

# Large dimensional analysis of LS-SVM transfer learning: Application to POLSAR classification

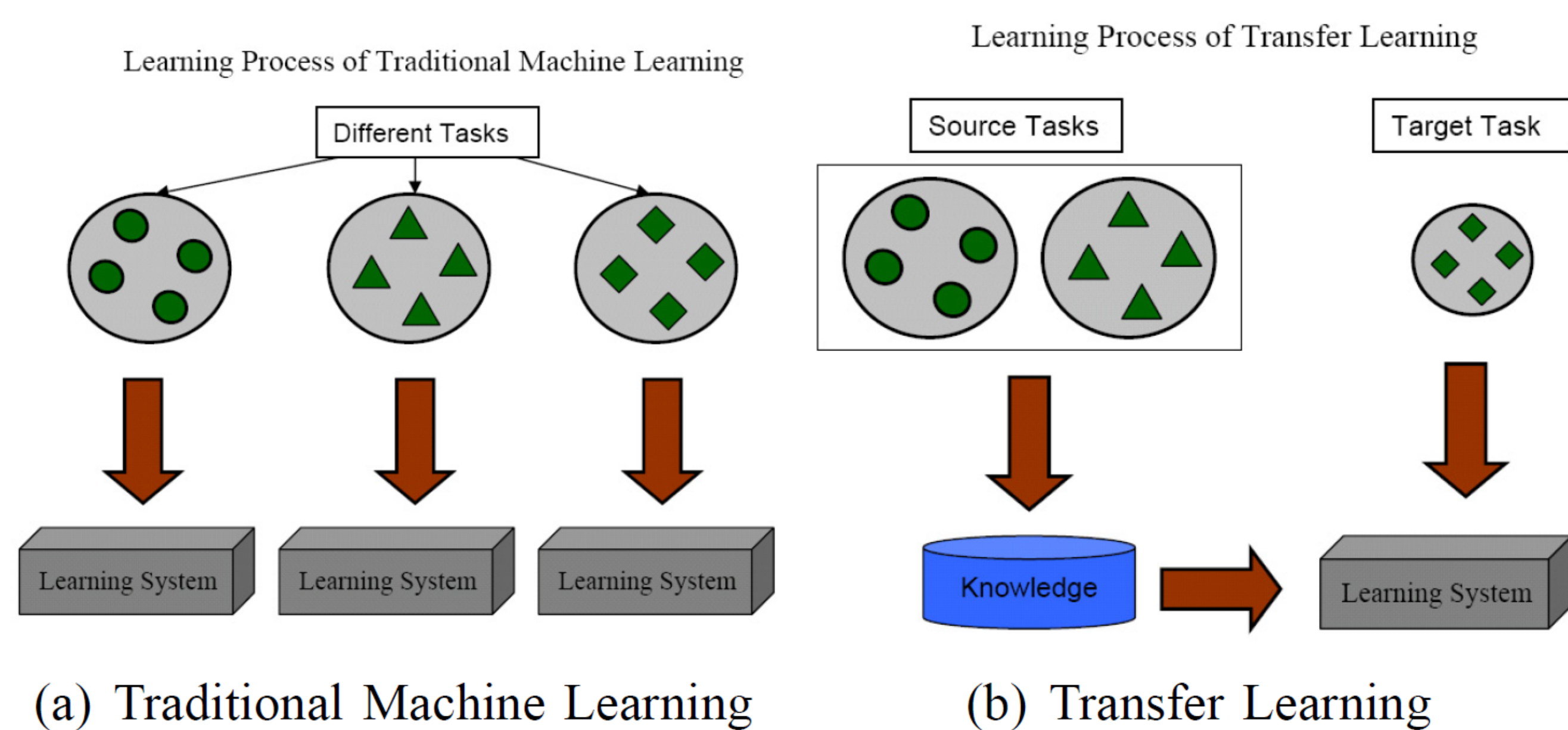
Cyprien DOZ\*, Romain COUILLET\*\*, Chengfang REN\*, Jean-Philippe OVARLEZ\*,\*\*\*

\*SONDRA, \*\*GIPSA-Lab, \*\*\*ONERA

## Abstract

- Analysis, Interpretation and Improvement of transfer learning [1] with Random Matrix Theory [2–4]
- Application to environmental monitoring : label optimization and performance guarantees

## Transfer learning framework [1]



1.  $[\mathbf{x}_1^T, \dots, \mathbf{x}_{n_T}^T]$ : target data (annotated) **insufficient**.  
 ➔ failing supervised learning
2.  $[\mathbf{x}_1^T, \dots, \mathbf{x}_{n_T}^T] \leftarrow [\mathbf{x}_1^S, \dots, \mathbf{x}_{n_S}^S]$ : source data **similar**
3. new learning set:  $[\mathbf{x}_1, \dots, \dots, \mathbf{x}_n]$ ,  $n = n_S + n_T$

## Model

- Gaussian mixture  $[\mathbf{x}_1, \dots, \dots, \mathbf{x}_n]$  in 4 classes  $\mathcal{C}_{S_1}, \mathcal{C}_{T_1}, \mathcal{C}_{S_2}, \mathcal{C}_{T_2}$

$$\mathbf{x}_i \in \mathcal{C}_a \longleftrightarrow \mathbf{x}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_a), \mathbf{C}_a \in \mathbb{R}^{p \times p}$$

for  $a \in \{S_1, T_1, T_2, S_2\}$ ,  $n_a = |\mathcal{C}_a|$ ,  $c_a = n_a/n$ .

- High dimension hypotheses:  $n, p \rightarrow \infty$ ,  $p/n < \infty$
- Non trivial classification:  $\|\mathbf{C}_a\| = O(1)$  and  $\text{tr}(\mathbf{C}_b - \mathbf{C}_a) = O(\sqrt{p})$ .

## Large Dimensional Asymptotic

**Theorem 1** (Gaussian Approximation). Let  $\mathbf{x} \in \mathcal{C}_a$ ,  $a \in \{T_1, T_2\}$  and  $g(\mathbf{x})$  LS-SVM decision function. Then,

$$n V_a^{-\frac{1}{2}} (g(\mathbf{x} | \mathbf{x} \in \mathcal{C}_a) - E_a) \xrightarrow{d} \mathcal{N}(0, 1), \quad (1)$$

$$\text{with } E_a = \gamma \ell^T \mathbf{P}_c \mathbf{m}_a + *, \quad (2)$$

$$\mathbf{m}_a = \frac{f''(\tau)}{p} \mathbf{t} \mathbf{t}_a + 2 \frac{f''(\tau)}{p^2} \mathbf{t} \mathbf{c}_a,$$

and

$$V_a = 2\gamma^2 \ell^T (\mathbf{P}_c \mathbf{W}_1^a \mathbf{P}_c + \mathbf{J}^T \mathbf{P} \mathbf{W}_2^a \mathbf{P} \mathbf{J}) \ell, \quad (3)$$

$$\text{where } \mathbf{W}_1^a = p^{-3} (f''(\tau))^2 \mathbf{t} \mathbf{t}^T \text{tr} \mathbf{C}_a^2,$$

$$\mathbf{W}_2^a = 2 p^{-2} n^{-2} (f''(\tau))^2 \text{diag}(\mathbf{J} \mathbf{t} \mathbf{c}_a),$$

$$\mathbf{t} = [t_1, \dots, t_k]^T, \mathbf{t}_{c_a} \triangleq [\text{tr}(\mathbf{C}_a \mathbf{C}_1), \dots, \text{tr}(\mathbf{C}_a \mathbf{C}_k)]^T \mathbf{C}^\circ \triangleq \sum_{a=1}^k \frac{n_a}{n} \mathbf{C}_a$$

$$\text{and } \tau \triangleq \frac{2}{p} \text{tr} \mathbf{C}^\circ > 0, t_a \triangleq p^{-1/2} \text{tr}(\mathbf{C}_a - \mathbf{C}^\circ),$$

$$\mathbf{P} \triangleq \mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T, \mathbf{P}_c \triangleq \frac{1}{n} \mathbf{J}^T \mathbf{P} \mathbf{J} = \text{diag}(\mathbf{c}) - \mathbf{c} \mathbf{c}^T.$$

## Labels optimisation

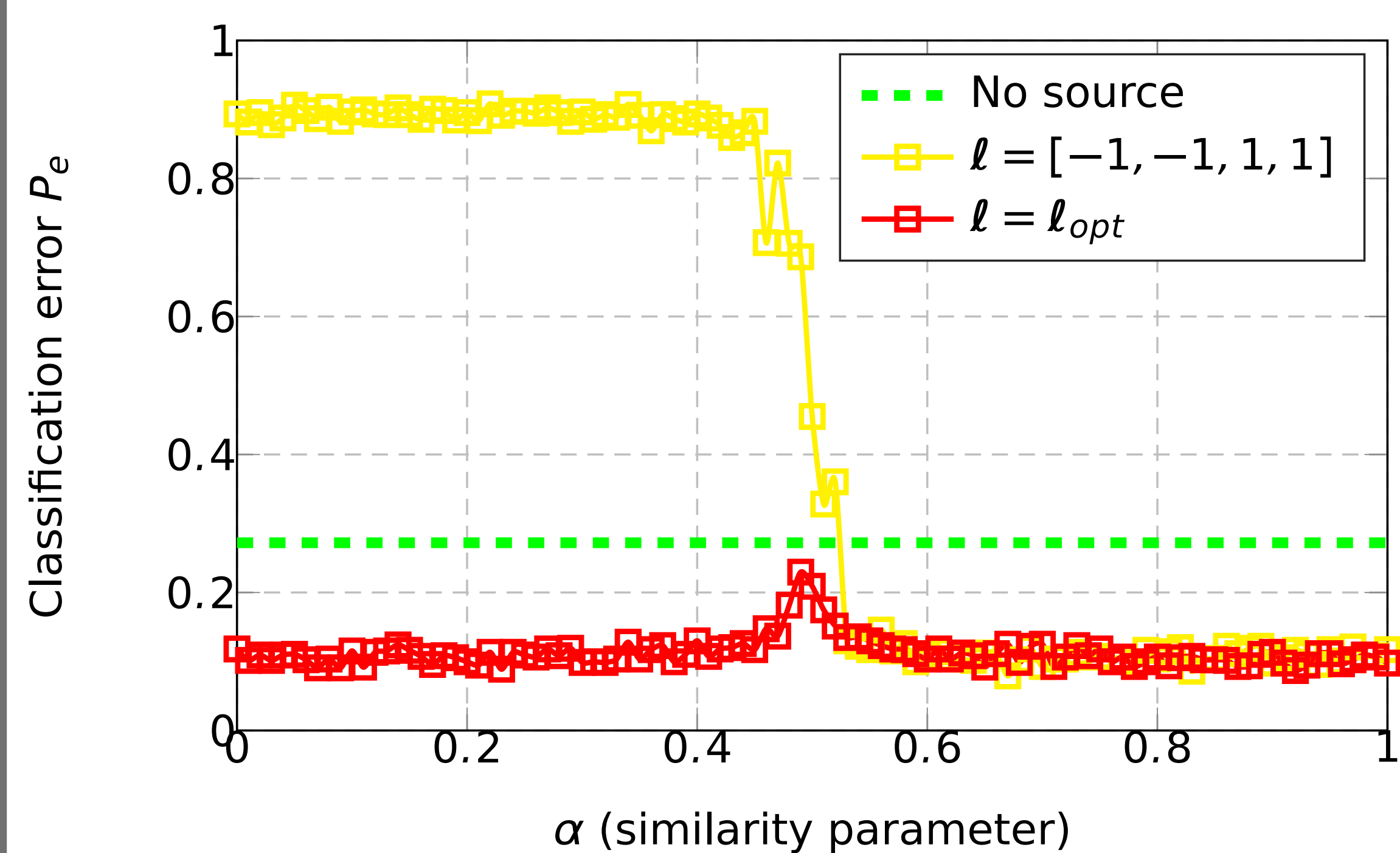


Figure 1: Classification performance on simulated data, related by parameter  $\alpha$ , s.t.  $\mathbf{C}_{S_a} = \alpha \mathbf{C}_{T_a} + (1 - \alpha) \mathbf{C}_{T_b}$ , for various label strategies;  $p = 512$ ,  $n_{S_1} = n_{S_2} = 508$ ,  $n_{T_1} = n_{T_2} = 4$ , polynomial kernel  $f$ .

Avoid negative transfer. Optimal label benefit from sources.

## Importance of source similarity

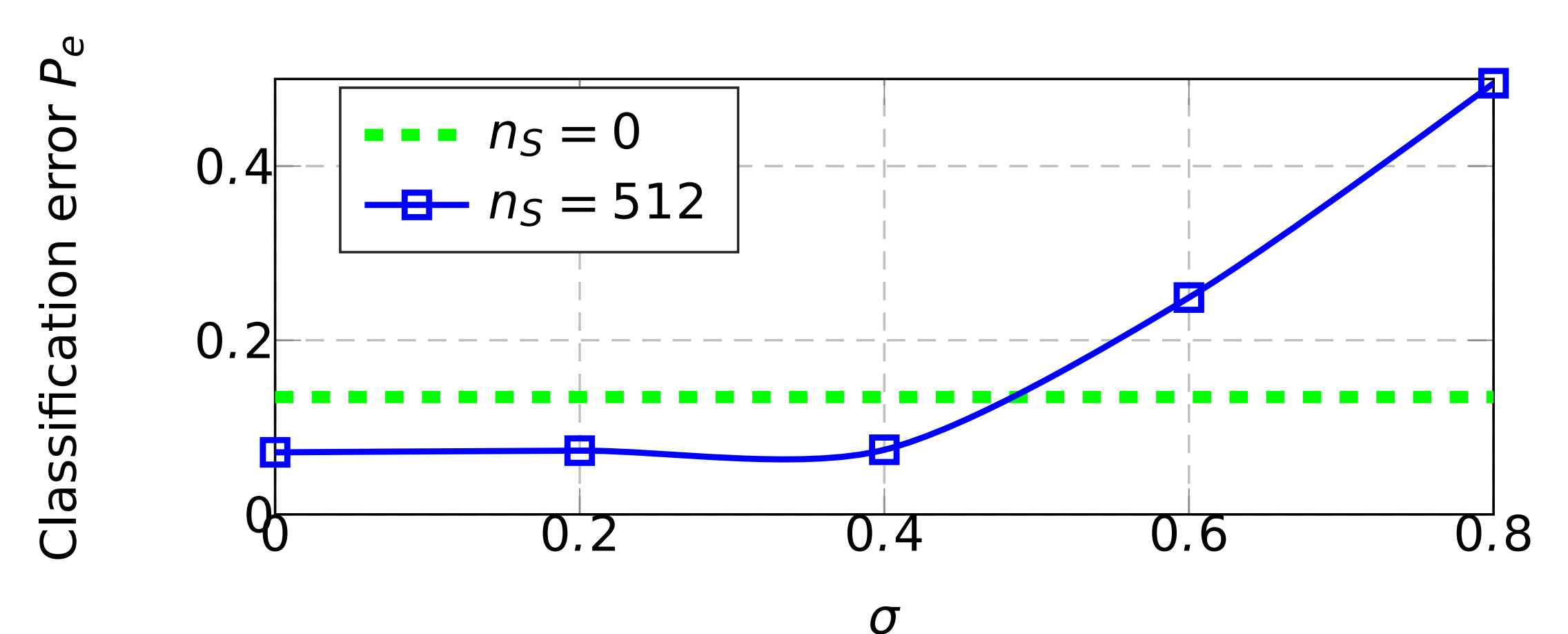


Figure 2: Classification between two homogeneous target zones in POLSAR data (Brétigny) with  $\mathcal{N}(0, \sigma^2)$ -noisy sources;  $n_T = 32$ ,  $p = 216$ , polynomial kernel  $f$ .

## Analysis and Improvement tracks

- LS-SVM Gaussian Approximation
  - ➔ Error probability, classification threshold
- Adaptability
  - Data centered  $\implies f'(\tau) = 0$
  - ➔ Error Optimisation (labels)
    - Adaptation criteria on source data
      - \* Noise
      - \* Diversity
      - \* Similarity

## References

1. S. J. Pan et al., *IEEE Transactions on Knowledge and Data Engineering* **22**, 1345–1359 (2010).
2. R. Couillet et al., *Electronic Journal of Statistics*, 1393–1454 (2016).
3. Z. Liao et al., *IEEE Transactions on Signal Processing* **67**, 1065–1074 (2019).
4. M. Tiomoko et al., *IEEE ICASSP'20* (2020).