A. Omitted Proofs

Lemma 12. *The algorithm maintains the invariant that* $\vec{0} \le \vec{x} \le \vec{z}$.

Proof. We show the lemma by induction on the number of updates (lines 20 and 21). Consider an iteration of the inner while loop. If the algorithm executes line 23, we have $\vec{x} = \vec{z}$ at the end of the iteration. Therefore we may assume that the algorithm does not execute line 23. Let $\vec{x}'$ and $\vec{z}'$ be the updated vectors after performing the updates on line 20 and line 21, respectively. Let $\vec{x}$ and $\vec{z}$ denote the vectors right before the update. By the induction hypothesis, we have $\vec{0} \le \vec{x} \le \vec{z}$.

For each coordinate $i \in [n]$, we have:

$$\vec{x}'_i = \begin{cases} \vec{x}_i + \eta(1 - \vec{x}_i) = \eta + (1-\eta)\vec{x}_i & \text{if } i \in S(\eta) \\ \vec{x}_i & \text{otherwise} \end{cases}$$

$$\vec{z}'_i = \begin{cases} \vec{z}_i + \eta(1 - \vec{z}_i) = \eta + (1-\eta)\vec{z}_i & \text{if } i \in S \\ \vec{z}_i & \text{otherwise} \end{cases}$$

Since $\vec{0} \le \vec{x} \le \vec{z}$, $1 - \eta \ge 0$, and $S(\eta) \subseteq S$, we have $\vec{0} \le \vec{x}' \le \vec{z}'$, as needed. □

Lemma 13. *Consider an iteration of the inner while loop. Let $\vec{x}$ and $\vec{z}$ be the respective vectors before the updates on lines 20–23, and let $\vec{x}'$ and $\vec{z}' = \vec{z}(\eta)$ be the respective vectors after the updates. For each coordinate $i \in S(\eta)$, we have $\nabla_i f(\vec{x}') \ge \nabla_i f(\vec{z}') \ge \frac{v}{1-\vec{z}'_i} > 0$.*

Proof. Note that the update rule on line 21 sets $\vec{z}' = \vec{z}(\eta)$. Since $\vec{x}' \le \vec{z}'$ (Lemma 12), DR-submodularity implies that $\nabla f(\vec{x}') \ge \nabla f(\vec{z}')$. Let $i \in S(\eta)$. By the definition of $S(\eta)$, we have $\vec{z}_i(\eta) \le 1 - (1-\epsilon)^j < 1$ and $(1 - \vec{z}_i(\eta))\nabla_i f(\vec{z}(\eta)) \ge v > 0$, which implies that $\nabla_i f(\vec{z}(\eta)) \ge \frac{v}{1-\vec{z}_i(\eta)} > 0$. □

Lemma 14. *Consider the vectors and sets defined on line 15. For all η and η' such that $0 \le \eta \le \eta' \le \epsilon$, we have:*

(1) $\vec{z}(\eta) \le \vec{z}(\eta')$,

(2) $S(\eta) \supseteq S(\eta')$.

Proof. (1) For every $i \notin S$, we have $\vec{z}_i(\eta) = \vec{z}_i(\eta') = \vec{z}_i$. For every $i \in S$, we have:

$$\vec{z}_i(\eta) \overset{(a)}{=} \vec{z}_i + \eta(1 - \vec{z}_i) \overset{(b)}{\le} \vec{z}_i + \eta'(1 - \vec{z}_i) \overset{(c)}{=} \vec{z}_i(\eta')$$

where (a) and (c) are due to $i \in S$, (b) is due to $\eta \le \eta'$ and $1 - \vec{z}_i \ge 0$ (since $i \in S$, $\vec{z}_i \le 1 - (1-\epsilon)^j \le 1$).

(2) Let $i \in S(\eta')$. By (1) and DR-submodularity, we have $\nabla f(\vec{z}(\eta)) \ge \nabla f(\vec{z}(\eta'))$. Since $i \in S(\eta')$, we have $1 - z_i(\eta') \ge 0$ (since $z_i(\eta') \le 1 - (1-\epsilon)^j \le 1$), and $\nabla_i f(\vec{z}(\eta')) \ge 0$ (since $1 - z_i(\eta') \ge 0$ and $(1 - z_i(\eta'))\nabla_i f(\vec{z}(\eta')) \ge v > 0$). Therefore

$$\begin{aligned} \vec{g}_i(\eta) &\overset{(a)}{=} (1-\eta)(1-\vec{z}_i)\nabla_i f(\vec{z}(\eta)) \\ &\overset{(b)}{\ge} (1-\eta')(1-\vec{z}_i)\nabla_i f(\vec{z}(\eta')) \\ &\overset{(c)}{=} \vec{g}_i(\eta') \\ &\overset{(d)}{\ge} v \end{aligned}$$

where (a) and (c) are due to $i \in S$; (b) is due to $\eta \le \eta'$, $1 - \vec{z}_i \ge 0$, and $\nabla_i f(\vec{z}(\eta)) \ge \nabla_i f(\vec{z}(\eta')) \ge 0$; (d) is due to $i \in S(\eta')$.

□

A.1. Proof of Lemma 3

Proof of Lemma 3. We show that the invariants are maintained using induction on the number of iterations of the inner while loop in phase j. Let $\vec{z}$ be the vector right before the update on line 21 and let $\vec{z}'$ be the vector right after the update. By the induction hypothesis, we have $\vec{z}_i \le 1 - (1-\epsilon)^j + \epsilon^2$. If $i \notin S$, we have $\vec{z}'_i = \vec{z}_i$, and the invariant is maintained. Therefore we may assume that $i \in S$. By the definition of S, we have $\vec{z}_i \le 1 - (1-\epsilon)^j$. We have $\vec{z}'_i = \vec{z}_i + \eta(1 - \vec{z}'_i) \le \vec{z}_i + \eta \le \vec{z}_i + \epsilon^2$. Thus the invariant is maintained.

Next, we show the upper bound on the ℓ_1 norm. Note that $\vec{z}' = \vec{z}(\eta) \le \vec{z}(\eta_2)$, where η is the step size chosen on line 19. Thus we have $\|\vec{z}'\|_1 \le \|\vec{z}(\eta_2)\|_1 \le \epsilon j k$, where the last inequality is by the choice of η_2. □

A.2. Proof of Theorem 10

Proof of Theorem 10. By Lemma 3, the algorithm maintains the invariant that, at the end of phase j, we have $\|\vec{x}\|_1 \le \|\vec{z}\|_1 \le \epsilon j k$. Thus, at the end of the algorithm, we have $\|\vec{x}\|_1 \le k$.

Next, we show the approximation guarantee. Let $\vec{x}^{(0)} = \vec{0}$ and let $\vec{x}^{(j)}$ be the solution $\vec{x}$ at the end of phase j. We will show by induction on j that:

$$f(\vec{x}^{(j)}) \ge \epsilon j (1-\epsilon)^j f(\vec{x}^*) - 8 j \epsilon^2 f(\vec{x}^*)$$

The above inequality clearly hods for $j = 0$. Consider $j \ge 1$. By Theorem 4, we have

$$\begin{aligned} &f(\vec{x}^{(j)}) \ge f(\vec{x}^{(j-1)}) \\ &\qquad + (1-5\epsilon)\epsilon((1-\epsilon)^j f(\vec{x}^*) - f(\vec{x}^{(j)}) - 3\epsilon f(\vec{x}^*)) \end{aligned}$$

Hence

$f(\vec{x}^{(j)})$

$$
\begin{aligned}
&\geq (1-\epsilon) f(\vec{x}^{(j-1)}) \\
&+ (1-5\epsilon)\epsilon((1-\epsilon)^j f(\vec{x}^*) - 3\epsilon f(\vec{x}^*)) \\
&\geq (1-\epsilon) f(\vec{x}^{(j-1)}) + \epsilon(1-\epsilon)^j f(\vec{x}^*) - 8\epsilon^2 f(\vec{x}^*) \\
&\overset{(a)}{\geq} (1-\epsilon)\Big(\epsilon(j-1)(1-\epsilon)^{j-1} f(\vec{x}^*) - 8(j-1)\epsilon^2 f(\vec{x}^*)\Big) \\
&+ \epsilon(1-\epsilon)^j f(\vec{x}^*) - 8\epsilon^2 f(\vec{x}^*) \\
&\geq \epsilon j (1-\epsilon)^j f(\vec{x}^*) - 8j\epsilon^2 f(\vec{x}^*)
\end{aligned}
$$

where (a) is by the inductive hypothesis.

Thus it follows by induction that:

$$
f(\vec{x}^{(1/\epsilon)}) \geq ((1-\epsilon)^{1/\epsilon} - 8\epsilon) f(\vec{x}^*) \geq \left(\frac{1}{e} - O(\epsilon)\right) f(\vec{x}^*),
$$

as needed. □

B. Approximate Step Sizes

In this section, we show how to extend the idealized algorithm (Algorithm 1) and its analysis. In order to obtain an efficient algorithm, we find the step size η_1 approximately using t-ary search, as described below. The modified algorithm is given in Algorithm 2. On line 18 of Algorithm 2, the Θ notation hides a sufficiently small constant so that $\delta \leq \epsilon/N$, where N is the total number of iterations of the algorithm (as we discuss later in this section, the analysis of the number of iterations given in Theorem 11 still holds and thus $N = O(\log(n)\log(1/\epsilon)/\epsilon^3)$.)

Finding η_1 on line 20. As in the description of the algorithm, we let η_1^* be the maximum $\eta \in [0, \epsilon^2]$ such that $|S(\eta)| \geq (1-\epsilon)|S|$ and we let δ be the value on line 18. As shown in Lemma 14, for every $\eta \leq \eta'$, we have $S(\eta) \supseteq S(\eta')$, and thus $|S(\eta)|$ is non-increasing as a function of η. Note that $S(0) = S$ and thus $|S(0)| \geq (1-\epsilon)|S|$. We first check whether $|S(\epsilon^2)| \geq (1-\epsilon)|S|$; if so, we have $\eta_1^* = \epsilon^2$ and we return $\eta_1 = \epsilon^2$. Therefore we may assume that $|S(\epsilon^2)| < (1-\epsilon)|S|$ and thus $\eta_1^* \in [0, \epsilon^2)$. Starting with the interval $[0, \epsilon^2]$, we perform t-ary search, and we stop once we reach an interval $[a, b]$ of length at most δ. We return $\eta_1 = b$. Note that we have $\eta_1^* \leq \eta_1 \leq \eta_1^* + \delta$.

The arity of the t-ary search gives us different trade-offs between the number of parallel rounds and the total running time. The t-ary search takes $\log_t(\epsilon^2/\delta)$ parallel rounds and $t\log_t(\epsilon^2/\delta)$ evaluations of f and ∇f. If we use binary search ($t = 2$), the number of rounds is $\log_2(\epsilon^2/\delta) = O(\log\log n + \log(1/\epsilon))$ and the number of evaluations of f and ∇f is also $O(\log\log n + \log(1/\epsilon))$. If we take $t = \Theta(\log n/\epsilon)$, the number of rounds is $O(1)$ and the number of evaluations of f and ∇f is $O(\log n/\epsilon)$.

Next, we show how to extend the analysis given in Sections 4 and 5. We first note that the upper bound on the total

Algorithm 2 Algorithm for $\max_{\vec{x}\in[0,1]^n \colon \|\vec{x}\|_1 \leq k} f(\vec{x})$, where f is a non-negative DR-submodular function.

1: $M : f(\vec{x}^*) \leq M \leq (1+\epsilon) f(\vec{x}^*)$
2: $\vec{x} \leftarrow \vec{0}$
3: $\vec{z} \leftarrow \vec{0}$
4: **for** $j = 1$ to $1/\epsilon$ **do**
5: ⟨⟨ Start of phase j ⟩⟩
6: $\vec{x}_{\text{start}} \leftarrow \vec{x}$
7: $\vec{z}_{\text{start}} \leftarrow \vec{z}$
8: $v_{\text{start}} \leftarrow \frac{1}{k}(((1-\epsilon)^j - 2\epsilon)M - f(\vec{x}))$
9: $v \leftarrow v_{\text{start}}$
10: **while** $v > \epsilon v_{\text{start}}$ and $\|\vec{z}\|_1 < \epsilon jk$ **do**
11: $\vec{g} = (\vec{1} - \vec{z}) \circ \nabla f(\vec{z})$
12: $S = \{i \in [n] \colon \vec{g}_i \geq v$ and $\vec{z}_i \leq 1 - (1-\epsilon)^j$ and $\vec{z}_i - (\vec{z}_{\text{start}})_i < \epsilon(1 - (\vec{z}_{\text{start}})_i)\}$
13: **if** $S = \emptyset$ **then**
14: $v \leftarrow (1-\epsilon)v$
15: **else**
16: For a given $\eta \in [0, \epsilon^2]$, we define:

$$
\begin{aligned}
\vec{z}(\eta) &= \vec{z} + \eta(\vec{1} - \vec{z}) \circ \vec{1}_S \\
\vec{g}(\eta) &= (\vec{1} - \vec{z}(\eta)) \circ \nabla f(\vec{z}(\eta)) \\
S(\eta) &= \{i \in S \colon \vec{g}(\eta)_i \geq v\} \\
T(\eta) &= \{i \in S \colon \vec{g}(\eta)_i > 0\}
\end{aligned}
$$

17: ⟨⟨ $\delta \leq \epsilon/N$, where N is the total number of iterations of the algorithm ⟩⟩
18: Let $\delta = \Theta\left(\frac{\epsilon^4}{\log(n)\log(1/\epsilon)}\right)$
19: ⟨⟨ Let η_1^* be the maximum $\eta \in [0, \epsilon^2]$ such that $|S(\eta)| \geq (1-\epsilon)|S|$ ⟩⟩
20: Using t-ary search, find $\eta_1 \in [0, \epsilon^2]$ such that $\eta_1^* \leq \eta_1 \leq \eta_1^* + \delta$
21: ⟨⟨ $\eta_2 = \min\left\{\epsilon^2, \frac{\epsilon jk - \|\vec{z}\|_1}{|S| - \|\vec{z} \circ \vec{1}_S\|_1}\right\}$ ⟩⟩
22: Let η_2 be the maximum $\eta \in [0, \epsilon^2]$ such that $\|\vec{z}(\eta)\|_1 \leq \epsilon jk$
23: $\eta \leftarrow \min\{\eta_1, \eta_2\}$
24: $\vec{x} \leftarrow \vec{x} + \eta(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta-\delta)}$
25: $\vec{z} \leftarrow \vec{z} + \eta(\vec{1} - \vec{z}) \circ \vec{1}_S$
26: **if** $f(\vec{z}) > f(\vec{x})$ **then**
27: $\vec{x} \leftarrow \vec{z}$
28: **end if**
29: **end if**
30: **end while**
31: **end for**
32: **return** $\vec{x}$

number of iterations given in Theorem 11 still holds, since we have $\eta \geq \eta^* := \min\{\eta_1^*, \eta_2\}$ and $T(\eta - \delta) \supseteq T(\eta^*)$. Therefore it only remains to show that the approximate search only introduces an overall $O(\epsilon)$ additive error in the approximation guarantee. Since $\delta \leq \epsilon/N$, where N is the total number of iterations of the algorithm, it suffices to show that the error is $O(\delta) f(\vec{x}^*)$ in each iteration.

We start with the following lemma:

Lemma 15. *Let $\alpha, \beta \in \mathbb{R}$ and $\vec{u}, \vec{v} \in \mathbb{R}^n$. Suppose that $0 \leq \beta \leq \alpha \leq 1$, $\vec{u} \leq \vec{v} \leq \alpha \vec{1}$, and $\vec{v} - \vec{u} \leq \beta \vec{1}$. Then $f(\vec{u}) - f(\vec{v}) \leq \frac{\beta}{1-\alpha+\beta} f(\vec{u})$.*

Proof. For $t \geq 0$, let $\vec{w}(t) := \vec{u} + (\vec{v} - \vec{u})t$. The conditions in the lemma statement ensure that $\vec{w}((1-\alpha+\beta)/\beta) \leq \vec{1}$. Using that f is concave in non-negative directions and f is non-negative, we obtain:

$$\begin{aligned} f(\vec{v}) &= f(\vec{w}(1)) \\ &\geq \left(1 - \frac{\beta}{1-\alpha+\beta}\right) f(\vec{w}(0)) \\ &\quad + \frac{\beta}{1-\alpha+\beta} f\left(\vec{w}\left(\frac{1-\alpha+\beta}{\beta}\right)\right) \\ &\geq \left(1 - \frac{\beta}{1-\alpha+\beta}\right) f(\vec{w}(0)) \\ &= \left(1 - \frac{\beta}{1-\alpha+\beta}\right) f(\vec{u}) \end{aligned}$$

The lemma now follows by rearranging the above inequality. □

We now fix an iteration of the algorithm (an iteration of the inner while loop) that updates $\vec{x}$ and $\vec{z}$ on lines 24–27. Let $\vec{x}, \vec{z}$ denote the vectors right before the update on lines 24–27. We define:

$$\begin{aligned} \eta^* &:= \min\{\eta_1^*, \eta_2\} \\ \vec{x}' &:= \vec{x} + \eta(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta-\delta)} \\ \vec{a} &:= \vec{x} + (\eta - \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta-\delta)} \\ \vec{b} &:= \vec{x} + \eta^*(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta^*)} + (\eta - \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta-\delta) \setminus T(\eta^*)} \end{aligned}$$

Note that we have $\eta - \delta \leq \eta^* \leq \eta$ and thus it follows from Lemma 14 that $T(\eta - \delta) \supseteq T(\eta^*) \supseteq T(\eta)$.

We start by applying Lemma 15. Let $\vec{u} = \vec{a}$ and $\vec{v} = \vec{x}'$. We have $\vec{x}' \leq (1 - (1-\epsilon)^j + \epsilon^2)\vec{1}$, and thus we can take $\alpha = 1 - (1-\epsilon)^j + \epsilon^2 \leq 1 - (1-\epsilon)^{1/\epsilon} + \epsilon^2 \approx \frac{1}{e} + \epsilon^2$. We have $\vec{a} \leq \vec{x}'$ and $\vec{x}' - \vec{a} \leq \delta \vec{1}$, and thus we can take $\beta = \delta$. It follows from Lemma 15 that:

$$f(\vec{x}') - f(\vec{a}) \geq -O(\delta) f(\vec{a}) \geq -O(\delta) f(\vec{x}^*),$$

where in the second inequality we have used that $\vec{a}$ is feasible.

Next, we have:

$$\begin{aligned} f(\vec{a}) - f(\vec{x}) &\overset{(a)}{\geq} \langle \nabla f(\vec{a}), \vec{a} - \vec{x} \rangle \\ &= \langle \nabla f(\vec{a}), (\eta - \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta-\delta)} \rangle \\ &\overset{(b)}{\geq} \langle \nabla f(\vec{a}), (\eta - \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta^*)} \rangle \\ &\overset{(c)}{\geq} \langle \nabla f(\vec{z}(\eta^*)), (\eta - \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta^*)} \rangle \\ &\overset{(d)}{\geq} \langle \vec{g}(\eta^*), (\eta - \delta)\vec{1}_{T(\eta^*)} \rangle \end{aligned}$$

In (a), we used that f is concave in non-negative directions and $\vec{a} \geq \vec{x}$. We can show (b) as follows. As noted earlier, $T(\eta^*) \subseteq T(\eta - \delta)$. Since $\vec{a} \leq \vec{z}(\eta - \delta)$, we have $\nabla f(\vec{a}) \geq \nabla f(\vec{z}(\eta - \delta))$ by DR-submodularity, and thus $\nabla f(\vec{a})$ is non-negative on the coordinates in $T(\eta - \delta)$. In (c), we have used that $\vec{a} \leq \vec{z}(\eta - \delta) \leq \vec{z}(\eta^*)$ and thus $\nabla f(\vec{a}) \geq \nabla f(\vec{z}(\eta^*))$ by DR-submodularity. In (d), we used that $\nabla f(\vec{z}(\eta^*))$ is non-negative on the coordinates in $T(\eta^*)$ and $\vec{1} - \vec{x} \geq \vec{1} - \vec{z}(\eta^*) \geq \vec{0}$.

Similarly, we have:

$$\begin{aligned} f(\vec{b}) - f(\vec{a}) &\overset{(a)}{\geq} \langle \nabla f(\vec{b}), \vec{b} - \vec{a} \rangle \\ &\overset{(b)}{=} \langle \nabla f(\vec{b}), (\eta^* - \eta + \delta)(\vec{1} - \vec{x}) \circ \vec{1}_{T(\eta^*)} \rangle \\ &\overset{(c)}{\geq} \langle \vec{g}(\eta^*), (\eta^* - \eta + \delta)\vec{1}_{T(\eta^*)} \rangle \end{aligned}$$

In (a), we used that f is concave in non-negative directions and $\vec{b} \geq \vec{a}$. In (b), we used that $T(\eta - \delta) \supseteq T(\eta^*)$. We can show (c) as follows. Since $\eta - \delta \leq \eta^*$ and $T(\eta - \delta) \subseteq S$, we have $\vec{b} \leq \vec{z}(\eta^*)$. Thus $\nabla f(\vec{b}) \geq \nabla f(\vec{z}(\eta^*))$ by DR-submodularity and $\nabla f(\vec{b})$ is non-negative on the coordinates of $T(\eta^*)$.

By combining the inequalities above, we obtain:

$$f(\vec{x}') - f(\vec{x}) \geq \langle \vec{g}(\eta^*), \eta^* \vec{1}_{T(\eta^*)} \rangle - O(\delta) f(\vec{x}^*)$$

Thus we see that the gain obtained in the iteration is the one required by the proof of Theorem 4 apart from the additive loss of $O(\delta) f(\vec{x}^*)$. By propagating the additive loss through the proof of Theorem 4, we obtain a total loss of $O(\delta N) f(\vec{x}^*)$, where N is the total number of iterations. As noted above, $O(\delta N) = O(\epsilon)$, as needed.

C. DR-submodular Algorithms

In this section, we give the pseudocode of the sequential and parallel algorithms evaluated in our experiments. The sequential algorithm we used is the continuous greedy algorithm shown in Algorithm 3. The algorithm is a variant of the measured continuous greedy algorithm that was studied in previous works (Feldman et al., 2011; Chekuri et al., 2015; Bian et al., 2017). This variant obtains higher function value in practice, since it allows for the possibility of filling up more of the available budget, and this is what we observed in our experiments as well. The state of the art parallel algorithm for non-monotone DR-submodular maximization subject to a cardinality constraint is the algorithm of (Ene et al., 2019); Algorithm 4 gives the pseudocode of this algorithm specialized to a single cardinality constraint.

Algorithm 3 A variant of the measured continuous greedy algorithm for $\max_{\vec{x}\in[0,1]^n:\ \|\vec{x}\|_1\le k} f(\vec{x})$, where f is a non-negative DR-submodular function.

$\vec{x} \leftarrow \vec{0}$
$\langle\langle$ In our experiments, we used $\eta = \epsilon/n$ $\rangle\rangle$
$\eta \leftarrow \epsilon/n^3$
$T \leftarrow 1/\eta$
for $t = 1$ to T **do**
    $\vec{d} \leftarrow \arg\max_{\vec{z}\in[0,1]^n:\ \vec{z}\le\vec{1}-\vec{x}, \|\vec{z}\|_1\le k}\langle\nabla f(\vec{x}), \vec{z}\rangle$
    $\vec{x} \leftarrow \vec{x} + \eta\vec{d}$
end for
return $\vec{x}$

Algorithm 4 The algorithm of (Ene et al., 2019) specialized to a single cardinality constraint. The algorithm solves the problem $\max_{\vec{x}\in[0,1]^n:\ \|\vec{x}\|_1\le k} f(\vec{x})$, where f is a non-negative DR-submodular function. The algorithm takes as input a target value M such that $f(\vec{x}^*) \le M \le (1+\epsilon)f(\vec{x}^*)$.

$\eta \leftarrow \frac{\epsilon}{2\log(n+1)}$
$\vec{x} \leftarrow \frac{\epsilon}{n}\vec{1}$
$\vec{z} \leftarrow \vec{x}$
$\langle\langle$ MWU weights for the $(n+1)$ constraints $\vec{z}_i \le 1$ for all $i \in [n]$ and $\frac{1}{k}\langle\vec{z}, \vec{1}\rangle \le 1$ $\rangle\rangle$
$\vec{w}_i \leftarrow \exp(\vec{z}_i/\eta)$ for all $i \in [n]$
$\vec{w}_{n+1} \leftarrow \exp(\|\vec{z}\|_1/(\eta k))$
$t \leftarrow \eta\ln(\|\vec{w}\|_1)$
while $t < 1-\epsilon$ **do**
    $\lambda \leftarrow M\cdot(e^{-t} - 2\epsilon) - f(\vec{x})$
    $\vec{c} \leftarrow (\vec{1}-\vec{x}) \circ \nabla f((1+\eta)\vec{x}) \vee \vec{0}$
    $\vec{m}_i \leftarrow \left(1 - \lambda\cdot\frac{1}{\vec{c}_i}\cdot\frac{1}{\|\vec{w}\|_1}\left(\vec{w}_i + \frac{1}{k}\vec{w}_{n+1}\right)\right) \vee 0$ for all $i \in [n]$ with $\vec{c}_i \ne 0$, and $\vec{m}_i = 0$ if $\vec{c}_i = 0$
    $\vec{d} \leftarrow \eta\vec{x}\circ\vec{m}$
    if $\vec{d} = \vec{0}$ **then**
        break
    end if
    $\vec{x} \leftarrow \vec{x} + \vec{d}\circ(\vec{1}-\vec{x})$
    $\vec{z} \leftarrow \vec{z} + \vec{d}$
    $\langle\langle$ Update the weights $\rangle\rangle$
    $\vec{w}_i \leftarrow \exp(\vec{z}_i/\eta)$ for all $i \in [n]$
    $\vec{w}_{n+1} \leftarrow \exp(\|\vec{z}\|_1/(\eta k))$
end while
return $\vec{x}$