV	pFDR	CV	pFDR	CV
VI.R	VI.L	.1101	**3.6144**	**.0915**	**3.7060**	.9966	3.9374	.2715	**3.7033**
VI.R	BA3b,OP4.R	.1413	4.1552	**.0767**	**3.1041**	.9966	4.5693	.9912	**2.6882**
VI.R	BA3b,OP4.L	**.0959**	**2.6447**	**.0438**	**2.6062**	.9966	4.0495	.9880	8.2120
VI.L	VI.R	.1537	5.1530	.1963	7.1669	.9966	**3.5852**	.9880	10.9397
VI.L	BA3b,OP4.R	.1570	5.8323	.1963	6.7362	.9966	**2.9687**	.9880	7.0685
VI.L	BA3b,OP4.L	.1537	4.9825	**.0438**	**2.5106**	.9966	6.4292	.9912	**2.7151**
BA3b,OP4.R	VI.R	.1570	5.6009	.1963	6.4526	.9966	4.0842	.3050	5.1349
BA3b,OP4.R	VI.L	.5562	46.6850	.4677	81.3787	.9966	**2.3304**	.6447	11.2962
BA3b,OP4.R	BA3b,OP4.L	**.0985**	**3.2781**	.0915	4.0262	.9966	5.8199	.9779	37.1344
BA3b,OP4.L	VI.R	.1537	4.9899	.1961	5.8293	.9966	4.1199	.2715	**3.9087**
BA3b,OP4.L	VI.L	.3076	11.4084	.0915	4.0120	.9966	**3.6104**	.2715	4.3630
BA3b,OP4.L	BA3b,OP4.R	**.0985**	**3.2164**	.0915	3.8329	.9966	4.1509	.6447	14.2828

This application was designed to test two related questions. First, does Dual-flow detect stronger directed interactions during tongue movement than during rest in physiologically plausible motor circuits? Second, compared with GCap, GC, and VAR-LiNGAM, does Dual-flow provide greater sensitivity when the fitted residuals for tongue-motion trials show stronger departure from Gaussianity? In the main text, we therefore focus on whether the tongue condition—which showed the strongest departure from Gaussianity among the movement conditions—also exhibits the clearest practical benefit from distribution-aware capacity estimation.

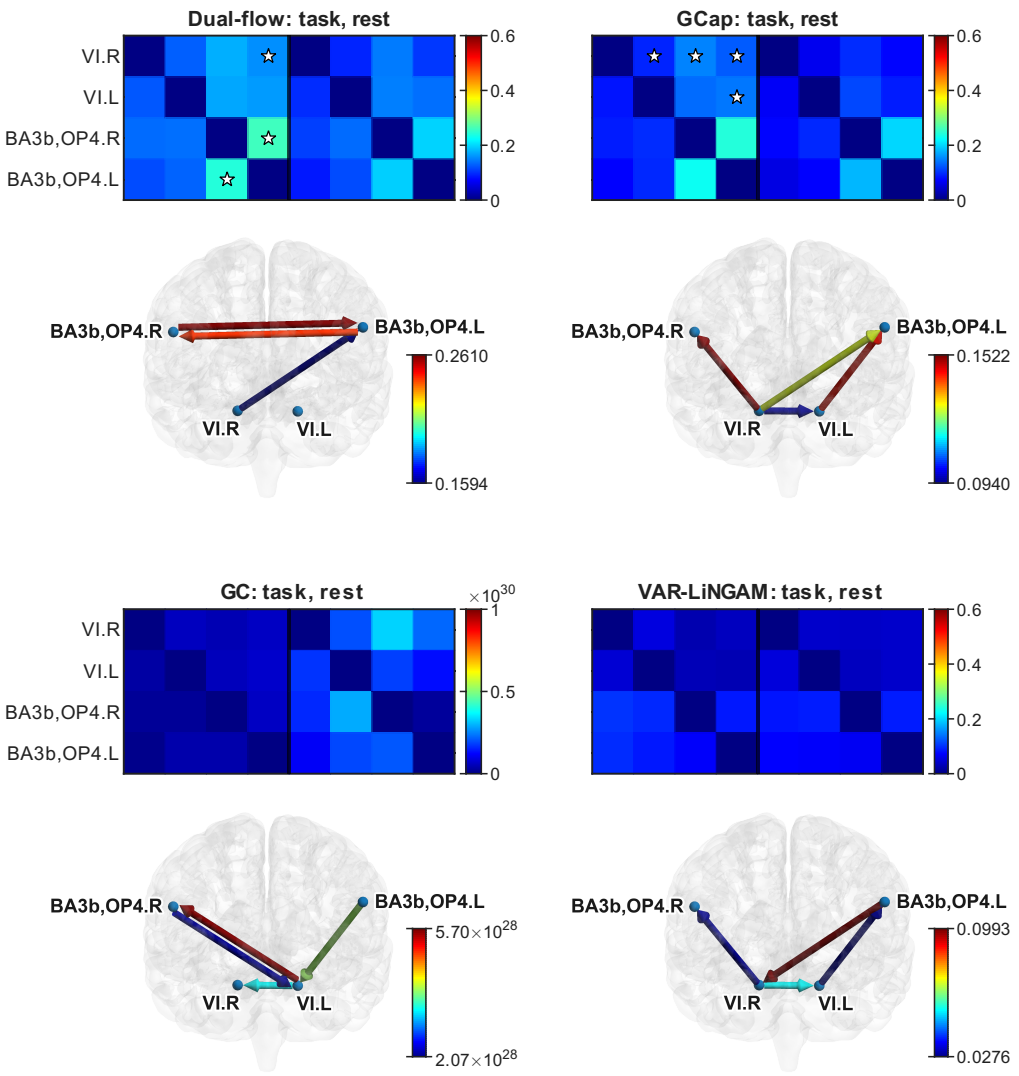


Figure 8: **Connectivity matrices of Dual-flow (upper left), GCap (upper right), GC (lower left), and VAR-LiNGAM (lower right) during tongue motion and rest.** Significant task-related directed interactions are marked with stars and visualized in BrainNet Viewer (Xia et al., 2013). Since no edges met the task-sensitivity criteria for GC or VAR-LiNGAM, the four nonsignificant edges with the lowest FDR-adjusted p-values are visualized for each method. GIMME was excluded because the task segments were shorter than its documented minimum timeseries length.

D.3 Cross-Modal Correspondence of Time-Varying EC Estimates

To assess the physiological relevance of time-varying EC estimates, we analyzed the concurrent rat BOLD–LFP recordings and six LFP band-limited power (BLP) time series described in Appendix D.1. We used 50-s sliding windows, following Thompson et al. (2013), with 50% overlap between consecutive windows (25-s steps). For a fair comparison, all estimators were fitted independently within each window using the same window length and overlap. For each estimator, we estimated directed connectivity from left to right S1 in both modalities, then pooled the windowed estimates across scans and correlated the BLP-derived and corresponding BOLD-derived connectivity sequences. We assessed cross-modal correspondence using 10,000 bootstrap resamples of the 22 scans, retaining all paired windowed estimates within each resampled scan. Statistical significance was assessed using one-sided bootstrap tests of positive correlation, with Benjamini–Hochberg correction across the six bands within each method.

The band-specific correspondence between Dual-flow-derived BOLD and BLP EC estimates matches that previously reported using sliding-window correlation (Thompson et al., 2013, Fig. 2A), with significant correspondence in theta, both beta bands, and gamma. GCap showed a similar pattern but did not reach significance in gamma ($q = 0.068$), whereas GC and VAR-LiNGAM showed no significant correspondence in any band (Fig. 9). This analysis provides cross-modal support, beyond simulation, that time-resolved capacity-based EC estimates capture shared temporal variation in directed interactions across hemodynamic and electrophysiological measurements.

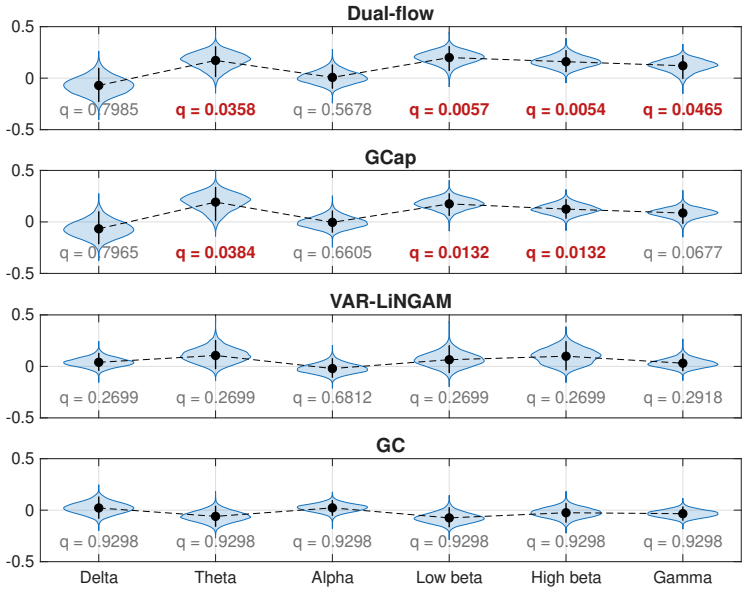


Figure 9: **Cross-modal correspondence of time-resolved EC estimates across six LFP frequency bands.** Each panel shows one estimator. Violins represent correlations between BLP- and BOLD-derived EC estimates (left S1→right S1), pooled across windows, over 10,000 bootstrap resamples of the 22 scans. Black markers and vertical bars indicate bootstrap means and 95% percentile intervals. Values below the violins are one-sided bootstrap q values after Benjamini–Hochberg correction across bands within each method; significant values ($q < 0.05$) appear in bold red.

E Ground-Truth Effective-Connectivity Validation in Brain-Like Simulations

E.1 Simulation Design

We conducted controlled stress tests using a DCM–Balloon simulation framework adapted from Smith et al. (2011), with prespecified ground-truth directed EC matrices. Linear neural dynamics were driven by two-state Markov-switching Gaussian inputs, with off/on means of 0/1, a common SD of 0.05, and off-to-on/on-to-off transition rates of 0.1/0.4 s^{-1}. A nonlinear Balloon–Windkessel model transformed neural activity into BOLD signals. We generated 5-min recordings with $TR = 1$ s and added temporally correlated measurement noise using an AR(1) process ($\rho = 0.4$) with two-component Gaussian-mixture innovations. The resulting noise was scaled to an overall SNR of 10 dB across observed ROIs and realizations.

We evaluated ten conditions spanning common drivers, indirect pathways, hidden nodes, feedback, modular organization, heterogeneous hemodynamics, and increasing network size. Topologies were drawn or adapted from prior benchmarks (Xu et al., 2017; Smith et al., 2011; Sanchez-Romero et al., 2019) or constructed for this study; their sources, directed edges, coupling weights, and modifications are specified in Table 4. We generated 50 realizations per condition, keeping the ground-truth connectivity matrix fixed while varying neural inputs and measurement noise. Hemodynamic parameters additionally varied across realizations in the conditions marked with an asterisk.

In the heterogeneous-HRF conditions, regional BOLD delay offsets received Gaussian perturbations with SD $= 0.10$ s, and regional transit times received multiplicative lognormal perturbations with approximately 5% CV before clipping. Delay offsets were bounded to ± 1.50 s, and transit times to 0.65–1.40 s. These manipulations were motivated by reported between-region and between-subject HRF heterogeneity (Handwerker et al., 2004; Smith et al., 2011).

Table 4: Simulation topologies and manipulations. ROIs are given as observed/total, so that conditions with a hidden node have fewer observed than total ROIs. Asterisks mark conditions in which hemodynamic response functions varied across realizations.

Condition	ROIs	Edges	Ground-truth topology	Manipulation tested	Topology source
CD-e	3/3	2	$1 \to 2$, $1 \to 3$; weights 0.4/0.4	Equal-coupling common driver	Xu et al. (2017)
CD-u	3/3	2	$1 \to 2$, $1 \to 3$; weights 0.8/0.4	Unequal-coupling common driver	Adapted from Xu et al. (2017)
Diamond	5/5	5	$1 \to 2$, $1 \to 3$, $2 \to 4$, $3 \to 4$, $4 \to 5$; weights 0.4	Parallel indirect paths and convergence	Present study
Diamond (H)	4/5	5 total; 3 evaluated	Same diamond; ROI 1 simulated but omitted from the analysis	Hidden common driver of ROIs 2 and 3	Present-study extension
Chain	5/5	5	$1 \to 2 \to 3 \to 4 \to 5$ and $1 \to 5$; weights 0.4	Indirect path with direct shortcut	Smith S5 (Smith et al., 2011)
Chain (HRF)*	5/5	5	Same as Chain	Regional HRF delays $[-0.60, 0.70]$ s; transit multipliers $[0.85, 1.18]$	Smith S5 with HRF manipulation
Feedback	5/5	6	$1 \to 2 \to 3 \to 4 \to 5$, $3 \to 2$, $5 \to 3$; weights 0.4	Reciprocal loop $2 \leftrightarrow 3$ and cycle $3 \to 4 \to 5 \to 3$	Present-study recurrent extension
Mod10	10/10	11	Two S5 subnetworks connected by $3 \to 8$	Two-module organization	Smith S10 (Smith et al., 2011)
Comp12*	10/12	22 total; 18 evaluated	Module A: $1 \to 2 \to 3 \to 4 \to 5$, $3 \to 2$, $5 \to 3$, $1 \to 4$, $2 \to 5$. Module B: analogous over ROIs 6–10. Cross-module: $5 \to 6$, $9 \to 2$. Hidden drivers: $11 \to \{2, 5\}$, $12 \to \{7, 10\}$; weights 0.4	Indirect paths, feedback, modular structure, two hidden drivers, and heterogeneous HRFs combined	Present-study combined topology
Macq28*	28/28	52	Fixed weighted SmallDegree matrix; weights 0.468–0.551; 10 cycles including five reciprocal pairs	Larger recurrent network with independently varying regional HRFs	SmallDegree topology (Sanchez-Romero et al., 2019), pruned from a 28-node subnetwork of the macaque cortical connectome (Markov et al., 2014)

E.2 Evaluation Procedure and Performance

Evaluation Procedure. All methods were evaluated on the same 50 simulated datasets per condition against the same prespecified directed-edge ground truth. We evaluated all $N(N-1)$ ordered edges independently, allowing neither, one, or both directions within each ROI pair to be present. For each realization, AUROC and AUPRC were computed from all ordered-edge scores. Precision, sensitivity, and false-positive rate (FPR) were computed using a method-specific threshold selected from the other 49 realizations within the same condition by maximizing the Matthews correlation coefficient (leave-one-simulation-out, LOSO-MCC). The same calibration procedure was applied to every method-condition pair, excluding the evaluated realization from threshold selection.

Performance and reliability. Table 5 reports the mean and standard deviation of all five metrics across the 50 realizations per condition. Dual-flow achieved the highest mean AUROC and AUPRC, the two threshold-independent ranking metrics, among the evaluated methods across all ten conditions (AUROC: .858–.949; AUPRC: .521–.891), including Comp12 (.879/.724) and Macq28 (.884/.521). Macq28 contains only 52 true edges among 756 candidates, yielding a chance-level AUPRC reference of .069. Across 50 realizations, AUROC/AUPRC were $.888 \pm .108/.857 \pm .132$ for CD-e and $.858 \pm .104/.830 \pm .112$ for CD-u. Threshold-dependent metrics—precision, sensitivity, and FPR—often showed greater variability, particularly in these common-driver conditions. Each realization contains only two true and four null edges, so sensitivity and FPR change in increments of .50

Table 5: Directed-edge recovery in brain-like simulations. Entries are mean ± SD across 50 realizations. Precision, sensitivity, and FPR use LOSO-MCC-calibrated thresholds. Bold marks the best mean in each metric within a condition (highest for AUROC, AUPRC, precision, and sensitivity; lowest for FPR). GIMME was not evaluated on Macq28 because its 28 ROIs exceed the recommended range of 3–20 ROIs (Beltz & Gates, 2017).

Condition	Method	AUROC	AUPRC	Precision	Sensitivity	FPR
CD-e	Dual-flow	**.888±.108**	**.857±.132**	**.887±.206**	.620±.216	.070±.134
	GCap	.700±.129	.623±.155	.472±.069	**.960±.137**	.545±.120
	VAR-LiNGAM	.725±.315	.739±.273	.618±.404	.570±.378	.200±.226
	GIMME	.498±.018	.333±.000	.000±.000	.000±.000	**.005±.035**
	GC	.580±.221	.572±.201	.332±.206	.550±.323	.555±.249
CD-u	Dual-flow	**.858±.104**	**.830±.112**	.890±.273	.470±.120	.040±.093
	GCap	.668±.130	.579±.125	.455±.089	**.860±.227**	.525±.154
	VAR-LiNGAM	.713±.318	.726±.270	.568±.400	.570±.404	.220±.206
	GIMME	.750±.000	.667±.000	**1.000±.000**	.500±.000	**.000±.000**
	GC	.610±.187	.576±.166	.395±.128	.810±.265	.630±.184
Diamond	Dual-flow	**.874±.076**	**.785±.120**	**.853±.168**	.532±.234	.040±.049
	GCap	.777±.050	.541±.065	.418±.072	**.820±.168**	.404±.154
	VAR-LiNGAM	.751±.126	.618±.186	.605±.217	.560±.236	.131±.089
	GIMME	.505±.032	.259±.047	.040±.198	.012±.063	**.003±.013**
	GC	.607±.157	.453±.165	.374±.173	.508±.195	.312±.144
Diamond (H)	Dual-flow	**.910±.079**	**.844±.120**	**.927±.154**	.560±.218	.022±.045
	GCap	.799±.070	.591±.103	.462±.066	**.847±.204**	.344±.135
	VAR-LiNGAM	.787±.172	.675±.220	.620±.247	.660±.247	.149±.111
	GIMME	.502±.037	.260±.049	.000±.000	.000±.000	**.009±.030**
	GC	.589±.233	.494±.226	.394±.279	.507±.303	.311±.200
Chain	Dual-flow	**.938±.042**	**.876±.069**	**.764±.166**	.688±.262	.103±.098
	GCap	.761±.056	.504±.065	.430±.076	**.744±.218**	.355±.165
	VAR-LiNGAM	.715±.160	.603±.200	.606±.326	.420±.253	.101±.096
	GIMME	.522±.051	.287±.071	.230±.419	.052±.097	**.008±.022**
	GC	.603±.100	.428±.118	.347±.183	.384±.217	.260±.141
Chain (HRF)	Dual-flow	**.939±.057**	**.891±.091**	.783±.184	.776±.216	.101±.122
	GCap	.795±.059	.615±.086	.395±.068	**.876±.127**	.472±.154
	VAR-LiNGAM	.693±.134	.511±.144	.354±.137	.568±.266	.337±.155
	GIMME	.605±.024	.409±.036	**.990±.071**	.208±.040	**.001±.009**
	GC	.752±.099	.631±.118	.804±.320	.208±.090	.032±.062
Feedback	Dual-flow	**.889±.071**	**.866±.091**	**.769±.160**	**.790±.183**	.202±.162
	GCap	.613±.043	.534±.051	.476±.040	.748±.144	.555±.134
	VAR-LiNGAM	.563±.147	.545±.141	.511±.243	.305±.173	.192±.108
	GIMME	.333±.006	.400±.000	.000±.000	.000±.000	**.000±.000**
	GC	.514±.111	.495±.104	.411±.058	.750±.182	.715±.165
Mod10	Dual-flow	**.949±.027**	**.815±.073**	**.802±.151**	.616±.205	.029±.030
	GCap	.883±.026	.469±.055	.390±.054	**.782±.149**	.178±.063
	VAR-LiNGAM	.701±.098	.393±.146	.422±.170	.396±.148	.084±.046
	GIMME	.517±.034	.153±.055	.230±.384	.038±.064	**.005±.010**
	GC	.630±.098	.266±.090	.219±.095	.411±.173	.218±.083
Comp12	Dual-flow	**.879±.025**	**.724±.052**	.708±.113	.581±.083	.066±.036
	GCap	.852±.023	.644±.039	.542±.063	**.656±.077**	.143±.041
	VAR-LiNGAM	.692±.062	.428±.080	.409±.099	.426±.124	.160±.063
	GIMME	.566±.005	.313±.019	**.870±.188**	.088±.028	**.005±.007**
	GC	.736±.044	.483±.060	.446±.093	.441±.158	.157±.090
Macq28	Dual-flow	**.884±.018**	**.521±.042**	**.573±.118**	.453±.105	**.029±.018**
	GCap	.846±.021	.391±.037	.427±.082	.403±.074	.043±.019
	VAR-LiNGAM	.578±.038	.110±.022	.102±.020	.305±.078	.200±.046
	GC	.679±.029	.145±.022	.122±.020	**.568±.125**	.315±.102

and .25, respectively. LOSO-MCC calibration yielded precision/FPR of .887/.070 for CD-e and .890/.040 for CD-u, with sensitivity of .620 and .470, respectively. This coarse metric resolution partly explains the larger threshold-dependent SDs, while mean edge-ranking performance remained strong. GCap achieved higher sensitivity in several conditions but also substantially higher FPR, whereas GIMME generally showed low FPR alongside low sensitivity. Overall, Dual-flow maintained strong directed-edge recovery across individual stressors, their coexistence, and the larger network.

F COMPUTATIONAL REQUIREMENTS

Dual-flow performs iterative optimization of the generator and observer flows, while independent directed-pair evaluations can be executed concurrently. As a representative benchmark, we evaluated the Mod10 workload comprising 900 directed EC evaluations (10 realizations $\times$ 90 directed pairs) using 4 $\times$ NVIDIA L40S GPUs on an AMD EPYC 7763 64-core system with 1.0 TiB RAM, with eight pair evaluations scheduled concurrently per GPU. Across five seeded repetitions, the complete workload required 1272.74 ± 122.57 s ($\sim$21.2 min) of wall time, corresponding to an amortized 1.41 ± 0.14 s per directed EC evaluation. We used $L = 256$, for which capacity estimates were stable across the evaluated training-sequence lengths (Appendix C.2). The current implementation is well suited to hypothesis-driven analyses of prespecified circuits and moderate-sized subnetworks, an established practice in effective-connectivity studies across cognitive and clinical neuroimaging (Smith et al., 2012; Deshpande & Hu, 2012; Bielczyk et al., 2019).