Annexe de
Vote de majorité *a priori* contraint pour de la classification binaire

Aurélien Bellet[1], Amaury Habrard[2], Emilie Morvant[3], et Marc Sebban[2]

[1]euh... ???
[2]Université Jean Monnet de Saint-Étienne, Laboratoire Hubert Curien, UMR CNRS 5516
[3]Aix-Marseille Univ., LIF-QARMA, UMR CNRS 7279

27 mai 2013

1 Tools

Theorem 4 (Markov's inequality)**.** *Let Z be a random variable and $t \geq 0$, then :*

$$P(|Z| \geq t) \leq \mathbf{E}(|Z|)/t.$$

Theorem 5 (Jensen's inequality)**.** *Let X be an integrable real-valued random variable and $g(\cdot)$ convex, then :*

$$g(\mathbf{E}[Z]) \leq \mathbf{E}[g(Z)].$$

Lemme 1 (from inequalities (1) and (2) of [Mau04])**.** *Let $m \geq 8$, and $X = (X_1, .., X_m)$ be a vector of i.i.d. random variables, $0 \leq X_i \leq 1$. Then :*

$$\sqrt{m} \leq \mathbf{E} \exp(m \,\mathrm{kl}(\tfrac{1}{m} \textstyle\sum_{i=1}^n X_i \big\| \mathbf{E}[X_i])) \leq 2\sqrt{m},$$

where $\mathrm{kl}(a\|b) = a \ln \frac{a}{b} + (1-a) \ln \frac{1-a}{1-b}$.

2 Proof of Proposition 2

Proposition 2. *For all distributions Q on $\mathcal{H}$, there exists a $\mathbf{P}$-aligned distribution Q' on the auto-complemented $\mathcal{H}$ that provides the same majority vote as Q, and that has the same empirical and true C-bound values.*

Démonstration. Let Q be a distribution over $\mathcal{H}$, let M be defined as $M = \max_{k' \in \{1,\ldots,n\}} \frac{1}{P_{k'}} |Q_{k'+n} - Q_{k'}|$, and let Q' be defined as $Q'_k = \frac{P_k}{2} + \frac{Q_k - Q_{k+n}}{2M}$, where by convention $(k+n)+n = k$ and $P_{k+n} = P_k$. First, let us show that Q' is actually $\mathbf{P}$-aligned on the auto-complemented $\mathcal{H}$, that is $\forall k \in \{1, \ldots, n\}$, $Q'_k \leq P_k$ and $Q'_k + Q'_{k+n} = P_k$. We have :

$$\begin{aligned}
& Q'_k && \leq P_k \\
\Leftrightarrow\quad & \frac{P_k}{2} + \frac{Q_k - Q_{k+n}}{2M} && \leq P_k \\
\Leftrightarrow\quad & \frac{Q_k - Q_{k+n}}{M} && \leq P_k \\
\Leftrightarrow\quad & \frac{1}{P_k}(Q_k - Q_{k+n}) && \leq \max_{k' \in \{1,\ldots,n\}} \frac{1}{P_{k'}} |Q_{k'+n} - Q_{k'}|,
\end{aligned}$$

which always holds.
Moreover :

$$\begin{aligned}Q'_k + Q'_{k+n} &= \frac{P_k}{2} + \frac{Q_k - Q_{k+n}}{2M} + \frac{P_{k+n}}{2} + \frac{Q_{k+n} - Q_k}{2M} \\ &= P_k + \frac{Q_k - Q_{k+n} + Q_{k+n} - Q_k}{2M} \\ &= P_k.\end{aligned}$$

Then, let us show that using Q' does not restrict the set of possible majority votes :

$$\begin{aligned}\mathop{\mathbf{E}}_{h\sim Q'} h(x) &= \sum_{k=1}^{2n} Q'_k h_k(\mathbf{x}) \\ &= \sum_{k=1}^{n} (Q'_k - Q'_{k+n}) h_k(\mathbf{x}) \\ &= \frac{1}{M}\sum_{k=1}^{n} (Q_k - Q_{k+n}) h_k(\mathbf{x}) \\ &= \frac{1}{M}\sum_{k=1}^{2n} Q_k h_k(\mathbf{x}) \\ &= \frac{1}{M}\mathop{\mathbf{E}}_{h\sim Q} h(\mathbf{x}).\end{aligned}$$

Therefore, we deduce that $\forall \mathbf{x} \in \mathcal{X}$, $B_{Q'}(\mathbf{x}) = B_Q(\mathbf{x})$ and since the constant term $\frac{1}{M}$ is present in both first and second moments $\mathcal{M}^D_{Q'}$ and $\mathcal{M}^D_{Q'^2}$, it vanishes in the C-bound. Hence, $C^D_{Q'} = C^D_Q$ regardless of the distribution D over $\mathcal{X} \times \mathcal{Y}$. □

3 Proof of the Algorithm P-MinCq

Proof of Algorithm 2. **The Objective Function.** In the following, we show how to obtain Eq. (6) from the definition of the second moment $\mathcal{M}^S_{Q^2}$ of the Q-margin over S.

$$\begin{aligned}&\mathcal{M}^S_{Q^2} = \mathop{\mathbf{E}}_{(h,h')\sim Q^2} \mathcal{M}^S_{h,h'} \\ &= \sum_{k=1}^{2n}\sum_{k'=1}^{2n} Q_k Q_{k'} \mathcal{M}^S_{h_k,h_{k'}} \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \Big[Q_k Q_{k'} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'}(\mathbf{x}) + Q_{k+n} Q_{k'} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_{k+n}(\mathbf{x})h_{k'}(\mathbf{x}) + Q_k Q_{k'+n} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'+n}(\mathbf{x}) + Q_{k+n} Q_{k'+n} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_{k+n}(\mathbf{x})h_{k'+n}(\mathbf{x})\Big] \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} Q_k Q_{k'} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'}(\mathbf{x}) - Q_{k+n} Q_{k'} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'}(\mathbf{x}) - Q_k Q_{k'+n} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'}(\mathbf{x}) + Q_{k+n} Q_{k'+n} \mathop{\mathbf{E}}_{(\mathbf{x},y)\sim S} h_k(\mathbf{x})h_{k'}(\mathbf{x}) \\ &\qquad\qquad\qquad \text{(because } h_{k+n} = -h_k\text{)} \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \mathcal{M}^S_{h_k,h_{k'}} [Q_k Q_{k'} - (P_k - Q_k)Q_{k'} - Q_k(P_{k'} - Q_{k'}) + (P_k - Q_k)(P_{k'} - Q_{k'})] \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \mathcal{M}^S_{h_k,h_{k'}} [4Q_k Q_{k'} - 2P_k Q_{k'} - 2P_{k'} Q_k + P_k P_{k'}] \\ &= 4\sum_{k=1}^{n}\sum_{k'=1}^{n} Q_k \mathcal{M}^S_{h_k,h_{k'}} Q_{k'} - 4\sum_{k=1}^{n}\sum_{k'=1}^{n} P_k \mathcal{M}^S_{h_k,h_{k'}} Q_{k'} + \sum_{k=1}^{n}\sum_{k'=1}^{n} P_k P_{k'} \mathcal{M}^S_{h_k,h_{k'}} \\ &= 4[(\mathbf{Q} - \mathbf{P})^{\mathbf{T}} \mathbf{M_S} \mathbf{Q}] + C_1,\end{aligned}$$

where $C_1 = \sum_{k=1}^{n}\sum_{k'=1}^{n} P_k P_{k'} \mathcal{M}_{h_k,h_{k'}}$ and the multiplicative value 4 can be considered as constant w.r.t. Q. Therefore, we get Eq. (6) of the optimization problem.

The Margin Constraint. We now show how to obtain the Eq. (7) from $\mathcal{M}_Q^{\mathcal{S}}$.

$$\begin{aligned}
\mathcal{M}_Q^{\mathcal{S}} =& \mathop{\mathbf{E}}_{h\sim Q} \mathcal{M}_h^S \\
=& \sum_{k=1}^{2n} Q_k \mathcal{M}_{h_k}^S \\
=& \sum_{k=1}^{n} (Q_k - Q_{k+n}) \mathcal{M}_{h_k}^S \\
=& \sum_{k=1}^{n} (2Q_k - P_k) \mathcal{M}_{h_k}^S \\
=& \mathbf{m_S^T}(2\mathbf{Q} - \mathbf{P}),
\end{aligned}$$

where $\mathbf{m_S^T} = (\mathcal{M}_{h_1}, \ldots, \mathcal{M}_{h_n})^T$. Replacing $\mathcal{M}_Q^{\mathcal{S}}$ by μ, we get Eq. (7) of the optimization problem. □

4 Proof of Theorem 3

We first recall the theorem.

Theorem 3. *For any distribution D over $\mathcal{X}\times\mathcal{Y}$, any $m \geq 8$, any $\delta \in (0, 1]$, with probability at least $1-\delta$ over any sample S from D^m, for any auto-complemented family $\mathcal{H}^S$ of B-bounded real value functions of sample compression size at most $|\mathbf{j}^{\max}| < \frac{m}{2}$ and for all $\mathbf{P}$-aligned distribution Q on $\mathcal{H}^S$:*

$$\left|\mathcal{M}_Q^D - \mathcal{M}_Q^S\right| \leq \frac{2B\sqrt{\frac{|\mathbf{j}^{\max}|}{B\delta} + \ln\left(\frac{2\sqrt{m}}{\delta}\right)}}{\sqrt{2(m-|\mathbf{j}^{\max}|)}}, \tag{9}$$

$$\left|\mathcal{M}_{Q^2}^D - \mathcal{M}_{Q^2}^S\right| \leq \frac{2B^2\sqrt{\frac{2|\mathbf{j}^{\max}|}{B^2\delta} + \ln\left(\frac{2\sqrt{m}}{\delta}\right)}}{\sqrt{2(m-2|\mathbf{j}^{\max}|)}}. \tag{10}$$

Proof of Equation (9)**.** Let S be any training sequence of size m. Suppose that $\mathcal{H}^S$ is auto-complemented. Moreover, a distribution on $\mathcal{H}^S$ is $\mathbf{P}$-aligned if for any $(\mathbf{j},\sigma) \in \mathbf{J}_m \times \Omega_{S_\mathbf{j}}$ we have :

$$Q(h_S^{(\sigma,+)}) + Q(-h_S^{(\sigma,+)}) = Q(h_S^{(\sigma,+)}) + Q(h_S^{(\sigma,-)}) = P(h_S^{(\sigma,+)}) + P(h_S^{(\sigma,-)}) = P(h_S^{(\sigma,+)}) + P(-h_S^{(\sigma,+)}).$$

It implies that :

$$\mathcal{M}^D_{h_S^{(\sigma,+)}} = -\mathcal{M}^D_{h_S^{(\sigma,-)}},$$

and :

$$\left(\mathcal{M}^S_{h_{S_\mathbf{j}}^{(\sigma,+)}} - \mathcal{M}^D_{h_{S_\mathbf{j}}^{(\sigma,+)}}\right)^2 = \left(-\mathcal{M}^S_{h_{S_\mathbf{j}}^{(\sigma,-)}} - (-\mathcal{M}^D_{h_{S_\mathbf{j}}^{(\sigma,-)}})\right)^2 = \left(\mathcal{M}^S_{h_{S_\mathbf{j}}^{(\sigma,-)}} - \mathcal{M}^D_{h_{S_\mathbf{j}}^{(\sigma,-)}}\right)^2.$$

Similarly as in [McA03], we now consider the following Laplace transform :

$$X_P = \mathop{\mathbf{E}}_{h_{S_\mathbf{j}}^{\omega}\sim P} \exp\left(\frac{m-|\mathbf{j}|}{2B^2}\left(\mathcal{M}^S_{h_{S_\mathbf{j}}^{\omega}} - \mathcal{M}^D_{h_{S_\mathbf{j}}^{\omega}}\right)^2\right).$$

Remark that $f(a,b) = \frac{1}{2B^2}(a-b)^2$ is convex because its Hessian matrix is positive semi-definite. For lightening the proof reading, we denote $m_{\mathbf{j}} = \dfrac{m-|\mathbf{j}|}{2B^2}$.

For any $\mathbf{P}$-aligned distribution Q, we have :

$$
\begin{aligned}
2X_P =& \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim P} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} P(h^{(\sigma,+)}_{S_{\mathbf{j}}}) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,+)}_{S_{\mathbf{j}}} + \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} P(h^{(\sigma,-)}_{S_{\mathbf{j}}}) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,-)}_{S_{\mathbf{j}}}\\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} \big(P(h^{(\sigma,+)}_{S_{\mathbf{j}}}) + P(-h^{(\sigma,+)}_{S_{\mathbf{j}}})\big) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,+)}_{S_{\mathbf{j}}}\\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} \big(Q(h^{(\sigma,+)}_{S_{\mathbf{j}}}) + Q(-h^{(\sigma,+)}_{S_{\mathbf{j}}})\big) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,+)}_{S_{\mathbf{j}}}\\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} Q(h^{(\sigma,+)}_{S_{\mathbf{j}}}) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,+)}_{S_{\mathbf{j}}} + \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}\in\mathcal{H}^S} Q(h^{(\sigma,-)}_{S_{\mathbf{j}}}) \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}}\big)^2\right) dh^{(\sigma,-)}_{S_{\mathbf{j}}}\\
=& 2 \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
=& 2X_Q.
\end{aligned}
$$

Using Markov's inequality (Theorem 4 in Supplemental Material) we have :

$$
\mathop{\mathbf{Pr}}_{S\sim D^m}\left(X_P \le \frac{1}{\delta}\mathop{\mathbf{E}}_{S\sim D^m} X_P\right) \ge 1-\delta.
$$

Taking the logarithm on each side of the innermost inequality, for any $\delta \in (0,1]$, with a probability at least $1-\delta$ over the choice of $S\sim D^m$, for all $\mathbf{P}$-aligned distribution Q on $\mathcal{H}^S$, we get :

$$
\ln\left[\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\right] \le \ln\left[\frac{1}{\delta}\mathop{\mathbf{E}}_{S\sim D^m} X_P\right].
$$

We apply Jensen's inequality (Theorem 5 in Supplemental Material) on the concave function $\ln(\cdot)$:

$$
\ln\left[\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\right] \ge \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2.
$$

Recall that $|\mathbf{j}^{\max}|$ is the maximal size of the compression sample. Then by again applying the Jensen's inequality on the convex function $(m-|\mathbf{j}^{\max}|)f(a,b) = \frac{m-|\mathbf{j}^{\max}|}{2B^2}(a-b)^2 = m_{\mathbf{j}}(a-b)^2$ for the right side of the previous inequality, we have :

$$
\begin{aligned}
\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} m_{\mathbf{j}}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2 &= \frac{m}{2B^2}\left(\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} - |\mathbf{j}|\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&\ge \frac{m-|\mathbf{j}^{\max}|}{2B^2}\left(\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim Q} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&\ge \frac{m-|\mathbf{j}^{\max}|}{2B^2}\big(\mathcal{M}^S_Q - \mathcal{M}^D_Q\big)^2.
\end{aligned}
$$

Then :

$$
\mathop{\mathbf{Pr}}_{S\sim D^m}\left(\frac{m-|\mathbf{j}^{\max}|}{2B^2}\big(\mathcal{M}^S_Q - \mathcal{M}^D_Q\big)^2 \le \ln\left[\frac{1}{\delta}\mathop{\mathbf{E}}_{S\sim D^m} X_P\right]\right) \ge 1-\delta.
$$

We thus have to bound $\mathop{\mathbf{E}}_{S\sim D^m} X_P$. We consider $\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}}$ the empirical margin computed on the examples of the learning sample S that are not in the compression sequence $S_{\mathbf{j}}$. While $\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}}$ may contain some bias, $\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}}$ is an arithmetic mean of truly *i.i.d.* $(m-|\mathbf{j}|)$ random variables. Note also that these two random variables have very close values. We have :

$$0 \le m\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - (m-|\mathbf{j}|)\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}} \le B|\mathbf{j}|,$$

then :

$$-B|\mathbf{j}| \le -|\mathbf{j}|\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}} \le m\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - m\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}} \le |\mathbf{j}| - |\mathbf{j}|\mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}} \le B|\mathbf{j}|,$$

and thus :

$$\left|\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{S\setminus S_{\mathbf{j}}}_{h^{\omega}_{S_{\mathbf{j}}}}\right| \le \frac{B|\mathbf{j}|}{m}. \tag{11}$$

Given a compression sequence $S_{\mathbf{j}}$, we denote by $\bar{\mathbf{j}}$ the vector of indices that are not in $\mathbf{j}$. Then :

$$\begin{aligned}
\mathop{\mathbf{E}}_{S\sim D^m} X_P &= \mathop{\mathbf{E}}_{S\sim D^m}\ \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}\sim P} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&= \mathop{\mathbf{E}}_{\mathbf{j}\sim P}\ \mathop{\mathbf{E}}_{S_{\mathbf{j}}\sim D^{|\mathbf{j}|}}\ \mathop{\mathbf{E}}_{\omega\sim P_{S_{\mathbf{j}}}}\ \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right).
\end{aligned}$$

For all $\mathbf{j}\in\mathbf{J}_m$, $S_{\mathbf{j}}\in\mathcal{Z}^{|\mathbf{j}|}$, $\omega\in\Omega'_{S_{\mathbf{j}}}\times\{+,-\}$, we have :

$$\begin{aligned}
\mathop{\mathbf{E}}_{S\sim D^m} X_P &= \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&= \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(m_{\mathbf{j}}\big(\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} + \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&\le \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left[m_{\mathbf{j}}\left(\big[\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}}\big]^2 + 2\big|\mathcal{M}^{S}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}}\big|\big|\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big| + \big[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big]^2\right)\right].
\end{aligned}$$

From Equation (11) and since exp(.) is increasing, we obtain :

$$\mathop{\mathbf{E}}_{S\sim D^m} X_P \le \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left[m_{\mathbf{j}}\left(\left[\frac{B|\mathbf{j}|}{m}\right]^2 + 2\frac{B|\mathbf{j}|}{m} + \big[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big]^2\right)\right].$$

Since we suppose that for all $\mathbf{j}$: $|\mathbf{j}| \le |\mathbf{j}^{\max}| \le m$, then :

$$\frac{m-|\mathbf{j}|}{2B}\left(\left[\frac{|\mathbf{j}|}{m}\right]^2 + 2\frac{|\mathbf{j}|}{m}\right) \le |\mathbf{j}^{\max}|\left(\frac{m-|\mathbf{j}|}{2B}\left[\frac{|\mathbf{j}|}{m^2} + \frac{2}{m}\right]\right) \le \frac{|\mathbf{j}^{\max}|}{B}.$$

Then :

$$\begin{aligned}
\mathop{\mathbf{E}}_{S\sim D^m} X_P &\le \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(\frac{|\mathbf{j}^{\max}|}{B} + m_{\mathbf{j}}\big(\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big)^2\right)\\
&\le \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(m_{\mathbf{j}}\big[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}} - \mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}\big]^2\right)\\
&\le \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left(2(m-|\mathbf{j}|)\left[\Big(\frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}}}}{2B}\Big) - \Big(\frac{1}{2} - \frac{\mathcal{M}^{D}_{h^{\omega}_{S_{\mathbf{j}}}}}{2B}\Big)\right]^2\right).
\end{aligned}$$

By definition $2(a-b)^2 \leq \mathrm{kl}(a\|b) = a\ln\frac{a}{b} + (1-a)\ln\frac{1-a}{1-b}$ is valid for any $a, b \in [0,1]$ provided that if $a=0$ then so is b and if $a=1$ then so is b. Since the elements of $\mathcal{H}^S$ are B-bounded and $S_{\bar{\mathbf{j}}}$ is drawn *i.i.d.* from D, we have :

$$\mathcal{M}^D_{h^\omega_{S_\mathbf{j}}} = -B \Rightarrow \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}} = -B, \quad \text{and} \quad \mathcal{M}^D_{h^\omega_{S_\mathbf{j}}} = B \Rightarrow \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}} = B.$$

Then :

$$\frac{1}{2} - \frac{\mathcal{M}^D_{h^\omega_{S_\mathbf{j}}}}{2B} = 0 \Rightarrow \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}}}{2B} = 0, \quad \text{and} \quad \frac{1}{2} - \frac{\mathcal{M}^D_{h^\omega_{S_\mathbf{j}}}}{2B} = 1 \Rightarrow \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}}}{2B} = 1.$$

Moreover since :

$$0 \leq \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}}}{2B} \leq 1, \quad \text{and} \quad 0 \leq \frac{1}{2} - \frac{\mathcal{M}^D_{h^\omega_{S_\mathbf{j}}}}{2B} \leq 1,$$

we have :

$$\mathop{\mathbf{E}}_{S\sim D^m} X_P \leq \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} \exp\left((m-|\mathbf{j}|)\,\mathrm{kl}\left(\frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}}}}{2B} \middle\| \frac{1}{2} - \frac{\mathcal{M}^D_{h^\omega_{S_\mathbf{j}}}}{2B}\right)\right).$$

We apply Maurer's Lemma (Lemma 1 in Supplemental Material) :

$$\begin{aligned}\mathop{\mathbf{E}}_{S\sim D^m} X_P &\leq \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}|}} 2\sqrt{(m-|\mathbf{j}|)} \\ &\leq \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + 2\sqrt{(m-|\mathbf{j}|)} \leq \exp\left(\frac{|\mathbf{j}^{\max}|}{B}\right) + 2\sqrt{m}.\end{aligned}$$

Finally :

$$\mathop{\mathbf{Pr}}_{S\sim D^m}\left(\begin{array}{l}\text{for all } \mathbf{P}\text{-aligned distribution } Q \text{ on } \mathcal{H}^S, \\ |\mathcal{M}^D_Q - \mathcal{M}^S_Q| \leq \dfrac{2B\sqrt{\dfrac{|\mathbf{j}^{\max}|}{B\delta} + \ln\left(\dfrac{2\sqrt{m}}{\delta}\right)}}{\sqrt{2(m-|\mathbf{j}^{\max}|)}}\end{array}\right) \geq 1-\delta$$

□

Proof of Equation (10)**.** Using similar arguments as the beginning of the proof Equation (9), we have :

$$\begin{aligned}\left(\mathcal{M}^S_{h^{(\sigma,+)}_{S_\mathbf{j}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_\mathbf{j}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}}\right)^2 =& \left(\mathcal{M}^S_{h^{(\sigma,-)}_{S_\mathbf{j}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_\mathbf{j}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}}\right)^2 \\ =& \left(\mathcal{M}^S_{h^{(\sigma,+)}_{S_\mathbf{j}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_\mathbf{j}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}}\right)^2 \\ =& \left(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}'}}, h^{(\sigma,-)}_{S_\mathbf{j}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}'}}, h^{(\sigma,-)}_{S_\mathbf{j}}}\right)^2.\end{aligned}$$

Similarly as in [McA03], we now consider the following Laplace transform :

$$X_P = \mathop{\mathbf{E}}_{h^\omega_{S_\mathbf{j}}, h^\omega_{S'_\mathbf{j}}\sim P^2} \exp\left(\frac{m-|\mathbf{j}\cup\mathbf{j}'|}{2B^4}\left(\mathcal{M}^S_{h^\omega_{S_\mathbf{j}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^\omega_{S_\mathbf{j}}, h^{\omega'}_{S_{\mathbf{j}'}}}\right)^2\right).$$

For lightening the proof reading, we denote $m_{\mathbf{j}\cup\mathbf{j}'} = \dfrac{m - |\mathbf{j}\cup\mathbf{j}'|}{2B^4}$. Remark that $f(a,b) = \frac{1}{2B^4}(a-b)^2$ is convex. For any $\mathbf{P}$-aligned distribution Q, we have :

$$
\begin{aligned}
4X_P =& \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega}_{S'_{\mathbf{j}}} \sim P^2} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} \big)^2 \Big) \\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} P(h^{(\sigma,+)}_{S_{\mathbf{j}}}) P(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} P(h^{(\sigma,-)}_{S_{\mathbf{j}}}) P(h^{(\sigma,-)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,-)}_{S_{\mathbf{j}}} h^{(\sigma,-)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} P(h^{(\sigma,-)}_{S_{\mathbf{j}}}) P(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,-)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} P(h^{(\sigma,+)}_{S_{\mathbf{j}}}) P(h^{(\sigma,-)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,-)}_{S_{\mathbf{j}'}} \\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} \big(P(h^{(\sigma,+)}_{S_{\mathbf{j}}}) + P(-h^{(\sigma,+)}_{S_{\mathbf{j}}}) \big) \big(P(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) + P(-h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \big) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}}
\end{aligned}
$$

$$
\begin{aligned}
4X_P =& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} \big(Q(h^{(\sigma,+)}_{S_{\mathbf{j}}}) + Q(-h^{(\sigma,+)}_{S_{\mathbf{j}}}) \big) \big(Q(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) + Q(-h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \big) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}} \\
=& \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} Q(h^{(\sigma,+)}_{S_{\mathbf{j}}}) Q(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} Q(h^{(\sigma,-)}_{S_{\mathbf{j}}}) Q(h^{(\sigma,-)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,-)}_{S_{\mathbf{j}}} h^{(\sigma,-)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} Q(h^{(\sigma,-)}_{S_{\mathbf{j}}}) Q(h^{(\sigma,+)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,-)}_{S_{\mathbf{j}}}, h^{(\sigma,+)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,-)}_{S_{\mathbf{j}}} h^{(\sigma,+)}_{S_{\mathbf{j}'}} \\
&+ \int_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S'_{\mathbf{j}}} \in (\mathcal{H}^S)^2} Q(h^{(\sigma,+)}_{S_{\mathbf{j}}}) Q(h^{(\sigma,-)}_{S_{\mathbf{j}'}}) \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{(\sigma,+)}_{S_{\mathbf{j}}}, h^{(\sigma,-)}_{S_{\mathbf{j}'}}} \big)^2 \Big) dh^{(\sigma,+)}_{S_{\mathbf{j}}} h^{(\sigma,-)}_{S_{\mathbf{j}'}} \\
=& 4 \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega}_{S'_{\mathbf{j}}} \sim Q^2} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} \big)^2 \Big) \\
=& 4X_Q.
\end{aligned}
$$

Now, by Markov's inequality (Theorem 4) we have :

$$
\mathop{\mathbf{Pr}}_{S\sim D^m} \left(X_P \le \frac{1}{\delta} \mathop{\mathbf{E}}_{S\sim D^m} X_P \right) \ge 1-\delta.
$$

By taking the logarithm on each side of the innermost inequality, for any $\delta \in (0,1]$, with a probability at least $1-\delta$ over the choice of $S \sim D^m$, for all $\mathbf{P}$-aligned distribution Q on $\mathcal{H}^S$ we have :

$$
\ln \left[\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}} \sim Q^2} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} \big)^2 \Big) \right] \le \ln \left[\frac{1}{\delta} \mathop{\mathbf{E}}_{S\sim D^m} X_P \right].
$$

We apply Jensen's inequality (Theorem 5) on ln(.) :

$$
\ln \left[\mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}} \sim Q^2} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} \big)^2 \Big) \right] \ge \mathop{\mathbf{E}}_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}} \sim Q^2} m_{\mathbf{j}\cup\mathbf{j}'} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}}, h^{\omega'}_{S_{\mathbf{j}'}}} \big)^2.
$$

Recall that $|\mathbf{j}^{\max}| < \frac{m}{2}$ the maximal size of the compression sample. Then by again applying the Jensen's inequality on the convex function $(m-|\mathbf{j}^{\max}|)f(a,b) = \frac{m-|\mathbf{j}^{\max}|}{2B^4}(a-b)^2 = m_{\mathbf{j}\cup\mathbf{j}'}(a-b)^2$ for the right side of the previous inequality, we have :

$$\begin{aligned}\underset{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}\sim Q^2}{\mathbf{E}} m_{\mathbf{j}\cup\mathbf{j}'}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2 &= \frac{m}{2B^4}\left(\underset{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}\sim Q^2}{\mathbf{E}} (-|\mathbf{j}\cup\mathbf{j}'|)\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\right)\\ &\geq \frac{m-2|\mathbf{j}^{\max}|}{2B^4}\left(\underset{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}\sim Q^2}{\mathbf{E}} \big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\right)\\ &\geq \frac{m-2|\mathbf{j}^{\max}|}{2B^4}\big(\mathcal{M}^S_{Q^2} - \mathcal{M}^D_{Q^2}\big)^2.\end{aligned}$$

Then :

$$\underset{S\sim D^m}{\mathbf{Pr}}\left(\frac{m-2|\mathbf{j}^{\max}|}{2B^4}\big(\mathcal{M}^S_{Q^2} - \mathcal{M}^D_{Q^2}\big)^2 \leq \ln\left[\frac{1}{\delta}\underset{S\sim D^m}{\mathbf{E}} X_P\right]\right) \geq 1-\delta.$$

We thus have to bound $\underset{S\sim D^m}{\mathbf{E}} X_P$. We consider $\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}$ the empirical second moment of the margin computed on the examples of the learning sample S that are not in the compression sequence $S_{\mathbf{j}}$. While $\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}$ may contain some bias, $\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}$ is an arithmetic mean of truly *i.i.d.* $(m-|\mathbf{j}\cup\mathbf{j}'|)$ random variables. We can also note that these two random variables have very close values. We have :

$$0 \leq m\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - (m-|\mathbf{j}\cup\mathbf{j}'|)\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} \leq B^2|\mathbf{j}\cup\mathbf{j}'|,$$

then :

$$-B^2|\mathbf{j}\cup\mathbf{j}'| \leq -|\mathbf{j}\cup\mathbf{j}'|\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} \leq m\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - m\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} \leq |\mathbf{j}\cup\mathbf{j}'| - |\mathbf{j}\cup\mathbf{j}'|\mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} \leq B^2|\mathbf{j}\cup\mathbf{j}'|,$$

thus :

$$\left|\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{S\setminus(S_{\mathbf{j}}\cup S_{\mathbf{j}'})}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\right| \leq \frac{B^2|\mathbf{j}\cup\mathbf{j}'|}{m}. \tag{12}$$

Given two compression sequences $S_{\mathbf{j}}$ and $S_{\mathbf{j}'}$, Let $\bar{\mathbf{j}}$ be the vector of indices that are not in $\mathbf{j}\cup\mathbf{j}'$. Then :

$$\begin{aligned}\underset{S\sim D^m}{\mathbf{E}} X_P &= \underset{S\sim D^m}{\mathbf{E}}\ \underset{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}\sim P^2}{\mathbf{E}} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\Big)\\ &= \underset{\mathbf{j},\mathbf{j}'\sim P^2}{\mathbf{E}}\ \underset{S_{\mathbf{j}},S_{\mathbf{j}'}\sim D^{|\mathbf{j}|}\times D^{|\mathbf{j}'|}}{\mathbf{E}}\ \underset{\omega,\omega'\sim P_{S_{\mathbf{j}}}\times P_{S_{\mathbf{j}'}}}{\mathbf{E}}\ \underset{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}}{\mathbf{E}} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\Big).\end{aligned}$$

For all $\mathbf{j},\mathbf{j}' \in (\mathbf{J}_m)^2$, $S_{\mathbf{j}},S_{\mathbf{j}'} \in \mathcal{Z}^{|\mathbf{j}|}\times\mathcal{Z}^{|\mathbf{j}'|}$, $\omega,\omega' \in (\Omega'_{S_{\mathbf{j}}}\times\{+,-\})\times(\Omega'_{S_{\mathbf{j}'}}\times\{+,-\})$, we have :

$$\begin{aligned}&\underset{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}}{\mathbf{E}} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\Big)\\ =&\underset{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}}{\mathbf{E}} \exp\Big(m_{\mathbf{j}\cup\mathbf{j}'}\big(\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} + \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big)^2\Big)\\ \leq&\underset{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}}{\mathbf{E}} \exp\Big[m_{\mathbf{j}\cup\mathbf{j}'}\Big(\big[\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big]^2 + 2\big|\mathcal{M}^S_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big|\big|\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big|\\ &\qquad + \big[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^D_{h^{\omega}_{S_{\mathbf{j}}},h^{\omega'}_{S_{\mathbf{j}'}}}\big]^2\Big)\Big].\end{aligned}$$

From Equation (12), since exp(.) is increasing we obtain :

$$\mathop{\mathbf{E}}_{S\sim D^m} X_P \leq \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} \exp\left[m_{\mathbf{j}\cup\mathbf{j}'}\left(\left[\frac{B^2|\mathbf{j}\cup\mathbf{j}'|}{m}\right]^2 + 2\frac{B^2|\mathbf{j}\cup\mathbf{j}'|}{m} + \left[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}\right]^2\right)\right].$$

Since we suppose that for all $\mathbf{j}$ we have $|\mathbf{j}| \leq |\mathbf{j}^{\max}| \leq \frac{m}{2}$, we can easily compute :

$$m_{\mathbf{j}\cup\mathbf{j}'}\left(\left[\tfrac{|\mathbf{j}\cup\mathbf{j}'|}{m}\right]^2 + 2\tfrac{|\mathbf{j}\cup\mathbf{j}'|}{m}\right) \leq 2|\mathbf{j}^{\max}|\left[m_{\mathbf{j}\cup\mathbf{j}'}\left(\tfrac{|\mathbf{j}\cup\mathbf{j}'|}{m^2} + \tfrac{2}{m}\right)\right] \leq \tfrac{2|\mathbf{j}^{\max}|}{B^2}.$$

Then :

$$\begin{aligned}
\mathop{\mathbf{E}}_{S\sim D^m} X_P &\leq \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} \exp\left[\frac{2|\mathbf{j}^{\max}|}{B^2} + m_{\mathbf{j}\cup\mathbf{j}'}\left[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}\right]^2\right] \\
&\leq \exp\left[\frac{2|\mathbf{j}^{\max}|}{B^2}\right] + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} \exp\left[m_{\mathbf{j}\cup\mathbf{j}'}\left[\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} - \mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}\right]^2\right] \\
&\leq \exp\left[\frac{2|\mathbf{j}^{\max}|}{B^2}\right] \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} \exp\left[2(m-|\mathbf{j}\cup\mathbf{j}'|)\left[\left(\frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B}\right) - \left(\frac{1}{2} - \frac{\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B}\right)\right]^2\right].
\end{aligned}$$

We know $2(a-b)^2 \leq \mathrm{kl}(a\|b)$ is valid for any $a, b \in [0,1]$ provided that if $a = 0$ then so is b and if $a = 1$ then so is b. Since the elements of $\mathcal{H}^S$ are B-bounded and $S_{\bar{\mathbf{j}}}$ is *i.i.d.* from D, we have :

$$\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} = -B^2 \Rightarrow \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} = -B^2, \quad \text{and} \quad \mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} = B^2 \Rightarrow \mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}} = B^2.$$

Then :

$$\frac{1}{2} - \frac{\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} = 0 \Rightarrow \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} = 0, \quad \text{and} \ \frac{1}{2} - \frac{\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} = 1 \Rightarrow \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} = 1.$$

Since :

$$0 \leq \frac{1}{2} - \frac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} \leq 1, \quad \text{and} \quad 0 \leq \frac{1}{2} - \frac{\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2} \leq 1,$$

we have :

$$\mathop{\mathbf{E}}_{S\sim D^m} X_P \leq \exp\left[\frac{2|\mathbf{j}^{\max}|}{B^2}\right] + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} \exp\left[(m-|\mathbf{j}\cup\mathbf{j}'|)\,\mathrm{kl}\left(\tfrac{1}{2} - \tfrac{\mathcal{M}^{S_{\bar{\mathbf{j}}}}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2}\,\middle\|\,\tfrac{1}{2} - \tfrac{\mathcal{M}^{D}_{h^\omega_{S_\mathbf{j}},h^{\omega'}_{S_{\mathbf{j}'}}}}{2B^2}\right)\right].$$

By applying Maurer's Lemma (Lemma 1), we obtain :

$$\begin{aligned}
\mathop{\mathbf{E}}_{S\sim D^m} X_P &\leq \exp\left(\frac{2|\mathbf{j}^{\max}|}{B^2}\right) + \mathop{\mathbf{E}}_{S_{\bar{\mathbf{j}}}\sim D^{m-|\mathbf{j}\cup\mathbf{j}'|}} 2\sqrt{(m-|\mathbf{j}\cup\mathbf{j}'|)} \\
&\leq \exp\left(\frac{2|\mathbf{j}^{\max}|}{B^2}\right) + 2\sqrt{(m-|\mathbf{j}\cup\mathbf{j}'|)} \\
&\leq \exp\left(\frac{2|\mathbf{j}^{\max}|}{B^2}\right) + 2\sqrt{m}.
\end{aligned}$$

Finally, we obtain :

$$\Pr_{S\sim D^m}\left(\begin{array}{l}\text{for all } \mathbf{P}\text{-aligned distribution } Q \text{ on } \mathcal{H}^S, \\ |\mathcal{M}^D_{Q^2} - \mathcal{M}^S_{Q^2}| \leq \dfrac{2B^2\sqrt{\dfrac{2|\mathbf{j}^{\max}|}{B^2\delta} + \ln\left(\dfrac{2\sqrt{m}}{\delta}\right)}}{\sqrt{2(m-2|\mathbf{j}^{\max}|)}}\end{array}\right) \geq 1-\delta$$

□

Références

[Mau04] A. Maurer. A note on the pac bayesian theorem. *CoRR*, cs.LG/0411099, 2004.

[McA03] D.A. McAllester. Simplified PAC-bayesian margin bounds. In *COLT*, 2003.