Appendix: Supplementary material

A. Proof of Theorem 1

(i) In $(\mathbb{E}^d, \overset{\phi_\alpha}{+}, \overset{\phi_\alpha}{\cdot})$ the determinant of matrix $\boldsymbol{R}$ is defined as $|\boldsymbol{R}|_\alpha = \phi_\alpha^{-1}|\Phi_\alpha(\boldsymbol{R})|$. Then, we deduce that the characteristic polynomial associated with matrix $\boldsymbol{R}$ is:

$$\phi_\alpha^{-1}|\Phi_\alpha[\boldsymbol{R} \overset{\Phi_\alpha}{-} \lambda I_d]| = 0$$

Then,

$$\begin{aligned}
&\phi_\alpha^{-1}|\Phi_\alpha[\boldsymbol{R} \overset{\Phi_\alpha}{-} \lambda \overset{\Phi_\alpha}{\cdot} I_d]| = 0\\
\iff &\phi_\alpha^{-1}|\Phi_\alpha \circ \Phi_\alpha^{-1}[\Phi_\alpha(\boldsymbol{R}) - \Phi_\alpha(\lambda \overset{\Phi_\alpha}{\cdot} I_d)]| = 0\\
\iff &\phi_\alpha^{-1}|\Phi_\alpha(\boldsymbol{R}) - \Phi_\alpha(\lambda \overset{\Phi_\alpha}{\cdot} I_d)]| = 0\\
\iff &\phi_\alpha^{-1}|\Phi_\alpha(\boldsymbol{R}) - \phi_\alpha(\lambda) I_d)]| = 0\\
\iff &|\Phi_\alpha(\boldsymbol{R}) - \phi_\alpha(\lambda) I_d)]| = \phi_\alpha(0) = 0
\end{aligned}$$

Making the following change of variable $\tilde{\lambda} = \phi_\alpha(\lambda)$, then:

$$|\Phi_\alpha(\boldsymbol{R}) - \tilde{\lambda} I_d)]| = 0$$

Consequently, solving for the eigenvalues of $\Phi_\alpha(\boldsymbol{R})$ in $\mathbb{R}^d$ yields the eigenvalues of $\boldsymbol{R}$ in $\mathbb{E}^d$, that are $\Lambda = \{\phi_\alpha^{-1}(\tilde{\lambda}_1), \cdots, \phi_\alpha^{-1}(\tilde{\lambda}_d)\}$.
(ii) Let $\Lambda_\alpha = \{\tilde{\lambda}_1, \cdots, \tilde{\lambda}_d\}$ be the set of eigenvalues of $\Phi_\alpha(\boldsymbol{R})$ in $\mathbb{R}^d$ and $\mathscr{B}_\alpha = \{\tilde{b}_1, \cdots, \tilde{b}_d\}$ the set of corresponding eigenvectors. From the eigenvalue equation in $\mathbb{R}^d$, we obtain:

$$\Phi_\alpha(\boldsymbol{R})\tilde{b}_j = \tilde{\lambda}_j \tilde{b}_j \iff \Phi_\alpha^{-1}[\Phi_\alpha(\boldsymbol{R})\tilde{b}_j] = \Phi_\alpha^{-1}(\tilde{\lambda}_j \tilde{b}_j)$$

Equivalently,

$$\boldsymbol{R} \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_j) = \phi_\alpha^{-1}(\tilde{\lambda}_j) \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_j)$$

Then, we obtain the eigenvalue equation in the vector space $(\mathbb{E}^d, \overset{\phi_\alpha}{+}, \overset{\phi_\alpha}{\cdot})$. Therefore, solving for the eigenvectors $\tilde{b}_j$ in $\mathbb{R}^d$ yields the eigenvectors $\Phi_\alpha^{-1}(\tilde{b}_j)$ in $\mathbb{E}^d$, and this ends the proof.

(iii) $\mathscr{B} = \Phi_\alpha^{-1}(\mathscr{B}_\alpha) = \{\Phi_\alpha^{-1}(\tilde{b}_1), \cdots, \Phi_\alpha^{-1}(\tilde{b}_d)\}$ is a basis of $(\mathbb{E}^d, \overset{\phi_\alpha}{+}, \overset{\phi_\alpha}{\cdot})$.

Let us prove first that $\{\Phi_\alpha^{-1}(\tilde{b}_1), \cdots, \Phi_\alpha^{-1}(\tilde{b}_d)\}$ is a linearly independent sequence of vectors. Let us take $(a_1, a_2, \cdots, a_d) \in \mathbb{E}^d$. Let us assume that:

$$a_1 \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_1) \overset{\Phi_\alpha}{+} \cdots \overset{\Phi_\alpha}{+} a_d \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_d) = \Phi_\alpha^{-1}(0)$$

Since $\alpha > 0$, $\Phi_\alpha^{-1}(0) = (0, \cdots, 0)$, and so:

$$\begin{cases} a_1 \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_1) = 0 \\ a_2 \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_2) = 0 \\ \cdots \\ a_d \overset{\Phi_\alpha}{\cdot} \Phi_\alpha^{-1}(\tilde{b}_d) = 0 \end{cases} \Longrightarrow \begin{cases} a_1 = 0 \\ a_2 = 0 \\ \cdots \\ a_d = 0 \end{cases}$$

Then, we deduce that $\{\Phi_\alpha^{-1}(\tilde{b}_1), \cdots, \Phi_\alpha^{-1}(\tilde{b}_d)\}$ is a linearly independent sequence of vectors of $\mathbb{E}^d$. In addition, $\{\Phi_\alpha^{-1}(\tilde{b}_1), \cdots, \Phi_\alpha^{-1}(\tilde{b}_d)\}$ is a spanning sequence of vectors of $\mathbb{E}^d$, therefore $\{\Phi_\alpha^{-1}(\tilde{b}_1), \cdots, \Phi_\alpha^{-1}(\tilde{b}_d)\}$ is a basis of $(\mathbb{E}^d, \overset{\phi_\alpha}{+}, \overset{\phi_\alpha}{\cdot})$.

(iv.a) $\tilde{b}_k \perp \tilde{b}_j \Longleftrightarrow \Phi_\alpha^{-1}(\tilde{b}_k) \overset{\phi_\alpha}{\perp} \Phi_\alpha^{-1}(\tilde{b}_j)$, for $\alpha > 0$.
[⇒] Let us assume that $\tilde{b}_k \perp \tilde{b}_j$. By definition **1**, the ϕ_α-inner product yields,

$$\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j) \rangle_\alpha = \phi_\alpha^{-1} \circ \langle \tilde{b}_k, \tilde{b}_j \rangle$$

Since $\tilde{b}_k \perp \tilde{b}_j$, then

$$\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j) \rangle_\alpha = \phi_\alpha^{-1}(0)$$

For all $\alpha > 0$, we have $\phi_\alpha^{-1}(0) = 0$, thus:

$$\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j) \rangle_\alpha = \phi_\alpha^{-1} \circ \langle \tilde{b}_k, \tilde{b}_j \rangle = 0$$

Finally, by Definition **1**, $\Phi_\alpha^{-1}(\tilde{b}_k) \overset{\phi_\alpha}{\perp} \Phi_\alpha^{-1}(\tilde{b}_j)$, that is, $b_k \overset{\phi_\alpha}{\perp} b_j$.
[⇐] Suppose $b_k \overset{\phi_\alpha}{\perp} b_j$, then $\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j) \rangle_\alpha = 0$. Therefore,

$$\begin{aligned} & \langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j) \rangle_\alpha = 0 \\ \Longleftrightarrow & \overset{\Phi_\alpha}{\sum_{l \in [\![d]\!]}} \phi_\alpha^{-1}(\tilde{b}_{k,l}) \overset{\phi_\alpha}{\cdot} \phi_\alpha^{-1}(\tilde{b}_{j,l}) = 0 \\ \Longleftrightarrow & \phi_\alpha^{-1} \circ \sum_{l \in [\![d]\!]} \tilde{b}_{k,l} \cdot \tilde{b}_{j,l} = 0 \\ \Longleftrightarrow & \sum_{l \in [\![d]\!]} \tilde{b}_{k,l} \cdot \tilde{b}_{j,l} = \phi_\alpha(0) \end{aligned}$$

If $\alpha > 0$, then $\phi_\alpha(0) = 0$, that is, $\tilde{b}_k \perp \tilde{b}_j$, and this ends the proof.

(iv.b) $\Phi_\alpha^{-1}(\tilde{b}_k) \overset{\phi_\alpha}{\perp} \Phi_\alpha^{-1}(\tilde{b}_j) \Longleftrightarrow \phi_\alpha^{-1}(\tilde{b}_k' \tilde{b}_j) = 0$, for $\alpha > 0$.

[⇒] Suppose $\Phi_\alpha^{-1}(\tilde{b}_k) \overset{\phi_\alpha}{\perp} \Phi_\alpha^{-1}(\tilde{b}_j)$, then $\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j)\rangle_\alpha = 0$. Therefore,

$$\begin{aligned} &\langle \Phi_\alpha^{-1}(\tilde{b}_k), \Phi_\alpha^{-1}(\tilde{b}_j)\rangle_\alpha = 0 \\ \iff &\sum_{l\in[\![d]\!]}^{\Phi_\alpha} \phi_\alpha^{-1}(\tilde{b}_{k,l}) \overset{\phi_\alpha}{\cdot} \phi_\alpha^{-1}(\tilde{b}_{j,l}) = 0 \\ \iff &\phi_\alpha^{-1} \circ \sum_{l\in[\![d]\!]} \tilde{b}_{k,l} \cdot \tilde{b}_{j,l} = 0 \end{aligned}$$

[⇐] Straightforward.

(v) $\|\tilde{b}_j\|_2 = 1 \iff \|\Phi_\alpha^{-1}(\tilde{b}_j)\|_2^\alpha = 1$
[⇒] We have:

$$\phi_\alpha^{-1}(\|\tilde{b}_j\|_2) = \phi_\alpha^{-1}(1)$$

By definition of ϕ_α^{-1}, $\phi_\alpha^{-1}(1) = 1^{\frac{1}{\alpha}} = 1$, then,

$$\phi_\alpha^{-1} \circ \|\tilde{b}_j\|_2 = 1$$

On the other hand, the Φ_α-Euclidean norm is given by:

$$\begin{aligned} \|\Phi_\alpha^{-1}(x)\|_2^\alpha &= \sqrt{\langle \Phi_\alpha^{-1}(x), \Phi_\alpha^{-1}(x)\rangle_\alpha} \\ &= \sqrt{\phi_\alpha^{-1} \circ \langle x, y\rangle} \\ &= \phi_\alpha^{-1} \circ \|x\|_2 \end{aligned} \tag{3}$$

So, $\|\Phi_\alpha^{-1}(\tilde{b}_j)\|_2^\alpha = 1$.

[⇐] Let us assume that $\|\Phi_\alpha^{-1}(\tilde{b}_j)\|_2^\alpha = 1$. By, Eq.(3), $\|\Phi_\alpha^{-1}(x)\|_2^\alpha = \phi_\alpha^{-1} \circ \|x\|_2$, then:

$$\|\Phi_\alpha^{-1}(\tilde{b}_j)\|_2^\alpha = \phi_\alpha^{-1} \circ \|\tilde{b}_j\|_2 = 1$$

Since $\phi_\alpha(1) = 1$, we finally get that $\|b_j\|_2 = 1$.

(vi) Let $\tilde{\boldsymbol{B}} = [\tilde{b}_1, \cdots, \tilde{b}_d]$ be the matrix of the eigenvectors issued from the standard PCA based on $\Phi_\alpha(\boldsymbol{R})$. Then, the principal components issued from the standard PCA in $\mathbb{R}^d$ is:

$$\tilde{\boldsymbol{F}} = \Phi_\alpha(\boldsymbol{Z})\tilde{\boldsymbol{B}} = \Phi_\alpha(\boldsymbol{Z})\Phi_\alpha(\boldsymbol{B})$$

On the other hand, in $\mathbb{E}^d$, the principal components are given by:

$$\boldsymbol{F} = \boldsymbol{Z} \overset{\Phi_\alpha}{\cdot} \boldsymbol{B} = \Phi_\alpha^{-1}[\Phi_\alpha(\boldsymbol{Z})\Phi_\alpha(\boldsymbol{B})]) = \Phi_\alpha^{-1}[\tilde{\boldsymbol{F}}]$$

B. Proof of Proposition 1

(i) Let us begin with the inner product of x and y over $\mathbb{E}^d$ denoted by $\langle x, y\rangle_\alpha$:

$$\begin{aligned}
\langle x, y\rangle_\alpha &= \overset{\phi_\alpha}{\sum_{k\in[\![d]\!]}} x_k \overset{\phi_\alpha}{\cdot} y_k \\
&= \overset{\phi_\alpha}{\sum_{k\in[\![d]\!]}} \phi_\alpha^{-1}(\phi_\alpha(x_k)\cdot\phi_\alpha(y_k)) \\
&= \phi_\alpha^{-1} \circ \sum_{k\in[\![d]\!]} \phi_\alpha(x_k)\cdot\phi_\alpha(y_k) \\
&= \phi_\alpha^{-1} \circ \langle \Phi_\alpha(x), \Phi_\alpha(y)\rangle
\end{aligned}$$

Also,

$$\begin{aligned}
\cos_\alpha^2(x, y) &= \left[\frac{\langle x, y\rangle_\alpha}{\|x\|_2^\alpha \overset{\phi_\alpha}{\cdot} \|y\|_2^\alpha}\right]^2 \\
&= \left[\frac{\phi_\alpha^{-1} \circ \langle \Phi_\alpha(x), \Phi_\alpha(y)\rangle}{\phi_\alpha^{-1}(\|\Phi_\alpha(x)\|_2 \cdot \|\Phi_\alpha(y)\|_2)}\right]^2 \\
&= [\phi_\alpha^{-1} \circ \cos(\Phi_\alpha(x), \Phi_\alpha(y))]^2 \\
&= \phi_\alpha^{-1} \circ \cos^2(\Phi_\alpha(x), \Phi_\alpha(y))
\end{aligned}$$

(ii) The cosine in $\mathbb{E}^d$ is given by the ϕ_α-Pearson correlation coefficient:

$$\begin{aligned}
\mathrm{cor}_\alpha(\mathbf{f}_k, \mathbf{z}_k) &= \frac{\langle \mathbf{f}_k, \mathbf{z}_k\rangle_\alpha}{\|\mathbf{f}_k\|_2^\alpha \overset{\phi_\alpha}{\cdot} \|\mathbf{z}_k\|_2^\alpha} \\
&= \frac{\phi_\alpha^{-1} \circ \langle \Phi_\alpha(\mathbf{f}_k), \Phi_\alpha(\mathbf{z}_k)\rangle}{\phi_\alpha^{-1} \circ (\|\Phi_\alpha(\mathbf{f}_k)\|_2 \cdot \|\Phi_\alpha(\mathbf{z}_k)\|_2)} \\
&= \phi_\alpha^{-1} \circ \mathrm{cor}(\Phi_\alpha(\mathbf{f}_k), \Phi_\alpha(\mathbf{z}_k)),
\end{aligned}$$

with $\mathrm{cor}(\cdot)$ the usual Pearson correlation coefficient.
(iii) By Theorem **1**, the eigenvectors $\tilde{b}_k$ in $(\mathbb{R}^d, +, \cdot)$ are $b_k = \Phi^{-1}(\tilde{b}_k)$ in the transformed space $\mathbb{E}^d$. Then, the inertia of the principal component $\mathbf{f}_k = \boldsymbol{Z} \overset{\Phi_\alpha}{\cdot} b_k$ in $\mathbb{E}^d$ is:

$$b_k^\intercal \overset{\Phi_\alpha}{\cdot} \boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k$$

On the other hand, we define the following Lagrangian:

$$L(b_k, \lambda_k) = b_k^\intercal \overset{\Phi_\alpha}{\cdot} \boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k \overset{\phi_\alpha}{-} \lambda_k \overset{\phi_\alpha}{\cdot} \left[1 \overset{\phi_\alpha}{-} b_k^\intercal \overset{\phi_\alpha}{\cdot} b_k\right]$$

Maximizing $L(\cdot)$ yields,

$$\boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k = \lambda_k \overset{\Phi_\alpha}{\cdot} b_k,$$

which is the eigenvalue equation in the transformed space $\mathbb{E}^d$. Rewriting the expression provides:

$$\begin{aligned}
&\boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k = \lambda_k \overset{\Phi_\alpha}{\cdot} b_k \\
\iff\ & b_k^\intercal \overset{\Phi_\alpha}{\cdot} \boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k = b_k^\intercal \overset{\Phi_\alpha}{\cdot} \lambda_k \overset{\Phi_\alpha}{\cdot} b_k \\
\iff\ & b_k^\intercal \overset{\Phi_\alpha}{\cdot} \boldsymbol{R} \overset{\Phi_\alpha}{\cdot} b_k = \lambda_k = \phi_\alpha^{-1}(\tilde{\lambda}_k)
\end{aligned}$$

Since $\mathbb{V}(\tilde{\mathbf{f}}_k) = \tilde{\lambda}_k$, the result follows.

C. Proof of Proposition 2

(i) By definition, a U-statistics is an unbiased estimator of a parameter $\theta(F)$ expressed as [27]

$$\theta(F) = \int \cdots \int h(x_1, \cdots, x_m) dF(x_1) \cdots dF(x_m)$$

where the function $h(\cdot)$ is called a kernel, which has to be symmetric in its arguments. The corresponding U-statistics of $\theta(F)$ on a sample $X_1, \cdots, X_n$ of size $n \geq m$ is:

$$U = \frac{1}{\binom{n}{m}} \sum_c h(X_{i_1}, \cdots, X_{i_m}),$$

where c denotes the number of combinations $\binom{n}{m}$ of m elements $\{i_1, \cdots, i_m\}$ among $\{1, \cdots, n\}$. Let us define the sample variance as a U-statistics with the following kernel h_1 and the following change of variable $\Phi_\alpha(\mathbf{f}_k) = X$,

$$h_1(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$$

Then, the unbiased estimator of the sample variance of $\Phi_\alpha(\mathbf{f}_k)$ is given by:

$$U_1 = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} h_1(X_i, X_j) = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) = \mathrm{var}(X)$$

Let $\Phi_\alpha(\mathbf{z}_k) = Y$, then the unbiased estimator of the sample variance of $\Phi_\alpha(\mathbf{z}_k)$ is given by:

$$U_2 = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} h_1(Y_i, Y_j) = \frac{1}{n-1} \left(\sum_{i=1}^n Y_i^2 - n\bar{Y}^2 \right) = \mathrm{var}(Y)$$

We now define the sample covariance with $V_i = (X_i, Y_i)$ ordered pairs of samples. We define the following kernel on pairs:

$$h_2(V_1, V2) = \frac{1}{2}(X_1 - X_2)(Y_1 - Y_2)$$

Then, the unbiased estimator of the covariance is given by the following U-statistics:

$$\begin{aligned}U_3 =& \frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} h_2(V_i,V_j) = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \frac{1}{2}(X_i - X_j)(Y_i - Y_j) \\ =& \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})\end{aligned}$$

Note that,

$$\mathrm{cor}_\alpha(\mathbf{f}_k,\mathbf{z}_k) = \phi_\alpha^{-1} \circ \frac{\mathrm{cov}(\Phi_\alpha(\mathbf{f}_k),\Phi_\alpha(\mathbf{z}_k))}{\sqrt{\mathrm{cov}(\Phi_\alpha(\mathbf{f}_k),\Phi_\alpha(\mathbf{f}_k))\mathrm{cov}(\Phi_\alpha(\mathbf{z}_k),\Phi_\alpha(\mathbf{z}_k))}} = \phi_\alpha^{-1} \circ \frac{U_3}{\sqrt{(U_1U_2)}}$$

We now invoke Slutzky's theorem, see [29], Theorem 10.4:

Slutzky's Theorem: The asymptotic distribution of a function of several (dependent) U-statistics.

Let $U' = (U_1, \cdots, U_t)$ be t U-statistics based on a sample $X_1, \cdots, X_n$ of size n, with U_i corresponding to y_i (with kernel h_i), $i = 1, \cdots, t$. If the function $g(y) = g(y_1, \cdots, y_t)$ does not involve n and is continuous together with its partial derivatives in some neighborhood of the point $y = \theta = (\theta_1, \cdots, \theta_n)$ and if $\mathbb{E}(h_i^2(X_{1i}, \cdots, X_{mi}))$ exist for all i then the distribution of $\sqrt{n}(g(U') - g(y))$ tends to the normal distribution as $n \to \infty$.

Clearly, on the basis of the previous theorem, $\mathrm{cor}_\alpha(\mathbf{f}_k,\mathbf{z}_k)$ is asymptotically normal.

(ii) The same demonstration applies for $\cos^2_\alpha(\mathbf{f}_i,\mathbf{f}_k)$, just note that $\cos^2_\alpha = \mathrm{cor}^2_\alpha$.

D. Additional experiments

MNIST classification: Accuracy according to 1 to 5 principal components

# P. components	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.7$	$\alpha = 0.9$	PCA
1	0.142	0.174	**0.241**	0.176	0.175	0.174
2	0.128	0.248	**0.274**	0.261	0.267	0.261
3	0.149	0.286	**0.334**	0.268	0.265	0.259
4	0.177	0.342	**0.379**	0.282	0.258	0.250
5	0.185	0.421	**0.454**	0.383	0.336	0.316

Fashion-MNIST classification: Accuracy according to 1 to 5 principal components

# P. components	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.7$	$\alpha = 0.9$	PCA
1	0.202	0.214	**0.217**	0.152	0.120	0.105
2	0.263	0.353	**0.360**	0.237	0.186	0.178
3	0.279	0.394	**0.419**	0.287	0.217	0.191
4	0.314	**0.450**	0.441	0.291	0.226	0.193
5	0.350	**0.509**	0.500	0.349	0.260	0.230

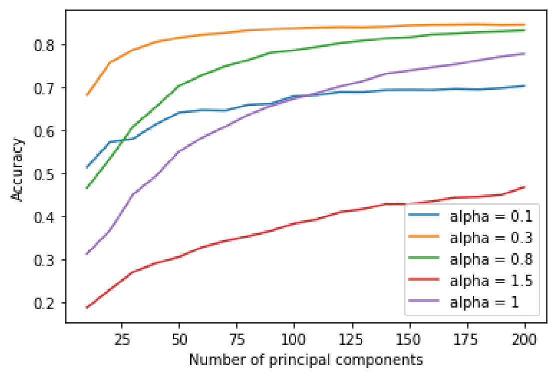

Fig. 13. Accuracy on Fashion-MNIST

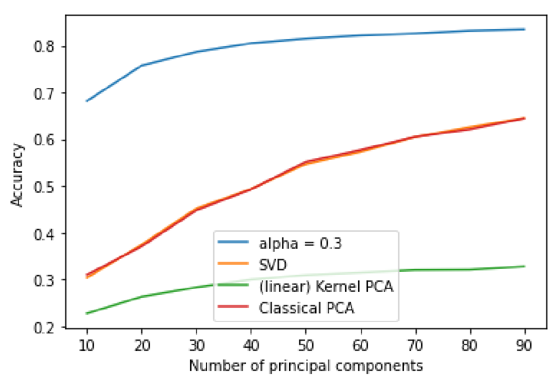

Fig. 14. Accuracy on Fashion-MNIST

It is noteworthy that the kernel PCA has not been fine-tuned (to reduce computation time), which explains its lower performance.