Online Matroid Embeddings

Andrés Cristi[1] Paul Dütting[2] Robert Kleinberg[3] Renato Paes Leme[2]
Neel Patel[1]

[1]EPFL, andres.cristi@epfl.ch, neel.patel@epfl.ch
[2]Google Research, duetting@google.com, renatoppl@google.com
[3]Cornell University, rdk@cs.cornell.edu

October 2025

Abstract

We introduce the notion of an online matroid embedding, which is an algorithm for mapping an unknown matroid that is revealed in an online fashion to a larger-but-known matroid. We establish the existence of such an embedding for binary matroids, and use it to relate variants of the binary matroid secretary problem to each other, showing that seemingly simpler problems are in fact equivalent to seemingly harder ones (up to constant-factors). Specifically, we show this to be the case for the version of the matroid secretary problem in which the matroid is not known in advance, and where it is known in advance. We also show that the version with known matroid structure, is equivalent to the problem where weights are not fully adversarial but drawn from a known pairwise-independent distribution.

1 Introduction

A common setup in online algorithms is to have a matroid whose structure is revealed to the algorithm one element at a time. The algorithm processes the elements of the ground set in sequence, and at each point in time, it has access to the dependencies between the elements that have already arrived. Typical examples include the famous matroid secretary problem (MSP) [BIK07, BIKK07, Lac14, FSZ18] and matroid prophet inequalities [CHMS10, KW12].

In such problems, is there any advantage in knowing the matroid structure in advance? Imagine the following situation: we are processing an unknown matroid $\mathbf{M}$; however, we know a fixed (potentially very large) matroid **BigM** that has an isomorphic copy of every possible matroid $\mathbf{M}$, and we can construct this embedding online. We will show that if such an object exists, then we can reduce the version of the problem where the matroid is revealed online to the version of the problem where the matroid structure is known, by assuming our matroid is **BigM**. Moreover, if the on-the-fly embedding maintains uniform random order, then the existence of such an online embedding implies that in that class, the matroid secretary problem with *unknown* structure is no harder than the matroid secretary problem with *known* structure.

Important recent progress on the MSP has established that it is equivalent to the matroid prophet secretary problem with correlated distributions [Dug21, Dug20]. Though seemingly unrelated, another consequence of the existence of such online embeddings will be that—for certain matroids— this equivalence holds even if we impose pairwise independence.

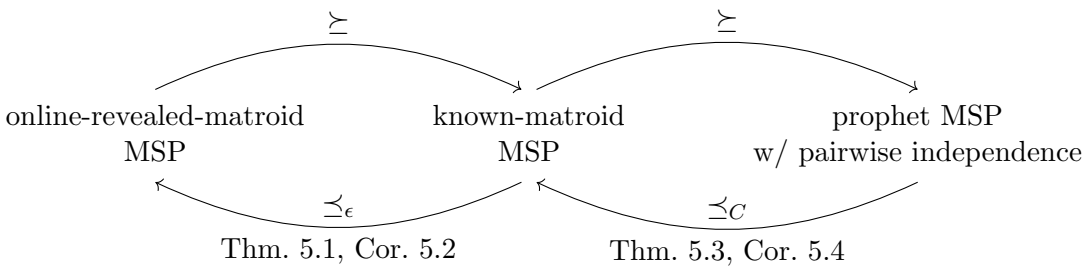


Figure 1: Reductions for binary matroids. We use $P \succeq Q$ to indicate that P is harder than Q, and we use $\succeq_\epsilon$ and $\succeq_C$ to designate an additive ε or multiplicative factor C loss in approximation.

1.1 Our Contribution

Online Matroid Embedding Our main conceptual contribution is to define the notion of an *online matroid embedding* (OME) in Section 3. For a given class of matroids $\mathcal{C}$ and a host matroid **BigM**, we define an OME as a set of matroid monomorphisms, i.e., mappings that preserve the matroid structure, from any matroid in the class $\mathcal{C}$ into **BigM** that can be constructed sequentially, only using calls to an independence oracle over the set of elements observed so far.

The use of embeddings in algorithm design is an idea that has been successfully explored in other contexts, most notably, metric embeddings both in classic settings [Bou85, LLR95, Bar98] as well as more recently in online settings [IMSZ10, BFU20, NR25]. While our motivation and main application is the matroid secretary problem, we believe that understanding maps between matroids preserving structure is an important mathematical question in its own right that can enable other algorithmic applications beyond the matroid secretary problem.

Consequences for the MSP We use the concept of online matroid embeddings to gain insights into the complexity of the matroid secretary problem (MSP) on *binary matroids* (see Section 2). Specifically, we relate different variants of the problem to each other and show that seemingly simpler ones are actually equivalent to harder ones (up to constant factors). See Figure 1 for an overview of the reductions that we establish in this paper.

The three variants we are interested in are: (1) the online-revealed-matroid MSP, where the matroid is a priori unknown to the online algorithm and the algorithm has access to an independence oracle on the already arrived elements, (2) the known-matroid MSP, where the structure of the matroid is known to the algorithm in advance, and (3) the prophet MSP, where the matroid structure is known in advance and additionally the weights of the elements are drawn from a known, but possibly correlated distribution.

Clearly, the online-revealed-matroid MSP is harder than the known-matroid MSP and the known-matroid MSP is harder than the prophet MSP, in the sense that if we have an α-approximation for one problem, then we also have an α-approximation for the other. Two main implications of our work are "inverses" of these statements for binary matroids, that hold up to a constant-factor loss, and apply even if we impose pairwise-independence in the prophet MSP. Such reductions between different average-case problems are notoriously difficult to achieve, as they need to ensure or maintain rather stringent assumptions on the input distribution that are essential for the target algorithm to be applicable in a meaningful way, and the required properties are easily disrupted.

Step 1: A Reduction From Online-Revealed Matroid MSP to Known-Matroid MSP. In Theorem 5.1 we show that the existence of an OME for a class of matroids enables a reduction from the

online-revealed-matroid MSP to the known-matroid MSP. The challenge in proving this is to show that the online embedding can be used in a way that (almost) maintains uniform random arrival order. More precisely, let $\mathbf{M}$ be the unknown matroid that is revealed to the algorithm in an online fashion and let f be an OME into $\mathbf{BigM}$. Now consider the reduction: upon arrival of an element-weight pair (e, w_e) in matroid $\mathbf{M}$ at iteration t, we construct the corresponding element-weight pair $(f(e), w_e)$ as an input to MSP on $\mathbf{BigM}$ at iteration t. However, this is not a valid input to known-matroid MSP on $\mathbf{BigM}$ as it does not construct a random arrival order over the elements that are not in the image of $\mathbf{M}$.

To overcome the shortcoming of the above simple reduction, we interleave the elements in $\mathbf{BigM}$ that are in the image of $\mathbf{M}$ with the remaining elements in $\mathbf{BigM}$. In the proof of Theorem 5.1, our main technical argument shows that, while this interleaving does not ensure uniformity of the arrival order of the elements in $\mathbf{BigM}$, it leads to an arrival order over elements in $\mathbf{BigM}$ that is close to uniformly random arrival order in total variation distance (Section 5.1). We then complete the reduction with a coupling argument that shows that the existence of an α-competitive algorithm for the known-matroid MSP implies the existence of an $(\alpha - \epsilon)$-competitive algorithm for the online-revealed-matroid MSP.

Together with the existence of an OME that maps binary matroids into the complete binary matroid (see Section 4 and discussion below), our reduction implies that an algorithm for the MSP over binary matroids cannot meaningfully use any advance information about the matroid (Corollary 5.2). This is in contrast to all known $O(1)$-competitive algorithms for special cases of binary matroids [KP09, DK14].

Step 2: A Reduction from Prophet MSP w/ Pairwise Independence to Known-Matroid MSP. In Theorem 5.3, we show that the existence of an OME from a class of matroids $\mathcal{C}$ to $\mathbf{BigM}$ satisfying a 2-transitivity property (that is satisfied by complete binary matroids, see definition in Section 2), allows to translate an α-competitive algorithm for prophet MSP with pairwise-independent weight distribution on $\mathbf{BigM}$ into a $C \cdot (\alpha - o(1))$-competitive algorithm for known-matroid MSP on matroid $\mathbf{M} \in \mathcal{C}$, for some constant $C > 0$.

To establish Theorem 5.3, we build on [Dug21, Dug20] and show how to reduce prophet MSP with arbitrary correlation on $\mathbf{M} \in \mathcal{C}$ to prophet MSP with pairwise-independent weight distribution on $\mathbf{BigM}$.[1] To prove this, we first show that any uniformly random automorphism $f : \mathbf{BigM} \to \mathbf{BigM}$ satisfies the following property: for any pair of elements e, e', $\Pr[f(e) = e'] = \frac{1}{n}$ and for any two pairs of independent elements e_1, e_2 and e'_1, e'_2, $\Pr[f(e_1) = e'_1 \wedge f(e_2) = e'_2] = \frac{1}{n \cdot (n-1)}$, where $n = |\mathbf{BigM}|$ (Lemma 3.2). This property allows us to construct an "almost pairwise independent" randomized OME $f' : \mathbf{M} \to \mathbf{BigM}$ by simply composing the given OME with a uniformly random automorphism on $\mathbf{BigM}$. The resulting weight distribution is approximately pairwise independent in the sense that for any pair of elements $e, e' \in \mathbf{BigM}$ and weights w, w' it holds that $|\Pr[w(e) = w] \cdot \Pr[w(e') = w'] - \Pr[w(e) = w' \wedge w(e') = w']| = O\left(\frac{1}{n^2}\right)$.

To conclude the proof, we show that there exists a pairwise-independent distribution that is close to the induced weight distribution over $\mathbf{BigM}$ (Theorem 5.9). This is one of the most technical results of the paper (see discussion below, and Section B). Combining Theorem 5.9 with a coupling argument similar to the one in our other reduction completes the proof.

We note that constructing an exactly k-wise independent distribution from an approximately k-wise independent distribution has been studied in previous work [AGM03, AL12, AAK$^+$07]. However, their techniques focus on a set of *Bernoulli* random variables with identical marginals

[1] A technical detail that we are ignoring here is that our reduction is from a restricted version of the prophet MSP with arbitrary correlation, which results in an additional constant-factor loss.

[AGM03, AAK+07] or "uniformity" of the underlying random variables [AL12]—both conditions do not hold in our setting as the weight distribution in the prophet MSP instance can be arbitrarily correlated. In fact, in both works [AGM03, AL12], they show that if the random variables $X_1, \ldots, X_n$ satisfy "uniformity" and $|\mathbb{E}[X_i \cdot X_j] - \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]| \leq \varepsilon$ then there exists pairwise independent random variables $\tilde{X}_1, \ldots, \tilde{X}_n$ within a distance of $O(n^2 \cdot \varepsilon)$ — which is not enough for our purpose as $\varepsilon = \Theta(1/n^2)$ in our case.

To obtain Theorem 5.9, we construct an explicit pairwise-independent weight distribution over **BigM** by a sequence of "small" perturbations to a naturally induced "almost" pairwise-independent distribution. At each step, we perturb the original distribution such that $\omega(1)$ many pairs of random variables end up being independent (Procedure 1 and Procedure 2 in Section B) while always decreasing the pairwise correlation of the rest of the pairs (Lemma B.1). Then the main technical work is devoted to showing that the total deviation through our procedures is in the order of $\varepsilon \cdot o(n^2)$ (Section B.3 and Section B.4), which, combined with the fact that $\varepsilon = O(1/n^2)$ leads to the desired result. We believe that our idea of sequentially constructing small perturbations would find further applications to obtain exact pairwise (or k-wise) independent distributions from their approximate counterparts in other settings.

Constructing OMEs In Section 4 (Theorem 4.2 and Theorem 4.4), we provide a complete analysis for binary matroids. Namely, for the class of binary matroids **M** with n elements, there is an online matroid embedding into **BigM** the complete binary matroid $\mathbb{F}_2^n$. We also develop a technique for making the OME order-independent: we use properties of the automorphism group of $\mathbb{F}_2^n$ together with randomization to ensure that the images of the elements in **M** are not correlated with the arrival order. This technique is in fact more general (Theorem 4.6) and can be applied whenever the group of automorphisms of the host matroid is "sufficiently rich" in a sense that the theorem statement makes precise.

The key property of binary matroids that we exploit to establish these results is that in $\mathbb{F}_2^n$, there is a unique element that completes a circuit, in the sense that there cannot be two circuits of the same size that intersect in all but one element of each.

We refer to online matroid embeddings where both **M** and **BigM** are of the same class as "within-class" OMEs. In Section 6.1 we explore whether such "within-class" OMEs can exist for graphic matroids. We show that such embeddings cannot exist, in fact we show that graphic matroids cannot be embedded in an online-fashion to regular matroids. To rule out the existence of such an online embedding, we show that if it would exist, then **BigM** must contain an isomorphic copy of $\mathbb{F}_2^n$. However, $\mathbb{F}_2^n$ contains an isomorphic copy of the Fano plane which is not representable over $\mathbb{F}_3$ [Tut58]. Hence **BigM** can't be regular.

We believe that the lack of online matroid embeddings for graphic matroids/regular matroids that "don't leave the class" may shed light on why progress on the known-matroid MSP for graphic and regular matroids has not extended to the online-revealed version of these problems, and more generally the MSP for general binary matroids.

In Section 6.2 we give another example of an OME, namely for laminar matroids. We show (in Theorem 6.6) how to embed the class of laminar matroids **M** with at most n elements into **BigM** which is a complete linear matroid of rank n over any field with sufficiently many elements.

Finally in Section 6.3 we show an impossibility result of constructing an OME for the class of all matroids. This is shown by studying finite projective planes and showing that for those matroids, elements that haven't arrived yet impose non-trivial constraints on the already arrived elements. As a corollary we obtain an impossibility of constructing an OME for the class of all matroids representable over fields of characteristic at least 7.

Approximate OMEs In Section 7 we extend the notion of an OME to allow distortion, i.e., the map approximately preserves the rank. We observe that the α-partition property in [AKKG23] and [DKP24] can be viewed as an approximate matroid embedding into the free matroid.

First, combining our formalism with their lower bounds on embedding into the free matroid, we also provide a lower bound on the distortion needed to embed the complete binary matroid into a graphic matroid (Corollary 7.3). This result is an example of the power of the formalism: with the right definitions, extending the lower bound to larger classes becomes a simple corollary.

In Theorem 7.6 we show the tightness of the $\Omega(n/\log n)$ lower bound in [DKP24] of the distortion of embedding the complete binary matroid into the free matroid by constructing an embedding achieving this distortion. We also show that there is no constant approximate online embedding of the class of graphical matroids into a free matroid when the underlying matroid is not known upfront (Theorem 7.7). Therefore, any constant competitive algorithm for unknown graphical matroid secretary that relies on constructing an online embedding of the graph into a free-matroid has to exploit the random arrival order of the underlying elements or develop new techniques that do not rely on online embedding into a free matroid.

1.2 Discussion and Significance of Results

We believe that the existence or non-existence of (approximate) online matroid embeddings can shed new light on different classes of matroids and how they relate to each other. In this work, we demonstrate two implications for the matroid secretary problem.

Our first implication (Theorem 5.1) offers the first formalization of the intuition that, in general, advance knowledge of the matroid structure should not help in the design of a constant-competitive algorithm for the matroid secretary problem. In light of this, it would be interesting to develop algorithms for classes of matroids for which constant-competitive algorithms exist when the algorithm has advance knowledge of the matroid structure [e.g., KP09, DK14].

Our second implication (Theorem 5.3), in turn, presents a novel "line of attack" for obtaining such an algorithm for the class of binary matroids (for which no constant-competitive algorithm is known). While it was already known that it suffices to find such an algorithm for the secretary prophet version with correlated weights [Dug21, Dug20], general correlated weight distributions offer little additional structure. Our result shifts the challenge away from intractable arbitrary correlations, towards the better-understood realm of pairwise independent distributions. Pairwise independent distributions admit powerful tools like concentration inequalities and have found application in areas such as hashing and constructions of pseudo-random generators (for more details, see surveys [LW$^+$06, Vad12]), as well as prophet inequalities [CGLW22].

1.3 Related Work

Matroid Secretary Problem The matroid secretary problem was first studied in [BIK07, BIKK07, BIKK18], who gave a $O(\log(\mathsf{rank}))$-competitive algorithm for general matroids. This bound was improved to $O(\sqrt{\log(\mathsf{rank})})$ in [CL12], and the state-of-the-art is a $O(\log\log(\mathsf{rank}))$-competitive algorithm [Lac14, FSZ18]. The algorithms of [Lac14, FSZ18] only uses independence oracle calls on subsets of the elements revealed so far.

For graphic matroids there is a $O(1)$-competitive algorithm, provided that the graphic matroid is known in advance [KP09]. The same is true for the more general class of regular matroids [DK14]. Laminar matroids also admit an $O(1)$-competitive algorithm [IW11, JSZ13]. Some evidence for the difficulty of the matroid secretary problem for general binary matroids can be found in [LMP22]

and [AKKG23], showing that binary matroids are not (b, c)-decomposable, ruling out a promising approach to obtaining an $O(1)$-competitive algorithm for this class.

Oveis Gharan and Vondrák [OGV13] systematized the study of matroid secretary problem variants, establishing a notation for classifying problem variants according to whether the elements arrive in adversarial or random order, whether the assignment of weights to elements is adversarial or random, and whether or not the matroid structure is known in advance. In their nomenclature, the main question addressed in our work is whether the RO-AA-MK variant is equivalent to the RO-AA-MN variant for matroids in general, or for specific classes of matroids. Interestingly, for variants with adversarial arrival order but random weight assignment, [OGV13] demonstrates a stark qualitative difference in approximability: the AO-RA-MK model (when the matroid structure is known in advance) admits a 64-competitive algorithm for all matroids, whereas the AO-RA-MN model (when the number of elements is known in advance but the matroid structure is revealed online) has no constant-competitive algorithm even for the class of *rank one matroids*!

Very recently, [SSZ25] gave a $O(1)$-competitive algorithm for the matroid secretary problem in the random assignment model when the matroid structure is not known in advance, and instead is only revealed over time. In a similar spirit, [SSZ23] presents an online contention resolution scheme for graphic matroids, that uses almost no advance information about the graph. However, they assume that the endpoints of the edges are revealed upon their arrival which leads to an obvious OME into a graphical matroid.

Matroid Prophet Inequalities The matroid prophet inequality problem was first studied in [HKS07]. An asymptotically optimal $(1 - o(1))$-competitive algorithm for k-uniform matroids was given in [Ala14]. A tight $O(1)$-competitive algorithm for the matroid prophet inequality problem was given in [KW12], also see [DFKL20] for the problem of maximizing submodular functions subject to matroid constraints. Constant-factor competitive algorithms can also be obtained via online contention resolution schemes (OCRS) [FSZ21]. Random-order versions of the matroid prophet inequality problem are studied in [EHKS18].

To the best of our knowledge, all these algorithms exploit that the matroid structure is known in advance. An additional difficulty for reductions of the type we present in this paper, is that typically these algorithms need to know the identity of the distribution that a certain element's weight is drawn from. For the i.i.d. case this is obviously not an obstacle, and so our reductions apply. We believe that extensions of our techniques might shed further light on the variant of the matroid prophet inequality problem, in which the matroid is revealed online.

Metric Embeddings and Distortion An important inspiration for this work comes from the literature on metric embeddings. A classic result in this context is Bourgain's theorem [Bou85]. The algorithmic importance of such embeddings, and Bourgain's theorem in particular, was first highlighted in the seminal papers of [LLR95, Bar98].

Since then metric embeddings have found applications in a host of algorithmic problems, see, e.g., the survey of [Ind01] and Chapter 15 of [Mat02]. Closer to our notion of online matroid embeddings is a recent line of work on online metric embeddings [IMSZ10, BFU20, NR25] in which points of a metric space are presented one at a time to an algorithm who must then decide on a mapping to the host metric space. The main difference is that instead of preserving a matroid structure, those papers try to minimize metric distortion.

2 Matroids, Morphisms, and K-representations

Throughout the paper, we will use $[n]$ to denote the set of integers $\{1, 2, \ldots, n\}$.

Matroid Definition A matroid $\mathbf{M}$ is composed by a ground set M and a rank function $\mathsf{rank}_{\mathbf{M}} : 2^M \to \mathbb{Z}_+$ satisfying the following properties:

- $\mathsf{rank}_{\mathbf{M}}(\emptyset) = 0$;
- $\mathsf{rank}_{\mathbf{M}}(S \cup \{i\}) - \mathsf{rank}_{\mathbf{M}}(S) \in \{0, 1\}, \forall S, \{i\} \subseteq M$
- $\mathsf{rank}_{\mathbf{M}}(S \cup T) + \mathsf{rank}_{\mathbf{M}}(S \cap T) \leq \mathsf{rank}_{\mathbf{M}}(S) + \mathsf{rank}_{\mathbf{M}}(T), \forall S, T \subseteq M$ (submodularity)

It follows from the second condition that $\mathsf{rank}_{\mathbf{M}}(S) \leq |S|$. Whenever $|S| = \mathsf{rank}_{\mathbf{M}}(S)$ we say that S is an independent set of the matroid. Otherwise, we say that S is dependent. A minimal dependent set is called a *circuit*, i.e., $C \subseteq M$ is a circuit if C is dependent but every strict subset $S \subsetneq C$ is independent. We say that a matroid has rank r if $r = \max_S \mathsf{rank}_{\mathbf{M}}(S)$.

We say that an element $x \in \mathbf{M}$ is a loop if $\mathsf{rank}_{\mathbf{M}}(\{x\}) = 0$. We say that a matroid is loop-free if every set of one element is independent. Given a set $S \subseteq \mathbf{M}$ we define the span as $\mathsf{span}_{\mathbf{M}}(S) = \{x \in \mathbf{M}; \mathsf{rank}_{\mathbf{M}}(S \cup \{x\}) = \mathsf{rank}_{\mathbf{M}}(S)\}$.

Matroid Morphisms We will use the same notation to refer to a matroid and its ground set. Given two matroids $\mathbf{M}$ and $\mathbf{N}$ we will define a morphism $f : \mathbf{M} \to \mathbf{N}$ to be a map between their ground sets that preserves rank, i.e.:

$$\mathsf{rank}_{\mathbf{N}}(f(S)) = \mathsf{rank}_{\mathbf{M}}(S), \forall S \subseteq \mathbf{M}.$$

Whenever the matroid morphism is an injective map, we will say it is a *matroid monomorphism* or a *matroid embedding*. Whenever it is bijective, we will say it is a *matroid isomorphism*. An isomorphism from a matroid to itself is called an automorphism. (Aside: this paragraph defines the category of matroids in the sense of category theory. However, we won't use any other fact from category theory other than borrowing its very convenient language.)

We refer to the set of automorphisms $\mathbf{M} \to \mathbf{M}$ as $\mathrm{Aut}(\mathbf{M})$, which forms a group under composition, i.e., given $f, g \in \mathrm{Aut}(\mathbf{M})$, then $f \circ g \in \mathrm{Aut}(\mathbf{M})$ (and $\circ$ satisfies the group axioms).

Element Copies Given a matroid $\mathbf{M}$ and an integer k we will define the matroid $\mathbf{M}_{[k]}$ by creating k copies of each element of $\mathbf{M}$. Formally, the ground set of $\mathbf{M}_{[k]}$ is $\{(u, j); u \in \mathbf{M}, j \in [k]\}$. The rank function of $\mathbf{M}_{[k]}$ is induced by the projection $\phi : \mathbf{M}_{[k]} \to \mathbf{M}$ that maps $(u, j) \mapsto u$, i.e., $\mathsf{rank}_{\mathbf{M}_{[k]}}(S) = \mathsf{rank}_{\mathbf{M}}(\phi(S))$. By definition, the projection ϕ is a matroid morphism from $\mathbf{M}_{[k]} \to \mathbf{M}$.

If $\mathbf{N}$ is a matroid of at most n elements, every morphism $f : \mathbf{N} \to \mathbf{M}$ can be written as: $f = \phi \circ f'$ where $f' : \mathbf{N} \to \mathbf{M}_{[n]}$ is a monomorphism.

Direct Sum Given two matroids $\mathbf{M}$ and $\mathbf{N}$, we define their direct sum $\mathbf{M} \oplus \mathbf{N}$ as the matroid whose ground set is the disjoint union of the ground sets of $\mathbf{M}$ and $\mathbf{N}$ and $\mathsf{rank}_{\mathbf{M} \oplus \mathbf{N}}(S) = \mathsf{rank}_{\mathbf{M}}(S \cap \mathbf{M}) + \mathsf{rank}_{\mathbf{N}}(S \cap \mathbf{N})$ for all S in the disjoint union of ground sets.

Graphic Matroids We will define a few special classes of interest. We start with *graphic matroids*. Given a graph with edge set E, we can define a matroid with ground set E by defining the $\mathsf{rank}(S)$ of a subset $S \subseteq E$ as the maximum number of edges in S that don't form a cycle. We say that a matroid $\mathbf{M}$ is graphic if it is isomorphic to the matroid obtained from an undirected graph as we just described.

As an example, consider the matroid $\mathbf{M}$ with ground set $\{a, b, c\}$ and rank function such that $\mathsf{rank}(S) = |S|$. The matroid is graphic since it is isomorphic to the matroid that can be obtained from any of the graphs in Figure 2. An important thing to note, however, is that the matroid description contains no information about vertices. It only tells us which sets of edges are independent and which are not. As we can see in the figure, this is typically not enough to fully determine the graph structure.

Figure 2: Two graphs that generate the same matroid on their edge set

K-representable Matroids Let K be a field (e.g. $\mathbb{Q}, \mathbb{R}, \mathbb{F}_p$) and let K^d be the vector space formed by d-dimensional vectors with coordinates in K. We say that a subset of vectors $u_1, \ldots, u_k \in K^d$ is independent if the unique solution to $\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k = 0$ for $\alpha_i \in K$ is $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Any subset of K^d together with the independency relation above defines a matroid. From now on, we will use the notation K^d to represent both the vector space and the corresponding matroid. We say that a matroid $\mathbf{M}$ is K-representable if there is a matroid morphism $\mathbf{M} \to K^d$ for some integer d.

If a matroid $\mathbf{M}$ is K-representable for every field K we say that $\mathbf{M}$ is a *regular matroid*. Every graphic matroid is representable over any field by mapping an edge (u, v) to the vector $e_u - e_v$ where e_u the the u-th unit vector. (This is true even over $\mathbb{F}_2$ where $e_u - e_v = e_u + e_v$.)

For example, the matroids in Figure 2 can be represented by the vectors $(1, -1, 0, 0)$, $(0, 1, -1, 0)$, $(0, 0, 1, -1)$. As it is the case for graphic matroids, the matroid description has no information about vectors and the representation is again not unique. An equally good representation is $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.

We will be specially interested in *binary matroids* which are matroids that are representable over $\mathbb{F}_2$ (the finite field of 2 elements where addition and multiplication are performed mod 2).

Laminar matroids A family of sets, $\mathcal{A}$, is called *laminar* if it satisfies the property that for any $A, A' \in \mathcal{A}$, at least one of the sets $A \cap A'$, $A \setminus A'$, $A' \setminus A$ is empty. A *laminar matroid* $\mathbf{M}$ is one for which there exists a laminar family of sets $\mathcal{A}$ consisting of subsets of the ground set of $\mathbf{M}$ and a function $c : \mathcal{A} \to \mathbb{Z}_+$, such that the independent sets of $\mathbf{M}$ are precisely those sets $I \subseteq M$ such that $|I \cap A| \leq c(A)$ for all $A \in \mathcal{A}$.

Uniform Matroid We will denote by $\mathbf{U}_{n,r}$ the uniform matroid of n elements and rank r. This is the matroid with ground set $[n]$ and whose rank function is $\mathsf{rank}_{\mathbf{U}_{n,r}}(S) = \min(r, |S|)$. We call the $\mathbf{Fr}_n := \mathbf{U}_{n,n}$ the free matroid of rank n, i.e., a matroid of n elements in which every set is independent.

Trivial Matroid Let $\mathbf{T}$ be the trivial matroid which has ground set $\{0\}$ and rank function $\mathsf{rank}_{\mathbf{T}}(S) = 0$ for all sets S.

2-transitive Matroid We say that a simple matroid $\mathbf{M}$ (loop-free and no parallel elements) is 2-transitive if for any pair of independent sets of size two, $\{e_1, e_1'\}, \{e_2, e_2'\} \in \mathbf{M}$ there exists an automorphism $f \in \mathrm{Aut}(\mathbf{M})$ satisfying $f(e_1) = e_1'$ and $f(e_2) = e_2'$. There are several matroids that satisfies the 2-transitive property including complete affine matroids, complete projective matroids, free matroids, and their truncations [Kan85].

3 Online Matroid Embeddings

We are interested in studying matroids whose structure is revealed to an algorithm in an online fashion. For that, it will be useful to take into account the order in which elements are processed, which we will represent by an indexing of the ground set: $\pi : [n] \to \mathbf{M}$.

A matroid $\mathbf{N}$ is a restriction of matroid $\mathbf{M}$ if the ground set of $\mathbf{N}$ is a subset of the ground set of $\mathbf{M}$ and $\mathsf{rank}_{\mathbf{M}}$ coincides with $\mathsf{rank}_{\mathbf{N}}$ on the ground set of $\mathbf{N}$.

Given a matroid with ordered ground set specified by a pair $(\mathbf{M}, \pi)$, we say that $(\mathbf{M}', \pi')$ is a prefix-restriction of $(\mathbf{M}, \pi)$ if $n' = |\mathbf{M}'| < |\mathbf{M}|$, $\mathbf{M}'$ is the restriction of $\mathbf{M}$ to $\pi([n'])$ and π' is the restriction of π to $[n']$.

Let $\mathcal{C}$ be a class of matroids that is closed under restriction (e.g., the class of all matroids, graphic matroids, binary matroids, K-representable matroids, matroids of rank at most r). An *online matroid morphism* (OMM) for class $\mathcal{C}$ consists of a host matroid $\mathbf{BigM}$, together with matroid morphisms

$$f_{\mathbf{M},\pi} : \mathbf{M} \to \mathbf{BigM}$$

for every $\mathbf{M} \in \mathcal{C}$ and every indexing $\pi : [n] \to \mathbf{M}$ of the ground set of $\mathbf{M}$, such that for every prefix-restriction $(\mathbf{M}', \pi')$ of $(\mathbf{M}, \pi)$, the map $f_{\mathbf{M}',\pi'}$ is the restriction of $f_{\mathbf{M},\pi}$ to the ground set of $\mathbf{M}'$.

If all morphisms $f_{\mathbf{M},\pi}$ are monomorphisms, we say that they form an *online matroid embedding* (OME). Given an online matroid morphism it is easy to construct an online matroid embedding by copying the elements of $\mathbf{BigM}$.

Lemma 3.1. *Let $\mathcal{C}$ be a class of matroids, where each matroid $\boldsymbol{M} \in \mathcal{C}$ has at most n elements and $f_{\boldsymbol{M},\pi} : \boldsymbol{M} \to \boldsymbol{BigM}$ form an online matroid morphism. Then there is an online matroid embedding $f'_{\boldsymbol{M},\pi} : \boldsymbol{M} \to \boldsymbol{BigM}_{[n]}$.*

Proof. We define $f'_{\mathbf{M},\pi}$ as follows: for each $u \in \mathbf{M}$ if $u = \pi(k)$ let $f'_{\mathbf{M},\pi}(u) = (f_{\mathbf{M},\pi}(u), j)$ where $j = |i \in [k]; f_{\mathbf{M},\pi}(\pi(i)) = u|$. The functions $f'_{\mathbf{M},\pi}$ are injective by construction and they are matroid morphisms by the definition of $\mathbf{BigM}_{[n]}$. In fact: $f_{\mathbf{M},\pi} = \phi \circ f'_{\mathbf{M},\pi}$ where ϕ is the natural projection $\mathbf{BigM}_{[n]} \to \mathbf{BigM}$. Finally note that they can be constructed online since the identity of the copy used is only a function of the set of elements that arrived up to this point. □

With this definition we can ensure that an online algorithm is able to construct a monomorphism from an unknown matroid $\mathbf{M}$ to $\mathbf{BigM}$ in an online fashion. Consider a matroid $\mathbf{M}$ for which the elements arrive according to π. At each time t, we can observe the structure of the matroid $\mathbf{M}_t$ which is the restriction of $\mathbf{M}$ to $\pi([t])$. Let π_t be the restriction of π to $[t]$. If we have an online matroid embedding, we can first construct $f_{\mathbf{M}_1,\pi_1}$, then extend to $f_{\mathbf{M}_2,\pi_2}$ and so forth.

It will also be convenient to define a randomized online matroid morphism (embedding) which for every matroid $\mathbf{M} \in \mathcal{C}$ and ordering π specifies a distribution over (mono)morphisms $f_{\mathbf{M},\pi} : \mathbf{M} \to \mathbf{BigM}$ such that for every prefix-restriction $(\mathbf{M}', \pi')$ the distribution of the restriction of $f_{\mathbf{M},\pi}$ to the ground set of $\mathbf{M}'$ coincides with the distribution of $f_{\mathbf{M}',\pi'}$.

Finally, we say that a randomized online matroid embedding is order-independent if the distribution of $f_{\mathbf{M},\pi}$ doesn't depend on π. In other words, for any two orderings π and π', the morphisms $f_{\mathbf{M},\pi}$ and $f_{\mathbf{M},\pi'}$ are equally distributed.

Uniform Order-Independent Embedding Let f be an order-independent online embedding from $\mathbf{M} \to \mathbf{BigM}$. We consider an order independent-randomized embedding $g : \mathbf{M} \to \mathbf{BigM}$ by composing f with uniformly random automorphism $f' \in \mathrm{Aut}(\mathbf{BigM})$, i.e. $g = f' \circ f$.

Interestingly, whenever **BigM** satisfies 2-transitive property then the embedding g maps each element of $e \in \mathbf{M}$ uniformly at random over the matroid **BigM**. In addition, for any independent set of pair of elements $\{e, e'\} \subseteq \mathbf{M}$ and pair of elements $\{\tilde{e}, \tilde{e}'\} \subseteq \mathbf{BigM}$, the events $\{g(e) = \tilde{e}\}$ and $\{g(e') = \tilde{e}'\}$ are 'almost' independent. More formally,

Lemma 3.2. *Given a simple host matroid (loop-free and no parallel elements) **BigM** such that for any two pairs of distinct elements $\{e_1, e_2\}$ and $\{e'_1, e'_2\}$ there exists an automorphism $f' \in \mathrm{Aut}(\boldsymbol{BigM})$ satisfying $f'(e_1) = e'_1$ and $f'(e_2) = e'_2$, then a uniformly random automorphism f sampled from $\mathrm{Aut}(\boldsymbol{BigM})$ satisfies:*

1. *For any $e, e' \in \boldsymbol{BigM}$, $\Pr[f(e) = e'] = \frac{1}{n}$.*
2. *For any two pairs of elements e_1, e_2 and e'_1, e'_2 (s.t. $e_1 \neq e_2$ and $e'_1 \neq e'_2$), we have*

$$\Pr[f(e_1) = e'_1 \wedge f(e_2) = e'_2] = \frac{1}{n \cdot (n-1)}$$

Proof. Consider the action of the group of automorphisms $\mathrm{Aut}(\mathbf{BigM})$ on the set of pairs of distinct elements $P = \{(e_1, e_2) : e_1, e_2 \in \mathbf{BigM}, e_1 \neq e_2\}$. For two pairs $(e_1, e_2), (e'_1, e'_2) \in P$, the set of automorphisms $f \in \mathrm{Aut}(\mathbf{BigM})$ that satisfy $f(e_1, e_2) = (e'_1, e'_2)$ is nonempty, and therefore, it is a coset of the subgroup of stabilizers of (e_1, e_2) (i.e., the set of automorphisms that satisfy $f(e_1, e_2) = (e_1, e_2)$). Since all cosets of a subgroup must have the same size, and because all pairs $(e'_1, e'_2) \in P$ define a different coset, for a uniformly drawn automorphism f,

$$\Pr[f(e_1, e_2) = (e'_1, e'_2)] = \frac{1}{|P|} = \frac{1}{n \cdot (n-1)}.$$

An analogous argument gives that $\Pr[f(e) = e'] = 1/n$. □

Online vs. Offline Embeddings The difficulty of constructing an online matroid embedding is that the elements of **BigM** corresponding to certain elements of **M** must be chosen before the full matroid structure of **M** is known. If we merely wanted to construct a matroid **BigM** that contains an isomorphic copy of every matroid in $\mathcal{C}$, that would be very easy: **BigM** could be taken to be the direct sum of all the matroids in $\mathcal{C}$.

4 OMEs for Binary Matroids

Before we discuss how to use online matroid embeddings in online algorithms, it is important to show first that they exist in non-trivial cases. For that, we will provide a complete analysis for

binary matroids. Recall that $\mathbb{F}_2^n$ is the complete binary matroid of rank n and that graphic matroids and regular matroids are special cases of binary matroids. Our first result is the existence of an OME for this class. This will be done by showing the existence of an OMM and using Lemma 3.1 to convert an OMM to an OME.

Our first step is to show a lemma that the matroid $\mathbb{F}_2^n$ is special in the sense that its group of matroid automorphisms coincides with its group of vector space automorphisms:

Lemma 4.1. *A mapping $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a matroid automorphism iff it is an automorphism of vector spaces.*

Proof. An automorphism of vector spaces $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a bijection such that for any vectors $v_1, \ldots, v_k \in \mathbb{F}_2^n$ it holds that $A(\sum_{i=1}^k v_i) = \sum_{i=1}^k Av_i$. This in particular implies that a set of vectors $v_1, \ldots, v_k$ is independent iff the vectors $Av_1, \ldots, Av_k$ are independent. This is because a coefficient vector $(\alpha_1, \ldots, \alpha_k)$ satisfies the equation $\sum_{i=1}^k \alpha_i v_i = 0$ if and only if it satisfies $\sum_{i=1}^k \alpha_i Av_i = A(\sum_{i=1}^k \alpha_i v_i) = 0$, so the first equation has only trivial solutions if and only if the second equation has only trivial solutions.

For the opposite direction, if A is a matroid automorphism and $e_1, \ldots, e_n$ is the standard basis of $\mathbb{F}_2^n$ then $Ae_1, \ldots, Ae_n$ must be linearly independent elements of $\mathbb{F}_2^n$. Now, take any vector $v = \sum_{i \in S} e_i$. Since $\{v\} \cup \{e_i; i \in S\}$ forms a circuit, then $\{Av\} \cup \{Ae_i; i \in S\}$ must form a circuit. Since the only non-zero constant in $\mathbb{F}_2$ is 1, it must hold that: $Av + \sum_{i \in S} Ae_i = 0$ and hence $Av = \sum_{i \in S} Ae_i$. Hence A is also an automorphism of vector spaces. □

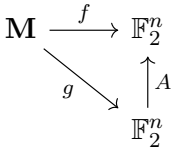
Theorem 4.2. *Let $\mathcal{C}$ be the class of binary matroids of at most n elements and let **BigM** be the complete binary matroid $\mathbb{F}_2^n$. Then there exists an OMM for $\mathcal{C}$ into **BigM**.*

Proof. Given a binary matroid $\mathbf{M}$ we construct a mapping $f : \mathbf{M} \to \mathbb{F}_2^n$ as follows. We keep a counter k initially set to 1. For each element a we process, if it is independent of the previously arrived elements (i.e. there are no circuits containing a and the elements seen so far), we set $f(a) = e_k$ and increment k. Otherwise, a forms a circuit with a set of previously arrived elements $u_1, \ldots, u_m$ for some integer $m \geq 0$. This means that their image $f(a), f(u_1), \ldots, f(u_m)$ must be a minimal $\mathbb{F}_2$-linearly dependent set. Since the only non-zero constant in $\mathbb{F}_2^n$ is 1, then it must hold that:

$$f(a) + f(u_1) + \ldots + f(u_m) = 0$$

and hence we can map: $f(a)$ to $f(u_1) + \ldots + f(u_m)$ (recall that $1 = -1$ in $\mathbb{F}_2$).

Finally, we need to argue that f is a matroid morphism. Observe that if $\mathbf{M}$ is a binary matroid, then there exists a morphism $g : \mathbf{M} \to \mathbb{F}_2^n$. Let $\{b_1, \ldots, b_r\}$ be the elements of $\mathbf{M}$ such that $f(b_i) = e_i$. By the fact that g is matroid morphism, $g(b_1), \ldots, g(b_r)$ are linearly independent elements in $\mathbb{F}_2^n$. By Lemma 4.1 there is an automorphism $A \in \mathrm{Aut}(\mathbb{F}_2^n)$ that takes $g(b_i)$ to e_i. Since matroid morphisms compose, $Ag : \mathbf{M} \to \mathbb{F}_2^n$ is matroid morphism.



Finally, we argue that $f(a) = Ag(a)$ for all a in $\mathbf{M}$. We show this by induction. For each element processed by the algorithm, if it is independent from previously arrived elements, then $f(a) = Ag(a)$ by construction. Otherwise, there are previously arrived elements such that $a, u_1, \ldots, u_m$ form a

circuit. Hence their image under Ag must be a linearly dependent set of $\mathbb{F}_2$-vectors, which means that:

$$Ag(a) = Ag(u_1) + \ldots + Ag(u_m) = f(u_1) + \ldots + f(u_m) = f(a)$$

where the second equality holds by induction. Since f coincides with Ag, f is a matroid morphism. □

Furthermore, there is a randomized OME that is order-independent. We will show it as a consequence of the following lemmas:

Lemma 4.3. *Given a binary matroid $\boldsymbol{M}$ and the complete binary matroid $\mathbb{F}_2^n$, if there are two matroid morphisms $f, g : \boldsymbol{M} \to \mathbb{F}_2^n$, then there exist an automorphism $A \in Aut(\mathbb{F}_2^n)$ such that $f = A \circ g$.*

Proof. Let r be the rank of $\mathbf{M}$ and let $\{b_1, \ldots, b_r\}$ be a basis of $\mathbf{M}$. Then $\{f(b_1), f(b_2), \ldots, f(b_r)\}$ and $\{g(b_1), g(b_2), \ldots, g(b_r)\}$ are both sets of independent vectors in $\mathbb{F}_2^n$. Then there exists an automorphism A of vector spaces (and hence a matroid automorphism) that sends $g(b_i)$ to $f(b_i)$. For any other element in $v \in \mathbf{M}$ consider any circuit formed with a subset of the basis. If $\{v\} \cup \{b_i; i \in S\}$ is a circuit then it must be the case that: $f(v) = \sum_{i \in S} f(b_i)$ and $g(v) = \sum_{i \in S} g(b_i)$. Given that A is a automorphism of vector spaces, then: $Ag(v) = \sum_{i \in S} Ag(b_i) = \sum_{i \in S} f(b_i) = f(v)$. □

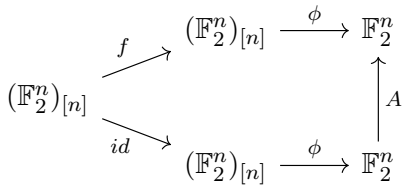
Theorem 4.4. *There is a order-independent randomized OMM from the class of binary matroids $\mathcal{C}$ into the complete binary matroid.*

Proof. Let $f_{\mathbf{M},\pi}$ be the online matroid morphism constructed in Theorem 4.2 and consider $A \circ f_{\mathbf{M},\pi}$ when A is drawn uniformly at random from $\mathrm{Aut}(\mathbb{F}_2^n)$. It is clear that for every fixed A the morphisms $A \circ f_{\mathbf{M},\pi}$ still form an OMM. We only need to check that they are order-independent. To see that, observe that if π and π' are two different orderings of the ground set of $\mathbf{M}$ then $f_{\mathbf{M},\pi}$ and $f_{\mathbf{M},\pi'}$ are two morphisms $\mathbf{M} \to \mathbb{F}_2^n$. By the previous lemma, there is $A_0 \in \mathrm{Aut}(\mathbb{F}_2^n)$ such that $f_{\mathbf{M},\pi} = A_0 \circ f_{\mathbf{M},\pi'}$. Now, the distribution of $A \circ f_{\mathbf{M},\pi}$ for a random $A \sim \mathrm{Aut}(\mathbb{F}_2^n)$ is the same distribution as $A \circ A_0 \circ f_{\mathbf{M},\pi'}$ which is the same distribution of $A \circ f_{\mathbf{M},\pi'}$, since $A_0 \circ A$ is also uniformly distributed over $\mathrm{Aut}(\mathbb{F}_2^n)$. □

Extending to Copies In the following section we will be needing an online matroid embedding. For that reason, we need to extend the last two theorems to deal with copies. The extension is rather simple: we only need to observe that a matroid autormorphism of the commplete binary matroid with n copies of each element $(\mathbb{F}_2^n)_{[n]}$ can be decomposed into an automorphism $A \in \mathrm{Aut}(\mathbb{F}_2^n)$ and indexings of the identities of the copies.

Lemma 4.5. *If $f \in Aut((\mathbb{F}_2^n)_{[n]})$ then there exists $A \in Aut(\mathbb{F}_2^n)$ and indexings $\sigma_u : [n] \to [n]$ for each $u \in \mathbb{F}_2^n$ such that $f((u, j)) = (Au, \sigma_u(j))$*

Proof. Let id be the identity map and $\phi : (\mathbb{F}_2^n)_{[n]} \to \mathbb{F}_2^n$ the natural projection. Now, $\phi \circ f$ and ϕ are two matroid morphisms from $(\mathbb{F}_2^n)_{[n]} \to \mathbb{F}_2^n$ so by Theorem 4.3 there is $A \in \mathrm{Aut}(\mathbb{F}_2^n)$ such that $\phi \circ f = A \circ \phi$ (see the commutative diagram below). This means in particular that $f((u, j)) = (Au, \sigma_u(j))$ for some indexings σ_u.



□

With that, Theorem 4.4 automatically extends to the matroid with copies $(\mathbb{F}_2^n)_{[n]}$ by taking a random automorphism from group $\mathrm{Aut}((\mathbb{F}_2^n)_{[n]})$.

Remark A (Single Orbit Morphisms) Theorem 4.4 is not particular to binary matroids. The only fact it uses is that all the morphisms of the OMM belong to the same orbit under the action of the automorphism group Aut(**BigM**). We can state it more generally as follows. The proof is identical to Theorem 4.4, so we omit it here.

Theorem 4.6 (Generalization of Theorem 4.4)**.** *Let $f_{\boldsymbol{M},\pi}$ be an OMM of class $\mathcal{C}$ into **BigM** such that given two orderings π and π' of the ground set of **M**, there is an automorphism $A \in Aut(\boldsymbol{BigM})$ such that $f_{\boldsymbol{M},\pi'} = A \circ f_{\boldsymbol{M},\pi}$. Then there is an order independent randomized OMM of class $\mathcal{C}$ into **BigM**.*

Remark B (Other Fields) In this section we repeatedly use the fact that $\mathbb{F}_2$ has only one non-zero constant, so whenever we identify a circuit in the matroid, we know exactly what is the linear dependency between the elements in the corresponding vector field. This is no longer true even in slightly larger fields like $\mathbb{F}_3$. If vectors $u, v, w \in \mathbb{F}_3^n$ form a circuit, it could be that: $w = \pm u \pm v$ in the representation. As a consequence, given two matroid morphisms $f, g : \mathbf{M} \to \mathbb{F}_3^n$ there may not exist a vector-space automorphism A of $\mathbb{F}_3^n$ such that $f = Ag$. In the previous example, if $f(u) = g(u)$, $f(v) = g(v)$ but $f(w) = f(u) + f(v)$ but $g(w) = g(u) - g(v)$ no such automorphism can exist.

5 OMEs and the Matroid Secretary Problem

In this section, we use OMEs to explore the complexity of the matroid secretary problem (MSP) on binary matroids. We consider three versions of the problem, each making different assumptions on the data generation process and what's known to the algorithm. We will use OMEs to establish equivalences between these problems, showing that seemingly simpler problems are actually equivalent to harder ones (up to constants).

Three Versions of the MSP. We consider the following three versions of the MSP, and aim to establish the relations in Figure 1. In all three variants, the goal is an algorithm for selecting elements that form an independent set, and whose combined weight is in expectation an α-approximation to the weight of the optimal basis.

- **Online-revealed-matroid MSP:** In this version of the problem, there is an underlying matroid $\mathbf{M}$, which is a priori unknown to the algorithm. The algorithm has only access to the number of elements $n = |\mathbf{M}|$ and to a promise that $\mathbf{M} \in \mathcal{C}$ for a class of matroids $\mathcal{C}$. For each element $u \in \mathbf{M}$, an adversary determines a weight $w_u \in \mathbb{R}_+$. The algorithm processes pairs (u, w_u) in random order at each time, but it only knows the rank function restricted to the subset of elements that have already arrived. Upon seeing the element, the algorithm must irrevocably decide whether to accept that element or not, subject to the constraint that the set of accepted elements must be an independent set of $\mathbf{M}$.
- **Known-matroid MSP:** In this version, the matroid $\mathbf{M}$ is known to the algorithm ahead of time. The only information missing is the weight of each element, which is again chosen

adversarially. As before elements arrive in random order, and the algorithm must make immediate accept/reject decisions, with the restriction that the chosen set of elements must be an independent set of $\mathbf{M}$.

- **Prophet MSP:** In this version, the matroid $\mathbf{M}$ is again known to the algorithm ahead of time. However, this time the weight of each element is drawn from a known distribution $\mathcal{D}$, which can potentially sample weights in a correlated manner. As in the other versions elements are then presented to the algorithm in random order, and the algorithm aims to select an independent set of high weight in an online manner.

Note that the first version is clearly harder than the second and the second version is clearly harder than the third, in the sense that an α-approximation to the harder problem immediately implies an α-approximation to the simpler one. We derive approximate "inverses" of these comparisons from the existence of OMEs, even if we restrict prophet MSP to pairwise-independent distributions.

Our Reductions. We first use OMEs to show an (essentially exact) "inverse" of the comparison between the online-revealed-matroid MSP and the known-matroid MSP, implying that for binary matroids the latter is as hard as the former.

Theorem 5.1. *If a class $\mathcal{C}$ of matroids admits a randomized order-independent online matroid embedding into matroid $\boldsymbol{BigM}$, then an α-approximation to the known-matroid MSP for $\boldsymbol{BigM}_{[n]}$ implies that for every $\epsilon > 0$ there is a $(\alpha - \epsilon)$-approximation to the online-revealed-matroid MSP for $\mathcal{C}$.*

For the case of binary matroids that we previously discussed, the matroids $\mathbf{BigM}$ and $\mathbf{BigM}_{[n]}$ themselves are binary in which case we can obtain the following corollary:

Corollary 5.2. *For binary matroids, there is no gap in approximability between the known-matroid MSP and the online-revealed-matroid MSP.*

As our second result, we use the existence of OMEs into a 2-transitive host matroid as a tool to establish the approximate equivalence of known-matroid MSP and prophet MSP with pairwise-independent distributions.

Theorem 5.3. *Suppose a class $\mathcal{C}$ of matroids admits a randomized order-independent online matroid embedding into matroid $\boldsymbol{BigM}$ which is 2-transitive. Then an α-approximation to the prophet MSP with pairwise-independent weight distributions for $\boldsymbol{BigM}_{[k]}$, implies that for some constant $C > 0$ there is a $C \cdot (\alpha - o(1))$-approximation to the known-matroid MSP for $\mathcal{C}$.*

Noting that the full binary matroid is 2-transitive, and there exists an order-independent OME from the class of binary matroids into a full binary matroid, we obtain the following corollary.

Corollary 5.4. *For binary matroids, there is a constant-factor gap in approximability between the known-matroid MSP and the prophet MSP with pairwise-independent weight distributions.*

We note that we can also chain the two reductions, and this way relate the online-revealed-matroid MSP to the Prophet MSP with pairwise-independent distributions. The rest of this section is devoted to the proofs of the reductions.

5.1 Proof of Theorem 5.1

Let $\mathbf{M}$ be the unknown matroid that is revealed to the algorithm in an online fashion and let $f_{\mathbf{M},\pi}$ be an order-independent randomized OME into $\mathbf{BigM}$. Let $n = |\mathbf{M}|$, $N = |\mathbf{BigM}|$ and d, k be two larger integers (to be specified later) where k is a multiple of d.

An instance of the MSP consists of a sequence of weighted elements from $\mathbf{M}$ that are presented to the algorithm in random order. Our goal is to map it on the fly to a random instance of the MSP on $\mathbf{BigM}_{[k]}$. The main difficulty is, as usual, doing it online and preserving the random order. Our strategy will be to first provide an *offline reduction* which preserves random order and obtains the desired approximation, but can't be implemented online. After that we will provide a *mostly-online implementation* of this reduction, i.e. a procedure that samples from the same distribution generated by the offline reduction and that with $1 - \epsilon$ probability can be implemented online. With the remaining ϵ probability, the process *raises a flag*. Raising a flag will indicate that from that point on, the reduction can no longer be implemented online. Algorithmically, we will stop the algorithm whenever we raise a flag and obtain zero reward. Finally, we will show that the probability of raising a flag is very small for large values of k and d.

Offline Reduction We will view a weighted element of $\mathbf{M}$ as a pair (u, w_u) with $u \in \mathbf{M}$ and $w_u \in \mathbb{R}_+$. In the offline reduction, we assume we have access to the entire matroid $\mathbf{M}$ and the entire sequence of weights. Now, we will produce a distribution of instances of $\mathbf{BigM}_{[k]}$ as follows.

For each element v in $\mathbf{BigM}$ sample k different i.i.d. timestamps t_{vj} for $j \in [k]$ from the Uniform($[0, 1]$) distribution. Those timestamps specify the arrival time of each of the k copies of the elements in $\mathbf{BigM}$ and induce a random ordering over the ground set of $\mathbf{BigM}_{[k]}$. For the matroid $\mathbf{M}$, sample a random embedding $f : \mathbf{M} \to \mathbf{BigM}$ from the OME. (Since the embedding is order independent, we don't need to know the arrival order of elements in $\mathbf{M}$ to sample such embedding). For each $u \in \mathbf{M}$, pick a random copy of $f(u)$ and set its weight to w_u. For the remaining elements, set the weight equal to zero.

This random input is clearly in random order as it is equivalent to starting with k copies of the elements of $\mathbf{BigM}$ where all but one copy has weight zero if that corresponds to an element of $\mathbf{M}$ and randomly permuting those elements. Now, feed this instance to the α-competitive algorithm for the MSP on $\mathbf{BigM}_{[k]}$. From the set selected by the algorithm, discard any element with zero weight chosen by the algorithm. The elements with non-zero weight chosen in $\mathbf{BigM}$ correspond to an independent set of $\mathbf{M}$ with the same weight. Hence, in expectation, we select an α-approximation to the optimal basis of $\mathbf{M}$.

Mostly-online Implementation The drawback of the previous reduction is that it can't be implemented online as we are assuming we know everything in advance. We will describe the same sampling procedure in a way that with high probability we can generate the instance as we go. In the sampling procedure, we will also define an event *raise a flag* which will mean that we can't generate that instance online as we learn the structure of the matroid $\mathbf{M}$.

The process will again start by sampling i.i.d. timestamps t_{vj} for $v \in \mathbf{BigM}$ and $j \in [k]$ from Uniform($[0, 1]$). In addition, we will also sample n additional timestamps from Uniform($[0, 1]$) sort them in increasing order and denote the sorted list by $T_1, \ldots, T_n$.

We will now divide the interval $[0, 1]$ into intervals $I_i = [\frac{i-1}{d}, \frac{i}{d})$ for $i \in [d]$. If more than one timestamp T_s falls in the same interval I_i, we will raise a flag. We will count how many of the timestamps $t_{v,j}$ fall in each interval:

$$X_{vi} = |\{j \in [k]; t_{vj} \in I_i\}|$$

With that, also define:

$$A_{vi} = \min\left(\frac{k}{d}, X_{vi}\right) \qquad B_{vi} = \max\left(0, X_{vi} - \frac{k}{d}\right)$$

Now, process the elements of the matroid $\mathbf{M}$ according to order π (which will be sampled at random). As we process the s-th element $u = \pi(s) \in \mathbf{M}$, we will map it to an element in $v = f_{\mathbf{M},\pi}(u) \in \mathbf{BigM}$ using the randomized OME. Now, we will apply the following procedure to choose a copy of v in $\mathbf{BigM}_{[k]}$ to assign weight w_u:

- find the interval I_i containing T_s.
- with probability $A_{vi}d/k$, choose one of the X_{vi} timestamps t_{vj} in interval I_i
- with remaining probability (if any), raise a flag and choose a different interval $I_{i'}$ with probability proportional to $B_{vi'}$ and choose a timestamp t_{vj} in that interval.

We assign weight w_u to the element with the chosen timestamp and zero weight to others. Now, we will show the following facts.

Lemma 5.5. *The* mostly online implementation *samples sequences with the same probability as the* offline reduction*.*

Proof. Observe that if we ignore the weights, the order of the elements of $\mathbf{BigM}$ is the same in both processes since they are determined by the timestamps t_{vj}. What we are left to argue is that for each v we select uniformly random timestamp t_{vj} to assign the non-negative weight. For that, observe that since $f_{\mathbf{M},\pi}$ is order independent, it has the same distribution as if we first sampled a monomorphism $f : \mathbf{M} \to \mathbf{BigM}$, and then we sampled an independent uniform indexing π for the arrival order of the elements in $f(\mathbf{M})$. This implies that when we assign the timestamps $T_1, \ldots, T_n$ according to π, the resulting distribution is the same as if we assigned i.i.d. Uniform$[0,1]$ timestamps T_v to each element $v \in f(\mathbf{M})$, and therefore, the interval T_v lands in is uniformly chosen and independent across elements $v \in f(\mathbf{M})$. Now fix a certain timestamp t_{vj} and let I_i be the interval containing it. We will show that the probability that this timestamp is selected is exactly $1/k$.

Consider two cases: either $X_{vi} \leq k/d$ in which case the probability of sampling t_{vj} is the probability that the timestamp T_v is in I_i (which is $1/d$), times the probability we decide to sample a timestamp inside I_i (which is $X_{vi}d/k$), times the probability that out of those, we choose t_{vj} (which is $1/X_{vi}$). The total probability is:

$$\frac{1}{d} \cdot \frac{X_{vi}d}{k} \cdot \frac{1}{X_{vi}} = \frac{1}{k}$$

In the case where $X_{vi} > k/d$, then it is possible that we sample t_{vj} also when T_v is outside I_i. The probability that we sample t_{vj} and T_v is in I_i is:

$$\frac{1}{d} \cdot 1 \cdot \frac{1}{X_{vi}} = \frac{1}{dX_{iv}}$$

The probability that we sample when it is outside is the probability that we choose a different interval $I_{i'}$, raise a flag and then move to interval I_i, which is:

$$\sum_{i'} \frac{1}{d}\left(1 - \frac{A_{vi'}d}{k}\right) \cdot \frac{B_{vi}}{\sum_{i'} B_{vi'}} \cdot \frac{1}{X_{vi}} = \frac{1}{k}\frac{(k - \sum_{i'} A_{vi'})}{\sum_{i'} B_{vi'}} \cdot \frac{B_{vi}}{X_{vi}} = \frac{1}{k}\frac{B_{vi}}{X_{vi}}$$

because $\sum_{i'} A_{vi'} + \sum_{i'} B_{vi'} = \sum_{i'} X_{vi'} = k$. Taking those two probabilities together, we have:

$$\frac{1}{dX_{iv}} + \frac{1}{k}\frac{B_{vi}}{X_{vi}} = \frac{1}{k}.$$

□

Lemma 5.6. *If no flags were raised, we can produce the instance on the fly as we process* $\boldsymbol{M}$*.*

Proof. Let $i_1 < \ldots < i_n$ be the indices of the intervals such that $T_s \in I_{i_s}$. Since no flag was raised, then each T_s landed in a different interval and the s-th element that arrives from matroid $\mathbf{M}$ is mapped to a copy inside I_{i_s}. This enables the following online reduction: once the s-th element arrives we can decide the weights of all the elements in intervals $I_{i_{s-1}+1}$ to I_{i_s} and feed to the MSP algorithm for $\mathbf{BigM}_{[k]}$. In this sub-sequence there will be at most one element on non-zero weight which corresponds to the arriving element of $\mathbf{M}$. We can observe if that element was selected in $\mathbf{BigM}_{[k]}$ and if so, we can select it in $\mathbf{M}$. □

Lemma 5.7. *For any* n *and* ϵ*, there are large enough* k *and* d*, such that the probability that we raise a flag is at most* ϵ*.*

Proof. The first event in which we raise a flag is when two timestamps T_s land in the same interval. The probability that this happens is at most n^2/d. Now, note that for each interval i and each of the n elements v in $\mathbf{BigM}$ that have non-zero weights, we have by the Chernoff bound that:

$$\mathbb{P}\left(X_{vi} \leq (1-\delta)\frac{k}{d}\right) \leq \exp\left(-\frac{\delta^2 k}{2d}\right)$$

Hence with probability at most $nd\exp\left(-\frac{\delta^2 k}{2d}\right)$, the timestamps t_{vj} are such that the probability we raise a flag when we try to choose a timestamp in the same interval as T_s is more than δn. Taking the union bound of those events, we get:

$$\frac{n^2}{d} + nd\exp\left(-\frac{\delta^2 k}{2d}\right) + \delta n$$

Taking $\delta = \epsilon/(3n)$, $d = 3n^2/\epsilon$ and k large enough, we get that the total probability of raising a flag is at most ϵ. □

Taking those lemmas together, we can conclude the proof of Theorem 5.1. For that, let Alg be an α-competitive algorithm for $\mathbf{BigM}_{[k]}$ and let Y represent the sequence of the MSP sampled by the offline reduction. Let's represent by $\mathsf{Alg}(Y)$ the weight of the elements selected by Alg and Opt the weight of the optimal basis. By the fact that the offline reduction produces an instance in random order, we know that $\mathbb{E}[\mathsf{Alg}(?'Y)] \geq \alpha\mathsf{Opt}$.

Our online reduction, will attempt to construct Y on the fly. If we raise the flag, we will stop the algorithm and pretend we had zero reward. If not, we will continue the reduction and collect $\mathsf{Alg}(Y)$ reward. We will denote by Flag the event that the flag was raised and by $\overline{\mathsf{Flag}}$ its complement. Our total reward will be:

$$\mathbb{E}[\mathsf{Alg}(Y)\cdot\mathbf{1}\{\overline{\mathsf{Flag}}\}] = \mathbb{E}[\mathsf{Alg}(Y)] - \mathbb{E}[\mathsf{Alg}(Y)\cdot\mathbf{1}\{\mathsf{Flag}\}] \geq \mathbb{E}[\mathsf{Alg}(Y)] - \mathsf{Opt}\cdot\mathbb{P}[\mathsf{Flag}] \geq (\alpha-\epsilon)\mathsf{Opt}.$$

5.2 Proof of Theorem 5.3

Let $\mathbf{M}$ be the unknown matroid with $|\mathbf{M}| = n$ that is revealed to the algorithm in an online fashion, which admits an order-independent randomized OME into $\mathbf{BigM}$ with $|\mathbf{BigM}| = M$. First, we obtain the following simple reduction that allows us to focus on a special class of the prophet MSP in which each element takes a weight from the set of weights W with $|W| = m = O(n^2)$. In addition, we can restrict the weight distribution such that each element has a distinct weight from the weight class.

Lemma 5.8. *If there exists an α-approximation to the prophet MSP on $\boldsymbol{M}$ with weight distribution $\mathcal{D}$ supported over the set of weights W with $\mathsf{rank}(\boldsymbol{M}) = d$, $|\boldsymbol{M}| = n$, $|W| = O(n^2)$ with $\max_{w \in W} w \leq 1$ and $\mathbb{E}_{\mathcal{D}}[\boldsymbol{OPT}(\boldsymbol{M})] \in \left[\frac{1}{16}, 1\right]$ such that for all $w \in W$, there exists at most one element assigned weight of w with probability one, then there exists an $\left(\frac{\alpha}{256} - \frac{1}{32d}\right)$-approximation to prophet MSP on matroid $\boldsymbol{M}$ with any arbitrary weight distribution.*

The proof of the above lemma simply follows from Sublemma-4.2 from [Dug21] that reduces any arbitrary prophet MSP with $O(\log(|\mathsf{rank}(\mathbf{M})|))$ many weights and $\mathbb{E}_{\mathcal{D}}[\mathbf{OPT}(\mathbf{M})] \in \left[\frac{1}{16}, 1\right]$. We then add distinct noise of the order of $O\left(\frac{1}{n^2}\right)$ to ensure that the weight of each element is distinct. The full proof of the reduction is delegated to Section A.

For simplicity, we let $W = \{w_1, \ldots, w_m\}$ and consider the prophet MSP on matroid $\mathbf{M}$ and weight distribution $\mathcal{D}$ supported over the set of weights W satisfying the conditions from Lemma 5.8.

Extending BigM with Copies We let $\mathbf{BigM}_{[m \cdot N]}$ be a matroid with $m \cdot N$ parallel copies of each element of $\mathbf{BigM}$ with $|\mathbf{BigM}| = M$ and integer $N = \Omega\left(2^{M^2}\right)$. We divide the set of $N \cdot m$ copies into m sets of size N, each part corresponding to weight class w_i. We use N_i to denote the set of labels corresponding to weight w_i for all $i \in [m]$ with $|N_i| = N$. We sometimes denote N_i by $[N] = \{1, 2, \ldots, N\}$ whenever it is clear from the context which w_i we are referring to. In addition, the ℓ-th copy of the weight class corresponding to weight w_i of element $\mathbf{v} \in \mathbf{BigM}$ is denoted as $\mathbf{v}^{i,\ell}$.

Reduction to "Almost" Pairwise Independent Prophet MSP We first define the weight distribution $\mathcal{D}^*$ over $\mathbf{BigM}_{[m \cdot N]}$ in Definition 1, which is "almost" pairwise independent. Then in Theorem 5.9, we show an existence of exact pairwise independent weight distribution $\tilde{\mathcal{D}}$ over $\mathbf{BigM}_{[m \cdot N]}$ which is "close" to the distribution defined in Definition 1 in total-variation distance. This allows us to utilize the fact that any algorithm $\mathcal{A}$ can not distinguish between the almost pairwise independent weight distribution $\mathcal{D}^*$ and $\tilde{\mathcal{D}}$ with high probability. Finally, we complete the proof of Theorem 5.3.

We begin by defining the almost pairwise independent weight distribution over $\mathbf{BigM}_{[m \cdot N]}$.

Definition 1 (Almost P.W. Independent Distribution)**.** *Consider the weight distribution $\mathcal{D}^*$ over the elements of $\boldsymbol{BigM}_{[m \cdot N]}$ defined as follows:*

1. *Given an order independent OMM $f' : \boldsymbol{M} \to \boldsymbol{BigM}$, we sample a random automorphism $f'' \in \mathrm{Aut}(\boldsymbol{BigM})$ and obtain an order independent matroid morphism $f = f'' \circ f'$.*
2. *For any $\mathbf{v} \in \boldsymbol{M}$ with $w(\mathbf{v}) = w_i$, let $\mathbf{u} = f(\mathbf{v})$. We sample $\ell \sim \mathrm{Unif}(N_i)$ and assign the weight of $w(\mathbf{u}^{i,\ell}) = w_i$.*
3. *We assign the weight of the rest of the elements of $\boldsymbol{BigM}_{[m \cdot N]}$ to be zero.*

We first observe that when N is much larger than M, the distribution in the above definition is almost pairwise independent. To see this, we first observe that for any $i \in [m], \ell \in N_i$ and $\mathbf{u} \in \mathbf{BigM}$, $\mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ can potentially either take a weight of w_i or zero. For simplicity, now consider any two distinct elements $\mathbf{u}_{i,\ell}, \mathbf{u}'_{j,\ell'} \in \mathbf{BigM}_{[m \cdot N]}$ and weight distribution $\mathcal{D}$ such that there always exists a pair of elements $\mathbf{v}, \mathbf{v}'' \in \mathbf{M}$ that are assigned weights of w_i, w_j, respectively. Since f' is a random automorphism $\mathrm{Aut}(\mathbf{BigM})$, we have $\Pr[w(\mathbf{u}_{i,\ell}) = w_i] = \Pr[f(\mathbf{v}) = \mathbf{u}] \cdot \frac{1}{N} = \frac{1}{M \cdot N}$. On the other hand, we have

$$\Pr[w(\mathbf{u}_{i,\ell}) = w_i \wedge w(\mathbf{u}_{j,\ell}) = w_j] = \Pr[f'(\mathbf{v}) = \mathbf{v} \wedge f'(\mathbf{v}') = \mathbf{u}'] \cdot \frac{1}{N^2} = \frac{1}{M(M-1)} \cdot \frac{1}{N^2},$$

which is close to the product $\Pr[w(\mathbf{u}_{i,\ell}) = w_i] \cdot \Pr[w(\mathbf{u}_{j,\ell}) = w_j] = \frac{1}{M^2 \cdot N^2}$. However, in general, we can not guarantee that $\mathcal{D}$ will always assign weights w_i, w_j to some pair of elements of $\mathbf{v}, \mathbf{v}'$ of $\mathbf{M}$. In addition, the above argument also fails if we have $i = j$ or $\mathbf{u} = \mathbf{u}'$ as in both of these cases, $\Pr[w(\mathbf{u}_{i,\ell}) = w_i \wedge w(\mathbf{u}_{j,\ell}) = w_j] = 0$. Intuitively, we can circumvent these pairwise correlation issues by taking N large enough as it makes pairwise correlations small enough. More precisely,

$$|\Pr[w(\mathbf{u}_{i,\ell}) = w_i \wedge w(\mathbf{u}_{j,\ell}) = w_j] - \Pr[w(\mathbf{u}_{i,\ell}) = w_i] \cdot \Pr[w(\mathbf{u}_{j,\ell}) = w_j]| = O\left(\frac{1}{M^2 \cdot N^2}\right).$$

Using this observation, we prove the following technical theorem.

Theorem 5.9. *For $N = \Omega\left(2^{M^2}\right)$, let weight distribution $\mathcal{D}^*$ over $\boldsymbol{BigM}_{[m \cdot N]}$ be defined as in Definition 1, then there exist a pairwise-independent weight distribution $\tilde{\mathcal{D}}$ over $\boldsymbol{BigM}_{[m \cdot N]}$ such that $\mathrm{TV}_{\mathcal{D}^*, \tilde{\mathcal{D}}} \leq O\left(\frac{m^3}{M}\right)$.*

The proof of the above theorem is highly technical and constructs an explicit $\tilde{\mathcal{D}}$ by a sequence of small perturbations to $\mathcal{D}^*$. For the sake of the uninterrupted flow of the presentation, we delegate it to Appendix B. We emphasize that the choice of $N = \Omega(2^{M^2})$ is required due to the limitations of the techniques developed to prove Theorem 5.9. We conjecture that one can prove the similar theorem for $N = \Omega(\mathrm{Poly}(M))$, which we leave as an intriguing technical open problem.

Proof of Theorem 5.3 We now complete the proof of the main theorem.

Proof of Theorem 5.3. To prove the main theorem, we first prove the following: if there exists an α-approximate algorithm to the prophet MSP instance with a pairwise independent weight distribution for the matroid $\mathbf{BigM}_{[m \cdot N]}$, then there exists an $(\alpha - o(1))$-approximate algorithm for the prophet MSP instance for the matroid $\mathbf{M}$ with weight distribution $\mathcal{D}$ supported over the set of weights W with $\mathsf{rank}(\mathbf{M}) = d$, $|\mathbf{M}| = n$ and $|W| = O(n^2)$ with $\max_{w \in W} w \leq 1$ and $\mathbb{E}_{\mathcal{D}}[\mathbf{OPT}(\mathbf{M})] \in \left[\frac{1}{16}, 1\right]$ such that for all $w \in W$, there exists at most one element assigned weight of w with probability one. Combining this with Lemma 5.8, we will conclude the proof of the Theorem 5.3

Given an order-independent matroid morphism $f' : \mathbf{M} \to \mathbf{BigM}$, we sample a random automorphism $f'' \in \mathrm{Aut}(\mathbf{BigM})$ and obtain an order-independent matroid morphism $\tilde{f} = f'' \circ f'$. Given $\tilde{f} : \mathbf{M} \to \mathbf{BigM}$, we obtain $f : \mathbf{M} \to \mathbf{BigM}_{[m]}$ that maps each $\mathbf{v} \in \mathbf{M}$ to $\mathbf{u}^i \in \mathbf{BigM}_{[m]}$ iff $\tilde{f}(\mathbf{v}) = \mathbf{u}$ and $w(\mathbf{v}) = w_i$. We then consider $\mathbf{BigM}_{[m \cdot N]}$, i.e. matroid $\mathbf{BigM}_{[m]}$ with N many copies of each element.

We first consider an offline reduction as follows: for all $\mathbf{v}^i \in \mathbf{BigM}_{[m]}$, sample N many independent arrival times from $\mathrm{Unif}[0, 1]$ denoting the uniformly random arrival times of elements of

$\mathbf{BigM}_{[m\cdot N]}$. Given any $\mathbf{v} \in \mathbf{M}$ and pair $(\mathbf{v}, w(\mathbf{v}) = w_i)$, with $\mathbf{u}^i = f(\mathbf{v})$, let $w(\mathbf{u}^{i,\ell}) = w_i$ uniformly random from $\ell \in [N]$.

Since f does not require any information about the arrival order, the above-described offline reduction is a valid instance of matroid prophet secretary over $\mathbf{BigM}_{[m\cdot N]}$. In addition, the induced weight assignment over $\mathbf{BigM}$ due to offline reduction is identical to the distribution $\mathcal{D}^*$ defined in Definition 1.

Given the uniformly random arrival of elements of $\mathbf{M}$, we construct an "almost online implementation" of the above offline reduction similar to the proof of Theorem 5.1. We let N be large enough ($\Omega(n^2/\varepsilon)$) such that the probability of "almost online implementation" raising a flag is at most ε. In fact, for the proof of Theorem 5.9, we let $N = \Omega(2^{M^2})$ which satisfies the required condition.

We let $\mathcal{A}$ be an α-approximate algorithm for the pairwise-independent prophet MSP on matroid $\mathbf{BigM}_{[m\cdot N]}$. Since there exists a pairwise independent weight distribution $\tilde{\mathcal{D}}$ over $\mathbf{BigM}$ within the total variation distance of $O\left(\frac{m^3}{M}\right)$, the algorithm $\mathcal{A}$ can not distinguish the weight distribution $\mathcal{D}^*$ from $\tilde{\mathcal{D}}$ with probability at least $1 - O\left(\frac{m^2}{M}\right)$.

Let $\mathcal{E}$ be the event when the algorithm $\mathcal{A}$ can not distinguish between $\mathcal{D}^*$ and $\tilde{D}$. We note that when event $\mathcal{E}$ does not hold, the offline optimal can be bounded by

$$\mathbb{E}[\mathbf{OPT}(\mathbf{BigM}_{m\cdot N}) \mid \mathcal{E}^c] \le \mathsf{rank}(\mathbf{BigM}_{[m\cdot N]}) \cdot \max_{w_i \in W} w_i \le n \cdot 1 = n,$$

where the second inequality follows because $\mathsf{rank}(\mathbf{BigM}_{[m\cdot N]}) = n$ and $w_i \le 1$ for all $i \in [m \cdot N]$.

Now, let S be the selected set of elements of $\mathbf{BigM}$ by $\mathcal{A}$ w.r.t. weight distribution $\mathcal{D}^*$. We can bound,

$$\begin{aligned}
\mathbb{E}[w(S)] &= \mathbb{E}[w(S) \mid \mathcal{E}] \cdot \Pr[\mathcal{E}] + \mathbb{E}[w(S) \mid \mathcal{E}^c] \cdot \Pr[\mathcal{E}^c] \\
&\le \mathbb{E}[w(S) \mid \mathcal{E}] + \mathbb{E}[\mathbf{OPT}(\mathbf{BigM}_{[M\cdot m]}) \mid \mathcal{E}^c] \cdot \frac{m^3}{M} \\
&\le \mathbb{E}[w(S) \mid \mathcal{E}] + \frac{m^3 \cdot n}{M}.
\end{aligned}$$

Above, the first inequality holds because $\Pr[\mathcal{E}^c] \le \frac{m^2}{M}$ and the second inequality holds because $\mathbb{E}[\mathbf{OPT}(\mathbf{BigM}_{[m\cdot N]}) \mid \mathcal{E}^c] \le n$. Due to our reduction, the performance of the algorithm on the original Prophet MSP instance $\mathbb{F}'$ is lower bounded by $\mathbb{E}[w(S) \mid \mathcal{E}]$, next we lower bound the expectation $\mathbb{E}[w(S) \mid \mathcal{E}]$,

$$\begin{aligned}
\mathbb{E}[w(S) \mid \mathcal{E}] &\ge \mathbb{E}[w(S)] - \frac{m^3 \cdot n}{M} \\
&\ge \alpha \cdot \mathbb{E}_{\tilde{\mathcal{D}}}[\mathbf{OPT}(\mathbf{BigM}_{[m\cdot N]})] - o(1) \cdot \mathbf{OPT}(\mathbf{M}) \\
&\ge \mathbf{OPT}(\mathbf{M}) \cdot \left(\alpha - \frac{\alpha \cdot m^3}{M}\right) - o(1) \cdot \mathbf{OPT}(\mathbf{M}) \ge (\alpha - o(1)) \cdot \mathbf{OPT}(\mathbf{M}).
\end{aligned}$$

Above, the second inequality holds because $\mathcal{A}$ is an α approximate algorithm for Prophet MSP with pairwise independent prior and $\frac{m^3 \cdot n}{M} = O(m^4/2^m) = o(1)$. The third inequality holds because $\mathrm{TV}_{\mathcal{D}^*,\tilde{\mathcal{D}}} \le O(m^3/M)$. Finally, since the performance of the reduction on the original matroid secretary is

$$\mathbb{E}[w(S) \mid \mathcal{E}] \cdot \Pr[\text{Reduction does not Flag}] \ge (1 - \varepsilon) \cdot (\alpha - o(1)) \cdot \mathbf{OPT}(\mathbf{M}).$$

This concludes the proof. □

6 OMEs Beyond Binary Matroids

We next explore OMEs beyond binary matroids. In Section 6.1, we show that there cannot be an OME that embeds graphical matroids into graphical matroids or, more generally, regular matroids. Then, in Section 6.2, we give an OME that embeds laminar matroids into laminar matroids. Finally, in Section 6.3, we show that there is no universal host matroid, that allows embedding of all matroids on n elements of a given rank.

6.1 Graphic and Regular Matroids

For graphic matroids, there exists an elegant $2e$-approximation algorithm for the MSP in the known-matroid case by Korula-Pál [KP09]. Their algorithm assumes that when an edge arrives, the algorithm learns the pair of vertices it connects as well as the weight. In other words, the algorithm processes in each step $(u, v), w_{uv}$ where u and v are vertices. Even though the full graph is not known in advance, there is enough information about the graph structure to randomly decompose the problem into instances of the single-item secretary problem.

In the online-revealed matroid case, the algorithm has only access to an oracle that tells which subsets of previously arrived edges are independent (contain no cycles). For example, if 3 edges arrive and the algorithm knows that they are all independent, it is impossible for the algorithm to know if they form a path, a star or if they share not endpoints (recall Figure 2). Korula-Pál heavily relies on having vertex information and doesn't easily extend to this model. In fact, we are not aware of any $O(1)$-approximation algorithm for the MSP for graphic matroids in the online-revealed-matroid setting.

In the remainder of this section we investigate whether we can obtain such an algorithm using a OME. A natural idea is to try to construct an embedding where the host matroid **BigM** is itself a graphic matroid. If one could do that, it would be possible to combine the reduction in Section 5 with the Korula-Pál algorithm to obtain an algorihtm for the graphic MSP in the online-revealed-matroid setting. Unfortunately, we show below that no such embedding exists:

Theorem 6.1. *If $\mathcal{C}$ is the class of graphic matroids, there is no online matroid embedding into a host matroid **BigM** where **BigM** is also graphic.*

Proof. Assume that such embedding exists and let $G = (V, E)$ be the graph representing **BigM**. Now, consider two graphic matroids $\mathbf{M}_1$ and $\mathbf{M}_2$ represented respectively by the graphs in the left and right of Figure 3. Observe that when restricted to $\{a, b, c\}$, the matroids are identical since every non-empty subset of elements is independent. When an OMM observes the restriction to those three elements, it has no way to know in which matroid we are in, so it needs to map those three edges to the same edges of the graph G representing **BigM**. Let's denote those edges by a', b', c'. Let also d', e', f' be the edges in G that the edges of the left matroid are mapped to and let g' be the edge that the g edge in the right matroid is mapped to. Note that $\{a', b', d'\}$, $\{b', c', e'\}$ and $\{a', c', f'\}$ must form cycles in G. The only way that this is possible while keeping $\{a', b', c'\}$ independent is if edges a', b', c' all share an endpoint. However, edges a', b', c', g' must form a cycle in G as well, which is not possible if the first three edges share an endpoint. □

Graphic matroids are a special case of regular matroids, for which there exists an $O(1)$-competitive algorihtm by Dinitz-Kortsarz [DK14]. It is then tempting to construct an embedding from graphic into a host matroid **BigM** that is regular. However, that is again not possible:

Lemma 6.2. *If a host matroid **BigM** is regular and admits an online matroid embedding of all rank n graphic matroids, then **BigM** must also admit an online matroid embedding of all rank n binary matroids.*

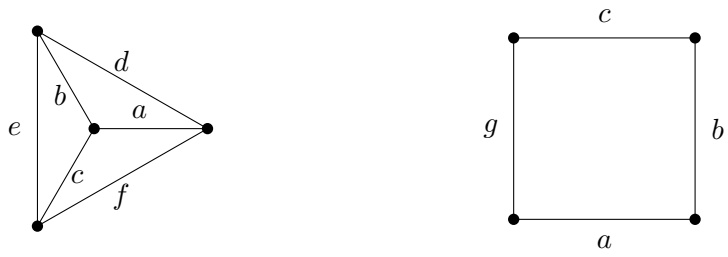


Figure 3: Two graphic matroids whose restriction to $\{a, b, c\}$ coincide.

Proof. We will show that **BigM** contains an isomorphic copy of $\mathbb{F}_2^n$. We first define for each $S \subseteq [n]$ a graphic $\mathbf{M}_S$ with ground set $[n+1]$ represented by the graph where the only cycle is formed by the edges with labels in $S \cup \{n+1\}$. Equivalently, the rank function is given by

$$\mathsf{rank}_{\mathbf{M}_S}(T) = \begin{cases} |T| - 1 & \text{for } T \supseteq S \cup \{n+1\} \\ |T| & \text{otherwise} \end{cases}$$

Let also $\mathbf{M}'$ be the matroid on $[n]$ such that $\mathsf{rank}_{\mathbf{M}'}(T) = |T|$ for all subsets T. Assuming all ground serts are ordered according to the labels, note that $\mathbf{M}'$ is a prefix of all matroids $\mathbf{M}_S$. Now, let $b_1, \ldots, b_n \in \mathbf{BigM}$ be the image of the ground set of $\mathbf{M}'$ by the online embedding. Since **BigM** is regular, it admits an $\mathbb{F}_2$-representation, so we can think of $b_1, \ldots, b_n$ as linearly independent vectors in $\mathbb{F}_2^n$.

Since this is an online embedding, it can be extended to a morphism $\mathbf{M}_S \to \mathbf{BigM}$ for each $S \subseteq [n]$. So, for each S, let b_S be the element in **BigM** $[n+1]$ maps to. If we view b_S as an $\mathbb{F}_2$-vector, we must have $b_S = \sum_{i \in S} b_i$ in $\mathbb{F}_2^n$. As a consequence, the vectors $\{b_S; S \subseteq [n]\}$ form an isomorphic copy of $\mathbb{F}_2^n$. □

We can now derive the following theorem as as corollary:

Theorem 6.3. *If $\mathcal{C}$ is the class of graphic matroids, there is no online matroid embedding into a host matroid **BigM** where **BigM** is regular.*

Proof. By the previous lemma, if such embedding exists then **BigM** must contain an isomorphic copy of $\mathbb{F}_2^n$. However, $\mathbb{F}_2^n$ contains an isomorphic copy of the Fano plane which is not representable over $\mathbb{F}_3$ [Tut58]. Hence **BigM** can't be regular. □

The reader will notice that Theorem 6.1 can be derived as a trivial corollary of Theorem 6.3 since graphic matroids are regular. Nevertheless, we find the more direct proof of Theorem 6.1 enlightening and opted to keep it.

6.2 Laminar Matroids

Another important class of matroids that admits an online embedding is the class of laminar matroids. We begin this section by recalling the definition of a laminar matroid and proving a useful structural lemma about them. We then present an OMM for the class of laminar matroids with at most n elements, using a host matroid **BigM** which is a complete linear matroid of rank n over any field with sufficiently many elements.

Lemma 6.4. *If $f : \boldsymbol{M} \to \boldsymbol{N}$ is a matroid monomorphism and $\boldsymbol{N}$ is laminar, then $\boldsymbol{M}$ is laminar as well.*

Proof. Denote the ground sets of $\mathbf{M}, \mathbf{N}$ by M, N, respectively. If $\mathcal{A}$ is a laminar family of subsets of N such that $\mathbf{N}$ is a laminar matroid with respect to the function $c : \mathcal{A} \to \mathbb{Z}_+$, then the family of sets $f^{-1}(\mathcal{A})$ consisting of the sets $f^{-1}(A)$ for each $A \in \mathcal{A}$ is a laminar family, and $\mathbf{M}$ is a laminar matroid with respect to the function $\tilde{c} : f^{-1}(\mathcal{A}) \to \mathbb{Z}_+$ defined by

$$\tilde{c}(B) = \min\{c(A); A \in \mathcal{A} \text{ and } B = f^{-1}(A)\}. \qquad \square$$

Let us also recall the definitions of the span and flat in a matroid. Given a subset S of the ground set of a matroid $\mathbf{M}$, we define the $\mathsf{span}_{\mathbf{M}}(S) = \{x \in \mathbf{M}; \mathsf{rank}_{\mathbf{M}}(S \cup \{x\}) = \mathsf{rank}_{\mathbf{M}}(S)\}$. From the properties of the rank function, this is the equivalent to the maximal set containing S that has the same rank as S. We say that a subset S is a flat if $S = \mathsf{span}(S)$. For the complete K-representable matroid K^n, the span coincides with the notion of the linear span of vector fields and the flats are linear subspaces.

The span of laminar matroids has the following useful property:

Theorem 6.5 (Fife and Oxley [FO17])**.** *A matroid is laminar if and only if, for all circuits C_1 and C_2 with $C_1 \cap C_2 \neq \emptyset$, either $\mathsf{span}(C_1) \subseteq \mathsf{span}(C_2)$ or $\mathsf{span}(C_2) \subseteq \mathsf{span}(C_1)$.*

Theorem 6.6. *Let $\mathcal{C}$ be the class of laminar matroids of at most n elements and $\mathbb{F}$ be a field with at least 2^n elements. Then there is an OMM from $\mathcal{C}$ into $\mathbb{F}^n$.*

Proof. The OMM is defined inductively. If $\mathbf{M}$ is an empty matroid then $f_{\mathbf{M},\pi}$ is the trivial morphism from the empty matroid to $\mathbf{BigM} = \mathbb{F}^n$. Otherwise let $n = |\mathbf{M}|$, let $(\mathbf{M}', \pi')$ denote the restriction of $(\mathbf{M}, \pi)$ to $[n-1]$, let g denote the morphism $f_{\mathbf{M}',\pi'}$, and let $u = \pi(n)$. The morphism $f = f_{\mathbf{M},\pi}$ is defined as follows. For $u' \neq u$ we set $f(u') = g(u')$. If u doesn't belong to any circuit with previously arrived elements, define $f(u)$ to be any element outside the linear span of $f(\mathbf{M}')$. If u is in some circuit, define $A \subseteq \mathbf{M}$ to be the intersection of $\mathsf{span}(C)$ over all circuits C in $\mathbf{M}$ containing u. By Theorem 6.5, A is the span of some circuit in $\mathbf{M}$ and hence a flat. Now, choose $f(u)$ to be an element in the linear subspace $V = \mathsf{span}_{\mathbb{F}^n}(f(A \setminus \{u\}))$ that is not contained in $V_S = \mathsf{span}_{\mathbb{F}^n}(f(S))$ for every subset S of $\mathbf{M}$ such that $S \cup \{u\}$ is independent on $\mathbf{M}$.

To justify that $f(u)$ is well-defined, we must argue that the linear subspace V contains at least one element that is not in V_S for every subset S such that $S \cup \{u\}$ is independent. We will use a counting argument that consists of showing that $V_S \cap V$ is a proper linear subspace of V for every such S, and then observing that a vector space over a field with at least 2^n elements cannot be expressed as the union of 2^n or fewer proper linear subspaces. To show that $V_S \cap V$ is a proper linear subspace of V we argue by contradiction. Let C be a circuit containing u such that $A = \mathsf{span}_{\mathbf{M}}(C)$. As $S \cup \{u\}$ is independent, there must be some $u' \in C \setminus \{u\}$ such that $S \cup \{u'\}$ is also independent. Then, since $g : \mathbf{M}' \to \mathbf{BigM}$ is a matroid morphism, $g(u')$ is linearly independent of $g(S)$. In particular, $g(u') \in V \setminus V_S$ and hence $V_S \cap V$ is a proper linear subspace of V as claimed.

Having justified that f is well-defined, we must show that it is a matroid morphism. Consider any set S in $\mathbf{M}$. To show $\mathsf{rank}_{\mathbf{BigM}}(f(S)) = \mathsf{rank}_{\mathbf{M}}(S)$ we will proceed by case analysis.

1. If $u \notin S$ then $f(S) = g(S)$ and we use the induction hypothesis that $g = f_{\mathbf{M}',\pi'}$ is a matroid morphism.

2. If $u \in S$ and $\mathsf{rank}_{\mathbf{M}}(S) = \mathsf{rank}_{\mathbf{M}}(S \setminus \{u\}) + 1$ then let B be a maximal independent subset of $S \setminus \{u\}$. Since g is a matroid morphism, every $u' \in S \setminus \{u\}$ lies in the linear span of $g(B)$. Since $B \cup \{u\}$ is independent in $\mathbf{M}$, by construction $f(u)$ lies outside the linear span of $g(B)$. Hence, $\mathsf{rank}_{\mathbf{BigM}}(S) = \mathsf{rank}_{\mathbf{BigM}}(B \cup \{u\}) = |B| + 1 = \mathsf{rank}_{\mathbf{M}}(S)$.

3. If $u \in S$ and $\mathsf{rank}_{\mathbf{M}}(S) = \mathsf{rank}_{\mathbf{M}}(S \setminus \{u\})$ then there is a circuit $C \subseteq S$ containing u. The $\mathsf{span}(C \setminus \{u\})$ is a flat A' that contains u, so A' must be a superset of A by Theorem 6.5. Then we have the following chain of containments.

$$\mathsf{span}_{\mathbb{F}^n}(g((S \setminus \{u\})) \supseteq \mathsf{span}_{\mathbb{F}^n}(g((C \setminus \{u\})) = \mathsf{span}_{\mathbb{F}^n}(g((A')) \supseteq \mathsf{span}_{\mathbb{F}^n}(g(A \setminus \{u\})).$$

Since $f(u)$, by construction, belongs to $\mathsf{span}_{\mathbb{F}^n}(g(A \setminus \{u\}))$, it belongs to $\mathsf{span}_{\mathbb{F}^n}(g(S \setminus \{u\}))$ and therefore

$$\mathsf{rank}_{\mathbb{F}^n}(f(S)) = \mathsf{rank}_{\mathbb{F}^n}(g(S \setminus \{u\})) = \mathsf{rank}_{\mathbf{M}'}(S \setminus \{u\}) = \mathsf{rank}_{\mathbf{M}}(S). \qquad \square$$

Similarly to the situation of graphic matroids, we can embed laminar matroids into a large linear matroid, but there is no embedding for which **BigM** is also laminar:

Theorem 6.7. *If $\mathcal{C}$ is the class of laminar matroids of at most n elements, there is no online matroid embedding into a host matroid **BigM** where **BigM** itself is laminar.*

Proof. Given a set S let $\mathbf{U}_{n,r}(S)$ denote the uniform matroid with n elements and rank r defined on ground set S. Using this notation, consider the following three matroids on ground set $\{a, b, c, d\}$:

$$\mathbf{U}_{3,2}(\{a, b, d\}) \oplus \mathbf{U}_{1,1}(\{c\}) \qquad \mathbf{U}_{3,2}(\{a, c, d\}) \oplus \mathbf{U}_{1,1}(\{b\}) \qquad \mathbf{U}_{3,2}(\{b, c, d\}) \oplus \mathbf{U}_{1,1}(\{a\})$$

It is easy to check that all three of those matroids are laminar. Moreover, the restriction to $\{a, b, c\}$ is the matroid $\mathbf{U}_{3,3}(\{a, b, c\})$. Assume now that there is an OMM into a laminar matroid **BigM** and let a', b', c' be the elements it maps to. There must be some set in the laminar family of **BigM** that separates those elements, otherwise we can't extend it to the three matroids above. For example, in the first matroid we can't swap the roles of a and c preserving rank. Finally, there is at most one two-element subset of $\{a', b', c'\}$ that can be formed by intersecting one of the sets in the laminar family with $\{a', b', c'\}$. If the two-element subset that spans d is not that one, then you have no way of extending the embedding to include d. $\square$

6.3 Matroids of Bounded Rank

Finally, we show that there is no universal host matroid **BigM** such that all matroids of at most n elements admit an OMM into **BigM**. In fact, we show that even if we restrict to matroids of rank 3 this is not possible. The geometric intuition is that once we reach rank 3, we start being able to represent finite projective planes, where elements that haven't arrived yet impose non-trivial constraints on the already arrived elements. Before we get there, it is useful to analyze ranks 1 and 2.

Let $\mathcal{M}_{n,r}$ be the matroids of rank at most r defined on at most n elements. We will show that for $r = 1, 2$ it is trivial to construct an online matroid embedding. As usual, we will construct an OMM and convert it to an OME using Lemma 3.1.

Lemma 6.8. *There is an online matroid morphism from $\mathcal{M}_{n,1}$ into $\boldsymbol{U}_{1,1} \oplus T$.*

Proof. The matroid $\mathbf{U}_{1,1} \oplus T$ has ground set $\{0, 1\}$ and the only independent set is $\{1\}$. Given a matroid $\mathbf{M} \in \mathcal{M}_{n,1}$ and an element e of the ground set of $\mathbf{M}$ we map it to 1 if $\mathsf{rank}_{\mathbf{M}}(\{e\}) = 1$ and to 0 otherwise. The mapping is constructed online since it only depends on the rank of a single-set element consisting of the element we are processing. It is also simple to check that it is a morphism since $\mathbf{M}$ has rank 1. $\square$

Lemma 6.9. *There is an online matroid morphism from $\mathcal{M}_{n,2}$ into $\boldsymbol{U}_{n,2} \oplus T$.*

Proof. The matroid $\mathbf{U}_{n,2} \oplus T$ has ground set $\{0, 1, 2, \ldots, n\}$ and the independent sets are subsets of at most 2 elements that don't contain zero. To construct an embedding $f : \mathbf{M} \to \mathbf{U}_{n,2} \oplus T$, when we process an element e, we first check if $\mathsf{rank}_{\mathbf{M}}(\{e\}) = 0$. If so, we map it to 0. Otherwise, for every element a processed before e, we check if $\mathsf{rank}_{\mathbf{M}}(\{a, e\}) = 1$. If so, we map $f(e) = f(a)$. Otherwise, we map e to the first unused index in $[n]$. □

Once we reach rank 3, the situation becomes a lot more interesting, as there exist many non-trivial matroids like finite projective planes (e.g. Fano plane). The richness of the space of rank 3 matroids will also imply that an online matroid embedding no longer exists.

Before we prove it formally, let's give some geometric intuition. Given points on the plane $\mathbb{R}^2$, we can define a matroid of rank 3 as follows: (i) every set of one point is independent (ii) every pair of different points is independent; (iii) every triple of points is independent iff it they are not collinear; (iv) no other set is independent.

Now, consider the following online problem: we are presented with labels a, b, c, d, e, f, g in this order. Each label represents a point, but we are not told which point it is. Instead, we are told the dependency relation between them. As they arrive, we are asked to map each label to a $\mathbb{R}^2$ such that the dependency relations are satisfied.

Take the following instance of this problem: the first 6 points a, b, c, d, e, f arrive, we are told that all triples are independent, i.e., neither of them is collinear. Figure 4 shows two possible ways to place those points in the plane satisfying those dependencies. In the first arrangement, the segments $[a, b], [c, d], [e, f]$ all intersect in a single point. In the second arrangement, they don't.

Now, the last label g arrives. If the algorithm chose the arrangement on the left, we can ask to place g such that $\{a, g, b\}, \{c, g, d\}$ are dependent, but $\{e, f, g\}$ are independent. There is no way to satisfy those dependencies in the figure on the left, since g must be in the intersection of the $[a, b]$ and $[c, d]$ but the only way to do so is by also being in the line $[e, f]$. On the other hand, if the algorithm chose the arrangement on the right in Figure 4, we can ask to place g such that $\{a, g, b\}, \{c, g, d\}$ and $\{e, f, g\}$ are dependent. There is no way to satisfy those dependencies in the figure on the right since the lines $[a, b]$, $[c, d]$ and $[e, f]$ must meet on the same point.

Those two dependencies are satisfiable offline, as we can see in Figure 5. The difficulty is that the dependencies on points that have not arrived pose constraints on the relative position of points that have already arrived.

Figure 4: Two sets of points that induce the same matroid on $\{a, b, c, d, e, f\}$

Our goal with the previous discussion is to provide an intuition for the following proof. Note that while embedding a matroid of rank 3, we are not restricted to $\mathbb{R}^2$. We could in principle embed it in a matroid that is not representable over any field. However, geometric intuition can now be turned into a combinatorial proof of the following statement:

Theorem 6.10. *There is no host matroid **BigM** for which there is an online matroid morphism from $\mathcal{M}_{n,3}$ into **BigM**.*

Figure 5: Once the dependency of point g with respect the remaining points is specified, this constrains the possible geometric arrangements of points $\{a, b, c, d, e, f\}$.

Proof. Define matroids $\mathbf{M}_1$ and $\mathbf{M}_2$ on elements $\{a, b, c, d, e, f, g\}$ represented by the points in Figure 5. In those pictures, a triple of elements is independent iff the corresponding points are not collinear.

Those points induce the same matroid on the first 6 points, which is simply the uniform matroid $\mathbf{U}_{6,3}$, but the matroids differ when we one considers element g.

- in $\mathbf{M}_1$, the sets $\{a, b, g\}$, $\{c, d, g\}$ and $\{e, f, g\}$ are dependent
- in $\mathbf{M}_2$, the sets $\{a, b, g\}$ and $\{c, d, g\}$ are dependent but $\{e, f, g\}$ is independent

Now suppose **BigM** is a matroid and we're trying to construct an online embedding of a matroid $\mathbf{M}$ into **BigM**, where $\mathbf{M}$ is a rank-3 matroid that could either be $\mathbf{M}_1$ or $\mathbf{M}_2$. When the first six elements of $\mathbf{M}$ arrive, we have no way of distinguishing whether the input sequence is going to be $\mathbf{M}_1$ or $\mathbf{M}_2$. The online embedding algorithm chooses some function h mapping $\{a, b, c, d, e, f\}$ to a six-element subset of **BigM**. Denote the images of a, b, c, d, e, f in **BigM** by capital letters, for example $h(a) = A$. Now suppose there are two different extensions of h to the domain a, b, c, d, e, f, g, denoted by h_1 and h_2, such that h_i is an embedding of $\mathbf{M}_i$ into **BigM** for each i. Let $G_1 = h_1(g)$ and $G_2 = h_2(G)$. Then the following must hold:

- the sets $\{A, B, G_1\}$, $\{C, D, G_1\}$ and $\{E, F, G_1\}$ are dependent in **BigM**
- the sets $\{A, B, G_2\}$ and $\{C, D, G_2\}$ are dependent but $\{E, F, G_2\}$ is independent in **BigM**

The sets $\{A, B, G1, G2\}$ and $\{C, D, G1, G2\}$ both have rank 2 in **BigM**, whereas their union has rank 3 since it contains the rank-3 set $\{A, B, C, D\}$. By submodularity, the set $\{G1, G2\}$ must have rank 1. Again by submodularity,

$$\mathsf{rank}_{\mathbf{BigM}}(\{E, F, G1, G2\}) \leq \mathsf{rank}_{\mathbf{BigM}}(\{E, F, G1\}) + \mathsf{rank}_{\mathbf{BigM}}(\{G1, G2\}) - \mathsf{rank}_{\mathbf{BigM}}(\{G1\})$$

which evalutes to $2 + 1 - 1 = 2$. This contradicts the fact that $\{E, F, G2\}$ has rank 3. □

Corollary 6.11. *For every field K of characteristic $p \geq 7$, there is no host matroid **BigM** for which there is an online matroid embedding from all K-representable matroids into **BigM**.*

Proof. Observe that the matroids $\mathbf{M}_1$ and $\mathbf{M}_2$ in the proof of Theorem 6.10 are representable in every field of characteristic at least 7. For example, they can be represented respectively by the columns of the following matrices:

$$\begin{bmatrix} 0 & 2 & 1 & 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 & 0 & 2 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 2 & 1 & 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 & 1 \end{bmatrix}$$

□

We know that because of the results in Section 4, Corollary 6.11 is false for $\mathbb{F}_2$. We leave it as an open question whether there is an online matroid embedding for class of $\mathbb{F}_3$ and $\mathbb{F}_5$-representable matroids.

7 Approximate Embeddings

Like metric embeddings, it is useful to extend the notion of matroid embedding to allow *distortion*. We define an (offline) β-approximate embedding $f : \mathbf{M} \to \mathbf{N}$ as a map between ground sets that approximately preserves rank:

$$\frac{1}{\beta} \cdot \mathsf{rank}_{\mathbf{M}}(S) \leq \mathsf{rank}_{\mathbf{N}}(f(S)) \leq \mathsf{rank}_{\mathbf{M}}(S), \forall S \subseteq \mathbf{M}$$

We also define a randomized β-approximate matroid embedding as a family of functions $f_r : \mathbf{M} \to \mathbf{N}_r$ indexed by a random variable $r \sim R$ such:

$$\mathsf{rank}_{\mathbf{N}_r}(f_r(S)) \leq \mathsf{rank}_{\mathbf{M}}(S) \text{ a.s. } \forall S \subseteq \mathbf{M} \tag{1}$$

$$\mathbb{E}[\mathsf{rank}_{\mathbf{N}_r}(f_r(S))] \geq \frac{1}{\beta}\mathsf{rank}_{\mathbf{M}}(S), \forall S \subseteq \mathbf{M} \tag{2}$$

Their online counterparts can be defined in the natural way: given a class of matroids $\mathcal{C}$ and host matroid **BigM** then an online β-approximate randomized embedding is a family of functions: $f_{\mathbf{M},\pi,r} : \mathbf{M} \to \mathbf{BigM}$ defined for each $\mathbf{M} \in \mathcal{C}$, $\pi : \mathbf{M} \to [n]$ and $r \in R$ such that: (i) it is a randomized β-approximate matroid embedding for each fixed $\mathbf{M}, \pi$; (ii) satisfied the prefix-restriction property defined in Section 3.

For example, if we can design a randomized order-independent β-approximate embedding for class $\mathcal{C}$ into a matroid **BigM** and there is a known α-competitive algorithm for the known-matroid MSP on **BigM**, then we can use the reduction in Section 5 to convert it into an $(\alpha\beta)$-competitive algorithm for the online-revealed-matroid MSP on class $\mathcal{C}$.

It is worth noting that every loop-free matroid $\mathbf{M}$ of rank n admits a trivial n-approximate embedding into the free matroid $\mathbf{Fr}_1 := U_{1,1}$. In the next paragraph, we observe below that we can't obtain better than n for any graphic host matroid. This is related to the notion called *partition property*.

Offline Embedding and the Partition Property The notion of the α-partition property was defined by Abdolazimi et al [AKKG23] to generalize a property exploited by Korula-Pál [KP09] in their algorithm for the graphic matroid secretary problem. Translating it to our notation, we say that a matroid $\mathbf{M}$ satisfies the α-partition property if there is an (offline) α-approximate randomized embedding into a free matroid (i.e., a matroid where every non-empty subset is independent). The name partition property comes from the fact that partition matroids are the class of matroids that admit an (exact) morphism to a free matroid. Their respective lower and upper bounds translate to the following results:

Theorem 7.1 (Korula-Pal [KP09])**.** *If $\boldsymbol{M}$ is a graphic matroid, then it admits a 2-approximate randomized embedding into a free matroid.*

Theorem 7.2 (Abdolazimi et al. [AKKG23], Dughmi et al. [DKP24])**.** *If $\boldsymbol{M}$ is the complete binary matroid of rank n and if it admits an β-approximate randomized embedding into a free matroid, then $\beta \geq \Omega(n/\log n)$. Moreover, there exist a linear matroid $\boldsymbol{M}$ of rank n such that β-approximate randomized embedding into a free matroid only exist for $\beta \geq \Omega(n)$.*

Since composing an β-approximate embedding with a α-approximate embedding we obtain an $(\alpha\beta)$-approximate embedding, we can strenghten the lower bound to also allow for embedding into graphic matroids.

Corollary 7.3. *If $\boldsymbol{M}$ is the complete binary matroid of rank n and if it admits an β-approximate randomized embedding into a graphic matroid, then $\beta \geq \Omega(n/\log n)$. Moreover, if there is a linear matroid $\boldsymbol{M}$ such that it admits a β-approximate randomized embedding into a graphic matroid, then $\beta = \Omega(n)$.*

As we discussed, constructing an n-approximate embedding is trivial. For binary matroids, we complement the lower bound of Dughmi et al. [DKP24] with an algorithm to construct a $O(n/\log n)$ embedding into the free matroid. Moreover, this embedding can be computed online.

Matching Upper Bound that is Also Online Computable We now construct a randomized approximate embedding from a binary matroid to $\mathbf{Fr}_n$. Using Theorem 4.2 we can identify every received element $x \in \mathbf{M}$ with a vector in $\mathbb{F}_2^n \setminus \{0\}$. To map x to a partition matroid, we sample a random basis $b_1, \ldots, b_n$ of $\mathbb{F}_2^n$ and define the function $f_b : \mathbb{F}_2^n \setminus \{0\} \to [n]$ that maps each $x \in \mathbb{F}_2^n$ to the smallest index i such that b_i is in the unique circuit in $\{x, b_1, \ldots, b_n\}$.

We first show that this embedding doesn't increase the rank (first property of an approximate embedding in equation (1)):

Lemma 7.4. *Given any loop-free matroid $\boldsymbol{M}$ of rank n and a basis $b_1, \ldots, b_n$ of $\boldsymbol{M}$, let $f(x)$ be the smallest index i such that b_i is in the unique circuit formed by $\{x, b_1, \ldots, b_n\}$. The function $f_b : \boldsymbol{M} \to \boldsymbol{Fr}_n$ is such that $\mathsf{rank}_{\boldsymbol{Fr}}(f_b(S)) \leq \mathsf{rank}_{\boldsymbol{M}}(S), \forall S \subseteq \boldsymbol{M}$.*

Proof. Given S, choose $S' \subseteq S$ be such that $\mathsf{rank}_{\mathbf{Fr}}(f_b(S)) = \mathsf{rank}_{\mathbf{Fr}}(f_b(S')) = |S'|$. We will show that S' is independent in $\mathbf{M}$ and hence $|S'| = \mathsf{rank}_{\mathbf{M}}(S') \leq \mathsf{rank}_{\mathbf{M}}(S)$.

To show that, observe that for any element $x \in \mathbf{M}$, $f_b(x) = j$ iff $x \in \mathsf{span}_{\mathbf{M}}(\{b_j, b_{j+1}, \ldots, b_n\}) \setminus \mathsf{span}_{\mathbf{M}}(\{b_{j+1}, \ldots, b_n\})$. If that is the case, then: $\mathsf{span}_{\mathbf{M}}(\{x, b_{j+1}, \ldots, b_n\}) = \mathsf{span}_{\mathbf{M}}(\{b_j, b_{j+1}, \ldots, b_n\})$.

Now, construct $x_1, \ldots, x_n$ such that x_j is the element in S' that maps to j if there is such element and $x_j = b_j$ otherwise. Then $f_b(x_j) = j$. We can show by induction that $\mathsf{span}_{\mathbf{M}}(\{x_j, \ldots, x_n\}) = \mathsf{span}_{\mathbf{M}}(\{b_j, \ldots, b_n\})$, by replacing elements of $b_1, \ldots, b_n$ one by one by their corresponding x_j element and observing that the spans are preserved. In particular, $\{x_1, \ldots, x_n\}$ must be independent since they span the entire matroid. As a consequence the original set S' is independent. □

We now show a lower bound on the expected rank of $f_b(S)$ which will lead to the second property of an approximate embedding:

Lemma 7.5. *For the complete binary matroid $\mathbb{F}_2^n$, the embedding f_b for a random basis $b_1, \ldots, b_n$ of $\mathbb{F}_2^n$ satisfies:*

$$\mathbb{E}[\mathsf{rank}_{\boldsymbol{Fr}}(f_b(S))] \geq (1 - 1/e)\log_2(\mathsf{rank}_{\boldsymbol{M}}(S)) - O(1), \forall S \subseteq [n]$$

Proof. Step 1: Sampling the basis We consider the following procedure for sampling an uniformly random basis $b_1, \ldots, b_n$ of $\mathbb{F}_2^n$. For each i we sample an uniform independent vector from $\mathbb{F}_2^n$ and call it b'_i. Now, if b'_i is not in $\mathsf{span}(b_1, \ldots, b_{i-1})$ then we set $b'_i = b_i$. Otherwise we resample until we get an element not in the span of the previous elements and set it as b_i. Nevertheless, we still record b'_i as the first element sampled such that $b'_1, \ldots, b'_n$ are i.i.d. vectors. Now, let's define $\mathcal{E}$ as the event that $b_i = b'_i$ for $i \leq n/2$. We observe that:

$$\mathbb{P}(\mathcal{E}) = (1 - 2^{-n})(1 - 2^{-n+1}) \ldots (1 - 2^{-n/2-1}) \geq 1 - \textstyle\sum_{k=1}^{n/2} 2^{-n+k-1} \geq 1 - 2^{-n/2}$$

Finally, define B to be the $n \times n$ matrix over $\mathbb{F}_2$ that has b_i as the i-th column. Since $b_1, \ldots, b_n$ is a basis, B is invertible.

Step 2: Change of basis Fix a set $S \subseteq \mathbb{F}_2^n$ such that $r = \min(\mathsf{rank}_{\mathbb{F}_2^n}(S), n/2)$ and let $S' = \{s_1, \ldots, s_r\}$ be an independent subset of S. Our goal is to bound $\mathbb{E}[\mathsf{rank}_{\mathbf{Fr}}(f_b(S'))]$.

Let also X be a $n \times n$ invertible matrix over $\mathbb{F}_2$ where the first r columns correspond to $s_1, \ldots, s_r$. Define now $\tilde{b}_i = XB^{-1}e_i$ (where e_i is the standard basis). Since the columns of B are uniformly random and X is invertible, $\tilde{b}_1, \ldots, \tilde{b}_n$ is also a uniformly random basis. With that we observe that:

$$\mathbb{E}[\mathsf{rank}(f_{\tilde{b}_1..\tilde{b}_n}(S))] \geq \mathbb{E}[\mathsf{rank}(f_{\tilde{b}_1..\tilde{b}_n}(S'))] = \mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b_1, \ldots, b_r\}))]$$

since $b_i = BX^{-1}s_i$, $e_i = BX^{-1}\tilde{b}_i$ and $z \mapsto BX^{-1}z$ is a matroid isomorphism of $\mathbb{F}_2^n$.

Step 3: Conditioning on $\mathcal{E}$ The vectors $b_1, \ldots, b_r$ are not sampled i.i.d. but the vectors $b'_1, \ldots, b'_r$ are and conditioned on $\mathcal{E}$ they are the same. So we write:

$$\begin{aligned}
\mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b_1, \ldots, b_r\}))] &\geq \mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b_1, \ldots, b_r\})) \mid \mathcal{E}] \cdot \mathbb{P}[\mathcal{E}] \\
&= \mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b'_1, \ldots, b'_r\})) \mid \mathcal{E}] \cdot \mathbb{P}[\mathcal{E}] \\
&\geq \mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b'_1, \ldots, b'_r\}))] - n/2^{n/2}
\end{aligned}$$

Step 4: Bounding the rank in the free matroid The rank in the free matroid is simply the number of elements in the image, so we can re-write:

$$\begin{aligned}
\mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b'_1, \ldots, b'_r\}))] &= \sum_{j=1}^{n} \mathbb{P}[j \in f_{e_1..e_n}(\{b'_1, \ldots, b'_r\}))] \\
&= \sum_{j=1}^{n} \left(1 - \prod_{i=1}^{r} \mathbb{P}[j \neq f_{e_1..e_n}(b'_i)]\right) \\
&= \sum_{j=1}^{n} \left(1 - (1 - 2^{-j})^r\right)
\end{aligned}$$

since $\mathbb{P}[j \neq f_{e_1..e_n}(b'_i)]$ is the probability that the first non-zero entry of the random vector b'_i is not j, which happens with probability $(1 - 2^{-j})$. Finally observe that for $j \leq \log_2(r)$ we have $(1 - 2^{-j})^r \leq (1 - 1/r)^r \leq 1/e$. Putting this together we obtain that:

$$\mathbb{E}[\mathsf{rank}(f_{e_1..e_n}(\{b'_1, \ldots, b'_r\}))] \geq (1 - 1/e)\log_2(r) \geq (1 - 1/e)\log_2(\mathsf{rank}_{\mathbb{F}_2^n}(S)/2)$$

This completes the proof. □

Theorem 7.6. *For any loop-free binary matroid $\boldsymbol{M}$ there is a $O(n/\log n)$-online approximate embedding to the partition matroid.*

Proof. For a loop-free binary matroid $\mathbf{M}$ use the composition of the embedding f_b for a random basis b of $\mathbb{F}_2^n$ from the previous lemma with the online morphism $\mathbf{M} \to \mathbb{F}_2^n$ in Theorem 4.2. Since both can be computed online, their composition can also be computed online.

By Lemma 7.4 the embedding satisfies the first property of an approximate embedding. By Lemma 7.5 a set of S of rank r in $\mathbf{M}$ is mapped to a set with expected rank $\max(1, c\log(r) - c')$ for constants c, c'. Finally, observe that $r/\max(1, c\log(r) - c') \leq O(n/\log(n))$. □

No Constant-Approximate Embedding from Graphical Matroid to Free Matroid Next, we show that there is no constant approximate online embedding of a graphical matroid into a free matroid when the elements are revealed online with access to an independence oracle. This is in contrast to the result of [KP09] that allows us to construct an approximate offline embedding.

Theorem 7.7. *There is no constant approximate online embedding from the class of graphical matroids into a free matroid.*

Proof. First, we construct the underlying graph that is revealed online to an algorithm that embeds the corresponding graphical matroid into a free matroid. Suppose the underlying graph has two disjoint paths, $e_1, e_2, \ldots, e_d$ and $f_1, \ldots, f_d$. Let the endpoints of the edge $e_i = (v_i, v_{i+1})$ and $f_i = (u_i, u_{i+1})$. We consider two special edges $e_* = (v_1, u_1)$ and $f_* = (v_{d+1}, u_{d+1})$. In addition, we have a set of edges $g_i = (v_i, u_i)$ for all $i \in \{2, \ldots, d\}$. We make d many identical copies of the above graph and k-th copy of vertices and corresponding edges are denoted as $u_i^k, v_i^k, f_i^k, g_i^k$ and f_*^k, e_*^k for all $i \in [d]$ and $k \in [d]$.

To prove the lemma, we need to define an arrival order π over the edges such that any algorithm that maps these edges into a free matroid, and equivalently, a simple partition matroid, can not obtain $O(1)$-approximate embedding.

Suppose the arrival order of edges is as follows: first the set of edges $e_1^k, \ldots, e_d^k; f_1^k, \ldots, f_d^k$ and e_*^k arrives in uniformly random order for $k \in [d]$. Since these edges form an independent set, the algorithm can not distinguish the identity of these edges. Therefore, any constant approximate algorithm needs to map the arrived edges into $O(d^2)$ many disjoint parts of the underlying simple partition matroid. Let the algorithm map the arrived edges into $C \cdot d^2$ many parts. We first observe that at most $C^2 \cdot d^2$ many parts with more than $\frac{1}{C^2}$ edges are assigned to them by the algorithm. We let P be the set of parts that contains at most $\frac{1}{C^2}$ many edges. In addition, since the algorithm can not distinguish between the arrived edges, we can assume that the algorithm adds these edges into $C \cdot d^2$ many parts uniformly at random.

Next, for any fix k, we lower bound the probability that the algorithm embeds all $e_1^k, \ldots, e_d^k$; $f_1^k, \ldots, f_d^k$ and e_*^k into separate parts — denoted as the event $\mathcal{E}$. We let $i_1, \ldots, i_d$, $j_1, \ldots, j_d$ and ℓ^* be the parts in which algorithm embeds the edges $e_1^k, \ldots, e_d^k$, $f_1^k, \ldots, f_d^k$ and e_*^k respectively. We can lower the probability of the event $\mathcal{E}$ by $i_q, j_q \in P$ for all $q \in [d]$ intersecting with the event $\mathcal{E}$. Since each part in P contains at most $\frac{1}{C^2}$ many edges, we can lower bound the probability of the event $\mathcal{E}$ by,

$$\prod_{i=1}^{d+1} \left(1 - \frac{1}{(C - C^2)d^2 - \frac{i}{C^2}}\right) \geq \left(1 - \frac{d}{(C - C^2)d^2 - \frac{d}{C^2}}\right) \geq \left(1 - O\left(\frac{1}{d}\right)\right).$$

For the rest of the proof, we condition on the event that the algorithm embeds all $e_1^k, \ldots, e_d^k$; $f_1^k, \ldots, f_d^k$ and e_*^k into separate parts. Therefore, from now on, we also drop the superscript k as it is fixed in the rest of the proof. In addition, we let P_i, Q_i and P_* be the parts that contains the edges e_i, f_i and e_* for $i \in [d]$.

Next, let the edge f_* arrive followed by the edge e_*. Since, the edges $e_1, \ldots, e_d; f_1, \ldots, f_d$ and e_*, f_* form a cycle, the algorithm is forced to add the edge e_* into one of the existing parts. Since all the edges are symmetric to the algorithm, at this stage, we can assume that the algorithm adds the edge f_* into one of the parts uniformly at random that contains one of the edges $e_1, \ldots, e_d$; $f_1, \ldots, f_d$ and e_*. We define an event $\mathcal{F}$ as the event where the edge f_* does not belong to the parts $P_d, P_{d-1}, \ldots, P_{d-\sqrt{d}}$ and $Q_d, Q_{d-1}, \ldots, Q_{d-\sqrt{d}}$. We note that $\Pr[\mathcal{F}] \geq 1 - \frac{2}{\sqrt{d}}$. For the rest of the sketch, we condition on the event $\mathcal{F}$.

Next, we define the arrival order of the edges as: $g_d, g_{d-1}, g_{d-2}, \ldots, g_{d-\sqrt{d}}$. Upon arrival of the edge g_i for any $i \in \{d, \ldots, d-\sqrt{d}\}$, we consider the following cycles

$$\mathcal{C} = (e_*, e_1, \ldots, e_{i-1}, g_i, u_{i-1}, u_{i-2}, \ldots, u_1)$$
$$\mathcal{C}' = (g_i, v_i, v_{i+1}, \ldots, v_d, f_*, u_d, u_{d-1}, \ldots, u_i).$$

Under the event $\mathcal{F}$, we claim that the edge g_i has to be added to the part P^* that contains the edge f_*.

Suppose the algorithm does not add the edge g_i to the part P^*. In this case, we have either $g_i \in \{P_1, \ldots, P_{i-1}, Q_1, \ldots, Q_{i-1}, \bar{P}\} \setminus P^*$ or $g_i \in \{P_{i+1}, \ldots, P_d, Q_{i+1}, \ldots, Q_d\}$. In the first case, we can observe that all the edges in C' belong to different parts which breaks the assumption that the algorithm comes up with a valid embedding. In the second case, all the edges in C are in different parts which again breaks the same assumption. Therefore, the edge g_i has to be in the part P^*. We can further observe that if g_i is mapped to the part P^* then it leads to a valid (approximate) embedding.

The above argument shows that all the edges $g_i : i \in \{d-\sqrt{d}, \ldots, d\}$ have to be in the same part while conditioned on $\mathcal{E}$ and $\mathcal{F}$ while the rank of $g_i : i \in \{d-\sqrt{d}, \ldots, d\}$ is $\sqrt{d}$. Since $\Pr[\mathcal{E} \cap \mathcal{F}] \geq 1 - O(1/\sqrt{d})$, the algorithm is at most $\sqrt{d}$-approximate. This rules out any constant approximate algorithm. □

References

[AAK+07] Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie. Testing k-wise and almost k-wise independence. In *STOC 2007*, pages 496–505, 2007.

[AGM03] Noga Alon, Oded Goldreich, and Yishay Mansour. Almost k-wise independence versus k-wise independence. *Inf. Process. Lett.*, 88(3):107–110, 2003.

[AKKG23] Dorna Abdolazimi, Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. Matroid partition property and the secretary problem. In *ITCS 2023*, pages 2:1–2:9, 2023.

[AL12] Noga Alon and Shachar Lovett. Almost k-wise vs. k-wise independent permutations, and uniformity for general group actions. In *APPROX'12*, pages 350–361, 2012.

[Ala14] Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM J. Comput.*, 43(2):930–972, 2014.

[Bar98] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *STOC 1998*, pages 161–168, 1998.

[BFU20] Yair Bartal, Nova Fandina, and Seeun William Umboh. Online probabilistic metric embedding: A general framework for bypassing inherent bounds. In *SODA 2020*, pages 1538–1557, 2020.

[BIK07] Moshe Babaioff, Nicole Immorlica, and Robert Kleinberg. Matroids, secretary problems, and online mechanisms. In *SODA 2007*, pages 434–443, 2007.

[BIKK07] Moshe Babaioff, Nicole Immorlica, David Kempe, and Robert Kleinberg. A knapsack secretary problem with applications. In *APPROX-RANDOM 2007*, pages 16–28, 2007.

[BIKK18] Moshe Babaioff, Nicole Immorlica, David Kempe, and Robert Kleinberg. Matroid secretary problems. *J. ACM*, 65(6):35:1–35:26, 2018.

[Bou85] Jean Bourgain. On lipschitz embedding of finite metric spaces in hilbert space. *Isr. J. Math.*, 52:46–52, 1985.

[CGLW22] Ioannis Caragiannis, Nick Gravin, Pinyan Lu, and Zihe Wang. Relaxing the independence assumption in sequential posted pricing, prophet inequality, and random bipartite matching. In *WINE 2022*, pages 131–148, 2022.

[CHMS10] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In Leonard J. Schulman, editor, *STOC 2010*, pages 311–320, 2010.

[CL12] S. Chakraborty and O. Lachish. Improved competitive ratio for the matroid secretary problem. In *SODA 2012*, page 1702–1712, 2012.

[DFKL20] Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM J. Comput.*, 49(3):540–582, 2020.

[DK14] Michael Dinitz and Guy Kortsarz. Matroid secretary for regular and decomposable matroids. *SIAM J. Comput.*, 43(5):1807–1830, 2014.

[DKP24] Shaddin Dughmi, Yusuf Hakan Kalayci, and Neel Patel. Limitations of stochastic selection problems with pairwise independent priors. In *STOC 2024*, pages 479–490, 2024.

[Dug20] Shaddin Dughmi. The outer limits of contention resolution on matroids and connections to the secretary problem. In *ICALP 2020*, volume 168 of *LIPIcs*, pages 42:1–42:18, 2020.

[Dug21] Shaddin Dughmi. Matroid secretary is equivalent to contention resolution. In *ITCS 2022*, volume 215 of *LIPIcs*, pages 58:1–58:23, 2021.

[EHKS18] Soheil Ehsani, MohammadTaghi Hajiaghayi, Thomas Kesselheim, and Sahil Singla. Prophet secretary for combinatorial auctions and matroids. In *SODA 2018*, pages 700–714, 2018.

[FO17] Tara Fife and James Oxley. Laminar matroids. *Eur. J. Comb.*, 62:206–216, 2017.

[FSZ18] Moran Feldman, Ola Svensson, and Rico Zenklusen. A simple $O(\log \log(\text{rank}))$-competitive algorithm for the matroid secretary problem. *Math. Oper. Res.*, 43(2):638–650, 2018.

[FSZ21] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes with applications to bayesian selection problems. *SIAM J. Comput.*, 50(2):255–300, 2021.

[HKS07] Mohammad Taghi Hajiaghayi, Robert D. Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI 2007*, pages 58–65, 2007.

[IMSZ10] Piotr Indyk, Avner Magen, Anastasios Sidiropoulos, and Anastasios Zouzias. Online embeddings. In *APPROX-RANDOM 2010*, pages 246–259, 2010.

[Ind01] Piotr Indyk. Algorithmic applications of low-distortion geometric embeddings. In *FOCS 2001*, pages 10–33, 2001.

[IW11] Sungjin Im and Yajun Wang. Secretary problems: Laminar matroid and interval scheduling. In *SODA 2011*, pages 1265–1274, 2011.

[JSZ13] Patrick Jaillet, José A Soto, and Rico Zenklusen. Advances on matroid secretary problems: Free order model and laminar case. In *IPCO 2013*, pages 254–265, 2013.

[Kan85] William M Kantor. Homogeneous designs and geometric lattices. *J. Comb. Theory Ser. A*, 38(1):66–74, 1985.

[KP09] Nitish Korula and Martin Pál. Algorithms for secretary problems on graphs and hypergraphs. In *ICALP 2009*, pages 508–520, 2009.

[KW12] Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *STOC 2012*, pages 123–136, 2012.

[Lac14] Oded Lachish. O(log log rank) competitive ratio for the matroid secretary problem. In *FOCS 2014*, pages 326–335, 2014.

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15:215–245, 1995.

[LMP22] Marilena Leichter, Benjamin Moseley, and Kirk Pruhs. On the impossibility of decomposing binary matroids. *Oper. Res. Lett.*, 50(5):623–625, 2022.

[LW$^+$06] Michael Luby, Avi Wigderson, et al. Pairwise independence and derandomization. *Foundations and Trends® in Theoretical Computer Science*, 1(4):237–301, 2006.

[Mat02] Jiri Matousek. *Lectures on Discrete Geometry*. Springer, New York, NY, 2002.

[NR25] Ilan Newman and Yuri Rabinovich. Online embedding of metrics. *Isr. J. Math.*, 2025. Forthcoming.

[OGV13] Shayan Oveis Gharan and Jan Vondrák. On variants of the matroid secretary problem. *Algorithmica*, 67:472–497, 2013.

[SSZ23] Richard Santiago, Ivan Sergeev, and Rico Zenklusen. Simple random order contention resolution for graphic matroids with almost no prior information. In *SOSA 2023*, pages 84–95, 2023.

[SSZ25] Richard Santiago, Ivan Sergeev, and Rico Zenklusen. Constant-competitiveness for random assignment matroid secretary without knowing the matroid. *Math. Program.*, 210(1):815–846, 2025.

[Tut58] William Thomas Tutte. A homotopy theorem for matroids. I, II. *Trans. Am. Math. Soc.*, 88(1):144–174, 1958.

[Vad12] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, 2012.

A Reduction to a Special Case of Prophet MSP

In this appendix, we provide a proof of Lemma 5.8. First, we present a simple reduction from [Dug21].

Lemma A.1 (Sublemma-4.2 from [Dug21]). *If there exists an α-approximate Prophet MSP on matroid $\boldsymbol{M}$ with $\boldsymbol{Rank}(\boldsymbol{M}) = d$ and weight distribution $\mathcal{D}$ supported over $\left\{\frac{1}{256\cdot d}, \frac{2}{256\cdot d}, \frac{2^2}{256\cdot d}, \ldots, 1\right\}$ with the offline optimum $\mathbb{E}_{\mathcal{D}}[\boldsymbol{OPT}(\boldsymbol{M})] \in \left[\frac{1}{16}, 1\right]$ then there exists $\frac{\alpha}{256}$-approximate prophet MSP on matroid $\boldsymbol{M}$ with any arbitrary weight distribution.*

Given the simple reduction in Lemma A.1, we prove the following simple reduction that allow us to focus on the weight distribution that assigns distinct weight to each element of the matroid. We consider weight class,

$$W = \left\{\frac{1}{256\cdot d} + \frac{i-1}{256d\cdot n^2} : i \in \{0, 1, \ldots, 256d\cdot n^2 + 1\}\right\} \cup \left\{\frac{1}{256\cdot d}, \frac{2}{256\cdot d}, \frac{2^2}{256\cdot d}, \ldots, 1\right\}.$$

Above, n denotes the set of elements of the matroid $\mathbf{M}$.

Given any prophet MSP instance consists of matroid $\mathbf{M}$ and arbiritary weight distribution, we reduce it to a prophet MSP instance $\mathbb{F}$ consists of matroid $\mathbf{M}$ and weight distribution $\mathcal{D}$ supported on $\left\{\frac{1}{256\cdot d}, \frac{2}{256\cdot d}, \frac{2^2}{256\cdot d}, \ldots, 1\right\}$ with the offline optimum $\mathbb{E}_{\mathcal{D}}[\mathbf{OPT}(\mathbf{M})] \in \left[\frac{1}{16}, 1\right]$. Then we construct prophet MSP instnace $\mathbb{F}'$ over the same matroid but slightly perturbed weight distribution as follows: we find a random permutation π over the set of elements of the matroid. We sample the weights of the elements using the distribution $\mathcal{D}$. Then for all $i \in [n]$, we subtract $\frac{i-2}{256\cdot d\cdot n^2}$ from the weight of element e, iff the element e appears at the i-th position on the permutation π.

We observe that in the perturbed matroid secretary instance, the weight distribution is supported over the set of weights W. In addition, for any $w \in W$, there is at most one element whose weight is assigned to be w. This simply follows from the fact that for any two elements e, e' with distinct weight in the draw from distribution $\mathcal{D}$ will be assigned a different weight as the prtrubution to both elements is smaller than $\frac{1}{256d\cdot n}$. In addition, the perturbation is distinct for two distinct elements $e, e' \in V$. Therefore, the pair of elements e, e' with the identical weight in the draw from $\mathcal{D}$ are assigned different weights in the perturbed instance.

Proof of Lemma 5.8. To prove the lemma, we prove the following statement: if there is an α approximate algorithm for perturb prophet matroid secretary instance $\mathbb{F}'$ then we can construct $\alpha - O\left(\frac{1}{d}\right)$-approximate algorithm for prophet matroid secretary instance $\mathbb{F}$.

Suppose, we are given an algForithm $\mathcal{A}$ that is α-approximate for the instance $\mathbb{F}'$. Our goal is to design an algorithm for the instance $\mathbb{F}$ using $\mathcal{A}$ that is $\alpha - \frac{1}{d}$-approximate. We give the following online reduction from $\mathbb{F}$ to $\mathbb{F}'$: in the instance $\mathbb{F}$, upon arrival of the element e at the position $i \in [n]$, we perturb the weight of the element e by subtracting $\frac{i-1}{256\cdot d\cdot n^2}$. We note that since the arrival order of the elements is uniformly at random, the modified weight distribution is identical to that of weight distribution in the perturbed instance. Hence, we feed the perturbed weight of the arrived element to the algorithm $\mathcal{A}$. We then select an arrived element e iff the algorithm $\mathcal{A}$ selects the element e in the perturbed instance.

Let S, $\mathbf{OPT}$ and $\mathbf{OPT}'$ be the selected elements by $\mathcal{A}$, the optimal value of the instance $\mathbb{F}$, and the optimal value of the perturbed instance $\mathbb{F}'$. We denote the modified weight as $w'(\cdot)$. By construction, we have,

$$w(S) \geq w'(S) \geq \alpha\cdot\mathbf{OPT}' \geq \alpha\cdot\left(1 - \frac{1}{32\cdot d}\right)\cdot\mathbf{OPT}.$$

Above, the first inequality holds because $w'(e) \leq w(e)$. The second inequality holds because $\mathcal{A}$ is α-approximate for thee perturbed distribution. The final inequality holds because the total decrease in the weight $\sum_{e \in V}(w(e) - w'(e)) \leq \frac{1}{512 \cdot d}$ and $\mathbf{OPT} \geq \frac{1}{16}$ by assumption on $\mathcal{D}$. This concludes the proof. □

B Exact from Approximate Pairwise Independence

In this section, we present the proof of Theorem 5.9, and construct an exact pairwise-independent weight distribution on $\mathbf{BigM}_{[m \cdot N]}$ that is close to the weight distribution $\mathcal{D}^*$ defined in Definition 1 on $\mathbf{BigM}_{[m \cdot N]}$. First, in Section B.1, we present our construction that iteratively applies small perturbations as defined in Procedure 1 and Procedure 2 to turn the weight distribution $\mathcal{D}^*$ into an exact pairwise-independent weight distribution $\tilde{\mathcal{D}}$. Then in Section B.2, we obtain tight upper and lower bounds on the parameters of Procedure 1 and Procedure 2. Afterwards, in Sections B.3 and B.4, we bound the total perturbations due to Procedures 1 and 2, respectively. Finally, Section B.5 shows how this implies the theorem.

B.1 Constructing an Exact-Pairwise Independent Distribution

Before we present the procedure to obtain an exact pairwise independent weight distribution, we recall the notations from Theorem 5.9. We have a prophet MSP instance on matroid $\mathbf{M}$ and weight distribution $\mathcal{D}$ supported over the set of weights $W = \{w_1, w_2, \ldots, w_m\}$ satisfying the conditions from Lemma 5.8. Given the prophet MSP instance we let $\mathcal{D}^*$ be the weight distribution over $\mathbf{BigM}$ defined in Definition 1. For any $i, j \in [m]$, we let,

$$p_i = \Pr_{w \sim \mathcal{D}}[\exists \mathbf{v} \in \mathbf{M} : w(\mathbf{v}) = w_i]$$
$$p_{ij} = \Pr_{w \sim \mathcal{D}}[\exists \mathbf{v} \in \mathbf{M} : w(\mathbf{v}) = w_i \wedge \exists \mathbf{v}' \in \mathbf{M} : w(\mathbf{v}') = w_j].$$

Since $\mathcal{D}$ is arbitrarily correlated over the support W, we can potentially have $p_{ij} \neq p_i \cdot p_j$.

We first observe that for any $i \in [m], \ell \in N_i$ and $\mathbf{u} \in \mathbf{BigM}$, $\mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ can potentially take either weight of w_i or zero. Therefore, we define Bernoulli random variables $X^{\mathbf{u}}_{i\ell}$ for all $\mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ such that $X^{\mathbf{u}}_{i\ell} = 1$ iff $w(\mathbf{u}^{i,\ell}) = w_i$. We note that the Bernoulli random variables $X^{\mathbf{u}}_{i\ell}$ for all $\mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ contain the full information of the weight distribution over $\mathbf{BigM}$.

We next observe that,

$$\begin{aligned}
\mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1] &= \Pr\left[\exists \mathbf{v} \in \mathbf{M} : w(\mathbf{v}) = w_i \wedge f(\mathbf{v}) = \mathbf{u}\right] \cdot \Pr_{\ell' \sim \mathrm{Unif}(N_i)}[\ell' = \ell] \\
&= \Pr\left[\exists \mathbf{v} \in \mathbf{M} : w(\mathbf{v}) = w_i\right] \cdot \Pr\left[f(\mathbf{v}) = \mathbf{u}\right] \cdot \frac{1}{N} \\
&= \frac{p_i}{M \cdot N}.
\end{aligned}$$

Above, the second equality holds because the weight assignment to $\mathbf{M}$ is independent of the matroid morphism f. The second equality holds because $f = f'' \circ f'$ and f'' is a uniformly random automorphism from $\mathrm{Aut}(\mathbf{BigM})$, which implies $\Pr[f(\mathbf{v}) = \mathbf{u}] = \frac{1}{|\mathbf{BigM}|} = \frac{1}{M}$. Similarly, we compute the pairwise joint distribution of the Bernoulli random variables. First, for any elements $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, $i, j \in [m]$ and $\ell \in [N_i], \ell' \in [N_j]$, we have,

$$\mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1 \wedge X^{\mathbf{u}'}_{j,\ell'} = 1] = 0$$

if $i = j$ or $\mathbf{u} = \mathbf{u}'$ because there can be at most one element $\mathbf{v} \in \mathbf{M}$ with $w(\mathbf{v}) = w_i$ and f is a matroid morphism and two distinct elements from $\mathbf{M}$ with weight w_i and w_j can not be mapped to $\mathbf{u} \in \mathbf{BigM}$ via f.

Finally, for any pair of elements $\mathbf{u}_{i,\ell}, \mathbf{u'}_{j,\ell}$, for $i \neq j$ and $\mathbf{u} \neq \mathbf{u}'$, we have,

$$\begin{aligned}\mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1 \wedge X^{\mathbf{u}'}_{j,\ell'} = 1] &= \Pr\left[\exists \mathbf{v} \in \mathbf{M} : w(\mathbf{v}) = w_i \wedge \exists \mathbf{v}' \in \mathbf{M} : w(\mathbf{v}') = w_j \wedge f(\mathbf{v}) = \mathbf{u} \wedge f(\mathbf{v}') = u'\right] \cdot \frac{1}{N^2} \\ &= p_{ij} \cdot \Pr[f(\mathbf{v}) = \mathbf{u} \wedge f(\mathbf{v}') = u'] \cdot \frac{1}{N^2} \\ &= \frac{p_{ij}}{M \cdot (M-1) \cdot N^2}.\end{aligned}$$

Above, the second equality holds because the weight assignment to $\mathbf{M}$ is independent of the matroid morphism f. The second equality holds because $f = f'' \circ f'$ and f'' is a uniformly random automorphism from $\mathrm{Aut}(\mathbf{BigM})$, which implies $\Pr[f(\mathbf{v}) = \mathbf{u} \wedge f(\mathbf{v}') = u'] = \frac{1}{M \cdot (M-1)}$. Given the definition of $X^{\mathbf{u}}_{i,\ell}$, since it captures the complete information about the weight distribution over $\mathbf{BigM}_{[m \cdot M]}$, we prove the existence of Bernoulli random variables $\tilde{X}^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell}$ with pairwise-independent joint distribution $\tilde{\mathbf{X}}$ s.t. $\mathrm{TV}_{\mathbf{X},\tilde{\mathbf{X}}} \leq O\left(\frac{m^3}{M}\right)$. This immediately implies the proof of Theorem 5.9.

It is difficult to give a closed form construction of $\tilde{\mathbf{X}}$ in a "simple" procedure, therefore, we give a sequential process that iteratively generates Bernoulli random variables $X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ for $k = 0, \ldots, k^*$ starting from $X^{\mathbf{u}}_{i,\ell}(0) = X^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$ such that at the end of the process, $\tilde{X}^{\mathbf{u}}_{i,\ell}(\bar{k}) : \forall \mathbf{u}^{i,\ell}$ are pairwise independent. For any $i, j \in [m]$, and distinct $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ $i \neq j$, we define bias

$$\varepsilon_{ij} = |\mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1 \wedge X^{\mathbf{u}'}_{j,\ell'} = 1] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1] \cdot \mathbb{E}[X^{\mathbf{u}'}_{j,\ell'} = 1]| = \frac{|\bar{p}_{ij} - p_i \cdot p_j|}{M^2 \cdot N^2},$$

where $\bar{p}_{ij} = \frac{M}{M-1} \cdot p_{ij}$. The above expression captures the closeness of the random variables $X^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[n \cdot M]}$ from being pairwise independent. The above bias ε_{ij} does not depend on the choice of vectors $\mathbf{u}, \mathbf{u}'$ and their corresponding labels ℓ, ℓ'. On the other hand, for $i \in [m]$ and distinct $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, we let,

$$\varepsilon_i = |\mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1 \wedge X^{\mathbf{u}'}_{i,\ell'} = 1] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell} = 1] \cdot \mathbb{E}[X^{\mathbf{u}'}_{i,\ell'} = 1]| = \frac{p_i^2}{M^2 \cdot N^2}.$$

Our high-level idea is to sequentially transform the random variables starting from $\{X^{\mathbf{u}}_{i,\ell}(0) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}\}$ distributed as $\mathcal{D}(0) = \mathcal{D}^*$ to $\{X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell}\}$ for bounded k such that at each step, the resultant distribution becomes "closer" to being pairwise independent. More formally, we prove the following lemma.

Lemma B.1. *Let $\{\tilde{X}^{\mathbf{u}}_{i,\ell} : \mathbf{u}_{i,\ell} \in \boldsymbol{BigM}_{[m \cdot N]}\}$ be the random variables obtained after applying Procedure 1 and Procedure 2 to $\{X^{\mathbf{u}}_{i,\ell} : \mathbf{u}_{i,\ell} \in \boldsymbol{BigM}_{[m \cdot N]}\}$, then $\{\tilde{X}^{\mathbf{u}}_{i,\ell} : \mathbf{u}_{i,\ell} \in \boldsymbol{BigM}_{[m \cdot N]}\}$ are pairwise-independent.*

Proof of Lemma B.1. We define initial bias $\varepsilon_{ij}(0) = \varepsilon_{ij}$ and $\varepsilon_i(0) = \varepsilon_i$. We let the evolution of the weight distribution starting from $\mathcal{D}^* = \mathcal{D}(0) \to \mathcal{D}(1) \to \cdots \to \mathcal{D}(\bar{k})$, where $\bar{k} = O(m^2)$ in Procedure 1. For any $k > 0$, we let,

$$\begin{aligned}\varepsilon_i(k) &= |\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = 1 \wedge X^{\mathbf{u}'}_{i,\ell'}(k) = 1] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = 1] \cdot \mathbb{E}[X^{\mathbf{u}'}_{i,\ell'}(k) = 1]| \\ \varepsilon_{ij}(k) &= |\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = 1 \wedge X^{\mathbf{u}'}_{j,\ell'}(k) = 1] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = 1] \cdot \mathbb{E}[X^{\mathbf{u}'}_{j,\ell'}(k) = 1]|.\end{aligned}$$

Through the process of transforming Bernoulli random variables in Procedure 1, we make sure that $\varepsilon_{ij}(k)$ and $\varepsilon_i(k)$ do not depend on $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and their labels $\ell \in [N]$ and $\ell' \in [N]$ and rather only depends on the weight class $i, j \in [m]$.

Suppose we have completed k many iterations of Procedure 1 and have obtained distribution $\mathcal{D}(k)$ over $\{X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$. In addition, we have for any $i \in [m], \mathbf{u} \in \mathbf{BigM}$ and $\ell \in [N]$, we have $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}] = p_i(k)$ and any distinct $i, j \in [m], \mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and any labels $\ell, \ell' \in [N]$, we have $\mathbb{E}[X^{\mathbf{u}}_{i,\ell} \cdot X^{\mathbf{u}'}_{j,\ell'}] = p_{ij}(k)$. Given $X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$, and $i \neq j$ with $\varepsilon_{ij}(k) \neq 0$, we consider the two cases: $\varepsilon_{ij}(k) > 0$ and $\varepsilon_{ij}(k) < 0$ separately and resolve their pairwise correlations (See Procedure 1).

Procedure 1:
Initialize $k = 0$ **and** $\{X^{\mathbf{u}}_{i,\ell}(0) = X^{\mathbf{u}}_{i,\ell} : \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$
Run until for all distinct $i, j \in [m]$**,** $p_{ij}(k) = p_i(k) \cdot p_j(k)$**:**
Case-1 $(p_{ij}(k) > p_i(k) \cdot p_j(k))$**:** We define $X^{\mathbf{u}}_{i,\ell}(k+1) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$ as follows:

1. Assign $X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$.
2. With probability q_{ij}, sample $Z \sim \mathrm{Ber}(1/2)$ and do the following:
   (a) If $Z = 1$ then assign $X^{\mathbf{u}}_{i,\ell}(k+1) = 1$ with probability $\frac{1}{N}$ independently $\forall \mathbf{u} \in \mathbf{BigM}, \forall \ell \in [N]$ and $X^{\mathbf{u}}_{j,\ell}(k+1) = 0 : \forall \mathbf{u} \in \mathbf{BigM}, \forall \ell \in [N]$.
   (b) Otherwise assign $X^{\mathbf{u}}_{i,\ell}(k+1) = 0 : \forall \mathbf{u} \in \mathbf{BigM}_{[m\cdot N]}, \ell \in [N]$ and $X^{\mathbf{u}}_{j,\ell}(k+1) = 1$ with probability $\frac{1}{N}$ independently $\forall \mathbf{u} \in \mathbb{F}_2^d, \ell \in [N]$.

Case-2 $(p_{ij}(k) < p_i(k) \cdot p_j(k))$**:** We define $X^{\mathbf{u}}_{i,\ell}(k+1) : \forall \mathbf{u}^{i,\ell}$ as follows:

1. Assign $X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$.
2. With probability q_{ij}, let $X^{\mathbf{u}}_{i,\ell}(k+1) = 1 : \forall \mathbf{u} \in \mathbf{BigM}, \ell \in [N]$ and $X^{\mathbf{u}}_{j,\ell}(k+1) = 1 : \forall \mathbf{u} \in \mathbf{BigM}, \ell \in [N]$.

First, we can straightforwardly observe that while obtaining $\{X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$ we ensure that for any fix $i \in [N]$, $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and $\ell, \ell' \in [N]$, we have $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k+1)] = \mathbb{E}[X^{\mathbf{u}'}_{j,\ell'}(k+1)]$. In addition, given any fix distinct $i, j \in [m]$, for any distinct $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and labels $\ell, \ell' \in [N]$, $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k+1) \cdot X^{\mathbf{u}'}_{j,\ell'}(k+1)]$ does not depend on $\mathbf{u}, \mathbf{u}'$ and their labels $\ell, \ell' \in [N]$.

In both cases, we let q_{ij} such that we obtain $\varepsilon_{ij}(k+1) = 0$. We will define the closed form and an upper bound on q_{ij} by writing the conditions to obtain $\varepsilon_{ij}(k+1) = 0$ in Section B.2. However, we prove the following crucial property that shows that the above procedure stops after iterating over all pairs of indices $i \neq j$ and leads to the distribution where $X^{\mathbf{u}}_{i,\ell}(k)$ and $X^{\mathbf{u}'}_{j,\ell'}(k)$ are independent as long as $i \neq j$.

Claim B.2. *For any pair r, r' such that $r \neq r'$, we have $\varepsilon_{rr'}(k+1) \leq \varepsilon_{rr'}(k)$. In addition, for any $\mathbf{u} \in \boldsymbol{BigM}$, $\ell, \ell' \in [N]$ and distinct $r, r' \in [m]$, we have,*

$$|\mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k+1) \cdot X^{\mathbf{u}}_{r',\ell'}(k+1)] - \mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k+1)] \cdot \mathbb{E}[X^{\mathbf{u}}_{r',\ell'}(k+1)]| \leq \mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k) \cdot X^{\mathbf{u}}_{r',\ell'}(k)] - \mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k)] \cdot \mathbb{E}[X^{\mathbf{u}}_{r',\ell'}(k)]$$

Proof. We can easily observe that when $r, r' \notin \{i, j\}$ then $\varepsilon_{rr'}(k+1) = \varepsilon_{rr'}(k)$. In addition, when $r = i, r' = j$, the choice of q_{ij} ensures that $\varepsilon_{ij}(k+1) = 0 < \varepsilon_{ij}(k+1)$. Hence, we can focus on the

case, $r' = i$. We note that for any pair of $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$,

$$\begin{aligned}
&|\mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k+1) \cdot X^{\mathbf{u}}_{r',\ell'}(k+1)] - \mathbb{E}[X^{\mathbf{u}}_{r,\ell}(k+1)] \cdot \mathbb{E}[X^{\mathbf{u}}_{r',\ell'}(k+1)]| \\
&= |\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}'}_{r,\ell'}(k+1) = 1] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k+1) = 1] \cdot \mathbb{E}[X^{\mathbf{u}'}_{r,\ell'}(k+1) = 1]| \\
&= |(1 - q_{ij}) \cdot \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = X^{\mathbf{u}'}_{r,\ell'}(k) = 1] + \alpha \cdot q_{ij} \cdot \mathbb{E}[X^{\mathbf{u}'}_{r,\ell'}(k) = 1] \\
&- ((1 - q_{ij})\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) = 1] + q_{ij} \cdot \alpha) \cdot \mathbb{E}[X^{\mathbf{u}'}_{r,\ell'}(k) = 1]| \qquad (\alpha = 1 \text{ or } \alpha = 1/2) \\
&= (1 - q_{ij}) \cdot \varepsilon_{ir}(k).
\end{aligned}$$

Next, we observe that for the case when $\mathbf{u} \neq \mathbf{u}'$, $|\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k') \cdot X^{\mathbf{u}}_{r',\ell'}(k')] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k')] \cdot \mathbb{E}[X^{\mathbf{u}}_{r',\ell'}(k')]| = \varepsilon_{ir}(k')$ for $k' \in \{k, k+1\}$. This further implies that $\varepsilon_{i,r}(k+1) \leq \varepsilon_{i,r}(k)$. □

Claim B.2 implies that that once Procedure 1 iterates over all possible distinct pairs $i, j \in [m]$ over $\bar{k}$ many iterations, for the resultant , the resultant $X^{\mathbf{u}}_{i,\ell}$ and $X^{\mathbf{u}'}_{j,\ell'}$ are independent as long as $i \neq j$. Suppose the above procedure terminates after $\bar{k}$ many iterations. We note that $\bar{k} \leq \frac{m \cdot (m-1)}{2}$ due to Claim B.2. Once the above procedure terminates for all $i \neq j$, the only correlation we have left is between $X^{\mathbf{u}}_{i,\ell}$ and $X^{\mathbf{u}'}_{i,\ell'}$ for $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, $i \in [m]$ and $\ell, \ell \in [N]$ and between $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}}_{j,\ell'}$ for $\mathbf{u} \in \mathbf{M}$, $i, j \in [m]$ and $\ell, \ell' \in [N]$.

We now describe Procedure 2 to resolve the rest of the correlations, which is similar to our Procedure 1. In Procedure 2, we initialize $\{\tilde{X}^{\mathbf{u}}_{i,\ell}(0) = X^{\mathbf{u}}_{i,\ell}(\bar{k}) : \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}\}$ and at each iteration, we resolve the correlation between a pair of Bernoulli random variables which exhibits pairwise correlation. This is in contrast to Procedure 1 that resolved the pairwise correlation between the sets of random variables corresponding to pairs of weight classes $i, j \in [m]$.

Procedure 2:
Initialize $k = 0$ **and** $\{\tilde{X}^{\mathbf{u}}_{i,\ell}(0) = \tilde{X}^{\mathbf{u}}_{i,\ell}(\bar{k}) : \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}\}$
Run until $\{\tilde{X}^{\mathbf{u}}_{i,\ell}(k) : \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}\}$ **are Pairwise Independent:**
$\exists \tilde{X}^{\mathbf{u}}_{i,\ell}(k) \tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ **not independent for some** $i, j \in [m]$**,** $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ **and** $\ell, \ell' \in [N]$
Case-1: $(\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k) \cdot \tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] > \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)])$**:**

1. Assign $X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$.
2. With probability $q(k)$, sample $Z \sim \text{Ber}(1/2)$ and do the following:
   (a) If $Z = 1$ then assign $X^{\mathbf{u}}_{i,\ell}(k+1) = 1$ and $X^{\mathbf{u}'}_{j,\ell'}(k+1) = 0$.
   (b) Otherwise assign $X^{\mathbf{u}}_{i,\ell}(k+1) = 0$ and $X^{\mathbf{u}'}_{j,\ell'}(k+1) = 1$.

Case-2: $(\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k) \cdot \tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] < \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)])$**:**

1. Assign $X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}}_{i,\ell}(k) : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m \cdot N]}$.
2. With probability $q(k)$, let $X^{\mathbf{u}}_{i,\ell}(k+1) = X^{\mathbf{u}'}_{j,\ell'} = 1$.

At any iteration k of Procedure 2, let $\tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ are the selected random variables whose correlation is being resolved. We set $q(k)$ such that we end up with $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k) \cdot \tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] = \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]$. Similar to Claim B.2, we show that the correlation between all pair of random variables can only go lower.

Claim B.3. *For any $k \geq 0$ and pair of random variables $X(k), X'(k) \in \{\tilde{X}^{\mathbf{u}}_{i,\ell}(k) : \mathbf{u}^{i,\ell} \in \boldsymbol{BigM}_{[m\cdot N]}\}$, we have,*

$$|\mathbb{E}[X(k+1)\cdot X'(k+1)] - \mathbb{E}[X(k+1)]\cdot\mathbb{E}[X'(k+1)]| \leq |\mathbb{E}[X(k)\cdot X'(k)] - \mathbb{E}[X(k)]\cdot\mathbb{E}[X'(k)]|.$$

Proof. Let $\tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ are the selected random variables whose correlation is being resolved. For any $X(k), X'(k) \notin \{\tilde{X}^{\mathbf{u}}_{i,\ell}(k), \tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)\}$, the condition in the claim trivially satisfies. On the other hand, if $X(k) = \tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $X'(k) = \tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ then we have $\mathbb{E}[X(k+1)\cdot X'(k+1)] = \mathbb{E}[X(k+1)]\cdot\mathbb{E}[X'(k+1)]$ which implies the condition in the claim.

To complete the proof, we focus on the case when $X(k) = \tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $X'(k+1) \notin \{\tilde{X}^{\mathbf{u}}_{i,\ell}(k), \tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)\}$. The proof of this case is identical to the proof of Claim B.2 but we give a detailed proof for the sake of completeness. We observe that,

$$\begin{aligned}
&|\mathbb{E}[X(k+1)\cdot X'(k+1)] - \mathbb{E}[X(k+1)]\cdot\mathbb{E}[X'(k+1)]| \\
=&|(1-q(k))\cdot\mathbb{E}[X(k)\cdot X'(k)] + q(k)\cdot\alpha\cdot\mathbb{E}[X'(k)] - ((1-q(k))\cdot\mathbb{E}[X(k)] + q(k)\cdot\alpha)\cdot\mathbb{E}[X'(k)]| \\
=&(1-q(k))\cdot|\mathbb{E}[X(k)\cdot X'(k)] - \mathbb{E}[X(k)]\cdot\mathbb{E}[X'(k)]| \\
\leq&|\mathbb{E}[X(k)\cdot X'(k)] - \mathbb{E}[X(k)]\cdot\mathbb{E}[X'(k)]|.
\end{aligned}$$

In the above calculations, $\alpha = 1/2$ if $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)] > \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)]$ and $\alpha = 1$ otherwise. □

The above claim implies that at any iteration k, if $\tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ are the selected random variables whose correlation is being resolved, either $i = j$ or $\mathbf{u} = \mathbf{u}'$ since we have $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(0)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(0)] = \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(0)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(0)]$ for distinct $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and distinct $i, j \in [m]$ at the end of Procedure 1. This implies that Procedure 2 stops after iterative over pair of random variables $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}}_{j,\ell'}$ for $\mathbf{u} \in \mathbf{BigM}, i, j \in [m], \ell, \ell' \in [N]$ and $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}'}_{i,\ell'}$ for $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}, i \in [m], \ell, \ell' \in [N]$. This completes the proof fo the lemma that the resultant random variables $\{\tilde{X}^{\mathbf{u}}_{i,\ell} : \mathbf{u}_{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$ at the end of Procedure 1, followed by Procedure 2 are pairwise independent. □

B.2 Parameters of Procedure 1 and Procedure 2

Next, in order to complete the proof of Theorem 5.9, we first bound q_{ij} and $q(k)$, i.e, the probability by which Procedure 1 and Procedure 2 deviate from the original random variables $\{X^{\mathbf{u}}_{i,\ell} : \mathbf{u}_{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$. We next obtain upper and lower bound on the parameters of the Procedures 1 and 2, respectively.

Lemma B.4. *Let (i, j) be the indices selected at iteration k of Procedure 1, then we have,*

1. *When $p_{ij}(k) > p_i(k)\cdot p_j(k)$, we have $\frac{N\cdot\varepsilon_{ij}(k)}{2(p_i(k)+p_j(k))} \leq q_{ij} \leq \frac{N\cdot\varepsilon_{ij}(k)}{p_i(k)+p_j(k)}$, and*
2. *When $p_{ij}(k) < p_i(k)\cdot p_j(k)$, we have $\varepsilon_{ij} \leq q_{ij} \leq \frac{\varepsilon_{ij}(k)}{(1-p_i(k)-p_j(k))}$*

In addition, for $k \geq 0$, let $\tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $\tilde{X}^{\mathbf{u}'}_{j,\ell'}(k)$ are the selected random variables whose correlation is being resolved at the k-th iteration of Procedure 2. Let $p_1(k) = \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}]$ and $p_2(k) = \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell'}]$, then we have,

1. *When $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] > \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]$, we have*

$$\frac{|\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]|}{2(p_1(k)+p_2(k))} \leq q(k) \leq \frac{|\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]|}{p_1(k)+p_2(k)},$$

2. When $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] < \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]$*, we have*

$$|\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]| \le q(k) \le \frac{|\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)]|\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(k)]|}{1-p_1(k)-p_2(k)},$$

Proof. In both cases, we want to find q_{ij} such that $p_{ij}(k+1) = p_i(k+1)\cdot p_j(k+1)$. We analyze both cases of Procedure 1 separately,

Case-1 $(p_{ij}(k) > p_i(k)\cdot p_j(k))$ In this case, by construction, we have,

$$p_i(k+1) = (1-q_{ij})\cdot p_i(k) + \frac{q_{ij}}{2N} \quad \text{and} \quad p_{ij}(k+1) = (1-q_{ij})\cdot p_{ij}(k).$$

For simplicity in notations, we let $a_{ij} = \frac{1}{4N^2} - \frac{1}{2N}\cdot(p_i(k)+p_j(k))$ and $b_{ij} = \frac{1}{2N}\cdot(p_i(k)+p_j(k)) + \varepsilon_{ij}(k) + p_i(k)\cdot p_j(k)$. By constraint, $p_{ij}(k+1) = p_i(k+1)\cdot p_j(k+1)$, we get,

$$\begin{aligned}
&(1-q_{ij})\cdot p_{ij}(k) = \left((1-q_{ij})\cdot p_i(k) + \frac{q_{ij}}{2N}\right)\cdot\left((1-q_{ij})\cdot p_j(k) + \frac{q_{ij}}{2N}\right)\\
\implies &\left(\frac{1}{4N^2} - \frac{1}{2N}\cdot(p_i(k)+p_j(k))\right)\cdot q_{ij}^2 + \left(\frac{1}{2N}\cdot(p_i(k)+p_j(k)) + \varepsilon_{ij}(k) + p_i(k)\cdot p_j(k)\right)\cdot q_{ij} - \varepsilon_{ij}(k) = 0\\
\implies q_{ij} = &\frac{-b_{ij} + \sqrt{b_{ij}^2 + 4\cdot\varepsilon_{ij}(k)\cdot a_{ij}}}{2\cdot a_{ij}} \implies q_{ij} = \frac{\varepsilon_{ij}(k)}{b_{ij} + \sqrt{b_{ij}^2 + 4\cdot\varepsilon_{ij}(k)\cdot a_{ij}}} \le \frac{N\cdot\varepsilon_{ij}(k)}{p_i(k)+p_j(k)}.
\end{aligned}$$

Above the last inequality follows because $b_{ij} \ge \frac{1}{2}\cdot(p_i(k)+p_j(k))$. In addition, we get $q_{ij} \ge \frac{\varepsilon_{ij}}{2\cdot b_{ij}+2\varepsilon_{ij}\cdot a_{ij}} \ge \frac{\varepsilon_{ij}}{2\cdot(p_i(k)+p_j(k))}$. We conclude the proof of this case.

Case-2 $(p_{ij}(k) < p_i(k)\cdot p_j(k))$ In this case, by construction, we have,

$$p_i(k+1) = (1-q_{ij})\cdot p_i(k) + q_{ij} \quad \text{and} \quad p_{ij}(k+1) = (1-q_{ij})\cdot p_{ij}(k) + q_{ij}.$$

For simplicity in notations, we let $a_{ij} = 1 + p_i(k)\cdot p_j(k) - (p_i(k)+p_j(k))$ and $b_{ij} = (1 - p_i(k) - p_j(k) - p_i(k)\cdot p_j(k) + \varepsilon_{ij}(k))$. By constraint, $p_{ij}(k+1) = p_i(k+1)\cdot p_j(k+1)$, we get,

$$\begin{aligned}
&(1-q_{ij})\cdot p_{ij}(k) + q_{ij} = ((1-q_{ij})\cdot p_i(k) + q_{ij})\cdot((1-q_{ij})\cdot p_j(k) + q_{ij})\\
\implies &(1 + p_i(k)\cdot p_j(k) - (p_i(k)+p_j(k)))\cdot q_{ij}^2 - (1 - p_i(k) - p_j(k) - p_i(k)\cdot p_j(k) + \varepsilon_{ij}(k))\cdot q_{ij} + \varepsilon_{ij}(k) = 0\\
\implies q_{ij} = &\frac{b_{ij} - \sqrt{b_{ij}^2 - 4\cdot\varepsilon_{ij}(k)\cdot a_{ij}}}{2\cdot a_{ij}} \implies q_{ij} = \frac{\varepsilon_{ij}(k)}{b_{ij} + \sqrt{b_{ij}^2 - 4\cdot\varepsilon_{ij}(k)\cdot a_{ij}}}.
\end{aligned}$$

Above the last inequality follows because $b_{ij} \ge (1 - p_i(k) - p_j(k))$.

The proof of the second part for Procedure 2 is identical to the proof of claim for the first part by replacing $\varepsilon_{ij}(k) = |\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(0)\cdot\tilde{X}^{\mathbf{u}'}_{j,\ell}(0)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(0)]\cdot\mathbb{E}[\tilde{X}^{\mathbf{u}'}_{j,\ell}(0)]|$, $p_i(k), p_j(k)$ by $p_1(k), p_2(k)$ (defined in the statement of the lemma.) and q by $N\cdot q$. This concludes the proof. □

B.3 Analysis of Procedure 1

Next, we bound the marginals and bias in the set of random variables $\{X^{\mathbf{u}}_{i,\ell}(\bar{k}) : \mathbf{u}_{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}\}$ at the end of Procedure 1. Here, $\bar{k}$ denotes the number of iterations performed of Procedure 1. We prove the following claim that lower bounds the marginals $p_i(\bar{k})$ for all $i \in [m]$.

Claim B.5. *For any $k \geq 0$ and $i \in [m]$, we have that $p_i(k) \geq p_i(0)$ at the end of the k-th iteration of Procedure 1. In addition, we have that,*

1. *For any pair of distinct $i, j \in [m]$, $q_{ij} \leq \frac{1}{(M-1)}$ when $p_{ij}(0) > p_i(0) \cdot p_j(0)$ and $q_{ij} \leq \frac{1}{(M-1)^2 \cdot N^2}$ when $p_{ij}(0) < p_i(0) \cdot p_j(0)$.*
2. *For any $i \in [m]$, $p_i(\bar{k}) \leq p_i(0) + \frac{m-1}{N\cdot(M-1)}$ at the end of Procedure 1.*

Proof. We first observe that $p_i(k) \leq \frac{1}{2N}$ for large enough $N, M > 0$ which follows inductively. This implies that for any $k > 0$, we have $p_i(k+1) = p_i(k)$ if weight class i is not processed at the iteration k of Procedure 1. Otherwise if weight classes i, j is processed for some $j \in [m]$, we have $p_i(k+1) \geq (1 - q_{ij}) \cdot p_i(k) + q_{ij}/2 \geq p_i(k)$. This concludes the proof of the first part.

Second, we observe that for any distinct $i, j \in [m]$, we can bound q_{ij} as follows:

$$\begin{aligned} q_{ij} &\leq \frac{N \cdot \varepsilon_{ij}(k)}{(p_i(k) + p_j(k))} \leq \frac{\varepsilon_{ij}(0)}{(p_i(0) + p_j(0))} \\ &\leq \frac{N \cdot |\frac{M}{M-1} \cdot p_{ij} - p_i \cdot p_j|}{M^2 \cdot N^2 \cdot \min(p_i(0), p_j(0))} \leq \frac{N \cdot M}{M-1} \cdot \frac{\min(p_i(0), p_j(0))}{M \cdot N \cdot \min(p_i(0), p_j(0))} \\ &= \frac{1}{(M-1)}. \end{aligned}$$

Above, the first inequality holds due to Claim B.4, the second inequality holds because of Claim B.2 and the first part of the claim $p_i(k) \geq p_i(0)$. In the other case, we have,

$$q_{ij} \leq \frac{\varepsilon_{ij}(k)}{(1 - p_i(k) - p_j(k))} \leq \frac{\varepsilon_{ij}(0)}{1 - p_i(k) - p_j(k)} \leq \frac{N \cdot |\frac{M}{M-1} \cdot p_{ij} - p_i \cdot p_j|}{M^2 \cdot N^2 \cdot (1 - 1/M)} \leq \frac{1}{(M-1)^2 \cdot N^2}.$$

Finally, the probability $p_i(\bar{k})$ at the end of Procedure 1 can be bound by union bound on the event that at least one of the Bernoulli $Z \sim \mathrm{Ber}(q_{ij})$ for some $j \neq i \in [m]$ turns out to be 1 throughout Procedure 1. This implies,

$$p_i(\bar{k}) \leq p_i(0) + \frac{1}{N} \cdot \sum_{j \neq i} q_{ij} \leq p_i(0) + \frac{m-1}{(M-1) \cdot N}.$$

This completes the proof. □

Next, we analyze the type of remaining correlation after the end of Procedure 1. We recall that at the end of Procedure 1, the only correlation we have left is between $X^{\mathbf{u}}_{i,\ell}$ and $X^{\mathbf{u}'}_{i,\ell'}$ for $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, $i \in [m]$ and $\ell, \ell \in [N]$ and between $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}}_{j,\ell'}$ for $\mathbf{u} \in \mathbf{M}$, $i, j \in [m]$ and $\ell, \ell' \in [N]$. We make the following claim.

Claim B.6. *Let Procedure 1 run for $\bar{k}$ many iterations. The following statements holds:*

1. *For any $i \in [m]$ and pair of $\mathbf{u}, \mathbf{u}' \in \boldsymbol{BigM}$ and $\ell, \ell' \in [N]$, the random variables $X^{\mathbf{u}}_{i,\ell}(\bar{k})$ and $X^{\mathbf{u}'}_{i,\ell'}(\bar{k})$ are positively correlated. In addition,*

$$\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{i,\ell'}(\bar{k})] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k})] \cdot \mathbb{E}[X^{\mathbf{u}'}_{i,\ell'}(\bar{k})] \leq \frac{m}{4 \cdot N^2 \cdot (M-1)}.$$

2. *For any distinct $i, j \in [m]$, $\mathbf{u} \in \boldsymbol{BigM}$ and $\ell, \ell' \in [N]$, we have $X^{\mathbf{u}}_{i,\ell}(\bar{k})$ and $X^{\mathbf{u}'}_{j,\ell'}(\bar{k})$ are negatively correlated. In addition,*

$$\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k})] \cdot \mathbb{E}[X^{\mathbf{u}'}_{j,\ell}(\bar{k})] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell}(\bar{k})] \leq \frac{1}{N^2 \cdot M^2}.$$

Proof. We first consider any $i \in [m]$ and pair of $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$ and $\ell, \ell' \in [N]$, we observe that before we start Procedure 1, $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(0) \cdot X^{\mathbf{u}'}_{i,\ell'}(0)] = 0$ as there is at most one element $\mathbf{v} \in \mathbf{M}$ that takes weight w_i. Therefore, at iteration $\bar{k}$, $X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell}(\bar{k}) = 1$ holds iff $X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell}(\bar{k})$ with probability q_{ij} while resolving correlation between $i, j \in [m]$ where $p_{ij} < p_i \cdot p_j$ or with probability $\frac{q_{ij}}{4 \cdot N^2}$ while resolving correlation between $i, j \in [m]$ where $p_{ij} > p_i \cdot p_j$. This implies,

$$\begin{aligned}
\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell}(\bar{k})] &= \sum_{j \neq i : (p_{ij} < p_i \cdot p_j)} q_{ij} + \sum_{j \neq i : (p_{ij} > p_i \cdot p_j)} \frac{q_{ij}}{4N^2} \\
&\leq \sum_{j \neq i : (p_{ij} < p_i \cdot p_j)} \frac{1}{(M-1)^2 \cdot N^2} + \sum_{j \neq i : (p_{ij} > p_i \cdot p_j)} \frac{1}{4N^2} \cdot \frac{1}{(M-1)} \\
&\leq \frac{m}{4 \cdot N^2 \cdot (M-1)}.
\end{aligned}$$

Above, the first inequality follows from Claim B.5 and the last inequality follows from the fact that $\frac{1}{4N^2} \cdot \frac{1}{(M-1)} > \frac{1}{(M-1)^2 \cdot N^2}$ for $M > 3$. The positive correlation between $X^{\mathbf{u}}_{i,\ell}(\bar{k}), X^{\mathbf{u}'}_{j,\ell}(\bar{k})$ follows from the fact that Procedure 1 mixes the original distribution with positively correlated distribution among $X^{\mathbf{u}}_{i,\ell}(\bar{k}), X^{\mathbf{u}'}_{j,\ell}(\bar{k})$ with marginal higer than the original marginals. This above bound on $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell}(\bar{k})]$ concludes the proof of the first part.

Finally, for any distinct $i, j \in [m]$, $\mathbf{u} \in \mathbf{BigM}$ and $\ell, \ell' \in [N]$, we have $X^{\mathbf{u}}_{i,\ell}(\bar{k}) \cdot X^{\mathbf{u}'}_{j,\ell'}(\bar{k}) = 1$, only if while resolving correlation between i, j during Procedure 1, it ends up assigning $X^{\mathbf{u}}_{i,\ell}(k+1) = 1 : \forall \mathbf{u} \in \mathbf{BigM}, \ell \in [N]$ and $X^{\mathbf{u}}_{j,\ell}(k+1) = 1 : \forall \mathbf{u} \in \mathbf{BigM}, \ell \in [N]$. In this case, q_{ij} satisfies, $p_{ij}(k+1) = p_i(k) \cdot p_j(k+1)$. Since, $\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k) \cdot X^{\mathbf{u}'}_{j,\ell'}(k)] < p_{ij}(k)$ at round k, $X^{\mathbf{u}}_{i,\ell}(k), X^{\mathbf{u}'}_{j,\ell'}(k)$ remains negatively correlated at iteration k while the correlation between i, j is being resolved. During all the other iterations of Procedure 1, Claim B.2 ensures that $X^{\mathbf{u}}_{i,\ell}(\bar{k}), X^{\mathbf{u}'}_{j,\ell'}(\bar{k})$ remains negatively correlated. □

B.4 Analysis of Procedure 2

Given the structure and bound on the correlation between $X^{\mathbf{u}}_{i,\ell}$ and $X^{\mathbf{u}'}_{i,\ell'}$ for $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, $i \in [m]$ and $\ell, \ell' \in [N]$ and between $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}}_{j,\ell'}$ for $\mathbf{u} \in \mathbf{M}$, $i, j \in [m]$ and $\ell, \ell' \in [N]$ in Proposition B.6, we next show that the probability that Procedure 2 alters the original distribution is small which is crucial to bound the total variation distance between the original and the final distribution.

First, we observe that the order in which Procedure 2 resolves correlation does not affect its termination. Therefore, we first analyze the total probability of deviation from the original distribution while resolving the positive correlations. For simplicity, we assume that for the first $M \cdot N - 1$ many iterations of Procedure 2, we resolve correlation between $X^{\mathbf{u}}_{i,1}$ with $X^{\mathbf{u}'}_{i,\ell}$ for all $\mathbf{u}'^{i,\ell'} \neq \mathbf{u}^{i,1} \in \mathbf{BigM}_{[m \cdot N]}$. We first bound $\sum_{k \leq MN-1} q_k$. Before, we obtain the bound on $\sum_{k \leq MN-1} q_k$, we prove the following crucial claims:

Claim B.7. *For $k \leq MN - 1$, we have*

$$\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k)] = \prod_{p<k}(1 - q_p) \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0)] + \frac{1}{2} \cdot \sum_{p \leq k} q_p \cdot \left(\prod_{k \geq j > p} \left(1 - \frac{q_j}{2}\right) \right).$$

Proof. Since, fir $k < MN - 1$, we resolves the correlation between $X^{\mathbf{u}}_{i,1}$ with $X^{\mathbf{u}'}_{i,\ell}$, at each iteration $p < k$, we assign $X^{\mathbf{u}}_{i,1}(p) = 1$ independently at each round with probability $q_p/2$. Finally, we have $\tilde{X}^{\mathbf{u}}_{i,1}(k) = 1$ iff we have $\tilde{X}^{\mathbf{u}}_{i,1}(p) = 1$ and Procedure 2 does not assign $\tilde{X}^{\mathbf{u}}_{i,1}(j)$ for all $j > p$. Combining the above argument for the set of disjoint events $\tilde{X}^{\mathbf{u}}_{i,1}(p-1) = 0$ and $\tilde{X}^{\mathbf{u}}_{i,1}(j) = 1$ for all $j \geq p$, we obtain the proof of the claim. □

Next, we prove the monotonicity of the probability q_k for $k = 1, \ldots, MN - 1$.

Claim B.8. *For any $k < MN - 1$, we have $q_k > q_{k+1}$.*

Proof. This claim simply follows from the fact that $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k+1)] \geq \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k)]$, $\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k+1) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(k+1)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k+1)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(k+1)] = (1 - q_k) \cdot (\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{1,\ell}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(k)])$ and $q_{k+1} \leq \frac{\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k+1) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(k+1)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k+1)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(k+1)]}{\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k)]}$ due to Lemma B.4 and Claim B.2. □

Finally, we are now ready to bound the the probability q_k.

Lemma B.9. *For $k \leq MN - 1$, we have*

$$q_k \leq \frac{\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(0)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(0)]}{\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0)] + \frac{1}{4} \cdot \sum_{p \leq k} \frac{\varepsilon}{p_0 + \frac{p \cdot \varepsilon}{2 \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(0)]}}}.$$

Proof. First, we observe that for the first $MN - 1$ iterations of Procedure 2, we resolve correlation between $\tilde{X}^{\mathbf{u}}_{i,1}$ and $\tilde{X}^{\mathbf{u}'}_{i,\ell}$ therefore due to Claim B.2, at iteration k when we resolve correlation between $\tilde{X}^{\mathbf{u}}_{i,1}$ and $\tilde{X}^{\mathbf{u}'}_{i,\ell}$ for some $\ell \in [N]$ and $\mathbf{u}' \in \mathbf{BigM}$, we have,

$$\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(k)] = \left(\prod_{p<k}(1 - q_p)\right) \cdot \varepsilon,$$

for $\varepsilon = \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0) \cdot \tilde{X}^{\mathbf{u}'}_{i,\ell}(0)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}'}_{i,\ell}(0)]$. For the sake of simplicity, we define $p_0 := \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,1}(0)]$. Therefore, we can bound,

$$\begin{aligned}
q_k &\leq \frac{\left(\prod_{p<k}(1 - q_p)\right) \cdot \varepsilon}{\mathbb{E}[X^{\mathbf{u}}_{i,1}(k)]} = \frac{\left(\prod_{p<k}(1 - q_p)\right) \cdot \varepsilon}{\prod_{p<k}(1 - q_p) \cdot p_0 + \frac{1}{2} \cdot \sum_{p \leq k} q_p \cdot \left(\prod_{k \geq j > p} \left(1 - \frac{q_j}{2}\right)\right)} \\
&= \frac{\varepsilon}{p_0 + \frac{1}{2} \cdot \frac{\sum_{p \leq k} q_p \cdot \left(\prod_{k \geq j > p} \left(1 - \frac{q_j}{2}\right)\right)}{\prod_{p<k}(1 - q_p)}} \\
&\leq \frac{\varepsilon}{p_0 + \frac{1}{2} \cdot \frac{\sum_{p \leq k} q_p \cdot \left(\prod_{k \geq j > p}(1 - q_j)\right)}{\prod_{p<k}(1 - q_p)}}
\end{aligned}$$

Above, the first inequality holds due to Claim B.4, the first equality holds due to Claim B.7, the second equality holds by simple re-arrangement and the last inequality follows because $1 - \frac{q_j}{2} \geq 1 - q_j$. We now focus on lower-bounding the term $\frac{\sum_{p \leq k} q_p \cdot \left(\prod_{k \geq j > p}(1 - q_j)\right)}{\prod_{p<k}(1 - q_p)}$. We now again expand

the the expression,

$$\begin{aligned}\frac{\sum_{p\le k} q_p \cdot \left(\prod_{k\ge j>p}(1-q_j)\right)}{\prod_{p<k}(1-q_p)} &\ge \frac{1}{2}\cdot \frac{\sum_{p\le k} \frac{\varepsilon\cdot\left(\prod_{p<k}(1-q_j)\right)}{\prod_{j<p}(1-q_p)\cdot p_0+\frac{1}{2}\cdot\sum_{j\le p} q_j\cdot\left(\prod_{p\ge j'>j}\left(1-\frac{q_{j'}}{2}\right)\right)}}{\left(\prod_{p<k}(1-q_j)\right)} \\ &= \frac{1}{2}\cdot\sum_{p\le k}\frac{\varepsilon}{\prod_{j<p}(1-q_p)\cdot p_0+\frac{1}{2}\cdot\sum_{j\le p} q_j\cdot\left(\prod_{p\ge j'>j}\left(1-\frac{q_{j'}}{2}\right)\right)} \\ &\ge \frac{1}{2}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{1}{2}\cdot\sum_{j\le p} q_j} \\ &\ge \frac{1}{2}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}}.\end{aligned}$$

Above, the first inequality holds because of Claim B.4, the second equality holds because $(1-q_j)<1$. The last inequality holds due to Claim B.8, i.e. $q_j \le q_1$ and due to Claim B.4 $q_1 \le \frac{\varepsilon}{p_0}$. Combining this with the earlier inequality, we obtain,

$$q_k \le \frac{\varepsilon}{p_0+\frac{1}{4}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}}}$$

This completes the proof. □

Next, we analyze the bound on q_k obained in the previous lemma. First, wre observe that,

Claim B.10. *For $\varepsilon \in (0,1)$, $p_0 \in (0,1)$ and $\frac{\varepsilon}{p_0^2} > 1$, function $h(\varepsilon, p_0) = \frac{\varepsilon}{p_0+\frac{1}{4}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}}}$ is increasing in ε and decreasing in p_0.*

Proof. First we prove the function's monotonicity in ε. Absolutely, let's analyze the given function and prove that it's increasing with respect to ε. We let $D(\varepsilon) = p_0+\frac{1}{4}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}}$. We need to show that, $\frac{D(\varepsilon)-\varepsilon\cdot D'(\varepsilon)}{[D(\varepsilon)]^2} \ge 0$ which is eqvivelenet to proving that $D(\varepsilon)-\varepsilon\cdot D'(\varepsilon) > 0$. We observe that, $D'(\varepsilon) = \frac{1}{4}\cdot\sum_{p\le k}\frac{p_0}{[p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}]^2}$. This implies that,

$$p_0+\frac{1}{4}\cdot\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}} - \varepsilon\cdot\frac{1}{4}\cdot\sum_{p\le k}\frac{p_0}{[p_0+\frac{p\cdot\varepsilon}{2\cdot p_0}]^2} = p_0+\frac{1}{4}\sum_{p\le k}\left(\frac{p\varepsilon^2}{2p_0(p_0+\frac{p\varepsilon}{2p_0})^2}\right) > 0.$$

Next, we analyze its monotonicity property in p_0.

Let's analyze the function with respect to p_0. We let, $D(p_0) = p_0+\frac{1}{4}\sum_{p\le k}\frac{\varepsilon}{p_0+\frac{p\varepsilon}{2p_0}}$. Therefore, to complete the claim, we need to show that $D'(p_0) > 0$. We observe that $D'(p_0) = 1-\frac{\varepsilon}{4}\sum_{p\le k}\frac{1-\frac{p\varepsilon}{2p_0^2}}{\left(p_0+\frac{p\varepsilon}{2p_0}\right)^2}$. Since, $\frac{\varepsilon}{p_0^2} > 1$ (due to positive correlation at the beginning of Procedure 2), we get

$$D'(p_0) > 1-\frac{\varepsilon}{4}\sum_{p\le k}\frac{1-\frac{p}{2}}{\left(p_0+\frac{p\varepsilon}{2p_0}\right)^2} > 1.$$

Above the last inequality follows because $k > 2$. □

Finally, due to Claim B.6, we get the following lemma.

Lemma B.11. *Let Procedure 2 resolve positive correlations for the first $k^* \leq m \cdot M^2 \cdot N$ many iterations, then we have*

$$\sum_{k \leq k^*} q_k \leq O\left(\frac{m \cdot M^2}{\log(M \cdot N)}\right) \leq O\left(\frac{m}{M}\right).$$

Proof. First, we let the procedure run for each $\mathbf{u} \in \mathbf{BigM}$ and resolve positive correlations for all pairs $X^{\mathbf{u}}_{i,1}$ with $X^{\mathbf{u}'}_{i,\ell}$ for all $\mathbf{u}'^{i,\ell'} \neq \mathbf{u}^{i,1} \in \mathbf{BigM}_{[m \cdot N]}$. Since the correlation between the pair of random variables goes down after each iteration, we can bound the total deviation $\sum_{k \leq k^*} q_k \leq M \cdot \sum_{k=1}^{MN-1} q_k$. In other words, we need to bound the total deviation while resolving the correlation between for all $\mathbf{u}'^{i,\ell'} \neq \mathbf{u}^{i,1} \in \mathbf{BigM}_{[m \cdot N]}$ for the first considered element $\mathbf{u} \in \mathbf{BigM}$. We can bound,

$$\begin{aligned}
\sum_{k=1}^{MN-1} q_k &\leq \sum_{k=1}^{MN-1} \frac{\varepsilon}{p_0 + \frac{1}{4} \cdot \sum_{p \leq k} \frac{\varepsilon}{p_0 + \frac{p \cdot \varepsilon}{2 \cdot p_0}}} \\
&\leq \sum_{k=1}^{MN-1} \frac{\frac{m}{N^2(M-1)}}{\frac{1}{MN} + \frac{1}{4} \cdot \sum_{p \leq k} \frac{\frac{m}{N^2(M-1)}}{\frac{1}{MN} + p \cdot \frac{m}{2N}}} \\
&\leq \sum_{k=1}^{MN-1} \frac{\frac{m}{N^2(M-1)}}{\frac{1}{MN} + \frac{1}{2MN} \cdot \sum_{p \leq k} \frac{1}{p+1}} \\
&\leq \frac{2m}{N} \cdot \sum_{k=1}^{MN-1} \frac{1}{1 + \log k} \leq O\left(\frac{m \cdot M}{\log(MN)}\right).
\end{aligned}$$

Above, the first inequality holds due to Claim B.9. The second inequality holds due to Claim B.10 and $p_0 \geq \frac{1}{M \cdot N}$ and $\varepsilon \leq \frac{m}{(M-1) \cdot N}$. The third inequality holds because $\sum_{p \leq k} \frac{1}{p+1} \geq \frac{1}{\log k}$ and the final inequality holds because $\sum_{k=1}^{MN} \frac{1}{\log k}$ is the order of $O(MN/\log(MN))$. Since $N \geq \Omega\left(2^{M^3}\right)$, we conclude the lemma. □

Next, we analyze the total probability of deviation from the original distribution while resolving the negative correlations. This case is simpler because the correlation resolution requires to alter the distribution with a small probability.

Lemma B.12. *For any $k \geq 0$, $i \in [m]$, $\mathbf{u} \in \boldsymbol{BigM}$ and $\ell \in [N]$, we have $\mathbb{E}[\tilde{X}_{i,\ell}(k)] \geq \mathbb{E}[\tilde{X}_{i,\ell}(0)]$ at the end of the k-th iteration of Procedure 2. In addition, at iteration k, for distinct $i, j \in [m]$, $\mathbf{u} \in \boldsymbol{BigM}$ and $\ell, \ell' \in [N]$ if correlation between $\tilde{X}^{\mathbf{u}}_{i,\ell}(k)$ and $\tilde{X}^{\mathbf{u}}_{j,\ell}(k)$ being resolved then, $q_k \leq \frac{M \cdot m}{(M-1)^3 \cdot N^2}$.*

Proof. We first observe that $\mathbb{E}[\tilde{X}_{i,\ell}(k)] \leq \frac{1}{2}$ for large enough $N, M > 0$ which follows inductively. This implies that for any $k > 0$, we have $\mathbb{E}[\tilde{X}_{i,\ell}(k+1)] = \mathbb{E}[\tilde{X}_{i,\ell}(k)]$ if $\tilde{X}_{i,\ell}(k)$ is not processed at the iteration k of Procedure 1. Otherwise, we have $\mathbb{E}[\tilde{X}_{i,\ell}(k+1)] \geq (1 - q_k) \cdot \mathbb{E}[\tilde{X}_{i,\ell}(k)] + q_{ij}/2 \geq \mathbb{E}[\tilde{X}_{i,\ell}(k)]$. This concludes the proof of the first part. Next, Lemma B.4 implies that,

$$\begin{aligned}
q_k &\leq \frac{\mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k)] \cdot \mathbb{E}[\tilde{X}^{\mathbf{u}}_{j,\ell'}(k)] - \mathbb{E}[\tilde{X}^{\mathbf{u}}_{i,\ell}(k) \cdot X^{\mathbf{u}}_{j,\ell'}(k)]}{(1 - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k)] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k)])} \\
&\leq \frac{\mathbb{E}[X^{\mathbf{u}}_{i,\ell}(0)] \cdot \mathbb{E}[X^{\mathbf{u}}_{j,\ell'}(0)] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(0) \cdot X^{\mathbf{u}}_{j,\ell'}(0)]}{(1 - \mathbb{E}[X^{\mathbf{u}}_{i,\ell}(k)] - \mathbb{E}[X^{\mathbf{u}}_{i,\ell'}(k)])} \\
&\leq \frac{\frac{m}{(M-1)^2 \cdot N^2}}{1 - p_i(k) - p_j(k)} \leq \frac{M \cdot m}{(M-1)^3 \cdot N^2}.
\end{aligned}$$

Above, the first inequality follows from Claim B.4, the second inequality follows because of Claim B.2. □

B.5 Wrapping Things Up

Finally, we show that the resultant random variables at the end of the procedure are close to the original random variables in terms of total variation distance and complete the proof of Theorem 5.9.

Proof of Theorem 5.9. Let $\tilde{\mathbf{X}} := \tilde{X}^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$ be the random variables after applying Procedures 1 and 2 on the set of random variables $\mathbf{X} := X^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$. Due to Lemma B.1, we have that $\tilde{X}^{\mathbf{u}}_{i,\ell} : \forall \mathbf{u}^{i,\ell} \in \mathbf{BigM}_{[m\cdot N]}$ are pairwise independent. Finally, we can bound the total variation distance between $\mathbf{X}$ and $\tilde{\mathbf{X}}$.

Let k_1 and k_2 be the number of iteration performed by Procedures 1 and 2, respectively. We divide the iterations of Procedure 2 into two parts, during the first $\bar{k}_2$ many iterations, Procedure 2 resolves positive correlation between $X^{\mathbf{u}}_{i,\ell}$ and $X^{\mathbf{u}'}_{i,\ell'}$ for $\mathbf{u}, \mathbf{u}' \in \mathbf{BigM}$, $i \in [m]$ and during the last $k_2 - \bar{k}_2$ many iterations, it resolves negative correlation between $X^{\mathbf{u}}_{i,\ell}, X^{\mathbf{u}}_{j,\ell'}$ for $\mathbf{u} \in \mathbf{M}$, $i, j \in [m]$ and $\ell, \ell' \in [N]$.

$$\begin{aligned}
\mathrm{TV}_{\mathbf{X},\tilde{\mathbf{X}}} &\leq \sum_{i\neq j\in[m]} q_{ij} + \sum_{k=1}^{k_1} q_k + \sum_{k=k_1+1}^{\bar{k}_2} q_k + \sum_{k=\bar{k}_2+1}^{k_2} q_k \\
&\leq \frac{m^2}{(M-1)\cdot N} + \frac{m}{M^2}\cdot M + |k_2 - \bar{k}_2| \cdot \frac{M\cdot m}{(M-1)^3\cdot N^2} \\
&\leq \frac{2m^3}{M}.
\end{aligned}$$

The first inequality follows because it bounds the total probability that either Procedure 1 or Procedure 2 alters the original distribution $\mathbf{X}$. The second inequality follows because of Lemmas B.12 and B.11. The last inequality follows because $(k_2 - k_1) \leq M \cdot N^2 \cdot m^2$ combining with $N = \Omega(2^{M^2})$ and $M = \Omega(2^m)$. □