HAAR WAVELETS SYSTEMS FOR HEDGING FINANCIAL DERIVATIVES

P. J. CATUOGNO, S. E. FERRANDO, AND A. L. GONZALEZ

Abstract. We present a new discretization of financial instruments in terms of a martingale expansion constructed using Haar wavelets systems. In particular, expansions on these bases give the pointwise convergence needed for hedging derivatives. Examples of these systems are constructed which illustrate the discrete, spacewise, nature of the approximations. We describe natural conditions under which our Haar hedging strategy can be realized by means of a self financing portfolio consisting of binary options.

1. Introduction

Continuous models for the underlying asset are well established although in practice the hedging of options depending on this underlying is performed through a time discretization. In delta hedging the underlying itself is used for constructing the portfolio replication, this involves an implicit linear spatial approximation of the option. This approximate hedging gives a pointwise error, the quality of this error depends on the efficiency of this space-time approximation.
We present an approximation which is based on an explicit discretization of the probability space. The discretization iteratively defines atomic sigma algebras constructed via random variables which are representatives of the process, this is the basic ingredient to achieve pointwise convergence to the option being hedged. The main technical tool is the use of H-systems (or Haar systems) which are a basic generalization of the Haar wavelets on the interval $[0, 1]$, with Lebesgue measure, to a general probability space $(\Omega, \mathcal{B}, P)$. The choice of this particular wavelet is not arbitrary, besides giving an orthonormal basis of $L^2(\Omega, \mathcal{B}, P)$, partial sums of expansions in this basis form a martingale sequence. This last property is crucial for the Haar based hedging we propose, it allows for self financing strategies.
A successful deployment of our proposal involves the construction/definition of the above mentioned atomic sigma algebras. When the formalism is used for hedging, the probabilistic events involved need to have specific financial meaning, moreover, the efficiency of the Haar based hedging will depend on the way in which the atoms of the algebras are generated. A financial realization of the proposed mathematical constructs requires the existence of *binary options* associated to these atoms. We present examples where this requirement is naturally met.
We will restrict ourselves to arbitrary European type portfolios in a one dimensional setting. We hope it will be clear that the ideas and techniques are also meaningful in more general situations.

Date: November 28, 2004.
S.E.Ferrando thanks NSERC for its support.

We now describe informally the main ideas behind our approach, precise definitions will be found elsewhere in the document. Let $\{B_t\}$ denote a bank account and $\{S_t\}$ a price process where $T_0 \leq t \leq T$. Our constructions give the following identity valid for any random variable X,

$$\mathbf{E}(X|\mathcal{B}_n) = \mathbf{E}(X|u_0, u_1, \ldots, u_n) = \sum_{k=0}^{n} \langle X, u_n \rangle u_n, \ n \geq 0. \tag{1.1}$$

The simple functions u_n, the *Haar functions*, are an orthonormal set in $L^2(\Omega, \mathcal{F}, P)$, where $(\Omega, \mathcal{F})$ is the sigma algebra generated by the price process and P is the risk neutral measure. The sigma algebra $\mathcal{B}_n$ is generated by $u_0, \ldots, u_n$ and contains $n+1$ atoms, these atoms give a space-time discretization of the process and, under natural conditions, can be realized financially via binary options. It follows that (1.1) can be realized by means of a dynamic portfolio of binary options. The left hand side of (1.1) is a martingale which, under appropriate conditions, converges to X almost everywhere (a.e.). Therefore, we have a portfolio of binary options converging a.e. to X, moreover this portfolio can be implemented dynamically, via financial transactions, in a self financing way due to the martingale property. In short, we have a discrete, self-financing, hedging strategy to replicate X. This hedging strategy will be referred to as *Haar hedging* below.
To explain the usefulness of (1.1) it is better to give a brief indication of some basic ideas in modern computational harmonic analysis (CHA). The right hand side of (1.1) is just a rewrite of the left hand side in terms of the martingale differences, which always form an orthogonal set. The novelty is in the appearance of the inner products $\langle X, u_n \rangle$, these are a set of new coordinates with useful properties and information. In particular, these inner products can be efficiently computed via the multiresolution analysis algorithm (see Appendix A). Moreover, the setting is flexible enough so that the actual Haar functions u_n can be chosen via some optimization, see Section 5.2, in order to give efficient representations of X. Efficient representations of functional classes is a chief concern of CHA, see for example [7], these representations can be used, in signal processing settings, for several tasks, in particular compression and denoising, see for example [5] and [6]. Potentially, the representation (1.1) can be used for analogous tasks in finance, reference [14] describes possible uses of a Hilbert space basis for valuation and hedging. Nevertheless, our main motivation to search for efficient representations is the desire to achieve a small error of approximation to X while incurring in a small number of financial transactions when implementing the portfolio of binary options. This will keep the number of transactions in the Haar hedging portfolio realistically small. The number of transactions is, roughly speaking, the number of Haar functions in the approximation. Our approximations also open the possibility of reducing the transaction costs while achieving a small hedging error. Assuming the cost of a transaction is δ-proportional to the volume of transactions, the following definition is meaningful when studying transaction costs.

Definition 1. *Let $w \in \Omega$, and Π_1 and Π_2 be two approximating hedging portfolios for X. We say that Π_1 is more efficient than Π_2 (at w) if*

$$|\Pi_1(\omega) - X(\omega)| \leq |\Pi_2(\omega) - X(\omega)| \text{ and } VT(\Pi_1)(\omega) \leq VT(\Pi_2)(\omega),$$

where $VT(\Pi_i)(\omega)$ is the volume of transactions necessary to implement the portfolio Π_i at w.

Clearly, the above definition can be easily modified to require the inequalities to hold with large probability or in the mean. For technical reasons, this paper will not address the issue of minimizing the volume of transactions (while keeping a small hedging error) directly but instead concentrates in minimizing the *number of transactions* which is a more standard quantity in wavelet theory.

We now explain the empirical meaning of the representation (1.1) and compare it with "static" hedging and delta-hedging. Usually, static option replication involves hedging an option X with other options, see for example [4]. For simplicity, consider an option X that initiates at T_0 and expires at T, assuming one can perform a static hedging that does not involve short selling the cost of exact static hedging replication is $\delta\ V_{T_0}(X)$ where $V_{T_0}(X)$ is the risk neutral price of X. On the one hand, static hedging generally does not involve B_t, on the other hand, the bank account is a crucial ingredient in delta-hedging. We may assume that there are no transaction costs related to B_t, hence the transaction costs depend only on the volume of transactions associated to the price process, this opens the possibility to have a smaller cost than $\delta\ V_{T_0}(X)$. We note that $u_0 = \mathbf{1}_\Omega$ and therefore, it can be implemented by means of the bank account, the Haar functions are of the form $u_k = a\,\mathbf{1}_{A_0} + b\,\mathbf{1}_{A_1}$ where A_0 and A_1 ($A_0 \cap A_1 = \emptyset$) are atoms of $\mathcal{B}_i$ for some $i \leq k$ and $A = A_0 \cup A_1$ is an atom of $\mathcal{B}_{i-1}$. The simple functions u_k, for $k \geq 1$, are *wavelets*, namely $\int_\Omega u_k(\omega)\ dP(\omega) = 0$, which under natural conditions can be realized by means of binary options, involving short selling. It will be clear that $\langle X, u_k \rangle\ u_k$ approximates the oscillations of $X - \mathbf{E}_A(X)$ on A (the support of u_k) where $\mathbf{E}_A(X)$ denotes the expectation on A. In general, the events A_0 and A_1 will be level sets of financially relevant random variables, hence the wavelet u_k captures fluctuations in X due to these two financial events. In short, the financial meaning of (1.1) is the use of the bank account to capture the mean value of X and the use of binary options (involving short selling) to capture the oscillations of X about this mean value. Even though Haar hedging uses (binary) options to build the replicating portfolio, it will be misleading to call it a static type of hedging as we explain next. In general, each u_k is localized to its support, say the atom A, this atom will be localized in time to same interval $[s_a, t_a]$ (essentially, this means that A is generated by the random variables $\{S_t\}_{s_a \leq t \leq t_a}$) and will also be localized in space (it will be the level set of some appropriate random variable). This *localization* of the Haar functions, and hence of the binary options, has the effect that for a given unfolding path $w \in \Omega$ only the Haar functions in (1.1) whose support contain this w have to be implemented by the Haar hedging portfolio. This is the essence of dynamic hedging. The localization property opens the possibility, through the dynamic conditioning on the unfolding path, of obtaining efficient Haar hedging portfolios for general options X. It is also recognized in signal processing applications that localization of wavelets is a key property to represent discontinuities efficiently [5], we have observed this phenomena also in our numerical examples, see Section 6. Finally, in order to have a useful insight into our approach one can think that the linear approximation implicit in delta-hedging is replaced in Haar hedging by an appropriate simple function. This point of view clearly indicates the fundamental nature, relative to delta-hedging, of the newly proposed hedging.

One simplified way to put our contributions into perspective is to emphasize that we are studying the use of binary options for hedging general derivatives. More to the point, our contributions reveal a martingale structure underlying binary options.

This structure is then used to set up hedging strategies and related algorithms. Our contributions give evidence to the notion that digital options are fundamental building blocks for other options, this suggestion has already been made in [17].
The paper is organized as follows, in Section 2 we introduce H-systems, some basic properties and several examples, some of these examples will be used in later sections. Section 3 describes some fundamental constructions and computational properties of H-systems. Section 4 starts by describing simple, but revealing, examples of Haar hedging and continues with a general approach to hedging via H-systems and the self-financing property of Haar hedging is proved. Section 5 introduces two general and powerful approaches to obtain efficient H-system representations, in particular the *greedy splitting* algorithm is introduced which constructs adapted H-systems, for a given option X, under rather general conditions. Section 6 presents numerical output as well as comparisons with alternative approaches. Section 7 describes possible extensions and summarizes the main results of the paper. Appendix A states and proves the properties of a multiresolution analysis for H-systems. Appendix B recalls, for completeness, known results and inequalities for martingales applicable in our setting. We also discuss the issue of working with finite or infinite H-systems. As a final note, for the sake of simplicity, we did not qualify many statements with the phrase "up to a set of measure zero".

2. H-Systems

Let $(\Omega, \mathcal{B}, P)$ denote an arbitrary probability space. The notation $|| \, ||^2 = \langle , \rangle$ stands for the inner product on $L^2(\Omega, \mathcal{B}, P)$. The following Gundy's [8] definition is motivated by the standard Haar system of $L^2([0,1])$.

Definition 2. *An orthonormal system of functions $\{u_k\}_{k\geq 0}$ defined on Ω is called an H-system if and only if for any $X \in L^2(\Omega, \mathcal{B}, P)$*

$$X_{\mathcal{B}_n} \equiv \mathbf{E}(X|u_0, u_1, \ldots, u_n) = \sum_{k=0}^{n} \langle X, u_n \rangle u_n, \text{ for all } n \geq 0, \tag{2.1}$$

where $\mathcal{B}_n = \sigma(u_0, \ldots, u_n)$. The intended meaning of $k \geq 0$ in the above definition is to allow the system $\{u_k\}_{k\geq 0}$ to be finite or infinite. We caution the reader that we will attach the word *Haar* to several definitions and constructions even though they may refer to general H-systems, see also Definition 4. Underlying our proposed pointwise hedging approximations is the associated martingale convergence theorem which is stated in Appendix B. The following proposition, which is proven in [8], gives an alternative characterization of H-systems equivalent to Definition 2.

Proposition 1. *An orthonormal system $\{u_k\}_{k\geq 0}$ defined on Ω is an H-system if and only if the following three conditions hold:*

1. *Each u_k assumes at most two nonzero values with positive probability.*
2. *The σ-algebra $\mathcal{B}_n$ consists exactly of $n+1$ atoms.*
3. $\mathbf{E}(u_{k+1}|u_0, u_1, \ldots, u_k) = 0$; $k \geq 0$. *So the functions u_k are martingale differences.*

Corollary 1. *Assume $\{u_k\}_{k\geq 0}$ is an H-system. Then, for each $n \geq 0$, u_{n+1} takes two nonzero values (one positive and the other negative) only on one atom of $\mathcal{B}_n$ (hence this atom becomes its support). Consequently, $\mathcal{B}_{n+1}$ consists of n atoms from $\mathcal{B}_n$ and two more atoms obtained by splitting the remaining atom from $\mathcal{B}_n$.*

Proof. According to the property (2) of Proposition 1, $\mathcal{B}_0$ must be generated by exactly one atom, Ω, then $u_0 = 1$. By the orthonormality condition, and (1) from Proposition 1, u_1 must take two nonzero values, then it has to take each one on each atom of $\mathcal{B}_1$. Now, reasoning inductively, let $\{A_k : k = 0, ..., n\}$ be the set of atoms of $\mathcal{B}_n$. It is clear that $\{w \in \Omega : u_{n+1}(\omega) \neq 0\}$ intersects properly at most one A_k, otherwise $\mathcal{B}_{n+1}$ would have more than $n+2$ atoms. Since u_{n+1} assumes at most two nonzero values with positive probability and is orthonormal to u_k, $k = 0, ..., n$, it has to take two different nonzero values on that A_k. Moreover, since $\mathcal{B}_{n+1}$ must have $n+2$ atoms, u_{n+1} can not vanish on that A_k. Then we can conclude that u_{n+1} splits one atom of $\mathcal{B}_n$ to create two new atoms which will belong to $\mathcal{B}_{n+1}$. □

It is then obvious that an H-system naturally defines a *binary tree* of partitions, the next definition introduces, recursively, a convenient indexation of this tree.

Definition 3. *A sequence of partitions of Ω, $\mathcal{Q} := \{\mathcal{Q}_j\}_{j\geq 0}$, is called a binary sequence if $\mathcal{Q}_0 = \{A_{0,0} = \Omega\}$ and for $j \geq 1$ consider $A \in \mathcal{Q}_j$, if $A = A_{k,i} \in \mathcal{Q}_{j-1}$ then A preserves its index, otherwise (i.e. $A \notin \mathcal{Q}_{j-1}$) then there exists $A' \in \mathcal{Q}_j$ and $A_{k,i} \in \mathcal{Q}_{j-1}$ such that*

$$A_{k,i} = A \cup A' \tag{2.2}$$

then set $A_{k+1,2i} = A$ and $A_{k+1,2i} = A'$.

The index j in $A_{j,i}$ will be called the *scale parameter* (we will also call it the *level*), it indicates the number of times $A_{0,0}$ has been split to obtain $A_{j,i}$. The name scale is borrowed from wavelet theory where it indicates the extent of the spatial localization (or spatial *resolution*) of the wavelet, in our setting the analogy is strict in the case of equal probability splitting.
Although the essential computations introduced in this paper, in particular the ones described in Section 5.1 and Appendix A, can be carried out in a binary tree data structure, we will introduce a refinement of this data type by means of *weak dyadic partitions*, see Definition 7.

Definition 4. *We say that an H-system $\{u_k\}_{0\leq k\leq m}$ is a Haar system if $m = \infty$ or $m = 2^J - 1$ and each atom of $\sigma(u_0, \ldots, u_{2^j-1})$ is the union of two atoms of $\sigma(u_0, \ldots, u_{2^{j+1}-1})$ for all j such that $j < J - 1$.*

Definition 5. *Given $A \in \mathcal{B}$, $P(A) > 0$, a function ψ is called a Haar function on A if there exist $A_i \in \mathcal{B}$, $A_0 \cap A_1 = \emptyset$, $A = A_0 \cup A_1$, $\psi = a\, \mathbf{1}_{A_0} + b\, \mathbf{1}_{A_1}$ and*

$$\int_\Omega \psi(\omega)\ dP(\omega) = 0,\ \int_\Omega \psi^2(\omega)\ dP(\omega) = 1.$$

Examples:

Five examples of H-systems are described next, the first one is a classical one from wavelet theory, the remaining ones belong to the realm of finance. The examples illustrate basic constructions, some of the examples will be further developed once we introduce the notion of Haar hedging.

1. The Haar system in $L^2([0, 1])$:
We use the classical notation $L^2([0, 1])$ for the set of square integrable functions with respect to the Lebesgue measure of $[0, 1]$.

Let $\psi_H : \mathbb{R} \to \mathbb{R}$ be the Haar function

$$\psi_H(t) = \begin{cases} 1 & \text{if} \quad 0 \leq t \leq 1/2 \\ -1 & \text{if} \quad 1/2 < t \leq 1 \end{cases},$$

$$u_0 \equiv 1 \text{ and } u_{2^j+i}(t) = 2^{j/2}\psi_H(2^j t - i),$$

for $j = 0, 1, \ldots$ and $i = 0, \ldots, 2^j - 1$. The family of functions $\{u_k\}_{k=0}^{\infty}$ is known as "the Haar system". Observe that $\{u_k\}_{k=0}^{\infty} \subset L^2([0,1])$, notice that since $0 \leq 2^j t - i \leq 1$ implies $2^{-j}i \leq t \leq 2^{-j}(i+1)$ we obtain $supp\, u_{2^j+i} = [2^{-j}i, 2^{-j}(i+1)]$. From wavelet theory it is standard to denote, for each fix j, with V_j the subspace of $L^2([0,1])$ of piecewise constant functions on each interval $[2^{-j}i, 2^{-j}(i+1)]$, $i = 0, \ldots, 2^j - 1$. It is clear that

$$V_j \subset V_{j+1}$$

and

$$W_j = span\{u_{2^j+i} : i = 0, \ldots, 2^j - 1\} \subset L^2([0,1])$$

is the orthogonal complement of V_j in V_{j+1}. In other words,

$$V_{j+1} = W_j \oplus V_j,$$

moreover, for $j \geq 1$

$$V_j = V_0 \oplus \bigoplus_{k=0}^{j-1} W_k.$$

From this last expression it follows that $\{u_k\}_{k=0}^{2^j-1}$ is an orthonormal basis of V_j. Having in mind the definition of V_j, it is known that $\overline{\cup_{j=0}^{\infty} V_j} = L^2([0,1])$ hence $L^2([0,1]) = \overline{V_0 \oplus \bigoplus_{j=0}^{\infty} W_j}$ and then $\{u_k\}_{k=0}^{\infty}$ is an orthonormal basis of $L^2([0,1])$.

It is clear that $\{u_k\}_{k\geq 0}$ is an H-system, in fact is also a basis of $L^2[0,1]$. Moreover, if we choose $m = \infty$ or $m = 2^J - 1$, for some $J \geq 1$, we have that $\{u_k\}_{k=0}^{m}$ is also a Haar-system. In fact, for all j we have that the atoms of $\sigma(u_0, \ldots, u_{2^j-1})$, are

$$\{[2^{-j}i, 2^{-j}(i+1)]\}_{i=0}^{2^j-1}.$$

2. Haar-Systems for the binomial model:

Let S the price of an stock and $t_0, t_1, \ldots, t_n$ the trading dates. The price $S_{t_i} = S(t_i)$, $i = 0, 1, \ldots, n$, varies according to the rule

$$S_{t_{i+1}} = S_{t_i} H_{i+1}, \; i = 0, 1, \ldots, n,$$

where the H_i's are independent random variables such that

$$H_i = \begin{cases} U & \text{with probability } p \\ D & \text{with probability } q \end{cases},$$

where $0 < D < 1 < U$ and $p + q = 1$. The situation can be written in terms of the probability space $(\Omega, \mathcal{B}, P)$, where $\Omega := \{w : \{t_0, t_1, \ldots, t_n\} \to \{U, D\}\}$, $\mathcal{B} := \mathcal{P}(\Omega)$ and P the corresponding product probability measure. Then $S : \Omega \times \{t_0, t_1, \ldots, t_n\} \to \mathbb{R}$ and $S_t(\omega) := S(w,t) = S_0 \prod_{t_i \leq t} w(t_i)$.
Let us consider the sets $A_{j,i}$, $0 \leq j \leq n-1$ and $0 \leq i \leq 2^j - 1$ defined by $A_{0,0} = \Omega$ and

$$A_{j+1,2i+1} = A_{j,i} \cap \{\omega(t_{j+1}) = U\}, \; A_{j+1,2i} = A_{j,i} \cap \{\omega(t_{j+1}) = D\}. \tag{2.3}$$

From independence, it is clear that $P(A_{j+1,2i}) = q\ P(A_{j,i})$ and $P(A_{j+1,2i+1}) = p\ P(A_{j,i})$, consequently $P(A_{j,i}) = p^{i_0} \cdots p^{i_j} q^{1-i_0} \cdots q^{1-i_j}$ where $i = \sum_{l=0}^{j} i_l 2^l$ is the binary representation of i (with $j+1$ digits).
Define now, for $j = 0, ..., n-1$; $i = 0, ..., 2^j - 1$ the normalized functions

$$
\begin{aligned}
&u_0 \equiv 1, \\
&u_{2^j+i} = \tfrac{1}{\sqrt{P(A_{j,i})}} (\sqrt{\tfrac{p}{q}}\ 1_{A_{j+1,2i}} - \sqrt{\tfrac{q}{p}}\ 1_{A_{j+1,2i+1}}),
\end{aligned} \tag{2.4}
$$

see (3.2) for the normalization.
It is clear, by definition of the functions u_k, that the system $\{u_k\}_{0 \le k \le 2^n - 1}$ verifies trivially the items (1) and (3) of Proposition 1. Having in mind also the choice of the sets $A_{j,i}$, it follows that for any $j \ge 0$ and $i = 0, ..., 2^j - 1$ fixed, the atoms generating $\sigma(u_0, \ldots, u_{2^j+i})$ are

$$\{A_{j+1,0}, \ldots, A_{j+1,2i}, A_{j+1,2i+1}, A_{j,i+1}, \ldots, A_{j,2^j-1}\},$$

which are 2^j+i+1, thus (2) of the referred Proposition also holds, and consequently the system forms an H-system for $L^2(\Omega, \mathcal{B}, P)$. Now observing that for each $j \ge 0$ the atoms of $\sigma(u_0, \ldots, u_{2^j-1})$, are $\{A_{j+1,i} : i = 0, ..., 2^{j+1} - 1\}$ it follows that $\{u_k\}_{0 \le k \le 2^n - 1}$ is also a Haar-system. Particularly the sub-system $\{u_0, \ldots, u_{2^j-1}\}$ is an orthonormal basis of $L^2(\Omega, \mathcal{F}_{t_j}, P)$, where $\mathcal{F}_{t_j} = \sigma(S_{t_0}, ...S_{t_j})$.

3. A non Haar-System for the binomial model:
We are now going to construct another H-system for the binomial model. This time associated with a particular partition of the final σ-algebra $\sigma(S_{t_n})$. Let J be the smallest integer such that $n+1 \le 2^J$, then for $0 \le j \le J$ and $0 \le i \le 2^j - 1$ we define the sets $A_{j,i}$, as follows. For $i \ne 0$,

$$A_{j,i} = \{\omega \in \Omega : \frac{i}{2^j} < \frac{1}{n}|\omega|_U \le \frac{i+1}{2^j}\} \tag{2.5}$$

whenever this set is not empty, and for $i = 0$

$$A_{j,0} = \{\omega \in \Omega : 0 \le \frac{1}{n}|\omega|_U \le \frac{1}{2^j}\} \tag{2.6}$$

where $|\omega|_U$ is the number of t_i's such that $\omega(t_i) = U$. We calculate the probabilities

$$P(A_{j,i}) = \sum_{\frac{i}{2^j} < \frac{s}{n} \le \frac{i+1}{2^j}} \binom{n}{s} p^s q^{n-s} \text{ for } i \ne 0$$

and

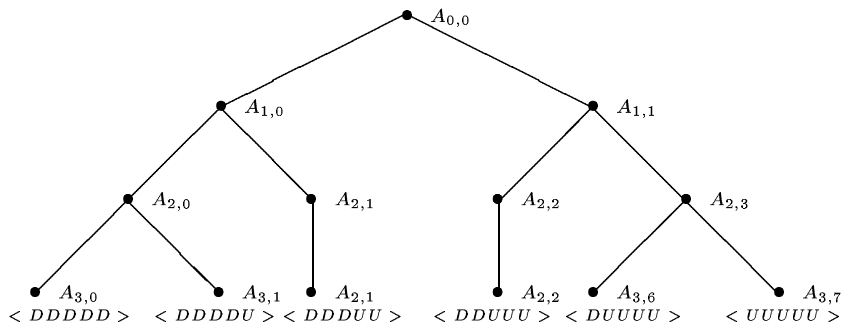
$$P(A_{j,0}) = \sum_{0 \le \frac{s}{n} \le \frac{1}{2^j}} \binom{n}{s} p^s q^{n-s}.$$

It is important to observe that $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$ or $A_{j,i} = A_{j+1,i}$. The corresponding H-system is given, through Theorem 1 and Remark 2 (to be introduced shortly), by

$$
\begin{aligned}
&v_0 \equiv 1, \\
&v_{j,i} = \tfrac{1}{\sqrt{P(A_{j,i})}} (\sqrt{\tfrac{P(A_{j+1,2i+1})}{P(A_{j+1,2i})}} 1_{A_{j+1,2i}} - \sqrt{\tfrac{P(A_{j+1,2i})}{P(A_{j+1,2i+1})}} 1_{A_{j+1,2i+1}})
\end{aligned} \tag{2.7}
$$

if $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$. It results in a Haar system only if $n = 2^J - 1$.

The tree illustration below corresponds to the H-system with $n = 5$. We have labelled the atoms of the final σ-algebra to clarify the situation, with e.g. $<DDDDU> = \{(D,D,D,D,U),(D,D,D,U,D),(D,D,U,D,D),(D,U,D,D,D),(U,D,D,D,D)\}$.



Observe that $A_{3,2} = A_{2,1}$ and $A_{3,4} = A_{2,2}$ because $A_{3,3} = A_{3,5} = \emptyset$.

4. The H-system associated to the Willinger´s almost sure approximation scheme:

W. Willinger introduces in [19] and [20] an approximation scheme for random variables and processes, we briefly indicate here that this scheme also gives an example of H-system specially designed to represent a given random variable. Let $(\Omega, \mathcal{F}, P)$ be a probability space and X be an integrable random variable. Let us consider the sets $A_{j,i}$, $0 \leq j$,$i = 0, \ldots, 2^j - 1$, defined by $A_{0,0} \equiv \Omega$ and

$$A_{j+1,2i+1} \equiv \{\omega \in A_{j,i} : X(\omega) \leq \frac{1}{P(A_{j,i})} \int_{A_{j,i}} X \, dP\}, \tag{2.8}$$

$$A_{j+1,2i} \equiv \{\omega \in A_{j,i} : X(\omega) > \frac{1}{P(A_{j,i})} \int_{A_{j,i}} X \, dP\}. \tag{2.9}$$

if $P(A_{j,i}) > 0$. We organize the family of sets $A_{j,i}$ by levels defining $\mathcal{R}_0 \equiv \{\Omega\}$ and recursively, for $A_{k,i} \in \mathcal{R}_j$ with $P(A_{k,i}) > 0$, $A_{k+1,2i}$, $A_{k+1,2i+1}$ will belong to $\mathcal{R}_{j+1}$, but if $P(A_{k,i}) = 0$, $A_{k,i}$ will be in $\mathcal{R}_{j+1}$ too. The sequence of partitions $\mathcal{R} \equiv \{\mathcal{R}_j\}_{j \geq 0}$ is an example of a special type of sequence of partitions, the weak-dyadic ones. Theorem 1 below, associates an H-system $\{u_{j,i}\}$ to this sequence of partitions. We observe three facts (we rely on notation from (3.1)): $\sigma(X) = \sigma(\cup_j \mathcal{R}_j)$, $\mathbb{E}(X \mid \sigma(\mathcal{R}_J)) = \sum_{0 \leq j \leq J} \sum_{(k,i) \in I_j} \langle X, u_{k,i} \rangle \, u_{k,i}$ and $X = \lim_{J \to \infty} \mathbb{E}(X \mid \sigma(\mathcal{R}_J))$ almost everywhere.

5. An H-System in the Black-Sholes model:

This example describes how to construct a basic class of Haar systems associated to the Black-Scholes model. It will follow that these systems can be used to approximate a general class of options of European type. The building block for the Black-Scholes model is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, Q)$ with filtration $(\mathcal{F}_t)_{T_0 \leq t \leq T}$. The splitting of atoms will be performed using the Brownian motion increments. The price process under the risk neutral measure P is given by $S_t : \Omega \to \mathbb{R}$, $T_0 \leq t \leq T$,

$$S_t(\omega) = S_{T_0} \exp(\nu(t - T_0) + \sigma\sqrt{(t - T_0)} \, W_t(\omega)),$$

where $\nu = (r - \sigma^2/2)$, and we have used the Gaussian random variables $W_t \sim \mathcal{N}(0,1)$ which are defined on $(\Omega, \mathcal{F}_t, P)$.

The construction will be based on two parameters, the first parameter n_T will turn out to be the number of transaction dates during the period $[T_0, T]$ and the second set of parameters $j_1, \ldots, j_{n_T}$ will be the scale or space discretizations associated to each trading date. For simplicity, the splitting of atoms will be in pieces of equal probability, this constrain can be easily removed. For convenience, we first introduce a "purely static" Haar system, considering $n_T = 1$, which is applicable to path *independent* European options. This system will be the building block for the more general construction with $n_T \geq 1$. Therefore, we first concentrate on the sigma algebra $\sigma(S_T) = S_T^{-1}(\mathcal{B}(0,\infty))$, due to $\sigma(S_T) = \sigma(S_T^{-1}((a_1, a_2]), 0 < a_1 < a_2 < \infty)$, the following equation specifies P on $\sigma(S_T)$, let $B = S_T^{-1}((a_1, a_2))$

$$P(B) = \frac{1}{\sigma\ \sqrt{2\pi(T-T_0)}} \int_{a_1}^{a_2} \exp\left[\frac{-\left(\ln(\frac{s}{S_{T_0}}) - \nu(T-T_0)\right)^2}{2\ \sigma^2(T-T_0)}\right] \frac{ds}{s}$$

From our previous notation, $W_T : \Omega \to \mathbb{R}$

$$P(W_T^{-1}(A)) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{y^2}{2}} dy,$$

for any Borel subset $A \subset \mathbb{R}$. This equation gives P on $\sigma(W_T) = W_T^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_T$, clearly, $\sigma(S_T) = \sigma(W_T)$. Denote the cumulative standard normal distribution by

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy.$$

Given an integer j, define the numbers $-\infty = c_0^j < c_1^j < \ldots < c_{2^j}^j = \infty$ such that

$$\Phi(c_{i+1}^j) - \Phi(c_i^j) = \frac{1}{2^j}, \text{ for all } i = 0, \ldots, 2^j - 1.$$

Whenever encountered, the inequality $\leq \infty$ should be interpreted to mean $< \infty$. We define the binary splitting of atoms inductively by setting $A_{0,0} = \Omega$ and for given j consider $0 \leq i \leq 2^j - 1$,

(2.10)

$$A_{j+1,2i} = \{w \in A_{j,i} |\ c_{2i}^{j+1} < W_T(\omega) \leq c_{2i+1}^{j+1}\} = \{w |\ c_{2i}^{j+1} < W_T(\omega) \leq c_{2i+1}^{j+1}\},$$

$$A_{j+1,2i+1} = \{w \in A_{j,i} |\ c_{2i+1}^{j+1} < W_T(\omega) \leq c_{2i+2}^{j+1}\} = \{w |\ c_{2i+1}^{j+1} < W_T(\omega) \leq c_{2i+2}^{j+1}\}.$$

Note that $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$, therefore we have defined a dyadic sequence of partitions $\mathcal{P} = \{\mathcal{P}_j\}_{j\geq 0}$ with $\mathcal{P}_j = \{A_{j,i}\}, i = 0, \ldots, 2^j - 1$ (see Definition 6 below), where the atoms satisfy

$$P(A_{j,i}) = \frac{1}{2^j}.$$

To have an alternative perspective on these atoms, notice they correspond to partitioning the range of S_T as follows,

$$A_{j,i} = \{w |\ a_i^j < S_T(\omega) \leq a_{i+1}^j\} \text{ where } i = 0, \ldots, 2^j - 1, \tag{2.11}$$

the real numbers $0 = a_0^j < \ldots < a_{2^j-1}^j < a_{2^j}^j = \infty$, satisfy

$$a_i^j = S_{T_0} \exp\left(c_i^j\ \sigma\sqrt{(T-T_0)} + \nu(T-T_0)\right).$$

Setting $m = 2^j$ and $\mathcal{B}_m = \sigma(\{A_{j,i} : i = 0, \ldots, m-1\})$ gives $\mathcal{B}_\infty = \sigma(\cup_{m\geq 0}\mathcal{B}_m) = \sigma(S_T)$. Being this a situation similar to the one in Example 3.

Associated with the sequence of partitions $\mathcal{P}$, we have the functions

$$u_0 \equiv 1,$$
$$u_{2^j+i} = a_{j,i}\mathbf{1}_{A_{j+1,2i}} + b_{j,i}\mathbf{1}_{A_{j+1,2i+1}},$$

defined for $0 \leq j$, and $i = 0, \ldots, 2^j - 1$, with $a_{j,i}$ and $b_{j,i}$ as in (3.3) and (3.4). It follows from Theorem 1 that $\{u_j\}_{j=0}^{2^J-1}$ is a Haar system capable of approximating any random variable in $L^2(\Omega, \sigma(S_T), P)$, choosing a sufficiently large J.

We are now ready to describe the construction of a finite Haar system for an arbitrary $n_T \geq 1$. The idea is simply to construct a Haar dyadic system by a concatenation of several Haar systems, each of them analogous to the case $n_T = 1$ but this later one now restricted to smaller time intervals. Given an arbitrary sequence of times $T_0 = t_0 < t_1 < \ldots < t_{n_T-1} < t_{n_T} = T$, we consider the Brownian motion increments $\sqrt{t_{i+1} - t_i}\; W_{t_i,t_{i+1}}$ where the random variables $W_{t_i,t_{i+1}} \sim \mathcal{N}(0,1)$ are independent. Fix a corresponding sequence of scales $\{j_i = j_{t_i}\}_{i=1}^{n_T}$, we will define the splitting of atoms on stages according to the time intervals $\{t_i, t_{i+1}\}$. For the first stage $\{t_0, t_1\}$ we define the binary splitting of atoms inductively by setting $A_{0,0} = \Omega$ and for $0 \leq j < j_1$, $i = 0, \ldots, 2^j - 1$, $A_{j+1,i}$ as in (2.10), using W_{t_0,t_1} instead of W_T.

For the second stage $\{t_1, t_2\}$, and as a model for the subsequents, consider $0 \leq j < j_2$ and $i = 0, \ldots, 2^{j_1+j} - 1$ as usual, let p and $0 \leq q < 2^{j+1}$ be respectively the quotient and residue in the integer division of i by 2^{j+1}, then define inductively the sets

$$\begin{aligned}
A_{j_1+j+1,2i} &= \{w \in A_{j_1+j,i} |\; c_{2q}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+1}^{j+1}\} \\
&= \{w \in A_{j_1,p} |\; c_{2q}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+1}^{j+1}\} \\
A_{j_1+j+1,2i+1} &= \{w \in A_{j_1+j,i} |\; c_{2q+1}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+2}^{j+1}\} \\
&= \{w \in A_{j_1,p} |\; c_{2q+1}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+2}^{t+1}\}.
\end{aligned}$$

Notice that $P(A_{j_1+1,i}) = 1/2^{j_1+1}$ by independence of W_{t_0,t_1} and W_{t_1,t_2}. The completion of a generic stage $\{t_k, t_{k+1}\}$ is done exactly as in the previous described stage replacing j_1 by $J_k = j_1 + \ldots + j_k$ and W_{t_1,t_2} by $W_{t_k,t_{k+1}}$. This induction must be continued until $k + 1 = n_T$.

We have defined, once again, a dyadic sequence of partitions (see Definition 6 below) $\{\mathcal{P}_j\}_{j\geq 0}$ with $\mathcal{P}_j = \{A_{j,i}\}, i = 0, \ldots, 2^j - 1$ and consequently, following the steps in the proof Theorem 1, there is a Haar system $\{u_j\}_{j=0}^{2^J-1}$ associated with it.

3. Constructions and Properties of H-Systems

This section introduces some elementary properties of H-sytems and partitions. It represents the computational core of our approach, we introduce most of the notation to be used in the rest of the paper as well as the main constructions. The reader who wishes to see financial applications first should then read the next two sections first and then refer to the present section as needed.

We have shown in the examples above that sequences of partitions of Ω and H-systems are mutually associated. A natural and computationally useful binary sequence of partitions are the dyadic ones.

Definition 6. *Let $\mathcal{P} := \{\mathcal{P}_j\}_{j\geq 0}$ be a sequence of partitions of Ω, we say that it is a dyadic sequence if*

$$\mathcal{P}_j = \{A_{j,i}\}_{i=0}^{2^j-1} \quad \textit{and} \quad A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}.$$

We observe that the Haar systems in Examples 1 and 2 induce a sequence of dyadic partitions, naturally associated with the multi-resolution analysis of wavelets theory, [9]. Example 3 shows H-systems which produce a sequence of binary partitions that is not dyadic. We call this type of sequence of partitions *weak dyadic.*

Definition 7. *We say that a binary sequence of partitions of Ω, $\mathcal{R} := \{\mathcal{R}_j\}_{j\geq 0}$, is a* weak dyadic sequence *if $\mathcal{R}_0 = \{\Omega\}$ and for $j \geq 1$, $\mathcal{R}_j$ satisfies:*

$A \in \mathcal{R}_j$ if

$$\textit{there exist another } A' \in \mathcal{R}_j \textit{ such that } A \cup A' \in \mathcal{R}_{j-1}$$

or

$$A \in \mathcal{R}_k \textit{ for all } k \geq j-1.$$

Notice that if $A \in \mathcal{R}_k$ for all $k \geq j-1$, the index of A in $\mathcal{R}_{j-1}$ is preserved. Therefore, we obtain sequences $\mathcal{R}_j$ where $A_{k,i} \in \mathcal{R}_j$ if and only if $k = j$ or otherwise $A_{k,i} \in \mathcal{R}_r$ for all $r \geq j-1$, in which case $k = j-1$. In other words, we collect, in $\mathcal{R}_j$, all atoms with the same scale parameter and also include those atoms which will not be further split. To this type of partitions we will associate a Multiresolution Analysis Algorithm (MRA) (see Appendix A) in complete analogy with wavelet theory and, in particular, allows the computation of inner products and the corresponding approximations to be organized by the scale parameter.
Observe that $\mathcal{R}_j$ can have at most 2^j members, and if $A_{k,i} \in \mathcal{R}_j$ then $k \leq j$ and $0 \leq i \leq 2^k - 1$. We remark that all information about the splitting of atoms is stored in this indexation. The figure displayed in Example 3 will clarify this indexing.
The following sets of indexes will be useful later. Consider $j \geq 0$ and let

$$\begin{aligned} I_j &\equiv \{i : A_{j,i} \in \mathcal{R}_j \text{ and } A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}\}, \text{ and} \\ K_j &\equiv \{(k,i) : A_{k,i} \in \mathcal{R}_j \}. \end{aligned} \tag{3.1}$$

Theorem 1. *Every H-system induces naturally a sequence of weak dyadic partitions and reciprocally.*

Proof. Let $\{u_k\}_{k\geq 0}$ be an H-system and $A_{0,0} = \Omega$. We define recursively the following sequence of partitions.

$$\mathcal{R}_0 = \{A_{0,0}\}.$$

Assuming $\mathcal{R}_j$ has been defined, we will generate $\mathcal{R}_{j+1}$. Consider a generic atom $A_{k,i} \in \mathcal{R}_j$, by Corollary 1 it is enough to consider the following cases:

- For $A_{k,i} \in \mathcal{R}_j$ with $k < j$ we add $A_{k,i}$ to $\mathcal{R}_{j+1}$
- If $k = j$ and $A_{j,i}$ is not the support of any u_r, we add $A_{j,i}$ to $\mathcal{R}_{j+1}$.
- If $k = j$ and $A_{j,i} = u_r$ for some u_r. Then we add

$$A_{j+1,2i} = u_k^{-1}((-\infty,0)) \text{ and } A_{j+1,2i+1} = u_k^{-1}((0,\infty))$$

to $\mathcal{R}_{j+1}$.

Clearly this sequence of partitions is weak dyadic. Reciprocally, let $\mathcal{R}$ be a sequence of weak dyadic partitions, We are going to define a family of Haar functions $\{\psi_{j,i}\}$, associated with $\mathcal{R}$. Consider $A_{j,i} \in \mathcal{R}_j$ such that $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$, let $\psi_{j,i}$ be defined on Ω by

$$\psi_{j,i}(\omega) = \begin{cases} a_{j,i} & \text{if} \quad w \in A_{j+1,2i}, \\ b_{j,i} & \text{if} \quad w \in A_{j+1,2i+1} \text{ and,} \\ 0 & \text{if} \quad w \notin A_{j,i}. \end{cases} \tag{3.2}$$

Where $a_{j,i}$ and $b_{j,i}$ are chosen requiring that $\psi_{j,i}$ is a Haar function. The above equations define (up to a sign) $\psi_{j,i}(\omega)$ for all $w \in \Omega$, indeed we choose

$$a_{j,i} = \sqrt{\frac{P(A_{j+1,2i+1})}{P(A_{j+1,2i})P(A_{j,i})}} \tag{3.3}$$

and

$$b_{j,i} = -\sqrt{\frac{P(A_{j+1,2i})}{P(A_{j+1,2i+1})P(A_{j,i})}}. \tag{3.4}$$

Give the natural order to the set $\mathcal{N} = \{2^j + i : j \geq 0, i \in I_j\}$ and let π be an order preserving isomorphism from $\mathcal{N}$ to an integer interval $[1, N]$ or $\mathbb{N}$. Defining

$$v_{\pi(2^j+i)} = u_{2^j+i} \equiv \psi_{j,i}$$

and $v_0 \equiv \phi_{0,0}(\omega) \equiv \mathbf{1}_\Omega(\omega)$, the resulting sequence $\{v_l\}_{l \geq 0}$ is orthonormal and trivially verifies (1), (2) and (3) of Proposition 1, thus, it is an H-system. ☐

Remark 1. *The above theorem applied to a dyadic sequence of partitions $\mathcal{P}$ implies that $\mathcal{P}_j = \{A_{j,i}\}_{i=0}^{2^j-1}$ with $I_j = [0, 2^j - 1]$, the system of functions $\{u_k\}_{k \geq 0}$ defined by*

$$u_{2^j+i} \equiv \psi_{j,i}$$

is actually a Haar-system. This holds because a given integer $k \geq 1$ can be written as $k = 2^{j_k} + i_k$ where j_k is the maximum integer satisfying $2^{j_k} \leq k$, resulting in consequence $i_k \in I_{j_k} = [0, 2^{j_k} - 1]$. Moreover, $\mathcal{P}_j$ is the set of atoms of $\sigma(u_0, \ldots, u_{2^j-1})$.

Remark 2. *It is also clear that if we start with the sequence of partitions $\mathcal{R}$ induced by an H-system $\{v_k\}_{k \geq 0}$, then the H-system built in Theorem 1 is a rearrangement $\{v_{\pi k}\}_{k \geq 0}$ of $\{v_k\}_{k \geq 0}$.*

Let $\mathcal{R} := \{\mathcal{R}_j\}_{j \geq 0}$ be a sequence of weak dyadic partitions of Ω and the sequence of functions $\{\psi_{j,i}\}$ associated with $\mathcal{R}$ in (3.2). We will now introduce the natural orthonormal basis of characteristic functions at level j, for each $A_{k,i} \in \mathcal{R}_j$, let

$$\phi_{k,i} \equiv \frac{\mathbf{1}_{A_{k,i}}}{\sqrt{P(A_{k,i})}}.$$

Given a random variable X, our aim is to study the relationship between the coefficients in this basis, which represent samples at level j, with the coefficients in the H-system $\{\psi_{j,i}\}$.
Taking $j \geq 0$, $\{\phi_{k,i}\}_{(k,i) \in K_j}$ is an orthonormal basis of the subspaces V_j of piecewise constant functions on the atoms of $\mathcal{R}_j$. The $\phi_{k,i}$ correspond, in our setting, to the scaled and translated scale functions from wavelet theory. Similarly the $\psi_{j,i}$ correspond to the wavelets [9]. We have the simple, but relevant, relations:

$$\phi_{j,i} = \sqrt{\frac{p_{j+1}[2i]}{p_j[i]}}\phi_{j+1,2i} + \sqrt{\frac{p_{j+1}[2i+1]}{p_j[i]}}\phi_{j+1,2i+1} \tag{3.5}$$

and

$$\begin{aligned} \psi_{j,i} &= a_{j,i}\mathbf{1}_{A_{j+1,2i}} + b_{j,i}\mathbf{1}_{A_{j+1,2i+1}} \\ &= a_{j,i}\sqrt{p_{j+1}[2i]}\phi_{j+1,2i} + b_{j,i}\sqrt{p_{j+1}[2i+1]}\phi_{j+1,2i+1} \end{aligned} \tag{3.6}$$

where $a_{j,i} = \sqrt{\frac{p_{j+1}[2i+1]}{p_{j+1}[2i]p_j[i]}}$ and $b_{j,i} = -\sqrt{\frac{p_{j+1}[2i]}{p_{j+1}[2i+1]p_j[i]}}$ were calculated in (3.3) and (3.4) respectively, and

$$p_j[i] \equiv P(A_{j,i}).$$

Observe that since

$$\sqrt{\frac{p_{j+1}[2i+1]}{p_j[i]}}a_{j,i}\sqrt{p_{j+1}[2i]} - \sqrt{\frac{p_{j+1}[2i]}{p_j[i]}}b_{j,i}\sqrt{p_{j+1}[2i+1]} = 1 \neq 0,$$

$\{\psi_{j,i}, \phi_{j,i}\}$ and $\{\phi_{j+1,2i}, \phi_{j+1,2i+1}\}$ span the same 2-dimensional subspace. Thus $\{\phi_{0,0}\} \cup \{\psi_{k,i}\}_{0\leq k\leq j-1, i\in I_k}$ is a basis of $L^2(\Omega, \sigma(\mathcal{R}_j), P) = V_j$, and moreover it is also orthonormal as the basis $\{\phi_{k,i}\}_{(k,i)\in K_j}$.

For $X \in L^2(\Omega)$ and $j \geq 0$, for simplicity set

$$X_j \equiv X_{\sigma(\mathcal{R}_j)} \equiv \mathbf{E}(X|\sigma(\mathcal{R}_j)).$$

Then we have the following expansions

$$X_j(\omega) = c_0[0]\ \phi_{0,0}(\omega) + \sum_{k=0}^{j-1}\sum_{i\in I_k} d_k[i]\ \psi_{k,i}(\omega) = \sum_{(k,i)\in K_j} c_k[i]\ \phi_{k,i}(\omega) \tag{3.7}$$

where

$$c_k[i] = \langle X_j, \phi_{k,i}\rangle \quad \text{and} \quad d_k[i] = \langle X_j, \psi_{k,i}\rangle.$$

Given that the conditional expectation X_j of X is constant on each $A_{k,i}$, we have that for $w \in A_{k,i}$

$$c_k[i] = \langle X_j, \phi_{k,i}\rangle = \frac{1}{\sqrt{p_k[i]}}\int_{A_{k,i}} X_j dP = \frac{1}{\sqrt{p_k[i]}}\int_{A_{k,i}} X dP = \langle X, \phi_{k,i}\rangle. \tag{3.8}$$

Analogously, we have that $d_k[i] = \langle X, \psi_{k,i}\rangle$. Moreover we can state the following

Proposition 2. *Given $X \in L^2(\Omega, \mathcal{B}, P)$ and a sequence $\mathcal{R} = \{\mathcal{R}_j\}_0^J$ of weak dyadic partitions of Ω, for each $j' < j \leq J$, the following holds*

$$X_j = X_{j'} + \sum_{k=j'}^{j-1}\sum_{i\in I_k} d_k[i]\psi_{k,i} \tag{3.9}$$

and

$$\sum_{(k,i)\in K_j} c_k^2[i] = \sum_{(k,i)\in K_{j'}} c_k^2[i] + \sum_{k=j'}^{j-1}\sum_{i\in I_k} d_k^2[i]. \tag{3.10}$$

Proof. For each $j < J$ we have that $V_j = span\ \{\phi_{k,i} : (k,i) \in K_j\}$, let $W_j \equiv span\ \{\psi_{j,i} : i \in I_j\}$. It is clear that $X_j \in V_j$ and $V_{j-1} \subset V_j$.
By definition $\psi_{j-1,i} \in V_j \cap V_{j-1}^{\perp}$, also as we have noted before, $\{\psi_{j-1,i}, \phi_{j-1,i}\}$ and $\{\phi_{j,2i}, \phi_{j,2i+1}\}$ span the same subspace, thus $V_j = V_{j-1} \oplus W_{j-1}$. This is one reason we have used the classical wavelet notation.

Since $\mathbf{E}(X|\sigma(\mathcal{R}_j))$ and $\mathbf{E}(X|\sigma(\mathcal{R}_{j-1}))$ are the orthogonal projections of X onto V_j and V_{j-1} respectively, we see that $\mathbf{E}(X|\sigma(\mathcal{R}_j)) - \mathbf{E}(X|\sigma(\mathcal{R}_{j-1}))$ is the orthogonal projection of X onto W_{j-1} and then we have the expansion

$$X_j = X_{j-1} + \sum_{i \in I_{j-1}} d_{j-1}[i]\psi_{j-1,i},$$

from which (3.9) follows inductively. Equation (3.10) is a direct consequence of (3.7) and (3.9). □

The precedent proposition, with the aid of (3.5) and (3.6), also gives a relation between the coefficients $c_j[i]$ and $d_j[i]$, which permit us to have expansions on all coarser "levels j", starting from the correspondent to $\{\phi_{k,i}\}_{(k,i)\in K_J}$ on a finer level J. These are the fundamentals of the multiresolution algorithm for H-systems, it is an adaptation of the well known algorithm for wavelet theory given for S. Mallat [15] to our probabilistic setting. This algorithm produces a relation between the samples of X, $x_k[i] = X_j(\omega)$, $\omega \in A_{k,i}$, for $(k,i) \in K_j$, and the coefficients $d_k[i]$. This algorithm is described in Appendix A.

4. Haar Hedging

This section introduces two hedging strategies, *Haar hedging* and *binary hedging* which will be denoted by HΠ and BΠ respectively. HΠ is based on the lef side of (3.7) and BΠ is based on its right side. They are both realized financially via *binary options* whenever these options are available for trade. For its financial implementation, HΠ will require short selling of some of the binary options. The values of the constructed approximating portfolios are the same but the financial transactions differ, in particular the volume of transactions are not equal. We define the *number of transactions* as the number of binary options needed to implement the given portfolio, both portfolios will have the same number of transactions. The new hedging strategies may require a potentially large number of transactions, this crucial problem is tackled in Section 5. We start with two revealing examples before the formal developments.

4.1. Basic Examples.

Here we describe, for the sake of clarity, two of the simplest examples of Haar hedging strategies applied to a path independent European option X. The new notions defined below will be defined formally and in a more general context in Section 4.2.

Recall Example 5 from Section 2, consider first the Haar system with $n_T = 1$, fixed j and define $n = 2^j$, the atoms at this finest scale are given by (2.11), namely, $A_{j,i} = \{w|\ a^j_i < S_T(\omega) \leq a^j_{i+1}\}$. The *binary option* with payoff $\mathbf{1}_{A_{j,i}}$ is available for trading either as a *double digital option*, [21] pages 409-410, or as a linear combination, involving short selling, of two cash or nothing options. We will use, unambiguously, the notation $\mathbf{1}_A$ for the payoff of the binary option as well for the actual option itself when speaking, for example, of a portfolio of binary options. Recall that X_j, given by (3.7), is a simple function on the atoms $A_{j,i}$. We will associate $\mathbf{1}_{A_{0,0}} = \mathbf{1}_\Omega$ with a bank account. The *binary hedging portfolio strategy* consists on purchasing an amount

$$X_j(\omega) = \frac{\langle X, \mathbf{1}_{A_{j,i}}\rangle}{P(A_{j,i})}, \text{ where } w \in A_{j,i}, \tag{4.1}$$

of each binary option $\mathbf{1}_{A_{j,i}}$, $i = 0, \ldots, 2^j - 1$. We denote this portfolio with BΠ. The cost of setting up this portfolio is

$$V_{T_0}(\mathrm{B\Pi}) = e^{-t(T-T_0)} \sum_{i=0}^{2^j-1} \langle X, \mathbf{1}_{A_{j,i}} \rangle = e^{-t(T-T_0)} \int_\Omega X(\omega)\ dP(\omega) = V_{T_0}(X),$$

hence the proposed binary hedging strategy is self financing. This means, of course, that only the value of the option X, namely $V_{T_0}(X)$, is needed to set BΠ. Similarly, the *Haar hedging portfolio strategy* consists in purchasing an amount

$$X_j(\omega) - \mathbf{E}(X), \text{ where } w \in A_{j,i}, \tag{4.2}$$

of each binary option $\mathbf{1}_{A_{j,i}}$, $i = 0, \ldots, 2^j-1$. Moreover, we also invest $e^{-r(T-T_0)}\mathbf{E}(X)$ in the bank account. We denote this portfolio with HΠ, as is the case with the portfolio BΠ, HΠ is also self financing. We formally think of these two strategies as predictable processes, HΠ takes values in $\mathbb{R}^{n+1}$, where the values are given by (4.2) plus the bank account investment, and BΠ takes values in $\mathbb{R}^n$, the values for this portfolio are given by (4.1). These two strategies are *static* because all binary options have to be purchased only at time T_0. Therefore, these hedging portfolios are constant for $T_0 \leq t < T$. The number of transactions in the present case is $n = 2^j$ for both portfolios.
Notice that we have not mentioned the Haar functions u_k, $k = 0, \ldots, 2^j - 1$, which offer an alternative way of setting HΠ. It should be clear that each of the Haar functions can be realized financially purchasing two binary options of the form $\mathbf{1}_{A_{k,i}}$ with $0 \leq k < j$. It is important to add that *short selling* have to be allowed for this to take place.

Remark 3. *Recalling (3.7), we see that the financial implementation of $X_j - \mathbf{E}(X)$ could be done in terms of these Haar functions. It should be clear that this alternative is more inefficient in terms of both, the number and volume of transactions. These observations can be generalized, the idea is that whenever several Haar functions can be combined into a single simple function one should implement this later one directly using the corresponding binary options.*

The case $n_T > 1$ is considered next, for pedagogical reasons we set $j = j_1 = \ldots = j_{n_T} = 1$. Notice that that pointwise convergence is only achieved in the limit $j \to \infty$, hence considering $n_T \to \infty$ alone will not result in strong convergence. Assume that a sequence of rebalancing times $T_0 = t_0 < t_1 < \ldots < t_{n_T} = T$ is given, we will define two hedging portfolios, $\mathrm{H\Pi}_t$ and $\mathrm{B\Pi}_t$, which will be $\mathbb{R}^3$ and $\mathbb{R}^2$ valued (predictable) processes respectively. These processes will be defined to be constant on the intervals $t_{i-1} \leq t < t_i$. Details are only provided for $\mathrm{H\Pi}_t$, setting up $\mathrm{B\Pi}_t$ is completely analogous. Given the premiun $V_{T_0}(X)$, i.e. the risk neutral value of X, we would like to set up, at time t_0, the Haar hedging portfolio $\mathrm{H\Pi}_{t_0}$ so that at time t_1 it has a value equal to $V_{t_1}(\mathrm{H\Pi}_{t_0})(\omega) = e^{-r(T-t_1)}\mathbf{E}(X|u_0, u_1)(\omega)$. The *Haar hedging portfolio strategy* consists then on purchasing an amount

$$e^{-r(T-t_1)}\left(X_1(\omega) - \mathbf{E}(X)\right) = e^{-r(T-t_1)}\langle X, u_1\rangle\ a_{1,0} \text{ where } w \in A_{1,0},$$

of the binary option $\mathbf{1}_{A_{1,0}}$ and $a_{1,0}$ is given by (3.3). We also have to purchase an amount

$$e^{-r(T-t_1)}\left(X_1(\omega) - \mathbf{E}(X)\right) = e^{-r(T-t_1)}\langle X, u_1\rangle\ b_{1,0} \text{ where } w \in A_{1,1},$$

of the binary option $\mathbf{1}_{A_{1,1}}$ and $b_{1,0}$ is given by (3.4). Finally, $\mathbf{E}(X)e^{-r(T-t_0)}$ is invested in the bank account. Given that $\mathbf{E}(u_1|\mathcal{F}_{T_0}) = \mathbf{E}(u_1) = 0$, the cost of setting up $\mathrm{H}\Pi_{t_0}$ equals $V_{T_0}(X)$, hence the first step of the strategy is self financing. Now we describe, inductively, how to rebalance the portfolio $\mathrm{H}\Pi_{t_{s-2}}$ conditionally on the information at a generic time t_{s-1}. Given that the sets $A_{s-1,i}$, $i = 0, \dots, 2^{s-1}-1$, belong to $\mathcal{F}_{t_{s-1}}$ and they form a partition of Ω we may assume there exists $A_{s-1,i_0} \in \{A_{s-1,i}\}_{i=0}^{2^{s-1}-1}$ such that $w \in A_{s-1,i_0}$. Now according to the constructions described in Example 5, Section 2, we know that A_{s-1,i_0} will be split into two sets which we call C and D. More explicitly, we know that

$$C = A_{s-1,i_0} \cap \{w| -\infty < W_{t_{s-1},t_s} \leq 0\} \text{ and } D = A_{s-1,i_0} \cap \{w|0 < W_{t_{s-1},t_s} < \infty\}.$$

For notational simplicity set $A = A_{s-1,i_0}$ and $u_A = a\,\mathbf{1}_C + b\,\mathbf{1}_D$ is the corresponding Haar function on A. As indicated, we proceed inductively assuming that at time t_{s-1} the value of the Haar hedging portfolio is

$$V_{t_{s-1}}(\mathrm{H}\Pi_{t_{s-2}})(\omega) = e^{-r(T-t_{s-1})}\mathbf{E}(X|u_0, \dots, u_{2^{s-1}-1})(\omega) \tag{4.3}$$

at the given w. This capital is used to rebalance $\mathrm{H}\Pi_{t_{s-2}}$ to $\mathrm{H}\Pi_{t_{s-1}}$ by depositing $V_{t_{s-1}}(\mathrm{H}\Pi_{t_{s-2}})(\omega)$ in the bank account and purchasing an amount

$$e^{-r(T-t_s)}\langle X, u_A\rangle\; a \tag{4.4}$$

of the binary option $\mathbf{1}_C$. We also have to purchase an amount

$$e^{-r(T-t_s)}\langle X, u_A\rangle\; b \tag{4.5}$$

of the binary option $\mathbf{1}_D$. Using the fact that u_A has mean equal to 0, we can see that the value at t_{s-1} of the new portfolio $\mathrm{H}\Pi_{t_{s-1}}$ (i.e. $V_{t_{s-1}}(\mathrm{H}\Pi_{t_{s-1}})(\omega)$) equals $V_{t_{s-1}}(\mathrm{H}\Pi_{t_{s-2}})(\omega)$ hence our portfolio rebalancing was self financing. Moreover the value of $\mathrm{H}\Pi_{t_{s-1}}$ at t_s equals

$$V_{t_s}(\mathrm{H}\Pi_{t_{s-1}})(\omega) = e^{-r(T-t_s)}\langle X, u_A\rangle u_A(\omega) + e^{r(t_s-t_{s-1})}V_{t_{s-1}}(\mathrm{H}\Pi_{t_{s-2}})(\omega)) = \tag{4.6}$$
$$e^{-r(T-t_s)}\mathbf{E}(X|u_0, \dots, u_{2^{s-1}-1}, u_A)(\omega),$$

after a comparison with equation (4.3) it follows that we have proven the inductive step, namely (4.6) gives the correct value for the Haar hedging portfolio at the step t_s. We then have proved that

$$V_T(\mathrm{H}\Pi_{t_{n_T-1}})(\omega) = X_{n_T}(\omega).$$

The portfolio $\mathrm{B}\Pi_t$ is completely analogous and can be understood from the case $n_T = 1$, the key difference being the fact that the bank account is not used. More details can be found in the description of the general case in Section 4.2. We remark that both, $\mathrm{H}\Pi$ and $\mathrm{B}\Pi$, have a number of transactions equal to 2^{n_T}.

4.2. **Formal Developments.**

We will work in a frictionless market model $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{T_0 \leq t \leq T}, P)$ with the usual assumptions, we refer to [1] for background. Let $B = (B(t) = e^{rt})$ be the *bond* and a non-negative adapted stochastic process $S = (S_t)_{T_0 \leq t \leq T}$, the *price process.* We assume that P is the risk neutral measure, that is, the discounted price process $(e^{-r(T-t)}S_t)$ is a martingale. Let $\mathcal{R} = \{\mathcal{R}_j\}_{j \geq 0}$ be a sequence of weak-dyadic partitions as described in Definition 7, associated, via Theorem 1, with the H-system $\{\psi_{j,i}\}$ defined on Ω, and an European derivative X in $L^2(\Omega, \sigma(\cup_{j\geq 0}\mathcal{R}_j), P)$.

As a sufficient condition for the atoms in $\mathcal{R}$ to be used in a dynamic hedging portfolio we will impose a natural association between the martingale property of the H-system and a sequence of rebalancing times. In particular, in order to define dynamic hedging strategies, we will use the concept of *time support* of events.

Definition 8. *Let $E \in \mathcal{F}_T$, set $s_E = \sup\{s \in [T_0, T] : E \in \sigma(S_r : r \geq s)\}$ and $t_E = \inf\{t \in [T_0, T] : E \in \mathcal{F}_t\}$. We then say that E is localized to the time interval $[s_E, t_E]$ and call $[s_E, t_E]$ the time support of E. We denote the time support of E by $t - supp(E)$.*

The following definition is an extension to partitions of the notion of time localization of events.

Definition 9. *Let $\mathcal{P} \subset \mathcal{F}_T$ be a partition of Ω. $\mathcal{P}$ is said to be localized (in time) to the interval $[a, b]$ if $t - supp(B) \subset [a, b]$ or $t - supp(B) \subset [T_0, a]$ for each $B \in \mathcal{P}$. Moreover, define the $t - supp(\mathcal{P})$ as the intersection of the all intervals $[a, b]$ such that $\mathcal{P}$ is localized to that interval.*

The definition below is the cornerstone of our dynamic hedging strategy based on H-systems.

Definition 10. *Let $\mathcal{R} = \{\mathcal{R}_j\}_{J \geq j \geq 0}$ be a sequence of weak-dyadic partitions, we say that $\mathcal{R}$ is localized to the time sequence $t_0 = T_0 < \ldots < t_n = T$ if there exist a sequence $j_1 < \ldots < j_n = J$ such that $t - supp(\mathcal{R}_{j_s}) = [t_{s-1}, t_s]$ for $s = 1, \ldots, n$. We call the sequence $j_1 < \ldots < j_n = J$ the levels of localization of $\mathcal{R}$.*

The financial blocks underlying $\mathcal{R}$ are the binary options $\mathbf{B}_{j,i} = (\mathbf{1}_{A_{j,i}}(t) \equiv \mathbf{1}_{[t_{s+1},T]}(t)\mathbf{1}_{A_{j,i}})$, with $j_s \leq j \leq j_{s+1}$ which are acquired at time t_s and reach its maturity at time t_{s+1}. These binary options have payoff $\mathbf{1}_{A_{j,i}}$ at time t_{s+1}.
To have a financial realization of the hedging we are proposing we need to assume $\mathcal{R}$ to be admissible as defined in the next definition.

Definition 11. Assumption on Financial Realization: *The weak dyadic partition $\mathcal{R}$ is called admissible if for any integer j and each atom $A_{k,i} \in \mathcal{R}_j$ the binary options $\mathbf{B}_{k,i}$ are available for trading, in particular, short selling is possible.*

For clarity of exposition, when defining the Haar hedging portfolio, we will further define the *Haar obligations* as follows: $\mathbf{\Psi}_{j,i} = (\mathbf{\Psi}_{j,i}(t) \equiv \mathbf{1}_{[t_{s+1},T]}(t)\psi_{j,i})$, with $j_s \leq j \leq j_{s+1}$ which are obligations at time t_{s+1} that are acquired at time t_s. Obviously, the Haar obligations can be realized in terms of the binary options $\mathbf{B}_{j,i}$, see Remark 3.

Next we will define two hedging strategies via self-financing portfolios, of static and dynamic types, to replicate an European option using H-systems. In fact, we introduce two strategies, HΠ associated to Haar obligations and another BΠ associated to binary options. The examples in Section 4.1 are special cases of the formalism to be introduced.

Haar Hedging Portfolio. $\mathrm{H\Pi}_{\mathcal{R}}(X) = (\mathrm{H\Pi}_{\mathcal{R}}(X)_t)$ will be a predictable, vector valued, stochastic processes constant on the intervals $t_{s-1} \leq t < t_s$. The portfolio $\mathrm{H\Pi}_{\mathcal{R}}(X)_t$ is rebalanced at times t_{s-1} replicating $e^{-r(T-t_s)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))$ for $s = 1, \ldots, n$. As indicated in the Introduction and it is clear from the examples in

Section 4.1, this portfolio approximates fluctuations of the option about its mean value by means of the Haar functions. Taking $n = 1$ the construction gives, as a special case, an example of static hedging. At each time t_{s-1} we will specify how much to invest in the bond and how much to invest in the Haar obligations available at that rebalancing time, this will specify the coordinates of the vector $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$. Here are the coordinates of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ for $t \in [t_0, t_1)$

$$e^{-r(T-t_0)}\mathbf{E}(X) \text{ invested in the bond and}$$

$$e^{-r(T-t_1)}\ d_j[i] \text{ invested in } \Psi_{j,i}\ j = 0, \ldots j_1 - 1, i \in I_j, \tag{4.7}$$

where the coefficients $d_j[i]$ are given by (3.7).
Observe that the purchasing value of this portfolio is $V_{t_0}(\mathrm{H}\Pi_{\mathcal{R}}(X)) = e^{-r(T-t_0)}\mathbf{E}(X)$. The following (inductive) step will be to rebalance the portfolio at time t_{s-1}, assume that at this time we are in the event A_{k_0,i_0} with $(k_0, i_0) \in K_{j_{s-1}}$, and the value of this portfolio is $e^{-r(T-t_{s-1})}\frac{1}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} XdP$. There are two cases to consider, the event is split or not at the next level.

I) In the first case, here are the coordinates of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ for $t \in [t_{s-1}, t_s)$

$$(e^{-r(T-t_{s-1})}\frac{1}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} X\ dP) \text{ invested in the bond and}$$

$$e^{-r(T-t_s)}\ d_j[i] \text{ invested in } \Psi_{j,i}\ j = j_{s-1}, \ldots j_s - 1, i \in I_j^{i_0}, \tag{4.8}$$

where $I_j^{i_0} = I_j \cap [2^{(j-j_{s-1}+1)}i_0, 2^{(j-j_{s-1}+1)}(i_0+1) - 1]$. Recall that the obligations $\Psi_{j,i}$ expire at time t_s.

II) In the second case, we need only to invest

$$e^{-r(T-t_{s-1})}\frac{1}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} X\ dP, \tag{4.9}$$

in the bond, and this specifies the portfolio for all future times i.e. $t \in [t_{s-1}, T)$. The quantity of Haar obligations involved in this dynamic portfolio is at most $2^{j_1} + 2^{j_2-j_1} + \ldots + 2^{j_n-j_{n-1}}$. Now we are in conditions to establish the following theorem.

Theorem 2. *The portfolio* $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ *is self-financing and replicates* $e^{-r(T-t_s)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))$ *at* $s = 1, \ldots, n$.

Proof. We proceed by induction on s. For $s = 1$ the portfolio $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ is given by (4.7) when $t \in [t_0, t_1)$. It is clear from (3.7) that $\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_0}$ replicates $e^{-r(T-t_1)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_1}))$ and is self-financing because $V_{t_0}(\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_0}) = e^{-r(T-t_0)}\mathbf{E}(X)$ since $\mathbf{E}(\psi_{j,i}) = 0$. For convenience, we will use the notation $t^- = t - \epsilon$, $\epsilon > 0$. For the inductive step, at time t_{s-1} we are in some event A_{k_0,i_0} with $(k_0, i_0) \in K_{j_{s-1}}$, and assume

$$V_{t_{s-1}}(\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_{s-1}^-})(\omega) = e^{-r(T-t_{s-1})}\ \mathbf{E}(X|\sigma(\mathcal{R}_{j_{s-1}}))(\omega) = \frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} XdP$$

for $\omega \in A_{k_0,i_0}$. The rebalancing of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ at t_{s-1} is given by (4.8), for all $t \in [t_{s-1}, t_s)$, if A_{k_0,i_0} splits at the next level or by (4.9) with $t \in [t_{s-1}, T]$ if A_{k_0,i_0} does not split any further. The purchasing of $\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_{s-1}}$ is self-financing since the value of the portfolio given by (4.8) or (4.9) is $\frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} XdP$. Consider again case I), and $t = t_s$, by (3.9) and (4.8) we compute

$$V_{t_s}(\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_s^-}) = (\frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} X\ dP)\ e^{r(t_s-t_{s-1})}\ \mathbf{1}_{A_{k_0,i_0}} +$$

$$e^{-r(T-t_s)}\sum_{j=j_{s-1}}^{j_s-1}\sum_{i\in I_j^{i_0}} d_j[i]\ V_{t_s}(\Psi_{j,i}(t_s)) =$$

$$(\frac{e^{-r(T-t_s)}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} X\ dP)\ \mathbf{1}_{A_{k_0,i_0}} + e^{-r(T-t_s)}\sum_{j=j_{s-1}}^{j_s-1}\sum_{i\in I_j^{i_0}} d_j[i]\ \psi_{j,i} =$$

$$e^{-r(T-t_s)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))\ \text{ a. e. on }\ A_{k_0,i_0}.$$

For the case II), we have

$$V_{t_s}(\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_s^-}) = (\frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} X\ dP)\ e^{r(t_s-t_{s-1})}\mathbf{1}_{A_{k_0,i_0}} =$$

$$e^{-r(T-t_s)}\ \mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))\ \text{ a. e. on }\ A_{k_0,i_0}.$$

□

Now, we will present the dynamic strategy $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$. Let $\mathcal{R} = \{\mathcal{R}_j\}$ be a weak dyadic sequence of partitions localized in the sequence of times $t_0 = T_0 < \ldots < t_n = T$, and $X \in L^2(\Omega, \sigma(\cup_{j\geq 0}\mathcal{R}_j), P)$. We will show how to construct a self-financing portfolio $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$ to hedge X.
The portfolio $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$ will be also rebalanced at times $t_0, \ldots, t_{n-1}$, replicating $e^{-r(T-t_s)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))$ for $s = 1, \ldots, n$. We recall that $x_k[i]$ are the coefficients of X in the basis $\{\mathbf{1}_{A_{k,i}} : (k,i) \in K_j\}$, see (A.1).
We formalize $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$ as a vector valued process which is constant on the intervals $t_{s-1} \leq t < t_s$. At time t_0 it is defined, for $t \in [t_0, t_1)$, by specifying its coordinates, namely how much to invest in each of the binary options,

$$e^{-r(T-T_0)}\ x_k[i]\ \mathbf{B}_{k,i} \text{ where } (k,i) \in K_{j_1}. \tag{4.10}$$

The cost of purchasing this portfolio is $V_{t_0}(\mathrm{B}\Pi_{\mathcal{R}}(X)) = e^{-r(T-t_0)}\mathbf{E}(X)$. The inductive step will be to rebalance the portfolio at time t_{s-1}. Assume that at this time we are in the event A_{k_0,i_0} with $(k_0, i_0) \in K_{j_{s-1}}$, and the value of this portfolio is $\frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} XdP$. There are two cases to consider, the event splits or it does not split at the next level. In the first case, for $t_{s-1} \leq t < t_s$, we need to specify the coordinates of $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$, namely,

$$e^{-r(T-t_{s-1})}\ x_k[i]\ \mathbf{B}_{k,i} \text{ where } (k,i) \in K_{j_s}^{i_0}, \tag{4.11}$$

and $K_{j_s}^{i_0} = \{(k,i) \in K_{j_s} : 2^{j_s-j_{s-1}}i_0 \leq i \leq 2^{j_s-j_{s-1}}(i_0+1) - 1\}$.
In the second case, we invest the value of the current portfolio in the bond, namely

$$\frac{e^{-r(T-t_{s-1})}}{P(A_{k_0,i_0})}\int_{A_{k_0,i_0}} XdP, \tag{4.12}$$

and this specifies $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$ for all $t \in [t_s, T)$. In an analogous way to the done for $\mathrm{H}\Pi_{\mathcal{R}}(X)$ is easy to prove that the strategy $\mathrm{B}\Pi_{\mathcal{R}}(X)$ is self-financing and replicates

$e^{-r(T-t_s)}\mathbf{E}(X|\sigma(\mathcal{R}_{j_s}))$ at $s = 1, \ldots, n$. It should be clear that the hedging strategies $\mathrm{B}\Pi_{\mathcal{R}}(X)$ and $\mathrm{H}\Pi_{\mathcal{R}}(X)$ can be intermixed at different time intervals $[t_{s-1}, t_s)$.

Below we especialize the formalism to simple situations.
Consider the case of hedging an European option in the binomial model. Let us consider the sequence of partitions $\mathcal{R} = \{\mathcal{R}_j\}_{0\le j\le n}$ where $\mathcal{R}_j \equiv \{A_{j,i}\}_{0\le i\le 2^j-1}$ and $A_{j,i}$ as defined in (2.3). It is clear that the sequence of partitions $\mathcal{R} = \{\mathcal{R}_j\}_{0\le j\le n}$ is dyadic and localized at times $t_0, \ldots, t_n$. Therefore, we can hedge any European derivative X, with the strategies $\mathrm{H}\Pi_{\mathcal{R}}(X)$ and $\mathrm{B}\Pi_{\mathcal{R}}(X)$. The Δ-Hedging strategy $\Pi_\Delta(X) = \Delta\ S + B$ is generally used to hedging derivatives [1], its value at $\omega \in A_{j,i}$ is

$$\Pi_\Delta(X)_{t_j}(\omega) = \mathbf{E}(X \mid \sigma(\mathcal{R}_j))(\omega) = e^{-r(T-t_j)}x_j[i] = \Delta_j[i]\ S_j[i] + B_j[i].$$

Computing the amount to hedge on the stock gives

$$\Delta_j[i] = e^{-r(T-t_{j+1})}\frac{x_{j+1}[2i+1] - x_{j+1}[2i]}{S_{j+1}[2i+1] - S_{j+1}[2i]}, \tag{4.13}$$

where the notation $S_j[i]$ refers to the stock values in the binomial model. In order to compare the different hedging strategies, we observe that

$$\Delta_j[i] = \frac{-e^{-r(T-t_{j+1})}}{S_j[i]\ (U-D)\sqrt{p\ q\ P(A_{j,i})}}\ d_j[i].$$

The relation between the portfolio $\mathrm{B}\Pi_{\mathcal{R}}(X)$ and Δ-Hedging is given by equation (4.13) since the coefficients of the portfolio $\mathrm{B}\Pi_{\mathcal{R}}(X)$, corresponding to the binary option $\mathbf{B}_{j,i}$, are equal to $e^{T-t_j}\ x_j[i]$.
Now we consider the sequence of partitions $\mathcal{R} = \{\mathcal{R}_j\}_{0\le j\le n}$ from Example 3 of Section 2 where $\mathcal{R}_j \equiv \{A_{j,i}\}_{i\in I_j}$ and $A_{j,i}$ were defined in (2.5) and (2.6). It is easy to verify that $\mathcal{R}$ is a sequence of weak dyadic partitions localized in $[T_0, T]$. Then the hedging strategies $\mathrm{H}\Pi_{\mathcal{R}}(X)$ and $\mathrm{B}\Pi_{\mathcal{R}}(X)$ are static. We observe that these strategies can only hedge derivatives in $L^2(\sigma(S_T))$.

Consider now hedging an European options in the Black-Sholes model. We have defined, in the Example 4 of section 2, a dyadic sequence of partitions $\{\mathcal{P}_j\}_{j\ge 0}$ with $\mathcal{P}_j = \{A_{j,i}\}, i = 0, \ldots, 2^j - 1$. We observe that this sequence is localized in the times $t_0 = T_0 < \ldots < t_n = T$. Given a derivative $X \in L^2(\sigma(\mathcal{P}_J))$, we have two hedging strategies $\mathrm{H}\Pi_{\mathcal{P}}(X)$ and $\mathrm{B}\Pi_{\mathcal{P}}(X)$. Complete details of special cases of this example have been developed in Section 4.1.

5. Optimizing H-systems Approximations

As outlined in the Introduction, a main goal is to study approximations given by H-systems in order to obtain efficient hedging strategies. Namely, while keeping a small approximation error we seek to minimize transaction costs. In this paper we settle for minimizing the number of transactions, the reason being that this last quantity is more easily described in terms of inner products. See Section 6 for some numerical information and a discussion on transaction costs. Minimizing the number of transactions has an intrinsic interest beyond the issue of transaction costs, namely, it keeps the number of transactions in the Haar hedging portfolio realistically small and, at the same time, isolates the most relevant binary options

needed in the portfolio implementation. The present section makes clear the relevance of the use of Haar functions to obtain hedging strategies with a small number of transactions.
In this section we present two approaches to optimize H-systems, loosely speaking we try to capture as much as possible of the L^2 norm of the option X with as few Haar functions as possible. We notice that the first approach, *compression*, could be combined with the second approach, the *greedy splitting* algorithm.

5.1. **Compressing H-systems.** Our next discussion can be applied to an arbitrary H-system, nevertheless we prefer to phrase the discussion in terms of the Haar systems introduced in Section 2, more specifically Example 5 from that section. The system with $n_T = 1$ is a purely static approximation and requires $2^{j_{n_T}}$ transactions. We will employ some of the notation used in Appendix A. Consider $n_T > 1$ and $J = j_1 = \ldots = j_{n_T}$ (for simplicity), we then have a Haar hedging portfolio which is dynamic and hence allows us to condition on the path as it unfolds. Still, this portfolio requires $n_T 2^j$ transactions (along each path) which, even by taking $j < J$, gives the problem of an unrealistic number of transactions. A solution to this problem is to introduce some kind of optimization, in the present case it will be an aposteriori optimization of our martingale expansion. Consider first the case of $n_T = 1$, we do this optimization by *compressing* our expansion, namely, we sort the inner products $|d_k[i]|, k = 0, \ldots, J-1,\ i = 0, \ldots, 2^k - 1$ and $|x_0[0]|$ in decreasing order and keep a small number of them, say R, to synthesize a new approximation. More precisely, let $n = 2^J$ and X_n denote our martingale approximation, let $u_{k_i}, i = 0, \ldots, n$ be a new indexing for our Haar system $\{u_k\}, k = 0, \ldots, n$, such that $|\langle X, u_{k_{i+1}}\rangle| \geq |\langle X, u_{k_i}\rangle|$. So our compressed approximation, which we will denote by X_n^c, is given by

$$X_n^c = \sum_{i=0}^{R-1} \langle x, u_{k_i}\rangle u_{k_i}, \tag{5.1}$$

X_n^c approximates, in the space L^2, the martingale expansion X_n with the following error

$$||X_n - X_n^c||^2 = \sum_{i=R}^{2^J} |\langle x, u_{k_i}\rangle|^2. \tag{5.2}$$

We have taken R to be fixed because we later intend to compare against a delta hedging approximation. Alternatively, one may decide how many terms R to include by specifying an epsilon level of error via (5.2). All the above computations can be performed with a computational cost of order n. We note that the compression operation can be also performed for Haar systems with parameter $n_T > 1$, to this end, introduce the compression parameters $R_i,\ i = 0, \ldots, n_T - 1$, which will play the role of R in (5.1) (for each of the stages $\{t_i, t_{i+1}\}$). Therefore a compressed approximation for this system will contain $\sum_{i=0}^{n_T-1} R_i$ Haar functions along each path. Therefore, when comparing the performance of this class of models against a delta hedging in which the Black-Scholes portfolio is rebalanced R times along each sampled path, one should need to consider $R = \sum_{i=0}^{n_T-1} R_i$.
In Section 6 we compare the performance of the above approximation against a delta hedging approximation in which the portfolio is rebalanced a fixed number of times along each sampled path.

One can see the effect of compression in Table 1. For example, on the one hand, using $R = 32$ an error of 0.05 can be attained for the parameter $J = 16$ (hence we are choosing the most relevant 32 Haar functions out of a set of 2^{16} Haar functions), on the other hand, using no compression when $J = 8$ (which gives $R = 2^8 = 256$, hence using almost ten more times Haar functions) gives a larger error of 0.07.

5.2. **Greedy Splitting Algorithm.** We briefly describe here an alternative to the above mentioned aposteriori optimization. This algorithm iteratively generates an H-system by sequentially splitting one atom at a time. For a given $X \in L^2$ the idea is to generate an H-system by reducing the L^2 norm of X as much as possible in each iteration of the basic splitting. Due to the general nature of the algorithm, the atoms constructed by the algorithm may not correspond to readily tradable binary options. In particular, several transactions with available binary options may be needed to implement one of the constructed atoms.
To give an intuitive understanding of the algorithm described below define the $0-residue$ by $R^0X = X - \mathbf{E}(X)$, and continue inductively defining the $nth.-residue$ by $R^nX = R^{n-1}X - \langle X, u_n\rangle\ u_n$, notice that $X = \sum_{k=0}^{n} \langle X, u_k\rangle u_k + R^nX$. Due to the Pythagorean relationship $||R^nX||^2 = ||R^{n-1}X||^2 - |\langle X, u_n\rangle|^2$, the greedy-splitting algorithm can then be described by indicating that it maximizes $|\langle X, u_n\rangle|^2$, or equivalently, it minimizes $||R^nX||^2$, under the constrain of constructing an H-system. The maximization is *greedy* because it is only one look ahead, namely it searches for one Haar function at a time. The maximization of $|\langle X, u_n\rangle|^2$ requires a combinatorial number of computations if implemented naively. Under appropriate conditions it can be shown that this combinatorial explosion can be avoided and the construction of $u_1, \ldots, u_n$ requires a number of computations of order $O(n\ \log(n))$ where the constant of proportionality is related to the cost of maximizing a one dimensional continuous function on a close interval.
In the remaining of this section we describe formally the *greedy splitting algorithm*. The algorithm is based on Proposition 3 and constructs an H-system which approximates a given $X \in L^2(\Omega, \mathcal{B}, P)$. First we mention some notation to be used in the remaining of this section, let $\mathcal{B}_A \equiv \{B \cap A : B \in \mathcal{B}\}$; $X_A \equiv X|A$ (the restriction of X to A) and $P_A \equiv \frac{1}{P(A)}P$. It is clear that $X_A \in L^2(A, \mathcal{B}_A, P_A)$ and $F_{X_A}(t) = P_A(X_A \leq t) = \frac{1}{P(A)}P(\{X \leq t\} \cap A)$, where F_X denotes the distribution function of X. F_X^{-1} denotes the right continuous inverse of F_X. The norm $||Y||_A^2 = \langle Y, Y\rangle_A$ denotes the inner product in $L^2(A, \mathcal{B}_A, P_A)$. Expectation on $(A, \mathcal{B}_A, P_A)$ will be denoted with $\mathbf{E}_A$.
We do describe first the main steps, then we indicate how to set up these steps computationally and finally we give precise mathematical statements that justify the computational setup. We will define the H-system implicitly by describing a binary partition $\mathcal{Q} = \{\mathcal{Q}_j\}_{j\geq 0}$. Start by setting $\mathcal{Q}_0 = \{A_{0,0} = \Omega\}$ and assume, inductively, that $\mathcal{Q}_k,\ k \leq j$ has been constructed. We need some intermediate definitions in order to define $\mathcal{Q}_{j+1}$, for a given measurable set A define,

$$\mathcal{C}_A = \{\psi :\ \psi \text{ is a Haar function on } A\}, \tag{5.3}$$

under appropriate conditions, it follows from Proposition 3 that there exists $\psi_A^m = a\ \mathbf{1}_{A_0^m} + b\ \mathbf{1}_{A_1^m}$ in $\mathcal{C}_A$ satisfying

$$|\langle X, \psi_A^m\rangle| = \sup_{\psi\in\mathcal{C}_A} |\langle X - \frac{1}{P(A)}\int_A X,\ \psi\rangle| = \sup_{\psi\in\mathcal{C}_A} |\langle X,\ \psi\rangle|. \tag{5.4}$$

Select now $\hat{A} \in \mathcal{Q}_j$, such that

$$|\langle X, \psi^m_{\hat{A}} \rangle| \geq |\langle X, \psi^m_A \rangle| \tag{5.5}$$

for all $A \in \mathcal{Q}_j$. According to the indexing of binary partitions introduced in Definition 3, $\hat{A} = A_{k,i}$ for some index (k,i), $k \leq j$, now define $A_{k+1,2i} = \hat{A}^m_0$ and $A_{k+1,2i+1} = \hat{A}^m_1$. Finally, set $\mathcal{Q}_{j+1} = \mathcal{Q}_j \backslash \{\hat{A}\} \cup \{A_{k+1,2i}, A_{k+1,2i+1}\}$. It follows from the developments below that we have $|\mathcal{Q}_{j+1}| = |\mathcal{Q}_j| + 1$ (where $|\mathcal{S}|$ denotes cardinality of a set $\mathcal{S}$) unless $\mathbf{E}(X|\mathcal{Q}_j) = X$ in which case the algorithm terminates. We now indicate how to carry out the key computations implicit in the algorithm, the crucial point is that our approach avoids the exponential explosion in the number of computations involved in a direct approach at computing ψ^m_A. Two optimizations are defined (one over subsets of $\mathcal{C}_A$ and another one over $[0,1]$) in such a way that they may be carried out under the sole assumption $X \in L^2$, extra hypothesis are needed for these optimizations to deliver ψ^m_A in (5.4).

Let $A \in \mathcal{Q}_j$, it follows from the above description of the algorithm that the key step is to indicate how to compute ψ^m_A, hence in the discussions below an arbitrary $A \in \mathcal{Q}_j$ will be assumed to be fixed once and for all.

Observe that any $\psi \in \mathcal{C}_A$ can be written, in the form

$$\psi = a\ \mathbf{1}_{A_0} + b\ \mathbf{1}_{A_1} \tag{5.6}$$

for some $u \in (0,1)$, $A_0 \subset A$ with $P_A(A_0) = u$, and $A_1 = A \backslash A_0$. With this notation

$$\langle X, \psi \rangle = (a-b)\langle X,\ \mathbf{1}_{A_0}\rangle + b\langle X,\ \mathbf{1}_A\rangle = b\ P(A) \left(\mathbf{E}_A(X_A) - \frac{1}{u}\langle X_A,\ \mathbf{1}_{A_0}\rangle_A \right). \tag{5.7}$$

Noticing that $b = \pm\sqrt{\frac{u}{P(A)\ (1-u)}}$, in order to calculate the supremum in (5.4) we define

$$\lambda_A(u) \equiv \sup_{\{A_0 : P_A(A_0) = u\}} \sqrt{\frac{P(A)\ u}{(1-u)}} \left(\mathbf{E}_A(X_A) - \frac{1}{u}\langle X_A, \mathbf{1}_{A_0}\rangle_A \right). \tag{5.8}$$

Observe that we have chosen $b > 0$ in (5.7). It is clear that if $\psi \in \mathcal{C}_A$ then $\psi \in \mathcal{C}_{A,u} \equiv \{\psi \in \mathcal{C}_A : P(A_0) = u\}$ for some $u \in (0,1)$. Therefore

$$\sup_{\psi \in \mathcal{C}_A} |\langle X, \psi \rangle| = \sup_{u \in (0,1)} \sup_{\psi \in \mathcal{C}_{A,u}} |\langle X, \psi \rangle| = \sup_{u \in (0,1)} \lambda_A(u). \tag{5.9}$$

Under appropriate conditions we will prove

$$\sup_{\psi \in \mathcal{C}_A} |\langle X, \psi \rangle| = \lambda_A(u^\star) = \langle X, \psi_{u^\star} \rangle,$$

for some $u^\star \in (0,1)$ and $\psi_{u^\star} \in \mathcal{C}_{A,u^\star}$. We will need a series of intermediate results.

Lemma 1. *Assume $X \in L^2(\Omega, \mathcal{B}, P)$ and $A \in \mathcal{B}$. Then the fuction λ_A, in (5.8), is well defined for all $u \in (0,1)$ and in fact*

$$|\lambda_A(u)| \leq \sqrt{P(A)} \|X\|_A. \tag{5.10}$$

Moreover, assuming F_X to be continuous, the Haar function ψ_u defined by

$$\psi_u = -\sqrt{\frac{P(A)\ (1-u)}{u}} \mathbf{1}_{\{X_A \leq F^{-1}_{X_A}(u)\}} + \sqrt{\frac{P(A)\ u}{1-u}} \mathbf{1}_{\{X_A \geq F^{-1}_{X_A}(u)\}} \tag{5.11}$$

satisfies

$$\lambda_A(u) = \langle X, \psi_u \rangle.$$

Proof.

$$\begin{aligned}\left|\mathbf{E}_A(X_A) - \tfrac{1}{u}\langle X_A, \mathbf{1}_{A_0}\rangle_A\right| &= \left|\int_A X(\mathbf{1}_A - \tfrac{1}{u}\mathbf{1}_{A_0})dP_A\right| \leq \|X_A\|_A \|\mathbf{1}_A - \tfrac{1}{u}\mathbf{1}_{A_0}\|_A \\ &= \|X_A\|_A(\|\mathbf{1}_{A_1}\|_A^2 + \|(1-\tfrac{1}{u})\mathbf{1}_{A_0}\|_A^2)^{1/2} \\ &= \|X_A\|_A\sqrt{\tfrac{(1-u)}{u}}.\end{aligned}$$

Consequently (5.10) holds. For the last assertion consider

$$\begin{aligned}\lambda_A(u) &= \sup_{\{A_0:P(A_0)=u\}} \sqrt{\tfrac{P(A)\ u}{(1-u)}} \left(\mathbf{E}_A(X_A) - \tfrac{1}{u}\langle X_A, \mathbf{1}_{A_0}\rangle_A\right) = \\ &= \sqrt{\tfrac{P(A)\ u}{(1-u)}} \left(\mathbf{E}_A(X_A) - \tfrac{1}{u}\ \inf_{\{A_0:P(A_0)=u\}} \langle X_A, \mathbf{1}_{A_0}\rangle_A\right).\end{aligned}$$

In order to find the minimizer of $\langle X_A, \mathbf{1}_{A_0}\rangle_A$, for fixed u, we apply the bathtub principle from [13] to the probability space $(A, \mathcal{B}_A, P_A)$ and remark that continuity of F_X implies continuity of F_{X_A}. The application of this principle and the continuity of F_{X_A} immediately give $A_0 = \{X_A \leq F_{X_A}^{-1}(u)\}$ and $P_A(A_0) = u$. This gives (5.11) and concludes the proof. □

The following result gives sufficient conditions under which $\lambda_A(u)$ is continuous on $[0,1]$ and also for its supremum to be realized for some $u^\star \in (0,1)$.

Lemma 2. *Assume F_X to be continuous then $\lambda_A(u)$ is continuous on $(0,1)$. Furthermore, if $X \in L^\infty$ then $\lambda_A(u)$ is continuous on $[0,1]$ and*

$$\lim_{u\to 0^+} \lambda_A(u) = \lim_{u\to 1^-} \lambda_A(u) = 0. \tag{5.12}$$

Proof. According to Lemma 1, we have

$$\lambda_A(u) = \sqrt{\frac{P(A)\ u}{(1-u)}} \left(\mathbf{E}_A(X_A) - \frac{1}{u}\langle X_A, \mathbf{1}_{A_0}\rangle_A\right),$$

where $A_0 = \{X_A \leq F_{X_A}^{-1}(u)\}$. Notice that $\int_{A_0} X_A(\omega)dP_A(\omega) = \int_{\{F_{X_A}(X_A)\leq u\}} X_A(\omega)dP_A(\omega)$ is continuous for $u \in (0,1)$, this proves continuity of $\lambda_A(u)$ on $(0,1)$. Notice that $P_A(A_0) = u$, hence

$$-||X_A||_\infty \sqrt{u} \leq \frac{\sqrt{u}}{u}\int_{A_0} X(\omega)dP_A(\omega) \leq \sqrt{u}F_{X_A}^{-1}(u), \tag{5.13}$$

both left and right sides of the above equation converge to 0 because X_A is a bounded function, therefore, $\lambda_A(0^+) = 0$.
In order to evaluate $\lambda_A(1^-)$ define $A_1 = A\backslash A_0$ and write

$$\lambda_A(u) = \sqrt{\frac{(1-u)}{P(A)u}}\left(\frac{1}{(1-u)}\int_{A_1} X(\omega)dP(\omega) - \int_A (X(\omega)dP(\omega)\right). \tag{5.14}$$

Observe that $P_A(A_1) = 1-u$, hence

$$\sqrt{1-u}\ F_{X_A}^{-1}(u) \leq \frac{\sqrt{1-u}}{1-u}\int_{A_1} X(\omega)dP_A(\omega) \leq ||X_A||_\infty\sqrt{1-u}, \tag{5.15}$$

both left and right sides of the above equation converge to 0 because X_A is a bounded function. It follows from (5.14) that we have stablished $\lambda_A(1^-) = 0$. □

Proposition 3. *Assume $X \in L^2(\Omega, \mathcal{B}, P)$ and that the hypothesis in Lemma 2 are satisfied. Then there exist $u^* \in (0,1)$ such that $\psi_A^m \equiv \psi_{u^*} \in \mathcal{C}_{A,u^*}$, where ψ_{u^*} is given by (5.11), verifies*

$$\langle X, \psi_A^m \rangle = \sup_{\psi \in \mathcal{C}_A} |\langle X, \psi \rangle|. \tag{5.16}$$

Proof. Since λ_A is continuous, by Lemma 2, let u^* be its maximizer and $\psi_A^m \equiv \psi_{u^*}$. Consider now $\psi \in \mathcal{C}_A$, we know that

$$\psi = b\,\sqrt{P(A)}\,\left(-\frac{(1-u)}{u}\mathbf{1}_{A_0} + \mathbf{1}_{A_1}\right)$$

with $b \equiv \pm\sqrt{\frac{u}{(1-u)}}$. If $b < 0$

$$\begin{aligned}\psi &= \sqrt{\tfrac{P(A)\,(1-u)}{u}}\,\mathbf{1}_{A_0} - \sqrt{\tfrac{P(A)\,u}{(1-u)}}\,\mathbf{1}_{A_1} \\ &= \sqrt{\tfrac{P(A)\,u'}{(1-u')}}\,\left(-\tfrac{(1-u')}{u'}\mathbf{1}_{A_0'} + \mathbf{1}_{A_1'}\right),\end{aligned}$$

where $u' = 1-u$, $A_0' = A_1$ and $A_1' = A_0$, thus $\psi \in \mathcal{C}_{A,u'}$ with $b' = \sqrt{\frac{u'}{(1-u')}} > 0$. Thus, ψ belongs to some $\mathcal{C}_{A,u}$ with $b > 0$. Consequently

$$-\langle X, \psi_A^m \rangle \leq \langle X, \psi \rangle \leq \langle X, \psi_A^m \rangle.$$

□

Essentially, whenever F_X is continuous and X bounded, we are able to prove that the greedy splitting algorithm satisfies some optimality properties. Moreover, the computations can be reduced to maximizing a continuous functional on $[0,1]$ where several powerful algorithms are available. Analogous results to Propositions 3 and 2 are also possible for the case when X is a purely discrete random variable. In this last case we need to do a search over the set of $u_n \in (0,1)$ such that there exists s_n with $F_{X_A}(s_n) = u_n$. It may not be possible to avoid an exhaustive search for this optimization. The case of general X is also solved by the bathtub principle mentioned above but it requires the use of "Haar" functions outside of the scope of our paper. If X is a path independent European option that is an increasing function of S_T, the atoms constructed by the greedy splitting algorithm will correspond to binary options associated to a single interval $S_T \in [a,b]$, for other type of options this is not necessarily the case.

The algorithm converges pointwise and in L^p by the simple fact that it constructs a martingale. Under appropriate conditions, it is possible to prove that it actually converges to X. We plan to study questions related to this algorithm, as well as software realizations, in a future publication.

6. Numerical Examples

In this section we present output from a computer implementation based on Example 5 from Section 2. More specifically, we concentrate on the case where we have a Haar dyadic system, Definition 6, whose sequence of partitions $\mathcal{P}_j = \{A_{j,i}\}$ are constructed via the increments of the Brownian motion and are characterized through the parameters n_T and $j_1, \ldots, j_{n_T}$. We will also use compression as described in Section 5.1 and some of the definitions and notions introduced in Section 4.

To indicate the potential improvements that can be expected for this example it is enough to consider the case of $n_T = 1$, therefore, all the atomic sigma algebras

$\mathcal{B}_n$ are included in $\sigma(W_T)$ and $\mathcal{B}_\infty = \sigma(W_T)$. The case $n_T > 1$ is essentially a concatenation of several steps where each step is algorithmically equivalent to the case $n_T = 1$. Moreover, the errors along these steps accumulate as is the case with delta hedging.

We compare the errors in the approximations as well as the volume of transactions as a function of the number of transactions. We find generic cases where Haar systems outperform delta hedging, moreover, in these examples, the improvements have a simple intuitive financial meaning. Our numerical output uses the parameter R, introduced in (5.1), R is (essentially) equal to the number of Haar hedging transactions plus one. This is just a peculiarity of our software and it can be understood by noticing that the bank account u_0 may or may not be chosen during the compression step (in practice it is one of the largest contributing inner products). In short, the parameter R is equal to the number of times the Black-Scholes portfolio is rebalanced when performing delta hedging and equals the number of Haar functions used in the final approximation when performing Haar hedging. We rebalance the Black-Scholes portfolio at uniformly spaced time intervals.

Here we will give the initial data for the MRA (this algorithm is described in Appendix A) for the H-system $\{u_{2^j+i}\}$ in Example 5 and X an European option. As previously remarked, computations can be carried out by specifying the finest scale J. We will then perform compression by only keeping the R Haar functions, including also u_0, with the largest inner products. Fixed an acceptable error $\epsilon > 0$, we approximate X specifying the finest scale J, in such way that the conditional expectation satisfies

$$\sup |X(\omega) - \mathbf{E}(X|\sigma(\{A_{J,i} : 0 \leq i \leq 2^J - 1))(\omega)| < \epsilon,$$

this is possible because every bounded random variable can be approximated by simple functions supported on atoms of probability $\frac{1}{2^J}$. As a matter of convenience, according to computational costs, we have used $J = 14$ or $J = 16$. The input to the MRA is obtained by computing

$$x_J[i] = 2^J \int_{A_{J,i}} X(\omega) dP(\omega), \tag{6.1}$$

or, more conveniently, for the case of continuous $X(\omega) = X(S_T(\omega))$, by first computing

$$\begin{aligned} s_J[i] &= 2^J \textstyle\int_{A_{J,i}} S_T(\omega) dP(\omega) = \\ &= \frac{2^J}{\sqrt{2\pi}} \int_{c_i^J}^{c_{i+1}^J} S_{T_0} e^{(\nu(T-T_0)+\sigma\sqrt{(T-T_0)}\; y)}\; e^{-\frac{y^2}{2}}\; dy = \\ &= S_{T_0} e^{(\nu(T-T_0))}\; e^{\frac{b^2}{2}}\; 2^J \left(\Phi(c_{i+1}^J - b) - \Phi(c_i^J - b)\right), \end{aligned} \tag{6.2}$$

where $b = \sigma\sqrt{(T-T_0)}$ and $\nu = (r - \frac{\sigma^2}{2})$. Therefore, by taking J sufficiently large, we can use the approximation $x_J[i] \approx X(s_J[i])$. We recall that $p_J[i] = P(A_{J,i}) = \frac{1}{2^J}$.

For the sake of clarification, consider the European call $X(\omega) = (S_T(\omega) - K)_+$ where $T \equiv t_n$ is the time of exercise and K is the strike price. Clearly X is unbounded, but $\lim_{c\to\infty} X\mathbf{1}_{\{X\leq c\}} = X$ a.e., hence one can always consider an approximation of a desired quality.

Next we comment on the output displays; numerical values were obtained by sampling $S_T(\omega)$, the limited range in these values (x-axis on most displays) correspond to these sampled values (after sorting). Consider first a single European call $X(\omega) = (S_T(\omega) - K)_+$ as above, values of parameters are indicated in the text surrounding the figures. In Figures 1, 2 and 3 we present the Black-Scholes and Haar approximations with $R = 1,\ 2,\ 20$ respectively. Notice how Figure 1 shows the Haar approximation with $u_1 = 1/2\ (1_{A_{1,0}} - 1_{A_{1,1}})$ which happens to give the largest inner product. Figure 2 shows the Haar approximation when u_0 is added, giving the second largest inner product in this example. Figure 4 shows the estimation of the L^2 norm of the errors as a function of R.
As a second example we consider a portfolio built as a linear combination of European calls and puts as follows, $X = (S_T - K_1) + (S_T - K_2) - (S_T - K_3)$. values of parameters are indicated in the text surrounding Figure 5. Finally, Figure 6 shows the estimation of the L^2 norm of the errors as a function of R.
Tables 2 and 3 show the volume of transactions for the Haar hedging portfolio, and for the binary hedging portfolio (see Section 4), which for the case $n_T = 1$ are both constant quantities, and the volume of transactions for the Black-Scholes portfolio. Using the notation X_n^c from (5.1), the volume of transactions for the Haar hedging portfolio, is equal to

$$(6.3)\quad VT(\text{H}\Pi) = e^{-r(T-T_0)}||X_n^c - \mathbf{E}(X)||_{L^1} = e^{-r(T-T_0)} \int_\Omega |X_n^c(\omega) - \mathbf{E}(X)|dP(\omega).$$

The volume of transactions for the portfolio of binary options is

$$(6.4)\qquad VT(\text{B}\Pi) = e^{-r(T-T_0)}||X_n^c||_{L^1} = e^{-r(T-T_0)} \int_\Omega |X_n^c(\omega)|dP(\omega).$$

Analogous expressions for (6.3) and (6.4) are also possible for the case $n_T > 1$. On the other hand, letting

$$\varphi_{t_i} = \frac{\partial V_{t_i}(X)}{\partial S_{t_i}},$$

the volume of transactions for a Black-Scholes portfolio with rebalancing dates $\{t_i\}, i = 0, \ldots, R-1$ is

$$(6.5)\qquad \sum_{i=0}^{R-1} [\ |\varphi_{t_i} - \varphi_{t_{i-1}}|\ S_{t_i} + (B_{t_i} - B_{t_{i-1}} e^{r(t_i - t_{i-1})})_+\ ],$$

with $\varphi_{t_{-1}} = B_{t_{-1}} \equiv 0$. We have used equally spaced rebalancing dates starting at $t_0 = T_0$. Given that (6.5) is a random quantity we will report the average (`AverageVolTrBS`) over many samples.
The smaller the oscillations of X around $\mathbf{E}(X)$ the smaller $VT(\text{H}\Pi)$ will be compared to $VT(\text{B}\Pi)$. Notice the difference in magnitudes with `AverageVolTrBS`. The volume of transactions offer a clear numerical evidence of the different nature between Haar hedging and delta heding. In both cases, Haar hedging and binary hedging, the explicit use of space discretization implies that the volume of transactions is essentially the same when the number of transactions increases. For delta hedging, its relience on time discretization implies much larger volumes of transactions when the rebalancing frequency is increased to reduce the hedging error.
We now comment on our choice of examples. It is expected, and it is confirmed by our experience with numerical examples, that the Haar approximation outperforms (in the sense of smaller error for equal value of R) the Black-Scholes approximation

whenever the payoff, or its derivative, contains discontinuities. Moreover it is important that the Haar functions are adapted to these discontinuities, for example, we can choose u_1 such that it is supported in the union of $A_{1,0} = \{S_T < K\}$ and $A_{1,0} = \{S_T \geq K\}$ for the case of the European call. Our examples reflect these choices, for example S_{T_0} was taken close to K so as the discontinuity in the first derivative of the European call becomes problematic for Black-Scholes approximation and can be reproduced efficiently by the Haar expansion. An extreme example of this kind will be the case of a digital option where, of course, the Haar expansions have no bearing as a hedging tool.
One can, of course, easily find situations where delta hedging ouperforms Haar hedging as, for example, a position in a European call which is well in or out of the money. This is a situation where the linear approximation in delta hedging becames very efficient. It may be interesting to see under what conditions delta hedging and Haar hedging are complementary and to investigate how to combine both techniques.

7. Conclusions and Extensions

We have introduced a basic and general new framework to represent European options. Key ingredients are the flexibility given by the possible space and time discretizations which can be adapted to the given option and the potential for financial realization of these discretizations. From a theoretical point of view, the approach is as fundamental as delta hedging and it is reasonable to think that can be extended to settings where this last technique is available. Some of the computational tools introduced can also be used even when an actual financial realization is not available, pricing computations is an example,. We have emphasized the issue of *efficient* representations of a given option X, this notion isolates a few binary options with small approximating error. The representation in terms of Haar functions was created with this goal in mind, Section 5 provides examples of how this tool could be deployed.
Further empirical and theoretical work is needed to assess the realm of applications where the new constructions offer a financial advantage. The techniques could likely be extended to the setting of American options and higher dimensional models.

Appendix A. Multiresolution Analysis for H-systems and Applications

Here we complete the computational material and describe the Multiresolution Analysis Algorithm (MRA) for a general H-system. The intended use of this algorithm is for pricing and hedging derivatives. It is known that the price of one european option X at time T in a market without arbitrage opportunities is given by $e^{-rT}\mathbf{E}(X)$ where r is the interest rate of a risk free bond and the expectation is taken with respect to the risk neutral measure. The MRA for an H-system $\{u_n\}$ is a useful tool for calculate this expectation, furthermore it is a fast way to obtain the expansions of X at all levels of approximation in the associated H-system. Fix $X \in L^2(\Omega, \mathcal{B}, P)$ and assume a sequence $\mathcal{R} = \{\mathcal{R}_j\}_{j=0}^{J}$ of weak dyadic partitions of Ω is given.
To facilitate the definition of the MRA algorithm we introduce a signal vector with the following notation

$$\mathbf{x}_j = (x_k[i])_{(k,i)\in K_j}, \text{ where,}$$

(A.1)
$$x_k[i] = \frac{1}{P(A_{k,i})} \int_{A_{k,i}} X(\omega)\ dP(\omega) = X_j(\omega) \text{ for any } w \in A_{k,i}, \text{ with } (k,i) \in K_j.$$

For purposes of the algorithm specification we will set a finest scale J and initialize our input signal vector with $\mathbf{x}_J$.

Proposition 4. *The following equations hold for* $1 \le j \le J$ *and* $i \in I_j$.

$$x_{j-1}[i] = \frac{1}{p_{j-1}[i]} \left(p_j[2i]\ x_j[2i] + p_j[2i+1]\ x_j[2i+1]\right), \tag{A.2}$$

$$d_{j-1}[i] = \sqrt{\frac{p_j[2i]\ p_j[2i+1]}{p_{j-1}[i]}} \left(x_j[2i] - x_j[2i+1]\right). \tag{A.3}$$

Also we have the reconstruction equations, starting with $x_0[0]$ *and* $d_j[i]$, $1 \le j \le J-1, i \in I_j$.

$$x_j[2i] = x_{j-1}[i] + \sqrt{\frac{p_j[2i+1]}{p_{j-1}[i]\ p_j[2i]}}\ d_{j-1}[i], \tag{A.4}$$

$$x_j[2i+1] = x_j[2i] - \sqrt{\frac{p_{j-1}[i]}{p_j[2i]\ p_j[2i+1]}}\ d_{j-1}[i]. \tag{A.5}$$

Proof. From (3.8) and (3.5) we have that

$$\begin{aligned} c_{j-1}[i] &= \langle X, \phi_{j-1,i}\rangle = \langle X, \sqrt{\tfrac{p_j[2i]}{p_{j-1}[i]}}\phi_{j,2i} + \sqrt{\tfrac{p_j[2i+1]}{p_{j-1}[i]}}\phi_{j,2i+1}\rangle = \\ &= \tfrac{1}{\sqrt{p_{j-1}[i]}}(\sqrt{p_j[2i]}c_j[2i] + \sqrt{p_j[2i+1]}c_j[2i+1]) \end{aligned}$$

and, for $\omega \in A_{j-1,i}$

$$c_{j-1}[i] = \frac{1}{\sqrt{p_{j-1}[i]}} \int_{A_{j-1,i}} X\ dP = \sqrt{p_{j-1}[i]}X_j(\omega) = \sqrt{p_{j-1}[i]}x_j[i],$$

from where (A.2) follows. Similarly, combining (3.8) and (3.6), (A.3) can be obtained. Finally (A.4) and (A.5) follows adding and subtracting (A.2) and (A.3). □

The above propositions contain all the essential information to set up the algorithm, here are the details.

MRA Description.
We will need to keep, besides the information on inner products and sample values, the probabilities at different scales. To avoid misunderstandings lets clarify the notation. At scales $j < J$ we will compute the vector

$$(\mathbf{x}_j, \mathbf{d}_j, \mathbf{d}_{j+1}, \ldots, \mathbf{d}_{J-1}), \tag{A.6}$$

where each vector $\mathbf{d}_j$ has length $|I_j|$. Similarly, starting with the probability input vector $\mathbf{p}_J = (p_k[i])_{(k,i)\in K_J}$, at scale $j < J$ we will compute the vector

$$(\mathbf{p}_j, \mathbf{dp}_j, \mathbf{dp}_{j+1}, \ldots, \mathbf{dp}_{J-1}), \tag{A.7}$$

where each vector $\mathbf{dp}_j$ has length $|I_j|$.
We describe how the different data structures are updated when going from step j to $j-1$ $(1 \le j \le J)$ this is called the *analysis* part of the algorithm, and how

to go from step $j-1$ to j in the *synthesis* part of the algorithm. For the analysis algorithm we proceed as follows. For $1 \leq j \leq J$ and $i \in I_{j-1}$

$$p_{j-1}[i] = p_j[2i] + p_j[2i+1] \text{ and } dp_{j-1}[i] = p_j[2i], \tag{A.8}$$

$x_{j-1}[i]$ and $d_{j-1}[i]$ are given respectively by (A.2) and (A.3).
Analogously, for the synthesis algorithm,

$$p_j[2i] = dp_{j-1}[i] \text{ and } p_j[2i+1] = p_{j-1}[i] - dp_{j-1}[i], \tag{A.9}$$

$x_j[2i]$ and $x_j[2i+1]$ are given respectively by (A.4) and (A.5)

Applications.

In that follows we consider the H-systems developed in the examples of Section 2, and show some the interest in the MRA for financial computations.

1. Path dependent options in the binomial model.

We now illustrate the MRA for the Haar system $\{u_n\}$ in Example 2, for the case of a path dependent option. We consider the arithmetic Asian call $X(\omega) = (A_T(\omega) - K)_+$ where $T \equiv t_n$ is the time of exercise, K is the strike price and $A_T(\omega)$ is the average of the values of the stock. In this case the signal vector for initializing the MRA is

$$\mathbf{x}_J = (x_J[i])_{\{0 \leq i \leq 2^J - 1\}}.$$

because $\{u_{j,i}\}$ is actually a Haar system. $x_J[i]$ can be calculated observing that the path of stock prices is stored in the index. In fact, writing i in its binary representation with $J+1$ figures, $i = \sum_{l=0}^{J} i_l 2^l$ we have that

$$x_J[i] = \left(\frac{S_0}{J+1} \sum_{l=0}^{J} U^{i_0} \cdots U^{i_l} D^{1-i_0} \cdots D^{1-i_l} - K \right)_+$$

Recall that the probability vector $\mathbf{p}_J = (p_J[i])_{\{0 \leq i \leq 2^J - 1\}}$ is given by $p_J[i] = p^{i_0} \cdots p^{i_J} q^{1-i_0} \cdots q^{1-i_J}$. The analysis algorithm have the equations:

$$x_{j-1}[i] = q \; x_j[2i] + p \; x_j[2i+1],$$

$$d_{j-1}[i] = \sqrt{p_{j-1}[i] \; p \; q} \; \left(x_j[2i] - x_j[2i+1] \right),$$

where $p_{j-1}[i] = p_j[2i] + p_j[2i+1]$. Clearly, $e^{-rT} x_0[0]$ is the price of X in this model. Finally, we observe that the synthesis algorithm have the equations:

$$x_j[2i] = x_{j-1}[i] + \sqrt{\frac{p}{p_{j-1}[i] \; q}} \; d_{j-1}[i],$$

$$x_j[2i+1] = x_j[2i] - \sqrt{\frac{1}{p_{j-1}[i] pq}} \; d_{j-1}[i].$$

2. European options in the binomial model.

We now indicate the initial data of the MRA for the H-system $\{v_{j,i}\}$ in Example 3 for a, path independent, European option. Consider the European call $X(\omega) = (S_T(\omega) - K)_+$ where $T \equiv t_n$ is the time of exercise and K is the strike price. In this case the signal vector to initialize the MRA is

$$\mathbf{x}_J = ((S_0 U^{\alpha(i,k)} D^{n-\alpha(i,k)} - K)_+)_{(k,i) \in K_J}$$

where $\alpha(i,k)$ is the integer part of $\frac{n(i+1)}{2^k}$. In this case, the probability vector is given by $p_k[i] = \binom{n}{\alpha(i,k)} p^{\alpha(i,k)} q^{n-\alpha(i,k)}$ for $(k,i) \in K_J$. It is important to observe that the MRA is most effective to calculate the value $V_{T_0}(X) \equiv e^{-rT}x_0[0]$, of the derivative X, than the usual pricing algorithm in this model.

Appendix B. Martingale Inequalities and Hedging Errors

In this section we review some well known results, in particular we state the fundamental martingale convergence theorem. We also review a few of the basic martingale inequalities that are available for assesing the quality of the approximations while performing Haar hedging.

Here is a martingale convergence theorem relevant to our setting. Assume a sequence of increasing sigma algebras (not necessarily atomic ones) $\mathcal{B}_n$, $n = 0, 1, \ldots$ is given relative to a probability space $(\Omega, \mathcal{B}, P)$. We also use the notation $\mathcal{B}_\infty = \sigma(\cup_{n\geq 0}\mathcal{B}_n)$.

Theorem 3. *Let $p \in [1, \infty)$ be a given real number. For every $X \in L^p$, the sequence $X_n = \mathbf{E}(Z|\mathcal{B}_n)$ is a martingale which converges a.s. and in L^p to $X_\infty = \mathbf{E}(X|\mathcal{B}_\infty)$.*

When applying the above result to a given option X one needs to argue separately for $\sigma(X) \subseteq \mathcal{B}_\infty$, this inclussion should be clear in our simple examples but in general it requires to prove that the atomic approximation satisfies some kind of Vitali covering property, [18] chapter 7. Usually, convergence results on martingale theory are described for an infinite sequence, this is the case for the above theorem. It is trivial to make most statements also applicable to a martingale indexed by a finite index set of integers but this new version of the result is not informative. Nevertheless, some results can also be made informative for the finite index case. For example, in some of our applications, we may start with a finite martingale converging to $\mathbf{E}(X|\sigma(\mathcal{R}_j))$ where $\mathcal{R}_j = \{A_{k,i}, (k,i) \in K_j\}$ is a partition of Ω satisfying

$$|X(\omega) - 1/P(A_{k,i}) \int_{A_{k,i}} X(\omega)dP(\omega)| < \epsilon_{(k,i)} \text{ for all } w \in A_{k,i} \tag{B.1}$$

and for all $(k,i) \in K_j$. Hence, convergence statements and supporting inequalities will give information on convergence to $\mathbf{E}(X|\sigma(\mathcal{R}_j))$ which is useful if the $\epsilon_{(k,i)}$ are taken sufficiently small.

We state the following results using the setting and notation introduced in Section 2 even though the results may hold under more general conditions. Recall the notation $X_n = \mathbf{E}(X|\mathcal{B}_n)$ and set $\xi_k = \langle X, u_k\rangle u_k$.

To simplify the statements we assume $X \in L^2$ in the remaining of this section. The simplest inequality valid for any orthogonal system is Chebychev's inequality which gives information on weak convergence.

Proposition 5. *For any $\lambda > 0$,*

$$P(|X - X_n| \geq \lambda) \leq \frac{\sum_{k=n+1}^{\infty} |\langle X, u_k\rangle|^2}{\lambda^2}. \tag{B.2}$$

The next inequalities are directly related to pointwise convergence and do depend crucially on martingale properties.

Theorem 4. *Fix $n < m$ and $\epsilon > 0$ then,*

$$P(|\sum_{k=n+1}^{j} \xi_k| \leq \epsilon;\ \textit{for all } j,\ n+1 \leq j \leq m) \geq 1 - \frac{1}{\epsilon^2} \sum_{k=n+1}^{m} |\langle X, u_k\rangle|^2 \geq$$

$$1 - \frac{1}{\epsilon^2}(||X||^2 - \sum_{k=0}^{n} |\langle X, u_k \rangle|^2). \tag{B.3}$$

Moreover

$$P(\max_{n+1 \leq j \leq m} | \sum_{k=n+1}^{j} \xi_k | > \epsilon) \leq \frac{K}{\epsilon^2} \sum_{k=n+1}^{m} |\langle X, u_k \rangle|^2 \leq \frac{K}{\epsilon^2}(||X||^2 - \sum_{k=0}^{n} |\langle X, u_k \rangle|^2). \tag{B.4}$$

Clearly the above inequalities give information for the error of an H-system approximation, namely $e_n(X) = X - X_n = \sum_{k=n+1}^{\infty} \xi_k$.
The inequalities in the previous equations uniformly bound the error on a large portion of the space. The next inequalities, whenever applicable, are more detailed, [10], [11] pp.125. We first introduce some notation, set $W_n^2 = \sum_{j=n}^{\infty} \xi_j^2$, $T_n = \sum_{j=n}^{\infty} \xi_j$ and $t_n^2 = \mathbf{E}(T_n^2)$, notice that $t_n \to 0$ as $n \to \infty$.

Theorem 5. *Assume the following condition where the limit is a.s. (also the inequality is a.s)*

$$t_n^{-2} W_n^2 \to \eta^2 > 0. \tag{B.5}$$

Moreover if the following two conditions are satisfied ($I(\ )$ is the indicator function)

$$\sum_{j=1}^{\infty} t_j^{-1} \mathbf{E}(X_j I(|X_j| > \epsilon\ t_j)) < \infty \text{ for all } \epsilon > 0, \tag{B.6}$$

and

$$\sum_{j=1}^{\infty} t_j^{-4} \mathbf{E}(X_j^4 I(|X_j| \leq \delta\ t_j)) < \infty \text{ for some } \delta > 0, \tag{B.7}$$

Then a.s.,

$$-1 = \liminf_{n \to \infty} [\phi(W_n^2)]^{-1} T_n = \limsup_{n \to \infty} [\phi(W_n^2)]^{-1} T_n = 1, \tag{B.8}$$

where $\phi(t) = (2t \log|\log t|)^{1/2}$

Let $Z_n^2 = \sum_{j=n}^{\infty} \mathbf{E}(\xi_j^2 | \mathcal{B}_{j-1})$, then, if

$$\sum_{j=1}^{\infty} t_j^{-2} (\xi_j^2 - \mathbf{E}(\xi_j^2 | \mathcal{B}_{j-1})) \text{ converges a.e.} \tag{B.9}$$

we can replace W_n^2 by Z_n^2 in the above results.

Appendix C. Figures and Tables

TABLE 1. L^2 norm for errors, between X and X_J^c, in terms of number of transactions and scales. Single European Call. Values of parameters as in Figure 1.

No. of Transactions	J=6	J=8	J=10	J=12	J=14	J=16
$R = 8$	0.22	0.22	0.22	0.22	0.22	0.22
$R = 16$	0.15	0.10	0.10	0.10	0.10	0.10
$R = 32$	0.14	0.08	0.06	0.05	0.05	0.05
$R = 64$	0.14	0.08	0.05	0.03	0.02	0.02
$R = 128$	x	0.07	0.05	0.03	0.02	0.01
$R = 256$	x	0.07	0.05	0.03	0.01	0.00

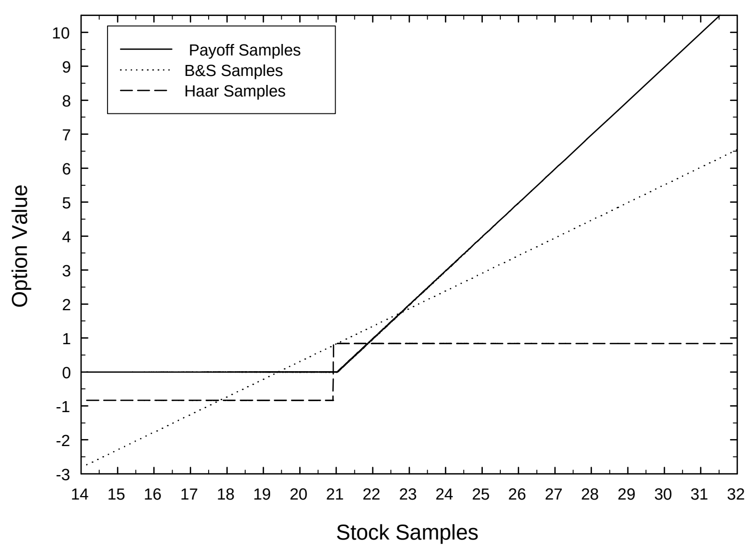


FIGURE 1. Approximations to single European Call using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $R = 1$, $S_{T_0} = 20$, $r = 0.05$, $\sigma = 0.1$, $T - T_0 = 1$, $K = 21$.

REFERENCES

[1] N. H. Bingham and R. Kiesel (1998), *Risk-Neutral Valuation. Pricing and Hedging of Financial Derivatives.* Springer-Verlag, London.

[2] J. Cai and S. E. Ferrando (March 2004), C++ Implementation of Haar Systems for Black-Scholes Model. *Technical Report, Ryerson University.*

[3] A. Cohen, W. Dahmen, I. Daubechies and R. DeVore (2001), Tree approximation and optimal encoding. *Appl. Comput. Harmonic Anal.*, **11**, no. 2., pp. 192-226.

[4] Emanuel Derman, Deniz Ergener and Iraj Kani (Summer 1995), Static options replication. *The Journal of Derivatives*, 78-95.

[5] D.L. Donoho (1993), Unconditional bases are optimal for data compression and for statistical estimation. *Appl. Comput. Harmonic Anal.*, **1**, no. 1., pp. 100-105.

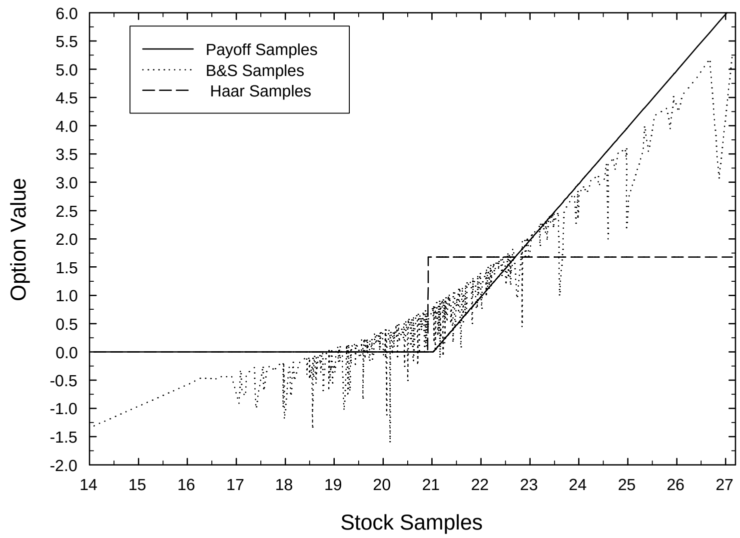


FIGURE 2. Same as in Figure 1 except $R = 2$.

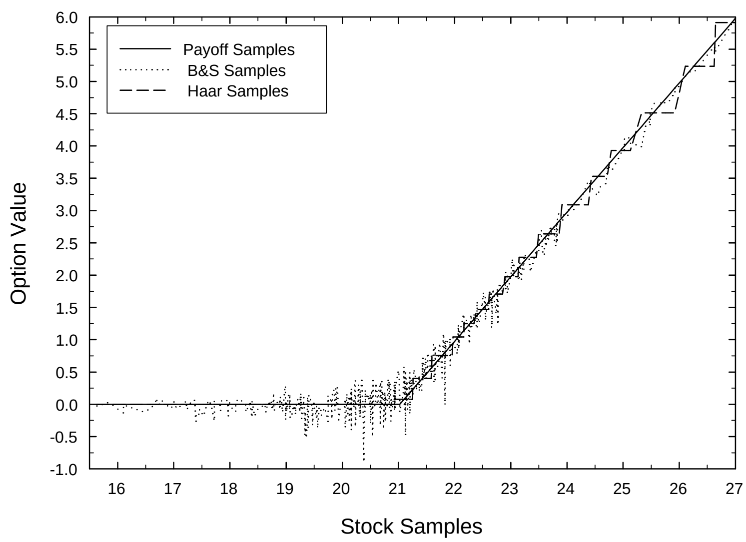


FIGURE 3. Same as in Figure 1 except $R = 20$.

[6] D.L. Donoho (1996), Unconditional bases and bit-level compression. *Appl. Comput. Harmonic Anal.*, **3**, no. 4., pp. 388-392.

[7] D.L. Donoho, M. Vetterli, R.A. DeVore and I Daubechies (1998), Data compression and harmonic analysis. *IEEE Trans. Inf. Theory*,**44** (6), pp. 2435-2476.

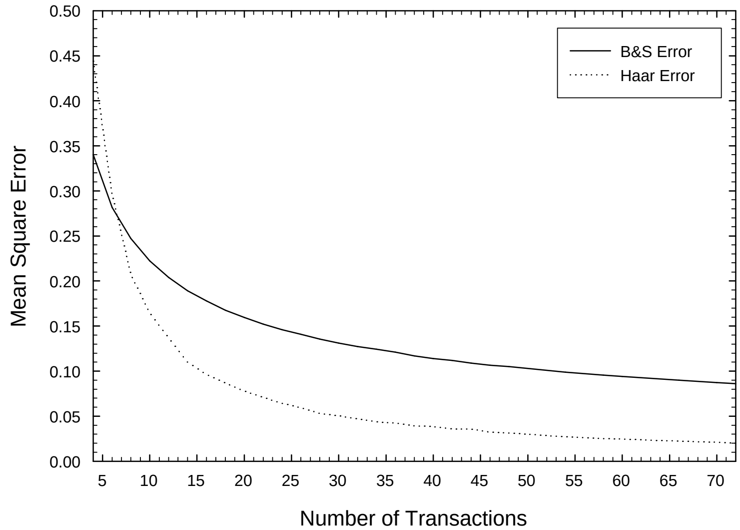


FIGURE 4. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter R. Values of the parameters used: as in Figure 1.

TABLE 2. Volume of Transactions for single European Call. Values of parameters as in Figures 1-4. $V_{T_0}(X) = 0.797$

No. of Transactions (R)	VT(BΠ)	VT(HΠ)	AverageVolTrBS
$R = 5$	0.79	0.92	29.58
$R = 10$	0.8	0.94	40.3
$R = 15$	0.76	0.92	48.1
$R = 20$	0.78	0.93	55.31
$R = 25$	0.79	0.93	60.52
$R = 30$	0.79	0.94	65.22

[8] Richard F. Gundy (1966), Martingale theory and pointwise convergence of certain orthogonal systems. *Trans. Amer. Math. Soc.*, **124**, 228-248.
[9] E. Hernandez and G. Weiss (1996):*A First Course on Wavelets.* CRC Press, Boca Raton Fl.
[10] C. C. Heyde (1977), On central limit and iterated logarithm supplements to the martingale convergence theorem. *J. Appl. Prob.* **14**, 758-775.
[11] P. Hall and C. C. Heyde (1980): *Martingale Limit Theory and Application.* Academic Press, New York.
[12] A. Jensen and A. la Cour-Harbo (2001): *Ripples in Mathematics. The Discrete Wavelet Transform.* Springer.
[13] E. Lieb and M. Loss (1997): *Analysis.* Graduate Studies in Mathematics 14 American Mathematical Society.
[14] D. B. Madan and F. Milne (July 1994), Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance*, **4**, No. 3, 223-245.

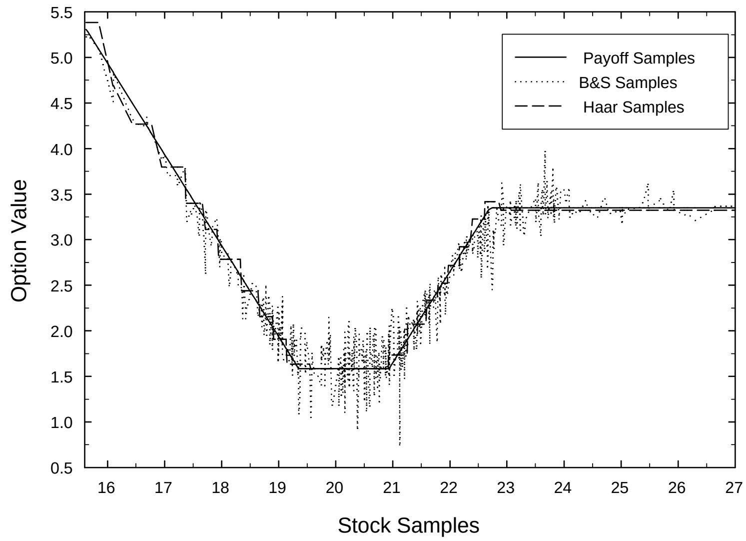


FIGURE 5. Approximations to portfolio X, constructed from two calls and one put, using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $R = 20$, $S_{T_0} = 20$, $r = 0.05$, $\sigma = 0.1$, $T - T_0 = 1$, $K_1 = 19$, $K_2 = 21$ $K_3 = 23$.

TABLE 3. Volume of Transactions for portfolio composed of two calls and one put. Values of parameters as in Figures 5-6. $V_{T_0}(X) = 2.3$

No. of Transactions (R)	VT(BΠ)	VT(HΠ)	AverageVolTrBS
$R = 5$	2.29	0.61	24.8
$R = 10$	2.31	0.66	38.88
$R = 15$	2.29	0.65	49.6
$R = 20$	2.29	0.67	58.35
$R = 25$	2.30	0.65	65.63
$R = 30$	2.29	0.65	72.13

[15] S. Mallat (1989), Multiresolution aproximations and wavelets orthonormal bases for $L^2(\mathbb{R})$. *Trans. of Amer. Math. Soc.*, **315**, 69-87.
[16] J. Neveu (1975): *Discrete-Parameter Martingales*, North-Holland.
[17] A. Pechtl (1995), Classified information. *RISK*, **8 (6)**, 59-61.
[18] G.A. Edgar and L. Sucheston (1992):*Stopping Times and Directed Processes.* Cambridge University Press.
[19] W. Willinger (1987), *Pathwise Stochastic Integration and Almost-Sure Aproximation of Stochastic Processes.*, Thesis, Cornell University.

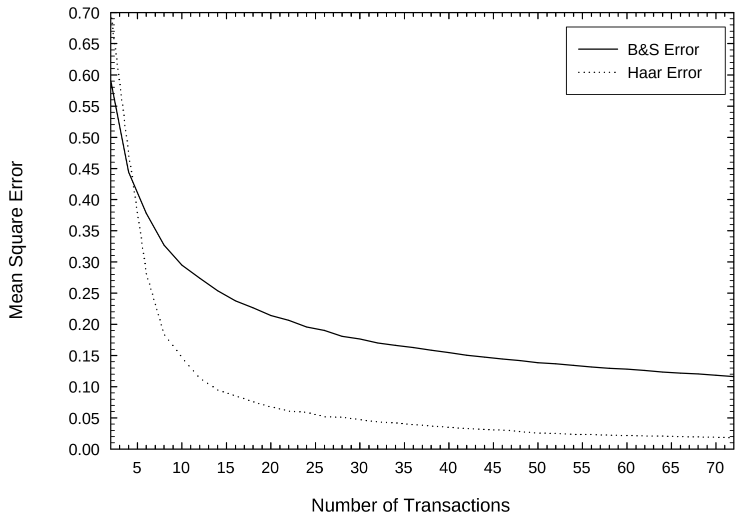


FIGURE 6. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter R. Values of the parameters used: as in Figure 5.

[20] W. Willinger and M.S. Taqqu (January 1991), Toward a convergence theory for continuous stochastic securities market models. *Mathematical Finance*, **Vol. 1**, No. 1, 55-99.

[21] Petre G. Zhang (1998): *Exotic Options*, second edition. World Scientific Publishing Co.

DEPARTAMENTO DE MATEMÁTICA. FACULTAD DE CIENCIAS EXACTAS Y NATURALES. UNIVERSIDAD NACIONAL DE MAR DEL PLATA. FUNES 3350, MAR DEL PLATA 7600, ARGENTINA.

E-mail address: pedrojc@mdp.edu.ar

DEPARTMENT OF MATHEMATICS, PHYSICS AND COMPUTER SCIENCE, RYERSON UNIVERSITY, 350 VICTORIA ST., TORONTO, ONTARIO M5B 2K3, CANADA.

E-mail address: ferrando@ryerson.ca

DEPARTAMENTO DE MATEMÁTICA. FACULTAD DE CIENCIAS EXACTAS Y NATURALES. UNIVERSIDAD NACIONAL DE MAR DEL PLATA. FUNES 3350, MAR DEL PLATA 7600, ARGENTINA.

E-mail address: algonzal@mdp.edu.ar