POISSON THICKENING

ORI GUREL-GUREVICH AND RON PELED

Abstract. Let X be a Poisson point process of intensity λ on the real line. A *thickening* of it is a (deterministic) measurable function f such that $X \cup f(X)$ is a Poisson point process of intensity λ' where $\lambda' > \lambda$. An *equivariant* thickening is a thickening which commutes with all shifts of the line. We show that a thickening exists but an equivariant thickening does not. We prove similar results for thickenings which commute only with integer shifts and in the discrete and multi-dimensional settings. This answers 3 questions of Holroyd, Lyons and Soo.

1. Introduction and Results

Let $\mathbb{M}$ be the space of locally finite sets in $\mathbb{R}$, endowed with the local topology and Borel σ-algebra. Let X be a Poisson process of intensity λ on $\mathbb{R}$, viewed as an element of $\mathbb{M}$. A measurable function $f : \mathbb{M} \to \mathbb{M}$ is called a *thickening* if $X \cup f(X)$ is a Poisson process of intensity $\lambda' > \lambda$. A thickening f is *equivariant* if $\sigma \circ f = f \circ \sigma$ for any shift operator $\sigma : \mathbb{R} \to \mathbb{R}$.

Theorem 1.1. *An equivariant thickening does not exist for any $\lambda' > \lambda$.*

Theorem 1.2. *A (non-equivariant) thickening exists for every $\lambda' > \lambda > 0$.*

Extending naturally the definition of thickening to Poisson processes in $\mathbb{R}^d$, our results extend also to this multi-dimensional setting and to the case where the thickening commutes only with a sub-group of all shifts.

Theorem 1.3. *For any dimension $d \geq 1$ and any pair of intensities $\lambda' > \lambda$:*

1. *There is no thickening which commutes with d linearly independent shifts in $\mathbb{R}^d$.*
2. *There exists a thickening which commutes with a $(d-1)$-dimensional space of shifts in $\mathbb{R}^d$.*

In other works on equivariant extension of processes, it is also common to have the function f depend on additional randomness and have the equality $\sigma \circ f = f \circ \sigma$ hold only in distribution. In our context, existence of such randomized equivariant thickenings is trivial.

A result analogous to Theorem 1.2 holds also in the discrete setting. Let $X = \{X_i\}_{i \in \mathbb{Z}}$ be a sequence of i.i.d. $\{0, 1\}$-valued random variables with

$\mathbb{E}(X_0) = \lambda$. A function $f : \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ is a *discrete thickening*, if the sequence $\{f(X)_i\}_{i\in\mathbb{Z}}$ is i.i.d. with $\mathbb{E}(f(X)_i) = \lambda' > \lambda$ and $f(X)_i \geq X_i$ for all $i \in \mathbb{Z}$. An equivariant function cannot increase the entropy of a process (Kolmogorov-Sinai theorem, see e.g. [4], Chapter 5), hence, there is no equivariant discrete thickening if $\lambda < \lambda' < 1 - \lambda$. The case $1 - \lambda < \lambda'$ was resolved by Ball [3], who showed the existence of such thickenings.

Theorem 1.4. *A (non-equivariant) discrete thickening exists.*

Our three theorems answer the 3 questions posed by Holroyd, Lyons and Soo in [1] where the related problem of Poisson splitting was addressed. The problem of Poisson thinning was considered earlier by Ball [2].

2. Non-existence of equivariant thickenings

Proof of Theorem 1.1. For simplicity we will set $2 = \lambda' > \lambda = 1$, but the proof works just as well for any $\lambda' > \lambda$.

Assume, in order to obtain a contradiction, that there exists an equivariant thickening $f : \mathbb{M} \to \mathbb{M}$". Let X be a Poisson process of intensity 1 and let $Y := X \cup f(X)$. We assume, WLOG, that $X \cap f(X) = \emptyset$, since we can always replace $f(X)$ by $f'(X) = f(X) \setminus X$, which also satisfies $Y = X \cup f'(X)$. Now split Y into two disjoint sets $Y = Y_1 \cup Y_2$ by randomly and independently assigning each point of Y to Y_1 or Y_2 with probability $\frac{1}{2}$.

Let Z be a Poisson process of intensity 2, and split it similarly into Z_1 and Z_2. The resulting distribution on (Z_1, Z_2) is simply the distribution of 2 independent Poisson processes of intensity 1. We will show that the distribution of (Y_1, Y_2) differs from that of (Z_1, Z_2) by constructing an event which has different probabilities under these two distributions. This will imply that the distribution of $Y = X \cup f(X)$ is different from that of Z, since the splitting process is the same.

Informally, we consider the possibility that the splitting (Y_1, Y_2) of Y coincides with $(X, f(X))$ on some large interval. On this event, there is another large interval on which $f(Y_1)$ is much more correlated with Y_2 than what we would get for $f(Z_1)$ and Z_2. The equivariance condition enters in ensuring that the probability of this event decays only exponentially fast in the length of the interval on which the correlation holds.

For a Borel $S \subset \mathbb{R}$, let $\mathcal{F}_S$ denote the Borel σ-algebra of the restriction of $\mathbb{M}$ to S, i.e. all the events which depend only on the points of the process which are in S.

For $\varepsilon > 0$ and $U \in \mathbb{M}$, let $A_t(U)$ be the event "U has at least one point in $[t, t+\varepsilon]$". We claim that

$$\mathbb{P}(A_0(f(X))) = \varepsilon(1 + o(1)) \tag{2.1}$$

as $\varepsilon \to 0$. To see this, note that

$$\mathbb{P}(A_0(f(X))) \le \mathbb{E}|f(X) \cap [0,\varepsilon]| = \mathbb{E}|Y \cap [0,\varepsilon]| - \mathbb{E}|X \cap [0,\varepsilon]| = \varepsilon \ ,$$

and

$$\mathbb{P}(A_0(f(X))) \ge \mathbb{P}(A_0(Y)) - \mathbb{P}(A_0(X)) = \varepsilon(1+o(1)) \ .$$

These two inequalities establish (2.1). We continue under the assumption that $\varepsilon > 0$ is small enough so that $\varepsilon/2 < \mathbb{P}(A_0(f(X))) < 2\varepsilon$.

Now, $A_0(f(X))$ can be $\varepsilon/4$-approximated by some event $B_0(X) \in \mathcal{F}_{[-r,r]}$ for some $r < \infty$, i.e. $B_0(X)$ depends only on the points of X in $[-r,r]$ and $\mathbb{P}(A_0(f(X)) \vartriangle B_0(X)) < \varepsilon/4$, where $\vartriangle$ denotes the symmetric difference. In particular $\mathbb{P}(B_0(X)) \le 4\varepsilon$ and if we define $C_0(U) = A_0(f(U)) \wedge B_0(U)$ then we also have $\mathbb{P}(C_0(X)) \ge \varepsilon/4$. If we define $B_t(U) = B_0(\sigma_t(U))$ and $C_t(U) = C_0(\sigma_t(U))$, where $\sigma_t : \mathbb{R} \to \mathbb{R}$ is translation by t, then by our equivariance assumption, $B_t(X)$ is measurable with respect to $\mathcal{F}_{[-r+t,r+t]}$ and approximates $A_t(f(X))$ similarly.

Consider the events $\{B_{i\varepsilon}(U)\}_{i=0}^{L-1}$ and note that they all belong to $\mathcal{F}_{[-r,L\varepsilon+r]}$. Let $b(U)$ be the number of these events that occur, so $\mathbb{E}(b(X)) = \mathbb{P}(B_0(X))L \le 4L\varepsilon$. By ergodicity of X with respect to the shift by ε, we get that $\mathbb{P}(b(X) < 5L\varepsilon) \to 1$ as $L \to \infty$.

Similarly, if $c(U)$ is the number of events that occur among $\{C_{i\varepsilon}(U)\}_{i=0}^{L-1}$, then $\mathbb{P}(c(X) > L\varepsilon/8) \to 1$ as $L \to \infty$ (although these events do not necessarily belong to $\mathcal{F}_{[-r,L\varepsilon+r]}$).

For $\varepsilon > 0$ and $U, V \in \mathbb{M}$ let $D_t(U,V) = B_t(U) \wedge A_t(V)$ and let $d(U,V)$ denote the number of events that occur among $\{D_{i\varepsilon}(U,V)\}_{i=0}^{L-1}$. Notice that $D_t(U, f(U)) = B_t(U) \wedge A_t(f(U)) = C_t(U)$. In particular, $\mathbb{P}(d(X,f(X)) > L\varepsilon/8) \to 1$ as $L \to \infty$.

Finally, let $E(U,V)$ be the event "$b(U) < 5L\varepsilon$ and $d(U,V) > L\varepsilon/8$". We claim that this event distinguishes between (Y_1,Y_2) and (Z_1,Z_2).

Define

$$\Omega_1 := \{X \cup f(X) \text{ has at most } 3(L\varepsilon + 2r) \text{ points in } [-r, L\varepsilon + r]\},$$
$$\Omega_2 := \{Y_1|_{[-r,L\varepsilon+r]} = X|_{[-r,L\varepsilon+r]}\}$$

and notice that Ω_2 is also equivalent to $\{Y_2|_{[-r,L\varepsilon+r]} = f(X)|_{[-r,L\varepsilon+r]}\}$.

As $L \to \infty$, we have

$$\mathbb{P}(\Omega_1) \to 1,$$
$$\mathbb{P}(\Omega_2 \mid \Omega_1) \ge 2^{-3(L\varepsilon+2r)},$$
$$\mathbb{P}(E(Y_1,Y_2) \mid \Omega_2) = \mathbb{P}(E(X,f(X))) \to 1$$

Hence, for sufficiently large L we have

$$\mathbb{P}(E(Y_1,Y_2)) \ge 2^{-1-3(L\varepsilon+2r)} \ . \tag{2.2}$$

On the other hand, since Z_2 is independent of Z_1, we have

$$\mathbb{P}(A_t(Z_2)|B_t(Z_1)) \leq \varepsilon \ .$$

Therefore, if $b(Z_1) < 5L\varepsilon$, Then the probability of $A_t(Z_2)$ happening more than $L\varepsilon/8$ times among these $b(Z_1)$ indices is bounded by $2^{5L\varepsilon}\varepsilon^{L\varepsilon/8}$, since $A_{i\varepsilon}(Z_2)$ and $A_{j\varepsilon}(Z_2)$ are independent when $i \neq j$. Consequently,

$$\mathbb{P}(E(Z_1, Z_2)) \leq 2^{5L\varepsilon}\varepsilon^{L\varepsilon/8}. \tag{2.3}$$

Comparing (2.2) with (2.3) for small enough $\varepsilon > 0$ and large enough L show that (Y_1, Y_2) and (Z_1, Z_2) do not have the same distribution, yielding a contradiction to the existence of f. ■

3. Existence of non-equivariant thickenings

The proofs of Theorems 1.2 and 1.4 are essentially the same. We will first prove Theorem 1.4 and then discuss the changes needed to prove Theorem 1.2.

Let $0 < \lambda < \lambda' < 1$ be fixed. Denote by $\mathbb{P}_p$ the distribution of a $\{0,1\}$-valued random variable with expectation p, and let $\mathbb{P}_p^I$ be a set of i.i.d. $\mathbb{P}_p$ random variables, indexed by I. Our goal is to construct $f : \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ such that if $X \sim \mathbb{P}_\lambda^{\mathbb{Z}}$ then $f(X) \sim \mathbb{P}_{\lambda'}^{\mathbb{Z}}$ and a.s. for all $i \in \mathbb{Z}$ we have $f(X)_i \geq X_i$.

Since we don't require equivariance, the specific choice of index set plays no role beyond its cardinality. That is, there is no difference between a (discrete) thickening on $\{0,1\}^{\mathbb{Z}}$, as in the statement of Theorem 1.4, and a thickening on $\{0,1\}^{\mathbb{N}}$ or $\{0,1\}^{\mathbb{N}\times\mathbb{N}}$ (which are defined analogously). To be more specific, let $n : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ be a bijection and let $h : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}\times\mathbb{N}}$ be the isomorphism defined by $h(X)_{ij} = X_{n(i,j)}$. If f is a thickening of $\mathbb{P}_\lambda^{\mathbb{N}}$ into $\mathbb{P}_{\lambda'}^{\mathbb{N}}$, then $h \circ f \circ h^{-1}$ is a thickening of $\mathbb{P}_\lambda^{\mathbb{N}\times\mathbb{N}}$ into $\mathbb{P}_{\lambda'}^{\mathbb{N}\times\mathbb{N}}$ and vice versa.

Another useful fact is that $\mathbb{P}_p^{\mathbb{N}}$ and $\mathbb{P}_q^{\mathbb{N}}$ are isomorphic (as measure spaces), for any $0 < p, q < 1$. Let g be such an isomorphism taking $\mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$ into $\mathbb{P}_\mu^{\mathbb{N}}$, where $\mu := \frac{\lambda'-\lambda}{1-\lambda}$. μ is chosen so that if $x \sim \mathbb{P}_\lambda$ and $y \sim \mathbb{P}_\mu$ are independent then $\max(x, y) \sim \mathbb{P}_{\lambda'}$.

We define an *extractor* to be a function $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ such that if $X \sim \mathbb{P}_\lambda^{\mathbb{N}}$ and $Y \sim \mathbb{P}_\mu^{\mathbb{N}}$ are independent then

$$f(X) \sim \mathbb{P}_{\frac{1}{2}} \qquad \text{and}$$
$$f(X) \text{ and } \max(X, Y) \text{ are independent.}$$

We remark that this is different from the extractor which is sometimes used in the computer science literature.

How are extractors useful? First, notice that given independent $X \sim \mathbb{P}_\lambda^{\mathbb{N}}$ and $Y \sim \mathbb{P}_\mu^{\mathbb{N}}$, by rearranging indices (using the function n above) one can extract infinitely many bits from X, i.e. one can get a function $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$, such that $f(X) \sim \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$ independently of $\max(X,Y)$. Second, by applying g we can extract a sequence distributed $\mathbb{P}_{\mu}^{\mathbb{N}}$. Now, to thicken $\mathbb{P}_{\lambda}^{\mathbb{N}\times\mathbb{N}}$ into $\mathbb{P}_{\lambda'}^{\mathbb{N}\times\mathbb{N}}$, define $F : \{0,1\}^{\mathbb{N}\times\mathbb{N}} \to \{0,1\}^{\mathbb{N}\times\mathbb{N}}$ by

$$F(X)^i = \max(X^i, g(f(X^{i+1})))$$

where for $U \in \{0,1\}^{\mathbb{N}\times\mathbb{N}}$ we write U^i for $U(i,\cdot)$.

Claim 3.1. *If f is an extractor, F is a thickening.*

Proof. Fix j and consider the distribution of $\{F(X)^i\}_{i=1}^j$. We will show by induction on $0 \le k \le j$ that the bits in $\{F(X)^i\}_{i=j-k+1}^j$, $f(X^{j-k+1})$ and $\{X^i\}_{i=1}^{j-k}$ are all independent. For $k = j$ this implies the desired result.

For the base case, $k = 0$, we need to show that $f(X^{j+1})$ and $\{X^i\}_{i=1}^j$ are independent, which is trivially true.

Assume the induction hypothesis holds for a particular k. Since $f(X^{j-k+1})$ and X^{j-k} are independent we get that $F(X)^{j-k} = \max(X^{j-k}, g(f(X^{j-k+1})))$ is distributed $P_{\lambda'}^{\mathbb{N}}$ and is independent of $\{F(X)^i\}_{i=j-k+1}^j$ and $\{X^i\}_{i=1}^{j-k-1}$. Since f is an extractor, we get that $f(X^{j-k})$ is independent of $F(X)^{j-k} = \max(X^{j-k}, g(f(X^{j-k+1})))$ and of course also of $\{F(X)^i\}_{i=j-k+1}^j$ and $\{X^i\}_{i=1}^{j-k-1}$. ∎

All that is left, then, is to construct an extractor. Unfortunately, such an object does not exist.

Lemma 3.2. *There is no extractor.*

Proof. Let f be an extractor. We will reach a contradiction by showing that $f(X)$ is independent of $\{X_i\}_{i=1}^k$ for any integer k.

Fix $k \ge 1$. For $U \in \{0,1\}^{\mathbb{N}}$, let $A(U)$ be the event $\wedge_{i\le k}(U_i = 0)$. Since f is an extractor, $f(X)|A(Z) \sim \mathbb{P}_{\frac{1}{2}}$, but $A(Z) = A(X) \wedge A(Y)$, and X and Y are independent, so the distribution of $X|A(Z)$ is the same as $X|A(X)$, so $f(X)|A(X) \sim \mathbb{P}_{\frac{1}{2}}$.

Now, for $1 \le j \le k$, let $A_j(U)$ be the event $\wedge_{i\le k, i\ne j}(U_i = 0) \wedge (U_j = 1)$. Again, $f(X)|A_j(Z) \sim \mathbb{P}_{\frac{1}{2}}$, but now $X|A_j(Z)$ is $\frac{\lambda}{\lambda'}X|A_j(X) + (1-\frac{\lambda}{\lambda'})X|A(X)$ (that is, $X_i = 0$ for $i \le k, i \ne j$ and $X_j \sim \mathbb{P}_{\lambda/\lambda'}$), since $\mathbb{P}(X_j = 1|Z_j = 1) = \frac{\lambda}{\lambda'}$. We already know that $f(X)|A(X) \sim \mathbb{P}_{\frac{1}{2}}$, so we conclude that also $f(X)|A_j(X) \sim \mathbb{P}_{\frac{1}{2}}$.

Proceed by induction on the number of 1's among $\{Z_i\}_{i=1}^k$ to show that conditioned on any sequence of values for $\{X_i\}_{i=1}^k$, $f(X)$ is distributed $\mathbb{P}_{\frac{1}{2}}$. ∎

Fortunately, one can make do with something that is only almost an extractor, though the way it is used will be a bit more complicated. An ε-*extractor* is a function $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ such that if $X \sim \mathbb{P}_\lambda^{\mathbb{N}}$ and $Y \sim \mathbb{P}_\mu^{\mathbb{N}}$ are independent then

$$f(X) \sim \mathbb{P}_{\frac{1}{2}} \qquad \text{and}$$
$$\mathbb{E} d_{\mathrm{TV}}\Big(\mathcal{L}\big(f(X) \mid \max(X,Y)\big), \mathbb{P}_{\frac{1}{2}}\Big) < \varepsilon,$$

where $\mathcal{L}(f(X) \mid \max(X,Y))$ is the law of $f(X)$ conditioned on $\max(X,Y)$ and $d_{\mathrm{TV}}(\mathcal{L}_1, \mathcal{L}_2)$ is the total variation distance between the laws $\mathcal{L}_1$ and $\mathcal{L}_2$. That is, observing $\max(X,Y)$ gives us little information on $f(X)$. Learning from our previous experience, we first verify the existence of ε-extractors.

Lemma 3.3. *For any $\varepsilon > 0$ there is an ε-extractor.*

Proof. Fix $\varepsilon > 0$. For an integer $k \geq 1$, let a_k be the parity of the first k values of X, i.e. $a_k := \bigoplus_{i=1}^k X_i$. Let $\ell_k := \sum_{i=1}^k \max(X_i, Y_i)$. Then it is readily verified, using the Fourier transform, that $\mathbb{E}(a_k | \max(X,Y)) \sim \mathbb{P}_{(1-(1-2\frac{\lambda}{\lambda'})^{\ell_k})/2}$. Let ℓ' be such that $(1-2\frac{\lambda}{\lambda'})^{\ell'} < \varepsilon/2$ and fix k large enough so that $\mathbb{P}(\ell < \ell') < \varepsilon/2$. Then, observing that $d_{\mathrm{TV}}\Big(\mathcal{L}\big(a_k \mid \max(X,Y)\big), \mathbb{P}_{\frac{1}{2}}\Big) = |\mathbb{P}\big(a_k = 1 \mid \max(X,Y)\big) - \frac{1}{2}|$, we have

$$\mathbb{E} d_{\mathrm{TV}}\Big(\mathcal{L}\big(a_k \mid \max(X,Y)\big), \mathbb{P}_{\frac{1}{2}}\Big) < \mathbb{P}(\ell_k \geq \ell')\varepsilon/2 + \mathbb{P}(\ell_k < \ell')1 \leq \varepsilon.$$

Let m be the minimal positive integer such that $X_{k+2m} \neq X_{k+2m+1}$ and let $b_k := X_{k+2m}$. Then $b_k \sim \mathbb{P}_{\frac{1}{2}}$ and is independent of a_k, both unconditionally and conditionally on $\max(X,Y)$. Therefore, $f(X) := a_k \oplus b_k$ satisfies both requirements of being an ε-extractor. ∎

Of course, we cannot simply replace the extractors from the proof of Claim 3.1 with ε-extractors, since one might learn something about the output bits of the ε-extractors by observing the thickening of the bits from which they were extracted. We will therefore introduce a "correction" mechanism for these bits.

Given an ε-extractor f, a *corrector* for f is a function $f' : \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \to \{0,1\}$ such that when $X \sim \mathbb{P}_\lambda^{\mathbb{N}}$, $Y \sim \mathbb{P}_\mu^{\mathbb{N}}$ and $Z \sim \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$ are independent, the following properties hold:

$$f(X) \oplus f'(X,Y,Z) \sim \mathbb{P}_{\frac{1}{2}},$$
$$f(X) \oplus f'(X,Y,Z) \text{ and } \max(X,Y) \text{ are independent},$$
$$\mathbb{E}(f'(X,Y,Z)) < \varepsilon.$$

Claim 3.4. *For any ε-extractor, there is a corrector.*

Proof. Let f be an ε-extractor. Let $U : \{0,1\}^{\mathbb{N}} \to [0,1]$ be defined by $U(Z) = \sum_{i=1}^{\infty} Z_i 2^{-i}$ so that $U(Z)$ is a uniform random variable on $[0,1]$. Now for $x, y, z \in \{0,1\}^{\mathbb{N}}$ we define f' as

$$f'(x,y,z) = \begin{cases} 1 & U(z)(\mathbb{P}(f(X) = f(x) \mid \max(X,Y) = \max(x,y))) > \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

Given $m \in \{0,1\}^{\mathbb{N}}$, what is the probability that $f(X) \oplus f'(X,Y,Z) = 0$, conditioned on $\max(X,Y) = m$?

Let $p = \mathbb{P}(f(X) = 0 \mid \max(X,Y) = m)$, and assume, WLOG that $p \geq \frac{1}{2}$. Then

$$\mathbb{P}(f(X) = 0, f'(X,Y,Z) = 0 \mid \max(X,Y) = m) = p\frac{1}{2p} = \frac{1}{2}$$

and

$$\mathbb{P}(f(X) = 1, f'(X,Y,Z) = 1 \mid \max(X,Y) = m) = (1-p)0$$

Hence, $f(X) \oplus f'(X,Y,Z) \sim \mathbb{P}_{\frac{1}{2}}$, independently of $\max(X,Y)$.

To see that the third requirement is satisfied, notice that

$$\mathbb{P}(f'(X,Y,Z) = 1 | \max(X,Y) = m) = d_{TV}(\mathcal{L}(f(X) | \max(X,Y) = m), \mathbb{P}_{\frac{1}{2}}) \ .$$

Taking expectations, we get

$$\mathbb{P}(f'(X,Y,Z) = 1) = \mathbb{E}(d_{TV}(\mathcal{L}(f(X) | \max(X,Y)), \mathbb{P}_{\frac{1}{2}})) < \varepsilon$$

since f is an ε-extractor. ∎

As before, we need more than a single bit. An *ε-extractor into* $\{0,1\}^{\mathbb{N}}$ is a function $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ such that if $X \sim \mathbb{P}_{\lambda}^{\mathbb{N}}$ and $Y \sim \mathbb{P}_{\mu}^{\mathbb{N}}$ are independent then

$$f(X) \sim \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}} \qquad \text{and}$$

$$\mathbb{E}d_{\mathrm{TV}}\Big(\mathcal{L}\big(f(X) \mid \max(X,Y)\big), \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}\Big) < \varepsilon.$$

To construct an ε-extractor into $\{0,1\}^{\mathbb{N}}$ we take a sequence of functions $f_i : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ such that f_i is an $\varepsilon 2^{-i}$-extractor and define

$$f(X)_i = f_i(h(X)^i) \ ,$$

where we recall that h is an isomorphism taking $\{0,1\}^{\mathbb{N}}$ into $\{0,1\}^{\mathbb{N}\times\mathbb{N}}$.

For f, an ε-extractor into $\{0,1\}^{\mathbb{N}}$, one calls $f' : \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ a *corrector*, if when $X \sim \mathbb{P}_{\lambda}^{\mathbb{N}}$, $Y \sim \mathbb{P}_{\mu}^{\mathbb{N}}$ and $Z \sim \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$ are independent, the following properties hold:

$$f(X) \oplus f'(X,Y,Z) \sim \mathbb{P}_{\frac{1}{2}}^{\mathbb{N}},$$

$$f(X) \oplus f'(X,Y,Z) \text{ and } \max(X,Y) \text{ are independent},$$

$$\mathbb{P}(f'(X,Y,Z) \neq (0,0,\ldots)) < \varepsilon.$$

Existence of correctors can be proved by the methods of Claim 3.4. However, if the ε-extractor is constructed as above, as a sequence of $\varepsilon 2^{-i}$-extractors, then one can take a corresponding sequence of correctors to get a corrector for this ε-extractor.

Given an ε-extractor, f, and an associated corrector, f', one defines the *corrected extractor*, $f'' : \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$, to be $f''(X,Y,Z) := f(X) \oplus f'(X,Y,Z)$. Corrected extractors are very similar to extractors. The difference is that they depend, though rather weakly, on extra bits, and that they exist.

We are now prepared to prove our theorem.

Proof of Theorem 1.4. First, by using h we transfer the problem to thickening $\mathbb{P}_{\lambda}^{\mathbb{N}\times\mathbb{N}}$ into $\mathbb{P}_{\lambda'}^{\mathbb{N}\times\mathbb{N}}$.

For $i \in \mathbb{N}$, let f_i be a $\frac{1}{2^i}$-extractor into $\mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$, Let f_i' be a corresponding corrector, and let f_i'' be the resulting corrected extractor.

We would like to make the following definitions: for $i \in \mathbb{N}$

$$\begin{aligned} Y^i &:= g(h(f''_{i+1}(X^{i+1}, Y^{i+1}, Z^{i+1}))^0), \\ Z^i &:= h(f''_{i+1}(X^{i+1}, Y^{i+1}, Z^{i+1}))^1, \\ F(X)^i &:= \max(X^i, Y^i). \end{aligned}$$

Then the same argument as in Claim 3.1 would show F is a thickening. Alas, this is not well defined since for each i, (Y^i, Z^i) depend on (Y^{i+1}, Z^{i+1}) and so on ad infinitum. However, since corrections are rare, there is a way to make sense of the above definitions, as follows.

For $n \in \mathbb{N}$ define Y_n and Z_n by

$$(Y_n)^i := \begin{cases} g(h(f_{i+1}(X^{i+1}))^0) & \text{if } i \geq n \\ g(h(f''_{i+1}(X^{i+1}, (Y_n)^{i+1}, (Z_n)^{i+1}))^0) & \text{if } i < n \end{cases}$$

$$(Z_n)^i := \begin{cases} h(f_{i+1}(X^{i+1}))^1 & \text{if } i \geq n \\ h(f''_{i+1}(X^{i+1}, (Y_n)^{i+1}, (Z_n)^{i+1}))^1 & \text{if } i < n \end{cases}$$

In other words, we use f instead of f'' above n. Since f depends only on X, this yields, for any $n \in \mathbb{N}$, well defined sequences, Y_n and Z_n.

Claim 3.5. *Y_n and Z_n a.s. converge (pointwise) as $n \to \infty$.*

Proof. What is the probability that $Y_n, Z_n = Y_{n+1}, Z_{n+1}$?

First, notice that $\big((Y_n)^i, (Z_n)^i\big) = \big((Y_{n+1})^i, (Z_{n+1})^i\big)$ for any $i > n$ and if $\big((Y_n)^n, (Z_n)^n\big) = \big((Y_{n+1})^n, (Z_{n+1})^n\big)$ then we have $\big((Y_n)^i, (Z_n)^i\big) = \big((Y_{n+1})^i, (Z_{n+1})^i\big)$ for all i, by backward induction on i.

Using that f_{n+1} is a $2^{-(n+1)}$-extractor, f'_{n+1} is a corrector for f_{n+1} and the definition of the corrected extractor f''_{n+1}, we get for each $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}\Big(\big((Y_n)^n, (Z_n)^n\big) \neq \big((Y_{n+1})^n, (Z_{n+1})^n\big)\Big) = \\ = \mathbb{P}(f_{n+1}(X^{n+1}) \neq f''_{n+1}(X^{n+1}, (Y_{n+1})^{n+1}, (Z_{n+1})^{n+1})) = \\ = \mathbb{P}(f'_{n+1}(X^{n+1}, (Y_{n+1})^{n+1}, (Z_{n+1})^{n+1}) \neq (0, 0, \dots) < \frac{1}{2^{n+1}}.\end{aligned}$$

The sum of these probabilities is finite, hence there exists a.s. an $m \in \mathbb{N}$ such that $\big(Y_n, Z_n\big) = \big(Y_{n+1}, Z_{n+1}\big)$ for all $n > m$, proving the claim. ■

Let Y and Z be the limits of Y_n and Z_n respectively. Let $F(X)^i := \max(X^i, Y^i)$. Now F is easily seen to be a thickening by the inductive argument of Claim 3.1. ■

To adapt this argument to prove Theorem 1.2 one needs to construct an ε-extractor for a Poisson process instead of $\{0,1\}^{\mathbb{N}}$. To do this let a be the parity of the number of points in $X|_{[-r,r]}$ and let $b := \operatorname{sgn}(\min(X|_{(r,\infty)}) + \max(X|_{(-\infty,-r)}))$. Then for r large enough $a \oplus b$ is an ε-extractor. Note that Lemma 3.2 also holds in this context; the proof is by induction on the number of points of $Z|_{[-r,r]}$.

Two other ingredients are needed: The first, an isomorphism $h : \mathbb{M} \to \mathbb{M} \times \mathbb{N}$, can be induced from an isomorphism $n : \mathbb{R} \to \mathbb{R} \times \mathbb{N}$. The second is an isomorphism $g : \{0,1\}^{\mathbb{N}} \to \mathbb{M}$, taking $\mathbb{P}_{\frac{1}{2}}^{\mathbb{N}}$ into a Poisson process of intensity 1.

4. Thickenings in higher dimensions

As noted in the introduction, the proof of Theorem 1.1 also applies when we require f to be equivariant only with respect to integer shifts (and by scaling, for any lattice of shifts). The reason for this is that we may allow $\{A_{i\varepsilon}(f(X))\}_{i=0}^{\lfloor 1/\varepsilon \rfloor -1}$ to be approximated by $\{B_{i\varepsilon}(X)\}_{i=0}^{\lfloor 1/\varepsilon \rfloor -1}$, where B_t are not necessarily shifts of each other as long as they all belong to $\mathcal{F}_{[-r,r]}$ for some $r < \infty$. This can be done since there are only finitely many of these events (specifically, $\lfloor 1/\varepsilon \rfloor$). We can then shift those events by $1, 2, \dots, n$ to get n/ε events belonging to $\mathcal{F}_{[-r,n+r]}$. The proof then proceeds as before.

The same proof also holds in dimension $d > 1$, when the function is equivariant with respect to d linearly independent shifts. In this case we consider events of the form "there is a point of U in $[t_1, t_1 + \varepsilon] \times [t_2, t_2 + \varepsilon] \times \cdots \times [t_d, t_d + \varepsilon]$", and get an event with probability which decays only exponentially in the volume of the region for $X \cup f(X)$, whereas for a Poisson process it decays exponentially fast with constants depending on ε.

It is easy to see that there is a thickening equivariant with respect to $d-1$ independent shifts. By applying a linear transformation, we may assume without loss of generality that these shifts are by the first $d-1$ unit vectors. One can then split $\mathbb{R}^d$ (or $\mathbb{Z}^d$) into strips of $1\times 1\times\cdots\times\mathbb{R}$, and use the same non-equivariant thickening in each of these strips. The existence of a non-equivariant thickening in a strip is guaranteed either by constructing it directly, by the methods of Theorem 1.2, or by noting that the strip and $\mathbb{R}$ are isomorphic as measure spaces, and this induces an isomorphism between the corresponding Poisson processes.

It is only slightly harder to see how to construct a thickening equivariant with respect to *any* shift of some $d-1$ dimensional space. Let us illustrate how to get a thickening in $\mathbb{R}^2$ which is equivariant with respect to all horizontal shifts. Consider the strip $\mathbb{R}\times[0,1]$ and let $\{x_i\}_{i\in\mathbb{Z}}$ be the set of x-coordinates of points in the original process in this strip, indexed in increasing order. Of course, there are countably many ways to index, but the actual indices won't matter, only the order. Now take each vertical strip $(x_i, x_{i+1})\times\mathbb{R}$, and apply a non-equivariant thickening to it, taking into account that there won't be any points in $(x_i, x_{i+1})\times[0,1]$. The resulting function is, of course, a thickening and equivariant with respect to horizontal shifts.

References

[1] A. E. Holroyd, R. Lyons, and T. Soo, *Poisson Splitting by Factors*. preprint available at http://arxiv.org/abs/0908.3409.

[2] K. Ball, *Poisson Thinning by Monotone Factors*, Electronic Communications in Probability **10** (2005), 60-69.

[3] K. Ball, *Monotone factors of i.i.d. processes*, Israel J. Math. **150** (2005), 205–227.

[4] K. Petersen, *Ergodic Theory, volume 2 of Cambridge Studies in Advanced Mathematics*, Cambridge University Press, 1983.

Ori Gurel-Gurevich

Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.

E-mail address: origurel@microsoft.com

Ron Peled

Courant Institute of Mathematical Sciences, 251 Mercer st., New York, NY 10012, USA,

E-mail address: peled@cims.nyu.edu

URL: http://www.cims.nyu.edu/~peled