EFFICIENT HEDGING USING A DYNAMIC PORTFOLIO OF BINARY OPTIONS

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

Abstract. We present a discretization of financial instruments in terms of martingale expansions constructed using Haar wavelets systems. Expansions on these bases give the pointwise convergence needed in several applications. We work out the details of an application to hedging an European portfolio of options and describe natural conditions under which our Haar hedging strategy can be realized by means of a self financing portfolio consisting of binary options. The efficiency of the hedging is studied by analyzing the volume of transactions required to construct the approximating portfolio and by providing numerical comparisons with delta hedging.

Key Words: Heding; Portfolio of Options; Haar Wavelets; Martingales.

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 value 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. The area of nonlinear approximation (see [7] and [10] and the references given there) studies efficient representations of functional classes. For specific functional classes, wavelets have been proven optimal for the task of compression (efficient storage), noise removal, fast computation, etc.

The use of wavelets techniques in finance has been directed towards time series processing, (see for instance [13] and [25]) and the fast numerical solution, via the Galerkin method, of Black-Scholes equation (see [21] and [22] 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 $\mathcal{F} = \{\mathcal{F}_t\}$ is the filtration

Date: July 15, 2009.

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.

generated by a given stochastic process and P a probability measure of interest. It is possible, and necessary for some of the applications, that the basis functions $\{u_k\}$ 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.

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 [18] describes possible uses of a Hilbert space basis for valuation and hedging. Our contribution is different, the paper introduces a framework that allows the construction of Hilbert space bases with the aim to provide efficient approximations. The efficiency is achieved, in part, thanks to the adaptability (in the sense of measure theory) with respect to the underlying process and the localization of the basis functions. We show how to use our approximations to construct portfolios of binary options for hedging general financial claims and provide bounds for the volume of transactions required to implement the approximating portfolios. These upper bounds rely on the property of localization of our basis functions and are presented in Section 5. Some of these characteristics are in contrast with other approaches to hedging ([1], [6] and [16]) which, similarly to our approach, use portfolios of simple options to hedge complex portfolios.
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 $\langle X, u_k \rangle$ 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). Most importantly, the setting is flexible enough so that the actual Haar functions $\{u_k\}$ can be chosen via an optimization as proposed in [4] in order to give efficient representations of X. Efficient representations of functional classes is a chief concern of computational harmonic analysis, see for example [10], [8] and [9].

The rest of the paper is organized as follows, Section 2 defines H-systems and develops some of the relationships between H-systems and sequences of partitions. Section 3 summarizes Willinger's main result on existence of atomic discretization of stochastic processes and connects them to our setup of H-systems. Section 3.1 introduces a useful example of an H-system in a basic financial setting. Section 4 motivates and develops our main application to hedging a given European portfolio of options, it also outlines other applications. Section 5 presents several upper bounds for the volume of transactions required to implement the approximating portfolios described in Section 4. Section 6 presents numerical examples and comparissons with delta hedging. Section 7 summarizes the main results of the paper. Appendix A presents notation and formulae required for a multiresolution analysis algorithm. Appendix B presents a simple example as a complement. Appendix C presents tables and figures from the numerical experiments. As a technical note, and for matters of convenience, we will supress writing a.e. (almost everywhere) from many statements.

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 [15] definition is motivated by the standard Haar system on $L^2([0,1])$.

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 $X \in L^2(\Omega, \mathcal{A}, P)$*

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

where $\mathcal{A}_m = \sigma(u_0, \ldots, u_m)$. 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 $\mathcal{A}_\infty = \sigma(\cup_{m\geq 0}\mathcal{A}_m)$. In applications we will make use of the pointwise convergence of (2.1) which holds due to the martingale convergence theorem [23]. Moreover, if $p \in [1, \infty)$ is a given real number then, for every $X \in L^p$, the sequence $X_{\mathcal{A}_m} = \mathbf{E}(X|\mathcal{A}_m)$ converges a.s. and in L^p to $X_\infty = \mathbf{E}(X|\mathcal{A}_\infty)$. Convergence to X holds whenever $\sigma(X) \subseteq \mathcal{A}_\infty$.

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 [15], 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 $\mathcal{A}_m$ consists exactly of $m+1$ 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. We also introduce some of the notation to be used in the rest of the paper. The reader who wishes to see financial applications first should refer to Section 4.

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

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

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

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 figure displayed in Appendix B should clarify the indexation.

Given a binary sequence of partitions $\mathcal{Q}_k$, $k \geq 0$, define the associated trees

$$\mathcal{T}_n \equiv \cup_{k=0}^{n} \mathcal{Q}_k, \text{ and } \mathcal{T} \equiv \cup_{n \geq 0} \mathcal{T}_n.$$

For every internal node $A \in \mathcal{T}$ we have, using the indexation introduced in Definition 3, $A = A_{k,i}$ and a corresponding Haar function at that node

$$\psi_A = a_{k,i} \mathbf{1}_{A_{k+1},2i} + b_{k,i} \mathbf{1}_{A_{k+1},2i+1}. \tag{2.2}$$

Given $A \in \mathcal{T}$ we have the natural associated tree $\mathcal{T}_A$.

The following definition refines Definition 3, it constructs partitions by collecting atoms with the same scale parameter j. Atoms at lower levels, which complete a partition and will not be further split, are also included.

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 can associate a Multiresolution Analysis algorithm (MRA) (see Appendix A) in complete analogy with wavelet theory which, 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 throughout the paper 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 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 $j < J - 1$).*

The proof of the following result is provided in [4].

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

Indeed, from [4], a multiresolution sequence of partitions$\{\mathcal{R}_j\}_{j\geq 1}$ has associated the H-system given by $u_0 = \phi_{0,0} \equiv \mathbf{1}_\Omega$ and$\{u_{2^j+i} \equiv \psi_{A_{j,i}} : j \geq 1, i \in I_j\}$.

Remark 1. *We will use H-systems to approximate stochastic processes, in fact for every partition of the time interval* $[0, T]$ *we will have a finite H-system. Details are provided in Section 3.*

Section 3 briefly describes general results on atomic approximations of stochastic processes. These results will guarantee the availability of H-systems for further developments in the paper. As indicated, optimized constructions of H-systems are also possible, an approach is fully developed in [4].

3. Atomic Discretizations of Stochastic Processes

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 [4], reference [14] 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 [28] and [29], 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. *$I^\xi = \{0 = t(\xi, 0) < ... < t(\xi, N_\xi) = T\}$, where $N_\xi < \infty$.*
2. *For each t, $\mathcal{F}^\xi_t$ is a finitely generated sub σ-algebra of $\mathcal{F}_t$, with atomic partition $\mathcal{P}^\xi_t$.*
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. The sequence $I^{(n)}$ of index satisfies:*

$$\lim_{n\to\infty} |I^{(n)}| = 0 \tag{3.1}$$

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 $[0, T]$, For each $0 \leq t \leq T$,

$$\mathcal{F}^{(n)}_t \uparrow \mathcal{F}_t, \tag{3.2}$$

and

$$P(\{\omega \in \Omega : \lim_{n\to\infty} \sup_{0\leq t\leq T} |S_t(\omega) - \xi^{(n)}_t(\omega)| = 0\}) = 1. \tag{3.3}$$

The fundamental result of W. Willinger ([28] pp 55, 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

$$\mathcal{R}^{\xi}_{j_m} = \mathcal{P}^{\xi}_{t_m} \text{ for } 0 = j_0 < j_1 < ... < j_N. \tag{3.4}$$

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 $\{\mathcal{R}^{\xi}_j\}_{j\geq 0}$. Clearly, 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 indexes $(I^{(n)} = \{0 = t^n_0 < ... < t^n_{N_n} = T\})$ 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. Using Theorem 1 construct for each n an H-system $(\mathcal{H}^{(n)} = \{\psi^n_{j,i}\})$ associated to the sequence of partitions $\{\mathcal{R}^{(n)}_j\}_{j\geq 0}$ in (3.4). 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 $t \in [t^n_m, t^n_{m+1})$. □

To be clear, we briefly remark on the connection between Proposition 2 and (2.1). The functions u_k appearing in (2.1) are the functions $\psi^n_{j,i}$ introduced in Proposition 2, the key remark being the extra parameter n which corresponds to the time interval discretizations $I^{(n)}$. Using notation introduced in Definition 1, we actually have $\mathcal{A}_n \equiv \sigma(A : A \in \mathcal{P}^{\xi^{(n)}}_{t_m}, m = 0, \ldots, N^{(n)})$ and Lemma 1 proves that $\mathcal{F} = \mathcal{A}_\infty$. This guarantees the pointwise convergence of our approximations (given by (2.1)) for any X satisfying $\sigma(X) \subseteq \mathcal{F} \equiv$ to the completion of $\sigma(S_t : t \in [0, T])$.

Willinger's results assume stochastic processes with continuous paths. In principle, our developments in the paper do not require completeness of the market model used, results from [27] could be used to extend our approach to processes with jumps.

3.1. **Example. An H-System in the Black-Sholes model.** We describe a simple example of an H-system in a familiar financial context. The example is a Haar system, namely, it is generated by dyadic partitions, see Definition 5. Appendix B, illustrates an H-system associated to a sequence of binary partitions that is not dyadic and illustrates the case of *multiresolution* sequence of partitions.

The 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 $0 \leq i \leq 2^j - 1$,
(3.5)

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

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

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 (3.5), 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

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

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. Application to Hedging

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 work 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 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 [16]. 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 [24].

Let us explain the basic idea in this section, the simple functions u_k, 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 $\mathcal{A}_m$ is generated by $u_0, \ldots, u_m$ and contains $m+1$ 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.

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. Section 5 provides the definitions of $VT(\Pi)$ for the approximating portfolios Π put forward in this paper.

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

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

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

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

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 Section 5. It is also recognized in signal processing applications that localization of wavelets is a key property to represent discontinuities efficiently [8]; this insight is reflected in our Proposition 5, see also the explanations at the end of Section 5.

4.2. **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 exists 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

$$\mathbf{B}_{j,i} = (\mathbf{1}_{A_{j,i}}(t) \equiv \mathbf{1}_{[t_{s+1}, T]}(t) \mathbf{1}_{A_{j,i}}), \quad j_s \leq j \leq j_{s+1}, \tag{4.0}$$

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 obligation $\Psi_{j,i}$ is realized in terms of the binary options $\mathbf{B}_{j+1,k}$, $k = 2i, 2i + 1$.

Next we will define two hedging strategies via self-financing portfolios, of static and dynamic types, to replicate an European option using H-systems. One of the strategies, denoted by HΠ, is associated to Haar obligations and the other strategy, denoted by BΠ, is associated directly to binary options. The constructions require the availability of $\mathcal{R} = \{\mathcal{R}_j\}$, a multiresolution sequence of partitions, localized in the sequence of times $t_0 = T_0 < \ldots < t_n = T$, and $X \in L^2(\Omega, \sigma(\cup_{j\geq 0}\mathcal{R}_j), P)$. Notation from Appendix A will be used.

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 $t \in [t_0, t_1)$

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

$$e^{-r(T-t_1)}\ d_j[i] \text{ invested in } \Psi_{j,i}\ j = 0, \ldots j_1 - 1, i \in I_j, \tag{4.1}$$

where the coefficients $d_j[i]$ are given by (A.3).
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.7)). 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, $k_0 = j_{s-1}$ (see the comment after Definition 4), the coordinates of $\mathrm{H}\Pi_{\mathcal{R}}(X)_t$ for $t \in [t_{s-1}, t_s)$ are

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

$$e^{-r(T-t_s)}\ d_j[i] \text{ invested in } \Psi_{j,i}\ j = j_{s-1}, \ldots j_s - 1, i \in I_j^{i_0}, \tag{4.2}$$

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

$$e^{-r(T-t_{s-1})}x_{k_0}[i_0], \tag{4.3}$$

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 (4.1) when $t \in [t_0, t_1)$. It is clear from (A.2) 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 (4.2), for all $t \in [t_{s-1}, t_s)$, if A_{k_0,i_0} splits at the next level or by (4.3) 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 (4.2) or (4.3) is $e^{-r(T-t_{s-1})}x_{k_0}[i_0]$. Consider again case I), and $t = t_s$, by (A.5) and (4.2) 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

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

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

□

Characteristic Functions Portfolio. 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.7).
We formalize $\mathrm{B\Pi}_{\mathcal{R}}(X)_t$ as a vector valued process which is constant on the intervals $t_{s-1} \le t < t_s$. 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,

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

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 $t_{s-1} \le t < t_s$, we need to specify the coordinates of $\mathrm{B\Pi}_{\mathcal{R}}(X)_t$, namely,

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

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

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

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 to the example in Section 3.1 and to the example from Appendix B.

4.3. **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 ($\mathcal{F}^n_{t_m^n}$) such that

(1) $I^{(n)} \subset I^{(n+1)}$ and $\cup_n I^{(n)}$ is dense in $[0, T]$,
(2) $\mathcal{F}^n_{t_m^n} \subset \mathcal{F}^{n+1}_{t_m^n} \subset \mathcal{F}_{t_m^n}$,
(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 $\xi_t^{(n)} = \mathbf{E}(S_{t_m^n} \mid \mathcal{F}^n_{t_m^n})$ for $t \in [t_m^n, t_{m+1}^n)$.

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

$$U^n_{j+1} = \max(Z^n_j, e^{r(t^n_{j+1} - t^n_j})\mathbf{E}(U^n_{j+1} | \mathcal{F}^n_{t_j^n})).$$

The numerical problem is to calculate the conditional expectation $\mathbf{E}(U_{j+1} | \mathcal{F}^n_{t_j})$. 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.

4.4. **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 [11] and [17]. By Proposition 2 we know there exists a sequence of finite H-systems ($\mathcal{H}^{(n)} = \{\psi^n_{j,i}\}$) and two sequences of finite indexes ($I^{(n)} = \{0 = t^n_0 < ... < t^n_{N_n} = T\}$) and ($J^{(n)} = \{0 = j^n_0 < ... < j^n_{M_n}\}$) such that

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

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

$$\mathbf{E}(X|\mathcal{F}^n_t) = \sum_{t^n_j \leq t} \langle X, \psi^n_{j,i}\rangle \psi^n_{j,i}.$$

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

$$a^n_{MC}(j, i, M) = \frac{1}{M} \sum_{m=1}^{M} X(w^m) \psi^n_{j,i}(w^m),$$

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

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

5. Volume of Transactions

Here we describe upper bounds, in the mean and pointwise, for the volume of transactions. We believe the quality of these upper bounds provide evidence of the efficiency of our approximations as far as volume of transactions are concerned. A discussion concerning the meaning as well as the quality of our upper bounds is provided at the end of the present Section.

We will assume the existence of a skeleton approximation (and, hence, continuity of S_t will be assumed) as described in Section 3, this assumption is to make sure we have available an H-system that approximates the process. The results provide upper bounds, in the mean and pointwise, for the volume of transactions required to implement the portfolios $B\Pi$ and $H\Pi$ (introduced in Section 4). They rely on the following structural aspects of our approximations: they approximate the relevant conditional expectations, the martingale property is used in the self financing nature of the portfolios and the orthogonality of the martingale differences, and the localized property of the basis functions is used throughout (as in (5.10)).

Fix an arbitrary sequence of times t_s, $s = 0, \ldots, N$ and let $\mathcal{F}_{t_s}^n$, n being variable, represent the discrete skeleton atomic sigma algebras approximating $\mathcal{F}_{t_s}$ (see Definition 7). Notice that the parameter N was implicit in the skeleton approximations from Section 3, here N will be fixed and n can increase to ∞ providing the pointwise convergence $\mathbf{E}(.|\mathcal{F}_{t_s}^n) \to \mathbf{E}(.|\mathcal{F}_{t_s})$ for all $s = 1, \ldots, N$. Fixing N apriori in such a way assumes the time discretization is fine enough in order to provide a good approximation to $\mathbf{E}(X|\mathcal{F}_t)$ at all times $t \in [0,T]$. This is possible given that we are assuming S_t has continuous paths and so (3.3) is available.

We also assume t_0 is the initial time and $\mathcal{F}_{t_0}$ the trivial sigma algebra and $t_N = T$- the final time. We will use $\{A_{l,t_s}\}$, $l = 0, \ldots, L_{t_s}^n$, to denote the collection of atoms generating $\mathcal{F}_{t_s}^n$. We will take $L_{t_s}^n = 2^{j_s^n} - 1$, where we have used the notation from Proposition 2, notice that $L_{t_0}^n = 0$ and $A_{0,t_0} = \Omega$.

In order to simplify the notation and developments, we assume that the given sets A_{l,t_s} split in the next level.

Definition 13. *Consider the the binary options* $\mathbf{B}_{k,t_s} = (\mathbf{1}_{A_{k,t_s}}(t) \equiv \mathbf{1}_{[t_{s-1},T]}(t)\mathbf{1}_{A_{k,t_s}})$, $k = 0, \ldots, L_{t_s}^n$ *as described in (4.0). The volume of transactions at time* t_{s-1}, *denoted* $VT_{t_{s-1}}$, *necessary to implement a portfolio* Π *made up as a linear combination* $\Pi = \sum a_{k,t_s}\mathbf{B}_{k,t_s}$ *of such binary options, is given by*

(5.1)
$$VT_{t_{s-1}}(\Pi)(w) = \sum |a_{k,t_s}|\, VT_{t_{s-1}}(\mathbf{B}_{k,t_s}) \equiv \sum |a_{k,t_s}|\, e^{-r(t_s - t_{s-1})}\, \mathbf{E}(\mathbf{1}_{A_{k,t_s}}|\mathcal{F}_{t_{s-1}})(w).$$

Note that, as defined above, the volume of transactions for a random variable Π depends on the representation, i.e. using different binary options, providing an equivalent representation of Π, will result on a different value for the volume of transactions.

Lemma 2. *Given the binary option* $\mathbf{B}_{k,t_s} = \mathbf{1}_{A_{k,t_s}}$*, with* $A_{k,t_s} \in \mathcal{F}^n_{t_s}$ *and* $A_{k,t_s} \subseteq A_{l,t_{s-1}} \in \mathcal{F}^n_{t_{s-1}}$*, then*

$$\mathbf{E}(\mathbf{1}_{A_{k,t_s}} | \mathcal{F}^n_{t_{s-1}}) = \frac{P(A_{k,t_s})}{P(A_{l,t_{s-1}})}\ \mathbf{1}_{A_{l,t_{s-1}}} \tag{5.2}$$

and so,

$$\mathbf{E}\left(VT_{t_{s-1}}(\mathbf{B}_{k,t_s}) - e^{-r(t_s-t_{s-1})}\mathbf{E}(\mathbf{1}_{A_{k,t_s}} | \mathcal{F}^n_{t_{s-1}})\right) = 0. \tag{5.3}$$

Proof. The proof follows by noticing that $A_{l,t_{s-1}}$ is an atom of the atomic sigma algebra $\mathcal{F}^n_{t_{s-1}}$. □

We will analyze the volume of transactions needed to implement the binary portfolio ($B\Pi$) and the Haar portfolio ($H\Pi$) introduced in Section 4.2. The notation will be designed, as much as possible, so the developments cover both cases simultaneously. When stating results that apply to both portfolios we will designate either of them by Π.

Define

$$D_{t_{s-1},t_s} \equiv \mathbf{E}(X|\mathcal{F}^n_{t_s}) - \mathbf{E}(X|\mathcal{F}^n_{t_{s-1}}),\ s = 2, \ldots, N,$$

and

$$D_{t_0,t_1} \equiv \mathbf{E}(X|\mathcal{F}^n_{t_1}) - \mathbf{E}(X|\mathcal{F}^n_{t_0}) \text{ for the } H\Pi \text{ portfolio and}$$
$$D_{t_0,t_1} \equiv \mathbf{E}(X|\mathcal{F}^n_{t_1}) \text{ for the } B\Pi \text{ portfolio.}$$

Lemma 3. *The following expression provides the volume of transactions required for a financial implementation of the portfolio(s)* Π*.*

$$VT_{t_0}(\Pi)(w) = \sum_{s=1}^{N} \sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w)\ VT_{t_{s-1}}(D_{t_{s-1},t_s})(w). \tag{5.4}$$

Proof. To establish (5.4) for the $H\Pi$ portfolio, we argue by induction. Recall that $A_{t_0} = \Omega$, given that $e^{-r(T-t_0)}\mathbf{E}(X)$ is invested in the bond; at time t_1 the Haar portfolio provides the approximation $\mathbf{E}(X|\mathcal{F}^n_{t_1})$. Therefore, only the difference D_{t_0,t_1} requires a financial realization using binary options; the associated volume of transactions is given by $VT_{t_0}(D_{t_0,t_1})(w)$. Reasoning inductively, at time t_{s-1}, the value of the Haar hedging strategy is $\mathbf{E}(X|\mathcal{F}^n_{t_{s-1}})$. Therefore, in order to achieve the approximation $\mathbf{E}(X|\mathcal{F}^n_{t_s})$ at time t_s, only the difference D_{t_{s-1},t_s} requires a financial realization using binary options; the associated volume of transactions is given by $VT_{t_{s-1}}(D_{t_{s-1},t_s})(w)$. This completes the inductive argument and gives (5.4). An analogous reasoning provides the argument for the case of the $B\Pi$ portfolio. □

In order to obtain pointwise upper bounds be will need to replace $\mathbf{E}(\mathbf{1}_{A_{k,t_s}}|\mathcal{F}_{t_{s-1}})(w)$ by $\mathbf{E}(\mathbf{1}_{A_{k,t_s}}|\mathcal{F}^n_{t_{s-1}})(w)$ in the expressions for $VT_{t_{s-1}}(D_{t_{s-1},t_s})(w)$. The resulting expressions for $VT_{t_0}(\Pi)$ and $VT_{t_{s-1}}(D_{t_{s-1},t_s})$ will be denoted by $\hat{VT}_{t_0}(\Pi)$ and $\hat{VT}_{t_{s-1}}(D_{t_{s-1},t_s})$, these values will be called *approximate* values. Notice that we have pointwise convergence

$$\lim_{n\to\infty} \hat{VT}_{t_{s-1}}(D_{t_{s-1},t_s}) = VT_{t_{s-1}}(D_{t_{s-1},t_s}) \tag{5.5}$$

for all $s = 1, \ldots, N$. Similarly for $\hat{VT}_{t_0}(\Pi)$ and $VT_{t_0}(\Pi)$. Therefore we can concentrate in establishing pointwise error bounds for the approximated quantities.

Equation (5.3) will be used in several instances to justify the equality $\mathbf{E}(VT(D_{t_{s-1},t_s})) = \mathbf{E}(\hat{VT}(D_{t_{s-1},t_s}))$, from which $\mathbf{E}(VT_{t_0}(H\Pi)) = \mathbf{E}(\hat{VT}_{t_0}(H\Pi))$ follows.

5.1. **Characteristic Functions Representations.** Introducing the orthonormal system

$$\phi_{k,t_s} \equiv \frac{1}{\sqrt{P(A_{k,t_s})}}\ \mathbf{1}_{A_{k,t_s}}, \tag{5.6}$$

we can write

$$D_{t_{s-1},t_s}(w) = \sum_{l=0}^{L^n_{t_{s-1}}} \sum_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} \langle D_{t_{s-1},t_s}, \phi_{k,t_s}\rangle\ \phi_{k,t_s}. \tag{5.7}$$

From (5.6), (5.7) and (5.1) we obtain the following Lemma.

Lemma 4. *For* $w \in A_{l,t_{s-1}}$

$$VT_{t_{s-1}}(D_{t_{s-1},t_s})(w) = e^{-r(t_s - t_{s-1})} \sum_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} \frac{|\langle D_{t_{s-1},t_s}, \phi_{k,t_s}\rangle|}{\sqrt{P(A_{k,t_s})}}\ \mathbf{E}(\mathbf{1}_{A_{k,t_s}} | \mathcal{F}_{t_{s-1}})(w). \tag{5.8}$$

The following is a mean upper bound for the total volume of transactions associated to the portfolio $B\Pi$.

Theorem 3.

$$\mathbf{E}(VT_{t_0}(B\Pi)) \leq \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \int_\Omega |D_{t_{s-1},t_s}|. \tag{5.9}$$

Proof. Notice that

$$\cup_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} A_{k,t_s} = A_{l,t_{s-1}}, \tag{5.10}$$

and

$$\cup_{l=0}^{L^n_{t_{s-1}}} A_{l,t_{s-1}} = \Omega, \tag{5.11}$$

$A_{k,t_s} \cap A_{k',t_s} = \emptyset$.

Integrating $\sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w)\ VT_{t_{s-1}}(D_{t_{s-1},t_s})$ in (5.4) over Ω and making use of (5.3) gives

$$\mathbf{E}\Big(\sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w)\ VT_{t_{s-1}}(D_{t_{s-1},t_s})\Big) = \tag{5.12}$$

$$e^{-r(t_s - t_{s-1})}\Big(\sum_{l=0}^{L^n_{t_{s-1}}} \sum_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} |\langle D_{t_{s-1},t_s}, \phi_{k,t_s}\rangle|\ \sqrt{P(A_{k,t_s})}\Big)$$

$$\leq e^{-r(t_s - t_{s-1})}\Big(\sum_{l=0}^{L^n_{t_{s-1}}} \sum_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} \int_{A_{k,t_s}} |D_{t_{s-1},t_s}(w)|\ dP(w)$$

$$= e^{-r(t_s - t_{s-1})} \int_\Omega |D_{t_{s-1},t_s}(w)|\ dP(w).$$

Adding up this equation over $s = 1, \ldots, N$ we obtain (5.9). □

Corollary 2.

$$\mathbf{E}(VT_{t_0}(B\Pi)) \leq ||\mathbf{E}(X|\mathcal{F}_T^n) - \mathbf{E}(X)|| \sqrt{\sum_{s=1}^{N} e^{-2r(t_s - t_{s-1})}}. \tag{5.13}$$

Proof. Notice that the martingale differences $D_{t_s,t_{s+1}}$ are orthogonal, hence

$$\sum_{s=1}^{N} ||D_{t_{s-1},t_s}||^2 = ||\sum_{s=1}^{N} D_{t_{s-1},t_s}||^2 = ||\mathbf{E}(X|\mathcal{F}_T^n) - \mathbf{E}(X)||^2. \tag{5.14}$$

Applying the Cauchy-Scwartz inequality twice

$$\sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \int_\Omega |D_{t_{s-1},t_s}(w)| \, dP(w) \leq \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \sqrt{\int_\Omega |D_{t_{s-1},t_s}(w)|^2 \, dP(w)} \leq$$

$$\sqrt{\sum_{s=1}^{N} ||D_{t_{s-1},t_s}||^2} \sqrt{\sum_{s=1}^{N} e^{-2r(t_s - t_{s-1})}},$$

therefore (5.9) combined with (5.14) gives (5.13). □

Equation (5.13) growths with N and it is not clear that $\sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \int_\Omega |D_{t_{s-1},t_s}|$, in equation (5.9), can be bounded independently of N.

Proposition 3. *Consider an arbitrary path $A_{l_s} \equiv A_{l_s,t_s}$ for a set of indexes l_s $s = 1, \ldots, N$. Let C_1 be an upper bound for the martingale differences along this path, i.e. $|D_{t_{s-1},t_s}(w)| \leq C_1$ a.e. on $\cap A_{l_s}$, then*

$$\hat{VT}_{t_0}(B\Pi)(w) \leq C_1 \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \;\; a.e. \;\, on \;\, \cap A_{l_s}. \tag{5.15}$$

Proof. From the definition of *approximate* values for the volume of transactions we have,

$$\hat{VT}_{t_0}(B\Pi)(w) = \sum_{s=1}^{N} \sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w) \; \hat{VT}_{t_{s-1}}(D_{t_{s-1},t_s})(w) = \tag{5.16}$$

$$\sum_{s=1}^{N} \sum_{l=0}^{L^n_{t_{s-1}}} e^{-r(t_s - t_{s-1})} \, \mathbf{1}_{A_{l,t_{s-1}}}(w) \sum_{\{k, A_{k,t_s} \subseteq A_{l,t_{s-1}}\}} \frac{|\langle D_{t_{s-1},t_s}, \phi_{k,t_s}\rangle|}{\sqrt{P(A_{k,t_s})}} \, \mathbf{E}(\mathbf{1}_{A_{k,t_s}}|\mathcal{F}^n_{t_{s-1}})(w).$$

It follows from (5.16) and (5.2) that the associated *approximate* pointwise volume of transactions for w in the given path, i.e. $w \in \cap_{s=1}^{N} A_{l_s,t_s}$, is:

$$\hat{VT}_{t_0}(B\Pi)(w) = \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \sum_{\{k, A_{k,t_s} \subseteq A_{l_{s-1}}\}} \frac{|\langle D_{t_{s-1},t_s}, \phi_{k,t_s}\rangle|}{\sqrt{P(A_{k,t_s})}} \, \mathbf{E}(\mathbf{1}_{A_{k,t_s}}|\mathcal{F}^n_{t_{s-1}})(w) =$$

$$\sum_{s=1}^{N} \frac{e^{-r(t_s - t_{s-1})}}{P(A_{l_{s-1}})} \sum_{\{k, A_{k,t_s} \subseteq A_{l_{s-1}}\}} | \int_{A_{k,t_s}} D_{t_{s-1},t_s}(w) \; dP(w)|.$$

Now, using $|D_{t_{s-1},t_s}(w)| \; \mathbf{1}_{A_{k,t_s}} \leq C_1$ for almost every $w \in \cap_{s=1}^{N} A_{l_s,t_s}$ and (5.10) we obtain (5.15). □

5.2. **Haar Functions Representation.** The previous bounds for the volume of transactions did not use the Haar coefficients. In this section we deal with the Haar portfolio $H\Pi$, this amounts to represent the martingale differences using the Haar functions instead of the characteristic functions. We will use the following notation, let $\mathcal{T}_{A_{l,t_{s-1}}}$ be the tree with root at atom $A_{l,t_{s-1}}$, moreover, nodes in $\mathcal{T}_{A_{l,t_{s-1}}}$ are elements from $\mathcal{F}^n_{t_s} \setminus \mathcal{F}^n_{t_{s-1}}$. We will write $A \in \mathcal{T}_{A_{l,t_{s-1}}}$ to indicate that A is one of the nodes of the tree (possible its root) *but not a leaf.* Noticing that the disjoint atoms $A_{l,t_{s-1}}$, $l = 0, \ldots, L^n_{t_{s-1}}$ generate $\mathcal{F}^n_{t_{s-1}}$, we have the decomposition:

$$D_{t_{s-1},t_s}(w) = \sum_{l=0}^{L^n_{t_{s-1}}} \sum_{A_{j,i} \in \mathcal{T}_{A_{l,t_{s-1}}}} \langle D_{t_{s-1},t_s}, \psi_{A_{j,i}} \rangle \; \psi_{A_{j,i}}, \tag{5.17}$$

where, of course $\psi_{A_{j,i}} = a_{j,i} \; \mathbf{1}_{A_{j+1,2i}} + b_{j,i} \; \mathbf{1}_{A_{j+1,2i+1}}$ is the Haar function at node $A_{j,i}$ (we have used notation introduced in (2.2)).

The volume of transactions associated to such a decomposition is given by:

(5.18)

$$VT^H_{t_{s-1}}(D_{t_{s-1},t_s})(w) = e^{-r(t_s - t_{s-1})} \sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w) \sum_{A_{j,i} \in \mathcal{T}_{A_{l,t_{s-1}}}} |\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}} \rangle| \times$$

$$\left(|a_{j,i}| \mathbf{E}(\mathbf{1}_{A_{j+1,2i}} | \mathcal{F}_{t_{s-1}})(w) + |b_{j,i}| \mathbf{E}(\mathbf{1}_{A_{j+1,2i+1}} | \mathcal{F}_{t_{s-1}})(w) \right).$$

As previously indicated, the volume of transactions depends on the representation and we have used the notation VT^H to highlight this fact. The approximate version of (5.18) is $\hat{VT}^H_{t_{s-1}}(D_{t_{s-1},t_s})$ and is obtained by replacing $\mathcal{F}_{t_{s-1}}$ by $\mathcal{F}^n_{t_{s-1}}$ in (5.18).

In the computations that follow we assume that our approximating martingale is *regular* ([26]), namely for all parents nodes $A_{j,i}$ and its children we have:

$$P(A_{j+1,k}) \leq \delta P(A_{j,i}) \;, k = 2i, 2i+1,$$

where $\delta > 0$ is the same for all nodes. In order to exploit the representation in terms of Haar functions, and obtain improved upper bounds for the volume of transactions, we will need to assume certain decay of the H-system inner products. The scope of these assumptions as well as the relevance of the results obtained will be explained at the end of the present section.

Here is our first result on pointwise bounds for $H\Pi$.

Proposition 4. *Assume the approximating martingale given by the H-system is regular with* $\frac{1}{2} \leq \delta \leq \frac{\sqrt{2}}{2}$, *and that the coefficients* $|\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}} \rangle|$ *satisfy*

(5.19)

$|\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}} \rangle| \leq C_2 \; P(A_{j,i})^{3/2},$ *at all nodes* $A_{j,i} \in \mathcal{T}_{A_{l,t_{s-1}}}$ *and all roots* $A_{l,t_{s-1}}$

then

$$\hat{VT}_{t_0}(H\Pi)(w) \leq \frac{2 \; \delta \; C_2}{1 - 2\delta^2} \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w) \; P(A_{l,t_{s-1}}). \tag{5.20}$$

Proof. Using the approximated version of (5.18), the regularity assumption and the constraints $\mathbf{E}(\psi) = 0$ and $\mathbf{E}(\psi^2) = 1$ we obtain:

$$\hat{VT}^H_{t_{s-1}}(D_{t_{s-1},t_s})(w) \leq 2 \; \delta \; e^{-r(t_s - t_{s-1})} \; \times \tag{5.21}$$

$$\sum_{l=0}^{L^n_{t_{s-1}}} \frac{\mathbf{1}_{A_{l,t_{s-1}}}(w)}{P(A_{l,t_{s-1}})} \sum_{A_{j,i}\in \mathcal{T}_{A_{l,t_{s-1}}}} \sqrt{P(A_{j,i})}\ |\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}}\rangle|.$$

Using (5.19), we obtain

$$\sum_{A_{j,i}\in \mathcal{T}_{A_{l,t_{s-1}}}} \sqrt{P(A_{j,i})}\ |\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}}\rangle| \leq C_2 \sum_{A_{j,i}\in \mathcal{T}_{A_{l,t_{s-1}}}} P(A_{j,i})^2. \tag{5.22}$$

Under the assumption $\frac{1}{2} \leq \delta \leq \frac{\sqrt{2}}{2}$ and considering $r_s = j^n_s - j^n_{s-1}$ (as indicated previously, we make use of notation introduced elsewhere in the paper, see (3.4) and Theorem 2), we obtain

$$\sum_{A_{j,i}\in \mathcal{T}_{A_{l,t_{s-1}}}} P(A_{j,i})^2 \leq \sum_{j=0}^{r_s} \sum_{i=0}^{2^j-1} \left(\delta^j P(A_{l,t_{s-1}}\right)^2 = P(A_{l,t_{s-1}})^2 \sum_{j=0}^{r_s} (2\delta^2)^j \tag{5.23}$$

$$\leq P(A_{l,t_{s-1}})^2 \frac{1}{1-2\delta^2}.$$

Finally, using (5.21), (5.22) and (5.23)

$$\hat{VT}^H_{t_{s-1}}(D_{t_{s-1},t_s})(w) \leq \frac{2\ \delta\ C_2}{(1-2\delta^2)}\ e^{-r(t_s-t_{s-1})} \sum_{l=0}^{L^n_{t_{s-1}}} \mathbf{1}_{A_{l,t_{s-1}}}(w)\ P(A_{l,t_{s-1}}). \tag{5.24}$$

Equation (5.20) then follows by using (5.4). □

Corollary 3. *Consider the same hypothesis as in Proposition 4, then*

$$\mathbf{E}(VT_{t_0}(H\Pi)) \leq \frac{2\ \delta\ C_2}{(1-2\delta^2)^2}.$$

Proof. Using (5.3) allows to reduce the problem for computing an upper bound for $\mathbf{E}(VT_{t_0}(H\Pi))$ to computing an upper bound for $\mathbf{E}(\hat{VT}_{t_0}(H\Pi))$, to this end we will make use of (5.20). Here is the computation

$$\mathbf{E}(VT_{t_0}(H\Pi)) = \mathbf{E}(\hat{VT}_{t_0}(H\Pi)) \leq \frac{2\ \delta\ C_2}{1-2\delta^2} \sum_{s=1}^{N}\ e^{-r(t_s-t_{s-1})} \sum_{l=0}^{L^n_{t_{s-1}}} P^2(A_{l,t_{s-1}}) \leq$$

$$\frac{2\ \delta\ C_2}{1-2\delta^2}(1+(2\delta^2)^{j^n_1}+\ldots+(2\delta^2)^{j^n_N}) \leq \frac{2\ \delta\ C_2}{(1-2\delta^2)^2}.$$

□

The upper bound (5.20) for a single path $w \in \cap A_{l_s}$, amounts to dropping the sum $\sum_{l=0}^{L^n_{t_{s-1}}}$ in (5.20). This fact plus a computation similar to the one employed in Corollary 3 gives the following result.

Corollary 4. *Consider an arbitrary path $A_{l_s} \equiv A_{l_s,t_s}$ for a set of indexes l_s, $s = 1, \ldots, N$, $w \in \cap A_{l_s,t_s}$ and the same hypothesis as in Proposition 4 (in this case (5.19) is only required at the nodes of the given path), then*

$$\hat{VT}_{t_0}(H\Pi)(w) \leq \frac{2\ \delta\ C_2}{1-2\delta^2} \sum_{s=1}^{N}\ e^{-r(t_s-t_{s-1})}\ P(A_{l_{s-1},t_{s-1}}) \leq \frac{2\ \delta\ C_2}{(1-2\delta^2)(1-\delta)}.$$

The following result uses a weaker assumption on the decay of the inner products $|\langle D_{t_{s-1},t_s}, \psi_{A_{j,i}}\rangle|$. The result relies on the underlying tree $\mathcal{T}_n$ associated to the approximating H-system, see the text surrounding (2.2).

Proposition 5. *Consider the same hypothesis as in Proposition 4 but relaxing (5.19) as follows: assume (5.19) holds at all nodes but for each t_s, $s = 1, \ldots, N$, there are nodes A^k_{l,t_s}, $k = 1, \ldots, K$ (where K an absolute constant), which are atoms in $\mathcal{F}^n_{t_s}$, satisfying: for each such a node there is at most a path of nodes $A^{r+1}_{t_s} \subseteq A^r_{t_s}$, $A^0_{t_s} \equiv A^k_{l,t_s}$, $A^r_{t_s} \in \mathcal{T}_{A^k_{l,t_s}}$ and for each $k = 1, \ldots, K$ we have*

$$|\langle D_{t_{s-1},t_s}, \psi_{A^r_{t_{s-1}}}\rangle| \leq C_3\ P(A^r_{t_{s-1}})^{1/2}, \text{ holds for all } A^r_{t_{s-1}} \text{ in such a path.} \tag{5.25}$$

then

$$\hat{VT}_{t_0}(H\Pi)(w) \leq \frac{2\ \delta\ C_3}{1-\delta} \sum_{s=1}^{N} e^{-r(t_s - t_{s-1})} \sum_{k=1}^{K} \mathbf{1}_{A^k_{l,t_{s-1}}}(w) + \tag{5.26}$$
$$\frac{2\ \delta\ C_2}{1-2\delta^2} \sum_{s=1}^{N} e^{-r(t_s-t_{s-1})} \sum_{A_{l,t_{s-1}} \neq A^k_{l,t_{s-1}}} P(A_{l,t_{s-1}})\ \mathbf{1}_{A_{l,t_{s-1}}}(w),$$

where the notation $\sum_{A_{l,t_{s-1}} \neq A^k_{l,t_{s-1}}}$ excludes summation over the atoms $A^k_{l,t_{s-1}}$, $k = 1, \ldots, K$. Moreover,

$$\mathbf{E}\left(VT_{t_0}(H\Pi)\right) \leq \frac{2\ \delta\ K\ C_3}{(1-\delta)^2} + \frac{2\ \delta C_2}{(1-2\delta^2)^2}. \tag{5.27}$$

Proof. Consider $A^k_{l,t_{s-1}} \in \mathcal{F}^n_{t_{s-1}}$ and $r_s \equiv j^n_s - j^n_{s-1}$ we estimate (the sets $A^r_{t_{s-1}}$ are the path elements with root $A^k_{l,t_{s-1}}$ as described in the hypothesis)

$$\sum_{A^r_{t_{s-1}} \subseteq A^0_{t_{s-1}} \equiv A^k_{l,t_{s-1}}} \sqrt{P(A^r_{t_{s-1}})}\ |\langle D_{t_{s-1},t_s}, \psi_{A^r_{t_{s-1}}}\rangle| \tag{5.28}$$
$$\leq C_3\ P(A^k_{l,t_{s-1}}) \sum_{j=0}^{r_s} \delta^j \leq C_3\ P(A^k_{l,t_{s-1}})\ \frac{1}{1-\delta}.$$

Finally, adding over s, using (5.21) (5.20), (5.28) and (5.4) we obtain (5.26). The mean upper bound (5.27) follows by integrating (5.26) and noticing that $\sum_{s=1}^{N} P(A^k_{l,t_{s-1}}) \leq \sum_{s=1}^{N} \delta^{j^n_s} \leq \frac{1}{1-\delta}$. □

We comment next on the relevance and meaning of the hypothesis employed in Propositions 4 and 5. Notice first that (5.25) holds under the assumption that D_{t_{s-1},t_s} is bounded, we may assume this hypothesis is satisfied in practice as one could consider a bounded approximation to the payoff X which holds in a large portion of the space.

Our hypothesis on the inner products' decay reflects well known results in wavelet theory, in particular the usual case of Haar wavelets on an interval $[a, b]$ (which are a special case of our setup). We will briefly explain the meaning of these results and the relationship to our context. The results we are referring to relate smoothness and singularities of the function to decay of inner products. In our context singularities and smoothness of the functions involved are notions defined by viewing these functions as dependent on S_{t_s}, where S_t is the underlying process. Define $\gamma(x) \equiv D_{t_{s-1},t_s}(w)$ where $x \equiv S_{t_s}(w)$. Theorem 6.4 in [20] says that if γ

is Lipschitz of order $\alpha = 1$, Haar wavelet's inner products decay as $s^{3/2}$ where s is the scale parameter. In particular, this result applies if γ has one continuous derivative. The inner products will decay as $s^{1/2}$ in the presence of singularities (in particular if the derivative is discontinuous). These results are local and make use of the wavelets being localized to the extent of their scales s. A location at which there is a singularity will have the effect that inner products of wavelets having support at this location will have a decay of order $s^{1/2}$, this is the reason for our hypothesis (5.25) to hold for an atom and all of its children along a single path.

Notice that the above wavelet results rely on analytical tools such as translations and dilations, which are available in $\mathbb{R}$, and its relation to Lebesgue measure. The decay (5.19) is not easily related to smoothness properties of $\gamma(x)$ in our general setup. The reason being that the H-systems we have considered in the paper are generated by general partitions. Constructing H-systems with the property that the decay of its inner products reflects the smoothness of the function being analyzed is under present investigation.

The localization of our approximations provide already interesting bounds like (5.9). Bounds like the ones in Corollaries 3 and 4 are excellent and exploit the Haar representation. Proposition 5 shows that the Haar representation can incorporate singularities in a localized way and at the same time continue taking advantage of faster decay of inner products away from singularities. These properties are not available in delta hedging where discontinuities creates several problems with hedging (see, for example, [12]).

6. Numerical Examples

In this section we present output from a computer implementation (detailed in [3]) based on the Brownian motion example from Section 3.1. 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 below and some of the definitions and notions introduced elsewhere in the paper.

We analyze the error of the approximations as well as the volume of transactions. To indicate the potential improvements that can be expected for this example we will only 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 $n_T = 1$. Errors along these steps accumulate as is the case with delta hedging. General expressions for the volume of transactions for our portfolios are presented in Section 5.

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

$$X^c_{(m)} = \sum_{i=0}^{m-1} \langle x, u_{k_i}\rangle u_{k_i}. \tag{6.1}$$

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 introduced above, 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 some information on this algorithm and associated notation) for the H-system $\{u_{2^j+i}\}$ associated to geometric Brownian motion described in Section 3.1 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

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

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

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

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

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

Next we comment on the output displays; numerical values were obtained by sampling $S_T(\omega)$, the limited range in these values (x-axis on most displays) correspond to these sampled values (after sorting). Consider first a single European call $X(\omega) = (S_T(\omega) - K)_+$ as above, values of parameters are indicated in the text surrounding the figures. In Figures 1, 2 and 3 we present the Black-Scholes and Haar approximations with $m = 1,\ 2,\ 20$ 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. Table 1 indicates the effect of optimizing the number of transactions (i.e. compression, as described above), for example we obtain a smaller error 0.10, by chosing $m = 16$ -Haar functions out of 2^{16} basis elements, than the error 0.14, obtained by choosing $m = 64$ -Haar functions out of 2^6 elements.

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 $X^c_{(m)}$ from (6.1), it is easy to show that the volume of transactions for the Haar hedging portfolio (as defined in Section 5), is equal to

(6.3)

$$VT_{t_0}(\text{H}\Pi) = e^{-r(T-T_0)}||X^c_{(m)} - \mathbf{E}(X)||_{L^1} = e^{-r(T-T_0)} \int_\Omega |X^c_{(m)}(\omega) - \mathbf{E}(X)| dP(\omega).$$

The volume of transactions for the portfolio of binary options is

$$VT_{t_0}(\text{B}\Pi) = e^{-r(T-T_0)}||X^c_{(m)}||_{L^1} = e^{-r(T-T_0)} \int_\Omega |X^c_{(m)}(\omega)| dP(\omega). \tag{6.4}$$

Expressions for (6.3) and (6.4) when $n_T > 1$ can be obtained from the general developments in Section 5. 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, a detailed analytical analysis is provided in Section 5.

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, the Haar functions can be 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 $A_{1,1} = \{S_T \geq K\}$ 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. A main goal of the paper is to bring forward the importance of efficient hedging strategies, these alternatives representations should be contrasted with delta hedging in order to compare trade offs and limitations of this last technique. To this end, the paper emphasizes efficiency in relation to volume of transactions and gives theoretical as well as numerical evidence that the suggested new hedging strategy has several advantages over delta hedging.

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 also 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 elsewhere in 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.

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

$$X_j \equiv X_{\mathcal{R}_j} \equiv \mathbf{E}(X|\sigma(\mathcal{R}_j)). \tag{A.1}$$

Then we have the following expansions

$$X_j(\omega) = c_0[0]\ \phi_{0,0}(\omega) + \sum_{k=0}^{j-1} \sum_{i \in I_k} d_k[i]\ \psi_{k,i}(\omega) = \sum_{(k,i) \in K_j} c_k[i]\ \phi_{k,i}(\omega) \tag{A.2}$$

where

$$c_k[i] = \langle X_j, \phi_{k,i} \rangle \quad \text{and} \quad d_k[i] = \langle X_j, \psi_{k,i} \rangle. \tag{A.3}$$

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

$$c_k[i] = \langle X_j, \phi_{k,i} \rangle = \frac{1}{\sqrt{p_k[i]}} \int_{A_{k,i}} X_j dP = \frac{1}{\sqrt{p_k[i]}} \int_{A_{k,i}} X dP = \langle X, \phi_{k,i} \rangle, \tag{A.4}$$

where we have made use of the array notation $p_j[i] \equiv P(A_{j,i})$. Analogously, we have that $d_k[i] = \langle X, \psi_{k,i} \rangle$. Moreover, we can state the following proposition.

Proposition 6. *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*

$$X_j = X_{j'} + \sum_{k=j'}^{j-1} \sum_{i \in I_k} d_k[i] \psi_{k,i} \tag{A.5}$$

and

$$\sum_{(k,i) \in K_j} c_k^2[i] = \sum_{(k,i) \in K_{j'}} c_k^2[i] + \sum_{k=j'}^{j-1} \sum_{i \in I_k} d_k^2[i]. \tag{A.6}$$

For a proof of the above Proposition and further details of the multiresolution algorithm for H-systems, see [4]. The algorithm is an adaptation of the well known algorithm for wavelet theory given for S. Mallat [19] to our probabilistic setting. This algorithm produces a relation between the samples of X, namely,

$$x_k[i] = X_j(\omega),\ \omega \in A_{k,i}, \text{for } (k,i) \in K_j, \tag{A.7}$$

and the coefficients $d_k[i]$.

Appendix B. Complements

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 $\{H_i\}_{i=1}^n$ 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

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

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

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

From Theorem 1 it follows 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 $\mathcal{F}_{t_j} = \sigma(S_{t_0}, ...S_{t_j})$.

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 $n + 1 \leq 2^J$, then for $0 \leq j \leq J$ and $0 \leq i \leq 2^j - 1$ we define the sets $A_{j,i}$, as follows. For $i \neq 0$,

$$A_{j,i} = \{\omega \in \Omega : \frac{i}{2^j} < \frac{1}{n}|\omega|_U \leq \frac{i+1}{2^j}\} \tag{B.3}$$

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

$$A_{j,0} = \{\omega \in \Omega : 0 \leq \frac{1}{n}|\omega|_U \leq \frac{1}{2^j}\} \tag{B.4}$$

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

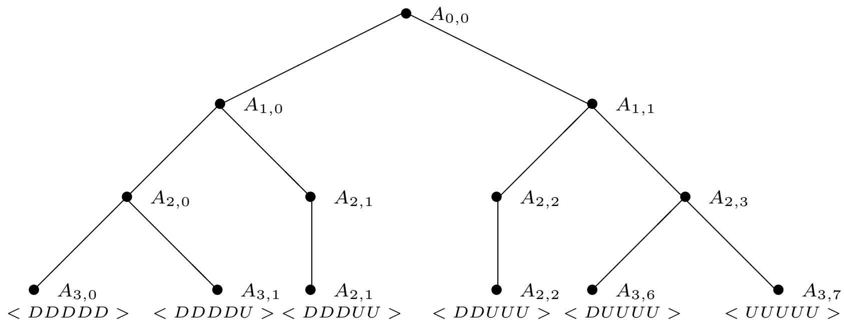
$$P(A_{j,0}) = \sum_{0 \leq \frac{s}{n} \leq \frac{1}{2^j}} \binom{n}{s} p^s q^{n-s}.$$

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, by using Theorem 1, namely by

$$\begin{aligned} &v_0 \equiv 1, \\ &v_{j,i} = \tfrac{1}{\sqrt{P(A_{j,i})}}\Big(\sqrt{\tfrac{P(A_{j+1,2i+1})}{P(A_{j+1,2i})}} 1_{A_{j+1,2i}} - \sqrt{\tfrac{P(A_{j+1,2i})}{P(A_{j+1,2i+1})}} 1_{A_{j+1,2i+1}}\Big) \end{aligned} \tag{B.5}$$

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 $X^c_{(m)}$, 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

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)	$VT(\mathrm{B\Pi})$	$VT(\mathrm{H\Pi})$	`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

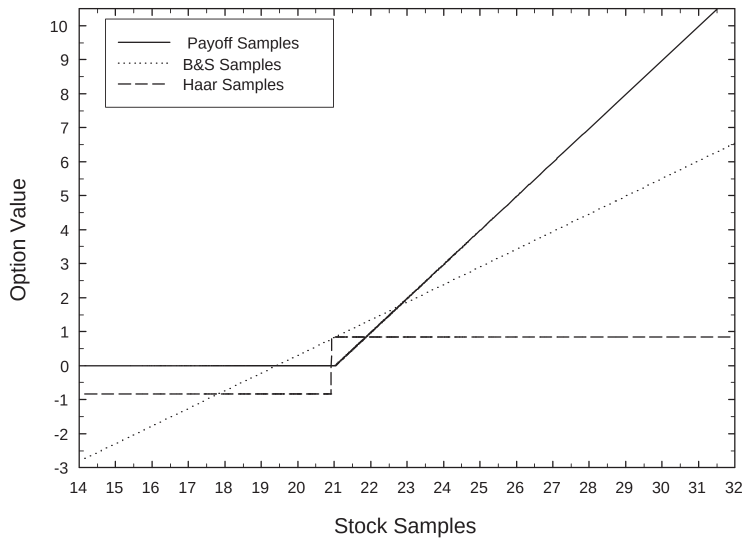


FIGURE 1. Approximations to single European Call using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $m = 1,\ S_{T_0} = 20,\ r = 0.05,\ \sigma = 0.1,\ T - T_0 = 1,\ K = 21$.

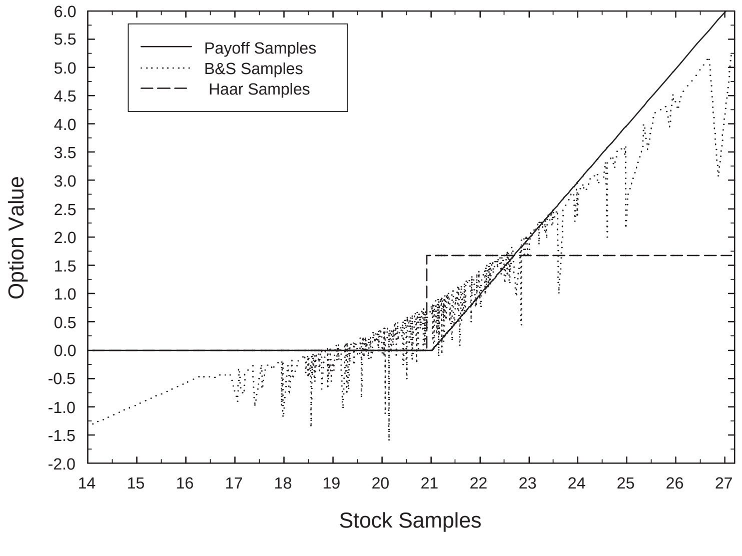


FIGURE 2. Same as in Figure 1 except $m = 2$.

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, 21-52.

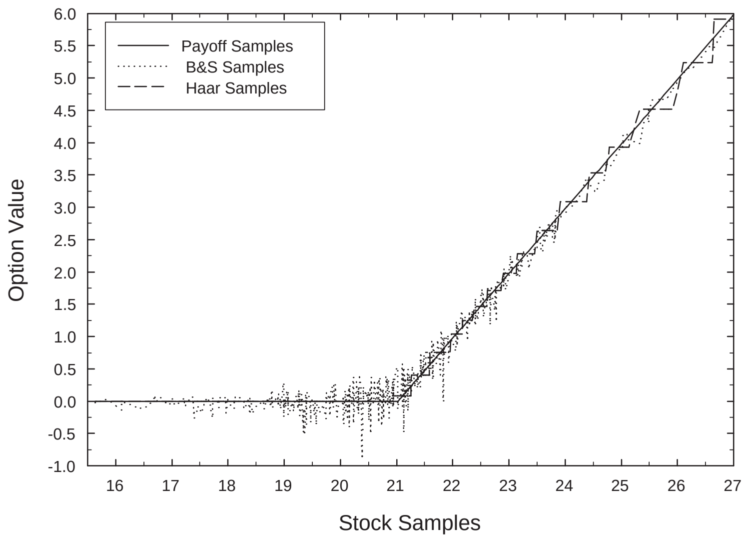


FIGURE 3. Same as in Figure 1 except $m = 20$.

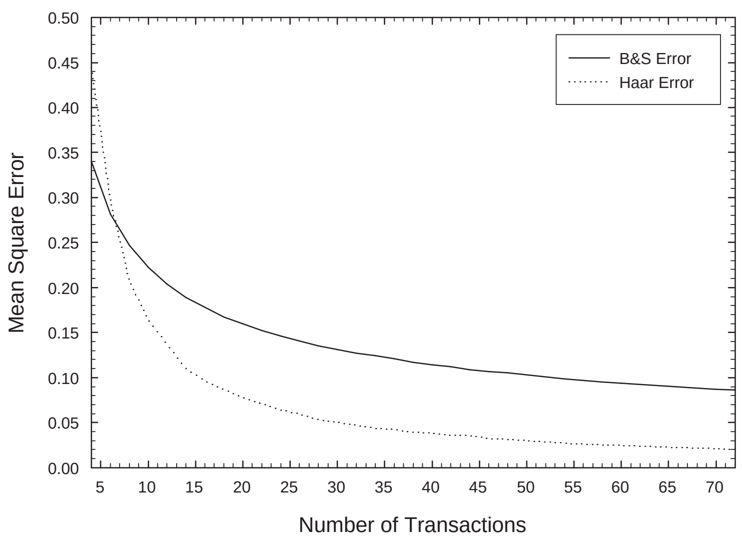


FIGURE 4. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter m. Values of the parameters used: as in Figure 1.

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

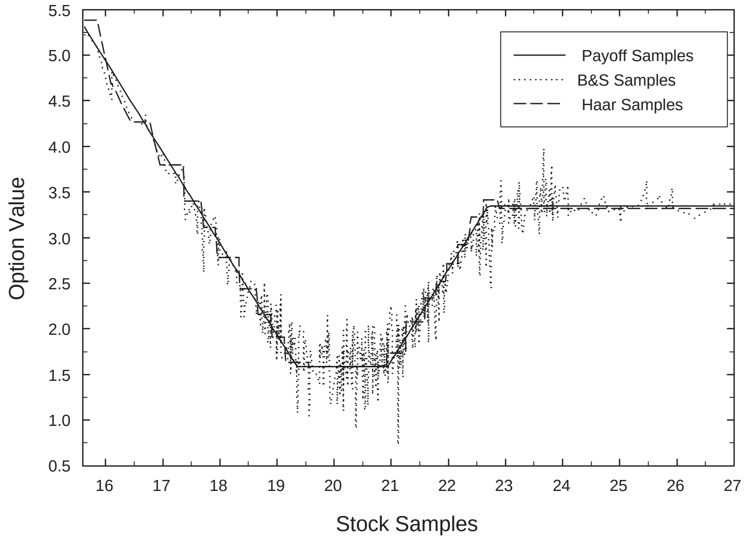


FIGURE 5. Approximations to portfolio X, constructed from two calls and one put, using delta hedging and the Haar system constructed via Brownian motion increments. Values of the parameters used: $m = 20$, $S_{T_0} = 20$, $r = 0.05$, $\sigma = 0.1$, $T - T_0 = 1$, $K_1 = 19$, $K_2 = 21$ $K_3 = 23$.

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

[3] J. Cai and S. E. Ferrando (March 2004), C++ Implementation of Haar Systems for Black-Scholes Model. *Technical Report, Ryerson University.*
[4] P.J. Catuogno, S.E. Ferrando and A.L. Gonzalez (December, 2008), *Adaptive martingale approximations*. Journal of Fourier Analysis and Applications . Volume 14, Numbers 5-6, 712-743.
[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.

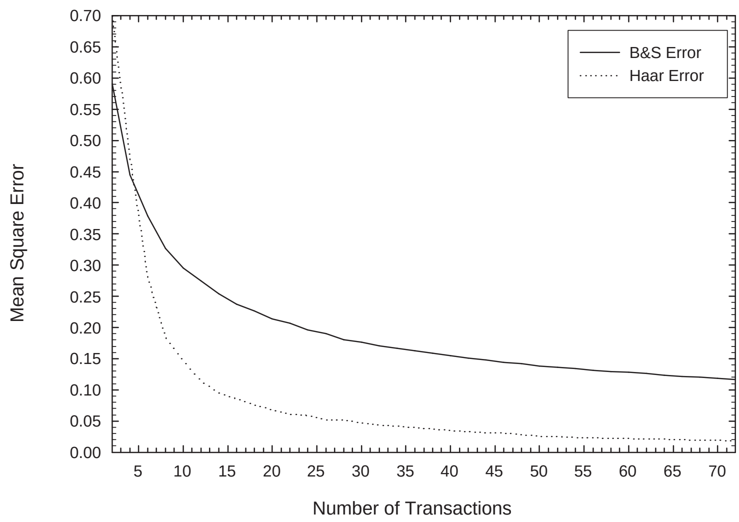


FIGURE 6. L^2 norm of the errors between the option X and delta hedging and Haar approximations respectively. The plot is in terms of the parameter m. Values of the parameters used: as in Figure 5.

[7] Ronald A. DeVore (Summer 1998), Nolinear approximation. *Acta Numerica*, 1-99.

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

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

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

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

[12] C. Gallus (1999), Exploding hedging errors for digital options. *Finance and Stochastic*, 3, 187-201-.

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

[14] 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.

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

[16] D.G. Hobson (1998), Robust hedging of the lookback option. *Finance and Stochastics*, 2, 329-347.

[17] P. Jaeckel (2002), Monte Carlo Methods in Finance. John Wiley and Sons.

[18] 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.

[19] S. Mallat (1989), Multiresolution aproximations and wavelets orthonormal bases for $L^2(\mathbb{R})$. *Trans. of Amer. Math. Soc.*, 315, 69-87.

[20] S. Mallat (1999). *A Wavelet Tour of Signal Processing.* Academic Press, second edition.

[21] 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.

[22] 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.

[23] J. Neveu (1975): *Discrete-Parameter Martingales*, North-Holland.
[24] A. Pechtl (1995), Classified information. *RISK*, 8 (6), 59-61.
[25] D. Percival and A. Walden (2000): *Wavelets Methods for Time Series Analysis*, Cambridge University Press, Cambridge.
[26] W.F Stout (1974):*Almost Sure Convergence*. Academic Press.
[27] A. Szimayer and R.A¿ Maller (2007), Finite approximation schemes for Levy processes, and their application to optimal stopping problems. *Stochastic Processes and Their Applications*, Vol. 117, 1422-1447.
[28] W. Willinger (1987), *Pathwise Stochastic Integration and Almost-Sure Aproximation of Stochastic Processes.*, Thesis, Cornell University.
[29] W. Willinger and M.S. Taqqu (January 1991), Toward a convergence theory for continuous stochastic securities market models. *Mathematical Finance*, Vol. 1, No. 1, 55-99.

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

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