RANDOM BASKET QUOTES: MINIMAX DISTRIBUTIONAL RECOVERY AND PRICE ESTIMATION

BY TONGSEOK LIM[1,a] AND YICHEN ZHANG[1,b]

[1]*Department of Quantitative Methods, Purdue University,* [a]*lim336@purdue.edu;* [b]*yichen@purdue.edu*

We study the recovery of a compactly supported multivariate probability law from random, possibly noisy basket-call prices. Exact quotes define a convex set of admissible laws, whose Wasserstein diameter equals the largest remaining price interval over Lipschitz claims. After a finite-dimensional moment correction, the L^2 pricing discrepancy is proportional to a homogeneous Sobolev norm of order $-(d+3)/2$. For densities in an H^α ball, we establish the sharp minimax expected W_1 risk $n^{-(\alpha+1)/d} + (\varsigma^2/n)^{(\alpha+1)/(2\alpha+2d+3)}$, uniformly over bounded noise levels, including exact quotes. A bounded systematic quote error of size ϵ adds the term $\epsilon^{(\alpha+1)/(\alpha+(d+3)/2)}$. The upper bound is attained by a probability-valued estimator combining local reconstruction of the pricing surface with residual-corrected Fourier moments. For a given payoff, we characterize minimax price estimation at fixed positive noise through an L^2 modulus of continuity. Its convex dual is a penalized static replication problem, which yields an affine estimator attaining the minimax rate. Finally, long-only basket quotes identify the joint law, yet the exchange option $(x_1 - x_2)_+$ has minimax risk of exact order $(\log n)^{-(\alpha+3/2)}$, uniformly down to zero noise. In contrast, polynomial claims admit the parametric rate at fixed positive noise.

1. Introduction. Many statistical problems observe an unknown probability measure μ only through random linear functionals. A design point Z_i is drawn, and the value $\int \phi_{Z_i} \,\mathrm{d}\mu$ of a known test function is recorded, exactly or with noise. Exact observations determine a convex set of measures consistent with the data, and noisy observations define a statistical inverse problem. The difficulty of recovering μ, or a single functional $\int g \,\mathrm{d}\mu$, depends on the geometry of the test functions, the design, the probability constraints on the unknown, and the loss. We study these effects for a family of test functions arising in multivariate pricing.

We consider basket calls at one fixed maturity. A direction $\boldsymbol{u} \in S^{d-1}$ and strike k define the payoff $(\boldsymbol{u}\cdot\boldsymbol{x} - k)_+$, whose price under a law $\mu \in \mathcal{P}(\mathbb{R}^d)$ is $C_\mu(\boldsymbol{u}, k) = \int (\boldsymbol{u}\cdot\boldsymbol{x} - k)_+ \,\mu(\mathrm{d}\boldsymbol{x})$. We observe the direction, strike, and price of each sampled basket. A finite collection of such prices imposes linear moment constraints on the pricing law, leaving an interval of admissible prices for an unquoted payoff. We treat recovery of the whole law and estimation of a single claim as two uses of this common information structure.

For distributional recovery, we measure error in W_1, the 1-Wasserstein distance. This metric controls all Lipschitz payoffs, and the W_1 diameter of a quote fiber equals the largest width of their admissible price intervals. The pricing operator has smoothing order $s = (d+3)/2$: after correcting for the first moments, its L^2 pricing discrepancy is proportional to a homogeneous negative Sobolev norm of order $-s$. We use this identity to establish sharp inverse moduli for arbitrary probability laws and Sobolev density classes. These results quantify how discrepancies between pricing surfaces translate into uncertainty about the underlying law.

Our first main result determines the minimax expected risk over a Sobolev density class $\mathcal{F}_\alpha(L) \subset H^\alpha(\mathbb{R}^d)$ supported strictly inside the observed strike range. Under uniform i.i.d.

MSC2020 subject classifications. Primary 62C20, 62G05; secondary 44A12, 91G20.

Keywords and phrases. Minimax estimation, Option pricing, Random design, Statistical inverse problems, Wasserstein distance.

quote locations and independent errors of standard deviation ς, Theorem 5.1 establishes the sharp minimax rate

$$\mathfrak{R}_n\big(\mathcal{F}_\alpha(L),\varsigma\big) \asymp n^{-\frac{\alpha+1}{d}} + \Big(\frac{\varsigma^2}{n}\Big)^{\frac{\alpha+1}{2\alpha+2d+3}}, \qquad 0 \le \varsigma \le \overline{\varsigma}.$$

The constants are uniform in n and the noise interval. The first term is the information limit imposed by finitely many exact random quotes; the second is the noise-dependent contribution at fixed positive noise. They have the same order at $\varsigma \asymp n^{-(2\alpha+d+3)/(2d)}$. The theory also accommodates bounded systematic quote errors: an error of size ϵ, such as a mid-quote bias, adds $\epsilon^{(\alpha+1)/(\alpha+(d+3)/2)}$ to the rate. Unlike independent noise, this contribution does not vanish as the number of quotes grows. Even fixing all one-dimensional marginals leaves these three terms unchanged (Corollary 5.9).

A probability-valued estimator combining local polynomial reconstruction of the pricing surface C_μ with residual correction of inverse Fourier features attains the rate uniformly over the noise level. For the lower bound, Theorem 4.6 establishes a general limitation of exogenous random linear information, while localized alternatives capture the additional effects of noise and systematic bias.

Our second main result concerns the price $\int g \,\mathrm{d}\mu$ of a given payoff g. For a convex class $\mathcal{C}$ of laws, let $\omega_g(\epsilon)$ be the largest price difference between laws in $\mathcal{C}$ whose pricing surfaces differ by at most ϵ in the L^2 norm of the design. Theorem 6.3 shows that, at fixed positive noise, the minimax risk is comparable to $\omega_g(\varsigma/\sqrt{n})$. Moreover, the modulus has the dual representation

$$\omega_g(\epsilon) = \inf_w \big[\operatorname{osc}_{\mathcal{C}}\big(g - \mathcal{A}^*(wQ)\big) + \epsilon \,\|w\|_{L^2(Q)}\big],$$

where $\mathcal{A}^*(wQ)$ is a static basket portfolio and $\operatorname{osc}_{\mathcal{C}}$ measures the remaining price uncertainty over $\mathcal{C}$. This duality yields a near-minimax affine estimator whose weights are determined by the replicating portfolio.

The distinction between recovering the law and estimating a particular price is especially striking under long-only directional coverage. Although these quotes identify the joint law, their statistical accuracy depends strongly on the payoff. Polynomial claims admit the parametric rate $n^{-1/2}$ at fixed positive noise, whereas Theorem 7.13 establishes minimax risk of exact order $(\log n)^{-(\alpha+3/2)}$ for the exchange option $(x_1 - x_2)_+$, uniformly over bounded noise levels, including exact quotes. Thus identification of the joint law does not imply a polynomial recovery rate for every claim. More generally, Theorem 7.8 shows that a kink $(\boldsymbol{e}\cdot\boldsymbol{x} - k)_+$ is Lipschitz stable when its direction is quoted, while an unquoted direction can lead to a sharp logarithmic pricing modulus.

1.1. *Related Literature.* **Random information and exact quotes.** The term $n^{-(\alpha+1)/d}$ is the error of recovering an H^α density in W_1 from n random linear functionals, a question of information-based complexity [21, 34]. For point evaluations of functions in the Sobolev space W_p^s, i.i.d. points attain the optimal rate of L_q-approximation when $q < p$ [28, 29]. Corollary 5.8 gives the corresponding result for basket quotes. The estimator attains the rate of optimal information from i.i.d. quotes without logarithmic loss, and it does not use the design density. The lower bound in Theorem 4.6 holds for any exogenous random information and any randomized estimator, with no linear or convex structure on the class or the loss. Our results connect exact random information and noisy statistical recovery through a single sharp minimax rate, including regimes in which the noise level decreases with the sample size. Bayes risks under a design-independent prior also give the rates of [36] for noisy random linear observations under elliptical constraints. For deterministic information, the argument

reduces to the Bernstein-width bounds of [34]. Noise-aware comparisons of optimal recovery and minimax estimation appear in [16].

Tomography. The pricing operator integrates the Radon transform twice in the strike variable, making basket quotes tomographic data with the basket direction specifying the projection angle. Statistical tomography studies density recovery from random projection data [9, 23, 27], with integrated or pointwise squared error. Related work treats unobserved projection angles [35] and inverse problems in density and white-noise models [25]. Statistical inverse learning studies regularized estimation under source conditions [5].

Three features distinguish our setting. The unknown is a probability law, and recovery is measured in W_1. The admissible laws form a convex set whose diameter equals the largest price uncertainty over Lipschitz claims. The pricing operator admits a negative-Sobolev identity that yields sharp inverse moduli (Theorems 3.1 and 3.2). Finally, random sampling of quote locations imposes an information limit even when prices are observed exactly. Wasserstein losses appear in direct density estimation through Besov duality [33, 39] and in deconvolution [8, 14]. Unlike these settings, the basket transform is neither direct nor a convolution.

Single prices and static replication. For a linear functional over a convex class, the minimax risk is determined by a modulus of continuity and attained by an affine rule. This is the optimal-recovery theory of [30] in the statistical form given by [17], with further work on convergence rates, nonconvex classes, indirect observations and convex programming in [18], [7], [20] and [24]. Theorem 6.3 establishes this modulus characterization for random-design regression with arbitrary bounded continuous test functions and convex weakly closed classes of laws. Under basket quotes the affine rule is a static portfolio of quoted calls and the modulus is a penalized replication problem. Spanning of multi-asset payoffs by basket calls is treated in [6], and the one-dimensional recovery problem from single-asset options in [1, 2]. Static-arbitrage bounds for basket options are obtained by convex optimization in [4] and by super- and subreplicating portfolios in [13, 22, 32, 38]. The penalized replication problem is the statistical form of this duality.

Missing directions. Derivative families that determine the joint law are studied in [31]. Long-only quotes identify the law by the homogeneous-polynomial criterion of [12], a refinement of the Cramér–Wold theorem [3, 11]. Weak convergence from positive-measure sets of projections is studied in [10] under moment assumptions, and compact support permits the direct argument in Theorem 2.2. Section 7 quantifies the instability of this identification. The missing cone around the spread direction corresponds to the missing wedge of limited-angle tomography [19, 26, 37]. The same long-only quotes determine polynomial claims at the rate $n^{-1/2}$ at every positive noise level, and a kink on a quoted direction is Lipschitz stable. Section 7 establishes the exact order of this instability for the exchange option: $(\log(1/\epsilon))^{-(\alpha+3/2)}$ for the modulus and $(\log n)^{-(\alpha+3/2)}$ for the minimax risk.

Organization. Section 2 specifies the observation experiments and establishes identification and consistency. Section 3 develops the operator estimates, and Section 4 treats exact sampling and information lower bounds. Section 5 proves the minimax rate uniformly over bounded noise levels. Section 6 relates admissible sets to price intervals and characterizes minimax estimation of a single price, and Section 7 analyzes directional information and the exchange option. Section 8 illustrates the rates numerically. Section 9 concludes and discusses open problems.

2. Observation Model and Identification.

We specify the probability classes and observation experiments, then characterize directional identification and the almost-sure limit of the admissible sets.

2.1. *Quotes and Admissible Laws.* We absorb discounting into the quoted prices and use a single terminal pricing law. Fix $d \geq 2$ and $R > 0$. Write $B_R = \{\boldsymbol{x} \in \mathbb{R}^d : |\boldsymbol{x}| \leq R\}$ and $\mathcal{Z} = S^{d-1} \times [-R, R]$. For $\mu \in \mathcal{P}(B_R)$, define

$$C_\mu(\boldsymbol{u}, k) = \int_{B_R} (\boldsymbol{u} \cdot \boldsymbol{x} - k)_+ \, \mu(\mathrm{d}\boldsymbol{x}). \tag{2.1}$$

We observe n independent quotes (Z_i, V_i), where

$$Z_i = (U_i, K_i) \overset{\text{i.i.d.}}{\sim} Q \in \mathcal{P}(\mathcal{Z}), \qquad V_i = C_\mu(Z_i), \qquad i = 1, \ldots, n. \tag{2.2}$$

We use the product metric $\rho((\boldsymbol{u}, k), (\boldsymbol{v}, l)) = |\boldsymbol{u} - \boldsymbol{v}| + \frac{|k-l|}{R}$. For every $\mu, \nu \in \mathcal{P}(B_R)$,

$$0 \leq C_\mu \leq 2R, \qquad \mathrm{Lip}_\rho(C_\mu) \leq R, \qquad \|C_\mu - C_\nu\|_\infty \leq W_1(\mu, \nu). \tag{2.3}$$

For a deterministic design $z = (z_1, \ldots, z_n)$ and a feasible price vector $v = (v_1, \ldots, v_n)$, define

$$\mathcal{K}_n(z, v) = \{\nu \in \mathcal{P}(B_R) : C_\nu(z_i) = v_i, \ 1 \leq i \leq n\}, \tag{2.4}$$

$$\mathcal{K}_n(\mu) = \{\nu \in \mathcal{P}(B_R) : C_\nu(Z_i) = C_\mu(Z_i), \ 1 \leq i \leq n\}. \tag{2.5}$$

We call $\mathcal{K}_n(z, v)$ the quote fiber of the price vector v, and $\mathcal{K}_n(\mu)$ the random quote fiber of μ. Both sets are convex and compact in the W_1 topology whenever they are nonempty.

2.2. *Regularity Classes and Statistical Risk.* Let σ be normalized surface measure on S^{d-1}, write $\omega_{d-1} = \left|S^{d-1}\right|$, and define

$$Q_0(\mathrm{d}\boldsymbol{u}\, \mathrm{d}k) = \sigma(\mathrm{d}\boldsymbol{u}) \frac{\mathrm{d}k}{2R}, \qquad s = \frac{d+3}{2}. \tag{2.6}$$

For a signed measure η supported on B_R, let $\mathcal{A}\eta = C_\eta$ denote the linear extension of (2.1). We use the Fourier convention

$$\widehat{\eta}(\boldsymbol{\xi}) = \int e^{-\mathrm{i}\boldsymbol{\xi} \cdot \boldsymbol{x}} \, \eta(\mathrm{d}\boldsymbol{x}), \qquad \|\eta\|^2_{\dot{H}^{-s}} = (2\pi)^{-d} \int_{\mathbb{R}^d} |\widehat{\eta}(\boldsymbol{\xi})|^2 \, |\boldsymbol{\xi}|^{-2s} \, \mathrm{d}\boldsymbol{\xi}.$$

We also use the inhomogeneous Sobolev norm $\|f\|^2_{H^\beta(\mathbb{R}^d)} = (2\pi)^{-d} \int_{\mathbb{R}^d} (1 + |\boldsymbol{\xi}|^2)^\beta \left|\widehat{f}(\boldsymbol{\xi})\right|^2 \mathrm{d}\boldsymbol{\xi}$, $\beta \in \mathbb{R}$. The observation model is

$$Z_i \overset{\text{i.i.d.}}{\sim} Q_0, \qquad Y_i = C_\mu(Z_i) + \varsigma \xi_i, \qquad 0 \leq \varsigma \leq \bar{\varsigma} < \infty, \tag{2.7}$$

with errors ξ_i that are i.i.d., independent of the design, with $\mathbb{E}\xi_1 = 0$ and $\mathbb{E}\xi_1^2 = 1$. Positive-noise lower bounds assume an absolutely continuous error density p_ξ with finite Fisher information $I_\xi = \int (p_\xi')^2 / p_\xi$. Gaussian errors have $I_\xi = 1$. For the recovery results, fix $0 < r < R$ and assume that the unknown law μ is supported in B_r. The estimator may use the noise level ς. The gap between r and R permits bounded extension in the strike coordinate. We define

$$\mathcal{F}_\alpha(L) = \left\{ f \geq 0 : \int f \, \mathrm{d}\boldsymbol{x} = 1, \quad \mathrm{supp}\, f \subseteq B_r, \quad \|f\|_{H^\alpha(\mathbb{R}^d)} \leq L \right\}, \quad \alpha \geq 0. \tag{2.8}$$

The zero extension of f outside B_r belongs to $H^\alpha(\mathbb{R}^d)$. For a class $\mathcal{C} \subseteq \mathcal{P}(B_r)$, define

$$\mathfrak{R}_n(\mathcal{C}, \varsigma) = \inf_{\widehat{\mu}} \sup_{\mu \in \mathcal{C}} \mathbb{E}_{\mu, \varsigma} W_1(\widehat{\mu}, \mu), \tag{2.9}$$

where the infimum is over measurable estimators based on the observations $(Z_i, Y_i)_{i=1}^n$ and taking values in $\mathcal{P}(B_r)$. The constants in rate comparisons may depend on the displayed class parameters, on I_ξ, and on fixed coordinate charts on S^{d-1}. Dependence on a sampling lower bound or a fixed positive noise level is stated where it enters.

2.3. *Directional Identification and Set Consistency.*

DEFINITION 2.1. The design distribution Q identifies $\mathcal{P}(B_R)$ if equality of C_μ and C_ν on $\operatorname{supp} Q$ implies $\mu = \nu$ for all $\mu, \nu \in \mathcal{P}(B_R)$.

Full support on $\mathcal{Z}$ is sufficient, since equality of call functions for all strikes determines the projected law in each direction. To include possible atoms at the support endpoints, extend the call function by zero for $k \geq R$ and by $\int \boldsymbol{u} \cdot \boldsymbol{x}\, \mu(\mathrm{d}\boldsymbol{x}) - k$ for $k \leq -R$. Its mean is recovered from the value at $-R$. The second distributional derivative of this extension is the projected measure. Equality in all directions then determines the characteristic function on $\mathbb{R}^d$.

Let $E \subseteq S^{d-1}$ be the closed support of the direction distribution, and assume the strike distribution has full support on $[-R, R]$. For a product design, equality of pricing surfaces almost everywhere is equivalent, by continuity, to equality on $E \times [-R, R]$.

THEOREM 2.2 (Directional identification). *The following statements are equivalent.*

(i) *The pricing surfaces on $E \times [-R, R]$ determine every law in $\mathcal{P}(B_R)$ uniquely.*
(ii) *No nonzero homogeneous real polynomial on $\mathbb{R}^d$ vanishes on every direction in E.*

If (ii) fails, there are distinct smooth probability densities supported strictly inside B_R with identical call prices for every direction in E and every strike. Every nonempty relatively open set of directions satisfies (ii), a finite set of directions fails (ii) when $d \geq 2$, and in dimension two every infinite set of directions satisfies (ii).

THEOREM 2.3 (Set consistency). *For every $\mu \in \mathcal{P}(B_R)$, almost surely the decreasing random compact sets $\mathcal{K}_n(\mu)$ converge in Hausdorff distance induced by W_1 to*

$$\mathcal{K}_\infty(\mu; Q) = \{\nu \in \mathcal{P}(B_R) : C_\nu = C_\mu \text{ on } \operatorname{supp} Q\}. \tag{2.10}$$

If Q identifies $\mathcal{P}(B_R)$, then $\sup_{\nu \in \mathcal{K}_n(\mu)} W_1(\nu, \mu) \longrightarrow 0$ almost surely. Hence, every feasible selection is strongly consistent.

3. Geometry of the Pricing Operator. How close must two probability laws be if their basket-call prices are close? We address this question through the pricing operator $\mathcal{A}$ defined in Section 2, for which $\mathcal{A}(\mu - \nu) = C_\mu - C_\nu$. The results below relate pricing discrepancies to Sobolev norms and Wasserstein distance, and provide representations of smooth functionals used in the estimation procedures. We first consider differences of laws with the same mean. Such differences have zero total mass and zero first moments, and the following theorem expresses their $L^2(Q_0)$ pricing discrepancy exactly in Fourier and spatial form.

THEOREM 3.1 (Exact operator identity). *Let η be a finite signed measure supported on B_R with $\eta(B_R) = 0$ and $\int \boldsymbol{x}\, \eta(\mathrm{d}\boldsymbol{x}) = 0$. Then*

$$\|\mathcal{A}\eta\|^2_{L^2(Q_0)} = \frac{1}{2\pi R \omega_{d-1}} \int_{\mathbb{R}^d} |\widehat{\eta}(\boldsymbol{\xi})|^2 \, |\boldsymbol{\xi}|^{-d-3} \, \mathrm{d}\boldsymbol{\xi} = \frac{(2\pi)^{d-1}}{R\omega_{d-1}} \, \|\eta\|^2_{\dot{H}^{-s}}. \tag{3.1}$$

There is also the spatial identity

$$\|\mathcal{A}\eta\|^2_{L^2(Q_0)} = \frac{m_{3,d}}{24R} \iint |\boldsymbol{x} - \boldsymbol{y}|^3 \, \eta(\mathrm{d}\boldsymbol{x}) \eta(\mathrm{d}\boldsymbol{y}), \quad m_{3,d} = \frac{\Gamma(d/2)}{\sqrt{\pi}\, \Gamma((d+3)/2)}. \tag{3.2}$$

The equal-mean restriction is not needed for the stability bounds below: Lemma B.1 controls the difference of means by the pricing discrepancy, allowing a smooth correction before applying Theorem 3.1.

THEOREM 3.2 (Inverse moduli). *There is $C = C(d, R) > 0$ so that for all $\mu, \nu \in \mathcal{P}(B_R)$,*

$$W_1(\mu, \nu) \le C \, \|C_\mu - C_\nu\|_{L^2(Q_0)}^{2/(d+3)} . \tag{3.3}$$

If $\mu = f \, \mathrm{d}\boldsymbol{x}$, $\nu = g \, \mathrm{d}\boldsymbol{x}$, and $\|f\|_{H^\alpha}, \|g\|_{H^\alpha} \le L$ for $\alpha \ge 0$, then, with $D = \|C_\mu - C_\nu\|_{L^2(Q_0)}$,

$$W_1(\mu, \nu) \le C_{d,R,\alpha} \big[(1 + L)^{(s-1)/(\alpha+s)} D^{(\alpha+1)/(\alpha+s)} + D \big]. \tag{3.4}$$

The exponent in (3.3) is optimal. For every $\alpha \ge 0$ and every sufficiently large L, the exponent $(\alpha + 1)/(\alpha + s)$ in (3.4) is also optimal over the indicated density class.

Recovery from sampled quotes also requires Sobolev bounds on the pricing surface. We use a fixed Hilbert norm on the standard L^2 Sobolev space $H^t(\mathcal{Z})$.

LEMMA 3.3 (Forward and inverse Sobolev bounds). *Let $\eta \in H^\alpha(\mathbb{R}^d)$ be supported on B_r and satisfy $\int \eta \, \mathrm{d}\boldsymbol{x} = 0$. Then*

$$\|\mathcal{A}\eta\|_{H^{\alpha+s}(\mathcal{Z})} \le C_{d,r,R,\alpha} \, \|\eta\|_{H^\alpha(\mathbb{R}^d)} . \tag{3.5}$$

For any two probability densities f, g supported on B_r for which the right side is finite,

$$W_1(f \, \mathrm{d}\boldsymbol{x}, g \, \mathrm{d}\boldsymbol{x}) \le C_{d,r,R} \, \|\mathcal{A}(f - g)\|_{H^{s-1}(\mathcal{Z})} . \tag{3.6}$$

Finally, smooth linear functionals can be written as weighted integrals of the pricing surface, plus a constant. The $L^2(Q_0)$ norm of the weight function bounds the estimator variance.

LEMMA 3.4 (Inverse features). *For every real $\psi \in C_c^\infty(\mathbb{R}^d)$ there are a constant a_ψ and a function $\widetilde{G}_\psi \in L^2(Q_0)$ such that, for all $\mu \in \mathcal{P}(B_r)$,*

$$\int \psi \, \mathrm{d}\mu = a_\psi + \int_{\mathcal{Z}} \widetilde{G}_\psi(\boldsymbol{u}, k) C_\mu(\boldsymbol{u}, k) \, \mathrm{d}Q_0, \qquad |a_\psi| + \|\widetilde{G}_\psi\|_{L^2(Q_0)} \le C \, \|\psi\|_{H^s(\mathbb{R}^d)} . \tag{3.7}$$

The construction extends to complex ψ by linearity.

4. Exact Recovery and Random Information. How much uncertainty remains about the law after finitely many exact basket quotes have been observed? We first bound the Wasserstein diameter of the quote fiber in terms of the coverage of the observed locations, without imposing smoothness on the law. We then obtain sharper bounds on Sobolev density classes and construct an estimator that does not require the class radius. Finally, we establish lower bounds showing the limits of recovery from finitely many exact linear observations. Extensions of the coverage bounds to stopping times and adaptive quote arrivals are given in Appendix C. Throughout this paper, we set $\operatorname{diam}_{W_1} \varnothing = 0$.

4.1. *Recovery Without Smoothness Assumptions.* For $z = (z_i)_{i=1}^n$, define its fill distance

$$h(z) = \sup_{w \in \mathcal{Z}} \min_{1 \le i \le n} \rho(w, z_i), \qquad h_n = h(Z_1, \dots, Z_n). \tag{4.1}$$

$h(z)$ is the largest distance from a possible quote location to the nearest observed location. As $\mathcal{Z} = S^{d-1} \times [-R, R]$ has dimension d, for $0 < t < 1$, it admits a t-net of cardinality at most $C_d t^{-d}$, and every ρ-ball of radius t centered in $\mathcal{Z}$ has Q_0-probability at least $c_d t^d$.

If two laws agree at the observed locations, the Lipschitz bound in (2.3) controls the difference of their prices elsewhere by the fill distance. The inverse modulus in Theorem 3.2 then bounds their Wasserstein distance.

THEOREM 4.1 (Distribution-free exact recovery bound). *For any deterministic design z,*

$$\sup_{v\in\mathbb{R}^n} \operatorname{diam}_{W_1} \mathcal{K}_n(z,v) \le C_{d,R} h(z)^{2/(d+3)}. \tag{4.2}$$

If $Q \ge q_* Q_0$ for some $q_* > 0$, then with probability at least $1-\delta$,

$$\sup_{v\in\mathbb{R}^n} \operatorname{diam}_{W_1} \mathcal{K}_n(Z,v) \le C_{d,R,q_*} \Big(\frac{\log(en/\delta)}{n}\Big)^{\frac{2}{d(d+3)}}. \tag{4.3}$$

If instead $|C_\mu(z_i) - v_i| \le \epsilon$ and $|C_\nu(z_i) - v_i| \le \epsilon$ for every i, the same argument gives

$$W_1(\mu,\nu) \le C_{d,R}(\epsilon + h(z))^{2/(d+3)}.$$

Proposition 4.4 below gives the corresponding bound on Sobolev classes.

The quantitative coverage assumption $Q \ge q_* Q_0$ cannot be replaced by full support alone if a rate uniform over design distributions is desired. For an arbitrary full-support Q, the coverage inequality (C.1) in Appendix C holds with cq_*t^d replaced by $\beta_Q(t) = \inf_{w\in\mathcal{Z}} Q(B_\rho(w,t))$. This quantity is positive for each $t > 0$, by compactness and a finite covering by smaller balls, but it need not have a uniform power-law lower bound. The following example shows the resulting limitation.

PROPOSITION 4.2 (No rate uniform over full-support designs). *Fix $z_0 \in \mathcal{Z}$ and $0 < \epsilon < 1$, and let $Q_\epsilon = (1-\epsilon)\delta_{z_0} + \epsilon Q_0$, which has full support. For exact quotes at locations Z_i i.i.d. Q_ϵ and estimators $\widehat{\mu}$ based on $(Z_i, V_i)_{i=1}^n$ with values in $\mathcal{P}(B_r)$,*

$$\inf_{\widehat{\mu}} \sup_{\mu\in\mathcal{P}(B_r)} \mathbb{E}_{\mu,Q_\epsilon} W_1(\widehat{\mu},\mu) \ge r(1-\epsilon)^n.$$

Hence the supremum of this minimax risk over full-support designs is at least r for every n.

4.2. *Recovery on Sobolev Classes.* Sobolev regularity yields a sharper bound on the uncertainty remaining after exact quotes. The difference of two feasible pricing surfaces vanishes at every observed location. A Sobolev sampling inequality, combined with the forward and inverse bounds of Lemma 3.3, converts these equalities into a bound of order $h(z)^{\alpha+1}$ on the Wasserstein distance. For $\mathcal{F}_\alpha(L)$ in (2.8), let

$$\mathcal{K}_{n,\alpha,L}(z,v) = \{ f\,\mathrm{d}\boldsymbol{x} : f \in \mathcal{F}_\alpha(L),\ C_f(z_i) = v_i,\ 1 \le i \le n \}.$$

THEOREM 4.3 (Exact recovery on Sobolev classes). *There are constants $C, h_* > 0$, depending only on d, r, R, α, such that every design with $h(z) \le h_*$ satisfies*

$$\sup_{v\in\mathbb{R}^n} \operatorname{diam}_{W_1} \mathcal{K}_{n,\alpha,L}(z,v) \le CLh(z)^{\alpha+1}. \tag{4.4}$$

If $Q \ge q_* Q_0$ for some $q_* > 0$, then with probability at least $1-\delta$,

$$\sup_{v\in\mathbb{R}^n} \operatorname{diam}_{W_1} \mathcal{K}_{n,\alpha,L}(Z,v) \le C_{q_*} L \Big(\frac{\log(en/\delta)}{n}\Big)^{\frac{\alpha+1}{d}}. \tag{4.5}$$

We also allow each quoted price to be known only up to a bounded error. For $\epsilon \ge 0$, let

$$\mathcal{K}^\epsilon_{n,\alpha,L}(z,v) = \{ f\,\mathrm{d}\boldsymbol{x} : f \in \mathcal{F}_\alpha(L),\ |C_f(z_i) - v_i| \le \epsilon,\ 1 \le i \le n \}.$$

PROPOSITION 4.4 (Bounded quote errors). *Let $\vartheta = (\alpha+1)/(\alpha+s)$. There are $C, h_* > 0$, depending only on d, r, R, α, such that every design with $h(z) \le h_*$ satisfies*

$$\sup_{v\in\mathbb{R}^n} \operatorname{diam}_{W_1} \mathcal{K}^\epsilon_{n,\alpha,L}(z,v) \le C\left[Lh(z)^{\alpha+1} + L^{1-\vartheta}\epsilon^\vartheta\right]. \tag{4.6}$$

If $Q \ge q_* Q_0$ for some $q_* > 0$, then with probability at least $1-\delta$, simultaneously for all v and ϵ, the right side can be replaced by $C_{q_*}\big[L\big(\frac{\log(en/\delta)}{n}\big)^{(\alpha+1)/d} + L^{1-\vartheta}\epsilon^\vartheta\big]$.

The two terms in the upper bound represent uncertainty due to finite sampling and bounded quote errors, respectively. They are of the same order when $\epsilon \asymp Lh(z)^{\alpha+s}$; above this scale, the quote-error term is larger. Section 5 studies the corresponding minimax expected risk when bounded systematic errors are combined with random quote noise.

For a specified smoothness index α, define the minimum-norm interpolant

$$\widehat{f}_n = \arg\min_f \left\{ \|f\|_{H^\alpha}^2 : f \geq 0, \ \operatorname{supp} f \subseteq B_r, \ \int f \,\mathrm{d}\boldsymbol{x} = 1, \ C_f(Z_i) = V_i \ (1 \leq i \leq n) \right\}.$$

If the exact quotes are generated by a density $f_0 \in \mathcal{F}_\alpha(L)$, then minimality gives $\|\widehat{f}_n\|_{H^\alpha} \leq \|f_0\|_{H^\alpha} \leq L$. Hence $\widehat{f}_n \in \mathcal{F}_\alpha(L)$, although the optimization requires the smoothness index α but not the radius L. Existence, uniqueness, and measurability are shown in Appendix C. On the event in (4.5), $\sup_{f_0 \in \mathcal{F}_\alpha(L)} W_1(\widehat{f}_n \,\mathrm{d}\boldsymbol{x}, f_0 \,\mathrm{d}\boldsymbol{x}) \leq CL\big(\frac{\log(en/\delta)}{n}\big)^{(\alpha+1)/d}$, where, for each f_0, the estimator $\widehat{f}_n$ is computed from the exact quotes generated by f_0.

For i.i.d. locations with $Q \geq q_* Q_0$, the proof of Theorem C.6 gives the worst-case expected error bound $\sup_{f_0 \in \mathcal{F}_\alpha(L)} \mathbb{E}_{f_0} W_1(\widehat{f}_n \,\mathrm{d}\boldsymbol{x}, f_0 \,\mathrm{d}\boldsymbol{x}) \leq CL\big(\frac{\log(en)}{n}\big)^{(\alpha+1)/d}$. Corollary 4.7 below shows that this estimator is minimax optimal up to a factor $(\log(en))^{(\alpha+1)/d}$.

4.3. *Lower Bounds for Exact Linear Information.* We now ask whether a different choice of quotes or estimator could improve the recovery rate. The next theorem shows that, for any fixed collection of n continuous payoffs, two densities in $\mathcal{F}_\alpha(L)$ have identical prices but remain separated in W_1 by at least $cn^{-(\alpha+1)/d}$. Thus the polynomial exponent in (4.5) cannot be improved for uniform recovery from a fixed design.

THEOREM 4.5 (Deterministic uniform lower bound). *For every $\alpha \geq 0$, there is $L_0 < \infty$ such that for each $L > L_0$ there is $c > 0$ with the following property. For every $n \geq 1$ and every collection of continuous payoffs $\phi_1, \ldots, \phi_n$ on B_r, there are $f^+, f^- \in \mathcal{F}_\alpha(L)$ satisfying*

$$\int \phi_i f^+ \,\mathrm{d}\boldsymbol{x} = \int \phi_i f^- \,\mathrm{d}\boldsymbol{x} \quad (1 \leq i \leq n), \qquad W_1(f^+ \,\mathrm{d}\boldsymbol{x}, f^- \,\mathrm{d}\boldsymbol{x}) \geq cn^{-(\alpha+1)/d}. \tag{4.7}$$

Hence for any deterministic quote design, the supremum of fiber diameters is at least $cn^{\frac{-\alpha-1}{d}}$.

The densities f^+, f^- in Theorem 4.5 may depend on the measurements, whereas minimax expected risk fixes the density before averaging over the design. To obtain a lower bound for random measurements, we use a design-independent prior whose average risk bounds the worst-case risk below. Unlike (4.5), which holds for all feasible price vectors on a common design event, this expected-risk criterion requires no logarithmic loss for exact quotes; see Theorem 5.1. The next theorem supplies the corresponding information lower bound.

THEOREM 4.6 (Lower bound for exact linear observations). *Let d be a pseudometric on a set $\mathcal{X}$, and let $\{f_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in [-1,1]^M\} \subseteq \mathcal{X}$ be continuous in $\boldsymbol{\theta}$ for d and satisfy*

$$\mathsf{d}(f_{\boldsymbol{\theta}}, f_{\boldsymbol{\theta}'}) \geq \delta \left\|\boldsymbol{\theta} - \boldsymbol{\theta}'\right\|_1, \qquad \boldsymbol{\theta}, \boldsymbol{\theta}' \in [-1,1]^M,$$

for some $\delta > 0$. Let the data consist of n observations $V_i = \mathsf{L}_i(f) + a_i$, where the linear functionals $\mathsf{L}_1, \ldots, \mathsf{L}_n$ and the shift $a \in \mathbb{R}^n$ have a joint distribution that does not depend on f, and $\boldsymbol{\theta} \mapsto \mathsf{L}_i(f_{\boldsymbol{\theta}})$ is affine for every realization. Then every measurable estimator $\widehat{f} \in \mathcal{X}$ based on $(V_i, \mathsf{L}_i, a_i)_{i=1}^n$ and independent randomization satisfies

$$\sup_{\boldsymbol{\theta} \in [-1,1]^M} \mathbb{E}_{f_{\boldsymbol{\theta}}} \mathsf{d}(\widehat{f}, f_{\boldsymbol{\theta}}) \geq \frac{\delta(M-n)}{4\pi^2}. \tag{4.8}$$

The proof uses a fixed product prior to convert coefficient uncertainty left by a rank $\le n$ observation matrix into expected loss. This Bakhvalov-type argument is related to Bernstein-width bounds [34] and Bayesian bounds for random linear recovery [36]. It allows random measurements and randomized estimators without requiring $\mathcal{X}$ to be linear or convex. Theorem 4.6, applied to the family in Lemma C.4, gives a lower bound for random basket quotes.

For a design distribution Q_n on $\mathcal{Z}^n$ independent of the unknown law, let

$$\mathfrak{R}_n^{\mathrm{ex}}(\mathsf{Q}_n,\alpha,L)=\inf_{\widehat{\mu}}\sup_{f\in\mathcal{F}_\alpha(L)}\mathbb{E}_{f,\mathsf{Q}_n}W_1(\widehat{\mu},f\,\mathrm{d}\boldsymbol{x}), \tag{4.9}$$

where the observations are $Z\sim\mathsf{Q}_n$ and $V_i=C_f(Z_i)$, and estimators take values in $\mathcal{P}(B_r)$. The locations Z_i need not be independent.

COROLLARY 4.7 (Exact random-design lower bound). *Fix $\alpha\ge 0$ and $L>L_0$ as in Theorem 4.5. There is $c>0$, independent of Q_n, such that*

$$\mathfrak{R}_n^{\mathrm{ex}}(\mathsf{Q}_n,\alpha,L)\ge cn^{-(\alpha+1)/d}. \tag{4.10}$$

The bound remains valid for any n exogenously chosen continuous linear payoffs.

Thus the rate $n^{-(\alpha+1)/d}$ is a lower limit for worst-case expected recovery error even when the observations are exact and the measurement distribution can be chosen freely, provided that it is independent of the unknown law. Section 5 shows that i.i.d. basket quotes with adequate coverage attain this rate without a logarithmic loss.

5. Minimax Recovery Across Noise Levels. How does recovery improve as quotes become more accurate, and what limits remain when their number increases? We distinguish independent quote noise from a bounded systematic bias. The former can be reduced by averaging, whereas the latter need not disappear even with arbitrarily many observations.

Specifically, let $e:\mathcal{Z}\to\mathbb{R}$ be a Borel function with $\sup_{z\in\mathcal{Z}}|e(z)|\le\epsilon$, and observe

$$Y_i=C_\mu(Z_i)+e(Z_i)+\varsigma\xi_i,\qquad i=1,\dots,n, \tag{5.1}$$

where Z_i are i.i.d. with law Q_0 and the errors ξ_i satisfy the assumptions of (2.7). The function e specifies the systematic bias at each direction–strike pair. It is fixed before sampling and may depend on the unknown law, but not on the realized locations or noise. No smoothness or mean-zero condition is imposed on e. For $\mathcal{C}\subseteq\mathcal{P}(B_r)$, define

$$\mathfrak{R}_n(\mathcal{C},\varsigma,\epsilon)=\inf_{\widehat{\mu}}\sup_{\mu\in\mathcal{C}}\sup_{|e|\le\epsilon}\mathbb{E}_{\mu,e,\varsigma}W_1(\widehat{\mu},\mu),$$

where estimators take values in $\mathcal{P}(B_r)$. As before, we identify a density class with its associated probability laws. In particular, $\mathfrak{R}_n(\mathcal{C},\varsigma,0)=\mathfrak{R}_n(\mathcal{C},\varsigma)$. Throughout this section, write

$$t=\alpha+s,\qquad D_\alpha=2t+d=2\alpha+2d+3,\qquad \vartheta=(\alpha+1)/(\alpha+s).$$

THEOREM 5.1 (Sharp rate uniformly in the noise level). *Fix $\alpha\ge 0$, $L>L_0$ as in Theorem 4.5, and finite nonnegative bounds $\overline{\varsigma}$ and $\overline{\epsilon}$. There are constants $0<c<C<\infty$, independent of n,ς,ϵ, such that, for every $n\ge 1$, $0\le\varsigma\le\overline{\varsigma}$, and $0\le\epsilon\le\overline{\epsilon}$,*

$$c\Big[n^{-\frac{\alpha+1}{d}}+\Big(\frac{\varsigma^2}{n}\Big)^{\frac{\alpha+1}{D_\alpha}}+\epsilon^{\vartheta}\Big]\le\mathfrak{R}_n(\mathcal{F}_\alpha(L),\varsigma,\epsilon)\le C\Big[n^{-\frac{\alpha+1}{d}}+\Big(\frac{\varsigma^2}{n}\Big)^{\frac{\alpha+1}{D_\alpha}}+\epsilon^{\vartheta}\Big]. \tag{5.2}$$

The upper bound holds for independent, centered, unit-variance errors, without a density or tail assumption. The constants depend on $d,r,R,\alpha,L,\overline{\varsigma},\overline{\epsilon}$, and, for the lower bound, on I_ξ.

The three terms describe distinct limitations. Even exact quotes leave uncertainty of order $n^{-(\alpha+1)/d}$ because only finitely many linear measurements are observed. Independent noise adds the second term, which decreases with both the sample size and the noise level. The last term is the worst-case uncertainty caused by systematic bias. Since the sum is comparable to its largest term, the largest of these three contributions determines the recovery rate.

COROLLARY 5.2 (Transition scales). *Under the assumptions of Theorem 5.1, the exact-information term and the independent-noise term have the same order when* $\varsigma \asymp n^{-\frac{2\alpha+d+3}{2d}}$. *The exact-information term and the systematic-bias term have the same order when* $\epsilon \asymp n^{-\frac{2\alpha+d+3}{2d}}$. *If* $\epsilon = 0$ *and* $\varsigma_n = n^{-\gamma}$ *with* $\gamma \geq 0$, *then the minimax risk has order* $n^{-\beta}$, *where* $\beta = \min\left\{\frac{\alpha+1}{d}, \frac{(2\gamma+1)(\alpha+1)}{2\alpha+2d+3}\right\}$. *For example, when* $d = 2$ *and* $\alpha = 1$, *the risk is comparable to* $n^{-1} + (\varsigma^2/n)^{2/9} + \epsilon^{4/7}$, *and both transition scales above are* $n^{-7/4}$.

REMARK 5.3 (An irreducible error from systematic bias). For a fixed $\epsilon > 0$, increasing the number of quotes cannot make the worst-case recovery error tend to zero. Two different laws can have pricing surfaces within 2ϵ of one another, so that admissible biases make their observed mean prices identical. The term ϵ^{ϑ} quantifies the resulting ambiguity on $\mathcal{F}_\alpha(L)$.

5.1. *Direct Fourier Reconstruction.* We first describe a direct estimator for unbiased quotes, that is, $e = 0$. This provides a baseline and introduces the Fourier reconstruction used below. Lemma 3.4 expresses each smooth moment as an integral of the pricing surface. Replacing this integral by a sample average gives, for $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$\widehat{\theta}_\psi = a_\psi + \frac{1}{n}\sum_{i=1}^{n} \widetilde{G}_\psi(Z_i) Y_i. \tag{5.3}$$

This estimator is unbiased for $\int \psi \, \mathrm{d}\mu$, and

$$\mathbb{E}_{\mu,\varsigma}\left|\widehat{\theta}_\psi - \int \psi \, \mathrm{d}\mu\right|^2 \leq \frac{4R^2+\varsigma^2}{n}\left\|\widetilde{G}_\psi\right\|^2_{L^2(Q_0)} \leq \frac{C}{n}\|\psi\|^2_{H^s}. \tag{5.4}$$

Indeed, $\mathbb{E}(Y_i^2 \mid Z_i) = C_\mu(Z_i)^2 + \varsigma^2 \leq 4R^2 + \varsigma^2$.

To reconstruct the law from these moments, work on the torus $\mathbb{T}_T^d = \mathbb{R}^d/(T\mathbb{Z}^d)$ with $T > 4R$, represented by the cube $[-T/2, T/2)^d$ with opposite faces identified. The ball B_R lies strictly inside this cube. Moreover, torus geodesic distance agrees with Euclidean distance between points of B_r. Choose a real $\chi \in C_c^\infty((-T/2,T/2)^d)$ with $\chi = 1$ on B_r, and set $\psi_{\boldsymbol{k}}(\boldsymbol{x}) = \chi(\boldsymbol{x})e^{-2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}/T}$, $\boldsymbol{k} \in \mathbb{Z}^d$. Then $\|\psi_{\boldsymbol{k}}\|_{H^s} \leq C(1+|\boldsymbol{k}|)^s$. Write $a_{\boldsymbol{k}} = a_{\psi_{\boldsymbol{k}}}$ and $\widetilde{G}_{\boldsymbol{k}} = \widetilde{G}_{\psi_{\boldsymbol{k}}}$. For $\boldsymbol{k} \neq 0$, let $\widehat{\theta}_{\boldsymbol{k}}$ be (5.3) with $\psi = \psi_{\boldsymbol{k}}$; it estimates $\theta_{\boldsymbol{k}}(\mu) = \int e^{-2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}/T}\,\mathrm{d}\mu(\boldsymbol{x})$. Set $\widehat{\theta}_{\boldsymbol{0}} = 1$. Choose the features consistently with complex conjugation, so that $\widehat{\theta}_{-\boldsymbol{k}} = \overline{\widehat{\theta}_{\boldsymbol{k}}}$.

Fix a real even $w \in C_c^\infty(\mathbb{R}^d)$ equal to one on the unit ball and zero outside the ball of radius two. For a frequency cutoff $M \geq 1$, define

$$\widetilde{f}_M(\boldsymbol{x}) = \frac{1}{T^d}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} w(\boldsymbol{k}/M)\widehat{\theta}_{\boldsymbol{k}} e^{2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}/T}. \tag{5.5}$$

$\widetilde{f}_M$ is finite and real. Orthogonality of the Fourier modes gives $\int_{\mathbb{T}_T^d} \widetilde{f}_M(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = w(0)\widehat{\theta}_{\boldsymbol{0}} = 1$. However, $\widetilde{f}_M$ need not be nonnegative, nor need it be supported in B_r. It therefore defines a signed measure of mass one, not necessarily a probability law. We enforce the probability and support constraints by projection. For a finite signed measure ζ of mass zero on $\mathbb{T}_T^d$, put

$$\|\zeta\|_{\mathrm{Lip}^*} = \sup_{\mathrm{Lip}(F)\leq 1}\left|\int F \,\mathrm{d}\zeta\right|,$$

where Lipschitz constants use torus geodesic distance. For probabilities supported on B_r, this norm of their difference equals their Euclidean W_1 distance. Choose $\widehat{\mu}_M \in \mathcal{P}(B_r)$ satisfying

$$\left\|\widehat{\mu}_M - \widetilde{f}_M \,\mathrm{d}\boldsymbol{x}\right\|_{\mathrm{Lip}^*} \le \inf_{\nu\in\mathcal{P}(B_r)} \left\|\nu - \widetilde{f}_M \,\mathrm{d}\boldsymbol{x}\right\|_{\mathrm{Lip}^*} + \frac{1}{n}. \tag{5.6}$$

A measurable choice is constructed in the proof of Proposition 5.4 in the Appendix.

PROPOSITION 5.4 (Direct reconstruction bounds). *Under (2.7), for every $M \ge 1$ and $n \ge 1$, the estimator in (5.6) satisfies, with constants uniform over $0 \le \varsigma \le \overline{\varsigma}$,*

$$\sup_{f\in\mathcal{F}_\alpha(L)} \mathbb{E}_{f,\varsigma} W_1(\widehat{\mu}_M, f\,\mathrm{d}\boldsymbol{x}) \le C\left[LM^{-\alpha-1} + n^{-1/2}M^{d+1/2}\right] + \frac{1}{n}, \tag{5.7}$$

$$\sup_{\mu\in\mathcal{P}(B_r)} \mathbb{E}_{\mu,\varsigma} W_1(\widehat{\mu}_M, \mu) \le C\left[M^{-1} + n^{-1/2}M^{d+1/2}\right] + \frac{1}{n}. \tag{5.8}$$

Taking $M \asymp n^{1/D_\alpha}$ in (5.7) gives the rate $n^{-(\alpha+1)/D_\alpha}$. The next result shows that this rate is optimal when the noise level is fixed and positive and there is no systematic bias.

PROPOSITION 5.5 (Lower bound at fixed positive noise). *Fix $\varsigma > 0$, $\alpha \ge 0$, and $L > L_0$ as in Theorem 4.5. Under the finite-Fisher-information assumption on the error density, there is $c > 0$ such that $\mathfrak{R}_n(\mathcal{F}_\alpha(L), \varsigma) \ge cn^{-(\alpha+1)/(2\alpha+2d+3)}$, where c may depend on ς.*

5.2. *Estimating the Pricing Surface Locally.* The direct estimator does not fully exploit highly accurate quotes: even when $\varsigma = 0$, its sample average fluctuates because the quote locations are random. To reduce this variation, we first estimate the entire pricing surface from part of the data. We call this preliminary estimate a *pilot*; the remaining observations will correct its error rather than estimate the original surface from scratch. The pilot fits polynomials to nearby direction–strike observations. Its bandwidth h is the nominal side length of the local fitting regions: smaller h resolves finer variation, whereas larger h pools more observations. Where a region contains too few suitably placed observations, the construction enlarges it until a stable fit is possible. The Appendix gives the precise construction.

LEMMA 5.6 (Pricing-surface pilots). *Under (2.7), for $m \ge 1$ and $\varsigma \ge 0$, there are measurable estimators $\widehat{C}_m$ taking values in $[0, 2R]$ with the following properties. For $\alpha \ge 0$ and $t = \alpha + s$, a local polynomial construction of fixed coordinatewise degree at least $\lceil t \rceil$ satisfies*

$$\sup_{f\in\mathcal{F}_\alpha(L)} \mathbb{E}_{f,\varsigma} \left\|\widehat{C}_m - C_f\right\|^2_{L^2(Q_0)} \le C\left[L^2 m^{-2t/d} + \varsigma^2\right]. \tag{5.9}$$

A construction of any fixed coordinatewise degree at least one satisfies

$$\sup_{\mu\in\mathcal{P}(B_r)} \mathbb{E}_{\mu,\varsigma} \left\|\widehat{C}_m - C_\mu\right\|^2_{L^2(Q_0)} \le C\left[m^{-3/d} + \varsigma^2\right]. \tag{5.10}$$

More generally, for any prescribed $h \in [m^{-1/d}, 1]$, the corresponding estimators $\widehat{C}_{m,h}$ satisfy

$$\begin{aligned} \sup_{f\in\mathcal{F}_\alpha(L)} \mathbb{E}_{f,\varsigma} \left\|\widehat{C}_{m,h} - C_f\right\|^2_{L^2(Q_0)} &\le C\Big[L^2h^{2t} + \frac{\varsigma^2}{mh^d}\Big], \\ \sup_{\mu\in\mathcal{P}(B_r)} \mathbb{E}_{\mu,\varsigma} \left\|\widehat{C}_{m,h} - C_\mu\right\|^2_{L^2(Q_0)} &\le C\Big[h^3 + \frac{\varsigma^2}{mh^d}\Big]. \end{aligned} \tag{5.11}$$

The two bounds concern the respective polynomial constructions above. Before pointwise truncation to $[0, 2R]$*, each estimator is linear in* $(Y_i)_{i=1}^m$ *conditional on* $(Z_i)_{i=1}^m$*. All assertions remain valid for i.i.d. locations with law* $Q \geq q_* Q_0$*, with constants depending on* $q_* > 0$*; the estimators do not require knowledge of* Q *or a density of* Q*. Under (5.1), the same bounds hold with an additional term* $C\epsilon^2$ *on each right-hand side.*

In (5.11), the first term is approximation error and the second is the contribution of independent noise. The choice $h \asymp m^{-1/d}$ gives (5.9) and (5.10). At this fine resolution the pilot alone need not average away the noise; its role is to make the subsequent residuals small when quotes are accurate.

5.3. *Residual Correction and Probability-Valued Recovery.* For $e = 0$, the variance of a direct moment estimate has the exact decomposition

$$\operatorname{Var}(\widehat{\theta}_\psi) = \frac{1}{n} \operatorname{Var}\Big(\widetilde{G}_\psi(Z) C_\mu(Z)\Big) + \frac{\varsigma^2}{n} \left\|\widetilde{G}_\psi\right\|^2_{L^2(Q_0)},$$

The first term is caused by random quote locations and remains nonzero for exact quotes. Thus the direct bound in Proposition 5.4 does not improve to the exact-information rate merely by setting $\varsigma = 0$. Residual correction removes most of this location-induced variation. For $n \geq 2$, we split the sample into groups of sizes $n_0 = \lfloor n/2 \rfloor$ and $n_1 = n - n_0$, and construct a pilot $\widehat{C}$ from the first group. For $\boldsymbol{k} \neq 0$, define

$$\widehat{\theta}^{\mathrm{cv}}_{\boldsymbol{k}} = a_{\boldsymbol{k}} + \int_{\mathcal{Z}} \widetilde{G}_{\boldsymbol{k}}(z) \widehat{C}(z) \, \mathrm{d}Q_0(z) + \frac{1}{n_1} \sum_{i=n_0+1}^{n} \widetilde{G}_{\boldsymbol{k}}(Z_i) \Big(Y_i - \widehat{C}(Z_i)\Big), \quad \widehat{\theta}^{\mathrm{cv}}_{\mathbf{0}} = 1. \tag{5.12}$$

The integral accounts for the estimated surface using the known design Q_0; only the residual is averaged over the second group. If the pilot equaled C_μ and the quotes were exact and unbiased, these residuals would vanish. More generally, conditional on the first group, the corrected moment has mean $\theta_{\boldsymbol{k}}(\mu) + \int_{\mathcal{Z}} \widetilde{G}_{\boldsymbol{k}}(z) e(z) \, \mathrm{d}Q_0(z)$. In particular, pilot approximation does not introduce bias into the corrected moment when $e = 0$.

Use (5.5) with these corrected moments to obtain $\widetilde{f}^{\mathrm{cv}}_M$, and choose $\widehat{\mu}^{\mathrm{cv}}_M \in \mathcal{P}(B_r)$ satisfying

$$\left\|\widehat{\mu}^{\mathrm{cv}}_M - \widetilde{f}^{\mathrm{cv}}_M \, \mathrm{d}\boldsymbol{x}\right\|_{\mathrm{Lip}^*} \leq \inf_{\nu \in \mathcal{P}(B_r)} \left\|\nu - \widetilde{f}^{\mathrm{cv}}_M \, \mathrm{d}\boldsymbol{x}\right\|_{\mathrm{Lip}^*} + e^{-n}. \tag{5.13}$$

The corrected construction preserves complex conjugation and total mass. A measurable projection is obtained as in Proposition 5.4. The smaller projection tolerance e^{-n} avoids limiting the attainable rate when α is large.

PROPOSITION 5.7 (Pilot-to-law risk bound). *Let the data follow (5.1). Suppose* $\widehat{C}$ *is based only on the first group and* $\mathbb{E}\|\widehat{C} - C_\mu\|^2_{L^2(Q_0)} \leq v_n(\mu)$*. For every* M, n *and* $f \in \mathcal{F}_\alpha(L)$*,*

$$\mathbb{E}_{f,e} W_1(\widehat{\mu}^{\mathrm{cv}}_M, f \, \mathrm{d}\boldsymbol{x}) \leq C\left[L M^{-\alpha-1} + \epsilon M^{s-1} + \sqrt{\frac{\varsigma^2 + \epsilon^2 + v_n(f)}{n}} \, M^{d+\frac{1}{2}}\right] + e^{-n}. \tag{5.14}$$

For every $\mu \in \mathcal{P}(B_r)$*,*

$$\mathbb{E}_{\mu,e} W_1(\widehat{\mu}^{\mathrm{cv}}_M, \mu) \leq C\left[M^{-1} + \epsilon M^{s-1} + \sqrt{\frac{\varsigma^2 + \epsilon^2 + v_n(\mu)}{n}} \, M^{d+\frac{1}{2}}\right] + e^{-n}. \tag{5.15}$$

The three contributions in (5.14) are Fourier approximation, amplified systematic bias, and random fluctuation. Unlike the direct bound, the last contribution is controlled by the pilot error and the noise level rather than by the size of the pricing surface itself.

PROOF SKETCH OF THEOREM 5.1. Lemma 5.6, with $m = n_0$ and $h \asymp n^{-1/d}$, gives $v_n(f) \le C[L^2 n^{-2t/d} + \varsigma^2 + \epsilon^2]$ uniformly over the class and admissible biases. Substitution into (5.14) and the cutoff $M \asymp \min\left\{n^{1/d}, (n/\varsigma^2)^{1/D_\alpha}, (L/\epsilon)^{1/t}\right\}$ give the upper bound. Entries involving division by zero are interpreted as $+\infty$; the cutoff is rounded and truncated below at one, with bounded-parameter cases handled by $W_1 \le 2r$.

For the lower bound, the exact-information argument of Theorem 4.6 applies even after independent noise is added. The testing argument of Proposition 5.5, with spatial scale $(\varsigma^2/n)^{1/D_\alpha}$, gives the noise-dependent term uniformly over positive bounded noise levels. Finally, Lemma D.2 provides two laws separated by order ϵ^{ϑ} whose pricing surfaces differ uniformly by at most 2ϵ. Choosing opposite biases to make their mean prices coincide gives the systematic term. The maximum of the three lower bounds controls their sum up to a constant. Full details are in the Appendix. □

5.4. *General Designs and Additional Marginal Information.* The integral in (5.12) uses the known uniform design. The same recovery rates can nevertheless be attained without knowing the sampling law, provided that every region receives sufficiently many quotes. For this purpose, we smooth a local estimate of the pricing surface and project it onto the pricing surfaces generated by the density class.

For $h \in [n^{-1/d}, 1]$, construct $\widehat{C}_{n,h}$ from all observations as in Lemma 5.6. Let J_h be a fixed linear smoothing operator that removes frequencies above order h^{-1} in coordinate charts, as constructed in the Appendix. Define $\widehat{C}^{\mathrm{proj}}_{n,h} = \operatorname{argmin}_{G \in \{C_f : f \in \mathcal{F}_\alpha(L)\}} \|G - J_h \widehat{C}_{n,h}\|_{H^{s-1}(\mathcal{Z})}$, and let $\widehat{\mu}_{n,h}$ be the law with this pricing surface. The projection is well defined and measurable: its target is a compact convex subset of the Hilbert space $H^{s-1}(\mathcal{Z})$, and Lemma 3.3 makes the inverse map to laws continuous. Write $\widehat{\mu}_n = \widehat{\mu}_{n,h}$ at the selected bandwidth.

Under $Q \ge q_* Q_0$, the construction gives

$$\sup_{f \in \mathcal{F}_\alpha(L)} \sup_{|e| \le \epsilon} \mathbb{E}_{f,e,\varsigma} W_1(\widehat{\mu}_{n,h}, f\,\mathrm{d}\boldsymbol{x}) \le C\left[Lh^{\alpha+1} + \varsigma n^{-1/2} h^{-d-1/2} + \epsilon h^{-(s-1)}\right]. \tag{5.16}$$

The terms represent approximation, independent noise, and systematic bias. A bandwidth of order $\min\left\{1, \max\left\{n^{-1/d}, (\varsigma^2/n)^{1/D_\alpha}, (\epsilon/L)^{1/t}\right\}\right\}$ yields the upper rate in Theorem 5.1. The local construction rounds h upward to a power of two, changing it by at most a factor of two. This discretization does not affect the rate. The estimator uses the observed locations and prices and may use $\alpha, L, \varsigma, \epsilon$ and the support, but does not require knowledge of Q.

COROLLARY 5.8 (General i.i.d. designs). *Under the class assumptions of Theorem 5.1, every i.i.d. design $Q \ge q_* Q_0$ with $q_* > 0$ satisfies*

$$\mathfrak{R}^{\mathrm{ex}}_n(Q^{\otimes n}, \alpha, L) \asymp n^{-(\alpha+1)/d}. \tag{5.17}$$

Under (5.1) with Q in place of Q_0, the upper bound in (5.2) also holds, uniformly over $0 \le \varsigma \le \overline{\varsigma}$ and $0 \le \epsilon \le \overline{\epsilon}$. If additionally $Q \le q^* Q_0$ for some finite q^*, the matching lower bound holds as well. Constants may depend on the stated design bounds.

Single-asset option prices provide additional information: observing their full strike curves determines each coordinate marginal law. One might therefore expect that knowing all these marginals would improve recovery of the joint law. The next result shows that, on a suitable fixed-marginal subclass, the remaining uncertainty about dependence is already sufficient to produce all three terms of the minimax rate.

COROLLARY 5.9 (Fixed marginal laws). *Let f_0 be the base density of Lemma C.4. Let $\mathcal{P}_0$ consist of the laws $f\,\mathrm{d}\boldsymbol{x}$ with $f \in \mathcal{F}_\alpha(L)$ whose one-dimensional coordinate marginals equal those of $f_0\,\mathrm{d}\boldsymbol{x}$. Under the assumptions of Theorem 5.1, its lower bound remains valid with $\mathcal{F}_\alpha(L)$ replaced by $\mathcal{P}_0$. Its estimator attains the matching upper bound on $\mathcal{P}_0$.*

REMARK 5.10 (What known marginals do and do not resolve). Corollary 5.9 concerns exact knowledge of the full coordinate marginals, not merely finitely many single-asset quotes. Even this additional information does not improve the worst-case rate on the specified subclass: the alternatives used for all three lower bounds have identical coordinate marginals. The remaining difficulty is therefore recovery of dependence. This is a rate statement for that subclass, not a claim that marginal information is unhelpful for every law or every payoff.

5.5. *Arbitrary Probability Laws.* The preceding sharp rate assumes a Sobolev density. The universal pilot in (5.10) also applies to atomic and singular laws. When $e = 0$, combining it with the same residual correction yields the following bounds without a density assumption.

THEOREM 5.11 (Bounds over all laws uniformly in the noise level). *Fix a finite $\bar{\varsigma} \geq 0$. Under (2.7), there are constants $c, C > 0$, independent of $n \geq 1$ and $0 \leq \varsigma \leq \bar{\varsigma}$, such that*

$$c\left[n^{-\frac{1}{d}} + \left(\frac{\varsigma^2}{n}\right)^{\frac{1}{2d+3}}\right] \leq \mathfrak{R}_n(\mathcal{P}(B_r), \varsigma) \leq C\left[n^{-\frac{d+3}{d(2d+3)}} + \left(\frac{\varsigma^2}{n}\right)^{\frac{1}{2d+3}}\right]. \tag{5.18}$$

For every fixed $c_0 > 0$, the two sides have the same order on the regime $\varsigma^2 \geq c_0 n^{-3/d}$, with comparison constants allowed to depend on c_0.

Thus the noise-dominated rate is sharp, but a gap remains for exact or very low-noise quotes over arbitrary laws. For example, when $d = 2$ and $\varsigma = 0$, $cn^{-1/2} \leq \mathfrak{R}_n(\mathcal{P}(B_r), 0) \leq Cn^{-5/14}$. Closing this gap requires a sharper recovery argument for laws that need not have a smooth density; it is not resolved by the Sobolev-class result above.

6. Price Intervals and Minimax Estimation of a Price. Recovering the law controls the prices of all Lipschitz claims, but estimating the price of a specified payoff can be substantially easier. We first express the uncertainty left by exact quotes as a price interval. We then introduce two claim-specific moduli: one for uniformly bounded quote discrepancies, and one for discrepancies measured in the design norm. The latter characterizes minimax estimation at a fixed positive noise level. Its dual formulation yields both an affine estimator and confidence intervals through a penalized static replication problem.

6.1. *Price Intervals and Wasserstein Diameter.* For $\mathcal{K} \subseteq \mathcal{P}(B_R)$ and $g \in C(B_R)$, define

$$I_g(\mathcal{K}) = \left[\min_{\nu \in \mathcal{K}} \int g \, \mathrm{d}\nu, \ \max_{\nu \in \mathcal{K}} \int g \, \mathrm{d}\nu\right].$$

If $\mathcal{K}$ is compact and convex, every value in this interval is the price of g under some law in $\mathcal{K}$. In particular, $I_g(\mathcal{K}_n(z, v))$ is the range of prices consistent with the observed exact quotes and the support constraint. Its width measures the remaining uncertainty about the claim g.

PROPOSITION 6.1 (Uncertainty equals Wasserstein diameter). *For compact $\mathcal{K} \subseteq \mathcal{P}(B_R)$,*

$$\sup_{\mathrm{Lip}(g) \leq 1} \left(\max_{\nu \in \mathcal{K}} \int g \, \mathrm{d}\nu - \min_{\nu \in \mathcal{K}} \int g \, \mathrm{d}\nu\right) = \mathrm{diam}_{W_1} \mathcal{K}. \tag{6.1}$$

For $\mathcal{K} = \mathcal{K}_n(\mu)$ and $g \in C(B_R)$, the endpoints of $I_g(\mathcal{K}_n(\mu))$ converge a.s. to the corresponding endpoints of $I_g(\mathcal{K}_\infty(\mu; Q))$. If Q identifies $\mathcal{P}(B_R)$, both endpoints converge to $\int g \, \mathrm{d}\mu$.

Thus a Wasserstein diameter bound is also a bound on every 1-Lipschitz claim's price interval. More explicitly, if $Q \geq q_* Q_0$, Theorem 4.1 gives, with probability at least $1 - \delta$,

$$|I_g(\mathcal{K}_n(Z, v))| \leq C_{d,R,q_*} \left(\frac{\log(en/\delta)}{n}\right)^{2/[d(d+3)]}$$

for every price vector v and g with $\mathrm{Lip}(g) \le 1$. $|I|$ denotes interval length. The event depends only on the quote locations: once these locations provide adequate coverage, the bound holds for all such claims and price vectors simultaneously. No smoothness restriction is imposed on the admissible laws. If instead the laws are restricted to $\mathcal{F}_\alpha(L)$, Theorem 4.3 gives

$$|I_g(\mathcal{K}_{n,\alpha,L}(Z,v))| \le CL\Big(\frac{\log(en/\delta)}{n}\Big)^{(\alpha+1)/d}$$

on an event of probability at least $1-\delta$, again simultaneously over v and 1-Lipschitz g. These are worst-case bounds over claims; the uncertainty for a specified payoff may be smaller.

6.2. *Uniform Quote Discrepancies and Static Replication.* Let $E \subseteq S^{d-1}$ be a nonempty closed set of quoted directions, put $\mathcal{Z}_E = E \times [-R,R]$, and let $\mathcal{M}(\mathcal{Z}_E)$ denote the finite signed measures on $\mathcal{Z}_E$. Write $\phi_z(\boldsymbol{x}) = (\boldsymbol{u}\cdot\boldsymbol{x} - k)_+$ for $z = (\boldsymbol{u}, k)$. A signed measure $m \in \mathcal{M}(\mathcal{Z}_E)$ specifies a static portfolio of basket calls with payoff

$$\mathcal{A}^* m(\boldsymbol{x}) = \int_{\mathcal{Z}_E} \phi_z(\boldsymbol{x})\, m(\mathrm{d}z), \qquad \boldsymbol{x} \in B_R.$$

This payoff is $\|m\|$-Lipschitz, where $\|m\|$ is the total variation norm, and $\int \mathcal{A}^* m \,\mathrm{d}\eta = \int_{\mathcal{Z}_E} C_\eta \,\mathrm{d}m$ for every $\eta \in \mathcal{M}(B_R)$. The payoffs obtainable from cash and quoted calls form

$$V_E = \{c + \mathcal{A}^* m : c \in \mathbb{R},\ m \in \mathcal{M}(\mathcal{Z}_E)\} \subset C(B_R). \tag{6.2}$$

If E has positive surface measure, let Q_E be the product of normalized surface measure on E and the uniform strike distribution on $[-R,R]$. Define the uniform-quote pricing modulus

$$\varpi_g(\epsilon) = \sup\Big\{ \int g \,\mathrm{d}(\mu-\nu) :\ \mu,\nu \in \mathcal{P}(B_R),\ \sup_{\mathcal{Z}_E} |C_\mu - C_\nu| \le \epsilon \Big\} \ \text{ for } g \in C(B_R),\ \epsilon \ge 0.$$

Let $\varpi_g^{\alpha,L}(\epsilon)$ denote the same supremum restricted to laws with densities in $\mathcal{F}_\alpha(L)$. The modulus is the largest price difference compatible with quote discrepancies of at most ϵ at every quoted direction and strike. Write $\operatorname{osc} h = \max_{B_R} h - \min_{B_R} h$.

LEMMA 6.2 (Replication bound). *For every $g \in C(B_R)$, $m \in \mathcal{M}(\mathcal{Z}_E)$, and $\epsilon \ge 0$,*

$$\varpi_g(\epsilon) \le \operatorname{osc}(g - \mathcal{A}^* m) + \epsilon \,\|m\|\,. \tag{6.3}$$

In particular, $\varpi_g(\epsilon) \le \epsilon\,\|m\|$ whenever $g = c + \mathcal{A}^* m \in V_E$.

The first term bounds the price uncertainty of the unreplicated payoff; the second bounds the effect of quote discrepancies on the portfolio. Cash does not affect either price differences or oscillation. Exact replication therefore gives a linear bound in the quote tolerance, while approximate replication trades residual uncertainty against the size of the portfolio. Section 7 uses this bound to distinguish claims that are stable under restricted directional information.

6.3. *The Design Modulus and Minimax Price Estimation.* Fix $K \in \{B_r, B_R\}$ and a nonempty convex weakly closed class $\mathcal{C} \subseteq \mathcal{P}(K)$. The cases used below are $\mathcal{P}(B_R)$, $\mathcal{P}(B_r)$, and $\mathcal{C}_{\alpha,L} = \{f\,\mathrm{d}\boldsymbol{x} : f \in \mathcal{F}_\alpha(L)\}$. Let Q be a known probability measure on $\mathcal{Z}_E$, and observe

$$Z_i \overset{\text{i.i.d.}}{\sim} Q, \qquad Y_i = C_\mu(Z_i) + \varsigma\xi_i, \qquad i = 1,\dots,n,$$

where $\varsigma > 0$ is fixed and the errors are independent of the design, centered, and have unit variance, as in (2.7). There is no systematic bias in this experiment. For $g \in C(K)$, put

$$\mathfrak{R}_n^Q(g;\mathcal{C}) = \inf_{\widehat{p}} \sup_{\mu\in\mathcal{C}} \mathbb{E}_\mu \left| \widehat{p} - \int g \,\mathrm{d}\mu \right|,$$

where the infimum is over real-valued estimators based on $(Z_i, Y_i)_{i=1}^n$. The relevant discrepancy for this experiment is measured in $L^2(Q)$ rather than uniformly on $\mathcal{Z}_E$. Define

$$\omega_g(\epsilon; \mathcal{C}) = \sup\left\{ \int g \,\mathrm{d}(\mu - \nu) : \ \mu, \nu \in \mathcal{C}, \ \|C_\mu - C_\nu\|_{L^2(Q)} \le \epsilon \right\}, \qquad \epsilon \ge 0,$$

and define its price-range width over $\mathcal{C}$ by $\mathrm{osc}_{\mathcal{C}}(h) = \max_{\mu\in\mathcal{C}} \int h \,\mathrm{d}\mu - \min_{\mu\in\mathcal{C}} \int h \,\mathrm{d}\mu$ for $h \in C(K)$. For $\mathcal{C} = \mathcal{P}(K)$, this width is $\max_K h - \min_K h$. For $w \in L^2(Q)$, the portfolio

$$\mathcal{A}^*(wQ)(\boldsymbol{x}) = \int_{\mathcal{Z}_E} \phi_z(\boldsymbol{x}) w(z) \, Q(\mathrm{d}z)$$

is bounded in absolute value by $2R\|w\|_{L^1(Q)}$ and is $\|w\|_{L^1(Q)}$-Lipschitz on K.

THEOREM 6.3 (Minimax estimation of a price). *Let $g \in C(K)$.*

(i) *For every $\epsilon \ge 0$,*

$$\omega_g(\epsilon; \mathcal{C}) = \inf_{w \in L^2(Q)} \left\{ \mathrm{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ)) + \epsilon \|w\|_{L^2(Q)} \right\}, \tag{6.4}$$

and the supremum defining $\omega_g(\epsilon; \mathcal{C})$ is attained.

(ii) *Suppose additionally, for the lower bound, that $I_\xi < \infty$ as in Section 2. For every $n \ge 1$,*

$$\frac{1}{8\sqrt{1 \vee I_\xi}} \, \omega_g\left(\frac{\varsigma}{\sqrt{n}}; \mathcal{C}\right) \le \mathfrak{R}_n^Q(g; \mathcal{C}) \le \sqrt{1 + \frac{4R^2}{\varsigma^2}} \, \omega_g\left(\frac{\varsigma}{\sqrt{n}}; \mathcal{C}\right). \tag{6.5}$$

The upper bound is attained up to an arbitrarily small additive error by $\widehat{p}_w = c_w + \frac{1}{n}\sum_{i=1}^n w(Z_i) Y_i$, where w is a deterministic near-minimizer in (6.4) at $\epsilon = \varsigma/\sqrt{n}$, and $c_w = \frac{1}{2}\left\{\max_{\mu\in\mathcal{C}} \int (g - \mathcal{A}^*(wQ)) \,\mathrm{d}\mu + \min_{\mu\in\mathcal{C}} \int (g - \mathcal{A}^*(wQ)) \,\mathrm{d}\mu\right\}$.

The dual objective balances residual price uncertainty against the $L^2(Q)$ size of the portfolio weights. The midpoint c_w centers the residual price range, giving worst-case absolute bias at most $\frac{1}{2}\mathrm{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ))$. Moreover, for $\mu \in \mathcal{C}$, $\mathrm{Var}_\mu(\widehat{p}_w) \le \frac{4R^2+\varsigma^2}{n} \|w\|_{L^2(Q)}^2$. This variance includes both quote noise and variation due to the random locations. Consequently, the comparison in (6.5) is a fixed-positive-noise result: its upper comparison constant diverges as $\varsigma \downarrow 0$, so setting $\varsigma = 0$ does not characterize the exact-quote risk.

The two moduli serve different purposes. Uniform quote control implies $L^2(Q)$ control, so $\varpi_g(\epsilon) \le \omega_g(\epsilon; \mathcal{P}(B_R))$ and $\varpi_g^{\alpha,L}(\epsilon) \le \omega_g(\epsilon; \mathcal{C}_{\alpha,L})$. Thus a lower bound for the uniform-quote modulus also supplies a lower bound for the statistical modulus on the same class. The converse comparison requires additional assumptions on the design. Likewise, a representing measure m in Lemma 6.2 need not have an $L^2(Q)$ density. Membership in V_E alone therefore does not give an $L^2(Q)$ replication bound; this must be established separately.

Theorem 6.3 is a random-design version of the connection between optimal recovery and affine estimation in [17]. Its proof also applies to a measurable family of continuous test functions ϕ_z on a compact metric space with $\sup_{z,x} |\phi_z(x)| \le B$. In that setting, B^2 replaces $4R^2$ in the risk and variance bounds, and the same modulus, dual formulation, and affine construction apply to any nonempty convex weakly closed class of probability laws.

6.4. *Confidence Intervals and Optimal Length.* The same bias and variance bounds give confidence intervals without a Gaussian or tail assumption on the errors. The weights below are chosen from the known class and design, independently of the observed sample.

COROLLARY 6.4 (Confidence intervals for a price). *Under the setting of Theorem 6.3, let $g \in C(K)$ and $0 < \delta < \frac{1}{4}$. For a deterministic $w \in L^2(Q)$, let $\widehat{p}_w$ and c_w be as above, and put $\widehat{I}_w = [\widehat{p}_w - \ell_w,\ \widehat{p}_w + \ell_w]$ and $\ell_w = \frac{1}{2}\operatorname{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ)) + \sqrt{\frac{4R^2+\varsigma^2}{n\delta}}\,\|w\|_{L^2(Q)}$.*

(i) *For every such w, $\inf_{\mu\in\mathcal{C}} \Pr_\mu\left(\int g\,\mathrm{d}\mu \in \widehat{I}_w\right) \geq 1-\delta$. For every $\eta > 0$, one can choose w so that $\ell_w \leq \delta^{-1/2}\sqrt{1+\frac{4R^2}{\varsigma^2}}\left\{\omega_g\left(\frac{\varsigma}{\sqrt{n}};\mathcal{C}\right)+\eta\right\}$. The full interval length is $|\widehat{I}_w| = 2\ell_w$.*

(ii) *If the error density has finite Fisher information I_ξ, every data-measurable interval $\widehat{I}$ satisfying $\inf_{\mu\in\mathcal{C}} \Pr_\mu\left(\int g\,\mathrm{d}\mu \in \widehat{I}\right) \geq 1-\delta$ obeys $\sup_{\mu\in\mathcal{C}} \mathbb{E}_\mu|\widehat{I}| \geq \frac{\frac{1}{2}-2\delta}{\sqrt{1\vee I_\xi}}\,\omega_g\left(\frac{\varsigma}{\sqrt{n}};\mathcal{C}\right)$.*

For fixed $\delta \in (0, 1/4)$ and $\varsigma > 0$, these bounds show that the smallest possible worst-case expected length among all intervals with uniform coverage at least $1-\delta$ is comparable to $\omega_g(\varsigma/\sqrt{n};\mathcal{C})$. The intervals constructed above attain this order. The comparison constants depend only on δ, R/ς, and I_ξ, not on n, g, or $\mathcal{C}$. This is an optimality statement for worst-case expected length under uniform coverage, not for the length at each individual law. Unlike the exact-quote price intervals of Section 6.1, these are statistical confidence intervals.

7. Directional Pricing and Exchange Options. Long-only basket quotes identify the joint law, yet the accuracy of an inferred price depends strongly on the payoff. Polynomial claims admit parametric estimation at fixed positive noise, while an exchange option in an unquoted spread direction has only logarithmic recovery. We first explain why missing directions make inversion unstable, then identify claims that remain well priced. We next establish sharp logarithmic moduli for unquoted kinks, compare uncertainty across classes of claims, and extend the exchange-price result to exact and noisy quotes.

7.1. *Missing Projection Directions.* Identifying a law does not imply a stable inverse. The relevant geometry concerns projection lines rather than oriented directions: $t \mapsto \widehat{\mu}(t\boldsymbol{u})$ covers both signs of $\boldsymbol{u}$. Thus a hemisphere meets every projection line, whereas long-only directions miss lines with mixed signs, including those near the spread direction $(\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}$.

THEOREM 7.1 (Missing-cone instability). *Suppose there are $\boldsymbol{e} \in S^{d-1}$, $\kappa > 0$ such that*

$$|\boldsymbol{u}\cdot\boldsymbol{e}| \leq 1-\kappa \qquad \text{for all } \boldsymbol{u} \in E. \tag{7.1}$$

For every $\alpha \geq 0$ and every sufficiently large L, there are densities $f_N^+, f_N^- \in \mathcal{F}_\alpha(L)$ satisfying

$$W_1(f_N^+\,\mathrm{d}\boldsymbol{x}, f_N^-\,\mathrm{d}\boldsymbol{x}) \geq cN^{-\alpha-1}, \tag{7.2}$$

whereas, for each $A > 0$, there is $C_A < \infty$ such that

$$\sup_{\boldsymbol{u}\in E,\, k\in[-R,R]} \left|C_{f_N^+}(\boldsymbol{u},k) - C_{f_N^-}(\boldsymbol{u},k)\right| \leq C_A N^{-A}. \tag{7.3}$$

Consequently, for every $\gamma > 0$, no finite constant C can make

$$W_1(f\,\mathrm{d}\boldsymbol{x}, g\,\mathrm{d}\boldsymbol{x}) \leq C\Big(\sup_{\boldsymbol{u}\in E,\, k\in[-R,R]} |C_f(\boldsymbol{u},k) - C_g(\boldsymbol{u},k)|\Big)^\gamma$$

hold uniformly over $f, g \in \mathcal{F}_\alpha(L)$. The same conclusion holds with an L^p pricing norm for any probability design supported on $E \times [-R,R]$ and $1 \leq p \leq \infty$.

COROLLARY 7.2 (Long-only basket directions). *Let $E = \{\boldsymbol{u} \in S^{d-1} : u_j \geq 0 \text{ for all } j\}$. Call prices on $E \times [-R,R]$ identify the joint law, but admit no uniform Hölder inverse bound as in Theorem 7.1 on a sufficiently rich bounded Sobolev class.*

7.2. *Stable Claims and Polynomial Prices.* Instability of the joint law need not make every claim difficult to price. A payoff represented by cash and a portfolio of quoted calls inherits the stability bound of Lemma 6.2. The next proposition gives approximation bounds.

PROPOSITION 7.3 (Stable claims). *Let $\boldsymbol{v} \in E$, $G : [-R, R] \to \mathbb{R}$, and $g(\boldsymbol{x}) = G(\boldsymbol{v} \cdot \boldsymbol{x})$.*

(i) *If G' has bounded variation, then $\varpi_g(\epsilon) \le (|G'(-R+)| + |G''|((-R, R)))\,\epsilon$.*
(ii) *If G is 1-Lipschitz, then $\varpi_g(\epsilon) \le 4\sqrt{2R\epsilon} + \epsilon$.*
(iii) *If E has nonempty relative interior in S^{d-1}, every polynomial g belongs to V_E defined in (6.2) and is Lipschitz stable, i.e., $\varpi_g(\epsilon) \le C_g\epsilon$.*

EXAMPLE 7.4. Let $E = \{\boldsymbol{u} \in S^{d-1} : u_j \ge 0 \ \forall j\}$ and $\boldsymbol{v} = (\boldsymbol{e}_1 + \boldsymbol{e}_2)/\sqrt{2} \in E$. The squared spread $(x_1 - x_2)^2 = 2x_1^2 + 2x_2^2 - 2(\boldsymbol{v} \cdot \boldsymbol{x})^2$ is a static portfolio of long-only calls by Proposition 7.3(i), hence it is Lipschitz stable. In particular, index and single-asset second moments determine the covariance, and hence the second moment of a spread.

EXAMPLE 7.5 (Polynomial claims from long-only quotes). Let E have nonempty relative interior, let $Q \ge q_* Q_E$ be known, and fix $\varsigma > 0$. For every nonconstant polynomial g,

$$cn^{-\frac{1}{2}} \le \mathfrak{R}_n^Q\left(g; \mathcal{P}(B_r)\right) \le Cn^{-\frac{1}{2}}.$$

The derivation is in the Appendix. Since $\omega_g(\epsilon; \mathcal{P}(B_r)) \asymp \epsilon$ for small ϵ, Corollary 6.4 gives confidence intervals of length of order $n^{-1/2}$ for these polynomial claims.

The coverage of the restricted quote domain determines how well these stable claims can be priced. For a finite design $z \subset \mathcal{Z}_E$, let $h_E(z) = \sup_{w \in \mathcal{Z}_E} \min_i \rho(w, z_i)$ be the fill distance. For the restricted-design measure Q_E defined in Section 6.2, we use the regularity condition

$$\mathcal{H}^{d-1}(E \cap B(\boldsymbol{u}, t)) \ge c_E t^{d-1}, \qquad \boldsymbol{u} \in E,\ 0 < t \le 1, \tag{7.4}$$

where $\mathcal{H}^{d-1}$ is surface measure. The long-only set satisfies (7.4).

COROLLARY 7.6 (Finitely many quotes). *For every continuous g and feasible price vector v, the price interval of g over $\mathcal{K}_n(z, v)$ has width at most $\varpi_g(2Rh_E(z))$. If (7.4) holds and the quote locations are i.i.d. Q with $Q \ge q_* Q_E$, then with probability at least $1 - \delta$,*

$$h_E(Z_1, \dots, Z_n) \le C_{d, q_*, c_E}\left(\frac{\log(en/\delta)}{n}\right)^{1/d}.$$

On this event, simultaneously for every $g = c + \mathcal{A}^* m \in V_E$ and v, the price interval has width at most $2RC\,\|m\|\,(\log(en/\delta)/n)^{1/d}$. The exponent $1/d$ exceeds the exponent $2/(d(d+3))$ in the bound of Theorem 4.1 for all Lipschitz claims under full directional coverage.

The missing-direction phenomenon also affects statistical recovery of the full law. In the noisy model of Section 6.3, with Q any probability measure on $\mathcal{Z}_E$ and $\varsigma > 0$ fixed, define

$$\mathfrak{R}_n^Q(\alpha, L) = \inf_{\widehat{\mu}} \sup_{f \in \mathcal{F}_\alpha(L)} \mathbb{E}_{f,Q} W_1(\widehat{\mu}, f\,\mathrm{d}\boldsymbol{x}), \quad \text{where} \quad \widehat{\mu} = \widehat{\mu}((Z_i, Y_i)_{i=1}^n).$$

COROLLARY 7.7 (Absence of a polynomial noisy rate). *Suppose E satisfies (7.1). For every $\alpha \ge 0$, $\gamma > 0$ and every sufficiently large L, $\lim_{n\to\infty} n^\gamma \mathfrak{R}_n^Q(\alpha, L) = +\infty$. In particular, this holds for noisy long-only basket quotes.*

7.3. *Quoted and Unquoted Kinks.* We now turn from recovery of the law to the price of one kinked payoff. If its direction is quoted, directly or with the opposite orientation, the claim is statically replicable and its price is Lipschitz stable. Otherwise, the price can be identified but is much less stable. The next theorem gives the exact logarithmic order of this loss for a strike inside the support. Throughout, $E \subseteq S^{d-1}$ is closed with nonempty relative interior, $\boldsymbol{e} \in S^{d-1}$, and $g_k(\boldsymbol{x}) = (\boldsymbol{e}\cdot\boldsymbol{x} - k)_+$. For a finite signed measure η on B_R, $\eta_{\boldsymbol{e}}$ denotes its image under $\boldsymbol{x} \mapsto \boldsymbol{e}\cdot\boldsymbol{x}$. Its Fourier transform satisfies $\widehat{\eta_{\boldsymbol{e}}}(t) = \widehat{\eta}(t\boldsymbol{e})$.

THEOREM 7.8 (Quoted and unquoted kinks). *If $\boldsymbol{e} \in E$ or $-\boldsymbol{e} \in E$, then $\varpi_{g_k}(\epsilon) \le 2\epsilon$ for all $|k| \le R$ and $\epsilon \ge 0$. If $\pm\boldsymbol{e} \notin E$ and $|k| < r$, then for every $\alpha \ge 0$ and every sufficiently large L there are $\epsilon_0 \in (0,1)$ and $0 < c < C < \infty$ such that, for $0 < \epsilon \le \epsilon_0$,*

$$c\Big(\log\frac{1}{\epsilon}\Big)^{-(\alpha+\frac{3}{2})} \le \varpi_{g_k}^{\alpha,L}(\epsilon) \le C\Big(\log\frac{1}{\epsilon}\Big)^{-(\alpha+\frac{3}{2})}. \tag{7.5}$$

The lower bound holds for every closed E satisfying (7.1) with direction $\boldsymbol{e}$.

PROOF SKETCH. The quoted-direction bound follows from Lemma 6.2 and the put–call parity identity $(y-k)_+ = (-y+k)_+ + y - k$. For the upper bound in (7.5), express $\int g_k \,\mathrm{d}\eta$ through the cosine transform of $\eta_{\boldsymbol{e}}$ and split at frequency N. Lemma F.2 bounds the low-frequency part by $C\epsilon e^{cN}$, while Lemma F.1 bounds the tail by $CLN^{-(\alpha+3/2)}$. Taking N to be a sufficiently small multiple of $\log(1/\epsilon)$ proves the upper bound.

For the lower bound, Lemmas F.5 and F.4 give a localized oscillatory perturbation with price gap of order $N^{-(\alpha+3/2)}$ and quote discrepancy at most Ce^{-cN} on $\mathcal{Z}_E$. A nonnegative envelope preserves positivity and the uniform H^α bound. Taking N to be a sufficiently large multiple of $\log(1/\epsilon)$ proves the lower bound. Full details are in the Appendix. □

The upper bound guarantees price error at most a when $\epsilon \le \exp(-Ca^{-1/(\alpha+3/2)})$; the lower bound requires $\exp(-ca^{-1/(\alpha+3/2)})$. Since adding a payoff in V_E changes the modulus by at most a linear term in ϵ, the bounds in Theorem 7.8 extend to $g_E + c\,g_k$ with $g_E \in V_E$ and $c \neq 0$. The functional exponent refines the logarithmic stability scale associated with limited-data Radon transforms [26]. For $|k| \ge r$ the payoff g_k is zero or affine on B_r, and its modulus on $\mathcal{F}_\alpha(L)$ is at most a constant multiple of ϵ by Lemma F.3.

EXAMPLE 7.9. Let E be the long-only directions and $\boldsymbol{e} = (\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}$, so that $\pm\boldsymbol{e} \notin E$. Since $x_j = \phi_{(\boldsymbol{e}_j, -R)} - R \in V_E$, Theorem 7.8 gives moduli of order $(\log(1/\epsilon))^{-(\alpha+3/2)}$ over $\mathcal{F}_\alpha(L)$ for the exchange option $(x_1 - x_2)_+ = \sqrt{2}(\boldsymbol{e}\cdot\boldsymbol{x})_+$, the spread call $(x_1 - x_2 - K)_+$ with $|K| < \sqrt{2}r$, the best-of payoff $\max(x_1, x_2) = x_2 + (x_1 - x_2)_+$, and the worst-of payoff $\min(x_1, x_2) = x_1 - (x_1 - x_2)_+$. All four prices are identified by the full long-only pricing surface, but become Lipschitz stable when the spread direction $\boldsymbol{e}$ is quoted. Together with Example 7.4, this shows that polynomial and kinked payoffs of the same unquoted spread can have very different pricing stability.

COROLLARY 7.10 (Minimax pricing of a kink from noisy quotes). *Let E be closed and satisfy the missing-cone condition with direction $\boldsymbol{e}$, let Q be a probability measure on $\mathcal{Z}_E$, $\varsigma > 0$ fixed, $|k| < r$, $\alpha \ge 0$, L sufficiently large, $g_E \in V_E$, $c \neq 0$, and $g = g_E + c\,g_k$.*

(i) *There are $c' > 0$ and n_0 such that, for $n \ge n_0$, $\mathfrak{R}_n^Q(g; \mathcal{C}_{\alpha,L}) \ge c'(\log n)^{-(\alpha+\frac{3}{2})}$.*
(ii) *If E also has nonempty relative interior and satisfies (7.4), and $Q \ge q_* Q_E$, there are $C < \infty$ and n_0 such that, for $n \ge n_0$, $\mathfrak{R}_n^Q(g; \mathcal{C}_{\alpha,L}) \le C(\log n)^{-(\alpha+\frac{3}{2})}$.*

Hence $\mathfrak{R}_n^Q(g; \mathcal{C}_{\alpha,L}) \asymp (\log n)^{-(\alpha+3/2)}$ under the assumptions of (ii).

At fixed positive noise, Corollary 6.4 also gives confidence intervals for the exchange price of length of order $(\log n)^{-(\alpha+3/2)}$ when $q_* Q_E \le Q$, since the proof of Corollary 7.10 bounds $\omega_{g_k}(\varsigma/\sqrt{n}; \mathcal{C}_{\alpha,L})$ above and below by constant multiples of $(\log n)^{-(\alpha+3/2)}$.

7.4. *Claim-Indexed Uncertainty.* By (6.1), the W_1-diameter of $\mathcal{K} \subseteq \mathcal{P}(B_R)$ is the largest price uncertainty among all 1-Lipschitz claims. The preceding results show why it is useful to measure uncertainty over a specified class of claims rather than over all Lipschitz payoffs.

For $\mathcal{G} \subseteq C(B_R)$ closed under $g \mapsto -g$, let $d_{\mathcal{G}}(\mu,\nu) = \sup_{g\in\mathcal{G}} \int g \,\mathrm{d}(\mu-\nu)$. Then every compact $\varnothing \neq \mathcal{K} \subseteq \mathcal{P}(B_R)$ satisfies $\operatorname{diam}_{d_{\mathcal{G}}} \mathcal{K} = \sup_{g\in\mathcal{G}} \left(\max_{\nu\in\mathcal{K}} \int g \,\mathrm{d}\nu - \min_{\nu\in\mathcal{K}} \int g \,\mathrm{d}\nu\right)$. The single claim $\mathcal{G} = \{\pm g\}$ gives the width of the price interval of g, whose worst case under quote errors is the modulus ϖ_g of Section 6.2. The class $\mathcal{G} = \{1\text{-Lipschitz functions}\}$ gives W_1. Between these extremes, for a target set $D \subseteq S^{d-1}$ of basket directions let

$$\mathcal{G}_D = \{\boldsymbol{x} \mapsto G(\boldsymbol{v}\cdot\boldsymbol{x}) :\ \boldsymbol{v} \in D,\ G : [-R,R] \to \mathbb{R} \text{ 1-Lipschitz}\}, \qquad d_D = d_{\mathcal{G}_D}.$$

Then $d_D(\mu,\nu) = \sup_{\boldsymbol{v}\in D} W_1(\mu_{\boldsymbol{v}}, \nu_{\boldsymbol{v}})$ is the max-sliced Wasserstein distance over D [15], where $\mu_{\boldsymbol{v}}$ is the law of $\boldsymbol{v}\cdot X$ under μ. Let $\varpi_D(\epsilon)$ be the supremum of $d_D(\mu,\nu)$ over $\mu,\nu \in \mathcal{P}(B_R)$ satisfying $\sup_{\mathcal{Z}_E} |C_\mu - C_\nu| \le \epsilon$, and $\varpi_D^{\alpha,L}(\epsilon)$ its restriction to densities in $\mathcal{F}_\alpha(L)$.

THEOREM 7.11 (Directional moduli). (a) *For every nonempty D and closed E, $\varpi_D(\epsilon) \ge (\sqrt{2}\pi)^{-1}\sqrt{R\epsilon}$ for $0 < \epsilon \le R/(2\pi^2)$. If $D \subseteq E$, then $\varpi_D(\epsilon) \le 4\sqrt{2R\epsilon} + \epsilon$ for all $\epsilon > 0$.*

(b) *Suppose $\boldsymbol{e} \in D$ and E satisfies (7.1) with direction $\boldsymbol{e}$. Then for every $\alpha \ge 0$ and sufficiently large L there are $c, \epsilon_0 > 0$ such that $\varpi_D^{\alpha,L}(\epsilon) \ge c(\log(1/\epsilon))^{-(\alpha+1)}$ for $0 < \epsilon \le \epsilon_0$.*

(c) *If E has nonempty relative interior, $\pm\boldsymbol{e} \notin E$, and $D \subseteq E \cup \{\boldsymbol{e}\}$, then $\varpi_D^{\alpha,L}(\epsilon) \le C(\log(1/\epsilon))^{-(\alpha+1)}$ for $0 < \epsilon \le \epsilon_0$. If moreover $\boldsymbol{e} \in D$ and L is sufficiently large, then $\varpi_D^{\alpha,L}(\epsilon) \asymp (\log(1/\epsilon))^{-(\alpha+1)}$.*

Even full directional coverage cannot improve the exponent $1/2$ in (a) for arbitrary one-dimensional Lipschitz claims. The exponent $\alpha+1$ concerns worst-case uncertainty over all Lipschitz payoffs along an unquoted direction, whereas $\alpha + \frac{3}{2}$ in (7.5) concerns a fixed kink. The Lipschitz class includes oscillatory payoffs that can vary with the quote tolerance.

COROLLARY 7.12. *Assume (7.4) and i.i.d. quote locations with $Q \ge q_* Q_E$. With probability at least $1-\delta$, simultaneously for all feasible price vectors, the unrestricted admissible set satisfies $\operatorname{diam}_{d_D} \mathcal{K}_n \le C(\log(en/\delta)/n)^{1/(2d)}$ whenever $D \subseteq E$. If E has nonempty relative interior, $\pm\boldsymbol{e} \notin E$, and $D \subseteq E \cup \{\boldsymbol{e}\}$, the Sobolev-restricted admissible set satisfies $\operatorname{diam}_{d_D} \mathcal{K}_{n,\alpha,L} \le C(\log n)^{-(\alpha+1)}$ for $n \ge n_0(\delta)$.*

TABLE 1
Pricing moduli for different classes of claims.

Claims	Information	Modulus	Result
Static portfolios $c + \mathcal{A}^* m$	Any direction set E	$\le \|m\|\,\epsilon$	Lem. 6.2
Lipschitz payoffs on a quoted basket	$D \subseteq E$, all laws	Order $\epsilon^{1/2}$	Thm. 7.11(a)
A kink on an unquoted spread	Long-only; $\mathcal{F}_\alpha(L)$	$(\log(1/\epsilon))^{-(\alpha+3/2)}$	Thm. 7.8
All Lipschitz payoffs on that spread	Long-only; $\mathcal{F}_\alpha(L)$	$(\log(1/\epsilon))^{-(\alpha+1)}$	Thm. 7.11(b,c)
All Lipschitz payoffs (W_1)	Full directions; all laws	$\le C\epsilon^{2/(d+3)}$	Thm. 3.2

7.5. *Exact Quotes and Uniformity in the Noise Level.* The preceding minimax bound assumes fixed positive noise. Does the exchange price become easier to estimate when the quotes are exact? Under two-sided bounds on the sampling density, the answer is no: the same logarithmic rate holds uniformly down to zero noise. The key difficulty is not measurement noise but the information missing from finitely many long-only quotes.

Throughout this subsection E has nonempty relative interior and satisfies (7.4), $\pm\boldsymbol{e} \notin E$, and there is $\boldsymbol{w} \in S^{d-1}$ with $\boldsymbol{w} \cdot \boldsymbol{e} = 0$ and $\min_{\boldsymbol{u}\in E} \boldsymbol{u} \cdot \boldsymbol{w} > 0$. Derivatives in the direction $\boldsymbol{w}$ then leave the price of g_k unchanged and enter every quoted call price with the positive factor $\boldsymbol{u} \cdot \boldsymbol{w}$. For the long-only directions and $\boldsymbol{e} = (\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}$ one can take $\boldsymbol{w} = d^{-1/2}(1, \dots, 1)$, with $\min_{\boldsymbol{u}\in E} \boldsymbol{u} \cdot \boldsymbol{w} \geq d^{-1/2}$. We write $\mathfrak{R}_n^Q(g; \mathcal{C}, \varsigma)$ for the risk defined in Section 6.3, now displaying its dependence on the noise level and allowing $\varsigma = 0$.

THEOREM 7.13 (Pricing an unquoted kink uniformly in the noise level). *Let $q_* Q_E \leq Q \leq q^* Q_E$, $|k| < r$, $\alpha \geq 0$, L sufficiently large, and $\bar{\varsigma} < \infty$. There are $0 < c < C < \infty$ and n_0, independent of n and ς, such that for all $n \geq n_0$ and $0 \leq \varsigma \leq \bar{\varsigma}$,*

$$c(\log n)^{-(\alpha+\frac{3}{2})} \leq \mathfrak{R}_n^Q\big(g_k; \mathcal{C}_{\alpha,L}, \varsigma\big) \leq C(\log n)^{-(\alpha+\frac{3}{2})}.$$

Overall, long-only quotes identify the joint law, but identification alone gives no useful polynomial rate for an unquoted spread payoff. Theorem 7.13 shows that even exact quotes yield only $(\log n)^{-(\alpha+3/2)}$ minimax risk for the exchange price. Yet the same quotes estimate polynomial prices at the parametric rate. Adding the spread direction $(\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}$ makes all four payoffs in Example 7.9 Lipschitz stable. Thus pricing stability depends on a claim's direction relative to the quoted directions, not merely on quote precision.

8. Numerical Experiments. We illustrate three consequences of the theory in dimension $d = 2$ with $R = 1$: the benefit of residual correction, recovery under nonuniform designs, and the effect of adding a spread direction when pricing a claim. Call prices and Fourier moments are evaluated analytically, with Gaussian quote noise $N(0, \varsigma^2)$. The reference design is $Q_0 = \mathrm{Unif}(S^1) \otimes \mathrm{Unif}([-1, 1])$. We compute W_1 error by exact Euclidean transport between cell masses on an 81×81 grid. The recovery and calibration densities have the form

$$f_{\beta,\boldsymbol{c},A}(\boldsymbol{x}) = \frac{\beta+1}{\pi|\det A|}\big(1 - \|A^{-1}(\boldsymbol{x} - \boldsymbol{c})\|^2\big)_+^{\beta}, \qquad A = \mathsf{U}_\varphi \,\mathrm{diag}(a, b),$$

where U_φ is the rotation through angle φ. All three laws use $\beta = 0.6$ and have H^1 densities supported in B_r with $r = 0.85$.

- *Ellipse.* A single density with $\boldsymbol{c} = (0.08, -0.04)$ and $(a, b, \varphi) = (0.64, 0.38, 0.61)$.
- *Mixture.* Two components with weights $(0.58, 0.42)$; centers $(-0.25, -0.13)$ and $(0.27, 0.18)$; (a, b, φ) equal to $(0.30, 0.21, -0.35)$ and $(0.31, 0.40, 0.80)$, respectively.
- *Calibration.* A mixture with weights $(0.65, 0.35)$; centers $(-0.12, 0.13)$ and $(0.24, -0.14)$; (a, b, φ) equal to $(0.46, 0.36, -0.45)$ and $(0.26, 0.32, 0.20)$, respectively.

8.1. *Residual correction.* For $\alpha = 1$, the direct Fourier bound in Proposition 5.4 gives $n^{-2/9}$, whereas the minimax benchmark in Theorem 5.1 is $n^{-1} + (\varsigma^2/n)^{2/9}$. Residual correction is designed to recover the faster exact-quote behavior. Figure 1 compares the two estimators across the exact-quote and noisy regimes. We use torus side 5, noise levels $\varsigma \in \{0, 0.05\}$, and cutoffs $M_{\mathrm{dir}} = 1.4\, n^{1/9}$ and $M_{\mathrm{cv}} = \min\{0.1\sqrt{n},\, 0.8(n/\varsigma^2)^{1/9}\}$, with the second term interpreted as infinity when $\varsigma = 0$. The pilot is a local quartic fit on the 60 nearest quotes. We average the two estimates obtained by exchanging the sample halves. After a smooth support cutoff and mass correction, we project onto probability histograms supported in $B_{0.85}$ in a

discrete H^{-1} norm, using $\|\cdot\|_{\mathrm{Lip}^*} \le C\|\cdot\|_{H^{-1}}$. The spatial grid has $J = \max(64, 8\lceil M\rceil)$ cells across $[-1, 1]$. Results are averaged over 100 replications.

With exact quotes, residual correction reduces mean W_1 error by about two orders of magnitude at the largest sample size. The corrected curves decrease faster than the n^{-1} reference line. This does not contradict the worst-case minimax result, since the simulated densities belong to H^α for every $\alpha < 1.1$. At $\varsigma = 0.05$, Theorem 5.1 predicts the slower rate $n^{-2/9}$, which closely matches direct Fourier recovery. For $n \ge 2048$, the corrected curves have empirical exponents 0.32 and 0.27 for the ellipse and mixture. At the largest sample size, residual correction reduces the ellipse and mixture errors by factors of about 4.5 and 2.4, respectively.

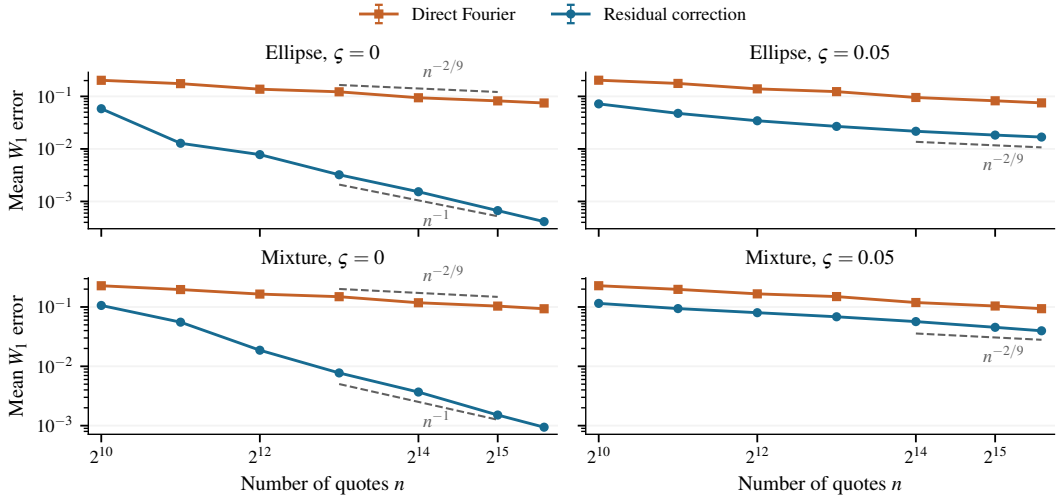


FIG 1. *Mean W_1 error of the probability estimates. Dashed lines give theoretical rates.*

8.2. *Recovery under general designs.* Corollary 5.8 extends the rate $n^{-1} + (\varsigma^2/n)^{2/9}$ to designs with densities bounded above and away from zero relative to Q_0. We compare recovery under Q_0 and $Q_1 = \frac{1}{2}Q_0 + \frac{1}{2}Q_c$, constructing $\widehat{\mu}_h$ by smoothing and projection onto a finite mixture dictionary. The estimator uses the observed quote locations, not the design density. Its bandwidth is selected by comparing candidate reconstructions, with a threshold motivated by the stochastic term in (5.16). Under Q_c, the angle has a von Mises distribution with mean $\pi/4$ and concentration 2, and the strike is $2U - 1$ for an independent $U \sim \mathrm{Beta}(2, 2)$.

For bandwidth h, we use a local quartic pilot with $\max(60, \lceil 4nh^2\rceil)$ neighbors. Its values are truncated to $[0, 2R]$ as in Lemma 5.6. For the three sample sizes in Table 2, the pricing grids have 48×81, 64×113 and 80×129 equally spaced direction–strike nodes, respectively. We use an angular Fourier and strike cosine basis for smoothing. Projection minimizes the discrete $H^{3/2}$ distance to the smoothed surface over convex mixtures of $f_{0.6,\boldsymbol{c},aI}$. The dictionary uses a from $(0.12, 0.22, 0.36, 0.54, 0.72)$ on square grids of spacing $(0.12, 0.16, 0.20, 0.20, 0.20)$, restricted to component supports in $B_{0.9}$. At sample size n, we add radius $a_n = 0.12(n/2048)^{-1/4}$ on a grid of spacing $a_n/2$.

We select the largest h in $\mathcal{H}_n = \{2^{j/4} : j \in \mathbb{Z}\} \cap [n^{-1/2}, 1/4]$ satisfying $W_1(\widehat{\mu}_h, \widehat{\mu}_{h'}) \le 0.35\,\widehat{\varsigma} n^{-1/2}(h')^{-5/2}$ for every $h' < h$ in $\mathcal{H}_n$. The common multiplier was chosen from a fixed grid by mean W_1 on 16 independent calibration datasets, using the calibration law specified above and both designs at $n = 4096$ and 8192. We estimate ς by the root mean square of residuals from leave-one-out quartic fits on the 60 nearest neighbors, after dividing

each residual by $(1+\sum_j a_{ij}^2)^{1/2}$, where a_{ij} are the local regression weights. Table 2 reports results from 50 replications. At each noise level $\varsigma \in \{0.005, 0.05\}$, the oracle error is the mean of the smallest candidate W_1 error on each replication.

Increasing n from $2,048$ to $16,384$ reduces the exact-quote errors by factors of 6.5 to 9.1, close to the factor 8 predicted by n^{-1}. At $\varsigma = 0.005$ and 0.05, the factors are 1.7 to 1.8 and 1.8 to 1.9, respectively, compared with $8^{2/9} \approx 1.6$ from the noise term. The mean errors under Q_1 differ from those under Q_0 by at most 26%. At the largest sample size, the selected errors are at most 12% above the oracle errors.

TABLE 2
Recovery under the two designs. Entries are 10^3 times mean W_1 error. Parentheses give Monte Carlo standard errors.

Law	Design	n	Selected, ς			Oracle, ς	
			0	0.005	0.05	0.005	0.05
Ellipse	Q_0	2,048	2.24 (0.02)	11.79 (0.19)	34.27 (1.20)	10.95	31.82
		8,192	0.56 (<0.01)	8.22 (0.14)	23.04 (0.77)	7.75	20.82
		16,384	0.35 (<0.01)	6.91 (0.11)	19.48 (0.52)	6.52	17.50
	Q_1	2,048	2.60 (0.03)	11.58 (0.25)	37.97 (1.26)	10.62	33.06
		8,192	0.62 (<0.01)	8.60 (0.18)	25.26 (0.92)	7.93	22.83
		16,384	0.37 (<0.01)	6.81 (0.10)	19.95 (0.50)	6.30	18.58
Mixture	Q_0	2,048	5.62 (0.04)	16.96 (0.20)	65.02 (1.23)	16.19	58.99
		8,192	1.22 (<0.01)	11.32 (0.23)	45.68 (0.80)	10.85	39.85
		16,384	0.72 (<0.01)	9.52 (0.11)	35.28 (0.86)	9.27	32.73
	Q_1	2,048	7.06 (0.10)	18.04 (0.25)	66.64 (1.39)	16.36	57.29
		8,192	1.48 (0.01)	11.62 (0.21)	47.54 (0.87)	11.30	43.05
		16,384	0.77 (<0.01)	10.12 (0.08)	37.34 (0.80)	9.44	33.93

8.3. *Quoted directions and claim prices.* We compare $E = \{(\cos\theta, \sin\theta) : 0 \le \theta \le \pi/2\}$ with $E \cup \{\boldsymbol{e}\}$, where $\boldsymbol{e} = (1,-1)/\sqrt{2}$. The claims are the squared spread $(\boldsymbol{e}\cdot\boldsymbol{x})^2$ and the exchange option $g_+(\boldsymbol{x}) = (\boldsymbol{e}\cdot\boldsymbol{x})_+$. We take $\mathcal{C}$ to be the convex hull of 299 component laws with radial densities $f_{3,\boldsymbol{c},aI}$, $a \in \{0.12, 0.24, 0.40\}$, and centers on a grid of spacing 0.12, retaining supports inside $B_{0.92}$.

Table 3 bounds the uniform-quote modulus $\varpi_g(\epsilon)$ on $\mathcal{C}$. The dual of a linear program on a direction–strike grid certifies upper bounds through the finite-class form of the replication bound in Lemma 6.2. For lower bounds, we mix both measures in the LP solution with their midpoint using the same weight. Derivative bounds between grid points determine the weight needed to keep the absolute quote difference at most ϵ over every quoted direction and strike.

TABLE 3
Lower and upper bounds for pricing moduli, multiplied by 10^3.

ϵ	Squared spread		Exchange option	
	Long-only	With spread	Long-only	With spread
10^{-2}	$[84.79, 86.53]$	$[41.39, 43.72]$	$[132.4, 135.2]$	$[9.638, 10]$
10^{-3}	$[8.829, 9.008]$	$[4.39, 4.48]$	$[67.59, 68.98]$	$[0.9488, 1]$
10^{-4}	$[0.9086, 0.9328]$	$[0.4603, 0.4709]$	$[28.17, 28.79]$	$[0.09417, 0.1]$

The squared-spread modulus decreases nearly linearly with ϵ under both designs, as expected from polynomial replication. By contrast, the exchange modulus decreases much more slowly with long-only quotes. Once the spread direction is included, the exchange payoff itself is quoted, and Lemma 6.2 gives the upper bound ϵ. The numerical intervals then decrease linearly with ϵ. These computations illustrate the directional effect of Theorem 7.8, whose exact logarithmic order concerns the full Sobolev class, not the finite mixture class used here.

9. Conclusion and Open Problems. Random basket quotes link distributional recovery and the estimation of individual claim prices through one linear inverse operator. Its negative-Sobolev geometry gives sharp Wasserstein recovery rates on Sobolev density classes, uniformly from exact quotes to bounded noise levels. A bounded systematic pricing bias creates an additional error that does not vanish with more observations. For a specified claim, a modulus of continuity and its static-replication dual characterize minimax estimation at fixed positive noise. Under long-only directions, the joint law is identified, yet the exchange price is recovered only at a logarithmic rate, even from exact quotes.

Arbitrary laws. For exact quotes without a density assumption, Theorem 5.11 leaves a gap between the exponents $1/d$ and $(d+3)/(d(2d+3))$. A possible route to closing it is a pricing-surface pilot with expected squared $L^2(Q_0)$ error bounded by $C(m^{-1-3/d} + \varsigma^2)$ uniformly over $\mathcal{P}(B_r)$. If such a pilot is available, Proposition 5.7, with $M \asymp \min\{n^{1/d}, (n/\varsigma^2)^{1/(2d+3)}\}$, would give the matching rate $n^{-1/d} + (\varsigma^2/n)^{1/(2d+3)}$.

Post-design uniformity. Theorem 4.3 controls all Sobolev-restricted quote fibers on a common random-design event, with a logarithmic factor. The minimax expected-risk rate of Theorem 5.1 has no such factor. Whether it is necessary for the post-design supremum remains open; the two criteria are distinguished in Section 4.3.

Adaptation. The bandwidth achieving the upper bound in Theorem 5.1 depends on α, L, ς, and ϵ. Section 8 uses a Lepski-type selection rule, but does not establish an adaptive minimax theorem. Proving such a result, even up to logarithmic factors for $\epsilon = 0$, and allowing heteroskedastic or price-dependent errors are natural next steps.

REFERENCES

[1] AÏT-SAHALIA, Y. and DUARTE, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics* **116** 9–47.

[2] AÏT-SAHALIA, Y. and LO, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* **53** 499–547.

[3] BÉLISLE, C., MASSÉ, J. C. and RANSFORD, T. (1997). When is a probability measure determined by infinitely many projections? *Annals of Probability* **25** 767–786.

[4] BERTSIMAS, D. and POPESCU, I. (2002). On the relation between option and stock prices: A convex optimization approach. *Operations Research* **50** 358–374.

[5] BLANCHARD, G. and MÜCKE, N. (2018). Optimal rates for regularization of statistical inverse learning problems. *Foundations of Computational Mathematics* **18** 971–1013.

[6] BOSSU, S., CRÉPEY, S. and NGUYEN, H. D. (2025). Spanning multi-asset payoffs with ReLUs. *Mathematical Finance* **35** 682–707.

[7] CAI, T. T. and LOW, M. G. (2004). Minimax estimation of linear functionals over nonconvex parameter spaces. *Annals of Statistics* **32** 552–576.

[8] CAILLERIE, C., CHAZAL, F., DEDECKER, J. and MICHEL, B. (2011). Deconvolution for the Wasserstein metric and geometric inference. *Electronic Journal of Statistics* **5** 1394–1423.

[9] CAVALIER, L. (2000). Efficient estimation of a density in a problem of tomography. *Annals of Statistics* **28** 630–647.

[10] CHOLAQUIDIS, A. and HERNANDEZ-BANADIK, M. (2026). Weak convergence from projections on a positive-measure set of directions. arXiv:2604.10310.

[11] CRAMÉR, H. and WOLD, H. (1936). Some theorems on distribution functions. *Journal of the London Mathematical Society* **11** 290–294.
[12] CUESTA-ALBERTOS, J. A., FRAIMAN, R. and RANSFORD, T. (2007). A sharp form of the Cramér–Wold theorem. *Journal of Theoretical Probability* **20** 201–209.
[13] D'ASPREMONT, A. and EL GHAOUI, L. (2006). Static arbitrage bounds on basket option prices. *Mathematical Programming* **106** 467–489.
[14] DEDECKER, J. and MICHEL, B. (2013). Minimax rates of convergence for Wasserstein deconvolution with supersmooth errors in any dimension. *Journal of Multivariate Analysis* **122** 278–291.
[15] DESHPANDE, I., HU, Y. T., SUN, R., PYRROS, A., SIDDIQUI, N., KOYEJO, S., ZHAO, Z., FORSYTH, D. and SCHWING, A. G. (2019). Max-sliced Wasserstein distance and its use for GANs. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
[16] DEVORE, R., NOWAK, R. D., PARHI, R., PETROVA, G. and SIEGEL, J. W. (2025). Optimal recovery meets minimax estimation. arXiv:2502.17671.
[17] DONOHO, D. L. (1994). Statistical estimation and optimal recovery. *Annals of Statistics* **22** 238–270.
[18] DONOHO, D. L. and LIU, R. C. (1991). Geometrizing rates of convergence, II. *Annals of Statistics* **19** 633–667.
[19] FRIKEL, J. and QUINTO, E. T. (2013). Characterization and reduction of artifacts in limited angle tomography. *Inverse Problems* **29** 125007.
[20] GOLDENSHLUGER, A. and PEREVERZEV, S. V. (2000). Adaptive estimation of linear functionals in Hilbert scales from indirect white noise observations. *Probability Theory and Related Fields* **118** 169–186.
[21] HINRICHS, A., KRIEG, D., NOVAK, E., PROCHNO, J. and ULLRICH, M. (2020). On the power of random information. In *Multivariate Algorithms and Information-Based Complexity* (F. J. Hickernell and P. Kritzer, eds.) 43–64. Berlin: De Gruyter.
[22] HOBSON, D., LAURENCE, P. and WANG, T. H. (2005). Static-arbitrage upper bounds for the prices of basket options. *Quantitative Finance* **5** 329–342.
[23] JOHNSTONE, I. M. and SILVERMAN, B. W. (1990). Speed of estimation in positron emission tomography and related inverse problems. *Annals of Statistics* **18** 251–280.
[24] JUDITSKY, A. B. and NEMIROVSKI, A. S. (2009). Nonparametric estimation by convex programming. *Annals of Statistics* **37** 2278–2300.
[25] KLEMELÄ, J. and MAMMEN, E. (2010). Empirical risk minimization in inverse problems. *Annals of Statistics* **38** 482–511.
[26] KOCH, H., RÜLAND, A. and SALO, M. (2021). On instability mechanisms for inverse problems. *Ars Inveniendi Analytica*.
[27] KOROSTELEV, A. P. and TSYBAKOV, A. B. (1993). *Minimax Theory of Image Reconstruction*. *Lecture Notes in Statistics* 82. New York: Springer.
[28] KRIEG, D., NOVAK, E. and SONNLEITNER, M. (2022). Recovery of Sobolev functions restricted to iid sampling. *Mathematics of Computation* **91** 2715–2738.
[29] KRIEG, D. and SONNLEITNER, M. (2024). Random points are optimal for the approximation of Sobolev functions. *IMA Journal of Numerical Analysis* **44** 1346–1371.
[30] MICCHELLI, C. A. and RIVLIN, T. J. (1977). A survey of optimal recovery. In *Optimal Estimation in Approximation Theory* (C. A. Micchelli and T. J. Rivlin, eds.) New York: Plenum.
[31] MOLCHANOV, I. and SCHMUTZ, M. (2014). Multiasset derivatives and joint distributions of asset prices. In *Inspired by Finance* (Y. Kabanov, M. Rutkowski and T. Zariphopoulou, eds.) Springer.
[32] NEUFELD, A., PAPAPANTOLEON, A. and XIANG, Q. (2023). Model-free bounds for multi-asset options using option-implied information and their exact computation. *Management Science* **69** 2051–2068.
[33] NILES-WEED, J. and BERTHET, Q. (2022). Minimax estimation of smooth densities in Wasserstein distance. *Annals of Statistics* **50** 1519–1540.
[34] NOVAK, E. and WOŹNIAKOWSKI, H. (2008). *Tractability of Multivariate Problems, Volume I: Linear Information*. *EMS Tracts in Mathematics* 6. Zürich: European Mathematical Society.
[35] PANARETOS, V. M. (2009). On random tomography with unobservable projection angles. *Annals of Statistics* **37** 3272–3306.
[36] PATHAK, R., WAINWRIGHT, M. J. and XIAO, L. (2024). Noisy recovery from random linear observations: Sharp minimax rates under elliptical constraints. *Annals of Statistics* **52** 2816–2850.
[37] QUINTO, E. T. (1993). Singularities of the X-ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$. *SIAM Journal on Mathematical Analysis* **24** 1215–1225.
[38] TANKOV, P. (2011). Improved Fréchet bounds and model-free pricing of multi-asset options. *Journal of Applied Probability* **48** 389–403.
[39] WEED, J. and BACH, F. (2019). Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** 2620–2648.

APPENDIX A: PROOFS FOR SECTION 2

We use $H^2(P,P') = \int(\sqrt{\mathrm{d}P} - \sqrt{\mathrm{d}P'})^2$ for the squared Hellinger distance.

LEMMA A.1 (Hellinger bound for quote experiments). *Let Q be a design, $\varsigma > 0$, and let $P_\mu^{(n)}$ denote the law of n i.i.d. triples $(Z_i, C_\mu(Z_i) + \varsigma\xi_i)$ with $Z_i \sim Q$. If ξ_1 has a density with finite Fisher information I_ξ, then*

$$\mathrm{TV}\left(P_\mu^{(n)}, P_\nu^{(n)}\right)^2 \le H^2\left(P_\mu^{(n)}, P_\nu^{(n)}\right) \le \frac{nI_\xi}{4\varsigma^2}\,\|C_\mu - C_\nu\|_{L^2(Q)}^2 .$$

PROOF. Since $\sqrt{p_\xi}$ is absolutely continuous with $\int((\sqrt{p_\xi})')^2 = I_\xi/4$, the Cauchy–Schwarz inequality gives $\int(\sqrt{p_\xi(y-a)} - \sqrt{p_\xi(y)})^2\,\mathrm{d}y \le a^2 I_\xi/4$ for every shift a. Conditioning on Z_1 and rescaling by ς bounds the squared Hellinger distance of one triple by $I_\xi\,\|C_\mu - C_\nu\|_{L^2(Q)}^2/(4\varsigma^2)$. Since Hellinger affinities multiply over independent coordinates, H^2 is subadditive, and $\mathrm{TV} \le H$. □

PROOF OF THEOREM 2.2. The compact-support argument follows the homogeneous-polynomial criterion of [5]. Suppose (ii) holds and $C_\mu = C_\nu$ on $E \times [-R,R]$. The projected measures agree in each direction in E, by the strike-extension argument preceding the theorem. For every integer $m \ge 0$, the homogeneous polynomial

$$p_m(\boldsymbol{u}) = \int (\boldsymbol{u}\cdot\boldsymbol{x})^m\,(\mu-\nu)(\mathrm{d}\boldsymbol{x})$$

vanishes on E and therefore vanishes identically. Expansion in monomials shows that all mixed moments of $\mu - \nu$ vanish. Polynomials are uniformly dense in the continuous functions on B_R, by the real Stone–Weierstrass theorem. Thus $\mu = \nu$.

Conversely, suppose a nonzero homogeneous polynomial p vanishes on E. Multiplication by a nonzero homogeneous factor, if necessary, ensures that its degree is at least two. Let $g \in C_c^\infty$ be a nonzero real function supported in a small cube inside B_R and define $\eta = p(\partial)g$. The function η is not identically zero: its Fourier transform is $p(\mathrm{i}\boldsymbol{\xi})\widehat{g}(\boldsymbol{\xi})$, and a nonzero polynomial cannot vanish on an open set on which $\widehat{g}$ is nonzero. Integration by parts shows that η has zero mass and first moments. For each $\boldsymbol{u} \in E$ and $t \in \mathbb{R}$,

$$\widehat{\eta}(t\boldsymbol{u}) = \mathrm{i}^{\deg p} t^{\deg p} p(\boldsymbol{u})\widehat{g}(t\boldsymbol{u}) = 0.$$

Therefore every corresponding projected signed measure is zero. Choose a smooth probability density f_0 positive on $\operatorname{supp}\eta$ and a sufficiently small $\epsilon > 0$. The distinct densities $f_0 \pm \epsilon\eta$ have identical projected laws, hence identical call prices, in all directions of E.

A homogeneous polynomial vanishing on a relatively open sphere patch vanishes on the associated open cone and hence is identically zero. For finite $E = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_m\}$, choose nonzero $\boldsymbol{v}_j$ perpendicular to $\boldsymbol{u}_j$ and use $p(\boldsymbol{\xi}) = \prod_j (\boldsymbol{v}_j \cdot \boldsymbol{\xi})$. This polynomial is nonzero and vanishes on E. In dimension two, a nonzero homogeneous polynomial has finitely many zeros on the real projective line, and an infinite set of distinct unoriented directions therefore satisfies (ii). □

PROOF OF THEOREM 2.3. Choose a countable base for the compact metric space $\operatorname{supp} Q$. Every nonempty relative open set in this base has positive Q probability and is visited almost surely by the i.i.d. sample. Taking the intersection of these probability-one events shows that $\{Z_i : i \ge 1\}$ is dense in $\operatorname{supp} Q$. Continuity in (2.3) gives

$$\bigcap_{n\ge 1} \mathcal{K}_n(\mu) = \mathcal{K}_\infty(\mu; Q).$$

Suppose the Hausdorff convergence fails on this event. Since $\mathcal{K}_\infty \subseteq \mathcal{K}_n$, there are $\epsilon > 0$, integers $n_j \uparrow \infty$, and $\nu_j \in \mathcal{K}_{n_j}(\mu)$ with $\operatorname{dist}_{W_1}(\nu_j, \mathcal{K}_\infty) \geq \epsilon$. A subsequence converges in W_1 to some ν. For each m, all sufficiently late terms belong to the closed set $\mathcal{K}_m(\mu)$, hence $\nu \in \mathcal{K}_m(\mu)$. Thus $\nu \in \mathcal{K}_\infty$, which contradicts the displayed distance bound. Under identification, $\mathcal{K}_\infty = \{\mu\}$. □

APPENDIX B: PROOFS FOR SECTION 3

LEMMA B.1. *If $\eta(B_R) = 0$ and $\boldsymbol{a} = \int \boldsymbol{x}\, \eta(\mathrm{d}\boldsymbol{x})$, then*

$$\mathcal{A}\eta(\boldsymbol{u}, k) - \mathcal{A}\eta(-\boldsymbol{u}, -k) = \boldsymbol{u} \cdot \boldsymbol{a}, \qquad \boldsymbol{a} = 2d \int_{\mathcal{Z}} \boldsymbol{u}\, \mathcal{A}\eta(\boldsymbol{u}, k)\, \mathrm{d}Q_0. \tag{B.1}$$

Moreover, $|\boldsymbol{a}| \leq 2\sqrt{d}\, \|\mathcal{A}\eta\|_{L^2(Q_0)}$.

PROOF. The identity $t_+ - (-t)_+ = t$ gives the first equality. The transformation $(\boldsymbol{u}, k) \mapsto (-\boldsymbol{u}, -k)$ preserves Q_0, and $\int \boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}\, \mathrm{d}\sigma = I_d/d$, giving the second. The odd part of $\mathcal{A}\eta$ under this transformation is $\boldsymbol{u} \cdot \boldsymbol{a}/2$. Its $L^2(Q_0)$ norm is $|\boldsymbol{a}|/(2\sqrt{d})$ and cannot exceed the norm of $\mathcal{A}\eta$. □

Fix a smooth probability density ϱ supported strictly inside B_R and define $\tau_j = -\partial_j \varrho$, $1 \leq j \leq d$. Integration by parts gives

$$\int \tau_j\, \mathrm{d}\boldsymbol{x} = 0, \qquad \int x_l \tau_j(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \mathbf{1}_{\{j=l\}}. \tag{B.2}$$

Let η be a finite signed measure on B_R with $\eta(B_R) = 0$, and let $\boldsymbol{a} = \int \boldsymbol{x}\, \eta(\mathrm{d}\boldsymbol{x})$. Then

$$\eta_0 = \eta - \sum_{j=1}^{d} a_j \tau_j\, \mathrm{d}\boldsymbol{x} \tag{B.3}$$

has zero mass and zero first moments. If $\eta = \mu - \nu$ with $\mu, \nu \in \mathcal{P}(B_R)$ and $D = \|\mathcal{A}\eta\|_{L^2(Q_0)}$, then

$$\|\eta_0\|_{\mathrm{TV}} \leq C_{d,R}, \qquad \|\mathcal{A}\eta_0\|_{L^2(Q_0)} \leq C_{d,R} D. \tag{B.4}$$

These follow from the bound $|\boldsymbol{a}| \leq 2\sqrt{d}\, D$ of Lemma B.1, from $D \leq 2R$, and from the smooth functions τ_j.

PROOF OF THEOREM 3.1. For each $\boldsymbol{u}$, let $\eta_{\boldsymbol{u}}$ be the pushforward of η under $\boldsymbol{x} \mapsto \boldsymbol{u} \cdot \boldsymbol{x}$. The zero mass and first moment imply $\mathcal{A}\eta(\boldsymbol{u}, k) = 0$ for $|k| \geq R$. Its second distributional derivative in k is $\eta_{\boldsymbol{u}}$. Fourier transformation in k gives

$$\widehat{\mathcal{A}\eta}(\boldsymbol{u}, t) = -\frac{\widehat{\eta}(t\boldsymbol{u})}{t^2}, \qquad t \neq 0. \tag{B.5}$$

Since $\widehat{\eta}(\boldsymbol{\xi}) = O(|\boldsymbol{\xi}|^2)$ near zero, the expressions below are integrable there. At infinity, boundedness of $\widehat{\eta}$ gives integrability. Plancherel's identity and polar coordinates yield

(B.6)

$$\|\mathcal{A}\eta\|^2_{L^2(Q_0)} = \frac{1}{4\pi R} \int_{S^{d-1}} \int_{\mathbb{R}} \frac{|\widehat{\eta}(t\boldsymbol{u})|^2}{t^4}\, \mathrm{d}t\, \mathrm{d}\sigma(\boldsymbol{u}) = \frac{1}{2\pi R \omega_{d-1}} \int_{\mathbb{R}^d} |\widehat{\eta}(\boldsymbol{\xi})|^2\, |\boldsymbol{\xi}|^{-d-3}\, \mathrm{d}\boldsymbol{\xi}.$$

This proves (3.1).

For $a, b \in [-R, R]$, direct integration gives

$$\int_{-R}^{R} (a-k)_+(b-k)_+ \,\mathrm{d}k = \frac{|a-b|^3}{12} + \frac{R^3}{3} + \frac{R^2}{2}(a+b) + Rab - \frac{a^3+b^3}{12} + \frac{a^2b+ab^2}{4}.$$

Substitute $a = \boldsymbol{u}\cdot\boldsymbol{x}$ and $b = \boldsymbol{u}\cdot\boldsymbol{y}$ and integrate against $\eta\otimes\eta$. Every polynomial term vanishes because each contains either a factor of total mass or a factor of first moment. Thus

$$\int_{-R}^{R} |\mathcal{A}\eta(\boldsymbol{u},k)|^2 \,\mathrm{d}k = \frac{1}{12}\iint |\boldsymbol{u}\cdot(\boldsymbol{x}-\boldsymbol{y})|^3 \,\eta(\mathrm{d}\boldsymbol{x})\eta(\mathrm{d}\boldsymbol{y}).$$

Rotation invariance gives $\int |\boldsymbol{u}\cdot\boldsymbol{x}|^3 \,\mathrm{d}\sigma = m_{3,d}\,|\boldsymbol{x}|^3$. The value of $m_{3,d}$ follows by integrating the density proportional to $(1-t^2)^{(d-3)/2}$ of the first coordinate of a uniform direction. Division by $2R$ proves (3.2). □

PROOF OF THEOREM 3.2. Let $\eta = \mu - \nu$ and form η_0 as in (B.3). A 1-Lipschitz function on B_R, normalized to vanish at the origin, has a compactly supported Lipschitz extension F with

$$\|F\|_{H^1} + \|F\|_\infty + \mathrm{Lip}(F) \le C_{d,R}. \tag{B.7}$$

Let $F_\epsilon = F * \phi_\epsilon$, where ϕ is a smooth compactly supported probability density. Then

$$\|F - F_\epsilon\|_\infty \le C\epsilon, \qquad \|F_\epsilon\|_{\dot H^s} \le C\epsilon^{-(s-1)}\,\|F\|_{H^1}.$$

For the second inequality, use $\sup_{\boldsymbol{\xi}} |\boldsymbol{\xi}|^{s-1}\left|\widehat{\phi}(\epsilon\boldsymbol{\xi})\right| \le C\epsilon^{-(s-1)}$. Theorem 3.1 and (B.4) give

$$\begin{aligned}\left|\int F\,\mathrm{d}\eta\right| &\le \|\eta_0\|_{\mathrm{TV}}\,\|F-F_\epsilon\|_\infty + \|\eta_0\|_{\dot H^{-s}}\,\|F_\epsilon\|_{\dot H^s} + \sum_j |a_j|\left|\int F\tau_j\,\mathrm{d}\boldsymbol{x}\right| \\ &\le C\left(\epsilon + D\epsilon^{-(s-1)} + D\right).\end{aligned}$$

Choose $\epsilon = D^{1/s}$ for $0 < D \le 1$. The cases $D = 0$ and $D > 1$ follow by a limit and the bound $W_1 \le 2R$, respectively. Taking the supremum over Lipschitz tests proves (3.3).

For the density assertion, $\|\eta_0\|_{H^\alpha} \le C(1+L)$ and $\|\eta_0\|_{H^{-s}} \le \|\eta_0\|_{\dot H^{-s}} \le CD$. Weighted Hölder's inequality in the Fourier integral gives

$$\|\eta_0\|_{H^{-1}} \le \|\eta_0\|_{H^{-s}}^{(\alpha+1)/(\alpha+s)}\,\|\eta_0\|_{H^\alpha}^{(s-1)/(\alpha+s)}. \tag{B.8}$$

Pairing with (B.7) and adding the mean correction proves (3.4).

For sharpness, take a nonnegative $\zeta_0 \in C_c^\infty(\mathbb{R}^d)$ supported in a small cube inside B_R with positive integral. Choose a smooth probability density f_0 that is bounded below by a positive constant on that support. For integers $N \ge 1$, define

$$H_N(\boldsymbol{x}) = bN^{-\alpha-2}\partial_1^2\left(\zeta_0(\boldsymbol{x})\cos(Nx_1)\right), \qquad f_N^{\pm} = f_0 \pm H_N. \tag{B.9}$$

For a sufficiently small $b > 0$, these are nonnegative probability densities. Integration by parts gives zero mass and first moments for H_N. The product rule gives $\|H_N\|_\infty \le CbN^{-\alpha}$, and Fourier translation gives $\|H_N\|_{H^\alpha} \le Cb$.

The Fourier transform is

$$\widehat{H}_N(\boldsymbol{\xi}) = -\frac{b}{2}N^{-\alpha-2}\xi_1^2\left(\widehat{\zeta}_0(\boldsymbol{\xi}-N\boldsymbol{e}_1) + \widehat{\zeta}_0(\boldsymbol{\xi}+N\boldsymbol{e}_1)\right).$$

On $|\boldsymbol{\xi}| \ge N/2$, the weighted integral in (3.1) is bounded by $Cb^2N^{-2\alpha-2s}$, since $4-2s = 1-d \le 0$ and the translated transforms have uniformly bounded L^2 norms. On $|\boldsymbol{\xi}| < N/2$,

Schwartz decay bounds each translated transform by $C_P N^{-P}$ for every P; the radial integral of $|\boldsymbol{\xi}|^{4-2s}$ over this ball is $O(N)$. Choosing $P \geq d/2$ gives the same upper bound. On a small ball about $N\boldsymbol{e}_1$, continuity and $\widehat{\zeta}_0(0) > 0$ give the matching lower bound. Therefore

$$\|\mathcal{A}H_N\|_{L^2(Q_0)} \asymp N^{-\alpha-s}. \tag{B.10}$$

The function $F_N(\boldsymbol{x}) = N^{-1}\cos(Nx_1)$ is 1-Lipschitz, and

$$\left| \int F_N H_N \, \mathrm{d}\boldsymbol{x} \right| = bN^{-\alpha-1} \int \zeta_0(\boldsymbol{x}) \cos^2(Nx_1) \, \mathrm{d}\boldsymbol{x} \geq cN^{-\alpha-1}$$

for all sufficiently large N. A single integration by parts against any Lipschitz test also gives the upper bound $CN^{-\alpha-1}$. Thus

$$W_1(f_N^+ \, \mathrm{d}\boldsymbol{x}, f_N^- \, \mathrm{d}\boldsymbol{x}) \asymp N^{-\alpha-1}.$$

Together with (B.10), this prevents an increase of the exponent in (3.4). The case $\alpha = 0$ proves sharpness of (3.3). □

PROOF OF LEMMA 3.3. Choose the functions in (B.2) supported strictly inside B_r. For $\eta_0 = \eta - \sum_j (\int x_j \eta \, \mathrm{d}\boldsymbol{x}) \tau_j$, both mass and first moments vanish. The moment functionals are bounded on functions supported on B_r in every Sobolev norm, by pairing with smooth cutoffs times $1, x_1, \ldots, x_d$. The finite sum removed from η has a smooth pricing surface on $\mathcal{Z}$. It suffices to estimate $\mathcal{A}\eta_0$, extended by zero outside $[-r, r]$ in the strike coordinate.

Begin with a smooth compactly supported function v with zero mass and first moments. Equation (B.5) applies, and the estimates below extend the operator by density. Let m be a nonnegative integer. The product Sobolev H^m norm on $S^{d-1} \times \mathbb{R}$ is bounded by the L^2 norms of derivatives obtained from at most m strike derivatives and smooth angular vector fields. Applying q strike derivatives and j angular derivatives to (B.5), with $q + j \leq m$, gives a finite sum of terms

$$b(\boldsymbol{u}) t^{q+l-2} \widehat{p_l v}(t\boldsymbol{u}), \qquad 0 \leq l \leq j, \tag{B.11}$$

where b is a smooth function on the sphere and p_l is a polynomial of degree at most l. Terms arising from differentiation of angular coefficients have the same form with a smaller l. On $|t| \geq 1$, polar coordinates bound the squared L^2 norm of (B.11) by

$$C \int_{|\boldsymbol{\xi}| \geq 1} |\widehat{p_l v}(\boldsymbol{\xi})|^2 \, |\boldsymbol{\xi}|^{2(q+l)-d-3} \, \mathrm{d}\boldsymbol{\xi} \leq C \, \|p_l v\|_{H^{m-s}}^2 \leq C \, \|v\|_{H^{m-s}}^2 . \tag{B.12}$$

For the last step, multiply p_l by a smooth cutoff equal to one on the common compact support; multiplication by this function is bounded on each Sobolev space in question.

On $|t| \leq 1$, use moment cancellation before differentiating:

$$\frac{\widehat{v}(t\boldsymbol{u})}{t^2} = -\int_0^1 (1-a) \left\langle v, (\boldsymbol{u} \cdot \boldsymbol{x})^2 e^{-\mathrm{i}at\boldsymbol{u}\cdot\boldsymbol{x}} \right\rangle \mathrm{d}a. \tag{B.13}$$

A cutoff on the compact support is understood in this pairing. Every required angular derivative of the test function, for $|t| \leq 1$, is bounded in H^{s-m} uniformly in $a, t, \boldsymbol{u}$. Strike derivatives in the original variable multiply by t^q after Fourier transformation. Hence the low-frequency contribution to the H^m norm is bounded by $C \, \|v\|_{H^{m-s}}$ as well. Combining the two regions yields

$$\|\mathcal{A}v\|_{H^m(S^{d-1}\times\mathbb{R})} \leq C_m \, \|v\|_{H^{m-s}(\mathbb{R}^d)} . \tag{B.14}$$

Interpolation is carried out on unconstrained spaces. Choose a smooth cutoff χ with compact support and equal to one on B_r, and a smooth function τ_0 with mass one and mean zero. Define the bounded finite-rank correction

$$\Pi w = w - \left(\int w \, \mathrm{d}\boldsymbol{x} \right) \tau_0 - \sum_j \left(\int x_j w \, \mathrm{d}\boldsymbol{x} \right) \tau_j$$

for compactly supported distributions w. Applying (B.14) to the operator $v \mapsto \mathcal{A}\Pi(\chi v)$ gives bounded maps from $H^{m-s}(\mathbb{R}^d)$ to $H^m(S^{d-1} \times \mathbb{R})$ at every integer $m \geq 0$. Interpolation between neighboring Hilbert Sobolev spaces gives the map from H^α to $H^{\alpha+s}$. For an input in H^α with $\alpha \geq 0$, approximation on a compact support in L^2 also gives convergence in L^1, and hence uniform convergence of the call integrals. Thus the extended operator agrees with the operator in (2.1). Adding the smooth mean correction proves (3.5).

For (3.6), let $\eta = f - g$, $\boldsymbol{a} = \int \boldsymbol{x} \eta \, \mathrm{d}\boldsymbol{x}$, and $\eta_0 = \eta - \sum_j a_j \tau_j$. From (B.5) and polar coordinates,

$$\int_{S^{d-1}} \int_{\mathbb{R}} |t|^{2(s-1)} \left| \widehat{\mathcal{A}\eta_0}(\boldsymbol{u}, t) \right|^2 \mathrm{d}t \, \mathrm{d}\sigma(\boldsymbol{u}) = \frac{2}{\omega_{d-1}} \int_{\mathbb{R}^d} |\widehat{\eta}_0(\boldsymbol{\xi})|^2 \, |\boldsymbol{\xi}|^{-2} \, \mathrm{d}\boldsymbol{\xi}. \tag{B.15}$$

The left side is bounded by the squared H^{s-1} norm on the full cylinder. Since $\mathcal{A}\eta_0$ vanishes for $|k| \geq r$, multiplying a bounded extension from $\mathcal{Z}$ by a smooth strike cutoff supported in $(-R, R)$ and equal to one on $[-r, r]$ shows that this full-cylinder norm is bounded by $C \, \|\mathcal{A}\eta_0\|_{H^{s-1}(\mathcal{Z})}$. Lemma B.1 gives

$$|\boldsymbol{a}| \leq C \, \|\mathcal{A}\eta\|_{L^2(Q_0)}, \qquad \|\mathcal{A}\eta_0\|_{H^{s-1}(\mathcal{Z})} \leq C \, \|\mathcal{A}\eta\|_{H^{s-1}(\mathcal{Z})}.$$

Pairing η_0 with the extension in (B.7), using (B.15) and $\|\eta_0\|_{H^{-1}} \leq \|\eta_0\|_{\dot{H}^{-1}}$, bounds every Lipschitz test by $C \, \|\mathcal{A}\eta\|_{H^{s-1}(\mathcal{Z})}$. The mean correction has the same bound. Taking the supremum proves (3.6). □

Set $a_{d,R} = R\omega_{d-1}(2\pi)^{1-d}$.

PROOF OF LEMMA 3.4. Define G_ψ on the full cylinder by its Fourier transform in k:

$$\widehat{G}_\psi(\boldsymbol{u}, t) = -a_{d,R} \, |t|^{d+1} \, \widehat{\psi}(t\boldsymbol{u}). \tag{B.16}$$

Plancherel and polar coordinates give

$$\int_{S^{d-1}} \int_{\mathbb{R}} |G_\psi(\boldsymbol{u}, k)|^2 \, \frac{\mathrm{d}k}{2R} \, \mathrm{d}\sigma(\boldsymbol{u}) = a_{d,R} \, \|\psi\|^2_{\dot{H}^s(\mathbb{R}^d)}. \tag{B.17}$$

For every finite signed measure η with zero mass and first moments supported on B_R, (B.5), (B.16), and the same polar calculation show

$$\int \psi \, \mathrm{d}\eta = \int_{\mathcal{Z}} G_\psi \mathcal{A}\eta \, \mathrm{d}Q_0. \tag{B.18}$$

The full-cylinder integral reduces to $\mathcal{Z}$, since $\mathcal{A}\eta$ vanishes outside $[-R, R]$. The full-cylinder Parseval integral equals

$$\frac{a_{d,R}}{4\pi R} \int_{S^{d-1}} \int_{\mathbb{R}} |t|^{d-1} \, \widehat{\psi}(t\boldsymbol{u}) \widehat{\eta}(-t\boldsymbol{u}) \, \mathrm{d}t \, \mathrm{d}\sigma(\boldsymbol{u}) = (2\pi)^{-d} \int \widehat{\psi}(\boldsymbol{\xi}) \widehat{\eta}(-\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi},$$

which is the required pairing. The Fourier integral converges absolutely since $\widehat{\psi}$ is rapidly decreasing and $\widehat{\eta}$ is bounded. The pairing of G_ψ and $\mathcal{A}\eta$ is a Plancherel pairing in L^2 of the full cylinder, since $G_\psi \in L^2$ by (B.17) and $\mathcal{A}\eta \in L^2$ by (3.1).

Let $\boldsymbol{m} = \int \boldsymbol{x}\,\mu(\mathrm{d}\boldsymbol{x})$ and use $\eta = \mu - \delta_0 - \sum_j m_j \tau_j\,\mathrm{d}\boldsymbol{x}$ in (B.18). Define

$$b_{\psi,j} = \int \psi\tau_j\,\mathrm{d}\boldsymbol{x} - \int_{\mathcal{Z}} G_\psi \mathcal{A}\tau_j\,\mathrm{d}Q_0, \qquad a_\psi = \psi(0) - \int_{\mathcal{Z}} G_\psi C_{\delta_0}\,\mathrm{d}Q_0. \tag{B.19}$$

Then $\int \psi\,\mathrm{d}\mu = a_\psi + \int G_\psi C_\mu\,\mathrm{d}Q_0 + \boldsymbol{m}\cdot\boldsymbol{b}_\psi$. For a probability measure, the parity identity reads $C_\mu(\boldsymbol{u},k) - C_\mu(-\boldsymbol{u},-k) = \boldsymbol{u}\cdot\boldsymbol{m} - k$. After multiplication by $\boldsymbol{u}$ and integration, the strike term vanishes and

$$\boldsymbol{m} = 2d\int_{\mathcal{Z}} \boldsymbol{u} C_\mu(\boldsymbol{u},k)\,\mathrm{d}Q_0.$$

Thus (3.7) holds with

$$\widetilde{G}_\psi(\boldsymbol{u},k) = G_\psi(\boldsymbol{u},k) + 2d\boldsymbol{b}_\psi\cdot\boldsymbol{u}. \tag{B.20}$$

Equation (B.17), Cauchy–Schwarz in (B.19), and the smoothness of τ_j bound $\left\|\widetilde{G}_\psi\right\|_2$ by $C\,\|\psi\|_{H^s}$. Finally, $s > d/2$ gives $|\psi(0)| \le C\,\|\psi\|_{H^s}$, proving the bound for a_ψ. □

APPENDIX C: PROOFS FOR SECTION 4

LEMMA C.1 (Random coverage). *Assume $Q \ge q_* Q_0$ as measures, where $q_* > 0$. There are constants $c, C > 0$ depending on d such that*

$$\Pr(h_n > 2t) \le Ct^{-d}\exp\left(-cq_* n t^d\right), \qquad 0 < t < 1. \tag{C.1}$$

Consequently, for every $0 < \delta < 1$,

$$\Pr\left(h_n > C_{d,q_*}\left(\frac{\log(en/\delta)}{n}\right)^{\frac{1}{d}}\right) \le \delta, \tag{C.2}$$

with a constant enlarged to include the cases where the displayed threshold is not small. Also $h_n = O((\log n/n)^{1/d})$ almost surely.

PROOF. Let $w_1,\dots,w_M$ be a t-net with $M \le Ct^{-d}$. If every ball $B_\rho(w_j,t)$ contains a sampled point, then $h_n \le 2t$. For each j, its probability of being empty is at most $\left(1 - cq_* t^d\right)^n \le \exp(-cq_* n t^d)$. The union bound gives (C.1). Taking t^d to be a sufficiently large multiple of $\log(en/\delta)/n$ gives (C.2). Taking $\delta = n^{-2}$ and applying the Borel–Cantelli lemma gives the almost-sure assertion. □

PROOF OF PROPOSITION 4.2. With probability $(1-\epsilon)^n$ all n locations equal $z_0 = (\boldsymbol{u}_0, k_0)$, and the data are n copies of $(z_0, C_\mu(z_0))$. Let $\boldsymbol{w}$ be a unit vector orthogonal to $\boldsymbol{u}_0$, and put $\boldsymbol{x} = r\boldsymbol{w}$ and $\boldsymbol{y} = -r\boldsymbol{w}$. The point masses $\delta_{\boldsymbol{x}}$ and $\delta_{\boldsymbol{y}}$ have the same call price $(-k_0)_+$ at z_0 and $W_1(\delta_{\boldsymbol{x}}, \delta_{\boldsymbol{y}}) = 2r$. The event depends only on the design, and on it the two laws produce identical data. The triangle inequality gives $\mathbb{E}_{\boldsymbol{x}} W_1(\widehat{\mu}, \delta_{\boldsymbol{x}}) + \mathbb{E}_{\boldsymbol{y}} W_1(\widehat{\mu}, \delta_{\boldsymbol{y}}) \ge 2r(1-\epsilon)^n$, and the larger of the two risks is at least $r(1-\epsilon)^n$. □

PROOF OF THEOREM 4.1. If $\mu, \nu \in \mathcal{K}_n(z, v)$, the function $C_\mu - C_\nu$ is zero at all z_i and has ρ-Lipschitz constant at most $2R$. Therefore

$$\|C_\mu - C_\nu\|_{L^2(Q_0)} \le \|C_\mu - C_\nu\|_\infty \le 2Rh(z).$$

Apply (3.3) and take the supremum over the pair. The probability event in Lemma C.1 depends only on the design, giving simultaneous validity over the price vectors. □

LEMMA C.2 (Sampling inequalities). *Let $t > d/2$ and $0 \le q \le t$. There are $h_* > 0$ and $C > 0$ such that every finite set $z \subset \mathcal{Z}$ with fill distance $h \le h_*$ and every $G \in H^t(\mathcal{Z})$ satisfy*

$$\|G\|_{L^2(\mathcal{Z})} \le C\left(h^t \|G\|_{H^t(\mathcal{Z})} + \max_{z_i \in z} |G(z_i)|\right). \tag{C.3}$$

If G vanishes on z, then

$$\|G\|_{H^q(\mathcal{Z})} \le C h^{t-q} \|G\|_{H^t(\mathcal{Z})}. \tag{C.4}$$

PROOF. The Euclidean L^2 sampling inequality on a bounded Lipschitz domain, in the case $p = q = 2$, $l = 0$ with discrete exponent ∞ of Theorem 1.1 in Corrigan, Wallin, and Wanner [4], which restates the sampling theorem of Arcangéli and coauthors, gives $\|G\|_{L^2} \le C(h^t \|G\|_{H^t} + \max_i |G(z_i)|)$ for every real $t > d/2$.

To transfer this estimate to $\mathcal{Z}$, use a finite collection of sphere charts with smaller subcharts covering the sphere and smooth cutoffs supported in the larger charts. In each chart, apply the Euclidean result to the cutoff times G. At original sample points within the chart, this product is bounded by $\max_i |G(z_i)|$. In the portion outside the cutoff support, add a finite h-net of points, where the product is zero. The positive distance between the smaller and larger chart boundaries ensures, for sufficiently small h, a chart fill distance at most Ch. The strike endpoints are covered by Lipschitz half-boxes, to which the same Euclidean result applies. Bounded coordinate maps and the finite partition of unity give (C.3). If G vanishes on z, its right side is $Ch^t \|G\|_{H^t(\mathcal{Z})}$, and Hilbert interpolation gives

$$\|G\|_{H^q(\mathcal{Z})} \le C \|G\|_{L^2(\mathcal{Z})}^{1-q/t} \|G\|_{H^t(\mathcal{Z})}^{q/t},$$

which proves (C.4). □

PROOF OF THEOREM 4.3. For two feasible densities, $G = \mathcal{A}(f - g)$ vanishes at every sampled quote and, by (3.5), satisfies $\|G\|_{H^{\alpha+s}(\mathcal{Z})} \le 2CL$. Since $\alpha + s > d/2$, Lemma C.2 applies with $t = \alpha + s$ and $q = s - 1$. Lemma 3.3 then gives

$$W_1(f\,\mathrm{d}\boldsymbol{x}, g\,\mathrm{d}\boldsymbol{x}) \le C \|G\|_{H^{s-1}(\mathcal{Z})} \le Ch(z)^{\alpha+1} \|G\|_{H^{\alpha+s}(\mathcal{Z})} \le CLh(z)^{\alpha+1}.$$

This proves (4.4). Lemma C.1 gives (4.5). When its fill-distance threshold exceeds h_*, use $W_1 \le 2r$ and the positive lower bound on the L^2 norm of a probability density supported on B_r to enlarge the constant. □

PROOF OF PROPOSITION 4.4. For two feasible densities f, g, write $G = \mathcal{A}(f - g)$ and $t = \alpha + s$. Lemma 3.3 gives $\|G\|_{H^t(\mathcal{Z})} \le CL$, and $\max_i |G(z_i)| \le 2\epsilon$. Inequality (C.3) gives $\|G\|_{L^2(\mathcal{Z})} \le C(Lh^t + \epsilon)$. Interpolating between orders 0 and t at order $s - 1$, and using $1 - (s-1)/t = \vartheta$ and $t\vartheta = \alpha + 1$,

$$\|G\|_{H^{s-1}(\mathcal{Z})} \le C \|G\|_{L^2(\mathcal{Z})}^{\vartheta} \|G\|_{H^t(\mathcal{Z})}^{1-\vartheta} \le C(Lh^t + \epsilon)^{\vartheta} L^{1-\vartheta} \le C\left[Lh^{\alpha+1} + L^{1-\vartheta}\epsilon^{\vartheta}\right].$$

Inequality (3.6) proves (4.6), and Lemma C.1 gives the random version. □

COROLLARY C.3 (Simultaneous control and stopping). *Under the assumptions of Theorem 4.3, with probability at least $1 - \delta$, its diameter bound holds for every $n \ge 1$ simultaneously with right side*

$$CL\left(\frac{\log(en^3/\delta)}{n}\right)^{\frac{\alpha+1}{d}}. \tag{C.5}$$

It therefore remains valid at any almost surely finite stopping time based on the observed triples. For $0 < t < h_*$*, the coverage stopping time* $\tau_t = \inf\{n \geq 1 : h_n \leq t\}$ *satisfies*

$$\mathbb{E}\tau_t \leq C_{d,q_*} t^{-d} \log(e/t). \tag{C.6}$$

At τ_t*, every Sobolev-restricted quote fiber has* W_1 *diameter at most* $CLt^{\alpha+1}$.

PROOF OF COROLLARY C.3. Apply Theorem 4.3 at sample size n with failure probability $6\delta/(\pi^2 n^2)$ and take a union bound. Since the resulting event holds for all sample sizes and price vectors, it also holds at the stopping time. Monotonicity of h_n and (C.1) give

$$\Pr(\tau_t > n) \leq \min\left(1, Ct^{-d} e^{-cq_* n t^d}\right).$$

Summing this bound over $n \geq 0$, splitting the sum at a constant multiple of $t^{-d}\log(e/t)$, proves (C.6). The last assertion is the deterministic bound (4.4). □

For the minimum-norm interpolant (4.7), if the true density is in $\mathcal{F}_\alpha(L)$, the feasible set is nonempty and the minimizer has norm at most L. A minimizing sequence is bounded in the Hilbert space H^α, and its weak limit preserves support, nonnegativity, mass, and the finite integral constraints. Lower semicontinuity gives existence and strict convexity gives uniqueness. The minimizer is a measurable function of the observations, by the measurable minimum theorem applied to the weakly compact metrizable norm sublevels and their countable union.

LEMMA C.4 (Localized perturbations). *Fix* $\alpha \geq 0$*. There are a smooth probability density* f_0 *supported on* B_r*, a nonzero smooth function* ψ *supported in the unit cube with zero mass and first moments, and constants* $b_0, C, c > 0$ *with the following property. For each sufficiently small* $\Delta > 0$*, one can place* $M \asymp \Delta^{-d}$ *disjoint translated cubes of side* Δ *in a region where* f_0 *is bounded below. For any* $\boldsymbol{\theta} \in [-1,1]^M$ *and* $0 < b \leq b_0$*, define*

$$H_{\boldsymbol{\theta}}(\boldsymbol{x}) = b\Delta^\alpha \sum_{j=1}^{M} \theta_j \psi\left(\frac{\boldsymbol{x} - \boldsymbol{x}_j}{\Delta}\right). \tag{C.7}$$

Then $f_{\boldsymbol{\theta}} = f_0 + H_{\boldsymbol{\theta}}$ *is a probability density with the one-dimensional coordinate marginals of* f_0*,* $\|H_{\boldsymbol{\theta}}\|_{H^\alpha} \leq Cb$*, and*

$$W_1\left(f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x}, f_{\boldsymbol{\theta}'}\,\mathrm{d}\boldsymbol{x}\right) \geq cb\Delta^{\alpha+d+1}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}'\right\|_1, \qquad \boldsymbol{\theta}, \boldsymbol{\theta}' \in [-1,1]^M. \tag{C.8}$$

PROOF. Take $\psi = \partial_1\partial_2\zeta$ for a nonzero smooth bump ζ compactly supported inside the unit cube. Integration by parts gives its zero mass and first moments, and integrating out x_1 or x_2 shows that every one-dimensional coordinate marginal of ψ vanishes. Choose a smooth function φ compactly supported in the same cube with $\mathrm{Lip}(\varphi) \leq 1$ and $\int \varphi\psi\,\mathrm{d}\boldsymbol{x} > 0$. Such a function is obtained by taking a sufficiently small positive multiple of ψ itself. Place the cubes with a fixed proportion of their side length between the supports. Choose f_0 smooth and positive on the containing cube. Small b_0 guarantees positivity, and the zero mass of every bump preserves total mass.

If k is an integer with $k \geq \alpha$ and $k \geq 1$, disjoint supports give

$$\|H_{\boldsymbol{\theta}}\|_{L^2} \leq Cb\Delta^\alpha, \qquad \|H_{\boldsymbol{\theta}}\|_{H^k} \leq Cb\Delta^{\alpha-k}.$$

Indeed, the squared norm of a derivative of order $l \leq k$ is bounded by $Cb^2\Delta^{2\alpha-2l+d}\sum_j \theta_j^2 \leq Cb^2\Delta^{2\alpha-2l}$. Fourier interpolation yields $\|H_{\boldsymbol{\theta}}\|_{H^\alpha} \leq Cb$; for $\alpha = 0$ the L^2 bound suffices.

The smooth function

$$F(\boldsymbol{x}) = \Delta \sum_j \operatorname{sgn}(\theta_j - \theta'_j)\varphi\left(\frac{\boldsymbol{x} - \boldsymbol{x}_j}{\Delta}\right)$$

has gradient norm at most one everywhere, since the summands have disjoint supports. Its integral against $H_{\boldsymbol{\theta}} - H_{\boldsymbol{\theta}'}$ is $b\Delta^{\alpha+d+1}(\int \varphi\psi)\left\|\boldsymbol{\theta} - \boldsymbol{\theta}'\right\|_1$, proving (C.8). □

PROOF OF THEOREM 4.5. Use Lemma C.4 with a grid containing M cells, where $2n \le M \le C_d n$ and $\Delta \asymp n^{-1/d}$. Set

$$T_{ij} = \int \phi_i(\boldsymbol{x})\psi\left(\frac{\boldsymbol{x} - \boldsymbol{x}_j}{\Delta}\right) \mathrm{d}\boldsymbol{x}, \qquad S = \ker(T) \subseteq \mathbb{R}^M.$$

Then $\dim S \ge M - n$. Choose a vertex $\boldsymbol{\theta}$ of the compact polytope $S \cap [-1,1]^M$. At least $\dim S$ coordinates of $\boldsymbol{\theta}$ have absolute value one. Otherwise, there is a nonzero vector in S whose coordinates vanish at all the saturated coordinates of $\boldsymbol{\theta}$; sufficiently small perturbations in both signs of this vector stay in the polytope, contradicting the vertex property. Hence

$$T\boldsymbol{\theta} = 0, \qquad \sum_j |\theta_j| \ge M - n \ge M/2.$$

Let $L_0 = \|f_0\|_{H^\alpha}$, and choose $b > 0$ sufficiently small that the positivity condition holds and $L_0 + Cb \le L$. The densities $f^\pm = f_0 \pm H_{\boldsymbol{\theta}}$ are in $\mathcal{F}_\alpha(L)$, and $T\boldsymbol{\theta} = 0$ gives all the quote equalities. Equation (C.8) yields

$$W_1(f^+ \mathrm{d}\boldsymbol{x}, f^- \mathrm{d}\boldsymbol{x}) \ge cb\Delta^{\alpha+d+1}(M - n) \ge c'\Delta^{\alpha+1} \ge c'' n^{-(\alpha+1)/d}.$$

The constants can be reduced to include any finitely many small n outside the initial grid construction. □

LEMMA C.5 (Bayesian bound for exact linear observations). *Let $T \in \mathbb{R}^{n\times M}$ be deterministic. Let $\boldsymbol{\Theta}$ have independent coordinates with density*

$$p_0(t) = \cos^2\left(\frac{\pi t}{2}\right)\mathbf{1}_{\{-1<t<1\}}. \tag{C.9}$$

For every measurable map $\widehat{\boldsymbol{\theta}} : \mathbb{R}^n \to [-1,1]^M$, based on the exact observation $T\boldsymbol{\Theta}$,

$$\mathbb{E}\left\|\widehat{\boldsymbol{\theta}}(T\boldsymbol{\Theta}) - \boldsymbol{\Theta}\right\|_2^2 \ge \frac{M - \operatorname{rank}(T)}{\pi^2}, \qquad \mathbb{E}\left\|\widehat{\boldsymbol{\theta}}(T\boldsymbol{\Theta}) - \boldsymbol{\Theta}\right\|_1 \ge \frac{M - \operatorname{rank}(T)}{2\pi^2}. \tag{C.10}$$

PROOF. The function p_0 integrates to one, vanishes at its support endpoints, and satisfies

$$\int_{-1}^1 p_0'(t)\,\mathrm{d}t = 0, \qquad \int_{-1}^1 \frac{p_0'(t)^2}{p_0(t)}\,\mathrm{d}t = \pi^2 \int_{-1}^1 \sin^2\left(\frac{\pi t}{2}\right)\mathrm{d}t = \pi^2.$$

Write $p(\boldsymbol{\theta}) = \prod_{j=1}^M p_0(\theta_j)$ and $\boldsymbol{S}(\boldsymbol{\theta}) = \nabla \log p(\boldsymbol{\theta})$ on the open cube. Independence and the preceding integrals give

$$\mathbb{E}\boldsymbol{S}(\boldsymbol{\Theta}) = 0, \qquad \mathbb{E}\left[\boldsymbol{S}(\boldsymbol{\Theta})\boldsymbol{S}(\boldsymbol{\Theta})^{\mathsf{T}}\right] = \pi^2 I_M. \tag{C.11}$$

Let $\boldsymbol{e}(\boldsymbol{\theta}) = \widehat{\boldsymbol{\theta}}(T\boldsymbol{\theta}) - \boldsymbol{\theta}$. For any unit vector $\boldsymbol{v} \in \ker(T)$ and any coordinate j, Fubini's theorem along lines parallel to $\boldsymbol{v}$ yields

$$\int_{\mathbb{R}^M} \widehat{\theta}_j(T\boldsymbol{\theta})\,\partial_{\boldsymbol{v}} p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = 0. \tag{C.12}$$

Indeed, $\widehat{\theta}_j(T(\boldsymbol{\theta}+t\boldsymbol{v}))$ is constant in t, whereas the integral of $\partial_t p(\boldsymbol{\theta}+t\boldsymbol{v})$ on each such line is zero. The product density is continuously differentiable with compact support, and the decoder is bounded, which justifies the integration. Integration by parts also gives

$$\int_{\mathbb{R}^M} \theta_j \partial_{\boldsymbol{v}} p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = -v_j.$$

Subtracting and using $\partial_{\boldsymbol{v}} p = p\,\boldsymbol{v}\cdot\boldsymbol{S}$ proves

$$\mathbb{E}\left[e_j(\boldsymbol{\Theta})\,\boldsymbol{v}\cdot\boldsymbol{S}(\boldsymbol{\Theta})\right] = v_j. \tag{C.13}$$

Let P be the orthogonal projection onto $\ker(T)$ and let $k=\operatorname{tr}(P)=M-\operatorname{rank}(T)$. Sum (C.13) over an orthonormal basis of the kernel, after taking its inner product with the corresponding basis vector. Equations (C.11) and (C.13) give

$$\mathbb{E}\left[(P\boldsymbol{e})\cdot(P\boldsymbol{S})\right] = k, \qquad \mathbb{E}\,\|P\boldsymbol{S}\|_2^2 = \pi^2 k.$$

For $k>0$, Cauchy–Schwarz therefore implies

$$k^2 \le \mathbb{E}\,\|P\boldsymbol{e}\|_2^2\;\pi^2 k \le \mathbb{E}\,\|\boldsymbol{e}\|_2^2\;\pi^2 k.$$

The case $k=0$ is immediate. Since both the coefficient vector and the decoder lie in $[-1,1]^M$, each coordinate error has absolute value at most two. Thus $\|\boldsymbol{e}\|_2^2 \le 2\,\|\boldsymbol{e}\|_1$, proving both inequalities in (C.10). □

PROOF OF THEOREM 4.6. Draw $\boldsymbol{\Theta}$ from the product prior (C.9), independently of the design and of the randomization. Fix $\varepsilon>0$ and choose a measurable $\widehat{\boldsymbol{\theta}}_\varepsilon\in[-1,1]^M$ with $\mathsf{d}(\widehat{f},f_{\widehat{\boldsymbol{\theta}}_\varepsilon}) \le \inf_{\boldsymbol{\theta}}\mathsf{d}(\widehat{f},f_{\boldsymbol{\theta}})+\varepsilon$; such a choice follows by enumerating a countable dense subset of the cube and selecting the first coefficient vector within ε of the infimum over that subset, the objective being continuous in $\boldsymbol{\theta}$ and measurable in the data. The triangle inequality and the separation hypothesis give

$$\delta\left\|\widehat{\boldsymbol{\theta}}_\varepsilon-\boldsymbol{\Theta}\right\|_1 \le \mathsf{d}(f_{\widehat{\boldsymbol{\theta}}_\varepsilon},f_{\boldsymbol{\Theta}}) \le 2\,\mathsf{d}(\widehat{f},f_{\boldsymbol{\Theta}})+\varepsilon.$$

Conditional on $(\mathsf{L}_1,\dots,\mathsf{L}_n,a)$ and the randomization, the data are $a'+T\boldsymbol{\Theta}$ with a fixed vector $a'\in\mathbb{R}^n$ and a fixed matrix $T\in\mathbb{R}^{n\times M}$ of rank at most n, and $\widehat{\boldsymbol{\theta}}_\varepsilon$ is a measurable function of $T\boldsymbol{\Theta}$. Lemma C.5 gives a conditional expectation of $\left\|\widehat{\boldsymbol{\theta}}_\varepsilon-\boldsymbol{\Theta}\right\|_1$ of at least $(M-n)/(2\pi^2)$, since the conditional prior is still (C.9). Taking expectations, letting $\varepsilon\downarrow 0$, and bounding the prior average by the supremum proves (4.8). □

THEOREM C.6 (Exact random-design minimax lower bound). *Fix $\alpha\ge 0$ and $L>L_0$ as in Theorem 4.5. There is $c>0$, independent of Q_n, such that*

$$\mathfrak{R}_n^{\mathrm{ex}}(\mathsf{Q}_n,\alpha,L) \ge c\,n^{-(\alpha+1)/d}. \tag{C.14}$$

For an i.i.d. design with $Q\ge q_*Q_0$, there is also $C<\infty$ such that

$$c\,n^{-\frac{\alpha+1}{d}} \le \mathfrak{R}_n^{\mathrm{ex}}(Q^{\otimes n},\alpha,L) \le C\left(\frac{\log(en)}{n}\right)^{\frac{\alpha+1}{d}}. \tag{C.15}$$

The lower assertion remains valid for any n continuous linear payoffs whose joint selection is exogenous.

PROOF OF THEOREM C.6 AND COROLLARY 4.7. Use the localized family of Lemma C.4 with $2n \le M \le C_d n$, $\Delta \asymp n^{-1/d}$, and a sufficiently small $b > 0$. All densities $f_{\boldsymbol{\theta}} = f_0 + H_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in [-1,1]^M$, lie in $\mathcal{F}_\alpha(L)$.

Conditional on a realized design z, the observations are $V_i = C_{f_0}(z_i) + (T_z\boldsymbol{\theta})_i$ with

$$(T_z)_{ij} = b\Delta^\alpha \int_{B_r} (\boldsymbol{u}_i \cdot \boldsymbol{x} - k_i)_+ \psi\left(\frac{\boldsymbol{x} - \boldsymbol{x}_j}{\Delta}\right) \mathrm{d}\boldsymbol{x},$$

which is affine in $\boldsymbol{\theta}$, and $\boldsymbol{\theta} \mapsto f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x}$ is continuous for W_1. Theorem 4.6 with $\mathsf{d} = W_1$ and the separation (C.8), $\delta = cb\Delta^{\alpha+d+1}$, gives

$$\sup_{f \in \mathcal{F}_\alpha(L)} \mathbb{E}_{f,\mathbb{Q}_n} W_1(\widehat{\mu}, f\,\mathrm{d}\boldsymbol{x}) \ge \frac{cb}{4\pi^2}\Delta^{\alpha+d+1}(M-n) \ge c_3 n^{-\frac{\alpha+1}{d}} \tag{C.16}$$

for every estimator, since $M - n \ge M/2 \asymp \Delta^{-d}$. The same calculation replaces the basket payoff by any specified continuous linear payoff, proving the final assertion.

For the upper bound, use the minimum-norm interpolant in (4.7). Write $\beta = (\alpha+1)/d$. For $n \ge 2$, apply (4.5) with $\delta = n^{-\beta-2}$ and use $W_1 \le 2r$ on its complement. Uniformly over the density class,

$$\mathbb{E}_f W_1(\widehat{\mu}, f\,\mathrm{d}\boldsymbol{x}) \le CL\left(\frac{\log(en^{\beta+3})}{n}\right)^\beta + 2rn^{-\beta-2} \le C'\left(\frac{\log(en)}{n}\right)^\beta.$$

Enlarging the constant includes $n = 1$. This proves (C.15). □

REMARK C.7 (Adaptive and dependent quote arrivals). The coverage argument of Lemma C.1 uses independence only through the probability that a fixed ball is missed. Suppose that the conditional law of each new quote location, given the past quotes and any selection randomization, dominates $q_i Q_0$ as a measure for deterministic $q_i \in [0,1]$. Iterating the conditional miss probability for each ball of a t-net gives $\Pr(h_n > 2t) \le Ct^{-d}\exp(-ct^d\Lambda_n)$ with $\Lambda_n = \sum_{i \le n} q_i$, and the deterministic bound (4.4) then yields the analogue of (4.5) with n replaced by $1 + \Lambda_n$ and $\log(en/\delta)$ replaced by $\log(en^3/\delta)$, simultaneously over n.

APPENDIX D: PROOFS FOR SECTION 5

The torus $\mathbb{T}^d_T$, the cutoff χ, the tests $\psi_{\boldsymbol{k}}$, the features $\widetilde{G}_{\boldsymbol{k}}$ and the constants t, D_α, ϑ are as in Section 5.

PROOF OF PROPOSITION 5.4. For the projection in (5.6), fix a countable W_1-dense subset of $\mathcal{P}(B_r)$ and choose its first element whose objective is within $1/n$ of the infimum over that subset. The objectives are measurable functions of the finitely many estimated moments, and continuity in ν makes this infimum equal to the infimum in (5.6). The construction works for any positive tolerance, including e^{-n} in (5.13).

Let $f_M = \mathbb{E}_\mu \widetilde{f}_M$, a signed filtered version of μ. For a mass-zero trigonometric density g on the torus, normalization of a Lipschitz test at a point bounds its H^1 norm by a constant depending only on the torus. Sobolev duality therefore gives

$$\|g\,\mathrm{d}\boldsymbol{x}\|_{\mathrm{Lip}^*} \le C\,\|g\|_{H^{-1}(\mathbb{T}^d_T)}. \tag{D.1}$$

Orthogonality of the Fourier basis and (5.4) give

(D.2)

$$\mathbb{E}_\mu \left\|\widetilde{f}_M - f_M\right\|^2_{H^{-1}(\mathbb{T}^d_T)} \le \frac{C}{n}\sum_{|\boldsymbol{k}| \le 2M} (1+|\boldsymbol{k}|^2)^{-1}(1+|\boldsymbol{k}|)^{2s} \le \frac{C}{n}M^{2s+d-2} = \frac{C}{n}M^{2d+1}.$$

If $\mu = f\,\mathrm{d}\boldsymbol{x}$ with $f \in \mathcal{F}_\alpha(L)$, its zero extension, regarded as a periodic function, has $H^\alpha(\mathbb{T}_T^d)$ norm at most CL. The multiplier $1 - w(\boldsymbol{k}/M)$ vanishes for $|\boldsymbol{k}| \le M$, and hence

$$\|f - f_M\|^2_{H^{-1}(\mathbb{T}_T^d)} \le CM^{-2\alpha-2}\,\|f\|^2_{H^\alpha(\mathbb{T}_T^d)} \le CL^2M^{-2\alpha-2}. \tag{D.3}$$

For an arbitrary probability measure, let Φ_M be the periodic convolution kernel with Fourier multiplier $w(\boldsymbol{k}/M)$. Its integral is one. Poisson summation expresses it as the periodization of a rescaled Schwartz function, and therefore

$$\int_{\mathbb{T}_T^d} \operatorname{dist}(\boldsymbol{x}, 0)\,|\Phi_M(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \le CM^{-1}. \tag{D.4}$$

Indeed, the corresponding integral on $\mathbb{R}^d$ is M^{-1} times the finite first absolute moment of that Schwartz function; folding to the torus only decreases distance. Since $f_M\,\mathrm{d}\boldsymbol{x} = \Phi_M * \mu$, each 1-Lipschitz F satisfies

$$\begin{aligned}\left|\int F\,\mathrm{d}\mu - \int F f_M\,\mathrm{d}\boldsymbol{x}\right| &= \left|\int\int \Phi_M(\boldsymbol{z})\,(F(\boldsymbol{x}) - F(\boldsymbol{x}+\boldsymbol{z}))\,\mathrm{d}\boldsymbol{z}\,\mu(\mathrm{d}\boldsymbol{x})\right| \\ &\le CM^{-1}.\end{aligned}$$

The true μ is feasible for the projection in (5.6). The triangle inequality yields

$$W_1(\widehat{\mu}_M, \mu) \le 2\left\|\widetilde{f}_M\,\mathrm{d}\boldsymbol{x} - \mu\right\|_{\mathrm{Lip}^*} + \frac{1}{n}.$$

Combine this inequality with (D.1)–(D.4) and Cauchy–Schwarz to obtain both conclusions. □

PROOF OF PROPOSITION 5.5. Use the sign family $f_{\boldsymbol{\theta}} = f_0 + H_{\boldsymbol{\theta}}$ from Lemma C.4, with a sufficiently small $b > 0$ such that every density lies in $\mathcal{F}_\alpha(L)$. The full observation law under $f_{\boldsymbol{\theta}}$ is denoted by $P^{(n)}_{\boldsymbol{\theta}}$. If $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ differ only in coordinate j, the density difference is

$$2b\Delta^\alpha\theta_j\psi\left(\frac{\boldsymbol{x} - \boldsymbol{x}_j}{\Delta}\right).$$

It has zero mass and first moments. Scaling its Fourier transform in (3.1) gives

$$\|C_{f_{\boldsymbol{\theta}}} - C_{f_{\boldsymbol{\theta}'}}\|^2_{L^2(Q_0)} = c_\psi b^2\Delta^{2\alpha+2d+3}, \tag{D.5}$$

where $0 < c_\psi < \infty$. To verify the exponent, the transform contributes $\Delta^{2\alpha+2d}$, the frequency substitution contributes Δ^{-d}, and the weight $|\boldsymbol{\xi}|^{-d-3}$ contributes Δ^{d+3}. The integral is finite at zero by the two vanishing moments of ψ and at infinity by smoothness.

Lemma A.1 gives

$$H^2\left(P^{(n)}_{\boldsymbol{\theta}}, P^{(n)}_{\boldsymbol{\theta}'}\right) \le \frac{nI_\xi}{4\varsigma^2}\,\|C_{f_{\boldsymbol{\theta}}} - C_{f_{\boldsymbol{\theta}'}}\|^2_{L^2(Q_0)} = \frac{nI_\xi c_\psi b^2}{4\varsigma^2}\Delta^{2\alpha+2d+3}. \tag{D.6}$$

Choose $\Delta \asymp n^{-1/(2\alpha+2d+3)}$ with a sufficiently small proportionality constant, so that (D.6) is at most $1/16$. A grid with $M \asymp \Delta^{-d}$ disjoint cells is available, and the total variation distance between adjacent observation laws is at most $1/4$.

For an arbitrary estimator $\widehat{\mu}$, choose a closest member $f_{\widehat{\boldsymbol{\theta}}}$ of the finite family in W_1, resolving ties deterministically. Since $W_1(\widehat{\mu}, f_{\widehat{\boldsymbol{\theta}}}\,\mathrm{d}\boldsymbol{x}) \le W_1(\widehat{\mu}, f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x})$, (C.8), with $\left\|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right\|_1 = 2\sum_j \mathbf{1}_{\{\widehat{\theta}_j \ne \theta_j\}}$, and the triangle inequality imply

$$W_1(\widehat{\mu}, f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x}) \ge \frac{c}{2}b\Delta^{\alpha+d+1}\sum_{j=1}^M \mathbf{1}_{\{\widehat{\theta}_j \ne \theta_j\}}.$$

Average over the 2^M sign vectors. For each coordinate j, pair every vector with the vector obtained by flipping its jth sign. For any binary decision, the sum of its two error probabilities is at least one minus the total variation distance of the corresponding laws. The adjacent bound therefore gives

$$2^{-M}\sum_{\boldsymbol{\theta}} P_{\boldsymbol{\theta}}^{(n)}(\widehat{\theta}_j \neq \theta_j) \geq \frac{3}{8}.$$

It follows that

$$\sup_{\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{\theta}} W_1(\widehat{\mu}, f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x}) \geq cb\Delta^{\alpha+d+1}M \geq c'\Delta^{\alpha+1} \geq c''n^{-(\alpha+1)/(2\alpha+2d+3)}.$$

Adjusting the constants covers the finitely many small n. □

The Fourier estimator of Section 5.1 uses the known uniform design. For a known density $q = \mathrm{d}Q/\mathrm{d}Q_0$ bounded below by a positive constant, replace $\widetilde{G}_\psi(Z_i)Y_i$ by $\widetilde{G}_\psi(Z_i)Y_i/q(Z_i)$. The upper proof is unchanged up to constants. Matching lower bounds under nonuniform designs follow from the same testing construction if q is also bounded from above.

D.1. Pointwise Control of Inverse Features. We control the weighted sum of squared inverse features uniformly over quote locations. The bound transfers integrated pricing-surface error directly to the negative-Sobolev variance of the reconstructed law.

Use the torus, real cutoff χ, and Fourier tests $\psi_{\boldsymbol{k}}$ defined above.

LEMMA D.1 (Spectral envelope). *For every $M \geq 1$, there is a jointly measurable version of the inverse features such that*

$$\sup_{z\in\mathcal{Z}} \sum_{0<|\boldsymbol{k}|\leq 2M} (1+|\boldsymbol{k}|^2)^{-1} \left|\widetilde{G}_{\boldsymbol{k}}(z)\right|^2 \leq CM^{2d+1}. \tag{D.7}$$

The constant depends only on the fixed support, torus, and smooth cutoffs.

PROOF. Let $\boldsymbol{\omega}_{\boldsymbol{k}} = 2\pi\boldsymbol{k}/T$. The uncorrected feature from (B.16) satisfies

$$\widehat{G}_{\boldsymbol{k}}(\boldsymbol{u}, t) = -a_{d,R}\,|t|^{d+1}\,\widehat{\chi}(t\boldsymbol{u} + \boldsymbol{\omega}_{\boldsymbol{k}}).$$

Fourier inversion defines a continuous function of $(\boldsymbol{u}, k)$, since the integrand is integrable. Write $b = \boldsymbol{u}\cdot\boldsymbol{\omega}_{\boldsymbol{k}}$ and $a = \mathrm{dist}(\boldsymbol{\omega}_{\boldsymbol{k}}, \mathbb{R}\boldsymbol{u})$. For every sufficiently large P, Schwartz decay gives

$$\begin{aligned}\sup_k |G_{\boldsymbol{k}}(\boldsymbol{u}, k)| &\leq C_P \int_{\mathbb{R}} |t|^{d+1} \left(1 + \sqrt{(t+b)^2 + a^2}\right)^{-P} \mathrm{d}t \\ &\leq C'_P(1+|\boldsymbol{k}|)^{d+1} \int_{\mathbb{R}} (1+|v|)^{d+1}(1+|v|+a)^{-P}\,\mathrm{d}v.\end{aligned}$$

In the second line $v = t + b$ and $|b| \leq C\,|\boldsymbol{k}|$. Choosing P sufficiently large for each A and integrating after $v = (1+a)w$ gives

$$\sup_k |G_{\boldsymbol{k}}(\boldsymbol{u}, k)| \leq C_A(1+|\boldsymbol{k}|)^{d+1}\,(1+\mathrm{dist}(\boldsymbol{k}, \mathbb{R}\boldsymbol{u}))^{-A}. \tag{D.8}$$

Uniformly in $\boldsymbol{u}$, the number of integer lattice points with $|\boldsymbol{k}| \leq 2M$ and distance at most $j+1$ from $\mathbb{R}\boldsymbol{u}$ is at most

$$C(M+j+2)(j+2)^{d-1} \leq C'M(j+2)^d.$$

To verify this, place disjoint unit cubes about these lattice points. Their union lies in a tube of length at most $CM + C(j+1)$ and radius at most $C(j+1)$; its volume gives the bound. Summing over integer distance shells, for $2A > d+1$,

$$\sum_{|\boldsymbol{k}|\le 2M} (1+\operatorname{dist}(\boldsymbol{k}, \mathbb{R}\boldsymbol{u}))^{-2A} \le CM.$$

Equation (D.8) consequently gives

$$\sup_{z\in\mathcal{Z}} \sum_{0<|\boldsymbol{k}|\le 2M} (1+|\boldsymbol{k}|^2)^{-1} |G_{\boldsymbol{k}}(z)|^2 \le CM^{2d}\, M = CM^{2d+1}. \tag{D.9}$$

For the correction in (B.20), the proof of Lemma 3.4 gives $|\boldsymbol{b}_{\psi_{\boldsymbol{k}}}| \le C \|\psi_{\boldsymbol{k}}\|_{H^s} \le C(1+|\boldsymbol{k}|)^s$. Hence

$$\sup_{\boldsymbol{u}\in S^{d-1}} \sum_{0<|\boldsymbol{k}|\le 2M} (1+|\boldsymbol{k}|^2)^{-1} |2d\boldsymbol{b}_{\psi_{\boldsymbol{k}}} \cdot \boldsymbol{u}|^2 \le CM^{2s+d-2} = CM^{2d+1}.$$

Combine this with (D.9) and $|x+y|^2 \le 2|x|^2 + 2|y|^2$. This proves (D.7). □

PROOF OF LEMMA 5.6. We describe one coordinate cube $D = [0,1]^d$. The sampling law dominates a positive multiple of Lebesgue measure on this cube; samples outside the cube are ignored. The same construction will be applied in finitely many sphere–strike charts.

Fix a coordinatewise polynomial degree p. Choose $J = (p+1)^d$ tensor-product interpolation nodes strictly inside D, with mutually disjoint small neighborhoods $A_1, \dots, A_J$. Write $b_1, \dots, b_J$ for a basis of the coordinatewise polynomial space. The matrix $B = (b_j(a_l))_{l,j}$ at the reference nodes is invertible. By continuity, the neighborhoods can be chosen so small that every matrix whose lth row is an average of $(b_j(x))_j$ over points in A_l has inverse norm at most $2\|B^{-1}\|$. For a dyadic cube D' of side a, transport these neighborhoods and basis functions by its affine coordinate map. Call D' good if every transported neighborhood contains at least one of the m samples. Each neighborhood has sampling probability at least ca^d. Thus

$$\Pr(D' \text{ is not good}) \le J\exp(-cma^d). \tag{D.10}$$

On a good cube, fit a polynomial by equating its average over the actual sample points in each neighborhood to the corresponding average observed price. This fit reproduces every polynomial in the chosen space exactly. Moreover, on the rescaled unit cube its absolute value is bounded by a constant times the largest absolute data value. Its conditional variance at each point is at most

$$C\varsigma^2 \sum_{l=1}^{J} \frac{1}{\mathsf{N}_l} \le CJ\varsigma^2, \tag{D.11}$$

where $\mathsf{N}_l \ge 1$ is the count in neighborhood l. Distinct neighborhoods in this fit use disjoint samples.

Choose a dyadic base side $h \ge m^{-1/d}$. A prescribed bandwidth in $[m^{-1/d}, 1]$ is rounded upward to a dyadic side, by a factor smaller than two. We retain the notation h for this side. For each base cube, use the fit on its smallest good dyadic ancestor, including itself. If no ancestor is good, use zero on that base cube. The resulting piecewise polynomial operator, denoted by P_m, is a measurable function of the design and is linear in the data.

Let $G \in H^t(D)$, $t > d/2$, and choose $p \ge \lceil t \rceil$. On a cube D' of side a, polynomial approximation and Sobolev embedding give a polynomial P in the fitted space with

$$\|G - P\|_{L^\infty(D')} \le Ca^{t-d/2} |G|_{H^t(D')}. \tag{D.12}$$

Here $|\cdot|_{H^t}$ is the top-order derivative seminorm for integer t, and the fractional seminorm of the derivatives of order $\lfloor t \rfloor$ otherwise. To obtain (D.12), on the unit cube subtract a bounded polynomial projection of degree $t-1$ for integer t, or degree $\lfloor t \rfloor$ for noninteger t. The generalized Poincaré inequality bounds the full H^t norm of this remainder by its top-order seminorm. The embedding $H^t(D) \subset C(D)$ bounds the supremum norm. Affine rescaling then supplies $a^{t-d/2}$.

Stability and polynomial reproduction give the same bound, up to a fixed constant, for the error of the fit to exact values of G on a good D'. If a base cube uses an ancestor of side $a = 2^\ell h$ with $\ell \geq 1$, the preceding ancestor of side $a/2$ was not good. Its selection probability is therefore at most

$$\pi_\ell = \min\left(1, J \exp\left[-cm(2^{\ell-1}h)^d\right]\right), \qquad \pi_0 = 1.$$

At each level, each cube D' contains at most a^d/h^d base cubes. Applying (D.12) on their possible parent, then summing the squared integrated errors, gives

$$\begin{aligned}\mathbb{E}\,\|P_m G - G\|^2_{L^2(D)} &\leq C \sum_{\ell \geq 0} \pi_\ell (2^\ell h)^{2t} \sum_{D' \text{ at level } \ell} |G|^2_{H^t(D')} + Ce^{-cm}\,\|G\|^2_{L^2(D)} \\ &\leq Ch^{2t}\,\|G\|^2_{H^t(D)}\,. \end{aligned} \tag{D.13}$$

The sums at each level are over disjoint cubes and are bounded by the global seminorm, also at fractional orders. The numerical series is finite uniformly in m, since $mh^d \geq 1$ and exponential decay in $2^{\ell d}$ dominates $2^{2t\ell}$. If the root cube is not good, (D.10) bounds its contribution by the last term. This term is also at most $Ch^{2t}\,\|G\|^2_{H^t}$. For the sharper noise bound, let N be the count in one neighborhood of a cube of side a, with sampling probability $p_a \geq ca^d$. Since $\mathsf{N} \sim \mathrm{Bin}(m, p_a)$,

$$\mathbb{E}\left[\frac{\mathbf{1}_{\{\mathsf{N}\geq 1\}}}{\mathsf{N}}\right] \leq 2\mathbb{E}\left[\frac{1}{\mathsf{N}+1}\right] = \frac{2[1-(1-p_a)^{m+1}]}{(m+1)p_a} \leq \frac{C}{ma^d}. \tag{D.14}$$

For a base cube, drop its ancestor-selection indicator when bounding each nonnegative noise-variance contribution in (D.11). Applying (D.14) at every possible ancestor and then integrating over the base cubes gives

$$\mathbb{E}\,\|P_m(\varsigma\boldsymbol{\xi})\|^2_{L^2(D)} \leq C\varsigma^2 \sum_{\ell \geq 0} \frac{1}{m(2^\ell h)^d} \leq \frac{C'\varsigma^2}{mh^d}. \tag{D.15}$$

Conditional on the design, $P_m(\varsigma\boldsymbol{\xi})$ has mean zero. Thus its variance adds to the expected squared bias without a cross term.

For the sphere, take a finite collection of larger smooth coordinate cubes and smaller cubes whose images cover S^{d-1}, with a smooth partition of unity supported in the larger charts. Include the full strike interval by an affine coordinate. Chart Jacobians and their reciprocals are bounded, and Q_0 has a coordinate density bounded below. Construct P_m in each larger chart and combine the outputs using the partition of unity. Coordinate Sobolev norms are bounded by the corresponding global norm. Squared errors and conditional variances of the finite sum are bounded by a constant multiple of the sum of the individual bounds. The strike endpoints are faces of the coordinate cubes and require no samples outside the observed interval. This proves the analogue of (D.13) on $\mathcal{Z}$.

For $f \in \mathcal{F}_\alpha(L)$, subtract a smooth probability density f_* supported inside B_r. Lemma 3.3 applies to $f - f_*$, and C_{f_*} is smooth on $\mathcal{Z}$. Consequently

$$\|C_f\|_{H^t(\mathcal{Z})} \leq C(1+L) \leq C'L.$$

The last inequality uses the positive lower bound on the L^2 norm of a probability density on B_r. Since C_f takes values in $[0, 2R]$, truncating the output pointwise to this interval decreases its squared error. This proves the first inequality in (5.11); taking $h \asymp m^{-1/d}$ proves (5.9).

For (5.10), fix $\boldsymbol{x} \in B_r$ and write $G_{\boldsymbol{x}}(\boldsymbol{u}, k) = (\boldsymbol{u} \cdot \boldsymbol{x} - k)_+$. In each coordinate chart, its nonsmooth set is the graph $k = \boldsymbol{u} \cdot \boldsymbol{x}$. The volume of its a-neighborhood is at most Ca, uniformly in $\boldsymbol{x}$: the defining function $\boldsymbol{u} \cdot \boldsymbol{x} - k$ has a nonzero strike derivative and uniformly bounded angular derivatives. Cubes of side a meeting this graph have total volume at most Ca. On such cubes, the Lipschitz property and reproduction of constants give a fit error at most Ca. On the remaining cubes, $G_{\boldsymbol{x}}$ is either zero or a smooth function with uniformly bounded second derivatives; reproduction of affine polynomials gives an error at most Ca^2. At each level, the squared integrated error is therefore bounded by

$$C\left(a\, a^2 + a^4\right) \le Ca^3.$$

Multiplication by the same selection probabilities π_ℓ and summation over levels yield

$$\sup_{\boldsymbol{x} \in B_r} \mathbb{E}\, \|P_m G_{\boldsymbol{x}} - G_{\boldsymbol{x}}\|^2_{L^2(Q_0)} \le Ch^3. \tag{D.16}$$

Before truncation, linearity gives $P_m C_\mu = \int P_m G_{\boldsymbol{x}}\, \mu(\mathrm{d}\boldsymbol{x})$ for exact inputs. Jensen's inequality in $L^2(Q_0)$ and Tonelli's theorem imply

$$\mathbb{E}\, \|P_m C_\mu - C_\mu\|^2_{L^2(Q_0)} \le \int \mathbb{E}\, \|P_m G_{\boldsymbol{x}} - G_{\boldsymbol{x}}\|^2_{L^2(Q_0)}\, \mu(\mathrm{d}\boldsymbol{x}) \le Ch^3.$$

The noise contribution is bounded by (D.15), and truncation again decreases the error. This proves the second inequality in (5.11), and $h \asymp m^{-1/d}$ gives (5.10). For $Q \ge q_* Q_0$, every neighborhood in (D.10) still has probability at least $cq_* a^d$. □

PROOF OF PROPOSITION 5.7. Let $\mathcal{D}_0$ denote the first-group observations. Conditional on $\mathcal{D}_0$, Lemma 3.4 and independence give

$$\mathbb{E}\left[\widehat{\theta}^{\mathrm{cv}}_{\boldsymbol{k}} \mid \mathcal{D}_0\right] = a_{\boldsymbol{k}} + \int \widetilde{G}_{\boldsymbol{k}}(C_\mu + e)\, \mathrm{d}Q_0 = \theta_{\boldsymbol{k}}(\mu) + b_{\boldsymbol{k}}, \qquad b_{\boldsymbol{k}} = \int \widetilde{G}_{\boldsymbol{k}}\, e\, \mathrm{d}Q_0. \tag{D.17}$$

Write $R_0 = C_\mu + e - \widehat{C}$, so that $\mathbb{E}\, \|R_0\|^2_{L^2(Q_0)} \le 2v_n(\mu) + 2\epsilon^2$. The conditional variance obeys

$$\mathbb{E}\left[\left|\widehat{\theta}^{\mathrm{cv}}_{\boldsymbol{k}} - \theta_{\boldsymbol{k}}(\mu) - b_{\boldsymbol{k}}\right|^2 \mid \mathcal{D}_0\right] \le \frac{1}{n_1} \int \left|\widetilde{G}_{\boldsymbol{k}}(z)\right|^2 \left[R_0(z)^2 + \varsigma^2\right]\, \mathrm{d}Q_0(z). \tag{D.18}$$

Apply (D.7) inside the integral, and then take expectation over $\mathcal{D}_0$. Let f_M be the deterministic filtered version of μ and

$$B(\boldsymbol{x}) = \frac{1}{T^d} \sum_{\boldsymbol{k} \ne 0} w(\boldsymbol{k}/M)\, b_{\boldsymbol{k}}\, e^{2\pi \mathrm{i} \boldsymbol{k} \cdot \boldsymbol{x}/T}.$$

This gives

$$\mathbb{E}\left\|\widetilde{f}^{\mathrm{cv}}_M - f_M - B\right\|^2_{H^{-1}(\mathbb{T}^d_T)} \le \frac{CM^{2d+1}}{n}\left[\varsigma^2 + \epsilon^2 + v_n(\mu)\right]. \tag{D.19}$$

Since B has zero mass, we may restrict to mean-zero 1-Lipschitz functions F on the torus. Then $\widehat{F}(\mathbf{0}) = 0$ and $\|F\|_{H^1(\mathbb{T}^d_T)} \le C$. Since $\psi \mapsto \widetilde{G}_\psi$ is linear,

$$\int FB\, \mathrm{d}\boldsymbol{x} = \int_{\mathcal{Z}} \widetilde{G}_\Psi\, e\, \mathrm{d}Q_0, \qquad \Psi = \chi\,(\Phi_M * F),$$

with Φ_M the convolution kernel in (D.4). Lemma 3.4, $|e| \le \epsilon$, and $\|\Phi_M * F\|_{H^s(\mathbb{T}^d_T)} \le CM^{s-1}\|F\|_{H^1(\mathbb{T}^d_T)}$ give

$$\|B\,\mathrm{d}\boldsymbol{x}\|_{\mathrm{Lip}^*} \le C\epsilon M^{s-1}. \tag{D.20}$$

The smooth and arbitrary-law bias bounds are exactly (D.3) and (D.4). The true law is feasible in (5.13), so

$$W_1(\widehat{\mu}^{\mathrm{cv}}_M, \mu) \le 2\left\|\widetilde{f}^{\mathrm{cv}}_M\,\mathrm{d}\boldsymbol{x} - \mu\right\|_{\mathrm{Lip}^*} + e^{-n}.$$

Use (D.1), (D.19), (D.20), and Cauchy–Schwarz to obtain both inequalities. □

LEMMA D.2 (Curved wave packets). *Let $\alpha \ge 0$, let f_0 be the base density of Lemma C.4, and let $L > L_0$ as in Theorem 4.5. There are $c, C > 0$ and N_0 such that for every integer $N \ge N_0$ there are $f^{\pm}_N \in \mathcal{F}_\alpha(L)$, whose one-dimensional coordinate marginals equal those of f_0, with*

$$W_1(f^+_N\,\mathrm{d}\boldsymbol{x}, f^-_N\,\mathrm{d}\boldsymbol{x}) \ge cN^{-\alpha-1}, \qquad \sup_{\boldsymbol{u}\in S^{d-1}, k\in\mathbb{R}} \left|C_{f^+_N}(\boldsymbol{u},k) - C_{f^-_N}(\boldsymbol{u},k)\right| \le CN^{-\alpha-s}.$$

PROOF. Fix $\boldsymbol{x}_1$ and $\rho_1 > 0$ such that f_0 is bounded below on $B(\boldsymbol{x}_1, \rho_1)$, let $\boldsymbol{a} = (\boldsymbol{e}_1 + \boldsymbol{e}_2)/\sqrt{2}$, and let $\boldsymbol{x}_c = \boldsymbol{x}_1 - 8\rho_1\boldsymbol{a}$. Let $\chi \in C^\infty_c(B(\boldsymbol{x}_1, \rho_1))$ be nonnegative and nonzero, $\varphi(\boldsymbol{x}) = |\boldsymbol{x} - \boldsymbol{x}_c|$, and

$$\eta_N = bN^{-\alpha-2}\partial_1\partial_2\left[\chi\cos(N\varphi)\right], \qquad f^{\pm}_N = f_0 \pm \eta_N.$$

Since the function $\chi\cos(N\varphi)$ has fixed support and derivatives of order m bounded by C_mN^m, $\|\eta_N\|_{H^m} \le C_m bN^{m-\alpha}$ for integers m, interpolation gives $\|\eta_N\|_{H^\alpha} \le Cb$, and $\|\eta_N\|_\infty \le CbN^{-\alpha}$. Since η_N is a derivative $\partial_1\partial_2$ of a compactly supported function, it has zero mass and first moments, and integrating out x_1 or x_2 shows that its one-dimensional coordinate marginals vanish. Choosing b so small that $f_0 \pm \eta_N \ge 0$ and $L_0 + Cb \le L$, where $L_0 = \|f_0\|_{H^\alpha}$, gives $f^{\pm}_N \in \mathcal{F}_\alpha(L)$.

Separation. Let $\tilde{\chi} \in C^\infty_c(B(\boldsymbol{x}_1, 2\rho_1))$ equal one on $B(\boldsymbol{x}_1, \rho_1)$ with $0 \le \tilde{\chi} \le 1$, and $F = (2N)^{-1}\tilde{\chi}\cos(N\varphi)$. Since $|\nabla\varphi| = 1$, F is 1-Lipschitz for $N \ge \|\nabla\tilde{\chi}\|_\infty$. On $\operatorname{supp}\chi$, the unit vector $\nabla\varphi$ makes an angle at most $\arcsin\frac{1}{8}$ with $\boldsymbol{a}$, hence $\partial_1\varphi\,\partial_2\varphi \ge \frac{1}{2}\cos(2\arcsin\frac{1}{8}) \ge \frac{1}{4}$, and $\partial_1\partial_2 F = -\frac{N}{2}\cos(N\varphi)\,\partial_1\varphi\,\partial_2\varphi - \frac{1}{2}\sin(N\varphi)\,\partial_1\partial_2\varphi$. Integrating by parts twice,

$$\begin{aligned}\int F\eta_N\,\mathrm{d}\boldsymbol{x} &= bN^{-\alpha-2}\int \chi\cos(N\varphi)\,\partial_1\partial_2 F\,\mathrm{d}\boldsymbol{x}\\ &= -\frac{b}{4}N^{-\alpha-1}\int \chi\,\partial_1\varphi\,\partial_2\varphi\left(1+\cos(2N\varphi)\right)\mathrm{d}\boldsymbol{x}\\ &\quad -\frac{b}{4}N^{-\alpha-2}\int \chi\,\partial_1\partial_2\varphi\,\sin(2N\varphi)\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

One integration by parts along $\nabla\varphi$ bounds both oscillatory integrals by C/N. Hence $\left|\int F\eta_N\,\mathrm{d}\boldsymbol{x}\right| \ge \frac{b}{32}N^{-\alpha-1}\int\chi$ for large N, and since $\pm F$ are 1-Lipschitz, $W_1(f^+_N\,\mathrm{d}\boldsymbol{x}, f^-_N\,\mathrm{d}\boldsymbol{x}) \ge 2\left|\int F\eta_N\,\mathrm{d}\boldsymbol{x}\right| \ge \frac{b}{16}N^{-\alpha-1}\int\chi$.

Quotes. For $\boldsymbol{u} \in S^{d-1}$ and $k \in \mathbb{R}$, the distributional derivative $\partial_1\partial_2$ of $\boldsymbol{x} \mapsto (\boldsymbol{u}\cdot\boldsymbol{x} - k)_+$ is u_1u_2 times surface measure on the hyperplane $H_{\boldsymbol{u},k} = \{\boldsymbol{u}\cdot\boldsymbol{x} = k\}$. Hence

$$C_{f^+_N}(\boldsymbol{u},k) - C_{f^-_N}(\boldsymbol{u},k) = 2bN^{-\alpha-2}u_1u_2\int_{H_{\boldsymbol{u},k}}\chi\cos(N\varphi)\,\mathrm{d}S,$$

and $|u_1 u_2| \le 1$. Let $\delta = \operatorname{dist}(\boldsymbol{x}_c, H_{\boldsymbol{u},k})$, and use coordinates $\boldsymbol{y} \in \mathbb{R}^{d-1}$ on $H_{\boldsymbol{u},k}$ centered at the foot of the perpendicular from $\boldsymbol{x}_c$, so that $\varphi = (\delta^2 + |\boldsymbol{y}|^2)^{1/2}$. On $\operatorname{supp}\chi$, $|\boldsymbol{x} - \boldsymbol{x}_c| \in [7\rho_1, 9\rho_1]$. If $\delta \le 4\rho_1$, the tangential gradient $\boldsymbol{y}/\varphi$ of φ has norm at least $(1 - 16/49)^{1/2}$ on $H_{\boldsymbol{u},k} \cap \operatorname{supp}\chi$, and repeated integration by parts gives $O(N^{-m})$ for every m. If $\delta > 4\rho_1$, the phase has the single critical point $\boldsymbol{y} = 0$, with Hessian $\delta^{-1} I_{d-1}$ and $|\nabla\varphi| \ge |\boldsymbol{y}|/(9\rho_1)$ on the support, and the phase and amplitude have derivatives bounded uniformly in $(\boldsymbol{u}, k)$. The stationary phase estimate [7, Theorem 7.7.5] gives $\left|\int_{H_{\boldsymbol{u},k}} \chi \cos(N\varphi)\, \mathrm{d}S\right| \le C N^{-(d-1)/2}$ uniformly in $(\boldsymbol{u}, k)$. Thus the quote difference is at most $CbN^{-\alpha-2-(d-1)/2} = CbN^{-\alpha-s}$. □

The pair in Lemma D.2 lies in the fiber $\mathcal{K}^{\epsilon}_{n,\alpha,L}(z, v)$ of every design, with v the average of the two quote vectors and $\epsilon = CN^{-\alpha-s}$. Hence, for each $L > L_0$, the exponent ϑ of ϵ in (4.6) is optimal.

PROOF OF THEOREM 5.1. Before clipping to $[0, 2R]$, the pilot of Lemma 5.6 is linear in the observed prices, and on each cube its fit is bounded by a constant times the largest absolute data value in that cube. The error surface therefore changes the unclipped pilot by at most $C\epsilon$ uniformly on $\mathcal{Z}$, and clipping does not increase the distance to C_f. Hence (5.9) holds with the additional term $C\epsilon^2$. For $n \ge 2$, Proposition 5.7 gives

$$\sup_{f \in \mathcal{F}_\alpha(L)} \sup_{|e| \le \epsilon} \mathbb{E} W_1(\widehat{\mu}^{\mathrm{cv}}_M, f\, \mathrm{d}\boldsymbol{x}) \le C \Big[L M^{-\alpha-1} + \epsilon M^{s-1} + \big(\varsigma n^{-\frac{1}{2}} + \epsilon n^{-\frac{1}{2}} + L n^{-\frac{1}{2}-\frac{t}{d}}\big) M^{d+\frac{1}{2}} \Big] + e^{-n}. \tag{D.21}$$

Take

$$M = \max\left(1, \left\lfloor \min\left(n^{\frac{1}{d}}, \left(\frac{n}{\varsigma^2}\right)^{\frac{1}{D_\alpha}}, \left(\frac{L}{\epsilon}\right)^{\frac{1}{\alpha+s}} \right) \right\rfloor \right), \tag{D.22}$$

where the second and third entries are $+\infty$ at $\varsigma = 0$ and $\epsilon = 0$. If $M \le n^{1/d}$, then

$$n^{-\frac{1}{2}-\frac{t}{d}} M^{d+\frac{1}{2}} = M^{-\alpha-1} n^{-D_\alpha/(2d)} M^{D_\alpha/2} \le M^{-\alpha-1}, \tag{D.23}$$

and $\epsilon n^{-1/2} M^{d+1/2} \le \epsilon M^{s-1}$, since $M^{d+1/2-(s-1)} = M^{d/2} \le n^{1/2}$. Likewise, $M \le (n/\varsigma^2)^{1/D_\alpha}$ implies $\varsigma n^{-1/2} M^{d+1/2} \le M^{-\alpha-1}$, and $M \le (L/\epsilon)^{1/(\alpha+s)}$ implies $\epsilon M^{s-1} \le L M^{-\alpha-1}$. Thus (D.21) is at most a constant times

$$n^{-\frac{\alpha+1}{d}} + \left(\frac{\varsigma^2}{n}\right)^{\frac{\alpha+1}{D_\alpha}} + \epsilon^{\vartheta}.$$

The exponentially small projection tolerance is absorbed for every α. For the finitely many remaining sample sizes, and when the minimum in (D.22) is below one, the estimator δ_0 and the bound $W_1 \le 2r$ suffice after increasing the constant.

For the lower bound, the error surface $e = 0$ is admissible, hence $\mathfrak{R}_n(\mathcal{F}_\alpha(L), \varsigma, \epsilon) \ge \mathfrak{R}_n(\mathcal{F}_\alpha(L), \varsigma)$. For the first lower term, use the localized family of Theorem C.6, with $M_0 \asymp n$ coefficients. The data are $C_{f_0}(Z_i) + (T_Z \boldsymbol{\theta})_i + \varsigma\xi_i$, hence Theorem 4.6 applies with the exogenous shift $a_i = \varsigma\xi_i$ and gives

$$\mathfrak{R}_n(\mathcal{F}_\alpha(L), \varsigma) \ge c n^{-(\alpha+1)/d} \tag{D.24}$$

uniformly in ς, including zero.

For $\varsigma > 0$, use the localized sign family of Proposition 5.5. Adjacent alternatives at spatial scale Δ satisfy, by Lemma A.1,

$$H^2(P_{\boldsymbol{\theta}}^{(n)}, P_{\boldsymbol{\theta}'}^{(n)}) \le \frac{nI_\xi c_\psi b^2}{4\varsigma^2}\Delta^{D_\alpha}.$$

Choose $\Delta = c_0(\varsigma^2/n)^{1/D_\alpha}$. A fixed $c_0 > 0$, depending on $\overline{\varsigma}$ and the support geometry, ensures both that the bump construction applies for every n and that this distance is at most $1/16$. There are $M_0 \asymp \Delta^{-d}$ cells. The paired sign-testing argument in the proof of Proposition 5.5, with (C.8) and (D.6), gives

$$\mathfrak{R}_n(\mathcal{F}_\alpha(L), \varsigma) \ge c\Delta^{\alpha+1} = c'\left(\frac{\varsigma^2}{n}\right)^{\frac{\alpha+1}{D_\alpha}}. \tag{D.25}$$

For the systematic term, let $\epsilon_1 > 0$ be small and $0 < \epsilon \le \epsilon_1$, and take the pair of Lemma D.2 with $N = \lceil (C/\epsilon)^{1/(\alpha+s)} \rceil$, so that $\left| C_{f_N^+} - C_{f_N^-} \right| \le \epsilon$ on $\mathcal{Z}$. The error surfaces $e_\pm = \mp\frac{1}{2}(C_{f_N^+} - C_{f_N^-})$ satisfy $|e_\pm| \le \epsilon$, and $C_{f_N^+} + e_+ = C_{f_N^-} + e_-$. The data therefore have the same law under (f_N^+, e_+) and (f_N^-, e_-), and every estimator has risk at least $\frac{1}{2}W_1(f_N^+ \,\mathrm{d}\boldsymbol{x}, f_N^- \,\mathrm{d}\boldsymbol{x}) \ge cN^{-\alpha-1} \ge c'\epsilon^\vartheta$ under one of them. For $\epsilon_1 < \epsilon \le \overline{\epsilon}$, monotonicity in ϵ gives the bound with a smaller constant. The largest of the three lower bounds is at least a third of their sum. □

PROOF OF COROLLARY 5.2. Equate the powers in (5.2) and use $2\alpha + 2d + 3 = 2[\alpha + (d+3)/2] + d$ and $(\alpha+1)/\vartheta = \alpha + (d+3)/2$. Substitution of $\varsigma_n = n^{-\gamma}$ gives the second assertion. □

PROOF OF COROLLARY 5.8. Set $q = s - 1$. There is a linear smoothing operator J_h on $\mathcal{Z}$ such that

$$\|J_h G\|_{H^q} \le Ch^{-q}\|G\|_{L^2}, \qquad \|J_h G - G\|_{H^q} \le Ch^{t-q}\|G\|_{H^t}. \tag{D.26}$$

To construct it, use a finite chart partition η_j, larger chart cutoffs χ_j equal to one on $\operatorname{supp}\eta_j$, and extension operators E_j from the coordinate cubes to Euclidean space. Each E_j is bounded simultaneously in L^2 and H^t. On a cube, finite linear combinations of reflections at each face, matching normal derivatives through an integer order above t, followed by a smooth cutoff, give such an extension; interpolation supplies the intermediate orders. In coordinates, the jth summand of $J_h G$ is

$$\chi_j \mathcal{F}^{-1}\left[w(h\boldsymbol{\xi})\,\mathcal{F}\left(E_j(\eta_j G)\right)(\boldsymbol{\xi})\right],$$

restricted to its chart and extended by zero off that chart. The sum of the unsmoothed summands is G. The multiplier bounds

$$(1+|\boldsymbol{\xi}|^2)^{q/2}\,|w(h\boldsymbol{\xi})| \le Ch^{-q}, \qquad (1+|\boldsymbol{\xi}|^2)^{(q-t)/2}\,|1 - w(h\boldsymbol{\xi})| \le Ch^{t-q}$$

and boundedness of smooth multiplication and coordinate changes prove (D.26). Strike endpoints are handled by the same cube extensions; no boundary condition on G is imposed.

For a dyadic bandwidth $h \in [n^{-1/d}, 1]$, let $\widehat{C}_{n,h}$ be the pilot of Lemma 5.6 built from all n observations, and set $\widetilde{C}_n = J_h\widehat{C}_{n,h}$. Let $\mathcal{S} = \{C_f : f \in \mathcal{F}_\alpha(L)\} \subset H^q(\mathcal{Z})$. The set $\mathcal{S}$ is convex since $\mathcal{F}_\alpha(L)$ is convex and $\mathcal{A}$ is linear. The class $\mathcal{F}_\alpha(L)$ is weakly compact in H^α. By Lemma 3.3 and the compact embedding of $H^t(\mathcal{Z})$ into $H^q(\mathcal{Z})$, its image $\mathcal{S}$ is compact. Let $\widehat{C}_n$ be the metric projection of $\widetilde{C}_n$ onto $\mathcal{S}$ in $H^q(\mathcal{Z})$. This projection is unique and nonexpansive. The inverse bound (3.6) makes $C_f \mapsto f\,\mathrm{d}\boldsymbol{x}$ continuous from $\mathcal{S}$ to W_1. Hence the law $\widehat{\mu}_n = f\,\mathrm{d}\boldsymbol{x}$ with $C_f = \widehat{C}_n$ is a measurable probability-valued estimator.

As in the proof of Theorem 5.1, the error surface changes the unclipped pilot by at most $C\epsilon$ uniformly on $\mathcal{Z}$. Hence (5.11) holds with the additional term $C\epsilon^2$, and (D.26) gives

$$\begin{aligned}\mathbb{E}\left\|\widetilde{C}_n - C_f\right\|_{H^q}^2 &\le 2\mathbb{E}\left\|J_h(\widehat{C}_{n,h} - C_f)\right\|_{H^q}^2 + 2\left\|J_h C_f - C_f\right\|_{H^q}^2 \\ &\le C\left[L^2 h^{2(t-q)} + \frac{\varsigma^2}{nh^{d+2q}} + \frac{\epsilon^2}{h^{2q}}\right].\end{aligned} \tag{D.27}$$

Since $C_f \in \mathcal{S}$ and the projection is nonexpansive, $\left\|\widehat{C}_n - C_f\right\|_{H^q} \le \left\|\widetilde{C}_n - C_f\right\|_{H^q}$, and (3.6) gives (5.16). Take h of order $\max(n^{-1/d}, (\varsigma^2/n)^{1/D_\alpha}, (\epsilon/L)^{1/(\alpha+s)})$, rounded to a dyadic side and capped at one. At $h \ge (\varsigma^2/n)^{1/D_\alpha}$ the second term is at most $(\varsigma^2/n)^{(\alpha+1)/D_\alpha}$, since $\frac{1}{2}D_\alpha - d - \frac{1}{2} = \alpha + 1$. At $h \ge (\epsilon/L)^{1/(\alpha+s)}$ the third term is at most $L^{1-\vartheta}\epsilon^{\vartheta}$, since $1 - (s-1)/(\alpha+s) = \vartheta$. The first term at the chosen h is a constant multiple of the largest of the three rates, which gives the upper side of (5.2). Boundedly many small sample sizes and a capped h are handled by $W_1 \le 2r$. At zero noise and $\epsilon = 0$, $h \asymp n^{-1/d}$ proves the upper side of (5.17).

The exact lower bound is Theorem C.6, and the conditioning argument for (D.24) preserves it under noise. If $Q \le q^* Q_0$, adjacent sign alternatives satisfy

$$H^2(P^{(n)}_{\boldsymbol{\theta},Q}, P^{(n)}_{\boldsymbol{\theta}',Q}) \le \frac{nI_\xi}{4\varsigma^2}\left\|C_{f_{\boldsymbol{\theta}}} - C_{f_{\boldsymbol{\theta}'}}\right\|^2_{L^2(Q)} \le \frac{nI_\xi q^* c_\psi b^2}{4\varsigma^2}\Delta^{2\alpha+2d+3}.$$

The scale choice and testing proof of (D.25) therefore apply with constants depending on q^*. The argument for the systematic term in the proof of Theorem 5.1 applies to every design, since the two data laws coincide. □

PROOF OF COROLLARY 5.9. By Lemma C.4, every $f_{\boldsymbol{\theta}} = f_0 + H_{\boldsymbol{\theta}}$ has the coordinate marginals of f_0, so the proofs of (D.24) and (D.25) apply verbatim on $\mathcal{P}_0$. For the systematic term, the pair of Lemma D.2 lies in $\mathcal{P}_0$, and the last argument in the proof of Theorem 5.1 applies. The upper bounds hold since $\mathcal{P}_0 \subseteq \mathcal{F}_\alpha(L)$. □

With exact quotes, $\varsigma = 0$ and $\epsilon = 0$, an estimate can also be made exactly price feasible without changing its risk order. For any estimator $\widehat{\mu}$, choose a nearest law $\overline{\mu}$ to $\widehat{\mu}$ in the unrestricted exact quote fiber on B_r. A nearest law exists by compactness. The constraint graph is closed in the design, price vector, and candidate law. Its compact sections form a measurable correspondence on the set of feasible data. The measurable minimum theorem therefore supplies a measurable choice. Since the true law belongs to this fiber,

$$W_1(\overline{\mu}, \mu) \le W_1(\overline{\mu}, \widehat{\mu}) + W_1(\widehat{\mu}, \mu) \le 2W_1(\widehat{\mu}, \mu).$$

Theorem 5.1 and Corollary 5.8 carry no logarithmic factor since the local estimate (D.13) averages the bad-cell probabilities for each pricing surface, whereas (4.5) controls all feasible price vectors on a common design event, as discussed in Section 4.3.

PROOF OF THEOREM 5.11. The upper bound follows from (5.10) and (5.15):

$$\sup_{\mu} \mathbb{E}W_1(\widehat{\mu}^{\mathrm{cv}}_M, \mu) \le C\left[M^{-1} + n^{-\frac{1}{2}}\left(\varsigma + n^{-3/(2d)}\right)M^{d+\frac{1}{2}}\right] + e^{-n}.$$

Choose M, up to integer rounding and a lower truncation at one, of order

$$\left(\frac{n}{\varsigma^2 + n^{-3/d}}\right)^{\frac{1}{2d+3}}.$$

The two displayed terms are then bounded by $C[(\varsigma^2 + n^{-3/d})/n]^{1/(2d+3)}$. Small n are covered as in Theorem 5.1. Apply that theorem with $\alpha = 0$ and one sufficiently large L to obtain the lower side, since this density class is contained in $\mathcal{P}(B_r)$. If $\varsigma^2 \geq c_0 n^{-3/d}$, the noise term bounds the exact upper term up to a fixed factor, while it is itself a lower bound. This proves the final assertion. □

APPENDIX E: PROOFS FOR SECTION 6

PROOF OF PROPOSITION 6.1. The left side of (6.1) equals

$$\sup_{\mathrm{Lip}(g)\leq 1} \sup_{\nu,\nu'\in\mathcal{K}} \int g \,\mathrm{d}(\nu - \nu') = \sup_{\nu,\nu'\in\mathcal{K}} \sup_{\mathrm{Lip}(g)\leq 1} \int g \,\mathrm{d}(\nu - \nu') = \sup_{\nu,\nu'\in\mathcal{K}} W_1(\nu, \nu').$$

For the convergence of endpoints, the functional $\nu \mapsto \int g \,\mathrm{d}\nu$ is continuous on the compact space $\mathcal{P}(B_R)$ and hence uniformly continuous. The Hausdorff convergence in Theorem 2.3 therefore gives convergence of its maximum and minimum over the admissible sets. □

PROOF OF LEMMA 6.2. Let μ, ν be admissible in the definition of $\varpi_g(\epsilon)$ and write $\eta = \mu - \nu$. Then $\eta(B_R) = 0$ and $\|\eta\| \leq 2$. For $h = g - \mathcal{A}^* m$ and c the midpoint of the range of h, $\int h \,\mathrm{d}\eta = \int (h - c) \,\mathrm{d}\eta \leq \frac{1}{2} \operatorname{osc}(h) \,\|\eta\| \leq \operatorname{osc}(h)$. Fubini's theorem gives $\int \mathcal{A}^* m \,\mathrm{d}\eta = \int C_\eta \,\mathrm{d}m \leq \epsilon \,\|m\|$, since $|C_\eta| \leq \epsilon$ on $\mathcal{Z}_E$. Adding the two bounds and taking the supremum over admissible pairs proves (6.3). □

LEMMA E.1 (Continuity under identification). *For every $g \in C(B_R)$, the modulus ϖ_g is nondecreasing. If the pricing surfaces on $\mathcal{Z}_E$ identify $\mathcal{P}(B_R)$, then $\varpi_g(\epsilon) \to 0$ as $\epsilon \downarrow 0$.*

PROOF. Monotonicity is immediate from the definition. Suppose that identification holds and that there are $\delta > 0$, $\epsilon_j \downarrow 0$, and $\mu_j, \nu_j \in \mathcal{P}(B_R)$ with $\sup_{\mathcal{Z}_E} \left| C_{\mu_j} - C_{\nu_j} \right| \leq \epsilon_j$ and $\int g \,\mathrm{d}(\mu_j - \nu_j) \geq \delta$. By compactness of $\mathcal{P}(B_R)$ in W_1, a subsequence satisfies $\mu_j \to \mu$ and $\nu_j \to \nu$. By (2.3),

$$\sup_{\mathcal{Z}_E} |C_\mu - C_\nu| \leq \epsilon_j + W_1(\mu_j, \mu) + W_1(\nu_j, \nu) \longrightarrow 0,$$

hence $C_\mu = C_\nu$ on $\mathcal{Z}_E$ and $\mu = \nu$ by identification. Continuity of $\nu \mapsto \int g \,\mathrm{d}\nu$ in W_1 gives $\int g \,\mathrm{d}(\mu - \nu) \geq \delta$, a contradiction. □

PROOF OF THEOREM 6.3. (i) Let $D_{\mathcal{C}} = \{\mu - \nu : \mu, \nu \in \mathcal{C}\}$. Since $\mathcal{C}$ is a closed subset of the compact space $\mathcal{P}(K)$, $D_{\mathcal{C}}$ is convex, symmetric, and weak-$*$ compact in the finite signed measures on K, and $\operatorname{osc}_{\mathcal{C}}(h) = \sup_{\eta \in D_{\mathcal{C}}} \int h \,\mathrm{d}\eta$. Apply the Fenchel–Rockafellar theorem [6, Ch. III, Theorem 4.1] to $V = L^2(Q)$, $Y = C(K)$, the bounded operator $\Lambda w = \mathcal{A}^*(wQ)$, $F(w) = \epsilon \,\|w\|_{L^2(Q)}$, and $G(h) = \operatorname{osc}_{\mathcal{C}}(g - h)$. Both functions are convex and finite, and G is 2-Lipschitz in the uniform norm. Hence the infimum in (6.4) equals $\max_\eta \{-F^*(\Lambda^* \eta) - G^*(-\eta)\}$ over the finite signed measures η on K. The functional $\Lambda^* \eta$ acts by $w \mapsto \int C_\eta w \,\mathrm{d}Q$ and has norm $\|C_\eta\|_{L^2(Q)}$, hence $F^*(\Lambda^* \eta)$ is 0 if $\|C_\eta\|_{L^2(Q)} \leq \epsilon$ and $+\infty$ otherwise. Substituting $q = g - h$,

$$G^*(-\eta) = -\int g \,\mathrm{d}\eta + \sup_{q \in C(K)} \left\{ \int q \,\mathrm{d}\eta - \sup_{\eta' \in D_{\mathcal{C}}} \int q \,\mathrm{d}\eta' \right\}.$$

The last supremum is 0 if $\eta \in D_{\mathcal{C}}$. If $\eta \notin D_{\mathcal{C}}$, the Hahn–Banach theorem separates η from the weak-$*$ closed convex set $D_{\mathcal{C}}$ by some $q \in C(K)$, and scaling q makes the supremum $+\infty$.

The dual problem is therefore $\sup\{\int g\,\mathrm{d}\eta : \eta \in D_{\mathcal{C}},\ \|C_\eta\|_{L^2(Q)} \le \epsilon\} = \omega_g(\epsilon;\mathcal{C})$. Its constraint set is weak-$*$ compact, since $\eta \mapsto C_\eta(z)$ is weak-$*$ continuous for each z and Fatou's lemma makes $\eta \mapsto \|C_\eta\|_{L^2(Q)}$ lower semicontinuous, hence the supremum is attained.

(ii) Let $a_\xi = \sqrt{1 \vee I_\xi}$ and let $\mu, \nu \in \mathcal{C}$ satisfy $\|C_\mu - C_\nu\|_{L^2(Q)} \le \varsigma/(a_\xi\sqrt{n})$. By Lemma A.1, the joint laws of the n triples have total variation distance at most $\frac{1}{2}$. Let $a = \int g\,\mathrm{d}\mu$, $b = \int g\,\mathrm{d}\nu$, and $\phi = \mathbf{1}\{|\widehat{p} - a| \ge |a-b|/2\}$. Then $|\widehat{p} - a| \ge \frac{1}{2}|a-b|\,\phi$ and $|\widehat{p} - b| \ge \frac{1}{2}|a-b|\,(1-\phi)$, so

$$\mathbb{E}_\mu|\widehat{p} - a| + \mathbb{E}_\nu|\widehat{p} - b| \ge \tfrac{1}{2}|a-b|\,(1 - \mathrm{TV}) \ge \tfrac{1}{4}|a-b|\,,$$

and the larger of the two risks is at least $|a-b|/8$. The supremum over such pairs gives $\mathfrak{R}_n^Q(g;\mathcal{C}) \ge \omega_g(\varsigma/(a_\xi\sqrt{n});\mathcal{C})/8$. Since $\mathcal{C}$ is convex, convex combinations of feasible pairs are feasible. Hence ω_g is concave with $\omega_g(0) \ge 0$, and $\omega_g(\epsilon/a_\xi) \ge \omega_g(\epsilon)/a_\xi$.

For the upper bound, fix $w \in L^2(Q)$ and $c \in \mathbb{R}$. Fubini's theorem gives $\mathbb{E}\widehat{p}_w = c + \int wC_\mu\,\mathrm{d}Q = c + \int \mathcal{A}^*(wQ)\,\mathrm{d}\mu$, hence the bias of $\widehat{p}_w$ at μ is $\int (c + \mathcal{A}^*(wQ) - g)\,\mathrm{d}\mu$, which has absolute value at most $\frac{1}{2}\operatorname{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ))$ when c is the midpoint of the price range of $g - \mathcal{A}^*(wQ)$ over $\mathcal{C}$. Since $0 \le C_\mu \le 2R$,

$$\operatorname{Var}_\mu(\widehat{p}_w) \le \frac{1}{n}\mathbb{E}\left[w(Z_1)^2 Y_1^2\right] \le \frac{4R^2 + \varsigma^2}{n}\,\|w\|_{L^2(Q)}^2\,.$$

Hence, with $\varkappa_\varsigma = \sqrt{1 + 4R^2/\varsigma^2} \ge 1$,

$$\begin{aligned}\sup_{\mu\in\mathcal{C}} \mathbb{E}_\mu\left|\widehat{p}_w - \int g\,\mathrm{d}\mu\right| &\le \tfrac{1}{2}\operatorname{osc}_{\mathcal{C}}\left(g - \mathcal{A}^*(wQ)\right) + \varkappa_\varsigma\frac{\varsigma}{\sqrt{n}}\,\|w\|_{L^2(Q)} \\ &\le \varkappa_\varsigma\left[\operatorname{osc}_{\mathcal{C}}\left(g - \mathcal{A}^*(wQ)\right) + \frac{\varsigma}{\sqrt{n}}\,\|w\|_{L^2(Q)}\right].\end{aligned}$$

Taking the infimum over w and using (6.4) proves the upper bound. □

PROOF OF COROLLARY 6.4. (i) The proof of Theorem 6.3 bounds the bias of $\widehat{p}_w$ by $\frac{1}{2}\operatorname{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ))$ and its variance by $(4R^2 + \varsigma^2)\,\|w\|_{L^2(Q)}^2/n$, uniformly over $\mathcal{C}$. Chebyshev's inequality gives the coverage. The length $2\ell_w$ is at most $2\delta^{-1/2}\sqrt{1 + 4R^2/\varsigma^2}\,[\operatorname{osc}_{\mathcal{C}}(g - \mathcal{A}^*(wQ)) + (\varsigma/\sqrt{n})\,\|w\|_{L^2(Q)}]$, and (6.4) gives the bound.

(ii) Let $a_\xi = \sqrt{1 \vee I_\xi}$, and let $\mu, \nu \in \mathcal{C}$ satisfy $\|C_\mu - C_\nu\|_{L^2(Q)} \le \varsigma/(a_\xi\sqrt{n})$, so that the observation laws have total variation distance at most $\frac{1}{2}$ by Lemma A.1. With $a = \int g\,\mathrm{d}\mu$ and $b = \int g\,\mathrm{d}\nu$,

$$\Pr_\mu(a \in \widehat{I},\ b \in \widehat{I}) \ge \Pr_\mu(a \in \widehat{I}) - \Pr_\nu(b \notin \widehat{I}) - \tfrac{1}{2} \ge \tfrac{1}{2} - 2\delta,$$

hence $\mathbb{E}_\mu\left|\widehat{I}\right| \ge (\frac{1}{2} - 2\delta)\,|a-b|$. The supremum over such pairs and the concavity of ω_g established in the proof of Theorem 6.3 give the bound. □

APPENDIX F: PROOFS FOR SECTION 7

Throughout this section, E, $\boldsymbol{e}$, g_k and $\eta_{\boldsymbol{e}}$ are as in Section 7.3.

PROOF OF THEOREM 7.1. Rotate the construction in (B.9) to the direction $\boldsymbol{e}$:

$$H_N(\boldsymbol{x}) = bN^{-\alpha-2}(\boldsymbol{e}\cdot\nabla)^2\left(\zeta_0(\boldsymbol{x})\cos(N\boldsymbol{e}\cdot\boldsymbol{x})\right), \qquad f_N^{\pm} = f_0 \pm H_N.$$

The positivity, moment cancellation, and H^α bounds in the proof of Theorem 3.2 are unchanged under rotation. Testing against $N^{-1}\cos(N\boldsymbol{e}\cdot\boldsymbol{x})$ proves (7.2).

Condition (7.1) implies that there is $c_\kappa > 0$ with

$$|t\boldsymbol{u} \pm N\boldsymbol{e}| \geq c_\kappa(|t| + N), \qquad t \in \mathbb{R},\ \boldsymbol{u} \in E. \tag{F.1}$$

To see this, the squared left side is at least $t^2 + N^2 - 2(1-\kappa)\,|t|\,N$, which is bounded below by a positive multiple of $(|t|+N)^2$. The two derivatives in the perturbation cancel the denominator in (B.5), giving

$$\widehat{\mathcal{A}H_N}(\boldsymbol{u},t) = \frac{b}{2}N^{-\alpha-2}(\boldsymbol{u}\cdot\boldsymbol{e})^2\left(\widehat{\zeta}_0(t\boldsymbol{u}-N\boldsymbol{e}) + \widehat{\zeta}_0(t\boldsymbol{u}+N\boldsymbol{e})\right).$$

For every $P > 1$, Schwartz decay and (F.1) imply

$$\sup_{\boldsymbol{u}\in E}\int_{\mathbb{R}}\left|\widehat{\mathcal{A}H_N}(\boldsymbol{u},t)\right|\,\mathrm{d}t \leq C_P N^{-\alpha-2}\int_{\mathbb{R}}(1+N+|t|)^{-P}\,\mathrm{d}t \leq C'_P N^{-\alpha-1-P}.$$

Fourier inversion gives (7.3), with P chosen sufficiently large for each A. If a Hölder inequality with exponent γ held, choose A with $A\gamma > \alpha + 1$ and combine it with (7.2); the resulting inequality fails as $N \to \infty$. Every L^p norm under a probability design is bounded by the supremum norm, giving the final assertion. □

PROOF OF COROLLARY 7.2. The set E contains a relatively open patch, hence Theorem 2.2 gives identification. For $\boldsymbol{e} = (\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}$ and $\boldsymbol{u} \in E$, $|\boldsymbol{u}\cdot\boldsymbol{e}| = |u_1 - u_2|/\sqrt{2} \leq 1/\sqrt{2}$. Thus (7.1) holds with $\kappa = 1 - 1/\sqrt{2}$. □

PROOF OF PROPOSITION 7.3. (i) For $y \in [-R, R]$, the spanning formula [2, 3] gives $G(y) = G(-R) + G'(-R+)(y+R) + \int_{(-R,R)}(y-k)_+ G''(\mathrm{d}k)$, and $\boldsymbol{v}\cdot\boldsymbol{x} + R = \phi_{(\boldsymbol{v},-R)}(\boldsymbol{x})$ on B_R. Hence $g = G(-R) + \mathcal{A}^*m$ with $\|m\|$ equal to the stated constant, and Lemma 6.2 applies.

(ii) Let $N = \lceil\sqrt{2R/\epsilon}\rceil$ and $h = 2R/N \leq \sqrt{2R\epsilon}$. The piecewise linear interpolant I of G on the uniform grid of mesh h satisfies $\|G - I\|_\infty \leq h$, and its slopes $a_1, \ldots, a_N$ lie in $[-1, 1]$. By (i), $I(\boldsymbol{v}\cdot\boldsymbol{x}) = c + \mathcal{A}^*m$ with $\|m\| \leq |a_1| + \sum_{j\geq 2}|a_j - a_{j-1}| \leq 2N - 1$. Lemma 6.2 gives $\varpi_g(\epsilon) \leq 2h + \epsilon(2N-1) \leq 4\sqrt{2R\epsilon} + \epsilon$.

(iii) By (i), $(\boldsymbol{v}\cdot\boldsymbol{x})^p \in V_E$ for every $\boldsymbol{v} \in E$ and $p \geq 0$. Let U be a nonempty relatively open subset of E. Under the Fischer inner product $\langle \boldsymbol{x}^a, \boldsymbol{x}^b\rangle = a!\,\mathbf{1}\{a = b\}$ on homogeneous polynomials of degree p, $\langle q, (\boldsymbol{v}\cdot\boldsymbol{x})^p\rangle = p!\,q(\boldsymbol{v})$. A homogeneous q orthogonal to every $(\boldsymbol{v}\cdot\boldsymbol{x})^p$ with $\boldsymbol{v} \in U$ vanishes on U, hence on the open cone it generates, hence identically. Thus these powers span all homogeneous polynomials of degree p. □

DERIVATION FOR EXAMPLE 7.5. For the upper bound, we construct $w \in L^2(Q)$ with $g = c + \mathcal{A}^*(wQ)$ on B_r. For $\boldsymbol{v} \in E$ and integers $p \geq 2$, the univariate spanning formula gives

$$(\boldsymbol{v}\cdot\boldsymbol{x})^p = (-r)^p + p(-r)^{p-1}(\boldsymbol{v}\cdot\boldsymbol{x} + r) + \int_{-r}^{r}(\boldsymbol{v}\cdot\boldsymbol{x} - k)_+\,p(p-1)k^{p-2}\,\mathrm{d}k.$$

The linear case is represented using the support buffer:

$$\boldsymbol{v}\cdot\boldsymbol{x} = \frac{1}{R-r}\int_{-R}^{-r}\phi_{(\boldsymbol{v},k)}(\boldsymbol{x})\,\mathrm{d}k - \frac{R+r}{2}, \qquad \boldsymbol{x} \in B_r.$$

Consequently, for every $p \geq 1$ there are a constant c_p and $w_p \in L^\infty[-R, R]$, independent of $\boldsymbol{v}$, such that $(\boldsymbol{v}\cdot\boldsymbol{x})^p = c_p + \int\phi_{(\boldsymbol{v},k)}(\boldsymbol{x})w_p(k)\,\mathrm{d}k$. For a nonempty relatively open $U \subseteq E$, the map

$$\beta \longmapsto \int_U (\boldsymbol{v}\cdot\boldsymbol{x})^p\beta(\boldsymbol{v})\,\mathcal{H}^{d-1}(\mathrm{d}\boldsymbol{v})$$

from $L^2(U)$ to the homogeneous polynomials of degree p is surjective. Indeed, a polynomial Fischer-orthogonal to its range vanishes on U, hence identically, as in Proposition 7.3(iii). Every polynomial therefore equals $c + \int_{\mathcal{Z}_E} \phi_z w_E(z)\, Q_E(\mathrm{d}z)$ with $w_E \in L^2(Q_E)$. Set $w = w_E\, \mathrm{d}Q_E/\mathrm{d}Q$. Then

$$\|w\|^2_{L^2(Q)} \le q_*^{-1} \|w_E\|^2_{L^2(Q_E)}, \qquad \omega_g(\epsilon; \mathcal{P}(B_r)) \le \epsilon \|w\|_{L^2(Q)},$$

and Theorem 6.3 gives the upper rate. For the lower rate, choose μ, ν with $\int g\, \mathrm{d}(\mu - \nu) \neq 0$, and write $\eta = \mu - \nu$. The representation implies $\|C_\eta\|_{L^2(Q)} > 0$. The pair

$$\mu_t = \tfrac{1}{2}(\mu + \nu) + \tfrac{t}{2}\eta, \qquad \nu_t = \tfrac{1}{2}(\mu + \nu) - \tfrac{t}{2}\eta, \qquad 0 \le t \le 1,$$

gives $\omega_g(\epsilon; \mathcal{P}(B_r)) \ge c\epsilon$ for all sufficiently small ϵ. Equation (6.5) proves the lower rate, with the constant reduced to include the finitely many remaining sample sizes.

To estimate correlation, fix $v_* > 0$ and restrict attention to the nonempty class

$$\mathcal{C}(v_*) = \{\mu \in \mathcal{P}(B_r) : \mathrm{Var}_\mu(X_j) \ge v_*, \quad j = 1, 2\}.$$

Let $m_j = \mathbb{E}_\mu X_j$, $s_j = \mathbb{E}_\mu X_j^2$, and $q_{12} = \mathbb{E}_\mu X_1 X_2$. The polynomial estimators above, clipped respectively to $[-r, r]$, $[0, r^2]$, and $[-r^2, r^2]$, give estimates $\widehat{m}_j, \widehat{s}_j, \widehat{q}_{12}$ whose expected absolute errors are uniformly $O(n^{-1/2})$. Define

$$\widehat{v}_j = \max\left(\frac{v_*}{2}, \widehat{s}_j - \widehat{m}_j^2\right), \qquad \widehat{\rho} = \mathrm{clip}_{[-1,1]}\left(\frac{\widehat{q}_{12} - \widehat{m}_1\widehat{m}_2}{\sqrt{\widehat{v}_1\widehat{v}_2}}\right).$$

For $v_j = s_j - m_j^2 \ge v_*$,

$$|\widehat{v}_j - v_j| \le |\widehat{s}_j - s_j| + 2r\,|\widehat{m}_j - m_j|,$$

and

$$|\widehat{q}_{12} - \widehat{m}_1\widehat{m}_2 - (q_{12} - m_1 m_2)| \le |\widehat{q}_{12} - q_{12}| + r\sum_{j=1}^{2} |\widehat{m}_j - m_j|.$$

The map $(c, v_1, v_2) \mapsto c/(v_1 v_2)^{1/2}$ is Lipschitz on the bounded range in question with $v_1, v_2 \ge v_*/2$. Clipping cannot increase its error relative to a correlation in $[-1, 1]$. Thus

$$\sup_{\mu \in \mathcal{C}(v_*)} \mathbb{E}_\mu\, |\widehat{\rho} - \mathrm{Corr}_\mu(X_1, X_2)| \le C_{r, v_*, Q, \varsigma} n^{-\frac{1}{2}}.$$

□

PROOF OF COROLLARY 7.6. For $\mu, \nu \in \mathcal{K}_n(z, v)$, $C_\mu - C_\nu$ vanishes on z and is $2R$-Lipschitz in ρ by (2.3), hence $\sup_{\mathcal{Z}_E} |C_\mu - C_\nu| \le 2R h_E(z)$. For the random bound, (7.4) gives $Q_E(B_\rho(w, t)) \ge c t^d$ for $w \in \mathcal{Z}_E$ and $t \le 1$, and $\mathcal{Z}_E$ has a t-net of cardinality at most $C_d t^{-d}$. The proof of Lemma C.1 applies verbatim. □

PROOF OF COROLLARY 7.7. Let $g_0(\boldsymbol{x}) = (\boldsymbol{e} \cdot \boldsymbol{x})_+$, which is 1-Lipschitz, so every estimator $\widehat{\mu}$ satisfies $\left|\int g_0\, \mathrm{d}\widehat{\mu} - \int g_0\, \mathrm{d}\mu\right| \le W_1(\widehat{\mu}, \mu)$ and $\mathfrak{R}_n^Q(\alpha, L) \ge \mathfrak{R}_n^Q(g_0; \mathcal{C}_{\alpha,L})$. Corollary 7.10(i) gives $\mathfrak{R}_n^Q(g_0; \mathcal{C}_{\alpha,L}) \ge c(\log n)^{-(\alpha+3/2)}$ for large n,which proves the claimed divergence. The long-only assertion follows from Corollary 7.2. □

LEMMA F.1. *There is $C = C(d, r, \alpha)$ such that every $\eta \in H^\alpha(\mathbb{R}^d)$ supported in B_r satisfies*

$$\int_{\mathbb{R}} |\widehat{\eta}(t\boldsymbol{e})|^2\, (1 + t^2)^\alpha\, \mathrm{d}t \le C\, \|\eta\|^2_{H^\alpha}.$$

PROOF. Rotate so that $\boldsymbol{e} = \boldsymbol{e}_1$, write $\boldsymbol{x} = (y, \boldsymbol{x}')$, and let $\mathcal{F}_y\eta(t, \boldsymbol{x}')$ denote the Fourier transform of η in y alone. Since η is supported in B_r, $\widehat{\eta}(t\boldsymbol{e}) = \int_{|\boldsymbol{x}'|\le r} \mathcal{F}_y\eta(t, \boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'$, and Cauchy–Schwarz gives $|\widehat{\eta}(t\boldsymbol{e})|^2 \le C_{d,r}\int |\mathcal{F}_y\eta(t, \boldsymbol{x}')|^2\,\mathrm{d}\boldsymbol{x}'$. Plancherel's identity in $\boldsymbol{x}'$ and $(1+t^2)^\alpha \le (1 + t^2 + |\boldsymbol{\xi}'|^2)^\alpha$ give

$$\int_{\mathbb{R}} (1+t^2)^\alpha\, |\widehat{\eta}(t\boldsymbol{e})|^2\,\mathrm{d}t \le C_{d,r}\int_{\mathbb{R}^d} (1+|\boldsymbol{\xi}|^2)^\alpha\, |\widehat{\eta}(\boldsymbol{\xi})|^2\,\mathrm{d}\boldsymbol{\xi} \le C\,\|\eta\|_{H^\alpha}^2 .$$

□

LEMMA F.2 (Extrapolation of powers). *Let $U \subseteq S^{d-1}$ be closed with nonempty relative interior and let $\boldsymbol{e} \in S^{d-1}$. There is $\Lambda = \Lambda(U, \boldsymbol{e}) \ge 1$ such that for every integer $i \ge 1$ there is a finite signed measure m_i on U with $\|m_i\| \le \Lambda^i$ and*

$$(\boldsymbol{e}\cdot\boldsymbol{x})^i = \int_U (\boldsymbol{v}\cdot\boldsymbol{x})^i\, m_i(\mathrm{d}\boldsymbol{v}), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

PROOF. Choose $\boldsymbol{u}_0$ in the relative interior of U with $\boldsymbol{u}_0\cdot\boldsymbol{e} \ne 0$, which is possible since $\boldsymbol{e}^\perp \cap S^{d-1}$ has empty relative interior, and $\rho \in (0,1]$ such that the closed cap $U_0 = \{\boldsymbol{v} \in S^{d-1} : |\boldsymbol{v} - \boldsymbol{u}_0| \le \rho\}$ lies in U. Let H_i be the space of homogeneous polynomials of degree i on $\mathbb{R}^d$. The powers $(\boldsymbol{v}\cdot\boldsymbol{x})^i$, $\boldsymbol{v} \in U_0$, form a compact set spanning H_i, by the argument of Proposition 7.3(iii). By the bipolar theorem in H_i, the least total variation of a representing measure on U_0 equals $\sup|\lambda((\boldsymbol{e}\cdot\boldsymbol{x})^i)|$ over linear functionals λ on H_i with $|\lambda((\boldsymbol{v}\cdot\boldsymbol{x})^i)| \le 1$ for $\boldsymbol{v} \in U_0$, and the infimum is attained. Every such λ has the form $\lambda(P) = \langle q, P\rangle/i!$ for some $q \in \mathsf{H}_i$ under the Fischer inner product, and then $\lambda((\boldsymbol{v}\cdot\boldsymbol{x})^i) = q(\boldsymbol{v})$. It therefore suffices to show $|q(\boldsymbol{e})| \le \Lambda^i \sup_{U_0}|q|$ for $q \in \mathsf{H}_i$.

Let $\boldsymbol{p}_{\boldsymbol{e}} = \boldsymbol{e}/(\boldsymbol{e}\cdot\boldsymbol{u}_0)$ and $D = |\boldsymbol{p}_{\boldsymbol{e}} - \boldsymbol{u}_0|$. Since $\boldsymbol{p}_{\boldsymbol{e}}\cdot\boldsymbol{u}_0 = 1$, the vector $\boldsymbol{p}_{\boldsymbol{e}} - \boldsymbol{u}_0$ is orthogonal to $\boldsymbol{u}_0$. If $D = 0$, then $\boldsymbol{e} = \pm\boldsymbol{u}_0$ and $|q(\boldsymbol{e})| \le \sup_{U_0}|q|$. Otherwise let $P(s) = q(\boldsymbol{u}_0 + sD^{-1}(\boldsymbol{p}_{\boldsymbol{e}} - \boldsymbol{u}_0))$, a polynomial of degree at most i. For $|s| \le \rho$, the point $\boldsymbol{y} = \boldsymbol{u}_0 + sD^{-1}(\boldsymbol{p}_{\boldsymbol{e}} - \boldsymbol{u}_0)$ satisfies $|\boldsymbol{y}|^2 = 1 + s^2$ and $|\boldsymbol{y}/|\boldsymbol{y}| - \boldsymbol{u}_0|^2 = 2 - 2(1+s^2)^{-1/2} \le s^2$, hence $\boldsymbol{y}/|\boldsymbol{y}| \in U_0$ and, by homogeneity,

$$|P(s)| = |\boldsymbol{y}|^i\,|q(\boldsymbol{y}/|\boldsymbol{y}|)| \le 2^{i/2}\sup_{U_0}|q|\,, \qquad |s| \le \rho.$$

If $D \le \rho$, this bounds $|P(D)|$. If $D > \rho$, Chebyshev's extremal property [1, Chapter 2] gives $|P(D)| \le |T_i(D/\rho)|\sup_{|s|\le\rho}|P(s)|$, where T_i is the Chebyshev polynomial, and $|T_i(x)| \le (2x)^i$ for $x \ge 1$. Since $q(\boldsymbol{e}) = (\boldsymbol{e}\cdot\boldsymbol{u}_0)^i q(\boldsymbol{p}_{\boldsymbol{e}}) = (\boldsymbol{e}\cdot\boldsymbol{u}_0)^i P(D)$ and $|\boldsymbol{e}\cdot\boldsymbol{u}_0| \le 1$,

$$|q(\boldsymbol{e})| \le \Big(\sqrt{2}\max(1, 2D/\rho)\Big)^i \sup_{U_0}|q|\,,$$

which proves the lemma with $\Lambda = \sqrt{2}\max(1, 2D/\rho)$. □

LEMMA F.3. *Let η be a finite signed measure on B_R with $\eta(B_R) = 0$ and $\sup_{\mathcal{Z}_E}|C_\eta| \le \epsilon$. For $|k| \le R$ and every integer $j \ge 1$,*

$$\left|\int (\boldsymbol{e}\cdot\boldsymbol{x} - k)^j\,\eta(\mathrm{d}\boldsymbol{x})\right| \le \frac{3j}{R}\,\varrho^j\epsilon, \qquad \varrho = R(\Lambda+1),$$

where $\Lambda = \Lambda(E, \boldsymbol{e})$ is the constant of Lemma F.2.

PROOF. For $\boldsymbol{v} \in E$ and $i \ge 1$, Proposition 7.3(i) with $G(y) = y^i$ writes $(\boldsymbol{v}\cdot\boldsymbol{x})^i$ as a constant plus $\mathcal{A}^*m$ with $\|m\| \le 3iR^{i-1}$, hence $\left|\int (\boldsymbol{v}\cdot\boldsymbol{x})^i\,\mathrm{d}\eta\right| \le 3iR^{i-1}\epsilon$. Lemma F.2 with $U = E$ gives $\left|\int (\boldsymbol{e}\cdot\boldsymbol{x})^i\,\mathrm{d}\eta\right| \le 3i\Lambda^i R^{i-1}\epsilon$. Expanding $(\boldsymbol{e}\cdot\boldsymbol{x} - k)^j$ binomially, the constant term integrates to zero, and the remaining terms sum to at most $(3j/R)(|k| + \Lambda R)^j\epsilon$. □

LEMMA F.4 (Compactly supported profiles). *Let $d' \geq 1$, let $\zeta_0 \in C_c^\infty(\mathbb{R}^{d'})$ be non-negative with $\int \zeta_0 > 0$, and let $a_\perp > 0$. For $N \geq 1$ let $u_N = (N/2a_\perp)\mathbf{1}_{[-a_\perp/N, a_\perp/N]}$, $\kappa_N(\boldsymbol{x}) = \prod_{j=1}^{d'} u_N^{*N}(x_j)$, and $\zeta_N = \zeta_0 * \kappa_N$. Then $\zeta_N \geq 0$ is smooth, supported in $\operatorname{supp}\zeta_0 + [-a_\perp, a_\perp]^{d'}$, with $\int \zeta_N = \int \zeta_0$ and $\|\zeta_N\|_{H^{\alpha'}} \leq \|\zeta_0\|_{H^{\alpha'}}$ for every $\alpha' \geq 0$. For every $c_0 > 0$ and $J \geq 0$ there are $C, c > 0$, independent of N, with*

$$\left|\widehat{\zeta_N}(\boldsymbol{\xi})\right| \leq Ce^{-cN}(1+|\boldsymbol{\xi}|)^{-J}, \qquad |\boldsymbol{\xi}| \geq c_0 N.$$

PROOF. κ_N is a probability density supported in $[-a_\perp, a_\perp]^{d'}$, which gives the support, sign, mass, and Sobolev assertions by Young's inequality. Its transform is $\widehat{\kappa}_N(\boldsymbol{\xi}) = \prod_j \operatorname{sinc}(a_\perp \xi_j/N)^N$ with $\operatorname{sinc} v = \sin v/v$. If $|\boldsymbol{\xi}| \geq c_0 N$, some coordinate satisfies $|\xi_j| \geq c_0 N/\sqrt{d'}$, hence $|a_\perp \xi_j/N| \geq \delta := a_\perp c_0/\sqrt{d'}$ and $|\widehat{\kappa}_N(\boldsymbol{\xi})| \leq q_\delta^N$ with $q_\delta = \sup_{|v|\geq\delta} |\operatorname{sinc} v| < 1$. Since $\left|\widehat{\zeta_0}(\boldsymbol{\xi})\right| \leq C_J(1+|\boldsymbol{\xi}|)^{-J}$, the claim holds with $c = \log(1/q_\delta)$. □

LEMMA F.5. *Fix $r' > 0$. For each $N \geq 1$ there is $\mathsf{K}_N \in C_c^\infty(-r', r')$ with $\mathsf{K}_N(0) = 1$ such that, with constants depending only on r' and m,*

(a) $\|\mathsf{K}_N\|_{L^2}^2 \leq C/N$ *and* $\left\|\mathsf{K}_N^{(m)}\right\|_\infty \leq C_m N^m$;

(b) $\left\|\widehat{\mathsf{K}}_N\right\|_\infty \leq C/N$;

(c) $\left|\widehat{\mathsf{K}}_N(\tau)\right| \leq C_m N^m e^{-r'N/256}(1+|\tau|)^{-m}$ *for* $|\tau| \geq N/8$.

PROOF. Let $\chi_0 \in C_c^\infty(-1,1)$ with $0 \leq \chi_0 \leq 1$ and $\chi_0 = 1$ on $[-\frac{1}{2}, \frac{1}{2}]$. Set $\mathsf{v}_N^2 = N/(32r')$, $\mathcal{W}_N(y) = e^{-\mathsf{v}_N^2 y^2}\chi_0(y/r')$, $B_N(y) = \sin(Ny/16)/(Ny/16)$, and $\mathsf{K}_N = \mathcal{W}_N B_N$. Then $\mathsf{K}_N(0) = 1$, $|\mathsf{K}_N| \leq |B_N|$, and $\|B_N\|_{L^2}^2 = 16\pi/N$; the derivative bounds follow from Leibniz's rule. Next, $\widehat{\mathsf{K}}_N = (2\pi)^{-1}\widehat{\mathcal{W}}_N * \widehat{B}_N$ with $\widehat{B}_N = (16\pi/N)\mathbf{1}_{[-N/16, N/16]}$, and $\widehat{\mathcal{W}}_N = \mathsf{g}_N + \mathcal{R}_N$. Here $\mathsf{g}_N(\tau) = (\sqrt{\pi}/\mathsf{v}_N)e^{-\tau^2/(4\mathsf{v}_N^2)}$ has $\|\mathsf{g}_N\|_{L^1} = 2\pi$, and $\mathcal{R}_N$ is the transform of the smooth function $e^{-\mathsf{v}_N^2 y^2}(\chi_0(y/r') - 1)$, supported in $|y| \geq r'/2$. Its derivatives give $|\mathcal{R}_N(\tau)| \leq C_m \mathsf{v}_N^m e^{-\mathsf{v}_N^2 r'^2/8}(1+|\tau|)^{-m}$, and $\mathsf{v}_N^2 r'^2/8 = r'N/256$. Hence $\left\|\widehat{\mathcal{W}}_N\right\|_{L^1} \leq C$, which gives (b). For $|\tau| \geq N/8$ and $|\zeta| \leq N/16$, $|\tau - \zeta| \geq \max(N/16, |\tau|/2)$, hence $\mathsf{g}_N(\tau - \zeta) \leq C_m \mathsf{v}_N^m e^{-r'N/64}(1+|\tau|)^{-m}$; since $(2\pi)^{-1}\left\|\widehat{B}_N\right\|_{L^1} = 1$, this gives (c). □

PROOF OF THE QUOTED-DIRECTION BOUND IN THEOREM 7.8. If $\boldsymbol{e} \in E$, then $g_k = \phi_{(\boldsymbol{e},k)}$. If $-\boldsymbol{e} \in E$, put–call parity gives $g_k = \phi_{(-\boldsymbol{e},-k)} - \phi_{(-\boldsymbol{e},-R)} + R - k$. Lemma 6.2 gives $\varpi_{g_k}(\epsilon) \leq 2\epsilon$. □

PROOF OF THE UPPER BOUND IN (7.5). If $\pm\boldsymbol{e} \notin E$, compactness of E gives the missing-cone condition with $\kappa = 1 - \max_{\boldsymbol{u}\in E} |\boldsymbol{u}\cdot\boldsymbol{e}| > 0$. Let $f_1, f_2 \in \mathcal{F}_\alpha(L)$ with $\sup_{\mathcal{Z}_E} |C_{f_1} - C_{f_2}| \leq \epsilon$, and let $\eta = (f_1 - f_2)\,\mathrm{d}\boldsymbol{x}$, hence $\|\eta\|_{H^\alpha} \leq 2L$ and $\operatorname{supp}\eta \subseteq B_r$. Since η has zero mass, $(y-k)_+ = \frac{1}{2}(y-k) + \frac{1}{2}|y-k|$ and $|y| = \pi^{-1}\int_{\mathbb{R}}(1-\cos ty)t^{-2}\,\mathrm{d}t$ give

$$\int g_k\,\mathrm{d}\eta = \frac{M_1}{2} - \frac{1}{2\pi}\int_{\mathbb{R}} \frac{\mathsf{c}(t)}{t^2}\,\mathrm{d}t, \qquad M_1 = \int \boldsymbol{e}\cdot\boldsymbol{x}\,\eta(\mathrm{d}\boldsymbol{x}), \quad \mathsf{c}(t) = \int \cos(t(y-k))\,\eta_{\boldsymbol{e}}(\mathrm{d}y),$$

with Fubini's theorem justified by $|1-\cos ty|\,t^{-2} \leq \min(y^2, 2t^{-2})$. Lemma F.3 with $j = 1$ and $k = 0$ gives $|M_1| \leq 3(\Lambda+1)\epsilon$.

For $|t| \le N$, expand $\mathsf{c}(t) = \sum_{j\ge1}(-1)^j t^{2j}\mathsf{m}_{2j}/(2j)!$ with $\mathsf{m}_j = \int (\boldsymbol{e}\cdot\boldsymbol{x} - k)^j \,\mathrm{d}\eta$. Lemma F.3 gives

$$\frac{|\mathsf{c}(t)|}{t^2} \le \frac{3\varrho^2\epsilon}{R}\sum_{j\ge1}\frac{(\varrho\,|t|)^{2j-2}}{(2j-1)!} \le \frac{3\varrho^2\epsilon}{R}e^{\varrho|t|}, \qquad \int_{|t|\le N}\frac{|\mathsf{c}(t)|}{t^2}\,\mathrm{d}t \le \frac{6\varrho\epsilon}{R}e^{\varrho N}.$$

For $|t| > N$, $|\mathsf{c}(t)| \le |\widehat{\eta}(t\boldsymbol{e})|$, and Cauchy–Schwarz with Lemma F.1 gives

$$\int_{|t|>N}\frac{|\mathsf{c}(t)|}{t^2}\,\mathrm{d}t \le \left(\int|\widehat{\eta}(t\boldsymbol{e})|^2\,(1+t^2)^{\alpha}\,\mathrm{d}t\right)^{\frac12}\left(\int_{|t|>N} t^{-4-2\alpha}\,\mathrm{d}t\right)^{\frac12} \le CLN^{-\alpha-\frac32}.$$

Taking $N = \log(1/\epsilon)/(2\varrho)$ gives $\left|\int g_k\,\mathrm{d}\eta\right| \le C\left(\epsilon^{1/2} + L(\log(1/\epsilon))^{-(\alpha+3/2)}\right)$, which proves the upper bound in (7.5) for sufficiently small ϵ. □

PROOF OF THE LOWER BOUND IN (7.5). Write $\boldsymbol{x} = y\boldsymbol{e} + \boldsymbol{x}'$ with $\boldsymbol{x}' \perp \boldsymbol{e}$ and $p = \alpha + \frac32$. Choose $r' > 0$ and $r_\perp > 0$ with $r_\perp^2 + (|k| + r')^2 < r^2$. Let ψ_N be the profile of Lemma F.4 on $\boldsymbol{e}^\perp \cong \mathbb{R}^{d-1}$, built from a nonnegative $\psi_0 \in C_c^\infty$ and an $a_\perp$ for which every ψ_N is supported in $\{|\boldsymbol{x}'| < r_\perp\}$; write $M_\psi = \int\psi_N = \int\psi_0 > 0$. Let K_N be the window of Lemma F.5 with this r', and $\Phi_N(y) = \cos(Ny)\mathsf{K}_N(y)$.

Envelope. In Lemma F.5, $\left|B_N^{(j)}(y)\right| \le C_j N^j(1 + N\,|y|)^{-1}$ for $j \le 2$, and the Gaussian factor with its first two derivatives is bounded by $C_j N^{j/2}e^{-c_1 N y^2}$, $c_1 = 1/(64r')$. Leibniz's rule gives $|\Phi_N''(y)| \le CN^2(1+N^2y^2)^{-1/2}e^{-c_1Ny^2}$ on $\operatorname{supp}\mathsf{K}_N$. Let $\bar\chi \in C_c^\infty(-r', r')$ equal one on $\operatorname{supp}\mathsf{K}_N$ and

$$E_N(y) = C_0 N^{\frac12-\alpha}(1+N^2y^2)^{-\frac12}e^{-c_1Ny^2}\bar\chi(y),$$

with C_0 fixed so that $E_N \ge N^{-p}\,|\Phi_N''|$. Writing $E_N(y) = C_0N^{1/2-\alpha}h_N(Ny)$ with $h_N(t) = (1+t^2)^{-1/2}e^{-c_1t^2/N}\bar\chi(t/N)$, every derivative of h_N is bounded in $L^2(\mathbb{R})$ uniformly in N, and $\|h_N(N\cdot)\|_{H^\alpha}^2 \le N^{2\alpha-1}\,\|h_N\|_{H^\alpha}^2$ gives $\sup_N\|E_N\|_{H^\alpha} < \infty$, while $\int E_N \le CN^{-\alpha-1/2}\log(2+N)$.

Densities. Fix a smooth probability density f_* supported in B_r and set

$$\eta_N(\boldsymbol{x}) = bN^{-p}\psi_N(\boldsymbol{x}')\Phi_N''(y-k), \qquad m_N = bM_\psi\int E_N,$$

$$f_N^{\pm} = (1-m_N)f_* + b\,\psi_N(\boldsymbol{x}')E_N(y-k) \pm \eta_N.$$

For small b, $m_N \le \frac12$ for all N. Both functions are nonnegative since $E_N \ge N^{-p}\,|\Phi_N''|$ and $\psi_N \ge 0$; η_N has zero mass and zero first moments, by integration by parts in y, so both have mass one; their supports lie in B_r. Since $(1 + \left|\boldsymbol{\xi}'\right|^2 + \tau^2)^\alpha \le (1 + \left|\boldsymbol{\xi}'\right|^2)^\alpha(1+\tau^2)^\alpha$, products satisfy $\|a(\boldsymbol{x}')b(y)\|_{H^\alpha(\mathbb{R}^d)} \le C\,\|a\|_{H^\alpha(\mathbb{R}^{d-1})}\,\|b\|_{H^\alpha(\mathbb{R})}$. Lemma F.4 bounds $\|\psi_N\|_{H^\alpha}$; on $\{||\tau| - N| \le N/8\}$, Lemma F.5(a) gives $\int\tau^4(1+\tau^2)^\alpha\left|\widehat\Phi_N\right|^2 \le CN^{2\alpha+3}$, and Lemma F.5(c) makes the remaining contribution bounded, so $\|\Phi_N''\|_{H^\alpha} \le CN^p$. Hence $\left\|f_N^{\pm}\right\|_{H^\alpha} \le \|f_*\|_{H^\alpha} + Cb$, and for L large and b small, $f_N^{\pm} \in \mathcal{F}_\alpha(L)$ for all $N \ge 1$.

Price gap. Integrating by parts twice in y,

$$\int g_k(f_N^+ - f_N^-)\,\mathrm{d}\boldsymbol{x} = 2bN^{-p}M_\psi\Phi_N(0) = 2bM_\psi N^{-p}.$$

Quote discrepancy. As in (B.5), for $\boldsymbol{u} \in E$ write $\boldsymbol{u}'$ for the component of $\boldsymbol{u}$ orthogonal to $\boldsymbol{e}$; the missing-cone condition gives $|\boldsymbol{u}'| \ge \kappa' = (1-(1-\kappa)^2)^{1/2}$. Then

$$\widehat{\mathcal{A}\eta_N}(\boldsymbol{u},t) = bN^{-p}(\boldsymbol{u}\cdot\boldsymbol{e})^2\,\widehat\psi_N(t\boldsymbol{u}')\,\widehat\Phi_N(t\,\boldsymbol{u}\cdot\boldsymbol{e})\,e^{-\mathrm{i}kt\,\boldsymbol{u}\cdot\boldsymbol{e}},$$

where $\widehat{\Phi}_N(\tau) = \frac{1}{2}(\widehat{\mathsf{K}}_N(\tau - N) + \widehat{\mathsf{K}}_N(\tau + N))$. For $|t| \le N/2$, $|t\,\boldsymbol{u}\cdot\boldsymbol{e} \mp N| \ge N/2$, and Lemma F.5(c) bounds the integral over this region by $CN^{m+1}e^{-r'N/256}$. For $|t| > N/2$, $|t\boldsymbol{u}'| \ge \kappa' N/2$, and Lemma F.4 with $c_0 = \kappa'/2$ and $J = 2$ gives $\left|\widehat{\psi}_N(t\boldsymbol{u}')\right| \le Ce^{-cN}(1 + \kappa'\,|t|)^{-2}$, while $\left\|\widehat{\Phi}_N\right\|_\infty \le C/N$. Fourier inversion in t gives

$$\epsilon_N := \sup_{\mathcal{Z}_E} \left| C_{f_N^+} - C_{f_N^-} \right| \le Ce^{-cN}.$$

For small ϵ, choose the integer N with $Ce^{-cN} \le \epsilon < Ce^{-c(N-1)}$. Then $\epsilon_N \le \epsilon$ and $N \le C' \log(1/\epsilon)$, so monotonicity of ϖ and the price gap give $\varpi_{g_k}^{\alpha,L}(\epsilon) \ge 2bM_\psi N^{-p} \ge c(\log(1/\epsilon))^{-p}$.

The exponent $\alpha + \frac{3}{2}$ is the Sobolev scale of the longitudinal packet Φ_N''. For $\alpha < \frac{1}{2}$ its amplitude N^{2-p} exceeds any fixed baseline, and the ridge $b\,\psi_N E_N$, whose H^α norm stays bounded since the ridge has width N^{-1}, restores positivity. □

PROOF OF COROLLARY 7.10. (i) If $g_E = c'' + \mathcal{A}^* m_E$, then $\omega_g(\epsilon; \mathcal{C}_{\alpha,L}) \ge \varpi_g^{\alpha,L}(\epsilon) \ge |c|\ \varpi_{g_k}^{\alpha,L}(\epsilon) - \|m_E\|\,\epsilon$. The lower bound in (6.5) with $\epsilon = \varsigma/\sqrt{n}$ and the lower bound in (7.5) give the assertion once $\varsigma/\sqrt{n} \le \epsilon_0$.

(ii) By the upper bound in (6.5), it suffices to show $\omega_g(\epsilon; \mathcal{C}_{\alpha,L}) \le C(\log(1/\epsilon))^{-(\alpha+3/2)}$ for small ϵ. Let $G = C_\mu - C_\nu$ with $\mu, \nu \in \mathcal{C}_{\alpha,L}$ and $\|G\|_{L^2(Q)} \le \epsilon$, and let $A = \sup_{\mathcal{Z}_E} |G|$. By (2.3), G is $2R$-Lipschitz in ρ, so $|G| \ge A/2$ on the ρ-ball of radius $A/(4R)$ about a maximizer. For small ϵ this radius is at most one, and (7.4) gives $\epsilon^2 \ge \|G\|_{L^2(Q)}^2 \ge q_* c (A/4R)^d (A/2)^2$, hence $A \le C\epsilon^{2/(d+2)}$. Thus

$$\omega_g(\epsilon; \mathcal{C}_{\alpha,L}) \le \varpi_g^{\alpha,L}(C\epsilon^{2/(d+2)}) \le |c|\ \varpi_{g_k}^{\alpha,L}(C\epsilon^{2/(d+2)}) + C\,\|m_E\|\,\epsilon^{2/(d+2)},$$

and the upper bound in (7.5) gives the bound. At $\epsilon = \varsigma/\sqrt{n}$, $\log(1/\epsilon) \asymp \log n$. □

LEMMA F.6 (Comparison with exact quotes). *For every $\varsigma \ge 0$, $\mathfrak{R}_n^Q(g; \mathcal{C}, 0) \le \mathfrak{R}_n^Q(g; \mathcal{C}, \varsigma)$.*

PROOF. Adding $\varsigma\xi_i'$, with ξ_i' independent copies of ξ_1, to exact quotes produces data with the law of the experiment at level ς. Let $\widehat{p}$ be an estimator for level ς, truncated to the range of $\int g\,\mathrm{d}\mu$ over $\mathcal{C}$, which does not increase its risk. Then $\widetilde{p}(z, v) = \mathbb{E}_{\xi'}\widehat{p}(z, v + \varsigma\xi')$ is a measurable function of exact quotes, and Jensen's inequality for the absolute loss shows that its risk at every μ is at most that of $\widehat{p}$. □

LEMMA F.7 (Bayesian bound for a linear functional). *Let $T \in \mathbb{R}^{n\times M}$ be deterministic, let $\boldsymbol{\Theta}$ have the product density (C.9), and let $\boldsymbol{a} \in \mathbb{R}^M$. For every unit vector $\boldsymbol{v} \in \ker(T)$ and every measurable $\widehat{\psi} : \mathbb{R}^n \to \mathbb{R}$,*

$$\mathbb{E}\left[\left(\widehat{\psi}(T\boldsymbol{\Theta}) - \boldsymbol{a}\cdot\boldsymbol{\Theta}\right)^2\right] \ge \frac{(\boldsymbol{a}\cdot\boldsymbol{v})^2}{\pi^2}.$$

PROOF. It suffices to treat bounded $\widehat{\psi}$, since truncation at levels $\pm\sup|\boldsymbol{a}\cdot\boldsymbol{\theta}|$ over the cube does not increase the error. As in the proof of Lemma C.5, $\int \widehat{\psi}(T\boldsymbol{\theta})\,\partial_{\boldsymbol{v}} p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = 0$ since $\widehat{\psi}(T\boldsymbol{\theta})$ is constant along lines parallel to $\boldsymbol{v}$, while $\int (\boldsymbol{a}\cdot\boldsymbol{\theta})\,\partial_{\boldsymbol{v}} p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} = -\boldsymbol{a}\cdot\boldsymbol{v}$. With $\partial_{\boldsymbol{v}} p = p\,\boldsymbol{v}\cdot\boldsymbol{S}$, this gives $\mathbb{E}[(\widehat{\psi}(T\boldsymbol{\Theta}) - \boldsymbol{a}\cdot\boldsymbol{\Theta})\,\boldsymbol{v}\cdot\boldsymbol{S}(\boldsymbol{\Theta})] = \boldsymbol{a}\cdot\boldsymbol{v}$. Cauchy–Schwarz and $\mathbb{E}(\boldsymbol{v}\cdot\boldsymbol{S})^2 = \pi^2$, from (C.11), prove the claim. □

LEMMA F.8 (Perturbations with zero exchange price). *Let $0 < r_0 < r_1 < r$ and suppose $Q \le q^* Q_E$. Fix a nonnegative $\zeta \in C_c^\infty(\mathbb{R}^d)$ supported in the ball of radius $\frac{1}{4}$ about the origin, with $\int \zeta = 1$, and fix $\varepsilon_0 > 0$. There are C_3 and n_0, depending only on $d, R, r_0, r_1, \alpha, q^*, \zeta, \varepsilon_0$, E, and $\boldsymbol{w}$, such that for every $n \ge n_0$ the following holds. There are functions $b_1, \dots, b_K \in C_c^\infty(B_{r_1})$ with $K \le n^{C_3}$, depending on n but not on the design, such that*

(a) $\int b_j \,\mathrm{d}\boldsymbol{x} = 0$ *and* $\int g_k b_j \,\mathrm{d}\boldsymbol{x} = 0$ *for all j and all $k \in \mathbb{R}$;*

(b) $\left\|\sum_j \theta_j b_j\right\|_\infty + \left\|\sum_j \theta_j b_j\right\|_{H^\alpha} \le \varepsilon_0$ *for every $\boldsymbol{\theta} \in [-1,1]^K$;*

(c) *there is an event G_n, depending only on $Z_1, \dots, Z_n$, with $\Pr(G_n) \ge \frac{3}{4}$, on which every finite signed measure η supported in B_{r_0} with zero mass and zero first moments admits $\boldsymbol{c} \in \mathbb{R}^K$ with*

$$\sum_{j=1}^{K} c_j C_{b_j}(Z_i) = C_\eta(Z_i) \quad (1 \le i \le n), \qquad |\boldsymbol{c}|_2 \le n^{C_3} \max_{i \le n} |C_\eta(Z_i)| .$$

PROOF. Set $w_* = \min_{\boldsymbol{u} \in E} \boldsymbol{u} \cdot \boldsymbol{w} > 0$.

The family. Let $\Delta \in (0,1)$ and $\beta > 0$ be chosen below, let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_K$ enumerate the points of $\Delta\mathbb{Z}^d$ in $B_{r_1 - \Delta}$, so that $K \le C\Delta^{-d}$, and put $\zeta_j(\boldsymbol{x}) = \zeta((\boldsymbol{x} - \boldsymbol{x}_j)/\Delta)$ and $b_j = \beta\Delta\, \partial_{\boldsymbol{w}} \zeta_j$. The supports of the ζ_j are disjoint subsets of B_{r_1}. Integrating by parts, $\int b_j = 0$ and

$$\int g_k b_j \,\mathrm{d}\boldsymbol{x} = -\beta\Delta\,(\boldsymbol{w} \cdot \boldsymbol{e}) \int \mathbf{1}\{\boldsymbol{e} \cdot \boldsymbol{x} > k\} \zeta_j \,\mathrm{d}\boldsymbol{x} = 0,$$

which is (a). The functions b_j have sup norm at most $\beta \|\nabla\zeta\|_\infty$ and disjoint supports, and for every integer $m \ge 0$ their sum satisfies $\left\|\sum_j \theta_j b_j\right\|_{H^m}^2 \le \sum_j \|b_j\|_{H^m}^2 \le C_m \beta^2 \Delta^{-2m}$. Interpolation gives $\left\|\sum_j \theta_j b_j\right\|_{H^\alpha} \le C\beta\Delta^{-\alpha}$, and the choice $\beta = c\,\varepsilon_0 \Delta^\alpha$ with a small constant c proves (b). For $z = (\boldsymbol{u}, k) \in \mathcal{Z}_E$, integration by parts and $\partial_{\boldsymbol{w}} (\boldsymbol{u} \cdot \boldsymbol{x} - k)_+ = (\boldsymbol{u} \cdot \boldsymbol{w}) \mathbf{1}\{\boldsymbol{u} \cdot \boldsymbol{x} > k\}$ give

$$C_{b_j}(z) = -\beta\Delta\,(\boldsymbol{u} \cdot \boldsymbol{w}) \int \mathbf{1}_{H_z^+} \zeta_j \,\mathrm{d}\boldsymbol{x}, \qquad H_z^+ = \{\boldsymbol{x} : \boldsymbol{u} \cdot \boldsymbol{x} > k\}. \tag{F.2}$$

The event. Let $\gamma = R/(8q^* n)$ and $\rho_0 = (8 C_* q^* n^2)^{-1/d}$, where C_* is such that $Q_E(B_\rho(z,t)) \le C_* t^d$ for all $z \in \mathcal{Z}_E$ and $t > 0$. Let G_n be the event that no K_i satisfies $r_1 - \gamma \le |K_i| < r_1$ and that $\rho(Z_i, Z_{i'}) \ge \rho_0$ for all $i \ne i'$. Since the strike marginal of Q has density at most $q^*/(2R)$, the first condition fails with probability at most $n q^* \gamma / R = \frac{1}{8}$, and the second fails with probability at most $n^2 q^* C_* \rho_0^d = \frac{1}{8}$.

Reduction. Write $H_i = \{\boldsymbol{x} : U_i \cdot \boldsymbol{x} = K_i\}$ and $H_i^+ = \{\boldsymbol{x} : U_i \cdot \boldsymbol{x} > K_i\}$. On G_n, every index belongs to exactly one of $I = \{i : |K_i| < r_1 - \gamma\}$, $I_+ = \{i : K_i \le -r_1\}$, and $I_0 = \{i : K_i \ge r_1\}$. Let η be as in (c) and $t_i = C_\eta(Z_i)$. For $n \ge n_0$ we have $r_0 < r_1 - \gamma$. Hence H_i misses $\operatorname{supp}\eta$ for $i \in I_+ \cup I_0$, and the vanishing mass and first moments of η give $t_i = 0$. By (F.2), $C_{b_j}(Z_i) = 0$ for $i \in I_0$, and $C_{b_j}(Z_i) = -\beta\Delta(U_i \cdot \boldsymbol{w}) s_j$ with $s_j = \int \zeta_j = \Delta^d$ for $i \in I_+$. Hence it suffices to solve $M\boldsymbol{c} = (t_I, 0)$, where M has the rows $(C_{b_j}(Z_i))_j$ for $i \in I$ and the row $(s_j)_j$. If $\sigma = \min\{\left|M^{\mathsf{T}}(\boldsymbol{\lambda}, \lambda'')\right| : |\boldsymbol{\lambda}|^2 + \lambda''^2 = 1\}$ is positive, the minimum-norm solution satisfies $|\boldsymbol{c}| \le \sigma^{-1} |t_I| \le \sigma^{-1} \sqrt{n} \max_i |t_i|$. It remains to show $\sigma \ge n^{-C}$ on G_n.

Separated points. Fix $i \in I$. On G_n, $H_i \cap B_{r_1 - \gamma/2}$ is a $(d-1)$-dimensional disk of radius at least $(r_1\gamma/2)^{1/2}$, hence of area at least $A_n = c_d (r_1 \gamma)^{(d-1)/2}$ for $\gamma \le r_1/2$. Let $\varrho > 0$ and $i' \ne i$. Since $\boldsymbol{u} \cdot \boldsymbol{w} \ge w_*$ on E, $|U_i + U_{i'}| \ge 2w_*$, hence $\sin\angle(U_i, U_{i'}) = \frac{1}{2}|U_i - U_{i'}|\,|U_i + U_{i'}| \ge w_* |U_i - U_{i'}|$. If $|U_i - U_{i'}| \ge \rho_0/4$, the set of $\boldsymbol{x} \in H_i \cap B_R$ with $|U_{i'} \cdot \boldsymbol{x} - K_{i'}| < 2\varrho$ is a slab of width at most $16\varrho/(w_* \rho_0)$, of area at most $C_d R^{d-2} \varrho/(w_* \rho_0)$.

If $|U_i - U_{i'}| < \rho_0/4$, then $|K_i - K_{i'}| \geq 3R\rho_0/4$ on G_n, and every $\boldsymbol{x} \in H_i \cap B_R$ satisfies $|U_{i'} \cdot \boldsymbol{x} - K_{i'}| \geq |K_i - K_{i'}| - R\,|U_i - U_{i'}| \geq R\rho_0/2$, which exceeds 2ϱ when $\varrho \leq R\rho_0/4$. Choosing $\varrho = c\min(\gamma, w_*\rho_0 A_n/(nR^{d-2}))$ with a small constant c, the excluded area is at most $A_n/2$. There is therefore $\boldsymbol{p}_i \in H_i \cap B_{r_1-\gamma/2}$ with $\operatorname{dist}(\boldsymbol{p}_i, H_{i'}) \geq 2\varrho$ for all $i' \in I \setminus \{i\}$. For $i' \notin I$, the same holds since $H_{i'}$ misses B_{r_1}. Since the 2ϱ-neighborhoods of all $H_{i'}$ cover at most $Cn\varrho$ of the volume of $B_{r_1/2}$, there is also $\boldsymbol{p}_0 \in B_{r_1/2}$ at distance at least 2ϱ from every $H_{i'}$. The balls $B_i = B(\boldsymbol{p}_i, \varrho)$, $i \in I \cup \{0\}$, lie in $B_{r_1-\gamma/4}$, are pairwise disjoint, and B_i meets no hyperplane other than H_i.

Lower bound on σ. Let $\Delta = \varrho/(8\sqrt{d})$ and fix $(\boldsymbol{\lambda}, \lambda'')$ of unit norm. By (F.2), $M^{\mathsf{T}}(\boldsymbol{\lambda}, \lambda'') = (\langle F, \zeta_j\rangle)_j$ with

$$F = -\beta\Delta \sum_{i \in I} \lambda_i (U_i \cdot \boldsymbol{w}) \mathbf{1}_{H_i^+} + \lambda''.$$

On B_i, $i \in I$, the function F takes two values that differ by $\beta\Delta\,|\lambda_i|\,(U_i \cdot \boldsymbol{w}) \geq \beta\Delta w_*\,|\lambda_i|$, one on each side of H_i. On B_0 it is a constant v_0, and $|\lambda'' - v_0| \leq \beta\Delta\sqrt{n}\,|\boldsymbol{\lambda}|$. Each side of H_i in B_i contains at least $c_d(\varrho/\Delta)^d$ of the supports of the ζ_j, on each of which $\langle F, \zeta_j\rangle = \Delta^d F$, and B_0 contains as many. Disjointness of the balls therefore gives

$$\left|M^{\mathsf{T}}(\boldsymbol{\lambda}, \lambda'')\right|^2 \geq c\,\varrho^d\Delta^d \left[\frac{w_*^2(\beta\Delta)^2}{2}\,|\boldsymbol{\lambda}|^2 + v_0^2\right] \geq c\,\varrho^d\Delta^d \min\left\{\frac{w_*^2(\beta\Delta)^2}{4}, \frac{w_*^2}{8n}\right\},$$

where the last step uses $v_0^2 \geq \frac{1}{2}\lambda''^2 - (\beta\Delta)^2 n\,|\boldsymbol{\lambda}|^2$ with weight $w_*^2/(4n)$. Since γ, ρ_0, A_n, ϱ, Δ, and β depend only on n and are fixed powers of n up to constants, the family $b_1, \dots, b_K$ is fixed before the design, $\sigma \geq n^{-C}$, and $K \leq C\Delta^{-d} \leq n^C$. □

PROOF OF THEOREM 7.13. *Lower bound.* By Lemma F.6 it suffices to treat $\varsigma = 0$. Let $p = \alpha + \frac{3}{2}$ and use the objects of the proof of the lower bound in (7.5) with integer $N \geq 1$ to be chosen: the packet η_N, the ridge $b\,\psi_N E_N$, and m_N. Choose $r_0 < r_1 < r$ so that $\operatorname{supp}\eta_N \subseteq B_{r_0}$ for all N, and choose f_* positive on B_{r_1}, with minimum c_* there. Let $b_1, \dots, b_K$ and G_n be as in Lemma F.8 with $\varepsilon_0 = c_*/4$. For $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_K) \in [-1, 1]^{K+1}$ set

$$f_{\boldsymbol{\theta}} = (1 - m_N) f_* + b\,\psi_N(\boldsymbol{x}') E_N(y - k) + \theta_0\eta_N + \sum_{j=1}^{K} \theta_j b_j.$$

The first and last terms together are at least $c_*/4$ on B_{r_1} and nonnegative elsewhere, and the middle two terms are nonnegative since $E_N \geq N^{-p}\,|\Phi_N''|$. Hence $f_{\boldsymbol{\theta}}$ is a probability density supported in B_r, and its H^α norm is at most $\|f_*\|_{H^\alpha} + Cb + \varepsilon_0 \leq L$. By Lemma F.8(a) and the price computation in the proof of the lower bound in (7.5),

$$\int g_k f_{\boldsymbol{\theta}}\,\mathrm{d}\boldsymbol{x} = p_* + a_0\theta_0, \qquad a_0 = bM_\psi N^{-p},$$

for a constant p_* not depending on $\boldsymbol{\theta}$. The exact quotes are $C_{f_0}(Z_i) + (T\boldsymbol{\theta})_i$, where T has first column $t = (C_{\eta_N}(Z_i))_i$ and remaining columns $(C_{b_j}(Z_i))_i$.

Draw $\boldsymbol{\Theta}$ from the product prior (C.9), independently of the design and of any estimator randomization, and condition on both. On G_n, Lemma F.8(c) applied to η_N yields $\boldsymbol{c}$ with $\sum_j c_j C_{b_j}(Z_i) = t_i$ and $|\boldsymbol{c}| \leq n^{C_3} \sup_{\mathcal{Z}_E} |C_{\eta_N}| \leq Cn^{C_3} e^{-cN}$, by the bound on ϵ_N in the proof of the lower bound in (7.5). Choose $N = \lceil C_4 \log n \rceil$ with C_4 large enough that $|\boldsymbol{c}| \leq 1$. Then $\boldsymbol{v} = (1, -\boldsymbol{c})/\,|(1, -\boldsymbol{c})|$ is a unit vector in $\ker(T)$ with $(a_0, 0, \dots, 0) \cdot \boldsymbol{v} \geq a_0/\sqrt{2}$. Clip an arbitrary estimator $\widehat{p}$ to $[p_* - a_0, p_* + a_0]$, which does not increase its error. Lemma F.7 gives a conditional mean squared error at least $a_0^2/(2\pi^2)$. Since the clipped error is at most $2a_0$, the

conditional mean absolute error is at least $a_0/(4\pi^2)$ on G_n. Averaging over the design and bounding the prior average by the supremum,

$$\mathfrak{R}_n^{Q}\big(g_k;\mathcal{C}_{\alpha,L},0\big)\ge \Pr(G_n)\frac{a_0}{4\pi^2}\ge \frac{bM_\psi}{16\pi^2}\lceil C_4\log n\rceil^{-p}.$$

Upper bound. Let $h=n^{-1/(4d)}$, let $\{w_l\}$ be an h-net of $\mathcal{Z}_E$ in ρ of cardinality at most Ch^{-d}, and let $\mathcal{I}_l=\{i:\rho(Z_i,w_l)\le h\}$. Put $\widehat{C}_l$ equal to the average of Y_i over $\mathcal{I}_l$, or zero if $\mathcal{I}_l$ is empty. Let $\widehat{f}\in\mathcal{F}_\alpha(L)$ satisfy

$$\max_l\left|C_{\widehat{f}}(w_l)-\widehat{C}_l\right|\le \inf_{f'\in\mathcal{F}_\alpha(L)}\max_l\left|C_{f'}(w_l)-\widehat{C}_l\right|+\frac{1}{n},$$

chosen measurably by the countable-dense-set construction in the proof of Proposition 5.4, and estimate the price by $\int g_k\widehat{f}$. By (7.4) and $Q\ge q_*Q_E$, each ball $B_\rho(w_l,h)$ has Q-probability at least cq_*h^d. A Chernoff bound gives an event of probability at least $1-n^{-1}$ on which every $|\mathcal{I}_l|\ge cq_*nh^d$. On this event, the averages $\bar{\xi}_l$ of the errors over $\mathcal{I}_l$ have conditional variances at most $(cq_*nh^d)^{-1}$, hence $\mathbb{E}\max_l\varsigma\left|\bar{\xi}_l\right|\le\bar{\varsigma}(Ch^{-d}/(q_*nh^d))^{1/2}\le C\bar{\varsigma}n^{-1/4}$, and Markov's inequality gives $\max_l\varsigma\left|\bar{\xi}_l\right|\le n^{-1/8}$ outside an event of probability at most $Cn^{-1/8}$. Since C_f is R-Lipschitz in ρ, on the remaining event the true density f satisfies $\max_l\left|C_f(w_l)-\widehat{C}_l\right|\le\delta_n:=Rh+n^{-1/8}$. Hence $\max_l\left|C_{\widehat{f}}(w_l)-C_f(w_l)\right|\le 2\delta_n+1/n$, and the $2R$-Lipschitz property extends this to $\sup_{\mathcal{Z}_E}\left|C_{\widehat{f}}-C_f\right|\le\epsilon_n:=2\delta_n+1/n+2Rh\le Cn^{-1/(8d)}$. Theorem 7.8 bounds the price error by $\varpi_{g_k}^{\alpha,L}(\epsilon_n)\le C(\log(1/\epsilon_n))^{-p}\le C'(\log n)^{-p}$ for $n\ge n_0$. Off these events, the error is at most $2R$, which contributes at most $Cn^{-1/8}$. □

PROOF OF THEOREM 7.11. (a) Fix $\boldsymbol{v}\in D$. Let $N\ge1$ be an integer, $p=R/N$, and let μ and ν be supported on the segment $\{t\boldsymbol{v}:|t|\le R/2\}$ with densities $R^{-1}(1+\cos(2\pi t/p))$ and R^{-1} in $t\in[-R/2,R/2]$. The difference $\eta=\mu_{\boldsymbol{v}}-\nu_{\boldsymbol{v}}$ has zero mass and zero mean, since the interval contains whole periods and the cosine is even. Its call function $\Xi(l)=\int(t-l)_+\eta(\mathrm{d}t)$ satisfies $\Xi''=\eta$ and vanishes for $|l|\ge R/2$, hence $\Xi(l)=(p^2/(4\pi^2R))\big((-1)^N-\cos(2\pi l/p)\big)$ on $[-R/2,R/2]$ and $\sup|\Xi|=\epsilon_N:=p^2/(2\pi^2R)$. For $\boldsymbol{u}\in S^{d-1}$ and $\lambda=\boldsymbol{u}\cdot\boldsymbol{v}$, the call difference $C_\mu(\boldsymbol{u},k)-C_\nu(\boldsymbol{u},k)=\int(\lambda t-k)_+\eta(\mathrm{d}t)$ equals $\lambda\Xi(k/\lambda)$ if $\lambda>0$, equals $|\lambda|\,\Xi(-k/|\lambda|)$ if $\lambda<0$ by put–call parity and the vanishing mass and mean of η, and vanishes if $\lambda=0$. Hence $\sup_{\mathcal{Z}}|C_\mu-C_\nu|\le\epsilon_N$, while $W_1(\mu_{\boldsymbol{v}},\nu_{\boldsymbol{v}})=\int|\Xi'|=p/\pi^2=\pi^{-1}\sqrt{2R\epsilon_N}$. Monotonicity of ϖ_D between consecutive ϵ_N, whose square roots differ by the factor $N/(N+1)\ge\frac12$, gives the lower bound. The upper bound is Proposition 7.3(ii), which is uniform over 1-Lipschitz G and $\boldsymbol{v}\in E$.

(b) Use the perturbation (B.9) rotated to $\boldsymbol{e}$, as in the proof of Theorem 7.1, with integer $N\ge1$ and the bump ζ_N of Lemma F.4 with $d'=d$, with common support in B_r; the uniform Sobolev bounds on ζ_N keep the densities $f_N^{\pm}$ in $\mathcal{F}_\alpha(L)$. The transform $\widehat{\mathcal{A}H_N}(\boldsymbol{u},t)=\frac{b}{2}N^{-\alpha-2}(\boldsymbol{u}\cdot\boldsymbol{e})^2(\widehat{\zeta}_N(t\boldsymbol{u}-N\boldsymbol{e})+\widehat{\zeta}_N(t\boldsymbol{u}+N\boldsymbol{e}))$, the bound $|t\boldsymbol{u}\pm N\boldsymbol{e}|\ge c_\kappa(|t|+N)$ for $\boldsymbol{u}\in E$, and Lemma F.4 with $c_0=c_\kappa$ and $J=2$ give $\sup_{\mathcal{Z}_E}\left|C_{f_N^+}-C_{f_N^-}\right|\le Ce^{-cN}$. The test claim $-N^{-1}\cos(N\boldsymbol{e}\cdot\boldsymbol{x})$ lies in $\mathcal{G}_D$, and its integral against $f_N^+-f_N^-$ is $2bN^{-\alpha-1}\int\zeta_N\cos^2(N\boldsymbol{e}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=bN^{-\alpha-1}(\int\zeta_N+\mathrm{Re}\,\widehat{\zeta}_N(2N\boldsymbol{e}))\ge cN^{-\alpha-1}$ for large N. Choosing $N\asymp\log(1/\epsilon)$ gives the bound.

(c) For $\boldsymbol{v}\in D\cap E$ use (a). For $\boldsymbol{v}=\boldsymbol{e}$, let $f_1,f_2\in\mathcal{F}_\alpha(L)$ have discrepancy at most ϵ on $\mathcal{Z}_E$, let $\eta=(f_1-f_2)\,\mathrm{d}\boldsymbol{x}$, and let $\Psi(y)=\eta_{\boldsymbol{e}}((-\infty,y])$. Then $W_1(\mu_{\boldsymbol{e}},\nu_{\boldsymbol{e}})=\int_{-r}^{r}|\Psi|\le\sqrt{2r}\,\|\Psi\|_{L^2}$ and $\widehat{\Psi}(\tau)=\widehat{\eta}(\tau\boldsymbol{e})/(\mathrm{i}\tau)$. Lemma F.3 with $k=0$ bounds the moments $\mathsf{m}_j=\int(\boldsymbol{e}\cdot\boldsymbol{x})^j\,\mathrm{d}\eta$

by $(3j/R)\varrho^j\epsilon$, hence $|\widehat{\eta}(\tau \boldsymbol{e})| \leq \sum_{j\geq 1} |\tau|^j |\mathsf{m}_j| / j! \leq (3\varrho/R)\epsilon |\tau| e^{\varrho|\tau|}$. Hence the frequencies $|\tau| \leq N$ contribute at most $C\epsilon e^{\varrho N}$ to $\left\|\widehat{\Psi}\right\|_{L^2}$, and Lemma F.1 bounds the frequencies $|\tau| > N$ by $CLN^{-\alpha-1}$. Taking $N = \log(1/\epsilon)/(2\varrho)$ gives $W_1(\mu_{\boldsymbol{e}}, \nu_{\boldsymbol{e}}) \leq C(\epsilon^{1/2} + L(\log(1/\epsilon))^{-(\alpha+1)})$. With (b), this proves the two-sided bound when $\boldsymbol{e} \in D$. □

PROOF OF COROLLARY 7.12. Two laws in the same admissible set have discrepancy at most $2Rh_E$ on $\mathcal{Z}_E$, as in Corollary 7.6, which also bounds h_E. Apply Theorem 7.11. □

REFERENCES

[1] BORWEIN, P. and ERDÉLYI, T. (1995). *Polynomials and Polynomial Inequalities*. New York: Springer.

[2] BREEDEN, D. T. and LITZENBERGER, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business* **51** 621–651.

[3] CARR, P. and MADAN, D. (2001). Optimal positioning in derivative securities. *Quantitative Finance* **1** 19–37.

[4] CORRIGAN, A., WALLIN, J. and WANNER, T. (2009). A sampling inequality for fractional order Sobolev semi-norms using arbitrary order data. arXiv:0801.4097.

[5] CUESTA-ALBERTOS, J. A., FRAIMAN, R. and RANSFORD, T. (2007). A sharp form of the Cramér–Wold theorem. *Journal of Theoretical Probability* **20** 201–209.

[6] EKELAND, I. and TÉMAM, R. (1976). *Convex Analysis and Variational Problems*. Amsterdam: North-Holland.

[7] HÖRMANDER, L. (1983). *The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis*. *Grundlehren der mathematischen Wissenschaften* 256. Berlin: Springer.