Suppose we start with the careful analysis of Hedge. for $\ell^t \in [-1,1]$, $\forall \varepsilon \le 1$

$$\sum_t \langle p^t, \ell^t \rangle \le \sum_t \langle e_i, \ell^t \rangle + \frac{\log N}{\varepsilon} + \varepsilon \sum_t \langle p^t, (\ell^t)^2 \rangle \quad \forall i$$

component-wise squared

$$(*) \qquad \le \sum_t \langle e_i, \ell^t \rangle + \frac{\log N}{\varepsilon} + \varepsilon \sum_t \langle p^t, |\ell^t| \rangle \quad \forall i$$

component-wise absolute value

~~But for any~~

Now suppose the real loss vectors are $L^1, L^2 \ldots L^T$, each in $[-\gamma, \rho]^N$ $\boxed{1 \le \gamma \le \rho \text{ (say)}}$

define $\ell^t = L^t/\rho \in [-1,1]^N$

define $\varepsilon' = \varepsilon/\gamma \le 1$

And use (*) with these loss vectors, this ε' to get

$$\sum_t \langle p^t, \ell^t \rangle \le \sum_t \langle e_i, \ell^t \rangle + \frac{\log N}{\varepsilon'} + \varepsilon' \sum_t \langle p^t, |\ell^t| \rangle \quad \forall i$$

- Multiply by ρ, and substitute $\frac{\varepsilon}{\gamma} = \varepsilon'$

$$\sum_t \langle p^t, L^t \rangle \le \sum_t \langle e_i, L^t \rangle + \frac{\gamma \rho \log N}{\varepsilon} + \frac{\varepsilon}{\gamma} \sum_t \langle p^t, |L^t| \rangle \quad \forall i$$

- Next: for any vector $\bar{u} \in [-\gamma, \rho]^N$,

  $-\langle p^t, \bar{u} \rangle + \langle p^t, |\bar{u}| \rangle = \langle p^t, (-\bar{u} + |\bar{u}|) \rangle$
  $\le 2\gamma$ (since the negative entries are $\ge -\gamma$)

$$\Rightarrow \sum_t \langle p^t, L^t \rangle \le \sum_t \langle e_i, L^t \rangle + \frac{\gamma \rho \log N}{\varepsilon} + \frac{\varepsilon}{\gamma}\left[\left(\sum_t \langle p^t, L^t \rangle\right) + 2\gamma T\right]$$

$$\Rightarrow (1 - \varepsilon/\gamma) \sum_t \langle p^t, L^t \rangle \le \sum_t \langle e_i, L^t \rangle + \frac{\gamma \rho \log N}{\varepsilon} + 2\varepsilon T$$

$$\Rightarrow \text{Our loss} \le \frac{1}{(1 - \varepsilon/\gamma)}\left[\text{opt loss} + \frac{\gamma \rho \log N}{\varepsilon} + 2\varepsilon T\right]$$

$$\Rightarrow \text{Avg loss} \le \text{Opt loss}\,(1 - \underset{(\text{as } \gamma \ge 1)}{O(\varepsilon)}) + \frac{O(\gamma \rho \log N)}{\varepsilon T} + O(\varepsilon)$$

setting $T \ge \frac{\gamma \rho \log N}{\varepsilon^2}$ gives Avg loss $\le (1-\varepsilon)$ OPT loss $+ O(\varepsilon)$