Proofs

We first present several lemmas to deal with various terms concerning the bias and variance components in (8), (12), and (18). In the proofs, we use C to denote a generic positive constant that can take different values at different places.

Lemma 1. *Under assumptions (A1), (A2) and (A4),* $E[(B(T) - A_0 X)^{\otimes 2}]$ *has eigenvalues bounded away from zero and infinity.*

Proof of Lemma 1. By the smoothness assumption (A4), there exists $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_p) \in R^{K \times p}$ such that

$$|E[X_j | T = t] - \boldsymbol{\gamma}_j^{\mathrm{T}} \mathbf{B}(t)| \leq C K^{-d'}. \tag{27}$$

Since $E[X_j|T = t]$ is assumed to be smooth for $t \in [0, 1]$, it is also bounded on $[0, 1]$. Then by (27), $\boldsymbol{\gamma}_j^{\mathrm{T}} \int_0^1 \mathbf{B}(t)\mathbf{B}^{\mathrm{T}}(t) dt \boldsymbol{\gamma}_j$ is bounded by a constant. By well-known properties of splines, eigenvalues of $\int_0^1 \mathbf{B}(t)\mathbf{B}^{\mathrm{T}}(t)dt$ are bounded away from zero, and thus we can deduce that $\|\boldsymbol{\gamma}_j\|$ is bounded. In turn this implies that the operator norm of $\boldsymbol{\gamma}$ is bounded.

Then we show that the operator norm of

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\boldsymbol{\gamma}^{\mathrm{T}} & \mathbf{I} \end{pmatrix} \tag{28}$$

is bounded. This is easily shown by definition, since

$$\begin{aligned} & \left\| \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\boldsymbol{\gamma}^{\mathrm{T}} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \right\|^2 \\ = & \left\| \begin{pmatrix} \mathbf{u} \\ \mathbf{v} - \boldsymbol{\gamma}^{\mathrm{T}} \mathbf{u} \end{pmatrix} \right\|^2 \\ \leq & \|\mathbf{u}\|^2 + 2(\|\mathbf{v}\|^2 + \|\boldsymbol{\gamma}\|_{op}^2 \|\mathbf{u}\|^2) \leq C(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2), \end{aligned}$$

where $\|.\|_{op}$ for a matrix denotes its operator norm (and we shall use $\|.\|$ to denote its Frobenius norm). Note that the inverse of (28) is $\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \boldsymbol{\gamma}^{\mathrm{T}} & \mathbf{I} \end{pmatrix}$ which also has a bounded operator norm.

Let

$$\mathbf{A} = E\left[\begin{pmatrix} B(T)B^{\mathrm{T}}(T) & B(T)X^{\mathrm{T}} \\ XB^{\mathrm{T}}(T) & XX^{\mathrm{T}} \end{pmatrix}\right] = E\left[\begin{pmatrix} B(T) \\ X \end{pmatrix}\begin{pmatrix} B^{\mathrm{T}}(T) & X^{\mathrm{T}} \end{pmatrix}\right].$$

Premultiplying $\mathbf{A}$ by (28) and post-multiplying $\mathbf{A}$ by the transpose of (28), we get the matrix

$$E\left[\begin{pmatrix} B(T) \\ X - \boldsymbol{\gamma}^{\mathrm{T}}B(T) \end{pmatrix}\begin{pmatrix} B^{\mathrm{T}}(T) & X^{\mathrm{T}} - B^{\mathrm{T}}(T)\boldsymbol{\gamma} \end{pmatrix}\right].$$

The operator norm for the difference between the above and

$$E\left[\begin{pmatrix} B(T) \\ X - E[X|T] \end{pmatrix}\begin{pmatrix} B^{\mathrm{T}}(T) & X^{\mathrm{T}} - E[X^{\mathrm{T}}|T] \end{pmatrix}\right]$$

is (using operator norm is bounded by the maximum row sum of absolute values of entires) $O(K^{-d'}) = o(1)$. The displayed matrix above is block diagonal and the eigenvalues of both blocks are bounded and bounded away from zero by assumptions (A2) and (A3). This proves $\mathbf{A}$ has eigenvalues bounded away from zero and infinity.

Since $A_0 = E[B(T)X^{\mathrm{T}}](E[XX^{\mathrm{T}}])^{-1}$, $E[(B(T) - A_0X)^{\otimes 2}] = E[B(T)B^{\mathrm{T}}(T)] - E[B(T)X^{\mathrm{T}}](E[XX^{\mathrm{T}}])^{-1}E[XB^{\mathrm{T}}(T)]$ has eigenvalues bounded away from zero by Lemma 2. $\square$

Lemma 2. *If a symmetric matrix* $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{pmatrix}$ *has all eigenvalues inside the interval* $[c, C]$ *for some* $0 < c < C < \infty$*, then* $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}$ *also has all its eigenvalues inside the interval* $[c, C]$.

Proof of Lemma 2. Obviously eigenvalues of $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}$ are no larger than that of $\mathbf{C}$ and thus bounded by C. Next, we have the identity

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix}$$

Thus for any vector $\mathbf{b}$ with dimension same as that of $\mathbf{C}$, we have

$$\begin{aligned}
& \mathbf{b}^{\mathrm{T}}(\mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B})\mathbf{b} \\
= \;& (\mathbf{0}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}) \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & \mathbf{C}-\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix} \\
= \;& (\mathbf{0}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}) \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix} \\
= \;& (-\mathbf{b}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}, \mathbf{b}^{\mathrm{T}}) \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{pmatrix} \begin{pmatrix} -\mathbf{A}^{-1}\mathbf{B}\mathbf{b} \\ \mathbf{b} \end{pmatrix} \\
\geq \;& \|\mathbf{b}\|^2 c,
\end{aligned}$$

which completes the proof.

□

Lemma 3. *For any unit vector* $\mathbf{a} \in R^K$, $\|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X}\| = O_p(1)$*. If* $\mathbf{a}$ *has bounded support,* $\|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X}\| = O_p(K^{-1/2})$.

Proof of Lemma 3. The first statement is easily obtained by the bound $\|(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\|_{op} = O_p(1/n)$ (using for example Lemma 3 in Wang et al. (2011)) and $\|\mathbf{Z}^{\mathrm{T}}\mathbf{X}\| = O_p(n)$ (using that $E[B_k(T_i)X_{ij}] = E[B_k(T_i)E[X_{ij}|T_i]] \leq CE[B_k(T_i)] = O(1/\sqrt{K})$).

For the second statement, we assume $\mathbf{a} = (1, 0, \ldots, 0)^{\mathrm{T}}$ without loss of generality. We also condition on the event that eigenvalues of $(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}/n)^{-1}$ are bounded. By Lemma 6.3 of Zhou et al. (1998), the (j, j') entry of $(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}$ is bounded by $(C/n)\gamma^{|j-j'|}$ for some $\gamma < 1$. This means the j-th component of $\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}$ is bounded by $(C/n)\gamma^j$. Thus, denoting the j-th column of $\mathbf{X}$ as $\mathbf{X}_j$,

$$|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X}_j|$$

$$\leq \quad C\sum_{k=1}^{K}\gamma^k\sum_{i=1}^{n}B_k(T_i)|X_{ij}|/n. \tag{29}$$

Since $E[B_k(T)|X_j|] \leq CK^{-1/2}$, the above is $O_p(K^{-1/2})$ and the lemma is proved. $\square$

Lemma 4. *For any unit vector* $\mathbf{a} \in R^K$, $|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}| = O_p(K^{-d})$. *If* $\mathbf{a}$ *has bounded support,* $|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}| = O_p(K^{-d-1/2})$.

Proof of Lemma 4. Using that the components of $\mathbf{R}$ are bounded by CK^{-d}, the proof is basically the same as the proof of Lemma 3 and thus omitted. $\square$

Lemma 5. *Suppose now* $\boldsymbol{\theta}_0$ *is chosen as in* (19). *For any unit vector* $\mathbf{a} \in R^K$, $|\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}| = O_p(n^{-1/2}K^{-d+1/2})$. *If* $\mathbf{a}$ *is bounded support, it has order* $O_p(n^{-1/2}K^{-d})$.

Proof of Lemma 5. For general $\mathbf{a}$, this is easily proved by noting that $\sum_i B(T_i)r_i$ has mean zero and thus is $O_p(\sqrt{n}K^{-d})$ and thus $\|\mathbf{Z}^{\mathrm{T}}\mathbf{R}\| = O_p(\sqrt{n}K^{-d+1/2})$. For $\mathbf{a}$ with a bounded support, we can improve the rate as follows. Suppose without loss of generality $\mathbf{a} = (1, 0, \ldots, 0)^{\mathrm{T}}$. By Lemma 6.3 of Zhou et al. (1998), the (j, j') entry of $(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}$ is bounded by $(C/n)\gamma^{|j-j'|}$ for some $\gamma < 1$. Thus

$$\begin{aligned} & |\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}| \\ \leq \quad & (C/n)\sum_k \gamma^k|\sum_i B_k(T_i)r_i|. \end{aligned}$$

Let $b_k = |\sum_i B_k(T_i)r_i|$ and then $Eb_k^2 \leq CnK^{-2d}$ and by Cauchy-Schwarz inequality $E[b_kb_{k'}] \leq CnK^{-2d}$. Thus

$$\begin{aligned} & E\left[\left(\sum_k \gamma^k|\sum_i B_{k'}(T_i)r_i|\right)^2\right] \\ = \quad & \sum_{k,k'}\gamma^{k+k'}E[b_kb_{k'}] \\ \leq \quad & CnK^{-2d}, \end{aligned}$$

which implies $\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R} = O_p(n^{-1/2}K^{-d})$. $\square$

Lemma 6.

$$\|(\mathbf{X}^{\mathrm{T}} - \mathbf{X}^{\mathrm{T}}\mathbf{P}_{\mathbf{Z}})\mathbf{R}\| = o_p(\sqrt{n}).$$

Proof of Lemma 6.

$$\begin{aligned}
& \|(\mathbf{X}^{\mathrm{T}} - \mathbf{X}^{\mathrm{T}}\mathbf{P}_{\mathbf{Z}})\mathbf{R}\| \\
\leq\ & \|(\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}\mathbf{R}\| + \|(E[\mathbf{X}^{\mathrm{T}}|\mathbf{T}] - \mathbf{X}^{\mathrm{T}}\mathbf{P}_{\mathbf{Z}})\mathbf{R}\| \\
\leq\ & \|(\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}\mathbf{R}\| + \|(E[\mathbf{X}|\mathbf{T}] - \mathbf{X})^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}\| + \|(E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}(\mathbf{I} - \mathbf{P}_{\mathbf{Z}})\mathbf{R}\| \\
\leq\ & \|(\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}\mathbf{R}\| + \|(E[\mathbf{X}|\mathbf{T}] - \mathbf{X})^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{R}\| + \|(E[\mathbf{X}|\mathbf{T}] - \mathbf{Z}\boldsymbol{\gamma})^{\mathrm{T}}(\mathbf{I} - \mathbf{P}_{\mathbf{Z}})\mathbf{R}\|,
\end{aligned}$$

where $E[\mathbf{X}|\mathbf{T}] = (E[X|T = T_1], \ldots, E[X|T = T_n])^{\mathrm{T}}$, $\boldsymbol{\gamma}$ is such that $\|E[\mathbf{X}|\mathbf{T}] - \mathbf{Z}\boldsymbol{\gamma}\| = O_p(\sqrt{n}K^{-d'})$, and the last equality is due to that $(\mathbf{I} - \mathbf{P}_{\mathbf{Z}})\mathbf{Z} = 0$. Since $E[X|T]$ is the projection of X to the space of square-integrable functions of T, $X - E[X|T]$ is orthogonal to this space and thus $(\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}\mathbf{R}$ and $(E[\mathbf{X}|\mathbf{T}] - \mathbf{X})^{\mathrm{T}}\mathbf{Z}$ both have mean zero. Thus by direct calculation of their variances the first and the second terms in the displayed above are $O_p(\sqrt{n}K^{-d})$ and $O_p(\sqrt{n}K^{-d+1/2})$, respectively, using that $\|\mathbf{Z}^{\mathrm{T}}\mathbf{R}\| = O_p(nK^{-d})$. The third term above is bounded by $\|(E[\mathbf{X}|\mathbf{T}] - \mathbf{Z}\boldsymbol{\gamma})\| \cdot \|\mathbf{R}\| = O_p(nK^{-d-d'})$. This completes the proof since we assumed $\sqrt{n}K^{-d-d'} = o(1)$. $\square$

Lemma 7. $\|(\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\otimes 2}/n - E[(X - E[X|T])^{\otimes 2}]\| = o_p(1)$.

Proof of Lemma 7. Let $E[\mathbf{X}|\mathbf{T}] = (E[X|T = T_1], \ldots, E[X|T = T_n])^{\mathrm{T}}$. Law of large numbers implies

$$\left\|\frac{1}{n}(\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}(\mathbf{X} - E[\mathbf{X}|\mathbf{T}]) - E[(X - E[X|T])^{\otimes 2}]\right\| = o_p(1).$$

Since $(\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}(\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X}) = (\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}\mathbf{X}$, we have

$$\begin{aligned}
& (\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}(\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X}) - (\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}(\mathbf{X} - E[\mathbf{X}|\mathbf{T}]) \\
=\ & (\mathbf{X} - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}\mathbf{X} - (\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}\mathbf{X} + (\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}E[\mathbf{X}|\mathbf{T}] \\
=\ & (E[\mathbf{X}|\mathbf{T}] - \mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}\mathbf{X} + (\mathbf{X} - E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}E[\mathbf{X}|\mathbf{T}]
\end{aligned}$$

$$= \mathbf{X}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}(\mathbf{X}-E[\mathbf{X}|\mathbf{T}]) + \mathbf{X}^{\mathrm{T}}(\mathbf{I}-\mathbf{P}_{\mathbf{Z}})E[\mathbf{X}|\mathbf{T}] + (\mathbf{X}-E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}E[\mathbf{X}|\mathbf{T}]$$

$$= \mathbf{X}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}(\mathbf{X}-E[\mathbf{X}|\mathbf{T}]) + \mathbf{X}^{\mathrm{T}}(\mathbf{I}-\mathbf{P}_{\mathbf{Z}})(E[\mathbf{X}|\mathbf{T}]-\mathbf{Z}\boldsymbol{\gamma}) + (\mathbf{X}-E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}E[\mathbf{X}|\mathbf{T}],$$

where $\boldsymbol{\gamma}$ is such that $\|E[\mathbf{X}|\mathbf{T}]-\mathbf{Z}\boldsymbol{\gamma}\| = O_p(\sqrt{n}K^{-d'})$, and the last equality is due to that $(\mathbf{I}-\mathbf{P}_{\mathbf{Z}})\mathbf{Z}=0$. Using $\|E[\mathbf{X}|\mathbf{T}]-\mathbf{Z}\boldsymbol{\gamma}\| = O_p(\sqrt{n}K^{-d'})$ and that $\|\mathbf{I}-\mathbf{P}_{\mathbf{Z}}\|_{op} \le 1$, the second term above is $O_p(nK^{-d'})$. Since $\mathbf{Z}^{\mathrm{T}}(\mathbf{X}-E[\mathbf{X}|\mathbf{T}])$ and $(\mathbf{X}-E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}E[\mathbf{X}|\mathbf{T}]$ both have mean zero by the characterization of conditional expectation as a projection, the first term and the third term above are $O_p(\sqrt{nK})$ and $O_p(\sqrt{n})$, respectively. These bounds show that

$$\left\|\frac{1}{n}(\mathbf{X}-\mathbf{P}_{\mathbf{Z}}\mathbf{X})^{\mathrm{T}}(\mathbf{X}-\mathbf{P}_{\mathbf{Z}}\mathbf{X}) - \frac{1}{n}(\mathbf{X}-E[\mathbf{X}|\mathbf{T}])^{\mathrm{T}}(\mathbf{X}-E[\mathbf{X}|\mathbf{T}])\right\| = o_p(1),$$

which completes the proof. $\square$

Lemma 8. *$\|(\mathbf{Z}-\mathbf{P}_{\mathbf{X}}\mathbf{Z})^{\otimes 2}/n - E[(B(T)-A_0X)^{\otimes 2}]\| = o_p(1)$ if $K^2 logn/n \to 0$; $\|(\mathbf{Z}-\mathbf{P}_{\mathbf{X}}\mathbf{Z})^{\otimes 2}/n - E[(B(T)-A_0X)^{\otimes 2}]\| = o_p(K^{-1/2})$ if $K^3 logn/n \to 0$.*

Proof of Lemma 8. Let $\mathbf{X}A_0^{\mathrm{T}} = (A_0X_1,\ldots,A_0X_n)^{\mathrm{T}}$. We first show

$$\left\|\frac{1}{n}(\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}(\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}}) - E[(B(T)-A_0X)^{\otimes 2}]\right\| = o_p(1). \tag{30}$$

Since $|B_k(t)B_{k'}(t)| \le K$ and $E|B_k(T)B_{k'}(T)|^2] \le CK$ (because B_k is bounded by $\sqrt{K}$ and it has a support on an interval with length $O(1/K)$), by Bernstein's inequality,

$$P\left(\left|\frac{1}{n}\sum_i B_k(T_i)B_{k'}(T_i) - E[B_k(T)B_{k'}(T)]\right| > t\right) \le C\exp\{-C\frac{n^2t^2}{nK+nKt}\},$$

and thus

$$\|\mathbf{Z}^{\otimes 2}/n - E[B(T)^{\otimes 2}]\| = o_p(1),$$

if $K\log n/n \to 0$.

We have $\|X_{ij}B_k(T_i)\|_{\psi_1} \le C\|X_{ij}\|_{\psi_2}\|B_k(T_i)\|_{\psi_2} \le C\sqrt{K}$, where ψ_1 and ψ_2 are the sub-exponential norm and the sub-gaussian norm, respectively (see for example Definition 5.7 and 5.13 in Vershynin (2011)). Again by Bernstein's inequality,

$$P\left(\left|\frac{1}{n}\sum_i X_{ij}B_k(T_i) - E[X_{ij}B_k(T_i)]\right| > t\right) \le C\exp\{-C\frac{nt^2}{K+\sqrt{K}t}\},$$

and thus

$$\|\mathbf{Z}^{\mathrm{T}}\mathbf{X}A_0/n - E[B(T)XA_0^{\mathrm{T}}]\| = o_p(1),$$

if $K^2 \mathrm{log} n/n \to 0$.

Furthermore,

$$\|A_0(\mathbf{X}^{\mathrm{T}}\mathbf{X}/n - E[XX^{\mathrm{T}}])A_0^{\mathrm{T}}\| = o_p(1)$$

by law of large numbers and that $\|A_0\|_{op}$ is bounded. The latter can be shown as follows. Since $A_0 = E[B(T)X^{\mathrm{T}}](E[XX^{\mathrm{T}}])^{-1}$, we only need to show $\|E[B(T)X^{\mathrm{T}}]\|_{op}$ is bounded. In turn, this is seen by $|\mathbf{a}^{\mathrm{T}}E[B(T)X^{\mathrm{T}}]\mathbf{b}| \leq (\mathbf{a}^{\mathrm{T}}E[B(T)B(T)^{\mathrm{T}}]\mathbf{a})^{1/2}(\mathbf{b}^{\mathrm{T}}E[XX^{\mathrm{T}}]\mathbf{b})^{1/2}$. Thus we complete the proof of (30).

Since $(\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z}) = (\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}\mathbf{Z}$, we have

$$\begin{aligned}
& (\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z}) - (\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}}) \\
= \ & (\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}\mathbf{Z} - (\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\mathbf{Z} + (\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\mathbf{X}A_0^{\mathrm{T}} \\
= \ & (\mathbf{X}A_0^{\mathrm{T}} - \mathbf{P_X Z})^{\mathrm{T}}\mathbf{Z} + (\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\mathbf{X}A_0^{\mathrm{T}} \\
= \ & -\mathbf{Z}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}}) + (\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\mathbf{X}A_0^{\mathrm{T}}.
\end{aligned}$$

Since $\mathbf{X}^{\mathrm{T}}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})$ has mean zero, and $\|\mathbf{Z}^{\mathrm{T}}\mathbf{X}\| = O_p(n)$ as shown in the proof of Lemma 3, the two terms above are $O_p(\sqrt{nK})$. Thus

$$\left\| \frac{1}{n}(\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z}) - \frac{1}{n}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}}) \right\| = o_p(1),$$

which completes the proof. $\square$

Lemma 9.

$$\left(\mathbf{a}^{\mathrm{T}}\left(nE[(B(T) - A_0X)^{\otimes 2}]\right)^{-1}\mathbf{a}\right)^{-1/2}\mathbf{a}^{\mathrm{T}}\left((\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z})\right)^{-1}(\mathbf{Z}-\mathbf{P_X Z})^{\mathrm{T}}\boldsymbol{\epsilon} \xrightarrow{d} N(0,1).$$

Proof of Lemma 9. Obviously $\mathbf{a}^{\mathrm{T}}(nE[(B(T) - A_0X)^{\otimes 2}])^{-1}\mathbf{a} \asymp 1/n$ by Lemma 1. We have

$$\mathbf{a}^{\mathrm{T}}((\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z}))^{-1}(\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}\boldsymbol{\epsilon} - \mathbf{a}^{\mathrm{T}}((\mathbf{Z} - \mathbf{P_X Z})^{\mathrm{T}}(\mathbf{Z} - \mathbf{P_X Z}))^{-1}(\mathbf{Z} - \mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\epsilon}$$

$$
\begin{aligned}
&= \mathbf{a}^{\mathrm{T}}((\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z}))^{-1}(\mathbf{X}A_0^{\mathrm{T}}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}\boldsymbol{\epsilon} \\
&= \mathbf{a}^{\mathrm{T}}((\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z}))^{-1}(\mathbf{X}A_0^{\mathrm{T}}-\mathbf{Z})^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\epsilon} \\
&= O_p(\sqrt{K}/n) = o_p(n^{-1/2}),
\end{aligned}
$$

since its (conditional) second moment is

$$
\sigma^2\left\|\mathbf{a}^{\mathrm{T}}((\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z}))^{-1}(\mathbf{X}A_0^{\mathrm{T}}-\mathbf{Z})^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1/2}\right\|^2 = O_p(K/n^2),
$$

using Lemma 8 and that $\|(\mathbf{X}A_0^{\mathrm{T}}-\mathbf{Z})^{\mathrm{T}}\mathbf{X}\| = O_p(\sqrt{nK})$ (note $(\mathbf{X}A_0^{\mathrm{T}}-\mathbf{Z})^{\mathrm{T}}\mathbf{X}$ has mean zero). Furthermore,

$$
\begin{aligned}
& \mathbf{a}^{\mathrm{T}}((\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z}))^{-1}(\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\epsilon} - \mathbf{a}^{\mathrm{T}}(nE[(B(T)-A_0X)^{\otimes 2}])^{-1}(\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\epsilon} \\
&= \mathbf{a}^{\mathrm{T}}((\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z}))^{-1}\left(E[(B(T)-A_0X)^{\otimes 2}] - (\mathbf{Z}-\mathbf{P_X}\mathbf{Z})^{\mathrm{T}}(\mathbf{Z}-\mathbf{P_X}\mathbf{Z})/n\right) \\
&\quad (E[(B(T)-A_0X)^{\otimes 2}])^{-1}(\mathbf{Z}-\mathbf{X}A_0)^{\mathrm{T}}\boldsymbol{\epsilon} \\
&= o_p(n^{-1/2}),
\end{aligned}
$$

using again Lemma 8 and that $\|\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}}\| \le \|\mathbf{Z}\| + \|\mathbf{X}\|\|A_0\|_{op} = O_p(\sqrt{nK})$. The proof would be complete if we show

$$
(\sigma^2\mathbf{a}^{\mathrm{T}}(nE[(B(T)-A_0X)^{\otimes 2}])^{-1}\mathbf{a})^{-1/2}\mathbf{a}^{\mathrm{T}}(E[(B(T)-A_0X)^{\otimes 2}])^{-1}(\mathbf{Z}-\mathbf{X}A_0^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\epsilon} \to N(0,1).
$$

By Lemma 3.1 of Huang (2003), we only need to show that

$$
\frac{\max_i \alpha_i^2}{\sum_i \alpha_i^2} = o_p(1),
$$

where $\alpha_i = \mathbf{a}^{\mathrm{T}}(E[(B(T)-A_0X)^{\otimes 2}])^{-1}(B(T_i)-A_0X_i)$. It is easy to see that $\sum_i \alpha_i^2$ is of order n while $|\alpha_i| \le \|B(T_i)-A_0X_i\| = O_p(\sqrt{K})$ since $B_k(T_i)$ is bounded by $\sqrt{K}$ and only at most s basis functions B_k are nonzero at the point T_i (s is the order of the splines), thus the above displayed is verified and the proof of the Lemma is complete.
$\square$

Proof of Theorem 1. Most of the results are derived based on the expression (11). (20) is standard and well-known and thus omitted from the proof. The stochastic terms (12) are order $O_p(n^{-1/2})$ whether or not $\mathbf{a}$ has bounded support, by (17) and Lemmas 3 and 7. The first bias term in (18) is dealt with in Lemma 4 while the second bias term in (18) is dealt with in Lemmas 3, 6 and 7 which is $o_p(n^{-1/2})$ whether $\mathbf{a}$ has bounded support or not. These proved (21) and (22).

When $\boldsymbol{\theta}_0$ is chosen as in (9), the bias term is of order $o_p(n^{-1/2})$ by Lemma 5 which established (23).

Asymptotic normality (24) is obtained by Lemma 9 and (8) and noting that the second term in the bias (18) is $o_p(n^{-1/2})$. The equivalence between (24) and (25) when $\mathbf{a}$ has bounded support is due to that (17) is $o_p(1/n)$ by Lemmas 3 and 7. When $\boldsymbol{\theta}_0$ is chosen as in (9), the bias term can be removed from (24) and (25) since the bias term is of order $o_p(n^{-1/2})$ by Lemma 5. $\square$