This supplementary document contains the omitted proofs of lemmas in Section 1 of the main paper. We restate the lemmas and prove them here.

**Lemma 0.1.** If $\hat{A}$ is a matrix with orthonormal columns such that $\text{range}(\hat{A}) = \text{range}\left( \sqrt[2]{A} \right)$ and if $U_1$ comprises the first $n$ rows of $\hat{A}$, then $\|U_1\|_2^2 = \text{sd}_A(A)$ and $\|U_1\|_2^2 = 1/(1 + \lambda/\sigma_i^2) \leq 1$.

*Proof.* See Lemma 12 of (Avron et al., 2017) for a proof.

**Lemma 0.2.** If $A'$ is the sub-matrix of $A$ formed by taking rows of $A$, then $\text{sd}_A(A') \leq \text{sd}_A(A)$.

*Proof.* The minimax characterization of singular values is as follows:

\[
\sigma_i(A) = \max_{U: \dim(U) = i} \min_{x : \|x\|_2 = 1} \|Ax\|_2 \tag{0.1}
\]

For any vector $x$, we have $\|A'x\|_2 \leq \|Ax\|_2$ (Since, vector $Ax$ has all the elements of $A'x$ and more). Hence, for any subspace $U$, $\min_{x : \|x\|_2 = 1} \|A'x\| \leq \min_{x : \|x\|_2 = 1} \|Ax\|_2$. Now, it is easy to see that

\[
\max_{U: \dim(U) = i} \min_{x : \|x\|_2 = 1} \|A'x\|_2 \leq \max_{U: \dim(U) = i} \min_{x : \|x\|_2 = 1} \|Ax\|_2 \Rightarrow \sigma_i(A') \leq \sigma_i(A) \tag{0.2}
\]

Now,

\[
\text{sd}_A(A') = \sum_{i=1}^{\text{rank}(A')} \frac{1}{1 + \lambda/\sigma_i^2(A')} \leq \sum_{i=1}^{\text{rank}(A')} \frac{1}{1 + \lambda/\sigma_i^2(A')} \quad \text{(Using (0.2))}
\]

\[
\leq \sum_{i=1}^{\text{rank}(A)} \frac{1}{1 + \lambda/\sigma_i^2(A)} = \text{sd}_A(A) \tag{0.3}
\]

**Lemma 0.3.** For any $r \geq 1$, $\text{sd}_{\lambda/r}(A) \leq \min(r \cdot \text{sd}_A(A), \text{rank}(A))$.

*Proof.* By definition of statistical dimension, $\text{sd}_A(A) \leq \text{rank}(A)$. We also have for all $r \geq 1$, $1/(1 + \lambda/\sigma_i^2) \leq r/(1 + \lambda/\sigma_i^2)$. By summing the inequality for all $i$, we get $\text{sd}_{\lambda/r}(A) \leq r \cdot \text{sd}_A(A)$.