Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Stochastic Approximation Algorithms for Trailing Stop

G. Yin, Q. Zhang, and C. Zhuang

Abstract— **Trailing stops are often used in stock trading to limit a maximum-possible loss and to lock in a profit. In this venue, it is important to identify the optimal trailing stop percentage, which is difficult to find and no apparent analytic technique can be applied directly. This work develops stochastic approximation algorithms to estimate the optimal trailing stop percentage. A modification using projection is also proposed to ensure the approximation sequence constructed to stay in a bounded region. Convergence of the algorithm is obtained. Moreover, interval estimates are constructed. Simulation examples are presented to compare our algorithm with Monte Carlo methods. Finally, real market data are used to demonstrate the algorithms.**

Index Terms— **Trailing stop, stochastic approximation, stochastic optimization.**

I. INTRODUCTION

Decision in selling a stock is crucial in successful investing in equity markets. The selling strategy can be determined by either a target level or a stop-loss limit. In this paper, we focus on the stop-loss side of the equation. In equity trading, a stop-loss order is an order placed with a broker to sell the equity when the stock price drops to a certain level. A stop loss is designed to limit the investor's loss on a security position. The advantage of a stop-loss order is that one need not monitor the market constantly on how the a stock is performing. A disadvantage is that the stop-loss order cannot help the investor to lock in his profits after a substantial rise in price. A key in successful trading is to "cut the losses and let the profit run." Such needs give rise to the so-called trailing stop. Before proceeding further, let describe the idea of trailing stop briefly. We set the stop price at a trailing stop percentage h as $[(1-h)\times$ maximal price of the stock up to time $t]$, where $0 < h < 1$ is known as the stop percentage. The first time the stock price reaches the stop price, we sale the stock. The trailing stop maintains a stop-loss order at a precise percentage below the market price. The stop-loss order is adjusted continually based on fluctuations of the market price, which always maintains the same percentage below the market price. The trader is then "guaranteed" to know the exact minimum profit that his or her position will garner.

As the market price advances, the stop price also rises. Should the price decline, the stop price does not change, and the position is closed whenever the stop price is reached. For example, assume an investor buy the Coca Cola (KO) on January 2, 1998 at the price of \$57.31. If he/she sets the trailing stop at 15% below the market price, then the initial stop price is \$48.71 (15% below the market price). When the stock price advances from \$57.31, the trailing stop (marked in Fig. 1 by dashed line) rises accordingly and stays 15% off the highest price. On August 27, 1998, the price drops from the peak back to \$64.24 (the first time below the trailing stop price). The investor should close his/her position at this point, resulting in a raw return of 12.09%. The closing price of KO and the corresponding trailing stop levels are given in Figure 1. Clearly, trailing stop is an effective tool helping

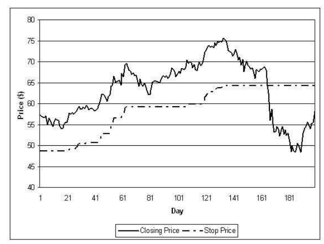

Fig. 1. The Prices of Coca Cola from Jan 12, 1998, to October 15, 1998

the investor to lock in the profit when the market moves against his or her position. Traditionally, a trailing stop percentage is determined based on the trader's predilection toward aggressive or conservative trading. In stock investing, deciding what constitutes appropriate profits (or acceptable losses) is perhaps the most difficult aspect of establishing a trailing-stop system.

Research using mathematical models on trailing stops is scarce in the literature. Glynn and Iglehart [4] studied the problem in both discrete and continuous time. In the continuous-time case, they considered a diffusion model and showed that the optimal strategy is not to sell at all, i.e., $h_* = 100\%$, which corresponds to the so-called buy and hold strategy. The stock price can become negative in their model. It also appears difficult to extend their results to a reasonable market model such as the geometric Brownian motion model.

It is the purpose of this paper to study optimal trailing stops. Here, the main issue is to determine the optimal stop percentage h_*. We develop a stochastic optimization approach. It provides a systematic way to compute the optimal trailing stop percentage h_*. The SA approach is effective in real time because of its recursive form. A main advantage of the SA approach is that there is no price model needed and one only needs the stock prices to come up with desired percentage.

This research was supported in part by the National Science Foundation.
G. Yin is with Department of Mathematics, Wayne State University, Detroit, MI 48202. gyin@math.wayne.edu

Q. Zhang is with Department of Mathematics, Boyd GSRC, The University of Georgia, Athens, GA 30602-7403. qingz@math.uga.edu

C. Zhuang is with Department of Mathematics, Boyd GSRC, The University of Georgia, Athens, GA 30602-7403. czhuang@math.uga.edu

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

To some extent, the iterates obtained using the recursive algorithm can be thought of as a point estimator of the true optimal trailing percentage. It would be nice if we also provided the quality of the estimation sequence. We do this by considering a confidence interval estimate. For related work on stopping rules for Robbins-Monro type stochastic approximation algorithms for root finding, we refer the reader to [8]. Here the purposes of our interval estimates are two folds. (i) They provide a practically useful range of estimations, and ensures that the limit confidence level is the desired one. (ii) They give an implementable stopping criterion for the iterates; with large probability, the iterates will be terminated when the criterion is met. Crucial to the development of the confidence estimate is asymptotic distribution of a scaled sequence of estimation errors. Furthermore, instead of examining the discrete iterates directly, we focus on continuous-time interpolations leading to diffusion limits. Among other things, a random time change argument is used to deduce the result when the deterministic iteration number is replaced by a random variable. For comparison purposes, we use a Monte Carlo method to obtain optimal percentage rates. In addition, we demonstrate our results using real market data.

The rest of this paper is arranged as follows. Section 2 begins with the precise formulation of the problem. Also provided are the recursive algorithm and its variations. Section 3 studies the convergence of the underlying algorithm. Section 4 concentrates on interval estimates. To demonstrate the feasibility and efficiency of the algorithms, numerical experiments using simulations and real market data are given in Section 5. These algorithms provide sound estimates of optimal trailing stop percentage; they can be easily implemented in real time. Finally we conclude this paper with some further remarks in Section 6.

II. FORMULATION

Normally, when one treats stock market models, one always proposes some models for the stock prices, for example, the well-known geometric Brownian motion model, the mean-reversion model, and the jump-diffusion model etc. Compared with the traditional approach, in our formulation, we shall not require the stock price $S(t)$ to be any of the models above. In fact, we treat model free case and only assume the stock price being observable. Based on the observed stock price, for a given time t, we define the stop price at a trailing stop percentage h with $0 < h < 1$ as

$$T_h(t) = (1-h)S_{\max}(t), \tag{II.1}$$

where $S_{\max}(t) = \max\{S(u) : 0 \le u \le t\}$. Let

$$\tau = \inf\{t > 0 : S(t) \le T_h(t)\}. \tag{II.2}$$

Then τ is the first time the stock price reaches the stop price. We aim to find the optimal trailing stop percentage $h_* \in [a,1]$ with $a > 0$ that maximize a suitable objective function. Thus the problem is

$$\text{Find argmax } J(h) = E[\Phi(S(\tau))\exp(-\rho\tau)], \quad h \in [a,1]. \tag{II.3}$$

Here $a > 0$ is a reasonable lower bound for the trailing stop percentage, $\rho > 0$ is an appropriate discount rate, and the reward function $\Phi(S) = \frac{S-S_0}{S_0}$. In general, an analytic solution is difficult to obtain even if $S(t)$ follows a particular model, for example, a geometric Brownian motion. Our contribution is to devise a numerical approximation procedure that estimates the optimal trailing stop percentage h. We will use a stochastic approximation procedure to resolve the problem by constructing a sequence of estimates of the optimal trailing stop percentage h, using $h_{n+1} = h_n + \{\text{step size}\}\{\text{gradient estimate of } J(h)\}$. Moreover, in accordance with (II.3), we need to make sure the iterate $h_n \in [a,1]$.

A. Recursive Algorithm

Let us begin with a simple noisy finite difference scheme. The only provision is that $S(t)$ can be observed. Associated with the iteration number n, denote the trailing stop percentage by h_n. Beginning at an arbitrary initial guess, we construct a sequence of estimates $\{h_n\}$ recursively as follows. We figure out τ_n, the first time when the stock price declines under the stop price as

$$\tau_n = \inf\{t > 0 : S(t) \le T_{h_n}(t)\}. \tag{II.4}$$

Define a process ξ_n that includes the random effect from observed $S(t)$, or the random effect from a simulation, and the stopping time τ_n, $\xi_n = (S(\tau_n), \tau_n)'$, where $S(\tau_n)$ denotes the stock price process $S(t)$ stopped at stopping time τ_n. We loosely call $\{\xi_n\}$ the sequence of noise. Let $\widetilde{J}(h,\xi)$ be the observed value of the objective function $J(h)$ with noise ξ. With the values $h \pm \delta_n$, define $Y_n^{\pm}$ as $Y_n^{\pm}(h, \xi_n^{\pm}) = \widetilde{J}(h \pm \delta_n, \xi_n^{\pm})$. $\xi_n^{\pm}$ being the two different collective noises taken at the trailing stop percentages $h \pm \delta_n$, where δ_n is the finite difference sequence satisfying $\delta_n \to 0$ as $n \to \infty$. We shall write $Y_n^{\pm} = Y_n^{\pm}(h, \xi_n^{\pm})$. For simplicity, in what follows, we often use ξ_n to represent both ξ_n^+ and ξ_n^- if there is no confusion. The gradient estimate at iteration n is given by

$$D\widehat{J}(h_n, \xi_n) \overset{\text{def}}{=} (Y_n^+ - Y_n^-)/(2\delta_n). \tag{II.5}$$

Then the recursive algorithm is

$$h_{n+1} = h_n + \varepsilon_n D\widehat{J}(h_n, \xi_n), \tag{II.6}$$

where ε_n is a sequence of real numbers known as step sizes. A frequently used choice of step size and finite difference sequences is $\varepsilon_n = O(1/n)$ and $\delta_n = O(1/n^{1/6})$. Throughout this paper, this is our default choice of step size and finite difference sequences.

To proceed, define

$$\begin{aligned}
\rho_n &= (Y_n^+ - Y_n^-) - E_n(Y_n^+ - Y_n^-), \\
\eta_n &= [E_n Y_n^+ - J(h_n + \delta_n)] - [E_n Y_n^- - J(h_n - \delta_n)], \\
\beta_n &= \frac{J(h_n + \delta_n) - J(h_n - \delta_n)}{2\delta_n} - J_h(h_n),
\end{aligned} \tag{II.7}$$

where E_n denotes the conditional expectation with respect to $\mathcal{F}_n$, the σ-algebra generated by $\{h_1, \xi_j^{\pm} : j < n\}$,

$J_h(h_n) = (\partial/\partial h)J(h_n)$. In the above, η_n and β_n represent the noise and bias, and $\{\rho_n\}$ is a martingale difference sequence. We separate the noise into two parts, uncorrelated noise ρ_n and correlated noise η_n. It is reasonable to assume that after taking the conditional expectations, the resulting function is smooth. With the above definitions, algorithm (II.6) can be rewritten as

$$h_{n+1} = h_n + \varepsilon_n J_h(h_n) + \varepsilon_n \frac{\rho_n}{2\delta_n} + \varepsilon_n \beta_n + \varepsilon_n \frac{\eta_n(h_n, \xi_n)}{2\delta_n}. \tag{II.8}$$

B. Projection Algorithms

The use of projections in the algorithms stems from two reasons. First, for the purpose of computations, it is more convenient if one uses projections to force the iterates to remain in a bounded region. In addition, the problems under consideration may well be constrained so that the iterates will be in a given set. Current problem under consideration is such an example (the iterates need to stay in the interval $[a, 1]$). For example, one might choose a lowest trailing stop percentage of 10% to ensure the holding position will not be closed due to the normal fluctuations of daily stock price. Obviously, there is a upper bound for the optimal trailing stop percentage, 100%. To solve the Problem (II.3) with constrains, we construct the following stochastic approximation algorithm with a projection

$$h_{n+1} = \Pi[h_n + \varepsilon_n D\widetilde{J}(h_n, \xi_n)], \tag{II.9}$$

where $\varepsilon_n = 1/n$, $\delta_n = \delta/(n^{1/6})$ and $\Pi[x]$ is a projection given by $\Pi[h] = \begin{cases} a, & \text{if } h < a, \\ 1, & \text{if } h > 1, \\ h, & \text{otherwise.} \end{cases}$ As explained in Kushner and Yin [6], The projection algorithm (II.9) can be rewritten as

$$h_{n+1} = h_n + \varepsilon_n D\widehat{J}(h_n, \xi_n) + \varepsilon_n r_n. \tag{II.10}$$

where $\varepsilon_n r_n = h_{n+1} - h_n - \varepsilon_n D\widehat{J}(h_n, \xi_n)$ is the real number with the shortest distance needed to bring $h_n + \varepsilon_n D\widehat{J}(h_n, \xi_n)$ back to the constraint set $[a, 1]$ if it is outside this set.

III. Convergence

This section is devoted to the study of convergence of the recursive algorithm. We will show that h_n defined in (II.9) is closely related to an ordinary differential equation (ODE). The stationary points of ODE are the optimal trailing stop percentage that we are seeking. The details asymptotic analysis can be worked out by virtue of the approach given in [9]; see also [6, Chapters 5 and 8]. Thus we shall be brief and only summarize the results via the following proposition.

To carry out the study of convergence, we define the following: $t_n = \sum_{i=1}^{n-1} \varepsilon_i$, $m(t) = \max\{n : t_n \le t\}$, $h^0(t) = h_n$ for $t \in [t_n, t_{n+1})$, $h^n(t) = h^0(t + t_n)$, $\widetilde{r}^0(t) = \sum_j^{m(t)-1} \varepsilon_j r_j$ and $\widetilde{r}^n(t) = \widetilde{r}^0(t + t_n) - \widetilde{r}^0(t_n)$. Note that $h^0(\cdot)$ is a piecewise constant process and $h^n(\cdot)$ is its shift. To proceed, we use the following conditions.

(A1) The second derivative $J_{hh}(\cdot)$ is continuous.

(A2) For each h belongs to a bounded set, $E|Y_n^{\pm}|^2 < \infty$, and the sequence $\{\eta_j(h, \xi_j)\}$ is a bounded ϕ-mixing sequence with mixing rate $\widetilde{\phi}_k$ such that $\sum_k \widetilde{\phi}_k^{1/2} < \infty$.

The following theorem and its corollary can be proved as in [9]. We thus omit the details.

Theorem 3.1:. Assume (A1)–(A2). Suppose the differential equation

$$\dot{h} = J_h(h) + r(t) \tag{III.11}$$

has a unique solution for each initial condition. Then $(h^n(\cdot), \widetilde{r}^n(\cdot))$ converges weakly to $(h(\cdot), \widetilde{r}(\cdot))$, the solution to (III.11) with $\widetilde{r}(t) = \int_0^t r(s)ds$, and $r(t) = 0$ when $h(t) \in [a, 1]$.

Corollary 3.2:. Suppose that (III.11) has a unique stationary point $h_* \in (a, 1)$ being globally asymptotically stable in the sense of Liapunov, and that $\{s_n\}$ is a sequence of real numbers such that $s_n \to \infty$. Then the weak limit of $h^n(s_n + \cdot)$ is h_*.

Combining the estimates obtained thus far and using the results in [6, Chapter 6], we obtain the w.p.1 convergence of the algorithm. We state the result below.

Theorem 3.3: Under the conditions of Theorem 3.1, $h^n(\cdot)$ converges w.p.1 to $h(\cdot)$ that is the solution of (III.11). Moreover, if (III.11) has a unique stationary point $h_* \in (a, 1)$ being globally asymptotically stable in the sense of Liapunov, and that $\{s_n\}$ is a sequence of real numbers such that $s_n \to \infty$. Then $h^n(s_n + \cdot) \to h_*$ w.p.1.

IV. Interval Estimates

This section is devoted to obtaining interval estimates as well as a piratically useful stopping rule for the recursive computation. Roughly, with prescribed confidence level, we wish to show that with large probability (probability close to 1), a sequence of scaled and centered estimates and a stopped sequence converge weakly to a diffusion process. Based on this result, we will then be able to build confidence interval for the iterates. To proceed, for simplicity of notation, we take $\varepsilon_n = 1/n$ and $\delta_n = \delta_0/n^{1/6}$. In the analysis to follow, for simplicity and without loss of generality, we take $\delta_0 = 1$. To carry out the subsequent study, in addition to the conditions of Theorem 3.1, we also assume an additional condition.

(A3) $J_h(h) = J_{hh}(h_*)(h - h_*) + o(|h - h_*|^2)$, where $J_{hh}(h_*) - (1/2) < 0$. In addition, $k^{2/3}E(h_k - h_*)^2 = O(1)$ and the bound holds uniformly in k.

Define $\rho_n^* = [Y(h_*, \xi_n^+) - Y(h_*, \xi_n^-)] - E_n[Y(h_*, \xi_n^+) - Y(h_*, \xi_n^-)]$. That is, ρ_n^* is ρ_n with the argument h_n replaced by h_*. The detailed development of the interval estimates can be outlined as follows. Suppose that we can show that $n^{1/3}(h_n - h_*)$ is asymptotically normal with mean zero and asymptotic variance σ^2. Choose α, such that $0 < \alpha < 1$ and $1 - \alpha$ is the desired confidence coefficient. Given $\varepsilon > 0$, then the asymptotic normality implies that

$$P\Big(\frac{n^{1/3}|h_n - h_*|}{\sigma} \le z_{\alpha/2}\Big) \to 1 - \alpha \quad \text{as} \quad n \to \infty. \tag{IV.12}$$

This will lead to the desired confidence interval estimator. Then we require the length of the interval $|h_n - h_*|$ be small enough in that for any $\varepsilon > 0$, for sufficiently large n, we can make $\sigma z_{\alpha/2}/n^{1/3} < \varepsilon$ or equivalently $n > \lfloor \sigma z_{\alpha/2}/\varepsilon \rfloor$. Define

$$M^n_{\varepsilon,\alpha} = \lfloor \frac{\sigma z_{\alpha/2}}{\varepsilon} \rfloor,\ \mu_{\varepsilon,\alpha} = \inf\{n : M^n_{\varepsilon,\alpha} \le n\}, \quad \text{(IV.13)}$$

where $\lfloor z \rfloor$ denotes the greatest integer that is less than or equal to z. Then $\mu_{\varepsilon,\alpha}$ is a stopping rule for the iterating sequence $\{h_n\}$. Denote $I_{\mu_{\varepsilon,\alpha}} = \Big[h_{\mu_{\varepsilon,\alpha}} - \sigma \frac{z_{\alpha/2}}{n^{2/3}}, h_{\mu_{\varepsilon,\alpha}} + \sigma \frac{z_{\alpha/2}}{n^{2/3}}\Big]$. We shall show that as the length of the interval shrinks, i.e., $\varepsilon \to 0$, $P\{h \in I_{\mu_{\varepsilon,\alpha}}$ and $|I_{\mu_{\varepsilon,\alpha}}| \le \varepsilon\} \to 1-\alpha$, where $|I_{\mu_{\varepsilon,\alpha}}|$ denotes the length of the interval $I_{\mu_{\varepsilon,\alpha}}$.

Remark 4.1: In view of definition (IV.13), we obtain the following result. Note that as $\varepsilon \to 0$, $M_{\varepsilon,\alpha} \to \infty$ and $\mu_{\varepsilon,\alpha} \to \infty$ w.p.1. Moreover, the definitions of $M_{\varepsilon,\alpha}$ and $\mu_{\varepsilon,\alpha}$ implies that as $\varepsilon \to 0$, $\mu_{\varepsilon,\alpha}/M_{\varepsilon,\alpha} \to 1$ w.p.1. This will be used in what follows.

Since $h_* \in (a, 1)$, h_* is not on the boundary of the projection. Thus we drop the projection part or the reflection term $\varepsilon_n r_n$ in what follows for simplicity. We simply assume that $h_n \in (a, 1)$ for all n large enough. To obtain the desired result, our plan is as follows. We first establish an asymptotic equivalence. Then we define a sequence of interpolated processes, show that the limit of the interpolation is a diffusion process, and further obtain the diffusion limit for a sequence involving the $\mu_{\varepsilon,\alpha}$ defined above. To proceed, let $v_n = n^{1/3}(h_n - h_*)$. We derive the following result.

Lemma 4.2: Assume (A1)–(A3).

$$v_{n+1} = \sum_{k=1}^{n} \frac{1}{2k^{1/2}} A_{n,k}(\rho^*_k + \eta_j(h_*)) + o(1), \quad \text{(IV.14)}$$

$$v_{\mu_{\varepsilon,\alpha}+1} = \sum_{k=m}^{\mu_{\varepsilon,\alpha}} \frac{1}{2k^{1/2}} A_{\mu_{\varepsilon,\alpha},k}(\rho^*_k + \eta_j(h_*)) + o(1), \quad \text{(IV.15)}$$

where $o(1) \to 0$ in probability as $\varepsilon \to 0$, and $A_{jk} = \begin{cases} \prod_{l=k+1}^{j}(I + \frac{J_{hh}(h_*)}{l}), & \text{if } j > k; \\ I, & \text{if } j = k. \end{cases}$

By virtue of the above lemma, instead of working with the discrete expression directly, we shall first examine interpolations of appropriate sequences. Let $\widetilde{W}_n(\cdot)$ and $W_n(\cdot)$ be defined by

$$\begin{aligned} \widetilde{W}_n(t) &= \sum_{k=1}^{\lfloor nt \rfloor} \frac{1}{2k^{1/2}} A_{\lfloor nt \rfloor k}[\rho^*_k + \eta_k(h_*)], \quad \text{for } t \in [0,1], \\ W_n(t) &= \frac{1}{2}\widetilde{W}_n(t), \end{aligned} \quad \text{(IV.16)}$$

where $\lfloor z \rfloor$ denotes the greatest integer which is less than or equal to z. Note that $\widetilde{W}_n(\cdot) \in D[0,1]$ and so is $W_n(\cdot)$, where the $D[0,1]$ is the space of functions that are right continuous and have left hand limits endowed with the Skorohod topology. For definitions and general notion of weak convergence, see [3], [6].

We complete the proof by employing the idea of random change of time. As a result, $W_{\mu_{\varepsilon,\alpha}}(\cdot)$ converges weakly to $W(\cdot)$ is established. Define $B_n(t) = \sum_{j=1}^{\lfloor nt \rfloor} \frac{1}{j^{1/2}}[\rho^*_j + \eta_j(h_*)]$. In view of (IV.16), a summation by parts yields

$$\widetilde{W}_n(t) = B_n(t) + j_{hh}(h_*) \sum_{k=1}^{\lfloor nt \rfloor - 1} \frac{1}{(k+1)} A_{\lfloor nt \rfloor,(k+1)} B_n(\frac{k}{n}). \quad \text{(IV.17)}$$

(A4) $B_n(\cdot)$ converges weakly to $B(\cdot)$, a Brownian motion with covariance $t\sigma^2_B$

Theorem 4.3: Under assumptions (A1)–(A4), $W_n(\cdot)$ converges weakly to $W(\cdot)$, a diffusion process given by $W(t) = \frac{1}{2}\int_0^t \exp\left(-J_{hh}(h_*)(\log u - \log t)\right) dB(u)$, where $B(\cdot)$ is the Brownian motion with covariance $t\sigma^2_B$.

Remark 4.4: Note that Theorem 4.3 allows us to have a handle on the estimation errors. Note that it follows from Theorem 4.3, setting $t = 1$, we have $(n+1)^{1/3}(h_{n+1} - h_*) \sim N(0, \sigma^2)$, where σ^2 is given by $\sigma^2 = E[W(1)]^2 \frac{\sigma^2_B}{8H}$, where $H = \frac{1}{2} - J_{hh}(h_*)$. The asymptotic variance σ^2 together with the scaling factor $n^{1/3}$ provides us with a rate of convergence result. We shall show that the weak convergence result still holds if n is replaced by $\mu_{\varepsilon,\alpha}$.

Theorem 4.5: If the conditions of Theorem 4.3 are satisfied, then

$$W_{\mu_{\varepsilon,\alpha}}(t) = \frac{1}{4} \sum_{k=1}^{[\mu_{\varepsilon,\alpha} t]} \frac{1}{k^{1/2}} A_{[\mu_{\varepsilon,\alpha} t]k}[\rho^*_k + \eta_k(h_*)] \quad \text{(IV.18)}$$

converges weakly to $W(t)$.

Thus, $W_{\mu_{\varepsilon,\alpha}}(1)$ converges in distribution to $N(0, \sigma^2)$ as $\varepsilon \to 0$, and hence, $(\mu_{\varepsilon,\alpha})^{1/3}(h_{\mu_{\varepsilon,\alpha}+1} - h_*)$ converges in distribution to $N(0, \sigma^2)$, as $\varepsilon \to 0$, where σ^2 is the asymptotic variance.

Remark 4.6: In the process of constructing the desired confidence interval, we used a sequence $\{\sigma_n\}$, and we assumed that $\sigma^2_n \to \sigma^2$, the asymptotic variance. Here, we illustrate how a sequence of consistent estimates $\{\sigma^2_n\}$ can be constructed. In view of the form of the asymptotic variance, to obtain a sequence of consistent estimates $\{\sigma^2_n\}$, we need only construct two consistent sequences, one for estimating $J_{hh}(h_*)$, the other one for estimating σ^2_B.

A consistent sequence of estimates for $J_{hh}(h_*)$ can be constructed by means of two-sided finite difference scheme similar to the estimate for $D\widehat{J}(h_n, \xi_n)$. That is, we construct a finite difference estimate of the derivative of $D\widehat{J}(h, \xi)$ with respect to h. Let assumptions (A1)–(A4) be satisfied. Then, a sequence of estimate $\{D_n\}$ can be constructed, and it is a sequence of consistent estimates of $J_{hh}(h_*)$.

In view of the form of σ^2_B, define $\zeta_{n,i} = \frac{1}{n}\sum_{k=1}^{n} Y_k Y_{k+i}$ $i \ge 0$ and $\zeta_n = \zeta_{n,0} + 2\sum_{i=1}^{n} \zeta_{n,i}$. Recall that if a process is ϕ-mixing, then it is ergodic. By this ergodicity, noting the noise involves a martingale difference sequence and a mixing sequence, it can be shown that $\zeta_n \to \sigma^2_B$ as $n \to \infty$. Moreover, the implementation can be made recursive. Finally, let $A_n = \frac{1}{2} - D_n$, with the constructions of D_n and ζ_n, we can define σ^2_n as $\sigma^2_n = \frac{\zeta_n}{8A_n}$.

V. Numerical results

In this section, we report our simulation and numerical experiment results. We first compare our algorithm with the

Monte Carlo simulations. Then we test our algorithm using real market data and compare our results to those using a moving average crossing method, which is well studied in the literature.

A. Simulation study

Due to the absence of an analytic solution to Problem (II.3), we use Monte Carlo method to generate optimal trailing stop percentage h. By comparing the results of stochastic approximation algorithm, we demonstrate that the algorithm constructed indeed provides good approximation results. Assume that the stock price follows a geometric Brownian motion given by $\frac{dS(t)}{S(t)} = \mu dt + \sigma dw(t)$ with $S(0) = S_0$ being the initial price, where μ and σ represent the expected rate of return and volatility, respectively. We generate random samples of $S(t)$ for given values of μ, σ, and S_0. Then we compute the optimal trailing stop percentage h. We take $S_0 = 100$. As shown in Table 1, one can see the optimal values of h increase in σ. For example, when μ is fixed at 10%, h rises from 8.00% to 16.75% as σ increases from 10% to 20%. Intuitively, one should set a higher h to avoid being stopped out (or forced to sell) from a position due to normal price fluctuations when σ is larger. On the other hand, the dependence of h on μ is not obvious. For instance, with a fixed σ at 20%, h varies in the range from 16.00% to 16.75%. These relations are also shown in Figures 2 and 3.

$\sigma \setminus \mu$	10.00	11.00	12.00	13.00	14.00	15.00	16.00
10.00	8.00	7.75	7.50	7.50	7.00	7.00	6.75
15.00	12.25	12.00	12.25	12.00	11.75	11.75	11.50
20.00	16.75	17.25	16.75	16.50	16.00	15.75	16.25
30.00	25.25	25.00	24.75	24.50	24.50	25.25	25.00

Table 1: Optimal trailing stop percentage using Monte Carlo Method for given expected rate of return μ and volatility σ.

We now can draw the graph of objective function $J(h)$ for fixed μ and σ. For example, the graph of $J(h)$ with $\mu = 20\%$ and $\sigma = 35\%$ is shown in Figure 4. Figure 5 is the graph of $J(h)$ with $\mu = 12\%$ and $\sigma = 25\%$. From these two figures, one can see the smoothness of $J(h)$ and hence our assumption of A1 is reasonable.

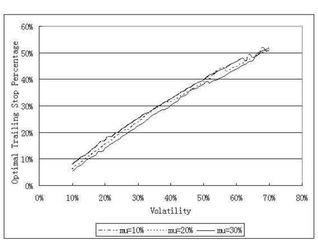

Fig. 2. Optimal trailing stop percentage using Monte Carlo Method against stock price volatility given different expected rate of return μ.

We use the stochastic approximation to compute the optimal values of h. In the following approach, the sequences $\{\varepsilon_n\}$ and $\{\delta_n\}$ are chosen to be $\varepsilon_n = 1/(n + k_0)$ and

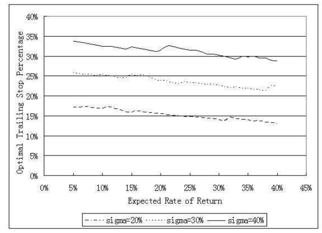

Fig. 3. Optimal trailing stop percentage using Monte Carlo Method against expected rate of return given different stock price volatilities σ.

$\delta_n = 1/(n^{1/6} + k_1)$, respectively, where k_0 and k_1 are some positive integers. We choose $k_0 = 1$, $k_1 = 10$, $\rho = 0.04$, and the lower bound $a = 5\%$. When the trailing stop percentage is set at h, instead of the finite difference approximation of the gradient given in the algorithm, we may take random samples of size n_0 with random noise sequence $\{\xi^{\pm}_{n,l}\}_{l=1}^{n_0}$ such that $\widehat{J}(h, \xi^{\pm}_n) \stackrel{\text{def}}{=} \frac{\widetilde{J}(h,\xi^{\pm}_{n,1})+\cdots+\widetilde{J}(h,\xi^{\pm}_{n,n_0})}{n_0}$. We assume that $E\widehat{J}(h, \xi^{\pm}_n) = J(h)$ for each h. Then for each h, $\widehat{J}(h, \xi^{\pm}_n)$ is an estimate of $J(h)$. In the simulation study, we can use independent random samples to estimate the expected value of $\Phi(S(\tau_n)) \exp(-\rho\tau_n)$. The law of large numbers implies that $\widehat{J}(h, \xi^{\pm}_n)$ converges to $J(h)$ w.p.1 as $n_0 \to \infty$. Recall that n_0 is the number of random samples used in each iteration. The iterates stop whenever $|h_{n+1} - h_n| < 0.0005$. Several different initial guesses are used. We take $n_0 = 1000$. Table 2 shows the results for $\mu = 10\%$ and $\sigma = 20\%$. The optimal trailing stop percentage calculated by Monte Carlo method is MC $= 16.75\%$. In Table 2, SA1 is the optimal trailing stop percentage calculated by stochastic approximation with averages of samples.

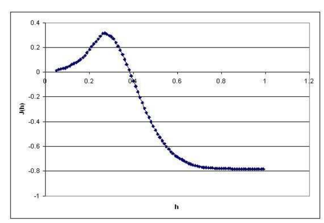

Fig. 4. The objective function $J(h)$ against the trailing stop percentage h for fixed expected rate of return and volatility at $\mu = 20\%$ and $\sigma = 35\%$.

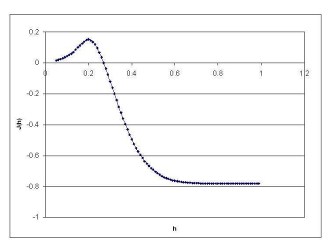

Fig. 5. The objective function $J(h)$ against the trailing stop percentage h for fixed expected rate of return and volatility at $\mu = 12\%$ and $\sigma = 25\%$.

It can be seen from Table 2 that the estimates are insensitive to the initial guesses, the algorithm leads to accurate estimation of the optimal value. Indeed, for $\sigma \in [10\%, 70\%]$

Initial guess	5.00%	17.00%	28.00%	40.00%	50.00%
SA1	16.82%	16.95%	16.94%	16.94%	16.94%
|MC-SA1|	0.07%	0.20%	0.19%	0.19%	0.19%

Table 2: Estimates using stochastic approximation with averages of samples (SA1) for fixed expected rate of return and volatility at $\mu = 10\%$ and $\sigma = 20\%$, where MC is the optimal trailing stop percentage calculated by Monte Carlo Method.

and $\mu \in [5\%, 40\%]$, the average value of $|MC - SA1|$ is only 1.34%.

Next we use the similar method without averages of samples. In this case, $n_0 = 1$, all other parameters remain unchanged. The results are shown in Table 3, where $SA2$ denotes the optimal trailing stop percentage calculated by stochastic approximation without averages of samples. Again, the estimates are insensitive to the initial guesses. However, the bias $|$MC-SA2$|$ is larger. For $\sigma \in [10\%, 70\%]$ and $\mu \in [5\%, 40\%]$, the average value of $|$MC-SA2$|$ is 3.10%.

Initial guess	5.00%	17.00%	28.00%	40.00%	50.00%
SA2	18.02%	18.02%	18.02%	18.02%	18.02%
|MC-SA2|	1.27%	1.27%	1.27%	1.27%	1.27%

Table 3: Estimates using stochastic approximation without averages of samples (SA2) for fixed expected rate of return and volatility at $\mu = 10\%$ and $\sigma = 20\%$, where MC is the optimal trailing stop percentage calculated by Monte Carlo Method.

Compared to the Monte Carlo method, the SA methods take much less time to calculate the optimal trailing stop percentage. We run this algorithm on a Sun Fire 880 serve with 8GB memory, generally, it takes about 30 to 60 seconds to obtain the estimated optimal trailing stop percentage. With the Monte Carlo method, it takes at least 20 minutes for the corresponding computation.

B. Using real market data

In what follows, we use the SA to compute the trailing stops. We use NASDAQ-100 components during the period from January 2,1995 - December 31, 2001. Here we consider two trading strategies. Since we need the 50-day moving averages of prices, we start our trading strategies on the fiftieth trading day after January 2, 1995.

Strategy 1. If the stock price on the fiftieth trading day after January 2, 1995 is below the 50-day moving average, buy the stock. Otherwise, buy the stock when price is up-crossing 50-day moving averages. And sell stock when price is down-crossing 50-day moving averages. If the latter doesn't happen, then sell the stock on the last day of the period, December 31, 2001.

Strategy 2. The entry point is exactly the same as described in Strategy 1. Then use trailing stop technique with the percentage calculated via the stochastic approximation method. If price doesn't hit the stop price before December 31, 2001, sell stock on that day.

For example, let us assume we started collecting stock prices for Cadence Design Systems Inc (CDNS) on January 2, 1995. Then March 15, 1995 is the first day we have the 50-day moving average. It happens the closing price on that day is greater than the 50-day moving average, therefore we buy the CDNS for the price \$5.75. On June 2, 1995, the closing price of CDNS is \$7.33, which is less than the 50-day moving average. Therefore, Strategy 1 suggests to sell CDNS at \$7.33, resulting a raw return of 27.48%. However, using the trailing stop technique, Strategy 2 suggest to hold till July 10, 1996. The closing price on that day is \$15.69, resulting a raw return of 172.88%. The daily close prices, their 50 day moving average, and the trailing stop curve are plotted in Figure 6.

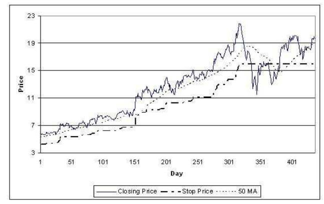

Fig. 6. The prices of Cadence Design Systems Inc from March 15, 1995, to December 2, 1996

We perform the same experiments for NASDAQ-100 components if prices are available. Table 3 reports the average rate of returns from Strategies 1 and 2. The average rate of return from Strategy 1 is 11.45% while the The average rate of return from Strategy 2 is 71.45%. It is easy to see that the Strategy 2 outperforms the Strategy 1 on average.

Moving average	Trailing stop
11.45%	71.45%

Table 4: Average Rate of Returns from Different Trading Strategies

References

[1] P. Billingsley, *Convergence of Probability Measures*, J. Wiley, New York, 1968.
[2] V.S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*, Hindustan Publishing Agency and Cambridge University Press, Dehli, 2008.
[3] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley, New York, 1986.
[4] P.W. Glynn and D.L. Iglehart, Trading Securities Using Trailing Stops, *Management Science*, **41**, (1995), 1096-1106.
[5] H.J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.
[6] H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd Ed., Springer, New York, 2003.
[7] P. L'Ecuyer and G. Yin, Budget-dependent convergence rate of stochastic approximation, *SIAM J. Optim.*, **8** (1998), 217-247.
[8] G. Yin, A stopping rule for the Robbins-Monro method, *J. Optim. Theory Appl.*, **67** (1990), 151-173.
[9] G. Yin, R.H. Liu and Q. Zhang, Recursive algorithms for stock liquidation: A stochastic optimization approach, *SIAM J.*, **13**, (2002), 240-263.
[10] G. Yin, Q. Zhang, F. Liu, R.H. Liu, and Y. Cheng, Stock liquidation via stochastic approximation using NASDAQ daily and intra-day data, *Math. Finance*, **16** (2006), 217-236.
[11] Q. Zhang, Stock trading: An optimal selling rule, *SIAM J. Control. Optim.*, **40**, (2001), 64-87.