Homework 2
IST 597
Physics-Informed Machine Learning
Subarna Pudasaini (sfp5828@psu.edu)

Question 1

Implement Levy and Lindenbaum's streaming SVD algorithm and deploy it for the MNIST dataset where images are streamed in batches of 50. Levy and Lindenbaum use smaller QR and SVD decompositions - use modified Gram-Schmidt and Golub-Reinsch for these.

Ans:

Code: hw2.ipynb

Code from scratch: from_scratch.py

Question 2

Assess run-times for varying batch-size in a plot with x-axis as batch size and y-axis as average run-time for 10 trials. Assume a forget-factor of 1.0 and visualize the evolution of the first three singular vectors with each batch as a video (mp4/gif) and upload to Canvas.

Compare the singular vectors obtained by your implementation of Levy and Lindenbaum (after processing all batches) to np.linalg.svd qualitatively for different values of the forget factor.

Ans:

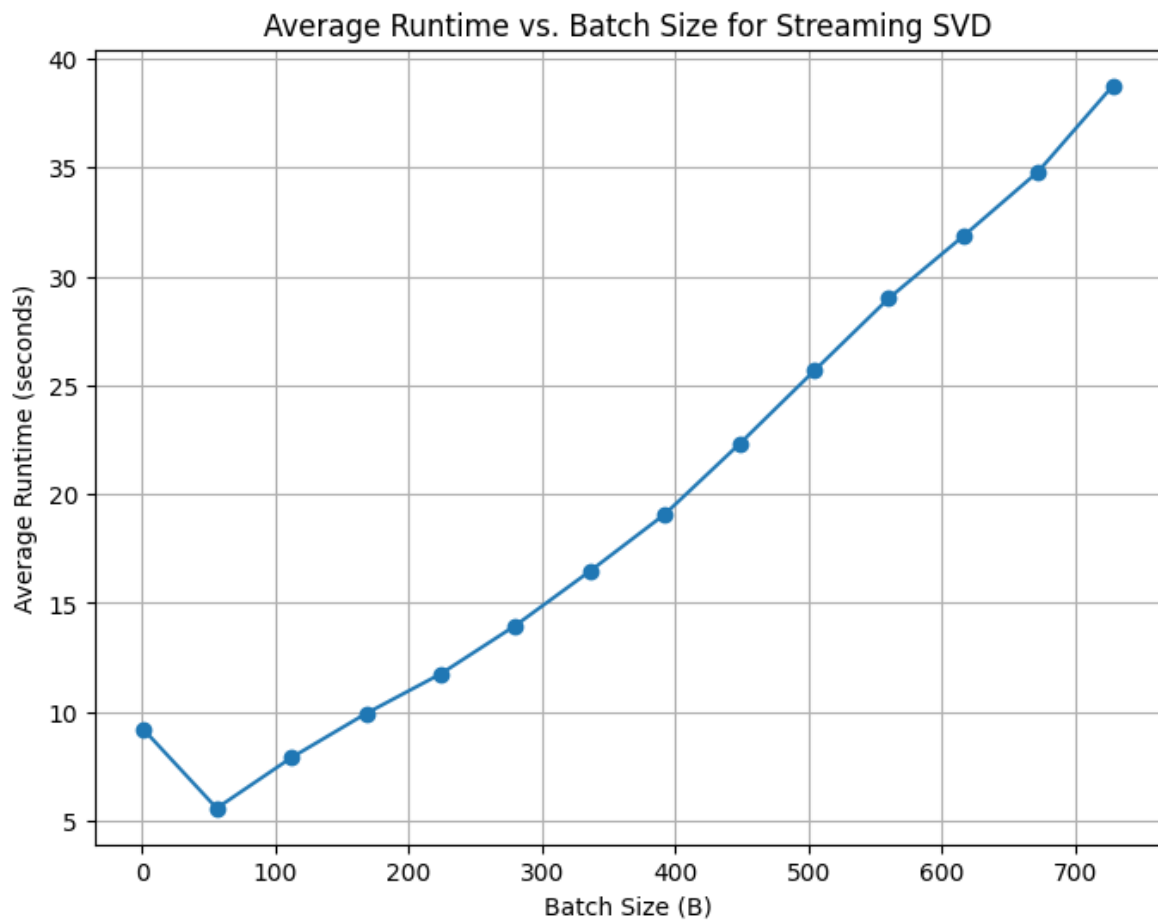*All the figures are also uploaded to the canvas.*

Runtime vs. Batch-size

Figure 1: Runtime vs. Batch-size

Comparison of singular vectors: Levy and Lindenbaum vs. np.linalg.svd
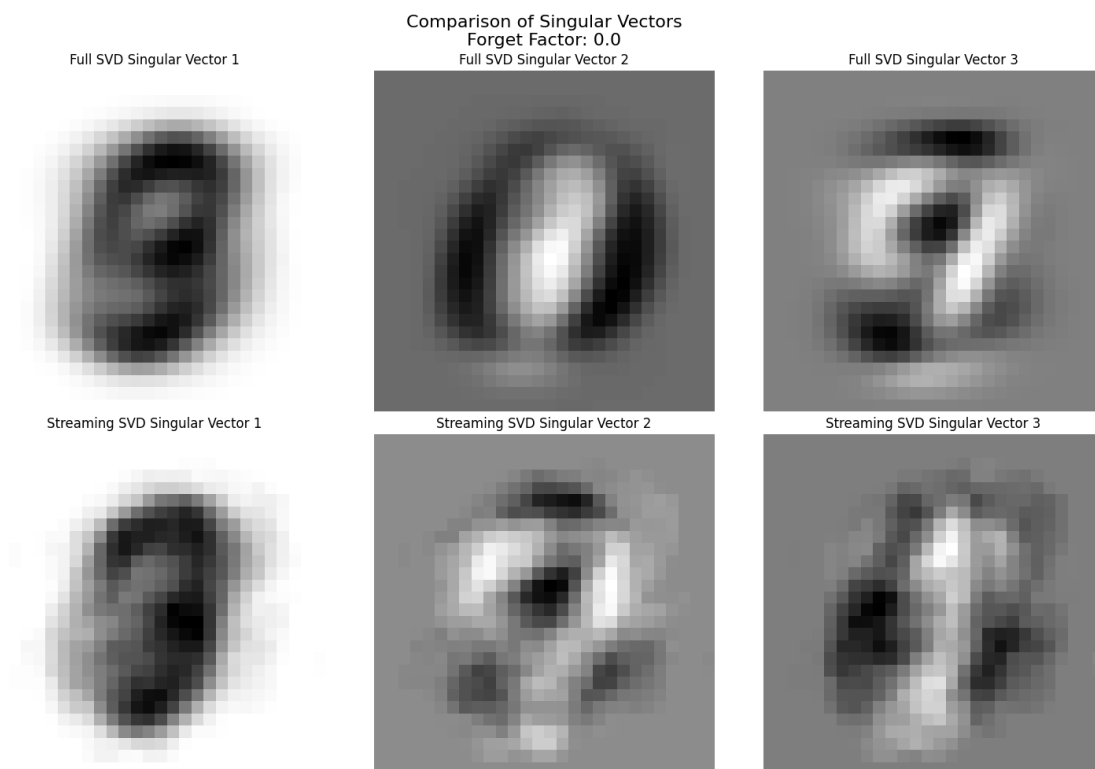

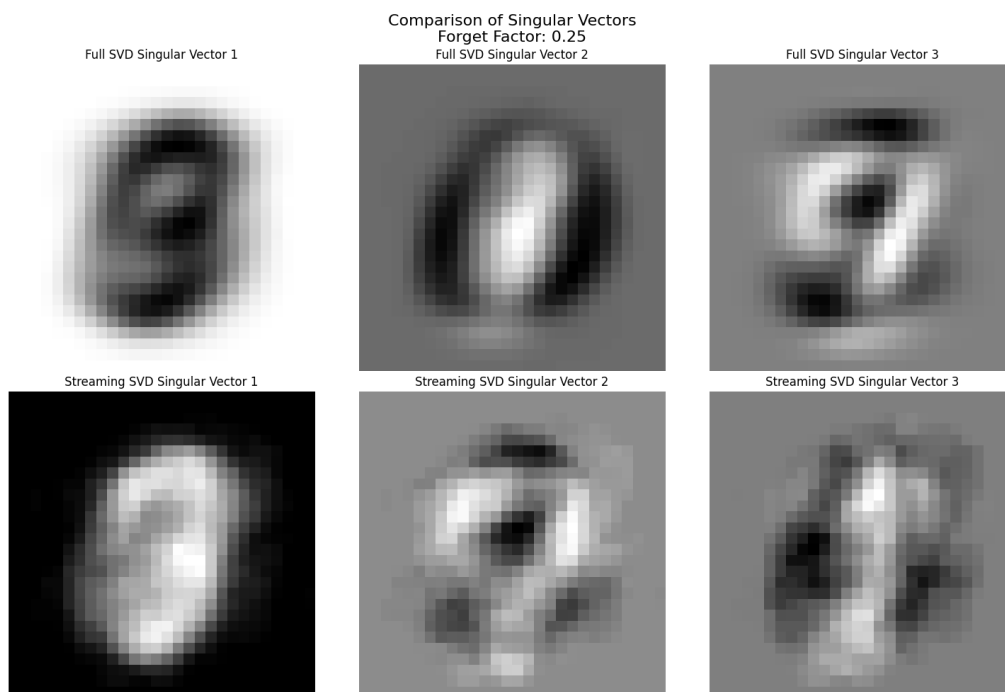Figure 2: SVD comparison; forget factor = 0

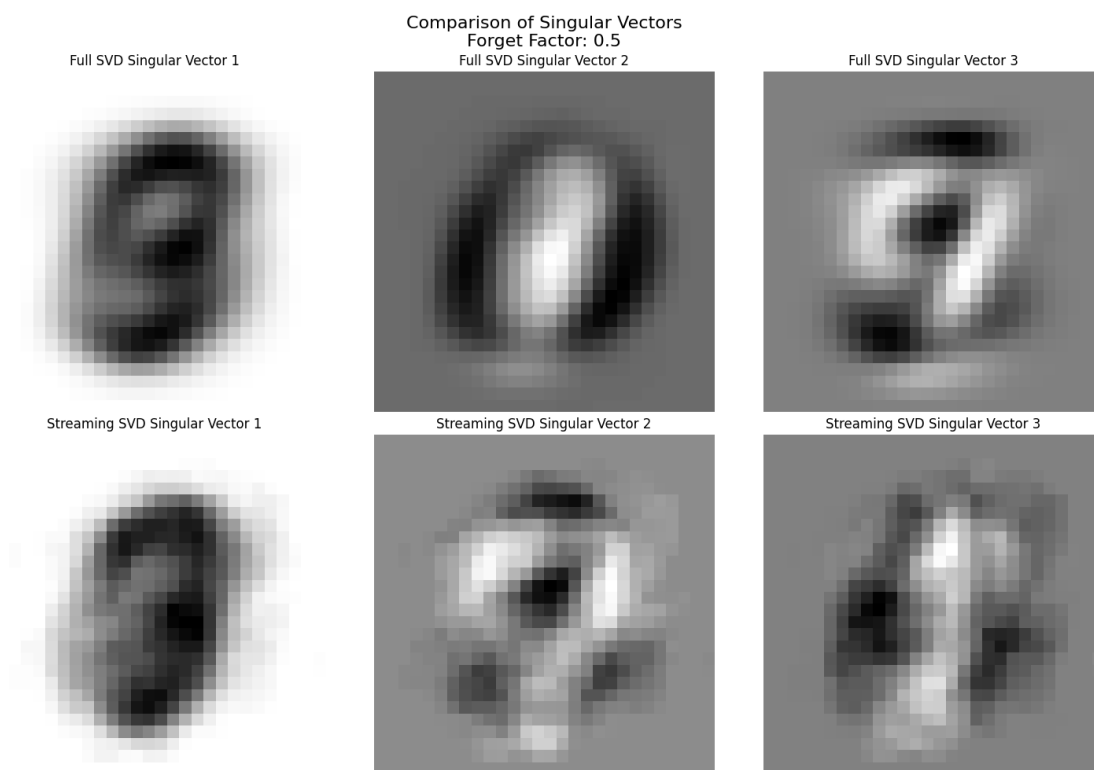Figure 3: SVD comparison; forget factor = 0.25

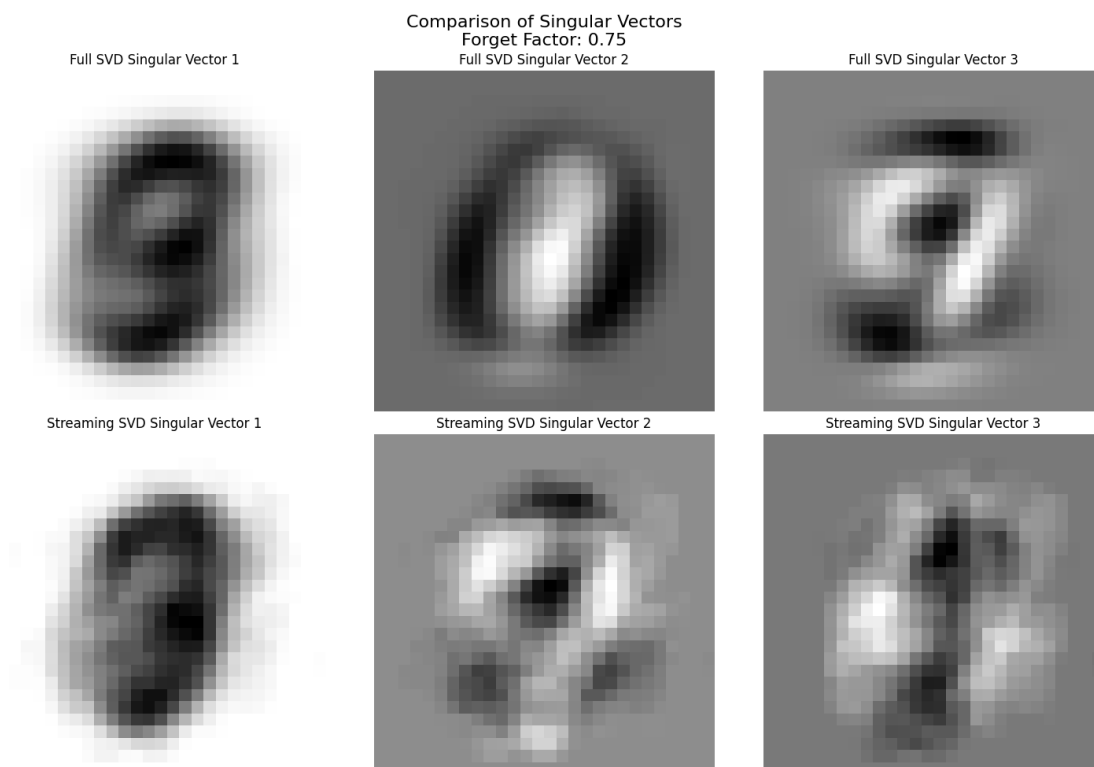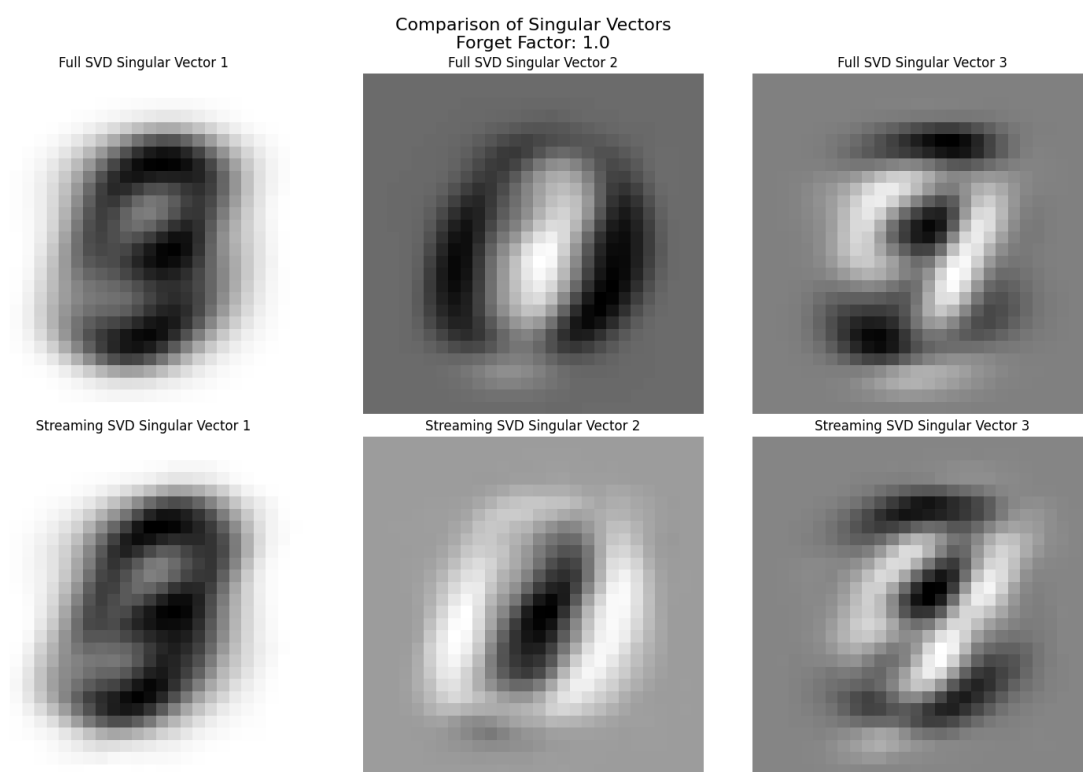Figure 4: SVD comparison; forget factor = 0.5

Figure 5: SVD comparison; forget factor = 0.75

Figure 6: SVD comparison; forget factor = 1.0