Complexity of the Prototype Maps Algorithm

The number of operations of the Prototype Maps Algorithm is divided as follows:

1. Step 1 is the sample of $(\tilde{x}^k)_{k\in\{1,\dots,n_{\text{maps}}\}}$ i.i.d. of density function g: this step is negligible

2. Step 2 is the computation of $(\hat{y}(\tilde{x}^k))_{1\leq k\leq n_{\text{maps}}}$, done with the following method:

   - Perform FPCA to work in a basis of n_{pc} maps with the 3 following steps:
     - The decomposition of the training maps $(y_{x^i}(z))_{i=1,\dots,n_{\text{train}}}$ in the Φ basis basis is $\mathcal{O}(s^2)$ where s^2 is the number of pixels in the maps
     - The selection of the most important vectors of the basis is $\mathcal{O}(n_{\text{train}}\tilde{K})$
     - The PCA step is $\mathcal{O}(min(n_{\text{train}}^3, \tilde{K}^3))$
   - Gaussian process regression on this basis: the prediction of the coordinates at n_{maps} new points is $\mathcal{O}(n_{\text{train}}^3 + nn_{\text{train}}^2 n_{\text{pc}}))$, n_{train}^3 referring to the inversion of $k(\mathbb{X}_{\text{train}}, \mathbb{X}_{\text{train}})$
   - Inverse FPCA to get n_{maps} predicted maps:
     - The computation of the coefficients $\tilde{\alpha}$ in the wavelets basis for the n_{maps} new inputs is $\mathcal{O}(n_{\text{maps}}n_{\text{pc}}\tilde{K} + (K-\tilde{K})n_{\text{train}})$
     - The inverse wavelets transform is $\mathcal{O}(s^2)$

   Therefore complexity of step 2 is $\mathcal{O}(n_{\text{maps}}n_{\text{train}}^2 n_{\text{pc}})$

3. Step 3 is the computation of $(\frac{f_X(\tilde{x}^k)}{g(\tilde{x}_k)})_{1\leq k\leq n_{\text{maps}}}$: $\mathcal{O}(n_{\text{maps}})$

4. Step 4 is the initialization of $\Gamma_\ell^{[0]}$ which is negligible

5. Step 5 refers to the iterations of the Lloyd's algorithm and for each iteration k. Computing $\gamma_j^{[k+1]} = \hat{E}_n(\Gamma_\ell^{[k]}, j, \hat{y}), j = 1, \dots, \ell$, requires to attribute every predicted map to a Voronoi cell, which is $\mathcal{O}(\ell s^2 n_{\text{maps}})$ and then compute the estimators which is $\mathcal{O}(n_{\text{maps}})$

6. Step 8 is the re-estimation $\hat{P}_{\tilde{n}}(\hat{\Gamma}_\ell^\star, j, \hat{y})$ at the end of the Lloyd's algorithm, complexity is $\mathcal{O}(\tilde{n}n_{\text{train}}^2 n_{\text{pc}})$ for the computation of $(\hat{y}(\tilde{x}^k))_{1\leq k\leq \tilde{n}}$ and $\mathcal{O}(\ell s^2\tilde{n})$ for the computation of the estimator.

Finally, the overall complexity is $\mathcal{O}((\tilde{n}+n_{\text{maps}})n_{\text{train}}^2 n_{\text{pc}} + \ell s^2(n_{\text{maps}}n_{\text{it}}+\tilde{n}))$, with $\ell = 5$ the number of prototype maps, $s^2 = 64^2$ the number of pixels of a map, $n_{\text{pc}} = 2$ (or 6 for Campbell) the number of maps in the FPCA basis, n_{it} around 50 the number of iterations of the algorithm, $n_{\text{train}} = 1300$ the number of maps in the training database, $n_{\text{maps}} = 10^6$ the number of predicted maps computed in the algorithm, and $\tilde{n} = 10^7$ the number of maps sampled at the end of the algorithm the re-compute the probabilities.