A. Missing proofs from Section 3

A.1. Single-dimensional case

Proof of Claim 1. It is easy to see that the function is monotone by induction. Assuming that for any $j,k \le i-1$, if $\pi_j \le \pi_k$ then $\mathcal{M}_f^\pi(\pi_j) \le \mathcal{M}_f^\pi(\pi_k)$, we show that this property also holds for any $j,k \le i$. Indeed, by definition $\mathcal{M}_f^\pi(\pi_i) \le H_i \le \mathcal{M}_f^\pi(\pi_j)$ for any $\pi_j \ge \pi_i$ with $j < i$. Similarly, $\mathcal{M}_f^\pi(\pi_i) \ge L_i \ge \mathcal{M}_f^\pi(\pi_j)$ for any $\pi_j \le \pi_i$ with $j < i$.

To see that $\mathcal{M}_f^\pi(x) \le \max_{y\le x} f(y)$ for any $x \in [m]$, notice that if $f(\pi_i) \ge \mathcal{M}_f^\pi(\pi_i)$ this is trivially true. We thus only need to argue that this is true when $\mathcal{M}_f^\pi(\pi_i) = L_i$. In this case, there is some $j < i$ with $\pi_j < \pi_i$, such that $L_i = \mathcal{M}_f^\pi(\pi_j)$. Again by induction, $\mathcal{M}_f^\pi(\pi_j) \le \max_{y\le\pi_j} f(y)$ and thus $L_i \le \max_{y\le\pi_j} f(y) \le \max_{y\le\pi_i} f(y)$. ☐

Proof of Claim 2. We first argue that it suffices to show the statement for functions f that take values in $\{0,1\}$ instead of $[0,1]$.

For a function g, denote by $[g > t]$ the indicator function that $g(x) > t$. One can check that $[\mathcal{M}_f^\pi > t] = \mathcal{M}_{[f>t]}^\pi$ as the definition of $\mathcal{M}_f^\pi$ only involves comparisons. Since the value $\mathcal{M}_f^\pi(x) = \int_0^1 [\mathcal{M}_f^\pi > t](x)dt$, we get that $\mathcal{M}_f^\pi(x) = \int_0^1 \mathcal{M}_{[f>t]}^\pi(x)dt$.

We have $\mathbb{E}_{x\sim\mathbb{U}([m])}[f'(x)] = \mathbb{E}_{x\sim\mathbb{U}([m])}[\mathbb{E}_\pi[\mathcal{M}_f^\pi]] = \int_0^1 \mathbb{E}_{x\sim\mathbb{U}([m])}[\mathbb{E}_\pi[\mathcal{M}_{[f>t]}^\pi(x)]]dt$. Therefore if for any boolean-valued function $g : [m] \to \{0,1\}$ we show that $\mathbb{E}_{x\sim\mathbb{U}([m])}[\mathbb{E}_\pi[\mathcal{M}_g^\pi(x)]] = \mathbb{E}_{x\sim\mathbb{U}([m])}[g(x)]$ then we obtain the required statement, as we can use $g = [f > t]$ for any $t \in [0,1]$.

We move on to prove the statement for functions f taking values in $\{0,1\}$. The description of the constructed f' becomes much simpler in this case. Starting from the interval $\{1,...,m\}$ the algorithm first selects an element i uniformly at random. If $f(i) = 0$, it sets $f'(j) = 0$ for any $j \le i$ and recursively solves the problem in the interval $\{i+1,...,m\}$. If $f(i) = 1$, it sets $f'(j) = 1$ for any $j \ge i$ and recursively solves the problem in the interval $\{1,...,i-1\}$.

Let $\{l,...,r\}$ be the interval at the current iteration. Let $V(l,r)$ be the value of this set defined as

$$V(l,r) = \sum_{x=l}^{r} f(x) - \sum_{x=l}^{r} \mathbb{E}[f'(x)]$$

where the expectation is taken over the randomness of f' on the interval $\{l,...,r\}$.

We will prove by induction on $r-l$ that $V(l,r) = 0$. In the case that $r = l$ we will select the only point x_l with probability 1, so $V(l,r) = f(x) - f(x) = 0$. We assume that $V(l,r) = 0$ for any l,r with $r-l \le m-1$.

Let $y \sim U(\{l,...,r\})$ be the uniformly random chosen point from $\{l,...,r\}$. We distinguish two cases for $f(y)$. If $f(y) = 0$ we obtain

$$\begin{aligned}\sum_{x=l}^{r} f(x) - \sum_{x=l}^{r} \mathbb{E}[f'(x)|y] &= \sum_{x=l}^{y} f(x)\\ &= \sum_{x=l}^{r} \mathbb{1}_{\{f(x)=0,f(y)=1,x\ge y\}}\end{aligned}$$

where the first equality follows, since conditional on y, $f'(x) = 0$ for $x \in \{l,...,y\}$ and by the induction hypothesis $\sum_{x=y+1}^{r} f(x) - \sum_{x=y+1}^{r} \mathbb{E}[f'(x)|y] = 0$. Similarly, if $f(y) = 1$ we obtain

$$\begin{aligned}\sum_{x=l}^{r} f(x) - \sum_{x=l}^{r} \mathbf{E}[f'(x)|y] &= \sum_{x=y+1}^{r} (f(x)-1)\\ &= -\sum_{x=l}^{r} \mathbb{1}_{\{f(x)=1,f(y)=0,x\le y\}}\end{aligned}$$

Overall, we have

$$\begin{aligned}V(l,r) &= \frac{1}{r-l+1}\sum_{y=l}^{r}\left(\sum_{x=l}^{r} f(x) - \sum_{x=l}^{r}\mathbb{E}[f'(x)|y]\right)\\ &= \frac{1}{r-l+1}\sum_{y=l}^{r}\left(\sum_{x=l}^{r}\mathbb{1}_{\{f(x)=0,f(y)=1,x\ge y\}}\right.\\ &\qquad\left.-\sum_{x=l}^{r}\mathbb{1}_{\{f(x)=1,f(y)=0,x\le y\}}\right) = 0\end{aligned}$$

The intuition behind this fact is that the expected loss incurred by turning 1's into 0's is exactly balanced by the expected gain by turning 0's to 1's. ☐

Proof of Claim 3. Fix a point $x \in [m]$ and a random permutation π.

The oracle for $\mathcal{M}_f^\pi$ keeps track of an interval $\{l_i,...,r_i\}$ and makes a query only when the next point in the permutation lies in this interval. As the permutation is chosen uniformly at random, the next point is chosen uniformly in $\{l_i,...,r_i\}$ and lies in the smaller interval $\{\frac{3l_i+r_i}{4},...,\frac{l_i+3r_i}{4}\}$ with probability $1/2$. Every time this happens, the algorithm discards at least $\frac{r_i-l_i}{4}$ of the elements. As this shrinks the interval by a constant factor, it can happen at most $O(\log m)$ times. By Hoeffding's inequality, the probability that after $O(\log m) + \sqrt{O(\log m \cdot \log(1/\delta))}$ iterations the interval size is still greater than 1 is at most δ. Since $O(\log m) + \sqrt{O(\log m \cdot \log(1/\delta))} = O(\log(m/\delta))$ we get that the number of oracle queries is at most $O(\log(m/\delta))$

to evaluate $A'(x)$ with probability $1-\delta$. To get the required bound for every $x \in [m]$, we set $\delta \to \delta/m$ and take a union bound on the probabilities of error. This only increases the number of oracle queries by a constant factor, so the bound of $O(\log(m/\delta))$ is still accurate. □

A.2. Extending to many dimensions (general d)

To establish the result of Theorem 1, we now extend our construction to the more general case with $d \geq 1$. We apply our single-dimensional construction from Section 3.1 to fix monotonicity in each direction separately starting with the first.

We set $f_0 = f$. For every $i \in [d]$, based on the function f_{i-1} that is monotone in the first $i-1$ coordinates we obtain a function f_i that is monotone in the first i coordinates. To do this we apply our single dimensional construction at every single-dimensional slice $f_{i-1}(\cdot, x_{-i})$ of f_{i-1} for all choices $x_{-i} \in \mathbb{R}^d$ of the coordinates other than i. Importantly, we use the same randomness at every slice, for the choices of the points in the intervals $I_1, ..., I_m$ when performing the discretization to $m = \frac{1}{\varepsilon}$ points as well as for the chosen permutation π over the discrete domain $[m]$. This allows us to fix the monotonicity in coordinate i while preserving monotonicity in the first $i-1$ coordinates. It is easy to see that the discretization preserves the monotonicity. We now argue that using the same permutation for every slice also maintains the monotonicity. The following lemma shows that any two functions where one is smaller than the other, preserve the same ordering after their monotonization.

Lemma 1. *Let $f, g : [m] \to [0,1]$ such that $f(x) \leq g(x)$, for all $x \in [m]$. For any permutation π, it holds that $\mathcal{M}_f^\pi(x) \leq \mathcal{M}_g^\pi(x)$, for all $x \in [m]$.*

Proof. As argued in Claim 2, it suffices to show the statement for boolean valued functions $f, g : [m] \to \{0,1\}$. Let i be the first point where $\mathcal{M}_f^\pi(\pi_i) \neq \mathcal{M}_g^\pi(\pi_i)$. By the definition of $\mathcal{M}$, $H_i^{(g)} = H_i^{(f)}$ and $L_i^{(g)} = L_i^{(f)}$ and thus it must be that $\mathcal{M}_f^\pi(\pi_i) = 0$ and $\mathcal{M}_g^\pi(\pi_i) = 1$. By monotonicity $\mathcal{M}_f^\pi(x) = 0$ for all $x \leq \pi_i$ and $\mathcal{M}_g^\pi(x) = 1$ for all $x \geq \pi_i$. Thus, $\mathcal{M}_f^\pi(x) \leq \mathcal{M}_g^\pi(x)$, for all $x \in [m]$. □

This allows us to obtain a chain of oracles $f = f_0, f_1, ..., f_d = f'$ where f_i is monotone in the first i coordinates. Evaluating f_i requires only $O(\log \frac{d}{\varepsilon})$ queries to oracle f_{i-1} and gets error at most ε/d. Thus, to evaluate $f' = f_d$, at most $O\left(\log \frac{d}{\varepsilon}\right)^d$ queries to oracle f are required for error ε.

B. Missing proofs from Section 4

Proof of Claim 4. This claim follows as with high probability, the transformation $\mathcal{M}_{f^1_{S,T}}(T)$ cannot distinguish between $f^0_{S,T}$ and $f^1_{S,T}$. To see this, let $R_1, ..., R_q$ be the queries performed by $\mathcal{M}_{f^0_{S,T}}(T)$. If all of those queries satisfy $f^0_{S,T}(R_i) = f^1_{S,T}(R_i)$ then it must be that $\mathcal{M}_{f^1_{S,T}}(T) = \mathcal{M}_{f^0_{S,T}}(T)$.

Initially observe that for any R where either $R \not\subseteq T$ or $|R| < \frac{4d}{10}$ we have that $f^0_{S,T}(R) = f^1_{S,T}(R)$. On the contrary, given the set T, for any $R \subseteq T$ such that $|R| \geq \frac{4d}{10}$, the functions $f^1_{S,T}$ and $f^0_{S,T}$ differ only when $|R \setminus S| \leq d/10$, therefore

$$\mathbf{Pr}[f^0_{S,T}(R) \neq f^1_{S,T}(R)] = \mathbf{Pr}\left[|R \setminus S| \leq \frac{d}{10}\right]$$

Since S is a random subset of T excluding each element independently with probability $1/3$, we get that the expected value of $\mathbb{E}[|R \setminus S|] = |R|/3 \geq \frac{4d}{30}$. By the Hoeffding inequality we get

$$\Pr\left[|R \setminus S| \leq \mathbb{E}[|R \setminus S|] - \frac{d}{30}\right] \leq \exp\left\{-\frac{d}{450}\right\}$$

The claim then follows by a union bound on all q queries. □

Proof of Claim 5. Let $R_1, ..., R_q$ be the queries performed by $\mathcal{M}_{f^1_{S,T}}(S)$. If all of those queries satisfy $f^1_{S,T}(R_i) = f^1(R_i)$ then it must be that, $\mathcal{M}_{f^1_{S,T}}(S) = \mathcal{M}_{f^1}(S)$.

For any query R, in order for $f^1_{S,T}(R) \neq f^1(R)$, it must be that $R \subseteq T$, $|R| \geq \frac{4d}{10}$ and $|R \setminus S| > \frac{d}{10}$.

Given the set S, for any $|R| \geq \frac{4d}{10}$ and $|R \setminus S| > \frac{d}{10}$, we have $\Pr[R \subseteq T] = \Pr[(R \setminus S) \subseteq T] < 2^{-\frac{d}{10}}$ since given the set S, the set T is created by including each coordinate in $[d] \setminus S$ with probability $1/2$. The claim then again follows by a union bound on all q queries. □