Efficient Coding in the Economics of Human Brain Connectomics

Dale Zhou[1], Christopher W. Lynn[2], Zaixu Cui[3], Rastko Ciric[4], Graham L. Baum[5], Tyler M. Moore[3], David R. Roalf[3,6], John A. Detre[7], Ruben C. Gur[3,6], Raquel E. Gur[3,6], Theodore D. Satterthwaite[3,6] †, and Danielle S. Bassett[2,3,7,8,9,10,11] †

[1]Neuroscience Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania
[3]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania
[4]Department of Bioengineering, Schools of Engineering and Medicine, Stanford University, Stanford, CA 94305
[5]Department of Psychology and Center for Brain Science, Harvard University, Cambridge MA USA
[6]Penn-Children's Hospital of Philadelphia Lifespan Brain Institute
[7]Department of Neurology, Perelman School of Medicine, University of Pennsylvania
[8]Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania
[9]Department of Electrical & Systems Engineering, School of Engineering and Applied Sciences, University of Pennsylvania
[10]Santa Fe Institute, Santa Fe, NM 87501 USA
[11]To whom correspondence should be addressed: dsb@seas.upenn.edu
†Co-senior authors

January 14, 2020

Abstract: 150 words
Main text: 5,466 words
8 figures
12 supplementary figures

1

Abstract

In systems neuroscience, most models posit that brain regions communicate information under constraints of efficiency. Yet, metabolic and information transfer efficiency across structural networks are not understood. In a large cohort of youth, we find metabolic costs associated with structural path strengths supporting information diffusion. Metabolism is balanced with the coupling of structures supporting diffusion and network modularity. To understand efficient network communication, we develop a theory specifying minimum rates of message diffusion that brain regions should transmit for an expected fidelity, and we test five predictions from the theory. We introduce compression efficiency, which quantifies differing trade-offs between lossy compression and communication fidelity in structural networks. Compression efficiency evolves with development, heightens when metabolic gradients guide diffusion, constrains network complexity, explains how rich-club hubs integrate information, and correlates with cortical areal scaling, myelination, and speed-accuracy trade-offs. Our findings elucidate how network structures and metabolic resources support efficient neural communication.

1   Introduction

Darwin described the law of compensation as the concept that "to spend on one side, nature is forced to economise on the other side," [1]. In the economics of brain connectomics, natural selection optimizes network architecture for versatility, resilience, and efficiency under constraints of metabolism, materials, space, and time [2, 1, 3]. Networks – composed of nodes representing cortical regions and edges representing white matter tracts – strike evolutionary compromises between costs and adaptations [2, 1, 4, 5, 6, 7], whereby disruptions may contribute to the development of neuropsychiatric disorders [8, 9, 10, 11]. To understand how the brain efficiently balances resource constraints with pressures of information processing, models of information diffusion in brain networks are necessary. Such models have gained traction [12, 13, 14, 5, 15, 16], but it is unknown how a network of brain regions efficiently transmits messages to targets in the presence of countless alternative routes that are spatially embedded in diverse architectures of connectivity [3, 16, 17].

Novel brain network communication models are needed because the predominant theories of shortest path routing and diffusion have been criticized as infeasible or inefficient [3, 15, 16]. In shortest path routing, neural signals travel from source to target using either the fewest connections or the shortest spatial distance [16]. Shortest path routing assumes biologically infeasible global information of path length or greedy selection of distances. Diffusion models assume an inefficient process of random propagation from source to target. In contrast to these models, the efficient coding hypothesis proposes that the brain represents information in a metabolically economical or compressed form by taking advantage of redundancy in the structure of information [18, 2]. Coding efficiency characterizes low-dimensional neural representations and dynamics supporting cognition [19, 20, 21]. New models should therefore demonstrate metabolic and information transfer efficiency that predictably differ according to variation in brain network structure across the protracted development of structural connectivity [3, 12, 5, 22, 17, 7, 16].

We develop a brain network communication model of efficient coding by information diffusion (Figure 1A). We apply our model to 1,042 youth (aged 8-23 years) in the Philadelphia Neurodevelopmental Cohort who underwent diffusion tensor imaging (DTI) and arterial-spin labeling (ASL; see Supplementary Figure 1) [23]. To operationalize metabolic expenditure, we use ASL, which measures cerebral blood flow (CBF) and is correlated with glucose expenditure and ATP consumption (Figure 1B) [24]. We join work modeling efficient coding with rate-distortion theory [25], a branch of information theory that provides the mathematical foundations of lossy compression [26]. By assuming that the minimal amount of noise is achieved by signals that

diffuse along shortest paths [12, 14, 16], we calculate the optimal rate of signal transmission to communicate between brain regions with an expected transmission fidelity in the capacity-limited structural network. Specifically, we define the expected signal distortion as the probability of *not* propagating along the shortest path. In developing the framework, we seek to understand how network structure and metabolic resources support and constrain the efficient transmission of information.

To evaluate the validity of our efficient diffusion model, we assess five published predictions of rate-distortion theory and information diffusion (Figure 1C) [27, 25]. As we will describe in detail, hypotheses of information diffusion models posit how network structure guides propagating signals in support of metabolic efficiency, transmission fidelity, and information integration [28, 12, 13, 29]. Hypotheses of rate-distortion theory posit that the trade-off between message fidelity and compression governs predictable differences in the efficiency of information broadcasting across networks [27, 25]. In evaluating the validity of our efficient diffusion model, we introduce *compression efficiency*, which quantifies how much structural networks prioritize lossy compression versus communication fidelity. To demonstrate the utility of our model, we use compression efficiency to test the hypothesis that diffusing information is integrated and broadcast by the brain's highly connected regions or hubs [29]. Finally, we use compression efficiency to explain individual variation in the speed-accuracy trade-off of cognitive performance, and we contrast its explanatory power with that of competing measures [30]. Our model advances the current understanding of how efficiency, noise, and information integration are associated with metabolic resources and network architecture.

2 Results

2.1 Metabolic running costs of network communication architectures

We sought to distinguish how brain metabolism is associated with structural signatures of shortest path routing versus diffusion signaling. Although shortest path routing is hypothesized to reduce metabolic cost, existing evidence for this hypothesis remains sparse [28]. To quantify the extent to which a person's brain is structured to support shortest path routing, we used the *global efficiency*, a commonly computed measure of the average shortest path strength between all pairs of brain regions. Intuitively, global efficiency represents the ease of information transfer by the strength of direct connections in a network. As an operationalization of metabolic running cost, we considered CBF, which is correlated with glucose consumption. To test the
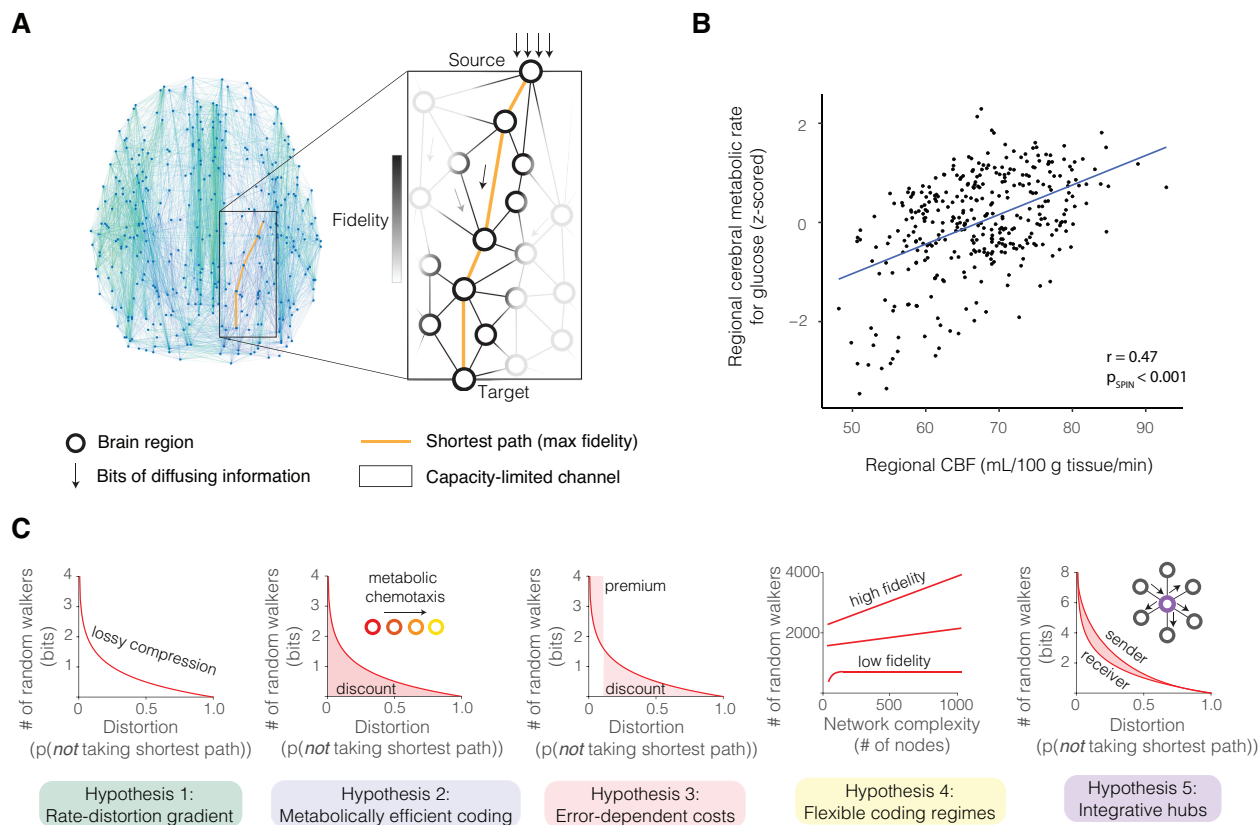
Figure 1: **The efficient diffusion model and associated hypotheses. (A)** To model efficient coding, we apply rate-distortion theory, a branch of information theory that provides the foundations of lossy compression, to a previously proposed network measure called *resource efficiency* [12]. Resource efficiency models probabilistic diffusion by shortest paths, solving for the number of random walkers required for at least one to propagate by the shortest path given an expected probability. By assuming that the minimal amount of noise is achieved by signals that diffuse along shortest paths, we calculate the optimal information rate to communicate between brain regions with an expected transmission fidelity in the rate-limited (or capacity-limited) structural network. **(B)** Across the brain, regional CBF measured from 1,042 participants in our study correlates with the regional cerebral metabolic rate for glucose acquired from published maps of 33 healthy adults (Pearson's correlation coefficient $r = 0.47, df = 358, p_{\text{SPIN}} < 0.001$), replicating prior findings and supporting the operationalization of metabolic expenditure using CBF [24]. We use CBF to investigate the relationship between metabolic demands and network organization supporting either shortest path routing or diffusion. **(C)** Importantly, our model generates five predictions from the rate-distortion theory and information diffusion literature [25, 27, 29]. First, information transfer should produce a characteristic rate-distortion gradient in empirical and artificial networks, where exponentially increasing information rates are required to minimize signal distortion. Second, information transfer efficiency should improve with manipulations of the system architecture designed to facilitate signal propagation, where information costs decrease when message diffusion is biased with regional differences in metabolic rates. Third, the information rate should vary as a function of the costs of error, with discounts when costs are low and premiums when costs are high. Fourth, brain network complexity should flexibly support communication regimes of varying costs and fidelity, where a high-fidelity regime predicts information rates that exponentially increase as the network grows more complex, and a low-fidelity regime predicts asymptotic information rates indicative of lossy compression. Fifth and finally, structural hubs should integrate incoming signals to efficiently broadcast information, where hubs (compared to other brain regions) have more compressed input rates and higher transmission rates for equivalent input-output fidelity.
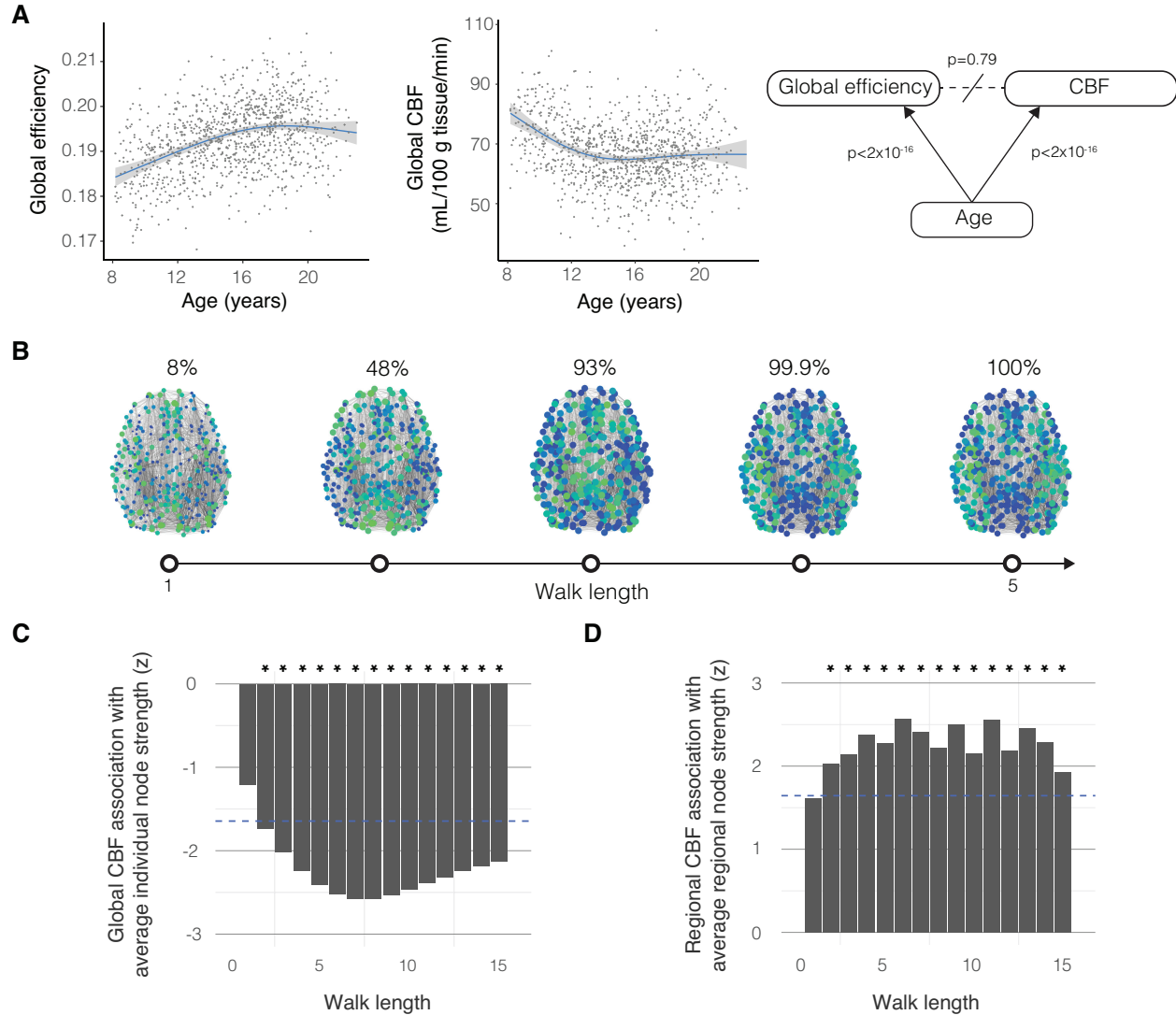
Figure 2: **Metabolic running costs support brain network architectures for diffusion. (A)** Global efficiency increases with development while cerebral perfusion declines ($F_{\text{global efficiency, age}} = 50$, estimated $df = 3.46$, $p < 2 \times 10^{-16}$; $F_{\text{CBF,age}} = 69.22$, estimated $df = 3.74$, $p < 2 \times 10^{-16}$). Contrary to prior reports, the relationship between CBF and global efficiency is confounded by age ($r = 0.01, df = 1039, p = 0.79$). Therefore, the claim of reduced metabolic cost associated with shortest path routing is weakened. **(B)** All brain regions are accessible to signals randomly walking across 5 edges or more. Bluer nodes represent greater normalized node strengths, reflecting the accessibility of the brain region to a random walker diffusing along the structural connectome with differing walk lengths. Shortest path routing predicts that only shortest path walks of length 5 or less are associated with CBF. Diffusion models predict that longer random walk paths are also associated with CBF. **(C)** Global CBF is negatively correlated with the strengths of structural paths with lengths of at least 7 when controlling for age, sex, age-by-sex interactions, degree, density, and in-scanner motion ($t = -1.59$ to -2.81, estimated model $df = 11.45$, FDR-corrected $p < 0.05$). Metabolic expenditure is related to structure supporting signal diffusion, rather than shortest path routing. **(D)** Regional CBF is positively correlated with high-integrity connections comprising the structural scaffolds supporting differing walk lengths ($\rho = 0.10$ to 0.15, $df = 358$, FDR-corrected $p < 0.05$), controlling for age, sex, age-by-sex interaction, degree, density, and in-scanner motion. Asterisks denote statistical significance following correction for multiple comparisons. Together, these data suggest that metabolic cost is associated with a regional profile of white matter path strengths that support diffusing messages. Individuals with greater path strengths tend to have lower global metabolic expenditure, while brain regions with greater path strengths tend to have greater metabolic expenditure.

spatial correlation between CBF and glucose consumption, we used a spatial permutation test that generates a null distribution of randomly rotated brain maps that preserves the spatial covariance structure of the original data; the p-value reflecting significance is denoted p_{SPIN} (Method 7.7.6). We replicated the previously reported linear association with glucose consumption (Figure 1B; Pearson's correlation coefficient $r = 0.47, df = 358, p < 0.001; p_{\mathrm{SPIN}} < 0.001$) [24]. Therefore, if shortest path routing is linked to a decrease in metabolic expenditure, then we should observe a negative correlation between global efficiency and CBF. Moreover, the negative correlation should explain variance in the data above and beyond the developmental effects of age. In a sensitivity analysis excluding age, but controlling for mean gray matter density, sex, mean degree, network density, and in-scanner motion, the global efficiency was correlated with CBF ($r = -0.20, df = 1039, p < 0.001$). However, we found that global efficiency was positively correlated with age (Figure 2A; $F = 50$, estimated $df = 3.46$, $p < 2 \times 10^{-16}$), while CBF was negatively correlated with age ($F = 69.22$, estimated $df = 3.74$, $p < 2 \times 10^{-16}$), and after controlling for age we do not find a significant relationship between global efficiency and CBF ($r = 0.01, df = 1039, p = 0.79$). Therefore, age confounds the relationship between global efficiency and CBF, and the data does not support the claim that shortest path routing is associated with reduced metabolic expenditure.

Rather than being driven by shortest path routing, metabolic expenditure could instead be associated with communication by diffusion. Each brain region can reach every other brain region via diffusion along paths of 5 connections (Figure 2B). A diffusing message will likely not take the most efficient paths and must instead rely on the structural strengths of longer paths. Hence, if brain metabolism is associated with communication by diffusion, then CBF should correlate with the strength of the white matter paths greater than length 5. We computed the strength of connections across different diffusion distances using the matrix exponent of the structural network (see Method 7.5.2 and Supplementary Figure 1B). We then tested the association between longer paths and metabolic expenditure across both individuals (Figure 2C) and regions (Figure 2D). Considering variation across individuals, we found that the average node strengths for walks of length 2 to 15 were negatively correlated with CBF ($t = -1.59$ to -2.81, estimated model $df = 11.45$, FDR-corrected $p < 0.05$), controlling for age, sex, age-by-sex interaction, average node degree, network density, and in-scanner motion (Figure 2C). The negative correlations between CBF and the average connection strengths suggest that the greater the connection integrity, the lower the metabolic expenditure. When examining variation across brain regions, we found that node strengths comprising walks of lengths 2 to 15 were positively correlated with CBF (Spearman's rank correlation coefficient $\rho = 0.10$ to 0.15, $df = 358$,

FDR-corrected $p < 0.05$), controlling for age, sex, age-by-sex interaction, average node degree, network density, and in-scanner motion (Figure 2D). Brain regions with greater path strengths tended to have higher metabolic expenditure. The convergent findings of an association between CBF and path strengths across individuals and regions suggest that metabolic running costs are linked to diffusion signaling and not to shortest path routing.

2.2   Adaptive trade-offs between metabolism and network architecture

Communication between brain regions or modules requires reliable broadcasting of information with an expected fidelity. Although our data does not link metabolic expenditure to shortest path routing, communication of information diffusing along shorter paths should nevertheless confer advantages in speed and signal fidelity compared to longer paths. To test this hypothesis, we investigated whether brain metabolism is associated with network structures that support diffusion over shorter paths. Specifically, we assessed the association between CBF and path transitivity, a measure of the density of connections re-accessing shortest paths, thereby guiding diffusion along efficient pathways (Figure 3A). Prior reports have demonstrated that path transitivity in structural networks is positively correlated with fMRI BOLD functional connectivity [13], a finding that we replicate in our own data (Supplementary Figure 2). Path transitivity requires more connections and presumably incurs greater metabolic running costs associated with both the structural connections and increased functional connectivity [31]. When considering variation across individuals, we find that greater path transitivity is associated with greater CBF (Supplementary Figure 3A; $t = 2.27$, estimated model $df = 11.45$, $p = 0.02$; controlling for age, sex, age-by-sex interaction, degree, density, and in-scanner motion). This result suggests that brain networks may strike a compromise between metabolic cost and the signaling advantages of path transitivity. Next, we sought to assess whether the relationship between brain metabolism and path transitivity was moderated by development. We found that the interaction between path transitivity and age was positively associated with CBF ($F = 24.6$, estimated $df = 3.13$, $p < 2 \times 10^{-6}$; Supplementary Figure 3B). Increased metabolic expenditure associated with greater path transitivity was prominent during adolescence, when global CBF tends to decrease [32].

We expanded our analysis of compromises between brain metabolism and network topology by considering multiple trade-offs. Specifically, we considered variations in metabolic cost, path transitivity, and modularity across individuals (Figure 3B). We found that the relationship between path transitivity and global
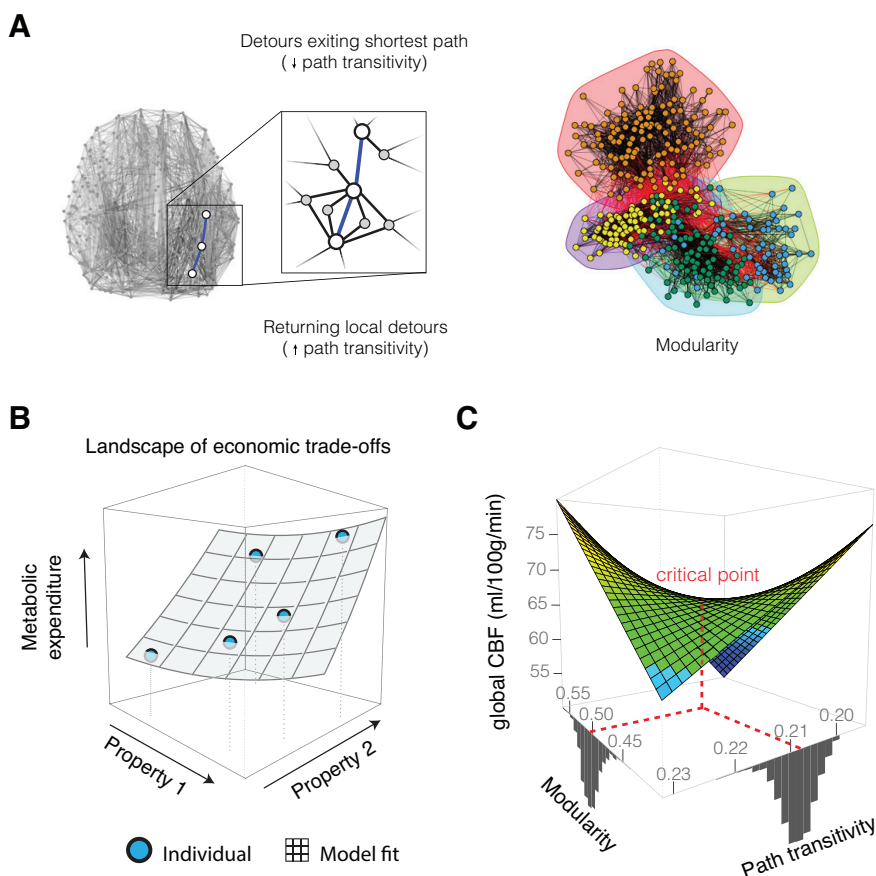
Figure 3: **Trade-offs between modularity and diffusion architecture locally optimize metabolic running cost. (A)** Greater path transitivity enhances the ability of diffusing signals to re-access the shortest path when the number of closed paths (triangles) returning to nodes on the path is high; path transitivity of the brain's structural network is statistically associated with fMRI BOLD activity [13]. Modularity is a feature of brain organization across species whereby regions cluster into highly intraconnected communities; modularity is thought to confer efficient use of physical materials, segregate information representation, and support efficient information transfer among brain regions. **(B)** The 3-dimensional fitness landscape conveys adaptive trade-offs associated with brain metabolism in the context of evolutionary constraints on efficiency. **(C)** When we consider variation across individuals, we find that economical network architectures are adaptively balanced with metabolism. Metabolic cost is associated with the interaction of path transitivity and modularity ($t = 2.56$, estimated model $df = 13.43$, $p = 0.01$), controlling for age, sex, age-by-sex interaction, degree, density, and in-scanner motion. The mean path transitivity and modularity across all individuals (red dotted lines projected from frequency histograms) approach a saddle point (critical point), defined as a point on the surface that is both a relative minimum and a maximum along different axes. The saddle point suggests that adaptive compromises in network architecture are constrained by dual objectives. Along one axis, the objective is minimizing metabolic expenditure by coupling modularity with path transitivity. Along the other axis, the objective is maximizing metabolic expenditure by decoupling modularity from path transitivity. This landscape suggests the existence of compromises which balance adaptations of functional flexibility and spatial efficiency with material and metabolic costs.

CBF is moderated by modularity ($t = 2.56$, estimated model $df = 13.43$, $p = 0.01$). When we consider variation across individuals, we find a saddle point function of metabolic costs, where the means for both path transitivity and modularity fall at the critical point (Figure 3C). A saddle point suggests that adaptive compromises in network architecture are constrained by dual objectives. Along one axis, the objective is minimizing metabolic expenditure by coupling modularity with path transitivity. Along the other axis, the objective is maximizing metabolic expenditure by decoupling modularity from path transitivity. Brain networks may negotiate multiple trade-offs between metabolism and structure such that most brain networks reside around a critical point with locally optimal metabolic savings when network structure is coupled (saddle point), whereas a smaller fraction of brain networks exhibit globally optimal metabolic savings when network structure is decoupled (global minima).

2.3 Systems-level efficient coding as lossy compression

To understand how the brain balances the transmission rate of diffusing signals and signal distortion across different network architectures, we propose a measure called *compression efficiency*, which synthesizes shortest path routing and diffusion (Figure 1A and 4A). We have thus far described how individual differences in metabolic running costs and brain architecture suggest that the brain communicates by diffusion. We formalize a rate-distortion model of efficient diffusion by assuming that the minimal amount of noise is achieved by signals that diffuse along shortest paths (Figure 1A and 4B). We define distortion as the probability of a diffusing signal *not* taking the shortest path. To understand how the brain balances information rate and distortion, we measure resource efficiency: the number of resources required for at least one resource to randomly walk along the shortest path to a target cortical region, with an expected probability (Figure 4C; Method 7.5.5). Just as rate-distortion theory predicts the minimum information rate needed to achieve a specified signal distortion transmitting through a capacity-limited channel, resource efficiency predicts the minimum number of resources needed to achieve a specified level of signal distortion resulting from diffusion across the structural connectome (Figure 4D). The information-theoretic trade-off between information rate and signal distortion is defined by individually different rate-distortion gradients (Figure 4E). As distortion increases, the information rate decays exponentially. By analogy with rate-distortion theory, here we consider the extent to which the brain's structural connectome prioritizes compression versus fidelity. We refer to this tradeoff as the *compression efficiency* (Figure 4E), and define it as the slope of the exponential rate-distortion gradient (Method 7.5.6).
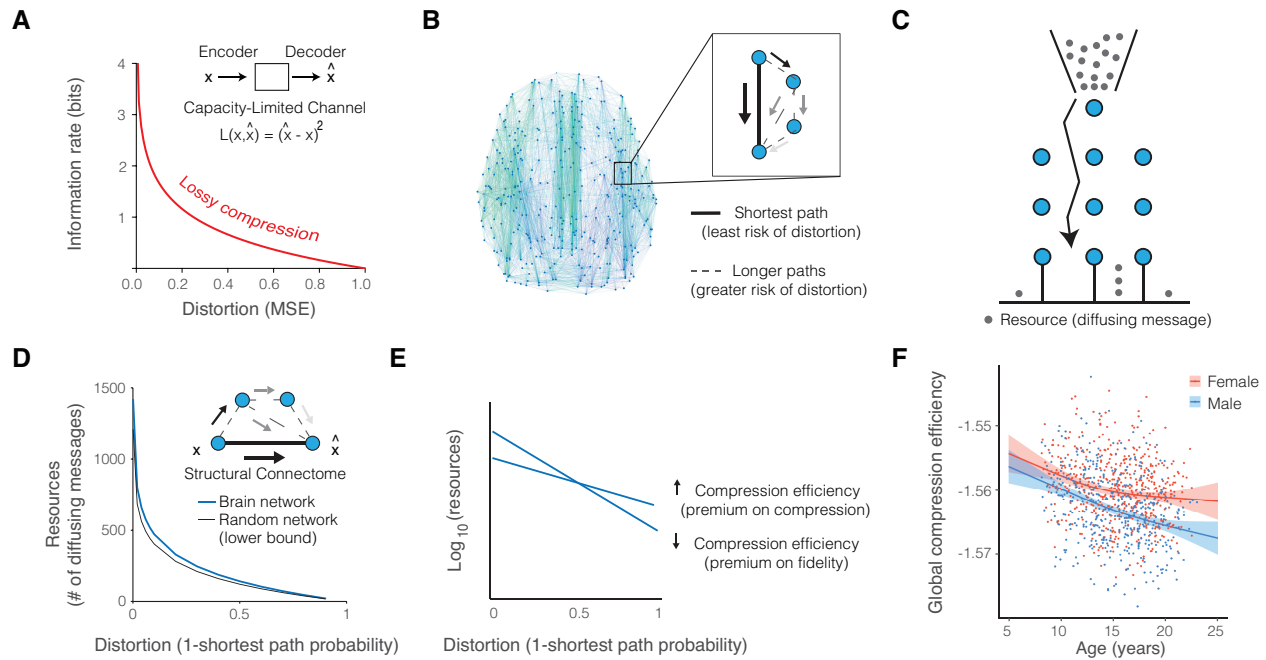
Figure 4: **Neurodevelopment places a premium on fidelity. (A)** Rate-distortion theory is a mathematical framework that defines the required amount of information for an expected level of signal distortion during communication through capacity-limited channels. Loss functions, such as the mean squared error, are defined to map encoded and decoded signals. In a given channel, compressed messages demand lower information rates at the cost of fidelity. **(B)** In brain networks, we define the distortion function as the probability of a diffusing signal *not* taking the shortest path. This definition is reasonable because the temporal delay, signal mixing, and decay introduced by longer paths collectively increase distortion. **(C)** The Galton board depicts the problem of determining the minimum number of resources (diffusing messages) that the starting node should prepare to transmit in order for one resource to propagate by the specified path given an expected probability. Formalizing the shortest path distortion function allows analytical solutions to determine the corresponding minimum rate of information using the metric of resource efficiency. Recall that the resource efficiency solves for the number of randomly walking resources (or messages) required for at least one walker to propagate by the shortest path with an expected probability. **(D)** In applying our theory to the structural connectome, we observe a characteristic curve that is consistent with that predicted by the rate-distortion function. Just as an artificial pixelated image is governed by the same information-theoretic rules as a naturalistic image, the synthetic random networks exhibit the rate-distortion gradient also observed in brain networks, supporting the first prediction of rate-distortion theory. The blue curve depicts the loess fit of the mean rate-distortion function across all individuals. The black curve depicts the mean rate-distortion function for Erdös-Renyí random networks whose edges maintain the weights from empirical measurements. The resources required of brain connectomes and random networks differ by distortion ($F = 10 \times 10^5, df = 29120, p < 2 \times 10^{-16}$), highlighting the established graph theoretical notion that random networks increase shortest path accessibility but demand a heftier materials cost. **(E)** We measure the differing slopes of the semi-log plot of the rate-distortion gradients as differing *compression efficiency*. Then, we interpret variation across individuals in the language of rate-distortion theory. For example, consider two brain networks functioning at the same low level of distortion. The brain network with the flatter slope between resources and distortion has greater compression efficiency because the network architecture confers resource discounts. In comparison, the brain network with a steeper slope has reduced compression efficiency because the network architecture pays a premium for the same expected fidelity. **(F)** When we consider variation across individuals, we find that compression efficiency decreases with age and differs by sex (Supplementary Figure 4; $F = 27.54$, estimated $df = 2.17, p < 0.001$), suggesting that neurodevelopment places a premium on high-fidelity network communication.

To evaluate the roles of resource efficiency and rate-distortion theory in the brain, we assess five previously published predictions of rate-distortion theory and information diffusion (Figure 1C) [25, 29]. The first prediction of rate-distortion theory is that communication systems should produce an information rate that is an exponential function of distortion. Moreover, artificial networks should be governed by the same information-theoretic rules as empirical networks. To test this prediction, we computed the resource efficiency of each individual, with the probability of diffusion along the shortest path ranging from 10% to 99.9% (Figure 4D). We designed artificial communication systems as Erdös-Renyí random networks (Method 7.6), which predominantly transfer information by short paths [33]. We observed an exponential gradient in individual brain networks and the Erdös-Renyí random networks, consistent with the first prediction of rate-distortion theory. Furthermore, the random networks, which are composed of more short connections than empirical brain networks, incurred a decreased resource cost compared to empirical brain networks (Figure 4D and Supplementary Figure 4; $F = 10 \times 10^5, df = 29120, p < 2 \times 10^{-16}$), consistent with the intuition that a greater prevalence of short connections in the random network translates to greater likelihood of shortest path diffusion [33]. Rate-distortion trade-offs vary as a function of age and sex, where individual differences in compression efficiency (Figure 4E) were negatively correlated with age ($F = 27.54$, estimated $df = 2.17$, $p < 0.001$), suggesting that neurodevelopment places a premium on fidelity (Figure 4F). Compression efficiency differs on average by sex, in parallel with the sex-specific developmental trajectories of CBF (Supplementary Figures 5-6). The data, therefore, indicate that resource efficiency gradients – that is, the trade-offs between resources and distortion – vary predictably across different network architectures.

The second prediction of rate-distortion theory is that manipulations of the system architecture designed to facilitate signal propagation should reduce resource costs. Efficient coding characterizes the constraint on information costs by metabolic costs. Whereas we have so far only provided information about local connection strengths to diffusing signals, we now modify edge weights to additionally describe regional changes in local brain metabolic rate (Figure 5A; Method 7.5.7). Metabolic chemotaxis describes a mechanism of diffusion in which random propagation is biased along gradients of increasing or decreasing metabolic resources. Depending on information processing demands, metabolic efficiency could be characterized by connected regions with high metabolic resources that attract chemosensitive neural signals or that repel them to connected regions with lower cost. To assess if chemosensitive neural signals support compression efficient coding, we modify our information diffusion model to allow brain regions to either attract random walkers with

increased probability or repel walkers with decreased probability as a function of greater metabolic expenditure in the source and target regions. If chemotactic diffusion along metabolic gradients reduces the required resources compared to diffusion along the original network of structural connection strengths (Figure 5B), then chemotaxis will appear to support compression efficiency. In the structural networks biased to attract or repel diffusing signals by regions of high cerebral blood flow, we observed that the rate-distortion gradient differs between unbiased diffusion and chemotactic diffusion when distortion is less than 60% (Figure 5C; $F = 6 \times 10^5, df = 29120$, all p-values corrected using the Holm-Bonferroni method for family-wise error rate; $p < 0.05$ at 60% distortion, $p < 0.01$ at 50% distortion, and $p < 0.001$ at distortion less than or equal to 40%). The differences arise from reduced resource requirements introduced by additional information from regional CBF (attract: $t = 20.9, df = 1993.9, p < 0.001$; repel: $t = 22.69, df = 1970.7, p < 0.001$; Figure 5D), supporting the second prediction of rate-distortion theory that metabolic rates support the efficient communication of information among brain regions. Together, these results support the use of rate-distortion theories of capacity-limited, efficient coding to model brain network communication by diffusion.

The third prediction of rate-distortion theory states that the information rate should vary as a function of the costs of errors in empirical systems that interact with the environment. If errors are more costly for networks operating at a high fidelity with exponentially greater information rates, then we should observe a resource rate surpassing the minimum predicted by rate-distortion theory. In contrast, if errors are less costly for networks operating at low fidelity with exponentially reduced resource rates, then we should observe no more than the minimum predicted resources. We observed that brain networks commit more resources than required for very low levels of distortion, such as 0.1%, but allocate the predicted resources or fewer to guarantee levels of distortion between 2% and 90% (Figure 5C). Hence, the third prediction of rate-distortion theory was consistent with our observation of a premium placed on very low signal distortion and a discounted cost of greater distortion.

Next, we sought to test the fourth prediction of our model by identifying network properties that support a system that we might expect to transmit information in a high- or low-fidelity regime according to the environment in which the system operates. Systems functioning in high-fidelity regimes place premiums on accuracy, even given some expected level of error. Systems functioning in low-fidelity regimes are tolerant to noise and fix resource rates despite increasing complexity. With increasing complexity, a high-fidelity regime will continue to place a premium on accuracy, whereas a low-fidelity regime will tolerate noise in support of
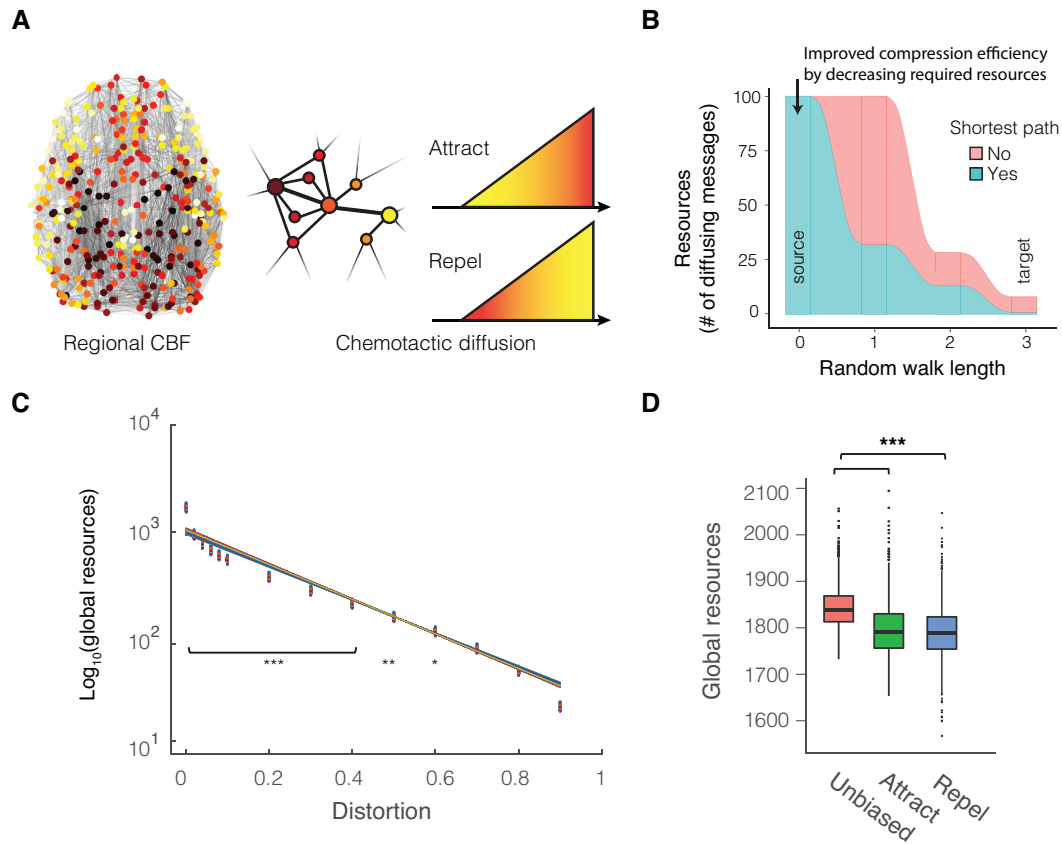
Figure 5: **Metabolic chemotaxis supports efficient coding. (A)** Chemotactic diffusion was modeled as a biased random walk with transition probabilities modified by regional CBF. A biased random walk attracting diffusing signals to regions of high CBF was modeled by constructing a biased structural connectivity matrix. The weight of an edge between region i and region j was given by the average, normalized CBF between pairs of brain regions multiplied by the weight of the structural connection. For biased random walks repelling diffusing signals from regions of high CBF, the weight of an edge was given by (1 minus the normalized inter-regional CBF) multiplied by the weight of the structural connection. **(B)** The second prediction of rate-distortion theory is that providing additional chemotactic information to the information processing system will reduce information cost. Specifically, we hypothesized that diffusion along metabolic gradients would increase compression efficiency. The alluvial diagram depicts our expected results: if metabolism acts as a medium supporting efficient coding, then we should observe reduced minimum resources in chemotactic diffusion compared to unbiased diffusion. **(C)** The variation of predicted minimum resources required for distortion levels less than 60% is explained by the interaction of the distortion level with the type of chemotactic diffusion (*** : $p < 0.001$), supporting the second prediction of rate-distortion theory. According to the third prediction of rate-distortion theory, we should expect asymmetries about the rate-distortion gradient where there are asymmetric costs of error. Specifically, the cost of error should be greater during neural dynamics requiring high-fidelity communication compared to low-fidelity communication. The level of distortion and type of random walk fully explain variance in resource efficiency ($R^2 = 0.99$, $F = 6 \times 10^5$, $df = 29120$, all p-values corrected using the Holm-Bonferroni method for family-wise error rate; $* : p < 0.05$, $** : p < 0.01$, $*** : p < 0.001$). Across individual brain networks, the resources required for 0.1% distortion are greater than predicted by the rate-distortion gradient, reflecting a premium placed on very high fidelity signaling. At greater levels of distortion, the resources required are less than predicted by the rate-distortion gradient. The premium cost of high-fidelity communication compared to discounts of low-fidelity communication supports the third prediction of rate-distortion theory. **(D)** In agreement with our hypothesis that chemotaxis supports efficient coding, the number of required resources decreased in chemotactic diffusion compared to unbiased diffusion ($t_{\text{unbiased,attract}} = 20.87$, $df = 1993.9$; $t_{\text{unbiased,repulse}} = 22.69$, $df = 1970.7$; ***: $p < 0.001$).

lossy compression. For an information-encoding channel aiming to achieve an expected level of signal distortion in a high-fidelity regime, rate-distortion theory predicts that the information rate, here operationalized as resources, should monotonically increase with the complexity of the information-encoding system (Figure 6A). To evaluate this prediction, we operationalized complexity as network size, or the number of nodes, to maintain consistency with the methods of existing predictions [27]. In a low-fidelity regime, the information rate should plateau as a function of network complexity. We observed that the minimum number of resources increases monotonically in brain networks reparcellated at different resolutions, consistent with a high-fidelity regime (Figure 6B and Supplementary Figure 7). Brain networks with a greater number of parcellated brain regions or modules will require exponentially greater information rates to achieve the same level of distortion as a brain with fewer parcels or modules.
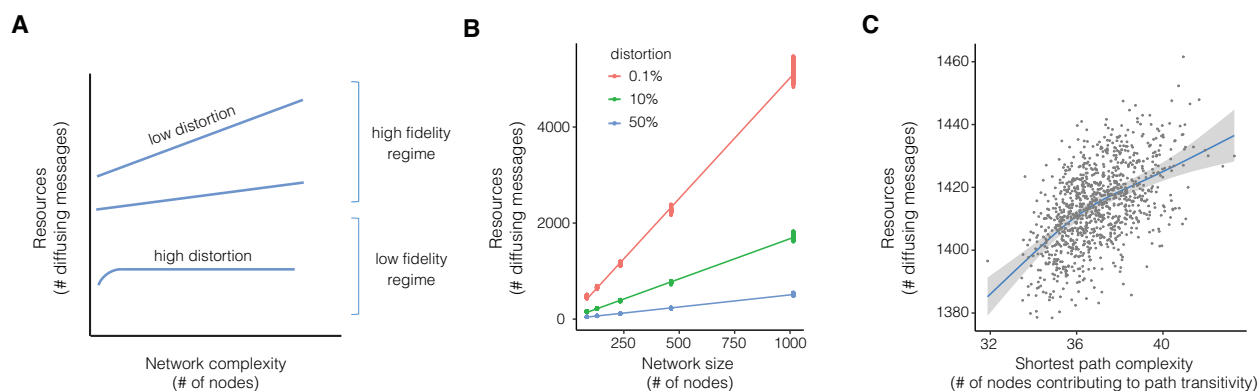

Figure 6: **Compression efficiency constrains brain network complexity. (A)** High versus low fidelity communication regimes make different predictions regarding the minimum resources required for an expected distortion and system complexity. We define system complexity as the network size, or the number of functionally and cytoarchitectonically distinct modules. In the high fidelity regime, the brain should place a premium on high fidelity signaling by allocating a monotonically increasing information rate to achieve low distortion despite increasing system complexity. In the low fidelity regime, the brain should place a premium on lossy compression by asymptotically capping the information rate despite increasing system complexity. **(B)** Resources scale monotonically with network size, consistent with the prediction of a high fidelity regime. Individual brain networks were reparcellated using the Lausanne brain atlases with 83, 129, 234, 463, and 1,015 parcels, respectively. Together, the monotonically increasing resource demand highlights a trade-off between high fidelity communication and network complexity, indicating an additional evolutionary constraint on brain network size and complexity. **(C)** Shortest path complexity is the average number of nodes comprising path transitivity. Resources scale asymptotically with inter-individual differences in path transitivity, consistent with a low-fidelity regime for lossy compression and storage savings. The non-linear model outperforms a linear model (non-linear $AIC = 7902$, linear $AIC = 7915$; non-linear $BIC = 7964$, linear $BIC = 7968$). Lossy compression of a signal by differing densities of path transitivity is perhaps akin to image compression by pixel resolution. Reduced path transitivity, or pixel resolution, tends to result in more compression, while greater path transitivity supports greater fidelity.

In addition to high-fidelity communication, a flexible system of communication may also transfer information in a low-fidelity regime to restrict information rates in noisy environments. We sought to investigate the properties of network architecture that support lossy compression, consistent with predictions of a low fidelity

regime. If the shortest path represents the structure supporting highest fidelity, then we hypothesized that path transitivity (Figure 3A), as longer approximations of shortest paths, supports lossy compression and a low fidelity regime. To remain consistent with the method of existing predictions [27], the complexity of the shortest path was defined as the number of nodes comprising the local detours re-accessing the shortest path in the measure of path transitivity. We found that the number of resources begins to plateau non-linearly as a function of shortest path complexity, consistent with a low-fidelity regime (Figure 6C). Model selection criteria support the non-linear form compared to a linear version of the same model (non-linear $AIC = 7902$, linear $AIC = 7915$; non-linear $BIC = 7964$, linear $BIC = 7968$). The non-linear fit of these data suggest that path transitivity supports neural communication that is tolerant to noise, consistent with the conception of path transitivity as local detours from the highest-fidelity shortest path.

2.4   Neurodevelopment and evolutionary constraints of compression efficiency

Motivated by our findings corroborating the validity of compression efficiency and that neurodevelopment places a premium on fidelity (Figure 4F), we sought to understand the association between compression efficiency and evolutionary properties of cortical organization. Evolutionarily new connections may support higher-order and flexible information processing [4], emerging from disproportionate expansion of the association cortex. The association cortex also contains reduced cortical myelin compared to sensorimotor cortices [34], which promotes efficient transmission and propagation speed while preserving communication fidelity [2]. To explore how compression efficiency relates to cortical areal expansion and myelination, we used published maps of cortical myelination (estimated using published maps of T2/T1w MRI measures with histological validation [34]) and areal scaling (estimated as allometric scaling coefficients defined by the non-linear ratios of surface area change to total brain size change over development; Figure 7B). To study the compression efficiency of brain regions sending or receiving messages, we computed the send and receive compression efficiency of brain regions (Figure 7A; Method 7.5.5). We found that brain regions with greater sender compression efficiency tend to have greater myelin content (Figure 7C; $r = 0.23$, $df = 358$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.04$), consistent with the understanding that myelination enhances the speed and efficiency of neural transmission, as regions with greater compression efficiency require a reduced rate of resources for a given fidelity. Brain regions that disproportionately expand in relation to total brain size during neurodevelopment tend to prioritize input fidelity ($r = -0.20$, $df = 358$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.03$) and transmission compression ($r = 0.14$, $df = 358$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.045$). In contrast, brain regions that are disproportionately out-scaled by total brain expansion may save material, space, and metabolic
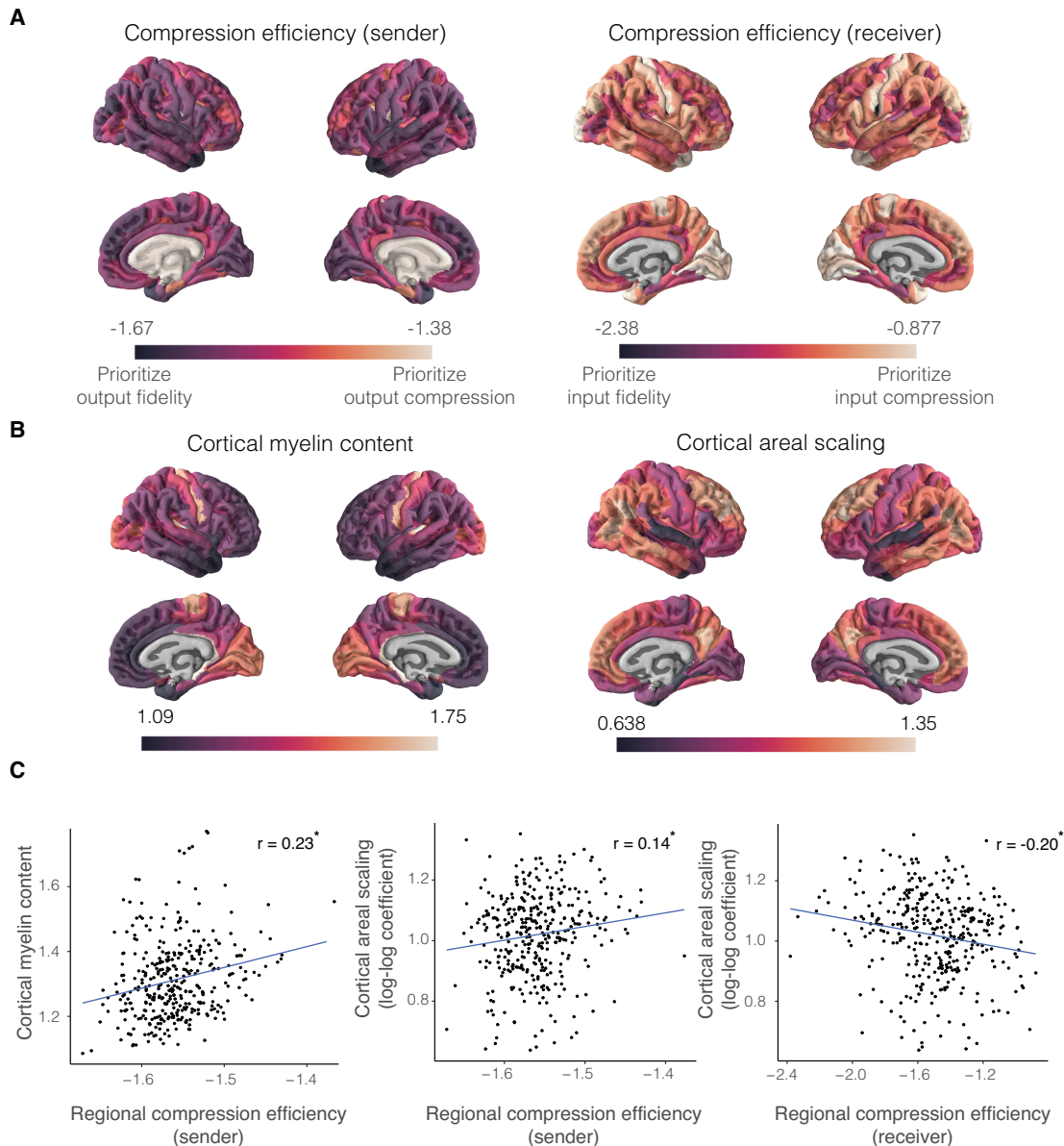
Figure 7: **Adaptive advantages of compression efficiency align with patterns of cortical myelination and areal scaling.** **(A)** Sender compression efficiency differs regionally across the cortex and describes the number of diffusing messages required to transmit information with specified signal fidelity (Supplementary Figure 8). Receiver compression efficiency describes the number of messages required to receive information with an expected signal fidelity. Regional values were averaged across individuals. **(B)** Brain regions differ in myelination and non-linear spatial scaling ratios of surface area change to total brain size change over development. Regional values were obtained from published maps [34, 7]. **(C)** Cortical regions with greater levels of myelin tend to have greater sender compression efficiency ($r = 0.23$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.04$), consistent with myelin reducing conduction delay and promoting the efficient trade-off between signal rate and fidelity to reduce the transmission rate while preserving fidelity. Cortical areal scaling in neurodevelopment reflects patterns of evolutionary remodeling. Brain regions that have higher sender compression efficiency tend to disproportionately expand in relation to total brain size during neurodevelopment ($r = 0.14$, $df = 358$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.045$). We observed that cortical regions with the lowest receiver compression efficiency, placing a premium on information processing fidelity, tend to disproportionately expand in relation to whole brain growth ($r = -0.20$, $df = 358$, $p_{\text{SPIN, Holm-Bonferroni}} = 0.03$). Positively scaling regions that prioritize compression-efficient broadcasting of messages arriving with high fidelity may reflect evolutionary expansion of brain regions with high information processing capacity, whereas negatively scaling regions that prioritize high-fidelity broadcasting of compressed messages may permit other modes of material, spatial, and metabolic cost efficiency.

17

resources by exploiting greater input compression efficiency and high-fidelity transmission of compressed messages.

2.5   Cognitive efficiency and efficient broadcasting of 'rich-club' structural hubs

The fifth and final hypothesis of our model posits that the structural hubs of the brain's highly interconnected rich club supports information integration of diffusing signals [29]. To explain the hypothesized information integration roles of rich-club structural hubs, we investigated the compression efficiency of messages diffusing into and out of hub regions compared to that of other regions. In order to identify the rich-club hubs, we computed the normalized rich-club coefficient and identified 43 highly interconnected structural hubs (Figure 8A). Next, we computed the send and receive compression efficiency of rich-club hubs compared to all other regions. In support of their hypothesized function, we found that the rich-club hubs tend to receive reduced rates of messages compared to other regions (Wilcox rank sum test, $W = 12829, p < 0.001$), suggesting prioritization of information compression (or integration). For the rich-club hub to transmit outgoing messages with a fidelity that is equivalent to the incoming messages, the rich-club hubs tend to transmit greater rates of messages compared to other brain regions ($W = 64, p < 0.001$), supporting the notion that rich-club hubs serve as high-fidelity information broadcasting sources. These contrasting roles of prioritizing input compression and output fidelity within rich-club hubs are consistent with the understanding of rich-club hubs as the information integration centers and broadcasters of the brain's network [29].

Given the distinct integrative and broadcasting role of rich-club hubs compared to other brain regions, we sought to evaluate the association between compression efficiency in rich-club hubs and cognitive performance in a diverse battery of tasks. In light of trade-offs between communication fidelity and information compression, compression efficiency should correlate with cognitive efficiency, defined as the speed-to-accuracy ratio in task performance. Moreover, studying the compression efficiency of communication in rich-club hubs compared to brain-wide structure can elucidate differing roles of network organization relevant to cognition. For example, a greater association of compression efficiency in rich-club hubs with cognitive efficiency, compared with other brain regions, would suggest that hubs are uniquely associated with cognition. In contrast, cognitive efficiency could correlate with compression efficiency in both rich-club hubs and other brain regions, consistent with the importance of diverse connectivity [35]. To assess the relationships between compression efficiency and cognitive efficiency, we used four independent cognitive domains that have been established by
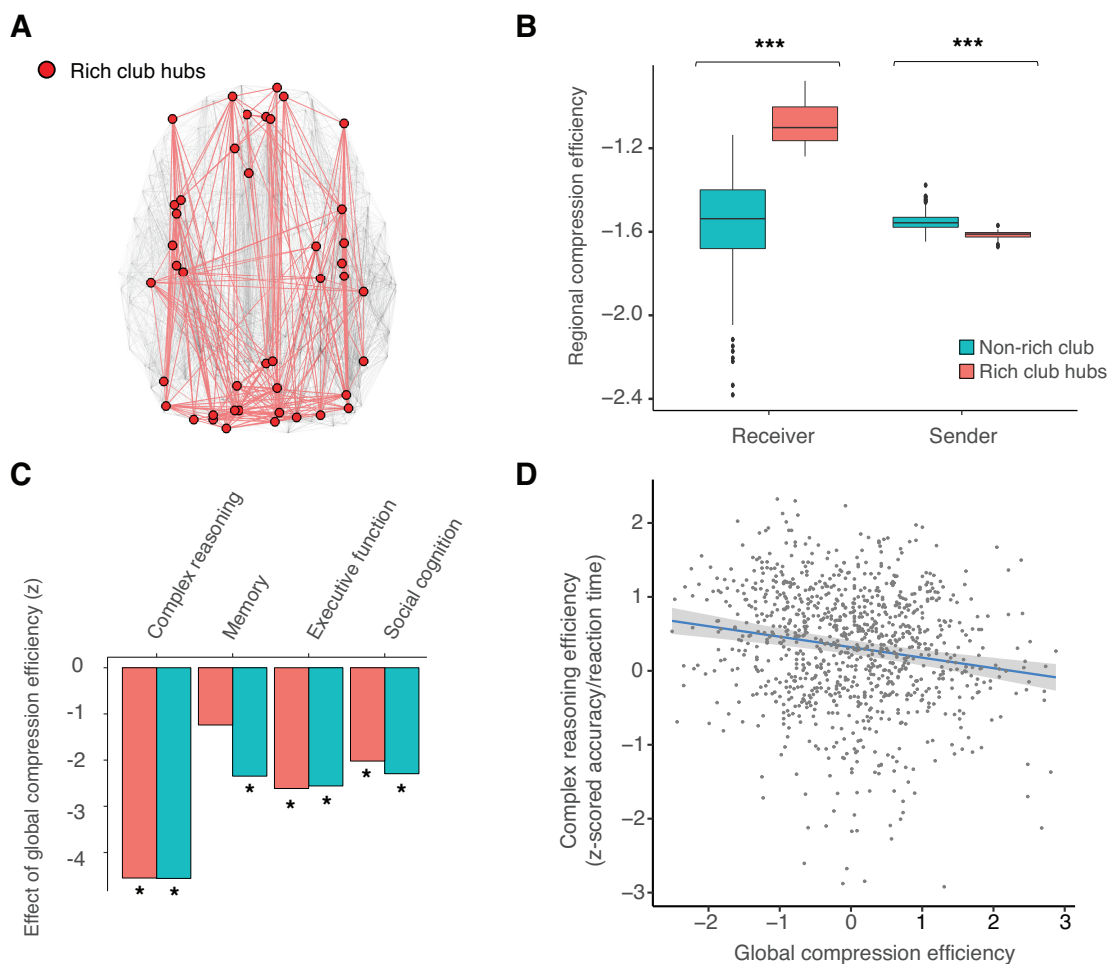
Figure 8: **Compression efficiency explains the integrative role of rich-club hubs and individual differences in cognitive efficiency. (A)** The rich club consists of highly interconnected structural hubs, thought to be the backbone of the brain connectome. Across participants, the rich club was comprised of 43 highly connected brain regions. **(B)** The receiver compression efficiency of rich clubs was greater than other brain regions ($W = 12829, p < 0.001$), suggesting that the organization of the structural network prioritizes integration and compression of information arriving to rich-club hubs. The sender compression efficiency of rich clubs was reduced compared to other brain regions, supporting the fidelity of information broadcasting ($W = 64, p < 0.001$). **(C)** Individuals having non-rich-club regions with decreased compression efficiency tended to exhibit greater efficiency (speed-to-accuracy trade-off) of complex reasoning (all p-values corrected using the Holm-Bonferroni method; $t = -4.72$, estimated model $df = 10.50$, $p = 2 \times 10^{-5}$), executive function ($t = -2.85$, estimated model $df = 10.64$, $p = 0.03$), and social cognition ($t = -2.30$, estimated model $df = 10.45$, $p = 0.04$), controlling for age, sex, age-by-sex interaction, degree, density, and in-scanner motion. Non-rich-club compression efficiency was not associated with memory efficiency ($t = -1.61$, estimated model $df = 9.85$, $p = 0.11$). Individuals with less rich-club compression efficiency tended to exhibit greater efficiency of complex reasoning ($p = 1 \times 10^{-7}$), memory efficiency ($p = 0.03$), social cognition ($p = 0.02$), and executive function ($p = 0.03$). Overall, individuals with brain structural networks prioritizing fidelity tended to perform with greater accuracy and/or speed in various cognitive functions. Asterisks depict significance following family-wise error correction. **(D)** Individual differences in complex reasoning efficiency were negatively associated with individual differences in compression efficiency ($t = -4.95$, estimated model $df = 11.52$, $p < 0.001$), controlling for global efficiency, age, sex, age-by-sex interaction, degree, density, and in-scanner motion. See Supplementary Figure 9 for additional scatterplots of the correlation between cognitive efficiency and compression efficiency. For comparison to a commonly used metric of shortest-path information integration, we include global efficiency in the model, which was positively correlated with complex reasoning efficiency ($t = 2.68$, estimated model $df = 11.52$, $p < 0.01$). Global efficiency was not partially correlated with compression efficiency ($r = -0.04, df = 997, p = 0.18$), controlling for age, sex, age-by-sex interaction, degree, density, and motion. Hence, reduced compression efficiency prioritizing fidelity explains individual differences in complex reasoning efficiency.

19

confirmatory factor analysis to assess individual variation in tasks of complex reasoning, memory, executive function, and social cognition 7.2. We found that the compression efficiency of the rich-club structural hubs was negatively associated with the cognitive efficiency of complex reasoning (all p-values corrected using the Holm-Bonferroni family-wise error method; $t = -4.72$, estimated model $df = 10.59$, $p = 2 \times 10^{-5}$), memory ($t = -2.60$, estimated model $df = 9.85$, $p = 0.04$), executive function ($t = -2.80$, estimated model $df = 10.69$, $p = 0.03$), and social cognition ($t = -2.55$, estimated model $df = 10.47$, $p = 0.04$). In our analyses, we controlled for age, sex, age-by-sex interaction, degree, density, and in-scanner motion. We further found that the compression efficiency of brain regions outside the rich club was negatively associated with the cognitive efficiency of complex reasoning ($t = -4.72$, estimated model $df = 10.50$, $p = 2 \times 10^{-5}$), executive function ($t = -2.85$, estimated model $df = 10.64$, $p = 0.03$), and social cognition ($t = -2.30$, estimated model $df = 10.45$, $p = 0.04$). Individuals with brain structural networks that prioritize fidelity tended to perform with greater cognitive efficiency in a diverse range of functions. We found relationships between cognitive efficiency and compression efficiency in both rich-club hubs and other brain regions, consistent with findings suggesting that diverse network connectivity profiles may play an equally important role in brain network communication as hubs [35]. Importantly, compression efficiency explained variation in cognitive efficiency even when controlling for the commonly used shortest-path measure of global efficiency (Figure 8D; compression efficiency $t = -4.95$, estimated model $df = 11.52$, $p < 0.001$; global efficiency $t = 2.68$, estimated model $df = 11.52$, $p < 0.01$). Taken together, compression efficiency explains the information integration and broadcasting role of the rich-club structural hubs [29], and individual differences in the speed-to-accuracy trade-off of cognitive functions.

3   Discussion

To constrain the expansive theoretical space of communication models, we investigated how principles of evolutionary efficiency constrain models of brain network communication [18, 2, 36, 37, 3, 12, 5, 38, 15, 16]. Specifically, we considered the brain structural connectome as a capacity-limited information channel performing lossy compression. We found metabolic expenditure correlated with structural signatures indicative of diffusion models, but not shortest path routing [3]. In developing an efficient diffusion model of communication, we introduced the notion of compression efficiency, which describes the prioritization of either communication fidelity or lossy compression in structural networks. Five predictions of rate-distortion theory and information diffusion adapted from prior literature corroborated our findings, supporting the validity of

compression efficiency [29, 27, 25]. Broadly, our work advances the study of brain network communication efficiency, information integration, and neural noise by reframing brain network communication as diffusing messages governed by rate-distortion and efficient coding theories.

Shortest path routing as a model of brain network communication was not supported by our data, consistent with the common acknowledgment that it is infeasible to expect a signal to have global knowledge of network structure to compute shortest paths [3, 5, 15, 16]. Rather, our observations agreed with the hypotheses that result from information diffusion along structural paths. We found that individuals whose brains are structured with high-integrity paths tended to have reduced metabolic cost, joining similar prior reports [34]. In our investigation of multiple trade-offs between network structure and metabolic cost, we discovered that brain metabolism reached a critical point as a function of path transitivity and modularity. Specifically, we observed a saddle point, where network structure was coupled or decoupled. In the decoupled axis, brain networks organized with, for example, high modularity and low path transitivity tended to exhibit optimal metabolic savings. In the coupled axis where increases in modularity are linked with increases in path transitivity, brain networks can achieve locally optimal metabolic savings around the average brain network, where deviations incur metabolic costs. Brain networks may reconfigure to place premiums on network structures thought to support functional versatility and resilience at the expense of cost efficiency, or *vice versa* [3, 16].

Turning from shortest-path based measures, such as global efficiency and betweenness centrality, our findings motivate future studies of information integration in structural brain networks that instead adopt metrics of the structural signatures and processes of information diffusion that emphasize neuroanatomically specific processes [39, 15, 16]. For example, we found that chemotactic diffusion along metabolic gradients supports efficient coding by enhancing compression efficiency, offering a potential biological medium for greedy navigation by shortest spatial distances [16]. Chemotactic attraction models the increased neural activity associated with metabolic expenditure [31], whereas repulsion models information bottlenecks redirecting flow away from congestion to less metabolically costly routes [14]. Furthermore, we found that compression efficiency was associated with cognitive efficiency above and beyond the contributions of global efficiency; the latter having been previously reported to explain variation in fluid intelligence [30]. Global efficiency remains a useful metric of local connection strength, pairwise wiring cost trade-offs, and shortest path structure accelerating diffusion [36, 3, 14], but falls short of explaining processes of information integration.

We offer an explanation of integrative processes arising from the connectivity of rich-club hubs. By measuring asymmetric send-receive message diffusion [40, 41] and modeling transmission rate as a function of expected fidelity, we showed that hubs are compression-efficient receivers and high-fidelity senders. This finding adds to the understanding of hubs as sources and sinks for the early spreading of diffusing signals [14]. Rich-club hubs develop early and underpin information integration and broadcasting, possibly offsetting high metabolic, spatial, and material costs [29, 42, 9, 41]. However, an adaptationist explanation as such is challenging to falsify [38]. We introduced compression efficiency in the context of efficient coding to reframe explanations of evolutionary adaptation in terms of a heritable capacity to develop hubs under constraints of whole-brain network efficiency [38, 41]. Prior findings of greater metabolic costs in rich-club regions were exploratory [29] and conflict with research supporting the metabolic efficiency of the myelinated long-distance connections prevalent in the rich-club [3, 34, 6]. Despite our well-powered analysis and replication of several other findings (Supplementary Figures 2, 6, and 10) [24, 32, 13, 22], we were unable to replicate observations of high metabolic costs in the rich-club (Supplementary Figure 11). Although further investigation of the metabolic costs of rich-club hubs is warranted, our findings nevertheless reinforce a wealth of evidence emphasizing the importance of the development, resilience, and function of hubs in cognition and psychopathology [29, 9, 31, 14, 6, 10].

With the objective of efficiency, developmental processes may balance compression efficiency, cortical scaling, and myelination to adapt to differing environments. Cortical areal scaling highlights the problem of allocating limited materials, space, and metabolic resources to the disproportionate changes in surface area of brain regions in relation to total brain size [7]. Brain regions that prioritize high-fidelity broadcasting of compressed messages may save space, materials, and metabolic resources with decreased scaling in proportion to the growth of the whole brain. For example, we found this property of compression-efficient inputs and high-fidelity outputs in rich-club hubs, which appear in an adult configuration at birth [43]. In contrast, brain regions prioritizing compression-efficient broadcasting of high-fidelity messages tended to disproportionately expand, and brain networks prioritizing communication fidelity tended to support greater cognitive efficiency. These novel findings converge with theories positing that evolutionarily new connections support higher-order and flexible information processing [4, 34, 7], and that plastic white matter microarchitecture supports reasoning ability and speed [44]. Indeed, we found monotonically increasing information processing costs and capacity with greater network complexity. In addition to spatial scaling, developing brain networks may use myelination to modify connection strengths and efficiency [45, 34]. We found that brain regions

with greater myelination tended to have greater sender compression efficiency, consistent with evidence that myelin promotes propagation speed and efficient transmission rates while preserving communication fidelity [2, 34]. The objective of efficient coding in brain networks can be achieved by balancing communication fidelity and lossy compression in developmentally plastic brain networks and rich-club hubs [37, 32, 42, 6, 46].

We suggest that compression efficiency may represent an information processing constraint on brain size and complexity. Brain systems viewed as information processors exhibit recurring compromises between information efficiency and other resource costs at the cellular [18, 2] and circuit levels of the brain [47, 19]. At the neuronal level, an optimal strategy for distributed coding is to reduce population size while distributing activity among a fraction of cells [18, 2]. Brain networks may reach a similar compromise through information processing constraints on complexity (i.e. size) of the network and its modules, and increasing the number of endogenously active components, such as in the default-mode system. Efficient coding predicts that bit rate varies as a function of the number and redundancy of synapses [2, 27]. Transmitting the same message across many parallel paths improves fidelity and increases bit rates, but information rate increases sublinearly with the number of paths because the system is highly redundant, incurring greater metabolic costs [2]. We similarly found that individuals with greater path transitivity—more redundant and lossy alternative paths to the shortest, direct paths—tended to require sublinearly increasing computational costs and tended to have greater global metabolic expenditure. Taken together, our efficient diffusion model addresses the notable absence of biologically plausible and efficient inter-regional brain network communication models [3, 16].

Our work admits several theoretical and methodological limitations. First, regionally aggregated brain signals are not discrete Markovian messages and do not have goals like reaching specific targets. As in recent work, our model introduced a deliberately simplified but useful abstraction of macro-scale brain network communication [14]. Second, although we used resource efficiency in light of prior methodological decisions and information theory benchmarks [12], compression efficiency can be implemented using alternative approaches (Supplementary Figure 12). Several methodological limitations should also be considered. The accurate reconstruction of white-matter pathways using DTI and tractography remains limited [48]. Moreover, non-invasive measurements of CBF with high sensitivity and spatial resolution remain challenging. We acquired images using an ASL sequence providing greater sensitivity and approximately four times higher spatial resolution than prior developmental studies of CBF [32]. Lastly, our data was cross-sectional, limiting the inferences that we could draw about neurodevelopmental processes.

In summary, our study advances understanding of the adaptive trade-offs in brain metabolism and architecture that support efficient diffusion processes. In addition to advancing the biophysical realism of information transmission, our information-theoretical model naturally admits future applications to measurements of entropy, which have provided insight into information flow of brain activity [47, 39]. Our model may be applied to study neural circuits of Bayesian integration in brain networks, as rate-distortion models of perception and cognition have been suggested as extensions of conventional Bayesian approaches [25]. Moreover, our work distinguishing a low- and high-fidelity regime suggests our framework could be used to investigate dual-system models of information processing bounded by resource and capacity limitations that characterize fast but error-prone versus slow but deliberate regimes [49]. In the complementary learning systems theory, we posit that the hippocampus acts as a hub in plastic cortical networks which pass, distort, and reconstruct compressed signals [50]. Compression efficiency of hippocampal and sensory pathways should predict the speed, accuracy, and efficient cognitive coding of high-dimensional visuospatial stimuli in sensorimotor learning [25, 19]. Such studies could illuminate how the representational structure of information drives the selective loss of redundant or core information in convolutional feedforward network models of sensorimotor information processing where triangular structural motifs of path transitivity resemble feedforward loops [13, 19]. Lastly, our findings invite further development and application of well-studied information routing models and coding schemes to brain network communication [26, 51, 40]. The compression efficiency model is a useful starting point for the development of more sophisticated approaches of efficient systems-level information transfer, and is also a novel tool to test leading hypotheses of dysconnectivity [8], hubopathy [9, 10], disrupted information integration [52], and neural noise [53] in neuropsychiatric disorders.

4   Author Contributions

5   Acknowledgments

6   Competing Interests

The authors declare that they have no competing interests.

7   Materials and Methods

7.1   Participants

As described in detail elsewhere [23], diffusion tensor imaging (DTI) and arterial-spin labeling (ASL) data were acquired for the Philadelphia Neurodevelopmental Cohort (PNC), a large community-based study of neurodevelopment. The subjects used in this paper are a subset of the 1,601 subjects who completed the cross-sectional imaging protocol. We excluded participants with health-related exclusionary criteria (n=154) and with scans that failed a rigorous quality assurance protocol for DTI (n=162) [54]. We further excluded subjects with incomplete or poor ASL and field map scans (n=60). Finally, participants with poor quality T1-weighted anatomical reconstructions (n=10) were removed from the sample. The final sample contained 1042 subjects (mean age=15.35, SD=3.38 years; 467 males, 575 females). Study procedures were approved by the Institutional Review Board of the Children's Hospital of Philadelphia and the University of Pennsylvania. All adult participants provided informed consent; all minors provided assent and their parent or guardian provided informed consent.

7.2   Cognitive Assessment

All participants were asked to complete the Penn Computerized Neurocognitive Battery (CNB). The battery consists of 14 tests adapted from tasks typically applied in functional neuroimaging, and which measure cognitive performance in four broad domains [23]. The domains included: (1) executive control (i.e., abstraction and flexibility, attention, and working memory), (2) episodic memory (i.e., verbal, facial, and spatial), (3) complex cognition (i.e., verbal reasoning, nonverbal reasoning, and spatial processing), (4) social cognition (i.e., emotion identification, emotion intensity differentiation, and age differentiation), and (5) sensorimotor and motor speed. Performance was operationalized as z-transformed accuracy and speed. The speed scores were multiplied by -1 so that higher indicates faster performance, and efficiency scores were calculated as the mean of these accuracy and speed z-scores. The efficiency scores were then z-transformed again, to achieve mean $= 0$ and SD $= 1.0$ for all scores. Confirmatory factor analysis supported a model of four latent factors corresponding to the cognitive efficiency of executive function, episodic memory, complex cognition, and social cognition [55]. Hence, we used these four cognitive efficiency factors in our analyses.

7.3   Image Acquisition, Preprocessing, and Network Construction

Neuroimaging acquisition and pre-processing were as previously described [23]. We depict the overall workflow of the neuroimaging and network extraction pipeline in Figure 1A.

7.3.1   Diffusion Tensor Imaging

As was previously described [22, 56], DTI data and all other MRI data were acquired on the same 3T Siemens Tim Trio whole-body scanner and 32-channel head coil at the Hospital of the University of Pennsylvania. DTI scans were obtained using a twice-focused spin-echo (TRSE) single-shot EPI sequence (TR $= 8100$ ms, TE $= 82$ ms, FOV $= 240$ mm^2/240 mm^2; Matrix $=$ RL: 128/AP:128/Slices:70, in-plane resolution (x & y) 1.875 mm^2; slice thickness $= 2$ mm, gap $= 0$; FlipAngle $= 90°/180°/180°$, volumes $= 71$, GRAPPA factor $= 3$, bandwidth $= 2170$ Hz/pixel, PE direction $=$ AP). The sequence employs a four-lobed diffusion encoding gradient scheme combined with a 90-180-180 spin-echo sequence designed to minimize eddy current artifacts. The complete sequence consisted of 64 diffusion-weighted directions with $b = 1000$ s/mm^2 and 7 interspersed scans where $b = 0$ s/mm^2. Scan time was about 11 min. The imaging volume was prescribed in axial orientation covering the entire cerebrum with the topmost slice just superior to the apex of the brain [54].

7.3.2 Connectome construction

Cortical gray matter was parcellated according to the Glasser atlas [57], defining 360 brain regions as nodes for each subject's structural brain network, denoted as the weighted adjacency matrix $\mathbf{A}$. To assess multiple spatial scales, cortical and subcortical gray matter was parcellated according to the Lausanne atlas [58]. Together, 89, 129, 234, 463, and 1015 dilated brain regions defined the nodes for each subject's structural brain network in the analyses of Figure 6.

DTI data was imported into DSI Studio software and the diffusion tensor was estimated at each voxel [59]. For deterministic tractography, whole-brain fiber tracking was implemented for each subject in DSI Studio using a modified fiber assessment by continuous tracking (FACT) algorithm with Euler interpolation, initiating 1,000,000 streamlines after removing all streamlines with length less than 10mm or greater than 400mm. Fiber tracking was performed with an angular threshold of 45, a step size of 0.9375mm, and a fractional anisotropy (FA) threshold determined empirically by Otzu's method, which optimizes the contrast between foreground and background [59]. FA was calculated along the path of each reconstructed streamline. For each subject, edges of the structural network were defined where at least one streamline connected a pair of nodes. Edge weights were defined by the average FA along streamlines connecting any pair of nodes.

7.3.3 Arterial-Spin Labeling

CBF was quantified from control-label pairs using ASLtbx [60], as was previously described [32]. We consider f as CBF, δM as the difference of the signal between the control and label acquisitions, R_{1a} as the longitudinal relaxation rate of blood, τ as the labeling time, ω as the post-labeling delay time, α as the labeling efficiency, λ as the blood/tissue water partition coefficient, and M_0 as the approximated control image intensity. Together, CBF f can be calculated according to the equation:

$$f = \frac{\Delta M \lambda R_{1a} \exp\left(\omega R_{1a}\right)}{2 M_0 \alpha} \left[1 - \exp\left(-\tau R_{1a}\right)\right]^{-1}. \tag{1}$$

Because prior work has shown that the T1 relaxation time changes substantially in development and varies by sex, this parameter was set according to previously established methods, which enhance CBF estimation accuracy and reliability in pediatric populations [61, 62].

7.4 Brain Maps

7.4.1 Cortical Myelin

As described previously [63], cortical myelin content was calculated by dividing the T1w image signal by the T2w image signal. Specifically, we define the myelin content x^2 in the following manner:

$$\frac{\text{T1w}}{\text{T2w}} \approx \frac{x * b}{(1/x) * b} = x^2, \tag{2}$$

where x is the myelin contrast in the T1w image, $1/x$ is the myelin contrast in the T2w, and b is the receive bias field in both T1w and T2w images. We used a published atlas generated by this method [34].

7.4.2 Cortical Areal Scaling

As described previously [7], to estimate cortical areal scaling between the size of cortical regions and the total brain, regression coefficients β were estimated for $\log_{10}$(total cortical surface area) as a covariate predicting $\log_{10}$(vertex area) using spline regression models that incorporated effects of age and sex on vertex area [64]. We used the following relational form:

$$\log 10(\text{Vertex area}) \sim \text{s(age by } = \text{sex}) + \text{B}_{\text{I}}\left[\log 10\left(\text{total\_area}\right)\right]. \tag{3}$$

When β is 1, the scaling between total brain size and brain regions is linear. When β deviates greater or less than 1, scaling is non-linearly and disproportionately expanding or contracting. We used the published atlas generated using the same data as in our study [7, 23].

7.5 Network Statistics

7.5.1 Global Efficiency

In the context of the brain structural connectome, global efficiency represents the strength of the shortest paths between brain regions supporting efficient communication. In network neuroscience, global efficiency is commonly used as a metric of a brain network's capacity for shortest path routing [3, 12, 16]. We calculated the common global efficiency statistic [33], which is defined for a graph G as:

$$\mathcal{E}_{glob}(G) = \frac{1}{\mathcal{N}(\mathcal{N}-1)} \sum_{i \neq j \epsilon G} \frac{1}{d_{ij}}, \tag{4}$$

where $\mathcal{N}$ is the number of nodes and d_{ij} is the shortest distance between node i and node j. Intuitively, a high $\mathcal{E}$ value indicates greater potential capacity for global and parallel information exchange along shortest paths, and a low $\mathcal{E}$ value indicates decreased capacity for such information exchange [33].

7.5.2 Path Strengths

Beyond shortest paths between pairs of brain regions, we also sought to measure the strength of structural connections S comprising the paths of multiple connections. As global efficiency measures the capacity of brain networks for shortest path routing, path strengths measure the capacity for diffusion signaling. Path strengths are apt for assessing the network capacity for diffusion because paths can be represented as random walks $p = (i, j, \ldots, k)$, where p is a path and i, j, and k are nodes in the path. As in prior work [65], the strength of the weighted connections in a path, denoted $\omega(p)$, in the graph G with adjacency matrix $\mathbf{A}$ is defined as:

$$\omega(p) = [\mathbf{A}]_{i_0 i_1} [\mathbf{A}]_{i_1 i_2} \ldots [\mathbf{A}]_{i_{(l-1)} i_i}, \tag{5}$$

where the matrix products produce the strengths of all possible random walks according to the length of p, as depicted in the schematic Figure 1B. Then, for walks of length n, the strengths of the paths from node i to node j are defined as:

$$[\mathbf{A}^n]_{ij} = \sum_{p \in \mathcal{P}_{ij}^n} \omega(p), \tag{6}$$

where P_{ij}^n is the set of all walks from node i to node j with length n. When $n=1$, the matrix exponent produces a matrix with elements equal to d_{ij} from Equation 1, or the shortest distance between node i and node j. Intuitively, a high path strength represents structural paths that consist of higher integrity connections measured by DTI, whereas a low path strength indicates paths consisting of low integrity connections. To compute node strengths, the values for each node were summed. An average value was also calculated across node strengths per individual participant.

7.5.3 Path Transitivity

Shortest paths confer advantages in speed and signal fidelity when messages are transmitted by diffusion. Therefore, we sought to measure a property of brain network architecture supporting diffusion by shortest paths. Local detours which first leave and then re-access the shortest path serve to support such diffusion, and the potential for such local detours can be estimated using a measure called path transitivity (see Figure 3A, left) [13]. Path transitivity was previously used to predict functional BOLD activation comparably to

conventional distance or computational models of neural dynamics. To compute path transitivity, we first calculated the matching index for each pair of successive nodes i and j along the shortest path $\pi_{s \to t}$, with neighboring non-shortest path nodes k as:

$$m_{ij} = \frac{\sum_{k \neq i,j} (w_{ik} + w_{jk}) \Theta(w_{ik}) \Theta(w_{jk})}{\sum_{k \neq j} w_{ik} + \sum_{k \neq i} w_{jk}}, \tag{7}$$

where w is the connection weight, and $\Theta(w_{ik}) = 1$ if $w_{ik} > 0$, and 0 otherwise. Intuitively, the numerator is non-zero if and only if there are two locally detouring connections that make a closed triangle along the shortest path. If either of the two connections w_{ik} or w_{jk} does not exist, then the numerator is 0. With the denominator representing the strength of all cumulative connections of the shortest path nodes, the matching index fraction then represents the density of closed triangles (i.e., transitivity) around the shortest path.

Whereas the matching index is a pairwise measure of the density of locally returning detours, path transitivity generalizes the density across the shortest path. Using the computed matching index m_{ij} for each pairwise connection Ω from source node s to target node t by the set of shortest path edges $\pi_{s \to t}$, we compute path transitivity M as:

$$M(\pi_{s \to t}) = \frac{2 \sum_{i \in \Omega} \sum_{j \in \Omega} m_{ij}}{|\Omega|(|\Omega| - 1)}, \tag{8}$$

where the numerator sums the matching index m_{ij} for all edges in Ω, the scale factor of 2 indicates an undirected graph, and the denominator sums over all possible edges. Intuitively, a high path transitivity M indicates that the shortest path is more densely encompassed by locally detouring triangular motifs. Low path transitivity indicates that the shortest path is surrounded by connections that deviate from the shortest path without an immediate avenue of return.

7.5.4 Modularity

Modularity is a common architectural feature observed in neural systems across species. A single community contains brain regions that are more highly connected to each other than to brain regions located in other communities (see Figure 3A, right). Modularity of brain networks is spatially efficient, supports the development of executive function in youths, and supports flexibly adaptable functional activations according to distinct task demands [66, 67, 22, 68, 69]. To assess modularity, we apply a common community detection technique known as modularity maximization [70], in which we used a Louvain-like locally greedy algorithm [71] to maximize a modularity quality function for the adjacency matrix A. The modularity quality function

30

is defined as:

$$Q = \frac{1}{2\mu} \sum_{ij} \left(A_{ij} - \gamma P_{ij} \right) \delta(g_i, g_j), \tag{9}$$

where $\mu = \frac{1}{2}\sum_{ij} A_{ij}$ denotes the total weight of $\mathbf{A}$, A_{ij} encodes the weight of an *edge* between node i and node j in the structural connectivity matrix, $\mathbf{P}$ represents the expected strength of connections according to a specified null model [70], γ is a structural resolution parameter that determines the size of modules, and δ is the Kronecker function which is 1 if $g_i = g_j$ and zero otherwise. As in prior work, we set γ to the default value of 1 [68]. Intuitively, a high Q value indicates that the structural connectivity matrix contains communities, where nodes within a community are more densely connected to one another than expected under a null model. Modularity maximization is commonly used to detect community structure, and to quantitatively characterize that structure by assessing the strength and number of communities [22, 68, 70].

7.5.5 Resource Efficiency

A signal that diffuses along the shortest path between brain regions confers advantages in speed, reliability, and fidelity [3, 72, 16]. Following prior work, we sought to compute the number of random walkers beginning at node i that were required for at least one to travel along the shortest path to another node j with probability η [12, 72]. To begin, we consider the transition probability matrix by $\mathbf{U}$, defined as $\mathbf{U} = \mathbf{W}\mathbf{L}^{-1}$, where each entry W_{ij} of $\mathbf{W}$ describes the weight of the directed edge from node i to node j, and each entry L_{ii} of the diagonal matrix $\mathbf{L}$ is the strength of each node i, defined as $\sum_i W_{ij}$. Intuitively, each entry U_{ij} of $\mathbf{U}$ defines the probability of a random walker traveling from node i to node j in one step. Next, to compute the probability that a random walker travels from node i to node j along the shortest path, we define a new matrix $U'(i)$ that is equivalent to $\mathbf{U}$ but with the non-diagonal elements of row i set to zero and $U_{ii} = 1$ as an absorbant state. Then, the probability of randomly walking from i to j along the shortest path is given by:

$$1 - \sum_{n=1}^{N} [U'(i)^H]_{in}, \tag{10}$$

where H is the number of connections composing the shortest path from i to j. Similarly, the probability η of releasing r random walkers at node i and having at least one of them reach node j along the shortest path is given by:

$$\eta = 1 - \left(\sum_{n=1}^{N} [U'(i)^H]_{in} \right)^r. \tag{11}$$

Setting the above probability to some set value η, we can then solve for the number of random walkers r required to guarantee (with probability η) that at least one of them travels from i to j along the shortest path, denoted by:

$$r_{ij}(\eta) = \frac{\log(1 - \eta)}{\log \left(\sum_{n=1}^{N} [U'(i)^H]_{in} \right)}. \tag{12}$$

We refer to the number of random walkers r_{ij} as resources. In our analyses, we calculate resources r_{ij} over a range of values of η for each participant. Finally, to calculate the resource efficiency of each participant, the resource efficiency of an entire network is taken to be $1/(r_{ij}(\eta))$ averaged over all pairs of nodes i and j. With the right stochastic matrix U_i', the resource efficiency of brain regions as message senders is $1/(r_{ij}(\eta))$ averaged over i, while brain regions as message receivers is $1/(r_{ji}(\eta))$ averaged over j.

7.5.6   Compression Efficiency

Rate-distortion theory formalizes the study of information transfer as passing signals (messages) through a capacity-limited information channel. A signal x is encoded as $\hat{x}$ with a level of distortion D that depends on the information rate R. The greater the rate, the less the distortion. The rate-distortion function $R(D)$ defines the minimum information rate required to transmit a signal corresponding to a level of signal distortion (see Figure 4A). Lossy compression arises from the choice of the distortion function $d(x, \hat{x})$, which implicitly determines the relevant and irrelevant features of a signal. With the true signal x mapped to the compressed signal $\hat{x}$ described by $p(\hat{x}|x)$, the rate-distortion function is defined by minimizing the mutual information of the signal and compression over the expected distortion defined as $d(x, \hat{x})_{p(x, \hat{x})} = \sum_{x \in X} \sum_{\hat{x} \in \hat{X}} p(x, \hat{x}) d(x, \hat{x})$:

$$R(D) \equiv \min_{d(x, \hat{x})} I(X, \hat{X}) = \sum_{x \in \Omega_X} \sum_{\hat{x} \in \Omega_{\hat{X}}} P(\hat{x}|x) P(x) \log_2 \left(\frac{P(\hat{x}|x)}{P(\hat{x})} \right). \tag{13}$$

By minimizing the mutual information $I(X, \hat{X})$, we arrive at a probabilistic map from the signal to the compressed representation, where the information gain between the signal and compression is as small as possible (i.e., high fidelity) to favor the most compact representations.

Similar to the mathematical framework of rate-distortion theory, we sought to specify a distortion function reflecting communication over the brain's structural network. Prior work building models of perceptual and cognitive performance have inferred distortion functions through Bayesian inference of a loss function [73, 25]. For instance, the loss function could be the squared error denoting the residual values of the true signal minus the compression, $L = (\hat{x} - x)^2$ (Figure 4A). A neural rate-distortion theory has been theoretically developed [27], but remains empirically untested due in part to a lack of methodological tools at the level of brain systems. Moreover, it has been difficult to define a distortion function that incorporates both true signals x and compressed signals $\hat{x}$ in part because the measurements of these signals in human brain networks remains challenging. Here, we define an analogous framework of information transfer through capacity-limited channels in the structural network of the brain. Particularly, we build a distortion function from the simple intuition that the shortest path is the route that most reliably preserves signal fidelity, as depicted in (Figure 4B).

Given that a random walker propagating from node i along the shortest path to node j retains the greatest signal fidelity, we define the distortion function of any signal x from brain region i to a compressed representation $\hat{x}$ decoded in brain region j as:

$$d(x, \hat{x})_{ij} = (1 - \eta), \tag{14}$$

where η denotes the probability that a walker gets from node i to node j along the shortest path. A signal with greater probability η of propagating by the shortest path between brain region i and brain region j is at a lower risk of distortion (see Figure 4D). Intuitively, increased topological distance adds greater risk of signal distortion due to further transmission through capacity-limited channels (i.e. structural connections), temporal delay, and potential mixing with other signals. Given the measure of resources in Equation 12, we develop and test predictions of a novel definition of the rate $R(D)$; here, we define $R(D)$ as the resources $r_{ij}(\eta)$ required to achieve a tolerated level of distortion $d(x, \hat{x})_{ij}$:

$$R(D) \equiv r_{ij}(d(x, \hat{x})_{ij}), \tag{15}$$

as in (Figure 4D). When the log of resources $\log(r_{ij})$ is plotted against our metric of distortion $D = d \in 1 - \eta_{ij}$, the exponential gradient is depicted linearly (see Figure 4E). Because prior work focused on 50% distortion during analyses, we required the slope to intersect the mean midpoint rate at 50% distortion [12]. In addition

33

to the precedent offered by prior work, this requirement is also reasonable given that we sought to model both high and low distortions equitably. The slope denotes the minimum number of resources required to achieve a tolerated level of distortion, which we refer to as the *compression efficiency* (4E; bottom). A steeper slope (i.e., a more negative relation) reflects reduced compression efficiency, or prioritization of message fidelity. A flatter slope (i.e., a more positive relation) reflects increased compression efficiency, or prioritization of lossy compression. Individual variation in compression efficiency can be assessed by using the average resource efficiency across brain regions. When compression efficiency is computed for sets of brain regions by averaging across individuals, the slope can denote either messages sent from or arriving to a brain region by using the average resource efficiency over either all nodes j or all nodes i, respectively.

7.5.7 Biased Random Walk

Given the advantages of shortest path diffusion, we sought to assess how brain metabolism could support the reliability and fidelity of signaling. Chemotactic diffusion can be modeled as random walks over a structural connectivity matrix biased by regional CBF [74]. To model chemotactic diffusion of random walkers attracted to or repelled from brain regions of high CBF, we used analytical solutions to biased random walks. First, we defined the matrix $\mathbf{T}$ of CBF-biased transition probabilities as:

$$T_{ij}^{\alpha} = \frac{\alpha_i A_{ij}}{\sum_k \alpha_k A_{kj}}, \tag{16}$$

where the element of T_{ij} defines the transition probabilities of a random walker traversing edges of the structural connectivity matrix $\mathbf{A}$ which are multiplied by a bias term α. For random walkers attracted to brain regions of high CBF, the bias term α was defined as the average CBF value for each pair of brain regions. For random walkers repelled by regions of high CBF, the bias term α was defined as 1 minus the average CBF value for each pair of brain regions. Hence, a random walker propagates over the brain's structural connections with transition probabilities of T_{ij} that reflect the integrity of structural connections and the average level of CBF between pairs of brain regions. We then substituted the $\mathbf{U}$ matrix in the resources $r_{ij}(\eta)$ of Equation 12 with T_{ij} in Equation 16 to compute the number of resources required for a biased random walker to propagate by the shortest path with a specified probability.

7.5.8 Rich Club

Due to the importance of brain network hubs in the broadcasting of a signal [3, 75, 16], we sought to identify the set of high-degree brain regions in the rich club (see Figure 8A) [76]. To identify the subnetwork of rich club brain regions, we computed the weighted rich club coefficient $\Phi^z(k)$ as:

$$\phi^z(k) = \frac{Z_{>k}}{\sum_{l=1}^{E>k} z_l^{\,\text{ranked}}}, \tag{17}$$

where Z^{ranked} is a vector of ranked network weights, k is the degree, $Z_{>k}$ is the set of edges connecting the group of nodes with degree greater than k, and $E_{>k}$ is the number of edges connecting the group of nodes with degree greater than k. Hence, the rich-club coefficient $\Phi^z(k)$ is the ratio between the set of edge weights connected to nodes with degree greater than k and the strongest $E_{>k}$ connections. The rich-club coefficient was normalized by comparison to the rich club coefficient of random networks [76]. Random networks were created by rewiring the edges of each individual's brain network while preserving the degree distribution. The rich-club coefficient for the randomized networks $\Phi_{\text{random}}(k)$ was computed using Equation 17. Then, the normalized rich-club coefficient $\Phi_{\text{norm}}(k)$ was calculated as follows:

$$\phi_{\text{norm}}(k) = \frac{\phi(k)}{\phi_{\text{random}}(k)}, \tag{18}$$

where $\Phi_{\text{norm}}(k) > 1$ indicates the presence of a rich club organization. We tested the statistical significance of $\Phi_{\text{norm}}(k)$ using a 1-sample t-test at each level of k, with family-wise error correction for multiple tests over k. Each individual was assigned the value of their highest degree $> k$ rich club level and their nodes were ranked by rich club level. Over the group of individuals, the nodal ranks were averaged and the top 12% of nodes were selected as the rich club, following prior work [77].

7.6 Network Null Models

Random graphs are commonly used in network science to test the statistical significance of the role of some network topology against null models. We used randomly rewired graphs generated by shuffling each individual's empirical networks 20 times, as in prior work [78]. Furthermore, we generated Erdös-Renyí random networks for each individual brain network where the presence or absence of an edge was generated by a uniform probability calculated as the density of edges existing in the corresponding brain network. Edge weights were randomly sampled from the edge weight distribution of the brain network. While the randomly

rewired graphs retain empirical properties such as the degree and edge weight distributions of the individual brain networks, the Erdös-Renyí networks do not. Hence, the randomly rewired null network was used in all analyses where the degree distribution should be retained (e.g., normalized rich club coefficient), while the Erdös-Renyí network was used in analyses assessing the overall contribution of the brain network topology (e.g., compression efficiency).

Our tests using the randomly rewired network evaluate the null hypothesis that an apparent rich-club property of brain networks is a trivial result of topology characteristic of random networks with some empirical properties preserved, as in prior work [76, 75]. The alternative hypothesis is that the brain network has a rich-club organization beyond the level expected in the random networks. Our tests using the Erdös-Renyí network evaluate the null hypothesis that the rates in the rate-distortion function modeling information processing capacity in brain networks does not differ from the rates in the rate-distortion function of random networks. The alternative hypothesis is that the rate of the brain network's rate-distortion function differs from that of random networks, consistent with the notion that Erdös-Renyí networks have a greater prevalence of shortest paths compared to brain networks. We additionally used the Erdös-Renyí network to assess the hypothesis of rate-distortion theory that synthetic networks should exhibit the same information processing trade-offs (the monotonic rate-distortion gradient) as empirical brain networks [25]. We selected Erdös-Renyí networks to assess these hypotheses for two reasons. First, Erdös-Renyí networks do not retain core architectures of brain networks, such as modularity, and therefore reflect an extreme synthetic network. Second, Erdös-Renyí networks are commonly used as a benchmark for assessing shortest path prevalence due to the prominence of uniformly distributed direct pairwise connections [39, 16]. In light of the central assumption that shortest paths represent the route of highest signal fidelity in our definition of distortion, we used Erdös-Renyí networks to verify our intuition that compression efficiency should be greater in the Erdös-Renyí network than in brain networks.

7.7   Statistical Analyses

To assess the covariation of our measurements across individuals and brain regions, we used generalized additive models (GAMs) with penalized splines. GAMs allow for statistically rigorous modeling of linear and non-linear effects while minimizing over-fitting [64]. Throughout, the potential for confounding effects was addressed in our model by including covariates for age, sex, age-by-sex interaction, network degree, network density, and in-scanner motion.

7.7.1 Metabolic running costs associated with brain network architectures

We used penalized splines to estimate the nonlinear developmental patterns of global efficiency (Equation 4) and CBF, as in prior work [32, 22]. Then, we assessed the partial correlation between the residual variance (unexplained by covariates of age, sex, age-by-sex-interaction, degree, density, and motion) of global efficiency and CBF. The final models can be written as:

$$\text{Global efficiency} \sim \text{spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion}, \quad (19)$$

$$\text{CBF} \sim \text{spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion}, \quad (20)$$

and

$$\text{Residual(Global efficiency)} \sim \text{residual(CBF)}. \quad (21)$$

To evaluate the importance of age as a confound for the relationship between global efficiency and CBF, we also performed sensitivity analyses by removing selected covariates and re-assessing the model. In addition, for consistency with prior work [28], we performed the same analysis including covariates for gray matter volume and density.

To assess the relationship between CBF and the strength of structural paths supporting diffusion (Equation 6), we again used penalized splines. The final model can be written as:

$$\text{CBF} \sim \text{path strength} + \text{spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion}. \quad (22)$$

Assessments of path strengths were corrected for false discovery rate across the statistical tests performed over the discrete path lengths.

7.7.2 Trade-offs between modularity and diffusion architecture

Next, we sought to evaluate the metabolic running cost of brain network properties, in line with calls for investigation of the economic landscape of resource-constrained trade-offs between hallmark brain network architectures such as modularity (Equation 9) and new measures of brain network organization [3]. Following our findings that CBF is associated with structural properties supporting diffusion, we investigated path transitivity (Equation 8). We continued to use penalized splines to model the non-linear patterns of CBF

and brain properties of interest. The final model can be written as:

$$\text{CBF} \sim \text{path transitivity} + \text{modularity} + \text{path transitivity by modularity}$$
$$+ \text{spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion}. \tag{23}$$

To visualize the landscape of CBF as a function of modularity and path transitivity, we plotted the GAM model response function. We described the distribution of modularity and path transitivity across individuals using frequency histograms.

7.7.3   Compression efficiency and development

To assess the possibility of distinct compression efficiency of brain networks compared to random networks, we calculated the resource efficiency (Equation 12) at 14 levels of distortion and performed an analysis of variance (ANOVA) test. The ANOVA model can be written as:

$$\text{Resources} \sim \text{distortion} + \text{type of network}, \tag{24}$$

where the type of the network is a categorical variable designating if the network was a brain network or a random network.

To compute compression efficiency per individual brain network, we used a polynomial regression function to find the best linear fit to the monotonic rate-distortion function according to the prediction of a linear rate-distortion gradient in semi-log space (log(resources) as a function of distortion). Next, we used a GAM model to assess the non-linear patterns of compression efficiency in development, which we can formally write as follows:

$$\text{Compression efficiency} \sim \text{spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion}. \tag{25}$$

7.7.4   Compression efficiency of chemotaxis

To compute the compression efficiency of chemotactic diffusion, we modified the model of Equation 24 to instead calculate resource efficiency using the biased random walk matrices from Equation 16. The model

can be written:

$$\text{Resources} \sim \text{distortion} + \text{type of random walk} + \text{distortion by type of random walk}, \qquad (26)$$

where the type of random walk is a categorical variable designating unbiased random walks using the structural network, attraction-biased random walks using the structural network biased with CBF, and repulsion-biased random walks using the structural network biased with (1 minus CBF). To assess the hypothesis that resources differ according to the type of random walk, we performed t-tests while controlling for family-wise error rate across multiple comparisons.

7.7.5 Compression efficiency in a low or high fidelity regime

Next, we sought to test the predictions of a high or low fidelity communication regime. In a high fidelity regime, minimum resources given an expected distortion should increase monotonically as a function of network complexity. To assess whether the relationship between resources and network complexity (operationalized here as network size) is monotonic, we used a linear model written as:

$$\text{Log}_2(\text{resources}) \sim \log_{10}(\text{network size}). \qquad (27)$$

In a low fidelity regime, minimum resources given an expected distortion should plateau as a function of complexity. We hypothesized that path transitivity is a property of structural networks that supports lossy compression and storage savings. The complexity of the shortest path was defined as the number of nodes contributing to path transitivity (Equation 8). To assess whether the resources non-linearly plateau as a function of shortest path complexity, we used a GAM model written as follows:

$$\text{Log}_2(\text{resources}) \sim \text{spline(shortest path complexity)}. \qquad (28)$$

7.7.6 Compression efficiency and patterns of neurodevelopment

To explore how compression efficiency might relate to patterns of cortical myelination and areal scaling, we assessed the Spearman's correlation coefficient between myelination or scaling and send or receive compression efficiency. To further test correspondence between brain maps, we used a spatial permutation test, which generates a null distribution of randomly rotated brain maps that preserve the spatial covariance structure

of the original data [79]. We refer to the p-value of this statistical test as p_{SPIN}. Finally, we applied the conservative Holm-Bonferroni correction for family-wise error across these tests.

7.7.7   Compression efficiency and rich-club hubs

Given the assumed integrative and broadcasting function of rich-club hubs, we sought to evaluate whether compression efficiency differed in rich-club hubs compared to other brain regions. We used the Wilcoxon rank-sum test to compare regional compression efficiency of either receiving or sending messages. Moreover, we assessed whether there was a difference between CBF in the rich-club hubs compared to other brain regions. Lastly, we tested the correlation of compression efficiency in the rich-club hubs and other brain regions with cognitive efficiency. To model non-linear patterns of cognitive efficiency, we used penalized splines controlling for potentially confounding covariates. The final model can be written as:

$$\text{Cognitive efficiency} \sim \text{compression efficiency} + \text{spline(age)}$$
$$+ \text{ sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion.} \tag{29}$$

Due to previous report of the relationship between cognition and global efficiency (Equation 1), we determined that compression efficiency and global efficiency were not collinear and therefore conducted a sensitivity analysis including global efficiency as a covariate. The model was written as:

$$\text{Cognitive efficiency} \sim \text{compression efficiency} + \text{global efficiency}$$
$$+ \text{ spline(age)} + \text{sex} + \text{spline(age by sex)} + \text{degree} + \text{density} + \text{motion.} \tag{30}$$

7.8   Citation Diversity Statement

Recent work in neuroscience and other fields has identified a bias in citation practices such that papers from women and other minorities are under-cited relative to the number of such papers in the field [80, 81, 82, 83, 84, 85]. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, and other factors. We used automatic classification of gender based on the first names of the first and last authors [80], with possible combinations including male/male, male/female, female/male, female/female. Excluding self-citations to the senior authors of our current paper, the references contain 58.0% male/male, 8.7% male/female, 21.7% female/male, 7.2% female/female, and 4.3% unknown categorization. We look forward to future work that could help us to better understand how to support equitable practices in science.

References

[1] M. J. West-Eberhard, *Developmental plasticity and evolution.* Oxford University Press, 2003.

[2] S. B. Laughlin, "Energy as a constraint on the coding and processing of sensory information," *Current opinion in neurobiology*, vol. 11, no. 4, pp. 475–480, 2001.

[3] E. Bullmore and O. Sporns, "The economy of brain network organization," *Nature Reviews Neuroscience*, vol. 13, no. 5, p. 336, 2012.

[4] R. L. Buckner and F. M. Krienen, "The evolution of distributed association networks in the human brain," *Trends in cognitive sciences*, vol. 17, no. 12, pp. 648–665, 2013.

[5] A. Avena-Koenigsberger, J. Goñi, R. Solé, and O. Sporns, "Network morphospace," *Journal of the Royal Society Interface*, vol. 12, no. 103, p. 20140881, 2015.

[6] K. J. Whitaker, P. E. Vértes, R. Romero-Garcia, F. Váša, M. Moutoussis, G. Prabhu, N. Weiskopf, M. F. Callaghan, K. Wagstyl, T. Rittman, *et al.*, "Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome," *Proceedings of the National Academy of Sciences*, vol. 113, no. 32, pp. 9105–9110, 2016.

[7] P. Reardon, J. Seidlitz, S. Vandekar, S. Liu, R. Patel, M. T. M. Park, A. Alexander-Bloch, L. S. Clasen, J. D. Blumenthal, F. M. Lalonde, *et al.*, "Normative brain size variation and brain shape diversity in humans," *Science*, vol. 360, no. 6394, pp. 1222–1227, 2018.

[8] A. Di Martino, D. A. Fair, C. Kelly, T. D. Satterthwaite, F. X. Castellanos, M. E. Thomason, R. C. Craddock, B. Luna, B. L. Leventhal, X.-N. Zuo, *et al.*, "Unraveling the miswired connectome: a developmental perspective," *Neuron*, vol. 83, no. 6, pp. 1335–1353, 2014.

[9] N. A. Crossley, A. Mechelli, J. Scott, F. Carletti, P. T. Fox, P. McGuire, and E. T. Bullmore, "The hubs of the human connectome are generally implicated in the anatomy of brain disorders," *Brain*, vol. 137, no. 8, pp. 2382–2395, 2014.

[10] L. L. Gollo, J. A. Roberts, V. L. Cropley, M. A. Di Biase, C. Pantelis, A. Zalesky, and M. Breakspear, "Fragility and volatility of structural hubs in the human connectome," *Nature neuroscience*, vol. 21, no. 8, p. 1107, 2018.

[11] A. N. Kaczkurkin, T. M. Moore, M. E. Calkins, R. Ciric, J. A. Detre, M. A. Elliott, E. B. Foa, A. G. de la Garza, D. R. Roalf, A. Rosen, *et al.*, "Common and dissociable regional cerebral blood flow differences associate with dimensions of psychopathology across categorical diagnoses," *Molecular psychiatry*, vol. 23, no. 10, p. 1981, 2018.

[12] J. Goni, A. Avena-Koenigsberger, N. V. de Mendizabal, M. P. van den Heuvel, R. F. Betzel, and O. Sporns, "Exploring the morphospace of communication efficiency in complex networks," *PLoS One*, vol. 8, no. 3, p. e58070, 2013.

[13] J. Goñi, M. P. van den Heuvel, A. Avena-Koenigsberger, N. V. de Mendizabal, R. F. Betzel, A. Griffa, P. Hagmann, B. Corominas-Murtra, J.-P. Thiran, and O. Sporns, "Resting-brain functional connectivity predicted by analytic measures of network communication," *Proceedings of the National Academy of Sciences*, vol. 111, no. 2, pp. 833–838, 2014.

[14] B. Mišić, R. F. Betzel, A. Nematzadeh, J. Goni, A. Griffa, P. Hagmann, A. Flammini, Y.-Y. Ahn, and O. Sporns, "Cooperative and competitive spreading dynamics on the human connectome," *Neuron*, vol. 86, no. 6, pp. 1518–1529, 2015.

[15] A. Avena-Koenigsberger, B. Mišić, R. X. Hawkins, A. Griffa, P. Hagmann, J. Goñi, and O. Sporns, "Path ensembles and a tradeoff between communication efficiency and resilience in the human connectome," *Brain Structure and Function*, vol. 222, no. 1, pp. 603–618, 2017.

[16] A. Avena-Koenigsberger, B. Misic, and O. Sporns, "Communication dynamics in complex brain networks," *Nature Reviews Neuroscience*, vol. 19, no. 1, p. 17, 2018.

[17] J. Stiso and D. S. Bassett, "Spatial embedding imposes constraints on neuronal network architectures," *Trends in cognitive sciences*, 2018.

[18] W. B. Levy and R. A. Baxter, "Energy efficient neural codes," *Neural computation*, vol. 8, no. 3, pp. 531–543, 1996.

[19] E. Tang, M. G. Mattar, C. Giusti, D. M. Lydon-Staley, S. L. Thompson-Schill, and D. S. Bassett, "Effective learning is accompanied by high-dimensional and efficient representations of neural activity," *Nature neuroscience*, vol. 22, no. 6, p. 1000, 2019.

[20] J. M. Shine, M. Breakspear, P. T. Bell, K. A. E. Martens, R. Shine, O. Koyejo, O. Sporns, and R. A. Poldrack, "Human cognition involves the dynamic integration of neural activity and neuromodulatory systems," *Nature neuroscience*, vol. 22, no. 2, p. 289, 2019.

[21] M. L. Mack, A. R. Preston, and B. C. Love, "Ventromedial prefrontal cortex compression during concept learning," *Nature Communications*, vol. 11, no. 1, pp. 1–11, 2020.

[22] G. L. Baum, R. Ciric, D. R. Roalf, R. F. Betzel, T. M. Moore, R. T. Shinohara, A. E. Kahn, S. N. Vandekar, P. E. Rupert, M. Quarmley, *et al.*, "Modular segregation of structural brain networks supports the development of executive function in youth," *Current Biology*, vol. 27, no. 11, pp. 1561–1572, 2017.

[23] T. D. Satterthwaite, M. A. Elliott, K. Ruparel, J. Loughead, K. Prabhakaran, M. E. Calkins, R. Hopson, C. Jackson, J. Keefe, M. Riley, *et al.*, "Neuroimaging of the philadelphia neurodevelopmental cohort," *Neuroimage*, vol. 86, pp. 544–553, 2014.

[24] R. C. Gur, J. D. Ragland, M. Reivich, J. H. Greenberg, A. Alavi, and R. E. Gur, "Regional differences in the coupling between resting cerebral blood flow and metabolism may indicate action preparedness as a default state," *Cerebral Cortex*, vol. 19, no. 2, pp. 375–382, 2008.

[25] C. R. Sims, "Efficient coding explains the universal law of generalization in human perception," *Science*, vol. 360, no. 6389, pp. 652–656, 2018.

[26] T. Berger, "Rate distortion theory, a mathematical basis for data compression (prentice-hall," *Inc. Englewood Cliffs, New Jersey*, 1971.

[27] S. E. Marzen and S. DeDeo, "The evolution of lossy compression," *Journal of The Royal Society Interface*, vol. 14, no. 130, p. 20170166, 2017.

[28] B. Várkuti, M. Cavusoglu, A. Kullik, B. Schiffler, R. Veit, Ö. Yilmaz, W. Rosenstiel, C. Braun, K. Uludag, N. Birbaumer, *et al.*, "Quantifying the link between anatomical connectivity, gray matter volume and regional cerebral blood flow: an integrative mri study," *PLoS One*, vol. 6, no. 4, p. e14801, 2011.

[29] M. P. van den Heuvel, R. S. Kahn, J. Goñi, and O. Sporns, "High-cost, high-capacity backbone for global brain communication," *Proceedings of the National Academy of Sciences*, vol. 109, no. 28, pp. 11372–11377, 2012.

[30] M. P. Van Den Heuvel, C. J. Stam, R. S. Kahn, and H. E. H. Pol, "Efficiency of functional brain networks and intellectual performance," *Journal of Neuroscience*, vol. 29, no. 23, pp. 7619–7624, 2009.

[31] X. Liang, Q. Zou, Y. He, and Y. Yang, "Coupling of functional connectivity and regional cerebral blood flow reveals a physiological basis for network hubs of the human brain," *Proceedings of the National Academy of Sciences*, vol. 110, no. 5, pp. 1929–1934, 2013.

[32] T. D. Satterthwaite, R. T. Shinohara, D. H. Wolf, R. D. Hopson, M. A. Elliott, S. N. Vandekar, K. Ruparel, M. E. Calkins, D. R. Roalf, E. D. Gennatas, *et al.*, "Impact of puberty on the evolution of cerebral perfusion during adolescence," *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8643–8648, 2014.

[33] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical review letters*, vol. 87, no. 19, p. 198701, 2001.

[34] M. F. Glasser, M. S. Goyal, T. M. Preuss, M. E. Raichle, and D. C. Van Essen, "Trends and properties of human cerebral cortex: correlations with cortical myelin content," *Neuroimage*, vol. 93, pp. 165–175, 2014.

[35] M. Bertolero, B. Yeo, and M. D'esposito, "The diverse club," *Nature communications*, vol. 8, no. 1, p. 1277, 2017.

[36] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLoS computational biology*, vol. 3, no. 2, p. e17, 2007.

[37] H. Johansen-Berg, "Behavioural relevance of variation in white matter microstructure," *Current opinion in neurology*, vol. 23, no. 4, pp. 351–358, 2010.

[38] M. Rubinov, "Constraints and spandrels of interareal connectomes," *Nature communications*, vol. 7, p. 13812, 2016.

[39] O. Sporns, "Network attributes for segregation and integration in the human brain," *Current opinion in neurobiology*, vol. 23, no. 2, pp. 162–171, 2013.

[40] G. Hahn, A. Ponce-Alvarez, G. Deco, A. Aertsen, and A. Kumar, "Portraits of communication in neuronal networks," *Nature Reviews Neuroscience*, vol. 20, no. 2, pp. 117–127, 2019.

[41] A. Avena-Koenigsberger, X. Yan, A. Kolchinsky, M. P. van den Heuvel, P. Hagmann, and O. Sporns, "A spectrum of routing strategies for brain networks," *PLOS Computational Biology*, vol. 15, pp. 1–24, 03 2019.

[42] P. E. Vértes, A. Alexander-Bloch, and E. T. Bullmore, "Generative models of rich clubs in hebbian neuronal networks and large-scale human brain networks," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1653, p. 20130531, 2014.

[43] S. Oldham and A. Fornito, "The development of brain network hubs," *Developmental cognitive neuroscience*, 2018.

[44] E. Ferrer, K. J. Whitaker, J. S. Steele, C. T. Green, C. Wendelken, and S. A. Bunge, "White matter maturation supports the development of reasoning ability through its influence on processing speed," *Developmental Science*, vol. 16, no. 6, pp. 941–951, 2013.

[45] J. Scholz, M. C. Klein, T. E. Behrens, and H. Johansen-Berg, "Training induces changes in white-matter architecture," *Nature neuroscience*, vol. 12, no. 11, p. 1370, 2009.

[46] S. G. Vij, J. S. Nomi, D. R. Dajani, and L. Q. Uddin, "Evolution of spatial and temporal features of functional brain networks across the lifespan," *NeuroImage*, vol. 173, pp. 498–508, 2018.

[47] T. Watanabe, S. Hirose, H. Wada, Y. Imai, T. Machida, I. Shirouzu, S. Konishi, Y. Miyashita, and N. Masuda, "A pairwise maximum entropy model accurately describes resting-state human brain networks," *Nature communications*, vol. 4, p. 1370, 2013.

[48] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, M. P. van den Heuvel, and M. Breakspear, "Connectome sensitivity or specificity: which is more important?," *Neuroimage*, vol. 142, pp. 407–420, 2016.

[49] F. Lieder and T. L. Griffiths, "Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources," *Behavioral and Brain Sciences*, pp. 1–85, 2019.

[50] A. C. Schapiro, N. B. Turk-Browne, M. M. Botvinick, and K. A. Norman, "Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 372, no. 1711, p. 20160049, 2017.

[51] V. Balasubramanian, D. Kimber, and M. J. B. Ii, "Metabolically efficient information processing," *Neural computation*, vol. 13, no. 4, pp. 799–815, 2001.

[52] L. M. Hernandez, J. D. Rudie, S. A. Green, S. Bookheimer, and M. Dapretto, "Neural signatures of autism spectrum disorders: insights into brain network dynamics," *Neuropsychopharmacology*, vol. 40, no. 1, p. 171, 2015.

[53] I. Dinstein, D. J. Heeger, L. Lorenzi, N. J. Minshew, R. Malach, and M. Behrmann, "Unreliable evoked responses in autism," *Neuron*, vol. 75, no. 6, pp. 981–991, 2012.

[54] D. R. Roalf, M. Quarmley, M. A. Elliott, T. D. Satterthwaite, S. N. Vandekar, K. Ruparel, E. D. Gennatas, M. E. Calkins, T. M. Moore, R. Hopson, *et al.*, "The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort," *Neuroimage*, vol. 125, pp. 903–919, 2016.

[55] T. M. Moore, S. P. Reise, R. E. Gur, H. Hakonarson, and R. C. Gur, "Psychometric properties of the penn computerized neurocognitive battery.," *Neuropsychology*, vol. 29, no. 2, p. 235, 2015.

[56] E. Tang, C. Giusti, G. L. Baum, S. Gu, E. Pollock, A. E. Kahn, D. R. Roalf, T. M. Moore, K. Ruparel, R. C. Gur, *et al.*, "Developmental increases in white matter network controllability support a growing diversity of brain dynamics," *Nature communications*, vol. 8, no. 1, p. 1252, 2017.

[57] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, *et al.*, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, p. 171, 2016.

[58] L. Cammoun, X. Gigandet, D. Meskaldji, J. P. Thiran, O. Sporns, K. Q. Do, P. Maeder, R. Meuli, and P. Hagmann, "Mapping the human connectome at multiple scales with diffusion spectrum mri," *Journal of neuroscience methods*, vol. 203, no. 2, pp. 386–397, 2012.

[59] F.-C. Yeh, T. D. Verstynen, Y. Wang, J. C. Fernández-Miranda, and W.-Y. I. Tseng, "Deterministic diffusion fiber tracking improved by quantitative anisotropy," *PloS one*, vol. 8, no. 11, p. e80713, 2013.

[60] Z. Wang, G. K. Aguirre, H. Rao, J. Wang, M. A. Fernández-Seara, A. R. Childress, and J. A. Detre, "Empirical optimization of asl data analysis using an asl data processing toolbox: Asltbx," *Magnetic resonance imaging*, vol. 26, no. 2, pp. 261–269, 2008.

[61] W.-C. Wu, V. Jain, C. Li, M. Giannetta, H. Hurt, F. W. Wehrli, and D. J. Wang, "In vivo venous blood t1 measurement using inversion recovery true-fisp in children and adults," *Magnetic resonance in medicine*, vol. 64, no. 4, pp. 1140–1147, 2010.

[62] V. Jain, J. Duda, B. Avants, M. Giannetta, S. X. Xie, T. Roberts, J. A. Detre, H. Hurt, F. W. Wehrli, and D. J. Wang, "Longitudinal reproducibility and accuracy of pseudo-continuous arterial spin–labeled perfusion mr imaging in typically developing children," *Radiology*, vol. 263, no. 2, pp. 527–536, 2012.

[63] M. F. Glasser and D. C. Van Essen, "Mapping human cortical areas in vivo based on myelin content as revealed by t1-and t2-weighted mri," *Journal of Neuroscience*, vol. 31, no. 32, pp. 11597–11616, 2011.

[64] S. N. Wood, "Stable and efficient multiple smoothing parameter estimation for generalized additive models," *Journal of the American Statistical Association*, vol. 99, no. 467, pp. 673–686, 2004.

[65] C. O. Becker, S. Pequito, G. J. Pappas, M. B. Miller, S. T. Grafton, D. S. Bassett, and V. M. Preciado, "Spectral mapping of brain functional connectivity from diffusion imaging," *Scientific reports*, vol. 8, no. 1, p. 1411, 2018.

[66] O. Sporns and R. F. Betzel, "Modular brain networks," *Annual review of psychology*, vol. 67, pp. 613–640, 2016.

[67] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. W. Moore, and E. T. Bullmore, "Efficient physical embedding of topologically complex information processing networks in brains and computer circuits," *PLoS computational biology*, vol. 6, no. 4, p. e1000748, 2010.

[68] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7641–7646, 2011.

[69] M. A. Bertolero, B. T. Yeo, and M. D'Esposito, "The modular and integrative functional architecture of the human brain," *Proceedings of the National Academy of Sciences*, vol. 112, no. 49, pp. E6798–E6807, 2015.

[70] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

[71] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.

[72] "Chapter 7 - paths, diffusion, and navigation," in *Fundamentals of Brain Network Analysis* (A. Fornito, A. Zalesky, and E. T. Bullmore, eds.), pp. 207 – 255, San Diego: Academic Press, 2016.

[73] C. R. Sims, "Rate–distortion theory and human perception," *Cognition*, vol. 152, pp. 181–198, 2016.

[74] W. Alt, "Biased random walk models for chemotaxis and related diffusion approximations," *Journal of mathematical biology*, vol. 9, no. 2, pp. 147–177, 1980.

[75] M. P. van den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends in cognitive sciences*, vol. 17, no. 12, pp. 683–696, 2013.

[76] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani, "Detecting rich-club ordering in complex networks," *Nature physics*, vol. 2, no. 2, p. 110, 2006.

[77] G. Collin, O. Sporns, R. C. Mandl, and M. P. van den Heuvel, "Structural and functional aspects relating to cost and benefit of rich club organization in the human cerebral cortex," *Cerebral cortex*, vol. 24, no. 9, pp. 2258–2267, 2013.

[78] S. Maslov and K. Sneppen, "Specificity and stability in topology of protein networks," *Science*, vol. 296, no. 5569, pp. 910–913, 2002.

[79] A. F. Alexander-Bloch, H. Shou, S. Liu, T. D. Satterthwaite, D. C. Glahn, R. T. Shinohara, S. N. Vandekar, and A. Raznahan, "On testing for spatial correspondence between maps of human brain structure and function," *Neuroimage*, vol. 178, pp. 540–551, 2018.

[80] J. D. Dworkin, K. A. Linn, E. G. Teich, P. Zurn, R. T. Shinohara, and D. S. Bassett, "The extent and drivers of gender imbalance in neuroscience reference lists," *bioRxiv*, 2020.

[81] D. Maliniak, R. Powers, and B. F. Walter, "The gender citation gap in international relations," *International Organization*, vol. 67, no. 4, pp. 889–922, 2013.

[82] N. Caplar, S. Tacchella, and S. Birrer, "Quantitative evaluation of gender bias in astronomical publications from citation counts," *Nature Astronomy*, vol. 1, no. 6, p. 0141, 2017.

[83] P. Chakravartty, R. Kuo, V. Grubbs, and C. McIlwain, "# communicationsowhite," *Journal of Communication*, vol. 68, no. 2, pp. 254–266, 2018.

[84] Y. Thiem, K. F. Sealey, A. E. Ferrer, A. M. Trott, and R. Kennison, "Just Ideas? The Status and Future of Publication Ethics in Philosophy: A White Paper," tech. rep., 2018.

[85] M. L. Dion, J. L. Sumner, and S. M. Mitchell, "Gendered citation patterns across political science and social science methodology fields," *Political Analysis*, vol. 26, no. 3, pp. 312–327, 2018.