Supplementary Material: Online Stochastic Linear Optimization under One-bit Feedback

Lijun Zhang ZHANGLJ@LAMDA.NJU.EDU.CN
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China

Tianbao Yang TIANBAO-YANG@UIOWA.EDU
Department of Computer Science, The University of Iowa, Iowa City, IA 52242, USA

Rong Jin JINRONG.JR@ALIBABA-INC.COM
Alibaba Group, Seattle, USA

Yichi Xiao XIAOYC@LAMDA.NJU.EDU.CN
Zhi-Hua Zhou ZHOUZH@LAMDA.NJU.EDU.CN
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China

A. More Strategies for Efficient Implementations

Enlarging the Confidence region For a positive definite matrix $A \in \mathbb{R}^{d\times d}$, we define

$$\|\mathbf{x}\|_{1,A} = \|A^{1/2}\mathbf{x}\|_1.$$

When studying SLB, Dani et al. (2008) propose to enlarge the confidence region from $\mathcal{C}_{t+1} = \left\{\mathbf{w} : \|\mathbf{w}-\mathbf{w}_{t+1}\|_{Z_{t+1}} \leq \sqrt{\gamma_{t+1}}\right\}$ to $\widetilde{\mathcal{C}}_{t+1} = \left\{\mathbf{w} : \|\mathbf{w}-\mathbf{w}_{t+1}\|_{1,Z_{t+1}} \leq \sqrt{d\gamma_{t+1}}\right\}$ such that the computational cost could be reduced. This idea can be direct incorporated to our OL^{2}M. Let $\mathcal{E}_{t+1}$ be the set of extremal points of $\widetilde{\mathcal{C}}_{t+1}$. With this modification, (11) becomes

$$(\mathbf{x}_{t+1}, \widehat{\mathbf{w}}_{t+1}) = \underset{\mathbf{x}\in\mathcal{D},\mathbf{w}\in\widetilde{\mathcal{C}}_{t+1}}{\operatorname{argmax}} \ \mathbf{x}^\top \mathbf{w} = \underset{\mathbf{x}\in\mathcal{D},\mathbf{w}\in\mathcal{E}_{t+1}}{\operatorname{argmax}} \ \mathbf{x}^\top \mathbf{w}$$

which means we just need to enumerate over the $2d$ vertices in $\mathcal{E}_{t+1}$. Following the arguments in Dani et al. (2008), it is straightforward to show that the regret is only increased by a factor of $\sqrt{d}$.

Lazy Updating Abbasi-yadkori et al. (2011) propose a lazy updating strategy which only needs to solve (11) $O(\log T)$ times. The key idea is to recompute $\mathbf{x}_t$ whenever $\det(Z_t)$ increases by a constant factor $(1 + c)$. While the computation cost is saved dramatically, the regret is only increased by a constant factor $\sqrt{1+c}$. We provide the lazy updating version of OL^{2}M in Algorithm 2.

B. Proof of Lemma 1

Let $\mu(x) = \frac{\exp(x)}{1+\exp(\mathbf{x})}$. It is easy to verify that $\forall x \in [-R, R]$,

$$\frac{1}{2(1+\exp(R))} \leq \mu'(\mathbf{x}) = \frac{\exp(x)}{(1+\exp(x))^2} \leq \frac{1}{4} \tag{23}$$

Note that for any $-R \leq a \leq b \leq R$, we have

$$\mu(b) = \mu(a) + \int_a^b \mu'(x)dx \tag{24}$$

Algorithm 2 OL^2M with Lazy Updating

Input: Regularization Parameter λ, Constant c
$Z_1 = \lambda I$, $\mathbf{w}_1 = 0$, $\tau = 1$
for $t = 1, 2, \ldots$ **do**
  if $\det(Z_t) > (1+c)\det(Z_\tau)$ **then**

$$(\mathbf{x}_t, \widehat{\mathbf{w}}_t) = \underset{\mathbf{x}\in\mathcal{D}, \mathbf{w}\in\mathcal{C}_t}{\operatorname{argmax}} \ \mathbf{x}^\top \mathbf{w}$$

    $\tau = t$
  end if
  $\mathbf{x}_t = \mathbf{x}_\tau$
  Submit $\mathbf{x}_t$ and observe $y_t \in \{\pm 1\}$
  Solve the optimization problem in (8) to find $\mathbf{w}_{t+1}$
end for

Combining (23) with (24), we have

$$\frac{1}{2(1+\exp(R))}(b-a) \leq \mu(b) - \mu(a) \leq \frac{1}{4}(b-a)$$

Let

$$\mathbf{x}_* = \underset{\mathbf{x}\in\mathcal{D}}{\operatorname{argmax}}\, \mathbf{x}^\top \mathbf{w}_* = \underset{\mathbf{x}\in\mathcal{D}}{\operatorname{argmax}}\, \frac{\exp(\mathbf{x}^\top \mathbf{w}_*)}{1+\exp(\mathbf{x}^\top \mathbf{w}_*)}$$

Since $-R \leq \mathbf{x}_t^\top \mathbf{w}_* \leq \mathbf{x}_*^\top \mathbf{w}_* \leq R$, we have

$$\frac{1}{2(1+\exp(R))}\left(\mathbf{x}_*^\top \mathbf{w}_* - \mathbf{x}_t^\top \mathbf{w}_*\right) \leq \frac{\exp(\mathbf{x}_*^\top \mathbf{w}_*)}{1+\exp(\mathbf{x}_*^\top \mathbf{w}_*)} - \frac{\exp(\mathbf{x}_t^\top \mathbf{w}_*)}{1+\exp(\mathbf{x}_t^\top \mathbf{w}_*)} \leq \frac{1}{4}\left(\mathbf{x}_*^\top \mathbf{w}_* - \mathbf{x}_t^\top \mathbf{w}_*\right)$$

which implies (7).

C. Proof of Lemma 2

We first show that the one-dimensional logistic loss $\ell(x) = \log(1+\exp(-x))$ is $\frac{1}{2(1+\exp(R))}$-strongly convex over domain $[-R, R]$. It is easy to verify that $\forall x \in [-R, R]$,

$$\ell''(x) = \frac{\exp(x)}{(1+\exp(x))^2} \geq \frac{1}{2(1+\exp(R))}$$

implying the strongly convexity of $\ell(\cdot)$. From the property of strongly convex, for any $a, b \in [-R, R]$ we have

$$\ell(b) \geq \ell(a) + \ell'(a)(b-a) + \frac{\beta}{2}(b-a)^2. \tag{25}$$

Notice that for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{B}_R$, we have

$$y_t \mathbf{x}_t^\top \mathbf{w}_1,\ y_t \mathbf{x}_t^\top \mathbf{w}_2 \in [-R, R],$$

since $y_t \in \{\pm 1\}$ and $\|\mathbf{x}_t\|_2 \leq 1$. Substituting $a = y_t \mathbf{x}_t^\top \mathbf{w}_1$ and $b = y_t \mathbf{x}_t^\top \mathbf{w}_2$ into (25), we have

$$\ell(y_t \mathbf{x}_t^\top \mathbf{w}_2) \geq \ell(y_t \mathbf{x}_t^\top \mathbf{w}_1) + \frac{\beta}{2}(y_t \mathbf{x}_t^\top \mathbf{w}_2 - y_t \mathbf{x}_t^\top \mathbf{w}_1)^2 + \ell'(y_t \mathbf{x}_t^\top \mathbf{w}_1)(y_t \mathbf{x}_t^\top \mathbf{w}_2 - y_t \mathbf{x}_t^\top \mathbf{w}_1).$$

We complete the proof by noticing

$$f_t(\mathbf{w}_1) = \ell(y_t \mathbf{x}_t^\top \mathbf{w}_1),\ f_t(\mathbf{w}_2) = \ell(y_t \mathbf{x}_t^\top \mathbf{w}_2), \text{ and } \nabla f_t(\mathbf{w}_1) = \ell'(y_t \mathbf{x}_t^\top \mathbf{w}_1) y_t \mathbf{x}_t.$$

D. Proof of Lemma 3

Lemma 3 follows from a more general result stated below.

Lemma 7. *Let M be a positive definite matrix, and*

$$\mathbf{y} = \arg\min_{\mathbf{w}\in\mathcal{W}} \langle \mathbf{w}, \mathbf{g}\rangle + \frac{1}{2}\|\mathbf{w}-\mathbf{x}\|_M^2,$$

where $\mathcal{W}$ is a convex set. Then for all $\mathbf{w}\in\mathcal{W}$, we have

$$\langle \mathbf{x}-\mathbf{w}, \mathbf{g}\rangle \leq \frac{\|\mathbf{x}-\mathbf{w}\|_M^2 - \|\mathbf{y}-\mathbf{w}\|_M^2}{2} + \frac{1}{2}\|\mathbf{g}\|_{M^{-1}}^2.$$

Proof. Since $\mathbf{y}$ is the optimal solution to the optimization problem, from the first-order optimality condition (Boyd & Vandenberghe, 2004), we have

$$\langle \mathbf{g} + M(\mathbf{y}-\mathbf{x}), \mathbf{w}-\mathbf{y}\rangle \geq 0, \ \forall \mathbf{w}\in\mathcal{W}. \tag{26}$$

Based on the above inequality, we have

$$\begin{aligned}
&\|\mathbf{x}-\mathbf{w}\|_M^2 - \|\mathbf{y}-\mathbf{w}\|_M^2 \\
=&\mathbf{x}^\top M\mathbf{x} - \mathbf{y}^\top M\mathbf{y} + 2\langle M(\mathbf{y}-\mathbf{x}), \mathbf{w}\rangle \\
\overset{(26)}{\geq}&\mathbf{x}^\top M\mathbf{x} - \mathbf{y}^\top M\mathbf{y} + 2\langle M(\mathbf{y}-\mathbf{x}), \mathbf{y}\rangle - 2\langle \mathbf{g}, \mathbf{w}-\mathbf{y}\rangle \\
=&\|\mathbf{y}-\mathbf{x}\|_M^2 + 2\langle \mathbf{g}, \mathbf{y}-\mathbf{x}+\mathbf{x}-\mathbf{w}\rangle \\
=&2\langle \mathbf{g}, \mathbf{x}-\mathbf{w}\rangle + \|\mathbf{y}-\mathbf{x}\|_M^2 + 2\langle \mathbf{g}, \mathbf{y}-\mathbf{x}\rangle
\end{aligned}$$

Combining with the following inequality

$$\|\mathbf{y}-\mathbf{x}\|_M^2 + 2\langle \mathbf{g}, \mathbf{y}-\mathbf{x}\rangle \geq \min_{\mathbf{w}} \|\mathbf{w}\|_M^2 + 2\langle \mathbf{g}, \mathbf{w}\rangle = -\|\mathbf{g}\|_{M^{-1}}^2,$$

we have

$$\|\mathbf{x}-\mathbf{w}\|_M^2 - \|\mathbf{y}-\mathbf{w}\|_M^2 \geq 2\langle \mathbf{g}, \mathbf{x}-\mathbf{w}\rangle - \|\mathbf{g}\|_{M^{-1}}^2.$$

□

E. Proof of Lemma 4

For each $\mathbf{w}\in\mathbb{R}^d$, we introduce a discrete probability distribution $p_{\mathbf{w}}$ over $\{\pm 1\}$ such that

$$p_{\mathbf{w}}(i) = \frac{1}{1+\exp(-i\mathbf{x}_t^\top \mathbf{w})}, \ i\in\{\pm 1\}.$$

Then, it is easy to verify that

$$\bar{f}_t(\mathbf{w}) = -\sum_{i\in\{\pm 1\}} p_{\mathbf{w}_*}(i)\log p_{\mathbf{w}}(i).$$

As a result

$$\begin{aligned}
&\bar{f}_t(\mathbf{w}) - \bar{f}_t(\mathbf{w}_*) \\
=&\sum_{i\in\{\pm 1\}} p_{\mathbf{w}_*}(i)\log p_{\mathbf{w}_*}(i) - \sum_{i\in\{\pm 1\}} p_{\mathbf{w}_*}(i)\log p_{\mathbf{w}}(i) \\
=&\sum_{i\in\{\pm 1\}} p_{\mathbf{w}_*}(i)\log \frac{p_{\mathbf{w}_*}(i)}{p_{\mathbf{w}}(i)} = D_{KL}(p_{\mathbf{w}_*}\|p_{\mathbf{w}}) \geq 0
\end{aligned}$$

where $D_{KL}(\cdot\|\cdot)$ is the Kullback–Leibler divergence between two distributions (Cover & Thomas, 2006).

F. Proof of Lemma 5

We need the Bernstein's inequality for martingales (Cesa-Bianchi & Lugosi, 2006), which is provided in Appendix J. Form our definition of $\bar{f}_i(\cdot)$ in (16), it is clear

$$b_i = [\nabla \bar{f}_i(\mathbf{w}_i) - \nabla f_i(\mathbf{w}_i)]^\top (\mathbf{w}_i - \mathbf{w}_*)$$

is a martingale difference sequence. Furthermore,

$$|b_i| \leq \left|[\nabla \bar{f}_i(\mathbf{w}_i)]^\top (\mathbf{w}_i - \mathbf{w}_*)\right| + \left|[\nabla f_i(\mathbf{w}_i)]^\top (\mathbf{w}_i - \mathbf{w}_*)\right| \leq 2|\mathbf{x}_i^\top (\mathbf{w}_i - \mathbf{w}_*)| \leq 2\|\mathbf{w}_i - \mathbf{w}_*\|_2 \leq 4R.$$

Define the martingale $B_t = \sum_{i=1}^t b_i$. Define the conditional variance Σ_t^2 as

$$\begin{aligned}\Sigma_t^2 =& \sum_{i=1}^t \mathrm{E}_{y_i}\left[\left([\nabla \bar{f}_i(\mathbf{w}_i) - \nabla f_i(\mathbf{w}_i)]^\top (\mathbf{w}_i - \mathbf{w}_*)\right)^2\right] \\ \leq & \sum_{i=1}^t \mathrm{E}_{y_i}\left[\left(\nabla f_i(\mathbf{w}_i)^\top (\mathbf{w}_i - \mathbf{w}_*)\right)^2\right] \leq \underbrace{\sum_{i=1}^t \left(\mathbf{x}_i^\top (\mathbf{w}_i - \mathbf{w}_*)\right)^2}_{:=A_t},\end{aligned}$$

where the first inequality is due to the fact that $\mathrm{E}[(\xi - \mathrm{E}[\xi])^2] \leq \mathrm{E}[\xi^2]$ for any random variable ξ.

In the following, we consider two different scenarios, i.e., $A_t \leq \frac{4R^2}{t}$ and $A_t > \frac{4R^2}{t}$.

$A_t \leq \frac{4R^2}{t}$ In this case, we have

$$B_t \leq \sum_{i=1}^t |b_i| \leq 2\sum_{i=1}^t |\mathbf{x}_i^\top (\mathbf{w}_i - \mathbf{w}_*)| \leq 2\sqrt{t\sum_{i=1}^t \left(\mathbf{x}_i^\top (\mathbf{w}_i - \mathbf{w}_*)\right)^2} \leq 4R. \tag{27}$$

$A_t > \frac{4R^2}{t}$ Since A_t in the upper bound for Σ_t^2 is a random variable, we cannot apply Bernstein's inequality directly. To address this issue, we make use of the peeling process (Bartlett et al., 2005). Note that we have both a lower bound and an upper bound for A_t, i.e., $4R^2/t < A_t \leq 4R^2 t$. Then,

$$\begin{aligned}& \Pr\left[B_t \geq 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t\right] \\ =& \Pr\left[B_t \geq 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t, \frac{4R^2}{t} < A_t \leq 4R^2 t\right] \\ =& \Pr\left[B_t \geq 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t, \Sigma_t^2 \leq A_t, \frac{4R^2}{t} < A_t \leq 4R^2 t\right] \\ \leq & \sum_{i=1}^m \Pr\left[B_t \geq 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t, \Sigma_t^2 \leq A_t, \frac{4R^2 2^{i-1}}{t} < A_t \leq \frac{4R^2 2^i}{t}\right] \\ \leq & \sum_{i=1}^m \Pr\left[B_t \geq \sqrt{2\frac{4R^2 2^i}{t}\tau_t} + \frac{8}{3}R\tau_t, \Sigma_t^2 \leq \frac{4R^2 2^i}{t}\right] \leq \; m e^{-\tau_t},\end{aligned}$$

where $m = \lceil 2\log_2 t \rceil$, and the last step follows the Bernstein's inequality for martingales. By setting $\tau_t = \log \frac{2mt^2}{\delta}$, with a probability at least $1 - \delta/[2t^2]$, we have

$$B_t \leq 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t. \tag{28}$$

Combining (27) and (28), with a probability at least $1 - \delta/[2t^2]$, we have

$$B_t \leq 4R + 2\sqrt{A_t \tau_t} + \frac{8}{3}R\tau_t.$$

We complete the proof by taking the union bound over $t > 0$, and using the well-known result

$$\sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6} \leq 2.$$

G. Proof of Lemma 6

We have

$$\|\mathbf{x}_i\|^2_{Z^{-1}_{i+1}} = \frac{2}{\beta}\langle Z^{-1}_{i+1}, Z_{i+1} - Z_i\rangle \leq \frac{2}{\beta} \log \frac{\det(Z_{i+1})}{\det(Z_i)},$$

where the inequality follows from Lemma 12 in Hazan et al. (2007). Thus, we have

$$\sum_{i=1}^{t} \|\mathbf{x}_i\|^2_{Z^{-1}_{i+1}} \leq \frac{2}{\beta} \sum_{i=1}^{t} \log \frac{\det(Z_{i+1})}{\det(Z_i)} = \frac{2}{\beta} \log \frac{\det(Z_{t+1})}{\det(Z_1)}.$$

H. Proof of Corollary 2

Recall that

$$Z_{t+1} = Z_1 + \frac{\beta}{2} \sum_{i=1}^{t} \mathbf{x}_t \mathbf{x}_t^\top$$

and $\|\mathbf{x}_t\|_2 \leq 1$ for all $t > 0$. From Lemma 10 of Abbasi-yadkori et al. (2011), we have

$$\det(Z_{t+1}) \leq \left(\lambda + \frac{\beta t}{2d}\right)^d.$$

Since $\det(Z_1) = \lambda^d$, we have

$$\log \frac{\det(Z_{t+1})}{\det(Z_1)} \leq d \log\left(1 + \frac{\beta t}{2\lambda d}\right).$$

I. Proof of Theorem 3

The proof is standard and can be found from Dani et al. (2008) and Abbasi-yadkori et al. (2011). We include it for the sake of completeness.

Let $\mathbf{x}_* = \operatorname{argmax}_{\mathbf{x}\in\mathcal{D}} \mathbf{x}^\top \mathbf{w}_*$. Recall that in each round, we have

$$(\mathbf{x}_t, \widehat{\mathbf{w}}_t) = \underset{\mathbf{x}\in\mathcal{D}, \mathbf{w}\in\mathcal{C}_t}{\operatorname{argmax}} \ \mathbf{x}^\top \mathbf{w}.$$

We decompose the instantaneous regret at round t as follows

$$\begin{aligned}
&\mathbf{x}_*^\top \mathbf{w}_* - \mathbf{x}_t^\top \mathbf{w}_* \\
\leq &\mathbf{x}_t^\top \widehat{\mathbf{w}}_t - \mathbf{x}_t^\top \mathbf{w}_* = \mathbf{x}_t^\top (\widehat{\mathbf{w}}_t - \mathbf{w}_t) + \mathbf{x}_t^\top (\mathbf{w}_t - \mathbf{w}_*) \\
\leq &\left(\|\widehat{\mathbf{w}}_t - \mathbf{w}_t\|_{Z_t} + \|\mathbf{w}_t - \mathbf{w}_*\|_{Z_t}\right) \|\mathbf{x}_t\|_{Z_t^{-1}} \leq 2\sqrt{\gamma_t}\|\mathbf{x}_t\|_{Z_t^{-1}}.
\end{aligned}$$

On the other hand, we always have

$$\mathbf{x}_*^\top \mathbf{w}_* - \mathbf{x}_t^\top \mathbf{w}_* \leq \|\mathbf{x}_* - \mathbf{x}_t\|_2 \|\mathbf{w}_*\|_2 \leq 2R.$$

Thus, the total regret can be upper bounded by

$$\begin{aligned}
&T \max_{\mathbf{x}\in\mathcal{D}} \mathbf{x}^\top \mathbf{w}_* - \sum_{t=1}^T \mathbf{x}_t^\top \mathbf{w}_* \\
\leq& 2\sum_{t=1}^T \min\left(\sqrt{\gamma_t}\|\mathbf{x}_t\|_{Z_t^{-1}}, R\right) \\
\leq& 2\sqrt{\gamma_T}\sum_{t=1}^T \min\left(\|\mathbf{x}_t\|_{Z_t^{-1}}, R\right) \\
=& 2\sqrt{\frac{2}{\beta}\gamma_T}\sum_{t=1}^T \min\left(\sqrt{\frac{\beta}{2}}\|\mathbf{x}_t\|_{Z_t^{-1}}, \sqrt{\frac{\beta}{2}}R\right) \\
\leq& 2\max\left(1, \sqrt{\frac{\beta}{2}}R\right)\sqrt{\frac{2}{\beta}\gamma_T}\sum_{t=1}^T \min\left(\sqrt{\frac{\beta}{2}}\|\mathbf{x}_t\|_{Z_t^{-1}}, 1\right) \\
\leq& 2\max\left(1, \sqrt{\frac{\beta}{2}}R\right)\sqrt{\frac{2T}{\beta}\gamma_T}\sqrt{\sum_{t=1}^T \min\left(\frac{\beta}{2}\|\mathbf{x}_t\|^2_{Z_t^{-1}}, 1\right)}.
\end{aligned}$$

To proceed, we need the following results from Lemma 11 in Abbasi-yadkori et al. (2011),

$$\sum_{t=1}^T \min\left(\frac{\beta}{2}\|\mathbf{x}_t\|^2_{Z_t^{-1}}, 1\right) \leq 2\sum_{t=1}^T \log\left(1 + \frac{\beta}{2}\|\mathbf{x}_t\|^2_{Z_t^{-1}}\right)$$

and

$$\begin{aligned}
&\det(Z_{T+1}) = \det\left(Z_T + \frac{\beta}{2}\mathbf{x}_T\mathbf{x}_T^\top\right) \\
=&\det(Z_T)\det\left(I + \frac{\beta}{2}Z_T^{-1/2}\mathbf{x}_T\mathbf{x}_T^\top Z_T^{-1/2}\right) \\
=&\det(Z_T)\left(1 + \frac{\beta}{2}\|\mathbf{x}_T\|^2_{Z_T^{-1}}\right) = \det(Z_1)\prod_{t=1}^T\left(1 + \frac{\beta}{2}\|\mathbf{x}_t\|^2_{Z_t^{-1}}\right).
\end{aligned}$$

Combining the above inequations, we have

$$T \max_{\mathbf{x}\in\mathcal{D}} \mathbf{x}^\top \mathbf{w}_* - \sum_{t=1}^T \mathbf{x}_t^\top \mathbf{w}_* \leq 4\max\left(1, \sqrt{\frac{\beta}{2}}R\right)\sqrt{\frac{\gamma_T T}{\beta}\log\frac{\det(Z_{T+1})}{\det(Z_1)}}.$$

J. Bernstein's Inequality for Martingales

Theorem 4. *Let $X_1, \ldots, X_n$ be a bounded martingale difference sequence with respect to the filtration $\mathcal{F} = (\mathcal{F}_i)_{1\leq i\leq n}$ and with $|X_i| \leq K$. Let*

$$S_i = \sum_{j=1}^i X_j$$

be the associated martingale. Denote the sum of the conditional variances by

$$\Sigma_n^2 = \sum_{t=1}^n \mathrm{E}\left[X_t^2|\mathcal{F}_{t-1}\right].$$

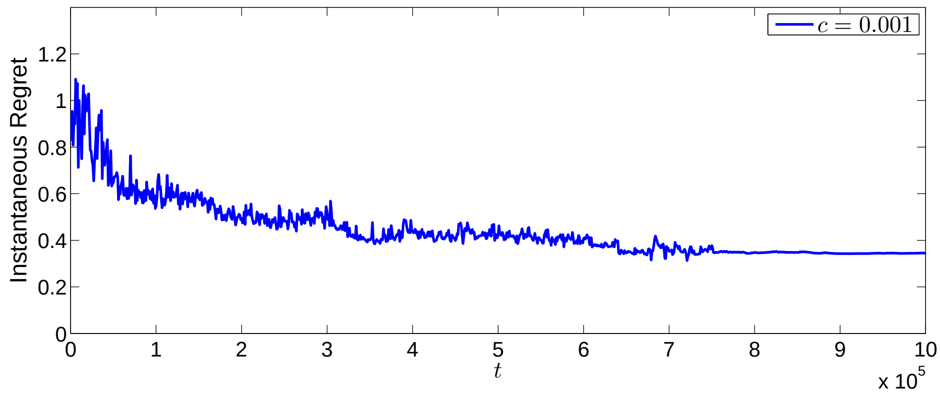
Then for all constants t, $\nu > 0$,

$$\Pr\left[\max_{i=1,\ldots,n} S_i > t \text{ and } \Sigma_n^2 \leq \nu\right] \leq \exp\left(-\frac{t^2}{2(\nu + Kt/3)}\right),$$

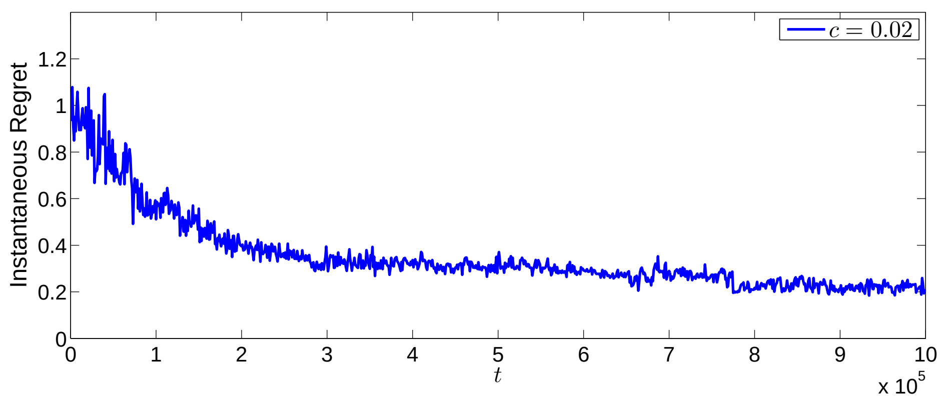
and therefore,

$$\Pr\left[\max_{i=1,\ldots,n} S_i > \sqrt{2\nu t} + \frac{2}{3}Kt \text{ and } \Sigma_n^2 \leq \nu\right] \leq e^{-t}.$$

K. Instantaneous regret of OL^{2}M when $\mathcal{D}$ is the unit ball in $\mathbb{R}^{100}$

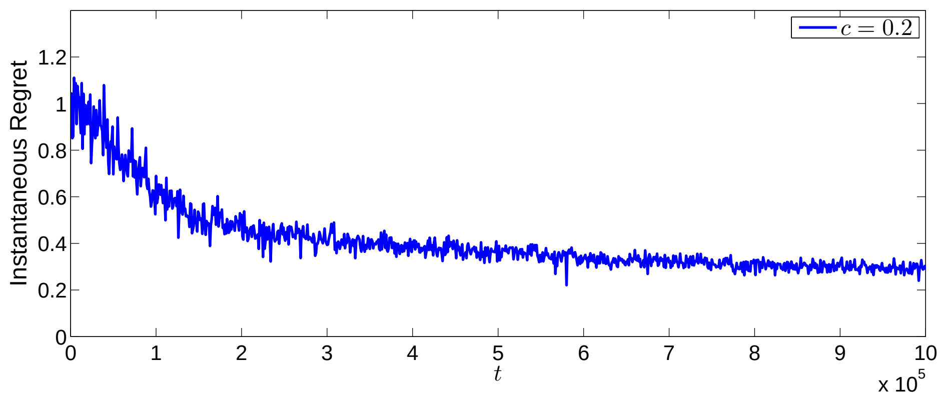


(a) $c = 0.001$

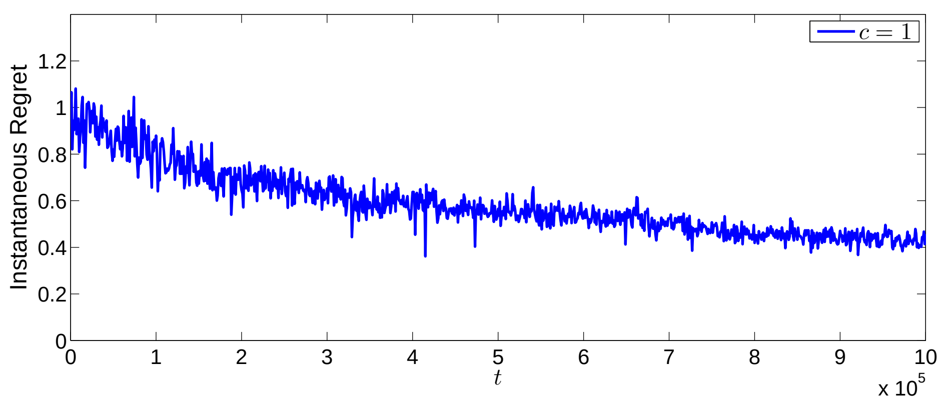


(b) $c = 0.02$

Figure 4. Instantaneous regret of OL^{2}M when $\mathcal{D}$ is the unit ball in $\mathbb{R}^{100}$.



(a) $c = 0.2$



(b) $c = 1$

Figure 5. Instantaneous regret of $\mathrm{OL}^2\mathrm{M}$ when $\mathcal{D}$ is the unit ball in $\mathbb{R}^{100}$.