Water transport on infinite graphs

Olle Häggström [*]
Chalmers University of Technology

Timo Hirscher [†]
Goethe-University Frankfurt am Main

January 18, 2018

Abstract

If the nodes of a graph are considered to be identical barrels – featuring different water levels – and the edges to be (locked) water-filled pipes in between the barrels, consider the optimization problem of how much the water level in a fixed barrel can be raised with no pumps available, i.e. by opening and closing the locks in an elaborate succession. This model is related to an opinion formation process, the so-called Deffuant model. We consider the initial water profile to be given by i.i.d. $\mathrm{unif}(0,1)$ random variables, investigate the supremum of achievable water levels at a given node – or to be more precise, the support of its distribution – and ask in which settings it becomes degenerate, i.e. reduces to a single value. This turns out to be the case for all infinite connected quasi-transitive graphs, with exactly one exception: the two-sided infinite path.

Keywords: Water transport, graph algorithms, optimization, infinite path, Deffuant model.

1   Introduction

Motivation and model

Imagine a plane on which rainwater is collected in identical rain barrels, some of which are connected through pipes (that are already water-filled). All the pipes feature locks that are normally closed. If a lock is opened, the contents of the two barrels which are connected via this pipe start to level, see Figure 1. If one waits long enough, the water levels in the two barrels will be exactly the same, namely lie at the average $\frac{a+b}{2}$ of the two water levels (a and b) before the pipe was unlocked.
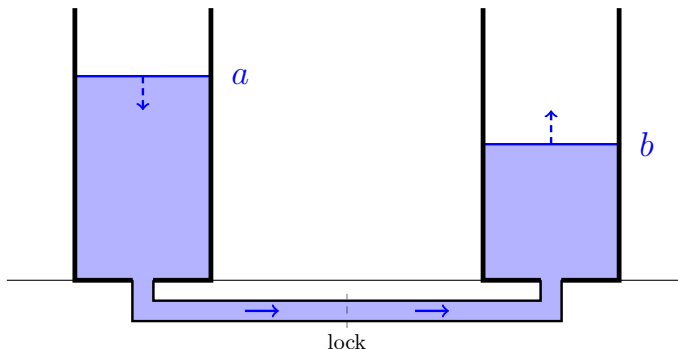

Figure 1: Leveling water stages after just having opened a lock.

After a rainy night in which the barrels accumulated various amounts of precipitation we might be interested in maximizing the water level in one fixed barrel by opening and closing some of the locks in carefully chosen order.

In order to mathematically model the setting, consider an undirected graph $G = (V, E)$, which is either finite or infinite with bounded maximum degree. Furthermore, we can assume without loss of generality that G is connected and simple, that means having neither loops nor multiple edges. Every vertex is understood to represent one of the barrels and the pipes correspond to the edges in the graph. The barrels themselves are considered to be identical, having a fixed capacity, say 1 for simplicity.

Making a move

Given some initial profile $\{\eta_0(u)\}_{u \in V} \in [0, 1]^V$, the system is considered to evolve in discrete time and in each round we can open one of the locked pipes and transport water from the fuller barrel into the emptier one. If we stop early, the two levels might not have completely balanced out giving rise to the following *update rule* for the water profile: If in round k the pipe $e_k = \langle x, y \rangle$ connecting the two barrels at sites x and y, with levels $\eta_{k-1}(x) = a$ and $\eta_{k-1}(y) = b$ respectively, is opened and closed after a certain period of time, we get

$$\begin{aligned} \eta_k(x) &= a + \mu_k\,(b - a) \\ \eta_k(y) &= b + \mu_k\,(a - b) \end{aligned} \qquad (1)$$

for some $\mu_k \in (0, \frac{1}{2}]$, which we assume can be chosen freely by an appropriate choice of how long the pipe is left open. All other levels stay unchanged, i.e. $\eta_k(w) = \eta_{k-1}(w)$ for all $w \in V \setminus \{x, y\}$.

Such a move can be described by the tuple $(e_k, \mu_k) \in E \times (0, \frac{1}{2}]$, consisting of chosen edge and transferred fraction. A *finite sequence of moves* is accordingly the concatenation of $(e_1, \mu_1), \ldots, (e_T, \mu_T)$, in chronological order, and leads to a final water profile $\{\eta_T(u)\}_{u \in V}$. The optimization problem is to maximize the

terminal value $\eta_T(v)$ over the number of rounds $T \in \mathbb{N}_0$ and move sequences. Here and in the sequel, we will denote by $\mathbb{N}_0$ the set of nonnegative integers. The quantity of interest is defined as follows:

Definition 1

For a graph $G = (V, E)$, an initial water profile $\{\eta_0(u)\}_{u \in V}$ and a fixed *target vertex* $v \in V$, let $\kappa_G(v)$ be defined as the supremum over all water levels that are achievable at v with finite move sequences, i.e.

$$\kappa_G(v) := \sup\{r \geq 0 : \text{there exists } T \in \mathbb{N}_0 \text{ and a move sequence s.t. } \eta_T(v) = r\}.$$

Connections to other concepts and synopsis

Some readers, familiar with mathematical models for social interaction processes, might note that (1) basically looks like the update rule in the opinion formation process given by the so-called *Deffuant model* for consensus formation in social networks (as described in the introduction of [3]), leaving aside the fact that here μ can change from update to update and the bounded confidence restriction is omitted. This similarity, however, is no coincidence: The situation described in the context above arises naturally in the analysis of the Deffuant model when dealing with the question of how extreme the opinion of a fixed agent can get (given an initial opinion profile on a specified network graph) if the interactions take place appropriately.

In order to tackle this question, Häggström [2] invented a non-random pairwise averaging procedure, which he proposed to call *Sharing a drink* (SAD). This process – which is the main focus of the preparatory Section 2 – was originally considered only on the (two-sided) infinite path, i.e. the graph $G = (V, E)$ with $V = \mathbb{Z}$ and $E = \{\langle v, v+1 \rangle, \ v \in \mathbb{Z}\}$, but can immediately be generalized to any graph (see Definition 2 below) and is dual to the water transport described above in a sense to be made precise in Lemma 2.1.

In the main part, Section 3, we consider the initial water levels to be i.i.d. random variables, uniformly distributed on $[0, 1]$, and detect a remarkable threshold phenomenon: While the border-line – marking where the distribution becomes degenerate – goes neatly between finite and infinite graphs for the average of water levels, in terms of $\kappa_G(v)$, the infinite path $\mathbb{Z}$ crosses this border: The highest achievable water level at a fixed vertex of $\mathbb{Z}$ depends on the initial profile in a nondegenerate sense, just like on any finite graph.

If an infinite graph contains a neighbor-rich half-line (see Definition 4), however, this dependence becomes degenerate: For any vertex $v \in V$, the value $\kappa_G(v)$ almost surely equals 1 (the essential supremum of the marginal distribution, cf. Theorem 3.5).

In the class of quasi-transitive graphs (see Definition 3 below), this makes $\mathbb{Z}$ quite unique: It constitutes the only exception among all infinite, connected, quasi-transitive graphs to the effect that $\kappa_G(v)$ is a nondegenerate random variable (cf. Corollary 3.2). In what is arguably our most striking result, Theorem 3.1, we classify all quasi-transitive graphs according to the support of $\kappa_G(v)$. We show that it is always of the form $[\alpha, 1]$ with α equal to either 0, $\frac{1}{2}$ or 1,

where $\alpha = 0$ corresponds to finite, $\alpha > 0$ to infinite graphs and the intermediate value $\alpha = \frac{1}{2}$ is attained only in the case $G = \mathbb{Z}$.

In addition, this result has implications in view of the connection to opinion formation as well: It indicates one reason why the analysis of the Deffuant model on $\mathbb{Z}^d$ with $d \geq 2$ has so far encountered more resistance as compared to the case of $d = 1$ (see the concluding Remark part (d) for details).

2   Connection to the SAD-procedure

In this section we collect some preliminaries needed for our main results in Section 3. Let us first recall from [2] the formal definition of the SAD-procedure:

Definition 2
For a graph $G = (V, E)$ and some fixed vertex $v \in V$, define $\{\xi_0(u)\}_{u \in V}$ by setting

$$\xi_0(u) = \delta_v(u) := \begin{cases} 1 & \text{for } u = v \\ 0 & \text{for } u \neq v. \end{cases}$$

In each time step, an edge $\langle x, y \rangle$ is chosen and the profile $\{\xi_0(u)\}_{u \in V}$ updated according to the rule (1) with $\{\xi_k(u)\}_{u \in V}$ in place of $\{\eta_k(u)\}_{u \in V}$. One can interpret this process as a full glass of water initially placed at vertex v (all other glasses being empty), which is then repeatedly shared among neighboring vertices by (at each time step) choosing a pair of neighbors and pouring a μ_k-fraction of the difference from the glass containing more water into the one containing less. Let us refer to this interaction process as *Sharing a drink (SAD)*.

Just as in [2], the SAD-procedure can be used to describe the composition of the contents in the water barrels after finitely many rounds of opening and closing pipe locks. The following lemma corresponds to Lemma 3.1 in [2], albeit in a more general graph context. Since the two dual processes (water transport and SAD) evolve in discrete time in our setting, its proof simplifies somewhat.

Lemma 2.1
Consider an initial profile of water levels $\{\eta_0(u)\}_{u \in V}$ on a graph $G = (V, E)$ and fix a vertex $v \in V$. For $T \in \mathbb{N}_0$ define the SAD-procedure that starts with $\xi_0 = \delta_v$ (see Definition 2) and is dual to the chosen move sequence in the water transport problem in the following sense: If in round $k \in \{1, \ldots, T\}$ the water profile is updated according to (1), the update in the SAD-profile at time $T - k \in \{0, \ldots, T - 1\}$ takes place along the same edge and with the same choice of μ_k. Then the terminal value at v is given by the convex combination

$$\eta_T(v) = \sum_{u \in V} \xi_T(u) \, \eta_0(u). \tag{2}$$

PROOF: We prove the statement by induction on T. For $T = 0$, the statement is trivial and there is nothing to show. For the induction step fix $T \in \mathbb{N}$ and

assume the first pipe opened to be $e = \langle x, y \rangle$. According to (1) we get

$$
\eta_1(u) = \begin{cases} \eta_0(u) & \text{if } u \notin \{x, y\} \\ (1 - \mu_1)\, \eta_0(x) + \mu_1\, \eta_0(y) & \text{if } u = x \\ (1 - \mu_1)\, \eta_0(y) + \mu_1\, \eta_0(x) & \text{if } u = y. \end{cases}
$$

Let us consider $\{\eta_1(u)\}_{u \in V}$ as some initial profile $\{\eta'_0(u)\}_{u \in V}$. By induction hypothesis we get

$$
\begin{aligned}
\eta'_{T-1}(v) &= \sum_{u \in V} \xi'_{T-1}(u)\, \eta'_0(u) \\
&= \sum_{u \in V \setminus \{x,y\}} \xi'_{T-1}(u)\, \eta_0(u) \\
&\quad + \Big((1 - \mu_1)\, \xi'_{T-1}(x) + \mu_1\, \xi'_{T-1}(y) \Big)\, \eta_0(x) \\
&\quad + \Big((1 - \mu_1)\, \xi'_{T-1}(y) + \mu_1\, \xi'_{T-1}(x) \Big)\, \eta_0(y),
\end{aligned} \tag{3}
$$

where $\eta'_{T-1}(v) = \eta_T(v)$ by construction and $\{\xi'_k(u)\}_{u \in V}$, $0 \le k \le T-1$, are the intermediate water profiles of the SAD-procedure corresponding to the move sequence after round 1. As by definition the original SAD-procedure arises from the shortened one by adding an update at time $T - 1$ along edge e with parameter μ_1, we find

$$
\xi_T(u) = \begin{cases} \xi_{T-1}(u) & \text{if } u \notin \{x, y\} \\ (1 - \mu_1)\, \xi_{T-1}(x) + \mu_1\, \xi_{T-1}(y) & \text{if } u = x \\ (1 - \mu_1)\, \xi_{T-1}(y) + \mu_1\, \xi_{T-1}(x) & \text{if } u = y, \end{cases}
$$

as well as $\xi_k = \xi'_k$ for all $k \in \{0, \ldots, T-1\}$, in particular $\xi'_{T-1}(u) = \xi_{T-1}(u)$ for all $u \in V$. This together with (3) establishes the claim. $\qquad\square$

Note that, although we are going to consider random initial profiles of water levels, the statement of Lemma 2.1 deals with a deterministic duality that does not involve any randomness (once the initial profile and the move sequence are fixed).

Let us prepare one more auxiliary result, which turns out to follow readily from the energy argument used in the proof of Thm. 2.3 in [2]:

Lemma 2.2
Given an initial profile of water levels $\{\eta_0(u)\}_{u \in V}$ on a graph $G = (V, E)$, fix a finite set $A \subseteq V$ and a set $E_A \subseteq E$ of edges inside A that connects A. If we open the pipes in E_A – and no others – in repetitive sweeps for times long enough such that $\mu_k \ge \varepsilon$ for some fixed $\varepsilon > 0$ in each round (cf. (1)), then the water levels at vertices in A approach a balanced average, i.e. converge to the common value $\frac{1}{|A|} \sum_{u \in A} \eta_0(u)$. For all $v \in A$, the corresponding dual SAD-profiles started with $\xi_0 = \delta_v$ converge uniformly to $\frac{1}{|A|}\, \mathbb{1}_A$.

PROOF: Let us define the energy after round k inside A by

$$W_k(A) = \sum_{u \in A} \big(\eta_k(u)\big)^2.$$

A short calculation reveals that an update of the form (1) reduces the energy by $2\mu_k(1 - \mu_k)(b - a)^2$, where the updated water levels were a and b respectively. If μ_k is bounded away from 0, the fact that $W_k(A) \geq 0$ for all k entails that the difference in water levels $|b - a|$ before a pipe is opened can be larger than any fixed positive value only finitely many times. In effect, since any pipe in E_A is opened repetitively we must have $|\eta_k(u) - \eta_k(v)| \to 0$ as $k \to \infty$ for all edges $\langle u, v \rangle \in E_A$. As the updates are average preserving, the first part of the claim follows from the fact that E_A connects A.

The second part of the lemma follows by applying the same argument to the dual SAD-procedure. $\qquad\square$

To round out these preliminary considerations, let us collect some properties of SAD-profiles derived in [2] into a single lemma for convenience.

Lemma 2.3
Consider the SAD-procedure on a path, started in vertex v, i.e. with $\xi_0(u) = \delta_v(u)$, $u \in V$.

(a) The SAD-profiles achievable on paths are all unimodal.

(b) If the vertex v only shares the water to one side, it will remain a mode of the SAD-profile.

(c) The supremum over all achievable SAD-profiles started with δ_v at another vertex w equals $\frac{1}{d+1}$, where d is the graph distance between v and w.

In this context, we call a vertex v a *mode* of the SAD-profile $\{\xi_t(u)\}_{u \in V}$ if $\xi_t(v) \geq \xi_t(u)$ for all $u \in V$. On a cycle-free graph (i.e. a tree), the profile is said to be *unimodal* if the values $\xi_t(u)$ along any path starting in a mode are nonincreasing.

The results in [2] actually all deal with the two-sided infinite path, but it is evident how the arguments used immediately transfer to finite paths. Part (a) hereby corresponds to Lemma 2.2 in [2], part (b) to Lemma 2.1 and part (c) to Thm. 2.3. The argument Häggström [2] used to prove the statement in (c) for the two-sided infinite path can in fact be generalized to prove the result for trees without much effort, as was done by Shang (see Prop. 6 in [5]).

3 In terms of water transport, $\mathbb{Z}$ behaves almost like a finite graph

In this section, we want to analyze the water transport problem given random initial water levels. More precisely, we will consider connected, simple

graphs $G = (V, E)$ with bounded degree and i.i.d. $\mathrm{unif}(0, 1)$ initial water levels $\{\eta_0(u)\}_{u \in V}$. The supremum $\kappa_G(v)$ of achievable water levels at a fixed target vertex $v \in V$ (cf. Def. 1) naturally depends on the initial water levels, which makes it a random variable as well. The question we want to address is in which cases the distribution of $\kappa_G(v)$ is degenerate.

Let us back away from $\kappa_G(v)$ for a moment and briefly discuss the global average of water levels across the graph (for infinite G defined as the limit of averages along a given sequence of nested subsets of V, which finally include every fixed vertex). As a side note, let us mention that this global average (if it exists) will not change with time as updates are average-preserving, cf. (1). In our i.i.d. $\mathrm{unif}(0, 1)$ setting, for any G it a.s. exists and the picture is clear-cut: On a finite graph, the average is a nondegenerate random variable, while on an infinite graph it almost surely equals the expected value $\frac{1}{2}$, according to the strong law of large numbers.

For $\kappa_G(v)$, in contrast, the regime of nondegeneracy extends a bit into the realm of infinite graphs, but not much. How this is to be understood can be made precise, if we restrict ourselves to quasi-transitive graphs – Theorem 3.1 below classifies them completely in terms of $\kappa_G(v)$.

Definition 3
Let $G = (V, E)$ be a simple graph. A bijection $f : V \to V$ with the property that $\langle f(u), f(v) \rangle \in E$ if and only if $\langle u, v \rangle \in E$ is called a *graph automorphism*. G is said to be *(vertex-) transitive* if for any two vertices $u, v \in V$ there exists a graph automorphism f that maps u on v, i.e. $f(u) = v$. If the vertex set V can be partitioned into finitely many classes such that for any two vertices u, v belonging to the same class there exists a graph automorphism that maps u on v, the graph G is called *quasi-transitive*.

Note that the idea of quasi-transitivity becomes nontrivial only for infinite graphs, as all finite graphs are quasi-transitive by definition. In what follows, we will denote the support of the distribution $\mathcal{L}(X)$ of a random variable X by

$$\mathrm{supp}(X) := \big\{ r \in \mathbb{R}, \ \forall \varepsilon > 0 : \ \mathbb{P}\big(|X - r| < \varepsilon\big) > 0 \big\}.$$

Theorem 3.1
Consider a connected, quasi-transitive graph $G = (V, E)$, a fixed vertex $v \in V$ and the initial water levels to be i.i.d. $\mathrm{unif}(0, 1)$. Then $\mathrm{supp}\big(\kappa_G(v)\big) = [\alpha, 1]$, where

(a) $\alpha = 0$, if G is finite,

(b) $\alpha = \frac{1}{2}$, if G is the two-sided infinite path $\mathbb{Z}$ and

(c) $\alpha = 1$, in all other cases, i.e. if G is infinite and has at least one vertex of degree at least 3.

With the intention of phrasing the special role of $\mathbb{Z}$ in the class of infinite (connected) quasi-transitive graphs more distinctly, we can formulate the following corollary of the above theorem:

Corollary 3.2
Given i.i.d. $\mathrm{unif}(0,1)$ initial water levels, the only infinite, connected, quasi-transitive graph G, for which $\kappa_G(v)$ is not almost surely equal to 1 is the two-sided infinite path $\mathbb{Z}$.

Before setting out to prove Theorem 3.1, let us approach the case of finite G with help of two toy examples: the paths on 2 and 3 nodes. For paths, Lemma 2.3 is the key to pin down the distribution of $\kappa_G(v)$, a task that evidently becomes involved rather quickly even for minimal examples.

Examples
(a) Consider the simplest nontrivial graph G, consisting of a single edge, and let the initial water levels be given by two random variables U_1 and U_2. Trivially, we have
$$U_1 \leq \kappa_G(1) \leq \max\{U_1, U_2\}.$$

In fact, with Lemma 2.3 in place, it doesn't take much thought to conclude $\kappa_G(1) = \max\{U_1, \frac{U_1+U_2}{2}\}$; the former referring to the case $U_1 \geq U_2$, the latter to $U_1 < U_2$. If we assume U_1 and U_2 to be independent and uniformly distributed on $[0,1]$, a short calculation (cf. the left part of Figure 2) reveals the distribution function
$$F_{\kappa_G(1)}(x) = \begin{cases} \frac{3}{2}x^2 & \text{for } 0 \leq x \leq \frac{1}{2} \\ x - \frac{1}{2}(1-x)^2 & \text{for } \frac{1}{2} \leq x \leq 1, \end{cases}$$

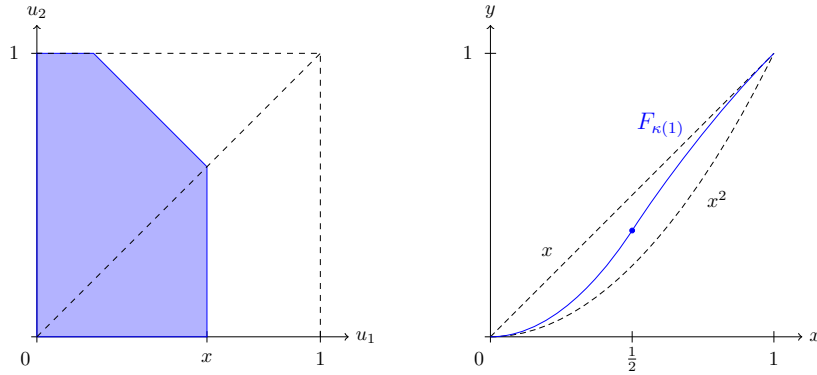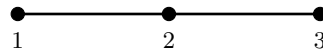which indeed lies in between $F_{U_1}(x) = x$ and $F_{\max\{U_1,U_2\}}(x) = x^2$, see the right part of Figure 2.

Figure 2: On the left a visualization of $\mathbb{P}(\kappa_G(1) \leq x)$, on the right the distribution function of $\kappa_G(1)$.

(b) The simplest nontransitive graph is the path on three vertices, i.e.
$$G = (\{1, 2, 3\}, \{\langle 1, 2\rangle, \langle 2, 3\rangle\}).$$

8

Simple conclusions based on Lemma 2.3 reveal that the suprema of achievable water levels are given by the random variables

$$\kappa_G(1) = \max\left\{\eta_0(1), \frac{\eta_0(1)+\eta_0(2)}{2}, \frac{\eta_0(1)+\eta_0(2)+\eta_0(3)}{3}\right\},$$
$$\kappa_G(2) = \max\left\{\eta_0(2), \frac{\eta_0(1)+\eta_0(2)}{2}, \frac{\eta_0(2)+\eta_0(3)}{2}, \frac{\eta_0(1)}{2} + \frac{\eta_0(2)+\eta_0(3)}{4},\right.$$
$$\left.\frac{\eta_0(3)}{2} + \frac{\eta_0(1)+\eta_0(2)}{4}\right\}.$$

Writing $e_1 = \langle 1, 2\rangle$, $e_2 = \langle 2, 3\rangle$, the five values whose maximum gives $\kappa_G(2)$ correspond respectively to: doing nothing, opening e_1, opening e_2, opening first e_2 then e_1 as well as opening first e_1 then e_2. For independent $\mathrm{unif}(0,1)$ initial water levels, a tedious but elementary calculation reveals the (already somewhat cumbersome) distribution function

$$F_{\kappa_G(1)}(x) = \begin{cases} \frac{8}{3}x^3 & x \in [0, \frac{1}{3}] \\ -\frac{11}{6}x^3 + \frac{9}{2}x^2 - \frac{3}{2}x + \frac{1}{6} & x \in [\frac{1}{3}, \frac{1}{2}] \\ -\frac{23}{6}x^3 + \frac{13}{2}x^2 - 2x + \frac{1}{6} & x \in [\frac{1}{2}, \frac{2}{3}] \\ \frac{2}{3}x^3 - \frac{5}{2}x^2 + 4x - \frac{7}{6} & x \in [\frac{2}{3}, 1]. \end{cases} \quad \text{for}$$

What now follows is a series of results that, in the end, will be combined to prove Theorem 3.1. In order to settle part (a) of the theorem, it is enough to observe that for a finite graph $G = (V, E)$ in general, (2) implies

$$\eta_0(v) \le \kappa_G(v) \le \max\{\eta_0(u), \ u \in V\}. \tag{4}$$

Indeed, this immediately entails the following statement, which is slightly stronger than Theorem 3.1 (a):

Proposition 3.3
Consider $G = (V, E)$ to be a finite connected graph, $v \in V$ a fixed vertex and the initial water levels to be i.i.d. $\mathrm{unif}(0,1)$. Then

$$\mathbb{P}\big(0 < \kappa_G(v) < 1\big) = 1 \quad and \quad \mathrm{supp}\big(\kappa_G(v)\big) = [0, 1].$$

PROOF: For ease of notation write $V = \{1, \ldots, n\}$ and $\vec{\eta} := (\eta_0(1), \ldots, \eta_0(n))$. Since $\vec{\eta} \sim \mathrm{unif}(0,1)^n$, it readily follows from (4) that $\kappa_G(v) \in (0,1)$ a.s. and we can further deduce for arbitrary $r \in [0, 1]$ and $\varepsilon > 0$:

$$\mathbb{P}\big(\kappa_G(v) \in (r - \varepsilon, r + \varepsilon)\big) \ge \mathbb{P}\big(\vec{\eta} \in (r - \varepsilon, r + \varepsilon)^n\big) \ge \varepsilon^n > 0. \qquad \square$$

This simple argument applies to the average of initial water levels just as well. As mentioned above, when dealing with infinite graphs, however, $\kappa_G(v)$ shows a different behavior than the global average: In order to determine whether the supremum of achievable water levels at a given vertex v is a.s. constant or not, we have to investigate the structure of the infinite graph a bit more closely.

What turns out to be crucial is, whether the graph contains an infinite self-avoiding path with sufficiently many extra neighbors as a subgraph (to be made

more precise in Definition 4 below). If it does, the distribution of $\kappa_G(v)$ becomes degenerate for all $v \in V$; more specifically, with probability 1, one can push the water level at v arbitrarily close to 1, the essential supremum of the marginal distribution (cf. Theorem 3.5).

The two-sided infinite path, however, is too lean to feature such a substructure and behaves therefore much more like a finite graph, in the sense that the distribution of $\kappa_G(v)$ is nondegenerate. This claim will be established by the following proposition, which is slightly stronger than Theorem 3.1 (b):

Proposition 3.4
Consider i.i.d. $\mathrm{unif}(0,1)$ *initial water levels on the two-sided infinite path $\mathbb{Z}$, i.e. the graph with $V = \mathbb{Z}$ and $E = \{\langle v, v+1 \rangle,\ v \in \mathbb{Z}\}$. Then for any $v \in V$:*

$$\mathbb{P}\left(\tfrac{1}{2} < \kappa_{\mathbb{Z}}(v) < 1\right) = 1 \quad and \quad \mathrm{supp}\left(\kappa_{\mathbb{Z}}(v)\right) = [\tfrac{1}{2}, 1].$$

PROOF: To begin with, given i.i.d. $\mathrm{unif}(0,1)$ initial water levels, $\kappa_G(v) > \frac{1}{2}$ a.s. holds for any infinite (connected) graph G: While the strong law of large numbers guarantees $\kappa_G(v) \geq \frac{1}{2}$ (cf. Lemma 2.2), with little more effort one can even get the strict inequality: Take $\{v_k,\ k \in \mathbb{N}\} \subseteq V$ such that $v_1 = v$ and for all $k \in \mathbb{N}$, $\{v_1, \ldots, v_k\}$ is a connected vertex-set of size k. Further, define $X_k = \eta_0(v_k) - \frac{1}{2}$ for all $k \in \mathbb{N}$. Then $(S_n)_{n \in \mathbb{N}}$, defined by $S_n = \sum_{k=1}^{n} X_k$, is an i.i.d. symmetric random walk, whence $\limsup_{n \to \infty} S_n = \infty$ almost surely, see e.g. Thm. 4.1.2 in [1]. In particular, there almost surely exists an $n \in \mathbb{N}$ with $\frac{1}{n} \cdot \sum_{k=1}^{n} \eta_0(v_k) > \frac{1}{2}$. Again by Lemma 2.2, this ensures $\mathbb{P}\left(\kappa_G(v) > \frac{1}{2}\right) = 1$.

In the case $G = \mathbb{Z}$, and for any fixed $\varepsilon > 0$, with positive probability the vertex v is what Häggström [2] calls two-sidedly ε-flat with respect to the initial profile (cf. Lemma 4.3 there), i.e.

$$\frac{1}{m+n+1} \sum_{u=v-m}^{v+n} \eta_0(u) \in \left[\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon\right] \quad \text{for all } m, n \in \mathbb{N}_0. \tag{5}$$

Lemma 6.3 in [2] states that in this situation, the water level at v is bound to stay within the interval $[\frac{1}{2} - 6\varepsilon, \frac{1}{2} + 6\varepsilon]$, irrespectively of future updates. Consequently, we arrive at $\frac{1}{2} \in \mathrm{supp}\left(\kappa_{\mathbb{Z}}(v)\right)$. By the same token, we get in fact $\mathrm{supp}\left(\kappa_{\mathbb{Z}}(v)\right) = [\frac{1}{2}, 1]$: Consider two i.i.d. $\mathrm{unif}(0,1)$ initial configurations $\{\eta_0(u)\}_{u \in V}$ and $\{\eta_0'(u)\}_{u \in V}$, which coincide for all $u \neq v$, while $\eta_0'(v)$ is independent of everything else. For fixed $r \in (\frac{1}{2}, 1]$, choose $0 < \varepsilon < \frac{1}{2}(r - \frac{1}{2})$ and let A be the event that v is two-sidedly ε-flat with respect to $\{\eta_0(u)\}_{u \in V}$ and $\eta_0'(v) \in [r - \varepsilon, r]$. By independence and Lemma 4.3 in [2], A has positive probability. Conditioned on the event A, again Lemma 6.3 in [2] together with $\eta_0(v) \in [\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]$ and (2) force all achievable water levels at v starting from the profile $\{\eta_0'(u)\}_{u \in V}$ to stay below $r + 7\varepsilon$. Together with the trivial lower bound from (4), we conclude $\mathbb{P}\left(\kappa_G(v) \in (r - \varepsilon, r + 7\varepsilon)\right) > 0$, hence $r \in \mathrm{supp}\left(\kappa_{\mathbb{Z}}(v)\right)$.

In order to establish the claim, we are left to show that $\kappa_{\mathbb{Z}}(v) < 1$ almost surely. Let us assume $\mathbb{P}\left(\kappa_{\mathbb{Z}}(v) = 1\right) > 0$ for contradiction and fix a finite set $I \subseteq \mathbb{Z}$ of nodes. If $\kappa_{\mathbb{Z}}(v) = 1$, there exist finite move sequences achieving water levels at v arbitrarily close to 1. Consequently, we can pick a sequence

of finite move sequences with terminal values at v converging to 1. Due to the fact that $\max\{\eta_0(u), \ u \in I\} < 1$ a.s., the dual SAD-profiles corresponding to the elements of this sequence must tend to 0 uniformly on I (cf. Lemma 2.1). As a consequence, $\mathbb{1}_{\{\kappa_{\mathbb{Z}}(v)=1\}}$ is independent of $\{\eta_0(u), \ u \in I\}$, which makes $\{\kappa_{\mathbb{Z}}(v) = 1\}$ a tail event. By Kolmogorov's zero-one law and our assumption, it has to be an almost sure event, which contradicts $\frac{1}{2} \in \mathrm{supp}\big(\kappa_{\mathbb{Z}}(v)\big)$. $\qquad\square$

Moving away from $\mathbb{Z}$, let us next properly define the type of infinite subgraph, which (if contained in G) makes the distribution of $\kappa_G(v)$ degenerate:

Definition 4
Let $G = (V, E)$ be an infinite, connected, simple graph. It is said to contain a *neighbor-rich half-line*, if there exists a subgraph of G consisting of a half-line

$$H = \big(\{v_k, \ k \in \mathbb{N}\}, \{\langle v_k, v_{k+1}\rangle, \ k \in \mathbb{N}\}\big)$$

and distinct vertices $\{u_n, \ n \in \mathbb{N}\}$ from $V \setminus \{v_k, \ k \in \mathbb{N}\}$ such that there is an injective function $f : \mathbb{N} \to \mathbb{N}$ with the following two properties (cf. Figure 3):

(i) For all $n \in \mathbb{N}$: $\langle u_n, v_{f(n)}\rangle \in E$, i.e. the vertices u_n and $v_{f(n)}$ are neighbors in G.

(ii) The function f is growing slowly, in the sense that $\sum_{n=1}^{\infty} \frac{1}{f(n)}$ diverges.
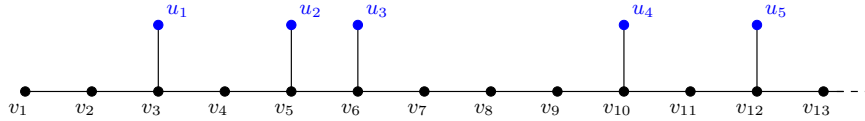

Figure 3: The beginning part of a neighbor-rich half-line.

Note that we can always take the function f to be (strictly) increasing – by relabeling $\{u_n, \ n \in \mathbb{N}\}$ if necessary. Furthermore, if G is connected and contains a neighbor-rich half-line, we can choose any vertex $v \in V$ to be its beginning vertex: If v_k is the vertex with highest index at shortest distance to v in H, replace $(v_1, \ldots, v_k)$ by a shortest path from v to v_k in H. The altered half-line will still be neighbor-rich, since for all $M, N \in \mathbb{N}$ and f as above:

$$\sum_{n=1}^{\infty} \frac{1}{f(n)} = \infty \quad \Longleftrightarrow \quad \sum_{n=M}^{\infty} \frac{1}{f(n) + N} = \infty.$$

With this notion in hand, we can state and prove the following result:

Theorem 3.5
Consider an infinite (connected) graph $G = (V, E)$, a fixed vertex $v \in V$ and the initial water levels to be i.i.d. $\mathrm{unif}(0, 1)$. If G contains a neighbor-rich half-line, then $\kappa_G(v) = 1$ almost surely.

Before we get down to the proof of this, let us provide a standard auxiliary result which will come in useful:

Lemma 3.6

For $\varepsilon > 0$, let $(Y_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence having Bernoulli distribution with parameter ε. If the function $f : \mathbb{N} \to \mathbb{N}$ is strictly increasing and such that $\sum_{n=1}^{\infty} \frac{1}{f(n)}$ diverges, then

$$\sum_{n=1}^{\infty} \frac{Y_n}{f(n)} = \infty \quad almost\ surely.$$

PROOF: Let us define

$$X_k = \sum_{n=1}^{k} \frac{Y_n - \varepsilon}{f(n)} \quad \text{for all } k \in \mathbb{N}.$$

As the increments are independent, bounded and centered, this defines a martingale with respect to the natural filtration. Furthermore, by independence of the $(Y_n)_{n\in\mathbb{N}}$,

$$\mathbb{E}\left(X_k^2\right) = \sum_{n=1}^{k} \frac{\mathbb{E}\left(Y_n - \varepsilon\right)^2}{f(n)^2} = (\varepsilon - \varepsilon^2) \cdot \sum_{n=1}^{k} \frac{1}{f(n)^2} \leq \varepsilon \frac{\pi^2}{6}.$$

By the L^p-convergence theorem (see for instance Thm. 5.4.5 in [1]) there exists a random variable X such that X_k converges to X almost surely and in L^2. Having finite variance, X must be a.s. real-valued and due to

$$\sum_{n=1}^{k} \frac{Y_n}{f(n)} - X_k = \varepsilon \cdot \sum_{n=1}^{n} \frac{1}{f(n)},$$

the divergence of $\sum_{n=1}^{\infty} \frac{1}{f(n)}$ forces $\sum_{n=1}^{\infty} \frac{Y_n}{f(n)} = \infty$ almost surely. $\square$

PROOF OF THEOREM 3.5: Given a graph G with the properties stated and a vertex v, we can choose a neighbor-rich half-line H with $v = v_1$ and the set of extra neighbors $\{u_n\}_{n\in\mathbb{N}}$ as described in and after Definition 4. The initial water levels at $\{u_n\}_{n\in\mathbb{N}}$ are i.i.d. unif$(0,1)$, of course.

Depending on the random initial profile, let us define the following SAD-procedure starting at v: Fix $\varepsilon, \delta > 0$ and let $\{N_k\}_{k\in\mathbb{N}}$ be the increasing (random) sequence of indices, chosen such that the initial water level at u_{N_k} is at least $1-\varepsilon$ for all k. Then define the SAD-procedure – starting with $\xi_0(v) = 1$, $\xi_0(u) = 0$ for all $u \in V\setminus\{v\}$ – such that first all vertices along the path $(v_1, v_2, \ldots, v_{f(N_1)}, u_{N_1})$ exchange liquids sufficiently often to get

$$\xi_{t_1}(u_{N_1}) \geq \frac{1}{f(N_1) + 2} \quad \text{for some } t_1 > 0,$$

and never touch u_{N_1} again. Note that by Lemma 2.2, $\xi_k(u_{N_1})$ can be pushed as close to $\frac{1}{f(N_1)+1}$ as desired in this way. At time t_1, the total amount of water in

the glasses at $v_1, v_2, \ldots, v_{f(N_1)}$ equals $1 - \xi_{t_1}(u_{N_1})$ and we will repeat the same procedure along $(v_1, v_2, \ldots, v_{f(N_2)}, u_{N_2})$ to get

$$\xi_{t_2}(u_{N_2}) \geq \frac{1}{f(N_2) + 2} \cdot \left(1 - \xi_{t_1}(u_{N_1})\right) \quad \text{for some } t_2 > t_1$$

and iterate this.

Note that the ith iteration decreases the total amount of water localized at vertices of the half-line H by a factor bounded from above by $1 - \frac{1}{f(N_i)+2}$. More precisely, after m iterations of this kind, the total amount of water on H equals $1 - \sum_{k=1}^{m} \xi_{t_k}(u_{N_k})$, which can be further bounded by:

$$\begin{aligned}
1 - \sum_{k=1}^{m} \xi_{t_k}(u_{N_k}) &\leq \prod_{k=1}^{m} \left(1 - \frac{1}{f(N_k) + 2}\right) \\
&\leq \exp\left(-\sum_{k=1}^{m} \frac{1}{f(N_k) + 2}\right),
\end{aligned} \tag{6}$$

where the standard inequality $1 - x \leq \mathrm{e}^{-x}$ was used in the second step. Defining $Y_n := \mathbb{1}_{\{\eta_0(u_n) \geq 1 - \varepsilon\}}$ for all $n \in \mathbb{N}$ we get $(Y_n)_{n \in \mathbb{N}}$ i.i.d. $\mathrm{Ber}(\varepsilon)$ and can rewrite the limit of the sum in the exponent as follows:

$$\sum_{k=1}^{\infty} \frac{1}{f(N_k) + 2} = \sum_{n=1}^{\infty} \frac{Y_n}{f(n) + 2}.$$

This allows us to conclude from Lemma 3.6 that the exponent in (6) tends a.s. to $-\infty$ as $m \to \infty$. Consequently, $m, T \in \mathbb{N}$ can be chosen large enough such that with probability $1 - \delta$ it holds that

$$\sum_{k=1}^{m} \xi_{t_k}(u_{N_k}) \geq 1 - \varepsilon \quad \text{and} \quad t_m \leq T.$$

Given this event, the move sequence corresponding to the SAD-procedure just described – adding no further updates after time t_m, i.e. $\mu_t = 0$ for $t_m < t \leq T$ – then ensures (cf. Lemma 2.1) that

$$\eta_T(v) \geq \sum_{k=1}^{m} \xi_T(u_{N_k}) \, \eta_0(u_{N_k}) \geq (1 - \varepsilon)^2,$$

forcing $\kappa_G(v) \geq (1-\varepsilon)^2$ with probability at least $1-\delta$. Since $\delta > 0$ was arbitrary, this implies $\kappa_G(v) \geq (1-\varepsilon)^2$ a.s. and letting ε go to 0 then establishes the claim. $\qquad \square$

Let us now assemble the established results to conclude our complete classification of quasi-transitive graphs.

PROOF OF THEOREM 3.1: As mentioned above, Propositions 3.3 and 3.4 imply parts (a) and (b) respectively. In view of Theorem 3.5, to prove part (c), we only

have to verify, that an infinite, connected, quasi-transitive graph that is not the two-sided infinite path contains a neighbor-rich half-line. Since G is infinite (and by our general assumptions both connected and having finite maximal degree) a compactness argument guarantees the existence of a half-line H on the vertices $\{v_k, \ k \in \mathbb{N}\}$ such that $v_1 = v$ and the graph distance from v_k to v is $k - 1$ for all k.

Let us consider the function $d : V \to \mathbb{N}_0$, where $d(u)$ is the graph distance from the node u to a vertex of degree at least 3 being closest to it. Since G is quasi-transitive, connected and not the two-sided infinite path, d is finite and can take on only finitely many values, which is why it has to be bounded, by a constant $c \in \mathbb{N}$ say. Consequently, G can not contain stretches of more than $2c$ linked vertices of degree 2. For this reason, there must be a vertex among $v_3, \ldots, v_{2c+3}$, say $v_{f(1)}$, having a neighbor u_1 outside of H. In the same way, we can find a vertex u_2 outside H having a neighbor $v_{f(2)}$ among $v_{2c+6}, \ldots, v_{4c+6}$ and in general some u_n not part of H but linked to a vertex $v_{f(n)} \in \{v_k, \ k \in \mathbb{N}\}$ with

$$(n - 1)\,(2c + 3) + 3 \leq f(n) \leq n\,(2c + 3) \quad \text{for all } n \in \mathbb{N}.$$

This choice ensures that $v_{f(m)}$ and $v_{f(n)}$ are at graph distance at least 3 for $m \neq n$, which forces the set $\{u_n, \ n \in \mathbb{N}\}$ to consist of distinct vertices. Due to

$$\sum_{n=1}^{\infty} \frac{1}{f(n)} \geq \frac{1}{2c + 3} \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

H is a neighbor-rich half-line in the sense of Definition 4 as desired, which concludes the proof. $\qquad\square$

Remarks

(a) Note that the essential property of the initial water levels, needed in the proof of Theorem 3.5, was independence. The argument can immediately be generalized to the situation where the initial water levels are independently (but not necessarily identically) distributed on a bounded interval $[0, C]$ and we have some weak form of uniformity, namely:

For every $\delta > 0$, there exists some $\varepsilon > 0$ such that for all $v \in V$:

$$\mathbb{P}\big(\eta_0(v) > C - \delta\big) \geq \varepsilon.$$

The sequence $Y_n := \mathbb{1}_{\{\eta_0(u_n) \geq C - \delta\}}$, $n \in \mathbb{N}$, similar to the one defined in the proof of Theorem 3.5 will no longer be i.i.d. $\mathrm{Ber}(\varepsilon)$, but an appropriate coupling will ensure that

$$\sum_{n=1}^{\infty} \frac{Y_n}{n} \geq \sum_{n=1}^{\infty} \frac{Z_n}{n} \quad \text{a.s.,}$$

where $(Z_n)_{n \in \mathbb{N}}$ is an i.i.d. sequence of $\mathrm{Ber}(\varepsilon)$ random variables. Accordingly, we get $\kappa_G(v) = C$ a.s. even in this generalized setting.

(b) To see how surprising Corollary 3.2 is, consider the case of the quasi-transitive graph G_M, obtained by taking $\mathbb{Z}$ and adding a single twig (an edge leading to a single vertex and nothing more) to every Mth vertex of $\mathbb{Z}$. Intuitively, one would think that for large M, e.g. $M = 10^{100}$, the distribution of $\kappa_{G_M}(v)$ at a vertex v far from any of the twigs would be close to the nondegenerate distribution of $\kappa_{\mathbb{Z}}(v)$. But this is not the case, since according to Theorem 3.1, the distribution of $\kappa_{G_M}(v)$ is concentrated at 1. In this sense the transition is discontinuous.

(c) Note that the second part of Theorem 3.1 trivially implies $\kappa_G(v) = 1$ a.s. for any graph G containing an infinite, connected, quasi-transitive subgraph other than $\mathbb{Z}$ (and any vertex $v \in V$). It is worth emphasizing that Theorem 3.1 does, however, not capture the full statement of Theorem 3.5: If we take G to consist of the two-sided infinite path $\mathbb{Z}$ and a single twig added to every node that corresponds to a prime number, the only infinite, connected, quasi-transitive subgraph of G is $\mathbb{Z}$ itself. However, since G contains a neighbor-rich half-line, Theorem 3.5 states that $\kappa_G(v) = 1$ a.s. for i.i.d. unif$(0, 1)$ initial water levels and any target vertex v.

(d) As alluded to in the introduction, the results from Section 3 do have implications for the Deffuant model for opinion formation, whose analysis originally inspired the optimization problem investigated here. The qualitative change of behavior of $\kappa_G(v)$ for a fixed vertex $v \in \mathbb{Z}^d$, when switching from the two-sided infinite path ($d = 1$) to higher-dimensional grids ($d \geq 2$), constitutes a substantial hindrance in the analysis of the Deffuant model: As proved in Lemma 6.3 in [2], for i.i.d. unif$(0, 1)$ initial values on $\mathbb{Z}$, almost surely there exist nodes that are stuck with a value close to the mean of the marginal distribution (irrespectively of future interactions of neighboring nodes), while Theorem 3.1 tells us that this is no longer true in higher dimensions. These vertices played however a central role in the two analyses of the Deffuant model on $\mathbb{Z}$, done by Lanchier [4] and Häggström [2] respectively, which is why their ideas can not be transferred to the case $d \geq 2$ immediately.

Acknowledgement

References

[1] Durrett, R., *"Probability: Theory and Examples (4th edition)"*, Cambridge University Press, 2010.

[2] Häggström, O., A pairwise averaging procedure with application to consensus formation in the Deffuant model, *Acta Applicandae Mathematicae*, Vol. 119 (1), pp. 185-201, 2012.

[3] Häggström, O. and Hirscher, T., Further results on consensus formation in the Deffuant model, *Electronic Journal of Probability*, Vol. 19, 2014.

[4] Lanchier, N., The critical value of the Deffuant model equals one half, *Latin American Journal of Probability and Mathematical Statistics*, Vol. 9 (2), pp. 383-402, 2012.

[5] Shang, Y., Deffuant model with general opinion distributions: First impression and critical confidence bound, *Complexity*, Vol. 19 (2), pp. 38-49, 2013.

Olle Häggström
Department of Mathematical Sciences,
Chalmers University of Technology,
412 96 Gothenburg, Sweden.
olleh@chalmers.se

Timo Hirscher
Institute for Mathematics,
J.W. Goethe University,
60054 Frankfurt a.M., Germany.
hirscher@math.uni-frankfurt.de