HAAR SYSTEMS FOR EFFICIENT HEDGING AND APPROXIMATION OF DERIVATIVES

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

ABSTRACT. We present a new discretization of financial instruments in terms of martingale expansions constructed using Haar wavelets systems. Examples of these systems are constructed which illustrate the discrete, space-wise, nature of the approximations. Expansions on these bases give the pointwise convergence needed in several applications, in particular, we work out the details of an application to hedging an European portfolio of options. We describe natural conditions under which our Haar hedging strategy can be realized by means of a self financing portfolio consisting of binary options. We emphasize the issue of efficient approximations and formalize a notion of optimality to approximate portfolios of 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 to construct the portfolio replication, this involves an implicit linear spatial approximation of the option. This approximate hedging gives a pointwise error the quality of which depends on the efficiency of this space-time approximation. We note that an *efficient* portfolio replication will aim to reduce the number and volume of transactions for a given approximation error. Efficiency is also important in the pricing of complex path dependent derivatives when using the Monte Carlo technique. In this situation, an efficient approximation will aim at minimizing the number of computations maintaining a certain level of error.

As hinted above, the notion of *efficiency* depends on the application at hand. Despite of this, there are theoretical guidelines on how to approach the problem under a variety of settings. The area of nonlinear approximation (see [7] and [11] and the references given there) studies efficient representation of functional classes. For specific functional classes, wavelets have been proven optimal for the task of compression (efficient storage), noise removal, fast computation, etc.

Presently, the use of wavelets techniques in finance has been directed towards time series processing, (see for instance [14] and [33]) and the fast numerical solution, via the Galerkin method, of Black-Scholes equation (see [28] and [29] for a recent account of these issues). These approaches make use of standard constructions of orthonormal basis of wavelets on the real line or other related higher dimensional (analytical) spaces. Our approach is different from the above as we carry our wavelet construction directly on the probability space $(\Omega, \mathcal{F}, P)$, where

Date: June 28, 2005.

The research of P. J. Catuogno is supported in part by a FAPESP grant. The research of S.E. Ferrando is supported in part by an NSERC grant.

$\mathcal{F} = \{\mathcal{F}_t\}$ is the filtration generated by a given random process and P a probability measure of interest. It is possible, and necessary for some of the applications, that the orthonormal basis $\{u_k\}$ be adapted to $\mathcal{F}$, this allows pathwise approximations which are needed for the hedging applications and applications to simulations. In our approach, the functions $\{u_k\}$ will take only two nonzero values, so they will be Haar-like. This restriction can be relaxed but, in that case, the financial interpretation of the expansions will be less natural.

This paper introduces a framework that allows the construction of Hilbert space bases which give optimal speed of convergence for the space of portfolios associated to a given collection of financial options. To achieve this optimality, we exploit the adaptability (in the sense of measure theory) of the options with respect to the underlying process. As an application, we construct approximating portfolios of binary options for hedging general financial claims. The approximations are efficient in the sense that they require a small number and volume of transactions. This is achieved thanks to the property of localizations of our wavelets. Some of these characteristics are in contrast with other approaches to hedging ([1], [6] and [22]) which, similarly to our approach, use portfolios of simple options to hedge complex portfolios.

In the literature there are several research streams that use Hilbert space basis for approximation of contingent claims. For example, [5] uses eigenfunction expansions to price options in a general setting. Reference [26] describes possible uses of a Hilbert space basis for valuation and hedging. Our contribution is different from these approaches as we emphasize the following (related) topics for our approximations: efficiency, the approximations are given by an orthonormal basis (with respect to the probability measure of interest), our basis can be adapted to a collection of options or underlying process, they are martingales and the computations can be assembled in a convenient algorithm. Moreover, the basis elements can be easily interpreted in terms of basic financial instruments and transactions and they can be optimized, for the sake of efficiency, for specific tasks. Our approach can be succinctly characterized by indicating that we are interested in the construction of approximating set of functions for efficient representation and computation of financial instruments.

To indicate the essence of our approach, we point to (2.1). The right hand side of (2.1) 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 writing of the conditional expectation as a Fourier expansion, the inner products $\langle X, u_n \rangle$ 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, in order to give efficient representations of X. Efficient representations of functional classes is a chief concern of computational harmonic analysis, see for example [11], [9] and [10].

Finally, we believe that the research community is well aware of the formal connections between martingales and wavelets (see for instance [12] and [30]). In particular, R. Gundy has made use of martingale theory in a wavelets setting [17] and has exposed related points of views [18]. An important reference in this regard is [15] which is very much related to our constructions.

The paper is organized as follows, Section 2 defines H-systems, develops the relationship between H-systems and sequences of partitions and states a result on existence of H-systems for an important class of stochastic processes. Section 3 introduces simple examples of H-systems in basic financial settings. Section 4 motivates and develops our main application to hedging a given European portfolio of options (we also outline other applications.) Section 5 formalizes the way in which H-systems are best basis for a class of financial portfolios (Theorem 4). Optimized constructions of H-systems are also described. Section 6 presents numerical examples. Section 7 summarizes the main results of the paper. Appendix A presents notation, formulae and the Multiresolution Analysis Algorithm, which are needed in computations. Appendix B presents a simple example as a complement. Appendix C presents tables and figures from the numerical experiments

2. H-Systems

Let $(\Omega, \mathcal{A}, P)$ denote an arbitrary probability space. The notation $||\ ||^2 = \langle , \rangle$ stands for the inner product on $L^2(\Omega, \mathcal{A}, P)$. The following Gundy's [16] definition is motivated by the standard Haar system of $L^2([0,1])$.

Definition 1. *An orthonormal system of functions $\{u_k\}_{k\geq 0}$ defined on Ω is called an H-system if and only if for any $X \in L^2(\Omega, \mathcal{A}, P)$*

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

where $\mathcal{A}_n = \sigma(u_0, \ldots, u_n)$. 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 also use the notation $\mathcal{A}_\infty = \sigma(\cup_{n\geq 0}\mathcal{A}_n)$. In applications we will make use of the pointwise convergence of (2.1) which holds due to the martingale convergence theorem [31]. Moreover, if $p \in [1, \infty)$ is a given real number then, for every $X \in L^p$, the sequence $X_{\mathcal{A}_n} = \mathbf{E}(X|\mathcal{A}_n)$ converges a.s. and in L^p to $X_\infty = \mathbf{E}(X|\mathcal{A}_\infty)$.

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 5. The following proposition, which is proven in [16], gives an alternative characterization of H-systems equivalent to Definition 1.

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 $\mathcal{A}_n$ 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 $\mathcal{A}_n$ (hence this atom becomes its support). Consequently, $\mathcal{A}_{n+1}$ consists of n atoms from $\mathcal{A}_n$ and two more atoms obtained by splitting the remaining atom from $\mathcal{A}_n$.*

In view of the above proposition and its corollary, the functions in an H-system are natural generalizations of classic Haar functions, as the next definition states.

Definition 2. *Given $A \in \mathcal{A}$, $P(A) > 0$, a function ψ is called a Haar function on A if there exist $A_i \in \mathcal{A}$, $A_0 \cap A_1 = \emptyset$, $A = A_0 \cup A_1$, $\psi = a\ \mathbf{1}_{A_0} + b\ \mathbf{1}_{A_1}$ and*

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

2.1. **Basic Properties of H-Systems.** This section introduces some elementary properties of H-systems and partitions. Along with Appendix A, 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 refer to Section 4.1.
It should be clear, from Corollary 1, that an H-system naturally defines a *binary tree* of partitions, these are formally introduced in the next definition.

Definition 3. *A sequence of partitions of Ω, $\mathcal{Q} := \{\mathcal{Q}_j\}_{j\geq 0}$, is called a* binary sequence of partitions *if for $j \geq 0$, the members of $\mathcal{Q}_j$ have positive probability, $\mathcal{Q}_0 = \{\Omega\}$, and for $j \geq 1$, $A \in \mathcal{Q}_j$ if and only if it is also a member of $\mathcal{Q}_{j-1}$ or there exists another member A' of $\mathcal{Q}_j$ such that $A \cup A' \in \mathcal{Q}_{j-1}$.*

We set $A_{0,0} := \Omega$, hence $\mathcal{Q}_0 = \{A_{0,0}\}$. For $j \geq 1$, if $A \in \mathcal{Q}_j$ and $A = A_{k,i} \in \mathcal{Q}_{j-1}$ then A preserves its index. Otherwise (i.e. $A \notin \mathcal{Q}_{j-1}$, and not yet indexed) then there exists $A_{k,i} \in \mathcal{Q}_{j-1}$ and $A' \in \mathcal{Q}_j$ 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 $A_{j,i}$. Notice that $\mathcal{Q}_j$ can have at most 2^j members, and if $A_{k,i} \in \mathcal{Q}_j$ then $k \leq j$ and $0 \leq i \leq 2^k - 1$. The name scale is borrowed from wavelet theory where it indicates the extent of the localization (or *resolution*) of the wavelet. The figure displayed in the example of appendix B will clarify the indexation. The information about the splitting of atoms is stored in the indexation, it allows to rearrange a given binary sequence of partitions so as to collect all atoms with the same scale parameter j. Atoms at lower levels, which complete a partition and will not be further split, are also included. This will be formalized in the next definition.

Definition 4. *A binary sequence of partitions $\mathcal{R} = \{\mathcal{R}_j\}$ will be called a mutiresolution sequence (of partitions) if each $A_{k,i}$ belonging to $\mathcal{R}_j$, with $j > k$, also belongs to $\mathcal{R}_{j'}$ for all $j' \geq j$.*

Observe that if $\mathcal{R}$ is a multiresolution sequence of partitions and $A_{k,i} \in \mathcal{R}_j$ with $k < j$, $A_{k,i}$ has not been split since level k and will not be further split, while if $k = j$, $A_{j,i}$ comes from the splitting of an atom of $\mathcal{R}_{j-1}$. To this type of partitions we will associate a Multi-resolution 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.
The following sets of indexes will be used shortly and in Appendix A, consider $j \geq 0$ and let

$$\begin{gathered} I_j \equiv \{i : A_{j,i} \in \mathcal{R}_j \text{ and } A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}\}, \text{ and} \\ K_j \equiv \{(k,i) : A_{k,i} \in \mathcal{R}_j\ \}. \end{gathered} \tag{2.3}$$

Natural and computationally useful binary sequences of partitions are the dyadic ones, these are sequences $\{\mathcal{Q}_j\}_{j\geq 0}$ such that each atom of $\mathcal{Q}_{j-1}$ split into two atoms of $\mathcal{Q}_j$. Since the usual Haar wavelet system is associated with this kind of sequences, we introduce the following general definition.

Definition 5. *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 (or for all $j<J-1$).*

Theorem 1. *Every H-system induces naturally a multiresolution sequence of 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:
- If $k<j$ 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 $A_{j,i}=\text{supp}u_r$ for some u_r. Then we add

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

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

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

$$\psi_{j,i}(\omega)=\begin{cases} a_{j,i} & \text{if} \quad \omega\in A_{j+1,2i},\\ b_{j,i} & \text{if} \quad \omega\in A_{j+1,2i+1} \text{ and},\\ 0 & \text{if} \quad \omega\notin A_{j,i}.\end{cases} \tag{2.4}$$

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 $\omega\in\Omega$, indeed we choose

$$a_{j,i}=\sqrt{\frac{P(A_{j+1,2i+1})}{P(A_{j+1,2i})P(A_{j,i})}} \tag{2.5}$$

and

$$b_{j,i}=-\sqrt{\frac{P(A_{j+1,2i})}{P(A_{j+1,2i+1})P(A_{j,i})}}. \tag{2.6}$$

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 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. *If $\mathcal{R}$ is a multiresolution sequence 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 $\{v_k\}_{k\geq 0}$.*

In order to explain and justify our use of Definition 1 and Theorem 1 we will need some results on discrete approximations of continuous stochastic processes. This is presented in Proposition 2 below, it gives an existence result that can be employed in our applications. Alternative constructions of H-systems are presented in Section 5, also, reference [15] describes ways to construct H-systems associated with nested partitions.

Let $(\Omega, \mathcal{A}, P)$ be a complete probability space and $S = (S_t : 0 \leq t \leq T)$ be a continuous stochastic process defined on this probability space. Let $\mathcal{F} = \{\mathcal{F}_t : 0 \leq t \leq T\}$ be the filtration where $\mathcal{F}_t$ is the completion of $\sigma(S_r : 0 \leq r \leq t)$. Following W. Willinger [34], we introduce the notion of skeleton-approach for stochastic processes.

Definition 6. *A* continuous-time skeleton approach *of S is a triple $(I^\xi, \mathcal{F}^\xi, \xi)$, consisting of a index-set I^ξ, a filtration $\mathcal{F}^\xi = \{\mathcal{F}^\xi_t : 0 \leq t \leq T\}$ the skeleton filtration and a $\mathcal{F}^\xi$-adapted process $\xi = (\xi : 0 \leq t \leq T)$ such that verifies:*

(1) *$I^\xi = \{0 = t(\xi, 0) < \ldots < t(\xi, N_\xi) = T\}$, where $N_\xi < \infty$.*
(2) *For each t, $\mathcal{F}^\xi_t$ is a finitely generated sub σ-algebra of $\mathcal{F}_t$, with atoms $\mathcal{P}^\xi_t$.*
(3) *For $t \in [0, T] - I^\xi$, we set $\mathcal{F}^\xi_t = \mathcal{F}^\xi_{t(\xi,k)}$ if $t \in [t(\xi, k), t(\xi, k+1))$for some $0 \leq k < N_\xi$.*
(4) *For each $0 \leq t \leq T$, $\xi_t = \mathbf{E}(S_t \mid \mathcal{F}^\xi_t)$.*

Definition 7. *A sequence $(I^{(n)}, \mathcal{F}^{(n)}, \xi^{(n)})$ of continuous time skeletons of S will be called a* continuous-time skeleton approximation *of S if the following three properties hold.*

(1) *The sequence $I^{(n)}$ of index satisfies*

$$\lim_{n\to\infty} |(I^{(n)}| = 0$$

where $|(I^{(n)}| \equiv \max\{|t(\xi^{(n)}, k) - t(\xi^{(n)}, k-1)| : 1 \leq k \leq N^{(n)}\}$, and $I \equiv \cup_n I^{(n)}$ is a dense subset of $[0, T]$,
(2) *For each $0 \leq t \leq T$, $\mathcal{F}^{(n)}_t \uparrow \mathcal{F}_t$,*
(3) *$P(\{\omega \in \Omega : \lim_{n\to\infty} sup_{0\leq t\leq T} |S_t(\omega) - \xi^{(n)}_t(\omega)| = 0\}) = 1$.*

The fundamental result of W. Willinger ([34] pp 52, Lemma 4.3.1) is stated next, it guarantees the existence of continuous-time skeleton approximations for continuous processes. These discrete pathwise approximations are finite in space and time.

Lemma 1. *There exist a continuous-time skeleton approximation for S.*

Each continuous time skeleton $(I^{(\xi)}, \mathcal{F}^{\xi}, \xi)$ of S determines a sequence of nested finite partitions $\{\mathcal{P}^{\xi}_{t_m}\}$. Clearly, there exists a multiresolution sequence of partitions $\{\mathcal{R}^{\xi}_j\}_{j\geq 0}$ such that $\mathcal{R}_{j_m} = \mathcal{P}^{\xi}_{t_m}$ for $0 = j_0 < j_1 < ... < j_N$. Now, we can construct a finite family of H-systems associated to the continuous time skeleton $(I^{(\xi)}, \mathcal{F}^{\xi}, \xi)$ of S applying Theorem 1 to the multiresolution sequences $\{\mathcal{R}^{\xi}_j\}_{j\geq 0}$. Obviously, these H-systems are adapted to the filtration $\mathcal{F}^{\xi}_{t_m}$, that is $\psi_{j,i} \in \mathcal{F}^{\xi}_{t_m}$ for $j \leq j_m$.

Proposition 2. *Let $(\Omega, \mathcal{A}, P)$ be a complete probability space and $S = (S_t : 0 \leq t \leq T)$ be a continuous stochastic process defined on this probability space. Let $\mathcal{F} = \{\mathcal{F}_t : 0 \leq t \leq T\}$ be the filtration where $\mathcal{F}_t$ is the completion of $\sigma(S_r : 0 \leq r \leq t)$. Then there exist a sequence of finite H-systems $(\mathcal{H}^{(n)} = \{\psi^n_{j,i}\})$ and two sequences of finite index $(I^{(n)} = \{0 = t^n_0 < ... < t^n_{N_n} = T\})$ and $(J^{(n)} = \{0 = j^n_0 < ... < j^n_{M_n}\})$ such that*

(1) *$\psi^n_{j,i} \in \mathcal{F}_{t^n_m}$ for $j \leq j^n_m$.*
(2) *For each $0 \leq t \leq T$,*

$$\lim_{n\to\infty} \sup\{|S_t - \xi^{(n)}_t| : 0 \leq t \leq T\} = 0 \ \ a.e.$$

where $\xi^{(n)}_t = \sum_{j\leq j^n_m} \langle S_t, \psi^n_{j,i}\rangle \ \psi^n_{j,i}$ for $t \in [t^n_m, t^n_{m+1})$.

Proof. Let $(I^{(n)}, \mathcal{F}^{(n)}, \xi^{(n)})$ be a continuous-time skeleton approximation of S. We construct for each n an H-system $(\mathcal{H}^{(n)} = \{\psi^n_{j,i}\})$ associated to the sequence of partitions $\{\mathcal{R}^{(n)}_j\}_{j\geq 0}$, as we explain above. In order to conclude the proof it is sufficient to observe that $\xi^{(n)}_t = \mathbf{E}(S_t \mid \mathcal{F}^{(n)}_t) = \sum_{j\leq j^n_m} \langle S_t, \psi^n_{j,i}\rangle \ \psi^n_{j,i}$ for $t \in [t^n_m, t^n_{m+1})$. □

3. Examples

To be specific, and for the reader's convenience, we describe some simple examples of H-systems in familiar financial contexts. Both examples are Haar systems, namely, they are generated by dyadic partitions, see Definition 5. More general constructions follow from the developments in Section 5. Another example, in Appendix B, constructs an H-system associated to a sequence of binary partitions that is not dyadic and illustrates the case of *multiresolution* sequence of partitions.

1. **Haar-Systems for the binomial model:**

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

$$S_{t_{i+1}} = S_{t_i} H_{i+1}, \ i = 0, 1, ..., n-1,$$

where $\{H_i\}^n_{i=1}$ is an independent set of 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 setting can be formalized in terms of the probability space $(\Omega, \mathcal{A}, P)$, where $\Omega := \{\omega : \{t_1, ..., t_n\} \to \{U, D\}\}$, $\mathcal{A} \equiv \mathcal{P}(\Omega)$ and P the corresponding product probability measure. Then $S : \Omega \times \{t_0, t_1, ..., t_n\} \to \mathbb{R}$, $S_0(\omega) := S(\omega, t_0) = S_0$ and $S_t(\omega) := S(\omega, t) = S_0 \prod_{t_i \leq t} \omega(t_i)$.

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

$$A_{j+1,2i+1} = A_{j,i} \cap \{\omega(t_{j+1}) = U\},\ A_{j+1,2i} = A_{j,i} \cap \{\omega(t_{j+1}) = D\}. \tag{3.1}$$

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

$$\begin{aligned} &u_0 \equiv 1, \\ &u_{2^j+i} = \tfrac{1}{\sqrt{P(A_{j,i})}}(\sqrt{\tfrac{p}{q}}\ 1_{A_{j+1,2i}} - \sqrt{\tfrac{q}{p}}\ 1_{A_{j+1,2i+1}}), \end{aligned} \tag{3.2}$$

It is clear, from Theorem 1, that $\{u_k\}_{0 \leq k \leq 2^n - 1}$ is an H-system for $L^2(\Omega, \mathcal{A}, P)$. Observe that for each $j \geq 0$ 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 $\{u_k\}_{0 \leq k \leq 2^n - 1}$ 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 $\mathcal{F}_{t_j} = \sigma(S_{t_0}, ...S_{t_j})$.

2. 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 underlying process 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 $[T_0, T]$ (see Section 6) 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. It is convenient to introduce first 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 $0 \leq i \leq 2^j - 1$,

(3.3)

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

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

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 $\mathcal{P}_j = \{A_{j,i}\}, i = 0, \ldots, 2^j - 1$, where the atoms satisfy

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

Setting $m = 2^j$ and $\mathcal{A}_m = \sigma(\{A_{j,i} : i = 0, \ldots, m-1\})$ gives $\mathcal{A}_\infty = \sigma(\cup_{m \geq 0} \mathcal{A}_m) = \sigma(S_T)$. Notice that the above atoms correspond to partitioning the range of S_T. It follows from Theorem 1 that there 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 (3.3), 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 $\{t_k, t_{k+1}\}$, $1 \leq k \leq n_T - 1$ is done setting $J_k = j_1 + \ldots + j_k$. Consider $1 \leq j \leq j_{k+1}$ and $i = 0, \ldots, 2^{J_k+j} - 1$, let $i = p2^j + q$ (p and q are respectively the quotient and the residue in the integer division of i by 2^j). Then define the sets

$$A_{J_k+j,i} \quad = \quad \{\omega \in A_{J_k,p} \mid c_q^j < W_{t_k,t_{k+1}}(\omega) \leq c_{q+1}^j\}.$$

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

4. Applications to Hedging and Pathwise Simulation

This section illustrates how H-systems can be applied in financial mathematics. It develops in detail a theory of hedging based on binary options, the martingale property of the H-system is put to use in this theory. There is also a brief description of the use of our approximations as control variate for Monte Carlo simulations and an outline of an application to American options. For the sake of simplicity, 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 [2] for background. Let $B = (B(t) = e^{rt})$ be the *bond* and a non-negative adapted continuous 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 multiresolution partitions as described in Definition 3, associated, via Theorem 1, with the H-system $\{\phi_{0,0}, \psi_{j,i}\}$ defined on Ω, and an European derivative X in $L^2(\Omega, \sigma(\cup_{j\geq 0}\mathcal{R}_j), P)$.

4.1. **Haar Hedging.**

Motivations and Meaning: A sample of references describing hedging with options is given by [1], [6] and [22]. In contrast to previous results, our approach is general, in the sense that allows for general underlyings *and* options types, and, more importantly, our approximations address the issue of the number and volume of transactions. We would like to mention that the idea of using binary options for approximations has been previously treated in [32].

Let us explain the basic idea in this section, 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 $\mathcal{A}_n$ 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 (2.1) can be realized by means of a dynamic portfolio of binary options. The left hand side of (2.1) 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.

The goal of this section is to find pathwise approximations to X with small error and at the same time incurring in a small number of financial transactions (when

implementing the associated 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 to reduce the transaction costs while achiving 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 8. *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 corresponds to a more standard quantity in wavelet theory.
We now explain the empirical meaning of the representation (2.1) and compare it with "static" hedging and briefly comment on the relationship to delta-hedging. Usually, static option replication involves hedging an option X with other options, see for example [6]. For simplicity, consider an option X that initiates at T_0 and expires at T with $V_{T_0}(X)$ denoting the risk neutral price of X. Lets study an example that shows a key problem with the standard static hedging. Consider a digital option with payoff $X = \mathbf{1}_{S_T \geq K}$, approximate this digital option with the following portfolio

$$\Pi = \frac{1}{K_2 - K_1} (X_1 - X_2) \tag{4.1}$$

where we go long on a European call $X_1 = (S_T - K_1)_+$ with strike K_1 and short on a European call $X_2 = (S_T - K_2)_+$ with strike K_2, and $K_1 < K < K_2$. We obtain a better and better approximation to X by considering $(K_2 - K_1) \to 0$. By risk neutrality we then have $V_{T_0}(X) \approx V_{T_0}(\Pi)$ but the volume of transactions for Π (which in this static example is a constant) is equal to

$$VT(\Pi) = \frac{1}{K_2 - K_1} [V_{T_0}(X_1) + V_{T_0}(X_2)] \tag{4.2}$$

which can be arbitrarily large as $(K_2 - K_1) \to 0$. In short, when decreasing the error of approximation we have the undesirable effect of increasing the volume of transactions. This is due to the fact that the approximation $X \approx \Pi$ is obtained by cancellation of (unbounded) terms and each term entering in this approximation will contribute separately to the volume of transactions. The discontinuity in X just exacerbates this phenomena.

We now explain how our proposed Haar hedging overcomes the above type of problem. First 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 $\mathcal{A}_i$ for some $i \leq k$ and $A = A_0 \cup A_1$ is an atom of $\mathcal{A}_{i-1}$. 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 is 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 (2.1) 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 (2.1) 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. This localization is also the key for our approximations to have a small volume of transactions, see numerical examples in Section 6. It is also recognized in signal processing applications that localization of wavelets is a key property to represent discontinuities efficiently [9], we have observed this phenomena also in our numerical examples.

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.

Formal Developments:
As a sufficient condition for the atoms in a multiresolution sequence 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 9. *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 10. *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 there exist $B \in \mathcal{P}$ such that $t - supp(B) \subset [a, b]$, and for all $B \in \mathcal{P}$ $t - supp(B) \subset [a, b]$ or $t - supp(B) \subset [T_0, a]$. 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 11. *Let $\mathcal{R} = \{\mathcal{R}_j\}_{J \geq j \geq 0}$ be a sequence of multiresolution 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 $j_1, \ldots, j_n$ the levels of localization of $\mathcal{R}$.*

The financial blocks underlying $\mathcal{R}$ are the binary options

$$\mathbf{B}_{j,i} = (\mathbf{1}_{A_{j,i}}(t) \equiv \mathbf{1}_{[t_{s+1},T]}(t)\mathbf{1}_{A_{j,i}}), \quad j_s \leq j \leq j_{s+1}, \tag{4.3}$$

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 12. Assumption on Financial Realization: *The multiresolution 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: $\Psi_{j,i} = (\Psi_{j,i}(t) \equiv \mathbf{1}_{[t_{s+1},T]}(t)\psi_{j,i})$, 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 (4.3).

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 3 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 re-balanced 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 previously indicated, 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 re-balancing 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 $t \in [t_0, t_1)$

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

$$e^{-r(T-t_1)}\, d_j[i] \text{ invested in } \Psi_{j,i}\ j = 0, \ldots j_1 - 1, i \in I_j, \tag{4.4}$$

where the coefficients $d_j[i]$ are given by (A.6).
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 re-balance 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 $e^{-r(T-t_{s-1})}x_{k_0}[i_0]$ (where we used the notation from (A.10)). There are two cases to consider, the event is split or not at the next level.

I) In the case A_{k_0,i_0} splits, $k_0 = j_{s-1}$ as we remarked before, the coordinates of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ for $t \in [t_{s-1}, t_s)$ are

$$e^{-r(T-t_{s-1})} x_{k_0}[i_0] \text{ invested in the bond and}$$

$$e^{-r(T-t_s)}\ d_j[i] \text{ invested in } \Psi_{j,i}\ j = j_{s-1}, \dots j_s - 1, i \in I_j^{i_0}, \tag{4.5}$$

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

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

$$e^{-r(T-t_{s-1})} x_{k_0}[i_0], \tag{4.6}$$

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 (4.4) when $t \in [t_0, t_1)$. It is clear from (A.5) 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} the process is in some event A_{k_0,i_0} with $(k_0, i_0) \in K_{j_{s-1}}$, and assume

$$V_{t_{s-1}}(\mathrm{H}\Pi_{\mathcal{R}}(X)_{t_{s-1}^-})(\omega) = e^{-r(T-t_{s-1})}\ \mathbf{E}(X|\sigma(\mathcal{R}_{j_{s-1}}))(\omega) = e^{-r(T-t_{s-1})} x_{k_0}[i_0]$$

for $\omega \in A_{k_0,i_0}$. The re-balancing of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ at t_{s-1} is given by (4.5), for all $t \in [t_{s-1}, t_s)$, if A_{k_0,i_0} splits at the next level or by (4.6) 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 (4.5) or (4.6) is $e^{-r(T-t_{s-1})} x_{k_0}[i_0]$. Consider again case I), and $t = t_s$, by (A.8) and (4.5) we compute

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

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

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

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

For the case II), we have

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

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

□

Now, we will present the dynamic strategy $\mathrm{B}\Pi_{\mathcal{R}}(X)_t$. Let $\mathcal{R} = \{\mathcal{R}_j\}$ be a multiresolution 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 re-balanced 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 the samples $x_k[i]$ are the coefficients of X in the basis $\{\mathbf{1}_{A_{k,i}} : (k,i) \in K_j\}$, see (A.10).
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,

$$e^{-r(T-T_0)}\ x_k[i]\ \mathbf{B}_{k,i} \text{ where } (k,i) \in K_{j_1}. \tag{4.7}$$

The cost of purchasing this portfolio is $V_{t_0}(\mathrm{B}\Pi_{\mathcal{R}}(X)) = e^{-r(T-t_0)}\mathbf{E}(X) = e^{-r(T-t_0)}x_0[0]$.
The inductive step will be to re-balance the portfolio at time t_{s-1}. Assume that at this time the price process is 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 $e^{-r(T-t_{s-1})}x_{k_0}[i_0]$. 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,

$$e^{-r(T-t_{s-1})}\ x_k[i]\ \mathbf{B}_{k,i} \text{ where } (k,i) \in K_{j_s}^{i_0}, \tag{4.8}$$

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

$$e^{-r(T-t_{s-1})}x_{k_0}[i_0], \tag{4.9}$$

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 $[t_{s-1}, t_s)$.

It is a simple exercise to apply the above theory for the Examples 1 and 2 from Section 3. We provide numerical examples in Section 6.

4.2. **American options.** This section illustrates how H-systems can be applied in financial mathematics to evaluate American options. We consider the previous setting of a frictionless market model with the usual assumptions, and an American derivative $Z = (Z_t)$. We know that there exists a continuous-time skeleton approximation for S. We will use it in order to approach the value of Z. In fact, we have that there exist a sequence of finite indexes $(I^{(n)} = \{0 = t_0^n < ... < t_{N_n}^n = T\})$ and filtrations $(\mathcal{F}_{t_m^n})$ such that

(1) $I^{(n)} \subset I^{(n+1)}$ and $\cup_n I^{(n)}$ is dense in $[0,T]$,
(2) $\mathcal{F}_{t_m^n} \subset \mathcal{F}_{t_m^n}^{n+1} \subset \mathcal{F}_{t_m^n}$,
(3) For each $0 \leq t \leq T$,

$$\lim_{n\to\infty} \sup\{|S_t - \xi_t^{(n)}| : 0 \leq t \leq T\} = 0 \ \ a.e.$$

where $\xi_t^{(n)} = \mathbf{E}(S_{t_m^n} \mid \mathcal{F}_{t_m^n})$ for $t \in [t_m^n, t_{m+1}^n)$.

Let $Z_j^n \equiv \mathbf{E}(Z_{t_j^n}|\mathcal{F}_{t_j^n}^n)$. We can consider (Z_j^n) as the American option obtained by projection of Z into the finite market $(\Omega, \mathcal{F}_{t_j^n}^n, P)$. Recall that the value of this option is calculated by the backward algorithm, $U_{N_n}^n = Z_{N_n}^n$ and

$$U_{j+1}^n = \max(Z_j^n, e^{r(t_{j+1}^n - t_j^n)}\mathbf{E}(U_{j+1}^n|\mathcal{F}_{t_j^n}^n)).$$

The numerical problem is to calculate the conditional expectation $\mathbf{E}(U_{j+1}|\mathcal{F}_{t_j}^n)$. It is here where the H-system can be of help. In fact, if we have the Haar-Fourier expansion of U_{j+1} it is then easy to compute the conditional expectation. In the case that we want to calculate this conditional expectation by montecarlo, we only need to compute the Haar expansion along the sampled path, this involves a small number of Haar functions (proportional to the length of the path) thanks to their localization.

4.3. **Pathwise Simulation.** We want to compute the value of $\mathbf{E}(X|\mathcal{F}_t)$ via Monte Carlo simulation. We only provide the general references [13] and [23]. By Proposition 2 we know there exists a sequence of finite H-systems $(\mathcal{H}^{(n)} = \{\psi_{j,i}^n\})$ and two sequences of finite indexes $(I^{(n)} = \{0 = t_0^n < ... < t_{N_n}^n = T\})$ and $(J^{(n)} = \{0 = j_0^n < ... < j_{M_n}^n\})$ such that

(1) $\psi_{j,i}^n \in \mathcal{F}_{t_m^n}^n$ for $j \leq j_m^n$.
(2) $(\psi_{j,i}^n)_{t_j^n \leq t}$ is an orthonormal basis of $L^2(\Omega, \mathcal{F}_t^n, P)$.

We replace $\mathcal{F}_t$ by $\mathcal{F}_t^n$ and we will concentrate in computing $\mathbf{E}(X|\mathcal{F}_t^n)$. Using the H-system, we have the representation

$$\mathbf{E}(X|\mathcal{F}_t^n) = \sum_{t_j^n \leq t} \langle X, \psi_{j,i}^n \rangle \psi_{j,i}^n.$$

In order to calculate the coefficients $\langle X, \psi_{j,i}^n \rangle$, we construct the Monte Carlo estimator

$$a_{MC}^n(j,i,M) = \frac{1}{M}\sum_{m=1}^{M} X(w^m)\psi_{j,i}^n(w^m),$$

where w^m are the sampled paths and m is the sampling index. Finally, we obtain the following Monte Carlo estimator of $\mathbf{E}(X|\mathcal{F}_t^n)$,

$$A_{MC}^n(X,t,M) = \sum_{t_j^n \leq t} a_{MC}^n(j,i,M)\psi_{j,i}^n.$$

5. H-Systems as Best Basis, Optimized Constructions

As outlined in the Introduction, a main application is to use H-systems to obtain efficient hedging strategies. Namely, while keeping a small approximation error we seek to minimize transaction costs, see Definition 8. In other applications we seek to minimize the number of computations. For technical reasons, this section concentrates in minimizing the error in the m-term nonlinear expansion defined below, (5.1). In other words, for a given error level we seek to minimize the number of Haar functions needed to achieve such an error. In applications to hedging this translates into reducing the number of transactions. See Section 6 for some numerical information 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.

5.1. **Best Basis and Compression.** We describe how H-systems provide best basis for portfolios built from a given class of options. First we need to recall some definitions from nonlinear approximation theory ([7, 8, 9, 10, 11]). Let V be a Banach space with norm $||\ ||_V$ and let $B = (b_k)$ be a Schauder basis for V (assume $||b_k||_V = 1$).

If $f \in V$, the error of the best m-term approximation, using the basis B, is given by

$$\sigma_m(f)_V \equiv \sigma_m(f, B)_V \equiv \inf_S ||f - S||_V,\ m \geq 0, \tag{5.1}$$

where $S = \sum_{k \in \Delta} a_k\ b_k$ with $|\Delta| \leq m$.
Assume we are given a class of functions $\mathcal{X} \subseteq V$ which we are interested in approximating efficiently. Furthermore, assume this class is provided with a norm $||\ ||_{\mathcal{X}}$. Given such a class, define

$$\sigma_m(\mathcal{X}_0)_V \equiv \sigma_m(\mathcal{X}_0, B)_V \equiv \sup_{f \in \mathcal{X}_0} \sigma_m(f, B)_V, \tag{5.2}$$

where $\mathcal{X}_0$ is the unit ball (relative to $||\ ||_{\mathcal{X}}$) of $\mathcal{X}$.

The following definition will be key to our developments. Let $\mathcal{B}$ be a collection of basis B as described above.

Definition 13. *B is a best basis for $\mathcal{X}$, relative to $\mathcal{B}$, if*

$$\sigma_m(\mathcal{X}_0, B)_V = O(m^{-\alpha}),\ m \to \infty \tag{5.3}$$

and no other basis $B' \in \mathcal{B}$ satisfies

$$\sigma_m(\mathcal{X}_0, B')_V = O(m^{-\beta}),\ m \to \infty \tag{5.4}$$

for a value $\beta > \alpha$.

It follows from the above framework that a main goal is to identify functional classes $\mathcal{X}$, (with an associated norm $||\ ||_{\mathcal{X}}$), with interest in applications and corresponding best bases.
Let $\mathcal{FI}$ be a collection of random variables defined on a common probability space $(\Omega, \mathcal{A}, P)$. To illustrate the formalism described below, consider the situation where $\mathcal{FI}$ is a collection of *financial instruments*. We will then look for the best orthonormal basis to approximate the family of portfolios (linear combinations) built from $\mathcal{FI}$. We could also consider $\mathcal{FI}$ to consist of the value processes of a family of financial instruments. In this case, the family of portfolios will have to be defined with the appropriate hypothesis of predictability on the coefficients. The results below could be recast in this extended setting due to the generality of Burkholder's result. For simplicity, we will not present this more general case.

We will need Burkholder's inequality about martingale differences which we recall next. A martingale difference $\{e_k\}$ is said to be subordinated to another martingale difference $\{d_k\}$ if and only $|e_k| \leq |d_k|$ for all k.

Theorem 3 (Burkholder). *Let $\{e_k\}$ and $\{d_k\}$ be two sequences of martingale differences defined on a common probability space. If $\{e_k\}$ is subordinated to $\{d_k\}$ and*

both are L^p-sequences, for some $1 < p < \infty$, then there exists an absolute constant c_p such that

$$||\sum_k e_k||_p \leq c_p \; ||\sum_k d_k||_p. \tag{5.5}$$

Definition 14. *Let H denote a Hilbert space and B an orthonormal basis of H. Given a normed functional class $(\mathcal{X}, ||\ ||_{\mathcal{X}})$, with $\mathcal{X} \subseteq H$, we say that B is an unconditional basis for $\mathcal{X}$ if for any $f = \sum_k \langle f, u_k \rangle u_k$ (with convergence in $||\ ||_{\mathcal{X}}$) and $\tilde{f} \equiv \sum_k m_k \langle f, u_k \rangle u_k$, with $\{m_k\}$ an arbitrary sequence of real numbers with $m_k \in [-1, 1]$, we have*

$$||\tilde{f}||_{\mathcal{X}} \leq c_0 ||f||_{\mathcal{X}}, \tag{5.6}$$

for some absolute constant c_0.

Assume we are given an H-system $B = \{u_k\}$, with associated atomic sigma algebras $\mathcal{A}_n$, therefore, B is an orthonormal basis of $L^2(\Omega, \mathcal{A}_\infty, P)$. We will make use of the spaces $L^p(\mathcal{C}) = L^p(\Omega, \mathcal{C}, P)$, where p is a given integer in the range $1 < p < \infty$ and $\mathcal{C}$ is a given sigma algebra. We require the following condition.

Definition 15. *We say that the two objects, $\mathcal{FI}$ and $B = \{u_k\}$, are admissible if*

$$\sigma(\mathcal{FI}) \subseteq \mathcal{A}_\infty \text{ and } \mathcal{FI} \subseteq L^p(\sigma(\mathcal{FI})). \tag{5.7}$$

In particular, this requirement makes the inner products $\langle \gamma, u_k \rangle$, where $\gamma \in \mathcal{FI}$ and $u_k \in B$, well defined. A fact that we will use frequently below is that if $g \in L^p(\Omega, \sigma(\mathcal{A}_\infty), P)$, $1 \leq p < \infty$, then $\sum_{k=0}^n \langle g, u_k \rangle u_k$ converges to g pointwise and in the L^p norm ([31]).

We augment $\mathcal{FI}$ to $\mathcal{FI}'$ defined by

$$\mathcal{FI}' = \mathcal{FI} \cup \{u_k : \exists \gamma \in \mathcal{FI} \text{ such that } \langle \gamma, u_k \rangle \neq 0\}. \tag{5.8}$$

For a given integer p and constant $c > 0$, we will define a functional class $\mathcal{X}$ endowed with the norm $||\ ||_{\mathcal{X}} = ||\ ||_p$.

Definition 16. $\mathcal{X} = \{f \in L^p(\sigma(\mathcal{A}_\infty)) : ||f||_2 \leq c, \; \forall \epsilon > 0 \; \exists$ *real numbers d_i and $\delta_i \in \mathcal{FI}', i = 1, \ldots, n$, such that $||f - \sum_i d_i \delta_i||_p \leq \epsilon\}$.*

Notice that $\mathcal{X}$ is (essentially) the L^p-closure of $span(\mathcal{FI}')$; for technical reasons we also need to intersect with the L^2 ball of radius c.

Use $\mathcal{B}$ to denote the collection of all orthonormal basis in $L^2(\Omega, \mathcal{A}_\infty, P)$; here is our main result.

Theorem 4. *Assume the objects $\mathcal{FI}$ and $B = \{u_k\}$, as described previously, are admissible and $1 < p < \infty$. Then, the H-system $B = \{u_k\}$ is a best basis, relative to $\mathcal{B}$, for the class $(\mathcal{X}, ||\ ||_p)$ from Definition 16.*

Proof. According to the main result in [9] it is enough to stablish that B is unconditional for $\mathcal{X}$. To this end consider $f = \sum_k \langle f, u_k \rangle u_k \in \mathcal{X}$, as mentioned before, this expansion converges in $L^p(\sigma(\mathcal{A}_\infty))$ and pointwise. Given a small $\epsilon_1 > 0$ (to be fixed later) consider $v = \sum_{i=1}^n d_i \delta_i$, with $\delta_i \in \mathcal{FI}'$, so that $||f - v||_p \leq \epsilon_1$. Set $v_1 = \sum_{k=0}^{n_1} \langle v, u_k \rangle u_k$ where n_1 is chosen so that $||v - v_1||_p \leq \epsilon_1$. Consider $\tilde{f} = \sum_k m_k \langle f, u_k \rangle u_k$, notice that $||\tilde{f}||_2 \leq c$ and that $\tilde{f} \in L^p(\sigma(\mathcal{A}_\infty))$ by a simple application of Burkholder's inequality. Define

$$\tilde{v}_1 = \sum_{k=0}^{n_1} m_k \langle v, u_k \rangle u_k = \sum_{k=0}^{n_1} m_k (\sum_{i=1}^{n} d_i \langle \delta_i, u_k \rangle u_k). \tag{5.9}$$

Without loss of generality we may assume that the inner products $\langle \delta_i, u_k \rangle$ appearing in (5.9) are non zero, it then follows that the basis elements u_k appearing in (5.9) are in $\mathcal{FI}'$. Therefore, in order to check that $\tilde{f} \in \mathcal{X}$ it is enough to prove that for any $\epsilon > 0$ we have

$$||\tilde{f} - \tilde{v}_1||_p \leq \epsilon. \tag{5.10}$$

According to our notation we have $\tilde{v} = \sum_{k=0}^{\infty} m_k \langle v, u_k \rangle u_k$. In order to establish (5.10) we proceed as follows

$$||\tilde{f} - \tilde{v}_1||_p \leq ||\tilde{f} - \tilde{v}||_p + ||\tilde{v} - \tilde{v}_1||_p =$$

$$||\sum_{k=0}^{\infty} (m_k \langle f - v, u_k \rangle u_k)||_p + ||\sum_{k=n_1+1}^{\infty} m_k \langle v, u_k \rangle u_k||_p \leq \tag{5.11}$$

$$c_p \; ||f - v||_p + c_p \; ||\sum_{k=n_1+1}^{\infty} \langle v, u_k \rangle u_k||_p,$$

where c_p is the constant appearing in Burkholder's inequality which we applied twice to the martingale differences $e_k = m_k \langle f - v, u_k \rangle u_k$, $d_k = \langle f - v, u_k \rangle u_k$ and $e_k = m_k \langle v, u_k \rangle u_k$, $d_k = \langle v, u_k \rangle u_k$. Therefore by setting $\epsilon_1 = \frac{\epsilon}{c_p \; 2}$, equation(5.10) follows from (5.11). Finally (5.6) follows from Burkholder's inequality with $c_0 = c_p$. □

In the previous theorem we think of the functional class $\mathcal{X}$ as (modelling) the space of portfolios (linear combinations) of the given class $\mathcal{FI}$ which, for example, could be taken to be the collection of all Calls and Puts available for trading in the market. Theorem 4 tell us what type of orthonormal basis are best basis for this space of portfolios. Recall that best basis means optimal decay of the size of the rearranged inner products. The norm used to measure the error is the L^2 norm, it could be interesting to introduce other measures of nonlinear approximation (analogously to [10] and [8]). Missing in Theorem 4 is an estimate for the decay parameter α in Definition 13. In general, the value of α will depend on the actual H-system used. Section 5.2 presents constructions of H-systems which optimize the value of this parameter.

Compression: Notice that in the case when V is a Hilbert space and having a best basis for $\mathcal{X}$ allows us to realize (5.3) with the following simple procedure which we term *compression*. Let $u_{k_i}, i = 0, \ldots$ be a new indexing for our Haar system $\{u_k\}, k = 0, \ldots$, such that $|\langle X, u_{k_{i+1}} \rangle| \geq |\langle X, u_{k_i} \rangle|$. So our m-term compressed approximation, which we will denote by $X^c_{(m)}$, is given by

$$X^c_{(m)} = \sum_{i=0}^{m-1} \langle x, u_{k_i} \rangle u_{k_i}. \tag{5.12}$$

Therefore, in order to realize the best m-term approximation we order the (absolute value of) inner products by decreasing size and keep the m largest of them; we can then perform the associated reconstruction. Details of the computations involved are presented in Appendix A.

5.2. **Optimized Constructions.** Given a functional class $\mathcal{X}$, implicit in the presentation of Section 5.1 is the construction of associated best basis B such that the best basis rate parameter α in (5.3) is large. In this section we present an interesting way to approach this problem by means of H-systems. The procedure performing the constructions will be called *greedy splitting algorithm.*
Consider $\mathcal{X} = \{X^p,\ 1 \le p \le P\}$, to be a *finite* collection of random variables defined in a common probability space $(\Omega, \mathcal{A}, P)$. For example we could take $X^p = S_{t_p}$ or X^p could be a collection of options, with different expire dates etc., or the sequence X^p could be the value process of a portfolio of options.

We will define the H-system implicitly by describing a binary sequence (of partitions) $\mathcal{Q} = \{\mathcal{Q}_j\}_{j\ge 0}$. Start by setting $\mathcal{Q}_0 = \{A_{0,0} = \Omega\}$ and assume, inductively, that $\mathcal{Q}_k$, $k \le j$, has been constructed. We need some intermediate definitions in order to define $\mathcal{Q}_{j+1}$. For a given measurable set A define

$$\mathcal{C}_A = \{\psi : \text{ such that } \psi \text{ is a Haar function on } A, \text{ see Definition 2}\}. \tag{5.13}$$

Under appropriate conditions on a given random variable X, it follows that there exists $\psi_A^m = a\mathbf{1}_{A_0^m} + b\mathbf{1}_{A_1^m} \in \mathcal{C}_A$ (we will say that A_0^m and A_1^m are the *best split of* A, for the given X) satisfying

$$|\langle X, \psi_A^m\rangle| = \sup_{\psi\in\mathcal{C}_A}|\langle X - \frac{1}{P(A)}\int_A X,\ \psi\rangle| = \sup_{\psi\in\mathcal{C}_A}|\langle X,\ \psi\rangle|, \tag{5.14}$$

where *sup* is an abreviation for *supremum.* Select now $\hat{A} \in \mathcal{Q}_j$, such that there exists X^{p_j} which satisfies

$$|\langle X^{p_j}, \psi_{\hat{A}}^m\rangle| \ge |\langle X, \psi_A^m| \text{ for all } X \in \mathcal{X} \text{ and for all } A \in \mathcal{Q}_j. \tag{5.15}$$

According to the indexing of partitions previously introduced in the paper (Definition 3), $\hat{A} = A_{k,i}$ for some index (k,i), $k \le j$, now define $A_{k+1,2i} = \hat{A}_0^m$ and $A_{k+1,2i+1} = \hat{A}_1^m$. Finally, set $\mathcal{Q}_{j+1} = \mathcal{Q}_j\backslash\{\hat{A}\} \cup \{A_{j+1,2i}, A_{j+1,2i+1}\}$. Therefore, 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.
To study the convergence and the implementation of the above procedure we embark on some mathematical developments, most results are presented without proof. First we mention some notation to be used in the remaining of this section, let $\mathcal{A}_A \equiv \{B \cap A : B \in \mathcal{A}\}$; $X_A \equiv X|A$ (the restriction of X to A) and $P_A \equiv \frac{1}{P(A)}P$. It is clear that $X_A \in L^2(A, \mathcal{A}_A, P_A)$ and $F_{X_A}(t) = P_A(X_A \le t) = \frac{1}{P(A)}P(\{X \le t\} \cap A)$, 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 $L^2(A, \mathcal{A}_A, P_A)$. Expectation on $(A, \mathcal{A}_A, P_A)$ will be denoted with $\mathbf{E}_A$. In order to evaluate the supremum in (5.14) we observe that any $\psi \in \mathcal{C}_A$ can be written, in the form

$$\psi = a\ \mathbf{1}_{A_0} + b\ \mathbf{1}_{A_1} \tag{5.16}$$

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

$$\langle X, \psi\rangle = (a-b)\langle X,\ \mathbf{1}_{A_0}\rangle + b\langle X,\ \mathbf{1}_A\rangle = b\ P(A)\left(\mathbf{E}_A(X_A) - \frac{1}{u}\langle X_A,\ \mathbf{1}_{A_0}\rangle_A\right). \tag{5.17}$$

Noticing that $b = \pm\sqrt{\frac{u}{P(A)\ (1-u)}}$, in order to calculate the supremum in (5.14) we define

$$\lambda_A(u) \equiv \sup_{\{A_0 : P_A(A_0)=u\}} \sqrt{\frac{P(A)\ u}{(1-u)}}\ \left(\mathbf{E}_A(X_A) - \frac{1}{u}\langle X_A, \mathbf{1}_{A_0}\rangle_A\right). \tag{5.18}$$

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 for a given random variable X,

$$\sup_{\psi\in\mathcal{C}_A} |\langle X, \psi\rangle| = \sup_{u\in(0,1)} \sup_{\psi\in\mathcal{C}_{A,u}} |\langle X, \psi\rangle| = \sup_{u\in(0,1)} \lambda_A(u). \tag{5.19}$$

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 2. *Assume $X \in L^2(\Omega, \mathcal{A}, P)$ and $A \in \mathcal{A}$. Then the fuction λ_A, in (5.18), is well defined for all $u \in (0,1)$ and in fact*

$$|\lambda_A(u)| \le \sqrt{P(A)}\|X\|_A. \tag{5.20}$$

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

$$\psi_u = -\sqrt{\frac{P(A)\ (1-u)}{u}}\mathbf{1}_{\{X_A \le F^{-1}_{X_A}(u)\}} + \sqrt{\frac{P(A)\ u}{1-u}}\mathbf{1}_{\{X_A \ge F^{-1}_{X_A}(u)\}} \tag{5.21}$$

satisfies

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

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 3. *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*

$$\lim_{u\to 0^+} \lambda_A(u) = \lim_{u\to 1^-} \lambda_A(u) = 0. \tag{5.22}$$

Proposition 3. *Assume $X \in L^2(\Omega, \mathcal{A}, P)$ and that the hypothesis in Lemma 3 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 (5.21) and verifies*

$$\langle X, \psi_A^m\rangle = \sup_{\psi\in\mathcal{C}_A} |\langle X, \psi\rangle|. \tag{5.23}$$

Proof. Since λ_A is continuous, by Lemma 3, 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

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

□

Finally, we have the following result for the greedy splitting algorithm.

Theorem 5. *Assume that each $X \in \mathcal{X}$ satisfies: $X \in L^\infty(\Omega, \mathcal{A}, P)$ and F_X is continuous. Then, the greedy splitting algorithm satisfies $\mathcal{A}_\infty = \sigma(X^1, \ldots, X^P) = \sigma(\mathcal{X})$.*

Under the same Hypothesis as in Theorem 5 we also obtain the following theorem.

Theorem 6. *The H-system constructed by the above greedy splitting algorithm is an unconditional basis for each $L^p(\Omega, \mathcal{A}_\infty, P)$ for $1 < p < \infty$.*

6. Numerical Examples

In this section we present output from a computer implementation based on Example 2 from Section 3. More specifically, we concentrate on the case where we have a Haar system, Definition 5, whose sequence of dyadic 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 4.1.
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 $\mathcal{A}_n$ are included in $\sigma(W_T)$ and $\mathcal{A}_\infty = \sigma(W_T)$. 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 m, as used in (5.12), m 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 m 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 (see Appendix A for a description of this algorithm and associated notation) for the H-system $\{u_{2^j+i}\}$ in Example 2 of Section 3 and X an European option. Computations can be carried out by specifying the finest scale J. We will then perform compression by only keeping the m 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 $\lim_{c\to\infty} X\mathbf{1}_{\{X\le c\}} = X$ 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 $m = 1,\ 2,\ 20$ 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 m.

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 m.

Tables 2 and 3 show the volume of transactions for the Haar hedging portfolio HΠ, and for the binary hedging portfolio BΠ (see Section 4.1), which for the case $n_T = 1$ are both constant quantities, and the volume of transactions for the Black-Scholes portfolio. Using the notation $X_{(m)}^c$ from (5.12), it is easy to show that the volume of transactions for the Haar hedging portfolio, is equal to

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

The volume of transactions for the portfolio of binary options is

$$VT(\mathrm{B}\Pi) = e^{-r(T-T_0)} ||X_{(m)}^c||_{L^1} = e^{-r(T-T_0)} \int_\Omega |X_{(m)}^c(\omega)| dP(\omega). \tag{6.4}$$

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 $\{t_i\}, i = 0, \ldots, m-1$ is

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

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 hedging. 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 reliance 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 m) 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 instance, 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.

Naturally, it is easy to 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 contingent claims. Key ingredients are the flexibility given by the possible space and time discretizations which can be adapted to a given class of options 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 other settings where this last technique is available. Some of the computational tools introduced can also be used even when an actual financial realization (of the approximation) is not available, pricing computations is an example. We have emphasized the issue of *efficient* representations of a given class

of options $\mathcal{X}$, 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 or computational advantage. The techniques could likely be extended to the setting of higher dimensional models.

Appendix A. Multiresolution Analysis Algorithm

First we introduce notation and algebraic relationships needed to set up computations in the multiresolution algorithm and in the rest of the paper.

Let $\mathcal{R} := \{\mathcal{R}_j\}_{j\geq 0}$ be a sequence of multiresolution partitions of Ω. 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 next aim is to study the relationship between the coefficients in this basis, which are proportional to samples at level j, with the coefficients in the H-system $\{\phi_{0,0}, \psi_{j,i}\}$ associated with $\mathcal{R}$ in Theorem 1.
Recalling the notation from (2.3), for each $j \geq 0$, $\{\phi_{k,i}\}_{(k,i)\in K_j}$ there is an orthonormal basis of the subspace 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 [19]. We have the simple, but relevant, relations:

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

and

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

where $a_{j,i}$ and $b_{j,i}$ were calculated in (2.5) and (2.6) respectively, and we have used the array notation

$$p_j[i] \equiv P(A_{j,i}). \tag{A.3}$$

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

$$X_j \equiv X_{\mathcal{R}_j} \equiv \mathbf{E}(X|\sigma(\mathcal{R}_j)). \tag{A.4}$$

Then we have the following expansions

$$X_j(\omega) = c_0[0]\ \phi_{0,0}(\omega) + \sum_{k=0}^{j-1} \sum_{i \in I_k} d_k[i]\ \psi_{k,i}(\omega) = \sum_{(k,i) \in K_j} c_k[i]\ \phi_{k,i}(\omega) \tag{A.5}$$

where

$$c_k[i] = \langle X_j, \phi_{k,i} \rangle \quad \text{and} \quad d_k[i] = \langle X_j, \psi_{k,i} \rangle. \tag{A.6}$$

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

$$c_k[i] = \langle X_j, \phi_{k,i} \rangle = \frac{1}{\sqrt{p_k[i]}} \int_{A_{k,i}} X_j dP = \frac{1}{\sqrt{p_k[i]}} \int_{A_{k,i}} X dP = \langle X, \phi_{k,i} \rangle. \tag{A.7}$$

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

Proposition 4. *Given $X \in L^2(\Omega, \mathcal{A}, P)$ and a sequence of multiresolution partitions $\mathcal{R} = \{\mathcal{R}_j\}_{j=0}^J$. Then for each $j' < j \leq J$, the following holds*

$$X_j = X_{j'} + \sum_{k=j'}^{j-1} \sum_{i \in I_k} d_k[i] \psi_{k,i} \tag{A.8}$$

and

$$\sum_{(k,i) \in K_j} c_k^2[i] = \sum_{(k,i) \in K_{j'}} c_k^2[i] + \sum_{k=j'}^{j-1} \sum_{i \in I_k} d_k^2[i]. \tag{A.9}$$

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 (A.8) follows inductively. Equation (A.9) is a direct consequence of (A.5) and (A.8). □

The precedent proposition, with the aid of (A.1) and (A.2), 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 [27] to our probabilistic setting. This algorithm produces a relation between the samples of X, namely,

$$x_k[i] = X_j(\omega),\ \omega \in A_{k,i}, \text{for } (k,i) \in K_j, \tag{A.10}$$

and the coefficients $d_k[i]$. This algorithm is described next.

We explain how to perform analysis, compression and synthesis, the recursive procedures are simple generalizations of the equal probability splitting Haar multiresolution analysis (see [24]) to our setting.

A.1. **Analysis.** Assume we are given the orthonormal basis associated to a indexed binary sequence of partitions. Given a finite sequence of binary partitions $\mathcal{Q} = \{\mathcal{Q}_j\}_{j=0}^n$, we can see it as a binary tree whose nodes are associated to a unique atom of $\cup_{j=0}^n \mathcal{Q}_j$, partially ordered by inclusion. Notice that the leafs of the tree are the non split atoms. We want to remark again that the indexation we introduced captures this structure, see Definition 3. Recall that if $\{\mathcal{R}_j\}_{j=0}^J$ is the indexed rearrangement of $\mathcal{Q}$, J is the maximum scale of its atoms, and $\mathcal{R}_J = \mathcal{Q}_n$. In this case we have all atoms of the same scale together, this is analogous to the wavelet analysis where coarser approximations are organized by scales. Of course if an atom does not split any more it will be in all subsequent partitions and in the tree appears as a terminal node at some level.

Assume we have an input signal $X : \Omega \to \mathbb{R}$, we will describe how to compute, using the notation introduced in (A.4) and (A.5), its expansion in the associated H-system

$$X_{\mathcal{R}_J}(\omega) = c_0[0]\ \phi_{0,0}(\omega) + \sum_{j=0}^{J-1} \sum_{i \in I_j} d_j[i]\ \psi_{j,i}(\omega) \tag{A.11}$$

These computations are called the *analysis* part of the algorithm, it involves computing all the inner products $d_j[i] = \langle X, \psi_{j,i} \rangle \equiv d_{A_{j,i}}$ and computing the values of $X_{\mathcal{R}_J}(\omega)$ on the atoms of $\mathcal{R}_J$. Notice $X_{\mathcal{R}_J}$ is a simple function constant on each of the atoms of $\mathcal{R}_J$. The inputs to the *analysis* formulas are the numbers

$$P(A) \text{ and } \frac{1}{P(A)} \int_A X(w) dP(w) = X_{\mathcal{R}_J}(\omega) \text{ for all } A \in \mathcal{R}_J, \omega \in A. \tag{A.12}$$

Therefore, our finest approximation is just the discretization, of $X(\omega)$, given by averaging X over the finest atoms, namely the elements of $\mathcal{R}_J$. As in classic Multiresolution analysis, we will compute the inner products $\langle X, \psi_{j-1,i} \rangle$ from $\langle X, \psi_{j,i} \rangle$. To define the formulas we need intermediate node variables, labelled `x`, `d` and `p`, their simple meaning is explained below. Here are the recursive formulas (bottom up recursion, i.e., we start at the leafs)

$$\mathtt{p_{father}} = \mathtt{p_{Lchildren}} + \mathtt{p_{Rchildren}}, \tag{A.13}$$

$$\mathtt{x_{father}} = \frac{1}{\mathtt{p_{father}}} (\mathtt{p_{Lchildren}}\ \mathtt{x_{Lchildren}} + \mathtt{p_{Rchildren}}\ \mathtt{x_{Rchildren}}) \tag{A.14}$$

$$\mathtt{d_{father}} = \sqrt{\frac{\mathtt{p_{Lchildren}}\ \mathtt{p_{Rchildren}}}{\mathtt{p_{father}}}}\ (\mathtt{x_{Lchildren}} - \mathtt{x_{Rchildren}}). \tag{A.15}$$

To be able to run this recursive formula we just need to initialize the `x` and `p` variables at the terminal nodes, this is done by using the inputs,

$$\mathtt{p_{leaf}} = \mathtt{P(A)}, \text{ where the atom } A \in \mathcal{R}_J \text{ is associated to the given } \textit{leaf}, \tag{A.16}$$

$$\mathtt{x_{leaf}} = \mathtt{X}_{\mathcal{R}_J}(\omega), \text{ where } \omega \in A \in \mathcal{R}_J \text{ is the atom associated to the given } \textit{leaf}. \tag{A.17}$$

It is easy to see that the meaning of the intermediate variable x at node $A \in \cup_{j=0}^{J} \mathcal{R}_j$ is

$$x_A = \frac{1}{P(A)} \int_A X(w) dP(w) = \sqrt{P(A)}\, c_A, \tag{A.18}$$

where $c_{A_{k,i}} \equiv c_k[i]$. The meaning of the variable d at node $A \in \cup_{k=0,\ldots,n-1} \mathcal{Q}_k$ is

$$d_A = \langle X, \psi_A \rangle. \tag{A.19}$$

Also

$$p_{nodeA} = P(A). \tag{A.20}$$

The above recursion gives $\langle X, \psi_{0,0} \rangle = \frac{1}{P(\Omega)} \int_\Omega X(w) dP(w) = \int_\Omega X(w) dP(w) = x_{root}$.

Therefore, analysis gives us the inner products and, if we are interested, it gives coarser approximations ("zoom-outs") to X given by the values of the x intermediate variables. To be more precise at each level j, $0 \le j \le J$, of the tree we have the approximations

$$X_{\mathcal{R}_j}(\omega) = x_A = \frac{1}{P(A)} \int_A X(w) dP(w) \text{ where } \omega \in A \in \mathcal{R}_j, \tag{A.21}$$

so, the simple function $X_{\mathcal{R}_j}$ is the coarser approximation associated to the partition $\mathcal{R}_j$, completely analogous to the wavelet approximations at different scales.

Transform compression: notice that the orthonormal basis associated to a sequence of dyadic partitions is coded in a binomial tree. Naturally the analysis (inner products and coarser approximations) of an input X is also coded in another binomial tree, which we will call the *analysis tree* (of X). Transform compression is just the process of setting the smallest inner products in the analysis tree to zero. So, we sort them and only keep a percentage of the largest ones.

Synthesis: The input to synthesis is the analysis tree of X containing the $\mathtt{d}_A$ values (which are equal to the inner product at node A or zero) where A is a non-terminal node. Also the $\mathtt{p}_A$ (i.e. the probability at node A) values are needed for all nodes A. We also need as input $\mathtt{x}_{A_{0,0}}$; the output will be $\mathtt{x}_A$ for all nodes A (including the leafs). Here are the reconstruction formulas (top-down recursion, i.e. we start at the top)

$$\mathtt{x}_{\mathtt{Lchildren}} = \mathtt{x}_{\mathtt{father}} + \sqrt{\frac{\mathtt{p}_{\mathtt{Rchildren}}}{\mathtt{p}_{\mathtt{father}}\ \mathtt{p}_{\mathtt{Lchildren}}}}\ \mathtt{d}_{\mathtt{father}} \tag{A.22}$$

$$\mathtt{x}_{\mathtt{Rchildren}} = \mathtt{x}_{\mathtt{Lchidren}} - \sqrt{\frac{\mathtt{p}_{\mathtt{father}}}{\mathtt{p}_{\mathtt{Lchidren}}\ \mathtt{p}_{\mathtt{Rchildren}}}}\ \mathtt{d}_{\mathtt{father}}. \tag{A.23}$$

Appendix B. Complements

Example of non Haar-System for the binomial model:
Here we present 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 $i \neq 0$,

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

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

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

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

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

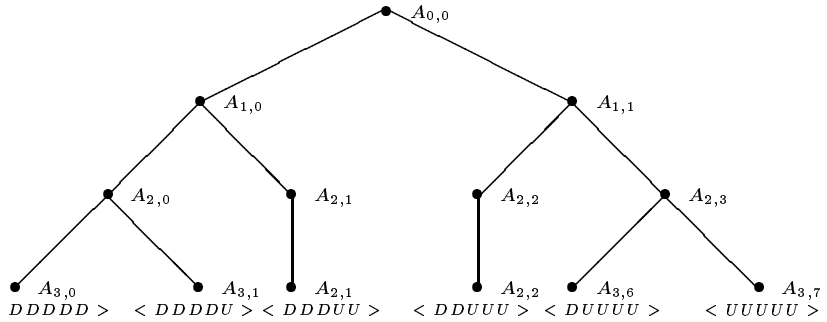
and

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

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, namely by

$$\begin{aligned} &v_0 \equiv 1, \\ &v_{j,i} = \frac{1}{\sqrt{P(A_{j,i})}}\left(\sqrt{\frac{P(A_{j+1,2i+1})}{P(A_{j+1,2i})}} 1_{A_{j+1,2i}} - \sqrt{\frac{P(A_{j+1,2i})}{P(A_{j+1,2i+1})}} 1_{A_{j+1,2i+1}}\right) \end{aligned} \tag{B.3}$$

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 $A_{3,3} = A_{3,5} = \emptyset$.

Appendix C. Figures and Tables

Table 1. L^2 norm for errors, between X and $X^c_{(J)}$, 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

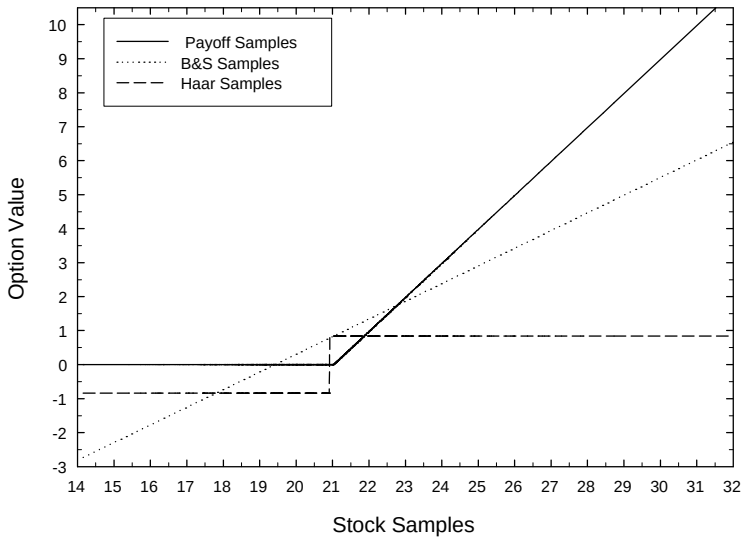


FIGURE 1. Approximations to single European Call using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $m = 1$, $S_{T_0} = 20$, $r = 0.05$, $\sigma = 0.1$, $T - T_0 = 1$, $K = 21$.

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`
$R = 5$	0.78	0.88	53.32
$R = 10$	0.79	0.93	107.1
$R = 15$	0.78	0.96	157.2
$R = 20$	0.75	0.91	213.8
$R = 25$	0.71	0.91	258.8
$R = 30$	0.74	0.93	317

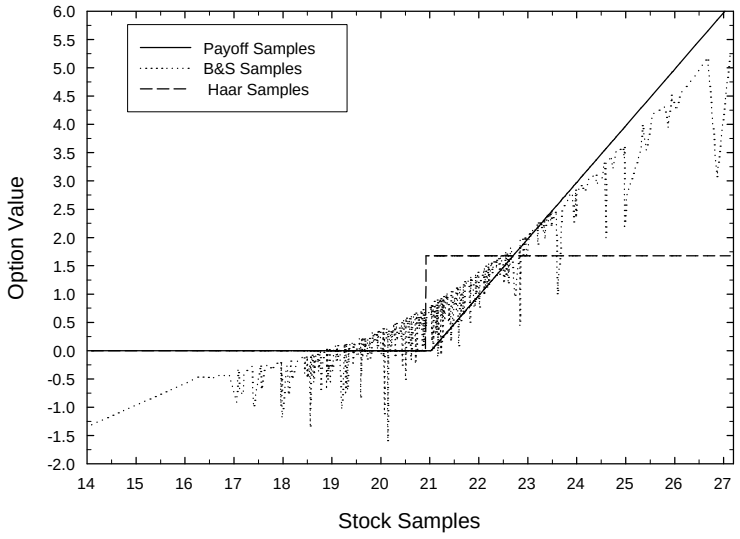


FIGURE 2. Same as in Figure 1 except $m = 2$.

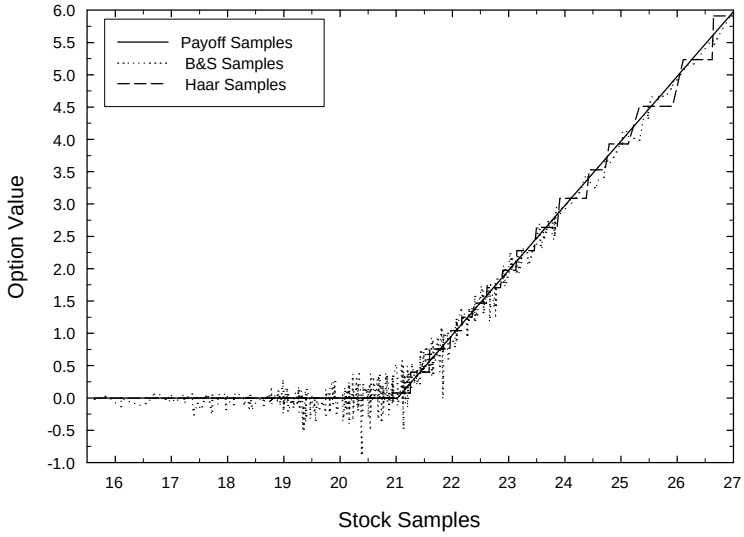


FIGURE 3. Same as in Figure 1 except $m = 20$.

REFERENCES

[1] M. Avellaneda and A. Paras (1996), Managing the volatility risk of portfolios of derivative securities: the Lagrangian uncertain volatility model. *Applied Mathematical Finance*, 3, pp. 21-52.

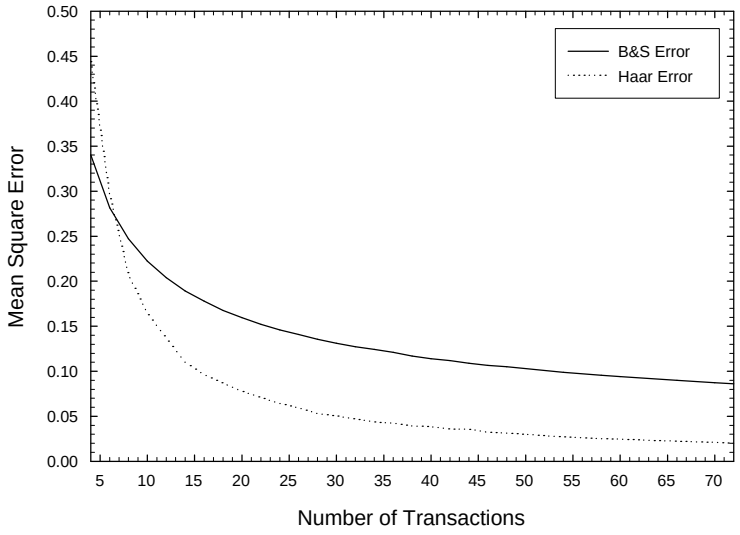


FIGURE 4. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter m. Values of the parameters used: as in Figure 1.

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`
$R = 5$	2.3	0.6	27.1
$R = 10$	2.31	0.65	59.43
$R = 15$	2.29	0.66	90.75
$R = 20$	2.29	0.67	121.59
$R = 25$	2.23	0.63	152.03
$R = 30$	2.25	0.65	184.9

[2] N. H. Bingham and R. Kiesel (1998), *Risk-Neutral Valuation. Pricing and Hedging of Financial Derivatives.* Springer-Verlag, London.
[3] J. Cai and S. E. Ferrando (March 2004), C++ Implementation of Haar Systems for Black-Scholes Model. *Technical Report, Ryerson University.*
[4] 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.
[5] D. Dadydov and V. Linetsky (March-April 2003), Pricing options on scalar diffusions: an eigenfunction expansion approach. *Operations Research*, Vol. 51, No 2, 185-209.
[6] Emanuel Derman, Deniz Ergener and Iraj Kani (Summer 1995), Static options replication. *The Journal of Derivatives*, 78-95.
[7] Ronald A. DeVore (Summer 1998), Nolinear approximation. *Acta Numerica*, 1-99.

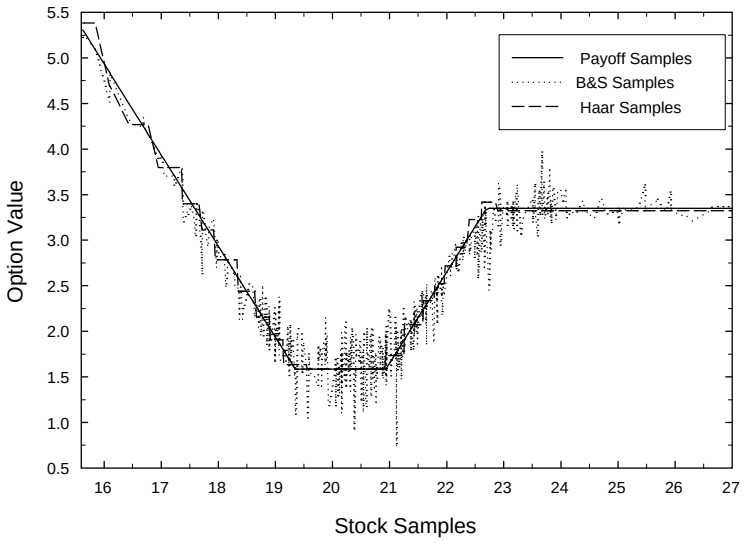


FIGURE 5. Approximations to portfolio X, constructed from two calls and one put, using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $m = 20$, $S_{T_0} = 20$, $r = 0.05$, $\sigma = 0.1$, $T - T_0 = 1$, $K_1 = 19$, $K_2 = 21$ $K_3 = 23$.

[8] Ronald A. DeVore, Guergana Petrova and Vladimir Temlyakov (May 2003), Best basis selection for approximation in L_p *Foundations of Computational Mathematics*, **3**, no. 2., pp. 161 - 185.

[9] 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.

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

[11] 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.

[12] D.E. Dutkay and P. E.T. Jorgensen (July 2004), Martingales, endomorphisms, and covariant systems of operators in Hilbert space. Preprint `arXiv:math.CA/ 0407331`.

[13] G. Fishman (1996), Monte Carlo. Springer Verlag.

[14] R. Gençay, F. Selçuk and B. Whitcher (2002), *An Introduction to Wavelets and other Filtering Methods in Finance and Economics.* Academic Press, San Diego.

[15] M. Girardi and W. Sweldens (1997), A new class of unbalanced Haar wavelets that form an unconditional basis basis for L^p on general measure spaces. *The Journal of Fourier Analysis and Applications*, **3**, 457-474.

[16] Richard F. Gundy (1966), Martingale theory and pointwise convergence of certain orthogonal systems. *Trans. Amer. Math. Soc.*, **124**, 228-248.

[17] Richard F. Gundy (2000), Low-Pass Filters, Martingales, and Multiresolution analysis. *App. Comp. Har. Anal.*, **9**, 204-219.

[18] Richard F. Gundy (2002), Wavelets and Probability. *Lecture Notes presented at workshop "Wavelets and Applications", Barcelona*, July 1-6, 2002.

[19] E. Hernandez and G. Weiss (1996):*A First Course on Wavelets.* CRC Press, Boca Raton Fl.

[20] C. C. Heyde (1977), On central limit and iterated logarithm supplements to the martingale convergence theorem. *J. Appl. Prob.* **14**, 758-775.

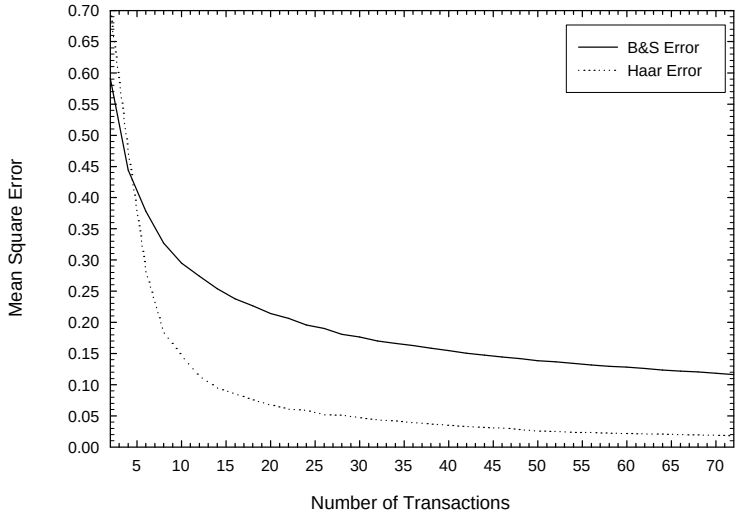


FIGURE 6. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter m. Values of the parameters used: as in Figure 5.

[21] P. Hall and C. C. Heyde (1980): *Martingale Limit Theory and Application.* Academic Press, New York.
[22] D.G. Hobson (1998), Robust hedging of the lookback option. *Finance and Stochastics*, 2, pp. 329-347.
[23] P. Jaeckel (2002), Monte Carlo Methods in Finance. John Wiley and Sons.
[24] A. Jensen and A. la Cour-Harbo (2001): *Ripples in Mathematics. The Discrete Wavelet Transform.* Springer.
[25] E. Lieb and M. Loss (1997): *Analysis.* Graduate Studies in Mathematics 14 American Mathematical Society.
[26] 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.
[27] S. Mallat (1989), Multiresolution aproximations and wavelets orthonormal bases for $L^2(\mathbb{R})$. *Trans. of Amer. Math. Soc.*, **315**, 69-87.
[28] A. Matache, T. Von Petersdoff and C. Schwab (2004), Fast deterministic pricing of options on Levy driven assets. *M2AN Math. Model. Numer. Anal.* **38**, 1, 37-71.
[29] A. Matache, C. Shwab and T. Wihler (2004), Fast numerical solution of parabolic integro differential equations with applications in finance. Preprint 1954, IMA, University of Minnesota.
[30] Y. Meyer, Personal communication *3rd. European Congress of Mathematics held in Barcelona July 10-15, 2000.*
[31] J. Neveu (1975): *Discrete-Parameter Martingales*, North-Holland.
[32] A. Pechtl (1995), Classified information. *RISK*, **8 (6)**, 59-61.
[33] D. Percival and A. Walden (2000): *Wavelets Methods for Time Series Analysis*, Cambridge University Press, Cambridge.
[34] W. Willinger (1987), *Pathwise Stochastic Integration and Almost-Sure Aproximation of Stochastic Processes.*, Thesis, Cornell University.

Departamento de Matemática. Universidade Estadual de Campinas, 13.081-97-Campinas, SP, Brazil.
E-mail address: pedrojc@ime.unicamp.br

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