Appendix

This appendix contains several pieces of exposition that were removed from the main text due to space constraints. First, in Appendix A, we provide several properties of Fenchel conjugates, which we hope will serve as a useful reference. In Appendix B, we provide proofs for results that were stated without proof in the main text. In Appendix C, we re-prove one of these results in the vector case for the convenience of readers who do not wish to read through matrix manipulations. In Appendix D, we generalize the exponentiated gradient results from the simplex case to the unconstrained case. Finally, in Appendix E, we show how to adaptively control the step size η in our algorithms to obtain regret almost as good as if the optimal η were known in advance.

A. Properties of Fenchel Conjugates

Throughout this paper we make extensive use of properties of Fenchel conjugates. We provide them here for reference. In all cases we assume that ψ is a convex function. We assume that the argument w to ψ is constrained to lie in some convex set S.

A.1. General Properties

Definition. The Fenchel conjugate $\psi^*(\beta)$ of a function $\psi(w)$ is defined as $\psi^*(\beta) \stackrel{\text{def}}{=} \sup_{w \in S} w^\top \beta - \psi(w)$.

Gradient. Let w be the maximizing vector in the preceding definition. Then w is a subgradient of ψ^* at β. If ψ^* is differentiable then

$$\nabla \psi^*(\beta) = \arg\max_{w \in S} w^\top \beta - \psi(w), \tag{30}$$

and in particular $\nabla \psi^*(\beta) \in S$ for all β.

Translations. For any vector c, define $\psi_c(w)$ to be $\psi(w) - w^\top c$. Then $\psi_c^*(\beta) = \psi^*(\beta + c)$.

A.2. Calculations (vector case)

Simplex. Let $S = \Delta_n$ and $\psi(w) = \sum_{i=1} w_i \log(w_i)$. This choice of ψ is also called the negative entropy (as well as, somewhat confusingly, an entropic regularizer). Then we have $\psi^*(\beta) = \log(\sum_{i=1}^n \exp(\beta_i))$ and $\nabla \psi^*(\beta)_i = \frac{\exp(\beta_i)}{\sum_{j=1}^n \exp(\beta_j)}$.

To see the latter, we note that applying the KKT conditions to $w^\top \beta - \psi(w)$ implies that the maximizer (and hence the gradient $\nabla \psi^*(\beta)$) satisfies $\beta_i = \log(w_i) + 1 + \lambda$ for some scalar λ, hence $w_i \propto \exp(\beta_i)$, and so $\nabla \psi^*(\beta_i) = w_i = \frac{\exp(\beta_i)}{\sum_j \exp(\beta_j)}$. Computing $\psi^*(\beta)$ now only involves evaluating $w^\top \beta - \psi(w)$ at its maximizing value, yielding (where we define Z_β as $\sum_j \exp(\beta_j)$)

$$\psi^*(\beta) = \sum_{i=1}^n \left[\beta_i \frac{\exp(\beta_i)}{Z_\beta} - \log(\exp(\beta_i)/Z_\beta) \frac{\exp(\beta_i)}{Z_\beta} \right] \tag{31}$$

$$= \sum_{i=1}^n \frac{\exp(\beta_i)}{Z_\beta} \log(Z_\beta) \tag{32}$$

$$= \log(Z_\beta), \tag{33}$$

which completes the calculation.

Non-negative orthant. If instead S is the non-negative orthant and we now take $\psi(w) = \sum_{i=1}^n w_i \log(w_i)$, we will have $\psi^*(\beta) = \sum_{i=1}^n \exp(\beta_i)$ and $\nabla \psi^*(\beta)_i = \exp(\beta_i)$.

To see this, again apply the KKT conditions to $w^\top \beta - \psi(w)$, which imply that the maximizing value of w satisfies

$\beta_i = \log(w_i)$, and hence $\nabla \psi^*(\beta)_i = w_i = \exp(\beta_i)$. Evaluating $w^\top \beta - \psi(w)$ at this point yields

$$\psi^*(\beta) = \sum_{i=1}^{n} [\beta_i \exp(\beta_i) - \beta_i \exp(\beta_i) + \exp(\beta_i)] \tag{34}$$

$$= \sum_{i=1}^{n} \exp(\beta_i), \tag{35}$$

thus completing the calculation.

A.3. Calculations (matrix case)

Trace constrained. Let $S = \{W \mid W \succeq 0, \operatorname{tr}(W) = 1\}$ and let $\psi(W) = \operatorname{tr}(W \log(W))$. This choice of ψ is called the von-Neumann entropy. We have $\psi^*(B) = \log(\operatorname{tr}(\exp(B)))$ and $\nabla \psi^*(B) = \frac{\exp(B)}{\operatorname{tr}(\exp(B))}$. Note that in this case $\psi^*(B)$ is defined as $\sup_{W \in S} \operatorname{tr}(WB) - \operatorname{tr}(W \log(W))$.

To calculate $\nabla \psi^*$, note that the KKT conditions yield $B = \log(W) + (1 + \lambda)I$ for the maximizing value of W. Thus $W \propto \exp(B)$ and hence $\nabla \psi^*(B) = W = \frac{\exp(B)}{\operatorname{tr}(\exp(B))}$. Defining Z_B to be $\operatorname{tr}(\exp(B))$ and plugging back in yields

$$\psi^*(B) = \operatorname{tr}(B \exp(B))/Z_B - \operatorname{tr}(\exp(B)[B - \log(Z_B)I])/Z_B \tag{36}$$

$$= \operatorname{tr}(\exp(B)) \log(Z_B)/Z_B \tag{37}$$

$$= \log(Z_B), \tag{38}$$

which completes the calculations for the trace-constrained case.

Trace unconstrained. Let $S = \{W \mid W \succeq 0\}$ and let $\psi(W) = \operatorname{tr}(W \log(W) - W)$. We have $\psi^*(B) = \operatorname{tr}(\exp(B))$ and $\nabla \psi^*(B) = \exp(B)$.

To calculate $\nabla \psi^*$, note that the KKT conditions yield $B = \log(W)$ and hence $\nabla \psi^*(B) = W = \exp(B)$. Plugging back in to ψ^* yields

$$\psi^*(B) = \operatorname{tr}(B \exp(B)) - \operatorname{tr}(\exp(B) \log(\exp(B)) - \exp(B)) \tag{39}$$

$$= \operatorname{tr}(\exp(B)), \tag{40}$$

which completes the calculations for the unconstrained case.

B. Deferred Proofs

In this section we prove all results stated in the main text that were deferred to the supplementary material.

Proof of Proposition 2.2. We will construct two sequences $(z_t)_{t=1}^{T}$ such that exponentiated gradient with any fixed step size η will perform poorly ($\Omega(\sqrt{T})$) on at least one of them. Our constructed sequences will involve $n = 2$ experts. In both sequences, the first expert has $z_{t,1} = 0$ for all t, and $z_{t,2}$ will satisfy $\sum_{t=1}^{T} z_{t,2} \geq 0$ to ensure quasi-realizability.

Sequence 1. The second expert has loss $z_{t,2} = (-1)^{t-1}$. Then $\sum_{t=1}^{T} z_{t,2}$ is either 0 or 1 depending on the parity of T, and in particular is non-negative. On odd-numbered rounds, $w_t = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}^\top$, and on even-numbered rounds, $w_t = \begin{bmatrix} \frac{1}{1+\exp(-\eta)} & \frac{1}{1+\exp(\eta)} \end{bmatrix}^\top$. Assume that $\eta \leq 1$. The total loss (and hence regret) of the learner is then at least

$$\sum_{k=1}^{\lfloor \frac{T}{2} \rfloor} \frac{1}{2} - \frac{1}{1 + \exp(\eta)} = \left\lfloor \frac{T}{2} \right\rfloor \left(\frac{1}{2} - \frac{1}{1 + \exp(\eta)} \right) \tag{41}$$

$$\geq \left\lfloor \frac{T}{2} \right\rfloor \left(\frac{1}{2} - \frac{1}{2 + 2\eta} \right) \tag{42}$$

$$\geq \frac{1}{4} \left\lfloor \frac{T}{2} \right\rfloor \eta. \tag{43}$$

So, for any $\eta \leq 1$, there is a quasi-realizable sequence with regret at least $\frac{1}{4} \lfloor \frac{T}{2} \rfloor \eta$. Since (41) can be seen to be an increasing function of η, we have a lower bound of $\frac{1}{4} \lfloor \frac{T}{2} \rfloor \min(\eta, 1)$. The point is that for large η, the learner will pay heavily because it switches around too much.

Sequence 2. On the other hand, we consider the sequence given by $z_{t,2} = 1$ for all t. Then $w_{t,2} = \frac{1}{1+\exp((t-1)\eta)}$, which for $t \leq \lceil \frac{1}{\eta} \rceil$ is at least $\frac{1}{1+e}$. Therefore, the regret of the learner on this sequence is at least $\frac{1}{1+e} \min\left(T, \frac{1}{\eta}\right)$. The point is that for small η, the learner will pay heavily because it can't decrease the weight on expert 2 fast enough.

Combining these together, we see that the first sequence inflicts a regret of $\Omega(\sqrt{T})$ whenever $\eta \geq 1/\sqrt{T}$, whereas the second sequence inflicts a regret of $\Omega(1/\sqrt{T})$ whenever $\eta \leq 1/\sqrt{T}$. Since one of these two conditions on η must always be satisfied, one of these sequences will always inflict regret $\Omega(1/\sqrt{T})$, thus proving the proposition. $\qquad\square$

Proof of Proposition 4.1. As noted in the main text, the proof parallels Proposition 3.3, with the main new tool being the Golden-Thompson inequality, which says that $\text{tr}(\exp(A+B)) \leq \text{tr}(\exp(A)\exp(B))$ (Golden, 1965; Thompson, 1965).

When $\psi(W) = \text{tr}(W \log(W))$ and W is constrained to have trace 1, we have $\psi^*(B) = \log(\text{tr}(\exp(B)))$ and $\nabla \psi^*(B) = \frac{\exp(B)}{\text{tr}(\exp(B))}$, so that $\nabla \psi^*(B_t - \eta M_t)$ matches W_t as given in the proposition. So, again, we are performing an instance of Algorithm 2 and it suffices to check that the condition of Corollary 3.2 is satisfied for $A_t = (Z_t - M_t)^2$. To do so, we use the Golden-Thompson inequality together with the fact that $-X - X^2 \preceq \log(I - X)$ for $-\frac{1}{2}I \preceq X \preceq \frac{1}{2}I$. We have

$$
\psi^*(B_t - \eta Z_t - \eta^2 A_t)
$$
$$
= \log(\text{tr}(\exp(B_t - \eta Z_t - \eta^2 (Z_t - M_t)^2)))
$$
$$
\leq \log(\text{tr}(\exp(B_t - \eta M_t) \exp(-\eta(Z_t - M_t) - \eta^2(Z_t - M_t)^2)))
$$
$$
\leq \log(\text{tr}(\exp(B_t - \eta M_t)(I - \eta(Z_t - M_t))))
$$
$$
= \log(\text{tr}(\exp(B_t - \eta M_t)) - \eta \, \text{tr}(\exp(B_t - \eta M_t)(Z_t - M_t)))
$$
$$
\leq \log(\text{tr}(\exp(B_t - \eta M_t))) - \eta \frac{\text{tr}(\exp(B_t - \eta M_t)(Z_t - M_t))}{\text{tr}(\exp(B_t - \eta M_t))}
$$
$$
= \psi^*(B_t - \eta M_t) - \eta \langle \nabla \psi^*(B_t - \eta M_t), Z_t - M_t \rangle.
$$

This verifies the condition of Corollary 3.2, so that we have a regret bound of $\frac{\psi^*(0) + \psi(U)}{\eta} + \eta \sum_{t=1}^{T} \text{tr}(U A_t)$. Finally, noting that $\psi^*(0) = \log(n)$, $\psi(U) = \text{tr}(U \log(U)) \leq 0$, and $A_t = (Z_t - M_t)^2$ completes the proof. $\qquad\square$

Proof of Lemma 4.2. Write $M' = M^* + D$. Then we have

$$
\delta(M')^2 + \sum_{t=1}^{T}(Z_t - M')^2 \tag{44}
$$

$$
= \delta(M^* + D)^2 + \sum_{t=1}^{T}(Z_t - M^* - D)^2 \tag{45}
$$

$$
= \delta(M^*)^2 + \sum_{t=1}^{T}(Z_t - M^*)^2 + \left[\delta M^* + \sum_{t=1}^{T}(M^* - Z_t)\right]D + D\left[\delta M^* + \sum_{t=1}^{T}(M^* - Z_t)\right] + (T+\delta)D^2 \tag{46}
$$

$$
= \delta(M^*)^2 + \sum_{t=1}^{T}(Z_t - M^*)^2 + (T+\delta)D^2 \tag{47}
$$

$$
\succeq \delta(M^*)^2 + \sum_{t=1}^{T}(Z_t - M^*)^2, \tag{48}
$$

which completes the lemma. □

Proof of Lemma 4.3. The proof is structurally identical to the vector case (see Hazan (2011) for a proof of the vector case). We will prove the lemma by induction on T. Note that the lemma is equivalent to showing that

$$\psi(M_1) + \sum_{t=1}^{T} f_t(M_{t+1}) \leq_{\mathcal{K}} \psi(M) + \sum_{t=1}^{T} f_t(M) \tag{49}$$

for all M. In the base case $T = 0$, we have

$$\psi(M_1) \leq_{\mathcal{K}} \psi(M), \tag{50}$$

which follows from the fact that M_1 is a global minimizer of ψ and hence $\psi(M_1) \leq_{\mathcal{K}} \psi(M)$ for all M. For the inductive step, suppose that

$$\sum_{t=1}^{T-1} f_t(M_{t+1}) \leq_{\mathcal{K}} \psi(M) + \sum_{t=1}^{T-1} f_t(M) \tag{51}$$

for all M, and invoke this for the particular choice $M = M_{T+1}$. Then we have

$$\psi(M_1) + \sum_{t=1}^{T} f_t(M_{t+1}) = \psi(M_1) + \left[\sum_{t=1}^{T-1} f_t(M_{t+1})\right] + f_T(M_{T+1}) \tag{52}$$

$$\leq_{\mathcal{K}} \psi(M_{T+1}) + \left[\sum_{t=1}^{T-1} f_t(M_{T+1})\right] + f_T(M_{T+1}) \tag{53}$$

$$= \psi(M_{T+1}) + \sum_{t=1}^{T} f_t(M_{T+1}) \tag{54}$$

$$\leq_{\mathcal{K}} \psi(M) + \sum_{t=1}^{T} f_t(M) \tag{55}$$

for all M, where we use the fact that M_{T+1} is a global minimizer of $\psi(M) + \sum_{t=1}^{T} f_t(M)$ for the last inequality. This completes the induction and hence the proof. □

Proof of Corollary 4.4. The key tool is the *matrix Young's inequality*: $AB + BA \preceq \frac{1}{\gamma}A^2 + \gamma B^2$ for all symmetric A, B and all $\gamma > 0$. (This follows immediately upon expanding $(A/\sqrt{\gamma} - \sqrt{\gamma}B)^2 \succeq 0$.) We then note that, by Lemma 4.2, M_t obeys Lemma 4.3 with $\psi(M) = M^2$, $f_t(M) = (M - Z_t)^2$, and $\mathcal{K}$ the cone of positive semidefinite matrices. Therefore:

$$\sum_{t=1}^{T}(Z_t - M_t)^2 - (Z_t - \bar{Z})^2 \preceq \bar{Z}^2 + \sum_{t=1}^{T}(Z_t - M_t)^2 - (Z_t - M_{t+1})^2 \tag{56}$$

$$= \bar{Z}^2 + \sum_{t=1}^{T}\left[Z_t(M_{t+1} - M_t) + (M_{t+1} - M_t)Z_t + M_t^2 - M_{t+1}^2\right] \tag{57}$$

$$= \bar{Z}^2 + M_1^2 - M_{T+1}^2 + \sum_{t=1}^{T}\left[Z_t(M_{t+1} - M_t) + (M_{t+1} - M_t)Z_t\right] \tag{58}$$

$$= \bar{Z}^2 + M_1^2 - M_{T+1}^2 + \sum_{t=1}^{T}\frac{1}{t+1}\left[Z_t(Z_t - M_t) + (Z_t - M_t)Z_t\right] \tag{59}$$

$$\text{(since } M_{t+1} = \tfrac{1}{t+1}Z_t + \tfrac{t}{t+1}M_t\text{)}$$

$$\preceq I + \sum_{t=1}^{T} \frac{Z_t^2}{\gamma(t+1)^2} + \gamma(Z_t - M_t)^2 \tag{60}$$

$$\preceq I + \frac{I}{\gamma} + \gamma \sum_{t=1}^{T}(Z_t - M_t)^2. \tag{61}$$

(For the second-to-last inequality, note that $M_1 = 0$ and hence $M_1^2 - M_{T+1}^2 \preceq 0$.) Re-arranging yields

$$\sum_{t=1}^{T}(Z_t - M_t)^2 \leq \frac{1}{1-\gamma}\left(\frac{1+\gamma}{\gamma}I + \sum_{t=1}^{T}(Z_t - \bar{Z})^2\right). \tag{62}$$

Setting γ to $\frac{1}{2}$ gives the desired result. Note that by instead setting γ to $\frac{\epsilon}{2}$, we can replace the constants 2 and 6 by $1 + \epsilon$ and $\frac{6}{\epsilon}$ for any $\epsilon \leq 1$. $\qquad\square$

C. Improved Variance Bound

We claimed in Section 3 that we could obtain a regret bound in terms of $2V_i + 6$ by using the optimistic prediction based on $m_t = \frac{1}{t}\sum_{s=1}^{t-1} z_s$. The following proposition establishes this. Its proof is essentially the same as that of Corollary 4.5, and in fact is implied by Corollary 4.5. The only purpose of this section is to keep proofs accessible to readers who prefer not to read through algebraic manipulations of matrices.

Proposition C.1. *Suppose that we choose $m_{t,i} = \frac{1}{t}\sum_{s=1}^{t-1} z_{s,i}$ and that $\|z_s\|_\infty \leq 1$. Then for all i and all $0 < \epsilon \leq 1$ we have*

$$\sum_{t=1}^{T}(z_{t,i} - m_{t,i})^2 \leq 2\sum_{t=1}^{T}(z_{t,i} - m_i^*)^2 + 6. \tag{63}$$

Proof. Note that $m_{t,i}$ is the minimizer of $m_i^2 + \sum_{s=1}^{t-1}(z_{s,i} - m_i)^2$. Therefore, by the FTRL Lemma (Hazan, 2011), we have

$$\sum_{t=1}^{T}(z_{t,i} - m_{t,i})^2 - (z_{t,i} - m_i^*)^2 \leq (m_i^*)^2 + \sum_{t=1}^{T}(z_{t,i} - m_{t,i})^2 - (z_{t,i} - m_{t+1,i})^2 \tag{64}$$

$$= (m_i^*)^2 + \sum_{t=1}^{T} 2z_{t,i}(m_{t+1,i} - m_{t,i}) + m_{t,i}^2 - m_{t+1,i}^2 \tag{65}$$

$$= (m_i^*)^2 + m_{1,i}^2 - m_{T+1,i}^2 + \sum_{t=1}^{T}\frac{2}{t+1}z_{t,i}(z_{t,i} - m_{t,i}) \tag{66}$$

$$\leq 1 + \sum_{t=1}^{T}\frac{z_{t,i}^2}{\gamma(t+1)^2} + \gamma(z_{t,i} - m_{t,i})^2 \tag{67}$$

$$\leq 1 + \frac{1}{\gamma} + \gamma\sum_{t=1}^{T}(z_{t,i} - m_{t,i})^2. \tag{68}$$

Re-arranging yields

$$\sum_{t=1}^{T}(z_{t,i} - m_{t,i})^2 \leq \frac{1}{1-\gamma}\left(\frac{1+\gamma}{\gamma} + \sum_{t=1}^{T}(z_{t,i} - m_i^*)^2\right). \tag{69}$$

Setting γ to $\frac{1}{2}$ then yields the desired result. $\qquad\square$

D. Bounds for Exponentiated Gradient in the Unconstrained Case

The main text contained an analysis of adaptive versions of the exponentiated gradient and matrix exponentiated gradient algorithms. However, this analysis was for the case that the weights were constrained to the simplex (or that $\mathrm{tr}(W) = 1$ in the case of matrices). In Section 2 we promised to include an analysis of these algorithms in the unconstrained case, and we do so here. Note that this "unconstrained case" still has the constraint $w \geq 0$ (or $W \succeq 0$ for matrices), although this is not a serious limitation since we can split w into its positive and negative components (see Kivinen & Warmuth (1997) for details).

The updates and proofs are almost identical. The major difference is in the initialization, where to obtain good bounds we need to initialize $\beta_{1,i}$ to $-\log(n)$ rather than 0 (in the matrix case, we need to initialize B_1 to $-\log(n)I$). The complete algorithms are shown below:

Exponentiated Gradient:

$$
\begin{aligned}
\beta_{1,i} &= -\log(n) \\
w_{t,i} &= \exp(\beta_{t,i} - \eta m_{t,i}) \\
\beta_{t+1,i} &= \beta_{t,i} - \eta z_{t,i} - \eta^2(z_{t,i} - m_{t,i})^2
\end{aligned}
\tag{70}
$$

Matrix Exponentiated Gradient:

$$
\begin{aligned}
B_1 &= -\log(n)I \\
W_t &= \exp(B_t - \eta M_t) \\
B_{t+1} &= B_t - \eta Z_t - \eta^2(Z_t - M_t)^2
\end{aligned}
\tag{71}
$$

We have the following regret bounds in the vector and matrix cases:

Proposition D.1. *For $\|z_t\|_\infty \leq 1$, $\|m_t\|_\infty \leq 1$, and $0 < \eta \leq \frac{1}{4}$, the unconstrained exponentiated gradient updates (70) achieve the bound*

$$
\mathrm{Regret}(u) \leq \frac{1 + (\log(n) - 1)\|u\|_1 + \sum_{i=1}^n u_i \log(u_i)}{\eta} + \eta \sum_{i=1}^n u_i \sum_{t=1}^T (z_{t,i} - m_{t,i})^2.
\tag{72}
$$

Proposition D.2. *For $\|Z_t\|_{\mathrm{op}} \leq 1$, $\|M_t\|_{\mathrm{op}} \leq 1$, and $0 < \eta \leq \frac{1}{4}$, the unconstrained matrix exponentiated gradient updates (71) achieve the bound*

$$
\mathrm{Regret}(U) \leq \frac{1 + (\log(n) - 1)\,\mathrm{tr}(U) + \mathrm{tr}(U\log(U))}{\eta} + \eta \sum_{t=1}^T \mathrm{tr}(U(Z_t - M_t)^2).
\tag{73}
$$

The proofs are basically identical to the proofs of Propositions 3.3 and 4.1, but we include them for completeness.

Proof of Proposition D.1. We note that, for $\psi(w) = \sum_{i=1}^n w_i \log(w_i) - w_i$ and w constrained to be non-negative, $\psi^*(\beta) = \sum_{i=1}^n \exp(\beta_i)$ and $\nabla \psi^*(\beta_t - \eta m_t)$ is equal to w_t as defined in the proposition. It therefore suffices to check that the condition of Corollary 3.2 is satisfied with $a_{t,i} = (z_{t,i} - m_{t,i})^2$. We have

$$
\psi^*(\beta_t - \eta z_t - \eta^2 a_t) = \sum_{i=1}^n \exp(\beta_{t,i} - \eta z_{t,i} - \eta^2(z_{t,i} - m_{t,i})^2)
\tag{74}
$$

$$
= \sum_{i=1}^n \exp(\beta_{t,i} - \eta m_{t,i}) \exp(-\eta(z_{t,i} - m_{t,i}) - \eta^2(z_{t,i} - m_{t,i})^2)
\tag{75}
$$

$$
\leq \sum_{i=1}^n \exp(\beta_{t,i} - \eta m_{t,i})(1 - \eta(z_{t,i} - m_{t,i}))
\tag{76}
$$

$$= \sum_{i=1}^{n} \exp(\beta_{t,i} - \eta m_{t,i}) - \eta \sum_{i=1}^{n} \exp(\beta_{t,i} - \eta m_{t,i})(z_{t,i} - m_{t,i}) \tag{77}$$

$$= \psi^*(\beta_t - \eta m_t) - \eta \nabla \psi^*(\beta_t - \eta m_t)^\top (z_t - m_t). \tag{78}$$

The one inequality we made use of was $\exp(-x - x^2) \le 1 - x$ for $|x| < \frac{1}{2}$. This verifies the condition of Corollary 3.2, yielding a regret bound of $\frac{\psi^*(\beta_1) + \psi(u) - u^\top \beta_1}{\eta} + \eta \sum_{i=1}^{n} u^\top a_t$. Finally, we note that $\psi^*(\beta_1) = \sum_{i=1}^{n} \exp(-\log(n)) = 1$, $\psi(u) - u^\top \beta_1 = \sum_{i=1}^{n} u_i \log(u_i) + (\log(n) - 1)u_i$, and $a_{t,i} = (z_{t,i} - m_{t,i})^2$, which completes the proof. $\qquad \square$

Proof of Proposition D.2. When $\psi(W) = \text{tr}(W \log(W) - W)$ and W is constrained to be positive semidefinite, we have $\psi^*(B) = \log(\text{tr}(\exp(B)))$ and $\nabla \psi^*(B) = \exp(B)$, so that $\nabla \psi^*(B_t - \eta M_t)$ matches W_t as given in the proposition. So, again, it suffices to check that the condition of Corollary 3.2 is satisfied for $A_t = (Z_t - M_t)^2$. To do so, we need to make use of the Golden-Thompson inequality $\text{tr}(\exp(A + B)) \preceq \text{tr}(\exp(A) \exp(B))$ (Golden, 1965; Thompson, 1965), together with the fact that $-X - X^2 \preceq \log(I - X)$ for $-\frac{1}{2}I \preceq X \preceq \frac{1}{2}I$. We then have

$$\psi^*(B_t - \eta Z_t - \eta^2 A_t) = \text{tr}(\exp(B_t - \eta Z_t - \eta^2(Z_t - M_t)^2)) \tag{79}$$

$$\le \text{tr}(\exp(B_t - \eta M_t) \exp(-\eta(Z_t - M_t) - \eta^2(Z_t - M_t)^2)) \tag{80}$$

$$\le \text{tr}(\exp(B_t - \eta M_t)(I - \eta(Z_t - M_t))) \tag{81}$$

$$= \text{tr}(\exp(B_t - \eta M_t)) - \eta \text{tr}(\exp(B_t - \eta M_t)(Z_t - M_t)) \tag{82}$$

$$= \psi^*(B_t - \eta M_t) - \eta \langle \nabla \psi^*(B_t - \eta M_t), Z_t - M_t \rangle. \tag{83}$$

This verifies the condition of Corollary 3.2, so that we have a regret bound of $\frac{\psi^*(B_1) + \psi(U) - \text{tr}(B_1 U)}{\eta} + \eta \sum_{t=1}^{T} \text{tr}(U A_t)$. Finally, noting that $\psi^*(B_1) = \text{tr}(\frac{1}{n}I) = 1$, $\psi(U) - \text{tr}(B_1 U) = \text{tr}(U \log(U)) + (\log(n) - 1)\text{tr}(U)$, and $A_t = (Z_t - M_t)^2$ completes the proof. $\qquad \square$

E. Adaptive Step Size

In this section we show how to obtain an adaptive version of Algorithm 2, which relies on the standard doubling trick. The adaptive algorithm is given as Algorithm 3. The regret bound of this procedure when applied to learning from experts is worse than in the non-adaptive case, depending (in the language of Figure 1 and (24)) on $\max_i D_i$ rather than D_{i^*} (in other words, the maximum path length of any expert rather than the path length of the best expert).

The algorithm basically calls Algorithm 2 repeatedly with different step sizes, halving the step size every time the regret exceeds a certain bound. For this algorithm we require a bound B on the inner product term $u^\top z_t$ and a bound C on the regularizer term in the regret bound. Cesa-Bianchi et al. (2007) proposed an adaptive step size scheme in the learning from experts setting that does not require knowledge of B. It would be interesting to apply the same ideas here, but we have not tried to do so, although the exposition given below follows Section 3.1 of the same paper.

The regret of Algorithm 3 is bounded in the following theorem:

Theorem E.1. *Let* $u_t \in \arg\min_u u^\top \sum_{s=1}^{t} z_s$ *and let* $Q_t = u_t^\top \sum_{s=1}^{t} a_s$. *Let* $Q = \max(B, \max_{t=1}^{T} Q_t)$. *Then the regret of Algorithm 3 is bounded as*

$$\text{Regret} \le B \left[1 + \log \left(\frac{Q}{B} \right) \right] + 10\sqrt{CQ}. \tag{84}$$

Proof. First note η is monotonically non-increasing across rounds, and decays by a factor of 2 every time it changes. We can group the rounds based on what value of η was used in that round; in this way, Algorithm 3 is equivalent to running several sub-algorithms, each of which is an instance of Algorithm 2. The total regret is then bounded above by the sum of the regrets of these individual algorithms.

Now consider the rounds when η is equal to $2^{-j}\sqrt{\frac{C}{B}}$. Let t_j be the final such round. By construction, we must have $u_{t-1}^\top \sum_{s=1}^{t-1} a_s \le 4^{j+1}B$, or else we would have already decreased η by the next factor of 2. Let Regret_j denote the regret

---

Algorithm 3 Adaptive Step Size Mirror Descent

Given: convex regularizer ψ, corrections a_t, hints m_t, and β
Let B be any bound on $\max_{t=1}^{T} u^\top z_t$
Let C be any upper bound on $\psi^*(\beta_1) + \psi(u) - u^\top \beta_1$
$Q, \eta, t \leftarrow B, \sqrt{\frac{C}{B}}, 1$
while there are rounds remaining **do**
$\quad \beta_t \leftarrow \beta$
$\quad$ **while** $\sqrt{\frac{C}{Q}} \geq \frac{\eta}{2}$ **do**
$\qquad$ Choose $w_t = \nabla \psi^*(\beta_t - \eta_t m_t)$
$\qquad$ Observe z_t and suffer loss $w_t^\top z_t$
$\qquad$ Update $\beta_{t+1} = \beta_t - \eta_t z_t - \eta_t^2 a_t$
$\qquad$ Let $u_t \in \arg\min_u u^\top \sum_{s=1}^{t} z_s$
$\qquad Q \leftarrow \max(Q, u_t^\top \sum_{s=1}^{t} a_s)$
$\qquad t \leftarrow t + 1$
$\quad$ **end while**
$\quad \eta \leftarrow \frac{\eta}{2}$
end while

---

of the sub-algorithm on this set of rounds. Note that it is bounded above by B plus the regret on all but the last of these rounds. Then we have

$$\text{Regret}_j \leq B + \frac{C}{\eta} + \eta u_{t-1}^\top \sum_{s=1}^{t-1} a_s \tag{85}$$

$$= B + 2^j \sqrt{CB} + 2^{-j} 4^{j+1} \sqrt{CB} \tag{86}$$

$$= B + 5 \cdot 2^j \sqrt{CB} \tag{87}$$

$$\leq B + 5 \sqrt{CQ_{t_j}}. \tag{88}$$

Note that $\sqrt{Q_{t_j}} \geq 2\sqrt{Q_{t_{j-1}}}$ by construction. Then we have

$$\text{Regret} \leq \sum_j \text{Regret}_j \tag{89}$$

$$\leq \sum_j B + 5\sqrt{CQ_{t_j}} \tag{90}$$

$$\leq B\left\lceil 1 + \log\left(\frac{Q}{B}\right)\right\rceil + 10\sqrt{CQ}, \tag{91}$$

as was to be shown. $\qquad\square$