Through all the proofs, we denote the set $\{1, \ldots, m\}$ by $[m]$.

A. Proof of Theorem 1

We first state a few properties of the function R:

Proposition 1. *For any vector $\mathbf{M}$ of N dimensions and integer r,*

Property 1. $R(\mathbf{M}, r) = a + R((M_1 - a, \ldots, M_N - a), r)$ *for any real number a and $r \geq 0$.*

Property 2. *$R(\mathbf{M}, r)$ is non-decreasing in M_i for each $i = 1, \ldots, N$.*

Property 3. *If $r > 0$, $R(\mathbf{M}, r) - R(\mathbf{M}, r-1) \leq 1/N$.*

Property 4. *If $r > 0$, and $P_i = \frac{1}{N} + R(\mathbf{M} + \mathbf{e}_i, r-1) - R(\mathbf{M}, r)$ for each $i = 1, \ldots, N$, then $\mathbf{P} = (P_1, \ldots, P_N)$ is a distribution in the simplex $\Delta(N)$.*

Proof of Proposition 1. We omit the proof for Property 1 and 2, since it is straightforward. We prove Property 3 by induction. For the base case $r = 1$, let $S = \{j : M_j = \min_i M_i\}$. If $|S| = 1$, then $R(\mathbf{M}+\mathbf{e}_i, 0)$ is $R(\mathbf{M}, 0)$ for $i \notin S$ and $R(\mathbf{M}, 0)+1$ otherwise. If $|S| > 1$, then $R(\mathbf{M}+\mathbf{e}_i, 0)$ is simply $R(\mathbf{M}, 0)$ for all i. In either case, we have

$$
\begin{aligned}
R(\mathbf{M}, 1) &= \frac{1}{N}\sum_{i=1}^{N} R(\mathbf{M} + \mathbf{e}_i, 0) \leq \frac{1}{N}(1 + \sum_{i=1}^{N} R(\mathbf{M}, 0)) \\
&= \frac{1}{N} + R(\mathbf{M}, 0),
\end{aligned}
$$

proving the base case. Now for $r > 1$, by definition of R and induction,

$$
\begin{aligned}
&R(\mathbf{M}, r) - R(\mathbf{M}, r-1) \\
=&\frac{1}{N}\sum_{j=1}^{N} (R(\mathbf{M} + \mathbf{e}_i, r-1) - R(\mathbf{M} + \mathbf{e}_i, r-2)) \\
\leq&\frac{1}{N}\sum_{j=1}^{N} \frac{1}{N} = \frac{1}{N},
\end{aligned}
$$

completing the induction. For Property 4, it suffices to prove $P_i \geq 0$ for each i and $\sum_{i=1}^{N} P_i = 1$. The first part can be shown using Property 2 and 3:

$$
\begin{aligned}
P_i &= \frac{1}{N} + R(\mathbf{M} + \mathbf{e}_i, r-1) - R(\mathbf{M}, r) \\
&\geq \frac{1}{N} + R(\mathbf{M}, r-1) - (\frac{1}{N} + R(\mathbf{M}, r-1)) = 0.
\end{aligned}
$$

The second part is also easy to show by definition of R:

$$
\begin{aligned}
\sum_{i=1}^{N} P_i &= 1 + \sum_{i=1}^{N} R(\mathbf{M} + \mathbf{e}_i, r-1) - NR(\mathbf{M}, r) \\
&= 1 + NR(\mathbf{M}, r) - NR(\mathbf{M}, r) = 1.
\end{aligned}
$$

□

Proof of Theorem 1. First inductively prove $V(\mathbf{M}, r) = r/N - R(\mathbf{M}, r)$ for any $r \geq 0$. The base case $r = 0$ is trivial by definition. For $r > 0$,

$$
\begin{aligned}
V(\mathbf{M}, r) &= \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} (\mathbf{P}\cdot\mathbf{Z} + V(\mathbf{M} + \mathbf{Z}, r-1)) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max_{i\in[N]} (P_i + V(\mathbf{M} + \mathbf{e}_i, r-1)) \\
&\qquad\qquad (\text{LS} = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max_{i\in[N]} \left(P_i + \frac{r-1}{N} - R(\mathbf{M} + \mathbf{e}_i, r-1)\right) \\
&\qquad\qquad (\text{by induction})
\end{aligned}
$$

Denote $P_i + (r-1)/N - R(\mathbf{M} + \mathbf{e}_i, r-1)$ by $g(\mathbf{P}, i)$. Notice that the average of $g(\mathbf{P}, i)$ over all i is irrelevant to $\mathbf{P}$: $\frac{1}{N}\sum_{i=1}^{N} g(\mathbf{P}, i) = r/N - R(\mathbf{M}, r)$. Therefore, $\max_i g(\mathbf{P}, i) \geq r/N - R(\mathbf{M}, r)$ for any $\mathbf{P}$, and

$$
V(\mathbf{M}, r) = \min_{\mathbf{P}} \max_{i} g(\mathbf{P}, i) \geq r/N - R(\mathbf{M}, r). \quad (11)
$$

On the other hand, from Proposition 1, we know that $P_i^* = 1/N + R(\mathbf{M} + \mathbf{e}_i, r-1) - R(\mathbf{M}, r)$ $(i \in [N])$ is a valid distribution. Also,

$$
\begin{aligned}
V(\mathbf{M}, r) &= \min_{\mathbf{P}} \max_{i} g(\mathbf{P}, i) \leq \max_{i} g(\mathbf{P}^*, i) \\
&= \max_{i} \left(\frac{r}{N} - R(\mathbf{M}, r)\right) = \frac{r}{N} - R(\mathbf{M}, r).
\end{aligned} \quad (12)
$$

So from Eq. (11) and (12) we have $V(\mathbf{M}, r) = r/N - R(\mathbf{M}, r)$, and also $P_i^* = 1/N + R(\mathbf{M} + \mathbf{e}_i, r-1) - R(\mathbf{M}, r) = V(\mathbf{M}, r) - V(\mathbf{M} + \mathbf{e}_i, r-1)$ realizes the minimum, and thus is the optimal strategy.

It remains to prove $V(\mathbf{0}, T) \leq c_N\sqrt{T}$. Let $\mathbf{Z}_1, \ldots, \mathbf{Z}_T$ be independent uniform random variables taking values in $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$. By what we proved above,

$$
V(\mathbf{0}, T) = \frac{T}{N} - \mathbb{E}[\min_{i\in[N]} \sum_{t=1}^{T} Z_{t,i}] = \mathbb{E}[\max_{i\in[N]} \sum_{t=1}^{T} (1/N - Z_{t,i})].
$$

Let $y_{t,i} = 1/N - Z_{t,i}$. Then each $y_{t,i}$ is a random variable that takes value $1/N$ with probability $1 - 1/N$ and $1/N - 1$ with probability $1/N$. Also, for a fixed i, $y_{1,i}, \ldots, y_{T,i}$ are independent (note that this is not true for $y_{t,1}, \ldots, y_{t,N}$ for a fixed t). It is shown in Lemma 3.3 of Berend & Kontorovich (2013) that each $y_{t,i}$ satisfies

$$
\mathbb{E}[\exp(\lambda y_{t,i})] \leq \exp(\frac{\lambda^2\sigma^2}{2}), \ \forall\lambda > 0,
$$

where $\sigma^2 = (N-1)/N^2$ is the variance of $y_{t,i}$. So if we let $Y_i = \sum_{t=1}^{T} y_{t,i}$, by the independence of each term, we have $\forall\lambda > 0$,

$$
\mathbb{E}[\exp(\lambda Y_i)] = \mathbb{E}[\prod_{t=1}^{T} \exp(\lambda y_{t,i})] = \prod_{t=1}^{T} \mathbb{E}[\exp(\lambda y_{t,i})]
$$

$$\leq \exp(\frac{\lambda^2\sigma^2 T}{2}).$$

Now using Lemma A.13 from Cesa-Bianchi & Lugosi (2006), we arrive at

$$\mathbb{E}[\max_{i\in[N]} Y_i] \leq \sigma\sqrt{2T\ln N} = c_N\sqrt{T}.$$

We conclude the proof by pointing out

$$V(\mathbf{0},T) = \mathbb{E}[\max_{i\in[N]}\sum_{t=1}^{T}(1/N - Z_{t,i})] = \mathbb{E}[\max_{i\in[N]} Y_i] \leq c_N\sqrt{T}.$$

□

As a direct corollary of Proposition 1 and Theorem 1, below we list a few properties of the function V for later use.

Proposition 2. *If* $\mathbf{LS} = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$*, then for any vector* $\mathbf{M}$ *and integer* r*,*

Property 5. $V(\mathbf{M}, r) = V((M_1 - a, \ldots, M_N - a), r) - a$ *for any real number* a *and* $r \geq 0$.

Property 6. $V(\mathbf{M}, r)$ *is non-increasing in* M_i *for each* $i = 1, \ldots, N$.

Property 7. $V(\mathbf{M}, r)$ *is non-decreasing in* r.

B. Proof of Theorem 2

Proof. Define $\bar{V}_t(\mathbf{M}) = \mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t]$ and $q(t', t) = \Pr[T = t'|T \geq t]$. We will prove an important property of the $\bar{V}$ function:

$$\bar{V}_t(\mathbf{M}) = \min_{\mathbf{P}\in\Delta(N)}\max_{i\in[N]}(P_i + q(t,t)V(\mathbf{M}+\mathbf{e}_i, 0) + (1-q(t,t))\,\bar{V}_{t+1}(\mathbf{M}+\mathbf{e}_i)). \tag{13}$$

This equation shows that $\bar{V}_t(M)$ is the conditional expectation of the regret given $T \geq t$, starting from cumulative loss vector $\mathbf{M}$ and assuming both the learner and the adversary are optimal. This is similar to the function V in the fixed horizon case, and again the value of the game $\inf_{\mathbf{Alg}}\sup_{\mathbf{Z}_{1:\infty}}\mathbb{E}_{T\sim Q}[\mathbf{Reg}(L_T, \mathbf{M}_T)]$ is simply $\bar{V}_1(\mathbf{0})$.

To prove Eq. (13), we plug the definition of $\bar{V}_{t+1}(\mathbf{M}+\mathbf{e}_i)$ into the right hand side and get $\min_{\mathbf{P}}\max_i g(\mathbf{P}, i)$ where $g(\mathbf{P}, i) = P_i + q(t,t)V(\mathbf{M}+\mathbf{e}_i, 0) + (1-q(t,t))\,\mathbb{E}[V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t+1]$. Using the fact that for any $t' \geq t+1$,

$$\begin{aligned}&(1-q(t,t))q(t', t+1)\\ =&\Pr[T > t|T \geq t]\Pr[T = t'|T \geq t+1]\\ =&\Pr[T = t'|T \geq t] = q(t', t),\end{aligned}$$

$g(\mathbf{P}, i)$ can be simplified in the following way:

$$\begin{aligned}&g(\mathbf{P}, i) \\ =&P_i + q(t,t)V(\mathbf{M}+\mathbf{e}_i, 0) + \\ &(1-q(t,t))\sum_{T=t+1}^{\infty}(q(T, t+1)V(\mathbf{M}+\mathbf{e}_i, T-t)) \\ =&P_i + q(t,t)V(\mathbf{M}+\mathbf{e}_i, 0) + \\ &\sum_{T=t+1}^{\infty}(q(T,t)V(\mathbf{M}+\mathbf{e}_i, T-t)) \\ =&P_i + \mathbb{E}[V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t].\end{aligned} \tag{14, 15}$$

Also, the average of $g(\mathbf{P}, i)$ over all i is independent of $\mathbf{P}$:

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N} g(\mathbf{P}, i) \\ =&\frac{1}{N} + \frac{1}{N}\sum_{i}^{N}\mathbb{E}[V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t] \\ =&\mathbb{E}\left[\frac{1}{N} + \frac{1}{N}\sum_{i}^{N} V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t\right] \\ =&\mathbb{E}\left[\frac{1}{N} + \frac{1}{N}\sum_{i}^{N}\left(\frac{T-t}{N} - R(\mathbf{M}+\mathbf{e}_i, T-t)\right)|T \geq t\right] \\ =&\mathbb{E}[\frac{T-t+1}{N} - R(\mathbf{M}, T-t+1)|T \geq t] \quad \text{(by definition of R)} \\ =&\mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t],\end{aligned}$$

which implies

$$\min_{\mathbf{P}\in\Delta(N)}\max_{i\in[N]} g(\mathbf{P}, i) \geq \mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t]. \tag{16}$$

On the other hand, let $\mathbf{P}^* = \mathbb{E}[\mathbf{P}^T|T \geq t]$, where $P_i^T = V(\mathbf{M}, T-t+1) - V(\mathbf{M}+\mathbf{e}_i, T-t)$. $\mathbf{P}^*$ is a valid distribution since $\mathbf{P}^T$ is a distribution for any T. Also, by plugging into Eq. (15),

$$\begin{aligned}g(\mathbf{P}^*, i) &= \mathbb{E}[V(\mathbf{M}, T-t+1) - V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t] \\ &\quad + \mathbb{E}[V(\mathbf{M}+\mathbf{e}_i, T-t)|T \geq t] \\ &= \mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t].\end{aligned}$$

Therefore,

$$\begin{aligned}\min_{\mathbf{P}\in\Delta(N)}\max_{i\in[N]} g(\mathbf{P}, i) &\leq \max_{i\in[N]} g(\mathbf{P}^*, i) \\ &= \mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t].\end{aligned} \tag{17}$$

Eq. (16) and (17) show that $\min_{\mathbf{P}}\max_i g(\mathbf{P}, i) = \mathbb{E}[V(\mathbf{M}, T-t+1)|T \geq t]$, which agrees with the left hand side of Eq. (13). We thus prove $\inf_{\mathbf{Alg}}\sup_{\mathbf{Z}_{1:\infty}}\mathbb{E}_{T\sim Q}[\mathbf{Reg}(L_T, \mathbf{M}_T)] = \mathbb{E}[V(\mathbf{0}, T)|T \geq 1] = \mathbb{E}_{T\sim Q}[\inf_{\mathbf{Alg}}\sup_{\mathbf{Z}_{1:T}}\mathbf{Reg}(L_T, \mathbf{M}_T)]$, and $\mathbf{P}^*$ is the optimal strategy. □

C. Proof of Theorem 3

To prove Theorem 3, we need to find out what $V(\mathbf{0}, T)$ is under the general loss space $[0,1]^2$. Note that Theorem 1 only tells us the case of the basis vector loss space. Fortunately, it turns out that they are the same if $N = 2$. To be more specific, we will show later in Theorem 10 that if $N = 2$ and $\mathbf{LS} = [0,1]^2$, then $V(\mathbf{0}, T) = T/2 - R(\mathbf{0}, T)$, which can be further simplified as

$$V(\mathbf{0}, T) = \frac{T}{2} - \frac{1}{2^T}\sum_{m=0}^{T}\binom{T}{m}\min\{m, T-m\} = \frac{T}{2^T}\binom{T-1}{\lfloor \frac{T}{2} \rfloor}.$$

We can now prove Theorem 3 using this explicit scaling factor, denoted by $S(T)$ for simplicity.

Proof of Theorem 3. Again, solving Eq. (4) is equivalent to finding the value function $\tilde{V}$ defined on each state of the game, similar to the functions V and $\bar{V}$ we had before. The difference is that $\tilde{V}$ should be a function of not only the index of the current round t and the cumulative loss vector $\mathbf{M}$, but also the cumulative loss L for the learner. Moreover, to obtain a base case for the recursive definition, it is convenient to first assume that T is at most T_0, where T_0 is some fixed integer. Under these conditions, we define $\tilde{V}_t^{T_0}(L, \mathbf{M})$ recursively as:

$$\tilde{V}_{T_0}^{T_0}(L, \mathbf{M}) \triangleq \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} \frac{\mathbf{Reg}(L + \mathbf{P}\cdot\mathbf{Z}, \mathbf{M}+\mathbf{Z})}{V(\mathbf{0}, T_0)},$$

$$\tilde{V}_{t}^{T_0}(L, \mathbf{M}) \triangleq \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} \max\left\{\frac{\mathbf{Reg}(L + \mathbf{P}\cdot\mathbf{Z}, \mathbf{M}+\mathbf{Z})}{V(\mathbf{0}, t)}, \tilde{V}_{t+1}^{T_0}(L + \mathbf{P}\cdot\mathbf{Z}, \mathbf{M}+\mathbf{Z})\right\},$$

which is the scaled regret starting from round t with cumulative loss L for the learner and $\mathbf{M}$ for the actions, assuming both the learner and the adversary will play optimally from this round on. The value of the game $\tilde{V}$ is now $\lim_{T_0\to+\infty} \tilde{V}_1^{T_0}(0, \mathbf{0})$.

To simplify this value function, we will need three facts. First, the base case can be related to $V(\mathbf{M}, 1)$:

$$\begin{aligned}
&\tilde{V}_{T_0}^{T_0}(L, \mathbf{M}) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} \frac{\mathbf{Reg}(L + \mathbf{P}\cdot\mathbf{Z}, \mathbf{M}+\mathbf{Z})}{V(\mathbf{0}, T_0)} \\
&= \left(L + \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} \mathbf{Reg}(\mathbf{P}\cdot\mathbf{Z}, \mathbf{M}+\mathbf{Z})\right)\Big/ V(\mathbf{0}, T_0) \\
&= \left(L + \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}} \mathbf{P}\cdot\mathbf{Z} + V(\mathbf{M}+\mathbf{Z}, 0)\right)\Big/ V(\mathbf{0}, T_0) \\
&= \frac{L + V(\mathbf{M}, 1)}{V(\mathbf{0}, T_0)}.
\end{aligned}$$

Second, for any L and $\mathbf{M}$, one can inductively show that

$$\tilde{V}_t^{T_0}(L, \mathbf{M}) = \tilde{V}_t^{T_0}(L - R(\mathbf{M}, 0), \mathbf{M}'), \tag{18}$$

where $M_i' = M_i - R(\mathbf{M}, 0)$. (We omit the details since it is straightforward.)

Third, when $\mathbf{M} = \mathbf{0}$, by symmetry, one has with $\mathbf{P}_u = (\frac{1}{N}, \ldots, \frac{1}{N})$

$$\begin{aligned}
&\tilde{V}_t^{T_0}(L, \mathbf{0}) \\
&= \max_{\mathbf{Z}\in\mathbf{LS}} \max\left\{\frac{\mathbf{Reg}(L + \mathbf{P}_u\cdot\mathbf{Z}, \mathbf{Z})}{V(\mathbf{0}, t)}, \tilde{V}_{t+1}^{T_0}(L + \mathbf{P}_u\cdot\mathbf{Z}, \mathbf{Z})\right\} \\
&\geq \max\left\{\frac{L + \frac{1}{N}}{V(\mathbf{0}, t)}, \tilde{V}_{t+1}^{T_0}(L + \frac{1}{N}, \mathbf{e}_1)\right\}.
\end{aligned} \tag{19}$$

Now we can make use of the condition $N = 2$ to lower bound $\tilde{V}$. The key point is to consider a restricted adversary who can only place one unit more loss on one of the action than the other, if not stopping the game. Clearly the value of this restricted game serves as a lower bound of $\tilde{V}$. Specifically, consider the value of $\tilde{V}_t^{T_0}(L, \mathbf{e}_1)$ for $t \leq T_0 - 2$:

$$\begin{aligned}
&\tilde{V}_t^{T_0}(L, \mathbf{e}_1) \\
&\geq \min_{p\in[0,1]} \max\left\{\frac{\mathbf{Reg}(L + p, 2\mathbf{e}_1)}{S(t)}, \tilde{V}_{t+1}^{T_0}(L + 1 - p, \mathbf{e}_1 + \mathbf{e}_2)\right\} \\
&\qquad \text{(restricted adversary)} \\
&= \min_{p\in[0,1]} \max\left\{\frac{L+p}{S(t)}, \tilde{V}_{t+1}^{T_0}(L - p, \mathbf{0})\right\} \quad \text{(by Eq. (18))} \\
&\geq \min_{p\in[0,1]} \max\left\{\frac{L+p}{S(t)}, \frac{L + 1/2 - p}{S(t+1)}, \tilde{V}_{t+2}^{T_0}(L + \frac{1}{2} - p, \mathbf{e}_1)\right\} \\
&\qquad \text{(by Eq. (19))} \\
&\geq \min_{p\in\mathbb{R}} \max\left\{\frac{L+p}{S(t)}, \tilde{V}_{t+2}^{T_0}(L + \frac{1}{2} - p, \mathbf{e}_1)\right\}
\end{aligned}$$

Therefore, if we assume T_0 is even without loss of generality and define function $G_t^{T_0}(L)$ recursively as:

$$G_{T_0}^{T_0}(L) \triangleq \tilde{V}_{T_0}^{T_0}(L, \mathbf{e}_1) = \frac{L + V(\mathbf{e}_1, 1)}{S(T_0)} = \frac{L}{S(T_0)}$$

$$G_t^{T_0}(L) \triangleq \min_{p\in\mathbb{R}} \max\left\{\frac{L+p}{S(t)}, G_{t+2}^{T_0}(L + \frac{1}{2} - p)\right\},$$

then it is clear that $\tilde{V}_t^{T_0}(L, \mathbf{e}_1) \geq G_t^{T_0}(L)$, and thus by (19),

$$\tilde{V}_1^{T_0}(0, \mathbf{0}) \geq \max\{1, \tilde{V}_2^{T_0}(\frac{1}{2}, \mathbf{e}_1)\} \geq \max\{1, G_2^{T_0}(\frac{1}{2})\}.$$

It remains to compute $G_2^{T_0}(\frac{1}{2})$. By some elementary computations and the fact that for two linear functions $h_1(p)$ and $h_2(p)$ of different signs of slopes,

$\min_p \max\{h_1(p), h_2(p)\} = h_1(p^*)$ where p^* is such that $h_1(p^*) = h_2(p^*)$, one can inductively prove that for $t = 2, 4, \ldots, T_0$,

$$G_t^{T_0}(L) = \frac{2^{\frac{T_0-t}{2}}(L+\frac{1}{2}) - \frac{1}{2}}{S(T_0) + \sum_{k=1}^{(T_0-t)/2}\left(2^{k-1}S(T_0-2k)\right)}.$$

Plugging $S(t) = \frac{t}{2^t}\binom{t-1}{\lfloor t/2 \rfloor}$ and letting $T_0 \to \infty$, we arrive at

$$\begin{aligned}
&\lim_{T_0\to\infty} G_2^{T_0}(1/2) \\
&= \lim_{T_0\to\infty}\left(\sum_{k=1}^{T_0/2-1}\left(2^{k-T_0/2}S(T_0-2k)\right)\right)^{-1} \\
&= \lim_{T_0\to\infty}\left(\sum_{k=1}^{T_0/2-1}\left(\frac{S(2k)}{2^k}\right)\right)^{-1} \\
&= \left(\sum_{j=0}^{\infty}\frac{j}{8^j}\binom{2j}{j}\right)^{-1}.
\end{aligned}$$

Define $G(x) = \sum_{j=0}^{\infty}\binom{2j}{j}x^j$ and $F(x) = x \cdot G'(x)$. Note that what we want to compute above is exactly $1/F(\frac{1}{8})$. Lehmer (1985) showed that $G(x) = (1-4x)^{-1/2}$. Therefore, $F(x) = 2x \cdot (1-4x)^{-3/2}$ and

$$\lim_{T_0\to\infty} G_2^{T_0}(1/2) = 1/F(1/8) = \sqrt{2}.$$

We conclude the proof by pointing out

$$\begin{aligned}
\tilde{V} &= \lim_{T_0\to\infty}\tilde{V}_1^{T_0}(0, \mathbf{0}) \\
&\geq \max\{1, \lim_{T_0\to\infty} G_2^{T_0}(1/2)\} = \sqrt{2}.
\end{aligned}$$

□

As we mentioned at the beginning of this section, the last thing we need to show is that the value $V(\mathbf{0}, T)$ is the same under the two loss spaces. In fact, we will prove stronger results in the following theorem claiming that this is true only if $N = 2$.

Theorem 10. *Let* $\mathbf{LS}_1, \mathbf{LS}_2, \mathbf{LS}_3$ *be the three loss spaces* $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}, \{0,1\}^N$ *and* $[0,1]^N$ *respectively, and* $V_{\mathbf{LS}}(\mathbf{0}, T)$ *be the value of the game* $V(\mathbf{0}, T)$ *under the loss space* $\mathbf{LS}$*. If* $N > 2$*, we have for any* T*,*

$$V_{\mathbf{LS}_1}(\mathbf{0}, T) < V_{\mathbf{LS}_2}(\mathbf{0}, T) = V_{\mathbf{LS}_3}(\mathbf{0}, T).$$

However, the three values above are the same if $N = 2$*.*

Proof. We first inductively show that for any $\mathbf{M}$ and r, $V_{\mathbf{LS}_2}(\mathbf{M}, r) = V_{\mathbf{LS}_3}(\mathbf{M}, r)$ and $V_{\mathbf{LS}_3}(\mathbf{M}, r)$ is convex in $\mathbf{M}$. For the base case $r = 0$, by definition, $V_{\mathbf{LS}_2}(\mathbf{M}, 0) = V_{\mathbf{LS}_3}(\mathbf{M}, 0) = -\min_i M_i$. Also, for any two loss vectors $\mathbf{M}$ and $\mathbf{M}'$, and $\lambda \in [0,1]$,

$$\begin{aligned}
&V_{\mathbf{LS}_3}(\lambda\mathbf{M} + (1-\lambda)\mathbf{M}', 0) \\
=& -\min_i\left(\lambda M_i + (1-\lambda)M_i'\right) \\
\leq& -\min_i\left(\lambda M_i\right) - \min_i\left((1-\lambda)M_i'\right) \\
=& \lambda V_{\mathbf{LS}_3}(\mathbf{M}, 0) + (1-\lambda)V_{\mathbf{LS}_3}(\mathbf{M}', 0),
\end{aligned}$$

showing $V_{\mathbf{LS}_3}(\mathbf{M}, 0)$ is convex in $\mathbf{M}$. For $r > 0$,

$$V_{\mathbf{LS}_3}(\mathbf{M}, r) = \min_{\mathbf{P}\in\Delta(N)}\max_{\mathbf{Z}\in\mathbf{LS}_3}\left(\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1)\right).$$

Notice that $\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1)$ is equal to $\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_2}(\mathbf{M}+\mathbf{Z}, r-1)$ and is convex in $\mathbf{Z}$ by induction. Therefore the maximum is always achieved at one of the corner points of $\mathbf{LS}_3$, which is in $\mathbf{LS}_2$. In other words,

$$\begin{aligned}
V_{\mathbf{LS}_3}(\mathbf{M}, r) &= \min_{\mathbf{P}\in\Delta(N)}\max_{\mathbf{Z}\in\mathbf{LS}_2}\left(\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_2}(\mathbf{M}+\mathbf{Z}, r-1)\right) \\
&= V_{\mathbf{LS}_2}(\mathbf{M}, r).
\end{aligned}$$

On the other hand, by introducing a distribution Q over all the elements in $\mathbf{LS}_2$, we have

$$\begin{aligned}
&V_{\mathbf{LS}_3}(\mathbf{M}, r) \\
=& \min_{\mathbf{P}\in\Delta(N)}\max_{\mathbf{Z}\in\mathbf{LS}_2}\left(\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1)\right) \\
=& \min_{\mathbf{P}\in\Delta(N)}\max_{Q}\mathbb{E}_{\mathbf{Z}\sim Q}\left[\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1)\right] \\
=& \max_{Q}\min_{\mathbf{P}\in\Delta(N)}\mathbb{E}_{\mathbf{Z}\sim Q}\left[\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1)\right] \\
=& \max_{Q}\left(\mathbb{E}_{\mathbf{Z}\sim Q}V_{\mathbf{LS}_3}(\mathbf{M}+\mathbf{Z}, r-1) + \min_{\mathbf{P}\in\Delta(N)}\mathbf{P}\cdot\mathbb{E}_{\mathbf{Z}\sim Q}[\mathbf{Z}]\right)
\end{aligned}$$

where we switch the min and max by Corollary 37.3.2 of Rockafellar (1970). Note that the last expression is the maximum over a family of linear combinations of convex functions in $\mathbf{M}$, which is still a convex function in $\mathbf{M}$, completing the induction step. To conclude, $V_{\mathbf{LS}_2}(\mathbf{0}, T) = V_{\mathbf{LS}_3}(\mathbf{0}, T)$ for any N and T.

We next prove if $N = 2, V_{\mathbf{LS}_1}(\mathbf{0}, T) = V_{\mathbf{LS}_2}(\mathbf{0}, T)$. Again, we inductively prove $V_{\mathbf{LS}_1}(\mathbf{M}, r) = V_{\mathbf{LS}_2}(\mathbf{M}, r)$ for any $\mathbf{M}$ and r. The base case is clear. For $r > 0$, let $P_i^* = V_{\mathbf{LS}_1}(\mathbf{M}, r) - V_{\mathbf{LS}_1}(\mathbf{M}+\mathbf{e}_i, r-1)$ $(i = 1, 2)$. By induction,

$$\begin{aligned}
&V_{\mathbf{LS}_2}(\mathbf{M}, r) \\
=& \min_{\mathbf{P}\in\Delta(2)}\max_{\mathbf{Z}\in\mathbf{LS}_2}\left(\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_1}(\mathbf{M}+\mathbf{Z}, r-1)\right) \\
\leq& \max_{Z_1, Z_2\in\{0,1\}}\left(P_1^* Z_1 + P_2^* Z_2 + V_{\mathbf{LS}_1}(\mathbf{M}+(Z_1, Z_2), r-1)\right)
\end{aligned}$$

$$
\begin{aligned}
&= \max\{V_{\mathbf{LS}_1}(\mathbf{M}, r-1), 1 + V_{\mathbf{LS}_1}(\mathbf{M}+(1,1), r-1), \\
&\qquad V_{\mathbf{LS}_1}(\mathbf{M}, r)\} \\
&= \max\{V_{\mathbf{LS}_1}(\mathbf{M}, r-1), V_{\mathbf{LS}_1}(\mathbf{M}, r)\} \\
&\qquad\qquad\qquad \text{(by Property 5 in Proposition 2)} \\
&= V_{\mathbf{LS}_1}(\mathbf{M}, r). \qquad \text{(by Property 7 in Proposition 2)}
\end{aligned}
$$

However, it is clear that $V_{\mathbf{LS}_2}(\mathbf{M}, r) \geq V_{\mathbf{LS}_1}(\mathbf{M}, r)$. Therefore, $V_{\mathbf{LS}_1}(\mathbf{M}, r) = V_{\mathbf{LS}_2}(\mathbf{M}, r)$.

Finally, to prove $V_{\mathbf{LS}_1}(\mathbf{0}, T) < V_{\mathbf{LS}_2}(\mathbf{0}, T)$ for $N > 2$, we inductively prove $V_{\mathbf{LS}_1}((T-r)\mathbf{e}_1, r) < V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1, r)$ for $r = 1, \dots, T$. For the base case $r = 1$, $V_{\mathbf{LS}_1}((T-1)\mathbf{e}_1, 1) = 1/N - R((T-1)\mathbf{e}_1, 1) = 1/N$, while

$$
\begin{aligned}
&V_{\mathbf{LS}_2}((T-1)\mathbf{e}_1, 1) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}_2} (\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_2}((T-1)\mathbf{e}_1 + \mathbf{Z}, 0)) \\
&\geq \min_{\mathbf{P}\in\Delta(N)} \max_{i\in[N]} (1 - P_i + V_{\mathbf{LS}_2}((T-1)\mathbf{e}_1 + \mathbf{1} - \mathbf{e}_i, 0)) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max\{-P_1, 1-P_2, \dots, 1-P_N\}.
\end{aligned}
$$

We claim that the value of the last minimax expression above, denoted by v, is $(N-2)/(N-1)$, which is strictly greater than $1/N$ if $N > 2$ and thus proves the base case. To show that, notice that for any $\mathbf{P} \in \Delta(N)$, there must exist $i \in \{2, \dots, N\}$ such that $P_i \leq 1/(N-1)$ and

$$
\max\{-P_1, 1-P_2, \dots, 1-P_N\} \geq 1 - P_i \geq \frac{N-2}{N-1},
$$

showing $v \geq (N-2)/(N-1)$. On the other hand, the equality is realized by the distribution $\mathbf{P}^* = (0, \frac{1}{N-1}, \dots, \frac{1}{N-1})$.

For $r > 1$, we have

$$
\begin{aligned}
&V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1, r) \\
&= \min_{\mathbf{P}\in\Delta(N)} \max_{\mathbf{Z}\in\mathbf{LS}_2} (\mathbf{P}\cdot\mathbf{Z} + V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1 + \mathbf{Z}, r-1)) \\
&\geq \min_{\mathbf{P}\in\Delta(N)} \max_{i\in[N]} (P_i + V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1)) \\
&\geq \min_{\mathbf{P}\in\Delta(N)} \frac{1}{N}\sum_{i=1}^{N} (P_i + V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1)) \\
&= \frac{1}{N} + \frac{1}{N}\sum_{i=1}^{N} V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1) \\
&> \frac{1}{N} + \frac{1}{N}\sum_{i=1}^{N} V_{\mathbf{LS}_1}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1) \\
&= V_{\mathbf{LS}_1}((T-r)\mathbf{e}_1, r).
\end{aligned}
$$

Here, the last strict inequality holds because for $i = 1$, $V_{\mathbf{LS}_2}((T-r+1)\mathbf{e}_1, r-1) > V_{\mathbf{LS}_1}((T-r+1)\mathbf{e}_1, r-1)$ by induction; for $i \neq 1$, it is trivial that $V_{\mathbf{LS}_2}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1) \geq V_{\mathbf{LS}_1}((T-r)\mathbf{e}_1 + \mathbf{e}_i, r-1)$. Therefore, we complete the induction step and thus prove $V_{\mathbf{LS}_1}(\mathbf{0}, T) < V_{\mathbf{LS}_2}(\mathbf{0}, T)$. □

D. Proof of Theorem 5

The proof (and the one of Theorem 8) relies heavily on a common technique to approximate a sum using an integral, which we state without proof as the following claim.

Claim 1. *Let $f(x)$ be a non-increasing nonnegative function defined on $\mathbb{R}_+$. Then the following inequalities hold for any integer $0 < j \leq k$.*

$$
\int_j^{k+1} f(x)\,dx \leq \sum_{i=j}^{k} f(i) \leq \int_{j-1}^{k} f(x)\,dx
$$

Proof of Theorem 5. By Theorem 4, it suffices to upper bound $\bar{V}_1(\mathbf{0})$ and $\sum_{t=1}^{T_s} q_t \bar{V}_{t+1}(\mathbf{0})$. Let $S_t = \sum_{t'=t}^{\infty} 1/t'^d$. By applying Claim 1 multiple times, we have

$$
\frac{1}{S_t} \leq \left(\int_t^{\infty} \frac{dx}{x^d}\right)^{-1} = t^{d-1}(d-1); \tag{20}
$$

$$
q_t = \frac{1}{S_t \cdot t^d} \leq \frac{d-1}{t};
$$

$$
\begin{aligned}
&\bar{V}_1(\mathbf{0}) = \mathbb{E}[V(\mathbf{0}, T) | T \geq 1] \\
&\leq \frac{c_N}{S_1}\sum_{T=1}^{\infty} \frac{1}{T^{d-\frac{1}{2}}} && \text{(by Assumption 2)} \\
&\leq \frac{c_N}{S_1}\left(1 + \int_1^{\infty} \frac{dx}{x^{d-\frac{1}{2}}}\right) && \text{(by Claim 1)} \\
&= \frac{c_N(d-\frac{1}{2})}{S_1(d-\frac{3}{2})} \\
&\leq \frac{c_N(d-1)(d-\frac{1}{2})}{d-\frac{3}{2}} = O(1). && \text{(by Eq. (20))}
\end{aligned}
$$

For

$$
\begin{aligned}
\bar{V}_{t+1}(\mathbf{0}) &= \mathbb{E}[V(\mathbf{0}, T-t) | T \geq t+1] \\
&= \frac{c_N}{S_{t+1}}\sum_{k=1}^{\infty} \frac{\sqrt{k}}{(t+k)^d},
\end{aligned}
$$

Claim 1 does not readily apply since the function $g(k) = \sqrt{k}/(t+k)^d$ is increasing on $[0, t/(2d-1)]$ and then decreasing on $[t/(2d-1), \infty)$. However, we can still apply the claim to these two parts separately. Let $x_0 = \lfloor t/(2d-1) \rfloor$ and $x_1 = \lceil t/(2d-1) \rceil$. For simplicity, assume $1 \leq x_0 < x_1$ and $g(x_0) \leq g(x_1)$ (other cases hold similarly). Then we have

$$
\bar{V}_{t+1}(\mathbf{0}) = \frac{c_N}{S_{t+1}}\left(g(x_1) + \sum_{k=1}^{x_0} g(k) + \sum_{k=x_1+1}^{\infty} g(k)\right)
$$

$$\leq \frac{c_N}{S_{t+1}}\left(g(x_1) + \int_0^{x_1} g(x)dx + \int_{x_1}^{\infty} g(x)dx\right)$$
$$= \frac{c_N}{S_{t+1}}\left(g(x_1) + \frac{\Gamma(d-\frac{3}{2})}{2\Gamma(d)} \cdot \frac{\sqrt{\pi}}{t^{d-\frac{3}{2}}}\right)$$
$$\leq (d-1)c_N\sqrt{\pi} \cdot \frac{\Gamma(d-\frac{3}{2})}{2\Gamma(d)} \cdot \sqrt{t} + o(\sqrt{t}).$$

So finally we have

$$\sum_{t=1}^{T_s} q_t \bar{V}_{t+1}(\mathbf{0})$$
$$\leq (d-1)^2 c_N\sqrt{\pi} \cdot \frac{\Gamma(d-\frac{3}{2})}{2\Gamma(d)} \sum_{t=1}^{T_s}\left(\frac{1}{\sqrt{t}} + o(\frac{1}{\sqrt{t}})\right)$$
$$\leq \frac{\Gamma(d-\frac{3}{2})}{\Gamma(d)}(d-1)^2 c_N\sqrt{\pi T_s} + o(\sqrt{T_s}),$$

which proves the theorem. □

E. Proof of Theorem 6

Proof. Let $\Phi_t^T = \sqrt{\|\mathbf{W}_{t-1}\|^2 + (T-t+1)}$ be the potential function for this setting. The key property of the minimax algorithm Eq. (6) shown by Abernethy et al. (2008a) is the following:

$$\mathbf{x}_t^T \cdot \mathbf{w}_t \leq \Phi_t^T - \Phi_{t+1}^T.$$

Based on this property, the loss of our algorithm after T_s rounds is

$$\sum_{t=1}^{T_s} \mathbb{E}[\mathbf{x}_t^T | T \geq t] \cdot \mathbf{w}_t = \sum_{t=1}^{T_s} \mathbb{E}[\mathbf{x}_t^T \cdot \mathbf{w}_t | T \geq t]$$
$$\leq \sum_{t=1}^{T_s} \mathbb{E}[\Phi_t^T - \Phi_{t+1}^T | T \geq t].$$

Now define $U_t = \mathbb{E}[\Phi_t^T | T \geq t]$ and $q_t = \Pr[T < t+1 | T \geq t]$. By the fact that $f_{T\geq t}(t') = (1-q_t) f_{T \geq t+1}(t')$ for any $t' \geq t+1$, where $f_{T\geq t}$ and $f_{T\geq t+1}$ are conditional density functions, we have

$$\sum_{t=1}^{T_s} \mathbb{E}[\mathbf{x}_t^T | T \geq t] \cdot \mathbf{w}_t$$
$$\leq \sum_{t=1}^{T_s} \left(U_t - \mathbb{E}[\Phi_{t+1}^T | T \geq t]\right)$$
$$= \sum_{t=1}^{T_s} \left(U_t - \int_t^{t+1} \Phi_{t+1}^T f_{T\geq t}(T)dT - (1-q_t)U_{t+1}\right)$$
$$\leq \sum_{t=1}^{T_s} \left(U_t - \Phi_{t+1}^t \int_t^{t+1} f_{T\geq t}(T)dT - (1-q_t)U_{t+1}\right)$$
$$(\because \Phi_{t+1}^T \text{ increases in } T)$$
$$= \sum_{t=1}^{T_s} (U_t - U_{t+1} + q_t(U_{t+1} - \|\mathbf{W}_{t-1}\|))$$
$$(\because \Phi_{t+1}^t = \|\mathbf{W}_{t-1}\|)$$
$$= U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} q_t \mathbb{E}\left[\sqrt{\|\mathbf{W}_{t-1}\|^2 + (T-t)} - \|\mathbf{W}_{t-1}\| \mid T \geq t+1\right]$$
$$\leq U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} q_t \mathbb{E}\left[\sqrt{T-t} \mid T \geq t+1\right].$$
$$(\because \sqrt{a+b} - \sqrt{a} \leq \sqrt{b})$$

Note that $U_{T_s+1} \geq \|\mathbf{W}_T\|$, and thus it remains to plug in the distribution and compute U_1 and $\sum_{t=1}^{T_s} q_t \mathbb{E}[\sqrt{T-t} \mid T \geq t+1]$, which is almost the same process as what we did in the proof of Theorem 5 if one realizes $q_t \leq (d-1)/t$ also holds here. In a word, the regret can be bounded by

$$\frac{\Gamma(d-\frac{3}{2})}{\Gamma(d)}(d-1)^2\sqrt{\pi T_s} + o(\sqrt{T_s}),$$

which is $\pi\sqrt{T_s} + o(\sqrt{T_s})$ if $d = 2$. The explicit form in Eq. (7) comes from a direct calculation. □

F. Proof of Lemma 2 and Theorem 7

Proof of Lemma 2. The results follow by a direct calculation. The conditional distribution of $\boldsymbol{\xi}_t$ given T is $\frac{1}{\Delta_T^N}\mathbf{1}\{\boldsymbol{\xi} \in [0, \Delta_T]^N\}$. Let $S_t = \int_t^\infty 1/T^d dT = \left((d-1)t^{d-1}\right)^{-1}$. The marginal distribution for $\boldsymbol{\xi}$ that has negative coordinates is clearly 0. Otherwise, with $\bar{t} = \max\{t, \frac{\|\boldsymbol{\xi}\|_\infty^2}{bN}\}$ one has

$$f_t(\boldsymbol{\xi}) = \frac{1}{S_t}\int_t^\infty \frac{1}{T^d \Delta_T^N}\mathbf{1}\{\boldsymbol{\xi} \in [0, \Delta_T]^N\}dT$$
$$= \frac{1}{S_t}\int_{\bar{t}}^\infty \frac{1}{T^d \Delta_T^N} dT$$
$$= \frac{(d-1)t^{d-1}}{(\sqrt{bN})^N}\int_{\bar{t}}^\infty \frac{1}{T^{d+N/2}} dT$$
$$= \frac{d-1}{d-1+N/2}\Delta_t^{-N}\min\left\{1, \left(\frac{\Delta_t}{\|\boldsymbol{\xi}\|_\infty}\right)^{2d-2+N}\right\}.$$

□

Proof of Theorem 7. Applying Theorem 4.2 of Cesa-Bianchi & Lugosi (2006), the pseudo-regret of the FPL algorithm is bounded by

$$\mathbb{E}[\max_i \xi_{T_s,i}] + \sum_{t=1}^{T_s} \mathbb{E}[\max_i(\xi_{t-1,i} - \xi_{t,i})]$$

$$+\sum_{t=1}^{T_s}\int_{\mathbb{R}^N} F_t(\boldsymbol{\xi})(f_t(\boldsymbol{\xi})-f_t(\boldsymbol{\xi}-\mathbf{Z}_t))d\boldsymbol{\xi},$$

where we define $\boldsymbol{\xi}_0 = \mathbf{0}$ and $F_t(\boldsymbol{\xi}) = Z_{t,I_{\boldsymbol{\xi}}}$ with $I_{\boldsymbol{\xi}} \in \arg\min_i(M_{t-1,i}+\xi_i)$. Now the key observation is that the pseudo-regret remains the same if we replace random variables $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_{T_s}$ with $\boldsymbol{\xi}'_1,\ldots,\boldsymbol{\xi}'_{T_s}$ as long as $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}'_t$ have the same marginal distribution for any t. Specifically, we can let $\boldsymbol{\xi}'_{T_s} = \boldsymbol{\xi}_{T_S}$, and for $1 < t \le T_s$, let $\boldsymbol{\xi}'_{t-1} = \boldsymbol{\xi}'_t$ with probability $S_t/S_{t-1} = (1-1/t)^{d-1}$ (recall $S_t = \int_t^\infty 1/T^d dT$), or with $1-S_t/S_{t-1}$ probability be obtained by first drawing $T\in[t-1,t]$ according to density $f(T)\propto 1/T^d$, and then drawing a point uniformly in $[0,\Delta_T]^N$. It is clear that $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}'_t$ have the same marginal distribution. So the pseudo-regret can be in fact bounded by three terms:

$$A = \mathbb{E}[\max_i \xi_{T_s,i}],$$

$$B = \sum_{t=1}^{T_s}\mathbb{E}[\max_i(\xi'_{t-1,i}-\xi'_{t,i})],$$

$$C = \sum_{t=1}^{T_s}\int_{\mathbb{R}^N} F_t(\boldsymbol{\xi})(f_t(\boldsymbol{\xi})-f_t(\boldsymbol{\xi}-\mathbf{Z}_t))d\boldsymbol{\xi}.$$

A can be further bounded by

$$\frac{1}{S_{T_s}}\int_{T_s}^\infty \frac{\Delta_T}{T^d}dT = \frac{d-1}{d-3/2}\sqrt{bT_sN}.$$

For B, by construction of $\boldsymbol{\xi}'_t$, we have

$$\begin{aligned}
B &\le \sum_{t=2}^{T_s}\left(\frac{\Delta_t}{S_{t-1}}\int_{t-1}^t \frac{dT}{T^d} + \frac{S_t}{S_{t-1}}\cdot 0\right)\\
&= \sum_{t=2}^{T_s}\frac{\Delta_t}{t^{d-1}}\left(t^{d-1}-(t-1)^{d-1}\right)\\
&\le \sum_{t=2}^{T_s}\frac{\Delta_t}{t^{d-1}}\cdot(d-1)t^{d-2} \qquad \text{(by convexity)}\\
&\le 2(d-1)\sqrt{bT_sN}.
\end{aligned}$$

For C, let $H = \{\boldsymbol{\xi} : f_t(\boldsymbol{\xi}) > f_t(\boldsymbol{\xi}-\mathbf{Z}_t)\}$. Since $0 \le F_t(\boldsymbol{\xi}) \le 1$, we have $C \le \sum_{t=1}^{T_s}\int_H f_t(\boldsymbol{\xi})d\boldsymbol{\xi}$. Now observe that when $\min_i \xi_i \ge 0$, $f_t(\boldsymbol{\xi})$ is non-increasing in each ξ_i. So the only possibility that $f_t(\boldsymbol{\xi}) > f_t(\boldsymbol{\xi}-\mathbf{Z}_t)$ holds is when there exists an i such that ξ_i is strictly smaller than $Z_{t,i}$. That is

$$H = \{\boldsymbol{\xi} : \min_i \xi_i \ge 0 \text{ and } \exists i, s.t.\ \xi_i < Z_{t,i}\}$$

So we have

$$\begin{aligned}
C &\le \sum_{t=1}^{T_s}\frac{1}{S_t}\int_t^\infty \frac{dT}{T^d}\int_H \frac{\mathbf{1}\{\boldsymbol{\xi}\in[0,\Delta_T]^N\}}{\Delta_T^N}d\boldsymbol{\xi}\\
&\le \sum_{t=1}^{T_s}\frac{1}{S_t}\int_t^\infty \frac{N}{T^d}\frac{Z_{t,i}\Delta_T^{N-1}}{\Delta_T^N}dT\\
&\le \frac{d-1}{d-1/2}\sqrt{\frac{N}{b}}\sum_{t=1}^{T_s}\frac{1}{\sqrt{t}}\\
&\le \frac{2(d-1)}{\sqrt{b}(d-1/2)}\sqrt{T_sN}.
\end{aligned}$$

Combining A, B and C proves the theorem. □

G. Proof of Theorem 8

Proof. We will first show that

$$\mathbf{Reg}(L_{T_s},\mathbf{M}_{T_s}) \le \underbrace{(\ln N)\cdot\mathbb{E}\left[\frac{1}{\eta_T}|T\ge T_S+1\right]}_{A} + \underbrace{\frac{1}{8}\sum_{t=1}^{T_s}\mathbb{E}[\eta_T|T\ge t]}_{B}. \tag{21}$$

Let $\Phi_t^T = \frac{1}{\eta_T}\ln\left(\sum_{i=1}^N \exp(-\eta_T M_{t-1,i})\right)$. The key point of the proof for the non-adaptive version of the exponential weights algorithm is to use Φ_t^T as a "potential" function, and bound the change in potential before and after a single round (Cesa-Bianchi & Lugosi, 2006). Specifically, they showed that

$$\mathbf{P}_t^T\cdot\mathbf{Z}_t \le \frac{\eta_T}{8} + \Phi_t^T - \Phi_{t+1}^T.$$

We also base our proof on this inequality. The total loss of the learner after T_s rounds is

$$\begin{aligned}
L_{T_s} &= \sum_{t=1}^{T_s}\mathbb{E}[\mathbf{P}_t^T|T\ge t]\cdot\mathbf{Z}_t = \sum_{t=1}^{T_s}\mathbb{E}[\mathbf{P}_t^T\cdot\mathbf{Z}_t|T\ge t]\\
&\le B + \sum_{t=1}^{T_s}\mathbb{E}[\Phi_t^T-\Phi_{t+1}^T|T\ge t].
\end{aligned}$$

Define $U_t = \mathbb{E}[\Phi_t^T|T\ge t]$. We do the following transformation:

$$\begin{aligned}
&\mathbb{E}[\Phi_t^T-\Phi_{t+1}^T|T\ge t]\\
=&U_t - E_T[\Phi_{t+1}^T|T\ge t]\\
=&U_t - q_t\Phi_{t+1}^t - (1-q_t)U_{t+1}\\
=&U_t - U_{t+1} + q_t(U_{t+1}-\Phi_{t+1}^t)\\
=&U_t - U_{t+1} + q_t\cdot\mathbb{E}[\Phi_{t+1}^T-\Phi_{t+1}^t|T\ge t+1]\\
=&U_t - U_{t+1} + q_t\cdot\mathbb{E}[F_{T,t}(\mathbf{M}_t)|T\ge t+1],
\end{aligned}$$

where we define

$$F_{T,t}(\mathbf{M}) = \frac{\ln\left(\sum_i\exp(-\eta_T M_i)\right)}{\eta_T} - \frac{\ln\left(\sum_i\exp(-\eta_t M_i)\right)}{\eta_t}.$$

A key observation is

$$\max_{\substack{\mathbf{M}\in\mathbb{R}_+^N \\ \eta_T<\eta_t}} F_{T,t}(\mathbf{M}) = \frac{\ln N}{\eta_T} - \frac{\ln N}{\eta_t}, \tag{22}$$

which can be verified by a standard derivative analysis that we omit. (An alternative approach using KL-divergence can be found in Chapter 2.5 of Bubeck, 2011.)

We further define another potential function $\bar{\Phi}_t^T = (\ln N)/\eta_T$ and also $\bar{U}_t = \mathbb{E}[\bar{\Phi}_t^T | T \geq t]$. Note that the new potential $\bar{\Phi}_t^T$ has no dependence on t and thus $\bar{\Phi}_t^T = \bar{\Phi}_{t'}^T$ for any t, t'. We now have

$$\begin{aligned}
&\sum_{t=1}^{T_s} \mathbb{E}[\Phi_t^T - \Phi_{t+1}^T | T \geq t] \\
&= \sum_{t=1}^{T_s} \left(U_t - U_{t+1} + q_t \cdot \mathbb{E}[\Phi_{t+1}^T - \Phi_{t+1}^t | T \geq t+1]\right) \\
&= \underbrace{U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} \left(q_t \cdot \mathbb{E}[\Phi_{t+1}^T - \Phi_{t+1}^t | T \geq t+1]\right)}_{C}
\end{aligned} \tag{23}$$

$$\begin{aligned}
&\leq U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} \left(q_t \cdot \mathbb{E}[\frac{\ln N}{\eta_T} - \frac{\ln N}{\eta_t} | T \geq t+1]\right) \\
&\qquad \text{(by Eq. (22))} \\
&= \underbrace{\bar{U}_1 - \bar{U}_{T_s+1} + \sum_{t=1}^{T_s} \left(q_t \cdot \mathbb{E}[\bar{\Phi}_{t+1}^T - \bar{\Phi}_{t+1}^t | T \geq t+1]\right)}_{D} \\
&\quad + \bar{U}_{T_s+1} - U_{T_s+1}. \qquad (\because U_1 = \bar{U}_1)
\end{aligned}$$

Notice that D has the exact same form as C except for a different definition of the potential, and also Eq. (23) is an equality. Therefore, by a reverse transformation, we have

$$\begin{aligned}
&\sum_{t=1}^{T_s} \mathbb{E}[\Phi_t^T - \Phi_{t+1}^T | T \geq t] \\
=&\sum_{t=1}^{T_s} \mathbb{E}[\bar{\Phi}_t^T - \bar{\Phi}_{t+1}^T | T \geq t] + \bar{U}_{T_s+1} - U_{T_s+1} \\
=&\bar{U}_{T_s+1} - U_{T_s+1} \qquad (\because \bar{\Phi}_t^T = \bar{\Phi}_{t+1}^T)
\end{aligned}$$

$\bar{U}_{T_s+1}$ is exactly A in Eq. (21), and U_{T_s+1} can be related to the loss of the best action:

$$\begin{aligned}
U_{T_s+1} &= \mathbb{E}\left[\frac{1}{\eta_T} \ln \sum_{i=1}^N \exp(-\eta_T M_{T_s,i}) \mid T \geq T_s + 1\right] \\
&\geq \mathbb{E}\left[\frac{1}{\eta_T} \ln \exp(-\eta_T R(M_{T_s}, 0)) \mid T \geq T_s + 1\right] \\
&= -R(M_{T_s}, 0).
\end{aligned}$$

The regret is therefore

$$\begin{aligned}
\mathbf{Reg}(L_{T_s}, \mathbf{M}_{T_s}) &= L_{T_S} - R(M_{T_s}, 0) \\
&\leq A + B - U_{T_s+1} - R(M_{T_s}, 0) \\
&\leq A + B,
\end{aligned}$$

proving Eq. (21).

The rest of the proof is merely to plug in the distribution and $\eta_T = \sqrt{(b \ln N)/T}$, and upper bound Eq. (21) using Claim 1. Adopting the notation $S_t = \sum_{t'=t}^{\infty} 1/t'^d$ and the result of Eq. (20) in the proof of Theorem 5, we have

$$\begin{aligned}
A &= \frac{\sqrt{\ln N}}{S_{T_s+1}\sqrt{b}} \sum_{T=T_s+1}^{\infty} \frac{1}{T^{d-1/2}} \\
&\leq \frac{(d-1)\sqrt{\ln N}}{\sqrt{b}} (T_s+1)^{d-1} \cdot \\
&\quad \left(\int_{T_s+1}^{\infty} \frac{dx}{x^{d-1/2}} + \frac{1}{(T_s+1)^{d-1/2}}\right) \\
&= \frac{d-1}{(d-3/2)\sqrt{b}} \sqrt{T_s \ln N} + o(\sqrt{T_s \ln N});
\end{aligned}$$

$$\begin{aligned}
B &= \frac{\sqrt{b \ln N}}{8} \sum_{t=1}^{T_s} \frac{1}{S_t} \sum_{T=t}^{\infty} \frac{1}{T^{d+1/2}} \\
&\leq \frac{(d-1)\sqrt{b \ln N}}{8} \sum_{t=1}^{T_s} t^{d-1} \left(\int_t^{\infty} \frac{dx}{x^{d+1/2}} + \frac{1}{t^{d+1/2}}\right) \\
&\leq \frac{(d-1)\sqrt{b \ln N}}{8} \sum_{t=1}^{T_s} \left(\frac{1}{(d-1/2)\sqrt{t}} + \frac{1}{t^{d+3/2}}\right) \\
&\leq \frac{\sqrt{b}(d-1)}{4(d-1/2)} \sqrt{T_s \ln N} + o(\sqrt{T_s \ln N}).
\end{aligned}$$

Combining the bounds above for A and B proves the theorem. □

H. Proof of Theorem 9

Proof. The main idea resembles the one of Theorem 8, but the details are much more technical. Let us first define several notations:

$$\begin{aligned}
S_t &\triangleq \int_{m_t}^{\infty} \frac{dm}{m^d} = \frac{1}{(d-1)m_t^{d-1}}, \\
q_t &\triangleq \Pr[m < m_t | m \geq m_{t-1}] = \frac{1}{S_{t-1}} \int_{m_{t-1}}^{m_t} \frac{dm}{m^d} \\
&= 1 - \left(\frac{m_{t-1}}{m_t}\right)^{d-1}, \\
Y_t^m &\triangleq \sum_{i=1}^N \exp(-\eta_m M_{t-1,i}),
\end{aligned}$$

$$\Phi_t^m \triangleq \left(1+\frac{1}{\eta_m}\right)\ln Y_t^m, \quad U_t \triangleq \mathbb{E}[\Phi_t^m | m \ge m_{t-1}].$$

The proof starts from the following property of the exponential weights algorithm (Cesa-Bianchi & Lugosi, 2006):

$$\begin{aligned}\mathbf{P}_t^m \cdot \mathbf{Z}_t &\le \frac{1}{1-e^{-\eta_m}}\left(\ln Y_t^m - \ln Y_{t+1}^m\right)\\ &\le \Phi_t^m - \Phi_{t+1}^m. \qquad (\because \eta_m \ge \ln(1+\eta_m))\end{aligned}$$

By the fact that $f_{m\ge m_{t-1}}(m') = (1-q_t)f_{m\ge m_t}(m')$ for any $m' \ge m_t$, where $f_{m\ge m_{t-1}}$ and $f_{m\ge m_t}$ are conditional density functions, the loss of the learner after T_s rounds L_{T_s} is

$$\begin{aligned}&\sum_{t=1}^{T_s}\mathbb{E}[\mathbf{P}_t^m\cdot\mathbf{Z}_t | m \ge m_{t-1}]\\ \le&\sum_{t=1}^{T_s}\mathbb{E}[\Phi_t^m - \Phi_{t+1}^m | m \ge m_{t-1}]\\ =&\sum_{t=1}^{T_s}\left(U_t - \int_{m_{t-1}}^{m_t}\Phi_{t+1}^m f_{m\ge m_{t-1}}(m)dm + (1-q_t)U_{t+1}\right)\\ \le&\sum_{t=1}^{T_s}\left(U_t - \Phi_{t+1}^{m_{t-1}}\int_{m_{t-1}}^{m_t} f_{m\ge m_{t-1}}(m)dm + (1-q_t)U_{t+1}\right)\\ =&U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} q_t(U_{t+1} - \Phi_{t+1}^{m_{t-1}}),\end{aligned}$$

Here the last inequality holds because Φ_t^m is increasing in m. To show this, we consider the following

$$\begin{aligned}&\left(1+\frac{1}{\eta}\right)\ln\sum_{i=1}^N \exp(-\eta a_i)\\ =&\left(1+\frac{1}{\eta}\right)\left(-\eta a_1 + \ln\sum_{i=1}^N\exp(-\eta(a_i-a_1))\right)\\ =&-(\eta+1)a_1 + \left(1+\frac{1}{\eta}\right)\ln\sum_{i=1}^N\exp(-\eta(a_i-a_1)),\end{aligned}$$

where $\eta, a_1, \ldots, a_N$ are positive numbers. Since $\ln\sum_i \exp(-\eta(a_i - a_1)) \ge 0$, the expression above is decreasing in η, which along with the fact that η_m decreases in m shows that Φ_t^m increases in m.

We now compute U_1 and U_{T_s+1}:

$$\begin{aligned}U_1 &= \mathbb{E}[(1+\sqrt{m/\ln N})\ln N \mid m \ge 1]\\ &= \ln N + \frac{d-1}{d-3/2}\sqrt{\ln N}\end{aligned}$$

$$\begin{aligned}U_{T_s+1} &= \mathbb{E}\left[(1+1/\eta_m)\ln\sum_i \exp(-\eta_m M_{T_s,i}) \mid m \ge m_{T_s}\right]\\ &\ge \mathbb{E}[(1+1/\eta_m)(-\eta_m m^*) \mid m \ge m_{T_s}]\\ &= -m^*\left(1+\mathbb{E}[\eta_m \mid m \ge m_{T_s}]\right)\\ &= -m^*\left(1+\frac{d-1}{d-1/2}\sqrt{\frac{\ln N}{m_{T_s}}}\right)\\ &\ge -m^* - \frac{d-1}{d-1/2}\sqrt{m^*\ln N}\\ &\qquad (\because m_{T_s} = m^*+1)\end{aligned}$$

For $U_{t+1} - \Phi_{t+1}^{m_{t-1}} = \mathbb{E}[\Phi_{t+1}^m - \Phi_{t+1}^{m_{t-1}} \mid m \ge m_t]$, we first upper bound the part inside the expectation:

$$\begin{aligned}&\Phi_{t+1}^m - \Phi_{t+1}^{m_{t-1}}\\ =&\left(\frac{\ln Y_{t+1}^m}{\eta_m} - \frac{\ln Y_{t+1}^{m_{t-1}}}{\eta_{m_{t-1}}}\right) + (\eta_{m_{t-1}} - \eta_m)\min_i M_{t,i}\\ &+ \ln\frac{\sum e^{-\eta_m(M_{t,i}-\min_i M_{t,i})}}{\sum e^{-\eta_{m_{t-1}}(M_{t,i}-\min_i M_{t,i})}}.\end{aligned}$$

The first term above is at most $\left(\frac{1}{\eta_m} - \frac{1}{\eta_{m_{t-1}}}\right)\ln N = \sqrt{\ln N}(\sqrt{m} - \sqrt{m_{t-1}})$ by Eq. (22). The second term is at most $\sqrt{\ln N}(\frac{1}{\sqrt{m_{t-1}}} - \frac{1}{\sqrt{m}})m_{t-1}$ since $\min_i M_{t,i} = m_t - 1 \le m_{t-1}$, and the last term is at most $\ln N$ since the numerator is at most N while the denominator is at least 1. Therefore, we have

$$\begin{aligned}&U_{t+1} - \Phi_{t+1}^{m_{t-1}}\\ \le&\ln N + \sqrt{\ln N}\cdot\mathbb{E}[\sqrt{m} - \frac{m_{t-1}}{\sqrt{m}} \mid m \ge m_t]\\ =&\ln N + \sqrt{\ln N}\left(\frac{d-1}{d-3/2}\sqrt{m_t} - \frac{d-1}{d-1/2}\frac{m_{t-1}}{\sqrt{m_t}}\right)\\ \le&\ln N + \sqrt{\ln N}\left(\frac{d-1}{d-3/2}\sqrt{m_t} - \frac{d-1}{d-1/2}\frac{m_t-1}{\sqrt{m_t}}\right)\\ =&\ln N + \frac{(d-1)\sqrt{m_t\ln N}}{(d-3/2)(d-1/2)} + \frac{d-1}{d-1/2}\sqrt{\frac{\ln N}{m_t}}.\end{aligned}$$

It remains to compute $\sum_{t=1}^{T_s} q_t(U_{t+1} - \Phi_{t+1}^{m_{t-1}})$, which, using the above, can be done by computing $A = \sum_{t=1}^{T_s} q_t$, $B = \sum_{t=1}^{T_s} q_t\sqrt{m_t}$ and $C = \sum_{t=1}^{T_s} q_t/\sqrt{m_t}$. By inequality $1 - x \le -\ln x$ for any $x > 0$, we have

$$\begin{aligned}A &= \sum_{t=1}^{T_s}\left(1-\left(\frac{m_{t-1}}{m_t}\right)^{d-1}\right)\\ &\le -(d-1)\sum_{t=1}^{T_s}(\ln m_{t-1} - \ln m_t)\\ &= (d-1)\ln(m^*+1).\end{aligned}$$

For B, we first show $q_t\sqrt{m_t} \le 2(d-1)(\sqrt{m_t} - \sqrt{m_{t-1}})$, which is equivalent to

$$\frac{q_t\sqrt{m_t}}{\sqrt{m_t}-\sqrt{m_{t-1}}} = \frac{\left(\frac{m_t}{m_{t-1}}\right)^{d-1} - 1}{\left(\frac{m_t}{m_{t-1}}\right)^{d-1} - \left(\frac{m_t}{m_{t-1}}\right)^{d-3/2}} \le 2(d-1)$$

if $m_t \neq m_{t-1}$ (it is trivial otherwise). Define $h(x) = (x^{d-1}-1)/(x^{d-1}-x^{d-3/2})$ for $x \in [1,2]$ (note that m_t/m_{t-1} is within this interval). One can verify that $h'(x) < 0$ and thus $h(x) \le \lim_{x\to 1} h(x) = 2(d-1)$. So we prove $q_t\sqrt{m_t} \le 2(d-1)(\sqrt{m_t}-\sqrt{m_{t-1}})$ and

$$\begin{aligned} B &\le 2(d-1)\sum_{t=1}^{T_s}(\sqrt{m_t}-\sqrt{m_{t-1}}) \\ &= 2(d-1)(\sqrt{m_{T_s}}-1) \le 2(d-1)\sqrt{m^*}. \end{aligned}$$

A simple comparison of B and C shows $C = o(\sqrt{m^*})$. We finally conclude the proof by combining all we have

$$\begin{aligned} &\mathbf{Reg}(L_{T_s}, \mathbf{M}_{T_s}) \\ &\le U_1 - U_{T_s+1} + \sum_{t=1}^{T_s} q_t(U_{t+1} - \Phi_{t+1}^{m_{t-1}}) - m^* \\ &= (1+(d-1)\ln(m^*+1))\ln N \\ &\quad + \left(\frac{d-1}{d-1/2} + \frac{2(d-1)^2}{(d-3/2)(d-1/2)}\right)\sqrt{m^*\ln N} \\ &\quad + o(\sqrt{m^*\ln N}) \\ &= \frac{3(d-7/6)(d-1)}{(d-3/2)(d-1/2)}\sqrt{m^*\ln N} \\ &\quad + (1+(d-1)\ln(m^*+1))\ln N + o(\sqrt{m^*\ln N}). \end{aligned}$$

□

I. Examples

The first example shows that the results stated in Theorem 2 can not generalize to other loss spaces.

Example 1. *Consider the following Hedge setting:* $N = 3, \mathbf{LS} = \{\mathbf{1}-\mathbf{e}_1, \mathbf{1}-\mathbf{e}_2, \mathbf{1}-\mathbf{e}_3\}$ *where* $\mathbf{1} = (1,1,1)$. *Suppose the adversary picked* $\mathbf{1}-\mathbf{e}_1$ *and* $\mathbf{1}-\mathbf{e}_2$ *for the first two rounds and we are now on round* $t = 3$ *with* $\mathbf{M}_2 = (1,1,2)$. *Also the conditional distribution of the horizon given* $T \ge 3$ *is* $\Pr[T=3] = \Pr[T=4] = 1/2$. *Let* $\mathbf{P}^*$ *be the minimax strategy for this round and* $\mathbf{P}^T$ *be the minimax strategy assuming the horizon to be* T. *Then* $\mathbf{P}^* \neq \mathbb{E}[\mathbf{P}^T|T\ge 3]$, *and also*

$$\begin{aligned} &\inf_{\mathbf{Alg}}\sup_{\mathbf{Z}_{3:\infty}} \mathbb{E}[\mathbf{Reg}(L_T,\mathbf{M}_T)|T\ge 3] \\ &\neq \mathbb{E}[\inf_{\mathbf{Alg}}\sup_{\mathbf{Z}_{3:T}} \mathbf{Reg}(L_T,\mathbf{M}_T)|T\ge 3]. \end{aligned} \tag{24}$$

Proof. Recall the V function we had in Section 3. Ignoring the loss for the learner for the first two rounds (which is the same for both sides of Eq. (24)), we point out that the right hand side of Eq. (24) is essentially

$$\frac{1}{2}V(\mathbf{M}_2,1) + \frac{1}{2}V(\mathbf{M}_2,2),$$

and the left hand side, denoted by V', is

$$\min_{\mathbf{P}}\max_{\mathbf{Z}}(\mathbf{P}\cdot\mathbf{Z} + \frac{1}{2}V(\mathbf{M}_2+\mathbf{Z},0) + \frac{1}{2}V(\mathbf{M}_2+\mathbf{Z},1)).$$

Also $\mathbf{P}^*$ and $\mathbf{P}^T$ are the distributions that realize the minimum in the definition of V' and $V(\mathbf{M}_2, T-2)$ respectively. Below we show the values of these quantities without giving full details:

$$\begin{aligned} V(\mathbf{M}_2,1) &= \min_{\mathbf{P}}\max_i\{1-P_i+V(\mathbf{M}_2+\mathbf{1}-\mathbf{e}_i,0)\} \\ &= \min_{\mathbf{P}}\max\{-P_1,-P_2,-P_3-1\} \\ &= -1/2, \end{aligned}$$

with $\mathbf{P}^3 = (1/2,1/2,0)$;

$$\begin{aligned} V(\mathbf{M}_2,2) &= \min_{\mathbf{P}}\max_i\{1-P_i+V(\mathbf{M}_2+\mathbf{1}-\mathbf{e}_i,1)\} \\ &= \min_{\mathbf{P}}\max\{-P_1,-P_2,-P_3-1/3\} \\ &= -4/9, \end{aligned}$$

with $\mathbf{P}^4 = (4/9,4/9,1/9)$;

$$\begin{aligned} V' &= \min_{\mathbf{P}}\max_i\Big(1-P_i+\frac{1}{2}V(\mathbf{M}_2+\mathbf{1}-\mathbf{e}_i,0) \\ &\qquad + \frac{1}{2}V(\mathbf{M}_2+\mathbf{1}-\mathbf{e}_i,1)\Big) \\ &= \min_{\mathbf{P}}\max\{-P_1,-P_2,-P_3-2/3\} \\ &= -1/2, \end{aligned}$$

with $\mathbf{P}^* = (1/2,1/2,0)$. We thus conclude that

$$\mathbb{E}[\mathbf{P}^T|T\ge 3] = (17/36, 17/36, 1/18) \neq \mathbf{P}^*$$

and

$$\mathbb{E}[V(\mathbf{M}_2, T-2)|T\ge 3] = -17/36 \neq V'.$$

□

The next two examples show that the idea of "treating the current round as the last round" does not work for minimax algorithms.

Example 2. *Consider the following Hedge setting:* $N = 2, \mathbf{LS} = [0,1]^2$ *and the horizon* T *is a even number. Suppose on round* t, *the learner chooses* $\mathbf{P}_t$ *using the minimax algorithm assuming horizon* $T = t$. *Then the adversary can make the regret after* T *rounds to be* $T/4$ *by choosing* $\mathbf{e}_1$ *and* $\mathbf{e}_2$ *alternatively.*

Proof. As shown in Theorem 10, when $N = 2$, the minimax algorithm with $\mathbf{LS} = [0,1]^2$ is the same as the one with $\mathbf{LS} = \{\mathbf{e}_1, \mathbf{e}_2\}$, which we already know from Theorem 1. If the learner treats the current round as the last round, then $P_{t,1}$ is

$$\begin{aligned} & V(\mathbf{M}_{t-1}, 1) - V(\mathbf{M}_{t-1} + \mathbf{e}_1, 0) \\ =& \frac{1}{2}\big(1 + \min\{M_{t-1,1} + 1, M_{t-1,2}\} \\ & - \min\{M_{t-1,1}, M_{t-1,2} + 1\}\big). \end{aligned}$$

Hence, for any round t where t is odd, we have $\mathbf{M}_{t-1} = (\frac{t-1}{2}, \frac{t-1}{2})$ and thus $P_{t,1} = P_{t,2} = 1/2$ and the learner suffers loss $1/2$. For any round t where t is even, we have $\mathbf{M}_{t-1} = (\frac{t}{2}, \frac{t}{2} - 1)$ and thus $P_{t,1} = 0, P_{t,2} = 1$ and the learner suffers loss 1 since the adversary will choose $\mathbf{e}_2$ for this round. Finally, at the end of T rounds, the loss of the best action is clearly $T/2$. So the regret would be $3T/4 - T/2 = T/4$. □

Example 3. *Consider the online linear optimization problem described in Section 6.1. If horizon T is even and the learner predicts using the minimax algorithm Eq (6) with T replaced with t. Then the adversary can make the regret to be $\sqrt{2}T/4$ after T rounds by choosing $\mathbf{e}_1$ and $-\mathbf{e}_1$ alternatively.*

Proof. For any round t where t is odd, we have $\mathbf{W}_{t-1} = \mathbf{0}$ and thus $\mathbf{x}_t = \mathbf{0}$. So the loss for this round is 0. For any round t where t is even, we have $\mathbf{W}_{t-1} = \mathbf{e}_1$ and thus $\mathbf{x}_t = -\frac{\sqrt{2}}{2}\mathbf{e}_1$. So the loss for this round is $\sqrt{2}/2$ since the adversary will pick $-\mathbf{e}_1$. At the end of T rounds, since $\mathbf{W}_T = \mathbf{0}$, the regret will simply be $\sqrt{2}T/4$. □