3 Error rates for a randomly stopped FIEM

sec:FIEM:complexity

3.1 Error rates: a general result

Given a maximal number of iterations $K_{\max}$, we address the choice of a random strategy to stop the algorithm at some (random) time $K \in \{1, \ldots, K_{\max}\}$. The quality criterion relies on an upper bound of the mean value of

$$\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2$$

which can be seen as an error when replacing a root of the function $s \mapsto h(s) = \bar{s} \circ \mathsf{T}(s) - s$ (see (8)) by a current estimate $\widehat{S}^k$. Under H4, this criterion is also related to the distance of the gradient of the Lyapunov function $V \stackrel{\text{def}}{=} F \circ \mathsf{T}$ (see Lemma 1) to zero:

$$\|\dot{V}(\widehat{S}^k)\|^2.$$

We derive explicit controls in Theorem 4 for a family of random stopping rules and then discuss the choice of this distribution.

H4. *1. The functions ϕ, ψ and R are continuously differentiable on Θ. T is continuously differentiable on $\mathcal{S}$.*

hyp:regV:C1 *2. For any $s \in \mathcal{S}$, $B(s) \stackrel{\text{def}}{=} (\dot{\phi \circ \mathsf{T}})(s)$ is a symmetric $q \times q$ matrix and there exist $0 < v_{min} \leq v_{max} < \infty$ such that for all $s \in \mathcal{S}$, the spectrum of $B(s)$ is in $[v_{min}, v_{max}]$.*

hyp:Tmap:smooth *3. For any $i \in \{1, \ldots, n\}$, $\bar{s}_i \circ \mathsf{T}$ is globally Lipschitz on $\mathcal{S}$ with constant L_i.*

hyp:regV:DerLip *4. The function $s \mapsto B^T(s)\,(\bar{s} \circ T(s) - s)$ is globally Lipschitz on $\mathcal{S}$ with constant $L_{\dot{V}}$.*

Under the additional assumptions that for any $s \in \mathcal{S}$, $\tau \mapsto L(s, \tau) \stackrel{\text{def}}{=} \overline{\psi}(\tau) - \langle s, \phi(\tau)\rangle + \mathsf{R}(\tau)$ is twice continously differentiable on Θ, $q \leq d$ and $\text{rank}(\dot{\mathsf{T}}(s)) = q$, then $B(s)$ is a symmetric matrix and its minimal eigenvalue is positive (see Lemma 8). Lemma 9 in Section 3.4.5 provides a sufficient condition for H4-item 3, condition which was essentially given in Karimi et al. (2019b).

Theorem 4 provides an explicit control for the sum of two terms: *(i)* some cumulated weighted errors $\mathbb{E}\left[\|h(\widehat{S}^k)\|^2\right]$ along a FIEM sequence $\{\widehat{S}^k,\ k \in \mathbb{N}\}$ of length $K_{\max}$; and *(ii)* some cumulated weighted errors when approximating $\bar{s} \circ \mathsf{T}(\widehat{S}^k)$ by $n^{-1} \sum_{i=1}^n \mathsf{S}_{k+1,i}$ along a FIEM sequence $\{\widehat{S}^k,\ k \in \mathbb{N}\}$ of length $K_{\max}$.

We then provide two applications of this result, both of them consisting in tuning some design parameters (the choice of the stepsize sequence $\{\gamma_k,\ k \in \mathbb{N}\}$ for example) in order to have non-negative weights with sum equal to one, in the cumulated errors $\mathbb{E}\left[\|h(\widehat{S}^k)\|^2\right]$. With these applications, Theorem 4 provides an explicit control on how far the algorithm is from the limiting set (caracterized by $\{s \in \mathcal{S}, h(s) = 0\}$) when stopped at a random time $K \in \{1, \ldots, K_{\max}\}$. The "distance" to this set is given by $\mathbb{E}\left[\|h(\widehat{S}^K)\|^2\right]$.

theo:FIEM:NonUnifStop

Theorem 4. *Assume H1item 1-item 2, H2item 1-item 1 and H3 and H4-item 1 to H4-item 4. Define* $L^2 \overset{\text{def}}{=} n^{-1} \sum_{i=1}^n L_i^2$.

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

For any positive numbers $\beta_1, \ldots, \beta_{K_{\max}-1}$*, we have*

$$\sum_{k=0}^{K_{\max}-1} \alpha_k \ \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] + \sum_{k=0}^{K_{\max}-1} \delta_k \mathbb{E}\left[\|\frac{1}{n}\sum_{i=1}^n \mathsf{S}_{k+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^k)\|^2\right] \\ \leq \mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right],$$

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

$$\alpha_k \overset{\text{def}}{=} \gamma_{k+1} v_{min} - \frac{L_{\dot{V}}}{2}\gamma_{k+1}^2 \left(1 + L^2 \Lambda_k\right), \quad \delta_k \overset{\text{def}}{=} \frac{L_{\dot{V}}}{2}\gamma_{k+1}^2 \left(1 + \frac{L^2 \Lambda_k}{(1 + \beta_{k+1}^{-1})}\right),$$

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

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

pageref:sketch

Proof. The detailed proof is provided in Section 3.4; we just give here a sketch of proof. Set

$$H_{k+1} \overset{\text{def}}{=} \bar{s}_{J_{k+1}} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k + \lambda_{k+1}\left\{\frac{1}{n}\sum_{i=1}^n \mathsf{S}_{k+1,i} - \mathsf{S}_{k+1,J_{k+1}}\right\},$$

so that $\widehat{S}^{k+1} = \widehat{S}^k + \gamma_{k+1} H_{k+1}$. In Lemma 11, it is proved that V is regular enough so that

$$V(\widehat{S}^{k+1}) - V(\widehat{S}^k) - \gamma_{k+1}\left\langle H_{k+1}, \dot{V}(\widehat{S}^k)\right\rangle \leq \gamma_{k+1}^2 \frac{L_{\dot{V}}}{2}\|H_{k+1}\|^2.$$

Taking the expectation, using a contraction inequality satisfied by V (see Lemma 11), and applying Lemma 3, yield

$$\mathbb{E}\left[V(\widehat{S}^{k+1})\right] - \mathbb{E}\left[V(\widehat{S}^k)\right] + \gamma_{k+1}\left(v_{\min} - \gamma_{k+1}\frac{L_{\dot{V}}}{2}\right)\mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] \\ \leq \gamma_{k+1}^2 \frac{L_{\dot{V}}}{2}\mathbb{E}\left[\|H_{k+1} - \mathbb{E}\left[H_{k+1}|\mathcal{F}_{k+1/2}\right]\|^2\right].$$

By summation from $k = 0$ to $k = K_{\max} - 1$, it holds

$$\sum_{k=0}^{K_{\max}-1} \gamma_{k+1}\left(v_{\min} - \gamma_{k+1}\frac{L_{\dot{V}}}{2}\right)\mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] \leq \mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^K_{\max})\right] \\ + \frac{L_{\dot{V}}}{2}\sum_{k=0}^{K_{\max}-1} \gamma_{k+1}^2 \mathbb{E}\left[\|H_{k+1} - \mathbb{E}\left[H_{k+1}|\mathcal{F}_{k+1/2}\right]\|^2\right]. \tag{12}$$

eq:condition:lambdak

Finally, in Lemma 12 and Proposition 13, we prove that the last term on the RHS is upper bounded by

$$\frac{L_{\dot V} L^2}{2} \sum_{k=0}^{K_{\max}-1} \gamma_{k+1}^2 \Lambda_k \ \Big\{ \mathbb{E}\left[\| \bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k \|^2 \right] - (1+\beta_{k+1}^{-1})^{-1} \ \mathbb{E}\left[\| \frac{1}{n} \sum_{i=1}^n \mathsf{S}_{k+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^k) \|^2 \right] \Big\} \ .$$

□

When $\alpha_k \geq 0$, we have by Lemma 11,

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

so that Theorem 4 also provides an upper bound for the gradient of the objective function along the FIEM path.

In Sections 3.2 and 3.3, we discuss how to choose some design parameters so that for any $k \in \{0, \ldots, K_{\max} - 1\}$, the coefficient α_k is non-negative and such that $A_{K_{\max}} \stackrel{\text{def}}{=} \sum_{k=0}^{K_{\max}-1} \alpha_k$ is positive and maximal. We then deduce from Theorem 4 that

$$\sum_{k=0}^{K_{\max}-1} \frac{\alpha_k}{A_{K_{\max}}} \mathbb{E}\left[\| \bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k \|^2 \right] \leq \frac{1}{A_{K_{\max}}} \Big\{ \mathbb{E}\left[V(\widehat{S}^0) \right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}}) \right] \Big\} , \tag{13}$$

eq:FIEM:error

which gives an upper bound on a weighted cumulated distance to $\{s \in \mathcal{S} : h(s) = \bar{s} \circ \mathsf{T}(s) - s = 0\}$ expressed by $\mathbb{E}\left[\| \bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k \|^2 \right]$. Note that the LHS is equal to

$$\mathbb{E}\left[\| \bar{s} \circ \mathsf{T}(\widehat{S}^K) - \widehat{S}^K \|^2 \right] \tag{14}$$

eq:FIEM:randomlecture

where $K \in \{0, \ldots, K_{\max} - 1\}$ is a r.v. with distribution $\{\alpha_k / A_{K_{\max}}, 0 \leq k \leq K_{\max} - 1\}$, and sampled independently of $\{\widehat{S}^k, 0 \leq k \leq K_{\max} - 1\}$. Therefore, the RHS in Theorem 4 (and therefore, in the following propositions), provides an explicit control for a random stopping rule of FIEM.

3.2 A uniform random stopping rule with constant step-sizes

sec:FIEM:errorrate:case1

In Proposition 5, we propose to choose constant step sizes γ_k, depending on n; this strategy yields to the uniform weights $\alpha_k / A_{K_{\max}} = 1/K_{\max}$ in (13) and to the uniform distribution for the stopping time K in (14).

Set

$$
\begin{aligned}
\mathsf{E}_0 &\stackrel{\text{def}}{=} \frac{1}{v_{\max}^2 K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\dot{V}(\widehat{S}^k)\|^2\right] = \frac{1}{v_{\max}^2}\mathbb{E}\left[\|\dot{V}(\widehat{S}^K)\|^2\right] \ , \\
\mathsf{E}_1 &\stackrel{\text{def}}{=} \frac{1}{K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^k) - \widehat{S}^k\|^2\right] = \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^K) - \widehat{S}^K\|^2\right] \ , \\
\mathsf{E}_2 &\stackrel{\text{def}}{=} \frac{1}{K_{\max}} \sum_{k=0}^{K_{\max}-1} \mathbb{E}\left[\|\frac{1}{n}\sum_{i=1}^{n} \mathsf{S}_{k+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^k)\|^2\right] \\
&= \mathbb{E}\left[\|\frac{1}{n}\sum_{i=1}^{n} \mathsf{S}_{K+1,i} - \bar{s} \circ \mathsf{T}(\widehat{S}^K)\|^2\right] \ ,
\end{aligned}
$$

where $K \sim \mathcal{U}(\{1, \ldots, n\})$ is independent of $\mathcal{F}_{K_{\max}}$.

coro:optimal:sampling

Proposition 5 (following of Theorem 4)**.** *Let* $\mu \in (0,1)$*. There exists* $C \in (0,1)$ *such that for any* $n \geq 2$ *and* $K_{\max} \geq 1$*, we have*

$$
\mathsf{E}_0 \leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L}\frac{\sqrt{C}\mathsf{E}_2}{v_{\min} n^{2/3}} \leq \frac{n^{2/3}}{K_{\max}} \frac{L}{\sqrt{C}(1-\mu)v_{\min}} \left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) \ ,
$$

where the FIEM sequence $\{\widehat{S}^k, k \in \mathbb{N}\}$ *is obtained with* $\gamma_\ell = \sqrt{C} n^{-2/3} L^{-1}$*. The constant* C *can be chosen such that*

$$
\sqrt{C}\left(\frac{1}{n^{2/3}} + \frac{1}{1-n^{-1/3}}\left(\frac{1}{n} + \frac{1}{1-C}\right)\right) \leq \frac{2L}{L_{\dot{V}}}\mu v_{\min} \ . \tag{15}
$$

eq:C:uniform

Upon noting that in (15), the LHS is an increasing function of C which is lower bounded by the increasing function $x \mapsto \sqrt{x}/(1-x)$, any constant C satisfying (15) is upper bounded by $C^+ \in (0,1)$ solving

$$
\sqrt{x} L_{\dot{V}} - 2L\mu v_{\min}(1-x) = 0 \ .
$$

There exist similar results in the literature, in the case $p_k = 1/K_{\max}$ for any k. In Karimi et al. (2019b), FIEM is run with a constant step size sequence equal to

$$
\gamma_{\mathrm{KM}} = \frac{v_{min}}{\max(6, 1+4v_{\min}) \ \max(L_{\dot{V}}, L_1, \ldots, L_n) \ n^{2/3}};
$$

and the upper bound is as in Proposition 5 where the constant

$$
C_{\mathrm{GFM}} \stackrel{\text{def}}{=} \frac{L}{\sqrt{C}(1-\mu)v_{\min}}
$$

is replaced with (see (Karimi et al., 2019b, Theorem 2))

$$
C_{\mathrm{KM}} \stackrel{\text{def}}{=} (\max(6, 1+4v_{\min}))^2 \ \max(L_{\dot{V}}, L_1, \ldots, L_n). \tag{16}
$$

eq:C:KM

Definition of C from an asymptotic point of view. Proposition 5 indicates how to fix the constant C which plays a role in the definition of the stepsize sequence $\{\gamma_k,\ k \in \mathbb{N}\}$ and in the control of the errors E_i. Based on an asymptotic point of view, another strategy which is only available for n large enough can be derived (see Section 3.4.2): choose

$$C_\star \stackrel{\text{def}}{=} \frac{1}{4}\left(\frac{v_{\min} L}{L_{\dot{V}}}\right)^{2/3} \; ; \tag{17}$$ eq:C:optimal:asymptotique

for any $\mu_\star \in (0,1)$, there exists $N_\star$ (depending upon $v_{\min}$, L, $L_{\dot{V}}$) such that for $n \geq N_\star$,

$$\mathsf{E}_0 \leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L}\frac{\sqrt{C}\mathsf{E}_2}{v_{\min} n^{2/3}} \leq \mu_\star \, \frac{8}{3}\frac{n^{2/3}}{K_{\max}}\frac{L^{2/3}L_{\dot{V}}^{1/3}}{v_{\min}^{4/3}}\left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) \, .$$

The definitions of C in Proposition 5 and of $C_\star$ in (17) are most often of no numerical interest since in many applications $v_{\min}$, L or $L_{\dot{V}}$ is unknown. Nevertheless, they attest that given a tolerance $\varepsilon > 0$, there exists a constant M depending upon $v_{\min}, L, L_{\dot{V}}$ such that

$$K_{\max} = M n^{2/3}\varepsilon^{-1} \Longrightarrow \mathsf{E}_0 \leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L}\frac{\sqrt{C}\mathsf{E}_2}{v_{\min} n^{2/3}} \leq \varepsilon\left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) \, .$$

In Proposition 6, we propose to choose constant step sizes γ_k depending both on n and $K_{\max}$; we obtain another control of the errors E_i computed along a FIEM path obtained with $\gamma_\ell \propto n^{-1/3}K_{\max}^{-1/3}$.

coro:optimal:sampling:Ketn **Proposition 6** (following of Theorem 4)**.** *Let $\mu \in (0,1)$. Choose $\lambda \in (0,1)$ and $C > 0$ such that*

$$\sqrt{C}\left(1 + C\left(1 + \frac{1}{1-\lambda}\right)\right) \leq \frac{2L}{L_{\dot{V}}}\mu v_{\min} \, .$$

Then for any $n, K_{\max} \geq 1$ such that $n^{1/3}K_{\max}^{-2/3} \leq \lambda/C$, we have

$$\mathsf{E}_0 \leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L}\frac{\sqrt{C}}{v_{\min}}\frac{\mathsf{E}_2}{(nK_{\max})^{1/3}} \leq \frac{n^{1/3}}{K_{\max}^{2/3}}\frac{L\left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right)}{\sqrt{C}(1-\mu)v_{\min}} \, ,$$

where the FIEM sequence $\{\widehat{S}^k, k \in \mathbb{N}\}$ is obtained with $\gamma_\ell = \sqrt{C}n^{-1/3}K_{\max}^{-1/3}L^{-1}$. There exists a constant $M \in (1, +\infty)$ depending upon $v_{\min}, L, L_{\dot{V}}, \mu$ such that for any $\varepsilon \in (0,1)$, we have

$$\frac{n^{1/3}}{K_{\max}^{2/3}}\frac{L}{\sqrt{C}(1-\mu)v_{\min}} \leq \varepsilon \, ,$$

by setting $K_{\max} = M\sqrt{n}\varepsilon^{-3/2}$.

Note that when $K_{\max} \propto \sqrt{n}$ then $\gamma_\ell \propto 1/\sqrt{n}$: this second upper bound is obtained with a slower step size than what was required in Proposition 5.

Definition of C from an asymptotic point of view. Proposition 6 indicates how to fix the constant C which plays a role in the definition of the stepsize sequence $\{\gamma_k,\ k \in \mathbb{N}\}$ and in the control of the errors E_i. Based on an asymptotic point of view, another strategy which is only available for n large enough can be derived (see Section 3.4.3): choose

$$C_\star \stackrel{\text{def}}{=} \left(\frac{v_{\min} L}{2 L_{\dot{V}}}\right)^{2/3} (1-\lambda_\star)^{2/3} , \tag{18}$$

eq:C:optimal:asymptotique:

where $\lambda_\star$ is the unique solution in $(0,1)$ of

$$\kappa \left(\frac{v_{\min} L}{2 L_{\dot{V}}}\right)^{2} (1-\lambda_\star)^{2} - \lambda_\star^3 = 0$$

for $\kappa > 0$. For any $\mu_\star \in (0,1)$ and $\kappa > 0$, there exists $N_\star$ (depending upon $v_{\min}$, L, $L_{\dot{V}}$) such that for $n \geq N_\star$ such that $n^{1/3} K_{\max}^{-2/3} \leq \kappa$,

$$\begin{aligned} \mathsf{E}_0 &\leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L} \frac{\sqrt{C}\mathsf{E}_2}{v_{\min}(n K_{\max})^{1/3}} \\ &\leq \frac{\mu_\star}{(1-\lambda_\star)^{1/3}} \ \frac{n^{1/3}}{K_{\max}^{-2/3}} \frac{4}{3} \left(\frac{2L^2 L_{\dot{V}}}{v_{\min}^4}\right)^{1/3} \left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) . \end{aligned}$$

The definitions of C in Proposition 6 and of $C_\star$ in (18) are most often of no numerical interest since in many applications $v_{\min}$, L or $L_{\dot{V}}$ is unknown. But here again, they attest that given a tolerance $\varepsilon > 0$, there exists a constant M depending upon $v_{\min}, L, L_{\dot{V}}$ such that

$$\begin{aligned} & K_{\max} = M n^{1/2} \varepsilon^{-3/2} \\ & \quad \Longrightarrow \mathsf{E}_0 \leq \mathsf{E}_1 + \frac{L_{\dot{V}}}{2L} \frac{\sqrt{C}\mathsf{E}_2}{v_{\min}(n K_{\max})^{1/3}} \leq \varepsilon \left(\mathbb{E}\left[V(\widehat{S}^0)\right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}})\right]\right) . \end{aligned}$$

As a corollary of Proposition 5 and Proposition 6, we have two upper bounds of the errors $\mathsf{E}_1, \mathsf{E}_2$: the first one is $O(n^{2/3} K_{\max}^{-1})$ and the second one is $O(n^{1/3} K_{\max}^{-2/3})$. Given a tolerance $\varepsilon > 0$, the first or second strategy will be chosen depending on how $n^{1/2}\varepsilon^{-3/2}$ and $n^{2/3}\varepsilon^{-1}$ compare i.e. how $\sqrt{\varepsilon}$ and $n^{-1/6}$ compare.

3.3 A non-uniform random stopping rule

sec:FIEM:errorrate:case2

For a $\{0, \ldots, K_{\max}-1\}$-valued random variable K, define

$$\begin{aligned} \mathsf{E}_3 &\stackrel{\text{def}}{=} \frac{1}{v_{\max}^2} \mathbb{E}\left[\|\dot{V}(\widehat{S}^K)\|^2\right] , \\ \mathsf{E}_4 &\stackrel{\text{def}}{=} \mathbb{E}\left[\|\bar{s} \circ \mathsf{T}(\widehat{S}^K) - \widehat{S}^K\|^2\right] . \end{aligned}$$

Given a distribution $p_0, \ldots, p_{K_{\max}-1}$ for the r.v. K, we show how to fix the step sizes $\gamma_1, \ldots, \gamma_{K_{\max}}$ in order to deduce from Theorem 4 a control of the errors E_3 and E_4. The proof of Proposition 7 is in Section 3.4.1.

For $C \in (0,1)$ and $n \geq 2$, define the function $F_{n,C}$

$$F_{n,C}(x) \mapsto \frac{1}{Ln^{2/3}} x \left(v_{\min} - x f_n(C) \right) \,,$$
$$f_n(C) \stackrel{\text{def}}{=} \frac{L_{\dot{V}}}{2L} \left(\frac{1}{n^{2/3}} + \frac{1}{1-n^{-1/3}} \left(\frac{1}{n} + \frac{1}{1-C} \right) \right) \,;$$

$F_{n,C}$ is positive, increasing and continuous on $(0, v_{\min}/(2f_n(C))]$.

coro:given:sampling

Proposition 7 (following of Theorem 4)**.** *Let K be a $\{0, \ldots, K_{\max}-1\}$-valued random variable with positive weights $p_0, \ldots, p_{K_{\max}-1}$. Let $C \in (0,1)$ solving*

$$2\sqrt{C} f_n(C) = v_{\min} \,. \tag{19}$$

eq:FIEM:NonUnifStep:C

For any $n \geq 2$ and $K_{\max} \geq 1$, we have

$$\mathsf{E}_3 \leq \mathsf{E}_4 \leq n^{2/3} \, \max_k p_k \, \frac{4Lf_n(C)}{v_{\min}^2} \left(\mathbb{E}\left[V(\widehat{S}^0) \right] - \mathbb{E}\left[V(\widehat{S}^{K_{\max}}) \right] \right) \,,$$

where the FIEM sequence $\{\widehat{S}^k, k \in \mathbb{N}\}$ is obtained with

$$\gamma_{k+1} = \frac{1}{n^{2/3}L} \; F_{n,C}^{-1} \left(\frac{p_k}{\max_\ell p_\ell} \frac{\sqrt{C} v_{\min}}{4L} \frac{1}{n^{2/3}} \right).$$

The constant C satisfying (19) is upper bounded by the unique point C^+ solving

$$v_{\min} L (1-x) - \sqrt{x} L_{\dot{V}} = 0 \,;$$

thus showing that $L_{\dot{V}}(1-C^+)^{-1}/(2L) \leq f_n(C) \leq \sup_n f_n(C^+) < \infty$

Note that since $\sum_k p_k = 1$, we have $\max_k p_k \geq 1/K_{\max}$ thus showing that among the dstributions, this term is minimal with the uniform distribution. In that case, the results in Proposition 7 can be compared to the results of Proposition 5: the control evolves as $n^{2/3}/K_{\max}$; the constant C solving the equality in (15) in the case $\mu = 1/2$ is the same as the constant C solving (19), and as a consequence, it is easily seen by using the equation (19), that

$$\frac{4Lf_n(C)}{v_{\min}^2} = \frac{2L}{\sqrt{C} v_{\min}} = \frac{L}{\sqrt{C}(1-\mu) v_{\min}} \,, \qquad \mu = 1/2.$$

Finally, when p_k is constant, the step sizes given by Proposition 7 are constant as in Proposition 5; and they are equal since $F_{n,C}^{-1}(v_{\min}^2 n^{-2/3}/(4Lf_n(C))) = \sqrt{C} = v_{\min}/(2f_n(C))$.

Analogue de la proposition 6 à faire.