Solution for Exercise 1 :

Pick η_t as $\eta_t = \frac{\gamma}{M-m}\sqrt{\frac{\ln N}{t}}$ where γ is to be determined by the analysis

Hoeffding's lemma : $\delta_t \leqslant \frac{\eta_t}{8}(M-m)^2$

So that :
$$R_T = \sum_{t,j} p_{jt}\,\ell_{jt} - \min_k \sum_{t=1}^{T} \ell_{kt}$$
$$\leqslant \frac{\ln N}{\eta_T} + \sum_{t=1}^{T} \delta_t \leqslant \frac{M-m}{\gamma}\sqrt{T\ln N} + \gamma\frac{(M-m)}{8}\sum_{t=1}^{T}\sqrt{\frac{\ln N}{t}}$$

Now,
$$\sum_{t=1}^{T} 1/\sqrt{t} \leqslant \int_0^T \frac{1}{\sqrt{t}}\,dt = \left[2\sqrt{t}\right]_0^T = 2\sqrt{T}$$

$$R_T \leqslant (M-m)\sqrt{T\ln N}\left(\underbrace{\frac{1}{\gamma} + \frac{2\gamma}{8}}_{\text{optimal value } \gamma^* = 2}\right) = (M-m)\sqrt{T\ln N}$$

better constant than with the doubling trick

We only lose a $\sqrt{2}$ factor wrt bound when T is known
($\sqrt{2}$ is the price for adaptivity in T).

Solution for Exercise 2 :

If you find one (even in several weeks !) please send it to me ! The reward for a significant contribution will be given by bonus points at the exam.

Solution for Exercise 3.

(1) $v_t = \sum_j p_{jt} \Big(\ell_{jt} - \sum_k p_{kt} \ell_{kt} \Big)^2$ is some variance; thus, by the variational formula for the variance

$$= \min_{\mu \in \mathbb{R}} \sum_j p_{jt} (\ell_{jt} - \mu)^2$$

pick $\mu = 0$

$$\leqslant \sum_j p_{jt} \ell_{jt}^2$$

$\ell_{jt} \geqslant 0$ and $\ell_{jt} \leqslant M$, thus $\ell_{jt}^2 \leqslant M \ell_{jt}$

$$\leqslant M \sum_j p_{jt} \ell_{jt}$$

Hence, summing over $t = 1, \ldots T$:

$$\sum_{t=1}^T v_t = \sum_{t=1}^T \sum_{j=1}^N p_{jt} \Big(\ell_{jt} - \sum_{k=1}^N p_{kt} \ell_{kt} \Big)^2 \leqslant M \sum_{t=1}^T \sum_j p_{jt} \ell_{jt}$$

(2) The theorem for EWA tuned with $\eta_t = \dfrac{\ln N}{\sum_{s \leqslant t-1} \delta_s}$ ensures that for this algorithm (since $m = 0$):

$$\sum_{t,j} p_{jt} \ell_{jt} - \min_k \sum_t \ell_{kt} \leqslant 2 \underbrace{\sqrt{\sum_t v_t \ln N}}_{\leqslant 2\sqrt{M \sum_{t,j} p_{jt} \ell_{jt} \ln N}} + M \Big(2 + \frac{4}{3} \ln N \Big) \qquad (*)$$

This, denoting $x = \sqrt{\sum_{t,j} p_{jt} \ell_{jt}}$, we have the 2nd order inequality:

$$x^2 \leqslant \Big(2\sqrt{M \ln N} \Big) x + \Big(\min_k \sum_{t=1}^T \ell_{kt} + M \Big(2 + \frac{4}{3} \ln N \Big) \Big)$$

A lemma that we saw in class says that if $x^2 \leqslant b + a\sqrt{x}$ $(a, b \geqslant 0)$

then $x \leqslant a + \sqrt{b}$

This means here that

$$\sqrt{\sum_{t,j} p_{jt} \ell_{jt}} \leqslant 2\sqrt{M \ln N} + \underbrace{\sqrt{\min_k \sum_t \ell_{kt} + M \Big(2 + \frac{4}{3} \ln N \Big)}}_{\leqslant \sqrt{\min_k \sum_t \ell_{kt}} + \sqrt{M \left(2 + \frac{4}{3} \ln N \right)}}$$

We could take $(\ )^2$ of both sides

but it is slighty more elegant to substitute this inequality into $(*)$

We get:

$$\underbrace{\sum_{t,j} p_{jt}\, \ell_{jt} - \min_k \sum_t \ell_{kt} \leqslant 4M\ln N + 2M\sqrt{\left(2+\tfrac{4}{3}\ln N\right)\ln N} + M\left(2+\tfrac{4}{3}\ln N\right)}_{\text{let's make these more readable!}} + 2\sqrt{M \min_k \sum_t \ell_{kt} \ln N}$$

(3)

We may assume that $N \geqslant 2$, in which case $\ln N \geqslant 0{,}693$ and $1 \leqslant \frac{3}{2}\ln N$

Then,

$$4M\ln N + 2M\sqrt{\Big(\underbrace{2}_{\leqslant 2\times\frac{3}{2}\ln N}+\tfrac{4}{3}\ln N\Big)\ln N} + M\Big(\underbrace{2}_{\leqslant 2\times\frac{3}{2}\ln N}+\tfrac{4}{3}\ln N\Big) \leqslant M\ln N \times \underbrace{\left(4+2\sqrt{3+\tfrac{4}{3}} + 3+\tfrac{4}{3}\right)}_{\leqslant 13}$$

Final bound:

$$\sum_{t,j} p_{jt}\,\ell_{jt} - \min_k \sum_{t=1}^{T} \ell_{kt} \leqslant \underbrace{13M\ln N}_{\substack{\text{error suffered when}\\ \exists k \mid \sum_t \ell_{kt} = 0\\ \text{i.e., all } \ell_{kt}=0\\ \text{given that losses are non-negative}}} + 2\sqrt{M \min_k \sum_t \ell_{kt} \ln N}$$

Solution for Exercise 4.

Assume that $\eta \leqslant 1/2M$: $-\eta \ell_{ks} \geqslant -1/2$, positive weights p_{jt}, algorithm well defined

(1)
$$\underbrace{-\eta \sum_{j=1}^{N} p_{jt}\ell_{jt}}_{\geqslant -1/2} \underset{\ln(1+u)\leqslant u}{\geqslant} \ln\Big(1-\eta\sum_j p_{jt}\ell_{jt}\Big)$$
$$= \ln\Big(\sum_j p_{jt}(1-\eta\ell_{jt})\Big)$$
$$\underset{\text{def. of } p_{jt}}{=} \ln \frac{\sum_j \prod_{s=1}^{t}(1-\eta\ell_{js})}{\sum_{k\leqslant N}\prod_{s=1}^{t-1}(1-\eta\ell_{ks})}$$

Telescoping sum:
$$-\eta\sum_{t=1}^{T}\sum_{j=1}^{N} p_{jt}\ell_{jt} \underset{\substack{\text{convention:}\\ \text{empty } \prod \text{ equals } 1}}{\geqslant} \ln\frac{\sum_{k\leqslant N}\prod_{t=1}^{T}(1-\eta\ell_{kt})}{N}$$

We lower bound $\ln\Big(\sum_{k\leqslant N}\prod_{t=1}^{T}(1-\eta\ell_{kt})\Big)$:

$\forall j$,
$$\geqslant \ln\Big(\prod_{t=1}^{T}(1-\eta\ell_{jt})\Big) = \sum_{t=1}^{T}\ln(1-\eta\ell_{jt})$$
$$\geqslant -\eta\ell_{jt} - \eta^2\ell_{jt}^2$$

$-\eta\ell_{jt}\geqslant -1/2$ and $\ln(1+u)\geqslant u-u^2 \ \forall u\geqslant -1/2$

Summarizing what we got so far:

$$\forall j,\quad -\eta\sum_{t=1}^{T}\sum_k p_{kt}\ell_{kt} \geqslant -\ln N - \eta\sum_{t=1}^{T}\ell_{jt} - \eta^2\sum_{t=1}^{T}\ell_{jt}^2$$

Thus, $\forall j$,
$$\sum_{t=1}^{T}\sum_{k=1}^{N} p_{kt}\ell_{kt} - \sum_{t=1}^{T}\ell_{jt} \leqslant \frac{\ln N}{\eta} + \eta\sum_{t=1}^{T}\ell_{jt}^2.$$

We did so under the assumption that $\forall j,t$, $-\eta\ell_{jt}\geqslant -1/2$ valid as soon as $\ell_{jt}\leqslant M$ and $\eta\leqslant 1/2M$.

(2) Waiting for your solutions! (But keep in mind that it's a difficult problem...)