Proceedings of Machine Learning Research 60:1–15, 2017 Conformal and Probabilistic Prediction and Applications

# Online Aggregation of Unbounded Signed Losses Using Shifting Experts

**Vladimir V. V'yugin** VYUGIN@IITP.RU
*Institute for Information Transmission Problems*
*Skolkovo Institute of Science and Technology*
*Moscow, Russia*

**Editor:** Alex Gammerman, Vladimir Vovk, Zhiyuan Luo, and Harris Papadopoulos

## Abstract

For the decision theoretic online (DTOL) setting, we consider methods to construct algorithms that suffer loss not much more than of any sequence of experts distributed along a time interval (shifting experts setting). We present a modified version of the method of Mixing Past Posteriors which uses as basic algorithm AdaHedge with adaptive learning rate. Due to this, we combine the advantages of both algorithms: regret bounds are valid in the case of signed unbounded losses of the experts, also, we use the shifting regret which is a more optimal characteristic of the algorithm. All results are obtained in the adversarial setting – no assumptions are made about the nature of data source.

We present results of numerical experiments for the case where losses of the experts cannot be bounded in advance.

**Keywords:** On-line learning, unbounded losses, signed losses, adaptive learning rate, algorithm Hedge, mixing past posteriors, shifting experts, share updates

## 1. Introduction

We consider the decision-theoretic on-line learning model in which a master (aggregating) algorithm has to combine losses (or predictions) from a set of experts (see e.g. Littlestone and Warmuth 1994, Freund and Schapire 1997, Vovk 1990, Vovk 1998, Cesa-Bianchi and Lugosi 2006 and so on).

Learning proceeds in trials $t = 1, 2, \dots, T$. In each time moment $t$ the aggregating algorithm presents a prediction in form of a vector of weights $w_t = (w_{1,t}, \dots, w_{N,t})$ assigned to all $N$ experts, where $\sum\limits_{i=1}^{N} w_{i,t} = 1$ and $w_{i,t} \geq 0$ for all $i$. The weight $w_{i,t}$ of an expert $i$ is an estimate of the "quality" of the expert's predictions at step $t$. After that, experts $i = 1, \dots, N$ reveal their losses $l_t = (l_t^1, \dots, l_t^N)$; the aggregating algorithm suffers the loss $h_t = (w_t \cdot l_t) = \sum\limits_{i=1}^{N} w_{i,t} l_t^i$. The cumulative loss suffered by any expert $i$ for the first $T$ steps is defined $L_T^i = \sum\limits_{t=1}^{T} l_t^i$ and the cumulative loss of the aggregating algorithm is $H_T = \sum\limits_{t=1}^{T} h_t$. The important characteristics of the algorithm is a regret $R_T^i = H_T - L_T^i$ with respect to an expert $i$ or minimax regret $R_T = \max_{1 \leq i \leq N} R_T^i$.

The aggregating algorithm updates the experts weights at the end of each trial using losses of the experts and of the algorithm in hindsight. The goal of the algorithm is to design weight updates that guarantee that the loss of the aggregating algorithm is never much larger than the loss of the best expert or the best convex combination of the losses of the experts. So here the best expert or a convex combination of experts serves as a comparator.

A comparison vector is a vector $q = (q_1, \dots, q_N)$ such that $q_1 + \dots + q_N = 1$ and all its components are nonnegative. We compare the cumulative loss of the aggregating logarithm $H_T$ and cumulative convex combination of losses of the experts $\sum\limits_{t=1}^{T}(q \cdot l_t)$.

In classical setting (Freund and Schapire 1997, Vovk 1990 and so on), the process of expert $i$ weights updating is based on the method of exponential weighting with a constant or variable learning rate $\eta$:

$$w_{i,t+1} = \frac{w_{i,t}e^{-\eta l_t^i}}{\sum\limits_{j=1}^{N} w_{j,t}e^{-\eta l_t^j}}.$$

A more challenging goal is to learn well when the comparator $q$ changes over time, i.e the algorithm competes with the cumulative sum $\sum\limits_{t=1}^{T}(q_t \cdot l_t)$, where comparison vector $q_t$ changes over time. An important partial case is when $q_t$ are unit vectors, and so, the sequence of trials is partitioned into segments. In each segment the loss of the algorithm is compared to the loss of a particular expert and this expert changes at the beginning of a new segment. The goal of the the aggregation algorithm is to do almost as well as the sum of losses of experts forming the best partition. Algorithms and bounds of for shifting comparators were presented by Herbster and Warmuth (1998). This method called Fixed Share was generalized by Bousquet and Warmuth (2002) to the method of Mixing Past Posteriors (MPP) in which an arbitrary mixing schemes are used.

Most papers in DTOL setting either consider bounded losses or assume the existence of a specific loss function (see Vovk 1990, Cesa-Bianchi and Lugosi 2006). For example, $l_t^i \in [0, 1]$ for all $i$ and $t$. But in some practical applications, this assumption is too restrictive. Unrestricted gains and losses are typical for financial games where we are faced with unpredictably large values of time series. We allow losses at any step to be unbounded and signed. The notion of a specific loss function is not used.

Algorithm AdaHedge presented by Rooij et al. (2014) is among a few algorithms that do not have similar restrictions. This algorithm is a classical algorithm Hedge of Freund and Schapire (1997) which is completely parameterless and tunes the learning rate $\eta$ in terms of a direct measure of past performance. It develops the Hedge strategy presented by Cesa-Bianchi et al. (2007). In Rooij et al. (2014) an upper bound for regret of this algorithm is presented which is free from boundness assumptions for losses of the experts. Let $l_t^- = \min\limits_i l_t^i,\ l_t^+ = \max\limits_i l_t^i$ be smallest and greatest losses of the experts at a time step $t$. Note that these losses can be of arbitrary sign. Define $L_T^+ = \sum\limits_{t=1}^{T} l_t^+,\ L_T^- = \sum\limits_{t=1}^{T} l_t^-$. Let also $s_t = l_t^+ - l_t^-,\ S_T = \max\{s_1, \dots, s_T\}$.

Let $L^*_T = \min_{1\le i\le N} L^i_T$ be the minimal expert's loss and $R_T = H_T - L^*_T$ be the minimax regret of aggregation algorithm. By Rooij et al. (2014) the regret of the AdaHedge algorithm is bounded

$$R_T \le 2\sqrt{S_T \frac{(L^*_T - L^-_T)(L^+_T - L^*_T)}{L^+_T - L^-_T}\ln N} + \left(\frac{16}{3}\ln N + 2\right) S_T. \tag{1}$$

In case where experts' losses are bounded the upper bound (1) is of the form $O(\sqrt{T\ln N})$.

The goal of this paper is to combine advantages of both algorithms – MPP and AdaHedge. Note that a version of MPP algorithm presented by Bousquet and Warmuth (2002) uses a constant learning rate. At the same time, the algorithm AdaHedge uses adaptive learning rate which is tuned online. In this paper we present a version of algorithm MPP with an adaptive learning parameter. Due to this, we combine algorithms AdaHedge and MPP. We also obtain an upper bound for regret of combined algorithm of order $O(\ln(NT)B_T)$, where $B_T$ denotes the right-hand side of the inequality (1).

All results are obtained in the adversarial setting – no assumptions are made about the nature of data source.

In Section 2 we present our main result – a version of the algorithm MPP with a variable learning rate and an upper bound for regret of this modified version. Results of numerical experiments with unbounded and volatile data are presented in Section 3.

Some details of the proof from Rooij et al. (2014) are presented as a supplementary material in Section 4.

## 2. Algorithm MPP with a variable learning rate

Recall some details of the method of Mixing Past Posteriors (MPP) by Bousquet and Warmuth (2002). Relative entropy is the basic notion of this analysis; it can be considered as a measure of progress.

For any $n$ denote by $\Gamma_n$ the simplex of all probability distributions on a set of cardinality $n$. Let $D(p\|q) = \sum_{i=1}^{n} p_i \ln\frac{p_i}{q_i}$ be the relative entropy, where $p = (p_1,\dots,p_n)$ and $q = (q_1,\dots,q_n)$ are elements of $\Gamma_n$. We define $0\ln 0 = 0$.

For two vectors $p$ and $q$, we write $p > q$, if the $>$ relationship holds componentwise. We use 0 to denote the all-zero vector. The proofs for MPP rely on the following simple inequalities for the relative entropy. For all $p, q, w \in \Gamma_n$ such that $q, w > 0$

$$D(p\|q) \le D(p\|w) + \ln\left(\sum_{i=1}^{n} p_i \frac{w_i}{q_i}\right).$$

If $q \ge \mu w$ for some $\mu > 0$ then $D(p\|q) \le D(p\|w) + \ln\frac{1}{\mu}$. In particular, for $p = w$ we have $D(w\|q) \le \ln\frac{1}{\mu}$ for $q \ge \mu w$. Indeed, from concavity of the logarithm $D(p\|q) - D(p\|w) = \sum_{i=1}^{n} p_i \ln\frac{w_i}{q_i} \le \ln\left(\sum_{i=1}^{n} p_i\frac{w_i}{q_i}\right)$. If $q \ge \mu w$ then $\sum_{i=1}^{n} p_i\frac{w_i}{q_i} \le \sum_{i=1}^{n} p_i\frac{w_i}{\mu w_i} = \frac{1}{\mu}$.

Put $w_{i,1} = w^m_{i,0} = \frac{1}{N}$ for $i = 1, \dots, N$, $\eta_1 = \infty$.
FOR $t = 1, \dots, T$
Predict with weights $w_t = (w_{i,1}, \dots, w_{i,N})$.
Receive losses of the experts $l_t = (l^1_t, \dots, l^N_t)$.
Compute the aggregating algorithm loss $h_t = \sum\limits_{i=1}^{N} w_{i,t} l^i_t$.
Update experts weights and learning parameter in three stages:
**Loss Update**
Define $w^m_{i,t} = \frac{w_{i,t} e^{-\eta_t l^i_t}}{\sum\limits_{j=1}^{N} w_{j,t} e^{-\eta_t l^j_t}}$ for $1 \le i \le N$.
**Mixing Update**
Choose a mixing scheme $\beta^{t+1} = (\beta^{t+1}_0, \dots, \beta^{t+1}_t)$ and define future experts weights $w_{i,t+1} = \sum\limits_{s=0}^{t} \beta^{t+1}_s w^m_{i,s}$ for $1 \le i \le N$.
**Learning Parameter Update**
Define mixloss $m_t = -\frac{1}{\eta_t} \ln \sum\limits_{i=1}^{N} w_{i,t} e^{-\eta_t l^i_t}$, $\delta_t = h_t - m_t$ and mixability gap $\Delta_t = \Delta_{t-1} + \delta_t$.
Define the learning rate $\eta_{t+1} = 1/\Delta_t$.
ENDFOR

Figure 1: Adaptive MPP

The following inequality will be used below. Let $p \in \Gamma_n$ and $q = \sum\limits_{i=0}^{t} \beta_i w_i$, where $w_i \in \Gamma_n$, $w_i > 0$ for $0 \le i \le t$, $\beta = (\beta_0, \dots, \beta_t) \in \Gamma_{t+1}$ and $\beta > 0$. Then

$$D(p\|q) \le D(p\|w_s) + \ln \frac{1}{\beta_s} \tag{2}$$

for any $0 \le s \le t$. In particular, for $p = w_s$, $D\left(w_s \| \sum\limits_{i=0}^{n} \beta_i w_i\right) \le \ln \frac{1}{\beta_s}$ for any $s$.

In what follows the vector $w^m_t = (w^m_{1,t}, \dots, w^m_{N,t})$ presents the normalized experts weights at step $t$. The corresponding posteriori probability distribution $w_{t+1} = (w_{1,t+1}, \dots, w_{N,t+1})$ for step $t+1$ is defined as a convex combination $w_{t+1} = \sum\limits_{s=0}^{t} \beta^{t+1}_s w^m_s$ with weights $\beta^{t+1}_s$, $0 \le s \le t$, where $w^m_s = (w^m_{1,s}, \dots, w^m_{N,s})$.

By the method MPP a mixing scheme is defined by a vector $\beta^{t+1} = (\beta^{t+1}_0, \dots, \beta^{t+1}_t)$, where $\sum\limits_{s=0}^{t} \beta^{t+1}_s = 1$ and $\beta^{t+1}_s \ge 0$ for $0 \le s \le t$.

We use a notion of mixloss $m_t = -\frac{1}{\eta_t} \ln \sum\limits_{i=1}^{N} w_{i,t} e^{-\eta_t l^i_t}$, the cumulative mixloss is defined $M_T = \sum\limits_{t=1}^{T} m_t$. Also, $\delta_t = h_t - m_t$ is the mixability gap and $\Delta_t = \Delta_{t-1} + \delta_t$ is the cumulative mixability gap. From the convexity of the exponent $m_t \le h_t$ for all $t$; also, $\Delta_t \le \Delta_{t+1}$.

Let $H_T = \sum\limits_{t=1}^{T} h_t$ be the cumulative loss of the algorithm for the first $T$ steps. By definition $H_T = M_T + \Delta_T$.

Algorithm MPP with adaptive learning rate is presented on Figure 1.

We will use the following mixing schemes by Bousquet and Warmuth (2002):

**Example 1.** A version of Fixed Share by Herbster and Warmuth (1998) with a variable learning rate is defined by the following mixing scheme. Let a sequence $1 \ge \alpha_1 \ge \alpha_2 \ge \dots$ of parameters be given. Define $\beta_t^{t+1} = 1 - \alpha_{t+1}$ and $\beta_0^{t+1} = \alpha_{t+1}$ ($\beta_s^{t+1} = 0$ for $0 < s < t$). The corresponding prediction for step $t+1$ is defined

$$w_{i,t+1} = \frac{\alpha_{t+1}}{N} + (1 - \alpha_{t+1}) w_{i,t}^m$$

for all $1 \le i \le N$.

**Example 2.** Uniform Past by Bousquet and Warmuth 2002 with a variable learning rate. Put $\beta_t^{t+1} = 1 - \alpha_{t+1}$ and $\beta_s^{t+1} = \frac{\alpha_{t+1}}{t}$ for $0 \le s < t$. The corresponding prediction for step $t+1$ is defined

$$w_{i,t+1} = \alpha_{t+1} \sum_{s=0}^{t-1} \frac{w_{i,s}^m}{t} + (1 - \alpha_{t+1}) w_{i,t}^m$$

for all $i$ and $t$;

Let $l_t = (l_t^1, \dots, l_t^N)$ be a vector of losses of the experts at step $t$. The convex combinations of these losses at step $t$ is $(q_t \cdot l_t) = \sum\limits_{i=1}^{N} q_{i,t} l_t^i$, where $q_t = (q_{1,t}, \dots, q_{N,t})$ is a comparison vector for step $t$, $q_t \in \Gamma_N$. In particular, for unit comparison vector $q_t = (0, \dots, 1, \dots, 0)$ we have $(q_t \cdot l_t) = l_t^i$, where 1 is the $i$th component of $q_t$.

Following Kivinen and Warmuth (1999) we consider a more general notion of regret, which is the difference between cumulative loss of the algorithm and the loss of a given sequence of cumulative convex combinations $q_1, \dots, q_T$ of losses of the experts

$$R_T = \sum_{t=1}^{T} h_t - \sum_{t=1}^{T} (q_t \cdot l_t).$$

If $q_1 = \dots = q_T = e_i$, where $e_i = (0, \dots, 1, \dots, 0)$ is the unit vector, $R_T$ is the regret with respect to an expert $i$.

The derivation of upper bounds for regret is based on the following lemma of Bousquet and Warmuth (2002).

**Lemma 1** *For any comparison vector* $q_t \in \Gamma_N$,

$$m_t = (q_t \cdot l_t) + \frac{1}{\eta}(D(q_t \| w_t) - D(q_t \| w_t^m)). \tag{3}$$

*Proof.* By (2) the inequality (3) is obtained as follows:

$$m_t - \sum_{i=1}^N q_{i,t} l_t^i = \sum_{i=1}^N q_{i,t}\left(\frac{1}{\eta_t}\ln e^{-\eta_t l_t^i} + m_t\right) =$$

$$\frac{1}{\eta_t}\sum_{i=1}^N q_{i,t}\left(\ln e^{-\eta_t l_t^i} - \ln\sum_{j=1}^N w_{j,t} e^{-\eta_t l_t^j}\right) =$$

$$\frac{1}{\eta_t}\sum_{i=1}^N q_{i,t}\ln\frac{e^{-\eta_t l_t^i}}{\sum\limits_{j=1}^N w_{j,t}e^{-\eta_t l_t^j}} =$$

$$\frac{1}{\eta_t}\sum_{i=1}^N q_{i,t}\ln\frac{w_{i,t}^m}{w_{i,t}} = \frac{1}{\eta_t}(D(q_t\|w_t) - D(q_t\|w_t^m)).$$

Let us apply Lemma 1 for mixing schemes of Example 1 (Fixed Share): $\beta_t^{t+1} = 1-\alpha_{t+1}$, $\beta_s^{t+1} = 0$ for $0 < s < t$, and $\beta_0^{t+1} = \alpha_{t+1}$. Also, $w_{i,t+1} = \frac{\alpha_{t+1}}{N} + (1-\alpha_{t+1})w_{i,t}^m$ for all $i$ and $t$. In what follows, define $\alpha_t = \frac{1}{t+1}$ for all $t = 1, 2, \dots$.

Let a sequence $l_t = (l_t^1, \dots, l_t^N)$ of losses of the experts and a sequence of comparison vectors $q_t = (q_{1,t}, \dots, q_{N,t})$ be given online for $t = 1, 2, \dots$. Assume that $T$ be an arbitrary and the comparison vector $q_t$ changes $k$ times for $1 \le t \le T$.

We let $t_1 < t_2 < \dots t_k$ be the subsequence of indices in the sequence of comparators $q_1, \dots, q_T$, where shifting occurs: $q_{t_j} \ne q_{t_{j-1}}$ and $q_t = q_{t-1}$ for all other steps, where $t > 1$. Define also $t_0 = 1$ and $t_{k+1} = T + 1$. We apply Lemma 1 for the distribution $\beta^{t+1}$ from Example 1. Recall that $w_{i,1} = w_{i,0}^m = \frac{1}{N}$ for $i = 1, \dots, N$.

Summing (3) on time interval where $q_t = q_{t-1}$ for $t_j + 1 \le t \le t_{j+1} - 1$, we obtain

$$\sum_{t=t_j+1}^{t_{j+1}-1} m_t = \sum_{t=t_j+1}^{t_{j+1}-1}(q_t\cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1}\frac{1}{\eta_t}(D(q_t\|w_t) - D(q_t\|w_t^m)) =$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}(q_t\cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1}\frac{1}{\eta_{t-1}}D(q_t\|w_t) - \frac{1}{\eta_t}D(q_t\|w_t^m) +$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)D(q_t\|w_t) \le \tag{4}$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}(q_t\cdot l_t) + \sum_{t=t_j+1}^{t_{j+1}-1}\frac{1}{\eta_{t-1}}D(q_t\|w_t) - \frac{1}{\eta_t}D(q_t\|w_t^m) +$$

$$+\sum_{t=t_j+1}^{t_{j+1}-1}\delta_{t-1}\left(\ln N+\ln\frac{1}{\alpha_t}\right)\le \qquad (5)$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}(q_t\cdot l_t)+\sum_{t=t_j+1}^{t_{j+1}-1}\left(\frac{1}{\eta_{t-1}}D(q_t\|w^m_{t-1})-\frac{1}{\eta_t}D(q_t\|w^m_t)\right)+$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}\frac{1}{\eta_{t-1}}\ln\frac{1}{1-\alpha_t}+\sum_{t=t_j+1}^{t_{j+1}-1}\delta_{t-1}\ln\frac{1}{\alpha_t}+\ln N\sum_{t=t_j+1}^{t_{j+1}-1}\delta_{t-1}= \qquad (6)$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}(q_t\cdot l_t)+\left(\frac{1}{\eta_{t_j}}D(q_{t_j}\|w^m_{t_j})-\frac{1}{\eta_{t_{j+1}-1}}D(q_{t_j}\|w^m_{t_{j+1}-1})\right)+$$

$$\sum_{t=t_j+1}^{t_{j+1}-1}\frac{1}{\eta_t}\ln\frac{1}{1-\alpha_t}+\sum_{t=1}^{t_{j+1}-1}\delta_{t-1}\ln\frac{1}{\alpha_t}+\ln N\sum_{t=t_j+1}^{t_{j+1}-1}\delta_{t-1}. \qquad (7)$$

In transition from (4) to (5) the inequality $w_{i,t}\ge\frac{\alpha_t}{N}$ was used, then

$$D(q_t\|w_t)=\sum_{i=1}^N q_{i,t}\ln\frac{q_{i,t}}{w_{i,t}}\le\sum_{i=1}^N q_{i,t}\ln q_{i,t}-\sum_{i=1}^N q_{i,t}\ln\frac{\alpha_t}{N}\le\ln N+\ln\frac{1}{\alpha_t}. \qquad (8)$$

In transition from (5) to (6) we use the inequality (2), where $s=t-1$,

$$D(q_t\|w_t)\le D(q_t\|w^m_{t-1})+\ln\frac{1}{1-\alpha_t}.$$

In transition from (6) to (7) the entropy terms within the sections telescope and only for the beginning and the end of each section a positive and a negative entropy term remains, respectively.

For the beginnings of the $k$ sections $t=t_1,\dots,t_k$ define $s=0$, $\beta_0^{t_j}=\alpha_{t_j}$ in the inequality (2), then

$$m_{t_j}\le(q_{t_j}\cdot l_{t_j})+\frac{1}{\eta_{t_j}}D(q_{t_j}\|w^m_0)-\frac{1}{\eta_{t_j}}D(q_{t_j}\|w^m_{t_j})+\frac{1}{\eta_{t_j}}\ln\frac{1}{\alpha_{t_j}}. \qquad (9)$$

Summing all these inequalities, we obtain

$$M_T-\sum_{t=1}^T(q_t\cdot l_t)\le\sum_{j=1}^k\left(\frac{1}{\eta_{t_j}}D(q_{t_j}\|w^m_0)-\frac{1}{\eta_{t_{j+1}-1}}D(q_{t_j}\|w^m_{t_{j+1}-1})\right)+ \qquad (10)$$

$$\sum_{t=1}^T\frac{1}{\eta_t}\ln\frac{1}{1-\alpha_t}+\sum_{t=1}^T\delta_{t-1}\ln\frac{1}{\alpha_t}+\ln N\sum_{t=1}^T\delta_{t-1}+\sum_{j=1}^k\Delta_{t_j-1}\ln\frac{1}{\alpha_{t_j}}\le$$

$$((k+2)\ln(T+1)+(k+1)\ln N+1)\Delta_T. \qquad (11)$$

In transition from (10) to (11) we use inequality $D(q\|w^m_T)\ge 0$ for all $q$ and equality $D(q\|w^m_0)=\ln N$. Then $\sum\limits_{j=1}^k\frac{1}{\eta_{t_j}}D(q_{t_j}\|w^m_0)\le k\ln N\Delta_T$. For $\alpha_t=\frac{1}{t+1}$ we use inequalities

$$\sum_{t=1}^T\frac{1}{\eta_t}\ln\frac{1}{1-\alpha_t}\le\frac{1}{\eta_T}(\ln T+1)=\Delta_{T-1}(\ln T+1)\le\Delta_T(\ln T+1).$$

Therefore, we have proved the following theorem.

**Theorem 2** *Let $\alpha_t = \frac{1}{t+1}$ for all $t$ and mixing scheme from Example 1 was used. Then for any $T$, for any sequence of losses of the experts, and for any sequence of comparison vectors $q_t \in \Gamma_N$ given online with no more than $k$ changes,*

$$M_T \le \sum_{t=1}^{T}(q_t \cdot l_t) + ((k+2)\ln(T+1) + (k+1)\ln N + 1)\Delta_T.$$

*Besides,*

$$H_T \le \sum_{t=1}^{T}(q_t \cdot l_t) + ((k+2)\ln(T+1) + (k+1)\ln N + 2)\Delta_T.$$

Denote $\gamma(T) = (k+2)\ln(T+1) + (k+1)\ln N + 2$.

The corresponding bound for Example 2 can be obtained analogously. Since $w_{i,t} \ge \frac{\alpha_t}{Nt}$, the inequality (8) can be replaced by inequality $D(q_t \| w_t) \le \ln N + \ln T + \ln \frac{1}{\alpha_t}$. Also, the the last term of the inequality (9) can be replaced by $\frac{1}{\eta_{t_j}} \ln \frac{1}{t\alpha_{t_j}}$. As a result, we get $\gamma(T) = (2k+3)\ln(T+1) + (k+1)\ln N + 2$.

We use notations $l_t^- = \min_{1\le i\le N} l_t^i$, $l_t^+ = \max_{1\le i\le N} l_t^i$, $s_t = l_t^+ - l_t^-$, $L_T^- = \sum\limits_{t=1}^{T} l_t^-$, $L_T^+ = \sum\limits_{t=1}^{T} l_t^+$, $S_T = \max_{1\le t\le T} s_t$, and $L_T^{(k)} = \sum\limits_{t=1}^{T}(q_t \cdot l_t)$, where $q_t$ is a comparison vector and $k$ is the number of $t \le T$ such that $q_t \ne q_{t-1}$.

The proof of Theorem 8 from Rooij et al. (2014) can be easily modified (see Section 4) to obtain the following upper bound for $\Delta_T$: for any $T$ and $1 \le k \le T$,

$$\Delta_T \le \sqrt{S_T \frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}} + \left(\gamma(T) + \frac{5}{3}\right) S_T. \tag{12}$$

Therefore, we have the main result of this paper:

**Theorem 3** *For any $T$ and for any sequence of comparison vectors $q_t \in \Gamma_N$ with no more than $k$ changes given online,*

$$R_T^{(k)} = H_T - L_T^{(k)} \le \gamma(T)\sqrt{S_T \frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}} + \gamma(T)\left(\gamma(T) + \frac{5}{3}\right) S_T, \tag{13}$$

*where $\gamma(T) = (k+2)\ln(T+1) + (k+1)\ln N + 2$ for scheme of Example 1.*

In particular, the same upper bound holds for compound experts $i_1, \dots, i_T$ (sequence of experts, $1 \le i_t \le N$ for $1 \le t \le T$). The complexity $s(i_1, \dots, i_T)$ of this compound expert is equal to the number of $1 < t \le T$ such that $i_{t-1} \ne i_t$. Let also, the comparison vectors $q_t$ be unit vectors and $q_{t-1} \ne q_t$ for $\le k$ distinct $1 < t \le T$. In this case, since for some sequence $\{q_t\}$ of unit vectors

$$L_T^{(k)} = \sum_{t=1}^{T}(q_t \cdot l_t) = \min_{s(i_1,\dots,i_T)\le k} \sum_{t=1}^{T} l_t^{i_t},$$

we can rewrite the bound (13) in the form

$$H_T \leq \min_{s(i_1,\dots,i_T)\leq k} \sum_{t=1}^{T} l_t^{i_t} +$$

$$\gamma(T)\sqrt{S_T\frac{(L_T^+ - L_T^{(k)})(L_T^{(k)} - L_T^-)}{L_T^+ - L_T^-}} + \gamma(T)\left(\gamma(T)+\frac{5}{3}\right)S_T,$$

where $\gamma(T) = (k+2)\ln(T+1) + (k+1)\ln N + 2$ for scheme of Example 1.

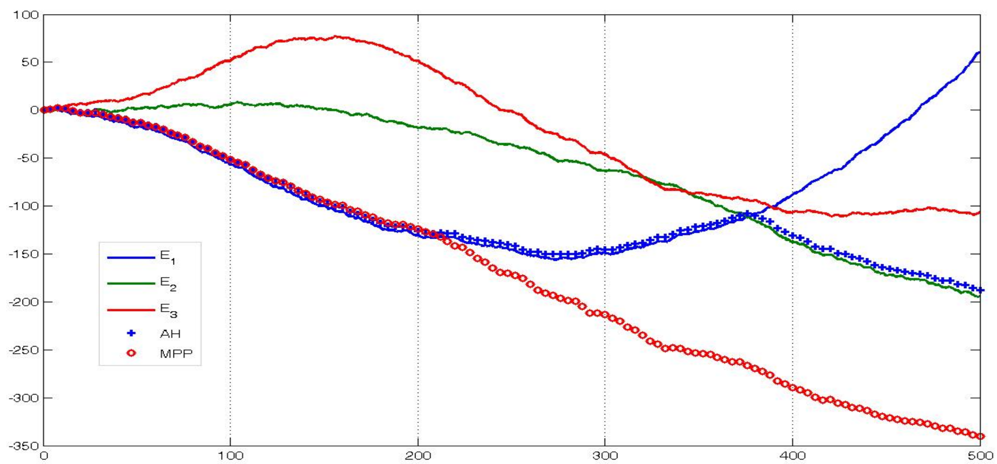


Figure 2: Results of the experiment which was performed on artificial data. Three blue, green, and red lines – experts $E_1$, $E_2$, $E_3$ cumulative losses, AdaHedge losses – thick blue line, MPP losses – thick red line.

## 3. Numerical experiments

In this section we present results of experiments for scheme of Example 1.

The first experiment is performed on artificial data, where one-step losses of experts are signed and unbounded. The cumulative losses of three experts and losses of aggregating algorithms AdaHedge and MPP are presented on Figure 2. In this case MPP algorithm significantly outperforms AdaHedge.

The next three experiments are performed on financial data which was downloaded from website `http://finam.ru`. In this case losses of the experts cannot be bounded in advance. A characteristic feature of these examples is that the strong volatility of the experts' losses was observed. To be closer to financial terminology, we represent results in terms of incomes (gains) instead of losses: we interpreted loss as a negative gain. The first of these experiments is based on ideas of Delbaen and Schachermayer (1994) and Vovk (2003) (see also a similar example in V'yugin 2013).

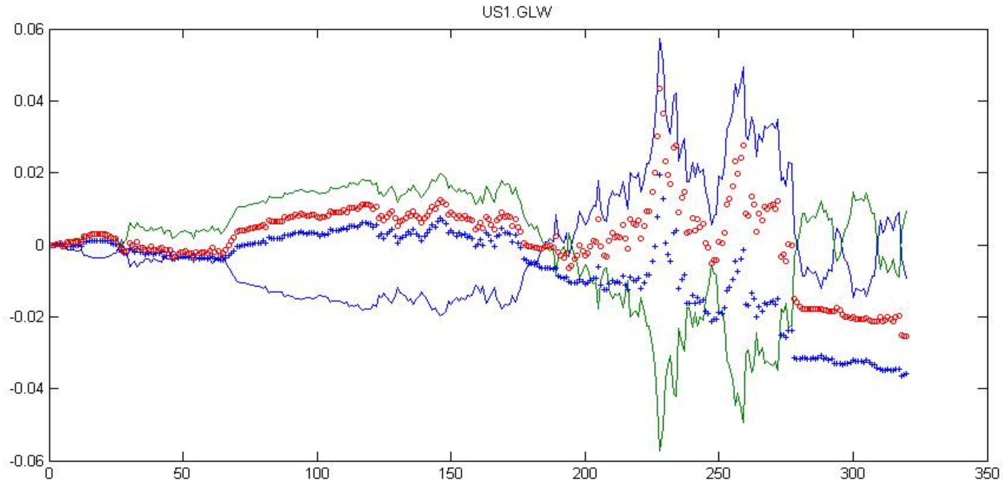


Figure 3: Zero-Sum game for US1.GLW stock. Two symmetric green and blue lines – experts income, AdaHedge relative income – thick blue line, MPP relative income – thick red line.

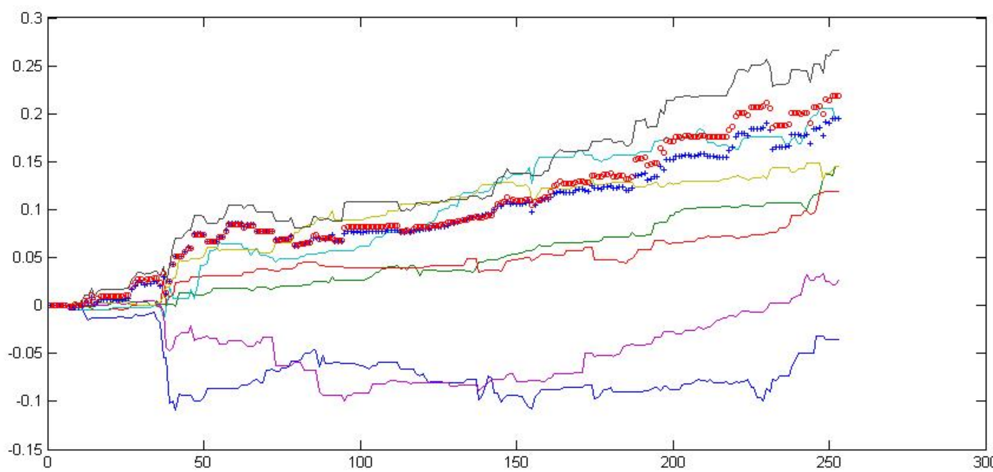


Figure 4: Russian stocks: AdaHedge income – thick blue line, MPP income – thick red line. Vertical line show a relative cumulative income or loss for each expert and hedging trading strategies, horizontal lines show time in days.

Let $S_0, S_1, \ldots, S_T$ be a sequence stock prices. Let also, $\Delta S_t = S_{t+1} - S_t$. Experts 1 and 2 represent two concurrent methods of buying and selling shares of this stock. The true equality $(S_T - S_0)^2 - \sum_{t=0}^{T-1} (\Delta S_t)^2 = \sum_{t=0}^{T-1} 2(S_t - S_0)\Delta S_t$ leads to the two experts strategies.

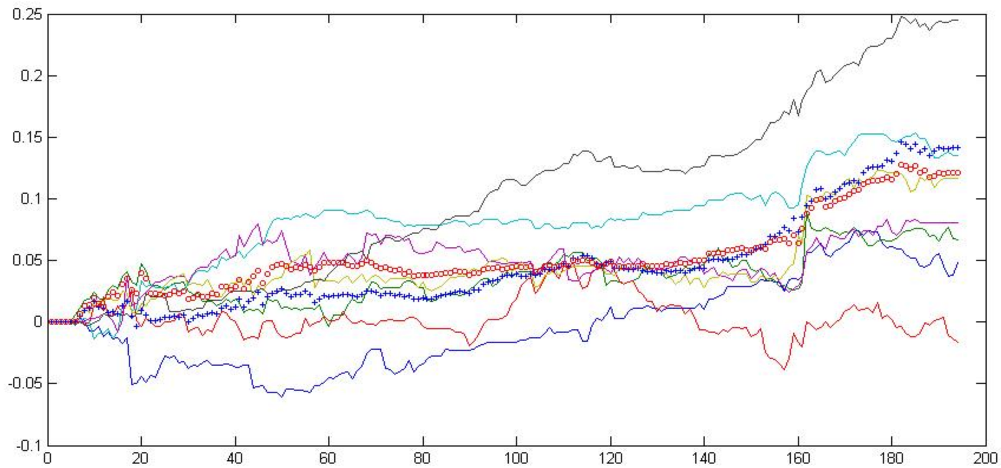


Figure 5: BATS Electronic Market stocks: AdaHedge income – thick blue line, MPP income – thick red line. Vertical line show a relative cumulative income or loss for each expert and hedging trading strategies, horizontal lines show time in days.

At the beginning of step $t$ Experts 1 and 2 hold the number of shares $C_t^1 = 2C(S_t - S_0)$ and $C_t^2 = -C_t^1$, where $C$ is an arbitrary positive constant; at the end of this step they earn the incomes $s_t^1 = 2C(S_t - S_0)\Delta S_t$ and $s_t^2 = -s_t^1$. The cumulative income of Expert 1 in first $T$ steps is $s_{1:T}^1 = \sum\limits_{t=1}^{T} s_t^1 = 2C\left((S_T - S_0)^2 - \sum\limits_{t=1}^{T-1}(\Delta S_t)^2\right)$, and cumulative income of Expert 2 is $s_{1:T}^2 = -s_{1:T}^1$. The number of shares $C_t^1$ and $C_t^2 = -C_t^1$ can be positive or negative, i.e, short selling is allowed. The one-step gains $s_t^1$ and $s_t^2 = -s_t^1$ are unbounded and can be positive or negative. The hedging strategy AdaHedge or MPP at any step divide his investment in proportion to the weights of expert strategies. At any step, the hedging strategy use the same amount of money for investment. Thus, we postpone extra income in cash or borrow money. Figure 3 show results of experiment performed with stock US1.GLW of BATS Electronic Market (trading performing twice at each working day of 2015).

Next two experiments are performed with two sets of stocks which were downloaded from website `http://finam.ru`. These sets are Russian stocks with tickers: VTBR, GAZP, LKOH, NVTK, ROSN, SBER, TRNFP and BATS Electronic Market stocks with tickers: US1T, US1AXP, US1BA, US1GE, US2GOOG, US1JPM, US1KO (trading performing at each working day of 2014). Seven Expert strategies are represented by daily incomes of Universal Algorithmic Trading strategy which operates for each of these stocks separately. We refer readers to V'yugin (2013) for details of this strategy. As in the first experiment, at each step the aggregating strategy AdaHedge or MPP distribute the same amount of money among these experts and receive income or suffer loss in proportion to the weights of expert strategies; we postpone extra income in cash or borrow money when suffer loss. Figures 4 and 5 show cumulative incomes of all strategies by days of trading period in 2014.

Vertical lines on plots show a proportion of cumulative income or loss for each expert and hedging trading strategy, horizontal lines show time in days.

The version of MPP algorithm presented in this paper gives us a tool for aggregating the losses of experts in difficult cases that arise in practice. The results of experiments show that MPP algorithm significantly outperforms AdaHedge on artificial data and that MPP algorithm slightly outperforms AdaHedge algorithm even in cases where losses (gains) of experts are volatile, signed, and unbounded in advance.

## Acknowledgments

This work was supported by the Russian Science Foundation, project no. 14-50-00150.

Author thanks Vladimir Trunov for performing experiments.

The author is grateful to anonymous referees for corrections and suggestions on improving the presentation of this paper.

## References

O. Bousquet, M. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*. 3:363–396, 2002.

F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.

Y. Freund, R.E. Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*. 55:119–139, 1997.

N. Cesa-Bianchi, G. Lugosi. Prediction, Learning, and Games. Cambridge University Press. 2006.

N. Cesa-Bianchi, Y. Mansour, and G. Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2/3):321352, 2007.

M. Herbster, M. Warmuth. Tracking the best expert. *Machine Learning*, 32(2):151178, 1998.

J. Kivinen, M.K. Warmuth. Averaging expert prediction. In Paul Fisher and Hans Ulrich Simon, editors, Computational Learning Theory: 4th European Conference (EuroColt '99), 153–167. Springer, 1999.

N. Littlestone, M. Warmuth. The weighted majority algorithm. *Information and Computation*. 108:212–261, 1994.

S. de Rooij, T. van Erven, P. Grunwald, W. Koolen. Follow the Leader. If You Can, Hedge If You Must. *Journal of Machine Learning Research*. 15:1281–1316, 2004.

V. Vovk, Aggregating strategies. In M. Fulk and J. Case, editors, Proceedings of the 3rd Annual Workshop on Computational Learning Theory, 371–383. San Mateo, CA, Morgan Kaufmann, 1990.

V. Vovk, A game of prediction with expert advice. *Journal of Computer and System Sciences.* 56(2):153–173, 1998.

V. Vovk. Derandomizing stochastic prediction strategies. *Machine Learning.* 35(3):247-282, 1999.

V. Vovk. A game-theoretic explanation of the $\sqrt{dt}$ effect. Working paper #5, Available online at http://www.probabilityandfinance.com, 2003.

V. V'yugin. Online Learning in Case of Unbounded Losses Using Follow the Perturbed Leader Algorithm. *Journal of Machine Learning Research* 12:241–266, 2011.

V. V'yugin. Universal Algorithm for Trading in Stock Market Based on the Method of Calibration. *Lecture Notes in Artficial Intellinges*, 8139:53–67, 2013.

## 4. Supplementary material

We use a slightly different procedure for experts weights update than the one what was used by Rooij et al. (2014). Nevertheless, the proof does not change. In this section we reproduce the derivation of the upper bound (12) of Theorem 8 by Rooij et al. (2014) with an exception that experts weights update and some constants are changed.

The corresponding derivation is based on Bernstein inequality. Let $V_T = \sum\limits_{t=1}^{T} v_t$ be cumulative variance of losses of the experts, where

$$v_t = Var_{j\sim w_t}[l_t^j] = E_{j\sim w_t}[(l_t^j - E_{j\sim w_t}[(l_t^j)])^2] = \sum_{j=1}^{N} w_{j,t}(l_t^j - h_t)^2$$

and $w_t = (w_{1,t}, \dots, w_{N,t})$ be the experts weights.

Denote $s_t = l_t^+ - l_t^-$, where $l_t^+ = \max_{1\le j\le N} l_t^j$, $l_t^- = \min_{1\le j\le N} l_t^j$.

**Lemma 4** *The difference $\delta_t = h_t - m_t$ satisfies*

$$\delta_t \le \frac{e^{s_t\eta_t} - 1 - s_t\eta_t}{\eta_t s_t^2} v_t. \tag{14}$$

*Proof.* The proof is based on Bernstein inequality (see Lemmas 3-5 of Cesa-Bianchi and Lugosi (2006)).

Let $X \in (-\infty, 1)$ be a random variable, where $EX = 0$, $EX^2 = \sigma^2$. Bernstein inequality says that for any $\eta > 0$, $\ln Ee^{\eta X} \le \sigma^2(e^\eta - \eta - 1)$. In what follows, we use also the following form of this inequality. Let $X \in [0, 1]$ be a random variable and $\sigma = \sqrt{EX^2 - (EX)^2}$. The for any $\eta > 0$, $\ln E[e^{-\eta(X-EX)}] \le \sigma^2(e^\eta - \eta - 1)$. Consider a random variable $X$ taking values $l_t^j$ with probabilities $w_{j,t}$, where $j = 1, \dots, N$, and its transformation $X_t^j = \frac{l_t^j - l_t^-}{s_t} \in [0, 1]$. Then Bernstein inequality has the form: $\ln E_{j\sim w_t}\left(e^{-\eta\left(X_t^j - EX_t\right)}\right) \le \sigma^2\left(e^\eta - 1 - \eta\right)$ for all

$\eta > 0$. Rewrite this inequality in more details, where $\eta = s_t\eta_t$:

$$\ln E_{j\sim w_t}\left(e^{-\eta\left(X_t^j - EX_t\right)}\right) = \ln\left(\sum_{j=1}^N w_{t,j} e^{-s_t\eta_t\left(\frac{l_t^j - l_t^-}{s_t} - \sum_{i=1}^N w_{i,t}\frac{l_t^i - l_t^-}{s_t}\right)}\right) =$$

$$\ln\left(\frac{\sum_{j=1}^N w_{j,t} e^{-\eta_t\left(l_t^j - l_t^-\right)}}{e^{-\eta_t\sum_{j=1}^N w_{j,t}\left(l_t^j - l_t^-\right)}}\right) = \ln\left(\frac{\sum_{j=1}^N w_{j,t} e^{-\eta_t l_t^j}}{e^{-\eta_t\sum_{j=1}^N w_{j,t} l_t^j}}\right) =$$

$$\ln\sum_{j=1}^N w_{j,t} e^{-\eta_t l_t^j} + \eta_t\sum_{j=1}^N w_{j,t} l_t^j = \eta_t(h_t - m_t) = \eta_t\delta_t \le \sigma^2\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right) =$$

$$Var_{j\sim w_t}[X_t^j]\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right) = \frac{1}{s_t^2} Var_{j\sim w_t}[l_t^j]\left(e^{s_t\eta_t} - 1 - s_t\eta_t\right).$$

From this, we obtain the needed inequality $\delta_t \le \frac{e^{s_t\eta_t} - 1 - s_t\eta_t}{\eta_t s_t^2} v_t$. Lemma is proved.

Rewrite (14) in the form

$$\delta_t \le \frac{g(s_t\eta_t)}{s_t} v_t, \text{ where } g(x) = \frac{e^x - x - 1}{x}. \tag{15}$$

Recall that $\eta_t = \frac{1}{\Delta_{t-1}}$.

**Lemma 5** *It holds* $(\Delta_T)^2 \le V_T + \frac{5}{3} S_T\Delta_T$, *where* $S_T = \max_{1\le t\le T}(l_t^+ - l_t^-)$.

*Proof.* It holds

$$(\Delta_T)^2 = \sum_{t=1}^T\left(\Delta_t^2 - \Delta_{t-1}^2\right) = \sum_{t=1}^T\left((\Delta_{t-1} + \delta_t)^2 - \Delta_{t-1}^2\right) =$$

$$\sum_{t=1}^T\left(2\delta_t\Delta_{t-1} + \delta_t^2\right) = \sum_{t=1}^T\left(\frac{2\delta_t}{\eta_t} + \delta_t^2\right) \le$$

$$\sum_{t=1}^T\left(\frac{2\delta_t}{\eta_t} + s_t\delta_t\right) \le 2\sum_{t=1}^T\frac{\delta_t}{\eta_t} + S_T\Delta_T. \tag{16}$$

Now obtain $\frac{\delta_t}{\eta_t}$ using (15):

$$\frac{1}{2}v_t \ge \frac{\delta_t s_t}{2g(s_t\eta_t)} = \frac{\delta_t}{\eta_t} + A,$$

$$A = \frac{\eta_t s_t\delta_t - 2g(s_t\eta_t)\delta_t}{2g(s_t\eta_t)\eta_t} = \frac{s_t\left(\eta_t^2\delta_t s_t^2 - 2\delta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)\right)}{2s_t\eta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)} =$$

$$= s_t\delta_t\frac{\frac{1}{2}(s_t\eta_t)^2 - e^{s_t\eta_t} + s_t\eta_t + 1}{s_t\eta_t\left(e^{s_t\eta_t} - s_t\eta_t - 1\right)} = -\varphi(s_t\eta_t)\, s_t\delta_t,$$

where $\varphi(x)=\frac{e^x-\frac{1}{2}x^2-x-1}{xe^x-x^2-x}$. Using Taylor expansion, we obtain $\varphi(x)\le 1/3$. Then

$$\frac{\delta_t}{\eta_t}\le\frac{1}{3}s_t\delta_t+\frac{1}{2}v_t. \tag{17}$$

Substituting the (17) in the inequality (16) and, summing up, we get

$$(\Delta_T)^2\le V_T+\frac{5}{3}S_T\Delta_T.$$

Lemma is proved.

Now we estimate $V_T$. By definition $v_t\le(l_t^+-h_t)(h_t-l_t^-)\le\frac{s_t^2}{4}$.

**Lemma 6** *Let $k\le T$, $S_T=\max_{1\le t\le T}s_t$ and $L_T^{(k)}\le H_T$. Then $V_T\le S_T\frac{(L_T^+-L_T^{(k)})(L_T^{(k)}-L_T^-)}{L_T^+-L_T^-}+\gamma(T)S_T\Delta_T$.*

*Proof.* The following inequalities are valid:

$$V_T=\sum_{t=1}^T v_t\le\sum_{t=1}^T(l_t^+-h_t)(h_t-l_t^-)\le S_T\sum_{t=1}^T\frac{(l_t^+-h_t)(h_t-l_t^-)}{s_t}=$$
$$S_T T\sum_{t=1}^T\frac{1}{T}\frac{(l_t^+-h_t)(h_t-l_t^-)}{(l_t^+-h_t)+(h_t-l_t^-)}\le S_T\frac{(L_T^+-H_T)(H_T-L_T^-)}{L_T^+-L_T^-}.$$

Here an instance of Jensen inequality was used (see for details Rooij et al. (2014)). Lemma is proved.

By Theorem 2 $H_T\le L_T^{(k)}+\gamma(T)\Delta_T$. Assume that $L_T^{(k)}\le H_T$. Using these inequalities and Lemma 6 we obtain

$$V_T\le S_T\frac{(L_T^+-H_T)(H_T-L_T^-)}{L_T^+-L_T^-}\le$$
$$S_T\frac{(L_T^+-L_T^k)(L_T^{(k)}+\gamma(T)\Delta_T-L_T^-)}{L_T^+-L_T^-}\le$$
$$S_T\frac{(L_T^+-L_T^{(k)})(L_T^{(k)}-L_T^-)}{L_T^+-L_T^-}+\gamma(T)S_T\Delta_T. \tag{18}$$

Denote $A_T=\frac{(L_T^+-L_T^{(k)})(L_T^{(k)}-L_T^-)}{L_T^+-L_T^-}$. By inequality (18) and Lemma 5 $\Delta_T^2\le S_TA_T+\left(\gamma(T)+\frac{5}{3}\right)S_T\Delta_T$. This is an inequality of the form $\Delta_T^2\le a+b\Delta_T$, where $a=S_TA_T$, $b=(\gamma(T)+\frac{5}{3})S_T$. Solving this inequality with respect to $\Delta_T$, we have

$$\Delta_T\le\frac{1}{2}b+\frac{1}{2}\sqrt{b^2+4a}\le\frac{1}{2}b+\frac{1}{2}\left(\sqrt{b^2}+\sqrt{4a}\right)=$$
$$\sqrt{a}+b=\sqrt{S_TA_T}+\left(\gamma(T)+\frac{5}{3}\right)S_T.$$

Then $R_T^{(k)}=H_T-L_T^{(k)}\le\gamma(T)\Delta_T\le\gamma(T)\sqrt{S_TA_T}+\gamma(T)\left(\gamma(T)+\frac{5}{3}\right)S_T$,where $\gamma(T)=(k+2)\ln(T+2)+(k+1)\ln N+2$, and inequality (13) is valid.

If $H_T\le L_T^{(k)}$ then $R_T^{(k)}\le 0$ and the inequality (13) is valid again.