Multivariate Conformal Prediction for Brain Volumetry From Accelerated MRI

Huaqing Jin

Department of Statistics and Data Science
Tsinghua University

huaqingjin@mail.tsinghua.edu.cn

Joint work with **Dr. Fei Jiang** and **Prof. Yanyuan Ma**

2025 年 10 月 18 日

Overview

1. Background

2. Methodology
   - Multivariate Conformal Prediction
   - Semiparametric MCP
   - Full Pipeline

3. Application on ADNI Data

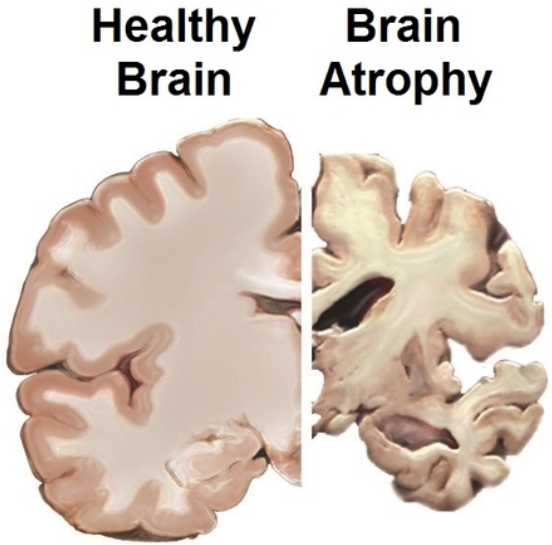
Background

Brain Atrophy

- One of most well-established neuroimaging biomarkers in neurodegenerative diseases (Scheltens et al., 2021).
- Examples: Hippocampal atrophy for Alzheimer's disease (AD) (Rao et al., 2022).

Brain Atrophy

- One of most well-established neuroimaging biomarkers in neurodegenerative diseases (Scheltens et al., 2021).
- Examples: Hippocampal atrophy for Alzheimer's disease (AD) (Rao et al., 2022).
- Characterized by **volume loss** in specific brain regions.

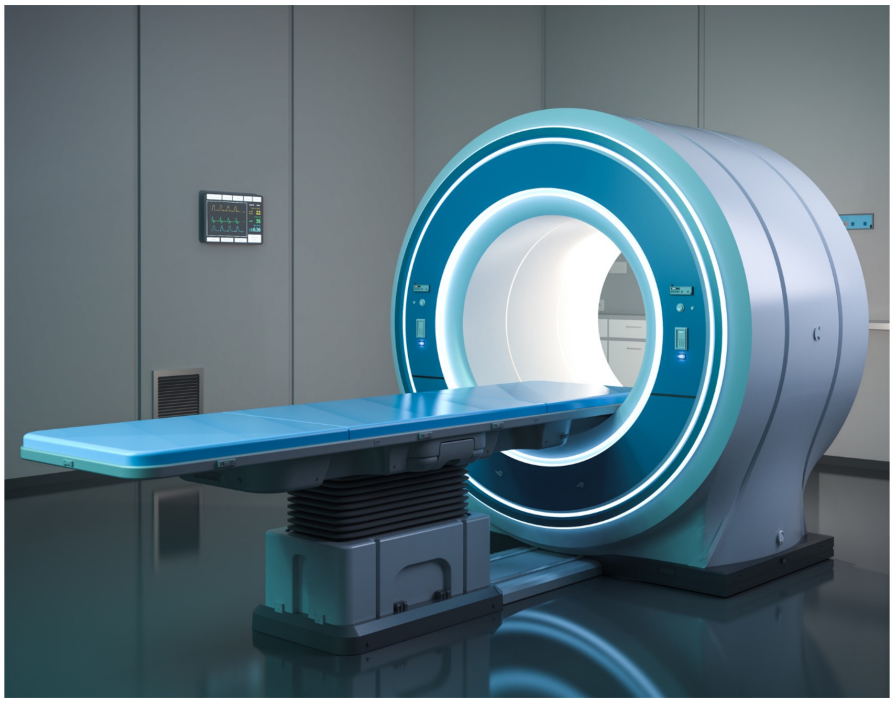


MRI

- **MRI** (Magnetic Resonance Imaging, 磁共振成像): Gold standard for brain atrophy assessment.

MRI

- **MRI** (Magnetic Resonance Imaging, 磁共振成像): Gold standard for brain atrophy assessment.
- High-resolution MRI scans are required ⇒ Long acquisition time.



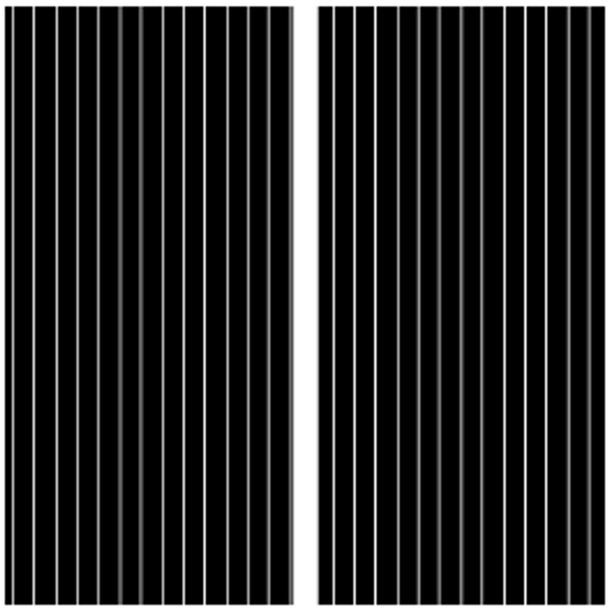
Accelerating MRI Acquisition

- **k-space data**: Raw MRI measurements from the MRI machine (frequency domain, Fourier transform of the image-space data).

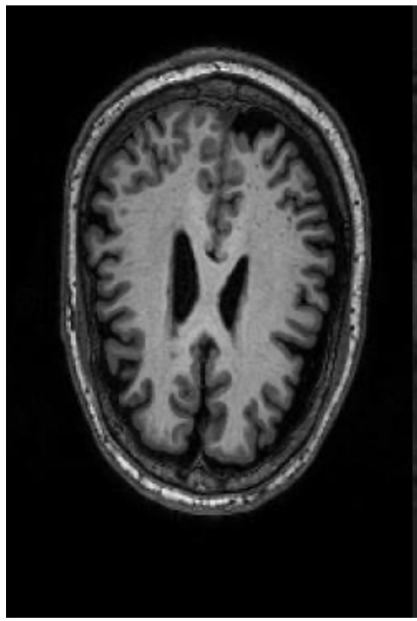
Accelerating MRI Acquisition

- **k-space data**: Raw MRI measurements from the MRI machine (frequency domain, Fourier transform of the image-space data).
- Undersample k-space data to reduce acquisition time.

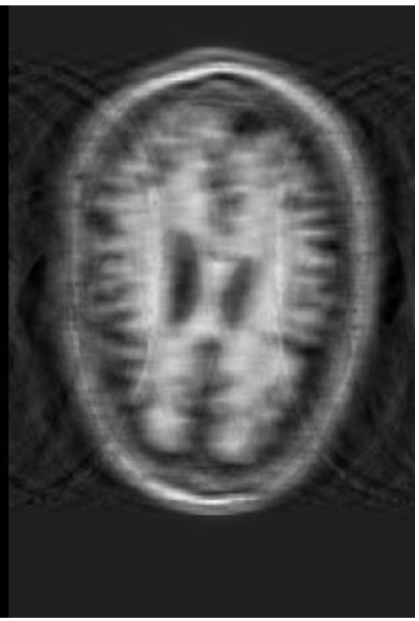
8-fold acceleration



original



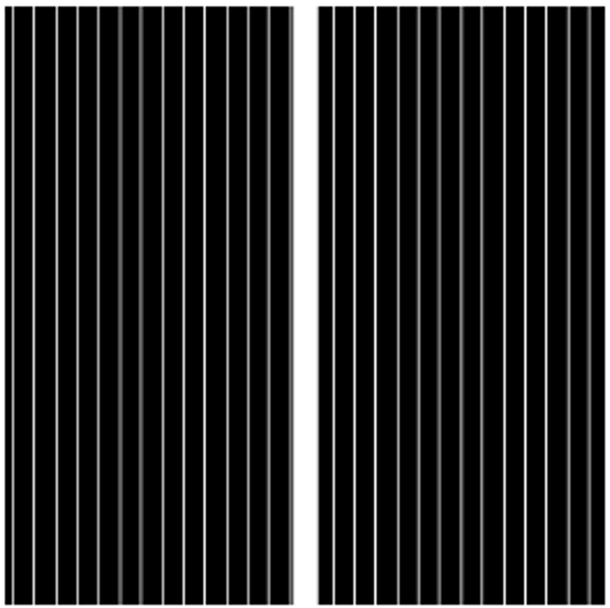
undersampled



Accelerating MRI Acquisition

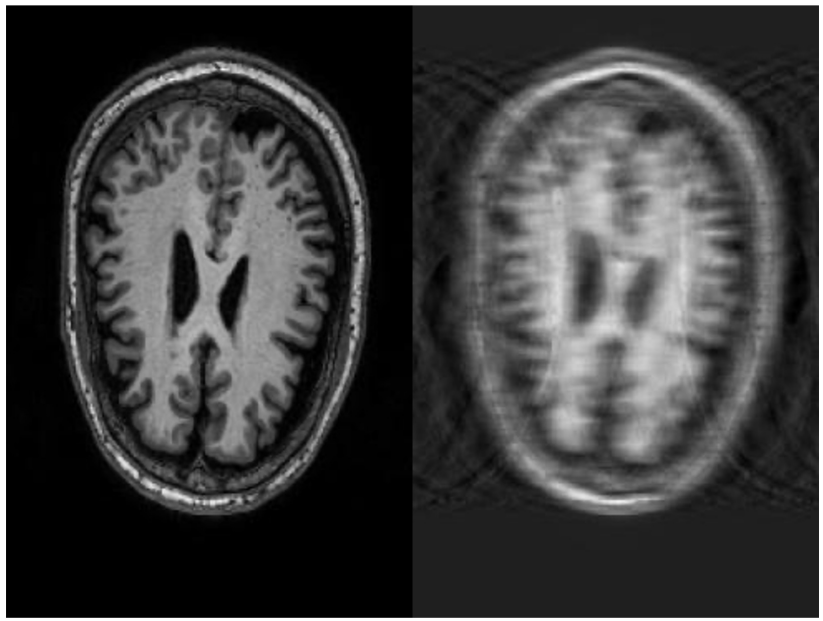
- **k-space data**: Raw MRI measurements from the MRI machine (frequency domain, Fourier transform of the image-space data).
- Undersample k-space data to reduce acquisition time.

8-fold acceleration



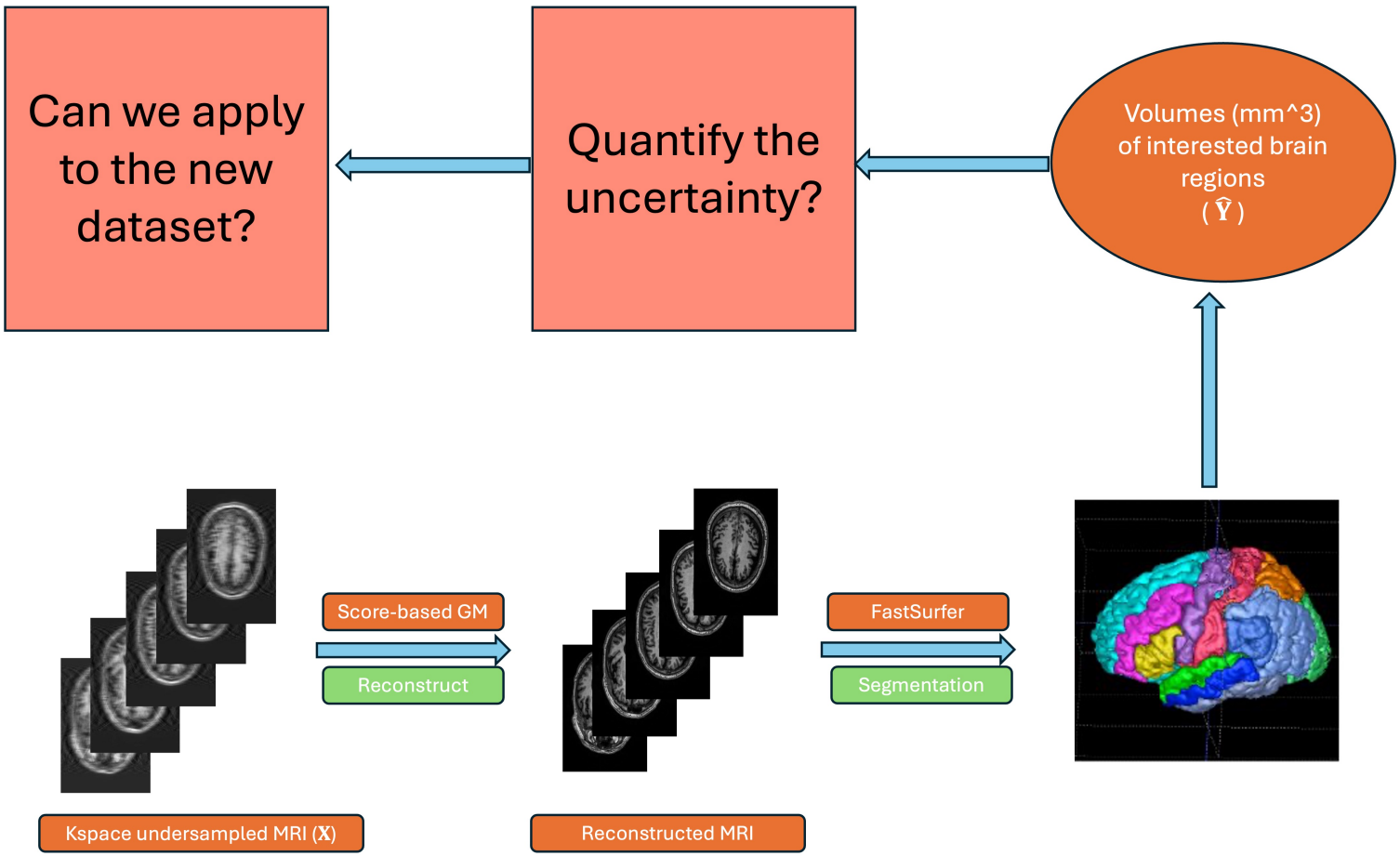
original

undersampled



- Reconstruction: Score-based generative model (Song et al., 2022).

Problem


Can we apply to the new dataset?
Quantify the uncertainty?
Volumes (mm^3) of interested brain regions ($\hat{\mathbf{Y}}$)
Score-based GM
Reconstruct
FastSurfer
Segmentation
Kspace undersampled MRI ($\mathbf{X}$)
Reconstructed MRI

Conformal Prediction

Covariates $\mathbf{X} \in \mathcal{R}^{d_1}$ and Response $Y \in \mathcal{R}$. Let $\mathcal{D} = \{(\mathbf{X}_i, Y_i)\}_{i=1}^n$ be the observed data and $\mathbf{X}_{n+1}$ be the new covariate, the goal is to construct the prediction set $\mathcal{C}(\mathbf{X}_{n+1})$ such that

$$\Pr(Y_{n+1} \in \mathcal{C}(\mathbf{X}_{n+1})) \geq 1 - \alpha,$$

where $(1 - \alpha)$ is the target coverage probability.

- Little assumptions: $\mathcal{D} \cup (\mathbf{X}_{n+1}, Y_{n+1})$ is exchangeable.
- Statistically-guaranteed prediction set.

Split Conformal Prediction (SCP)

- $\widehat{f}(\mathbf{X})$: a predictive model based on $\mathcal{D}^{\mathrm{tr}}$.
- $S(\mathbf{X}, Y)$: a nonconformity score function quantifies the deviation of an observation $(\mathbf{X}, Y)$ from $\widehat{f}(\mathbf{X})$. $S(\mathbf{X}, Y) = |Y - \widehat{f}(\mathbf{X})|$.

Split Conformal Prediction (SCP)

- $\widehat{f}(\mathbf{X})$: a predictive model based on $\mathcal{D}^{\mathrm{tr}}$.
- $S(\mathbf{X}, Y)$: a nonconformity score function quantifies the deviation of an observation $(\mathbf{X}, Y)$ from $\widehat{f}(\mathbf{X})$. $S(\mathbf{X}, Y) = |Y - \widehat{f}(\mathbf{X})|$.

Procedure:

1. Split $\mathcal{D}$ into $\mathcal{D}^{\mathrm{tr}}$ and $\mathcal{D}^{\mathrm{cal}}$.
2. Train $\widehat{f}(\mathbf{X})$ on $\mathcal{D}^{\mathrm{tr}}$.
3. Calculate $S(\mathbf{X}, Y)$ on $\mathcal{D}^{\mathrm{cal}}$.
4. Get $\widehat{\epsilon}_{1-\alpha}$: $(1-\alpha)$th quantile of $\{S(\mathbf{X}_i, Y_i) : (\mathbf{X}_i, Y_i) \in \mathcal{D}^{\mathrm{cal}}\} \cup \{\infty\}$.
5. Construct the prediction set $\mathcal{C}(\mathbf{X}_{n+1}) = \{y : S(\mathbf{X}_{n+1}, y) \leq \widehat{\epsilon}_{1-\alpha}\}$.

Methodology

Multivariate Conformal Prediction

Interested in the volumes from multiple brain regions.

Multivariate Conformal Prediction

Interested in the volumes from multiple brain regions.

- $\mathbf{X} \in \mathcal{R}^{d_1}$: undersampled MRI data.
- $\mathbf{Y} \in \mathcal{R}^{d_2}$: volumes from multiple brain regions.
- $\mathbf{f}(\mathbf{X}) = \mathrm{E}(\mathbf{Y} \mid \mathbf{X})$: conditional expectation of $\mathbf{Y}$ given $\mathbf{X}$.
- $\widehat{\mathbf{f}}(\mathbf{X})$: learned pipeline to estimate volumes from undersampled MRI data.

MCP

Let $Y_j, \widehat{f}_j$ denote the jth element of $\mathbf{Y}, \widehat{\mathbf{f}}$ respectively and $\mathbf{M}(\mathbf{X}) \equiv \text{diag}[\text{var}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{X}) \mid \mathbf{X}\}]$.
Define the nonconformity score function as

$$\begin{aligned} S(\mathbf{X}, \mathbf{Y}) &\equiv \|\mathbf{M}(\mathbf{X})^{-1/2}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{X})\}\|_2 \\ &= \left(\sum_{j=1}^{d_2} [\text{var}\{Y_j - \widehat{f}_j(\mathbf{X}) \mid \mathbf{X}\}]^{-1} \{Y_j - \widehat{f}_j(\mathbf{X})\}^2 \right)^{1/2}, \end{aligned}$$

MCP

Let $Y_j, \widehat{f}_j$ denote the jth element of $\mathbf{Y}, \widehat{\mathbf{f}}$ respectively and $\mathbf{M}(\mathbf{X}) \equiv \text{diag}[\text{var}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{X}) \mid \mathbf{X}\}]$.
Define the nonconformity score function as

$$\begin{aligned} S(\mathbf{X}, \mathbf{Y}) &\equiv \|\mathbf{M}(\mathbf{X})^{-1/2}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{X})\}\|_2 \\ &= \left(\sum_{j=1}^{d_2} [\text{var}\{Y_j - \widehat{f}_j(\mathbf{X}) \mid \mathbf{X}\}]^{-1} \{Y_j - \widehat{f}_j(\mathbf{X})\}^2 \right)^{1/2}, \end{aligned}$$

Let $\widehat{\epsilon}_{\text{MCP}}$ be the $(1-\alpha)$th quantile of $\{S(\mathbf{X}_i, \mathbf{Y}_i) : i = 1, \ldots, n\} \cup \{\infty\}$.
The $(1-\alpha)$ prediction set for a future response $\mathbf{Y}_{n+1}$ associated with a new under-sampled MRI data $\mathbf{X}_{n+1}$ is

$$\mathcal{C}_{\text{MCP}}(\mathbf{X}_{n+1}) = \{\mathbf{y} : S(\mathbf{X}_{n+1}, \mathbf{y}) \leq \widehat{\epsilon}_{\text{MCP}}\}$$

Semiparametric MCP

- MCP discards the known structural information of the score $S_i = S(\mathbf{X}_i, \mathbf{Y}_i)$.

Semiparametric MCP

- MCP discards the known structural information of the score $S_i = S(\mathbf{X}_i, \mathbf{Y}_i)$.
- MCP: Find root of $\sum_{i=1}^{n} [I\{S(\mathbf{X}_i, \mathbf{Y}_i) \leq \epsilon\} - 1 + \alpha] = 0$ w.r.t. ϵ.

Semiparametric MCP

- MCP discards the known structural information of the score $S_i = S(\mathbf{X}_i, \mathbf{Y}_i)$.
- MCP: Find root of $\sum_{i=1}^{n} [I\{S(\mathbf{X}_i, \mathbf{Y}_i) \leq \epsilon\} - 1 + \alpha] = 0$ w.r.t. ϵ.
- **SMCP**: Use **influence function** and solve $\sum_{i=1}^{n} \widehat{\phi}(\mathbf{X}_i, \mathbf{Y}_i, \epsilon) = 0$ to get $\widehat{\epsilon}_{\mathrm{SMCP}}$, where

$$\widehat{\phi}(\mathbf{x}, \mathbf{y}, \epsilon) \equiv \underbrace{\epsilon^{-1}\widehat{\mathbf{M}}(\mathbf{x})^{-1}\mathrm{E}[I\{S(\mathbf{x}, \mathbf{Y}) = \epsilon\}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{x})\} \mid \mathbf{x}] \circ \{\mathbf{y} - \widehat{\mathbf{f}}(\mathbf{x})\}}_{\psi_1(\mathbf{x},\mathbf{y},\epsilon)} + \underbrace{[I\{S(\mathbf{x}, \mathbf{y}) \leq \epsilon\} - 1 + \alpha]}_{\psi_2(\mathbf{x},\mathbf{y},\epsilon)}$$

Semiparametric MCP

- MCP discards the known structural information of the score $S_i = S(\mathbf{X}_i, \mathbf{Y}_i)$.
- MCP: Find root of $\sum_{i=1}^{n} [I\{S(\mathbf{X}_i, \mathbf{Y}_i) \leq \epsilon\} - 1 + \alpha] = 0$ w.r.t. ϵ.
- **SMCP**: Use **influence function** and solve $\sum_{i=1}^{n} \widehat{\phi}(\mathbf{X}_i, \mathbf{Y}_i, \epsilon) = 0$ to get $\widehat{\epsilon}_{\mathrm{SMCP}}$, where

$$\widehat{\phi}(\mathbf{x}, \mathbf{y}, \epsilon) \equiv \underbrace{\epsilon^{-1}\widehat{\mathbf{M}}(\mathbf{x})^{-1}\mathrm{E}[I\{S(\mathbf{x}, \mathbf{Y}) = \epsilon\}\{\mathbf{Y} - \widehat{\mathbf{f}}(\mathbf{x})\} \mid \mathbf{x}] \circ \{\mathbf{y} - \widehat{\mathbf{f}}(\mathbf{x})\}}_{\psi_1(\mathbf{x},\mathbf{y},\epsilon)} + \underbrace{[I\{S(\mathbf{x}, \mathbf{y}) \leq \epsilon\} - 1 + \alpha]}_{\psi_2(\mathbf{x},\mathbf{y},\epsilon)}$$

- The $(1 - \alpha)$ prediction set for a future response $\mathbf{Y}_{n+1}$ associated with a new under-sampled MRI data $\mathbf{X}_{n+1}$ is

$$\mathcal{C}_{\mathrm{SMCP}}(\mathbf{X}_{n+1}) = \{\mathbf{y} : S(\mathbf{X}_{n+1}, \mathbf{y}) \leq \widehat{\epsilon}_{\mathrm{SMCP}}\}.$$

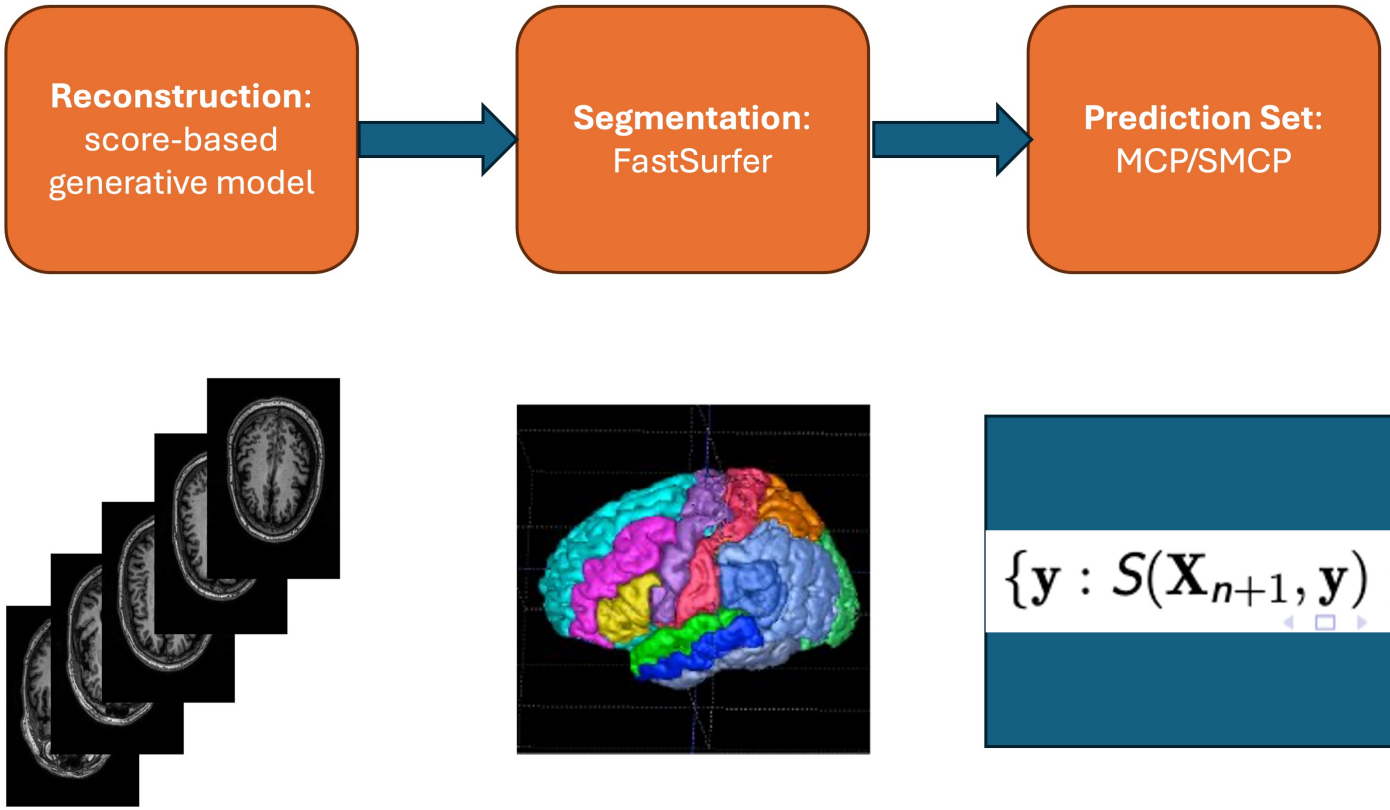
Remarks

- Let ϵ be the $(1-\alpha)$th quantile of $S(\mathbf{X}, \mathbf{Y}) \Rightarrow \widehat{\epsilon}_{\mathrm{SMCP}} = \epsilon + o_p(1)$ and $\widehat{\epsilon}_{\mathrm{MCP}} = \epsilon + o_p(1)$.

Remarks

- Let ϵ be the $(1-\alpha)$th quantile of $S(\mathbf{X}, \mathbf{Y}) \Rightarrow \widehat{\epsilon}_{\mathrm{SMCP}} = \epsilon + o_p(1)$ and $\widehat{\epsilon}_{\mathrm{MCP}} = \epsilon + o_p(1)$.
- The coverage of $\mathcal{C}_{\mathrm{SMCP}}(\mathbf{X}_{n+1})$ and $\mathcal{C}_{\mathrm{MCP}}(\mathbf{X}_{n+1})$ are asymptotically valid.

Full Pipeline



Hypothesis Test for Future Data

A batch of observations $(\mathbf{X}_{n+1}, \mathbf{Y}_{n+1}), \ldots, (\mathbf{X}_{n+m}, \mathbf{Y}_{n+m})$.

Can we apply the pipeline to them?

Hypothesis Test for Future Data

A batch of observations $(\mathbf{X}_{n+1}, \mathbf{Y}_{n+1}), \ldots, (\mathbf{X}_{n+m}, \mathbf{Y}_{n+m})$.

Can we apply the pipeline to them?

Construct the test statistic

$$T \equiv V^{-1/2} m^{-1/2} \sum_{i=1}^{m} \left[I\{S(\mathbf{X}_{n+i}, \mathbf{Y}_{n+i}) \leq \widehat{\epsilon}_j\} - (1-\alpha) \right],$$

where $j = \mathrm{MCP}, \mathrm{SMCP}$ and V is the asymptotic variance.

Hypothesis Test for Future Data

A batch of observations $(\mathbf{X}_{n+1}, \mathbf{Y}_{n+1}), \ldots, (\mathbf{X}_{n+m}, \mathbf{Y}_{n+m})$.

Can we apply the pipeline to them?

Construct the test statistic

$$T \equiv V^{-1/2} m^{-1/2} \sum_{i=1}^{m} \left[I\{S(\mathbf{X}_{n+i}, \mathbf{Y}_{n+i}) \leq \widehat{\epsilon}_j\} - (1-\alpha) \right],$$

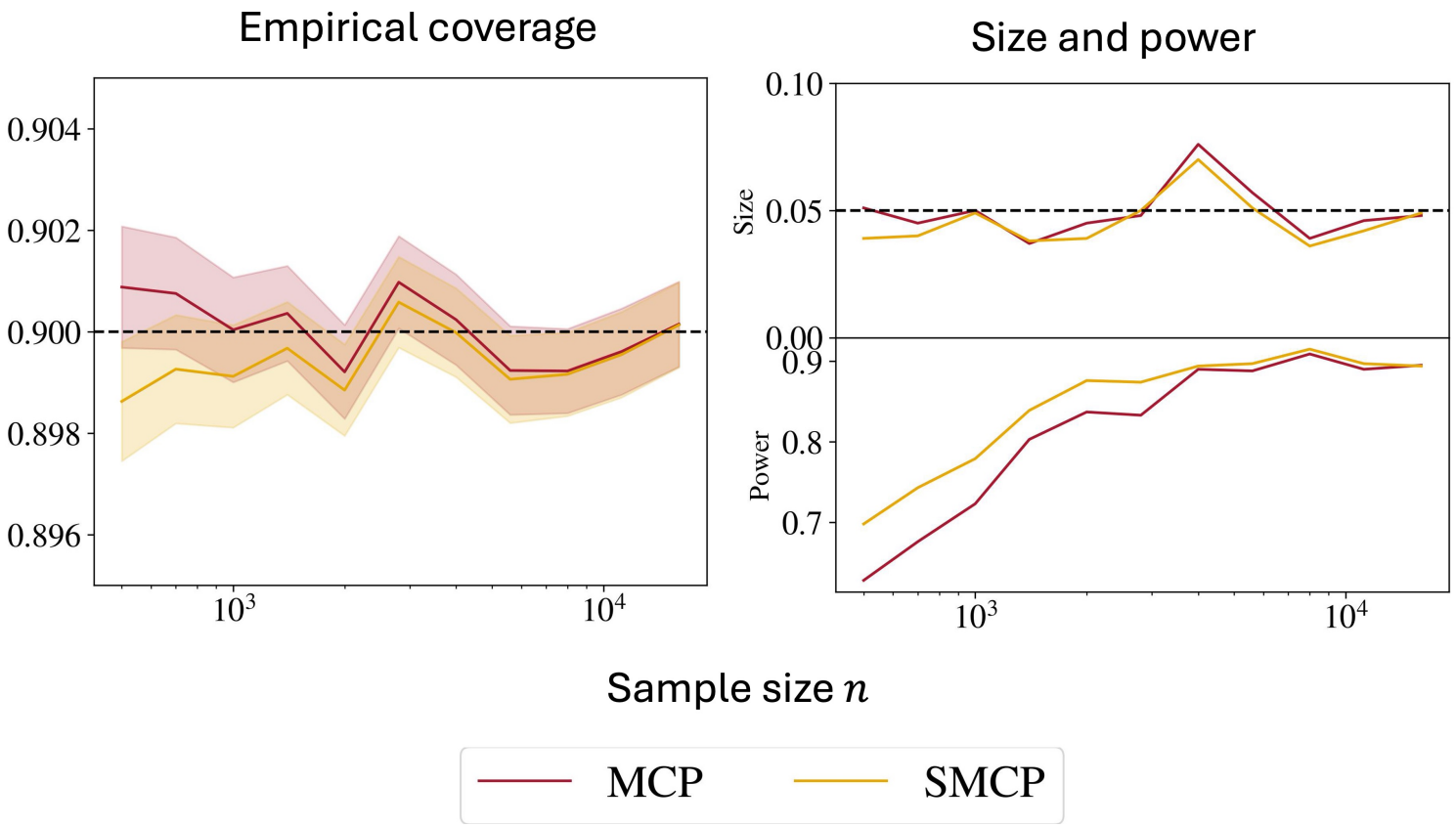
where $j = \mathrm{MCP}, \mathrm{SMCP}$ and V is the asymptotic variance.
H_0: the new data follow the same joint distribution of previous data (i.e., the pipeline is applicable to the new data)

$$\text{Under } H_0,\ T \xrightarrow{d} N(0,1)$$

Given a significance level γ, we reject H_0 if $|T| > z_{1-\gamma/2}$.

Simulation

Target coverage probability is $(1-\alpha) = 0.9$ and $\gamma = 0.05$



Application on ADNI Data

ADNI Data

- 3693 MRI volumes from AD subjects: 1.5T (2567) and 3T (1126).
- Can the pipeline from 1.5T MRI data be applied to 3T MRI data?
- 6 Braak stages with different ROIs. (Combine stages 1 and 2)

Empirical Coverage

- Pipeline from 1.5T MRI data and test the applicability to 3T MRI data.
- Sample 1000 scans from 1.5T MRI data to estimate CP cutoffs and apply to 100 1.5T and 100 3T scans, respectively.
- Target coverage probability $1 - \alpha = 0.90$ and repeat 1000 times.

Empirical Coverage

- Pipeline from 1.5T MRI data and test the applicability to 3T MRI data.
- Sample 1000 scans from 1.5T MRI data to estimate CP cutoffs and apply to 100 1.5T and 100 3T scans, respectively.
- Target coverage probability $1 - \alpha = 0.90$ and repeat 1000 times.

Stage	1 and 2	3	4	5	6
MCP	0.897	0.903	0.897	0.896	0.901
SMCP	0.894	0.901	0.904	0.897	0.899

Hypothesis Test

表 1: Hypothesis test results when significance level $\gamma = 0.05$ under 1000 simulations.

Stage		1 and 2	3	4	5	6
MCP	Size	0.034	0.045	0.034	0.044	0.029
SMCP		0.041	0.047	0.036	0.061	0.036
MCP	Power	0.808	0.088	0.216	0.224	0.045
SMCP		0.832	0.097	0.272	0.330	0.059
Hotelling's T^2	p-value	< 0.001	0.003	< 0.001	< 0.001	0.062

- Hotelling's T^2 test to test the difference between the volumes from 1.5T and 3T MRI data in different stages.

Reference

Rao, Y. L., Ganaraja, B., Murlimanju, B., Joy, T., Krishnamurthy, A., and Agrawal, A. (2022). Hippocampus and its involvement in alzheimer's disease: a review. *3 Biotech* **12,** 55.

Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., Cummings, J., and van der Flier, W. M. (2021). Alzheimer's disease. *The Lancet* **397,** 1577–1590.

Song, Y., Shen, L., Xing, L., and Ermon, S. (2022). Solving inverse problems in medical imaging with score-based generative models. In *International Conference on Learning Representations*.

Thank You