Assumption 2 *(i) $\{\epsilon_i, \mathcal{F}_{i-1}\}_{i=1}^n$ with $\mathcal{F}_{i-1} = \sigma(X_i, \epsilon_{i-1}, X_{i-1}, \ldots)$ is a strictly stationary martingale difference sequence, (ii) $E[\epsilon_i^2|\mathcal{F}_{i-1}]$ is uniformly bounded for all $i \geq 1$, almost surely, (iii) $E[|\epsilon_i|^{2+\delta}] < \infty$ for some $\delta > 0$.*

Let $N(\mathcal{D}_n, \epsilon)$ denote the internal ϵ-covering number of $\mathcal{D}_n$ with respect to the Euclidean norm (i.e. the minimum number of points $x_1, \ldots, x_m \in \mathcal{D}_n$ such that the collection of ϵ-balls centered at each of $x_1, \ldots, x_m$ cover $\mathcal{D}_n$).

Assumption 3 *(i) $\mathcal{D}_n$ is compact, convex, has nonempty interior, and $\mathcal{D}_n \subseteq \mathcal{D}_{n+1}$ for all n, (ii) there exists $\nu_1, \nu_2 > 0$ such that $N(\mathcal{D}_n, \epsilon) \lesssim n^{\nu_1} \epsilon^{-\nu_2}$.*

Define $\zeta_{K,n} \equiv \sup_x \|b_w^K(x)\|$ and $\lambda_{K,n} \equiv \left[\lambda_{\min}(E[b_w^K(X_i) b_w^K(X_i)'])\right]^{-1/2}$.

Assumption 4 *(i) there exist $\omega_1, \omega_2 \geq 0$ s.t. $\sup_{x \in \mathcal{D}_n} \|\nabla b_w^K(x)\| \lesssim n^{\omega_1} K^{\omega_2}$, (ii) there exist $\varpi_1 \geq 0, \varpi_2 > 0$ s.t. $\zeta_{K,n} \lesssim n^{\varpi_1} K^{\varpi_2}$, (iii) $\lambda_{\min}(E[b_w^K(X_i) b_w^K(X_i)']) > 0$ for each K and n.*

Assumptions 1 and 2 trivially nest i.i.d. sequences, but also allow the regressors to exhibit quite general weak dependence. Note that Assumption 2(ii) reduces to $\sup_x E[\epsilon_i^2|X_i = x] < \infty$ in the i.i.d. case. Suitable choice of δ in Assumption 2(iii) for attainability of the optimal uniform rate will be explained subsequently. Strict stationarity of $\{\epsilon_i\}$ in Assumption 2 may be dropped provided the sequence $\{|\epsilon_i|^{2+\delta}\}$ is uniformly integrable. However, strict stationarity is used to present simple sufficient conditions for the asymptotic normality of functionals of $\widehat{h}$ in Section 3.

Assumption 3 is trivially satisfied when $\mathcal{X}$ is compact and $\mathcal{D}_n = \mathcal{X}$ for all n. More generally, when $\mathcal{X}$ is noncompact and $\mathcal{D}_n$ is an expanding sequence of compact subsets of $\mathcal{X}$ as described above, Assumption 3(ii) is satisfied provided each $\mathcal{D}_n$ is contained in an Euclidean ball of radius $r_n \lesssim n^\nu$ for some $\nu > 0$.[6]

Assumption 4 is a mild regularity condition on the sieve basis functions. When $\mathcal{X}$ is compact and rectangular this assumption is satisfied by all the widely used series (or linear sieve bases) with $\lambda_{K,n} \lesssim 1$, and $\zeta_{K,n} \lesssim \sqrt{K}$ for tensor-products of univariate polynomial spline, trigonometric polynomial or wavelet bases, and $\zeta_{K,n} \lesssim K$ for tensor-products of power series or orthogonal polynomial bases (see,

[6]By translational invariance we may assume that $\mathcal{D}_n$ is centered at the origin. Then $\mathcal{D}_n \subseteq \mathcal{R}_n = [-r_n, r_n]^d$. We can cover $\mathcal{R}_n$ with $(r_n/\epsilon)^d$ ℓ^∞-balls of radius ϵ, each of which is contained in an Euclidean ball of radius $\epsilon\sqrt{d}$. Therefore, $N(\mathcal{D}_n, \epsilon) \leq (\sqrt{d} r_n)^d \epsilon^{-d} \lesssim n^{\nu d} \epsilon^{-d}$.