Limit theorems for fixed points of Ewens random permutations

Oleksii Galganov
based on recent joint work
with Andrii Ilienko

National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»,
Institute for Applied System Analysis,
Ukraine, Kyiv

The Skorokhod readings, 7 June 2022

Outline

Ewens permutations

Consider the following probability measure on symmetric group S_n:

$$\mathbb{P}(\{\pi\}) = \frac{\theta^{\mathrm{c}(\pi)}}{\theta_{(n)}}, \ \pi \in \mathrm{S}_n, \tag{1}$$

where $\theta > 0$ is a fixed parameter, $\mathrm{c}(\pi)$ denotes the number of cycles in π and $\theta_{(n)} = \theta(\theta+1)\ldots(\theta+n-1)$.

(1) is known as the *Ewens distribution* on S_n. For $\theta = 1$, it reduces to the uniform distribution, which assigns probability $\frac{1}{n!}$ to each $\pi \in \mathrm{S}_n$.

Ewens permutations

Let σ be distributed as (1), and $\mathrm{c}_j(\sigma)$ stands for the number of cycles of length j in σ. For $c_1, ..., c_n \in \mathbb{N}_0$ with $\sum_{j=1}^{n} jc_j = n$ formula (1) results in

$$\mathbb{P}\left(\mathrm{c}_1(\sigma) = c_1, ..., \mathrm{c}_n(\sigma) = c_n\right) = \frac{n!}{\theta_{(n)}} \prod_{j=1}^{n} \left(\frac{\theta}{j}\right)^{c_j} \frac{1}{c_j!}. \tag{2}$$

(2) is known as the celebrated *Ewens sampling formula* (Ewens, 1972) which originally came from population genetics and subsequently found numerous applications in combinatorial stochastic processes, Bayesian nonparametrics, inductive inference, etc. (Crane, 2016)

Ewens permutations

As is customary in the literature, we will in what follows denote the Ewens distribution (1) by $\mathrm{ESF}(n, \theta)$ and write $\sigma \sim \mathrm{ESF}(n, \theta)$.

There is a classic result (Arratia, Barbour, Tavaré, 1992) about limiting distribution of cycle counts for Ewens permutation:

$$(\mathrm{c}_1(\sigma_n), \mathrm{c}_2(\sigma_n), \mathrm{c}_3(\sigma_n), \ldots) \xrightarrow{d} (X_1, X_2, X_3, \ldots), \ n \to \infty, \tag{3}$$

where $(X_j, j \geq 1)$ are independent $\mathrm{Pois}(\theta/j)$ random variables.

We will focus on the limiting behavior of fixed points, the total number of which is, accordingly to (3), asymptotically $\mathrm{Pois}(\theta)$.

Vague convergence to Poisson process

Let's define a sequence of point processes $(P_n, n \geq 1)$ on $[0,1]$ by

$$P_n = \sum_{i=1}^{n} \delta_{\frac{i}{n}} \mathbb{1} \{\sigma_n(i) = i\}, \tag{4}$$

where $\sigma_n \sim \mathrm{ESF}(n, \theta)$, and δ_x denotes Dirac measure at x.

Specifically, $P_n([0,1]) = c_1(\sigma_n)$ is the number of fixed points of σ_n.

In which sense can P_n converge and what can limit be equal to?

Vague convergence to Poisson process

Definition (vague topology)

Let $(\mu_n, n \geq 1)$ be a sequence of measures on a metric space E. This sequence *converges vaguely* to the measure ξ if $\int_E f \mathrm{d}\mu_n \to \int_E f \mathrm{d}\mu$ for each non-negative continuous compactly supported test function f.

Definition (vague convergence in distribution)

Let $(\xi_n, n \geq 1)$ be a sequence of point processes on a metric space E. This sequence *vaguely converges in distribution* to the point process ξ if $\mathbb{E}\varphi(\xi_n) \to \mathbb{E}\varphi(\xi)$ for each bounded and continuous w.r.t. vague topology $\varphi : M_p(E) \to \mathbb{R}$, which is denoted by $\xi_n \xrightarrow{vd} \xi$.

Vague convergence to Poisson process

In case of $E = [0, 1]$, vague convergence in distribution is closely related to the distributional convergence of cumulative processes in the Skorokhod J_1-topology (see, e.g., Kallenberg, 2017).

Theorem

Let $(\xi_n, n \geq 1)$ be a sequence of point processes on $[0, 1]$. Then $\left(\xi_n([0, \cdot\,]) \xrightarrow{J_1, d} \xi([0, \cdot\,])\right) \Rightarrow \left(\xi_n \xrightarrow{vd} \xi\right)$ and two statement are equivalent, if $\xi(\{0\}) = 0\ a.s.$ and ξ is simple, i.e. $\mathbb{P}\left(\forall\, x \in [0, 1] : \xi(\{x\}) \leq 1\right) = 1$.

Vague convergence to Poisson process

Since the limit in our core result is the homogeneous Poisson point process, we first recall the corresponding definition.

Definition

Let $\lambda > 0$, and $\mathcal{E}$ and Leb denote the Borel sigma-algebra on $E \subset \mathbb{R}^d$ and the Lebesgue measure on E, respectively. The *homogeneous Poisson point process with rate* λ is a point process N defined by

1. For all $F \in \mathcal{E}$, $N(F) \sim \mathrm{Pois}\,(\lambda \cdot \mathrm{Leb}(F))$ if $\mathrm{Leb}(F) < \infty$ and $N(F) = \infty$ $a.s.$ otherwise.
2. $(N(F_i), 1 \leq i \leq k)$ are independent for mutually disjoint sets $F_1, ..., F_k \in \mathcal{E}$.

Vague convergence to Poisson process

Theorem (convergence of P_n)

The sequence of point processes $(P_n, n \geq 1)$ defined by (4) vaguely converges in distribution to the homogeneous Poisson point process N with rate θ: $P_n \xrightarrow{vd} N$.

As N satisfies all equivalence conditions of the preceding theorem, then also $P_n \xrightarrow{J_1,d} N$ holds.

It follows from the well-known property of the homogeneous Poisson process that, as $n \to \infty$, unordered (normalized by n) positions of fixed points become *asymptotically independent* and uniformly distributed conditionally on their total number.

Vague convergence to Poisson process

Before we briefly describe the idea of the proof, let's take a look at a vd-convergence criterion. (Kallenberg, 2017)

Theorem

Let $(\xi_n, n \geq 1)$ *and* ξ *be point processes on a metric space* E *with Borel sigma-algebra* $\mathcal{E}$, *where* ξ *is simple, and fix a dissecting ring* $\mathcal{U} \subset \hat{\mathcal{E}}_\xi$, *where* $\hat{\mathcal{E}}_\xi = \{B \in \mathcal{E} : \xi(\partial B) = 0 \ a.s.\}$, *and a dissecting semi-ring* $\mathcal{I} \subset \mathcal{U}$. *Then* $\xi_n \xrightarrow{vd} \xi$ *iff*

1. $\lim\limits_{n\to\infty} \mathbb{P}\left(\xi_n(U) = 0\right) = \mathbb{P}\left(\xi(U) = 0\right)$ *for* $U \in \mathcal{U}$;
2. $\limsup\limits_{n\to\infty} \mathbb{P}\left(\xi_n(I) > 1\right) \leq \mathbb{P}\left(\xi(I) > 1\right)$ *for* $I \in \mathcal{I}$.

For our case, $\mathcal{U} = \mathcal{I}$ to be a family of disjoint unions of $\langle \alpha, \beta \rangle \subset [0, 1]$ will be enough, $\langle$ being either (or [.

Vague convergence to Poisson process

The proof is based on establishing an explicit formula for $\mathbb{P}\left(P_n(I) = k\right)$, where $I = \bigvee_{j=0}^{m} \langle \alpha_j, \beta_j \rangle \subset [0,1]$ and showing that $\mathbb{P}\left(P_n(I) = k\right) \to \mathbb{P}\left(N(I) = k\right)$.

For the sake of simplicity, we restrict ourselves here only to the case $I = [0, \gamma]$. A combinatorial argument based on the inclusion-exclusion principle leads to

$$\mathbb{P}\left(P_n([0,\gamma]) = k\right) = C^k_{\lceil \gamma n \rceil} \sum_{i=0}^{\lceil \gamma n \rceil - k} (-1)^i C^i_{\lceil \gamma n \rceil - k} \frac{\theta^{i+k}\theta_{(n-k-i)}}{\theta_{(n)}} =$$

$$= \frac{\theta^k}{k!} \sum_{i=0}^{\lceil \gamma n \rceil - k} (-1)^i \frac{\theta^i}{i!} \frac{(\lceil \gamma n \rceil)! \cdot \theta_{(n-k-i)}}{(\lceil \gamma n \rceil - k - i)! \cdot \theta_{(n)}}.$$

Then, after routine calculations and careful asymptotic analysis we come to $\mathbb{P}\left(P_n(I) = k\right) \to \frac{(\gamma\theta)^k}{k!} e^{-\gamma\theta} = \mathbb{P}\left(N(I) = k\right)$. Since both conditions for Kallenberg's vd-convergence criterion are fulfilled, it proves that $P_n \xrightarrow{vd} N$ as $n \to \infty$.

Limit theorems for statistics

Even though $P_n \xrightarrow{vd} N$ itself is an interesting result, it becomes even more fruitful if we consider the continuous mapping theorem.

In our case, it can shortly be formulated as following:

Theorem

If $\xi_n \xrightarrow{vd} \xi$ and $\varphi : M_p([0,1]) \to \mathbb{R}$ s.t. $\xi \in C_\varphi$ $a.s.$, then $\varphi(\xi_n) \xrightarrow{d} \varphi(\xi)$.

Thus, we can obtain a limit theorem for any statistics of fixed points, provided that it can be represented as a continuous function w.r.t vague topology.

It's worth recalling a property of N: positions of its points are uniform and independent conditioned on their number. In what follows, we will refer to this property as *CUI*.

Limit theorems for statistics: min and max fixed points

Theorem

Let m_n and M_n denote the smallest and the largest fixed points of $\sigma_n \sim \mathrm{ESF}\,(n, \theta)$, respectively, with $m_n = n + 1$ and $M_n = 0$ in case there are no fixed points at all. Then $\frac{m_n}{n} \xrightarrow{d} m$ and $\frac{M_n}{n} \xrightarrow{d} M$, with m and M having CDFs

$$F_m(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\theta x}, & 0 \le x < 1, \\ 1, & x \ge 1, \end{cases} \quad F_M(x) = \begin{cases} 0, & x < 0, \\ e^{\theta(x-1)}, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

Furthermore, $\frac{\mathbb{E}m_n}{n} \to \mathbb{E}m = \frac{1-e^{-\theta}}{\theta}$ and $\frac{\mathbb{E}M_n}{n} \to \mathbb{E}M = \frac{e^{-\theta}+\theta-1}{\theta}$.

This result can be proven using CUI and explicit distributions of the first and last order statistics of $\mathrm{U}\,(0, 1)$.

Limit theorems for statistics: min and max spacings

Let's consider the spacings between fixed points of σ_n. We also introduce the initial and final spacings by placing to virtual fixed points at 0 and $n+1$. Recalling CUI, we can study the distribution of spacings for $\mathrm{U}(0,1)$.

Let $S_i^{[n+1]}, i = 1, \ldots, n+1$ denote the spacings between n i.i.d. $\mathrm{U}(0,1)$ variables, including those starting at 0 and ending at 1. It is known (e.g., Rényi, 1953 or Holst, 1980) that for i.i.d. $X_i \sim \mathrm{Exp}(1), i = 1, \ldots, n+1$,

$$\left(S_1^{[n+1]}, S_2^{[n+1]}, \ldots, S_{n+1}^{[n+1]}\right) \stackrel{d}{=} \left(\frac{X_1}{\sum_{i=1}^{n+1} X_i}, \frac{X_2}{\sum_{i=1}^{n+1} X_i}, \ldots, \frac{X_{n+1}}{\sum_{i=1}^{n+1} X_i}\right),$$

$$\left(S_{(1)}^{[n+1]}, S_{(2)}^{[n+1]}, \ldots, S_{(n+1)}^{[n+1]}\right) \stackrel{d}{=} \left(\frac{X_{(1)}}{\sum_{i=1}^{n+1} X_i}, \frac{X_{(2)}}{\sum_{i=1}^{n+1} X_i}, \ldots, \frac{X_{(n+1)}}{\sum_{i=1}^{n+1} X_i}\right),$$

Limit theorems for statistics: min and max spacings

$$X_{(i)} \overset{d}{=} \frac{X_{n+1}}{n+1} + \frac{X_n}{n} + \cdots + \frac{X_{n-i+2}}{n-i+2} = \sum_{k=0}^{i-1} \frac{X_{n+1-k}}{n+1-k}.$$

One can also prove the following equalities:

$$S_{(i)}^{[n+1]} \overset{d}{=} \frac{\frac{X_{n+1}}{n+1} + \frac{X_n}{n} + \cdots + \frac{X_{n-i+2}}{n-i+2}}{\sum_{j=1}^{n+1} X_j} = \frac{\sum_{k=0}^{i-1} \frac{X_{n+1-k}}{n+1-k}}{\sum_{j=1}^{n+1} X_j}, \ i = 1, \ldots, n+1.$$

Together with CUI, equalities $S_{(1)}^{[n+1]} \overset{d}{=} \frac{X_{n+1}}{(n+1)\sum_{i=1}^{n+1} X_i}$ and $S_{(n+1)}^{[n+1]} = \frac{\sum_{i=1}^{n+1} \frac{X_i}{i}}{\sum_{i=1}^{n+1} X_i}$ lead to the following result:

Limit theorems for statistics: min and max spacings

Theorem

Let δ_n and Δ_n denote the smallest and the largest spacings between fixed points of $\sigma_n \sim \mathrm{ESF}(n, \theta)$, respectively. Then $\frac{\delta_n}{n} \xrightarrow{d} \delta$ and $\frac{\Delta_n}{n} \xrightarrow{d} \Delta$ with

$$\delta \overset{d}{=} \frac{X_{\nu+1}}{(\nu+1)\sum_{i=1}^{\nu+1} X_i}, \quad \Delta \overset{d}{=} \frac{\sum_{i=1}^{\nu+1} \frac{X_i}{i}}{\sum_{i=1}^{\nu+1} X_i},$$

where $(X_i, i \geq 1)$ are independent $\mathrm{Exp}(1)$ variables and $\nu \sim \mathrm{Pois}(\theta)$ is independent of $(X_i, i \geq 1)$.
Furthermore, $\frac{\mathbb{E}\delta_n}{n} \to \mathbb{E}\delta = \frac{e^{-\theta}}{\theta} \int_0^\theta \frac{e^t - 1}{t} \mathrm{d}t$ and $\frac{\mathbb{E}\Delta_n}{n} \to \mathbb{E}\Delta = \frac{1}{\theta} \int_0^\theta \frac{1 - e^{-t}}{t} \mathrm{d}t$.

Limit theorems for statistics: sum of fixed points

Theorem

Let S_n *denote the sum of the fixed points of* $\sigma_n \sim \mathrm{ESF}\,(n,\theta)$*. Then* $\frac{S_n}{n} \xrightarrow{d} S$ *with*

$$F_S(x) = e^{-\theta} \sum_{k=0}^{\lfloor x \rfloor} (-1)^k \frac{1}{k!} \left(\theta(x-k)\right)^{\frac{k}{2}} I_k\left(2\sqrt{\theta(x-k)}\right),$$

and Laplace transform given by

$$\mathbb{E}e^{-pS} = \exp\left\{-\theta\left(1 + \frac{1}{p}(e^{-p} - 1)\right)\right\}.$$

Here $I_k(z) = \left(\frac{1}{2}z\right)^k \sum_{m=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^m}{m!(k+m)!} = \frac{1}{\pi}\int_0^\pi e^{z\cos t}\cos{(kt)}\mathrm{d}t$ is the modified Bessel function of the first kind.

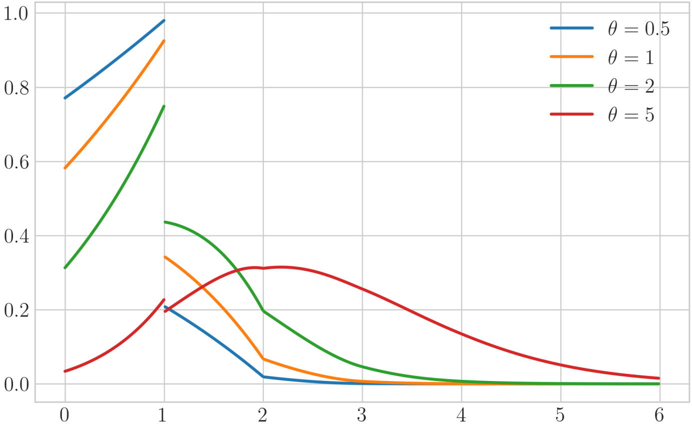
Limit theorems for statistics: sum of fixed points

In order to get the explicit distribution of S, we use CUI with some rather involved calculus, and Laplace functional of N (e.g., Resnick, 2008). The density of its (normalized) absolutely continuous component $\widehat{S}$ is $f_{\widehat{S}}(x) = \frac{\theta e^{-\theta}}{1-e^{-\theta}} \sum_{k=0}^{\lfloor x \rfloor} \frac{(-1)^k}{k!} f_k(x)$, where

$$f_k(x) = (\theta(x-k))^{\frac{k}{2}-1} \left(k I_k \left(2\sqrt{\theta(x-k)} \right) + \sqrt{\theta(x-k)} I_{k+1} \left(2\sqrt{\theta(x-k)} \right) \right).$$

$\widehat{S}$ can be considered as the distributional limit of the sum of fixed points conditioned on the event that there is at least one such point.

Limit theorems for statistics: sum of fixed points



Plots of $f_{\widehat{S}}(x)$ for different values of θ.

Modelling methods overview

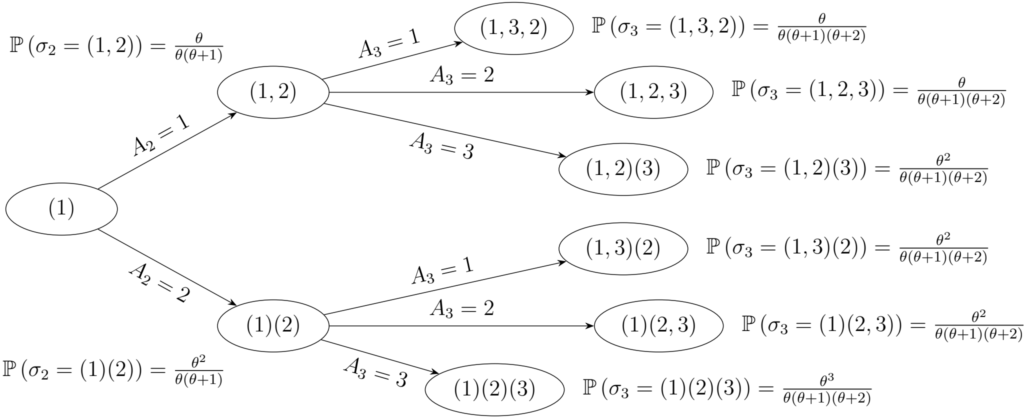
As it is computationally inefficient to sample Ewens permutations directly from S_n, modelling can be done by means of various couplings, which can also bring some useful properties. (Arratia, Barbour, Tavaré, 1992)

1. *Chinese restaurant process (CRP)* is used to generate a sequence $(\sigma_i, i \geq 1)$ of $\sigma_i \sim \mathrm{ESF}\,(i, \theta)$ on a common probability space. Let $(A_i, i \geq 1)$ be a sequence of independent random variables defined by

$$\mathbb{P}\,(A_i = j) = \begin{cases} \frac{\theta}{\theta+i-1}, & j = i, \\ \frac{1}{\theta+i-1}, & j = 1, 2, ..., i-1. \end{cases}$$

   First cycle starts with 1. As the first $n-1$ integers have been assigned to cycles, n starts a new cycle with probability $\mathbb{P}\,(A_n = n)$, or place to the right of j with probability $\mathbb{P}\,(A_n = j)$.

Modelling methods overview



An example of how CRP generates Ewens permutations for $n = 3$.

Modelling methods overview

2. *Feller coupling* only allows to construct σ_n for fixed n. Let $B_1, B_2, ..., B_n$ be independent Bernoulli random variables defined by

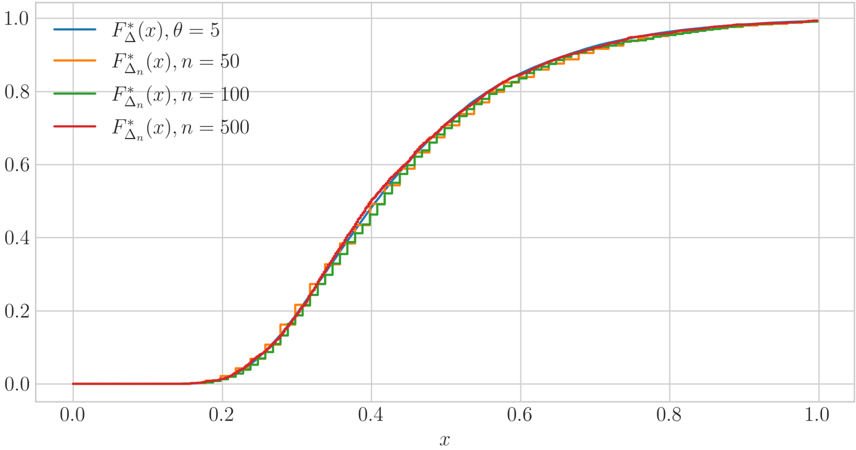
$$\mathbb{P}\left(B_i = 1\right) = \frac{\theta}{\theta + i - 1},\ i = 1, 2, ..., n,$$

so $B_i = \mathbb{1}\{A_i = i\}$ for A_i from CRP. By construction,

$$\mathrm{c}_j(\sigma_n) = \sum_{i=1}^{n-j} B_i(1 - B_{i+1})...(1 - B_{i+j-1})B_{i+j} + B_{n-j+1}(1 - B_{n-j+2})...(1 - B_n),$$

which, for example, can be used to establish bounds for Wasserstein distance between $\mathrm{c}_j(\sigma_n)$ and limiting Poisson variable. (Arratia, Barbour, Tavaré, 1992)

Modelling methods overview



Plots of CDFs of Δ and $\frac{\Delta_n}{n}$ modelled with CRP for $\theta = 5$.

Perspectives

1. *Related results for short cycles.* As it was already mentioned, $\mathrm{c}_j(\sigma_n) \xrightarrow{d} X_j \sim \mathrm{Pois}\,(\theta/j)\,,\ j \in \mathbb{N}$. It is reasonable to expect vd-convergence of point processes associated with cycles of fixed length j to a θ/j-rate homogeneous Poisson point process on some appropriate space.
2. *Functional limit theorems.* As an example, let $X_n(t) = P_{\lceil nt \rceil}([0,1])$, $t \in [0,1]$. It is already known that their one-dimensional distributions weakly converge to $\mathrm{Pois}\,(\theta)$ as $n \to \infty$. What can be said about finite-dimensional convergence? About convergence in $\mathcal{D}_{[0,1]}$? In order for all these questions to make sense, we need to couple all the permutations on a common probability space. As it was mentioned, it can be done by means of the Chinese restaurant process.

References

Alfréd Rényi.
On the theory of order statistics.
Acta Mathematica Academiae Scientiarum Hungarica, 4(3):191–231, Sep 1953.

Warren J. Ewens.
The sampling theory of selectively neutral alleles.
Theoretical Population Biology, 3(1):87–112, 1972.

Lars Holst.
On the lengths of the pieces of a stick broken at random.
Journal of Applied Probability, 17(3):623–634, 1980.

Richard Arratia, A. D. Barbour, and Simon Tavaré.
Poisson process approximations for the Ewens sampling formula.
The Annals of Applied Probability, 2(3):519 – 535, 1992.

References

Harry Crane.
The ubiquitous Ewens sampling formula.
Statistical Science, 31(1):1 – 19, 2016.

Richard Arratia, A. D. Barbour, and Simon Tavaré.
Logarithmic combinatorial structures: A probabilistic approach.
EMS Monogr. Math. Zürich: European Mathematical Society (EMS), 2003.

Olav Kallenberg.
Random measures, theory and applications, volume 77 of *Probability Theory and Stochastic Modelling*.
Cham: Springer, 2017.

Sidney I. Resnick.
Extreme Values, Regular Variation and Point Processes.
Springer Science+Business Media New York, 2008.

And special thanks to the Armed Forces of Ukraine for making this presentation possible.