We thank the reviewers for their valuable insights and comments, and for pointing out corrections in the manuscript. We will implement their advice and address all of their remarks in the final version of the manuscript.

One of the referees raised a concern about Assumption 1 (random uniform distribution of the arms). Random design is a classical modelling in many different topics (matrix completion, random design regression, supervised classification). In the bandit literature, the problem of contextual bandits, similar in some respects to the Finite Continuum-Armed Bandit (FCAB) problem, has been studied in Perchet and Rigolet (2013) under the assumption that contexts be randomly distributed in $[0,1]^d$. We underline that the FCAB problem was originally motivated by a pair matching problem, where one aims at discovering unobserved edges on a graph. A typical modelling in this setting is that each node is characterised by a random feature ξ_i, and that given those features the probability that two nodes i and j are linked by an edge is given by a function of ξ_i and ξ_j. This is the case in the well known graphon model; other models of graphs with random covariates have also been considered in Deshpande et al. (2018).

More restrictive, is the assumption of uniform distribution of the arms. Yet, our proofs can readily be extended to cover any distribution with a positive density on $[0,1]$. We chose to focus on the uniform distribution to avoid burdening our proofs with additional technical details. As stated, our aim is to consider the problem of resource allocation with limited budget, which is of practical importance and has not been addressed in the literature, and to expose as simply as possible the differences and similarities between this problem and related problems (among which is the Continuum-Armed Bandits (CAB)).

We thank the referees for raising a very interesting question regarding the dependency of our regret bound on p. We can already address this question without modifying our proofs. In our paper, we show that in the FCAB, when the budget T is a fixed proportion p of the number of arms N, the regret scales as $T^{1/3}\log(T)^{4/3}$. On the other hand, as p decreases, we expect the problem to reduce to a CAB. In the limit where p is sufficiently small, we therefore expect the regret to scale as $\sqrt{T}\log(T)$. Our answer indicates that this is indeed the case, and that a smooth transition occurs between those two settings as $p = T/N$ goes to 0. We present our results before providing a sketch of proof.

To highlight the dependency of R_T on T, we consider regimes where $T = cN^\alpha$ for some $\alpha \in [0,1]$ and $c \in (0, 1/2]$ (the choice $c \le 1/2$ reflects the fact that we are interested in settings where T may be small compared to N, and is arbitrary). Defining $\epsilon_N = (2\log\log(N)/3 - \log(c))/\log(N)$, we obtain that for $\alpha \in (2/3+\epsilon_N, 1]$, $R_T \le CT^{1/(3\alpha)}\log(T)^{4/3}$ with large probability for some constant C depending on Q, L and c. This bound is obtained for the optimal choice of the parameters $K = N^{1/3}\log(N)^{-2/3}$ and $\delta = N^{-4/3}$. Our lower bound on the regret matches this rate up to poly-logarithmic factors. The setting considered in our paper corresponds to the case where $\alpha = 1$, and $R_T = O(T^{1/3}\log(T)^{4/3})$. As expected, as α decreases, the regret increases. When $\alpha \to 2/3 + \epsilon_N$, the regret is of the order $\sqrt{T}\log(T)$, and $K = \sqrt{T}/\log(T)$. We underline that these values correspond respectively to the regret and the optimal value for K in the CAB. This should come as no surprise, as $\alpha = 2/3 + \epsilon_N$ corresponds to a transition from a setting where the finiteness is predominant, to a setting where the difficulty of the problem is that of a CAB.

To understand this transition, note that when $\alpha = 2/3 + \epsilon_N$, we have $T = N/K$. We recall that all intervals contain approximately N/K arms. Thus, when $\alpha > 2/3 + \epsilon_N$, $T > N/K$, the oracle strategy exhausts all arms in the best interval, and it must select arms in other intervals, so the finiteness is a constraining issue. On the contrary, if $\alpha \le 2/3 + \epsilon_N$, no interval is ever exhausted by any strategy, and our problem becomes very close to the CAB. The oracle strategy only selects arms from the interval with highest mean reward. The analysis of the problem becomes much simpler, as results can be directly inferred from Auer et al. (2007) by noticing that Algorithm UCBF never exhausts any intervals, and is therefore a variant of Algorithm UCBC. In this case, the optimal choice for the number of intervals remains $K = \sqrt{T}/\log(T)$, and yields a regret bound $R_T = O(\sqrt{T}\log(T))$ (choosing as previously $K = N^{1/3}\log(N)^{-2/3}$ would produce too many intervals, which we would not have time to explore). The following table summarises these results for $T = cN^\alpha$.

α	$0 \le \alpha \le 2/3 + \epsilon_N$	$2/3 + \epsilon_N < \alpha \le 1$
Optimal K	$K = \sqrt{T}/\log(T)$	$K = N^{1/3}\log(N)^{-2/3}$
Regret	$R_T = O(\sqrt{T}\log(T))$	$R_T = O(T^{1/(3\alpha)}\log(T)^{4/3})$

Sketch of proof: Along the proof of Theorem 1 (at line 486 in the Appendix), we show that for $\delta = N^{-4/3}$ and the optimal choice $K = N^{1/3}\log(N)^{-2/3}$, with large probability $R_T \le CN^{1/3}\log(N)^{4/3}$ for some C depending only on L and Q when $K > p^{-1} \wedge (1-p)^{-1}$. This bound can be rephrased as $R_T \le CT^{1/(3\alpha)}\log(T)^{4/3}$, where C depends on L, Q and c. We assumed that $T \le N/2$, so $p \le 1-p$, and the condition $K > p^{-1} \wedge (1-p)^{-1}$ is met if $\alpha > 2/3 + \epsilon_N$. On the other hand, Theorem 2 shows that with positive probability, $R_T \ge 0.01T^{1/3}p^{-1/3} = 0.01T^{1/(3\alpha)}c^{-1/(3\alpha)}$ when $N \ge N_{L,p}$, where $N_{L,p}$ is defined at line 504 in the Appendix. This condition is met when N is larger than some absolute constant, and $\alpha \ge 2/3 + C/\log(N)$ for some constant C depending on L.