A Supplementary Proofs

A.1 Exp2 Regret Proofs

First, we directly analyze Exp2's regret for the two kinds of feedback.

A.1.1 Full Information

Lemma 19. *Let $L_t(X) = X^\top l_t$. If $|\eta L_t(X)| \leq 1$ for all $t \in [T]$ and $X \in \{0,1\}^n$, the Exp2 algorithm satisfies for any X:*

$$\sum_{t=1}^T p_t^\top L_t - \sum_{t=1}^T L_t(X) \leq \eta \sum_{t=1}^T p_t^\top L_t^2 + \frac{n \log 2}{\eta}$$

Proof. (Adapted from [11] Theorem 1.5) Let $Z_t = \sum_{Y \in \{0,1\}^n} w_t(Y)$. We have:

$$\begin{aligned} Z_{t+1} &= \sum_{Y \in \{0,1\}^n} \exp(-\eta L_t(Y)) w_t(Y) \\ &= Z_t \sum_{Y \in \{0,1\}^n} \exp(-\eta L_t(Y)) p_t(Y) \end{aligned}$$

Since $e^{-x} \leq 1 - x + x^2$ for $x \geq -1$, we have that $\exp(-\eta L_t(Y)) \leq 1 - \eta L_t(Y) + \eta^2 L_t(Y)^2$ (Because we assume $|\eta L_t(X)| \leq 1$). So,

$$\begin{aligned} Z_{t+1} &\leq Z_t \sum_{Y \in \{0,1\}^n} (1 - \eta L_t(Y) + \eta^2 L_t(Y)^2) p_t(Y) \\ &= Z_t(1 - \eta p_t^\top L_t + \eta^2 p_t^\top L_t^2) \end{aligned}$$

Using the inequality $1 + x \leq e^x$,

$$Z_{t+1} \leq Z_t \exp(-\eta p_t^\top L_t + \eta^2 p_t^\top L_t^2)$$

Hence, we have:

$$Z_{T+1} \leq Z_1 \exp(-\sum_{t=1}^T \eta p_t^\top L_t + \sum_{t=1}^T \eta^2 p_t^\top L_t^2)$$

For any $X \in \{0,1\}^n$, $w_{T+1}(X) = \exp(-\sum_{t=1}^T \eta L_t(X))$. Since $w(T+1)(X) \leq Z_{T+1}$ and $Z_1 = 2^n$, we have:

$$\exp(-\sum_{t=1}^T \eta L_t(X)) \leq 2^n \exp(-\sum_{t=1}^T \eta p_t^\top L_t + \sum_{t=1}^T \eta^2 p_t^\top L_t^2)$$

Taking the logarithm on both sides manipulating this inequality, we get:

$$\sum_{t=1}^T p_t^\top L_t - \sum_{t=1}^T L_t(X) \leq \eta \sum_{t=1}^T p_t^\top L_t^2 + \frac{n \log 2}{\eta}$$

□

Theorem 3. *In the full information setting, if $\eta = \sqrt{\frac{\log 2}{nT}}$, Exp2 attains the regret bound:*

$$E[\mathcal{R}_T] \leq 2n^{3/2} \sqrt{T \log 2}$$

Proof. Using $L_t(X) = X^\top l_t$ and applying expectation with respect to the randomness of the player to definition of regret, we get:

$$\begin{aligned} E[\mathcal{R}_T] &= \sum_{t=1}^T \sum_{X \in \{0,1\}^n} p_t(X) L_t(X) - \min_{X^\star \in \{0,1\}^n} \sum_{t=1}^T L_t(X^\star) \\ &= \sum_{t=1}^T p_t^\top L_t - \min_{X^\star \in \{0,1\}^n} \sum_{t=1}^T L_t(X^\star) \end{aligned}$$

Applying Lemma 19, we get $E[\mathcal{R}_T] \leq \eta \sum_{t=1}^T p_t^\top L_t^2 + n \log 2 / \eta$. Since $|L_t(X)| \leq n$ for all $X \in \{0,1\}^n$, we get $\sum_{t=1}^T p_t^\top L_t^2 \leq Tn^2$.

$$E[\mathcal{R}_T] \leq \eta T n^2 + \frac{n \log 2}{\eta}$$

Optimizing over the choice of η, we get the regret is bounded by $2n^{3/2}\sqrt{T \log 2}$ if we choose $\eta = \sqrt{\frac{\log 2}{nT}}$.

To apply Lemma 19, $|\eta L_t(X)| \leq 1$ for all $t \in [T]$ and $X \in \{0,1\}^n$. Since $|L_t(X)| \leq n$, we have $\eta \leq 1/n$. □

A.1.2 Bandit

Lemma 20. *Let $\tilde{L}_t(X) = X^\top \tilde{l}_t$, where $\tilde{l}_t = P_t^{-1} X_t X_t^\top l_t$. If $|\eta \tilde{L}_t(X)| \leq 1$ for all $t \in [T]$ and $X \in \{0,1\}^n$, the Exp2 algorithm with uniform exploration satisfies for any X*

$$\sum_{t=1}^T q_t^\top L_t - \sum_{t=1}^T L_t(X) \leq \eta \mathbb{E}[\sum_{t=1}^T q_t^\top \tilde{L}_t^2] + \frac{n \log 2}{\eta} + 2\gamma n T$$

Proof. We have that:

$$\begin{aligned} \sum_{t=1}^T q_t^\top \tilde{L}_t - \sum_{t=1}^T \tilde{L}_t(X) &= (1-\gamma)(\sum_{t=1}^T p_t^\top \tilde{L}_t - \sum_{t=1}^T \tilde{L}_t(X)) \\ &\quad + \gamma(\sum_{t=1}^T \mu^\top \tilde{L}_t - \sum_{t=1}^T \tilde{L}_t(X)) \end{aligned}$$

Since the algorithm essentially runs Exp2 using the losses $\tilde{L}_t(X)$ and $|\eta \tilde{L}_t(X)| \leq 1$, we can apply Lemma 19:

$$\begin{aligned} \sum_{t=1}^T q_t^\top \tilde{L}_t - \sum_{t=1}^T \tilde{L}_t(X) &\leq (1-\gamma)(\frac{n \log 2}{\eta} + \eta \sum_{t=1}^T p_t^\top \tilde{L}_t^2) \\ &\quad + \gamma(\sum_{t=1}^T \mu^\top \tilde{L}_t - \sum_{t=1}^T \tilde{L}_t(X)) \end{aligned}$$

Apply expectation with respect to X_t. Using the fact that $\mathbb{E}[\tilde{l}_t] = l_t$ and $\mu^\top L_t - L_t(X) \leq 2n$:

$$\begin{aligned}\sum_{t=1}^{T} q_t^\top L_t - \sum_{t=1}^{T} L_t(X) &\leq (1-\gamma)(\frac{n\log 2}{\eta} + \eta\mathbb{E}[\sum_{t=1}^{T} p_t^\top \tilde{L}_t^2]) \\ &+ \gamma(\sum_{t=1}^{T} \mu^\top L_t - \sum_{t=1}^{T} L_t(X)) \\ &\leq \eta\mathbb{E}[\sum_{t=1}^{T} q_t^\top \tilde{L}_t^2] + \frac{n\log 2}{\eta} + 2\gamma n T\end{aligned}$$

□

Theorem 4. *In the bandit setting, if* $\eta = \sqrt{\frac{\log 2}{9n^2 T}}$ *and* $\gamma = 4n^2\eta$*, Exp2 with uniform exploration on* $\{0,1\}^n$ *attains the regret bound:*

$$\mathbb{E}[\mathcal{R}_T] \leq 6n^2\sqrt{T\log 2}$$

Proof. Applying expectation with respect to the randomness of the player to the definition of regret, we get:

$$\begin{aligned}\mathbb{E}[\mathcal{R}_T] &= \mathbb{E}[\sum_{t=1}^{T} L_t(X_t) - \min_{X^\star\in\{0,1\}^n} L_t(X^\star)] \\ &= \sum_{t=1}^{T} q_t^\top L_t - \min_{X^\star\in\{0,1\}^n} \sum_{t=1}^{T} L_t(X^\star)\end{aligned}$$

Applying Lemma 20

$$\mathbb{E}[\mathcal{R}_T] \leq \eta\mathbb{E}[\sum_{t=1}^{T} q_t^\top \tilde{L}_t^2] + \frac{n\log 2}{\eta} + 2\gamma n T$$

We follow the proof technique of [6] Theorem 4. We have that:

$$\begin{aligned}q_t^\top \tilde{L}_t^2 &= \sum_{X\in\{0,1\}^n} q_t(X)(X^\top \tilde{l}_t)^2 \\ &= \sum_{X\in\{0,1\}^n} q_t(X)(\tilde{l}_t^\top X X^\top \tilde{l}_t) \\ &= \tilde{l}_t^\top P_t \tilde{l}_t \\ &= l_t^\top X_t X_t^\top P_t^{-1} P_t P_t^{-1} X_t X_t^\top l_t \\ &= (X_t^\top l_t)^2 X_t^\top P_t^{-1} X_t \\ &\leq n^2 X_t^\top P_t^{-1} X_t = n^2 \mathrm{Tr}(P_t^{-1} X_t X_t^\top)\end{aligned}$$

Taking expectation, we get $E[q_t^\top \tilde{L}_t^2] \leq n^2\mathrm{Tr}(P_t^{-1}\mathbb{E}[X_t X_t^\top]) = n^2\mathrm{Tr}(P_t^{-1}P_t) = n^3$. Hence,

$$\mathbb{E}[\mathcal{R}_T] \leq \eta n^3 T + \frac{n\log 2}{\eta} + 2\gamma n T$$

However, in order to apply Lemma 20, we need that $|\eta X^\top \tilde{l}_t| \leq 1$. We have that

$$|\eta X^\top \tilde{l}_t| = \eta|(X_t^\top l_t) X^\top P_t^{-1} X_t| \leq 1$$

As $|X_t^\top l_t| \leq n$ and $|X_t^\top X| \leq n$, we get $\eta n|X^\top P_t^{-1} X_t| \leq \eta n|X^\top X_t|\|P_t^{-1}\| \leq \eta n^2\|P_t^{-1}\| \leq 1$. The matrix $P_t = (1-\gamma)\Sigma_t + \gamma\Sigma_\mu$. The smallest eigenvalue of Σ_μ is 1/4[8]. So $P_t \succeq \frac{\gamma}{4} I_n$ and $P_t^{-1} \preceq \frac{4}{\gamma} I_n$. We should have that $\frac{4n^2\eta}{\gamma} \leq 1$. Substituting $\gamma = 4n^2\eta$ in the regret inequality, we get:

$$\begin{aligned}\mathbb{E}[\mathcal{R}_T] &\leq \eta n^3 T + 8\eta n^3 T + \frac{n\log 2}{\eta} \\ &\leq 9\eta n^3 T + \frac{n\log 2}{\eta}\end{aligned}$$

Optimizing over the choice of η, we get $\mathbb{E}[\mathcal{R}_T] \leq 2n^2\sqrt{9T\log 2}$ when $\eta = \sqrt{\frac{\log 2}{9n^2 T}}$. □

A.2 Lower Bounds

A.2.1 Full Information Lower bound

In the game between player and adversary, the players strategy is to pick some probability distribution $p_t \in \Delta(\{0,1\}^n)$ for $t = 1\ldots T$. The adversary picks a density q_t over loss vectors $l \in [-1,1]^n$ for $t = 1\ldots T$. So player picks $X_t \sim p_t$ and adversary picks $l_t \sim q_t$. The min max expected regret is:

$$\inf_{p_1\ldots p_T}\sup_{q_1\ldots q_t}\mathbb{E}_{l_t\sim q_t}\mathbb{E}_{X_t\sim p_t}\left[\sum_{t=1}^{T} l_t^\top X_t - \min_X \sum_{t=1}^{T} l_t^\top X\right]$$

Let $\mathbb{E}_{X_t\sim p_t} = x_t$.

$$\inf_{p_1\ldots p_T}\sup_{q_1\ldots q_t}\mathbb{E}_{l_t\sim q_t}[\sum_{t=1}^{T} l_t^\top x_t - \min_X \sum_{t=1}^{T} l_t^\top X]$$

Theorem 12. *For any learner there exists an adversary producing* L_∞ *losses such that the expected regret in the full information setting is:*

$$\mathbb{E}\left[\mathcal{R}_T\right] = \Omega\left(n\sqrt{T}\right).$$

Proof. We choose q_t to be the density such that $l_{t,i}$ is a Rademacher random variable, ie, $l_{t,i} = +1$ w.p. 1/2 and $l_{t,i} = -1$ w.p 1/2 for all $t = 1\ldots T$ and $i = [n]$. So,

$$\begin{aligned}&\inf_{p_1\ldots p_T}\sup_{q_1\ldots q_t}\mathbb{E}_{l_t\sim q_t}\left[\sum_{t=1}^{T} l_t^\top x_t - \min_X \sum_{t=1}^{T} l_t^\top X\right] \\ &\geq \inf_{p_1\ldots p_T}\mathbb{E}_{l_t}\left[\sum_{t=1}^{T} l_t^\top x_t - \min_X \sum_{t=1}^{T} l_t^\top X\right]\end{aligned}$$

For our choice of q_t, we have $\mathbb{E}_{l_t}[l_t^\top x_t] = 0$. So,

$$\inf_{p_1 \dots p_T} \mathbb{E}_{l_t}[\sum_{t=1}^{T} l_t^\top x_t - \min_X \sum_{t=1}^{T} l_t^\top X]$$
$$= \inf_{p_1 \dots p_T} \mathbb{E}_{l_t}[-\min_X \sum_{t=1}^{T} l_t^\top X]$$
$$= \mathbb{E}_{l_t}[\max_X \sum_{t=1}^{T} l_t^\top X]$$

Simplifying this, we get:

$$\begin{aligned}\mathbb{E}_{l_t}[\max_X \sum_{t=1}^{T} l_t^\top X] &= \mathbb{E}_{l_t}[\max_{X_1 \dots X_n} \sum_{t=1}^{T} \sum_{i=1}^{n} l_{t,i} X_i] \\ &= \mathbb{E}_{l_t}[\sum_{i=1}^{n} \max_{X_i} \sum_{t=1}^{T} l_{t,i} X_i] \\ &= \sum_{i=1}^{n} \mathbb{E}_{l_{t,i}}[\max_{X_i} \sum_{t=1}^{T} l_{t,i} X_i] \\ &= n\mathbb{E}_Y[\max_x \sum_{t=1}^{T} Y_t x]\end{aligned}$$

Here Y is a Rademacher random vector of length T and $x \in \{0, 1\}$. We have that

$$\max_x \left[\sum_{t=1}^{T} Y_t x\right] = \begin{cases} 0 & \text{If } \sum_{t=1}^{T} Y_t \leq 0 \\ \sum_{t=1}^{T} Y_t & \text{otherwise} \end{cases}$$

So

$$\begin{aligned}\mathbb{E}_Y \left[\max_x \sum_{t=1}^{T} Y_t x\right] &= \mathbb{E}_Y[\sum_{t=1}^{T} Y_t | \sum_{t=1}^{T} Y_t > 0] \\ &= \frac{1}{2}\mathbb{E}_Y \left|\sum_{t=1}^{T} Y_t\right|\end{aligned}$$

Using Khintchine's inequality, we have positive constants A and B such that:

$$A\left(\sum_{t=1}^{T} |1|^2\right)^{1/2} \leq \mathbb{E}_Y \left|\sum_{t=1}^{T} Y_t\right| \leq B\left(\sum_{t=1}^{T} |1|^2\right)^{1/2}$$

Hence, the regret is lower bounded by $\Omega(n\sqrt{T})$. □

A.2.2 Bandit Lower bound

Theorem 13. *For any learner there exists an adversary producing L_∞ losses such that the expected regret in the Bandit setting is:*

$$\mathbb{E}\left[\mathcal{R}_T\right] = \Omega\left(n^{3/2}\sqrt{T}\right).$$

Proof. We consider 2^n stochastic adversaries indexed by $X \in \{0, 1\}^n$. Adversary X draws losses as follows:

$$l_{t,i} = \begin{cases} \begin{cases} +1 \text{ w.p } \frac{1}{2} + \epsilon \\ -1 \text{ w.p } \frac{1}{2} - \epsilon \end{cases} & \text{if } X_i = 0 \\ \begin{cases} +1 \text{ w.p } \frac{1}{2} - \epsilon \\ -1 \text{ w.p } \frac{1}{2} + \epsilon \end{cases} & \text{if } X_i = 1 \end{cases}$$

Let $\tilde{l}_t = [X_1^\top l_1, X_2^\top l_2, \dots, X_t^\top l_t]$. We consider deterministic algorithms, ie X_t is a deterministic function of $\tilde{l}_{t-1}$. So, the only the adversary's randomness remains. The obtained result can be extended to randomized algorithms via application of Fubini's Theorem. Let E_X denote the expectation conditioned on adversary X. When playing against adversary X, the vector X is the best action in expectation. The expected regret when playing against adversary X.

$$\begin{aligned}\mathbb{E}_X[\mathcal{R}_T] &= \mathbb{E}_X\left[\sum_{t=1}^{T} l_t^\top X_t - \min_{X^\star} \sum_{t=1}^{T} l_t^\top X^\star\right] \\ &\geq \mathbb{E}_X\left[\sum_{t=1}^{T} l_t^\top X_t - \sum_{t=1}^{T} l_t^\top X\right] \\ &= 2\epsilon \sum_{i=1}^{n} \mathbb{E}_X[\sum_{t=1}^{T} \mathbf{1}(X_{i,t} \neq X_i)] \\ &= 2\epsilon T \sum_{i=1}^{n} \left(1 - \frac{\mathbb{E}_X[\sum_{t=1}^{T} \mathbf{1}(X_{i,t} = X_i)]}{T}\right)\end{aligned}$$

$\sum_{t=1}^{T} \mathbf{1}(X_{i,t} = X_i)/T$ is the empirical mean of playing X_i. Let J_i be a Bernoulli random drawn according to this mean. Hence,

$$\mathbb{E}_X[\mathcal{R}_T] \geq 2\epsilon T \sum_{i=1}^{n} (1 - \mathbb{P}_X(J_i = X_i))$$

Taking the average over adversaries:

$$\begin{aligned}\mathbb{E}[\mathcal{R}_T] &= \frac{1}{2^n} \sum_X \mathbb{E}_X[\mathcal{R}_T] \\ &\geq 2\epsilon T \sum_{i=1}^{n} \left(1 - \frac{1}{2^n} \sum_X \mathbb{P}_X(J_i = X_i)\right)\end{aligned}$$

Let $X^{\oplus i}$ be the vector X with the i'th bit flipped. Using Pinsker's inequality, we have that:

$$\mathbb{P}_X(J_i = X_i) \leq \mathbb{P}_{X^{\oplus i}}(J_i = X_i) + \sqrt{\frac{1}{2} KL(\mathbb{P}_{X^{\oplus i}} \| \mathbb{P}_X)}$$

Taking the summation, and using the concavity of square root:

$$\frac{1}{2^n} \sum_X \mathbb{P}_X(J_i = X_i) \leq \frac{1}{2} + \sqrt{\frac{1}{2} \frac{1}{2^n} \sum_X KL(\mathbb{P}_{X^{\oplus i}} \| \mathbb{P}_X)}$$

The sequence of observed losses $\tilde{l}_T \in \{-n, \ldots, +n\}^T$ determines the empirical distribution of plays. Let $\mathbb{P}_X^T$ be the law of $\tilde{l}_T$ when playing against adversary X. So, using the chain rule of Kullback Leibler divergence:

$$\begin{aligned}
&KL(\mathbb{P}_{X^{\oplus i}} \| \mathbb{P}_X) \leq KL(\mathbb{P}^T_{X^{\oplus i}} \| \mathbb{P}^T_X) \\
&= KL(\mathbb{P}^1_{X^{\oplus i}} \| \mathbb{P}^1_X) \\
&+ \sum_{t=2}^{T} \sum_{\tilde{l}_{t-1}} \mathbb{P}^{t-1}_{X^{\oplus i}}(\tilde{l}_{t-1}) KL(\mathbb{P}^t_{X^{\oplus i}}(\cdot|\tilde{l}_{t-1}) \| \mathbb{P}^t_X((\cdot|\tilde{l}_{t-1})) \\
&= KL(\mathcal{B}_0 \| \mathcal{B}'_0)\mathbf{1}(X_{1,i} = X_i) \\
&+ \sum_{t=T}^{t} \sum_{\tilde{l}_{t-1}: X_{t,i} = X_i} \mathbb{P}^{t-1}_{X^{\oplus i}}(\tilde{l}_{t-1}) KL(\mathcal{B}_{\tilde{l}_{t-1}} \| \mathcal{B}'_{\tilde{l}_{t-1}})
\end{aligned}$$

Here, $\mathcal{B}_{\tilde{l}_{t-1}}, \mathcal{B}'_{\tilde{l}_{t-1}}$ are sums of at most n Bernoulli random variables such that their means agree on all coordinates except i. Using Lemma 24 from [2], we get that:

$$KL(\mathcal{B}_{\tilde{l}_{t-1}} \| \mathcal{B}'_{\tilde{l}_{t-1}}) \leq \frac{16\epsilon^2}{n}$$

Substituting this back into the previous expression, we get:

$$\begin{aligned}
KL(\mathbb{P}_{X^{\oplus i}} \| \mathbb{P}_X) &\leq \frac{16\epsilon^2}{n} \sum_{t=1}^{T} \sum_{\tilde{l}_{1:t-1}: X_{t,i} = X_i} \mathbb{P}^t_{X^{\oplus i}}(\tilde{l}_{1:t-1}) \\
&\leq \frac{16\epsilon^2}{n} \sum_{t=1}^{T} \mathbb{E}_{X^{\oplus i}}(\mathbf{1}(X_{i,t} = X_i)) \\
&= \frac{16\epsilon^2}{n} T \mathbb{P}_{X^{\oplus i}}(J_i = X_i)
\end{aligned}$$

Taking the summation, we have:

$$\frac{1}{2^n} \sum_X KL(\mathbb{P}_{X^{\oplus i}} \| \mathbb{P}_X) \leq \frac{8\epsilon^2}{n} T$$

Substituting this in the regret inequality:

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_T] &\geq 2\epsilon T \sum_{i=1}^{n} \left(1 - \frac{1}{2} - \sqrt{\frac{4\epsilon^2 T}{n}}\right) \\
&= 2\epsilon T n \left(\frac{1}{2} - 2\epsilon\sqrt{\frac{T}{n}}\right)
\end{aligned}$$

Optimizing over ϵ, we get that $\mathbb{E}[\mathcal{R}_T] = \Omega(n^{3/2}\sqrt{T})$ ☐

A.3 $\{-1, +1\}^n$ Hypercube Case

Lemma 21. *Exp2 on $\{-1, +1\}^n$ with losses l_t is equivalent to Exp2 on $\{0, 1\}^n$ with losses $2l_t$ while using the map $2X_t - \mathbf{1}$ to play on $\{-1, +1\}^n$.*

Proof. Consider the update equation for Exp2 on $\{-1, +1\}^n$

$$p_{t+1}(Z) = \frac{\exp(-\eta \sum_{\tau=1}^t Z^\top l_\tau)}{\sum_{W \in \{-1,+1\}^n} \exp(-\eta \sum_{\tau=1}^t W^\top l_\tau)}$$

$Z \in \{-1, +1\}^n$ can be mapped to a $X \in \{0, 1\}^n$ using the bijective map $X = (Z + \mathbf{1})/2$. So:

$$\begin{aligned}
p_{t+1}(Z) &= \frac{\exp(-\eta \sum_{\tau=1}^t (2X - \mathbf{1})^\top l_\tau)}{\sum_{Y \in \{0,1\}^n} \exp(-\eta \sum_{\tau=1}^t (2Y - \mathbf{1})^\top l_\tau)} \\
&= \frac{\exp(-\eta \sum_{\tau=1}^t X^\top (2l_\tau))}{\sum_{Y \in \{0,1\}^n} \exp(-\eta \sum_{\tau=1}^t Y^\top (2l_\tau))}
\end{aligned}$$

This is equivalent to updating the Exp2 on $\{0, 1\}^n$ with the loss vector $2l_t$. ☐

Theorem 14. *Exp2 on $\{-1, +1\}^n$ using the sequence of losses l_t is equivalent to PolyExp on $\{0, 1\}^n$ using the sequence of losses $2\tilde{l}_t$. Moreover, the regret of Exp2 on $\{-1, 1\}^n$ will equal the regret of PolyExp using the losses $2\tilde{l}_t$.*

Proof. After sampling X_t, we play $Z_t = 2X_t - \mathbf{1}$. So $\Pr(X_t = X) = \Pr(Z_t = 2X - \mathbf{1})$. In full information, $2\tilde{l}_t = 2l_t$ and in the bandit case $\mathbb{E}[2\tilde{l}_t] = 2l_t$. Since $2\tilde{l}_t$ is used to update the algorithm, by Lemma 21 we have that $\Pr(X_{t+1} = X) = \Pr(Z_{t+1} = 2X - \mathbf{1})$. By equivalence of Exp2 to PolyExp, the first statement follows immediately.

Let $Z^\star = \min_{Z \in \{-1,+1\}^n} \sum_{t=1}^T Z^\top l_t$ and $2X^\star = Z^\star + \mathbf{1}$. The regret of Exp2 on $\{-1, +1\}^n$ is:

$$\begin{aligned}
\sum_{t=1}^{T} l_t^\top (Z_t - Z^\star) &= \sum_{t=1}^{T} l_t^\top (2X_t - \mathbf{1} - 2X^\star + \mathbf{1}) \\
&= \sum_{t=1}^{T} (2l_t)^\top (X_t - X^\star)
\end{aligned}$$

☐