Optimal split of orders across liquidity pools: a stochastic algorithm approach

Sophie Laruelle, Charles-Albert Lehalle, Gilles Pagès

Optimal split of orders across liquidity pools: a stochastic algorithm approach

Sophie Laruelle * Charles-Albert Lehalle † Gilles Pagès ‡

First Version : 6th October 2009
This Version : May 27, 2010

Abstract

Evolutions of the trading landscape lead to the capability to exchange the same financial instrument on different venues. Because of liquidity issues, the trading firms split large orders across several trading destinations to optimize their execution. To solve this problem we devised two stochastic recursive learning procedures which adjust the proportions of the order to be sent to the different venues, one based on an optimization principle, the other on some reinforcement ideas. Both procedures are investigated from a theoretical point of view: we prove *a.s.* convergence of the optimization algorithm under some light ergodic (or "averaging") assumption on the input data process. No Markov property is needed. When the inputs are i.i.d. we show that the convergence rate is ruled by a Central Limit Theorem. Finally, the mutual performances of both algorithms are compared on simulated and real data with respect to an "oracle" strategy devised by an "insider" who *a priori* knows the executed quantities by every venues.

Keywords *Asset allocation, Stochastic Lagrangian algorithm, reinforcement principle, monotone dynamic system.*

2001 AMS classification: 62L20, secondary: 91B32, 62P05

1 Introduction

The trading landscape have seen a large number of evolutions following two regulations: Reg NMS in the US and MiFID in Europe. One of their consequences is the capability to exchange the same financial instrument on different trading venues. New trading destinations appeared to complement the trading capability of primary markets as the NASDAQ and the NYSE in the US, or EURONEXT, the London Stock Exchange and Xetra in Europe. Such alternative venues are called "Electronic Communication Network" (ECN) in the US and Multilateral Trading Facilities

*Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: sophie.laruelle@upmc.fr

†Head of Quantitative Research, Crédit Agricole Cheuvreux, CALYON group ; 9 quai Paul Doumer, 92920 Paris La Défense. E-mail: clehalle@cheuvreux.com

‡Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: gilles.pages@upmc.fr

(MTF) in Europe. Each trading venue differentiates from the others at any time because of the fees or rebate it demands to trade and the liquidity it offers.

As the concerns about consuming liquidity increased with the financial crisis, trading firms use Smart Order Routers (SOR) as a key element in the process of optimizing their execution of large orders. Such devices are dedicated to split orders across trading destinations as a complement to the temporal slicing coming from the well known balance between the need to trade rapidly (to minimize market risk) and trading slow (to avoid market impact).

If the temporal slicing has been studied since the end of the nineties [1] with recent advances to adapt it to sophisticated investment strategies [19], this kind of spatial slicing (across trading destinations) has been mainly studied by economists from the point of view of its global market efficiency [8] rather than from one investor's point of view.

The complexity of spreading an order between N trading destinations comes from the fact that you never knows the quantity D_i available on the i^{th} trading venue to execute your order of size V during a time interval δt at your given price. If the fraction $r_i V$ of your order that you sent to the i^{th} liquidity pool is higher than D_i: you will loose time and may loose opportunity to execute $r_i V - D_i$ in an another pool; on another hand if $r_i V$ is lower than D_i: you will loose money if this pool fees are cheap, and an opportunity to execute more quantity here. The only way to optimize such a split on real time is to adjust on the fly the proportions $(r_i)_i$ according to the result of your previous executions.

This paper is an in depth analysis of the optimal split of orders. The illustrations and most of the vocabulary come from the "Dark pool" case, where the price S is not chosen by the trader (it is the market "mid point" price) and the answer of the pool is immediate (i.e. $\delta t = 0$). Dark pools are MTFs that do not publish pre-trade informations, so an efficient use of the results of the previous executions (namely the realizations of the $\min(D_i^t, r_i^t V^t)$ for any i and all t in the past) is crucial. The results exposed here solve the problem of simultaneously splitting orders and using the information coming back from the pools to adjust the proportions to send for the next order, according to a criteria linked to the overall quantity executed (*i.e.* a linear combination of the $\min(D_i, r_i V)$).

The resulting trading strategy (which optimality is proven here) can be considered as an extension of the one conjectured by Almgren in [2]. It may also be related to the class of multi-armed bandit recursive learning procedures, recently brought back to light in several papers (see [14, 22], [15, 16]; which in turn belongs to the wide family of "recursive stochastic algorithms" also known as "stochastic approximation" and extensively investigated in the applied probability literature (see [13], [3], [7], etc)).

In fact, we introduce two learning algorithms one based on an optimization under constraints principle and a second algorithm based on a reinforcement principle for which we establish the existence of an equilibrium. We extensively investigate the first one, considering successively the classical – although unrealistic – case where the inputs (requests, answers) are i.i.d. and a setting in which the input only share some averaging properties. In the i.i.d. setting we establish *a.s.* convergence of the procedure and a Central Limit Theorem relying on classical results from Stochastic Approximation Theory. By *averaging setting* (also referred as *ergodic setting*), we mean that the inputs of the procedure has *a.s.* an averaging property with respect to a distribution ν at a given rate, say $n^{-\beta}$, $\beta > 0$, for a wide enough class of Borel functions. Typically, in our problem, these inputs are the successive $N+1$-tuples $(V^n, D_i^n, i = 1, \ldots, N)$, $n \geq 1$. Typically, if we denote this

input sequence inputs by $(Y_n)_{n\geq 1}$, we will assume that, for every $f \in \mathcal{V}_{\beta,p}$,

$$\frac{1}{n}\sum_{k=1}^{n} f(Y_k) - \int_{\mathbb{R}_+^{N+1}} f d\nu = O(n^{-\beta}) \quad \mathbb{P}\text{-}a.s \text{ and in } L^p(\mathbb{P}).$$

Usually, $\mathcal{V}_{\beta,p}$ is supposed to contain at least bounded continuous function $g : \mathbb{R}_+^{N+1} \to \mathbb{R}$ and subsequently all bounded ν-$a.s.$ continuous functions. This will be enough for our purpose in this paper (Stochastic approximation in this general framework is investigate in [17]). But the key point to be noted here is that *no Markov assumption is needed on this input sequence* $(Y_n)_{n\geq 1}$. These assumptions are hopefully light enough to be satisfied by real data since it can be seen as a kind of "light" ergodicity at a given rate. In a Markovian framework it could be related to the notion of "stability" in the literature, see [7].

Thus, this setting includes stationary α-mixing processes (satisfying an Ibragimov condition) like those investigated in [6] (in [5] weaker dependence assumptions are made in the chapter devoted to stochastic approximation but the perturbation is supposed to be additive and non causal which is not at all the case in our problem). As concerns the second procedure for which no Lyapunov function seems to be (easily) made available, we establish the existence of an equilibrium and show the ODE related to the algorithm is a competitive system in the terminology of monotonous differential systems extensively studied by Hirsch et al. (see *e.g.* [12]). The behaviour of such competitive systems is known to be the most challenging, even when the equilibrium point is unique (which is not the case here).

Both procedures are compared in the final section, using simulated and real data. Further numerical tests and applications are ongoing works in CA Cheuvreux.

The paper is organized as follows: in Section 2, we make precise the modeling of splitting orders among several venues in the framework of *Dark pools*, first in static then in a dynamic way. This leads to an optimization problem under constraints. In Section 3, we study the execution function of one dark pool and introduce the recursive stochasic algorithm resulting from the optimization problem. In Section 4 we analyze in depth this algorithm ($a.s.$ convergence and weak rate) when the "innovations" (data related to the orders, the executed quantities and the market price) are assumed i.i.d. In Section 5 we extend the $a.s.$ convergence result to a more realistic framework where these innovations are supposed to share some appropriate averaging properties (*e.g.* satisfied by α-mixing processes satisfying Ibragimov's condition). Section 6 is devoted to the second learning procedure, based this time on reinforcement principle, introduced in [4]. We make a connexion with the theory of (competitive) monotonous dynamical systems. Finally, in Section 7, we present several simulations results on simulated and real data to evaluate the performances of both procedures with respect to an "oracle" strategy of an "insider" who could know *a priori* the executed quantities by every dark pool.

NOTATIONS: • For every $N \geq 1$, set $\mathcal{I}_N := \{1, 2, \ldots, N\}$, $\mathcal{P}_N := \{r = (r_i)_{1\leq i\leq n} \in \mathbb{R}_+^N \mid \sum_{i=1}^N r_i = 1\}$. Let $\mathbf{1}^\perp := \{u \in \mathbb{R}^N \mid \sum_{i\in\mathcal{I}_N} u_i = 0\}$.
• δ_{ij} denotes the Kronecker symbol.
• $\langle . | . \rangle$ denotes the canonical inner product on $\mathbb{R}^d$ and $|\,.\,|$ the derived Euclidean norm.
• $int(A)$ denotes the interior of a subset A of $\mathbb{R}^d$.
• δ_a denotes the Dirac mass at $a \in \mathbb{R}^d$.

2 A simple model for the execution of orders by dark pools

2.1 Static modelling

As mentioned in the introduction, we will focus in this paper on the splitting order problem in the case of (competing) *dark pools*. The execution policy of a dark pool differs from a primary market: thus a dark pool proposes bid/ask prices with no guarantee of executed quantity at the occasion of an over the counter transaction. Usually its bid price is lower than the bid price offered on the regular market (and the ask price is higher). Let us temporarily focus on a buying order sent to several dark pools. One can model the impact of the existence of N dark pools ($N \geq 2$) on a given transaction as follows: let $V > 0$ be the random volume to be executed and let $\theta_i \in (0,1)$ be the *discount factor* proposed by the dark pool $i \in \{1,\ldots,N\}$. We will make the assumption that this discount factor is deterministic or at least known prior to the execution. Let r_i denote the percentage of V sent to the dark pool i for execution and let $D_i \geq 0$ be the quantity of securities that can be delivered (or made available) by the dark pool i at price $\theta_i S$ where S denotes the bid price on the primary market (this is clearly an approximation since on the primary market, the order will be decomposed into slices executed at higher and higher prices following the order book). The rest of the order has to be executed on the primary market, at price S. Then the cost C of the executed order is given by

$$\begin{aligned} C &= S\sum_{i=1}^{N}\theta_i \min(r_iV, D_i) + S\left(V - \sum_{i=1}^{N}\min(r_iV, D_i)\right) \\ &= S\left(V - \sum_{i=1}^{N}\rho_i \min(r_iV, D_i)\right) \end{aligned}$$

where $\rho_i = 1 - \theta_i > 0$, $i = 1,\ldots,N$. At this stage, one may wish to minimize the mean execution cost C, *given the price* S. This amounts to solving the following (conditional) maximization problem

$$\max\left\{\sum_{i=1}^{N}\rho_i\,\mathbb{E}\left(\min(r_iV, D_i)\,|\,S\right),\ r \in \mathcal{P}_{_N}\right\}. \tag{2.1}$$

However, none of the agents being insiders, they do not know the price S when the agent decides to buy the security and when the dark pools answer to their request. This means that one may assume that $(V, D_1, \ldots, D_n)$ and S are independent so that the maximization problem finally reads

$$\max\left\{\sum_{i=1}^{N}\rho_i\mathbb{E}\left(\min(r_iV, D_i)\right),\ r \in \mathcal{P}_{_N}\right\} \tag{2.2}$$

where we assume that all the random variables $\min(V, D_1)$, $\ldots$, $\min(V, D_{_N})$ are integrable (otherwise the problem is meaningless).

An alternative choice could be to include the price S of the security into the optimization which leads to the mean maximization problem

$$\max\left\{\sum_{i=1}^{N}\rho_i\,\mathbb{E}\left(S\min(r_iV, D_i)\right),\ r \in \mathcal{P}_{_N}\right\} \tag{2.3}$$

(with the appropriate integrability assumption to make the problem consistent). It is then convenient to *include the price S into both random variables V and D_i* by considering $\widetilde{V} := V\,S$ and $\widetilde{D}_i := D_i S$ instead of V and D_i which leads again to the maximization problem (2.2).

If one considers symmetrically a selling order to be executed, the dark pool is supposed to propose a higher ask price $\theta_i S$, $\theta_i > 1$, than the order book. The seller aims at maximizing the execution global (mean) price of the transaction. This yields to the same formal optimization problem, this time with $\rho_i = \theta_i - 1$, $i = 1, \ldots, N$.

All these considerations lead us to focus on the abstract optimal allocation problem (2.2) which explains why the price variable S will no longer appear explicitly in what follows.

2.2 The dynamical aspect

In practice, there is no *a priori* assumption – or information available – on the joint distribution of $(V, D_1, \ldots, D_N)$ under $\mathbb{P}$. So the only reasonable way to provide a procedure to solve this allocation problem is to devise an *on-line learning algorithm based on historical data*, namely the results of former transactions with the dark pools on this security executed in the past. This underlines that our agent dealing with the dark pools is a financial institution like a bank, a broker or possibly a large investor which often – that means at least daily – faces some large scale execution problems on the same securities.

This means that we will have to make some assumptions on the dynamics of these transactions *i.e.* on the data input sequence $(V^n, D_1^n, \ldots, D_N^n)_{n\geq 1}$ supposed to be defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Our basic assumption on the sequence $(D_i^n, V^n, i = 1, \leq, N)_{n\geq 1}$ is of statistical – or ergodic – nature: we ask this sequence to be ν-averaging ($a.s.$ and in $L^p(\mathbb{P})$), at least on bounded continuous functions, where ν is a distribution on $(\mathbb{R}_+^{N+1}, \mathcal{B}or(\mathbb{R}_+^{N+1}))$. This leads to the following formal assumption:

$$(ERG)_\nu \equiv \begin{cases} (i) & \text{the sequence } (V^n, D_i^n, i = 1, \ldots, N)_{n\geq 1} \text{ is averaging } i.e. \\ & \mathbb{P}\text{-}a.s. \quad \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \delta_{(V^k, D_1^k, \ldots, D_N^k)} \overset{(\mathbb{R}_+^{N+1})}{\Longrightarrow} \nu, \\ (ii) & \sup_n \mathbb{E}(V^n)^2 < +\infty. \end{cases}$$

where $\overset{(\mathbb{R}_+^{N+1})}{\Longrightarrow}$ denotes the weak convergence of probability measures on $\mathbb{R}_+^{N+1}$. For convenience, we will denote $(V, D_1, \ldots, D_N)$ the canonical random vector on $\mathbb{R}_+^{N+1}$ so that we can write $\nu = \mathcal{L}(V, D_1, \ldots, D_N)$.

Assumption (ii) on the marginal distribution of the sequence $(V^n)_{n\geq 1}$ is mainly technical. In fact standard arguments from weak convergence theory show that combining (i) and (ii) implies

$$\frac{1}{n}\sum_{k=1}^{n} V^k \longrightarrow \mathbb{E}\,V \quad \text{as} \quad n \to \infty$$

($\sup_n \mathbb{E}(V^n)^{1+\varepsilon} < +\infty$ would be enough). An important subcase is the the (IID) setting

$$(IID) \equiv \begin{cases} (i) & \text{the sequence } (V^n, D_1^n, \ldots, D_N^n)_{n\geq 1} \text{ is i.i.d. with distribution } \nu = \mathcal{L}(V, D_1, \ldots, D_N), \\ (ii) & V \in L^2(\mathbb{P}). \end{cases}$$

This more restrictive assumption is undoubtedly less realistic from a modeling point of view but it remains acceptable as a first approximation. It is the most common framework to apply the standard Stochastic Approximation machinery (*a.s.* convergence, asymptotically normal fluctuations, etc). So, its interest may be considered at least as pedagogical. The (*ERG*) setting is slightly more demanding in terms of assumptions and needs more specific methods of proof. It will be investigated as a second step, using some recent results established in [18] which are well suited to the specificities of our problem (in particular we will not need to assume the existence of a solution to the Poisson equation related to the procedure like in the reference book [3]).

3 Optimal allocation: a stochastic Lagrangian algorithm

3.1 The mean execution function of a dark pool

In view of the modeling section, we need to briefly describe the precise behaviour of the mean execution function $\varphi : [0,1] \to \mathbb{R}_+$ of a single dark pool.

Let (V, D) be an $\mathbb{R}_+^2$-valued random vector defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ representing the *global volume* to be executed and the *deliverable quantity* (by the dark pool) respectively. Throughout this paper we will assume the following consistency assumption

$$V > 0 \quad \mathbb{P}\text{-}a.s. \quad \text{and} \quad \mathbb{P}(D > 0) > 0. \tag{3.1}$$

The *a.s.* positivity of V means that we only consider true orders. The fact that D is not identically 0 means that the dark pool does exist in practice. The "rebate" coefficient ρ is specific to the dark pool.

To define in a consistent way the mean execution function of a dark pool we only need to assume that $V \in L^1(\mathbb{P})$ (although more stringent integrability assumptions are made throughout the paper).

Here the *mean execution function* $\varphi : [0,1] \to \mathbb{R}_+$ of the dark pool is defined by

$$\forall\, r \in [0,1], \qquad \varphi(r) = \rho\, \mathbb{E}(\min(rV, D)) \tag{3.2}$$

where $\rho > 0$. The function φ is finite, non-identically 0. It is clearly a concave non-decreasing bounded function. Furthermore, one easily checks that its right and left derivatives are given at every $r \in [0,1]$ by

$$\varphi_l'(r) = \rho\, \mathbb{E}\left(\mathbf{1}_{\{rV \le D\}} V\right) \qquad \text{and} \qquad \varphi_r'(r) = \rho\, \mathbb{E}\left(\mathbf{1}_{\{rV < D\}} V\right). \tag{3.3}$$

In particular,

$$\varphi'(0) = \rho\, \mathbb{E}(V \mathbf{1}_{\{D>0\}}) > 0$$

and if

$$\textit{the (right continuous) distribution function of } \tfrac{D}{V} \textit{ is continuous on } \mathbb{R}_+, \tag{3.4}$$

then

$$\varphi \textit{ is everywhere differentiable on the unit interval } [0,1] \textit{ with } \varphi' = \varphi_l' \textit{ on } (0,1].$$

Assumption (3.4) means that the distribution of $\frac{D}{V}$ has no atom except possibly at 0. It can be interpreted as the fact that a dark pool has no "quantized" answer to an order.

More general models of execution functions in which the rebate ρ and the deliverable quantity D may depend upon the quantity to be executed rV are briefly discussed further on.

3.2 Design of the stochastic Lagrangian algorithm

Let V be the quantity to be executed by N dark pools. For every dark pool $i \in \mathcal{I}_N$ the available quantity D_i is defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ as V. We assume that all couples (V, D_i) satisfy the consistency assumption (3.1).

To each dark pool $i \in \mathcal{I}_N$ is attached a (bounded concave) mean execution function φ_i of type (3.2), introduced in Section 2.1, or (8.1), (8.3) studied in Section 8.

Then for every $r = (r_1, \ldots, r_N) \in \mathcal{P}_N$,

$$\Phi(r_1, \ldots, r_N) := \sum_{i=1}^{N} \varphi_i(r_i). \tag{3.5}$$

In order to design the algorithm we will need to extend the mean execution function φ (whatever its form is) as a concave function on the whole real line by setting

$$\varphi(r) = \left(r - \frac{r^2}{2}\right)\varphi'(0) \quad \text{if } r < 0 \quad \text{and} \quad \varphi(r) = \varphi(1) + \varphi'(1)\log r \quad \text{if } r > 1. \tag{3.6}$$

Based on the extension of the functions φ_i defined by (3.6), we can formally extend Φ on the whole affine hyperplane spanned by $\mathcal{P}_N$ *i.e.*

$$\mathcal{H}_N := \{r \in \mathbb{R}^N \mid \sum_i r_i = 1\}.$$

As announced, we aim at solving the following maximization problem

$$\max_{r \in \mathcal{P}_N} \Phi(r)$$

but we will also have to deal for algorithmic purpose with the same maximization problem when r runs over $\mathcal{H}_N$.

Before stating a rigorous result, let us a have a look at a Lagrangian approach that only takes into account the affine constraint that is $\max_r \Phi(r) - \lambda \sum_i r_i$. Straightforward formal computations suggest that

$$r^* \in \operatorname{argmax}_{\mathcal{P}_N} \Phi \text{ iff } \varphi_i'(r_i^*) \text{ is constant when } i \text{ runs over } \mathcal{I}_N$$

or equivalently if

$$\forall\, i \in \mathcal{I}_N, \qquad \varphi_i'(r_i^*) = \frac{1}{N}\sum_{j=1}^{N} \varphi_j'(r_j^*). \tag{3.7}$$

In fact this statement is not correct in full generality because the Lagrangian method does not provide a necessary and sufficient condition for a point to be a maximum of a (concave) function; thus, it does not take into account the case where Φ reaches its maximum on the boundary $\partial\mathcal{P}_N$ where the above condition on the derivatives may fail. So, an additional assumption is necessary to make it true as established in the proposition below.

Proposition 3.1 *Assume that (V, D_i) satisfies the consistency assumptions (3.1) and (3.4) for every $i \in \mathcal{I}_N$.*

(a) Assume that the functions φ_i defined by (3.2) satisfy the following assumption

$$(\mathcal{C}) \quad \equiv \quad \min_{i\in\mathcal{I}_N} \varphi_i'(0) \geq \max_{i\in\mathcal{I}_N} \varphi_i'\left(\frac{1}{N-1}\right).$$

Then $\mathrm{argmax}_{\mathcal{P}_N}\Phi$ *is a compact convex set and*

$$\mathrm{argmax}_{\mathcal{P}_N}\Phi = \{r \in \mathcal{P}_N,\, |\, \varphi_i'(r_i) = \varphi_1'(r_1),\ i = 1, \ldots, N\}.$$

Furthermore $\mathrm{argmax}_{\mathcal{H}_N}\Phi = \mathrm{argmax}_{\mathcal{P}_N}\Phi$.

(b) If the functions φ_i satisfy the slightly more stringent assumption,

$$(\mathcal{C}_<) \quad \equiv \quad \min_{i\in\mathcal{I}_N} \varphi_i'(0) > \max_{i\in\mathcal{I}_N} \varphi_i'\left(\frac{1}{N-1}\right).$$

then

$$\mathrm{argmax}_{\mathcal{H}_N}\Phi = \mathrm{argmax}_{\mathcal{P}_N}\Phi \subset int(\mathcal{P}_N).$$

Remarks. • If $N = 2$, one checks that Assumption $(\mathcal{C})$ is also necessary to derive the conclusion of item (a).

• As a by-product of the proof below we have the following more precise result on the optimal allocation r^*: if $r^* \in \mathrm{argmax}_{\mathcal{P}_N}$ and $\mathcal{I}_0(r^*) := \{i \in \mathcal{I}_N \,|\, r_i^* = 0\}$, then

$$\max_{i\in\mathcal{I}_0(r^*)} \varphi_i'(0) \leq \min_{i\in\mathcal{I}_0(r^*)^c} \varphi_i'(0).$$

Interpretation and comments: • In the case of a "regular" mean execution function, Assumption $(\mathcal{C})$ is a kind of *homogeneity assumption on the rebates* made by the involved dark pools. If we assume that $\mathbb{P}(D_i = 0) = 0$ for every $i \in \mathcal{I}_N$ (all dark pools buy or sell at least one security with the announced rebate), then $(\mathcal{C})$ reads

$$\min_{i\in\mathcal{I}_N} \rho_i \geq \max_{i\in\mathcal{I}_N} \left(\rho_i \frac{\mathbb{E}\, V \mathbf{1}_{\{\frac{V}{N-1} \leq D_i\}}}{\mathbb{E}\, V}\right)$$

since $\varphi_i'(0) = \rho_i\, \mathbb{E}\, V$. In particular,

Assumption $(\mathcal{C})$ is always satisfied when all the ρ_i's are equal

(all dark pools propose the same rebates).

• Assumption $(\mathcal{C})$ is in fact our main assumption in terms of modeling. It may look somewhat difficult to satisfy when the rebates are not equal. But the crucial fact in order to preserve the generality of what follows is that it contains *no assumption about the dependence between the volume V and the "answers" D_i from the dark pools.*

Proof. (a) The function Φ is continuous on a compact set hence $\mathrm{argmax}_{\mathcal{P}_N}\Phi$ is not empty. Let $r\in\mathrm{argmax}_{\mathcal{P}_N}\Phi$ and $\mathcal{I}_0(r):=\{i\in\mathcal{I}_N\,|\,r_i=0\}$. Clearly $\mathcal{I}_0(r)\neq\mathcal{I}_N$ so that $\mathrm{card}\,\mathcal{I}_0(r)\leq N-1$. Let $u\in\mathbf{1}^\perp$ such that $u_i>0$, $i\in\mathcal{I}_0(r)$. Then $t\mapsto\Phi(r+tu)$ defined on the right neighbourhood of 0 reaches its maximum at 0 so that its derivative at 0 is non-positive. Specifying the vector u yields

$$\forall\, i\in\mathcal{I}_0(r),\ \forall\, j\in\mathcal{I}_0(r)^c,\quad \varphi_i'(0)\leq\varphi_j'(r_j).$$

Now if $u\in\mathbf{1}^\perp$ with $u_i=0$, $i\in\mathcal{I}_0(r)$, then the $t\mapsto\Phi(r+tu)$ is defined on a neighbourhood of 0 and reaches its maximum at $t=0$ so that its derivative is 0 at 0; specifying the vector u yields

$$\forall\, i,\ j\in\mathcal{I}_0(r)^c,\quad \varphi_i'(r_i)=\varphi_j'(r_j).$$

Now, there exists at least one index $i_1\in\mathcal{I}_0(r)^c$ such that $r_{i_1}\geq\frac{1}{|\mathcal{I}_0(r)^c|}\geq\frac{1}{N-1}$. Hence $\varphi_{i_1}'(r_{i_1})\leq\varphi_{i_1}'(\frac{1}{N-1})$ which implies in turn that for every $i_0\in\mathcal{I}_0(r)$, $\varphi_{i_0}'(0)\leq\varphi_{i_1}'(r_{i_1})\leq\varphi_{i_1}'(\frac{1}{N-1})$. Finally Assumption $(\mathcal{C})$ implies that these inequalities hold as equalities so that

$$\forall\, i\in\mathcal{I}_N,\quad \varphi_i'(r_i)=\varphi_1'(r_1).$$

Conversely, let $r\in\mathcal{P}_N$ satisfying the above equalities. Then, for every $r'\in\mathcal{P}_N$, the function $t\mapsto\Phi(tr'+(1-t)r)$ is concave on $[0,1]$ with a right derivative equal to 0 at $t=0$. So it is maximum at $t=0$ *i.e.* $\Phi(r)\geq\Phi(r')$.

Now we pass to the the maximization over $\mathcal{H}_N$. Since it is an affine space and Φ is concave, it is clear, *e.g.* by considering Φ as a function of $(r_1,\ldots,r_{N-1})$, that

$$\mathrm{argmax}_{\mathcal{H}_N}\Phi=\{r\in\mathcal{H}_N,\,|\,\varphi_i'(r_i)=\varphi_1'(r_1),\ i=1,\ldots,N\}$$

(which is non-empty since it contains at least $\mathrm{argmax}_{\mathcal{P}_N}$). Now let $r\in\mathcal{H}_N\backslash\mathcal{P}_N$. Assume there exists $i_0\in\mathcal{I}_N$ such that $r_{i_0}<0$. Then there always exists an index $i_1\in\mathcal{I}_N$ such that $r_{i_1}\geq\frac{1-r_{i_0}}{N-1}>\frac{1}{N-1}$. Consequently

$$\varphi_{i_0}'(r_{i_0})=(1-r_{i_0})\varphi_{i_0}'(0)>\varphi_{i_0}'(0)\geq\min_i\varphi_i'(0)\geq\max_i\varphi_i'\left(\frac{1}{N-1}\right)\geq\varphi_{i_1}'\left(\frac{1}{N-1}\right)\geq\varphi_{i_1}'(r_{i_1})$$

which contradicts the equality of these two derivatives. Consequently all r_i's are non-negative so that $r\in\mathcal{P}_N$.

(b) If $\mathcal{C}_<$ holds, the above proof shows that $\mathcal{I}_0(r)=\emptyset$ so that $\mathrm{argmax}_{\mathcal{P}_N}\Phi_N\subset int(\mathcal{P}_N)$. □

3.3 Design of the stochastic algorithm

Now we are in position to devise the stochastic algorithm for the optimal allocation among the dark pools, taking advantage of the characterization of $\mathrm{argmax}_{\mathcal{P}_N}\Phi$. In fact we will simply use the obvious remark that N numbers $a_1,\ldots,a_N$ are equal if and only if they are all equal to their arithmetic mean $\frac{a_1+\cdots+a_N}{N}$.

We consider the mean execution function as defined by (3.2). We assume from now on that the continuity assumption (3.4) holds so that the representation (3.3) of its derivative can be taken as its right or its left derivative on $(0,1]$ (and its right derivative only at 0).

Using this representation (3.3) for all the derivatives φ_i' yields that, if Assumption $(\mathcal{C})$ is satisfied, then $\mathrm{argmax}_{\mathcal{H}_N}\Phi = \mathrm{argmax}_{\mathcal{P}_N}\Phi$ and

$$r^* \in \mathrm{argmax}_{\mathcal{P}_N}\Phi \Longleftrightarrow \forall\, i \in \{1,\ldots,N\},\ \mathbb{E}\left(V\left(\rho_i \mathbf{1}_{\{r_i^* V \le D_i\}} - \frac{1}{N}\sum_{j=1}^{N}\rho_j \mathbf{1}_{\{r_j^* V \le D_j\}}\right)\right) = 0.$$

However, the set $\mathcal{P}_N$ is not stable for the "naive" zero search algorithm naturally derived from the above characterization, we are led to devise the procedure on the hyperplane $\mathcal{H}_N$.

Consequently, this leads to devise the following zero search procedure

$$r^n = r^{n-1} + \gamma_n H(r^{n-1}, V^n, D_1^n, \ldots, D_N^n),\ n \ge 1, \quad r^0 \in \mathcal{P}_N, \tag{3.8}$$

where, for every $i \in \mathcal{I}_N$, every $r \in \mathcal{H}_N$, every $V > 0$ and every $D_1, \ldots, D_N \ge 0$,

$$\begin{aligned} H_i(r, V, D_1, \ldots, D_N) &= V\Big(\rho_i \mathbf{1}_{\{r_i V \le D_i\}\cap\{r_i \in [0,1]\}} - \frac{1}{N}\sum_{j=1}^{N}\rho_j \mathbf{1}_{\{r_j V \le D_j\}\cap\{r_j \in [0,1]\}} \\ &\qquad + R_i(r, V, D_1, \ldots, D_N)\Big) \end{aligned} \tag{3.9}$$

and the "innovation" $(V^n, D_1^n, \ldots, D_N^n)_{n\ge1}$ is a sequence of random vectors with non negative components such that, for every $n \ge 1$, $(V^n, D_i^n, i = 1, \le, N) \stackrel{d}{=} (V, D_i, i = 1, \le, N)$ and the remainder terms R_i have a mean-reverting effect to pull back the algorithm into $\mathcal{P}_N$. They are designed from the extension (3.6) of the derivative functions φ_i' outside the unit interval $[0,1]$; to be precise, for every $i \in \mathcal{I}_N$,

$$\begin{aligned} R_i(r, V, D_1, \ldots, D_N) &= \rho_i\left((1-r_i)\mathbf{1}_{\{D_i>0\}\cap\{r_i<0\}} + \frac{1}{r_i}\mathbf{1}_{\{V\le D_i\}\cap\{r_i>1\}}\right) \\ &\quad - \frac{1}{N}\sum_{j=1}^{N}\rho_j\left((1-r_j)\mathbf{1}_{\{D_j>0\}\cap\{r_j<0\}} + \frac{1}{r_j}\mathbf{1}_{\{V\le D_j\}\cap\{r_j>1\}}\right). \end{aligned}$$

3.4 Interpretation and implementability of the procedure

▷ Implementability. The vector $(r_i^n)_{1\le i\le N}$ in (3.8) represents the dispatching of the orders among the N dark pools to be sent at time $n+1$ by the investor. It is computed at time n. On the other hand V^n represents the volume to be executed (or its monetary value if one keeps in mind that we "plugged" the price into the volume) and the D_i^n the "answer" of dark pool i, still at time n.

The point is that the investor does have no access to the quantities D_i^n. However, he/she knows what he/she receives from dark pool i, *i.e.* $\min(D_i^n, r_i^{n-1}V^n)$. As a consequence, the investor has access to the event

$$\{\min(D_i^n, r_i^{n-1}V^n) = r_i^{n-1}V^n\} = \{r_i^{n-1}V^n \le D_i^n\}$$

which in turn makes possible the updating of the procedure although he/she has no access to the true value of D_i^n.

So, except for edge effects outside the simplex $\mathcal{P}_N$, the procedure as set can be implemented on real data.

▷ Interpretation. As long as r is a true allocation vector, *i.e.* lies in the simplex $\mathcal{P}_N$, the interpretation of the procedure is the following: assume first that all the factors ρ_i are equal (to 1). Then the dark pools which fully executed the sent orders ($r_i V \leq D_i$) are rewarded proportionally to the numbers of dark pools which *did not fully executed* the request they received. Symmetrically, the dark pools which could not execute the whole request are penalized proportionally to the number of dark pools which satisfied the request.

Thus, if only one dark pool, say dark pool 1, fully executes the request at time n, its pourcentage will be increased for the request at time $n+1$ by $\gamma_n(1-\frac{1}{N})V^n$ *i.e.* it will asked to execute $r_1^n = r_1^{n-1} + \gamma_n(1-\frac{1}{N})V^n$ % of the total order V^{n+1}. The other $N-1$ dark pools will be penalized symmetrically: the pourcentage r_i^n of the total request V^{n+1} each of them will receive at time $n+1$ will be reduced by $\gamma_n \frac{1}{N} V^n$.

If k dark pools totally execute their request at time n and the $N-k$ other fail, the pourcentages of V^{n+1} that the "successful" dark pools will receive for execution at time $n+1$ will be increased by $\gamma_n(1-\frac{k}{N})V^n$, each of the $N-k$ "failing dark pools" being reduced by $\gamma_n \frac{k}{N} V^n$.

If no dark pool was able to satisfy their received request at time n, none will be penalized and if all dark pools fully execute the received orders, none will be rewarded.

In short, the dark pools are rewarded or penalized by comparing their mutual performances. When the "attractivity" coefficents ρ_i are not equal, the reasoning is the same but weighted by these attractivities.

▷ Practical implementation. One may force the above procedure to stay in the simplex $\mathcal{P}_N$ by projecting, once updated, the procedure on $\mathcal{P}_N$ each time it exits the simplex. This amounts to replace the possibly negative r_i by 0, the $r_i > 1$ by 1 and to renormalize the vector r by dividing it by the sum of its terms.

Furhermore, to avoid that the algorithm leaves too often the simplex, one may simply normalize the step γ_n by considering the predictable step

$$\tilde{\gamma}_n = \gamma_n \times \frac{n-1}{V^1 + \cdots + V^{n-1}} \approx \frac{\gamma_n}{\mathbb{E}V}.$$

4 The (IID) setting: $a.s$ convergence and CLT

Theorem 4.1 *Assume that (V, D) satisfy (3.1), that $V \in L^2(\mathbb{P})$ and that Assumption $(\mathcal{C})$ holds. Assume furthermore that the distribution of $\frac{D}{V}$ satisfies the continuity Assumption (3.4). Let $\gamma := (\gamma_n)_{n\geq 1}$ be a step sequence satisfying the usual decreasing step assumption*

$$\sum_{n\geq 1} \gamma_n = +\infty \quad \text{and} \quad \sum_{n\geq 1} \gamma_n^2 < +\infty.$$

Let $(V^n, D_1^n, \ldots, D_N^n)_{n\geq 1}$ be an i.d.d. sequence defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then, there exists an $\mathrm{argmax}_{\mathcal{P}_N}\Phi$*-valued random variable r^* such that*

$$r^n \longrightarrow r^* \quad a.s.$$

If the functions φ_i satisfy $(\mathcal{C}_<)$ then $\mathrm{argmax}_{\mathcal{P}_N}\Phi \subset int(\mathcal{P}_N)$.

Proof of the theorem. In this setting, the algorithm is (non homogenous) Markov discrete time process with respect to the natural filtration $\mathcal{F}_n := \sigma(r^0, (V^k, D_1^k, \ldots, D_N^k),\, 1 \leq k \leq n)$ with the following canonical representation

$$\begin{aligned} r^{n+1} &= r^n + \gamma_{n+1} H(r^n, V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1}),\ r^0 \in \mathcal{P}_N \\ &= r^n + \gamma_{n+1} h(r^n) + \gamma_{n+1} \Delta M_{n+1} \end{aligned}$$

where, for every $r \in \mathcal{H}_N$,

$$h(r) := \mathbb{E}\, H(r, V, D_1, \ldots, D_N) = \left(\varphi_i'(r_i) - \frac{1}{N} \sum_{j=1}^{N} \varphi_j'(r_j) \right)_{1 \leq i \leq N}$$

is the so-called *mean* function of the algorithm, and

$$\begin{aligned} \Delta M_{n+1} &= H(r^n, V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1}) - \mathbb{E}(H(r^n, V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1}) \,|\, \mathcal{F}_n) \\ &= H(r^n, V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1}) - h(r^n) \end{aligned}$$

since $(V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1})$ is independent of $\mathcal{F}_n$.

One derives from Proposition 3.1(a) that the mean function h of the algorithm satisfies $\{h = 0\} = \operatorname{argmax}_{\mathcal{P}_N}$ and that, for every $r \in \mathcal{H}_N \setminus \{h = 0\}$ and every $r^* \in \{h = 0\}$,

$$\langle h(r) \,|\, r - r^* \rangle = \langle h(r) - h(r^*) \,|\, r - r^* \rangle = \sum_{i=1}^{N} \underbrace{(\varphi_i'(r_i) - \varphi_i'(r_i^*))(r_i - r_i^*)}_{\leq 0} < 0 \tag{4.1}$$

simply because each function φ_i' is non-increasing which implies that each term of the sum is non-positive. The sum is not zero otherwise $\varphi'(r_i) = \varphi'(r_i^*)$ as soon as $r_i \neq r_i^*$ which would imply $h(r) = 0$.

The random vector V being square integrable, it is clear that $H_i(r, V, D_1, \ldots, D_N)$ satisfies the linear growth assumption

$$\forall\, i \in \mathcal{I}_N,\ \forall\, r \in \mathcal{H}_N, \quad \|H_i(r, V, D_1, \ldots, D_N)\|_2 \leq 2\, (\max_j \rho_j) \|V\|_2 (N + |r|)$$

At this stage one may conclude using a simple variant of the standard Robbins-Monro Theorem (like that established in [20]): there exists a random variable r^* taking values in $\{h = 0\}$ such that $r^n \to r^*$. □

4.1 Rate of convergence

Our aim in this section is to show that the assumptions of the regular Central Limit Theorem (CLT) for stochastic approximation procedures are fulfilled. For a precise statement, we refer (among others) to [3] (Theorem 13 p.332). For the sake of simplicity, we will assume that the mean function h has a single zero denoted r^*. The following lemma provides a simple criterion to ensure this uniqueness.

Lemma 4.1 *Assume that all the functions φ_i, $i \in \mathcal{I}_N$, are decreasing (strictly). Then*

$$\{h = 0\} = \mathrm{argmax}_{\mathcal{P}_N} \Phi = r^* \in int(\mathcal{P}_N).$$

Proof. In particular $(\mathcal{C}_<)$ is satisfied so that $\mathrm{argmax}_{\mathcal{P}_N} \Phi \subset r^* \in int(\mathcal{P}_N)$. If r, $r' \in \{h = 0\}$, $r \neq r'$, it follows from (4.1) that $\varphi_i'(r_i) = \varphi_i'(r_i')$ for some index i such that $r_i \neq r_i'$. $\square$

The second ingredient needed to establish a CLT will be the Hessian of function Φ. To ensure its existence we will make one further assumption on a generic random couple (V, D), keeping in mind that $\mathbb{P}(D > 0) > 0$, but that $\mathbb{P}(D = 0)$ may possibly be positive too. Namely, assume that the distribution function of (V, D) *given* $\{D > 0\}$ is absolutely continuous with a probability density f defined on $(0, +\infty)^2$. Furthermore we make the following assmptions on f:

$$\begin{cases} (i) & \text{for every } v > 0,\ u \mapsto f(v, u) \text{ is continuous and positive on } (0, \infty), \\ (ii) & \forall\, \varepsilon \in (0, 1), \quad \sup\limits_{\varepsilon V \leq u \leq V/\varepsilon} f(V, u) V^2 \in L^1(\mathbb{P}). \end{cases} \tag{4.2}$$

Note that (ii) is clearly always satisfied when $V \in L^2(\mathbb{P})$ and f is bounded. The conditional distribution function of D given $\{D > 0\}$ and V is given by

$$F_D(u \,|\, V = v, D > 0) := \mathbb{P}(D \leq u \,|\, V = v, D > 0) = \int_0^u f(v, u')du',\ u > 0,\ v > 0,$$

Lemma 4.2 (a) *Assume (V, D) satisfies the above assumption (4.2). Then the mean execution function $\varphi(u) := \rho\, \mathbb{E}(\min(uV, D))$ is concave, twice differentiable on $\mathbb{R}_+$ and for every $u > 0$,*

$$\varphi''(u) = -\rho\, \mathbb{E}\left(V^2 \mathbf{1}_{\{D>0\}} f(V, uV)\right) < 0.$$

(b) *If (V, D_i) satisfies the above assumption (4.2) for every $i \in \mathcal{I}_N$, then the function $\tilde{h}$ defined on $\mathbb{R}_+^N$ by $\tilde{h}(u_1, \ldots, u_N) = \left(\varphi_i'(u_i) - \frac{1}{N}\sum_{1 \leq j \leq N} \varphi_j'(u_j)\right)_{1 \leq i \leq N}$ is differentiable on $(0, \infty)^N$ and admits a continuous extension on $\mathbb{R}_+^N$ given by*

$$D\tilde{h}(u) = -\frac{1}{N}\Big[- a_j(u_j) + N a_i(u_i)\, \delta_{ij} \Big]_{1 \leq i,j \leq N} \qquad \textit{with} \qquad a_i(u) = -\varphi_i''(u) > 0.$$

(c) *Let $A := [-a_j + N a_i \delta_{ij}]_{1 \leq i,j \leq N}$, $a_1, \ldots, a_N > 0$ and let $\underline{a} = \min_i a_i$. Its kernel $\mathrm{Ker}(A)$ is one dimensional, $A(\mathbb{R}^N) = \mathbf{1}^\perp$ and $A_{|\mathbf{1}^\perp}$ is bijective. Every non-zero eigenvalue λ (with eigenspace E_λ) satisfies*

$$\Re(\lambda) \geq N \times \underline{a} \quad \textit{and} \quad E_\lambda \subset \mathbf{1}^\perp.$$

Proof. (a) is a straightforward application of the Lebesgue differentiation Theorem for expectation.
(b) is a consequence of (a).
(c) The transpose A^t of A has a strict dominating diagonal structure *i.e.* $A^t_{ii} > 0$, $A^t_{ij} < 0$, $i \neq j$ and $\sum_j A^t_{ij} = 0$ for every i. Consequently, it follows from Gershgorin's Lemma (see [9]) that 0 is an eigenvalue of order 1 of A^t (with $\mathbf{1}$ as an eigenvector and that all other eigenvalues have (strictly)

positive real parts). Consequently $\mathrm{Ker}(A)$ is one dimensional. The fact that $A(\mathbb{R}^N) \subset \mathbf{1}^\perp$ is obvious so that this inclusion holds as an equality by the dimension formula. Hence all the eigenvectors not in $\mathrm{Ker}(A)$ are in $\mathbf{1}^\perp$. Set $\tilde{a}_i - a_i - \underline{a} \geq 0$, $i = 1, \ldots, N$. Then $\tilde{A}^t$ has a dominating diagonal structure so that all its eigenvalues have non-negative real parts. Now if λ is an eigenvalue of A, it is obvious that $\lambda - N\underline{a}$ is an eigenvalue of $\tilde{A}$. Consequently $\Re(\lambda) \geq N\underline{a}$. □

Theorem 4.2 *Assume that the assumptions of Theorem 4.1 holds and that* $\mathrm{argmax}\,\Phi$ *is reduced to a single point* $r^* \in \mathcal{P}_N$ *so that* $r^n \to r^*$ $\mathbb{P}$*-a.s. as* $n \to \infty$*. Furthermore, suppose that Assumption (4.2) holds for every* (V, D_i), $i \in \mathcal{I}_N$ *and that* $V \in L^{2+\delta}(\mathbb{P})$, $\delta > 0$*. Set*

$$\gamma_n = \frac{c}{n},\ n \geq 1 \quad \textit{with} \quad c > \frac{1}{2\Re e(\lambda_{\min})}$$

where $\lambda_{\min}$ *denotes the eigenvalue of* $A^\infty := -Dh(r^*)_{|\mathbf{1}^\perp}$ *with the lowest real part. Then*

$$\frac{r^n - r^*}{\sqrt{\gamma_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0; \Sigma^\infty)$$

where the asymptotic covariance matrix Σ^∞ *is given by*

$$\Sigma^\infty = \int_0^\infty e^{u(A^\infty - \frac{Id}{2c})} C^\infty e^{u(A^\infty - \frac{Id}{2c})^t} du$$

where

$$C^\infty = \mathbb{E}\left(H(r^*, V, D_1, \ldots, D_N) H(r^*, V, D_1, \ldots, D_N)^t\right)_{|\mathbf{1}^\perp}$$

and $(A^\infty - \frac{Id}{2c})^t$ *stands for the* transpose operator *of* $A^\infty - \frac{Id}{2c} \in \mathcal{L}(\mathbf{1}^\perp)$.

Remark. The above claim is consistent since $u \mapsto H(r, v, \delta_1, \ldots, \delta_N)^t\, u$ preserves $\mathbf{1}^\perp$.

Proof. First note that, since $r^* \in int(\mathcal{P}_N)$, the above Lemma 4.2($b$) shows that (still making the confusion between the linear operator $Dh(r^*)$ and its matrix representation in the canonical basis)

$$Dh(r^*) = -\frac{1}{N}\left[-a_j(r_j^*) + N a_i(r_i^*)\,\delta_{ij}\right]_{1 \leq i,j \leq N} \qquad \text{with } a_i(r) = \rho_i \mathbb{E}(V^2 \mathbf{1}_{\{D_i > 0\}} f(V, rV)) > 0$$

Then, Lemma 4.2 (c) implies that $-Dh(r^*)_{|\mathbf{1}^\perp}$ has eigenvalues with positive real parts, all lower bounded by $\min_i a_i(r_j^*) > 0$.

At this stage, one can apply the CLT for stochastic algorithms defined on $\mathbf{1}^\perp$ (see *e.g.* [3], p.341). □

5 The (ERG) setting: convergence

For the sake of simplicity, although it is not really necessary, we will assume throughout this section that

$$\mathrm{argmax}_{\mathcal{P}_N}\, \Phi = \{r^*\} \subset int(\mathcal{P}_N)$$

possibly because all the execution functions φ_i are decreasing so that, following the former Lemma 4.1.

So we assume that the sequence $(V^n, D_i^n, i = 1, \dots, N)_{n\geq 1}$ satisfies $(ERG)_\nu$ with a limiting distribution ν such that, for every $i \in \mathcal{I}_N$, its marginal $\nu_i = \mathcal{L}(V, D_i)$ satisfies the consistency assumption (3.1) and the continuity assumption (3.4). We will also need to make a specific assumption: there exists $\varepsilon_0 > 0$ such that

$$\left\{ \begin{array}{ll} (i) & \mathbb{P}(V \geq \varepsilon_0) > 0 \\ (ii) & \mathrm{supp}\Big(\mathcal{L}\Big(\frac{D_i}{V}, i = 1, \dots, N \,|\, \{V \geq \varepsilon_0\}\Big)\Big) \text{ is a neighbourhood of } \mathcal{P}_N \text{ in } \mathbb{R}_+^N. \end{array} \right. \tag{5.1}$$

This assumption means that *all allocations across the pools lying in the neihbourhood of* $\mathcal{P}_N$ can be executed.

On the other hand, it follows from $(ERG)_\nu$ and some standard weak convergence arguments that

$$\forall\, i \in \mathcal{I}_N,\ \forall\, u \in \mathbb{R}_+, \quad \frac{1}{n}\sum_{k=1}^n V^k \mathbf{1}_{\{uV^k \leq D_i^k\}} - \mathbb{E}(V\mathbf{1}_{\{uV \leq D_i\}}) \overset{a.s.\&L^2}{\longrightarrow} 0 \quad \text{as } n \to \infty,$$

since the (non-negative) functions $f_u(v, \delta) := v\mathbf{1}_{\{uv\leq\delta\}}$, $u > 0$, are $\mathbb{P}_{(V,D_i)}$-*a.s.* continuous and $O(v)$ as $v \to +\infty$ by (3.4). Moreover this *a.s.* convergence holds uniformly on compact sets with respect to u since $u \mapsto \mathbb{E}V\mathbf{1}_{\{uV\leq D_i\}}$ is continuous, still owing to (3.4). Our specific assumption is to require a rate in the above *a.s.* and $L^2(\mathbb{P})$-convergence. Namely, we assume that there exists an exponent $\alpha_i \in (0, 1]$ such that

$$\forall\, u \in \mathbb{R}_+, \quad \frac{1}{n}\sum_{k=1}^n V^k \mathbf{1}_{\{uV^k < D_i^k\}} - \mathbb{E}(V\mathbf{1}_{\{uV<D_i\}}) = O(n^{-\alpha_i}) \quad a.s. \text{ and in } L^2(\mathbb{P}). \tag{5.2}$$

This assumption *e.g.* from the more general assumption that, for every $i \in \mathcal{I}_N$, the marginal $\nu_i = \mathcal{L}(V, D_i)$ satisfies (3.4) and

$$(V^n, D_i^n) \text{ is } \nu_i\text{-averaging at rate } \alpha_i$$

on a subspace $\mathcal{V}_{\alpha_i,2}$ containing all the functions f_u.

Note that, when the sequence $(V^n, D_i^n, i = 1, \dots, N)_{n\geq 1}$ is i.i.d. with distribution ν then elementary martingale arguments show that the whole sequence is ν-averaging at rate $\frac{1}{2} - \eta$ for every $\eta \in (0, 1/2)$ on $\mathcal{V}_{\frac{1}{2}-\eta,2} = L^2(\nu)$ (and all $f_u \in L^2(\nu)$, $u > 0$, since $V \in L^2(\mathbb{P})$). So, the theorem below almost embodies the *a.s.* convergence theorem established in the (IID) setting (except for the integrability assumption on V).

Now we are in position to state the main convergence result of this section. We rely on the extension of Robbins-Siegmund Lemma proposed in [18]. For the reader's convenience it is recalled in the Appendix.

Theorem 5.1 *Let $(V^n, D_1^n, \dots, D_N^n)_{n\geq 0}$ be a sequence of input satisfying $(ERG)_\nu$ and such that, for every $i \in \mathcal{I}_N$, the marginal distribution $\nu_i = \mathcal{L}(V, D_i)$ satisfies the consistency assumptions (3.1) and (3.4). Suppose furthermore that, the sequence $(V^n, D_i^n)_{n\geq 1}$ satisfies the rate assumption (5.2). If the step sequence $(\gamma_n)_{n\geq 1}$ satisfies*

$$\sum_{n\geq 1} \gamma_n = +\infty, \quad \gamma_n = o(n^{\underline{\alpha}-1}) \quad \text{and} \quad \sum_{n\geq 1} n^{1-\underline{\alpha}} \max(\gamma_n^2, |\gamma_n - \gamma_{n+1}|) < \infty$$

where $\underline{\alpha} := \min_{i\in\mathcal{I}_N} \alpha_i \in (0, 1]$, then the algorithm defined by (3.9) a.s. converges towards $r^* = \mathrm{argmax}_{\mathcal{P}_N} \Phi$.

TECHNICAL COMMENT. The above condition on the step sequence $(\gamma_n)_{n\geq 1}$ is satisfied as soon as $\gamma_n = \frac{c}{n^\beta}$ with $\beta \in (1-\underline{\alpha}, 1]$.

Proof. STEP 1. First, we aim at applying the extended Robbins-Siegmund Lemma established in [18] (see Appendix, Theorem A.1 for its statement) for stochastic algorithms with ν-averaging inputs dynamics in presence of a Lyapunov function. We will consider the case $p = 2$ and $\beta \in (0, \underline{\alpha}]$. We set $G = -H$ and $\Delta M^n \equiv 0$ and we consider the input $Y^n = (V^{n+1}, D_1^{n+1}, \ldots, D_N^{n+1})$, $n \geq 0$. Let $L(r) = \frac{1}{2}|r - r^*|^2$ be our candidate as a Lyapunov function.

First note that it follows from (3.9) that the function H satisfies the growth assumption (A.4) since

$$\forall r \in \mathcal{H}_n,\ \forall y \in \mathbb{R}^{N+1}, \quad |H(r,y)| \leq C_H g(y)(1+|r|)$$

where $C_H > 0$ and $g(v, \delta_1, \ldots, \delta_N) = v$.

In view of the ergodic assumption (5.2) and the fact that r^* lies in $\mathcal{P}_N$, it is clear from its definition that $H(r^*, .) \in \mathcal{V}_{\beta,2}$ for every $\beta \in (0, \underline{\alpha}]$.

At this stage it remains to check the "weak local Lyapunov" assumption (A.5) for $G = -H$. This fact is obvious since, for every $r \in \mathcal{H}_N$ and every input $y = (v, \delta_1, \ldots, \delta_N) \in (0, +\infty) \times \mathbb{R}^N$,

$$\langle H(r,y) - H(r^*,y) | r - r^* \rangle = \sum_{i=1}^N (\widetilde{H}_i(r_i, v, \delta_i) - \widetilde{H}_i(r_i^*, v, \delta_i))(r_i - r_i^*) \leq 0$$

where

$$\widetilde{H}_i(u, v, \delta_i) = \rho_i v \left(\mathbf{1}_{\{uv \leq \delta_i\}} \mathbf{1}_{[0,1]}(u) + (1-u)\mathbf{1}_{\delta_i > 0, u < 0} + \frac{1}{u}\mathbf{1}_{v \leq \delta_i, u > 1} \right),\ i \in \mathcal{I}_N \tag{5.3}$$

is clearly non-increasing with respect to u.

At this stage, using that $\sup_{n\geq 1} \mathbb{E}(V^n)^2 < +\infty$, we can apply our extended Robbins-Siegmund lemma also that

$$|r^n - r^*| \xrightarrow{a.s.} L_\infty < +\infty\ a.s. \text{ and } \quad \sum_{n\geq 1} \gamma_n \langle r^n - r^* \,|\, G(r^n, Y^n) - G(r^*, Y^n) \rangle < +\infty\ a.s. \tag{5.4}$$

STEP 2. At this stage it suffices to show that r^* is $a.s.$ a limiting point of $(r^n)_{n\geq 0}$ since $|r^n - r^*|$ converges to $L_\infty < +\infty\ a.s.$

Let η denote a positive real number such that, for every $i \in \mathcal{I}_N$, $[r_i^* - \eta, r_i^* + \eta] \subset (0,1)$. One derives from (5.3) and the monotonicity of $\widetilde{H}_i(u, v, \delta)$ in $u \in \mathbb{R}$ that for every $i \in \mathcal{I}_N$ and every $r \in \mathcal{H}_N$,

$$(\widetilde{H}_i(r_i, v, \delta_i) - \widetilde{H}_i(r_i^*, v, \delta_i))(r_i^* - r_i) \geq \rho_i v \eta \mathbf{1}_{\{r_i > r_i^* + \eta\}} \mathbf{1}_{\{\delta / v \in J_\eta\}}$$

where $J_\eta = (r_i^*, r_i^* + \eta)$. As a consequence

$$\langle G(r,y) - G(r^*,y) \,|\, r - r^* \rangle \geq \varepsilon_0 \underline{\rho}\, \eta \mathbf{1}_{\{v \geq \varepsilon_0\}} \mathbf{1}_{y \in O_\eta} \sum_{i \in \mathcal{I}_N} \mathbf{1}_{r_i > r_i^* + \eta}.$$

where $\underline{\rho} = \min_i \rho_i$ and the open set O_η is defined by

$$O_\eta = \left\{ y = (v, \delta_1, \ldots, \delta_N) \in (\varepsilon_0, +\infty) \times \mathbb{R}_+^N \text{ s.t. } \frac{\delta_i}{v} \in J_\eta,\ i \in \mathcal{I}_N \right\}.$$

Now, one derives from (5.4) that

$$\sum_n \gamma_n \mathbf{1}_{O_\eta}(Y^n) \sum_{i\in\mathcal{I}_N} \mathbf{1}_{r_i^n > r_i^* + \eta} < +\infty \quad a.s.$$

Now Assumption (5.1) implies that $\nu(O_\eta) > 0$. Furthermore $\nu(\partial O_\eta) = 0$ owing to the continuity assumption so that $(ERG)_\nu$ implies

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{O_\eta}(Y^k) \longrightarrow \nu(O_\eta) > 0 \quad a.s.$$

An Abel transform and the facts that the sequence γ_n is non-increasing and $\sum_{n\geq 1}\gamma_n = +\infty$ classically implies that

$$\sum_{n\geq 1} \gamma_n (\mathbf{1}_{O_\eta}(Y^k) - \nu(O_\eta)) \quad a.s. \text{ converge}$$

so that

$$\sum_n \gamma_n \mathbf{1}_{O_\eta}(Y^n) = +\infty \quad a.s.$$

In turn, this implies that

$$\liminf_n \sum_{i\in\mathcal{I}_N} \mathbf{1}_{\{r_i^n > r_i^* + \eta\}} = 0 \quad a.s.$$

This holds of course for a sequence of real numbers η^ℓ decreasing to 0.

Let $\mathcal{R}_\infty$ be the set of limiting values of the sequence $(r^n)_{n\geq 0}$. It is is $a.s.$ non-empty since the sequence $(r^n)_{n\geq 0}$ is bounded. Then $\mathcal{R}_\infty$ is $a.s.$ compact and it follows from what precedes that $\mathcal{R}_\infty \cap \prod_{1\leq i\leq N}(-\infty, r_i^* + \eta^\ell] \neq \emptyset$ (and is compact). Hence a decreasing intersection of non-empty compact sets being a (non-empty) compact set $\mathcal{R}_\infty \cap \prod_{1\leq i\leq N}(-\infty, r_i^*] \neq \emptyset$. On the other hand $\mathcal{R}_\infty \subset \mathcal{H}_N$ since the algorithm is $\mathcal{H}_N$-valued. But $\prod_{1\leq i\leq N}(-\infty, r_i^*] \cap \mathcal{H}_N = \{r^*\}$. Consequently r^* is a limiting point of the algorithm which implies that it is its true $a.s.$ limit. □

Application to α-mixing stationary data. If $(V^n, D_i^n, i = 1, \ldots, N)_{n\geq 1}$ is a stationary α-mixing sequence which mixing coefficients $(\alpha_n)_{n\geq 1}$ satisfy Ibragimov's condition for some $\delta > 0$:

$$\sum_{n\geq 1} \alpha_n^{\frac{2}{2+\delta}} < +\infty$$

(which is satisfied in case of geometric α-mixing) then the sequence $(V^n, D_i^n, i = 1, \ldots, N)_{n\geq 1}$ is ν-averaging where ν is the stationary marginal distribution of the sequence (supposed to satisfy (3.1) and (3.4)) at rate β for every $\beta \in (0, 1/2)$. To be precise, $L^2(\nu) \subset \mathcal{V}_{0^+,2}$ and

$$L^{2+\delta}(\nu) \subset \bigcap_{0<\beta<\frac{1}{2}} \mathcal{V}_{\beta,2}.$$

In particular, all the functions $f_u(v, \delta) := v\mathbf{1}_{\{uv\leq\delta\}}$, $u \geq 0$, lie in every $\mathcal{V}_{\beta,2}$, $0 < \beta < \frac{1}{2}$, so that the rate condition (5.2) is satisfied.

As concerns the stationary assumption on the input data sequence, it can be considered as realistic if one think of execution objectives given on a daily basis.

EXAMPLE: An exponential discrete time Ornstein-Ulhenbeck model for $(V^n, D_1^n, \ldots, D_N^n)$.

$$V^n = v^0 e^{X_0^n}, \qquad D_i^n = d_i^0 e^{X_i^n},\ i = 1, \ldots, N,\ n \geq 1,$$

where v^0, d_1^0, $\ldots$, d_N^0 are positive real numbers and the sequence $(X^n)_{n\geq 1}$ satisfies the linear auto-regressive dynamics

$$X^{n+1} = m + AX^n + B\,\Xi^{n+1}, n \geq 1,$$

with $m \in \mathbb{R}^{N+1}$, $A \in \mathcal{M}(N+1, N+1, \mathbb{R})$, $\|\!|A|\!\| < 1$, $B \in \mathcal{M}(N+1, M)$ with $\mathrm{rank}(B) = N+1$ $(\leq M)$ and $(\Xi^n)_{n\geq 1}$ is an i.i.d. sequence of $\mathcal{N}(0; Id_M)$-distributed random variables. We assume that the sequence is stationary *i.e.* that the distribution of X^1 is the (Gaussian) invariant distribution (with covariance matrix C solution to the Lyapunov equation $C - ACA^t = BB^t$ where ^{t} stands for transpose). Then (see [6], p. 99), the sequence $(X^n)_{n\geq 0}$ is geometrically α-mixing and subsequently so is $(V^n, D_1^n, \ldots, D_N^n)_{n\geq 1}$ (with respect to its natural filtration). Furthermore, it is clear that its distribution ν satisfies the dispersion assumption (5.1) the process $(X_0^1, X_i^1 - X_0^1, i = 1, \ldots, N)$ is a non-degenerate Gaussian distribution over $\mathbb{R}^N$ since B has full rank $N + 1$.

6 An alternative procedure based on a reinforcement principle.

Recently, inspired by the discussion developed by Almgren and Harts in [2] about liquidity estimation, Berenstein and Lehalle devised a "smart routing" recursive procedure of requests to be executed by a pool of N dark pools (see [4]). This procedure is not based on the opimization of a potential function but on a intuitive reinforcement mechanism. Let I_i^n be the profit induced by the execution of the order sent to dark pool i at time n. The proportion r_i^n of the global order V^{n+1} to be sent to dark pool i for execution at time $n+1$ is defined as proportional to this profit *i.e.* by

$$\forall\, i \in \mathcal{I}_N, \qquad r_i^n := \frac{I_i^n}{\sum_j I_j^n}.$$

The updating of the random vector I^n is as follows

$$\forall\, n \geq 0,\ \forall\, i \in \mathcal{I}_N, \quad I_i^{n+1} = I_i^n + \rho_i \min\left(r_i^n V^{n+1}, D_i^{n+1}\right), \quad I_i^0 = 0.$$

The first equation models the idea of "reinforcement" since the proportion of orders sent for execution to dark pool i is proportional to the historical performances of this dark pool since the beginning of the procedure.

The second equation describes in a standard way – like in the optimization algorithm – the way dark pools execute orders.

Elementary computations show that the algorithm can be written directly in a recursive way in terms of a new vector valued variable

$$X^n = \frac{I^n}{n}, \qquad n \geq 1,$$

since

$$X_i^{n+1} = X_i^n - \frac{1}{n+1}\left(X_i^n - \rho_i \min\left(r_i^n V^{n+1}, D_i^{n+1}\right)\right), \quad i \in \mathcal{I}_N.$$

This is a standard form a stochastic algorithm (with step $\gamma_n = \frac{1}{n}$).

Furthermore, note that setting $\underline{\rho} := \min_i \rho_i$,

$$\sum_{i\in\mathcal{I}_N} I_i^n \geq \underline{\rho} \min\left(\frac{1}{N}V^n, \min_{i\in\mathcal{I}_N} D_i^n\right)$$

since $r_i^n \geq \frac{1}{N}$ for at least one dark pool $i \in \mathcal{I}_N$. Consequently, as soon as the sequence $(V^n, D_1^n, \ldots, D_N^n)$ is stationary and ergodic

$$\liminf_n \sum_{i\in\mathcal{I}_N} X_i^n \geq \underline{\rho} \lim_n \frac{1}{n}\sum_{k=1}^n \min\left(\frac{1}{N}V^k, \min_{i\in\mathcal{I}_N} D_i^k\right) = \underline{\rho}\,\mathbb{E}\min\left(\frac{1}{N}V, \min_{i\in\mathcal{I}_N} D_i\right) \ a.s.$$

So if we make the natural assumption that

$$\mathbb{E}\min\left(\frac{1}{N}V, \min_{i\in\mathcal{I}_N} D_i\right) > 0$$

then, *a.s.*, the algorithm X^n cannot converge to 0.

If we make the additional assumption that the sequence $(V^n, D_1^n, \ldots, D_N^n)$ is *i.i.d.* then the algorithm is a discrete time (non homogenous) $\mathcal{F}_n$-Markov process with respect to the filtration $\mathcal{F}_n = \sigma(V^k, D_1^k, \ldots, D_N^k, k = 1, \ldots, n)$, $n \geq 0$, so that it admits the canonical representation

$$X_i^{n+1} = X_i^n - \gamma_{n+1}\left(X_i^n - \varphi_i(r_i^n)\right) + \gamma_{n+1}\Delta M_i^{n+1} \quad i \in \mathcal{I}_N, \quad n \geq 0,$$

where $\gamma_n = \frac{1}{n}$ and

$$\Delta M_i^n = \rho_i \min\left(r_i^{n-1}V^n, D_i^n\right) - \varphi_i(r_i^{n-1}), \quad i \in \mathcal{I}_N, \quad n \geq 1,$$

is an $\mathcal{F}_n$-martingale increment. Furthermore it is L^2-bounded as soon as $V \in L^2$.

In fact the specific difficulties induced by this algorithm are more in relation with its mean function

$$h : x \longmapsto \left(x_i - \varphi_i\left(\frac{x_i}{\sum_j x_j}\right)\right)_{1\leq i\leq N} \tag{6.1}$$

than with the martingale "disturbance term" $\gamma_{n+1}\Delta M^{n+1}$. Our first task will be to prove under natural assumptions the existence of a non degenerate equilibrium point. Then we will show why this induces the existence of many parasitic equilibrium points.

6.1 Existence of an equilibrium

In this section, we will need to introduce a new function associated to a generic order V and a generic dark pool with characteristics (ρ, D).

$$\psi(u) := \frac{\varphi(u)}{u}, \quad u > 0, \qquad \psi(0) = \varphi'(0) = \rho\,\mathbb{E}\,V\mathbf{1}_{\{D>0\}}. \tag{6.2}$$

If Assumption (3.1) holds then $\psi(0) < +\infty$ and ψ is continuous at 0. It follows from the concavity of φ and $\varphi(0) = 0$ that ψ is non-increasing. It is continuous as soon as φ is *e.g.* if Assumption (3.4) holds true.

Proposition 6.1 *Let $N \geq 1$. Assume that Assumption (3.1) holds for every couple (V, D_i), $i \in \mathcal{I}_N$.*
(a) There exists a $x^* \in \mathbb{R}_+^N$ such that

$$\sum_{i \in \mathcal{I}_N} x_i^* > 0 \quad \text{and} \quad \varphi_i \left(\frac{x_i^*}{\sum_{j \in \mathcal{I}_N} x_j^*} \right) = x_i^*, \quad i \in \mathcal{I}_N. \tag{6.3}$$

(b) Let ψ_i be the functions associated to dark pool $i \in \mathcal{I}_N$ by (6.2). Assume that for every $i \in \mathcal{I}_N$, ψ_i is (continuous and) decreasing on $[0, \infty)$ and that

$$\sum_{i \in \mathcal{I}_N} \psi_i^{-1}(\min_{i \in \mathcal{I}_N} \varphi_i'(0)) < 1. \tag{6.4}$$

Then there exists $x^* \in int(\mathcal{P}_N)$ satisfying (6.3).

Proof. (a) We define for every $r = (r_1, \ldots, r_N) \in \mathcal{P}_N$

$$\Psi(r) := \left(\frac{\varphi_i(r_i)}{\sum_{j \in \mathcal{I}_N} \varphi_j(r_j)} \right)_{i \in \mathcal{I}_N}.$$

This function maps the compact convex set $\mathcal{P}_N$ into itself. Furthermore it is continuous since, on the one hand, for every $i \in \mathcal{I}_N$, φ_i is continuous owing to the fact that (V, D_i) satisfies (3.1) and, on the other hand,

$$\sum_{j \in \mathcal{I}_N} \varphi_j(r_j) \geq \min_{j \in \mathcal{I}_N} \varphi_j \left(\frac{1}{N} \right) > 0.$$

Indeed, for every $i \in \mathcal{I}_N$,

$$\varphi_j \left(\frac{1}{N} \right) \geq \frac{1}{N} \mathbb{E} \min(V, D_i) > 0$$

since $V > 0$ $\mathbb{P}$-*a.s.* and $\mathbb{P}(D_j = 0) < 1$. Then it follows from the Brouwer Theorem that the function Ψ has a fixed point r^*. Set for every $i \in \mathcal{I}_N$,

$$x_i^* = r_i^* \sum_{j \in \mathcal{I}_N} \varphi_j(r_j).$$

It follows immediately from this definition that

$$\forall\, i \in \mathcal{I}_N, \quad x_i^* = \varphi_i(r_i^*)$$

which in turn implies that $\sum_{j \in \mathcal{I}_N} \varphi_j(r_j^*) = \sum_{j \in \mathcal{I}_N} x_j^*$ so that $r_i^* = \frac{x_i^*}{\sum_{j \in \mathcal{I}_N} x_j^*}$, $i \in \mathcal{I}_N$.

(b) For every $i \in \mathcal{I}_N$ we consider the inverse of ψ_i defined on the interval $(0, \varphi_i'(0)]$. This function is decreasing continuous and $\lim_{v \to 0} \psi_i^{-1}(v) = +\infty$. Then, let Θ be the continuous function defined by

$$\forall\, \theta \in (0, \min_{i \in \mathcal{I}_N} \varphi_i'(0)], \quad \Theta(\theta) = \sum_{i \in \mathcal{I}_N} \psi_i^{-1}(\theta).$$

We know that $\lim_{\theta\to 0}\Theta(\theta)=+\infty$ and we derive from Assumption (6.4) that $\Theta(\min_{i\in\mathcal{I}_N}\varphi_i'(0))\le 1$. So, owing to the (strict) monotonicity of θ, there exists $\theta^*\in(0,\min_{\in\mathcal{I}_N}\varphi_i'(0))$ such that $\Theta(\theta^*)=1$. Set

$$r_i^*=\psi_i^{-1}(\theta^*),\quad i\in\mathcal{I}_N.$$

Then $r^*:=(r_1^*,\dots,r_N^*)\in int(\mathcal{P}_N)$ since $\sum_i r_i^*=1$ by definition of θ^*. If $r_{i_0}^*=0$, then $\theta^*=\psi_{i_0}(0)=\min_{i\in\mathcal{I}_N}\varphi_i'(0)$ which is impossible. □

Corollary 6.1 *Assume that all the functions ψ_i are continuous and decreasing. If furthermore, the rebate coefficients ρ_i are equal (to 1) and if $\mathbb{P}(D_i=0)=0$ for every $i\in\mathcal{I}_N$ then there exists an equilibrium point lying in $int(\mathcal{P}_N)$.*

Proof. Under the above assumptions $\varphi_i'(0)=\mathbb{E}\,V>0$. Consequently

$$\psi_i^{-1}(\min_{i\in\mathcal{I}_N}\varphi_i'(0))=\psi_i^{-1}(\psi_i(0))=0<1.\quad\square$$

COMMENTS. Unfortunately there is no hope to prove that all the equilibrium points lie in the interior of $\mathcal{P}_N$ since one may always adopt an execution strategy which boycotts a given dark pool or, more generally, N_0 dark pools. So it seems hopeless to get uniqueness of the equilibrium point. To be more precise, under the assumptions of claim (b) of the above Proposition 6.1, there exists at least one strategy involving a subset of $N-N_0$ dark pools $N_0=0,\dots,N-1$ (one dark pool is needed at least). Elementary combinatorial arguments show that there are *at least* 2^N-1 equilibrium points.

So, from a theoretical point of view, we are facing a situation where there may be many parasitic equilibrium points, some of them being clearly parasitic. However it is quite difficult to decide *a priori*, even if we make the unrealistic assumption that we know all the involved distributions, which equilibrium points are parasitic.

This is a typical situation encountered when dealing with procedures devised from a reinforcement principle.

However, one may reasonably hope that some of them are so-called "traps", that means equilibrium points which are repulsive at least in one noisy direction so that the algorithm escapes from it. Another feature described below suggests that a theoretical study of the convergence behaviour of this procedure would need a specific extra work.

The next natural question is to wonder whether an equilibrium x^* of the algorithm – namely a zero of h – is (at least) a target for the algorithm *i.e.* is attractive for the companion ODE, $\dot{x}=-h(x)$.

Proposition 6.2 *An equilibrium x^* satisfying (6.3) is locally uniformly attractive as soon as*

$$\sum_{j\in\mathcal{I}_N}\frac{x_j^*}{(\bar{x}^*)^2}\varphi_j'\left(\frac{x_j^*}{\bar{x}^*}\right)<1-\frac{1}{\bar{x}^*}\max_{i\in\mathcal{I}_N}\varphi'\left(\frac{x_i^*}{\bar{x}^*}\right)$$

where $\bar{x}^*=\sum_{i\in\mathcal{I}_N}x_i^*$.

Remark. In fact the following inequalities are satisfied by any equilibrium x^*:

$$1 - \frac{1}{\bar{x}^*}\varphi_i'\left(\frac{x_i^*}{\bar{x}^*}\right) > 0, \quad i \in \mathcal{I}_N.$$

This follows from the convexity of the function $\xi \mapsto \xi - \varphi_i\left(\frac{\xi}{\bar{x}^*}\right)$ which is zero at x_i^* with positive derivative. As a consequence, the right hand side in the above sufficient condition is always positive which makes this criterion more realistic.

Proof. Elementary computations show that the differential $Dh(x)$ of h at $x \in \mathbb{R}_+^N$ is given by

$$\forall\, i,\, j \in \mathcal{I}_N, \qquad \frac{\partial h_i}{\partial x_j}(x) = \delta_{ij}\left(1 - \frac{1}{\bar{x}}\varphi_i'\left(\frac{x_i}{\bar{x}}\right)\right) + \frac{x_i}{\bar{x}^2}\varphi_i'\left(\frac{x_i}{\bar{x}}\right).$$

As a consequence all the diagonal terms of $Dh(x^*)$ are positive. The above condition for all the eigenvalues of $Dh(x)$ to have positive real parts follows from a standard application of Gershgorin's Lemma to the transpose of $Dh(x)$. □

6.2 A competitive system

But once again, even if we could show that all equilibrium points are noisy traps, the convergence would not follow for free since this algorithm is associated to a so-called *competitive system*. A competitive differential system $\dot{x} = h(x)$ is a system in which the field $h : \mathbb{R}^N \to \mathbb{R}^N$ is differentiable and satisfies

$$\forall x \in \mathbb{R}^N,\ \forall\, i,\, j \in \mathcal{I}_N,\ i \neq j, \qquad \frac{\partial h_i}{\partial x_j}(x) > 0.$$

As concerns Almgren and Harts's algorithm, the mean function h is given by (6.1), and under the standard differentiability assumption on the functions φ_i's,

$$\forall x \in \mathbb{R}^N, \quad \frac{\partial h_i}{\partial x_j}(x) = \varphi_i'\left(\frac{x_i}{x_1 + \cdots + x_N}\right)\frac{x_i}{(x_1 + \cdots + x_N)^2} > 0.$$

These systems are known to have possibly a non converging behaviour even in presence of a single (attracting) equilibrium. This is to be compared to their *cooperative* counterparts (with negative non-diagonal partial derivatives) whose flow converge uniformly on compact sets toward the single equilibrium in that case. This property can be transferred to the stochastic procedure by the mean if the so-called ODE method which shows that the algorithm almost behaves like some trajectories of the Ordinary differential Equation associated to its mean field h (see *e.g.* [13, 3, 7] for an introduction). Cooperativeness and competitiveness are in fact some criterions which ensure some generalized monotonicity properties on the flow of the ODE viewed as a function of its space variable. For some background on cooperative and competitive systems we refer to [11, 12] and the references therein.

7 Numerical tests

The aim of this section is to compare the behaviour of both algorithms on different data sets : simulated i.i.d. data, simulated α-mixing data and (pseudo-)real data.

Two natural situations of interest can be considered *a priori*: *abundance* and *shortage*. By "abundance" we mean $\mathbb{E}V \leq \sum_{i=1}^{N} \mathbb{E}D_i$ (in average, the requested volume is lower than the available one). The "shortage" setting is the reverse situation where $\mathbb{E}V > \sum_{i=1}^{N} \mathbb{E}D_i$.

In fact, in the "abundance" setting, both our procedures (optimization and reinforcement) tend to remain "frozen" at their starting allocation value (usually uniform allocation) and they do not provide a significant improvement with respect to more naive approaches. By contrast the shortage setting is by far more commonly encountered on true markets and turns out to be much more challenging for our allocation procedures, so from now on we will focus on this situation.

Our first task is to define a reference strategy. To this end, we introduce an "oracle strategy" devised by an insider who knows all the values V^n and D_i^n before making his/her optimal execution requests to the dark pools. It can be described as follows: assume for simplicity that the rebates are ordered *i.e.* $\rho_1 > \rho_2 > \cdots > \rho_N$. Then, it is clear that the "oracle" startegy yields the following cost reduction (CR) of the execution at time $n \geq 1$,

$$\mathrm{CR}^{oracle} := \begin{cases} \displaystyle\sum_{i=1}^{i_0-1} \rho_i D_i^n + \rho_{i_0} \left(V^n - \sum_{i=1}^{i_0-1} D_i^n \right), & \text{if } \displaystyle\sum_{i=1}^{i_0-1} D_i^n \leq V^n < \sum_{i=1}^{i_0} D_i^n \\ \displaystyle\sum_{i=1}^{N} \rho_i D_i^n, & \text{if } \displaystyle\sum_{i=1}^{N} D_i^n < V^n. \end{cases}$$

Now, we introduce indexes to measure the performances of our recursive allocation procedures.

- **Relative cost reduction (w.r.t. the regular market):** they are defined as the ratios between the cost reduction of the execution using dark pools and the cost resulting from an execution on the regular market for the three algorithms, *i.e.*, for every $n \geq 1$,
  - *Oracle:* $\dfrac{CR^{oracle}}{V^n}$
  - *Recursive "on-line" algorithms:* (with *algo* = *opti, reinf*). $\dfrac{CR^{algo}}{V^n} = \dfrac{\sum_{i=1}^{N} \rho_i \min\left(r_i^n V^n, D_i^n\right)}{V^n}$
- **Performances (w.r.t. the oracle):** the ratios between the relative cost reductions of our allocation algorithms and that of the oracle, *i.e.* for every $n \geq 1$

$$\frac{CR^{opti}}{CR^{oracle}} \quad \text{and} \quad \frac{CR^{reinf}}{CR^{oracle}}$$

which seems a more realistic measure of the performance of our allocation procedures since the oracle strategy cannot be beaten.

Since these relative cost reductions are strongly fluctuating (with variables V^n and D_i^n in fact), we will plot *the moving average* of these ratios (on the running period of interest) and express them in pourcentage.

Moreover, when we simulate the data, we have chosen 10^4 simulations because it corresponds approximatively to the number of pseudo-real data observed within a day.

The choice of the gain parameter is the following (in the different settings considered below)

$$\gamma_n = \frac{c}{n}, \quad n \geq 1$$

where c equals to some units.

7.1 The (IID) setting

We consider here simulated data in the i.i.d. setting, where the quantity V and D_i, $i \in \mathcal{I}_N$, are log-normal variables and $N = 3$. The variables V and D_i, $i \in \mathcal{I}_N$, satisfy the assumptions of the CLT and we have the rate of convergence at least of the optimization algorithm.

The shortage setting is specified as follows:

$$\mathbb{E}V = \frac{3}{2}\sum_{i=1}^{N} \mathbb{E}D_i$$

with

$$\mathbb{E}D_i = i,\ 1 \leq i \leq N, \quad \mathrm{Var}(V) = 1, \mathrm{Var}(D_i) = 1, 1 \leq i \leq N \quad \text{and } \rho = \begin{pmatrix} 0.01 \\ 0.03 \\ 0.05 \end{pmatrix}.$$

The running means of the performances are computed from the very beginning for the first 100 data, and by a moving average on a window of 100 data.

The initial value for both algorithms is set at $r_i^0 = \frac{1}{N}$, $1 \leq i \leq N$.

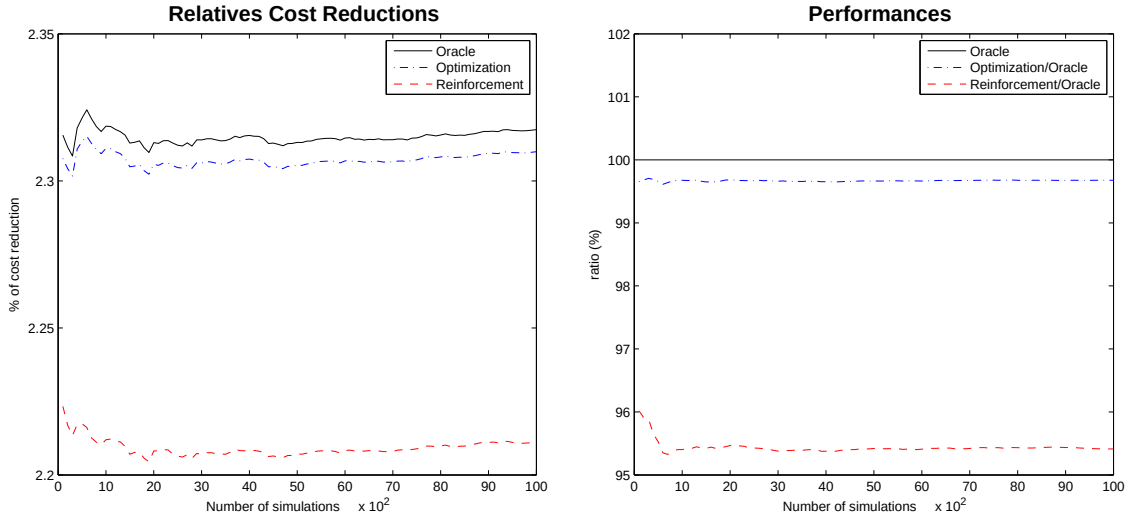

Figure 1: *Shortage setting* Case $N = 3$, $m_V = \frac{3}{2}\sum_{i=1}^{N} m_{D_i}$, $m_{D_i} = i$, $\sigma_V = 1$, $\sigma_{D_i} = 1$, $1 \leq i \leq N$.

As expected, the optimization procedure outperforms the reinforcement one and both procedures quickly converge (see Figure 1) with respect to the data set size. Note that the allocation coefficients (not reproduced here) generated by the two algorithms are significantly different. A more interesting feature is that the performances of the optimization procedure almost replicate those of the "oracle". Further simulations suggest that the optimization algorithm also seems more robust when the variances of the random variables fluctuate.

7.2 The (ERG) setting

We consider here simulated data in the ergodic setting, where the quantity V and D_i, $i \in \mathcal{I}_N$, are exponentials of an Ornstein-Uhlenbeck process, *i.e.*

$$X^{n+1} = m + AX^n + B\Xi^{n+1},$$

where $\|A\| < 1$, $BB^* \in GL(d, \mathbb{R})$ and

$$m = \begin{pmatrix} m_1 \\ \vdots \\ m_{N+1} \end{pmatrix} \in \mathbb{R}^{N+1}, \quad \Xi^{n+1} = \begin{pmatrix} \Xi_1^{n+1} \\ \vdots \\ \Xi_{N+1}^{n+1} \end{pmatrix} \sim \mathcal{N}(0, I_{N+1}) \text{ i.i.d.}, \quad e^{X^n} = \begin{pmatrix} V^n \\ D_1^n \\ \vdots \\ D_N^n \end{pmatrix}.$$

We are still interested in the shortage situation. The initial value of the algorithms is $r_i^0 = \frac{1}{N}$, $1 \leq i \leq N$ and we set

$$\rho = \begin{pmatrix} 0.01 \\ 0.03 \\ 0.05 \end{pmatrix}, \quad m = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0.7 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.3 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.2 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.1 \end{pmatrix}, \quad B = \begin{pmatrix} 0.02 & 0 & 0 & 0 \\ 0.01 & 0.9 & 0 & 0 \\ 0.01 & 0.01 & 0.6 & 0 \\ 0.01 & 0.01 & 0.01 & 0.3 \end{pmatrix}.$$

The running means of the performances are computed from the very beginning for the first 100 data, and by a moving average on a window of 100 data.

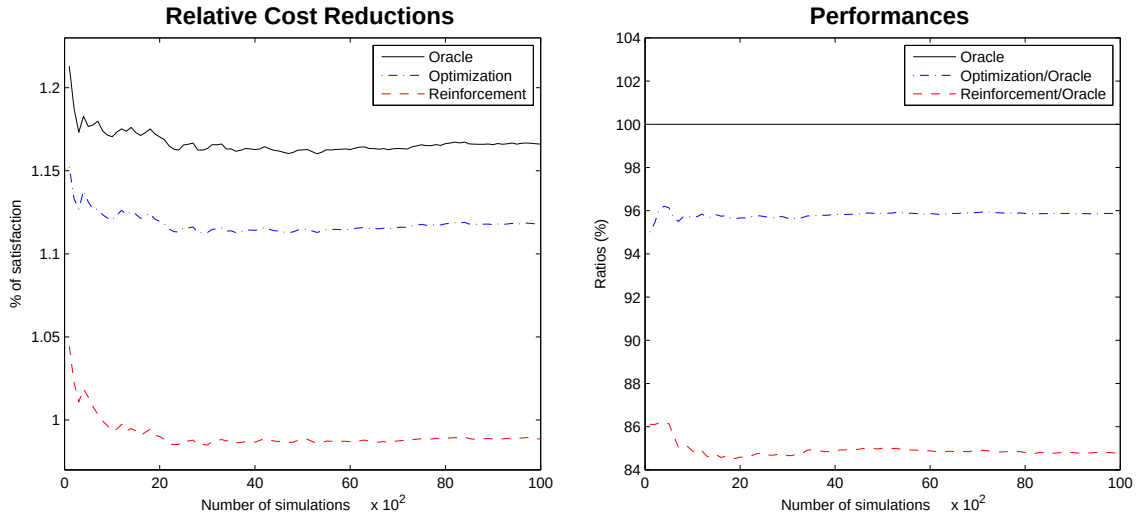

Figure 2: *Shortage setting*: Case $N = 3$, $m_V \geq \sum_{i=1}^N m_{D_i}$, $\sigma_V = 1.21$, $\sigma_D = (8.21, 3.05, 1.07)'$.

We observe in this ergodic setting a very similar behaviour to the i.i.d. one, with maybe a more significant advantage for the optimization approach (see Figure 2 right): the difference between the performances of both algorithms reaches 11% in favour of the optimization algorithm.

7.3 The pseudo-real data setting

Firstly we explain how the data have been created. We have considered for V the traded volumes of a very liquid security – namely the asset BNP – during an 11 day period. Then we selected the N most correlated assets (in terms of traded volumes) with the original asset. These assets are denoted S_i, $i = 1, \dots, N$ and we considered their traded volumes during the same 11 day period. Finally, the available volumes of each dark pool i have been modelled as follows using the mixing function

$$\forall 1 \leq i \leq N, \quad D_i := \beta_i \left((1 - \alpha_i)V + \alpha_i S_i \frac{\mathbb{E}V}{\mathbb{E}S_i} \right)$$

where $\alpha_i,\ i = 1, \dots, N$ are the mixing coefficients, $\beta_i,\ i = 1, \dots, N$ some scaling parameters and $\mathbb{E}V$ and $\mathbb{E}S_i$ stand for the empirical mean of the data sets of V and S_i.

The shortage situation corresponds to $\sum_{i=1}^N \beta_i < 1$ since it implies $\mathbb{E}\left[\sum_{i=1}^N D_i\right] < \mathbb{E}V$.

The simulations presented here have been made with four dark pools ($N = 4$). Since the data used here covers 11 days and it is clear that unlike the simulated data, these pseudo-real data are not stationary: in particular they are subject to daily changes of trend and volatility (at least). To highlight this resulting changes in the response of the algorithms, we have specified the days by drawing vertical doted lines. The dark pool pseudo-data parameters are set to

$$\beta = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.2 \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} 0.4 \\ 0.6 \\ 0.8 \\ 0.2 \end{pmatrix}$$

and the dark pool trading (rebate) parameters are set to

$$\rho = \begin{pmatrix} 0.01 \\ 0.02 \\ 0.04 \\ 0.06 \end{pmatrix}.$$

The mean and variance characteristics of the data sets of $(V^n)_{n\geq 1}$ and $(D_i^n)_{n\geq 1}$, $i = 1, \dots, 4$ are the following:

	V	D_1	D_2	D_3	D_4
Mean	955.42	95.54	191.08	286.63	191.08
Variance	2.01×10^6	9.05×10^3	4.29×10^4	4.73×10^5	5.95×10^4

Firstly, we benchmarked both algorithms on the whole data set (11 days) as though it were stationary without any resetting (step, starting allocation, etc.). In particular, the running means of the performances are computed from the very beginning for the first 1500 data, and by a moving average on a window of 1500 data. As a second step, we proceed on a daily basis by resetting the parameters of both algorithms (the initial profit for the reinforcement algorithm (*i.e.* $I_i = 0$, $1 \leq i \leq N$) and the step parameter γ_n of the optimization procedure) at the beginning of every day. The performances of both algorithms are computed on each day.

▷ **Long-term optimization** We observe that, except for the first and the fourth days where they behave similarly, the optimization algorithm is more performing than the reinforcement one. Its performance is approximately 30 % higher on average (see Figure 3).

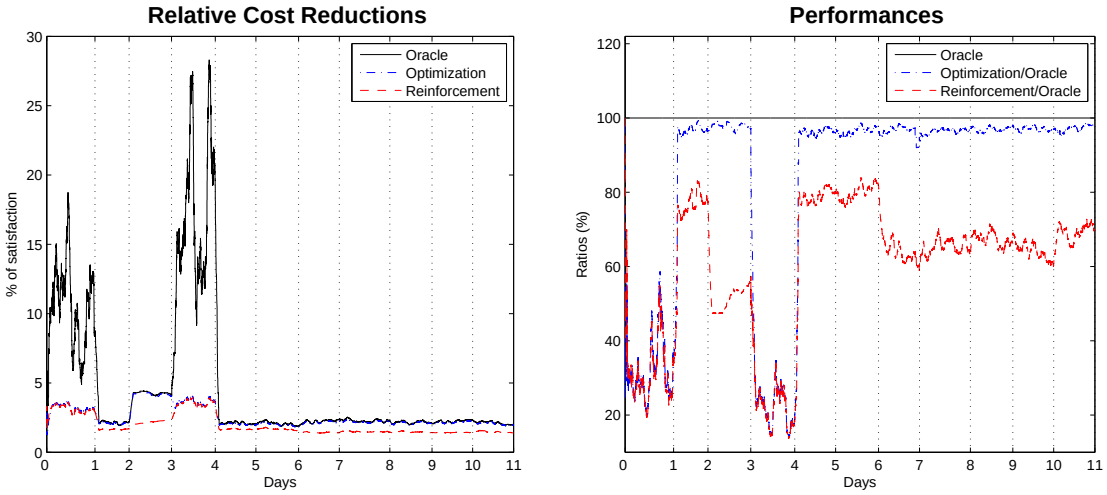

Figure 3: *Long term optimization*: Case $N = 4$, $\sum_{i=1}^{N} \beta_i < 1$, $0 < \alpha_i \leq 0.2$ and $r_i^0 = 1/N$, $1 \leq i \leq N$.

This test confirms that the statistical features of the data are strongly varying from one day to another (see Figure 3), so there is no hope that our procedures converge in standard sense on a long term period. Consequently, it is necessary to switch to a short term monitoring by resetting the parameters of the algorithms on a daily basis as detailed below.

▷ **Daily resetting of the procedure** We consider now that we reset each day all the parameters of the algorithm, namely we reset the step γ_n at the beginning of each day and the satisfaction parameters and we keep the allocation coefficients of the precedent day. We obtains the following results

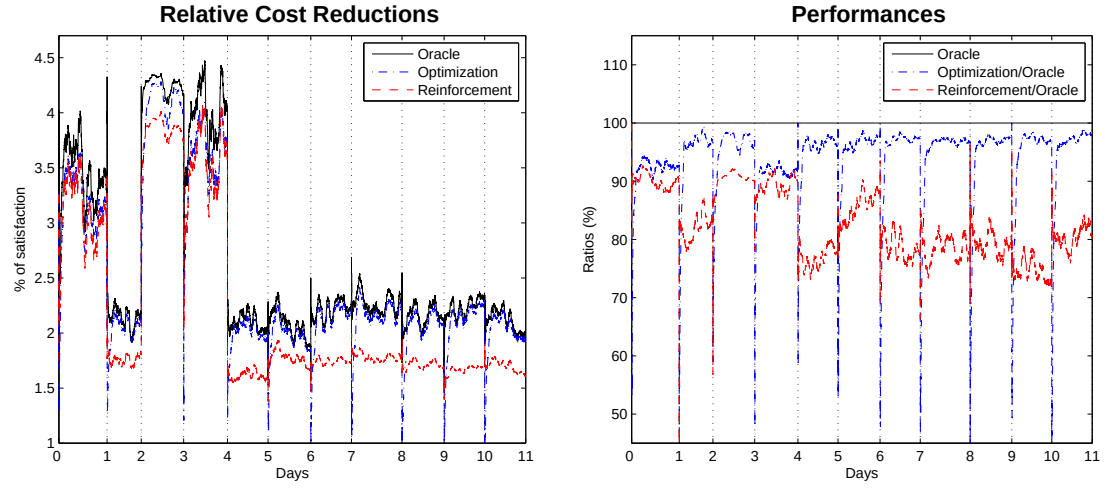

Figure 4: *Daily resetting of the algorithms parameters*: Case $N = 4$, $\sum_{i=1}^{N} \beta_i < 1$, $0 < \alpha_i \leq 0.2$ and $r_i^0 = 1/N$ $1 \leq i \leq N$.

We observe (see Figure 4) that the optimization algoritm still significantly outperforms the reinforcement one, reaching more 95 % of the performance of the oracle. Furthermore, although not represented here, the allocation coefficients look more stable.

8 Provisional remarks

8.1 Toward more general mean execution functions

One natural idea is to take into account that the rebate may depend on the quantity rV sent to be executed by the dark pool. The mean execution function of the dark pool can be modeled by

$$\forall\, r \in [0,1], \qquad \varphi(r) = \mathbb{E}(\rho(rV)\min(rV, D)) \tag{8.1}$$

where the rebate function ρ is a non-negative, bounded, non-decreasing right differentiable function.

For the sake of simplicity, we assume that (V, D) satisfies (3.4). The right derivative of φ reads

$$\varphi'_r(r) = \mathbb{E}\left(\rho'_r(rV)V\min\left(rV, D\right)\right) + \mathbb{E}\left(\rho(rV)V\mathbf{1}_{\{rV<D\}}\right), \tag{8.2}$$

with in particular $\varphi'(0) = \rho(0)\,\mathbb{E}(V\mathbf{1}_{\{D>0\}}) > 0$ as above. The main gap is to specify the function ρ so that φ remains concave which is the key assumption to apply the convergence theorem. Unfortunately the choice for ρ turns out to strongly depend on the (unknown) distribution of the random variable D. Let us briefly consider the case where V and D are independent and D has an exponential distribution $\mathcal{E}(\lambda)$.

First note that the function g defined by $g(u) := \mathbb{E}(u \wedge D)$, $u \geq 0$ is given by

$$\forall\, u \geq 0, \quad g(u) = \frac{1 - e^{-u\lambda}}{\lambda}$$

so that, owing to the independence of V and D,

$$\forall\, r \geq 0, \quad \varphi(r) = \mathbb{E}(\rho(rV)g(rV)).$$

At this stage, φ will be concave as soon as the function $\rho\, g$ is so. Among all possible choices, elementary computations show that a class of possible choices is to consider $\rho = g^\theta$ with $\theta \in (0, \lambda]$. Of course this may appear as not very realistic since the rebate function is a structural feature of the different dark pools.

However several numerical experiments not reproduced here testify that both algorithms are robust to a realistic choice for the function ρ *e.g.* a non-decreasing and stepwise constant.

Another natural extension is to model the fact that the dark pool may take into account the volume rV to decide which quantity will really executed rather than simply the *a priori* deliverable quantity D. One reason for such a behaviour is that the dark pool may wish to preserve the possibility of future transactions with other clients.

One way to model this phenomenon is to introduce a *delivery function* $\psi : \mathbb{R}_+^2 \to \mathbb{R}_+$, non-decreasing and concave w.r.t. its first variable and satisfying $0 \leq \psi(x, y) \leq y$, so that the new mean execution function is as follows:

$$\varphi(r) = \rho\,\mathbb{E}\left(\min(rV, \psi(rV, D))\right). \tag{8.3}$$

It is clear that the function φ is concave (as the minimum of two concave functions) and bounded. In this case, the first (right) derivative of φ reads

$$\varphi'_r(r) = \rho\,\mathbb{E}\left(V\left(\mathbf{1}_{\{rV<\psi(rV,D)\}} + \psi'_x(rV,D)\mathbf{1}_{\{rV\geq\psi(rV,D)\}}\right)\right) \tag{8.4}$$

where ψ'_x denotes the right derivative with respect to x. In particular $\varphi'_r(0) = \rho\,\mathbb{E}(V\mathbf{1}_{\{D>0\}}) > 0$.

As concerns the implementations resulting from these new execution functions, the adaptation is straightforward. Note for the optimization procedure under constraints that, firstly, the "edge" functions R_i functions are not impacted by the type of the execution function. On the other hand the definition of the functions H_i or the updating of the variables I^n for the reinforcement procedure should be adapted in accordance with the representations (8.2) and (8.4) of the new mean execution function φ.

EXAMPLE: We consider for modelling the quantity delivered by the dark pool i a function where we can define a minimal quantity required to begin to consum D_i, namely

$$\psi_i(rV, D_i) = D_i\mathbf{1}_{\{rV>s_iD_i\}}$$

where s_i is a parameter of the dark pool i assumed to be deterministic.

Pseudo-real data setting

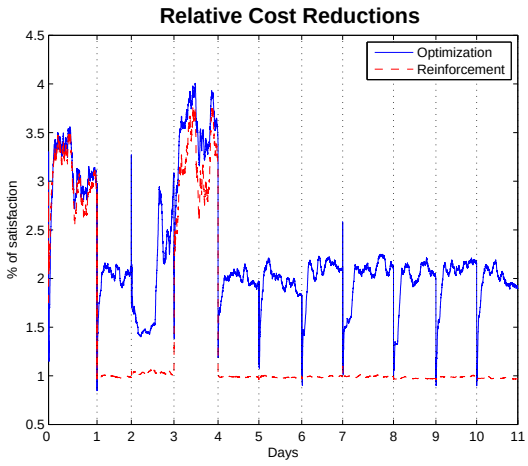

Figure 5: *Shortage setting*: Case $N = 4$, $\sum_{i=1}^{N}\beta_i < 1$, $0 < \alpha_i \leq 0.2$ and $r_i^0 = 1/N$, $1 \leq i \leq N$, $s = (0.3, 0.2, 0.2, 0.3)^t$.

8.2 Optimization *vs* reinforcement ?

For practical implementation what conclusions can be drawn from our investigations on both procedures. Both reach quickly a stabilization/convergence phase close to optimality. The reinforcement algorithm leaves the simplex structurally stable which means the proposed dispatching at each time step is realistic whereas the stochastic Lagrangian algorithm in its present form sometimes needs to be corrected from time to time. This can be corrected by adding a projection on the simplex at

each step. We did not consider this variant from a theoretical point of view to keep our convergence proofs more elementary.

In a high volatility context, the stochastic Lagrangian algorithm clearly prevails with performances that turn out to be significantly better. This optimization procedure also relies on established convergence results in a rather general framework (stationary α-mixing input data). However, given the computational cost of these procedures which is close to zero, a possible strategy is to implement them in parallel to get a synergistic effect. In particular, one may the reinforcement algorithm – which step parameter is structurally fixed equal to $\frac{1}{n}$ – can be used to help tuning the constant c in the gain parameter $\gamma_n = \frac{c}{n}$ of the stochastic Lagrangian. Doing so one may start with a small constant c, preventing a variance explosion of the procedure. Then based one may increase slowly this constant until the Lagragian outperforms the reinforcement procedure.

References

[1] Almgren, R. F. and Chriss N. (2000). Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5-39.

[2] Almgren R. and Harts B. (2008). A dynamic algorithm for smart order routing. Technical report, StreamBase.

[3] Benveniste M., Métivier M., Priouret P. (1987). *Adaptive Algorithms and Stochastic Approximations*, Springer Verlag, Berlin, 365p.

[4] Berenstein R. (2008). *Algorithmes stochastiques, microstructure et exécution d'ordres*, Master 2 internship report (Quantitative Research, Dir. C.-A. Lehalle, CA Cheuvreux), *Probabilités & Finance*, UPMC-École Polytechnique.

[5] Dedecker J., Doukhan P., Lang G., León J.R., Louhichi S., Prieur C. (2007). *Weak dependence* (with Examples and Apllications), Lecture notes in Statistcs, **190**, Springer, Berlin, 317p.

[6] Doukhan P. (1991). *Mixing: Properties and Examples*, Lect. Notes Statist. **85**, Springer, New York, 142p.

[7] Duflo M. (1997). *Random Iterative Models*. Series: Stochastic Modelling and Applied Probability, Vol. 34. XVI, 385 p., Springer, Berlin.

[8] Foucault, T. and Menkveld, A. J. (2006). Competition for order flow and smart order routing systems, *Journal of Finance*, **63**(1):119-158.

[9] Gantmacher F. R. (1959). *The theory of matrices*, **1-2**, Chelsea, New York, 374p.

[10] Gál I.S. and Koksma J.F. (1950). Sur l'ordre de grandeur des fonctions sommables, *Indigtiones Math.*, **12**:192-207.

[11] Hirsch M., Smith H. (2004). *Monotone Dynamical systems*, monography, 136p.

[12] Hirsch M., Smith H. (2005). *Monotone Dynamical systems*, Handbook of differential equations: ordinary differential equations. Vol. II, Elsevier B. V., Amsterdam, 2005, 239-357.

[13] Kushner H.J., Yin G.G. (1997). Stochastic approximation and recursive algorithms and applications, New York, Springer, 496p; and 2^{nd} edition, 2003.

[14] Lamberton D., Pagès G., Tarrès P. (2004). When can the two-armed bandit algorithm be trusted? *The Annals of Applied Probability*, **14**(3):1424-1454.

[15] Lamberton D., Pagès G. (2008). How fast is the bandit? *Stoch. Anal. and Appl.*, **26**:603-623.

[16] Lamberton D., Pagès G. (2008) A penalized bandit algorithm, *EJP*, **13**:341-373.

[17] Laruelle S., PhD Thesis, in progress.

[18] Laruelle S., Pagès G. (2009). Stochastic Approximation with Ergodic Innovation, in progress.

[19] Lehalle C.-A. (2009). Rigorous strategic trading: Balanced portfolio and mean-reversion. *The Journal of Trading*, **4**(3):40-46.

[20] Lemaire V., Pagès G. (2008). Unconstrained recursive importance sampling, pre-pub LPMA 1231, to appear in *Annals of Applied Probability*.

[21] Pagès G. (2001). Sur quelques algorithmes récursifs pour les probabilités numériques, *ESAIM P&S*, **5**:141-170.

[22] Tarrès P. (2001). Traps of stochastic algorithms and vertex-reinforced random walks, thèse d'Université, ENS Cachan, 180p.

Appendix

A Robbins-Zygmund with averaging innovation

We consider an algorithm of the following form

$$\theta_{n+1} = \theta_n - \gamma_{n+1}(G(\theta_n, Y_n) + \Delta M_{n+1}), \quad n \geq 0 \tag{A.1}$$

where G is a Borel function from $\mathbb{R}^d \times \mathbb{R}^q$ to $\mathbb{R}^d$, $(Y_n)_{n\geq 0}$ is a sequence of $\mathbb{R}^q$-valued random vectors adapted to a filtration $(\mathcal{F}_n)_{n\geq 1}$, θ_0 is an $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random vector independent of $(Y_n)_{n\geq 1}$, all defined on the same probability space $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$, $(\Delta M_n)_{n\geq 1}$ is a sequence of $\mathcal{F}_n$-martingale increments and $(\gamma_n)_{n\geq 1}$ is a non-increasing sequence of positive real numbers going to 0 as n goes to infinity.

We will say that $(Y_n)_{n\geq 0}$ is ν-averaging (under $\mathbb{P}$) on a class of functions $\mathcal{V}_{0^+,p} \subset L^p(\nu)$ if, for every $p \in [1, +\infty)$,

$$\forall f \in \mathcal{V}_{0^+,p}, \quad \frac{1}{n}\sum_{k=0}^{n-1} f(Y_k) \longrightarrow \int_{\mathbb{R}^q} f d\nu \quad \mathbb{P}\text{-}a.s. \text{ and in } L^p(\mathbb{P}). \tag{A.2}$$

Let $\beta \in (0,1)$, let $p \in [1, \infty)$. We denote by $\mathcal{V}_{\beta,p}$ the class of functions whose convergence rate in (A.2) $\mathbb{P}$-$a.s.$ and in $L^p(\mathbb{P})$ is $O(n^{-\beta})$, namely

$$\mathcal{V}_{\beta,p} = \left\{ f : \mathbb{R}^q \to \mathbb{R} \,|\, \frac{1}{n}\sum_{k=1}^{n} f(Y_k) - \int f d\nu \overset{\mathbb{P}\text{-}a.s. \ \&\ L^p(\mathbb{P})}{=} O(n^{-\beta}) \right\}. \tag{A.3}$$

Now we are in a position to state the convergence theorem.

Theorem A.1 *(A Robbins-Zygmund like Lemma) Let $G : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$ be a Borel function, let $(Y_n)_{n\geq 0}$ be an $\mathcal{F}_n$-adapted ν-averaging sequence of $\mathbb{R}^q$-valued random vectors and let $(\Delta M_n)_{n\geq 1}$ be a sequence of $\mathcal{F}_n$-martingale increments. Assume that there exists a continuously differentiable function $L : \mathbb{R}^d \to \mathbb{R}_+$ satisfying*

$$\nabla L \text{ is Lipschitz continuous and } |\nabla L|^2 \leq C\,(1 + L) \tag{A.4}$$

such that the function G satisfies the following local weak mean-reverting *assumption:*

$$\forall\, \theta \in \mathbb{R}^d,\ \forall\, y \in \mathbb{R}^q, \quad \langle \nabla L(\theta) \,|\, G(\theta, y) - G(\theta^*, y) \rangle \geq 0. \tag{A.5}$$

Suppose there exists $\beta \in (0,1)$, $p \in [1,\infty)$ *such that*

$$G(\theta^*, \cdot) \in \mathcal{V}_{\beta,p}. \tag{A.6}$$

Moreover, assume that G *satisfies the following linear growth assumption*

$$\forall \theta \in \mathbb{R}^d, \forall y \in \mathbb{R}^q, \quad |G(\theta, y)| \leq \varphi(y)(1 + L(\theta))^{\frac{1}{2}} \tag{A.7}$$

and

$$\mathbb{E}(|\Delta M_{n+1}|^2 \,|\, \mathcal{F}_n) \leq \varphi^2(Y_n)(1 + L(\theta_n)) \tag{A.8}$$

where the function φ *satisfies* $\sup_n \|\varphi(Y_n)\|_{2\vee\frac{p}{p-1}} < +\infty$ *(convention* $\frac{1}{0} = +\infty$*).*

Let $\gamma = (\gamma_n)_{n\geq 1}$ *be a non-negative, non-increasing sequence of gain parameters satisfying*

$$\sum_{n\geq 1} \gamma_n = +\infty, \quad n^{1-\beta}\gamma_n \longrightarrow 0, \quad \textit{and} \quad \sum_{k\geq 1} k^{1-\beta} \max\left(\gamma_k^2, |\Delta\gamma_{k+1}|\right) < +\infty. \tag{A.9}$$

Assume that θ_0 *is* $\mathcal{F}_0$*-adapted and* $L(\theta_0) < +\infty$, $\mathbb{P}$*-a.s. Then, the recursive procedure defined by (A.1) satisfies*

$$L(\theta_n) \xrightarrow{a.s.} L_\infty < +\infty \;\; \textit{and} \;\; \sum_{n\geq 0} \gamma_{n+1} \langle \nabla L(\theta_n) \,|\, G(\theta_n, Y_n) - G(\theta^*, Y_n)\rangle < +\infty \; a.s.$$