

An Algorithmic Solution to the Blotto Game using Multi-marginal Couplings

VIANNEY PERCHET, Crest, ENSAE and Criteo AI LAB

PHILIPPE RIGOLLET, Department of Mathematics, MIT

THIBAUT LE GOUIC, Aix Marseille University, Centrale Marseille, CNRS, I2M

We describe an efficient algorithm to compute solutions for the general two-player Blotto game on n battle-fields with heterogeneous values. While explicit constructions for such solutions have been limited to specific, largely symmetric or homogeneous, setups, this algorithmic resolution covers the most general situation to date: value-asymmetric game with asymmetric budget with sufficient symmetry and homogeneity. The proposed algorithm rests on recent theoretical advances regarding Sinkhorn iterations for matrix and tensor scaling. An important case which had been out of reach of previous attempts is that of heterogeneous but symmetric battlefield values with asymmetric budget. In this case, the Blotto game is constant-sum so optimal solutions exist, and our algorithm samples from an ε -optimal solution in time $\tilde{O}(n^2 + \varepsilon^{-4})$, independently of budgets and battlefield values. In the case of asymmetric values where optimal solutions need not exist but Nash equilibria do, our algorithm samples from an ε -Nash equilibrium with similar complexity but where implicit constants depend on various parameters of the game such as battlefield values.

1 INTRODUCTION

A century ago, Emile Borel published his seminal paper on *the theory of play and integral equations with skew symmetric kernels* [Bor21], see also [?, Page 157]. While perhaps not as conspicuous, it predates von Neumann’s monumental work on *the theory of games of strategy* [von28] by several years. In this work, Borel describes what is now called the Blotto game and introduces the notions of strategy, mixed strategies and even foresees the fruitful interactions between game theory and economics that are to be observed throughout the century. As such, the Blotto game is considered to be the genesis of modern game theory [Fré53, Nak06]. Despite its prestigious pedigree, equilibrium strategies for this game are only known in special cases.

Blotto is a resource-allocation game in which two players compete over n different battlefields by simultaneously allocating resources to each battlefield. The following two additional characteristics are perhaps the most salient features of the Blotto game:

- (1) *Winner-takes-all*: For each battlefield, the player allocating the most resources to a given battlefield wins the battlefield.
- (2) *Fixed budget*: each player is subject to a fixed—and deterministic—budget that mixed strategies should satisfy almost surely.

Despite its apparent simplicity the Blotto game captures a variety of practical situations that extend far beyond the context of the above military terminology. These include political strategy [Mye93, LP02, MMT05], network security [LHSR15, FSM21], and various forms of practical auction markets [MS15, HM17].

The goal of this paper is to efficiently construct a Nash equilibrium for this game or, when they exist, an optimal strategy.

Prior work. Despite its century-long existence, Nash equilibria for the Blotto game are only known under various restrictions on the main parameters of the problem: the budget of each player and the value given to each battlefield.

- *Budget*. A large fraction of the literature considers the case where the players have *symmetric budgets*, starting with the original problem of Borel [Bor21] and in most of the main contributions throughout the twentieth century [Bor21, BV38, Gro50, GW50, Las02, Tho18]. The case of symmetric budgets is well understood except in the setup where players may disagree on the value of battlefield that was recently introduced [KR21].
- *Battlefields*. When the two players have a different budget the situation becomes more complex as the poorest will have to forfeit some battlefields. In this case, only partial results are known. To understand what “partial” means, recall that full generality of the battlefield values occurs when (i) players may assign a different value to a given battlefield—we say that the values are *asymmetric*—and (ii) these values may vary across battlefield—we say that the battlefields are *heterogeneous*. Partial results are known for symmetric values. Even under this simplifying assumption, the case of heterogeneous battlefields remains poorly understood, except in the case of two battlefields [MM15]. In the case of more than two battlefields, Nash equilibria are known for homogeneous battlefields [Rob06] or under stringent assumptions on the battlefield values [SLS14] that essentially reduce to the homogeneous case.

We refer the reader to Table 1 for a survey of recent advances. While we tackle the most general setup to date, we stress that an important case was not covered by prior literature: the case of asymmetric budget, heterogeneous and symmetric values. Indeed, in this case, the game is constant-sum and optimal strategies exist. Our results also cover the case of asymmetric values introduced very

recently in [KR21] but this setup leads to only Nash equilibria rather than optimal strategies. Whenever possible, we conflate the two setups and simply refer to a *solution* to the Blotto.

A discrete version of the Blotto game where both budgets and allocations are required to be integral was also introduced in Borel's original paper [Bor21]. Explicit optimal solutions were provided in [Har08] for the homogeneous and symmetric version of the discrete game; see also [HVL12] for partial solutions in the asymmetric-value case. More recently, this discrete version has seen significant computational advances [ADH⁺19, Bea22]. Conceptually, this line of work is close to the present paper in the sense that it provides an algorithm to sample from approximate solutions. Moreover, the discrete Blotto game can be seen as a discretization of the continuous version of interest here and that could be quantified using the arguments of Section 4. However carrying out this analysis, for instance based on [Bea22, Theorem 4.2], leads to worse dependence on n and ϵ compared to Theorem 12 here. More strikingly, the complexity bound of Theorem 12 does not depend on budgets or battlefield values while this dependence is polynomial in the bounds for discrete Blotto. The two lines of work also differ in more profound ways. First and foremost, the approach employed here is fundamentally different: it aims at mixing known solutions for the related Lotto game while solutions to the discrete Blotto games are more agnostic so that it is unclear what the marginals of the resulting strategy are. In particular, the present approach allows us sample from ϵ -Nash equilibria in the asymmetric-value case whereas this setup is currently out of reach for solutions to the discrete Blotto game.

Finally, note that our approach also yields new (existential) results for the discrete Blotto game. Since they are not the focus of our contribution, they are relegated to the appendix.

Our contributions. All of the above solutions for two-player games have consisted in constructing explicit solutions. Because of the budget constraints, these strategies can be decomposed in two parts: *marginal distributions* that indicate which (random) strategy to play on each battlefield and a *coupling* that correlates the marginal strategies in such a way to ensure that the budget constrained is satisfied almost surely.

The first question may be studied independently of the second by considering what is known as the (General) *Lotto* game [BC80]. In this game the budget constraint need only be enforced in *expectation* with respect to the randomization of the mixed strategies. While this setup lacks a defining characteristic of the Blotto game (fixed budget), it has the advantage of landing itself to more amenable computations. Indeed, unlike the Blotto game, a complete solution to the Lotto game was recently proposed in [KR21] where the authors describe an explicit Nash equilibrium in the most general case: asymmetric budget, asymmetric and heterogeneous values.

In light of this progress a natural question is whether the marginal solutions discovered in [KR21] can be *coupled* in such a way that the budget constraint is satisfied almost surely. We provide a positive answer to this question by appealing to an existing result from the theory of *joint mixability* [WW16]. Mixability asks the following question: Can n random variables X_1, \dots, X_n with prescribed marginal distributions $X_i \sim P_i$ be coupled in such a way that $\text{var}(X_1 + \dots + X_n) = 0$. Joint mixability is precisely the step required to go from a Lotto solution to a Blotto one by coupling the marginals of the Lotto solution in such a way that the budget constraint is satisfied.

In this paper we exploit a simple and new connection between joint mixability and the theory of *multi-marginal couplings* that has recently received a regain of interest in the context of optimal transport [AC11, DMGN17, ABA20]. In multi-marginal optimal transport, the goal is to optimize a cost over the space of couplings with given marginals. Unlike the case of two marginals that arises in traditional optimal transport, this question raises significant computational challenges and often leads to NP-hardness [ABA21]. In the language of optimization, joint mixability merely asks if the set of constraints is nonempty. We propose an algorithmic solution to the Blotto problem

by efficiently constructing a coupling that satisfies the budget constraint almost surely and can be easily sampled from. Our construction relies on three key steps: first we reduce the problem to a small number of marginals to bypass the inherent NP hardness of multi-marginal problems, second we discretize the marginals and finally, we employ a multi-marginal version of the Sinkhorn algorithm [Sin64, SK67] to construct a coupling of the discretized marginals. After a simple smoothing step, we produce a sampling with continuous marginals that are close to the ones prescribed by the Lotto solutions and from which it is straightforward to sample. Furthermore, we quantify the combined effect of discretization error and of the Sinkhorn algorithm on the value of the game, effectively leading to an *approximate* Nash equilibrium and even to an approximately optimal solution in the case of symmetric values.

We exhibit tight—or near-tight in the asymmetric values case—conditions for the mixability of *specific* Lotto solutions into Blotto solutions; see Corollaries 6 and 7 below. While these conditions are reasonable and cover most cases, some heavily skewed games, either in terms of budget asymmetry or values inhomogeneity, are not covered by our results. We leave it as an open question to exhibit Lotto strategies that can be mixed into Blotto ones even for such games.

The rest of this paper is organized as follows. In the next section, we recall the solution for the Lotto game and show that they can be turned into solutions for the Blotto game. This existential result simply appeals to existing results of joint mixability. We move from an existential to an algorithmic result in Section 3 by proceeding in three steps: first we reduce the problem to the case $n = 4$, then we discretize the problem and finally we apply Sinkhorn algorithm to couple the resulting marginals in an appropriate fashion. The main product of Section 3 is Algorithm 8 which shows how to sample from an approximate solution to the Blotto game. Finally, we provide a detailed complexity analysis for this algorithm in Section 4, showing in particular, that it runs in time polynomial in the parameters of the Blotto game and the approximation error ε . Finally, our techniques also yield new results for the discrete Blotto game largely studied by [Har08, HVLS12] that are of independent interest. We postpone them to the appendix.

NOTATION. For any integer n , define $[n] = \{1, \dots, n\}$. We use 1 to denote an all-ones vector or tensor. Note that the dimension of this vector will be clear from the context but may vary across occurrences. For any two vectors x, y , we denote their entrywise (Hadamard) product $x \odot y$ and their entrywise division $x \oslash y$ whenever y has only nonzero entries. For any two real numbers a, b we denote by $a \vee b$ their maximum and by $a \wedge b$ their minimum.

2 SOLUTIONS FOR BLOTTO AND LOTTO GAMES

The goal of this section is to describe the Blotto game and its connection to the Lotto game for which explicit solutions are known. We first recall a solution for the Lotto game derived in [KR21] and show that it can be readily turned into a Blotto solution using the theory of joint mixability.

2.1 The Blotto game

The classical two-player *Blotto* game is formalized as follows. Two players, respectively denoted by A and B , are competing over $n \geq 2$ battlefields denoted by $i \in [n]$. Since we focus on two-player games where both players obey the same rules, it will be convenient when describing the game to denote by $P \in \{A, B\}$ either player and by \bar{P} the other player so that $(P, \bar{P}) \in \{(A, B), (B, A)\}$.

The *datum* of a Blotto game is as follows. Player $P \in \{A, B\}$ has a total budget of $T_P > 0$ to allocate across the n battlefields. Moreover, she evaluates battlefield $i \in [n]$ to $v_{P,i} > 0$ which may differ from $v_{\bar{P},i}$. Without loss of generality, we assume that $T_A \geq T_B$ to break symmetry and that

$$\sum_{i \in [n]} v_{A,i} = \sum_{i \in [n]} v_{B,i} = 1 \tag{1}$$

Indeed, multiplying the value of all battlefields by the same constant has no impact on the players' strategies.

The *rules* of the Blotto game are as follows. A pure strategy for player P is an allocation vector $x_P = (x_{P,1}, \dots, x_{P,n})$ where $x_{P,i} > 0$ is the amount allocated to battlefield $i \in [n]$. A mixed strategy for player P is a probability distribution over pure strategies. A salient feature of the Blotto game is that a player P is constrained to playing strategies that satisfy the budget constraint: $x_{P,1} + \dots + x_{P,n} \leq T_P$. In turn, admissible mixed strategies for the Blotto games are random vectors $X_P \in \mathbb{R}^n$ such that

$$\sum_{i=1}^n X_{P,i} \leq T_P \quad \text{almost surely.} \quad (2)$$

Given two pure strategies x_P and $x_{\bar{P}}$ for players P and \bar{P} respectively, player P wins battlefield $i \in [n]$ if $x_{P,i} > x_{\bar{P},i}$ and receives a reward $v_{P,i} > 0$. Ties $x_{P,i} = x_{\bar{P},i}$ are broken arbitrarily as they are irrelevant for our analysis.

The existence of Nash equilibria is a consequence of standard game theoretic arguments [Ren99]. Unfortunately, these general results say little about the structure of equilibrium strategies. At the end of this section, we make partial progress towards this question by describing the marginals of such equilibrium strategies. However, these remain existential results in essence.

This is in stark contrast with the associated Lotto game, described in the following section, where the hard budget constraint is dropped in favor of a constraint in expectation, and whose explicit solutions have been computed.

2.2 The associated Lotto game

A *Lotto* game has the same data and rules as its associated Blotto game except for the almost sure budget constraint (2) which is relaxed to the following *expected* budget constraint:

$$\sum_{i=1}^n \mathbb{E}[X_{P,i}] \leq T_P \quad (3)$$

This relaxation greatly simplifies the game. In fact, Kovenock and Roberson [KR21] have recently elicited an explicit characterization of a non-trivial Nash equilibrium for the most general version of the Lotto game to date; see Table 1. In the rest of Section 2.2, we describe their solution in details since it is the basis for ours.

Finding an optimal strategy for the Lotto game amounts to finding a stationary point for an optimization problem subject to constraints of the form (3). Because of linearity of expectation, the associated Lagrangian is decomposed as the sum of n terms, one per battlefield, that are each mathematically equivalent to an “all-pay” auction whose solutions are well known.

More explicitly, Nash equilibria of the Lotto problem depend on two parameters $\gamma \geq 0$ and $\lambda \geq 0$, that are set later on. First, given any $\gamma \geq 0$, consider the subsets of battlefields $\mathcal{N}(\gamma)$ that are at least γ -times more valuable to A than to B :

$$\mathcal{N}(\gamma) = \{i \in [n], \frac{v_{A,i}}{v_{B,i}} \geq \gamma\}.$$

Given a scaling parameter $\lambda \geq 0$ to be defined later, the mixed strategy of player A at equilibrium prescribes to allocate a (random) budget of $X_{A,i}$ to battlefield i with distribution given by:

	Continuous Strategy	Asymm. Budget	Heterogeneous Values	Asymm. Values	More than 3 Battlefields	Complete Results
[BV38]	✓	✗	✗	✗	✗	✓
[GW50]	✓	✗	✗	✗	✓	✓
[Wei12]	✓	✗	✗	✗	✓	✓
[Gro50]	✓	✗	✓	✗	✓	✓
[Las02]	✓	✗	✓	✗	✓	✓
[Tho18]	✓	✗	✓	✗	✓	✓
[GW50]	✓	✓	✓	✗	✗	✓
[Rob06]	✓	✓	✗	✗	✓	✓
[MM15]	✓	✓	✓	✗	✗	✓
[SLS14]	✓	✓	✓	✗	✓	✗
[KR21]	✓	✓	✓	✓	✓	✗
This paper	✓	✓	✓	✓	✓	✓

Table 1. Variants of the continuous Blotto game and their solutions. The last column, “complete results” indicates whether results obtained hold with possibly strong assumptions on the different values (for instance, there always exist more than 3 battlefields with the exact same value).

$$X_{A,i} \sim \begin{cases} \text{Unif} \left[0, \frac{\gamma v_{B,i}}{\lambda} \right] & \text{if } i \in \mathcal{N}(\gamma) \\ \left(1 - \frac{v_{A,i}}{\gamma v_{B,i}} \right) \delta_0 + \frac{v_{A,i}}{\gamma v_{B,i}} \text{Unif} \left[0, \frac{v_{A,i}}{\lambda} \right] & \text{if } i \notin \mathcal{N}(\gamma). \end{cases},$$

where δ_0 denotes the Dirac point mass at 0.

The strategy of player B is given by

$$X_{B,i} \sim \begin{cases} \left(1 - \frac{\gamma v_{B,i}}{v_{A,i}} \right) \delta_0 + \frac{\gamma v_{B,i}}{v_{A,i}} \text{Unif} \left[0, \frac{\gamma v_{B,i}}{\lambda} \right] & \text{if } i \in \mathcal{N}(\gamma), \\ \text{Unif} \left[0, \frac{v_{A,i}}{\lambda} \right] & \text{if } i \notin \mathcal{N}(\gamma) \end{cases}.$$

Note that the strategy of A and B are the same except that the roles of $v_{A,i}$ and $\gamma v_{B,i}$ are switched. In that sense, γ plays the role of an “exchange” rate that accounts for discrepancies between budgets and valuations across the two players.

It remains to find the parameters γ and λ using the budget constraints. For this set of strategies, saturating the total budget constraint (2) readily yields the following two equations:

$$\lambda T_A = \sum_{i \in \mathcal{N}(\gamma)} \frac{\gamma v_{B,i}}{2} + \sum_{i \notin \mathcal{N}(\gamma)} \frac{(v_{A,i})^2}{2\gamma v_{B,i}} = \frac{1}{2} \sum_{i=1}^n (\gamma v_{B,i}) \wedge \frac{(v_{A,i})^2}{\gamma v_{B,i}}, \quad (4)$$

$$\lambda T_B = \sum_{i \in \mathcal{N}(\gamma)} \frac{(\gamma v_{B,i})^2}{2v_{A,i}} + \sum_{i \notin \mathcal{N}(\gamma)} \frac{v_{A,i}}{2} = \frac{1}{2} \sum_{i=1}^n \frac{(\gamma v_{B,i})^2}{v_{A,i}} \wedge v_{A,i}. \quad (5)$$

Any pair (γ, λ) solving the above system of two equations yields a Nash equilibrium. It remains to show that such solutions may be computed efficiently. Observe that eliminating λ from the equations yields the following nonlinear equation in γ :

$$\begin{aligned}
f(\gamma) &:= \gamma^3 \left(T_A \sum_{i \in N(\gamma)} \frac{v_{B,i}^2}{v_{A,i}} \right) - \gamma^2 T_B \sum_{i \in N(\gamma)} v_{B,i} + \gamma T_A \sum_{i \notin N(\gamma)} v_{A,i} - T_B \sum_{i \notin N(\gamma)} \frac{v_{A,i}^2}{v_{B,i}} \\
&= \gamma T_A \sum_{i=1}^n \left(\gamma^2 \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} \right) - T_B \sum_{i=1}^n \frac{v_{A,i}}{v_{B,i}} \left(\gamma^2 \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} \right) = 0
\end{aligned} \tag{6}$$

Any solution γ^* to this equation readily yields a unique λ^* by plugging it into either (4) or (5); both equations will yield the same solution by (6). In turn, the existence and efficient computation of solutions γ^* to (6) are ensured by the following proposition. The bounds on γ^* presented in the following proposition depend on the distance between the vectors of battlefield values $v_A := (v_{A,1}, \dots, v_{A,n})$ and $v_B := (v_{B,1}, \dots, v_{B,n})$. Interestingly the natural measure of distance that emerges is the χ^2 -divergence that commonly arises in information theory and statistics; see e.g. [PW22]. The χ^2 -divergence $\chi^2(u||v)$ between two probability vectors $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$ is defined by

$$\chi^2(u||v) = \sum_{i=1}^n \frac{u_i^2}{v_i} - 1 = \sum_{i=1}^n \left(\frac{u_i}{v_i} - 1 \right)^2 v_i.$$

It is clear that $\chi^2(u||v) \geq 0$ with equality if and only if $u = v$.

PROPOSITION 1. *Equation (6) has the following properties:*

- (1) *It always has at least one and at most $3n + 3$ solutions γ^* .*
- (2) *Any solution γ^* satisfies*

$$\frac{T_B}{T_A} \frac{1}{1 + \chi^2(v_B||v_A)} \leq \gamma^* \leq \frac{T_B}{T_A} (1 + \chi^2(v_A||v_B))$$

- (3) *Computing all solutions can be done in $O(n \log n)$ operations.*

The proof is based on standard computations, hence postponed to Appendix B.2, with the associated Algorithm 2.

REMARK 2. *In case of symmetric values, that is when $v_{A,i} = v_{B,i} = v_i$, the game is constant-sum and each player has a then unique¹ optimal strategy given by a unique pair (γ^*, λ^*) [KR21]. In fact, in that case, the unique (γ^*, λ^*) can be computed analytically as $\gamma^* = T_B/T_A$ and $\lambda^* = T_B/(2T_A^2)$; this can be easily seen from Proposition 1 (point 2.), since $\chi^2(v_A||v_B) = \chi^2(v_B||v_A) = 0$. With these parameters, the optimal strategy of player A is to choose $X_{A,i}$ uniformly at random on $[0, 2T_A v_i]$ and that of player B is to forfeit each battlefield with probability $1 - \frac{T_B}{T_A}$ and, to choose $X_{B,i}$ uniformly at random on $[0, 2T_A v_i]$ on battlefield i if not forfeited.*

2.3 From Lotto to Blotto

In the previous section, we described how to compute solutions of a Lotto game. To turn a strategy for the Lotto game into a strategy for the Blotto game, one can couple the marginal strategies of a Lotto game, effectively turning the constraint (3) on the expected budget into the almost sure budget constraint (2).

Stated otherwise, a solution to the associated Lotto game induces a solution to the original Blotto game if the random variables $\{X_{A,i}\}_{i \in [n]}$ (and similarly $\{X_{B,i}\}_i$) are jointly mixable [WW16].

¹In the case of the Lotto game, it is natural to call a strategy an equivalence class of strategies with the same marginals.

DEFINITION 3. A family of k random variables Z_1, \dots, Z_k with finite expectations is jointly mixable if there exists a coupling π such that if $(Z_1, \dots, Z_k) \sim \pi$,

$$\sum_{i=1}^k Z_i = \sum_{i=1}^k \mathbb{E}[Z_i], \quad \text{almost surely.}$$

In that case, the coupling π is called a joint mix.

Obviously, not all random k -tuples variables are jointly mixable. Take for example Z_1, Z_2 and Z_3 to be Bernoulli with parameter 1/2. Then $\mathbb{E}[Z_1] + \mathbb{E}[Z_2] + \mathbb{E}[Z_3] = 3/2$ whereas there is no coupling of the Z_i s such that their sum equals a fractional number.

While the full characterization of jointly mixable distribution is a complex question, some conditions, either sufficient or necessary, for joint mixability have been derived. The following proposition is a simple extension of a result of [WW16] (see also, [Zim20]) on the mixability of distributions with monotone densities.

PROPOSITION 4. For $i = 1, \dots, k$, let $p_i \in [0, 1]$, $b_i > 0$ be fixed parameters and let Z_i be a random variable with distribution given by the following mixture:

$$Z_i \sim (1 - p_i)\delta_0 + p_i \text{Unif } [0, b_i]. \quad (7)$$

Then Z_1, \dots, Z_k are jointly mixable if and only if

$$\max_{1 \leq i \leq k} b_i \leq \frac{1}{2} \sum_{i=1}^k p_i b_i \quad (8)$$

This proposition is a consequence of few computations; its proof is delayed to Section B.3.

We are now in a position to state the main result of this section: the marginal distributions of the Lotto game described above are jointly mixable into a solution to the Blotto game. To that end, we instantiate Proposition 4 to the parameters of the marginal distributions described in Section 2.2.

THEOREM 5. Let γ^*, λ^* be the parameters of Nash equilibrium of the Lotto game described in Section (2.2). Then the marginal distributions can be coupled into a Nash equilibrium for the corresponding Blotto game if and only if

$$\max_{i \in [n]} (\gamma^* v_{B,i} \wedge v_{A,i}) \leq \lambda^* T_B. \quad (9)$$

Inequality (9) is simply an instantiation of (8), hence details are postponed to Section B.3

Condition (9) of the previous theorem relies on the values γ^*, λ^* that define the solution of the Lotto game. In light of the bounds obtained in Proposition 1, these parameters may be eliminated to produce a sufficient condition for the existence of said solution for Blotto games with symmetric values. Recall that in this case, the game is constant-sum so a solution is, in fact, an optimal strategy. This result is captured in the following corollary which is a straightforward consequence of Theorem 5.

COROLLARY 6. Assume symmetric values: $v_A = v_B = v$. Then the marginal distributions of the optimal Lotto strategy described in Section 2.2 with $\gamma^* = T_B/T_A$ and $\lambda^* = T_B/(2T_A^2)$ can be coupled into an optimal strategy for the corresponding Blotto game if and only if

$$\max_{i \in [n]} v_i \leq \frac{T_B}{2T_A}.$$

In fact a sufficient condition may be derived in the case of non-symmetric values.

COROLLARY 7. *Assume that battlefield are balanced in the sense that there exists $r \in (0, 1)$ such that*

$$\chi^2(v_A \| v_B) \vee \chi^2(v_B \| v_A) \leq r^2$$

Then, the marginal distributions of the optimal Lotto strategy described in Section 2.2 can be coupled into an optimal strategy for the corresponding Blotto game as long as

$$\max_{i \in [n]} v_{B,i} \leq \frac{T_B}{2T_A}(1 - r).$$

The proof of this result is based solely on computations; it is postponed to Section B.4

Note that the result of Corollary 7 is tight in the sense that if $r \rightarrow 0$ it recovers the result of Corollary 6. It is unclear whether the dependence in r is sharp in our result and it is an interesting question to address in future work.

Under rather general conditions, the above two corollaries show the *existence* of solutions with marginal distributions of the Lotto game derived in [KR21]. It remains to show that such a coupling may be realized efficiently. This is done in the next section.

3 AN EFFICIENT ALGORITHM TO COMPUTE SOLUTIONS

Deriving solutions, either optimal strategies in the constant-sum setting or Nash equilibria, remains one of the major open problems surrounding the Blotto game. Previous attempts at this task have focused on deriving an *explicit* coupling between marginals. This is possible in specific cases. For example, several explicit couplings between n random variables $X_i \sim \text{Unif}[0, 1], i = 1, \dots, n$ are known [KS06, RU02]. In particular, this provides a solution to some Blotto problems with sufficient symmetry. However, this explicit approach fails for more general problems, and, in particular in the important case of asymmetric budget such as the one covered in Corollary 6. In this paper, we take another route by describing the efficient Algorithm LOTTO2BLOTO, whose pseudo-code is postponed to the Appendix B.1, that computes an ε -approximate solution with time complexity which is polynomial in n and $1/\varepsilon$.

In light of the previous section, our goal is to find an algorithm that efficiently computes a coupling between the marginal Lotto strategies described above. This task faces two major hurdles.

On the one hand, the continuous nature of the marginals described above does not lend itself to efficient algorithms which typically work with discrete quantities. Instead, we propose to simply discretize the marginals at a scale of order $\varepsilon > 0$. In particular, this prevents us from replicating exactly the marginals of the Lotto game but we can show that the error employed in said discretization remains of the same order once propagated to the utility of a given player.

On the other hand, the mere description of a coupling between n discrete marginals on $O(1/\varepsilon)$ atoms is an object of size $O(1/\varepsilon^n)$, which is exponential in the number n of battlefields. To overcome this limitation, we develop a careful scheme that allows us to reduce the problem to the case of 4 marginals instead of n .

Finally, we employ recent developments in computational optimal transport, to couple our 4 marginals using a variant of Sinkhorn iterations [Sin64, SK67, Cut13].

3.1 Reductions

The typical size of a coupling with n marginals is exponential in n . While this issue is, in general, hopeless to overcome, we can exploit some of the structure of the problem at hand. Indeed, a similar principle has been recently employed in multi-marginal optimal transport to devise polynomial-time algorithms under additional structure [ABA20]. More specifically, we reduce our problem to the case where there are only four marginals which remain mixable if the original marginals are mixable.

This reduction is done in two steps. Recall that the marginals for the Lotto game described in Section 2.2 are either uniform distributions or mixtures of a uniform distribution with a Dirac point mass at zero. Our first step reduces to the case where $n - 1$ marginals are uniform and only one is a mixture as above. In our second step, we further reduce to the case where there are three uniform marginals and one mixture.

Throughout this section we focus on player A for brevity. Reductions for player B are analogous.

3.1.1 Step 1: reduction to a single mixture. The marginal distributions described in Section 2.2 consist of $|\mathcal{N}(\gamma^*)|$ uniform distributions and $n - |\mathcal{N}(\gamma^*)|$ mixtures of a uniform distribution and a Dirac point mass at 0 and our goal is to efficiently couple them into a joint mix coupling π that has these marginals and satisfies the Blotto budget constraint. For clarity, we also regard uniform distributions as mixture distributions albeit with weight zero on the point mass. Otherwise, we say that a distribution is a *strict* mixture. The goal of this first step is to reduce this coupling problem to the case where there are $n - 1$ uniform distributions and one single strict mixture. To that end, we show that such a coupling π may be obtained as a mixture of n joint mixes $\pi_k, k = 1, \dots, n$:

$$\pi = \sum_{k=1}^n q_k \pi_k, \quad q_k \geq 0, \quad \sum_k q_k = 1$$

where the marginal distributions of π_k consist of at most one strict mixture, the rest being uniform distributions. Moreover, this decomposition can be computed efficiently as the solution of a simple greedy procedure.

LEMMA 8. *Let γ^*, λ^* be the parameter of a solution for the Lotto game and assume that the mixability condition (9) holds. Then, there exists a family π_1, \dots, π_n of couplings and a set of non negative weights $q_k \geq 0, \sum_k q_k = 1$ such that*

(1) *The marginal distributions of $(X_{A,1}^k, \dots, X_{A,n}^k) \sim \pi_k$ are given by*

$$X_{A,i}^k \sim (1 - p_i^{(k)})\delta_0 + p_i^{(k)} \text{Unif} \left[0, \frac{v_{A,i} \wedge (\gamma^* v_{B,i})}{\lambda^*} \right]$$

for some $p_i^{(k)} \in [0, 1]$, $i, k \in [n]$ with at most one $p_i^{(k)}$ in $(0, 1)$ for each k .

(2) *Each coupling $\pi_k, k = 1, \dots, n$ is a joint mix*

(3) *The mixture of couplings*

$$\pi = \sum_{k=1}^n q_k \pi_k. \tag{10}$$

is a solution for the Blotto game.

(4) *The total complexity of computing the weights $q_k, p_i^{(k)}, i, k \in [n]$ scales as $O(n^2 \log n)$.*

Note that the mixture of couplings π in (10) is necessarily a joint mix as a mixture of joint mixes. To sample from it, Player A , simply samples π_k with probability q_k and plays according to the strategy prescribed by it.

The geometric proof of this lemma is delayed to Section B.5, along with the pseudo-code of associated Algorithm DECOMP.

LEMMA 9. *Fix $\ell \in \mathbb{R}_+^n$, $T \in \mathbb{R}$ and let $\mathcal{H} = \{x \in \mathbb{R}^n \text{ s.t. } \langle \ell, x \rangle = T\}$ be an affine hyperplane. For any point $x \in C_n = \mathcal{H} \cap [0, 1]^n$, define $\text{supp}(x) := \{i \in [n], x_i \in (0, 1)\}$; then there exist $\theta \in [0, 1]$, an extreme point $y \in C_n$ and another vector $\bar{y} \in C_n$ satisfying $\text{supp}(\bar{y}) \subsetneq \text{supp}(x)$ (in particular, \bar{y} belongs to some m -face of C_n where $m = |\text{supp}(\bar{y})| < |\text{supp}(x)|$) such that*

$$x = \theta y + (1 - \theta)\bar{y}.$$

The overall complexity of computing y, \bar{y} and θ is of order $O(n \log n)$.

The proof of this Lemma is based on simple geometric arguments, and is postponed to Section B.6 with the associated pseudo-code of Algorithm EXTREMIZE.

3.1.2 Step 2. Reduction to four random variables. The previous step reduces the joint mixability problem of n general mixtures, to a simpler one where at most one strict mixture is involved. Still, computing—in fact even describing—a coupling of n variables requires generically exponential (in n) time and memory. To overcome this limitation, we reduce the number of random variables from n to a constant number.

The following Lemma states that each π_k can be realized as the coupling of 3 new uniform random variables and a strict mixture, thus reducing the mixability question from n to only 4 random variables. A careful inspection of the proof of Lemma 10 below indicates that the reduction may lead to three marginals rather than four. In that case, two marginals are uniform and one is a strict mixture. To handle this case, some adjustments are needed; in particular—and obviously—with the size of the resulting coupling. However, extensions from four to three marginals are straightforward and we omit this case for clarity.

LEMMA 10. *Fix $n \geq 4$, $k \in [n]$, and assume without loss of generality that the last marginal of the coupling π_k from Lemma 8 is a strict mixture. Then $(X_{A,1}, \dots, X_{A,n}) \sim \pi_k$ may be constructed from three uniform random variables Y_1, Y_2, Y_3 and a partition $\mathcal{I}_0 \sqcup \mathcal{I}_1 \sqcup \mathcal{I}_2 \sqcup \mathcal{I}_3 = [n-1]$ as follows. Set $X_{A,i} = 0$ for all $i \in \mathcal{I}_0$, and*

$$X_{A,i} = \theta_i Y_j, \quad i \in \mathcal{I}_j, \quad j \in \{1, 2, 3\},$$

where $\theta_1, \dots, \theta_{n-1} \in [0, 1]$ are such that

$$\sum_{i \in \mathcal{I}_j} \theta_i = 1, \quad j = 1, 2, 3.$$

In particular, it holds that

$$\sum_{i=1}^{n-1} X_{A,i} = Y_1 + Y_2 + Y_3 \quad \text{almost surely},$$

and $(Y_1, Y_2, Y_3, X_{A,n})$ are jointly mixable. The support of Y_j is $[0, b_j^*]$ where

$$b_j^* = \sum_{i \in \mathcal{I}_j} \frac{Y^* v_{B,i}}{\lambda^*} \wedge \frac{v_{A,i}}{\lambda^*}$$

Moreover, the θ_i 's, the sets \mathcal{I}_j , and the parameters of the distributions of Y_1, Y_2, Y_3 can each be computed in constant time.

The proof of this Lemma, based on standard mixability arguments, is postponed to Section B.7, with the pseudo-code of the corresponding Algorithm REDUC.

Note that any joint mix of $(Y_1, Y_2, Y_3, X_{A,n})$ readily yields a joint mix of $(X_{A,1}, \dots, X_{A,n})$ by defining $X_{A,i} = \theta_i Y_{j(i)}$, where $j(i) \in [3]$ is the unique integer such that $i \in \mathcal{I}_{j(i)}$.

3.2 Discretization

The problem of finding a solution for the Blotto game has been reduced to the construction of a coupling of (at most) four random variables, three of them being uniform over some intervals and the fourth one being a mixture between a Dirac mass at zero and some uniform distribution.

Throughout this section we denote these random variables as Y_1, \dots, Y_4 for simplicity; in the notation of the previous section, they correspond to Y_1, Y_2, Y_3 , and $X_{A,n}$ respectively.

Unfortunately, even in this simple case, finding explicit, closed-form, couplings appears to be possible only under stringent additional conditions that limit the scope of the Blotto game. To overcome this limitation, we take an algorithmic approach, describing an efficient way to find an approximate solution. To that end, we obviously need to work with discrete random variables and describe here a coupling between these discretized random variables.

Let $(Y_1, Y_2, Y_3, Y_4) \sim \omega$ be jointly mixed so that

$$Y_1 + Y_2 + Y_3 + Y_4 = T_A. \quad (11)$$

Moreover, let $h > 0$ be some (small) discretization parameter. Define the quantized random variables \tilde{Y}_i by

$$\tilde{Y}_i = \left\lfloor \frac{Y_i}{h} \right\rfloor, \quad i = 1, 2, 3, \quad \tilde{Y}_4 = \frac{T_A}{h} - \left\lfloor \frac{T_A - Y_4}{h} \right\rfloor. \quad (12)$$

Our goal is to compute any of the joint distributions $D(\omega)$ of the vector $\tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \tilde{Y}_4)$ when ω ranges over joint mixes.

As a first step towards this goal, note that these discretized random variables need not be jointly mixable. Indeed, in general we have $\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 \leq \lfloor (Y_1 + Y_2 + Y_3)/h \rfloor$ but equality may fail to hold because of discretization errors. To account for these, let $\varepsilon \in \{0, 1, 2\}$ be defined as

$$\varepsilon = (T_A/h - \tilde{Y}_4) - (\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3), \quad (13)$$

and consider the augmented random vector $\tilde{Y}_+ = (\tilde{X}, \varepsilon)$. In light of (13), $\tilde{Y}_+ \in \mathbb{R}^5$ lives almost surely on a four dimensional subspace. As such, its distribution may be represented by a 4-tensor (Γ_{ijke}) with entries given by

$$\Gamma_{ijke} = \mathbb{P}(\tilde{Y}_1 = i, \tilde{Y}_2 = j, \tilde{Y}_3 = k, \varepsilon = e).$$

In particular, note that $e \in \{0, 1, 2\}$ while i, j, k each range in a set of integers of size $\Theta(1/h)$. Using (13) we can read off the distribution of \tilde{Y}_4 from this tensor.

This tensor is subject to four sets of linear constraints, one for each of the marginal constraints given in (12). They are given by

$$\sum_{jke} \Gamma_{ijke} = \mathbb{P}(\tilde{Y}_1 = i) \quad \forall i, \quad \sum_{ike} \Gamma_{ijke} = \mathbb{P}(\tilde{Y}_2 = j) \quad \forall j, \quad \sum_{ije} \Gamma_{ijke} = \mathbb{P}(\tilde{Y}_3 = k) \quad \forall k,$$

and, in light of (13), by

$$\sum_{i+j+k+e=\ell} \Gamma_{ijke} = \mathbb{P}(T_A/h - \tilde{Y}_4 = \ell) \quad \forall \ell.$$

Note that indeed, any draw from a distribution that satisfies the above constraints yields a random vector $(\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \varepsilon)$. Defining \tilde{Y}_4 by solving (13) yields a vector $\tilde{Y} \sim D(\omega)$ for some joint mix ω defined as above. In other words, \tilde{Y} is indeed the discretization of random variables drawn from a joint mix (though it need not be jointly mixable itself).

Since the random variables Y_j , for $j \in [3]$, constructed at the previous step have a support equal to $[0, b_j^*]$ where $b_j^* = \sum_{i \in I_j} \frac{Y_i^* v_{B,i}}{\lambda^*} \wedge \frac{v_{A,i}}{\lambda^*}$, the reduced (to 4 random variables) and discretized problem reduces to finding some tensor (Γ_{ijke}) with $3 \cdot \left\lfloor \frac{b_1^*}{h} + 1 \right\rfloor \cdot \left\lfloor \frac{b_2^*}{h} + 1 \right\rfloor \cdot \left\lfloor \frac{b_3^*}{h} + 1 \right\rfloor$ entries satisfying at most $4 \cdot \left\lfloor \frac{\max b_j^*}{h} + 1 \right\rfloor$ linear constraints. Although this can be done simply via linear programming (hence polynomially in h^{-1} , more precisely in $\tilde{O}(1/h^{8.5})$ with Vaidya's algorithm), a quite efficient

and more popular way is to use a variant of Sinkhorn-Knopp algorithm that quickly finds approximated solutions. This is more relevant as this linear program is already some approximation of the original problem, hence there is no point of solving it exactly.

The pseudo-code of the Algorithm DISCRETIZE can be found in Section B.8

3.3 Tensor scaling using Sinkhorn iterations

In light of the previous sections, we have reduced our problem to that of finding coupling in the form of a 4-tensor $(\Gamma_{ijke}, i \in [d_1], j \in [d_2], k \in [d_3], e \in \{0, 1, 2\})$ with non-negative entries subject to marginal constraints. We approach this problem from a computational perspective and propose an algorithm that converges rapidly to a feasible solution. To describe this algorithm, recall that its input are four probability vectors $\tilde{\mu}_j \in \mathbb{R}^{d_j}, j = 1, \dots, 4$, with $d_4 = d_1 + d_2 + d_3 + 2$ that represent the probability mass functions of the discretized random variable $\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, T_A/h - \tilde{Y}_4$ defined in the previous section: $\tilde{\mu}_{j, \cdot} = \mathbb{P}(\tilde{Y}_j = \cdot), j = 1, \dots, 3, \tilde{\mu}_{4, \cdot} = \mathbb{P}(T_A/h - \tilde{Y}_4 = \cdot)$.

The linear constraints take the form

$$\bar{\Gamma}_i^{(1)} := \sum_{jke} \Gamma_{ijke} = \tilde{\mu}_{1,i} \quad \forall i \in [d_1], \quad (14)$$

$$\bar{\Gamma}_j^{(2)} := \sum_{ike} \Gamma_{ijke} = \tilde{\mu}_{2,j} \quad \forall j \in [d_2], \quad (15)$$

$$\bar{\Gamma}_k^{(3)} := \sum_{ije} \Gamma_{ijke} = \tilde{\mu}_{3,k} \quad \forall k \in [d_3], \quad (16)$$

$$\bar{\Gamma}_\ell^{(4)} := \sum_{i+j+k+e=\ell} \Gamma_{ijke} = \tilde{\mu}_{4,\ell} \quad \forall \ell \in [d_4]. \quad (17)$$

Denote by \mathcal{G} the set of tensors $\Gamma = (\Gamma_{ijke})$ that satisfy these constraints.

To solve this problem, we propose to project the all-ones tensor 1 onto \mathcal{G} using the Kullback-Leibler (KL) divergence. Recall that the KL divergence between two nonnegative tensors Γ, Γ' is given by

$$\text{KL}(\Gamma \parallel \Gamma') = \sum_{ijke} \Gamma_{ijke} \log \left(\frac{\Gamma_{ijke}}{\Gamma'_{ijke}} \right)$$

In particular, $\text{KL}(\Gamma \parallel 1)$ is simply the (negative) entropy $H(\Gamma)$ of Γ and we aim to solve the convex optimization problem

$$\min_{\Gamma \in \mathcal{G}} H(\Gamma) = \sum_{ijke} \Gamma_{ijke} \log (\Gamma_{ijke}).$$

While many algorithms are available to solve this problem [Bub15], its specific structure can be exploited efficiently. Indeed, first order optimality conditions imply that any optimal Γ must be of the form

$$\Gamma_{ijke} = \xi_{1,i} \cdot \xi_{2,j} \cdot \xi_{3,k} \cdot \xi_{4,i+j+k+e}, \quad (18)$$

for some scaling vectors $\xi_j \in (0, \infty)^{d_j}, j = 1, \dots, 4$. This representation readily calls for an iterative tensor scaling algorithm similar to the Sinkhorn algorithm [Sin64, SK67, Cut13]. Tensor scaling has been investigated in more classical setups [LHCJ20, ABA20] that slightly differ from the present setup because the fourth marginal constraint takes a special form. Nevertheless, the implementation of Algorithm SINKHORN remains straightforward and is presented in Section B.9. Its analysis is also a straightforward extension of that for the traditional matrix case [AWR17]. More specifically, following the exact same lines as the one of Theorem 4.3 in [LHCJ20], we readily get the following result.

PROPOSITION 11. Define

$$\tilde{\mu}_{\min} = \min_{i \in [4], j, \tilde{\mu}_{i,j} \neq 0} \tilde{\mu}_{i,j}.$$

Algorithm SINKHORN terminates and returns a tensor Γ such that $\sum_{i=1}^4 \|\bar{\Gamma}^{(i)} - \tilde{\mu}_i\|_1 \leq \eta$ after at most $32\eta^{-1}(1 - \log \tilde{\mu}_{\min})$ iterations. Moreover, each marginal $\bar{\Gamma}^{(i)}$ of Γ has positive entries that sum to one and hence is a probability vector.

What have we accomplished so far? Through several reductions and a tensor scaling algorithm, given the datum of a Blotto game, we are able to compute a joint distribution that corresponds to an approximate solution. In Section 4, we evaluate the accuracy of this approximation in terms of the value of the game by showing that the various approximations (discretization and numerical precision of the algorithm) do not blow up when propagated back into the reductions. Before that, we investigate an important operational question: how to sample a strategy from the resulting coupling Γ .

3.4 From coupling to sampling

Finding an efficient construction of (approximate) equilibria or optimal strategies is only relevant if it can be associated to some efficient sampling method so that a player may query a sampler and receive the allocation $(X_{A,1}, \dots, X_{A,n})$ that they should play on each battlefield. In light of the various reduction steps employed above, it is sufficient to sample a 4-tuple

$$(\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \varepsilon) \in [0, b_1^*/h] \times [0, b_2^*/h] \times [0, b_3^*/h] \times \{0, 1, 2\}$$

from the output Γ of Algorithm SINKHORN. Indeed, from $(\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \varepsilon)$, we obtain the random variables $Y_i, i = 1, \dots, 4$ that are approximately distributed from the joint mix ω as follows.

To ensure that the marginal distributions are continuous, let $U \sim \text{Unif}[0, 1]$ and define

$$Y'_1 = (\tilde{Y}_1 + \frac{\varepsilon}{3} + \frac{U}{3})h \wedge b_1^*, \quad Y'_2 = (\tilde{Y}_2 + \frac{\varepsilon}{3} + \frac{U}{3})h \wedge b_2^*, \quad Y'_3 = (\tilde{Y}_3 + \frac{\varepsilon}{3} + \frac{U}{3})h \wedge b_3^*.$$

To correct for potential boundary effects, define $S = Y'_1 + Y'_2 + Y'_3$ and

$$\zeta = \mathbb{1}\{S > T_A\} \frac{T_A - S}{3} + \mathbb{1}\{S < T_A - b_4^*\} \frac{T_A - b_4^* - S}{3}.$$

Then take $\bar{Y}_1 = Y'_1 + \zeta$, $\bar{Y}_2 = Y'_2 + \zeta$, $\bar{Y}_3 = Y'_3 + \zeta$, and $\bar{Y}_4 = T_A - (Y_1 + Y_2 + Y_3)$.

We call this procedure the *smoothing* procedure. Finally, as mentioned before, just define $X_{A,i} = \theta_i \bar{Y}_{j(i)}$, where $j(i) \in [3]$ is the unique integer such that $i \in \mathcal{I}_{j(i)}$.

Note that the random variable $U \sim \text{Unif}[0, 1]$ is superfluous and theoretical results would follow by taking $U = 0$. Its role is simply to ensure, for cosmetic reasons, that the random marginal distributions are continuous apart from the potential point mass at zero.

It remains to sample $(\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \varepsilon)$ from the output Γ of Algorithm SINKHORN. This is quite straightforward in light of the factored form of Γ . Indeed, recall that the coupling output by Algorithm SINKHORN has the form (18).

$$\Gamma_{ijke} = \xi_{1,i} \cdot \xi_{2,j} \cdot \xi_{3,k} \cdot \xi_{4,i+j+k+e}, \quad \forall i \in [d_1], j \in [d_2], k \in [d_3], e \in \{0, 1, 2\}$$

As a consequence, we can draw from Γ as follows:

- (1) Set $\tilde{Y}_1 = i \in [d_1]$ with probability proportional to $\xi_{1,i}$
- (2) Set $\tilde{Y}_2 = j \in [d_2]$ with probability proportional to $\xi_{2,j}$
- (3) Set $\tilde{Y}_3 = k \in [d_3]$ with probability proportional to $\xi_{3,k}$
- (4) Conditionally on $(\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3)$, set $\varepsilon = e \in \{0, 1, 2\}$ with probability proportional to $\xi_{4,\tilde{Y}_1+\tilde{Y}_2+\tilde{Y}_3+e}$.

The pseudo-code of Algorithm SAMPLE can be found in Section B.10.

4 APPROXIMATION ERRORS AND COMPUTATIONAL COMPLEXITY

The construction of the previous section relies on various approximations, each of them inducing some error that can be mitigated at the cost of additional computational complexity by tuning the discretization parameter h of Section 3.2 and the tolerance parameter η in Algorithm SINKHORN. In this section we study the computational complexity required to reach an ε -approximate solution.

4.1 From approximate strategies to approximate solutions

Note that the very notion of “approximate solution” strongly depends on whether the problem is value-symmetric ($v_{A,i} = v_{B,i}$ for all i) or -asymmetric ($v_{A,i} \neq v_{B,i}$ for some i). Indeed, in the former case, the game is constant-sum and optimal strategies do exist. This is no longer true in the latter case where only Nash equilibria are considered. As a consequence, we can consider approximation of a single optimal strategies in value-symmetric games, while we will have to consider approximations of a pair of equilibrium strategies in value-asymmetric ones. In the following two sections, we consider each case separately. In the remaining, we shall focus the analysis on Player A, but it is almost identical for player B; hence we do not repeat it for the sake of clarity.

4.1.1 The value-symmetric case. A value-symmetric Blotto game, where $v_{A,i} = v_{B,i} = v_i$ for all i is constant-sum and optimal strategies exist for each player. In particular, this allows us to provide strong approximation guarantees by controlling how sub-optimal the expected utility of a player is.

To check this well-known fact on our specific instance, consider the utility of player A. Set two equilibrium parameters $\gamma^* = T_B/T_A \leq 1$ and $\lambda^* = T_B/(2T_A^2)$ (see Corollary 6) defining an optimal strategy and observe that $\mathcal{N}(\gamma^*) = [n]$ since $\gamma^* \leq 1$. For $i \in [n]$, let $X_{A,i} \sim \text{Unif}[0, 2T_A v_i]$ denote the amount allocated by player A to battlefield i according to this optimal strategy and denote by $F_{A,i}$ its cumulative distribution function (cdf). The expected utility (a.k.a. reward) of player A if player B chooses allocation $x_B = (x_{B,i})_i$ depends only on the sequence $F_A = (F_{A,1}, \dots, F_{A,n})$ of marginal cdfs rather than the whole coupling. It is given by

$$\mathcal{U}_A(F_A, x_B) = \sum_{i=1}^n v_{A,i} \mathbb{P}(X_{A,i} > x_{B,i}) = \sum_{i=1}^n v_i (1 - F_{A,i}(x_{B,i})) = 1 - \sum_{i=1}^n v_i \wedge \frac{x_{B,i}}{2T_A} \geq 1 - \frac{T_B}{2T_A}.$$

where we used the fact that the v_i ’s sum to 1 and the $x_{B,i}$ ’s sum to at most T_B . Moreover, if B employs the mixed strategy $X_{B,1}, \dots, X_{B,n}$ described in Corollary 6, the utility of player A, denoted $\mathcal{U}_A(F_A, F_B)$, changes as follows. Let $U_{A,i}, U_{B,i} \sim \text{Unif}[0, 2T_A v_i]$ be a sequence of uniform random variables such that $U_{A,i}$ is independent of $U_{B,i}$. In particular, $\mathbb{P}(U_{A,i} > U_{B,i}) = .5$ and

$$\begin{aligned} \mathcal{U}_A(F_A, F_B) &= \left(1 - \frac{T_B}{T_A}\right) \sum_{i=1}^n v_i (1 - F_{A,i}(0)) + \frac{T_B}{T_A} \sum_{i=1}^n v_i \mathbb{P}(U_{A,i} > U_{B,i}) \\ &= \left(1 - \frac{T_B}{T_A}\right) + \frac{T_B}{2T_A} = 1 - \frac{T_B}{2T_A}. \end{aligned}$$

In particular, the strategy of player A given in Corollary 6 is optimal and its optimal utility is given by $1 - T_B/(2T_A)$.

To estimate the cost of the various approximations incurred by player A, let $X_{A,i}^{h,\eta} \sim P_{A,i}^{h,\eta}$ denote the strategy on battlefield $i \in [n]$. The notation is meant to emphasize that the approximation error stems from two sources: the precision level η of Algorithm SINKHORN and the grid size h of discretization procedure in Section 3.2. In particular, we write $P_{A,i}^{0,0} := P_{A,i}$. Cognizant of this approximation error, player A may take advantage of the suboptimality of the strategy of player A and respond with best-response strategy denoted $X_B^{h,\eta} = (X_{B,1}^{h,\eta}, \dots, X_{B,n}^{h,\eta})$. As a result, the

suboptimality gap of player A 's expected utility is controlled as follows

$$\begin{aligned} \mathcal{U}_A(F_A, F_B) - \mathcal{U}_A(F_A^{h,\eta}, F_B^{h,\eta}) &= \sum_{i=1}^n v_i \mathbb{P}(X_{A,i} > X_{B,i}) - \sum_{i=1}^n v_i \mathbb{P}(X_{A,i}^{h,\eta} > X_{B,i}^{h,\eta}) \\ &\leq \sum_{i=1}^n v_i \mathbb{P}(X_{A,i} > X_{B,i}^{h,\eta}) - \sum_{i=1}^n v_i \mathbb{P}(X_{A,i}^{h,\eta} > X_{B,i}^{h,\eta}) \\ &\leq \sum_{i=1}^n v_i \sup_{x>0} \left[\mathbb{P}(X_{A,i} > x) - \mathbb{P}(X_{A,i}^{h,\eta} > x) \right] \end{aligned}$$

where in the first inequality, we use the fact that X_B is an optimal response for B when A plays X_A .

It will be convenient in the sequel to further bound the above quantity using the ∞ -Wasserstein distance—see [San15, Section 5.5.1]—between $P_{A,i}$ and $P_{A,i}^{h,\eta}$, denoted $W_\infty(P_{A,i}, P_{A,i}^{h,\eta})$. Indeed, to show that $W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq \omega$ for some $\omega \geq 0$, it is sufficient to exhibit a coupling of $X_{A,i}, X_{A,i}^{h,\eta}$ such that $|X_{A,i} - X_{A,i}^{h,\eta}| < \omega$ almost surely. Below, we often do so implicitly as such couplings are, in all instances, trivial.

Fix $\omega > 0$ and assume $W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq \omega$. Then for any $\omega \geq 0$, we have

$$\mathbb{P}(X_{A,i} > x) - \mathbb{P}(X_{A,i}^{h,\eta} > x) \leq \mathbb{P}(x < X_{A,i} < x + \omega) \leq \frac{\lambda^* \omega}{\gamma^* v_i},$$

where we used the fact that $X_{A,i} \sim \text{Unif}[0, \gamma^* v_i / \lambda^*]$. The above two displays together yield that the suboptimality gap for player A is controlled as

$$\mathcal{U}_A(F_A, F_B) - \mathcal{U}_A(F_A^{h,\eta}, F_B^{h,\eta}) \leq \frac{\lambda^*}{\gamma^*} \sum_{i=1}^n W_\infty(P_{A,i}, P_{A,i}^{h,\eta}).$$

Recall that Step 1 in the reduction consists in decomposing $P_{A,i}$ as a mixture of (at most) n other distributions, i.e., $P_{A,i} = \sum_k q_k P_{A,i}^{(k)}$. Accordingly, we also have constructed $P_{A,i}^{h,\eta}$ as a mixture $P_{A,i}^{h,\eta} = \sum_k q_k P_{A,i}^{h,\eta,(k)}$. It follows readily from the definition of W_∞ that that

$$W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq \sum_k q_k W_\infty(P_{A,i}^{(k)}, P_{A,i}^{h,\eta,(k)}).$$

In particular controlling each term on the right-hand side uniformly in k results in the same control on the desired error. Therefore, without loss of generality, we may assume that $P_{A,i}^{(k)} = P_{A,i}$ and $P_{A,i}^{h,\eta,(k)} = P_{A,i}^{h,\eta}$ so as to keep the notation light. Moreover, as above, we assume without loss of generality the last marginal $P_{A,n}$ is the only strict mixture.

4.1.2 The value-asymmetric case. When values are asymmetric, the game is no longer constant-sum and we shift our focus from optimal strategies to Nash equilibria. In this context the notion of approximation is more subtle and has to be carried out jointly for both players.

For any set of marginal cdfs $G_A = (G_{A,1}, \dots, G_{A,n})$ and $G_B = (G_{B,1}, \dots, G_{B,n})$, denote by $\mathcal{U}_A(G_A, G_B)$ the expected utility of player A if player A plays according to strategy G_A while player B plays according to strategy G_B .

A pair (F_A, F_B) is a Nash equilibrium if

$$\mathcal{U}_A(F_A, F_B) \geq \mathcal{U}_A(G_A, F_B), \text{ for all admissible } G_A.$$

Writing

$$\begin{aligned}\mathcal{U}_A(F_A, F_B) &= \mathcal{U}_A(F_A, F_B) - \mathcal{U}_A(F_A^{h,\eta}, F_B) + \mathcal{U}_A(F_A^{h,\eta}, F_B) - \mathcal{U}_A(F_A^{h,\eta}, F_B^{h,\eta}) + \mathcal{U}_A(F_A^{h,\eta}, F_B^{h,\eta}), \\ \mathcal{U}_A(G_A, F_B) &= \mathcal{U}_A(G_A, F_B) - \mathcal{U}_A(G_A, F_B^{h,\eta}) + \mathcal{U}_A(G_A, F_B^{h,\eta})\end{aligned}$$

it readily follows from the above two displays that

$$\mathcal{U}_A(F_A^{h,\eta}, F_B^{h,\eta}) + \varepsilon \geq \mathcal{U}_A(G_A, F_B^{h,\eta}), \text{ for all admissible } G_A,$$

where, using similar computations as above,

$$\varepsilon \leq \sum_{i=1}^n v_{A,i} \left[\frac{\lambda^*}{\gamma^* v_{B,i}} W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) + \frac{2\lambda^*}{v_{A,i}} W_\infty(P_{B,i}, P_{B,i}^{h,\eta}) \right]$$

4.2 Control of the errors

In both cases, symmetric or asymmetric values, a control of the approximation error follows from controlling $W_\infty(P_{A,i}, P_{A,i}^{h,\eta})$. In the rest of this section, we slightly abuse notation and write $W_\infty(X, Y)$ when the distributions of the random variables X and Y are clear from the context.

Recall that for any $i \in [n]$ since $X_{A,i} = \theta_i Y_{k(i)}$ and similarly for the approximate versions, for some fixed $\theta_i \in [0, 1]$, we have for any $h, \eta \geq 0$ that

$$W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) = \theta_i W_\infty(Y_{j(i)}, \bar{Y}_{j(i)})$$

where Y_j is the result of the reductions and is defined in Lemma 10 while \bar{Y}_j is the output the smoothing procedure and is defined in Section 3.4.

Recall that the discrepancy between \bar{Y}_j and the target Y_j stems from three approximations: discretization error ($h > 0$), numerical error ($\eta > 0$), and the error due to smoothing step. The error coming from the smoothing step is easy to control: for any $j \in [4]$, we have $W_\infty(Y_j, \bar{Y}_j) \leq W_\infty(Y_j, h\tilde{Y}_j) + h$. We have proved that

$$W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq \theta_i W_\infty(Y_{j(i)}, h\tilde{Y}_{j(i)}) + \theta_i h \tag{19}$$

As a result, it is sufficient to control the discretization error and the numerical error at the level of the variables \tilde{Y}_j . To emphasize the presence of these errors, we employ the same notation as for $P_{A,i}^{h,\eta}$ and write $\tilde{Y}_j = \tilde{Y}_j^{h,\eta}$ for $h > 0, \eta \geq 0$. By the triangle inequality, we have

$$W_\infty(Y_j, h\tilde{Y}_j) \leq \underbrace{W_\infty(Y_j, h\tilde{Y}_j^{h,0})}_{\text{discretization error}} + \underbrace{W_\infty(h\tilde{Y}_j^{h,0}, h\tilde{Y}_j^{h,\eta})}_{\text{numerical error}}.$$

The discretization error is trivial to control. Indeed, in light of the coupling provided by (12), we get that

$$W_\infty(Y_j, h\tilde{Y}_j^{h,0}) \leq h \tag{20}$$

Finally, to control the numerical error $W_\infty(h\tilde{Y}_j^{h,0}, h\tilde{Y}_j^{h,\eta})$, recall that the tolerance $\eta > 0$ in Algorithm SINKHORN controls the ℓ_1 error between the current marginals and the targets. Hence, we need to bound the ∞ -Wasserstein distance by the ℓ_1 distance. This is quite straightforward since $\tilde{Y}_j^{h,\eta}$ has bounded support for $\eta \geq 0$. Indeed, recall from Lemma 10 that for any $j \in [4]$ we have that $\tilde{Y}_j^{h,\eta} \in [0, b_j^*/h]$, where

$$b_j^* = \frac{1}{\lambda^*} \sum_{i \in I_j} (\gamma^* v_{B,i}) \wedge v_{A,i}.$$

Hence, $W_\infty(h\tilde{Y}_j^{h,0}, h\tilde{Y}_j^{h,\eta}) \leq b_j^*\eta$, and we have proved that

$$W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq 2\theta_i h + \theta_i b_{k(i)}^* \eta.$$

In particular, for the value-symmetric case, since

$$b_{k(i)} = \frac{\gamma^*}{\lambda^*} \sum_{l \in I_{k(i)}} v_l \leq \frac{\gamma^*}{\lambda^*},$$

we get the following simple bound

$$W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq 2\theta_i h + \frac{\theta_i \gamma^* \eta}{\lambda^*}.$$

Since $\sum_i \theta_i = 4$, the suboptimality gap for player A is controlled as

$$\mathcal{U}_A(F_A) - \mathcal{U}_A(F_A^{h,\eta}) \leq \frac{\lambda^*}{\gamma^*} \sum_{i=1}^n W_\infty(P_{A,i}, P_{A,i}^{h,\eta}) \leq 8 \frac{\lambda^*}{\gamma^*} h + 4\eta. \quad (21)$$

In the value-asymmetric case, we get a suboptimality for player A smaller than

$$\varepsilon_A = \left(16 + \frac{8}{\gamma^*} \max_i \frac{v_{A,i}}{v_{B,i}} \right) \lambda^* h + \left(8\gamma^* + 4 \max_i \frac{v_{A,i}}{v_{B,i}} \right) \eta \quad (22)$$

and, with symmetric arguments, a suboptimality for player B smaller than

$$\varepsilon_B = \left(16 + 8 \max_i \frac{v_{B,i}}{v_{A,i}} \right) \frac{\lambda^* h}{\gamma^*} + \left(8 + 4 \max_i \frac{v_{B,i}}{v_{A,i}} \right) \eta.$$

4.3 Computational complexity

In this section, we tally the complexity required to achieve either ε -suboptimality gap in the value-symmetric case or an ε -Nash equilibrium in the value-asymmetric case.

Before making this distinction recall the various steps that were employed, together with their computational complexity.

Lotto Step. Computing one (and, actually, all) pair of parameters (γ, λ) requires $\mathcal{O}(n \log n)$ operations, see Proposition 1.

Step 1: Computing all couplings π_k and their associated convex weights q_k requires $\mathcal{O}(n^2)$ operations; see Lemma 8.

Step 2: Given some coupling π_k computed at step 1, the reduction from n to only 4 random variables requires $\mathcal{O}(n^2)$ operations; see Lemma 10.

Step 3: The discretization step is computationally costless.

Step 4: The Sinkhorn algorithm requires $\mathcal{O}(\frac{\log(1/\tilde{\mu}_{\min})}{\eta})$ operations; see Proposition 11. Since the marginal distributions μ_i are h -discretizations of either uniform on an interval of size at most $\sum \frac{\gamma^* v_{B,i}}{\lambda^*} \leq \frac{\gamma^*}{\lambda^*}$, or uniform on interval of length $\frac{v_{A,i}}{\lambda^*}$ with weight $\frac{v_{A,i}}{\gamma^* v_{B,i}}$, it holds that $\tilde{\mu}_{\min} \geq \frac{h\lambda^*}{\gamma^*}$. At each iteration of Sinkhorn, all the components of the tensor are computed, hence a complexity, per iteration, in $\mathcal{O}\left(\left(\frac{\gamma^*}{h\lambda^*}\right)^3\right)$.

Step 5: The sampling cost comes from the reconstruction of the budget allocation from the 4 random variables, constructed at Step 2 and sampled from the coupling computed at step 4.

The sampling step has a linear cost with respect to the discretization size $\mathcal{O}(\frac{\gamma^*}{h\lambda^*})$ while the reconstruction complexity scales linearly with respect to the number of battlefields $\mathcal{O}(n)$.

We are now in a position to state our main theorems. We begin with the symmetric-value case for player A . The result for player B is completely analogous and therefore omitted.

THEOREM 12. Consider the two-player Blotto game on n battlefields with symmetric values v_1, \dots, v_n where player A has budget T_A and player B has budget $T_B \leq T_A$ and assume that these data satisfy the conditions of Corollary 6. Fix $\varepsilon > 0$ and let $\eta = \varepsilon/4$ and $h = \varepsilon T_A/8$. Then the procedure described in Algorithm SAMPLE samples from an ε -suboptimal strategy for player A in time

$$O\left(n^2 + \frac{\log(1/\varepsilon)}{\varepsilon^4}\right)$$

PROOF. Note first that the preprocessing cost associated to steps 1 through 4 is $O(n^2)$.

To compute the cost of Sinkhorn iterations, observe that the parameters η and h are chosen in such a way that each term on the right-hand side of (21) is equal to $\varepsilon/2$:

$$8 \frac{\lambda^*}{\gamma^*} h = 4\eta = \frac{\varepsilon}{2}$$

Hence,

$$\tilde{\mu}_{\min} \geq \frac{h\lambda^*}{\gamma^*} = \frac{\varepsilon}{16}, \quad \text{and} \quad \left(\frac{\gamma^*}{h\lambda^*}\right)^3 = \left(\frac{16}{\varepsilon}\right)^3.$$

Moreover, since $\eta = \varepsilon/8$, we get that the total complexity of Sinkhorn iterations is

$$O\left(\left(\frac{\gamma^*}{h\lambda^*}\right)^3 \frac{\log(1/\tilde{\mu}_{\min})}{\eta}\right) = O\left(\varepsilon^{-3} \frac{\log(1/\varepsilon)}{\varepsilon}\right) = O(\varepsilon^{-4} \log(1/\varepsilon))$$

Finally, the last step has a total cost of $O(n + \varepsilon)$ which is negligible with respect to the combination of previous steps. \square

It is worth noting that in the value-symmetric case, the computational complexity of our procedure is *independent* of the datum of the problem (budgets and values) under the normalization (1). Note that this normalization merely scales the utility and should of course affect the desired accuracy parameter ε .

We now move to the asymmetric-value case and characterize the complexity of our procedure to compute an ε -Nash equilibrium for the Blotto game. As above, we focus on player A only.

THEOREM 13. Consider the Blotto game on n battlefields with asymmetric values $v_{P,1}, \dots, v_{P,n}$, $P \in \{A, B\}$, where player $P \in \{A, B\}$ has budget T_P , with $T_B \leq T_A$ and assume that these data satisfy the conditions of Corollary 6. Define

$$\mathfrak{m} = \max_i \frac{v_{A,i}}{v_{B,i}} \vee \frac{v_{B,i}}{v_{A,i}}.$$

Fix $\varepsilon > 0$ and let

$$\eta = \frac{\varepsilon}{24\mathfrak{m}}, \quad h = \frac{\gamma^*}{\lambda^*} \frac{\varepsilon}{48\mathfrak{m}}.$$

Then Algorithm LOTTO2BLOTO samples from an ε -Nash equilibrium in time

$$O\left(n^2 + \left(\frac{\mathfrak{m}}{\varepsilon}\right)^4 \log\left(\frac{\mathfrak{m}}{\varepsilon}\right)\right)$$

PROOF. Note first that the preprocessing cost associated to steps 1 through 4 is $O(n^2)$.

To compute the cost of Sinkhorn iterations, observe that the parameters η and h are chosen in such a way that each term on the right-hand side of (22) are smaller than $\varepsilon/2$:

$$\left(16 + \frac{8}{\gamma^*} \mathfrak{m}\right) \lambda^* h, (4\gamma^* + 2\mathfrak{m}) \eta \leq \frac{\varepsilon}{2}$$

Hence,

$$\tilde{\mu}_{\min} \geq \frac{h\lambda^*}{\gamma^*} = \frac{\varepsilon}{48\mathfrak{m}}, \quad \text{and} \quad \left(\frac{\gamma^*}{h\lambda^*}\right)^3 = \left(\frac{48\mathfrak{m}}{\varepsilon}\right)^3.$$

Together with the prescribed value of η and since $\gamma^* \leq \mathfrak{m}$ because of Proposition 1, we get that the total complexity of Sinkhorn iterations is

$$O\left(\left(\frac{\gamma^*}{h\lambda^*}\right)^3 \frac{\log(1/\tilde{\mu}_{\min})}{\eta}\right) = O\left(\left(\frac{\mathfrak{m}}{\varepsilon}\right)^4 \log\left(\frac{\mathfrak{m}}{\varepsilon}\right)\right)$$

Finally, the last step has a total cost of $O(n+\varepsilon)$ which is negligible with respect to the combination of previous steps. \square

Acknowledgments. Vianney Perchet acknowledges support from the French National Research Agency (ANR) under grant number (ANR-19-CE23-0026 as well as the support grant, as well as from the grant “Investissements d’Avenir” (LabEx Ecodec/ANR-11-LABX-0047). Philippe Rigollet is supported by NSF grants IIS-1838071, DMS-2022448, and CCF-2106377.

REFERENCES

- [ABA20] Jason M Altschuler and Enric Boix-Adsera. Polynomial-time algorithms for multimarginal optimal transport problems with structure. *arXiv preprint arXiv:2008.03006*, 2020.
- [ABA21] Jason Altschuler and Enric Boix-Adsera. Hardness results for multimarginal optimal transport problems. *Discrete Optim. (to appear)*, 2021.
- [AC11] Martial Aguech and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.
- [ADH⁺19] AmirMahdi AhmadiNejad, Sina Dehghani, MohammadTaghi Hajiaghayi, Brendan Lucier, Hamid Mahini, and Saeed Seddighin. From duels to battlefields: Computing equilibria of blotto and other games. *Mathematics of Operations Research*, 44(4):1304–1325, 2019.
- [AWR17] Jason Altschuler, Jonathan Weed, and Philippe Rigollet. Near-linear time approximation algorithms for optimal transport via sinkhorn iteration. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4–9 December 2017, Long Beach, CA, USA*, pages 1961–1971, 2017.
- [BC80] Robert M. Bell and Thomas M. Cover. Competitive optimality of logarithmic investment. *Mathematics of Operations Research*, 5(2):161–166, 1980.
- [Bea22] Daniel Beaglehole. An efficient approximation algorithm for the Colonel Blotto game. *arXiv preprint arXiv:2201.10758*, 2022.
- [Bor21] Emile Borel. La théorie du jeu et les équations intégrales à noyau symétrique. *C. R. Acad. Sci., Paris*, 173:1304–1308, 1921.
- [Bub15] Sébastien Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3–4):231–357, 2015.
- [BV38] Emile Borel and Jean Ville. *Applications de la théorie des probabilités aux jeux de hasard: cours professé à la Faculté des sciences de Paris*. Gauthier-Villars, 1938.
- [Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 2292–2300. Curran Associates, Inc., 2013.
- [DMGN17] Simone Di Marino, Augusto Gerolin, and Luca Nenna. Optimal transportation theory with repulsive costs. *Topological Optimization and Optimal Transport: In the Applied Sciences*, 17:204–256, 2017.
- [Fré53] Maurice Fréchet. Emile Borel, initiator of the theory of psychological games and its application. *Econometrica*, 21:95–96, 1953.
- [FSM21] Aidin Ferdowsi, Walid Saad, and Narayan B. Mandayam. Colonel Blotto game for sensor protection in inter-dependent critical infrastructure. *IEEE Internet of Things Journal*, 8(4):2857–2874, 2021.
- [Gro50] Oliver Gross. The symmetric Blotto game. Technical report, Rand Corporation, 1950.
- [GW50] Oliver Gross and Robert Wagner. A continuous Colonel Blotto game. Technical report, Rand Corporation, 1950.
- [Har08] Sergiu Hart. Discrete Colonel Blotto and general Lotto games. *International Journal of Game Theory*, 36(3):441–460, 2008.

[HM17] Mohammad Hajimirsadeghi and Narayan B. Mandayam. A dynamic Colonel Blotto game model for spectrum sharing in wireless networks. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 287–294, 2017.

[HVL12] Rafael Hortala-Vallve and Aniol Llorente-Saguer. Pure strategy Nash equilibria in non-zero sum Colonel Blotto games. *International Journal of Game Theory*, 41(2):331–343, 2012.

[KR21] Dan Kovenock and Brian Roberson. Generalizations of the general Lotto and Colonel Blotto games. *Economic Theory*, 71(3):997–1032, 2021.

[KS06] Martin Knott and Cyril Smith. Choosing joint distributions so that the variance of the sum is small. *J. Multivar. Anal.*, 97(8):1757–1765, sep 2006.

[Las02] Jean-François Laslier. How two-party competition treats minorities. *Review of Economic Design*, 7(3):297, 2002.

[LHCJ20] Tianyi Lin, Nhat Ho, Marco Cuturi, and Michael I. Jordan. On the complexity of approximating multimarginal optimal transport, 2020.

[LHSR15] Mina Labib, Sean Ha, Walid Saad, and Jeffrey H. Reed. A Colonel Blotto game for anti-jamming in the internet of things. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2015.

[LP02] Jean-François Laslier and Nathalie Picard. Distributive politics and electoral competition. *Journal of Economic Theory*, 103(1):106–130, 2002.

[MM15] Scott T. Macdonell and Nick Mastronardi. Waging simple wars: a complete characterization of two-battlefield Blotto equilibria. *Economic Theory*, 58(1):183–216, 2015.

[MMT05] Jennifer Merolla, Michael Munger, and Michael Tofias. In play: A commentary on strategies in the 2004 u.s. presidential election. *Public Choice*, 123(1/2):19–37, 2005.

[MS15] Antonia Maria Masucci and Alonso Silva. Defensive resource allocation in social networks. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 2927–2932. IEEE, 2015.

[Mye93] Roger B. Myerson. Incentives to cultivate favored minorities under alternative electoral systems. *The American Political Science Review*, 87(4):856–869, 1993.

[Nak06] Mikio Nakayama. The dawn of modern theory of games. In *Advances in mathematical economics*. Vol. 9, pages 73–97. Tokyo: Springer, 2006.

[PW22] Yury Polyanskiy and Yihong Wu. *Information Theory Information Theory: From Coding to Learning*. Cambridge University Press, Cambridge, 2022.

[Ren99] Philip J. Reny. On the existence of pure and mixed strategy nash equilibria in discontinuous games. *Econometrica*, 67(5):1029–1056, 1999.

[Rob06] Brian Roberson. The Colonel Blotto game. *Economic Theory*, 29(1):1–24, 2006.

[RU02] Ludger Rüschenhoff and Ludger Uckelmann. Variance minimization and random variables with constant sum. In *Distributions with given marginals and statistical modelling*, pages 211–222. Springer, 2002.

[San15] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015.

[Sin64] Richard Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The annals of mathematical statistics*, 35(2):876–879, 1964.

[SK67] Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967.

[SLS14] Galina Schwartz, Patrick Loiseau, and Shankar S. Sastry. The heterogeneous Colonel Blotto game. In *2014 7th International Conference on NETwork Games, COntrol and OPTimization (NetGCoop)*, pages 232–238, 2014.

[Tho18] Caroline Thomas. N-dimensional Blotto game with heterogeneous battlefield values. *Economic Theory*, 65(3):509–544, 2018.

[Tsy09] Alexandre B. Tsybakov. *Introduction to nonparametric estimation*. Springer Series in Statistics. Springer, New York, 2009.

[von28] John von Neumann. Zur Theorie der Gesellschaftsspiele. *Math. Ann.*, 100:295–320, 1928.

[Wei12] Jonathan Weinstein. Two notes on the Blotto game. *The BE Journal of Theoretical Economics*, 12(1), 2012.

[WW16] Bin Wang and Ruodu Wang. Joint mixability. *Mathematics of Operations Research*, 41(3):808–826, 2016.

[Zim20] Alexander P. Zimin. On existence of measure with given marginals supported on a hyperplane, 2020.

A THE DISCRETE CASE

During the construction of an approximated solution of the classical Blotto game, we had to resort to some discretization, and we implicitly proved that some discrete random variables were jointly mixable. Quite unfortunately, this result can not be directly generalized to solve the discrete, pure count Blotto game (where Lotto solutions were computed explicitly [Har08]).

We recall that in the discrete pure count Blotto problem, the budget of each player T_A and T_B are non-negative integers and that the amount allocated by players to battlefields, $x_{A,i}$ and $x_{B,j}$, are also non-negative integers. Finally, explicit solutions of the Lotto problem are only available in the pure count problem, when $v_{A,i} = v_{B,i} = 1$, hence we shall focus on this case. In discrete Lotto/Blotto, the probability of equal forces $x_{A,i} = x_{B,i}$ is positive and if this happens, we will still assume that both players wins the battlefield with probability 1/2.

[Har08] described optimal strategies in the associated Lotto game with the following additional notations. The average budget per battlefield are denoted by $a = T_A/n$ and $b = T_B/n$, the uniform distribution on even integers between 0 and $2m$ is denoted by $U_e^m = \text{Unif}\{0, 2, \dots, 2m\}$ while $U_o^m = \text{Unif}\{1, 3, \dots, 2m-1\}$ is the uniform distribution on even integers between 1 and $2m-1$. Those strategies are described in [Har08, Fig. 1].

He also introduced the term of a “feasible” distribution, to indicate that n random variable of that distribution are jointly mixable. He then proved the following

PROPOSITION 14. • If $T_A = mn$ then U_o^m is “feasible” if and only if T_A and n have the same parity
• If $T_A = mn$ then U_e^m is “feasible” if and only if X_A is even
• If $T_A = mn + r$ with $1 \leq r \leq n-1$ then $(1 - \frac{r}{n})U_e^m + \frac{r}{n}U_o^{m+1}$ is “feasible”

As a consequence, [Har08] characterized Nash equilibrium of the discrete Blotto problem in the following three cases:

- (1) If $T_A = T_B$ (because of the first and third rows of [Har08, Fig. 1])
- (2) If $mn < T_B < T_A < (m+1)n$ for some $m \in \mathbb{N}$ (because of the third row of [Har08, Fig. 1])
- (3) If $mn = T_B < T_A < (m+1)n$ for some $m \in \mathbb{N}$, if X_B is even (because of the fourth row of [Har08, Fig. 1])

Now, let us state the following Proposition 15 that will imply the above 3 points. The proof, rather technical (yet algorithmic) is postponed.

PROPOSITION 15. *Discrete random variables $\text{Unif}\{0, \dots, \ell_i\}$ are jointly mixable if and only if their continuous counterparts $\text{Unif}[0, \ell_i]$ are jointly mixable and $\sum_i \ell_i$ is even.*

This proposition allows us to describe solutions of the discrete Blotto game.

THEOREM 16. *Assume X_B is even and X_A has the same parity than n , then the optimal strategies of the discrete Blotto game are given by [Har08, Fig. 1]. Moreover those strategies can be computed using the same algorithmic approach as in Proposition 15.*

PROOF. We will prove the theorem first when $T_A = mn$ for some $m \in \mathbb{N}$ and then when $T_B < mn < T_A$, again for some $m \in \mathbb{N}$. The other cases are already covered [Har08].

1. If $T_B < mn = T_A$, then the optimal strategy of Player A is U_o^m , which is feasible if T_A and n have the same parity. Indeed, since $U_o^m = \text{Unif}\{1, \dots, 2m-1\} = 1 + 2\text{Unif}\{0, \dots, m-1\}$, the problem reduces to mixing n random variables of law $\text{Unif}\{0, \dots, m-1\}$; this requires that $n(m-1) = T_A - n$ is even.

Player B marginals are $(1 - \frac{b}{m})\delta_0 + \frac{b}{m}U_{o/e}^m$. So it remains to prove that those random variables are jointly mixable. In particular, this holds if $(1 - \frac{b}{m})\delta_0 + \frac{b}{m}\text{Unif}\{0, \dots, 2m\}$ are jointly mixable by choosing appropriate weights on U_o^m and U_e^m .

In the continuous case (i.e., if uniform distributions are over $[0, m]$ instead of $\{0, m\}$), [Rob06] constructed explicit couplings between such random variables by reducing to coupling of uniform continuous random variables. We apply the exact same techniques, yet we just need to ensure that the intervals of the continuous uniform distributions start and end on integers

and that the sums of lengths are always even. But this immediately happens as soon as X_A/n is an integer.

Indeed, all the couplings introduced in [Rob06] involves uniform random variables over intervals $[\alpha_i, \beta_i]$ where $\alpha_i, \beta_i \in \{0, 2m, X_B - 2m, bn - 4m, \dots\}$. As a consequence, all interval lengths are even if $T_B = bn$ is even, which implies that the discrete uniform variables over $\{\alpha_i, \dots, \beta_i\}$ are jointly mixable.

2. In the second case, the strategy of player B is to jointly mix n random variables of distribution $(1 - \frac{b}{m})\delta_0 + \frac{b}{m}\text{Unif}\{0, 2, \dots, 2m\}$ which is equivalent to mixing n distributions $(1 - \frac{b}{m})\delta_0 + \frac{b}{m}\text{Unif}\{0, 1, \dots, m\}$. Using again, exactly as above, the construction based on the continuous couplings of [Rob06], this is possible as soon as T_B is even.

On the other hand, Player A marginals are $(1 - \alpha)U_o^m + \alpha U_o^{m+1}$, thus we need to prove that $(1 - \alpha)\text{Unif}\{0, m-1\} + \alpha\text{Unif}\{0, m\}$ are jointly mixable. To ensure this, one just need to select $n\alpha$ battlefields at random and to allocate $\text{Unif}\{0, m\}$ on them (and $\text{Unif}\{0, m-1\}$ on the $n - n\alpha$ remaining battlefields). As a consequence, we end up in mixing $n\alpha$ uniforms $\text{Unif}\{0, m\}$ and $n - n\alpha$ uniforms $\text{Unif}\{0, m-1\}$, which is possible as soon as $n\alpha m + (n - n\alpha)(m-1) = T_A - n$ is even, i.e., if T_A and n have the same parity.

These two claims give the result. \square

We finally prove Proposition 15

PROOF OF PROPOSITION 15. The proof of the necessary part of the condition is identical to the continuous case. The only difference is the fact that $\sum \ell_i$ must be even. This is naturally implied by the fact that $\sum X_i$ is always an integer, hence $\sum \mathbb{E}X_i = \sum \ell_i/2$ should also be an integer if these random variables are jointly mixable.

It only remains to prove this statement for $n \geq 3$ as the case $n = 2$ is trivial. Moreover, proving the statement for $n > 3$ can be be reduced to the case $n = 3$ with a simple induction over n , as in the continuous case [WW16]. Indeed, since $\text{Unif}\{0, \dots, \ell_1\}$, $\text{Unif}\{0, \dots, \ell_2\}$ and $\text{Unif}\{0, \dots, \ell_1 + \ell_2\}$ are jointly mixable (because they satisfy the condition of the Proposition for $n = 3$), it is possible to reduce the joint mixability of n uniform to only 3. As a consequence, we will solely focus on $n = 3$ and and we assume wlog that $\ell_1 \leq \ell_2 \leq \ell_3$

The proof will be based on another induction on the maximal size ℓ_3 . We will distinguish three cases, depending on whether $\ell_1 = \ell_2 = \ell_3$, or $\ell_1 < \ell_2 = \ell_3$, or $\ell_2 < \ell_3$.

First case: $\ell_1 = \ell_2 = \ell_3$ (in particular, this implies that ℓ_1, ℓ_2, ℓ_3 are even since $\ell_1 + \ell_2 + \ell_3 = 3\ell_1$ is even by assumption).

If $\ell_1 = 2$, then the uniform coupling on the triplets $(0, 1, 2)$, $(1, 2, 0)$ and $(2, 0, 1)$ proves joint mixability. If $\ell_1 = 4$, then considering the uniform distribution on the following set

$$\left\{ (3, 3, 0); (3, 0, 3); (0, 3, 3); (4, 1, 1); (1, 4, 1); (1, 1, 4); (0, 2, 4); (2, 4, 0); (4, 0, 2); (2, 2, 2) \right\}$$

is sufficient to prove joint mixability. From now on, we shall assume that $\ell_1 = \ell_2 = \ell_3 \geq 6$.

Consider the coupling defined by

$$\begin{aligned} X_1^{1,-} &\sim \delta_0, \quad X_2^{1,-} \sim \text{Unif}\{\frac{\ell_1}{2} + 1, \ell_1 - 1\} \quad \text{and} \\ X_3^{1,-} &= \frac{3}{2}\ell_1 - X_2^{1,-} \sim \text{Unif}\{\frac{\ell_1}{2} + 1, \ell_1 - 1\}. \end{aligned}$$

Then it immediately follows that $X_1^{1,-} + X_2^{1,-} + X_3^{1,-} = \frac{3}{2}\ell_1$. Similarly, this property holds with the following alternative coupling

$$\begin{aligned} X_1^{1,+} &\sim \delta_{\ell_1}, \quad X_2^{1,+} \sim \text{Unif}\{1, \frac{\ell_1}{2} - 1\} \quad \text{and} \\ X_3^{1,+} &= \frac{1}{2}\ell_1 - X_2^{1,+} \sim \text{Unif}\{1, \frac{\ell_1}{2} - 1\}. \end{aligned}$$

We define similarly the coupling $(X_i^{2,\pm})_{i \in [3]}$ and $(X_i^{3,\pm})_{i \in [3]}$ where the role of X_1 is exchanged with X_2 and, respectively, X_3 . We also define the last coupling $X_i^0 = \delta_{\frac{\ell_i}{2}}$.

Then X_i can be decomposed as follows:

$$X_i = \frac{1}{\ell_1 + 1} \left\{ \sum_{k \in [3]} X_i^{k,-} + \sum_{k \in [3]} X_i^{k,+} + \frac{2}{\frac{\ell_i}{2} - 1} X_i^0 \right\} + \frac{\ell_1 - 5}{\ell_1 + 1} Y_i,$$

where $Y_i \sim \text{Unif}\{1, \dots, \ell_1\}$. A simple induction gives the joint mixability of $\{X_1, X_2, X_3\}$

Second case: $\ell_1 < \ell_2 = \ell_3$ (in particular, this implies that ℓ_1 is even). There are two specifics cases $\ell_1 = 2$ and $\ell_2 \in \{3, 4\}$ that are constructed explicitly as follows.

If $\ell_1 = 2$ and $\ell_2 = \ell_3 = 3$, a joint mixability coupling is

$$\frac{1}{6} \left(\delta_{(0,3,1)} + \delta_{(0,1,3)} + \delta_{(2,2,0)} + \delta_{(2,0,2)} \right) + \frac{1}{12} \left(\delta_{(1,3,0)} + \delta_{(1,0,3)} + \delta_{(1,1,2)} + \delta_{(1,2,1)} \right)$$

while if $\ell_1 = 2$ and $\ell_2 = \ell_3 = 4$, a valid coupling is

$$\begin{aligned} &\frac{2}{15} \left(\delta_{(2,3,0)} + \delta_{(2,0,3)} + \delta_{(0,4,1)} + \delta_{(0,1,4)} + \delta_{(1,2,2)} \right) \\ &+ \frac{1}{15} \left(\delta_{(1,0,4)} + \delta_{(1,4,0)} + \delta_{(0,3,2)} + \delta_{(2,2,1)} + \delta_{(1,1,3)} \right) \end{aligned}$$

For the other cases, we consider similar couplings as above, i.e.,

$$\begin{aligned} X_3^{3,-} &\sim \delta_0, \quad X_2^{3,-} \sim \text{Unif}\{\ell_2 - \frac{\ell_1}{2}, \ell_2 - 1\} \quad \text{and} \\ X_1^{3,-} &= \ell_2 + \frac{\ell_1}{2} - X_2^{3,-} \sim \text{Unif}\{\frac{\ell_1}{2} + 1, \ell_1\} \end{aligned}$$

and also

$$\begin{aligned} X_3^{3,+} &\sim \delta_{\ell_3}, \quad X_2^{3,+} \sim \text{Unif}\{1, \frac{\ell_1}{2}\} \quad \text{and} \\ X_1^{3,+} &= \frac{\ell_1}{2} - X_2^{3,+} \sim \text{Unif}\{0, \frac{\ell_1}{2} - 1\}. \end{aligned}$$

We define $X_i^{2,\pm}$ similarly.

If $\ell_1 \geq 4$, we introduce the following random variables (the case $\ell_1 = 2$ is detailed just below)

$$\begin{aligned} Y_1^1 &\sim \delta_{\frac{\ell_1}{2}}, Y_1^2 \sim \text{Unif}\{0, \ell_1\}, Y_1^3 \sim \text{Unif}\{0, \ell_1\} \\ \text{and } Y_1^4 &\sim \text{Unif}\{0, \ell_1\} \end{aligned}$$

and similarly

$$\begin{aligned} Y_2^1 &\sim \text{Unif}\{1, \ell_2 - 1\}, Y_2^2 \sim \text{Unif}\{\frac{\ell_1}{2} + 1, \ell_2 - \frac{\ell_1}{2} - 1\}, Y_2^3 \sim \text{Unif}\{1, \ell_2 - 1\} \\ \text{and } Y_2^4 &\sim \text{Unif}\{1, \ell_2 - 1\} \end{aligned}$$

and

$$Y_3^1 \sim \text{Unif}\{1, \ell_3 - 1\}, Y_3^2 \sim \text{Unif}\{1, \ell_3 - 1\}, Y_3^3 \sim \text{Unif}\{\frac{\ell_1}{2} + 1, \ell_2 - \frac{\ell_1}{2} - 1\}$$

$$\text{and } Y_3^4 \sim \text{Unif}\{1, \ell_3 - 1\}$$

So that we can decompose

$$X_i = \frac{1}{\ell_3 + 1} (X_i^{3,-} + X_i^{3,+} + X_i^{2,-} + X_i^{2,+}) + p_1 Y_i^1 + p_2 Y_i^2 + p_3 Y_i^3 + p_4^i$$

where the probability are defined by

$$p_1 = \frac{4}{\ell_3 + 1} \frac{1}{\ell_1}, p_3 = p_2 = \frac{2}{\ell_3 + 1} \frac{\ell_2 - \ell_1 - 1}{\ell_1} \text{ and } p_4 = 1 - \frac{4}{\ell_3 + 1} - p_1 - p_2 - p_3$$

Those couplings are well defined and satisfy the theorem length condition, hence we will be able to proceed by induction.

It remains to consider the case where $\ell_1 = 2$ (plugging $\ell_1 = 2$ in the above construction would give $p_4 < 0$ which is obviously impossible).

We define the variables

$$Z_1^1 \sim \delta_1 \text{ and } Z_1^2 \sim \text{Unif}\{0, 2\}$$

and similarly

$$Z_2^1 \sim \text{Unif}\{2, \ell_2 - 2\}, \text{ and } Z_2^2 \sim \text{Unif}\{2, \ell_2 - 2\}$$

and

$$Z_3^1 \sim \text{Unif}\{2, \ell_2 - 2\} \text{ and } Z_3^2 \sim \text{Unif}\{2, \ell_3 - 2\}$$

So that we can decompose

$$X_i = \frac{1}{\ell_3 + 1} (X_i^{3,-} + X_i^{3,+} + X_i^{2,-} + X_i^{2,+}) + q_1 Z_i^1 + q_2 Z_i^2$$

where the probability are defined by

$$q_1 = \frac{2}{\ell_3 + 1} \text{ and } q_2 = 1 - \frac{4}{\ell_3 + 1} - q_1 = 1 - \frac{6}{\ell_3 + 1},$$

and the result also follows by induction.

Third case: $\ell_1 < \ell_2 < \ell_3$. We are going to proceed by induction (on the maximal length) as before, and we consider the following couplings

$$X_3^- = \delta_0, \quad X_2^- \sim \text{Unif}\{\frac{\ell_3 + \ell_2 - \ell_1}{2}, \ell_2\}$$

$$\text{and } X_1^- = \frac{\ell_3 + \ell_2 + \ell_1}{2} - X_2^- = \text{Unif}\{0, \frac{\ell_1 + \ell_2 - \ell_3}{2}\}$$

and the similar one

$$X_3^- = \delta_{\ell_3}, \quad X_2^- \sim \text{Unif}\{0, \frac{\ell_1 + \ell_2 - \ell_3}{2}\}$$

$$\text{and } X_1^- = \frac{\ell_3 + \ell_2 + \ell_1}{2} - X_2^- = \text{Unif}\{\frac{\ell_1 + \ell_3 - \ell_2}{2}, \ell_1\}$$

It remains to introduce the following random variables

$$Y_1^1 \sim \text{Unif}\{\frac{\ell_1 + \ell_2 - \ell_3}{2} + 1, \frac{\ell_1 + \ell_3 - \ell_2}{2} - 1\}, \quad Y_1^2 \sim \text{Unif}\{0, \ell_1\}$$

$$\text{and } Y_1^3 \sim \text{Unif}\{0, \ell_1\}$$

and similarly

$$Y_2^1 \sim \text{Unif}\{0, \ell_2\}, \quad Y_2^2 \sim \text{Unif}\left\{\frac{\ell_1 + \ell_2 - \ell_3}{2} + 1, \frac{\ell_2 + \ell_3 - \ell_1}{2} - 1\right\}$$

and $Y_2^3 \sim \text{Unif}\{0, \ell_2\}$

and

$$Y_3^1 \sim \text{Unif}\{1, \ell_3 - 1\}, \quad Y_3^2 \sim \text{Unif}\{1, \ell_3 - 1\}$$

and $Y_3^3 \sim \text{Unif}\{1, \ell_3 - 1\}$.

So that we can decompose

$$X_i = \frac{1}{\ell_3 + 1} (X_i^- + X_i^+) + p_1 Y_i^1 + p_2 Y_i^2 + p_3 Y_i^3$$

where the probability are defined by

$$p_1 = \frac{1}{\ell_3 + 1} \frac{\ell_3 - \ell_2 - 1}{\frac{\ell_1 + \ell_2 + \ell_3}{2} + 1}, \quad p_2 = \frac{1}{\ell_3 + 1} \frac{\ell_3 - \ell_1 - 1}{\frac{\ell_1 + \ell_2 + \ell_3}{2} + 1}$$

and $p_3 = 1 - \frac{2}{\ell_3 + 1} - p_1 - p_2$

The proof relies on a simple induction by noticing that $\{Y_1^1, Y_2^1, Y_3^1\}$ satisfies the theorem condition (as soon as $p_1 > 0$) since these three random variables are uniform over intervals of respective lengths $\ell_3 - \ell_2 - 2$, ℓ_2 and $\ell_3 - 2$ (in particular, ℓ_2 is the maximum of these 3 quantities, necessarily $\ell_2 = \ell_3 - 1$ and $p_1 = 0$). Similarly $\{Y_1^2, Y_2^2, Y_3^2\}$ and $\{Y_1^3, Y_2^3, Y_3^3\}$ satisfy the theorem condition as well. \square

B OMITTED PROOFS AND ALGORITHMS

B.1 Algorithm LOTTO2BLOTO

We provide the simple pseudo-code of the main algorithm; it is decomposed into several procedures described in subsequent sections.

ALGORITHM 1: LOTTO2BLOTO

Data: Length & weight vectors $b \in \mathbb{R}_+^n$ and $p \in [0, 1]^n$, budget T , approximation levels η, h ;

Result: Allocation vector $X \in \mathbb{R}_+^n$;

$\{(p^{(j)}, q_j)\}_j \leftarrow \text{DECOMP}(b, T, p)$ Sample $j^* = j$ with probability q_j ; $(I_0, I_1, I_2, I_3, I_4) \leftarrow \text{REDUC}(b, p^{(j^*)})$ for $i \in [4]$ do $b_i^* \leftarrow \sum_{j \in I_i} b_j; p_j^* \leftarrow \sum_{j \in I_i} p_j$; end $(\mu_1, \mu_2, \mu_3, \mu_4) \leftarrow \text{DISCRETIZE}((b_1^*, b_2^*, b_3^*, b_4^*), p_4^*, h)$ $(\xi_1, \xi_2, \xi_3, \xi_4) \leftarrow \text{SINKHORN}((\mu_1, \mu_2, \mu_3, \mu_4), \eta)$ $X \leftarrow \text{SAMPLE}((\xi_1, \xi_2, \xi_3, \xi_4), T, b, (I_0, I_1, I_2, I_3, I_4), h)$	▷ Couplings with one strict mixture ▷ Reduction to 4 random variables ▷ Appropriate Discretization ▷ Numerical computations ▷ Reconstruction and sampling ▷ A valid allocation
--	---

Output: X

B.2 Proof of Proposition 1 and associated Algorithm LOTTO

ALGORITHM 2: LOTTO

Data: Battlefield values $v_{A,i}$ and $v_{B,i}$, budgets T_A and T_B ;

Result: Vector $b = (b_i)_{i \in [n]}$ and probability vector $p \in [0, 1]^n$;

$\gamma \leftarrow 0, \lambda \leftarrow 0, k \leftarrow 0$

▷ Initialization

Sort $\left\{ \frac{v_{A,i}}{v_{B,i}}; i \in [n] \right\}$

▷ Sort to get degree 3 polynomial equation to solve

while $\gamma = 0$ **do**

$S_k \leftarrow \text{roots}\left(\gamma T_A \sum_{i=1}^k v_{A,i} + \gamma^3 T_A \sum_{i=k+1}^n \frac{v_{B,i}^2}{v_{A,i}} - T_B \sum_{i=1}^k \frac{v_{A,i}^2}{v_{B,i}} - T_B \gamma^2 \sum_{i=k+1}^n v_{B,i}\right);$

for $g \in S_k$ **do**

$\gamma \leftarrow g \mathbb{1}\{\gamma \in [\frac{v_{A,k}}{v_{B,k}}, \frac{v_{A,k+1}}{v_{B,k+1}}]\}$ ▷ Keep valid root of degree 3 polynomial

end

$k \leftarrow k + 1$ ▷ Move to next interval between two ratios $\frac{v_{A,k}}{v_{B,k}}$

end

$\lambda \leftarrow \frac{1}{T_A} \frac{1}{2} \sum_{i=1}^n (\gamma v_{B,i}) \wedge \frac{(v_{A,i})^2}{\gamma v_{B,i}};$

for $i \in [n]$ **do**

$b_i \leftarrow \frac{\gamma v_{B,i}}{\lambda} \wedge \frac{v_{A,i}}{\lambda}, p_i \leftarrow \frac{v_{A,i}}{\gamma v_{B,i}} \wedge 1$ ▷ Description of Lotto strategies

end

Output: (b, p) ▷ lengths of uniforms and associated weights

PROOF. First, reorder the battlefields by increasing reward ratios $v_{A,i}/v_{B,i}$ so we can assume that

$$\frac{v_{A,1}}{v_{B,1}} \leq \frac{v_{A,2}}{v_{B,2}} \leq \dots \leq \frac{v_{A,n}}{v_{B,n}}$$

Recall that $f(\gamma)$ is the left-hand side of (6). We first observe that f is continuous on \mathbb{R} . Indeed, f is obviously continuous on each open interval

$$\left(\frac{v_{A,i}}{v_{B,i}}, \frac{v_{A,i+1}}{v_{B,i+1}} \right)$$

To check that it is continuous at $\gamma_{i_0} := v_{A,i_0}/v_{B,i_0}$, note that for γ in a small enough neighborhood of γ_{i_0} , we have

$$\mathcal{N}(\gamma) = \begin{cases} \{i_0 + 1, \dots, n\} & \text{if } \gamma > \gamma_{i_0} \\ \{i_0, \dots, n\} & \text{if } \gamma \leq \gamma_{i_0}. \end{cases}$$

Hence f is left-continuous at γ_{i_0} . Moreover, we can check right-continuity by observing that

$$\lim_{\gamma \rightarrow \gamma_{i_0}^+} f(\gamma) = \gamma_{i_0}^3 \left(T_A \sum_{i=i_0+1}^n \frac{v_{B,i}^2}{v_{A,i}} \right) - \gamma_{i_0}^2 T_B \sum_{i=i_0+1}^n v_{B,i} + \gamma_{i_0} T_A \sum_{i=1}^{i_0} v_{A,i} - T_B \sum_{i=1}^{i_0} \frac{v_{A,i}^2}{v_{B,i}} = f(\gamma_{i_0})$$

where the last identity can be readily checked by substitution.

Having proved that f is continuous, note that for γ large enough we have that $\mathcal{N}(\gamma) = \emptyset$ so that

$$f(\gamma) = \gamma T_A \sum_{i=1}^n v_{A,i} - T_B \sum_{i=1}^n \frac{v_{A,i}^2}{v_{B,i}} \xrightarrow{\gamma \rightarrow \infty} +\infty.$$

Moreover, for $\gamma > 0$ small enough, we have $\mathcal{N}(\gamma) = [n]$ so that

$$\frac{f(\gamma)}{\gamma^2} = \gamma \left(T_A \sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \right) - T_B \sum_{i=1}^n v_{B,i} \xrightarrow{\gamma \rightarrow 0} -T_B \sum_{i=1}^n v_{B,i} < 0.$$

Hence, by the intermediate value theorem, there exists $\gamma^* > 0$ such that $f(\gamma^*) = 0$. Moreover, observe that if one sets $\mathcal{N}(\gamma) = [N]$ for $N \in \{0, \dots, n\}$, with the convention that $[0] = \emptyset$, Equation (6) becomes a polynomial equation of degree three with at most three solutions denoted $\gamma_{N,i}$, $i = 1, 2, 3$. Hence, $\gamma^* \in \bigcup_{N=0}^n \{\gamma_{N,1}, \gamma_{N,2}, \gamma_{N,3}\}$ can take at most $3n + 3$ values.

We now check the bounds on the possible values of γ^* . To that end, recall from (6) that

$$f(\gamma) = \sum_{i=1}^n \left(\gamma T_A - T_B \frac{v_{A,i}}{v_{B,i}} \right) \cdot \left(\gamma^2 \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} \right).$$

To prove the an upper bound on γ^* , observe that

$$\gamma T_A - T_B \frac{v_{A,i}}{v_{B,i}} \geq 0 \iff \gamma \geq \frac{T_B}{T_A} \frac{v_{A,i}}{v_{B,i}} \iff \gamma^2 \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} \geq \left(\frac{T_B}{T_A} \right)^2 v_{A,i} \wedge v_{A,i} = \left(\frac{T_B}{T_A} \right)^2 v_{A,i},$$

where the last identity follows from the assumption that $T_B \leq T_A$.

It yields

$$f(\gamma) \geq \left(\frac{T_B}{T_A} \right)^2 \sum_{i=1}^n \left(\gamma T_A - T_B \frac{v_{A,i}}{v_{B,i}} \right) v_{A,i} = \left(\frac{T_B}{T_A} \right)^2 \cdot (\gamma T_A - T_B (1 + \chi^2(v_A \| v_B)))$$

Therefore, if

$$\gamma > \frac{T_B}{T_A} (1 + \chi^2(v_A \| v_B)),$$

then $f(\gamma) > 0$, which yields the desired upper bound on γ^* .

To prove the lower bound on γ^* , we proceed essentially in the same fashion:

$$\gamma T_A - T_B \frac{v_{A,i}}{v_{B,i}} \leq 0 \implies \gamma \leq \frac{T_B}{T_A} \frac{v_{A,i}}{v_{B,i}} \implies \gamma^2 \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} = \gamma^2 \frac{v_{B,i}^2}{v_{A,i}}.$$

It yields

$$f(\gamma) \leq \gamma^2 \sum_{i=1}^n \left(\gamma T_A - T_B \frac{v_{A,i}}{v_{B,i}} \right) \frac{v_{B,i}^2}{v_{A,i}} = \gamma^2 ((1 + \chi^2(v_B \| v_A)) \cdot \gamma T_A - T_B)$$

Therefore, if

$$\gamma < \frac{T_B}{T_A} \frac{1}{1 + \chi^2(v_B \| v_A)},$$

then $f(\gamma) < 0$, which yields the desired lower bound on γ^* .

To complete the proof of the proposition, it remains to observe that the computational complexity is dominated by sorting reward ratios which costs $\mathcal{O}(n \log n)$ operations since finding roots of degree three polynomials for each of $\mathcal{O}(n)$ polynomials costs $\mathcal{O}(1)$ time.

e if $X_B = X_A$ and $\frac{v_{A,1}}{v_{B,1}} = \frac{v_{A,n}}{v_{B,n}}$, which implies that $v_{A,i} = v_{B,i}$ (as they both sum to 1). \square

B.3 Proofs of Proposition 4 and Theorem 5

B.3.1 Proof of Proposition 4.

PROOF. Note first that (8) is necessary. Indeed, denote by $\max \in [k]$ any index such that $b_{\max} = \max_i b_i$ and assume Z_1, \dots, Z_k are coupled so that

$$\sum_{i \in [k]} Z_i = \sum_{i \in [k]} \mathbb{E}[Z_i] = \frac{1}{2} \sum_{i \in [k]} p_i b_i \quad \text{a.s.}$$

We have for $\delta > 0$ small enough,

$$p_{\max} \delta = \mathbb{P}(Z_{\max} > b_{\max}(1 - \delta)) \leq \mathbb{P}\left(\sum_{i \in [k]} Z_i > b_{\max}(1 - \delta)\right) = \mathbb{P}\left(\frac{1}{2} \sum_{i \in [k]} p_i b_i > b_{\max}(1 - \delta)\right).$$

Hence

$$\frac{1}{2} \sum_{i \in [k]} p_i b_i > b_{\max}(1 - \delta)$$

since the above deterministic inequality holds with positive probability, it holds with probability one for all $\delta > 0$. Letting $\delta \rightarrow 0$ yields (8).

To show that (8) is sufficient, recall from [WW16, Theorem 3.2] that a collection of random variables Z_1, \dots, Z_k where Z_i is a continuous random variable, with non-increasing density function and supported on the interval $[0, b_i]$ are jointly mixable if and only if

$$\max_{i \in [k]} b_i \leq \sum_{i \in [k]} \mathbb{E}[Z_i] \leq \sum_{i \in [k]} b_i - \max_{i \in [k]} b_i.$$

In particular, if $\mathbb{E}[Z_i] = p_i b_i / 2$, these two inequalities reduce to (8).

Unfortunately the Z_i s do not have a density so we use the following approximation. For $i \in [k]$, let $\varepsilon < b_i$ and

$$Z_i^\varepsilon \sim (1 - q_i^\varepsilon) \text{Unif}[0, \varepsilon] + q_i^\varepsilon \text{Unif}[0, b_i],$$

where $q_i^\varepsilon \in (0, 1)$ is chosen precisely so that Z_i^ε has the same expectation and the same support as Z_i . Moreover, Z_i^ε has a monotone decreasing density and hence the Z_i^ε are jointly mixable under condition (8) resulting in a coupling π^ε over the product space $\prod_{i \in [k]} [0, b_i]$. By Prokhorov's theorem, letting $\varepsilon \rightarrow 0$ implies that Z_1, \dots, Z_k are jointly mixable.

□

B.3.2 Proof of Theorem 5.

PROOF. Recall that the marginal strategy Z_i of player A on battlefield i has a distribution of the form (7) with

$$\begin{aligned} p_i &= 1, & b_i &= \frac{\gamma^* v_{B,i}}{\lambda^*}, & \text{if } i \in \mathcal{N}(\gamma^*) \\ p_i &= \frac{v_{A,i}}{\gamma^* v_{B,i}}, & b_i &= \frac{v_{A,i}}{\lambda^*}, & \text{if } i \notin \mathcal{N}(\gamma^*) \end{aligned}$$

Plugging this values into (8) yields

$$\max_{i \in \mathcal{N}(\gamma^*)} \frac{\gamma^* v_{B,i}}{\lambda^*} \vee \max_{i \notin \mathcal{N}(\gamma^*)} \frac{v_{A,i}}{\lambda^*} \leq \frac{1}{2} \left(\sum_{i \in \mathcal{N}(\gamma^*)} \frac{\gamma^* v_{B,i}}{\lambda^*} + \sum_{i \notin \mathcal{N}(\gamma^*)} \frac{v_{A,i}^2}{\lambda^* \gamma^* v_{B,i}} \right) = T_A,$$

where the last equality follows from the saturation of the budget constraint in (4). It is easy to check that

$$\max_{i \in \mathcal{N}(\gamma^*)} \gamma^* v_{B,i} \vee \max_{i \notin \mathcal{N}(\gamma^*)} v_{A,i} = \max_{i \in [n]} (\gamma^* v_{B,i} \wedge v_{A,i})$$

Similarly, we get using (5) that

$$\max_{i \in [n]} (\gamma^* v_{B,i} \wedge v_{A,i}) \leq \lambda^* T_B,$$

is a necessary and sufficient condition for mixability the Lotto strategy of player B into a Blotto strategy. The proof can then be concluded by recalling that $T_A \geq T_B$. \square

B.4 Proof of Corollary 7

PROOF. Using (5), we get

$$\frac{\lambda^* T_B}{\gamma^*} = \frac{1}{2\gamma^*} \sum_{i=1}^n \frac{(\gamma^* v_{B,i})^2}{v_{A,i}} \wedge v_{A,i} \geq \frac{1}{2} \left(\gamma^* \wedge \frac{1}{\gamma^*} \right) \sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i}.$$

Next, observe that

$$\begin{aligned} \sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} &= \sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \mathbb{1}(v_{A,i} > v_{B,i}) + \sum_{i=1}^n v_{A,i} \mathbb{1}(v_{A,i} \leq v_{B,i}) \\ &= \sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \mathbb{1}(v_{A,i} > v_{B,i}) + 1 - \sum_{i=1}^n v_{A,i} \mathbb{1}(v_{A,i} > v_{B,i}) \\ &= 1 - \sum_{i=1}^n v_{A,i} \left(1 - \left(\frac{v_{B,i}}{v_{A,i}} \right)^2 \right) \mathbb{1}(v_{A,i} > v_{B,i}). \end{aligned}$$

Noting now that $1 - x^2 \leq 2 - 2x$ for $x \in [0, 1]$, we get for $x = v_{B,i}/v_{A,i}$ that

$$\sum_{i=1}^n \frac{v_{B,i}^2}{v_{A,i}} \wedge v_{A,i} \geq 1 - \sum_{i=1}^n v_{A,i} \left(1 - \frac{v_{B,i}}{v_{A,i}} \right) \mathbb{1}(v_{A,i} > v_{B,i}) = 1 - \text{TV}(v_A, v_B),$$

where TV denotes the total variation distance and is defined as

$$\text{TV}(v_A, v_B) = \frac{1}{2} \sum_{i=1}^n |v_{A,i} - v_{B,i}|.$$

Using Pinsker's and Jensen's inequalities, see, e.g., [Tsy09, Chapter 2], we get the classical result

$$\text{TV}(v_A, v_B) \leq \frac{1}{2} \sqrt{\chi^2(v_B \| v_A)} \leq \frac{r}{2}$$

Next, note that Proposition 1 yields

$$\gamma^* \wedge \frac{1}{\gamma^*} \geq \frac{T_B}{T_A} \frac{1}{1 + \chi^2(v_B \| v_A)} \wedge \frac{T_A}{T_B} \frac{1}{1 + \chi^2(v_A \| v_B)} \geq \frac{T_B}{T_A} \frac{1}{1 + r^2}.$$

Hence we have established that

$$\frac{\lambda^* T_B}{\gamma^*} \geq \frac{T_B}{2T_A} \frac{1 - r/2}{1 + r^2} \geq \frac{T_B}{2T_A} (1 - r).$$

Together with (9), this completes our proof. \square

ALGORITHM 3: DECOMP

Data: length vector $b \in \mathbb{R}_+^n$, budget T , weights vector $p = (p_i)_{i \in [n]} \in [0, 1]^n$;

Result: (at most) n vectors $p^{(k)} \in [0, 1]^n$ and weights $q_k \in [0, 1]$;

```

 $k \leftarrow 0$  ▷ Initialization
while  $\text{supp}(p) > 2$  do
   $k \leftarrow k + 1$ ;
   $(p^{(k)}, \bar{p}^{(k)}, \theta_k) \leftarrow \text{EXTREMIZE}(b, T, p)$  ▷ Find an extreme point  $p^{(k)}$ 
   $p \leftarrow \bar{p}^{(k)}$ ;
   $q_k \leftarrow q_{k-1} \frac{\theta_k}{\theta_{k-1}} (1 - \theta_{k-1})$  ▷ Compute convex weight of  $p^{(k)}$ 
end
Output:  $\{(p^{(j)}, q_j) ; j \in [k]\}$  ▷ (a coupling with its convex weight)

```

B.5 Proof of Lemma 8 and Algorithm DECOMP

PROOF. Throughout this proof we write

$$p_i = \begin{cases} 1, & \text{if } i \in \mathcal{N}(\gamma^*) \\ \frac{v_{A,i}}{\gamma^* v_{B,i}}, & \text{if } i \notin \mathcal{N}(\gamma^*) \end{cases}$$

so that the i th marginal of π is given by

$$(1 - p_i)\delta_0 + p_i \text{Unif} \left[0, \frac{\gamma^* v_{B,i} \wedge v_{A,i}}{\lambda^*} \right].$$

Note that (10) consists in representing π as a convex combination of couplings. To obtain this representation we are going to appeal to Carathéodory's theorem. However the latter requires finite dimension so we first make the following observation: the map $p^{(k)} = (p_1^{(k)}, \dots, p_n^{(k)}) \mapsto \pi_k$, (resp. $p = (p_1, \dots, p_n) \mapsto \pi$) is linear and injective. Therefore, (10) is sufficient to produce the decomposition

$$p = \sum_k q_k p^{(k)}. \quad (23)$$

Furthermore, we need π_k to be satisfy the budget constraint (3) of the Lotto game. The saturation of this constraint translates into the constraint $p^k \in \mathcal{H}$, $k = 1, \dots, n$, where \mathcal{H} is the affine hyperplane in \mathbb{R}^n defined by

$$\mathcal{H} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i \frac{\gamma^* v_{B,i} \wedge v_{A,i}}{2\lambda^*} = T_A \right\} \quad (24)$$

As a result, we must ensure that $p^{(k)} \in C_n = \mathcal{H} \cap [0, 1]^n$; see Figure 1 for a representation of this constraint set. Since this set has dimension $n - 1$, Carathéodory's theorem ensures the existence of a decomposition (23) where $p^{(k)}$ are extreme points of C_n . In particular, each such extreme point has at most one coordinate in $(0, 1)$; this completes the proof of point 1.

Note that (23) readily ensures that π is a solution for the Lotto game. Moreover, since $p^{(k)} \in \mathcal{H}$ by construction, we have that marginals of π_k automatically satisfy the mixability condition. Hence, we can choose π_k to be a joint mix; this completes the proof of point 2. As stated above, this readily implies that π defined in (10) is a joint mix and hence a solution for the Blotto game; this completes the proof of point 3.

It remains to find an efficient algorithm that outputs decomposition (10). The algorithm is initialized at $p \in C_n = \mathcal{H} \cap [0, 1]^n$. From there, we use Lemma 9 to construct two points $p^{(1)}, \bar{p}^{(1)} \in$

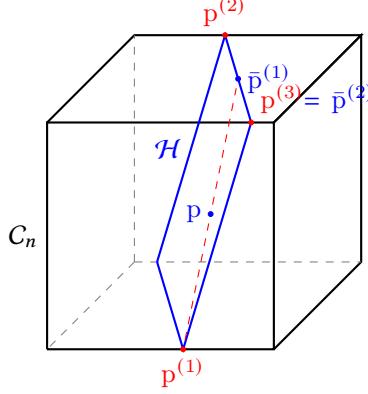


Fig. 1. Starting from p , the first step computes a first extreme point $p^{(1)}$. Then p is written as a convex combination of $p^{(1)}$ and $\bar{p}^{(1)}$. The latter is decomposed (iteratively) into $p^{(2)}$ and then $p^{(3)}$.

$\mathcal{H} \cap \partial[0,1]^n$ such that $p = \theta_1 p^{(1)} + (1 - \theta_1) \bar{p}^{(1)}$, for some $\theta_1 \in [0, 1]$, where $p^{(1)}$ is an extreme point of C_n and $\bar{p}^{(1)}$ belongs to a face \mathcal{F}_{n_1} of dimension $n_1 < n$. We repeat this procedure as follows. Define by $C_{n_1} = \mathcal{H} \cap \mathcal{F}_{n_1}$ so that $\bar{p}^{(1)} \in C_{n_1}$. Hence, using again Lemma 9, it may be decomposed as $\bar{p}^{(1)} = \theta_2 p^{(2)} + (1 - \theta_2) \bar{p}^{(2)}$ for some $\theta_2 \in [0, 1]$ and $p^{(2)}$ is an extreme point of C_n while $\bar{p}^{(2)} \in C_{n_2}$, where C_{n_2} is the intersection of \mathcal{H} with a face \mathcal{F}_{n_2} of $[0, 1]^n$ of dimension $n_2 < n_1$. Moreover,

$$p = \theta_1 p^{(1)} + (1 - \theta_1) \theta_2 p^{(2)} + (1 - \theta_1)(1 - \theta_2) \bar{p}^{(2)}$$

Iterating this procedure yields the decomposition

$$p = \sum_{i=1}^{N+1} \theta_i \prod_{j \leq i-1} (1 - \theta_j) p^{(i)},$$

with the convention that $\theta_{N+1} = 1$ and where $N \leq n$ and the $p^{(j)}$'s are extreme points of C_n . Moreover, one can readily check that for any sequence $\theta_1, \dots, \theta_N \in [0, 1]$, $\theta_{N+1} = 1$, it holds

$$\sum_{i=1}^{N+1} \theta_i \prod_{j \leq i-1} (1 - \theta_j) = 1,$$

which gives the result by defining $q_k = \theta_k \prod_{j \leq k-1} (1 - \theta_j)$.

As we appealed at most n times to Lemma 9, the overall complexity of this algorithm is of order $\mathcal{O}(n^2 \log n)$. \square

B.6 Proof of Lemma 9 and Algorithm DECOMP

PROOF. Assume that the coordinates of ℓ are sorted as $\ell_1 \leq \ell_2 \leq \dots \leq \ell_n$. For any subset $\mathcal{S} \subset [n]$, we denote its associated indicator vector by $\mathbb{1}_{\mathcal{S}} = (\mathbb{1}_{\{i \in \mathcal{S}\}})_i \in \{0, 1\}^n$.

First, notice that there exists some $\varepsilon_k \in [0, 1]$ such that the following vector y belongs to $C_n \cap \mathcal{H}$:

$$y := x \odot \mathbb{1}_{\overline{\text{supp}}(x)} + (\underbrace{1, 1, \dots, 1}_{k-1 \text{ terms}}, \varepsilon_k, 0, \dots, 0) \odot \mathbb{1}_{\text{supp}(x)},$$

ALGORITHM 4: EXTREMIZE**Data:** vector $b = (b_i)_{i \in [n]} \in \mathbb{R}^n$, budget $T \in \mathbb{R}_+$, vector $x \in \mathbb{R}^n$;**Result:** 2 vectors y, \bar{y} , convex weight θ ; $t \leftarrow 0; k \leftarrow 0, z \leftarrow x \odot \mathbb{1}_{\overline{\text{supp}}(x)}$ ▷ Initialization**while** $t < T$ **do** $| \quad k \leftarrow k + 1; z \leftarrow z + \mathbb{1}_{\{k\}} \odot \mathbb{1}_{\text{supp}(x)}; t \leftarrow z^\top b$ ▷ Find the index with non {0, 1} component**end** $\epsilon' \leftarrow \frac{t-T}{b_k}; y \leftarrow z - \epsilon' \mathbb{1}_{\{k\}}$ ▷ Compute explicitly the extreme point y $\delta \leftarrow \min_{i \in S} \frac{\mathbb{1}_{\{x_i \geq y_i\}} - x_i}{x_i - y_i}; \theta \leftarrow \frac{\delta}{1+\delta};$ $\bar{y} \leftarrow x + \eta(x - y);$ **Output:** (y, \bar{y}, θ) ▷ (two vectors, one convex weight)

where $\overline{\text{supp}}(x) = [n] \setminus \text{supp}(x)$ is the complement of the support of x . The reason is simply that

$$x \odot \mathbb{1}_{\overline{\text{supp}}(x)} \leq x \leq x \odot \mathbb{1}_{\overline{\text{supp}}(x)} + (1, \dots, 1) \odot \mathbb{1}_{\text{supp}(x)},$$

where the inequalities are component wise.

We now introduce $\bar{y} = x + \delta(x - y)$, where $\delta \in \mathbb{R}_+$ is defined by

$$\delta = \delta_{i^*} := \min_{i \in \text{supp}(x)} \delta_i, \quad \text{with } \delta_i := \frac{\mathbb{1}_{\{x_i \geq y_i\}} - x_i}{x_i - y_i}.$$

Then \bar{y} belongs to \mathcal{H} as x and y are two vectors of this affine hyperplane. Moreover, as $\eta_{i^*} = \min \eta_i$, it also holds that $\bar{y} \in [0, 1]^n$ and $\bar{y}_{i^*} = \mathbb{1}_{\{x_{i^*} \geq y_{i^*}\}}$, and therefore $\text{supp}(\bar{y}) \subset \text{supp}(x) \setminus \{i^*\}$. Finally, one just needs to define $\theta = \delta/(1 + \delta)$.

The construction of those quantities requires sorting the coordinates of ℓ , finding the k -th coordinate of y , say, by binary search (computing the value of ϵ_k is immediate as the value of $y^\top \ell$ is linear in p_k) and finding i^* . □

B.7 Proof of Lemma 10 and Algorithm REDUC

PROOF. Write for simplicity $X_i = X_{A,i}$ and

$$X_i \sim (1 - p_i)\delta_0 + p_i \text{Unif}[0, b_i],$$

where $p_i \in \{0, 1\}$ for $i \in [n-1]$, $p_n \in [0, 1]$, and $b_i = (\gamma^* v_{B,i} \wedge v_{A,i})/\lambda^*$. If $p_i = 0$, then X_i is almost surely equal to 0 and this random variable can be removed from the analysis; we might therefore assume that $p_i > 0$ for all i .

The joint mixability condition (8) rewrites as

$$2 \max_k b_k \leq \sum_{i=1}^{n-1} b_i + p_n b_n \tag{25}$$

Without loss of generality, assume that $b_1 \leq b_2 \leq \dots \leq b_{n-1}$ and define $X_{12} \sim \text{Unif}[0, b_1 + b_2]$. Note that $\{X_1, X_2, -X_{12}\}$ are jointly mixable (for instance, consider $X_1 = \frac{b_1}{b_1 + b_2} X_{12}$ and $X_2 = \frac{b_2}{b_1 + b_2} X_{12}$), and $\{X_{12}, X_3, \dots, X_n\}$ are also jointly mixable if

$$2 \max_{k \geq 3} \{b_k, b_1 + b_2\} \leq \sum_{i=1}^{n-1} b_i + p_n b_n \tag{26}$$

ALGORITHM 5: REDUC

Data: length vector $b \in \mathbb{R}_+^n$, weights $p \in [0, 1]^n$;

Result: 5 sets of indices;

$I_0 \leftarrow \emptyset, I_1 \leftarrow [n], I_2 \leftarrow \emptyset, I_3 \leftarrow \emptyset, I_4 \leftarrow \emptyset$ ▷ Initialization of sets

for $i \in [n]$ **do**

- if** $p_i = 1$ **then**
 - $I_0 \leftarrow I_0 \cup \{i\}, I_1 \leftarrow I_1 \setminus \{i\}$ ▷ Dirac masses not treated, index removed
- else**
 - if** $p_i \in (0, 1)$ **then**
 - $I_n \leftarrow \{i\}, I_1 \leftarrow I_1 \setminus \{i\}$ ▷ At most one strict mixture
 - else**
 - $I(i) \leftarrow \{i\}$ ▷ End of initialization
 - end**
- end**

end

while $|I_1| > 3$ **do**

- $i_1 \leftarrow \arg \min_{i \in I_1} b_i, I_1 \leftarrow I_1 \setminus \{i_1\}, i_2 \leftarrow \arg \min_{i \in I_1} b_i$ ▷ i_1, i_2 : 2 smallest uniform lengths
- $b_{i_2} \leftarrow b_{i_2} + b_{i_1}, I(i_2) \leftarrow I(i_2) \cup I(i_1)$ ▷ 2 uniforms combined in 1

end

$m = |I_1|$, denote $I_1 = \{i_1^*, \dots, i_m^*\}$;

for $j \in [m]$ **do**

- $I_j \leftarrow I(i_j^*)$ ▷ (At most) 3 sets of combined indices

end

Output: $(I_0, I_1, I_2, I_3, I_4)$ ▷ (Dirac, 3 sets of uniforms - possibly empty-, a strict mixture)

On the one hand, if $b_1 + b_2 \leq b_{n-1}$ then (26) follows readily from (25). On the other hand, if $b_1 + b_2 > b_{n-1}$ and $n - 1 \geq 4$ then

$$2(b_1 + b_2) \leq b_1 + b_2 + b_3 + b_4 \leq b_1 + b_2 + \sum_{k=3}^{n-1} b_k + t^{(n)} b_n$$

so that (25) is satisfied. The result follows from sequentially iterating this construction. Indeed, recall that X_{12} is a variable created by mixing together X_1 and X_2 . At a further step, this variable might be mixed again with another one, ending up with a set of, at least, 3 variables mixed together. At the end of the procedure, there are at most 3 remaining random variables Y_j , that are joint mix of (disjoints) subsets of X_i .

We denote by I_j the set of indices of variables that are mixed to form Y_j and, for any index $i \in I_j$, we just need to define $\theta_i = \frac{b_i}{\sum_{k \in I_j} b_k}$. □

B.8 Algorithm DISCRETIZE

B.9 Algorithm SINKHORN

ALGORITHM 6: DISCRETIZE**Data:** 4 lengths $b_i^* \in \mathbb{R}_+$, positive weight $p_4 \in (0, 1)$, discretization parameter h ;**Result:** 4 vectors $\mu_i \in \mathbb{R}_+^{d_i}$ with $d_i = \lfloor \frac{b_i^*}{h} \rfloor + 1$;**for** $i \in [3]$ **do**

$$\mu_i \leftarrow \frac{h}{b_i^*} \left(\underbrace{1, \dots, 1}_{\lfloor \frac{b_i^*}{h} \rfloor \text{ times}}, \frac{b_i^*}{h} - \lfloor \frac{h}{b_i^*} \rfloor \right)$$

end

$$\mu_4 \leftarrow p_4 \frac{h}{b_4^*} \left(\frac{b_4^*}{h} - \lfloor \frac{h}{b_4^*} \rfloor, \underbrace{1, \dots, 1}_{\lfloor \frac{b_4^*}{h} \rfloor \text{ times}} \right) + (1 - p_4)(0, \dots, 0, 1);$$

Output: $(\mu_1, \mu_2, \mu_3, \mu_4)$ ▷ (4 discrete distributions)**ALGORITHM 7: SINKHORN****Data:** Marginals $\mu_j \in \mathbb{R}^{d_j}$, $j = 1, \dots, 4$;Precision level $\eta > 0$;**Result:** A Coupling Γ , represented by $(\xi_1, \xi_2, \xi_3, \xi_4) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_3} \times \mathbb{R}^{d_4}$, with marginals such that $\sum_{i=1}^4 \|\bar{\Gamma}^{(i)} - \mu_i\|_1 \leq \eta$;

$$\Gamma \leftarrow 1 \in \mathbb{R}^{d_1 \times d_2 \times d_3 \times 3} \quad \text{Initialization}$$

$$\xi_1 \leftarrow 1 \in \mathbb{R}^{d_1}; \xi_2 \leftarrow 1 \in \mathbb{R}^{d_2}; \xi_3 \leftarrow 1 \in \mathbb{R}^{d_3}; \xi_4 \leftarrow 1 \in \mathbb{R}^{d_4};$$

while $\sum_{i \in [4]} \|\mu_i - \Gamma_i\|_1 > \eta$ **do**

$$\tau \leftarrow \underset{i \in [4]}{\operatorname{argmax}} \text{KL}(\mu_i \|\Gamma_i);$$

$$\xi_\tau \leftarrow \xi_\tau \odot \mu_\tau \oslash \bar{\Gamma}_\tau$$

$$\Gamma_{ijke} = \xi_{1,i} \cdot \xi_{2,j} \cdot \xi_{3,k} \cdot \xi_{4,i+j+k+e}, \quad \forall i, j, k, e$$

▷ Matrix scaling

▷ Computation of coupling

end**Output:** $(\xi_1, \xi_2, \xi_3, \xi_4)$ **B.10 Algorithm SAMPLE**

ALGORITHM 8: SAMPLE

Data: Coupling $(\xi_1, \xi_2, \xi_3, \xi_4)$, budget $T \in \mathbb{R}_+$, length vector $b \in \mathbb{R}_+^n$, sets of indices $(\mathcal{I}_0, \mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4)$, precision h ;

Result: An allocation vector $X \in \mathbb{R}_+^n$;

Set $\{j_4\} = \mathcal{I}_4$; $b_4^* \leftarrow b_{j_4}$ ▷ Initialization

Sample $\tilde{Y}_1 \sim \xi_1$, $\tilde{Y}_2 \sim \xi_2$, $\tilde{Y}_3 \sim \xi_3$ ▷ Sampling from computed coupling

Sample $\varepsilon = e \in \{0, 1, 2\}$ with proba. $\xi_{4, \tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 + e}$;

Sample $U \sim \text{Unif}[0, 1]$ ▷ smoothing

for $i \in [3]$ **do**

$b_i^* \leftarrow \sum_{j \in \mathcal{I}_i} b_j$;

$Y'_i \leftarrow (\tilde{Y}_i + \frac{\varepsilon}{3} + \frac{U}{3})h \wedge b_i^*$

end

$S \leftarrow Y'_1 + Y'_2 + Y'_3$;

$\zeta \leftarrow \mathbb{1}\{S > T\} \frac{T-S}{3} + \mathbb{1}\{S < T - b_4^*\} \frac{T-b_4^*-S}{3}$;

for $i \in [3]$ **do**

$\bar{Y}_i \leftarrow Y'_i + \zeta$;

for $j \in \mathcal{I}_i$ **do**

$X_j = \frac{b_j}{b_i^*} \bar{Y}_i$

end

end

for $j \in \mathcal{I}_0$ **do**

$X_j = 0$ ▷ Dirac masses set at 0

end

$X_{j_4} = T - (\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3)$ ▷ Budget saturation

Output: X
