Benefits of Being Bayesian: Motor Imagery Electroencephalogram Classification

Ethan Davis

University of Washington

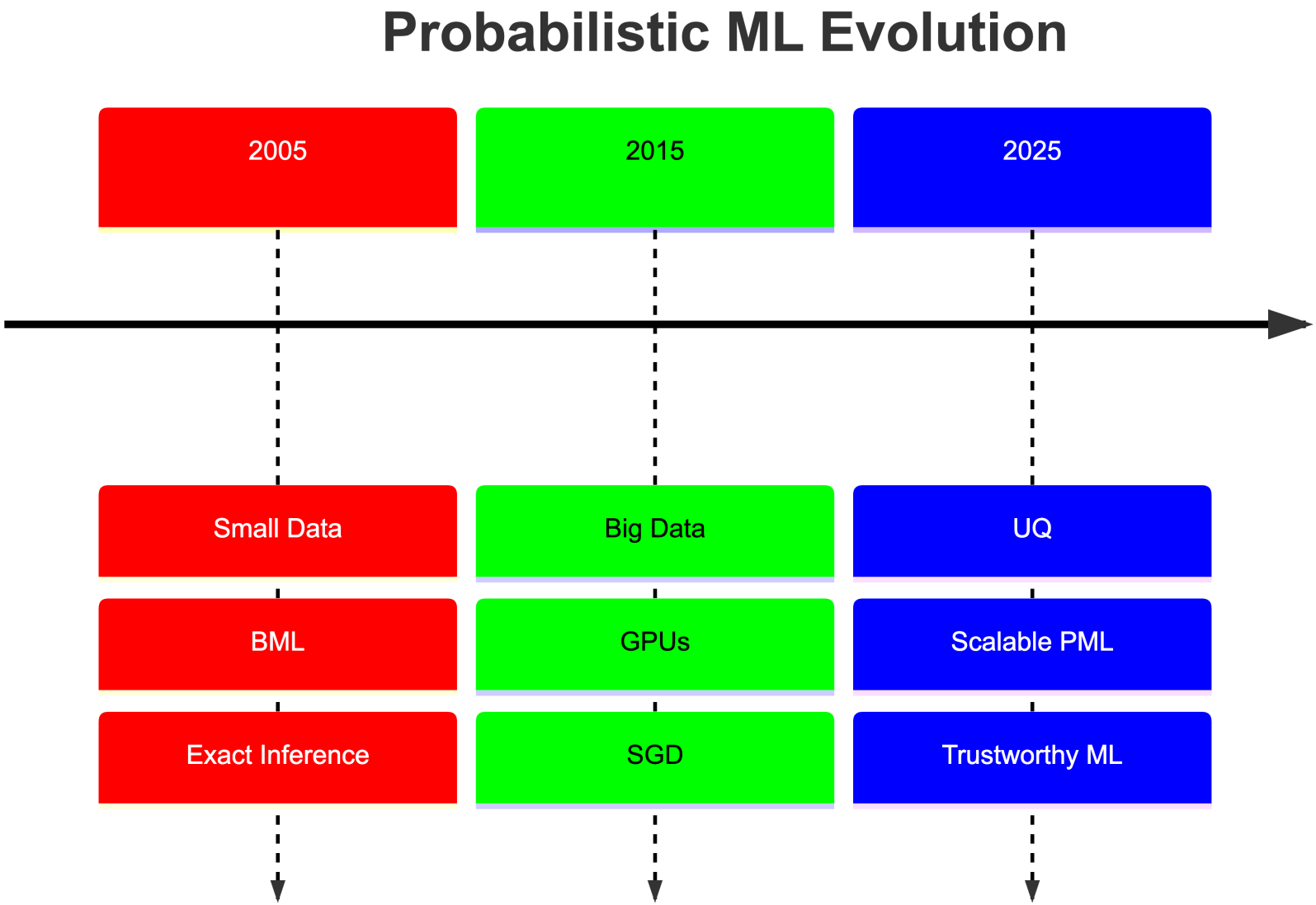
June 2026

Reading Committee:
Erika Parsons, Ph.D., Chair
Michael Stiber, Ph.D.
Elizabeth Field, Ph.D.

Outline

Outline

Interests

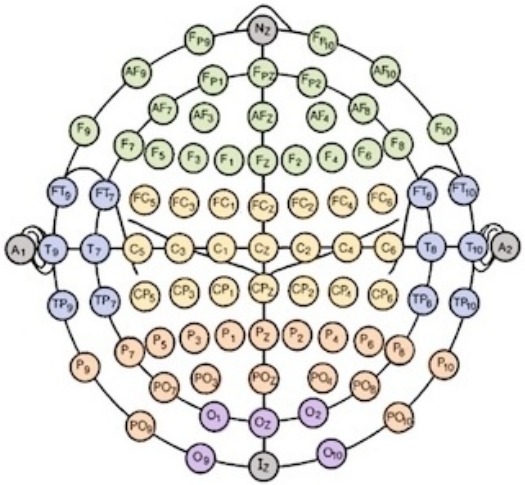

Probabilistic ML Evolution
2005
2015
2025
Small Data
Big Data
UQ
BML
GPUs
Scalable PML
Exact Inference
SGD
Trustworthy ML

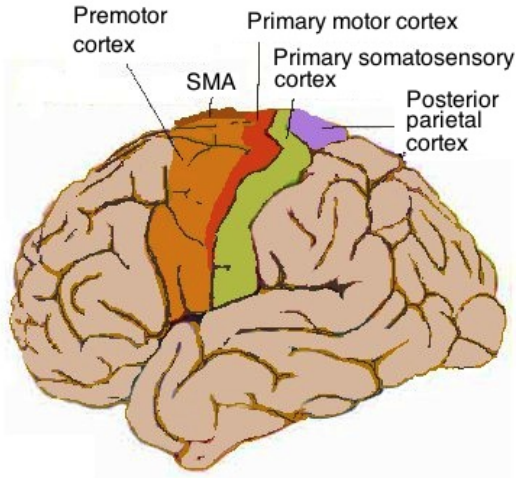
Glossary

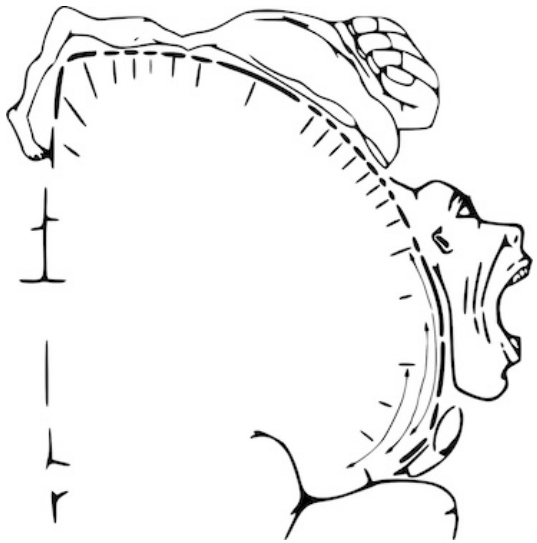
Electroencephalogram (EEG) Measures electrical brain activity via scalp electrodes

Motor Imagery (MI) Mental simulation of movement without physical execution

Brain-Computer Interface (BCI) System translating brain signals into commands



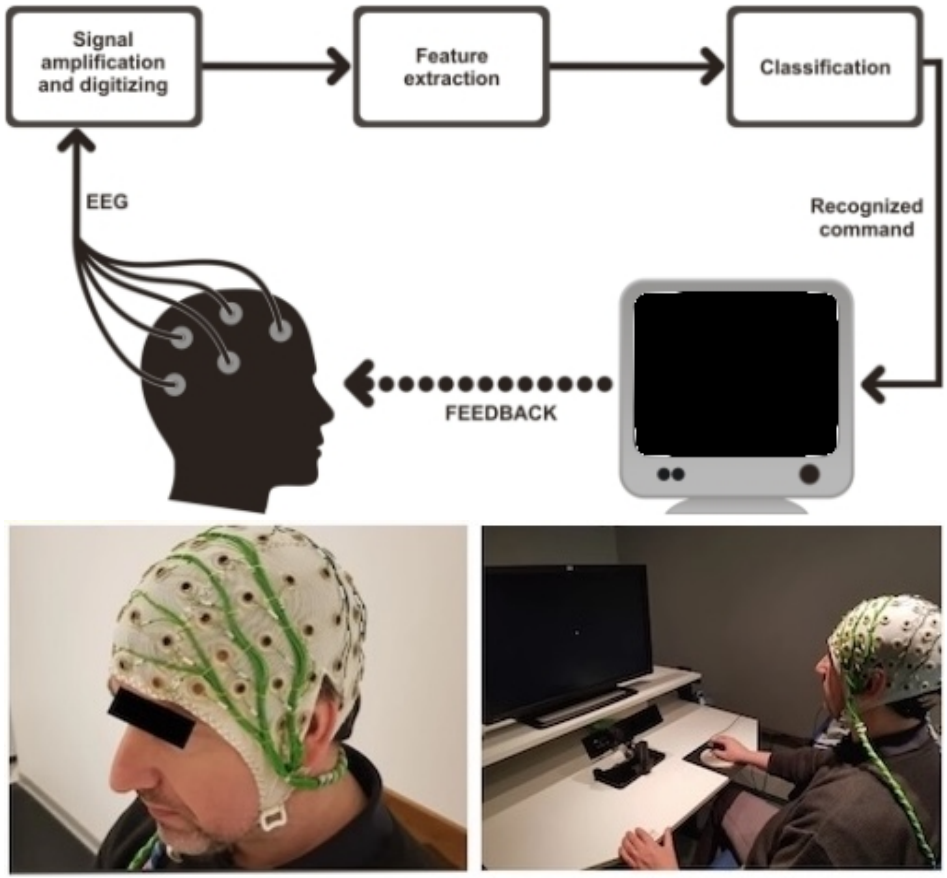




Standard 10-10 montage, sensorimotor cortex, and cortical homunculus [3, 12, 22]

Motivation

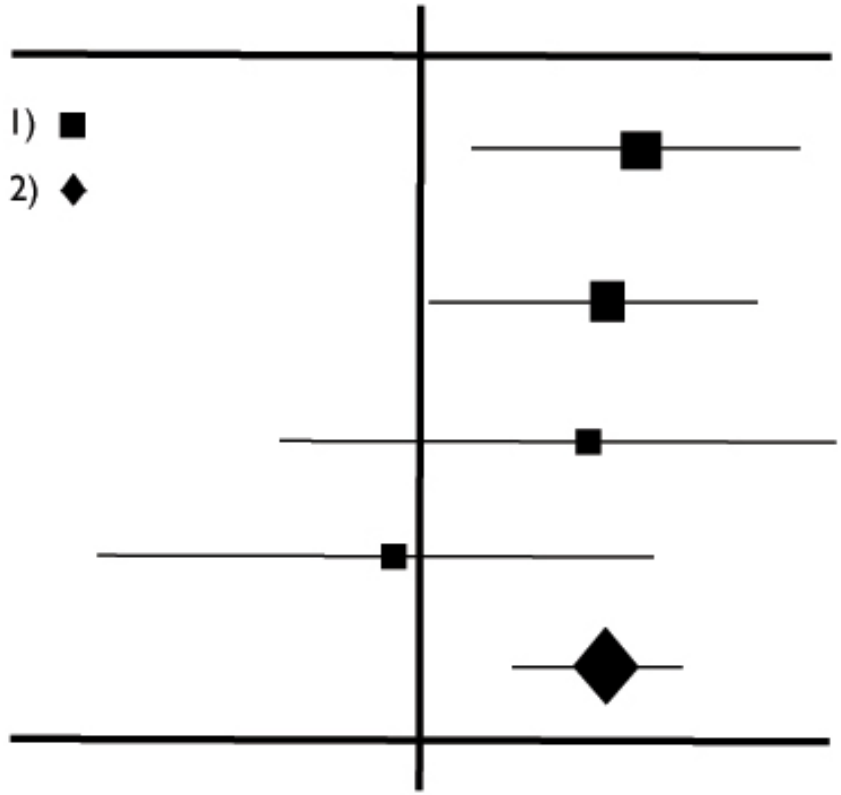
- Clinical value in BCI (e.g., motor rehabilitation, neuroprosthetics)
- Classification is challenging (e.g., low SNR, nonstationarity)
- Bayesian ML explores complex, potentially important properties
  - The literature: Frequentist vs. Bayesian tests are informal, primary studies [26]
- Opportunity to research large-scale meta-analysis of frequentist and Bayesian ML for MI-EEG classification



BCI pipeline [11, 3]

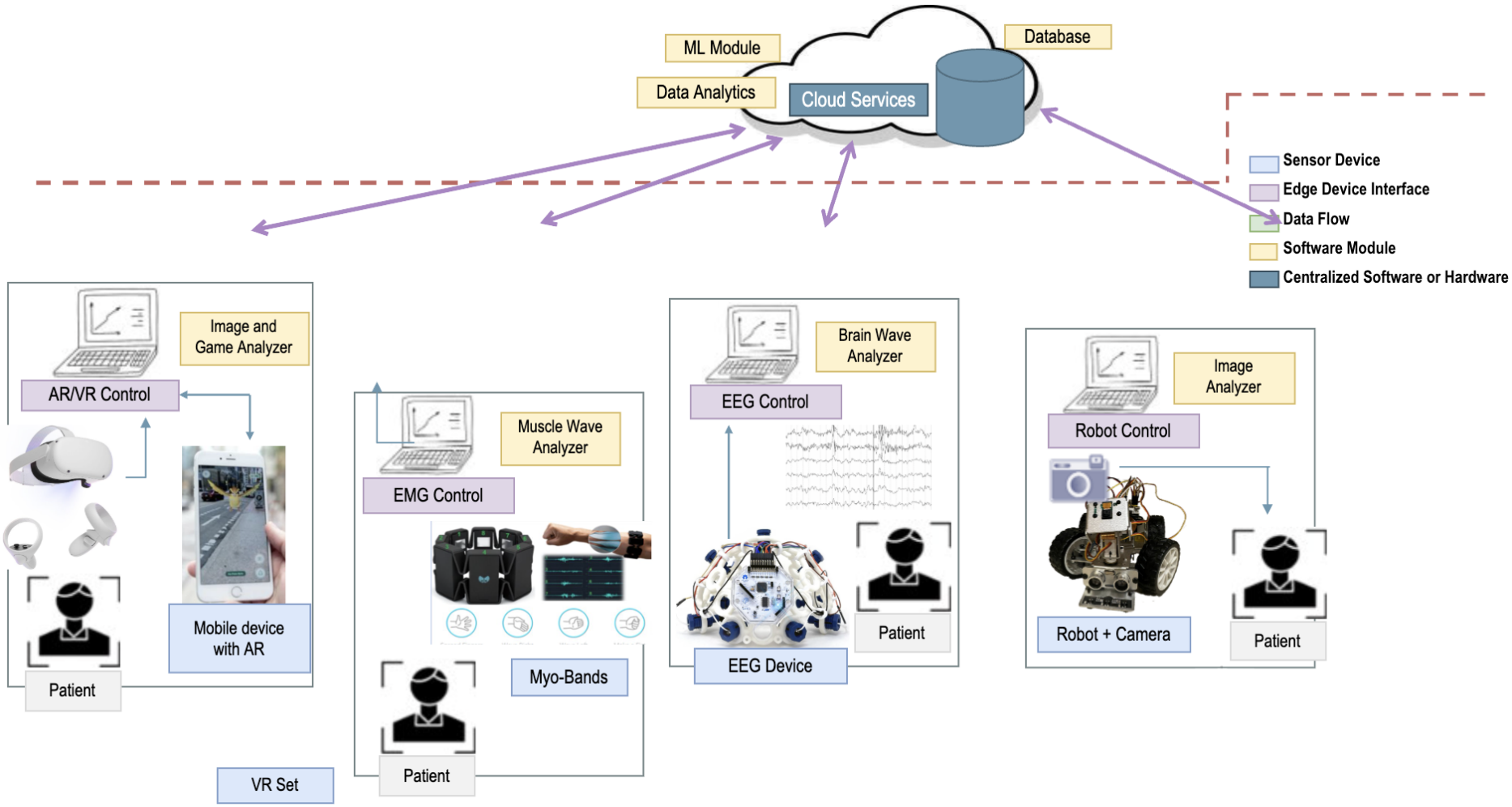
Hypothesis

H_0 The precision-weighted pooled effect of paired differences in Bayesian minus frequentist ML predictions from left/right hand MI-EEG classification is not statistically significant for any discrimination, calibration, or proper scoring metric.



A meta-analysis forest plot [21]

SPANER Lab



Smart Platform for Augmented Neurology Rehabilitation (SPANER)

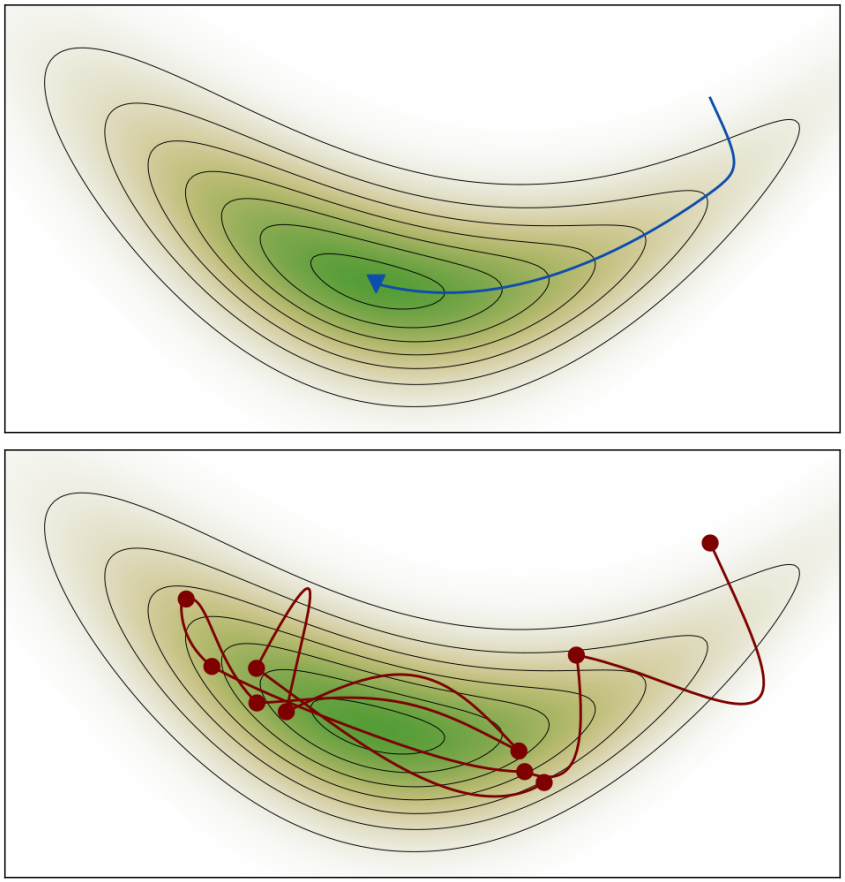
Bayesian Machine Learning

$p(y|\boldsymbol{x}, \hat{\boldsymbol{\theta}})$ such that

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{n=1}^{N} \log p(y_n|\boldsymbol{x}_n, \boldsymbol{\theta})$$

$$p(y|\boldsymbol{x}, \mathcal{D}) = \int p(y|\boldsymbol{x}, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathcal{D}) \mathrm{d}\boldsymbol{\theta}$$

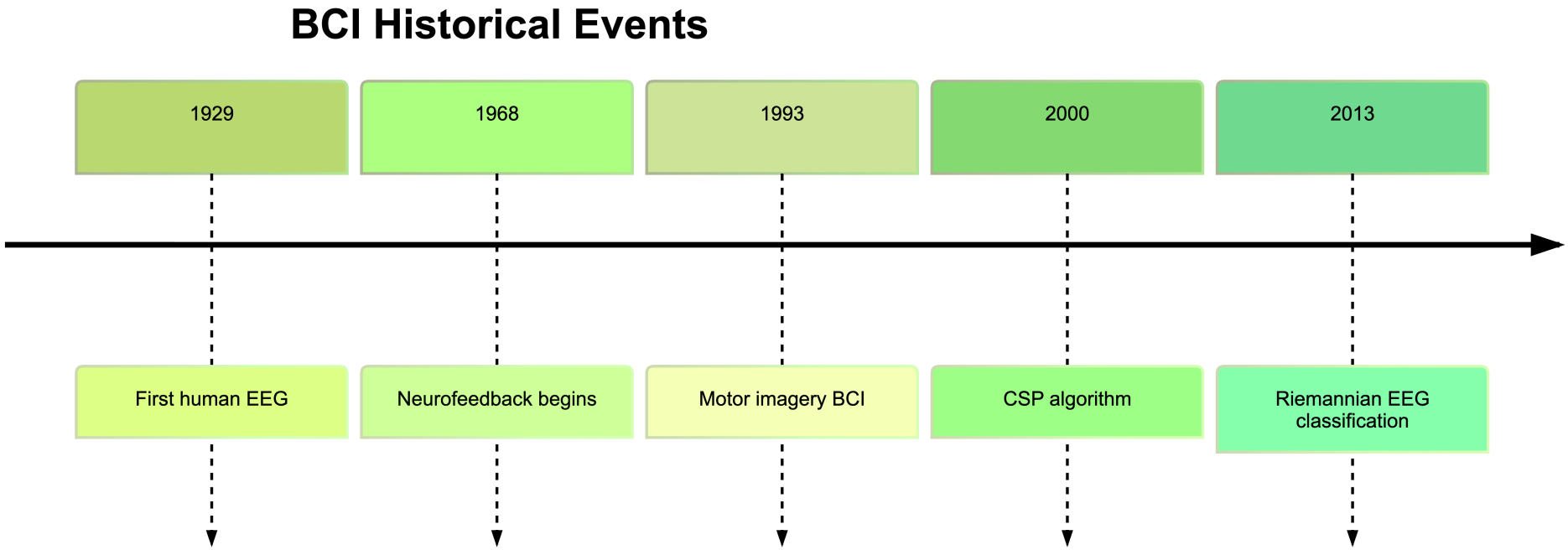
where $p(\boldsymbol{\theta}|\mathcal{D}) \propto p(\boldsymbol{\theta}) p(\mathcal{D}|\boldsymbol{\theta})$



Gradient descent maximum likelihood & Hamiltonian Monte Carlo [15, 16]

Motor Imagery Electroencephalogram

- Origins relied on time and frequency domain engineering (e.g., power & energy features, Hjorth parameters), highly complex classifiers
- Current methodology spans temporal and spatial filters (e.g., CSP), Riemannian geometric statistics on SPD manifold, and deep learning



EEG research timeline [17]

Outline

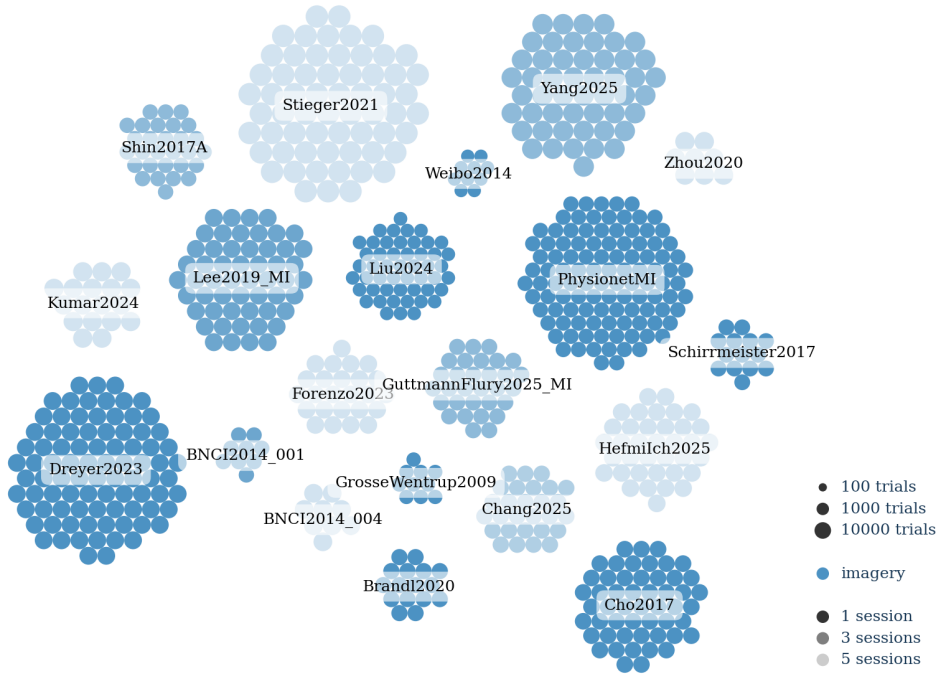
Datasets

- Mother of all BCI Benchmarks (MOABB) centralizes datasets
- Selection criteria yielded 20 diverse datasets (PyPI v1.5)
  1. Left/right hand MI-EEG
  2. At least 8 subjects
  3. No subset/duplicate studies

Session Several runs, device is attached

Run Series of trials without interruption

Trial Atomic event (e.g., left/right hand MI)

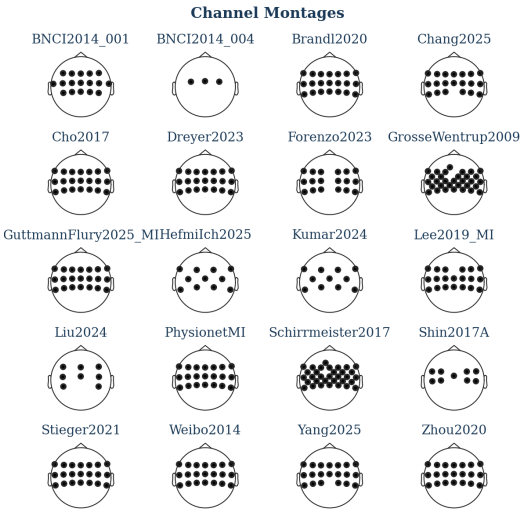


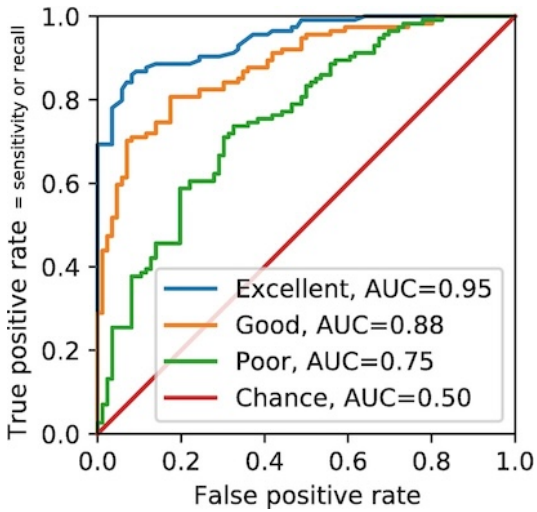
Sample characteristics scale comparison

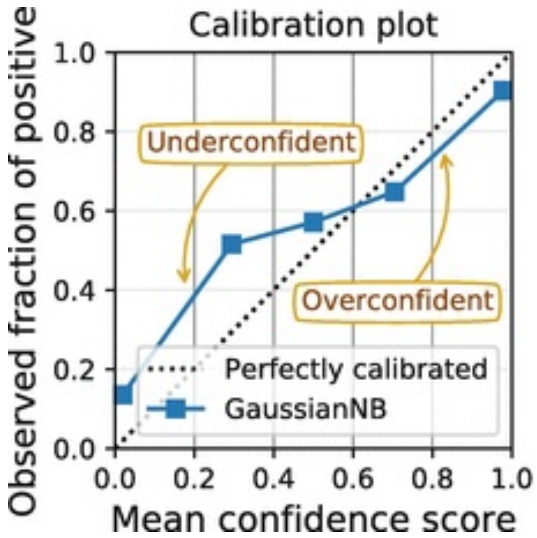
No. Trials				
Min	Max	Median	Mean	Std
1580	43422	5719	8923.9	10534.83

Experiment Assumptions

- Dimensionality reduction based on 10-05 sensorimotor electrodes
- 8–32 Hz high/low pass temporal filter and 128 Hz downsampling
- Discrimination (AUROC, MCC): ↑ranking ability ↓ignores magnitude
- Calibration (ECE, MCE): ↑reliability ↓uninformative predictions
- Proper Scoring (NLL, BS): Rewards true beliefs, cannot be gamed







Channel montages per dataset and performance & confidence metrics [25]

Frequentist Baselines

Raw Traditional temporal & spatial filters, analysis

Riemannian Covariance estimation, Riemannian statistics

Deep Learning Learns temporal & spatial features directly

Selected top two pipelines from each category making six frequentist models

Highest Performers

Pipeline	AUROC
TS+LR	81.479
TS+SVM	81.352
ShallowConvNet	78.029
CSP+SVM	77.72
CSP+LDA	77.629
DeepConvNet	75.402

MOABB's literature [2]. 19 pipelines by 10 left/right hand MI-EEG datasets. Rows ordered by highest to lowest AUROC averaged over datasets.

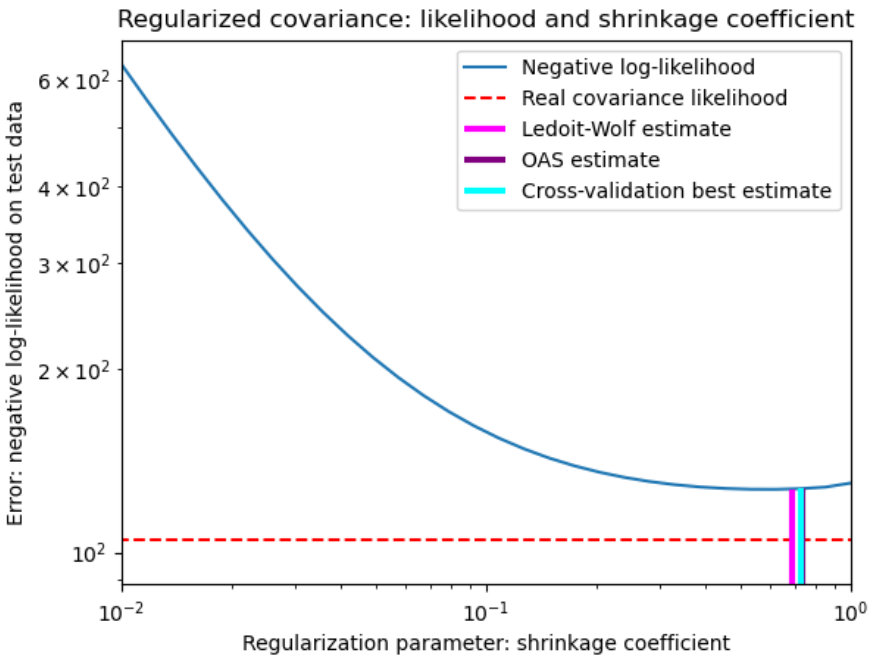
Bayesian Models

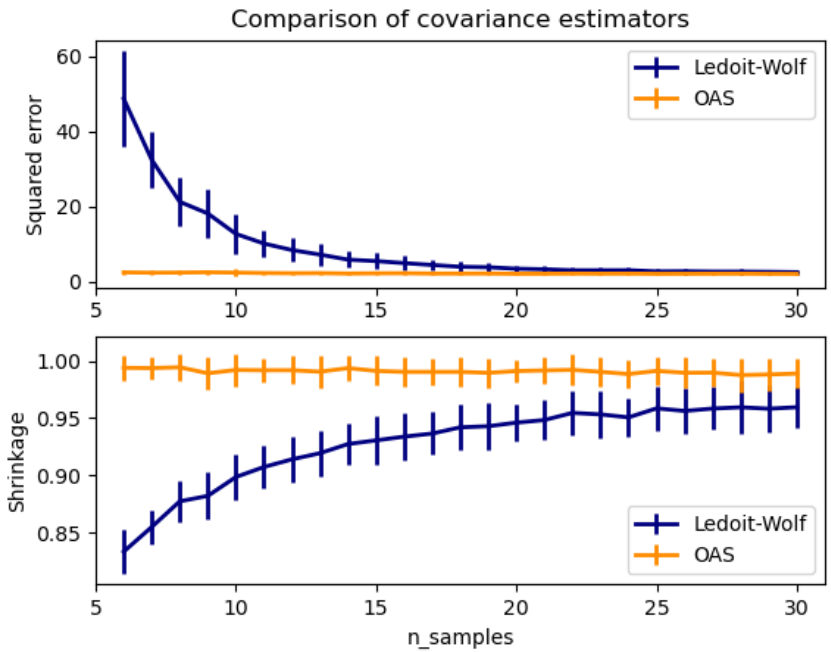
Pairwise Pipelines		
Category	Frequentist	Bayesian
Raw	CSP+LDA	CSP+BLDA
Raw	CSP+SVM	CSP+GP
Riemannian	TS+LR	TS+BLR
Riemannian	TS+SVM	TS+GP
Deep Learning	SCNN	BSCNN
Deep Learning	DCNN	BDCNN

- No-U-Turn Sampler (NUTS)
  - Extends Hamiltonian Monte Carlo (HMC) and MCMC
  - chains=4, warmup=1000, draws=1000, acceptance=0.95
- Input is centered & scaled, generally $\mathcal{N}(0, 1)$ is weakly informative

Covariance Estimation

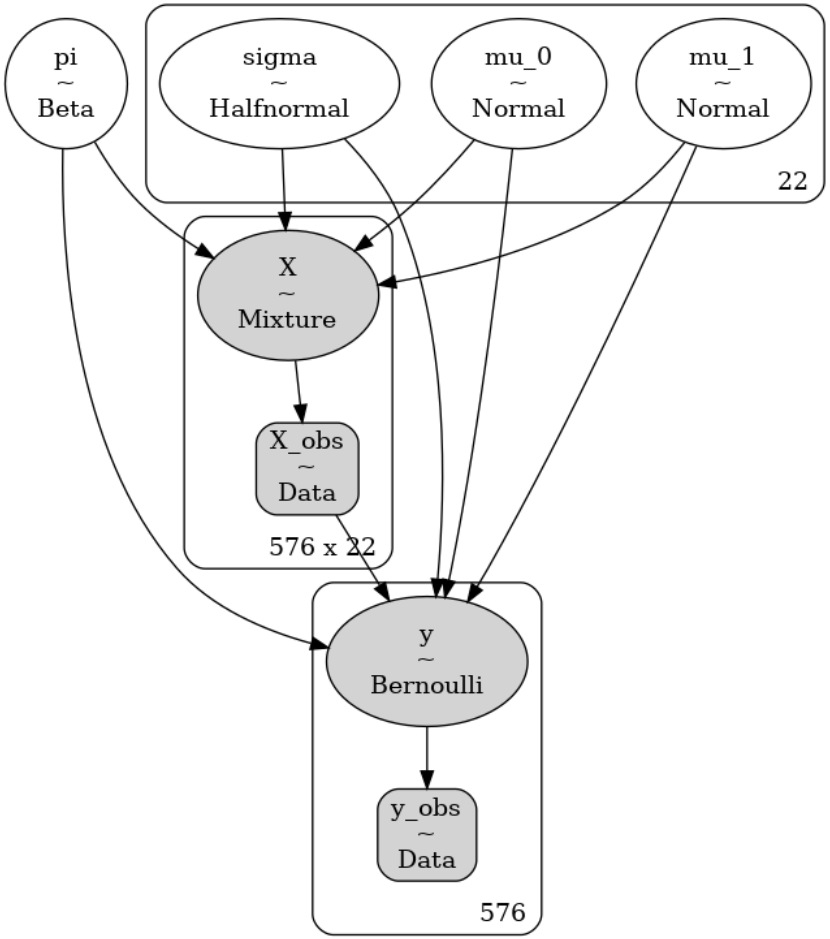
- Shrinkage: Reduces condition number $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$ for numerical analysis
- Assumed closed-form Oracle Approximating Shrinkage (OAS)
  - Used by CSP and TS algorithms
- Minimizes MSE between estimated and real covariance matrix



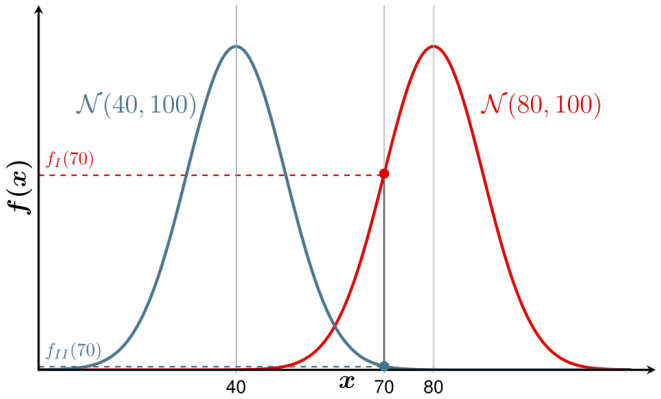


Comparison of Ledoit-Wolf and OAS [20]

Linear Discriminant Analysis



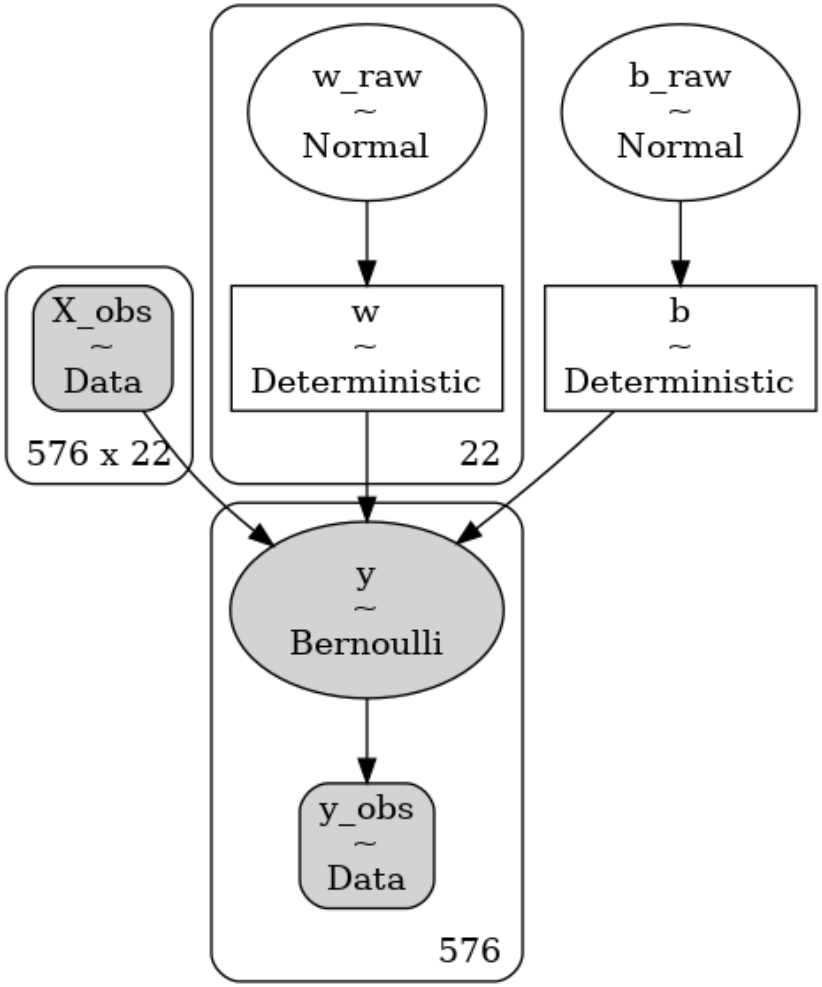
Based on PyMC's GMM [7]



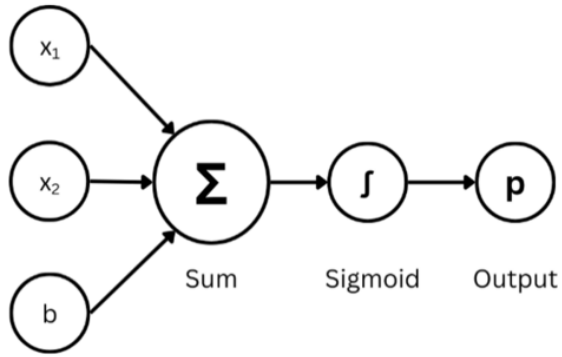
LDA is a supervised GMM [18]

- Generative model
- Frequentist solver used SVD
- $\pi \sim \text{Beta}(2, 2)$
- $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1 \sim \mathcal{N}(0, 1)$
- $\boldsymbol{\sigma} \sim \mathcal{HN}(1)$

Logistic Regression



- LBFGS solver, 1000 iterations
  - Uses 2nd order Hessian approx.
- $\boldsymbol{w}, b \sim \mathcal{N}(0, 1)$
- LeCun initialization: $\sigma_w = \frac{1}{\sqrt{D}}$ where D is the fan-in [19]
- Non-centered parameterization

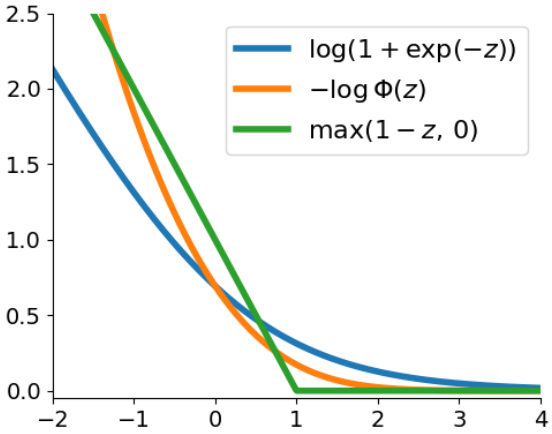


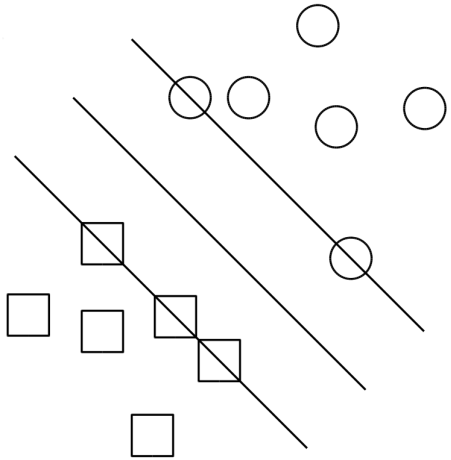
Weighted sum through sigmoid [24]

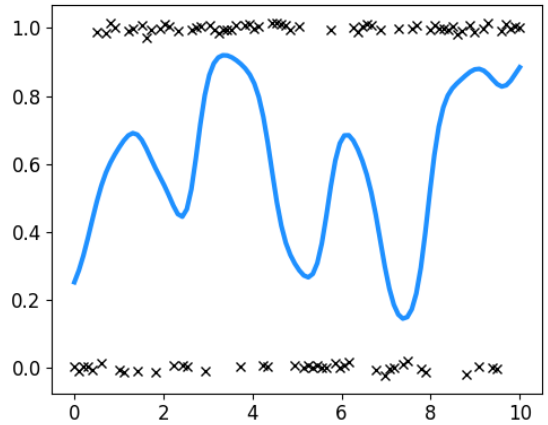
Support Vector Machines & Gaussian Process Classifiers

- Kernel methods: GPC MAP and SVM yield similar solutions [23]
- Post-hoc Platt scaling applied to SVMs for probabilistic calibration
- Non-centered parameterization applied via Cholesky factorization
- Used PyMC's DTC sparse approximation with 100 inducing points

$$q_{\mathrm{DTC}}(\boldsymbol{f}_X, \boldsymbol{f}_*) = \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \boldsymbol{Q}_{X,X} & \boldsymbol{Q}_{X,*} \\ \boldsymbol{Q}_{*,X} & \boldsymbol{K}_{*,*} \end{pmatrix}\right)$$

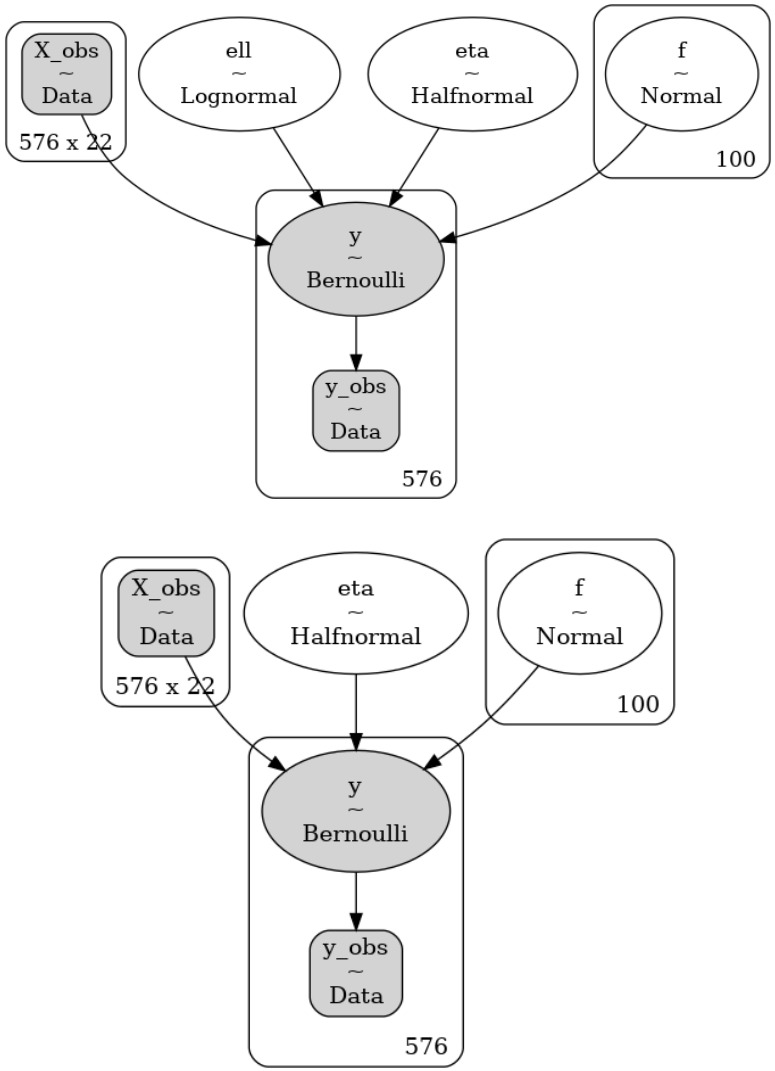






Loss functions logistic, probit, and hinge, an SVM margin, and a GPC

RBF and Linear Kernel: SVMs & GPCs



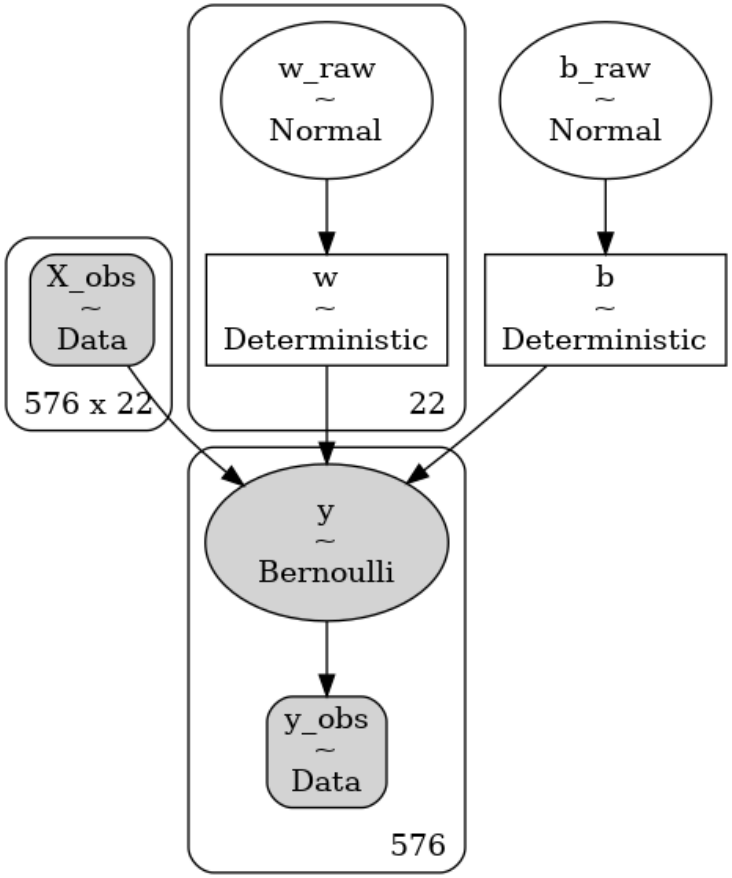
- CSP features → RBF kernel
  - Low-dimensional features
  - RBF is a good default: flexible, smooth, tractable
- $\eta \sim \mathcal{HN}(1),\ \ell \sim \mathcal{LN}(0, 0.5)$

$$k(x, x') = \eta^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right)$$

- TS features → linear kernel
  - Linear approximation of the geodesic decision boundary
- $\eta \sim \mathcal{HN}(1)$

$$k(x, x') = \eta^2 (x \cdot x')$$

Convolutional Neural Networks



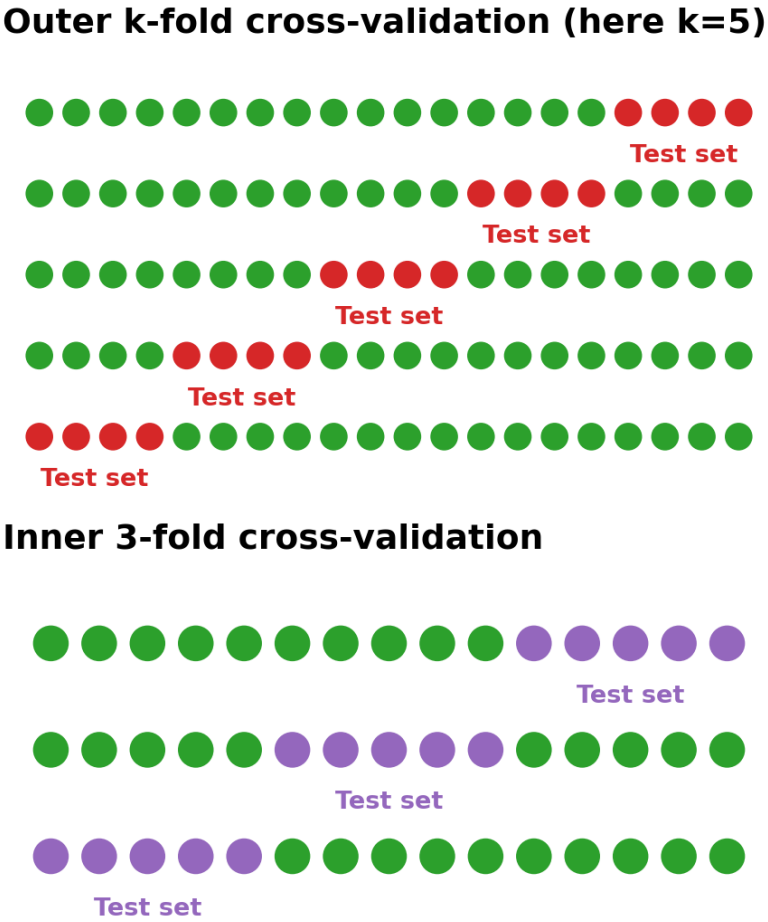
Bayesian last layer (BLL) model, reintroduces Bayesian logistic regression

- Cross-entropy loss function
  - Minimizes KL-divergence of true vs. pred. distributions
- AdamW optimizer algorithm
  - Applies weight decay directly to the weights, decouples LR
- LR = 1e-3, WD = 1e-2, max epochs = 300, batch size = 64
- 80-20 training validation split
- Early stopping on validation loss
  - patience=150, threshold=1e-4
- LR scheduler on validation loss
  - patience = 50, factor = 0.5, min LR = 1e-6

Cross Subject

- Highest variation, lowest SNR
- Realistic generalization for transfer learning research
- By default MOABB uses both outer & inner cross-validation
  - Outer k-fold CV: Avg scores
  - Inner 3-fold CV: Max scores, hyperparameter optimization
- Nested CV is good practice when training is fast and datasets are small [25]

$$\text{No. folds} = \begin{cases} \text{No. subjects} & \text{if No. subjects} < 10, \\ 5 & \text{if No. subjects} < 50, \\ 10 & \text{Otherwise.} \end{cases}$$



MOABB cross subject evaluation

Multilevel Meta-Analysis

1. Sampling error acknowledged (i.e., aleatoric uncertainty)
2. Three categories of pipelines
3. One primary study per dataset

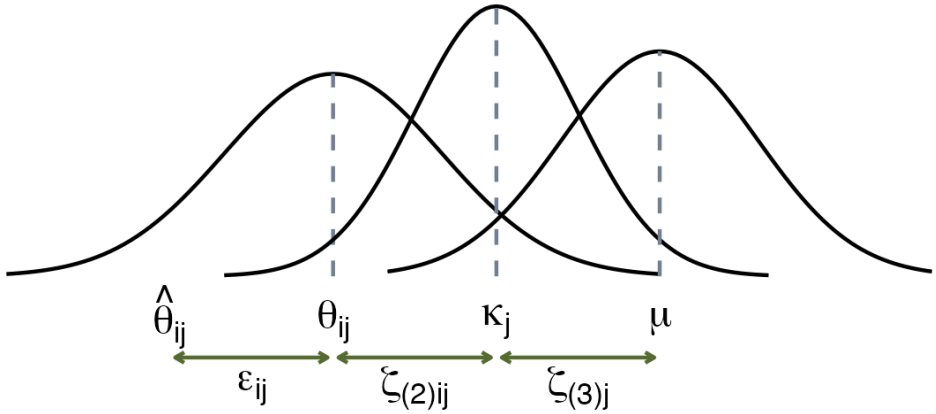
$$\text{(Level 1)} \quad \hat{\theta}_{ij} = \theta_{ij} + \epsilon_{ij}$$

$$\text{(Level 2)} \quad \theta_{ij} = \kappa_j + \zeta_{(2)ij}$$

$$\text{(Level 3)} \quad \kappa_j = \mu + \zeta_{(3)j}$$

$$\hat{\theta}_{ij} = \mu + \zeta_{(3)j} + \zeta_{(2)ij} + \epsilon_{ij}$$

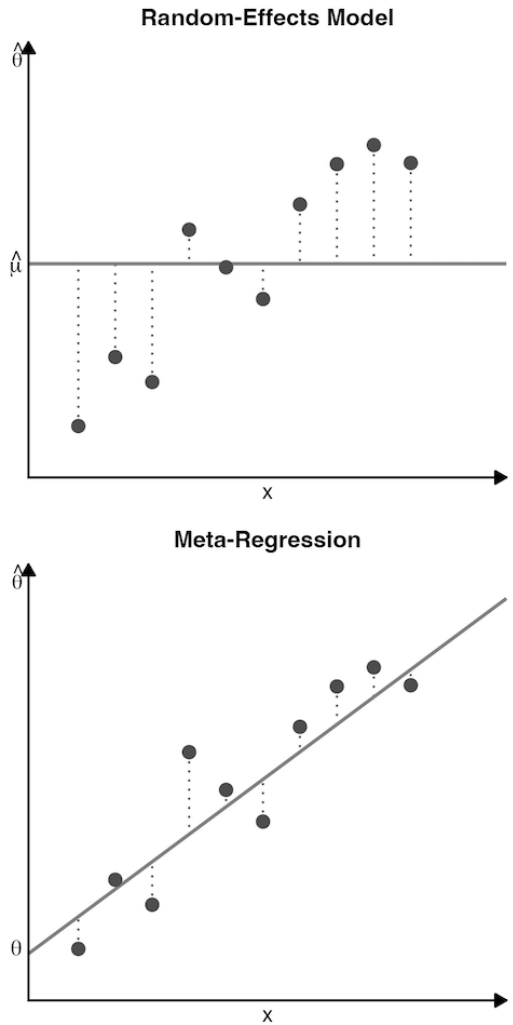
MOABB Limitation: Uses Stouffer's z-score to combine p-values of paired prediction differences across datasets



3-Level Meta-Analysis: $\hat{\theta}_{ij}$ estimates the true effect size θ_{ij} for some draw i in cluster j. κ_j is the average effect size of cluster j, and μ the overall population effect. ϵ_{ij} is the sampling error, $\zeta_{(2)ij}$ the within-cluster heterogeneity, and $\zeta_{(3)j}$ the between-cluster heterogeneity [10].

Intercept-Only & Meta-Regression

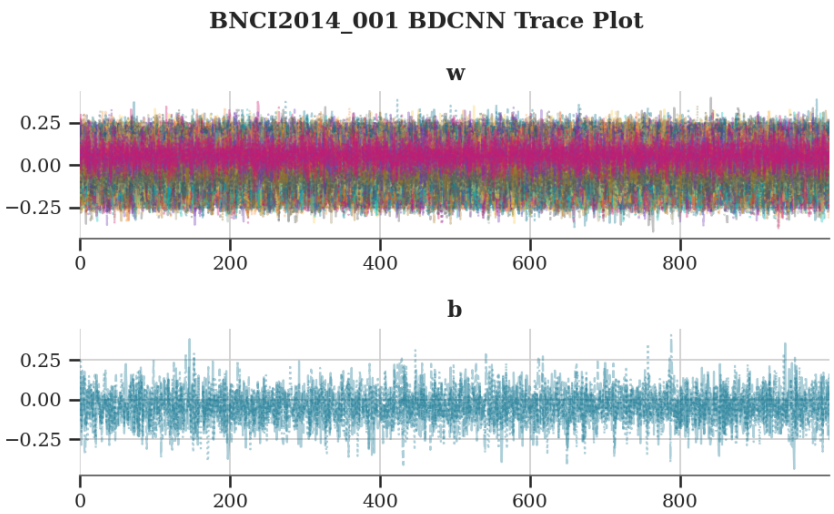
- One meta-analysis per prediction metric (AUROC, MCC, NLL, BS, ECE, MCE)
- Moderator analysis on sample size, not computational costs which would be too relative
- Assumed CHE, applied RVE
  - CR2 bias-reduced linearization
  - SA: $\rho = 0.2, 0.4, 0.6, 0.8$
- Meta-regression also used CWB with 2000 bootstrap iterations
- KH adjustment based on Student's t-distribution
- Restricted maximum likelihood



Outline

Convergence Diagnostics

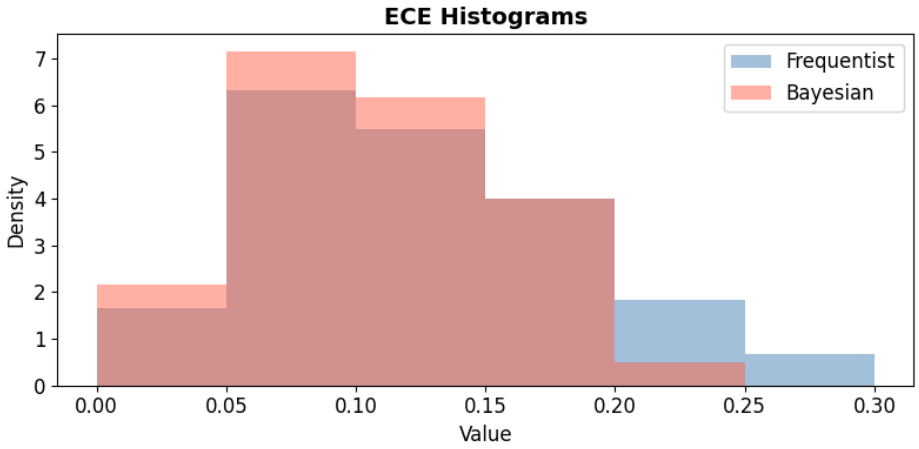
- Strong diagnostics, reassuring for further data analysis
- Per dataset+pipeline: total divergences ≤ 7, mean MCSE ≤ 0.06, std MCSE ≤ 0.03
- Trace, autocorrelation, and rank plots via our Figshare DOI [4]

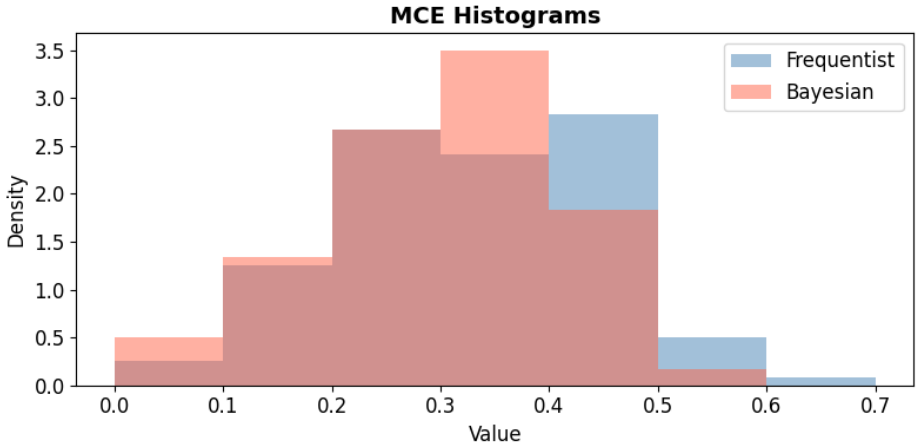


Limits Exceeded	
Threshold	Total
min bulk ESS < 400	7/120 (5.83%)
min bulk ESS < 400	1/120 (0.83%)
max $\hat{R} > 1.01$	2/120 (1.67%)
divergences > 0	4/120 (3.33%)

Portion of dataset and pipeline combinations exceeding MCMC convergence thresholds over k-fold cross-validation

Data Exploration



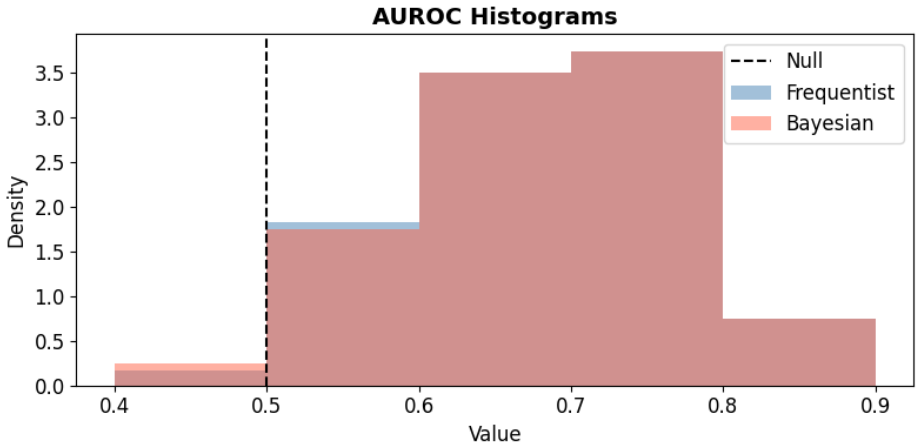


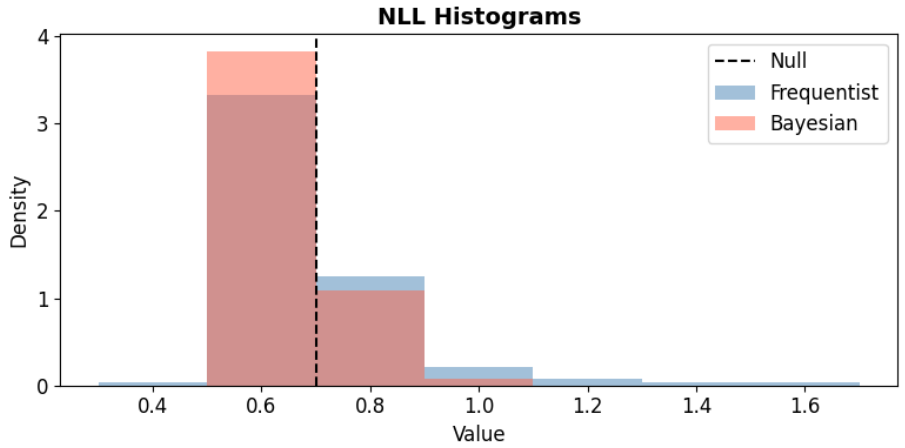
Threshold Crossings

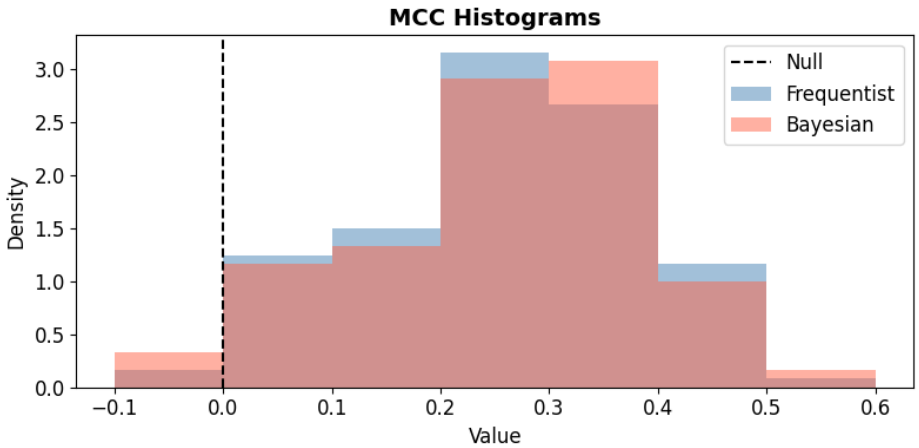
Limit	Freq.	Bayes.
AUROC < 0.5	1.7%	2.5%
MCC < 0.0	2.5%	3.3%
NLL > 0.7	32.5%	23.3%
BS > 0.25	26.7%	15.8%

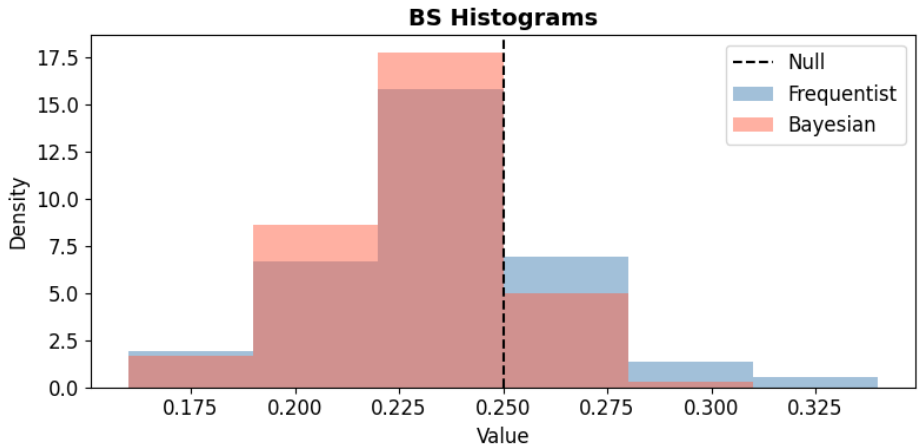
This shows that discrimination is much easier than probability calibration in left/right hand MI-EEG classification

Data Exploration (Cont.)









Intercept-Only: Pooled Effect Size

Intercept-Only: Pooled Effect Size

Metric	Estimate	SE	$t(59)$	p	95% CI LB	95% CI UB
AUROC	0.0028	0.0014	1.9616	0.0545	-0.0001	0.0056
MCC	0.0035	0.0025	1.4178	0.1615	-0.0014	0.0084
NLL	0.0024	0.0024	1.0159	0.3138	-0.0024	0.0073
BS	-0.0007	0.0005	-1.5461	0.1274	-0.0016	0.0002
ECE	-0.0058	0.0016	-3.6658	**0.0005**	-0.0090	-0.0027
MCE	-0.0272	0.0085	-3.1992	**0.0022**	-0.0443	-0.0102

RVE accounting for CHE yielded similar conclusions for significant effects

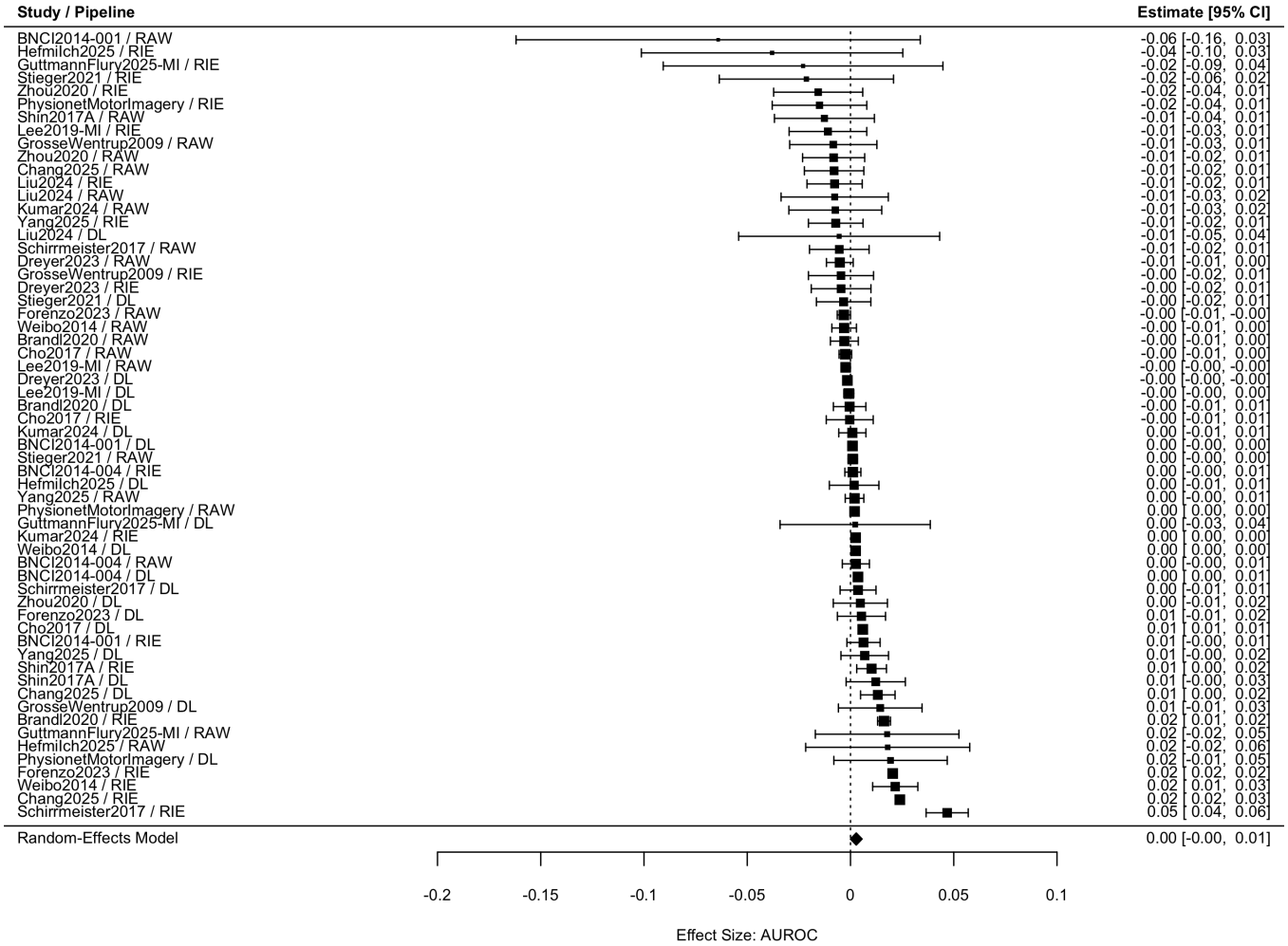
Intercept-Only: Variance Components

Intercept-Only: Variance Components

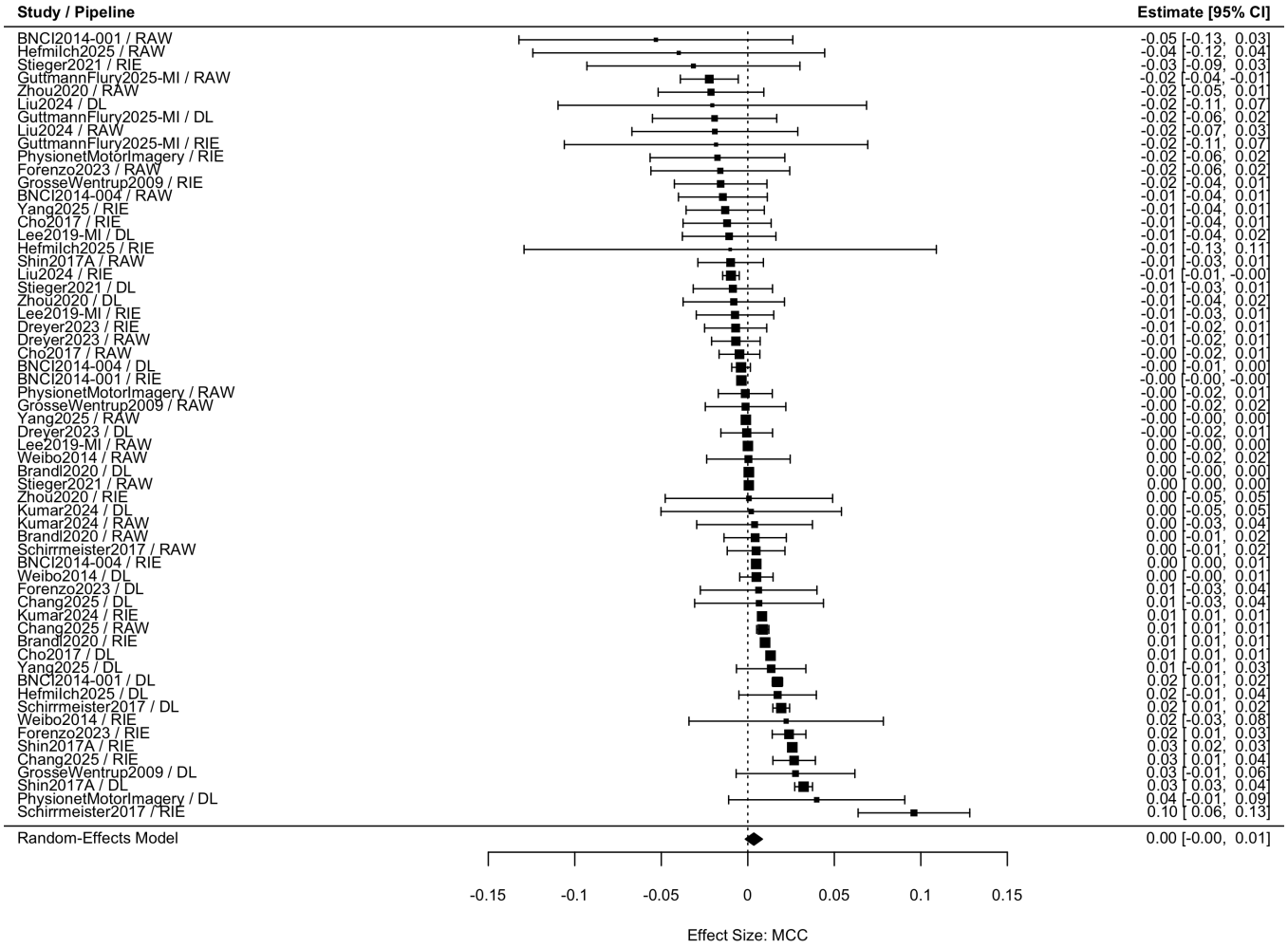
Metric	$Q(59)$	I_3^2	I_2^2	95% PI LB	95% PI UB
AUROC	2848.6744	0.0000	98.7439	-0.0154	0.0210
MCC	2240.8959	30.8506	68.5230	-0.0222	0.0292
NLL	2866.4233	0.0000	97.7473	-0.0231	0.0280
BS	384.4948	0.0000	80.4427	-0.0051	0.0037
ECE	694.9828	0.0000	95.8991	-0.0227	0.0111
MCE	114816.5782	0.9936	98.9607	-0.1431	0.0886

For all metrics Cochran's Q yielded a p-value < 0.0001. Higgins & Thomspon's I^2 statistic measures the percentage of variation caused by true heterogeneity rather than sampling error. Prediction intervals bound the effect of a new study.

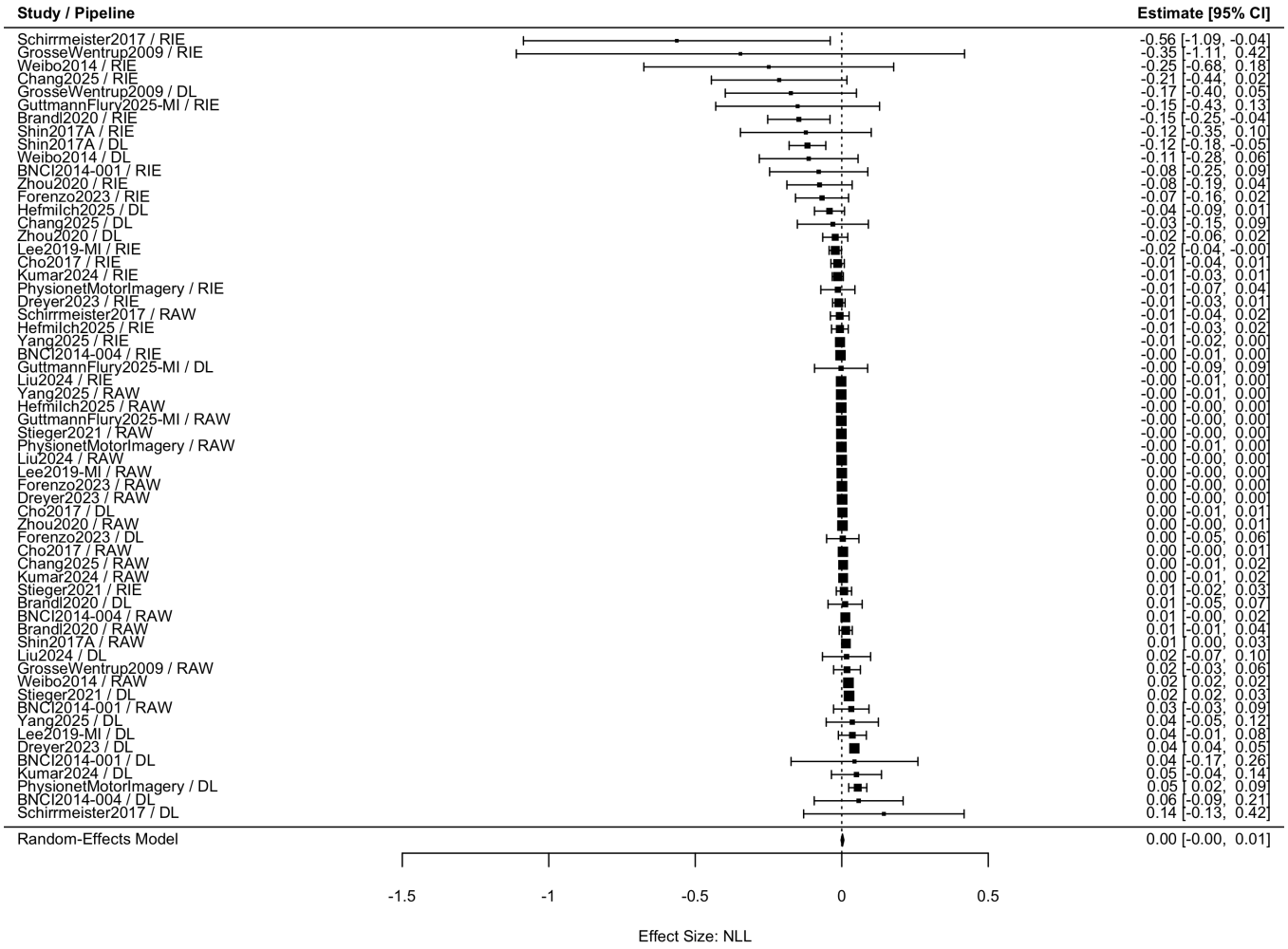
Intercept-Only: AUROC Forest Plot



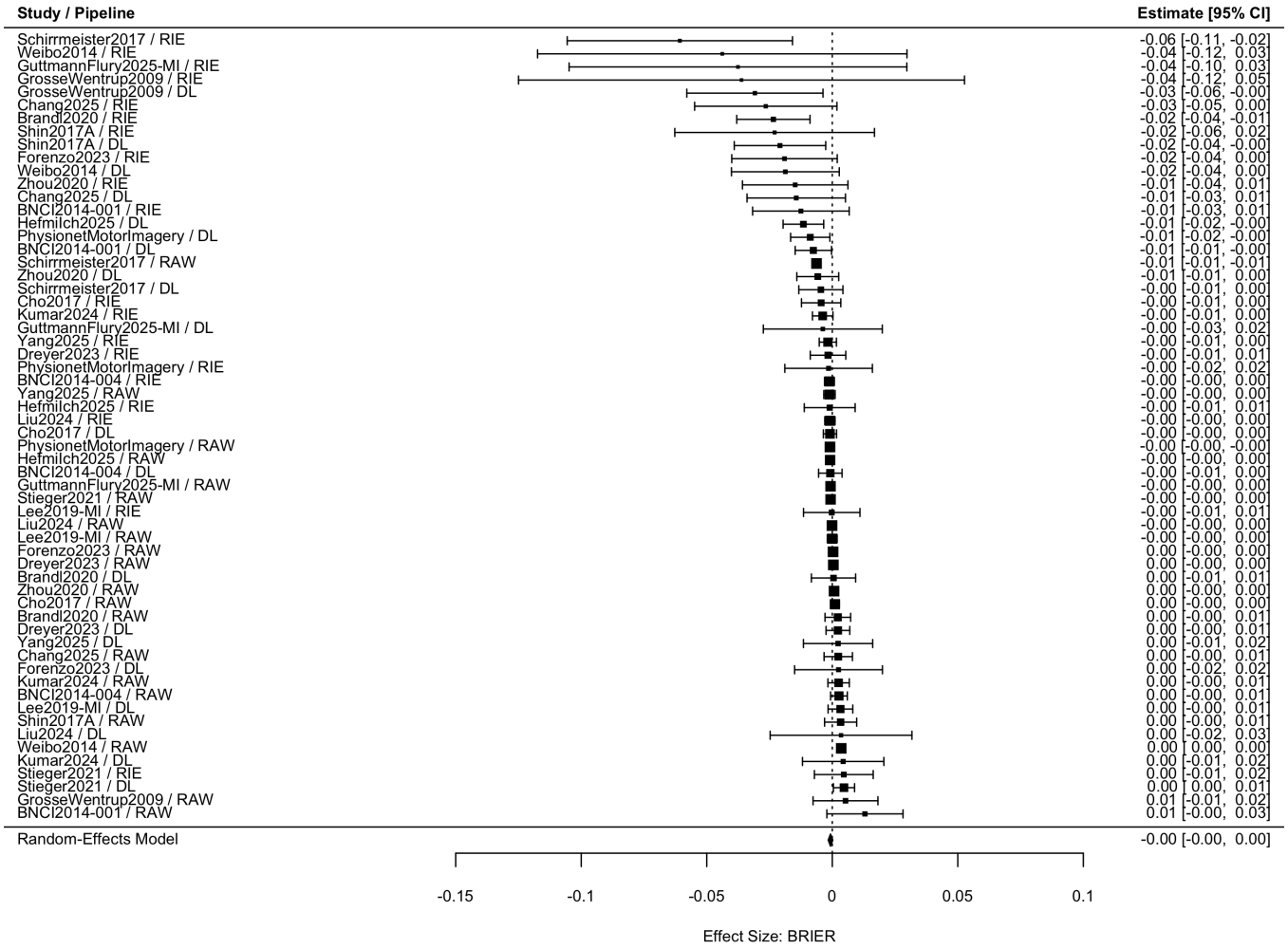
Intercept-Only: MCC Forest Plot



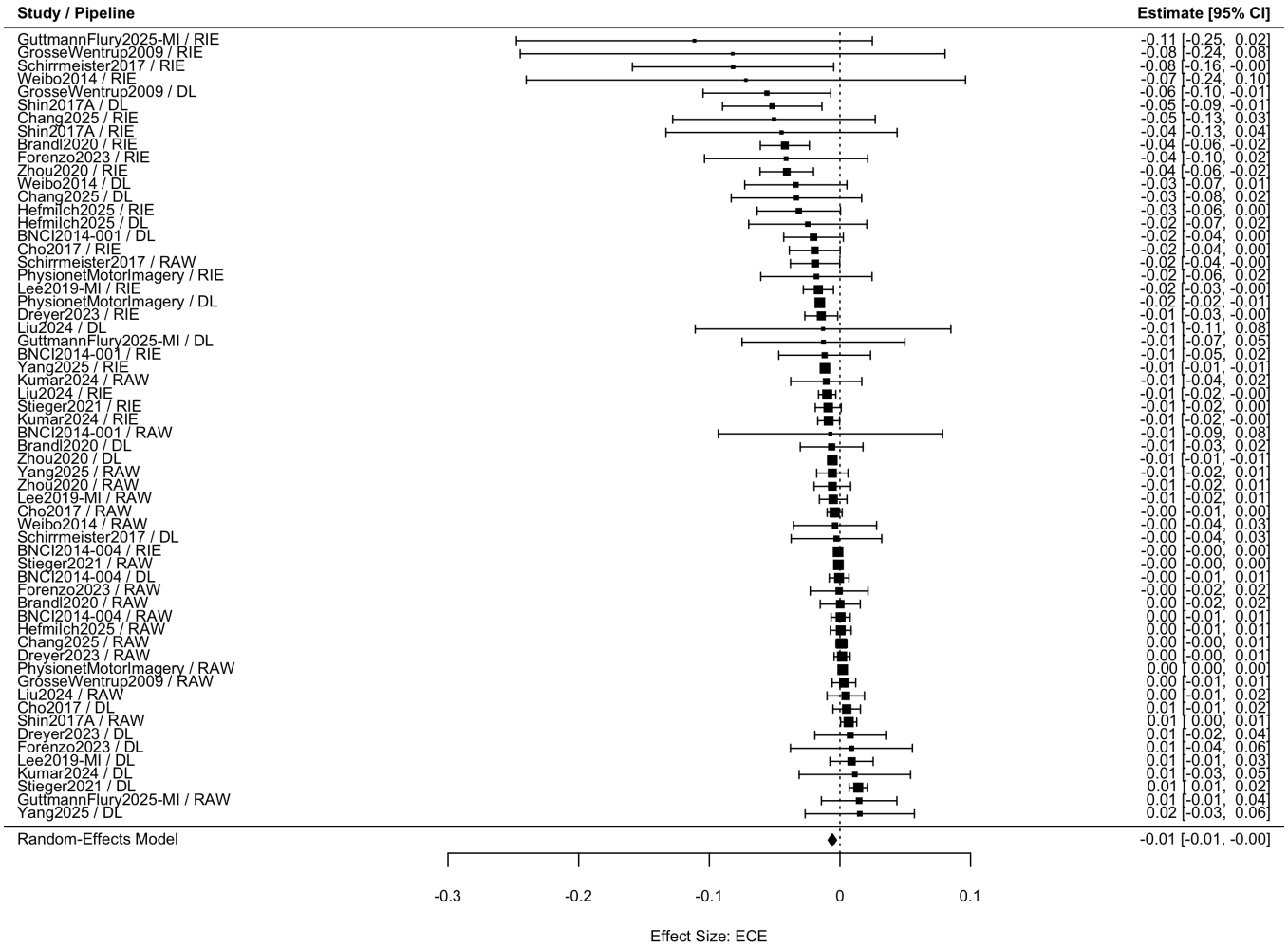
Intercept-Only: NLL Forest Plot



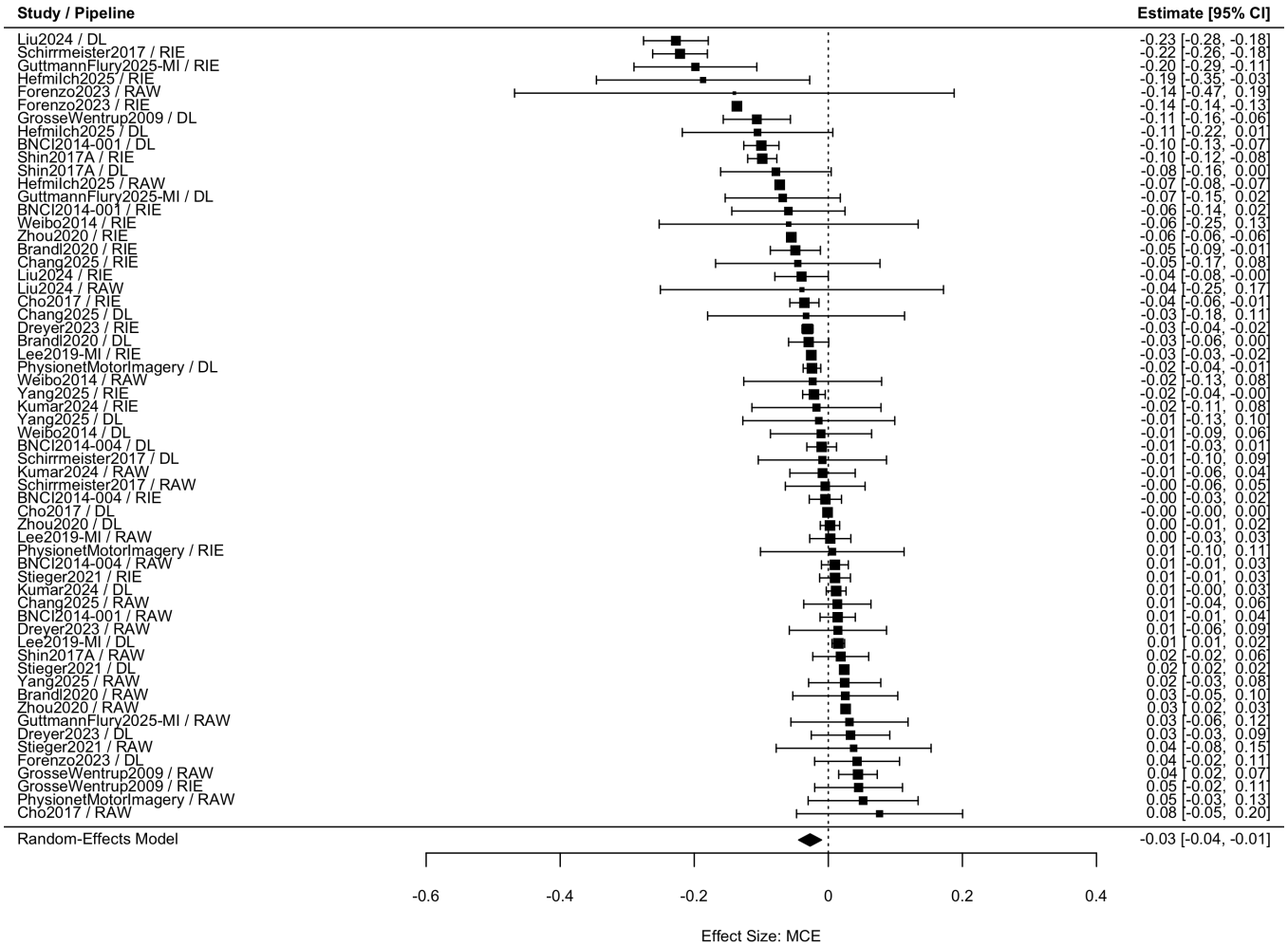
Intercept-Only: BS Forest Plot



Intercept-Only: ECE Forest Plot



Intercept-Only: MCE Forest Plot



Meta-Regression: Pooled Effect Size

Meta-Regression: Pooled Effect Size

Metric	Model	Estimate	SE	$t(58)$	p	95% CI LB	95% CI UB
AUROC	β_0	0.0029	0.0014	2.0949	**0.0406**	0.0001	0.0057
AUROC	β_1	-0.0023	0.0014	-1.5664	0.1227	-0.0051	0.0006
MCC	β_0	0.0037	0.0024	1.5460	0.1275	-0.0011	0.0085
MCC	β_1	-0.0030	0.0024	-1.2895	0.2024	-0.0078	0.0017
NLL	β_0	0.0017	0.0025	0.6815	0.4983	-0.0033	0.0067
NLL	β_1	0.0024	0.0020	1.1546	0.2530	-0.0017	0.0064
BS	β_0	-0.0008	0.0005	-1.7524	0.0850	-0.0018	0.0001
BS	β_1	0.0006	0.0004	1.2949	0.2005	-0.0003	0.0014
ECE	β_0	-0.0065	0.0017	-3.8480	**0.0003**	-0.0098	-0.0031
ECE	β_1	0.0018	0.0014	1.3316	0.1882	-0.0009	0.0046
MCE	β_0	-0.0280	0.0082	-3.4297	**0.0011**	-0.0444	-0.0117
MCE	β_1	0.0189	0.0082	2.3215	**0.0238**	0.0026	0.0353

Estimated intercept (β_0) and slope (β_1). Based on RVE, we conservatively concluded AUROC β_0 was not significant. Given CWB, we concluded β_1 was generally not significant per standard deviation increase in sample size.

Computational Costs

- CodeCarbon was used during evaluation for compute profiling
- Median training duration was 5.78× higher for Bayesian ML
- 7.19× more energy consumed
- Training conducted in WA, USA
  - 3.8^{th} %ile USA grid intensity
- On average, fitting Bayesian ML cost more than daily charge of smartphone 19.178 Wh [1]
  - Frequentist ML cost less

Training Costs: Duration (s)			
Pipeline	Q1	Median	Q3
CSPLDA	7.064	7.753	9.129
CSPBLDA	228.296	329.301	563.299
CSPSVM	7.647	10.779	16.034
CSPGP	193.5	308.212	334.453
TSLR	7.807	10.945	14.939
TSBLR	40.932	55.223	71.469
TSSVM	15.037	44.645	183.5
TSGP	112.267	121.267	137.281
SCNN	34.614	70.052	111.64
BSCNN	150.83	208.774	300.183
DCNN	41.389	62.06	104.022
BDCNN	110.889	168.541	247.232

Computational Costs (Cont.)

Training Costs: Emissions (g CO_2 eq)			
Pipeline	Q1	Median	Q3
CSPLDA	0.059	0.065	0.077
CSPBLDA	2.125	3.151	4.894
CSPSVM	0.064	0.09	0.133
CSPGP	2.911	4.871	5.331
TSLR	0.065	0.102	0.136
TSBLR	0.539	0.794	1.089
TSSVM	0.125	0.369	1.515
TSGP	1.621	1.79	1.983
SCNN	0.43	0.912	1.604
BSCNN	2.126	3.557	5.294
DCNN	0.492	0.773	1.292
BDCNN	1.481	2.44	3.989

Training Costs: Energy (Wh)			
Pipeline	Q1	Median	Q3
CSPLDA	0.696	0.767	0.903
CSPBLDA	25.076	37.176	57.74
CSPSVM	0.757	1.061	1.571
CSPGP	34.349	57.479	62.902
TSLR	0.772	1.198	1.605
TSBLR	6.356	9.37	12.852
TSSVM	1.472	4.351	17.877
TSGP	19.125	21.124	23.392
SCNN	5.07	10.767	18.921
BSCNN	25.088	41.968	62.465
DCNN	5.809	9.118	15.242
BDCNN	17.48	28.789	47.07

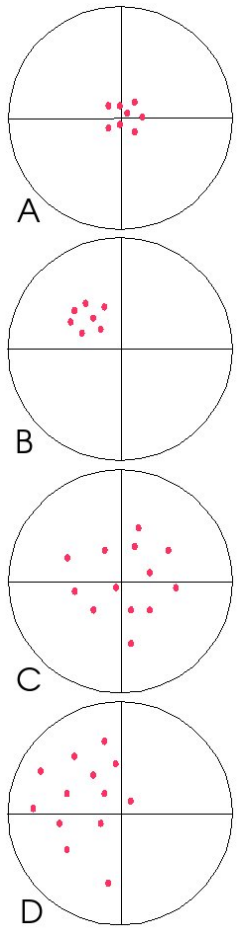
Outline

Interpretation & New Hypothesis

Probabilistic Forecasts calibration = statistical reliability, sharpness = prediction concentration [8]

- Bayesian effect on AUROC, MCC was not significant
  → Discrimination too coarse to see sharpness change
- Proper scoring effect was not significant, calibration effect was → Overconfident frequentist models, inferred change in sharpness countered calibration
- Significant yet marginal calibration (ECE = -0.58%, MCE = -2.72%), high relative cost but absolute low
  → Bayesian ML generally not justified in static models

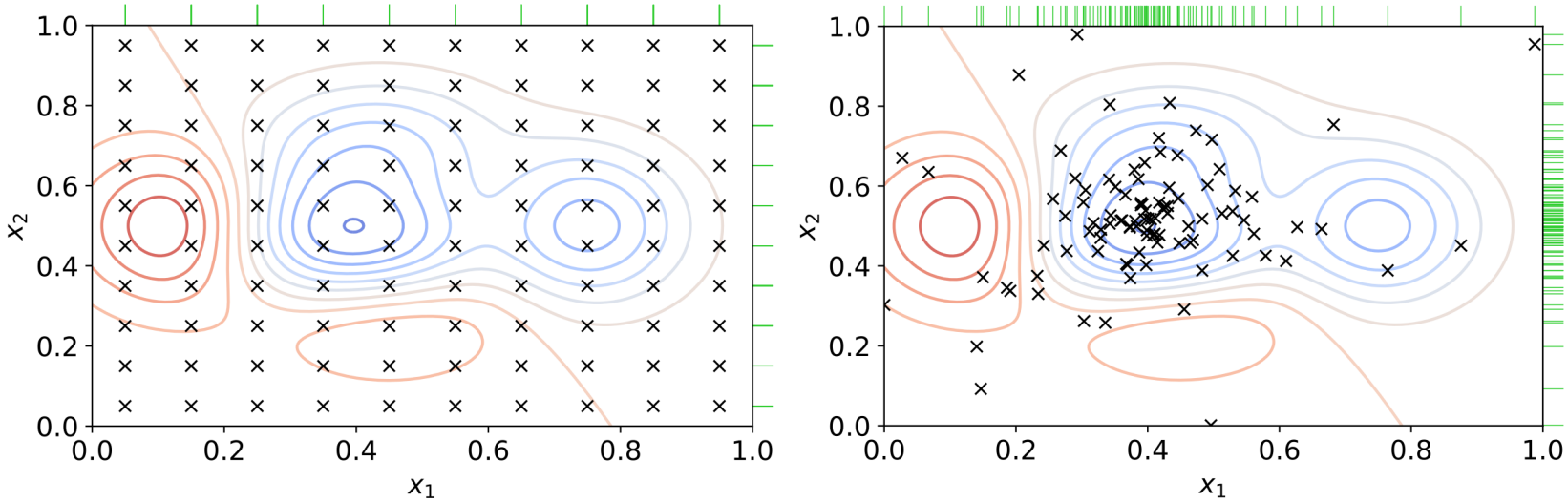
New Hypothesis The main asset of Bayesian ML is modeling epistemic uncertainty in adaptive learning



Accuracy vs. precision [9]

Limitations

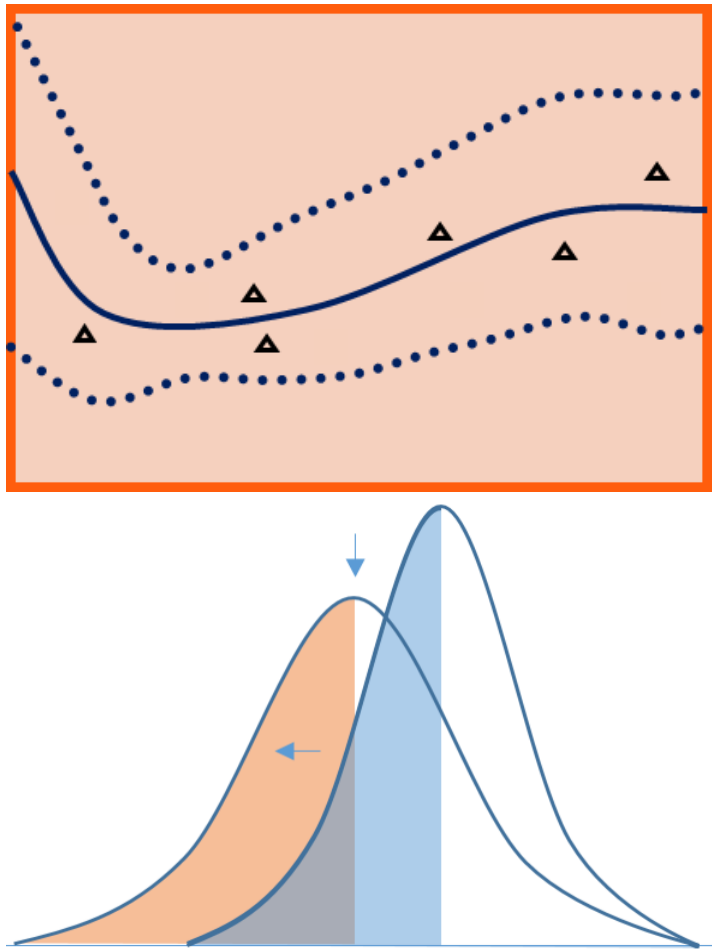
- Few pipelines per category (2 data points) likely underestimated variance components yielding somewhat anti-conservative inferences
  - Yet a sensitivity analysis not grouping by category gave similar results
- Absence of formal hyperparameter optimization reduces robustness
  - However choices were informed by literature and standardization
- Conclusions based on multiple MAs and a small number of studies
  - Our results motivate multivariate MA via SEMs and Bayesian MA



HPO using grid search and tree-structured Parzen estimators (TPEs) [5, 6]

Future Work

- Research generated hypothesis
  - The BCI literature: TL[a] benefits from Bayesian ML [13]
  - SL[b] research directions: RL, dynamical systems, HMMs
- Study theoretical generalizability
  - Statistical learning theory
- Rework our experiment design
  - Explicit analysis of sharpness
    - Measure Shannon entropy
  - More pipelines per category
    - GNNs for deep learning

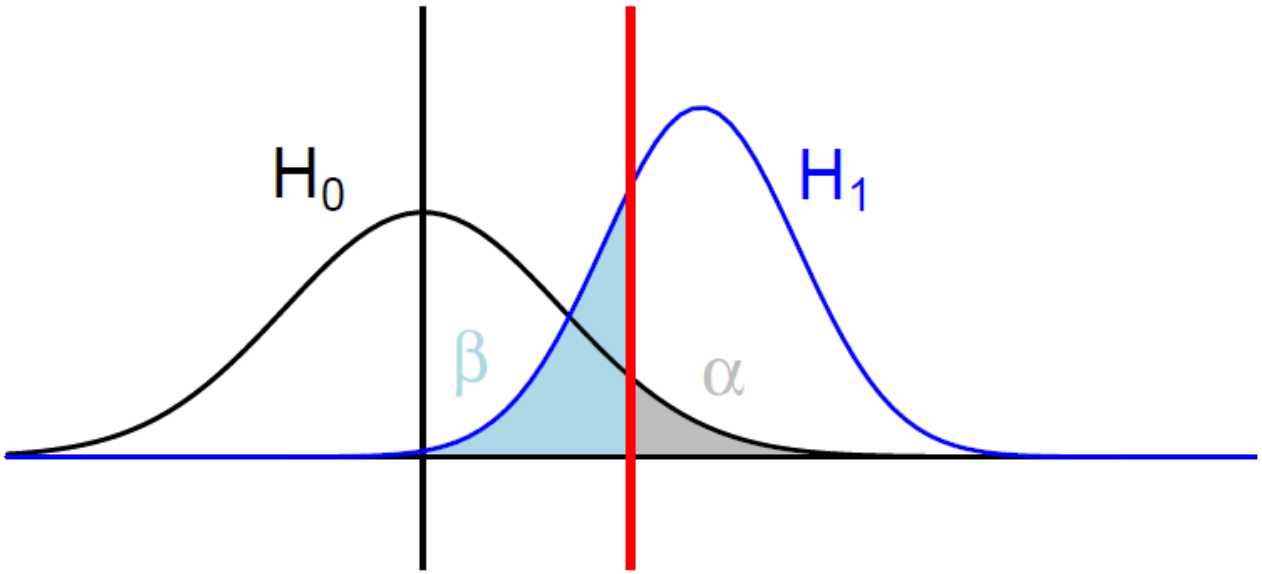


(↑) Prediction with uncertainty quantification [14], (↓) Distribution shift

[a] TL = transfer learning
[b] SL = sequential learning

Outline

Summary



- Statistically significant yet marginal effect size for calibration metrics
- Bayesian ML is computationally inexpensive for MI-EEG classification
- Generally Bayesian ML is not justified for static models in BCI
- Hypothesis: Primary advantage of Bayesian ML is adaptive learning

Digital Object Identifiers

GitHub https://doi.org/10.5281/zenodo.20415596

Data https://doi.org/10.6084/m9.figshare.32176734

Questions

Outline

References I

[1] La librairie ADEME. *Panel usages électrodomestiques année 6*. Mar. 2026. URL: https://librairie.ademe.fr/batiment/9155-panel-usages-electrodomestiques-annee-6.html.

[2] Sylvain Chevallier et al. "The largest EEG-based BCI reproducibility study for open science: the MOABB benchmark". working paper or preprint. Apr. 2024. URL: https://universite-paris-saclay.hal.science/hal-04537061.

[3] Marie-Constance Corsi. "Electroencephalography and Magnetoencephalography". In: *Machine Learning for Brain Disorders*. Ed. by Olivier Colliot. New York, NY: Springer US, 2023, pp. 285–312. ISBN: 978-1-0716-3195-9. DOI: 10.1007/978-1-0716-3195-9_9. URL: https://doi.org/10.1007/978-1-0716-3195-9_9.

References II

[4] Ethan Davis. *Bayesian and Frequentist Motor Imagery EEG Classification: Model Scores, Compute Profiling, and MCMC Convergence Diagnostics*. May 2026. DOI: 10.6084/m9.figshare.32176734.v2. URL: https://figshare.com/articles/dataset/Bayesian_and_Frequentist_Motor_Imagery_EEG_Classification_Model_Scores_Compute_Profiling_and_MCMC_Convergence_Diagnostics/32176734.

[5] Alexander Elvers. *Hyperparameter Optimization using Grid Search.svg*. https://commons.wikimedia.org/wiki/File:Hyperparameter_Optimization_using_Grid_Search.svg. Licensed under CC BY-SA 4.0. 2019.

References III

[6] Alexander Elvers. *Hyperparameter Optimization using Tree-Structured Parzen Estimators.svg*. https://commons.wikimedia.org/wiki/File:Hyperparameter_Optimization_using_Tree-Structured_Parzen_Estimators.svg. Licensed under CC BY-SA 4.0. 2019.

[7] Abe Flaxman, Thomas Wiecki, and Benjamin T. Vincent. *Gaussian Mixture Model*. Ed. by PyMC Team. 2023. DOI: 10.5281/zenodo.5654871.

References IV

[8] Tilmann Gneiting, Fadoua Balabdaoui, and Adrian E. Raftery. "Probabilistic forecasts, calibration and sharpness". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69.2 (2007), pp. 243–268. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2007.00587.x.

[9] Guam commonswiki. *Accuracy and precision example.jpg*. https://commons.wikimedia.org/wiki/File:Accuracy_and_precision_example.jpg. Licensed under CC BY-SA 3.0. 2005.

[10] Mathias Harrer et al. *Doing Meta-Analysis With R: A Hands-On Guide*. 1st. Boca Raton, FL and London: Chapman & Hall/CRC Press, 2021. ISBN: 9780367610074.

References V

[11] HgDeviasse. *3112189 pone.0020674.g001*. https://commons.wikimedia.org/wiki/File:3112189_pone.0020674.g001.png. Licensed under public domain. 2014.

[12] Iamozy. *Human motor cortex.jpg*. https://commons.wikimedia.org/wiki/File:Human_motor_cortex.jpg. Licensed under CC BY-SA 3.0. 2014.

[13] Vinay Jayaram et al. "Transfer Learning for BCIs". In: *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. Ed. by Chang S. Nam, Anton Nijholt, and Fabien Lotte. CRC Press, 2018. Chap. 22, pp. 425–441.

[14] Jiangzhenjerry. *Bias Correction.png*. https://commons.wikimedia.org/wiki/File:Bias_Correction.png. Licensed under CC BY-SA 3.0. 2012.

References VI

[15] Justinkunimune. *Gradient descent maximum likelihood.svg*. https://commons.wikimedia.org/wiki/File:Gradient_descent_maximum_likelihood.svg. Licensed under CC0 1.0. 2024.

[16] Justinkunimune. *Hamiltonian Monte Carlo.svg*. https://commons.wikimedia.org/wiki/File:Hamiltonian_Monte_Carlo.svg. Licensed under CC0 1.0. 2024.

[17] Fabien Lotte, Chang S. Nam, and Anton Nijholt. "Introduction: Evolution of Brain–Computer Interfaces". In: *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. Ed. by Chang S. Nam, Anton Nijholt, and Fabien Lotte. CRC Press, 2018. Chap. Introduction, pp. 1–88.

References VII

[18] MartinThoma. *Lda-gauss-1.svg*. https://commons.wikimedia.org/wiki/File:Lda-gauss-1.svg. Licensed under CC0 1.0. 2014.

[19] Kevin P. Murphy. *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023. URL: http://probml.github.io/book2.

[20] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[21] Produnis. *Forestplot01.jpg*. https://commons.wikimedia.org/wiki/File:Forestplot01.jpg. Licensed under CC BY-SA 3.0. 2006.

[22] Ralf@ark.in-berlin.de. *Motor homunculus.svg*. https://commons.wikimedia.org/wiki/File:Motor_homunculus.svg. Licensed under CC BY-SA 4.0. 2016.

References VIII

[23] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, Nov. 2005. ISBN: 978026256834. DOI: 10.7551/mitpress/3206.001.0001. eprint: https://direct.mit.edu/book-pdf/2514321/book_9780262256834.pdf. URL: https://doi.org/10.7551/mitpress/3206.001.0001.

[24] Shazam7734. *Logistic Regression.png*. https://commons.wikimedia.org/wiki/File:Logistic_Regression.png. Licensed under CC0 1.0. 2026.

References IX

[25] Gael Varoquaux and Olivier Colliot. "Evaluating Machine Learning Models and Their Diagnostic Value". In: *Machine Learning for Brain Disorders*. Ed. by Olivier Colliot. New York, NY: Springer US, 2023, pp. 601–630. ISBN: 978-1-0716-3195-9. DOI: 10.1007/978-1-0716-3195-9_20. URL: https://doi.org/10.1007/978-1-0716-3195-9_20.

[26] Yu Zhang. "Bayesian Learning for EEG Analysis". In: *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. Ed. by Chang S. Nam, Anton Nijholt, and Fabien Lotte. CRC Press, 2018. Chap. 21, pp. 407–424.