



# Hipster random walks

L. Addario-Berry<sup>1</sup>  · H. Cairns<sup>2</sup> · L. Devroye<sup>3</sup> · C. Kerriou<sup>1</sup> · R. Mitchell<sup>1</sup>

Received: 18 September 2019 / Revised: 25 May 2020 / Published online: 19 June 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

We introduce and study a family of random processes on trees we call hipster random walks, special instances of which we heuristically connect to the min-plus binary trees introduced by Robin Pemantle and studied by Auffinger and Cable (Pemantle's Min-Plus Binary Tree, 2017. [arXiv:1709.07849](https://arxiv.org/abs/1709.07849) [math.PR]), and to the critical random hierarchical lattice studied by Hambly and Jordan (Adv Appl Probab 36(3):824–838, 2004. <https://doi.org/10.1239/aap/1093962236>). We prove distributional convergence for the processes, after rescaling, by showing that their evolutions can be understood as a discrete analogues of certain convection–diffusion equations, then using a combination of coupling arguments and results from the numerical analysis literature on convergence of numerical approximations of PDEs.

**Keywords** Recursive distributional equations · Random trees · Numerical analysis · Burgers' equation · Porous medium equation · PDEs · Interacting particle systems

**Mathematics Subject Classification** 60F05 · 65M75 · 60B10 · 60G18

---

✉ L. Addario-Berry  
[louigi.addario@mcgill.ca](mailto:louigi.addario@mcgill.ca)

H. Cairns  
[ahc238@cornell.edu](mailto:ahc238@cornell.edu)

L. Devroye  
[lucdevroye@gmail.com](mailto:lucdevroye@gmail.com)

C. Kerriou  
[celine.kerriou@mail.mcgill.ca](mailto:celine.kerriou@mail.mcgill.ca)

R. Mitchell  
[rivka.mitchell@mail.mcgill.ca](mailto:rivka.mitchell@mail.mcgill.ca)

<sup>1</sup> Department of Mathematics and Statistics, McGill University, Montréal, Canada

<sup>2</sup> Department of Mathematics, Cornell University, Ithaca, USA

<sup>3</sup> School of Computer Science, McGill University, Montréal, Canada

### 1 Introduction

Let  $\mathcal{T}$  denote the complete rooted infinite binary tree. The root receives label  $\emptyset$ ; children of node  $v$  receive labels  $v0$  and  $v1$ . In this way generation- $n$  nodes of  $\mathcal{T}$  are labeled by the set  $\mathcal{L}_n := \{0, 1\}^n$ . For  $n \geq 1$ , write  $\mathcal{T}_n$  for the binary tree consisting of the root of  $\mathcal{T}$  and its first  $n$  generations of descendants. The *leaves* of  $\mathcal{T}_n$  are the nodes  $\mathcal{L}_n$  in generation  $n$ ; its *internal nodes* are precisely the nodes of  $\mathcal{T}_{n-1}$ .

Fix any assignment  $F = (f_v, v \in \mathcal{T})$  of binary functions  $f_v : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  to the nodes of  $\mathcal{T}$ . Then for any  $n \geq 1$ , the functions  $F_n = (f_v, v \in \mathcal{T}_n)$  may be viewed as turning  $\mathcal{T}_n$  into a recursively constructed function of arity  $2^n$ , with inputs at the leaves of  $\mathcal{T}_n$  and output at the root of  $\mathcal{T}_n$ . More precisely, given real values  $\mathbf{z} = (z_v, v \in \mathcal{T})$  for  $n \geq 1$  let  $F_n^{\mathbf{z}} : \mathcal{T}_n \rightarrow \mathbb{R}$  be given by

$$F_n^{\mathbf{z}}(v) = \begin{cases} z_v & \text{if } v \in \mathcal{L}_n \\ f_v(F_n^{\mathbf{z}}(v0), F_n^{\mathbf{z}}(v1)) & \text{if } v \in \mathcal{T}_{n-1}. \end{cases} \tag{1}$$

When either the functions comprising  $F$  are random, or the inputs are random (or both), then  $F_n$  is itself a random function; a large number of problems in probability can be phrased in terms of such random functions. As a very simple example, fix  $p \in (0, 1)$ , and independently for each  $v \in \mathcal{T}$  define  $f_v$  by

$$f_v(x, y) = \begin{cases} 1 & \text{with probability } p, \\ x + y + 1 & \text{with probability } 1 - p. \end{cases}$$

Then  $F_n^1(\emptyset)$  is distributed as the total number of individuals in the first  $n$  generations of a branching process with offspring distribution given by  $p_0 = p, p_2 = 1 - p$ . Here  $\mathbf{1}$  assigns value 1 to all nodes of  $\mathcal{T}$ ; below we likewise write  $\mathbf{0}$  for the vector assigning values 0 to all nodes of  $\mathcal{T}$ .

This work establishes two distributional limit theorems for the output of systems of random functions on  $\mathcal{T}$  which we dub *hipster random walks*.

*Hipster random walk* Let  $(A_v, v \in \mathcal{T})$  be independent Bernoulli(1/2) random variables, and let  $(D_v, v \in \mathcal{T})$  be IID random variables, independent of  $(A_v, v \in \mathcal{T})$ . Then set

$$f_v(x, y) = xA_v\mathbf{1}_{[x \neq y]} + y(1 - A_v)\mathbf{1}_{[x \neq y]} + (x + D_v)\mathbf{1}_{[x=y]}. \tag{2}$$

In other words, if  $x \neq y$  then  $f_v(x, y)$  flips a fair coin to decide whether to output  $x$  or  $y$ . If  $x = y$  then  $f_v(x, y)$  outputs  $x + D_v$ . We call  $(D_v, v \in \mathcal{T})$  the *steps* of the hipster random walk.

The name is inspired by the following intuitive picture, which is based on the stereotype that hipsters don't want to be observed liking popular things. In our setting, the "things" in question are potential random walk locations. Imagine that for each node  $v$ , one of  $v0$  or  $v1$  is hipper than the other; which one is hipper is determined randomly using  $A_v$ . If  $x \neq y$  then the hipper individual doesn't have any new company at their current location and stays put. If  $x = y$  then the hipper individual detects new

company, takes this as a sign that their current location is becoming popular, and so decides to leave (moves to  $x + D_v$ ). The output of  $f_v(x, y)$  is the new location of the hipper of  $v_0$  and  $v_1$ .

In this work we focus on two specific choices for the common law of the steps. *Totally asymmetric  $q$ -lazy simple hipster random walk* This is the hipster random walk with steps  $(C_v, v \in \mathcal{T})$  which are independent Bernoulli( $q$ ) random variables. In this case, definition (1) yields functions  $B_n^z : \mathcal{T}_n \rightarrow \mathbb{R}$  given by

$$B_n^z(v) = \begin{cases} z_v & \text{if } v \in \mathcal{L}_n \\ B_n^z(v_0)A_v + B_n^z(v_1)(1 - A_v) & \text{if } v \in \mathcal{T}_{n-1}, B_n^z(v_0) \neq B_n^z(v_1) \\ B_n^z(v_0) + C_v & \text{if } v \in \mathcal{T}_{n-1}, B_n^z(v_0) = B_n^z(v_1). \end{cases} \quad (3)$$

*Symmetric simple hipster random walk* This is the hipster random walk with steps  $(R_v, v \in \mathcal{T})$  satisfying  $\mathbf{P}\{R_v = 1\} = \mathbf{P}\{R_v = -1\} = 1/2$ . In this case, definition (1) yields functions  $G_n^z : \mathcal{T}_n \rightarrow \mathbb{R}$  given by

$$G_n^z(v) = \begin{cases} z_v & \text{if } v \in \mathcal{L}_n \\ G_n^z(v_0)A_v + G_n^z(v_1)(1 - A_v) & \text{if } v \in \mathcal{T}_{n-1}, G_n^z(v_0) \neq G_n^z(v_1) \\ G_n^z(v_0) + R_v & \text{if } v \in \mathcal{T}_{n-1}, G_n^z(v_0) = G_n^z(v_1). \end{cases} \quad (4)$$

Our main results are contained in the following two theorems.

**Theorem 1** Let  $\mathbf{Z} = (Z_v, v \in \mathcal{T})$  be IID integer random variables. Next fix  $q \in (0, 1]$ , and for  $n \geq 1$  let  $B_n = B_n^z(\emptyset)$  be the output of the  $n$ -step totally asymmetric  $q$ -lazy hipster random walk on input  $(Z_v, v \in \mathcal{L}_n)$ . Then

$$\frac{B_n}{(4qn)^{1/2}} \xrightarrow{d} B,$$

where  $B$  is Beta(2, 1)-distributed.

**Theorem 2** Let  $\mathbf{Z} = (Z_v, v \in \mathcal{T})$  be IID integer random variables, and for  $n \geq 1$  let  $G_n = G_n^z(\emptyset)$  be the output of the  $n$ -step symmetric simple hipster random walk on input  $(Z_v, v \in \mathcal{L}_n)$ . Then

$$\frac{G_n}{(36n)^{1/3}} + \frac{1}{2} \xrightarrow{d} G,$$

where  $G$  is Beta(2, 2)-distributed.

### 1.1 Related work

Fix  $p \in (0, 1)$  and independently for each  $v \in \mathcal{T}$ , define  $f_v$  by

$$f_v(x, y) = \begin{cases} x + y & \text{with probability } p, \\ \min(x, y) & \text{with probability } 1 - p. \end{cases}$$

Write  $M_n^z$  for the resulting random functions (again obtained by applying definition (1)). The study of this model was proposed by Robin Pemantle [8], who conjectured that there exists a constant  $c$  such that when  $p = 1/2$ ,

$$\frac{\ln M_n^1(\emptyset)}{(cn)^{1/2}} \xrightarrow{d} B, \tag{5}$$

where  $B$  is Beta(2,1)-distributed. This was recently proved by Auffinger and Cable [1], who identified the correct constant as  $c = \pi^2/3$ , and dubbed this model *Pemantle's min-plus binary tree*.

There is an obvious similarity between (5) and the convergence in Theorem 1 and, indeed, there is a heuristic connection between the models. By (5), we know that  $M_n$  is growing at a stretched exponential rate. In view of this, it is natural to consider what is happening at a log scale. Write  $\beta_0 = \log M_n^1(0)$  and  $\beta_1 = \log M_n^1(1)$  for the log (base 2, here and below) of the inputs to the root.

Most of the time  $\log M_n^1(0)$  and  $\log M_n^1(1)$  take radically different values (since they are identically distributed and exhibit random fluctuations on the scale  $n^{1/2}$ ). In this (typical) case, with  $\beta_\emptyset = \log M_n^1(\emptyset)$  we have

$$\beta_\emptyset = \begin{cases} \log(M_n^1(0) + M_n^1(1)) \approx \max(\beta_0, \beta_1) & \text{with probability } 1/2, \\ \log \min(M_n^1(0), M_n^1(1)) \approx \min(\beta_0, \beta_1) & \text{with probability } 1/2. \end{cases}$$

In other words, when  $M_n^1(0)$  and  $M_n^1(1)$  are extremely different, the output at the root just looks like the value of a random child.

On the other hand, will occasionally have  $M_n^1(0) \approx M_n^1(1)$ . In this case, the dynamics look rather different; for example, when  $M_n^1(0) = M_n^1(1)$  we have

$$\beta_\emptyset = \begin{cases} \log(M_n^1(0) + M_n^1(1)) = \beta_0 + 1 & \text{with probability } 1/2, \\ \log \min(M_n^1(0), M_n^1(1)) = \beta_0 & \text{with probability } 1/2. \end{cases}$$

So in this case, at the log scale, the output of the min-plus binary tree just looks like the common value of the children plus a Bernoulli(1/2) increment. This looks very much like the dichotomy for the totally asymmetric hipster random walk: when the children have different values, output the value of a random child; when they have the same value, output that value plus a random increment.

The analogy isn't perfect, because when  $M_n^1(0)$  and  $M_n^1(1)$  take similar but not identical values, the behaviour of the min-plus binary tree interpolates between the two cases. This "smearing out" creates a speed-up relative to the totally asymmetric 1/2-lazy simple hipster random walk (the constants in the rescaling are  $(\log e) \cdot \sqrt{\pi^2/3}$  and  $\sqrt{2}$ , respectively).

Another related model, called the *random hierarchical lattice*, was proposed by Hambly and Jordan [4], which in the language of this work may be described as

follows. Fix  $p \in (0, 1)$ , and independently for each  $v \in \mathcal{T}$ , define  $f_v$  by

$$f_v(x, y) = \begin{cases} x + y & \text{with probability } p, \\ \frac{xy}{x+y} & \text{with probability } 1 - p. \end{cases}$$

A natural interpretation of this is as follows. View the inputs to  $v$  as electrical networks with effective resistances  $x$  and  $y$ . Then at node  $v$  the resistors are combined in series or in parallel, with probability  $p$  or  $1 - p$  respectively; the output is the new, combined network.

Write  $R_n^z$  for the resulting random functions. Hambly and Jordan show that almost surely,  $R_n^1(\emptyset) \rightarrow \infty$  when  $p > 1/2$  and  $R_n^1(\emptyset) \rightarrow 0$  when  $p < 1/2$ , and conjecture that  $R_n(p)$  almost surely grows exponentially when  $p = 1/2$ .

By analogy with the min-plus tree, it seems plausible to conjecture that when  $p = 1/2$ , the random variables  $R_n^1(\emptyset)$  are again growing at a stretched exponential scale. In order to make a more precise guess at the phenomenology, we reprise the argument from the case of the min-plus binary tree.

In the setting of this paper, if  $|\log x - \log y|$  is large then  $\log(x + y) \approx \max(\log(x), \log(y))$ , and  $\log(xy/(x + y)) \approx \min(\log(x), \log(y))$ . In this case,  $\log f_v(x, y)$  just looks like the value of a random child. On the other hand, when  $x = y$  then  $\log(x + y) = \log(x) + 1$  and  $\log(xy/(x + y)) = \log(x) - 1$ , so  $\log f_v(x, y)$  looks like the common log-value of the children plus a random walk step. Again, there is an interpolation between these two extremes, but the heuristic suggests that at a log scale, when  $p = 1/2$  the random hierarchical lattice should look somewhat like a symmetric simple hipster random walk. In view of the validity of such a heuristic in the case of Pemantle's min-plus binary tree, we are led to conjecture that there exists  $c > 0$  such that (in the case  $p = 1/2$ ), with  $R_n = R_n^1(\emptyset)$ ,

$$\frac{\log R_n}{(cn)^{1/3}} + 1/2 \xrightarrow{d} R,$$

where  $R$  is Beta(2, 2)-distributed. This is in slight disagreement with a prediction of Hambly and Jordan [4], who write "In the case  $p = 1/2$  ... we also believe that there is an almost sure exponential growth rate for the resistance". The truth of our conjecture would imply that the growth rate is in fact stretched exponential. However, even the weaker conjecture that

$$R_n \xrightarrow{d} \frac{1}{2}\delta_0 + \frac{1}{2}\delta_\infty,$$

i.e., that  $\mathbf{P}\{R_n \leq r\} \rightarrow 1/2$  for any  $r \in (0, \infty)$ , is open at this point.

### 1.2 Our approach

Recall that  $B_n^z$  and  $G_n^z$  are our notation for the totally asymmetric and symmetric simple hipster random walks, respectively. For a probability measure  $\mu$  supported by

$\mathbb{Z}$  we write  $\mathbf{P}_\mu$  for the probability measure under which the entries of  $\mathbf{z}$  are IID with law  $\mu$ , independent of all the other random variables describing the system.

The utility of taking random input values  $\mathbf{z}$  is that the measures  $\mathbf{P}_\mu$  endow the random functions  $B_n^{\mathbf{z}}$  and  $G_n^{\mathbf{z}}$  with a sort of projective consistency: under  $\mathbf{P}_\mu$ , for all  $0 \leq j \leq n$ , the random variables  $(B_n^{\mathbf{z}}(v), v \in \mathcal{L}_j)$  are IID with the law of  $B_{n-j}^{\mathbf{z}}(\emptyset)$ . An exactly analogous statement holds for  $G_n^{\mathbf{z}}$ . This allows us to write recurrences in  $n$  for the output distribution at the root in both models.

We first derive the recurrence for the totally asymmetric  $q$ -lazy simple hipster random walk. Write

$$p_j^n(\mu) = \mathbf{P}_\mu \{B_n^{\mathbf{z}}(\emptyset) = j\}.$$

Provided we assume that  $\mu$  is supported on  $\mathbb{Z}$ , then the distribution of  $B_n^{\mathbf{z}}(\emptyset)$  under  $\mathbf{P}_\mu$  is determined by the values  $(p_j^n(\mu))_{n \in \mathbb{Z}}$ .

In what follows, we will write  $B_n(v) = B_n^{\mathbf{z}}(v)$  when the law of  $\mathbf{z}$  is clear from context. By considering the values at the children of the root, using (3) we have

$$\begin{aligned} p_j^{n+1}(\mu) &= \mathbf{P}_\mu \{B_{n+1}(\emptyset) = j\} \\ &= \frac{1}{2} \mathbf{P}_\mu \{B_{n+1}(0) = j\} \mathbf{P}_\mu \{B_{n+1}(1) \neq j\} \\ &\quad + \frac{1}{2} \mathbf{P}_\mu \{B_{n+1}(0) \neq j\} \mathbf{P}_\mu \{B_{n+1}(1) = j\} \\ &\quad + q \cdot \mathbf{P}_\mu \{B_{n+1}(0) = B_{n+1}(1) = j - 1\} \\ &\quad + (1 - q) \cdot \mathbf{P}_\mu \{B_{n+1}(0) = B_{n+1}(1) = j\} \\ &= p_j^n(\mu)(1 - p_j^n(\mu)) + q \cdot (p_{j-1}^n(\mu))^2 + (1 - q) \cdot (p_j^n(\mu))^2. \end{aligned}$$

For the second equality we use the fact that, under  $\mathbf{P}_\mu$ ,  $B_{n+1}(0)$  and  $B_{n+1}(1)$  are independent and have the law of  $B_n(\emptyset)$ . After rearrangement this yields the identity

$$p_j^{n+1}(\mu) - p_j^n(\mu) = -q \cdot (p_j^n(\mu))^2 - p_{j-1}^n(\mu)^2. \tag{6}$$

This is a discrete analog of the inviscid Burgers' equation,

$$\partial_t u = -q \cdot \partial_x (u^2). \tag{7}$$

Next consider the symmetric simple hipster random walk, and write

$$q_j^n(\mu) = \mathbf{P}_\mu \{G_n^{\mathbf{z}}(\emptyset) = j\}.$$

We again write  $G_n(v) = G_n^{\mathbf{z}}(v)$  when the law of  $\mathbf{z}$  is clear. By considering the values at the children of the root, using (4) have

$$\begin{aligned} q_j^{n+1}(\mu) &= \mathbf{P}_\mu \{G_{n+1}(\emptyset) = j\} \\ &= \frac{1}{2} \mathbf{P}_\mu \{G_{n+1}(0) = j\} \mathbf{P}_\mu \{G_{n+1}(1) \neq j\} \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \mathbf{P}_\mu \{G_{n+1}(0) \neq j\} \mathbf{P}_\mu \{G_{n+1}(1) = j\} \\
 & + \frac{1}{2} \cdot \mathbf{P}_\mu \{G_{n+1}(0) = G_{n+1}(1) = j - 1\} \\
 & + \frac{1}{2} \cdot \mathbf{P}_\mu \{G_{n+1}(0) = G_{n+1}(1) = j + 1\} \\
 & = q_j^n(\mu)(1 - q_j^n(\mu)) + \frac{1}{2} q_{j-1}^n(\mu)^2 + \frac{1}{2} q_{j+1}^n(\mu)^2,
 \end{aligned}$$

where for the second equality we have again used projective consistency. Rearrangement now gives

$$q_j^{n+1}(\mu) - q_j^n(\mu) = \frac{1}{2} \left( (q_{j+1}^n(\mu))^2 - 2(q_j^n(\mu))^2 + (q_{j-1}^n(\mu))^2 \right), \tag{8}$$

a discrete analogue of the porous medium equation for groundwater infiltration [9],

$$\partial_t u = \frac{1}{2} \cdot \partial_{xx}(u^2). \tag{9}$$

The preceding development shows that for both of the models under consideration, the evolution of the probability distribution as  $n$  varies is a discrete analogue of a PDE. As such, it's natural to expect the behaviour of the PDE to predict that of the finite system. Indeed, if  $B$  is Beta(2, 1)-distributed then  $(4qt)^{1/2}B$  has density

$$\frac{x}{2qt} \mathbf{1}_{[0 \leq x \leq (4qt)^{1/2}]}, \tag{10}$$

which solves (7) at its points of differentiability. Similarly, if  $G$  is Beta(2, 2)-distributed then  $(36t)^{1/3}(G - 1/2)$  is supported by  $[-(36t)^{1/3}, (36t)^{1/3}]$  and has density

$$\frac{1}{(36t)^{1/3}} \cdot 6 \left( \frac{1}{2} + \frac{x}{(36t)^{1/3}} \right) \left( \frac{1}{2} - \frac{x}{(36t)^{1/3}} \right) = \frac{3}{4} \left( \left( \frac{2}{9t} \right)^{\frac{1}{3}} - \left( \frac{2x^2}{9t} \right) \right) \tag{11}$$

on that interval; this solves (9) wherever it is differentiable.

It isn't *a priori* obvious that this perspective is useful, for multiple reasons. First, the PDEs under consideration are degenerate convection–diffusion equations, for which neither existence nor uniqueness of solutions is clear. (The “solutions” above already have points of non-differentiability so do not make sense classically; on the other hand, once one abandons classical solutions, uniqueness is in general lost.) Second, even if one can identify the “correct” PDE solutions, it isn't obvious that the behaviour of the finite systems will correctly approximate the limiting PDEs.

Showing that a discrete difference equation provides a good approximation for an associated PDE is a problem that sits squarely within the area of numerical analysis. It turns out that, by viewing (6) and (8) as numerical approximation schemes we are able to use results from the rigorous numerical analysis literature to prove distributional convergence of the associated random variables.

In fact, in their initial form the numerical analysis results are not strong enough for our purposes, as they establish convergence in an integrated sense which doesn't give us access to the distribution of the hipster random walks at fixed times. However, we are able to strengthen the numerical approximation theorems using coupling arguments together with carefully chosen initial conditions for the hipster random walks. The coupling arguments are slightly surprising, so we briefly describe them. Write  $\mu$  for the law of the entries of the input field  $\mathbf{Z}$ . Suppose  $\mathbf{Z}$  is replaced by another input field  $\mathbf{W}$  whose entries have some law  $\nu$ , and that for some  $\alpha \in (0, 1)$  there exists a coupling  $(z, w)$  of  $\mu$  and  $\nu$  such that  $\mathbf{P}\{z > w\} \leq \alpha$ . Then the totally asymmetric dynamics (3) may be coupled so that for any  $n$ ,

$$\mathbf{P}\left\{B_n^{\mathbf{Z}}(\emptyset) > B_n^{\mathbf{W}}(\emptyset)\right\} \leq \alpha,$$

and likewise the symmetric dynamics (4) may be coupled so that for any  $n$ ,

$$\mathbf{P}\left\{G_n^{\mathbf{Z}}(\emptyset) > G_n^{\mathbf{W}}(\emptyset)\right\} \leq \alpha.$$

For the precise statements, see Lemmas 10 and 11, below.

The remainder of the paper is structured as follows. In Sect. 2 we describe the setting of the numerical approximation theorem we will use, as well as the theorem itself (Theorem 3). We also state propositions which verify that the dynamics we study may be recast within the framework of Theorem 3; the proofs of these propositions appear in ‘‘Appendix A’’. In Sect. 3 we prove ‘‘integrated’’ versions of Theorems 1 and 2. In Sect. 4 we state the coupling lemmas mentioned above (their proofs are also deferred to ‘‘Appendix A’’), and use them to prove Theorems 1 and 2. Finally, Sect. 5 contains several suggestions for interesting avenues of research related to hipster random walks and their ilk.

## 2 Finite approximation schemes for degenerate convection–diffusion equations

This section summarizes the principal result of [2], which is the main tool we use to study the asymptotic behaviour of the recurrence relations (7) and (9). In [2], Evje and Karlsen consider convection–diffusion initial value problems of the form

$$\begin{cases} \partial_t u + \partial_x f(u) - \partial_{xx} K(u) = 0, & (x, t) \in \mathbb{R} \times (0, T) \\ u(x, 0) = u_0(x). \end{cases} \tag{12}$$

The problem is specified by the choice of the (measurable) functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $K : \mathbb{R} \rightarrow [0, \infty)$ , which are respectively called the *convection flux* and the *diffusion flux*, and by the choice of initial condition  $u_0 : \mathbb{R} \rightarrow \mathbb{R}$ . Burgers' equation (7) is obtained by taking  $f(u) = f_B^{(q)}(u) := qu^2$  and  $K = K_B^{(q)} \equiv 0$ . The porous membrane equation (9) is obtained by taking  $f = f_P \equiv 0$  and  $K(u) = K_P(u) := \frac{1}{2}u^2$ .



Evje and Karlsen provide sufficient conditions for the convergence of certain numerical approximation schemes to solutions of (12). The “solutions” in question are not everywhere differentiable, so must be understood in a weak sense, which we now explain in detail. (We impose stronger conditions on our solutions than those in [2], since they are easier to state and hold in the cases we consider in the current work.)

Recall that for a signed measure  $\mu$  on a measurable space  $(M, \mathcal{B})$ , there are unique non-negative measures  $\mu^+, \mu^-$  on  $(M, \mathcal{B})$  such that  $\mu = \mu^+ - \mu^-$ ; this is the *Jordan decomposition* of  $\mu$ . The *variation* of  $\mu$  is the (unsigned) measure  $|\mu| := \mu^+ + \mu^-$ .

A measurable function  $z : \mathbb{R} \times [0, T]$  is *locally integrable* if for all compact sets  $S \subset \mathbb{R} \times [0, T]$ , the function  $z|_S$  is integrable. It has *bounded variation* if it is locally integrable and its partial derivatives  $\partial_x z$  and  $\partial_t z$ , considered as signed Borel measures, satisfy  $|\partial_x z|(\mathbb{R} \times (0, T)) + |\partial_t z|(\mathbb{R} \times (0, T)) < \infty$ . Finally,  $z$  lies in the Hölder space  $C^{1, \frac{1}{2}}(\mathbb{R} \times [0, T])$  if it is bounded and additionally there is  $C > 0$  such that for all  $(x, s), (y, t) \in \mathbb{R} \times [0, T]$ ,

$$|z(x, s) - z(y, t)| \leq C(|y - x| + |t - s|^{1/2}).$$

Let  $u : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$  be a bounded measurable function. We say  $u$  is a solution of (12) if the following conditions hold.

1. The function  $u|_{\mathbb{R} \times (0, T)}$  has bounded variation and  $u(\cdot, 0) \equiv u_0$ .
2. The function  $K(u)$  is bounded and Lipschitz continuous on  $\mathbb{R} \times [0, T]$ .
3. For all non-negative  $\phi \in C^\infty(\mathbb{R} \times [0, T])$  with compact support and with  $\phi|_{t=T} \equiv 0$ , and for all  $c \in \mathbb{R}$ ,

$$\int_{\mathbb{R}} \int_{(0, T)} ((u - c) \cdot \partial_t \phi + (f(u) - f(c) - \partial_x K(u)) \cdot \partial_x \phi) \cdot \text{sgn}(u - c) dt dx + \int_{\mathbb{R}} |u_0 - c| \phi(x, 0) dx \geq 0 \tag{13}$$

Here and elsewhere,  $\text{sgn}(x) := \mathbf{1}_{[x > 0]} - \mathbf{1}_{[x < 0]}$ . In [2], such a function  $u$  is called a *BV entropy weak solution* of (12).

The following may help understand the content of (13). Imagine that a smooth solution  $u$  of (12) existed, and fix any bounded smooth function  $\phi : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$  with compact support. Using integration by parts, we have

$$\int_{\mathbb{R}} \int_{(0, T)} u \cdot \partial_t \phi dt dx = \int_{\mathbb{R}} [u(x, \cdot) \phi(x, \cdot)]_{t=0}^T dx - \int_{\mathbb{R}} \int_{(0, T)} \phi \cdot \partial_t u dt dx.$$

By (12) we also have  $\partial_t u = \partial_{xx} K(u) - \partial_x f(u)$ . Thus, if  $\phi|_{t=T} = 0$  then the right-hand side is

$$\begin{aligned} & - \int_{\mathbb{R}} u_0 \cdot \phi(x, 0) dx - \int_{(0, T)} \int_{\mathbb{R}} \phi \cdot (\partial_{xx} K(u) - \partial_x f(u)) dx dt \\ & = - \int_{\mathbb{R}} u_0 \cdot \phi(x, 0) dx + \int_{(0, T)} \int_{\mathbb{R}} \partial_x \phi \cdot (\partial_x K(u) - f(u)) dx dt, \end{aligned}$$

the equality following from integration by parts and the fact that  $\phi$  has compact support. This yields the identity

$$\int_{\mathbb{R}} \int_{(0,T)} u \cdot \partial_t \phi + (f(u) - \partial_x K(u)) \cdot \partial_x \phi \, dt dx + \int_{\mathbb{R}} u_0 \cdot \phi(x, 0) dx = 0, \tag{14}$$

which is an integrated form of (12). Unfortunately, for many PDEs, there is no solution of (12) in the classical sense as the “obvious” candidate is non-differentiable. On the other hand, passing to the integrated form yields too much flexibility—solutions exist but are not unique.

One common way to single out a “physically relevant” solution of (12) is to first add a diffusive term  $\epsilon u_{xx}$  to the PDE in (12), and find integrated solutions  $u^{(\epsilon)}$  to the modified PDE. The smoothing effect of the diffusive term will often yield uniqueness of  $u^{(\epsilon)}$ ; one may then hope to define  $u := \lim_{\epsilon \downarrow 0} u^{(\epsilon)}$ . Informally, the addition of such a viscosity term is meant to enforce that any “shocks” (discontinuities of the solution or of its derivatives) propagate at “physically meaningful speeds”.

There are many versions of such arguments for different families of PDEs; one of the casualties of this approach is that the equality in (14) does not always persist in the  $\epsilon \rightarrow 0$  limit. Its replacement by an inequality in some sense encodes the idea that shocks inhibit information transmission (i.e., they are entropy increasing), but we have not found a convincing informal explanation of why this is so. For further details on and applications of this perspective, we refer the reader to [3,5–7,10,11].

For a given initial condition  $u_0 : \mathbb{R} \rightarrow \mathbb{R}$  and real  $\Delta_x, \Delta_t > 0$ , we define  $(U_j^n)_{n \in \mathbb{N}, j \in \mathbb{Z}} = (U_j^n(u_0, \Delta_x, \Delta_t))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  via the following discretization of (12).

$$\begin{aligned} \frac{U_j^{n+1} - U_j^n}{\Delta_t} &= -\frac{f(U_j^n) - f(U_{j-1}^n)}{\Delta_x} + \frac{K(U_{j+1}^n) - 2K(U_j^n) + K(U_{j-1}^n)}{(\Delta_x)^2} \\ U_j^0 &= \frac{1}{\Delta_x} \int_{j\Delta_x}^{(j+1)\Delta_x} u_0(x) dx. \end{aligned} \tag{15}$$

The first identity inductively specifies  $U_j^{n+1}$  for  $j \in \mathbb{Z}$  and  $n \geq 0$ ; the second defines  $U_j^0$  for  $j \in \mathbb{Z}$ . We refer to (15) as an *approximation scheme* for (12). Given an interval  $I \subseteq \mathbb{R}$ , we say the approximation scheme is *monotone on I* if the function  $S : \mathbb{R}^3 \rightarrow \mathbb{R}$  defined by

$$S(u^-, u, u^+) = \frac{u}{\Delta_t} - \frac{1}{\Delta_x} (f(u) - f(u^-)) + \frac{1}{(\Delta_x)^2} (K(u^+) - 2K(u) + K(u^-)) \tag{16}$$

satisfies  $S(I \times I \times I) \subseteq I$ , and is non-decreasing in each argument on  $I \times I \times I$ . Equivalently, in the first equation in (15), the value of  $U_j^{n+1}$  is a monotone function of  $U_{j+1}^n, U_j^n$  and  $U_{j-1}^n$ , provided those values all lie in  $I$ .

**Theorem 3** ([2]) *Suppose  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $K : \mathbb{R} \rightarrow [0, \infty)$  are continuously differentiable. Fix a bounded variation function  $u_0 : \mathbb{R} \rightarrow \mathbb{R}$  with compact support*

and such that  $f(u_0) - K'(u_0)$  also has bounded variation. Then there is a unique BV entropy weak solution  $u$  of the corresponding convection–diffusion equation (12).

Next, fix sequences  $(\Delta_x^M)_{M \geq 1}$  and  $(\Delta_t^M)_{M \geq 1}$  decreasing to zero, such that the corresponding approximation schemes  $(U_j^n(u_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  are monotone on an interval  $I \subseteq \mathbb{R}$ . Let  $u^M : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  be the function which, for all  $j \in \mathbb{Z}$  and  $n \in \mathbb{N}$ , takes the value  $U_j^n(u_0, \Delta_x^M, \Delta_t^M)$  on the half-open rectangle

$$[j\Delta_x^M, (j + 1)\Delta_x^M) \times [n\Delta_t^M, (n + 1)\Delta_t^M).$$

If  $u_0(\mathbb{R}) \subseteq I$ , then  $u^M$  converges pointwise almost everywhere to  $u$ , and for all compacts  $C \subset \mathbb{R} \times [0, \infty)$ ,  $\int_C |u^M - u| dx dt \rightarrow 0$  as  $M \rightarrow \infty$ . Moreover, the sequence of functions  $(K(u^M))_{M \geq 1}$  converges uniformly on compacts to  $K(u)$ .

In [2], the approximation schemes are required to be monotone on  $\mathbb{R}$ ; however, the above formulation requiring monotonicity on an interval is in fact an immediate consequence of the proof in [2].

The next two propositions verify that (10) and (11) indeed describe the BV entropy weak solutions of the convection–diffusion equation (12), for the relevant choices of  $f$  and  $K$ , and that the corresponding approximation schemes are monotone provided we take a suitably fine-meshed discretization. The proofs of these propositions, which boil down to careful applications of the divergence theorem together with case analysis (based on the value of  $c$  in (13)), appear in “Appendix A”.

The first of the propositions relates to Burgers’ equation, which corresponds to the totally asymmetric hipster random walk. For this model  $f(u) = f_B(u) := q \cdot u^2$  and  $K = K_B \equiv 0$ .

**Proposition 4** *Let  $q \in (0, 1)$ . Fix  $\varepsilon > 0$  and  $T > 0$  and define  $u_B : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$  by*

$$u_B(x, t) = \frac{x}{2q(t + \varepsilon)} \mathbf{1}_{[x \in [0, \sqrt{4q(t+\varepsilon)}]]}$$

Then  $u_B$  is the BV entropy weak solution to the initial value problem

$$\partial_t u + \partial_x (q \cdot u^2) = 0,$$

with initial condition  $u_0(x) = u_B(x, 0)$ . Moreover, the following holds. Fix  $M > 0$ , let  $\Delta_x^M = \frac{1}{M}$  and  $\Delta_t^M = \frac{1}{M^2}$ , and consider the approximation scheme  $(U_j^n(u_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  given by

$$\begin{aligned} \frac{U_j^{n+1} - U_j^n}{\Delta_t^M} &= -q \cdot \left( \frac{(U_j^n)^2 - (U_{j-1}^n)^2}{\Delta_x^M} \right), \quad j \in \mathbb{Z}, n \geq 1 \\ U_j^0 &= \frac{1}{\Delta_x^M} \int_{j\Delta_x^M}^{(j+1)\Delta_x^M} u_0(x) dx, \quad j \in \mathbb{Z} \end{aligned} \tag{17}$$

which is obtained from (15) by taking  $K \equiv K_B$  and  $f \equiv f_B$ . Then for  $M$  sufficiently large, the approximation scheme  $(U_j^n(u_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  is monotone on  $[0, (q\varepsilon)^{-1/2}]$ .

The second of the propositions concerns the porous medium equation, which corresponds to the symmetric simple hipster random walk. For this model we defined  $f = f_P \equiv 0$  and  $K(u) = K_P(u) = \frac{1}{2}u^2$ .

**Proposition 5** Fix  $\varepsilon > 0$  and  $T > 0$  and define  $v_P : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$  by

$$v_P(x, t) = \max \left( \frac{3}{4} \left( \left( \frac{2}{9(t + \varepsilon)} \right)^{\frac{1}{3}} - \left( \frac{2x^2}{9(t + \varepsilon)} \right) \right), 0 \right).$$

Then  $v_P$  is the BV entropy weak solution to the initial value problem

$$\partial_t v - \partial_{xx} \frac{v^2}{2} = 0,$$

with initial condition  $v_0(x) = v_P(x, 0)$ . Moreover, the following holds. Fix  $M > 0$ , let  $\Delta_x^M = \frac{1}{M}$  and  $\Delta_t^M = \frac{1}{M^3}$ , and consider the approximation scheme  $(U_j^n(v_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  given by

$$\begin{aligned} \frac{U_j^{n+1} - U_j^n}{\Delta_t^M} &= \frac{(U_{j+1}^n)^2 - 2(U_j^n)^2 + (U_{j-1}^n)^2}{2(\Delta_x^M)^2}, & j \in \mathbb{Z}, n \geq 1 \\ U_j^0 &= \frac{1}{\Delta_x^M} \int_{j\Delta_x^M}^{(j+1)\Delta_x^M} v_0(x) dx, & j \in \mathbb{Z} \end{aligned} \tag{18}$$

which is obtained from (15) by taking  $K \equiv K_P$ , and  $f \equiv f_P$ . Then for  $M$  sufficiently large, the approximation scheme  $(U_j^n(v_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  is monotone on  $[0, (3/4)(2/(9\varepsilon))^{1/3}]$ .

### 3 Integrated versions of Theorems 1 and 2

The approximation schemes in Propositions 4 and 5 differ from the recurrences for the hipster random walks, namely (6) and (8), by factors involving the spatial and discretizations,  $\Delta_x^M$  and  $\Delta_t^M$ . However, the form of those factors is such that we still have easily verified exact relations between the values  $U_j^n$  and the distributions of the hipster random walks. These relations are summarized in the next two propositions. Fix a non-negative measurable function  $\rho : \mathbb{R} \rightarrow [0, \infty)$  with  $\int_{\mathbb{R}} \rho(x) dx < \infty$ . For  $M > 0$  and  $j \in \mathbb{Z}$ , define a measure  $\mu_{\rho, M}$  on  $\mathbb{Z}$  by

$$\mu_{\rho, M}(\{j\}) = \int_{j/M}^{(j+1)/M} \rho(x) dx. \tag{19}$$

Next, for  $j \in \mathbb{Z}$  let  $u_j^0 = u_j^0(\rho, M) = M \cdot \mu_{\rho, M}(\{j\})$ , and for  $n \geq 1$  define  $(u_j^n)_{j \in \mathbb{Z}} = (u_j^n(\rho, M))_{j \in \mathbb{Z}}$  via the recurrence

$$M \cdot (u_j^{n+1} - u_j^n) = -q \cdot \left( (u_j^n)^2 - (u_{j-1}^n)^2 \right).$$

Note that this is equivalent to the recurrence in (17) since, in that recurrence,  $\Delta_x^M = 1/M$  and  $\Delta_t^M = 1/M^2$ . The following proposition connects the evolution of  $u_j^n$  with the totally asymmetric hipster random walk. Its proof is a straightforward inductive argument and is omitted.

**Proposition 6** *Suppose that  $\rho$  is a probability density function on  $\mathbb{R}$ . Fix  $M > 0$  and let  $\mu = \mu_{\rho, M}$  be as in (19). Then for all  $n \in \mathbb{N}$  and  $j \in \mathbb{Z}$ ,*

$$u_j^n(\rho, M) = M \cdot \mathbf{P}_\mu \{ B_n^Z(\emptyset) = j \}.$$

Next, fix  $M$  and  $\rho$  as above, and for  $j \in \mathbb{Z}$  let  $v_j^0 = v_j^0(\rho, M) = M \cdot \mu_{\rho, M}(\{j\})$ , where  $\mu_{\rho, M}(\{j\})$  is again given by (19). Then, for  $n \geq 1$ , define  $(v_j^n)_{j \in \mathbb{Z}} = (v_j^n(\rho, M))_{j \in \mathbb{Z}}$  by the recurrence

$$M \cdot (v_j^{n+1} - v_j^n) = \frac{1}{2} \left( (v_{j+1}^n)^2 - 2(v_j^n)^2 + (v_{j-1}^n)^2 \right).$$

This is equivalent to the recurrence in (18), since in (18) we have  $\Delta_x^M = 1/M$  and  $\Delta_t^M = 1/M^3$ .

**Proposition 7** *Suppose that  $\rho$  is a probability density function on  $\mathbb{R}$ . Fix any  $M > 0$  and let  $\mu = \mu_{\rho, M}$  be as in (19). Then for all  $n \in \mathbb{N}$  and  $j \in \mathbb{Z}$ ,*

$$v_j^n(\rho, M) = M \cdot \mathbf{P}_\mu \{ G_n^Z(\emptyset) = j \}.$$

The proof of Proposition 7 is also an easy induction and is omitted.

Having stated these results, we are prepared to prove weakenings of Theorems 1 and 2. We must weaken the theorems in two ways. First, rather than starting from arbitrary inputs, we choose initial distributions which are fine-mesh discretization of the initial conditions for which we understand the solutions to the associated initial value problems. In other words, in the totally asymmetric case we will start from a discretization of a scaled Beta(2, 1) distribution, and in the symmetric case we will start from a discretized Beta(2, 2) distribution. Second, our conclusions concern the distribution of trees of a *random* rather than fixed height. The reason for this is that the almost sure convergence provided by Theorem 3 is two-dimensional (it concerns the space-time field of values  $(U_j^n)_{n \geq 0, j \in \mathbb{Z}}$ ). Fixing the height of the tree corresponds to considering the PDE approximation at a fixed time; but Theorem 3 doesn't *a priori* guarantee the absence of "pathological" times at which the discrete approximations are badly-behaved.

**Proposition 8** Fix  $q \in (0, 1)$  and  $\varepsilon \in (0, 1)$ . Then for  $M > 0$  let  $\mu^M = \mu_\varepsilon^M$  be the probability measure on  $\mathbb{Z}$  defined by

$$\mu^M(\{j\}) = \int_{j/M}^{(j+1)/M} \frac{x}{2q\varepsilon} \mathbf{1}_{[x \in [0, \sqrt{4q\varepsilon}]]} dx .$$

Next fix  $0 \leq \ell < r$  and, under  $\mathbf{P}_{\mu^M}$ , let  $W$  be a Uniform $[\ell, r]$  random variable, independent of  $\mathbf{z}$ . Then

$$\frac{B_{\lfloor WM^2 \rfloor}^{\mathbf{z}}}{(4q(W + \varepsilon))^{1/2} M} \xrightarrow{d} B \text{ as } M \rightarrow \infty$$

where  $B$  is a Beta(2, 1) random variable.

**Proof** (Proof of Proposition 8)

The joint law of  $W$  and  $B_{\lfloor WM^2 \rfloor}^{\mathbf{z}}$  can be given explicitly as

$$\mathbf{P}_{\mu^M} \left\{ B_{\lfloor WM^2 \rfloor}^{\mathbf{z}} = j, W \in dt \right\} = \frac{dt}{r - \ell} \mathbf{P}_{\mu^M} \left\{ B_{\lfloor tM^2 \rfloor}^{\mathbf{z}}(\emptyset) = j \right\} \mathbf{1}_{[t \in [\ell, r]]} .$$

In the proof we write  $\mathbf{P}$  instead of  $\mathbf{P}_{\mu^M}$  and  $B_k$  instead of  $B_k^{\mathbf{z}}(\emptyset)$ , for succinctness. For  $b \in (0, 1)$ , we have

$$\begin{aligned} & \mathbf{P} \left\{ B_{\lfloor WM^2 \rfloor} \leq (4q(W + \varepsilon))^{1/2} Mb \right\} \\ &= \int_{\ell}^r \mathbf{P} \left\{ B_{\lfloor WM^2 \rfloor} \leq (4q(W + \varepsilon))^{1/2} Mb, W \in dt \right\} \\ &= \frac{1}{r - \ell} \int_{\ell}^r \mathbf{P} \left\{ B_{\lfloor tM^2 \rfloor} \leq (4q(t + \varepsilon))^{1/2} Mb \right\} dt . \end{aligned} \tag{20}$$

Define  $u_0(x) = (x/(2q\varepsilon)) \mathbf{1}_{[x \in [0, \sqrt{4q\varepsilon}]]}$  and for  $0 \leq t \leq 4q$  let

$$u(x, t) = \frac{x}{2q(t + \varepsilon)} \mathbf{1}_{[x \in [0, \sqrt{4q(t+\varepsilon)}]]} .$$

Then  $u \equiv u_B$  and  $u_0(x) = u_B(x, 0)$ , where  $u_B$  is as in Proposition 4, applied with  $T = 4q$  and  $t_0 = \varepsilon$ . Also write  $U_j^n = U_j^n(u_0, \Delta_x^M, \Delta_t^M)$ , where  $(U_j^n(u_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  is again as in Proposition 4.

Note that  $\mu^M(\{j\}) = \int_{j/M}^{(j+1)/M} u_0(x) dx$ , and by Proposition 6, for all  $n \in \mathbb{N}$  and  $j \geq 0$  we have  $U_j^n = M \cdot \mathbf{P} \{ B_n = j \}$ . Now let  $u^M : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  be the function which takes the value  $U_j^n$  on  $[j/M, (j + 1)/M) \times [n/M^2, (n + 1)/M^2)$  for  $n, j \in \mathbb{N}$ . Then for  $b \in [0, 1]$  we have

$$\int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} u^M(x, t) dx dt = \int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} M \cdot \mathbf{P} \left\{ B_{\lfloor tM^2 \rfloor} = \lfloor Mx \rfloor \right\} dx dt ,$$

so

$$\int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} u^M(x, t) dx dt \geq \int_{\ell}^r \mathbf{P} \left\{ B_{\lfloor tM^2 \rfloor} < \lfloor Mb(4q(t+\varepsilon))^{1/2} \rfloor \right\} dt \quad (21)$$

and

$$\int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} u^M(x, t) dx dt \leq \int_{\ell}^r \mathbf{P} \left\{ B_{\lfloor tM^2 \rfloor} \leq \lfloor Mb(4q(t+\varepsilon))^{1/2} \rfloor \right\} dt. \quad (22)$$

Since  $u_B$  is the solution to the initial value problem  $\partial_t u + q\partial_x u^2 = 0$  with initial condition  $u_0 = u_{B,0}$ , it follows by Theorem 3 that

$$\lim_{M \rightarrow \infty} \int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} u^M(x, t) dx dt = \int_{\ell}^r \int_0^{b(4q(t+\varepsilon))^{1/2}} u(x, t) dx dt = \int_{\ell}^r b^2 = (r - \ell)b^2.$$

Combining this with (20), (21) and (22), we obtain that for all  $b \in (0, 1)$ ,

$$\lim_{M \rightarrow \infty} \mathbf{P} \left\{ B_{\lfloor WM^2 \rfloor} \leq (4q(W + \varepsilon))^{1/2} Mb \right\} = b^2.$$

For  $B$  a Beta(2, 1) random variable,  $\mathbf{P} \{ B \leq b \} = b^2$ , so the result follows. □

**Proposition 9** Fix  $\varepsilon \in (0, 1)$ . For  $M > 0$  let  $\nu^M = \nu_{\varepsilon}^M$  be the probability measure on  $\mathbb{Z}$  defined by

$$\nu^M(\{j\}) = \int_{j/M}^{(j+1)/M} \max \left( \frac{3}{4} \left( \left( \frac{2}{9\varepsilon} \right)^{1/3} - \left( \frac{2x^2}{9\varepsilon} \right) \right), 0 \right) dx.$$

Next fix  $0 \leq \ell < r$  and, under  $\mathbf{P}_{\nu^M}$ , let  $W$  be a Uniform $[\ell, r]$  random variable, independent of  $\mathbf{z}$ . Then

$$\frac{G_{\lfloor WM^3 \rfloor}^{\mathbf{z}}}{(36(W + \varepsilon))^{1/3} M} + \frac{1}{2} \xrightarrow{d} G \text{ as } M \rightarrow \infty$$

where  $G$  is a Beta(2, 2) random variable.

**Proof** (Proof of Proposition 9) Similarly to the previous case, the joint law of  $W$  and  $G_{\lfloor WM^3 \rfloor}^{\mathbf{z}}$  is given by

$$\mathbf{P}_{\nu^M} \{ G_{\lfloor WM^3 \rfloor}^{\mathbf{z}} = j, W \in dt \} = \frac{dt}{r - \ell} \mathbf{P}_{\nu^M} \left\{ G_{\lfloor tM^3 \rfloor}^{\mathbf{z}}(\emptyset) = j \right\} \mathbf{1}_{[t \in [\ell, r]]}.$$

Again, we write  $\mathbf{P}$  instead of  $\mathbf{P}_{v,M}$  and  $G_k$  instead of  $G_k^z(\emptyset)$ . For  $0 \leq a < b \leq 1$ , we have

$$\mathbf{P} \left\{ \frac{G_{\lfloor WM^3 \rfloor}}{(36(W + \varepsilon))^{1/3} M} + \frac{1}{2} \in [a, b] \right\} = \frac{1}{r - \ell} \int_{\ell}^r \mathbf{P} \{ G_{\lfloor tM^3 \rfloor} \in [\lfloor M(a - 1/2)c \rfloor, \lfloor M(b - 1/2)c \rfloor] \} dt, \tag{23}$$

where  $c = (36(t + \varepsilon))^{1/3}$ .

Define  $v_0(x) = \max \left( \frac{3}{4} \left( \left( \frac{2}{9\varepsilon} \right)^{1/3} - \left( \frac{2x^2}{9\varepsilon} \right) \right), 0 \right)$  and for  $t \in [0, \lceil r \rceil]$  let

$$v(x, t) = \max \left( \frac{3}{4} \left( \left( \frac{2}{9(t + \varepsilon)} \right)^{1/3} - \left( \frac{2x^2}{9(t + \varepsilon)} \right) \right), 0 \right).$$

Then  $v \equiv v_P$  and  $v_0(x) = v_P(x, 0)$ , where  $v_P$  is as in Proposition 5, applied with  $T = \lceil r \rceil$ , say. We also let  $U_j^n = U_j^n(v_0, \Delta_x^M, \Delta_t^M)$  be as defined in Proposition 5.

Note that  $v^M(\{j\}) = \int_{j/M}^{(j+1)/M} v_0(x) dx$ , and by Proposition 7, for all  $n \in \mathbb{N}$ , and  $j \in \mathbb{Z}$  we have that  $U_j^n = M \cdot \mathbf{P}\{G_n = j\}$ . Let  $v^M : \mathbb{R} \times (0, \infty) \rightarrow \mathbb{R}$  be the function which takes the value  $U_j^n$  on  $[j/M, (j + 1)/M) \times [n/M^3, (n + 1)/M^3)$  for  $n \in \mathbb{N}$ ,  $j \in \mathbb{Z}$ . Then for  $0 \leq a < b \leq 1$ ,

$$\begin{aligned} & \int_{\ell}^r \int_{(a-1/2)(36(t+\varepsilon))^{1/3}}^{(b-1/2)(36(t+\varepsilon))^{1/3}} v^M(x, t) dx dt \\ &= \int_{\ell}^r \int_{(a-1/2)(36(t+\varepsilon))^{1/3}}^{(b-1/2)(36(t+\varepsilon))^{1/3}} M \cdot \mathbf{P}\{G_{\lfloor tM^3 \rfloor} = \lfloor Mx \rfloor\} dx dt. \end{aligned}$$

As in the proof of Proposition 5 it follows that

$$\begin{aligned} & \int_{\ell}^r \int_{(a-\frac{1}{2})(36(t+\varepsilon))^{1/3}}^{(b-\frac{1}{2})(36(t+\varepsilon))^{1/3}} v^M(x, t) dx dt \\ & \geq \int_{\ell}^r \mathbf{P} \{ G_{\lfloor tM^3 \rfloor} \in (\lfloor M(a - 1/2)c \rfloor, \lfloor M(b - 1/2)c \rfloor] \} dt, \tag{24} \end{aligned}$$



and

$$\int_{\ell}^r \int_{\ell}^{(b-\frac{1}{2})(36(t+\varepsilon))^{1/3}} v^M(x, t) dx dt \leq \int_{\ell}^r \mathbf{P} \{ G_{\lfloor tM^3 \rfloor} \in ([M(a - 1/2)c], [M(b - 1/2)c]) \} dt. \tag{25}$$

By Proposition 5  $v_P$  is the BV entropy weak solution to the initial value problem  $\partial_t v - \partial_{xx} v^2/2 = 0$  with initial condition  $v_0$ , by Theorem 3,

$$\begin{aligned} \lim_{M \rightarrow \infty} \int_{\ell}^r \int_{\ell}^{(b-\frac{1}{2})(36(t+\varepsilon))^{1/3}} v^M(x, t) dx dt &= \int_{\ell}^r \int_{\ell}^{(b-\frac{1}{2})(36(t+\varepsilon))^{1/3}} v(x, t) dx dt \\ &= \int_{\ell}^r 3(b^2 - a^2) - 2(b^3 - a^3) dt \\ &= (r - \ell)(3(b^2 - a^2) - 2(b^3 - a^3)). \end{aligned}$$

Combining this with (23), (24), (25), we obtain that for all  $b \in (0, 1)$ ,

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{G_{\lfloor WM^3 \rfloor}}{(36(W + \varepsilon))^{1/3} M} + \frac{1}{2} \in [a, b] \right\} &= 3(b^2 - a^2) - 2(b^3 - a^3) \\ &= \int_a^b 6x(1 - x) dx. \end{aligned}$$

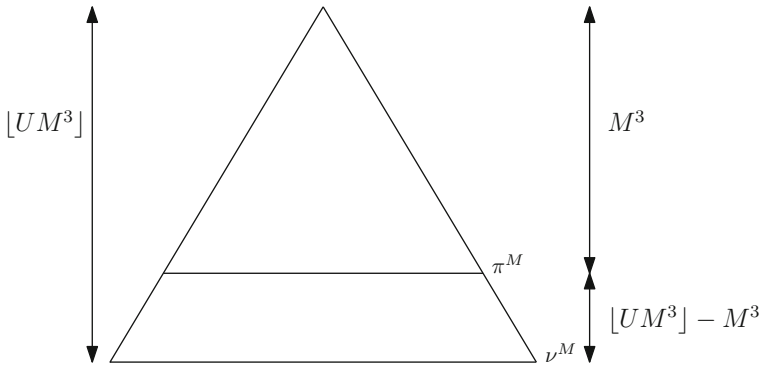
Since the density for a Beta(2, 2) random variable is  $6x(1 - x)\mathbf{1}_{[x \in [0, 1]]}$ , the result follows. □

### 4 Proofs of Theorems 1 and 2.

In order to strengthen Propositions 8 and 9 to remove the time averaging, we shall use the following coupling lemmas.

**Lemma 10** Fix two probability distributions  $\mu, \nu$  on  $\mathbb{Z}$  and a coupling  $(X, Y)$  of  $\mu$  and  $\nu$ , and write  $\alpha = \mathbf{P}(X > Y)$ . Fix  $k \geq 1$ , let  $\mu_k$  be the law of  $B_k^Z(\emptyset)$  under  $\mathbf{P}_\mu$  and let  $\nu_k$  be the law of  $B_k^Z(\emptyset)$  under  $\mathbf{P}_\nu$ . Then there exists a coupling  $(X', Y')$  of  $\mu_k, \nu_k$  such that  $\mathbf{P}(X' > Y') = \alpha$ .

**Lemma 11** Fix two probability distributions  $\mu, \nu$  on  $\mathbb{Z}$  and a coupling  $(X, Y)$  of  $\mu$  and  $\nu$ , and write  $\alpha = \mathbf{P}(X > Y)$ . Fix  $k \geq 1$ , let  $\mu_k$  be the law of  $G_k^Z(\emptyset)$  under  $\mathbf{P}_\mu$  and



**Fig. 1** If the level- $\lfloor UM^3 \rfloor$  inputs are  $\nu^M$ -distributed then the resulting level- $M^3$  outputs are  $\pi^M$ -distributed. In other words, for all  $\ell \in \mathbb{N}$  and  $v \in \mathcal{L}_{M^3}$  we have  $\mathbf{P}_{\nu^M} \{G_{\lfloor UM^3 \rfloor}^{\mathbf{z}}(v) = \ell\} = \pi^M\{\ell\} = \mathbf{P}_{\pi^M} \{z_v = \ell\}$ , and so also  $\mathbf{P}_{\nu^M} \{G_{\lfloor UM^3 \rfloor}^{\mathbf{z}}(\emptyset) = \ell\} = \mathbf{P}_{\pi^M} \{G_{M^3}^{\mathbf{z}}(\emptyset) = \ell\}$

let  $\nu_k$  be the law of  $G_k^{\mathbf{z}}(\emptyset)$  under  $\mathbf{P}_\nu$ . Then there exists a coupling  $(X', Y')$  of  $\mu_k, \nu_k$  such that  $\mathbf{P}(X' > Y') \leq \alpha$ .

Both lemmas are proved by the explicit construction of a coupling with the claimed property. In ‘‘Appendix A’’ we prove Lemma 11 in detail, then briefly explain how to modify the construction to prove Lemma 10, since the constructions are nearly identical.

**Proof** (Proof of Theorem 2.) We aim to prove that for any field of IID random variables  $\mathbf{Z} = (Z_v, v \in T)$ , we have

$$(36n)^{-1/3} G_n^{\mathbf{z}}(\emptyset) + 1/2 \xrightarrow{d} G_\infty,$$

where, here and later in the proof,  $G_\infty$  denotes a Beta(2, 2)-distributed random variable. We first handle the case that  $\mathbf{Z} = \mathbf{0}$ , or equivalently that the random variables  $Z_v$  are  $\delta_0$ -distributed. At the end of the proof we explain how to extend from this case to general input distributions.

Fix  $\varepsilon \in (0, 1)$  and let  $U$  be a Uniform $[1, 1 + \varepsilon]$  random variable, independent of all other randomness in the system. We recall the definition of  $\nu^M$  from Proposition 9: for  $M > 0$ ,  $\nu^M$  is the probability measure on  $\mathbb{Z}$  such that for all  $j \in \mathbb{Z}$

$$\nu^M(\{j\}) = \int_{j/M}^{(j+1)/M} \max\left(\frac{3}{4} \left(\left(\frac{2}{9\varepsilon}\right)^{1/3} - \left(\frac{2x^2}{9\varepsilon}\right)\right), 0\right) dx.$$

Now for  $M > 0$  such that  $M^3 \in \mathbb{N}$ , let  $\pi^M$  be the law of  $G_{\lfloor UM^3 \rfloor - M^3} := G_{\lfloor UM^3 \rfloor - M^3}^{\mathbf{z}}(\emptyset)$  under  $\mathbf{P}_{\nu^M}$ ; for  $v \in \mathcal{L}_{M^3}$ , this is also the law of  $G_{\lfloor UM^3 \rfloor}^{\mathbf{z}}(v)$  under  $\mathbf{P}_{\nu^M}$ . We will use the fact that the law of  $G_{\lfloor UM^3 \rfloor} = G_{\lfloor UM^3 \rfloor}(\emptyset)$  under  $\mathbf{P}_{\nu^M}$  is the same as the law of  $G_{M^3} = G_{M^3}(\emptyset)$  under  $\mathbf{P}_{\pi^M}$ ; see Fig. 1.

Since  $U - 1$  is Uniform $[0, \varepsilon]$ , by Proposition 9 applied with  $\ell = 0, r = \varepsilon$  we have that

$$\frac{G_{\lfloor UM^3 \rfloor - M^3}}{(36(U - 1 + \varepsilon))^{1/3} M} + \frac{1}{2} \xrightarrow{d} G_\infty \tag{26}$$

as  $M \rightarrow \infty$  along values with  $M^3 \in \mathbb{N}$ .

Therefore

$$\begin{aligned} & \mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} > \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \\ & \leq \mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} > \frac{(36(U - 1 + \varepsilon))^{1/3} M}{2} \right\} \\ & = \mathbf{P}_{v^M} \left\{ \frac{G_{\lfloor UM^3 \rfloor - M^3}}{(36(U - 1 + \varepsilon))^{1/3} M} + \frac{1}{2} > 1 \right\}. \end{aligned}$$

By (26), the final probability tends to 0 as  $M \rightarrow \infty$ , so

$$\mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} > \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \rightarrow 0 \text{ as } M \rightarrow \infty,$$

and by a similar analysis one finds that

$$\mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} < -\frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \rightarrow 0 \text{ as } M \rightarrow \infty.$$

Fixing  $\alpha > 0$ , we can therefore choose  $M_0$  large enough that for  $M \geq M_0$ ,

$$\mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} > \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \leq \alpha, \tag{27}$$

and

$$\mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} < -\frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \leq \alpha.$$

Using the definition of  $\pi^M$ , the second inequality implies that for  $M \geq M_0$ , under  $\mathbf{P}_{\pi^M}$ , the inputs  $(z_v, v \in \mathcal{L}_{M^3}) =: (G_{M^3}(v), v \in \mathcal{L}_{M^3})$  are such that

$$\mathbf{P}_{\pi^M} \left\{ z_v < -\frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} = \mathbf{P}_{v^M} \left\{ G_{\lfloor UM^3 \rfloor - M^3} < -\frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \leq \alpha \tag{28}$$

Note that we may view (28) as stating that there exists a coupling  $(X, Y)$  of  $\delta_0$ , a Dirac mass at 0, and  $\pi_M$ , such that  $\mathbf{P} \left\{ Y + \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} < X \right\} \leq \alpha$ . We can then apply

Lemma 11 to find a coupling  $(X_M, Y_M)$  of  $G_{M^3}$  under  $\mathbf{P}_{\delta_0}$  and  $G_{M^3}$  under  $\mathbf{P}_{\pi^M}$  such that

$$\mathbf{P} \left\{ Y_M + \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} < X_M \right\} \leq \alpha.$$

For all  $x \in \mathbb{R}$ , this gives

$$\begin{aligned} & \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} > x \right\} \\ &= \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ X_M > (36)^{1/3} M \left( x - \frac{1}{2} \right) \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ Y_M + \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \geq X_M > (36)^{1/3} M \left( x - \frac{1}{2} \right) \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ Y_M + \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} > (36)^{1/3} M \left( x - \frac{1}{2} \right) \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{Y_M}{(36)^{1/3} M} + \frac{1}{2} > x - \frac{(2\varepsilon)^{1/3}}{2} \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\pi^M} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} > x - \frac{(2\varepsilon)^{1/3}}{2} \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\nu^M} \left\{ \frac{G_{\lfloor UM^3 \rfloor}}{(36)^{1/3} M} + \frac{1}{2} > x - \frac{(2\varepsilon)^{1/3}}{2} \right\}. \end{aligned}$$

Combining this result with the fact that  $(1 + 2\varepsilon)/(U + \varepsilon) > 1$ , we get that for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} & \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \\ &\geq \liminf_{M \rightarrow \infty} \mathbf{P}_{\nu^M} \left\{ \frac{G_{\lfloor UM^3 \rfloor}}{(36)^{1/3} M} + \frac{1}{2} \leq x - \frac{(2\varepsilon)^{1/3}}{2} \right\} - \alpha, \\ &\geq \liminf_{M \rightarrow \infty} \mathbf{P}_{\nu^M} \left\{ (1 + 2\varepsilon)^{1/3} \left( \frac{G_{\lfloor UM^3 \rfloor}}{(36(U + \varepsilon))^{1/3} M} + \frac{1}{2} \right) \leq x - \frac{(2\varepsilon)^{1/3}}{2} \right\} - \alpha \\ &= \mathbf{P} \left\{ G_\infty \leq \frac{1}{(1 + 2\varepsilon)^{1/3}} \left( x - \frac{(2\varepsilon)^{1/3}}{2} \right) \right\} - \alpha; \end{aligned} \tag{29}$$

the last equality follows from Proposition 9.

Similarly, we may view (27) as stating that for  $M$  sufficiently large there exists a coupling  $(X, Y)$  with  $X$  having distribution  $\pi^M$  and  $Y$  having distribution  $\delta_0$ , such that  $\mathbf{P} \left\{ X - \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} > Y \right\} \leq \alpha$ . For  $M$  large we may thus apply Lemma 11 to find

a coupling  $(X_M, Y_M)$  of  $G_{M^3}$  under  $\mathbf{P}_{\pi^M}$  and  $G_{M^3}$  under  $\mathbf{P}_{\delta_0}$  such that

$$\mathbf{P} \left\{ X_M - \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} > Y_M \right\} \leq \alpha. \tag{30}$$

This is a different coupling from the one used just above, but we allow ourselves to recycle the notation  $(X_M, Y_M)$  as the previous coupling plays no further role. (Note that the marginals of the coupling have switched places.)

It follows from (30) that for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} & \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \\ &= \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ Y_M \leq (36)^{1/3} M \left( x - \frac{1}{2} \right) \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ X_M - \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} < Y_M \leq (36)^{1/3} M \left( x - \frac{1}{2} \right) \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ X_M \leq (36)^{1/3} M \left( x - \frac{1}{2} \right) + \frac{(36 \cdot 2\varepsilon)^{1/3} M}{2} \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\pi^M} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x + \frac{(2\varepsilon)^{1/3}}{2} \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\nu^M} \left\{ \frac{G_{\lfloor UM^3 \rfloor}}{(36(U + \varepsilon))^{1/3} M} + \frac{1}{2} \leq x + \frac{(2\varepsilon)^{1/3}}{2} \right\} \\ &= \alpha + \mathbf{P} \left\{ G_\infty \leq x + \frac{(2\varepsilon)^{1/3}}{2} \right\} \end{aligned}$$

where the final inequality holds since  $U + \varepsilon > 1$ , and the last equality again holds by Proposition 9. Using this in combination with (29) gives

$$\begin{aligned} \mathbf{P} \left\{ G_\infty \leq \frac{1}{(1 + 2\varepsilon)^{1/3}} \left( x - \frac{(2\varepsilon)^{1/3}}{2} \right) \right\} - \alpha &\leq \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \\ &\leq \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \\ &\leq \alpha + \mathbf{P} \left\{ G_\infty \leq x + \frac{(2\varepsilon)^{1/3}}{2} \right\}, \end{aligned}$$

for all  $x \in \mathbb{R}$ . Since  $\varepsilon \in (0, 1)$  was arbitrary we can let  $\varepsilon \rightarrow 0$  to obtain that for all  $x \in \mathbb{R}$

$$\begin{aligned} \mathbf{P}\{G_\infty \leq x\} - \alpha &\leq \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \\ &\leq \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{G_{M^3}}{(36)^{1/3} M} + \frac{1}{2} \leq x \right\} \leq \alpha + \mathbf{P}\{G_\infty \leq x\}. \end{aligned}$$

Since  $\alpha > 0$  was also arbitrary, we may take  $\alpha \rightarrow 0$  to get that under  $\mathbf{P}_{\delta_0}$ , as  $M \rightarrow \infty$  along  $M^3 \in \mathbb{N}$ ,

$$\frac{G_{M^3}}{(36)^{1/3}M} + \frac{1}{2} \xrightarrow{d} G_\infty.$$

This handles the case that  $\mathbf{Z} = \mathbf{0}$ ; we finish the proof by explaining how to extend to general input distributions. First notice that for any  $k \in \mathbb{Z}$ , for all  $n \in \mathbb{N}$  the collection of random variables  $(G_n^{\mathbf{0}}(v) + k, v \in \mathcal{T}_n)$  has the same law as  $(G_n^{\mathbf{k}}(v), v \in \mathcal{T}_n)$ , where  $\mathbf{k}$  denotes an input vector where all entries equal  $k$ . It follows that

$$\frac{G_{M^3}^{\mathbf{k}}(\emptyset)}{(36)^{1/3}M} + \frac{1}{2} \stackrel{d}{=} \frac{G_{M^3}^{\mathbf{0}}(\emptyset) + k}{(36)^{1/3}M} + \frac{1}{2} \xrightarrow{d} G_\infty,$$

since  $k/((36)^{1/3}M) \rightarrow 0$  as  $M \rightarrow \infty$ . This proves convergence whenever all inputs take a common, fixed value.

Now suppose the entries of  $\mathbf{Z}$  are IID with some common law  $\xi$ . Fix  $\beta > 0$  and let  $K = K(\beta)$  be large enough that  $\xi([-K, K]) > 1 - \beta$ . Then for all  $v \in \mathcal{T}$  we have  $\mathbf{P}\{Z_v > K\} < \beta$ , so by Lemma 11, for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{G_{M^3}^{\mathbf{Z}}(\emptyset)}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} &= \limsup_{M \rightarrow \infty} \mathbf{P}_\xi \left\{ \frac{G_{M^3}}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} \\ &< \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_K} \left\{ \frac{G_{M^3}}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} + \beta \\ &= \mathbf{P}\{G_\infty \leq x\} + \beta, \end{aligned}$$

the last equality holding since we already established distributional convergence for constant input. It likewise follows that

$$\begin{aligned} \liminf_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{G_{M^3}^{\mathbf{Z}}(\emptyset)}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} &= \liminf_{M \rightarrow \infty} \mathbf{P}_\xi \left\{ \frac{G_{M^3}}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} \\ &> \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_{-K}} \left\{ \frac{G_{M^3}}{(36)^{1/3}M} + \frac{1}{2} \leq x \right\} - \beta \\ &= \mathbf{P}\{G_\infty \leq x\} - \beta; \end{aligned}$$

combining the two preceding displays and taking  $\beta \rightarrow 0$ , the result follows. □

**Proof** (Proof of Theorem 1) We aim to prove that for any field of IID random variables  $\mathbf{Z} = (Z_v, v \in \mathcal{T})$ , we have

$$(4qn)^{-1/2} B_n^{\mathbf{Z}}(\emptyset) \xrightarrow{d} B_\infty,$$

where, here and later in the proof,  $B_\infty$  denotes a Beta(2, 1)-distributed random variable. We restrict our attention to the case that  $\mathbf{Z} = \mathbf{0}$ ; the extension to general input

distributions proceeds exactly as in the proof of Theorem 2, using Lemma 10 in place of Lemma 11.

Fix  $\varepsilon \in (0, 1)$ , and recall the definition of  $\mu^M$  from Proposition 8: for  $M > 0$ ,  $\mu^M$  is the probability measure on  $\mathbb{Z}$  such that for all  $j \in \mathbb{Z}$

$$\mu^M(\{j\}) = \int_{j/M}^{(j+1)/M} \frac{x}{2q\varepsilon} \mathbf{1}_{[x \in [0, \sqrt{4q\varepsilon}]]} dx.$$

Next, let  $U$  be a Uniform $[1, 1 + \varepsilon]$  random variable, and for  $M > 0$  such that  $M^2 \in \mathbb{N}$ , let  $\pi^M$  be the law of  $B_{\lfloor UM^2 \rfloor - M^2}$  under  $\mathbf{P}_{\mu^M}$ ; this is also the law of  $B_{\lfloor UM^2 \rfloor}^Z(v)$  under  $\mathbf{P}_{\mu^M}$  for nodes  $v \in \mathcal{L}_{M^2}$ . Also, the law of  $B_{\lfloor UM^2 \rfloor}$  under  $\mathbf{P}_{\mu^M}$  is the same as the law of  $B_{M^2}$  under  $\mathbf{P}_{\pi^M}$ . (We have used the shorthand  $B_n = B_n^Z(\emptyset)$  repeatedly in this paragraph.)

Since  $U - 1$  is Uniform $[0, \varepsilon]$ , by Proposition 8 applied with  $\ell = 0, r = \varepsilon$  we have that

$$\frac{B_{\lfloor UM^2 \rfloor - M^2}}{(4q(U - 1 + \varepsilon))^{1/2} M} \xrightarrow{d} B, \tag{31}$$

as  $M \rightarrow \infty$  along values with  $M^2 \in \mathbb{N}$  where  $B$  is a Beta(2, 1) random variable. Therefore

$$\begin{aligned} & \mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} > (4q \cdot (2\varepsilon))^{1/2} M \right\} \\ & \leq \mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} > (4q(U - 1 + \varepsilon))^{1/2} M \right\} \\ & = \mathbf{P}_{\mu^M} \left\{ \frac{B_{\lfloor UM^2 \rfloor - M^2}}{(4q(U - 1 + \varepsilon))^{1/2} M} > 1 \right\}. \end{aligned}$$

By (31), the final probability tends to 0 as  $M \rightarrow \infty$  along values with  $M^2 \in \mathbb{N}$ , so

$$\mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} > (4q \cdot (2\varepsilon))^{1/2} M \right\} \rightarrow 0 \text{ as } M \rightarrow \infty. \tag{32}$$

Therefore, for any  $\alpha > 0$ , we can choose  $M_0$  large enough such that for  $M \geq M_0$ ,

$$\mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} > (4q \cdot (2\varepsilon))^{1/2} M \right\} \leq \alpha.$$

Also, since the dynamics are monotone non-decreasing, for all  $M$  we have

$$\mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} < 0 \right\} = 0. \tag{33}$$

Under  $\mathbf{P}_{\pi^M}$ , for  $M \geq M_0$  the inputs  $(z_v, v \in \mathcal{L}_{M^2}) =: (B_{M^2}(v), v \in \mathcal{L}_{M^2})$  are such that

$$\mathbf{P}_{\pi^M} \left\{ z_v > (4q \cdot (2\varepsilon))^{1/2} M \right\} = \mathbf{P}_{\mu^M} \left\{ B_{\lfloor UM^2 \rfloor - M^2} > (4q \cdot (2\varepsilon))^{1/2} M \right\} \leq \alpha, \tag{34}$$

and

$$\mathbf{P}_{\pi^M} \{z_v < 0\} = \mathbf{P}_{\mu^M} \{B_{\lfloor UM^2 \rfloor - M^2} < 0\} = 0. \tag{35}$$

By (34) we can apply Lemma 10 to find a coupling  $(X_M, Y_M)$  of  $B_{M^2}$  under  $\mathbf{P}_{\pi^M}$  and  $B_{M^2}$  under  $\mathbf{P}_{\delta_0}$  such that

$$\mathbf{P} \left\{ X_M > (4q \cdot (2\varepsilon))^{1/2} M + Y_M \right\} \leq \alpha.$$

Then for all  $x \in \mathbb{R}$ , we obtain the bound

$$\begin{aligned} & \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} \\ &= \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{Y_M}{(4q)^{1/2} M} \leq x \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{X_M}{(4q)^{1/2} M} - (2\varepsilon)^{1/2} < \frac{Y_M}{(4q)^{1/2} M} \leq x \right\} \\ &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P} \left\{ \frac{X_M}{(4q)^{1/2} M} < x + (2\varepsilon)^{1/2} \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\pi^M} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} < x + (2\varepsilon)^{1/2} \right\} \\ &= \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\mu^M} \left\{ \frac{B_{\lfloor UM^2 \rfloor}}{(4q)^{1/2} M} < x + (2\varepsilon)^{1/2} \right\}. \end{aligned}$$

Using that  $U + \epsilon > 1$ , then using Proposition 8 applied with  $\ell = 1$ , and  $r = 1 + \varepsilon$ , this gives

$$\limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} \tag{36}$$

$$\begin{aligned} &\leq \alpha + \limsup_{M \rightarrow \infty} \mathbf{P}_{\mu^M} \left\{ \frac{B_{\lfloor UM^2 \rfloor}}{(4q(U + \epsilon))^{1/2} M} < x + (2\varepsilon)^{1/2} \right\} \\ &= \alpha + \mathbf{P} \left\{ B_\infty \leq x + (2\varepsilon)^{1/2} \right\}. \end{aligned} \tag{37}$$

Likewise, using (35) and Lemma 10 we may find a (different) coupling  $(X_M, Y_M)$  of  $B_{M^2}$  under  $\mathbf{P}_{\delta_0}$  and  $B_{M^2}$  under  $\mathbf{P}_{\pi^M}$  such that  $\mathbf{P} \{X_M > Y_M\} = 0$ . In other words,  $B_{M^2}$  under  $\mathbf{P}_{\delta_0}$  is stochastically dominated by  $B_{M^2}$  under  $\mathbf{P}_{\pi^M}$ . It follows that for all  $x \in \mathbb{R}$ ,

$$\liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} \geq \liminf_{M \rightarrow \infty} \mathbf{P}_{\pi^M} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\}.$$



Using that  $(1 + 2\epsilon)/(U + \epsilon) > 1$ , then using Proposition 8 applied with  $\ell = 1$ , and  $r = 1 + \epsilon$ , we get that for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} &\geq \liminf_{M \rightarrow \infty} \mathbf{P}_{\pi^M} \left\{ (1 + 2\epsilon)^{1/2} \frac{B_{M^2}}{(4q(U + \epsilon))^{1/2} M} \leq x \right\} \\ &= \liminf_{M \rightarrow \infty} \mathbf{P}_{\mu^M} \left\{ (1 + 2\epsilon)^{1/2} \frac{B_{\lfloor UM^2 \rfloor}}{(4q(U + \epsilon))^{1/2} M} \leq x \right\} \\ &= \mathbf{P} \left\{ (1 + 2\epsilon)^{1/2} B_\infty \leq x \right\}. \end{aligned}$$

Combining this with (36) gives

$$\begin{aligned} &\mathbf{P} \left\{ (1 + 2\epsilon)^{1/2} B_\infty \leq x \right\} \\ &\leq \liminf_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} \\ &\leq \limsup_{M \rightarrow \infty} \mathbf{P}_{\delta_0} \left\{ \frac{B_{M^2}}{(4q)^{1/2} M} \leq x \right\} \leq \alpha + \mathbf{P} \left\{ B_\infty \leq x + (2\epsilon)^{1/2} \right\}. \end{aligned}$$

Since  $\alpha > 0$  and  $\epsilon \in (0, 1)$  were arbitrary, it follows that  $((4q)^{1/2} M)^{-1} B_{M^2}^0 \xrightarrow{d} B_\infty$ , as required.  $\square$

### 5 Conclusion

There are several natural avenues for extensions of our results which are deserving of study. The first three points below relate specifically to hipster random walks.

- The robustness of Theorems 1 and 2 with respect to the law of the inputs is due to the fact that, by the coupling lemmas, changes to the input law have an essentially additive effect on the dynamics, and this effect vanishes after rescaling.

We can say less about robustness with respect to changes in the step distribution. In general, for a hipster random walk with bounded steps  $(D_v, v \in \mathcal{T})$ , we would expect that if the steps are centred then one should expect a version of Theorem 2 to hold (with a normalizing constant depending on the step distribution), whereas if the steps have non-zero mean then a version of Theorem 1 should hold.

As a special case of the second assertion, one might try to extend Theorem 1 to hipster random walks with non-negative, bounded integer steps. If the steps take values in  $\{0, 1, \dots, M\}$  with

$$\mathbf{P} \{D_v = i\} = c_i,$$

then one would obtain the recurrence

$$r_k^{n+1} = r_k^n(1 - r_k^n) + \sum_{i=0}^M c_i (r_{k-i}^n)^2.$$

Another natural special case to consider is that of asymmetric simple random walk, where  $\mathbf{P}\{D_v = 1\} = q = 1 - \mathbf{P}\{D_v = -1\}$ , for  $q \neq 1/2$ . This case may even be accessible with a variant of the techniques of the current work, since the resulting finite difference scheme appears to fit within the framework of [2]. However, although the resulting PDE has the same long-term behaviour as Burgers' equation, at any finite time the solution behaves like a mixture of Burgers' equation and the porous membrane equation, and this seems to complicate the analysis.

- For unbounded step distributions—and in particular for heavy-tailed step distributions—one should be able to construct more exotic behaviour. It would be quite interesting to understand whether there is a dictionary between the possible behaviours of the hipster random walk and the various solutions of the associated PDEs.
- It would be quite natural to try to relax the requirement that in order for a random displacement to occur, the child values must be precisely equal. We believe it would be very interesting to understand the behaviour of models of the following form. Fix  $C > 0$ , and set

$$F_n^z(v) = \begin{cases} z_v & \text{if } v \in \mathcal{L}_n \\ F_n^z(v_0) & \text{if } v \in \mathcal{T}_{n-1}, |F_n^z(v_0) - F_n^z(v_1)| > C, \\ F_n^z(v_0) + D_v & \text{if } v \in \mathcal{T}_{n-1}, |F_n^z(v_0) - F_n^z(v_1)| \leq C. \end{cases} \quad (38)$$

Here, as before,  $\mathbf{z}$  is a vector of fixed or random input values, and  $(D_v, v \in \mathcal{T})$  are IID random variables. In this model, if the values at the children of  $v$  differ by more than  $C$  then the output is the value of the left child (one could instead output the value of a random child as was done above, but the symmetry of the tree implies that the law of  $F_n^z(v)$  is the same in either case). If the child values differ by at most  $C$ , then the output is the left child's value plus a random displacement.

We conclude with some further potential research directions, in the spirit of this work but not specifically related to hipster random walks.

- Instead of a hipster random walk, one may consider a fomo random walk (fomo stands for “fear of missing out”). Here the combination rule is

$$f_v(x, y) = \begin{cases} x & \text{if } x = y \\ xA_v + y(1 - A_v) + D_v & \text{if } x \neq y, \end{cases}$$

where as before the  $A_v$  are Bernoulli(1/2)-distributed. For this dynamics, walkers are happy when they have company, and only move when they find themselves alone. The recurrence relation for the fomo symmetric simple random walk (when

the  $D_v$  are centred  $\pm 1$  random variables) can be written as

$$r_k^{n+1} - r_k^n = \frac{1}{2} (r_{k-1}^n - 2r_k^n + r_{k+1}^n) - \frac{1}{2} \left( (r_{k-1}^n)^2 - 2(r_k^n)^2 + (r_{k+1}^n)^2 \right),$$

which one may suppose converges to a solution of the PDE  $\partial_t u = (\partial_{xx} u - \partial_{xx}(u^2))/2$ . The presence of a diffusive term should make this model's analysis somewhat more straightforward.

- More generally, what conditions on a discrete difference equation imply that it can be interpreted as describing the evolution of the distribution function for an integer-valued recursive distributional equation? Conversely, which integer RDEs yield difference equations which may be interpreted as numerical schemes (and fruitfully analyzed using techniques from numerical analysis)? Also: can this approach be of any use in settings where integrality is not preserved (such as that of the random hierarchical lattice)?
- This paper imports theorems from numerical analysis to prove probabilistic results; perhaps information can also flow in the other direction. The development of any reasonably general techniques for analyzing such probabilistic systems would seem likely to simultaneously establish new stability/convergence results for numerical approximations of PDEs. Thus far, we are not aware of any theorems in the numerical analysis literature which have been proved in such a way.

**Acknowledgements** We thank Lia Bronsard for directing us to the reference [2]. We further thank Rustum Choksi, Jessica Lin, Pascal Maillard and Robin Vacus for useful conversations, and an anonymous referee for their their helpful comments on the paper.

**Funding** LAB was supported by National Sciences and Engineering Research Council (NSERC) Discovery Grant 643473. LAB and LD were supported by Fonds de Recherche du Québec - Nature et Technologies (FRQNT) Team Grant No. 206470. LD was supported by an NSERC Discovery Grant. RM was supported by NSERC USRA Grant No. 539343 and by FRQNT Grant No. 282161.

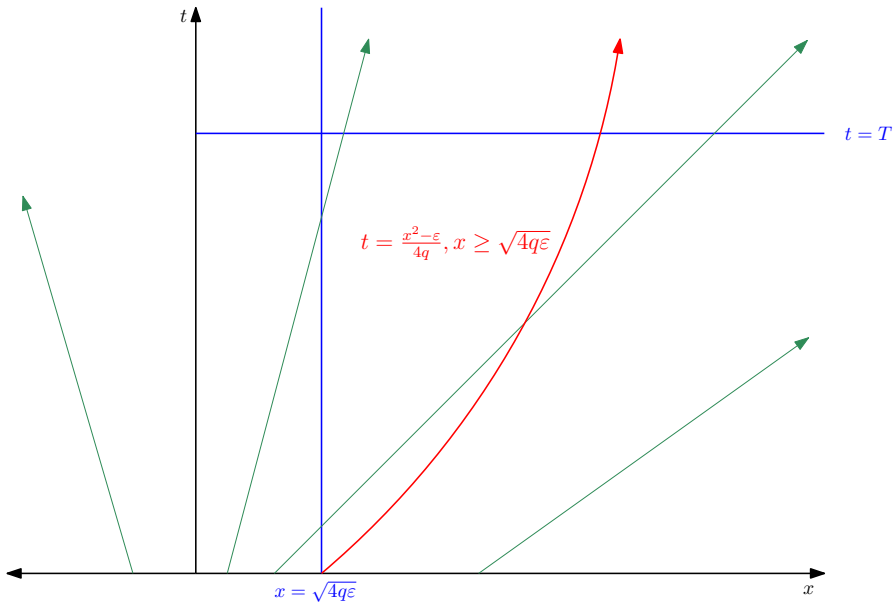
## A Remaining proofs

In this section we prove Propositions 4 and 5, and Lemmas 10 and 11.

**Proof** (Proof of Proposition 4) We first verify monotonicity. Since  $K_B \equiv 0$ , the function  $S : \mathbb{R}^3 \rightarrow \mathbb{R}$  defined by (16) is

$$\begin{aligned} S(u^-, u, u^+) &= \frac{u}{\Delta_t^M} - \frac{q}{\Delta_x^M} (u^2 - (u^-)^2) \\ &= M^2 \cdot u - qM \cdot (u^2 - (u^-)^2). \end{aligned}$$

The function  $S$  is nondecreasing in all of its arguments on  $[0, M/(2q)]$ , so for  $M$  sufficiently large, the approximation scheme  $(U_j^n(u_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  is monotone on  $[0, 1/(\sqrt{q\varepsilon})]$ .



**Fig. 2** The region of integration and different “sign-change” regimes for different values of  $c$ . Read from left to right, the straight, green lines represent equations of the form  $t = x/(2qc) - \epsilon$  for  $c \leq 0$ ,  $c \in (0, 1/\sqrt{q(T + \epsilon)})$ ,  $c \in (1/\sqrt{q(T + \epsilon)}, 1/\sqrt{q\epsilon}]$  and  $c > 1/\sqrt{q\epsilon}$ , respectively

We now turn to the first claim of the proposition. The function  $u = u_B$  is clearly of bounded variation since it has bounded Lipschitz constant in the compact region

$$\{(x, t) : 0 \leq x \leq \sqrt{4q(t + \epsilon)}, 0 \leq t \leq T\}$$

and is zero outside this region. Also, by definition  $u(\cdot, 0) \equiv u_0$  and it is clear that  $f(u_0) - K'(u_0)$  is of bounded variation. To prove the proposition, it then remains to show that

$$\int_{\mathbb{R} \times [0, T]} \text{sgn}(u - c) \cdot \left( (u - c)\partial_t \phi + q(u^2 - c^2)\partial_x \phi \right) dt dx + \int_{\mathbb{R}} |u_0 - c| \phi(x, 0) dx \geq 0. \quad (39)$$

for all  $c \in \mathbb{R}$  and all non-negative  $\phi \in C^\infty(\mathbb{R} \times [0, T])$  with compact support such that  $\phi|_{t=T} \equiv 0$ . Before beginning the analysis, note that  $\phi, \partial_t \phi, \partial_x \phi$  and  $u$  are all bounded on  $\mathbb{R} \times [0, T]$  hence there is no issue when changing the order of integration. The proof naturally splits into four cases according to whether  $c \leq 0$ ,  $c \in (0, \sqrt{1/(q(T + \epsilon))}]$ ,  $c \in (\sqrt{1/(q(T + \epsilon))}, \sqrt{1/(q\epsilon)}]$ , and  $c > \sqrt{1/(q\epsilon)}$ ; see Fig. 2.

The most involved case is when  $c \in (\frac{1}{\sqrt{q(T + \epsilon)}}, \frac{1}{\sqrt{q\epsilon}}]$ . We will provide a full proof only for this case. Define the two regions

$$R^- := \left\{ (x, t) : x \leq \min(2qc(t + \epsilon), \sqrt{4q(t + \epsilon)}), t \in [0, T] \right\},$$

and

$$\mathbb{R}^+ := \left\{ (x, t) : 2qc(t + \varepsilon) \leq x \leq \sqrt{4q(t + \varepsilon)}, t \in [0, T] \right\},$$

and write  $n_{\mathbb{R}^-}$ ,  $n_{\mathbb{R}^+}$  for their respective outward normal vectors.

For a  $C^1$  function  $F = (F^{(x)}, F^{(y)}) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  we write  $\text{div}(F) = \partial_x F^{(x)} + \partial_y F^{(y)} : \mathbb{R}^2 \rightarrow \mathbb{R}$  for the divergence of  $F$ . Remark that for  $(x, t)$  lying in the interior of either  $R^+$  or  $R^-$ ,

$$\begin{aligned} \text{div} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) &= \partial_x \left( q(u^2 - c^2)\phi \right) + \partial_t \left( (u - c)\phi \right) \\ &= \partial_x (q \cdot u^2)\phi + q(u^2 - c^2) \cdot \partial_x \phi + \partial_t u \cdot \phi + (u - c) \cdot \partial_t \phi \\ &= q(u^2 - c^2) \cdot \partial_x \phi + (u - c) \cdot \partial_t \phi. \end{aligned}$$

We can therefore rewrite the left hand side of (39) as

$$\begin{aligned} &\int_{R^+} \text{div} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) - \int_{R^-} \text{div} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \\ &\quad + \int_{\mathbb{R}} |u_0 - c|\phi(x, 0)dx, \end{aligned}$$

and by applying the divergence theorem, this can in turn be written as

$$\begin{aligned} &\int_{\partial R^+} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^+} - \int_{\partial R^-} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^-} \\ &\quad + \int_{\mathbb{R}} |u_0 - c|\phi(x, 0)dx \\ &\geq \int_{T^+} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^+} + \int_L \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^+} \\ &\quad - \left( \int_{T^-} \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^-} + \int_L \left( q(u^2 - c^2)\phi, (u - c)\phi \right) \cdot n_{R^-} \right), \end{aligned} \tag{40}$$

where

$$L := \left\{ (x, t) : x = 2qc(t + \varepsilon), 0 \leq t \leq \frac{4q}{c^2} - \varepsilon \right\},$$

$$T^+ := \left\{ (x, t) : x = \sqrt{4q(t + \varepsilon)}, 0 \leq t \leq \frac{1}{c^2q} - \varepsilon \right\},$$

and

$$T^- := \left\{ (x, t) : x = \sqrt{4q(t + \varepsilon)}, \frac{1}{c^2q} - \varepsilon < t \leq T \right\}.$$

Note that  $u - c \equiv 0$  on  $L$ , and therefore we can rewrite (40) as

$$\int_{T^+} (q(u^2 - c^2)\phi, (u - c)\phi) \cdot n_{R^+} - \int_{T^-} (q(u^2 - c^2)\phi, (u - c)\phi) \cdot n_{R^-}.$$

On  $T^+$  and  $T^-$  we have  $n_{R^+} = n_{R^-}$ , where  $n_{R^+} = \left( \left| \frac{(t+\varepsilon)}{t+\varepsilon+q} \right|^{\frac{1}{2}}, - \left| \frac{q}{t+\varepsilon+q} \right|^{\frac{1}{2}} \right)$ . This yields

$$\begin{aligned} & \int_{T^+} (q(u^2 - c^2)\phi, (u - c)\phi) \cdot n_{R^+} - \int_{T^-} (q(u^2 - c^2)\phi, (u - c)\phi) \cdot n_{R^-} \\ &= \int_{T^+} \frac{c \cdot \phi}{|t + \varepsilon + q|^{1/2}} (\sqrt{q} - qc\sqrt{t + \varepsilon}) - \int_{T^-} \frac{c \cdot \phi}{|t + \varepsilon + q|^{1/2}} (\sqrt{q} - qc\sqrt{t + \varepsilon}) \\ &\geq \int_{T^+} \frac{c \cdot \phi}{|t + \varepsilon + q|^{1/2}} \left( \sqrt{q} - q\sqrt{\frac{1}{q}} \right) - \int_{T^-} \frac{c \cdot \phi}{|t + \varepsilon + q|^{1/2}} \left( \sqrt{q} - q\sqrt{\frac{1}{q}} \right) \\ &= 0. \end{aligned}$$

The final inequality holds because  $c \cdot |4(t + t_0) + 4q|^{-\frac{1}{2}}\phi > 0$ , on  $T^+$  we have  $c\sqrt{t + \varepsilon} \leq \sqrt{\frac{1}{q}}$ , and on  $T^-$ ,  $c\sqrt{t + \varepsilon} \geq \sqrt{\frac{1}{q}}$ . From this result we conclude that inequality (39) holds in the case  $c \in \left( \sqrt{\frac{4q}{T+\varepsilon}}, \sqrt{\frac{4q}{\varepsilon}} \right]$ .

As inequality (39) is satisfied for all cases of  $c \in \mathbb{R}$ , we conclude that  $u_B$  is the BV entropy weak solution to (12) with  $f = f_B$ ,  $K = K_B$  and  $u_0 = u_B(\cdot, 0)$ .  $\square$

**Proof** (Proof of Proposition 5) Again, we begin by verifying monotonicity. Since  $f_P \equiv 0$ , the function  $S : \mathbb{R}^3 \rightarrow \mathbb{R}$  defined by (16) is

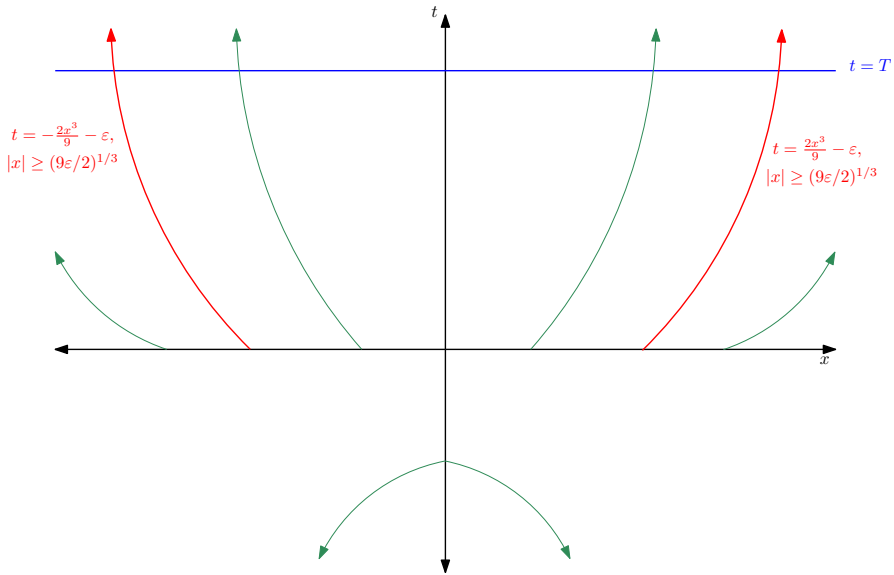
$$\begin{aligned} S(v^-, v, v^+) &= \frac{v}{\Delta_t^M} + \frac{1}{2(\Delta_x^M)^2} ((v^+)^2 - 2v^2 + (v^-)^2) \\ &= M^3 v - \frac{v^2}{2} M^2 + \frac{M^2}{4} ((v^+)^2 + (v^-)^2). \end{aligned}$$

The function  $S$  is non-decreasing in all of its arguments on  $[0, M]$ , so for sufficiently large  $M$  the approximation scheme  $(U_j^n(v_0, \Delta_x^M, \Delta_t^M))_{n \in \mathbb{N}, j \in \mathbb{Z}}$  is monotone on  $[0, (3/4)(2/(9\varepsilon))^{1/3}]$ .

We now turn to the first claim of the proposition. Similarly to in the proof of (4) it is clear that the functions,  $v \equiv v_P$ , and  $f(v_0) - \partial_x K(v_0)$  with  $v(\cdot, 0) \equiv v_0$  are of bounded variation. Therefore, it remains to show that

$$\int_{\mathbb{R} \times [0, T]} \operatorname{sgn}(v - c) \cdot ((v - c)\partial_t \phi - vv_x \partial_x \phi) dt dx + \int_{\mathbb{R}} |v_0 - c| dx \geq 0, \quad (41)$$

for all  $c \in \mathbb{R}$  and all non-negative  $\phi \in C^\infty(\mathbb{R} \times [0, T])$  with compact support such that  $\phi|_{t=T} \equiv 0$ . The proof splits into three cases:  $c \leq 0$ ,  $c \in (0, A(T + \varepsilon)^{-\frac{1}{3}})$ , and  $c \geq A(T + \varepsilon)^{-\frac{1}{3}}$  where  $A = (\frac{9}{2})^{\frac{2}{3}} \cdot \frac{1}{6}$ ; see Fig. 3.



**Fig. 3** The region of integration and different “sign-change” regimes for different values of  $c$ . The thinner, green curves represent equations of the form  $x = (((9(t + \epsilon))/2)^{2/3} - 6c(t + \epsilon))^{1/2}$ , where for the outermost curves in the upper quadrants  $c < 0$ , for the innermost curves in the upper quadrants  $c \in (0, A(T + \epsilon)^{-1/3})$ , and for the curve in the lower quadrants  $c \geq A(T + \epsilon)^{-1/3}$

In this setting the most involved case under consideration is when  $c \in (0, A/(T + \epsilon)^{-1/3}]$ . We provide a full proof for this case, with the other cases following by similar arguments.

Let  $\phi \in C^\infty(\mathbb{R} \times [0, T])$  be a non-negative function with compact support such that  $\phi|_{t=T} \equiv 0$  and define the regions

$$R^+ := \left\{ (x, t) : |x| \leq \left( \left( \frac{9(t + \epsilon)}{2} \right)^{2/3} - 6c(t + \epsilon) \right)^{1/2}, 0 \leq t \leq T \right\},$$

$$R^- := \left\{ (x, t) : \left( \left( \frac{9(t + \epsilon)}{2} \right)^{2/3} - 6c(t + \epsilon) \right)^{1/2} \leq |x| \leq \left( \frac{9(t + \epsilon)}{2} \right)^{1/3}, 0 \leq t \leq T \right\},$$

and

$$R_0 := \left\{ (x, t) : |x| \geq \left( \frac{9(t + \epsilon)}{2} \right)^{1/3}, 0 \leq t \leq T \right\}.$$

Observe for  $(x, t)$  in the interior of any of  $R^+, R^-$  or  $R_0$ .

$$\begin{aligned} \operatorname{div}(-vv_x\phi, (v - c)\phi) &= \partial_x(-vv_x\phi) + \partial_t((v - c)\phi) \\ &= -(v_x)^2 \cdot \phi - vv_{xx} \cdot \phi - vv_x \partial_x \phi + v_t \cdot \phi + (v - c)\partial_t \phi \end{aligned}$$

$$\begin{aligned}
 &= - \left( \partial_{xx} \left( \frac{v^2}{2} \right) \right) \cdot \phi - vv_x \partial_x \phi + v_t \cdot \phi + (v - c) \partial_t \phi \\
 &= -vv_x \partial_x \phi + (v - c) \partial_t \phi.
 \end{aligned}$$

We can therefore rewrite the left hand side of (41) as

$$\begin{aligned}
 &\int_{\mathbb{R}^+} \operatorname{div} (-vv_x \phi, (v - c)\phi) - \int_{\mathbb{R}^-} \operatorname{div} (-vv_x \phi, (v - c)\phi) \\
 &\quad - \int_{\mathbb{R}_0} \operatorname{div} (-vv_x \phi, (v - c)\phi) + \int_{\mathbb{R}} |v_0 - c| \phi(x, 0) dx.
 \end{aligned}$$

Applying the divergence theorem, this can in turn be written as

$$\begin{aligned}
 &\int_{\partial \mathbb{R}^+} (-vv_x \phi, (v - c)\phi) \cdot n_{\mathbb{R}^+} - \int_{\partial \mathbb{R}^-} (-vv_x \phi, (v - c)\phi) \cdot n_{\mathbb{R}^-} \\
 &\quad - \int_{\partial \mathbb{R}_0} (-vv_x \phi, (v - c)\phi) \cdot n_{\mathbb{R}_0} + \int_{\mathbb{R}} |v_0 - c| \phi(x, 0) dx, \tag{42}
 \end{aligned}$$

where  $n_{\mathbb{R}^+}$ ,  $n_{\mathbb{R}^-}$  and  $n_{\mathbb{R}_0}$  are the outward normal vectors on  $\mathbb{R}^+$ ,  $\mathbb{R}^-$ , and  $\mathbb{R}_0$  respectively. Now, let

$$\begin{aligned}
 L^+ &:= \left\{ (x, t) : x = \left( \left( \frac{9(t + \varepsilon)}{2} \right)^{2/3} - 6c(t + \varepsilon) \right)^{1/2}, 0 \leq t \leq T \right\}, \\
 L^- &:= \left\{ (x, t) : x = - \left( \left( \frac{9(t + \varepsilon)}{2} \right)^{2/3} - 6c(t + \varepsilon) \right)^{1/2}, 0 \leq t \leq T \right\}, \\
 T^+ &:= \left\{ (x, t) : x = \left( \frac{9(t + \varepsilon)}{2} \right)^{1/3}, 0 \leq t \leq T \right\}, \\
 T^- &:= \left\{ (x, t) : x = - \left( \frac{9(t + \varepsilon)}{2} \right)^{1/3}, 0 \leq t \leq T \right\}.
 \end{aligned}$$

Note that on the regions  $L^+$  and  $L^-$  we have that  $v - c \equiv 0$ , and on the regions  $T^+$  and  $T^-$ ,  $v \equiv 0$ . We can then bound (42) from below by

$$\begin{aligned}
 &\left( \int_{L^+} (-vv_x \phi, 0) \cdot n_{\mathbb{R}^+} + \int_{L^-} (-vv_x \phi, 0) \cdot n_{\mathbb{R}^+} \right) \\
 &\quad - \left( \int_{L^+} (-vv_x \phi, 0) \cdot n_{\mathbb{R}^-} + \int_{L^-} (-vv_x \phi, 0) \cdot n_{\mathbb{R}^-} + \int_{T^+} (0, -c\phi) \cdot n_{\mathbb{R}^-} \right. \\
 &\quad \left. + \int_{T^-} (0, -c\phi) \cdot n_{\mathbb{R}^-} \right) \\
 &\quad - \left( \int_{T^+} (0, -c\phi) \cdot n_{\mathbb{R}_0} + \int_{T^-} (0, -c\phi) \cdot n_{\mathbb{R}_0} \right)
 \end{aligned}$$



$$\begin{aligned}
 &= \left( \int_{L^+} (-vv_x\phi, 0) \cdot n_{R^+} - \int_{L^+} (-vv_x\phi, 0) \cdot n_{R^-} \right) \\
 &\quad + \left( \int_{L^-} (-vv_x\phi, 0) \cdot n_{R^+} - \int_{L^-} (-vv_x\phi, 0) \cdot n_{R^-} \right) \\
 &\quad + \left( \int_{T^+} (0, -c\phi) \cdot n_{R^-} - \int_{T^+} (0, -c\phi) \cdot n_{R_0} \right) \\
 &\quad + \left( \int_{T^-} (0, -c\phi) \cdot n_{R^-} - \int_{T^-} (0, -c\phi) \cdot n_{R_0} \right) \\
 &= 2 \cdot \int_{L^+} (-vv_x\phi, 0) \cdot n_{R^+} + 2 \cdot \int_{L^-} (-vv_x\phi, 0) \cdot n_{R^+}.
 \end{aligned}$$

The last equality follows from the fact that

$$n_{R^-} = n_{R^+} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ on } L^+, L^-$$

and

$$n_{R_0} = n_{R^-} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ on } T^+, T^-.$$

Direct calculation gives that  $n_{R^+} = (B^{-1}, DB^{-1})$  on  $L^+$ , and  $n_{R^+} = (-B^{-1}, DB^{-1})$  on  $L^-$ , where

$$B = \left( \frac{9}{4} \left( \left( \frac{2}{9(t+\varepsilon)} \right)^{1/3} - 2c \right)^2 \cdot \left( \left( \frac{9(t+\varepsilon)}{2} \right)^{2/3} - 6c(t+\varepsilon) \right)^{-1} \right)^{1/2}$$

and

$$D = \frac{3}{2} \left( \left( \frac{2}{9(t+\varepsilon)} \right)^{1/3} - 6c \right) \cdot \left( \left( \frac{9(t+\varepsilon)}{2} \right)^{2/3} - 6c(t+\varepsilon) \right)^{-1/2}.$$

We thus have  $(-vv_x\phi, 0) \cdot n_{R^+} = -vv_x\phi B^{-1} \geq 0$  on  $L^+$ , and  $(-vv_x\phi, 0) \cdot n_{R^+} = vv_x\phi B^{-1} \geq 0$  on  $L^-$ . Combining the above results, it follows that for all non-negative  $\phi \in C^\infty(\mathbb{R} \times [0, T])$  with compact support such that  $\phi|_{t=T} \equiv 0$ , and for all  $c \in (0, 4q(T + \varepsilon)^{-\frac{1}{3}})$ ,

$$\int_{\mathbb{R} \times [0, T]} \text{sgn}(v - c) \cdot ((v - c)\partial_t \phi - vv_x \partial_x \phi) dt dx + \int_{\mathbb{R}} |v_0 - c| dx \geq 0.$$

The same inequality follows for the other ranges of  $c$ , and by similar arguments, and we can therefore conclude that  $v_p$  is the BV entropy weak solution to (12) with  $f = f_p$ ,  $K = K_p$ , and  $v_0 = v_p(\cdot, 0)$ . □

**Proof** (Proof of Lemma 11.) By induction, it suffices to prove the lemma when  $k = 1$ , and we now restrict our attention to this setting.

Let  $(A, B)$  and  $(C, D)$  be independent of one another, with each pair distributed according to the coupling  $(X, Y)$ . Then  $A$  and  $C$  are independent and  $\mu$ -distributed, and  $B$  and  $D$  are independent and  $\nu$ -distributed. We shall couple the symmetric simple hipster random walk dynamics with input  $(A, C)$  to those with input  $(B, D)$ , via a case-by-case construction of the coupling dynamics.

Define the events  $E_1 = \{A \leq B, C \leq D\}$ ,  $E_2 = \{A \leq B, C > D\}$ ,  $E_3 = \{A > B, C \leq D\}$ , and  $E_4 = \{A > B, C > D\}$ . For  $i \in \{1, 2, 3, 4\}$  we consider the following sub-cases:

$$\begin{aligned} & (i) E_i \cap \{A = C\} \cap \{B \neq D\} \quad (ii) E_i \cap \{A \neq C\} \cap \{B \neq D\} \\ & (iii) E_i \cap \{A \neq C\} \cap \{B = D\}, \text{ and } (iv) E_i \cap \{A = C\} \cap \{B = D\}. \end{aligned}$$

We begin by constructing the coupling for  $E_1$ . In case (i), we construct the coupling  $(X', Y')$  as

$$\begin{cases} Y' = \max(B, D) \iff X' = A + 1 \\ Y' = \min(B, D) \iff X' = A - 1, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_1 \cap \{A = C\} \cap \{B \neq D\}\} = 0$ .

For case (ii), we construct the coupling as

$$\begin{cases} Y' = D \iff X' = C \\ Y' = B \iff X' = A, \end{cases}$$

giving that  $\mathbf{P}\{X' > Y' \mid E_1 \cap \{A \neq C\} \cap \{B \neq D\}\} = 0$ .

Further, for case (iii) the coupling is

$$\begin{cases} Y' = D + 1 \iff X' = \max(A, C) \\ Y' = D - 1 \iff X' = \min(A, C), \end{cases}$$

which again gives  $\mathbf{P}\{X' > Y' \mid E_1 \cap \{A \neq C\} \cap \{B = D\}\} = 0$ .

Lastly, for case (iv), the coupling is

$$\begin{cases} Y' = D + 1 \iff X' = A + 1 \\ Y' = D - 1 \iff X' = A - 1, \end{cases}$$

giving that  $\mathbf{P}\{X' > Y' \mid E_1 \cap \{A = C\} \cap \{B = D\}\} = 0$ .

Combining cases (i) through (iv) for  $i = 1$ , we obtain that

$$\mathbf{P}\{X' > Y' \mid E_1\} = 0. \tag{43}$$

Note that for  $i \in \{2, 3\}$ , the event  $E_i \cap \{A = C\} \cap \{B = D\}$  is empty, so it suffices to study sub-cases (i) through (iii).

We next construct the coupling on  $E_2$ . For case (i), we note that  $E_2 \cap \{A = C\} \cap \{B \neq D\} = \{D < A = C \leq B\}$ . In this case the coupling is

$$\begin{cases} Y' = D \iff X' = A + 1 \\ Y' = B \iff X' = A - 1, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_2 \cap \{A = C\} \cap \{B \neq D\}\} = 1/2$ . For case (ii) the coupling is

$$\begin{cases} Y' = D \iff X' = C \\ Y' = B \iff X' = A, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_2 \cap \{A \neq C\} \cap \{B \neq D\}\} = \frac{1}{2}$ . Next, for case (iii),  $E_2 \cap \{A \neq C\} \cap \{B = D\} = \{A \leq B = D < C\}$ , and the coupling is

$$\begin{cases} Y' = D + 1 \iff X' = A \\ Y' = D - 1 \iff X' = C, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_2 \cap \{A \neq C\} \cap \{B = D\}\} = 1/2$ . Combining cases (i) through (iii) with  $i = 2$ , we obtain that

$$\mathbf{P}\{X' > Y' \mid E_2\} = \frac{1}{2}. \tag{44}$$

We now construct the coupling on the event  $E_3$ . For case (i),  $E_3 \cap \{A = C\} \cap \{B \neq D\} = \{B < A = C \leq D\}$ . The coupling is

$$\begin{cases} Y' = D \iff X' = A - 1 \\ Y' = B \iff X' = A + 1, \end{cases}$$

giving  $\mathbf{P}\{X' > Y' \mid E_3 \cap \{A = C\} \cap \{B \neq D\}\} = 1/2$ . For case (ii), the coupling is

$$\begin{cases} Y' = D \iff X' = C \\ Y' = B \iff X' = A, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_3 \cap \{A \neq C\} \cap \{B \neq D\}\} = \frac{1}{2}$ .

Lastly, for case (iii),  $E_3 \cap \{A \neq C\} \cap \{B = D\} = \{A > B = D \geq C\}$ . The coupling is

$$\begin{cases} Y' = D + 1 \iff X' = C \\ Y' = D - 1 \iff X' = A, \end{cases}$$

which gives  $\mathbf{P}\{X' > Y' \mid E_3 \cap \{A \neq C\} \cap \{B = D\}\} = 1/2$ .

Combining cases (i) through (iii) with  $i = 3$ , we obtain that

$$\mathbf{P}\{X' > Y' \mid E_3\} = \frac{1}{2}. \tag{45}$$

Finally, we construct the coupling on the event  $E_4$  arbitrarily (for example, by making independent choices for the two processes). This gives

$$\mathbf{P}\{X' > Y' \mid E_4\} \leq 1. \tag{46}$$

Since both  $(A, B)$  and  $(C, D)$  are distributed according to the coupling  $(X, Y)$ , we have  $\mathbf{P}(A > B) = \alpha = \mathbf{P}(C > D)$ . Since  $(A, B)$  and  $(C, D)$  are independent, we also have that

$$\mathbf{P}\{E_2\} = \mathbf{P}\{A \leq B, C > D\} = \mathbf{P}\{A > B, C \leq D\} = \mathbf{P}\{E_3\} = \alpha(1 - \alpha),$$

and  $\mathbf{P}\{E_4\} = \mathbf{P}\{A > B, C > D\} = \alpha^2$ . Combining this with (43), (44), (45) and (46), and using the law of total probability, we obtain that

$$\begin{aligned} \mathbf{P}\{X' > Y'\} &= \sum_{i=1}^4 \mathbf{P}\{X' > Y' \mid E_i\} \cdot \mathbf{P}\{E_i\} \\ &\leq \alpha(1 - \alpha) + \alpha^2 = \alpha, \end{aligned}$$

as required. □

**Proof** (Proof of Lemma 10) As noted above, the construction of a coupling with the claimed property is essentially identical to the construction from the proof of Lemma 11. To obtain it from that construction, simply replace all instances of  $A - 1$  by  $A$  (in cases  $E_1(i), E_1(iv), E_2(i)$  and  $E_3(i)$ ) and all instances of  $D - 1$  by  $D$  (in cases  $E_1(iii), E_1(iv), E_2(iii)$  and  $E_3(iii)$ ). We leave the detailed verification to the reader. □

## References

1. Auffinger, A., Cable, D.: Pemantle’s Min-Plus Binary Tree (2017). [arXiv:1709.07849](https://arxiv.org/abs/1709.07849) [math.PR]
2. Evje, S., Karlsen, K.H.: Monotone difference approximations of BV solutions to degenerate convection–diffusion equations. *SIAM J. Numer. Anal.* **37**(6), 1838–1860 (2000). <https://doi.org/10.1137/S0036142998336138>

3. Godlewski, E., Raviart, P.A.: Numerical approximation of hyperbolic systems of conservation laws. In: Marsden, J.E., Sirovich, L. (eds.) *Applied Mathematical Sciences*, vol. 118. Springer, New York (1996). <https://doi.org/10.1007/978-1-4612-0713-9>
4. Hambly, B.M., Jordan, J.: A random hierarchical lattice: the series–parallel graph and its properties. *Adv. Appl. Probab.* **36**(3), 824–838 (2004). <https://doi.org/10.1239/aap/1093962236>
5. Harten, A., Hyman, J.M., Lax, P.D.: On finite-difference approximations and entropy conditions for shocks. *Commun. Pure Appl. Math.* **29**(3), 297–322 (1976). <https://doi.org/10.1002/cpa.3160290305>. (with an appendix by B. Keyfitz)
6. Kruřkov, S.N.: First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)* **81**(123), 228–255 (1970)
7. Lax, P.D.: Hyperbolic systems of conservation laws and the mathematical theory of shock waves. In: *Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics*, No. 11. Society for Industrial and Applied Mathematics, Philadelphia (1973)
8. Pemantle, R.: *Columbia–Courant Joint Probability Seminar Series, 2016-10-14* (2016). <http://web.archive.org/web/20161029170321/>, <http://www.math.nyu.edu:80/seminars/probability/ColumbiaCourant2016.html>
9. Vazquez, J.L.: *The Porous Medium Equation*, vol. 14. Oxford University Press, Oxford (2006). <https://doi.org/10.1093/acprof:oso/9780198569039.001.0001>
10. Volpert, A.I.: Spaces BV and quasilinear equations. *Mat. Sb. (N.S.)* **73**(115), 255–302 (1967)
11. Volpert, A.I.: Generalized solutions of degenerate second-order quasilinear parabolic and elliptic equations. *Adv. Differ. Equ.* **5**(10–12), 1493–1518 (2000)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.