Supplementary material:
Online Adaptive Principal Component Analysis and Its extensions

The supplementary material contains proofs of the main results of the paper along with supporting results.

Before presenting the proofs, we need the following lemma from previous literature:

Lemma 4. *(Freund & Schapire, 1997) Suppose* $0 \leq L \leq \tilde{L}$ *and* $0 < R \leq \tilde{R}$*. Let* $\beta = g(\tilde{L}/\tilde{R})$ *where* $g(z) = 1/(1 + \sqrt{2/z})$*. Then*

$$\frac{-L \ln \beta + R}{1 - \beta} \leq L + \sqrt{2\tilde{L}\tilde{R}} + R$$

Additionally, we need the following classic bound on traces for postive semidefinite matrices. See, e.g. (Tsuda et al., 2005).

Lemma 5. *For any positive semi-definite matrix* A *and any symmetric matrices* B *and* C*,* $B \preceq C$ *implies* $\mathrm{Tr}(AB) \leq \mathrm{Tr}(AC)$*.*

A. Proof of Theorem 1

Proof. Fix $1 \leq r \leq s \leq T$. We set $\mathbf{q_t} = \mathbf{q} \in \mathcal{B}^n_{n-k}$ for $t = r, \ldots, s$ and 0 elsewhere. Thus, we have that $\|\mathbf{q_t}\|_1$ is either 0 or 1.

According to Lemma 1, for both cases of $\mathbf{q_t}$, we have

$$\|\mathbf{q_t}\|_1 \mathbf{w_t}^T \ell_\mathbf{t} (1 - \exp(-\eta)) - \eta \mathbf{q_t}^T \ell_\mathbf{t} \leq \sum_{i=1}^n q_{t,i} \ln(\frac{v_{t+1,i}}{\hat{w}_{t,i}}) \tag{32}$$

The analysis for $\sum_{i=1}^n q_{t,i} \ln(\frac{v_{t+1,i}}{\hat{w}_{t,i}})$ follows the Proof of Proposition 2 in (Cesa-Bianchi et al., 2012b). We describe the steps for completeness, since it is helpful for understanding the effect of the fixed-share step, Eq.(4b). This analysis will be crucial for the understanding how the fixed-share step can be applied to PCA problems.

$$\sum_{i=1}^n q_{t,i} \ln(\frac{v_{t+1,i}}{\hat{w}_{t,i}}) = \underbrace{\sum_{i=1}^n \left(q_{t,i} \ln \frac{1}{\hat{w}_{t,i}} - q_{t-1,i} \ln \frac{1}{v_{t,i}}\right)}_{A} + \underbrace{\sum_{i=1}^n \left(q_{t-1,i} \ln \frac{1}{v_{t,i}} - q_{t,i} \ln \frac{1}{v_{t+1,i}}\right)}_{B} \tag{33}$$

For the expression of A, we have

$$A = \sum_{i: q_{t,i} \geq q_{t-1,i}} \left((q_{t,i} - q_{t-1,i}) \ln \frac{1}{\hat{w}_{t,i}} + q_{t-1,i} \ln \frac{v_{t,i}}{\hat{w}_{t,i}}\right) + \sum_{i: q_{t,i} < q_{t-1,i}} \left(\underbrace{(q_{t,i} - q_{t-1,i}) \ln \frac{1}{v_{t,i}}}_{\leq 0} + q_{t,i} \ln \frac{v_{t,i}}{\hat{w}_{t,i}}\right) \tag{34}$$

Based on the update in Eq.(4), we have $1/\hat{w}_{t,i} \leq n/\alpha$ and $v_{t,i}/\hat{w}_{t,i} \leq 1/(1-\alpha)$. Plugging the bounds into the above equation, we have

$$A \leq \underbrace{\sum_{i: q_{t,i} \geq q_{t-1,i}} (q_{t,i} - q_{t-1,i})}_{= D_{TV}(\mathbf{q_t}, \mathbf{q_{t-1}})} \ln \frac{n}{\alpha} + \underbrace{\Big(\sum_{i: q_{t,i} \geq q_{t-1,i}} q_{t-1,i} + \sum_{i: q_{t,i} < q_{t-1,i}} q_{t,i}\Big)}_{= \|\mathbf{q_t}\|_1 - D_{TV}(\mathbf{q_t}, \mathbf{q_{t-1}})} \ln \frac{1}{1-\alpha}. \tag{35}$$

Telescoping the expression of B, substituting the above inequality in Eq.(33), and summing over $t = 2, \ldots, T$, we have

$$\sum_{t=2}^T \sum_{i=1}^n q_{t,i} \ln \frac{v_{t+1,i}}{\hat{w}_{t,i}} \leq m(\mathbf{q_{1:T}}) \ln \frac{n}{\alpha} + \Big(\sum_{t=2}^T \|\mathbf{q_t}\|_1 - m(\mathbf{q_{1:T}})\Big) \ln \frac{1}{1-\alpha} + \sum_{i=1}^n q_{1,i} \ln \frac{1}{v_{2,i}}. \tag{36}$$

Adding the $t = 1$ term to the above inequality, we have

$$\sum_{t=1}^T \sum_{i=1}^n q_{t,i} \ln \frac{v_{t+1,i}}{\hat{w}_{t,i}} \leq \|\mathbf{q_1}\|_1 \ln(n) + m(\mathbf{q_{1:T}}) \ln \frac{n}{\alpha} + \Big(\sum_{t=1}^T \|\mathbf{q_t}\|_1 - m(\mathbf{q_{1:T}})\Big) \ln \frac{1}{1-\alpha}. \tag{37}$$

Now we bound the right side, using the choices for $\mathbf{q_t}$ described at the beginning of the proof. If $r \geq 2$, $m(\mathbf{q_{1:T}}) = 1$, and $\|\mathbf{q_1}\|_1 = 0$. If $r = 1$, $m(\mathbf{q_{1:T}}) = 0$, and $\|\mathbf{q_1}\|_1 = 1$. Thus, $m(\mathbf{q_{1:T}}) + \|\mathbf{q_1}\|_1 = 1$, and the right part can be upper bounded by $\ln \frac{n}{\alpha} + T \ln \frac{1}{1-\alpha}$.

Combine the above inequality with Eq.(32), set $\mathbf{q_t} = \mathbf{q} \in \mathcal{B}^{\mathrm{n}}_{\mathrm{n\text{-}k}}$ for $t = r, \ldots, s$ and 0 elsewhere, and multiply both sides by $n-k$, we have

$$\begin{aligned}(1-\exp(-\eta))\sum_{t=r}^{s}(n-k)\mathbf{w_t}^T\ell_\mathbf{t}-\eta\sum_{t=r}^{s}(n-k)\mathbf{q}^T\ell_\mathbf{t}\\ \leq (n-k)\ln\tfrac{n}{\alpha}+(n-k)T\ln\tfrac{1}{1-\alpha}\end{aligned} \tag{38}$$

If we set $\alpha = 1/(1+(n-k)T)$, then the right part can be upper bounded by $(n-k)\ln(n(1+(n-k)T))+1$, which equals to D as defined in the Theorem 1. Thus, the above inequality can be reformulated as

$$\sum_{t=r}^{s}(n-k)\mathbf{w_t}^T\ell_\mathbf{t} \leq \frac{\eta\sum_{t=r}^{s}(n-k)\mathbf{q}^T\ell_\mathbf{t}+D}{1-\exp(-\eta)} \tag{39}$$

Since the above inequality holds for arbitrary $\mathbf{q}\in\mathcal{B}^{\mathrm{n}}_{\mathrm{n\text{-}k}}$, we have

$$\sum_{t=r}^{s}(n-k)\mathbf{w_t}^T\ell_\mathbf{t} \leq \frac{\eta\min_{\mathbf{q}\in\mathcal{B}^{\mathrm{n}}_{\mathrm{n\text{-}k}}}\sum_{t=r}^{s}(n-k)\mathbf{q}^T\ell_\mathbf{t}+D}{1-\exp(-\eta)} \tag{40}$$

We will apply the inequality in Lemma 4 to upper bound the right part in Eq.(40). With $\min_{\mathbf{q}\in\mathcal{B}^{\mathrm{n}}_{\mathrm{n\text{-}k}}}\sum_{t=r}^{s}(n-k)\mathbf{q}^T\ell_\mathbf{t}\leq L$ and $\eta=\ln(1+\sqrt{2D/L})$, we have

$$\sum_{t=r}^{s}(n-k)\mathbf{w_t}^T\ell_\mathbf{t}-\min_{\mathbf{q}\in\mathcal{B}^{\mathrm{n}}_{\mathrm{n\text{-}k}}}\sum_{t=r}^{s}(n-k)\mathbf{q}^T\ell_\mathbf{t}\leq\sqrt{2LD}+D \tag{41}$$

Since the above inequality always holds for all intervals, $[r,s]$, the result is proved by maximizing the left side over $[r,s]$. □

B. Proof of Theorem 3

Proof. In the proof, we will use two cases of Q_t: $Q_t\in\mathcal{B}^n_{n-k}$, and $Q_t=0$.

We first apply the eigendecomposition to Q_t as $Q_t = \widetilde{D}\operatorname{diag}(\mathbf{q_t})\widetilde{D}^T$, where $\widetilde{D}=[\tilde{\mathbf{d}}_\mathbf{1},\ldots,\tilde{\mathbf{d}}_\mathbf{n}]$. Since in the adaptive setting, Q_{t-1} is either equal to Q_t or 0, they share the same eigenvectors and Q_{t-1} can be expressed as $Q_{t-1}=\widetilde{D}\operatorname{diag}(\mathbf{q_{t-1}})\widetilde{D}^T$.

According to Lemma 2, the following inequality is true for both cases of Q_t:

$$\begin{aligned}\|\mathbf{q_t}\|_1\operatorname{Tr}(W_t\mathbf{x_t}\mathbf{x_t}^T)(1-\exp(-\eta))-\eta\operatorname{Tr}(Q_t\mathbf{x_t}\mathbf{x_t}^T)\\ \leq -\operatorname{Tr}(Q_t\ln\widehat{W}_t)+\operatorname{Tr}(Q_t\ln V_{t+1})\end{aligned} \tag{42}$$

The next steps extend proof of Proposition 2 in (Cesa-Bianchi et al., 2012b) to the matrix case.

We analyze the right part of the above inequality, which can be expressed as:

$$-\operatorname{Tr}(Q_t\ln\widehat{W}_t)+\operatorname{Tr}(Q_t\ln V_{t+1})=\bar{A}+\bar{B} \tag{43}$$

where $\bar{A}=-\operatorname{Tr}(Q_t\ln\widehat{W}_t)+\operatorname{Tr}(Q_{t-1}\ln V_t)$, and $\bar{B}=-\operatorname{Tr}(Q_{t-1}\ln V_t)+\operatorname{Tr}(Q_t\ln V_{t+1})$.

We will first upper bound the $\bar{A}$ term, and then telescope the $\bar{B}$ term.

$\bar{A}$ can be expressed as:

$$\begin{aligned}\bar{A}=&\sum_{i:q_{t,i}\geq q_{t-1,i}}\Big(\underbrace{-\operatorname{Tr}\big((q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T-q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)\ln\widehat{W}_t\big)}_{①}\\ &+\underbrace{\operatorname{Tr}(q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T\ln V_t)-\operatorname{Tr}(q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T\ln\widehat{W}_t)}_{②}\Big)\\ &+\sum_{i:q_{t,i}<q_{t-1,i}}\Big(\underbrace{-\operatorname{Tr}\big((q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T-q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)\ln V_t\big)}_{③}\\ &+\underbrace{\operatorname{Tr}(q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T\ln V_t)-\operatorname{Tr}(q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T\ln\widehat{W}_t)}_{④}\Big)\end{aligned} \tag{44}$$

For ①, it can be expressed as:

$$\begin{aligned}①&=\operatorname{Tr}\big((q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T-q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^\mathbf{T})\ln\widehat{W}_t^{-1}\big)\\ &\leq\operatorname{Tr}\big((q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T-q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)\ln\tfrac{n}{\alpha}\big)\\ &=(q_{t,i}-q_{t-1,i})\ln\tfrac{n}{\alpha}.\end{aligned} \tag{45}$$

The inequality holds because the update in Eq.(14b) implies $\ln\widehat{W}_T^{-1}\preceq I\ln\frac{n}{\alpha}$ and furthermore, $(q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T-q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)$ is positive semi-definite. Thus, Lemma 5, gives the result.

The expression for ② can be bounded as

$$\begin{aligned}②&=\operatorname{Tr}(q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T\ln(V_t\widehat{W}_t^{-1}))\\ &\leq q_{t-1,i}\ln\tfrac{1}{1-\alpha}\end{aligned} \tag{46}$$

where the equality is due to the fact that V_t and $\widehat{W}_t$ have the same eigenvectors. The inequality follows since $\ln(V_t\widehat{W}_t^{-1})\preceq I\ln\frac{1}{1-\alpha}$, due to the update in Eq.(14b), while $q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T$ is positive semi-definite. Thus Lemma 5 gives the result.

The bound ③ can be expressed as:

$$③=\operatorname{Tr}\big((-q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T+q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)\ln V_t\big)\leq 0 \tag{47}$$

Here, the inequality follows since $\ln V_t\preceq 0$ and and $(-q_{t,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T+q_{t-1,i}\tilde{\mathbf{d}}_\mathbf{i}\tilde{\mathbf{d}}_\mathbf{i}^T)$ is positive semi-definite. Thus, Lemma 5 gives the result.

For ④, we have ④ $\leq q_{t,i} \ln \frac{1}{1-\alpha}$, which follows the same argument used to bound the term ②.

Thus, $\bar{A}$ can be upper bounded as follows:

$$\bar{A} \leq \underbrace{\sum_{i:q_{t,i} \geq q_{t-1,i}} (q_{t,i} - q_{t-1,i}) \ln \frac{n}{\alpha}}_{=D_{TV}(\mathbf{q_t}, \mathbf{q_{t-1}})} + \underbrace{\Big(\sum_{i:q_{t,i} \geq q_{t-1,i}} q_{t-1,i} + \sum_{i:q_{t,i} < q_{t-1,i}} q_{t,i} \Big)}_{=\|\mathbf{q_t}\|_1 - D_{TV}(\mathbf{q_t}, \mathbf{q_{t-1}})} \ln \frac{1}{1-\alpha} \tag{48}$$

Then we telescope the $\bar{B}$ term, substitute the above inequality for $\bar{A}$ into Eq.(43), and sum over $t = 2, \ldots, T$ to give:

$$\begin{aligned} &\sum_{t=2}^{T} \Big(-\operatorname{Tr}(Q_t \ln \widehat{W}_t) + \operatorname{Tr}(Q_t \ln V_{t+1}) \Big) \\ &\leq m(\mathbf{q_{1:T}}) \ln \tfrac{n}{\alpha} + \Big(\sum_{t=2}^{T} \|\mathbf{q_t}\|_1 - m(\mathbf{q_{1:T}}) \Big) \ln \tfrac{1}{1-\alpha} - \operatorname{Tr}(Q_1 \ln V_2) \end{aligned} \tag{49}$$

Adding the $t = 1$ term to the above inequality, we have

$$\begin{aligned} &\sum_{t=1}^{T} \Big(-\operatorname{Tr}(Q_t \ln \widehat{W}_t) + \operatorname{Tr}(Q_t \ln V_{t+1}) \Big) \\ &\leq \|\mathbf{q_1}\|_1 \ln(n) + m(\mathbf{q_{1:T}}) \ln \tfrac{n}{\alpha} + \Big(\sum_{t=1}^{T} \|\mathbf{q_t}\|_1 - m(\mathbf{q_{1:T}}) \Big) \ln \tfrac{1}{1-\alpha} \end{aligned} \tag{50}$$

For the above inequality, we set $Q_t = Q \in \mathcal{B}_{n-k}^n$ for $t = r, \ldots, s$ and 0 elsewhere, which makes $\mathbf{q_t} = \mathbf{q} \in \mathcal{B}_{n-k}^n$ for $t = r, \ldots, s$ and 0 elsewhere. If $r \geq 2$, $m(\mathbf{q_{1:T}}) = 1$, and $\|\mathbf{q_1}\|_1 = 0$. If $r = 1$, $m(\mathbf{q_{1:T}}) = 0$, and $\|\mathbf{q_1}\|_1 = 1$. Thus, $m(\mathbf{q_{1:T}}) + \|\mathbf{q_1}\|_1 = 1$, and the right part can be upper bounded by $\ln \frac{n}{\alpha} + T \ln \frac{1}{1-\alpha}$.

The rest of the steps follow exactly the same as in the proof of Theorem 1. □

C. Proof of Lemma 3

Proof. We first deal with the term $\operatorname{Tr}(Q_t \ln V_{t+1})$. According to the update in Eq.(24a), we have

$$\begin{aligned} \operatorname{Tr}(Q_t \ln V_{t+1}) &= \operatorname{Tr}\left(Q_t \ln \left(\tfrac{\exp(\ln Y_t - \eta C_t)}{\operatorname{Tr}(\exp(\ln Y_t - \eta C_t))} \right) \right) \\ &= \operatorname{Tr}\big(Q_t(\ln Y_t - \eta C_t) \big) - \ln\Big(\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big) \Big), \end{aligned} \tag{51}$$

since $Q_t \in \mathcal{B}_1^n$ and $\operatorname{Tr}(Q_t) = 1$.

As a result, we have $\operatorname{Tr}(Q_t \ln V_{t+1}) - \operatorname{Tr}(Q_t \ln Y_t) = -\eta \operatorname{Tr}(Q_t C_t) - \ln\Big(\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big) \Big)$.

Thus, to prove the inequality in Lemma 3, it is enough to prove the following inequality

$$\eta \operatorname{Tr}(Y_t C_t) - \frac{\eta^2}{2} + \ln\Big(\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big) \Big) \leq 0 \tag{52}$$

Before we proceed, we need the following lemmas:

Lemma 6 (Golden-Thompson inequality)**.** *For any symmetric matrices A and B, the following inequality holds:*

$$\operatorname{Tr}\big(\exp(A + B) \big) \leq \operatorname{Tr}\big(\exp(A) \exp(B) \big)$$

Lemma 7 (Lemma 2.1 in (Tsuda et al., 2005))**.** *For any symmetric matrix A such that $0 \preceq A \preceq I$ and any $\rho_1, \rho_2 \in \mathbb{R}$, the following holds:*

$$\exp\big(A\rho_1 + (I - A)\rho_2 \big) \preceq A \exp(\rho_1) + (I - A) \exp(\rho_2)$$

Then we apply the Golden-Thompson inequality to the term $\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big)$, which gives us the inequality below:

$$\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big) \leq \operatorname{Tr}(Y_t \exp(-\eta C_t)). \tag{53}$$

For the term $\exp(-\eta C_t)$, by applying the Lemma 7 with $\rho_1 = -\eta$ and $\rho_2 = 0$, we will have the following inequality:

$$\exp(-\eta C_t) \preceq I - C_t(1 - \exp(-\eta)). \tag{54}$$

Thus, we will have

$$\operatorname{Tr}(Y_t \exp(-\eta C_t)) \leq 1 - \operatorname{Tr}(Y_t C_t)(1 - \exp(-\eta)), \tag{55}$$

and

$$\operatorname{Tr}\big(\exp(\ln Y_t - \eta C_t) \big) \leq 1 - \operatorname{Tr}(Y_t C_t)(1 - \exp(-\eta)), \tag{56}$$

since $Y_t \in \mathcal{B}_1^n$ and $\operatorname{Tr}(Y_t) = 1$.

Thus, it is enough to prove the following inequality

$$\eta \operatorname{Tr}(Y_t C_t) - \frac{\eta^2}{2} + \ln\Big(1 - \operatorname{Tr}(Y_t C_t)(1 - \exp(-\eta)) \Big) \leq 0 \tag{57}$$

Since $\ln(1 - x) \leq -x$, we have

$$\ln\Big(1 - \operatorname{Tr}(Y_t C_t)(1 - \exp(-\eta)) \Big) \leq -\operatorname{Tr}(Y_t C_t)(1 - \exp(-\eta)). \tag{58}$$

Thus, it suffices to prove the following inequality:

$$\big(\eta - 1 + \exp(-\eta) \big) \operatorname{Tr}(Y_t C_t) - \frac{\eta^2}{2} \leq 0 \tag{59}$$

Note that by using convexity of $\exp(-\eta)$, $\eta - 1 + \exp(-\eta) \geq 0$.

By applying Lemma 5 with $A = Y_t$, $B = C_t$, and $C = I$, we have $\operatorname{Tr}(Y_t C_t) \leq \operatorname{Tr}(Y_t) = 1$. Thus, when $\eta \geq 0$, it is enough to prove the following inequality

$$\eta - 1 + \exp(-\eta) - \frac{\eta^2}{2} \leq 0. \tag{60}$$

This inequality follows from convexity of $\frac{\eta^2}{2} - \exp(-\eta)$ over $\eta \geq 0$. □

D. Proof of Theorem 6

Proof. First, since $0 \preceq C_t \preceq I$, we have $\max_{i,j} |C_t(i,j)| \leq 1$.

Before we proceed, we need the following lemma from (Warmuth & Kuzmin, 2006)

Lemma 8 (Lemma 1 in (Warmuth & Kuzmin, 2006))**.** *Let $\max_{i,j} |C_t(i,j)| \leq \frac{r}{2}$, then for any $\mathbf{u_t} \in \mathcal{B}_1^n$, any constants a and b such that $0 \leq a \leq \frac{b}{1+rb}$, and $\eta = \frac{2b}{1+rb}$, we have*

$$a{\mathbf{y_t}}^T C_t \mathbf{y_t} - b{\mathbf{u_t}}^T C_t \mathbf{u_t} \leq d(\mathbf{u_t}, \mathbf{y_t}) - d(\mathbf{u_t}, \mathbf{v_{t+1}})$$

Now we apply Lemma 8 under the conditions $r = 2$, $a = \frac{b}{2b+1}$, $\eta = 2a$, and $b = \frac{c}{2}$.

Recall that $d(\mathbf{u_t}, \mathbf{y_t}) - d(\mathbf{u_t}, \mathbf{v_{t+1}}) = \sum_i u_{t,i} \ln\left(\frac{v_{t+1,i}}{y_{t,i}}\right)$. Combining this with the inequality in Lemma 8 and the fact that $\|\mathbf{u_t}\|_1 = 1$, we have

$$a\,\|\mathbf{u_t}\|_1\, {\mathbf{y_t}}^T C_t \mathbf{y_t} - b{\mathbf{u_t}}^T C_t \mathbf{u_t} \leq \sum_i u_{t,i} \ln\left(\frac{v_{t+1,i}}{y_{t,i}}\right) \tag{61}$$

Note that the above inequality is also true when $\mathbf{u_t} = 0$.

Note that the right side of the above inequality is the same as the right part of the Eq.(32) in the proof of Theorem 1.

As a result, we will use the same steps as in the proof of Theorem 1. Then we will set $\mathbf{u_t} = \mathbf{u} = \operatorname{argmin}_{\mathbf{q}\in\mathcal{B}_1^n} \sum_{t=r}^{s} \mathbf{q}^T C_t \mathbf{q}$ for $t = r, \dots, s$, and 0 elsewhere. Summing from $t = 1$ up to T, gives the following inequality:

$$a\Big[\sum_{t=r}^{s} {\mathbf{y_t}}^T C_t \mathbf{y_t}\Big] - b\Big[\min_{\mathbf{u}\in\mathcal{B}_1^n} \sum_{t=r}^{s} \mathbf{u}^T C_t \mathbf{u}\Big] \leq \ln\frac{n}{\alpha} + T\ln\frac{1}{1-\alpha} \tag{62}$$

Since $\alpha = 1/(T+1)$, $T\ln\frac{1}{1-\alpha} \leq 1$. Then the above inequality becomes

$$a\Big[\sum_{t=r}^{s} {\mathbf{y_t}}^T C_t \mathbf{y_t}\Big] - b\Big[\min_{\mathbf{u}\in\mathcal{B}_1^n} \sum_{t=r}^{s} \mathbf{u}^T C_t \mathbf{u}\Big] \leq \ln\big((1+T)n\big) + 1 \tag{63}$$

Plugging in the expressions of $a = c/(2c+2)$, $b = c/2$, and $c = \frac{\sqrt{2\ln\big((1+T)n\big)+2}}{\sqrt{L}}$ we will have

$$\begin{aligned}
&\sum_{t=r}^{s} {\mathbf{y_t}}^T C_t \mathbf{y_t} - \min_{\mathbf{u}\in\mathcal{B}_1^n} \sum_{t=r}^{s} \mathbf{u}^T C_t \mathbf{u} \\
&\leq c\Big[\min_{\mathbf{u}\in\mathcal{B}_1^n} \mathbf{u}^T C_t \mathbf{u}\Big] + 2\tfrac{c+1}{c}\big(\ln\big((1+T)n\big) + 1\big) \\
&\leq cL + 2\tfrac{c+1}{c}\big(\ln\big((1+T)n\big) + 1\big) \\
&= 2\sqrt{2L\Big(\ln\big((1+T)n\big) + 1\Big)} + 2\ln\big((1+T)n\big)
\end{aligned} \tag{64}$$

Since the inequality holds for any $1 \leq r \leq s \leq T$, the proof is concluded by maximizing over $[r, s]$ on the left. □