Di Tella, Paolo; Haubold, Martin; Keller-Ressel, Martin

Semistatic and sparse variance-optimal hedging

Mathematical Finance

Provided in Cooperation with:
John Wiley & Sons

Suggested Citation: Di Tella, Paolo; Haubold, Martin; Keller-Ressel, Martin (2019) : Semistatic and sparse variance-optimal hedging, Mathematical Finance, ISSN 1467-9965, Wiley, Hoboken, NJ, Vol. 30, Iss. 2, pp. 403-425, https://doi.org/10.1111/mafi.12235

Received: 1 September 2017 | Accepted: 1 August 2019

DOI: 10.1111/mafi.12235

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics WILEY

ORIGINAL ARTICLE

Semistatic and sparse variance-optimal hedging

Paolo Di Tella | **Martin Haubold** | **Martin Keller-Ressel**

Institute for Mathematical Stochastics, TU Dresden, Germany

Correspondence
Martin Keller-Ressel, Institute for Mathematical Stochastics, TU Dresden, 01069 Dresden, Germany.
Email: Martin.Keller-Ressel@tu-dresden.de

Funding information
German Research Foundation (DFG) under grant ZUK 64 and KE 1736/1-1.

Abstract

We consider the problem of hedging a contingent claim with a "semistatic" strategy composed of a dynamic position in one asset and static (buy-and-hold) positions in other assets. We give general representations of the optimal strategy and the hedging error under the criterion of variance optimality and provide tractable formulas using Fourier integration in case of the Heston model. We also consider the problem of optimally selecting a sparse semistatic hedging strategy, i.e., a strategy that only uses a small subset of available hedging assets and discuss parallels to the variable-selection problem in linear regression. The methods developed are illustrated in an extended numerical example where we compute a sparse semistatic hedge for a variance swap using European options as static hedging assets.

1 | INTRODUCTION

Semistatic hedging strategies are strategies that are composed of a dynamic (i.e., continuously rebalanced) position in one asset and of static (i.e., buy-and-hold) positions in other assets. Such hedging strategies have appeared in mathematical finance in several different contexts: the hedging of Barrier options (cf. Carr, 2011), model-free hedging approaches based on martingale optimal transport (cf. Beiglböck, Henry-Labordère, & Penkner, 2013), and—most relevant in our context—the semistatic replication of variance swaps by Neuberger's formula (cf. Neuberger, 1994). Compared with fully dynamic strategies, semistatic strategies have the advantage that no rebalancing costs or liquidity risks are associated with the static part of the strategy and hence even assets with limited liquidity can be used as static hedging assets.

Remarkably, for certain hedging problems, semistatic strategies allow for perfect replication even in incomplete markets—at least theoretically. Again, the most prominent example is the replication formula for a variance swap, given by Neuberger (1994) and Carr and Madan (2001) and discussed in more detail in Section 2.4: In any continuous martingale model, a variance swap can be replicated by dynamic hedging in the underlying and a static portfolio of European put- and call-options. This very replication formula is at the heart of the computation of the volatility index `VIX`, whose value is determined from a discretization of Neuberger's replicating option portfolio (cf. Chicago Board Options Exchange (2019)).

However, Neuberger's result relies on certain idealizations: Most importantly, the static part of the strategy consists of *infinitesimally small* positions in an *infinite number* of puts and calls with strikes ranging from zero to infinity. Any practical implementation of this strategy therefore has to decide on a certain quantization of the theoretical strategy, i.e., how to assign noninfinitesimal weights to the actually tradable put- and call-options. Rather than doing this in an ad hoc manner, our goal is to determine how to *optimally* implement a semistatic hedging strategy when a finite number n of hedging assets is available. Our optimality criterion is the well-known variance-optimality criterion introduced by Schweizer (1984) and Föllmer and Sondermann (1986), i.e., we minimize the variance of the residual hedging error under the risk-neutral measure. As we show in Section 2, this criterion is perfectly compatible with semistatic hedging: The semistatic hedging problem separates into an inner problem, which is equivalent to the variance-optimal hedging problem with a single asset (as considered in Föllmer & Sondermann, 1986; Schweizer, 1984) and an outer problem, which is an n-dimensional quadratic optimization problem, cf. Theorem 2.3.

After having analyzed the general structure of the variance-optimal semistatic hedging problem, we turn to another question in Section 3: How many assets $d < n$ are enough to obtain a "reasonably small" hedging error? In case of Neuberger's formula for the variance swap—where infinitely many European options reduce the hedging error to zero—how good is using 12, 6, or even just 3 options? Beyond that, which 3 options should one select from, say, 30 that are available in the market? It turns out that this problem of finding a *sparse semistatic hedging* strategy is closely related to the well-known problem of *variable selection* in high-dimensional regression (cf. Hastie, Tibshirani, & Friedman, 2013, section 3.3), and more generally to sparse modeling approaches in statistics and machine learning.[1] Indeed, to solve the problem of optimal selection we will draw from methods developed in statistics, such as the LASSO, greedy forward selection and the method of Leaps-and-Bounds.

Finally, with the goal of a numerical implementation of sparse semistatic hedging in mind, we have to find tractable methods to compute variances and covariances of hedging errors, expressed mathematically as residuals in the Galtchouk-Kunita-Watanabe (GKW) martingale decomposition. Here, we build on results from Kallsen and Pauwels (2010) which enable us to calculate the GKW decomposition "semi-analytically," i.e., in terms of Fourier integrals, in several models of interest, such as the Heston model. The results of Kallsen and Pauwels (2010), which focus on calculation of the strategy and the hedging error in a classic variance-optimal hedging framwork, are however not sufficient for the *semistatic* hedging problem and we draw from some extensions that are developed in the technical companion paper, Di Tella, Haubold, and Keller-Ressel (2019).

We conclude in Section 5 with a detailed numerical example, implementing the sparse semistatic hedging problem for hedging a variance swap with put and call options in the Heston model. In particular, we compare the performance of the different solution methods for the subset selection problem, analyze the dependency of optimal hedging portfolio and hedging error on the number d of static hedging assets, and study the influence of the leverage parameter ρ on the optimal solution.

2 | VARIANCE-OPTIMAL SEMISTATIC HEDGING

To set up our model for the financial market, we fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F}$ satisfying the usual conditions. We fix a time horizon $T > 0$, assume that $\mathcal{F}_0$ is the trivial σ-algebra, and set $\mathcal{F} = \mathcal{F}_T$. We denote by $S = (S_t)_{t\geq 0}$ the price process of a traded asset and set interest rates to zero to simplify the exposition of results. Consistent with market practice, we assume that a unique risk-neutral pricing measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, can be determined by calibration to vanilla option prices, regardless of market completeness. Expectations $\mathbb{E}[.]$ denote expectations under this risk-neutral measure $\mathbb{Q}$, unless otherwise indicated. Finally, we make the technical assumption that S is a continuous square-integrable martingale under $\mathbb{Q}$. More generally, we denote by $\mathcal{H}^2 = \mathcal{H}^2(\mathbb{F})$ the set of real-valued $\mathbb{F}$-adapted square integrable $\mathbb{Q}$-martingales, which becomes a Hilbert space when equipped with the norm $\|X\|^2_{\mathcal{H}^2} := \mathbb{E}[X_T^2]$. We also set $\mathcal{H}^2_0 := \{X \in \mathcal{H}^2 : X_0 = 0\}$.

2.1 | Variance-optimal hedging

Before discussing semistatic hedging, we quickly review variance-optimal hedging of a claim H^0 in $L^2(\Omega, \mathcal{F}, \mathbb{Q})$, up to the time horizon $T > 0$, as discussed for instance in Föllmer and Sondermann (1986). We identify the claim H^0 with the martingale

$$H^0_t = \mathbb{E}[H^0|\mathcal{F}_t], \qquad t \in [0, T],$$

which is an element of $\mathcal{H}^2$. The set of all admissible dynamic strategies is denoted by

$$\mathrm{L}^2(S) := \left\{ \vartheta \text{ predictable and } \mathbb{R}\text{-valued: } \mathbb{E}\left[\int_0^T \vartheta_t^2 \mathrm{d}\langle S, S\rangle_t\right] < +\infty \right\},$$

where $\langle S, S\rangle$ denotes, as usual, the predictable quadratic variation of S. The variance-optimal hedge ϑ with initial capital c of the claim H^0 is the solution of

$$\epsilon^2 = \min_{\vartheta\in \mathrm{L}^2(S), c\in\mathbb{R}} \mathbb{E}\left[\left(c + \int_0^T \vartheta_t \mathrm{d}S_t - H^0_T\right)^2\right]. \tag{1}$$

The resulting quantity ϵ is the minimal hedging error. The minimization problem (1) can be interpreted as orthogonal projection (in $\mathcal{H}^2$) of the claim H^0 onto the closed subspace spanned by deterministic constants (corresponding to the initial capital c) and by $\mathcal{L}^2(S) := \{\int_0^T \vartheta_t \mathrm{d}S_t,\ \vartheta \in \mathrm{L}^2(S)\} \subset \mathcal{H}^2_0$, the set of claims attainable with strategies from $\mathrm{L}^2(S)$. The resulting orthogonal decomposition

$$H^0_t = c + \int_0^t \vartheta_s \mathrm{d}S_s + L_t\,, \tag{2}$$

of H^0 is known as the GKW decomposition of H^0 with respect to S, cf. Kunita and Watanabe (1967) and Ansel and Stricker (1993). From the financial mathematics perspective, (2) decomposes the claim H^0 into initial capital, hedgable risk, and unhedgable residual risk.

The orthogonality of L to $\mathcal{L}^2(S)$ in the Hilbert space sense implies orthogonality of L to S in the martingale sense, i.e., it holds that $\langle L, S\rangle = 0$. Hence, the variance-optimal strategy ϑ can be computed from (2) as

$$\langle H, S\rangle_t = \int_0^t \vartheta_s\, \mathrm{d}\langle S, S\rangle_s\,, \tag{3}$$

and ϑ can be expressed as the Radon-Nikodym derivative $\vartheta = \mathrm{d}\langle H, S\rangle/\mathrm{d}\langle S, S\rangle$.

2.2 | The variance-optimal semistatic hedging problem

We are now prepared to discuss the variance-optimal *semistatic* hedging problem and its solution. In addition to the contingent claim H^0, which is to be hedged, denote by $H = (H^1, \dots, H^n)^\top$ the vector of supplementary contingent claims, all assumed to be square-integrable random variables in $L^2(\Omega, \mathcal{F}, \mathbb{Q})$. Again, we associate to each H^i the martingale

$$H_t^i := \mathbb{E}\big[H^i \big| \mathcal{F}_t\big], \quad t \in [0,T], \quad i = 0, \dots, n. \tag{4}$$

The static part of the strategy can be represented by an element v of $\mathbb{R}^n$, where v_i represents the quantity of claim H^i bought at time $t = 0$ and held until time $t = T$. The dynamic part ϑ of the strategy is again represented by an element of $\mathrm{L}^2(S)$.

Definition 2.1 (Variance-Optimal Semistatic Hedging Problem)**.** The variance-optimal semistatic hedge $(\vartheta, v) \in \mathrm{L}^2(S) \times \mathbb{R}^n$ and the optimal initial capital $c \in \mathbb{R}$ are the solution of the minimization problem

$$\epsilon^2 = \min_{(\vartheta, v) \in \mathrm{L}^2(S) \times \mathbb{R}^n, c \in \mathbb{R}} \mathbb{E}\left[\left(c - v^\top \mathbb{E}\big[H_T\big] + \int_0^T \vartheta_t \mathrm{d}S_t - \big(H_T^0 - v^\top H_T\big)\right)^2\right]. \tag{5}$$

Note that $v^\top \mathbb{E}[H_T]$ is the cost of setting up the static part of the hedge and its terminal value is $v^\top H_T$. The dynamic part is self-financing and results in the terminal value $\int_0^T \vartheta_t \mathrm{d}S_t$. Adding the initial capital c and subtracting the target claim H_T^0 yields the above expression for the hedging problem.

To solve the variance-optimal semistatic hedging problem, we decompose it into an inner and an outer minimization problem and rewrite (5) as

$$\begin{cases} \epsilon^2(v) = \min_{\vartheta \in \mathrm{L}^2(S), c \in \mathbb{R}} \mathbb{E}\left[\left(c - v^\top \mathbb{E}\big[H_T\big] + \int_0^T \vartheta_t \mathrm{d}S_t - \big(H_T^0 - v^\top H_T\big)\right)^2\right], & \text{(inner prob.)} \\ \epsilon^2 = \min_{v \in \mathbb{R}^n} \epsilon(v)^2. & \text{(outer prob.)} \end{cases} \tag{6}$$

The inner problem is of the same form as (1), while the outer problem turns out to be a finite dimensional quadratic optimization problem. To formulate the solution, we write the GWK decompositions of the claims $(H^0, \dots, H^n)$ with respect to S as

$$H_t^i = H_0^i + \int_0^t \vartheta_s^i \mathrm{d}S_s + L_t^i, \quad i = 0, \dots, n. \tag{7}$$

Similarly to (3) we get

$$\vartheta^i = \frac{\mathrm{d}\langle H^i, S\rangle}{\mathrm{d}\langle S, S\rangle}, \quad i = 0, \dots, n, \tag{8}$$

and introduce the vector notation $\vartheta := (\vartheta^1, \dots, \vartheta^n)^\top$ for the strategies and $L := (L^1, \dots, L^n)^\top$ for the residuals in the GKW decomposition. Finally we formulate the following condition.

Definition 2.2 (Nonredundancy condition)**.** The supplementary claims $H = (H^1, \dots, H^n)^\top$ satisfy the *nonredundancy condition* if there is no $x \in \mathbb{R}^n \setminus \{0\}$ with $x^\top L_T = 0$, a.s.

Intuitively, existence of a nonzero x with $x^\top L_T = 0$ means that the number of supplementary assets can be reduced without changing the hedging error ϵ^2 in (5). We are now prepared to state our main result on the solution of the variance-optimal semistatic hedging problem.

Theorem 2.3. *Consider the variance-optimal semistatic hedging problem (5) and set*

$$A := \mathrm{Var}\left[L_T^0\right], \quad B := \mathrm{Cov}\left[L_T, L_T^0\right], \quad C := \mathrm{Cov}[L_T, L_T]. \tag{9}$$

Under the nonredundancy condition, C is invertible and the unique solution of the semistatic hedging problem is given by

$$c = \mathbb{E}\left[H_T^0\right], \quad \upsilon = C^{-1}B, \quad \vartheta^\upsilon = \vartheta^0 - \upsilon^\top \vartheta.$$

The minimal squared hedging error is given by

$$\varepsilon^2 = A - B^\top C^{-1} B.$$

Moreover, the elements of A, B, and C can be expressed as

$$\mathbb{E}\left[L_T^i L_T^j\right] = \mathbb{E}\left[\langle L^i, L^j\rangle_T\right] = \mathbb{E}\left[\langle H^i, H^j\rangle_T - \int_0^T \vartheta_t^i \vartheta_t^j \, \mathrm{d}\langle S, S\rangle_t\right], \quad i, j = 0, \dots, n. \tag{10}$$

Corollary 2.4. *If the nonredundancy condition does not hold true, then any solution $\upsilon \in \mathbb{R}^n$ of the linear system $C\upsilon = B$, together with $c = \mathbb{E}[H_T^0]$ and $\vartheta^\upsilon = \vartheta^0 - \upsilon^\top \vartheta$ is a solution of the semistatic hedging problem. The solution set is never empty, and the solution that minimizes the Euclidian norm of υ can be obtained by setting $\upsilon = C^\dagger B$, where $C^\dagger$ denotes the Moore–Penrose pseudo-inverse of C.*

Notice that the minimal squared hedging error ε^2 in Theorem 2.3 is the Schur complement of the block C in the 'extended covariance matrix'

$$\mathrm{Cov}\left[\left(L_T^0, L_T\right), \left(L_T^0, L_T\right)\right] = \begin{bmatrix} A & B^\top \\ B & C \end{bmatrix}.$$

In particular, if (L_T^0, L_T) has normal distribution, then ε^2 can be expressed as $\varepsilon^2 = \mathrm{Var}[L_T^0 | L_T]$.

Proof of Theorem 2.3 *and Corollary 2.4.* First, we consider the inner minimization problem in (6). This problem is equivalent to the variance-optimal hedging problem for the claim $H^\upsilon := H^0 - \upsilon^\top H$. The solution ϑ^υ is given by the GKW decomposition

$$\left(H_t^0 - \upsilon^\top H_t\right) = \left(H_0^0 - \upsilon^\top H_0\right) + \int_0^t \vartheta_s^\upsilon \mathrm{d}S_s + L_t^\upsilon, \quad t \in [0, T] \tag{11}$$

of the martingale $(H^0 - \upsilon^\top H)$ with respect to S. By (3) we obtain

$$\vartheta_t^\upsilon = \frac{\mathrm{d}\langle (H^0 - \upsilon^\top H), S\rangle_t}{\mathrm{d}\langle S, S\rangle_t} = \vartheta_t^0 - \upsilon^\top \vartheta_t,$$

using the bilinearity of the predictable quadratic covariation. Uniqueness of the GKW decomposition yields $L_t^\upsilon = L_t^0 - \upsilon^\top L_t$ and the squared hedging error is given by

$$\begin{aligned} \varepsilon(\upsilon)^2 = \mathbb{E}\left[(L^\upsilon)^2\right] = \mathbb{E}\left[\left(L_T^0 - \upsilon^\top L_T\right)^2\right] = \upsilon^\top \mathbb{E}\left[L_T L_T^\top\right] \upsilon - 2\upsilon^\top \mathbb{E}\left[L_T L_T^0\right] + \mathbb{E}\left[\left(L_T^0\right)^2\right] = \\ = \upsilon^\top C \upsilon - 2\upsilon^\top B + A. \end{aligned}$$

Thus, the outer optimization problem in (6) becomes

$$\varepsilon^2 = \min_{v \in \mathbb{R}^n} (v^\top C v - 2 v^\top B + A). \tag{12}$$

Because C is positive semidefinite, the first-order condition $Cv = B$ is necessary and sufficient for optimality of v. Under the nonredundancy condition, $\mathrm{Var}(x^\top L_T) > 0$ for any $x \in \mathbb{R}^n \setminus \{0\}$, hence C is positive definite and in particular invertible. The unique solution of the outer problem is therefore given by $v = C^{-1}B$, completing the proof of Theorem 2.3.

For the corollary, it remains to show that $Cv = B$ has a solution, even when the nonredundancy condition does not hold. A solution exists, if B is in the range of C, or equivalently, if B is in $(\ker C)^\perp$. By assumption $\ker C$ is nonempty, and we can choose come $x \in \ker C$, i.e., with $x^\top C = 0$. Because C is the covariance matrix of L_T is follows that $x^\top L_T = 0$, a.s. This implies that also $x^\top B = \mathrm{Cov}(x^\top L_T, L_T^0) = 0$, for all $x \in \ker C$ and hence that $B \in (\ker C)^\perp$. □

Finally, we compute the *hedge contribution* of a single supplementary asset H^{n+1}. By hedge contribution, we mean the reduction in squared hedging error that is achieved by adding the asset H^{n+1} to a given pool of supplementary assets $(H^1, \dots, H^n)$. We denote by ε_n^2 and ε_{n+1}^2 the minimal hedging error achieved with supplementary assets $(H^1, \dots, H^n)$ and $(H^1, \dots, H^{n+1})$, respectively.

Proposition 2.5 (Relative Hedge Contribution)**.** *Suppose that the non-redundancy condition holds true for all supplementary assets $H^1, \dots, H^{n+1}$. Then, the relative hedge contribution RHC_{n+1} of H^{n+1} is given by*

$$RHC_{n+1}^2 := \frac{\varepsilon_n^2 - \varepsilon_{n+1}^2}{\varepsilon_n^2} = \frac{\left(\mathrm{Cov}\left[L_T^{n+1}, L_T^0\right] - K^\top C^{-1} B\right)^2}{\left(\mathrm{Var}\left[L_T^{n+1}\right] - K^\top C^{-1} K\right)\left(A - B^\top C^{-1} B\right)} \in [0, 1], \tag{13}$$

where $K \in \mathbb{R}^n$ with $K_i = \mathrm{Cov}(L_T^i, L_T^{n+1})$ for $i = 1, \dots, n$.

Remark 2.6. The expression for the relative hedge contribution has an intuitive interpretation under the assumption that the residuals $(L_T^0, \dots, L_T^{n+1})$ have multivariate normal distribution. In this case the hedge contribution of H^{n+1} is equal to the *partial correlation* $\mathrm{Cor}(L_T^0, L_T^{n+1} | L_T)$ of L_T^0 and L_T^{n+1}, given L_T (cf. Muirhead, 2009, Ch. 5.3). Thus, roughly speaking, a supplementary asset has a high hedge contribution, if it is strongly correlated with H^0, even after conditioning on all claims that are attainable with semistatic strategies in S and $(H^1, \dots, H^n)$.

Proof. We set $B^{\mathrm{new}} := [B^\top, \mathrm{Cov}[L_T^{n+1}, L_T^0]]^\top$; $C^{\mathrm{new}} := (C_{i,j}^{\mathrm{new}})_{i,j=1,\dots,n+1}$, where $C_{i,j}^{\mathrm{new}} := \mathrm{Cov}[L_T^i, L_T^j]$, $i, j = 1, \dots, n+1$. Then, we have

$$C^{\mathrm{new}} = \begin{bmatrix} C & K \\ K^\top & \mathrm{Var}\left[L_T^{n+1}\right] \end{bmatrix},$$

which is invertible due to the non-redundancy condition. Write $M = (\mathrm{Var}[L_T^{n+1}] - K^\top C^{-1} K)$ for the Schur complement of C in C^{new}. Using the Schur complement, the inverse of the block matrix C^{new} can be written as

$$(C^{\mathrm{new}})^{-1} = \begin{bmatrix} C^{-1} + C^{-1} K K^\top C^{-1} M^{-1} & -C^{-1} K M^{-1} \\ -K^\top C^{-1} M^{-1} & M^{-1} \end{bmatrix},$$

cf. Horn and Johnson (2012), and applying Theorem 2.3 yields $\varepsilon_{n+1}(u)^2 = A - (B^{\text{new}})^\top (C^{\text{new}})^{-1} B^{\text{new}}$. Consequently,

$$\varepsilon_n(u)^2 - \varepsilon_{n+1}(u)^2 = (B^{\text{new}})^\top (C^{\text{new}})^{-1} B^{\text{new}} - B^\top C^{-1} B$$

and further algebraic manipulation yields (13). □

2.3 | Hedging under $\mathbb{Q}$ vs. $\mathbb{P}$

The hedging criteria in (1) and (5) of minimizing the risk-neutral variance may seem "unnatural" at first sight, compared with the arguably more natural criterion of minimizing the variance under $\mathbb{P}$. After all, the $\mathbb{P}$-variance represents a quantification of the "true" risk associated to a hedging portfolio, which may be different from its risk-neutral counterpart. This discrepancy has been much discussed in the literature, with a good overview to be found in Cont and Tankov (2004, section 10.4.3). We also remark that the variance-optimal hedging principle has indeed been extended to the "physical setting" under $\mathbb{P}$ (see the surveys of Pham, 2000 and Schweizer, 2001), albeit at the cost of an even more demanding mathematical theory. Instead of pursuing the path of variance-optimal hedging under $\mathbb{P}$, we present two arguments why the risk-neutral objectives (1) and (5) are a sensible choice after all, in particular with applications in mind:

- Minimizing the hedging error under $\mathbb{P}$ requires an accurate statistical estimation of the drift of the underlying asset(s), which is a challenging task from an econometric point of view. The risk-neutral pricing measure $\mathbb{Q}$ on the other hand can usually be inferred without difficulty from observed options prices ("model calibration"). Also the fact that an infinitude of equivalent local martingale measures exists in an incomplete market setting is of little concern in practice, as the choice of a parameterized model (e.g., the Heston stochastic volatility model) reduces the choice of $\mathbb{Q}$ to an optimization problem in a low-dimensional parameter space, which can easily be solved by numerically, cf. Cont (2010).
- Contrary to portfolio optimization, where long time horizons prevail and views on the growth rate of assets (i.e., on the "$\mathbb{P}$-world") are indispensable, the manager of a hedging portfolio remains focused on shorter time horizons and will avoid to make his/her choice of hedging portfolio dependent on subjective views on asset growth. Therefore, risk-neutral variance minimization may be in better correspondence to the actual objectives of the manager of a hedging portfolio.

2.4 | The variance swap and long/short constraints

We review the semistatic hedging of a variance swap with an infinite pool of European put- and call-options, as discussed in Neuberger (1994) and Carr and Madan (2001). We will apply the methods developed in this paper to this hedging problem in Section 5. It also serves as a motivation to add long/short constraints to the semistatic hedging problem.

Recall that a *variance swap* is a contingent claim on an underlying traded asset S, which at maturity T pays an amount $H_T^{\text{swap}} := [\log S, \log S]_T - k$, where $k \in \mathbb{R}$. Usually, k is chosen such that the value of the contract is zero at inception, and the corresponding value $k_* = \mathbb{E}[[\log S, \log S]_T]$ is called the *swap rate*. Recall that our only assumption on the discounted price process S is that it is a square-integrable strictly positive continuous martingale. Applying Itô's formula to $\log S_T$ we get

$$\log S_T = \log S_0 + \int_0^T \frac{1}{S_t}\,\mathrm{d}S_t - \frac{1}{2}\int_0^T \frac{1}{S_t^2}\mathrm{d}[S,S]_t. \tag{14}$$

Hence,

$$H_T^{\text{swap}} = \int_0^T \frac{1}{S_t^2} \mathrm{d}[S,S]_t - k = 2\int_0^T \frac{1}{S_t}\,\mathrm{d}S_t - 2\log\frac{S_T}{S_0} - k,$$

that is, to hedge the variance swap, it is enough to dynamically trade in the stock S and enter a static position in the "log-contract" with payoff $\log\frac{S_T}{S_0}$, cf. Neuberger (1994). Furthermore, from Carr and Madan (2001), we have

$$\log\frac{S_T}{S_0} = \frac{S_T - S_0}{S_0} - \int_0^{S_0} \frac{(K-S_T)^+}{K^2}\,\mathrm{d}K - \int_{S_0}^{\infty} \frac{(S_T-K)^+}{K^2}\,\mathrm{d}K\,, \tag{15}$$

which inserting into the above equation yields

$$H_T^{\text{swap}} = 2\int_0^T \left(\frac{1}{S_t} - \frac{1}{S_0}\right)\mathrm{d}S_t - k + 2\int_0^{S_0} \frac{(K-S_T)^+}{K^2}\,\mathrm{d}K + 2\int_{S_0}^{\infty} \frac{(S_T-K)^+}{K^2}\,\mathrm{d}K. \tag{16}$$

This equality can be interpreted as a semistatic replication strategy for the variance swap, which uses a dynamic position in S and a static portfolio of *infinitesimally small* positions in an *infinite number* of out-of-the-money puts and out-of-the-money calls. We make several observations:

(a) For any practical implementation the "infinitesimal portfolio" has to be discretized and portfolio weights have to be assigned to each put and call.

(b) As they are calls and puts on the same underlying asset, the static hedging assets are highly correlated.

(c) The static positions in puts and calls are long positions only.

To address point (a) different ad hoc discretizations of the integrals in (16) are possible (e.g., left or right Riemann sums, trapezoidal sums). However, it is not obvious which discretization is optimal in the sense of minimizing the hedging error. The choice of an optimal discretization in the variance-minimizing sense is precisely given by Theorem 2.3.

Point (b) suggests that given a moderate number (say 30) of puts and calls as static hedging assets, many of them will be redundant in the sense that their hedge contribution (given the other supplementary assets) is small. This observation motivates the sparse approach of the next section and will be confirmed numerically in the application Section 5.

Point (c) finally motivates the addition of short/long constraints, or more generally, linear constraints of the type

$$v^\top p \geq 0, \tag{17}$$

where $p \in \mathrm{R}^n$ is fixed, to the outer problem in (6). With these constraints, the outer problem is a linearly constrained quadratic optimization problem, which can still be efficiently solved by standard numerical software.

3 | SPARSE SEMISTATIC HEDGING

We now focus on the problem of optimal selection of static hedging assets, as outlined in the introduction and motivated in the previous section. Note that the subset selection only affects the static part of

the strategy and hence only the outer problem in (6). Recall the ℓ_1-norm $\|v\|_1 = \sum_{i=1}^n |v_i|$ on $\mathbb{R}^n$ and the (nonconvex) ℓ_0-quasinorm $\|v\|_0$, which counts the number of nonzero elements of v, cf. Foucart and Rauhut (2013).

Definition 3.1 (Sparse Variance-Optimal Semistatic Hedging Problem)**.** The *sparse variance-optimal semistatic hedge* $(\vartheta, v) \in \mathrm{L}^2(S) \times \mathbb{R}^n$ with effective portfolio size $d < n$ and its optimal initial capital $c \in \mathbb{R}$ are the solution of the minimization problem (6), with the outer problem replaced by

$$\epsilon^2 = \min_{v \in \mathbb{R}^n, v \geq 0} (v^\top C v - 2v^\top B + A), \quad \text{subj. to} \quad \|v\|_0 \leq d. \qquad (\ell_0\text{-constrained problem}) \tag{18}$$

The ℓ_1-relaxation of this problem is given by

$$\epsilon^2 = \min_{v \in \mathbb{R}^n, v \geq 0} (v^\top C v - 2v^\top B + A) + \lambda \|v\|_1, \qquad (\ell_1\text{-relaxation}). \tag{19}$$

where $\lambda > 0$ is a tuning parameter that replaces d. In both problems, we allow for long/short contains of the form (17).

The minimization problem (18) is of course equivalent to the extensively studied subset selection problem in linear regression and (19) to its convex relaxation in Lagrangian form, usually called LASSO. We refer to Hastie et al. (2013) for a general overview and to Tibshirani (1996) for the LASSO. We emphasize that

- The ℓ_0-constrained subset selection problem (18) is nonconvex and hard to solve exactly if the dimension n is high.
- The ℓ_1-penalized minimization problem (19) is convex and efficient numerical solvers exist even for large n. Its solution is usually a good approximation to the exact subset selection problem, but no guarantee of being close to the solution of (18) can be given in general.[2]

To illustrate the effect of the ℓ_1-penalty, denote by v_* the solution of the unpenalized hedging problem (5) and assume for a moment that all GKW-r esiduals $(L_T^0, \dots, L_T^d)$ are uncorrelated. This assumption is highly unrealistic in the hedging context, but leads to a simple form of the solution of the penalized problem, cf. Tibshirani (1996): It is given by $v' = \mathrm{sign}(v_*)(|v_*| - \lambda)^+$, i.e., all static positions are shrunk toward zero by λ and truncated when zero is reached. This nicely illustrates the sparsifying effect of the penalty and the role of λ.

While (19) is frequently used as a surrogate for (18), the following alternatives exist for solving (18) directly, or for approximating its solution. Again, we refer to Tibshirani (1996) for further details on the described methods:

Brute-Force Solve the quadratic optimization problem for each possible subset of cardinality d. As there are $\binom{n}{d}$ of these subsets, this approach is usually not efficient and becomes completely infeasible for large n.

Leaps-and-Bounds "Leaps-and-Bounds" is a branch-and-bound algorithm introduced by Furnival and Wilson (1974) for subset selection in linear regression, which gives an exact solution to (18) without testing all possible subsets. Essentially, a tree of nested subsets of hedging instruments (or regression variables) is formed, some branches of which can be discarded a priori, see Furnival and Wilson (1974) for details. While the worst case performance is no better than in the brute-force approach, it tends to be much faster in many practical settings.

Greedy Forward Selection A simple greedy approximation to (18) is to assume that the optimal subsets of different cardinality are nested. In the forward approach the problem (18) is first solved for $d = 1$, which is easy. Then, iteratively, the supplementary claim with the largest relative hedge contribution (see (2.5)) is added to the set of active static positions in each step. The same procedure could be used backward ("greedy backward selection") i.e., starting with $d = n$ and then removing iteratively the supplementary claim with the smallest hedge contribution. In general, no guarantee of being close to the exact solution of (18) can be given for these methods.

We will compare the practical performances of the different solution methods in Section 5.

4 | STOCHASTIC VOLATILITY MODELS WITH FOURIER REPRESENTATION

The final ingredient still missing for a numerical solution of the (sparse) semistatic hedging problem is an efficient method to compute the quantities A, B, and C from Theorem 2.3. One possible approach would be to compute transition densities of S and $H^0, \dots, H^n$ by Monte Carlo simulation and to compute the GKW decomposition by sequential backward regression, cf. Föllmer and Schweizer (1988). Due to the fact that the joint distribution of S and all price processes of supplementary claims is needed, we expect a heavy computational load in order to obtain reasonably high accuracy with this method. An interesting alternative method has been suggested by Kallsen and Pauwels (2010) (see also Hubalek, Kallsen, & Krawczyk, 2006; Pauwels, 2007) for the classic variance-optimal hedging problem (1) of European claims. This alternative is based on the well-known Fourier method for pricing of European claims, cf. Carr and Madan (2001); Raible (2000); Kallsen and Pauwels (2010).

4.1 | Fourier representation of strategies and hedging errors

We stay close to the framework of Kallsen and Pauwels (2010) and assume that the payoff of some option H is given by $H = f(X_T)$, where X is the log-price process of the underlying stock, i.e., we also assume $S = \exp(X)$. The payoff of a call for example can be written as $f(x) = (e^x - K)^+$, but it is not necessary to restrict to this specific case. Furthermore, we assume that the (rescaled) two-sided Laplace transform

$$\tilde{f}(u) = \frac{1}{2\pi \mathrm{i}} \int_{-\infty}^{+\infty} \exp(-ux) f(x) \mathrm{d}x \tag{20}$$

of the payoff exists at some $u = R \in \mathbb{R}$ and is integrable over the strip

$$S(R) := \{u \in \mathbb{C} : \operatorname{Re} u = R\}$$

in the complex plane. If the integrability condition $\mathbb{E}[e^{RX_T}] < \infty$ holds, then the risk-neutral price of the claim H at time $t \in [0, T]$ can be recovered by the Fourier-type integral

$$H_t = \int_{S(R)} H_t(u) \tilde{f}(u) \mathrm{d}u, \tag{21}$$

along $\mathcal{S}(R)$, where we denote the conditional moment generating function (analytically extended to the complex plane) of X_T by

$$H_t(u) := \mathbb{E}[e^{uX_T}|\mathcal{F}_t].$$

Note that $H_t(u)$ is well defined on $\mathcal{S}(R)$ due to the integrability condition imposed on X_T. In the important cases of European puts and calls, the two-sided Laplace transform $\tilde{f}$ is given by

$$\tilde{f}(u) = \frac{1}{2\pi \mathrm{i}} \frac{K^{1-u}}{u(u-1)},$$

with $R > 1$ for calls and $R < 0$ for puts, (cf. Hubalek et al., 2006, section 4).

The key insight, pioneered by Hubalek et al. (2006) for variance-optimal hedging in models with independent increments and by Kallsen and Pauwels (2010) and Pauwels (2007) for affine stochastic volatility models, is that the Fourier representation (21) of European claims can be extended to their GKW decomposition (2). More precisely, both the strategy ϑ and the hedging error $\epsilon^2 = \mathbb{E}[L_T^2]$ of the variance-optimal hedging problem (1) can be expressed in terms of Fourier-type integrals, similar to (21). For our problem of interest, the *semistatic* hedging problem (5), the results of Hubalek et al. (2006), Kallsen and Pauwels (2010), and Pauwels (2007) are not sufficient: To obtain the quantities A, B, and C of Theorem 2.3, we also need to compute the covariances $\mathbb{E}[L_T^i L_T^j]$ between the GKW residuals of different claims. In the companion paper, Di Tella et al. (2019), we extend the results of Hubalek et al. (2006), Kallsen and Pauwels (2010), and Pauwels (2007) to the semistatic hedging problem. Moreover, we show that the method can be used in any stochastic volatility models where the Fourier transform of the log-price X is known (e.g., the Heston, the 3/2 or the Stein–Stein model, cf. Lewis, 2000). Here, we only need a special case of the more general results in Di Tella et al. (2019), which is condensed into Theorem 4.1(i) below.

In order to formulate the representation result, we assume that a claim H^0 (e.g., a variance swap), and supplementary assets $H^1, \dots, H^n$ with Fourier representations (20) are given, and define for $u, u_1, u_2 \in \mathbb{C}$ complex-valued predictable processes of finite variation $\mathcal{A}, \mathcal{B}(u), \mathcal{C}(u_1, u_2)$ by

$$\mathrm{d}\mathcal{A} = \mathrm{d}\langle H^0, H^0\rangle - \vartheta^0\vartheta^0 \mathrm{d}\langle S, S\rangle, \qquad \mathcal{A}_0 = 0 \tag{22a}$$

$$\mathrm{d}\mathcal{B}(u) = \mathrm{d}\langle H^0, H(u)\rangle - \vartheta^0\vartheta(u)\mathrm{d}\langle S, S\rangle, \qquad \mathcal{B}_0(u) = 0, \tag{22b}$$

$$\mathrm{d}\mathcal{C}(u_1, u_2) = \mathrm{d}\langle H(u_1), H(u_2)\rangle - \vartheta(u_1)\vartheta(u_2)\mathrm{d}\langle S, S\rangle, \qquad \mathcal{C}_0(u_1, u_2) = 0. \tag{22c}$$

Theorem 4.1. *Let a stochastic volatility model with forward price process $S = e^X$ and variance process V be given, and let $T > 0$ be a fixed time horizon. Let H^0 be a variance swap with payoff $[X, X]_T = \int_0^T V_t \mathrm{d}t$ and let the supplementary assets $(H^1, \dots, H^n)$ be European puts or calls with Fourier representations given by (20). Assume that (S, V) are continuous square-integrable semimartingales and that there exist functions $h(u, t, V_t)$, $\gamma(t, V_t)$, continuously differentiable in the last component, such that*

$$H_t(u) = \mathbb{E}[e^{uX_T}|\mathcal{F}_t] = e^{uX_t} h(u, T-t, V_t), \quad F_t := \mathbb{E}\big[[X, X]_T - [X, X]_t \big| \mathcal{F}_t\big] = \gamma(T-t, V_t). \tag{23}$$

Then, the following holds true:

(i) The quantities A, B, and C in Theorem 2.3 can be represented as

$$A = \mathbb{E}\left[\mathcal{A}_T\right] \tag{24a}$$

$$B_i = \int_{S(R_i)} \mathbb{E}\left[\mathcal{B}_T(u)\right] \tilde{f}_i(u)\mathrm{d}u, \qquad i \in \{1, \dots, n\} \tag{24b}$$

$$C_{ij} = \int_{S(R_i)} \int_{S(R_j)} \mathbb{E}\left[C_T(u_1, u_2)\right] \tilde{f}_i(u_1)\tilde{f}_j(u_2)\mathrm{d}u_1\mathrm{d}u_2, \qquad (i, j) \in \{1, \dots, n\}^2. \tag{24c}$$

(ii) The processes (22) can be written as

$$\mathrm{d}\mathcal{A}_t = \left(\partial_v \gamma(T - t, V_t)\right)^2 dQ_t, \tag{25a}$$

$$\mathrm{d}\mathcal{B}_t(u) = e^{uX_t}\, \partial_v h(u, T - t, V_t)\, \partial_v \gamma(T - t, V_t)\, dQ_t \tag{25b}$$

$$\mathrm{d}C_t(u_1, u_2) = e^{(u_1+u_2)X_t}\, \partial_v h(u_1, T - t, V_t)\, \partial_v h(u_2, T - t, V_t)\, dQ_t, \tag{25c}$$

where

$$dQ = \mathrm{d}[V, V] - \frac{\mathrm{d}[X, V]}{\mathrm{d}[X, X]}\mathrm{d}[X, V]. \tag{26}$$

Proof. Part (i) of the theorem is technically demanding and follows from theorems 4.5, 4.6, and 4.8 in the companion paper Di Tella et al. (2019). In order to show part (ii), let Y be a $\mathbb{R}^n$-valued continuous semimartingale and let α, β be functions in $C^2(\mathbb{R}^n, \mathbb{C})$. Using Ito's formula (cf. Jacod & Shiryaev, 2003, theorem I.4.57) and the properties of quadratic covariation (cf. Jacod & Shiryaev, 2003, theorem I.4.49) we obtain the calculation rule

$$\mathrm{d}[\alpha(Y), \beta(Y)] = \sum_{i,j} \partial_{y_i}\alpha(Y)\partial_{y_j}\beta(Y)\mathrm{d}[Y_i, Y_j]. \tag{27}$$

Inserting the definition of the variance-optimal strategy $\vartheta(u) = \frac{\mathrm{d}\langle H(u), S\rangle}{\langle S, S\rangle}$ into (22) and recognizing that for continuous martingales predictable variation $\langle .,. \rangle$ and quadratic variation $[.,.]$ coincide, we obtain

$$dC(u_1, u_2) = \mathrm{d}[H(u_1), H(u_2)] - \frac{\mathrm{d}[H(u_1), S]}{\mathrm{d}[S, S]}\mathrm{d}[H(u_2), S]$$

for C and similar expressions for B and A. Using assumption (23) and applying (27) several times we obtain (25). □

4.2 | The Heston model

In the Heston model (cf. Heston, 1993) the risk neutral price process S is given by $S_t = S_0 \exp(X_t)$, $t \geq 0$, where

$$\mathrm{d}X_t = -\frac{1}{2}V_t\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t^1, \tag{28a}$$

$$dV_t = -\lambda(V_t - \kappa)dt + \sigma\sqrt{V_t}dW_t^2, \tag{28b}$$

where W^1 and W^2 are two Brownian motions such that $\langle W^1, W^2\rangle_t = \rho t$, $\rho \in [-1, 1]$; $\lambda, \sigma, \kappa > 0$.

The joint moment generating function of the Heston model is known explicitly and of the form

$$\mathbb{E}[\exp(uX_T + wV_T)] = \exp(\phi_T(u, w) + \psi_T(u, w)V_0 + uX_0), \tag{29}$$

well defined for real arguments in the set

$$\mathcal{D}_T := \{(u, w) \in \mathbb{R}^2 : \mathbb{E}[\exp(uX_T + wV_T)] < \infty\}, \tag{30}$$

and with analytic extension to the associated "complex strip"

$$S(\mathcal{D}_T) := \{(u, w) \in \mathbb{C}^2 : (\mathrm{Re}\, u, \mathrm{Re}\, w) \in \mathcal{D}_T\}.$$

To represent $\phi_t(u, w)$ and $\psi_t(u, w)$, we introduce $\Delta(u) = (\rho\sigma u - \lambda)^2 - \sigma^2(u^2 - u)$,

$$r_\pm = r_\pm(u, w) := \frac{1}{\sigma^2}\left(\lambda - \rho\sigma u \pm \sqrt{\Delta(u)}\right) \qquad g = g(u, w) = \frac{r_- - w}{r_+ - w}.$$

Then, the explicit expression of ψ_t is given, for $(u, w) \in \mathcal{D}_t$ by (cf. Alfonsi, 2015, proposition 4.2.1),

$$\psi_t(u, w) := \begin{cases} w + (r_- - w)\frac{1-\exp\left(-t\sqrt{\Delta}\right)}{1-g\exp(-t\sqrt{\Delta})}, & \Delta(u) \neq 0; \\ w + (r_- - w)^2\frac{\sigma^2 t}{2+\sigma^2 t(r_- - w)}, & \Delta(u) = 0, \end{cases} \tag{31}$$

with the convention

$$\frac{\exp(-t\sqrt{\Delta}) - g}{1 - g} := 1, \qquad \frac{1 - \exp(t\sqrt{\Delta})}{1 - g\exp(t\sqrt{\Delta})} := 0$$

whenever the denominator of g is equal to zero. Moreover, $\phi_t(u, w)$ is given by

$$\phi_t(u, w) := \begin{cases} \lambda\kappa r_- t - \frac{2\lambda\kappa}{\sigma^2}\log\left(\frac{1-g\exp(-t\sqrt{\Delta})}{1-g}\right) & \Delta(u) \neq 0; \\ \lambda\kappa r_- t - \frac{2\lambda\kappa}{\sigma^2}\log\left(1 + \frac{\sigma^2}{2}(r_- - w)t\right) & \Delta(u) = 0. \end{cases} \tag{32}$$

The following theorem specializes Theorem 4.1 to the Heston model and gives (up to integration) explicit expressions for the quantities A, B, and C from Theorem 2.3. The proof of the theorem is given in Appendix A.

Theorem 4.2. *Let (X, V) be given by the Heston model (28) and let the claim H^0 be a variance swap, i.e., with payoff $H_T^0 = [X, X]_T$ at maturity T. Let the supplementary claims $(H^1, \dots, H^n)$ be European puts and calls with payoffs f_i and two-sided Laplace transforms $\tilde{f}_i$, integrable along strips $S(R_i)$, as in (21). If $\mathbb{E}[e^{2R_i X_T}] < \infty$ for all $i = 1, \dots, n$ then the quantities A, B, and C, defined in Theorem 2.3 are given by*

$$A = \frac{\sigma^2(1-\rho^2)}{\lambda^2}\int_0^T \left(1 - e^{-\lambda(T-t)}\right)^2 \mathbb{E}\left[V_t\right] dt,$$

$$B_i = \frac{\sigma^2(1-\rho^2)}{\lambda}\int_0^T \int_{S(R_i)} \left(1 - e^{-\lambda(T-t)}\right)\psi_{T-t}(u)\mathbb{E}\left[H_t(u)V_t\right]\tilde{f}_i(u)\,du\,dt,$$

$$C_{ij} = \sigma^2(1-\rho^2)\cdot$$

$$\int_0^T \int_{S(R_i)}\int_{S(R_j)} \psi_{T-t}(u_1)\psi_{T-t}(u_2)\mathbb{E}\left[H_t(u_1,u_2)V_t\right]\tilde{f}_i(u_1)\tilde{f}_j(u_2)\,du_1\,du_2\,dt,$$

where

$$\mathbb{E}\left[V_t\right] = e^{-\lambda t}V_0 + \left(1 - e^{-\lambda t}\right)\kappa \tag{33a}$$

$$\mathbb{E}\left[H_t(u)V_t\right] = \left\{\partial_w\phi\left(u, \psi_{T-t}(u,0)\right) + V_0\partial_w\psi\left(u, \psi_{T-t}(u,0)\right)\right\}e^{uX_0}h(u,t,V_0), \tag{33b}$$

$$\mathbb{E}\left[H_t(u_1)H_t(u_2)V_t\right] = \left\{\partial_w\phi\left(u, q_{T-t}(u_1,u_2)\right) + V_0\partial_w\psi\left(u, q_{T-t}(u_1,u_2)\right)\right\}\cdot \tag{33c}$$

$$e^{(u_1+u_2)X_0}h(u_1,t,V_0)h(u_2,t,V_0),$$

with

$$q_t(u_1,u_2) = \psi_t(u_1,0) + \psi_t(u_2,0), \qquad h(u,t,V_0) = \exp\left(\phi_t(u,0) + V_0\psi_t(u,0)\right).$$

Remark 4.3. Note that the common leading factor $\sigma^2(1-\rho^2)$ of A, B, and C also becomes the leading factor of the minimal squared hedging error ϵ^2, cf. (12). This makes perfect sense, as it makes the hedging error roughly proportional to vol-of-vol σ and shows that the hedging error vanishes in the complete-market boundary cases $\rho = \pm 1$ of the Heston model. However, ρ and σ also appear inside ϕ, ψ and therefore their influence on ϵ^2 is not limited to the leading factor $\sigma^2(1-\rho^2)$ alone.

Remark 4.4. The domain D_T of finite moments in the Heston model has been described in Andersen & Piterbarg (cf. 2007, proposition 3.1) (see also Friz & Keller-Ressel, 2010). Using these results, the moment condition in Theorem 4.2 can be checked in the following way: Set $\chi(u) := \rho\sigma u - \lambda$, $\Delta(u) := \chi(u)^2 - \sigma^2(u^2 - u)$ and define

$$T_*(u) = \begin{cases} +\infty\,, & \Delta(u) \geq 0,\ \chi(u) < 0; \\ \frac{1}{\sqrt{\Delta(u)}}\log\left(\frac{\chi(u)+\sqrt{\Delta(u)}}{\chi(u)-\sqrt{\Delta(u)}}\right), & \Delta(u) \geq 0,\ \chi(u) > 0; \\ \frac{2}{\sqrt{-\Delta(u)}}\left(\arctan\frac{\sqrt{-\Delta(u)}}{\chi(u)} + \pi 1_{\{\chi(u)<0\}}\right), & \Delta(u) < 0\,. \end{cases} \tag{34}$$

By Andersen & Piterbarg (2007, prop. 3.1), the moment condition $\mathbb{E}[e^{2R_iX_T}] < \infty$ is equivalent to $T < T_*(2R_i)$.

5 | NUMERICAL RESULTS

The following numerical implementation should be considered in terms of a 'stylized financial market' setting, i.e., while we do not calibrate the model to current market data, we use parameters that are realistic in a market setting. More specifically, we use the Heston model parameters from Gatheral (2006):

$$\begin{aligned} \kappa &= 0.0354 \qquad \lambda = 1.3253 \qquad \rho = -0.7165 \\ \sigma &= 0.3877 \qquad V_0 = 0.0174. \end{aligned} \tag{35}$$

In Section 5.4 we vary the leverage parameter ρ, but keep all other parameters fixed. The current stock price is normalized to $S_0 = 100$ and we use a time-to-maturity of $T = 1$ (years) for the variance swap and the call options. The price of a variance swap (i.e., the swap rate $k_* = \mathbb{E}[[\log S, \log S]_T]$) can be readily calculated as

$$k_* = \int_0^T \mathbb{E}\big[V_t\big] dt = \kappa T + (V_0 - \kappa)\frac{1 - e^{-\lambda T}}{\lambda} = 0.025427.$$

The supplementary assets are OTM-puts and OTM-calls with strikes ranging from

$$K_{\min} = 50 \text{ to } K_{\max} = 150 \text{ in steps of } \Delta K = 5.$$

We focus on three aspects of the semistatic hedging problem:

- Comparing the different methods that were proposed in Section 3 to solve the sparse semistatic hedging problem;
- Analyzing the dependency of hedging error and optimal portfolio composition on effective portfolio size d;
- Analyzing the dependency of hedging error and optimal portfolio composition on the leverage parameter ρ.

5.1 | Comparison of methods

As a first step, we computed A, B, and the matrix C from Theorem 2.3 using the Fourier representation in Theorem 4.2 by adaptive integration in MATLAB. Next, we implemented the methods, described in Section 3, i.e.,

(1) Greedy forward selection (with and without short-sale constraints)

(2) Leaps-and-Bounds (with and without short-sale constraints)

(3) LASSO

in the statistical computing environment `R`; using the function `lars` in the package `lars` with option `type=''lasso''` for the computation of the LASSO solution. While computationally most demanding, the Leaps-and-Bounds solution can serve as a benchmark solution, as it is (up to numerical error) the exact solution of the sparse semistatic hedging problem (18). The other methods, in contrast, only return a "reasonably close" solution to (18). In all cases, we report the *relative hedging error* ϵ/k_*, i.e., the hedging error normalized by the price of the variance swap.

A challenge that is faced by all methods is the bad condition of the matrix C. With parameters chosen as in (35) the reciprocal condition number of C is 1.11×10^{-6}. While small, this number is still several orders of magnitude larger than the machine precision of 2.22×10^{-16} (double precision arithmetic) on the computer that was used. The bad condition of C is not surprising, as put and call options with neighboring strikes are highly correlated. This effect is likely amplified by the fact that C contains the correlations of the GKW residuals and not the correlations of the option prices themselves. While we have considered preconditioning of C, along the lines of Neumaier (1998), we have found that greedy

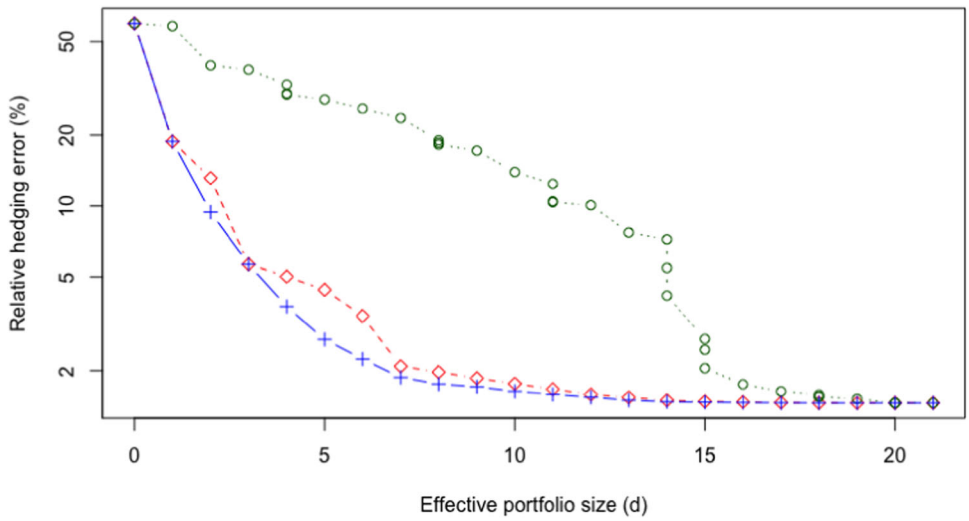

FIGURE 1 Relative hedging error (on log-scale) for sparse semistatic hedging of a variance swap with different effective portfolio sizes d. The plot compares the solutions obtained with the Leaps-and-Bounds method (blue crosses), greedy method (red diamonds) and LASSO (green circles) [Color figure can be viewed at wileyonlinelibrary.com]

forward selection and Leaps-and-Bounds perform well even without additional conditioning. Also the addition of short-sale constraints seems to have a regularizing effect on the methods.

Figure 1 shows the relative hedging error (as percentage of the variance swap price) attained with the optimal portfolio returned by Methods 1–3 for different effective portfolio sizes $d = 0 \dots 21$. Notice that the implementation of LASSO adds *and removes* supplementary assets from the active set, such that the graph can show multiple solutions for the same effective portfolio size (e.g., for $d = 15$). Focusing on the comparison of methods, we find that

- The Leaps-and-Bounds method returns the solution with the smallest hedging error, consistent with the fact that it solves (18) *exactly*. It is remarkably fast, but further numerical experiments indicate that its runtime is sensitive to the choice of model parameters.
- The greedy method is the fastest method and the residual hedging error of its solution is only slightly higher than the hedging error of the Leaps-and-Bounds solution. Moreover, the performance of the greedy method is stable with respect to parameter choice.
- The LASSO methods seems to be severely affected by the bad condition of C. This is not surprising, as it has been remarked, for instance, in Bühlmann & Van De Geer (2011, section 2.6) that the LASSO method has problems with highly correlated data.

Summing up, we can recommend the greedy method as fast, reliable, and easy to implement. The Leaps-and-Bounds method is useful as an efficient way to compute an exact benchmark solution. We cannot recommend LASSO, as it cannot deal well with the bad condition of C.[3] Interestingly, this observations are contrary to the usual wisdom in variable selection for regression problems, where greedy forward selection often has unstable performance and LASSO yields superior results (cf. Bühlmann & Van De Geer, 2011, chapter 2). We attribute these findings to the highly correlated nature of the matrix C, which is untypical in regression scenarios, but a natural feature of our hedging problem.

5.2 | Analysis of the hedging error

We return to Figure 1 to analyze the hedging error resulting from the sparse variance-optimal semistatic hedging problem (18) for different effective portfolio sizes d. We consider the benchmark solution returned by the Leaps-and-Bounds method with short-sale constraints. First, we note that dynamic hedging in the underlying S, *without* using any static positions in puts and calls ($d = 0$), results in a relative hedging error of 59.7%. This error is already reduced to 5.7% by just adding three supplementary assets ($d = 3$) and can be further reduced to 3.4% by selecting six supplementary assets ($d = 6$). Finally, the error levels off to 1.6% when the full range ($d = 21$) of puts and calls between $K_{\min} = 50$ and $K_{\max} = 150$ is used. Further substantial reductions of the hedging error can only be achieved by extending the range of available strikes; adding more options within the current range has only negligible effects.

The sharp decrease of the hedging error between $d = 0$ and $d = 3$ affirms the basic premise of *sparse semistatic hedging*: That selecting only a small number of supplementary assets already leads to a significant reduction of the hedging error. On the other hand, the poor performance of the LASSO solution shows that a suboptimal choice of supplementary assets does not result in a satisfactory reduction of the hedging error. In other words, it is important that the sparse subportfolios are chosen optimally, and not arbitrarily.

5.3 | Composition of the hedging portfolios

We now turn to the composition of the static hedging portfolio, i.e., the vector $v \in \mathbb{R}^n$ with the constraint $\|v\|_0 \leq d$, that is returned by the solution methods for the sparse semistatic hedging problem (18). Recall that the element v_i is the nominal size of the position in the supplementary asset H^i, with negative sign indicating a short position. In our setting, the elements of v can simply be indexed by the strike K of the corresponding put/call. The optimal portfolios returned by the different solution methods, along with their dependency on effective portfolio size d are shown in Figure 2. We make the following observations:

- With the exception of the put $K = 55$ only long positions are observed;
- Positions in OTM puts ($K < 100$) are larger than in OTM calls ($K > 100$), in line with Neuberger's replicating portfolio (16);
- The general pattern (going from effective portfolio size $d = 1$ to 21) for all methods can be described as follows: Start with an (approximately) ATM option. Proceed by selecting both OTM puts and calls, going outwards as d increases and putting more weight on OTM puts, until the limit $K_{\min} = 50$ is reached. Continue by adding OTM calls and by filling up the gaps from earlier stages.

We suspect that the rare short positions are numerical artifacts, rather than belonging to the true optimal solution of (18). Indeed, their effect on the hedging error is minuscule, and we hence recommend to use a priori short-sale constraints, in the case of hedging a variance swap.

Figure 2 gives a good overview of the portfolio composition, but it is difficult to assess the precise size of the individual positions v_i. For this reason, we provide in Figure 3 an additional plot of the portfolio weights v indexed by strike K for the optimal portfolios of effective sizes $d = 3, 6, 12$ in doubly logarithmic coordinates. Note that Neuberger's replicating portfolio (16) puts an infinitesimal weight of $v(K)dK = \frac{1}{K^2}dK$ on an option with strike K. In doubly logarithmic coordinates, this becomes

$$\log v(K) = -2 \log K,$$

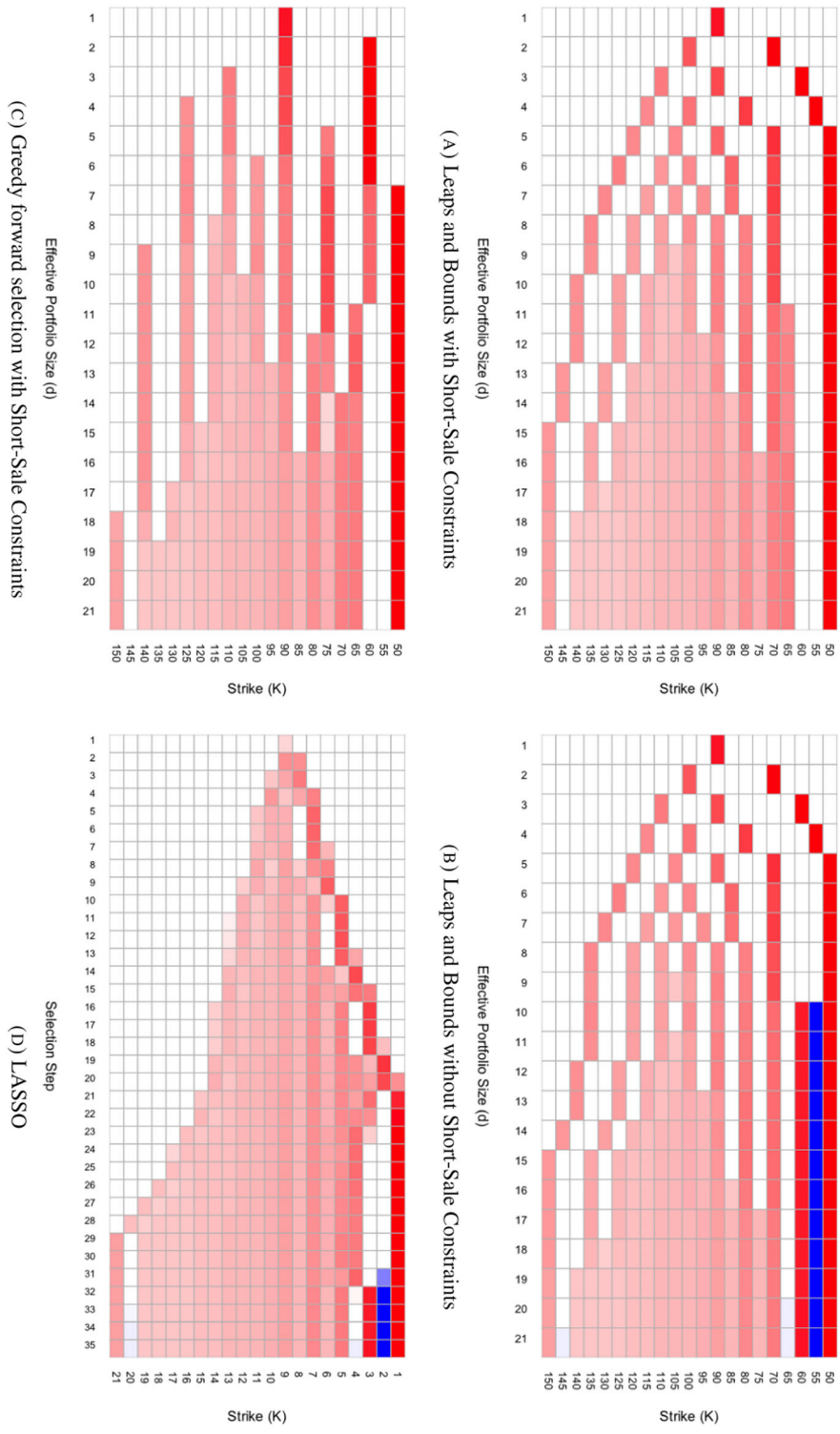

FIGURE 2 Composition of optimal hedging portfolios in dependency on effective portfolio size. Long positions are shown in red and short positions in blue; color saturation corresponds to position size $v(K)$. Different subplots correspond to different solution methods [Color figure can be viewed at wileyonlinelibrary.com]

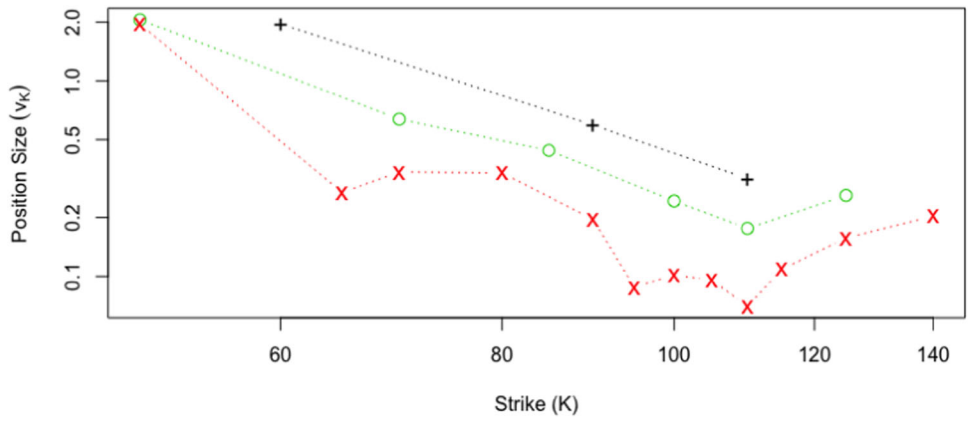

FIGURE 3 Portfolio weights v_K in the optimal hedging portfolios of effective size $d = 3$ (black crosses), $d = 6$ (green circles), and $d = 12$ (red x's) in doubly logarithmic coordinates [Color figure can be viewed at wileyonlinelibrary.com]

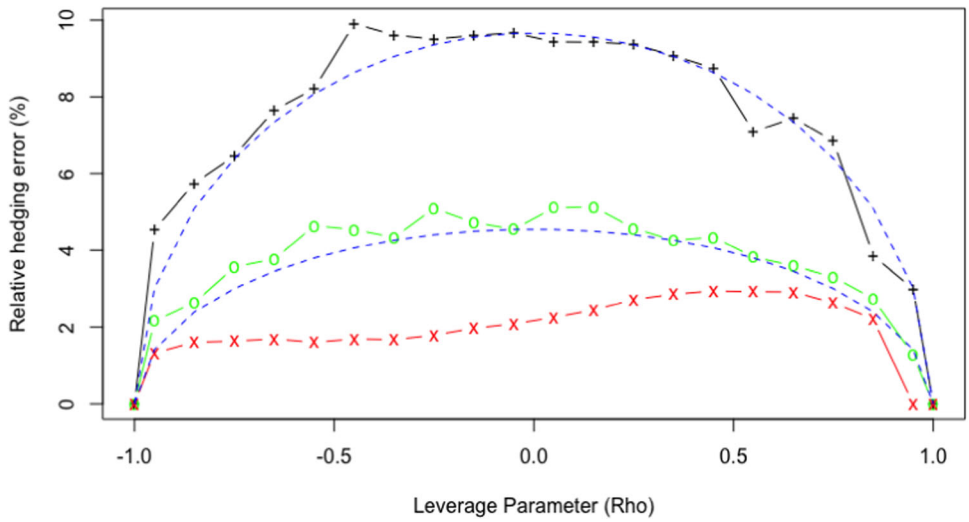

FIGURE 4 Relative hedging error attainable with a portfolio of effective size $d = 3$ (black crosses), $d = 6$ (green circles) and $d = 12$ (red x's) in relation to the leverage parameter ρ. Also shown are the graphs of $f(\rho) = c_d\sqrt{1-\rho^2}$ (blue dashes) with c_d chosen to fit the red and blue graphs [Color figure can be viewed at wileyonlinelibrary.com]

i.e., the portfolio weights should form a line of downward slope -2. Figure 3 shows reasonable agreement with this asymptotic result, even for effective portfolio size as small as $d = 3$. For $d = 12$ numerical errors from the bad condition of the matrix C seem to accumulate and could explain the unruly shape of the graph.

5.4 | The role of correlation

Finally, we turn to the role of the correlation parameter ρ, which is interesting for several reasons: First, the value of ρ does not affect the theoretical price of the variance swap. Second, ρ also does not affect the infinitesimally optimal strategy (16). Finally, ρ allows to tune the degree of market incompleteness,

as the Heston model becomes a complete market model in the boundary cases $\rho = \pm 1$. Despite of the first two points, it turns out that ρ has a significant effect on the attainable hedging error and the composition of the optimal portfolio in the sparse semistatic hedging problem. This influence can already be suspected from the leading factor $1 - \rho^2$ appearing in Theorem 4.2, which propagates to the (squared) hedging error itself, see also Remark 4.3. Indeed, as Figure 4 shows, the dependency of the relative heading error on ρ is very close to a "semi-circle law" $f(\rho) = c_d \sqrt{1 - \rho^2}$, with different constants c_d for different effective portfolio sizes d.

6 | CONCLUSIONS AND OUTLOOK

In this paper, we have introduced a general framework for semistatic and sparse variance-optimal hedging. We have also demonstrated the numerical feasibility of our approach in the concrete example of hedging a variance-swap in the Heston model. In future work, we think it would be interesting to examine the degree of model-dependency of these results, by implementing the same approach for other models with known Fourier transform, e.g., the $3/2$ model of Carr and Sun (2007) or the recent rough Heston model of El Euch and Rosenbaum (2019). Even in stochastic volatility models without an explicit Fourier transform, a combination of PDE and Monte Carlo techniques should in principle allow to compute the quantities A, B, and C in Theorem 2.3 and hence to determine optimal semistatic hedging strategies.

Finally, we identify two other directions, in which we see potential for relevant extensions of our results: In the setup of Section 2 we have assumed that all supplementary assets (H^i) as well as the hedging target H^0 mature at the same time. In practice, also assets with maturity mismatch can be useful for hedging. For example, a variance swap with fixed maturity could be hedged—not perfectly, but variance optimally—by a portfolio of European options with shorter, larger, or even mixed maturities. Moreover, because the supplementary assets (H^i) are typically less liquid than the underlying, it seems reasonable to include a bid–ask spread for these assets. Such bid–ask spreads should only lead to additional linear inequality constraints to the outer problem in (6). As a quadratic optimization problem with linear constraints, the hedging problem should remain numerically feasible, even under bid–ask spreads.

ACKNOWLEDGMENTS

MKR thanks Johannes Muhle-Karbe for discussions on variance-optimal semistatic hedging. We acknowledge funding from the German Research Foundation (DFG) under grant ZUK 64 (all authors) and KE 1736/1-1 (MKR, MH)

ENDNOTES

[1] These connections to linear regression should not come as a surprise: It has been noted already in Föllmer and Schweizer (1988) that variance-optimal hedging in discrete time is equivalent to a sequential linear regression problem.

[2] We remark that conditions for the perfect recovery of solutions of (18) by solving (19) can be given within the theoretical framework of compressive sensing, see, for instance, Foucart and Rauhut (2013).

[3] It should be said, in all fairness, that the `R`-function `lars` also provides the option "`stepwise`" instead of "`lasso`" which effectively corresponds to the greedy method.

REFERENCES

Alfonsi, A. (2015). *Affine diffusions and related processes: Simulation, theory and applications*. Cham, Switzerland: Springer.

Andersen, L. B. G., & Piterbarg, V. V. (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics*, *11*(1), 29–50.

Ansel, J.-P., & Stricker, C. (1993). Décomposition de Kunita-Watanabe. *Séminaire de probabilités de Strasbourg*, *27*, 30–32.

Beiglböck, M., Henry-Labordère, P., & Penkner, F. (2013). Model-independent bounds for option prices: A mass transport approach. *Finance and Stochastics*, *17*(3), 477–501.

Bühlmann, P., & Van De Geer, S. (2011). *Statistics for high-dimensional data: Methods, theory and applications*. Berlin, Heidelberg: Springer Science & Business Media.

Carr, P. (2011). Semi-static hedging of barrier options under Poisson jumps. *International Journal of Theoretical and Applied Finance*, *14*(07), 1091–1111.

Carr, P., & Madan, D. (2001). Towards a theory of volatility trading. In E. Jouini, J. Cvitanic and M. Musiela (Eds.), *Option pricing, interest rates and risk management*, *Handbooks Mathematical Finance* (pp. 458–476). Cambridge: Cambridge University Press.

Carr, P., & Sun, J. (2007). A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, *10*(2), 87–150.

Cont, R. (2010). Model calibration. In R. Cont (Ed.), *Encyclopedia of quantitative finance* (pp. 1210–1219). Wiley Online Library.

Cont, R., & Tankov, P. (2004). *Financial modelling with jump processes*. *Financial mathematics series*. Boca Raton, FL: Chapman & Hall/CRC.

Di Tella, P., Haubold, M., & Keller-Ressel, M. (2019). Variance-optimal semi-static hedging in stochastic volatility models with Fourier representation. *Journal of Applied Probability*, *56*(3), 787–809.

El Euch, O., & Rosenbaum, M. (2019). The characteristic function of rough Heston models. *Mathematical Finance*, *29*(1), 3–38.

Chicago Board Options Exchange (2019). *VIX White Paper*. http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-faqs

Filipović, D., & Mayerhofer, E. (2009). Affine diffusion processes: Theory and applications. *Radon Series on Computational & Applied Mathematics*, *8*, 1–40.

Föllmer, H., & Schweizer, M. (1988). Hedging by sequential regression: An introduction to the mathematics of option trading. *Astin Bulletin*, *18*(02), 147–160.

Föllmer, H., & Sondermann, D. (1986). Hedging of nonredundant contingent claims. In W. Hildenbrand and A. Mas-Colell (Eds.), *Contributions to mathematical economics* (pp. 205–223). North Holland, The Netherlands: Elsevier Science Publishing.

Foucart, S., & Rauhut, H. (2013). *A mathematical introduction to compressive sensing*. Basel, Switzerland: Birkhäuser.

Friz, P. K., & Keller-Ressel, M. (2010). Moment explosions. In R. Cont (Ed.), *Encyclopedia of quantitative finance* (pp. 1247–1253). Chichester, UK: Wiley.

Furnival, G. M., & Wilson, R. W. (1974). Regressions by leaps and bounds. *Technometrics*, *16*(4), 499–511.

Gatheral, J. (2006). *The volatility surface*. New Jersey: Hoboken.

Hastie, T., Tibshirani, R., & Friedman, J. (2013). *The elements of statistical learning* (2nd ed., Vol. 1). *Springer series in statistics*. Berlin: Springer.

Heston, S. (1993). A closed-form solution of options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*, 327–343.

Horn, R. A., & Johnson, C. R. (2012). *Matrix analysis*. Cambridge: Cambridge University Press.

Hubalek, F., Kallsen, J., & Krawczyk, L. (2006). Variance-optimal hedging for processes with stationary independent increments. *The Annals of Applied Probability*, *16*, 853–885.

Jacod, J., & Shiryaev, A. (2003). *Limit theorems for stochastic processes* (2nd ed., *Volume 288 of Grundlehren der mathematischen wissenschaften*. Berlin: Springer-Verlag.

Kallsen, J., & Pauwels, A. (2010). Variance-optimal hedging in general affine stochastic volatility models. *Advances in Applied Probability*, *42*(1), 83–105.

Kunita, H., & Watanabe, S. (1967). On square integrable martingales. *Nagoya Mathematical Journal*, *30*, 209–245.

Lewis, A. L. (2000). *Option valuation under stochastic volatility*. Newport Beach, CA: Finance Press.
Lukacs, E. (1960). *Characteristic functions*. London, UK: Charles Griffin & Co. Ltd.
Muirhead, R. J. (2009). *Aspects of multivariate statistical theory*, Vol. 197. Hoboken, NJ: John Wiley & Sons.
Neuberger, A. (1994). The log contract. *The Journal of Portfolio Management*, *20*(2), 74–80.
Neumaier, A. (1998). Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Review*, *40*(3), 636–666.
Pauwels, A. (2007). *Variance-optimal hedging in affine volatility models*. PhD thesis, TU Munich.
Pham, H. (2000). On quadratic hedging in continuous time. *Mathematical Methods of Operations Research*, *51*(2), 315–339.
Raible, S. (2000). *Lévy processes in finance: Theory, numerics and empirical facts*. PhD thesis, Albert-Ludwigs-Universität Freiburg.
Schweizer, M. (1984). *Varianten der Black-Scholes-Formel*. Master's thesis, ETH Zürich.
Schweizer, M. (2001). A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanic, & M. Musiela (Eds.), *Option pricing, interest rates and risk management* (pp. 538–574) Cambridge: Cambridge University Press.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*, 267–288.

How to cite this article: Di Tella P, Haubold M, Keller-Ressel M. Semistatic and sparse variance-optimal hedging. *Mathematical Finance*. 2020;30:403–425. https://doi.org/10.1111/mafi.12235

APPENDIX A: THE PROOF OF THEOREM 4.2

We show Theorem 4.2 with the help of two lemmas.

Lemma A.1. *The set $\mathcal{D}_T$ (from (30)) and the function $\psi_t(u,w)$ (from (31)) have the following properties:*

(a) The sets $\mathcal{D}_T$ are open and convex.
(b) If $(u,w) \in S(\mathcal{D}_T)$, then $(u, \psi_{T-t}(u,w)) \in S(\mathcal{D}_t)$.
(c) The functions $\phi_t(u,w)$ and $\psi_t(u,w)$ are analytic on $S(\mathcal{D}_t)$.
(d) If $(a,b) \in \mathcal{D}_T$, then $(a,b') \in \mathcal{D}_T$ for all $b' \leq b$.
(e) $\operatorname{Re}\psi_t(u,w) \leq \psi_t(\operatorname{Re}u, \operatorname{Re}w)$ for all $(u,w) \in S(\mathcal{D}_t)$

Proof. Properties (a), (b), and (c) are shown in Filipović and Mayerhofer (2009). For (d), note that $V_T \geq 0$ implies that

$$\mathbb{E}\left[\exp\left(aX_T + b'V_T\right)\right] \leq \mathbb{E}\left[\exp\left(aX_T + bV_T\right)\right]$$

for all $b' \leq b$. For (e), note that Jensen's inequality implies

$$\exp\left(\operatorname{Re}\phi_T(u,w) + \operatorname{Re}\psi_T(u,w)V_0 + \operatorname{Re}uX_0\right) = \left|\mathbb{E}\left[\exp\left(uX_T + wV_T\right)\right]\right| \leq$$
$$\leq \mathbb{E}[|\exp(uX_T + wV_T)|] = \exp\left(\phi_T(\operatorname{Re}u, \operatorname{Re}w) + \psi_T(\operatorname{Re}u, \operatorname{Re}w)V_0 + \operatorname{Re}uX_0\right).$$

As V_0 can be chosen arbitrarily large, (e) follows. □

Lemma A.2. *Let (X, V) be given by the Heston model (28) and assume that $(u, w) \in S(D_t)$. Then,*

$$\mathbb{E}\left[e^{uX_t+wV_t}V_t\right] = \left\{\partial_w\phi_t(u,w) + V_0\partial_w\phi_t(u,w)\right\}e^{X_0}h(u,t,V_0).$$

Proof. Fix $(a, b) \in D_t$ and consider $(u, w) \in S(D_t)$ of the form $(u = a + \mathrm{i}\, y, w = b + \mathrm{i}\, z)$. By assumption $K = \mathbb{E}[e^{aX_t+bV_t}]$ exists and is a number in $(0, \infty)$. Define a probability measure $\mathbb{M}$ on $(\Omega, \mathcal{F}_t)$ by $\frac{\mathrm{d}\mathbb{M}}{\mathrm{d}\mathbb{Q}}|_{\mathcal{F}_t} = \exp(aX_t + bV_t)/K$, i.e., by exponential tilting of $\mathbb{Q}$. Clearly, the characteristic function of (X_t, V_t) under $\mathbb{M}$ is given by

$$\mathbb{E}^{\mathbb{M}}\left[e^{iyX_t+izV_t}\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{uX_t+wV_t}\right] = \exp\left(\phi_t(u,w) + V_0\psi_t(t,u,w) + uX_0\right).$$

Due to the analyticity properties of $\phi_t(u, w)$ and $\psi_t(u, w)$, cf. Lemma A.1(c), all partial derivatives of the left-hand side with respect to (y, z) exist. Standard results on differentiability of characteristic functions (cf. Lukacs, 1960, section 2.3)) yield that

$$\mathbb{E}^{\mathbb{M}}\left[e^{iyX_t+izV_t}V_t\right] = -i\frac{\mathrm{d}}{\mathrm{d}z}\mathbb{E}^{\mathbb{M}}\left[e^{iyX_t+izV_t}\right] = \frac{\mathrm{d}}{\mathrm{d}w}\exp\left(\phi_t(u,w) + V_0\psi_t(u,w) + uX_0\right).$$

Transforming the left hand side back to $\mathbb{Q}$ yields the desired result. □

Proof of Theorem 4.2. First, we determine the relevant quantities of Proposition 4.1 in case of the Heston model. Using (29) and (28) we obtain

$$h(u, T-t, V_t) = \exp\left(\phi_{T-t}(u,0) + V_t\psi_{T-t}(u,0)\right), \quad \partial_v h(u, T-t, V_t) = \psi_{T-t}(u)h(u, T-t, V_t),$$

$$\partial_v\gamma(T-t, V_t) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right), \quad dQ = \sigma^2(1-\rho^2)V_t\, dt.$$

Therefore, by Proposition 4.1,

$$\mathrm{d}\mathcal{A} = \frac{\sigma^2(1-\rho^2)}{\lambda^2}\left(1 - e^{-\lambda(T-t)}\right)^2 V_t\, dt,$$

$$\mathrm{d}\mathcal{B}(u) = \frac{\sigma^2(1-\rho^2)}{\lambda}\left(1 - e^{-\lambda(T-t)}\right)\psi_{T-t}(u)H_t(u)V_t\, dt,$$

$$\mathrm{d}C(u_1, u_2) = \sigma^2(1-\rho^2)\psi_{T-t}(u_1)\psi_{T-t}(u_2)H_t(u_1)H_t(u_2)V_t\, dt.$$

If the expectations in (33) are finite, then an application of Theorem 4.1 yields the desired representations of A, B, C. Thus, it remains to show integrability and to evaluate the expectations in (33).

First, (33a) is easily obtained from the Heston SDE (28). To show (33b) we use Lemma A.1 and A.2. Let $u = x + iz$ be element of some strip $S(R_j)$ and note that the integrability condition on X_T implies $(x, 0) \in \mathcal{D}_T$. From Lemma A.1(b) we conclude that $(x, \psi_{T-t}(x, 0)) \in \mathcal{D}_t$. Now $\operatorname{Re}\psi_{T-t}(u, 0) \leq \psi_{T-t}(0, x)$, together with Lemma A.1(d) shows that also $(\operatorname{Re} u, \operatorname{Re}\psi_{T-t}(0, u)) \in \mathcal{D}_t$, which is equivalent to $(u, \psi_{T-t}(u, 0)) \in S(\mathcal{D}_t)$. Applying Lemma A.2 with $w = \psi_{T-t}(u, 0)$ yields (33b).

For (33c) we can use a similar argument: Write $u_1 = x_1 + iz_1$ and $u_2 = x_2 + iz_2$. The integrability condition on X_T implies that $(2x_1, 0)$ and $(2x_2, 0)$ are in $\mathcal{D}_T$. From Lemma A.1(b) we conclude that $(2x_1, \psi_{T-t}(2x_1, 0)) \in \mathcal{D}_t$, and similarly for x_2. Convexity of $\mathcal{D}_t$, see Lemma A.1(a), shows that $(x_1 + x_2, \frac{1}{2}q_{T-t}(2x_1, 2x_2)) \in \mathcal{D}_T$. Now convexity of ψ_{T-t} (and hence of q_{T-t}), together with Lemma A.1(d) yields that also $(x_1 + x_2, q_{T-t}(x_1, x_2)) \in \mathcal{D}_t$. To pass to complex arguments, note that Lemma A.1(e) implies that also $(u_1 + u_2, q_{T-t}(u_1, u_2)) \in S(\mathcal{D}_T)$. Hence, we may apply Lemma A.2 with $u = u_1 + u_2$ and $w = q_{T-t}(u_1, u_2)$, which yields (33c). □