Market Dynamics: On Directional Information Derived From (Time, Execution Price, Shares Traded) Transaction Sequences.

Vladislav Gennadievich Malyshkin*

Ioffe Institute, Politekhnicheskaya 26, St Petersburg, 194021, Russia

(Dated: February, 15, 2018)

$Id: DirectionalInformationFromTimePriceSharesTraded.tex,v 1.544 2024/05/13 07:07:43 mal E

A new approach to obtaining market–directional information, based on a non–stationary solution to the dynamic equation “future price tends to the value that maximizes the number of shares traded per unit time” [1] is presented. In our previous work[2], we established that it is the share execution flow ($I = dV/dt$) and not the share trading volume (V) that is the driving force of the market, and that asset prices are much more sensitive to the execution flow I (the dynamic impact) than to the traded volume V (the regular impact). In this paper, an important advancement is achieved: we define the “scalp–price” $\mathcal{P}$ as the sum of only those price moves that are relevant to market dynamics; the criterion of relevance is a high I. Thus, only “follow the market” (and not “little bounce”) events are included in $\mathcal{P}$. Changes in the scalp–price defined this way indicate a market trend change — not a bear market rally or a bull market sell–off; the approach can be further extended to non–local price change. The software calculating the scalp–price given market observations triples (time, execution price, shares traded) is available from the authors.

* malyshki@ton.ioffe.ru

I. INTRODUCTION

Introduced in [3], the ultimate market dynamics problem — finding evidence of existence (or proof of non–existence) of an automated trading machine consistently making positive P&L as a result of trading on a free market as an autonomous agent — can be formulated in its weak and strong forms:

- Weak form: Whether such an automated trading machine can exist at all using only legally available data. (It can definitely exist in an illegal form – e.g. when a brokerage uses client order flow information to frontrun their own clients. This type of strategies typically rely on using proprietary information about clients' Supply–Demand future disbalance and on the subsequent monetization of this information.)

- Strong form: Whether such an automated trading machine can exist and be based solely on transaction sequences – say, the historical time series of (time, execution price, shares traded) market observations triples. This information has supply and demand matched for every observation: at time t trader A sold v shares of some security at price P to trader B and received $v \cdot P$ dollars. Such a strategy can utilize only information about volume and execution flows.

We have shown in [1, 2] that it is share execution flow $I = dV/dt$, not share trading volume V, that is the driving force of the market (see the Figs. 2 and 3 of Ref. [2]: the asset price shows singularity at a high I, but there is no price singularity at the maximal volume price, the median of price–volume distribution).

In [1, 4], the concept of liquidity deficit trading was introduced: *open a position at low I, then close already opened position at high I*; this is the only strategy that avoids catastrophic P&L losses. This strategy is ideologically similar to a classic volatility trading strategy: *buy a straddle at low volatility, sell it at high volatility, never go short volatility to avoid catastrophic P&L loss*, but is different from it by incorporating asset price directional contribution: the decision is needed on whether to open a long or a short position at low I. In [3], the first attempt at finding a non–stationary solution to the dynamic equation by linking asset price and liquidity deficit via "impact–from–the–future" operator (adding to execution flow a contribution from not–yet–executed trades) was presented. In this paper, a different approach is developed.

Instead of adding not–yet–executed trades (impact–from–the–future), we now consider removing from consideration already executed trades (impact–from–the–past) corresponding to *high* I → *low* I transitions. A liquidity deficit trading strategy assumes that only *low* I → *high* I transitions will be captured by the trader. The *high* I → *low* I transitions are not to be used, as they are a major source of catastrophic risk. A typical market behavior after a liquidity excess (high I) event is to "bounce a little," then go in the original direction of the market. This creates an uncertainty of strategy. What does one bet on: "little bounce" or "follow the market"? In contrast, after a liquidity deficit (low I) event, the market can *only* go in the direction of the market trend, eliminating this uncertainty. This shows the importance of the assymetry of dynamic impact (price sensitivity to I [2]): *low* I → *high* I and *high* I → *low* I transitions are to be considered **separately**, as they lead to very different price behaviors. This asymmetry is the topic of this study. The scalp–function (34) is introduced to comprise only those price moves relevant to market dynamics (high I), which allows constructing scalp–price $\mathcal{P}$ (Fig. 2) containing only "follow the market" (and not "little bounce") events. A change in the scalp–price indicates a market trend change, not a bear market rally or a bull market sell–off.

II. BASIC MATHEMATICS

The key concept of the dynamic equation "future price tends to the value that maximizes the number of shares traded per unit time" [1, 3] is to find an averaging weight from the behavior of a market dynamics operator f (e.g. dV/dt, V/t, or dI/dt), then to estimate some directional indicator (e.g. price change, signed volume, etc.) using the obtained weight. Mathematically, the weight is considered in the form of an average depending on wavefunction $\psi(x) = \sum_{k=0}^{n-1} \alpha_k Q_k(x)$: $\psi^2(x(t))\omega(t)dt$, an important generalization of commonly–used parameter–independent fixed time scale averaging such as the exponential moving average : $\omega(t)dt$. The bases $Q_m(x(t))\omega(t)dt$ we use in this paper are listed in Section II of Ref. [3]). Here $\omega(t)$ is decaying exponent and $x(t)$ is either linear or exponential function on time:

$$\omega(t) = \exp\left(-(t_{now} - t)/\tau\right) \tag{1}$$

$$x(t) = \begin{cases} (t - t_{now})/\tau & \text{Laguerre basis} \\ \exp\left(-(t_{now} - t)/\tau\right) & \text{shifted Legendre basis} \end{cases} \tag{2}$$

The problem is then reduced to a generalized eigenvalue problem of operator $\|f\|$:

$$\left| f \middle| \psi_f^{[i]} \right\rangle = \lambda_f^{[i]} \left| \psi_f^{[i]} \right\rangle \tag{3}$$

$$\sum_{k=0}^{n-1} \langle Q_j \,|\, f \,|\, Q_k \rangle \, \alpha_k^{[i]} = \lambda_f^{[i]} \sum_{k=0}^{n-1} \langle Q_j \,|\, Q_k \rangle \, \alpha_k^{[i]} \tag{4}$$

$$\psi_f^{[i]}(x) = \sum_{k=0}^{n-1} \alpha_k^{[i]} Q_k(x) \tag{5}$$

The most general form of the averaging weight is a density matrix:

$$\|\rho\| = \sum_{i=0}^{n-1} \left|\psi_\rho^{[i]}\right\rangle \lambda_\rho^{[i]} \left\langle\psi_\rho^{[i]}\right| \tag{6}$$

$$f_\rho = \operatorname{Spur} \|f|\rho\| = \sum_{i=0}^{n-1} \left\langle \psi_\rho^{[i]} \,\middle|\, f \,\middle|\, \psi_\rho^{[i]} \right\rangle \lambda_\rho^{[i]} = \sum_{i=0}^{n-1} \left\langle \psi_f^{[i]} \middle| \rho \middle| \psi_f^{[i]} \right\rangle \lambda_f^{[i]} \tag{7}$$

The most promising result of Refs. [1, 3] is averaging with the weight in the state $\left|\psi_I^{[IH]}\right\rangle$ of the maximum execution rate $I = dV/dt$ on the past sample. This corresponds to the following density matrix and asset price:

$$\|\rho^{[IH]}\| = \left|\psi_I^{[IH]}\right\rangle \left\langle\psi_I^{[IH]}\right| \tag{8}$$

$$p^{[IH]} = \left\langle \psi_I^{[IH]} \middle| pI \middle| \psi_I^{[IH]} \right\rangle \Big/ \lambda_I^{[IH]} \tag{9}$$

Given a state $|\psi\rangle$, a number of values in this state can be calculated. Just a few examples. Let's define

$$V_s(t) = \int_t^{t_{now}} p^s(t') dV' \tag{10a}$$

$$T_s(t) = \int_t^{t_{now}} p^s(t') dt' \tag{10b}$$

Here, $V_0(t) = V(t_{now}) - V(t)$ is traded volume, $V_1(t)$ is traded capital, $V_1(t)/V_0(t)$ is volume–weighted average price, $T_0(t) = t_{now} - t$, and $T_1(t)/T_0(t)$ is time–weighted average price; these are the values for the time interval: between t and t_{now}. Then $p_{\{v,t\}}$ is {volume,time}–averaged price in the $|\psi\rangle$ state, $p_{\{V,T\}}$ is {volume,time} averaged aggregated price in the

$|\psi\rangle$ state, calculated using the aggregated moments (10). If $|\psi\rangle$ is localized at some given t, then, approximately, $p_{\{v,t\}}$ is the price at t and $p_{\{V,T\}}$ are {volume,time}–weighted price moving average calculated for the time interval between t and t_{now}:

$$p_v = \frac{\langle\psi\,|\,pI\,|\,\psi\rangle}{\langle\psi\,|\,I\,|\,\psi\rangle} \tag{11a}$$

$$p_t = \frac{\langle\psi\,|\,p\,|\,\psi\rangle}{\langle\psi\,|\,\psi\rangle} \tag{11b}$$

$$p_V = \frac{\langle\psi\,|\,V_1\,|\,\psi\rangle}{\langle\psi\,|\,V_0\,|\,\psi\rangle} \tag{11c}$$

$$p_T = \frac{\langle\psi\,|\,T_1\,|\,\psi\rangle}{\langle\psi\,|\,T_0\psi\rangle} \tag{11d}$$

Moments $\langle Q_m V_s\rangle$ and $\langle Q_m T_s\rangle$ can be calculated from moments $\langle Q_m\, p^s I\rangle$ and $\langle Q_m\, p^s\rangle$ and, more generally, moments $\left\langle Q_m \frac{dF}{dt}\right\rangle$ can be calculated from moments $\langle Q_m F\rangle$ using integration by parts (see the Appendices D and E below). In some cases, it is more convenient to directly integrate the wavefunction rather the individual basis functions as in (D1):

$$w_\psi(t) = \int\limits_{-\infty}^{t} \psi^2(x(t'))\omega(t')dt' = \omega(t)J(\psi^2(x(t))) \tag{12}$$

$$\int\limits_{-\infty}^{t_{now}} F(t)\psi^2(x(t))\omega(t)dt = -\int\limits_{-\infty}^{t_{now}} \frac{dF}{dt} w_\psi(t)dt \tag{13}$$

$$F(t_{now}) = 0 : \quad \text{Boundary condition} \tag{14}$$

For the bases we use $\omega(t) = \exp\left(-(t_{now} - t)/\tau\right)$ is monotonic, $x(t)$ is a simple function (linear or exponential), and $J(\cdot)$ in (12) is analytically–known polynomial–to–polynomial mapping function[1]:

$$J(P) = \begin{cases} \dfrac{1}{\exp(x)} \displaystyle\int\limits_{-\infty}^{x} P(x')\exp(x')dx' & \text{Laguerre basis} \\ \dfrac{1}{x} \displaystyle\int\limits_{0}^{x} P(x')dx' & \text{shifted Legendre basis} \end{cases} \tag{15}$$

Averaging with $\psi^2(x(t))\omega(t)dt$ weight gives the value in a pure state $|\psi\rangle$, averaging with $J(\psi^2(x(t)))\omega(t)dt$ weight gives the value in a mixed state: starting since $|\psi\rangle$ till "now".

[1] See the classes `com/polytechnik/freemoney/{WIntegratorLegendreShifted,WIntegratorLaguerre, WIntegratorMonomials}.getPsi2WIntegratedDt()` for numerical implementations)

This allows simultaneously calculate the values of operator pairs: $(\|V_0\|, \|I\|)$, $(\|V_1\|, \|pI\|)$, etc. in the state of a given $\psi(x)$. These operators are known explicitely and all their moments $\langle FQ_s\rangle$, $s = 0 \dots 2n-2$ can be obtained directly from sample, then matrix elements $\langle Q_j \,|\, F \,|\, Q_k\rangle$, $j,k = 0\dots n-1$ are obtained using basis functions multiplication operator (17). However, the situation is different when operator's moments $\langle FQ_s\rangle$, $s = 0 \dots 2n-2$ are not explicitly known, often available only through matrix elements $\langle Q_j \,|\, F \,|\, Q_k\rangle$ that are obtained from some algebra (e.g. an operator as a product and sum of other operators, or an operator with it's eigenvalues adjusted for some reasons, such as the technique of [5] where the eigenvalues (not the eigenvectors!) are adjusted for an effective identification of weak hydroacoustic signals). In this case the average $\int_{-\infty}^{t_{now}} F(t)J(\psi^2(x(t)))\omega(t)dt$ cannot be calculated directly. However, Theorem 3 of [6] establishes a mapping between a polynomial (such as $J(\psi^2(x))$) and a measure, this allows to obtain the moments of a measure that produces a density matrix providing the same average. This way operator's average in a mixed state can be obtained in a regular way as a Spur of the operator with the obtained density matrix even without explicit knowledge of the operator's moments $\langle FQ_s\rangle$, $s = 0\dots 2n-2$, see e.g. `com/polytechnik/trading/trading/DM_DI.java` that uses `com/polytechnik/utils/BasisFunctionsMultipliable.java:getMomentsOfMeasureProducingPolynomialInKK_MQQM` to obtain the density matrix[2].

What input data is required to obtain all the results of this paper? The $n \times n$ matrices $\langle Q_j \,|\, f \,|\, Q_k\rangle$ $(j,k = [0\dots n-1])$: are calculated from generalized moments $(m = [0\dots 2n-2])$:

$$\langle Q_m \rangle \tag{16a}$$

$$\langle Q_m\, I \rangle \tag{16b}$$

$$\langle Q_m\, pI \rangle \tag{16c}$$

$$\left\langle Q_m \frac{dp}{dt} \right\rangle \tag{16d}$$

by applying basis functions multiplication operator (Eq. (G1) of Ref. [3]):

$$Q_jQ_k = \sum_{m=0}^{j+k} c_m^{jk} Q_m \tag{17}$$

[2] Note that while integration density $J(\psi^2(x(t)))\omega(t)$ from (12) is always positive monotonically increasing with t on measure support (the $J(\psi^2(x))$ is a positive polynomial on x interval matching to $t \in (-\infty \dots t_{now}]$ interval), the eigenvalues of the density matrix corresponding to this positive polynomial is not necessary all positive. This means that $J(\psi^2(x))$ is not always a sum of squares.

All the (16) are calculated from (Time, Price, Shares traded) transaction sequence.

III. P&L AND OPTIMAL POSITION CHANGE

Given a directional density matrix $\|\rho\|$, how we do apply it? A naïve answer is to average a directional attribute with it, for example:

- Use price change operator $\|f\| = \|\frac{dp}{dt}\|$ (or $\|f\| = \|\frac{d^2p}{dt^2}\|$ with some boundary condition from the Appendix E), calculate Spur $\|f|\rho\|$; in a pure state $\|\rho\| = |\psi\rangle\langle\psi|$, hereof Spur $\|f|\rho\| = \langle\psi\,|\,f\,|\,\psi\rangle$. Other directional attributes (signed volume, spread multiplied by signed volume, time difference spent in the order book, etc.) can be also considered[4].

- The state determining the dynamics often corresponds to a large dI/dt. Because $dI \approx I(t+dt)-I(t) > 0$, $I = dV/dt$ is larger at the end of the interval. The asset price difference $p_v - p_t$, with volume dV and time dt averaged in a state with such an asymmetry, is proportional to the directional component, where $p_v = \langle\psi\,|\,pI\,|\,\psi\rangle\,/\,\langle\psi\,|\,I\,|\,\psi\rangle$, and $p_t = \langle\psi\,|\,p\,|\,\psi\rangle\,/\,\langle\psi\,|\,\psi\rangle$. Note that such a difference between volume– and time–averaged attribute $p_v - p_t$ carries directional information only in a state of large dI/dt, which makes an asymmetry of price averaging with dV and dt correspond to δp. This is not the case in other states, e.g. trying to use the difference between volume– and time–averaged price in the $\left|\psi_I^{[IH]}\right\rangle$ state was fruitless in [3], see Appendix A for a demonstration. It is now clear why: only the states with large dI/dt provide weight asymmetry required to obtain directional information using dV vs. dt averaging.

In [1] a P&L operator has been introduced in the Section II.E "P&L operator and trading strategy". Given a position change dS, the amount of shares bought ($dS > 0$) or sold ($dS < 0$) during time interval dt, the P&L is[3]:

$$\text{P\&L} = -\int p\,dS \tag{18}$$

$$0 = \int dS \tag{19}$$

[3] While the P&L is $-\int p\,dS$, Eq. (18), the $\int I\,dS$, can be tried as a directional indicator.

The constraint (19) means: total asset position should be zero in the beginning and in the end of a trading period. Formally,

$$dS = \frac{d}{dt}\left(w(t)\frac{dp}{dt}\right)dt \tag{20}$$

where $w(t)$ is an arbitrary positive function, provides positive P&L in (18) (integrate by parts and assume $\frac{dp}{dt} = 0$ at the boundary to satisfy (19)). Position increment dS of optimal P&L trading has a symmetry of the second derivative of price. Note that in (20) other than dp/dt attributes can be used, designate it as $\mathcal{F}$, for example: weighted price change $\mathcal{F} = \delta V \frac{dp}{dt}$ (price change multiplied by the volume traded at this price), signed volume, signed volume multiplied by spread, etc.

There is a dS answer of integral type:

$$dS = \omega(t)\int^{t} dt' \int^{t'} dt''\omega(t'')p(t'') \tag{21}$$

but it's non–local nature and the difficulty to choose integration limits to satisfy the constraint (19) make such an approach more difficult to implement. In the simplest form this approach is equivalent to buying below the median and selling above the median strategy considered in the Appendix E of Ref. [3].

A very promising idea is a "local trading strategy" for dS : in $\left|\psi_I^{[IH]}\right\rangle$ state buy at prices below the $p^{[IH]}$ from (9), sell above the $p^{[IH]}$. Corresponding $\|dS/dt\|$ operator is then:

$$dS = -\left(p - p^{[IH]}\right)dV \tag{22}$$

$$\left\|\frac{dS}{dt}\right\| = -\left\|\left(p - p^{[IH]}\right)I\right\| \tag{23}$$

$$\text{P\&L} = -\left\langle\psi_I^{[IH]}\middle|p\frac{dS}{dt}\middle|\psi_I^{[IH]}\right\rangle = \left\langle\psi_I^{[IH]}\middle|\left(p - p^{[IH]}\right)^2 I\middle|\psi_I^{[IH]}\right\rangle \tag{24}$$

For this dS, in the $\left|\psi_I^{[IH]}\right\rangle$ state, the (19) condition is satisfied, and the P&L has a meaning of price standard deviation (24).

IV. DIRECTIONAL INFORMATION: BEYOND THE WAVEFUNCTION

As we have discussed in [1, 2] the most interesting market behavior is observed at large I, optimization problem $I = \frac{\langle\psi|I|\psi\rangle}{\langle\psi|\psi\rangle} \xrightarrow[\psi]{} \max$ can be reduced to a generalized eigenvalue problem (4) for $\|I\|$ operator:

$$\left|I\middle|\psi_I^{[i]}\right\rangle = \lambda_I^{[i]}\left|\psi_I^{[i]}\right\rangle \tag{25}$$

While the enter/exit conditions can be easily obtained from (25) as in (B1), the directional information is a much trickier problem[3]. In [4], the importance of P&L dynamics was emphasized. In Section III above, several trading strategies (dS), retrospectively providing positive P&L are presented. The goal, however, is to build a strategy providing *future* positive P&L. Consider p_t (11b) in the $\left|\psi_I^{[IH]}\right\rangle$ state:

$$p_t^{[IH]} = \frac{\left\langle\psi_I^{[IH]}\middle|p\middle|\psi_I^{[IH]}\right\rangle}{\left\langle\psi_I^{[IH]}\middle|\psi_I^{[IH]}\right\rangle} \tag{26}$$

$$P^{last} - p_t^{[IH]} = \int dt \frac{dp}{dt} w_{\psi_I^{[IH]}}(t) \tag{27}$$

The (27) is just dp/dt integration with the weigh (12) for $\left|\psi_I^{[IH]}\right\rangle$: the sum of the derivative values with the proper weights give the last price minus the average. The (27) can be expressed via the $\langle Q_m \frac{dp}{dt}\rangle$ moments using an integration by parts of the Appendix D. The problem is reduced to calculation of the moments ($m = [0 \dots 2n-2]$) from observations[4] sample $l = [1 \dots M]$:

$$\left\langle Q_m \frac{dp}{dt}\right\rangle = \sum_{l=1}^{M} \left[p(t_l) - p(t_{l-1})\right] Q_m(x(t_l))\omega(t_l) \tag{28}$$

Then (28) can be substituted for (27) and the best directional answer of Ref. [1]: the last price minus the price in the $\left|\psi_I^{[IH]}\right\rangle$ state is obtained (the (11a) and (11b) are almost identical in the $\left|\psi_I^{[IH]}\right\rangle$ state). These answers are the most general form that can be obtained using the "pure wavefunction approach": all the answers are two quadratic forms ratio, possibly incoherently superposed to a density matrix (7). However, as we have discussed above, "not all observations are equal": only the events with a high I are important for market dynamics. Consider the expression (27) for a general attribute $\mathcal{F}$:

$$\text{DIR\_scalped} = \int dt \mathcal{F}(t) w_{\psi_I^{[IH]}}(t) \tag{29a}$$

$$\langle Q_m \mathcal{F}\rangle = \sum_{l=1}^{M} (t_l - t_{l-1}) \mathcal{F}_l Q_m(x(t_l))\omega(t_l) \tag{29b}$$

For

$$\mathcal{F}_l = \frac{dp}{dt} = \frac{p(t_l) - p(t_{l-1})}{t_l - t_{l-1}} \tag{30}$$

[4] Here the "right" sum is selected to simplify the recurrence by preserving the invariance of the time–grid. One can possibly use the "middle" sum with the $(t_l - t_{l-1})/2$ in the weight $\omega(t)$ and the basis $Q_m(x(t))$ functions argument.

the (29 is exactly the (27) and (28). Consider

$$\mathcal{F}(t) = \frac{dp}{dt}\mathcal{S}(t) \tag{31}$$

$$(t_l - t_{l-1})\mathcal{F}_l = (p(t_l) - p(t_{l-1}))\,\mathcal{S}_l = (t_l - t_{l-1})\,\frac{dp}{dt}\mathcal{S}_l \tag{32}$$

$$\mathcal{S}(t) : [0\dots1] \text{ bounded function} \tag{33}$$

$$\mathcal{S}_l = \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \textbf{: For } t \in [-\infty \dots t_l] \textbf{ interval} \tag{34}$$

Now price change is multiplied by a $[0\dots1]$ bounded scalp–function $\mathcal{S}(t)$ to select "the relevance to market dynamics" of any single observation moment t_l. This way, we can remove from consideration all "irrelevant" observations, as discussed in the introduction; the relevance is determined by estimating whether the current execution flow I_0 is extremely large. The answer obtained in [1, 7] is: for every t_l observation solve the (25) problem for the interval $[-\infty \dots t_l]$ and consider the projection (34) for time–shifted ($t_{now} = t_l$) problem (25). The calculations are straightforward. At time "now," look back at all $[1\dots M]$ market observations, calculate the sum (29b); for every term at t_l also "look back" to construct a separate set of matrices $\langle Q_j \,|\, Q_k \rangle$ and $\langle Q_j \,|\, I \,|\, Q_k \rangle$ for the interval $l' = [1\dots l]$ and calculate the scalp–function $\mathcal{S}$ from (34). This is a problem of $O(M^2)$ complexity when approached directly, but it can be optimized using recurrence relation for the moments calculated for different observation intervals $l' = [0\dots l]$, $l = [0\dots M]$. The major difference with the (27) is that the averaging (29a) can no longer be written in the density–matrix form (7) with the original $\left\langle Q_m \frac{dp}{dt} \right\rangle$ moments. The integration weight in (27) is obtained from the integration of (12). Using Theorem 3 from the Appendix A of Ref. [6], any polynomial $P(x)$ of $2n-2$ degree can be isomorphly mapped to a linear operator of the dimension n, thus the density matrix, corresponding to the $w_{\psi_I^{[IH]}}$ averaging (27), can be readily obtained. This is no longer the case for (29a) averaging. The scalp–function $\mathcal{S}$, while is easy to calculate numerically, does not allow to reduce (29a) averaging to a density matrix averaging (7) of the original moments (28); we now need the scalp–moments (32) to average them with the $w_{\psi_I^{[IH]}}$. This is similar to Bloch wavefunction in quantum mechanics, where the "true" wavefunction is considered as a product of slow and fast oscillating terms. Now we have a product of slow $w_\psi(t)$ and fast $\mathcal{S}(t)$ changing weights in (29a). The greatest advantage of such a transition from regular to scalp–moments, is that the averaging weight can be very sharp. Compare the I_0 in Fig. 1 with, calculated from the (16) input at fixed t_{now}, the "interpolated" $I(y(-\infty \le t \le t_{now}))$ in Fig. 6 of the Appendix A: even for the dimension $n = 12$ obtained wavefunction states are not sufficiently localized to select the sharp spikes in price changes at high I. In the same time the dimension $n = 12$ is perfectly OK for the execution flow I. The scalp–function (33) is a practical way to unify price and execution–flow dynamics within a single framework.

In the Fig. 1 the $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ projection (34) along with I_0 and $\lambda_I^{[IH]}$ are presented. One can clearly see that the (34) is a very good indicator of market activity, the effect we have noticed back in [1, 7]. Now, however, we know how to apply this knowledge: the criterion of current execution flow being extremely high (such as $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$) can be used as a scalp–function $\mathcal{S}$ when calculating the dp/dt moments in (29b): multiply each $p(t_l) - p(t_{l-1})$ by the scalp–function. This way only the relevant (high I) market moves will be accounted in the scalp–moments $\langle Q_m \mathcal{F} \rangle = \left\langle Q_m \mathcal{S} \frac{dp}{dt} \right\rangle$. Typical scalping is price spikes (relatively some "average"–like level) identification technique along with a set of rules to enter a trade and to take a profit/stoploss. As we have shown[2] the spikes in the execution flow, not in the price, are responsible for market dynamics.

The (29b) main idea is to accumulate, with the $Q_m(x(t))\omega(t)dt$ weight, a directional attribute, such as $p(t_l) - p(t_{l-1})$, (Ref. [1] result) *multiplied* by a scalp–function, such as (34) (this paper result); in practice this is just a directional attribute transform (53). Algorithmically, we need to listen for all trading events, and, for each coming event in sequence, obtain a directional attribute $\mathcal{F}_l$ from the regular moments, then calculate (29b) scalp–moments (recurrent optimization make it very efficient computationally) to obtain the directional information (29a). Important, that the value of $\mathcal{F}_l$ is *calculated* from already sampled moments and recent observations. This calculated value is now used to calculate it's moments "as it were directly observed from sample" (e.g. as it were a regular price change). This sampling technique, using *calculated* value as it were a new observable, can be called the "secondary sampling". It can be implemented in several ways:

- Tick trading. As a transaction sequence consider every tick (execution or limit order book event). For every tick l calculate[5]

$$\mathcal{F}_l = \frac{p(t_l) - p(t_{l-1})}{t_l - t_{l-1}} \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \tag{35}$$

[5] Most of $\mathcal{D}_l = p(t_l) - p(t_{l-1}) = 0$ as most trading occur at the same price. Also note that $p_M - p_{m-1} = \sum_{l=m}^{M} \mathcal{D}_l$. For a different weight in the sum obtain (27).

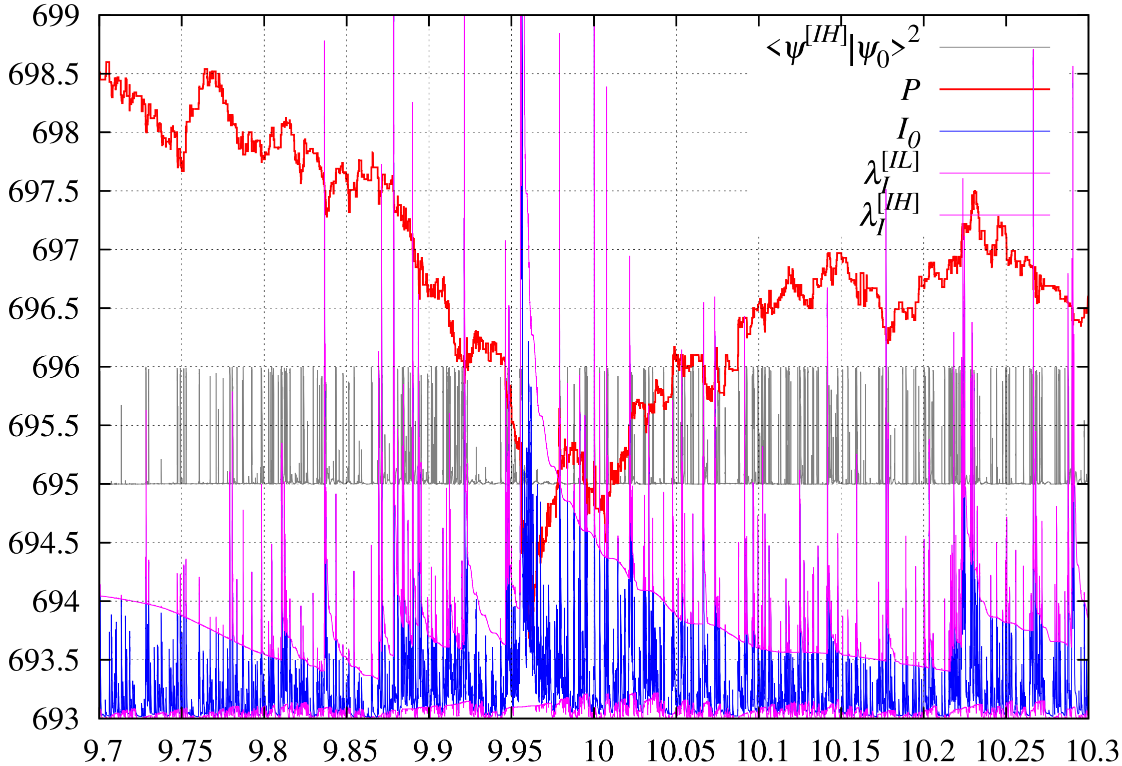


FIG. 1. The AAPL stock price on September, 20, 2012. The calculations in Shifted Legendre basis with $n = 12$ and τ=128sec. The I_0, $\lambda_I^{[IL]}$, $\lambda_I^{[IH]}$, and $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ projection (34) are presented. The execution flow I is scaled and shifted to 693, the projection is shifted to 695 to fit the chart. In between [9.92 . . . 9.94] the execution flow I_0 is small and the $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ is close to zero, thus make this interval non–contributing to scalp–moments. What will happen to them, when the $I_0 = \langle \psi_0 \,|\, I \,|\, \psi_0 \rangle$ is used as a scalp–function instead of the $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$? In the [9.92 . . . 9.94] interval the I_0, while being small, is not particularly zero and the contributions from this interval will propagate to (29b); moreover the $I \to I + \text{const}$ transform makes these contributions even larger. In the same time the (34) is almost zero in irrelevant to market dynamics intervals and is invariant with respect to $I \to I + \text{const}$ transform. Effectively the $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ is the definition of scalp: the condition of I_0 being high[7].

to obtain the "filtered by relevance" moments in (29b).

- Assuming we have all the ticks data[6] , instead of the price difference some average

[6] In practice, for US equity market, a sub–millisecond data can be obtained at reasonable cost. For other markets, such as fixed income, every tick data cannot be practically obtained. Even for currency trading

multiplied by the scalp–function can be used:

$$\mathcal{F}_l = \left\langle \psi_0 \left| \frac{dp}{dt} \right| \psi_0 \right\rangle \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \tag{36}$$

The (36) uses the $|\psi_0\rangle$ for the interval $t \in [-\infty \dots t_l]$ with $t_{now} = t_l$. The ψ_0 from (A1) has an internal time scale $1/\psi_0^2(x_0)$ (which is determined by the basis dimension n and scale τ), thus in (36) the dp/dt is averaged over the time $1/\psi_0^2(x_0)$. The result is very similar to price tick (35) approach, see the Fig. 7 below. A quite similar result can also be obtained with

$$\mathcal{F}_l = \psi_0^2(x_0) \left[\frac{\langle \psi_0 \,|\, pI \,|\, \psi_0 \rangle}{\langle \psi_0 \,|\, I \,|\, \psi_0 \rangle} - \frac{\langle \psi_0 \,|\, p \,|\, \psi_0 \rangle}{\langle \psi_0 \,|\, \psi_0 \rangle} \right] \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \tag{37}$$

this corresponds to described above approach of the difference between volume and time averaged price.

- The (35) and (36) are calculated in the $|\psi_0\rangle$ state. One can consider other states, the $\left|\psi_I^{[IH]}\right\rangle$ is of special interest

$$\mathcal{F}_l = \left\langle \psi_I^{[IH]} \left| \frac{dp}{dt} \right| \psi_I^{[IH]} \right\rangle \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \tag{38a}$$

$$\mathcal{F}_l = 2 \left[\left\langle \psi_I^{[IH]} \middle| pI \middle| \mathrm{ED}(\psi_I^{[IH]}) \right\rangle - \left\langle \psi_I^{[IH]} \middle| pI \middle| \psi_I^{[IH]} \right\rangle \left\langle \psi_I^{[IH]} \middle| \mathrm{ED}(\psi_I^{[IH]}) \right\rangle \right] \tag{38b}$$

$$\begin{aligned} \mathcal{F}_l &= \left\langle \psi_I^{[IH]} \left| \frac{d}{dt} pI \right| \left\langle \psi_I^{[IH]} \right\rangle \right\rangle \\ &= 2 \left[P^{last} \lambda_I^{[IH]} \left\langle \psi_I^{[IH]} \middle| \mathrm{ED}(\psi_I^{[IH]}) \right\rangle - \left\langle \psi_I^{[IH]} \middle| pI \middle| \mathrm{ED}(\psi_I^{[IH]}) \right\rangle \right] \end{aligned} \tag{38c}$$

An important feature of (38) is that some of these $\mathcal{F}_l$ expressions (38b) and (38c) are calculated from $\|pI\|$ operator variation and have: 1. the dimension of capital 2. the $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ factor entering due to the identity

$$\left(\psi_I^{[IH]}(x_0)\right)^2 = 2 \left\langle \psi_I^{[IH]} \middle| \mathrm{ED}(\psi_I^{[IH]}) \right\rangle = \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \psi_0^2(x_0) \tag{39}$$

- An ability to use an expression, calculated from the regular moments $\langle \cdot \rangle$, such as (38) is a very important generalization of price change directional attribute (30). The minimal

---

the fragmentation of the markets along with prohibitively high prices on sub–millisecond data, make any tick–trading approach practically unfeasible. However, as we have discussed in Ref. [3], the time scale of market opportunities (along with liquidity available!) expand well beyond sub–millisecond time scale, maximal scale is determined by the availability of high enough fluctuations in the execution flow I, at least an order of magnitude in $\lambda_I^{[IH]}/\lambda_I^{[IL]}$.

time–scale of such an attribute is $1/\psi_0^2(x_0)$, and the experiment shows that (35) and (36) produce very similar results. This makes promising to consider a directional attributes of more general form in the $|\psi_0\rangle$ state: calculate the $\mathcal{F}_l$, and use it as it were a regular price change. All the previously considered $\mathcal{F}_l$ were some kind a price change analogue. In Ref. [3] two new directional attributes have been introduced: skewness and probability correlation. Consider the skewness (Eq. (66) of Ref. [3]) calculated out of four input moments:

$$\pi_s = \langle\psi_0 \,|\, p^s I \,|\, \psi_0\rangle \tag{40}$$

$$s = 0, 1, 2, 3$$

$$\widetilde{\Gamma} = \frac{2\overline{p} - p_{\min} - p_{\max}}{p_{\min} - p_{\max}} \tag{41}$$

The idea is to build two–point Gauss quadrature (the $p_{\min}$, $p_{\max}$ are min/max nodes of the quadrature, Eq. (64) of Ref. [3], and $\overline{p} = \pi_1/\pi_0$) then to consider it's weight asymmetry as the asymmetry of the distribution. The weigh asymmetry (41) is actually proportional to the difference between the median estimator $(p_{\min} + p_{\max})/2$ and the average $\overline{p}$. One can use the skewness

$$\mathcal{F}_l = [p_{\max} - p_{\min}]\,\widetilde{\Gamma}\left\langle\psi_I^{[IH]}\,\middle|\,\psi_0\right\rangle^2 \tag{42}$$

as a directional attribute instead of price change. The (42) is calculated from the regular moments $\langle Q_k\rangle$, $\langle IQ_k\rangle$, $\langle pIQ_k\rangle$, $\langle p^2IQ_k\rangle$, and $\langle p^3IQ_k\rangle$, then the $\mathcal{F}_l$ is used *as it were observed* at $t = t_l$. This way we substitute price change by the skewness calculated at $1/\psi_0^2(x_0)$ scale. The scalp–function $\left\langle\psi_I^{[IH]}\,\middle|\,\psi_0\right\rangle^2$ makes only relevant to market dynamics observations to contribute.

- Variate the $\|pI\|$ in the $\left|\psi_I^{[IH]}\right\rangle$ state[7] with $|\psi_0\rangle$:

$$\mathcal{F}_l = 2\left[\left\langle\psi_I^{[IH]}\,\middle|\,pI\,\middle|\,\psi_0\right\rangle - \left\langle\psi_I^{[IH]}\,\middle|\,pI\,\middle|\,\psi_I^{[IH]}\right\rangle\left\langle\psi_I^{[IH]}\,\middle|\,\psi_0\right\rangle\right]\left\langle\psi_I^{[IH]}\,\middle|\,\psi_0\right\rangle \tag{43}$$

If $|\psi_0\rangle$ is the $\left|\psi_I^{[IH]}\right\rangle$ then (43) is zero and no directional information is available. The factor $\left\langle\psi_I^{[IH]}\,\middle|\,\psi_0\right\rangle$, which does not enter the (H3), is included in (43) as a scalp–function; this factor also provides proper sign invariance for $\psi \to -\psi$ transform:

[7] The expression has the meaning of capital change due to (39) identity. For single asset consideration it is convenient to divide (43) by $\lambda_I^{[IH]}$.

$\left\langle\psi_I^{[IH]}\middle|\psi_0\right\rangle$ is not squared as it is in (34). One can also use a higher degree of $\left\langle\psi_I^{[IH]}\middle|\psi_0\right\rangle$ factor in (43) to make the peaks sharper.

- Similar to (37), but with $\left|\psi_I^{[IH]}\right\rangle$ and $|\psi_0\rangle$

$$\mathcal{F}_l = \psi_0^2(x_0)\left[\frac{\langle\psi_0\,|\,pI\,|\,\psi_0\rangle}{\langle\psi_0\,|\,I\,|\,\psi_0\rangle} - \frac{\left\langle\psi_I^{[IH]}\middle|pI\middle|\psi_I^{[IH]}\right\rangle}{\left\langle\psi_I^{[IH]}\middle|I\middle|\psi_I^{[IH]}\right\rangle}\right]\left\langle\psi_I^{[IH]}\middle|\psi_0\right\rangle^2 \tag{44}$$

can be considered. This is price difference in $|\psi_0\rangle$ and $\left|\psi_I^{[IH]}\right\rangle$ states. Were it not for the scalp–function $\left\langle\psi_I^{[IH]}\middle|\psi_0\right\rangle^2$ this would be almost Ref. [1] answer: the difference between the last price and $p^{[IH]}$ (9). The scalp–function makes this difference to be accumulated only for the events of extremeny high I_0. The (43) and (44) are zero for $\left|\psi_I^{[IH]}\right\rangle$ being equal to $|\psi_0\rangle$, thus satisfy Ref. [3] Eq. (48) condition of "no directional information about the future available".

- All the $\mathcal{F}_l$ considered above are some kind of price change. Tick trading (35) is last price minus previous price, the other (e.g. (36), (37), etc.) are calculated from the regular moments. It is a promising path to combine tick and moments approaches.

$$\mathcal{F}_l = [p_l - p_l^*]\left\langle\psi_I^{[IH]}\middle|\psi_0\right\rangle^2 \tag{45}$$

$$p_l^* : \text{calculated from } \langle IQ_k\rangle \text{ and } \langle pIQ_k\rangle \text{ moments} \tag{46}$$

Estimating the p_l^* as (9), or (11) in the state $|\psi_0\rangle$ or $\left|\psi_I^{[IH]}\right\rangle$ will not provide a good answer, the (38) is a demonstration. A promising approach is to consider skweness like answer (41). Take (42) but consider it in a different basis of dimension two, replace the basis 1, $p(t)$ by $|\psi_0\rangle$, $\left|\psi_I^{[IH]}\right\rangle$, as these are the states that are localized and relevant to market dynamics:

$$\begin{aligned}
&\begin{pmatrix} \langle\psi_0\,|\,pI\,|\,\psi_0\rangle & \left\langle\psi_0\middle|pI\middle|\psi_I^{[IH]}\right\rangle \\ \left\langle\psi_I^{[IH]}\middle|pI\middle|\psi_0\right\rangle & \left\langle\psi_I^{[IH]}\middle|pI\middle|\psi_I^{[IH]}\right\rangle \end{pmatrix}\begin{pmatrix}\alpha_0^{[0,1]} \\ \alpha_1^{[0,1]}\end{pmatrix} = \\
&= \lambda_{p^*}^{[0,1]}\begin{pmatrix} \langle\psi_0\,|\,I\,|\,\psi_0\rangle & \left\langle\psi_0\middle|I\middle|\psi_I^{[IH]}\right\rangle \\ \left\langle\psi_I^{[IH]}\middle|I\middle|\psi_0\right\rangle & \left\langle\psi_I^{[IH]}\middle|I\middle|\psi_I^{[IH]}\right\rangle \end{pmatrix}\begin{pmatrix}\alpha_0^{[0,1]} \\ \alpha_1^{[0,1]}\end{pmatrix}
\end{aligned} \tag{47}$$

The $\lambda_{p^*}^{[0]}$ and $\lambda_{p^*}^{[1]}$ eigenvalues give the min/max price estimates, that can be obtained in a state of $|\psi_0\rangle$ and $\left|\psi_I^{[IH]}\right\rangle$ superposition. An answer similar to the skewness (42) can be used as an estimator of p_l being low/high:

$$\mathcal{F}_l = -zD_p \left\langle \psi_I^{[IH]} \,\middle|\, \psi_0 \right\rangle^2 \tag{48}$$

$$D_p = \frac{2p_l - \lambda_{p^*}^{[0]} + \lambda_{p^*}^{[1]}}{\lambda_{p^*}^{[0]} - \lambda_{p^*}^{[1]}} \tag{49}$$

$$z = \begin{cases} \frac{|p(t_l)-p(t_{l-1})|}{t_l - t_{l-1}} \\ \frac{V(t_l)-V(t_{l-1})}{t_l - t_{l-1}} \\ \lambda_{p^*}^{[1]} - \lambda_{p^*}^{[0]} \\ \dots \end{cases}$$

The $\mathcal{F}_l$ is proportional to the difference between p_l and $\frac{1}{2}\left[\lambda_{p^*}^{[0]} + \lambda_{p^*}^{[1]}\right]$ is similar to Eq. (95) of Ref. [3]. This is an approach generalizing tick and moments approaches. However, now the (49) is no longer $[-1 : 1]$ bounded (it would be if one replaces p_l by $\langle \psi_0 \,|\, pI \,|\, \psi_0 \rangle \big/ \langle \psi_0 \,|\, I \,|\, \psi_0 \rangle$). A moments–only answer (without the last price used explicitly) can be also obtained:

$$\mathcal{F}_l = zD_p \left\langle \psi_I^{[IH]} \,\middle|\, \psi_0 \right\rangle^2 \tag{50}$$

$$D_p = \frac{\left[\phi^{[1]}(x_0)\right]^2 - \left[\phi^{[0]}(x_0)\right]^2}{\left[\phi^{[1]}(x_0)\right]^2 + \left[\phi^{[0]}(x_0)\right]^2} \tag{51}$$

$$\phi^{[0,1]}(x) = \alpha_0^{[0,1]}\psi_0(x) + \alpha_1^{[0,1]}\psi_I^{[IH]}(x) \tag{52}$$

$$z = \begin{cases} \frac{|p(t_l)-p(t_{l-1})|}{t_l - t_{l-1}} \\ \frac{V(t_l)-V(t_{l-1})}{t_l - t_{l-1}} \\ \lambda_{p^*}^{[1]} - \lambda_{p^*}^{[0]} \\ \dots \end{cases}$$

The sign of (50) is determined by which one of (47) eigenfunctions $\phi^{[0,1]}(x)$ is greater at x_0, (A6) distance from Ref. [3]. The directional factor $\frac{\left[\phi^{[1]}(x_0)\right]^2 - \left[\phi^{[0]}(x_0)\right]^2}{\left[\phi^{[1]}(x_0)\right]^2 + \left[\phi^{[0]}(x_0)\right]^2}$ can be considered as probability correlation (Appendix C of Ref. [3]) between price and "distance to now". In (48) and (50) the scale factor $\left[\lambda_{p^*}^{[1]} - \lambda_{p^*}^{[0]}\right]$ by D_p is replaced by more general form z, what makes the scalp–price to preserve the singularities for a variety of D_p used.

As we have discussed in [1, 4], price and price changes are secondary to execution flow and cannot be used to determine market direction for the reason of insufficient information. The main idea behind the scalp–moments is to replace in the sum (28)

$$p(t_l) - p(t_{l-1}) \rightarrow (t_l - t_{l-1})\mathcal{F}_l \tag{53}$$

$$\mathcal{P}(t_M) = \sum_{l=1}^{M} (t_l - t_{l-1})\mathcal{F}_l \tag{54}$$

$$\text{DIR\_scalped} = \mathcal{P}^{last} - \left\langle \psi_I^{[IH]} \middle| \mathcal{P} \middle| \psi_I^{[IH]} \right\rangle = \int dt \mathcal{F}(t) w_{\psi_I^{[IH]}}(t) \tag{55}$$

where $\mathcal{F}_l$ contains not only price changes, but also execution flow information. A good $\mathcal{F}_l$ selection allows us to accumulate much more directional information in the scalp–moments $\langle Q_m \mathcal{F} \rangle$ compared to the information in the regular moments $\left\langle Q_m \frac{dp}{dt} \right\rangle$. If one sum all the $\mathcal{F}_l$ terms, the $\mathcal{P}$, a generalized price can be obtained (54). The $\mathcal{P}$ is defined within a constant (it is convenient to take the last "price" $\mathcal{P}^{last}$ equals to zero). The transition from price p to the scalp–price $\mathcal{P}$ makes all directional singularities expressed much more clearly. The directional information (29a) now take the (55) form, that is identical to (27), but instead of price p the scalp–price $\mathcal{P}$ is used. If a trader wants to watch the prices — he should be watching the scalp–price $\mathcal{P}$, a much more informative characteristic in terms or market trend, than the regular price p.

A. A Demonstration of Scalp–Price $\mathcal{P}$ Behavior

Before we go any further, let us demonstrate scalp–price (54) $\mathcal{P}(t)$ for a given $\mathcal{F}_l$. The results with $\mathcal{F}_l$ from (35), (36), and (38a) are very similar to each other, so we present only the scalp–price calculated from (36) terms; `.dp_to_use=F_dpdt0_SCALP` in `ScalpedMaxIProjection.java`. The regular price is a sum of all price changes (30), the scalp–price is a sum of relevant to market dynamics (high I) price changes (54). In Fig. 2 regular and scalp–price are presented. One can clearly see, that while the regular price has an erratic behavior due to whatever market moves, the scalp–price $\mathcal{P}$ has a more regular type of behavior. If scalp–price changes it's trend — the trend actually changes. The scalp–price $\mathcal{P}$ (54) is defined within a constant, and it is typically not a good idea to compare regular and scalp–price. However, if one takes an event in the past, where the price is equal to the last price, the change in the scalp–price gives marker direction, i.e. instead of comparing price and scalp–price, one needs to identify a situation of zero price change, then scalp–price change gives market directional information. From a market practitioner's perspective, plain observation

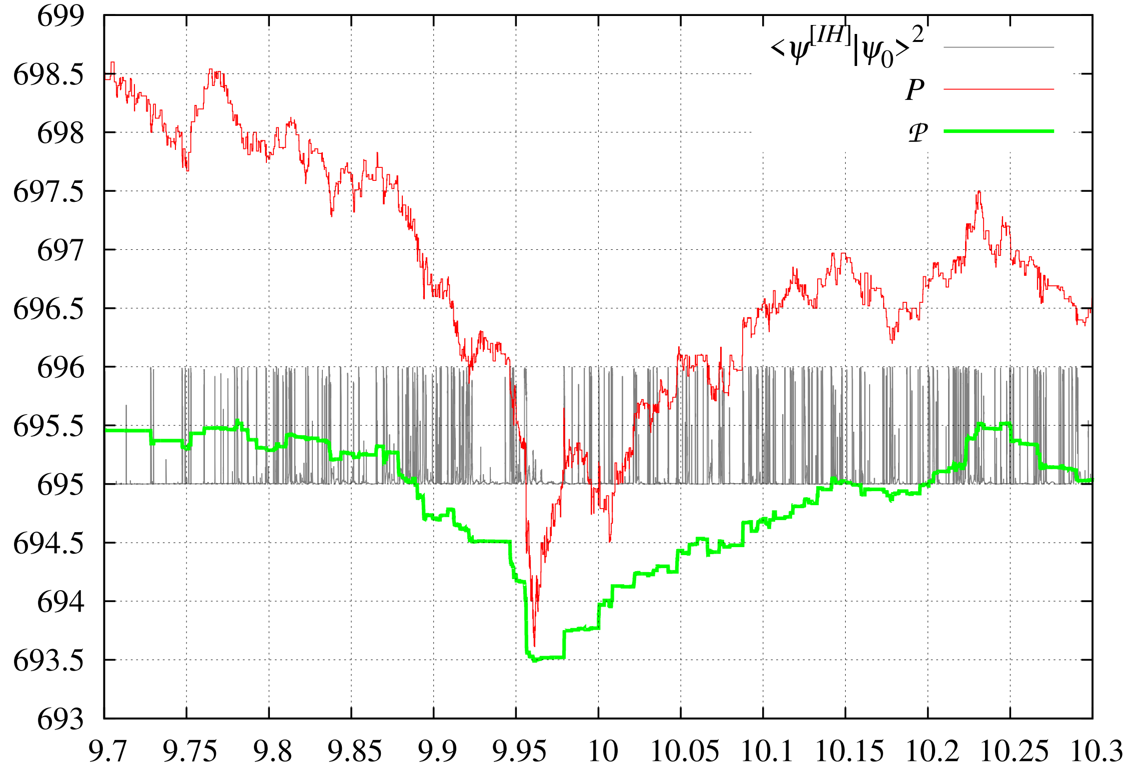


FIG. 2. Price and scalp–price $\mathcal{P}$ for $\mathcal{F}_l$ from (36) are presented. The $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ is used as a scalp–function $\mathcal{S}(t)$ (33). Scalp–price is shifted to fit the chart. See Appendix B 3 for data fields `T`, `p_last`, `shares`, `p_IH`, `I.wH_squared`, and `getSumFdt()` corresponding to: time, price, shares traded, $p^{[IH]}$ (9), scalp–function (34), and scalp–price $\mathcal{P}$ (54).

of the scalp–price is a good source of directional information. As we have discussed above, a typical price behavior after liquidity excess (high I) event is to bounce a little, then go in the original direction of the market. This gives a risk of on what to bet: "little bounce" or "follow the market". The $\mathcal{P}$, obtained from (36) $\mathcal{F}_l$, has no "little bounce" contributions; watching the $\mathcal{P}$ is actually watching pure market trend. If the price moves, and the scalp–price stays — this typically indicates a bear market rally or a bull market sell–off. The $\mathcal{P}$ is an integral attribute. The $\mathcal{F} = d\mathcal{P}/dt$ is a local attribute. One can try the

$$\left\langle \psi_I^{[IH]} \middle| \mathcal{F} \middle| \psi_I^{[IH]} \right\rangle = \left\langle \psi_I^{[IH]} \middle| \frac{d\mathcal{P}}{dt} \middle| \psi_I^{[IH]} \right\rangle \tag{56}$$

attribute (not show in Fig. 2, see `.F_IH` field of `ScalpedMaxIProjection.java` output), but the result is worse compared to the $\mathcal{P}$ result, no clear bear/bull market switch is observed. The situation is similar to the one in Fig. 2 of Ref. [3]: dp/dt chart in the $\left| \psi_I^{[IH]} \right\rangle$ state.

B. A Demonstration of the Directional Information

The directional information should be accumulated over an interval of a substantial duration for the reason of low information available in a single price change. However, the strategies as the last price minus the average will never work for the reasons of fixed time scale of price averaging. In [1], the time–scale of the state of maximal past I, the $\left|\psi_I^{[IH]}\right\rangle$, was introduced and the (27) answer was obtained. In this paper, the next critically–important step is made: instead of regular price p, the scalp–price $\mathcal{P}$ (it includes only high I events: only relevant to market dynamics price moves) is introduced and the (55) answer is obtained. In the Fig. 3 (bottom) two directional answers are presented. In the top chart moving average $\frac{\langle p \rangle}{\langle 1 \rangle}$ and $p_t^{[IH]}$ are presented. In the bottom chart the (27) (`.dp_to_use=F_SAMPLE_DP_NOSCALP`), and (55) (`.dp_to_use=F_dpdt0_SCALP`), they are normalized to the same integral taken with all $\mathcal{F}_l$ positive in (27) and (55). One can clearly see that:

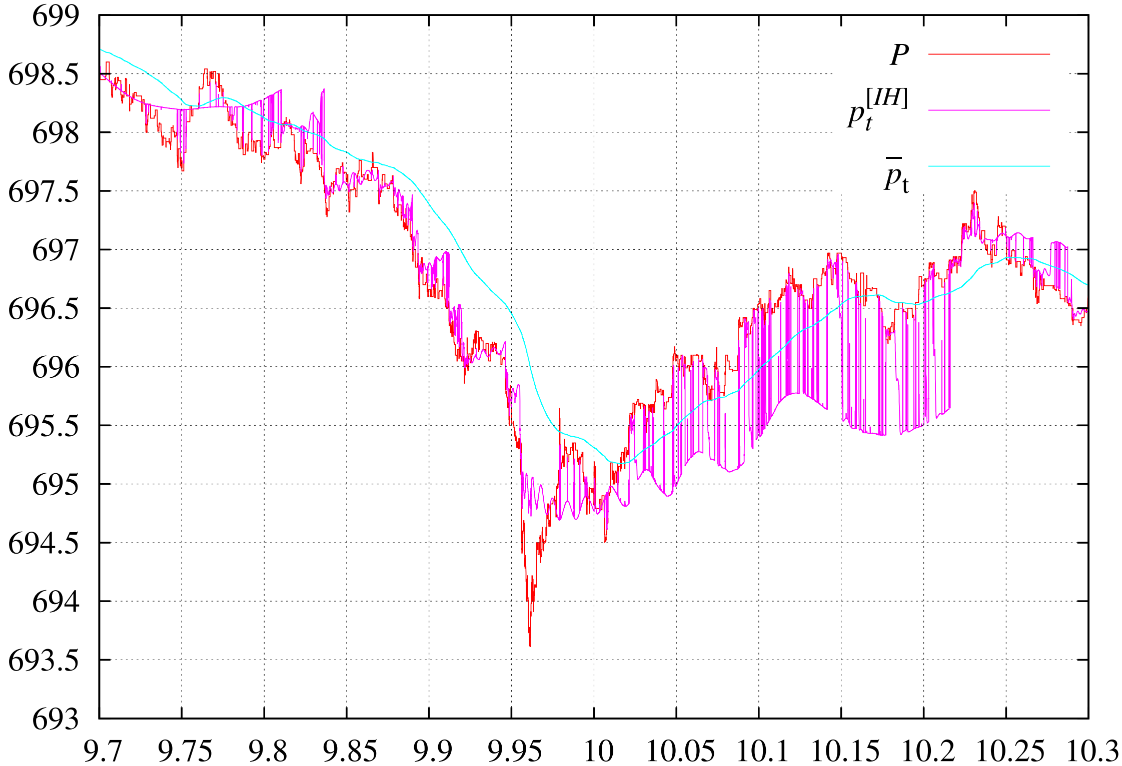
- When divided by the absolute variation, the non—scalped answer (27) is pretty small, and the scalped one (55) is much larger. This means that the price can be moved due to a variety of reasons, and only scalped price changes (31) are relevant to the market dynamics. Moreover, high I market moves are much more consistent.

- Look at $t \in [9.9 \dots 9.95]h$ interval. The price bounce around $p_t^{[IH]}$, what make it difficult to trade the direction as $P^{last} - p_t^{[IH]}$. In the same time the scalp–price (55) stays in the same sign, the scalp function $\mathcal{S}(t)$ (34) is about zero in this interval, see Fig. 2

- Look around $t = 10h$. The scalped answers captured all the relevant price changes and switched from bear to bull market. The execution flow I defines market sentiment.

C. On the Selection of $\mathcal{F}_l$

The selection of $\mathcal{F}_l$ to be summed to the scalp price $\mathcal{P}$ (54) is the most important question for directional attribute selection. Consider two choices. In Fig. 4 top $\mathcal{F}_l = \langle \psi_0 \,|\, dp/dt \,|\, \psi_0 \rangle$ from (36) is presented. One can see that dp/dt (green) has rather erratic behavior, that is caused by a variety of market moves, the sum of these moves gives the regular price p. But when each price move is multiplied by the scalp–function $\mathcal{S} = \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ from (34),



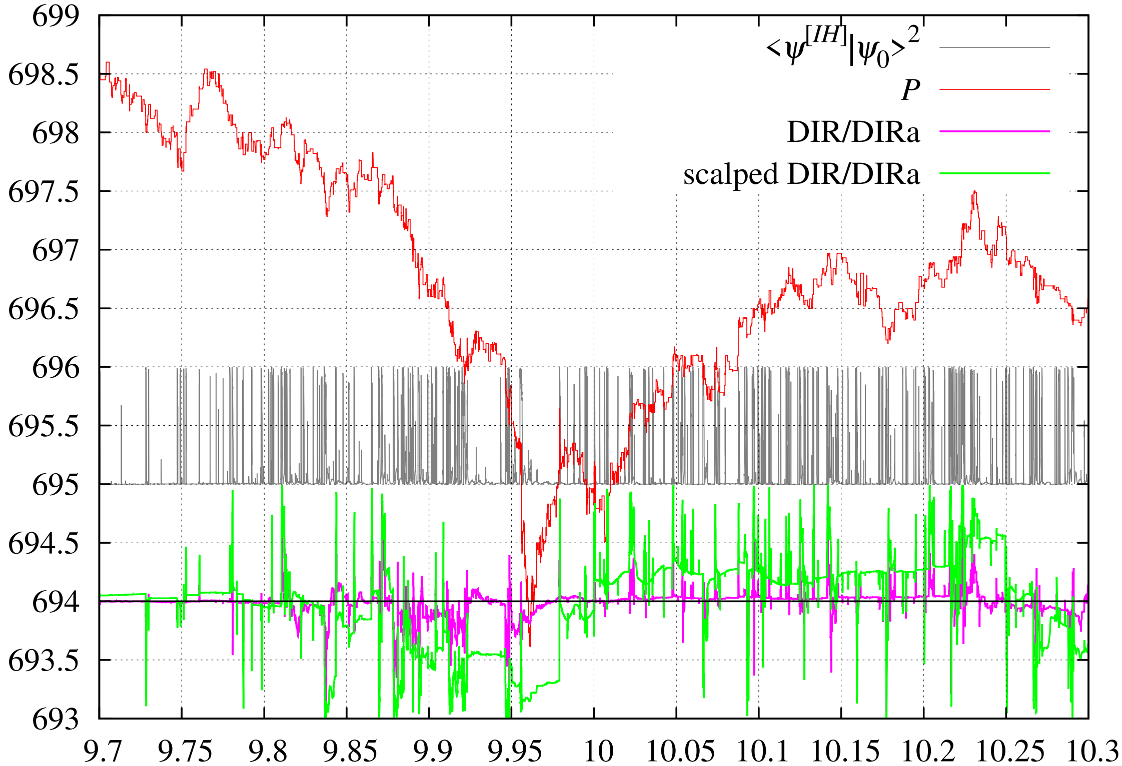


FIG. 3. Top: $\frac{\langle p \rangle}{\langle 1 \rangle}$ (moving average), and $p_t^{[IH]}$ (11b). Bottom:The DIR $= P^{last} - \left\langle \psi_I^{[IH]} \middle| p \middle| \psi_I^{[IH]} \right\rangle$ (27) and DIR_scalped $= \mathcal{P}^{last} - \left\langle \psi_I^{[IH]} \middle| \mathcal{P} \middle| \psi_I^{[IH]} \right\rangle$ (55); both DIRs are **normalized** to all $\mathcal{F}_l$ taken positive (normalized to total variation).

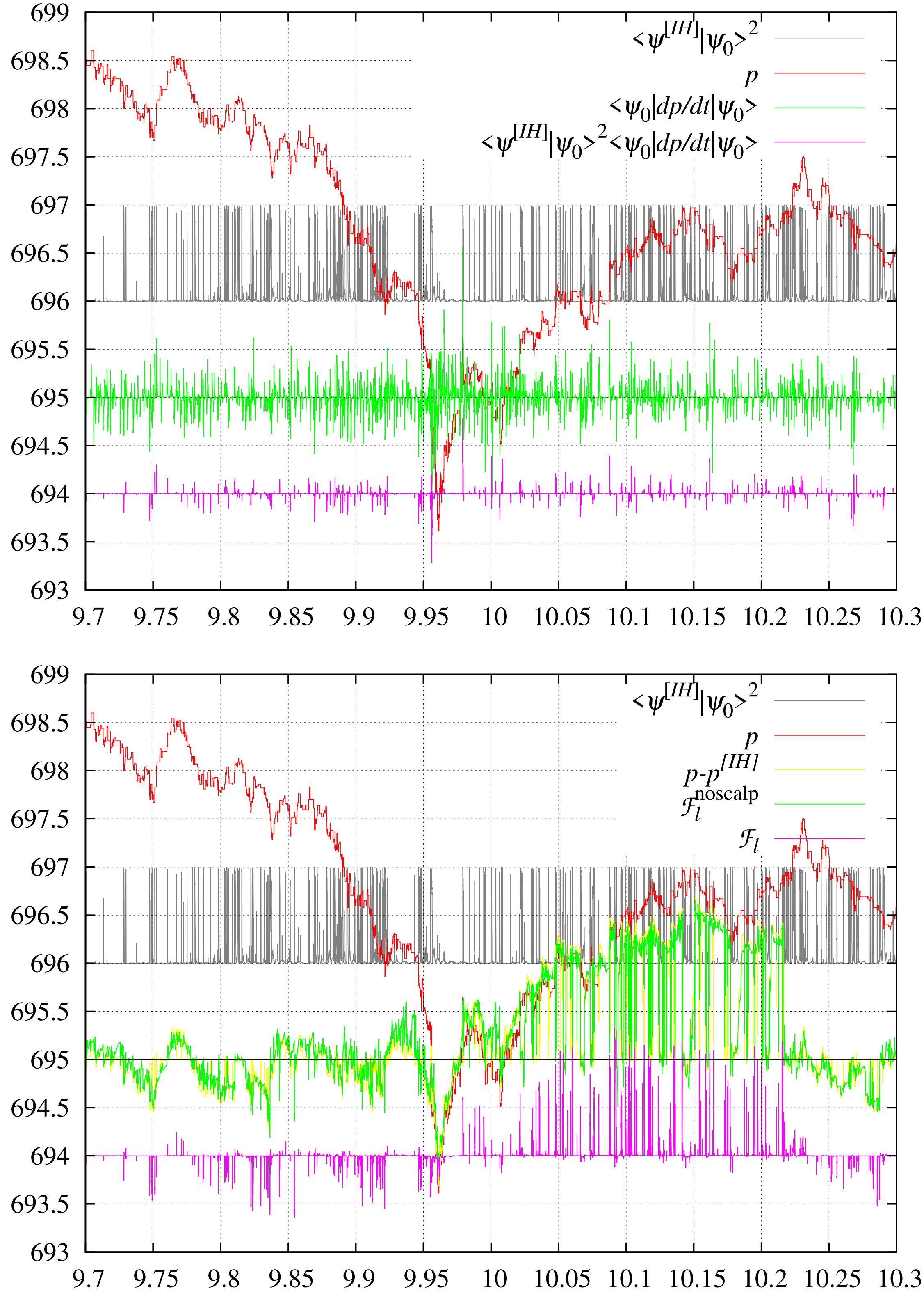


FIG. 4. Top: the scalp–function $\mathcal{S}$ (34), $\langle \psi_0 \,|\, dp/dt \,|\, \psi_0 \rangle$, and $\mathcal{F}_l$ from (36), Bottom: same $\mathcal{S}$ (34), $\mathcal{F}_l$ from (50) for $z = \lambda^{[1]}_{p^*} - \lambda^{[0]}_{p^*}$ with (pink) and without (green) scalp function multiplied; the $p - p^{[IH]}$ (yellow) is also presented. The values are shifted to 694, 695, and 696 levels and scaled to fit the chart.

this selects only high I market moves, what makes the directional behavior much more clear (pink), the sum now gives the scalp price $\mathcal{P}$ from the Fig. 2. But even in this simplistic case the scalp–price selects only "relevant" market moves.

A much more interesting behavior can be observed with $\mathcal{F}_l$ from (50). Selecting $|\psi_0\rangle$, $\left|\psi_I^{[IH]}\right\rangle$ basis, solving (47), then obtaining (50) — this accumulates much more directional information in $\mathcal{F}_l$. When $\left|\psi_I^{[IH]}\right\rangle$ and $|\psi_0\rangle$ are not close to each other, the (50) for $z = \lambda_{p^*}^{[1]} - \lambda_{p^*}^{[0]}$ is approximately equals to last price and $p^{[IH]}$ difference multiplied by the scalp–function (that is close to zero). When $\left|\psi_I^{[IH]}\right\rangle$ and $|\psi_0\rangle$ are close to each other the (50) does not vanish[8], i.e. the (50) does not vanish (like $p - p^{[IH]}$, yellow) when I_0 is extremely high. However, the $\mathcal{F}_l$ enters into the integral (55), and the selection of the z is a non–trivial task. The most important feature of the charts in Fig. 4 is that once we got a spike in the $\left|\psi_I^{[IH]}\right\rangle$ state — the trend is going to continue. These spikes are much greater in values (because of non–local price difference) compared to local price difference $p(t_l) - p(t_{l-1})$ of Eq. (35). This allows to collect much more directional information, than can can be typically obtained from price changes.

We can generalize this non–local price change approach. Consider $p_t^{[IH]}$ in Fig. 3 ($p^{[IH]}$ is very close to it). A typical behavior for $p^{[IH]}$ is to jump from some past value to last price when the execution flow I_0 becomes large, (34) is the criteria of I_0 largeness. How often these jumps occur is the criteria to determine market direction, see Fig. 4. These non–local structural changes in $\left|\psi_I^{[IH]}\right\rangle$ can be included to scalp–price, for every tick l calculate:

$$p^{[IH]}(t_l) \tag{57a}$$

$$\lambda_I^{[IH]}(t_l) \tag{57b}$$

$$\mathcal{S}(t_l) = \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2 \tag{57c}$$

All (57) values are calculated from the sequence: $(t_m, p(t_m), V(t_m)); m = 1 \dots l$

These are just Eq. (25) solution performed for every observation tick l using $m = 1 \dots l$ previous ticks as input data. This is the result we had obtained back in Ref. [1]. The **new** idea is to consider the $p^{[IH]}(t_l)$ as if it were the last price $p(t_l)$. This way one tick price change becomes non–local:

$$p(t_l) - p(t_{l-1}) \to p^{[IH]}(t_l) - p^{[IH]}(t_{l-1}) \tag{58}$$

[8] In this case $\mathcal{F}_l$ from (44) is almost zero, but $\mathcal{F}_l$ from (37) does not vanish, while providing a much smaller response than (50)

Depending on the execution flow, the $\left|\psi_I^{[IH]}\right\rangle$ may (or may not) change drastically at every tick. One tick non–local difference $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ can be much greater than one tick local price difference $p(t_l) - p(t_{l-1})$, see Fig. 3. As we discussed in the introduction only *low* $I \to$ *high* I to be considered:

$$\theta_{I+}(t_l) = \begin{cases} 1 & \text{if } \lambda_I^{[IH]}(t_l) \geq \lambda_I^{[IH]}(t_{l-1}) \\ 0 & \text{otherwise} \end{cases} \tag{59}$$

$$(t_l - t_{l-1})\mathcal{F}_l = z\left[p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})\right]\theta_{I+}(t_l) \tag{60}$$

With a number of possible options for z:

$$z = 1 \tag{61a}$$

$$z = \mathcal{S}(t_l)\left[V(t_l) - V(t_{l-1})\right] \tag{61b}$$

$$z = dt\mathcal{S}(t_l)\left[\lambda_I^{[IH]}(t_l) - \lambda_I^{[IH]}(t_{l-1})\right] \tag{61c}$$

$$\dots$$

This is the $\mathcal{F}_l$ to be used in (55). The (60) considers every *low* $I \to$ *high* I jump in $p^{[IH]}(t_l)$ (not in $p(t_l)$) as the source of the directional information.

In Fig. 5 a demonstration of non–local price change (58) is presented. Only $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ with positive $\lambda_I^{[IH]}(t_l) - \lambda_I^{[IH]}(t_{l-1})$ are presented (the Eq. (60) with $z = 1$ and $dt = 1$). On can clearly see that the non–local directional information is:

- Much greater than the local price change $p(t_l) - p(t_{l-1})$.
- The bull/bear market trend switch can be much better identified. The $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ with (59) constaint preserves the sign during extended intervals.
- The "bounce back" interval $t \in [9.9 \dots 9.95]h$ is clearly identified: it has no I spikes, the $\left|\psi_I^{[IH]}\right\rangle$ does not change, and the $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ is close to zero even without $\mathcal{S}$ multiplied!

This makes us to conclude, that non–local price change (58) taken with the constraint (59) provides a very promising possible directional indicator. Fig. 5 presents a non–local price answer (58) obtained from one tick $p^{[IH]}$ price change $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ as it were one tick regular price change $p(t_l) - p(t_{l-1})$. This answer is similar (but much better) than Fig. 4

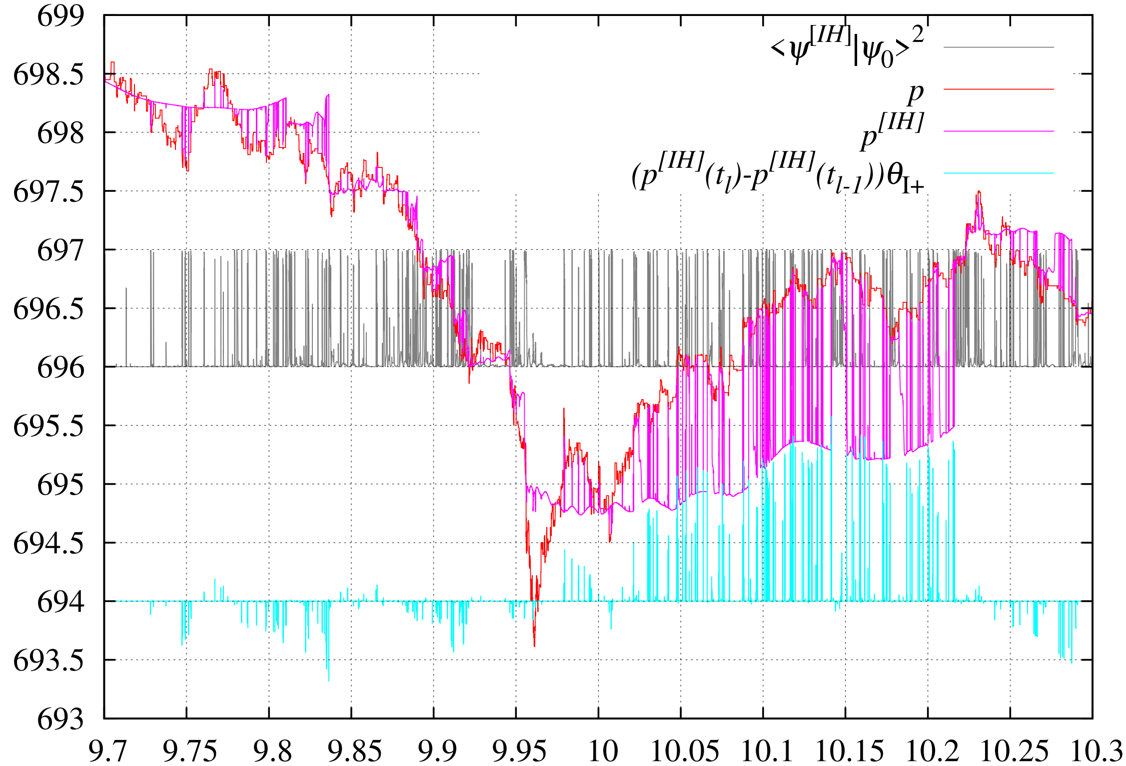


FIG. 5. The price, scalp–function $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$, $p^{[IH]}$ (9) (pink), and $\mathcal{F}_l$ from (60) without z term (blue). $\mathcal{F}_l$ is presented as one tick $p^{[IH]}$ change $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ multiplied by $\theta_{I+}(t_l)$ factor (59).

(bottom) answer, that is obtained from the regular moments by solving (47) $d = 2$ eigenvalue problem. The (60) is the directional indicator. However, because it enters the integral (54), the selection of proper integration weight z is required. This to be a subject of a separate study. In the simplest form a non–local answer can be obtained from (57) solution of (25) problem, then consider:

- Only $\lambda_I^{[IH]}(t_l) \geq \lambda_I^{[IH]}(t_{l-1})$ events: $\theta_{I+}(t_l) > 0$, Eq. (59), field (B2t).

- For such events consider $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ as it were one tick price change $p(t_l) - p(t_{l-1})$, Eq. (58), field (B2u). In Fig. 5 an example of such a non–local price changes is presented.

V. ON THE DIRECTIONAL INFORMATION CALCULATION

Given very interesting results of the previous section, let us formulate all the components, required to obtain directional information from (time, execution price, shares traded) market observations triples, and how these components can be improved.

- The state important for market dynamics. The answer we have is (8). Other states (such as considered in the Appendix C) can be also tried. In any case such a state is obtained from regular moments (16a) and (16b), solving some kind of $I \underset{\psi}{\longrightarrow} \max$ problem. The solution gives us open/close position signals and the scale for directional calculations.

- The problem to obtain the direction is way more complex, it requires scalp–moments (55). For the scalp–function $\mathcal{S}$ the best[7] answer is (34). For $\mathcal{F}_l$ several answers (35), (36), (37), and (38a) produce good results, that are very similar to each other, the non–local answer (60) is of special interest. The "varied" answers (38b), and (38c) are worse with and without scalp–function multiplied. The simplest practical abswer is the scalp–moments directional answer (55), as a scale one can use absolute variation: take all $\mathcal{F}_l$ positive in (55). However, a number of non–local answers of (60) type can be obtained utilizing (57) and (58).

VI. SPECULATIONS

The scalp–moments are price change moments *filtered* by high I events: I is the driving force of the market. The question arises whether a directional information can be obtained from the regular moments (16)? We are inclined to say no. A number of constrained (see Appendices F and G below) and unconstrained optimization problems have been tried (among many others) without any success at obtaining market directional information:

$$\max_{\psi} \frac{\left\langle \psi \,\middle|\, (p - p^{[IH]})^2 I \,\middle|\, \psi \right\rangle}{\langle \psi \,|\, \psi \rangle} \qquad \texttt{DynHPnL.java} \tag{62a}$$

$$\max_{\psi} \frac{\left\langle \psi \,\middle|\, (p - P^{last})^2 I \,\middle|\, \psi \right\rangle}{\langle \psi \,|\, \psi \rangle} \qquad \texttt{DynPnL.java},\ \texttt{PnLSensitivity.java} \tag{62b}$$

$$\max_{\psi} \min_{p_x} \frac{\langle \psi \,|\, (p - p_x)^2 I \,|\, \psi \rangle}{\langle \psi \,|\, \psi \rangle} \qquad \texttt{DynYp.java},\ \texttt{DIminP2maxI.java} \tag{62c}$$

$$\max_{\psi} \max_{p_x} \frac{\langle \psi \,|\, I \,|\, \psi \rangle}{\langle \psi \,|\, (p-p_x)^2 \,|\, \psi \rangle} \qquad \texttt{DynYp.java}, \texttt{DIminP2tmaxI.java} \tag{62d}$$

$$\max_{\psi} \min_{p_1,p_2} \frac{\langle \psi \,|\, (p-p_1)^2(p-p_2)^2 I \,|\, \psi \rangle}{\langle \psi \,|\, \psi \rangle} \qquad \text{Ref. [3] Section IX F, } \texttt{PnLdIV4.java} \tag{62e}$$

$$\max_{\psi} \frac{\left\langle \psi \,\middle|\, (p-p_t)^2 \, I \,\middle|\, \psi \right\rangle}{\langle \psi \,|\, \psi \rangle} \qquad \texttt{MaxPtPv2I.java}, \texttt{MinMaxPnLratioNorm.java} \tag{62f}$$

$$\max_{\psi} \frac{\left\langle \psi \,\middle|\, (p-p_v)^2 \, I \,\middle|\, \psi \right\rangle}{\left\langle \psi \,\middle|\, (p-p_t)^2 \,\middle|\, \psi \right\rangle} \qquad \texttt{MaxPnLratio.java} \text{ ; } \texttt{flag\_swap\_PtPv=false} \tag{62g}$$

$$\max_{\psi} \frac{\left\langle \psi \,\middle|\, (p-p_t)^2 \, I \,\middle|\, \psi \right\rangle}{\left\langle \psi \,\middle|\, (p-p_v)^2 \,\middle|\, \psi \right\rangle} \qquad \texttt{MaxPnLratio.java} \text{ ; } \texttt{flag\_swap\_PtPv=true} \tag{62h}$$

$$\max_{\psi} \frac{\langle \psi \,|\, I \,|\, I \,|\, \psi \rangle}{\langle \psi \,|\, (p-P^{last})^2 I \,|\, \psi \rangle} \qquad \texttt{MaxPPl2I.java}, \texttt{MaxPPl2Iinbasis.java} \tag{62i}$$

The regular moments answers are: 1. not "sufficiently sharp", see Appendix A, and 2. price changes sum is small relatively total variation, see Fig. 3. In the same time, when we go to the scalp–moments (32) these problems get solved.

When, in September 1997, I joined Columbus Advisors LLC (Greenwich CT), the fund had been doing Emerging Market sovereign fixed income convergence–divergence relative value spread trades. The following year, I studied a classic technical analysis book with the goal to program some of the rules algorithmically. However, I was not able to program even *a single rule* from the book. The reason was simple: any rule required a time scale to apply. Time scale selection is the main criterion separating good traders from bad, and the criterion which defines a trader's talent. The state (8) is an algorithmic criterion, that automatically determines the time scale. This criterion is actually very simple ideologically: look back to find an event of trading with maximal I. The time between this event and "now" is the time scale. The typical market practitioner's activity is to watch the difference between the last price and moving average calculated on the time scale obtained his feel. With a proper time scale, any strategy (like return to the moving average) would work, and Ref. [1] answer of last price minus $p^{[IH]}$ (9) was my first successful attempt.

Besides the time scale, the most important result of this paper is that "not all price moves are equal". We need to select only the high I price moves[9]. High execution rate requirement is the condition creating an asymmetry to separate the "bounce a little, then to go in the original direction of the market" and "go in the original direction of the market straight away" scenarios, such as to identify a bear market rally on steroids. The answer we obtained

[9] I think that the market impact concept is a dead end.

is the scalp–price (54). It does not have any "internal averaging", but in the same time it has all low I price changes removed! This way, the scalp–price has no "bounce a little" behavior. Only hardcore. Only directional. See the Fig. 2. The software is available[8] under the GPLv3 license.

ACKNOWLEDGMENTS

Vladislav Malyshkin would like to thank Aleksandr Bobyl from the Ioffe Institute for many discussions of whether the activity of a typical market participant can be reduced to some "simplified automata" with intelligence as an aggregate function of past transactions, and, even more importantly, the ability (or inability) to apply the obtained knowledge to market activity; Emilio J. Lamar for discussions of slow/fast market transitions — especially the spread–widening effect during fast markets; and Aleksandr Maslov from the Ioffe Institute for discussions on contradistinction of operators $\|dV/dt\|$ and $\|V/t\|$.

Appendix A: A demonstration of the difference between time and volume weighted price.

To demonstrate the difference consider localized at $x = y$ the wavefunction $\psi_y(x)$ (A1), producing Radon–Nikodym interpolating answer, Eq. (7) of Ref. [9], Different attributes (price, execution flow, etc.) are interpolated using the $\psi_y^2(x)\omega(x)dx$ weight:

$$\psi_y(x) = \frac{\sum\limits_{j,k=0}^{n-1} Q_j(x) G^{-1}_{jk} Q_k(y)}{\sqrt{\sum\limits_{j,k=0}^{n-1} Q_j(y) G^{-1}_{jk} Q_k(y)}} \tag{A1}$$

$$1 = \langle \psi_y \,|\, \psi_y \rangle \tag{A2}$$

$$I(y) = \langle \psi_y \,|\, I \,|\, \psi_y \rangle \Big/ \langle \psi_y \,|\, \psi_y \rangle \tag{A3}$$

$$p_t(y) = \langle \psi_y \,|\, p \,|\, \psi_y \rangle \Big/ \langle \psi_y \,|\, \psi_y \rangle \tag{A4}$$

$$p_v(y) = \langle \psi_y \,|\, pI \,|\, \psi_y \rangle \Big/ \langle \psi_y \,|\, I \,|\, \psi_y \rangle \tag{A5}$$

One can see that:

- For a large n (we use $n = 12$) the p_t and p_v are very similar.

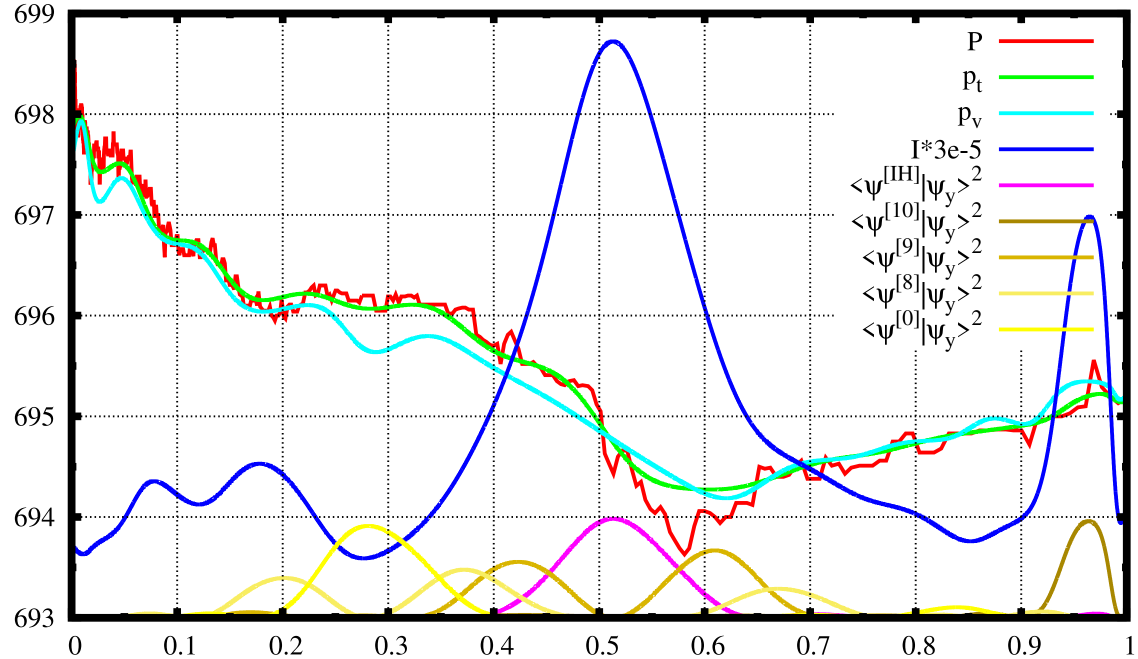


FIG. 6. The AAPL stock price on September, 20, 2012. Interpolation answers are calculated in Shifted Legendre basis with $n = 12$ and τ=128sec, for $0 \leq t \leq t_{now} = 9.98045$ hrs, $y = \exp\left((t - t_{now})/\tau\right)$, $y = [0 \ldots 1]$. Execution flow (A3), time (A4), and volume (A5) weighted prices are presented. One can clearly see the $p_v(y) - p_t(y)$ changes the sight at y, corresponding to a high I. The maximal eigenstate IH, (#11=$n-1$), $\left\langle \psi_y(x) \middle| \psi_I^{[IH]} \right\rangle^2$, pink, is typically a localized state. The projections $\left\langle \psi_y(x) \middle| \psi_I^{[i]} \right\rangle^2$ on four other eigenstates (#0, #8, #9, and #10), yellow, are presented as an example of delocalized states. The execution flow I and the projection are shifted to 693 to fit the chart.

- The projection $\left\langle \psi_y(x) \middle| \psi_I^{[IH]} \right\rangle^2$ is close to 1 for large I, i.e. the $\psi_I^{[IH]}(x)$ is typically a localized function, this is not the case for other states. See four other eigenstates projections (yellow).

- The $p_v - p_t$ changes the sign at large I. Only the states with a large dI/dt provide weight asymmetry required to obtain directional information using dV vs. dt averaging.

Appendix B: Computer Implementation

The codebase architecture is described in the Appendix G 3 of Ref. [3]. Relevant to this paper functionality consists of:

- Conversion of the transaction sequence of an observable f to a vector of moments $\langle fQ_m\rangle$, $m = [0 \dots 2n-2]$, several bases $Q_k(x)$ are implemented ($x = t$, $x = \exp\left(-(t_{now}-t)/\tau\right)$, and $x = p(t)$, see the Section II of Ref. [3]), the integration measure is always exponential decay: $d\mu = \exp\left(-(t_{now}-t)/\tau\right)dt$. See the classes `com/polytechnik/trading/{QVMDataL,QVMDataP,QVMData}.java`

- Using basis functions multiplication operator (Eq. (G1) of Ref. [3]), obtain the $\langle Q_j \,|\, f \,|\, Q_k\rangle$, $j,k = [0 \dots n-1]$ matrix from the moments $\langle fQ_m\rangle$, $m = [0 \dots 2n-2]$.

- There are a number of observables f possibly to consider (price, price change, execution flow, etc.). Depending on the approach used, a different set of observables is required. All the $\langle Q_j \,|\, f \,|\, Q_k\rangle$ matrices we possibly use in this paper are stored in the class `com/polytechnik/trading/SMomentsData.java`.

- If/When, in addition to a $\langle Q_j \,|\, f \,|\, Q_k\rangle$ matrix, the matrix corresponding to the derivative df/dt (or to the integral $\int^t f(t')dt'$), is required, then, for a basis with infinitesimal time–shift operator $\mathrm{ED}(Q(x))$ (E3), the result can be obtained using integration by parts, see Appendices D and E.

As a result of these preliminary steps the $n\times n$ matrices are obtained: $\langle Q_j \,|\, Q_k\rangle$, $\left\langle Q_j \,\middle|\, \frac{dp}{dt} \,\middle|\, Q_k\right\rangle$, $\langle Q_j \,|\, p \,|\, Q_k\rangle$, $\langle Q_j \,|\, p^2 \,|\, Q_k\rangle$, $\langle Q_j \,|\, p^3 \,|\, Q_k\rangle$, $\langle Q_j \,|\, I \,|\, Q_k\rangle$, $\langle Q_j \,|\, pI \,|\, Q_k\rangle$, $\langle Q_j \,|\, p^2I \,|\, Q_k\rangle$, and $\langle Q_j \,|\, p^3I \,|\, Q_k\rangle$. These are plain exponential moving–average of: an observable f multiplied by two basis functions product $Q_j(x)Q_k(x)$; for example if $f = p$, then $\langle Q_0 \,|\, p \,|\, Q_0\rangle$ is exponential moving average of price.

1. The `EVXData.java` implementation

The class `com/polytechnik/utils/EVXData.java` takes two matrices $\langle Q_j \,|\, f \,|\, Q_k\rangle$, $\langle Q_j \,|\, Q_k\rangle$ and basis functions operations class (extending the `com/polytechnik/utils/OrthogonalPolynomialsBasisFunctionsCalculatable.java`), solves generalized eigenvalue problem, such as (25) for $\|f\| = \|I\|$, and stores the result. The fields are:

$$\texttt{.sL} = \lambda_I^{[IL]} \tag{B1a}$$

$$\texttt{.sH} = \lambda_I^{[IH]} \tag{B1b}$$

$$\texttt{.s0} = \langle \psi_0 \,|\, I \,|\, \psi_0 \rangle \tag{B1c}$$

$$\texttt{.wL} = \left\langle \psi_I^{[IL]} \middle| \psi_0 \right\rangle \tag{B1d}$$

$$\texttt{.wH} = \left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle \tag{B1e}$$

The squares $\texttt{.wL}^2$ and $\texttt{.wH}^2$ are bounded to $[0:1]$, and are very good indicators of whether the I "now", the $I_0 = \langle \psi_0 \,|\, I \,|\, \psi_0 \rangle$, is large or small. Alternative estimator as the number of the eigenvalues above the I_0 can also be used[7]. The key concept of liquidity deficit trading[1, 4] is to open a position at low I_0, large $\left\langle \psi_I^{[IL]} \middle| \psi_0 \right\rangle^2$, then to close already opened position at high I_0, large $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$, the $\texttt{.wL}^2$ and $\texttt{.wH}^2$ are the indicators of these actions. The question is: whether to open a *long* or a *short* position at high $\texttt{.wL}^2$?

2. The `ScalpedMaxIProjection.java` implementation

The class `com/polytechnik/trading/ScalpedMaxIProjection.java` converts a transaction sequences to a set of $\langle f Q_m \rangle$ vectors, then to a set of $\langle Q_j \,|\, f \,|\, Q_k \rangle$ matrices, stored in the object of `com/polytechnik/trading/SMomentsData.java` type. Then it calls the `com/polytechnik/utils/EVXData.java` class that solves (25) eigenvalue problem. Having the $\left|\psi_I^{[IH]}\right\rangle$ and $|\psi_0\rangle$ states the scalp–function (34) and the $\mathcal{F}_l$ are obtained. Which one $\mathcal{F}_l$ to be used depends on the parameter `.dp_to_use`. The values `F_SAMPLE_DP_NOSCALP,F_SAMPLE_DP_SCALP,F_dpdt0_SCALP,F_varpIH_0_divI_SCALP,F_SKEWNESS_at_Pl_SCALP,F_PROBABILITYCORRELATION_SCALP` correspond to (30), (35), (36), (43), (48), and (50) respectively; there are several other options for `.dp_to_use`. The class `ScalpedMaxIProjection` is assumed to be called on every tick, and the internal state is preserved in the object of `com/polytechnik/trading/StateWIScalpMomentsSaver.java` class. The internal state contains an object of `com/polytechnik/freemoney/WIntegrator.java` type, that calculates the moments of the observable $\mathcal{F}_l$ (recurrent shift of the basis offset (t_{now}) for previously calculated moments allows the calculations to be performed extremely fast). The `WIntegrator` is called on every tick with the $\texttt{.Fdt} = (t_l - t_{l-1})\,\mathcal{F}_l$ (the choice of the

$\mathcal{F}_l$ depends on the `.dp_to_use` value) to accumulate scalped data. The scalp–moments are obtained by taking the `.Fdt` instead of the $p(t_l) - p(t_{l-1})$ when calculating the (29b) sum. The directional information is then obtained as (29a). The fields are:

`.p_offset`	Price offset. All prices are relatively this offset	(B2a)
`.pi_average`	Volume–weighted price exponential moving average $\dfrac{\langle pI\rangle}{\langle I\rangle}$	(B2b)
`.pt_average`	Time–weighted price exponential moving average $\dfrac{\langle p\rangle}{\langle 1\rangle}$	(B2c)
`.I`	An object of `EVXData.java` type, (25) solution	(B2d)
`.p_0`	The $\langle\psi_0 \,\vert\, pI \,\vert\, \psi_0\rangle \,/\, \langle\psi_0 \,\vert\, I \,\vert\, \psi_0\rangle$	(B2e)
`.pt_0`	The $\langle\psi_0 \,\vert\, p \,\vert\, \psi_0\rangle \,/\, \langle\psi_0 \,\vert\, \psi_0\rangle$	(B2f)
`.dpdt_0`	The $\left\langle \psi_0 \middle\vert \dfrac{dp}{dt} \middle\vert \psi_0 \right\rangle$	(B2g)
`.p_IH`	The (11a) in the $\left\vert\psi_I^{[IH]}\right\rangle$ state (8)	(B2h)
`.pt_IH`	The (11b) in the $\left\vert\psi_I^{[IH]}\right\rangle$ state (8)	(B2i)
`.pV_IH`	The (11c) in the $\left\vert\psi_I^{[IH]}\right\rangle$ state (8)	(B2j)
`.pT_IH`	The (11d) in the $\left\vert\psi_I^{[IH]}\right\rangle$ state (8)	(B2k)
`.var1pI_IH`	The (38b)	(B2l)
`.var1pI_IH_00`	The (43)	(B2m)
`pmin_0_IH`, `pmax_0_IH`	The eigenvalues $\lambda_{p^*}^{[0,1]}$ of (47)	(B2n)
`Skewness_0_IH`	The "skewness" (49)	(B2o)
`ProbabilityCorrelation_0_IH`	Directional factor $\dfrac{\left[\phi^{[1]}(x_0)\right]^2 - \left[\phi^{[0]}(x_0)\right]^2}{\left[\phi^{[1]}(x_0)\right]^2 + \left[\phi^{[0]}(x_0)\right]^2}$ (51)	(B2p)
`.I.wH`	When squared `.I.wH`^{2} gives the scalp function (34)	(B2q)
`.getFlFromRegularMoments()`	$\mathcal{F}_l$ when it is from the moments, `NaN` otherwise	(B2r)
`.sst.getSumFdt()`	The scalp–price (54) with an arbitrary offset	(B2s)
`.dIH`	The $\lambda_I^{[IH]}(t_l) - \lambda_I^{[IH]}(t_{l-1})$ difference	(B2t)

`.dp_IH`	The $p^{[IH]}(t_l) - p^{[IH]}(t_{l-1})$ difference	(B2u)
`.DIR`	The (29a)	(B2v)
`.aDIR`	The (29a) with all $\mathcal{F}_l$ taken positive	(B2w)

The liquidity deficit indicator (B2d) defines whether to open or to close a position. The directional indicator (B2v) from (29a) defines, when opening a position, whether to open a long or a short.

3. The `CallAMuseOfCashFlowAndLiquidityDeficitWithScalp.java` implementation

The class `com/polytechnik/algorithms/CallAMuseOfCashFlowAndLiquidityDeficitWithScalp.java` is "an interface" between transactions sequence input (a tab–separated file), liquidity deficit trading of the class `com/polytechnik/trading/ScalpedMaxIProjection.java`, and data output, saved as a tab–separated file. The parameters are read by the class `com/polytechnik/algorithms/MuseConfig.java`. This is an example of how to run the code:

```
java com/polytechnik/algorithms/CallAMuseOfCashFlowAndLiquidityDeficitWithScalp \
      --musein_file=aapl.csv \
      --musein_cols=15:1:4:5 \
      --museout_file=museout.dat \
      --n=12 \
      --tau=128 \
      --measure=ScalpedMaxIProjectionLegendreShifted
```

The parameters are:

- `--musein_file=aapl.csv` : Specify input tab–separated file with (time, execution price, shares traded) triples time series. If the file is `gzip`–ed and has the `.gz` extension, then internal decompression is performed.

- `--musein_cols=15:1:4:5` : Out of total 15 columns in the specified `--musein_file=aapl.csv` file, take the column #1 as time (nanoseconds since midnight), #4 (execution price), and #5 (shares traded), column index is base 0.

- `--museout_file=museout.dat` : Output file name.
- `--n=12` : Basis dimension. Typical values are: $n \in [4 \dots 16]$. The $m \in [0 \dots 2n-2]$ moments (in $Q_m(x)$ basis) are calculated to obtain $n \times n$ matrices.
- `--tau=128` : Exponent time (in seconds) for the measure used.
- `--measure=ScalpedMaxIProjectionLegendreShifted` : The measure. Possible values are: `{ScalpedMaxIProjectionLegendreShifted,ScalpedMaxIProjectionLaguerre,ScalpedMaxIProjectionMonomials}`, they correspond to the measures (11) and (4) of Ref. [3]. The `ScalpedMaxIProjectionLaguerre` and `ScalpedMaxIProjectionMonomials` use the same measure, but different basis $Q_k(x) = L_k(x)$, $x = -t/\tau$ and $Q_k(x) = x^k$, $x = t/\tau$ respectively. These two results should be *identical*, as the measure is the same, and all the calculations are $Q_k(x)$–basis invariant (but the numerical stability can be drastically different).

Output file is a tab–separated file with the columns (35 columns total), corresponding to the results of this paper. Field names are printed in the first line of the output file. The data can be processed by any common plotting software (such as gnuplot or matlab). Below is the description of the most noticeable fields:

- `T` : Time in nanoseconds since midnight (copied from input).
- `shares` : Shares traded (copied from input).
- `P_last` : Last execution price (copied from input).
- `{pi_average,pt_average}` : Regular exponential moving average of price with the given `--tau=128`, using volume/time as the weight.
- `I.{s0,sL,wL_squared,sH,wH_squared,Gamma0}` : Correspond to (B1) fields of $\left|I|\psi\right\rangle = \lambda\left|\psi\right\rangle$ eigenvalue problem (25), the solution with the given `--n=12`; the `I.wL` and `I.wH` are squared in the output, `Gamma0` $= \left(2I_0 - \lambda_I^{[IL]} - \lambda_I^{[IH]}\right) \Big/ \left(\lambda_I^{[IL]} - \lambda_I^{[IH]}\right)$ is the $\widetilde{\Gamma^0}$ skewness of I, Eq. (95) of Ref. [3]. The `I.wH_squared` is the scalp–function $\mathcal{S}(t)$ (34).
- `{p_IH,pt_IH,pV_IH,pT_IH}` Correspond to (11) prices, calculated in the state $\left|\psi_I^{[IH]}\right\rangle$ (8), the (B2) fields.

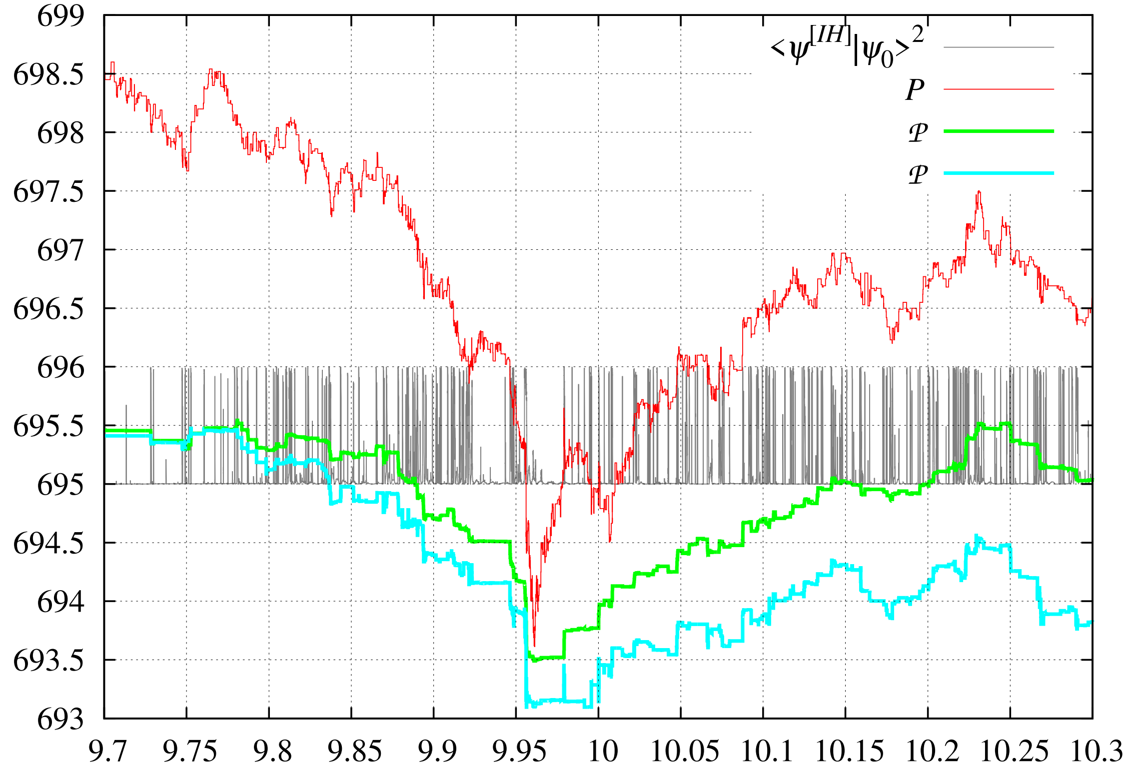


FIG. 7. The comparison of scalp–price $\mathcal{P}$ obtained from $\mathcal{F}_l$: from (36) (green: `.dp_to_use=F_dpdt0_SCALP`) from (35) (blue: `.dp_to_use=F_SAMPLE_DP_SCALP`). The $\left\langle \psi_I^{[IH]} \middle| \psi_0 \right\rangle^2$ is used as a scalp–function $\mathcal{S}(t)$ (33). The scalp–prices are shifted to fit the chart; they are defined (54) within a constant. If one use `.dp_to_use=F_SAMPLE_DP_NOSCALP` (30) the result will be exactly the price P, shifted by some initial level.

- `getFlFromRegularMoments()` The $\mathcal{F}_l$ when it is calculated from regular moments, `NaN` otherwise, the field (B2r).
- `getSumFdt()` The scalp–price $\mathcal{P}$ (54), corresponding to given `dp_to_use`, the field (B2s). See the Fig. 7 to compare the results for `.dp_to_use=F_dpdt0_SCALP` (36) and `.dp_to_use=F_SAMPLE_DP_SCALP` (35).
- `dIH,dp_IH` The $\lambda_I^{[IH]}$ and $p^{[IH]}$ change *per tick*, the fields (B2t) and (B2u). This is the starting point of non–local price change (58) study, Fig. 5.
- `{DIR,DIRa}` and etc. Correspond to (B2) fields of an object of `ScalpedMaxIProjection.java` type.

4. Installation and usage example

- Install java 1.8 or later.

- Download from [8] the archive `AMuseOfCashFlowAndLiquidityDeficit.zip` with the source code.

- Decompress and recompile the program:

```
unzip AMuseOfCashFlowAndLiquidityDeficit.zip
javac -g com/polytechnik/*/*java
```

- Run the test with the bundled file `dataexamples/aapl_old.csv.gz` data of Ref. [2]. The file contains only execution events, the (time, execution price, shares traded) market observations triples are in the 1:2:3 columns, column index is base 0; 28492 lines, 9 columns total.

```
java com/polytechnik/algorithms/CallAMuseOfCashFlowAndLiquidityDeficitWithScalp \
      --musein_file=dataexamples/aapl_old.csv.gz \
      --musein_cols=9:1:2:3 \
      --museout_file=museout.dat \
      --n=12 \
      --tau=128 \
      --measure=ScalpedMaxIProjectionLegendreShifted
```

  The code is run under 16 seconds, the output fields of the `museout.dat` are described in the Appendix B 3. The `I.wH_squared`, `getSumFdt()`, and `p_IH` are the scalp–function (34), scalp–price (54) (has an arbitrary offset), and $p^{[IH]}$ from (9). The default `.dp_to_use=F_PROBABILITYCORRELATION_SCALP` corresponds to (50).

- Download NASDAQ ITCH data file `S092012-v41.txt.gz` from [8], extract triples (time, execution price, shares traded) from NASDAQ ITCH data file:

```
java com/polytechnik/itch/DumpData2Trader \
      S092012-v41.txt.gz AAPL >aapl.csv
```

Execution data and limit order book edges are now saved to tab–separated file `aapl.csv` of 15 columns. The (time, execution price, shares traded) market observations triples are in the 1:4:5 columns, column index is base 0; 634205 lines, 15 columns total.

- Run the java command of the Appendix B 3 to obtain the `museout.dat` file of 634206 lines with: scalp–function (34), scalp–price (54) and $p^{[IH]}$ from (9) and `com/polytechnik/trading/ScalpedMaxIProjection.java` fields is created. The code is run under 5 minutes, much longer than that of previous run. The `--musein_file=aapl.csv` input file now contains much more events than the file `--musein_file=dataexamples/aapl_old.csv.gz`.

Appendix C: The state of maximal aggregated execution flow V/t

In our previous work[1, 3] the extremal state of $I = dV/dt$ operator have been considered. This answer has two critically important features:

- Uses execution flow I, as it is the driving force of the market.
- Has automatic time–scale selection (eigenvalue problem), huge advantage compared to any fixed time scale approach[1].

While this result is very promising, it has an issue of zero first variation of I. Consider the same approach, but with the operator V/t. Here V and t are measured *since* t_{now}, they are volume/time between t and t_{now}. The V/t is **aggregated** execution flow, the dV/dt is **local** execution flow. Put $f = V/t$ into (4) and obtain generalized eigenvalue problem to find the state $\left|\psi^{[\max]}_{V/t}\right\rangle$ of maximal $\lambda^{[\max]}_{V/t}$:

$$\left|V\middle|\psi^{[i]}_{V/t}\right\rangle = \lambda^{[i]}_{V/t}\left|t\middle|\psi^{[i]}_{V/t}\right\rangle \tag{C1}$$

$$\sum_{k=0}^{n-1}\langle Q_j\,|\,V\,|\,Q_k\rangle\,\alpha^{[i]}_k = \lambda^{[i]}_{V/t}\sum_{k=0}^{n-1}\langle Q_j\,|\,t\,|\,Q_k\rangle\,\alpha^{[i]}_k \tag{C2}$$

$$\psi^{[i]}_{V/t}(x) = \sum_{k=0}^{n-1}\alpha^{[i]}_k Q_k(x) \tag{C3}$$

The calculation of $\langle Q_j\,|\,V\,|\,Q_k\rangle$ and $\langle Q_j\,|\,t\,|\,Q_k\rangle$ matrix elements is described in the Appendix D. In (C2) the V and t have the sign changed to have positively definite right–hand–side matrix $\langle Q_j \,|\, t \,|\, Q_k \rangle$, $V = V_0$ (10a), $t = T_0$ (10b). The multiplication by V and t create, for $t \leq t_{now}$, two Radau–like measures: $(V(t_{now}) - V(t))\,\omega(t)dt$ and $(t_{now} - t)\,\omega(t)dt$. The problem (C2) finds the state $\psi_{V/t}^{[\max]}(x)$, corresponding to the maximal Radon–Nikodym derivative relatively two these measures, the maximal aggregated execution flow V/t. Previously [1] we have been considering the state $\psi_I^{[IH]}(x)$, corresponding to the maximal Radon–Nikodym derivative relatively the measures $\omega(t)dV$ and $\omega(t)dt$, the maximal local execution flow dV/dt. The eigenvectors $\left|\psi_{V/t}^{[i]}\right\rangle$ of $\|V/t\|$ operator have the following remarkable features:

Normalized to Radau–like measure $(t_{now} - t)\,\omega(t)dt$:

$$1 = \left\langle \psi_{V/t}^{[i]} \middle| t \middle| \psi_{V/t}^{[i]} \right\rangle \tag{C4}$$

In the $\left|\psi_{V/t}^{[i]}\right\rangle$ states aggregated V/t and local dV/dt execution flows are equal:

$$\lambda_{V/t}^{[i]} = \frac{\left\langle \psi_{V/t}^{[i]} \middle| V \middle| \psi_{V/t}^{[i]} \right\rangle}{\left\langle \psi_{V/t}^{[i]} \middle| t \middle| \psi_{V/t}^{[i]} \right\rangle} = \frac{\left\langle \psi_{V/t}^{[i]} \middle| I \middle| \psi_{V/t}^{[i]} \right\rangle}{\left\langle \psi_{V/t}^{[i]} \middle| \psi_{V/t}^{[i]} \right\rangle} \tag{C5}$$

For infinitesimal time–shift $\delta\psi = \mathrm{ED}(\psi_{V/t}^{[i]})$ the second variation (H4) of V/t is equal to the first variation (H3) of dV/dt:

$$\langle \delta\psi \,|\, V \,|\, \delta\psi \rangle - \lambda_{V/t}^{[i]} \langle \delta\psi \,|\, t \,|\, \delta\psi \rangle = \left\langle \delta\psi \middle| I \middle| \psi_{V/t}^{[i]} \right\rangle - \lambda_{V/t}^{[i]} \left\langle \delta\psi \middle| \psi_{V/t}^{[i]} \right\rangle \tag{C6}$$

Lemma. *In the state of maximal aggregated execution flow the dI/dt is positive.*

Proof. In the state of maximal V/t the second variation (H4) is negative. Because the first I variation (H3) with $\delta\psi = \mathrm{ED}(\psi_{V/t}^{[i]})$ corresponds to $-dI/dt$, this provides positive dI/dt. □

This lemma makes the state $\left|\psi_{V/t}^{[\max]}\right\rangle$ of maximal aggregated execution flow (the eigenvector of (C2), corresponding to the maximal $\lambda_{V/t}^{[\max]}$), a very promising one for the market dynamics to consider. The "aggregated" attributes (10) have been originally introduced in the Section (IX D) "Measure: The Period After Maximal Future I" of Ref. [3], but their application to skewness study was not a very successful back then.

Appendix D: On calculation of $\langle VQ_m \rangle$ moments from $\langle IQ_m \rangle$ moments

The $\langle VQ_m \rangle$ and $\langle tQ_m \rangle$, $m = [0 \dots 2n-2]$, moments, required to construct $\|V\|$ and $\|t\|$ operators in (C1), can be calculated directly from the sample. However, in practical application it is more convenient to calculate the $\langle IQ_m \rangle$ moments first, then to obtain the $\langle VQ_m \rangle$ moments using an integration by parts. For Shifted Legendre and Laguerre bases the integration by parts gives:

$$\int_{-\infty}^{t_{now}} VQ_m(x(t))\omega(t)dt = V(t_{now})Q_m(x(t_{now})) - \int_{-\infty}^{t_{now}} J(Q_m(x(t)))\omega(t)Idt \qquad \text{(D1)}$$

where $J(\cdot)$ is a polynomial to polynomial transforming function (12). The $\langle VQ_m \rangle$ then can be expressed as $\langle IQ_s \rangle$, $s = [0 \dots m]$, linear combination. This is possible *only* for the bases in question, in general case an integration by parts $\int_{-\infty}^{t} Q_m(x(t'))\omega(t')dt'$ cannot be reduced to a $J(Q_m(x(t)))\omega(t)$ form, and the $\langle VQ_m \rangle$ moments cannot be expressed via a linear combination of the $\langle IQ_s \rangle$ moments.

The boundary condition is straightforward, consider $V(t) - V(t_{now})$, that is zero at $t = t_{now}$. Use current volume $V(t_{now})$ as the starting value, then out–of–integral term in (D1) vanish, and past/future volume correspond to negative/positive volume values[10]. See the method `setFMoments` of `com/polytechnik/trading/{QVMDataLDirectAccess,QVMDataPDirectAccess,QVMDataDirectAccess}.java`, that calculates the $\langle VQ_m \rangle$ moments as a linear combination of the $\langle IQ_s \rangle$, $s = [0 \dots m]$ moments.

Appendix E: On calculation of $\|dI/dt\|$ operator matrix elements from operator $\|I\|$.

When we study an operators of execution rate change $\|dI/dt\|$, it's matrix elements cannot be calculated directly from sample. In general case the $\left\langle \frac{dI}{dt} Q_m \right\rangle$ moments can be calculated from $\langle IQ_m \rangle$ moments using integration by parts (D1) of the Appendix D, see the method `setDFMoments` of `com/polytechnik/trading/{QVMDataLDirectAccess,QVMDataPDirectAccess,QVMDataDirectAccess}.java`, that, for zero boundary condition, obtains $\left\langle \frac{dI}{dt} Q_m \right\rangle$ as a linear combination of $\langle IQ_s \rangle$, $s = [0 \dots m]$. However, for $\|dI/dt\|$, the boundary condition may take a variety of forms, and direct operator approach is often more convenient.

[10] It is sometimes convenient to change the sign of time and volume $V(t) - V(t_{now})$ as in (10), then past time and volume correspond to positive values and the right hand side matrix in (C1) is positively definite.

Consider e.g. generalized eigenvalue problem (4) for $\|dI/dt\|$ operator:

$$\left|\frac{dI}{dt}\middle|\psi^{[i]}_{\frac{dI}{dt}}\right\rangle = \lambda^{[i]}_{\frac{dI}{dt}}\left|\psi^{[i]}_{\frac{dI}{dt}}\right\rangle \tag{E1}$$

where the $\left\langle Q_j\left|\frac{dI}{dt}\right|Q_k\right\rangle$ matrix cannot be calculated directly from sample. For a $Q_k(x)$ basis with infinitesimal time–shift operator $\mathrm{ED}(Q_k(x))$,

$$\frac{d}{dt}\omega(x(t))\psi(x(t))\varphi(x(t)) = \omega(x)\left[\mathrm{ED}(\psi)\varphi + \psi\mathrm{ED}(\varphi)\right] \tag{E2}$$

$$\mathrm{ED}(\psi(x)) = \begin{cases} \dfrac{d\psi(x)}{dx} + \dfrac{1}{2}\psi(x) & \text{Laguerre basis} \\ x\dfrac{d\psi(x)}{dx} + \dfrac{1}{2}\psi(x) & \text{shifted Legendre basis} \end{cases} \tag{E3}$$

providing time-derivative of a polynomial multiplied by a weight is represented by the *same weight* multiplied by other polynomial, the matrix can be obtained from the $\langle Q_j\,|I|\,Q_k\rangle$ matrix using integration by parts[11], Eq. (35) of Ref. [3]:

$$\left\langle Q_j\left|\frac{dI}{dt}\right|Q_k\right\rangle = I^f Q_j(x_0)Q_k(x_0) - \langle \mathrm{ED}(Q_j)\,|\,I\,|\,Q_k\rangle - \langle Q_j\,|\,I\,|\,\mathrm{ED}(Q_k)\rangle \tag{E4}$$

This problem is an inverse one to considered in Appendix D, and requires a non–trivial boundary condition I^f. There are several options for I^f, that can be reasonably considered:

- The zero of $\|dI/dt\|$ in the $\left|\psi^{[IH]}_I\right\rangle$ state, $\left\langle \psi^{[IH]}_I\left|\frac{dI}{dt}\right|\psi^{[IH]}_I\right\rangle = 0$:

$$I^f = \lambda^{[IH]}_I \tag{E5}$$

- The zero of $\|dI/dt\|$ in the $|\psi_0\rangle$ state, $\left\langle \psi_0\left|\frac{dI}{dt}\right|\psi_0\right\rangle = 0$:

$$I^f = 2\frac{\langle \psi_0\,|\,I\,|\,\mathrm{ED}(\psi_0)\rangle}{\psi_0^2(x_0)} \tag{E6}$$

- The I_0 value:

$$I^f = \langle \psi_0\,|\,I\,|\,\psi_0\rangle \tag{E7}$$

- Zero value:

$$I^f = 0 \tag{E8}$$

[11] See java classes for Shifted Legendre and Laguerre $Q_k(x)$ bases implementation of infinitesimal time–shift operator $\mathrm{ED}(Q_k(x))$: the method `getEDPsi` of `com/polytechnik/freemoney/{WIntegratorLegendreShifted,WIntegratorLaguerre,WIntegratorMonomials}.java`. Also see the `com/polytechnik/trading/QQdidtMatrix.java` class, implementing the calculation of (E4) matrix for the (E5), (E6), (E7), and (E8), boundary conditions. This class uses `com/polytechnik/utils/VolMatrix.java` to calculate $\langle \mathrm{ED}(Q_j)\,|\,I\,|\,Q_k\rangle + \langle Q_j\,|\,I\,|\,\mathrm{ED}(Q_k)\rangle$, then adds boundary condition term $I^f Q_j(x_0)Q_k(x_0)$.

Regardless the I^f selection, the $\|I\|$ and $\|dI/dt\|$ operators have no common eigenvectors unless the $|\psi_0\rangle$ is the $\|I\|$ eigenvector, this degeneracy case was considered in Ref. [3]. The most critical degeneracy arise in the situation, when the state "now" and the state of "maximal past I" are the same:

$$|\psi_0\rangle = \left|\psi_I^{[IH]}\right\rangle \tag{E9}$$

An example of such a degeneracy can be the situation of huge volume traded "now" (at $x = x_0$).

Appendix F: Directional Information: $I \xrightarrow[\psi]{} \max$ Subject To the Constraint $\langle\psi \,|\, C \,|\, \psi\rangle = 0$.

Consider market dynamics split in two operators: $\|I\|$ (execution flow dynamics) and $\|C\|$ (price dynamics). The constrained $I \to \max$ problem is:

$$I = \frac{\langle\psi \,|\, I \,|\, \psi\rangle}{\langle\psi \,|\, \psi\rangle} \xrightarrow[\psi]{} \max \tag{F1a}$$

$$\text{subject to:} \quad 0 = \langle\psi \,|\, C \,|\, \psi\rangle \tag{F1b}$$

The constraint (F1b) is a requirement on price in the $|\psi\rangle$ state. There are a number of choices for the constraint operator $\|C\|$ selection:

$$\|C\| = \|\left(p - P^{last}\right) I\| \qquad \text{Price (11a) in the } |\psi\rangle \text{ state is equal to } P^{last} \tag{F2a}$$

$$\|C\| = \|V_1 - P^{last} V_0\| \qquad \text{Moving average price (11c) is equal to } P^{last} \tag{F2b}$$

$$\|C\| = \left\|\frac{d}{dt}\left[\left(p - P^{last}\right) I\right]\right\| \qquad \text{Price–execution flow changes match} \tag{F2c}$$

$$\|C\| = \left\|\frac{dp}{dt}\right\| \qquad \text{Price extremum} \tag{F2d}$$

$$\|C\| = \left\|\frac{d^2p}{dt^2}\right\| \qquad dp/dt \text{ extremum} \tag{F2e}$$

The maximization problem (F1a) with the quadratic constraint (F1b) can no longer be reduced to a regular eigenvalue problem such as (25). The solution exists only if $\|C\|$ operator has both: positive and negative eigenvalues. Ideologically the (F1b) constraint facilitates taking into account a typical market practitioner activity: look how the market used to behave in the past at prices near some level. Our previous paper [3] has been mostly devoted to skewness and probability correlation study in the unconstrained $I \to \max$ state $\left|\psi_I^{[IH]}\right\rangle$. The (F1b) constraint allows us, within the framework of a single formalism of constrained optimization, take into account the driving force of the market $I \to \max$ (F1a) and the reaction of the market participants on it (F1b). For mathematical properties and numerical solution of (F1) problem see Appendices F 1 and (F 2 below. Here we assume that the solution does exist, we denote it as $\left|\psi_I^{[\mathcal{M}]}\right\rangle$, and name: the state of price–matching maximal execution flow. The found state $\left|\psi_I^{[\mathcal{M}]}\right\rangle$ (it is just a pure state averaging weight $\left(\psi_I^{[\mathcal{M}]}(x(t))\right)^2 \omega(t)dt$, not even a density matrix (6)) is the state to obtain market directional information.

1. The `IstatesConditional.java` implementation

The optimization problem (F1a) with quadratic constraint (F1b) can be solved using Lagrange multipliers technique:

$$\max_{\psi} \langle\psi \,|\, I \,|\, \psi\rangle - \lambda(\langle\psi \,|\, \psi\rangle - 1) + \mu \langle\psi \,|\, C \,|\, \psi\rangle \tag{F3a}$$

$$1 = \langle\psi \,|\, \psi\rangle \tag{F3b}$$

$$0 = \langle\psi \,|\, C \,|\, \psi\rangle \tag{F3c}$$

$$|0\rangle = |I|\psi\rangle - \lambda\,|\psi\rangle + \mu\,|C|\psi\rangle \tag{F3d}$$

Were the constraint (F3c) to be of a linear type, instead of a quadratic one, the constrained optimization problem (F3a) can be reduced to a regular eigenvalue problem in a transformed basis[10]. However, for the quadratic constraint (F3c), such a one–step transform is not possible, and self–concordant procedure of iterational type is the simplest option:

- For an initial $|\psi\rangle$ find the coefficient α, such that:

$$|b\rangle = |C|\psi\rangle \tag{F4a}$$

$$0 = \langle\psi + \alpha b \,|\, C \,|\, \psi + \alpha b\rangle \tag{F4b}$$

  The (F4b) is a quadratic equation with respect to α, if no real solution exist — iterational process failed. If a success — obtain the solution, satisfying the (F3c) constraint:

$$\left|\widetilde{\psi}\right\rangle = |\psi\rangle + \alpha\,|b\rangle \tag{F5}$$

From two α solutions select the one with the maximal $\left\langle \widetilde{\psi} \middle| I \middle| \widetilde{\psi} \right\rangle \big/ \left\langle \widetilde{\psi} \middle| \widetilde{\psi} \right\rangle$. There are exist several good alternatives to (F4), see `com/polytechnik/utils/FindPsiConstrainedSingleQuadratic0.java` implementation for details.

- Put $\left| \widetilde{\psi} \right\rangle$ to (F3d), then left–multiply it by the vector $\left\langle \widetilde{\psi} \middle| C \right|$, obtain the Lagrange multiplier iteration μ:

$$\mu = -\frac{\left\langle \widetilde{\psi} \middle| C \middle| I \middle| \widetilde{\psi} \right\rangle}{\left\langle \widetilde{\psi} \middle| C \middle| C \middle| \widetilde{\psi} \right\rangle} \tag{F6}$$

- Construct an operator $\|\mathcal{I}\|$ and find all it's eigenvectors:

$$\|\mathcal{I}\| = \|I\| + \mu\|C\| \tag{F7}$$

$$\left| \mathcal{I} \middle| \psi^{[i]} \right\rangle = \lambda^{[i]} \left| \psi^{[i]} \right\rangle \tag{F8}$$

- Among all the $\left|\psi^{[i]}\right\rangle$ found select the $|\psi\rangle$, providing the maximal $\langle \psi \,|\, I \,|\, \psi \rangle$.

- Repeat the process of above for this new $|\psi\rangle$. If a solution exists, iterational procedure converges quickly (typically 5–7 iterations), unless $\|I\|$ and $\|C\|$ operators have several eigenvectors in common[12]. The result of this iterational process is the state of price–matching maximal execution flow $\left|\psi_I^{[\mathcal{M}]}\right\rangle$, the (F1) solution.

The class `com/polytechnik/utils/IstatesConditional.java` implements this procedure. It takes three matrices $\langle Q_j \,|\, Q_k \rangle$, $\langle Q_j \,|\, I \,|\, Q_k \rangle$, $\langle Q_j \,|\, C \,|\, Q_k \rangle$, and basis functions operations class (extending the `com/polytechnik/utils/OrthogonalPolynomialsBasisFunctionsCalculatable.java`), as constructor's arguments. Then it solves generalized eigenvalue problem (25) using the `EVXData.java` class to obtain an initial $|\psi\rangle$ and to reproduce the [1] results. Then ten iterations of above are performed to obtain the solution of (F1): $\left|\psi_I^{[\mathcal{M}]}\right\rangle$ and μ. The fields are:

`.I`	An object of `EVXData.java` type, (25) solution	(F9a)
`.flag_solution_exists`	Whether the (F1) solution exists for the input data	(F9b)

[12] Assume $\|I\|$ and $\|C\|$ operators have identical eigenvectors. Then the (F8) always produce the same eigenvectors, and the maximization problem (F1) is reduced to a linear programming problem relatively the projections squares.

`.psi_M`	$\left\|\psi_I^{[\mathcal{M}]}\right\rangle$ the (F1) solution; equals to 0 on failure	(F9c)
`.LagrangeMultiplier_M`	Lagrange multiplier μ, Eq. (F6)	(F9d)
`.i_M`	$\left\langle\psi_I^{[\mathcal{M}]}\middle\| I \middle\|\psi_I^{[\mathcal{M}]}\right\rangle$ execution flow in the $\left\|\psi_I^{[\mathcal{M}]}\right\rangle$ state	(F9e)
`.wr0_M`	$\left\langle\psi_I^{[\mathcal{M}]}\middle\|\psi_0\right\rangle^2$ a kind of "distance to now"	(F9f)

2. The `IstatesConditionalLocalized.java` implementation

When the global maximum of constrained $I \to \max$ problem is not required, and localized answer with $|\psi\rangle$ in (A1) form is considered as good enough, optimization problem (F1a) with quadratic constraint (F1b) can be easily solved. Substitute (A1) to (F3c) and obtain:

$$I = \frac{\langle\psi_y \,|\, I \,|\, \psi_y\rangle}{\langle\psi_y \,|\, \psi_y\rangle} \xrightarrow[y]{} \max \tag{F10a}$$

$$0 = \sum_{j,k,s,t=0}^{n-1} Q_j(y) G^{-1}_{jk} \langle Q_k \,|\, C \,|\, Q_s\rangle \, G^{-1}_{st} Q_t(y) \tag{F10b}$$

The (F10b) constraint is a polynomial of $2n-2$ degree, it has exactly $2n-2$ root, possibly complex. The classes extending the `com/polytechnik/trading/OrthogonalPolynomialsBasisFunctionsCalculatable.java` (see Appendix G 3 of Ref. [3]) provide an implementation for solving $P(y) = 0$ equation with a $P(y)$ in a given $Q_k(y)$ basis $P(y) = \sum_{m=0}^{2m-2} Q_m(y)$, the (F10b) is a polynomial of this form. Among $2n-2$ roots found select only the real roots, then among them select the state $\left|\psi_I^{[\mathcal{M}]}\right\rangle$, that provides the maximal $\left\langle\psi_I^{[\mathcal{M}]}\middle| I \middle|\psi_I^{[\mathcal{M}]}\right\rangle$. The situation is similar to the one of Appendix G, below, with the difference that $\left|\psi_I^{[\mathcal{M}]}\right\rangle$ is now selected among (A1) states with y from (F10b) real roots ($2n-2$ maximal number), not among n eigenvalues of some operator $\|\mathcal{C}\|$.

The class `com/polytechnik/utils/IstatesConditionalLocalized.java` implements this procedure. It takes three matrices $\langle Q_j \,|\, Q_k\rangle$, $\langle Q_j \,|\, I \,|\, Q_k\rangle$, $\langle Q_j \,|\, C \,|\, Q_k\rangle$, and basis functions operations class (extending the `com/polytechnik/utils/OrthogonalPolynomialsBasisFunctionsCalculatable.java`), as constructor's arguments. Then it solves $P(y) = 0$ polynomial roots problem (F10b) using `com/polytechnik/trading/OrthogonalPolynomialsBasisFunctionsCalculatable<T>:getPolynomialRootsFinderInBasis().findRoots(·)` method to obtain a set of y_m that are the roots of (F10b). Then corresponding $|\psi_{y_m}\rangle$ (A1) are constructed, and the one with the maximal $\langle\psi_{y_m} \,|\, I \,|\, \psi_{y_m}\rangle$ is selected: this is the "localized" $\left|\psi_I^{[\mathcal{M}]}\right\rangle$ solution. The fields are:

$$\begin{array}{lll} \texttt{.psi\_M} & \left|\psi_I^{[\mathcal{M}]}\right\rangle \text{ the (F1) localized solution of (A1) form; equals to 0 on failure} & \text{(F11a)} \\ \texttt{.y\_M} & \text{The "localization" point in (A1) of maximal } I \text{ (F10a), (F10b) root} & \text{(F11b)} \\ \texttt{.i\_M} & \text{The execution flow } \left\langle \psi_I^{[\mathcal{M}]} \middle| I \middle| \psi_I^{[\mathcal{M}]} \right\rangle & \text{(F11c)} \\ \texttt{.n\_roots} & \text{The number of real roots of (F10b)} & \text{(F11d)} \end{array}$$

Appendix G: Directional Information: $I \xrightarrow[\psi]{} \max$ in the States of Constraint Operator $\|\mathcal{C}\|$.

The constrained optimization of the Appendix F 1 above, while been very nice mathematically, does not provide a clear cut answer. There are two reasons: the difficulty to select an operator $\|C\|$ (F2) and the difficulty with (F8) Lagrange multiplier convergence, as $\|I\|$ and $\|C\|$ operators often have common eigenvectors. Consider a different, much more simplistic, constrained optimization approach:

$$I = \frac{\langle \psi \,|\, I \,|\, \psi \rangle}{\langle \psi \,|\, \psi \rangle} \xrightarrow[\psi]{} \max \tag{G1a}$$

$$|\psi\rangle : \text{is subject to being an eigenvector of } |\mathcal{C}|\psi\rangle = \lambda_{\mathcal{C}} \,|\psi\rangle \tag{G1b}$$

Here we also split the market dynamics in two operators: $\|I\|$ (execution flow dynamics) and $\|\mathcal{C}\|$ (price dynamics). But now we consider the $\|I\|$ only in the eigenstates of the operator $\|\mathcal{C}\|$. The operator $\|\mathcal{C}\|$ is selected in a way that it's derivative gives the constraint operator $\|C\|$, thus the $|\psi\rangle$ state of extremal $\|\mathcal{C}\|$ give zero of constraint operator $\|C\|$. Mathematically the problem (G1) is simple: find all n eigenvectors (G1b) of $\|\mathcal{C}\|$ first, then select the one, providing the maximal $\|I\|$ (G1a). The state of price–matching maximal execution flow $\left|\psi_I^{[\mathcal{M}]}\right\rangle$ is now plain (G1b) eigenvector, providing the maximal (G1a). There are a number of choices for the operator $\|\mathcal{C}\|$, selecting the states $|\psi\rangle$:

$$|pI|\psi\rangle = \lambda_{\mathcal{C}} \,|I|\psi\rangle \qquad \text{Price min/max} \tag{G2a}$$

$$|V_1|\psi\rangle = \lambda_{\mathcal{C}} \,|V_0|\psi\rangle \qquad \text{Moving average price (11c) is equal to the price (11a)} \tag{G2b}$$

The optimization with the constraint (G2a) is actually the pure dynamic impact approximation of Ref. [3]: price and execution flow operators are assumed to have the same eigenvectors. The (G2b) states, same as for the aggregated execution flow (C5) below, selects the states with the moving average price equals the price, a typical market practitioner point of attention. The problem (G1) uses the same input data moments (16) as the problem (F1).

Appendix H: The $\left|\psi_I^{[IH]}\right\rangle$ variation approach to positive and negative dI/dt states separation.

The separation of the states with positive and negative dI/dt can be developed based on $\left|\psi_I^{[IH]}\right\rangle$ variation. For example, in the Eq. (F3) of Ref. [3], the variation of I have been considered[13]:

$$I_{\psi+\delta\psi} = \frac{\langle \psi + \delta\psi \,|\, I \,|\, \psi + \delta\psi \rangle}{\langle \psi + \delta\psi \,|\, \psi + \delta\psi \rangle} = D0 + D1 + D2 + \dots \tag{H1}$$

$$D0 = \frac{\langle \psi \,|\, I \,|\, \psi \rangle}{\langle \psi \,|\, \psi \rangle} \tag{H2}$$

$$D1 = 2\left(\frac{\langle \psi \,|\, I \,|\, \delta\psi \rangle}{\langle \psi \,|\, \psi \rangle} - D0\frac{\langle \psi \,|\, \delta\psi \rangle}{\langle \psi \,|\, \psi \rangle}\right) \tag{H3}$$

$$D2 = \frac{\langle \delta\psi \,|\, I \,|\, \delta\psi \rangle}{\langle \psi \,|\, \psi \rangle} - D0\frac{\langle \delta\psi \,|\, \delta\psi \rangle}{\langle \psi \,|\, \psi \rangle} - 2\frac{\langle \psi \,|\, \delta\psi \rangle}{\langle \psi \,|\, \psi \rangle}D1 \tag{H4}$$

With $\delta\psi = -\mathrm{ED}(\psi_I^{[IH]}(x))$ variation (such a variation can be considered as a boundary condition alternative to (E5), (E6), (E7), or (E8)) obtain $\Delta_\psi P$ from the Eq. (31) of Ref. [3]. Any first variation (H3) in a $\left|\psi_I^{[i]}\right\rangle$ state is zero, any second variation (H4) in the state $\left|\psi_I^{[IH]}\right\rangle$ is negative. The first variation of the $\left|\psi_I^{[IH]}\right\rangle$ state can be written as $P(x)$ polynomial average:

$$P(x) = 2\psi_I^{[IH]}(x)\left[\mathrm{ED}(\psi_I^{[IH]}(x)) - \left\langle \psi_I^{[IH]} \,\middle|\, \mathrm{ED}(\psi_I^{[IH]}) \right\rangle \psi_I^{[IH]}(x)\right] \tag{H5}$$

$$D1 = \langle I\, P(x) \rangle = 0 \tag{H6}$$

In [6], we have have proved, that any polynomial $P(x)$ of $2n-2$ degree can be isomorphly mapped to a linear operator of the dimension n, the algorithm is presented in the Appendix A of Ref. [6]:

$$\rho(x,y) = \sum_{i=0}^{n-1} \lambda^{[i]}\psi^{[i]}(x)\psi^{[i]}(y) \tag{H7}$$

[13] See the class `com/polytechnik/utils/RayleighQuotient.java` of provided software, implementing the calculation of 0-th, 1-st, and 2-nd variations of two quadratic forms ratio.

$$P(x) = \rho(x, x) \tag{H8}$$

Then the $D1$ can be presented as a superposition of positive and negative terms:

$$0 = D1 = \sum_{i:\lambda^{[i]}>0} \lambda^{[i]} \left\langle \psi^{[i]} \middle| I \middle| \psi^{[i]} \right\rangle + \sum_{i:\lambda^{[i]}<0} \lambda^{[i]} \left\langle \psi^{[i]} \middle| I \middle| \psi^{[i]} \right\rangle \tag{H9}$$

This way the $P(x)$ average can be split in positive and negative contributions. Despite being a $\left|\psi_I^{[IH]}\right\rangle$ projection, the eigenvalues of (H7) are typically all non–zero, and corresponding density matrix is a mixed state:

$$\|\rho^+\| = \sum_{i:\lambda^{[i]}>0} \left|\psi^{[i]}\right\rangle \lambda^{[i]} \left\langle\psi^{[i]}\right| \tag{H10a}$$

$$\|\rho^-\| = \sum_{i:\lambda^{[i]}<0} \left|\psi^{[i]}\right\rangle \lambda^{[i]} \left\langle\psi^{[i]}\right| \tag{H10b}$$

For computer implementation see the class `com/polytechnik/trading/DIselDM.java` of provided software.

---

[1] V. G. Malyshkin and R. Bakhramov, Mathematical Foundations of Realtime Equity Trading. Liquidity Deficit and Market Dynamics. Automated Trading Machines, arXiv preprint arXiv:1510.05510 10.48550/arXiv.1510.05510 (2015).

[2] V. G. Malyshkin, Market Dynamics. On Supply and Demand Concepts, ArXiv e-prints (2016), `http://arxiv.org/abs/1602.04423`, arXiv:1602.04423.

[3] V. G. Malyshkin, Market Dynamics. On A Muse Of Cash Flow And Liquidity Deficit, ArXiv e-prints 10.48550/arXiv.1709.06759 (2017), arXiv:1709.06759 [q-fin.TR].

[4] V. G. Malyshkin and R. Bakhramov, Market Dynamics vs. Statistics: Limit Order Book Example, ArXiv e-prints 10.48550/arXiv.1603.05313 (2016), arXiv:1603.05313 [q-fin.TR].

[5] G. S. Malyshkin, The comparative efficiency of classical and fast projection algorithms in the resolution of weak hydroacoustic signals (Сравнительная эффективность классических и быстрых проекционных алгоритмов при разрешении слабых гидроакустических сигналов), Acoustical Physics **63**, 216 (2017), doi:10.1134/S1063771017020099 (eng) ; doi:10.7868/S0320791917020095 (рус).

[6] V. G. Malyshkin, On Lebesgue Integral Quadrature, arXiv preprint arXiv:1807.06007 10.48550/arXiv.1807.06007 (2018).

[7] A. V. Bobyl, V. V. Davydov, A. G. Zabrodskii, N. R. Kostik, V. G. Malyshkin, O. V. Novikova, D. M. Urishov, and E. A. Yusupova, The Spectral approach to timeserie bursts analysis (Спектральный подход к анализу всплесков временной последовательности), ISSN 0131-5226.Теоретический и научно-практический журнал. ИАЭП. , 77 (2018).

[8] V. G. Malyshkin, The code for polynomials calculation (2014), `http://www.ioffe.ru/LNEPS/malyshkin/code.html` and an alternative location.

[9] A. V. Bobyl, A. G. Zabrodskii, M. E. Kompan, V. G. Malyshkin, O. V. Novikova, E. E. Terukova, and D. V. Agafonov, Generalized Radon–Nikodym Spectral Approach. Application to Relaxation Dynamics Study., ArXiv e-prints 10.2139/ssrn.3229466 (2016), arXiv:1611.07386 [math.NA].

[10] G. H. Golub, Some modified matrix eigenvalue problems, Siam Review **15**, 318 (1973).