Appendix

Proof of Lemma 2 (generalized MD guarantee). Note that for any $\mathbf{w}^\star \in \mathcal{W}$,

$$
\begin{aligned}
\eta\left(\sum_{t=1}^n f_t(\mathbf{w}_t) - \sum_{t=1}^n f_t(\mathbf{w}^\star)\right) &\le \sum_{t=1}^n \langle \eta \nabla f_t(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}^\star\rangle \\
&= \sum_{t=1}^n \left(\left\langle \eta \nabla f_t(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}'_{t+1}\right\rangle + \left\langle \nabla f_t(\mathbf{w}_t), \mathbf{w}'_{t+1} - \mathbf{w}^\star\right\rangle\right) \\
&= \sum_{t=1}^n \left(\left\langle \eta \nabla f_t(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}'_{t+1}\right\rangle + \left\langle \nabla \Psi(\mathbf{w}_t) - \nabla\Psi(\mathbf{w}'_{t+1}), \mathbf{w}'_{t+1} - \mathbf{w}^\star\right\rangle\right) \\
&\le \sum_{t=1}^n \left(\|\eta \nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}} \left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star} + \left\langle \nabla \Psi(\mathbf{w}_t) - \nabla\Psi(\mathbf{w}'_{t+1}), \mathbf{w}'_{t+1} - \mathbf{w}^\star\right\rangle\right) \\
&\le \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q + \left\langle \nabla \Psi(\mathbf{w}_t) - \nabla\Psi(\mathbf{w}'_{t+1}), \mathbf{w}'_{t+1} - \mathbf{w}^\star\right\rangle\right)
\end{aligned}
$$

Using simple manipulation we can show that

$$\langle \nabla\Psi(\mathbf{w}_t) - \nabla\Psi(\mathbf{w}_{t+1}), \mathbf{w}_{t+1} - \mathbf{w}^\star\rangle = \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_t) - \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_{t+1}) - \Delta_\Psi(\mathbf{w}_{t+1}|\mathbf{w}_t)$$

where given any $\mathbf{w}, \mathbf{w}' \in \mathcal{B}$,

$$\Delta_\Psi(\mathbf{w}|\mathbf{w}') := \Psi(\mathbf{w}) - \Psi(\mathbf{w}') - \langle \nabla\Psi(\mathbf{w}'), \mathbf{w} - \mathbf{w}'\rangle$$

is the Bregman divergence between $\mathbf{w}$ and $\mathbf{w}'$ w.r.t. function Ψ. Hence,

$$
\begin{aligned}
&\eta\left(\sum_{t=1}^n f_t(\mathbf{w}_t) - \sum_{t=1}^n f_t(\mathbf{w}^\star)\right) \\
&\quad\le \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q + \left\langle \nabla \Psi(\mathbf{w}_t) - \nabla\Psi(\mathbf{w}'_{t+1}), \mathbf{w}'_{t+1} - \mathbf{w}^\star\right\rangle\right) \\
&\quad= \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q + \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_t) - \Delta_\Psi\left(\mathbf{w}^\star\middle|\mathbf{w}'_{t+1}\right) - \Delta_\Psi\left(\mathbf{w}'_{t+1}\middle|\mathbf{w}_t\right)\right) \\
&\quad\le \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q + \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_t) - \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_{t+1}) - \Delta_\Psi\left(\mathbf{w}'_{t+1}\middle|\mathbf{w}_t\right)\right) \\
&\quad= \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q - \Delta_\Psi\left(\mathbf{w}'_{t+1}\middle|\mathbf{w}_t\right)\right) + \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_1) - \Delta_\Psi(\mathbf{w}^\star|\mathbf{w}_{n+1}) \\
&\quad\le \sum_{t=1}^n \left(\frac{\eta^p}{p}\|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{1}{q}\left\|\mathbf{w}_t - \mathbf{w}'_{t+1}\right\|_{\mathcal{X}^\star}^q - \Delta_\Psi\left(\mathbf{w}'_{t+1}\middle|\mathbf{w}_t\right)\right) + \Psi(\mathbf{w}^\star)
\end{aligned}
$$

Now since Ψ is q-uniformly convex w.r.t. $\|\cdot\|_{\mathcal{X}^\star}$, for any $\mathbf{w}, \mathbf{w}' \in \mathcal{B}^\star$, $\Delta_\Psi(\mathbf{w}'|\mathbf{w}) \ge \frac{1}{q}\|\mathbf{w} - \mathbf{w}'\|_{\mathcal{X}^\star}^q$. Hence we conclude that

$$
\begin{aligned}
\sum_{t=1}^n f_t(\mathbf{w}_t) - \sum_{t=1}^n f_t(\mathbf{w}^\star) &\le \frac{\eta^{p-1}}{p}\sum_{t=1}^n \|\nabla f_t(\mathbf{w}_t)\|_{\mathcal{X}}^p + \frac{\Psi(\mathbf{w}^\star)}{\eta} \\
&\le \frac{\eta^{p-1}Bn}{p} + \frac{\sup_{\mathbf{w}\in\mathcal{W}}\Psi(\mathbf{w})}{\eta} \\
&\le \frac{\eta^{p-1}Bn}{p} + \frac{\sup_{\mathbf{w}\in\mathcal{W}}\Psi(\mathbf{w})}{\eta}
\end{aligned}
$$

Plugging in the value of $\eta = \left(\frac{\sup_{\mathbf{w}\in\mathcal{W}} \Psi(\mathbf{w})}{nB}\right)^{1/p}$ we get :

$$\sum_{t=1}^{n} f_t(\mathbf{w}_t) - \sum_{t=1}^{n} f_t(\mathbf{w}^\star) \leq 2\left(\sup_{\mathbf{w}\in\mathcal{W}} \Psi(\mathbf{w})\right)^{1/q} (Bn)^{1/p}$$

dividing throughout by n conclude the proof. $\square$

Lemma 10. *Let $1 < p \leq 2$ and $C > 0$ be fixed constants, the following statements are equivalent :*

1. *For all sequence of mappings $(\mathbf{x}_n)_{n\geq 1}$ where each $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ and any $\mathbf{x}_0 \in \mathcal{B}^\star$:*

$$\sup_n \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^{n} \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] \leq C^p \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{n\geq 1} \mathbb{E}\left[\|\mathbf{x}_n(\epsilon)\|_{\mathcal{X}}^p\right]\right)$$

2. *There exist a non-negative convex function Ψ on $\mathcal{B}$ with $\Psi(0) = 0$, that is q-uniformly convex w.r.t. norm $\|\cdot\|_{\mathcal{X}^\star}$ and for any $\mathbf{w} \in \mathcal{B}$, $\frac{1}{q}\|\mathbf{w}\|_{\mathcal{X}^\star}^q \leq \Psi(\mathbf{w}) \leq \frac{C^q}{q}\|\mathbf{w}\|_{\mathcal{W}}^q$.*

Proof. For any $\mathbf{x} \in \mathcal{B}^\star$ define $\Psi^* : \mathcal{B}^\star \mapsto \mathbb{R}$ as

$$\Psi^*(\mathbf{x}) := \sup\left\{\left(\frac{1}{C^p}\sup_n \mathbb{E}\left[\left\|\mathbf{x} + \sum_{i=1}^{n} \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)\right\}$$

where the supremum is over sequence of mappings $(\mathbf{x}_n)_{n\geq 1}$ where each $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ and the sequence is such that, $\sup_n \mathbb{E}\left[\|\mathbf{x} + \sum_{i=1}^{n} \mathbf{x}_i\|_{\mathcal{W}^\star}^p\right] < \infty$. Since supremum of convex functions is a convex function, it is easily verified that $\Psi^*(\cdot)$ is convex. Note that by the definition of M-type in Equation 8, we have that for any $\mathbf{x}_0 \in \mathcal{B}^\star$, $\Psi^*(\mathbf{x}_0) \leq \|\mathbf{x}_0\|_{\mathcal{X}}^p$. On the other hand, note that by considering the sequence of constant mappings, $\mathbf{x}_i = 0$ for all $i \geq 1$, we get that for any $\mathbf{x}_0 \in \mathcal{B}^\star$,

$$\Psi^*(\mathbf{x}_0) = \sup\left\{\left(\frac{1}{C^p}\sup_n \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^{n} \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)\right\} \geq \frac{1}{C^p}\|\mathbf{x}_0\|_{\mathcal{W}^\star}^p$$

Thus we can conclude that for any $\mathbf{x} \in \mathcal{B}^\star$, $\frac{1}{C^p}\|\mathbf{x}\|_{\mathcal{W}^\star}^p \leq \Psi^*(\mathbf{x}) \leq \|\mathbf{x}\|_{\mathcal{X}}^p$.

For any $\mathbf{x}_0, \mathbf{y}_0 \in \mathcal{B}^\star$, by definition of $\Psi^*(\mathbf{x}_0)$ and $\Psi^*(\mathbf{y}_0)$, for any $\gamma > 0$, there exist sequences $(\mathbf{x}_n)_{n\geq 1}$ and $(\mathbf{y}_n)_{n\geq 1}$ s.t. :

$$\Psi^*(\mathbf{x}_0) \leq \left(\frac{1}{C^p}\sup_n \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^{n} \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) + \gamma$$

and

$$\Psi^*(\mathbf{y}_0^{(j)}) \leq \left(\frac{1}{C^p}\sup_n \mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i=1}^{n} \epsilon_i \mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\geq 1} \mathbb{E}\left[\|\mathbf{y}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) + \gamma$$

In fact in the above two inequalities if the supremum over n were achieved at some finite n_0, by replacing the original sequence by one which is identical up to n_0 and for any $i > n_0$ using $\mathbf{x}_i(\epsilon) = 0$ (and similarly $\mathbf{y}_i(\epsilon) = 0$), we can in fact conclude that using these $\mathbf{x}$'s and $\mathbf{y}$'s instead,

$$\Psi^*(\mathbf{x}_0) \leq \left(\frac{1}{C^p}\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i\geq 1} \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) + \gamma \tag{11}$$

and

$$\Psi^*(\mathbf{y}_0^{(j)}) \le \left(\frac{1}{C^p}\mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i \mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\ge 1}\mathbb{E}\left[\|\mathbf{y}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) + \gamma \tag{12}$$

Now consider a sequence formed by taking $\mathbf{z}_0 = \frac{\mathbf{x}_0+\mathbf{y}_0}{2}$ and further let

$$\mathbf{z}_1 = \left(\frac{1+\epsilon_0}{2}\right)\frac{\mathbf{x}_0-\mathbf{y}_0}{2} + \left(\frac{1-\epsilon_0}{2}\right)\frac{\mathbf{y}_0-\mathbf{x}_0}{2} = \epsilon_0(\mathbf{x}_0-\mathbf{y}_0)$$

and for any $i \ge 2$, define

$$\mathbf{z}_i = \left(\frac{1+\epsilon_0}{2}\right)\epsilon_{i-1}\mathbf{x}_{i-1}(\epsilon) + \left(\frac{1-\epsilon_0}{2}\right)\epsilon_{i-1}\mathbf{y}_{i-1}(\epsilon)$$

where $\epsilon_0 \in \{\pm 1\}$ is drawn uniformly at random. That is essentially at time $i=1$ we flip a coin and decide to go with the sequence $(\mathbf{x}_n)_{n\ge 0}$ with probability $1/2$ and $(\mathbf{y}_n)_{n\ge 0}$ with probability $1/2$. Clearly using the sequence $(\mathbf{z}_n)_{n\ge 1}$, we have that,

$$\begin{aligned}
\Psi^*\left(\frac{\mathbf{x}_0+\mathbf{y}_0}{2}\right) &= \sup_{(\mathbf{z})_{n\ge 1}}\left\{\left(\frac{1}{C^p}\sup_n \mathbb{E}\left[\left\|\frac{\mathbf{x}_0+\mathbf{y}_0}{2} + \sum_{i=1}^n \mathbf{z}_i(\epsilon_0,\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\ge 1}\mathbb{E}\left[\|\mathbf{z}_i(\epsilon_0,\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/p}\right\}^p \\
&\ge \frac{1}{C^p}\mathbb{E}\left[\left\|\mathbf{z}_0 + \sum_{i\ge 1}\mathbf{z}_i(\epsilon_0,\epsilon)\right\|_{\mathcal{W}^\star}^p\right] - \sum_{i\ge 1}\mathbb{E}\left[\|\mathbf{z}_i(\epsilon_0,\epsilon)\|_{\mathcal{X}}^p\right] \\
&= \frac{1}{C^p}\frac{\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] + \mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right]}{2} - \sum_{i\ge 1}\mathbb{E}\left[\|\mathbf{z}_i(\epsilon_0,\epsilon)\|_{\mathcal{X}}^p\right] \\
&= \frac{1}{C^p}\frac{\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] + \mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right]}{2} - \sum_{i\ge 1}\mathbb{E}\left[\|\mathbf{z}_i(\epsilon_0,\epsilon)\|_{\mathcal{X}}^p\right] \\
&= \frac{1}{C^p}\frac{\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] + \mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right]}{2} - \left\|\frac{\mathbf{x}_0-\mathbf{y}_0}{2}\right\|_{\mathcal{X}}^p - \sum_{i\ge 2}\mathbb{E}\left[\|\mathbf{z}_i(\epsilon_0,\epsilon)\|_{\mathcal{X}}^p\right] \\
&= \frac{1}{C^p}\frac{\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] + \mathbb{E}\left[\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right]}{2} - \left\|\frac{\mathbf{x}_0-\mathbf{y}_0}{2}\right\|_{\mathcal{X}}^p \\
&\quad - \sum_{i\ge 1}\frac{\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right] + \mathbb{E}\left[\|\mathbf{y}_i(\epsilon)\|_{\mathcal{X}}^p\right]}{2} \\
&= \frac{\frac{1}{C^p}\mathbb{E}\left\|\mathbf{x}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p - \sum_{i\ge 1}\mathbb{E}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p + \frac{1}{C^p}\mathbb{E}\left\|\mathbf{y}_0 + \sum_{i\ge 1}\epsilon_i\mathbf{y}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p - \sum_{i\ge 1}\mathbb{E}\|\mathbf{y}_i(\epsilon)\|_{\mathcal{X}}^p}{2} \\
&\quad - \left\|\frac{\mathbf{x}_0-\mathbf{y}_0}{2}\right\|_{\mathcal{X}}^p \\
&\ge \frac{\Psi^*(\mathbf{x}_0) + \Psi^*(\mathbf{y}_0)}{2} - \left\|\frac{\mathbf{x}_0-\mathbf{y}_0}{2}\right\|_{\mathcal{X}}^p - \gamma
\end{aligned}$$

where the last step is obtained by using Equations 11 and 12. Since γ was arbitrary taking limit we conclude that for any $\mathbf{x}_0$ and $\mathbf{y}_0$,

$$\frac{\Psi^*(\mathbf{x}_0) + \Psi^*(\mathbf{y}_0)}{2} \le \Psi^*\left(\frac{\mathbf{x}_0+\mathbf{y}_0}{2}\right) + \left\|\frac{\mathbf{x}_0-\mathbf{y}_0}{2}\right\|_{\mathcal{X}}^p$$

Hence we have shown the existence of a convex function Ψ^* that is p-uniformly smooth w.r.t. norm $\|\cdot\|_{\mathcal{X}}$ such that $\frac{1}{C^p}\|\cdot\|_{\mathcal{W}^\star}^p \le \Psi^*(\cdot) \le \|\cdot\|_{\mathcal{X}}^p$. Using convex duality we can conclude that the convex conjugate Ψ of function Ψ^*, is q-uniformly convex w.r.t. norm $\|\cdot\|_{\mathcal{X}^\star}$ and is such that $\|\cdot\|_{\mathcal{X}}^q \le \Psi(\cdot) \le C^q\|\cdot\|_{\mathcal{W}}^q$. That 2 implies 1 can be easily verified using the smoothness property of Ψ^*.

□

The following sequence of four lemma's give us the essentials towards proving Lemma 5. They use similar techniques as in [16].

Lemma 11. *Let $1 < r \le 2$. If there exists a constant $D > 0$ such that any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\ge1}$, where $\mathbf{x}_n : \{\pm1\}^{n-1} \mapsto \mathcal{B}^\star$ satisfy :*

$$\forall n \in \mathbb{N}, \quad \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] \le D(n+1)^{1/r} \sup_{0\le i\le n}\sup_{\epsilon}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}$$

then for all $p < r$ and $\alpha_p = \frac{20D}{r-p}$ we can conclude that any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\ge1}$, where $\mathbf{x}_n : \{\pm1\}^{n-1} \mapsto \mathcal{B}^\star$ will satisfy :

$$\sup_n \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] \le \alpha_p \sup_{\epsilon}\left(\sum_{i\ge0}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right)^{1/p}$$

Proof. To begin with note that in the definition of type, if the supremum over n were achieved at some finite n_0, then by replacing the original sequence by one which is identical up to n_0 and then on for any $i > n_0$ using $\mathbf{x}_i(\epsilon) = 0$ would only tighten the inequality. Hence it suffices to only consider such sequences. Further to prove the statement we only need to consider finite such sequences (ie. sequences such that there exists some n so that for any $i > n$, $\mathbf{x}_i = 0$) and show that the inequality holds for every such n (every such sequence).

Restricting ourselves to such finite sequences, we now use the shorthand,
$S = \sup_\epsilon \left(\sum_{i=0}^n \|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right)^{1/p}$. Now define

$$\begin{aligned}
&I_k(\epsilon) = \left\{i \ge 0 \middle| \tfrac{S}{2^{(k+1)/p}} < \|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}} \le \tfrac{S}{2^{k/p}}\right\} \quad , \\
&T_0^{(k)}(\epsilon) = \inf\{i \in I_k(\epsilon)\} \ \text{ and} \\
&\forall m \in \mathbb{N},\ T_m^{(k)}(\epsilon) = \inf\{i > T_{m-1}^{(k)}(\epsilon), i \in I_k(\epsilon)\}
\end{aligned}$$

Note that for any $\epsilon \in \{\pm1\}^{\mathbb{N}}$,

$$S^p \ge \sum_{i\in I_k(\epsilon)}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p > \tfrac{S^p\,|I_k(\epsilon)|}{2^{(k+1)}}$$

and so we get that $\sup_\epsilon |I_k(\epsilon)| < 2^{k+1}$. From this we conclude that

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^n \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] &\le \sum_{k\ge0}\mathbb{E}\left[\left\|\sum_{i\in I_k(\epsilon)}\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] \\
&= \sum_{k\ge0}\mathbb{E}\left[\left\|\sum_{i\ge0}\mathbf{x}_{T_i^{(k)}(\epsilon)}\right\|_{\mathcal{W}^\star}\right] \\
&\le \sum_{k\ge0}\left(D\ \sup_{\epsilon}\{|I_k(\epsilon)|^{1/r}\}\sup_{\epsilon}\{\sup_{i\in I_k(\epsilon)}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}\}\right) \\
&\le \sum_{k\ge0}\left(D\ 2^{(k+1)/r}\sup_{\epsilon}\sup_{i\in I_k(\epsilon)}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X},\infty}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{k\geq 0}\left(D\,2^{(k+1)/r}\,2^{-k/p}S\right)\\
&= D\,2^{1/r}\sum_{k\geq 0}2^{k(\frac{1}{r}-\frac{1}{p})}\,S\\
&\leq \frac{2D}{1-2^{(\frac{1}{r}-\frac{1}{p})}}S\\
&\leq \frac{2D}{1-2^{-(r-p)/4}}S\\
&\leq \frac{12D}{r-p}S\\
&= \alpha_p \sup_{\epsilon}\left(\sum_{i=0}^{n}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right)^{1/p}
\end{aligned}
$$

□

Lemma 12. *Let $1 < r \leq 2$. If there exists a constant $D > 0$ such that any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\geq 1}$, where $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ satisfy :*

$$
\forall n \in \mathbb{N}, \quad \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^{n}\epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] \leq D(n+1)^{1/r}\sup_{0\leq i\leq n}\sup_{\epsilon}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}
$$

then for any $p < r$, any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any mapping $(\mathbf{x}_n)_{n\geq 1}$, where $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$:

$$
\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^{n}\epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > c\right) \leq 2\left(\frac{\alpha_p}{c}\right)^{p/(p+1)}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/(p+1)}
$$

Proof. For any $\mathbf{x}_0 \in \mathcal{B}^\star$ and sequence $(\mathbf{x}_n)_{n\geq 1}$ define

$$
V_n(\epsilon) = \sum_{i=0}^{n}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p
$$

For appropriate choice of $a > 0$ to be fixed later, define stopping time

$$
\tau(\epsilon) = \inf\{n \geq 0 | V_{n+1} > a^p\}
$$

Now for any $c > 0$ we have,

$$
\begin{aligned}
\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^{n}\epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > c\right) &\leq \mathbb{P}(\tau(\epsilon) < \infty) + \mathbb{P}\left(\tau(\epsilon) = \infty, \sup_n \left\|\sum_{i=0}^{n}\epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > c\right)\\
&\leq \mathbb{P}(\tau(\epsilon) < \infty) + \mathbb{P}\left(\tau(\epsilon) > 0,\ \sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > c\right) \qquad (13)
\end{aligned}
$$

As for the first term in the above equation note that

$$
\mathbb{P}(\tau(\epsilon) < \infty) = \mathbb{P}(\sup_n V_n > a^p) \leq \frac{\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]}{a^p} \qquad (14)
$$

To consider the second term of Equation 13 we note that $\left(\mathbb{1}_{\{\tau(\epsilon)>0\}}(\mathbf{x}_0 + \sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i \mathbf{x}_i(\epsilon))\right)_{n\geq 0}$ is a valid martingale (stopped process) and hence, $\left(\left\|\mathbb{1}_{\{\tau(\epsilon)>0\}}(\mathbf{x}_0 + \sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i \mathbf{x}_i(\epsilon))\right\|_{\mathcal{W}^\star}\right)_{n\geq 0}$ is a sub-matingale. Hence by Doob's inequality we conclude that,

$$\mathbb{P}\left(T>0,\ \sup_n \left\|\mathbf{x}_0+\sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}>c\right)\leq\frac{1}{c}\sup_n\mathbb{E}\left[\left\|\mathbb{1}_{\{\tau(\epsilon)>0\}}\left(\mathbf{x}_0+\sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i\mathbf{x}_i(\epsilon)\right)\right\|_{\mathcal{W}^\star}\right]$$

Applying conclusion of the previous lemma we get that

$$\begin{aligned}\mathbb{P}\left(T>0,\ \sup_n \left\|\mathbf{x}_0+\sum_{i=1}^{n\wedge\tau(\epsilon)}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}>c\right)&\leq\frac{\alpha_p}{c}\sup_\epsilon\left(\mathbb{1}_{\{\tau(\epsilon)>0\}}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i=1}^{\tau(\epsilon)}\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right)\right)^{1/p}\\&\leq\frac{\alpha_p}{c}(a^p)^{1/p}=\frac{\alpha_p\,a}{c}\end{aligned}$$

Plugging the above and Equation 14 into Equation 13 we conclude that:

$$\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0+\sum_{i=1}^{n}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}>c\right)\leq\frac{\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]}{a^p}+\frac{\alpha_p\,a}{c}$$

Using $a=\left(\frac{c}{\alpha_p}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)\right)^{1/(p+1)}$ we conclude that

$$\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0+\sum_{i=1}^{n}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}>c\right)\leq 2\left(\frac{\alpha_p}{c}\right)^{p/(p+1)}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/(p+1)}$$

This conclude the proof. □

Lemma 13. *Let $1<r\leq2$. If there exists a constant $D>0$ such that any $\mathbf{x}_0\in\mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\geq1}$, where $\mathbf{x}_n:\{\pm1\}^{n-1}\mapsto\mathcal{B}^\star$ satisfy :*

$$\forall n\in\mathbb{N},\quad \mathbb{E}\left[\left\|\mathbf{x}_0+\sum_{i=1}^{n}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right]\leq D(n+1)^{1/r}\sup_{0\leq i\leq n}\sup_\epsilon\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}$$

then for any $p<r$, any $\mathbf{x}_0\in\mathcal{B}^\star$ and any sequence $(\mathbf{x}_n)_{n\geq1}$ satisfies :

$$\begin{aligned}&\sup_{\lambda>0}\lambda^p\,\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0+\sum_{i=1}^{n}\epsilon_i\mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}>\lambda\right)\\&\quad\leq\max\left\{4^{\frac{p+1}{p}}\alpha_p\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{\frac{1}{p}},2^{2p+3}\log(2)\,\alpha_p^p\,\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)\right\}\end{aligned}$$

Proof. We shall use Proposition 8.53 of Pisier's notes which is restated below to prove this lemma. To this end consider any $\mathbf{x}_0\in\mathcal{B}^\star$ and any sequence $(\mathbf{x}_i)_{i\geq1}$. Given an $\epsilon\in\{\pm1\}^{\mathbb{N}}$, for any $j\in[M]$ and $i\in\mathbb{N}$ let $\epsilon_i^{(j)}=\epsilon_{(i-1)M+j}$. Let $\mathbf{z}_0=\mathbf{x}_0\,M^{-1/p}$ and define the sequence $(\mathbf{z}_i)_{i\geq1}$ as follows, for any $k\in\mathbb{N}$ given by $k=j+(i-1)M$ where $j\in[M]$ and $i\in\mathbb{N}$,

$$\mathbf{z}_k(\epsilon)=\mathbf{x}_i(\epsilon^{(j)})\,M^{-1/p}$$

Clearly,

$$\begin{aligned}\|\mathbf{z}_0\|_{\mathcal{X}}^p+\sum_{k\geq1}\mathbb{E}\left[\|\mathbf{z}_k(\epsilon)\|_{\mathcal{X}}^p\right]&=\|\mathbf{x}_0\|_{\mathcal{X}}^p+\frac{1}{M}\sum_{j=1}^{M}\sum_{k\geq1}\mathbb{E}\left[\left\|\mathbf{x}_k(\epsilon^{(j)})\right\|_{\mathcal{X}}^p\right]\\&=\|\mathbf{x}_0\|_{\mathcal{X}}^p+\sum_{i\geq1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\end{aligned}$$

By previous lemma we get that for any $c > 0$,

$$\mathbb{P}\left(\sup_n \left\|\mathbf{z}_0 + \sum_{i=1}^n \epsilon_i \mathbf{z}_i(\epsilon)\right\|_{\mathcal{W}^\star} > c\right) \le 2\left(\frac{\alpha_p}{c}\right)^{p/(p+1)} \left(\|\mathbf{z}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{z}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/(p+1)}$$
$$= 2\left(\frac{\alpha_p}{c}\right)^{p/(p+1)} \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/(p+1)}$$

Note that

$$\sup_n \left\|\mathbf{z}_0 + \sum_{i=1}^n \epsilon_i \mathbf{z}_i(\epsilon)\right\|_{\mathcal{W}^\star} = M^{-1/p} \sup_{j\in[M]} \sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i^{(j)} \mathbf{x}_i(\epsilon^{(j)})\right\|_{\mathcal{W}^\star}$$

Hence we conclude that

$$\mathbb{P}\left(\sup_{j\in[M]} M^{-1/p} \sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i^{(j)} \mathbf{x}_i(\epsilon^{(j)})\right\|_{\mathcal{W}^\star} > c\right) \le 2\left(\frac{\alpha_p}{c}\right)^{\frac{p}{(p+1)}} \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{\frac{1}{(p+1)}}$$

For any $j \in [M]$, defining $Z^{(j)} = \sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i^{(j)} \mathbf{x}_i(\epsilon^{(j)})\right\|_{\mathcal{W}^\star}$ and using Proposition 15 we conclude that for any $c > 0$,

$$\sup_{\lambda>0} \lambda^p\, \mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > \lambda\right)$$
$$\le \max\left\{c, 2c^p \log\left(\frac{1}{1 - 2\left(\frac{\alpha_p}{c}\right)^{\frac{p}{(p+1)}} \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{\frac{1}{(p+1)}}}\right)\right\}$$

Picking

$$c = 4^{\frac{p+1}{p}} \alpha_p \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{1/p}$$

we conclude that

$$\sup_{\lambda>0} \lambda^p\, \mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > \lambda\right)$$
$$\le \max\left\{4^{\frac{p+1}{p}} \alpha_p \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)^{\frac{1}{p}}, 2^{2p+3} \log(2)\, \alpha_p^p \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)\right\}$$

$\square$

Lemma 14. *Let $1 < r \le 2$. If there exists a constant $D > 0$ such that any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\ge 1}$, where $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ satisfy :*

$$\forall n \in \mathbb{N}, \quad \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}\right] \le D(n+1)^{1/r} \sup_{0\le i\le n} \sup_{\epsilon} \|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}$$

then for all $p < r$, we can conclude that any $\mathbf{x}_0 \in \mathcal{B}^\star$ and any sequence of mappings $(\mathbf{x}_n)_{n\ge 1}$ where each $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ will satisfy :

$$\sup_n \mathbb{E}\left[\left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] \le \left(\frac{1104\, D}{(r-p)^2}\right)^p \left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\ge 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)$$

That is the pair $(\mathcal{W}, \mathcal{X})$ is of martingale type p.

Proof. Given any $p < r$ pick $r > p' > p$, due to the homogeneity of the statement we need to prove, w.l.o.g. we can assume that

$$\|\mathbf{x}_0\|_{\mathcal{X}}^{p'} + \sum_{i \geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^{p'}\right] = 1$$

Hence by previous lemma, we can conclude that

$$\sup_{\lambda > 0} \lambda^{p'}\, \mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > \lambda\right) \leq p' 2^{2p'+3} \log(2)\, \alpha_{p'}^{p'} \leq (32\, \alpha_{p'})^{p'} \tag{15}$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] &\leq \inf_{a>0}\left\{a^{p'} + p\int_a^\infty \lambda^{p-1}\mathbb{P}\left(\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star} > \lambda\right) d\lambda\right\} \\
&\leq \inf_{a>0}\left\{a^p + p(32\, \alpha_{p'})^{p'} \int_a^\infty \lambda^{p-1-p'} d\lambda\right\} \\
&\leq \inf_{a>0}\left\{a^p + p(32\, \alpha_{p'})^{p'} \left[\frac{\lambda^{p-p'}}{p-p'}\right]_a^\infty\right\} \\
&\leq \inf_{a>0}\left\{a^p + (46\, \alpha_{p'})^{p'} \frac{a^{p-p'}}{p'-p}\right\} \\
&= 2\frac{(46\, \alpha_{p'})^p}{(p'-p)^{p/p'}} \leq 2\frac{(46\, \alpha_p)^p}{(p'-p)^{p/p'}}
\end{aligned}$$

Since $\|\mathbf{x}_0\|_{\mathcal{X}}^{p'} + \sum_{i \geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^{p'}\right] = 1$ and $p' > p$, we can conclude that $\|\mathbf{x}_0\|_{\mathcal{X}}^{p} + \sum_{i \geq 1} \mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^{p}\right] \geq 1$ and so

$$\begin{aligned}
\mathbb{E}\left[\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] &\leq 2\frac{(46\, \alpha_p)^p}{(p'-p)^{p/p'}}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) \\
&\leq 2\frac{(46\, \alpha_p)^p}{(p'-p)}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)
\end{aligned}$$

Since p' can be chosen arbitrarily close to r, taking the limit we can conclude that

$$\mathbb{E}\left[\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] \leq 2\frac{(46\, \alpha_p)^p}{(r-p)}\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)$$

Recalling that $\alpha_p = \frac{12D}{r-p}$ we conclude that

$$\begin{aligned}
\mathbb{E}\left[\sup_n \left\|\mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon)\right\|_{\mathcal{W}^\star}^p\right] &\leq \left(\frac{1104\, D}{(r-p)^{(p+1)/p}}\right)^p\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right) \\
&\leq \left(\frac{1104\, D}{(r-p)^2}\right)^p\left(\|\mathbf{x}_0\|_{\mathcal{X}}^p + \sum_{i\geq 1}\mathbb{E}\left[\|\mathbf{x}_i(\epsilon)\|_{\mathcal{X}}^p\right]\right)
\end{aligned}$$

This concludes the proof. □

We restate below a proposition from Pisier's note (in [17])

Proposition 15 (Proposition 8.53 of [17])**.** *Consider a random variable $Z \geq 0$ and a sequence $Z^{(1)}, Z^{(2)}, \ldots$ drawn iid from some distribution. For some $0 < p < \infty$, $0 < \delta < 1$ and $R > 0$,*

$$\sup_{M \geq 1} \mathbb{P}\left(\sup_{m \leq M} M^{-1/p} Z^{(m)} > R \right) \leq \delta \quad \Longrightarrow \quad \sup_{\lambda > 0} \lambda^p \, \mathbb{P}\left(Z > \lambda\right) \leq \max\left\{ R, 2R^p \log\left(\frac{1}{1-\delta} \right) \right\}$$

Proof of Lemma 5. By Theorem 4 and our assumption that $\mathcal{V}_n(\mathcal{W}, \mathcal{X}) \leq D n^{-(1-1/r)}$, we have that for any sequence $(\mathbf{x}_n)_{n \geq 1}$ such that $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{X}$ and any $n \geq 1$,

$$\mathbb{E}\left[\frac{1}{n} \left\| \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon) \right\|_{\mathcal{W}^\star} \right] \leq D n^{-\left(1 - \frac{1}{r}\right)}$$

Hence we can conclude for any sequence $(\mathbf{x}_n)_{n \geq 1}$ such that $\mathbf{x}_n : \{\pm 1\}^{n-1} \mapsto \mathcal{B}^\star$ and any $n \geq 1$,

$$\mathbb{E}\left[\left\| \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon) \right\|_{\mathcal{W}^\star} \right] \leq D n^{\frac{1}{r}} \sup_{1 \leq i \leq n} \sup_{\epsilon} \left\| \mathbf{x}_i(\epsilon) \right\|_{\mathcal{X}}$$

Hence for any $\mathbf{x}_0 \in \mathcal{B}^\star$, we have that

$$\begin{aligned}
\mathbb{E}\left[\left\| \mathbf{x}_0 + \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon) \right\|_{\mathcal{W}^\star} \right] &\leq \mathbb{E}\left[\left\| \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon) \right\|_{\mathcal{W}^\star} \right] + \left\| \mathbf{x}_0 \right\|_{\mathcal{W}^\star} \\
&\leq \mathbb{E}\left[\left\| \sum_{i=1}^n \epsilon_i \mathbf{x}_i(\epsilon) \right\|_{\mathcal{W}^\star} \right] + D \left\| \mathbf{x}_0 \right\|_{\mathcal{X}} \\
&\leq D n^{\frac{1}{r}} \sup_{1 \leq i \leq n} \sup_{\epsilon} \left\| \mathbf{x}_i(\epsilon) \right\|_{\mathcal{X}} + D \left\| \mathbf{x}_0 \right\|_{\mathcal{X}} \\
&\leq 2D(n+1)^{\frac{1}{r}} \sup_{0 \leq i \leq n} \sup_{\epsilon} \left\| \mathbf{x}_i(\epsilon) \right\|_{\mathcal{X}}
\end{aligned}$$

Now applying Lemma 14 completes the proof. □