1 PCA Question

Assume the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_d$ are all unit length vectors, i.e. $\forall i \in [d], \quad \|\mathbf{w}_i\|_2^2 = 1$ and for any $i \neq j$, $\mathbf{w}_i \perp \mathbf{w}_j$, that is: $\mathbf{w}_i^\top \mathbf{w}_j = 0$. Now say

$$\mathbf{x}_t = \mu + \sum_{j=1}^{d} \mathbf{y}_t[j]\mathbf{w}_j \quad \text{and} \quad \hat{\mathbf{x}}_t = \mu + \sum_{j=1}^{K} \mathbf{y}_t[j]\mathbf{w}_j$$

Question 1: For any K, we have shown that:

$$\frac{1}{n}\sum_{t=1}^{n} \|\hat{\mathbf{x}}_t - \mathbf{x}_t\|_2^2 = \sum_{j=K+1}^{d} \mathbf{w}_j^\top \Sigma \mathbf{w}_j$$

using this show that

$$\sum_{j=1}^{d} \mathbf{w}_j^\top \Sigma \mathbf{w}_j = \frac{1}{n}\sum_{t=1}^{n} \|\mathbf{x}_t - \mu\|_2^2$$

and hence conclude that finding orthonormal W that maximizes

$$\sum_{k=1}^{K} \frac{1}{n}\sum_{t=1}^{n} \left(y_t[k] - \frac{1}{n}\sum_{s=1}^{n} y_s[k] \right)^2$$

is equivalent to finding orthonormal W that minimizes

$$\frac{1}{n}\sum_{t=1}^{n} \|\hat{\mathbf{x}}_t - \mathbf{x}_t\|_2^2$$

2 Random Projection Question

Assume that the d dimensional vector $\mathbf{u}$ is obtained by setting each of its coordinates to either $+1$ or -1 based on coin flips. That is, for each $i \in [d]$, $\mathbf{u}[i] = \begin{cases} +1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$.
Now set $y_t = \mathbf{x}_t^\top \mathbf{u}$ for every $t \in [n]$.

Question: For some $t, s \in [n]$, what is:

1. $\mathbb{E}[y_t - y_s]$?
2. $\mathbb{E}[(y_t - y_s)^2]$?