Stock trading via feedback control: an extremum seeking approach

Simone Formentin, Fabio Previdi, Gabriele Maroni and Claudio Cantaro

Abstract— **In finance, the goal of technical analysis is to model stock dynamics to make reliable predictions of future prices. In recent years, Prof. B. Barmish and coauthors have proposed a paradigm shift in which stock trading is reformulated as a control design problem. From such a perspective, unpredictable price variations can be seen as external disturbances and do not need to be accurately modeled for the investment policy to be properly selected. Although very powerful, the tuning of the reactive scheme to stock trading is all but straightforward. In this paper, an Extremum Seeking approach is proposed to directly design the control action from data and make the overall strategy adaptive with respect to trend variations of the stock price. The method is extensively backtested on real stock data.**

I. INTRODUCTION

In November 1882, Charles Dow, Charles Bergstresser and Edward Jones founded *Dow Jones & Company*, nowadays one of the most important publishing firm in financial information[1]. During those years, Charles Dow published several analyses and articles on his own newspaper, the Wall Street Journal, aimed to a better comprehension of the financial markets and related investment strategies.

According to [13], [9], the contribution made by Charles Dow and his company may be considered the dawn of all those theories that, using specific information and historical prices, try to gain excess returns over the market average. These theories constitute the so-called *Technical Analysis*.

The main idea underlying all these methods is the use of past information (price, volume and open interest) to make reliable predictions on the equities' future prices. During the decades, several approaches have been developed based on, *e.g.*, charts, time series analyses (used also in econometrics and system identification), patterns recognition, volume and open interest indicators, and many others (see [9] for a review).

A common feature of all the existing approaches is that the investment policy is based on a model of the past stock price dynamics. However, describing the market behavior is never a trivial task and modeling errors may in turn lead to detrimental performance.

An innovative approach to stock trading was presented by B. Ross Barmish in [3] and follow-up works (e.g., [11], [12]). In these papers, it has been shown that gaining money against unpredictable price variations can be reformulated as a *control design problem* with a disturbance rejection goal. Within such a framework, the disturbance, namely the price variations, does not need to be modeled, thus the approach can be referred to as *model-free* [5], [6].

Notwithstanding the basic idea is as clever as simple, the tuning of the controller defining the best investment level for a single stock is all but a straightforward task. In [4], a two degrees of freedom controller named *Simultaneous Long-Short (SLS)* controller is implemented to combine both a long and a short strategy. In more detail, one control block is tuned to implement a long position, that is to perform well in all scenarios with a rising price (also known as *bull moments*), where the trader will profit from first buying and then selling the stock. The other control block is aimed to maximize the return when the price decreases (during the so-called *bear moments*), thus implementing a short investment position, where the trader profits by first selling and then buying the stock (uncovered sell). In [5], proportional controllers are used. The proportional gain selection is really a critical task and it is performed by means of simulations using past stock prices. The resulting control law is simple, but does not adapt to market changes and may provide bad worst-case performance as evidenced in [5]. In [12] the same authors extend their work introducing the *Initially Long-Short (ILS)* that consider a Proportional-Integral (PI) controller for the investment function, nevertheless the time-invariant nature of the controller might not adapt well to inversions of price trend.

In this paper, a different approach for control design in feedback stock trading is proposed, based on an *Extremum Seeking* rationale [1]. Such an approach appears to be very suitable for the problem at hand, for the following reasons: (i) its aim is to maximize the output of a system whose dynamics is unknown, like the excess return; (ii) it is intrinsically model-free; (iii) it provides a time-varying feedback gain, so it may adapt to market time-varying conditions; (iv) unlike many other adaptive methods, it is theoretically guaranteed to converge to a local optimum [10]. At the end of the paper, a real case study and extensive experiments on a significant number of stocks will show that the proposed approach may largely outperform the standard *Buy & Hold* strategy as well as the feedback scheme in [5] with a time-invariant SLS controller. Backtest overfitting [2] is avoided tuning the parameters of the methods on in-sample price data and testing them on out-of sample price data from the same stocks but coming from a different period of time, preventing ambiguity with respect to generalization properties.

The remainder of the paper is as follows. In Section II, the mathematical details of the feedback control approach

S. Formentin and C. Cantaro are with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, via G. Ponzio 34/5, 20133 Milano, Italy. Email to: simone.formentin@polimi.it. F. Previdi and G. Maroni are with the Department of Management,Information and Production Engineering, University of Bergamo, via Marconi 5, 24044 Dalmine (BG), Italy. Email to: fabio.previdi@unibg.it

[1]For more information, see: http://www.dj.com/history.asp.

to stock trading is presented and its current limitations are discussed. Section III illustrates how the Extremum Seeking rationale can be applied to the problem at hand, while Section IV shows experimental achievements. The paper is ended by some concluding remarks.

II. STOCK TRADING VIA FEEDBACK CONTROL

The whole analysis will be carried out under the hypothesis of *idealized frictionless market*. In this ideal setting performance certification theorems can be proved. This type of market is based on the following five assumptions (the same hypotheses of [5]):

A1. *Continuos trading*. The trader can continuously vary the investment level. In modern high frquency trading, it is indeed possible to run many trades per second, thus this assumption can be rated as mild.

A2. *Trader is price-taker and there is perfect liquidity*. The trader can transact as many shares as desired, without any gap between bid and ask prices. This situation is compatible with small and medium traders whereas the case of hedge funds that can directly influence the stock price through buying or selling large amounts of shares might violate it.

A3. *Zero interest and margins rates*. The risk-free rate of return $r \geq 0$ at which the trader's idle cash increases is considered equal to zero. Likewise, the margin interest owed to the broker is considered equal to zero. This assumption can be dropped modeling the impact of interests and margins on the account value and will be subject of study in later works.

A4. *No Transaction costs*. The commission structure for frequent traders with tens of thousands dollars of orders is nowadays adequately low and, hence, not a relevant issue for the purpose of this work[2].

A5. *One stock portfolio*. This is not a real constraint since the method can be applied to an entire portfolio of shares. Since portfolio optimization is out of the scopes of this work, the method will be applied to single stocks.

The approach introduced in [3] proposes to *react* to price variations instead of modeling them. The main philosophy is then the closed-loop strategy illustrated in Figure 1, representing the classic long position trading and which will be detailed next.

In the block denoted as *Gain/Loss Accounting*, the relationship between the cumulative gain $g(t)$, the price dynamics $p(t)$ and the investment level $I(t)$ is considered. In particular, such a dynamics can be deterministically described, if smooth prices are considered (the strategy will be valid also with non-smooth prices, but the method would be less easy to present). The mathematical model is as follows.

The incremental gain dg at time t depends on the incremental return dp (normalized with respect to the stock price p) and the investment level I as

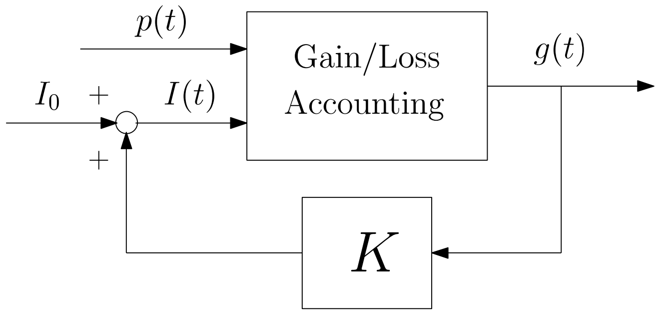

Fig. 1. Stock trading via feedback.

$$dg(t) = \frac{dp(t)}{p(t)} I(t), \tag{1}$$

where, e.g. for positive investments, dg is positive only if p increases, whereas it represents a loss if p decreases. Also, it can be set $\rho(t) = \frac{dp(t)}{p(t)}$ as the return function.

The investment level is modified of a quantity dI depending on a feedback measure of the derivative of the gain as

$$\frac{dI(t)}{dt} = K \frac{dg(t)}{dt} \tag{2}$$

where K is a suitable positive parameter working as a one degree of freedom proportional controller. Notice that, from (2), it follows that

$$I(t) = I_0 + K g(t), \tag{3}$$

where $I_0 = I(0)$ denotes the initial investment.

The above *long strategy* works well in bull scenarios, i.e., when the price rises. During market declines such an approach is clearly inadequate (see again (3)). In these so called bear scenarios, the above scheme should be modified setting the controller gain K and the initial investment I_0 to negative values, where a negative investment models a short trade, which is the practice of selling shares that are not directly owned but must be borrowed from a bank or a financial intermediary (broker). With this setting, a decrease in the price would produce an increasing of the gain $g(t)$ and the controll feedback will force an incrasing of the short position. This attidute towards the market is called *short strategy*. It is clear that, with this scheme, long and short positions can not coexist.

To overcome this issue, a controller able to apply both a long position, when the stock price increases, and a short position, when the price decreases, is needed. Therefore, in [4], the *Simultaneous Long-Short (SLS) Control* approach of Figure 2 has been introduced. In this scheme, two linear closed-loop schemes in parallel are set, one with $I_0 > 0$ and $K > 0$ and the second with $I_0 < 0$ and $K < 0$. The resulting overall strategy turns out to be a simple linear combination of the two control actions.

The above scheme shows interesting theoretical properties. First of all, the following Theorem [5] can be proven:

Theorem 1 (Arbitrage Theorem): In an idealized frictionless market with continuously differentiable price functions

[2]For further information about the effect of transaction costs on the economy, see [8].

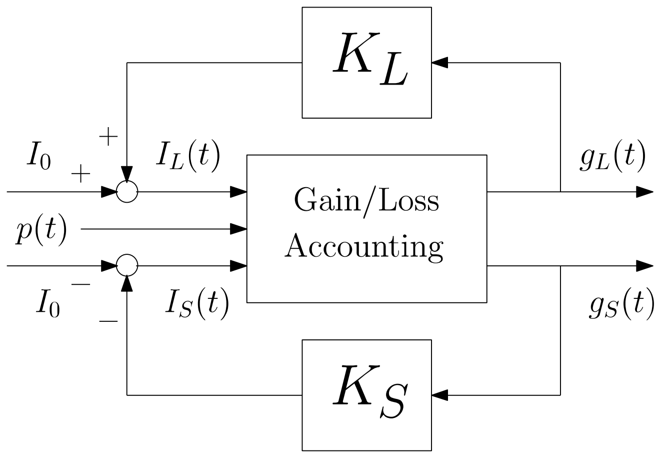

Fig. 2. Simultaneous Short-Long Controller: trading in both bull and bear markets.

in $[0, t]$, the trading gain/loss function resulting from the SLS controller is

$$g(t) = \frac{I_0}{K}\left[\left(\frac{p(t)}{p(0)}\right)^K + \left(\frac{p(t)}{p(0)}\right)^{-K} - 2\right], t \geq 0. \quad (4)$$

Moreover, except for the trivial break-even case with price $p(t) = p(0)$, it holds that

$$g(t) > 0. \quad (5)$$

Proof: See [5]. ■

Finally, in order to work within a realistic framework for trading, two additional boundary conditions will be also taken into account, namely a collateral requirements condition (also known as *leverage*) and a Value at Risk (VaR) condition.

The first condition is needed to ensure that the account value $V(t)$ has enough cash balance[3]. Such a condition can be formalized by introducing the upper bound $|I(t)| \leq \lambda V(t)$. Through this *leverage* condition, a guarantee for the brokers is introduced, that typically permit trades if $|I(t)| \leq 2V(t)$ is satisfied [5].

The second condition is needed to limit the maximum possible investment and can be formalized by saturating the controller action, namely by imposing $I(t) \leq I_{max}$.

One major limitation of the SLS scheme is the use of the time-invariant feedback controller K [5], as natural continuation of the work the problem of the development of an algorithm which dynamically adjust the value of K is addressed in this paper.

III. THE EXTREMUM SEEKING APPROACH

The Extremum Seeking approach is a feedback control methodology to achieve the maximum (or the minium) of the measured output, without needing the knowledge of the process under control [1]. The key idea behind the approach is that of estimating the gradient of the output with respect to the input parameters by perturbating *run-time* such parameters.

[3]Under our zero interest rate assumption, the expression for the account value is simply $V(t) = I_0 + g(t)$.

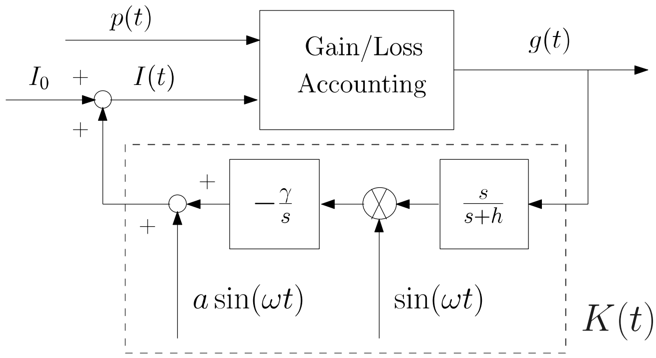

Fig. 3. The basic Extremum Seeking scheme for feedback stock trading.

The methodology is naturally suited for all systems, like the one considered here, in which the dynamics (of the price) is unknown or highly uncertain. Moreover, it allows the design parameters, i.e., the controller gain K in this case, to be time-varying and self-adapted on-line. A proof of convergence to the maximum (or minimum) output for a common implementation of the Extremum Seeking scheme is also provided in [10].

Figure 3 depicts the basic Extremum Seeking scheme, the following is a description of the functioning:

- the *dithering signal* $asin(\omega t)$, a small amplitude sinusoidal signal aimed at perturbing the control input, namely the investment level. This is useful to find out the direction of improvement of the performance;
- the *unknown plant* generates the output to maximize. In the present setting, it is represented by the gain/loss accounting block, where the inputs are the investments I and the price p, while the output is g;
- the output of the plant, the gain function, is high pass filtered by the *washout filter* described by the transfer function $s/(s + h)$ to reject his continuous component [7];
- the information about the gradient of the gain with respect to the investment level is obtained multypling the resulting filtered signal with the perturbation;
- the resulting signal is then integrated by the *integrator* γ/s which updates the control action in the direction of driving the gradient to zero. For $\gamma > 0$ the output is minimized, whereas it is maximized for $\gamma < 0$.

The stability proof in [10] shows that the solution will be within a neighborhood of the maximum (or minimum) of the output function but it will never reach it, due to the persistent perturbation signal $asin(\omega t)$. To get an estimation as close as possible to the optimum value, it will be needed to leverage on the parameters involved in the algorithm.

Notice that the resulting time-varying feedback gain $K(t)$ generated by composing the blocks in the blue box of Figure 3 must also here be composed by two elements: $K_S(t)$ for the short investment strategy and $K_L(t)$ for the long investment strategy. Therefore, the Extremum Seeking scheme is here applied to both the investment strategies and composed similarly to the SLS case in Figure 2 (the overall control gain will again be obtained as the sum of the two actions as $K(t) = K_L(t) + K_S(t)$). This new scheme proposed in this paper will be called Extremum Seeking Controller (ESC).

Notice that a, ω, γ and h are tuning knobs and need to be determined. In particular, γ must have positive values when applied for the long controller and negative values for the short one.

To suitably select such parameters, [1] suggests some qualitative guidelines:

- ω: perturbation frequency, it must be within the interval $[0, \pi]$ and it can be chosen depending on the closed loop system bandwidth;
- a: perturbation amplitude, it must be small enough to obtain small changes in the output function but large enough to assure reliable measure of the gradient of g;
- h: High pass filter, it must be designed such that $0 < h < 1$ and should be smaller than ω so that the filter removes the continuous component of the output without corrupting the estimation of the gradient.
- γ: the gain of the Extremum Seeking scheme; large values will speed up the convergence rate as well as the possibility of saturation conditions. For this reason, we should select γ large enough to obtain satisfying results under both ideal conditions and on real data, but also small enough to avoid steady state situations, namely values of gain K constant over the time [1].

Simply put, the tuning of the above knobs could be summarized into a managing problem of the trade-off between convergence rate and accuracy of the optimizer. In other words, a rapid convergence will imply that the investment will spend a shorter amount of time very close to the optimal investment strategy; on the other hand, a better tracking of the optimal trading policy can be achieved at the cost of a slower rate of adaptation when the price dynamics change.

It should be here also remarked that other dithers could be used instead of the sine wave as the perturbation signal. For instance, the square wave excitation is proved in [14] to provide the best convergence rate among all dithers of the same amplitude and frequency. Therefore, this signal will be used in the experimental examples of the next section.

Finally, notice that the higher level of flexibility of such a time-varying control strategy is paid in terms of theoretical guarantees like Theorem 1, which no longer apply. Nonetheless, convergence analysis of extremum seeking schemes is indeed feasible (see again [14]) and will be object of future works. Moreover, as illustrated in the following example, there are already some cases in which the assumptions of the theory with time-invariant controllers are not satisfied, and in such cases the extremum seeking control appears as the most suited solution.

IV. EXPERIMENTAL RESULTS

In this section, a numerical case study considering experimental data taken from a real stock is presented. Such an example has been selected within a round-trip period, where the traditional strategy is known to encounter some difficulties. After the case study, a more comprehensive simulation campaign, taking into account all DJIA's (Dow Jones Industrial Average) stocks between 2013 and 2015, is discussed, with the aim of providing a statistical assessment of the method.

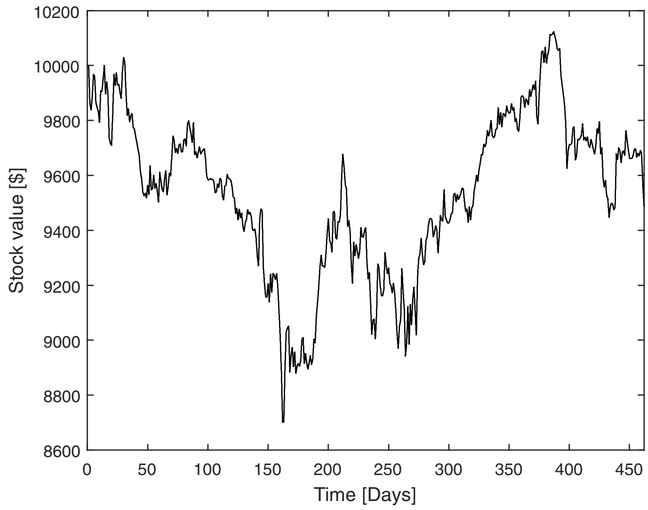

Fig. 4. Daily closing prices for the Exxon (XOM, NYSE) stock from January 2015 until October 2016.

A. The Exxon Mobil case study

Consider the Exxon Mobil Corporation (XOM, NYSE) stock from January 2015 until October 2016 in Figure 4. During this period, the equity presents different trends (both positive and negative), and, at the end of the period, the price is very close to the initial value.

Two reactive control strategies are here applied: the SLS control of [5] and the approach proposed in this paper. For a fair comparison, the same values for the initial account $V_0 = 10000\$$, the initial investment $I_0 = 5000\$$ and the maximum investment level $I_{max} = 20000\$$ are considered for both the techniques. Obviously, for the short-selling scheme, the initial and the limit investments must be considered with the negative sign.

For the SLS control, $K_{SLS} = 4$ is selected via backtesting using as in-sample data the previous year of prices of the same stock. For the Extremum Seeking controller, the following parameters are instead chosen with grid-search, by means of the same backtesting strategy: $\alpha = 0.04$, $\omega = 0.4\pi\ rad/sample$, $\gamma = 4$ and $h = 0.8$.

When the SLS controller is applied, the account value in Figure 5 is obtained. As expected, the controller increases the account in presence of either positive or negative trends. However, the overall return in such round-trip conditions is near to zero. In fact, although the SLS controller seems to properly react to the different bear moments between the $150th$ and $275th$ day, the account value stays around zero from day 275 until the end. This is due to the fact that

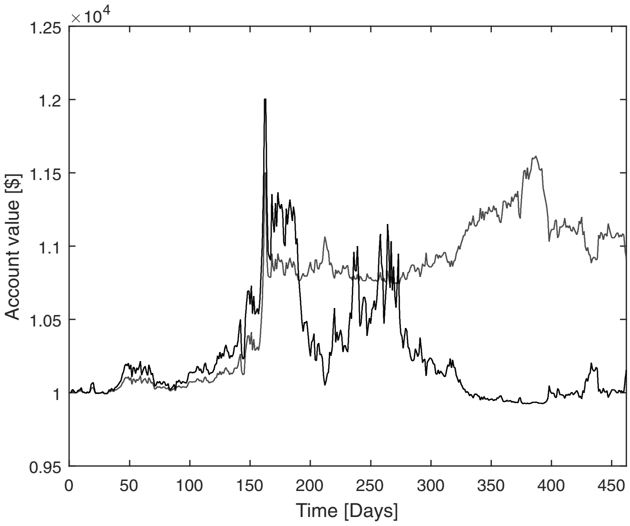

Fig. 5. Account values for the Exxon (XOM, NYSE) example: SLS control (black) and ESC control (red).

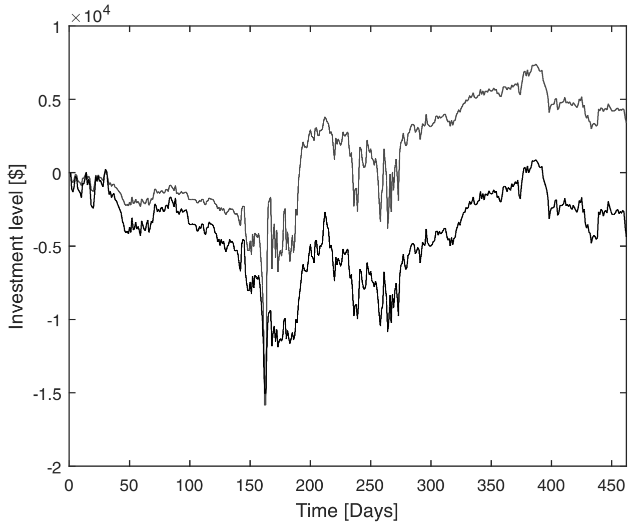

Fig. 6. Investment levels for the Exxon (XOM, NYSE) example: SLS control (black) and ESC control (red).

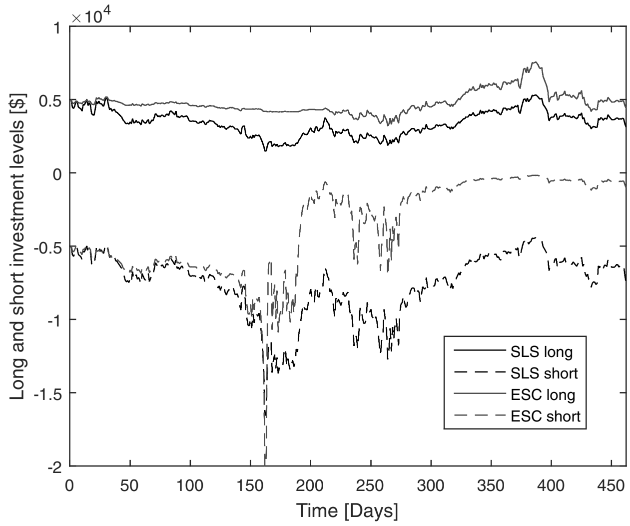

Fig. 7. Split of long/short investment levels for the Exxon (XOM, NYSE) example.

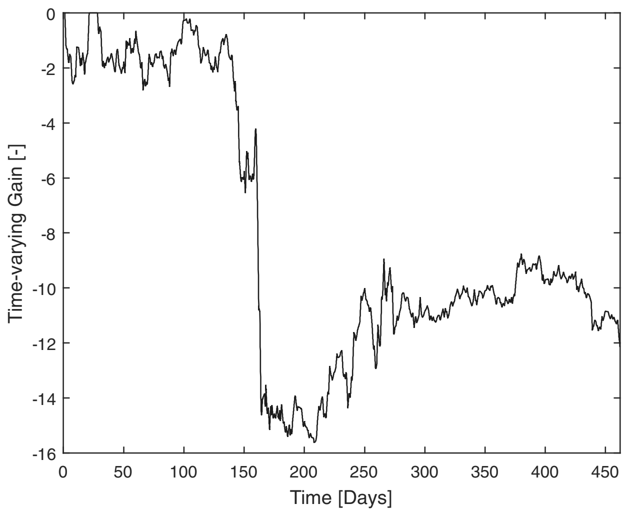

Fig. 8. Time variation of the overall feedback gain K of the ESC strategy for the Exxon (XOM, NYSE) example.

the SLS controller cannot rapidly change the short strategy, launched between day 150 and 275, into a long one. Ideally, the sign of the investment should change (from negative to positive) around the $200th$ day.

In Figure 6, the investment levels corresponding to the accounts of Figure 5 are illustrated (a split in long and short investment is instead shown in Figure 7).

Such a plot confirms that the ESC approach addresses more promptly the change of the trend, basically stabilizing the account notwithstanding the unexpected behavior. This property of the ESC scheme can be appreciated by analyzing the behavior of the overall feedback gain ($K(t) = K_L(t) + K_S(t)$) in Figure 8.

B. The DJIA's stocks between 2013 and 2015

The case study in the previous section is definitely of interest to show that the proposed approach *might* work well in some situations, but indeed is not general. Since when K is time-varying, the theoretical properties proven in [5] do not hold even within idealized markets, in this subsection a statistical evaluation of the approach will be proposed over all the DJIA's stocks in the period 2013-2015.

To do that, ESC and SLS schemes have been designed for each stock following the same approach adopted for the Exxon (XOM, NYSE) example. The final results of the ESC scheme are reported in Table I, in terms of average values of the final gain as compared to the SLS approach and a traditional Buy & Hold, but also considering the maximum and the minimum relative gain with respect to the other

	B&H	SLS
Mean	+9.26%	+3.76%
Upper limit	+63.17%	+49.78%
Lower limit	-4.05%	-2.66%

TABLE I

ACCOUNT VALUES OBTAINED BY THE EXTREMUM SEEKING APPROACH (ESC) IN COMPARISON TO TRADITIONAL SLS AND BUY AND HOLD.

strategies.

Interestingly, the simulations show that not only the ESC strategy provides better results in terms of mean gain, but it is never (significantly) lower than the non-adaptive strategy [5]. At the same time, it may outperform it - in the lucky case - of about $+50\%$. The performance improvement is even more significant as far as the traditional Buy & Hold approach is employed.

V. CONCLUSIONS

In this paper, the problem of controller tuning in reactive stock trading schemes has been tackled. In particular, an Extremum Seeking strategy has been proposed to obtain an adaptive controller, which is able to adjust the feedback mechanism both to increase the gain and follow unpredictable price trends.

The method has been tested on a real case study where traditional feedback trading does not work properly, namely throughout a round-trip period. Moreover, the scheme has been statistically assessed on a thorough simulation campaign over all the DJIA's stocks in the period 2013-2015.

Moreover, a restart strategy has been conceived and introduced in the SLS rationale. This avoids that the investement levels go to zero, so switching-off any investment, when in the case of round-trip price values. The effectiveness of this approach has been proven on a very well known example of round-trip price stock.

Future research will be devoted to the theoretical investigation of ESC-based stock trading, starting with idealized markets.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the valuable comments by Prof. B. Barmish.

REFERENCES

[1] Kartik B. Ariyur and Miroslav Krstic. *Real-time optimization by extremum-seeking control*. John Wiley & Sons, sep 2003.

[2] David Bailey, Jonathan Borwein, Marcos Lopez de Prado, and Qiji Zhu. Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance. 2014.

[3] B. Ross Barmish. On trading of equities: a robust control paradigm. *IFAC Proceedings Volumes*, 41(2):1621–1626, 2008.

[4] B. Ross Barmish. On performance limits of feedback control-based stock trading strategies. *Proceedings of the 2011 American Control Conference*, pages 3874–3879, 2011.

[5] B. Ross Barmish and James A. Primbs. On a new paradigm for stock trading via a model-free feedback controller. *IEEE Transactions on Automatic Control*, 61(3):662–676, 2016.

[6] Michael Heinrik Baumann. On stock trading via feedback control when underlying stock returns are discontinuous. *IEEE Transactions on Automatic Control*, 62(6):2987–2992, 2017.

[7] John H. Blakelock. *Automatic control of aircraft and missiles*. John Wiley & Sons, 1991.

[8] Ronald H Coase. The nature of the firm. *Economica*, 4(16):386–405, 1937.

[9] Robert D. Edwards, John Magee, and W.H. Charles Bassetti. *Technical analysis of stock trends*. CRC Press, 10 edition, 2012.

[10] Miroslav Krstić and Hsin-Hsiung Wang. Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, 36(4):595–601, 2000.

[11] Shirzad Malekpour and B Ross Barmish. On stock trading using a controller with delay: The robust positive expectation property. In *IEEE 55th Conference on Decision and Control (CDC)*, pages 2881–2887. IEEE, 2016.

[12] Shirzad Malekpour, James A Primbs, and B Ross Barmish. On stock trading using a pi controller in an idealized market: The robust positive expectation property. In *IEEE 52nd Annual Conference on Decision and Control (CDC)*, pages 1210–1216. IEEE, 2013.

[13] John J. Murphy. *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin, 2 edition, 1999.

[14] D. Nešić. Extremum seeking control: convergence analysis. *European Journal of Control*, 15(3):331–347, 2009.