The Strong Secretary Conjecture is True for Linear Matroids

Kristóf Bérczi [a] Shaddin Dughmi [b] Vasilis Livanos [c] José A. Soto [d]
Victor Verdugo [e]

September 17th, 2026

Abstract

We prove a $1/e$ guarantee for the matroid secretary problem on linear matroids, therefore settling the strong secretary conjecture in this class of matroids. The result holds both when the matroid is known in advance and when a linear representation over a finite field is given online. In the known-matroid model, the result extends more generally to matroids admitting a finitary modular extension. Each element of a fixed optimal basis is selected with probability at least $1/e$.

Concurrent Discovery Disclosure: The proof of the main result in this manuscript was obtained in a conversation with ChatGPT-6 Astra on Tuesday, September 15, 2026 at 1:02 AM PDT. We then prepared this manuscript for public release, with the intent of uploading it on the morning of Thursday, September 17, 2026. In the early morning hours of September 17, while finalizing the submission, we discovered the manuscript of [1], uploaded on September 16, 2026, which contains the same result via an essentially identical approach. We are sharing our manuscript nonetheless in case our exposition is of independent utility to the community, and we hope this experience stimulates broader discussion about concurrent discovery in the AI era.

1 Introduction

The *Matroid Secretary Problem*, introduced by Babaioff, Immorlica, and Kleinberg [5], extends the classical secretary problem from selecting a single element to selecting an independent set of a matroid. Each element e has a fixed nonnegative weight $w(e)$, and the elements arrive in uniformly random order, with weights revealed upon arrival. The algorithm must irrevocably decide whether to accept the element while keeping the selected set independent. The goal is to obtain a constant fraction of $w(\mathrm{OPT}(E))$.

Let E denote the ground set. For every $Y \subseteq E$, let $\mathrm{OPT}(Y)$ denote the maximum-weight basis of the matroid restricted to Y. We assume without loss of generality that the weights are distinct, so $\mathrm{OPT}(Y)$ is unique. We say that an algorithm is *c-probability-competitive* if

$$\mathbb{P}\left[e \in \mathrm{ALG}\right] \geq c \quad \text{for every } e \in \mathrm{OPT}(E).$$

[a]HUN-REN–ELTE Egerváry Research Group, Department of Operations Research, Eötvös Loránd University, and HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary. kristof.berczi@ttk.elte.hu

[b]Department of Computer Science, University of Southern California, Los Angeles, USA. shaddin@usc.edu

[c]Department of Computer Science, University of Southern California, Los Angeles, USA. vas.livanos@gmail.com

[d]Department of Mathematical Engineering and Center for Mathematical Modeling, Universidad de Chile, Santiago, Chile. jsoto@dim.uchile.cl

[e]Institute for Mathematical and Computational Engineering, and Department of Industrial and Systems Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile. victor.verdugo@uc.cl

For nonnegative weights, this implies an expected weight of at least c times $w(\mathrm{OPT}(E))$. In the classical secretary problem, which corresponds to a rank-one matroid, the optimal probability guarantee is $1/e$.

A matroid M on ground set E is *linear* if there is a field $\mathbb{F}$ and a collection of vectors $\{v_e : e \in E\}$ over $\mathbb{F}$, called a *representation* of M over $\mathbb{F}$, such that $I \subseteq E$ is independent in M if and only if $\{v_e : e \in I\}$ is linearly independent. We prove that the optimal $1/e$ guarantee is achievable in the *online-representation model*, in which a representation of the matroid over a finite field is revealed online. The algorithm knows the number of elements and the underlying finite field in advance, but not the representation. Upon arrival, each element reveals its weight and its vector, expressed in a common coordinate system over that field. Thus, the algorithm learns the representation one vector at a time as the elements arrive. In particular, the algorithm does not know the matroid, or a representation of it, before the arrival process begins. The algorithm is ordinal, using the weights only through relative comparisons.

Theorem 1. *In the online-representation model over a finite field, there is an ordinal online algorithm that always selects an independent set* ALG *and satisfies*

$$\mathbb{P}\left[e \in \mathrm{ALG}\right] \geq 1/e \quad \textit{for every } e \in \mathrm{OPT}(E).$$

Consequently, $\mathbb{E}[w(\mathrm{ALG})] \geq w(\mathrm{OPT}(E))/e$.

The same guarantee is achievable for every linear matroid in the *known-matroid model*. Here the entire matroid is known before the arrival process, while only the weights are revealed online. By a theorem of Rado [36], every finite matroid representable over some field also admits a representation over a finite field. Therefore, ignoring computational and query complexity, we may compute such a representation and apply Theorem 1. In particular, no representation need be supplied as part of the input, and the result applies to matroids representable over an arbitrary field.

The same argument extends to known matroids that are restrictions of finitary modular matroids, whose lattices of flats satisfy the modular rank identity. In particular, it applies to the fully modular extendable matroids studied by Bérczi, Gehér, Imolay, Lovász, Padró and Schwarcz [9].

The main idea behind our algorithm is to control how much of each subspace is spanned by the accepted vectors. For every subspace L, we maintain an upper bound on the expected dimension of $L \cap \langle \mathrm{ALG} \rangle$, where $\langle \mathrm{ALG} \rangle$ denotes the linear span of the vectors selected by the algorithm.

We then choose acceptance probabilities that preserve these bounds while ensuring that an arriving element is accepted with a prescribed probability whenever it belongs to $\mathrm{OPT}(X)$, where X is the set of elements revealed so far. The probability is over both the preceding arrival order and the internal randomness of the algorithm.

The above requirements placed on acceptance probabilities lead to a large linear program (LP), and the main technical step in our proof is to establish its feasibility. Duality and uncrossing reduce the feasibility argument to chains of subspaces, for each of which we give an explicit feasible solution. Since our algorithm maintains distributions over possible values of $\langle \mathrm{ALG} \rangle$ and solves large LPs, we do not claim a polynomial-time implementation.

1.1 Related Work

The classical secretary problem was studied by Lindley [34] and Dynkin [23]. Kleinberg [31] studied the multiple-choice secretary problem, in which up to k elements may be selected. Later, Babaioff, Immorlica, and Kleinberg [5] introduced the Matroid Secretary Problem; see also the journal version with Kempe [4]. They conjectured that a constant expected-weight (or utility-competitive) guarantee is achievable for every matroid.

For general rank-r matroids, Babaioff et al. [4,5] obtained an $\Omega(1/\log r)$ expected-weight guarantee. This was improved to $\Omega(1/\sqrt{\log r})$ by Chakraborty and Lachish [12], and then to $\Omega(1/\log\log r)$ by Lachish [33]. Feldman, Svensson, and Zenklusen [24] later gave a simpler algorithm achieving the same guarantee.

A stronger notion is probability-competitiveness, which requires a guarantee for each element of the optimum. Bateni, Hajiaghayi, and Zadimoghaddam [8] obtained an $\Omega(1/\log^2 r)$ probability-competitive guarantee, and Soto, Turkieltaub, and Verdugo [40] improved it to $\Omega(1/\log r)$. The strong Matroid Secretary Conjecture asks whether the optimal rank-one guarantee $1/e$ can be achieved under this criterion for every matroid [40].

Singla's algorithm. Singla [38] recently resolved the original Matroid Secretary Conjecture: he gave an ordinal algorithm that, even in the unknown-matroid model, selects each element of the optimum with probability at least $1/4$. This answers the long-standing question of whether arbitrary matroids admit a constant-factor competitive guarantee, while the strong Matroid Secretary Conjecture remains open.

Singla's proof proceeds through a matroid version of the two-sided Game of Googol of Correa, Cristi, Epstein, and Soto [15]. The algorithm maintains a configuration initialized by a random sample and updates it as elements arrive via reversible operations. For any fixed element, even after conditioning on the configuration at the moment it is presented, its arrival time can be treated as uniform with respect to the remaining elements. This makes it possible to condition on the current configuration without changing the distribution of the arrival order.

A substantial line of work had previously established constant guarantees for important classes of matroids, including transversal [17, 30, 40], graphic [3, 5, 7, 10, 32, 40], laminar [10, 27, 28, 29, 35, 40], and regular matroids [18]. For several of these classes, the known guarantees were improved over a sequence of works using different approaches.

Linear-programming approaches. Linear programming has also been used to design and analyze secretary algorithms. Buchbinder, Jain, and Singh [11] gave a finite LP characterization of secretary algorithms, and Chan, Chen, and Jiang [13] developed a continuous LP approach for generalized secretary problems. Related LP formulations have since been used in secretary problems with advice and in sample-driven optimal stopping [14, 21].

Matroid information model. The information available about the matroid gives rise to two standard variants [16, 26]. In the *known-matroid model*, the ground set and the entire independence structure are available before the arrival process, and only the weights are revealed online. In the *unknown-matroid model*, the algorithm knows the number of elements in advance but can query independence only for subsets of elements that have already arrived. In particular, it has no information about dependencies involving unseen elements. The online-representation model considered in this paper provides a different type of information: the representation is not known in advance, but each arriving element reveals its vector in a common coordinate system.

For many special classes of matroids, the natural online model similarly reveals a representation of each arriving element. For graphic matroids, an arriving edge reveals its endpoints, while for transversal matroids an arriving element reveals its neighborhood in the underlying bipartite representation. This additional information has been important in obtaining constant guarantees. For example, the $1/e$ guarantee of Kesselheim, Radke, Tönnis, and Vöcking [30] for transversal matroids uses this bipartite representation. Similarly, the strongest previous guarantees for graphic matroids use the graph representation [7, 10], whereas Dütting, Paes Leme, Pál, and Patel [22] recently obtained a $1/36$ utility-competitive guarantee using only independence queries on arrived elements. Our online-representation model takes the analogous viewpoint for linear matroids.

Cristi, Dütting, Kleinberg, Paes Leme, and Patel [16] introduced online matroid embeddings, which map a matroid revealed online into a larger known matroid. They constructed embeddings for binary and laminar matroids and obtained reductions from unknown to known matroids with an arbitrarily small loss in the expected-weight guarantee. They also proved limitations of this embedding approach for general matroids.

Other approaches and models. Dughmi [19, 20] used duality to establish an equivalence between constant-competitive matroid secretary algorithms and random-order contention resolution for possibly correlated inputs. On the negative side, Bahrani, Beyhaghi, Singla, and Weinberg [6] proved limitations of broad classes of greedy and partition-based algorithms. Abdolazimi, Karlin, Klein, and Oveis Gharan [2] showed that the partition-property approach cannot give a constant guarantee for general binary matroids. These barriers concern the specified algorithmic approaches rather than the existence of constant-competitive algorithms.

Constant guarantees were also known under other online models. Soto [39] studied random assignment of weights to elements, and Oveis Gharan and Vondrák [26] extended this line to adversarial arrival orders. Santiago, Sergeev, and Zenklusen [37] obtained a constant guarantee for random assignment and random arrival order without knowing the matroid. Jaillet, Soto, and Zenklusen [29] considered the free-order model, where the algorithm chooses the order in which elements are inspected. For more general objectives, Feldman and Zenklusen [25] reduced the submodular matroid secretary problem to the linear-objective problem. As observed by Singla [38], his theorem therefore also yields a constant guarantee for nonnegative submodular objectives.

2 The Online Representation Algorithm

In this section, we describe our algorithm for linear matroids in the online-representation model. Let n be the number of elements, and fix a sample size $1 \leq k < n$. The first k arrivals form the sample and are rejected. In the classical rank-1 secretary algorithm, if the best element arrives in position $i > k$, then it is selected precisely when the best of the first $i-1$ elements appears in the sample, which happens with probability $k/(i-1)$. We seek the same conditional probability in a matroid: if X is the set of the first i observed elements and $e \in \mathrm{OPT}(X)$ arrives last, we want to accept e with probability $k/(i-1)$. As in the rank-1 case, choosing k appropriately and averaging over the arrival position yields a $1/e$ guarantee. We obtain these acceptance probabilities by solving linear programs that simultaneously enforce the desired acceptance probabilities and control, for every subspace L, the expected dimension of $L \cap \langle \mathrm{ALG} \rangle$.

2.1 The Dimension Invariant Inequality

When no confusion may arise, we identify an observed element with its representing vector. We use the notation $\langle \cdot \rangle$ for *linear span*. If X is a set of observed elements, then $\langle X \rangle$ denotes the span of their representing vectors, and $\langle e \rangle$ denotes the subspace spanned by the vector representing e. We write $W \leq V$ when W is a subspace of V, and for subspaces W_1, W_2 we write $W_1 + W_2 \coloneqq \langle W_1 \cup W_2 \rangle$.

For a set X and an element e, we use the shorthand $X - e \coloneqq X \setminus \{e\}$ and $X + e \coloneqq X \cup \{e\}$. When an element e arrives, let X be the set of elements observed up to and including e. We call e *improving* if $e \in \mathrm{OPT}(X)$.

After the sample, the *state* of the algorithm consists of an observed set X, an accepted set ALG, and the span $\langle \mathrm{ALG} \rangle$ of the accepted vectors. In addition to its state, our algorithm maintains some auxiliary information for every set $Y \subseteq X$. Specifically, it stores a distribution μ_Y supported on the subspaces of $\langle Y \rangle$, which is defined recursively. Later we show that, conditional on Y being the set of the first $|Y|$ arrivals, $\mu_Y(W)$ is exactly the probability that $\langle \mathrm{ALG} \rangle = W$ after the elements of Y have been processed.

Achieving the desired acceptance probabilities is more subtle than in the rank-1 case because whether an element can be accepted depends on the span of the elements accepted before it. Accordingly, the algorithm stores acceptance probabilities $p_Y(e, W)$, for every $e \in Y$ and $W \in \mathrm{supp}(\mu_{Y-e})$. The value $p_Y(e, W)$ is used when e arrives last among the elements of Y and $\langle \mathrm{ALG} \rangle = W$ immediately before its arrival.

To compute these probabilities, however, it is not enough to maintain information only for the sets that occur as prefixes of the realized arrival order: computing p_Y requires μ_{Y-e} for every $e \in Y$. We therefore maintain (μ_Y, p_Y) for every $Y \subseteq X$.

For every $Y \subseteq X$ with $|Y| \leq k$, if the elements of Y were the first $|Y|$ arrivals, the algorithm would reject all of them as part of the sample. Therefore, we set $\mu_Y(\{0\}) = 1$, $\mu_Y(W) = 0$ for every $W \neq \{0\}$, and $p_Y(e, W) = 0$ for every $e \in Y$ and $W \in \mathrm{supp}(\mu_{Y-e})$. Now suppose that $|X| = i > k$. We want every improving element to be accepted with probability $k/(i-1)$, conditional on X being the set of the first i arrivals.

The property that makes this possible is the following inequality, which we call the *Dimension Invariant Inequality.* After processing a fixed set of elements X with $|X| \geq k$, we require that

$$\mathbb{E}_{W \sim \mu_X}\left[\dim(W \cap L)\right] \leq \left(1 - \frac{k}{|X|}\right) \dim(L) \quad \text{for every subspace } L. \tag{1}$$

To see why we use the factor $1 - k/|X|$, let $|Y| = i > k$ and suppose that the last element e to arrive from Y is improving. Apply (1) to $Y - e$ and $L = \langle e \rangle$. Since e is not a loop, $\langle e \rangle$ has dimension one. Moreover, for a random subspace W distributed according to μ_{Y-e}, $\dim(W \cap \langle e \rangle)$ is 1 if $e \in W$ and 0 otherwise. Hence (1) implies that $\mathbb{P}\left[e \notin W\right] \geq k/(i-1)$. Thus, before e arrives, it can be added to ALG without violating independence with at least the desired probability. We now need to choose the acceptance probabilities so that every improving element is accepted with exactly this probability and the Dimension Invariant Inequality continues to hold.

Notice that it suffices to guarantee the Dimension Invariant Inequality for $L \leq \langle X \rangle$. This is because every $W \in \mathrm{supp}(\mu_X)$ satisfies $W \leq \langle X \rangle$, and thus $W \cap L = W \cap (L \cap \langle X \rangle)$. Therefore, applying (1) to $L \cap \langle X \rangle$ gives the inequality for an arbitrary ambient subspace L.

Fix the current observed set X, and let $Y \subseteq X$ with $|Y| = i > k$. We next define μ_Y from the distributions μ_{Y-e} and the acceptance probabilities $p_Y(e, W)$. Suppose that μ_A has already been computed for every $A \subsetneq Y$.

Suppose also that acceptance probabilities $p_Y(e, W) \in [0, 1]$ have been chosen for every $e \in Y$ and $W \in \mathrm{supp}(\mu_{Y-e})$, and write p for this collection. For every subspace $U \leq \langle Y \rangle$, define

$$\mu_Y(U) = \frac{1}{i} \sum_{e \in Y} \sum_{W \in \mathrm{supp}(\mu_{Y-e})} \mu_{Y-e}(W) \left[(1 - p_Y(e, W)) \mathbf{1}_{\{U = W\}} + p_Y(e, W) \mathbf{1}_{\{U = W + \langle e \rangle\}}\right]. \tag{2}$$

Since each μ_{Y-e} is a probability distribution and $p_Y(e, W) \in [0, 1]$, the values in (2) are nonnegative and sum to one over all subspaces $U \leq \langle Y \rangle$. Thus, μ_Y is a probability distribution on the subspaces of $\langle Y \rangle$.

For every $e \in Y$ and every subspace L, let

$$u_e(L) = \mathbb{E}_{W \sim \mu_{Y-e}}\left[\dim(W \cap L)\right]$$

and

$$\Delta_e(W, L) = \dim((W + \langle e \rangle) \cap L) - \dim(W \cap L).$$

Thus $u_e(L)$ is the expected dimension of $W \cap L$ under μ_{Y-e}, and $\Delta_e(W,L)$ is the change in this dimension when W is replaced by $W + \langle e \rangle$.

Define

$$g_p(L) := -(i-k)\dim(L) + \sum_{e \in Y} u_e(L) + \sum_{e \in Y} \mathbb{E}_{W \sim \mu_{Y-e}} \left[p_Y(e,W)\Delta_e(W,L) \right]. \tag{3}$$

The following lemma relates $g_p(L)$ to the Dimension Invariant Inequality for a subspace L under the distribution μ_Y.

Lemma 1. *Let Y be a set with $|Y| = i > k$. Suppose that, for every $e \in Y$, μ_{Y-e} is a probability distribution supported on the subspaces of $\langle Y - e \rangle$. Let*

$$p = (p_Y(e,W) : e \in Y,\ W \in \operatorname{supp}(\mu_{Y-e}))$$

satisfy $p_Y(e,W) \in [0,1]$ for every $e \in Y$ and $W \in \operatorname{supp}(\mu_{Y-e})$. Define μ_Y by (2) and g_p by (3). Then, for every subspace L,

$$\mathbb{E}_{U \sim \mu_Y} [\dim(U \cap L)] = \frac{1}{i}((i-k)\dim(L) + g_p(L)).$$

Consequently, for every subspace L, the Dimension Invariant Inequality for L holds if and only if $g_p(L) \le 0$. In particular, if

$$g_p(L) \le 0 \qquad \forall\, L \le \langle Y \rangle,$$

then μ_Y satisfies the Dimension Invariant Inequality.

Proof. By (2) and the definitions of $u_e(L)$ and $\Delta_e(W,L)$,

$$\begin{aligned}
\mathbb{E}_{U \sim \mu_Y} [\dim(U \cap L)] &= \frac{1}{i} \sum_{e \in Y} \left(u_e(L) + \mathbb{E}_{W \sim \mu_{Y-e}} [p_Y(e,W)\Delta_e(W,L)] \right) \\
&= \frac{1}{i} \left((i-k)\dim(L) + g_p(L) \right),
\end{aligned}$$

where the second equality follows from (3). Since $|Y| = i$, the Dimension Invariant Inequality for a fixed subspace L is

$$\mathbb{E}_{U \sim \mu_Y} [\dim(U \cap L)] \le \frac{i-k}{i} \dim(L).$$

By the identity above, this holds if and only if $g_p(L) \le 0$.

Therefore, if $g_p(L) \le 0$ for every $L \le \langle Y \rangle$, the Dimension Invariant Inequality holds for every $L \le \langle Y \rangle$. For an arbitrary subspace L, every $U \in \operatorname{supp}(\mu_Y)$ satisfies $U \cap L = U \cap (L \cap \langle Y \rangle)$. Applying the inequality to $L \cap \langle Y \rangle$ and using $\dim(L \cap \langle Y \rangle) \le \dim(L)$ gives the result for L. □

We will choose the acceptance probabilities so that $p_Y(e,W) = 0$ whenever $e \notin \mathrm{OPT}(Y)$, that is, whenever e would not be improving if it arrived last among the elements of Y. Under this restriction, (3) becomes

$$g_p(L) = -(i-k)\dim(L) + \sum_{e \in Y} u_e(L) + \sum_{e \in \mathrm{OPT}(Y)} \mathbb{E}_{W \sim \mu_{Y-e}} \left[p_Y(e,W)\Delta_e(W,L) \right].$$

2.2 Linear Programming and the Secretary Algorithm

Fix the current observed set X, and let $Y \subseteq X$ with $|Y| = i > k$. Suppose that μ_A has already been computed for every $A \subsetneq Y$. We choose the acceptance probabilities $p_Y(e, W)$, for $e \in Y$ and $W \in \mathrm{supp}(\mu_{Y-e})$, by solving the following feasibility LP, whose variables are precisely these quantities $p_Y(e, W)$. We write p for their collection.

$$
\begin{aligned}
\mathrm{LP}(Y): \quad & 0 \le p_Y(e, W) \le 1 && \forall\, e \in Y,\ W \in \mathrm{supp}(\mu_{Y-e}), \\
& p_Y(e, W) = 0 && \forall\, e \in Y,\ W \in \mathrm{supp}(\mu_{Y-e}) \text{ with } e \notin \mathrm{OPT}(Y) \text{ or } e \in W, \\
& \mathbb{E}_{W \sim \mu_{Y-e}}\left[p_Y(e, W)\right] = \frac{k}{i-1} && \forall\, e \in \mathrm{OPT}(Y), \\
& g_p(L) \le 0 && \forall\, L \le \langle Y \rangle\,.
\end{aligned}
$$

The first constraint ensures that the variables are valid probabilities. The second restricts acceptance to elements of $\mathrm{OPT}(Y)$ and assigns zero acceptance probability whenever $e \in W$, so that accepting e preserves independence. For each $e \in \mathrm{OPT}(Y)$, the third sets the expected value of $p_Y(e, W)$ over $W \sim \mu_{Y-e}$ equal to the target value $k/(i-1)$.

Finally, by Lemma 1, the last constraint is equivalent to requiring that the distribution μ_Y defined by (2) satisfies the Dimension Invariant Inequality. Since the representation is over a finite field, $\langle Y \rangle$ has only finitely many subspaces, so $\mathrm{LP}(Y)$ is a finite linear program.

The following lemma shows that $\mathrm{LP}(Y)$ is feasible whenever the distributions μ_{Y-e} satisfy the Dimension Invariant Inequality (1). Its proof is deferred to Section 3.

Lemma 2. *Let Y be a set with $|Y| = i > k$. Suppose that, for every $e \in Y$, the distribution μ_{Y-e} satisfies the Dimension Invariant Inequality* (1). *Then* $\mathrm{LP}(Y)$ *is feasible.*

We fix a priori a rule for choosing a feasible solution of $\mathrm{LP}(Y)$, requiring the chosen solution to be rational whenever a rational feasible solution exists, and denote the chosen solution by p_Y. We then compute μ_Y from p_Y and the distributions μ_{Y-e} using (2). This completes the recursive construction of (μ_Y, p_Y).

This fixed choice ensures that (μ_Y, p_Y) depends only on Y and the data revealed for its elements, and not on the order in which the elements of Y arrived. We will show in Lemma 3 that, throughout the execution, the relevant linear programs have rational coefficients and admit rational feasible solutions. Consequently, the rule above always selects rational acceptance probabilities.

Throughout the execution, the algorithm maintains a table indexed by subsets Y of the observed elements. For each such Y, the table stores the pair (μ_Y, p_Y), where

$$\mu_Y = \big(\mu_Y(U) : U \le \langle Y \rangle\big)$$

is a probability distribution on the subspaces of $\langle Y \rangle$, and

$$p_Y = \big(p_Y(f, W) : f \in Y,\ W \in \mathrm{supp}(\mu_{Y-f})\big)$$

is the collection of acceptance probabilities associated with Y. Once a pair (μ_Y, p_Y) is computed, it is stored and reused in all subsequent calls to UpdateDistributions. Thus, when a new element e arrives, the subroutine only needs to compute the pairs corresponding to sets Y that contain e.

We can now describe the online algorithm. After each arrival, the subroutine UpdateDistributions computes the new entries of this table. The online algorithm then uses $p_X(e, \langle \mathrm{ALG} \rangle)$ to decide whether to accept the arriving element e. Since these probabilities are rational, the corresponding acceptance decisions can be implemented exactly.

Algorithm 1. Linear Matroid Secretary

Input: Number of elements n and sample size k

$X \leftarrow \emptyset$; $\text{ALG} \leftarrow \emptyset$
foreach *arriving element* e **do**
  Observe e's vector and its relative weight rank among the observed elements
  $X \leftarrow X + e$
  `UpdateDistributions`(X, e)
  if $|X| > k$ **then**
    Accept e with probability $p_X(e, \langle \text{ALG} \rangle)$
    if e *is accepted* **then**
      $\text{ALG} \leftarrow \text{ALG} + e$
return ALG

Algorithm 2. UpdateDistributions(X, e)

Input: The current observed set X, the newly arrived element e, together with the revealed vectors and relative weight order of the elements of X

foreach $Y \subseteq X$ *with* $e \in Y$, *in increasing order of* $|Y|$ **do**
  if $|Y| \leq k$ **then**
    Set μ_Y to the point mass at $\{0\}$ and $p_Y = 0$
  else
    Compute $\text{OPT}(Y)$ and enumerate the subspaces of $\langle Y \rangle$
    Solve $\text{LP}(Y)$ using $(\mu_{Y-f})_{f \in Y}$, and let p_Y be the chosen solution
    Compute μ_Y from (2)
  Store (μ_Y, p_Y)

Lemma 3. *Consider Algorithm 1 together with the subroutine* UpdateDistributions. *For every fixed set X, conditional on X being the set of the first $|X|$ arrivals, the algorithm is well-defined up to the processing of X, and the following properties hold:*

(i) μ_X *is the distribution of* $\langle \text{ALG} \rangle$ *after the elements of X have been processed.*

(ii) ALG *is always independent.*

(iii) If $|X| \geq k$, then μ_X satisfies the Dimension Invariant Inequality (1).

(iv) If $|X| = i > k$ and $e \in \text{OPT}(X)$, then, conditional on e arriving last among the elements of X, the algorithm accepts e with probability $k/(i-1)$.

(v) All entries of μ_X and p_X are rational.

Proof. We prove the well-definedness of the algorithm and the five statements simultaneously by induction on $|X|$.

If $|X| \leq k$, all elements belong to the sample and are rejected. Hence $\langle \text{ALG} \rangle = \{0\}$ and μ_X is the point mass at $\{0\}$. Moreover, $p_X = 0$, so all entries of μ_X and p_X are rational, proving (v). Thus, μ_X is the distribution of $\langle \text{ALG} \rangle$, proving (i), and ALG is independent, proving (ii). When $|X| = k$, (1) holds with equality, proving (iii).

Let $|X| = i > k$, and let a be the last element of X to arrive. When UpdateDistributions(X, a) processes X, every μ_{X-f} with $f \in X$ has already been computed: if $f = a$, it was computed before the arrival of a, while if $f \neq a$, the set $X - f$ has size $i - 1$ and, since sets are processed in increasing order of cardinality, it is processed earlier in the same call to UpdateDistributions. For $f \neq a$, the set $X - f$ need not be the set of the first $i-1$ arrivals. However, by the fixed choice rule, (μ_{X-f}, p_{X-f}) depends only on $X - f$ and the data revealed for its elements, so the induction hypothesis still applies. Hence, by (iii) and (v), each μ_{X-f} satisfies the Dimension Invariant Inequality and has rational entries. Lemma 2 therefore implies that $\mathrm{LP}(X)$ is feasible. Moreover, $\mathrm{LP}(X)$ has rational coefficients, so it admits a rational feasible solution. By our fixed choice rule, p_X has rational entries, and (2) then implies that μ_X has rational entries. Thus, p_X and μ_X are well-defined, and (v) holds.

Fix $e \in X$ and condition on e being the last element to arrive among the elements of X. The relative order of the elements in $X - e$ is then uniformly random. Thus, by induction, immediately before the arrival of e, $\langle \mathrm{ALG} \rangle$ has distribution μ_{X-e}.

Now fix $W \in \mathrm{supp}(\mu_{X-e})$. Conditional on $\langle \mathrm{ALG} \rangle = W$ immediately before the arrival of e, the algorithm rejects e with probability $1 - p_X(e, W)$, in which case $\langle \mathrm{ALG} \rangle$ remains equal to W, and accepts e with probability $p_X(e, W)$, in which case $\langle \mathrm{ALG} \rangle$ becomes $W + \langle e \rangle$. Therefore, conditional on e being the last element to arrive, the probability that $\langle \mathrm{ALG} \rangle = U$ after processing e is

$$\sum_{W \in \mathrm{supp}(\mu_{X-e})} \mu_{X-e}(W) \left[(1 - p_X(e,W)) \mathbf{1}_{\{U = W\}} + p_X(e,W) \mathbf{1}_{\{U = W + \langle e \rangle\}} \right].$$

Every element of X is equally likely to arrive last. Averaging the expression above over $e \in X$ gives exactly (2) with $Y = X$. Hence μ_X is the distribution of $\langle \mathrm{ALG} \rangle$ after processing the elements of X. This proves (i).

We next prove that ALG remains independent. Let e be the last element of X to arrive. By induction, ALG is independent before e arrives. If e is accepted, the second constraint of $\mathrm{LP}(X)$ implies that $e \notin \langle \mathrm{ALG} \rangle$. Therefore, $\mathrm{ALG} + e$ is independent. This proves (ii).

For the Dimension Invariant Inequality, the last constraint of $\mathrm{LP}(X)$ gives

$$g_{p_X}(L) \leq 0 \qquad \forall\, L \leq \langle X \rangle .$$

By Lemma 1, μ_X satisfies (1) for every $L \leq \langle X \rangle$. As observed in Section 2.1, this implies the inequality for every subspace L. This proves (iii).

Finally, fix $e \in \mathrm{OPT}(X)$ and condition on e being the last element to arrive among the elements of X. The relative order of the elements in $X - e$ is uniformly random. By induction, immediately before the arrival of e, $\langle \mathrm{ALG} \rangle$ has distribution μ_{X-e}. When $\langle \mathrm{ALG} \rangle = W$, the algorithm accepts e with probability $p_X(e, W)$, using fresh randomness. Therefore,

$$\mathbb{P}\left[e \in \mathrm{ALG} \mid e \text{ arrives last among } X\right] = \mathbb{E}_{W \sim \mu_{X-e}}\left[p_X(e,W)\right] = \frac{k}{i-1}, \tag{4}$$

where the last equality follows from the third constraint of $\mathrm{LP}(X)$. This proves (iv).

Observe that this does not mean that $p_X(e, W) = k/(i-1)$ for every $W \in \mathrm{supp}(\mu_{X-e})$. The third constraint only requires this equality after averaging over $W \sim \mu_{X-e}$. □

2.3 Proof of Theorem 1

We are now ready to prove that every element of $\mathrm{OPT}(E)$ is selected with probability at least $1/e$.

If $n = 1$ or $n = 2$, we do not run Algorithm 1. Instead, we accept the first non-loop element that arrives. Every element of $\mathrm{OPT}(E)$ is then selected with probability at least $1/n \geq 1/e$.

Hence assume that $n \geq 3$. We run Algorithm 1 with sample size $k = \lfloor n/e \rfloor$. Fix an element $e \in \mathrm{OPT}(E)$. For every $X \subseteq E$ containing e, we have $e \in \mathrm{OPT}(X)$. Indeed, since $e \in \mathrm{OPT}(E)$, it is not spanned by the elements of E of greater weight, and therefore it is not spanned by the elements of X of greater weight.

Fix $i > k$ and condition on e arriving in position i and on X being the set of the first i arrivals. Then e arrives last among the elements of X and $e \in \mathrm{OPT}(X)$. By Lemma 3(iv), the conditional probability that e is accepted is $k/(i-1)$. Since the value $k/(i-1)$ does not depend on X, we may drop the conditioning on X. Thus, conditional on e arriving in position i, it is accepted with probability $k/(i-1)$. Since the arrival position of e is uniform,

$$\mathbb{P}\left[e \in \mathrm{ALG}\right] = \frac{1}{n} \sum_{i=k+1}^{n} \frac{k}{i-1} = \frac{k}{n} \sum_{j=k}^{n-1} \frac{1}{j}. \tag{5}$$

Since $1/x$ is decreasing,

$$k \sum_{j=k}^{n-1} \frac{1}{j} = 1 + k \sum_{j=k+1}^{n-1} \frac{1}{j} \geq 1 + k \int_{k+1}^{n} \frac{1}{x}\, dx = 1 + k \ln \frac{n}{k+1}.$$

It remains to show that the last expression is at least n/e.

Since $k = \lfloor n/e \rfloor$, we have $ek \leq n < e(k+1)$. Let $g(x) = 1 + k\ln(x/(k+1)) - x/e$. On $[ek, e(k+1)]$ we have $g'(x) = k/x - 1/e \leq 0$, while $g(e(k+1)) = 0$. Hence $g(n) \geq 0$, and thus $1 + k\ln(n/(k+1)) \geq n/e$. Combining this with (5) gives $\mathbb{P}\left[e \in \mathrm{ALG}\right] \geq 1/e$. This proves Theorem 1.

3 LP Feasibility via Uncrossing

Throughout this section, fix a set X with $|X| = i > k$, together with a family of distributions

$$\boldsymbol{\mu} = \left(\mu_{X-e}\right)_{e \in X},$$

where each μ_{X-e} satisfies the Dimension Invariant Inequality (1).

Let $P = P(X, \boldsymbol{\mu})$ be the polytope obtained from $\mathrm{LP}(X)$ by dropping the constraints $g_p(L) \leq 0$. Since X and $\boldsymbol{\mu}$ remain fixed throughout this section, we simply write P.

We first show that P is non-empty. We then show that there exists a point $p \in P$ that also satisfies $g_p(L) \leq 0$ for every $L \leq \langle X \rangle$. For each $e \in \mathrm{OPT}(X)$, the Dimension Invariant Inequality applied to $\langle e \rangle$ gives

$$\mathbb{P}_{W \sim \mu_{X-e}}\left[e \notin W\right] \geq \frac{k}{i-1}. \tag{6}$$

For such an e, set $p(e, W) = 0$ when $e \in W$, and otherwise set

$$p(e, W) = \frac{k}{(i-1)\mathbb{P}_{W' \sim \mu_{X-e}}\left[e \notin W'\right]}.$$

The inequality above ensures that $p(e, W) \leq 1$, while by construction $\mathbb{E}_{W \sim \mu_{X-e}}\left[p(e, W)\right] = k/(i-1)$. For $e \notin \mathrm{OPT}(X)$, set $p(e, W) = 0$ for every W. Thus $p \in P$, so P is non-empty.

It remains to find a point $p \in P$ such that $g_p(L) \leq 0$ for every $L \leq \langle X \rangle$. We begin with a structural property of g_p.

Fix $p \in P$ and consider the following one-step experiment. Choose $e \in X$ uniformly, sample W according to μ_{X-e}, and accept e with probability $p(e, W)$. Let Z be the resulting subspace. By Lemma 1, applied with $Y = X$, we have

$$g_p(L) = -(i-k)\dim(L) + i\, \mathbb{E}_p\left[\dim(Z \cap L)\right].$$

We claim that g_p is supermodular on the lattice of subspaces. Namely, for any $L, H \leq \langle X \rangle$,

$$g_p(L) + g_p(H) \leq g_p(L \cap H) + g_p(L + H).$$

Indeed, for every fixed subspace Z,

$$\dim(Z \cap L) + \dim(Z \cap H) = \dim((Z \cap L) + (Z \cap H)) + \dim(Z \cap L \cap H).$$

Together with

$$(Z \cap L) + (Z \cap H) \leq Z \cap (L + H),$$

this shows that $L \mapsto \dim(Z \cap L)$ is supermodular. Taking expectations preserves supermodularity, while $\dim(L)$ is modular. Therefore, g_p is supermodular.

We now use this supermodularity to reduce the problem to chains of subspaces.

Lemma 4. *Suppose that for every chain*

$$\mathcal{C} = \{L_1, \ldots, L_s\}, \qquad L_1 < \cdots < L_s \leq \langle X \rangle\,,$$

there exists a $p \in P$ such that $g_p(L_j) \leq 0$ for every $j \in \{1, \ldots, s\}$. Then there exists a single $p \in P$ such that $g_p(L) \leq 0$ for every $L \leq \langle X \rangle$.

Proof. Since the field is finite, $\langle X \rangle$ has finitely many subspaces. Let Λ be the probability simplex on these subspaces. By the von Neumann minimax theorem,

$$v := \min_{p \in P} \max_{L \leq \langle X \rangle} g_p(L) = \max_{\lambda \in \Lambda} \min_{p \in P} \sum_L \lambda_L g_p(L). \tag{7}$$

Thus, the existence of a $p \in P$ such that $g_p(L) \leq 0$ for every $L \leq \langle X \rangle$ is equivalent to $v \leq 0$.

Among the maximizers on the right-hand side of (7), choose one that maximizes $\sum_L \lambda_L (\dim(L))^2$. Suppose that two incomparable subspaces L, H have positive mass in λ, and let $t = \min\{\lambda_L, \lambda_H\}$. Move mass t from each of L and H to $L \cap H$ and $L + H$ and let λ' denote the resulting point in Λ. By the supermodularity of g_p, we have $g_p(L) + g_p(H) \leq g_p(L \cap H) + g_p(L + H)$ for every $p \in P$. Hence the weighted sum does not decrease for any p, and neither does its minimum over P. Since the original λ is optimal, the new point λ' is also optimal.

On the other hand, $\sum_L \lambda_L (\dim(L))^2$ increases by $2t(\dim(L) - \dim(L \cap H))(\dim(H) - \dim(L \cap H)) > 0$, a contradiction. Therefore, the chosen optimal λ is supported on a chain $\mathcal{C}$. By the hypothesis of the lemma, there exists a $p \in P$ with $g_p(L) \leq 0$ for every $L \in \mathcal{C}$. Thus, $v \leq \sum_L \lambda_L g_p(L) \leq 0$. ☐

By Lemma 4, it is enough to fix an arbitrary chain $L_1 < \cdots < L_s \leq \langle X \rangle$ and construct a point $p \in P$ such that $g_p(L_j) \leq 0$ for every $j \in \{1, \ldots, s\}$.

Fix $e \in \mathrm{OPT}(X)$ and let

$$\mathcal{F}_e = \{W \in \operatorname{supp}(\mu_{X-e}) \mid e \notin W\}\,.$$

These are precisely the states in which e can be accepted.

Lemma 5. *For every $W \in \mathcal{F}_e$ and every subspace L, we have*

$$\Delta_e(W, L) = \mathbf{1}\left[e \in W + L\right].$$

Proof. Recall that $\Delta_e(W,L) = \dim((W+\langle e\rangle)\cap L) - \dim(W\cap L)$. Applying the equality $\dim(A\cap B) = \dim(A) + \dim(B) - \dim(A+B)$ to this expression twice, once for $(A,B) = (W+\langle e\rangle, L)$ and the other for $(A,B) = (W,L)$, yields

$$\begin{aligned}\dim((W+\langle e\rangle)\cap L) - \dim(W\cap L) &= (\dim(W+\langle e\rangle) + \dim(L) - \dim(W+\langle e\rangle + L))\\ &\quad - (\dim(W) + \dim(L) - \dim(W+L))\\ &= 1 - (\dim(W+\langle e\rangle + L) - \dim(W+L))\\ &= 1 - \mathbf{1}[e \notin W+L]\\ &= \mathbf{1}[e \in W+L],\end{aligned}$$

where the second equality follows from the fact that, since $W \in \mathcal{F}_e$, we have $e \notin W$ and thus $\dim(W+\langle e\rangle) - \dim(W) = 1$. $\square$

For each j, let

$$H_j = \{W \in \mathcal{F}_e \mid e \notin W + L_j\}.$$

Since $L_1 < \cdots < L_s$, we have $H_s \subseteq \cdots \subseteq H_1$. Order the states in $\mathcal{F}_e$ as $W_1, \ldots, W_t$ so that every H_j is an initial segment: first the states in H_s, then those in $H_{s-1} \setminus H_s$, and so on. For each W_ℓ, let $a_\ell = \mu_{X-e}(W_\ell)$ and assign acceptance probability greedily according to this order:

$$p(e, W_\ell) = \min\left\{1, \frac{\left[\frac{k}{i-1} - \sum_{h<\ell} a_h\right]^+}{a_\ell}\right\}, \tag{8}$$

where $[x]^+ := \max\{x, 0\}$. For the remaining states W, set $p(e,W) = 0$.

Recall that $e \in \mathrm{OPT}(X)$ is fixed throughout this construction. By (6) and the definition of $\mathcal{F}_e$,

$$\mu_{X-e}(\mathcal{F}_e) = \mathbb{P}_{W\sim\mu_{X-e}}[e \notin W] \geq \frac{k}{i-1}.$$

Hence the greedy assignment in (8) assigns total acceptance probability exactly $k/(i-1)$ to e.

We perform the construction above separately for each $e \in \mathrm{OPT}(X)$, and set $p(e,W) = 0$ for every $e \notin \mathrm{OPT}(X)$. This defines all the coordinates of a point p. Since, for each $e \in \mathrm{OPT}(X)$, the construction assigns total acceptance probability $k/(i-1)$ and sets $p(e,W) = 0$ whenever $e \in W$, we have $p \in P$.

Now fix $e \in \mathrm{OPT}(X)$. Since every H_j is an initial segment,

$$\sum_{W\in H_j} \mu_{X-e}(W) p(e,W) = \min\left\{\frac{k}{i-1}, \mu_{X-e}(H_j)\right\}. \tag{9}$$

For every L in the chain, let $q_e^p(L) = \mathbb{E}_{W\sim\mu_{X-e}}[p(e,W)\Delta_e(W,L)]$. Notice that

$$\begin{aligned}
q_e^p(L) &= \sum_{W \in \mathcal{F}_e} \mu_{X-e}(W) p(e,W) \Delta_e(W,L) \\
&= \sum_{W \in \mathcal{F}_e : e \in W+L} \mu_{X-e}(W) p(e,W) \\
&= \sum_{W \in \mathcal{F}_e} \mu_{X-e}(W) p(e,W) - \sum_{W \in \mathcal{F}_e : e \notin W+L} \mu_{X-e}(W) p(e,W) \\
&= \frac{k}{i-1} - \min\left\{ \frac{k}{i-1}, \mu_{X-e}(\{W \in \mathcal{F}_e \mid e \notin W+L\}) \right\} \\
&= \left(\frac{k}{i-1} - \mu_{X-e}(\{W \in \mathcal{F}_e \mid e \notin W+L\}) \right)^+ \\
&= \left(\frac{k}{i-1} - \mathbb{P}_{W \sim \mu_{X-e}} \left[e \notin W+L \right] \right)^+, \qquad (10)
\end{aligned}$$

where the second equality follows from Lemma 5, and the fourth follows from (9).

Let $d = \dim(L)$ and $y = u_e(L) + k/(i-1) - \mathbb{P}_{W \sim \mu_{X-e}} [e \notin W+L]$. We distinguish between the cases $e \notin L$ and $e \in L$. Suppose first that $e \notin L$. Then, we have

$$\begin{aligned}
y &= u_e(L) + \frac{k}{i-1} - 1 + 1 - \mathbb{P}_{W \sim \mu_{X-e}} [e \notin W+L] \\
&= u_e(L) + \frac{k}{i-1} - 1 + \mathbb{P}_{W \sim \mu_{X-e}} [e \in W+L] \\
&= u_e(L) - \left(1 - \frac{k}{i-1}\right) + \mathbb{E}_{W \sim \mu_{X-e}} [\dim(W \cap (L + \langle e \rangle)) - \dim(W \cap L)] \\
&= u_e(L) - \left(1 - \frac{k}{i-1}\right) + u_e(L + \langle e \rangle) - u_e(L) \\
&= u_e(L + \langle e \rangle) - \left(1 - \frac{k}{i-1}\right). \qquad (11)
\end{aligned}$$

Thus, using (10), (11) and the fact that $x + [y-x]^+ = \max\{x, y\}$ for $x = u_e(L)$, we have

$$\begin{aligned}
u_e(L) + q_e^p(L) &= u_e(L) + \left[\frac{k}{i-1} - \mathbb{P}_{W \sim \mu_{X-e}} [e \notin W+L] \right]^+ \\
&= \max\left\{ u_e(L), u_e(L + \langle e \rangle) - \left(1 - \frac{k}{i-1}\right) \right\}. \qquad (12)
\end{aligned}$$

The Dimension Invariant Inequality (1) for $X - e$, when applied to L and $L + \langle e \rangle$, gives

$$u_e(L) \le \left(1 - \frac{k}{i-1}\right) d \quad \text{and} \quad u_e(L + \langle e \rangle) \le \left(1 - \frac{k}{i-1}\right) \dim(L + \langle e \rangle) = \left(1 - \frac{k}{i-1}\right)(d+1).$$

Combining this with (12) yields

$$u_e(L) + q_e^p(L) = \max\left\{ u_e(L), u_e(L + \langle e \rangle) - \left(1 - \frac{k}{i-1}\right) \right\} \le \left(1 - \frac{k}{i-1}\right) d. \qquad (13)$$

If, on the other hand, $e \in L$, every feasible acceptance increases $\dim(\langle \mathrm{ALG} \rangle \cap L)$ by exactly one. Since the construction assigns total acceptance probability $k/(i-1)$ to e, we have $q_e^p(L) = k/(i-1)$. This, together with the bound on $u_e(L)$ above and (13), gives, in both cases,

$$u_e(L) + q_e^p(L) \le \left(1 - \frac{k}{i-1}\right) d + \frac{k}{i-1} \cdot \mathbf{1}[e \in L].$$

For $e \notin \mathrm{OPT}(X)$, we have $p(e, W) = 0$ for every W, and the Dimension Invariant Inequality gives

$$u_e(L) \leq \left(1 - \frac{k}{i-1}\right) d.$$

Since $\mathrm{OPT}(X)$ is independent, $|\mathrm{OPT}(X) \cap L| \leq d$. Summing these bounds over all elements in X gives

$$\begin{aligned}
\sum_{e \in X} u_e(L) + \sum_{e' \in \mathrm{OPT}(X)} q^p_{e'}(L) &\leq i \left(1 - \frac{k}{i-1}\right) d + \frac{k}{i-1} \cdot |\mathrm{OPT}(X) \cap L| \\
&\leq \left(i \left(1 - \frac{k}{i-1}\right) + \frac{k}{i-1}\right) d \\
&= (i-k)d.
\end{aligned}$$

Therefore $g_p(L) \leq 0$ for every L in the chain. Lemma 4 gives a single $p \in P$ satisfying all the constraints of $\mathrm{LP}(X)$. This proves Lemma 2.

4 Beyond Linear Matroids: Finitary Modular Extensions

We now turn to the known-matroid model. For linear matroids, the result follows from Theorem 1, since every finite linear matroid admits a representation over a finite field. We then show that the same guarantee holds more generally for matroids admitting a finitary modular extension.

Corollary 1. *In the known-matroid model, every linear matroid admits an ordinal* $1/e$*-probability-competitive algorithm.*

Proof. By a theorem of Rado [36], every finite matroid representable over some field is representable over a finite field. Since the matroid is known and we allow unbounded computational power, we may enumerate finite fields and matrices with columns indexed by E, checking each candidate against the independence structure of the matroid, until we find such a representation. We do not require this search to be efficient. Once a finite-field representation is fixed, Theorem 1 applies. □

Thus the representation need not be supplied as part of the input in the known-matroid model. In particular, the corollary applies to every finite linear matroid, regardless of the field over which it is initially known to be representable.

For a matroid N, we write cl_N for its *closure operator*. A set F is a *flat* if $\mathrm{cl}_N(F) = F$. For two flats F_1, F_2, their *meet* is $F_1 \cap F_2$ and their *join* is $F_1 \vee F_2 := \mathrm{cl}_N(F_1 \cup F_2)$. The pair F_1, F_2 is *modular* if $r_N(F_1) + r_N(F_2) = r_N(F_1 \cap F_2) + r_N(F_1 \vee F_2)$. The matroid is *modular* if every pair of its flats is modular. A possibly infinite matroid is *finitary* if every circuit is finite. We say that a finite matroid M on E admits a *finitary modular extension* if there is a finitary modular matroid M' on a ground set containing E such that $M'|E = M$.

The proof of Theorem 1 in fact uses less than a vector-space representation. The arguments only require the modular rank identities for flats. This yields the following more general consequence.

Corollary 2. *Let M be a known matroid admitting a finitary modular extension. Then M admits an ordinal* $1/e$*-probability-competitive algorithm.*

Proof. Suppose first that an extension M' is fixed and available to the algorithm. We work with the flats of M' contained in $\mathrm{cl}_{M'}(E)$, ordered by inclusion. Since E is finite, each such flat has rank at most $r_{M'}(E) < \infty$.

The construction of Section 2 carries over by replacing vector spans with closures in M', sums of subspaces with joins of flats, and dimension with $r_{M'}$. Thus, the state corresponding to an accepted set A is $\mathrm{cl}_{M'}(A)$, the initial state is $\mathrm{cl}_{M'}(\emptyset)$, and the Dimension Invariant Inequality becomes

$$\mathbb{E}_{W\sim\mu_X}\left[r_{M'}(W\cap L)\right] \le \left(1-\frac{k}{|X|}\right)r_{M'}(L)$$

for every flat $L \subseteq \mathrm{cl}_{M'}(E)$.

The rank identities in the feasibility proof remain valid in this setting. Indeed, modularity gives $r_{M'}(F_1) + r_{M'}(F_2) = r_{M'}(F_1 \cap F_2) + r_{M'}(F_1 \vee F_2)$ for any two flats $F_1, F_2 \subseteq \mathrm{cl}_{M'}(E)$. In particular, for every fixed flat W, the function assigning $r_{M'}(W \cap F)$ to a flat F is supermodular, since $(W \cap F_1) \vee (W \cap F_2) \subseteq W \cap (F_1 \vee F_2)$. Thus, the supermodularity and uncrossing arguments from Section 3 continue to hold. For the fixed-chain construction, let L be a flat in the fixed chain. If $e \notin W$, modularity gives $r_{M'}(\mathrm{cl}_{M'}(W \cup \{e\}) \cap L) - r_{M'}(W \cap L) = \mathbf{1}\left[e \in \mathrm{cl}_{M'}(W \cup L)\right]$, and, if $e \notin L$, it also gives $r_{M'}(W \cap \mathrm{cl}_{M'}(L \cup \{e\})) - r_{M'}(W \cap L) = \mathbf{1}\left[e \in \mathrm{cl}_{M'}(W \cup L)\right]$. Moreover, since $\mathrm{OPT}(X)$ is independent in $M = M'|E$, we have $|\mathrm{OPT}(X) \cap L| \le r_{M'}(L)$. These are exactly the properties used in the fixed-chain construction and in the proof of (13). Hence the construction and its proof carry over with closures and rank in place of spans and dimension.

There is, however, one difference from the finite-field setting. There may be infinitely many flats contained in $\mathrm{cl}_{M'}(E)$, so we must check that the acceptance LP still has only finitely many essential constraints. For a fixed observed set X, every state that can occur is of the form $\mathrm{cl}_{M'}(Z)$ for some $Z \subseteq X$. Hence the set

$$\mathcal{S}_X = \{\mathrm{cl}_{M'}(Z) \colon Z \subseteq X\}$$

is finite. For a flat $L \subseteq \mathrm{cl}_{M'}(E)$, define its signature relative to X by

$$\sigma_X(L) = \Big(r_{M'}(L), \left(r_{M'}(W \cap L)\right)_{W \in \mathcal{S}_X}\Big).$$

Every quantity appearing in the constraint $g_p(L) \le 0$ is determined by this signature. Indeed, both a predecessor state W and the state obtained after accepting an element belong to $\mathcal{S}_X$. Since all ranks in $\sigma_X(L)$ are integers between 0 and $r_{M'}(E)$, only finitely many signatures can occur. Thus there are only finitely many distinct constraints.

The chain-reduction argument can now be applied to these finitely many constraint types. Choose one representative flat for each signature. If two representatives carrying positive mass in an optimal minimax solution are incomparable, move their mass, exactly as in the proof of Lemma 4, to the chosen representatives of the signatures of their intersection and join. Flats with the same signature define the same constraint and have the same rank, so supermodularity ensures that the minimax objective does not decrease, while modularity gives the same strict increase in the secondary objective. Therefore an optimal solution can again be supported on a chain. The construction for a fixed chain uses only the Dimension Invariant Inequality and the properties established above, so the remainder of the proof of Lemma 2 carries over unchanged. We therefore obtain an ordinal randomized policy with the same $1/e$ probability guarantee as in Theorem 1.

This policy was constructed using the extension M'. Finally, we show that the extension need not be given to the algorithm. Since M is finite, there are only finitely many deterministic ordinal online policies that always maintain independence in M. A history is determined by the ordered sequence of labels seen so far, their relative weight order, and the previous accept/reject decisions. Fixing all random choices of the policy constructed above in advance therefore induces a probability distribution over this finite set of deterministic policies.

We can recover a suitable distribution using only the known matroid M. Consider the finite LP that chooses a probability distribution over deterministic ordinal policies and maximizes a variable

t. For every total order on E and every element of the corresponding optimal basis, require that its selection probability, averaged over all arrival orders, be at least t. All coefficients of this LP are determined by M. The randomized policy constructed from M' provides a feasible distribution with $t \geq 1/e$, so the optimum of the LP is at least $1/e$. We may therefore solve this LP before the arrivals, sample a deterministic policy from the resulting distribution, and follow it online. The resulting algorithm depends only on the known matroid M and has the required probability guarantee. □

For finite matroids, admitting a finitary modular extension is equivalent to being fully modular extendable. The nontrivial direction follows from [9, Lemma 3.2]. Hence Corollary 2 applies precisely to fully modular extendable matroids. By [9, Theorem 3.4], these are exactly the matroids whose connected components either have rank 3 or are skew-representable. In particular, the class covered by Corollary 2 strictly contains linear matroids; the rank-3 non-Pappus and non-Desargues matroids are not linear, but are fully modular extendable and hence are covered by the corollary. By contrast, the rank-4 Vámos matroid is not skew-representable and therefore does not admit a finitary modular extension.

AI Disclosure. The authors have been working on some of the ideas in this paper since 2024. In particular, the idea of assigning acceptance probabilities equal to $k/(i-1)$ in a correlated way predates the use of generative AI in this project, while the formulation of these probabilities through a linear program was motivated by the approach of Buchbinder et al. [11]. While trying to prove the statement for binary matroids, the authors discussed the problem with ChatGPT-6 Astra on September 14–15, 2026. On September 15, 2026 (1:02 AM PDT), Astra identified the supermodularity of g_p and proposed the uncrossing argument used to prove the feasibility of the LP. The resulting proof, which forms the basis of Section 3, was initially drafted with assistance from ChatGPT-6 Astra. The argument was subsequently checked in full by the authors, who are responsible for its verification and final presentation.

References

[1] H. Abdi, K. Banihashem, M. Hajiaghayi, and D. Mittal. On the strong matroid secretary conjecture and beyond, 2026. URL: `https://arxiv.org/abs/2609.19118`, `arXiv:2609.19118`.

[2] D. Abdolazimi, A. R. Karlin, N. Klein, and S. Oveis Gharan. Matroid Partition Property and the Secretary Problem. In *ITCS*, volume 251, pages 2:1–2:9, 2023.

[3] M. Babaioff, M. Dinitz, A. Gupta, N. Immorlica, and K. Talwar. Secretary problems: Weights and discounts. In *SODA*, pages 1245–1254, 2009.

[4] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. Matroid secretary problems. *Journal of the ACM*, 65(6):35:1–35:26, 2018. `doi:10.1145/3212512`.

[5] M. Babaioff, N. Immorlica, and R. Kleinberg. Matroids, secretary problems, and online mechanisms. In N. Bansal, K. Pruhs, and C. Stein, editors, *SODA*, pages 434–443, 2007.

[6] M. Bahrani, H. Beyhaghi, S. Singla, and S. M. Weinberg. Formal barriers to simple algorithms for the matroid secretary problem. In *WINE*, pages 280–298, 2021.

[7] K. Banihashem, M. Hajiaghayi, D. R. Kowalski, P. Krysta, D. Mittal, and J. Olkowski. Beating competitive ratio 4 for graphic matroid secretary. In *ESA*, pages 52:1–52:16, 2025.

[8] M. Bateni, M. Hajiaghayi, and M. Zadimoghaddam. Submodular secretary problem and extensions. *ACM Transactions on Algorithms*, 9(4):1–23, 2013.

[9] K. Bérczi, B. Gehér, A. Imolay, L. Lovász, C. Padró, and T. Schwarcz. Interaction between skew-representability, tensor products, extension properties, and rank inequalities. In *SODA*, pages 328–354, 2026.

[10] K. Bérczi, V. Livanos, J. A. Soto, and V. Verdugo. Matroid secretary via labeling schemes. In *IPCO*, pages 128–141, 2025.

[11] N. Buchbinder, K. Jain, and M. Singh. Secretary problems via linear programming. *Mathematics of Operations Research*, 39(1):190–206, 2014.

[12] S. Chakraborty and O. Lachish. Improved competitive ratio for the matroid secretary problem. In *SODA*, pages 1702–1712, 2012.

[13] T.-H. H. Chan, F. Chen, and S. H.-C. Jiang. Revealing optimal thresholds for generalized secretary problem via continuous LP: Impacts on online K-item auction and bipartite K-matching with random arrival order. In *SODA*, pages 1169–1188. SIAM, 2015.

[14] J. Correa, A. Cristi, B. Epstein, and J. A. Soto. Sample-driven optimal stopping: From the secretary problem to the i.i.d. prophet inequality. *Mathematics of Operations Research*, 49(1):441–475, 2024.

[15] J. R. Correa, A. Cristi, B. Epstein, and J. A. Soto. The two-sided game of googol and sample-based prophet inequalities. In *SODA*, pages 2066–2081, 2020.

[16] A. Cristi, P. Dütting, R. Kleinberg, R. Paes Leme, and N. Patel. Online Matroid Embeddings. In *ICALP*, volume 374, pages 69:1–69:24, 2026.

[17] N. B. Dimitrov and C. G. Plaxton. Competitive weighted matching in transversal matroids. *Algorithmica*, 62(1):333–348, 2012.

[18] M. Dinitz and G. Kortsarz. Matroid secretary for regular and decomposable matroids. *SIAM Journal on Computing*, 43(5):1807–1830, 2014.

[19] S. Dughmi. The outer limits of contention resolution on matroids and connections to the secretary problem. In *ICALP*, pages 42:1–42:18, 2020.

[20] S. Dughmi. Matroid secretary is equivalent to contention resolution. In *ITCS*, pages 58:1–58:23, 2022.

[21] P. Dütting, S. Lattanzi, R. P. Leme, and S. Vassilvitskii. Secretaries with advice. *Mathematics of Operations Research*, 49(2):856–879, 2024.

[22] P. Dütting, R. Paes Leme, M. Pál, and N. Patel. Graphic matroid secretary without the graph, 2026. arXiv:2608.11413. URL: `https://arxiv.org/abs/2608.11413`, `arXiv:2608.11413`.

[23] E. B. Dynkin. The optimum choice of the instant for stopping a Markov process. *Soviet Math. Dokl*, 4:627–629, 1963.

[24] M. Feldman, O. Svensson, and R. Zenklusen. A simple O(log log(rank))-competitive algorithm for the matroid secretary problem. *Mathematics of Operations Research*, 43(2):638–650, 2018.

[25] M. Feldman and R. Zenklusen. The submodular secretary problem goes linear. *SIAM Journal on Computing*, 47(2):330–366, 2018.

[26] S. O. Gharan and J. Vondrák. On variants of the matroid secretary problem. *Algorithmica*, 67(4):472–497, 2013.

[27] Z. Huang, Z. Parsaeian, and Z. Zhu. Laminar matroid secretary: Greedy strikes back. In *ESA*, pages 73:1–73:8, 2024.

[28] S. Im and Y. Wang. Secretary problems: Laminar matroid and interval scheduling. In *SODA*, pages 1265–1274, 2011.

[29] P. Jaillet, J. A. Soto, and R. Zenklusen. Advances on matroid secretary problems: Free order model and laminar case. In *IPCO*, pages 254–265, 2013.

[30] T. Kesselheim, K. Radke, A. Tönnis, and B. Vöcking. An optimal online algorithm for weighted bipartite matching and extensions to combinatorial auctions. In *ESA*, pages 589–600, 2013.

[31] R. Kleinberg. A multiple-choice secretary algorithm with applications to online auctions. In *SODA*, pages 630–631, 2005.

[32] N. Korula and M. Pál. Algorithms for secretary problems on graphs and hypergraphs. In *ICALP*, pages 508–520, 2009.

[33] O. Lachish. O(log log rank) competitive ratio for the matroid secretary problem. In *FOCS*, pages 326–335, 2014.

[34] D. V. Lindley. Dynamic programming and decision theory. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 10(1):39–51, 1961.

[35] T. Ma, B. Tang, and Y. Wang. The Simulated Greedy Algorithm for Several Submodular Matroid Secretary Problems. In *STACS*, pages 478–489, 2013.

[36] R. Rado. Note on independence functions. *Proceedings of the London Mathematical Society*, 7:300–320, 1957. `doi:10.1112/plms/s3-7.1.300`.

[37] R. Santiago, I. Sergeev, and R. Zenklusen. Constant-competitiveness for random assignment matroid secretary without knowing the matroid. *Mathematical Programming*, 210(1):815–846, 2025.

[38] S. Singla. The matroid secretary conjecture is true, 2026. arXiv:2609.14555v1. URL: `https://arxiv.org/abs/2609.14555`, `arXiv:2609.14555`.

[39] J. A. Soto. Matroid secretary problem in the random-assignment model. *SIAM Journal on Computing*, 42(1):178–211, 2013.

[40] J. A. Soto, A. Turkieltaub, and V. Verdugo. Strong algorithms for the ordinal matroid secretary problem. *Mathematics of Operations Research*, 46(2):642–673, 2021.