Supplement to the article "Binary Classification as a Phase Separation Process"

Rafael Monteiro

MONTEIRODASILVA-RAFAEL@AIST.GO.JP,
RAFAEL.A.MONTEIRO.MATH@GMAIL.COM

Mathematics for Advanced Materials Open Innovation Laboratory,
AIST, c/o Advanced Institute for Materials Research,
Tohoku University, Sendai, Japan

Abstract

Supplementary material to the article "Binary Classification as a Phase Separation Process", by Rafael Monteiro. We present the construction of the canonical basis matrix in the general case. Several plots with statistical data of the PSBC extend Figures 10 and 11 shown in the paper to parameters not covered there.

Contents

All the figures re constructed using the files in the Statistics folder; see "(Monteiro, 2021, PSBC_-using_statistical_files.ipynb)" for the code used to generate them. For further information on how to access the data in the Statistics folder see the "(Monteiro, 2021, README.pdf)".

1. Canonical basis matrix constructions, general case

We now generalize the canonical basis matrix construction presented in the paper, Remark 7. As before, let $1 \leq \mathrm{N_{pt}} \leq \mathrm{N_u}$. We recall that, therein, we considered $\mathrm{N_u} = Q\mathrm{N_{pt}}$, for $Q \in \mathbb{N}$. Here, we shall consider $\mathrm{N_u} = Q\mathrm{N_{pt}} + R$, for $Q \in \mathbb{N}$ and $0 \leq R \leq Q - 1$; both Q and R are unique, due to Euclid's algorithm.

Initially, set $\pi_0 := \emptyset$. For $1 \leq j \leq R$, construct sets π_j recursively with the $Q+1$ smallest elements in the set $\mathbb{G}_{\mathrm{N_u}} \setminus \left(\cup_{l=0}^{j-1}\pi_l\right)$. For the remaining $\mathrm{N_{pt}} - R$ sets, construct sets π_j recursively with the Q smallest elements in the set $\mathbb{G}_{\mathrm{N_u}} \setminus \left(\cup_{l=0}^{j-1}\pi_l\right)$.

In the end, define $\mathscr{B} := \left[e_{\pi_1}|\ldots|e_{\pi_{\mathrm{N_{pt}}}}\right]$ (recall notation in Section 1.3 of the paper).

2. Accuracies

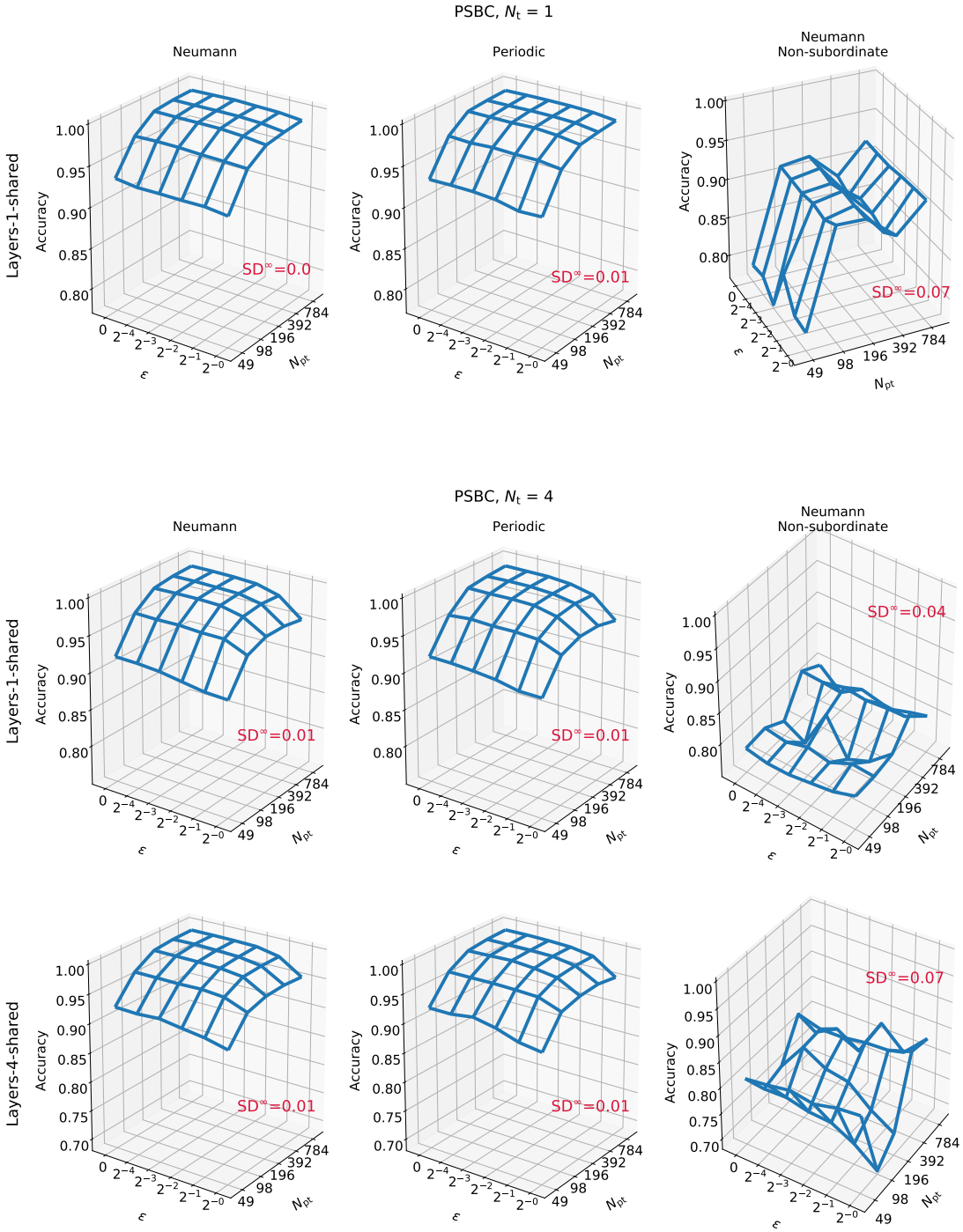

Figure 1: Accuracy for different values of viscosities ε and parameterization cardinalities N_{pt}. The classifier is applied to the "0" and "1" subset of the MNIST database. Each point on these surfaces is computed as an average of 5 experiments, with an associated standard deviation; SD^∞ indicates the largest among them. For convenience the positive part of the ε axis is in log-scale, whereas the N_{pt} axis is off scale, represented as an equally spaced grid. Values were rounded to 2 digits.

3. Best weights at their optimal epoch value

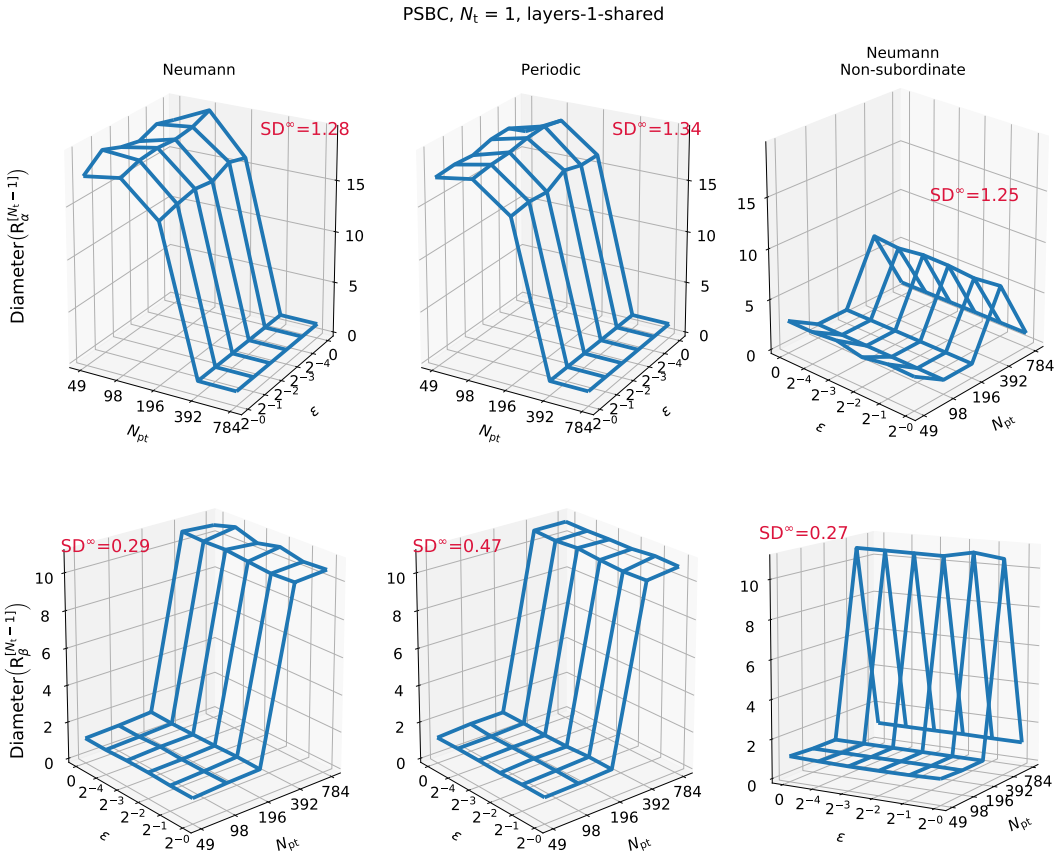

Figure 2: Average values attained by the diameter of the set $\mathcal{R}_\alpha^{\lfloor N_t-1 \rfloor} := \mathrm{conv}\left(\{0,1\} \cup_{m=0}^{N_t-1} \{\alpha^{\lfloor m \rfloor}\}\right)$ at the best epoch (according to an Early Stopping criterion); $\mathcal{R}_\beta^{\lfloor N_t-1 \rfloor}$ is defined similarly. Different values of viscosities ε and parameterization cardinalities N_{pt} have been tested. The classifier is applied to the "0" and "1" subset of the MNIST database. Each point on these surfaces is computed as an average of 5 experiments, with an associated standard deviation; SD^∞ indicates the largest among them. For the time evolution of these diameters through epochs, see the figures in the Supplementary Material and also the the accompanying movie file (Monteiro, 2021, Example_layers_snapshots.mp4, about 0.5 Mb). For convenience the positive part of the ε axis is in log-scale, whereas the N_{pt} axis is off scale, represented as an equally spaced grid. Values were rounded to 2 digits.

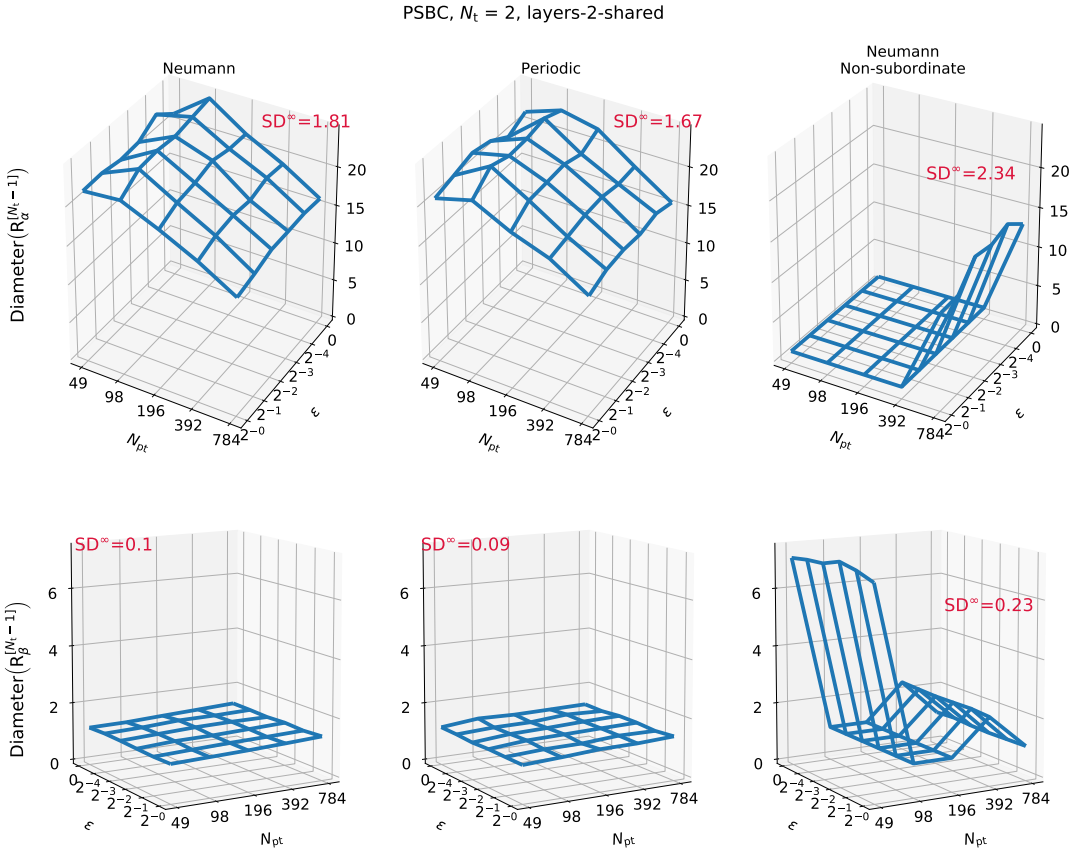

Figure 3: Average values attained by the diameter of the set $\mathcal{R}_\alpha^{\lfloor N_t-1 \rfloor} := \mathrm{conv}\left(\{0,1\} \cup_{m=0}^{N_t-1} \{\alpha^{\lfloor m \rfloor}\}\right)$ at the best epoch (according to an Early Stopping criterion); $\mathcal{R}_\beta^{\lfloor N_t-1 \rfloor}$ is defined similarly. Different values of viscosities ε and parameterization cardinalities N_{pt} have been tested. The classifier is applied to the "0" and "1" subset of the MNIST database. Each point on these surfaces is computed as an average of 5 experiments, with an associated standard deviation; SD^∞ indicates the largest among them. For the time evolution of these diameters through epochs, see the figures in the Supplementary Material and also the the accompanying movie file (Monteiro, 2021, Example_layers_snapshots.mp4, about 0.5 Mb). For convenience the positive part of the ε axis is in log-scale, whereas the N_{pt} axis is off scale, represented as an equally spaced grid. Values were rounded to 2 digits.

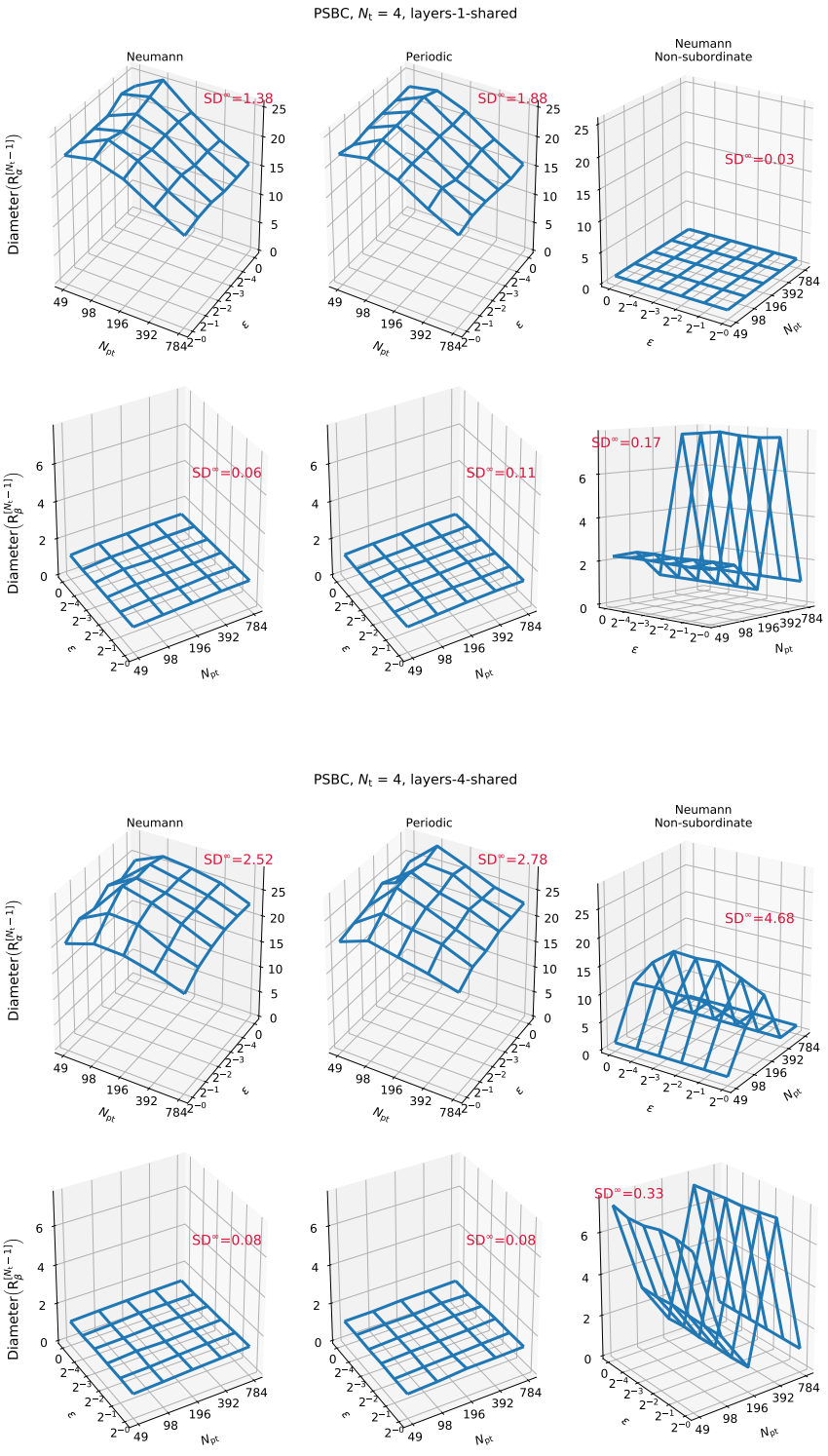

Figure 4: Average values attained by the diameter of the set $\mathcal{R}_\alpha^{\lfloor N_t-1 \rfloor} := \text{conv}\left(\{0,1\} \cup_{m=0}^{N_t-1} \{\alpha^{\lfloor m \rfloor}\}\right)$ at the best epoch (according to an Early Stopping criterion); $\mathcal{R}_\beta^{\lfloor N_t-1 \rfloor}$ is defined similarly. Different values of viscosities ε and parameterization cardinalities N_{pt} have been tested. The classifier is applied to the "0" and "1" subset of the MNIST database. Each point on these surfaces is computed as an average of 5 experiments, with an associated standard deviation; SD^∞ indicates the largest among them. For the time evolution of these diameters through epochs, see the figures in the Supplementary Material and also the the accompanying movie file (Monteiro, 2021, Example_layers_snapshots.mp4, about 0.5 Mb). For convenience the positive part of the ε axis is in log-scale, whereas the N_{pt} axis is off scale, represented as an equally spaced grid. Values were rounded to 2 digits.

4. Growth of weights through epochs ($\mathrm{N_t} = 2$)

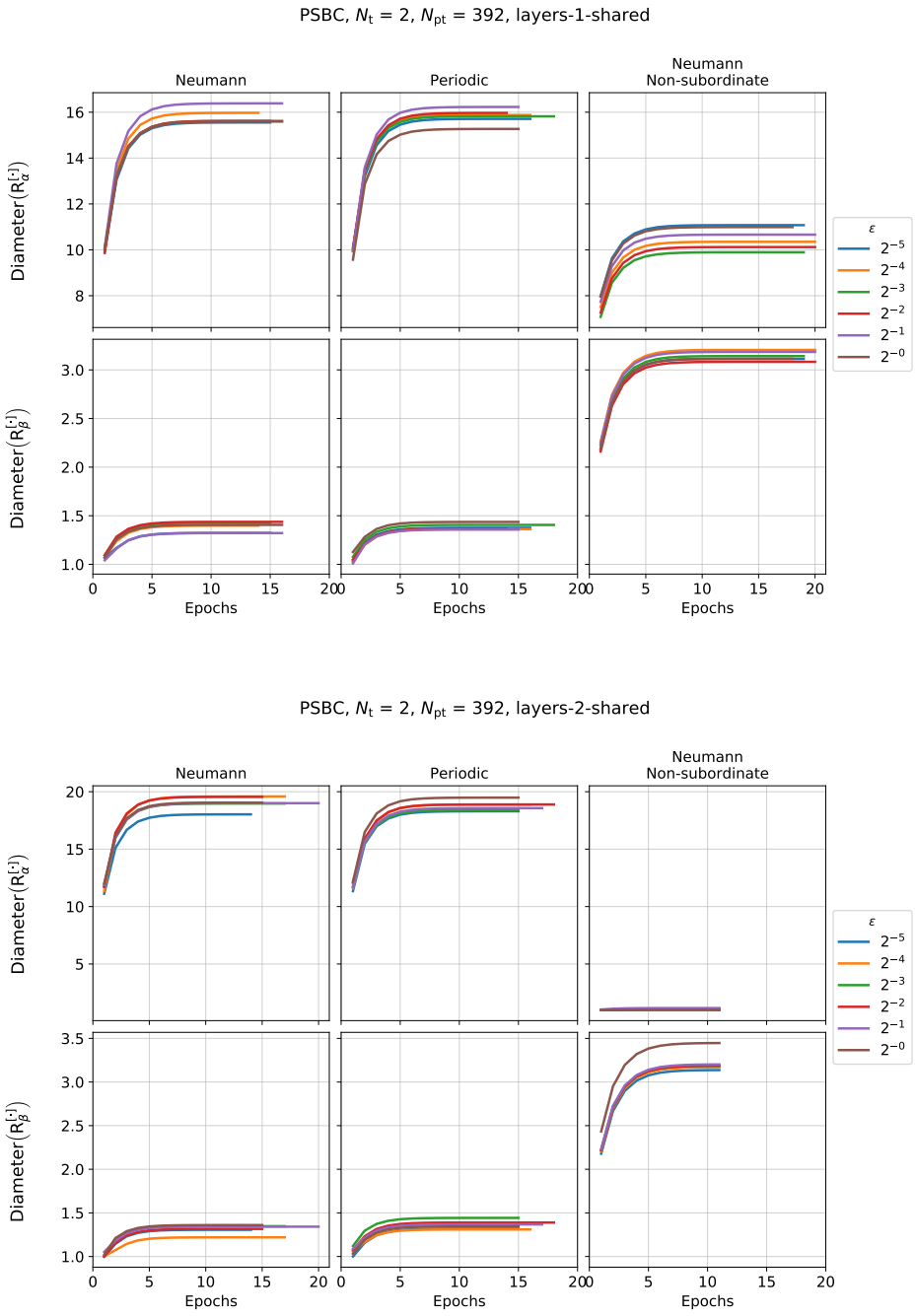

Figure 5: Evolution of the diameter of the set $\mathcal{R}_\alpha^{\lfloor \mathrm{N_t}-1 \rfloor} := \mathrm{conv}\left(\{0,1\} \cup_{m=0}^{\mathrm{N_t}-1} \{\alpha^{\lfloor m \rfloor}\}\right)$ during several realizations of the PSBC applied to digits "0" and "1" of the MNIST data set. These time series were collected at parameterization cardinality $\mathrm{N_{pt}} = 392$ but different values of the viscosities parameter ε. Notice that these time series have different lengths due to Early Stopping (details in Section A of the paper). See also the the accompanying movie file (Monteiro, 2021, Example_layers_snapshots.mp4, about 0.5 Mb).

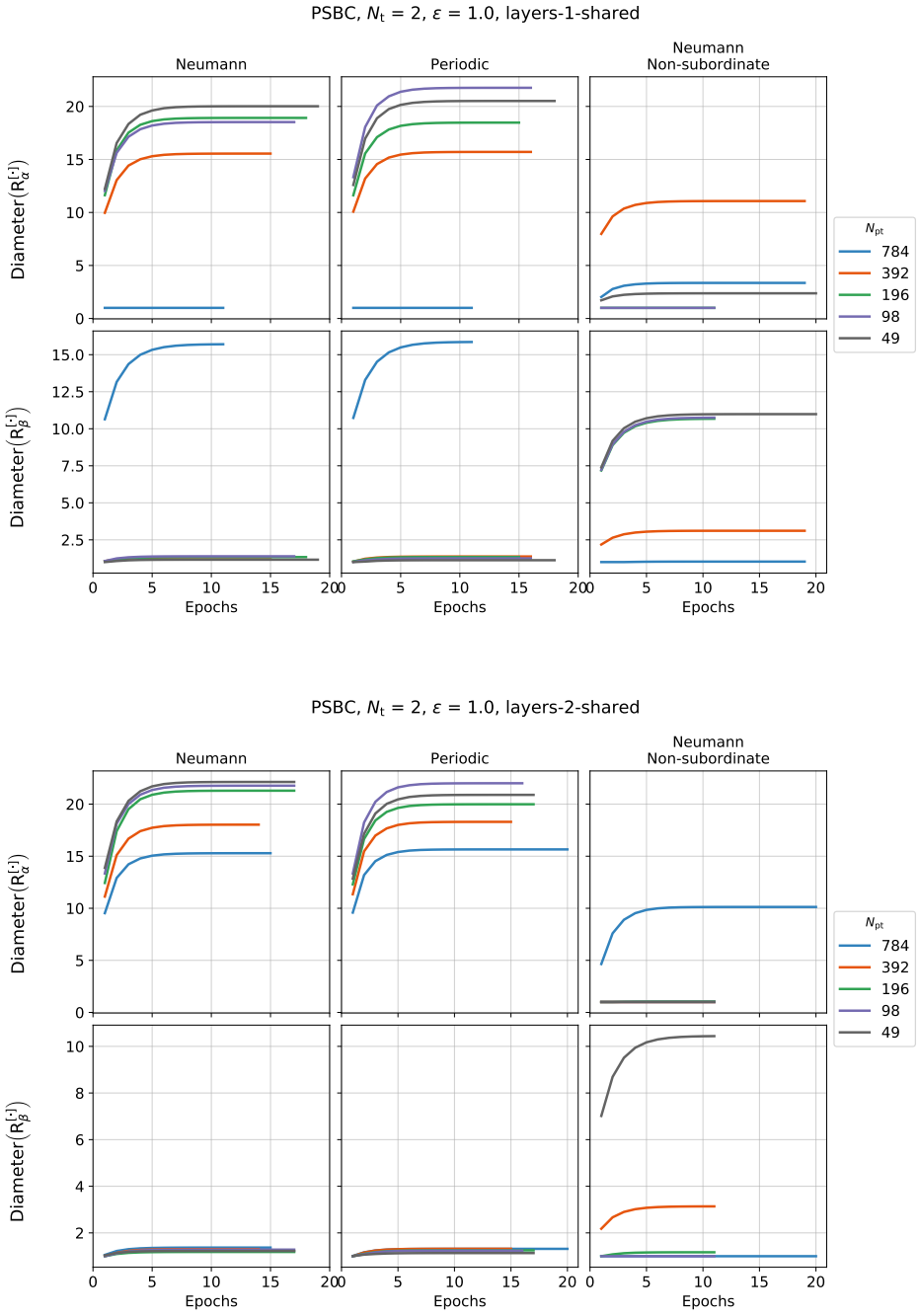

Figure 6: Evolution of the diameter of the set $\mathcal{R}_\alpha^{\lfloor N_t-1 \rfloor} := \text{conv}\left(\{0,1\} \cup_{m=0}^{N_t-1} \{\alpha^{\lfloor m \rfloor}\}\right)$ during several realizations of the PSBC applied to digits "0" and "1" of the MNIST data set. These time series were collected for the inviscid PSBC ($\varepsilon = 1.0$) at different values of the parameterization cardinalities N_{pt}. Notice that these time series have different lengths due to Early Stopping (details in Section A of the paper). See also the the accompanying movie file (Monteiro, 2021, Example_layers_snapshots.mp4, about 0.5 Mb).

References

Rafael Monteiro. Source code for the paper "Binary classification as a phase separation process". https://github.com/rafael-a-monteiro-math/Binary_classification_phase_separation, September 2021.