## EC.1. Electronic companion: additional models and deferred details

The main body develops the two core adversarial models in Part I—online fractional matching and edge-weighted matching with free disposal—and keeps the technical discussion focused on the reusable design-and-analysis template introduced there. This electronic companion records the broader development behind that template. In particular, it collects the richer deterministic extensions that would otherwise interrupt the flow of the main exposition, together with the deferred proofs whose full details are useful for reference.

Throughout the electronic companion, the notation table in Table 2 remains in force. Any additional symbol is defined locally at first use. The guiding philosophy is also unchanged: identify the right state, regularize the post-decision state rather than the current reward, extract the dual update from the KKT system, and then choose the regularizer so that approximate dual feasibility collapses to a one-dimensional identity or recursion.

More specifically, the electronic companion contains the following pieces.

1. **Batch arrival.** We study the $K$-stage arrival model, introduce the stage-dependent polynomial regularizers that yield the optimal factor $\Gamma(K) = 1 - (1 - 1/K)^K$, and record the recursive dual-fitting argument that replaces the scalar identity from the fully online setting (Section EC.2 and Section EC.5).
2. **Configuration allocation / whole-page optimization.** We move from scalar or reward-indexed states to a price-level state vector and show how direct convex duality replaces the support-graph decomposition that was available in matching (Section EC.3 and Section EC.6).
3. **Other extensions.** We explain how the same template specializes to AdWords, costly cancellations, and several nearby allocation models. The emphasis is on reusability of the main ideas rather than on developing a separate proof for every variant (Section EC.4).

The companion is therefore meant to be read in two ways. A first pass can follow only the model statements and the high-level proof roadmaps. A second pass can then use the supplementary sections as a self-contained reference for the detailed primal–dual calculations.

## EC.2. Batch arrival: multi-stage fractional matching

We next record the $K$-stage arrival model studied in Feng and Niazadeh [15]. Relative to the fully online setting of Section 3.2, the main new issue is temporal heterogeneity: a fixed regularizer is no longer appropriate across the entire horizon. Early stages should hedge more aggressively because much of the future is still unknown, whereas late stages should behave more like the residual greedy problem. The resulting analysis is still primal–dual, but the scalar identity from the one-arrival model is replaced by a recursion across stages.

### EC.2.1. How the universal recipe changes

The five-step recipe from Section 3.1 survives almost verbatim, with three modifications.

1. The state remains the offline load vector, but it is updated stage by stage rather than arrival by arrival.
2. The regularizer is stage dependent: the stage-$k$ program uses $F_k$ rather than a single function $F$.
3. The KKT conditions are now written for an entire batch. As a result, the relevant structure is no longer the one-arrival water-filling picture, but a decomposition of the positive-support graph of the batch.

The conceptual message is unchanged: the regularizer determines the online rule, and the KKT system determines the dual update.

### EC.2.2. Stage-dependent polynomial regularizers

The appropriate stage-wise regularizers are determined by a backward recursion.

PROPOSITION 3 **(Optimal stage-dependent regularizers)**. *Let $K \in \mathbb{N}$. Define functions $f_1, \dots, f_K$ on $[0,1]$ by*

$$f_K(x) = \mathbf{1}\{x = 1\}, \qquad f_k(x) = \max_{y \in [0,1-x]} (1-y) f_{k+1}(x+y) \qquad \forall k \in [K-1].$$

*Then, for every $k \in [K-1]$,*

$$f_k(x) = \left(1 - \frac{1-x}{K-k}\right)^{K-k}.$$

*Moreover:*

*(i) each $f_k$ is increasing on $[0,1]$, and therefore $F_k(x) := \int_0^x f_k(u)\,\mathrm{d}u$ is differentiable and strictly convex;*

*(ii) $f_k(x) \le \mathrm{e}^{x-1}$ for every $x \in [0,1]$;*

*(iii) $f_1$ converges pointwise to $\mathrm{e}^{x-1}$ as $K \to \infty$;*

*(iv)*

$$\Gamma(K) = 1 - \left(1 - \frac{1}{K}\right)^K = \min_{y \in [0,1]} \left[1 - (1-y) f_1(y)\right].$$

The proof is given in Section EC.5. The degree of $f_k$ is $K-k$, so the regularization intensity decreases with the stage index. In particular, when $K$ is large and the batches are correspondingly small, the polynomial family approaches the exponential rule from the fully online model.

### EC.2.3. The stage-wise algorithm

Let $y_j^{(k)}$ denote the cumulative load on offline node $j$ after stage $k$, with $y_j^{(0)} = 0$. After observing the batch $U_k$ and its edge set $E_k$, the algorithm solves

$$\begin{aligned} \max_{x_{tj}^{(k)} \ge 0} \quad & \sum_{(t,j) \in E_k} w_j x_{tj}^{(k)} - \sum_{j \in V} w_j F_k\left(y_j^{(k-1)} + \sum_{t:(t,j) \in E_k} x_{tj}^{(k)}\right) \\ \text{s.t.} \quad & \sum_{j:(t,j) \in E_k} x_{tj}^{(k)} \le 1 \qquad \forall t \in U_k, \\ & \sum_{t:(t,j) \in E_k} x_{tj}^{(k)} \le 1 - y_j^{(k-1)} \qquad \forall j \in V. \end{aligned} \tag{EC.1}$$

and then updates the state by

$$y_j^{(k)} = y_j^{(k-1)} + \sum_{t:(t,j)\in E_k} x_{tj}^{(k)}.$$

Here $x_{tj}^{(k)}$ is the fraction of batch vertex $t \in U_k$ assigned to offline node $j$ during stage $k$.

The unweighted case is a useful special case. There, the optimizer of (EC.1) is simply batched water-filling: within a batch, mass is pushed toward the least loaded feasible offline vertices until one either exhausts the batch supply or saturates a subset of offline vertices. The polynomial family is therefore not needed to describe the primal rule in the unweighted model; it is needed to analyze that rule and to obtain the sharp factor $\Gamma(K)$.

### EC.2.4. What the KKT conditions reveal

Fix a stage $k \le K-1$, and let $x^{(k)} = \{x_{tj}^{(k)}\}$ denote the optimizer of (EC.1). Consider the positive-support graph

$$G'_k := \Big(U_k \cup V,\ \{(t,j) \in E_k : x_{tj}^{(k)} > 0\}\Big).$$

Let

$$\mathcal{V}_{k,0} := \{j \in V : y_j^{(k)} = 1\},$$

and let $\mathcal{U}_{k,0}$ be the set of their neighbors in $G'_k$. Remove those vertices, and denote the connected components of the remaining support graph by $(\mathcal{U}_{k,\ell}, \mathcal{V}_{k,\ell})$, $\ell = 1, \dots, L_k$.

PROPOSITION 4 **(KKT decomposition of a batch).** *For every stage $k \le K-1$, there exist numbers $0 = c_{k,0} < c_{k,1} < \cdots < c_{k,L_k}$ such that:*

*(i)* ***Uniformity.*** *For every component $\ell \in \{0,1,\dots,L_k\}$ and every $j, j' \in \mathcal{V}_{k,\ell}$,*

$$w_j\big(1 - f_k(y_j^{(k)})\big) = w_{j'}\big(1 - f_k(y_{j'}^{(k)})\big) = c_{k,\ell}.$$

*(ii)* ***Saturation.*** *Every online node in $\mathcal{U}_{k,\ell}$, $\ell \ge 1$, is fully matched in stage $k$.*

*(iii)* ***Monotonicity.*** *If $c_{k,\ell} < c_{k,\ell'}$, then no edge of $E_k$ joins $\mathcal{U}_{k,\ell}$ to $\mathcal{V}_{k,\ell'}$.*

The proof is given in Section EC.5. This is the stage-wise analogue of the KKT picture from Section 3.2. In the one-arrival model the KKT conditions describe how a single request splits its mass. In the batched model they describe how the support graph decomposes into components with a common marginal "matchability level" $c_{k,\ell}$.

### EC.2.5. Dual construction and the scalar recursion

The offline dual benchmark is still (7). Guided by Proposition 4, we define the fitted dual solution stage by stage. For stages $k \le K-1$, set

$$\widehat{\alpha}_t := c_{k,\ell} \qquad \text{for every } t \in \mathcal{U}_{k,\ell},$$

and increment $\widehat{\beta}_j$ by

$$\Delta\widehat{\beta}_j^{(k)} := w_j\big(y_j^{(k)} - y_j^{(k-1)}\big) f_k\big(y_j^{(k)}\big) \qquad \forall j \in V. \tag{EC.2}$$

At the last stage, $F_K \equiv 0$, so the stage problem reduces to the residual maximum-weight matching LP. We therefore use an optimal dual solution of that residual LP for the final dual increment.

Approximate dual feasibility is governed by the scalar recursion

$$1 - (1-x) f_k(u_k + x) + \sum_{s=1}^{k-1} (u_{s+1} - u_s) f_s(u_{s+1}) \geq \Gamma(K) \tag{EC.3}$$

for every nondecreasing sequence $0 = u_1 \leq \cdots \leq u_k \leq 1 - x$. The proof is elementary but recursive, so we record it in Section EC.5. Equation (EC.3) is the stage-wise counterpart of the identity $1 - f(y) + F(y) = \Gamma$ from the fully online model.

THEOREM 8 **(Optimal competitive ratio under $K$-stage arrival).** *For $K$-stage fractional vertex-weighted matching, the stage-wise algorithm defined by* (EC.1) *is $\Gamma(K)$-competitive, where*

$$\Gamma(K) = 1 - \left(1 - \frac{1}{K}\right)^K.$$

*Proof.* The theorem is proved in full in Section EC.5. The argument follows exactly the same structure as in the main body. First, the KKT decomposition in Proposition 4 yields an online dual update for which the stage-wise dual increment matches the stage-wise primal reward exactly. Second, the scalar recursion (EC.3) implies approximate dual feasibility on every edge. Weak duality then gives the competitive ratio $\Gamma(K)$. □

REMARK 8. For $K = 2$, the theorem gives the sharp ratio $\Gamma(2) = 3/4$, already strictly larger than the fully online benchmark $1 - 1/\mathrm{e}$. As $K \to \infty$, $\Gamma(K) \downarrow 1 - 1/\mathrm{e}$, so the stage-wise model converges back to the online limit. The upper bound proved in Feng and Niazadeh [15] shows that this ratio is tight.

## EC.3. Configuration allocation

Configuration allocation is the most general deterministic model treated in Part I. It contains vertex-weighted matching, AdWords, edge-weighted matching with free disposal, and the whole-page optimization model of Devanur et al. [10] as special cases. The same convex-programming viewpoint still applies, but the state is now multi-dimensional and the clean support-graph decomposition from matching is no longer available. The natural replacement is direct convex duality.

### EC.3.1. Why this model is harder

Three new issues appear at once.

1. **Multiple price levels.** In matching, all mass assigned to a fixed offline node has the same value. In configuration allocation, advertiser $j$ may value different user–configuration pairs at very different prices.

2. **No useful graph decomposition.** In batch matching, the KKT system induced a graph decomposition of the positive-support edges. Here an advertiser may be tight at some price thresholds and slack at others, so a graph-only description no longer captures the relevant structure.
3. **Allocation and preemption are coupled.** At each stage, the algorithm must decide not only what new mass to keep, but also which previously retained mass to discard.

These are precisely the reasons the analysis moves from combinatorial structure to direct convex duality.

### EC.3.2. State, stage program, and interpretation of the variables

Assume there is a finite price universe

$$0 = \widehat{w}_0 < \widehat{w}_1 < \cdots < \widehat{w}_T.$$

For each advertiser $j$ and price index $\tau$, let $\eta_{j,\tau}^{(k-1)}$ denote the amount of mass retained by advertiser $j$ at price level $\widehat{w}_\tau$ before stage $k$.

Once batch $U_k$ arrives, the stage-$k$ convex program chooses three families of variables.

| Variable | Meaning |
|---|---|
| $z_{t,c}$ | fraction of user $t \in U_k$ assigned to configuration $c$ |
| $x_{t,c,j}$ | amount of new mass from $(t,c)$ that advertiser $j$ still pays for *after* stage $k$ |
| $y_{j,\tau}$ | amount of previously retained mass of advertiser $j$ at price level $\widehat{w}_\tau$ that survives stage $k$ |

The input data are the prices $w_{t,c,j}$, the consumptions $\xi_{t,c,j}$, and the pre-stage state $\eta_{j,\tau}^{(k-1)}$.

For any candidate solution $(x,y,z)$, define the cumulative post-stage mass of advertiser $j$ above threshold $\widehat{w}_\tau$ by

$$H_{j,\tau}(x,y) := \sum_{t\in U_k,\, c\in\mathcal{C}:\ w_{t,c,j}\geq \widehat{w}_\tau} x_{t,c,j} + \sum_{\tau'\geq\tau} y_{j,\tau'}. \tag{EC.4}$$

This is the price-level analogue of the cumulative state $Y_j(r)$ from Section 3.3.

The stage-wise regularized convex program is

$$\begin{aligned}\max_{x,\,y,\,z\geq 0}\quad & \sum_{t\in U_k}\sum_{c\in\mathcal{C}}\sum_{j\in V} w_{t,c,j}x_{t,c,j} \\ & -\sum_{j\in V}\sum_{\tau\in[T]} \widehat{w}_\tau\big(\eta_{j,\tau}^{(k-1)} - y_{j,\tau}\big) \\ & -\sum_{j\in V}\sum_{\tau\in[T]} (\widehat{w}_\tau - \widehat{w}_{\tau-1})F_k\big(H_{j,\tau}(x,y)\big) \qquad \text{(EC.5a)}\end{aligned}$$

subject to

$$\sum_{t\in U_k}\sum_{c\in\mathcal{C}} x_{t,c,j} + \sum_{\tau\in[T]} y_{j,\tau} \leq 1 \qquad \forall j\in V, \tag{EC.5b}$$

$$\sum_{c\in\mathcal{C}} z_{t,c} \leq 1 \qquad \forall t\in U_k, \tag{EC.5c}$$

$$x_{t,c,j} \leq \xi_{t,c,j} z_{t,c} \qquad \forall t\in U_k,\ \forall c\in\mathcal{C},\ \forall j\in V, \tag{EC.5d}$$

$$y_{j,\tau} \le \eta_{j,\tau}^{(k-1)} \qquad \forall j \in V,\ \forall \tau \in [T]. \tag{EC.5e}$$

The three lines of the objective have a direct interpretation: the first is the retained revenue from the new batch, the second subtracts the value of previously retained mass that is discarded in stage $k$, and the third regularizes the post-stage cumulative mass of each advertiser at every threshold. The state update is

$$\eta_{j,\tau}^{(k)} := \sum_{t \in U_k,\, c \in \mathcal{C}:\ w_{t,c,j} = \widehat{w}_\tau} x_{t,c,j} + y_{j,\tau},$$

so $\eta^{(k)}$ is exactly the retained price distribution after stage $k$.

### EC.3.3. Offline benchmark

The standard offline LP benchmark is

$$\begin{aligned}
\max \quad & \sum_{t \in U} \sum_{c \in \mathcal{C}} \sum_{j \in V} w_{t,c,j} x_{t,c,j} \\
\text{s.t.} \quad & \sum_{t \in U} \sum_{c \in \mathcal{C}} x_{t,c,j} \le 1 && \forall j \in V, \\
& \sum_{c \in \mathcal{C}} z_{t,c} \le 1 && \forall t \in U, \\
& x_{t,c,j} \le \xi_{t,c,j} z_{t,c} && \forall t \in U,\ \forall c \in \mathcal{C},\ \forall j \in V, \\
& x_{t,c,j}, z_{t,c} \ge 0.
\end{aligned} \tag{EC.6}$$

Its dual is

$$\begin{aligned}
\min \quad & \sum_{t \in U} \alpha_t + \sum_{j \in V} \beta_j \\
\text{s.t.} \quad & \gamma_{t,c,j} + \beta_j \ge w_{t,c,j} && \forall t, c, j, \\
& \alpha_t \ge \sum_{j \in V} \xi_{t,c,j} \gamma_{t,c,j} && \forall t, c, \\
& \alpha_t, \beta_j, \gamma_{t,c,j} \ge 0.
\end{aligned} \tag{EC.7}$$

### EC.3.4. How the universal recipe specializes

In this model, the five steps from Section 3.1 become:

1. **State:** the retained price-level vector $\eta^{(k-1)}$.
2. **Regularized stage problem:** solve (EC.5a)–(EC.5e).
3. **KKT signal:** read off the multipliers associated with the configuration, linking, and carry constraints.
4. **Dual construction:** use the Lagrangian dual of the stage program directly rather than a graph decomposition.
5. **Scalar inequality:** apply the same one-dimensional recursion from batch arrival, now threshold by threshold.

The next two subsections make these points precise (with proofs and details in Section EC.6).

### EC.3.5. Structural lemmas from convex duality

The first lemma records the KKT identities that later drive the offline-dual construction.

LEMMA 10 **(Stage-wise KKT relations).** *Fix a stage $k \le K-1$, and let $(x^*, y^*, z^*)$ be the optimizer of (EC.5a)–(EC.5e). There exist nonnegative multipliers $\lambda_t^{(k)}$, $\mu_{t,c,j}^{(k)}$, $\psi_{t,c,j}^{(k)}$, and $\phi_{t,c}^{(k)}$ such that:*

*(i) complementary slackness holds for the configuration-choice and linking constraints;*

*(ii) for every user–configuration pair $(t,c)$,*

$$\sum_{j\in V} \xi_{t,c,j}\mu_{t,c,j}^{(k)} - \lambda_t^{(k)} + \phi_{t,c}^{(k)} = 0;$$

*(iii) for every triple $(t,c,j)$,*

$$w_{t,c,j} - \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} (\widehat{w}_\tau - \widehat{w}_{\tau-1}) f_k\big(H_{j,\tau}(x^*,y^*)\big) - \mu_{t,c,j}^{(k)} + \psi_{t,c,j}^{(k)} = 0.$$

The next lemma identifies the preemption rule induced by optimality.

LEMMA 11 **(Optimal preemption discards the lowest retained prices first).** *Fix $k \le K-1$, and let $(x^*, y^*, z^*)$ be the stage optimizer. If*

$$\eta_{j,\tau}^{(k-1)} - y_{j,\tau}^* > 0,$$

*then*

$$H_{j,\tau}(x^*,y^*) = 1.$$

*Equivalently, whenever some previously retained mass at price level $\widehat{w}_\tau$ is discarded, the retained post-stage state is already full at that threshold.*

COROLLARY 1 **(Monotonicity of the price-level state).** *For every advertiser $j$, threshold $\tau$, and stage $k \le K-1$, we have:*

*(i)* ***Distribution monotonicity:***

$$H_{j,\tau}(x^*,y^*) \ge \sum_{\tau'\ge\tau} \eta_{j,\tau'}^{(k-1)};$$

*(ii)* ***Only the smallest retained prices are discarded:***

$$\sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^*\big) = \sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^*\big) f_k\big(H_{j,\tau}(x^*,y^*)\big).$$

The detailed proofs are deferred to Section EC.6. Conceptually, these statements play the same role as the graph decomposition in Section EC.2: they explain how the state evolves and therefore how the advertiser-side dual mass should be charged.

### EC.3.6. Dual construction and competitive ratio

Recall the offline dual benchmark (EC.7). Guided by Lemma 10, we construct a fitted dual solution stage by stage. For stages $k \le K-1$, set

$$\widehat{\alpha}_t := \lambda_t^{(k)}, \qquad \widehat{\gamma}_{t,c,j} := \mu_{t,c,j}^{(k)},$$

and define the advertiser-side increment by

$$\begin{aligned}\Delta\widehat{\beta}_j^{(k)} := \sum_{t\in U_k}\sum_{c\in\mathcal{C}} x_{t,c,j}^{(k)} \sum_{\tau:\,\widehat{w}_\tau \le w_{t,c,j}} (\widehat{w}_\tau - \widehat{w}_{\tau-1}) f_k\big(H_{j,\tau}^{(k)}\big) \\ - \sum_{\tau\in[T]} \widehat{w}_\tau\big(\eta_{j,\tau}^{(k-1)} - y_{j,\tau}^{(k)}\big).\end{aligned} \tag{EC.8}$$

At the final stage, $F_K \equiv 0$, so the stage problem is linear and we use an optimal dual solution of the residual LP to define the last advertiser-side increment.

The important point is that approximate dual feasibility reduces to exactly the same scalar recursion as in Section EC.2, but now one threshold at a time.

THEOREM 9 (**Competitive ratio for $K$-stage configuration allocation**). *Let $f_1,\dots,f_K$ be the polynomial family from Proposition 3. Then the stage-wise algorithm defined by* (EC.5a)–(EC.5e) *is $\Gamma(K)$-competitive for $K$-stage fractional configuration allocation.*

*Proof.* The full proof is given in Section EC.6. The argument follows the same two-step pattern as before. First, the KKT multipliers in Lemma 10 yield a dual update for which the stage-wise increment of the fitted dual objective matches the stage-wise primal reward exactly. Second, Lemma 11 and Corollary 1 reduce approximate dual feasibility to the scalar recursion from Section EC.2, applied separately at each threshold. Weak duality for (EC.7) then gives the factor $\Gamma(K)$. □

REMARK 9. This theorem subsumes the preceding adversarial models in a clean way. The reward-indexed state from Section 3.3 reappears here after discretizing the reward axis into price levels, and the stage-dependent regularizers from Section EC.2 reappear unchanged. What disappears is the helpful support-graph decomposition; direct convex duality replaces it.

## EC.4. Short extensions and further directions

This section collects nearby models that fit the same convex-programming viewpoint with only modest changes in state and scaling. The purpose is not to reproduce a full proof for each variant. Rather, the point is that once the state variable, the regularized one-step objective, and the KKT-based dual update are written down correctly, the earlier arguments can usually be reused with little additional work.

### EC.4.1. AdWords / online budgeted allocation

The AdWords problem is the canonical budget-allocation analogue of online matching. Advertiser $j$ has budget $B_j$ and submits bid $b_{tj}$ for query $t$. The platform may assign at most one advertiser to each arriving query, and if $(t,j)$ is selected then advertiser $j$ pays $b_{tj}$ and the remaining budget decreases accordingly.

The right state variable is the normalized budget consumption

$$y_j := \frac{1}{B_j}\sum_{s\le t} b_{sj}x_{sj} \in [0,1].$$

With this normalization, the arrival-wise regularized program is

$$\max_{x_{tj}\ge 0} \sum_j b_{tj}x_{tj} - \sum_j B_j F\left(y_j + \frac{b_{tj}}{B_j}x_{tj}\right) \quad \text{s.t. } \sum_j x_{tj} \le 1,\ \ y_j + \frac{b_{tj}}{B_j}x_{tj} \le 1\ \ \forall j. \tag{EC.9}$$

Geometrically, this is the same object as vertex-weighted matching: the unit-capacity state $y_j$ is replaced by normalized budget load, and the load increment induced by query $t$ is $b_{tj}/B_j$ rather than $x_{tj}$. This is exactly the scaling that appears in the inventory-balancing view of AdWords [6, 32]. The fitted dual construction is still arrival-by-arrival: after solving the convex program, one fixes the online dual for the arriving query and increments the advertiser-side dual variables according to the normalized budget consumption.

PROPOSITION 5 **(Buchbinder et al. [6], Mehta et al. [32])**. *With the exponential regularizer, the algorithm defined by* (EC.9) *is* $(1-1/\mathrm{e})$*-competitive for fractional online budgeted allocation in the small-bid regime.*

The proof is the same dual-fitting argument as in Section 3.2: the KKT system yields the online dual update, convexity of $F$ yields the lower bound on the advertiser-side dual variable, and the telescoping lower bound is written for normalized budget load. The only substantive change is that all loads are measured relative to budgets. In short, AdWords does not require a new primal–dual philosophy; it only requires the correct notion of load.

EXERCISE 1. Derive the KKT conditions for (EC.9) and show that the fitted dual update takes the form $\widehat{\alpha}_t := \alpha_t^*$ and $\widehat{\beta}_j \leftarrow \widehat{\beta}_j + x_{tj}^*(b_{tj} - \alpha_t^*)$. Verify that the analogue of the primal-gain-equals-dual-gain lemma from Section 3.2 holds with the same complementary-slackness argument.

### EC.4.2. Costly cancellations

In the free-disposal model of Section 3.3, previously allocated mass may be discarded at no cost. In many applications, however, revoking an allocation incurs a buyback cost. Suppose that disposing of one unit of allocation at reward level $r$ costs $\kappa r$ for some buyback parameter $\kappa \ge 0$. Then the cancellation term in (30) is scaled by $(1+\kappa)$:

$$\int_0^\infty (y_j(r) - x_j(r))r\ \mathrm{d}r \ \longrightarrow\ (1+\kappa)\int_0^\infty (y_j(r) - x_j(r))r\ \mathrm{d}r.$$

When $\kappa = 0$ we recover free disposal. As $\kappa$ increases, the algorithm becomes more reluctant to replace old allocations with new ones.

The convex-programming template extends directly. The stage program is unchanged except for the factor $(1+\kappa)$ in the disposal term, the KKT system yields the same style of threshold condition, and the dual-fitting proof again reduces to a one-dimensional identity. The effect of the buyback parameter is to decrease the advertiser-side dual increment and therefore the attainable competitive ratio. Ekbatani et al. [12] show that the optimal ratio varies continuously with $\kappa$, interpolating between $1 - 1/\mathrm{e}$ at $\kappa = 0$ and $1/2$ as $\kappa \to \infty$.

EXERCISE 2. Show that with buyback parameter $\kappa > 0$, the threshold from Lemma 6 becomes $G(r) = (1+\kappa)r - \int_0^r f(Y_j^{(t)}(\omega))\,\mathrm{d}\omega$. Verify that $G$ remains nondecreasing when $f(y) = \mathrm{e}^{y-1}$, and explain why the threshold is higher than in the free-disposal case.

EXERCISE 3. Assume $f(y) = \mathrm{e}^{y-1}$. Show that the slack-capacity case still yields the identity $\Gamma = 1 - f(y) + F(y)$, while the active-capacity case is weakened by the cancellation cost through the smaller $\beta$-increment. Recover the resulting competitive ratio as a function of $\kappa$; see Ekbatani et al. [12].

### EC.4.3. Vertex-weighted $b$-matching

Suppose offline node $j$ has capacity $B_j \geq 1$ rather than unit capacity, and each unit of that capacity is worth $w_j$. The offline LP benchmark is

$$\begin{aligned} \max \quad & \sum_{(t,j)\in E} w_j x_{tj} \\ \text{s.t.} \quad & \sum_{j:(t,j)\in E} x_{tj} \leq 1 \qquad \forall t, \\ & \sum_{t:(t,j)\in E} x_{tj} \leq B_j \quad \forall j, \\ & x_{tj} \geq 0. \end{aligned} \tag{EC.10}$$

The right state variable is the normalized load

$$\bar{y}_j := \frac{1}{B_j} \sum_{s:(s,j)\in E} x_{sj} \in [0,1].$$

After this normalization, the fully online regularized program becomes

$$\begin{aligned} \max_{x_{tj}\geq 0} \quad & \sum_{j\in N(t)} w_j x_{tj} - \sum_{j\in V} B_j w_j F\left(\bar{y}_j + \frac{x_{tj}}{B_j}\right) \\ \text{s.t.} \quad & \sum_{j\in N(t)} x_{tj} \leq 1, \\ & x_{tj} \leq B_j(1-\bar{y}_j) \qquad \forall j \in N(t). \end{aligned} \tag{EC.11}$$

The factor $B_j$ in front of the regularizer is exactly what preserves the geometry of the unit-capacity model: the penalty is applied to the fraction of capacity consumed rather than to the raw amount of consumed capacity.

The same normalization works in the batch-arrival model after replacing $F$ with the stage-dependent family $F_k$ and replacing $\bar{y}_j$ by the normalized pre-stage load. Every statement from the matching sections then carries over with $y_j$ replaced by $y_j/B_j$.

PROPOSITION 6. *In fractional vertex-weighted b-matching, the convex-regularization recipe yields the same optimal competitive ratios as in the unit-capacity case after normalizing loads by $B_j$. In particular, the exponential regularizer gives the classical $(1-1/\mathrm{e})$ guarantee in the fully online model, and the polynomial family from Section EC.2 gives the factor $\Gamma(K)=1-(1-1/K)^K$ in the $K$-stage model.*

The integral problem can then be handled by rounding. Under standard large-capacity assumptions, independent rounding of the fractional solution after a slight capacity reduction gives the usual loss of order $O(\sqrt{\log B_{\min}/B_{\min}})$, where $B_{\min} := \min_j B_j$.

### EC.4.4. Choice-based assortment and configuration views

The configuration-allocation model also captures assortment-style decisions. One may think of an arriving user $t$ as a customer and a feasible configuration $c \in \mathcal{C}_t$ as an offered assortment, ranked menu, or page layout. The coefficients $\xi_{t,c,j}$ specify how much resource $j$ is consumed when configuration $c$ is chosen, while $w_{t,c,j}$ specifies the associated revenue contribution. Under this interpretation, the stage-wise program from Section EC.3 is already an online assortment problem with resource constraints.

The modeling advantage is that one does not need a separate theorem for every such variant. Once an online assortment or whole-page decision can be written as "choose one configuration from a feasible family," the same price-level state, the same regularized stage program, and the same KKT-based dual construction apply. This is one reason the configuration model is a natural endpoint for Part I: many seemingly different online-allocation problems become different faces of the same convex-programming template.

For representative examples of this broader viewpoint, see Devanur et al. [10] for whole-page optimization and Aouad and Saban [2] for online assortment optimization on two-sided platforms.

## EC.5. Supplementary proofs for batch arrival

For this appendix only, it is convenient to index the stage loads one step earlier than in the main text. For a fixed run of the algorithm in Section EC.2, write

$$y_j^{(1)} := 0, \qquad y_j^{(k+1)} := y_j^{(k)} + \sum_{t:(t,j)\in E_k} x_{tj}^{(k)} \qquad (k=1,\dots,K),$$

where $x^{(k)} = \{x_{tj}^{(k)}\}$ is the optimizer of the stage-$k$ convex program. Thus $y_j^{(k)}$ is the load on offline node $j$ *before* stage $k$, while $y_j^{(k+1)}$ is the load after stage $k$.

### EC.5.1. Proof of Proposition 3

*Proof of Proposition 3.* Fix a stage $k \le K-1$, and write $n := K-k \ge 1$. Then

$$f_k(y) = \left(1 - \frac{1-y}{n}\right)^n.$$

Differentiating gives

$$f_k'(y) = \left(1 - \frac{1-y}{n}\right)^{n-1} \ge 0,$$

so $f_k$ is increasing on $[0,1]$. Since $F_k'(y) = f_k(y)$ and $f_k$ is nonconstant and increasing, $F_k$ is differentiable and strictly convex.

Next, set $z = (1-y)/n \in [0,1]$. The elementary bound $1 - z \le \mathrm{e}^{-z}$ yields

$$f_k(y) = \big(1-z\big)^n \le \mathrm{e}^{-nz} = \mathrm{e}^{y-1},$$

which proves the domination by the exponential rule.

For the pointwise convergence, let $k = 1$ and write $n = K-1$. Then

$$f_1(y) = \left(1 - \frac{1-y}{n}\right)^n \xrightarrow[n\to\infty]{} \mathrm{e}^{y-1}$$

by the standard limit $(1 - a/n)^n \to \mathrm{e}^{-a}$.

Finally, define

$$g(y) := (1-y) f_1(y) = (1-y)\left(1 - \frac{1-y}{K-1}\right)^{K-1}.$$

With the change of variable $u = 1-y$, this becomes

$$g(u) = u\left(1 - \frac{u}{K-1}\right)^{K-1}, \qquad u \in [0,1].$$

Differentiating,

$$g'(u) = \left(1 - \frac{u}{K-1}\right)^{K-2}\left(1 - \frac{Ku}{K-1}\right).$$

Hence $g$ is maximized at $u = (K-1)/K$, i.e., at $y = 1/K$. Substituting gives

$$\max_{y\in[0,1]} (1-y) f_1(y) = \frac{K-1}{K}\left(1 - \frac{1}{K}\right)^{K-1} = \left(1 - \frac{1}{K}\right)^K.$$

The last display is equivalent to

$$\Gamma(K) = 1 - \left(1 - \frac{1}{K}\right)^K = \min_{y\in[0,1]} \big[1 - (1-y) f_1(y)\big],$$

which is the required formula. □

The competitive-ratio proof relies on the following one-dimensional recursive inequality.

LEMMA 12. *Fix $K \in \mathbb{N}$, $k \in \{1, \dots, K-1\}$, and $x \in [0,1]$. For every nondecreasing sequence*

$$0 = u_1 \le u_2 \le \cdots \le u_k \le 1 - x,$$

*we have*

$$(1-x) f_k(u_k + x) - \sum_{s=1}^{k-1} (u_{s+1} - u_s) f_s(u_{s+1}) \le 1 - \Gamma(K)\,.$$

*Equivalently,*

$$1 - (1-x) f_k(u_k + x) + \sum_{s=1}^{k-1} (u_{s+1} - u_s) f_s(u_{s+1}) \ge \Gamma(K)\,.$$

*Proof.* We argue by induction on $k$.

For $k = 1$, the claim reduces to

$$(1-x) f_1(x) \le 1 - \Gamma(K)\,,$$

which is exactly the maximizing identity proved above with $y = x$.

Assume now that the statement holds for $k-1$, and consider stage $k$. By the recursive definition of the polynomials,

$$f_{k-1}(u_k) = \max_{y \in [0, 1-u_k]} (1-y) f_k(u_k + y).$$

Choosing $y = x$ is feasible because $u_k \le 1 - x$, hence

$$(1-x) f_k(u_k + x) \le f_{k-1}(u_k).$$

Therefore

$$\begin{aligned}(1-x) f_k(u_k + x) - \sum_{s=1}^{k-1} (u_{s+1} - u_s) f_s(u_{s+1}) &\le f_{k-1}(u_k) - \sum_{s=1}^{k-1} (u_{s+1} - u_s) f_s(u_{s+1}) \\ &= \big(1 - u_k + u_{k-1}\big) f_{k-1}(u_k) - \sum_{s=1}^{k-2} (u_{s+1} - u_s) f_s(u_{s+1}).\end{aligned}$$

Now apply the induction hypothesis to the shorter sequence

$$0 = u_1 \le u_2 \le \cdots \le u_{k-1}$$

with the choice $x' := u_k - u_{k-1}$. Since $u_k \le 1$, we have $u_{k-1} \le 1 - x'$, so the induction hypothesis yields the desired bound. □

### EC.5.2. Proof of Proposition 4

*Proof of Proposition 4.* Fix a stage $k \le K-1$, and abbreviate the optimizer of (EC.1) by $x = \{x_{tj}\}$. Introduce Lagrange multipliers $\lambda_t \ge 0$ for the online row constraints, $\theta_j \ge 0$ for the offline capacity constraints, and $\gamma_{tj} \ge 0$ for nonnegativity. The KKT system is

$$w_j\Big(1 - f_k\big(y_j^{(k)} + \sum_{t:(t,j)\in E_k} x_{tj}\big)\Big) - \lambda_t - \theta_j + \gamma_{tj} = 0 \qquad \forall (t,j) \in E_k, \tag{EC.12}$$

together with complementary slackness:

$$x_{tj} > 0 \implies \gamma_{tj} = 0, \tag{EC.13}$$

$$y_j^{(k)} + \sum_{t:(t,j)\in E_k} x_{tj} < 1 \implies \theta_j = 0, \tag{EC.14}$$

$$\sum_{j:(t,j)\in E_k} x_{tj} < 1 \implies \lambda_t = 0. \tag{EC.15}$$

Let $G'_k$ be the positive-support graph formed by the edges with $x_{tj} > 0$. Define

$$\mathcal{V}_{k,0} := \left\{ j \in V : y_j^{(k)} + \sum_{t:(t,j)\in E_k} x_{tj} = 1 \right\}$$

and let $\mathcal{U}_{k,0}$ be the set of neighbors of $\mathcal{V}_{k,0}$ inside $G'_k$. Remove these vertices and let $(\mathcal{U}_{k,\ell}, \mathcal{V}_{k,\ell})$, $\ell = 1, \dots, L_k$, be the connected components of the remaining support subgraph.

*Uniformity.* If $j \in \mathcal{V}_{k,0}$, then its post-stage load is one, so

$$w_j\big(1 - f_k(y_j^{(k+1)})\big) = w_j(1 - f_k(1)) = 0.$$

Set $c_{k,0} := 0$.

Now fix $\ell \ge 1$ and two offline vertices $j, j' \in \mathcal{V}_{k,\ell}$ that share an online neighbor $t \in \mathcal{U}_{k,\ell}$ in the positive-support graph. Since $x_{tj} > 0$ and $x_{tj'} > 0$, we have $\gamma_{tj} = \gamma_{tj'} = 0$ by (EC.13). Also, because $j, j' \notin \mathcal{V}_{k,0}$, both are unsaturated and (EC.14) gives $\theta_j = \theta_{j'} = 0$. Applying (EC.12) to $(t,j)$ and $(t,j')$ yields

$$w_j\big(1 - f_k(y_j^{(k+1)})\big) = \lambda_t = w_{j'}\big(1 - f_k(y_{j'}^{(k+1)})\big).$$

Since any two vertices in the same connected component can be joined by a path, the quantity above is constant on $\mathcal{V}_{k,\ell}$; call the common value $c_{k,\ell}$.

*Saturation.* Fix $\ell \ge 1$ and $t \in \mathcal{U}_{k,\ell}$. Choose a neighbor $j \in \mathcal{V}_{k,\ell}$ with $x_{tj} > 0$. Because $j \notin \mathcal{V}_{k,0}$, we have $y_j^{(k+1)} < 1$, and thus $f_k(y_j^{(k+1)}) < 1$. Hence

$$\lambda_t = w_j\big(1 - f_k(y_j^{(k+1)})\big) = c_{k,\ell} > 0.$$

Complementary slackness (EC.15) therefore implies $\sum_{j:(t,j)\in E_k} x_{tj} = 1$.

*Monotonicity.* Suppose there were an edge $(t,j') \in E_k$ with $t \in \mathcal{U}_{k,\ell}$ and $j' \in \mathcal{V}_{k,\ell'}$. Choose $j \in \mathcal{V}_{k,\ell}$ such that $x_{tj} > 0$; then $\gamma_{tj} = 0$. Applying (EC.12) to $(t,j)$ and $(t,j')$,

$$c_{k,\ell} - \lambda_t - \theta_j = 0, \qquad c_{k,\ell'} - \lambda_t - \theta_{j'} + \gamma_{tj'} = 0.$$

Thus

$$c_{k,\ell'} = \lambda_t + \theta_{j'} - \gamma_{tj'} \le \lambda_t \le c_{k,\ell}$$

whenever $\ell' \ge 1$, because then $\theta_{j'} = 0$. If $\ell' = 0$, then $c_{k,0} = 0$ anyway. Consequently, whenever $c_{k,\ell} < c_{k,\ell'}$, no edge can join $\mathcal{U}_{k,\ell}$ to $\mathcal{V}_{k,\ell'}$. □

### EC.5.3. Proof of Theorem 8

*Proof of Theorem 8.* Let $x^{(k)}$ be the optimizer of the stage-$k$ convex program for each $k = 1, \dots, K$, and let $(\mathcal{U}_{k,\ell}, \mathcal{V}_{k,\ell})$ and $c_{k,\ell}$ be the decomposition from Proposition 4 for every $k \le K-1$. Consider the weighted offline dual

$$\begin{aligned} \min \quad & \sum_{t \in U} \alpha_t + \sum_{j \in V} \beta_j \\ \text{s.t.} \quad & \alpha_t + \beta_j \ge w_j \qquad \forall (t,j) \in E, \\ & \alpha_t, \beta_j \ge 0. \end{aligned} \tag{EC.16}$$

We construct a feasible scaled dual solution.

For stages $k \le K-1$, set

$$\widehat{\alpha}_t := c_{k,\ell} \qquad \text{for every } t \in \mathcal{U}_{k,\ell},$$

and define the stage-wise increment of $\widehat{\beta}_j$ by

$$\Delta\widehat{\beta}_j^{(k)} := \big(y_j^{(k+1)} - y_j^{(k)}\big)\big(w_j - c_{k,\ell}\big) = w_j\big(y_j^{(k+1)} - y_j^{(k)}\big) f_k(y_j^{(k+1)}) \qquad \text{for } j \in \mathcal{V}_{k,\ell}. \tag{EC.17}$$

At the last stage $K$, the regularizer vanishes, so the stage problem is just the residual maximum-weight fractional matching problem with capacities $1 - y_j^{(K)}$. Let $(\bar{\alpha}_t)_{t \in U_K}, (\bar{\beta}_j)_{j \in V}$ be an optimal dual solution of that residual LP, i.e.,

$$\bar{\alpha}_t + \bar{\beta}_j \ge w_j \qquad \forall (t,j) \in E_K,$$

and

$$\sum_{t \in U_K} \bar{\alpha}_t + \sum_{j \in V} (1 - y_j^{(K)}) \bar{\beta}_j = \sum_{(t,j) \in E_K} w_j x_{tj}^{(K)}.$$

Set

$$\widehat{\alpha}_t := \bar{\alpha}_t \quad (t \in U_K), \qquad \Delta\widehat{\beta}_j^{(K)} := (1 - y_j^{(K)})\bar{\beta}_j,$$

and define $\widehat{\beta}_j := \sum_{k=1}^{K} \Delta\widehat{\beta}_j^{(k)}$.

*Primal and dual objective increments agree.* Fix a stage $k \le K-1$. Since only edges inside the support graph contribute, it is enough to compare the primal and dual contributions componentwise.

If $\ell = 0$, then $c_{k,0} = 0$, and hence

$$\sum_{t \in \mathcal{U}_{k,0}} \widehat{\alpha}_t + \sum_{j \in \mathcal{V}_{k,0}} \Delta\widehat{\beta}_j^{(k)} = \sum_{j \in \mathcal{V}_{k,0}} \sum_{t:(t,j)\in E_k} w_j x_{tj}^{(k)},$$

which is exactly the primal reward collected on that block.

If $\ell \ge 1$, then by saturation from Proposition 4, $\sum_{j:(t,j)\in E_k} x_{tj}^{(k)} = 1$ for every $t \in \mathcal{U}_{k,\ell}$. Therefore

$$\begin{aligned}
\sum_{t \in \mathcal{U}_{k,\ell}} \widehat{\alpha}_t + \sum_{j \in \mathcal{V}_{k,\ell}} \Delta\widehat{\beta}_j^{(k)} &= |\mathcal{U}_{k,\ell}|\, c_{k,\ell} + \sum_{j \in \mathcal{V}_{k,\ell}} \big(y_j^{(k+1)} - y_j^{(k)}\big)\big(w_j - c_{k,\ell}\big) \\
&= \sum_{j \in \mathcal{V}_{k,\ell}} w_j \big(y_j^{(k+1)} - y_j^{(k)}\big) + c_{k,\ell} \left( |\mathcal{U}_{k,\ell}| - \sum_{j \in \mathcal{V}_{k,\ell}} (y_j^{(k+1)} - y_j^{(k)}) \right) \\
&= \sum_{j \in \mathcal{V}_{k,\ell}} w_j \big(y_j^{(k+1)} - y_j^{(k)}\big).
\end{aligned}$$

The last equality uses conservation of mass inside the component: every edge carrying positive mass from a node in $\mathcal{U}_{k,\ell}$ ends in $\mathcal{V}_{k,\ell}$, and every node in $\mathcal{U}_{k,\ell}$ is fully matched. Therefore

$$\sum_{j \in \mathcal{V}_{k,\ell}} \big(y_j^{(k+1)} - y_j^{(k)}\big) = \sum_{t \in \mathcal{U}_{k,\ell}} \sum_{j \in \mathcal{V}_{k,\ell}} x_{tj}^{(k)} = |\mathcal{U}_{k,\ell}|.$$

So the displayed quantity is exactly the primal reward collected on that block. At stage $K$, equality of the primal and dual objectives follows from the optimality of the residual dual solution $(\bar{\alpha}, \bar{\beta})$. Summing over all stages gives

$$\sum_{t \in U} \widehat{\alpha}_t + \sum_{j \in V} \widehat{\beta}_j = \mathrm{ALG}.$$

*Approximate dual feasibility.* We show that every edge satisfies

$$\widehat{\alpha}_t + \widehat{\beta}_j \ge \Gamma(K)\, w_j.$$

Fix an edge $(t,j) \in E_k$ with $k \le K-1$. Suppose $t \in \mathcal{U}_{k,\ell}$ and $j \in \mathcal{V}_{k,\ell'}$. By monotonicity, $c_{k,\ell'} \le c_{k,\ell} = \widehat{\alpha}_t$. Hence, using (EC.17),

$$\begin{aligned}
\widehat{\alpha}_t + \sum_{s=1}^{k} \Delta\widehat{\beta}_j^{(s)} &\ge c_{k,\ell'} + \Delta\widehat{\beta}_j^{(k)} + \sum_{s=1}^{k-1} \Delta\widehat{\beta}_j^{(s)} \\
&= w_j\big(1 - f_k(y_j^{(k+1)})\big) + w_j\big(y_j^{(k+1)} - y_j^{(k)}\big) f_k(y_j^{(k+1)}) + \sum_{s=1}^{k-1} \Delta\widehat{\beta}_j^{(s)} \\
&= w_j \left[ 1 - \big(1 - y_j^{(k+1)} + y_j^{(k)}\big) f_k(y_j^{(k+1)}) + \sum_{s=1}^{k-1} \big(y_j^{(s+1)} - y_j^{(s)}\big) f_s(y_j^{(s+1)}) \right].
\end{aligned}$$

Apply Lemma 12 with

$$x := y_j^{(k+1)} - y_j^{(k)}, \qquad u_s := y_j^{(s)} \quad (s = 1, \dots, k),$$

and note that $u_k = y_j^{(k)}$ and $u_k + x = y_j^{(k+1)}$. This yields

$$\widehat{\alpha}_t + \sum_{s=1}^{k} \Delta\widehat{\beta}_j^{(s)} \geq \Gamma(K)\, w_j.$$

Since the remaining increments of $\widehat{\beta}_j$ are nonnegative, the full dual constraint is satisfied.

Now consider an edge $(t, j) \in E_K$ from the last stage. By construction,

$$\widehat{\alpha}_t + \Delta\widehat{\beta}_j^{(K)} = \bar{\alpha}_t + (1 - y_j^{(K)})\bar{\beta}_j \geq (1 - y_j^{(K)})w_j.$$

If $y_j^{(K)} = 0$, then this is already at least $\Gamma(K)\, w_j$. Otherwise, let $k' < K$ be the last stage with positive increment on node $j$, so $y_j^{(K)} = y_j^{(k'+1)}$. Since $f_{k'}(y) = \left(1 - \frac{1-y}{K-k'}\right)^{K-k'} \geq y$ for every $y \in [0, 1]$ by Bernoulli's inequality,

$$(1 - y_j^{(K)})w_j \geq \big(1 - f_{k'}(y_j^{(k'+1)})\big)w_j.$$

Therefore

$$\begin{aligned}
\widehat{\alpha}_t + \widehat{\beta}_j &\geq (1 - y_j^{(K)})w_j + \sum_{s=1}^{k'} \Delta\widehat{\beta}_j^{(s)} \\
&\geq \big(1 - f_{k'}(y_j^{(k'+1)})\big)w_j + \sum_{s=1}^{k'} \Delta\widehat{\beta}_j^{(s)} \\
&= w_j \left[1 - \big(1 - y_j^{(k'+1)} + y_j^{(k')}\big) f_{k'}(y_j^{(k'+1)}) + \sum_{s=1}^{k'-1} \big(y_j^{(s+1)} - y_j^{(s)}\big) f_s(y_j^{(s+1)})\right] \\
&\geq \Gamma(K)\, w_j
\end{aligned}$$

by one more application of Lemma 12. This proves approximate dual feasibility. Weak duality now gives $\mathrm{ALG} \geq \Gamma(K)\,\mathrm{OPT}$. □

## EC.6. Supplementary proofs for configuration allocation

Throughout this section, write

$$\Delta w_\tau := \widehat{w}_\tau - \widehat{w}_{\tau-1}$$

and, for each stage $k$, let

$$H_{j,\tau}^{(k)} := H_{j,\tau}(x^{(k)}, y^{(k)})$$

for the cumulative price-level state of advertiser $j$ at threshold $\widehat{w}_\tau$. We also set $H_{j,\tau}^{(0)} := 0$.

### EC.6.1. Proofs of the structural lemmas

We begin with the offline benchmark. The K-stage fractional configuration-allocation LP is

$$\begin{aligned}
\max \quad & \sum_{t\in U}\sum_{c\in\mathcal{C}}\sum_{j\in V} w_{t,c,j}x_{t,c,j} & \\
\text{s.t.} \quad & \sum_{t\in U}\sum_{c\in\mathcal{C}} x_{t,c,j} \le 1 & \forall j\in V,\\
& \sum_{c\in\mathcal{C}} z_{t,c} \le 1 & \forall t\in U,\\
& x_{t,c,j} \le \xi_{t,c,j} z_{t,c} & \forall t,c,j,\\
& x_{t,c,j}, z_{t,c} \ge 0. &
\end{aligned} \tag{EC.18}$$

Its dual is

$$\begin{aligned}
\min \quad & \sum_{t\in U}\alpha_t + \sum_{j\in V}\beta_j & \\
\text{s.t.} \quad & \gamma_{t,c,j} + \beta_j \ge w_{t,c,j} & \forall t,c,j,\\
& \alpha_t \ge \sum_{j\in V}\xi_{t,c,j}\gamma_{t,c,j} & \forall t,c,\\
& \alpha_t, \beta_j, \gamma_{t,c,j} \ge 0. &
\end{aligned} \tag{EC.19}$$

For the stage-wise analysis, fix a stage $k \le K-1$, and let $(x^*, y^*, z^*)$ be the optimizer of (EC.5a)–(EC.5e). Introduce Lagrange multipliers $\lambda_t \ge 0$ for the configuration-choice constraints, $\theta_j \ge 0$ for the advertiser-capacity constraints, $\mu_{t,c,j} \ge 0$ for the link constraints, $\pi_{j,\tau} \ge 0$ for the carry constraints, $\psi_{t,c,j} \ge 0$ for nonnegativity of $x_{t,c,j}$, $\phi_{t,c} \ge 0$ for nonnegativity of $z_{t,c}$, and $\iota_{j,\tau} \ge 0$ for nonnegativity of $y_{j,\tau}$.

*Proof of Lemma 10.* The KKT system consists of primal feasibility, dual feasibility, complementary slackness, and stationarity. The complementary-slackness relations needed later are

$$\lambda_t\Big(\sum_{c\in\mathcal{C}} z^*_{t,c} - 1\Big) = 0, \tag{EC.20}$$

$$\mu_{t,c,j}\big(x^*_{t,c,j} - \xi_{t,c,j} z^*_{t,c}\big) = 0, \tag{EC.21}$$

$$\pi_{j,\tau}\big(y^*_{j,\tau} - \eta^{(k-1)}_{j,\tau}\big) = 0, \qquad \iota_{j,\tau} y^*_{j,\tau} = 0, \tag{EC.22}$$

$$\psi_{t,c,j} x^*_{t,c,j} = 0, \qquad \phi_{t,c} z^*_{t,c} = 0. \tag{EC.23}$$

The stationarity relations with respect to $z_{t,c}$ and $x_{t,c,j}$ are

$$\sum_{j\in V}\xi_{t,c,j}\mu_{t,c,j} - \lambda_t + \phi_{t,c} = 0, \tag{EC.24}$$

$$w_{t,c,j} - \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau f_k(H^{(k)}_{j,\tau}) - \theta_j - \mu_{t,c,j} + \psi_{t,c,j} = 0. \tag{EC.25}$$

Similarly, stationarity with respect to $y_{j,\tau}$ gives

$$\widehat{w}_\tau - \sum_{\tau'\le\tau} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) - \theta_j - \pi_{j,\tau} + \iota_{j,\tau} = 0. \tag{EC.26}$$

It remains to show that $\theta_j = 0$ for every advertiser $j$ and stage $k \le K-1$. Suppose instead that $\theta_j > 0$. Then complementary slackness for the capacity constraint forces advertiser $j$ to be exactly full after stage $k$. Choose the smallest price index $\tau^\star$ that appears with positive retained mass after the stage, in the following sense: either $y^*_{j,\tau^\star} > 0$, or there exists some $(t,c)$ with $x^*_{t,c,j} > 0$ and $w_{t,c,j} = \widehat{w}_{\tau^\star}$. By minimality of $\tau^\star$, every retained unit of mass for advertiser $j$ has price at least $\widehat{w}_{\tau^\star}$. Since the total retained mass equals one, we must therefore have $H^{(k)}_{j,\tau'} = 1$ for every $\tau' \le \tau^\star$. Because $f_k(1) = 1$,

$$\sum_{\tau' \le \tau^\star} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) = \sum_{\tau' \le \tau^\star} \Delta w_{\tau'} = \widehat{w}_{\tau^\star}.$$

We now distinguish two cases.

If $y^*_{j,\tau^\star} > 0$, then $\iota_{j,\tau^\star} = 0$ by (EC.22). Plugging the previous display into (EC.26) with $\tau = \tau^\star$ gives

$$0 = \widehat{w}_{\tau^\star} - \sum_{\tau' \le \tau^\star} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) - \theta_j - \pi_{j,\tau^\star} = -\theta_j - \pi_{j,\tau^\star},$$

a contradiction.

If instead $y^*_{j,\tau^\star} = 0$, then by definition of $\tau^\star$ there exists some pair $(t,c)$ with $x^*_{t,c,j} > 0$ and $w_{t,c,j} = \widehat{w}_{\tau^\star}$. In that case $\psi_{t,c,j} = 0$ by (EC.23), and (EC.25) yields

$$0 = \widehat{w}_{\tau^\star} - \sum_{\tau' \le \tau^\star} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) - \theta_j - \mu_{t,c,j} = -\theta_j - \mu_{t,c,j},$$

again a contradiction.

Thus $\theta_j = 0$ for every advertiser $j$ and every stage $k \le K-1$. Substituting this fact into (EC.24) and (EC.25) yields exactly the identities stated in Lemma 10. □

*Proof of Lemma 11.* Assume that $\eta^{(k-1)}_{j,\tau} - y^*_{j,\tau} > 0$. Then the carry constraint at $(j,\tau)$ is slack, so $\pi_{j,\tau} = 0$ by complementary slackness. Since we already know that $\theta_j = 0$, (EC.26) gives

$$\widehat{w}_\tau - \sum_{\tau' \le \tau} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) + \iota_{j,\tau} = 0.$$

Hence

$$\sum_{\tau' \le \tau} \Delta w_{\tau'} f_k(H^{(k)}_{j,\tau'}) \ge \widehat{w}_\tau = \sum_{\tau' \le \tau} \Delta w_{\tau'}.$$

Because each term $f_k(H^{(k)}_{j,\tau'})$ lies in $[0,1]$, equality can hold only if $f_k(H^{(k)}_{j,\tau'}) = 1$ for every $\tau' \le \tau$. The polynomial $f_k$ is increasing and attains the value one only at the point one. Thus $H^{(k)}_{j,\tau} = 1$, as claimed. □

*Proof of Corollary 1.* Fix $(j,\tau)$. If

$$\sum_{\tau' \ge \tau} \left(\eta^{(k-1)}_{j,\tau'} - y^*_{j,\tau'}\right) = 0,$$

then no previously retained mass at levels $\tau, \tau+1, \dots, T$ is discarded, so

$$H_{j,\tau}^{(k)} = \sum_{t\in U_k,\, c\in\mathcal{C}:\ w_{t,c,j}\ge \widehat{w}_\tau} x_{t,c,j}^{*} + \sum_{\tau'\ge\tau} y_{j,\tau'}^{*} \ge \sum_{\tau'\ge\tau} \eta_{j,\tau'}^{(k-1)}.$$

This proves part (i), and part (ii) is immediate because both sides are zero.

Otherwise,

$$\sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^{*}\big) > 0.$$

Then there exists some $\tau' \ge \tau$ with $\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^{*} > 0$. By Lemma 11, $H_{j,\tau'}^{(k)} = 1$. Since cumulative states are nonincreasing in the threshold, $H_{j,\tau}^{(k)} \ge H_{j,\tau'}^{(k)} = 1$, and hence

$$H_{j,\tau}^{(k)} = 1 \ge \sum_{\tau'\ge\tau} \eta_{j,\tau'}^{(k-1)},$$

which again proves part (i). In this case $f_k(H_{j,\tau}^{(k)}) = f_k(1) = 1$, so part (ii) follows immediately as well. □

For the last stage we need one additional observation about the optimal LP dual.

LEMMA 13. *Fix an optimal extreme-point dual solution of the stage-$K$ linear program. Without loss of generality, $\bar{\theta}_j \in \{\widehat{w}_0, \dots, \widehat{w}_T\}$ for every advertiser $j$. Let $\bar{\theta}_j = \widehat{w}_{\tau_j^\star}$. Then, for every $\tau \in [T]$,*

$$\tau \le \tau_j^\star \implies \eta_{j,\tau}^{(K-1)} \bar{\pi}_{j,\tau} = 0, \tag{EC.27}$$

$$\tau > \tau_j^\star \implies \eta_{j,\tau}^{(K-1)} \big(\widehat{w}_\tau - \bar{\pi}_{j,\tau}\big) = \eta_{j,\tau}^{(K-1)} \bar{\theta}_j. \tag{EC.28}$$

*Proof.* At stage $K$, the problem is linear, and the dual constraints involving $(j,\tau)$ are

$$\bar{\theta}_j + \bar{\pi}_{j,\tau} \ge \widehat{w}_\tau, \qquad \bar{\pi}_{j,\tau} \ge 0.$$

Because the dual objective contains the term $\eta_{j,\tau}^{(K-1)} \bar{\pi}_{j,\tau}$, optimality sets $\bar{\pi}_{j,\tau}$ to its smallest feasible value whenever $\eta_{j,\tau}^{(K-1)} > 0$. If $\tau \le \tau_j^\star$, that minimum is zero, giving (EC.27). If $\tau > \tau_j^\star$, that minimum is $\widehat{w}_\tau - \bar{\theta}_j$, giving (EC.28). If $\eta_{j,\tau}^{(K-1)} = 0$, both displays are trivial. □

### EC.6.2. Proof of Theorem 9

*Proof of Theorem 9.* For each stage $k \le K-1$, let $(x^{(k)}, y^{(k)}, z^{(k)})$ be the optimizer of the stage program and let $\lambda^{(k)}, \mu^{(k)}, \psi^{(k)}, \phi^{(k)}$ be the multipliers supplied by Lemma 10. For the last stage $K$, let $(\bar{\lambda}, \bar{\theta}, \bar{\mu}, \bar{\pi})$ be an optimal extreme-point dual solution of the residual LP, chosen as in Lemma 13.

We define a candidate dual solution for the offline LP (EC.19) as follows.

- For every user $t \in U_k$ with $k \le K-1$, set $\widehat{\alpha}_t := \lambda_t^{(k)}$.
- For every triple $(t,c,j)$ with $t \in U_k$ and $k \le K-1$, set $\widehat{\gamma}_{t,c,j} := \mu_{t,c,j}^{(k)}$.
- For every advertiser $j$, define the stage-$k$ advertiser-dual increment

$$\Delta\widehat{\beta}_j^{(k)} := \sum_{t\in U_k} \sum_{c\in\mathcal{C}} x_{t,c,j}^{(k)} \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau f_k(H_{j,\tau}^{(k)}) - \sum_{\tau\in[T]} \widehat{w}_\tau \big(\eta_{j,\tau}^{(k-1)} - y_{j,\tau}^{(k)}\big) \qquad (k \le K-1). \tag{EC.29}$$

At stage $K$, set

$$\widehat{\alpha}_t := \bar{\lambda}_t, \qquad \widehat{\gamma}_{t,c,j} := \bar{\mu}_{t,c,j} \qquad (t \in U_K),$$

and

$$\Delta\widehat{\beta}_j^{(K)} := \bar{\theta}_j - \sum_{\tau\in[T]} \eta_{j,\tau}^{(K-1)}\big(\widehat{w}_\tau - \bar{\pi}_{j,\tau}\big). \tag{EC.30}$$

Finally define $\widehat{\beta}_j := \sum_{k=1}^{K} \Delta\widehat{\beta}_j^{(k)}$.

*Objective equality for stages* $k \le K-1$. Fix such a stage. By (EC.20) and (EC.24),

$$\begin{aligned}\sum_{t\in U_k} \widehat{\alpha}_t &= \sum_{t\in U_k} \lambda_t^{(k)} \sum_{c\in\mathcal{C}} z_{t,c}^{(k)} \\ &= \sum_{t\in U_k}\sum_{c\in\mathcal{C}}\sum_{j\in V} \xi_{t,c,j} z_{t,c}^{(k)} \mu_{t,c,j}^{(k)} - \sum_{t\in U_k}\sum_{c\in\mathcal{C}} \phi_{t,c}^{(k)} z_{t,c}^{(k)} \\ &= \sum_{t\in U_k}\sum_{c\in\mathcal{C}}\sum_{j\in V} x_{t,c,j}^{(k)} \mu_{t,c,j}^{(k)},\end{aligned}$$

where the last equality uses (EC.21) and (EC.23). Now apply (EC.25) with $\theta_j = 0$ and $\psi_{t,c,j}^{(k)} x_{t,c,j}^{(k)} = 0$:

$$x_{t,c,j}^{(k)} \mu_{t,c,j}^{(k)} = x_{t,c,j}^{(k)} w_{t,c,j} - x_{t,c,j}^{(k)} \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau f_k(H_{j,\tau}^{(k)}).$$

Summing over all $(t,c,j)$ and then adding (EC.29) gives

$$\sum_{t\in U_k} \widehat{\alpha}_t + \sum_{j\in V} \Delta\widehat{\beta}_j^{(k)} = \sum_{t\in U_k}\sum_{c\in\mathcal{C}}\sum_{j\in V} w_{t,c,j} x_{t,c,j}^{(k)} - \sum_{j\in V}\sum_{\tau\in[T]} \widehat{w}_\tau\big(\eta_{j,\tau}^{(k-1)} - y_{j,\tau}^{(k)}\big),$$

which is exactly the stage-$k$ contribution to the primal objective.

*Objective equality at stage* $K$. The stage-$K$ problem is linear. Its dual objective is

$$\sum_{t\in U_K} \bar{\lambda}_t + \sum_{j\in V} \bar{\theta}_j + \sum_{j\in V}\sum_{\tau\in[T]} \eta_{j,\tau}^{(K-1)} \bar{\pi}_{j,\tau}.$$

By strong duality, this equals the stage-$K$ primal objective, and by (EC.30)

$$\sum_{t\in U_K} \widehat{\alpha}_t + \sum_{j\in V} \Delta\widehat{\beta}_j^{(K)} = \sum_{t\in U_K} \bar{\lambda}_t + \sum_{j\in V} \bar{\theta}_j + \sum_{j\in V}\sum_{\tau\in[T]} \eta_{j,\tau}^{(K-1)} \bar{\pi}_{j,\tau} - \sum_{j\in V}\sum_{\tau\in[T]} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau,$$

which is again exactly the stage-$K$ primal contribution. Therefore

$$\sum_{t\in U} \widehat{\alpha}_t + \sum_{j\in V} \widehat{\beta}_j = \mathrm{ALG}.$$

*Nonnegativity of the advertiser dual increments.* For $k \le K-1$, reorder the summation in (EC.29):

$$\begin{aligned}\Delta\widehat{\beta}_j^{(k)} &= \sum_{\tau\in[T]} \Delta w_\tau \left( \sum_{t\in U_k,\, c\in\mathcal{C}:\ w_{t,c,j}\ge \widehat{w}_\tau} x_{t,c,j}^{(k)} \right) f_k(H_{j,\tau}^{(k)}) - \sum_{\tau\in[T]} \widehat{w}_\tau \big(\eta_{j,\tau}^{(k-1)} - y_{j,\tau}^{(k)}\big) \\ &= \sum_{\tau\in[T]} \Delta w_\tau \Big(H_{j,\tau}^{(k)} - \sum_{\tau'\ge\tau} y_{j,\tau'}^{(k)}\Big) f_k(H_{j,\tau}^{(k)}) - \sum_{\tau\in[T]} \Delta w_\tau \sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^{(k)}\big).\end{aligned}$$

By Corollary 1(ii),

$$\sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^{(k)}\big) = \sum_{\tau'\ge\tau} \big(\eta_{j,\tau'}^{(k-1)} - y_{j,\tau'}^{(k)}\big) f_k(H_{j,\tau}^{(k)}).$$

Substituting and simplifying,

$$\Delta\widehat{\beta}_j^{(k)} = \sum_{\tau\in[T]} \Delta w_\tau \big(H_{j,\tau}^{(k)} - H_{j,\tau}^{(k-1)}\big) f_k(H_{j,\tau}^{(k)}) \ge 0,$$

where the last inequality uses Corollary 1(i). For stage $K$, dual feasibility implies $\widehat{w}_\tau - \bar{\pi}_{j,\tau} \le \bar{\theta}_j$, and hence

$$\Delta\widehat{\beta}_j^{(K)} \ge \bar{\theta}_j - \bar{\theta}_j \sum_{\tau\in[T]} \eta_{j,\tau}^{(K-1)} \ge 0.$$

Thus every advertiser-dual increment is nonnegative.

*Feasibility of the $\alpha$-constraints.* For $k \le K-1$, (EC.24) gives

$$\widehat{\alpha}_t = \lambda_t^{(k)} = \sum_{j\in V} \xi_{t,c,j}\mu_{t,c,j}^{(k)} + \phi_{t,c}^{(k)} \ge \sum_{j\in V} \xi_{t,c,j}\widehat{\gamma}_{t,c,j} \qquad \forall c\in\mathcal{C}.$$

For stage $K$, the same inequality is one of the LP dual constraints. So the $\alpha$-constraints of (EC.19) are satisfied.

*Approximate feasibility for stages $k \le K-1$.* Fix a triple $(t,c,j)$ with $t\in U_k$, and let $\tau(t,c,j)$ denote the price index of $w_{t,c,j}$. Since $\psi_{t,c,j}^{(k)} \ge 0$, (EC.25) gives

$$\widehat{\gamma}_{t,c,j} = \mu_{t,c,j}^{(k)} \ge w_{t,c,j} - \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau f_k(H_{j,\tau}^{(k)}).$$

Combining this with the expression for $\Delta\widehat{\beta}_j^{(s)}$ derived above,

$$\widehat{\gamma}_{t,c,j} + \sum_{s=1}^{k} \Delta\widehat{\beta}_j^{(s)} \ge \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau \left[1 - f_k(H_{j,\tau}^{(k)}) + \sum_{s=1}^{k} \big(H_{j,\tau}^{(s)} - H_{j,\tau}^{(s-1)}\big) f_s(H_{j,\tau}^{(s)})\right].$$

For each threshold $\tau$ in the sum, apply Lemma 12 to the nondecreasing sequence

$$0 = H_{j,\tau}^{(0)} \le H_{j,\tau}^{(1)} \le \cdots \le H_{j,\tau}^{(k)}$$

with $x = H_{j,\tau}^{(k)} - H_{j,\tau}^{(k-1)}$. This yields

$$1 - f_k(H_{j,\tau}^{(k)}) + \sum_{s=1}^{k} \big(H_{j,\tau}^{(s)} - H_{j,\tau}^{(s-1)}\big) f_s(H_{j,\tau}^{(s)}) \ge \Gamma(K)\,.$$

Hence

$$\widehat{\gamma}_{t,c,j} + \sum_{s=1}^{k} \Delta\widehat{\beta}_j^{(s)} \ge \Gamma(K) \sum_{\tau:\, \widehat{w}_\tau \le w_{t,c,j}} \Delta w_\tau = \Gamma(K)\, w_{t,c,j}.$$

Since all later increments of $\widehat{\beta}_j$ are nonnegative, the full approximate constraint holds.

*Approximate feasibility at stage $K$.* Fix a triple $(t,c,j)$ with $t \in U_K$, and let $\tau^\dagger := \tau(t,c,j)$. Also write $\bar{\theta}_j = \widehat{w}_{\tau_j^\star}$ as in Lemma 13. We distinguish two cases.

*Case 1: $\tau^\dagger < \tau_j^\star$, i.e., $w_{t,c,j} < \bar{\theta}_j$.* By Lemma 13,

$$\Delta\widehat{\beta}_j^{(K)} = \bar{\theta}_j - \sum_{\tau > \tau_j^\star} \eta_{j,\tau}^{(K-1)} \bar{\theta}_j - \sum_{\tau \le \tau_j^\star} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.$$

Split the second sum at $\tau^\dagger$ and use $\widehat{w}_\tau \le \bar{\theta}_j$ for $\tau^\dagger \le \tau \le \tau_j^\star$:

$$\Delta\widehat{\beta}_j^{(K)} \ge \bar{\theta}_j \Big(1 - \sum_{\tau \ge \tau^\dagger} \eta_{j,\tau}^{(K-1)}\Big) - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.$$

Since $\bar{\mu}_{t,c,j} + \bar{\theta}_j \ge w_{t,c,j}$ and $\bar{\mu}_{t,c,j} \ge 0$,

$$\begin{aligned}\bar{\mu}_{t,c,j} + \Delta\widehat{\beta}_j^{(K)} &\ge \Big(1 - \sum_{\tau \ge \tau^\dagger} \eta_{j,\tau}^{(K-1)}\Big)(\bar{\mu}_{t,c,j} + \bar{\theta}_j) - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau \\ &\ge \Big(1 - \sum_{\tau \ge \tau^\dagger} \eta_{j,\tau}^{(K-1)}\Big) w_{t,c,j} - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.\end{aligned}$$

The right-hand side can be rewritten as

$$\sum_{\tau \le \tau^\dagger} \Delta w_\tau \big(1 - H_{j,\tau}^{(K-1)}\big),$$

which equals

$$\sum_{\tau \le \tau^\dagger} \Delta w_\tau \big(1 - f_{K-1}(H_{j,\tau}^{(K-1)})\big)$$

because $f_{K-1}(x) = x$.

*Case 2: $\tau^\dagger \ge \tau_j^\star$, i.e., $w_{t,c,j} \ge \bar{\theta}_j$.* Again by Lemma 13,

$$\Delta\widehat{\beta}_j^{(K)} = \bar{\theta}_j - \sum_{\tau > \tau_j^\star} \eta_{j,\tau}^{(K-1)} \bar{\theta}_j - \sum_{\tau \le \tau_j^\star} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.$$

Now use $\bar{\theta}_j \le \widehat{w}_\tau$ for $\tau_j^\star < \tau \le \tau^\dagger$ to obtain

$$\Delta\widehat{\beta}_j^{(K)} \ge \bar{\theta}_j - \sum_{\tau > \tau^\dagger} \eta_{j,\tau}^{(K-1)} \bar{\theta}_j - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.$$

Using again $\bar{\mu}_{t,c,j} + \bar{\theta}_j \ge w_{t,c,j}$,

$$\begin{aligned}\bar{\mu}_{t,c,j} + \Delta\widehat{\beta}_j^{(K)} &\ge \Big(1 - \sum_{\tau \ge \tau^\dagger} \eta_{j,\tau}^{(K-1)}\Big)(\bar{\mu}_{t,c,j} + \bar{\theta}_j) - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau \\ &\ge \Big(1 - \sum_{\tau \ge \tau^\dagger} \eta_{j,\tau}^{(K-1)}\Big) w_{t,c,j} - \sum_{\tau < \tau^\dagger} \eta_{j,\tau}^{(K-1)} \widehat{w}_\tau.\end{aligned}$$

The same telescoping rewrite as in Case 1 shows that the right-hand side equals

$$\sum_{\tau \le \tau^\dagger} \Delta w_\tau \big(1 - f_{K-1}(H_{j,\tau}^{(K-1)})\big).$$

Thus, in both cases,

$$\widehat{\gamma}_{t,c,j} + \Delta \widehat{\beta}_j^{(K)} \ge \sum_{\tau \le \tau(t,c,j)} \Delta w_\tau \big(1 - f_{K-1}(H_{j,\tau}^{(K-1)})\big).$$

Adding the earlier advertiser-dual increments and applying the same threshold-by-threshold recursion as above, we obtain

$$\widehat{\gamma}_{t,c,j} + \widehat{\beta}_j \ge \Gamma(K)\, w_{t,c,j}.$$

This proves approximate dual feasibility for every triple $(t, c, j)$. By weak duality for (EC.19), the constructed dual satisfies $\mathrm{ALG} \ge \Gamma(K)\,\mathrm{OPT}$. □