4 Perturbed FIEM

We here consider the case where the explicit computation of $\bar{s}_i$ is not available and has to be replaced at each step by an approximation.

4.1 Description of the algorithm

sec:montecarlo

Data: $K_{\max} \in \mathbb{N}$, $\widehat{S}^0 \in \mathcal{S}$ and $\widetilde{\mathsf{S}}_{0,i} \in \mathcal{S}$ for any $i \in \{1, \dots, n\}$.
Result: The P-FIEM sequence: $\widehat{S}^k, k = 0, \dots, K_{\max}$
for $k = 0, \dots, K_{\max} - 1$ **do**
- Sample $I_{k+1} \sim \mathcal{U}(\{1, \dots, n\})$ independently from the past ;
- Compute $\widetilde{\mathsf{S}}_{k+1,I_{k+1}}$, an approximation of $\bar{s}_{I_{k+1}} \circ \mathsf{T}(\widehat{S}^k)$ and set $\widetilde{\mathsf{S}}_{k+1,i} = \widetilde{\mathsf{S}}_{k,i}$ for $i \neq I_{k+1}$;
- Sample $J_{k+1} \sim \mathcal{U}(\{1, \dots, n\})$ independently from the past ;
- Compute $\tilde{s}_{k+1}$ an approximation of $\bar{s}_{J_{k+1}} \circ \mathsf{T}(\widehat{S}^k)$;
- Set $\widehat{S}^{k+1} = \widehat{S}^k + \gamma_{k+1} \left(\tilde{s}_{k+1} - \widehat{S}^k + \frac{1}{n} \sum_{i=1}^n \widetilde{\mathsf{S}}_{k+1,i} - \widetilde{\mathsf{S}}_{k+1,J_{k+1}} \right)$.

Algorithm 7: Perturbed FIEM algorithm

Define the error when approximating expectations of the form $\bar{s}_i \circ \mathsf{T}(\widehat{S}^k)$: for $k \geq 0$,

$$\varepsilon^{(0)} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n \|\widetilde{\mathsf{S}}_{0,i} - \bar{s}_i \circ \mathsf{T}(\widehat{S}^0)\|^2,$$

$$\eta_{k+1}^{(1)} \stackrel{\text{def}}{=} \widetilde{\mathsf{S}}_{k+1,I_{k+1}} - \bar{s}_{I_{k+1}} \circ \mathsf{T}(\widehat{S}^k), \quad \eta_{k+1}^{(2)} \stackrel{\text{def}}{=} \tilde{s}_{k+1} - \bar{s}_{J_{k+1}} \circ \mathsf{T}(\widehat{S}^k).$$

Note that the case addressed in Section 2.4 corresponds to the results in this section, applied with $\eta_{k+1}^{(2)} = \eta_{k+1}^{(1)} = 0$, and $\varepsilon^{(0)} = 0$.

4.2 Case of stochastic approximations

sec:MC

When the approximations are random, introduce the filtrations $\mathcal{F}_0 \stackrel{\text{def}}{=} \sigma(\widehat{S}^0, \widetilde{\mathsf{S}}_{0,\cdot})$ and for $k \geq 0$,

$$\mathcal{F}_{k+1/4} \stackrel{\text{def}}{=} \mathcal{F}_k \vee \sigma\left(I_{k+1}\right), \qquad \mathcal{F}_{k+1/2} \stackrel{\text{def}}{=} \mathcal{F}_{k+1/4} \vee \sigma\left(\widetilde{\mathsf{S}}_{k+1,\cdot}\right)$$
$$\mathcal{F}_{k+3/4} \stackrel{\text{def}}{=} \mathcal{F}_{k+1/2} \vee \sigma(J_{k+1}), \qquad \mathcal{F}_{k+1} \stackrel{\text{def}}{=} \mathcal{F}_{k+3/4} \vee \sigma(\tilde{s}_{k+1});$$

Note also that, for all $k \geq 0$, $\eta_{k+1}^{(1)}$ is $\mathcal{F}_{k+1/2}$-measurable and $\eta_{k+1}^{(2)}$ is $\mathcal{F}_{k+1}$-measurable. The approximations will be said *unbiased* if, with probability one, for any $k \geq 0$

$$\mathbb{E}\left[\eta_{k+1}^{(1)} | \mathcal{F}_{k+1/4}\right] = 0, \quad \text{and} \quad \mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_{k+3/4}\right] = 0.$$

As an example of stochastic approximation, consider the Monte Carlo case: the expectation

$$\bar{s}_i \circ \mathsf{T}(s) = \int_{\mathsf{Z}} s_i(z)\ p_i(z; \mathsf{T}(s)) \mu(\mathrm{d}z)$$

can be approximated by a Monte Carlo sum. It holds

$$\bar{s}_i \circ \mathsf{T}(s) \approx \frac{1}{m} \sum_{j=1}^{m} s_i(Z_j^{\mathsf{T}(s),i})$$

where $\{Z_j^{\mathsf{T}(s),i}, j \geq 0\}$ are i.i.d. samples with distribution $p_i(\cdot, \mathsf{T}(s))\, \mathrm{d}\mu$; or, when exact sampling is intractable, the points are from a Markov chain designed to be ergodic with unique invariant distribution $p_i(\cdot, \mathsf{T}(s))\, \mathrm{d}\mu$.

The Monte Carlo approximation is unbiased, for example, when for any $k \geq 0$, conditionally to $\mathcal{F}_k$, the samples $\{Z_j^{\mathsf{T}(\widehat{S}^k),i}, j \geq 0\}$ are i.i.d. under the distribution $p_i(z; \mathsf{T}(\widehat{S}^k))\, \mathrm{d}\mu(z)$.

4.3 A general result on the error rate

The following theorem is available whatever the approximations $\widetilde{\mathsf{S}}$ and $\tilde{s}$: they can be deterministic or random, and if such, possibly based on a Monte Carlo approximation. Th results are derived under a control on the errors $\eta_{k+1}^{(1)}$ and $\eta_{k+1}^{(2)}$ as described by H5. Typically, the controls exhibited below are of interest when m_k and $\overline{m}_k$ increase with k.

hyp:approx:MC

H5. *There exist positive sequences $\{m_k, k \geq 0\}$ and $\{\overline{m}_k, k \geq 0\}$, positive numbers $M^{(1)}$ and $M^{(2)}$ and $M_\nu^{(2)} \geq 0$ such that for all $k \geq 0$, the approximations $\tilde{s}_{k+1}$ and $\widetilde{\mathsf{S}}_{k+1,I_{k+1}}$ satisfy*

$$\mathbb{E}[\|\eta_{k+1}^{(1)}\|^2] \leq \frac{M^{(1)}}{\overline{m}_{k+1}}, \quad \mathbb{E}[\|\mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_{k+3/4}\right]\|^2] \leq \frac{M_\nu^{(2)}}{m_{k+1}^2}, \quad \mathbb{E}[\|\eta_{k+1}^{(2)}\|^2] \leq \frac{M^{(2)}}{m_{k+1}}.$$

Note that $M_\nu^{(2)} = 0$ iff the approximation is unbiased.

theo:PFIEM:NonUnifStop

Theorem 16. *Assume H1item 1-item 2, H2, H3 and H4-item 1 to H4-item 4. Define $L^2 \overset{\mathrm{def}}{=} n^{-1} \sum_{i=1}^n L_i^2$.*

Let $K_{\max}$ be a positive integer. Let $\{\gamma_k,\ k \in \mathbb{N}\}$ be a sequence of positive step sizes and consider the FIEM sequence $\{\widehat{S}^k,\ k \in \mathbb{N}\}$ obtained with $\lambda_{k+1} = 1$ for any k. Assume that $\widehat{S}^k \in \mathcal{S}$ for any $k \leq K_{\max}$.

Let $\nu, \bar{\nu} \in \{0, 1\}$ with the convention $\nu = 0$ iff the approximations are unbiased, and $\bar{\nu} = 0$ iff for any $k \geq 0$, $\|\eta_{k+1}^{(1)}\| = \|\eta_{k+1}^{(2)}\| = \varepsilon^{(0)} = 0$.

For any positive numbers $\beta_1, \cdots, \beta_{K_{\max}-1}$ and $\beta_0 \in \left(0, v_{\min}/v_{\max}^2\right)$, it holds

$$\sum_{k=0}^{K_{\max}-1} \alpha_k \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] + \sum_{k=0}^{K_{\max}-1} \delta_k \mathbb{E}\left[\|\frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathsf{S}}_{k+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^k)\|^2\right]$$
$$\leq \mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]$$
$$+ \xi_0(K_{\max}, n)\ \mathbb{E}\left[\varepsilon^{(0)}\right] + \bar{\nu}\,\Xi_1(\eta_\cdot^{(1)}, K_{\max}, n) + \bar{\nu}\,\Xi_2(\eta_\cdot^{(2)}, K_{\max}, n); \tag{36}$$

eq:theo:conclusion:perturb

for any $k = 0, \ldots, K_{\max} - 1$,

$$\alpha_k \stackrel{\text{def}}{=} \gamma_{k+1}\left(v_{\min} - \nu v_{\max}^2 \beta_0 - (1+\nu)\frac{L_{\dot{V}}}{2}\gamma_{k+1}\left\{1 + (1+\bar{\nu})(1+\nu)L^2\Lambda_k\right\}\right)$$
$$\delta_k \stackrel{\text{def}}{=} (1+\nu)\frac{L_{\dot{V}}}{2}\gamma_{k+1}^2\left(1 + (1+\bar{\nu})(1+\nu)\frac{\Lambda_k}{(1+\beta_{k+1}^{-1})}\right)$$

with $\Lambda_{K_{\max}-1} = 0$ and for $k = 0, \ldots, K_{\max} - 2$,

$$\Lambda_k \stackrel{\text{def}}{=} \left(1 + \frac{1}{\beta_{k+1}}\right)\sum_{j=k+1}^{K_{\max}-1} \gamma_{j+1}^2 \prod_{\ell=k+2}^{j}\left(1 - \frac{1}{n} + \beta_\ell + (1+\bar{\nu})(1+\nu)L^2\gamma_\ell^2\right);$$

ξ_0, Ξ_1 and Ξ_2 are non negative real numbers; their explicit expressions can be found in Section 4.6, Eqs (54), (55) *and* (56). *By convention, $\prod_{\ell \in \emptyset} a_\ell = 1$.*

The sketch of the proof of this theorem is on the same lines as the proof of Theorem 4: the main part of the proof consists in the computation of an upper bound for the moment $\mathbb{E}[\|\widehat{S}^{k+1} - \widehat{S}^k\|^2]$. The proof is given in Section 4.6.

The LHS in (36) is the sum of two terms: in some sense, the first one is a distance to the set $\{s \in \mathcal{S} : h(s) \stackrel{\text{def}}{=} \bar{s} \circ \mathsf{T}(s) - s = 0\}$; and the second one is a measure of the approximation of the sum $\bar{s} \circ \mathsf{T}(\widehat{S}^k)$ by $n^{-1}\sum_{i=1}^n \widetilde{\mathsf{S}}_{k+1,i}$. Therefore this LHS can be seen as a convergence analysis of the algorithm as soon as $\alpha_k \geq 0$ and $\delta_k \geq 0$. In the next section, we propose a choice of the stepsize sequence $\{\gamma_k, k \geq 1\}$ and of the positive numbers $\beta_0, \ldots, \beta_{K_{\max}-1}$ implying that $\alpha_k \geq 0$, $A_{K_{\max}} \stackrel{\text{def}}{=} \sum_{k=0}^{K_{\max}-1} \alpha_k > 0$ and $\delta_k \geq (1+\nu)L_{\dot{V}}\gamma_{k+1}^2/2$. As a consequence, we obtain an upper bound for

$$\mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^K) - \widehat{S}^K\|^2\right] + \mathcal{G}_{K_{\max}},$$

where K is a $\{0, \ldots, K_{\max} - 1\}$-valued random variable, independent of $\mathcal{F}_{K_{\max}}$, and with distribution $\alpha_k/A_{K_{\max}}$; and

$$\mathcal{G}_{K_{\max}} \stackrel{\text{def}}{=} (1+\nu)\frac{L_{\dot{V}}}{2}\sum_{k=0}^{K_{\max}-1} \gamma_{k+1}^2 \mathbb{E}\left[\|\frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathsf{S}}_{k+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^k)\|^2\right]. \tag{37}$$

eq:Gronde

Finally, note that when $\alpha_k \geq 0$, we have by Lemma 11

$$\frac{1}{v_{\max}^2}\sum_{k=0}^{K_{\max}-1} \alpha_k \mathbb{E}\left[\|\dot{V}(\widehat{S}^k)\|^2\right] \leq \sum_{k=0}^{K_{\max}-1} \alpha_k \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right];$$

hence, Theorem 16 (and therefore, all the corollaries in Section 4.4) also provides an explicit control of the gradient $\dot{V}$ of the objective function along the path of the algorithm.

4.4 Error rates for specific stopping rules

sec:coro:perturbed

In Proposition 17, we propose a definition of the step sizes γ_k yielding to $A_{K_{\max}}$ positive and maximal among the considered family of weights α_k (see the proof, section 4.7): the step sizes have to be constant, and yield to the uniform weights $\alpha_k / A_{K_{\max}} = 1/K_{\max}$ for any k.

coro:pFIEM:optimal

Proposition 17 (following Theorem 16)**.** *Let $C \in (0,1)$ satisfying*

$$v_{\min} \leq (1+\nu)\sqrt{\frac{1+\nu}{1+\bar{\nu}}} \frac{L_{\dot{V}}}{L} \sqrt{C} \left(\frac{1}{n^{2/3}} + \frac{C}{(\sqrt{2-C}-1)^2} \right); \tag{38}$$

eq:def:Cstar

the optimal choice $C_\star$ being the unique C satisfying the equality. By choosing the constant stepsizes

$$\gamma_k \overset{\text{def}}{=} \frac{2 v_{min}}{(1+\nu)^2 C_{\text{GFM}} n^{2/3}}, \qquad C_{\text{GFM}} \overset{\text{def}}{=} 2 L_{\dot{V}} \left(\frac{1}{n^{2/3}} + \frac{C}{(\sqrt{2-C}-1)^2} \right),$$

we obtain

$$\begin{aligned}
&\frac{1}{K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\dot{V}(\widehat{S}^k)\|^2 \right] + v_{\max}^2 \mathcal{G}_{K_{\max}} \\
&\leq \frac{v_{\max}^2}{K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2 \right] + v_{\max}^2 \mathcal{G}_{K_{\max}} \\
&\leq (1+\nu)^3 C_{\text{GFM}} \frac{n^{2/3}}{K_{\max}} \frac{v_{\max}^2}{v_{\min}^2} \left(\mathbb{E}\left[V(\widehat{S}^0) \right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}}) \right] \right) && \text{(39a)} \\
&+ \frac{C_0}{n^{2/3}} \{1 \wedge \frac{n}{K_{\max}}\} \mathbb{E}\left[\varepsilon^{(0)} \right] && \text{(39b)} \\
&+ \frac{C_0}{n^{5/3}} \{1 \wedge \frac{n}{K_{\max}}\} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\eta_{k+1}^{(1)}\|^2 \right] && \text{(39c)} \\
&+ C_1 \frac{1}{K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_k \right] \|^2 \right] && \text{(39d)} \\
&+ \frac{C_0}{2(1+\nu)} \frac{1}{n^{2/3} K_{\max}} \sum_{k=0}^{K_{\max}-1} \left(\mathbb{E}\left[\|\eta_{k+1}^{(2)}\|^2 \right] + \mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_{k+3/4} \right] \|^2 \right] \right); && \text{(39e)}
\end{aligned}$$

eq:coro:pfiem:opt:a

eq:coro:pfiem:opt:b

eq:coro:pfiem:opt:c

eq:coro:pfiem:opt:d

eq:coro:pfiem:opt:e

$\mathcal{G}_{K_{\max}}$ is given by (37) *and the constants C_0 and C_1 are given by*

$$C_0 \stackrel{\text{def}}{=} (1+\nu)\sqrt{\frac{1+\nu}{1+\bar{\nu}}}\frac{L_{\dot{V}}}{L}\frac{2v_{\max}^2}{v_{\min}}\left\{1 \wedge \frac{1}{\sqrt{2-C}(\sqrt{2-C}-1)}\frac{1}{n^{1/3}}\right\},$$
$$C_1 \stackrel{\text{def}}{=} (1+\nu)\frac{v_{\max}^4}{v_{\min}^2} + \sqrt{\frac{1+\nu}{1+\bar{\nu}}}2\frac{L_{\dot{V}}}{L}\frac{v_{\max}^2}{v_{\min}}\frac{1}{\sqrt{2-C}(\sqrt{2-C}-1)^2} + \frac{C_0}{2(1+\nu)n^{2/3}}.$$

It is easily seen that the RHS of (38) is an increasing function of C on $(0,1)$, which tends to zero when $C \to 0$ and to infinity when $C \to 1$, thus showing that $C_\star$ is unique (see Lemma 15).

Gers: Pierre: rajouter ici un commentaire sur la constante C qui n'explose pas quand $n \to \infty$. De même pour C_1

We now derive the upper bounds when the approximations $\widetilde{\mathsf{S}}_k$ and $\tilde{s}_k$ satisfy H5 and, for $u \geq 0$ and $\varepsilon > 0$, we discuss how to choose $K_{\max}, \overline{m}_k$ and m_k as a function of u, ε so that the RHS in Proposition 17 is upper bounded by $O(\varepsilon n^{-u})$. Then we have $\bar{\nu} = 1$, and if the Monte Carlo approximation is unbiased $\nu = M_\nu^{(2)} = 0$. We will use the inequality

$$\mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)}|\mathcal{F}_k\right]\|^2\right] \leq \mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)}|\mathcal{F}_{k+3/4}\right]\|^2\right].$$

The term in (39a) says that $K_{\max} \propto n^{u+2/3}/\varepsilon$. With this choice of $K_{\max}$, the term in (39b) is $O(n^{-2/3} \wedge \{\varepsilon/n^{u+1/3}\})$. If we choose $\overline{m}_k = \overline{m}$, then the term in (39c) is $\overline{m}^{-1}\, O(n^{-2/3} \wedge \{n^{u-1}\epsilon^{-1}\})$; and it is upper bounded by $O(\varepsilon n^{-u})$ by choosing $\overline{m} \gtrsim \{\varepsilon^{-1}n^{u-2/3}\} \wedge \{n^{2u-1}\varepsilon^{-2}\}$. Finally, let us consider $m_k = m$: the first term (39e) exists for both biased and unbiased approximations. It is controlled by $O(n^{-2/3}m^{-1})$ and is upper bounded by $O(\varepsilon n^{-u})$ by choosing $m \gtrsim n^{u-2/3}\varepsilon^{-1}$. When $M_\nu^{(2)} \neq 0$, the two conditional expectations in (39d) and (39e) are a term which is $O(1/m^2)$ and this term can be bounded by $O(\varepsilon n^{-u})$ by setting $m \gtrsim n^{u/2}\varepsilon^{-1/2}$.

To summarize, the RHS (39a) to (39e) is upper bounded by $O(\varepsilon n^{-u})$ by choosing

Gers: c'est le max ou le min pour $\bar{m}$? dans les deux discussions

$$K_{\max} \gtrsim \frac{n^{u+2/3}}{\varepsilon}, \quad \overline{m} \gtrsim \frac{1}{n^{2/3-u}\varepsilon} \wedge \frac{1}{n^{1-2u}\varepsilon^2}, \quad m \gtrsim \frac{1}{n^{2/3-u}\varepsilon}$$

in the unbiased case and

$$K_{\max} \gtrsim \frac{n^{u+2/3}}{\varepsilon}, \quad \overline{m} \gtrsim \frac{1}{n^{2/3-u}\varepsilon} \wedge \frac{1}{n^{1-2u}\varepsilon^2}, \quad m \gtrsim \frac{1}{n^{2/3-u}\varepsilon} \vee \frac{n^{u/2}}{\sqrt{\varepsilon}}$$

in the biased one. Not surprisingly, the biased Monte Carlo case requires stronger conditions on the Monte Carlo batch size than the unbiased one.

In the following statement, we propose a different application of Theorem 16. In Proposition 17, we gave an upper bound on the quantity $\mathbb{E}\left[\|h(\widehat{S}^K)\|^2\right]$, where

K acts as a stopping rule for the algorithm, sampled uniformly in $\{0, \ldots, K_{\max}-1\}$. In Proposition 18, we address the case when the distribution of K is chosen among any probability distribution on $\{0, \ldots, K_{\max}-1\}$.

Gers: Proposition qui suit NON RELUE.

coro:pFIEM:given:sampling

Proposition 18 (following Theorem 16)**.** *Let $p_0, \ldots, p_{K_{\max}-1}$ be non negative real numbers such that $\sum_{k=0}^{K_{\max}-1} p_k = 1$. Let $C \in (0,1)$ satisfying*

$$v_{\min} \leq \frac{(1+\nu)}{1+2\nu} \sqrt{\frac{1+\nu}{1+\bar{\nu}}} \frac{L_{\dot{V}}}{L} \sqrt{C} \left(\frac{1}{n^{2/3}} + \frac{C}{(\sqrt{2-C}-1)^2} \right);$$

the optimal choice $C_\star$ being the unique C satisfying the equality. Define

$$F_\star(g) \stackrel{\text{def}}{=} \frac{\sqrt{C}}{Ln^{2/3}} g \left(\sqrt{\frac{1+\nu}{1+\bar{\nu}}} \frac{v_{\min}}{1+2\nu} - g \frac{1}{1+\bar{\nu}} \frac{L_{\dot{V}}}{2L} \sqrt{C} \left(\frac{1}{n^{2/3}} + \frac{C}{\left(\sqrt{2-C}-1\right)^2} \right) \right),$$

$$g_\star \stackrel{\text{def}}{=} \frac{\sqrt{(1+\bar{\nu})(1+\nu)}}{1+2\nu} \frac{v_{\min} L}{L_{\dot{V}} \sqrt{C}} \left(\frac{1}{n^{2/3}} + \frac{C}{\left(\sqrt{2-C}-1\right)^2} \right)^{-1}.$$

$F_\star$ is positive, continuous and increasing on $(0, g_\star)$, and by choosing the step sizes

$$\gamma_k \stackrel{\text{def}}{=} \frac{\sqrt{C}}{\sqrt{(1+\bar{\nu})(1+\nu)} Ln^{2/3}} \; F_\star^{-1} \left(\frac{p_k}{\max_\ell p_\ell} \; F_\star(g_\star) \right)$$

$$= \frac{\sqrt{C}}{\sqrt{(1+\bar{\nu})(1+\nu)} Ln^{2/3}} \; F_\star^{-1} \left(\frac{p_k}{\max_\ell p_\ell} \; \frac{v_{min}\sqrt{C}}{2Ln^{2/3}} \sqrt{\frac{1+\nu}{1+\bar{\nu}}} \frac{g_\star}{1+2\nu} \right)$$

we obtain

$$
\begin{aligned}
&\sum_{k=0}^{K_{\max}-1} p_k \mathbb{E}\left[\|\dot{V}(\widehat{S}^k)\|^2\right] + v_{\max}^2 \mathcal{G}_{K_{\max}} \\
&\leq v_{\max}^2 \sum_{k=0}^{K_{\max}-1} p_k \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] + v_{\max}^2 \mathcal{G}_{K_{\max}} \\
&\leq \frac{(1+2\nu)^2}{1+\nu} C_{\mathrm{GFM}} n^{2/3} \max_k p_k \; \frac{v_{\max}^2}{v_{\min}^2} \; \left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) \\
&+ \bar{C}_0 \frac{K_{\max} \max_k p_k}{n^{2/3}} \{1 \wedge \frac{n}{K_{\max}}\} \mathbb{E}\left[\varepsilon^{(0)}\right] \\
&+ \bar{C}_0 \frac{K_{\max} \max_k p_k}{n^{5/3}} \{1 \wedge \frac{n}{K_{\max}}\} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\eta_{k+1}^{(1)}\|^2\right] \\
&+ \bar{C}_1 \max_k p_k \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_k\right]\|^2\right] \\
&+ \frac{\bar{C}_0}{2(1+\nu)} \frac{\max_k p_k}{n^{2/3}} \sum_{k=0}^{K_{\max}-1} \left(\mathbb{E}\left[\|\eta_{k+1}^{(2)}\|^2\right] + \mathbb{E}\left[\|\mathbb{E}\left[\eta_{k+1}^{(2)} | \mathcal{F}_{k+3/4}\right]\|^2\right]\right);
\end{aligned}
$$

$\mathcal{G}_{K_{\max}}$ *is given by* (37)*,* C_{GFM} *is given in Proposition 17 and the constants* $\bar{C}_0$ *and* $\bar{C}_1$ *are given by*

$$
\begin{aligned}
\bar{C}_0 &\stackrel{\text{def}}{=} \frac{(1+2\nu)^2}{(1+\bar{\nu})(1+\nu)} \frac{L_{\dot{V}} C_{\mathrm{GFM}}}{L^2} \frac{v_{\max}^2}{v_{\min}^2} \left\{1 \wedge \frac{1}{\sqrt{2-C}(\sqrt{2-C}-1)} \frac{1}{n^{1/3}}\right\}, \\
\bar{C}_1 &\stackrel{\text{def}}{=} \frac{(1+2\nu)^2}{\sqrt{(1+\bar{\nu})}(1+\nu)^{3/2}} \frac{3C_{\mathrm{GFM}}}{4L} \frac{v_{\max}^4}{v_{\min}^3} \\
&\quad + \frac{(1+2\nu)^2}{(1+\bar{\nu})(1+\nu)^2} \frac{L_{\dot{V}} C_{\mathrm{GFM}}}{L^2} \frac{v_{\max}^2}{v_{\min}^2} \frac{1}{\sqrt{2-C}(\sqrt{2-C}-1)^2} + \frac{\bar{C}_0}{2(1+\nu) n^{2/3}}.
\end{aligned}
$$

Notice that in the case of a uniform stopping time, meaning $\max_k p_k = 1/K_{\max}$, we recover a similar upper bound to the one in Proposition 17 regarding the dependency in n and $K_{\max}$, the constant being slightly different. More precisely, assuming we took the same constant C for both corollaries, we have:

$$
\bar{C}_0 \geq \frac{(1+2\nu)^3}{(1+\nu)^4} C_0 \quad \text{and} \quad \bar{C}_1 \geq \frac{3}{2} \frac{(1+2\nu)^3}{(1+\nu)^4} C_1
$$