Privacy-Preserving Visualization of Brain Functional Connectivity

Ye Tao[a], Anand D. Sarwate[a], Sandeep Panta[b], Jessica Turner[b,c], Sergey Plis[b], Vince D. Calhoun[b]

[a] *Department of Electrical and Computer Engineering at Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA*
[b] *TReNDS Center, Georgia State, Georgia Tech, and Emory, Atlanta, GA 30303, USA*
[c] *Department of Psychiatry and Behavioral Health at The Ohio State University, Columbus, OH 43210, USA*

Abstract

Data visualizations are an integral part of neuroimging research, supporting activities ranging from exploratory data analysis to the interpretation and communication of findings. While essential, visualizations can also reveal private information about individual participants. In this paper, we discuss how visualizations may inadvertently lead to privacy leakage and explore methods to mitigate such risks. Our work investigates ways to securely share visualizations that faithfully preserve the patterns supporting the derived insights from data analysis, rather than deriving conclusions from the visualizations themselves. We address the problem of privacy-preserving visualization under the framework of differential privacy, focusing on commonly used visualization methods for functional network connectivity. Several perturbation-based strategies are investigated for protecting correlation-related measures, with analyses of their privacy costs and the effects of pre- and post-processing. To achieve a better balance between privacy and visual utility, we propose workflows for connectogram and seed-based connectivity visualizations that preserve the qualitative structure of non-private results. Overall, this work illustrates how differential privacy can be effectively applied to neuroimaging visualization, highlighting its potential as a principled approach for safeguarding sensitive information.

Keywords:
Differential privacy, visualization, functional network connectivity

1. Introduction

Data visualization (Kostelnick, 2008; Bhattacharjee et al., 2020) is an integral part of the research workflow, serving distinct purposes at different stages of analysis. In the early phase of exploratory data analysis, it enables researchers to examine data distributions and uncover underlying patterns. At later stages, visualizations facilitate the communication of analytical findings and aid in the interpretation of results. While privacy concerns are often discussed in the context of sharing raw data or summary statistics (Sweeney, 1997; Narayanan and Shmatikov, 2008), visualizations derived from these data can also expose sensitive information, particularly in domains involving human health and behavior, yet this aspect has received comparatively less attention than data leakage itself.

A range of privacy attacks have demonstrated that anonymized or aggregated outputs can still lead to re-identification or data leakage. Classic examples include linking de-identified health records with publicly available information (Sweeney, 2002) and uncovering user identities from anonymized movie rating data (Narayanan and Shmatikov, 2008). Similar risks extend to visualizations, where adversaries may exploit graphical cues and/or external sources to infer sensitive information. A well-known real-world example is Strava's release of an anonymized global activity heatmap, which unintentionally exposed secret U.S. military bases (Allen, 2020). Comparable risks arise in scientific visualizations. For example, scatterplots or boxplots commonly used in neuroimaging exploratory data analysis may reveal actual data values. Likewise, when a research team publishes average fMRI functional network connectivity (FNC) matrices for healthy and diseased groups, re-identification becomes possible: if an attacker knows the FNC of all subjects except one, the group membership of the remaining individual can be inferred by comparing the published visualizations (see Figure 1). Although this represents an extreme case, it illustrates how visualization can lead to privacy leakage, as auxiliary information about individuals may be accessible through various external sources or linkage attacks. This example also motivates the need for privacy-preserving visualization methods that remain robust even when an attacker possesses substantial background knowledge. By contrast, even when the attacker has only limited knowledge, published visualizations can still enable inference if privacy is not explicitly considered. Concretely, one can adapt membership inference tests (Homer et al., 2008; Dwork et al., 2015) to the visualization setting. Representing a published vi-

sualization as v_{pub} generated from a group, a target individual's visualization as v_{target}, and a reference sample as v_{ref}, an adversary can compare similarity scores such as $\langle v_{\text{target}}, v_{\text{pub}} \rangle$ and $\langle v_{\text{ref}}, v_{\text{pub}} \rangle$, and determine whether the target belongs to the group based on their difference. Although such tests are abstract and rely on modeling assumptions, they demonstrate that releasing aggregate or visual summaries can still leak membership information in the absence of explicit privacy protection.

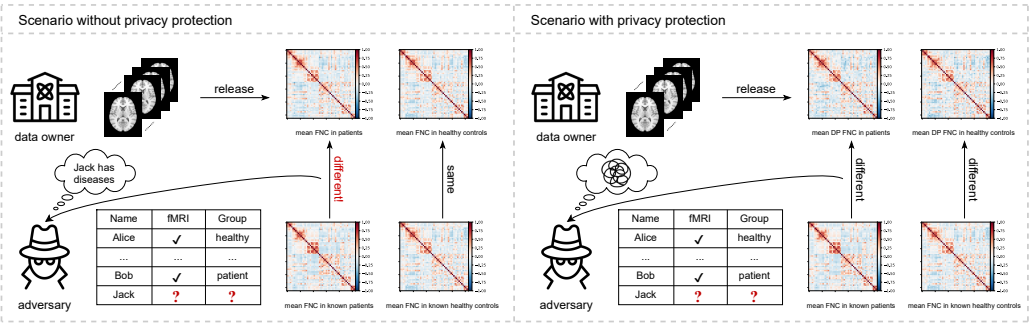


Figure 1: In a scenario without privacy protection, an adversary can infer information about unknown individuals. However, with differential privacy, the adversary cannot make such inferences.

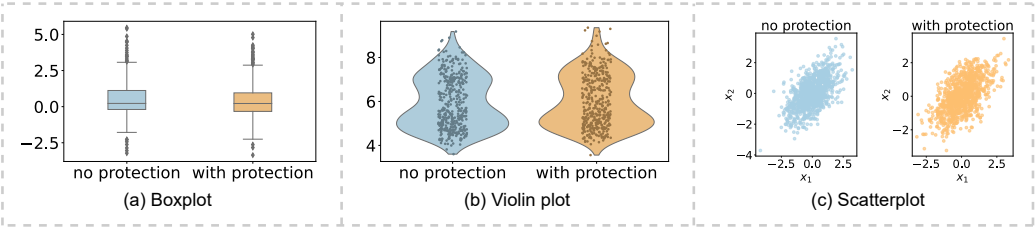


Figure 2: Comparison of visualizations with and without differential privacy protection, illustrating that differential privacy protects individual information while preserving the overall patterns and characteristics each plot is intended to communicate. Under differential privacy, the goal is to convey patterns rather than exact data values or make direct decisions; for example, boxplots show distribution and variability, violin plots reveal multimodal features, and scatter plots illustrate the increasing trend of one variable as the other grows.

Methods such as differential privacy (DP) (Dwork, 2006) offer protection against data leakage risks by introducing carefully calibrated randomness

into the released outputs, thereby preventing accurate inference about any individual even when an adversary possesses substantial background knowledge. This protection often comes at the cost of a loss of accuracy and utility, which raises an important question: can privacy-preserving visualizations still capture the same patterns and convey the intended information? Previous study on differentially private visualizations, including boxplots, violin plots, and scatterplots (Tao and Sarwate, 2025; Panavas et al., 2023), has demonstrated that these methods can effectively protect individual information while preserving the essential patterns intended to be communicated (see Figure 2). Motivated by the growing need for privacy protection and the previous findings, this paper examines how privacy considerations can be incorporated into neuroimaging visualizations. Specifically, we explore the application of DP to commonly used visualizations, including FNC, connectogram, and seed-based connectivity (SBC) plots (Van Den Heuvel and Pol, 2010; Mohanty et al., 2020; Joel et al., 2011; Chopra et al., 2023). Our main contributions are summarized as follows.

- We propose a series of novel methods for DP visualization in neuroimaging, including pre- and post-processing strategies to enhance visual quality.
- We systematically evaluate these methods across multiple real-world datasets, comparing the performance of different mechanisms and identifying their optimal composition strategies under varying privacy parameters.
- We develop end-to-end differentially private workflows for connectogram and SBC visualizations, illustrating group comparisons and significant connectivities.
- We assess the interpretability and utility of the proposed visualizations through both qualitative and quantitative analyses.

2. Background

2.1. Differential Privacy

As data collection, storage, and analysis capabilities have improved significantly, privacy risks have increased, driving a growing focus on privacy research over the past two decades. Earlier syntactic approaches to privacy,

such as de-identification and k-Anonymity (Garfinkel, 2015; Sweeney, 2002), are susceptible to linkage or membership inference attacks (Dwork et al., 2017; Merener, 2012; Ganta et al., 2008; Henriksen-Bulmer and Jeary, 2016; Fiore et al., 2020).

In contrast, differential privacy (Dwork et al., 2006; Dwork and Roth, 2014) has emerged as a state-of-the-art approach that provides a quantitative framework for privacy protection, addressing the vulnerabilities of earlier syntactic methods. It enables statistical analyses on datasets while ensuring that the privacy of individuals whose records are included is preserved. A differentially private algorithm or mechanism is a randomized function: the randomization is designed to make it difficult for an adversary observing the output to discern private information. The kind of information that DP protects is captured by the notion of neighboring databases. For two datasets D and D' that are neighbors, denoted as $D \sim D'$, the randomization should make it difficult for the adversary to distinguish, using a hypothesis test, whether the output was generated from D or D' (Wasserman and Zhou, 2010; Kairouz et al., 2015). Most commonly, we say $D \sim D'$ if they differ by the data of a single individual: this means that an adversary cannot reliably tell whether that individual was in the dataset or not. A full description of the key notations used in the paper is provided in Table 1. Some minor abuse of notation may occur, but it should be clear from the context.

Notation	**Description**
ϵ	Privacy budget
δ	Probability of privacy violation or failure
M	Randomized algorithm or mechanism
$D \sim D'$	Neighboring datasets
Δ_1, Δ_2	Global ℓ_1-sensitivity and ℓ_2-sensitivity of $f(D)$
Δ_u	Sensitivity of the utility function u
v	Brain voxel
r	A single ROI or r brain regions, as defined by context
t, T	Timepoint and the total number of timepoints, respectively
$S(\cdot, t)$	BOLD value for a voxel or an ROI at time t
X_{ij}	Correlation coefficient between the timeseries of r_i and r_j
$X^{(n)} \in \mathbb{R}^{r \times r}$	FNC matrix for subject n
$\bar{X}, \tilde{X}$	Average FNC and perturbed FNC

Table 1: Key notations used in the paper.

For an algorithm to be differentially private, it must guarantee that the adversary's test is unreliable for all possible $D \sim D'$. This unreliability is measured in DP using two parameters, ϵ and δ (see Definition 1). The parameter ϵ is the privacy budget that quantifies the maximum distance between the output distributions of algorithm M when it runs on any two neighboring datasets. A smaller ϵ will give stronger privacy but less accurate responses. The parameter δ represents the probability of information leakage. Typically, we are interested in values of δ that are far less than the inverse of the database size, $\delta \ll 1/N$. The definition of DP ensures that for all neighboring datasets D and D', the absolute value of the privacy loss (Dwork and Roth, 2014) is bounded by ϵ with a probability of at least $1 - \delta$.

Definition 1 ((ϵ, δ)-DP (Dwork and Roth, 2014)). *Let $\epsilon, \delta \geq 0$, a randomized algorithm $M : \mathcal{D} \to \mathcal{H}$ is (ϵ, δ)-differentially private if for all $\mathcal{Y} \subseteq \mathcal{H}$ and for all neighboring datasets $D, D' \in \mathcal{D}$,*

$$\mathbb{P}[M(D) \in \mathcal{Y}] \leq \exp(\epsilon)\mathbb{P}[M(D') \in \mathcal{Y}] + \delta.$$

To design a (ϵ, δ)-DP algorithm that approximates a desired function $f : \mathcal{D} \to \mathbb{R}^d$ (in our case, a data visualization), a common approach is to add random noise of appropriate magnitude to the output of $f(D)$. The amount of noise required depends not only on ϵ and δ but also on the sensitivity of the function, which quantifies the maximum change in the output when a single data point in the dataset is modified. The global ℓ_1 and ℓ_2 sensitivity of the function f are given by

$$\Delta_i = \max_{D \sim D'} \|f(D) - f(D')\|_i$$

for $i = 1$ or 2, where $\| \cdot \|_i$ denotes the ℓ_i-norm. "Pure" $(\epsilon, 0)$-DP can be achieved by applying the Laplace mechanism (Dwork and Roth, 2014), defined as $M_L(D) = f(D) + Z$, where Z follows the Laplace distribution $Z \sim \text{Laplace}(0, \Delta_1/\epsilon \cdot I)$. The identity matrix $I \in \mathbb{R}^d$ is included when $f(D)$ is high-dimensional. Alternative noise distributions have also been proposed (Geng et al., 2015; Alghamdi et al., 2022b). For any $\epsilon, \delta \in (0, 1)$, the classical Gaussian mechanism $M_G(D)$ guarantees "approximate" (ϵ, δ)-DP when $M_G(D) = f(D) + Z$, where $Z \sim \mathcal{N}(0, \sigma^2 \cdot I)$ and $\sigma = \Delta_2\sqrt{2\log(1.25/\delta)}/\epsilon$. However, this choice of σ can be overly conservative in the high privacy regime ($\epsilon \to 0$) and does not apply in the low privacy regime ($\epsilon \to \infty$). To address this, Balle and Wang (2018) proposed an "analytic" formula for the variance.

In the following sections and experiments, we use the analytic Gaussian mechanism rather than the classical Gaussian mechanism when considering noise from the Gaussian distribution.

Additive noise mechanisms are well-suited for privatizing real-valued functions. In other scenarios, algorithms may instead randomly select an output from a set of choices. In such cases, the exponential mechanism (McSherry and Talwar, 2007) enables $(\epsilon, 0)$-DP randomized selection from a discrete set. Sometimes, numerical outputs can be mapped to a discrete set. For example, in visualization, certain numerical values are mapped to the same color. Given a range $\mathcal{H}$, the exponential mechanism is defined with respect to a utility function $u : \mathcal{D} \times \mathcal{H} \to \mathbb{R}$, which assigns a utility score to each input-output pair. The exponential mechanism $M_E(D)$ selects and outputs an element $H \in \mathcal{H}$ with a probability proportional to $\exp\left(\frac{\epsilon u(D,H)}{2\Delta_u}\right)$, where $\Delta_u = \max_{H \in \mathcal{H}} \max_{D \sim D'} |u(D, H) - u(D', H)|$.

Mechanisms such as additive noise or randomized selection protect privacy in a single analysis. When applied multiple times, the accumulated privacy loss can be formally quantified using the composition property of DP. There are several types of composition rules: basic, advanced, exact, and Rényi differential privacy (RDP)-based compositions (Dwork et al., 2010; Dwork and Roth, 2014; Kairouz et al., 2015; Murtagh and Vadhan, 2015; Abadi et al., 2016; Mironov, 2017; Balle et al., 2020; Dong et al., 2022; Sommer et al., 2018; Alghamdi et al., 2022b,a; Koskela and Honkela, 2020; Gopi et al., 2021; Koskela et al., 2020; Ghazi et al., 2022; Doroshenko et al., 2022). A full discussion of composition rules is beyond the scope of this paper, but for the purposes of this work, advances in composition analyses can only improve the guarantees that we make. Another important property of DP is post-processing immunity: any function applied to the output of a differentially private mechanism cannot weaken its privacy guarantees. This property ensures that subsequent analyses or transformations remain safe without additional noise.

2.2. Privacy-preserving Visualizations

Anonymization and perturbation-based approaches. Early studies, such as Wang et al. (2018) and Chou et al. (2017), adopted an "anonymization" model to guarantee privacy. Such syntactic models, however, can still leak information through composition attacks (Narayanan and Shmatikov, 2008; Dwork and Roth, 2014; Ganta et al., 2008). Avraam et al. (2021) later pro-

posed probabilistic anonymization, which introduces random noise to individual values for creating privacy-preserving visualizations. While conceptually related to DP, these methods do not quantitatively measure privacy risk.

Algorithmic approaches to differentially private visualization. Differential privacy provides formal guarantees even in the presence of side information and has become the gold standard for privacy protection. Zhang et al. (2016) were among the first to systematically examine the challenges of privacy-preserving data visualization. Using two-dimensional location data as an example, they constructed grids over the data domain and applied the Laplace mechanism to compute noisy counts for each grid cell. Their work analyzed visual artifacts introduced by noise and proposed ways to communicate uncertainty in private visual outputs. Building on this work, Zhang et al. (2020) extended the Laplace mechanism to various visualization types, including bar charts, and designed strategies to adjust the amount of noise for improving visualization utility. This approach mitigates distortion while maintaining privacy guarantees. In a related study, Lee (2017) qualitatively examined the effects of Laplacian noise on various visualization types, such as bar graphs, pie charts, heatmaps, linear plots, and scatterplots. Using smart grid electricity consumption time-series data, the study explored how DP constraints affect visual cues and identified pre- and post-processing methods for restoring them.

Evaluation of visual perception and utility under DP. Panavas et al. (2023) focused on evaluating the perceptual utility of differentially private scatterplots. Their comparative study showed that factors such as algorithm choice (e.g., AHP, DAWA), privacy level, bin number, data distribution, and user task all influence the visual outcome. To address the lack of guidance in balancing these factors, they had experts evaluate 1200 DP scatterplots to assess their effectiveness in preserving aggregate patterns. The results established a ground truth for visual utility and identified multi-scale structural similarity (MS-SSIM) as an effective objective metric for optimizing parameter selection in DP scatterplots. Liu et al. (2019) examined the common belief that aggregating individual data into composite representations such as heatmaps protects privacy. However, their analysis of noise-free heatmaps using eye-tracking data showed that they do not guarantee privacy. To address this, they proposed two noise mechanisms that ensure privacy and analyzed their privacy-utility tradeoff. Their results indicated that the Gaussian noise

mechanism, compared to the Laplace mechanism, provides a more effective solution for preserving privacy in heatmaps. These findings have significant implications for the development of differentially private mechanisms in the eye-tracking domain.

Extensions to specialized visualization types. More recent work, motivated by Liu et al. (2019) and Steil et al. (2019), has looked at specific visualizations such as heatmaps and boxplots. Ghazi et al. (2023) presented an efficient DP algorithm for heatmaps. The core of their algorithm was a DP procedure that took a set of distributions as input and produced an output close to the average of the inputs in terms of earth mover's distance (EMD). EMD better captures perceptual similarity compared to other measures such as ℓ_1, ℓ_2, or Kullback-Leibler (KL) divergence. They also proved theoretical bounds on the error of their algorithm under a sparsity assumption, demonstrating that these bounds were near-optimal. Ramsay and Diaz-Rodriguez (2024) introduced a differentially private boxplot and evaluated its effectiveness in displaying location, scale, skewness, and tails of a given empirical distribution. They theoretically demonstrated that the location and scale of the boxplot are estimated with optimal sample complexity, and skewness and tails are estimated consistently. This is not always the case for a boxplot naively constructed from a single existing differentially private quantile algorithm.

System-level frameworks for DP visualization. Interactive systems for private data visualization have been proposed by Nanayakkara et al. (2022) and Budiu et al. (2022), focusing on the relationship between privacy budget and accuracy to support *data privacy management* rather than visualization itself. Zhou et al. (2022) introduced the PriVis model, a Bayesian network-based differential privacy framework, to generate private scatter, line, and bar charts while preserving user-preferred patterns, addressing a limitation of prior work, which often targeted specific visualization types or algorithms without systematically considering the importance of data patterns in the visual output. They also developed the DPVisCreator system to assist data custodians in implementing this approach and evaluated its effectiveness through quantitative analysis of pattern utility, case studies, and semi-structured expert interviews.

Limitations of prior work and motivation for our study. Despite these advances, prior work mainly focused on generic visualization types rather than

domain-specific cases such as brain functional connectivity, which may require different techniques. Most studies do not systematically compare DP mechanisms to minimize required noise, and pre- and/or post-processing methods often lack structured workflows, limiting their applicability to neuroimaging. Evaluations also rely heavily on qualitative assessment without exploring potential quantitative approaches to measure visual utility. In contrast, our work focuses on functional connectivity in neuroimaging and provides a systematic evaluation of DP mechanisms, composition, and pre- and/or post-processing methods.

2.3. Functional Connectivity in Neuroimaging

In this work, we focus on neuroimaging data, particularly functional magnetic resonance imaging (fMRI), due to its sensitivity, high dimensionality, and collaborative use across institutions. fMRI data can reveal aspects of an individual's cognitive state, making privacy protection critical. Differential privacy offers a principled framework for secure data sharing in such settings. Furthermore, the variety of visualization modalities employed in fMRI studies enables a comprehensive examination of privacy-preserving visualization techniques, with insights that can extend to other neuroimaging modalities and scientific domains.

Resting-state functional MRI (rs-fMRI) captures the blood oxygenation level dependent (BOLD) signal while subjects remain in the scanner without performing specific tasks. A series of MRI images is acquired over time, with each image recording BOLD signal intensity at each voxel. We denote the BOLD value for voxel v at time t as $S(v,t)$. Typical MRI images contain roughly $100,000$ voxels, which can be grouped into regions of interest (ROIs) using various parcellation schemes (Tzourio-Mazoyer et al., 2002; Craddock et al., 2012; Gordon et al., 2016; Du et al., 2020). ROI timeseries are obtained by averaging voxel-level signals within each ROI: $S(r,t) = \frac{1}{L}\sum_{i=1}^{L} S(v_i,t)$, where r is the ROI and L is the number of voxels in that ROI. Analyses at the voxel and ROI levels provide complementary insights into brain functional connectivity, motivating diverse visualization approaches to effectively present these results.

FNC quantifies correlations between activity in different brain regions (Mohanty et al., 2020). Various methods exist to compute these correlations (Marrelec et al., 2006; Hlinka et al., 2011); in this work, we typically use the Pearson correlation coefficient, which ranges from -1 to 1 and measures the linear relationship between ROI-based timeseries. For any two ROIs, the

correlation is given by

$$X_{ij} = \frac{\sum_{t=1}^{T} S(r_i, t) S(r_j, t)}{\sqrt{\sum_{t=1}^{T} S^2(r_i, t)} \sqrt{\sum_{t=1}^{T} S^2(r_j, t)}},$$

where T denotes the number of timepoints and $S(\cdot, t)$ is assumed to be centered at 0. The FNC for a subject n across r brain regions is represented by an $r \times r$ adjacency matrix $X^{(n)}$, with each entry corresponding to the correlation between a pair of regions. We focus on visualizing the average FNC across a group of N subjects (see Section 3.1), defined as

$$\bar{X} = \frac{1}{N} \sum_{n=1}^{N} X^{(n)}.$$

For visualization, the elements of $\bar{X}$ are normalized to the range $[0, 1]$, quantized to a set of color values $\mathcal{C}$, and mapped to RGB triples (R_{ij}, G_{ij}, B_{ij}). In addition to visualizing the average FNC matrix, a connectogram can be used to represent FNC results graphically, with nodes as brain regions and edges indicating connection strength (see Section 3.2).

While FNC captures large-scale network interactions using ROI-based timeseries, SBC provides detailed information about the connectivity of a specific seed region with the rest of the brain. SBC involves selecting a seed voxel or ROI and computing its correlation with the timeseries of all other brain voxels. The resulting connectivity patterns are typically visualized as SBC spatial maps (Joel et al., 2011) (see Section 4).

2.4. Challenges in DP Functional Connectivity Visualization

There are (at least) three potential issues to consider when using differential privacy for visualization: the mismatch between the specific computations needed in the visualization and existing (generic) DP methods, the issues of multiple data accesses coming from applications in exploratory data analysis, and the challenge of defining the utility of data visualizations (Bhattacharjee et al., 2020; Zhang et al., 2016; Zhou et al., 2022). When thinking about visualization, the notion of "utility" can differ from standard quantitative metrics. In some scenarios, a private visualization has high utility if an expert would draw the same qualitative conclusions as from the non-private (unprotected) visualization.

To address these challenges, we take the following approaches. For the computation mismatch between specific visualization tasks and existing generic DP methods, we adapt DP mechanisms specifically for FNC to provide a private estimate of the correlation matrix underlying the visualization. In the case of the connectogram, we introduce a DP workflow designed for group comparisons, and for SBC visualization, we propose a DP procedure to systematically identify and present statistically significant results. To manage multiple data accesses, we employ (adaptive) composition techniques (Kairouz et al., 2015; Dong et al., 2020, 2022; Mironov, 2017) to quantify and control cumulative privacy risk. Finally, to address the challenge of utility, we provide both qualitative assessments and quantitative measurements.

In addressing these issues, we outline a systematic process for designing privacy-preserving visualizations across different visualization types and application domains, emphasizing both visualization quality and data privacy. Although our focus is on neuroimaging, the *process* we outline can be extended to other domains. Visualization typically involves computing or aggregating statistics from raw data through multiple steps. The first step is to identify the statistics or aggregated quantities needed for visualization. Next, suitable pre-processing strategies are applied to reduce the noise required for privacy protection. Different visualization types may benefit from different pre-processing techniques. We then design the differentially private mechanism, which may involve selecting appropriate noise distributions, applying composition methods, or constructing privacy-preserving workflows. Finally, post-processing methods are used to mitigate the effects of perturbation. Following this process, we demonstrate how privacy-preserving visualizations are constructed in neuroimaging (see Sections 3 and 4).

3. FNC Visualization

3.1. FNC Matrix Visualization

In this section, we describe the privacy-preserving visualization process for the average FNC matrix, summarized in Table 2, together with pre- and post-processing techniques that enhance visual utility.

3.1.1. Matrix-level Differentially Private Algorithms

In FNC matrix visualization, the true outputs are the RGB triples corresponding to each pairwise correlation in the average matrix. To guarantee

Algorithm	Name	Strategy	Mechanism	Composition
matrix-level	Pure	add perturbation	Laplace	—
	Approx	add perturbation	Gaussian	—
	Exp	random mapping	Exponential	—
entry-level	$\mathsf{Pure_k}$	add perturbation	Laplace	exact
	$\mathsf{Pure_r}$	add perturbation	Laplace	RDP-based
	$\mathsf{Approx_k}$	add perturbation	Gaussian	exact
	$\mathsf{Approx_r}$	add perturbation	Gaussian	RDP-based
	$\mathsf{Exp_k}$	random mapping	Exponential	exact

Table 2: A summary of the mechanisms explored in this paper is provided. Although other noise distributions may perform better for specific tasks, we focus on Laplace and Gaussian mechanisms, which are the most widely studied, for clarity of exposition.

DP, we adopt the following strategy: first compute the average FNC matrix under DP, then apply standard quantization and color mapping. This strategy falls under *output perturbation*, where noise is added directly to the average FNC matrix:

$$\tilde{X} = \bar{X} + Z,$$

where the noise distribution Z is determined by the privacy parameters. After perturbation, the noisy outputs $\tilde{X}$ are truncated to the valid range $[-1, 1]$. We consider two mechanisms:

- Pure: $(\epsilon, 0)$-DP Laplace mechanism. Let $Z \sim \text{Laplace}(0, \Delta_1/\epsilon \cdot I)$, where $\Delta_1 = r(r-1)/N$ is the ℓ_1 sensitivity of $\bar{X}$.
- Approx: (ϵ, δ)-DP analytic Gaussian mechanism. Let $Z \sim \mathcal{N}(0, \sigma^2 \cdot I)$, where σ is the minimum standard deviation that satisfies the condition in (Balle and Wang, 2018, Theorem 8), based on the ℓ_2 sensitivity $\Delta_2 = \sqrt{2r(r-1)/N}$ of $\bar{X}$.

3.1.2. Entry-level Differentially Private Algorithms

For the matrix-level methods, we first calculate the overall sensitivity of the FNC matrix and then generate the corresponding perturbed matrix. More efficient approaches may be obtained by applying composition methods at the entry level. Given total privacy parameters (ϵ, δ) and sequence size $k = r(r-1)/2$, we consider the following variants:

- $\mathsf{Pure_k}$: We compute the per-entry privacy parameters $(\epsilon', 0)$ using exact composition and apply the Laplace mechanism to each entry with sensitivity $\Delta = 2/N$. Hence, $\tilde{X}_{ij} = \bar{X}_{ij} + Z_{ij}$, where $Z_{ij} \sim \text{Laplace}(0, \Delta/\epsilon')$. For one-dimensional outputs, the global ℓ_1 and ℓ_2 sensitivities are identical, so we use Δ to denote the sensitivity without distinction between norms.

- $\mathsf{Pure_r}$: Given the overall privacy guarantee, we compute the scale parameter b of the Laplace distribution using RDP-based composition. Then, $\tilde{X}_{ij} = \bar{X}_{ij} + Z_{ij}$, where $Z_{ij} \sim \text{Laplace}(0, b)$.

- $\mathsf{Approx_k}$: We compute the per-entry privacy parameters (ϵ', δ') using exact composition and apply the analytic Gaussian mechanism with sensitivity $\Delta = 2/N$. Thus, $\tilde{X}_{ij} = \bar{X}_{ij} + Z_{ij}$, where the variance of Z_{ij} satisfies the inequality in (Balle and Wang, 2018, Theorem 8).

- $\mathsf{Approx_r}$: We compute the variance σ^2 of the Gaussian distribution using RDP-based composition under the given privacy parameters. Then, $\tilde{X}_{ij} = \bar{X}_{ij} + Z_{ij}$, where $Z_{ij} \sim \mathcal{N}(0, \sigma^2)$.

Visualization often serves as a preparatory step in research or exploratory data analysis, which may require accessing raw data multiple times and generating visualizations from different intermediate statistics. In such cases, (adaptive) composition plays a critical role. The matrix-level Laplace mechanism provides a strong "pure" privacy guarantee ($\delta = 0$), but by allowing a slightly weaker guarantee ($\delta > 0$), we can enhance the privacy budget through composition methods, reducing the level of noise required. For the Gaussian mechanism, we compare the performance of matrix-level and entry-level analytic Gaussian mechanisms.

We also explored the Exponential mechanism (McSherry and Talwar, 2007), which performs a randomized mapping from $\bar{X}$ to the set of all RGB images $\mathcal{H}$, either as a whole (denoted Exp) or entry-by-entry (denoted $\mathsf{Exp_k}$). The Exp method provides $(\epsilon, 0)$-DP by sampling $H \in \mathcal{H}$ according to the utility function $u(\bar{X}, H) = -\|\bar{X} - H\|_2$, but this is computationally infeasible due to the size of $\mathcal{H}$. The entry-level approach $\mathsf{Exp_k}$ applies an $(\epsilon', 0)$ mechanism with exact composition to map each $\bar{X}_{ij}$ to a color $C \in \mathcal{C}$ using the utility function $-|\bar{X}_{ij} - C|$ with sensitivity $2/N$. However, this approach yielded substantially poorer visual quality compared to output perturbation–based methods.

Algorithm 1 Method for choosing the clipping bound b

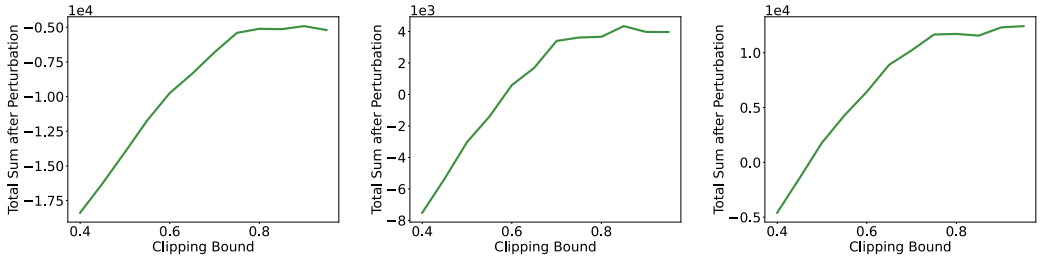
Input: A group of FNC matrices $X^{(1)}, X^{(2)}, ..., X^{(N)}$, privacy budget ϵ_b for each possible bound
Output: A list of $\{\tilde{s}_b\}$ for all possible b
for $b = 0.1, 0.2, ..., 1$ **do**
1. $s_b = 0$
2. **for** $i = 1, ..., r; j > i; n = 1, ..., N$ **do**
   If $|X_{ij}^{(n)}| \leq b$: $s_b = s_b + X_{ij}^{(n)}$;
   If $X_{ij}^{(n)} > b$ or $X_{ij}^{(n)} < -b$: $s_b = s_b + b$ or $s_b = s_b - b$, respectively.
3. $\tilde{s}_b = s_b + \text{Laplace}(0, br(r-1)/\epsilon_b)$

3.1.3. Pre-processing Approaches

Given privacy parameters ϵ and δ, some pre-processing approaches can help reduce the required noise. One effective approach is range reduction. Prior information of FNC suggest that extremely high correlations (e.g., $|\bar{X}_{ij}| \geq 0.9, i \neq j$) are rare. Thus, the range of $\bar{X}_{ij}$ can be constrained to $[-b, b]$, where $0 < b < 1$. Under this clipping, the ℓ_1 and ℓ_2 sensitivities of $\bar{X}$ are $\Delta_1 = br(r-1)/N$ and $\Delta_2 = b\sqrt{2r(r-1)}/N$, respectively, and each entry $\bar{X}_{ij}$ has sensitivity $\Delta = 2b/N$.

In this work, we employ a clipping method (Near and Abuah, 2021) to estimate the bounds of the FNC range (see Algorithm 1). The method establishes upper and lower bounds for $X_{ij}^{(n)}$ by gradually increasing the bound b from 0 to 1. If $X_{ij}^{(n)} > b$ or $X_{ij}^{(n)} < -b$, the value is clipped to b or $-b$, respectively. We then calculate the sum of all entries for each possible bound and apply the Laplace mechanism to ensure the privacy of the total sum. Finally, we select the bound that yields a stable perturbed sum (see Figure 3).

Alternative techniques can be used to estimate the bounds, such as leveraging prior domain knowledge, using public data, or applying different DP methods. For example, public datasets can help determine upper and lower bounds. Another approach is to estimate differentially private quantiles and use them as bounds (McSherry and Talwar, 2007; Smith, 2011). Liu et al. (2019) propose an algorithm that optimizes the expected mean squared error between perturbed aggregated data and its original counterpart to estimate bounds. Regardless of the method used, there is a tradeoff between information loss and the noise required for differential privacy. When the upper and



(a) FBIRN: $b = 0.75$ (b) BLSA: $b = 0.7$ (c) ADNI: $b = 0.75$

Figure 3: Clipping method for choosing b on three datasets. The total privacy budget allocated for bound estimation is 0.5. The search for b begins at 0.4 rather than 0, assuming that the bound of FNC values exceeds 0.4.

lower bounds are set closer to each other, less noise is needed to guarantee differential privacy due to a lower sensitivity. However, using smaller clipping bounds results in the removal of a significant amount of information.

3.1.4. Post-processing Approaches

Differential privacy has a post-processing invariance property, meaning that any data-independent operations on differentially private outputs do not incur additional privacy loss. Hence, post-processing can be used to "clean up" images without compromising privacy. This can help mitigate the effects of perturbation on FNC or any other visualizations and potentially improve visual utility.

For FNC visualization, we explore two post-processing approaches: Singular Value Decomposition (SVD) and Haar wavelet transform (Mulcahy, 1997). The SVD of the differentially private output is $\tilde{X} = U\Sigma V^{\intercal}$. We select the first s components after examining the magnitudes of the singular values. If the added noise does not destroy the visualization pattern, we are able to observe that the first s singular values are significantly larger than the remaining ones. For visualization, we use $\tilde{X}_s = U_s\Sigma_s V_s^{\intercal}$ after post-processing. In the Haar wavelet transform, we apply a transformation to obtain the matrix $W = P^{\intercal}\tilde{X}P$, where P is the invertible Haar transformation matrix. In many cases, the wavelet-transformed matrices are sparser than their original counterparts. We then threshold the coefficients in matrix W to generate a new coefficient matrix, denoted as W_w. Finally, we perform the inverse wavelet transform on W_w to obtain $\tilde{X}_w = (P^{\intercal})^{-1}W_w P^{-1}$, which represents the processed data used for visualization.

3.2. Connectogram Visualization

3.2.1. Non-private Workflow for Group Comparison

Connectogram visualizations effectively represent group differences in brain regions and functional connectivity, such as comparisons between healthy controls (HC) and schizophrenia (SZ) patients. The standard workflow for generating a connectogram first transforms each element of the FNC matrix into Fisher's Z-scores. The statistical significance of connectivity strength differences between HC and SZ groups is then assessed using a two-tailed two-sample t-test ($p < 0.05$ with Bonferroni correction) (Du et al., 2020).

3.2.2. Private Workflow for Group Comparison

We propose a novel differentially private workflow for connectogram visualization. A straightforward approach might be to add perturbation directly to the p-values or t-statistics. However, this is challenging because the sensitivity of the t-statistic is complex to compute, and added noise can substantially alter the p-values. Our experiments indicate that, under differential privacy, nonparametric tests (Couch et al., 2019) are more effective for constructing connectograms. Consequently, we adopt a private workflow based on the nonparametric absolute Kruskal-Wallis (AKW) test.

Consider a database of size N partitioned into g groups. Let N_i denote the size of group i, and r_{ij} the rank of the j-th element in group i. Define the mean rank of group i as $\bar{r}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} r_{ij}$, and the overall mean rank as $\bar{r} = \frac{N+1}{2}$. The absolute Kruskal-Wallis h-statistic is given by

$$h = (N-1)\frac{\sum_{i=1}^{g} N_i|\bar{r}_i - \bar{r}|}{\sum_{i=1}^{g}\sum_{j=1}^{N_i}|r_{ij} - \bar{r}|}$$

and has a sensitivity of 8 (Couch et al., 2019, Algorithm $\widetilde{\mathrm{KW}}_{\mathrm{P}}$).

Our workflow (see Figure 4) proceeds as follows: (1) Compute the AKW test on FNC matrices from two groups to obtain an h-statistic matrix. (2) For each h-statistic submatrix, which is partitioned by brain domains (Du et al., 2020, Table 2), calculate the average h-statistic. Apply the Report Noisy Max (RNM) algorithm (Dwork, 2006) to select the top p_d submatrices in a differentially private manner. The RNM algorithm is a differential privacy technique for private selection. Instead of selecting the top-ranked items based on their true values, RNM adds calibrated noise to each value and selects the top-ranked items based on the noisy values. The number of

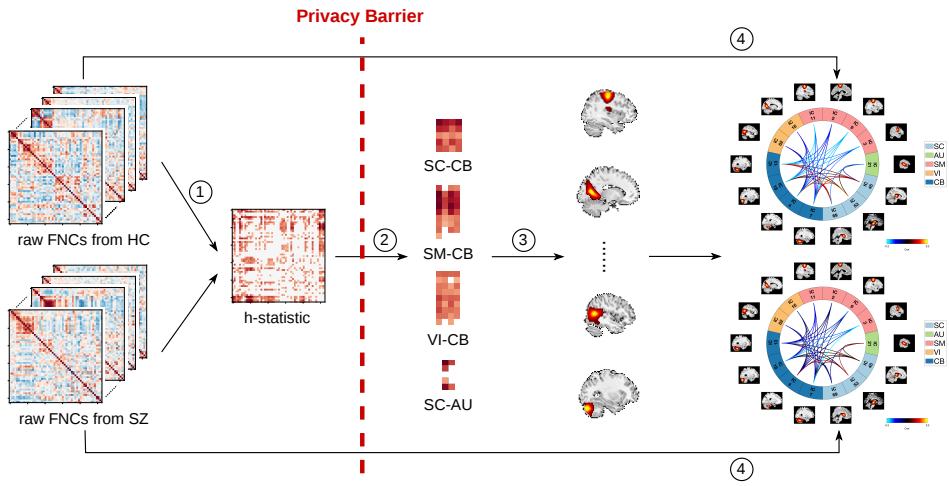


Figure 4: Privacy-preserving workflow for connectogram visualization. The privacy barrier refers to the stage where sensitive data (e.g., FNC values, h-statistics) are perturbed using differential privacy techniques. Data after the privacy barrier is considered privacy-preserving, and it is impossible to infer the true data from this released information.

selected items is determined by the privacy budget allocated at this stage. (3) Average the h-statistic for each brain region within the selected brain domains and apply RNM again to select the top p_r brain regions privately. (4) Finally, compute perturbed connectivities from the average FNC of each group. With privacy-preserving nodes and edges, the connectogram for group comparisons can be visualized.

4. SBC Visualization

4.1. Non-private Workflow for Significant Connectivity

SBC characterizes the connectivity pattern between a predefined seed and the entire brain. This involves computing Fisher-transformed bivariate correlation coefficients between the seed BOLD timeseries and the timeseries of each individual voxel (Joel et al., 2011). The correlation between the seed and any voxel v is defined as

$$\frac{\sum_{t=1}^{T} S(\cdot,t)S(v,t)}{\sqrt{\sum_{t=1}^{T} S^2(\cdot,t)}\sqrt{\sum_{t=1}^{T} S^2(v,t)}},$$

where $S(\cdot, t)$ represents the BOLD timeseries from the seed (either a single voxel or an ROI) at time t, and T is the number of timepoints in the experiment. The seed timeseries $S(\cdot, t)$ is assumed to be mean-centered. The correlation coefficient is then converted to a z-score using the Fisher Z-transform. A one-sample t-test is subsequently performed to determine whether the mean connectivity across subjects significantly differs from zero, indicating significant positive or negative connectivity.

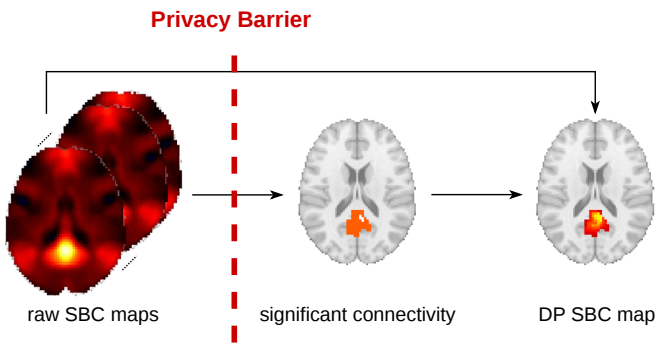


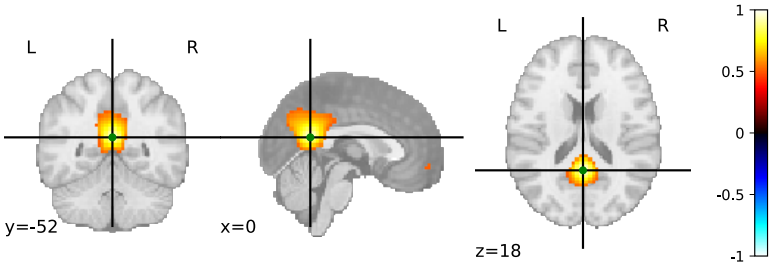
Figure 5: Privacy-preserving workflow for SBC map visualization.

4.2. Private Workflow for Significant Connectivity

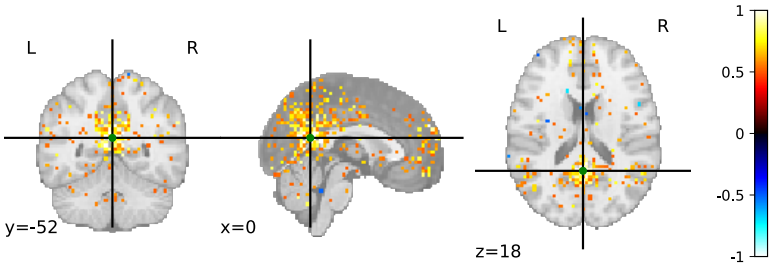
The privacy-preserving SBC visualization process is quite similar to connectogram visualization and also involves two main components (see Figure 5): selecting significant connectivities in a privacy-preserving manner and displaying z-scores with differential privacy. To identify statistically significant connectivities, we employ the Wilcoxon signed-rank (WSR) test, a nonparametric alternative to the one-sample t-test, as nonparametric tests are often more effective than their parametric counterparts. For a given database, let s_i denote the sign of the data x_i, and let r_i represent the rank of $|x_i|$, determined by the magnitude of the data. The WSR test statistic is then calculated as $w = \sum s_i r_i$. Connectivities are selected at a significance level of $p < 0.05$ after applying Bonferroni correction (Couch et al., 2019, Algorithm $\widetilde{\mathrm{WP}}_{\mathrm{P}}$).

To protect the z-scores, we use the analytic Gaussian mechanism to generate noise, which is added to the correlations before applying the Fisher Z-transform. Although narrowing the range of correlations could reduce the

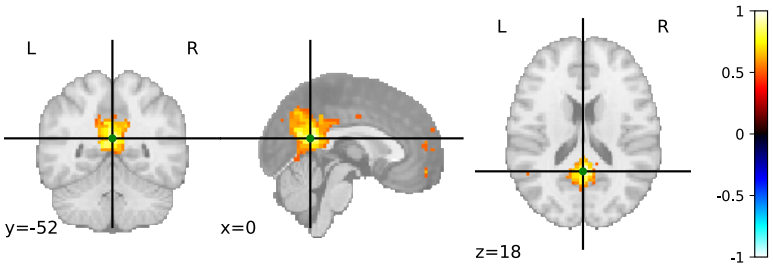
required noise, this approach is not suitable here because high correlations are expected (see Figure 6a). The privacy-protected correlation map exhibits noticeable salt-and-pepper noise (see Figure 6b), which is absent in FNC visualization. One likely reason is that connections within the same brain network tend to be strong, resulting in relatively stable and similar correlation values. To mitigate this noise, we apply a median filter (Erkan et al., 2018) as a post-processing method (see Figure 6c). After identifying significant connectivities and their corresponding values, we construct the privacy-preserving SBC map.



(a) Original SBC



(b) DP SBC



(c) DP SBC after applying a median filter

Figure 6: SBC maps on the NITRC dataset using the Posterior Cingulate Cortex (PCC) as the seed (indicated by the green dot), with parameters $N = 1082$, $\epsilon = 2.5$, and $\delta = N^{-3/2}$. Correlations with absolute values greater than 0.5 are displayed.

5. Experiments and Results

We aim to investigate the following research questions: (1) Which algorithm most effectively preserves functional connectivity patterns under the same level of privacy? (2) Can pre- and post-processing methods enhance visualization quality? (3) Does the differentially private connectogram workflow identify significant brain regions and connections as effectively as the non-private workflow? (4) Is the differentially private workflow for SBC visualization as effective as its non-private counterpart in identifying significant connectivities? (5) How do numerical utility measures, such as mean squared error (MSE), relate to visualization quality?

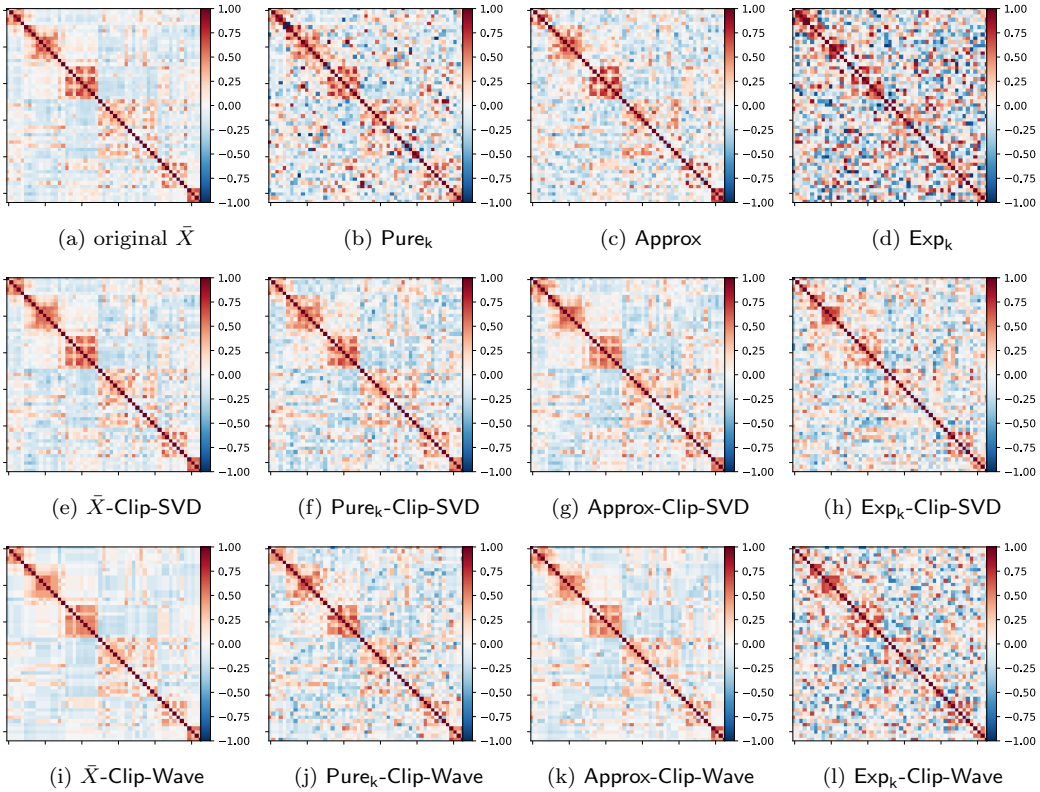
We conducted privacy-preserving FNC visualization experiments using datasets from the Function Biomedical Informatics Research Network (FBIRN), the Baltimore Longitudinal Study of Aging (BLSA), and the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Du et al., 2020; Sendi et al., 2021; Ferrucci, 2008). Our findings were consistent across all datasets. For connectogram visualizations, we used the FBIRN dataset to compare HC and SZ groups. For SBC visualizations, we employed the NITRC dataset.

5.1. Methods Comparison

In general, we found that Gaussian mechanisms preserved visual utility better than Laplace mechanisms, which in turn outperformed Exponential mechanisms (see Figure 7, Figure A.3, and Figure A.4). Overall, Approx-Clip-SVD achieved the best performance.

This superior performance can be explained from both structural and theoretical perspectives. At a structural level, the privacy-preserving FNCs produced by this method retain the same block-structured organization seen in the original matrix, with recognizable within-network modules along the diagonal and comparatively weaker cross-network structure off-diagonal. The method preserves the relative ordering of strong versus weak connectivity patterns across domains, so the same major network-network relationships remain visually interpretable even though individual entries show slight deviations. In contrast, less effective DP examples introduce noise that partially blurs or fragments these modules, reducing contrast and making the within-network blocks less coherent, which can obscure higher-level structure.

From a theoretical standpoint, because the ℓ_2 sensitivity is lower than the ℓ_1 sensitivity in high dimensions, the analytic Gaussian mechanism adds less noise than the Laplace mechanism, resulting in better performance in



(a) original $\bar{X}$ (b) $\mathsf{Pure}_{\mathsf{k}}$ (c) Approx (d) $\mathsf{Exp}_{\mathsf{k}}$

(e) $\bar{X}$-$\mathsf{Clip\text{-}SVD}$ (f) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$ (g) $\mathsf{Approx\text{-}Clip\text{-}SVD}$ (h) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$

(i) $\bar{X}$-$\mathsf{Clip\text{-}Wave}$ (j) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$ (k) $\mathsf{Approx\text{-}Clip\text{-}Wave}$ (l) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$

Figure 7: Visualization of the FBIRN dataset with $N = 1056$, clipping bound $b = 0.75$, number of components used in SVD set to 10, threshold used in the wavelet transform set to 0.2, and number of colors set to 256. The first row represents methods without pre- or post-processing, the second row represents methods with clipping and SVD post-processing, and the third row represents methods with clipping and wavelet post-processing. The overall privacy guarantee for all perturbed FNC matrices is $(\epsilon = 2, \delta = N^{-3/2})$-DP. As expected, the Laplace mechanism performs worse than the analytic Gaussian mechanism since the ℓ_1 sensitivity is larger than the ℓ_2 sensitivity in high dimensions. Even with composition methods applied to reduce noise, $\mathsf{Pure}_{\mathsf{k}}$ still does not outperform Approx. The Exponential mechanism is barely usable. Clipping and SVD post-processing significantly enhance the visual fidelity of private FNC matrices, whereas wavelet post-processing introduces additional visual distortion. This suggests that classical image processing techniques developed for natural images (e.g., photos) may not be suitable for this type of visualization.

terms of pattern preservation. The Exponential mechanism exhibits the worst visualization performance due to its high level of randomness. Specifically, the PDF of the output from $\mathsf{Pure}_{\mathsf{k}}$ for each entry is proportional to

$\exp(-N\epsilon'|z_{ij}-\bar{X}_{ij}|/2)$, while the PDF of the output from $\mathsf{Exp_k}$ is proportional to $\exp(-N\epsilon'|c-\bar{X}_{ij}|/4)$. In general, the Exponential mechanism is more suitable for tasks such as determining product prices, where small perturbations in the determined price may lead to significant changes in revenue (utility). However, for FNC matrices, small changes in correlation do not substantially affect color assignment.

In terms of composition, exact composition and RDP-based composition outperform advanced composition, regardless of the mechanism used. For the Laplace mechanism, $\mathsf{Pure_k}$ requires less noise than $\mathsf{Pure_r}$, whereas for the analytic Gaussian mechanism, $\mathsf{Approx_k}$ requires more noise.

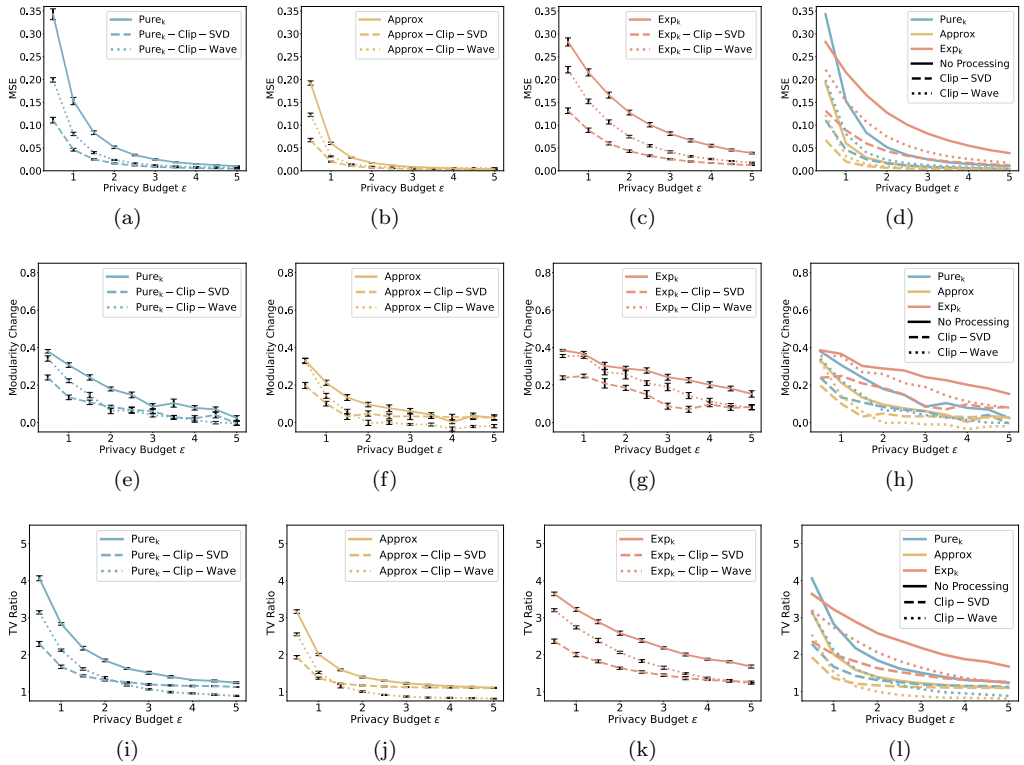


Figure 8: Quantitative evaluation of privacy-preserving FNC visualization on the FBIRN dataset. The first row reports MSE between the original and private average FNC matrices. The second row shows modularity change, defined as $Q_{\text{orginal}} - Q_{\text{DP}}$, where Q denotes signed Louvain modularity. The third row presents the TV ratio, computed as $\text{TV}_{\text{DP}}/\text{TV}_{\text{orginal}}$. Columns correspond to (a)(e)(i) $\mathsf{Pure_k}$-, (b)(f)(j) Approx-, (c)(g)(k) $\mathsf{Exp_k}$-related methods, and (d)(h)(l) all methods. Each configuration is repeated 50 times, and error bars are provided.

5.2. Processing Effect

We find that pre- and post-processing significantly enhance the visual utility of privacy-preserving data visualization. For FNC visualization, Figure 7g and 7k use different processing methods, and both exhibit the best visualization results, i.e., they are visually closest to the original average FNC shown in Figure 7a.

To quantitatively assess this observation, we first compute the MSE between the original average FNC and its private counterpart. As shown in Figure 8 (a)-(d), Figure A.1, and Figure A.2, the benefits of pre- and post-processing are more pronounced when the total privacy budget ϵ is small (high privacy) but become less significant as ϵ increases. The post-processing method based on SVD consistently achieves better performance than the wavelet transform.

While MSE captures pointwise numerical deviations, it does not reflect higher-level structural organization. Since FNC matrices are interpreted through block-structured functional networks, we additionally evaluate network modularity (Rubinov and Sporns, 2011). Modularity measures the strength of community structure relative to a random null model. A reduction in modularity after privacy perturbation indicates that intrinsic functional modules become less distinguishable, reflecting structural degradation.

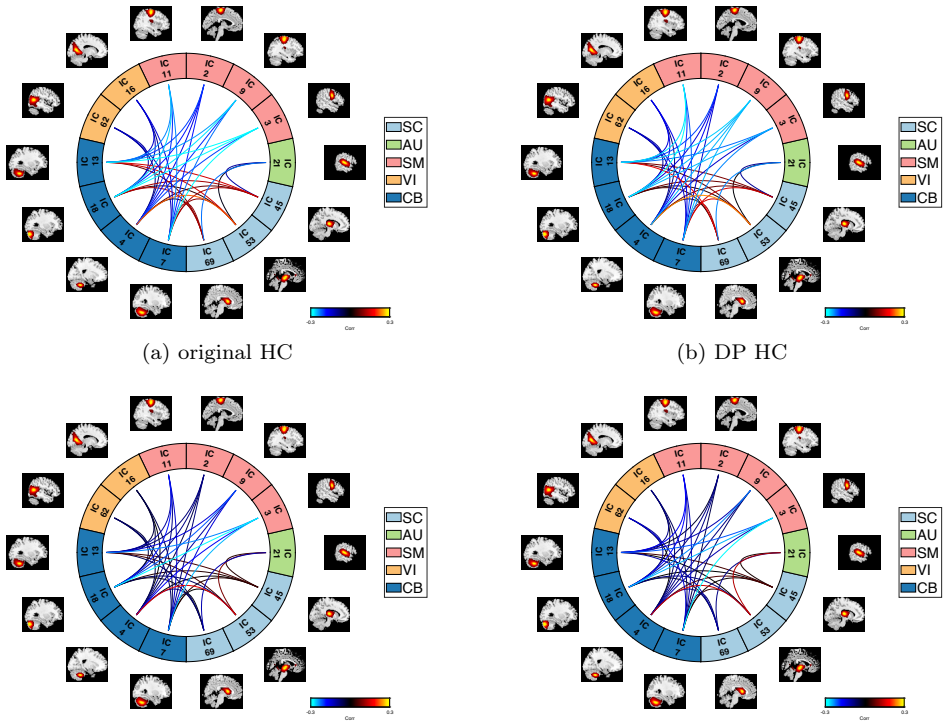
To further quantify local fluctuation patterns introduced by privacy noise, we compute the total variation (TV) (Michel et al., 2011) of the connectivity matrix and report the TV ratio between the private and original matrices. TV measures the aggregate magnitude of local differences between adjacent matrix entries, thereby characterizing high-frequency perturbations. A higher TV ratio indicates increased local irregularity and reduced visual smoothness.

As shown in Figure 8, Figure A.1, and Figure A.2, although SVD yields lower MSE across privacy budgets, wavelet-based post-processing better preserves structural properties and visual smoothness at larger ϵ. This divergence indicates that numerical, structural, and visual fidelity are not necessarily aligned.

For SBC visualization, pre-processing methods are not applicable. We apply the median filter technique as the post-processing method, resulting in a noticeable improvement in visualization quality (see Figure 6). These results suggest that different visualization types may require distinct pre- and/or post-processing methods.

5.3. Workflow Effectiveness

For connectogram visualization, we observe that our proposed workflow produces visualizations similar to its non-private counterpart (see Figure 9). Our work also finds that brain functional abnormalities in SZ are primarily located in the subcortical (SC), auditory (AU), sensorimotor (SM), visual (VI), and cerebellar (CB) domains Du et al. (2020). Given the importance of identifying these brain domains and regions rather than determining the exact connectivity values, we allocate more privacy budget to identify the brain domains and regions. Similar outcomes are observed in SBC visualization, where the privacy-preserving SBC map closely aligns with the SBC map without privacy protection. For instance, Figure 10 illustrates strong connectivities between the selected seed PCC and its neighboring voxels.



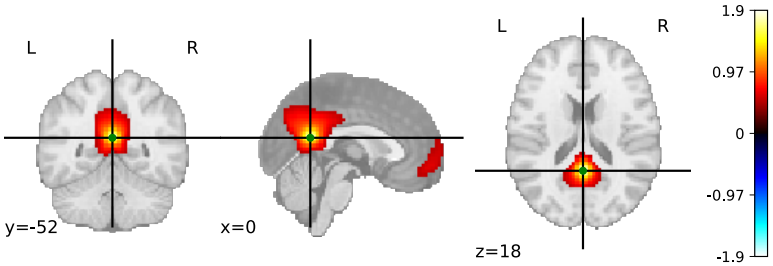
(a) original HC (b) DP HC

(c) original SZ (d) DP SZ

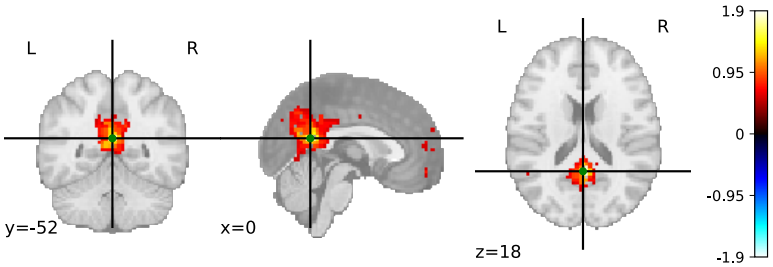
Figure 9: Privacy-preserving connectogram visualization with $\epsilon = 2$, $p_d = 5$, and $p_r = 14$. We allocate 45% of the privacy budget to identify brain domains, another 45% to pinpoint brain regions, and reserve the remaining 10% to protect the connectivities.

5.4. Visual Utility

It should be noted that there is no standard method for quantitatively assessing the quality of privacy-preserving visualizations. As discussed by Zhang et al. (2016), making a precise notion of "visual utility" is challenging. Errors are often measured using metrics such as the ℓ_1 or ℓ_2 norms, but these may not align with perceived quality. Moreover, in many cases, norm-based metrics may not be suitable for evaluating visualization quality. For example, in connectogram visualization, hypothesis testing errors (seeing irrelevant domains or regions, or failing to notice significant ones) may provide a more relevant evaluation approach, as the primary concern is identifying domains or regions that exhibit significant differences in group comparisons. When it comes to SBC visualization, identifying quantitative methods to assess visual utility becomes even more challenging.



(a) Original SBC



(b) DP SBC after applying a median filter

Figure 10: SBC maps on the NITRC dataset using the Posterior Cingulate Cortex (PCC) as the seed (indicated by the green dot), with parameters $N = 1082$, $\epsilon = 5$, and $\delta = N^{-3/2}$. We allocate 50% of the privacy budget to identify significant correlations and 50% to protect the z-scores.

We surveyed neuroimaging researchers (at various levels of experience) through a convenience sample of researchers interested in functional network

connectivity. We designed a survey consisting of five sections and collected 22 responses, with 18 participants consenting to the release of their responses. The full questionnaire, participant responses, and a summary of survey results are provided in the supplementary material. Our findings from this initial survey are as follows:

- Algorithm comparison: Participants found Approx-based methods, especially those with clipping and SVD, most similar to the original. In contrast, wavelet-based versions, although achieving lower MSE, were frequently rated as more visually distorted. This suggests that quantitative metrics like MSE alone may not fully capture perceptual quality.

- Privacy–utility tradeoff: As the privacy budget ϵ increased, visual quality tended to improve, showing clearer structures and fewer distortions.

- Consistency under the same privacy level: Most participants felt that visualizations generated using the same algorithm and privacy level were generally similar, despite differences in noise realizations.

- Connectogram workflow evaluation: Participants observed that allocating more privacy budget to brain region (node) identification improved perceived visualization quality. Conversely, reducing the budget for nodes increased visual distortions. This effect was not primarily caused by node misidentification itself but rather because node errors propagated to edges, resulting in larger and more noticeable changes in the central visual area.

- SBC workflow evaluation: Participants generally considered the privacy-preserving visualization usable and effective at preserving core connectivity patterns. However, some noted reduced smoothness and missing functional connections, suggesting that a higher privacy budget may be needed to better retain critical connections.

We believe this initial feedback provides some guidance for future research directions. In particular, we believe that for privacy-preserving visualization to become part of the regular workflow in future software systems for neuroimaging research, it is important to assess the visual utility as judged by the intended audience for the visualization.

6. Future Directions

Data visualizations are often designed to provide qualitative assessments of quantitative data. We discussed the risks of sensitive information leakage in visualizations, highlighting the growing importance of privacy protection in this area. Current research has identified various privacy attacks across fields such as machine learning (Rigaki and Garcia, 2023). Adversaries may seek to access sensitive information, including the training dataset, model parameters, hyperparameters, and architecture, etc. Inspired by these concerns, we believe that privacy attacks in visualizations are also an area worth exploring.

When visualizing sensitive data, the gap between qualitative assessments and quantitative precision offers many interesting options for guaranteeing privacy. In this study, we explored several approaches to providing privacy for visualizations related to functional network connectivity. We view this as a *starting point* for considering privacy in data visualization within neuroimaging and human health research more broadly. In these settings, individual data holders may not have a large sample size but in systems using federated learning, the differences among these methods would be much smaller as they would require less noise and provide better utility due to the increased sample size.

We did not explore the full gamut of postprocessing methods or transform-based privacy algorithms. For example, alternative processing methods such as Fourier transforms and bit plane slicing could be explored. Another approach involves transforming data to obtain sparse coefficients, applying privacy-preserving techniques to protect them, and then using the inverse transform to reconstruct the privacy-preserving data. In addition to refining these approaches, manipulating the choice of color map, the number of colors $|\mathcal{C}|$, and the quantization scheme can all potentially improve the visual effect.

Many visualizations seek to represent the uncertainty in estimates or observations. Uncertainty visualization has long been recognized as a critical component for scientific communication, with numerous studies highlighting its importance (MacEachren, 1992; MacEachren et al., 2005; Potter et al., 2010). Visualizing uncertainty and providing privacy can potentially present synergies: designing effective algorithms is an interesting avenue for future work. A starting point would be to develop privacy-preserving displays for error bars and confidence intervals. When privacy is included, we might find that other visual representations of uncertainty are more visually salient.

Data and Code Availability

This study utilizes publicly available datasets accessible directly from their respective repositories; signing a Data Use Agreement (DUA) may be required. The code is available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest that may bias or could be perceived to bias this work.

Acknowledgments

The work of authors was supported in part by the US National Institutes of Health under awards 2R01DA040487, R01DA049238, and R01MH121246, as well as by the National Science Foundation under award 2112455. The work of Sarwate was also supported by the National Science Foundation under award CNS-2148104 and is supported in part by funds from federal agency and industry partners as specified in the Resilient & Intelligent NextG Systems (RINGS) program.

The authors wish to thank the FBIRN, BLSA, and ADNI teams. In addition, the authors acknowledge the Neuro Bureau, the ADHD 200 consortium, and Virginia Tech's ARC. The authors also acknowledge the assistance of ChatGPT for English language editing.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K., Zhang, L., 2016. Deep learning with differential privacy, in: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security, pp. 308–318.

Alghamdi, W., Asoodeh, S., Calmon, F.P., Gomez, J.F., Kosut, O., Sankar, L., Wei, F., 2022a. The saddle-point accountant for differential privacy. arXiv preprint arXiv:2208.09595 .

Alghamdi, W., Asoodeh, S., Calmon, F.P., Kosut, O., Sankar, L., Wei, F., 2022b. Cactus mechanisms: Optimal differential privacy mechanisms in

the large-composition regime, in: 2022 IEEE International Symposium on Information Theory (ISIT), IEEE. pp. 1838–1843.

Allen, C.T., 2020. Finding the Enemy on the Data-Swept Battlefield of 2035. Military Review .

Avraam, D., Wilson, R., Butters, O., Burton, T., Nicolaides, C., Jones, E., Boyd, A., Burton, P., 2021. Privacy preserving data visualizations. EPJ data science 10, 2.

Balle, B., Barthe, G., Gaboardi, M., Hsu, J., Sato, T., 2020. Hypothesis testing interpretations and renyi differential privacy, in: International Conference on Artificial Intelligence and Statistics, PMLR. pp. 2496–2506.

Balle, B., Wang, Y.X., 2018. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising, in: International Conference on Machine Learning, PMLR. pp. 394–403.

Bhattacharjee, K., Chen, M., Dasgupta, A., 2020. Privacy-preserving data visualization: reflections on the state of the art and research opportunities, in: Computer Graphics Forum, Wiley Online Library. pp. 675–692.

Budiu, M., Thaker, P., et al., 2022. Overlook: Differentially private exploratory visualization for big data. Journal of Privacy and Confidentiality 12.

Chopra, S., Labache, L., Dhamala, E., Orchard, E.R., Holmes, A., 2023. A practical guide for generating reproducible and programmatic neuroimaging visualizations. Aperture Neuro 3, 1–20.

Chou, J.K., Bryan, C., Ma, K.L., 2017. Privacy preserving visualization for social network data with ontology information, in: 2017 IEEE Pacific Visualization Symposium (PacificVis), IEEE. pp. 11–20.

Couch, S., Kazan, Z., Shi, K., Bray, A., Groce, A., 2019. Differentially private nonparametric hypothesis testing, in: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, pp. 737–751.

Craddock, R.C., James, G.A., Holtzheimer III, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Human brain mapping 33, 1914–1928.

Dong, J., Durfee, D., Rogers, R., 2020. Optimal differential privacy composition for exponential mechanisms, in: International Conference on Machine Learning, PMLR. pp. 2597–2606.

Dong, J., Roth, A., Su, W.J., 2022. Gaussian differential privacy. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 84, 3–37.

Doroshenko, V., Ghazi, B., Kamath, P., Kumar, R., Manurangsi, P., 2022. Connect the dots: Tighter discrete approximations of privacy loss distributions. arXiv preprint arXiv:2207.04380 .

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., Salman, M., Abrol, A., Rahaman, M.A., Chen, J., et al., 2020. NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders. NeuroImage: Clinical 28, 102375.

Dwork, C., 2006. Differential privacy, in: International colloquium on automata, languages, and programming, Springer. pp. 1–12.

Dwork, C., McSherry, F., Nissim, K., Smith, A., 2006. Calibrating noise to sensitivity in private data analysis, in: Halevi, S., Rabin, T. (Eds.), Theory of Cryptography, Springer, Berlin, Heidelberg. pp. 265–284. URL: http://dx.doi.org/10.1007/11681878_14, doi:doi:10.1007/11681878_14.

Dwork, C., Roth, A., 2014. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9, 211–407.

Dwork, C., Rothblum, G.N., Vadhan, S., 2010. Boosting and differential privacy, in: 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, IEEE. pp. 51–60.

Dwork, C., Smith, A., Steinke, T., Ullman, J., 2017. Exposed! a survey of attacks on private data. Annual Review of Statistics and Its Application 4, 61–84.

Dwork, C., Smith, A., Steinke, T., Ullman, J., Vadhan, S., 2015. Robust traceability from trace amounts, in: 2015 IEEE 56th Annual Symposium on Foundations of Computer Science, IEEE. pp. 650–669.

Erkan, U., Gökrem, L., Enginoğlu, S., 2018. Different applied median filter in salt and pepper noise. Computers & Electrical Engineering 70, 789–798.

Ferrucci, L., 2008. The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future.

Fiore, M., Katsikouli, P., Zavou, E., Cunche, M., Fessant, F., Le Hello, D., Aivodji, U.M., Olivier, B., Quertier, T., Stanica, R., 2020. Privacy in trajectory micro-data publishing: a survey. Transactions on Data Privacy 13, 91–149.

Ganta, S.R., Kasiviswanathan, S.P., Smith, A., 2008. Composition attacks and auxiliary information in data privacy, in: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA. p. 265–273. URL: https://doi.org/10.1145/1401890.1401926, doi:doi:10.1145/1401890.1401926.

Garfinkel, S., 2015. De-identification of Personal Information. US Department of Commerce, National Institute of Standards and Technology.

Geng, Q., Kairouz, P., Oh, S., Viswanath, P., 2015. The staircase mechanism in differential privacy. IEEE Journal of Selected Topics in Signal Processing 9, 1176–1184.

Ghazi, B., He, J., Kohlhoff, K., Kumar, R., Manurangsi, P., Navalpakkam, V., Valliappan, N., 2023. Differentially private heatmaps, in: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 7696–7704.

Ghazi, B., Kamath, P., Kumar, R., Manurangsi, P., 2022. Faster privacy accounting via evolving discretization, in: International Conference on Machine Learning, PMLR. pp. 7470–7483.

Gopi, S., Lee, Y.T., Wutschitz, L., 2021. Numerical composition of differential privacy. Advances in Neural Information Processing Systems 34, 11631–11642.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and evaluation of a cortical area parcellation from resting-state correlations. Cerebral cortex 26, 288–303.

Henriksen-Bulmer, J., Jeary, S., 2016. Re-identification attacks—A systematic literature review. International Journal of Information Management 36, 1184–1192.

Hlinka, J., Paluš, M., Vejmelka, M., Mantini, D., Corbetta, M., 2011. Functional connectivity in resting-state fMRI: is linear correlation sufficient? Neuroimage 54, 2218–2225.

Homer, N., Szelinger, S., Redman, M., Duggan, D., Tembe, W., Muehling, J., Pearson, J.V., Stephan, D.A., Nelson, S.F., Craig, D.W., 2008. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS genetics 4, e1000167.

Joel, S.E., Caffo, B.S., Van Zijl, P.C., Pekar, J.J., 2011. On the relationship between seed-based and ICA-based measures of functional connectivity. Magnetic resonance in medicine 66, 644–657.

Kairouz, P., Oh, S., Viswanath, P., 2015. The composition theorem for differential privacy, in: International conference on machine learning, PMLR. pp. 1376–1385.

Koskela, A., Honkela, A., 2020. Learning rate adaptation for differentially private learning, in: International Conference on Artificial Intelligence and Statistics, PMLR. pp. 2465–2475.

Koskela, A., Jälkö, J., Honkela, A., 2020. Computing tight differential privacy guarantees using fft, in: International Conference on Artificial Intelligence and Statistics, PMLR. pp. 2560–2569.

Kostelnick, C., 2008. The visual rhetoric of data displays: The conundrum of clarity. IEEE Transactions on Professional Communication 51, 116–130.

Lee, H.B., 2017. Visualization and differential privacy. Master's thesis. University of Illinois at Urbana-Champaign.

Liu, A., Xia, L., Duchowski, A., Bailey, R., Holmqvist, K., Jain, E., 2019. Differential privacy for eye-tracking data, in: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, pp. 1–10.

MacEachren, A.M., 1992. Visualizing uncertain information. Cartographic perspectives , 10–19.

MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., 2005. Visualizing geospatial information uncertainty: What we know and what we need to know. Cartography and Geographic Information Science 32, 139–160.

Marrelec, G., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., Doyon, J., Benali, H., 2006. Partial correlation for functional brain interactivity investigation in functional MRI. Neuroimage 32, 228–237.

McSherry, F., Talwar, K., 2007. Mechanism design via differential privacy, in: 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), IEEE. pp. 94–103.

Merener, M.M., 2012. Theoretical results on de-anonymization via linkage attacks. Trans. Data Priv. 5, 377–402.

Michel, V., Gramfort, A., Varoquaux, G., Eger, E., Thirion, B., 2011. Total variation regularization for fMRI-based prediction of behavior. IEEE Transactions on Medical Imaging 30, 1328–1340.

Mironov, I., 2017. Rényi differential privacy, in: 2017 IEEE 30th computer security foundations symposium (CSF), IEEE. pp. 263–275.

Mohanty, R., Sethares, W.A., Nair, V.A., Prabhakaran, V., 2020. Rethinking measures of functional connectivity via feature extraction. Scientific reports 10, 1298.

Mulcahy, C., 1997. Image compression using the Haar wavelet transform. Spelman Science and Mathematics Journal 1, 22–31.

Murtagh, J., Vadhan, S., 2015. The complexity of computing the optimal composition of differential privacy, in: Theory of Cryptography Conference, Springer. pp. 157–175.

Nanayakkara, P., Bater, J., He, X., Hullman, J., Rogers, J., 2022. Visualizing privacy-utility trade-offs in differentially private data releases. arXiv preprint arXiv:2201.05964 .

Narayanan, A., Shmatikov, V., 2008. Robust de-anonymization of large sparse datasets, in: 2008 IEEE Symposium on Security and Privacy (sp 2008), IEEE. pp. 111–125.

Near, J.P., Abuah, C., 2021. Programming differential privacy. URL: https://uvm .

Panavas, L., Crnovrsanin, T., Adams, J.L., Ullman, J., Sargavad, A., Tory, M., Dunne, C., 2023. Investigating the visual utility of differentially private scatterplots. IEEE Transactions on Visualization and Computer Graphics .

Potter, K., Kniss, J., Riesenfeld, R., Johnson, C.R., 2010. Visualizing summary statistics and uncertainty, in: Computer Graphics Forum, Wiley Online Library. pp. 823–832.

Ramsay, K., Diaz-Rodriguez, J., 2024. Differentially Private Boxplots. arXiv preprint arXiv:2405.20415 .

Rigaki, M., Garcia, S., 2023. A survey of privacy attacks in machine learning. ACM Computing Surveys 56, 1–34.

Rubinov, M., Sporns, O., 2011. Weight-conserving characterization of complex functional brain networks. Neuroimage 56, 2068–2079.

Sendi, M.S., Zendehrouh, E., Ellis, C.A., Liang, Z., Fu, Z., Mathalon, D.H., Ford, J.M., Preda, A., van Erp, T.G., Miller, R.L., et al., 2021. Aberrant dynamic functional connectivity of default mode network in schizophrenia and links to symptom severity. Frontiers in Neural Circuits 15, 649417.

Smith, A., 2011. Privacy-preserving statistical estimation with optimal convergence rates, in: Proceedings of the forty-third annual ACM symposium on Theory of computing, pp. 813–822.

Sommer, D., Meiser, S., Mohammadi, E., 2018. Privacy loss classes: The central limit theorem in differential privacy. Cryptology ePrint Archive .

Steil, J., Hagestedt, I., Huang, M.X., Bulling, A., 2019. Privacy-aware eye tracking using differential privacy, in: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, pp. 1–9.

Sweeney, L., 1997. Weaving technology and policy together to maintain confidentiality. The Journal of Law, Medicine & Ethics 25, 98–110.

Sweeney, L., 2002. k-anonymity: A model for protecting privacy. International journal of uncertainty, fuzziness and knowledge-based systems 10, 557–570.

Tao, Y., Sarwate, A.D., 2025. Differentially Private Distribution Estimation Using Functional Approximation, in: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE. pp. 1–5.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Van Den Heuvel, M.P., Pol, H.E.H., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. European neuropsychopharmacology 20, 519–534.

Wang, X., Chen, W., Chou, J.K., Bryan, C., Guan, H., Chen, W., Pan, R., Ma, K.L., 2018. Graphprotector: A visual interface for employing and assessing multiple privacy preserving graph algorithms. IEEE transactions on visualization and computer graphics 25, 193–203.

Wasserman, L., Zhou, S., 2010. A Statistical Framework for Differential Privacy. Journal of the American Statistical Association 105, 375–389. URL: http://dx.doi.org/10.1198/jasa.2009.tm08651, doi:doi:10.1198/jasa.2009.tm08651.

Zhang, D., Hay, M., Miklau, G., O'Connor, B., 2016. Challenges of visualizing differentially private data. Theory and Practice of Differential Privacy 2016, 1–3.

Zhang, D., Sarvghad, A., Miklau, G., 2020. Investigating visual analysis of differentially private data. IEEE transactions on visualization and computer graphics 27, 1786–1796.

Zhou, J., Wang, X., Wong, J.K., Wang, H., Wang, Z., Yang, X., Yan, X., Feng, H., Qu, H., Ying, H., et al., 2022. Dpviscreator: Incorporating pattern constraints to privacy-preserving visualizations via differential privacy. IEEE Transactions on Visualization and Computer Graphics 29, 809–819.

Appendix

The appendix contains additional details and figures omitted from the main manuscript.

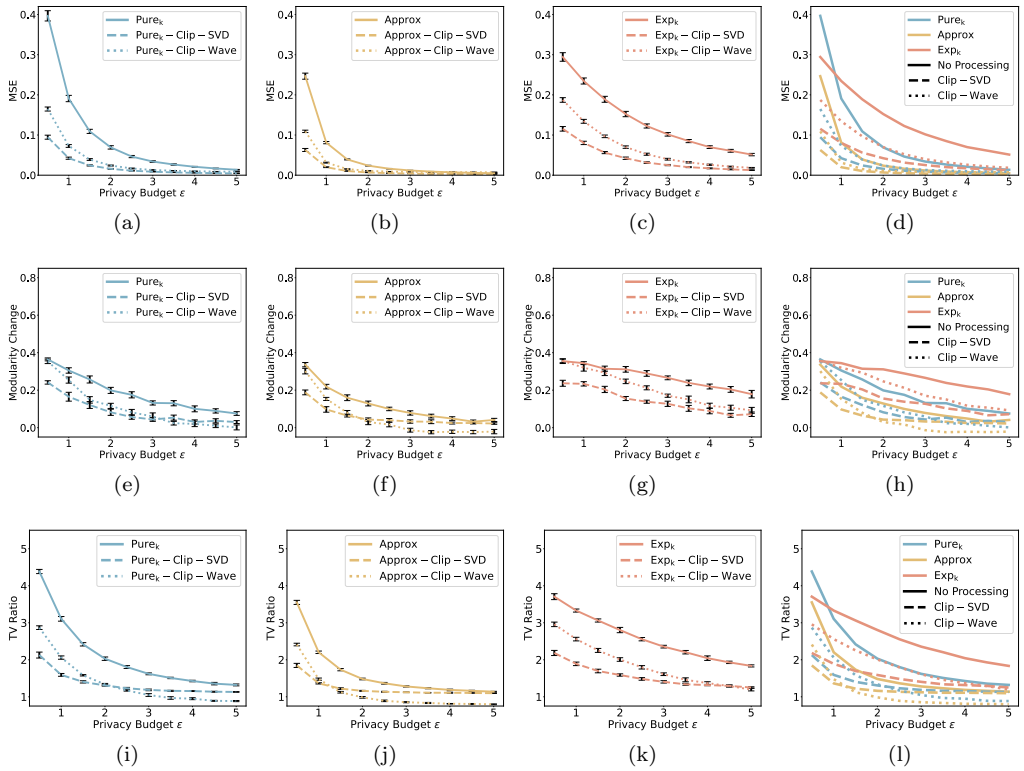


Figure A.1: Quantitative evaluation of privacy-preserving FNC visualization on the BLSA dataset. The first row reports MSE between the original and private average FNC matrices. The second row shows modularity change, defined as $Q_{\text{orginal}} - Q_{\text{DP}}$, where Q denotes signed Louvain modularity. The third row presents the TV ratio, computed as $\text{TV}_{\text{DP}}/\text{TV}_{\text{orginal}}$. Columns correspond to (a)(e)(i) Pure$_k$-, (b)(f)(j) Approx-, (c)(g)(k) Exp$_k$-related methods, and (d)(h)(l) all methods. Each configuration is repeated 50 times, and error bars are provided.

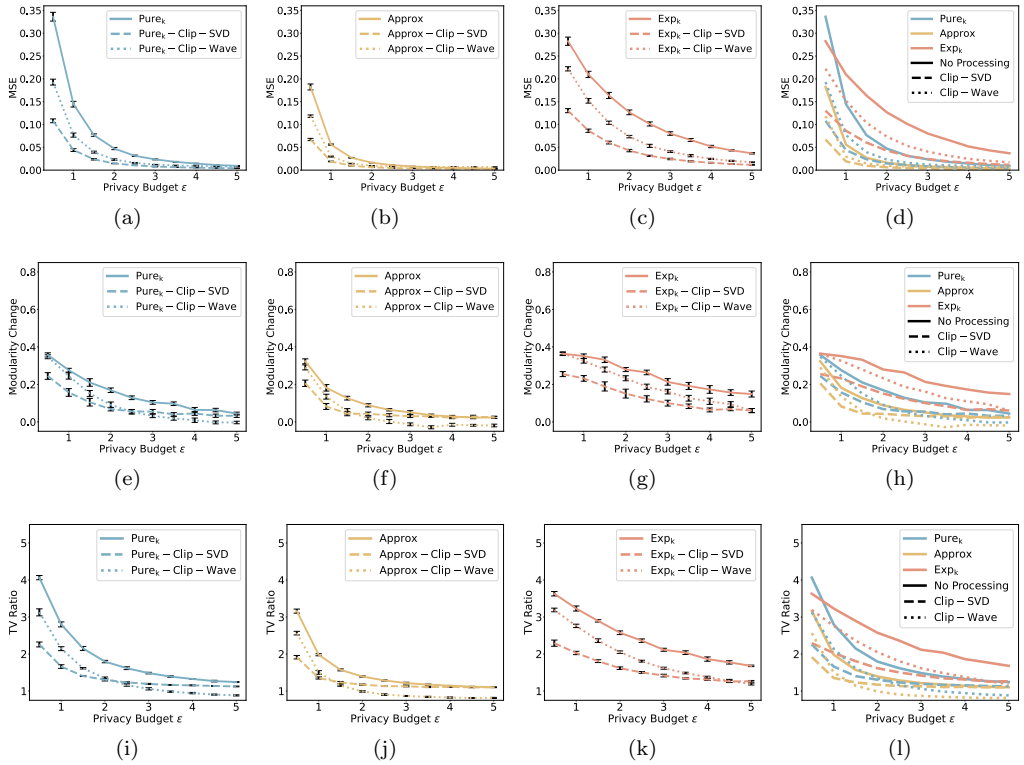
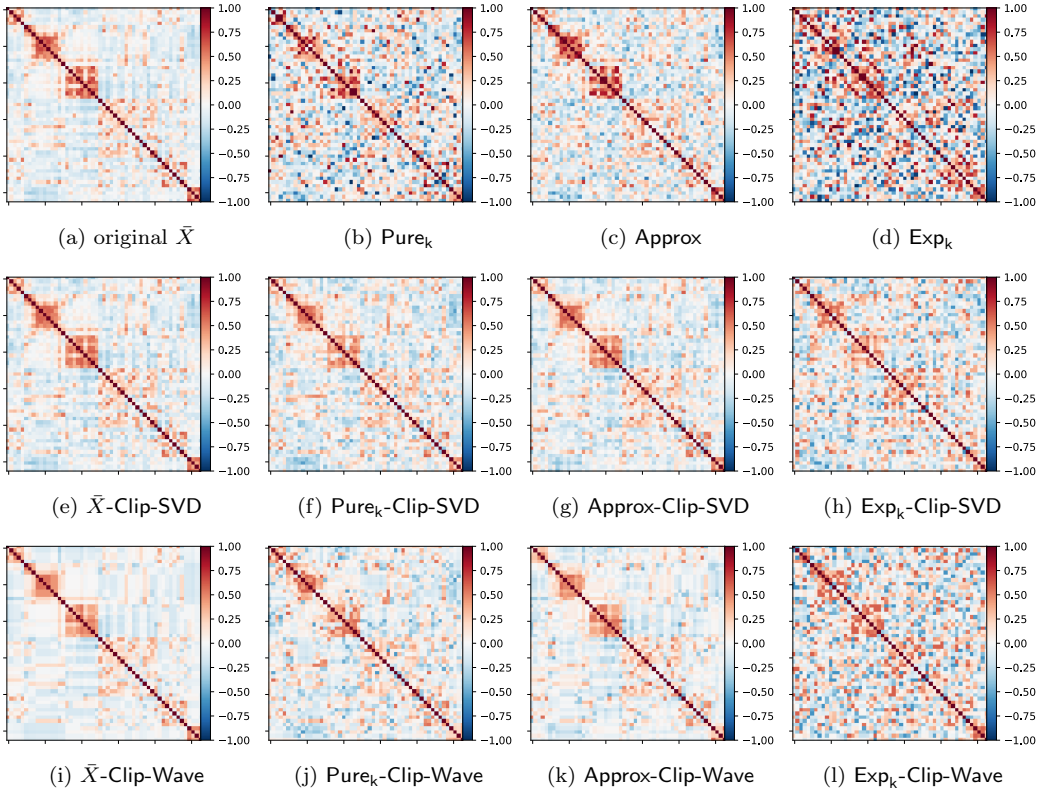


Figure A.2: Quantitative evaluation of privacy-preserving FNC visualization on the ADNI dataset. The first row reports MSE between the original and private average FNC matrices. The second row shows modularity change, defined as $Q_{\text{orginal}} - Q_{\text{DP}}$, where Q denotes signed Louvain modularity. The third row presents the TV ratio, computed as $\text{TV}_{\text{DP}}/\text{TV}_{\text{orginal}}$. Columns correspond to (a)(e)(i) $\mathsf{Pure_k}$-, (b)(f)(j) Approx-, (c)(g)(k) $\mathsf{Exp_k}$-related methods, and (d)(h)(l) all methods. Each configuration is repeated 50 times, and error bars are provided.

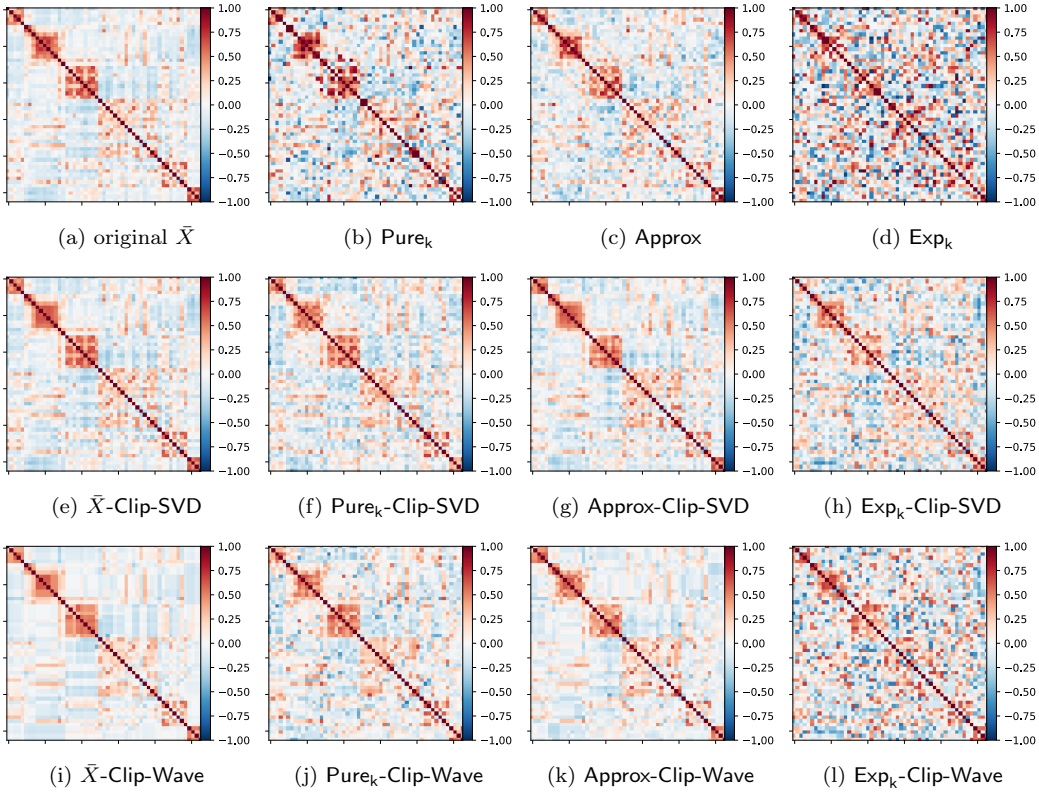


(a) original $\bar{X}$ (b) $\mathsf{Pure}_{\mathsf{k}}$ (c) Approx (d) $\mathsf{Exp}_{\mathsf{k}}$

(e) $\bar{X}$-$\mathsf{Clip\text{-}SVD}$ (f) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$ (g) $\mathsf{Approx\text{-}Clip\text{-}SVD}$ (h) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$

(i) $\bar{X}$-$\mathsf{Clip\text{-}Wave}$ (j) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$ (k) $\mathsf{Approx\text{-}Clip\text{-}Wave}$ (l) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$

Figure A.3: Visualization of the BLSA dataset with $N = 891$, $\epsilon = 2$, $\delta = N^{-3/2}$, clipping bound $b = 0.7$, number of components used in SVD set to 10, threshold used in the wavelet transform set to 0.2, and number of colors set to 256.



(a) original $\bar{X}$ (b) $\mathsf{Pure}_{\mathsf{k}}$ (c) Approx (d) $\mathsf{Exp}_{\mathsf{k}}$

(e) $\bar{X}$-$\mathsf{Clip\text{-}SVD}$ (f) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$ (g) $\mathsf{Approx\text{-}Clip\text{-}SVD}$ (h) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}SVD}$

(i) $\bar{X}$-$\mathsf{Clip\text{-}Wave}$ (j) $\mathsf{Pure}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$ (k) $\mathsf{Approx\text{-}Clip\text{-}Wave}$ (l) $\mathsf{Exp}_{\mathsf{k}}$-$\mathsf{Clip\text{-}Wave}$

Figure A.4: Visualization of the ADNI dataset with $N = 1091$, $\epsilon = 2$, $\delta = N^{-3/2}$, clipping bound $b = 0.75$, number of components used in SVD set to 10, threshold used in the wavelet transform set to 0.2, and number of colors set to 256.