Balazs, M., & Seppalainen, T. (2007). A convexity property of expectations under exponential weights. arXiv .

A convexity property of expectations under exponential weights

M. Balázs* T. Seppäläinen†

balazs@math.bme.hu seppalai@math.wisc.edu

December 18, 2013

Abstract

Please note: after completion of this manuscript we learned that our main results, Theorem 2.1 and 2.2, can be obtained as a special case of Proposition 3.2 on Page 23 of Karlin's book [6].

Take a random variable X with distribution μ with some finite exponential moments, and weight it by an exponential factor $e^{\theta X}$ to get the distribution μ^θ for the admissible θ-values. Define also the so-weighted expectation $\varrho(\theta) := \mathbf{E}^\theta X$ with inverse function $\theta(\varrho)$. This note proves that for a convex function Φ, $\mathbf{E}^{\theta(\varrho)}\Phi(X)$ is a convex function of ϱ, wherever it exists and is finite. Along the way we develop correlation inequalities for convex functions. Motivation for this result comes from equilibrium investigations of some stochastic interacting systems with stationary product distributions. In particular, convexity of the hydrodynamic flux function follows in some cases.

Keywords: Exponential weights, Gibbs measures, convexity, correlation inequalities, particle flux, zero range process, bricklayer process

2000 Mathematics Subject Classification: 60K35, 60E15

1 Introduction

Please note: after completion of this manuscript we learned that our main results, Theorem 2.1 and 2.2, can be obtained as a special case of Proposition 3.2 on Page 23 of Karlin's book [6].

*Budapest University of Technology and Economics, Institute of Mathematics, 1 Egry József u., H ép. V.7., Budapest, Hungary
M. Balázs was partially supported by the Hungarian Scientific Research Fund (OTKA) grants K60708, TS49835, F67729, and the Bolyai Scholarship of the Hungarian Academy of Sciences.
†University of Wisconsin-Madison, Mathematics Department, Van Vleck Hall, 480 Lincoln Dr, Madison WI 53706-1388, USA.
T. Seppäläinen was partially supported by National Science Foundation grant DMS-0402231.

Take a non-degenerate random variable X such that $\mathbf{E}e^{\theta X} < \infty$ for $\theta \in I$, for some open interval I. For these θ define the exponentially weighted distribution $\mathbf{E}^\theta(Y) := \mathbf{E}^\theta(Y e^{\theta X})/\mathbf{E}e^{\theta X}$. The exponentially weighted expectation of X is $\varrho(\theta) := \mathbf{E}^\theta X$. This function is strictly increasing due to the nondegeneracy assumption. We denote its inverse by $\theta(\varrho)$.

Let Φ be a convex function for which $\mathbf{E}^{\theta(\varrho)}\Phi(X)$ exists in an open interval of ϱ-values. The first result of this note is the convexity of the function

$$\varrho \mapsto \mathbf{E}^{\theta(\varrho)}\Phi(X).$$

Motivation for this result comes from a class of asymmetric stochastic interacting systems that includes the zero range process (ZRP) and the bricklayer process (BLP). We explain these informally before turning to precise statements.

The main application is related to the study of fluctuations of the current of particles $J^{(V)}(t)$ as seen by an observer moving at a fixed speed V. Equivalently, these are fluctuations of the height in the deposition formulation of the process. A key fact that underlies some of this work is that the variance of the current of a stationary process is linked to the deviations of a second class particle:

$$\mathbf{Var}(J^{(V)}(t)) = C(\varrho)\widehat{\mathbf{E}}|Q(t) - [Vt]|. \tag{1.1}$$

The variance on the left is taken in the stationary process at a fixed density ϱ. The expectation on the right is taken with an altered initial product distribution: the density ϱ invariant factor is put at each site other than the origin, while at the origin there is a different measure we denote by $\nu^{\theta(\varrho)}$, defined in (3.1) below. $Q(t)$ is the position of a second class particle in the system. $[Vt]$ is not the usual integer part but rather the integer between Vt and the origin that is closest to Vt.

Identity (1.1) has been known for the totally asymmetric simple exclusion process since the pioneering work of Ferrari and Fontes [5]. It was recently extended to the broader class of processes in [4]. For a complete discussion we refer the reader to [4], where the measures $\nu^{\theta(\varrho)}$ are defined in equation (2.6) and denoted by $\widehat{\mu}_\theta$.

The presently relevant point is that for coupling purposes it is important that the measures $\nu^{\theta(\varrho)}$ are stochastically monotone in the parameter ϱ. In the case of the asymmetric simple exclusion process (ASEP) this is immediately obvious. This fact was utilized in a recent coupling-intensive proof [3] that established the order $t^{2/3}$ for the variance in (1.1) for ASEP when the observer travels at the characteristic speed. The first step in a program to extend the variance bounds to ZRP and BLP is to develop the coupling framework. However, the stochastic monotonicity of the measures $\nu^{\theta(\varrho)}$ is not at all obvious for ZRP and BLP. This we derive from the main convexity result in Section 3.2.

The conserved quantity in these asymmetric processes (typically viewed as particle counts, but also discrete gradients of the interface height) satisfies a hydrodynamic scaling limit where the limiting evolution is the entropy solution of a scalar conservation law of the form

$$\partial_t\varrho + c\partial_x\mathcal{H}(\varrho) = 0.$$

(We refer the reader to [7] for general theory.) The key quantity is the flux $\mathcal{H}(\varrho)$ which is computed as the expected jump rate in the stationary process at particle density ϱ. The constant c is the mean increment of a particle that has decided to jump.

A second application of the main convexity result is to give an alternative proof of the convexity of the hydrodynamic flux function $\mathcal{H}$ for zero range and bricklayer processes when the jump rate is convex. (See [1] for the original proof.) Concavity of the hydrodynamic flux also follows for concave jump rates in the zero range process. Strict convexity or concavity are also discussed.

The characteristic speed referred to above is $V^{\varrho} = c\mathcal{H}'(\varrho)$. So the issue can also be framed as the monotonicity of this quantity in the particle density.

The rest of this note is organized as follows. We rewrite the convexity problem in terms of correlation inequalities of functions of X. These we handle via separation of positive and negative parts and further correlation inequalities. This is done in Section 2. In Section 3 we derive the consequences for stochastic interacting systems.

2 Derivatives and correlations

As in the introduction, let X be a nondegenerate real-valued random variable and Φ a convex function on some interval that contains the range of X. The standing assumption throughout this section is that for some open interval $I \subseteq \mathbb{R}$,

$$\mathbf{E}(\mathrm{e}^{\theta X}) < \infty \quad \text{and} \quad \mathbf{E}(X^2|\Phi(X)|\mathrm{e}^{\theta X}) < \infty \tag{2.1}$$

for all $\theta \in I$. Define the exponentially weighted distribution as $\mathbf{E}^{\theta}(Y) = \mathbf{E}^{\theta}(Y\mathrm{e}^{\theta X})/\mathbf{E}\mathrm{e}^{\theta X}$. The function $\varrho(\theta) = \mathbf{E}^{\theta}X$ is strictly increasing (justification below in Corollary 2.4). It has an inverse function $\theta(\varrho)$ defined in some nontrivial open interval J. The expectation $\mathbf{E}^{\theta(\varrho)}\Phi(X)$ is well defined for $\varrho \in J$.

Theorem 2.1. *The function $\varrho \mapsto \mathbf{E}^{\theta(\varrho)}\Phi(X)$ is convex on J.*

Our main interest lies in discrete distributions so we state a further condition for strict convexity for that case. Suppose the distribution μ of X is supported on a discrete subset $\mathbb{S}$ of $\mathbb{R}$. So $\mathbb{S}$ is either finite or countably infinite but locally finite. Let Φ be a function defined on $\mathbb{S}$. Extend Φ to a function on the smallest closed interval that contains $\mathbb{S}$ by connecting adjacent points on the graph of Φ with line segments. Assume the function Φ thus defined is convex. Say Φ is *strictly convex* at a point $z \in \mathbb{S}$ if the slope of the extended Φ jumps at z. Such a point z cannot be the maximum or minimum of $\mathbb{S}$ because we have not defined Φ outside the smallest interval that contains $\mathbb{S}$. If no such point exists then Φ is linear.

Theorem 2.2. *Consider the discrete case described in the paragraph above. The function $\varrho \mapsto \mathbf{E}^{\theta(\varrho)}\Phi(X)$ is strictly convex throughout its interval of definition if and only if $\mathbb{S}$ contains at least three points and Φ is strictly convex at some point of $\mathbb{S}$. In the complementary case the function $\varrho \mapsto \mathbf{E}^{\theta(\varrho)}\Phi(X)$ is linear.*

The remainder of this section covers the proofs. Throughout we only consider values $\theta \in I$ for which assumption (2.1) guarantees that the derivatives and other operations we perform are justified. In particular, since $|X|^k \le k!\varepsilon^{-k}(\mathrm{e}^{\varepsilon X} + \mathrm{e}^{-\varepsilon X})$, X has all moments under $\mathbf{E}^\theta$ for each $\theta \in I$ because $\theta \pm \varepsilon \in I$ for small enough $\varepsilon > 0$.

We start with a preliminary lemma, repeated from Lemma A2 of [2].

Lemma 2.3. *For any function φ, we have*

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbf{E}^\theta\varphi(X) = \mathbf{Cov}^\theta(\varphi(X),\, X)$$

provided the expectations exist in a neighborhood of θ.

Proof.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbf{E}^\theta\varphi(X) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\,\frac{\mathbf{E}(\varphi(X)\cdot\mathrm{e}^{\theta X})}{\mathbf{E}\mathrm{e}^{\theta X}}\\ &= \frac{\mathbf{E}(\varphi(X)\cdot X\cdot\mathrm{e}^{\theta X})}{\mathbf{E}\mathrm{e}^{\theta X}} - \mathbf{E}(\varphi(X)\cdot\mathrm{e}^{\theta X})\cdot\frac{\mathbf{E}(X\cdot\mathrm{e}^{\theta X})}{[\mathbf{E}\mathrm{e}^{\theta X}]^2}\\ &= \mathbf{Cov}^\theta(\varphi(X),\, X).\end{aligned}$$

□

Recall that we exclude the degenerate case where μ is supported on a single point.

Corollary 2.4.

$$\frac{\mathrm{d}\varrho(\theta)}{\mathrm{d}\theta} = \frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbf{E}^\theta X = \mathbf{Cov}^\theta(X,\, X) = \mathbf{Var}^\theta X > 0, \qquad \textit{and}$$

$$\frac{\mathrm{d}\mathbf{E}^{\theta(\varrho)}\varphi(X)}{\mathrm{d}\varrho} = \frac{\mathrm{d}\mathbf{E}^\theta\varphi(X)}{\mathrm{d}\theta}\cdot\frac{\mathrm{d}\theta(\varrho)}{\mathrm{d}\varrho} = \frac{\mathbf{Cov}^{\theta(\varrho)}(\varphi(X),\, X)}{\mathbf{Var}^{\theta(\varrho)}X}. \tag{2.2}$$

Now we proceed by rewriting the second derivative of $\mathbf{E}^{\theta(\varrho)}\Phi(X)$ in terms of covariances. We omit the notation (X) from $\Phi(X)$.

Lemma 2.5. *The following are equivalent:*

a) For any convex function Φ, $\mathbf{E}^{\theta(\varrho)}\Phi$ is a convex function of ϱ.

b) For any convex function Φ,

$$\mathbf{Cov}^\theta(\widetilde{\Phi}\cdot X,\, X)\cdot\mathbf{Cov}^\theta(X,\, X) \ge \mathbf{Cov}^\theta(\Phi,\, X)\cdot\mathbf{Cov}^\theta(\widetilde{X}\cdot X,\, X), \tag{2.3}$$

where $\widetilde{\cdot}$ stands for centering w.r.t. $\mathbf{E}^\theta$.

Proof. We write, as in the corollary,

$$\frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\theta(\varrho)}\Phi = \frac{\mathbf{Cov}^{\theta(\varrho)}(\Phi,\, X)}{\mathbf{Var}^{\theta(\varrho)}X}.$$

We need to see if this is nondecreasing in ϱ or, equivalently, nondecreasing in θ. That happens if and only if

$$\begin{aligned}
0 &\le \mathbf{Var}^\theta X \cdot \frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{Cov}^\theta(\Phi,\, X) - \mathbf{Cov}^\theta(\Phi,\, X)\cdot \frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{Var}^\theta X \\
&= \mathbf{Var}^\theta X \cdot \big[\mathbf{Cov}^\theta(\Phi X,\, X) - \mathbf{Cov}^\theta(\Phi,\, X)\cdot \mathbf{E}^\theta(X) - \mathbf{E}^\theta(\Phi)\cdot \mathbf{Cov}^\theta(X,\, X)\big] \\
&\quad - \mathbf{Cov}^\theta(\Phi,\, X)\cdot[\mathbf{Cov}^\theta(X^2,\, X) - 2\mathbf{E}^\theta X\cdot \mathbf{Cov}^\theta(X,\, X)] \\
&= \mathbf{Cov}^\theta(\widetilde{\Phi}X,\, X)\cdot \mathbf{Cov}^\theta(X,\, X) - \mathbf{Cov}^\theta(\Phi,\, X)\cdot \mathbf{Cov}^\theta(\widetilde{X}X,\, X). \quad \square
\end{aligned}$$

Next we concentrate on proving that part b) of the last lemma holds for *any distribution*. Therefore we omit the superscript θ.

Lemma 2.6. *Part b) of Lemma 2.5 is further equivalent to each of these two statements:*

c) For any convex function Φ that is uncorrelated with X, $\mathbf{Cov}(\Phi,\, X^2) \ge 0$.

d) For any convex function Φ,

$$\mathbf{Cov}(\Phi,\, X^2)\cdot \mathbf{Cov}(X,\, X) \ge \mathbf{Cov}(\Phi,\, X)\cdot \mathbf{Cov}(X^2,\, X). \tag{2.4}$$

Proof. Given a convex function Φ, let $\widehat{\Phi}(X) := \Phi(X) - C\cdot X$ with C chosen so that $\widehat{\Phi}$ is uncorrelated with X. $\widehat{\Phi}$ is also convex, and we note that (2.3) holds for Φ if and only if it holds for $\widehat{\Phi}$. Hence b) is equivalent to the statement obtained by restricting b) to convex functions that are uncorrelated with X. For such functions this statement becomes

$$\begin{aligned}
0 &\le \mathbf{Cov}(\widehat{\Phi}X,\, X) - \mathbf{E}\widehat{\Phi}\cdot \mathbf{Cov}(X,\, X) \\
&= \mathbf{Cov}(\widehat{\Phi},\, X^2) + \mathbf{E}\widehat{\Phi}\cdot \mathbf{E}(X^2) - \mathbf{E}(\widehat{\Phi}X)\cdot \mathbf{E}X - \mathbf{E}\widehat{\Phi}\cdot \mathbf{E}(X^2) + \mathbf{E}\widehat{\Phi}\cdot \mathbf{E}X\cdot \mathbf{E}X \\
&= \mathbf{Cov}(\widehat{\Phi},\, X^2).
\end{aligned}$$

Thus b) is equivalent to c).

Condition c) is a weakening of d), and we see that c) implies d) by determining the constant in the transformation that led to $\widehat{\Phi}$:

$$\mathbf{Cov}(\widehat{\Phi},\, X) = \mathbf{Cov}(\Phi,\, X) - C\cdot \mathbf{Cov}(X,\, X) = 0,$$

therefore

$$\widehat{\Phi} = \Phi - \frac{\mathbf{Cov}(\Phi,\, X)}{\mathbf{Cov}(X,\, X)}\cdot X.$$

Substituting this into $\mathbf{Cov}(\widehat{\Phi},\, X^2) \ge 0$ of c) leads to d). $\square$

Next we show that for part d) of the above lemma it suffices to consider the special case $\Phi(X) = |X|$.

Lemma 2.7. *Part d) of the above lemma is implied by this statement:*

e) For any distribution (with finite third absolute moments) we have

$$\mathbf{Cov}(|X|,\, X^2) \cdot \mathbf{Cov}(X,\, X) \geq \mathbf{Cov}(|X|,\, X) \cdot \mathbf{Cov}(X^2,\, X). \tag{2.5}$$

Proof. Consider functions of the form

$$\phi(x) = c + a \cdot [x - x_0]^+ - b \cdot [x - x_0]^- \tag{2.6}$$

for some $a > b$ and $x_0,\, c \in \mathbb{R}$. Notations $^+$ and $^-$ stand for positive and negative parts, respectively. These functions are convex. The first claim is that if (2.4) holds for functions of this special form, then it holds for any convex Φ.

This follows because Φ can be approximated from below in a pointwise fashion by a sequence of functions of this type:

$$g(x) = c - a_0[x - y_1]^- + a_1[x - y_1]^+ + \sum_{k=2}^{m}(a_k - a_{k-1})[x - y_k]^+$$

with $a_0 < a_1 < \cdots < a_m$ and $y_1 < \cdots < y_m$. The function g above is a sum of convex functions of type (2.6). To see the approximation, take points $z_0 < z_1 < \cdots < z_m$ and let a_i be the slope of a tangent to Φ at the point $(z_i, \Phi(z_i))$. Pick the z_i's so that the a_i's are strictly increasing. (This entails no loss of generality because a linear approximation to Φ is exact throughout any interval with constant slope.) Let g_i $(0 \leq i \leq m)$ be the linear function of slope a_i that passes through the point $(z_i, \Phi(z_i))$. Let y_i $(1 \leq i \leq m)$ be the x-coordinate of the point where the graphs of g_{i-1} and g_i intersect and set

$$c = \Phi(z_0) + a_0(y_1 - z_0) = \Phi(z_1) + a_1(y_1 - z_1).$$

Then it can be checked that g from above is the pointwise maximum of the g_i's, or equivalently, that $g = g_i$ on (y_i, y_{i+1}) with $y_0 = -\infty$ and $y_{m+1} = \infty$. By choosing the z_i's carefully one can create a sequence of convex functions $g^{(m)}$ such that $g^{(m)} \nearrow \Phi$ pointwise. By (2.1) monotone convergence applies to show $\mathbf{Cov}(g^{(m)},\, X^b) \to \mathbf{Cov}(\Phi,\, X^b)$ for $b = 1, 2$.

Thus we can derive (2.4) for Φ by checking it for each $g^{(m)}$. Since (2.4) is linear in Φ, it is then enough to know that it holds for each term of the type (2.6). This we now check.

With suitably chosen constants $A > 0$, B and C, the transformation

$$\phi(x) \mapsto A\phi(x) + Bx + C \tag{2.7}$$

turns ϕ of (2.6) into the function $|x - x_0|$. (Note that $a > b$ is needed for this.) The left and right-hand sides of (2.4) are, up to the multiplying factor A, invariant under these transformations. Hence (2.4) holds for ϕ if and only if it holds for $|x - x_0|$:

$$\mathbf{Cov}(|X - x_0|,\, X^2) \cdot \mathbf{Cov}(X,\, X) \ge \mathbf{Cov}(|X - x_0|,\, X) \cdot \mathbf{Cov}(X^2,\, X).$$

Introduce now $Y = X - x_0$, and write this inequality in the form

$$\mathbf{Cov}(|Y|,\, (Y + x_0)^2) \cdot \mathbf{Cov}(Y,\, Y) \ge \mathbf{Cov}(|Y|,\, Y) \cdot \mathbf{Cov}((Y + x_0)^2,\, Y).$$

Subtracting $2x_0 \cdot \mathbf{Cov}(|Y|,\, Y) \cdot \mathbf{Cov}(Y,\, Y)$ from both sides leads to e) (for the distribution of $Y = X - x_0$). □

Some elementary computations will now finish the proof of Theorem 2.1.

Lemma 2.8. *Part e) in Lemma 2.7 holds.*

Proof. For this proof, we introduce the positive and negative part moments:

$$P_i := \mathbf{E}\big((X^+)^i\big), \qquad N_i := \mathbf{E}\big((X^-)^i\big).$$

Expanding (2.5) gives

$$\begin{aligned}&\big[P_3 + N_3 - (P_1 + N_1) \cdot (P_2 + N_2)\big] \cdot \big[P_2 + N_2 - (P_1 - N_1)^2\big] \\ &\ge \big[P_2 - N_2 - (P_1 + N_1) \cdot (P_1 - N_1)\big] \cdot \big[P_3 - N_3 - (P_2 + N_2) \cdot (P_1 - N_1)\big].\end{aligned}$$

Somewhat tedious factoring shows that this is equivalent to

$$\begin{aligned}
&0 \le N_1 \cdot (P_3 P_1 - P_2 P_2) && (2.8)\\
&\quad + P_1 \cdot (N_3 N_1 - N_2 N_2) && (2.9)\\
&\quad + P_2 N_3 - P_1 P_1 N_3 - P_2 N_2 N_1 && (2.10)\\
&\quad + P_3 N_2 - P_3 N_1 N_1 - P_2 P_1 N_2. && (2.11)
\end{aligned}$$

We proceed by showing that each line above is non-negative. Clearly if $\mathbf{P}\{X > 0\}$ or $\mathbf{P}\{X \le 0\}$ is zero, then P_i's or N_i's are zero and the statement is trivially true. Assuming the contrary and dividing (2.8) by $[\mathbf{P}\{X > 0\}]^2$ makes conditional expectations out of the P_i's:

$$\frac{P_3 P_1 - P_2 P_2}{[\mathbf{P}\{X > 0\}]^2} = \mathbf{E}(X^3 \,|\, X > 0) \cdot \mathbf{E}(X \,|\, X > 0) - \mathbf{E}(X^2 \,|\, X > 0) \cdot \mathbf{E}(X^2 \,|\, X > 0).$$

To show that this is non-negative, introduce the expectation

$$\widehat{\mathbf{E}}(\cdot) := \frac{\mathbf{E}(\cdot \times X^2 \,|\, X > 0)}{\mathbf{E}(X^2 \,|\, X > 0)},$$

with which the previous formula becomes a constant multiple of

$$\widehat{\mathbf{E}} X \cdot \widehat{\mathbf{E}} \frac{1}{X} - 1 = -\widehat{\mathbf{Cov}}\Big(X,\, \frac{1}{X}\Big).$$

Notice that the $\widehat{}$ measure is concentrated on positive values, where $1/X$ is a decreasing function of X hence the above covariance is non-positive. A similar argument shows that (2.9) is non-negative.

Separate (2.10) into the sum of two terms:

$$\big[P_2\mathbf{P}\{X>0\}-P_1P_1\big]\cdot N_3+P_2\cdot\big[N_3\mathbf{P}\{X\le 0\}-N_2N_1\big].$$

Divide the first bracket by $[\mathbf{P}\{X>0\}]^2$ to get

$$\mathbf{E}(X^2\,|\,X>0)-[\mathbf{E}(X\,|\,X>0)]^2\ge 0.$$

Dividing the second bracket by $[\mathbf{P}\{X\le 0\}]^2$ leads to

$$\begin{aligned}\mathbf{E}(|X|^3\,\big|\,X\le 0)-\mathbf{E}(X^2\,\big|\,X\le 0)\cdot\mathbf{E}(|X|\,\big|\,X\le 0)\\ =\mathbf{Cov}(X^2,\,|X|\,\big|\,X\le 0)\ge 0\end{aligned}\tag{2.12}$$

since X^2 is an increasing function of $|X|$ on non-positive numbers. The term (2.11) is treated in a similar manner. □

Tracing the lemmas backward shows that we have verified part a) of Lemma 2.5 and thereby proved Theorem 2.1.

To prove Theorem 2.2, note first that in the complementary case $\Phi(X)=aX$ on $\mathbb{S}$, and then (2.2) implies that the derivative $\mathrm{d}\mathbf{E}^{\theta(\varrho)}\Phi(X)/\,\mathrm{d}\varrho$ is constant.

To prove the main statement of Theorem 2.2 we retrace some earlier steps. Let $x_0\in\mathbb{S}$ be a point of strict convexity whose existence is assumed. Namely,

$$\mathbf{P}\{X<x_0\}\mathbf{P}\{X=x_0\}\mathbf{P}\{X>x_0\}>0\tag{2.13}$$

and the slopes $b=\Phi'(x_0-)$ and $a=\Phi'(x_0+)$ satisfy $a>b$. Then we can write

$$\Phi(x)=\Phi(x_0)+a(x-x_0)^+-b(x-x_0)^-+\Psi(x)$$

for another convex function Ψ that vanishes on an interval around x_0. Since we already have Theorem 2.1 for Ψ, it suffices to prove strict convexity of $\varrho\mapsto\mathbf{E}^{\theta(\varrho)}\phi(X)$ for

$$\phi(x)=\Phi(x_0)+a(x-x_0)^+-b(x-x_0)^-.$$

After an application of the transformation (2.7) the question boils down to showing *strict* inequality in (2.5) for the new variable $Y=X-x_0$. For this it suffices to check that at least one of the quantities (2.8)–(2.11) is strictly positive. From (2.13) follows that each P_i and N_i is strictly positive. Schwarz inequality shows that $P_2P_2<P_3P_1$ if Y has two distinct strictly positive values, and $N_2N_2<N_3N_1$ if Y has two distinct strictly negative values. If both these requirements fail, then (2.13) forces Y to take one positive value, one negative value, and the value zero with positive probability. But then this makes the quantity in (2.12) strictly positive for Y.

Thus we conclude that strict inequality holds in (2.5) for $Y=X-x_0$, and strict convexity of $\varrho\mapsto\mathbf{E}^{\theta(\varrho)}\phi(X)$ follows. We have proved Theorem 2.2.

3 Application to stochastic interacting systems

To keep this note short we give a minimal possible introduction to the applications of the convexity result and refer the reader to [1] and [4] for the complete picture. Let $-\infty \le x^{\min} \le 0$ and $1 \le x^{\max} \le \infty$ be (possibly infinite valued) integers, and consider the discrete interval $I = (x^{\min} - 1,\, x^{\max} + 1) \cap \mathbb{Z}$. Fix a function $f : I \to \mathbb{R}^+$. For $I \ni x > 0$ we set

$$f(x)! := \prod_{y=1}^{x} f(y),$$

while for $I \ni x < 0$ let

$$f(x)! := \frac{1}{\prod\limits_{y=x+1}^{0} f(y)},$$

finally $f(0)! := 1$. Then we have

$$f(x)! \cdot f(x+1) = f(x+1)!$$

for all $x \in I$. Let

$$\bar{\theta} := \begin{cases} \log\left(\liminf\limits_{x\to\infty} \left(f(x)!\right)^{1/x}\right) & , \text{ if } x^{\max} = \infty \\ \infty & , \text{ else} \end{cases}$$

and

$$\underline{\theta} := \begin{cases} \log\left(\limsup\limits_{x\to\infty} \left(f(-x)!\right)^{-1/x}\right) & , \text{ if } x^{\min} = -\infty \\ -\infty & , \text{ else.} \end{cases}$$

We require f to be such that $\underline{\theta} < 0 < \bar{\theta}$. In this case

$$\mu(x) := \frac{\frac{1}{f(x)!}}{\sum\limits_{y\in I} \frac{1}{f(y)!}}$$

defines a probability measure on I, and the exponentially weighted version

$$\mu^{\theta}(x) := \frac{\frac{\mathrm{e}^{\theta x}}{f(x)!}}{\sum\limits_{y\in I} \frac{\mathrm{e}^{\theta y}}{f(y)!}}$$

is also well defined for any $\underline{\theta} < \theta < \bar{\theta}$. This latter is the marginal of a stationary product distribution of many stochastic interacting systems, see e.g. [4].

3.1 Convexity of hydrodynamic flux for zero range and bricklayer processes

In particular, the attractive zero range process is an example where $I = [0, \infty) \cap \mathbb{Z}$, $f(0) = 0 < f(1)$, and f is non-decreasing. The rate for a particle to jump from a site with x particles is $f(x)$. Its hydrodynamic (macroscopic) flux function $\mathcal{H} : \mathbb{R}^+ \to \mathbb{R}^+$ is given by

$$\mathcal{H}(\varrho) = \mathbf{E}^{\theta(\varrho)} f(X)$$

with the notation of the Introduction. The results of the previous section for f now read as follows:

Proposition 3.1. *If the jump rate f of the zero range process is convex (or concave), then the hydrodynamic flux $\mathcal{H}$ is also convex (or concave, respectively). Moreover, in this case $\mathcal{H}$ is strictly convex (or concave, respectively) if and only if f is not linear.*

The bricklayer process has $I = (-\infty, \infty) \cap \mathbb{Z}$ and f non-decreasing such that $f(x) \cdot f(1 - x) = 1$ for all $x \in \mathbb{Z}$. Its jump rate for a brick to be laid on a column between negative discrete gradients x on the left and y on the right is $f(x) + f(-y)$, see [4] for more details. The hydrodynamic flux function $\mathcal{H} : \mathbb{R} \to \mathbb{R}^+$ is now

$$\mathcal{H}(\varrho) = \mathbf{E}^{\theta(\varrho)} \big(f(X) + f(-Y) \big)$$

where X and Y are i.i.d. variables with distribution $\mu^{\theta(\varrho)}$. Notice that non-decreasingness and non-negativity of f on $\mathbb{Z}$ excludes concave functions with the exception of the constant one function. Our result for this process is

Proposition 3.2. *If the function f of the bricklayer process is convex and not constant one, then its hydrodynamic flux $\mathcal{H}$ is strictly convex.*

Parts of these two propositions were proved with coupling methods in [1].

3.2 Monotonicity of a special distribution

We come to the primary motivation of the note. As explained in the Introduction, the study of current fluctuations uses couplings of processes whose initial particle number at the origin obeys the following type of distribution:

$$\nu^{\theta(\varrho)}(y) = \frac{1}{\mathbf{Var}^{\theta(\varrho)} X} \sum_{x=y+1}^{x^{\max}} [x - \mathbf{E}^{\theta(\varrho)} X] \cdot \mu^{\theta(\varrho)}(x) \quad (x^{\min} \le y < x^{\max}). \tag{3.1}$$

(See [4, eqn. (2.6)] for the original definition.) To create couplings with useful monotonicity properties, one needs these distributions to be monotone in the parameter ϱ, in the sense of stochastic domination. This we can now derive as a consequence of the main result.

Proposition 3.3. *The family of measures $\nu^{\theta(\varrho)}$ is monotone in ϱ.*

Proof. By Corollary 2.4,

$$\begin{aligned}\nu^{\theta}(y) &= \frac{1}{\mathbf{Var}^{\theta} X} \cdot \mathbf{E}\big([X - \mathbf{E}^{\theta}(X)] \cdot \mathbf{1}\{X > y\}\big) \\ &= \frac{\mathbf{Cov}^{\theta}(X,\, \mathbf{1}\{X > y\})}{\mathbf{Var}^{\theta} X} = \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{P}^{\theta(\varrho)}\{X > y\}.\end{aligned}$$

Let us denote the $\nu^{\theta(\varrho)}$-expectation by $\mathbf{E}^{\nu,\,\theta(\varrho)}$. Monotonicity of the family $\nu^{\theta(\varrho)}$ is equivalent to the property that, for any bounded non-decreasing function φ,

$$0 \le \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\nu,\,\theta(\varrho)}\varphi(X).$$

We compute a different expression for this derivative. Passing the derivative through the sum in the third equality below is justified because the series involved are dominated by certain geometric series, uniformly over θ in small open neighborhoods. This follows from the definitions of $\underline{\theta}$ and $\bar{\theta}$ and the assumption $\underline{\theta} < 0 < \bar{\theta}$.

$$\begin{aligned}\mathbf{E}^{\nu,\,\theta(\varrho)}\varphi(X) &= \sum_{y=x^{\min}}^{x^{\max}} \varphi(y) \cdot \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{P}^{\theta(\varrho)}\{X > y\} \\ &= \sum_{y=x^{\min}}^{x^{\max}} \varphi(y) \cdot \frac{\mathrm{d}}{\mathrm{d}\varrho}[\mathbf{P}^{\theta(\varrho)}\{X > y\} - \mathbf{1}\{0 \ge y\}] \\ &= \frac{\mathrm{d}}{\mathrm{d}\varrho}\sum_{y=x^{\min}}^{x^{\max}} \varphi(y) \cdot [\mathbf{P}^{\theta(\varrho)}\{X > y\} - \mathbf{1}\{0 \ge y\}] \\ &= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\theta(\varrho)}\sum_{y=x^{\min}}^{x^{\max}} \varphi(y) \cdot [\mathbf{1}\{X > y\} - \mathbf{1}\{0 \ge y\}] \\ &= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\theta(\varrho)}\sum_{y=x^{\min}}^{x^{\max}} \varphi(y) \cdot [\mathbf{1}\{X > y > 0\} - \mathbf{1}\{0 \ge y \ge X\}] \\ &= \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\theta(\varrho)}\Big[\sum_{y=1}^{X-1}\varphi(y) - \sum_{y=X}^{0}\varphi(y)\Big] = \frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\theta(\varrho)}\Phi(X).\end{aligned}$$

Above we introduced the function

$$\Phi(x) = \sum_{y=1}^{x-1}\varphi(y) - \sum_{y=x}^{0}\varphi(y),$$

with the convention that empty sums are zero. To conclude the proof, notice that $\Phi(x+1) - \Phi(x) = \varphi(x)$. Thus a non-decreasing function φ determines a (non-strictly) convex function Φ with $\Phi(1) = 0$, and vice-versa. Hence Section 2 establishes that

$$\frac{\mathrm{d}}{\mathrm{d}\varrho}\mathbf{E}^{\nu,\,\theta(\varrho)}\varphi(X) = \frac{\mathrm{d}^2}{\mathrm{d}\varrho^2}\mathbf{E}^{\theta(\varrho)}\Phi(X) \ge 0. \qquad \square$$

Acknowledgment

We thank Omer Angel and Bálint Tóth for very illuminating discussions on the subject.

References

[1] M. Balázs. Growth fluctuations in a class of deposition models. *Ann. Inst. H. Poincaré Probab. Statist.*, 39:639–685, 2003.

[2] M. Balázs, F. Rassoul-Agha, T. Seppäläinen, and S. Sethuraman. Existence of the zero range process and a deposition model with superlinear growth rates. *Ann. Probab.*, 35(4), 2007.

[3] M. Balázs and T. Seppäläinen. Order of current variance and diffusivity in the asymmetric simple exclusion process. http://arxiv.org/abs/math.PR/0608400, 2006.

[4] M. Balázs and T. Seppäläinen. Exact connections between current fluctuations and the second class particle in a class of deposition models. *Journal of Stat. Phys.*, 127(2):431–455, 2007.

[5] P. A. Ferrari and L. R. G. Fontes. Current fluctuations for the asymmetric simple exclusion process. *Ann. Probab.*, 22:820–832, 1994.

[6] Samuel Karlin. *Total positivity. Vol. I.* Stanford University Press, Stanford, Calif, 1968.

[7] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems.* Springer-Verlag, Berlin, 1999.