Theory of Stochastic Processes
Vol. 26 (42), no. 1, 2022, pp. 21–26

Y. MINIAILYK

GAUSSIAN NOISE RELATED TO GENERALISED EHRENFEST MODEL

In this article we consider the generalization of Ehrenfest model, where at each moment of time not 1 but some k of n particles go from one box to another. We describe this process by a sequence of Bernoulli random vectors. We define related Bernoulli noise on a set of continuous functions for different times, and prove that it converges to Ornstein-Uhlenbeck sequence of Gaussian white noises when number of particles tends to infinity.

1. Generalization of Ehrenfest model

Ehrenfest model is the well-known Markov model from statistical mechanics. It was created in 1907 by Paul and Tatjana Ehrenfest to explain the second law of thermodynamics (see [1], pp. 311–314). In this model n particles are distributed between two boxes A and B. At the moments of time $m = 0, 1, ...$ a randomly chosen particle is changing its box to another. The process can be described in two ways. Firstly, with all possible 2^n states corresponding to relatively placed n particles and secondly, by the number of particles in box A. In the first way the model can be described by a random walk on hypercube, which have uniform stationary distribution. Authors of this model studied it in the second way. Let x_n be a number of particles in box A. x_n has $n + 1$ possible states: $0, 1, ..., n$ with probabilities of transition

$$P(i, i + 1) = \frac{n - i}{n} \qquad P(i, i - 1) = \frac{i}{n} \qquad P(i, j) = 0 \; otherwise \; (i = 0, 1, ..., n)$$

It is well-known that x_n has stationary binomial distribution, which can be gotten by projection of uniform stationary distribution on hypercube (see [2], p.397).

In this article we propose the generalization of Ehrenfest model, where not 1, but some k particles simultaneously change their boxes, where $1 \le k \le n - 1$. This model will be studied in two ways: as a random walk on hypercube and as a sequence of number of particles in box A, the same as in Ehrenfest model. Finally, we consider the Bernoulli noise related to our proposed model on $C([0, 1])$ for different times and study its limit when the number of particles goes to infinity. When the number of particles tends to infinity and $\frac{k}{n} \to \alpha$, where $1 \ge \alpha \ge 0$, the proposed sequence of Bernoulli noises converges to Ornstein-Uhlenbeck sequence of Gaussian white noises.

Let us describe the locations of particles by a sequence of ± 1 x_i^m. Then we can write the state of the system as $\vec{x}_n^m$, where $m \ge 0$ is time, n is the number of particles. For two states $\vec{i}$ and $\vec{j}$ when $P(\vec{i}, \vec{j}) > 0$ let us write $\vec{i} \sim \vec{j}$ and say that $\vec{i}$ is a neighbor of $\vec{j}$. Then notice, that $\vec{x}_n^m$ is a Markov chain with transition matrix

$$P(\vec{i}, \vec{j}) = \begin{cases} \frac{1}{\binom{n}{k}} \text{ if } \vec{i} \sim \vec{j} \\ 0 \text{ otherwise} \end{cases}$$

We say that two states $\vec{i}$ and $\vec{j}$ are connected if there exists an $r > 0$ such that $P^r(\vec{i}, \vec{j}) > 0$. Further, the transitions between states we will call moves.

2010 *Mathematics Subject Classification.* 60H40, 60F05, 60J10.
Key words and phrases. Markov chain, Ehrenfest model, Gaussian noise, characteristic function.

Lemma 1. *Any two states which differ only by* 2 *components are connected.*

Proof. It is enough to prove that for every state $\vec{x}_n^m$ there exist a sequence of moves that change signs only of 2 components x_i^m and x_j^m. At first we will change sign of x_i^m and other $k-1$ numbers in $\vec{x}_n^m$. Then we will change the sign of x_j^m and $k-1$ components from previous step. This completes the proof. □

Definition 1. *State $\vec{x}_n^m$ is called even(odd) if the number of* 1 *in this state is even(odd).*

Theorem 1. *a) If k is odd, then any two states are connected. b) If k is even, then any two even(odd) states are connected.*

Proof. a) Let $x_{i_1}^m$, $x_{i_2}^m$, ..., $x_{i_{k+1}}^m$ be $k+1$ distinct components of $\vec{x}_n^m$. At first move we change signs of $x_{i_2}^m$, ..., $x_{i_{k+1}}^m$. Then at every move s of we change signs of $x_{i_1}^m$ and $x_{i_s}^m$. So, at move $k+1$ we will only change the sign of $x_{i_1}^m$, because k is odd. Hence, any two states are connected.

b) Let's prove that in this case the parity of the state is invariant. Let N and $N^{'}$ be the number of 1 in states $\vec{x}_n^m$ and $\vec{x}_n^{m+1}$ respectively. Suppose b and a be the number of 1 and -1 that were changed respectively. Then $N^{'} = N + a - b = N + k - 2b$, as k is even then N an $N^{'}$ have the same parity. Including Lemma 1 we get the final result. □

Next we will find the stationary distribution of this chain by using

Lemma 2. *Consider the system with properties: 1) Any two of N states are connected. 2) For every state i is possible to move to another d states with $p_{ij} = \frac{1}{d}$ for some $d \in \mathbb{N}$. Then the stationary distribution is unique and equal to $\vec{\pi} = (\frac{1}{N}, \frac{1}{N}, ..., \frac{1}{N})$ [3]*

Theorem 2. *Suppose $\vec{\pi}$ is a stationary distribution of the sequence $\vec{x}_n^m$. a) If k is odd then $\vec{\pi} = (\frac{1}{2^n}, \frac{1}{2^n}, ..., \frac{1}{2^n})$. b) If k is even then $\pi_i = \frac{\alpha}{2^{n-1}}$ and $\pi_i = \frac{1-\alpha}{2^{n-1}}$, if the state i is even or odd respectively, where $1 \geq \alpha \geq 0$.*

Proof. a) Notice, that by Theorem 1 in this case we can apply Lemma 2 with $N = 2^n$, $d = \binom{n}{k}$.

b) Let α be the invariant probability that the state $\vec{x}_n^0$ is even. Then $1 - \alpha$ is the probability that the state $\vec{x}_n^0$ is odd. By the Theorem 1 we can apply Lemma 2 to the sets of even and odd states with $N = 2^{n-1}$, $d = \binom{n}{k}$. □

Let X_n^m be a number of particles in box A at the moment m. By projection of stationary distribution of $\vec{x}_n^m$ on the set of possible states of X_n^m we get

Theorem 3. *a) If k is odd, then there is a unique stationary distribution of the sequence X_n^m, given by*

$$\pi_i = \frac{\binom{n}{i}}{2^n}$$

b) If k is even, then all stationary distributions of the sequence X_n^m are given by

$$\pi_i = \begin{cases} \frac{\alpha\binom{n}{i}}{2^{n-1}}, & i \text{ is even} \\ \frac{(1-\alpha)\binom{n}{i}}{2^{n-1}}, & i \text{ is odd} \end{cases}$$

where $1 \geq \alpha \geq 0$.

2. Bernoulli noise and its convergence

Let f be a continuous function on $[0, 1]$. For a vector $\vec{u} = (u_1, u_2, ..., u_n)$ define product

$$(f, \vec{u}) = \sum_{k=1}^{n} f(\frac{k}{n}) \frac{u_k}{\sqrt{n}}$$

Let $\vec{\epsilon}$ be a Bernoulli random vector.

Definition 2. *The set of random variables*

$$\tilde{\epsilon} = \{(f, \vec{\epsilon}) \mid f \in C([0,1])\}$$

is called Bernoulli noise on $C([0,1])$.

Bernoulli noises were studied in detail in [4]. A lot of results can be found there, such as chaos representation property, Clark formula and other.

Let $\{\vec{x}_n^m\}$ be a sequence of the states of generalised Ehrenfest model for parameter k with initial uniform stationary distribution.

Definition 3. *The set of random variables*

$$\tilde{x}_n = \{\sum_{j=0}^{m}(f, \vec{x}_n^j) \mid f \in C([0,1])\}$$

is called Bernoulli noise related to generalised Ehrenfest model on $C([0,1])$.

We will consider the limit of the sequence of obtained Bernoulli noises when the number of particles and the number of transitions of particles per moment of time tend to infinity.

If k is fixed then for large enough n the difference between products become negligible, so, by CLT, the sequence of Bernolli noises converges to some white noise (which will be accurately proved later in case $\alpha = 0$). Let k be a function of n and $\frac{k}{n} \to \alpha$, $n \to \infty$, where $1 \geq \alpha \geq 0$.

Lemma 3. *Suppose* $\vec{\epsilon}^n = (\epsilon_1, \epsilon_2, ..., \epsilon_n) \in \{-1, 1\}^n$ *and* $a_i \in \mathbb{R}$ $i = 1, 2, ..., n$. *Then*

$$\mathbb{E}\Big(\cos(\sum_{k=1}^{n} \epsilon_k a_k)\Big) = \prod_{k=1}^{n} \cos(a_k)$$

Proof. We will prove Lemma by method of mathematical induction. For $n = 1$ Lemma is obviously true. Suppose it is still true for $n = m$. Let's prove it for $n = m + 1$.

$$\sum_{\vec{\epsilon}^{m+1}} \cos\Big(\sum_{k=1}^{m+1} \epsilon_k a_k\Big) = \sum_{\vec{\epsilon}^{m}} \Big(\cos\Big(\sum_{k=1}^{m} \epsilon_k a_k + a_{m+1}\Big) + \cos\Big(\sum_{k=1}^{m} \epsilon_k a_k - a_{m+1}\Big)\Big)$$

$$\sum_{\vec{\epsilon}^{m+1}} \cos\Big(\sum_{k=1}^{m+1} \epsilon_k a_k\Big) = 2\cos(a_{m+1}) \sum_{\vec{\epsilon}^{m}} \cos\Big(\sum_{k=1}^{m} \epsilon_k a_k\Big) = 2^{m+1} \prod_{k=1}^{m+1} \cos(a_k)$$

□

Theorem 4.

$$S_n = \sum_{j=1}^{m}(f, \vec{x}_n^j) \xrightarrow{d} N\Big(0, \Big(m + 1 + 2 \sum_{m \geq s > p \geq 0} (1 - 2\alpha)^{s-p}\Big) \int_0^1 f^2(x)dx\Big), n \to \infty$$

Proof. We will use characteristic functions. Let $\phi_{S_n}(t)$ be a characteristic function of S_n:

$$\phi_{S_n}(t) = \mathbb{E}(e^{iS_n t})$$

As the distribution of S_n is symmetric, then $\mathbb{E}(sin(S_n t)) = 0$, so

$$\phi_{S_n}(t) = \mathbb{E}\big(\cos(S_n t)\big)$$

Suppose

$$Y_k = \{\vec{y} \in \{-1, 1\}^n : exactly\ k\ components\ are\ equal\ to\ -1\},$$

then for every $\vec{x}_n^j$ there exists $(\vec{y}^1, \vec{y}^2, ..., \vec{y}^j) \in Y_k^m$ such that $x_r^j = x_r^0 y_r^1 y_r^2 ... y_r^j$ for $r = 1, 2, ..., n$. Suppose $N = \binom{n}{k}^m$, then

$$\mathbb{E}\big(\cos(S_n t)\big) = \frac{1}{N} \sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} \mathbb{E}\Big(\cos\big(\sum_{r=1}^{n} x_r^0 \frac{f(\frac{r}{n})}{\sqrt{n}} t \sum_{j=0}^{m} \prod_{q=1}^{j} y_r^q\big)\Big),$$

where $\prod_{q=1}^{j} y_r^q = 1$ for $j = 0$.

Notice, that we can apply Lemma 3 using $\vec{x}_n^0$ as $\vec{\epsilon}$

$$\phi_{S_n}(t) = \frac{1}{N} \sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} \prod_{r=1}^{n} \cos\Big(\frac{f(\frac{r}{n})}{\sqrt{n}} t \sum_{j=0}^{m} \prod_{q=1}^{j} y_r^q\Big) = \frac{1}{N} \sum_{i=1}^{N} \pi_n^i$$

For large enough n the values inside cosines are nearly 0, so we can take logarithm

$$ln(\pi_n^i) = \sum_{r=1}^{n} ln\Big(\cos\big(\frac{f(\frac{r}{n})}{\sqrt{n}} t \sum_{j=0}^{m} \prod_{q=1}^{j} y_r^q\big)\Big)$$

Using Maclaurin's expansion of $ln(x)$ we get

$$ln(\pi_n^i) = \sum_{r=1}^{n} -\frac{1}{2}\Big(\frac{f(\frac{r}{n})}{\sqrt{n}} t \sum_{j=0}^{m} \prod_{q=1}^{j} y_r^q\Big)^2 + O(\frac{1}{n}), n \to \infty$$

$$\lim_{n\to\infty} \phi_{S_n}(t) = \lim_{n\to\infty} \frac{1}{N} \sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} exp\Big(-\frac{1}{2} t^2 \sum_{r=1}^{n} \frac{f^2(\frac{r}{n})}{n} (\sum_{j=0}^{m} \prod_{q=1}^{j} y_r^q)^2\Big)$$

(1)

$$\lim_{n\to\infty} \phi_{S_n}(t) = \lim_{n\to\infty} \frac{1}{N} \sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} exp\Big(-\frac{1}{2} t^2 \sum_{r=1}^{n} \frac{f^2(\frac{r}{n})}{n} (m+1+2 \sum_{0\le p<s\le m} \prod_{q=p+1}^{s} y_r^q)\Big)$$

We will find the limit as $n \to \infty$ using squeeze theorem. We will find the lower bound using the AM–GM inequality

(2)

$$\lim_{n\to\infty} \phi_{S_n}(t) \ge \lim_{n\to\infty} exp\Big(-\frac{1}{2} t^2 \sum_{r=1}^{n} \frac{f^2(\frac{r}{n})}{n} (m+1+2 \sum_{0\le p<s\le m} \frac{1}{N} \sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} \prod_{q=p+1}^{s} y_r^q)\Big)$$

Further,

$$\sum_{(\vec{y}^1, \vec{y}^2, ..., \vec{y}^m) \in Y_k^m} \prod_{q=p+1}^{s} y_r^q = \binom{n}{k}^{m-s+p} \sum_{(\vec{y}^{p+1}, \vec{y}^{p+2}, ..., \vec{y}^s) \in Y_k^{s-p}} \prod_{q=p+1}^{s} y_r^q =$$

$$= \binom{n}{k}^{m-s+p} (\sum_{\vec{y} \in Y_k} y_r)^{s-p} = N(1 - 2\frac{k}{n})^{s-p}$$

Substituting this in (2) we get

$$\lim_{n\to\infty} \phi_{S_n}(t) \ge exp\Big(-\frac{1}{2} t^2 (m + 1 + 2 \sum_{m\ge s>p\ge 0} (1 - 2\alpha)^{s-p}) \int_0^1 f^2(x) dx\Big)$$

To find the upper bound let's rewrite (1) as

$$\lim_{n\to\infty}\phi_{S_n}(t)=exp\Big(-\frac{1}{2}t^2(m+1)\int_0^1 f^2(x)dx\Big)\times$$
$$\times\lim_{n\to\infty}\frac{1}{N}\sum_{(\vec{y}^1,\vec{y}^2,\dots,\vec{y}^m)\in Y_k^m} exp\Big(-t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\sum_{0\le p<s\le m}\prod_{q=p+1}^s y_r^q\Big)=$$
$$=exp\Big(-\frac{1}{2}t^2(m+1)\int_0^1 f^2(x)dx\Big)\times \tag{3}$$
$$\times\lim_{n\to\infty}\frac{1}{\binom{n}{k}^{m-1}}\sum_{(\vec{y}^1,\vec{y}^2,\dots,\vec{y}^{m-1})\in Y_k^{m-1}} exp\Big(-t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\sum_{0\le p<s\le m-1}\prod_{q=p+1}^s y_r^q\Big)\times$$
$$\times\frac{1}{\binom{n}{k}}\sum_{\vec{y}^m\in Y_k} exp\Big(\sum_{r=1}^n(-t^2\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-1}\prod_{q=p+1}^{m-1}y_r^q)y_r^m\Big)$$

By Maclourin's inequality the last term in (3) is less than

$$exp\Big(\sum_{r=1}^n -t^2\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-1}\prod_{q=p+1}^{m-1}y_r^q\Big)\Big(\frac{1}{n}\sum_{r=1}^n exp\big(2t^2\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-1}\prod_{q=p+1}^{m-1}y_r^q\big)\Big)^k \tag{4}$$

For any $C>1$ and large enough n, (4) is less than

$$C\,exp\Big(-(1-2\alpha)t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-1}\prod_{q=p+1}^{m-1}y_r^q\Big)=$$
$$=C\,exp\Big(-(1-2\alpha)t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-2}\prod_{q=p+1}^{m-1}y_r^q-(1-2\alpha)t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\Big)$$

Hence, for any $C>1$ and large enough n, the upper bound becomes

$$C\,exp\Big(-\frac{1}{2}t^2(m+1+2(1-2\alpha))\int_0^1 f^2(x)dx\Big)\times$$
$$\times\lim_{n\to\infty}\frac{1}{\binom{n}{k}^{m-2}}\sum_{(\vec{y}^1,\vec{y}^2,\dots,\vec{y}^{m-1})\in Y_k^{m-2}} exp\Big(-t^2\sum_{r=1}^n\frac{f^2(\frac{r}{n})}{n}\sum_{0\le p<s\le m-2}\prod_{q=p+1}^s y_r^q\Big)\times$$
$$\times\frac{1}{\binom{n}{k}}\sum_{\vec{y}^{m-1}\in Y_k} exp\Big(\sum_{r=1}^n(-(1+(1-2\alpha))t^2\frac{f^2(\frac{r}{n})}{n}\sum_{p=0}^{m-2}\prod_{q=p+1}^{m-2}y_r^q)y_r^{m-1}\Big)$$

By induction, it follows that for any $C>1$ and large enough n

$$\lim_{n\to\infty}\phi_{S_n}(t)\le C\,exp\Big(-\frac{1}{2}t^2\big(m+1+2\sum_{j=1}^m(m-j+1)(1-2\alpha)^j\big)\int_0^1 f^2(x)dx\Big)=$$
$$=C\,exp\Big(-\frac{1}{2}t^2\big(m+1+2\sum_{m\ge s>p\ge 0}(1-2\alpha)^{s-p}\big)\int_0^1 f^2(x)dx\Big)$$

Since $C>1$ is arbitrary, this finishes the proof. □

Theorem 4 shows, that the sequence of Bernoulli noise related to generalised Ehrenfest model converges to Gaussian white noise, as the number of particles tends to infinity.

Using the similar proof we can get more general resut

Theorem 5. *Let $\lambda_j \in \mathbb{R}$ for $j = 0, 1, ..., n$, then:*

$$S_n = \sum_{j=0}^{m} \lambda_j (f, \vec{x}_n^j) \xrightarrow{d} N\Big(0, \Big(\sum_{j=0}^{m} \lambda_j^2 + 2 \sum_{m \geq s > p \geq 0} \lambda_s \lambda_p (1-2\alpha)^{s-p}\Big) \int_0^1 f^2(x)dx\Big), n \to \infty$$

Consequence.

$$cov\big((f, \vec{x}_n^s), (f, \vec{x}_n^p)\big) \to (1-2\alpha)^{|s-p|} \int_0^1 f^2(x)dx, n \to \infty$$

Theorem 6. *Suppose $\xi_0 \sim N(0, \frac{1}{1-\alpha^2})$, $\{\eta_n\}$ is a sequence of standart normal independent random variables independent from ξ_0 and $\xi_n = \alpha\xi_{n-1} + \eta_n$, $n \geq 1, 1 > \alpha > 0$.*

Then $\xi_n \sim N(0, \frac{1}{1-\alpha^2})$ and $cov(\xi_s, \xi_p) = \frac{\alpha^{|s-p|}}{1-\alpha^2}$

Proof. Proof is standard. □

As Theorem 4 shows that the sequence of Bernoulli noises related to generalised Ehrenfest model converges to a sequence of Gaussian white noises, it's natural to assume correlation between products $(f, \vec{x}_n^j)$. Theorem 6 shows that correlation between white noises produced as limits of $(f, \vec{x}_n^j)$ satisfies autoregressive model of order 1.

In conclusion, the author thanks his supervisor, A. A. Dorogovtsev, for useful comments.

References

1. Paul und Tatjana Ehrenfest, *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*, Physikalische Zeitschrift **8** (1907).
2. William Feller, *An introduction to probability theory and its applications*, 3^{th} edition vol.1, London: John Wiley and Sons, Inc., 1968.
3. David A. Levin and Yuval Peres, *Markov Chains and Mixing Times: Second Edition*, American Mathematical Society, 2017. `https://doi.org/10.1007/s00283-018-9839-x`
4. N. Privault, *Stochastic analysis of Bernoulli processes*, Probab. Surv. **5** (2008), 435–483. `https://doi.org/10.1214/08-PS139`

Y. Miniailyk Richelieu Scientific Lyceum, Yelisavetynska str., 5, Odessa, 65026, Ukraine
E-mail address: `uragenesis206@gmail.com`