***SUPPLEMENTARY MATERIAL FOR*:**
NON-ASYMPTOTIC BOUNDS FOR VECTOR QUANTIZATION

By Clément Levrard

Université Paris Sud, UPMC and INRIA

APPENDIX A: SUPPLEMENTARY PROOFS FOR PROPOSITIONS 2.1 AND 2.2

A.1. Proof of Lemma 4.1

Since $\mathcal{H}$ is reflexive, according to the Banach-Alaoglu-Bourbaki theorem (see, e.g., Theorem 3.16 in [1]), combined with Tychonoff's theorem (see, e.g., Theorem 2.2.8 in [2]), $\mathcal{B}(0,R)^k$ is a compact subset of $\mathcal{H}^k$ for the weak topology. This proves i).

Let x be a fixed element of $\mathcal{H}$. Since $\mathbf{c} \mapsto \|x-c_i\|^2$ is weakly lower semi-continuous (see, e.g., Proposition 3.13 in [1]), $\mathbf{c} \mapsto \gamma(\mathbf{c},x)$ is weakly lower semi-continuous over $\mathcal{B}(0,M)^k$. Let $\mathbf{c}_n$ be a sequence of elements in $\mathcal{B}(0,M)^k$ such that $\mathbf{c}_n \rightharpoonup_{n\to\infty} \mathbf{c}$, for the weak topology, for some $\mathbf{c} \in \mathcal{B}(0,M)^k$. Then

$$\gamma(\mathbf{c},x) \leq \liminf_{n\to\infty} \gamma(\mathbf{c}_n,x).$$

Applying Fatou's Lemma (see, e.g., Lemma 4.3.3 in [2]) yields

$$R(\mathbf{c}) \leq \liminf_{n\to\infty} R(\mathbf{c}_n).$$

Hence ii) is proved. It is worth noting that this proves the existence of optimal codebooks for bounded distributions, and that $\mathcal{M}$ is closed for the weak topology. According to the centroid condition, $\mathcal{M} \subset \mathcal{B}(0,M)^k$. Thus, i) ensures that $\mathcal{M}$ is compact.

A.2. Proof of Lemma 4.2

Let x be in $V_i(\mathbf{c}^*) \cap V_j(\mathbf{c}) \cap \mathcal{B}(0,M)$, then $\|x-c_j\|^2 \leq \|x-c_i\|^2$, which leads to $\left\langle c_i - c_j, x - \frac{c_i+c_j}{2} \right\rangle \leq 0$. Since $\|x-c_i^*\| \leq \|x-c_j^*\|$, we may write

$$\|x-c_i\| \leq \|x-c_j\| + \|c_i-c_i^*\| + \|c_j-c_j^*\|.$$

Taking square on both sides leads to

$$\begin{aligned}\|x-c_i\|^2-\|x-c_j\|^2 \leq\; & 2\|x-c_j\|(\|c_i-c_i^*\|+\|c_j-c_j^*\|) \\ & + \left(\|c_i-c_i^*\|+\|c_j-c_j^*\|\right)^2 \\ \leq\; & 8M(\|c_i-c_i^*\|+\|c_j-c_j^*\|) \\ \leq\; & 8\sqrt{2}M\|\mathbf{c}-\mathbf{c}^*\|.\end{aligned}$$

Since $\|x-c_i\|^2-\|x-c_j\|^2 = -2\left\langle x-\frac{c_i+c_j}{2}, c_i-c_j\right\rangle$, (11) is proved.

To prove (12), remark that, since $x \in V_i(\mathbf{c}^*)$, $d(x, \partial V_i(\mathbf{c}^*)) \leq d(x, h^*_{i,j})$, where $h^*_{i,j}$ is the hyperplane defined by $\left\{x \in \mathcal{B}(0,M) | \|x-c_i^*\| = \|x-c_j^*\|\right\}$. Using quite simple geometric arguments, we deduce that

$$d(x, h^*_{i,j}) = \left|\left\langle x-\frac{c_i^*+c_j^*}{2}, \frac{c_i^*-c_j^*}{\left\|c_i^*-c_j^*\right\|}\right\rangle\right|.$$

The same arguments as in the proof of (11) guarantee that

$$\begin{aligned}\left|\left\langle x-\frac{c_i^*+c_j^*}{2}, \frac{c_i^*-c_j^*}{\left\|c_i^*-c_j^*\right\|}\right\rangle\right| &= \left\langle x-\frac{c_i^*+c_j^*}{2}, \frac{c_i^*-c_j^*}{\left\|c_i^*-c_j^*\right\|}\right\rangle \\ &\leq \frac{4\sqrt{2}M}{B}\|\mathbf{c}-\mathbf{c}^*\|.\end{aligned}$$

A.3. Proof of Lemma 4.3

Let $\tilde{\mathbf{c}}$ be a codebook satisfying the centroid condition, and such that there exists an optimal codebook $\mathbf{c}^*$ such that $\|\tilde{\mathbf{c}}-\mathbf{c}^*\| \leq Br_0/4\sqrt{2}M$. For $i = 1, \ldots, k$, we have

$$\begin{aligned}& P(x\mathbb{1}_{V_i(\tilde{\mathbf{c}})}(x)) - P(x\mathbb{1}_{V_i(\mathbf{c}^*)}(x)) \\ &\qquad = \sum_{j\neq i} P\left(x(\mathbb{1}_{V_i(\tilde{\mathbf{c}})\cap V_j(\mathbf{c}^*)}(x) - \mathbb{1}_{V_i(\mathbf{c}^*)\cap V_j(\tilde{\mathbf{c}})}(x))\right).\end{aligned}$$

Since $\mathbf{c}^*$ and $\tilde{\mathbf{c}}$ satisfy the centroid condition, it follows that

$$\begin{aligned}\sum_{i=1}^k \|P(V_i(\tilde{\mathbf{c}}))\tilde{c}_i - P(V_i(\mathbf{c}^*))c_i^*\| &= \sum_{i=1}^k \left\|P\left(x\mathbb{1}_{V_i(\tilde{\mathbf{c}})}(x)\right) - P(x\mathbb{1}_{V_i(\mathbf{c}^*)}(x))\right\| \\ &\leq 2M\sum_{i=1}^k\sum_{j\neq i} P(V_i(\mathbf{c}^*)\cap V_j(\tilde{\mathbf{c}})).\end{aligned}$$

According to (12) in Lemma 4.2, and to (3), we may write

$$\begin{aligned}\sum_{i=1}^{k}\sum_{j\neq l} P(V_i(\mathbf{c}^*)\cap V_j(\tilde{\mathbf{c}})) &\leq P\left(N_{\mathbf{c}^*}\left(\frac{4\sqrt{2}M}{B}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|\right)\right)\\ &\leq \frac{p_{min}}{16\sqrt{2}M}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|.\end{aligned}$$

This ensures that

$$\sum_{i=1}^{k}\|P(V_i(\tilde{\mathbf{c}}))\tilde{c}_i - P(V_i(\mathbf{c}^*))c_i^*\| \leq \frac{p_{min}}{8\sqrt{2}}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|.$$

For a fixed i, the triangle inequality yields

$$\begin{aligned}&\|P(V_i(\tilde{\mathbf{c}}))\tilde{c}_i - P(V_i(\mathbf{c}^*))c_i^*\|\\ &\qquad\geq P(V_i(\mathbf{c}^*))\|\tilde{c}_i - c_i^*\| - \|\tilde{c}_i\|\left|P(V_i(\mathbf{c}^*)) - P(V_i(\tilde{\mathbf{c}}))\right|.\end{aligned}$$

Since $B \leq 2M$ and $r_0 \leq 2M$, we have, for all $i = 1,\ldots,k$,

$$\|\tilde{c}_i\| \leq \|c_i^*\| + \|\tilde{\mathbf{c}}-\mathbf{c}^*\| \leq 2M.$$

Furthermore, using the same technique as above leads to

$$\begin{aligned}\sum_{i=1}^{k}\left|P(V_i(\mathbf{c}^*)) - P(V_i(\tilde{\mathbf{c}}))\right| &\leq 2\sum_{i=1}^{k}\sum_{j\neq l} P(V_i(\mathbf{c}^*)\cap V_j(\tilde{\mathbf{c}}))\\ &\leq \frac{p_{min}}{8\sqrt{2}M}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|.\end{aligned}$$

It follows that

$$\begin{aligned}\sum_{i=1}^{k}\|P(V_i(\tilde{\mathbf{c}}))\tilde{c}_i - P(V_i(\mathbf{c}^*))c_i^*\| &\geq p_{min}\sum_{i=1}^{k}\|\tilde{c}_i - c_i^*\| - \frac{p_{min}}{4\sqrt{2}}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|\\ &\geq \left(1-\frac{1}{4\sqrt{2}}\right)p_{min}\|\tilde{\mathbf{c}}-\mathbf{c}^*\|,\end{aligned}$$

which contradicts the upper bound, since $p_{min} > 0$.

A.4. Proof of Lemma 4.4

Let $\mathbf{c}$ be a codebook such that $c_i \notin \mathcal{B}(0, M)$, for some i. Denote by s the Hilbert space projection onto the closed convex set $\mathcal{B}(0, M + r)$ (see, e.g., Theorem 5.2 in [1]), and denote by $\bar{\mathbf{c}}$ the codebook such that $\bar{c}_j = c_j$ if $j \neq i$, and $\bar{c}_i = s(c_i)$. Theorem 5.2 in [1] ensures that

$$\|x - c_i\|^2 \geq \|x - s(c_i)\|^2 + \|c_i - s(c_i)\|^2,$$

for every x in $\mathcal{B}(0, M + r)$. Since P is M-bounded, it easily follows that $P\gamma(\bar{\mathbf{c}}, .) \leq P\gamma(\mathbf{c}, .)$, which in turn implies that

$$\inf_{\mathcal{H}^k \setminus \mathcal{B}^o(\mathcal{M},r)} P\gamma(\mathbf{c}, .) = \inf_{\mathcal{B}(0,M+r)^k \setminus \mathcal{B}^o(\mathcal{M},r)} P\gamma(\mathbf{c}, .).$$

Since $\mathcal{B}(0, M + r)^k \setminus \mathcal{B}^o(\mathcal{M}, r)$ is weakly compact, Lemma 4.1 provides $\mathbf{c}_r$ that achieves the infimum.

APPENDIX B: SUPPLEMENTARY PROOFS FOR PROPOSITION 3.1

B.1. Proof of Proposition 4.2

For simplicity, assume that σ is such that $\sigma_1 = \dots \sigma_{\frac{m}{2}} = +1$ and $\sigma_{\frac{m}{2}+1} = \dots = \sigma_m = -1$. Let $\mathcal{S}^-$ and $\mathcal{S}^+$ denote the sets of mistakes of σ', that is

$$\begin{cases} \mathcal{S}^- & = & \{i = 1, \dots, \frac{m}{2} | \quad \sigma'_i = -1\}, \\ \mathcal{S}^+ & = & \{i = \frac{m}{2} + 1, \dots, m | \quad \sigma'_i = +1\}. \end{cases}$$

Finally, let s^+ and s^- respectively denote $|\mathcal{S}^+|$ and $|\mathcal{S}^-|$. Since $\sum_{i=1}^m \sigma'_i = 0$, it is clear that $s^+ = s^- = s$.

Let $R_i(Q_{\sigma'})$ denote the contribution of $U_i \cup U'_i$ to the distortion, namely $R_i(Q_{\sigma'}) = P(\|x - Q_{\sigma'}(x)\|^2 \mathbb{1}_{U_i \cup U'_i}(x))$. Then, for i in $\mathcal{S}^-$, elementary calculations show that

$$R_i(Q_{\sigma'}) = R_i(Q_\sigma) + \frac{(1+\delta)\Delta^2}{4m}.$$

Symmetrically, for i in $\mathcal{S}^+$, we get

$$R_i(Q_{\sigma'}) = R_i(Q_\sigma) - \frac{(1-\delta)\Delta^2}{4m}.$$

Summing with respect to i and taking into account that $s^+ = s^- = s$ leads to

$$R(Q_{\sigma'}) = R(Q_\sigma) + s\frac{\Delta^2\delta}{2m}.$$

Since $s = \frac{\rho(\sigma,\sigma')}{4}$, (19) is proved.

The proof of the second part of Proposition 4.2 is based on elementary properties of distributions with finite supports, which are extended to the case where the source distributions are supported on small balls.

LEMMA B.1. *Let z_1 and z_2 be points in $\mathbb{R}^d$, denote by R the quantity $\|z_1 - z_2\|$ and by U_i the ball $\mathcal{B}(z_i, \rho)$. At last, let P denote the distribution with cone-shaped density*

$$\frac{2(d+1)}{V}\left(\mathbb{1}_{\|x-z_i\|\leq\rho}(\rho - \|x - z_i\|)\right),$$

over each ball U_i, where V denote the volume of the unit ball. Then, if

$$\rho \leq \frac{R}{10},$$

then the best 2-quantizer Q_2^* is such that $Q_2^*(U_i) = z_i$ for $i = 1, 2$. Furthermore, the best 1-quantizer Q_1^* is such that $Q_1^*(U_1 \cup U_2) = (z_1 + z_2)/2$.

PROOF OF LEMMA B.1. Proving that the best 1-quantizer is the quantizer with codebook $(z_1 + z_2)/2$ follows from straightforward calculation.

Let V_i denote the Voronoi cell associated with z_i in the Voronoi diagram generated by (z_1, z_2). Denote by Q_2^* the quantizer satisfying $Q_2^*(U_i) = z_i$, for $i = 1, 2$.

For any nearest neighbor quantizer Q, denote by $R_i(Q)$ the contribution of the cell i to the distortion of Q, that is $R_i(Q) = P(\|x - Q(x)\|^2 \mathbb{1}_{V_i}(x))$. Denote by V the volume of the unit ball, and by S its surface. Recalling that $S = d \times V$, an elementary calculation shows that

$$\begin{aligned} R_i(Q_2^*) &= \frac{1}{2}\frac{d+1}{\rho^{d+1}V}\int_0^\rho S(\rho r^{d+1} - r^{d+2})dr \\ &= \rho^2 \frac{d(d+1)}{2(d+2)(d+3)}. \end{aligned}$$

Let $m_i^{in} = |Q(U_i) \cap V_i|$ and $m_i^{out} = |Q(U_i) \cap V_i^c|$ denote the number of images of U_i sent inside and outside V_i. For a given i, there are three situations of interest, which are described below.

1. $m_i^{out} = 0$ and $m_i^{in} = 1$, then it is clear that $R_i(Q_2^*) \leq R_i(Q)$, with equality only if $Q(U_i) = z_i$.
2. $m_i^{out} = 0$ and $m_i^{in} = 2$, then $R_i(Q) \geq 0 = R_i(Q_2^*) - \rho^2 \frac{d(d+1)}{2(d+2)(d+3)}$.

3. $m_i^{out} \geq 1$, then there exists $z \in U_i$ such that $Q(z) \notin V_i$. Consequently, $\|z - Q(z)\| \geq \frac{R}{2} - \rho$. Let $z' \in U_i$, then

$$\|z' - Q(z')\| \geq \|z - Q(z')\| - 2\rho \geq \|z - Q(z)\| - 2\rho \geq \frac{R}{2} - 3\rho.$$

Hence we deduce

$$R_i(Q) \geq \frac{1}{2}(\frac{R}{2} - 3\rho)^2 = R_i(Q_2^*) + \frac{1}{2}\left((\frac{R}{2} - 3\rho)^2 - \rho^2 \frac{d(d+1)}{(d+2)(d+3)}\right).$$

Since Q is a 2-quantizer, it is easy to see that

$$|\left\{i; m_i^{in} \geq 2\right\}| \leq |\left\{i; m_i^{out} \geq 1\right\}|.$$

From this we deduce that

$$\begin{aligned} R(Q) = & \sum_{\{i; m_i^{in} \geq 2, m_i^{out} = 0\}} R_i(Q) + \sum_{\{i; m_i^{out} \geq 1\}} R_i(Q) + \sum_{\{i; m_i^{in} = 1, m_i^{out} = 0\}} R_i(Q) \\ \geq & \sum_{\substack{\{i; m_i^{in} \geq 2, m_i^{out} = 0\} \\ \cup \{i; m_i^{out} \geq 1\}}} R_i(Q_2^*) + \sum_{\{i; m_i^{in} = 1, m_i^{out} = 0\}} R_i(Q) \\ & + \frac{1}{2} |\left\{i; m_i^{out} \geq 1\right\}| \left(\left(\frac{R}{2} - 3\rho\right)^2 - 2\rho^2 \frac{d(d+1)}{(d+2)(d+3)}\right). \end{aligned}$$

Taking $\rho \leq \frac{R}{10}$ ensures that $(R/2 - 3\rho)^2 > 2\rho^2 \frac{d(d+1)}{(d+2)(d+3)}$. This proves that $R(Q) \leq R(Q_2^*)$, with equality only if, for $i = 1, 2$, $m_i^{out} = 0$ and $Q(U_i) = z_i$. □

Throughout the remainder of this subsection, a source distribution $P_{\sigma'}$ is fixed, so that $R(Q, P_{\sigma'})$ may be denoted by $R(Q)$. Taking $\rho = \frac{\Delta}{16}$ ensures that the conditions of Lemma B.1 are satisfied for $P_{\sigma' | U_i \cup U_i'}$. We turn now to the proof of Proposition 4.2.

Let Q be a nearest neighbor quantizer with k code points. The following construction provides Q_σ such that $R(Q_\sigma) \leq R(Q)$. Let V_i denote the union of the Voronoi cells associated with z_i and $z_i + \omega_i$, in the Voronoi diagram generated by the sequences z and ω. We adopt the following notation.

$$\left\{ \begin{array}{rcl} n_i(Q) & = & |Q(\mathcal{B}(0, M)) \cap V_i|, \\ n_i^{out}(Q) & = & |Q(V_i) \cap V_i^c|, \\ I_j(Q) & = & \{i; n_i(Q) = j\}, \\ i_j(Q) & = & |I_j(Q)|, \\ i_{\geq j}(Q) & = & \sum_{i \geq j} i_j(Q). \end{array} \right.$$

The first step is to add code points to empty cells. From the k-quantizer Q, a quantizer Q_1 is built as follows.

- If $n_i(Q) \geq 1$, then we take $Q_{1|V_i} \equiv Q_{|V_i}$.
- If $n_i(Q) = 0$, then we set $Q_1(U_i) = Q_1(U_i') = z_i + \frac{w_i}{2}$.

Note that Q_1 is a $(k + i_0(Q))$-quantizer. Let us denote $k_1 = k + i_0$ and $p_\pm = \frac{1 \pm \delta}{2m}$. Then $R(Q_1)$ can be bounded as follows.

Let i be an integer between 1 and m. We denote by $R_i(Q)$ the contribution of V_i to the risk $R(Q)$. If $i \in I_{\geq 1}$, then $R_i(Q) = R_i(Q_1)$. Otherwise, if $i \in I_0(Q)$, then

$$R_i(Q_1) = 2p_\pm \rho^2 \frac{d(d+1)}{(d+2)(d+3)} + p_\pm \frac{\Delta^2}{2}.$$

Furthermore, if $i \in I_0$, then $n_i^{out}(Q) \geq 1$, which ensures that, as in the proof of Lemma B.1,

$$R_i(Q) \geq p_\pm \left(2\Delta - 3\rho\right)^2 \geq 3p_\pm \Delta^2,$$

since $\rho = \frac{\Delta}{16}$. Thus we may write

$$\begin{aligned} R_i(Q) - R_i(Q_1) &\geq p_\pm \left[3\Delta^2 - 2\rho^2 - \frac{\Delta^2}{2}\right] \\ &\geq 2p_- \Delta^2. \end{aligned}$$

Summing all the contributions of the V_i's leads to

$$R(Q_1) \leq R(Q) - 2i_0(Q)p_- \Delta^2.$$

Next, we build the quantizer Q_2 according to the following rule:

- If $n_i(Q_1) \geq 2$, then $Q_2(U_i) = z_i$ and $Q_2(U_i') = z_i + w_i$.
- If $n_i(Q_1) = 1$, then $Q_2(U_i) = Q_2(U_i') = z_i + \frac{w_i}{2}$.

Since for $i = 1, \ldots, m$, $n_i(Q_1) \geq 1$, Q_2 is defined on every V_i. Notice that, since $I_j(Q_1) = I_j(Q)$ for $j \geq 2$, Q_2 has $k_2 = k + i_0(Q) - \sum_{p \geq 3} (p-2) i_p(Q)$ code points. The following Lemma gives a relation between $R(Q_2)$ and $R(Q_1)$.

Lemma B.2. *One has*

$$R(Q_2) \leq R(Q_1) + i_{\geq 3}(Q) \frac{p_+ \Delta^2}{128}.$$

Proof of Lemma B.2. Let i be an integer between 1 and m. Several cases may occur, as described below.

- Assume that $n_i(Q_1) = 1$.
  - If $n_i^{out}(Q_1) = 0$, then $R_i(Q_1) \geq R_i(Q_2)$, according to Lemma B.1.
  - If $n_i^{out}(Q_1) \geq 1$, then, using the same technique as employed to bound $R(Q_1)$ from above, $R_i(Q_1) - R_i(Q_2) \geq 2p_{\pm}\Delta^2$, which ensures that $R_i(Q_1) \geq R_i(Q_2)$.
- Assume that $n_i(Q_1) = 2$.
  - If $n_i^{out}(Q_1) = 0$, then $R_i(Q_1) \geq R_i(Q_2)$, according to Lemma B.1.
  - If $n_i^{out}(Q_1) \geq 1$, then, since $R_i(Q_2) = 2p_{\pm}\frac{\rho^2 d(d+1)}{(d+2)(d+3)} \leq p_{\pm}\frac{\Delta^2}{128}$, $R_i(Q_1) - R_i(Q_2) \geq 2\Delta^2 \geq 0$.
- At last, assume that $n_i(Q_i) \geq 3$. If $n_i^{out}(Q_1) \geq 1$, then $R_i(Q_1) \geq R_i(Q_2)$. If $n_i^{out}(Q_1) = 0$, then $R_i(Q_1) \geq 0 \geq R_i(Q_2) - p_{\pm}\frac{\Delta^2}{128}$. In both cases $R_i(Q_2) \leq R_i(Q_1) + p_{+}\frac{\Delta^2}{128}$.

Noticing that $I_{\geq 3}(Q_1) = I_{\geq 3}(Q)$, and summing the contributions $R_i(Q_2)$ leads to the result. □

The last step is to build a quantizer Q_σ from Q_2 with exactly k code points.

- If $k_2 = k$, set $Q_\sigma = Q_2$.
- If $k_2 < k$, choose $(k - k_2)$ cells V_i such that $n_i(Q_2) = 1$ (elementary calculation shows that there exist at least $k - k_2$ such V_i's). For every such V_i, set $Q_\sigma(U_i) = z_i$ and $Q_\sigma(U_i') = z_i + \omega_i$. Then

$$R(Q_\sigma) \leq R(Q_2) - (k - k_2)p_{-}\frac{\Delta^2}{2}.$$

- If $k_2 > k$, choose $(k_2 - k)$ cells V_i such that $n_i(Q_2) = 2$. For every such V_i, define $Q_\sigma(U_i) = Q_\sigma(U_i') = z_i + \frac{\omega_i}{2}$. Then

$$R(Q_\sigma) \leq R(Q_2) + (k_2 - k)p_{+}\frac{\Delta^2}{2}.$$

In both cases, Q_σ has exactly k code points. Finally, a result on the risk of Q_σ is given by the following proposition.

PROPOSITION B.1. *Let Q be a nearest neighbor quantizer and Q_σ be built as mentioned above. Then,*

$$R(Q_\sigma) \leq R(Q).$$

Moreover, if $Q \neq Q_\sigma$, then $R(Q) > R(Q_\sigma)$.

Proof of Proposition B.1. Since $\delta \leq \frac{1}{3}$, easy calculation ensures that $1 - \frac{p_-}{p_+} \leq \frac{1}{2}$.

Suppose that $k_2 \leq k$. Comparing the risk of Q to the risks of Q_1, Q_2 and Q_σ leads to

$$R(Q_\sigma) \leq R(Q) - 2i_0 p_- \Delta^2 + (i_0 + 2i_{\geq 3} - \sum_{p \geq 3} p i_p) p_- \frac{\Delta^2}{2} + i_{\geq 3} p_+ \frac{\Delta^2}{128}.$$

Since $\sum_{p\geq 3} p i_p \geq 3 i_{\geq 3}$, it is clear that

$$\begin{aligned} R(Q_\sigma) &\leq R(Q) - \frac{3}{2} p_- i_0 \Delta^2 + \Delta^2 i_{\geq 3} (\frac{p_+}{128} - \frac{p_-}{2}) \\ &\leq R(Q). \end{aligned}$$

Next, suppose that $k_2 > k$. Then

$$\begin{aligned} R(Q_\sigma) &\leq R(Q) + \left(i_0 + 2i_{\geq 3} - \sum_{p \geq 3} p i_p \right) p_+ \frac{\Delta^2}{2} + i_{\geq 3} p_+ \frac{\Delta^2}{128} - 2 i_0 p_- \Delta^2 \\ &\leq R(Q) + i_0 \frac{\Delta^2}{2} (p_+ - 4p_-) + p_+ i_{\geq 3} \Delta^2 (\frac{1}{128} - \frac{1}{2}), \end{aligned}$$

which yields $R(Q_\sigma) \leq R(Q)$.

If $i_0 > 0$ or $i_{\geq 3} > 0$, then the calculations above show that $R(Q_\sigma) < R(Q)$. If $i_0 = i_{\geq 3} = 0$, then, according to Lemma B.1, it is easy to see that $R(Q_\sigma) < R(Q)$ if $Q_\sigma \neq Q$. Now let τ be in $\{-1,1\}^{\frac{m}{2}}$ such that $\rho(\sigma(\tau), \sigma) = \min_{\tau'} \rho(\sigma(\tau'), \sigma)$. Recalling that $R(Q_\sigma) = R(Q_\sigma, P_{\sigma(\tau')})$, and using (19), it follows that

$$\begin{aligned} R(Q_{\sigma(\tau)}, P_{\sigma(\tau')}) &= R(Q_{\sigma(\tau')}, P_{\sigma(\tau')}) + \frac{\Delta^2 \delta}{8m} \rho(\sigma(\tau), \sigma(\tau')) \\ &\leq 2R(Q_{\sigma(\tau')}, P_{\sigma(\tau')}) + \frac{\Delta^2 \delta}{8m} \left(\rho(\sigma(\tau), \sigma) + \rho(\sigma, \sigma(\tau')) \right) \\ &\leq 2R(Q_\sigma, P_\sigma) + 2\frac{\Delta^2 \delta}{8m} \rho(\sigma(\tau'), \sigma) \\ &\leq 2R(Q_\sigma, P_{\sigma(\tau')}). \end{aligned}$$

□

B.2. Proof of Lemma 4.5

Let us introduce, for distributions P and Q with densities f and g the affinity

$$\alpha(P, Q) = \int \sqrt{fg(x)} d\lambda(x),$$

so that $H^2(P,Q) = 2(1-\alpha(P,Q))$. Elementary calculation shows that, if $\rho(\sigma,\sigma') = 4$, then

$$\alpha(P_\sigma, P_{\sigma'}) = 1 + \frac{2}{m}\left(\sqrt{1-\delta^2} - 1\right) \geq 1 - \frac{2\delta^2}{m}.$$

Hence we deduce

$$\begin{aligned} H^2(P_\sigma^{\otimes n}, P_{\sigma'}^{\otimes n}) &= 2(1-\alpha(P_\sigma^{\otimes n}, P_{\sigma'}^{\otimes n})) \\ &= 2(1-\alpha^n(P_\sigma, P_{\sigma'})) \\ &\leq \frac{4n\delta^2}{m}. \end{aligned}$$

Finally, we note that $\rho(\tau,\tau') = 2$ implies $\rho(\sigma(\tau),\sigma(\tau')) = 4$, for τ, τ' in $\{-1,+1\}^{\frac{m}{2}}$. This gives the result.

REFERENCES

[1] Brezis, H. (2011). *Functional analysis, Sobolev spaces and partial differential equations. Universitext*. Springer, New York. MR2759829 (2012a:35002)

[2] Dudley, R. M. (2002). *Real analysis and probability. Cambridge Studies in Advanced Mathematics* **74**. Cambridge University Press, Cambridge Revised reprint of the 1989 original. MR1932358 (2003h:60001)