A. On the branching factor

Proposition 2. $\kappa^u(\nu/2,\gamma) \leq \kappa^v(\nu,\gamma) \leq \kappa^u(2\nu,\gamma)$.

Proof. We prove, for any global optimum, for $h \geq 0$, $\mathcal{N}_h^v(\varepsilon) \leq \mathcal{N}_h^u(\varepsilon + \gamma^h/1-\gamma)$. Let a node $a \in A^h$ such that $v(a) \geq v^\star - \varepsilon$ then we have $u(a) \geq v(a) - \gamma^h/1-\gamma \geq v^\star - \varepsilon - \gamma^h/1-\gamma$. Similarly we have, for any global optimum, for $h \geq 0$, $\mathcal{N}_h^u(\varepsilon) \leq \mathcal{N}_h^v(\varepsilon + \frac{\gamma^h}{1-\gamma})$. Using the Definition 2 we obtain the claimed result. □

B. PlaTγPOOS is not using a budget larger than $n+1$

Notice, for any given depth $h \in [1:h_{\max}]$, PlaTγPOOS never uses more evaluations than $(p_{\max}+1)\frac{h_{\max}}{h}$ as

$$\begin{aligned}&\sum_{p=0}^{\lfloor \log_2(h_{\max}/\lceil h\gamma^{2h}\rceil)\rfloor} \left\lfloor \frac{h_{\max}}{h\lceil h2^p\gamma^{2h}\rceil}\right\rfloor \lceil h2^p\gamma^{2h}\rceil \\ &\leq (\lfloor \log_2(h_{\max}/\lceil h\gamma^{2h}\rceil)\rfloor + 1)\frac{h_{\max}}{h}\end{aligned}$$

Summing over the depths, PlaTγPOOS never uses more evaluations than the budget $n+1$ during its depth exploration as

$$\begin{aligned}&1 + (p_{\max}+1)\sum_{h=1}^{h_{\max}} \left\lfloor \frac{h_{\max}}{h} \right\rfloor \leq 1 + (p_{\max}+1)h_{\max}\sum_{h=1}^{h_{\max}} \frac{1}{h} \\ &= 1 + h_{\max}\overline{\log}(h_{\max})(p_{\max}+1) \leq 1 + h_{\max}(p_{\max}+1)^2 \\ &\leq n/2+1.\end{aligned}$$

We need to add the additional evaluation for the cross-validation at the end,

$$\sum_{p=0}^{p_{\max}}\sum_{t=0}^{h_{\max}} \frac{(t+1)\gamma^{2t}h_{\max}}{(1-\gamma^2)^2} \leq \sum_{p=0}^{p_{\max}} \left\lfloor \frac{n}{2(\log_2 n+1)^2}\right\rfloor \leq \frac{n}{2}.$$

Therefore, in total the budget is not more than $n/2+n/2+1 = n+1$. Again notice we use the budget of $n+1$ only for the notational convenience, we could also use $n/4$ for the evaluation in the end to fit under n (it's important that the amount of openings is *linear* in n).

C. Proofs of the lemmas

We first define and consider event ξ and prove it holds with high probability. The proof of the following lemmas can be found in Appendix C.

Lemma 2. *Let $\mathcal{C}$ be the set of sequence of actions evaluated by* PlaTγPOOS during one of its runs. $\mathcal{C}$ is a random quantity. Let ξ be the event under which all average estimates for the reward of the state-action pairs receiving at least one evaluation from PlaTγPOOS are within their classical confidence interval, then $P(\xi) \geq 1-\delta$, where

$$\begin{aligned}\xi \triangleq \Bigg\{&\forall a \in \mathcal{C}, \forall h \in [2:h(a)],\, p \in [0:p_{\max}]: \\ &\text{if } T_{a_{[h]}} \geq \left\lceil (h-1)2^p\gamma^{2(h-1)}\right\rceil, \\ &\text{then } |\widehat{u}(a) - u(a)| \leq b\sqrt{\frac{p_{\max}\log(4n/\delta)}{2^{p+1}}}\Bigg\}.\end{aligned}$$

Proof. The idea of the proof of this lemma follows the similar line as the proof of the statement given for `StoSOO` (Valko et al., 2013). The crucial point is that while we have potentially exponentially many combinations of cells that can be evaluated, given any particular execution we need to consider only a polynomial number of estimators, m, for which we can use a Azuma-Hoeffding concentration inequality.

We denote $\forall h \in [0 : h_{\max}]$, $p \in [0 : p_{\max}] : a^{i,h,p} \in \mathcal{C}$ the i-th evaluated node of depth h such that $\forall t \in [2 : h], T_{a^{i,h,p}_{[t]}} \geq \lceil (t-1)2^p\gamma^{2(t-1)} \rceil$. Note that in PlaT$\gamma$POOS we have $T_{a^{i,h,p}_{[1]}} = h_{\max}$.

Though $a^{i,h,p}$ is random, we study the quantity $\left|\widehat{u}(a^{i,h,p}) - u(a^{i,h,p})\right|$. We recall that

$$\widehat{u}(a^{i,h,p}) - u(a^{i,h,p}) = \sum_{t=0}^{h-1} \gamma^t (\widehat{r}_t(a^{i,h,p}) - r_t(a^{i,h,p})) \tag{2}$$

$$= \sum_{t=0}^{h-1} \gamma^t \sum_{s=0}^{T_{a^{i,h,p}_{[t+1]}}} \frac{\widehat{r}^{i,h,p}_{t,s} - r_t(a^{i,h,p})}{T_{a^{i,h,p}_{[t+1]}}} \tag{3}$$

This quantity is composed of the elements $\widehat{r}^{i,h,p}_{t,s} - r_t(a^{i,h,p})$ that form a martingale.

Therefore using a Azuma-Hoeffding concentration inequality with a union bound already on the values of T we have

$$\mathbb{P}\left(\widehat{u}(a^{i,h,p}) - u(a^{i,h,p}) \leq b\sqrt{\sum_{t=0}^{h-1} \frac{\gamma^{2t}\log(p_{\max}/\delta)}{2T_{a^{i,h,p}_{[t+1]}}}}\right) \geq 1 - \delta/p_{\max}$$

Moreover we have for all $h \geq t > 1$,

$$\frac{\gamma^{2t}}{T_{a^{i,h,p}_{[t+1]}}} \leq \frac{\gamma^{2t}}{\lceil t2^p\gamma^{2t}\rceil} \leq \frac{\gamma^{2t}}{t2^p\gamma^{2t}} = \frac{1}{t2^p} \tag{4}$$

For $t = 0$, $\frac{\gamma^{2t}}{T_{a^{i,h,p}_{[t+1]}}} = \frac{1}{h_{\max}} \leq= \frac{1}{2^p}$ for all $p \leq p_{\max}$.

Therefore we have

$$\mathbb{P}\left(\widehat{u}(a^{i,h,p}) - u(a^{i,h,p}) \leq b\sqrt{\frac{\log h_{\max}\log(?/\delta)}{2^{p+1}}}\right) \geq 1 - \delta/?$$

Then we had an extra union bound other all cells that is bounded by n

□

Lemma 3. *For any planning problem with associated (ν, ρ) (see Property 1), on event ξ, for any depths $h \in [h_{\max}]$, for any $p \in [0 : \lfloor \log_2(h_{\max}/(h^2\gamma^{2h}))\rfloor]$, we have $\perp_{h,p} = h$ if conditions (1) and (2) simultaneously hold true.*
(1) $b\sqrt{\log(4n/\delta)/2^{p+1}} \leq \nu\rho^h$
(2) For all $h' \in [h], h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h'}$.
Finally we have $\perp_{0,p} = 0$.

Proof. We place ourselves on event ξ defined in Lemma 2 and for which we proved that $P(\xi) \geq 1 - \delta$. We fix p.

We prove the statement of the lemma, given that event ξ holds, by induction in the following sense. For a given h and p, we assume the hypotheses of the lemma for that h and p are true and we prove by induction that $\perp_{h',p} = h'$ for $h' \in [h]$.
1° For $h = 0$, we trivially have that $\perp_{h,p} \geq 0$.
2° Now consider $h' > 0$, and assume $\perp_{h'-1,p} = h' - 1$ with the objective to prove that $\perp_{h',p} = h'$.

Therefore, at the end of the processing of depth $h' - 1$, during which we were opening the nodes of depth $h' - 1$ we managed to open an optimal node that we denote $a^{\star,h'-1} \in A^{\star,h'-1}$. Moreover if we consider all the sequence of actions b that one can build by appending any action in A to $a^{\star,h'-1} \in A^{\star,h'-1}$, we have for all such b that $T_{b_{[t]}} \geq \lceil (t-1)2^p\gamma^{t-1}\rceil$ for $t \in [h']$.

Note that by definition there exist an optimal infinite sequence of actions $a^\star \in A^\star$ such that $a^{\star,h'-1} = a^\star_{[h'-1]}$

During phase h' the $\left\lfloor h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right)\right\rfloor$ evaluated nodes from $A^{h'-1}$ with highest values $\{\widehat{u}(a^{h'-1,i})\}_{h'-1,i}$ are opened.

For the purpose of contradiction, let us assume that $a^\star_{[h']}$ is not one of them. This would mean that there exist at least $\left\lfloor h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right)\right\rfloor$ nodes from $A^{h'}$, distinct from $a^\star_{[h']}$, satisfying $\widehat{u}(a^{h',i}) \geq \widehat{u}(a^\star_{[h']})$ and each verifying $T_{a^{h',i}_{[t]}} \geq \lceil t2^p\gamma^t\rceil$ for $t \in [h']$. This means that, for these nodes we have: $u(a_{h',i}) + \nu\rho^{h'} \geq u(a_{h',i}) + \nu\rho^h \overset{\textbf{(a)}}{\geq} u(a_{h',i}) + b\sqrt{\log(4n/\delta)/2^{p+1}} \overset{\textbf{(b)}}{\geq} \widehat{u}(a_{h',i}) \geq \widehat{u}(a_{h',i^\star_{h'}}) \overset{\textbf{(b)}}{\geq} u(a_{h',i^\star_{h'}}) - b\sqrt{\log(4n/\delta)/2^{p+1}} \overset{\textbf{(a)}}{\geq} u(a_{h',i^\star_{h'}}) - \nu\rho^h \geq u(a_{h',i^\star_{h'}}) - \nu\rho^{h'}$, where **(a)** is by assumption of the lemma, **(b)** is because ξ holds. As $u(a_{h',i^\star_{h'}}) \geq v^\star - \nu\rho^{h'}$ by Proposition 1, this means we have $\mathcal{N}^u_{h'}(3\nu\rho^{h'}) \geq \left\lfloor h_{\max}/\left(h'\left\lceil h'2^p\gamma^{h'}\right\rceil\right)\right\rfloor + 1$ (the $+1$ is for $a_{h,i^\star_h}$). However, by assumption of the lemma $h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h'}$. It follows that in general $\mathcal{N}^u_{h'}(3\nu\rho^{h'}) > \left\lfloor C\kappa(\nu,\rho)^{h'}\right\rfloor$. This leads to having a contradiction with the $\kappa^u(\nu,\rho)$ with associated constant C as defined in Definition 2. Indeed, the condition $\mathcal{N}^u_{h'}(3\nu\rho^{h'}) \leq C\kappa(\nu,\rho)^{h'}$ in Definition 2 is equivalent to the condition $\mathcal{N}^u_{h'}(3\nu\rho^{h'}) \leq \left\lfloor C\kappa(\nu,\rho)^{h'}\right\rfloor$ as $\mathcal{N}^u_{h'}(3\nu\rho^{h'})$ is an integer.

□

Lemma 1. *For any planning problem with associated (ν,ρ) (see Property 1), on event ξ, for any depths $h \in [h_{\max}]$, for any $p \in [0 : \lfloor\log_2(h_{\max}/(h^2\gamma^{2h}))\rfloor]$, we have $\perp_{h,p} = h$ if conditions (1) and (2) simultaneously hold true.*
(1) $b\sqrt{\log(4n/\delta)/2^{p+1}} \leq \nu\rho^h$
(2) We distinguish cases and express the condition in each:

Case 1) $h2^p\gamma^{2h} \leq 1$:

$$\frac{h_{\max}}{h} = \frac{h_{\max}}{h\left\lceil h2^p\gamma^{2h}\right\rceil} \geq C\kappa(\nu,\rho)^h$$

And for all $h' \in [h]$, $\dfrac{h_{\max}}{h'^2 2^{p+1}\gamma^{2h'}} \geq C\kappa(\nu,\rho)^{h'}$.

Case 2) $h2^p\gamma^{2h} \geq 1$:

Case 2.1) $\gamma^2\kappa^u \geq 1$: $\dfrac{h_{\max}}{h^2 2^{p+1}\gamma^{2h}} \geq C\kappa(\nu,\rho)^h$

Case 2.2) $\gamma^2\kappa^u \leq 1$: $\dfrac{h_{\max}}{h^2 2^{p+1}} \geq C$.

Proof. To prove this statement we just need to show that we verify the hypotheses of Lemma 3. This means we need to prove that for all $h' \in [h]$, $h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h'}$.

We first consider the case 2) where $h2^p\gamma^h \geq 1$. If $h = 1$ we already know $\perp_{0,p} \geq 0$. Let us now look at the case $h > 1$. First notice that $h2^p\gamma^{2h} \geq 1$ gives $(h-1)2^p\gamma^{2(h-1)} \geq 1$. If $\gamma^2\kappa^u \geq 1$ we have that for all $h' \in [h-1]$,

$$\frac{h_{\max}}{h'^2 2^{p+1}} \geq \frac{h_{\max}}{h^2 2^{p+1}} \geq C\left(\gamma^2\kappa(\nu,\rho)\right)^h \geq C\left(\gamma^2\kappa(\nu,\rho)\right)^{h'}$$

If $h > 1$, and if $\gamma^2\kappa^u \leq 1$ we have that for all $h' \in [h-1]$,

$$\frac{h_{\max}}{h'2^{p+1}} \geq \frac{h_{\max}}{h2^{p+1}} \geq C \geq C\left(\gamma^2\kappa(\nu,\rho)\right)^{h'}.$$

For both $\gamma^2\kappa^u \leq 1$ and $\gamma^2\kappa^u \geq 1$, we then have,

$$\frac{h_{\max}}{h'\left\lceil h'2^p\gamma^{2h'}\right\rceil} \geq \frac{h_{\max}}{h'h'2^{p+1}\gamma^{2h'}}$$

as $h2^p\gamma^{2h} \geq 1$.

For both $\gamma^2\kappa^u \leq 1$ and $\gamma^2\kappa^u \geq 1$, the previous equations mean that for $h' \in [h-1]$, h' verifies $h_{\max}/h'^2 2^{p+1}\gamma^{2h'} \geq C\kappa(\nu,\rho)^{h'} \geq 1$. Therefore $p \leq \left\lfloor \log_2(h_{\max}/(h'^2\gamma^{2h'}))\right\rfloor$.

We now consider case 1) where $h2^p\gamma^{2h} \leq 1$. We prove by induction that for all $h' \in [h], h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h'}$.

1° By assumption of the lemma we say: $h_{\max}/\left(h\left\lceil h2^p\gamma^{2h}\right\rceil\right) \geq C\kappa(\nu,\rho)^h$

2° We further assume $h_{\max}/\left(h'\left\lceil h'2^p\gamma^{2h'}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h'}$ is true for some $h' \leq h$ with $h'2^p\gamma^{2h'} \leq 1$

We want to prove that either:

both $(h'-1)2^p\gamma^{2(h'-1)} \leq 1$

and $\frac{h_{\max}}{(h'-1)\left\lceil (h'-1)2^p\gamma^{2(h'-1)}\right\rceil} \geq C\kappa(\nu,\rho)^{h'-1}$

or $(h'-1)2^p\gamma^{2(h'-1)} \geq 1$

then $h_{\max}/\left((h'')\left\lceil (h'')2^p\gamma^{2(h'')}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h''}$ is already true for all $h'' \in [h']$. If $\left\lceil (h'-1)2^p\gamma^{2(h'-1)}\right\rceil = 1$ then we have

$$\frac{h_{\max}}{(h'-1)} \geq \frac{h_{\max}}{h'} \geq C\kappa(\nu,\rho)^{h'} \geq C\kappa(\nu,\rho)^{h'-1}$$

If $\left\lceil (h'-1)2^p\gamma^{2(h'-1)}\right\rceil > 1$, then we have that

$$h_{\max}/\left(h'\left\lceil (h'-1)2^p\gamma^{2(h'-1)}\right\rceil\right) \geq$$

$$h_{\max}/\left((h'-1)^2 2^{p+1}\gamma^{2(h'-1)}\right) \geq C\kappa(\nu,\rho)^{h'-1}$$

Using this inequality we can now use Case 2) to have that:
$h_{\max}/\left((h'')\left\lceil (h'')2^p\gamma^{2(h'')}\right\rceil\right) \geq C\kappa(\nu,\rho)^{h''}$ is already true for all $h'' \in [h']$.

The previous equations mean that for $h' \in [h-1]$, h' verifies $h_{\max}/\left(h'^2 2^{p+1}\gamma^{2h'}\right) \geq C\kappa(\nu,\rho)^{h'} \geq 1$. Therefore $p \leq \left\lfloor \log_2(h_{\max}/(h'^2\gamma^{2h'}))\right\rfloor$. □

D. Proof of Theorem 3 and Theorem 4

Theorem 3. *High-noise regime After n rounds, for any problem with associated (ν,ρ), and branching factor denoted $\kappa = \kappa^u(\nu,\rho)$, if the noise b is high enough to verify both high noise conditions as defined in the caption of Table 1, the simple regret of* `PlaTγPOOS` obeys

$$\mathbb{E}r_n = \begin{cases} \widetilde{\mathcal{O}}\left(\left(\frac{n}{b^2}\right)^{-\frac{1}{2}}\right) & if\ \gamma^2\kappa \leq 1, \\ \widetilde{\mathcal{O}}\left(\left(\frac{n}{b^2}\right)^{-\frac{\log(\frac{1}{\rho})}{\log(\frac{\gamma^2\kappa}{\rho^2})}}\right) & if\ \gamma^2\kappa > 1. \end{cases}$$

Theorem 4. *Low-noise regime After n rounds, for any problem with associated (ν,ρ), and branching factor denoted $\kappa=\kappa^u(\nu,\rho)$, if the noise b is low enough to verify both high noise conditions as defined in the caption of Table 1, the simple regret of* PlaTγPOOS obeys

$$\mathbb{E}r_n \;=\; \begin{cases}\widetilde{\mathcal{O}}\left(\nu\rho^n\right) & if\ \kappa=1,\\ \widetilde{\mathcal{O}}\left(\nu\left(\frac{n}{b^2}\right)^{-\frac{\log(\frac{1}{\rho})}{\log(\kappa)}}\right) & if\ \kappa>1.\end{cases}$$

Proof of Theorem 3 and Theorem 4. We first place ourselves on the event ξ defined in Lemma 2 and where it is proven that $P(\xi)\geq 1-\delta$. We bound the simple regret of PlaTγPOOS on ξ.

Step 1) General definition of the regret

We chose $\delta=\frac{4b(1-\delta)}{R_{\max}\sqrt{n}}$ for the bound. We consider a problem with associated (ν,ρ). For simplicity we write $\kappa=\kappa^u(\nu,\rho)$. We have for all $p\in[0:p_{\max}]$

$$\begin{aligned}
&v(a^n)+\frac{b}{1-\gamma^2}\sqrt{\frac{p_{\max}\log(R_{\max}n^{3/2}/b)}{2h_{\max}}}\\
&\geq u(a^n)+\frac{b}{1-\gamma^2}\sqrt{\frac{p_{\max}\log(R_{\max}n^{3/2}/b)}{2h_{\max}}}\overset{\textbf{(a)}}{\geq}\widehat{u}(a^n)\\
&\overset{\textbf{(c)}}{\geq}\widehat{u}(a^p)\overset{\textbf{(b)}}{\geq}\widehat{u}\left(a^p_{[\perp_{h_{\max},p}+1]}\right)\\
&\overset{\textbf{(a)}}{\geq}u\left(a^p_{[\perp_{h_{\max},p}+1]}\right)-\frac{b}{1-\gamma^2}\sqrt{\frac{p_{\max}\log(R_{\max}n^{3/2}/b)}{2h_{\max}}}\\
&\overset{\textbf{(d)}}{\geq}v^\star-\nu\rho^{\perp_{h_{\max},p}+1}-\frac{b}{1-\gamma^2}\sqrt{\frac{p_{\max}\log(R_{\max}n^{3/2}/b)}{2h_{\max}}}
\end{aligned}$$

where **(a)** is because the actions at time t, $a_t(n,p)$, of the *candidate* $a(n,p)$ have been evaluated $\frac{(t+1)\gamma^{2t}h_{\max}}{(1-\gamma)^2}$ times and because ξ holds, **(b)** is because $a^p_{[\perp_{h_{\max},p}+1]}\in\{a\in A^\bullet:\forall t\in[2:h(a)],T_{a_{[t]}}\geq\lceil(t-1)2^p\gamma^{2(t-1)}\rceil\}$ and $a^p=\underset{a\in A^\bullet:\forall t\in[2:h(a)],T_{a_{[t]}}\geq\lceil(t-1)2^p\gamma^{2(t-1)}\rceil}{\arg\max}\widehat{u}(a)$, **(c)** is because $a^n=\underset{\{a^p,p\in[0:p_{\max}]\}}{\arg\max}\widehat{u}(a^p)$, and **(d)** is by Assumption 1.

From the previous inequality we have $r_n=v^\star-Q^\star(x,a^n)\leq\nu\rho^{\perp_{h_{\max},p}+1}+2\frac{b}{1-\gamma^2}\sqrt{\frac{p_{\max}\log(R_{\max}n^{3/2}/b)}{2h_{\max}}}$, for $p\in[0:p_{\max}]$.

Step 2) Defining some important depths For the rest of proof we want to lower bound $\max_{p\in[0:p_{\max}]}\perp_{h_{\max},p}$. Lemma 3 and 1 provide some sufficient conditions on p and h to get lower bounds. These conditions are inequalities in which as p gets smaller (fewer samples) or h gets larger (more depth) these conditions are more and more likely not to hold. For our bound on the regret of PlaTγPOOS to be small, we want quantities p and h where the inequalities hold but using as few samples as possible (small p) and having h as large as possible. Therefore we are interested in determining when the inequalities flip signs which is when they turn to equalities. This is what we solve next.

We set the notation $g^{\delta,R_{\max}}_{n,b}=p_{\max}\log(R_{\max}n^{3/2}/b(1-\delta))$.

In its most general form we are interested in the real numbers $\widetilde{h}$ and $\widetilde{p}$ are such that $\widetilde{h}$ is the larger real number such that for all $h\leq\widetilde{h}'$

$$\frac{h_{\max}}{h^2 2^{\widetilde{p}+1}\gamma^{2h}}\geq C\kappa(\nu,\rho)^h \text{ while }\quad b\sqrt{\frac{g^{\delta,R_{\max}}_{n,b}}{2^{\widetilde{p}}}}=\nu\rho^{\widetilde{h}} \tag{5}$$

In the case $\gamma^2\kappa \geq 1$ we can simply solve the following equations. We denote $\widetilde{h}_1$, $\widetilde{p}_1$ the real numbers satisfying

$$\frac{h_{\max}\nu^2\rho^{2\widetilde{h}_1}}{2\widetilde{h}_1^2 b^2 g_{n,b}^{\delta,R_{\max}}\gamma^{2\widetilde{h}_1-2}} = C\kappa^{\widetilde{h}_1} \quad \text{and} \quad b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}_1}}} = \nu\rho^{\widetilde{h}_1}\gamma^2. \tag{6}$$

In the case $\gamma^2\kappa \leq 1$ the previous equation can possess two solutions where the largest of these two solutions will not verify Equation 5. Additionally the smallest solution might be hard to express in a close form when $\gamma^2\kappa \leq 1$. Therefore for simplicity we define for the case $\gamma^2\kappa \leq 1$, $\widetilde{h}_2$, $\widetilde{p}_2$ the real numbers satisfying

$$\frac{h_{\max}\nu^2\rho^{2\widetilde{h}_2}}{\widetilde{2}h_2^2 b^2 g_{n,b}^{\delta,R_{\max}}} = C \quad \text{and} \quad b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}_2}}} = \nu\rho^{\widetilde{h}_2}. \tag{7}$$

$\widetilde{h}_1$ and $\widetilde{p}_1$ are defined for the case $\gamma^2\kappa \geq 1$ while $\widetilde{h}_2$ and $\widetilde{p}_2$ are defined for the case $\gamma^2\kappa \leq 1$. Our approach is to solve Equation 6 and 7 and then verify that it gives a valid indication of the behavior of our algorithm in term of its optimal p and h. We have

$$\widetilde{h}_1 = \frac{2}{\log(\gamma^2\kappa/\rho^2)} W\left(\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)$$

$$\widetilde{h}_2 = \frac{2}{\log(1/\rho^2)} W\left(\log(1/\rho^2)/2\sqrt{\frac{\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)$$

where standard W is the Lambert W function.

However after a close look at the Equation 7, we notice that it is possible to get values of $\widetilde{p}$ and $\widetilde{h}$ which would lead to a number of evaluations $\widetilde{h}2^{\widetilde{p}}\gamma^{\widetilde{h}} < 1$. This actually corresponds to an interesting case when the noise has a small range and where we can expect to obtain an improved result, that is: obtain a regret rate close to the deterministic case. This low range of noise case then has to be considered separately.

Therefore, we distinguish two cases which corresponds to different noise regimes depending on the value of b. Looking at the equation on the right of (7), we have that $\widetilde{h}2^{\widetilde{p}}\gamma^{\widetilde{h}} < 1$ if $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} > 1$. Based on this condition we now consider the two cases. However for both of them we define some generic $\ddot{h}$ and $\ddot{p}$.

Case 1) $\gamma^2\kappa \geq 1$ **:** Note that in this case then $\kappa > 1$. We subdivide this case into multiple subcases:

Case 1.1) Noise regime $\frac{\nu^2\rho^{2\widetilde{h}_1}}{\gamma^{2\widetilde{h}_1}\widetilde{h}_1 b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$

Case 1.1.1) High-noise regime $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$

In this case, we denote $\ddot{h}_1 = \widetilde{h}_1$ and $\ddot{p}_1 = \widetilde{p}_1$. As $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$ by construction, we have $\widetilde{p}_1 \geq 0$. Using standard properties of the $\lfloor\cdot\rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor\widetilde{p}_1\rfloor+1}}} \leq b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}_1}}} \leq \nu\rho^{\widetilde{h}_1} \leq \nu\rho^{\lfloor\widetilde{h}_1\rfloor} \tag{8}$$

$$\begin{aligned}
\text{and,} \quad & \frac{h_{\max}}{\left\lfloor\widetilde{h}_1\right\rfloor\left\lfloor\widetilde{h}_1\right\rfloor 2^{\lfloor\widetilde{p}_1\rfloor+1}\gamma^{2\lfloor\widetilde{h}_1\rfloor}} \geq \frac{h_{\max}}{\left\lfloor\widetilde{h}_1\right\rfloor\left\lfloor\widetilde{h}_1\right\rfloor 2^{\widetilde{p}_1+1}\gamma^{2\lfloor\widetilde{h}_1\rfloor}} \\
\geq\ & \frac{h_{\max}}{\left\lfloor\widetilde{h}_1\right\rfloor\widetilde{h}_1 2^{\widetilde{p}_1+1}\gamma^{2\widetilde{h}_1-2}} = \frac{h_{\max}\nu^2\rho^{2\widetilde{h}_1}}{2\widetilde{h}_1^2 b^2 g_{n,b}^{\delta,R_{\max}}\gamma^{2\widetilde{h}_1}} \\
=\ & C\kappa^{\widetilde{h}_1} \geq C\kappa^{\lfloor\widetilde{h}_1\rfloor}.
\end{aligned}$$

We will verify that $\left\lfloor \ddot{h} \right\rfloor$ is a reachable depth by PlaTγPOOS in the sense that $\ddot{h} \leq h_{\max}$ and $\lfloor \ddot{p} \rfloor \leq \left\lfloor \log_2(h_{\max}/(h^2\gamma^{2h})) \right\rfloor$ and . As $\kappa < 1$, and $\ddot{h} \geq 0$ we have $\kappa^{\ddot{h}} \geq 1$. This gives $C\kappa^{\ddot{h}} \geq 1$. Finally as $\frac{h_{\max}}{\ddot{h}^2 2^{\ddot{p}}\gamma^{2\ddot{h}}} \geq C\kappa^{\ddot{h}}$, we have $\ddot{h}^2\gamma^{2\ddot{h}} \leq h_{\max}/2^{\ddot{p}}$.

Case 1.1.2) Low-noise regime 1 $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

We denote $\ddot{h} = \overline{h}_1$ and $\ddot{p} = \overline{p}_1$ where $\overline{h}$ and $\overline{p}$ verify,

$$\frac{h_{\max}}{2\overline{h}_1^2\gamma^{2\overline{h}_1}} = C\kappa^{\overline{h}_1} \quad \text{and} \quad \overline{p}_1 = 0. \tag{9}$$

Again, $\frac{h_{\max}}{2\overline{h}_1^2 2^{p_1}\gamma^{2\overline{h}_1}} \geq 1$.

$$\overline{h}_1 = \frac{2}{\log(\gamma^2\kappa)} W\left(\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}\right)$$

Using standard properties of the $\lfloor\cdot\rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor\overline{p}_1\rfloor+1}}} \leq b\sqrt{g_{n,b}^{\delta,R_{\max}}} < \nu\rho^{\widetilde{h}_1} \overset{\textbf{(a)}}{\leq} \nu\rho^{\overline{h}_1} \leq \nu\rho^{\lfloor\overline{h}_1\rfloor} \tag{10}$$

where **(a)** is because of the following reasoning. As we have $\frac{h_{\max}\nu^2\rho^{2\widetilde{h}_1}}{2\widetilde{h}_1^2 b^2 g_{n,b}^{\delta,R_{\max}}\gamma^{2\widetilde{h}_1}} = C\kappa^{\widetilde{h}_1}$ and $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$, then, $\frac{h_{\max}}{2\widetilde{h}_1^2\gamma^{2\widetilde{h}_1}} \leq C\kappa^{\widetilde{h}_1}$. From the inequality $\frac{h_{\max}}{2\widetilde{h}_1^2} \leq C\kappa^{\widetilde{h}_1}\gamma^{2\widetilde{h}_1}$ and the fact that $\overline{h}_1$ corresponds to the case of equality $\frac{h_{\max}}{2\overline{h}_1^2} = C\kappa^{\overline{h}_1}\gamma^{2\overline{h}_1}$, we deduce that $\overline{h}_1 \leq \widetilde{h}_1$, since the left term of the inequality decreases with h while the right term increases (as $\gamma^2\kappa \geq 1$). Having $\overline{h}_1 \leq \widetilde{h}_1$ gives $\rho^{\overline{h}_1} \geq \rho^{\widetilde{h}_1}$.

Moreover, the term $\log(\gamma^2\kappa)$ of $\overline{h}_1$ could lead to think that we could potentially obtain a better rate that in the deterministic case where the term is $\log(\gamma^2\kappa)$. However this is not true because as $\overline{h}_1$ is the solution of $\overline{h}_1 = \frac{h_{\max}}{2\overline{h}_1^2\gamma^{2\overline{h}_1}} = C\kappa^{\overline{h}_1}$ and we have by assumption in this case $\overline{h}_1\gamma 2\overline{h}_1 \geq 1$ then $\overline{h}_1 \leq h_3$ where h_3 is defined as the solution of $h_3 = \frac{h_{\max}}{2\overline{h}_3\gamma^{2h_3}} = C\kappa^{h_3}$. We have $h_3 = \frac{1}{\log(\kappa)} W\left(\log(\kappa)\frac{h_{\max}}{2C}\right)$. Therefore one can see that this rate is not better that the deterministic rates.

Case 1.2) Low noise regime 2 $\frac{\nu^2\rho^{2\widetilde{h}_1}}{\gamma^{2\widetilde{h}_1}\widetilde{h}_1 b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

We denote $\ddot{h} = \widehat{h}_1$ and $\ddot{p} = \widehat{p}_1$ where $\widehat{h}$ and $\widehat{p}$ verify,

$$\frac{h_{\max}}{2\widehat{h}_1} = C\kappa^{\widehat{h}_1} \quad \text{and} \quad \widehat{p}_1 = \max(0, \widetilde{p}_1)). \tag{11}$$

$$\widehat{h}_1 = \frac{1}{\log(\kappa)} W\left(\frac{h_{\max}\log(\kappa)}{2C}\right)$$

Using standard properties of the $\lfloor\cdot\rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor\widehat{p}_1\rfloor+1}}} \leq b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}_1}}} < \nu\rho^{\widetilde{h}_1} \overset{\textbf{(a)}}{\leq} \nu\rho^{\widehat{h}_1} \leq \nu\rho^{\lfloor\widehat{h}_1\rfloor} \tag{12}$$

where **(a)** is because of the following reasoning. As we have $\frac{h_{\max}\nu^2\rho^{2\widetilde{h}_1}}{2\widetilde{h}_1^2 b^2 g_{n,b}^{\delta,R_{\max}}\gamma^{2\widetilde{h}_1}} = C\kappa^{\widetilde{h}_1}$ and $\frac{\nu^2\rho^{2\widetilde{h}_1}}{\gamma^{2\widetilde{h}_1}\widetilde{h}_1 b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$, then, $\frac{h_{\max}}{2\widetilde{h}_1} \leq C\kappa^{\widetilde{h}_1}$. From the inequality $\frac{h_{\max}}{2\widetilde{h}_1} \leq C\kappa^{\widetilde{h}_1}$ and the fact that $\widehat{h}_1$ corresponds to the case of equality $\frac{h_{\max}}{2\widehat{h}_1} = C\kappa^{\widehat{h}_1}$, we deduce that $\widehat{h}_1 \leq \widetilde{h}_1$, since the left term of the inequality decreases with h while the right term increases . Having $\widehat{h}_1 \leq \widetilde{h}_1$ gives $\rho^{\widehat{h}_1} \geq \rho^{\widetilde{h}_1}$.

Case 2) $\gamma^2\kappa \leq 1$

Case 2.1) Noise regime $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$

Case 2.1.1) High-noise regime $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$

In this case, we denote $\ddot{h} = \widetilde{h}_2$ and $\ddot{p} = \widetilde{p}_2$. As $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$ by construction, we have $\widetilde{p}_2 \geq 0$. Using standard properties of the $\lfloor\cdot\rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor \widetilde{p}_2\rfloor+1}}} \leq b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}_2}}} = \nu\rho^{\widetilde{h}_2} \leq \nu\rho^{\lfloor \widetilde{h}_2\rfloor} \tag{13}$$

$$\begin{aligned}\text{and,}\quad & \frac{h_{\max}}{\left\lfloor \widetilde{h}_2\right\rfloor \left\lfloor \widetilde{h}_2\right\rfloor 2^{\lfloor \widetilde{p}_2\rfloor+1}} \geq \frac{h_{\max}}{\left\lfloor \widetilde{h}_2\right\rfloor \left\lfloor \widetilde{h}_2\right\rfloor 2^{\widetilde{p}_2+1}}\\ &\geq \frac{h_{\max}}{\widetilde{h}_2^2 2^{\widetilde{p}_2+1}} = \frac{h_{\max}\nu^2\rho^{2\widetilde{h}_2}}{2\widetilde{h}_2^2 b^2 g_{n,b}^{\delta,R_{\max}}}\\ &= C.\end{aligned}$$

We will verify that $\left\lfloor \ddot{h}\right\rfloor$ is a reachable depth by PlaTγPOOS in the sense that $\ddot{h} \leq h_{\max}$ and $\lfloor \ddot{p}\rfloor \leq \left\lfloor \log_2(h_{\max}/(h^2\gamma^{2h}))\right\rfloor$. As $\kappa < 1$, and $\ddot{h} \geq 0$ we have $\kappa^{\ddot{h}} \geq 1$. This gives $C\kappa^{\ddot{h}} \geq 1$. Finally as $\frac{h_{\max}}{\ddot{h}^2 2^{\ddot{p}}\gamma^{2\ddot{h}}} \geq C\kappa^{\ddot{h}}$, we have $\ddot{h}^2\gamma^{2\ddot{h}} \leq h_{\max}/2^{\ddot{p}}$.

Case 2.1.2) Low-noise regime 1 $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

We denote $\ddot{h} = \overline{h}_2$ and $\ddot{p} = \overline{p}_2$ where $\overline{h}$ and $\overline{p}$ verify,

$$\frac{h_{\max}}{2\overline{h}_2^2} = C \quad \text{and} \quad \overline{p}_2 = 0. \tag{14}$$

Again, $\frac{h_{\max}}{2\overline{h}_2^2 2^{p_2}\gamma^{2\overline{h}_2}} \geq 1$.

$$\overline{h}_2 = \sqrt{\frac{h_{\max}}{2C}}$$

Using standard properties of the $\lfloor\cdot\rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor \ddot{p}_2\rfloor+1}}} \leq b\sqrt{g_{n,b}^{\delta,R_{\max}}} < \nu\rho^{\widetilde{h}_1} \overset{\textbf{(a)}}{\leq} \nu\rho^{\overline{h}_2} \leq \nu\rho^{\lfloor \overline{h}_2\rfloor} \tag{15}$$

where **(a)** is because of the following reasoning. As we have $\frac{h_{\max}\nu^2\rho^{2\widetilde{h}_2}}{2\widetilde{h}_2^2 b^2 g_{n,b}^{\delta,R_{\max}}} = C$ and $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$, then, $\frac{h_{\max}}{2\widetilde{h}_2^2} \leq C$. From the inequality $\frac{h_{\max}}{2\widetilde{h}_2^2} \leq C$ and the fact that $\overline{h}_2$ corresponds to the case of equality $\frac{h_{\max}}{2\overline{h}_2^2} = C$, we deduce that $\overline{h}_2 \leq \widetilde{h}_2$, since the left term of the inequality decreases with h while the right term stays constant. Having $\overline{h}_2 \leq \widetilde{h}_2$ gives $\rho^{\overline{h}_2} \geq \rho^{\widetilde{h}_2}$.

Case 2.2) Low noise regime 2 $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

We denote $\ddot{h} = \widehat{h}_2$ and $\ddot{p} = \widehat{p}_2$ where $\widehat{h}$ and $\widehat{p}$ verify,

$$\frac{h_{\max}}{\widehat{h}_2} = C\kappa^{\widehat{h}_2} \tag{16}$$

$$\widehat{h}_2 = \frac{1}{\log(\kappa)} W\left(\frac{h_{\max}\log(\kappa)}{C}\right)$$

By construction, we have $\widetilde{h}_2 \leq \widetilde{h}$. We set

$$\widehat{p}_2 = \max(0, \widetilde{p})). \tag{17}$$

Using standard properties of the $\lfloor \cdot \rfloor$ function, we have

$$b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\lfloor \ddot{p}_2 \rfloor + 1}}} \leq b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2^{\widetilde{p}}}} = \nu\rho^{\widetilde{h}} \overset{\textbf{(a)}}{\leq} \nu\rho^{\widehat{h}_2} \leq \nu\rho^{\lfloor \widehat{h}_2 \rfloor} \tag{18}$$

where **(a)** is because of the following reasoning. As we have $\frac{h_{\max}\nu^2\rho^{2\widetilde{h}}}{\widetilde{h}^2 b^2 g_{n,b}^{\delta,R_{\max}}\gamma^{2\widetilde{h}}} = C\kappa^{\widetilde{h}}$ and $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$, then, $\frac{h_{\max}}{\widetilde{h}} \leq \kappa^{\widetilde{h}}$. From the inequality $\frac{h_{\max}}{\widetilde{h}} \leq C\kappa^{\widetilde{h}}$ and the fact that $\widehat{h}_2$ corresponds to the case of equality $\frac{h_{\max}}{2\widehat{h}_2} = C\kappa^{\widehat{h}_2}$, we deduce that $\widehat{h}_2 \leq \widetilde{h}$, since the left term of the inequality decreases with h while the right term increases. Having $\widehat{h}_2 \leq \widetilde{h}$ gives $\rho^{\widehat{h}_2} \geq \rho^{\widetilde{h}}$.

Step 3 Given these particular definitions of $\ddot{h}$ and $\ddot{p}$ in two distinct cases we now bound the regret.

We always have $\perp_{h_{\max},\lfloor \ddot{p} \rfloor} \geq 0$. If $\ddot{h} \geq 1$, as discussed above $\left\lfloor \ddot{h} \right\rfloor \in [h_{\max}]$, therefore $\perp_{h_{\max},\lfloor \ddot{p} \rfloor} \geq \perp_{\lfloor \ddot{h} \rfloor,\lfloor \ddot{p} \rfloor}$, as $\perp_{\cdot,\lfloor p \rfloor}$ is increasing for all $p \in [0, p_{\max}]$. Moreover on event ξ, and for the cases 1.1.1, 1.1.2, 2.1.1 and 2.1.2 described above, $\perp_{\lfloor \ddot{h} \rfloor,\lfloor \ddot{p} \rfloor} = \left\lfloor \ddot{h} \right\rfloor$ because of Lemma 1 (Case 2)) which assumptions on $\left\lfloor \ddot{h} \right\rfloor$ and $\lfloor \ddot{p} \rfloor$ are verified in each cases as detailed above and, in general, $\left\lfloor \ddot{h} \right\rfloor \in \left[\lfloor h_{\max}/2^{\ddot{p}} \rfloor\right]$ and $\lfloor \ddot{p} \rfloor \in [0 : p_{\max}]$. So, for the aforementioned cases, we have $\perp_{\lfloor h_{\max}/2^{\ddot{p}} \rfloor,\lfloor \ddot{p} \rfloor} \geq \left\lfloor \ddot{h} \right\rfloor$. Very similarly cases 1.2 and 2.2. lead to $\perp_{\lfloor h_{\max}/2^{\ddot{p}} \rfloor,\lfloor \ddot{p} \rfloor} \geq \left\lfloor \ddot{h} \right\rfloor$ by using Lemma 1 (Case 1)).

We bound the regret now discriminating on whether or not the event ξ holds. We have

$$\begin{aligned} r_n &\leq (1-\delta)\left(\nu\rho^{\perp_{h_{\max},\ddot{p}}+1} + 2\frac{b}{1-\gamma^2}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2h_{\max}}}\right) \\ &\quad + \delta \times \frac{R_{\max}}{1-\gamma} \leq \nu\rho^{\perp_{h_{\max},\ddot{p}}+1} + 2\frac{b}{1-\gamma^2}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{2h_{\max}}} + \frac{4b}{\sqrt{n}} \\ &\leq \nu\rho^{\perp_{h_{\max},\ddot{p}}+1} + 6\frac{b}{1-\gamma^2}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}}. \end{aligned}$$

We can now bound the regret in the two regimes.

Case 1) $\gamma^2\kappa \geq 1$ **:** Note that in this case then $\kappa > 1$. We subdivide this case into multiple subcases:

Case 1.1) Noise regime $\frac{\nu^2\rho^{2\widetilde{h}_1}}{\gamma^{2\widetilde{h}_1}\widetilde{h}_1 b^2 g_{n,b}^{\delta,R_{\max}}} \leq 1$

Case 1.1.1) High-noise regime In general,we have

$$r_n \;\leq\; \nu\rho^{\frac{2}{\log(\gamma^2\kappa/\rho^2)}W\left(\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)} + 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}}.$$

Moreover, as proved by Hoorfar & Hassani (2008), the Lambert $W(x)$ function verifies for $x \geq e$, $W(x) \geq \log\left(\frac{x}{\log x}\right)$.

Therefore, if $\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}} > e$ we have,

$$
\begin{aligned}
&r_n - 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \\
&\leq \nu\rho^{\frac{2}{\log(\gamma^2\kappa/\rho^2)}\left(\log\left(\frac{\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}}{\log\left(\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)}\right)\right)} \\
&= \nu e^{\frac{1}{\log(\gamma^2\kappa/\rho^2)}\left(\log\left(\frac{\log^2(\gamma^2\kappa/\rho^2)/2\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}{\log^2\left(\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)}\right)\right)\log(\rho)} \\
&= \nu\left(\frac{\log^2(\gamma^2\kappa/\rho^2)/2\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}{\log^2\left(\log(\gamma^2\kappa/\rho^2)/2\sqrt{\frac{\gamma^2\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)}\right)^{\frac{\log(\rho)}{\log(\gamma^2\kappa/\rho^2)}}.
\end{aligned}
$$

Case 1.1.2) Low-noise regime 1 $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

$$
r_n \;\leq\; \nu\rho^{\frac{2}{\log(\gamma^2\kappa)}W\left(\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}\right)} + 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}}.
$$

Moreover, as proved by Hoorfar & Hassani (2008), the Lambert $W(x)$ function verifies for $x \geq e$, $W(x) \geq \log\left(\frac{x}{\log x}\right)$. Therefore, if $\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}} > e$ we have,

$$
\begin{aligned}
&r_n - 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \\
&\leq \nu\rho^{\frac{2}{\log(\gamma^2\kappa)}\left(\log\left(\frac{\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}}{\log\left(\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}\right)}\right)\right)} \\
&= \nu e^{\frac{1}{\log(\gamma^2\kappa)}\left(\log\left(\frac{\log^2(\gamma^2\kappa)/2\frac{h_{\max}}{2C}}{\log^2\left(\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}\right)}\right)\right)\log(\rho)} \\
&= \nu\left(\frac{\log^2(\gamma^2\kappa)/2\frac{h_{\max}}{2C}}{\log^2\left(\log(\gamma^2\kappa)/2\sqrt{\frac{h_{\max}}{2C}}\right)}\right)^{\frac{\log(\rho)}{\log(\gamma^2\kappa)}}.
\end{aligned}
$$

We have $6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \leq 6\frac{\nu\rho^{\widetilde{h}_1}}{\sqrt{g_{n,b}^{\delta,R_{\max}}}}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \leq 6\nu\rho^{\widetilde{h}_1} \leq 6\nu\rho^{\overline{h}_1}$.

Therefore $r_n \leq \nu\rho^{\perp_{h_{\max},\overline{p}}+1} + 6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \leq 7\nu\rho^{\overline{h}_1}$.

Case 1.2) Low noise regime 2 $\frac{\nu^2\rho^{2\widetilde{h}_1}}{\gamma^{2\widetilde{h}_1}\widetilde{h}_1 b^2 g_{n,b}^{\delta,R_{\max}}} \geq 1$

$$r_n - 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le \nu\left(\frac{\frac{h_{\max}\log(\kappa)}{2C}}{\log\left(\frac{h_{\max}\log(\kappa)}{2C}\right)}\right)^{\frac{\log(\rho)}{\log(\kappa)}}.$$

Moreover if $\frac{\nu^2\rho^{2\widetilde{h}_1}}{b^2 g_{n,b}^{\delta,R_{\max}}} \ge 1$, we have $6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 6\frac{\nu\rho^{\widetilde{h}_1}}{\sqrt{g_{n,b}^{\delta,R_{\max}}}}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 6\nu\rho^{\widetilde{h}_1} \le 6\nu\rho^{\widehat{h}_1}$.

Therefore $r_n \le \nu\rho^{\perp_{h_{\max},\overline{p}}+1} + 6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 7\nu\rho^{\widehat{h}_1}$.

Case 2) $\gamma^2\kappa \le 1$
Case 2.1) Noise regime $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} \le 1$

Case 2.1.1) High-noise regime $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \le 1$

$$r_n - 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le \nu\left(\frac{\log^2(1/\rho^2)/2\frac{\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}{\log^2\left(\log(1/\rho^2)/2\sqrt{\frac{\nu^2 h_{\max}}{2Cb^2 g_{n,b}^{\delta,R_{\max}}}}\right)}\right)^{-\frac{1}{2}}.$$

Case 2.1.2) Low-noise regime 1 $\frac{\nu^2\rho^{2\widetilde{h}_2}}{b^2 g_{n,b}^{\delta,R_{\max}}} \ge 1$

Here with a similar reasoning as in the case 1.1.2) we have $r_n \le 7\nu\rho^{\overline{h}_1} \le 7\nu\rho^{\sqrt{\frac{h_{\max}}{2C}}}$.

Case 2.2) Low noise regime 2 $\frac{\nu^2\rho^{2\widetilde{h}}}{\gamma^{2\widetilde{h}}\widetilde{h}b^2 g_{n,b}^{\delta,R_{\max}}} \ge 1$

$$r_n - 6\frac{b}{1-\gamma}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le \nu\left(\frac{\frac{h_{\max}\log(\kappa)}{C}}{\log\left(\frac{h_{\max}\log(\kappa)}{C}\right)}\right)^{\frac{\log(\rho)}{\log(\kappa)}}.$$

Moreover if $\frac{\nu^2\rho^{2\widetilde{h}}}{b^2 g_{n,b}^{\delta,R_{\max}}} \ge 1$, we have $6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 6\frac{\nu\rho^{\widetilde{h}}}{\sqrt{g_{n,b}^{\delta,R_{\max}}}}\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 6\nu\rho^{\widetilde{h}} \le 6\nu\rho^{\widehat{h}_2}$.

Therefore $r_n \le \nu\rho^{\perp_{h_{\max},\overline{p}}+1} + 6b\sqrt{\frac{g_{n,b}^{\delta,R_{\max}}}{h_{\max}}} \le 7\nu\rho^{\widehat{h}_1}$.

Moreover if $\kappa = 1$ then $r_n \le 7\nu\rho^{\frac{h_{\max}}{C}}$ $\square$

E. Use of the budget

Remark 1. *The algorithm can be made anytime and unaware of n using the classic 'doubling trick'.*

Remark 2 (More efficient use of the budget)**.** *Because of the use of the floor functions $\lfloor \cdot \rfloor$, the budget used in practice can be significantly smaller than n. While this only affects numerical constants in the bounds, in practice, it can noticeably influence the performance. Therefore one should consider, for instance, having $h_{\max}$ replaced by $c \times h_{\max}$ with c been the largest number such that the budget is still smaller than n. Additionally, the use of the budget n could be slightly optimized by taking into account that the necessary number of pulls at depth h cannot be larger than K^h.*