Combinatorics, Probability and Computing (2021), **30**, pp. 374–397
doi:10.1017/S0963548320000425

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The minimum perfect matching in pseudo-dimension $0 < q < 1$

Joel Larsson

Department of Statistics, Lund University, Lund, Scania, Sweden
Email: joel.larsson@stat.lu.se

(Received 29 March 2016; revised 6 June 2017; accepted 2 July 2020; first published online 27 October 2020)

Abstract

It is known that for $K_{n,n}$ equipped with i.i.d. exp (1) edge costs, the minimum total cost of a perfect matching converges to $\zeta(2) = \pi^2/6$ in probability. Similar convergence has been established for all edge cost distributions of *pseudo-dimension* $q \geqslant 1$. In this paper we extend those results to all real positive q, confirming the Mézard–Parisi conjecture in the last remaining applicable case.

2020 MSC Codes: Primary: 05C80 (Random graphs); Secondary: 60C05 (Combinatorial probability)

1. Introduction

There has been substantial interest over the past few decades in the minimum matching problem: given a graph G, and a positive cost (or weight) associated to each edge of G, we want to find a perfect matching of minimal total cost $M(G)$. Of special interest is minimum matching on the complete graph K_n on n vertices or the complete bipartite graph $K_{n,n}$ on $n+n$ vertices, with random edge costs given by independent exp (1) variables. The latter is sometimes referred to as the *random assignment* problem. For this graph model, the lower bound $M(K_{n,n}) \geqslant 1 - o(1)$ (w.h.p.) is trivial: the cheapest edge from any given vertex has expected cost n^{-1}, and a perfect matching uses n edges. Similarly, $M(K_n) \geqslant 1/2 - o(1)$ w.h.p. The upper bound $\limsup_n M(K_{n,n}) \leqslant 3$ was established by Walkup [12], by finding a perfect matching using only fairly cheap edges. This was later improved to 2 by Karp [4].

Mézard and Parisi [8] conjectured that $M(K_{n,n})$ converges in probability to $\zeta(2) = \pi^2/6$ and $M(K_n)$ to $\zeta(2)/2$, based on heuristic replica symmetry calculations. Aldous [1] proved that the limit exists, and later confirmed the conjecture [2]. Both of these papers used what is sometimes called the 'objective method' [3], and worked with matchings on an infinite limit object. Parisi [10] further conjectured the more precise result that $\mathbb{E}[M(K_{n,n})] = \sum_{k=1}^{n} k^{-2}$. This was later established independently by Nair, Prabhakar and Shaw [9] and Linusson and Wästlund [6], both using inductive proofs. The proof was later simplified by Wästlund [13]. Salez and Shah [11] gave yet another proof of the Mézard–Parisi conjecture, using the objective method to analyse the behaviour of belief propagation on the limit object.

A more comprehensive overview of the existing literature and related problems can be found in a survey paper by Krokhmal and Pardalos [5].

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

A natural question is whether these results extend to other edge cost distributions. It turns out that only the scaling behaviour of the probability distribution near 0 matter. A random graph where edge costs are i.i.d. copies of the random variable ℓ is said to be of *pseudo-dimension q* if $\lim_{x\to 0} \mathbb{P}(\ell \leqslant x) \cdot x^{-q}$ exists and lies in $(0, \infty)$. The exponential distribution and the uniform distribution on $[0, 1]$ are both of pseudo-dimension 1, and the chi distribution with q degrees of freedom as well as the Weibull distribution with shape parameter q (*i.e.* the $(1/q)$th power of an exponential variable) are of pseudo-dimension q. In the paper by Mézard and Parisi [8], q was a real positive parameter, but most focus since then has been on the special case $q = 1$.

The motivation for the term pseudo-dimension is as follows. For $q \in \mathbb{N}$, a geometric graph model is given by embedding the vertices as n points chosen uniformly and independently at random in a hypercube $[0, 1]^q$, and setting the edge costs to be the corresponding Euclidean distances. The mean field approximation (*i.e.* the graph model where edge costs follow the same distribution, but are i.i.d.) of a geometric graph model of dimension q is a graph model of pseudo-dimension q.

For any graph G and probability measure ν on $\mathbb{R}_+$, let $G[\nu]$ denote G equipped with i.i.d. edge costs with distribution given by ν. If ν is of pseudo-dimension q, then the cost of the minimum matching on $K_{n,n}[\nu]$ can be shown to be of order $\Theta(n^{1-1/q})$, by a minor modification of [12]. This suggests studying the quantity $n^{-1+1/q}M(K_{n,n}[\nu])$. Does it converge in probability to a constant? This question was answered in the affirmative for $q \geqslant 1$ by Wästlund [14] (both for $K_{n,n}$ and K_n), but it remained open for $0 < q < 1$. Our main result is the following theorem, confirming the Mézard–Parisi conjecture for all $q > 0$.

Theorem 1.1. *For every $q \in (0, 1)$, there exists a $\beta = \beta(q)$ such that for any probability measure ν for which $c := \lim_{x\to 0} \nu(\{\ell \leqslant x\}) \cdot x^{-q}$ exists and $c \in (0, \infty)$,*

$$\frac{M(K_{n,n}[\nu])}{c^{1/q}n^{1-1/q}} \to \beta(q) \quad \textit{and} \quad \frac{M(K_n[\nu])}{c^{1/q}n^{1-1/q}} \to \frac{1}{2}\beta(q)$$

in probability as $n \to \infty$ (through even n in the latter case).

The reason for the factor $1/2$ in the graph model K_n is that perfect matchings in $K_{n,n}$ and K_n have n and $n/2$ edges respectively, but it turns out that the average cost of edges participating in the optimal perfect matching is (asymptotically) the same in both graph models.

We believe the theorem should hold in somewhat greater generality, *i.e.* for graph sequences G_n other than $K_{n,n}$ and K_n. As we will discuss in Section 3.6, the theorem only depends on the structure of the sparse subgraphs of G_n obtained by removing expensive edges. Roughly speaking, these subgraphs need to look like Galton–Watson trees locally, and be good global expanders. We therefore conjecture the following, which is most easily phrased in terms of graphons.[a]

Conjecture 1.2. Let $W\colon [0, 1]^2 \to [0, 1]$ be a connected graphon and let G_n be a sequence of dense graphs with $|V(G_n)| = 2n$ which converges to W.

If there is a $\delta > 0$ such that $\int_0^1 W(x, y)\, dx = \delta$ for almost all $y \in [0, 1]$, then $n^{-1+1/q}M(G_n[\nu])$ converges in probability to a constant (depending only on δ, $q = q(\nu)$, and $c = c(\nu)$).

In other words, if G_n admits a connected graphon limit and has degree $\delta n + o(n)$ at every vertex for a constant $\delta > 0$, then an analogue of Theorem 1.1 should hold for G_n. Special cases of this conjecture include the Erdős–Rényi graph $G_{n,\delta}$ and complete balanced k-partite graphs.

[a] A graphon is an analytic limit object for a sequence of dense graphs. For definitions and further details we refer to [7, Chapter 7].

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

2. Notation and definitions

We will assume $0 < q < 1$ and a large parameter $\lambda > 1$ is fixed, and often suppress dependence on them in our notation. We will restrict our attention to $0 < q < 1$, since Theorem 1.1 is already known to be true for $q \geqslant 1$. Although our proof strategy works for $q \geqslant 1$ too, some parts of our lemmas are trivial in that case, and assuming $0 < q < 1$ streamlines the proofs.

Unless otherwise stated, all functions f considered will be either $f\colon V(G) \to \Lambda$ for some graph G or $f\colon \Lambda \to \mathbb{R}$, where $\Lambda := [-\lambda/2, \lambda/2]$. For f and g functions on Λ, we will use $f \leqslant g$ to mean that $f(z) \leqslant g(z)$ for all $z \in \Lambda$. For $x \in \mathbb{R}$, we let $x_+ := \max(x, 0)$. We write $a_n \nearrow a$ if $a_n \to a$ and a_n is a non-decreasing sequence, and $a_n \searrow a$ if $-a_n \nearrow -a$.

An m-rooted graph is a graph where m of the vertices have been designated as root vertices. For an edge-weighted m-rooted graph G, the (k, λ)-truncation $G(k, \lambda)$ is the subgraph of G obtained as the union of k-neighbourhoods of the roots after all edges of weight more than λ have been removed. Equivalently, $G(k, \lambda)$ is the union of all paths from a root of length at most k that only use edges of weight at most λ. We say that an m-rooted random graph G is the λ-*local limit* of the m-rooted random graph sequence G_n if and only if, for every $k \in \mathbb{N}$, $G_n(k, \lambda)$ converges to $G(k, \lambda)$ in the total variation metric.[b] We will think of these edge weights as costs, and use the words cost and weight interchangeably.

Furthermore, we will let $|G|$ denote the number of *edges* of a graph G, and we will consider the edges of a rooted tree to be directed away from the root ϕ. By *path* we will mean a directed path away from the root. If u is the parent of v, we write $u \to v$. Let $|u|$ denote the distance to u from the root.

3. Proof strategy

In this section we will give an overview of our proof strategy, and how it relates to the proof for $q \geqslant 1$ in [14]. In broad terms, the strategy is to prove that the 'local' structure of the optimal matching is 'locally' determined. That is, for each edge uv we can approximate its expected contribution to the total cost of the optimal matching by only looking at some large but finite neighbourhoods of u and v.

As a first step, we switch to working with a rescaled model. Multiply all edge costs in $K_n[\nu]$ and $K_{n,n}[\nu]$ by $(n/c)^{1/q}$. Since minimum perfect matching is a linear programming problem, the only effect this has on the optimum is to multiply its cost by the same amount. One can think of this as changing the units of cost in such a way that the expected number of edges of cost at most 1 from a given vertex is $1 + o(1)$. Such rescaling allows us to work with λ-local limits more easily.

Let $\tilde{K}_n$ and $\tilde{K}_{n,n}$ be these rescaled models (suppressing dependence on ν), so that the quantities in Theorem 1.1 can be rewritten as

$$\frac{M(K_{n,n}[\nu])}{c^{1/q} n^{1-1/q}} = \frac{1}{n} M(\tilde{K}_{n,n}) \quad \text{and} \quad \frac{M(K_n[\nu])}{c^{1/q} n^{1-1/q}} = \frac{1}{n} M(\tilde{K}_n).$$

We also let $\ell = \ell(u, v)$ denote the edge cost of the edge uv in the rescaled model, and we say that it is *cheap* if $\ell(u, v) \leqslant \lambda$.

3.1 Local limit

We will be working with the λ-local limits of $\tilde{K}_n$ and $\tilde{K}_{n,n}$. By Proposition 2.2 of [14], the λ-local limit of $\tilde{K}_n$ rooted in m arbitrary vertices is m disjoint independent copies of a certain random tree

[b]The total variation distance between two edge-weighted random graphs is at most ε if they can be coupled in such a way that they are isomorphic with probability at least $1 - \varepsilon$ (where the isomorphism preserves roots and edge weights). For further details see [14, Sections 2.1–2.2].

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

T_λ^q which we will define shortly. It follows from the proof of this proposition that this is also the λ-local limit of $\tilde{K}_{n,n}$ rooted in m arbitrary vertices.

The tree T_λ^q is defined as the 1-rooted Galton–Watson tree with offspring distribution Poi (λ^q) and edge weights given by i.i.d. copies of the $[0, \lambda]$-valued random variable X with CDF $\mathbb{P}(X \leqslant t) = (t/\lambda)^q$. Note that since λ is a large number, this tree is supercritical, and infinite with probability $1 - o_\lambda(1)$. We will not reproduce the proof of [14, Proposition 2.2] here, but instead sketch an argument for why disjoint copies of T_λ^q form a plausible λ-local limit of $\tilde{K}_n$ and $\tilde{K}_{n,n}$.

Edge weight distribution. By the definition of pseudo-dimension and the rescaling of the edge weights, for any fixed $t > 0$ (*i.e.* not depending on n), $\mathbb{P}(\ell \leqslant t) = (1 + o(1))t^q/n$. For any $t \leqslant \lambda$,

$$\mathbb{P}(\ell \leqslant t \mid \ell \leqslant \lambda) = \frac{\mathbb{P}(\ell \leqslant t)}{\mathbb{P}(\ell \leqslant \lambda)} = (1 + o(1))(t/\lambda)^q,$$

and hence the edge weights of the λ-local limit should be i.i.d. with CDF $(t/\lambda)^q$.

Degree distribution. The probability that an edge is cheap is $p := \mathbb{P}(\ell \leqslant \lambda)$, independently of all other edges. In other words, the subgraph of $\tilde{K}_n$ consisting of all cheap edges is an Erdős–Rényi random graph $G_{n,p}$, and the analogous subgraph of $\tilde{K}_{n,n}$ is a bipartite Erdős–Rényi random graph $G_{n,n,p}$. The degree distributions of $G_{n,p}$ and $G_{n,n,p}$ are approximately Poisson with mean $p(n-1) = (1 + o(1))\lambda^q$ and $pn = (1 + o(1))\lambda^q$ respectively. Hence the λ-local limit of these graphs should have degree distribution Poi (λ^q).

Disjoint trees. Let $\mathcal{P}_k$ be the set of paths of length at most $k+1$ starting from any of the roots (in $G_{n,p}$ or $G_{n,n,p}$). A first moment calculation shows that the expected number of pairs $P, P' \in \mathcal{P}_k$ with $P \neq P'$ and $P \cap P' \neq \emptyset$ is $O(p^{2k+2} n^{2k+1}) = o(1)$. But if no such pair exists, then the k-neighbourhoods of the roots are disjoint trees. Hence the λ-local limit should consist of m disjoint trees.

We will frequently use the following equivalent construction of T_λ^q, where we generate the offspring of a vertex and the corresponding edge weights concurrently. For every vertex u, run an inhomogeneous Poisson point process on the time interval $[0, \lambda]$ with intensity qt^{q-1} at time t. Let m_q be the corresponding intensity measure, *i.e.* the measure on $[0, \lambda]$ such that $dm_q(t) = qt^{q-1}\, dt$. (Note that $\lambda^{-q} m_q$ is then the probability measure corresponding to the CDF $(t/\lambda)^q$ in our original definition of T_λ^q.) If $\ell_1, \ell_2, \ldots, \ell_j$ are the arrival times of the events in this process, we let $v_1, v_2, \ldots, v_j$ be the children of u, and give the edge uv_i weight ℓ_i.

Finally, while not necessary in order to understand our paper, it is worth mentioning that T_λ^q can also be constructed from the so-called 'Poisson-weighted infinite tree' (PWIT) which is often encountered in the literature on the objective method (*e.g.* [2], [3]). Let T be the tree obtained by raising all edge weights in the PWIT to the $(1/q)$th power. Then T_λ^q has the same distribution as $T(\infty, \lambda)$, *i.e.* the connected component of the root after all edges of weight more than λ are removed.

3.2 Exploration game

The game *Exploration* was introduced in [14]. This zero-sum perfect information game is played in the following way. On an edge-weighted rooted graph G, Alice and Bob take turns picking the next edge of a self-avoiding walk starting from the root. When it is a player's turn (Alice's, say), and the current vertex is u, she can take one of two actions.

(i) Pick any neighbour v of u that has not already been visited, and pay Bob the cost $\ell(u, v)$ of the edge uv. Bob then continues the game from v.
(ii) Quit the game, and pay Bob a penalty of $\lambda/2$, for some fixed parameter $\lambda > 0$.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

The payoff for Alice, once the game has finished, is the total amount Bob has paid to her minus the total amount she has paid to Bob. Each player's aim is to maximize their payoff. Note that it is always better to quit than to pick an edge with weight $\ell > \lambda$: even if the other player were to quit immediately after one picks this edge, the payoff for the last moves would be $-\ell + \lambda/2 < -\lambda/2$.

If the weighted graph G is finite, every game position has a well-defined game value $f = f(G, u)$.[c] If Alice starts by moving from u to v, the remainder of the game is equivalent to a game played on $G - u$ started on v, but with the roles of Alice and Bob reversed. By considering all possible options a player has from the vertex u, it is easy to see that

$$f(G, u) = \min\Big(\lambda/2, \min_{v\sim u} (\ell(u, v) - f(G - u, v))\Big), \tag{3.1}$$

where the second minimum is taken over all neighbours v of u. If the graph is a finite rooted tree T and we start the game at the root, no move can go from a vertex to its parent, so we may consider edges to be directed away from the root and forbid moves towards the root. But then $f(T, v) = f(T - u, v)$ if u is the parent of v, so we let $f(v) := f(T, v)$. Thus (3.1) can be slightly simplified to

$$f(u) = \min\Big(\lambda/2, \min_{v\leftarrow u} (\ell(u, v) - f(v))\Big), \tag{3.2}$$

where the second minimum is taken over all *children* v of u. If the tree is infinite, however, it is no longer clear that the function f is well-defined. Instead, we consider the set of functions f which satisfy (3.2) for all $u \in V(T)$, and call these 'game valuations'.

Wästlund [14] proved that for any $q > 0$, the limits of $n^{-1}M(\tilde{K}_n)$ and $n^{-1}M(\tilde{K}_{n,n})$ exist if, for all large λ, there exists a unique game valuation on (almost all realizations of) T_λ^q. We give an overview of this proof and why it works for both the complete and the complete bipartite graph in Section 3.5.

Wästlund proceeded to prove that the valuation was indeed unique for $q \geqslant 1$ ([14, Proposition 2.8]), but that proof did not extend to $0 < q < 1$. Therefore, in order to prove Theorem 1.1 it suffices to show the following.

Proposition 3.1. *For any $\lambda > 0$ and $q \in (0, 1)$, there is almost surely a unique game valuation on T_λ^q, i.e. a function $f\colon V(T_\lambda^q) \to \Lambda$ satisfying (3.2) for every $u \in V(T_\lambda^q)$.*

We will first show that the set of all game valuations on T_λ^q admits a bounded lattice ordering. The recursion (3.2) has a useful *monotonicity property*: if f and g are game valuations on a tree T such that $f(v) \leqslant g(v)$ for all children v of u, then $f(u) \geqslant g(u)$.

Let $f^k(u)$ be a game valuation on $V(T_\lambda^q(k, \lambda))$ (*i.e.* the tree T_λ^q truncated after k generations) satisfying (3.2) for all u with $|u| \leqslant k$. This valuation is almost surely unique, since $T_\lambda^q(k, \lambda)$ is almost surely finite. Note also that $f^k(u) = \lambda/2$ for all u with $|u| = k$, since these u have no offspring in $T_\lambda^q(k, \lambda)$.

Let us compare the valuations f^k and f^j for $j > k$. For any v with $|v| = k$, $f^j(v) \leqslant \lambda/2 = f^k(v)$, so by the monotonicity property, $f^k(u) \leqslant f^j(u)$ for all u with $|u| = k - 1$. Iterating this, we see that $f^k(u) \geqslant f^j(u)$ when $|u|$ has the same parity as k, and $f^k(u) \leqslant f^j(u)$ otherwise. This suggests defining the following partial orders on the set of functions on a rooted tree:

$$f \preceq_k g \iff \begin{cases} f(u) \leqslant g(u) & \forall u\colon |u| \text{ is odd and} \leqslant k, \\ f(u) \geqslant g(u) & \forall u\colon |u| \text{ is even and} \leqslant k. \end{cases} \tag{3.3}$$

[c] By the game value of $u \in V(G)$ we will mean the value of the exploration game on G, starting from u, to the second player. In other words, the net amount that Alice will pay to Bob, assuming optimal play by both.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

The previous inequalities between f^j and f^k can now be restated as follows: for any $j > k$, $f^k \preceq_k f^j$ if k is even and $f^k \succeq_k f^j$ if k is odd.

Lemma 3.2. *We say that $f \preceq g$ whenever $f \preceq_k g$ for all $k \in \mathbb{N}$. The order $\preceq$ is a bounded lattice order, with unique maximum f_A and minimum f_B given by the pointwise limits*

$$f_A(u) := \lim_{k\to\infty} f^{2k+1}(u) \quad \text{and} \quad f_B(u) := \lim_{k\to\infty} f^{2k}(u). \tag{3.4}$$

Proof. The order $\preceq$ is reflexive and transitive. It is also anti-symmetric, *i.e.* $f \preceq g \preceq f \Rightarrow f = g$, and hence a lattice order. We will prove that f_B exists and is the minimum; the argument for f_A is analogous.

Let g be any game valuation on $T^q_\lambda(2k, \lambda)$ (*i.e.* the first $2k$ levels of T^q_λ). For any v with $|v| = 2k$, $g(v) \leqslant \lambda/2 = f^{2k}(v)$, and hence $f^{2k} \preceq_{2k} g$. In particular, this holds for $g = f^{2k+2}$, *i.e.* $f^{2k} \preceq_{2k} f^{2k+2}$. So for any $u \in T^q_\lambda$, the sequence $(f^{2k}(u))_{2k>|u|}$ is monotone (non-decreasing or non-increasing depending on the parity of $|u|$) and hence the limit $f_B(u)$ exists.

To see that f_B is the minimum, let g be any game valuation on T^q_λ (not just on a truncation). Pick any $u \in T^q_\lambda$ with $|u|$ even, and consider $g(u)$. By the previous argument $f^{2k} \preceq_{2k} g$ for any k, and hence $f^{2k}(u) \geqslant g(u)$ for any $k > |u|/2$. Letting k tend to infinity, we see that $f_B(u) \geqslant g(u)$. Similarly, $f_B(u) \leqslant g(u)$ for all u with $|u|$ odd. Hence (3.3) is satisfied for any k, or in other words $f_B \preceq g$. ☐

The strategy will be to analyse the game where Alice and Bob play according to f_A and f_B respectively. One can show that if this game ends after finitely many moves, then $f_A = f_B$ (see the proof of either [14, Proposition 2.8] or Proposition 3.1). But by Lemma 3.2, all game valuations are sandwiched between f_A and f_B, so if $f_A = f_B$ there must be a unique game valuation.

3.3 Proof strategy for Proposition 3.1 for q ⩾ 1

We give here a short description of Wästlund's proof of an analogous statement of Proposition 3.1 for $q \geqslant 1$, in order to explain how our proof for $q > 0$ is similar, and yet differs from it significantly. One way to explain the proof idea is that there are two main components: (i) show that a game where both Alice and Bob play according to f_A must finish after finitely many moves, and (ii) show this game is not 'too different' from a game where Alice plays according to f_A and Bob according to f_B.

Let $u_0 := \phi, u_1, u_2, \ldots$ be the (finite or infinite) game path when both Alice and Bob play according to f_A. Let $Z_i := f_A(u_i)$ for $i \geqslant 0$, and note that these random variables are not independent. If $Z_i = \lambda/2$, then the f_A-optimal move from u_i is to quit and pay the penalty, and the game path is finite if and only if this happens for some i. It is not too hard to show that $\mathbb{P}(Z_i = \lambda/2) > 0$ uniformly in i, but since the Z_i are not independent this is not sufficient. However, conditional on Z_i, $(Z_0, \ldots, Z_{i-1})$ and $(Z_{i+1}, Z_{i+2}, \ldots)$ are independent. So in order to prove (i), it therefore suffices to show that there is an $\varepsilon > 0$ such that $\mathbb{P}(Z_{i+1} = \lambda/2 \mid Z_i = z) > \varepsilon$ for all $z \in \Lambda$.

A move is said to be δ-*reasonable* if and only if it is within δ of being f_A-optimal,[d] *i.e.* the move $u \to v$ is δ-reasonable if $\ell(u, v) - f_A(v) \leqslant f_A(u) + \delta$. It turns out that for any $\delta > 0$, if Bob plays according to f_B his moves will be δ-reasonable eventually – after a vertex w, say (see the proof of [14, Proposition 2.8]). Using (3.2) and the choice of f_A, Wästlund proved (statements equivalent to) the following more precise versions of (i) and (ii). Pick $u \in V(T)$ in a way that does not depend on the subtree rooted in u. Then

[d]In [14] reasonable moves were defined in terms of their deviations from f_B, but by focusing on f_A instead our notation becomes slightly more convenient later in Section 4.2.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

(i) there is an $\varepsilon > 0$ such that if $u \to v$ is the f_A-optimal move from u, then

$$\inf_{-\lambda/2 \leqslant z < \lambda/2} \mathbb{P}(f_A(v) = \lambda/2 \mid f_A(u) = z) > \varepsilon,$$

(ii) for any $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\sup_{z \in \Lambda} \mathbb{E}[\#(\delta\text{-reasonable but not } f_A\text{-optimal moves from } u) \mid f_A(u) = z] < \varepsilon/2.$$

Note that there is a.s. no f_A-optimal move $u \to v$ if $f_A(u) = \lambda/2$, which is why the range of the infimum in (i) excludes $\lambda/2$.

Now assume that Alice and Bob play according to f_A and f_B respectively. We will give a sketch of the argument that (i) and (ii) together imply that the game will end after finitely many moves [14, Proposition 2.5, Lemma 2.7].

Consider a vertex u at even distance from the root (so that it it Alice's turn at u), chosen in a way independent from the subtree rooted in u. Condition on $f_A(u)$. What is the expected number of vertices reachable by an f_A-optimal move by Alice followed by a δ-reasonable move by Bob?

Conditional on $f_A(u) = \lambda/2$, there are a.s. no such vertices, because Alice will a.s. quit at u. Conditional on $f_A(u) = z$ for some $z < \lambda/2$, there is an f_A-optimal move $u \to v$ which Alice will choose. Then by (i), $f_A(v) = \lambda/2$ with probability at least ε, so the expected number of f_A-optimal moves from v is at most $1 - \varepsilon$. By (ii), the expected number of δ-reasonable but not f_A-optimal moves from v is at most $\varepsilon/2$. In total, the expected number of δ-reasonable moves is at most $1 - \varepsilon/2$.

If we consider the tree $R(u)$ of game paths from u consisting of f_A-optimal play by Alice and δ-reasonable play by Bob, we see that its expected number of vertices at level k is at most $(1 - \varepsilon/2)^{\lfloor k/2 \rfloor}$. Hence $R(u)$ is almost surely finite. This in fact holds for all vertices simultaneously: since the tree T_λ^q has countable many vertices, there is almost surely no vertex u' with $R(u')$ infinite.

Now, recall that w is a vertex such that Bob plays δ-reasonably after w, whence $R(w)$ is guaranteed to contain the game path from w. But $R(w)$ is almost surely finite, and hence the game will end after finitely many moves.

3.4 Revised proof strategy for $0 < q < 1$

The main trouble that arises when trying to apply the argument above when $q \in (0, 1)$ is that the proof of statement (i) above fails for $q < 1$. Indeed, the statement is false for $q \leqslant 1/2$: see Remark 4.1.

Our aim is still to show that the probability of a player quitting at any given time is uniformly bounded away from 0 (when both players play according to f_A). Whether or not a player will quit the game at u_i is determined by the random variable $Z_i := f_A(u_i)$, but as mentioned earlier these random variables are not independent. This problem was sidestepped in Wästlund's proof by the conditioning in (i) above, but since that fails for small q we will instead need to understand the dependency between Z_i and Z_{i+1}. We do this by constructing a pair of linear operators (one for each parity of i) that map functions of the form $z \mapsto \mathbb{E}[\bullet \mid Z_{i+1} = z]$ to $z \mapsto \mathbb{E}[\bullet \mid Z_i = z]$. The statement that will correspond to (i) will be that the composition of these two operators is a contraction (Lemma 4.8). Having changed one major component of the proof, the second one (ii) is no longer compatible. The linear operators we construct can only provide information about the conditional expectation of random variables, so we must change our aim from proving that (some structure containing) the game path is almost surely finite, to proving that it has finite expected size. However, the expected size of the tree of reasonable moves (as defined in Section 3.3) does not appear to be finite.

We solve this by using a refined concept of reasonable moves, where we take into account not only single deviations by Bob from f_A, but instead consider the *sum* of these deviations along

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

a game path. This leads to a significantly smaller tree of reasonable moves (guaranteed to contain the game path) whose expected size we can bound recursively.

3.5 The connection between exploration and matching

Here we will briefly discuss the connection between the seemingly disparate topics of the exploration game and the minimum perfect matching problem. This is not strictly necessary in order to understand our proof of Proposition 3.1 (which is the novel result of this paper), but it gives some insight into how Proposition 3.1 implies Theorem 1.1. For a full proof of this implication, we refer the interested reader to Section 3 of [14]. Crucially, at no point in the proof of [14, Theorem 3.2] is the assumption $q \geqslant 1$ used: the proof only depends on there being a unique game valuation $f := f_A = f_B$ on T^q_λ.

We begin by defining the λ-*relaxed* (or λ-*diluted*) matching problem. A *partial matching* in a graph G is a subgraph $H \subseteq G$ where no two edges share a vertex. For a partial matching H on an edge-weighted graph G, we say that the λ-*relaxed* cost $c_\lambda(H)$ of it is the sum of the costs of all edges it contains, plus $\lambda/2$ for each unmatched vertex. In other words,

$$c_\lambda(H) := \sum_{uv \in E(H)} \ell(uv) + \sum_{u \notin V(H)} \lambda/2.$$

We let $M_\lambda(G) := \min_H c_\lambda(H)$, where the minimum is taken over all partial matchings. We also let $M^\star_\lambda(G)$ be the sum of edge costs of the matching H which minimizes $c_\lambda(H)$. Note that for any graph G, $M_\lambda(G)$ is an increasing function of λ, and $M^\star_\lambda(G) \leqslant M_\lambda(G) \leqslant M(G)$. There are two results from [14] that connect λ-relaxed matchings to perfect matchings and to the exploration game, respectively.

First, [14, Theorem 3.2] shows that the existence of a unique game valuation (Proposition 3.1) implies that $\beta_\lambda := \lim_n n^{-1} M^\star_\lambda(\tilde{K}_n)$ exists for any λ. A rough outline of the argument is as follows.

Let C_{uv} be the contribution of uv to $M^\star_\lambda(\tilde{K}_n)$, *i.e.* C_{uv} is the cost $\ell := \ell(u, v)$ of the edge uv if this edge participates in the minimal-cost λ-relaxed matching, and $C_{uv} = 0$ otherwise. Then $M^\star_\lambda(\tilde{K}_n) = \sum_{u,v} C_{uv}$, where the sum ranges over all pairs of vertices. It can be shown that for a finite graph, an arbitrary edge uv participates in the optimal λ-relaxed matching if and only if the move $u \to v$ is optimal in the exploration game on $\tilde{K}_n$ starting from u. Let $\ell := \ell(u, v)$, $Z_u := f(\tilde{K}_n - uv, u)$ and $Z_v := f(\tilde{K}_n - uv, v)$. Then the move $u \to v$ is optimal if and only if $Z_u \geqslant \ell - Z_v$, because by (3.1) the value (to the second player) of the move $u \to v$ is $\ell - Z_v$, while the value of the best move other than $u \to v$ is Z_u. Noting that $\{\ell \leqslant Z_u + Z_v\} \subseteq \{\ell \leqslant \lambda\}$, we can calculate the expected value of C_{uv}:

$$\mathbb{E}C_{uv} = \mathbb{E}[\chi_{\{\ell \leqslant Z_u + Z_v\}} \cdot \ell] = \mathbb{E}[\chi_{\{\ell \leqslant Z_u + Z_v\}} \cdot \ell \mid \ell \leqslant \lambda] \cdot \mathbb{P}(\ell \leqslant \lambda), \tag{3.5}$$

where χ_E denotes the indicator random variable for the event E. Let X_{uv} be defined as $(\ell \mid \ell \leqslant \lambda)$, *i.e.* the random variable given by the CDF $\mathbb{P}(X_{uv} \leqslant t) = \mathbb{P}(\ell \leqslant t \mid \ell \leqslant \lambda)$. The three random variables Z_u, Z_v and X_{uv} depend on n, but we will find their limits in probability as $n \to \infty$. Since (Z_u, Z_v) is independent from X_{uv}, we deal with them separately.

Recall that for any $t \leqslant \lambda$, $\mathbb{P}(\ell \leqslant t \mid \ell \leqslant \lambda) = (1 + o(1))(t/\lambda)^q$. Hence X_{uv} converges in distribution to a random variable X with CDF $\mathbb{P}(X \leqslant t) = (t/\lambda)^q$ (*i.e.* X has probability measure $\lambda^{-q} m_q$). The λ-local limit of $\tilde{K}_n - uv$ rooted at u and v is two disjoint independent copies of T^q_λ, and by the sandwiching argument in (3.4) together with Proposition 3.1, the pair of random variables (Z_u, Z_v) converge in probability to a pair (Z, Z') of i.i.d. copies of $f(T^q_\lambda, \phi)$. Hence (3.5) equals

$$(1 + o(1)) \cdot \mathbb{E}[\chi_{\{X \leqslant Z + Z'\}} \cdot X] \cdot \mathbb{P}(\ell \leqslant \lambda) = \frac{1 + o(1)}{n} \cdot \underbrace{\lambda^q \cdot \mathbb{E}[\chi_{\{X \leqslant Z + Z'\}} \cdot X]}_{=:\beta_\lambda}, \tag{3.6}$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

and in particular, $\mathbb{E}C_{uv}$ only depends on the distributions of Z, Z' and X. In other words, it depends on no structure of $\tilde{K}_n$ other than its λ-local limit and its edge cost distribution. Summing over all $\binom{n}{2}$ edges gives

$$\mathbb{E}M^\star_\lambda(\tilde{K}_n) = \sum_{uv} \mathbb{E}C_{uv} = (1+o(1))\frac{n}{2}\beta_\lambda.$$

Since rooting $\tilde{K}_n$ in any four vertices u, v, x, y gives a graph whose λ-local limit is four disjoint independent copies of T^q_λ, one can show with a similar calculation that

$$\mathbb{E}C_{uv}C_{xy} = (1+o(1))\mathbb{E}C_{uv}\mathbb{E}C_{xy}, \tag{3.7}$$

so by the second moment method, the sum $M^\star_\lambda(\tilde{K}_n) = \sum_{u,v} C_{uv}$ is concentrated around its mean $(1+o(1))(n/2)\beta_\lambda$.

Second, [14, Proposition 3.4] uses a variation on Pósa's extension–rotation method to show that if a graph sequence G_n with random edge costs $\ell(u, v)$ satisfies a certain expansion property, then a partial matching with few unmatched vertices can be extended to a perfect matching at a small extra cost. More precisely, if the partial matching has total cost M and leaves $\delta|V(G_n)|$ vertices unmatched for some $\delta > 0$, then there exists a perfect matching with total cost at most $M + o_\delta(1)\cdot|V(G_n)|$.

In [14, Lemma 3.5] it is shown that K_n satisfies this expansion property, and in [14, Proposition 3.1] it is shown that the fraction of unmatched vertices in the optimal λ-relaxed matching is $o_\lambda(1)$. Hence, by the extension–rotation argument, $M(\tilde{K}_n) \leqslant M^\star_\lambda(\tilde{K}_n) + o_\lambda(1)\cdot n$, and $M^\star_\lambda(G) \leqslant M(G)$ holds trivially for any weighted graph G. Since this holds for all large λ, $M(\tilde{K}_n) = (1+o(1))(n/2)\beta$ where $\beta := \limsup_{\lambda\to\infty} \beta_\lambda$.

The argument is nearly identical for $\tilde{K}_{n,n}$ as for $\tilde{K}_n$, except for the minor difference that the total number of edges is n^2 rather than $\binom{n}{2}$, and hence (w.h.p.)

$$M^\star_\lambda(\tilde{K}_{n,n}) := \sum_{u,v} C_{uv} = (1+o_n(1))n\beta_\lambda$$

(where the sum ranges over all u at the 'left' side of $K_{n,n}$ and v at the 'right' side). It is also worth noting that [14, Proposition 3.4] is only done explicitly for K_n, but the proof works without modification for $K_{n,n}$. In fact the crucial expander property in [14, Lemma 3.5] is stated in terms of a random subgraph of a bipartition[e] of K_n.

3.6 Generalizing to other graphs

A natural question now is: For what graph sequences G_n other than $\tilde{K}_n$ and $\tilde{K}_{n,n}$ does the above argument work? For the extension–rotation argument, we need the expander property of [14, Lemma 3.5]. For the calculations in (3.6) and (3.7) to work, the λ-local limit of G_n when rooting in four arbitrary vertices must be four disjoint independent copies of T^q_λ (possibly after rescaling the edge costs of G_n by some factor). But if this holds, the rest of the argument in [14, Theorem 3.2] also works. In fact, if one can show concentration of $M^\star_\lambda(G_n)$ in some way other than (3.7), it suffices to consider rooting in two arbitrary vertices.

A simple example is an Erdős–Rényi random graph $G_{n,p}$ with constant $p > 0$. It has the λ-local limit T^q_λ (if one adjusts the cost scaling factor appropriately), and verifying that it has the expander property should not be too difficult. This approach might also be feasible for $p \to 0$ slowly with n, but clearly not for $p < \log n/n$ (since then $G_{n,p}$ has isolated vertices with positive probability).

A necessary condition for G_n to have the λ-local limit T^q_λ at every vertex is that it is approximately regular. A graph that also has some 'self-similarity', in the sense that the subgraph induced

[e] The proof *would* need some modification for (say) the complete k-partite graph, $k \geqslant 3$.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

by a random subset of vertices 'looks like' the whole graph, should have decent expander properties, and its (k, λ)-truncations should 'look the same' rooted in any vertex. Any graphon is 'self-similar' in this way, and the graphon in Conjecture 1.2 is approximately regular. It therefore seems like a good candidate for G_n.

Some other graphs for which this proof method might work are quasi-random graphs and k-partite complete graphs.

4. Proof of main theorem

4.1 The tree T_λ^q conditional on the game valuation f_A

In order to be able to construct the linear operators mentioned in Section 3.4, we will change slightly how we generate the random tree T_λ^q and the game valuation f_A. Instead of first generating T_λ^q and then calculating f_A 'back from infinity', we will generate the tree and vertex labels f_A concurrently. This will require the following lemma from [14].

Let $F_A(z) := \mathbb{P}(f_A(\phi) \geqslant z)$. In a slight abuse of notation, we will also use F_A to refer to the probability measure on Λ of the random variable $f_A(\phi)$. Similarly, F_B will refer to both the function $z \mapsto \mathbb{P}(f_B(\phi) \geqslant z)$ and the corresponding measure. Furthermore, let the ℓf-square be the set $\{(\ell, f)\colon 0 \leqslant \ell \leqslant \lambda, |f| \leqslant \lambda/2\}$, and recall that m_q is the measure on $[0, \lambda]$ such that $dm_q = qt^{q-1}\, dt$.

Lemma 4.1. (Lemma 2.6 of [14][f]). *Let $u \in V(T_\lambda^q)$, let $v_1, v_2, \ldots v_k$ be its children, let $\ell_i := \ell(u, v_i)$, and let $f_i := f_A(v_i)$. Then the points (ℓ_i, f_i) constitute a two-dimensional inhomogeneous Poisson point process on the ℓf-square, with intensity given by measure $\mu_A := m_q \times F_A$ if $|u|$ is odd and $\mu_B := m_q \times F_B$ if $|u|$ is even.*

An immediate consequence of the lemma is that the f_A-optimal move from a vertex is a.s. unique: $\ell - f$ has continuous distribution, because its probability density function is given by the convolution of the function $t \mapsto qt_+^{q-1}$ and the measure $dF_A(-z)$.

To generate the tree T_λ^q concurrent with f_A, start by picking z according to the probability measure F_A, and assigning the root ϕ the game value $f_A(\phi) = z$. Then we generate the next generation of the tree by the Poisson point process of Lemma 4.1, conditioned on $\min(\lambda/2, \min_i(\ell_i - f_i)) = z$.

If $z = \lambda/2$, this is equivalent to conditioning on there being no point in the region $\{\ell - f < \lambda/2\}$ of the ℓf-square. Since the distribution of points in two disjoint regions are independent, the points in $\{\ell - f \geqslant \lambda/2\}$ are generated by an inhomogeneous Poisson point process according to the measure μ_B restricted to the region $\{\ell - f \geqslant \lambda/2\}$.

If $z < \lambda/2$, this is equivalent to conditioning on there being no point in the region $\{\ell - f < z\}$ and one special point on the line $\{\ell - f = z\}$. The points $\{\ell - f \geqslant z\}$ can be generated by restricting the intensity measure to $\{\ell - f \geqslant z\}$ as in the previous case. The line $\{\ell - f = z\}$ has zero μ_B-measure, so to pick a random point from it we condition on there being at least one point in the Poisson point process on the strip $\{z \leqslant \ell - f \leqslant z + \varepsilon\}$, and then let $\varepsilon \to 0$. Since $\mu_B = m_q \times F_B$, and m_q is absolutely continuous with respect to the Lebesgue measure, this is well-defined. In order to express the probability measure obtained in the limit explicitly (which we do in Lemma 4.4) we must first understand the measures F_A and F_B in more detail.

The following is proved in [14, page 1077] (as well as occurring in similar forms in *e.g.* [2], [8] and [11]), but we include the proof here because it helps in understanding some of our argument later on, in Lemma 4.8 and Lemma 4.7.

[f] The lemma in [14] only states that (ℓ_i, f_i) constitutes a Poisson point process, not what the intensity measure is. However, it is implicit in the proof of the lemma that μ_B is the correct measure when $|u|$ is even, and the other case is analogous.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Lemma 4.2. *Let V be the non-linear operator on functions on Λ defined by*

$$V(G)(z) := \exp\left(-\int_\Lambda q(z+t)_+^{q-1} G(t)\,dt\right).$$

Then $F_A = V(F_B)$ and $F_B = V(F_A)$.

Note that since $F_B(t) \leqslant 1$ for all t, the lemma implies that $F_A(t) \geqslant e^{-\lambda^q}$. Similarly, $F_B(t) \geqslant e^{-\lambda^q}$.

Proof. Recall that $F_A(z) := \mathbb{P}(f_A(\phi) \geqslant z)$. Since (by definition) $f_A(\phi) = \min(\lambda/2, \min_i(\ell_i - f_i))$, the event $\{f_A(\phi) \geqslant z\}$ happens if and only if there is no (ℓ_i, f_i) with $\ell_i - f_i < z$. By Lemma 4.1, the (ℓ_i, f_i) constitutes a Poisson point process, and the probability that no (ℓ_i, f_i) falls in the set $D_z := \{(\ell, f) \colon \ell - f < z\}$ is $\exp(-\mu_B(D_z))$. To calculate $\mu_B(D_z)$, first fix ℓ and let t be such that $z + t = \ell$. Then $\ell - f < z$ if and only if $f > t$. Integrating over all t gives

$$\mu_B(D_z) = \int_\Lambda q(z+t)_+^{q-1}\mathbb{P}(f > t)\,dt.$$

However, $\mathbb{P}(f > t) = \mathbb{P}(f \geqslant t)$ for all but at most countably many t, and $F_B(t) = \mathbb{P}(f \geqslant t)$ by definition. Hence $F_A(z) = \exp(-\mu_B(D_z)) = V(F_B)(z)$. The other case is analogous. □

The operator V is the composition of an integral operator with a continuous kernel and a smooth function applied pointwise, and we can use this to establish smoothness properties of F_A and F_B, as well as bound their derivatives.

Lemma 4.3. *Each of the two measures given by F_A and F_B on Λ is the sum of a point mass at $\lambda/2$ and a measure that is absolutely continuous with respect to the Lebesgue measure on $(-\lambda/2, \lambda/2)$. The functions F_A and F_B are continuously differentiable on $(-\lambda/2, \lambda/2)$, with derivative F_A' given by*

$$F_A'(z) = -F_A(z)\cdot\left(F_B(\lambda/2)\,q(z+\lambda/2)^{q-1} - \int_\Lambda q(z+t)_+^{q-1}F_B'(t)\,dt\right). \tag{4.1}$$

Furthermore, $\int_\Lambda q(z+t)_+^{q-1}F_A'(t)\,dt$ is a continuous function of z, and for some constant $\alpha > 1$ and all $|z| < \lambda/2$, we have the bounds

$$-F_A'(z) \leqslant \alpha(\lambda/2 - |z|)^{q-1}, \tag{4.2}$$

$$-\int_\Lambda q(z+t)_+^{q-1}F_A'(t)\,dt \leqslant \alpha\max\left((z+\lambda/2)^{2q-1}, |z-\lambda/2|^{q-1}\right). \tag{4.3}$$

Equation (4.3) also holds for $\lambda/2 < z \leqslant 3\lambda/2$, and analogous results hold for F_B'.

The proof of this lemma is largely a lengthy calculus exercise, and we postpone it to the end of the paper. We will often parametrize the diagonal line $\{(\ell, f) \colon \ell - f = z\}$ as $\{(z+t, t) \colon t \in \Lambda\}$. The measure μ_B has density

$$\rho_B^z(t) := q(z+t)_+^{q-1}\cdot(-F_B'(t)) \tag{4.4}$$

along such a diagonal for $t < \lambda/2$, and a point mass $q(z+\lambda/2)^{q-1}F_B(\lambda/2)$ at the end point $t = \lambda/2$ (and analogously for μ_A, ρ_A^z).

Lemma 4.4. *Let (ℓ, f) be a point in the inhomogeneous Poisson point process on the ℓf-square with intensity measure μ_A, conditioned to lie on the line $\ell - f = z$ (for some $z < \lambda/2$). Then the probability distribution of f is given by*

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

$$\mathbb{P}(f < x) = \frac{\int_{-\lambda/2}^{x} \rho_A^z(t)\,dt}{J_A^z}, \quad \mathbb{P}(f = \lambda/2) = \frac{q(z+\lambda/2)^{q-1} F_A(\lambda/2)}{J_A^z},$$

where

$$J_A^z := q(z+\lambda/2)^{q-1} F_A(\lambda/2) + \int_\Lambda \rho_A^z(t)\,dt.$$

Proof. Let η_ε be defined as ε^{-1} times the measure μ_A, restricted to the region

$$E_\varepsilon := \{(\ell, f)\colon z \leqslant \ell - f \leqslant z + \varepsilon, f \geqslant -z\},$$

and let η be the measure on E_0 which is given by $\rho_A^z(t)\,dt$ at the point $(z, z+t)$, and a point mass of $F_A(\lambda/2)$ at $(z, z+\lambda/2)$.

We will show that $\eta_\varepsilon \to \eta$, as $\varepsilon \to 0$, and that normalizing η gives the probability measure in the statement of the lemma. For $z < \lambda/2$,

$$\iint_{f<x} d\eta_\varepsilon = \int_{-\lambda/2}^{x} \int_{f+z}^{f+z+\varepsilon} \frac{1}{\varepsilon} \rho_A^{\ell-f}(f)\,d\ell\,df = -\int_{-\lambda/2}^{x} \frac{1}{\varepsilon} \int_{f+z}^{f+z+\varepsilon} q\ell^{q-1}\,d\ell f_A'(f)\,df.$$

Note that $\ell \mapsto q\ell^{q-1}$ is decreasing on $\mathbb{R}_+$, whence

$$\frac{1}{\varepsilon} \int_{f+z}^{f+z+\varepsilon} q\ell^{q-1}\,d\ell \nearrow q(z+f)^{q-1}.$$

By the monotone convergence theorem,

$$\iint_{f<x} d\eta_\varepsilon \to \int_{-\lambda/2}^{x} \rho_A^z(t)\,dt.$$

Similarly, the η_ε-measure of the line segment $\{f = \lambda/2, z \leqslant \ell - f \leqslant z + \varepsilon\}$ approaches $q(z+\lambda/2)^{q-1} F_A(\lambda/2)$ as $\varepsilon \to 0$. So $\eta_\varepsilon \to \eta$ as $\varepsilon \to 0$, and $J_A^z := \|\eta\| \in (0, \infty)$. (For a measure m, $\|m\|$ denotes the m-measure of the whole space on which m is defined.) It follows that $\eta_\varepsilon/\|\eta_\varepsilon\| \to \eta/\|\eta\|$, and $\eta_\varepsilon/\|\eta_\varepsilon\|$ is the probability measure for a random point picked according to μ_A in E_ε. □

We will also use inequality (4.3) of Lemma 4.3 in another (weaker) form, as a bound on the normalizing factor J_A^z:

$$\begin{aligned} J_A^z &\leqslant \alpha \max\left((z+\lambda/2)^{2q-1}, |z-\lambda/2|^{q-1}\right) + F_A(\lambda/2) q (z+\lambda/2)^{q-1} \\ &< \alpha\lambda^q \max\left((z+\lambda/2)^{q-1}, |z-\lambda/2|^{q-1}\right). \end{aligned} \tag{4.5}$$

4.2 (u, t)-reasonable moves

We will now introduce our new definition of reasonable moves, and show that the game path is reasonable according to this definition. For $v \in T_\lambda^q - \{\phi\}$, let $\delta(v)$ be how far from f_A-optimal it is to move to v from its parent u. More precisely, $\delta(v) := \ell(u, v) - f_A(u) - f_A(v)$. Note that $\delta(v) \geqslant 0$, since it follows from (3.2) that $f_A(u) \leqslant \ell(u, v) - f_A(v)$ for any v.

Definition 4.1. *We say that a (finite or infinite) path $P = uu_1u_2\ldots$ away from the root is (u, t)-*reasonable *if $\sum_{i=1}^{|P|} \delta(u_i) \leqslant t$ and $\delta(u_i) = 0$ whenever $|u_i|$ is odd. In other words, a path is (u, t)-reasonable if Alice's moves are f_A-optimal and Bob's deviations from f_A sum to at most t.*

Lemma 4.5. *The game path (when Alice plays according to f_A and Bob according to f_B) is $(\phi, 2\lambda)$-reasonable.*

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Proof of Lemma 4.5. Let P be the game path. Pick any length 2 subpath $(u \to v \to w) \subseteq P$, such that u is at even distance from ϕ.

Since u is at even distance from the root, it will be Alice's turn to move from u. She will choose the f_A-optimal move, *i.e.* she will move to a child v of u such that $f_A(u) = \ell(u, v) - f_A(v)$. In other words, $\delta(v) = 0$. This move may or may not be f_B-optimal, but $f_B(u) \leqslant \ell(u, v) - f_B(v)$ regardless. Thus[g]

$$f_A(u) - f_B(u) \geqslant [\ell(u, v) - f_A(v)] - [\ell(u, v) - f_B(v)] = f_B(v) - f_A(v),$$

Then it will be Bob's turn to move from v. He will choose the f_B-optimal move, *i.e.* he will move to a child w of v such that $f_B(v) = \ell(v, w) - f_B(v)$. This move may or may not be f_A-optimal, but by the definition of δ we have that $f_A(v) = \ell(v, w) - f_A(w) - \delta(w)$. Thus

$$f_B(v) - f_A(v) = [\ell(v, w) - f_B(w)] - [\ell(v, w) - f_A(w) - \delta(w)] = f_A(w) - f_B(w) + \delta(w),$$

and together with the move $u \to v$ and the fact that $\delta(v) = 0$, this gives that

$$f_A(u) - f_B(u) \geqslant f_A(w) - f_B(w) + \delta(v) + \delta(w).$$

Let $\phi = u_0 \to u_1 \to u_2 \to \cdots$ be the game path P. Pick $n \in \mathbb{N}$ such that $2n \leqslant |P|$ (P might be infinite, in which case we just pick any $n \in \mathbb{N}$). If we repeat the argument above with $(u, v, w) := (u_{2i-2}, u_{2i-1}, u_{2i})$, for all $1 \leqslant i \leqslant n$, we get that

$$\begin{aligned} f_A(\phi) - f_B(\phi) &\geqslant f_A(u_2) - f_B(u_2) + \delta(u_1) + \delta(u_2) \\ &\geqslant f_A(u_4) - f_B(u_4) + \delta(u_1) + \delta(u_2) + \delta(u_3) + \delta(u_4) \\ &\ \ \vdots \\ &\geqslant f_A(u_{2n}) - f_B(u_{2n}) + \sum_{i=1}^{2n} \delta(u_i). \end{aligned}$$

Since $|f_A|, |f_B| \leqslant \lambda/2$, this implies that

$$\sum_{i=1}^{2n} \delta(u_i) \leqslant 2\lambda.$$

Recall that $\delta(u_i) = 0$ for odd i, whence

$$\sum_{i=1}^{k} \delta(u_i) \leqslant 2\lambda \quad \text{for any } k \leqslant |P|.$$

Taking the supremum over such k, it follows that

$$\sum_{i=1}^{|P|} \delta(u_i) \leqslant 2\lambda. \qquad \square$$

Let $\Delta_t(u)$ be the union of all (u, t)-reasonable paths. The crucial property of $\Delta_t(u)$ is that the event $\{w \in \Delta_t(u)\}$ is determined by the first $|w|$ generations of T_λ^q together with vertex labels given by f_A. In other words, this event is independent from the descendants of w. The event $\{w \in P\}$, on the other hand, depends on both $f_A(v)$ and $f_B(v)$ for v descendants of w, and we do not even know *a priori* if $f_B(v)$ is determined by *any* finite subtree of T_λ^q.

But our aim is to bound $\mathbb{E}|P|$, and since $P \subseteq \Delta_{2\lambda}(\phi)$ it suffices to bound $\mathbb{E}|\Delta_{2\lambda}(\phi)|$. We will work with k-level truncations $\Delta_t^k(u) := \Delta_t(u)(k, \lambda)$, and recursively bound the expected value of

[g] A similar argument is used in [14, page 1076] to show that the difference $f_A(u_{2k}) - f_B(u_{2k})$ is monotone in k.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

$|\Delta_t^k(u)|$. Conditional on $f_A(u)$, the distribution of $\Delta_t^k(u)$ is the same for every u at even distance from the root, so we let

$$R_t^k(z) := \mathbb{E}[|\Delta_t^k(\phi)| \mid f_A(\phi) = z]. \tag{4.6}$$

We are now ready to state the following proposition, which essentially says that the tree of reasonable paths is finite almost surely.

Proposition 4.6. *There exists a family of continuous functions $(\psi_t)_{t\in[0,2\lambda]}$ on Λ such that $R_t^{2k}(z) < \psi_t(z)$ for all $z \in \Lambda$, $t \in [0, 2\lambda]$, and $k \in \mathbb{N}$, and satisfying $\sup_{z,t} \psi_t(z) < \infty$. In particular, $\mathbb{E}|\Delta_{2\lambda}(\phi)| < \sup_z \psi_{2\lambda}(z)$ is finite.*

4.3 Linear operators

To prove Proposition 4.6 we will need the following lemmas concerning certain linear operators. These operators relate functions of the form $z \mapsto \mathbb{E}[\bullet \mid f_A(u) = z]$ to $z \mapsto \mathbb{E}[\bullet \mid f_A(v) = z]$, whenever $u \to v$ is an f_A-optimal move.

Recall that

$$J_A^z := q(z + \lambda/2)^{q-1} F_A(\lambda/2) - \int_\Lambda q(z + t)_+^{q-1} \cdot F_A'(t)\, dt$$

is the measure of the diagonal line $\{\ell - f = z\}$ in the ℓf-square, according to the measure from Lemma 4.4.

Lemma 4.7. *Let the positive linear operator L_A on $\mathcal{C}(\Lambda)$ and the function $I_A \colon \Lambda \to [0, 1]$ be defined by*

$$L_A h(z) := \frac{\int_\Lambda h(t)\rho_A^z(t)\, dt}{J_A^z}, \tag{4.7}$$

$$I_A(z) := \frac{\int_\Lambda \rho_A^z(t)\, dt}{J_A^z} \tag{4.8}$$

on $(-\lambda/2, \lambda/2)$, and by their continuous extensions at $\pm\lambda/2$. Let L_B and I_B be defined similarly. Let u, v be such that $\phi \to u \to v$ are f_A-optimal moves. Then the following holds:

$$(L_B \circ L_A) R_t^k(z) = \mathbb{E}[|\Delta_t^k(v)| \mid f_A(\phi) = z]. \tag{4.9}$$

Furthermore, I_A satisfies these properties:

(i) *I_A is continuous,*
(ii) *$I_A(z) < 1$ for $z \in [-\lambda/2, \lambda/2)$, and*
(iii) *$I_A(\pm\lambda/2)$ are well-defined by continuous extension.*

Analogous statements hold for I_B.

Lemma 4.8. *$\|L_B \circ L_A\| < 1$, where $\|\cdot\|$ is the operator norm given by the ∞-norm on $\mathcal{C}(\Lambda)$.*

Proof of Lemma 4.7. Assume that the moves $\phi \to u$ and $u \to v$ are f_A-optimal. Let $Z_z := (f_A(v) \mid f_A(u) = z)$, and consider $\mathbb{E}[|\Delta_t^k(v)| \mid f_A(u) = z]$. Since (by definition)

$$R_t^k(z) = \mathbb{E}[|\Delta_t^k(v)| \mid f_A(v) = z],$$

we can write

$$\mathbb{E}[|\Delta_t^k(v)| \mid f_A(u) = z] = \mathbb{E} R_t^k(Z_z).$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

First, note that the operator that takes a function g to $z \mapsto \mathbb{E}g(Z_z)$ is linear: For any fixed z, the mapping $g \mapsto \mathbb{E}g(Z_z)$ is a linear functional, so the function $z \mapsto \mathbb{E}g(Z_z)$ depends linearly on the function g.

Now let us use the distribution of Z_z given by Lemma 4.4 to calculate $\mathbb{E}R_t^k(Z_z)$. Integrating over Λ gives that

$$\mathbb{E}R_t^k(Z_z) = \frac{q(z+\lambda/2)^{q-1}F_A(\lambda/2)R_t^k(\lambda/2) + \int_\Lambda R_t^k(s)\rho_A^z(s)\,ds}{J_A^z}.$$

Note that $R_t^k(\lambda/2) = 0$, since Alice's optimal move from a vertex with game value $\lambda/2$ will be to quit immediately. Thus

$$\mathbb{E}R_t^k(Z_z) = \int_\Lambda R_t^k(f)\rho_A^z(f)\,df/J_A^z,$$

which equals $L_A R_t^k(z)$ by (4.7). Note also that $\rho_A^z(t), \rho_B^z(t), J_A^z$ and J_B^z are positive for all z, t, so the operators L_A, L_B are positive. Applying the same method one more time gives the desired result for the first part of the lemma. For the second part, we verify that (i)–(ii) hold.

(i) The non-negative term $\int_\Lambda \rho_A^z(t)\,dt$ is continuous in z by Lemma 4.3, and so is the positive term $q(z+\lambda/2)^{q-1}F_A(\lambda/2)$. Hence both the numerator and denominator of (4.8) are continuous, and the denominator is non-zero, so I_A is continuous.

(ii) Both $q(z+\lambda/2)^{q-1}F_A(\lambda/2)$ and $\int_\Lambda \rho_A^z(t)\,dt$ are positive and finite for $|z| < \lambda/2$, so $I_A(z) < 1$ for such z.

(iii) Using (4.3), we see that for z near $-\lambda/2$,

$$I_A(z) = \frac{O((z+\lambda/2)^{2q-1})}{(z+\lambda/2)^{q-1} + O((z+\lambda/2)^{2q-1})} = O((z+\lambda)^q),$$

so that $\lim_{z\to-\lambda/2} I_A(z) = 0$. Near $\lambda/2$,

$$I_A(z) = \frac{\int_\Lambda \rho_A^z(t)\,dt}{q\lambda^{q-1}F_A(\lambda/2) + o(1) + \int_\Lambda \rho_A^z(t)\,dt} = 1 - \left(1 + \frac{\int_\Lambda \rho_A^z(t)\,dt}{q\lambda^{q-1}F_A(\lambda/2) + o(1)}\right)^{-1},$$

so $\lim_{z\to\lambda/2} I_A(z)$ will exist if $\lim_{z\to\lambda/2}\int_\Lambda \rho_A^z(t)\,dt$ exists (even if the latter limit is infinite).

The singularity of $\rho_A^z(t)$ at $t = -z$ moves as $z \to \lambda/2$, so we will instead work with the translate $\rho_A^z(t - z + \lambda/2)$ which has its singularity at $-\lambda/2$ for all z. Note also that

$$\int_\Lambda \rho_A^z(t)\,dt = \int_\Lambda \rho_A^z(t - z + \lambda/2)\,dt,$$

as the support of ρ_A^z is $[-z, \lambda/2] \subseteq \Lambda$, and translating by $-z + \lambda/2$ gives a function with support $[-\lambda/2, z] \subseteq \Lambda$. By (4.2), we have that

$$\rho_A^z(t - z + \lambda/2) = q(t+\lambda/2)^{q-1} \cdot F_A'(t - z + \lambda/2) \leqslant \begin{cases} \alpha q(t+\lambda/2)^{2q-2} & t \leqslant 0, \\ Kt > 0, & \end{cases}$$

for some constant K and all z sufficiently close to $\lambda/2$. Thus we have an upper bound on $\rho_A^z(t - z + \lambda/2)$ which is independent of z. For $q > 1/2$, this upper bound is integrable, and by dominated convergence it follows that

$$\lim_{z\to\lambda/2}\int_\Lambda \rho_A^z(t - z + \lambda/2)\,dt = \int_\Lambda \lim_{z\to\lambda/2}\rho_A^z(t - z + \lambda/2)\,dt = \int_\Lambda \rho_A^{\lambda/2}(t)\,dt < \infty.$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Hence $\lim_{z\to\lambda/2}\int_\Lambda \rho_A^z(t)\,dt$ exists (and is finite) for $q > 1/2$. For $q \leqslant 1/2$, we use (4.1) of Lemma 4.3 to lower-bound $-F_A'(t)$,

$$-F_A'(t) \geqslant F_A(t)F_B(\lambda/2)q(t+\lambda/2)^{q-1} \geqslant K' \cdot (t+\lambda/2)^{q-1}$$

for some $K' := qe^{-2\lambda^q} > 0$, whence

$$\int_\Lambda \rho_A^z(t)\,dt = \int_\Lambda q(t+z)^{q-1}\cdot(-F_A'(t))\,dt \geqslant K'q\int_\Lambda (t+z)^{q-1}(t+\lambda/2)^{q-1}\,dt.$$

The right-hand side goes to ∞ as $z \to \lambda/2$, since the singularity $(t+\lambda/2)^{2q-2}$ is not integrable. We conclude that $\lim_{z\to\lambda/2}\int_\Lambda \rho_A^z(t)\,dt$ exists for all q (finite for $q > 1/2$, infinite for $q \leqslant 1/2$), hence $I_A(\lambda/2)$ is well-defined. □

Remark 4.1. It follows from the proof of the previous lemma that $I_A(\lambda/2) = 1$ for $q \leqslant 1/2$, while $I_A(\lambda/2) < 1$ for $q > 1/2$. This implies that statement (i) in Section 3.3 is true only if $q > 1/2$.

Proof of Lemma 4.8. L_A is a substochastic operator,[h] and to be able to fully leverage this property we will factorize it into a stochastic operator that has almost all the structure of L_A and a sub-stochastic operator that is also a diagonal map. Start by defining the kernel $\kappa_A^z(t)$, as ρ_A^z normalized for $(z,t) \in (-\lambda/2, \lambda/2)^2$:

$$\kappa_A^z(t) := \frac{\rho_A^z(t)}{\int_\Lambda \rho_A^z(s)\,ds}. \tag{4.10}$$

Using this kernel, we write $L_A(h)(z)$ as $\int_\Lambda I_A(z)h(t)\kappa_A^z(t)\,dt$. The factor $I_A(z)$ does not depend on t, so it can be factored out of the integral. We can therefore write L_A as the composition of the operators S_A and D_A, defined by

$$S_A(h)(z) := \int_\Lambda h(t)\kappa_A^z(t)\,dt, \tag{4.11}$$

$$D_A(h)(z) := I_A(z)\cdot h(z). \tag{4.12}$$

For any function h, $\sup_t S_A(h)(t) \leqslant \sup_t h(t)$, so $\|S_A\| \leqslant 1$. Similarly, $\|S_B\| \leqslant 1$. In order to show that $\|L_B \circ L_A\| < 1$, we factorize $L_B \circ L_A$ into $D_B \circ S_B \circ D_A \circ S_A$. It then suffices to bound $\|D_A\|$, $\|D_B\|$ or $\|D_B \circ S_B \circ D_A\|$ away from 1, since by the definition of the operator norm and using that $\|S_A\|, \|S_B\| \leqslant 1$, we have

$$\|L_B \circ L_A\| \leqslant \|D_B \circ S_B \circ D_A\| \leqslant \|D_A\|\cdot\|D_B\|.$$

The proof of the lemma will be divided into two cases, depending on whether or not $I_A(\lambda/2) = I_B(\lambda/2) = 1$.

Case 1. $I_A(\lambda/2) < 1$ or $I_B(\lambda/2) < 1$. Assume without loss of generality that $I_A(\lambda/2) < 1$. Then $I_A(z) < 1$ for all z. By Lemma 4.7, I_A is a continuous function on a closed interval, so it attains its supremum τ, which must be less than 1. Thus $\|D_A\| \leqslant \tau < 1$.

Case 2. $I_A(\lambda/2) = I_B(\lambda/2) = 1$. $D_B \circ S_B \circ D_A$ is an integral operator with kernel given by $I_B(s)I_A(t)\kappa_B^s(t)$. In order to show that this integral operator has norm less than 1, we will bound the integral of its kernel along the line $s = z$, where $z \in (-\lambda/2, \lambda/2]$ is arbitrary but fixed. (We do not need to consider the case $z = -\lambda/2$, since $I_B(-\lambda/2) < 1$.)

[h] A positive linear operator T given by $T(h)(z) := \int h(t)\kappa(t,z)\,dt$ is said to be *stochastic* if $\int \kappa(t,z)\,dt = 1$ for every z and *substochastic* if $\int \kappa(t,z)\,dt \leqslant 1$ for every z.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

We have a good upper bound on I_A and I_B on any closed set not containing $z = \lambda/2$; in particular, on $[-\lambda/2, 0]$. We therefore bound the total mass of κ_B^z on $[0, \lambda/2]$. Since $I_B(\lambda/2) = 1$, we know that $\int_\Lambda \rho_B^z(t)\, dt \to \infty$ as $z \to \lambda/2$. But

$$\int_0^{\lambda/2} \rho_B^z(t)\, dt \leqslant q\lambda^{q-1} \quad \text{for any } z,$$

so $\int_0^{\lambda/2} \kappa_B^z(t)\, dt$ must vanish as $z \to \lambda/2$. Hence there exists $0 < \theta < \lambda/2$ such that, for all $z > \theta$,

$$\int_{-z}^{0} \kappa_B^z(t)\, dt > 1/2. \tag{4.13}$$

By (ii) of Lemma 4.7, we can find $\delta > 0$, such that when $t \leqslant \theta$ we have

$$I_A(t) < 1 - \delta \quad \text{and} \quad I_B(t) < 1 - \delta. \tag{4.14}$$

Then, for $-\lambda/2 \leqslant z \leqslant \theta$, we apply the bound to I_B to get

$$\int_\Lambda I_B(z) I_A(t) \kappa_B^z(t)\, dt \leqslant I_B(z) \cdot \int_\Lambda \kappa_B^z(t)\, dt \overset{(4.14)}{<} 1 - \delta,$$

while for $\theta < z \leqslant \lambda/2$ we apply it to I_A:

$$\int_\Lambda I_B(z) I_A(t) \kappa_B^z(t)\, dt \overset{(4.14)}{<} \int_{-z}^{0} (1-\delta)\kappa_B^z(t)\, dt + \int_0^{\lambda/2} \kappa_B^z(t)\, dt \overset{(4.13)}{<} 1 - \delta/2.$$

This means that the weight along each line of $D_B \circ S_B \circ D_A$ is at most $1 - \delta$ when $z \leqslant \theta$, and at most $1 - \delta/2$ otherwise. So in either case, $\|D_B \circ S_B \circ D_A\| \leqslant 1 - \delta/2$.

We conclude that $\|L_B \circ L_A\| < \max(\tau, 1 - \delta/2) < 1$. □

4.4 The game finishes in finite time

Armed with Lemmas 4.3, 4.7 and 4.8, we can proceed with the proof of Proposition 4.6.

Proof of Proposition 4.6. We will use the fact that the positive operator $L := L_B \circ L_A$ is a contraction (*i.e.* $\|L\| < 1$) to construct suitable ψ_t. Let $\mathbf{1}\colon \Lambda \to \mathbb{R}$ denote the function with constant value 1. We define the functions ψ_t (for any $t \in [0, 2\lambda]$ and some large constants $K, m > 0$ to be determined later) by

$$\psi_t := K \exp(mt) \cdot \sum_{k=0}^{\infty} L^k \mathbf{1}. \tag{4.15}$$

By Lemma 4.8, $\|L\| < 1$, so the above series is absolutely convergent, whence for all $z \in \Lambda$

$$1 \leqslant \frac{\psi_t(z)}{K \exp(mt)} \leqslant \frac{1}{1 - \|L\|}. \tag{4.16}$$

Because the series is absolutely convergent, we can apply the operator L to ψ_t by applying it termwise to the sum in (4.15):

$$L\psi_t(z) = K \exp(mt) \cdot \sum_{k=1}^{\infty} L^k \mathbf{1}(z) = \psi_t(z) - K \exp(mt). \tag{4.17}$$

Crucially, $L\psi_t$ is less than ψ_t, and with a sizeable margin. We will do induction on even k to establish the main claim, that

$$R_t^k(z) := \mathbb{E}[|\Delta_t^k| \mid f_A(\phi) = z] \leqslant \psi_t(z).$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

For $z=\lambda/2$ this is trivial, since there is no f_A-optimal move from ϕ when $f_A(\phi)=\lambda/2$ and hence $R_t^{2k}(\lambda/2)=0$. We therefore fix some $-\lambda/2\leqslant z<\lambda/2$ and $0\leqslant t\leqslant 2\lambda$, and consider the first two moves of the game conditional on $f_A(\phi)=z$.

Let $\phi\to u$ be the f_A-optimal move, which exists since $z<\lambda/2$. Let $u\to v_0$ be the f_A-optimal move from u, if such u_0 exists. Otherwise, add a dummy vertex u_0 without descendants.

Assume there are N of Bob's move options from u that are sub-optimal with respect to f_A but within t of being f_A-optimal. Let v_i, $1\leqslant i\leqslant N$, be the vertices those moves lead to, $t_i:=\delta(v_i)$, ℓ_i the cost of the edge (u,v_i), and let $f_i:=f_A(v_i)$. By assumption, $t_i\leqslant t$ for all $1\leqslant i\leqslant N$. Note that t_i, v_i, f_i and N are random functions of z.

For the base case $k=2$, the children of the root in $\Delta_t^2(\phi)$ are the f_A-optimal moves (by definition), and there is almost surely at most one such move (u). Its expected number of children, conditioned on any value of $f_A(u)$, is at most $1+\lambda^q$, so $R_t^2\leqslant 2+\lambda^q$. We let $K=2(2+\lambda^q)$, so that $R_t^2\leqslant K/2<\psi_t$, establishing the base case.

Next, assume $R_s^k<\psi_s$ for some even $k>2$ and all $0\leqslant s\leqslant 2\lambda$. We want to show that $R_t^{k+2}<\psi_t$ as well, and to do that we will bound the expected size of Δ_t^{k+2}. The tree Δ_t^{k+2} can be written as an edge-disjoint union of copies of Δ in the following way:

$$\Delta_t^{k+2}(\phi)=\Delta_t^2(\phi)\cup\Delta_t^k(v_0)\cup\bigcup_{i=1}^N\Delta_{t-t_i}^k(v_i). \tag{4.18}$$

Note that the trees $\Delta_t^k(v_0)$ and $\Delta_{t-t_i}^k(v_i)$, $1\leqslant i\leqslant N$, are independent conditional on $f_A(v_0), f_A(v_1),\ldots,f_A(v_N)$. We already have a bound for $\Delta_t^2(\phi)$, and we continue by bounding the conditional expected sizes of $\Delta_t^k(v_0)$ and $\bigcup_{i=1}^N\Delta_{t-t_i}^k(v_i)$. For $\Delta_t^k(v_0)$, by the definition of R_t^k and Lemma 4.7,

$$\begin{aligned}\mathbb{E}[|\Delta_t^k(v_0)|\mid f_A(\phi)=z]&=L(R_t^k)(z)\\&<L(\psi_t)(z),\quad\text{since } L \text{ is a positive operator}\\&\overset{(4.17)}{=}\psi_t(z)-K\exp(mt).\end{aligned} \tag{4.19}$$

Next we bound the expected size of the union of the trees $\Delta_{t-t_i}^k(v_i)$. To do this we condition first on the random variables N, f_i and t_i and then on the event $f_A(\phi)=z$, so that the first conditional expectation is itself a random variable:

$$\begin{aligned}\mathbb{E}\left[\left|\bigcup_{i=1}^N\Delta_{t-t_i}^k(v_i)\right|\mid f_A(\phi)=z\right]&=\mathbb{E}\left[\mathbb{E}\left[\left|\bigcup_{i=1}^N\Delta_{t-t_i}^k(v_i)\right|\mid N,t_i,f_i,1\leqslant i\leqslant N\right]\mid f_A(\phi)=z\right]\\&=\mathbb{E}\left[\sum_{i=1}^N R_{t-t_i}^k(f_i)\mid f_A(\phi)=z\right],\end{aligned} \tag{4.20}$$

since the subtree rooted in v_i, conditional on $f_A(v_i)$, is independent of $f_A(\phi)$. By the induction hypothesis with $s=t-t_i$,

$$\sum_{i=1}^N R_{t-t_i}^k(f_i)\leqslant\sum_{i=1}^N\psi_{t-t_i}(f_i).$$

Let σ_A be the Poisson random measure generated by μ_A (*i.e.* the counting measure of the points of the Poisson point process with intensity μ_A). It is a sum of Dirac measures, each corresponding to a point in the ℓf-square. Among these points, (ℓ_i,f_i), $1\leqslant i\leqslant N$, are exactly those that lie in the diagonal strip

$$D:=\{(\ell,f):\ z<\ell-f\leqslant z+t\}.$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Note that for any bounded μ_A-measurable function h, we have that $\mathbb{E}[\int h\, d\sigma_A] = \int h\, d\mu_A$ (which can be seen by approximating h by simple functions). The expression (4.20) is then at most

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{N} \psi_{t-t_i}(f_i) \,|\, f_A(\phi) = z\right] &= \mathbb{E}\left[\iint_D \psi_{z+t-l+f}(f)\, d\sigma_A(\ell, f)\right] \\ &= \iint_D \psi_{z+t-\ell+f}(f)\, d\mu_A(\ell, f) \\ &\overset{(4.16)}{\leqslant} \iint_D \frac{K \exp\left(m(z+t-\ell+f)\right)}{1-\|L\|}\, d\mu_A(\ell, f) \end{aligned} \tag{4.21}$$

The integrand is constant along diagonals $\ell - f = x$ for fixed $x \in (z, z+t]$. Recall that the one-dimensional measure of such a diagonal is J_A^x. Integrating along these diagonals first, we see that

$$\begin{aligned}(4.21) &\leqslant \frac{K}{1-\|L\|} \cdot \int_z^{z+t} J_A^x \exp\left(m(z+t-x)\right) dx \\ &\overset{(4.5)}{\leqslant} \frac{K \exp(mt)}{1-\|L\|} \cdot \int_z^{z+t} \alpha\lambda^q \cdot \left[(x+\lambda/2)^{q-1} + |x-\lambda/2|^{q-1}\right] \exp\left(m(z-x)\right) dx \\ &\leqslant K \exp(mt) \cdot \varepsilon_m \end{aligned} \tag{4.22}$$

for some ε_m which goes to zero as $m \to \infty$, and does not depend on k, t or z. We now have a bound on the expected size of each term in the right-hand side of (4.18). The bounds from (4.22) and (4.19) give that

$$\begin{aligned}R_t^{k+2}(z) &= \mathbb{E}\left[|\Delta_t^2(\phi)| + |\Delta_t^k(v)| + \sum_i |\Delta_{t-t_i}^k(v_i)| \,|\, f_A(\phi) = z\right] \\ &< (K/2) + (\psi_t(z) - K\exp(mt)) + (K\exp(mt)\varepsilon_m). \end{aligned} \tag{4.23}$$

Pick m large enough that $\varepsilon_m < 1/2$. The expression (4.23) is then at most $\psi_t(z)$, completing the inductive step. Hence $R_t^k \leqslant \psi_t$ for all even k and all $t \in [0, 2\lambda]$. □

Proof of Proposition 3.1. By Lemma 4.5, the game path P is $(\phi, 2\lambda)$-reasonable, and is therefore contained in the tree $\Delta_{2\lambda}(\phi)$ of *all* $(\phi, 2\lambda)$-reasonable paths. By Proposition 4.6, $\Delta_{2\lambda}(\phi)$ is almost surely finite, and hence the game finishes after finitely many steps. Thus P is the finite path $\phi = u_0 \to u_1 \to \cdots \to u_N$ for some u_i. For $1 \leqslant i \leqslant N$, let $\ell_i = \ell(v_{i-1}, v_i)$. Let S be the total payoff for Alice. (The total payoff for Bob is then $-S$.) Alice pays $\ell_1 + \ell_3 + \cdots$ to Bob, and Bob pays $\ell_2 + \ell_4 + \cdots$ to Alice, until one player decides to pay $\lambda/2$ and quit the game. Thus

$$S = -\ell_1 + \ell_2 - \ell_3 \cdots \pm \ell_N \mp \lambda/2,$$

where the $\pm$-sign depends on whether Alice or Bob is the one to quit, *i.e.* whether N is even or odd.

Claim 1. $f_A(\phi) \geqslant -S$

Proof of claim. Recall that $f_A(u_{i-1}) \leqslant \ell_i - f_A(u_i)$, with equality if $u_{i-1} \to u_i$ is f_A-optimal (which is always the case if i is odd). Using these inequalities along the game path P gives us

$$f_A(\phi) = \ell_1 - f_A(u_1) \geqslant \ell_1 - \ell_2 + f_A(u_2) = \cdots \geqslant \sum_{i=1}^{N} (-1)^{i+1} \ell_i + (-1)^N f_A(u_N).$$

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

If N is even, then Alice is the one that quits, which she would only have done if $f_A(u_N) = \lambda/2$. In that case

$$S = -\lambda/2 + \sum_{i=1}^{N} (-1)^i \ell_i = -f_A(u_N) + \sum_{i=1}^{N} (-1)^i \ell_i.$$

If on the other hand N is odd, then Bob is the one that quits, and (as for every vertex) $f_A(u_N) \leqslant \lambda/2$. Hence

$$S = \lambda/2 + \sum_{i=1}^{N} (-1)^i \ell_i \geqslant f_A(u_N) + \sum_{i=1}^{N} (-1)^i \ell_i.$$

In either case

$$-S \leqslant \sum_{i=1}^{N} (-1)^{i+1} \ell_i + (-1)^N f_A(u_N) \leqslant f_A(\phi).$$ □

By symmetry (reversing the roles of Alice and Bob in the proof, and δ instead measuring how far Alice deviates from what is f_B-optimal) we also have that $-f_B(\phi) \geqslant S$. Thus $f_A(\phi) \geqslant -S \geqslant f_B(\phi)$. But by the choice of f_A and f_B, we know that $f_A(\phi) \leqslant f_B(\phi)$, so we have that $f_A(\phi) = f_B(\phi)$. For any other $u \in V(T_\lambda^q)$, the subtree rooted in u has the same distribution as the whole T_λ^q, so a similar argument gives that $f_A(u) = f_B(u)$. Hence $f_A = f_B$. Since f_A and f_B are the maximum and minimum, respectively, in the lattice ordering of all valuations, this implies that the valuation is unique. □

We end the paper by giving the deferred proof of Lemma 4.3.

Proof of Lemma 4.3. Let $\mathcal{G}$ be the class consisting of all non-increasing functions $G\colon \Lambda \to (0, 1]$ which satisfy $G(-\lambda/2) = 1$. Recall that V is the non-linear operator defined by

$$V(G)(z) := \exp\left(-\int_\Lambda q(z+t)_+^{q-1} G(t)\, dt\right).$$

We will find the derivative of $V(G)$ for any $G \in \mathcal{G}$, and then show that $F_A, F_B \in \mathcal{G}$. Since $F_A = V(F_B)$ and $F_B = V(F_A)$ [14, page 1077], this will give us the derivatives F_A' and F_B'.

Claim 2. $V(\mathcal{G}) \subseteq \mathcal{G}$.

Proof. For any function $G\colon \Lambda \to \mathbb{R}$, we have that $V(G)(-\lambda/2) = 1$, because

$$V(G)(-\lambda/2) := \exp\left(-\int_\Lambda q(-\lambda/2+t)_+^{q-1} G(t)\, dt\right)$$

and $(-\lambda/2+t)_+^{q-1}$ vanishes for all $t \in \Lambda$.

Now pick a $G \in \mathcal{G}$. Note first that $V(G)$ is non-increasing since G is non-increasing and positive. Furthermore, $G \geqslant 0$ implies $V(G)(z) \leqslant 1$, and similarly $G \leqslant 1$ implies

$$V(G)(z) \geqslant \exp\left(-\int_\Lambda q(z+t)_+^{q-1}\, dt)\right) \geqslant \exp(-\lambda^q) > 0.$$

Thus $V(G) \in \mathcal{G}$, and the claim follows. □

Claim 3. $F_A, F_B \in \mathcal{G}$

Proof. First note that F_A and F_B are non-increasing by definition. Next recall that $F_A = V(F_B)$, $F_B = V(F_A)$ and that $V(G)(-\lambda/2) = 1$ for any real-valued function on Λ, whence $F_A(-\lambda/2) = F_B(-\lambda/2) = 1$. □

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Since any $G \in \mathcal{G}$ is bounded and monotone, it has bounded variation. We can therefore integrate with respect to the measure dG, in the sense of a Riemann–Stieltjes integral. However, Riemann–Stieltjes integration is usually defined for non-decreasing functions rather than the non-increasing function G here, and we must be careful with how Riemann–Stieltjes treats the endpoints of Λ. We therefore let $\tilde{G}$ be defined by $\tilde{G} := 1 - G$ on $[-\lambda/2, \lambda/2)$ and $\tilde{G}(\lambda/2) := 1$, and work with $d\tilde{G}$ rather than dG.

Claim 4. *For any $G \in \mathcal{G}$, $V(G)$ is differentiable on the interior of Λ, with derivative given by*

$$\frac{d}{dz} V(G)(z) = V(G)(z) \cdot \int_\Lambda q(z+t)_+^{q-1}\, d\tilde{G}(t). \tag{4.24}$$

Proof. To verify (4.24), start by integrating $\int_\Lambda q(z+t)_+^{q-1}\, d\tilde{G}(t)$ from $z = -\lambda/2$ to x (for some x with $|x| < \lambda/2$):

$$\begin{aligned}
\int_{-\lambda/2}^{x} \int_{-\lambda/2}^{\lambda/2} q(z+t)_+^{q-1}\, d\tilde{G}(t)\, dz &= \iint_{\substack{-\lambda/2 \leqslant t \leqslant \lambda/2,\\ -\lambda/2 \leqslant s-t \leqslant x}} q s_+^{q-1}\, d\tilde{G}(t)\, ds \\
&= \int_0^{\lambda/2 - x} q s^{q-1} G(x+s)\, ds \\
&= -\ln(V(G)(x)).
\end{aligned}$$

By the fundamental theorem of calculus, $\ln(V(G)(z))$ is differentiable, with derivative given by

$$\frac{d}{dz} \ln(V(G)(z)) = -\int_\Lambda q(z+t)_+^{q-1}\, d\tilde{G}(t).$$

This implies that $V(G)$ is also differentiable, with derivative given by (4.24), proving the claim. □

Claim 5. *Let the function g on Λ be defined by*

$$g(z) := (\lambda/2 - |z|)^{q-1}. \tag{4.25}$$

Then there exists a constant $a > 0$ such that if $G \in V(\mathcal{G})$ satisfies $-G' \leqslant ag$, then $-(V(G))' \leqslant ag$.

Proof. We need to calculate (and then estimate) $\frac{d}{dz} V(G)(z)$. Since $G \in V(\mathcal{G})$, G is differentiable, and therefore $d\tilde{G}(t) = -G'(t)\, dt$ for t in the interior of Λ. However, $\tilde{G}$ also has a point mass at $\lambda/2$, so for any $t \in \Lambda$ we have that

$$d\tilde{G}(t) = -G'(t)\, dt + G(\lambda/2)\, d\delta_{\lambda/2}(t),$$

where δ_x is a Dirac measure at x. Substituting this expression for $d\tilde{G}$ in (4.24),

$$\frac{d}{dz} V(G)(z) = -V(G)(z) \cdot \left(q(\lambda/2 + z)^{q-1} G(\lambda/2) - \int_\Lambda q(z+t)_+^{q-1} G'(t)\, dt \right). \tag{4.26}$$

The integrand on the right-hand side of (4.26) has a singularity at $t = -z$, while the function g has a singularity at $t = \lambda/2$. For some positive parameter $r < \min(\lambda/4, 2^{-4/q})$, we will deal separately with two cases: when these singularities are within $2r$ of each other, and when they are further apart. We will establish that the following inequality holds in both cases:

$$-\frac{d}{dz} V(G)(z) < g(z) \cdot (2q + 4ar^q) + qr^{q-1} \quad \text{for all } z, \tag{4.27}$$

from which it follows that

$$-\frac{d}{dz}V(G)(z) < a \cdot g(z) \quad \text{for all } z$$

by picking some $a > \max(8q, \lambda/r)$.

Case 1. $-\lambda/2 \leqslant z \leqslant -\lambda/2 + 2r$. We apply the bound $-G'(t) \leqslant a \cdot g(t)$, and use the fact that the resulting integrand is symmetric around $t = -z/2 + \lambda/4$:

$$\begin{aligned} -\int_\Lambda G'(t) \cdot q(z+t)_+^{q-1}\, dt &\leqslant a \int_\Lambda q(\lambda/2 - t)^{q-1}(z+t)_+^{q-1}\, dt \\ &\leqslant 2a(z/2 + \lambda/4)^{q-1} \cdot \int_0^{z/2+\lambda/4} qs^{q-1}\, ds \\ &= 2^{2-2q}a(z+\lambda/2)^{2q-1} \qquad (4.28) \\ &< 4ar^q g(z). \qquad (4.29) \end{aligned}$$

We will later be using the tighter bound in (4.28), but for now (4.29) suffices. Again using (4.26), this gives a bound on $\frac{d}{dz}V(G)(z)$:

$$-\frac{d}{dz}V(G)(z) \leqslant G(z) \cdot (G(\lambda/2) \cdot q(\lambda/2 + z)_+^{q-1} + 4ar^q g(z)) \leqslant g(z) \cdot (q + 4ar^q)$$

which is less than the bound from (4.27).

Case 2. $-\lambda/2 + 2r \leqslant z \leqslant \lambda/2$. We use the bound $-G'(t) \leqslant a \cdot g(t)$ for $-z < t < -z + r$:

$$-\int_\Lambda G'(t) \cdot q(z+t)_+^{q-1}\, dt \leqslant \int_{-z}^{-z+r} ag(t) \cdot q(z+t)^{q-1}\, dt - \int_{-z+r}^{\lambda/2} G'(t) \cdot q(z+t)^{q-1}\, dt. \quad (4.30)$$

If $g(t)$ is larger than $g(-z)$, for $-z \leqslant t \leqslant -z + r$, it can be at most twice as large, since g is increasing fastest at $\lambda/2 - 2r$ and $g(\lambda/2 - r) \leqslant 2g(\lambda/2 - 2r)$. Hence the first integral on the right-hand side of (4.30) is at most

$$\int_{-z}^{-z+r} 2g(-z) \cdot q(z+t)^{q-1}\, dt \leqslant 2g(z) \cdot r^q, \quad (4.31)$$

while the second integral on the right-hand side of (4.30) is at most

$$-\int_{-z+r}^{\lambda/2} G'(t) \cdot qr^{q-1}\, dt \leqslant qr^{q-1}, \quad (4.32)$$

since $q(z+t)^{q-1}$ is a decreasing function in t. Putting (4.31) and (4.32) together with (4.26), this gives us that $-\frac{d}{dz}V(G)(z)$ is at most

$$V(G)(z) \cdot (G(\lambda/2)q(\lambda/2 + z)^{q-1} + 2a \cdot g(z)r^q + qr^{q-1}) \leqslant 2g(z) \cdot (q + ar^q) + qr^{q-1},$$

which is also less than the bound from (4.27). □

Claim 6. $-F'_A, -F'_B \leqslant ag$ *(inequality (4.2) in the statement of the lemma).*

Proof. Note first that $F_A, F_B \in V(\mathcal{G})$, whence they are differentiable by a previous claim. Let $G_1(z) := 1$ for all $z \in \Lambda$, and $G_{k+1} := V(G_k)$. Then $G_1 \in \mathcal{G}$, and by induction $G_k \in \mathcal{G}$ for all $k \geqslant 1$. We know by [14, page 1077] that for any $z \in \Lambda$, $G_{2k}(z) \nearrow F_A(z)$, and similarly $G_{2k+1}(z) \searrow F_B(z)$.

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

Λ is compact and F_A is continuous, so by Dini's theorem $G_{2k} \to F_A$ uniformly. But since F_A is differentiable, uniform convergence of G_{2k} implies $G'_{2k} \to F'_A$. Similarly, $G'_{2k+1} \to F'_B$ as $k \to \infty$.

Noting that $-G'_1(z) = 0 < ag$, and $-G'_k(z) < ag \Rightarrow -G'_{k+1}(z) < ag$, by induction $-G'_k(z) < ag$ for all k. Since $G'_{2k} \to F'_A$, we have that $-F'_A \leqslant ag$, and similarly $-F'_B \leqslant ag$. □

The next step is to show that $\int_\Lambda \rho_A^z(t)\,dt$ is continuous in z on $(-\lambda/2, \lambda/2)$. (Recall that $\rho_A^z(t) := -q(t+z)_+^{q-1} F'_A(t)$.) It suffices to show that it is continuous on any closed sub-interval $I \subset (-\lambda/2, \lambda/2)$. Since $|F'_A|$ is bounded by $K_I := \sup_{t\in I} ag(t) < \infty$ on I, for any $x, y \in I$ we have

$$|F_A(x) - F_A(y)| < K_I \cdot |x-y|. \tag{4.33}$$

We will let $\varepsilon := \sqrt{|x-y|} \to 0$. Suppose (without loss of generality) that $x < y$ and $[x-\varepsilon, y+\varepsilon] \subseteq I$. We estimate the difference

$$\begin{aligned}
&\left|\int_{-x}^{\lambda/2} q(x+t)^{q-1} F'_A(t)\,dt - \int_{-y}^{\lambda/2} q(y+t)^{q-1} F'_A(t)\,dt\right| \\
&\overset{(4.33)}{\leqslant} 2\left|\int_{-x}^{-x+\varepsilon} K_I \cdot q(x+t)^{q-1}\,dt\right| + q\left|\int_{-y+\varepsilon}^{\lambda/2} ((x+t)^{q-1} - (y+t)^{q-1}) F'_A(t)\,dt\right| \\
&\leqslant 2K_I\varepsilon^q + 2\underbrace{|x-y|}_{\leqslant \varepsilon^2} \cdot q(1-q) \cdot \left|\int_{-y+\varepsilon}^{\lambda/2} \underbrace{(y+t)^{q-2}}_{\leqslant \varepsilon^{q-2}} F'_A(t)\,dt\right| \\
&= O(\varepsilon^q).
\end{aligned}$$

Hence $\int_\Lambda \rho_A^z(t)\,dt$ is continuous in z, and so is F'_B.

To establish the bound (4.3) for $\int_\Lambda \rho_A^z(t)\,dt$, we use $-F'_A \leqslant ag$. Then F_A satisfies the conditions necessary for (4.28) to hold for z near $-\lambda/2$ with $G = F_A$. For other z, note that the integrand is at most $-F'_B(z)$, for which the weaker bound ag suffices. In other words, for some constant b and any $-\lambda/2 < z < \lambda/2$, we have that

$$\int_\Lambda \rho_A^z(t)\,dt \leqslant b \max\left((\lambda/2 - z)^{q-1}, (z+\lambda/2)^{2q-1}\right).$$

Finally, for $z > \lambda/2$, note that $(z+t)^{q-1} \leqslant (z-\lambda/2)^{q-1}$, whence $\int_\Lambda \rho_A^z(t)\,dt$ is at most $q(z-\lambda/2)^{q-1}$. Setting $\alpha = \max(a, b, q)$ gives the desired result. □

References

[1] Aldous, D. J. (1992) Asymptotics in the random assignment problem. *Probab. Theory Rel. Fields* **93** 507–534.

[2] Aldous, D. J. (2001) The $\zeta(2)$ limit in the random assignment problem. *Random Struct. Algorithms* **18** 381–418.

[3] Aldous, D. and Steele, J. M. (2004) The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures* (H. Kesten, ed.), Vol. 110 of Encyclopaedia of Mathematical Sciences, pp. 1–72. Springer.

[4] Karp, R. M. (1987) An upper bound on the expected cost of an optimal assignment. In *Discrete Algorithms and Complexity*, Vol. 15 of Perspectives in Computing, pp. 1–4. Academic Press.

[5] Krokhmal, P. A. and Pardalos, P. M. (2009) Random assignment problems. *European J. Oper. Res.* **194** 1–17.

[6] Linusson, S. and Wästlund, J. (2004) A proof of Parisi's conjecture on the random assignment problem. *Probab. Theory Relat. Fields* **128** 419–440.

[7] Lovász, L. (2012) *Large Networks and Graph Limits*, Vol. 60 of Colloquium Publications. American Mathematical Society.

[8] Mézard, M. and Parisi, G. (1985) Replicas and optimization. *J. Physique Lett.* **46** L771–L778.

[9] Nair, C., Prabhakar, B. and Sharma, M. (2005) Proofs of the Parisi and Coppersmith–Sorkin random assignment conjectures. *Random Struct. Algorithms* **27** 413–443.

[10] Parisi, G. (1998) A conjecture on random bipartite matching. arXiv:cond-mat/9801176v1

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press

[11] Salez, J. and Shah, D. (2009) Belief propagation: an asymptotically optimal algorithm for the random assignment problem. *Math. Oper. Res.* **34** 468–480.
[12] Walkup, D. W. (1979) On the expected value of a random assignment problem. *SIAM J. Comput.* **8** 440–442.
[13] Wästlund, J. (2005) A proof of a conjecture of Buck, Chan, and Robbins on the expected value of the minimum assignment. *Random Struct. Algorithms* **26** 237–251.
[14] Wästlund, J. (2012) Replica symmetry of the minimum matching. *Ann. of Math.* **175** 1061–1091.

Cite this article: Larsson J (2021). The minimum perfect matching in pseudo-dimension $0 < q < 1$. *Combinatorics, Probability and Computing* **30**, 374–397. https://doi.org/10.1017/S0963548320000425

https://doi.org/10.1017/S0963548320000425 Published online by Cambridge University Press