SUPPLEMENTARY MATERIAL FOR "TIME FAIRNESS IN ONLINE KNAPSACK PROBLEMS"

Table 1: A summary of key notations.

Notation	**Description**
$j \in [n]$	Current position in sequence of items
$x_j \in \{0, 1\}$	Decision for jth item. $x_j = 1$ if accepted, $x_j = 0$ if not accepted
U	Upper bound on the value density of any item
L	Lower bound on the value density of any item
α	Fairness parameter, defined in Definition 3.4
w_j	(*Online input*) Item weight revealed to the player when jth item arrives
v_j	(*Online input*) Item value revealed to the player when jth item arrives

A RELATED WORK

We consider the online knapsack problem (OKP), a classical resource allocation problem wherein items with different weights and values arrive sequentially, and we wish to admit them into a capacity-limited knapsack, maximizing the total value subject to the capacity constraint.

Our work contributes to several active lines of research, including a rich literature on OKP first studied in Marchetti-Spaccamela and Vercellis (1995), with important optimal results following in Zhou et al. (2008). In the past several years, research in this area has surged, with many works considering variants of the problem, such as removable items Cygan et al. (2016), item departures Sun et al. (2022), and generalizations to multidimensional settings Yang et al. (2021). Closest to this work, several studies have considered the online knapsack problem with additional information or in a learning-augmented setting, including frequency predictions Im et al. (2021), online learning Zeynali et al. (2021), and advice complexity Böckenhauer et al. (2014).

In the literature on the knapsack problem, fairness has been recently considered in Patel et al. (2020) and Fluschnik et al. (2019). Unlike our work, both deal exclusively with the standard offline setting of knapsack, and consequently do not consider time fairness. (Patel et al., 2020) introduces a notion of *group fairness* for the knapsack problem, while (Fluschnik et al., 2019) introduces three notions of "individually best, diverse, and fair" knapsacks which aggregate the voting preferences of multiple voters. To the best of our knowledge, our work is the first to consider notions of fairness in the online knapsack problem.

In the broader literature on online algorithms and dynamic settings, several studies explore fairness, although most consider notions of fairness that are different from the ones in our work. Online fairness has been explored in resource allocation Manshadi et al. (2021); Sinclair et al. (2020); Bateni et al. (2016); Sinha et al. (2023), fair division Kash et al. (2013), refugee assignment Freund et al. (2023), matchings Deng et al. (2023); Ma et al. (2023), prophet inequalities Correa et al. (2021), organ allocation Bertsimas et al. (2013), and online selection Benomar et al. (2023). Banerjee et al. (2022) shows competitive algorithms for online resource allocation which seek to maximize the Nash social welfare, a metric which quantifies a trade-off between fairness and performance. Deng et al. (2023) also explicitly considers the intersection between predictions and fairness in the online setting. In addition to these problems, there are also several online problems adjacent to OKP in the literature which would be interesting to explore from a fairness perspective, including one-way trading El-Yaniv et al. (2001); Sun et al. (2022), bin packing Balogh et al. (2017); Johnson et al. (1974), and single-leg revenue management Balseiro et al. (2023); Ma et al. (2021).

In the online learning literature, several works consider fairness using regret as a performance metric. In particular, Talebi and Proutiere (2018) studies a stochastic multi-armed bandit setting where tasks must be assigned to servers in a proportionally fair manner. Several other works including Baek and Farias (2021); Patil et al. (2021); Chen et al. (2020) build on these results using different notions of fairness. For a general resource allocation problem, Sinha et al. (2023) presents a fair online resource allocation policy achieving sublinear regret with respect to an offline optimal allocation. Furthermore, many works in the regret setting explicitly consider fairness from the perspective of an α-fair utility function, including Sinha et al. (2023); Si Salem et al. (2022); Wang et al. (2022) , which partially inspires our *parameterized* definition of α-CTIF (Def. 3.4).

In the broader ML research community, fairness has seen burgeoning recent interest, particularly from the perspective of bias in trained models. Mirroring the literature above, multiple definitions of fairness in this setting have been proposed and studied, including equality of opportunity Hardt et al. (2016), conflicting notions of fairness Kleinberg et al. (2017), and inherent trade-offs between performance and fairness constraints Bertsimas et al. (2012). Fairness has also been extensively studied in impact studies such as Chouldechova (2017), which demonstrated the disparate impact of recidivism prediction algorithms used in the criminal justice system on different demographics.

B Proofs for Section 3 (Time Fairness)

Observation 3.2. *The* ZCL *algorithm (Zhou et al., 2008) is not* TIF.

Proof. Let $z_j \in [0, 1]$ be the knapsack's utilization when the jth item arrives. When the knapsack is empty, $z_j = 0$, and when the knapsack is close to full, $z_j = 1 - \epsilon$ for some small $\epsilon > 0$. Pick any instance with sufficiently many items, and pick $z_A < z_B$ such that at least one admitted item, say the kth one, satisfies $\Phi(z_A) \leq v_k < \Phi(z_B)$. Note that this implies that the kth item arrived between the utilization being at z_A and at z_B. Now, modify this same instance by adding a copy of the kth item after the utilization has reached z_B. Note that this item now has probability zero of being admitted. This implies that two items with the same value-to-weight ratio have different probabilities of being admitted into the knapsack, contradicting the definition of TIF. □

Theorem 3.3. *There is no nontrivial algorithm for* OKP *that guarantees* TIF *without additional information about the input. Further, even if the input length n or perfect frequency predictions as defined in Im et al. (2021) are known in advance, no nontrivial algorithm can guarantee* TIF.

Proof. We prove the three parts one by one.

1. WLOG assume that all items have the same value density $x = L = U$. Assume for the sake of a contradiction that there is some algorithm ALG which guarantees TIF, and suppose we have a instance $\mathcal{I}$ with n items (which we will set later).

   Since we assume that ALG guarantees TIF, consider $p(x)$, the probability of admitting an item with value density x. Let $p(x) = \mathtt{p}$. Note that $\mathtt{p} > 0$, and also that it cannot depend on the input length n. Note that we can have have all items of equal weight w, so that for $n \gg 3B/w\mathtt{p}$, the knapsack is full with high probability. But now consider modifying the input sequence from $\mathcal{I}$ to $\mathcal{I}'$, by appending a copy of itself, i.e., increasing the input with another n items of exactly the same type. The probability of admitting these new items must be (eventually) zero, even though they have the same value-to-weight ratio as the first half of the input (and, indeed, the same weights). Therefore, the algorithm violates the TIF constraint on the instance $\mathcal{I}'$, which is a contradiction.

2. Again WLOG assume that all items have the same value density $x = L = U$, so that $s_x = \sum_{i=1}^{n} w_i$. Assume for the sake of a contradiction that there is some competitive algorithm $\mathsf{ALG}^{\mathtt{FP}}$ which uses frequency predictions to guarantee TIF.

   Since we assume that $\mathsf{ALG}^{\mathtt{FP}}$ guarantees TIF, consider $p(x)$, the probability of admitting any item. Let $p(x) = \mathtt{p}$. Again, note that $\mathtt{p} > 0$, and also that it cannot depend on the input length n, but can depend on s_x this time.

   Consider two instances $\mathcal{I}$ and $\mathcal{I}'$ as follows: $\mathcal{I}$ consists exclusively of "small" items of (constant) weight $w_\delta \ll B$ each, whereas $\mathcal{I}'$ consists of a small, constant number of such small items followed by a single "large" item of weight $B - \delta/2$. Note that by taking enough items in $\mathcal{I}$, we can ensure the two instances have the same total weight $s(x)$. Therefore, $p(x)$ must be the same for these two instances, by assumption. Of course, the items all have value x by our original assumption.

Note that the optimal packing in both instances would nearly fill up the knapsack. However, $\mathcal{I}'$ has the property that any competitive algorithm must reject all the initial smaller items, as admitting any of them would imply that the large item can no longer be admitted. By making w_δ arbitrarily small, we can make the algorithm arbitrarily non-competitive.

The instance $\mathcal{I}$ guarantees that $p(x)$ is sufficiently large (i.e., bounded below by some constant), and so with high probability, at least one item in $\mathcal{I}'$ is admitted within the first constant number of items. Therefore, with high probability, $\mathsf{ALG}^{\mathrm{FP}}$ does not admit the large-valued item in $\mathcal{I}'$, and so it cannot be competitive.

3. Again WLOG assume that all items have the same value density $x = L = U$. Assume for the sake of a contradiction that there is some competitive algorithm $\mathsf{ALG}^{\mathrm{N}}$ which uses knowledge of the input length n to guarantee TIF.

   We will consider only input sequences of length n (assumed to be sufficiently large), consisting only of items with value density x. Again, since we assume that $\mathsf{ALG}^{\mathrm{FP}}$ guarantees TIF, consider $p(x)$, the probability of admitting any item. Let $p(x) = \mathrm{p}$. Again, note that $\mathrm{p} > 0$, and also that it must be the same for all input sequences of length n.

   Consider such an instance $\mathcal{I}$, consisting of identical items of (constant) weight w_c each. Suppose the total weight of the items is very close to the knapsack capacity B. Since the expected number of items admitted is $n\mathrm{p}$, the total value admitted is $x \cdot n\mathrm{p}$ on expectation. The optimal solution admits a total value of nx (since the total weight is close to B), and therefore, the competitive ratio is roughly $1/\mathrm{p}$. Since we assumed the algorithm was competitive, it follows that p must be bounded below by a constant.

   Now consider a different instance $\mathcal{I}'$, consisting of $3\log(n)/\mathrm{p}^2$ items of weight $w_\delta \ll B$, followed by $n - 3\log(n)/\mathrm{p}^2$ "large" items of weight $B - w_\delta/2$. Note that these are well-defined, as p is bounded below by a constant, and n is sufficiently large. The instance $\mathcal{I}'$ again has the property that any competitive algorithm must reject all the initial smaller items, as admitting any of them would imply that none of the large items can be admitted.

   However, by the coupon collector problem, with high probability $(1 - \mathrm{poly}(1/n))$, at least one of the $3\log(n)/\mathrm{p}^2$ small items is admitted, which contradicts the competitiveness of ALG^{N}. As before, by making w_δ arbitrarily small, we can make the algorithm arbitrarily non-competitive. □

Observation 3.5. *The* ZCL *algorithm (Zhou et al., 2008) is not* 1-CTIF.

Proof. This follows immediately from the fact that the threshold value in ZCL changes each time an item is accepted, which corresponds to the utilization changing. Consider two items with the same value density (close to L), where one of the items arrives first in the sequence, and the other arrives when the knapsack is roughly half-full, and assume that there is enough space in the knapsack to accommodate both items when they arrive. The earlier item will be admitted with certainty, whereas the later item will with high probability be rejected. So despite having the same value, the items will have a different admission probability purely based on their position in the sequence, violating 1-CTIF. □

C PROOFS FOR SECTION 4 (ONLINE FAIR ALGORITHMS)

Proposition 4.1. *Any constant threshold-based algorithm for* OKP *is* 1-CTIF. *Furthermore, any constant threshold-based* deterministic *algorithm for* OKP *cannot be better than* (U/L)*-competitive.*

Proof. Consider an arbitrary threshold-based algorithm ALG with constant threshold value ϕ. For any instance $\mathcal{I}$, and any item, say the jth one, in this instance, note that the probability of admitting the item depends entirely on the threshold ϕ, and nothing else, as long there is enough space in the knapsack to admit it. So for any value density $x \in [L, U]$, the admission probability $p(x)$ is just the indicator variable capturing whether there is space for the item or not.

For the second part, given a deterministic ALG with a fixed constant threshold $\phi \in [L, U]$, there are two cases. If $\phi > L$, the instance $\mathcal{I}$ consisting entirely of L-valued items induces an unbounded competitive ratio, as no items are admitted by ALG. If $\phi = L$, consider the instance $\mathcal{I}'$ consisting of m equal-weight items with value L followed by m items with value U, and take m large enough that the knapsack can become full with only L-valued items. ALG here admits only L-valued items, whereas the optimal solution only admits U-valued items, and so ALG cannot do better than the worst-case competitive ratio of $\frac{U}{L}$ for OKP. □

Proposition 4.2. *The* ZCL *algorithm is* $\frac{1}{\ln(U/L)+1}$-CTIF.

Proof. Consider the interval $\left[0, \frac{1}{\ln(\frac{U}{L})+1}\right]$, viewed as an utilization interval. An examination of the ZCL algorithm reveals that the value of the threshold is below L on this subinterval. But since we have a guarantee that the value-to-weight ratio is at least L, while the utilization is within this interval, the ZCL algorithm is exactly equivalent to the algorithm using the constant threshold L. By Proposition 4.1, therefore, the algorithm is 1-CTIF within this interval, and therefore is $\frac{1}{\ln(\frac{U}{L})+1}$-CTIF. □

Theorem 4.3. *For* $\alpha \in [1/\ln(U/L)+1, 1]$*, the baseline algorithm is* $\frac{U[\ln(U/L)+1]}{L\alpha[\ln(U/L)+1]+(U-L)(1-\ell)}$-*competitive and* α-CTIF *for* OKP.

Proof. To prove the competitive ratio of the parameterized baseline algorithm (call it BASE$[\alpha]$), consider the following:

Fix an arbitrary instance $\mathcal{I} \in \Omega$. When the algorithm terminates, suppose the utilization of the knapsack is z_T. Assume we obtain a value of BASE$[\alpha](\mathcal{I})$. Let $\mathcal{P}$ and $\mathcal{P}^\star$ respectively be the sets of items picked by BASE$[\alpha]$ and the optimal solution.

Denote the weight and the value of the common items (i.e., the items picked by both BASE and OPT) by $W = w(\mathcal{P} \cap \mathcal{P}^\star)$ and $V = v(\mathcal{P} \cap \mathcal{P}^\star)$. For each item j which is *not accepted* by BASE$[\alpha]$, we know that its value density is $< \Phi^\alpha(z_j) \leq \Phi^\alpha(z_T)$ since Φ^α is a non-decreasing function of z. Thus, using $B = 1$, we get

$$\mathsf{OPT}(\mathcal{I}) \leq V + \Phi^\alpha(z_T)(1 - W).$$

Since BASE$[\alpha](\mathcal{I}) = V + v(\mathcal{P} \setminus \mathcal{P}^\star)$, the inequality above implies that

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{BASE}[\alpha](\mathcal{I})} \leq \frac{V + \Phi^\alpha(z_T)(1 - W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)}.$$

Note that, by definition of the algorithm, each item j picked in $\mathcal{P}$ must have value density of at least $\Phi^\alpha(z_j)$, where z_j is the knapsack utilization when that item arrives. Thus, we have:

$$V \geq \sum_{j \in \mathcal{P} \cap \mathcal{P}^\star} \Phi^\alpha(z_j)\, w_j =: V_1,$$

$$v(\mathcal{P} \setminus \mathcal{P}^\star) \geq \sum_{j \in \mathcal{P} \setminus \mathcal{P}^\star} \Phi^\alpha(z_j)\, w_j =: V_2.$$

Since OPT$(\mathcal{I}) \geq$ BASE$[\alpha](\mathcal{I})$, we have:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{BASE}[\alpha](\mathcal{I})} \leq \frac{V + \Phi^\alpha(z_T)(1 - W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)} \leq \frac{V_1 + \Phi^\alpha(z_T)(1 - W)}{V_1 + v(\mathcal{P} \setminus \mathcal{P}^\star)} \leq \frac{V_1 + \Phi^\alpha(z_T)(1 - W)}{V_1 + V_2},$$

where the second inequality follows because $\Phi^\alpha(z_T)(1 - W) \geq v(\mathcal{P} \setminus \mathcal{P}^\star)$ and $V_1 \leq V$.

Note that $V_1 \leq \Phi^\alpha(z_T) w(\mathcal{P} \cap \mathcal{P}^\star) = \Phi^\alpha(z_T) W$, and by plugging in the actual values of V_1 and V_2, we get:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{BASE}[\alpha](\mathcal{I})} \leq \frac{\Phi^\alpha(z_T)}{\sum_{j \in \mathcal{P}} \Phi^\alpha(z_j)\, w_j}.$$

Based on the assumption that individual item weights are much smaller than 1, we can substitute for w_j with $\delta z_j = z_{j+1} - z_j$ for all j. This substitution gives an approximate value of the summation via integration.

$$\begin{aligned}
\sum_{j\in\mathcal{P}} \Phi^\alpha(z_j)\, w_j \approx \int_0^{z_T} \Phi^\alpha(z) dz \\
&= \int_0^\alpha L dz + \int_\alpha^{z_T} \left(\frac{Ue}{L}\right)^{\frac{z-\ell}{1-\ell}} \left(\frac{L}{e}\right) dz \\
&= L\alpha + \left(\frac{L}{e}\right)(1-\ell) \left[\frac{\left(\frac{Ue}{L}\right)^{\frac{z-\ell}{1-\ell}}}{\ln(Ue/L)}\right]_\alpha^{z_T} \\
&= L\alpha + \frac{\Phi^\alpha(z_T)}{\ln(Ue/L)} - \frac{\Phi^\alpha(\alpha)}{\ln(Ue/L)} - \frac{\ell\Phi^\alpha(z_T)}{\ln(Ue/L)} + \frac{\ell\Phi^\alpha(\alpha)}{\ln(Ue/L)} \\
&= L\left(\alpha - \frac{1}{\ln(Ue/L)} + \frac{\ell}{\ln(Ue/L)}\right) + \frac{\Phi^\alpha(z_T)}{\ln(Ue/L)} - \frac{\ell\Phi^\alpha(z_T)}{\ln(Ue/L)} \\
&= L\left(\alpha - \frac{1-\ell}{\ln(U/L)+1}\right) + \frac{(1-\ell)\Phi^\alpha(z_T)}{\ln(U/L)+1},
\end{aligned}$$

where the fourth equality has used the fact that $\Phi^\alpha(z_j) = (L/e)(Ue/L)^{\frac{z-\ell}{1-\ell}}$, and the fifth equality has used $\Phi^\alpha(\alpha) = L$. Substituting back in, we get:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{BASE}[\alpha](\mathcal{I})} \leq \frac{\Phi^\alpha(z_T)}{L\left(\alpha - \frac{1-\ell}{\ln(U/L)+1}\right) + \frac{(1-\ell)\Phi^\alpha(z_T)}{\ln(U/L)+1}} \leq \frac{U[\ln(U/L)+1]}{L\alpha[\ln(U/L)+1] - L(1-\ell) + U(1-\ell)}.$$

Thus, the baseline algorithm is $\frac{U[\ln(U/L)+1]}{L\alpha[\ln(U/L)+1]-L(1-\ell)+U(1-\ell)}$-competitive.

The fairness constraint of α-CTIF is immediate, because the threshold $\Phi^\alpha(z) \leq L$ in the interval $[0, \alpha]$, and so it can be replaced by the constant threshold L in that interval. Applying Proposition 4.1 yields the result. □

Theorem 4.5. *No α-CTIF deterministic online algorithm for* OKP *can achieve a competitive ratio smaller than $\frac{W\left(\frac{U(1-\alpha)}{L\alpha}\right)}{1-\alpha}$, where $W(\cdot)$ is the Lambert W function.*

Proof. For any α-CTIF deterministic online algorithm ALG, there must exist some utilization region $[a, b]$ with $b - a = \alpha$. Any item that arrives in this region is treated fairly, i.e., by definition of CTIF there exists a function $p(x) : [L, U] \to \{0, 1\}$ which characterizes the fair decisions of ALG. We define $v = \min\{x \in [L, U] : p(x) = 1\}$ (i.e., v is the lowest value density that ALG is willing to accept during the fair region).

We first state a lemma (proven afterwards), which asserts that the feasible competitive ratio for any α-CTIF deterministic online algorithm with $v > L$ is *strictly worse* than the feasible competitive ratio when $v = L$.

Lemma C.1. *For any α-CTIF deterministic online algorithm* ALG *for* OKP*, if the minimum value density v that* ALG *accepts during the fair region of the utilization (of length α) is $> L$, then it must have a competitive ratio $\beta' \geq W(\frac{U(1-\alpha)}{L\alpha} e^{\frac{1}{\alpha}})/(1-\alpha) - \frac{1}{\alpha}$, where $W(\cdot)$ is the Lambert W function.*

By Lemma C.1, it suffices to consider the algorithms that set $v = L$.

Given ALG, let $g(x) : [L, U] \to [0, 1]$ denote the *acceptance function* of ALG, where $g(x)$ is the final knapsack utilization under the instance $\mathcal{I}_x$. Note that for small δ, processing $\mathcal{I}_{x+\delta}$ is equivalent to first processing $\mathcal{I}_x$, and then processing m identical items, each with weight $\frac{1}{m}$ and value density $x + \delta$. Since this function is unidirectional (item acceptances are irrevocable) and deterministic, we must have $g(x + \delta) \geq g(x)$, i.e. $g(x)$ is non-decreasing in $[L, U]$. Once a batch of items with maximum value density U arrives, the rest of the capacity should be used, i.e., $g(U) = 1$.

For any algorithm with $v = L$, it will admit *all* items greedily for an α fraction of the knapsack. Therefore, under the instance $\mathcal{I}_x$, the online algorithm with acceptance function g obtains a value of $\mathsf{ALG}(\mathcal{I}_x) = \alpha L + g(r)r + \int_r^x udg(u)$, where $udg(u)$ is the value obtained by accepting items with value density u and total weight $dg(u)$, and r is defined as the lowest value density that ALG is willing to accept during the unfair region, i.e., $r = \inf_{x\in(L,U]:g(x)\geq\alpha} x$.

For any β'-competitive algorithm, we must have $r \leq \alpha\beta' L$ since otherwise the worst-case ratio is larger than β' under an instance $\mathcal{I}_x$ with $x = \alpha\beta' L + \epsilon$, ($\epsilon > 0$). To derive a lower bound of the competitive ratio, observe that it suffices WLOG to focus on algorithms with $r = \alpha\beta' L$. This is because if a β'-competitive algorithm sets $r < \alpha\beta' L$, then an alternative algorithm can postpone the item acceptance to $r = \alpha\beta' L$ and maintain β'-competitiveness.

Under the instance $\mathcal{I}_x$, the offline optimal solution obtains a total value of $\mathsf{OPT}(\mathcal{I}_x) = x$. Therefore, any β'-competitive online algorithm must satisfy:

$$\mathsf{ALG}(\mathcal{I}_x) = g(L)L + \int_L^x udg(u) = \alpha L + \int_r^x udg(u) \geq \frac{x}{\beta'}, \quad \forall x \in [L, U].$$

By integral by parts and Grönwall's Inequality (Theorem 1, p. 356, in Mitrinovic et al. (1991)), a necessary condition for the competitive constraint above to hold is the following:

$$g(x) \geq \frac{1}{\beta'} - \frac{\alpha L - g(r)r}{x} + \frac{1}{x}\int_r^x g(u)du \tag{2}$$

$$\geq \frac{1}{\beta'} - \frac{\alpha L}{r} + g(r) + \frac{1}{\beta'}\ln\frac{x}{r} = g(\alpha\beta' L) + \frac{1}{\beta'}\ln\frac{x}{\alpha\beta' L}, \quad \forall x \in [L, U]. \tag{3}$$

By combining $g(U) = 1$ with equation (3), it follows that any deterministic α-CTIF and β'-competitive algorithm must satisfy $1 = g(U) \geq g(\alpha\beta' L) + \frac{1}{\beta'}\ln\frac{x}{\alpha\beta' L} \geq \alpha + \frac{1}{\beta'}\ln\frac{x}{\alpha\beta' L}$. The minimal value for β' can be achieved when both inequalities are tight, and is the solution to $\ln(\frac{U}{\alpha\beta' L})/\beta' = 1 - \alpha$. Thus, $\frac{W\left(\frac{U - U\alpha}{L\alpha}\right)}{1-\alpha}$ is a lower bound of the competitive ratio, where $W(\cdot)$ denotes the Lambert W function. □

Proof of Lemma C.1 We use the same definition of the acceptance function $g(x)$ as that in Theorem 4.5. Based on the choice of v by ALG, we consider the following two cases.

Case I: when $v \geq \frac{U}{1+\alpha\beta'}$. Under the instance $\mathcal{I}_x$ with $x \in [L, v)$, the offline optimum is $\mathsf{OPT}(\mathcal{I}_x) = x$ and ALG can achieve $\mathsf{ALG}(\mathcal{I}_x) = Lg(L) + \int_L^x udg(u)$. Thus, any β'-competitive algorithm must satisfy:

$$\mathsf{ALG}(\mathcal{I}_x) = Lg(L) + \int_L^x udg(u) \geq \frac{x}{\beta'}, \quad x \in [L, v).$$

By integral by parts and Grönwall's Inequality (Theorem 1, p. 356, in Mitrinovic et al. (1991)), a necessary condition for the inequality above to hold is:

$$g(x) \geq \frac{1}{\beta'} + \frac{1}{x}\int_L^x g(u)du \geq \frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{x}{L}, \forall x \in [L, v).$$

Under the instance $\mathcal{I}_v$, to maintain α-CTIF , we must have $g(v) \geq \lim_{x\to v}[\frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{x}{L}] + \alpha = \frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{v}{L} + \alpha$. Thus, we have $1 \geq g(v) \geq \frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{v}{L} + \alpha$, which gives:

$$\beta' \geq \frac{1 + \ln\frac{v}{L}}{1-\alpha}. \tag{4}$$

This lower bound is achieved when $g(x) = \frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{x}{L}, x \in [L, v)$ and $g(v) = \frac{1}{\beta'} + \frac{1}{\beta'}\ln\frac{x}{L} + v$. In addition, the total value of accepted items is $\mathsf{ALG}(\mathcal{I}_v) = \frac{v}{\beta'} + \alpha v$.

Under the instance $\mathcal{I}_x$ with $x \in (v, U]$, we observe that the worst-case ratio is:

$$\frac{\mathsf{OPT}(\mathcal{I}_x)}{\mathsf{ALG}(\mathcal{I}_x)} \leq \frac{U}{\mathsf{ALG}(\mathcal{I}_v)} = \beta'\frac{U}{v(1+\alpha\beta')} \leq \beta'.$$

Thus, the lower bound of the competitive ratio is dominated by equation (4), and $\beta' \geq \frac{1+\ln\frac{v}{L}}{1-\alpha} \geq \frac{1+\ln\frac{U}{L(1+\alpha\beta')}}{1-\alpha}$.

Case II: when $L < v < \frac{U}{1+\alpha\beta'}$. In this case, we have the same results under instances $\mathcal{I}_x, x \in [L, v]$. In particular, $g(v) = \frac{1}{\beta'} + \frac{1}{\beta'} \ln \frac{x}{L} + v$ and $\mathsf{ALG}(\mathcal{I}_v) = \frac{v}{\beta'} + \alpha v$.

Under the instance $\mathcal{I}_x$ with $x \in (v, U]$, the online algorithm can achieve

$$\mathsf{ALG}(\mathcal{I}_x) = \mathsf{ALG}(\mathcal{I}_v) + \int_r^x u dg(u), x \in (v, U], \tag{5}$$

where r is the lowest value density that ALG admits outside of the fair region. Using the same argument as that in the proof of Theorem 4.5, WLOG we can consider $r = \beta' \cdot \mathsf{ALG}(\mathcal{I}_v) = v(1 + \alpha\beta'), (r < U)$. By integral by parts and Grönwall's Inequality, a necessary condition for equation (5) is:

$$\begin{aligned} g(x) &\geq \frac{1}{\beta'} - \frac{\mathsf{ALG}(\mathcal{I}_v) - g(r)r}{x} + \frac{1}{x}\int_r^x g(u)du \\ &\geq \frac{1}{\beta'} - \frac{\mathsf{ALG}(\mathcal{I}_v) - g(r)r}{r} + \frac{1}{\beta'} \ln \frac{x}{r} \\ &= g(r) + \frac{1}{\beta'} \ln \frac{x}{r}. \end{aligned}$$

Combining with $g(U) = 1$ and $g(r) \geq g(v)$, we have:

$$1 = g(U) \geq g(r) + \frac{1}{\beta'} \ln \frac{U}{r} \geq \alpha + \frac{1}{\beta'} + \frac{1}{\beta'} \ln \frac{v}{L} + \frac{1}{\beta'} \ln \frac{U}{v(1+\alpha\beta')},$$

and thus, the competitive ratio must satisfy:

$$\beta' \geq \frac{1 + \ln \frac{U}{L(1+\alpha\beta')}}{1-\alpha}. \tag{6}$$

Recall that under the instance $\mathcal{I}_x$ with $x \in [L, v)$, the worst-case ratio is:

$$\frac{\mathsf{OPT}(\mathcal{I}_x)}{\mathsf{ALG}(\mathcal{I}_x)} \leq \frac{1 + \ln \frac{v}{L}}{1-\alpha} < \frac{1 + \ln \frac{U}{L(1+\alpha\beta')}}{1-\alpha}.$$

Therefore, the lower bound is dominated by equation (6).

Summarizing above two cases, for any α-CTIF deterministic online algorithm, if the minimum value density v that it is willing to accept during the fair region is $> L$, then its competitive ratio must satisfy $\beta' \geq \frac{1+\ln \frac{U}{L(1+\alpha\beta')}}{1-\alpha}$, and the lower bound of the competitive ratio is $\frac{W\left(\frac{(1-\alpha)U}{L\alpha} e^{1/\alpha}\right)}{1-\alpha}$. It is also easy to verify that $\frac{W\left(\frac{(1-\alpha)U}{L\alpha} e^{1/\alpha}\right)}{1-\alpha} \geq \frac{W\left(\frac{U - U\alpha}{L\alpha}\right)}{1-\alpha}, \forall\ \alpha \in [\frac{1}{\ln(U/L)+1}, 1]$. Thus, for any α-CTIF algorithm, we focus on the algorithms where $v = L$ in order to minimize the competitive ratio. □

Theorem 4.6. *For any $\alpha \in [1/\ln(U/L)+1, 1]$, $\mathsf{ECT}[\alpha]$ is β-competitive and α-CTIF.*

Proof. Fix an arbitrary instance $\mathcal{I} \in \Omega$. When $\mathsf{ECT}[\alpha]$ terminates, suppose the utilization of the knapsack is z_T, and assume we obtain a value of $\mathsf{ECT}[\alpha](\mathcal{I})$. Let $\mathcal{P}$ and $\mathcal{P}^\star$ respectively be the sets of items picked by $\mathsf{ECT}[\alpha]$ and the optimal solution.

Denote the weight and the value of the common items (i.e., the items picked by both ECT and OPT) by $W = w(\mathcal{P} \cap \mathcal{P}^\star)$ and $V = v(\mathcal{P} \cap \mathcal{P}^\star)$. For each item j which is *not accepted* by $\mathsf{ECT}[\alpha]$, we know that its value density is $< \Psi^\alpha(z_j) \leq \Psi^\alpha(z_T)$ since Ψ^α is a non-decreasing function of z. Thus,

$$\mathsf{OPT}(\mathcal{I}) \leq V + \Psi^\alpha(z_T)(1 - W).$$

Since $\mathsf{ECT}[\alpha](\mathcal{I}) = V + v(\mathcal{P} \setminus \mathcal{P}^\star)$, the above inequality implies that

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{ECT}[\alpha](\mathcal{I})} \leq \frac{V + \Psi^\alpha(z_T)(1 - W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)}.$$

Note that, by definition of the algorithm, each item j picked in $\mathcal{P}$ must have value density at least $\Psi^\alpha(z_j)$, where z_j is the knapsack utilization when that item arrives. Thus, we have:

$$\begin{aligned} V &\geq \sum_{j \in \mathcal{P} \cap \mathcal{P}^\star} \Psi^\alpha(z_j)\, w_j =: V_1, \\ v(\mathcal{P} \setminus \mathcal{P}^\star) &\geq \sum_{j \in \mathcal{P} \setminus \mathcal{P}^\star} \Psi^\alpha(z_j)\, w_j =: V_2. \end{aligned}$$

Since $\mathsf{OPT}(\mathcal{I}) \geq \mathsf{ECT}[\alpha](\mathcal{I})$, we have:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{ECT}[\alpha](\mathcal{I})} \leq \frac{V + \Psi^\alpha(z_T)(1-W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)} \leq \frac{V_1 + \Psi^\alpha(z_T)(1-W)}{V_1 + v(\mathcal{P} \setminus \mathcal{P}^\star)} \leq \frac{V_1 + \Psi^\alpha(z_T)(1-W)}{V_1 + V_2},$$

where the second inequality follows because $\Psi^\alpha(z_T)(1-W) \geq v(\mathcal{P} \setminus \mathcal{P}^\star)$ and $V_1 \leq V$.

Note that $V_1 \leq \Psi^\alpha(z_T) w(\mathcal{P} \cap \mathcal{P}^\star) = \Psi^\alpha(z_T) W$, and by plugging in for the actual values of V_1 and V_2 we get:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{ECT}[\alpha](\mathcal{I})} \leq \frac{\Psi^\alpha(z_T)}{\sum_{j \in \mathcal{P}} \Psi^\alpha(z_j)\, w_j}.$$

Based on the assumption that individual item weights are much smaller than 1, we can substitute for w_j with $\delta z_j = z_{j+1} - z_j$ for all j. This substitution gives an approximate value of the summation via integration:

$$\sum_{j \in \mathcal{P}} \Psi^\alpha(z_j)\, w_j \approx \int_0^{z_T} \Psi^\alpha(z) dz$$

Now there are two cases to analyze – the case where $z_T = \alpha$, and the case where $z_T > \alpha$. Note that $z_T < \alpha$ is impossible, as this means $\mathsf{ECT}[\alpha]$ rejected some item that it had capacity for even when the threshold was at L, which is a contradiction. We explore each of the two cases in turn below.

Case I. If $z_T = \alpha$, then $\sum_{j \in \mathcal{P}} \Psi^\alpha(z_j)\, w_j \geq L\alpha$.
This follows because $\mathsf{ECT}[\alpha]$ is effectively greedy for at least α utilization of the knapsack, and so the admitted items during this portion must have value at least $L\alpha$. Substituting into the original equation gives us the following:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{ECT}[\alpha](\mathcal{I})} \leq \frac{\Psi^\alpha(z_T)}{L\alpha} \leq \beta.$$

Case II. If $z_T > \alpha$, then $\sum_{j \in \mathcal{P}} \Psi^\alpha(z_j)\, w_j \approx \int_0^\alpha L dz + \int_\alpha^{z_T} \Psi^\alpha(z) dz$.
Solving for the integration, we obtain the following:

$$\begin{aligned} \sum_{j \in \mathcal{P}} \Psi^\alpha(z_j)\, w_j &\approx \int_0^\alpha L dz + \int_\alpha^{z_T} \Psi^\alpha(z) dz \\ &= L\alpha + \int_\alpha^{z_T} U e^{\beta(z-1)} dz \;=\; L\alpha + \left[\frac{U e^{\beta(z-1)}}{\beta}\right]_\alpha^{z_T} \\ &= L\alpha - \frac{U e^{\beta(\alpha-1)}}{\beta} + \frac{U e^{\beta(z_T-1)}}{\beta} \;= L\alpha - \frac{\Psi^\alpha(\alpha)}{\beta} + \frac{\Psi^\alpha(z_T)}{\beta} = \frac{\Psi^\alpha(z_T)}{\beta}. \end{aligned}$$

Substituting into the original equation, we can bound the competitive ratio:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{ECT}[\alpha](\mathcal{I})} \leq \frac{\Psi^\alpha(z_T)}{\frac{1}{\beta}\Psi^\alpha(z_T)} = \beta,$$

and the result follows.

Furthermore, the value of β which solves the equation $x = \frac{U e^{x(\alpha-1)}}{L\alpha}$ can be shown as $\frac{W\left(\frac{U - U\alpha}{L\alpha}\right)}{1-\alpha}$, which matches the lower bound from Theorem 4.5. □

Theorem 4.11. *For any $\gamma \in (0, 1]$, and any $\mathcal{I} \in \Omega$,* $\mathsf{LA\text{-}ECT}[\gamma]$ *is $\frac{2}{\gamma}$-consistent.*

Proof. For consistency, assume that the black-box predictor $\rho_{\mathcal{I}}^{-1}(y)$ is accurate (i.e. $\hat{d}_\gamma = d_\gamma^\star$). Let ϵ denote the upper bound on any individual item's weight (previously assumed to be small).

In Lemma C.2, we describe $\mathsf{ORACLE}_\gamma^\star$, a competitive *semi-online algorithm* (Seiden et al., 2000; Tan and Wu, 2007; Kumar et al., 2019; Dwibedy and Mohanty, 2022) which is restricted to use a knapsack of size γ. Plainly, it is an algorithm that has full knowledge of the items in the instance, but must process items sequentially using a threshold it has to set in advance. Items still arrive in an online manner, decisions are immediate and irrevocable, and the order of arrival is unknown.

Lemma C.2. *There is a deterministic semi-online algorithm,* $\mathsf{ORACLE}_\gamma^\star$*, which is* 1-CTIF*, fills a knapsack of size* $\gamma \in [0, 1]$*, and has an approximation factor of* $2/(\gamma - \epsilon)$*. Moreover, no deterministic semi-online algorithm with an approximation factor less than* $2 - L/U$ *is* 1-CTIF.

Proof of Lemma C.2

Upper bound: Note that $\mathsf{ORACLE}_\gamma^\star$ can compute $d_\gamma^\star$ before any items arrive. Suppose $\mathsf{ORACLE}_\gamma^\star$ sets its threshold at $d_\gamma^\star$, and therefore admits any items with value density at or above $d_\gamma^\star$.

Recall the definition of the critical threshold $d_\gamma^\star$ from Definition 4.9. For an arbitrary instance $\mathcal{I}$, let V denote the value obtained by $\mathsf{OPT}(\mathcal{I})$, and $d_\gamma^\star$ gives the maximum value density such that the total value of items with value density $\geq d_\gamma^\star$ in $\mathsf{OPT}(\mathcal{I})$ is at least $\gamma V/2$.

Based on the definition of $d_\gamma^\star$, we know that either the total weight of items with value density $\geq d_\gamma^\star$ in $\mathsf{OPT}(\mathcal{I})$'s knapsack is strictly less than γ, or $\gamma d_\gamma^\star \geq \gamma V/2$. To verify this, we start by sorting $\mathsf{OPT}(\mathcal{I})$'s packed items in non-increasing order of value density.

Suppose a greedy approximation algorithm APX_γ iterates over this list in sorted order, packing items into a knapsack of size γ until it is full. Note that APX_γ packs $(\gamma - \epsilon)$ of the highest value density items from $\mathcal{I}$ into its knapsack.

By definition, $\mathsf{APX}_\gamma \geq (\gamma - \epsilon) \cdot V$. In the worst-case, where all items in $\mathsf{OPT}(\mathcal{I})$'s knapsack are the same value density, we have that $\mathsf{APX}_\gamma = (\gamma - \epsilon) \cdot V$.

Denote the value density of the last item packed by APX_γ as $\bar{d}$. For the sake of contradiction, assume that $d_\gamma^\star < \bar{d}$. Since APX_γ fills a $(\gamma - \epsilon)$ fraction of its knapsack with items of value density $\geq \bar{d}$ and obtains a value of at least $(\gamma - \epsilon)V > \frac{\gamma V}{2}$, this causes a contradiction: $d_\gamma^\star$ should be the *largest* value density such that the total value of items with value density $\geq d_\gamma^\star$ in $\mathsf{OPT}(\mathcal{I})$ is at least $\gamma V/2$, but the assumption $d_\gamma^\star < \bar{d}$ implies that the total value of items with value density $\geq \bar{d}$ is also at least $\gamma V/2$.

This further implies either of the following: (I) The true value of $d_\gamma^\star$ is $d_\gamma^\star > \bar{d}$ if APX_γ's knapsack contains enough items with value density $> \bar{d}$ and total weight $< (\gamma - \epsilon)$ such that their total value is at least $\gamma V/2$. (II) $d_\gamma^\star = \bar{d}$ if there are enough items of value density $\bar{d}$ in $\mathcal{I}$ such that $\gamma \bar{d} \geq \gamma V/2$.

Given this information, there are two possible outcomes for the items accepted by $\mathsf{ORACLE}_\gamma^\star$, listed below. In each, we show that the value obtained is at least $(\gamma - \epsilon)V/2$.

- If the total weight of items packed by $\mathsf{ORACLE}_\gamma^\star(\mathcal{I})$ is strictly less than $(\gamma - \epsilon)$, then we know the total weight of items with value density $\geq d_\gamma^\star$ in the optimal solution's knapsack is strictly less than $(\gamma - \epsilon)$, and that the total weight of items with value density $\geq d_\gamma^\star$ in the instance is also strictly less than $(\gamma - \epsilon)$. By definition of $d_\gamma^\star$, $\mathsf{ORACLE}_\gamma^\star(\mathcal{I})$ obtains a value of $\geq \gamma V/2$.
- If the total weight of items packed by $\mathsf{ORACLE}_\gamma^\star(\mathcal{I})$ is $\geq (\gamma - \epsilon)$, then we know that $\gamma d_\gamma^\star \geq \gamma V/2$. If this wasn't true, there would exist some other value density $\hat{d}$ in $\mathsf{OPT}(\mathcal{I})$'s knapsack such that $\hat{d} > d_\gamma^\star$ and the total value of items with value density $\geq \hat{d}$ in $\mathsf{OPT}(\mathcal{I})$'s knapsack would have value at least $\gamma V/2$. Thus, $\mathsf{ORACLE}_\gamma^\star(\mathcal{I})$ obtains a value of at least $(\gamma - \epsilon) d_\gamma^\star \geq (\gamma - \epsilon)V/2$.

Therefore, $\mathsf{ORACLE}_\gamma^\star$ admits a value of at least $(\gamma - \epsilon)V/2$, and its approximation factor is at most $2/(\gamma - \epsilon)$.

Lower bound: Consider an input formed by a large number of infinitesimal items of density L and total weight 1, followed by infinitesimal items of density U and total weight L/U. An optimal algorithm accepts all items of density U and fills the remaining space with items of density L, giving its knapsack a total value of $(L/U)U + (1 - L/U)L = 2L - L^2/U$. Any deterministic algorithm that satisfies 1-CTIF, however, must accept either all items of density L, giving its knapsack a value of L, or reject all items of density L, giving it a value of $(L/U)U = L$. In both cases, the approximation factor of the algorithm would be $\frac{2L - L^2/U}{L} = 2 - L/U$. □

We use $\mathsf{ORACLE}^\star_\gamma$ as a benchmark. Fix an arbitrary input $\mathcal{I} \in \Omega$. Let $\mathsf{LA\text{-}ECT}[\gamma]$ terminate filling z_T fraction of the knapsack and obtaining a value of $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$.

Let $\mathsf{ORACLE}^\star_\gamma$ terminate obtaining a value of $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$.

Now we consider two cases – the case where $z_T < \kappa + \gamma$, and the case where $z_T \geq \kappa + \gamma$. We explore each below.

Case I. If $z_T < \kappa + \gamma$, the following statements must be true:

- Since the threshold function $\Psi^{\hat{d}_\gamma}(z) \leq d^\star_\gamma$ for all values z less than $\kappa + \gamma$, any item accepted by $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$ must be accepted by $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$.
- Thus, $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I}) \geq \mathsf{ORACLE}^\star_\gamma(\mathcal{I})$, and $\frac{2}{\gamma - \epsilon}\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I}) \geq \mathsf{OPT}(\mathcal{I})$.

Note that **Case I** implies that as γ approaches 1, the value obtained by $\mathsf{LA\text{-}ECT}[1](\mathcal{I})$ is greater than or equal to that obtained by $\mathsf{ORACLE}^\star_1(\mathcal{I})$, and the competitive bound reduces to $\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[1](\mathcal{I})} \leq \frac{2}{1-\epsilon}$.

Case II. If $z_T \geq \kappa + \gamma$, then we know that any item accepted by $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$ must have been accepted by $\mathsf{LA\text{-}ECT}[\gamma]$.

Proof by contradiction: assume that $z_T \geq \kappa + \gamma$ and there exists some item accepted by $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$ that wasn't accepted by $\mathsf{LA\text{-}ECT}[\gamma]$. This implies that when the item arrived to $\mathsf{LA\text{-}ECT}[\gamma]$, the threshold function $\Psi^{\hat{d}_\gamma}(z)$ was greater than $d^\star_\gamma$, which is the minimum acceptable value density for any item accepted by $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$.

Since $\Psi^{\hat{d}_\gamma}(z) \leq \hat{d}_\gamma$ for all values $z \leq \kappa + \gamma$ and $z_T \geq \kappa + \gamma$ implies that $\mathsf{LA\text{-}ECT}[\gamma]$ saw *enough* items with value density $\geq d^\star_\gamma$ to fill a γ fraction of the knapsack, this causes a contradiction. Since items arrive in the same order to both $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$ and $\mathsf{LA\text{-}ECT}[\gamma]$, $\mathsf{ORACLE}^\star_\gamma(\mathcal{I})$'s knapsack would already be full by the time this item arrived.

This tells us that $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I}) \geq \mathsf{ORACLE}^\star_\gamma(\mathcal{I})$, and thus we have the following:

$$\frac{2}{\gamma - \epsilon}\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I}) \geq \mathsf{OPT}(\mathcal{I})$$

It follows in either case that $\mathsf{LA\text{-}ECT}[\gamma]$ is $\frac{2}{\gamma}$-consistent for accurate predictions.

□

Theorem 4.12. *For any $\gamma \in [0, 1]$, and any $\mathcal{I} \in \Omega$, $\mathsf{LA\text{-}ECT}[\gamma]$ is $\frac{1}{1-\gamma}(\ln(U/L) + 1)$-robust.*

Proof. Fix an arbitrary input sequence $\mathcal{I}$. Let $\mathsf{LA\text{-}ECT}[\gamma]$ terminate filling z_T fraction of the knapsack and obtaining a value of $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$. Let $\mathcal{P}$ and $\mathcal{P}^\star$ respectively be the sets of items picked by $\mathsf{LA\text{-}ECT}[\gamma]$ and the optimal solution.

Denote the weight and the value of the common items (items picked by both $\mathsf{LA\text{-}ECT}$ and OPT) by $W = w(\mathcal{P} \cap \mathcal{P}^\star)$ and $V = v(\mathcal{P} \cap \mathcal{P}^\star)$. For each item j which is *not* accepted by $\mathsf{LA\text{-}ECT}[\gamma]$, we know that its value density is $< \Psi^{\gamma,\hat{d}}(z_j) \leq \Psi^{\gamma,\hat{d}}(z_T)$ since $\Psi^{\gamma,\hat{d}}$ is a non-decreasing function of z. Thus, we have:

$$\mathsf{OPT}(\mathcal{I}) \leq V + \Psi^{\gamma,\hat{d}}(z_T)(1 - W).$$

Since $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I}) = V + v(\mathcal{P} \setminus \mathcal{P}^\star)$, the inequality above implies that:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{V + \Psi^{\gamma,\hat{d}}(z_T)(1-W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)}.$$

Note that each item j picked in $\mathcal{P}$ must have value density of at least $\Psi^{\gamma,\hat{d}}(z_j)$, where z_j is the knapsack utilization when that item arrives. Thus, we have that:

$$\begin{aligned} V &\geq \sum_{j \in \mathcal{P} \cap \mathcal{P}^\star} \Psi^{\gamma,\hat{d}}(z_j)\ w_j =: V_1, \\ v(\mathcal{P} \setminus \mathcal{P}^\star) &\geq \sum_{j \in \mathcal{P} \setminus \mathcal{P}^\star} \Psi^{\gamma,\hat{d}}(z_j)\ w_j =: V_2. \end{aligned}$$

Since $\mathsf{OPT}(\mathcal{I}) \geq \mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$, we have that

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{V + \Psi^{\gamma,\hat{d}}(z_T)(1-W)}{V + v(\mathcal{P} \setminus \mathcal{P}^\star)} \leq \frac{V_1 + \Psi^{\gamma,\hat{d}}(z_T)(1-W)}{V_1 + V_2}.$$

Note that $V_1 \leq \Psi^{\gamma,\hat{d}}(z_T) w(\mathcal{P} \cap \mathcal{P}^\star) = \Psi^{\gamma,\hat{d}}(z_T) W$, and so, plugging in the actual values of V_1 and V_2, we get:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{\Psi^{\gamma,\hat{d}}(z_T)}{\sum_{j \in \mathcal{P}} \Psi^{\gamma,\hat{d}}(z_j)\ w_j}.$$

Based on the assumption that individual item weights are much smaller than 1, we can substitute for w_j with $\delta z_j = z_{j+1} - z_j$ for all j. This substitution allows us to obtain an approximate value of the summation via integration:

$$\sum_{j \in \mathcal{P}} \Psi^{\gamma,\hat{d}}(z_j)\ w_j \approx \int_0^{z_T} \Psi^{\gamma,\hat{d}}(z) dz$$

Now we consider three separate cases – the case where $z_T \in [0, \kappa)$, the case where $z_T \in [\kappa, \kappa + \gamma)$, and the case where $z_T \in [\kappa + \gamma, 1]$. We explore each below.

Case I. If $z_T \in [0, \kappa)$, $\mathsf{OPT}(\mathcal{I})$ is bounded by $\Psi^{\gamma,\hat{d}}(z_T) \leq \hat{d}$. Furthermore,

$$\begin{aligned} \sum_{j \in \mathcal{P}} \Psi^{\gamma,\hat{d}}(z_j)\ w_j \approx \int_0^{z_T} \Psi^{\gamma,\hat{d}}(z) dz\ &= \int_0^{z_T} \left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}} \left(\frac{L}{e}\right) dz \\ &= (1-\gamma) \left(\frac{L}{e}\right) \left[\frac{\left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}}}{\ln(Ue/L)}\right]_0^{z_T} = (1-\gamma) \frac{\Psi^{\gamma,\hat{d}}(z_T)}{\ln(Ue/L)} \end{aligned}$$

Combined with the previous equation for the competitive ratio, this gives us the following:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{\Psi^{\gamma,\hat{d}}(z_T)}{(1-\gamma) \frac{\Psi^{\gamma,\hat{d}}(z_T)}{\ln(U/L)+1}} \leq \frac{\hat{d}}{(1-\gamma) \frac{\hat{d}}{\ln(U/L)+1}} = \frac{1}{1-\gamma} \left(\ln(U/L) + 1\right).$$

Case II. If $z_T \in [\kappa, \kappa + \gamma)$, $\mathsf{OPT}(\mathcal{I})$ is bounded by $\Psi^{\gamma,\hat{d}}(z_T) = \hat{d}$. Furthermore,

$$\int_0^{z_T} \Psi^{\gamma,\hat{d}}(z) dz = \int_0^{\kappa} \left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}} \left(\frac{L}{e}\right) dz + \int_\kappa^{z_T} \hat{d}\ dz \geq \int_0^{\kappa} \left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}} \left(\frac{L}{e}\right) dz.$$

Note that since the bound on $\mathsf{OPT}(\mathcal{I})$ can be the same (i.e. $\mathsf{OPT}(\mathcal{I}) \leq \hat{d}$), Case I is strictly worse than Case II for the competitive ratio, and we inherit the worse bound:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{\hat{d}}{(1-\gamma) \frac{\hat{d}}{\ln(U/L)+1}} = \frac{1}{1-\gamma} \left(\ln(U/L) + 1\right).$$

Case III. If $z_T \in [\kappa + \gamma, 1]$, $\mathsf{OPT}(\mathcal{I})$ is bounded by $\Psi^{\gamma,\hat{d}}(z_T) \leq U$. Furthermore,

$$\sum_{j \in \mathcal{P}} \Psi^{\gamma,\hat{d}}(z_j)\, w_j \approx \int_0^{z_T} \Psi^{\gamma,\hat{d}}(z) dz \; = \int_0^{z_T - \gamma} \left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}} \left(\frac{L}{e}\right) dz + \int_\kappa^{\kappa+\gamma} \hat{d}\; dz$$

$$= (1-\gamma) \left(\frac{L}{e}\right) \left[\frac{\left(\frac{Ue}{L}\right)^{\frac{z}{1-\gamma}}}{\ln(Ue/L)} \right]_0^{z_T - \gamma} + \gamma \hat{d}$$

$$= (1-\gamma) \frac{\Psi^{\gamma,\hat{d}}(z_T)}{\ln(Ue/L)} + \gamma \hat{d}.$$

Combined with the previous equation for the competitive ratio, this gives us the following:

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})} \leq \frac{\Psi^{\gamma,\hat{d}}(z_T)}{(1-\gamma)\frac{\Psi^{\gamma,\hat{d}}(z_T)}{\ln(U/L)+1} + \gamma\hat{d}} \leq \frac{\Psi^{\gamma,\hat{d}}(z_T)}{(1-\gamma)\frac{\Psi^{\gamma,\hat{d}}(z_T)}{\ln(U/L)+1}} = \frac{1}{1-\gamma}\left(\ln(U/L)+1\right).$$

Since we have shown that $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$ obtains at least $\frac{1}{1-\gamma}\left(\ln(U/L)+1\right)$ of the value obtained by $\mathsf{OPT}(\mathcal{I})$ in each case, we conclude that $\mathsf{LA\text{-}ECT}[\gamma](\mathcal{I})$ is $\frac{1}{1-\gamma}\left(\ln(U/L)+1\right)$-robust. □

Theorem 4.13. *For any learning-augmented online algorithm* ALG *which satisfies* 1-CTIF*, one of the following holds: Either* ALG*'s consistency is* $> 2\sqrt{U/L} - 1$*, or* ALG *has unbounded robustness. Furthermore, the consistency of any algorithm is lower bounded by* $2-\varepsilon^2/1+\varepsilon$*, where* $\varepsilon = \sqrt{L/U}$.

Proof. We begin by proving the first statement, which gives a *consistency-robustness* trade off for any learning-augmented ALG.

Lemma C.3. *One of the following statements holds for any* 1-CTIF *online algorithm* ALG *with any prediction:*

(i) ALG *has consistency worse (larger) than* $2\sqrt{U/L} - 1$.

(ii) ALG *has unbounded robustness.*

Proof of Lemma C.3. Let $\varepsilon = \sqrt{L/U}$ and $V = \sqrt{LU}$ and note that $\varepsilon V = L$ and $V/\varepsilon = U$.

Consider a sequence $\mathcal{I}$ that starts with a set of *red items* of density L and total size 1, continues with $1/\varepsilon$ "white" items, each of size ε and density $\varepsilon(1+\varepsilon)V$ (which is in $[L, U]$), and ends with one *black* item of size ε and density $V/\varepsilon = U$. The optimal solution rejects all red items and accepts all other items except one white item.
The optimal profit is thus $\mathsf{OPT}(\mathcal{I}) = (1+\varepsilon)V - \varepsilon(1+\varepsilon)V + V = (2-\varepsilon^2)V$.

Suppose the predictions are consistent with $\mathcal{I}$. Then, a 1-CTIF learning-augmented ALG has the following two choices:

- It accepts all red items. Then if the input is indeed $\mathcal{I}$, the consistency of ALG would be $\frac{(2-\varepsilon^2)\sqrt{LU}}{L} = (2-\varepsilon^2)\sqrt{U/L} = 2\sqrt{U/L} - \sqrt{L/U} > 2\sqrt{U/L} - 1$. In this case, **(i)** holds.
- It rejects all red items. Then, the input may be formed entirely by the red items (and the predictions are incorrect). The algorithm does not accept any item, and its robustness will be unbounded. In this case, **(ii)** holds. □

Note that $\mathsf{LA\text{-}ECT}[\gamma]$ satisfies **(ii)** when $\gamma \to 1$.

Next, we prove the final statement, which lower bounds the achievable consistency for any 1-CTIF algorithm. To do this, we consider a semi-online 1-CTIF algorithm ALG. It has full knowledge of the items in the instance, but must process items sequentially using a threshold it has to set in advance. Items still arrive in an online manner, decisions are immediate and irrevocable, and the order of arrival is unknown.

Lemma C.4. *Any semi-online* 1-CTIF *algorithm has an approximation factor of at least* $\frac{2-\varepsilon^2}{1+\varepsilon}$*, where* $\varepsilon = \sqrt{L/U}$.

Proof of Lemma C.4. As previously, let $V = \sqrt{LU}$ and note that $\varepsilon V = L$ and $V/\varepsilon = U$.

Consider an input sequence starting with $1/\varepsilon$ “white" items, each of size ε and density $\varepsilon(1+\varepsilon)V$ (which is in $[L, U]$). Note that white items have a total size of 1 (knapsack capacity) and a total value of $(1+\varepsilon)V$. Suppose the input continues with one *black* item of size ε and density $V/\varepsilon = U$. An optimal algorithm accepts all items in the input sequence except one white item. The optimal profit is thus $(1+\varepsilon)V - \varepsilon(1+\varepsilon)V + V = (2-\varepsilon^2)V$.

Given that the entire set of white items fits in the knapsack, any 1-CTIF algorithm must accept or reject them all. In the former case, the algorithm cannot accept the black item (the knapsack becomes full before processing the black item), and its profit will be $(1+\varepsilon)V$, resulting in an approximation factor of $\frac{2-\varepsilon^2}{1+\varepsilon}$. In the latter case, the algorithm can only accept the black item, and its approximation factor would be at least $2-\varepsilon^2$.

Combining the statements of Lemmas C.3 and C.4, the original statement follows. □