Positive Influence and Negative Dependence *

Devdatt Dubhashi[†]
Department of Computing Science
Chalmers University
SE 412 96 Göteborg, Sweden
e-mail: dubhashi@cs.chalmers.se

Johan Jonasson[‡]
Department of Mathematics
Chalmers University
SE 412 96 Göteborg, Sweden
e-mail: jonasson@math.chalmers.se

Desh Ranjan[§]
Department of Computer Science
New Mexico State University
Las Cruces, New Mexico 88003, USA
dranjan@cs.nmsu.edu

June 30, 2003

Abstract

We study negative dependence properties of a sampling process due to Srinivasan to produce distributions on level sets with given marginals. We give a simple proof that the distribution satisfies negative association. We also show that under linear match schedule it

*Work partially supported by the ESPRIT Basic Research Actions Program of the EC under contract No. 7141 (project ALCOM II)

†Work done partly while the author was visiting BRICS.

‡Work supported by the Swedish Research Council

§Work done partly while the author was visiting the Max–Planck–Institut für Informatik and BRICS.

satisfies the stronger condition of conditional negative association via a non-trivial application of the Feder-Mihail Theorem. This method involves the notion of a variable of positive influence. We give some results and counter-examples related to it which might shed some light on its role in a theory of negative dependence.

1 Introduction

Pemantle[13] points out that there is a need for a theory of negatively dependent events analogous to the existing natural and useful theory of positively dependent events. He gives a survey of various notions of negative dependence, three of which are reviewed below, and points out several applications of consequences of one of these concepts: negative association. These examples include natural stochastic processes and structures such as the uniform random spanning tree[11], the simple exclusion process[10], the random cluster model and occupancy numbers of competing bins[3]. Negative dependence may also be used to obtain information on the distribution of functionals such as the sum of the variables involved (or more generally, a non-decreasing function). Newman[12] shows that under negative dependence, one obtains a Central Limit Theorem (CLT) for stationary sequences of random variables. Dubhashi and Ranjan[3] show that under negative dependence conditions, one can employ classical tools of concentration of measure such as the Chernoff-Hoeffding (CH) bounds and related martingale inequalities.

In this paper, we contribute to the incipient theory of negative dependence by adding to the stock of examples of negative dependence. This class of examples involves distributions on level sets with given marginals that is introduced in the next subsection. We also add to the body of available techniques, by giving a non-trivial application of the Feder-Mihail Theorem [6] to this class of examples.

1.1 Randomized Rounding and Distributions on Level Sets

Recent results in approximation algorithms have suggested new applications of negative dependence via the very successful technique of randomized rounding [15]. Various hard combinatorial optimization problems can often be solved approximately via the following steps:

1. Formulate the problem as an integer linear program (ILP) involving decision variables taking binary values.

2. Relax the ILP to the corresponding linear program (LP) and solve it optimally. (There are several polynomial time algorithms to do this.)

3. Round the fractional values of the optimal LP solution by a randomized rounding rule that uses the optimal LP solution as probabilities: $X_i = 1,$ with probability x_i^* and $X_i = 0,$ with probability $1 - x_i^*$.

Often the ILP involves *hard cardinality constraints* corresponding to sharing of resources, of the form:

$$\sum_{i \in I} x_i = k.$$

If the randomized rounding step is performed independently for all variables, it is very likely that such constraints are violated. Srinivasan [16] uses this as motivation to introduce the problem of generating distributions on level sets with given marginals. A *level set* distribution is a distribution on $0, 1$ variables which is concentrated on the event that their sum is a fixed value k. The randomized rounding technique motivates the need to generate distribtuions on n $0 - 1$ variables such that:

P1 It is a level set distribution i.e. concentrated on the event $\sum_i X_i = k$ for some integer $k > 0$.

P2 The marginals are given: $P[X_i = 1] = p_i, i \in [n]$.

In addition, for applying standard tools of concentration of measure such as the CH bounds, we need that

P3 The distribution is negatively dependent.

Srinivasan [16] gave a simple process that generated such distributions and gave a proof a negative dependence property that is needed in his applications. This yields improved approximation algorithms for problems such as (a) low congestion multi-path routing (b) maximum coverage set cover (c) Group Steiner Tree problem. Gandhi, Khuller, Parthasarathy and Srinivasan[7] extended the method combining it with a pipage rounding technique due to Ageev and Sviridenko to develop new randomized rounding approaches to obtain fractional vectors on edge sets of bipartite graphs. This yielded richer random graph models for graphs with a given degree sequence, improved approximation algorithms for (a) throughput maximization in broadcast scheduling, (b) delay minimization in broadcast scheduling (c) capacitated vertex cover and (d) fair scheduling of jobs on unrelated parallel machines.

In this paper, we give the following results. First we give a very simple proof of a stronger negative dependence property of Srinivasan's distribution on level sets: namely we show that it satisfies negative association. Second, we give a proof of a (potentially) stronger negative depence condition, namely, conditional negative association via a non-trivial application of the Feder-Mihail Theorem [6]. This theorem can be used to establish negative dependence, somewhat paradoxically, via a notion of a variable of positive dependence. This concept could possibly play an important role in a theory of negative dependence, but at the moment it is not sufficiently well understood. A very basic and natural question in the general theory is: for which combinations of a function and a distribution does a variable of positive influence exist? Here we provide some first simple results related to this question and we give some general results and counter-examples that might shed more light on its role in a theory of negative dependence.

In § 2 we give the definitions of various concepts of negative dependence and positive influence. In § 3 we give general results and counter-examples on the existence of a variable of positive influence. In § 4, we establish negative and conditional negative assocation for various level set distributions using the Feder-Mihail Theorem. We give a simple proof of negative association for Srinivasans's level set distribution and then a non-trivial application of the Feder-Mihail Theorem to establish conditional negative association.

2 Concepts of Negative Dependence and Positive Influence

Intuitively, a set of variables is negatively dependent if, conditioned on a subset of the variables taking "high" values, it is more likely that the complementary subset of variables takes on "low" values. This notion can be formalised in a number of ways[3], out of which, we focus on the following two:

Definition 1 ((Conditional) Negative Association) *A set of random variables $X_1, \ldots X_n$ is* **negatively associated** *(NA) if for very two disjoint subsets $I, J \subseteq [n]$ of variables, and all non-decreasing functions f, g,*

$$E[f(X_i, i \in I)g(X_j, j \in J)] \leq E[f(X_i, i \in I)]E[g(X_j, j \in J)].$$

The variables are (CNA) conditionally negatively associated if conditioned on setting any subset of the variables to fixed values, the resulting conditional distribution on the remaing set of variables is NA.

These defintions are much stronger than the simple notion of **pairwise negative correlation**: $E[X_iX_j] \leq E[X_i]E[X_j]$ for distinct $i, j \in [n]$. The Feder-Mihail Theorem yields the surprising conclusion that this weak notion of negative dependence yields the very strong notion CNA whenever there is a variable of *positive influence*. Intuitively, a variable has a positive influence on a function if one expects the value of the function to increase if the value of the variable is increased, assuming that the other variables are chosen at random. To be more precise, positive influence is defined as follows [3]:

Definition 2 (Positive Influence) *Let $X_1, \ldots, X_n$ be real-valued random variables and μ be any probability distribution on $X_1, \ldots, X_n$. Let f be any real-valued function of $X_1, \ldots, X_n$. We say that X_j is a* **positive influence variable** *for (f, μ) if $E_\mu[f(X_1, \ldots, X_n)|X_j = t]$ is a non-decreasing function of t.*

Other notions of influence of variables on functions have been defined and used before. In particular, it is instructive to compare our definition of positive influence with that of influence in Bourgain et al [1]. The influence of a variable X_j on a function $f(X_1, \ldots, X_n)$ there is defined to be the probability that the univariate function of X_j that results by picking all the other X_i's at random is a non-constant.

For simplification of notation, when it is clear what distribution we are talking about or we are talking about an arbitrary fixed distribution, we will often omit μ from the subscript in the expectation.

Theorem 3 (Feder-Mihail) *Let $X_1, \ldots, X_n$ be random variable such that conditional distribution on a subset of the variables $X_i, i \in I$, for $I \subseteq [n]$ for any setting of the complementary subset of variables $X_i, i \notin I$ satisfies the following two conditions:*

Positive influence *For any monotone function f, there is a variable of positive influence.*

Pairwise Negative Correlation *Any two variables are negatively correlated.*

Then the variables $X_1, \ldots, X_n$ satisfy CNA (conditional negative association).

3 Positive Influence: General Results and Counter-Examples

Proposition 4 *A sufficient condition for the existence of a positive influence variable is*

$$cov(f, \sum_i X_i) \geq 0. \tag{3.1}$$

Proof. By linearity of expectation,

$$\begin{aligned} 0 &\leq \operatorname{cov}(f, \sum_i X_i) \\ &= \sum_i \operatorname{cov}(f, X_i) \end{aligned}$$

Hence there exists an $i \in [n]$ such that $\operatorname{cov}(f, X_i) \geq 0$, that is, i is a variable of positive influence. □

One of the most useful cases where condition (3.1) holds is distributions on level sets:

Corollary 5 *Suppose μ is concentrated on a fixed level $k \geq 0$ i.e. $\sum_i X_i = k$. Then there is a variable of positive influence.*

Another set of circumstances in which positive influenceholds for a pair (μ, f) is if either μ or f is symmetric.

Proposition 6 *Let f be symmetric i.e. $f(a_1, \ldots, a_n) = f(a_{\sigma(1)}, \ldots, a_{\sigma(n)})$ for any permutation σ. Then there is a variable of positive influence for any μ.*

Proof. Since f is symmetric and the X_i's are binary we have that f only depends on $\sum_i X_i$. Since f is increasing it follows that f is an increasing function of $\sum_i X_i$. Thus Chebyshev's inequality entails that $cov(f, \sum_i X_i) \geq 0$. Now apply the above proposition. □

Proposition 7 *Let μ be symmetric (exchangeable) i.e. $\mu(X_1 = a_1, \ldots, X_n = a_n) = \mu(X_1 = a_{\sigma(1)}, \ldots, X_n = a_{\sigma(n)})$ for any permutation σ. Then there is a variable of positive influence for any μ.*

Proof. The result will follow if it is established that the distribution of $(X_1, X_2, \ldots, X_n)$ given that $\sum_i X_i = k$ is stochastically increasing in k.

However the symmetry of μ implies that this conditional distribution is uniform over all k-subsets of $[n]$. This trivially implies the desired property. □

3.1 Counter-example: What one cannot hope for

One might begin thinking naively that if a function $f(X_1, \ldots, X_n)$ is non-decreasing then at least one of $X_1, \ldots, X_n$ must have a positive influence on the function f no matter what the probability distribution on $X_1, \ldots, X_n$ is. This is quickly seen to be false. We provide a counterexample here that shows that this is not true even for two 3-valued symmetric random variables and even if the functions are additionally restricted to be binary and symmetric.

Consider the following probability distribution on X, Y that take values in $\{0, 1, 2\}$ with,

$$\begin{aligned} \Pr[X = 0, Y = 2] &= p_2 \\ \Pr[X = 1, Y = 1] &= p_1 \\ \Pr[X = 2, Y = 0] &= p_2 \end{aligned}$$

where $p_1, p_2 > 0$ and $2p_2 + p_1 = 1$. Consider the non-decreasing binary symmetric function $f(X, Y)$ defined as follows:

$$f(X, Y) = \begin{cases} 1 & \text{if} X^2 + Y^2 > 3 \\ 0 & \text{otherwise} \end{cases}$$

Then neither X nor Y has a positive influence on f as $E[f(X, Y)|X = 0] = E[f(X, Y)|Y = 0] = 1$ whereas $E[f(X, Y)|X = 1] = E[f(X, Y)|Y = 1] = 0$.

3.2 Counter-example: Binary variables

Even when the variables are restricted to be binary, there exist probability distributions for which there are non-decreasing binary functions which don't have any positive influence variable. This is evidenced by the example below. Consider the distribution μ on six random variables

a_1	a_2	a_3	b_1	b_2	b_3	$\mu(X_1 = a_1, X_2 = a_2, X_3 = a_3, Y_1 = b_1, Y_2 = b_2, Y_3 = b_3)$
0	0	0	1	1	1	1/5
0	1	1	0	1	1	1/5
1	0	1	1	0	1	1/5
1	1	0	1	1	0	1/5
1	1	1	0	0	0	1/5

Consider the binary function $f(X_1, X_2, X_3, Y_1, Y_2, Y_3)$ defined as follows:

$$f(X_1,X_2,X_3,Y_1,Y_2,Y_3) = \begin{cases} 1 & \text{if } (X_1+X_2+X_3)^2 + (Y_1+Y_2+Y_3)^2 > 8 \\ 0 & \text{otherwise} \end{cases}$$

Then, for $i = 1,2,3$, $E[f|X_i = 0] = E[f|Y_i = 0] = \frac{1}{2}$ whereas $E[f|X_i = 1] = E[f|Y_i = 1] = \frac{1}{3}$. Hence, none of the variables is a positive influence variable for f under μ. Notice that neither μ nor f is symmetric.

4 Negative Dependence via the Feder-Mihail Theorem

We give four applications of the Feder-Mihail Theorem to distributions on binary random variables that are concentrated on level sets i.e. where the number of 1s is fixed. *Ipso facto* the positive influence property obtains for free. It remains to check the pairwise negative correlation property to establish CNA.

4.1 Example 1: Fermi-Dirac Statistics

A (n,k) Fermi-Dirac statistics is a distribution on n binary variables concentrated on the event that there are exactly k 1's. Conditioned on setting some of the variables, the resulting distribution is just another Fermi-Dirac distribution, and negative covariance holds for any Fermi-Dirac Distribution because of the simple log-concavity property of the binomial-coefficients: $\binom{n}{k-2}\binom{n}{k} \le \left(\binom{n}{k-1}\right)^2$. Thus we get CNA for Fermi-Dirac statistics.

4.2 Example 2: Conditional Poisson Sampling

Suppose we want to sample from a population of n individuals in such a way that the probability that individual $i \in [n]$ is included in the sample equals a prescribed value p_i. Assume the $\sum_i p_i = k$ for some $k \in [n-1]$.

In other words we want to produce a family $X_1, X_2, \ldots, X_n$ of 0/1 random variables such that $P(X_i = 1) = p_i$ for all $i \in [n]$, and $\sum_i X_i = k$.

Poisson sampling refers to the method of doing this by simply letting each invidual be included with its prescribed inclusion probability indepndently of the other individuals, i.e. taking the X_i's to be independent. However this will produce a sample of random size. If we insist on having exactly k individuals in our sample we can condition on the event $\sum_i X_i = k$. Then we get **conditional Poisson sampling** (CPS). CPS has many nice properties. E.g. it has maximal entropy in the class of all samling processes with fixed sample size k and the same inclusion probabilities, see [9]. The main drawback is that the conditional inclusion prababilities are in general not equal to the unconditional ones and in order to find out what unconditional inclusion probabilities to use one has to solve an overwhelming non-linear system of equations.

We now adopt the Feder-Mihail Theorem to show that CPS satisfies CNA. Note that a CPS sample conditioned on the value of some of the X_i's is another CPS sample with another sample size. Therefore, by the Feder-Mihail Theorem, all we need to show is pairwise negative correlations for CPS sampling, i.e. that for $i \neq j$, $P(X_i = 1|X_j = 1) \leq P(X_i = 1|X_j = 0)$. But the distributions of the X_l's given $X_j = 1$ and $X_j = 0$ respectively, are those of two CPS sample that are identical apart from that the sample size is $k - 1$ in the first case and k in the latter. Therefore it suffices to establish that a CPS sample is stochastically increasing in the sample size. There are many ways to do this; perhaps the most efficient (though not the most elementary) way is by a coupling argument that we will shortly describe here and that is described in more detail in [9].

First we note that the CPS distribution is the stationary distribution of the aperiodic irreducible Markov chain on all subsets of size k with the following behavior:

Let S_t denote the state of the chain at time t. Given S_t the state S_{t+1} is generated by (i) picking an individual i from S_t uniformly at random, (ii) picking an individual j from the whole population according to the probabilitis $\{\theta_l/c\}$ where $\theta_l = p_l/(1-p_l)$ and $c = \sum_{m=1}^{n} \theta_m$.

If $j \in S_t$, then let $S_{t+1} = S_t$ otherwise replace i with j, i.e. put $S_{t+1} = (S_t \setminus \{i\}) \cup \{j\}$. That this Markov chain indeed has the CPS sampling distribution given that the sample size is k is straightforward to verify using the detailed balance equations.

Second, we couple the Markov chain with a corresponding one, $\{S'_t\}$, for sample size $k - 1$ starting from a state $S'_0 \subset S_0$. Since the p_i's are the same for the second chain it is easily seen that the transitions of this chain can then be coupled to the ones of the $\{S_t\}$-chain in such a way that $S'_t \subset S_t$ for all t. Hence we also have the desired stochastic inequality for the two stationary distributions.

4.3 Example 3: The Uniform Spanning Tree

In the examples above, the conditional pairwise negative correlation property was rather easy to check. However, sometimes this can be quite non-trivial. One such eaxmple is the *uniform spanning tree* i.e. a tree T chosen uniformly at random from all possible spanning trees of a graph $G = (V, E)$. Let $X_e, e \in E$ be the indicator variables corresponding to whether an edge e is picked in T. Since a tree has exactly $|V| - 1$ edges, this is a level set distribution and the positive influence property comes for free. To establish (CNA), we need to verify the conditional pairwise negative correlation property. However, the only known proof of the conditional pairwise negative correlation property in this case is, very surprisingly, via a transfer-impedence theorem of Burton and Pemantle [2] that exploits an analogy with electrical networks.

5 Srinivasan's Sampling Process

Let us now look at what we shall call **Srinivasan's samling process** (SSP). SSP works by *fixing* the values of variables to 0 or 1 one by one. It does so by repeatedly applying the following rules to pairs of unfixed variables in some pre-defined order. As before, p_i denote the inclusion probability for individual i and note that $\sum_i p_i = k$, the desired sample size.

Initially start with each variable *unfixed*, and with value $X_i = p_i$. At each step, pick two unfixed variables X_i and X_j. There are now two cases:

$p_i + p_j \leq 1$: With probability $p_i/(p_i + p_j)$ let $X_j = 0$ and $X_i = p_i + p_j$. The variable X_j is *fixed*. With the complementary probability $p_j/(p_i + p_j)$ set $X_i = 0$ and $X_j = p_i + p_j$. In this case, the variable X_i is fixed.

$p_i + p_j > 1$: With probability $(1 - p_j)/(2 - p_i - p_j)$ let $X_i = 1$ and $X_j = (p_i + p_j) - 1$. The variable X_i is fixed. With the complementary probability $(1 - p_i)/(2 - p_i - p_j)$ put $X_j = 1$ and $X_i = (p_i + p_j) - 1$ and the variable X_j is fixed.

One of the main virtues of SSP is that it is extremely easy to describe and to implement and still, as we shall now see, it satisfies strong properties of negative dependence.

It should be noted that the exact probabilities specified above with which the probability mass is transfered from one variable to the other, are designed to make the marginal probability that i is included in the sample equal to

p_i. However this is not essential for the proof of negative dependence. For negative dependence purposes, we can view the SSP as carrying out a series of *matches* between a pair of individual unfixed variables X_i and X_j. Each such match m has a *winner* and a *loser* whose identities are determined by a 0/1 random variable Z_m whose distribution (independently of every other match) is given by some pre-determined rule involving the values of the contesting variables. Each such match results in a transfer of probability mass from the loser to the winner. The amount transferred depends on whether $p_i + p_j$ exceeds 1 as described above. If $p_i + p_j \leq 1$ we refer to a *good* match and otherwise to a *bad* match. Note that in a good match, the winner will be automatically qualified for a place in the final sample and the loser will go on playing for a place whereas in a bad match the loser will be once and for all ruled out from the sample and the winner will go on to play for a place. Let us call such a process setting the variables a *generalized Srinivasan Sampling Process*(GSSP).

Theorem 8 *The distribution produced by applying the GSSP to the variables in any pre-determined order produces a distribution satisfying NA.*

Proof. We use the conditional covariance formula: if X, Y and Z are three random variables, then

$$cov(X, Y) = \mathbf{E}[cov(X, Y|Z)] + cov(\mathbf{E}[X|Z], \mathbf{E}[Y|Z]).$$

We will use induction on n, the number of variables (i.e. the number of individuals in the population). First note that for $n = 2$ NA is trivial, so we can entirely focus on the induction step.

Let f and g be two increasing functions depending on disjoint sets of variables and let Z be the outcome of the match between the first two individuals, which we may without loss of generality call 1 and 2. i.e. let Z be the indicator that 1 wins that match.

By the conditional covariance formula

$$cov(f, g) = \mathbf{E}[cov(f, g|Z)] + cov(\mathbf{E}[f|Z], \mathbf{E}[g|Z]).$$

Given Z we face an GSSP on $n - 1$ individuals and so by the induction hypothesis the first term on the right is non-positive.

For the second term note that the distribution of $X_3, X_4, \ldots, X_n$ is independent of Z. Hence the second term will be 0 unless 1 and 2 are in the support of different functions so assume without loss of generality that 1 is in the support of f and 2 is in the support of g.

However then $\mathbf{E}[f|Z]$ is increasing in Z and $\mathbf{E}[g|Z]$ is decreasing in Z (here we also use that $X_3, \ldots, X_n$ are independent of Z) and so the second term is also non-positive. This establishes NA. $\square$

To establish CNA is less straightforward but we are lucky since the Feder-Mihail Theorem tells us that it is sufficient to prove conditional pairwise negative correlation. However we will still not be able to do this in full generality; we will need to restrict to the case of a *linear match schedule*, i.e. in the first match 1 plays 2, in the second math the variable from match 1 that remains unfixed plays 3, in the third match the variable that remains unfixed from match 2 plays 4 etc.

We need to prove that

$$cov(X_a, X_b | X_A = x_A) \leq 0$$

for any $A \subset [n]$, any $x_A \in \{0,1\}^A$ and any pair of individuals a and b ($a < b$). We will do this by showing that

$$P(X_b = 1 | X_a = 1, X_A = x_A) \leq P(X_b = 1 | X_a = 0, X_A = x_A).$$

Assume without loss of generality that $\{a, b\} \cap A = \emptyset$; the result is otherwise trivial.

Assign to each individual $i \in [n]$ for a given conditional SSP a *status*: we say that i has status 1 if it has been conditioned that $X_i = 1$, 0 if it has been conditioned that $X_i = 0$ and U (for "unknown") if i is not in the set of individuals conditioned on. We linearly order the set $\{1, 0, U\}$ by putting $0 \leq U \leq 1$.

Let us now start to compare the conditional SSP's where we on one hand condition on the event $\{X_a = 1, X_A = x_A\}$ and on the other on the event $\{X_a = 0, X_A = x_A\}$. For convenience we denote the first process "Process I" and the latter "Process II". We want to show that b has an at least as great chance to qualify for the sample in Process II as in Process I. We will do this by coupling the two processes so that as soon as b is included in the sample of Process I, then b is also included in the sample of Process II. Start be letting the two processes evolve independently for the first $a-1$ matches. Assume that match $a-1$ (i.e. the first match a plays) is bad. (The case when this match is good is analogously treated and is left to the reader.) In Process I where a has status 1, a must necessarily win this match. In particular this means that the opponent of $a+1$ in match number a will have status 1. (Conditioned on the opponent of $a+1$ in match a all that matters as far as far as the rest of the process is concerned is the status of the opponent of $a+1$.) In Process II a may or may not win match $a-1$, but in any case the opponent of $a+1$ in match a will trivially have a status less than or equal to 1, in particular less than or equal to the status of the opponent of $a+1$ in Process I. For $j = a+1, a+2, \ldots, b$, let V_j denote the status of the opponent of j in its first match, i.e. in match $j-1$, in Process I and let $W_{a+1}, W_{a+2}, \ldots, W_b$ denote the corresponding thing for Process II; we just showed that $V_{a+1} \geq W_{a+1}$.

We now claim that we can couple the two processes so that $V_j \geq W_j$ for all j. The claim is trivially true if $b = a+1$ so assume that $b > a+1$. We will prove the claim by showing that for $a+1 \leq j < b$, $V_j \geq W_j$ implies that we can couple so that $V_{j+1} \geq W_{j+1}$. If $V_j = W_j$ then we can take $V_{j+1} = W_{j+1}$ so we may assume that $V_j > W_j$, i.e. $(V_j, W_j) \in \{(U,0),(1,0),(1,U)\}$. We then face three cases depending on the status of j:

- j has status U. In this case $V_{j+1} \geq W_{j+1}$ whatever we do.
- j has status 0. Then the same goes when $W_j = 0$. When $V_j = 1$ and $W_j = U$, then either match $j-1$ is good in which case we will necessarily get $W_{j+1} = V_{j+1} = 0$ and if match $j-1$ is bad then we automatically get $V_{j+1} = 1 \geq W_{j+1}$.
- j has status 1. Here the analysis is analogous (but reversed): When $V_j = 1$ things work out whatever we do. When $W_j = 0$ and $V_j = U$ then if match $j-1$ is bad we necessarily get $V_{j+1} = W_{j+1} = 1$ and if match $j-1$ is good then we necessarily get $W_{j+1} = 0 \leq V_{j+1}$.

Now that we know that we can couple Process I and Process II so that $V_j \geq W_j$ for all j we know in particular that we will have $V_b \geq W_b$. Now, the conditional probability that b wins match $b-1$ given that the status, S, of its opponent is s is decreasing in s. To see this we may as well assume that the opponent of b is $b-1$. Let Z_b be the indicator that b wins the match, assume $s \in \{0,1\}$ and apply Baye's formula:

$$Pr(Z_b = 1 | X_{b-1} = s) = \frac{Pr(X_{b-1} = s | Z_b = 1)}{Pr(X_{b-1} = s)} Pr(Z_b = 1).$$

(Here Pr refers to the conditional distribution of $X_{b-1}, \ldots, X_n$ given $X_{A \setminus [b]} = x_{A \setminus [b]}$ which is the same for Process I as for Process II.) The ratio is trivially at most 1 when $s = 1$ and at least 1 when $s = 0$ and the claimed monotonicity follows as the case $s = U$ corresponds to not conditioning on X_{b-1} at all.

Thus we can couple the two processes so that b wins match $b-1$ for Process II as soon as the same thing happens for Process I. If match $b-1$ is bad this means that either b is in Process I ruled out from the sample whereas it still has a chance to qualify for the sample of Process II, in which case we simply let the two processes evolve independently for the remaining matches, b is ruled out from both samples, in which case we do the same thing, or b is still in the game for both processes in which case we can let b meet the same fate for the two processes. The case when match $b-1$ is good is analogous. We have shown:

Theorem 9 *The distribution produced by applying the GSSP to the variables in linear order produces a distribution satisfying CNA.*

6 Conclusion and Open Questions

The extra condition of linear match schedule was necessary for the proof of CNA for GSSP to work out, but we strongly believe that the result also holds for a general match schedule:

Conjecture 10 *GSSP satisfies CNA under any match schedule.*

It is clear that even though SSP has CNA it has some undesirably strong dependencies. E.g. if $p_1+p_2<1$ then 1 and 2 cannot both be in the resulting sample. One way to increase the entropy of the Process is of course to give the individuals a random order before starting the SSP. Let us call the resulting Process the Random Order Srinivasan Sampling Process (ROSSP). It is natural to believe that ROSSP also has CNA. However we do not know how to prove this.

All we can show here is pairwise negative correlations. The proof goes via the conditional covariance formula: Let R denote the random order of the individuals (i.e. R is an element of S_n, chosen uniformly at random). Then

$$cov(X_i, X_j) = \mathbf{E}[cov(X_i, X_j|R)] + cov(\mathbf{E}[X_i|R], \mathbf{E}[X_j|R]).$$

The first term i negative by CNA for SSP and the second term is 0 as the individual inclusion probabilities are not affected by R.

Note that in order to prove CNA it is by the Feder-Mihail Theorem sufficient to prove conditional pairwise negative correlations and one could easily belive that the same argument goes for that too. However the fact that we are conditioning means that the second term is no longer unaffected by R.

Conjecture 11 *ROSSP satisfies conditional negative association.*

The general problem of characterizing (f, ψ) pairs for which positive influence variables exist remains open. It will be interesting and useful if one could establish positive influence lemmas for other broad families (f, ψ) pairs.

References

[1] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson and N. Linial. "The influence of variables in product spaces", *Israel Journal of Mathematics* **77** (1992), 55–64.

[2] R. Burton and R. Pemantle, "Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances", *Ann. Probab.*21 no. 3, 1329–1371, 1993.

[3] D. Dubhashi and D. Ranjan, "Balls and Bins: A Study in Negative Dependence", *Random Structures and Algorithms*, 13 pp. 99 -124 1998.

[4] C.M. Fortuin and P.W. Kasteleyn and J. Ginibre, "Correlation inequalities on some partially ordered sets", *Communications in Mathematical Physics* **22**, 89–103 (1971).

[5] M.L. Eaton, "Lectures on Topics in Probability Inequalities", CWI Tract in Mathematics 35, Amsterdam 1987.

[6] T. Feder and M.Mihail, "Balanced Matroids", in *Proceedings of the 24th Annual STOC*, pp. 26–38, 1992.

[7] R. Gandhi, S. Khuller, S. Parthasarthy and A. Srinivasan, "Dependent Rounding in Bipartite Graphs", in *Proc. 43rd IEEE Symposium on Foundations of Computer Science (FOCS)*, 2002.

[8] K. Joag-Dev and F. Proschan (1983) "Negative association of random variables, with applications", *Annals of Statistics* **11** 286–295.

[9] J. Jonasson and O. Nerman (1997) "On maximum entropy πps-sampling with fixed sample size", Report, Chalmers University of Technology, 1997.

[10] T. Liggett, "The stochastic evolution of infinite systems of interacting particles", *lecture Notes in Mathematics* vol 598, Springer-Verlag, 1977.

[11] R, Lyons and Y. Peres, *Probability on Networks and Trees*, Book manuscript to appear in Cambridge University Press 2003, `http://www.math.gatech.edu/ rdlyons/prbtree/prbtree.html`

[12] C. Newman, "Asymptotic independence and limit theorems for positively and negatively associated random variables", in *Inequalities in Statistics and Probability*, edited by Y.L. Tong, I.M.S. Lecture notes-monograph series, vol 5, pp. 127–140.

[13] R. Pemantle, "Towards a Theory of Negative Dependence", *Journal of Mathematical Physics*, 41:3, pp. 1371 – 1390, 2000.

[14] S.C. Port. *Theoretical Probability for Applications*, John Wiley 1994.

[15] P. Raghavan and C.D. Thompson, "Randomized rounding: a technique for provably good algorithms and algorithmic proofs", *Combinatorica* 7, pp. 365–374, 1987.

[16] A. Srinivasan, "Distributions on level sets with applications to approximation algorithms", in *Proc. 42nd IEEE Symposium on Foundations of Computer Science (FOCS)*, 2001.