Research Statement

Ümit Işlak

My research is primarily concerned with probability theory and its interactions with other fields. In Sections 1 and 2 of my statement below, I discuss my contributions to the areas of concentration of measure inequalities and Stein's method, which are my main research interests during my PhD studies. Section 3 is devoted to a very brief discussion of three other research interests of mine which I enjoy working on and also plan to pursue further in my future career. These are Markov chain Monte Carlo methods, Markov chain theory and negative dependence of random variables.

1. Concentration of measure inequalities. My dissertation is on the concentration of measure inequalities. These inequalities, in a broad sense, explain the phenomenon that a function, with certain smoothness conditions, of a large number of independent random variables tends to concentrate its values in a relatively narrow range. Since the days when Markov proved his famous tail inequality, concentration of measure has been the focus of extensive research and turned out to be a valuable tool in many areas of mathematics and related fields. For an excellent book length treatment of concentration of measure, see [3].

First, for a motivating and historical example, let's consider the case where $X_1, X_2, ..., X_n$ are independent random variables with distribution $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$ and $S_n = \sum_{i=1}^n X_i$. For this simple case, it is well known by the strong law of large numbers that $S_n/n \to 0$ a.s. as $n \to \infty$, and it is natural to ask how rapid this convergence occurs. For instance, Bernstein's classical inequality answers this question by giving the bound $\mathbb{P}(|S_n/n| > t) \leq 2\exp(-nt^2/2(1+t/3))$, $t \geq 0$ which was later improved in several different ways. In a more general setup, we can similarly consider the concentration of a given random variable Y, or smooth functions of Y.

An important instance that often emerges in applications of concentration of measure inequalities is the case when Y is a function of dependent random variables. For such Y, Sourav Chatterjee introduced a novel approach in 2007 [4] with an inspiration from Stein's method that works out perfectly in various applications. First part of my results follows the work of Subhankar Ghosh and Larry Goldstein [11] who employed Chatterjee's approach by making use of size biased couplings which are very important for distributional approximations within dependent settings in Stein's method literature. We recall from [15] that for a non-negative, integrable random variable Y, another random variable Y^s is said to have the Y size biased distribution if $\mathbb{E}[Yf(Y)] = \mathbb{E}[Y]\mathbb{E}[f(Y^s)]$ for all functions f for which the expectations exist.

Letting $\mu = \mathbb{E}[Y] > 0$, and assuming the existence of a size biased coupling Y^s of Y satisfying $Y^s - Y \leq C$ for some constant $C > 0$, my first results were the following concentration of measure inequalities

$$\mathbb{P}(Y - \mu \leq -t) \leq \exp\left(-\frac{t^2}{2C\mu}\right) \text{ and } \mathbb{P}(Y - \mu \geq t) \leq \exp\left(\frac{1}{C}\left(t - (\mu + t)\log\left(\frac{\mu + t}{\mu}\right)\right)\right) \quad (0.1)$$

which hold for all $t \geq 0$.

These concentration bounds improve the results in [11] significantly for both tails. For the upper tail, the asymptotic order of the bound is improved to $\exp(-ct\log t)$ from $\exp(-ct)$. This yields the correct order [18], for example, for generalized matching problems in uniformly random permutations improving the corresponding results of Chatterjee in [4]. Also the upper tail inequality in (0.1) reveals the following Bernstein type bound as a corollary

$$\mathbb{P}(Y-\mu \geq t) \leq \exp\left(-\frac{t^2}{2C\mu + (2/3)Ct}\right)$$

which still has better constants than the ones given in [11].

For the lower tail, I was able to remove a monotonicity condition (that, $Y^s \geq Y$ a.s.) which was assumed as a hypothesis for the results of [11]. This is quite useful since in applications size biased coupling constructions need not be monotone as can be seen from various examples in [12]. In a recent paper [1], different proofs of my results were given where a slight improvement in the upper tail inequality was also provided.

As a further note, the boundness condition in the lower tail inequality of (0.1) is indeed relaxable to $\mathbb{E}[Y^s - Y|\mathbf{X}] \leq C$ when Y is a function of some underlying random variables $\mathbf{X} = (X_1, X_2, ..., X_n)$. In [18], I showed that this observation can be used to introduce a new type of dependence, which I call as the *layered dependence*, of random variables which helps one to understand functions of possibly globally dependent random variables. For instance, under this more general assumption one can obtain the Suen type correlation inequality $\mathbb{P}(Y=0) \leq \exp\left(-\frac{\mu}{2C}\right)$, that yields significant improvements to Svante Janson's results in [20] when the underlying random variables are strongly dependent.

In a subsequent work with Jay Bartroff and Larry Goldstein [2], we extended the scope of the applications of the results stated in previous paragraphs. Namely, this work provides a general framework for construction of bounded size biased couplings for counts based on multivariate occupancy models $\mathbf{M} = (M_\alpha)_{\alpha\in\mathcal{A}}$ with log concave marginals. For a flavor of our results, consider the Chung-Lu random graph $\mathcal{G}$ [8] where we have n vertices labeled as $1, 2, ..., n$ and a pair of vertices i and j are connected to each other with probability $p_{ij} = (w_i w_j / l_n) \wedge 1$, where w_i's are non-negative weights and $l_n = \sum_{i\in[n]} w_i$. Letting d_k's be fixed degrees of interest and M_k be the degree of vertex k in $\mathcal{G}$, our results in [2] can be used, for example, to obtain concentration bounds for $Y_{\geq} = \sum_{k=1}^n \mathbb{1}(M_k \geq d_k)$ using the inequalities in (0.1) with $C = 1 + \max_{k=1,...,n} d_k$. Several applications to other multivariate occupancy models such as multinomial allocation, random graphs, multivariate hypergeometric sampling and germ-grain models in stochastic geometry are also included in [2], and our results improve previous efforts in these directions in terms of generality and the constants in bounds.

In one other subsequent work with Subhankar Ghosh [13], we established multivariate analogues of the inequalities in (0.1) which are the first attempts in this direction that use couplings. Namely, under multivariate bounded coupling assumptions similar to the univariate case, we were able to prove upper bounds of orders $\exp(-\|\mathbf{t}\|_2^2)$ and $\exp(-\|\mathbf{t}\|_2)$ for the probabilities $\mathbb{P}(\mathbf{W} - \boldsymbol{\mu} \preceq -\mathbf{t})$ and $\mathbb{P}(\mathbf{W} - \boldsymbol{\mu} \succeq \mathbf{t})$, respectively, where $\mathbf{W} = (W_1, W_2, ..., W_k)$ is random vector with nonnegative coordinates, $\boldsymbol{\mu} = \mathbb{E}[\mathbf{W}]$ and $\mathbf{t} = (t_1, t_2, ..., t_k) \succeq \mathbf{0}$. Here $\|\cdot\|_2$ is the L^2 norm on $\mathbb{R}^k$, and the partial order $\preceq$ on $\mathbb{R}^k$ is the standard componentwise order. Although we believe that the order in the upper tail inequality can be improved as in the univariate case, we were not able to succeed in this yet

and we keep this as a problem for future research. We also have similar multivariate concentration bounds by making use of multidimensional exchangeable pairs and our current research is on forming a general framework for all of these results using the Stein coupling approach of [6].

As a final result on concentration of measure inequalities, in joint work with Larry Goldstein, we were able to show that one can also prove exponential tail bounds using zero biased couplings, another important coupling from Stein's method, in possibly dependent settings. Let's recall from [14] that for a mean zero random variable Y with positive, finite variance σ^2, a random variable Y^* is said to have Y-zero biased distribution if $\mathbb{E}[Yf(Y)] = \sigma^2\mathbb{E}[f'(Y^*)]$ for all absolutely continuous functions f for which the expectation of either side exists. It is well known that the distributional transformation $\mathcal{L}(Y) \to \mathcal{L}(Y^*)$ has the normal distribution as its unique fixed point and thus, under a bounded zero biased coupling assumption, it is natural to assume that the tail behavior of Y will be *close* to the tails of the normal distribution. One of our main results from [**?**] justifies this heuristics by establishing the inequality

$$\mathbb{P}(|Y| \geq t) \leq 2\exp\left(-\frac{t^2}{2(\sigma^2 + Ct)}\right), \quad t \geq 0 \tag{0.2}$$

whenever one can find a zero biased coupling of Y with $|Y^* - Y| \leq C$ for some $C \geq 0$. Later we were also able to show that the orders of the bounds given in (0.2) can actually be improved to $\exp(-ct\log t)$ under similar assumptions. Applications of this approach include the Hoeffding statistic $Y = \sum_{i=1}^n a_{i\pi(i)}$ where $A = (a_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$ and the permutation π has the uniform distribution over the symmetric group, and when its distribution is constant on cycle type. In particular, our bounds provide improvements to Chatterjee's corresponding results in [4] when π has uniform distribution. See [**?**] for details. We are now in the process of finding new applications, forming a multivariate framework and exploring connections to statistics of random matrices.

Finally, I would like to note that there are several questions I will pursue in the field of concentration of measure inequalities in the following years. Here is a short list.

- As noted above, the zero biased transformation $Y \to Y^*$ has the normal distribution as its unique fixed point and the statement $|Y^* - Y| \leq C$ for some constant C indicates that the tail behavior of Y should be *similar* to the tail behavior of a normal random variable in a certain sense. With this motivation, can we use other fixed point transformations from Stein's method literature to obtain concentration bounds? As another example, so far I have also obtained results with the equilibrium transformation which has the geometric distribution as its unique fixed point.
- Chen and Roellin introduced a class of couplings (known as Stein couplings) in [6] which generalizes various important coupling constructions from Stein's method literature such as size biasing and exchangeable pairs. Can we form a general framework in a multidimensional setting for the previous efforts using the constructions of Chen and Roellin?
- Can we prove Talagrand's inequality by using the coupling approach discussed above? Now there is some work in the literature done in this direction, however to the best of my knowledge a convincing answer has not been given yet.
- Proofs of the technical results stated above involve differential inequalities and the use of Gronwall type estimates from differential equations literature. There are several other coupling approaches that result in differential inequalities with time scales. Can one use similar estimates for obtaining useful concentration bounds for such cases?

2. Stein's method. My secondary main research interest during my PhD studies was on Stein's method. I already discussed the use of couplings from Stein's method for concentration of measure inequalities in the previous section. Here our focus will be on distributional approximations for random variables [23]. In this aspect, the method refers to a general framework for bounding the distance between the distribution of a random variable X and that of a random variable Z having some specified target distribution. The metrics for which this approach is applicable are of the form $d_{\mathcal{H}}(\mathcal{L}(X), \mathcal{L}(Z)) = \sup_{h\in\mathcal{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Z)]|$ for some suitable class of functions $\mathcal{H}$. Note that Kolmogorov, total variation and Wasserstein distances are all included as special cases of $d_{\mathcal{H}}$.

The first step in distributional approximations with Stein's method is to find an operator $\mathcal{A}$ such that $\mathbb{E}[(\mathcal{A}f)(X)] = 0$ for all f belonging to some sufficiently large class of functions $\mathcal{F}$ if and only if $\mathcal{L}(X) = \mathcal{L}(Z)$. Once we have found such an operator $\mathcal{A}$, as a second step, for given $h \in \mathcal{H}$ one attempts to solve the equation $(\mathcal{A}f)(x) = h(x) - \mathbb{E}[h(Z)]$. If for each $h \in \mathcal{H}$, a solution $f_h \in \mathcal{F}$ exists, then upon taking expectations, absolute values, and supremum, one finds that $d_{\mathcal{H}}(\mathcal{L}(X), \mathcal{L}(Z)) = \sup_{h\in H} |\mathbb{E}[h(X)] - \mathbb{E}[h(Z)]| = \sup_{h\in H} |\mathbb{E}[(\mathcal{A}f_h)(X)]|$. Surprisingly, it is often easier to work with the right-hand side of this equation and the techniques for analyzing distances between probability distributions in this regard are known as Stein's method. As an example of a characterizing operator $\mathcal{A}$, we recall from [23] that $X \sim N(0,1)$ if and only if $\mathbb{E}[Xf(X) - f'(X)] = 0$ for all absolutely continuous functions f with $\|f\|_\infty < \infty$. Also we note that the method has been used to attain rates of convergence for a variety of distributions other than normal, including the Poisson, exponential and geometric. For excellent surveys of the subject, see [5] and [22].

My involvement with Stein's method began by studying applications to random permutation statistics. My first result [19] was answering a question of Jason Fulman in [10] by proving the asymptotic normality of the number of inversions in a random permutation $\rho \in S_n$, that is $inv(\rho) = \sum_{i<j} \mathbb{1}(\rho(i) > \rho(j))$, where the permutation results from a biased riffle shuffle. Proof of this theorem makes use of inverse shuffles, and one also obtains a convergence rate of order $1/\sqrt{n}$ in the Kolmogorov distance. During the course of analysis of this problem, I have become very interested in non-parametric statistical tests, and kept working on various other statistics related to extreme points of random sequences throughout my PhD studies.

A more recent work of mine in this direction is on the length of longest alternating subsequences in uniformly random permutations and (possibly non-uniform) random words on finite alphabets. Due to a characterization [16] of longest alternating subsequences in terms of local extremes, obtaining convergence rates for the case of uniformly random permutations is straightforward using standard results on locally dependent random variables. Same problem for random words turns out to be rather hard as one may have repeated values in a random word causing a necessary change in the definitions of local extremes, thus making the dependence of the underlying random variables global. However, I was still able to prove asymptotic normality with convergence rates by using the right coupling approach. My current research is on studying statistics of multisets having symbols with fixed multiplicities (which is equivalent to riffle shuffles with fixed cuts). I am also planning to explore the connections of this latter study to problems related to DNA sequences in the future.

A second approach I focused in the area of Stein's method is the concentration inequality approach of Chen and Shao [7]. One of my results [17] here is a proof of the asymptotic normality of $\sum_{i=1}^{N_n} X_i$ where $X_1, ..., X_n$ are locally dependent random variables with certain assumptions and N_n is another non-negative integer valued random variable independent of X_i's. This result generalizes the work done by [7] for sums of independent random variables and I believe that it will be useful

for interesting non-parametric statistical tests in future research. One other question to follow in this direction will be extending my result on random sums to the case where N_n and X_i's are not independent which is very important as in many applications N_n's are actually stopping times.

There are various other problems I will be working on in the following years some of which are:

- Can we use Stein's method to study convergence rates for Markov chains? Persi Diaconis' work in [9] where he discusses a very simple example provides a good starting point for this study.
- The bounded coupling constructions we gave in [2] raised the possibility of studying distributional approximations for various interesting occupancy statistics. Is it possible to provide a unifying framework in this direction by using our constructions which simplifies and improves various previous efforts?
- Can we develop a Stein's method framework for Tracy-Widom distribution which appears in limits of important statistics in random permutations, random matrix theory and percolation?

3. Other research interests. In this section I will describe three of my other research interests which I enjoy working on, and very briefly will discuss my ongoing progress.

3.1 MCMC - Perfect sampling. My master's thesis was on the Propp-Wilson algorithm and its generalizations which are used to obtain exact samples from stationary distributions of certain Markov chains. I was able to show that these methods can be considered as corollaries of Letac's principle and analyzed how backward processes can be used to obtain convergence rates in Markovian processes. There are various drawbacks of the algorithms in the lines of Propp and Wilson's work such as the user impatience bias, correct choice of the iterative function system and uniform ergodicity requirement. My current research in this direction is on designing alternative algorithms that can be used in more general settings.

3.2 Markov chains. My research is closely related to Markov chain theory as described in the previous sections. Here I will briefly mention two other problems I am planning to work on in the following years. (1) As discussed above, I am planning to adapt Stein's method techniques for analysis of statistics driven by Markov chains. As an exemplary special case, I am interested in obtaining bounds to normality for finite or countable Markovian random words statistics, like the lengths of longest alternating subsequences. Asymptotic normality in this latter problem was shown in [16], but it would still be very interesting to obtain convergence rates as this will require one to deal with a global dependence situation which may provide a route to understand several other problems. (2) My second main interest in Markov chains lies in shuffling chains. My research so far is focused on mostly riffle shuffles and riffle shuffles with fixed cuts, but I am planning to extend my research by working on other different shuffling models and plenty of cut-off problems that are unsolved.

3.3 Negative dependence. I became interested in negative dependence [21] of random variables due to its close connection to my research in concentration of measure inequalities and Stein's method. There are several problems to attack in this field for the following years, most of which are related to the connections between log concavity, negative dependence and couplings from Stein's method. Besides these general considerations, so far I also worked on understanding weaker formulations of negative association, and finding applications of these different classes. I am planning to extend my research in these directions in the near future.

References

[1] Arratia, R. and Baxendale, P., *Bounded size bias coupling: a Gamma function bound, and universal Dickman-function behavior*, 2013, preprint, http://arxiv.org/abs/1306.0157.

[2] Bartroff, J., Goldstein, L., and Işlak, Ü., *Bounded size biased couplings for log concave distributions and concentration of measure for occupancy models*, 2013, preprint.

[3] Boucheron, S., Lugosi, G. and Massart, P., Concentration inequalities: A nonasymptotic theory of independence, Oxford University Press, 2013.

[4] Chatterjee, S., *Stein's method for concentration inequalities*, Probab. Theory Related Fields, 138, pp. 305-321, 2007.

[5] Chen, L. H. Y., Goldstein L. and Shao, Q. M., Normal Approximation by Stein's Method, Springer, New York, 2010.

[6] Chen, L. H. Y. and Roellin, A., *Stein couplings for normal approximation*, 2010, preprint.

[7] Chen, L. H. Y. and Shao, Q. M., *Normal approximation for nonlinear statistics using a concentration inequality approach*, Bernoulli 13, no. 2, 581-599, 2007.

[8] Chung, F. and Lu, L., Complex graphs and networks, American Mathematical Society, Providence, RI, 2006.

[9] Diaconis, P., *Stein's Method for Markov Chains: First Examples*, Stein's Method: Expository Lectures and Applications, 27-43, 2004.

[10] Fulman, J., *The combinatorics of biased riffle shuffles*, Combinatorica 18, no.2, 173-184, 1998.

[11] Ghosh, S. and Goldstein, L., *Concentration of measures via size biased couplings*, Probability Theory and Related Fields, 149, pp. 271-278, 2011.

[12] Ghosh, S., and Goldstein, L., *Applications of size biased couplings for concentration of measures*, Electronic Communications in Probability, 16, pp. 70-83, 2011.

[13] Ghosh, S. and Işlak, Ü., *Multivariate concentration inequalities with size biased couplings*, 2013, preprint, http://arxiv.org/abs/1310.5448.

[14] Goldstein, L. and Reinert, G., *Stein's method and the zero bias transformation with application to simple random sampling*, Ann. Appl. Probab., 7, pp. 935-952, 1997.

[15] Goldstein, L. and Rinott, Y., *Multivariate normal approximations by Stein's method and size bias couplings*, Journal of Applied Probability, vol 33, pp. 1-17, 1996.

[16] Houdre, C. and Restrepo, R., *A probabilistic approach to the asymptotics of the longest alternating subsequence*, Electron. J. Combin. 17, no1., 2010.

[17] Işlak, Ü., *Asymptotic behavior of random sums of locally dependent random variables*, 2013, preprint, http://arxiv.org/abs/1303.2386.

[18] Işlak, Ü., *Size biasing, layered dependence and correlation inequalities*, 2013, preprint.

[19] Işlak, Ü., *Descent-Inversion statistics in riffle shuffles*, 2013, preprint, arxiv.org/abs/1303.2387.

[20] Janson, S., *New versions of Suen's correlation inequality*, Random Structures Algorithms 13, no. 3-4, 467-483, 1998.

[21] Pemantle, Robin, *Towards a theory of negative dependence*, J. Math. Phys. 41, no. 3, 1371-1390, 2000.

[22] Ross, N., *Fundamentals of Stein's method*, Probability Surveys 8, 210-293, 2011.

[23] Stein, C., *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proc. Sixth Berkeley Symp. Math. Stat. Prob., pp. 583-602, 1972.