

# Contents

<b>1</b>	<b>Tournament Formats</b>	
1.1	Definitions . . . . .	3
<b>2</b>	<b>Brackets</b>	
2.1	Brackets and Rounds . . . . .	5
2.2	Bracket Signatures . . . . .	8
2.3	Proper Brackets . . . . .	13
2.4	Ordered Brackets . . . . .	20
2.5	Edwards's Theorem . . . . .	26
<b>3</b>	<b>Alternative Seedings</b>	
3.1	Reseeding . . . . .	34
3.2	Tiered Seeding . . . . .	38
3.3	The Palandromic Algorithm . . . . .	44
<b>4</b>	<b>Multibrackets</b>	
4.1	Simple Multibrackets . . . . .	50
4.2	Properness . . . . .	61
4.3	Multiple Elimination . . . . .	62
4.4	Semibrackets . . . . .	63
4.5	Swiss . . . . .	70
<b>5</b>	<b>References</b>	

# 1 Tournament Formats

## 1.1 Definitions

### Definition 1.1.1: Gameplay Function

A *gameplay function*  $g$  on a list of teams  $\mathcal{T}$  is a nondeterministic function  $g : \mathcal{T} \times \mathcal{T} \rightarrow \mathcal{T}$  with the following properties:

- $\mathbb{P}[g(t_1, t_2) = t_1] + \mathbb{P}[g(t_1, t_2) = t_2] = 1$ .
- $\mathbb{P}[g(t_1, t_2) = t_1] = \mathbb{P}[g(t_2, t_1) = t_1]$ .

A gameplay function represents a process in which two teams compete in a game, with one of them emerging as the winner. This model simplifies away effects like home-field advantage or teams improving over the course of a tournament: a gameplay function is fully described by a single probability for each pair of teams in the list.

### Definition 1.1.2: Tournament Format

A *tournament format* is an algorithm that takes as input a list of teams  $\mathcal{T}$  and a gameplay function  $g$  and outputs a champion  $t \in \mathcal{T}$ .

### Definition 1.1.3: Playing, Winning, and Losing

If a tournament format queries  $g$  on input  $(t_1, t_2)$  we say that  $t_1$  and  $t_2$  *played a game*. We say that the team that got outputted by  $g$  *won*, and the team that did not *lost*.

We also introduce some shorthand to help make mathematical notation more concise.

### Definition 1.1.4: $\mathbb{W}_{\mathcal{A}}(t, \mathcal{T})$

$\mathbb{W}_{\mathcal{A}}(t, \mathcal{T})$  is the probability that team  $t \in \mathcal{T}$  wins tournament format  $\mathcal{A}$  when it is run on the list of teams  $\mathcal{T}$ .

(This chapter will be fleshed out but I'm including the important definitions here for the sake of the next chapter.)

## 2 Brackets

## 2.1 Brackets and Rounds

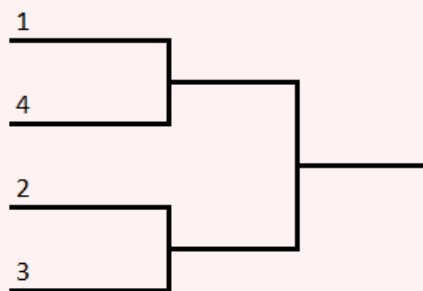
### Definition 2.1.1: Bracket

A *bracket* is a tournament format in which:

- Teams don't play any games after their first loss,
- Games are played until only one team has no losses, and that team is crowned champion, and
- The matchups between teams that have not yet lost are determined based on the ordering of the teams in  $\mathcal{T}$  in advance of the outcomes of any games.

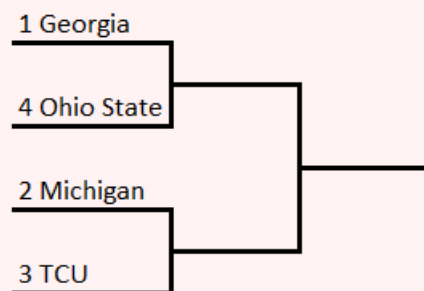
We can draw a bracket as a tree-like structure in the following way:

Figure 2.1.2: The 2023 College Football Playoff



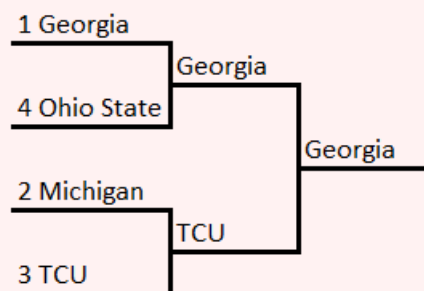
The numbers 1, 2, 3, and 4 indicate where the first, second, third and fourth team in  $\mathcal{T}$  are placed to start. In the actual 2023 College Football Playoff, the list of teams  $\mathcal{T}$  was Georgia, Michigan, TCU, and Ohio State, in that order, so the bracket was filled in like so:

**Figure 2.1.3: The 2023 CFP After Team Placement**



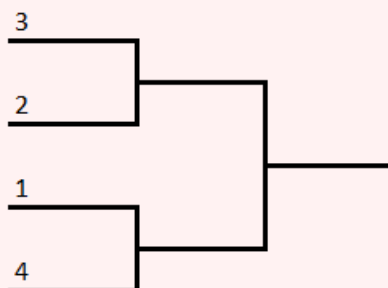
As games are played, we write the name of the winning teams on the corresponding lines. This bracket tells us that Georgia played Ohio State, and Michigan played TCU. Georgia and TCU won their respective games, and then Georgia beat TCU, winning the tournament.

**Figure 2.1.4: The 2023 CFP After Completion**



Rearranging the way the bracket is pictured, if it doesn't affect any of the matchups, does not create a new bracket. For example, Figure 2.1.5 is just another way to draw the same 2023 CFP Bracket.

Figure 2.1.5: Alternative Drawing of the 2023 CFP



One key piece of bracket vocabulary is the *round*.

**Definition 2.1.6: Round**

A *round* is a set of games such that the winners of each of those games have the same number of games remaining to win the tournament.

For example, the 2023 CFP has two rounds. The first round included the games Georgia vs Ohio State and Michigan vs TCU, and the second round was just a single game: Georgia vs TCU.

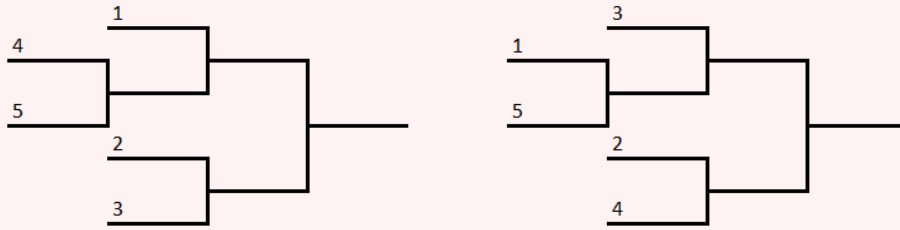
## 2.2 Bracket Signatures

### Definition 2.2.1: Shape

The *shape* of a bracket is the tree that underlies it.

For example, the following two brackets have the same shape:

Figure 2.2.2: Two Brackets with the Same Shape



### Definition 2.2.3: Bye

A team has a *bye* in round  $r$  if it plays no games in round  $r$  or before.

One way to describe the shape of a bracket is its signature.

### Definition 2.2.4: Bracket Signature

The *signature*  $[[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  of an  $r$ -team bracket  $\mathcal{A}$  is a list of natural numbers, such that  $a_i$  is the number of teams with  $i$  byes.

The signature of a bracket is defined only by its shape: the two brackets in Figure 2.2.2 have the same shape, so they also have the same signature.

The signatures of the brackets discussed in this section are shown in Figure 2.2.5. It's worth verifying the signatures we've seen so far and trying to draw brackets with the signatures we haven't yet before moving on.

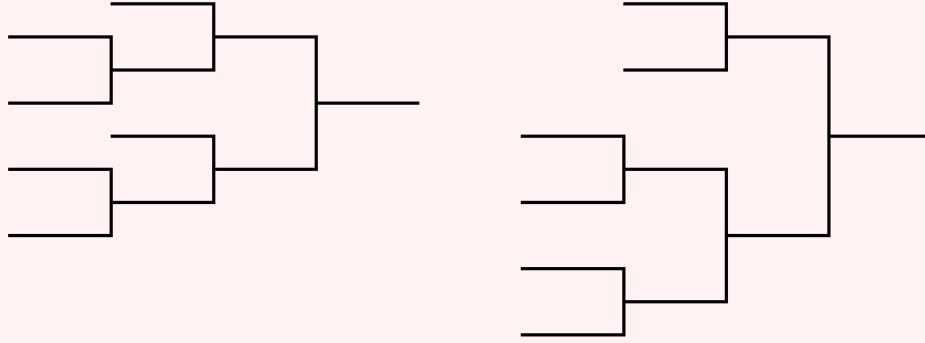


**Figure 2.2.5: The Signatures of Some Brackets**

Bracket	Signature
2023 College Football Playoff	$[[4; 0; 0]]$
The brackets in Figure 2.2.2	$[[2; 3; 0; 0]]$
The brackets in Figure 2.2.6	$[[4; 2; 0; 0]]$
2023 WCC Men's Basketball Tournament	$[[4; 2; 2; 2; 0; 0]]$

Two brackets with the same shape must have the same signature, but the converse is not true: two brackets with different shapes can have the same signature. For example, both bracket shapes depicted in Figure 2.2.6 have the signature  $[[4; 2; 0; 0]]$ .

**Figure 2.2.6: Two Shapes with the Signature  $[[4; 2; 0; 0]]$**



Despite this, bracket signatures are a useful way to talk about the shape of a bracket. Communicating a bracket's signature is a lot easier than communicating its shape, and much of the important information (such as how many games each team must win in order to win the tournament) is contained in the signature.

Bracket signatures have one more important property.

**Theorem 2.2.7**

Let  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  be a list of natural numbers. Then  $\mathcal{A}$  is a bracket

signature if and only if

$$\sum_{i=0}^r a_i \cdot \left(\frac{1}{2}\right)^{r-i} = 1.$$

*Proof.* Let  $\mathcal{A}$  be the signature for some bracket. Assume that every game in the bracket was a coin flip, and consider each team's probability of winning the tournament. A team that has  $i$  byes must win  $r - i$  games to win the tournament, and so will do so with probability  $\left(\frac{1}{2}\right)^{r-i}$ . For each  $i \in \{0, \dots, r\}$ , there are  $a_i$  teams with  $i$  byes, so (because any two teams winning are mutually exclusive)

$$\sum_{i=0}^r a_i \cdot \left(\frac{1}{2}\right)^{r-i}$$

is the probability that one of the teams wins, which is 1.

We prove the other direction by induction on  $r$ . If  $r = 0$ , then the only list with the desired property is  $[[1]]$ , which is the signature for the unique one-team bracket. For any other  $r$ , first note that  $a_0$  must be even: if it were odd, then

$$\begin{aligned} \sum_{i=0}^r a_i \cdot \left(\frac{1}{2}\right)^{r-i} &= \frac{1}{2^r} \cdot \sum_{i=0}^r a_i \cdot 2^i \\ &= \frac{1}{2^r} \cdot \left(a_0 + 2 \sum_{i=1}^r a_i \cdot 2^{i-1}\right) \\ &= k/2^r && \text{for some odd } k \\ &\neq 1. \end{aligned}$$

Now, consider the signature  $\mathcal{B} = [[\mathbf{a}_1 + \mathbf{a}_0/2; \mathbf{a}_2; \dots; \mathbf{a}_r]]$ . By induction, there exists a bracket with signature  $\mathcal{B}$ . But if we take that bracket and replace  $a_0/2$  of the teams with no byes with a game whose winner gets placed on that line, we get a new bracket with signature  $\mathcal{A}$ .  $\square$

The operation of transforming a bracket signature  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  into a bracket signature with one fewer round  $\mathcal{B} = [[\mathbf{a}_1 + \mathbf{a}_0/2; \mathbf{a}_2; \dots; \mathbf{a}_r]]$  that we

used at the end of the proof of Theorem 2.2.7 will become somewhat frequent as we often induct on the number of rounds in a bracket, so it has a name:

**Definition 2.2.8: The Successor Signature**

If  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$ , then the successor signature

$$\mathfrak{S}(\mathcal{A}) = [[\mathbf{a}_1 + \mathbf{a}_0/2; \mathbf{a}_2; \dots; \mathbf{a}_r]].$$

(The successor signature of zero-round signatures is undefined.)

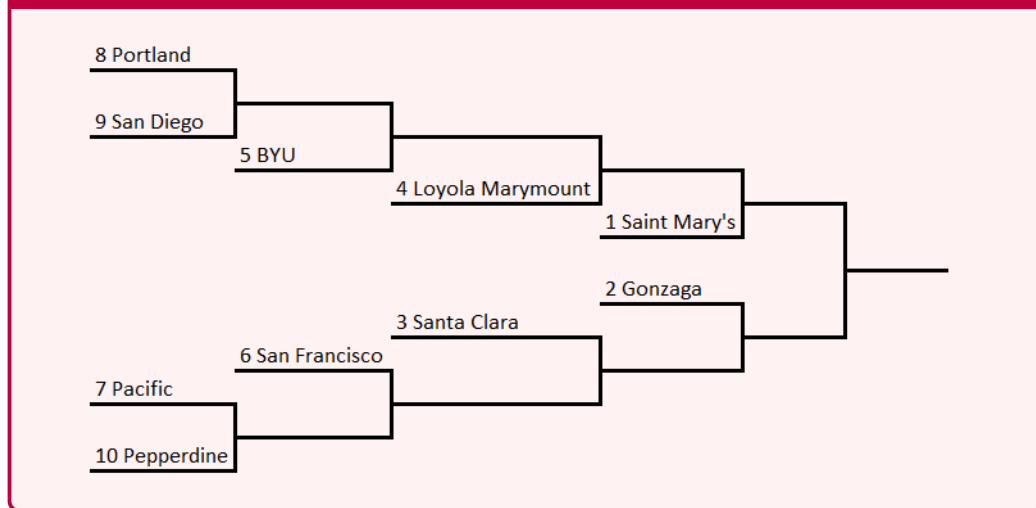
In the next few sections, we will use the language and properties of bracket signatures to describe the brackets that we work with. For now though, let's return to the 2023 College Football Playoff. The bracket used in the 2023 CFP has a special property that not all brackets have: it is *balanced*.

**Definition 2.2.9: Balanced Bracket**

A *balanced bracket* is a bracket in which none of the teams have byes.

The 2023 West Coast Conference Men's Basketball Tournament, on the other hand, is unbalanced:

**Figure 2.2.10: The 2023 WCC Men's Basketball Tournament**



Saint Mary's and Gonzaga each have three byes and so only need to win two games to win the tournament, while Portland, San Diego, Pacific, and

Pepperdine need to win five. Unsurprisingly, this format conveys a massive advantage to Saint Mary's and Gonzaga, but this was intentional: those two teams were being rewarded for doing the best during the regular season.

In many cases, however, it is undesirable to grant advantages to certain teams over others. One might hope, for any  $n$ , to be able to construct a balanced bracket for  $n$  teams, but unfortunately this is rarely possible.

#### Theorem 2.2.11

There exists an  $n$ -team balanced bracket if and only if  $n$  is a power of two.

*Proof.* A bracket is balanced if no teams have byes, which is true exactly when its signature is of the form  $\mathcal{A} = [[\mathbf{n}; \mathbf{0}; \dots; \mathbf{0}]]$  where  $n$  is the number of teams in the bracket. If  $n$  is a power of two, then by Theorem 2.2.7  $\mathcal{A}$  is indeed a bracket signature and so points to a balanced bracket for  $n$  teams. If  $n$  is not a power of two, however, then Theorem 2.2.7 tells us that  $\mathcal{A}$  is not a bracket signature, and so no balanced brackets exist for  $n$  teams.  $\square$

Given this, brackets are not a great option when we want to avoid giving some teams advantages over others unless we have a power of two teams. They are a fantastic tool, however, if doling out advantages is the goal, perhaps after some teams did better during the regular season and ought to be rewarded with an easier path in the bracket.

## 2.3 Proper Brackets

### Definition 2.3.1: Seeding

The *seeding* of an  $n$ -team bracket is the arrangement of the numbers 1 through  $n$  in the bracket.

Together, the shape and seeding fully specify a bracket.

### Definition 2.3.2: $i$ -seed

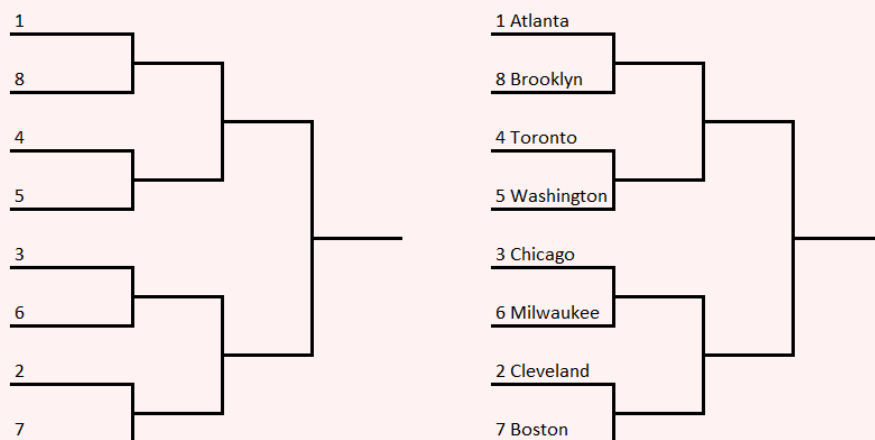
The  $i$ -seed is  $i$ th team in list of teams inputted to a bracket. Sometimes we use the notation  $t_i$  to refer to the  $i$ -seed in a bracket.

### Definition 2.3.3: Higher and Lower Seeds

Somewhat confusingly, convention is that smaller numbers are the *higher seeds*, and greater numbers are the *lower seeds*.

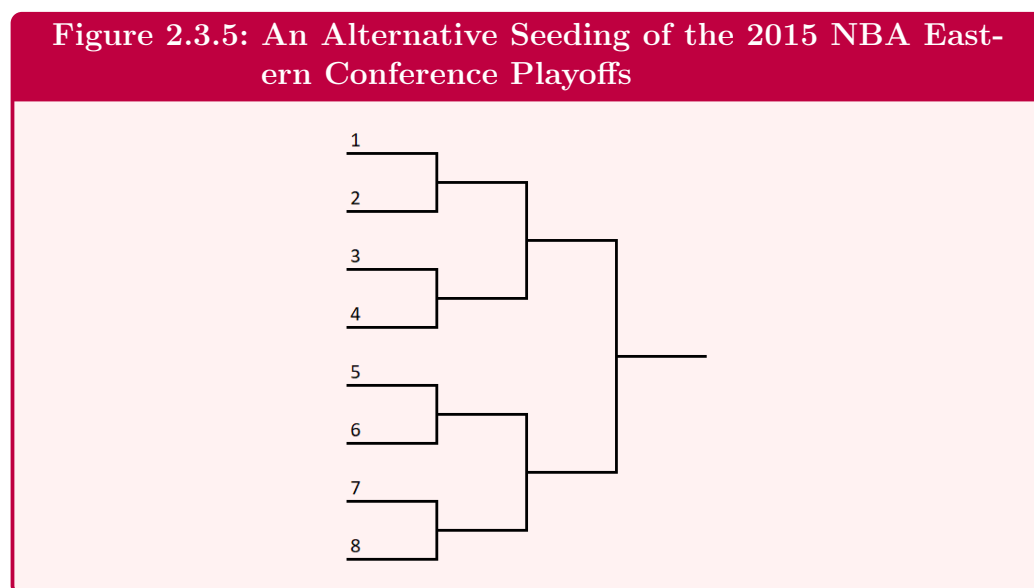
Seeding is typically used to reward better and more deserving teams. As an example, on the left is the 8-team bracket used in the 2015 NBA Eastern Conference Playoffs. At the end of the regular season, the top eight teams in the Eastern Conference were ranked and placed into the bracket as shown on the right.

Figure 2.3.4: 2015 NBA Eastern Conference Playoffs



Despite this bracket being balanced, the higher seeds are still at advantage: they have an easier set of opponents. Compare 1-seed Atlanta, whose first two rounds are versus 8-seed Brooklyn and then (most likely) 4-seed Toronto, versus 7-seed Boston, whose first two rounds are versus 2-seed Cleveland and then (most likely) 3-seed Chicago. Atlanta's schedule is far easier: despite them having the same number of games to win as Boston, Atlanta is expected to play lower seeds in each round than Boston will.

Thus, we've identified two ways in which brackets can convey an advantage onto certain teams: by giving them more byes, and by giving them easier (expected) opponents. Not every seeding of a bracket does this: for example, consider the following alternative seeding for the 2015 NBA Eastern Conference Playoffs.



This seeding does a very poor job of rewarding the higher-seeded teams: the 1- and 2-seeds are matched up in the first round, while the easiest road is given to the 7-seed, who plays the 8-seed in the first round and then (most likely) the 5-seed in the second. Since the whole point of seeding is to give the higher-seeded teams an advantage, we introduce the concept of a *proper seeding*.

### Definition 2.3.6: Chalk

We say a tournament *went chalk* if the higher-seeded team won every game during the tournament.

### Definition 2.3.7: Proper Seeding

A *proper seeding* of a bracket is one such that if the bracket goes chalk, in every round it is better to be a higher-seeded team than a lower-seeded one, where:

- (1) It is better to have a bye than to play a game.
- (2) It is better to play a lower seed than to play a higher seed.

### Definition 2.3.8: Proper Bracket

A *proper bracket* is a bracket that has been properly seeded.

It is clear that the actual 2015 NBA Eastern Conference Playoffs was properly seeded, while our alternative seeding was not.

A few quick lemmas about proper brackets:

### Lemma 2.3.9

In a proper bracket, if  $m$  teams have a bye in a given round, those teams must be seeds 1 through  $m$ .

*Proof.* If they did not, the seeding would be in violation of condition (1). □

### Lemma 2.3.10

If a proper bracket goes chalk, then after each round the  $m$  teams remaining will be the top  $m$  seeds.

*Proof.* We will prove the contrapositive. Assume that for some  $i < j$ , after some round,  $t_i$  has been eliminated but  $t_j$  is still alive. Let  $k$  be the seed of the team that  $t_i$  lost to. Because the bracket went chalk,

$k < i$ . Now consider what  $t_j$  did in that round. If they had a bye, then the bracket violates condition (1). Assume instead they played  $t_\ell$ . They beat  $t_\ell$ , so  $j < \ell$ , giving,

$$k < i < j < \ell.$$

In the round that  $t_i$  was eliminated,  $t_i$  played  $t_k$ , while  $t_j$  played  $t_\ell$ , violating condition (2). Thus, the bracket is not proper.  $\square$

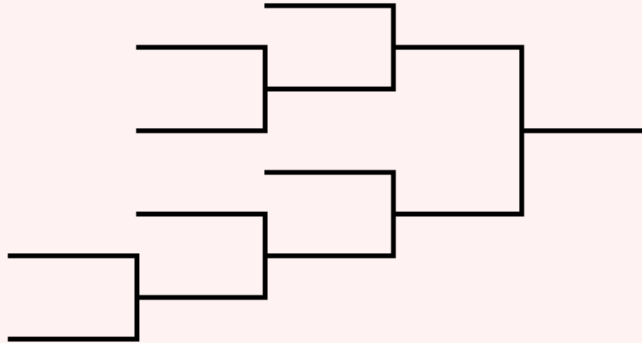
### Lemma 2.3.11

In a proper bracket, if  $m$  teams have a bye and  $k$  games are being played in a given round, then if the bracket goes chalk those matchups will be seed  $m + i$  vs seed  $(m + 2k + 1) - i$  for  $i \in \{1, \dots, k\}$ .

*Proof.* In the given round, there are  $m + 2k$  teams remaining. Theorem 2.3.10 tells us that (if the bracket goes chalk) those teams must be seeds 1 through  $m + 2k$ . Theorem 2.3.9 tells us that seeds 1 through  $m$  must have a bye, so the teams playing must be seeds  $m + 1$  through  $m + 2k$ . Then condition (2) tells us that the matchups must be exactly  $m + i$  vs seed  $(m + 2k + 1) - i$  for  $i \in \{1, \dots, k\}$ .  $\square$

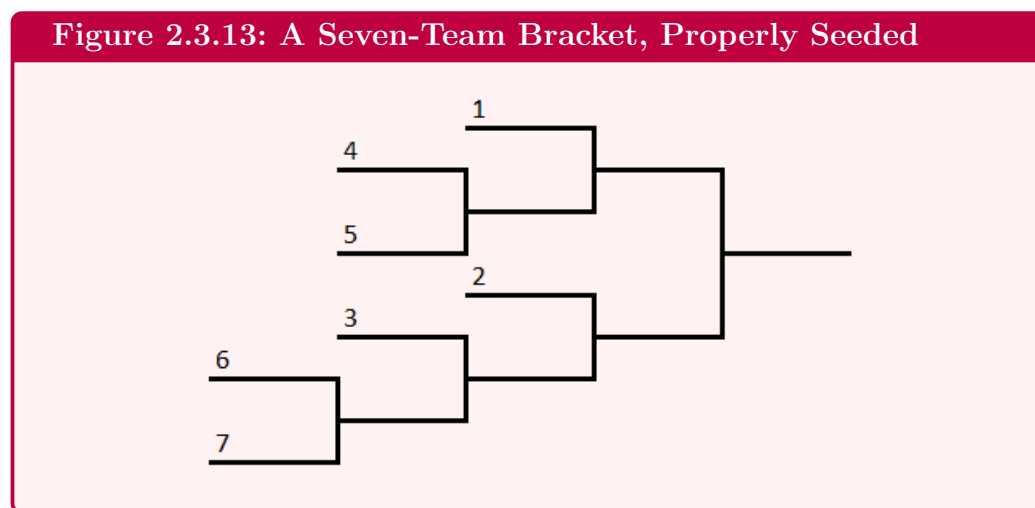
We can use Lemmas 2.3.9 through 2.3.11 to properly seed various bracket shapes. For example, consider the following seven-team shape:

Figure 2.3.12: A Seven-Team Bracket Shape





Lemma 2.3.9 tells us that the first-round matchup must be between the 6-seed and the 7-seed. Lemma 2.3.11 tells us that if the bracket goes chalk, the second-round matchups must be 3v6 and 4v5, so the 3-seed play the winner of the first-round matchup. Finally, we can apply Lemma 2.3.11 again to the semifinals to find that the 1-seed should play the winner of the 4v5 matchup, while the 2-seed should play the winner of the 3v(6v7) matchup. In total, our proper seeding looks like:

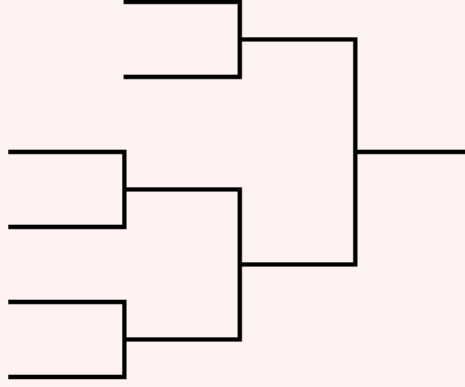


We can also quickly simulate the bracket going chalk to verify Lemma 2.3.10.

Lemmas 2.3.9 through 2.3.11 are quite powerful. It is not a coincidence that we managed to specify exactly what a proper seeding of the above bracket must look like with no room for variation: soon we will prove that the proper seeding for a particular bracket shape is unique.

But not every shape admits even this one proper seeding. Consider the following six-team shape:

**Figure 2.3.14: A Six-Team Bracket Shape**



This shape admits no proper seedings. Lemma 2.3.9 requires that the two teams getting byes be the 1- and 2-seed, but this violates Lemma 2.3.11 which requires that in the second round the 1- and 2-seeds do not play each other. So how can we think about which shapes admit proper seedings?

**Theorem 2.3.15: The Fundamental Theorem of Brackets**

There is exactly one proper bracket with each bracket signature.

*Proof.* Let  $\mathcal{A}$  be an  $r$ -round bracket signature. We proceed by induction on  $r$ . If  $r = 0$ , then the only possible bracket signature is  $[[1]]$ , and it points to the unique one-team bracket, which is indeed proper.

For any other  $r$ , the first-round matchups of a proper bracket with signature  $\mathcal{A}$  are defined by Theorem 2.3.11. Then if those matchups go chalk, we are left with a proper bracket of signature  $\mathfrak{S}(\mathcal{A})$ , which induction tells us exists admits exactly one proper bracket.

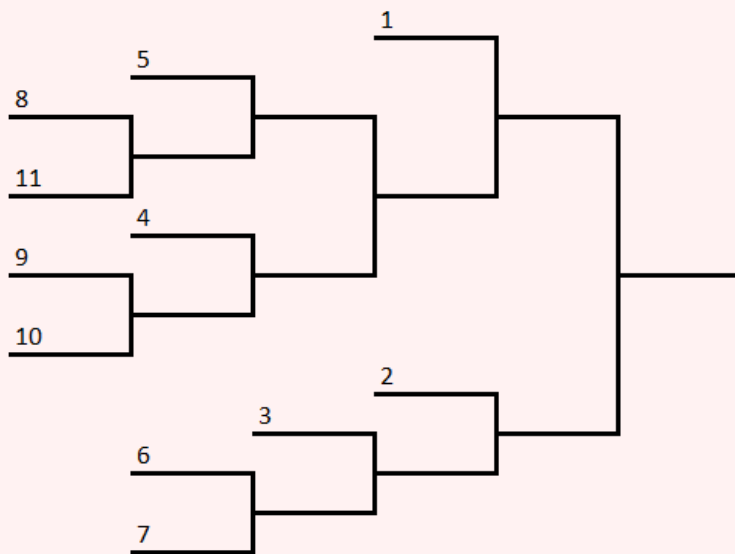
Thus both the first-round matchups and the rest of the bracket are determined, and by combining them we get a proper bracket with signature  $\mathcal{A}$ , so there is exactly one proper bracket with signature  $\mathcal{A}$ .  $\square$

The fundamental theorem of brackets means that we can refer to the

proper bracket  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  in a well-defined way, as long as

$$\sum_{i=0}^r a_i \cdot \left(\frac{1}{2}\right)^{r-i} = 1.$$

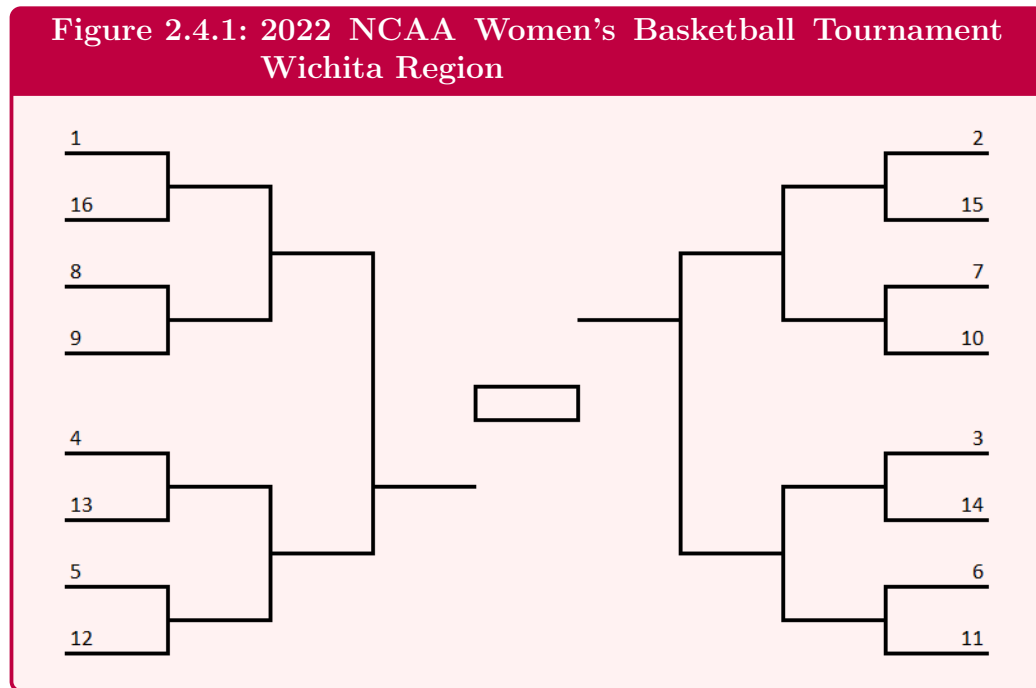
In practice, virtually every sports league that uses a traditional bracket uses a proper one: while different leagues take very different approaches to how many byes to give teams (compare the 2023 West Coast Conference Men's Basketball Tournament with the 2015 NBA Eastern Conference Play-offs), they are almost all proper. This makes bracket signatures a convenient labeling system for the set of brackets that we might reasonably encounter. They also are a powerful tool for specifying new brackets: if you are interested in (say) an eleven-team bracket where four teams get no byes, four teams get one bye, one team gets two byes and two teams get three byes, we can describe the proper bracket with those specs as  $[[4; 4; 1; 2; 0; 0]]$  and use Lemmas 2.3.9 through 2.3.11 to draw it with ease:

Figure 2.3.16: The Proper Bracket of Signature  $[[4; 4; 1; 2; 0; 0]]$ 

Due to these properties, we will almost exclusively discuss proper bracket from here on out: unless stated otherwise, assume all brackets are proper.

## 2.4 Ordered Brackets

Consider the proper bracket  $[[16; 0; 0; 0; 0]]$ , which was used in the 2022 NCAA Women’s Basketball Tournament Wichita Region, and is shown below. (Sometimes brackets are drawn in the manner below, with teams starting on both sides and the winner of each side playing in the championship game.)



The definition of a proper seeding ensures that as long as the bracket goes chalk (that is, higher seeds always beat lower seeds), it will always be better to be a higher seed than a lower seed. But what if it doesn’t go chalk?

One counter-intuitive fact about the NCAA Basketball Tournament is that it is probably better to be a 10-seed than a 9-seed. (This doesn’t violate the proper seeding property because 9-seeds have an easier first-round matchup than 10-seeds, and for further rounds, proper seedings only care about what happens if the bracket goes chalk, which would eliminate both the 9-seed and 10-seed in the first round.) Why? Let’s look at whom each seed-line matchups against in the first two rounds:

**Figure 2.4.2: NCAA Basketball Tournament 9- and 10-seed Schedules**

Seed	First Round	Second Round
9	8	1
10	7	2

The 9-seed has an easier first-round matchup, while the 10-seed has an easier second-round matchup. However, this isn't quite symmetrical. Because the teams are probably drawn from a roughly normal distribution, the expected difference in skill between the 1- and 2-seeds is far greater than the expected difference between the 7- and 8-seeds, implying that the 10-seed does in fact have an easier route than the 9-seed.

Nate Silver [6] investigated this matter in full, finding that in the NCAA Basketball Tournament, seed-lines 10 through 15 give teams better odds of winning the region than seed-lines 8 and 9. Of course this does not mean that the 11-seed (say) has a better chance of winning a given region than the 8-seed does, as the 8-seed is a much better team than the 11-seed. But it does mean that the 8-seed would love to swap places with the 11-seed, and that doing so would increase their odds to win the region.

This is not a great state of affairs: the whole point of seeding is confer an advantage to higher-seeded teams, and the proper bracket  $[[16; 0; 0; 0; 0]]$  is failing to do that. Not to mention that giving lower-seeded teams an easier route than higher-seeded ones can incentivize teams to lose during the regular season in order to try to get a lower but more advantageous seed.

To fix this, we need a stronger notion of what makes a bracket effective than properness. The issue with proper seedings is the false assumption that higher-seeded teams will always beat lower-seeded teams. A more nuanced assumption, initially proposed by H.A. David [2], might look like this:

**Definition 2.4.3: Strongly Stochastically Transitive**

A list of teams is *strongly stochastically transitive* if for each  $i, j, k$  such that  $j < k$ ,

$$\mathbb{P}[t_i \text{ beats } t_j] \leq \mathbb{P}[t_i \text{ beats } t_k].$$

A list of teams being strongly stochastically transitive (or SST) captures the intuition that each team ought to do better against lower-seeded teams

than against higher-seeded teams. A few quick implications of this definition are:

**Corollary 2.4.4**

If a list of teams is SST, then for each  $i < j$ ,  $\mathbb{P}[t_i \text{ beats } t_j] \geq 0.5$ .

**Corollary 2.4.5**

If a list of teams is SST, then for each  $i, j, k, \ell$  such that  $i < j$  and  $k < \ell$ ,

$$\mathbb{P}[t_i \text{ beats } t_\ell] \geq \mathbb{P}[t_j \text{ beats } t_k].$$

Note that not every set of teams can be seeded to be SST. Consider, for example, the game of rock-paper-scissors. Rock beats paper which beats scissors which beats rock, so no ordering of these “teams” will be SST. For our purposes, however, SST will work well enough.

Our new, nuanced alternative a proper bracket is an *ordered bracket*, first defined by Chen and Hwang [1].

**Definition 2.4.6: Ordered**

A tournament format  $\mathcal{A}$  is *ordered* if, for any SST list of teams  $\mathcal{T}$ , if  $i < j$ , then  $\mathbb{W}_{\mathcal{A}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{T})$

In an informal sense, a bracket being ordered is the strongest thing we can want without knowing more about why the tournament is being played. Depending on the situation, we might be interested in a format that almost always declares the most-skilled team as the winner, or in a format that gives each team roughly the same chance of winning, or anywhere in between. But certainly, better teams should win more, which is what the ordered bracket condition requires.

In particular, a bracket being ordered is a stronger claim than it being proper.

**Theorem 2.4.7**

Every ordered bracket is proper.

*Proof.* Let  $\mathcal{A}$  be an ordered  $n$ -team bracket with  $r$  rounds.

Consider a list of teams such that every team wins every game with probability 0.5. These teams are SST. A team that plays their first game in the  $i$ th round will win the tournament with probability  $(\frac{1}{2})^{r-i}$ , so teams that get more byes will have a higher probability to win the tournament than teams with fewer byes. This implies that higher-seeded teams must have more byes than lower-seeded teams, so in each round, the teams with byes must be the highest-seeded teams that are still alive. Thus, condition (1) is met.

We show that condition (2) is met by proving the stronger condition from Lemma 2.3.11: if  $m$  teams have a bye and  $k$  games are being played in round  $s$ , then if the bracket goes chalk, those matchups will be  $t_{m+i}$  vs  $t_{(m+2k+1)-i}$  for  $i \in \{1, \dots, k\}$ . We show this by strong induction on  $s$  and on  $i$ .

Assume that this is true for every round up until  $s$  and for all  $i < j$  for some  $j$ . Let  $\ell = (m + 2k + 1) - j$ . We want to show that if the bracket goes chalk,  $t_{m+j}$  will face off against seed  $t_\ell$  in the given round. Consider the following SST matchup table: every game is a coin flip, except for games involving a team seeded  $\ell$  or lower, in which case the higher seed always wins. Then, each team seeded between  $\ell - 1$  and  $m + j$  will win the tournament with probability  $(\frac{1}{2})^{r-s}$ , other than the team slated to play  $t_\ell$  in round  $s$  who wins with probability  $(\frac{1}{2})^{r-i-1}$ . In order for  $\mathcal{B}$  to be ordered, that team must be  $t_{m+j}$ .

Thus  $\mathcal{A}$  satisfies both conditions, and so is a proper bracket. □

With Theorem 2.4.7, we can use the language of bracket signatures to describe ordered brackets without worrying that two ordered brackets might share a signature. Now we examine three particularly important examples of ordered brackets.

We begin with the unique one-team bracket.

**Figure 2.4.8: The One-Team Bracket  $[[1]]$**



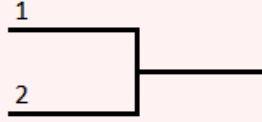
**Theorem 2.4.9**

The one-team bracket  $[[1]]$  is ordered.

*Proof.* Since there is only team, the ordered bracket condition is vacuously true.  $\square$

Next we look at the unique two-team bracket.

**Figure 2.4.10: The Two-Team Bracket  $[[2; 0]]$**



**Theorem 2.4.11**

The two-team bracket  $[[2; 0]]$  is ordered.

*Proof.* Let  $\mathcal{A} = [[2; 0]]$ . Then,

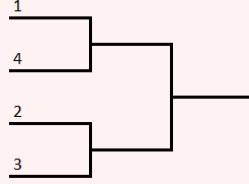
$$\mathbb{W}_{\mathcal{A}}(t_1, \mathcal{T}) = \mathbb{P}[t_1 \text{ beats } t_2] \geq 0.5 \geq \mathbb{P}[t_2 \text{ beats } t_1] = \mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T})$$

so  $\mathcal{A}$  is ordered.  $\square$

And thirdly, we show that the balanced four-team bracket is ordered, first proved by Horen and Riezman [4].



**Figure 2.4.12: The Four-Team Bracket  $[[4; \mathbf{0}; \mathbf{0}]]$**



**Theorem 2.4.13**

The four-team bracket  $[[4; \mathbf{0}; \mathbf{0}]]$  is ordered.

*Proof.* Let  $\mathcal{A} = [[4; \mathbf{0}; \mathbf{0}]]$  and let  $p_{ij} = \mathbb{P}[t_i \text{ beats } t_j]$ . Then,

$$\begin{aligned} \mathbb{W}_{\mathcal{A}}(t_1, \mathcal{T}) &= p_{14} \cdot (p_{23}p_{12} + p_{32}p_{13}) \\ &= p_{14}p_{23}p_{12} + p_{14}p_{32}p_{13} \\ &\geq p_{14}p_{23}p_{21} + p_{24}p_{41}p_{23} \\ &= p_{23} \cdot (p_{14}p_{21} + p_{41}p_{24}) \\ &= \mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T}) \end{aligned}$$

$$\begin{aligned} \mathbb{W}_{\mathcal{A}}(t_2, \mathcal{T}) &= p_{23} \cdot (p_{14}p_{21} + p_{41}p_{24}) \\ &\geq p_{32} \cdot (p_{14}p_{31} + p_{41}p_{34}) \\ &= \mathbb{W}_{\mathcal{A}}(t_3, \mathcal{T}) \end{aligned}$$

$$\begin{aligned} \mathbb{W}_{\mathcal{A}}(t_3, \mathcal{T}) &= p_{32} \cdot (p_{14}p_{31} + p_{41}p_{34}) \\ &= p_{32}p_{14}p_{31} + p_{32}p_{41}p_{34} \\ &\geq p_{42}p_{23}p_{41} + p_{32}p_{41}p_{43} \\ &= p_{41} \cdot (p_{23}p_{42} + p_{32}p_{43}) \\ &= \mathbb{W}_{\mathcal{A}}(t_4, \mathcal{T}) \end{aligned}$$

Thus the four-team bracket  $[[4; \mathbf{0}; \mathbf{0}]]$  is ordered. □

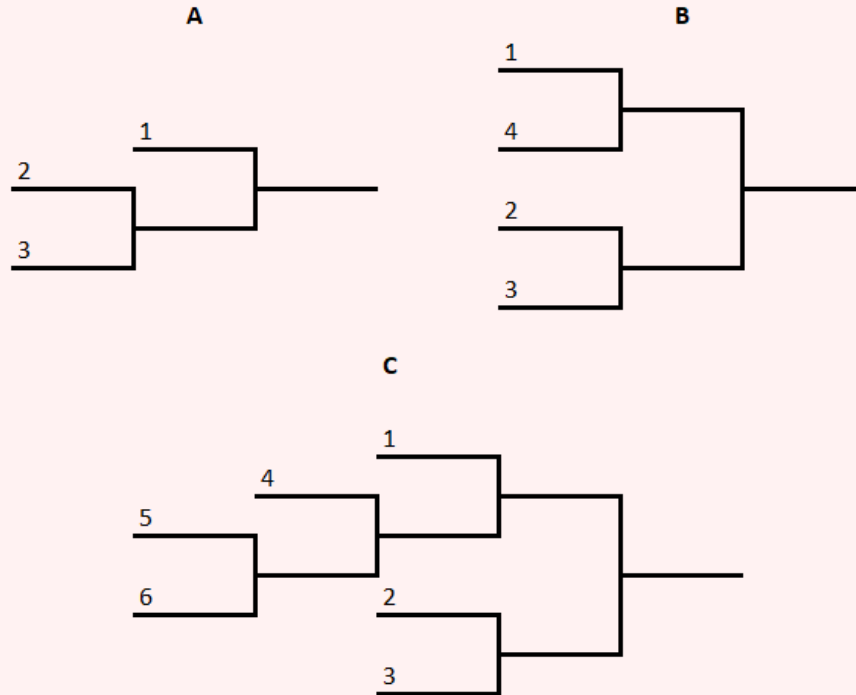
In the next section, we move on from describing particular ordered brackets in favor of a more general result.

## 2.5 Edwards's Theorem

We now attempt to completely specify the set of ordered brackets, culminating in Corollary 2.5.4 and Theorem 2.5.6, which together form Edwards's Theorem. Edwards [3] originally proved the theorem without access to the machinery of bracket signatures or proper brackets: we present a quicker proof that makes use of the fundamental theorem of brackets.

We begin with a theorem that allows us to combine two smaller ordered brackets together into a larger ordered one by having the winner of one of the brackets be treated as the lowest seed in the other. This is depicted in Figure 2.5.1.

**Figure 2.5.1: The Setup of Theorem 2.5.2**



### Theorem 2.5.2

If  $\mathcal{A}$  and  $\mathcal{B}$  are  $n$ - and  $m$ -team ordered brackets, respectively, we can construct an  $(n + m - 1)$ -team ordered bracket by replacing the lowest

seed in  $\mathcal{B}$  with the entire bracket  $\mathcal{A}$  (and lowering the seed of each team in  $\mathcal{A}$  by  $m - 1$ ).

*Proof.* Let  $\mathcal{C}$  be the bracket formed by merging  $\mathcal{A}$  and  $\mathcal{B}$ . Let  $\mathcal{T}$  be an SST list of teams  $n + m - 1$  teams, and let  $\mathcal{R}, \mathcal{S} \subset \mathcal{T}$  be the lowest  $n$  and the highest  $m - 1$  seeds of  $\mathcal{T}$  respectively. We divide proving that  $\mathcal{C}$  is ordered into proving three substatements:

1. For  $i < j < m$ ,  $\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})$
2.  $\mathbb{W}_{\mathcal{C}}(t_{m-1}, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_m, \mathcal{T})$
3. For  $m \leq i < j$ ,  $\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) \geq \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})$

Together, these show that  $\mathcal{C}$  is ordered.

We begin with the first sub-statement. Let  $i < j < m$ . Then,

$$\begin{aligned} \mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) &= \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_k\}) \\ &\geq \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_j, \mathcal{S} \cup \{t_k\}) \\ &= \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T}) \end{aligned}$$

The first and last equalities follow from the structure of  $\mathcal{C}$ , and the inequality follows from  $\mathcal{B}$  being ordered.

Now the second sub-statement.

$$\begin{aligned} \mathbb{W}_{\mathcal{C}}(t_{m-1}, \mathcal{T}) &= \sum_{k=m}^{n+m-1} \mathbb{W}_{\mathcal{A}}(t_k, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_{m-1}, \mathcal{S} \cup \{t_k\}) \\ &\geq \mathbb{W}_{\mathcal{A}}(t_m, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_{m-1}, \mathcal{S} \cup \{t_m\}) \\ &\geq \mathbb{W}_{\mathcal{A}}(t_m, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_m, \mathcal{S} \cup \{t_m\}) \\ &= \mathbb{W}_{\mathcal{C}}(t_m, \mathcal{T}) \end{aligned}$$

The equalities follow from the structure of  $\mathcal{C}$ , the first equality follows from probabilities being non-negative, and the second inequality

follows from  $\mathcal{B}$  being ordered.

Finally we show the third sub-statement. Let  $m \leq i < j$ . Then,

$$\begin{aligned}\mathbb{W}_{\mathcal{C}}(t_i, \mathcal{T}) &= \mathbb{W}_{\mathcal{A}}(t_i, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_i\}) \\ &\geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_i, \mathcal{S} \cup \{t_i\}) \\ &\geq \mathbb{W}_{\mathcal{A}}(t_j, \mathcal{R}) \cdot \mathbb{W}_{\mathcal{B}}(t_j, \mathcal{S} \cup \{t_j\}) \\ &= \mathbb{W}_{\mathcal{C}}(t_j, \mathcal{T})\end{aligned}$$

The equalities follow from the structure of  $\mathcal{C}$ , the first inequality from  $\mathcal{A}$  being ordered, and the second inequality from the teams being SST.

We have shown all three sub-statements, and so  $\mathcal{C}$  is ordered.  $\square$

### Corollary 2.5.3

If  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  and  $\mathcal{B} = [[\mathbf{b}_0; \dots; \mathbf{b}_s]]$  are ordered brackets, then  $\mathcal{C} = [[\mathbf{a}_0; \dots; \mathbf{a}_r + \mathbf{b}_0 - \mathbf{1}; \dots; \mathbf{b}_s]]$  is an ordered bracket as well.

We can then construct the set of brackets that we have thus far shown are ordered. We do this by starting with  $\{[[\mathbf{1}]], [[\mathbf{2}; \mathbf{0}]], [[\mathbf{4}; \mathbf{0}; \mathbf{0}]]\}$  and then repeatedly applying the above theorem on the set to expand it. In other words,

### Corollary 2.5.4

Any bracket signature formed by the following process is ordered:

1. Start with the list  $[[\mathbf{0}]]$  (note that this not yet a bracket signature).
2. As many times as desired, prepend the list with  $[[\mathbf{1}]]$  or  $[[\mathbf{3}; \mathbf{0}]]$ .
3. Then, add 1 to the first element in the list, turning it into a bracket signature.

Corollary 2.5.4 uses the tools that we have developed so far to demark a set of brackets as ordered. Somewhat surprisingly, this set is complete: any

bracket not reachable using the process in Corollary 2.5.4 is not ordered. To prove this we first show a lemma.

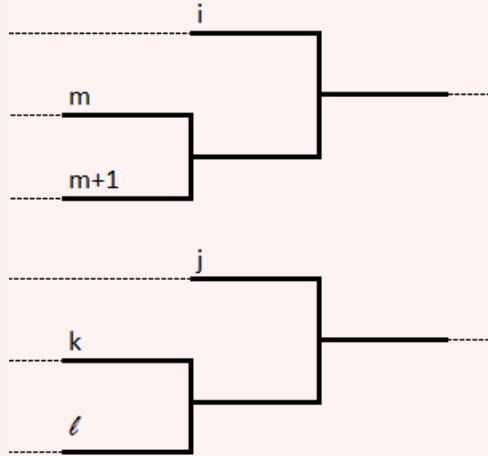
### Lemma 2.5.5

If  $\mathcal{A}$  is an  $r$ -round ordered bracket and two games are played in round  $s$ , then the winners of those games must play each other in round  $s+1$ . Further, this must be the only game of round  $s+1$ .

*Proof.* Let  $m$  be the number of teams that are still alive in  $\mathcal{A}$  after round  $s$ . Because  $\mathcal{A}$  is proper, we know that if the bracket goes chalk,  $t_m$  will play  $t_{m+1}$  in round  $s$ .

Now, assume for contradiction that two games are played in round  $s$  but the winners of those games do not play each other in the following round. So let  $k < \ell$  such that, if the bracket goes chalk,  $t_k$  plays  $t_\ell$  in round  $s$ , but the winner of that game doesn't play the winner of the  $t_m$  game in round  $s+1$ . In fact, let  $k$  be the lowest such seed. Finally, let  $i$  and  $j$  be such that if the bracket goes chalk, in round  $s+1$ ,  $t_i$  will play  $t_m$  and  $t_j$  will play  $t_k$ .

The situation so far (assuming all omitted games go chalk):



We can use  $\mathcal{A}$ 's properness to determine that  $i < j < k < m < m+1 < \ell$ .

Consider now the following SST set of matchups: games between two teams seeded  $\ell + 1$  or higher are coin flips, games between  $t_\ell$  and teams seeded between  $t_{\ell+1}$  and  $t_k$  are coin flips, and all other games are always won by the higher seed.

Let's calculate the probability of  $t_i$  and  $t_j$  winning the tournament.  $t_i$  will auto-win any games the have prior to round  $s + 1$ , and then have to win  $r - s$  coin flips in order to win the tournament. This happens with probability  $0.5^{r-s}$ .

$t_j$  will also win all of its games prior to round  $s + 1$ , also has to win a coin flip for each games in round  $s + 2$  or later. For round  $s + 1$ , however, half the time  $t_j$  will be matched up with  $t_k$ , which is a coin flip, but half the time they will be matched with  $t_k$ , which is an auto-win. Thus,  $t_j$  will win the tournament with probability  $0.75(0.5)^{r-s-1}$ .

Therefore,  $t_j$  has a better chance of winning the tournament than  $t_i$  does, despite  $t_i$  being higher seeded, so  $\mathcal{A}$  is not ordered, completing the contradiction. Thus if two games are played in round  $s$ , the winners of those games must play each other in round  $s + 1$ .

This immediately implies that at most two games can be played in each round of an ordered bracket.

Finally, we know that this can be the only game played in round  $s + 1$  because if there was another game, it would have to be between two teams that are higher-seeded than the two teams who won in the previous round, meaning that round violated condition (2).  $\square$

And now the main theorem,

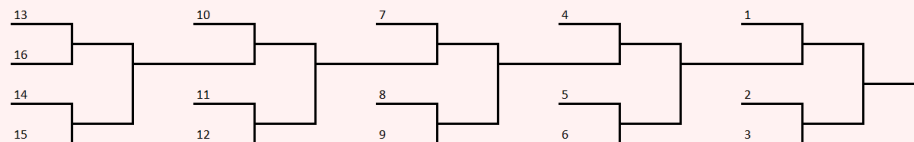
### **Theorem 2.5.6: Edwards's Theorem**

The only ordered brackets are those described by Corollary 2.5.4.

*Proof.* Corollary 2.5.4 describes all proper brackets in which each round either has only game, or has two games but is immediately followed by a round with only one game. A proper bracket not described by Corollary 2.5.4 would thus have to include a round with two or more games followed by another round with two or more games, violating Lemma 2.5.5.  $\square$

Edwards’s Theorem is both exciting and disappointing. On one hand, it means that we can fully describes the set of ordered brackets, making it easy to check whether a given bracket is ordered or not. On the other hand, it means that in an ordered bracket at most three teams can be introduced each round, so the length of the shortest ordered bracket on  $n$  teams grows linearly with  $n$  (rather than logarithmically as is the case for proper brackets). If we want a bracket on many teams to be ordered, we risk forcing lower-seeded teams to play large numbers of games, and we only permit the top seeded teams to play a few. For example, if the shortest ordered bracket that could’ve been used in the 2022 NCAA Women’s Basketball Wichita Region is  $[[4; 0; 3; 0; 3; 0; 3; 0; 3; 0; 0]]$ , which is played over a whopping ten rounds.

**Figure 2.5.7: The Shortest Sixteen-Team Ordered Bracket**



Because of this, few leagues use ordered brackets, and those who do usually have so few teams that every proper bracket is ordered (the 2023 College Football Playoffs, for example). Even the Korean Baseball Organization League, which uses a somewhat unconventional  $[[2; \mathbf{1}; \mathbf{1}; \mathbf{1}; \mathbf{0}]]$ , only sends five teams to the playoffs, and again every five-team proper bracket is ordered. If the KBO League ever expanded to the six-team bracket  $[[2; \mathbf{1}; \mathbf{1}; \mathbf{1}; \mathbf{1}; \mathbf{0}]]$ , we would have a case of an ordered bracket being used when a proper non-ordered bracket exists on the same number of teams.

### Open Question 2.5.8

Which brackets are ordered if  $\mathcal{T}$  is not SST?

Edward's Theorem assumes that that inputted list of teams are SST, and in particular, are in the correct order such that better teams are seeded higher. In real life, however, seeding is not that simple: if we knew who the best team was going into the tournament, we wouldn't need to conduct the tournament at all. Instead, seeding is often a guess as to the strengths of various teams based on the information known at the time, and in some cases, such as the 2023 Major League Baseball Playoffs, certain higher-seeded are generally agreed to be weaker than certain lower-seeded teams. Which brackets remain ordered under weaker assumptions than SST is still open.



### 3 Alternative Seedings

### 3.1 Reseeding

While Edwards’s Theorem is certainly a very powerful result, with a little clever engineering we can still construct a bracket-like tournament without requiring that the number of rounds grow linearly with the number of participating teams.

Ultimately, the reason that proper brackets are not, in general, ordered, is that lower-seeded teams are treated, if they win, as the team that they beat for the rest of the format. Consider again the proper bracket analyzed by Silver:  $[[\mathbf{16}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}]]$ . If an 11-seed wins in the first-round, they take on the schedule of a 6-seed for the rest of the tournament, while if the 9-seed wins, they take on the schedule of an 8-seed. Given that a 6-seed has an easier schedule than an 8-seed, it’s not hard to see why it might be preferable to be an 11-seed rather than a 9-seed.

*Reseeding* (poorly named) fixes this by resorting the match-ups every round: if an 11-seed keeps winning, they will have to play teams according to their seed, rather than getting an effective upgrade to 6-seed status.

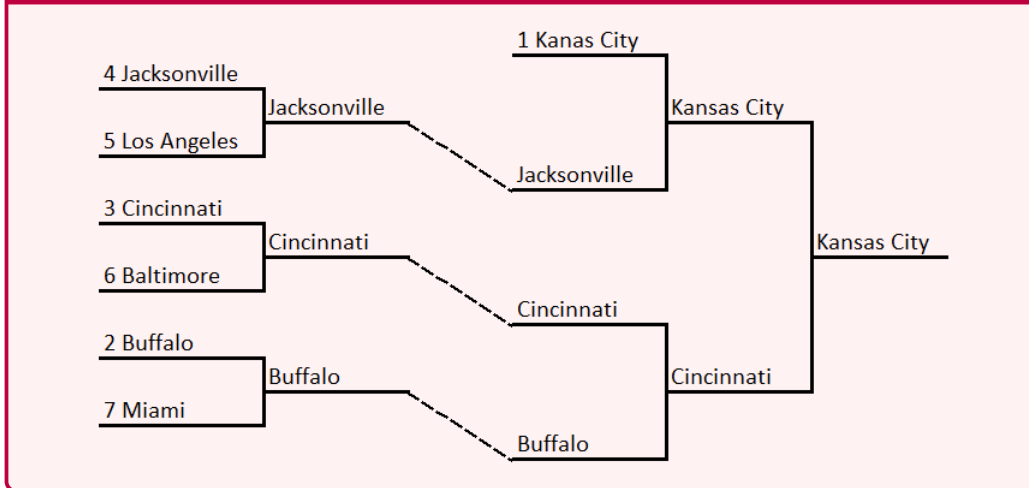
#### Definition 3.1.1: Reseeded Brackets

In a *reseeded* bracket, after each round, match-up the highest-seeded team with the lowest-seeded team, second highest vs second-lowest, etc.

Note that by Definition 2.1.1, a reseeded bracket is not a bracket at all, as matchups between teams that have not yet lost are not determined in advance of the outcomes of any games. However, because reseeded brackets act so similarly to traditional brackets, and because colloquially they are referred to as brackets, we opt to continue using the word “bracket” to describe them.

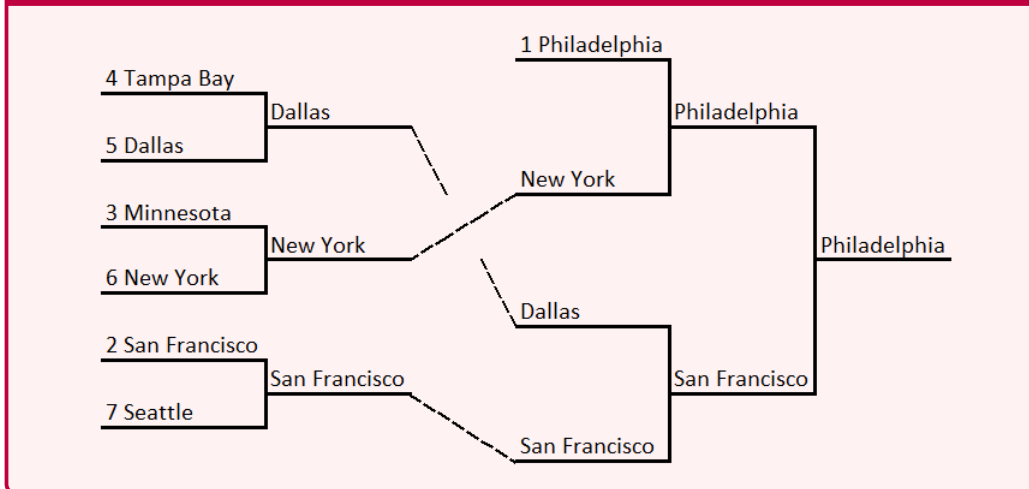
Both National Football League conferences use a reseeded bracket with signature  $[[\mathbf{6}; \mathbf{1}; \mathbf{0}; \mathbf{0}]]$ . If the first-round of the bracket goes chalk, then it looks just like a normal bracket:

**Figure 3.1.2: 2023 National Football League AFC Playoffs**



The dotted lines are drawn after the first round of games have been played: if there are some first-round upsets, then the bracket is rearranged to ensure that it is still better to be a higher seed rather than a lower seed.

**Figure 3.1.3: 2023 National Football League NFC Playoffs**



In the NFC, 6-seed New York upset 3-seed Minnesota. Had a conventional bracket been used, the semifinal matchups would have been 1-seed vs 5-seed and 2-seed vs 6-seed: the 2-seed would have had an easier draw than the 1-seed, while the 6-seed would have an easier draw than the 5-seed. Reseeding fixes this by matching 6-seed New York is with top-seed Philadelphia, and

2-seed San Francisco with 5-seed Dallas.

Reseeding is powerful technique. For one, the fundamental theorem still applies to reseeded brackets.

#### Theorem 3.1.4

There is exactly one proper reseeded bracket with each bracket signature.

*Proof.* The definition of properness ensures that there is only one way byes can be distributed such that a reseeded bracket can be proper. Additionally, because reseeded brackets have no additional parameters beyond which seeds get how many byes, there is no more than one reseeded bracket with each signature that could be proper. Finally, that bracket is indeed proper: if the bracket goes to chalk, the matchups will be the exact same as a traditional bracket, which by the fundamental theorem is a proper set of matchups.  $\square$

And, more excitingly, Edward's theorem does *not* apply (a fact first shown by Hwang [5]).

#### Lemma 3.1.5

If  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  is a proper reseeded bracket on  $n$  teams,  $\mathcal{T} + \{x, y\}$  is an SST list of  $n + 1$  teams,  $x$  and  $y$  are one seed apart in  $\mathcal{T} + \{x, y\}$  with  $x$  being the higher seed, and  $z \in \mathcal{T}$ , then

$$\mathbb{P}[z \text{ beats } y] \cdot \mathbb{W}_{\mathcal{A}}(z, \mathcal{T} + \{x\}) \geq \mathbb{P}[z \text{ beats } x] \cdot \mathbb{W}_{\mathcal{A}}(z, \mathcal{T} + \{y\}).$$

*Proof.*

$\square$

#### Theorem 3.1.6

All proper reseeded brackets are ordered.

*Proof.*

$\square$

However, reseeding is not without its drawbacks. In a reseeded bracket, teams and spectators alike don't know who they will play or where their next game will be until the entire previous round is complete. This can be

an especially big issue if parts of the bracket are being played in different locations on short turnarounds: in the 2022 NCAA Women’s Basketball Tournament, the first two rounds are played over a weekend on the various college campuses of the highest-seeded teams. It would cause problems if teams had to pack up and travel across the country because their opponent changed because of reseeding.

In addition, part of what makes the NCAA Basketball Tournament (affectionately known as “March Madness”) such a fun spectator experience is that fact that these matchups are known ahead of time. In “bracket pools,” groups of fans each fill out their own brackets, predicting who will win each game and getting points based on how many they get right. If it wasn’t clear where in the bracket the winner of a given game was supposed to go, this experience would be diminished.

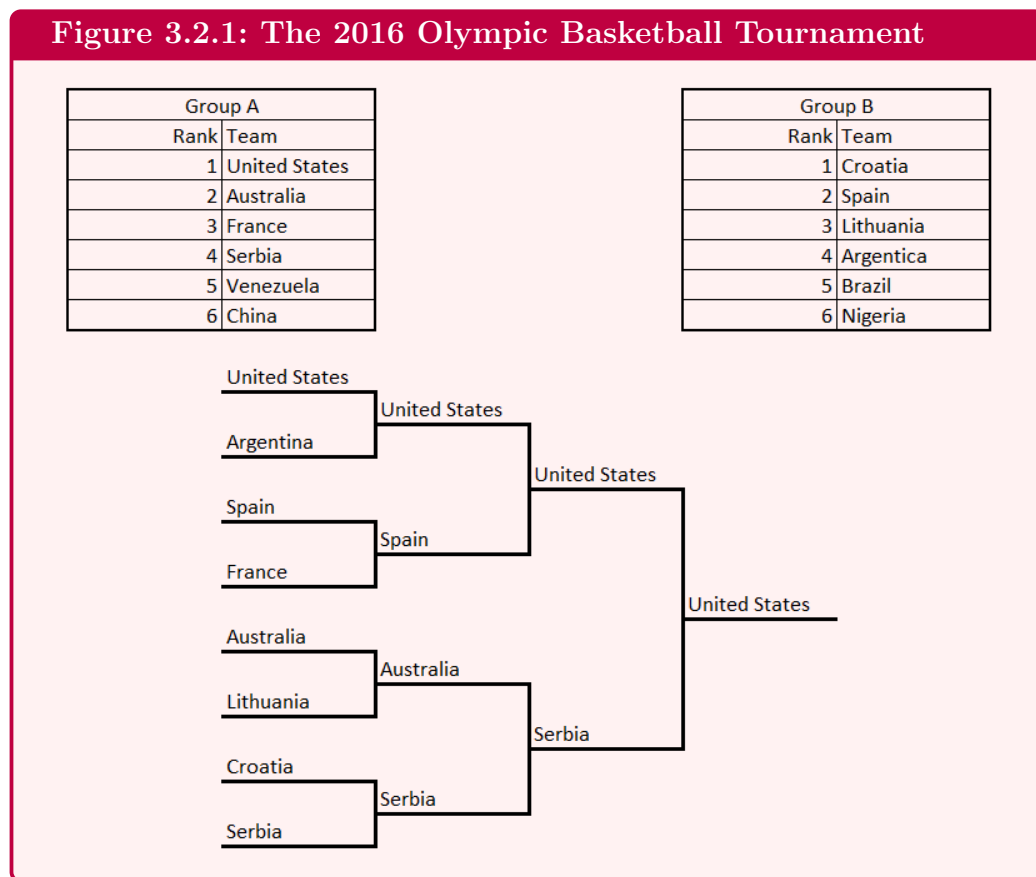
Finally, reseeding gives the top-seed(s) an even greater advantage than they already have: instead of playing against merely the *expected* lowest-seeded team(s) each round, they would get to play against the *actual* lowest-seeded team(s). In March Madness, “Cinderella Stories,” that is, deep runs by low seeds, would become much less common.

In many ways, the NFL conferences playoffs are a perfect place to use a reseeded bracket: games are played once a week, giving plenty of time for travel; only seven teams make the playoffs in each, so a huge March Madness-style bracket challenge is unlikely; as a professional league, the focus is far more on having the best team win and protecting Cinderella Stories isn’t as important; and because the bracket is only three rounds long, reseeding is only required once.

Other leagues with similar structures might consider adopting forms of reseeding to protect their incentives and competitive balance (looking at you, Major League Baseball), but in many cases, the traditional bracket structure is too appealing to adopt a reseeded one.

## 3.2 Tiered Seeding

Consider the 2016 Olympic Basketball Tournament. Twelve teams qualified for the Olympics, and they were divided into two groups of six teams each. Each group conducted a mini-tournament, ranking the teams in each group from first through sixth (the specifics of the mini-tournament are not relevant). Then, the bottom two teams in each group were eliminated, with the remaining eight teams (four from each group) entering the bracket [[8; 0; 0; 0]]. The entire format as it played out is displayed in Figure 3.2.1.



The seeding going into the bracket portion of the 2016 Olympic Basketball Tournament is a little different than the seedings that we have discussed so far. Rather than the ranking of the teams being a complete ordering, it is a partial one: teams are grouped into tiers, and the tiers are ranked. Two teams, one from each group, occupy each tier.

### Definition 3.2.2: Tiered Seeding

A *tiered seeding* is a partial ordering on the teams entering a tournament.

This is as opposed to traditional seedings, which are a complete ordering (although we can view a traditional seeding as a special example of a tiered seeding where each tier has a single team). When filling out a bracket using a tiered seeding, we continually assign the top remaining seeds to the teams in the top remaining tier. Recall the proper bracket  $[[8; 0; 0; 0]]$ :

Figure 3.2.3:  $[[8; 0; 0; 0]]$



The United States and Croatia, as the two teams in the top tier, are given seeds one and two. The two tier-two teams, Australia and Spain, get seeds three and four, and so on. The actual algorithm used for assigning the seeds to the teams within each tier can be arbitrary: in the particular case of the 2016 Olympic Basketball Tournament, teams from Group A were given the odd seeds and teams from Group B the evens.

We can describe a tiered seeding with a list of integers indicating how many teams are in each tier. The eight teams that advanced to bracket at the Olympics were divided into four pools of two teams each, so we write  $(2, 2, 2, 2)$ . A quick notational note: we list tier sizes in reverse order, with the size of the lowest tier coming first, and the size of the top tier coming

last. This is done to keep it consistent with bracket signatures, in which the lower-seeded teams are listed earlier, and higher-seeded teams that get more byes are listed later.

The tiered seeding  $(\mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2})$  interacts very nicely with the proper bracket  $[[\mathbf{8}; \mathbf{0}; \mathbf{0}; \mathbf{0}]]$ : there is no advantage for a team being assigned a particular seed within their tier.

#### Definition 3.2.4: Strongly Respectful

A bracket *strongly respects* a tiered seeding if, as long as teams' win probabilities are defined only by what tier they are in (that is, for all  $t_i, t_j$  that share a tier and  $t_k, t_\ell$  that share a tier,  $\mathbb{P}[t_i \text{ beats } t_k] = \mathbb{P}[t_j \text{ beats } t_\ell]$ ), then teams in the same tier have the same probability of winning the tournament.

Sometimes, it is not possible to generate a bracket that strongly respects a tiered seeding (for example, the tiered seeding  $(\mathbf{3}, \mathbf{1})$ ), so we also introduce the concept of a bracket weakly respecting a tiered seeding.

#### Definition 3.2.5: Weakly Respectful

A bracket *weakly respects* a tiered seeding if each team in a tier is given the same number of byes.

Although no signature strongly respects the tiered seeding  $(\mathbf{3}, \mathbf{1})$ , the bracket  $[[\mathbf{4}; \mathbf{0}; \mathbf{0}]]$  is preferable to  $[[\mathbf{2}; \mathbf{1}; \mathbf{1}; \mathbf{0}]]$  because at least it weakly respects it. The names of the two conditions come from strong respectfulness being a stronger condition than weak respectfulness.

#### Theorem 3.2.6

If a bracket strongly respects a tiered seeding then it weakly respects it as well.

*Proof.* If a bracket strongly respects a tiered seeding, then all teams within the same tier must have the same probability of winning the tournament if every game is a coin flip. If indeed every game is coinflip, two teams have the same chance of winning the tournament only if they have the same number of byes, so the bracket must weakly respect the



tiered seeding as well. □

Finally, weak and strong respectfulness give names to the intuition that teams in the same tier should be treated similarly. There is a separate intuition that teams in higher tiers should be treated better than teams in lower tiers. To that end, we introduce the notion of *proper respectfulness*.

### Definition 3.2.7: Properly Respectful

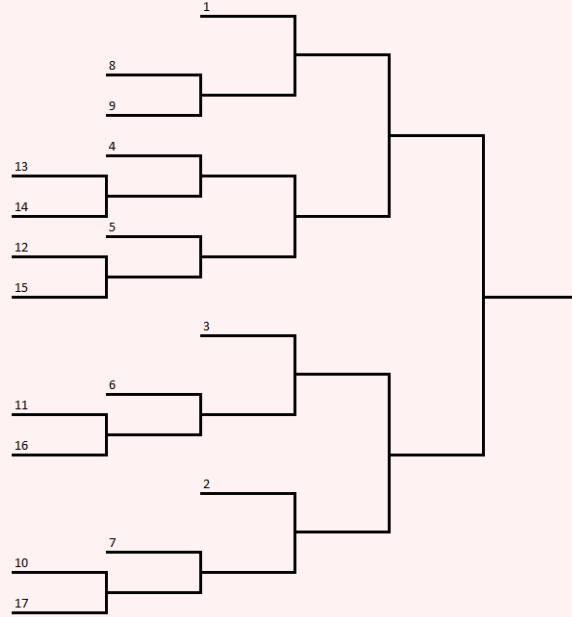
A bracket *properly respects* a tiered seeding  $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_m)$  if the bracket is proper and teams in tier  $i$  are given seeds  $1 + \sum_{j=1}^{i-1} b_j$  through  $\sum_{j=1}^i b_j$ .

If a bracket is not properly respectful, then lower tiers might grant more byes or an easier road to winning the tournament than higher tiers. The 2016 Olympic Basketball Bracket is properly respectful, as well as weakly and strongly respectful.

How can we tell whether an arbitrary bracket strongly, weakly, or properly respects an arbitrary tiered seeding? Proper respectfulness is the easiest: just check if the bracket is proper and then if the tiers are being assigned to the right seeds.

Checking if a bracket signature weakly respects a tiered seeding is also somewhat straightforward: we can simply matchup the signature with the tiered seeding to see if we ever have to split a tier across two different levels in the signature. For example, consider the bracket signature  $\mathcal{A} = [[8; 6; 3; 0; 0; 0]]$  and the tiered seeding  $(4, 4, 4, 4, 1)$ . The four teams in the second-highest tier are distributed over two levels: two of them (seeds 2 and 3) get two byes and two of them (seeds 4 and 5) get a single bye, so  $[[8; 6; 3; 0; 0; 0]]$  does not weakly respect  $(4, 4, 4, 4, 1)$ .

Figure 3.2.8:  $[[8; 6; 3; 0; 0; 0]]$



If a bracket signature *does* weakly respect a tiered seeding, we can combine the information of the bracket signature and the tiered seeding into a single list of lists called the *tiered signature*.

### Definition 3.2.9: Tiered Signature

If a bracket signature  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  weakly respects a tiered signature  $\mathcal{B}$ , then the *tiered signature* of the signature-seeding pair  $(\mathcal{A}, \mathcal{B})$  is a list  $\mathcal{C} = [[\mathcal{C}_0; \dots; \mathcal{C}_r]]$  where  $\mathcal{C}_i$  is the sublist of  $\mathcal{B}$  corresponding to the  $a_i$  teams that get  $i$  byes.

The bracket  $\mathcal{A} = [[8; 6; 3; 0; 0; 0]]$  weakly respects the tiered seeding  $\mathcal{B} = (4, 4, 4, 2, 2, 1)$ , and the associated tiered signature of this pair is

$$\mathcal{C} = [[(4, 4); (4, 2); (2, 1); (); (); ()].$$

The somewhat trivial tiered signature of the 2016 Olympic Basketball Tournament is

$$[[ (2, 2, 2, 2); (); (); ()].$$

Note that we can easily extract both the bracket signature and the tiered seeding from the tiered signature. For the former, sum each sublist, and for the latter, concatenate the sublists into a single list. Sometimes, we will refer a tiered signature as being strongly respectful as a shorthand for saying that the associated tiered seeding respects the associated bracket signature.

Checking for strong respectfulness seems to be much trickier than weak respectfulness. Somehow, we need to be able to verify that for any distribution of win probabilities, (as long as teams within the same tier have the same matchup table,) teams within the same tier have the same probability of winning the tournament. Luckily, there is a simple algorithm for doing just that, which we will explore in the next section.

### 3.3 The Palandromic Algorithm

In this section, we present an algorithm for verifying whether a bracket signature strongly respects a tiered seeding. We will first intuitively describe what the algorithm is doing and then we will formally specify it, before running the algorithm on a few examples and then finally proving its correctness.

The idea behind the algorithm is to ensure that in each round, teams of the same tier are being assigned opponents of the same tier. This is done by keeping track of the tiers of the teams that will be playing in each round, and ensuring that the round-specific tiered signatures are palandromic. Formally,

#### Definition 3.3.1: The Palandromic Algorithm for Tiered Signatures

Let  $\mathcal{A}$  be a bracket signature and  $\mathcal{B}$  be a tiered seeding. First, check if  $\mathcal{A}$  weakly respects  $\mathcal{B}$ . If it doesn't, then it certainly doesn't strongly respect it. If it does, then let  $\mathcal{C}$  be the tiered signature of  $(\mathcal{A}, \mathcal{B})$ .

We define  $\mathcal{F}$ , a recursive operator that maps a tiered signature to either **true** or **false**. Then, if  $\mathcal{F}(\mathcal{C})$  is true,  $\mathcal{A}$  strongly respects  $\mathcal{B}$ , otherwise it does not.

The operator  $\mathcal{F}$  is defined in the following way on  $\mathcal{C} = [[\mathcal{C}_0; \dots; \mathcal{C}_r]]$ .

- If  $r = 0$ , then  $\mathcal{F}(\mathcal{C})$  is **true**.
- Otherwise, if  $\mathcal{C}_0$  is not palandromic, then  $\mathcal{F}(\mathcal{C})$  is **false**.
- Otherwise, let  $\mathcal{D}_0$  be the right half of  $\mathcal{C}_0$  concatenated with  $\mathcal{C}_1$ , and  $\mathcal{D} = \mathcal{F}([\mathcal{D}_0; \mathcal{C}_2; \dots; \mathcal{C}_r])$ . Then,  $\mathcal{F}(\mathcal{C}) = \mathcal{F}(\mathcal{D})$ .

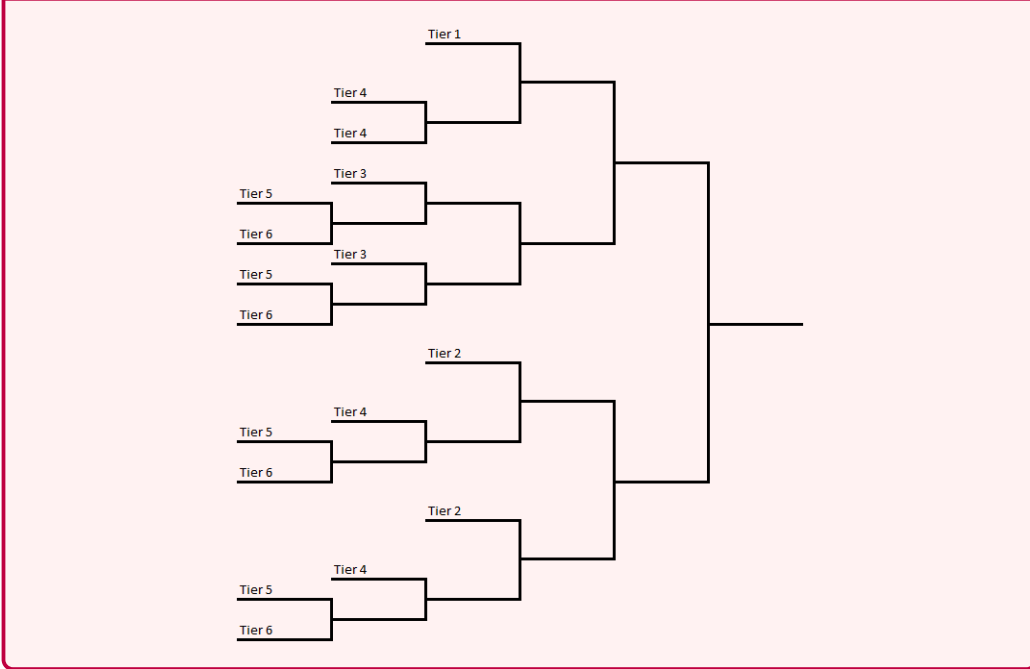
(For the last step, if  $\mathcal{C}_0$  has odd length, then the first element of  $\mathcal{D}_0$  is half of the middle element of  $\mathcal{C}_0$ . The middle element of  $\mathcal{C}_0$  will always be even because it is palandromic and its sum must be even.)

Let's go over a few examples. Consider the bracket signature  $\mathcal{A} = [[8; 6; 3; 0; 0; 0]]$  along with the tiered seeding  $\mathcal{B} = (4, 4, 4, 2, 2, 1)$ . As we verified earlier,  $\mathcal{A}$  weakly respects  $\mathcal{B}$ , so we can apply the palandromic algo-

rithm to check if it is strongly respectful.

$$\begin{aligned}
\mathcal{F}(\mathcal{C}) &= \mathcal{F}([(4, 4); (4, 2); (2, 1); (); (); ()]) \\
&= \mathcal{F}([(4, 4, 2); (2, 1); (); (); ()]) \\
&= \text{false} \text{ (because } (4, 4, 2) \text{ is not palandromic.)}
\end{aligned}$$

**Figure 3.3.2:**  $[(4, 4); (4, 2); (2, 1); (); (); ()]$



We can verify this result intuitively with the help of the bracket  $\mathcal{A}$ . In the second round, for example, two of the Tier 4 teams play each other, while two of them play the winner of a Tier 5 vs Tier 6 matchup. If the Tier 5 and 6 teams are much worse than the rest of the teams, it is not hard to imagine that the two Tier 4 teams who have to play each other are at a severe disadvantage.

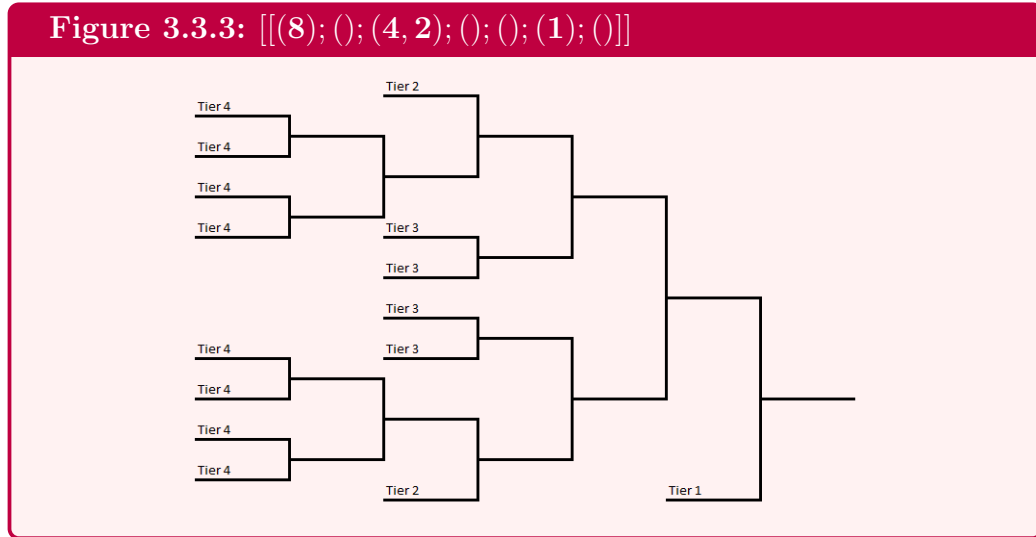
Let's instead consider the bracket signature  $\mathcal{A} = [[8; 0; 6; 0; 0; 1; 0]]$  along with the tiered seeding  $\mathcal{B} = (8, 4, 2, 1)$ .  $\mathcal{A}$  weakly respects  $\mathcal{B}$  with tiered signature

$$\mathcal{C} = [(8); (); (4, 2); (); (); (1); ()]$$

Applying the palandromic algorithm,

$$\begin{aligned}
\mathcal{F}(\mathcal{C}) &= \mathcal{F}([[(8); (); (4, 2); (); (); (1); ())]) \\
&= \mathcal{F}([[(4); (4, 2); (); (); (1); ())]) \\
&= \mathcal{F}([[(2, 4, 2); (); (); (1); ())]) \\
&= \mathcal{F}([[(2, 2); (); (1); ())]) \\
&= \mathcal{F}([[(2); (1); ())]) \\
&= \mathcal{F}([[(1, 1); ())]) \\
&= \mathcal{F}([[(1)]]]) \\
&= \mathbf{true}
\end{aligned}$$

So  $\mathcal{A}$  does strongly respect  $\mathcal{B}$ . This can also be seen intuitively by looking at the bracket: teams in each tier have the same exact path throughout the tournament.



Finally, we leave as an exercise to the reader to use the palandromic algorithm to show that the 2016 Olympic Basketball Tournament was strongly respectful.

Hopefully, these three examples have given a sense as to why the palandromic algorithm accurately ascertains whether a bracket signature strongly respects a tiered seeding. We will prove it by induction.

### Theorem 3.3.4

The palandromic algorithm correctly ascertains whether a bracket signature strongly respects a tiered seeding.

*Proof.* Let  $\mathcal{A}$  be a bracket signature with  $r$  rounds and  $\mathcal{B}$  be a tiered seeding. If  $\mathcal{A}$  doesn't weakly respect  $\mathcal{B}$ , then the palandromic algorithm will correctly say that it doesn't strongly respect  $\mathcal{B}$  either. Assume then that  $\mathcal{A}$  does weakly respect  $\mathcal{B}$ , where  $\mathcal{C} = [[\mathcal{C}_0; \dots; \mathcal{C}_r]]$  is the tiered signature of the pair  $(\mathcal{A}, \mathcal{B})$ .

We proceed by induction on  $r$ . If  $r = 0$ , then  $\mathcal{A} = [[1]]$ ,  $\mathcal{B} = (1)$ , and  $\mathcal{C} = [[(1)]]$ . The palandromic algorithm will correctly claim that  $\mathcal{A}$  strongly respects  $\mathcal{B}$  without any recursive calls.

For any other  $r$ , we will show that the palandromic algorithm returns **false** if and only if  $\mathcal{A}$  does not strongly respect  $\mathcal{B}$ .

Assume first that  $A$  does not strongly respect  $B$ . Then, for some tier, either teams in that tier are not all equally likely to make it out of the first round, or they are not all equally likely to win the bracket, conditional on having made it out of the first round. In the former case, this would be caused by teams in the same tier having first-round matchups in different tiers, meaning  $\mathcal{C}_0$  would not be palandromic, and so the palandromic algorithm would fail on its first iteration. In the latter case, this would imply that  $\mathcal{D} = [[\mathcal{D}_0; \mathcal{C}_2; \dots; \mathcal{C}_r]]$  is not a strongly respectful tiered signature, (where  $\mathcal{D}_0$  is the right half of  $\mathcal{C}_0$  concatenated with  $\mathcal{C}_1$ ), so by induction,  $\mathcal{F}(\mathcal{C}) = \mathcal{F}(\mathcal{D}) = \mathbf{false}$ . In either case, the palandromic algorithm correctly identifies that  $A$  does not strongly respect  $B$ .

Now, assume that the palandromic algorithm returns **false**. If it did so in the first iteration, then that means that there are two tiers  $T_0, T_1$  for which some but not all teams in  $T_0$  are matched up in the first-round against teams in  $T_1$ . Consider a list of teams such that teams in  $T_1$  always lose, and all other games are coin-flips. Then, the teams in  $T_0$  matched up against  $T_1$  teams in the first-round will

win the tournament with probability  $(0.5)^{r-1}$ , while the teams that are not will win with probability  $0.5^r$ , so  $\mathcal{A}$  does not strongly respect  $\mathcal{B}$ .

Meanwhile, if the palandromic algorithm failed at a later iteration, then by induction,  $\mathcal{D} = [[\mathcal{D}_0; \mathcal{C}_2; \dots; \mathcal{C}_r]]$  is not a strongly respectful tiered signature, (where  $\mathcal{D}_0$  is the right half of  $\mathcal{C}_0$  concatenated with  $\mathcal{C}_1$ ). However, if we consider a set of teams such that all of the first-round matchups in  $\mathcal{C}$  are guaranteed wins for the higher tier, then a team's probability of winning the entire bracket (as long as they are in a tier that will win in the first-round) is the same as their probability of winning  $\mathcal{D}$ . Because  $\mathcal{D}$  is not a strongly respectful tiered signature, some teams in the same tier have different tournament-win probabilities, so  $\mathcal{C}$  is also not strongly respectful. Thus,  $\mathcal{A}$  does not strongly respect  $\mathcal{B}$ .

So by induction, the palandromic algorithm claims that a bracket signature strongly respects a tiered seeding if and only if it truly does so.

□

With the palandromic algorithm in our back pocket, we can now quickly identify the relation between a given bracket signature and tiered seeding: whether it is strongly, weakly, properly, or not at all respectful. The concept of tiered seedings will show up in a few different places down the line: tiers are a powerful and generalizable tool for understanding tournament formats from Wimbledon to the NCAA Softball Tournament to the World Cup, as we shall investigate in the coming sections.



## 4 Multibrackets

## 4.1 Simple Multibrackets

While the bracket is a very powerful and important tournament design, it has two intimately related shortcomings. First, brackets lead teams to play wildly different numbers of games: in the bracket  $[[8; 0; 0; 0]]$ , for example, two teams will play three game, two teams will play two, and four teams will play only one.

And second, brackets tend to do a very poor job of ranking teams beyond just selecting a winner. Again considering the bracket  $[[8; 0; 0; 0]]$ , the first-place finish is obviously the winner of the bracket, and we can easily grant the loser of the championship game second place, but the two semifinals losers both might have a claim to the third place, and sorting through the fifth-through eight-place finishes is even trickier.

These problems are reflections of each other: the reason that ranking the lower-placing teams is so hard is because they play so few games. It's easier to differentiate between first and second because both teams have played three games, but differentiating between the four teams that have played only a single game is nigh impossible.

In some cases, these problems do not cause concern: perhaps we are only interested in crowning a champion and don't care about exactly who came in third, or maybe this bracket is being played at the conclusion of a long season and so teams playing variable numbers of games is not a big deal. But the interconnected nature of the two problems lets us solve them together, leveraging the extra games that lower-ranked teams have left in order to rank them.

To do this, we consider class of formats called *simple multibrackets*.

### Definition 4.1.1: Simple Multibracket

A *simple multibracket* is a sequence of brackets in which all teams begin in the first bracket, and then the losers of certain games in the upper brackets fall into the lower brackets rather than being eliminated outright, and teams place based on which bracket they won.

### Definition 4.1.2: Primary Bracket

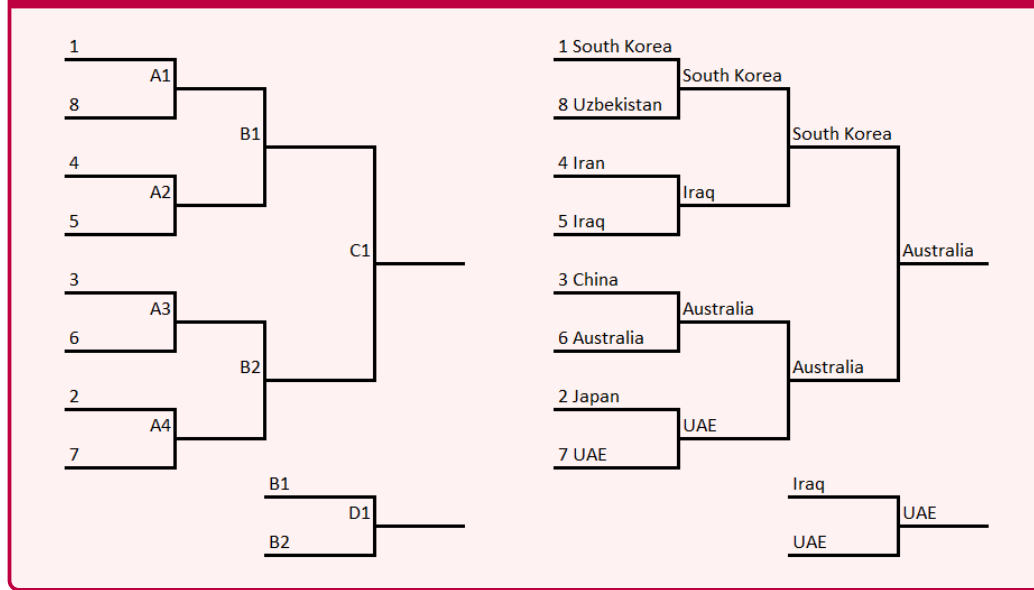
The first bracket in a simple multibracket is called the *primary bracket*.

### Definition 4.1.3: Order of a Simple Multibracket

The *order* of a Simple Multibracket is the number of brackets it contains.

Multibrackets of order one are just traditional brackets. The simplest example of a multibracket with order greater than one is the third-place game, in which the losers of the two semifinal games play each other for third. The 2015 Asian Football Confederation Asian Cup, whose bracket is of signature  $[[8; 0; 0; 0]]$ , employs a third-place game.

Figure 4.1.4: 2015 AFL Asian Cup



Each game in this figure is labeled. In the primary bracket, first-round games are **A1** through **A4**, while the semifinals are **B1** and **B2**, and the finals is game **C1**. The labeling system works as follows: each round in the first bracket is given a letter, and then each round in the second bracket, etc. Then games are assigned a number from the top of the bracket to the bottom within each round.

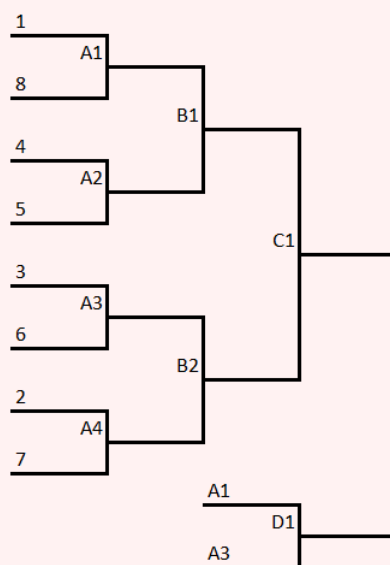
Thus, the third-place game is labeled **D1**: even though it could be played concurrently to the championship game, it is part of a different bracket and so we label it as a different round.

We indicate that the third-place game is to be played in between the

losers of games **B1** and **B2** by labeling the starting lines in the third-place game with those games. This is not ambiguous because the winner of those games always continue on in the original bracket, so such labels only refer to the losers.

In theory, the Asian Cup could have had any two teams play for third.

**Figure 4.1.5: 2015 AFL Asian Cup Alternative Format**



In this alternative, the third-place game is played between the loser of **A1** and **A3** instead of between the two semifinal losers. This is probably a bad idea: the losers of **B1** and **B2** each made it further in the main bracket than the losers of **A1** and **A3** did, and so they ought to be the teams playing the third-place game.

This matches very closely with the notion of a bracket being properly respectful of a tiered seeding:

#### **Definition 4.1.6: The Multibracket Tiered Seeding**

The *the multibracket tiered seeding* is a tiered seeding that groups the teams participating in a multibracket by the round of their most recent loss, with letters later in the alphabet corresponding to higher tiers.

#### Definition 4.1.7: Proper Multibracket

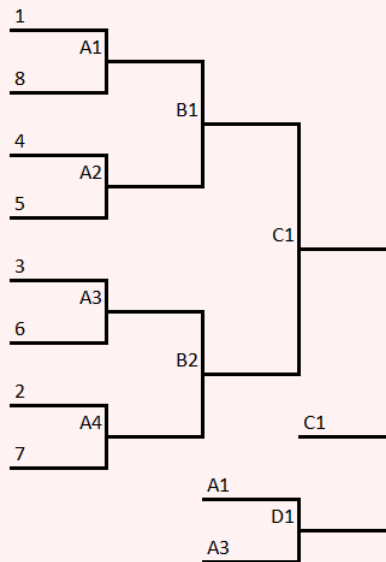
A multibracket is *proper* if its primary bracket is proper and the rest of its brackets are each properly respect the multibracket tiered seeding.

Since the losers of **B1** and **B2** lost more recently and so are in a higher tier than any of the losers of **A1** through **A4**, if the Asian Cup wants to be proper, its third-place game must be played between the losers of **B1** and **B2**.

Note that as a multibracket of order two, the Asian Cup is actually not quite proper: remember that a team's final place is based on which bracket it wins. Thus, in order for the Asian Cup to have a top-three, it needs to have order three. Otherwise, the winner of game **D1** would come in second place, leaving the loser of **C1** out in the cold and violating the properness condition.

In reality there is an implied one-team bracket in between the primary bracket and the third-place game. Rigorously, the Asian Cup multibracket is proper with order three and looks like this:

Figure 4.1.8: 2015 AFL Asian Cup, Rigorously



However for clarity, when drawing figures, we will omit brackets in a

simple multibracket of signature  $[[1]]$ . Much like the case with traditional brackets, we will almost always focus our study on *proper* multibrackets: unless stated otherwise, it is safe to assume that all multibrackets discussed are proper. Thus in a figure, if the loser of the championship of a bracket doesn't appear in the next bracket, there is an implied bracket of signature  $[[1]]$  between them.

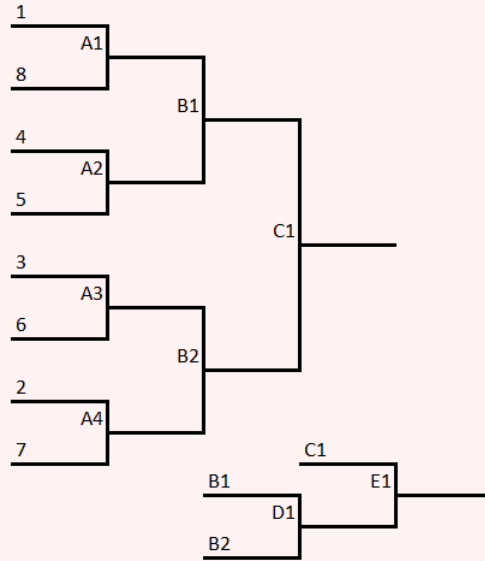
#### Definition 4.1.9: Multibracket Signature

The *signature* of a multibracket is simply the list of signatures of its brackets.

So the 2015 AFL Asian Cup has signature  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 0]]$ . But the multibracket with this signature is far from the only multibracket that the AFL could have used to give out gold, silver, and bronze. In fact, it's not clear the loser of **C1**, who comes in second place, is really more deserving than the winner of **D1**, who comes in third. One could imagine the UAE saying: we and South Korea both finished with two wins and one loss – a first round win, a win against Iraq, and a loss against Australia. The only reason that South Korea came in second and we came in third was because South Korea lucked out by having Australia on the other half of the bracket as them. That's not fair!

If the AFL took this complaint seriously, they could modify their format to have signature  $[[8; 0; 0; 0]] \rightarrow [[2; 1; 0]] \rightarrow [[1]]$ . (Again, the bracket of signature  $[[1]]$  is implied and left out of the figure.)

Figure 4.1.10:  $[[8; 0; 0; 0]] \rightarrow [[2; 1; 0]] \rightarrow [[1]]$



If the AFL used the format in Figure 4.1.10 in 2015, then South Korea and the UAE would have played each other for second place after all of the other games were completed. In some sense, this is a more equitable format than the one used in reality: we have the same data about the UAE and South Korea and so we ought to let them play for second place instead of having decided almost randomly.

However, swapping formats doesn't come without costs. For one thing, South Korea and the UAE would've had to play a fourth game: if the AFL had only three days to put on the tournament and teams can play at most one game a day, then the format in Figure 4.1.10 isn't feasible.

Another concern: what if Iraq had beaten the UAE when they played in game **D1**? Then the two teams with a claim to second place would have been South Korea and Iraq, except South Korea already beat Iraq! In this world, South Korea being given second place without having to win a rematch with Iraq seems more equitable than giving Iraq a second chance to win. To address this, one could imagine a format in which game **E1** is played only if it is not a rematch, although this would no longer be a multibracket and is a bit out of scope.

Ultimately, whether including game **E1** is worth it depends on the goal of the format. If there is a huge difference between the prizes for coming in

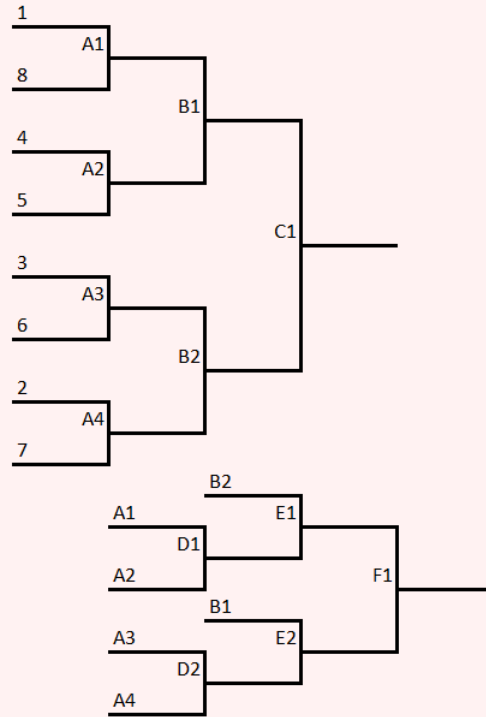
second and third, for example, if the top two finishing teams in the Asian Cup qualified for the World Cup, then **E1** is quite important. If, on the other hand, this is self-contained format played purely for bragging rights, **E1** could probably be left out. In reality, the 2015 AFL Asian Cup qualified only its winner to another tournament (the 2017 Confederations Cup), and gave medals to its top three, and so game **E1**, which distinguishes between second and third place, is probably unnecessary.

Let's imagine, however, that instead of just the champion, the top four teams from the Asian Cup advanced to the Confederations Cup. In this case, the format used in 2015 would be quite poor, as teams finish in the top-four based only on the result of their first-round game: the rest of the games don't even have to be played. (Formally, the simple multibracket  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 0]]$  has order three and so doesn't even assign a fourth-place, but it could easily be extended to the following multibracket of order four  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 0]] \rightarrow [[1]]$ , which has the property mentioned above.)

A better format for selecting the top four teams might look like this:



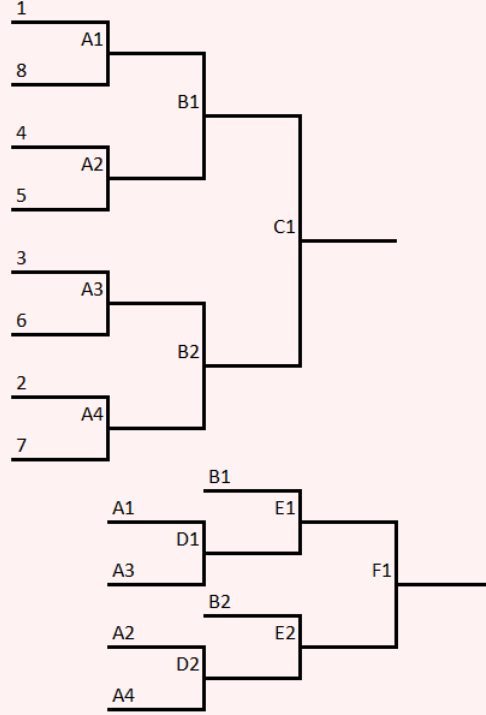
Figure 4.1.11:  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$



The simple multibracket in Figure 4.1.11 selects a top-four without having the selection be completely determined by the first-round games. In fact,  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$  has the attractive property that a team will finish in the top four if and only if it wins two of its first three games.

A few notes about what we've seen so far. First, note that the fundamental theorem of brackets doesn't apply to multibrackets: the signature of a proper multibracket does not uniquely determine it. For example, here is another proper multibracket with the same signature.

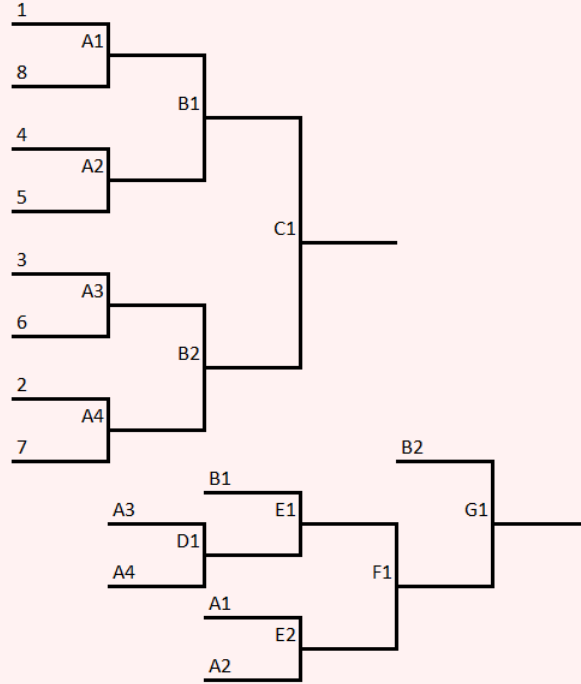
**Figure 4.1.12: Alternative  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$**



We generally prefer the format in Figure 4.1.11 to the one in Figure 4.1.12 because it ensures that no rematches can occur until game **F1**, but both are proper. (Further discussion of rematches in multibrackets can be found in later sections.)

A second thing of note is that  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$  is not the only simple multibracket signature of order four whose primary bracket is  $[[8; 0; 0; 0]]$ . Another option might be  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 3; 0; 1; 0]] \rightarrow [[1]]$ .

Figure 4.1.13:  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 3; 0; 1; 0]] \rightarrow [[1]]$



The bracket in Figure 4.1.13 is proper: both (or technically, all four) of its brackets are proper, and teams that lose in later rounds are given better spots in later brackets than teams that lost earlier. However it still feels unfair in some sense: the losers of **B1** and **B2** lost in the same round and so they ought to be on the same foot. Instead the loser of **B1** must win three games to ensure a top-three finish, while the loser of **B2** comes in fourth even if they lose their next game. Once again, we can use the already-established language of tiered seeding to express this.

#### Definition 4.1.14: Weakly Respectful Simple Multibrackets

A simple multibracket is *weakly respectful* if each of its brackets weakly respects the multibracket tiered seeding.

#### Definition 4.1.15: Strongly Respectful Simple Multibrackets

A simple multibracket is *strongly respectful* if each of its brackets strongly respects the multibracket tiered seeding.

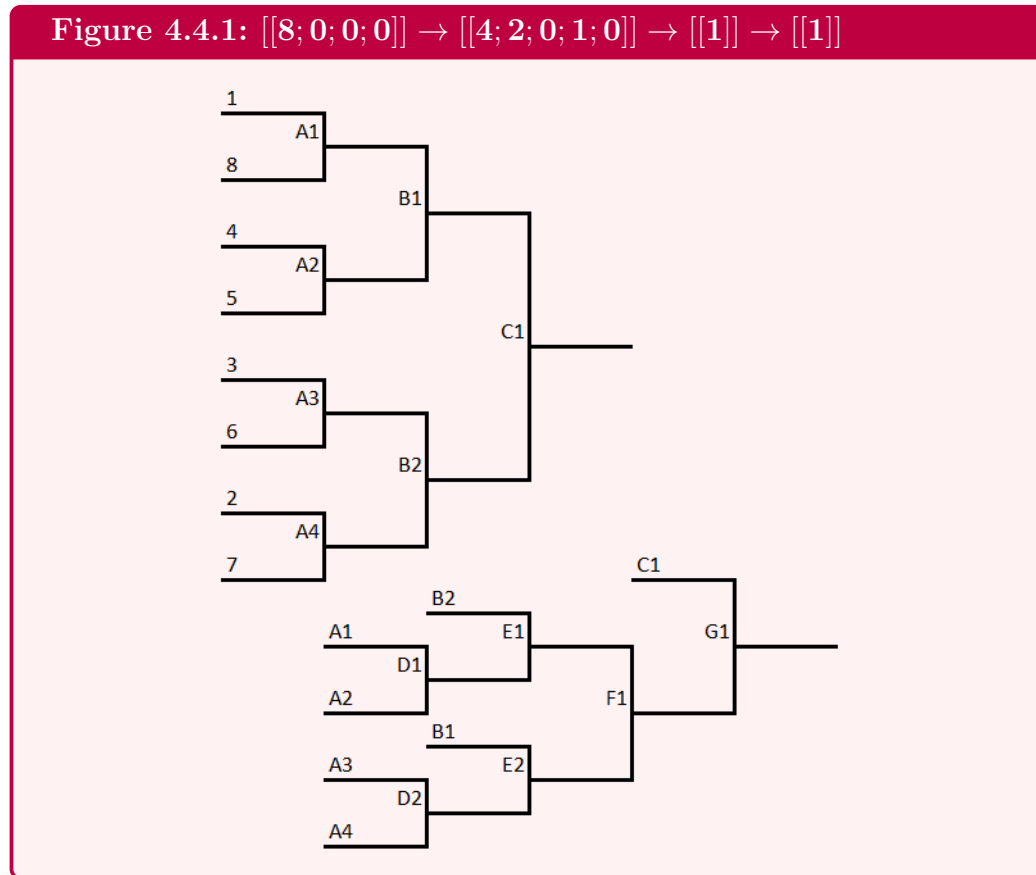
Thus, the multibrackets of signature  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[4; 2; 0; 0]] \rightarrow [[1]]$  are strongly respectful, while ones of signature  $[[8; 0; 0; 0]] \rightarrow [[1]] \rightarrow [[2; 3; 0; 1; 0]] \rightarrow [[1]]$  are not respectful at all. As general rule, we prefer multibrackets that are more respectful over ones that are less so, although there can be other factors that might convince us to use less respectful multibrackets, and under certain circumstances no strongly or weakly respectful brackets will be available.

## 4.2 Properness

### 4.3 Multiple Elimination

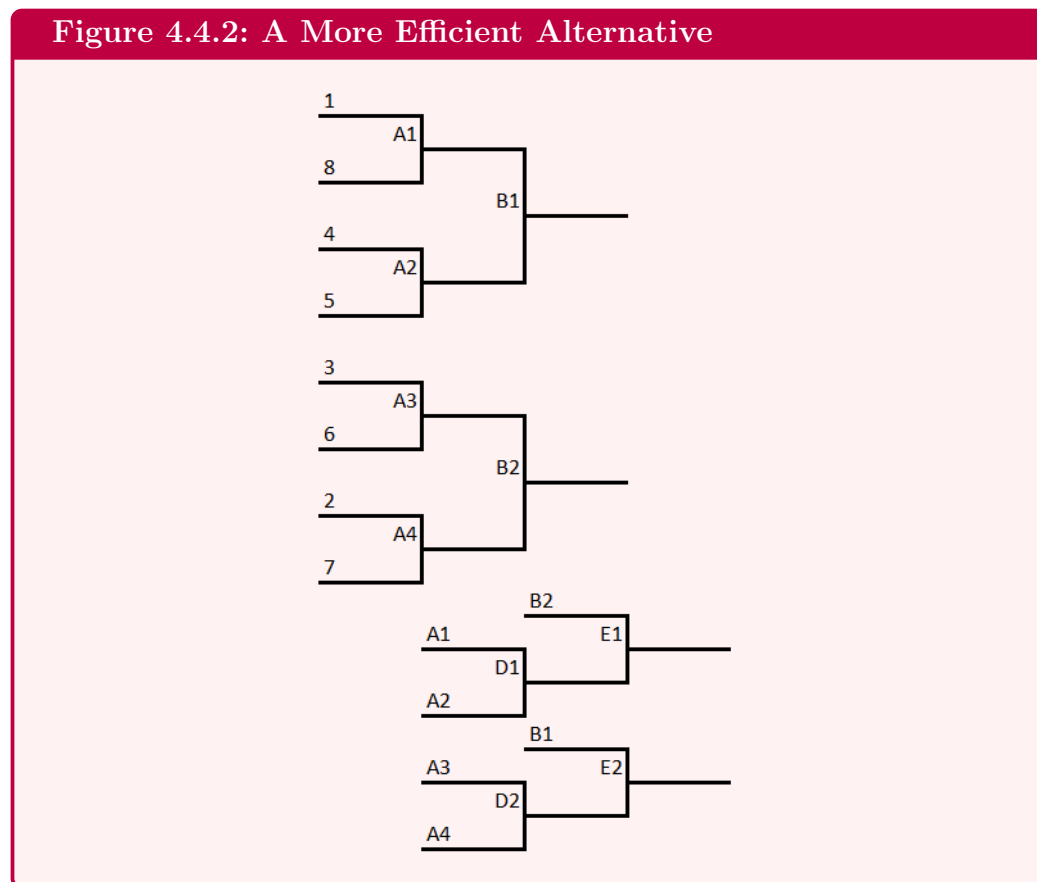
## 4.4 Semibrackets

One common application of multibrackets is when an  $n$ -team tournament is needed to select the top  $m$  teams to move on to the next stage of the competitions (perhaps this is the regional tournament, and the top  $m$  teams from this region qualify for nationals.) For a tournament such as this, a natural option would be an  $m$ -team multibracket of order  $n$ . This is exactly what the USA Ultimate Manual of Championship Series Tournament Formats [7] calls for in many circumstances. One such example is the following eight-team order four simple multibracket:



One observation about this format is that a few of the games played seem unnecessary. In particular, games **F1** and **G1** are downright silly: after round **E**, the four teams that advance to nationals are already set: the winner and loser of **C1**, and the winners of the two round **E** games. Perhaps more

subtly, game **C1** is also unnecessary: both the winner and loser will advance. A more efficient alternative might look like so.



The format in Figure 4.4.2 isn’t an example of anything we’ve defined thus far. It’s almost a simple multibracket but not quite: for one thing, some teams start in the “second bracket.” To formally describe what is going in Figure 4.4.2, we introduce the notion of semibracket.

**Definition 4.4.3: Semibracket**

A *semibracket* is a tournament format in which:

- Teams don’t play any games after their first loss,
- The matchups between teams that have not yet lost are determined based on the ordering of the teams in  $\mathcal{T}$  in advance of the



outcomes of any games.

All teams that finish a semibracket with no losses are declared co-champions.

#### Definition 4.4.4: Order of a Semibracket

The *order* of a semibracket is the number of co-champions it produces.

This is the same definition of a bracket but without the requirement that games be played until only one team is without losses. Note that semibrackets of order one are just traditional brackets. Figure 4.4.2 can be viewed then as being composed of two semibrackets of order two rather than of four brackets.

Many of the important properties of traditional brackets are true of semibrackets as well, up to and including the fundamental theorem. We state the key ones here without proof for brevity, though the proofs are analogous to those presented for traditional brackets.

#### Definition 4.4.5: Semibracket Signature

The *signature*  $[[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  of an  $r$ -team semibracket  $\mathcal{A}$  is a list of natural numbers, such that  $a_i$  is the number of teams with  $i$  byes.

#### Theorem 4.4.6

Let  $\mathcal{A} = [[\mathbf{a}_0; \dots; \mathbf{a}_r]]$  be a list of natural numbers. Then  $\mathcal{A}$  is a semibracket signature of order  $z$  if and only if

$$\sum_{i=0}^r a_i \cdot \left(\frac{1}{2}\right)^{r-i} = z.$$

Note that the semibracket signature summation sums to the order of the semibracket. As traditional brackets are just semibrackets of order one, this is consistent with Theorem 2.2.7.

#### Definition 4.4.7: Proper Semibracket

A *proper semibracket* is a semibracket that has been properly seeded.

#### Theorem 4.4.8

There is exactly one proper semibracket with each semibracket signature.

We can combine semibrackets into larger formats in the same way that we combined traditional brackets into simple multibrackets.

#### Definition 4.4.9: Multibracket

A *multibracket* is a sequence of semibrackets in which the losers of certain games in the upper semibrackets fall into the lower semibrackets rather than being eliminated outright, and teams place based on which semibracket they won.

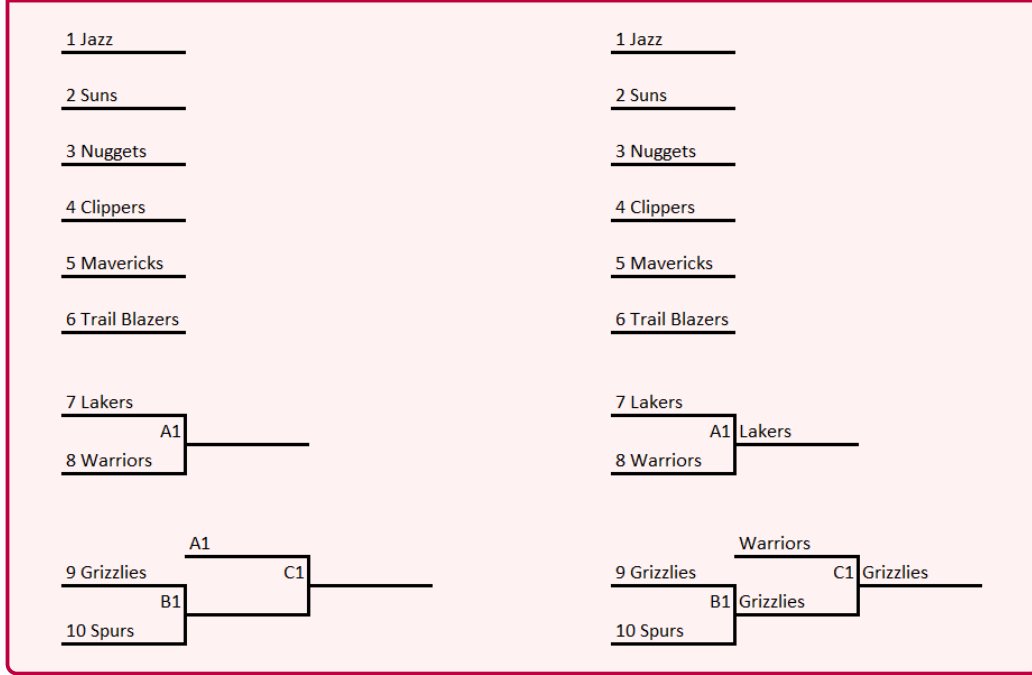
#### Definition 4.4.10: Order of a Semibracket

The *order* of a multibracket is the sum of the orders of the semibrackets it consists of.

Thus, saying that a multibracket is *simple* means that all of its semibrackets are of order one, and that all teams start in the primary bracket: in generalized multibrackets, some teams (in particular lower seeded teams) might start in lower brackets.

An example of a non-simple multibracket that took advantage of both properties is the 2021 NBA Western Conference Play-in Tournament, which was a ten-team multibracket with order eight and following signature:  $[[6]] \rightarrow [[2; 0]] \rightarrow [[2; 1; 0]]$ . The play-in tournament was used to whittle the top ten teams in the conference down to eight teams who would qualify for the playoffs.

Figure 4.4.11: 2021 NBA Western Conference Play-in



The top six seeds get slotted into the primary semibracket where they don't have to play any games in order to advance to playoffs. Seeds 7 and 8 each have to win one of their next two games to qualify, while seeds 9 and 10 have to win both of their next two.

It is now clear that the format in Figure 4.4.2 is also a multibracket, one of signature  $[[8; 0; 0]] \rightarrow [[4; 2; 0]]$ . Both that format and the one used in the 2021 NBA Western Conference Play-in share a property that the original format in Figure 4.4.1 does not: they are *efficient*.

#### Definition 4.4.12: Efficient

A multibracket is *efficient* if there are no games played within it such that both the winner and loser of that game are guaranteed to win a semibracket.

Identifying whether a (weakly proper) multibracket is efficient can be done by looking at its signature.

### Theorem 4.4.13

A weakly proper multibracket

$$\begin{aligned}\mathcal{A} &= \mathcal{A}_1 \rightarrow \dots \rightarrow \mathcal{A}_k \\ &= [[\mathbf{a}_{1,0}; \dots; \mathbf{a}_{1,r_1}]] \rightarrow \dots \rightarrow [[\mathbf{a}_{k,0}; \dots; \mathbf{a}_{k,r_k}]]\end{aligned}$$

is efficient if and only if for each  $i > 1$ ,

$$a_{i,r_i} = 0.$$

*Proof.* Let  $\mathcal{A}$  be a weakly proper multibracket.

Assume that for some  $i > 1$ ,  $a_{i,r_i} \neq 0$ . Thus there is at least one game  $\mathbf{G}$  such that the loser drops into  $\mathcal{A}_j$  and wins the semibracket without playing any more games. To show that  $\mathcal{A}$  is not efficient, we show that the winner of  $\mathbf{G}$  also wins a semibracket. If the winner of  $\mathbf{G}$  wins  $\mathcal{A}_j$  for  $j < i$ , then this is clearly true. But even if they don't, they will have lost more recently than the loser of  $\mathbf{G}$ , and so because  $\mathcal{A}$  is weakly proper, they too must enter into  $\mathcal{A}_j$  such that they win the semibracket without playing any more games. Thus,  $\mathcal{A}_j$  is not efficient.

Now assume that  $\forall i > 1$ ,  $a_{i,r_i} = 0$ . We will show by inducting on the semibrackets in  $\mathcal{A}$  in reverse that none of the semibrackets contain a game that violates the efficiency condition. Firstly,  $\mathcal{A}_k$  upholds the condition because any team that loses a game in  $\mathcal{A}_k$  doesn't fall into another semibracket, much less have a chance to win one.

Now we must show that if all of the semibrackets from  $\mathcal{A}_{j+1}$  to  $\mathcal{A}_k$  uphold the condition, then  $\mathcal{A}_j$  does as well. We let  $\mathbf{G}$  be a game in  $\mathcal{A}_j$ , and show that the loser of  $\mathbf{G}$  is not guaranteed to win a semibracket. If the loser of  $\mathbf{G}$  does not fall into another semibracket, then we are done. If they do, then because  $\forall i > 1$ ,  $a_{i,r_i} = 0$ , they will play another game. However, by induction, the loser of this game is not guaranteed to win a semibracket. Thus neither is the loser of  $\mathbf{G}$ .  $\square$

Efficient multibrackets are great tournament designs for tournaments whose primary goal is to select the top  $m$  teams to move on to the next

stage of the competitions, as discussed in the beginning of this section. They do so in exciting manner, with each spot in the top  $m$  being award as the winner of a particular game, efficiently, with no games being played between teams who have each already clinched spots, and fairly (as long as the format is as proper and respectful as desired.)

## 4.5 Swiss

## 5 References

- [1] Robert Chen and F. K. Hwang. Do stronger players win more knockout tournaments? *Journal of the American Statistical Association*, 73(363), 1978.
- [2] H.A. David. *The method of paired comparisons*, volume 12. London, 1963.
- [3] Christopher Edwards. *The Combinatorial Theory of Single-Elimination Tournaments*. PhD thesis, Montana State University, 1991.
- [4] Jeff Horen and Raymond Riezman. Comparing draws for single elimination tournaments. *Operations Research*, 33(2), 1985.
- [5] F. K. Hwang. New concepts in seeding knockout tournaments. *The American Mathematical Monthly*, 89(4), 1982.
- [6] Nate Silver. When 15th is better than 8th: The math shows the bracket is backward, 2011.
- [7] Eric Simon. The upa manual of championship series tournament formats, 2008.