ETH zürich

Binary Models for Motor-Imagery Brain—Computer Interfaces: Sparse Random Projection and Binarized SVM

AICAS 2020

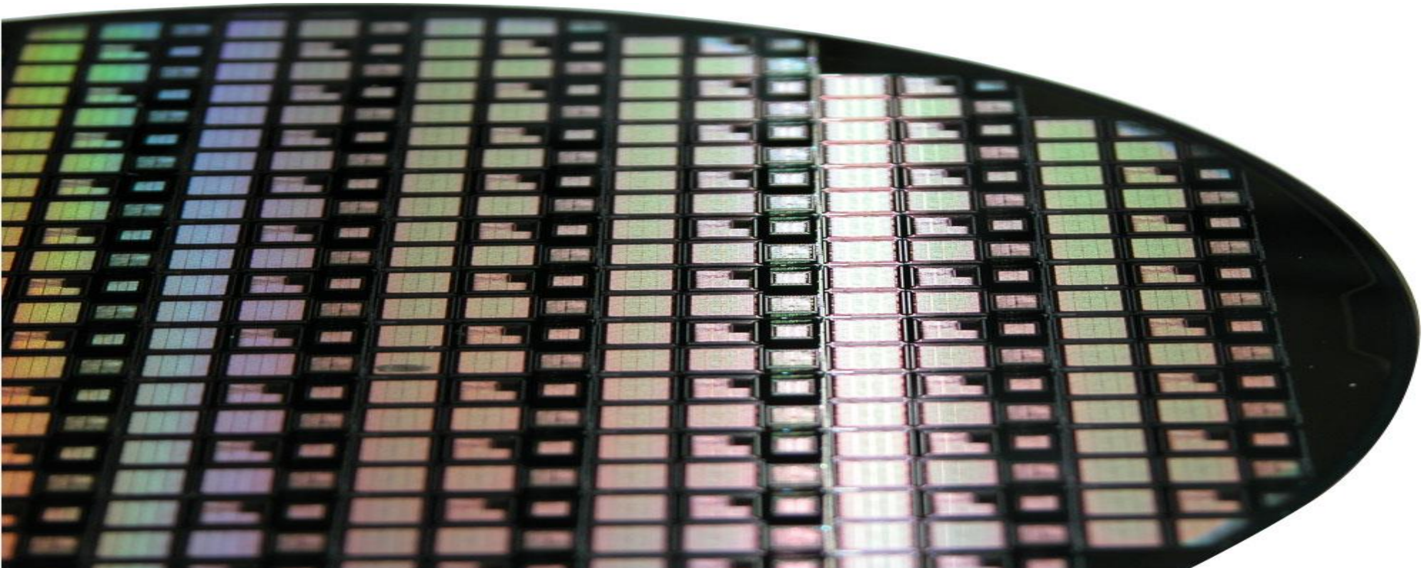
4 September 2020

Michael Hersche

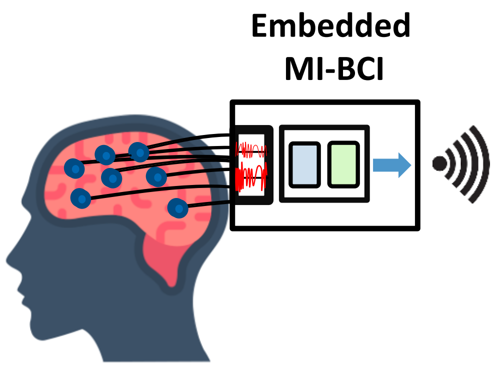
Luca Benini

Abbas Rahimi

Integrated Systems Laboratory
D-ITET ETH Zürich



Towards wearable embedded Motor-Imagery Brain – Computer Interfaces (MI-BCIs)



Why embedded MI-BCI?

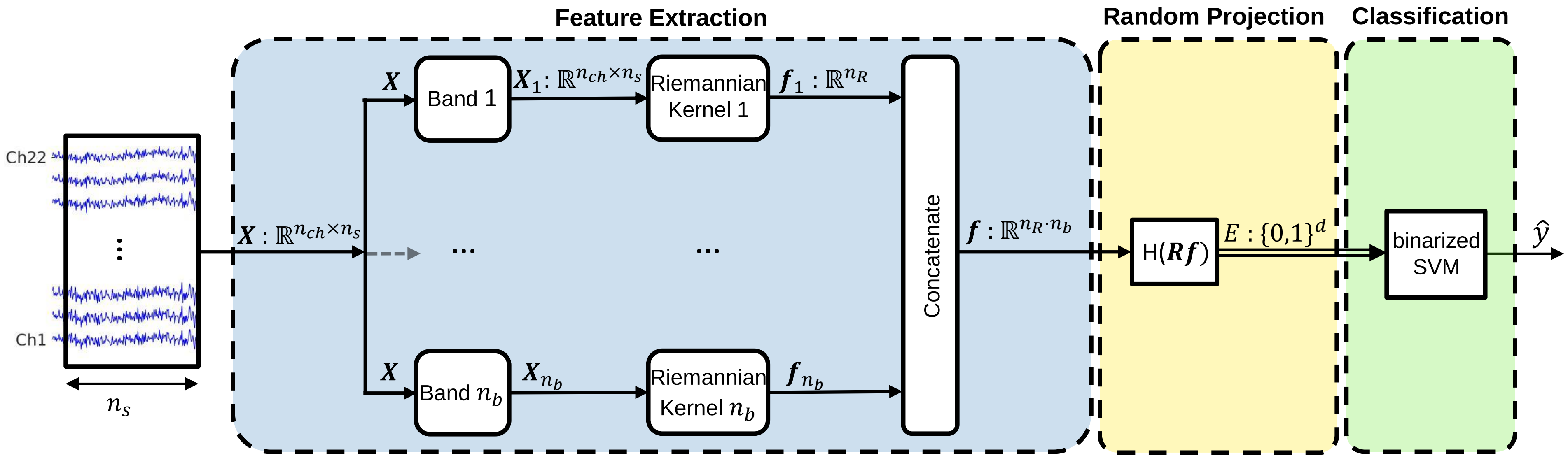
- User comfort
- Latency
- Security & privacy
- Long battery lifetime

Challenges for embedding

- Complexity
- Memory requirements of model

This work reduces the MI-BCI model size with random projection and binarized SVM

- Project Riemannian features to binary space with **random projection**
- Novel **binarized SVM** for classification
- Test on two different MI datasets
- Reduce the model size by **1.5-1.9x** with less than 1.27% accuracy loss



We test our models on two MI datasets

4-class: BCI Competition IV-2a

- 9 subjects
- 2 sessions per subject on 2 days: training and test set
- 288 trials per session and subject
- 4 MI tasks initiated by visual cue
  - Left hand/right hand/feet/tongue
- 22 EEG channels sampled with 250 Hz

3-class: Saeedi et al. [1]

- 5 subjects
- 4 sessions per subject on same day
  - 4-fold cross validation: 3 sessions for training and one for testing
- 45 trials per session and subject
- 3 MI tasks
  - Left hand/right hand/feet
- 16 EEG channels sampled with 512 Hz

[1] Saeedi, et al. "Adaptive Assistance for Brain-Computer Interfaces by Online Prediction of Command Reliability." *IEEE Computational Intelligence Magazine*. 2016.

Riemannian covariance method extracts features in unsupervised manner

- Spatial covariance matrix:

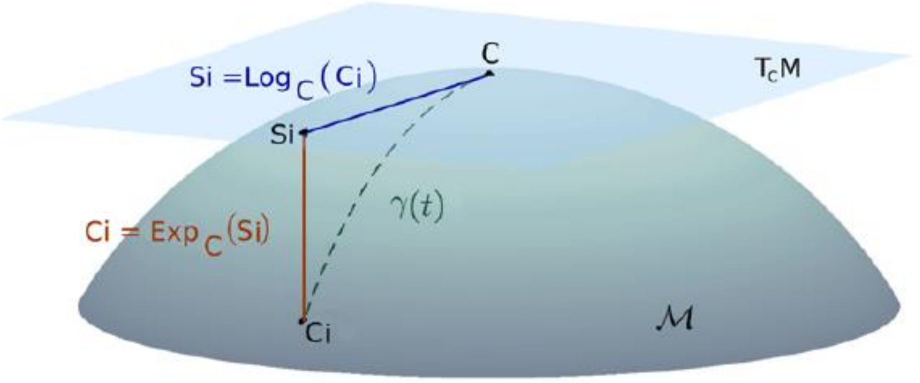
$$\boldsymbol{C} = \boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}} + \rho\boldsymbol{I} \qquad \boldsymbol{C} \in \mathbb{R}^{n_{ch} \times n_{ch}}$$

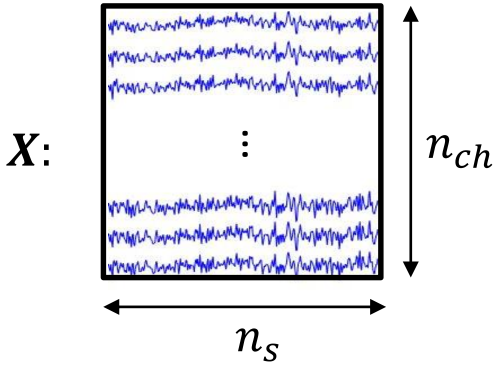
- Riemannian kernel:

$$\widetilde{\boldsymbol{C}} = \text{logm}(\boldsymbol{C}_{ref}^{-1/2}\boldsymbol{C}\boldsymbol{C}_{ref}^{-1/2})$$

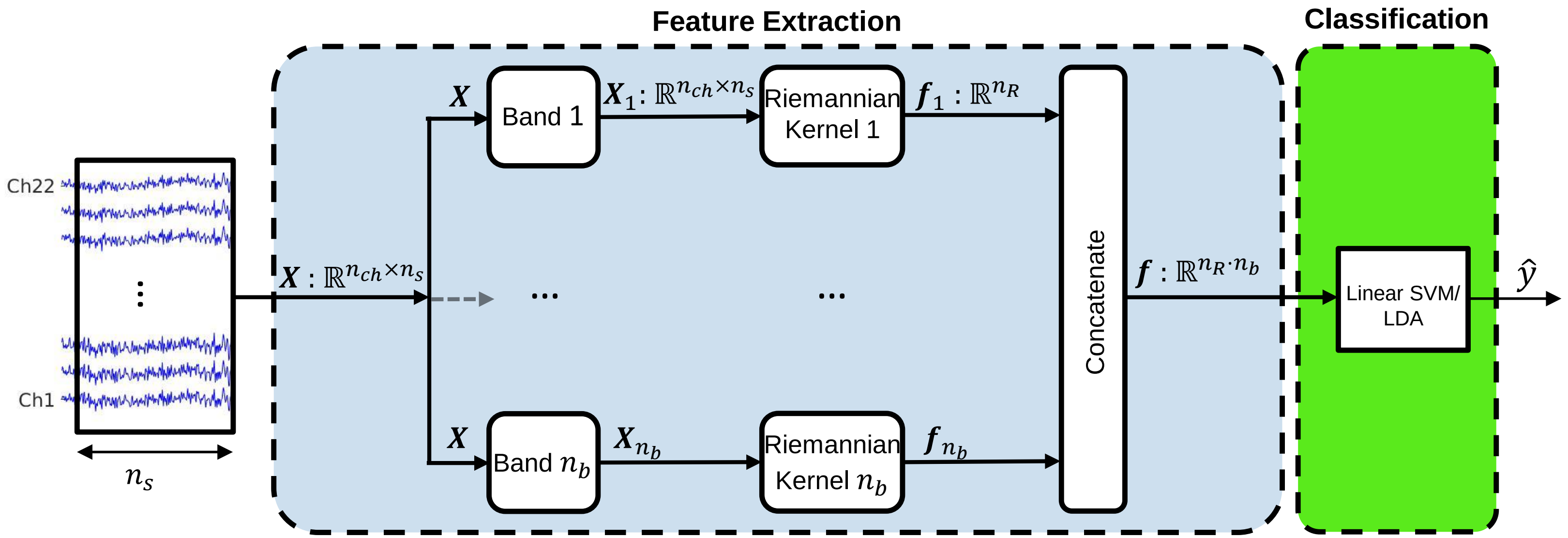
- Reference matrix $\boldsymbol{C_{ref}}$ average over all covariance matrices in training **(unsupervised)**
- Half vectorization yields $n_R = n_{ch}(n_{ch}+1)/2$ features:

$$\boldsymbol{f} = \text{vect}(\widetilde{\boldsymbol{C}})$$





State-of-the-art linear SVM or LDA on multi-spectral Riemannian features



4-class MI

- $n_b = 43$ multiscale frequency bands
- $n_R = \frac{22(22+1)}{2} = 253$ features/band

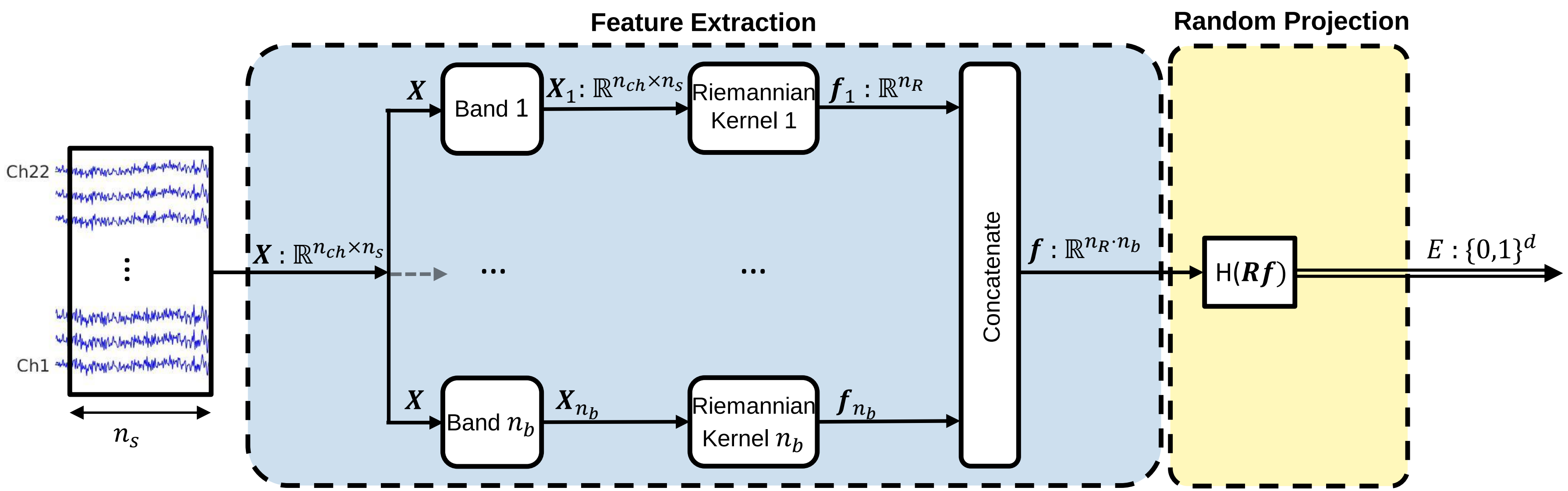
=> Total 10879 features

3-class MI

- $n_b = 13$ frequency bands
- $n_R = \frac{16(16+1)}{2} = 136$ features/band

=> Total 1768 features

Sparse bipolar random projection encodes feature vector to binary space



Sparse bipolar random projection encodes the whole feature vector at once

$$\boxed{E = \mathrm{H}(\boldsymbol{Rf})} \qquad E \in \{0,1\}^d$$

- Elementwise Heaviside function $\mathrm{H}(.)$
- Sparse bipolar random projection matrix

$$\boldsymbol{R} \in \mathbb{R}^{d\,\times} \quad \boldsymbol{R}_{i,j} = \begin{cases} 1, \mathrm{w.p.}\ \ s/2 \\ 0, \mathrm{w.p.}\ 1-s \\ -1, \mathrm{w.p.}\ s/2 \end{cases}$$

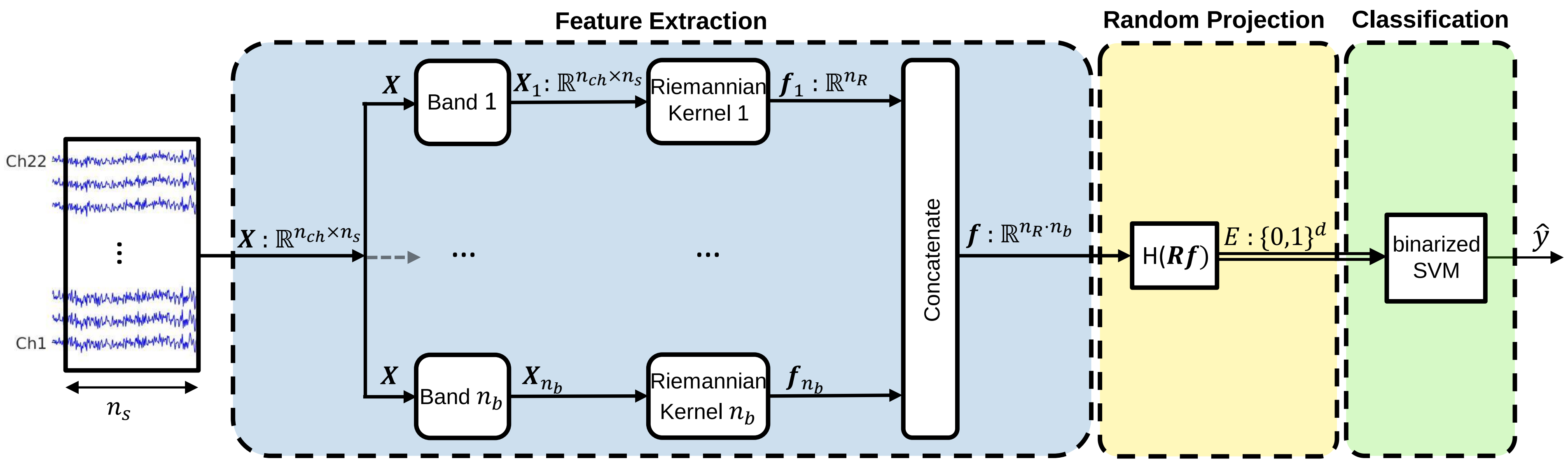
- 90% of the elements in $\boldsymbol{R}$ are 0

Implementation of random projection

- Traditional processor-based computation
  - ***R*** **regenerated with on device** random number generator
- Memristors [2]
- Optical devices [3]

[2] Du, et al. "Reservoir computing using dynamic memristors for temporal information processing." *Nature Communications. 2016*.
[3] Saade, et al. "Random projections through multiple optical scattering: Approximating Kernels at the speed of light." *IEEE ICASSP. 2016*

Binarized SVM for classification of projected binary features with **Hamming distance**



Key contribution: binarized SVM for classification of projected binary features with **Hamming distance**

Decision function original linear SVM

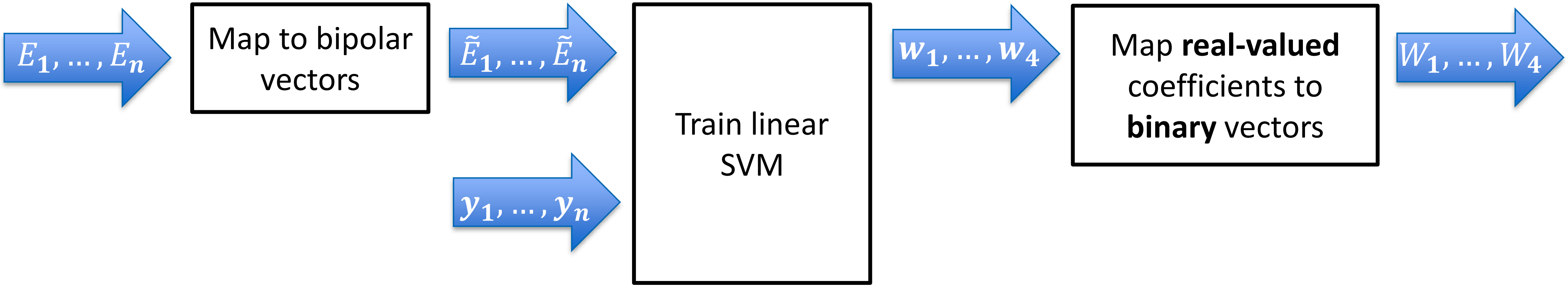
$$\hat{y} = \underset{i=\{1,2,3,4\}}{\operatorname{argmax}} \langle \boldsymbol{f}, \boldsymbol{w}_i \rangle + b_i$$

$\boldsymbol{w}_i$: support vector
b_i: bias

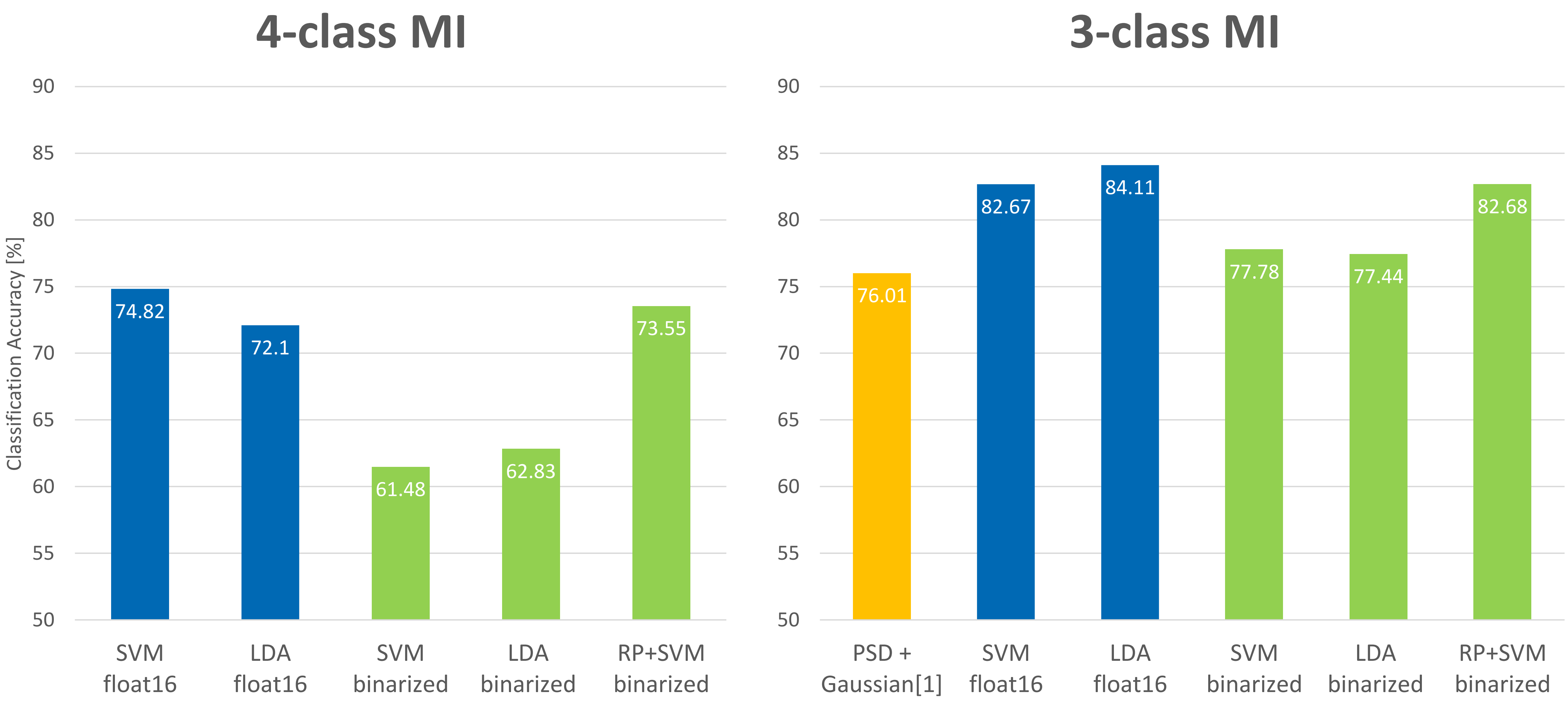
Decision function new **binarized SVM**

$$\hat{y} = \underset{i=\{1,2,3,4\}}{\operatorname{argmin}} \mathrm{d_h}(E, W_i)$$

W_i : binary support vector

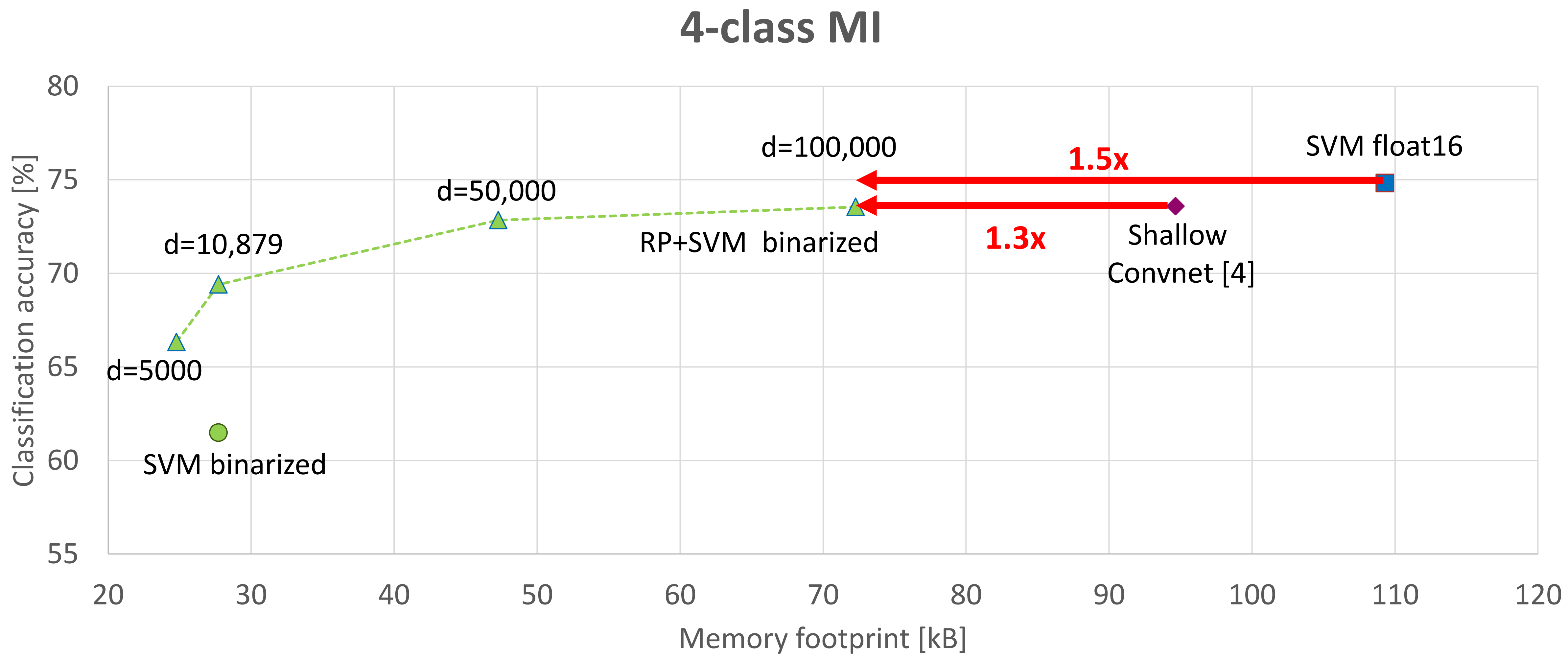


Classification results



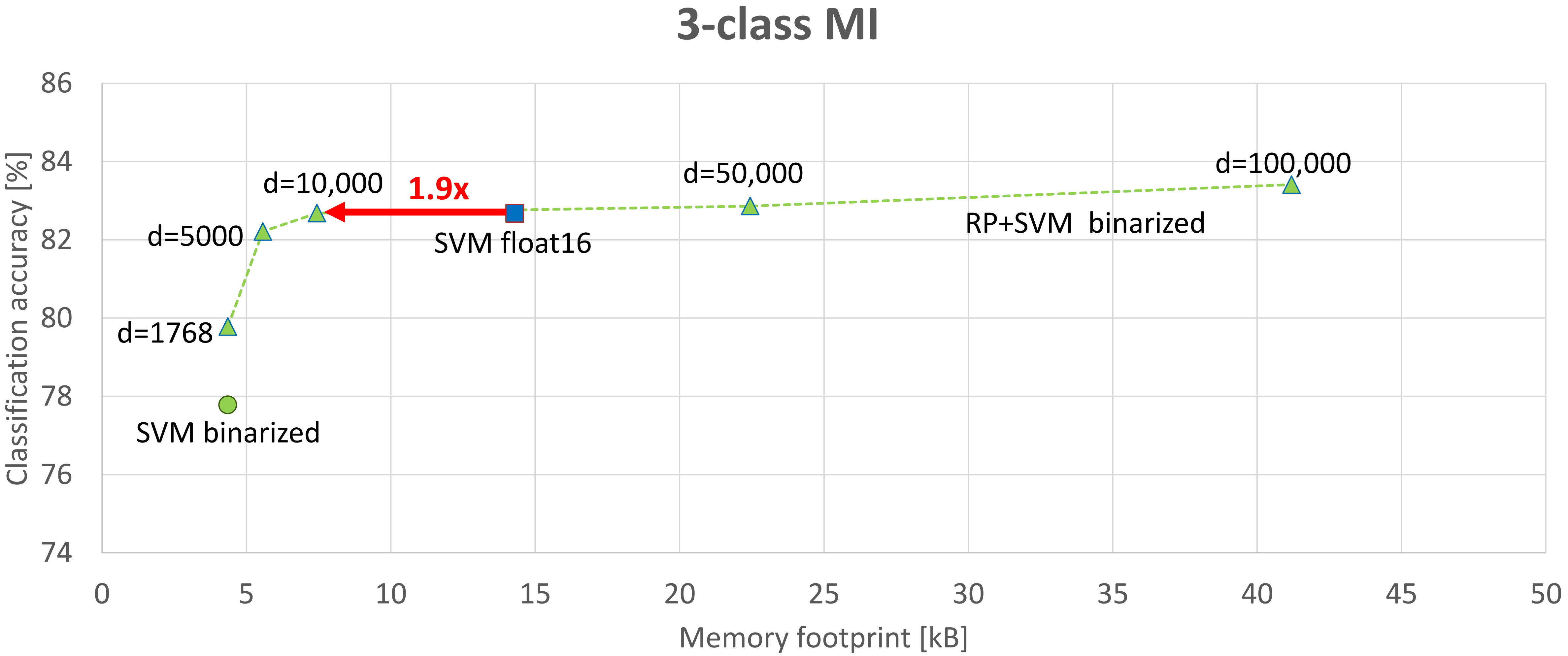
[1] Saeedi, et al. "Adaptive Assistance for Brain-Computer Interfaces by Online Prediction of Command Reliability." *IEEE Computational Intelligence Magazine*. 2016.

Our binarization method reduces the memory footprint of the model



[4] Schirrmeister, et al. “Deep learning with convolutional neural networks for EEG decoding and visualization.” *Human Brain Mapping*. 2017.

Similar results in 3-class MI



Conclusion

- Binarize multi-spectral Riemannian features with sparse bipolar random projection
- Novel binarized SVM
- Test on two MI datasets:
  - 4-class MI: 1.27% lower accuracy at **1.5x smaller model** size
  - 3-class MI: Same accuracy at **1.9x smaller model** size
- **Next:** Binarize CNN-based classifier with Memory Augmented Neural Networks (MANNs)

Code available!

https://github.com/MHersche/HDembedding-BCI

ETH zürich

Binary Models for Motor-Imagery Brain—Computer Interfaces: Sparse Random Projection and Binarized SVM

AICAS 2020

4 September 2020

Michael Hersche

Luca Benini

Abbas Rahimi

Integrated Systems Laboratory
D-ITET ETH Zürich