# Static Hedging and Pricing of Exotic Options With Payoff Frames

**Article** *in* Mathematical Finance · January 2019

DOI: 10.2139/ssrn.2501812



**2 authors:**

Justin Lars Kirkby
Georgia Institute of Technology
**63** PUBLICATIONS **773** CITATIONS
SEE PROFILE

Shi-Jie Deng
Georgia Institute of Technology
**50** PUBLICATIONS **1,839** CITATIONS
SEE PROFILE

# STATIC HEDGING AND PRICING OF EXOTIC OPTIONS WITH PAYOFF FRAMES

J. LARS KIRKBY AND SHIJIE DENG

Abstract. We develop a general framework for statically hedging and pricing European-style options with nonstandard terminal payoffs which can be applied to mixed static-dynamic and semi-static hedges for many path dependent exotic options including variance swaps and barrier options. The goal is achieved by separating the hedging and pricing problems to obtain replicating strategies. Once prices have been obtained for a set of basis payoffs, the pricing and hedging of financial securities with arbitrary payoff functions is accomplished by computing a set of"hedge coefficients" for that security. This method is particularly well suited for pricing baskets of options simultaneously, and is robust to discontinuities of payoffs. In addition, the method enables a systematic comparison of the value of a payoff (or portfolio) across a set of competing model specifications with implications for security design.

## 1. Introduction and Literature Review

To accommodate the growing demand for nonstandard derivative payoffs, markets have developed to facilitate their trading. The sources of risk on which these contingent payoffs depend have become increasingly diverse, as have the payoff structures specified in their contracts, which necessitates robust pricing and hedging strategies to preempt the admission of arbitrage.

One strand of research in the continuously-expanding derivatives pricing literature focuses on the static-hedging approach to pricing through statically or semi-statically replicating the complex derivative payoffs with simple payoffs which are more amenable to valuation (see [1, 7, 10, 12, 14, 36, 37] and the references therein). The critical feature that distinguishes static hedging from ordinary function approximation is that traditional numerical approaches take mesh refinement for granted, whereas static hedging is constrained by the availability of market payoffs, which in turn define the mesh. This distinction motivates the pursuit of approximation with limited mesh refinement, where"basis" functions are chosen according to their availability and liquidity in a given market. We show that by recasting the problem as one of projection onto a suitably chosen frame space, the optimal static hedge is efficiently obtained. Furthermore, such approximations are shown to produce exact representations in the limit as the mesh is refined, much like the integral representations provided by [12].

The contributions of our analysis are as follows. First, we advance a new theoretical framework for pricing contingent claims and studying their perfect static replication strategies, with implications for security design. Specifically, new financial instruments can be introduced according to the richness of payoffs they are able to synthesize, and in a way that generates standardized markets such as those for plain vanilla options. Frames provide the flexibility to study spaces of claims spanned by simpler securities.

Second, we provide a systematic scheme for hedging exotic derivatives including path-dependent options through a new means of static replication that can be implemented in markets with a reasonable spectrum of strikes on European options spanning practical trading ranges. By improving the accuracy of static replication, fewer instruments are required to achieve a desired hedge, and hedges are obtained with greater efficiency using methods developed in this work. Specifically, our approach consistently reduces the relative hedge error by more than half, while at a cost within fractions of a millisecond of basic interpolation. The risk reduction across a portfolio of exotic payoffs can be substantial. Not only is the hedging of nonlinear European

*Date*: 2017.

The authors would like to thank the editors and two anonymous referees, whose suggestions have greatly improved the quality of this manuscript.

options improved, but the method extends to procedures designed to hedge path dependent options. In particular, semi-static hedges for barrier and American options as well as mixed static-dynamic strategies employed for products such as realized variance swaps are improved by the new methods.

Finally, accurate prices for exotic payoffs are obtained by combining the static hedge with existing methods for pricing vanilla instruments. Prices computed from projected payoffs converge at a rate that is often several orders faster than when pricing the payoffs directly. Hence, once payoff coefficients are obtained, subsequent valuations are implemented at a fraction of the cost. This method is particularly well suited for pricing baskets of options simultaneously, and is robust to discontinuities of payoffs. In addition, the method enables a systematic comparison of the value of a payoff (or portfolio) across a set of competing model specifications.

Traditional approaches to static hedging typically impose restrictions on the underlying's risk-neutral dynamics, though superior results relative to dynamic hedging have been documented when the assumptions hold and even when these assumptions are relaxed. In [8], a put-call symmetry is established which yields a parity between call and put prices at different strikes, assuming a particular symmetry holds for the underlying or an auxiliary process. A powerful result (which inspired the present work) is found in [12], where a static integral representation is shown to hold for a large class of functions in terms of liquid assets along a spectrum of strikes. A similar representation is derived in [14] which relates call prices to a spectrum of nearer-expiry calls. Static hedging of exotic options has experienced great success in recent years (see [20] for an early account). Numerical and simulation studies demonstrating the superiority (in terms of replication error) and robustness to model misspecification relative to discrete delta hedging are given in [14], [21] and [37] (see also [41] for simulation studies involving Asian, barrier, lookback, and quanto options).

The approach to static hedging via orthonormal basis representation has been studied in [36] and [19], where special features of the underlying risk-neutral dynamics are used to construct an orthonormal basis for the claim space using the valuation operator. Option representation (spanning) in terms of characteristic functions is introduced and analyzed in [1]. As in [12], our approach is based on synthesizing a target payoff function with a set of simple, liquidly traded payoffs (that is, contingent claims on specific payoff forms), where hedging instruments are prescribed according to features of the physical payoff to be received. Once hedging is accomplished (or eschewed if prices are the only risk source), valuation follows either by observing market prices for the simple payoffs, or by specifying a model and then pricing the simple payoffs simultaneously.

The remainder of the paper is organized as follows. Section 2 presents several applications of pricing and hedging exotic derivatives and path-dependent options where our proposed framework can be effectively applied. Section 3 introduces the methodology, and develops the theory for frame-based pricing and hedging. In section 4, frame theory is applied to the current state of option markets by utilizing a basis formed from actively traded vanilla options, coined the butterfly basis, whereby we obtain an analytical representation of the "dual basis". The classes of payoff functions to which the methods apply are characterized in section 4.2. Calculation of hedge coefficients is discussed in section 5, where we develop a new method of function approximation with *alternative biorthogonal sequences.* Section 6 presents several applications, including static hedges for exotic European payoffs, mixed static-dynamic hedges for variance swaps and semi-static hedges for barrier options. Concluding remarks are provided in section 7. An additional set of numerical pricing experiments is conducted with respect to the butterfly basis in the Appendix, demonstrating the method's acceleration of option value convergence. Proofs are provided in the appendix as well. A supplemental appendix is also available in the authors' online version, containing additional experiments and complimentary results.

## 2. Motivating Applications

In this section we consider several financial applications to motivate the framework presented. In each of these applications the unifying strategy is to identify a static representation of some or all of the risk inherent in a financial position, expressed as a function $g(S_T)$ of the underlying risk-factor $S_T$, where $T$ indexes a static time horizon. Given a suitably chosen basis $\{\Psi_k(S_T)\}$,

we form an approximation $g(S_T) \approx \sum_k \alpha_k \Psi_k(S_T)$. If possible, we will decompose $\Psi_k(S_T)$ into tradable market securities, which provides an implementable hedge. Regardless, pricing of $g(S_T)$ is accomplished by pricing the component functions, either using a model or inferring their prices from traded instruments. As discussed in Section B.1, problems with multiple decision periods can be handled as well, by identifying a series of static exposures.

2.1. **Static Hedging: Nonlinear Risks.** Despite the prevalence of nonlinear risk in financial applications, its idiosyncratic nature often leads firms to seek customized over-the-counter payoffs to offset exposure. For example, demand elasticities faced by commodity producers cause revenues to vary nonlinearly with realized commodity prices. If the price risk of future revenues can be quantified it is possible to acquire protection in the form of financial contracts [45]. For instruments with interest rate sensitivities, convexity risk poses a similar problem which can be eliminated by an offsetting *power straddle* position (see [42] and the references therein), which has a terminal payoff $g(S_T) = (S_T - K)^2$, where $S_T$ is an underlying source of randomness. Additionally, power straddles can be used as tool for capturing implied volatility (*vega*) risk faced by options traders [40]. According to [42] power straddles face constant exposure to future implied volatility, allowing the holder to lock in future levels. In addition to power straddles, one often encounters the *powered call* option [27]

$$g(S_T) = \left(\max\{S_T - K, 0\}\right)^2 = \left([S_T - K]^+\right)^2,$$

and the *power call* option

$$g(S_T) = \max\{S_T^2 - K^2, 0\} = [S_T^2 - K^2]^+.$$

In particular, the power straddle can be decomposed in terms of a powered call and put option $(S_T - K)^2 = ([S_T - K]^+)^2 + ([K - S_T]^+)^2$, so any one of these contracts can be priced or hedged in terms of the other two. To avoid catastrophically large payoffs, one can consider the capped power payoffs. For example, a p-th order capped power call pays

$$g(S_T) = (S_T - K)^p \mathbb{1}_{[K \le S_T \le C]} + (C - K)^p \mathbb{1}_{[S_T > C]},$$

for some $C > K$. From the perspective of the option supplier, we develop methods to hedge the sale of general nonlinear contracts in terms of more liquid instruments.[1] Experiments illustrating the effectiveness of this framework for exotic European options are given in Appendix A.

2.2. **Semi-Static Hedging: Barrier and American Options.** Semi-static refers to a hedging strategy that requires finitely many trades during the life of a contract. In [6], an approach to semi-static hedging of barrier options using only European options maturing on the same date is developed that requires at most one transaction during the barrier option's life ( [8] extends the approach to include rolldown, ratchet and lookback options as well). This approach is used as well by [39] to price and hedge barrier and lookback options.

While a Black-Scholes economy is required for the hedge to provide perfect replication, simulation studies have shown superior performance in terms of hedge error variance relative to dynamic hedging strategies for specifications including Heston's stochastic volatility model, Merton's jump diffusion model, and the variance gamma model [37]. These results are further extended rigorously to local/stochastic volatility models and time-changed Lévy processes when a symmetry condition is satisfied [10].

Take for example the down-and-in claim which pays out $f(S_T)$ at time $T$ as long as a lower barrier $H < S_0$ is breached during $[0, T]$. Based on the result of [6, 7] (later extended to more general dynamics in [10]), we define the adjusted payoff

$$\tilde{f}^{DI}(S_T) = \left[f(S_T) + (S_T/H)^p \, f\left(H^2/S_T\right)\right] \mathbb{1}_{[S_T < H]},$$

where $p := 1 - 2(r - d)/\sigma^2$, with interest rate $r \ge 0$, dividend yield $d \ge 0$, and volatility $\sigma > 0$. If $S_t$ fails to reach the lower barrier for $t \in [0, T]$ this European payoff expires worthless, as does the down-and-in claim. However, if the barrier is hit at some $\tau_H \in [0, T]$, then the value of $(S_T/H)^p \, f\left(H^2/S_T\right) \mathbb{1}_{[S_T < H]}$ coincides at time $\tau_H$ with the value of $f(S_T)\mathbb{1}_{[S_T > H]}$. Thus, if the

[1] Moreover, rather than offer standardized markets for a plethora of nonstandard products, the framework we develop can be used to design standardized markets capable of approximating a multitude of nonstandard payoffs with high accuracy.

proceeds generated by selling $(S_T/H)^p f\left(H^2/S_T\right) \mathbb{1}_{[S_T<H]}$ at time $\tau_H$ are used to purchase the payoff $f(S_T)\mathbb{1}_{[S_T>H]}$, the position held over $[\tau_H, T]$ is the payoff $f(S_T)$, which matches that of the down-and-in claim. Hence the semi-static hedge coincides perfectly with the down-and-in position for any path of the underlying. For a concrete example, by in-out parity the adjusted payoff corresponding to the down-and-out and up-and-out contracts satisfy

$$\tilde{f}^{DO}(S_T) = h(S_T)\mathbb{1}_{[S_T>H]} - (S_T/H)^p h\left(H^2/S_T\right) \mathbb{1}_{[S_T\leq H]}$$

and

$$\tilde{f}^{UO}(S_T) = -(S_T/H)^p h\left(H^2/S_T\right) \mathbb{1}_{[S_T>H]} + h(S_T)\mathbb{1}_{[S_T\leq H]},$$

where $h(S_T) = (S_T - K)^+$ for a call and $h(S_T) = (K - S_T)^+$ for a put. In Section 6.4, we demonstrate the application of butterfly basis hedging to down-and-out put options.

In a similar pursuit, certain options with early exercise features such as American binary claims [13] offer perfect static replication strategies in a Black-Scholes economy. Given the pair of *stationary* securities $S_T^{\gamma\pm\epsilon}$ whose values are invariant over $[0, T]$, an American binary claim is perfectly replicated by the European payoff

$$\left[(S_T/H)^{\gamma+\epsilon} + (S_T/H)^{\gamma-\epsilon}\right] \mathbb{1}_E, \tag{1}$$

where $E = [S_T < H]$ for a put, $E = [S_T > H]$ for a call, $\gamma := \frac{1}{2} - \frac{r-d}{\sigma^2}$, and $\epsilon := \sqrt{\gamma^2 + \frac{2r}{\sigma^2}}$. In all of these cases, the success of a static or semi-static replicating strategy depends on the existence of a liquid market in nonlinear European payoffs. Of course no such markets exist, so the task of replication is again to find a suitable approximation for nonlinear payoffs using actively traded instruments. The analysis to follow is in response to these demands for accurate nonlinear payoff approximations.

## 3. Frames of Hedging Instruments and the Basis Theory

In what follows, we consider a generic European option market on the time $T$ realization of an underlying process (e.g an interest rate, equity, index, etc.), denoted $S_T$, in which a set of payoff forms $\mathcal{M}_T$ are currently traded. For example, $f \in \mathcal{M}_T$ where $f(S_T) = (S_T - K)^+$ would denote the payoff of a European call option on $S_T$. The present work analyzes how to "optimally" approximate a general payoff $h \notin \mathcal{M}_T$ by assembling traded payoffs from $\mathcal{M}_T$.

The standard approach to static replication, aside from simple linear interpolation, is to discretize an integral representation of the desired payoff provided by [12]. Assuming that $h(S_T)$ is twice differentiable,[2] Carr and Madan provide the integral representation

$$\begin{aligned} h(S_T) &= h(F_0) + h'(F_0)\cdot(S_T - F_0) \\ &\quad + \int_0^{F_0} h''(K)(K - S_T)^+ dK + \int_{F_0}^{\infty} h''(K)(S_T - K)^+ dK, \end{aligned} \tag{2}$$

which decomposes the payoff in terms of bonds, forwards (with current price $F_0$), and a continuum of calls and puts (see also [11]). To operationalize the integral representation, a discrete approximation is required (see [45] for details). Their method performs reasonably well for smooth payoffs but in practice it is ill-suited for discontinuities, which is one of the drawbacks addressed by our approach.

We develop a method which is similar to Carr and Madan's in that the payoff itself, instead of the valuation operator particular to a given model, is approximated by a discrete set of payoffs in $\mathcal{M}_T$. However, instead of working with market payoffs directly, we selectively "fuse" them together to form a set $\{\Psi_k\}_{k\in\mathcal{K}}$ of more amenable payoff forms.[3] By mandating $\Psi_k \in \mathcal{H} := L^2(\mathbb{R})$ (or $\mathcal{H}^+ := L^2(\mathbb{R}_+)$), we recast the problem of static replication in terms of optimal approximation (in the $L^2$ norm) from the set $\mathcal{M} := \overline{\text{span}}\{\Psi_k\}_{k\in\mathcal{K}}$. A careful design of the payoffs $\{\Psi_k\}_{k\in\mathcal{K}}$ will enable us to then re-express the optimal approximation from $\mathcal{M}$ in terms of the original elements of $\mathcal{M}_T$, thereby producing an optimal hedge in traded payoffs. After developing the theory in $L^2$, section 4.2 discusses the extension to unbounded payoffs using localized projections.

[2]This can be relaxed to "weakly" twice differentiable functions, for which kinks may exist in the payoff, by applying the theory of distributions.

[3]Here, $\mathcal{K}$ is just an arbitrary indexing which will later represent a collection of available strikes.

3.1. **Basis Theory and Frames for Hedging.** We present here only the rudiments of frame and basis theory that are necessary to understand our presentation. Rigorous introductions to the field include [28], [15] and [46].

Our primary objective is to approximate payoffs in $\mathcal{H}$ as linear combinations of the payoffs $\{\Psi_k\}_{k\in\mathcal{K}}$, $f \approx \sum_{k\in\mathcal{K}} \alpha_k \Psi_k$, for some coefficient functionals $\alpha_k = \alpha_k(f)$, which we refer to as *hedge coefficients*. Let $\|\cdot\|_2$ denote the $L^2$ norm. By restricting attention to *Bessel sequences* $\{\Psi_k\}_{k\in\mathcal{K}}$, which for some $B > 0$ satisfy $\sum_{k\in\mathcal{K}} |\langle f, \Psi_k\rangle|^2 \leq B\|f\|_2^2$, $\forall f \in \mathcal{H}$, the approximation $\sum_{k\in\mathcal{K}} \alpha_k \Psi_k$ converges unconditionally[4] for any $\{\alpha_k\}_{k\in\mathcal{K}} \in l^2(\mathcal{K})$. For $\{\Psi_k\}_{k\in\mathcal{K}}$ to admit useful representations, both a lower and upper bound must be satisfied for arbitrary $f \in \overline{\text{span}}\{\Psi_k\}_{k\in\mathcal{K}} := \mathcal{M}$. Specifically, $\{\Psi_k\}_{k\in\mathcal{K}} \subset \mathcal{H}$ is called a *frame sequence* (or a frame for $\mathcal{M}$) if for some dense subset $\widetilde{\mathcal{M}} \subset \mathcal{M}$, it satisfies

$$A\|f\|_2^2 \leq \sum_{k\in\mathcal{K}} |\langle f, \Psi_k\rangle|^2 \leq B\|f\|_2^2, \quad \forall f \in \widetilde{\mathcal{M}},$$

for some $0 < A \leq B$.[5] In fact, every frame sequence defines a bounded linear operator $T : l^2(\mathcal{K}) \to \mathcal{H}$ by $T\{c_k\} = \sum_{k\in\mathcal{K}} c_k \Psi_k$. The adjoint, $T^* : \mathcal{H} \to l^2(\mathcal{K})$ is given by $T^* f = \{\langle f, \Psi_k\rangle\}_{k\in\mathcal{K}}$. Upon composing $T$ with $T^*$, we obtain the *frame operator* $S : \mathcal{M} \to \mathcal{M}$ by $Sf = TT^* f = \sum_{k\in\mathcal{K}} \langle f, \Psi_k\rangle \Psi_k$, which is bounded, invertible and self-adjoint. Furthermore, for any $f \in \mathcal{M}$, $f = \sum_{k\in\mathcal{K}} \langle f, S^{-1}\Psi_k\rangle \Psi_k$, which is called the *frame representation*, and $\{\widetilde{\Psi}_k\}_{k\in\mathcal{K}} := \{S^{-1}\Psi_k\}_{k\in\mathcal{K}}$ the *cannonical dual*. For a general frame sequence, the canonical dual is unique although the representation is not.[6] Moreover, $\{\widetilde{\Psi}_k\}_{k\in\mathcal{K}}$ is also a frame for $\mathcal{M}$.

Frame sequences can be thought of as spanning sets which relax the unique representation requirement of bases. A frame sequence for which $\{c_k\}_{k\in\mathcal{K}} \equiv 0$ is implied whenever $\sum_{k\in\mathcal{K}} c_k \Psi_k = 0$ is called a *Riesz sequence* (or a Riesz basis for $\mathcal{M} = \overline{\text{span}}\{\Psi_k\}_{k\in\mathcal{K}}$).[7] In fact, a Riesz sequence is a (non-orthogonal) basis for its closed span. This implies that $f = \sum_{k\in\mathcal{K}} \langle f, \widetilde{\Psi}_k\rangle \Psi_k$ is the unique representation of any $f \in \mathcal{M}$, where the canonical dual $\{\widetilde{\Psi}_k\}_{k\in\mathcal{K}}$ is now *biorthogonal* to $\{\Psi_k\}_{k\in\mathcal{K}}$: $\langle \Psi_k, \widetilde{\Psi}_m\rangle = \delta_{k,m} = \mathbb{1}_{\{k=m\}}$, for any $k, m \in \mathcal{K}$.

Whenever $\{\Psi_k\}_{k\in\mathcal{K}}$ is a frame or Riesz sequence, the orthogonal projection $P_{\mathcal{M}} : \mathcal{H} \to \mathcal{M}$ of $\mathcal{H}$ onto $\mathcal{M}$ is given by $P_{\mathcal{M}} f = \sum_{k\in\mathcal{K}} \langle f, \widetilde{\Psi}_k\rangle \Psi_k$, $\forall f \in \mathcal{H}$. As for our objective, the $L^2$ optimal static hedge of $f \in \mathcal{H}$ in terms of the available payoffs in $\mathcal{M}$ is given by $P_{\mathcal{M}} f$. When $\{\Psi_k\}_{k\in\mathcal{K}}$ is a frame (so the representation is not necessarily unique), $P_{\mathcal{M}} f$ selects among all $L^2$ optimal hedges the one for which the coefficients are $l^2(\mathcal{K})$ minimal. That is, $\|\{\langle f, \widetilde{\Psi}_k\rangle\}\|_{l^2(\mathcal{K})} \leq \|\{\langle f, c_k\rangle\}\|_{l^2(\mathcal{K})}$ whenever $\sum_{k\in\mathcal{K}} \langle f, \widetilde{\Psi}_k\rangle \Psi_k = \sum_{k\in\mathcal{K}} c_k \Psi_k$, so $P_{\mathcal{M}} f$ uses the fewest assets (in the $l^2$ norm sense).

3.2. **Hedging with Frames of Simple Payoffs.** The simplest procedure for manufacturing a frame sequence is to take $\mathcal{M} = \overline{\text{span}}\{\phi_k\}_{k\in\mathcal{K}}$ where $\phi_k(S_T) := T_k\phi(S_T) = \phi(S_T - k)$ for some fixed payoff $\phi \in \mathcal{H}$, and $T_k$ the translation operator. The payoff $\phi$ is called the *generator* of the frame of translates $\{\phi_k\}_{k\in\mathcal{K}} = \{T_k\phi\}_{k\in\mathcal{K}}$, and the corresponding frame operator $S : \mathcal{M} \to \mathcal{M}$ is given by $Sf = \sum_{k\in\mathcal{K}} \langle f, T_k\phi\rangle T_k\phi$. Although our primary interest is hedging payoffs in $\mathcal{H}_+$, it is easier to modify representations for $\mathcal{H}$ than to frame $\mathcal{H}_+$ directly. Similarly, by allowing (at most) countable frame sequences, we take $\mathcal{K} = \mathbb{Z}$ unless otherwise specified, where non-integer translations will be handled shortly.

For frames of translates, the commuter relations $ST_k = T_k S$ and $S^{-1}T_k = T_k S^{-1}$ for all $k \in \mathbb{Z}$ imply an especially simple characterization of the canonical dual frame:

$$\{\widetilde{T_k\phi}\}_{k\in\mathbb{Z}} = \{S^{-1}T_k\phi\}_{k\in\mathbb{Z}} = \{T_k S^{-1}\phi\}_{k\in\mathbb{Z}} = \{T_k\widetilde{\phi}\}_{k\in\mathbb{Z}} := \{\widetilde{\phi}_k\}_{k\in\mathbb{Z}},$$

where $\widetilde{\phi} := S^{-1}\phi$ is the canonical *dual generator* corresponding to $\phi$.

A further generalization is to generate sequences of translated *dilations* of a single generator $\phi \in \mathcal{H}$ at a modified scale or *resolution*. When implementing a hedge in practice, the resolution

[4]Irrespective of how the $\Psi_k$ are ordered.

[5]If these bounds hold $\forall f \in \mathcal{H}$ (or a dense subset thereof), $\{\Psi_k\}_{k\in\mathcal{K}}$ is called a *frame*. In particular, a frame is complete for $\mathcal{H}$, $\overline{\text{span}}\{\Psi_k\}_{k\in\mathcal{K}} = \mathcal{H}$.

[6]There may be other coefficients $\alpha_k$ for which $f = \sum_{k\in\mathcal{K}} \alpha_k \Psi_k$, for $f \in \mathcal{M}$.

[7]This criterion, which is a stronger form of linear independence, is known as $\omega$-independence.

is fixed according to available market payoffs. However, to facilitate the theoretical development of frame based pricing and hedging, it becomes advantageous to partition the domain space dyadically in the following manner, consistent with the approach taken in the frame (and wavelet) literature.

3.2.1. *Frame Multiresolution Analysis.* Starting with a frame generator $\phi \in \mathcal{H}$, consider first the space $U_0 := \overline{\text{span}}\{T_k\phi\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{\phi_{0,k}\}_{k\in\mathbb{Z}}$, where $\phi_{0,k}(S_T) := \phi(S_T - k)$. To transition between approximations at different scales we define the operator $D$ by $Df(x) = 2^{1/2}f(2x)$. By composing $D$ with itself $j$ times, it follows that $D^j f(x) = 2^{j/2}f(2^j x)$. By slicing the translation granularity in half, the next finer space is $U_1 := D\overline{\text{span}}\{T_k\phi\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{\phi_{1,k}\}_{k\in\mathbb{Z}}$, where $\phi_{1,k}(S_T) := 2^{1/2}\phi(2S_T - k)$. Proceeding similarly we obtain

$$U_j := D^j\overline{\text{span}}\{T_k\phi\}_{k\in\mathbb{Z}} = \overline{\text{span}}\{\phi_{j,k}\}_{k\in\mathbb{Z}}, \tag{3}$$

where $\phi_{j,k}(S_T) := 2^{j/2}\phi(2^j S_T - k) = D^j T_k\phi(S_T)$.

As long as $\phi$ has been appropriately selected, with selection criteria provided shortly, any function in $\mathcal{H}$ can be approximated arbitrarily well for some fixed resolution level $j$ by forming linear combinations of $\phi_{j,k}$ as $k$ varies over $\mathbb{Z}$. Appropriately chosen generators $\phi$ induce a special structure on the space $\mathcal{H}$.

**Definition 3.1. (Frame Multiresolution Analysis)** A function $\phi \in \mathcal{H}$ which generates a frame sequence $\{T_k\phi\}_{k\in\mathbb{Z}}$ is said to generate a *frame multiresolution analysis* (FMRA)[8] if the spaces $\{U_j\}_{j\in\mathbb{Z}}$ defined in (3) satisfy

(i) $\cdots U_{-1} \subset U_0 \subset U_1 \cdots$, (ii) $\overline{\cup_j U_j} = \mathcal{H}$.

At each resolution, the canonical dual is given by $\{T_{2^{-j}k}\theta_j\}_{k\in\mathbb{Z}}$, where $\theta_j := S^{-1}D^j\phi$. For $f \in \mathcal{H}$, an approximation in terms of the scale $j$ frame sequence is given by the orthogonal projection operator

$$P_j f(S_T) = \sum_{k\in\mathbb{Z}} \langle f, T_{2^{-j}k}\theta_j\rangle \phi_{j,k}(S_T).$$

**Proposition 3.1.** *If $\phi$ generates an FMRA, then for each $j \in \mathbb{Z}$ the following hold:*

*(i) $\{\phi_{j,k}\}_{k\in\mathbb{Z}}$ is a frame for $U_j = \overline{span}\{\phi_{j,k}\}_{k\in\mathbb{Z}}$.*
*(ii) $\forall f \in \mathcal{H}$, $\|P_j f\|_2 \le \|f\|_2$ and $\lim_{j\to\infty}\|P_j f - f\|_2 = 0$.*

In summary, if we start with a payoff $\phi(S_T)$ which generates an FMRA, then scaled and shifted versions of $\phi$ enable an approximate hedge of any $f \in \mathcal{H}$ which approaches (in $L^2$) the true payoff as the support of $f$ is partitioned into a finer strike space. The next step is to price $f$ by pricing elements of the frame. Once frame elements have been priced, then at any fixed resolution, the pricing of arbitrary $f \in \mathcal{H}$ is reduced to the computation of its hedge coefficients.

3.3. **The Pricing Functional.** Consider the case of positive underlying asset prices, $S_T = S_0 e^{X_T}$, where $X_T$ is a process with known characteristic function $\widehat{\mu}_T(\xi)$.[9] We denote the density of $\ln(S_T)$ by $q_T$ (with characteristic function $\widehat{q}_T(\xi)$), and note that $\ln(S_T) = \ln(S_0) + X_T$, where $S_0$ is assumed to be known at the time of pricing. For a given risk neutral density $q_T$ (with $S_0$ fixed), define the *pricing functional* $\mathcal{V} : \mathcal{H} \to \mathbb{C}$ by

$$\mathcal{V}(f) = e^{-rT}\int_{\mathbb{R}} f(e^y)q_T(y)dy = e^{-rT}\int_{\mathbb{R}} f(S_0 e^x)\mu_T(x)dx.$$

A mild restriction on $q_T$ enables the pricing of any $f \in \mathcal{H}$.

**Theorem 3.1.** *Suppose $q_T$ is essentially bounded ($q_T \in L^\infty(\mathbb{R})$), and $|\widehat{q}_T(i)| < \infty$. Then the pricing functional satisfies $\mathcal{V} \in \mathcal{H}^*$, i.e $\mathcal{V}$ is a bounded linear functional over $\mathcal{H}$.*

[8]The reader should be aware that this definition is significantly shorter than equivalent ones given in most (especially older) treatments of FMRAs. See [15] chapter 13 for details.

[9]Note that $\widehat{\mu}_T(\xi)$ may depend on additional model parameters. The transform here is with respect to a single source of randomness, $X_T$.

For example, if $\widehat{q}_T \in L^1(\mathbb{R})$, it follows that for every continuity point of $\widehat{q}_T(x)$, $|q_T(\xi)| = \frac{1}{2\pi}|\int_{\mathbb{R}} e^{-ix\xi}\widehat{q}_T(x)dx| \leq \frac{1}{2\pi}\|\widehat{q}_T\|_1$, so $q_T \in L^\infty(\mathbb{R})$. In terms of a frame sequence $\{\phi_{j,k}\}_{k\in\mathbb{Z}}$ with dual $\{T_{2^{-j}k}\theta_j\}_{k\in\mathbb{Z}}$, the approximate pricing functional at resolution $j$ can be defined as

$$\mathcal{V}_j := \mathcal{V}\circ P_j, \quad \mathcal{V}_j f = e^{-rT}\int_{\mathbb{R}}\Big(\sum_k \langle f, T_{2^{-j}k}\theta_j\rangle \phi_{j,k}(e^y)\Big)q_T(y)dy.$$

**Theorem 3.2.** *Suppose the conditions of Theorem 3.1 hold, and let $\phi$ be a frame generator with $\theta_j := S^{-1}D_j\phi$. Then for each $j \in \mathbb{N}$, $\mathcal{V}_j : \mathcal{H} \to \mathbb{C}$ is a bounded linear functional satisfying*

$$\mathcal{V}_j f = \sum_k \langle f, T_{2^{-j}k}\theta_j\rangle \mathcal{V}\circ\phi_{j,k}, \qquad f \in \mathcal{H}.$$

*Moreover, if $\phi$ generates a FMRA, $\mathcal{V}_j \to \mathcal{V}$ uniformly in $\mathcal{H}$. That is,*

$$\lim_{j\to\infty}\sup_{f\in\mathcal{H},\|f\|_2=1}|\mathcal{V}_j f - \mathcal{V}f| = 0.$$

In particular, frame pricing is a (uniformly) consistent approximation to the pricing problem. Moreover, payoff representations are valid independently of the underlying dynamics, as long as $|\widehat{q}_T(i)| < \infty$, since any such pricing operator $\mathcal{V}$ yields a consistent approximation in terms of the frame elements. Likewise, admissible payoffs are specified independently of $\mathcal{V}$, and so can be used from one model to the next.

3.4. **Pricing and Hedging with Riesz Bases.** While frames are certainly sufficient to guarantee the representation properties we desire, and they provide enough structure to conduct fruitful analysis, the elegant dual structure of Riesz bases greatly simplifies their implementation.[10] Recall that for $\phi \in \mathcal{H}$, $\{T_k\phi\}_{k\in\mathbb{Z}}$ is a Riesz basis for $\mathcal{M} := \overline{\text{span}}\{T_k\phi\}_{k\in\mathbb{Z}}$ if

(1) $\exists\widetilde{\phi} \in \mathcal{M}$ such that $\{T_k\widetilde{\phi}\}_{k\in\mathbb{Z}}$ is biorthogonal to $\{T_k\phi\}_{k\in\mathbb{Z}}$, and
(2) $\exists B \geq A > 0$ such that $A\|f\|_2^2 \leq \sum_{k\in\mathbb{Z}}|\langle f, T_k\phi\rangle|^2 \leq B\|f\|_2^2$, $\forall f \in \mathcal{M}$.

By replacing the frame condition with the requirement that $\phi$ generates a Riesz sequence which satisfies (i) and (ii) of definition 3.1 for the spaces $\{U_j\}_{j\in\mathbb{Z}}$ defined in (3), we have what is called a *Riesz multiresolution analysis* (RMRA). For $f \in \mathcal{H}$, an approximation in terms of the scale $j$ Riesz sequence is given by the orthogonal projection operator

$$P_j f(S_T) = \sum_{k\in\mathbb{Z}}\langle f, \widetilde{\phi}_{j,k}\rangle\phi_{j,k}(S_T).$$

Similarly, to restrict attention to payoffs $f \in \mathcal{H}_+$, we define the $\mathcal{H}_+$-projection $P_{j^+} : \mathcal{H}_+ \to U_{j^+} := \overline{\text{span}}\{\phi_{j,k}\}_{k\geq 0}$ as $P_{j^+}f(S_T) = \sum_{k\geq 0}{}'\langle f, \widetilde{\phi}_{j,k}\rangle\phi_{j,k}(S_T)$, where $\sum'$ indicates that each $\phi_{j,k}$ has been restricted to $\mathbb{R}_+$ ($\phi_{j,k}(x)\mathbb{1}_{x\geq 0}$). By theorem 3.2, the approximate pricing functional at resolution $j$ can be defined in the RMRA case as $\mathcal{V}_j := \mathcal{V}\circ P_j$, $\mathcal{V}_j f = \sum_{k\in\mathbb{Z}}\langle f, \widetilde{\phi}_{j,k}\rangle\mathcal{V}\circ\phi_{j,k}$, for $f \in \mathcal{H}$. The value convergence can also be characterized formally by following:

**Corollary 3.3.** *If $\phi \in \mathcal{H}$ generates a Riesz sequence, then for each $j \in \mathbb{Z}$ the following hold:*

*(i) $\{\phi_{j,k}\}_{k\in\mathbb{Z}}$ is a Riesz basis for $U_j = \overline{span}\{\phi_{j,k}\}_{k\in\mathbb{Z}}$.*
*(ii) If $\phi$ generates an RMRA, then $\forall f \in \mathcal{H}$, $\|P_j f\|_2 \leq \|f\|_2$ and $\lim_{j\to\infty}\|P_j f - f\|_2 = 0$. Moreover, $\mathcal{V}_j \to \mathcal{V}$, uniformly in $f \in \mathcal{H}$.*
*(iii) For $f \in \mathcal{H}_+$, $\lim_{j^+\to\infty}\|P_{j^+}f - f\|_2 = 0$, and $\lim_{j^+\to\infty}|\mathcal{V}\circ P_{j^+}f - \mathcal{V}f| = 0$.*

When the basis elements (or their components) are priced in a market, valuation is reduced to an observation of market prices.

[10] In particular, the butterfly basis discussed in section 4 is a special class of FMRA in which the the generator $\phi$ is compactly supported and generates a Riesz sequence of translates.

3.4.1. *Integral Representation for Riesz Basis Projection.* A closed-form integral representation of the orthogonal projection coefficients for Riesz bases is now derived. Consistent with the financial literature, we define the FT of an $L^1$ or $L^2$ function as

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}} e^{i\xi x} f(x)dx.$$

If $\phi$ generates a Riesz sequence, then $\widetilde{\phi} := S^{-1}\phi$ is given by

$$\widehat{\widetilde{\phi}}(\xi) = \frac{\widehat{\phi}(\xi)}{\mathbf{\Phi}(\xi)}, \quad \text{where} \quad \mathbf{\Phi}(\xi) = \sum_{k\in\mathbb{Z}} \left|\widehat{\phi}(\xi + 2\pi k)\right|^2.$$

In fact, $\{T_k\phi\}_{k\in\mathbb{Z}}$ is a Riesz (resp. frame) sequence iff $A \leq \mathbf{\Phi}(\xi) \leq B$ for almost every $\xi \in [0, 2\pi]$ (resp. $\xi \in \mathcal{N}$, where $\mathcal{N} := [0, 2\pi]/\{\xi : \mathbf{\Phi}(\xi) = 0\}$) for some $A, B > 0$. For a frame, the dual generator is given by $\widehat{\phi}(\xi)/\mathbf{\Phi}(\xi) \cdot \mathbb{1}_{\mathcal{N}}(\xi)$. As outlined in the appendix, one way to determine $\mathbf{\Phi}$ is by Fourier series expansion.

To accommodate the resolution $a > 0$ found in practical settings (implied by the spacing between option strikes), we define the dilation operator $D_a$ which acts on $f \in \mathcal{H}$ according to $D_a f(x) = |a|^{1/2} f(ax)$. We now denote by $P_{\mathcal{M}_a} f = \sum_{k\in\mathbb{Z}} \langle f, \widetilde{\phi}_{a,k}\rangle \phi_{a,k}$ the projection of $f \in \mathcal{H}$ onto $\mathcal{M}_a := \overline{\text{span}}\{\phi_{a,k}\}_{k\in\mathbb{Z}}$, where $\phi_{a,k}(x) := D_a T_k \phi = a^{1/2}\phi(ax - k)$. An advantage of Riesz bases over general frames is that the dual of $\{D_a T_k \phi\}$ is simply $\{D_a T_k \widetilde{\phi}\}$ for $\widetilde{\phi}$ defined above, which yields the following.

**Theorem 3.4.** *Let $\{T_k\phi\}_{k\in\mathbb{Z}}$ be a Riesz sequence in $\mathcal{H}$ with real-valued generator $\phi$ and canonical dual $\widetilde{\phi}$. Further, define $\mathcal{M}_a = \overline{span}\{\phi_{a,k}\}_{k\in\mathbb{Z}}$, where $\phi_{a,k} := D_a T_k \phi$. The following hold:*

*(i) $\widetilde{\phi}_{a,k} := D_a T_k \widetilde{\phi}$ is the unique biorthogonal dual Riesz basis in $\mathcal{M}_a$ corresponding to $\{\phi_{a,k}\}_{k\in\mathbb{Z}}$.*
*(ii) For any real-valued $f \in \mathcal{H}$ the projection of $f$ onto $\mathcal{M}_a$ is given by*

$$P_{\mathcal{M}_a} f(S_T) = \sum_{k\in\mathbb{Z}} \beta_{a,k} \cdot \phi_{a,k}(S_T),$$

$$\beta_{a,k} := \frac{a^{1/2}}{2\pi} \int_{\mathbb{R}} e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi) d\xi = \frac{a^{1/2}}{\pi} \int_0^\infty \Re\left[ e^{-ik\xi} \frac{\widehat{f}(a\xi)\widehat{\phi}(-\xi)}{\mathbf{\Phi}(-\xi)} \right] d\xi.$$

Computationally speaking, in order to switch between scales a Riesz sequence is preferred since a single calculation of $\widetilde{\phi}$ reveals the dual at all scales.

3.4.2. *Error Bounds for Riesz Bases.* This section analyzes the convergence behavior for a particular class of Riesz basis, and for functions of sufficient regularity. We denote by $\mathcal{W}_2^m$ the Sobolev space of functions whose first $m$ (weak) derivatives are defined in $\mathcal{H} = L^2$. Similarly, denote by $\mathcal{W}_\infty^m$ the Sobolev space of functions with bounded (weak) derivatives up to order $m$, where

$$\|f\|_\infty = \lim_{p\to\infty} \left( \int_{-\infty}^{\infty} |f(x)|^p dx \right)^{1/p} = \sup_{x\in\mathbb{R}} |f(x)|.$$

For a Riesz sequence generated by $\phi$, the convergence rate of approximations from $\mathcal{M}_a$ to $f \in \mathcal{H}$ of the form

$$f_a(S_T) := a^{1/2} \sum_{k\in\mathbb{Z}} \vartheta_{a,k} \phi(aS_T - k), \quad \{\vartheta_{a,k}\}_{k\in\mathbb{Z}} \in l_2(\mathbb{Z}),$$

can be characterized in terms of $\widehat{\phi}$ and its derivatives. In particular, a Riesz generator $\phi$ is called an $m^{th}$ *order Riesz generator* if

$$\widehat{\phi}(0) = 1, \quad \text{and for } q = 0, ..., m-1, \quad \widehat{\phi}^{(q)}(2\pi k) = 0, \quad k \in \mathbb{Z}/\{0\}, \tag{4}$$

where $\widehat{\phi}^{(q)}$ denotes the $q$th derivative of $\phi$. The analysis of [44] establishes hedging error bounds and the speed of convergence with respect to $a$. Namely, if $(\phi, \widetilde{\phi})$ form a $m^{th}$ order Riesz generator/dual pair, then the projection error for $f \in \mathcal{W}_2^m$ satisfies

$$\inf_{f_a\in\mathcal{M}_a} \|f - f_a\|_2 \leq \|f - P_{\mathcal{M}_a} f\|_2 \leq C(\phi) a^{-m} \|f^{(m)}\|_2,$$

where $C(\phi)$ is independent of $f$. From the perspective of hedging, in which mesh refinement is limited, [43] observes that coarse scale least-squares approximations behave like an interpolation but with twice the order of accuracy. In particular, [43] finds that the projection error for $f \in \mathcal{W}_2^m$ satisfies

$$\|f - P_{\mathcal{M}_a} f\|_2 \leq B_{2,m}(\phi)\|f^{(2m)}\|_2 \cdot a^{-2m} + \sqrt{B_{2,m}(\phi)}\|f^{(m)}\|_2 \cdot a^{-m}, \tag{5}$$

for some $B_{2,m}(\phi)$ independent of $f$. While asymptotically the error decays as $\mathcal{O}(a^{-m})$, it exhibits $\mathcal{O}(a^{-2m})$ decay for coarse resolutions. Moreover, the swift initial decay translates into accelerated valuation algorithms when applied to projected payoffs.

We can also characterize the decay in terms of the $L^\infty$ norm in the Sobolev space $\mathcal{W}_\infty^m$. In general, the orthogonal projection operator $P_{\mathcal{M}_a}$ can be expressed as

$$P_{\mathcal{M}_a} f(x) = \int_{\mathbb{R}} f(y) a \cdot K(ax, ay) dy, \quad f \in \mathcal{H},$$

in terms of the reproducing kernel

$$K(x, y) = \sum_{k \in \mathbb{Z}} \phi(x - k)\widetilde{\phi}(y - k).$$

From this perspective, condition (4) is equivalent to

$$\int_{-\infty}^{\infty} K(x, y) dy = 1, \quad \int_{-\infty}^{\infty} (y - x)^q K(x, y) dy = 0, \quad q = 1, \ldots, m - 1. \tag{6}$$

From Proposition 3.3 of [43], for $\phi$ satisfying (6), it holds that

$$\|f - P_{\mathcal{M}_a} f\|_\infty \leq B_{\infty,m}(\phi)\|f^{(m)}\|_\infty \cdot a^{-m}, \quad \forall f \in W_\infty^m,$$

where

$$B_{\infty,m}(\phi) := \frac{1}{m!} \sup_x \left[ \int_{-\infty}^{\infty} |x - y|^m |K(x, y)| dy \right].$$

In particular, for functions of sufficient regularity, convergence is attained in the $L^2$ and $L^\infty$ norms.

## 4. The Butterfly Basis

Given the current state of financial markets, we study in detail a surprisingly effective method of payoff replication which utilizes the compactly supported, real-valued payoff generator $\varphi$:

$$\varphi(S_T) = (1 - |S_T|)^+ = (1 - |S_T|)\mathbb{1}_{[-1,1]}(S_T).$$

For a fixed $a > 0$, we define the scale-$a$ *butterfly basis*[11] as the sequence $\varphi_{a,k} := T_{\frac{k}{a}} D_a \varphi \equiv D_a T_k \varphi$ given by:

$$\begin{aligned} \varphi_{a,k}(S_T) &= a^{1/2} \left(1 - |a\left(S_T - k/a\right)|\right) \mathbb{1}_{[-\frac{1}{a}, \frac{1}{a}]}\left(S_T - k/a\right) \\ &= a^{1/2}(1 - |aS_T - k|)\mathbb{1}_{[\frac{k-1}{a}, \frac{k+1}{a}]}(S_T), \end{aligned}$$

where $1/a$ represents the step size of a uniform spacing of strikes. For short-term S&P 500 index options, a value of $a = 1/5$ provides a five dollar spacing between strikes, at least near the forward price. From [15]( p.289), it is easily verified that the butterfly basis on $\mathcal{H}$ with generator $\varphi(S_T) = (1 - |S_T|)\mathbb{I}_{[-1,1]}(S_T)$ generates an RMRA.

In most markets, asset prices (and the support of any contingent payoffs) are assumed to be positive. In this case, $\varphi_{a,0}$ is easily adjusted by truncating its support, while $\varphi_{a,k}$ for $k \geq 1$ is supported over $[\frac{k-1}{a}, \frac{k+1}{a}] \subset \mathbb{R}_+$. Specifically, in $\mathcal{H}_+$ we define the left boundary put

$$\varphi_{a,0}(S_T) = a^{1/2}(1 - aS_T)\mathbb{1}_{[0, \frac{1}{a}]}(S_T) = a^{3/2}\Big(\frac{1}{a} - S_T\Big)^+,$$

[11]Readers may recognize $\varphi$ as the hat, tent, or linear B-spline scaling function, depending on the context. Our terminology is a reference to the butterfly spread commonly used in option markets.

and we will represent arbitrary payoffs in $\mathcal{H}_+$ by linear combinations of butterfly basis elements:

$$f(S_T) \approx \beta_0 \varphi_{a,0}(S_T) + \sum_{k\geq 1} \beta_k \varphi_{a,k}(S_T), \tag{7}$$

for a set of hedging coefficients $\{\beta_{a,k}\}$ to be determined.When hedging $f \in \mathcal{H}_+$ with payoffs from $\mathcal{M}_a^+ := \overline{\text{span}}\{\varphi_{a,k}\}_{k\geq 0}$, we define the $\mathcal{H}_+$-projection $P_{\mathcal{M}_a^+}$ (analogous to $P_{j^+}$ from section 3.4) by starting with the true orthogonal projection $P_{\mathcal{M}_a}$ of $f$ onto $\mathcal{M}_a := \overline{\text{span}}\{\varphi_{a,k}\}_{k\in\mathbb{Z}}$ and truncating the basis to to $\varphi_{a,0} \cup \{\varphi_{a,k}\}_{k\geq 1}$. We use $\sum'$ to denote a series taken with respect to the truncated basis elements.

**Corollary 4.1.** *The butterfly basis, with generator $\varphi$ and scale parameter $a > 0$ fixed, is a Riesz basis for its closed span on $\mathcal{H}$. Furthermore, the $\mathcal{H}_+$-projection of $f \in \mathcal{H}_+$ onto $\mathcal{M}_a^+$ is given by*

$$P_{\mathcal{M}_a^+} f(S_T) = \sum_{k\geq 0}{}' \langle f, \widetilde{\varphi}_{a,k}\rangle \varphi_{a,k}(S_T) = \beta_{a,0}\varphi_{a,0}(S_T) + \sum_{k\geq 1} \beta_{a,k}\varphi_{a,k}(S_T),$$

*where for $k \geq 1$*

$$\beta_{a,k} = \frac{12a^{1/2}}{\pi} \Re\left[\int_0^\infty e^{-ik\xi} \widehat{f}(a\xi) \frac{\sin^2(\xi/2)}{\xi^2(2+\cos(\xi))} d\xi\right]. \tag{8}$$

Since the Butterfly generator $\varphi$ is a second order Riesz generator, the following error bound of hedging the payoff $f$ with the butterfly basis is obtained by applying equation (5),

$$\|f - P_{\mathcal{M}_a} f\|_2 \leq K_4(\varphi)\|f^{(4)}\|_2 \cdot a^{-4} + \sqrt{K_4(\varphi)}\|f^{(2)}\|_2 \cdot a^{-2}, \tag{9}$$

for a constant $K_4(\varphi)$, independent of $f$.

*Remark* 1. Note that for $f \in \mathcal{H}_+$, with the natural extension to $f \in \mathcal{H}$ by $f(x) = 0$ for $x < 0$, we have

$$\|f - P_{\mathcal{M}_a} f\|_2^2 = \int_{-\infty}^0 \left(P_{\mathcal{M}_a} f(x)\right)^2 dx + \int_0^\infty \left(f(x) - P_{\mathcal{M}_a^+} f(x)\right)^2 dx$$

since $P_{\mathcal{M}_a} f(x) = P_{\mathcal{M}_a^+} f(x)$ for $x \geq 0$. Hence, $\|f - P_{\mathcal{M}_a^+} f\|_2 \leq \|f - P_{\mathcal{M}_a} f\|_2$ for $f \in \mathcal{H}_+$, so error bounds given in terms $P_{\mathcal{M}_a} f$ (e.g. (9)) will also hold for $P_{\mathcal{M}_a^+} f$.

*Remark* 2. With the butterfly basis, upon replacing $\|f\|_2$ with $\|P_{\mathcal{M}_a^+} f - f\|_2$ in the proof of Theorem 3.1, and combining with the norm convergence in equation (9), we have for some $\zeta > 0$ that

$$|\mathcal{V} \circ P_{\mathcal{M}_a^+} f - \mathcal{V} f| \leq \zeta \|P_{\mathcal{M}_a^+} f - f\|_2 = \mathcal{O}(a^{-2}),$$

where $\mathcal{O}(a^{-2})$ holds for twice continuously differentiable payoffs. Hence, the convergence in option prices is governed by the projection convergence. As illustrated in the numerical pricing experiments, prices converge even faster than predicted.

4.1. **The Dual Butterfly Basis.** By expanding the dual generator $\widetilde{\phi}$ itself in the basis generated by $\phi$, we derive an analytical procedure for approximating hedging coefficients which can be calculated from $f(S_T)$ itself, rather than its Fourier transform. It also leads in some cases to analytical formulas for payoff projections, in which case the associated cost is reduced dramatically. Moreover, we will use this method as a control for the development of efficient approximations to the true projection.

**Corollary 4.2.** *Let $\phi \in \mathcal{H}$ be a symmetric real-valued Riesz generator with canonical dual generator $\widetilde{\phi}$. Then*

$$\widetilde{\phi} = \sum_{m\in\mathbb{Z}} \alpha_m T_m \phi, \quad \textit{where} \quad \alpha_m = \frac{1}{\pi}\int_0^\infty \cos(m\xi) \frac{\widehat{\phi}^2(\xi)}{\mathbf{\Phi}^2(\xi)} d\xi.$$

We can apply Corollary 4.2 to generate the dual of general Riesz generators when more explicit descriptions are unavailable. It also gives a procedure to check the validity of alternative

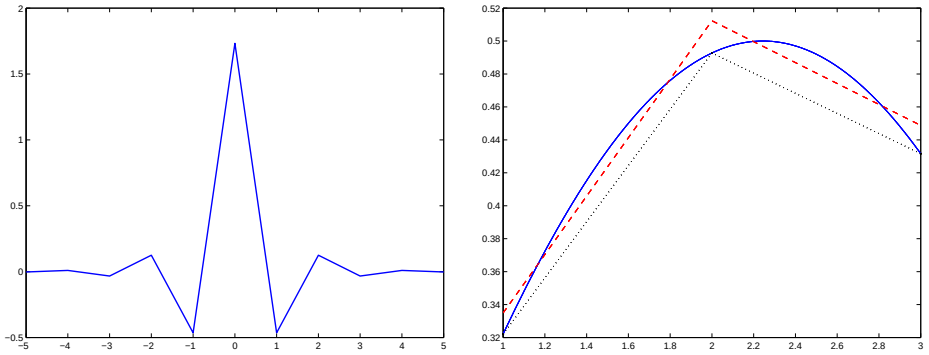

FIGURE 1. Left: cannonical dual, $\widetilde{\varphi}$. Right: comparison of dual method (dashed line) vs. linear interpolation (dotted line) for a continuous payoff (solid line).

descriptions as we show next for the butterfly basis. Applying Corollary 4.2 to the butterfly basis generator $\varphi$, we conclude that $\widetilde{\varphi}$ has coefficients

$$\alpha_m = \frac{36}{\pi}\int_0^\infty \cos(xm)\frac{(1-\cos(x))^2}{x^4(2+\cos(x))^2}dx,$$

the first five of which are given by

$$\alpha_0 = 1.73205,\ \alpha_1 = -0.46410,\ \alpha_2 = 0.12436,\ \alpha_3 = -0.03332,\ \alpha_4 = 0.00893.$$

In fact, for the butterfly basis we can obtain a much cleaner description of the dual using biorthogonality in a different way, which can be verified using the previous Corollary.

**Proposition 4.1.** *The coefficients $\alpha_m$ of $\widetilde{\varphi}$ with respect to the linear basis converge to zero exponentially in $m$. In particular,*

$$\widetilde{\varphi}(x) = \sum_{m\in\mathbb{Z}}\left(\frac{3}{\sqrt{3}}\left(\sqrt{3}-2\right)^{|m|}\right)\varphi(x-m). \tag{10}$$

Hence, the dual butterfly generator, illustrated in Figure 2, is a piecewise continuous function which is well approximated over a small, compact interval. In general, we can write

$$\begin{aligned} P_{\mathcal{M}_a^+}f(S_T) &= {\sum_{k\geq 0}}'\langle f,\widetilde{\varphi}_{a,k}\rangle\varphi_{a,k}(S_T) = {\sum_{k\geq 0}}'\langle f, D_aT_k\sum_{m\in Z}\alpha_m T_m\varphi\rangle\varphi_{a,k}(S_T)\\ &= {\sum_{k\geq 0}}'\Big(\sum_{m\in\mathbb{Z}}\alpha_m\langle f,\varphi_{a,k+m}\rangle\Big)\varphi_{a,k}(S_T). \end{aligned} \tag{11}$$

The hedge coefficients are given by $\beta_k = \sum_{m\in\mathbb{Z}}\alpha_m\theta_{a,k-m}$ where we define

$$\theta_{a,k} := \langle f,\varphi_{a,k}\rangle = \int_{[\frac{k-1}{a},\frac{k+1}{a}]} f(s)\varphi_{a,k}(s)ds. \tag{12}$$

Compared with (8), this provides an alternative description of $\beta_{a,k}$ when the Fourier transform $\hat{f}(\xi)$ is unknown. We discuss an approximation to $\beta_{a,k}$ in section 5.1 below.

4.2. **Localized Projections.** Given a complete description of the dual in section 4.1, we are able to characterize the classes of payoff functions for which butterfly projection methods apply, relaxing the ostensible requirement that $f$ belongs to $L^2(\mathbb{R})$. As a first example, consider the function $f(x) = x$ for $x \in \mathbb{R}$, which is clearly not in $L^2(\mathbb{R})$, and fix any interval $[L,R] = [k_L/a, k_R/a]$. Then with projection coefficients $\beta_{a,k} = \langle f,\widetilde{\varphi}_{a,k}\rangle = a^{-3/2}k$ (derived in section 4.3),

$$f(x) - \sum_{k_L\leq k\leq k_R}\beta_{a,k}\varphi_{a,k}(x) \equiv 0, \qquad \forall x\in[L,R]. \tag{13}$$

Hence, the "projection", localized to $[L,R]$, is a perfect local representation of the payoff. This localized projection can be used to form a meaningful approximation to a large class of payoffs.

We define two such classes of $L^2_{loc} := \{f : f \in L^2([L,R]), \forall L < R \in \mathbb{R}\}$, the first is given by

$$\bar{C} := \{f \in L^2_{loc} : \forall a > 0, \|f\|_2^{I^a_{m^+}} \leq C\|f\|_2^{I^a_m} \text{ eventually, where } |C\nu| < 1\},$$

where *eventually* is defined as the existence of $M$ (which can depend on $a$) such that $\forall |m| \geq M$ the bound holds, $I^a_m := [(m-1)/a, (m+1)/a]$, and $I^a_{m^+} := I^a_{m+1\cdot\text{sign}(m)}$. The second class, $\bar{C}(a)$, allows some threshold $\bar{a} > 0$ after which the bound holds for $a \geq \bar{a}$:

$$\bar{C}(a) := \{f \in L^2_{loc} : \exists \bar{a} > 0 \text{ s.th } \forall a \geq \bar{a},\ \|f\|_2^{I^a_{m^+}} \leq C\|f\|_2^{I^a_m} \text{ eventually, where } |C\nu| < 1\},$$

where $\nu := \sqrt{3}-2$ is the constant in Proposition 4.1. For functions in either class, the projection coefficients are finite, and as we now show, can be approximated arbitrarily closely using only local knowledge of $f$ (as required by the approximations developed in section 5). This further allows us to extend the idea of projection to non-$L^2$ payoffs.

Given a function $f$ in either class, with $[L,R] = [k_L/a, k_R/a]$ a fixed hedge interval, we define for $[k_{\bar{L}}/a, k_{\bar{R}}/a] = [\bar{L}, \bar{R}] \supset [L,R]$ the *localized projections*

$$\bar{P}_{\mathcal{M}_a} f(S_T) = \sum_{k_L \leq k \leq k_R} \beta_{a,k}\varphi_{a,k}(S_T), \qquad \beta_{a,k} = \langle f, \widetilde{\varphi}_{a,k}\rangle, \tag{14}$$

$$\widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}(S_T) = \sum_{k_L \leq k \leq k_R} \widehat{\beta}_{a,k}\varphi_{a,k}(S_T), \qquad \widehat{\beta}_{a,k} = \langle f\mathbb{1}_{[\bar{L},\bar{R}]}, \widetilde{\varphi}_{a,k}\rangle. \tag{15}$$

Here $\bar{P}_{\mathcal{M}_a} f$ is a localized projection which uses global information about $f$, while $\widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}$ uses only local information about $f$ to determine coefficients. The choice of $[\bar{L},\bar{R}]$ is discussed in section 5. For localized projection to be meaningful, $\widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}$ should be a consistent approximation to $\bar{P}_{\mathcal{M}_a} f$, which is indeed the case.

**Proposition 4.2.** *If $f \in \bar{C}$, $a > 0$ and $[L,R]$ is fixed, then $\forall \epsilon > 0$, $\exists [\bar{L},\bar{R}]$ such that*

$$\|\bar{P}_{\mathcal{M}_a} f - \widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}\|_2 < \epsilon.$$

*Similarly, if $f \in \bar{C}(a)$, $\exists \bar{a} > 0$ such that $\forall a > \bar{a}$ fixed and $\forall \epsilon > 0$, $\exists [\bar{L},\bar{R}]$ such that the bound holds. Moreover, both projections are finite. In particular, if $f \in \bar{C}$ then $f \cdot \widetilde{\varphi}_{a,k} \in L^2(\mathbb{R})$, $\forall a > 0, k \in \mathbb{Z}$, and if $f \in \bar{C}(a)$ then $\exists \bar{a} \geq 0$ such that $f \cdot \widetilde{\varphi}_{a,k} \in L^2(\mathbb{R})$, $\forall a \geq \bar{a}, k \in \mathbb{Z}$.*

*Remark* 3. As an example of the first class, let $f(x) := x^p$ and fix $m \geq 1$ and $a > 0$ (the argument for $m \leq -1$ is analogous). Then

$$\left(\|f\|_2^{I^a_{m^+}}\right)^2 = \int_{I^a_{m^+}} x^{2p}dx = \int_{I^a_m} \left(x + \frac{1}{a}\right)^{2p} dx = \int_{I^a_m} \left(x^{2p} + \sum_{k=0}^{2p-1} \binom{2p}{k} x^k \left(\frac{1}{a}\right)^{2p-k}\right) dx.$$

The first term in the rightmost integral is $\left(\|f\|_2^{I^a_m}\right)^2$, and by a choice of $M$ sufficiently large the second term can be dominated by $\epsilon \cdot \left(\|f\|_2^{I^a_m}\right)^2$ for some $\epsilon \in [0,1)$ and all $m \geq M$, since the highest power in $x$ is $2p-1$. Hence $\|f\|_2^{I^a_{m^+}} \leq \sqrt{1+\epsilon}\|f\|_2^{I^a_m}$. By our choice of $M$, we can ensure that $\nu\sqrt{1+\epsilon} < 1$, so $x^p \in \bar{C}$. From Proposition 4.2, we have that $x^p \cdot \widetilde{\varphi}_{a,k} \in L^2(\mathbb{R})$, so $f \cdot \widetilde{\varphi}_{a,k} \in L^2(\mathbb{R})$ for any polynomial $f(x) = \sum_{k=0}^p c_k x^k$ since $L^2(\mathbb{R})$ is a vector space. As an example of the second class, let $f(x) := e^{\tau x}$, $\tau > 0$. Then by a change of variable

$$\|f\|_2^{I^a_{m^+}} = \exp(\tau/a)\|f\|_2^{I^a_m},$$

and $a > 0$ can be chosen so that $|\exp(\tau/a)\nu| < 1$. Hence functions with exponential growth belong to $\bar{C}(a)$.

From Proposition 4.2, for $f \in \bar{C}$, and for $f \in \bar{C}(a)$ with $a$ sufficiently large, $\bar{P}_{\mathcal{M}_a} f$ is the limit in $L^2([L,R])$ of $\widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}$ as $[\bar{L},\bar{R}]$ increases to $(-\infty,\infty)$. Fortunately, most payoffs of practical interest belong to one of these classes, so localized projections obtained through truncated payoffs will converge. For payoffs with isolated singularities, well behaved approximations to the payoff near singularities can be used to obtain the coefficients of accurate hedges. This will be demonstrated for the log contract, which is well behaved away from the origin.

4.3. **Closed-Form Localized Projections.** In some cases, analytical formulas for the hedge coefficients $\beta_{a,k}$ can be derived from the representation in Proposition 4.1. For example, if

$$f(S_T) := (S_T - K)^p, \quad p \in \mathbb{N},$$

the coefficients in (12) satisfy

$$\theta_{a,k} = C(a,p)\left[(k - Ka - 1)^{(p+2)} - 2(k - Ka)^{(p+2)} + (k - Ka + 1)^{(p+2)}\right], \tag{16}$$

where $C(a,p) := \frac{a^{-(p+\frac{1}{2})}}{(p+1)(p+2)}$, which can be used to derive the coefficients for any polynomial at a negligible cost. However, in the particular case of polynomial payoffs, the exact beta coefficients are known analytically. Moreover, we can derive the $p$th moments of $\widetilde{\varphi}$, which will be used to develop efficient approximations to the true localized projection, as discussed in section 4.2.

**Corollary 4.3.** *The pth moment of $\widetilde{\varphi}$, $M^p$, is finite for all $p \in \mathbb{N}$, and is given by*

$$M^p = \int_R \widetilde{\varphi}(x)x^p dx = \begin{cases} \dfrac{6(\nu - 2 + \nu^{-1})}{\sqrt{3}(p+1)(p+2)} \cdot \dfrac{d^{p+2}}{d\zeta^{p+2}}\left[\dfrac{1}{1-\nu e^{\zeta}}\right|_{\zeta=0}, & p \text{ even} \\ 0, & p \text{ odd,} \end{cases} \tag{17}$$

*where $\nu = \sqrt{3} - 2$. Hence, the dual coefficients $\beta_{a,k}^{[p]}$ are known analytically for any polynomial:*

$$\beta_{a,k}^{[p]} := \int_R \widetilde{\varphi}_{a,k}(x)x^p dx = a^{-(p+\frac{1}{2})} \sum_{n=0}^{\lfloor \frac{p}{2} \rfloor} \binom{p}{2n} k^{p-2n} M^{2n}. \tag{18}$$

For example, the first four even moments are found to be

$$M^0 = 1, \quad M^2 = -1/6, \quad M^4 = 1/15, \quad M^6 = 17/84,$$

and higher order moments are easily derived. Note as well that for any payoff of the form $f(S_T) := (S_T - K)^p$, if $K = \bar{k}/a$ for some $\bar{k} \in \mathbb{N}$, the corresponding coefficients are found by shifting the index of those for $S_T^p$ by $\bar{k}$. Thus, for any $K = \bar{k}/a$, the localized projection of the power straddle $f(S_T) := (S_T - K)^2$ over $[L,R] = [k_L/a, k_R/a]$ is given simply by

$$\bar{P}_{\mathcal{M}_a} f(S_T) = a^{-5/2} \sum_{k=k_L}^{k_R} \left[(k - \bar{k})^2 - \frac{1}{6}\right] \varphi_{a,k}(S_T), \tag{19}$$

which corresponds to the localized projection in (14). Similarly, the coefficients of $f(S_T) := (S_T - K)^3$ are found as

$$\beta_{a,k} = a^{-7/2}(k - \bar{k})\left((k - \bar{k})^2 - \frac{1}{2}\right), \qquad k_L \le k \le k_R.$$

Given that polynomial payoffs offer exact formulas for the corresponding localized projection, an immediate consequence is the ability to transform a Taylor series (or other polynomial approximation method) into a portfolio of liquid contracts. Specifically, if $f \in C^{p+1}$,

$$f(S_T) = f(K) + \sum_{n=1}^{p} \frac{f^{(n)}(K)}{n!}(S_T - K)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad \xi \in (L,R).$$

By projecting the expansion onto $\{\varphi_{a,k}\}$, the coefficients for $k_L \le k \le k_R$ satisfy

$$\beta_{a,k} = a^{-1/2}\left[f(K) + \sum_{n=1}^{p}\left(\frac{f^{(n)}(K)a^{-n}}{n!}\right)\sum_{m=0}^{\lfloor \frac{n}{2} \rfloor}\binom{n}{2m}(k - \bar{k})^{n-2m} M^{2m}\right].$$

For near expiry contracts, low order projected Taylor expansions about $S_t$ can provide accurate approximations at a low cost.

## 5. Efficient Hedge Calculation

This section describes the implementation of the dual method for hedge coefficient calculation, and introduces a new method, which we call the *Alternative Biorthogonal Sequence* (ABS) method. The ABS method reduces the computational cost while retaining the accuracy of the dual method. This new approach to function approximation is especially promising for higher dimensions, with potential applications extending beyond finance.[12]

In this section, we focus on localized hedges of a payoff $f(S_T)$ over $[L,R]$, which can be applied to a large class non-$L^2$ payoff functions as discussed in section 4.2. The payoff may naturally be supported on $[L,R]$, or $[L,R]$ represents a truncation interval. In the latter case, an extension can be made at the boundaries $L$ and $R$ to form a global hedge (see section 6.5), after the local hedge is constructed. Information about $f$ on a larger interval $[\bar{L},\bar{R}] \supset [L,R]$ may be used to calculate hedge coefficients, where the choice of $[\bar{L},\bar{R}]$ depends on the method used as detailed below.

5.1. **The Dual Method.** The first method we consider for calculating hedge coefficients of the orthogonal projection is based on an approximation of the dual. From Proposition 4.1, with $\gamma \in \mathbb{N}_+$ we can approximate the biorthogonal dual on $[-(\gamma+1),\gamma+1]$ by

$$\widetilde{\varphi} \approx \sum_{|m|\le\gamma} \alpha_m T_m \varphi.$$

To obtain a hedge over $[L,R] = [k_L/a, k_R/a]$ for payoff $f(S_T)$, from (12) we obtain

$$f(S_T) \approx \sum_{k=k_L}^{k_R} \Big( \sum_{|m|\le\gamma} \alpha_m \langle f, \varphi_{a,k+m}\rangle \Big) \varphi_{a,k}(S_T) := f_a^\gamma(S_T). \tag{20}$$

Near the boundaries $L$ and $R$, we can use information about $f(S_T)$ outside of $[L,R]$ in our coefficient calculation. With $\gamma$ fixed (we find that $\gamma = 12$ is sufficient in general) the enlarged interval $[\bar{L},\bar{R}] = [L-(\gamma+1)/a, R+(\gamma+1)/a]$ contains the full knowledge of $f(S_T)$ that is used by this method to calculate coefficients, rather than treating $f(S_T)$ as if it were zero outside of $[L,R]$.[13]

If we calculate $\theta_a = (\theta_{a,k_L-\gamma}, \theta_{a,k_L-\gamma+1}, ..., \theta_{a,k_R+\gamma})$ defined by (12), $\beta_k$ is approximated by

$$\beta_k \approx \sum_{|m|\le\gamma} \alpha_m \theta_{a,k-m} = (\alpha * \theta)_k, \tag{21}$$

the circular convolution of $\theta$ with $\alpha = (\alpha_{-\gamma}, ..., \alpha_\gamma)$. The resulting implementation, which we refer to as the *dual method*, converges exponentially in $\gamma$ to the true projection over any hedge interval $[L,R]$, and for any resolution $a > 0$. As a result, the size of $\gamma$ (hence the required computational effort) increases slowly with higher resolutions. Figure 1 (right) demonstrates the local behavior of $f_a^\gamma(S_T)$ compared with interpolation when $\gamma = 12$.

**Proposition 5.1.** *Let $f$ be square integrable on $[L,R]$. With $f_a^\gamma(S_T)$ defined by (20), the deviation from the true orthogonal projection, $P_{\mathcal{M}_a}$, is characterized by:*

$$\|f_a^\gamma - P_{\mathcal{M}_a} f\|_j^{[L,R]} \le a^{1/2} C_j \cdot \tau(\gamma), \quad j = 1,2,$$

*where*

$$C_1 := (R-L+1)^{3/2} \|f\|_\infty^{[L,R]}, \quad C_2 := \frac{2}{\sqrt{3}}(R-L+1)\|f\|_\infty^{[L,R]},$$

*and*

$$\tau(\gamma) := \left(\frac{6}{1-\nu^2}\right)^{1/2} \nu^{\gamma+1}, \quad \nu := \sqrt{3}-2 \approx -0.268.$$

*Hence, for any resolution $a > 0$, the $L^1$ and $L^2$ errors over $[L,R]$ converge exponentially in $\gamma$.*

[12] In the supplemental appendix, we present a purely numerical method which works well for smooth payoffs and can be applied to more general Riesz bases for which $\mathcal{F}[\widetilde{\phi}]$ is known.

[13] For example, at the left boundary $k = k_L$, the approximation of $\widetilde{\varphi}_{a,k_L}$ is supported on $[L-(\gamma+1)/a, L+(\gamma+1)/a]$, and similarly for $\widetilde{\varphi}_{a,k_R}$, which is why we use $[\bar{L},\bar{R}]$. Equation (20) is therefore equivalent to the localized projection in (15).

In the next subsection, we introduce an alternative to the dual method which approximates the projection at a reduced cost.

5.2. **Alternative Biorthogonal Sequences (ABS).** As we noted before, the canonical *dual* to a Riesz sequence $\{\phi_{a,k}\} \subset \mathcal{H}$ at any scale is the unique biorthogonal sequence living within the *same* space as the Riesz sequence, $\mathcal{M}_a = \overline{\text{span}}\{\phi_{a,k}\}_k$, and is also a Riesz basis for $\mathcal{M}_a$. However, our search for biorthogonal *sequences* (as opposed to duals) is not limited to $\mathcal{M}_a$. For a given Riesz sequence $\{\phi_{a,k}\}$, we will refer to any sequence in $\mathcal{H}$ which is biorthogonal to $\{\phi_{a,k}\}$ (that is, $\langle \phi_{a,k}, \breve{\phi}_{a,j} \rangle = \mathbb{1}_{\{k=j\}}$) as an *alternative biorthogonal sequence* (ABS). By using an ABS instead of the canonical dual to approximate the hedging coefficients, an approximation to the true orthogonal projection is obtained, often at a significant reduction in computational effort. This is especially promising for higher dimensional extensions, which are left for future research. To ensure that convergent representations are still obtained, the ABS must satisfy a Bessel upper bound. This can be trivial to verify with the following proposition.

**Proposition 5.2.** *( [15], p.63) Let $\{\Psi_k\}_{k\in\mathcal{K}} \in \mathcal{H}$ be any sequence satisfying $\sum_{k\in\mathcal{K}} |\langle \Psi_j, \Psi_k \rangle| \leq B, \ \forall j \in \mathcal{K}$, where $B > 0$ is some constant. Then $\{\Psi_k\}$ is a Bessel sequence with bound $B$.*

Likewise, we will make use of the fact that a frame sequence, as opposed to a frame for all of $\mathcal{H}$, is actually a Bessel sequence on $\mathcal{H}$.

**Lemma 5.1.** *If $\{\Psi_k\}_{k\in\mathcal{K}}$ is a frame sequence in a Hilbert space $\mathcal{H}$, then $\{\Psi_k\}_{k\in\mathcal{K}}$ is a Bessel sequence (on all of $\mathcal{H}$). That is, for some $B > 0$, $\sum_{k\in\mathcal{K}} |\langle f, \Psi_k \rangle|^2 \leq B\|f\|_2^2, \ \forall f \in \mathcal{H}$.*

With the previous two results in hand, we can characterize an ABS approximation, denoted by $\breve{P}$, which is most appropriate for our purposes. With $\mathcal{M}_a := \overline{\text{span}}\{\phi_{a,k}\}_{k\in\mathbb{Z}}$, we have

$$\breve{P}_{\mathcal{M}_a} f := \sum_{k\in\mathbb{Z}} \langle f, \breve{\phi}_{a,k} \rangle \phi_{a,k}, \quad f \in \mathcal{H}. \tag{22}$$

**Proposition 5.3.** *Let $\phi$ be a Riesz generator in $\mathcal{H}$. If $\breve{\phi} \in \mathcal{H}$ is a compactly supported function such that the sequences of translates $\{\phi_{1,k}\}$ and $\{\breve{\phi}_{1,k}\}$ are biorthogonal, then for $a > 0$ the mapping $\breve{P}_{\mathcal{M}_a} : \mathcal{H} \to \mathcal{M}$ defined in equation (22) is a bounded linear projection operator on $\mathcal{M}$ which commutes with the orthogonal projection $P_{\mathcal{M}_a}$ of $\mathcal{H}$ onto $\mathcal{M}$. However, $\breve{P}_{\mathcal{M}_a} = P_{\mathcal{M}_a}$ iff $P_{\mathcal{M}_a}$ is the null operator on $\mathcal{M}_a^\perp$, that is $\breve{P}_{\mathcal{M}_a}(\mathcal{M}_a^\perp) = \{0\}$, which occurs iff $\breve{\phi}_1 \in \mathcal{M}_1$.*

By commutativity, the true projection preserves that of the ABS and conversely. While we can design an ABS to provide identical representations as the dual for polynomials of arbitrarily high order (as below), there will always be $f \in \mathcal{H}$ for which the ABS and orthogonal projections disagree. However, as shown in Proposition 5.5 below, ABS projections and value approximations converge to the true values, and they also converge at the same rate, up to a constant which depends on the ABS design.

*ABS Construction.* In the butterfly case, starting at the initial resolution, we begin our search for a viable ABS at one higher resolution and supported on $[-1, 1]$, posited to be a symmetric linear function of the form:

$$\breve{\varphi}(x) = \begin{cases} \lambda - 2(\lambda + \nu)x, & [0, 1/2) \\ -2\nu(1 - x), & [1/2, 1], \end{cases}$$

for some constants $\lambda, \nu$, where $\breve{\varphi}(x) = \breve{\varphi}(-x)$ for $x \in [-1, 0]$. To solve for $\lambda, \nu$, we use (i) $\int_0^1 \varphi(x)\breve{\varphi}(x)dx = 1/2$ (by symmetry, since $\int_{-1}^1 \varphi(x)\breve{\varphi}(x)dx = 1$) and (ii) $\int_0^1 \varphi(x)\breve{\varphi}(x-1)dx = 0$, which follows from symmetry and biorthogonality. These conditions yield the following.

**Result 5.1.** *The butterfly basis, with generator $\varphi$ and scale parameter $a > 0$ fixed, admits an ABS with generator*

$$\breve{\varphi}^{[1]}(x) = \begin{cases} 3 - 7|x|, & |x| < 1/2 \\ |x| - 1, & 1/2 \leq |x| \leq 1. \end{cases}$$

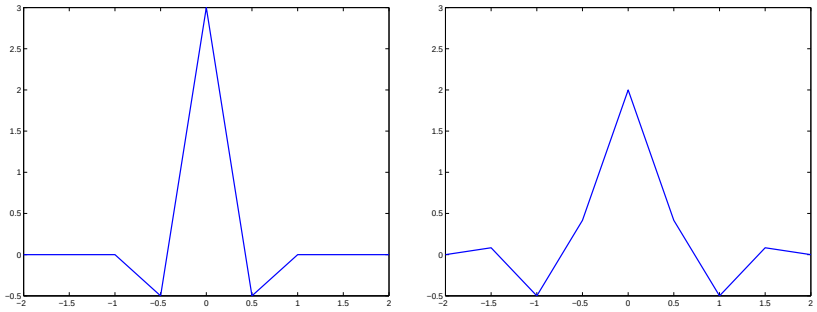

FIGURE 2. Left: $\text{ABS}_1$ generator supported on $[-1,1]$. Right: $\text{ABS}_2$ generator on $[-2,2]$.

We designed the ABS generator $\breve{\varphi}^{[1]}$ shown in Figure 2 to have narrow support, thereby reducing the computational burden of calculating the hedge coefficients, while mitigating boundary effects[14]. Note as well that by imposing symmetry of $\breve{\varphi}^{[1]}$ about the origin the following odd moments condition is satisfied:

$$\int_{\mathbb{R}} \breve{\varphi}^{[1]}(x) \cdot x^{2p+1} dx = \int_{\mathbb{R}} \widetilde{\varphi}(x) \cdot x^{2p+1} dx = 0, \quad p \in \mathbb{N}.$$

By increasing the support of $\breve{\varphi}$ to $[-2,2]$, we gain an additional degree of freedom with which we impose a second moment condition, which is solved for to obtain the ABS denoted by $\breve{\varphi}^{[2]}$:

$$\int_{\mathbb{R}} \breve{\varphi}^{[2]}(x) \cdot x^2 dx = \int_{\mathbb{R}} \widetilde{\varphi}(x) \cdot x^2 dx = -\frac{1}{6},$$

from Corollary 4.3 (where we recall the butterfly basis dual generator from (10)). We show in Proposition 5.4 that specifying an ABS by equating its first $\gamma$ moments to those of the dual is tantamount to obtaining the true projection[15] for all polynomials of degree $p \leq 2\gamma - 1$.

In general, we define an $\text{ABS}_\gamma$ generator as the ABS generator supported on $[-\gamma, \gamma]$ for which all moments $p \leq 2\gamma - 1$ coincide with the true dual (in addition to sharing all odd moments by the imposition of symmetry). The $\text{ABS}_\gamma$ generators are of the form

$$\breve{\varphi}^{[\gamma]}(x) = \sum_{|m| \leq 2\gamma - 1} c_m^{[\gamma]} \varphi(2x - m) = \sum_{|m| \leq 2\gamma - 1} 2^{-1/2} c_m^{[\gamma]} \varphi_{2,m}(x), \tag{23}$$

where $c_m^{[\gamma]}$ are determined as follows.

**Proposition 5.4.** *For any $\gamma \geq 2$, the unique set of coefficients $c_m^{[\gamma]}$, $|m| \leq 2\gamma - 1$, for which $\breve{\varphi}^{[\gamma]}$ is an $ABS_\gamma$ generator is given by the solution of the system*

$$\begin{aligned}
1 &= \lambda_0 c_0 + 2\lambda_1 c_1 + 2\lambda_2 c_2 \\
0 &= \lambda_2(c_{2k-2} + c_{2k+2}) + \lambda_0 c_{2k} + \lambda_1(c_{2k-1} + c_{2k+1}), \quad 1 \leq k \leq \gamma - 2 \\
0 &= \lambda_2 c_{2\gamma-4} + \lambda_0 c_{2\gamma-2} + \lambda_1(c_{2\gamma-3} + c_{2\gamma-1}) \\
0 &= \lambda_2 c_{2\gamma-2} + \lambda_1 c_{2\gamma-1} \\
M^{2k} &= \frac{2^{-2k}}{(2k+1)(2k+2)} \left( c_0 + \sum_{m=1}^{2\gamma-1} c_m \nu_m^{2k} \right), \quad 1 \leq k \leq \gamma - 1,
\end{aligned}$$

*where*

$$\nu_m^{2k} = (m-1)^{2k+2} - 2m^{2k+2} + (m+1)^{2k+2}, \quad m = 1, ..., 2\gamma - 1, \tag{24}$$

$$\lambda_0 = 5/12, \quad \lambda_1 = 3/12, \quad \lambda_2 = 1/24, \tag{25}$$

[14]We have found the ABS generated by $\breve{\varphi}^{[1]}$ to roughly double the approximation accuracy of interpolation in terms of hedging and pricing errors. However, the $\text{ABS}_2$ defined below provides a substantial improvement, especially for pricing, and the added cost is insignificant.

[15]We mean this in the sense of localized projections discussed in section 4.2. In particular, the projection coefficients $\beta_{a,k}$ for $a > 0$ and $k \in \mathbb{Z}$ will agree for all such polynomials, which themselves do not belong to $\mathcal{H}$.

*$c_k = c_{-k}$ $\forall 1 \le k \le 2\gamma - 1$, and $M^{2k}$ are the dual moments which are found by equation (17). Moreover, the ABS$_\gamma$ coefficients agree with the orthogonal projection coefficients from Corollary 4.1 for all polynomials of degree $p \le 2\gamma - 1$, and for all odd degrees.*

Application of Proposition 5.4 leads to the ABS$_2$ generator, which very closely approximates the coefficients obtained by the true dual for the functions tested.

**Result 5.2.** *The ABS$_2$ generator is given by $\breve{\varphi}^{[2]}(x) = \sum_{|m|\le 3} c_{|m|}\varphi(2x-m)$, where*

$$(c_0, c_1, c_2, c_3) = \left(2, \frac{5}{12}, -\frac{1}{2}, \frac{1}{12}\right),$$

*and is depicted in Figure 2.*

*ABS$_2$ Implementation.* We now outline the steps required to calculate coefficients for the ABS$_2$ method. From Result 5.2, the ABS$_2$ projection over $[L, R]$ is given by

$$f(S_T) \approx \sum_{k=k_L}^{k_R} \langle f, \breve{\varphi}^{[2]}_{a,k}\rangle \varphi_{a,k}(S_T), \quad \breve{\varphi}^{[2]}_{a,k}(x) = \frac{1}{\sqrt{2}} \sum_{|m|\le 3} c_{|m|}\varphi_{2a,2k-m}(x), \tag{26}$$

which requires the calculation of

$$\beta_{a,k} = \langle f, \breve{\varphi}^{[2]}_{a,k}\rangle = \sum_{|m|\le 3} c_{|m|}\theta_{2a,2k-m}, \quad k = k_L, \dots, k_R, \tag{27}$$

where $\theta_{2a,2k-m}$ is again defined by (12). Hence we use knowledge of $f$ on $[\bar{L}, \bar{R}] := [L - 2\Delta, R + 2\Delta]$ to determine ABS$_2$ coefficients.[16]

From (27), we find $\theta_{2a,j}$ for $j = 2k_L - 3, \dots, 2k_L + 2a(R-L) + 3$, which arise from $\theta_{2a,2k-m}$ as $k$ runs over $k = k_L, \dots, k_R$, and $m = -3, \dots, 3$. The first step is to calculate

$$\begin{aligned}\theta_{2a,j} &:= \frac{1}{\sqrt{2}} \int_{[\frac{j-1}{2a}, \frac{j+1}{2a}]} f(s)\varphi_{2a,j}(s)ds \\ &= 2a^{3/2}\left[\int_{[\frac{j-1}{2a}, \frac{j}{2a}]} f(s)\left(s - \frac{j-1}{2a}\right)ds + \int_{[\frac{j}{2a}, \frac{j+1}{2a}]} f(s)\left(\frac{j+1}{2a} - s\right)ds\right],\end{aligned}$$

where $\theta_{2a,0}$ is centered over 0. Using a Newton-Cotes rule reduces the number of required function evaluations by uniform sampling. As an example, a simple application of Simpson's rule with function values sampled over $\eta_{4a,l} := L + l/4a$, $l = -7, ..., 4a(R-L) + 7$ yields the approximation for $n = -3, -2, \dots, 2a(R-L) + 3$

$$\theta_{2a,2k_L+n} = \frac{a^{-1/2}}{6}\left[f(\eta_{4a,2n-1}) + f(\eta_{4a,2n}) + f(\eta_{4a,2n+1})\right] + \mathcal{O}(a^{-5}), \tag{28}$$

where the convergence rate is understood in the context of sufficiently smooth functions. The coefficients $\beta_{a,k}$ are then found from (27) for $k = k_L, \dots, k_R$.

*ABS Convergence.* Using a well designed ABS, the computational expense can be reduced, without compromising consistency as the strike space is refined. As long as $\{T_k\phi\}_k$ and $\{T_k\breve{\phi}\}_k$ are biorthogonal, by a change of variables $\langle \phi_{j,k}, \breve{\phi}_{j,m}\rangle = \langle T_k\phi, T_m\breve{\phi}\rangle = \delta_{k,m}$, so biorthogonality is preserved from one resolution to the next (recall the multiresolution notation from section 3.2.1). If $\breve{P}_j f = \sum_{k\in\mathbb{Z}} \langle f, \breve{\phi}_{j,k}\rangle \phi_{j,k}$ denotes the ABS projection operator defined on the refinement $U_j := \overline{\text{span}}\{\phi_{j,k}\}_{k\in\mathbb{Z}}$,

$$\begin{aligned}\breve{P}_j\breve{P}_j f &= \sum_m \Big\langle \sum_k \langle f, \breve{\phi}_{j,k}\rangle \phi_{j,k}, \breve{\phi}_{j,m}\Big\rangle \phi_{j,m} \\ &= \sum_m \Big(\sum_k \langle f, \breve{\phi}_{j,k}\rangle \delta_{k,m}\Big)\phi_{j,m} = \sum_k \langle f, \breve{\phi}_{j,k}\rangle \phi_{j,k},\end{aligned} \tag{29}$$

so $\breve{P}_j\breve{P}_j f = \breve{P}_j f$. We then have the following:

[16] At the left boundary $k_L$, $\theta_{2a,2k_L-m}$ for $m = 3$ requires $f(x)$ defined on $[L - 2\Delta, L + 2\Delta]$, while at the right boundary $k_R$, $\theta_{2a,2k_R+m}$ for $m = 3$ requires $f(x)$ defined on $[R - 2\Delta, R + 2\Delta]$. For a general ABS$_\gamma$ method, $[\bar{L}, \bar{R}] = [L - \gamma/a, R + \gamma/a]$.

**Proposition 5.5.** *Let $\phi \in \mathcal{H}$ be an RMRA generator, and $\breve{\phi}$ a compactly supported ABS generator. Then*

$$\lim_{j\to\infty} \|\breve{P}_j f - f\|_2 = 0, \quad \textit{and} \quad \lim_{j\to\infty} |\mathcal{V}\circ\breve{P}_j f - \mathcal{V}f| = 0,$$

*where $\breve{P}_j$ is a bounded linear (non-orthogonal) projection of $\mathcal{H}$ onto $U_j$, for all $j \in \mathbb{N}$. In fact, for each $f \in \mathcal{H}$, $\|\breve{P}_j f\|_2$ is uniformly bounded over $j \in \mathbb{Z}$. Moreover, for some $j' \in \mathbb{N}$, and all $j \geq j'$, we have $\|\breve{P}_j f - f\|_2 \leq C_1\|P_j f - f\|_2$ and $|\mathcal{V}\circ\breve{P}_j f - \mathcal{V}f| \leq C_2\|P_j f - f\|_2$ for some constants $C_1, C_2$ independent of $f$.*

From Proposition 5.5, not only does the ABS projection converge, but it does so at the same rate as the true projection (with same generator) up to a constant, and similarly for value approximations. For the butterfly basis, the convergence rate is $\mathcal{O}(2^{-2j})$.

5.3. **Choice of Butterfly Methods.** The main appeal of ABS methods is that rather than integrating the payoff against $\widetilde{\varphi}_{a,k}$, which has infinite (though safely truncated) support, we can obtain accurate approximations by integrating the payoff against an ABS dual with much narrower support. Hence, not only do the coefficients requires less work to obtain, ABS methods are more robust to the presence of discontinuities (which affect fewer coefficients) and are better equipped to handle isolated singularities than the dual method of section 5.1. From Proposition 5.5, ABS methods converge at the same rate as the true projection, up to a constant. Through experimentation, we find that the $\text{ABS}_2$ projector provides accurate approximations which closely match the true dual for smooth functions, and outperform with discontinuous payoffs (such as those in section 6.4). Given in addition its computational advantage, the $\text{ABS}_2$ projector is recommended.

## 6. Applications

We now demonstrate the effectiveness of the frame projection approach to hedging with several examples. After illustrating the decomposition of a butterfly basis hedge in terms of traded vanilla instruments, we consider exotic option hedging in an S&P 500 index market, a Russell 2000 index market, and a market for Henry Natural Gas. Simulation studies are used to analyze the hedging methodology for variance swaps and barrier options under the Heston and SABR stochastic volatility models, illustrating the application of static, semi-static, and mixed static-dynamic hedging. The application to exotic option pricing is considered in Appendix A.

6.1. **From Butterflies to Plain "Vanilla" Payoffs.** By construction, we can utilize the butterfly basis to approximate a given payoff $f$, which can then be expressed simply in terms of a payoff position in the underlying $S_T$, one strike of a put payoff $\psi_K^{put}(S_T) = (K - S_T)^+$, and call payoffs $\psi_K^{call}(S_T) = (S_T - K)^+$ with strikes along the support of $f$:

**Result 6.1.** *Let $f \in \mathcal{H}_+$, and suppose we have an order $\bar{K}+1$ approximation of $f$ in terms of the scale-$a$ butterfly basis given by*

$$f(S_T) \approx \beta_0\varphi_{a,0}(S_T) + \sum_{1\leq k\leq \bar{K}} \beta_k\varphi_{a,k}(S_T),$$

*where the $\beta_k$ for $k \geq 0$ are computed by any means. Noting that $\varphi_{a,0}(S_T) = \psi^{put}_{\frac{1}{a}}(S_T)$, we have the following static hedge in terms of positions in liquid assets:*

$$f(S_T) \approx a^{3/2}\Big[\beta_0\psi^{put}_{\frac{1}{a}}(S_T) + \beta_1 S_T + \sum_{1\leq k\leq \bar{K}+1} c_k^{call}\cdot\psi^{call}_{\frac{k}{a}}(S_T)\Big],$$

*where the call positions are given by*

$$c_k^{call} = \begin{cases} \beta_2 - 2\beta_1 & k = 1 \\ \beta_{k-1} - 2\beta_k + \beta_{k+1} & k \in \{2, ..., \bar{K}-1\} \\ \beta_{\bar{K}-1} - 2\beta_{\bar{K}} & k = \bar{K} \\ \beta_{\bar{K}} & k = \bar{K}+1. \end{cases}$$

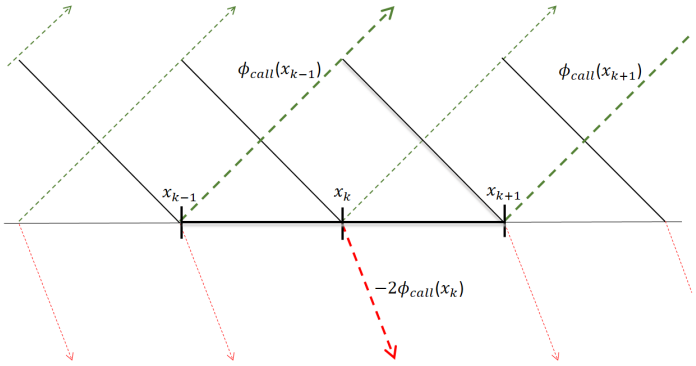


FIGURE 3. Butterfly basis in terms of call options along some grid $\{x_k\}$, where $\phi_{call}(x_k)$ is a call option with strike $x_k$.

Note too that we can avoid holding the asset itself if we use the put representation of $\varphi_{a,1}(S_T) = a^{3/2}\left[\psi^{put}_{\frac{2}{a}}(S_T) - 2\psi^{put}_{\frac{1}{a}}(S_T)\right]$. Figure 3 illustrates the butterfly basis decomposition into call option positions. In effect, the butterfly basis takes a set of vanilla payoffs which are not members of $\mathcal{H}$ and fashions a Riesz basis to which the theory and algorithms apply.[17]

Similarly, an out-of-the money butterfly representation (analogous to (2)) is simple to derive from the relationship

$$a^{-3/2}\varphi_{a,k}(S_T) = \psi^{put}_{\frac{k-1}{a}}(S_T) - 2\psi^{put}_{\frac{k}{a}}(S_T) + \psi^{put}_{\frac{k+1}{a}}(S_T). \tag{30}$$

With $S_0$ denoting the current market price of the underlying, by utilizing puts for butterfly elements with $k < S_0$ and calls for $k > S_0$, one can derive an out-of-the-money (OTM) representation. Specifically, we assume that vanilla options are available over[18] the interval $[L - 1/a, R + 1/a]$ with listed strikes $\{L - 1/a, L, L + 1/a, ..., R, R + 1/a\}$, where $[L, R]$ is the desired support for our hedge and $S_0 \in [L, R]$. We assume that $R = L + \bar{K}/a$ for some $\bar{K} \in \mathbb{N}$, so that $\bar{K} = (R - L)/a$. Now define $\beta^L_k := \beta_{k_L+k}$, where $k_L := L/a$, so $\beta^L_0$ corresponds to the basis element centered over $L$, $\beta^L_1$ corresponds to $L + 1/a$, and so on. With $\zeta^{call}_k$ ($\zeta^{put}_k$) denoting the portfolio holdings in a strike $L + k/a$ call (put), we obtain an OTM representation as

$$\begin{aligned}
&\zeta^{put}_k = a^{3/2}[\beta^L_{k-1} - 2\beta^L_k + \beta^L_{k+1}], \quad k = 0, ..., k^* - 1\\
&\zeta^{put}_{-1} = a^{3/2}\beta^L_0, \quad \zeta^{put}_{\frac{k^*}{a}} = a^{3/2}[\beta_{k^*-1} - 2\beta_{k^*}], \quad \zeta^{put}_{\frac{k^*+1}{a}} = a^{3/2}\beta_{k^*}\\
&\zeta^{call}_k = a^{3/2}[\beta_{k-1} - 2\beta_k + \beta_{k+1}], \quad k = k^* + 2, ..., \bar{K}\\
&\zeta^{call}_{\frac{k^*}{a}} = a^{3/2}\beta_{k^*+1}, \quad \zeta^{call}_{\frac{k^*+1}{a}} = a^{3/2}[\beta_{k^*+2} - 2\beta_{k^*+1}], \quad \zeta^{call}_{\frac{\bar{K}+1}{a}} = a^{3/2}\beta_{\bar{K}}
\end{aligned}$$

where $k^* = \lfloor (S_0 - L)/a \rfloor$. In practice, the set of available strikes will dictate our choice of which instruments to include in a hedge, namely our selection of $[L, R]$. For liquid products such as S&P 500 index options, active trading occurs for uniformly spaced strikes surrounding the ATM strike, although trading is not limited to contracts which have open interest (there is a higher transaction cost associated with less liquid strikes).

For example, on December 17, 2015, with a market close of 2,041.89 USD on the S&P 500 index, strikes for December 2015 options are quoted in increments of 5 USD from about 1,300 USD to 2,250 USD (see for example CBOE), straddling the current index level. Beyond these boundaries, the granularity increases, and likewise as the time to maturity grows, although a similarly fine strike space is observed for several months following the prompt month. At a maturity of two years (December 2017), a granularity of 25 or 50 USD is to be expected. As a

[17] While it is useful to unwind a butterfly hedge in terms of underlying options, trading the butterfly positions themselves has an advantage with respect to the margining practices used by clearing houses, since the finite risk associated with such products is well understood, and margins are set accordingly.

[18] Otherwise a representation purely in terms of calls or in terms of puts is available.

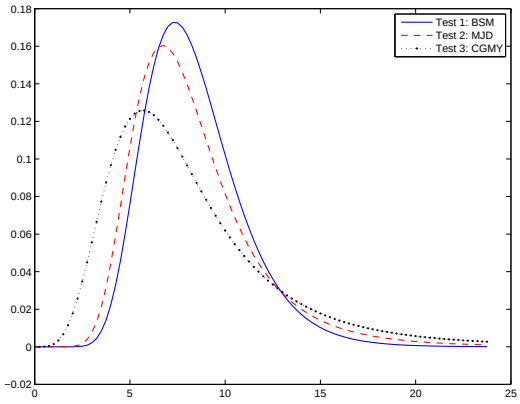


FIGURE 4. A plot of scaled butterfly prices, $a^{-1/2}\{\mathcal{V}\circ\varphi_{a,k}\}_{k\geq 0}$, for three test cases with $a=2^{10}$.

static hedge draws closer to maturity, the portfolio can be adjusted to add finer granularity if desired.

*Remark* 4. Given a representation in terms of vanilla instruments, the approximate value for an exotic payoff may be obtained directly from market quotes of vanilla instruments. In this case, smile effects are imputed in the prices obtained. One can also obtain model-based prices by pricing the vanillas (or butterflies) using a standard pricing method for the chosen model. As illustrated in Appendix A, value approximations of projected payoffs converge very quickly with respect to the resolution of the payoff projection. The set of butterfly basis values is illustrated in Figure 4 for the three test cases considered in the appendix, where, at fine resolutions a smooth risk-neutral density emerges. This reflects the Breeden and Litzenberger result [5], $\frac{\partial}{\partial K}\mathbb{P}[S_T\leq K]=e^{rT}\frac{\partial^2 P_T}{\partial K^2}(K;S_0)$, where $P_T(K;S_0)$ is the price of a strike-$K$ put, when one considers the relation between (30) and the second derivative.

### 6.2. **Static Hedging: Exotic European Options.**

*Example 1: Russell 2000 Market.* In this example, we consider a market for capped power straddles written on the Russell 2000 (RUT) index. With a closing price of 1,365.26 USD on March 10, 2017, European option strikes are listed on $[500,1700]$, at increments of 5 USD near the money, and 10 USD away from the money.[19] We consider a capped power straddle

$$f(S_T)=(S_T-K)^2\mathbb{1}_{|S_T-K|\leq C}+(C-K)^2\mathbb{1}_{|S_T-K|>C},$$

which reduces the extreme payoff risk (for the seller) from a standard power straddle (discussed in section 2.1), and fix $K=1365$ to be near the money and $C=80$. We simulate 10,000 paths each of a Heston and SABR model, with parameters listed in Table 1. Heston's model [29] prescribes the dynamics

$$\begin{cases} dS_t=\mu S_t dt+\sqrt{\nu_t}S_t dW_{1,t}\\ d\nu_t=\kappa(\theta-\nu_t)dt+\sigma_\nu\sqrt{\nu_t}dW_{2,t}, \end{cases} \tag{31}$$

where $W_{1,t}$ and $W_{2,t}$ are correlated Brownian motions, with correlation $\rho\in(-1,1)$. Similarly, the SABR model [26] posits

$$\begin{cases} dS_t=\mu S_t dt+\nu_t S_t^\beta dW_{1,t}\\ d\nu_t=\alpha\nu_t dW_{2,t}, \end{cases} \tag{32}$$

where $\beta\in(0,1)$. For each $\Delta$ fixed, hedges for both methods are formed from the strikes in $[L,R]=[K-C,K+C]$. For this example, the payoff is bounded outside of $[L,R]$, and it is simple to form a global hedge by flat extrapolating the payoff from each of the boundaries $L$ and $R$ using vanilla options (as discussed in the supplemental appendix). This global hedge matches the capped payoff outside of $[L,R]$. For payoffs of this type, which are linear outside of $[L,R]$, there is no tail risk since we can replicate them globally with vanilla options.

[19] The last few strikes on either boundary are at increments of 50 USD.

| | | Heston | | | | SABR | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\Delta$ | Method | Avg.Err. | MAD | Std | Max.Err. | Avg.Err. | MAD | Std | Max.Err. |
| 5 | $ABS_2$ | 0.006 | 1.308 | 1.671 | 4.319 | 0.025 | 1.599 | 1.853 | 4.162 |
| 5 | Interp | 3.324 | 3.324 | 2.353 | 6.250 | 4.192 | 4.192 | 1.853 | 6.250 |
| 10 | $ABS_2$ | 0.010 | 5.008 | 6.557 | 17.287 | -0.624 | 6.458 | 7.512 | 16.658 |
| 10 | Interp | 11.859 | 11.859 | 9.801 | 25.000 | 16.043 | 16.043 | 7.512 | 25.000 |
| 20 | $ABS_2$ | 0.196 | 18.125 | 24.995 | 69.439 | -21.547 | 29.201 | 27.277 | 66.630 |
| 20 | Interp | 34.770 | 34.770 | 39.699 | 100.000 | 45.533 | 45.533 | 27.581 | 100.000 |

TABLE 1. Capped power straddle simulated hedge error for Russell 2000 example: $S_0 = 1365.26$, $K = 1365$, $C = 80$, $T = 1/12$. Heston: $\mu = 0.02$, $\nu_0 = (0.14)^2$, $\sigma_v = 0.15$ , $\kappa = 3$, $\rho = -0.6$, $\theta = v_0$. SABR: $\mu = 0$, $\nu_0 = (0.14)^2$, $\alpha = 0.15$, $\beta = 0.9$, $\rho = -0.6$.

Table 1 documents the hedging errors in USD for $ABS_2$ and interpolation for the two stochastic volatility models for $S_T$. At each resolution available in the market the $ABS_2$ method outperforms interpolation in terms of Average Error, Mean Absolute Deviation (MAD), Standard Deviation (Std), and Maximum Absolute Error (Max.Err.). In particular, the MAD is significantly higher for interpolation at each resolution in both models. For this payoff, interpolation systematically overhedges (while for a concave payoff such as the log contract, it will always underhedge). In contrast, projection balances errors from over and under hedging, leading to a much smaller error on average (see Avg.Err. column). We also note that the maximum error is much larger for interpolation.

*Remark* 5. In practice, the choice of resolution is based on the application objectives, and the tolerance (as well as definition) for hedging error may vary. For any fixed resolution, frame projection provides a formula for the quantity of instruments to hold at each strike. The previous example considers four practical metrics: average error, MAD, standard deviation, and maximum error. These metrics can be computed for the error of the physical payoff, or based on the results of a simulated model of interest. If the resulting error tolerance in any given metric is not met, the resolution can be increased up to granularity of listed strikes.

*Example 2: S&P 500 Market.* The previous example shows the improved performance of frame projection for simulated asset paths. In general, we find that frame projection offers a consistent improvement over linear interpolation for arbitrary payoffs. To illustrate this, we provide three nonstandard exotic European payoff examples for the S&P 500 market from section 6.2 with $[L, R] = [1300, 2200]$ selected to match the available strikes. The first is a power straddle with rational exponent, the second is a Gaussian-style payoff, and the third is the product of a power straddle with a damping term $\log(S_T)/S_T$. For each of these payoffs, we consider strikes which are spaced uniformly $\Delta$ USD apart, with $\Delta \in \{100, 50, 25, 10, 5\}$, and record the relative absolute hedge error (RAHE) defined by

$$\text{RAHE} = \frac{\sum_{i=1}^{N_s} |\tilde{f}_{[L,R]}(x_i; a) - f(x_i)|}{\sum_{i=1}^{N_s} |f(x_i)|}, \tag{33}$$

where $\{x_i\}_{i=1}^{N_s}$ is a uniform sampling of $[L, R]$, and $\tilde{f}_{[L,R]}(x_i; a)$ is the prescribed approximation on $[L, R]$ at resolution $a$ ($\Delta := 1/a$). For this comparison, we have isolated the local error by restricting to $[L, R]$. Interpolation can be represented in the butterfly basis by

$$f^{Interp}_{[L,R]}(S_T; a) = \sum_{k_L \le k \le k_R} a^{-1/2} f\Big(\frac{k}{a}\Big) \varphi_{a,k}(S_T).$$

Comparing the $ABS_2$ method to linear interpolation in Table 2, we see that interpolation incurs about 2.6 times the error of the $ABS_2$ method, and the difference in computational effort is negligible. Experimenting with many smooth payoff forms, an interesting finding is that a constant of about 2.6 holds in general, while this constant holds for non-smooth payoffs once a threshold resolution is reached. For the first example payoff, which is non-smooth, the threshold is reached for $\Delta$ between 100 and 50. This indicates a systematic error reduction when using frame projection over interpolation, which holds in general. The alternative model-free method

| $f(S_T) = \lvert S_T - K\rvert^{6/5}$ | | | | | |
|---|---|---|---|---|---|
| Scale($\Delta$) | 100 | 50 | 25 | 10 | 5 |
| Interp | 1.335e-02 | 1.757e-03 | 4.703e-04 | 8.084e-05 | 2.111e-05 |
| $ABS_2$ | 1.056e-02 | 6.663e-04 | 1.788e-04 | 3.082e-05 | 8.061e-06 |
| Ratio | 1.264 | 2.637 | 2.630 | 2.623 | 2.619 |
| $f(S_T) = (R-L)^{1/2} \cdot \exp\big(-\big((S_T-K)/(R-L)\big)^2\big)$ | | | | | |
| Scale($\Delta$) | 100 | 50 | 25 | 10 | 5 |
| Interp | 1.739e-03 | 4.344e-04 | 1.086e-04 | 1.737e-05 | 4.342e-06 |
| $ABS_2$ | 6.739e-04 | 1.675e-04 | 4.181e-05 | 6.686e-06 | 1.671e-06 |
| Ratio | 2.580 | 2.594 | 2.597 | 2.598 | 2.598 |
| $f(S_T) = \log(S_T) \cdot (S_T - K)^2/S_T$ | | | | | |
| Scale($\Delta$) | 100 | 50 | 25 | 10 | 5 |
| Interp | 2.665e-02 | 6.667e-03 | 1.667e-03 | 2.667e-04 | 6.668e-05 |
| $ABS_2$ | 1.025e-02 | 2.565e-03 | 6.415e-04 | 1.027e-04 | 2.566e-05 |
| Ratio | 2.601 | 2.599 | 2.598 | 2.598 | 2.598 |

TABLE 2. RAHE of three payoffs on $[L, R] = [1300, 2200]$, corresponding to the S&P 500 example, with payoff strike set to $K = (L + R)/2$. Ratio gives the ratio of the RAHE for interpolation (Interp) over that of $ABS_2$. The errors are taken relative to a uniformly sampled "value" over the interval $[L, R]$. The values for the three payoffs are respectively 694.125, 27.677, 297.574.

of Carr-Madan is not reported, as frame projection and interpolation are orders of magnitude more accurate, even for smooth payoffs.

6.3. **Mixed Static-Dynamic Hedging: Variance Swaps.** Due to the inextricable presence of volatility in the trading of many derivative products, extensive markets have developed to exchange volatility contingent securities. An especially important product is the realized variance swap. Given a set of trading dates $\{t_0, ..., t_n\}$ at which an investor can trade in (generic) futures contracts with prices denoted by $F_i$, a standard variance swap specification is the terminal payoff $VS_n - K$, where

$$VS_n = \frac{N}{n} \sum_{i=1}^{n} \ln^2 \left( \frac{F_i}{F_{i-1}} \right), \tag{34}$$

and $N$ is a trading day count specified in the contract. From Carr and Lee [10], a semi-static hedging strategy can be used to offset a position in the variance swap. In terms of the simple returns $R_i := (F_i - F_{i-1})/F_{i-1}$, $i = 1, ..., n$, a Taylor series expansion yields

$$\ln^2 \left( \frac{F_i}{F_{i-1}} \right) = 2R_i - 2(\ln F_i - \ln F_{i-1}) - \frac{1}{3} R_i^3 + \mathcal{O}(R_i^4),$$

from which

$$VS_n = \frac{N}{n} \left[ \sum_{i=1}^{n} \frac{2}{F_{i-1}} (F_i - F_{i-1}) - 2(\ln F_n - \ln F_0) - \frac{1}{3} \sum_{i=1}^{n} R_i^3 + \sum_{i=1}^{n} \mathcal{O}(R_i^4) \right].$$

The first term in the approximation represents a dynamic position in futures contracts, namely a holding of

$$e^{-r(t_n - t_i)} \frac{N}{n} \frac{2}{F_{i-1}} \tag{35}$$

futures contracts during the period $t_{i-1}$ to $t_i$. Given the existence of a traded log contract (as advocated in [38]), $g(F_n) = -2(\ln F_n - \ln F_0)$ represents a static position initiated at $t_0$. In the absence of log contracts, the methods presented in this paper facilitate the accurate approximation of such contracts in terms of liquidly traded assets, namely call and or put options.

| | | Heston | | | SABR | | |
|---|---|---|---|---|---|---|---|
| $\Delta$ | Method | MAD | Std | Max.Err. | MAD | Std | Max.Err. |
| 0.10 | Exact | 0.107 | 0.152 | 1.344 | 0.127 | 0.198 | 2.824 |
| 0.10 | $ABS_2$ | 0.701 | 0.836 | 4.396 | 0.714 | 0.860 | 4.769 |
| 0.10 | Interp | 1.820 | 0.880 | 6.078 | 1.807 | 0.891 | 9.913 |
| 0.10 | CM | 10.081 | 12.730 | 52.988 | 10.554 | 13.400 | 71.287 |
| 0.20 | Exact | 0.107 | 0.154 | 3.264 | 0.125 | 0.193 | 2.782 |
| 0.20 | $ABS_2$ | 2.787 | 3.325 | 13.380 | 2.787 | 3.320 | 16.173 |
| 0.20 | Interp | 7.323 | 3.618 | 27.773 | 7.359 | 3.681 | 26.474 |
| 0.20 | CM | 15.583 | 19.623 | 92.848 | 16.400 | 20.596 | 81.585 |

TABLE 3. Variance swap hedge errors for Henry Natural Gas example: $F_0 = 3.07$, $T = 0.5$, $n = 100$, $\mu = 0.02$, $r = 0.02$. Heston: $\nu_0 = (0.14)^2$, $\sigma_v = 0.15$ , $\kappa = 3$, $\rho = -0.6$, $\theta = v_0$. SABR: $\nu_0 = (0.4)^2$, $\alpha = 0.4$, $\beta = 0.9$, $\rho = -0.6$.

*Example: Henry Natural Gas.* As an example, consider the case of NYMEX Henry Hub Natural Gas Futures, with December 2018 expiry and futures settlement price of $F_0 = 3.070$ USD per mmBtu (million British thermal units) on March 12, 2017. A contract unit is 10,000 mmBtu. Consider a six month realized variance swap, written on one unit of the December 2018 futures contract, with 100 monitoring dates prior to expiry of the variance swap.[20] Cash settled vanilla European option strikes are listed from 0.60 to 18.00 (price normalized), at increments of 0.05.

Table 3 compares the hedge performance of one contract unit of a variance swap written on $VS_n$, with $N/n := 1$ for simplicity. We simulate 10,000 paths each of a Heston and SABR model, with parameters listed in Table 3. At each of $n = 100$ monitoring dates, the portfolio is rebalanced according to the dynamic futures position in (35). The hedging strategy also requires a static payoff of $g(F_n) = -2(\ln F_n - \ln F_0)$. The method labeled "Exact" represents the strategy which holds this (untraded) log payoff, while $ABS_2$, Interp (interpolation), and CM (Carr-Madan) use their respective approximations for the log contract in terms of vanilla instruments with strikes in $[L, R] = [0.6, 18.00]$. Unlike the capped power straddle considered in section 6.2, in this case we cannot perfectly match the payoff outside of $[L, R]$. To mitigate tail risk, the $ABS_2$ and interpolation methods use a linearly extrapolated vanilla hedge at the boundaries $L$ and $R$, which is detailed in the supplemental appendix (this is one way to form a global hedge from a local one).[21] Finally, the outcome of each strategy is compared with the exact formula for realized variance (34) to determine the hedge error.[22]

From Table 3, $ABS_2$ outperforms interpolation and the CM approach at each resolution, which are chosen to match market availability. The baseline "Exact" method is not implementable in practice, and is used as a benchmark. For both the Heston and SABR dynamics, the $ABS_2$ results in an improved mean absolute deviation (MAD), standard error (Std), and maximum absolute error (Max.Err.). In particular, MAD is less than half that of interpolation.

6.4. **Semi-Static Hedging: Barrier Options.** As demonstrated in [7, 8], barrier options, which are actively traded in currency markets, are natural candidates for semi-static hedging. While the ability to *perfectly* hedge such contracts depends on very specific market conditions, such as a Black-Scholes-Merton (BSM) economy, and the assumption that portfolios can be liquidated immediately upon barrier breach and at the BSM price, imperfect hedging is still possible under relaxed assumptions. Moreover, in cases where delta hedging is known to fail dramatically such as for down-and-out put (DOP) and up-and-out-call (UOC) contracts [37], the performance of semi-static hedging is much more satisfactory, although it is still far from a perfect hedge.

[20]For simplicity, we assume uniform monitoring for this particular simulation.

[21]The Carr-Madan approach also results in a linearly extrapolated hedge given a finite set of strikes on $[L, R]$. The implementation of Carr-Madan follows from a discretization of equation (2) (see [45] for details).

[22]Note that the payoff $VS_n - K$ of a strike $K$ variance swap adds a cash position to each hedge, and has no affect on the hedge error.

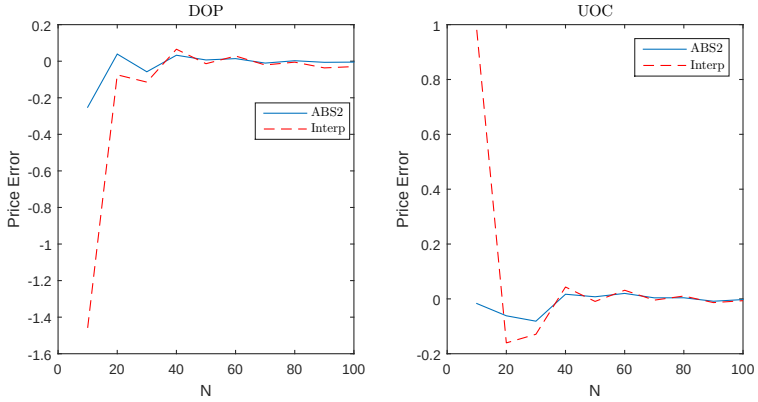


FIGURE 5. Pricing error for continuous monitoring as a function of the number of strikes $N$ for a DOP with $K = 100$ and $B = 80$ (left) and a UOC with $K = 100$ and $B = 120$ (right). Comparison of linear interpolation and $\text{ABS}_2$ when the approximated adjusted payoff is sold. BSM model: $\sigma = .18$, $r = .06$, $d = .02$, $T = 1$.

In the following set of experiments we consider the problem of semi-static hedging a DOP option, as discussed in section 2.2, where the adjusted payoff is defined as

$$\tilde{f}(S_T) = \begin{cases} (K - S_T)^+ & \text{if } S_T > B \\ -\left(\frac{S_T}{B}\right)^p \left(K - \frac{B^2}{S_T}\right)^+ & \text{if } S_T \le B \end{cases} \tag{36}$$

with strike $K$, barrier $B < K$, interest rate $r \ge 0$, dividend yield $d \ge 0$, volatility $\sigma > 0$, and $p := 1 - 2(r-d)/\sigma^2$. The adjusted payoff is constructed using a prescribed volatility (for now it is taken to be $\sigma$, the true market volatility), and a butterfly hedge is established in vanilla options using the $\text{ABS}_2$ methodology. In particular, we set $[L, R] = [B^2/K, K]$, fix a number $N$, and obtain the butterfly coefficients at strikes $L + k/a$ for $k = 0, ..., \bar{K} := N - 1$, where $a := (R - L)/\bar{K}$. In this example, the payoff is zero outside of $[L, R]$, and the hedge is entirely local. By setting the coefficients at $L$ ($k = 0$) and $R$ ($k = \bar{K}$) equal to zero, we obtain a hedge that requires $N$ strikes in vanilla options.

For each experiment, we simulate the underlying dynamics up to the first barrier breach, or contract expiry $T = 1$, whichever occurs first. If breach occurs, the vanilla portfolio is liquidated at the BSM prices, and is then discounted to obtain the profit and loss (P&L) for that simulation run (at breach, the true adjusted payoff has zero value). Otherwise, the portfolio results in a European-style payoff at time $T$, and the difference between this payoff and $(K - S_T)^+$ is discounted to the present to obtain the P&L.

Figure 5 illustrates the difference in pricing error (profit and loss) that occurs when selling the vanilla portfolio for each of the two methods linear interpolation and $\text{ABS}_2$ as a function of $N$. The left panel considers a DOP option with $K = 100$ and $B = 80$, and the right panel a UOC option with $K = 100$ and $B = 120$. In both cases, the time to maturity is $T = 1$ and the portfolio is sold exactly when the underlying touches the barrier, so the only source of error is in approximating the adjusted payoff. In both cases the $\text{ABS}_2$ methodology outperforms linear interpolation, for small (practical) values of $N$ as well as when $N$ is taken very large. The difference in errors is most pronounced for small values of $N$.

For the remaining experiments, monitoring is allowed every other day ($M = 252/2$ monitoring dates), so if a breach occurs the underlying can drift significantly away from the barrier prior to liquidating the vanilla portfolio, resulting in a nonzero P&L. Hence, there are now two sources of error: the error arising from approximating the adjusted payoff, as well as the error from a drift in the underlying away from the barrier at liquidation. With a finite monitoring frequency, even the true adjusted payoff results in an imperfect hedge. The simulation is then repeated 10,000 times, and relative P&Ls are recorded, where each P&L is relative to the discretely monitored DOP option value. In all experiments $r = 0.06$ and $d = 0.02$.

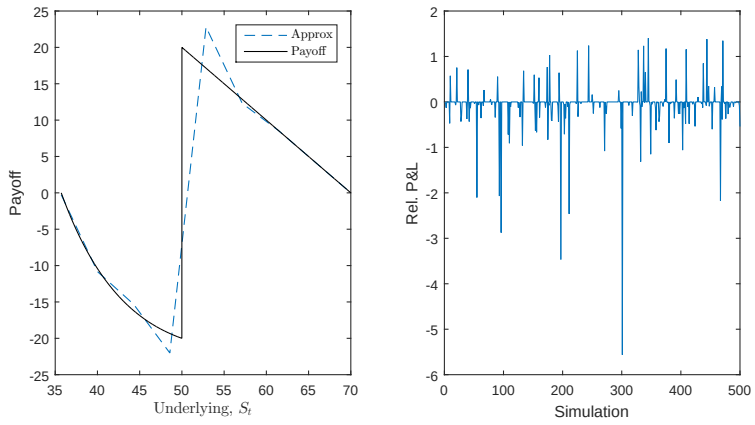


FIGURE 6. The left panel illustrates the DOP adjusted payoff for a contract with strike $K = 70$ and knock-out barrier $B = 50$, along with the $\text{ABS}_2$ approximation with $N = 9$ options, uniformly spaced within $(B^2/K, K)$. Additional parameters: $r = .06$, $d = .02$, $T = 1$, $M = 252/2$, $\sigma = .18$, $S_0 = 65$. The right panel illustrates the profit and loss for the semi-static hedging strategy in a discretely monitored BSM market, where each bar represents the outcome of a single simulation.

Figure 6 illustrates the butterfly approximation to the adjusted payoff (left) using $N = 9$ options, as well as the P&Ls for the first 500 simulations (right). Note that the previous example illustrates the error that occurs when the underlying is equal to the barrier upon liquidation, and the true adjusted payoff should have zero value. However, this value becomes nonzero as the underlying deviates from the barrier prior to liquidation. Moreover, in the absence of a breach, the adjusted payoff does not always coincide with the vanilla option at delivery. These errors can be relatively large, as seen in the right panel of Figure 6. Unlike the second source of error, the first of these can be mitigated by increasing the monitoring frequency.

*Remark* 6. Given that the alternative to semi-static hedging is a dynamic hedging strategy, we replicate an experiment in [37] to obtain a context for the errors observed with our framework.[23] A UOC contract with $K = 110$, $B = 140$ and $S_0 = 100$ is considered in a BSM model where $\sigma = 0.2$, and other parameters are as before. Relative to the price 2.277 (calculated in [37]), the delta hedging strategy has a mean error of -0.007 and standard deviation of 0.949. This compares to a mean error of -0.077 and a standard deviation of 0.5036 with the $\text{ABS}_2$ method using $N = 11$ options. While the mean error is somewhat larger, dynamic hedging is nearly twice as risky when measured by standard deviation. While not reported for dynamic hedging, the relative min and max deviation for $\text{ABS}_2$ were -9.34 and 1.89 respectively, so even with half the standard deviation of dynamic hedging, large errors can still be expected. Under the same conditions, the STR-static method of [37] performs similarly as frame projection in this case with a standard deviation of 0.523, but with more than twice the relative mean error at 0.164. The CAL-static method results in a mean error of 0.09, but with a standard deviation of 0.628, again both higher than frame projection.

Table 4 illustrates the performance of the semi-static hedging strategy for a DOP contract with $K = 70$ and $B = 60$, in terms of the mean error (mean), mean absolute deviation (MAD), minimum error (min), maximum error (max) and standard deviation of errors (std), all relative to the discretely monitored option value with the same monitoring frequency. As the number of vanilla options, $N$, increases, we expect that the approximation error and standard deviation should decline. While this is typically true, the convergence is far from monotone, and due to the discrete monitoring frequency there will be errors associated with the hedging policy even in the limit.

[23]Implementing the dynamic delta hedge is a delicate matter for DOP and UOC call options, especially as the underlying approaches the barrier (see [4] for examples). A careful implementation and description of such a strategy is outside the scope of the present work. Our focus is on introducing a new static hedging method, and illustrating its potential for application.

| N | 6 | 8 | 10 | 12 | 25 | 50 | 100 |
|---|---|---|---|---|---|---|---|
| mean | -0.11 | 0.043 | -0.003 | -0.033 | -0.031 | -0.027 | -0.026 |
| MAD | 0.14 | 0.070 | 0.037 | 0.043 | 0.034 | 0.028 | 0.027 |
| std | 0.30 | 0.162 | 0.165 | 0.178 | 0.151 | 0.134 | 0.140 |
| min | -4.26 | -2.936 | -3.285 | -3.636 | -3.546 | -3.630 | -3.775 |
| max | 0.85 | 0.768 | 0.764 | 0.760 | 0.743 | 0.413 | 0.364 |

TABLE 4. Hedge errors (P&L) of semi-static DOP hedge for BSM model, relative to discretely monitored value, 3.086, with breach percentage = 58.45. Parameters: $S_0 = 65$, $K = 70$, $B = 60$, $\sigma = 0.18$, $r = 0.06$, $d = 0.02$, $M = 252/2$.

| N | 6 | 8 | 10 | 12 | 25 | 50 | 100 |
|---|---|---|---|---|---|---|---|
| mean | -0.01 | -0.015 | -0.037 | -0.014 | -0.018 | -0.014 | -0.016 |
| MAD | 0.03 | 0.020 | 0.040 | 0.016 | 0.020 | 0.015 | 0.016 |
| std | 0.15 | 0.095 | 0.122 | 0.093 | 0.103 | 0.092 | 0.096 |
| min | -2.48 | -1.927 | -2.163 | -2.420 | -2.649 | -2.321 | -2.295 |
| max | 0.52 | 0.525 | 0.249 | 0.547 | 0.506 | 0.551 | 0.529 |

TABLE 5. Hedge errors (P&L) of semi-static DOP hedge for Heston's model, relative to discretely monitored value, 8.9836, with breach percentage = 52.56. Parameters: $S_0 = 65$, $K = 70$, $B = 50$, $\nu_0 = 0.1800$, $\sigma_v = 2.4400$, $\kappa = 0.3800$, $\rho = -0.5800$, $\theta = 0.1800$, $r = 0.06$, $d = 0.02$, $M = 252/2$

In the final set of experiments, we relax the BSM model assumption, and consider a stochastic volatility driven market with imperfect information. In particular, consider a Heston market with underlying and variance dynamics described by (31). The initial *variance* is $\nu_0 = 0.18$, correlation $\rho = -0.58$, and drift $\mu = 0.02$. While the market volatility is allowed to change, we assume that traders still price the underlying portfolio using Black-Scholes but at the prevailing observed market volatility.[24] Since a single volatility must be used to establish the original hedge, we take this to be the long term average volatility, $\sqrt{\theta} = 0.4243$. We consider the original contract parameters $K = 70$ and $B = 50$, with the remaining model parameters (obtained from a calibration to AUDJPY call options on September 16, 2008, [16, 35]) summarized in Table 5. Since the prevailing volatility at the time of liquidation differs from the volatility used establish the hedge, a new source of error is introduced in this case. Even so, the results are similar to what we have observed in a pure BSM market. In this case, even if the monitoring frequency becomes arbitrarily large, hedging errors will persist at large values of $N$.

6.5. **Global Hedges and Tail Risk.** In practice, no hedging method is capable of perfectly capturing the global behavior of arbitrary payoffs with a market-limited set of traded instruments. The tail behavior of an approximation is constrained by market availability. For model based pricing, there is no constraint on the truncation interval $[L, R]$. Hence an initial $[L_0, R_0]$ may be chosen, and expanded until value approximations converge within a given tolerance. Given closed-form expressions for coefficients, additional coefficients can be added incrementally until convergence its attained.

In the context of hedging, let $[L, R] = [L_0, R_0]$ contain the set of strikes which trade at time $t_0$. Once a local hedge is established over $[L, R]$ (using the $\text{ABS}_2$ method), which is restricted by market availability, a global hedge can be formed by extending the payoff linearly beyond the boundaries of $L$ and $R$ with vanilla instruments. If $g(S_T)$ denotes the approximation of a payoff $f(S_T)$ on $[L, R]^c$, given a model for $S_T$, the expected error with boundary extension satisfies

$$\begin{aligned}
&\left|\mathbb{E}\left[f(S_T) - \left(P_{\mathcal{M}_a^+} f(S_T)\mathbb{1}_{[L,R]}(S_T) + g(S_T)\mathbb{1}_{[L,R]^c}(S_T)\right)\right]\right| \\
&\leq \left|\mathbb{E}\left[\left(f(S_T) - P_{\mathcal{M}_a^+} f(S_T)\right)\mathbb{1}_{[L,R]}(S_T)\right]\right| + \left|\mathbb{E}\left[(f(S_T) - g(S_T))\mathbb{1}_{[L,R]^c}(S_T)\right]\right| \\
&\leq \zeta\|(f - P_{\mathcal{M}_a^+} f)\mathbb{1}_{[L,R]}\|_2 + \mathcal{E}_g([L,R]^c)
\end{aligned}$$

[24]Hence, while they can observe the level of volatility, the true process governing the volatility is unknown to the market.

for some $\zeta > 0$ (recall the discussion following Corollary 4.1), where $\mathcal{E}_g([L,R]^c)$ is the expected error of the tail approximation. Hence, while the local hedge restricted to $[L,R]$ can be thought of as roughly model independent (governed by the norm error of projection), truncation leads to model-dependent tail risk captured by $\mathcal{E}_g([L,R]^c)$. The tail approximation approach of linear boundary extension is illustrated in the applications below for variance swaps and capped power straddles. Implementation details for this simple extension approach are provided in the supplemental appendix.

*Remark* 7. In addition to tail extension, semi-static hedging can be used for portfolios of European options, reducing the tail risk associated with the initial truncation to $[L_0, R_0]$. Adjustments can be made at an arbitrary set of dates $\{t_m\}_{m=1}^M \subset (0,T)$, where the local hedge is adjusted to the interval $[L_m, R_m]$, depending on the realized value of $S(t_m)$. Moreover, the interior coefficients on $[L_{m-1}, R_{m-1}] \cap [L_m, R_m]$ remain unaltered (those corresponding to the intersection but not on the boundaries $L_m, R_m$), since the physical payoff does not change. At each $t_m$ the portfolio may be adjusted by unwinding or augmenting by out-of-the-money positions. Assuming $[L_m, R_m]$ is sufficiently wide, adjustments will be required infrequently.

## 7. Conclusions

In this work we propose a new theoretical framework for pricing contingent claims and studying their static replication strategies utilizing basic financial instruments whose payoffs form frames. This generates standardized markets such as those for plain vanilla options, and frames provide the flexibility to study spaces of claims spanned by simpler securities. We provide a systematic scheme for pricing/hedging exotic derivatives including path-dependent options through a new means of static replication that can be implemented in markets with a reasonable spectrum of strikes on European options spanning practical trading ranges. Numerical studies on the pricing of various exotic options demonstrate that this method is accurate in comparison to existing methods proposed in the literature. Namely, our approach outperforms alternative methods based on interpolation as well as the integral representations of Carr and Madan [12], providing more accurate hedges and faster converging value approximations which reduce the required number of basis elements to achieve a desired accuracy. Future research includes extensions to higher dimensional payoffs, as well as the design of frames to capture particular model features with sparse representations.

## Appendix A. Pricing Experiment

This section illustrates the application of frame projection for model-based pricing of exotic European options. We consider several exponential Lévy models for the underlying of the form $S_T = S_0 e^{X_T}$. Here $S_T$ is a non-dividend paying asset ($q = 0$), and $X_T$ is a Lévy process with known characteristic function $\widehat{\mu}_T(\xi) = \exp(T\psi_{RN}(\xi))$, where $\psi_{RN}$ is the risk-neutral *(Lévy) symbol* (see [3] for a development of exponential Lévy-based modeling in finance). We assume that the symbol is chosen (e.g. martingale adjustment) so that arbitrage-free prices are obtained after discounting. In this case, efficient pricing methods such that of [4] (which is an efficient realization of the inverse Fourier transform), the Hilbert transform method [24] or the PROJ method [30, 33] can be used to price the butterfly basis instruments simultaneously.[25]

As an example, we consider the capped powered call

$$f(S_T) = (S_T - K)^2 \mathbb{1}_{[K,C]}(S_T) + (C - K)^2 \mathbb{1}_{(C,\infty)}(S_T),$$

with resolutions set as in the Henry Natural Gas example of section 6.3. Once hedges are formed (independently of the pricing method), we price each of the underlying vanillas, and determine the hedge payoff's price using the same set of prices across each method in order to isolate the difference induced by the hedge itself. For all pricing experiments, reference prices are obtained by pricing the basis elements (ie the corresponding vanilla contracts) using interpolation

[25]Utilizing the theory of frame projection, the methodology developed in [30] obtains orthogonal projections of risk-neutral return densities for processes with known characteristic functions, enabling the efficient pricing of any finitely valued claim. Extensions to exotic option pricing are considered in [17, 18, 31, 32, 34].

| Model | $\Delta$ | 1.00 | 0.50 | 0.20 | 0.10 | 0.05 |
|---|---|---|---|---|---|---|
| BSM | Interp | 7.717e-02 | 1.851e-02 | 2.779e-03 | 7.074e-04 | 1.768e-04 |
| | $ABS_2$ | 4.153e-03 | 1.526e-04 | 5.698e-06 | 3.356e-07 | 2.112e-08 |
| MJD | Interp | 7.019e-02 | 1.694e-02 | 2.559e-03 | 6.500e-04 | 1.624e-04 |
| | $ABS_2$ | 5.158e-03 | 1.906e-04 | 5.829e-06 | 3.564e-07 | 2.281e-08 |
| CGMY | Interp | 6.464e-02 | 1.573e-02 | 2.414e-03 | 6.106e-04 | 1.526e-04 |
| | $ABS_2$ | 3.627e-03 | 1.198e-04 | 3.557e-06 | 2.379e-07 | 3.663e-08 |

TABLE 6. Capped powered call: $f(S_T) = (S_T - K)^2 \mathbb{1}_{[K,C]}(S_T) + (C - K)^2 \mathbb{1}_{(C,\infty)}(S_T)$. Price errors at spacing $\Delta$, $S_0 = 3.07$, $K = 3.15$, $C = 15$, $T = 1$, $r = 0.02$, $q = 0$. Ref. values: $[0.5468885, 1.0902305, 2.4033236]$. BSM model: $\sigma = 0.3$; MJD model: $\sigma = 0.3, \mu_J = 0.15, \sigma_J = 0.2, \lambda_J = 0.7$; CGMY Model: $C = 2, M = 5, G = 6, Y = 0.5$

coefficients on a fine mesh. The underlying vanilla prices are verified to the reported accuracy using each of the methods [22, 24, 30] (for BSM, analytical vanilla option prices are used).

Valuation errors are reported in Table 6, and $ABS_2$ performs substantially better at all resolutions, with rapid value convergence. For all resolutions and for each method tested, the marginal cost after basis prices are obtained is less than 0.2 milliseconds[26]. Moreover, for all resolutions hedge construction for $ABS_2$ is within .06 milliseconds of interpolation. Hence, as the cost of pricing the basis is by far the dominant expense (see [23, 30] for pricing of multiple strikes in Lévy and Heston models), ABS is able to reduce the total computational cost by reducing the number of basis elements that are needed. For example, in the BSM test, to reach an accuracy of $10^{-4}$, interpolation requires more than 10 times as many basis elements as $ABS_2$.

*Remark* 8. In contrast to the case of static hedging, in which the basis resolution is limited by available strikes, for option pricing in a theoretical model, we are free to choose the resolution as fine as required, as well as the truncation interval. It is therefore advantageous to obtain error bounds as a function of these free parameters. We leave a detailed investigation of these issues for future research.

A.1. **BSM and MJD.** The Black-Scholes-Merton (BSM) model is described by its volatility, $\sigma$, and is represented by the Lévy symbol $\psi_L(\xi) = -\sigma^2\xi^2/2$, prior to martingale adjustment. For this model we consider the test case

$$\text{Test 1 (BSM)}: \quad \sigma = 0.3 \tag{37}$$

By adding a Poisson jump process with normally distributed jump sizes, we arrive at

$$\psi_L(\xi) = -\sigma^2\xi^2/2 + \lambda_J \left(\exp\left(i\mu_J\xi - \sigma_J^2\xi^2/2\right) - 1\right),$$

where $\sigma$ is the diffusion volatility, $\mu_J, \sigma_J$ are the jump size mean and volatility, and $\lambda_J$ the rate of jump arrivals. For this model, we consider the test case

$$\text{Test 2 (MJD)}: \quad \sigma = 0.3, \quad \mu_J = 0.15, \quad \sigma_J = 0.2, \quad \lambda_J = 0.7 \tag{38}$$

A.2. **KoBoL.** The CGMY subclass of the general KoBoL model [2, 3] forms a four parameter family of exponential Lévy models [9]: $C \geq 0$ accounts for the activity level of jumps, $G, M \geq 0$ determine the skewness, and $Y < 2$ dictates the fine structure, where $Y < 0$ specifies a finite activity process, $0 \leq Y \leq 1$ a process with finite variation but infinite activity, and $1 \leq Y < 2$ a process of infinite activity and variation. The risk-neutral Lévy symbol is given by

$$\psi_{RN}(\xi) = i\xi(r - \psi_L(-i)) + C\Gamma(-Y)\left((M - i\xi)^Y - M^Y + (G + i\xi)^Y - G^Y\right),$$

where

$$\psi_L(-i) = C\Gamma(-Y)\left((M - 1)^Y - M^Y + (G + 1)^Y - G^Y\right),$$

and $\Gamma = \Gamma(y)$ is the Gamma function. Our third test is the CGMY model with parameters

$$\text{Test 3 (CGMY)}: \quad C = 2, \quad M = 5, \quad G = 6, \quad Y = 0.5, \tag{39}$$

[26]For all experiments conducted, the code is written in MATLAB 8.1, and the computer has an Intel(R) Core(TM) i5-3470T CPU, 2.90GHz with 3MB cache size.

which generates a density with much heavier tails than the BSM or MJD, as depicted in Figure 4. In terms of the original KoBoL parameterization, we can express the Lévy symbol $\psi_L(\xi)$ as

$$c\Gamma(-\nu)[(\lambda_+ + i\xi)^\nu - \lambda_+^\nu + (-\lambda_- - i\xi)^\nu - (-\lambda_-)^\nu], \tag{40}$$

where $C = c$, $M = \lambda_+$, $G = -\lambda_-$, and $Y = \nu$.

**Algorithm 1** $\text{ABS}_2$ coefficients

$\bar{K} := (R - L)/a; \quad \lambda := a^{-1/2}/6; \quad \widetilde{K} := 4a(R - L) + 14$
$c_0 := 2; \quad c_1 := 5/12; \quad c_2 := -1/2; \quad c_3 := 1/12$
$\{d_k\}_{k=0}^7 = \{c_3, c_2, c_1, c_0, c_1, c_2, c_3\}$
$z_k := f(L - 7/4a + k/4a), \quad k = 0, ..., \widetilde{K}$
**for** $k = 0, ..., \widetilde{K}/2 - 1$ **do**
  $\theta_k \leftarrow \lambda \cdot [z_{2k} + z_{2k+1} + z_{2k+2}]$
**end for**
**for** $j = 0, ..., \widetilde{K}/2 - 1$ **do**
  $B_k = \sum_{k=0}^7 \theta_{k+2(j-1)} \cdot d_k$
**end for**
**return** $\{B\}_{k=0}^{\bar{K}}$

## Appendix B. Algorithms

To ease implementation, this section summarizes the $\text{ABS}_2$ algorithm and the Carr-Madan approach. The Dual method, which is more involved, is omitted. We find that the $\text{ABS}_2$ method provides the same level of accuracy at a fraction of the cost. Moreover, it behaves better in the presence of payoff discontinuities and is easier to implement. We assume a hedge support of $[L, R]$ and a resolution of $a > 0$. Since the market dictates the size of $a$, the only user-supplied parameters (for hedging) are $L$ and $R$, which depend on the application. Assuming that $L > 1/a$, we take $\bar{K} := (R - L)/a$ to be one less than the number of basis elements. The $\text{ABS}_2$ algorithm summarized in Algorithm 1 takes a payoff function $f$ and returns a set of coefficients $\{B_k\}_{k=0}^{\bar{K}}$. In this case, $B_0$ corresponds to the basis element centered at $L$, $B_1$ corresponds to $L + 1/a$, and $B_k$ corresponds to $L + k/a$.

**Algorithm 2** Carr-Madan coefficients

$\bar{K} := (R - L)/a$
**for** $k = 0, ..., \bar{K}$ **do**
  $\widetilde{B}_k = .5 \cdot [f'(L + (k + 1)/a) - f'(L + (k - 1)/a)]$
**end for**
**return** $\{\widetilde{B}_k\}_{k=0}^{\bar{K}}$

A basic implementation of the Carr Madan method [12] is provided in Algorithm 2 where we assume the first and second order derivatives of the (smooth) payoff are available, otherwise a finite difference approximation may be substituted. For this method, we supply $L, R$ the resolution ($a > 0$) and $F_0$, which is the $T$-forward price of the underlying. Algorithm 2 returns the positions in each option with strikes $L, L + 1/a, ..., R - 1/a, R$, where strikes less than $F_0$ are taken in put options, and strikes greater than $F_0$ are taken in call options (the corresponding positions are denoted by $\widetilde{B}_k$, to distinguish them from positions in butterfly basis elements). In addition, we hold $f(F_0)$ units in a bond, and $f'(F_0)$ forward contracts. A derivation of the approximation given in Algorithm 2 can be found in [45].

## Appendix C. Auxiliary Results

One way to obtain an expression for $\mathbf{\Phi}$ is via Fourier series (FS) expansion:

**Theorem C.1.** *[15] Let $\phi \in L^2(\mathbb{R})$. Then $\mathbf{\Phi} \in L^1(0, 2\pi)$, and the Fourier coefficients of $\mathbf{\Phi}$ with respect to the orthonormal basis $\{\frac{1}{\sqrt{2\pi}}e^{-ikx}\}_{k\in\mathbb{Z}}$ are given by*

$$c_k = \sqrt{2\pi}\int_{-\infty}^{\infty} \phi(x)\overline{\phi(x-k)}dx, \qquad k \in \mathbb{Z}.$$

*Moreover, when $\phi$ is a compactly supported real-valued function, the FS expansion has only finitely many terms:*

$$\mathbf{\Phi}(\gamma) = \frac{c_0}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}}\sum_{k=1}^{N} c_k \cos(k\gamma),$$

*for some $N \le \lceil |supp(\phi)| \rceil$.*

The following proposition concerns the alteration of projection coefficients and its effect on convergence. The proof is omitted.

**Proposition C.1.** *Let $(\phi, \widetilde{\phi})$ be a Riesz generator-dual pair in $\mathcal{H}$. Fix an initial spacing $1/a$ and define $\{\alpha_j\}_{j\ge 0}$ by $\alpha_j = 2^j a$. Choose any finite set of points $\{x^m\}$ of the form $x^m = k_m/\alpha_{j_m}$, where $k_m \in \mathbb{Z}$ and $j_m \le J^* \in \mathbb{Z}_+$. For each $m$, choose a bounded sequence $\{v_j^m\}_{j\ge j_m}$ of modifications. Then $\forall \epsilon > 0$, $\exists J \ge J^* \in \mathbb{Z}_+$, such that $\forall j' \ge J \in \mathbb{Z}_+$*

$$\|\widetilde{P}_{J'}f - P_{J'}f\|_2 < \epsilon, \quad \forall f \in \mathbb{Z},$$

*where $P_{J'}$ is the orthogonal projection of $\mathcal{H}$ onto $J' := \overline{span}\{D_{\alpha_{j'}}T_k\phi\}_{k\in\mathbb{Z}}$, and $\widetilde{P}_{J'}$ is obtained by replacing for each $m$ and $\forall j' \ge j_m$ the coefficients of $D_{2^{j'}a}T_{2^{j'-j_m}k_m}\phi$ (those of the basis elements centered over $x_m$ at each resolution above $j_m$) by $v_{j'}^m/(\alpha_{j'})^{1/2}$.*

## Appendix D. Proofs

*Proof of Proposition 3.1.* While this result is standard in the literature, the proof of (ii) highlights the essence of a multiresolution analysis. Accordingly, fix any $\epsilon > 0$, and any $f \in \mathcal{H}$. By condition (ii) of the definition of FMRA, there exists $j \in \mathbb{Z}$ such that for some $h \in U_j$, $\|f - h\|_2 < \epsilon/2$. By condition (i) of this proposition and the fact that $h \in U_j$, $P_{j'}h(x) = h(x)$ for all $j' \ge j$. Hence

$$\begin{aligned}\|f - P_{j'}f\|_2 &= \|f - h + P_{j'}h - P_{j'}f\|_2 \\ &\le \|f-h\|_2 + \|P_{j'}(f-h)\|_2 \le 2\|f-h\|_2 < \epsilon.\end{aligned}$$

The multiresolution structure ensures that approximations can only improve with mesh refinement. □

*Proof of Theorem 3.1.* We have for $f \in \mathcal{H}$

$$e^{rT}|\mathcal{V}f| \le \int_{-\infty}^{\infty}|f(e^y)|q_T(y)dy = \int_0^{\infty}|f(x)|\frac{q_T(\ln(x))}{x}dx \le \|f\|_2\left[\int_0^{\infty}\frac{q_T^2(\ln(x))}{x^2}dx\right]^{1/2},$$

so by a change of variables

$$e^{rT}|\mathcal{V}f| \le \|f\|_2\left[\int_{-\infty}^{\infty}\frac{q_T^2(z)}{e^z}dz\right]^{1/2} \le \|f\|_2\|q_T\|_{\infty}^{1/2}\left[\int_{-\infty}^{\infty}q_T(z)e^{-z}dz\right]^{1/2}.$$

Thus, $|\mathcal{V}f| \le C\|f\|_2$, where $C := e^{-rT}\left(\|q_T\|_{\infty}\hat{q}_T(i)\right)^{1/2}$, and $\hat{q}_T(i) = e^{-\ln(S_0)}\widehat{\mu}_T(i)$. Linearity is clear. □

*Proof of Theorem 3.2.* To verify the first claim, note that by Proposition 3.1, $\|P_jf\|_2 \le \|f\|_2$, so $f \in \mathcal{H}$ implies $|\mathcal{V}\circ P_jf| \le \|\mathcal{V}\|_2\|P_jf\|_2 \le \|\mathcal{V}\|_2\|f\|_2$, so $\mathcal{V}_j : \mathcal{H} \to \mathbb{C}$. In fact, by the Uniform Boundedness Principle, $\sup_j\|\mathcal{V}_j\|_2 < \infty$, so $\mathcal{V}_j$ is a uniformly bounded class in $\mathcal{H}^*$. Since continuous linear functionals preserve convergent series,

$$\mathcal{V}_jf = \mathcal{V}\Big(\lim_{K\to\infty}\sum_{|k|\le K}\langle f, T_{2^{-j}k}\theta_j\rangle\phi_{j,k}\Big) = \lim_{K\to\infty}\sum_{|k|\le K}\langle f, T_{2^{-j}k}\theta_j\rangle\mathcal{V}\circ\phi_{j,k},$$

by Theorem 3.1. To prove the second claim, $\mathcal{V}$ bounded implies that by the Riesz Representation theorem, $\exists h \in \mathcal{H}$ for which $\mathcal{V}f = \langle f, h\rangle, \forall f \in \mathcal{H}$. Hence,

$$\begin{aligned}|\mathcal{V}_j f - \mathcal{V}f| = |\mathcal{V} \circ (P_j - I)f| &= |\langle (P_j - I)f, h\rangle| \\ &= |\langle f, (P_j - I)h\rangle| \leq \|f\|_2 \|(P_j - I)h\|_2,\end{aligned}$$

since $P_j$ is self-adjoint. Taking the supremum over $f \in \mathcal{H}$, $\|f\|_2 = 1$, uniform convergence is obtained as $\|(P_j - I)h\|_2 \to 0$. □

*Proof of Corollary 3.3.* The first two claims are immediate from previous results. To prove (iii), fix any $f \in \mathcal{H}_+$. Then

$$\begin{aligned}\|P_j f - P_{j^+} f\|_2 &\leq \|(P_j f - P_{j^+} f)\mathbb{1}_{x\geq 0}\|_2 + \|(P_j f - P_{j^+} f)\mathbb{1}_{x<0}\|_2 \\ &= \|(P_j f - P_{j^+} f)\mathbb{1}_{x\geq 0}\|_2 + \|P_j f \mathbb{1}_{x<0}\|_2 \\ &= \|(\langle f, \widetilde{\phi}_{j,0}\rangle - f(0)2^{-j/2}) \cdot \phi_{j,0}(x) \cdot \mathbb{1}_{x\geq 0}\|_2 + \|(P_j f - f)\mathbb{1}_{x<0}\|_2 \\ &\leq \|(\langle f, \widetilde{\phi}_{j,0}\rangle - f(0)2^{-j/2}) \cdot \phi_{j,0}(x)\|_2 + \|P_j f - f\|_2,\end{aligned}$$

which converges to zero by proposition C.1 and (ii) of this proposition. Hence

$$\|P_{j^+} f - f\|_2 \leq \|P_{j^+} f - P_j f\|_2 + \|P_j f - f\|_2 \to 0, \text{ as } j \to \infty.$$

Similarly,

$$|\mathcal{V} \circ P_{j^+} f - \mathcal{V}f| = |\mathcal{V}(P_{j^+} - I)f| \leq \|\mathcal{V}\|_2 \|(P_{j^+} - I)f\|_2.$$

But $\|(P_{j^+} - I)f\|_2 \to 0$, and the claim follows. □

*Proof of Theorem 3.4.* To prove (i), note that $D_a T_k \widetilde{\phi}$ is a frame sequence with the same bounds $0 < A \leq B$ as for $\{T_k \phi\}$. Indeed, $\langle D_a T_m \phi, D_a T_k \widetilde{\phi}\rangle = \langle \phi, T_{k-m}\widetilde{\phi}\rangle$, and $\text{span}\{D_a T_m \phi\}_m$ is of course dense in $\mathcal{M}_a$, so the bounds remain valid. From the biorthogonality of $\{T_k\phi\}$ and $\{T_k\widetilde{\phi}\}$, $\delta_{m,k} = \langle T_m\phi, T_k\widetilde{\phi}\rangle = \langle D_a T_m \phi, D_a T_k \widetilde{\phi}\rangle$, and the result follows since there is at most one sequence in $\mathcal{M}_a$ biorthogonal to $\{D_a T_k \phi\}$. For (ii), we have by frame representation and Parseval's identity

$$P_{\mathcal{M}_a} f = \sum_k \langle f, \widetilde{\phi}_{a,k}\rangle \phi_{a,k} = \frac{1}{2\pi} \sum_k \langle \widehat{f}, \mathcal{F}[D_a T_k \widetilde{\phi}]\rangle \phi_{a,k}.$$

With the modulation operator $E_b : \mathcal{H} \to \mathcal{H}$ defined by $(E_b f)(x) = e^{ibx} f(x)$,

$$\begin{aligned}\langle \widehat{f}, \mathcal{F}[D_a T_k \widetilde{\phi}]\rangle = \langle \widehat{f}, D_{\frac{1}{a}} E_k \widehat{\widetilde{\phi}}\rangle = \langle D_a \widehat{f}, E_k \widehat{\widetilde{\phi}}\rangle &= a^{1/2} \int_{-\infty}^{\infty} \widehat{f}(a\xi) \overline{e^{ik\xi} \widehat{\widetilde{\phi}}(\xi)} d\xi \\ &= a^{1/2} \int_{-\infty}^{\infty} e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi) d\xi.\end{aligned}$$

The reality of $f$ and $\widetilde{\phi}$ then implies

$$\langle f, \widetilde{\phi}_{a,k}\rangle = \Re\langle f, \widetilde{\phi}_{a,k}\rangle = a^{1/2} \Re \int_{-\infty}^{\infty} e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi) d\xi.$$

But for $z \in \mathbb{C}$, $\Re z = \Re \bar{z}$, so upon splitting the integral

$$\begin{aligned}\langle f, \widetilde{\phi}_{a,k}\rangle &= a^{1/2} \Re \int_{-\infty}^{0} \overline{e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi)} d\xi + a^{1/2} \Re \int_0^{\infty} e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi) d\xi \\ &= 2a^{1/2} \int_0^{\infty} \Re \left[ e^{-ik\xi} \widehat{f}(a\xi) \widehat{\widetilde{\phi}}(-\xi) d\xi \right].\end{aligned}$$

□

*Proof of Corollary 4.1.* A simple calculation yields $\widehat{\varphi}(\xi) = 4\sin^2(\xi/2)/\xi^2$. By Theorem C.1 in the appendix, using the fact that $\varphi$ is compactly supported we derive $\mathbf{\Phi}(\xi) = \frac{1}{3}(1 + 2\cos^2(\xi/2))$, from which

$$\widehat{\widetilde{\varphi}}(\xi) = \widehat{\varphi}(\xi)\mathbf{\Phi}(\xi)^{-1} = \frac{12\sin^2(\xi/2)}{\xi^2(1 + 2\cos^2(\xi/2))} = \frac{12\sin^2(\xi/2)}{\xi^2(2 + \cos(\xi))}.$$

Finally, noting that $\widehat{\widetilde{\varphi}}(\xi) = \widehat{\widetilde{\varphi}}(-\xi)$, the result follows from Theorem 3.4. □

*Proof of Corollary 4.2.* Since $\phi$ is real-valued and symmetric, the same is true of $\widehat{\phi}$. Further

$$\mathbf{\Phi}(\xi) = \sum_k |\widehat{\phi}(\xi + 2\pi k)|^2 = \sum_k |\widehat{\phi}(-\xi - 2\pi k)|^2 = \sum_k |\widehat{\phi}(-\xi + 2\pi k)|^2 = \mathbf{\Phi}(-\xi),$$

so $\widehat{\widetilde{\phi}}$ and hence $\widetilde{\phi}$ is real-valued and symmetric. The fact that $\widetilde{\phi}$ is the unique biorthogonal dual implies that $\widetilde{\phi} \in \overline{\text{span}}\{T_k\phi\}$, from which the representation $\widetilde{\phi} = \sum_k \langle \widetilde{\phi}, T_k\widetilde{\phi}\rangle T_k\phi$ is valid. Hence,

$$\langle \widetilde{\phi}, T_k\widetilde{\phi}\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{\widetilde{\phi}}(\xi)\overline{\widehat{\widetilde{\phi}}(\xi)e^{i\xi k}}d\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{\widetilde{\phi}}(\xi)\widehat{\widetilde{\phi}}(-\xi)e^{-i\xi k}d\xi.$$

Moreover, $\widetilde{\phi}$ real-valued implies $\langle \widetilde{\phi}, T_k\widetilde{\phi}\rangle = \Re\langle \widetilde{\phi}, T_k\widetilde{\phi}\rangle\ \forall k$. By the symmetry and reality of $\widehat{\widetilde{\phi}}$, $\Re[\widehat{\widetilde{\phi}}(\xi)\widehat{\widetilde{\phi}}(-\xi)e^{-i\xi k}] = (\widehat{\widetilde{\phi}}(\xi))^2\Re(e^{-i\xi k})$. Thus

$$\langle \widetilde{\phi}, T_k\widetilde{\phi}\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\widehat{\phi}^2(\xi)}{\mathbf{\Phi}^2(\xi)}\Re(e^{-i\xi k})d\xi = \frac{1}{\pi}\int_0^{\infty} \frac{\widehat{\phi}^2(\xi)}{\mathbf{\Phi}^2(\xi)}\cos(\xi k)d\xi,$$

since the integrand is even. □

*Proof of Proposition 4.1.* From biorthogonality, $\widetilde{\varphi}$ necessarily satisfies $\int_{\mathbb{R}} = \widetilde{\varphi}(x)\varphi(x-k)dx = 0$ for $|k| \geq 1$, or equivalently

$$\begin{aligned}
0 &= \int_{\mathbb{R}} \left(\sum_{m\in\mathbb{Z}} \alpha_m \varphi(x-m)\right)\varphi(x-k)dx \\
&= \alpha_{k-1}\int_{k-1}^{k} T_{k-1}\varphi \cdot T_k\varphi + \alpha_k \int_{k-1}^{k+1} T_k\varphi \cdot T_k\varphi + \alpha_{k+1}\int_{k}^{k+1} T_{k+1}\varphi \cdot T_k\varphi \\
&= \frac{1}{6}(\alpha_{k-1} + 4\alpha_k + \alpha_{k+1}),
\end{aligned}$$

so $\alpha_{k+1} = -4\alpha_k - \alpha_{k-1}$. If we posit the ansatz $\alpha_k = \alpha_0\nu^{|k|}$ for $|k| \geq 1$, the difference equation becomes $\alpha_0\nu^2 + 4\alpha_0\nu + \alpha_0 = 0$, which has a stationary (non-divergent) solution $\nu = \sqrt{3} - 2$. When $k = 0$, we can use the fact that $\alpha_{-1} = \alpha_1$ and the biorthogonality relation $\int_{\mathbb{R}} = \widetilde{\varphi}(x)\varphi(x)dx = 1$ to obtain the equation $\alpha_1 = 3 - 2\alpha_0$. Combined with the ansatz $\alpha_1 = \alpha_0\nu$, we find that $\alpha_0 = 3/\sqrt{3}$. Uniqueness of the biorthogonal dual gives the result, after verifying that $\widetilde{\varphi} \in L^2$ by a geometric series, with $\widetilde{\varphi}$ defined by equation (10). □

*Proof of Proposition 4.2.* We prove the case of $f \in \bar{C}$, the other being entirely similar. With $\epsilon_{a,k} := |\beta_{a,k} - \widehat{\beta}_{a,k}| = |\int_{[\bar{L},\bar{R}]^c} f \cdot \widetilde{\varphi}_{a,k}|$,

$$\begin{aligned}
\|\bar{P}_{\mathcal{M}_a} f - \widehat{P}_{\mathcal{M}_a} f_{[\bar{L},\bar{R}]}\|_2^2 &\leq \int \left(\sum_{k_L}^{k_R} \epsilon_{a,k}\varphi_{a,k}(x)\right)^2 dx \\
&\leq \int \left[\left(\sum_{k_L}^{k_R} \epsilon_{a,k}^2\right)^{1/2}\left(\sum_{k_L}^{k_R} \varphi_{a,k}(x)^2\right)^{1/2}\right]^2 dx \\
&= \sum_{k_L}^{k_R} \epsilon_{a,k}^2 \cdot \sum_{k_L}^{k_R}\int \varphi_{a,k}(x)^2 dx \leq \epsilon_a^* \cdot \frac{2}{3}a^2(R-L+1)^2,
\end{aligned}$$

where $\epsilon_a^* := \max_{k_L \le k \le k_R}\{\epsilon_{a,k}^2\}$. By Corollary 4.1, $\widetilde{\varphi}_{a,k} = \alpha_0 \sum_{m\in\mathbb{Z}} \nu^{|m|}\varphi_{a,k+m}$, from which we determine a bound on $\epsilon_a^*$. With $\bar{A}_k := \{m \in \mathbb{Z} : m + k \le k_{\bar{L}} \text{ or } m + k \ge k_{\bar{R}}\}$

$$\begin{aligned}
\epsilon_{a,k} &\le \alpha_0 \sum_{m\in\bar{A}_k} |\nu|^{|m|} \int |f(x)\varphi_{a,k+m}(x)|dx \\
&\le \sqrt{2} \sum_{m\in\bar{A}_k} |\nu|^{|m|} \|f\|_2^{I^a_{k+m}} \\
&\le \sqrt{2}\|f\|_2^{I^a_k} \sum_{m\in\bar{A}_k} |C\nu|^{|m|} \le \kappa \cdot \tau([\bar{L},\bar{R}]), \quad \forall k_L \le k \le k_R,
\end{aligned}$$

where $\kappa := \sqrt{2}\|f\|_2^{[L-1/a,R+1/a]}$ is independent of $[\bar{L},\bar{R}]$, and decreases in $a$, and $\tau([\bar{L},\bar{R}]) := \sum_{m\in\bar{A}_k} |C\nu|^{|m|}$ is finite since $|C\nu| < 1$. Hence, we can choose $[\bar{L},\bar{R}] = [k_{\bar{L}}/a, k_{\bar{R}}/a]$ such that this tail series is arbitrarily small. Finiteness of the coefficients, hence of the representation, follows similarly by splitting the expression for $\beta_{a,k}$ into a finite component and a convergent tail series. □

*Proof of Corollary 4.3.* From Proposition 4.1 with $\alpha_0 := 3/\sqrt{3}$,

$$\int_R \widetilde{\varphi}(x)x^p dx = \alpha_0 \sum_{m\in\mathbb{Z}} \nu^{|m|} \int_R \varphi(x-m)x^p dx = \alpha_0 \sum_{m\in\mathbb{Z}} \nu^{|m|}\theta_{1,m},$$

where $\theta_{1,m}$ is derived from equation (16). Thus, with $c_p := \alpha_0/(p+1)(p+2)$,

$$\begin{aligned}
M^p &= c_p \sum_{m\in\mathbb{Z}} \nu^{|m|} \left[(m-1)^{p+2} - 2m^{p+2} + (m+1)^{p+2}\right] \\
&= c_p \left(1 + (-1)^{p+2}\right) \left(1 + \sum_{m=1}^{\infty} \nu^m \left[(m-1)^{p+2} - 2m^{p+2} + (m+1)^{p+2}\right]\right) \\
&= c_p \left(1 + (-1)^{p+2}\right) \cdot \left(1 + \nu G^{p+2} - 2G^{p+2} + \nu^{-1}(G^{p+2} - \nu)\right),
\end{aligned}$$

where $G^q := \sum_{m=0}^{\infty} \nu^m m^q$. The result follows upon noting that $(1-\nu)G^q$ is the $q$th moment of a geometric random variable with success probability $1-\nu$, which is found by differentiating its moment generating function. To determine the coefficients,

$$\begin{aligned}
\int_R \widetilde{\varphi}_{a,k}(x)x^p dx &= a^{1/2} \int \widetilde{\varphi}(a(x-k/a))x^p \\
&= a^{-(p+\frac{1}{2})} \int \widetilde{\varphi}(x)(x+k)^p dx = a^{-(p+\frac{1}{2})} \sum_{n=0}^{p} \binom{p}{n} k^{p-n} \int \widetilde{\varphi}(x)x^n dx,
\end{aligned}$$

by a change of variables and binomial expansion. The result follows after eliminating odd moments. □

*Proof of Proposition 5.1.* We derive the $L^2$ error, with the $L^1$ error following similarly. Note first that

$$|\langle f, \varphi_{a,k}\rangle| \le \|f\|_\infty^{[L,R]} \int_{[\frac{k-1}{a}, \frac{k+1}{a}]} \varphi_{a,k}(s)ds = a^{-1/2}\|f\|_\infty^{[L,R]}.$$

Moreover, from equation (10) we have by a geometric series

$$\sum_{|m|>\gamma} |\alpha_m|^2 = 2\sum_{m>\gamma} \left(\frac{3}{\sqrt{3}}\right)^2 \nu^{2m} = \frac{6}{1-\nu^2}\nu^{2(\gamma+1)} = \tau(\gamma)^2.$$

Hence,

$$\begin{aligned}
\|f_a^\gamma - P_{\mathcal{M}_a} f\|_2^2 &= \int \Big| \sum_{k_L}^{k_R} \Big( \sum_{|m|\le\gamma} \alpha_m \langle f, \varphi_{a,k+m}\rangle \Big) \varphi_{a,k}(s) - \beta_{a,k}\varphi_{a,k}(s) \Big|^2 ds \\
&\le \int \left( \sum_{k_L}^{k_R} \Big| \sum_{|m|\le\gamma} \alpha_m \langle f, \varphi_{a,k+m}\rangle - \beta_{a,k} \Big| |\varphi_{a,k}(s)| \right)^2 ds \\
&\le \int \left( \sum_{k_L}^{k_R} |\varphi_{a,k}(s)| \sum_{|m|>\gamma} |\alpha_m| |\langle f, \varphi_{a,k+m}\rangle| \right)^2 ds \\
&\le \int \left( \sum_{k_L}^{k_R} |\varphi_{a,k}(s)| \left( \sum_{|m|>\gamma} |\alpha_m|^2 \right)^{1/2} \left( \sum_{|m|>\gamma} |\langle f, \varphi_{a,k+m}\rangle|^2 \right)^{1/2} \right)^2 ds \\
&\le \Big( \tau(\gamma) \|f\|_\infty^{[L,R]} \Big)^2 (R-L+1) \int \left( \sum_{k_L}^{k_R} |\varphi_{a,k}(s)| \right)^2 ds.
\end{aligned}$$

Expanding the square,

$$\begin{aligned}
\int \left( \sum_{k_L}^{k_R} |\varphi_{a,k}(s)| \right)^2 &\le \sum_{k_L}^{k_R} \int \varphi_{a,k}^2(s) ds + \sum_{\substack{k\neq j \\ k_L \le k,j \le k_R}} \int \varphi_{a,k}(s)\varphi_{a,j}(s) ds \\
&\le \frac{4}{3} a (R-L+1).
\end{aligned}$$

□

*Proof of Lemma 5.1.* Since $\mathcal{M} := \overline{\text{span}}\{\Psi_k\}_{k\in\mathcal{K}}$ is a closed subspace of $\mathcal{H}$, there exists a unique decomposition $\mathcal{H} = \mathcal{M} \bigoplus \mathcal{M}^\perp$, which we denote by $f = P_{\mathcal{M}} f + P_{\mathcal{M}^\perp} f$. Then for any $f \in \mathcal{H}$ and $k$ given, $\langle f, \Psi_k\rangle = \langle P_{\mathcal{M}} f, \Psi_k\rangle$. Thus if $B$ denotes the upper frame bound on $\mathcal{M}$,

$$\sum_{k\in\mathcal{K}} |\langle f, \Psi_k\rangle|^2 = \sum_{k\in\mathcal{K}} |\langle P_{\mathcal{M}} f, \Psi_k\rangle|^2 \le B \|P_{\mathcal{M}} f\|_2^2 \le B \|f\|_2^2,$$

since $P_{\mathcal{M}} f \in \mathcal{M}$. □

*Proof of Proposition 5.3.* To simplify notation, let $\breve{P} := \breve{P}_{\mathcal{M}_a}$. Idempotence of $\breve{P}$ at all scales is demonstrated in equation (29). As for continuity, the compact support of $\breve{\phi}$ implies that $\sum_k |\langle \breve{\phi}_{a,j}, \breve{\phi}_{a,k}\rangle| \le B$ from some $B > 0$ and for all $j \in \mathbb{Z}$. Hence by Proposition 5.2

$$\|\{\langle f, \breve{\phi}_{a,k}\rangle\}\|_{l^2(\mathbb{Z})}^2 = \sum_k |\langle f, \breve{\phi}_{a,k}\rangle|^2 \le B\|f\|_2^2 < \infty \quad \forall f \in \mathcal{H},$$

so $\{\langle f, \breve{\phi}_{a,k}\rangle\} \in l^2(\mathbb{Z})$, and the mapping $T_1 : f \to \{\langle f, \breve{\phi}_{a,k}\rangle\}$ is a well-defined, bounded linear map of $\mathcal{H}$ into $l^2(\mathbb{Z})$. The fact that $\{\phi_{a,k}\}$ is a Riesz sequence, and hence a Bessel sequence on all of $\mathcal{H}$ by Lemma 5.1, implies that $\sum_k c_k \phi_{a,k}$ converges unconditionally $\forall \{c_k\} \in l^2(\mathbb{Z})$. Moreover, $T_2 : \{c_k\} \to \sum_k c_k \phi_{a,k}$ is a well-defined bounded linear mapping of $l^2(\mathbb{Z})$ into $\mathcal{H}$. Hence the composition $\breve{P} = T_2 \circ T_1$ is a bounded linear map from $\mathcal{H}$ into $\mathcal{M}_a$. In fact, idempotence implies that $\breve{P}$ is onto $\mathcal{M}_a$.

Denoting the unique canonical dual generator by $\widetilde{\phi}$, the *orthogonal* projection of $\mathcal{H}$ onto $\mathcal{M}_a$ is given by $P_{\mathcal{M}_a} f = \sum_k \langle f, \widetilde{\phi}_{a,k}\rangle \phi_{a,k}$. If $f \in \mathcal{M}$, $f = P_{\mathcal{M}_a} f$ so by biorthogonality

$$\langle f, \breve{\phi}_{a,m}\rangle = \langle P_{\mathcal{M}_a} f, \breve{\phi}_{a,m}\rangle = \Big\langle \sum_k \langle f, \widetilde{\phi}_{a,k}\rangle \phi_{a,k}, \breve{\phi}_{a,m} \Big\rangle = \langle f, \widetilde{\phi}_{a,m}\rangle,$$

so that $\breve{P} f = P_{\mathcal{M}_a} f = f \ \forall f \in \mathcal{M}_a$. Thus we have for any $f, g \in \mathcal{H}$

$$\langle f, \breve{P} g\rangle = \langle P_{\mathcal{M}_a} f, \breve{P} g\rangle + \langle P_{\mathcal{M}_a^\perp} f, \breve{P} g\rangle = \langle P_{\mathcal{M}_a} f, \breve{P} g\rangle = \langle P_{\mathcal{M}_a} f, \breve{P}(P_{\mathcal{M}_a} g + P_{\mathcal{M}_a^\perp} g)\rangle.$$

Therefore $\langle f, \breve{P} g\rangle = \langle P_{\mathcal{M}_a} f, P_{\mathcal{M}_a} g + \breve{P} P_{\mathcal{M}_a{}^\perp} g\rangle$.

Although $P_{\mathcal{M}_a^\perp} g \in \mathcal{M}_a^\perp \subset \text{Range}(\breve{P})^\perp \Rightarrow \langle \breve{P} P_{\mathcal{M}_a^\perp} g, P_{\mathcal{M}_a^\perp} g\rangle = 0$, it is not true in general that $\breve{P} P_{\mathcal{M}_a^\perp} g = 0$. But $\langle P_{\mathcal{M}_a} f, \breve{P} P_{\mathcal{M}_a^\perp} g\rangle = \langle f, \breve{P} P_{\mathcal{M}_a^\perp} g\rangle$, so in order for $\langle f, \breve{P} g\rangle = \langle P_{\mathcal{M}_a} f, \breve{P} P_{\mathcal{M}_a} g\rangle$, it is necessary and sufficient that $\breve{P} P_{\mathcal{M}_a^\perp}(\mathcal{H}) = \{0\}$. In this case, $\langle f, \breve{g}\rangle = \langle P_{\mathcal{M}_a} f, P_{\mathcal{M}_a} g\rangle$, and likewise for the equality $\langle P_{\mathcal{M}_a} f, P_{\mathcal{M}_a} g\rangle = \langle \breve{P} f, g\rangle$, which is equivalent to self-adjointness of $\breve{P}$. Since idempotence holds, this is equivalent to $\breve{P} = P_{\mathcal{M}_a}$.

Finally, $\breve{P}(\mathcal{M}^\perp) = \{0\}$ is in turn equivalent to $\breve{\phi} \in \mathcal{M}$. Indeed, if $\breve{\phi} \in \mathcal{M}$, biorthogonality implies that $\{\breve{\phi}_{1,k}\}$ is the unique dual which satisfies $f = \breve{P} f + (I - \breve{P}) f = P_{\mathcal{M}} f + P_{\mathcal{M}^\perp} f$. Hence $\breve{P}(\mathcal{M}^\perp) = P_{\mathcal{M}}(\mathcal{M}_a^\perp) = \{0\}$. Conversely, if $\breve{P}(\mathcal{M}^\perp) = \{0\}$, then $\sum_k \langle f, \breve{\phi}_{1,k}\rangle \phi_{1,k} = 0 \ \forall f \in \mathcal{M}^\perp$. But $\{\phi_{1,k}\}$ is a Riesz sequence, so $\omega$-independence implies $\langle f, \breve{\phi}_{1,k}\rangle = 0 \ \forall k \in \mathbb{Z}$. Hence, $\{\breve{\phi}_{1,k}\} \in (\mathcal{M}^\perp)^\perp = \mathcal{M}$ since $\mathcal{M}$ is closed, so $\breve{\phi} \in \mathcal{M}$. □

*Proof of Proposition 5.4.* By symmetry, the biorthogonality condition $\langle \breve{\varphi}^{[\gamma]}(x), \varphi(x-k)\rangle = 0$ for $|k| \geq 1$ is equivalent to

$$0 = \int \breve{\varphi}^{[\gamma]}(x)\varphi(x-k)dx = \sum_{m=0}^{2\gamma-1} c_m^{[\gamma]} \lambda_{|m-k|}, \quad 1 \leq k \leq \gamma,$$

where $\lambda_{|m-k|} := \langle \varphi(x-m), \varphi(2x-k)\rangle$ can be shown to satisfy equation (25), and $\lambda_j = 0$ for $j \geq 3$. This gives the second through fourth equations above, which represent $\gamma$ equations in all. The first equation is derived similarly from $\langle \breve{\varphi}^{[\gamma]}(x), \varphi(x)\rangle = 1$, bringing the total to $\gamma + 1$ equations. Since $2\gamma - 1$ coefficients are needed, there are $\gamma - 1$ remaining degrees of freedom, which are consumed by the moment matching conditions:

$$M^{2k} = \int \breve{\varphi}^{[\gamma]}(x) x^{2k} dx = \frac{1}{\sqrt{2}} \left( c_0^{[\gamma]} \int \varphi(2x) x^{2k} dx + 2 \sum_{m=1}^{2\gamma-1} c_m^{[\gamma]} \int \varphi(2x-m) x^{2k} dx \right),$$

where the integrals are evaluated using equation (16). Linear independence is then easily verified from the corresponding matrix.

To verify the final claim, Let $f(x) := \sum_{k_L}^{k_R} \varphi_{1,k}(x) \in \mathcal{M}_1$, where $k_L := -\gamma, k_R := \gamma$. From Proposition 5.3, any ABS approximation is a projector onto $\mathcal{M}_1$, so in particular with $a = 1$ the coefficient corresponding to $k = 0$ in the $\text{ABS}_\gamma$ approximation satisfies

$$1 = \int f \breve{\varphi}^{[\gamma]} = \int_{-\gamma}^{\gamma} \sum_{k_L}^{k_R} \varphi_{1,k} \varphi^{[\gamma]} = \int_{-\gamma}^{\gamma} \sum_{k \in \mathbb{Z}} \varphi_{1,k} \breve{\varphi}^{[\gamma]} = \int_{-\gamma}^{\gamma} \breve{\varphi}^{[\gamma]} = \int_{\mathbb{R}} \breve{\varphi}^{[\gamma]},$$

where we have used $1 = \sum_{k \in \mathbb{Z}} \varphi_{1,k}$. Thus $\int_{\mathbb{R}} \breve{\varphi}^{[\gamma]} = M^0$ is satisfied for any $\gamma$. For $\gamma \geq 1$ fixed, by Corollary 4.3 with $p \leq 2\gamma - 1$,

$$\begin{aligned} \int x^p \widetilde{\varphi}_{a,k}(x) dx &= a^{-(p+\frac{1}{2})} \int (x+k)^p \widetilde{\varphi}(x) dx = a^{-(p+\frac{1}{2})} \sum_{n=0}^{\lfloor \frac{p}{2} \rfloor} \binom{p}{2n} k^{p-2n} M^{2n} \\ &= a^{-(p+\frac{1}{2})} \sum_{n=0}^{\lfloor \frac{p}{2} \rfloor} \binom{p}{2n} k^{p-2n} \int x^{2n} \breve{\varphi}^{[\gamma]}(x) dx = \int x^p \breve{\varphi}_{a,k}^{[\gamma]}(x) dx, \end{aligned}$$

which follows upon equating the moments, and then reversing the change of variables and binomial expansion. □

*Proof of Proposition 5.5.* By compactness of $\breve{\phi}$, we've seen that $\sum_k |\langle T_k \breve{\phi}, T_m \breve{\phi}\rangle| \leq B$ for some $B > 0$ and $\forall m \in \mathbb{Z}$. But

$$\sum_k |\langle T_k \breve{\phi}, T_m \breve{\phi}\rangle| = \sum_k |\langle D_j T_k \breve{\phi}, D_j T_m \breve{\phi}\rangle| \leq B,$$

from which

$$\|\{\langle f, \breve{\phi}_{j,k}\rangle\}\|_{l^2(\mathbb{Z})}^2 = \sum_k |\langle f, \breve{\phi}_{j,k}\rangle|^2 \leq B \|f\|_2^2, \quad \forall f \in \mathcal{H},$$

where $B$ is constant $\forall j \in \mathbb{Z}$. Hence the continuous linear map $A_{1,j} : f \to \{\langle f, \breve{\phi}_{j,k}\rangle\}$ is point-wise bounded over $f \in \mathcal{H}$, uniformly in $j \in \mathbb{Z}$. By the uniform boundedness principle (UBP), $\sup_{j\in\mathbb{Z}}\|A_{1,j}\| \le A_1$, for some $A_1 > 0$. Likewise, the fact that $\phi$ generates an RMRA implies the map $A_{2,j} : \{c_k\} \to \sum_k c_k\phi_{j,k}$ for $\{c_k\} \in l^2(\mathbb{Z})$ is point-wise bounded for each $f \in \mathcal{H}$, uniformly over $j \in \mathbb{Z}$. Again, the UBP implies $\sup_{j\in\mathbb{Z}}\|A_{2,j}\| \le A_2$, for some $A_2 > 0$. By composition, $\breve{P}_j = A_{2,j} \circ A_{1,j}$, and $\|\breve{P}_j\| \le \|A_{2,j}\|\|A_{1,j}\| \le A_2A_1$, so $\sup_{j\in\mathbb{Z}}\|\breve{P}_j\| \le A_2A_1$. Thus, $\breve{P}_j$ is a bounded, linear projection operator on $\mathcal{H}$.

Now let $P_j$ denote the orthogonal projection $P_jf = \sum_k\langle f, \widetilde{\phi}_{j,k}\rangle\phi_{j,k}$, and fix any $\epsilon > 0$. The fact that $\phi$ is an RMRA generator implies the existence of $j \in \mathbb{Z}$ such that, for some $h \in U_j$, $\|f - h\|_2 < \epsilon(1 + A_1A_2)/2$ for any $f \in \mathcal{H}$. Moreover, the RMRA structure ensures that $h \in U_{j'}$ for all $j' \ge j$, hence $\|f - P_{j'}f\|_2 \le \epsilon/(1 + A_1A_2)$, $\forall j' \ge j$ by the proof of proposition 3.1. Thus

$$\begin{aligned}\|\breve{P}_{j'}f - f\|_2 &\le \|\breve{P}_{j'}f - \breve{P}_{j'}P_{j'}f\|_2 + \|\breve{P}_{j'}P_{j'}f - f\|_2\\ &= \|\breve{P}_{j'}(I - P_{j'})f\|_2 + \|(P_{j'} - I)f\|_2\\ &\le (\|\breve{P}_{j'}\| + 1)\|(I - P_{j'})f\|_2 \le (A_2A_1 + 1)\|(I - P_{j'})f\|_2 < \epsilon,\end{aligned}$$

for all $j' \ge j$. Thus $\lim_{j\to\infty}\|\breve{P}_jf - f\|_2 = 0$, and $\lim_{j\to\infty}|\mathcal{V} \circ \breve{P}_jf - \mathcal{V}f| = 0$ follows from previous arguments. From the proof of Theorem 3.1, for some $\zeta > 0$ we have $|\mathcal{V} \circ \breve{P}_jf - \mathcal{V}f| \le \zeta\|\breve{P}_{j'}f - f\|_2 \le \zeta(A_2A_1 + 1)\|(f - P_{j'}f\|_2$. □

## References


[1] G. Bakshi and D. Madan. Spanning and derivative-security valuation. *Journal of Financial Economics*, 55:205–238, 2000.

[2] S. Boyarchenko and S. Levendorskii. Option pricing for truncated Levy processes. *International J. of Theoretical and Applied Finance*, 3:549–552, 2000.

[3] S.I. Boyarchenko and S.Z. Levendorskii. *Non-Gaussian Merton-Black-Scholes Theory*. World Scientific Publishing Co., River Edge, NJ, volume 9 of adv. ser. stat. sci. appl. probab. edition, 2002.

[4] S.I. Boyarchenko and S.Z. Levendorskii. Efficient variations of Fourier transform in applications to option pricing. *Journal of Computational Finance*, 18(2):57–90, 2015.

[5] D. Breeden and R. Litzenberger. Prices of state contingent claims implicit in option prices. *Journal of Business*, 51:621–651, 1978.

[6] P. Carr and A. Chou. Breaking barriers. *Risk*, 10:139–145, 1996.

[7] P. Carr and A. Chou. Hedging complex barrier options. Working paper, 2002.

[8] P. Carr, K. Ellis, and V. Gupta. Static hedging of exotic options. *Journal of Finance*, 53:1165–1190, 1998.

[9] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *J. Business*, 75:305–332, 2002.

[10] P. Carr and R. Lee. Put-call symmetry: Extensions and applications. *Mathematical Finance*, 19(4):523–560, October 2009.

[11] P. Carr, K. Lewis, and D. Madan. On the nature of options. Working Paper, 2000.

[12] P. Carr and D. Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 1:19–37, 2001.

[13] P. Carr and J. Picron. Static hedging of timing risk. *Journal of Derivatives*, 6, No.3:57–70, Spring 1999.

[14] P. Carr and L. Wu. Static hedging of standard options. *Journal of Financial Econometrics*, 12(1):3–46, 2014.

[15] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhauser Boston, 2003.

[16] I.J. Clark. *Foreign Exchange Option Pricing: A Practitioners Guide*. The Wiley Finance Series, John Wiley & Sons, Ltd, Chichester, UK, 2011.

[17] Z. Cui, J.L. Kirkby, and D. Nguyen. Equity-linked annuity pricing with cliquet-style guarantees in regime-switching and stochastic volatility models with jumps. *Insurance: Mathematics and Economics*, 74:46–62, 2017.

[18] Z. Cui, J.L. Kirkby, and D. Nguyen. A general framework for discretely sampled realized variance derivatives in stochastic volatility models with jumps. *European Journal of Operational Research*, 262(1):381–400, 2017.

[19] S. Darolles and J. Laurent. Approximating payoffs and pricing formulas. *Journal of Economic Dynamics and Control*, 24:1721–1746, 2000.

[20] E. Derman, D. Ergener, and I. Kani. Static options replication. *Journal of Derivates*, 2:78–95, 1995.

[21] B. Engelmann, M. Fengler, M. Nalholm, and P. Schwendner. Static versus dynamic hedges: an empirical comparison for barrier options. *Review of Derivative Research*, 9(3):239–264, 2006.

[22] F. Fang and C.W. Oosterlee. A novel pricing method for European options based on Fourier cosine series expansions. *SIAM J. Sci. Comput.*, 31:826–848, 2008.

[23] F. Fang and C.W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik*, 114:27–62, 2009.

[24] L. Feng and V. Linetsky. Pricing discretely monitored barrier options and defaultable bonds in Levy process models: a fast Hilbert transform approach. *Math. Finan.*, 18(3):337384, July 2008.
[25] J. Gatheral. *The Volatility Suface: a Practicioner's Guide*. John Wiley and Sons, Inc., New Jersey, 2006.
[26] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, 2002.
[27] E.G. Haug. *The Complete Guide to Option Pricing Formulas*. McGraw-Hill, second edition, 2006.
[28] C. Heil. *A Basis Theory Primer*. Birkhauser, expanded edition edition, 2011.
[29] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Studies*, 6:327–343, 1993.
[30] J.L. Kirkby. Efficient option pricing by frame duality with the fast Fourier transform. *SIAM J. Financial Mathematics*, 6(1):713–747, 2015.
[31] J.L. Kirkby. Robust barrier option pricing by frame projection under exponential Levy dynamics. Applied Math. Fin., Forthcoming, 2015.
[32] J.L. Kirkby. An efficient transform method for Asian option pricing. *SIAM J. Financial Mathematics*, 7(1):845–892, 2016.
[33] J.L. Kirkby. Robust option pricing with characteristic functions and the B-spline order of density projection. *J. Computational Finance*, 21(2):61–100, 2017.
[34] J.L. Kirkby, D. Nguyen, and Z. Cui. A unified approach to Bermudan and barrier options under stochastic volatility models with jumps. *Journal of Economic Dynamics and Control*, 80:75–100, 2017.
[35] Levendorskii. Efficient pricing and reliable calibration in the heston model. *International Journal of Theoretical and Applied Finance*, 15(7):12050 (44 pages), 2012.
[36] F. Milne and D. Madan. Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance*, 4(3):223–245, 1994.
[37] M. Nalholm and R. Poulsen. Static hedging and model risk for barrier options. *The Journal of Futures Markets*, 26(5):449–463, 2006.
[38] A. Neuberger. The log contract. *J. of Portfolio Management*, 20(2):74–80, 1994.
[39] R. Poulsen. Barrier options and their static hedges: simple derivations and extensions. *Quantitative Finance*, 6(4):327–335, August 2006.
[40] D. Shimko. Non linear risk management. *Financial Derivaties and Risk Management*, 1(4):36–38, 1995.
[41] R. Tompkins. Static versus dynamic hedging of exotic options: an evaluation of hedge performance via simulation. *Journal of Risk Finance*, 3(4):6–34, 2002.
[42] R. Tompkins. Power options: Hedging nonlinear risks. *The Journal of Risk*, 2, No.2:29–45, Winter 1999/2000.
[43] M. Unser and I. Daubechies. On the approximation power of convolution-based least squares versus interpolation. *IEEE Transactions on Signal Processing*, 45(7):1697–1711, July 1997.
[44] Michael Unser. Vanishing moments and the approximation power of wavelet expansions. In *Proceedings of the 1996 IEEE international conference on image processing*, 1996.
[45] S. Oren Y. Oum and S. Deng. Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Research Logistics*, 53:697–712, 2006.
[46] R. Young. *An Introduction to Nonharmonic Fourier Series*. Academic Press, New York, (revised) edition, 1980.

## Appendix A. Supplemental Material: Transform-Based Method

Direct application of Corollary 4.1 leads to a procedure which applies when the payoff form of $f(S_T)$ is known[27] on an interval $[\bar{L}, \bar{R}]$ containing $[L, R]$, the desired hedge support. While this method is not utilized in the present work, numerical experiments have been conducted to verify its accuracy, and it extends easily to alternative bases. We assume that for some $k_L$, $k_R \in \mathbb{N}$, the endpoints satisfy $L = k_L/a$, $R = k_R/a$. Two applications of the fast Fourier transform (FFT) are required, where the desired coefficients $c_k$, $k = k_L, ..., k_R$, are recovered from those corresponding to $[\bar{L}, \bar{R}]$. Hence, with $P \in \mathbb{N}_+$, we define $k_{\bar{R}} := k_{\bar{L}} + 2^P - 1$ so that $[k_{\bar{L}}/a, k_{\bar{R}}/a] = [\bar{L}, \bar{L} + (2^P - 1)/a)] = [\bar{L}, \bar{R}]$ straddles $[L, R]$.

By utilizing the function defined over $[\bar{L}, \bar{R}]$, coefficients corresponding to the projection over $[L, R]$ are obtained with superior accuracy. To begin, we define

$$\bar{f}(y) = f(\bar{L} + y) - f(\bar{L}) - y\Delta_f, \quad \text{where} \quad \Delta_f := \frac{f(\bar{R}) - f(\bar{L})}{\bar{R} - \bar{L}},$$

and determine the coefficients $\bar{c}_k$ of the projection

$$\bar{f}(S_T) \approx \sum_{k=0}^{N_F - 1} \bar{c}_k \varphi_{a,k}(S_T),$$

where $N_F := 2^P$. We then define

$$h(\xi) := \frac{a^2(1 - \cos(\xi/a))}{\xi^2(2 + \cos(\xi/a))},$$

which we sample over the grid $\xi_j := (j-1)\Delta_\xi$, $j = 1, ..., N_F$, where $\Delta_\xi = \frac{2\pi a}{N_F}$.We determine the coefficients of $\bar{f}$ over $[0, \bar{R} - \bar{L}]$, corresponding to the points $y_k = (k-1)/a$, $k = 1, ..., N_F$

$$\begin{aligned}
\bar{c}_{k-1} &\approx 12 \cdot \Re \int_0^{2\pi a} e^{-iy_k\xi} h(\xi) \hat{\bar{f}}(\xi) d\xi \\
&\approx 12 \cdot \Re \sum_{j=1}^{N_F} e^{-i\frac{(k-1)}{a}\xi_j} h(\xi_j) \hat{\bar{f}}(\xi_j) \Delta_\xi \\
&= \frac{24\pi}{N_F} \cdot \Re \sum_{j=1}^{N_F} e^{-i(k-1)(j-1)\frac{2\pi}{N_F}} h(\xi_j) \hat{\bar{f}}(\xi_j) = \frac{24\pi}{N_F} \cdot \Re \left[ \mathcal{D}_{k-1} \{ h(\xi_j) \hat{\bar{f}}(\xi_j) \}_{j=1}^{N_F} \right], \qquad (41)
\end{aligned}$$

where $\mathcal{D}$ denotes the discrete Fourier transform, and $\mathcal{D}^{-1}$ will denote its inverse. To determine $\{\hat{\bar{f}}(\xi_j)\}_{j=1}^{N_F}$, using the fact that $y_{N_F} = (2^P - 1)/a = \bar{R} - \bar{L}$ we compute

$$\begin{aligned}
\hat{\bar{f}}(\xi_j) &\approx \int_0^{\bar{R} - \bar{L}} e^{i\xi_j y} \bar{f}(y) dy \approx \frac{1}{a} \sum_{n=1}^{N_F} e^{a(j-1)\frac{2\pi}{N_F} y_n} \bar{f}(y_n) \\
&= \frac{1}{a} \sum_{n=1}^{N_F} e^{(j-1)(n-1)\frac{2\pi}{N_F}} \bar{f}(y_n) = \frac{N_F}{a} \mathcal{D}_j^{-1} \{ \bar{f}(y_n) \}_{n=1}^{N_F}. \qquad (42)
\end{aligned}$$

Combining equations (41) and (42), we arrive at

$$\bar{c}_{k-1} \approx \frac{24\pi}{a} \cdot \Re \left[ \mathcal{D}_{k-1} \left\{ h(\xi_j) \cdot D_j^{-1} \{ \bar{f}(y_n) \} \right\} \right]. \qquad (43)$$

From the set $\{\bar{c}_k\}_{k=0}^{2^P - 1}$, we extract the desired coefficients

$$c_{k_L + j} = \bar{c}_{k_0 + j}, \quad j = 0, ..., k_R - k_L,$$

where $k_0 := k_L - k_{\bar{L}}$. This procedure, which is applicable to general Reisz sequences of translates, is easy to implement and produces highly accurate approximations for smooth functions, but is not as robust to payoff discontinuities as the ABS and Dual methods.

[27] Alternatively, if the Fourier transform of $f$ is known, it can be used directly in place of $\hat{\bar{f}}(\xi)$ below.

## Appendix B. Supplemental Material: Additional Applications

B.1. **Hedging with Multiple Decision Periods.** The examples considered so far require portfolio rebalancings at up to two distinct times, once at initialization to establish the static hedge, and possibly an additional rebalancing at a single stopping time prior to expiry. As demonstrated in [7], the semi-static approach can be extended to multiple rebalancing (decision) periods. For example, consider a rolldown call option which is described by a set of barriers, $B_1 > B_2 > \ldots > B_n$, each less than $S_0$. If $S_t > B_1$ for $0 \le t \le T$, the option pays $(S_T - K_0)^+$ at maturity. However, if $B_1$ is breached priory to expiry, the prevailing strike rolls down from $K_0$ to $K_1 < K_0$, and likewise with each successive $B_i$ which is hit, the strike rolls down from $K_{i-1}$ to $K_i$. This contract, denoted $RDC$, can be represented as the following portfolio of down-and-out call options

$$RDC = DOC(K_0, B_1) + \sum_{i=1}^{n-1} \{DOC(K_i, B_{i+1}) - DOC(K_i, B_i)\}$$

where $DOC(K_j, B_m)$ represents a strike $K_j$ option with knock out barrier $B_m$. Hence, at initialization we establish a hedge portfolio of European payoffs

$$\tilde{f}(S_T; K_0, B_1) + \sum_{i=1}^{n-1} \left\{ \tilde{f}(S_T; K_i, B_{i+1}) - \tilde{f}(S_T; K_i, B_i) \right\}, \tag{44}$$

where the individual adjusted payoffs are defined by

$$\tilde{f}(S_T; K_j, B_m) = \begin{cases} (S_T - K_j)^+ & \text{if } S_T > B_m \\ -\left(\frac{S_T}{B_m}\right)^p \left(\frac{B_m^2}{S_T} - K_j\right)^+ & \text{if } S_T \le B_m \end{cases}, \tag{45}$$

and the butterfly methodology can be applied to each individually to obtain an aggregated position in butterfly (or vanilla) payoffs. For each $B_i$ that is breached, the positions in $\tilde{f}(S_T; K_j, B_{j+1})$ and $\tilde{f}(S_T; K_j, B_j)$ are liquidated at the market price, while the remaining positions are left untouched. Hence, this strategy faces at most $n$ rebalancings, including initialization. A similar strategy holds for ratchet options [7].

B.2. **Quasi-Analytical Hedges.** Many practical nonlinear payoffs can be expressed in the form $\Psi(S_T) = h \circ f(S_T)$, or $f \circ h(S_T)$ where $f(S_T)$ is a given payoff. For example, a simple generalization of the powered call is the payoff $(\max\{\lambda S_T - K, 0\})^2$. If analytical hedges of $f$ are known, then in some cases an analytical or quasi-analytical hedge of $\Psi$ is given by specifying the original payoff $f$ along with the type of transformation $h$ any transformation parameters, adding to the efficiency of pricing/hedging routines. The idea is to find formulas for transformed payoffs in terms of known or analytically given coefficients.

B.2.1. *Basic Transforms.* We start with the coefficients $\beta_{a,k}(\Psi)$ of some more obvious transforms when the coefficients $\beta_{a,k}$ are known for $f$:

$$\begin{aligned} &\text{Shift}: \quad \beta_{a,k}(T_{\frac{\bar{k}}{a}} f) = T_{\bar{k}} \beta_{a,k} = \beta_{a,k-\bar{k}} \\ &\text{Scale}: \quad \beta_{a,k}(\lambda f + c) = \lambda \beta_{a,k} + c a^{-1/2} \\ &\text{Dilation}: \quad \beta_{a,k}(f(\gamma \cdot)) = \gamma^{-1/2} \beta_{\frac{a}{\gamma},k} \end{aligned}$$

To avoid a resolution change, we fixed the shift $\bar{k}/a$ to remain at the market spacing, although this is not required in general. Similarly, in the case of dilation, we are finding the coefficients at the market spacing $a > 0$ in terms of those at a potentially unavailable resolution.

B.2.2. *Caps, Floors and Composition.* To avoid catastrophically large payouts or simply to reduce the option premium, caps are often introduced so that $\Psi(S_T) = f(S_T)\mathbb{1}_{f(S_T) \le C} + C\mathbb{1}_{f(S_T) > C} = \min\{f(S_T), C\}$. Floors serve a similar purpose of allowing for upside potential while restricting the maximum attainable loss for the option holder. These payoffs satisfy $\Psi^C(S_T) = \max\{C, f(S_T)\} = C\mathbb{1}_{f(S_T) \le C} + f(S_T)\mathbb{1}_{f(S_T) > C}$, the classic example being the standard vanilla option $\Psi(S_T) = (\alpha(S_T - K))^+ = \max\{0, \alpha(S_T - K)\}$, where $\alpha = \pm 1$ and $C = 0$.

For a capped payoff when $f(S_T)$ is monotone and $x^c$ satisfies $f(x^c) = C$, given the coefficients of $\beta_{a,k}$ of $f$, we set $k^c := \lfloor ax^c \rfloor$, $\lambda^c := ax^c - k^c = (x^c - k^c/a)/(1/a)$, and define

$$\beta^C_{a,k} = \begin{cases} \beta_{a,k} & k < k^c \\ \lambda^c \beta_{a,k^c} + (1-\lambda^c) C a^{-1/2} & k = k^c \\ C a^{-1/2} & k > k^c \end{cases}. \tag{46}$$

Hence, the coefficient at the "pasting" point $k^c$ is weighted according to its proximity to the two separate payoffs, $f(S_T)\mathbb{1}_{f(S_T)\le C}$ and $C\mathbb{1}_{f(S_T)>C}$, and on either side of $k_c$ the coefficients are set according to which payoff is active.

Of course caps and floors can be combined to form collar type payoffs,

$$\Psi^D(S_T) = \max\{F, \min\{f(S_T), C\}\} = F\mathbb{1}_{f(S_T)\le F} + f(S_T)\mathbb{1}_{f(S_T)\in(F,C)} + C\mathbb{1}_{f(S_T)\ge C},$$

where $f$ is typically monotone on $(F, C)$. With $f(x_F) = F$, $k_F := \lceil ax_F \rceil$, and $\lambda_F = ax_F - k_F$, an analogous formula for floored and capped payoffs is found:

$$\beta^D_{a,k} = \begin{cases} F a^{-1/2} & k < k_F \\ (1-\lambda_F)\beta_{a,k_F} + \lambda_F F a^{-1/2} & k = k_F \\ \beta_{a,k} & k_F < k < k^c \\ \lambda^c \beta_{a,k^c} + (1-\lambda^c) C a^{-1/2} & k = k^c \\ C a^{-1/2} & k > k^c \end{cases}. \tag{47}$$

Formulas for compositions of transforms can be specified as well. For example, the capped powered call is the composition of $(\min\{C, \cdot\})^2$ with $\min\{0, S_T - K\}$, where the payoff satisfies $([S_T - K]^+)^2\mathbb{1}_{[S_T\le K + C^{1/2}]} + C\mathbb{1}_{[S_T > K + C^{1/2}]}$ with $K = \bar{k}/a$ ($\bar{k} = k_F$) and $k^c := \lfloor a(K + C^{1/2}) \rfloor$. Hence, the coefficients for $\bar{k} < k < k^c$ are given by equation (19), $\beta^D_{a,k} = 0$ for $k < \bar{k}$ and $\beta^C_{a,k^c} = \lambda a^{-5/2}[(k^c - \bar{k})^2 - \frac{1}{6}] + (1-\lambda)a^{-1/2}C$.

Similarly, at a high enough resolution[28], the payoff $\Psi(S_T) = \max\{f_1(S_T), f_2(S_T)\}$ is accurately represented by the pairwise maximum of the coefficients $\beta^{\max}_{a,k} = \max\{\beta^1_{a,k}, \beta^2_{a,k}\}$. A call on a scaled maximum of the payoffs, $(\gamma\Psi(S_T))^+$ can be found by applying equation (47) with the dilation coefficients $\frac{1}{\gamma}\beta^{\max}_{\frac{a}{\gamma},k}$ in place of $\beta_{a,k}$. When $K = \gamma\bar{k}/a$, the coefficients of $(\gamma\Psi(S_T) - K)^+$ are found by the flooring formula with $F = 0$ and the set of beta coefficients $\frac{1}{\gamma}T_{\frac{\gamma\bar{k}}{a}}\beta^{\max}_{\frac{a}{\gamma},k} = \frac{1}{\gamma}\beta^{\max}_{\frac{a}{\gamma},k-\bar{k}}$. Of course an arbitrary $K$ can be specified, but the coefficients will need to be calculated at a different resolution.

B.2.3. *Piecewise Continuous Payoffs.* As an example, consider the case of a profit hedging commodity supplier whose future expenses vary nonlinearly with an unknown supply quantity, according to a set of tranche-dependent fixed and marginal costs (e.g. according to infrastructure utilization, outsourcing, stock-out, etc.). By capturing the correlation between price and quantity, the joint exposure can be mitigated by a payoff on realized price over $[L, R] = [k_L/a, k_R/a]$. Fixing tranches $I_m = [\frac{\tau_m}{a}, \frac{\tau_{m+1}}{a}]$, $m = 1, ..., M-1$, where $\mathcal{T} = \{\tau_1, ..., \tau_M\}$ are the left tranche boundaries, set to the nearest market strike, $\tau_1 := k_L, \tau_R := k_R$, and $\delta_m := f(\tau_m^+) - f(\tau_m^-)$, $m = 1, ..., M$ are the corresponding payoff jumps, we assume that to each $I_m$ there corresponds a payoff function $f^m := (f - \bar{\delta}_m)\mathbb{1}\{I_m\}$ with known coefficients $\beta^m_{a,k}$, where $\bar{\delta}_m := \sum_{j=1}^m \delta_j = f(\tau_m^+)$. The coefficients of $f = \sum_{m=1}^M (\bar{\delta}_m \mathbb{1}\{I_m\} + f^m)$ are given by

$$\beta^{pc}_{a,k} = \begin{cases} a^{-1/2}\bar{\delta}_m + \beta^m_{a,k} & k \in \mathcal{T}^c \cup \{k_L, k_R\} \\ a^{-1/2}[\bar{\delta}_{m-1} + \delta_m/2] + \dfrac{\beta^{m-1}_{a,k} + \beta^m_{a,k}}{2} & k = \tau_m \end{cases}.$$

Of course, the projection methods can still be applied directly to $f$, without specifying any known discontinuities. The ABS methods in particular are robust to payoff jumps, since their

[28]This qualitative statement is left to be specified by the user, with the condition that at low resolutions, ie large spacing between strikes, a method such as $ABS_2$ is used. For pricing resolutions, this approximation will generally be sufficient.

narrow support mitigates the effect of Gibbs phenomenon. When the jumps are known, and the payoffs corresponding to each tranch are analytically hedged, this approach is preferred.

B.3. **Global Hedges by Linear Extrapolation.** For the most part, we have used butterfly representations to obtain local hedges, which introduces tail risk for payoffs with unbounded support, such as uncapped power options. This section proposes a potential method to reduce tail risk for payoffs with unbounded support. Suppose that $f$ is differentiable at $L$ and $R$. To establish a global hedge for a general payoff $f$ from its local hedge on $[L, R]$, constructed from butterfly payoffs $\{\varphi_{a,k}\}_{k=0}^{\bar{K}}$, we can choose boundary instruments to match the payoff's slope at $L$ and $R$. Hence, we define the right boundary payoffs

$$\varphi^1_{a,\bar{K}}(S_T) = a^{3/2}\Big[S_T - \Big(L + \frac{\bar{K}-1}{a}\Big)\Big]^+ = a^{3/2}\cdot\psi^{call}_{R-\frac{1}{a}}(S_T)$$

$$\varphi^2_{a,\bar{K}}(S_T) = a^{3/2}\Big[S_T - \Big(L + \frac{\bar{K}}{a}\Big)\Big]^+ = a^{3/2}\cdot\psi^{call}_{R}(S_T),$$

which are used to approximate the right tail behavior of an unbounded payoff. Similarly define the left boundary payoffs

$$\varphi^1_{a,0}(S_T) = a^{3/2}\Big[\Big(L + \frac{1}{a}\Big) - S_T\Big]^+ = a^{3/2}\cdot\psi^{put}_{L+\frac{1}{a}}(S_T)$$

$$\varphi^2_{a,0}(S_T) = a^{3/2}\big[L - S_T\big]^+ = a^{3/2}\cdot\psi^{put}_{L}(S_T)$$

Starting with a local approximation on $[L, R]$, where the coefficients $\beta^1_{a,\bar{K}}$ of $\varphi^1_{a,\bar{K}}$ and $\beta^1_{a,0}$ of $\varphi^1_{a,0}$ are calculated in the usual manner, we augment the representation by linear extension at each boundary. To match boundary slopes, the coefficients $\beta^2_{a,\bar{K}}$ of $\varphi^2_{a,\bar{K}}$ and $\beta^2_{a,0}$ of $\varphi^2_{a,0}$ must satisfy

$$a^{1/2}(\beta^2_{a,\bar{K}} + \beta^1_{a,\bar{K}}) = f'(R)$$

and

$$a^{1/2}(\beta^2_{a,0} + \beta^1_{a,0}) = f'(L),$$

which yields

$$\beta^2_{a,\bar{K}} = a^{-1/2}f'(R) - \beta^1_{a,\bar{K}} \quad \text{and} \quad \beta^2_{a,0} = -a^{-1/2}f'(L) - \beta^1_{a,0}.$$

We then have the global static hedge

$$\begin{aligned} f(S_T) \approx\ & a^{3/2}\beta_{a,0}\cdot\psi^{put}_{L+\frac{1}{a}}(S_T) - \Big(af'(L) + a^{3/2}\beta_{a,0}\Big)\psi^{put}_{L}(S_T) \\ & + \sum_{k=1}^{\bar{K}-1}\beta_{a,k}\varphi_{a,k}(S_T) + a^{3/2}\beta_{a,\bar{K}}\cdot\psi^{call}_{R-\frac{1}{a}}(S_T) \\ & + \Big(af'(R) - a^{3/2}\beta_{a,\bar{K}}\Big)\cdot\psi^{call}_{R}(S_T), \end{aligned}$$

which gives an extended butterfly basis. This approach can be used to reduce, but not eliminate tail risk for unbounded payoffs, and $[L, R]$ can be selected in a model-dependent manner. For example, $[L, R]$ can be chosen so that the probability (or model-based value ascribed to the payoff) corresponding to $S_T \in [0, L) \cup (R, \infty)$ is below a fixed tolerance. At the same time, portfolio rebalancing can be performed in response to extreme moves in the underlying to further limit exposure.

B.4. **Pricing on the Level of Randomness.** While exponential (e.g. exponential Lévy) models are very popular in equity and other markets with strictly positive asset prices or indices, in some cases pricing on the level of randomness is more appropriate, and we can utilize duality for a Riesz basis to price a payoff when the risk-neutral characteristic function for the underlying

| Scale($\Delta$) | 2 | 1 | 0.5 | 0.25 | 0.1 | 0.01 |
|---|---|---|---|---|---|---|
| Interp | 2.000e-02 | 5.000e-03 | 1.250e-03 | 3.125e-04 | 5.000e-05 | 5.000e-07 |
| $\text{Dual}_{12}$ | 7.698e-03 | 1.924e-03 | 4.811e-04 | 1.203e-04 | 1.922e-05 | 2.254e-07 |
| $\text{ABS}_2$ | 7.698e-03 | 1.925e-03 | 4.811e-04 | 1.203e-04 | 1.925e-05 | 1.925e-07 |
| CM | 1.000e-02 | 7.500e-02 | 3.750e-02 | 1.875e-02 | 7.500e-03 | 7.500e-04 |

TABLE 7. RAHE of $f(S_T) = S_T^2$ on $[L, R] = [0, 10]$.

is known, as opposed to that of the log underlying. The pricing functional is then given by $\mathcal{V}(f) = e^{-rT}\int_{\mathbb{R}} f(s)p_T(s)ds$, $\forall f \in \mathcal{H}$. By duality, $\sum_k \langle f, \widetilde{\phi}_{j,k}\rangle \phi_{j,k} = \sum_k \langle f, \phi_{j,k}\rangle \widetilde{\phi}_{j,k}$. Hence

$$\begin{aligned}\mathcal{V}_j(f) = \mathcal{V}\left(\sum_k \langle f, \phi_{j,k}\rangle \widetilde{\phi}_{j,k}\right) &= e^{-rT}\sum_k \langle f, \phi_{j,k}\rangle \langle p_T, \widetilde{\phi}_{j,k}\rangle \\ &= \frac{e^{-rT}}{2\pi}\sum_k \langle f, \phi_{j,k}\rangle \langle \widehat{p}_T, \widehat{\widetilde{\phi}}_{j,k}\rangle,\end{aligned}$$

whereby the coefficients $\langle \widehat{p}_T, \widehat{\widetilde{\phi}}_{j,k}\rangle$ represent the projection coefficients of $p_T$, and the payoff is now integrated directly against $\phi_{j,k}$.

B.5. **Higher Dimensional Extensions.** In order to price multi-dimensional payoffs, for instance the basket call option with terminal payout $f(S_{1,T}, ..., S_{d,T}) = \left(\sum_{n=1}^d \gamma_n S_{n,T} - K\right)^+$, basis theory suggests the use of the tensor product basis $\prod_{n=1}^d \varphi_{a_n,k_n}(s_n)$, where each dimension is alloted a separate resolution. Fixing $a_1 = a_2 = a$ in the two dimensional case, and $\varphi^a_{k_1,k_2}(S_{1,T}, S_{2,T}) := \varphi_{a,k_1}(S_{1,T})\varphi_{a,k_2}(S_{2,T})$, the projection is formed by

$$\begin{aligned}P_{\mathcal{M}_a^2} f(S_{1,T}, S_{2,T}) &= \sum_{k_1} \varphi_{a,k_1}(S_{1,T}) \sum_{k_2} \varphi_{a,k_2}(S_{2,T}) \langle f(s_1, s_2), \widetilde{\varphi}^a_{k_1,k_2}(s_1, s_2)\rangle \\ &= \sum_{k_1,k_2} \beta^a_{k_1,k_2} \cdot \varphi^a_{k_1,k_2}(S_{1,T}, S_{2,T}),\end{aligned}$$

where $\widetilde{\varphi}^a_{k_1,k_2}(s_1, s_2) = \widetilde{\varphi}_{a,k_1}(s_1)\widetilde{\varphi}_{a,k_2}(s_2)$. In higher dimensions, the ABS schemes become even more essential to reduce the computational cost of $\beta^a_{k_1,k_2}$, where we simply replace the product dual by the product of one dimensional ABS elements. Given a joint pricing kernel (or joint characteristic function), prices for the basis elements are used to price $f$:

$$\mathcal{V} \circ f(S_{1,0}, S_{2,0}) \approx \sum_{k_1,k_2} \beta^a_{k_1,k_2} \cdot \mathcal{V} \circ \varphi^a_{k_1,k_2}(S_{1,0}, S_{2,0}).$$

Moreover, once payoff coefficients have been obtained for the payoff form, products on various asset pairs are priced by computing the corresponding basis prices, using the same set of $\beta^a_{k_1,k_2}$. Alternatively, higher dimensions offer the possibility of designing frames to efficiently price specific payoff forms. For example, a difference frame space could be tailored to the pricing of rainbow options such as $(\max\{S_{1,T}, S_{2,T}\} - K)^+$.

## APPENDIX C. SUPPLEMENTAL MATERIAL: ADDITIONAL EXPERIMENTS

This section considers several additional payoffs to further illustrate the performance of the proposed method. While the $\text{ABS}_2$ method is preferred in general, we present results for the $\text{ABS}_1$ method as well for comparison. We first demonstrate the effectiveness of the dual and $\text{ABS}_2$ methods with two practical examples. Table 7 presents hedge errors for the payoff $f(S_T) = S_T^2$ on $[0, 10]$. For the dual method, we take $\gamma = 12$, which corresponds to a dual approximation on the interval $[-13, 13]$. As expected from Proposition 5.4, the dual and ABS errors are nearly identical since the payoff is a second order polynomial. Moreover, the dual and ABS methods are more than twice as accurate as interpolation at each resolution, and many times more accurate than CM.

Table 8 illustrates this example with the payoff $f(S_T) = \ln(S_T)$ over $[1, 11]$. Note that we have avoided the singularity in our computation, since none of the methods are able to reasonably

| Scale($\Delta$) | 2 | 1 | 0.5 | 0.25 | 0.1 | 0.01 |
|---|---|---|---|---|---|---|
| Interp | 1.689e-02 | 4.487e-03 | 1.147e-03 | 2.885e-04 | 4.624e-05 | 4.626e-07 |
| $\text{Dual}_{12}$ | 8.778e-03 | 2.064e-03 | 4.853e-04 | 1.165e-04 | 1.816e-05 | 2.707e-07 |
| $\text{ABS}_2$ | 8.918e-03 | 2.088e-03 | 4.908e-04 | 1.174e-04 | 1.820e-05 | 1.784e-07 |
| CM | 2.866e-02 | 2.625e-02 | 1.174e-02 | 5.572e-03 | 2.163e-03 | 2.124e-04 |

TABLE 8. RAHE of $f(S_T) = \ln(S_T)$ on $[L, R] = [1, 11]$.

| | Test 1: BSM | | | Test 2: MJD | | |
|---|---|---|---|---|---|---|
| $\Delta$ | Interp | ABS1 | ABS2 | Interp | ABS1 | ABS2 |
| 2.00 | 2.385e-03 | 1.473e-03 | 6.359e-04 | 1.150e-03 | 6.036e-04 | 7.736e-05 |
| 1.00 | 6.414e-04 | 3.436e-04 | 4.518e-05 | 3.079e-04 | 1.592e-04 | 1.054e-05 |
| 0.50 | 1.631e-04 | 8.306e-05 | 2.839e-06 | 7.757e-05 | 3.928e-05 | 9.686e-07 |
| 0.25 | 4.095e-05 | 2.057e-05 | 1.775e-07 | 1.947e-05 | 9.822e-06 | 1.745e-07 |
| 0.10 | 6.560e-06 | 3.282e-06 | 4.595e-09 | 3.161e-06 | 1.614e-06 | 6.761e-08 |
| | Test 3: KoBoL (CGMY) | | | RAHE | | |
| $\Delta$ | Interp | ABS1 | ABS2 | Interp | ABS1 | ABS2 |
| 2.00 | 2.595e-04 | 3.866e-04 | 5.066e-04 | 1.092e-02 | 8.416e-03 | 7.854e-03 |
| 1.00 | 1.046e-04 | 6.853e-05 | 3.072e-05 | 2.850e-03 | 1.838e-03 | 1.375e-03 |
| 0.50 | 2.823e-05 | 1.434e-05 | 2.075e-07 | 7.264e-04 | 4.383e-04 | 2.991e-04 |
| 0.25 | 6.915e-06 | 3.122e-06 | 6.920e-07 | 1.827e-04 | 1.082e-04 | 7.161e-05 |
| 0.10 | 9.192e-07 | 2.976e-07 | 3.249e-07 | 2.924e-05 | 1.722e-05 | 1.129e-05 |

TABLE 9. $f(S_T) = [\sin(S_T)/2 + 3\ln(5 + S_T)]\mathbb{1}_{[0,40]}(S_T)$: Comparison of value approximation error and RAHE, with strikes in $[0, 40]$ at spacing $\Delta$, $S_0 = 8$, $T = 1$, reference values: $[7.39914, 7.33607, 7.19402, 9.264774]$.

hedge a payoff near a singularity without carefully adjusting the resolution. Moreover, to keep the singular behavior near $S_T = 0$ from affecting the coefficients, it is advantageous to use an auxiliary linear approximation of the form

$$\ln(S_T)\mathbb{1}[S_T \geq \tau] + [\ln(\tau) + (S_T - \tau)/\tau]\mathbb{1}[S_T < \tau],$$

which replaces the left singular tail by a linear Taylor approximation. While any value of $\tau \in (s, L]$ is reasonable, where $s$ denotes the singularity ($s = 0$ for $\ln(S_T)$), as $\tau$ approaches $s$ the approximation deteriorates (especially for the dual method, since the ABS has narrower support). We find that a value of $\tau = 1/2$ works well in this case[29].

To demonstrate the potential for pricing highly nonlinear contracts, we construct the payoff

$$f(S_T) = [\sin(S_T)/2 + 3\ln(5 + S_T)]\mathbb{1}_{[0,40]}(S_T), \quad S_0 = 8, T = 1.$$

Table 9 demonstrates again the superiority of ABS methods in terms of both pricing convergence and relative hedge errors. At a resolution $\Delta = .5$, the $\text{ABS}_2$ method is about two orders more accurate than interpolation, for each model tested, and its relative hedge error is less than half.

We also consider the amortizing option of [25]:

$$f(S_T) = \frac{(S_T - K)^+}{S_T}, \quad S_0 = 14, K = 15, T = .5,$$

the results of which are shown in Table 10. In the BSM case, $\text{ABS}_2$ significantly outperforms interpolation, where at the resolution $\Delta = 0.1$, an interpolation accuracy of $e - 05$ compares to an $\text{ABS}_2$ accuracy of $e-10$, demonstrating that while $\text{ABS}_2$ generally outperforms interpolation, the discrepancy can be extreme. For the CGMY model, with $\Delta = 2$, the accuracy of $\text{ABS}_2$ has already reached $e - 06$, and accuracy which is not reached by interpolation even for $\Delta = 0.1$.

[29]For the $\text{ABS}_2$, an approximation can be made arbitrarily close to $s$ by choosing $\Delta$ so that $s \leq L - 2\Delta$, and using the auxiliary payoff with $\tau(\Delta) := L - \Delta$. This can be applied to functions other than $\ln(S_T)$, and also at interior and right boundary singularities by similarly defining the auxiliary payoff.

| | Test 1: BSM | | | Test 2: MJD | | |
|---|---|---|---|---|---|---|
| $\Delta$ | Interp | $ABS_1$ | ABS2 | Interp | ABS1 | ABS2 |
| 2.00 | 6.433e-02 | 4.659e-02 | 2.730e-03 | 4.525e-02 | 3.239e-02 | 8.193e-04 |
| 1.00 | 3.810e-03 | 1.888e-03 | 1.136e-05 | 2.816e-03 | 1.396e-03 | 3.742e-06 |
| 0.50 | 9.558e-04 | 4.768e-04 | 6.488e-07 | 7.068e-04 | 3.537e-04 | 1.903e-06 |
| 0.25 | 2.391e-04 | 1.195e-04 | 3.774e-08 | 1.772e-04 | 8.904e-05 | 9.711e-07 |
| 0.10 | 3.827e-05 | 1.914e-05 | 8.865e-10 | 2.875e-05 | 1.465e-05 | 5.523e-07 |
| | Test 3: KoBoL (CGMY) | | | RAHE | | |
| $\Delta$ | Interp | ABS1 | ABS2 | Interp | ABS1 | ABS2 |
| 2.00 | 2.378e-02 | 1.711e-02 | 7.499e-06 | 2.838e-03 | 2.483e-03 | 2.481e-03 |
| 1.00 | 1.928e-03 | 9.754e-04 | 3.198e-05 | 2.983e-04 | 1.843e-04 | 1.286e-04 |
| 0.50 | 4.864e-04 | 2.490e-04 | 1.218e-05 | 7.460e-05 | 4.504e-05 | 3.048e-05 |
| 0.25 | 1.249e-04 | 6.556e-05 | 6.285e-06 | 1.865e-05 | 1.112e-05 | 7.403e-06 |
| 0.10 | 2.354e-05 | 1.406e-05 | 4.571e-06 | 2.984e-06 | 1.766e-06 | 1.163e-06 |

TABLE 10. Amortizing option $f(S_T) = \frac{(S_T-15)^+}{S_T}\mathbb{1}_{[0,50]}(S_T)$. Comparison of value approximation error and RAHE, with strikes in $[10, 50]$ at spacing $\Delta$, $S_0 = 14$, $T = .5$, reference values: $[0.04827, 0.05624, 0.07330, 0.42351]$.