Письма в ЖЭТФ, том 106, вып. 7, с. 451 – 458

Реализация квантового генератора случайных чисел, основанного на оптимальной группировке фотоотсчетов

К. А. Балыгин[a], *В. И. Зайцев*[a], *А. Н. Климов*[a,b], *С. П. Кулик*[a], *С. Н. Молотков*[c,d,e1)]

[a] *Физический факультет МГУ им. М.В. Ломоносова, 119991 Москва, Россия*

[b] *Институт общей физики им. А.М. Прохорова РАН, 119991 Москва, Россия*

[c] *Институт физики твердого тела РАН, 142432 Черноголовка, Россия*

[d] *Академия Криптографии Российской Федерации, 121552 Москва, Россия*

[e] *Факультет вычислительной математики и кибернетики МГУ им. М.В. Ломоносова, 119991 Москва, Россия*

Поступила в редакцию 7 июля 2017 г.
После переработки 7 сентября 2017 г.

При реализации квантовых генераторов случайных чисел принципиально важно иметь математически доказуемый и физически экспериментально проверяемый процесс измерений над системой, из которого генерируется исходная случайная последовательность. Это позволяет быть уверенным, что происхождение случайности действительно имеет квантовую природу. Экспериментально реализован квантовый генератор случайных чисел, использующий регистрацию квази однофотонного излучения матрицей SiPM (Silicon Photo-Multiplier), что позволяет надежно достичь пуассоновской статистики фотоотсчетов. Выбор и использование оптимальной группировки фотоотсчетов для исходной последовательности актов фотодетектирования и полиномиального по длине последовательности метода извлечения случайной последовательности 0 и 1 позволил достичь скорости выходной доказуемо случайной последовательности в 64 Мбит/с.

DOI: 10.7868/S0370274X17190109

Введение. Случайные числа широко используются в различных областях науки и техники, например, при вычислении многомерных интегралов, моделировании различных процессов методом Монте-Карло. Наиболее широкое применение случайные числа находят в криптографии. Случайные последовательности используются для секретных ключей в системах симметричного шифрования, генерации паролей, PIN-кодов для различных типов пластиковых карт, кодов аутентификации, вероятностных алгоритмов и систем квантового распределения ключей. Практически для всех упомянутых применений требуются случайные числа, полученные *исключительно с физических генераторов.*

Все генераторы случайных чисел можно разделить на два класса. Первый класс – генераторы, основанные на некоторых математических преобразованиях, как правило рекурентных, затравочного числа, обычно случайного. Такие генераторы выдают псевдослучайную последовательность. При известном алгоритме вся случайность связана с неопределенностью затравочного числа. Второй класс – физические генераторы. Случайная последовательность возникает как результат измерения состояния физической системы. Если эволюция системы описывается законами классической физики, то случайность связана только с неопределенностью начальных условий. Даже при сложном законе классической эволюции начальные условия, в принципе,могут быть восстановлены. После этого эволюция системы будет полностью предсказуемой. То есть с логической точки зрения последовательности также будут псевдослучайными, поскольку могут быть в принципе восстановлены по начальным условиям и известному закону эволюции системы.

В этом смысле только квантовые генераторы случайных чисел, представляющие отдельный тип физических генераторов, могут производить истинную случайную последовательность. Результаты измерений над квантовой системой, *приготовленной каждый раз в одном и том же состоянии, носят принципиально случайный характер.* Поэтому истинная случайность имеет место только в квантовой области.

[1)] e-mail: sergei.molotkov@gmail.com

Существует несколько типов измерений над квантовой системой, которые используются для получения случайных последовательностей. Наиболее широко используемым является процесс фотодетектирования. Сам по себе акт поглощения фотона является сугубо квантовым процессом [1, 2].

Было предложено большое число вариантов квантовых генераторов случайных чисел [3, 4–19], которые за исключением отдельных реализаций [4], можно разделить на две группы.

Первая группа работ в качестве исходной случайной величины использует непрерывную случайную величину. Такой величиной обычно является квадратурная компонента поля при гомодинном детектировании [11] или фаза лазерного излучения [12]. Обе измеряемые величины возникают как разность токов двух классических фотодетекторов, работающих не в режиме счета фотонов, а в линейном классическом режиме. Поэтому данный тип квантовых детекторов, хотя это обычно не произносится, занимает промежуточное положение между классическими физическими генераторами случайных чисел, использующими, например, оцифровку шума Джонсона–Найквиста, и истинно квантовыми.

Вторая группа в качестве источника физической случайности использует счет фотонов сильно ослабленного лазерного излучения при помощи лавинных однофотонных детекторов [3]. Такое излучение имеет пуассоновское распределение по числу фотонов [1, 2]. Дискретной случайной величиной является акт фотодетектирования – отсчет в некоторый временной интервал. Дискретные моменты времени привязаны к clock-генератору регистрирующей электроники. Возможны различные способы группировки фотоотсчетов [5–10]. Например, функция распределения временных интервалов τ между последовательными фотоотсчетами является экспоненциальной: $P(\tau) = e^{-\lambda\tau}$ (λ – некоторая константа, зависящая от интенсивности излучения, квантовой эффективности детектора). В работах [5–9] в качестве случайной дискретной переменной использовались случайные интервалы времени между последовательными отсчетами. В [5–9] случайными событиями были моменты времени от темновых отсчетов лавинного детектора. Однако такой способ дает низкую скорость генерации случайной последовательности.

Функция распределения фотоотсчетов не является равномерной по времени, поэтому требуется постобработка для получения случайной последовательности 0 и 1. Методы постобработки также можно разделить на две группы.

Первая группа – универсальные методы экстракции, которые теоретически работают, даже если отдельные акты измерений коррелированы между собой. Универсальные методы постобработки при известной $H_{\min}$ энтропии процесса позволяют получить последовательности 0 и 1 *сколь угодно близкие – ε близкие в смысле следового расстояния [13–18], к идеально случайной.* Такие методы используют сжатие (хеширование) исходной последовательности при помощи специального вида функций (например, универсальными хеш-функциями второго порядка). Правда, такие методы *требуют затравочной случайности для выбора хеш-функции.* Кроме того, эти методы *даже при известной статистике фотоотсчетов* дают последовательность, лишь сколь угодно близкую к идеальной. Универсальные методы постобработки чаще используются для выравнивания последовательностей, полученных при гомодинном детектировании [11], и в некоторых работах при выравнивании последовательностей при дискретном фотодетектировании [12, 18].

Вторая группа методов постобработки, *при известной статистике фотоотсчетов, получает блоки истинно случайных 0 и 1, а не сколь угодно близких к случайным, например, [7–9].* Однако данные методы постобработки работают, если гарантируется отсутствие корреляций между фотоотсчетами и известен тип статистики фотоотсчетов, что требует специальных мер при разработке таких квантовых генераторов. При этом невозможно извлечь из физического процесса случайности больше, чем в нем исходно содержится. Неучет этого обстоятельства может приводить к странным результатам. Например, в работе [6] число случайных бит не зависит от параметра физического процесса и от длины разбиения на временные блоки.

Любой процесс содержит некоторое количество истинно случайных 0 и 1, которое определяется энтропией (см. детали в [15])

$$H_{\min} = -\log \max_{X^{(n)}} P_{X^{(n)}}(x_1, x_2, \dots x_n), \qquad (1)$$

где $P_{X^{(n)}}(x_1, x_2, \dots, x_n)$ – функция распределения случайных исходов измерений x_i.

В общем случае последовательные исходы измерений x_i не являются некоррелированными $P_{X^{(n)}}(x_1, x_2, \dots, x_n) \neq P_X(x_1) \cdot P_X(x_2) \dots P_X(x_n)$. Как правило, функция распределения коррелированных событий неизвестна, поэтому приходится строить модельные предположения о ее виде. Причем коррелированность результатов измерений на практике связана в значительной степени с неидеальностями самой регистрирующей аппаратуры, в

частности фотодетекторов, которые имеют конечное мертвое время, собственные темновые шумы и имеют afterpulsing – повторные срабатывания детектора, связанные с рассасыванием лавины носителей, а не с регистрацией собственно фотонов.

Идея метода. Таким образом, для практической реализации квантового генератора случайных чисел надо решить следующие задачи.

(i) Выбрать способ фотодетектирования, который бы контролируемым образом обеспечивал независимость отдельных актов регистрации, и приводил к пуассоновской статистике фотоотсчетов.

(ii) Выбрать такую группировку фотоотсчетов в отдельные блоки, которая бы обеспечивала извлечение всей случайности, содержащейся в процессе регистрации квантовых состояний.

(iii) Выбрать такой способ постобработки, который бы гарантировал получение идеально случайной последовательности, а не ε, близкой к идеальной. Способ постобработки должен обеспечивать эффективную реализацию по числу операций при увеличении размеров обрабатываемых блоков последовательностей исходных фотоотсчетов.

(iv) Квантовый генератор случайных чисел, как единая система, должен быть простым и прозрачным для анализа.

Обеспечение пуассоновской статистики фотоотсчетов. В качестве реализации была выбрана схема с одним фотодетектором. В случае идеальных детекторов без темновых шумов, мертвого времени и afterpulsing статистика фотоотсчетов от ослабленного лазерного излучения должна иметь пуассоновский характер [1, 2]. Вероятность обнаружить m фотонов во временном окне T при пуассоновской статистике есть (см., например, [1, 2])

$$P_T(m) = e^{-\mu T}\frac{(\mu T)^m}{m!}.$$

Поскольку лавинные детекторы не различают число фотонов, то случайными событиями являются: отсутствие фотоотсчета во временном окне (такте) T – $\sqcup$ или фотоотсчет – $*$. При этом вероятность фотоотсчета (от одного, двух и более фотонов в окне T) есть $P(*) = 1 - e^{-\mu T}$, соответственно, отсутствие отсчета $P(\sqcup) = e^{-\mu T}$. Отметим, что параметр μ включает в себя среднее число фотонов в излучении и квантовую эффективность лавинного детектора. Темновые шумы также приводят к пуассоновскому процессу. Композиция двух пуассоновских процессов также есть пуассоновский процесс. Поэтому вероятность $P(*) = 1 - e^{-\mu T}$ есть наблюдаемая вероятность, которая включает несколько составляющих.

Главная проблема при фотодетектировании однофотонных сигналов одиночными лавинными фотодетекторами состоит в учете мертвого времени последних, что ограничивает тактовую частоту и темп фотоотсчетов. Тактовая частота не может превышать обратное время рассасывания лавины. Кроме того, эффекты afterpulsing после регистрации реального фотона могут приводить к паразитным отсчетам после регистрации фотона. Оба упомянутых эффекта нарушают идеальную пуассоновскую статистику фотоотсчетов.

Для устранения данных паразитных отсчетов были использованы не одиночные лавинные фотодетекторы, а матрицы лавинных детекторов, содержащие несколько тысяч таких детекторов – матрицы кремниевых фотоумножителей (SiPM). Поскольку среднее число фотонов за время тактового импульса составляет не более одной тысячной фотона на пиксел (см. ниже), то после регистрации отдельным фотодетектором фотона, вероятность того, что следующий фотон попадет в тот же самый детектор, крайне мала. Это позволяет увеличить тактовую частоту и темп поступления фотонов на матрицу. В этом случае мертвое время отдельного фотодетектора и тактовая частота оказываются развязанными.

Выбор оптимальной группировки фотоотсчетов. Из предыдущего раздела следует, что случайная величина – "есть отсчет", "нет отсчета" во временно́м окне – подчиняется распределению Бернулли, т.е. в каждом временно́м окне может быть только один акт регистрации. Как было показано ранее [10], оптимальной группировкой фотоотсчетов, которая позволяет извлечь всю случайность в асимптотическом пределе длинных блоков, является группировка, приводящая к статистике Ферми–Дирака.

Экстракция случайной последовательности 0 и 1 из пуассоновской последовательности фотоотсчетов. Под случайной последовательностью понимается последовательность 0 и 1, которые независимы в каждой позиции, и вероятность $P(0) = P(1) = \frac{1}{2}$. Экстракция истинно случайной последовательности из последовательности фотоотсчетов на формальном языке означает отображение последовательностей фотоотсчетов длины из n тактов с m отсчетами в истинно случайные блоки 0 и 1 длиной l, зависящей от n и m, $\{*, \sqcup\}^{n_m} \to \{0,1\}^l$ ($l \le n$). Важно, чтобы метод группировки позволял обрабатывать последовательность фотоотсчетов и извлекать случайную последовательность вычислительно эффективно, т.е. с полиномиальными вычислительными ресурсами по длине обрабатываемой последовательности.

Общая идея экстракции случайной последовательности 0 и 1 из пуассоновской последовательности фотоотсчетов состоит в следующем и выполняется в два этапа. Напомним, что пуассоновский процесс является однородным во времени. Для дальнейшего важно, что процедуре извлечения случайности требуется только статистическая независимость фотоотсчетов и не требуется знание самих вероятностей $p = P(\sqcup)$ и $1 - p = P(*)$.

Последовательность разбивается на ходу на блоки, содержащие одинаковое число тактов (имеющих одинаковую временну́ю длительность). Пусть длина блока – n тактов. В каждом блоке может быть m фотоотсчетов (*), $0 \le m \le n$. Всего типов таких блоков существует 2^n. Последовательности, содержащие одинаковое число тактов с отсчетами $*$ и пустых тактов $\sqcup$, имеют одинаковую вероятность. Вероятность последовательности с m отсчетами $(*)$ и $n - m$ пропусками $(\sqcup)$ имеют вероятность $(1 - P(*))^{n-m} P(*)^m$, соответственно, полное число равновероятных последовательностей в данном классе равно C_n^m.

Первый этап состоит в нумерации всех равновероятных последовательностей из одного класса [10]. Второй этап состоит в извлечении блока случайных 0 и 1 из двоичного номера последовательности в данном классе [10]. Например, если число последовательностей и соответственно номеров в данном классе есть степень двойки, то двоичное представление номера сразу дает случайный блок 0 и 1 (общий случай см. далее).

Возможен способ нумерации в “лоб”, когда сама последовательность является адресом номера в данном классе. Поскольку число последовательностей экспоненциально велико, то размер адресной таблицы также экспоненциально велик. Например, при длине обрабатываемого блока в 64 такта размер адресной таблицы $2^{64} \approx 10^{19}$. Такая таблица практически не реализуема. Скорость генерации случайных чисел при таком способе технически ограничена размером таблицы, который растет экспоненциально с длиной блока.

Существует элегантный способ нумерации последовательностей фотоотсчетов, который требует лишь полиномиальных ресурсов по длине последовательности и позволяет обрабатывать практически последовательности любой длины. Воспользуемся методами из теории арифметического кодирования, часто называемого кодированием без потерь. Выходной последовательности $i_1, i_2, \ldots, i_n$ сопоставляется номер, который определяется полиномиальным алгоритмом (см. детали в [20]). Пусть в последовательности имеется K отсчетов *. Обозначим номер позиции индексом j_m. Тогда имеется взаимно однозначное соответствие между последовательностью фотоотсчетов $(i_1, i_2, \ldots, i_n)$ и ее номером $Num(i_1, i_2, \ldots, i_n)$ $(0 \le Num(i_1, i_2, \ldots, i_n) \le C_n^K - 1)$:

$$Num(i_1, i_2, \ldots i_n) = C_{j_1-1}^1 + C_{j_2-1}^2 + \ldots + C_{j_k-1}^k, \quad C_j^l = 0, \quad j < l. \tag{2}$$

Биномиальные коэффициенты вычисляются заранее и помещаются в таблицу размером $n \times n$ в память FPGA (рис. 1). Положение и сам биномиальный коэффициент в таблице выбирается ”на ходу” по мере появления последовательности фотоотсчетов (см. рис. 1). При появлении первого отсчета в позиции j_1 выбирается число (биномиальный коэффициент) на пересечении первой строки j_1 и первого столбца матрицы. При появлении второго отсчета берется коэффициент в матрице на пересечении j_2 строки и второго столбца, и т.д. В итоге получается номер последовательности $Num(i_1, i_2, \ldots, i_n)$. Такой способ нумерации естественным образом укладывается на архитектуру FPGA.

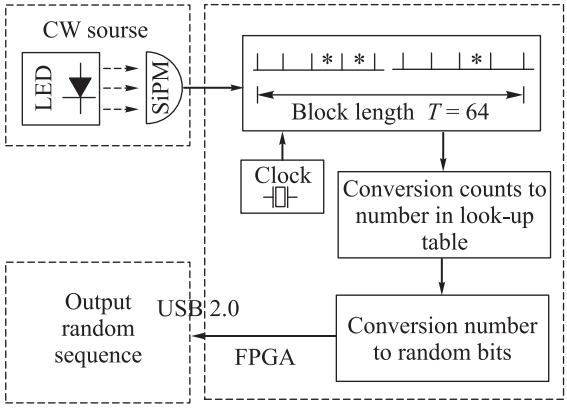

Рис. 1. Функциональная блок-схема генератора случайных чисел. Тактовая частота составляла 150 МГц. Длительность импульсов на полувысоте < 2 нс

Пусть последовательность принадлежит к некоторому классу, полное число последовательностей в классе – N_k.

Следующий этап – по номеру последовательности выдается блок случайных 0 и 1. Номера последовательностей находятся в диапазоне $0 \le Num \le N_k - 1$. Далее – номер текущей последовательности Num. Пусть бинарное представление числа последовательностей в классе $N_k = \sum_{i=0}^{i_{\max}} 2^{k_i}$. Процедура выполняется рекурсивно. Если номер текущей последовательности Num находится в интервале $2^{k_0} + 2^{k_1} + \ldots + 2^{k_{i-1}} \le Num \le 2^{k_0} + 2^{k_1} + \ldots + 2^{k_{i-1}} + 2^{k_i} - 1$, $(i \le i_{\max})$, тогда выходной случайной последова-

тельностью будет k_i младших разрядов бинарного представления Num. Число номеров последовательностей в этом диапазоне равно 2^{k_i}. Подчеркнем, что такой способ экстракции дает истинно случайную последовательность, а не ε близкую к случайной. Единственное условие для этого – пуассоновский характер последовательности фотоотсчетов, что должно достигаться аккуратной экспериментальной реализацией.

Квантовый генератор случайных чисел. Поскольку способ постобработки решающим образом определяется исходной физической случайной последовательностью, то физическая реализация генератора должна быть максимально прозрачной и простой, но не в ущерб качеству и математической обоснованности реализации[2)], для анализа и обоснования качества исходной последовательности фотоотсчетов.

Экспериментальная реализация. В качестве SiPM была выбрана матрица фирмы SENSL, с оптической площадкой 1×1 мм, и активной оптической площадкой отдельного пиксела 10×10 мкм. Матрица содержала $N_{\text{pic}} = 2880$ детекторов – пикселей. Квантовая эффективность η отдельного детектора составляла $\eta \approx 10\,\%$ на длине волны 0.405 мкм. Температура детектора была стабилизирована на уровне 25°C. Скорость темного счета 30 КГц/мм^2 активной площади, применялся детектор с активной площадью 1 мм^2. Источником излучения являлся светодиод (SLD3143VL) фирмы Sony Laser Diode с рабочей длиной волны излучения 0.405 мкм. Для обработки использовалась FPGA фирмы Intel FPGA (Altera). Таковая частота 150 МГц. Внешний интерфейс для выдачи результирующей случайной последовательности в непрерывном потоке – USB 2.0.

Оценка среднего числа фотонов на пиксел. Покажем, что генератор действительно работает в квантовом режиме. Оценим среднее число фотонов, падающих на SiPM. Реально наблюдаемой величиной является вероятность отсчета за один такт $P(*) = 0.207$. Данная вероятность равна $P(*) = \mu\eta N_{\text{pic}}$, μ – среднее число фотонов на один пиксел в SiPM за один такт, $\eta \approx 0.1$ – квантовая эффективность пиксела, $N_{\text{pic}} = 2880$ – число пикселей в матрице. В итоге получаем $\mu = \frac{P(*)}{\eta\, N_{\text{pic}}} \approx 0.7 \cdot 10^{-3}$ [(фотонов)/(в такт на пиксел)], т.е. приходится менее тысячной фотона на пиксел. Напомним, что для когеретного состояния с пуассоновской статистикой вероятность появления одного фотона $P(n = 1) = e^{-\mu}\mu \approx \mu$ ($\mu \ll 1$), соответственно, вероятность появления двух фотонов $P(n = 2) = e^{-\mu}\frac{\mu^2}{2} \approx 2.5 \cdot 10^{-7}$. Таким образом, реализован практически однофотонный режим[3)].

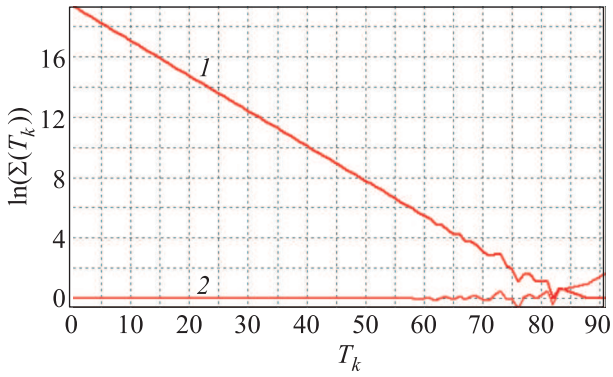

Рис. 2. (Цветной онлайн) Зависимость *1* – логарифм количества временны́х интервалов $\ln(\Sigma(T_k))$ в зависимости от длины интервала T_k, k – число тактов. Метки по оси абсцисс отвечают числу тактов на частоте 150 МГц. Зависимость *2* – логарифм количества временных интервалов с вычитанием наклона. Наблюдаемая вероятность отсчета за один такт $P(*) = 0.207$. Гистограмма строилась с накоплением, полное число событий в гистограмме $\sum_{k=0}^{128} \Sigma(T_k) = 1257768945$. Число отсчетов в первом ящике гистограммы 259338190, соответственно, $\ln(259338190) = 19.374$. При верхней границе по длительности интервалов в 128 тактов вероятность отсчетов практически равна нулю. Флуктуации при временном интервале между отсчетами при числе тактов > 70 связаны с редкими событиями за счет экспоненциальной зависимости функции распределения временны́х интервалов. Длина обрабатываемых блоков для извлечения случайных блоков была выбрана в 64 такта

Проверка пуассоновской статистики фотодетектирования. Таким образом, матрица детекторов представляет собой один эффективный счетчик фотонов, близкий к идеальному. При пуассоновской статистике интервалы между последовательными отсчетами представляют собой случайную величину, подчиняющуюся геометрическому распределению

$$P(T_k) = (1 - P(*))^{k-1} P(*), \qquad (3)$$

[2)] Например, в работе [12] декларировалась скорость генерации в несколько ГГц, при этом постобработка состояла в простой операции XOR для блоков исходной последовательности, сдвинутых друг относительно друга, что вряд ли можно как-то строго обосновать.

[3)] Отметим, что для учета возможного паразитного влияния эффекта cross-talk между соседними пикселями на статистику фотоотсчетов (см., например, [26]) было проведено дополнительное исследование. При наших рабочих уровнях потоков фотонов искажения статистики выявлено не было (рис. 2).

в котором T_k – число тактов между последовательными моментами регистрации.

При идеальной пуассоновской статистике логарифм вероятности $\ln(P(T_k))$ (k – число тактов) должен представлять собой линейную зависимость от k. На рис. 2 приведена экспериментальная гистограмма. Как видно из рис. 2, зависимость демонстрирует пуассоновскую статистику.

Скорость генерации результирующей случайной последовательности 0 и 1 в итоге составляла 63.98 Мбит/с.

Проверка статистических свойств случайных последовательностей. Существует несколько наборов статистических тестов [21–24]. Был выбран стандартный набор тестов NIST [24], прохождение которых является основанием для дальнейшего тестирования специальными тестами. Общая идеология проверки случайности последовательности вкратце сводится к следующему. Проверяется гипотеза H_0 – последовательность является истинно случайной. Если это так, то различные статистики S – группировки 0 и 1 также являются случайными величинами, распределения вероятностей различных статистик при длине последовательности $n \to \infty$ должны стремиться к некоторым эталонным распределениям для случайной последовательности. Статистика как случайная величина подвержена флуктуациям. Задается некоторый уровень значимости α и некоторое пороговое значение для каждой статистики. Если вероятность уклонения статистики превышает пороговое значение, то гипотеза о случайности отклоняется, последовательность отбрасывается. Это означает, что даже генератор идеальных случайных последовательностей может выдать последовательность, которая имеет такое уклонение.

Подсчитывается вероятность *P-value* – вероятность того, что даже идеальный случайный источник может выдать последовательность с таким уклонением статистики. Если $P > \alpha$, то гипотеза H_0 принимается, при $P < \alpha$ гипотеза отклоняется, последовательность считается не случайной.

Интерпретация P-value значения. При заданном уровне значимости α, *P-value* есть вероятность того, что даже идеальный генератор "имеет право" сгенерировать с такой вероятностью последовательность, которая будет выглядеть как не случайная для данного теста. Чем меньше *P-value*, тем с меньшей вероятностью идеальный генератор "имеет право" сгенерировать такую последовательность. Если вычисленное *P-value* больше α, то тест считается пройденным. Стандартное значение уровня значимости для $\alpha \in [0.001, 0.01]$ [24]. При выполнении тестов использовалась $\alpha = 0.01$.

Тест тестов. Доля последовательностей, прошедших тесты сама является случайной величиной, поэтому тоже испытывает флуктуации. Допустимый диапазон флуктуаций определяется дисперсией *P-value*. Поскольку *P-value* являются случайными величинами с распределением Бернулли – тест пройден, либо тест не пройден, то допустимый разброс *P-value* должен укладываться в "три сигма". Величина дисперсии для *P-value* есть $P(1-P)N$ (N – число тестируемых последовательностей). Согласно [24], при уровне значимости $\alpha = 0.01$ все *P-value* должны попадать в интервал "три сигма" $1-P\pm3\sqrt{\frac{P(1-P)}{N}} = 0.99 \pm \frac{0.297}{N}$.

Тест на однородность значений P-value. При большом числе тестируемых последовательностей (в предположении, что тесты проводятся в одних и тех же условиях) суммарное значение *P-value* по всем тестам есть сумма одинаково распределенных величин, которая распределена по гауссовскому нормальному закону [24]. В этом случае статистика

$$X_N^2 = \sum_{j=1}^{M} \frac{(\nu_j - Np_j)^2}{Np_j},$$

где ν_j – доля значений *P-value*, попадающих в j-ый интервал $[0, 1]$, p_j – истинная вероятность попадания в j-ый интервал. При большом числе тестируемых последовательностей распределение вероятностей статистики X_N^2 не зависит от распределения p_j и стремится к распределению Пирсона $\chi^2(N-1)$ с $(N-1)$-ой степенью свободы [25]. Рекомендуемое число интервалов $M = 10$ [24]. Пороговое значение при нулевой гипотезе (H_0 – последовательность случайна) допустимого разброса получается при заданном уровне (вероятности) значимости α из соотношения $\Pr\{X_N^2 > t_\alpha | \mathrm{H}_0\} = \alpha$, где $t_\alpha = \chi^2_{1-\alpha,N-1}$ – $(1-\alpha)$-ая квантиль распределения χ^2 с $(N-1)$-ой степенью свободы.

Другими словами, если уклонение статистики $X_N^2 > \chi^2_{1-\alpha,N-1}$, то нулевая гипотеза H_0 отвергается, и последовательность считается не случайной, так как уклонение статистики от допустимой нормы (при заданной вероятности – уровне значимости) превышено.

Тест на однородность *P-value* считается пройденным, если *P-value* от *P-value* $\hat{P} = \frac{\int_{X_K^2}^{\infty} dx e^{-x/2} x^{K/2-1}}{2^{K/2}\Gamma(K/2)}$, $K = N-1$, не менее $\hat{P} > 0.0001$. Тест на однородность считается пройденным, и принимается H_0 гипотеза – последовательность случайная.

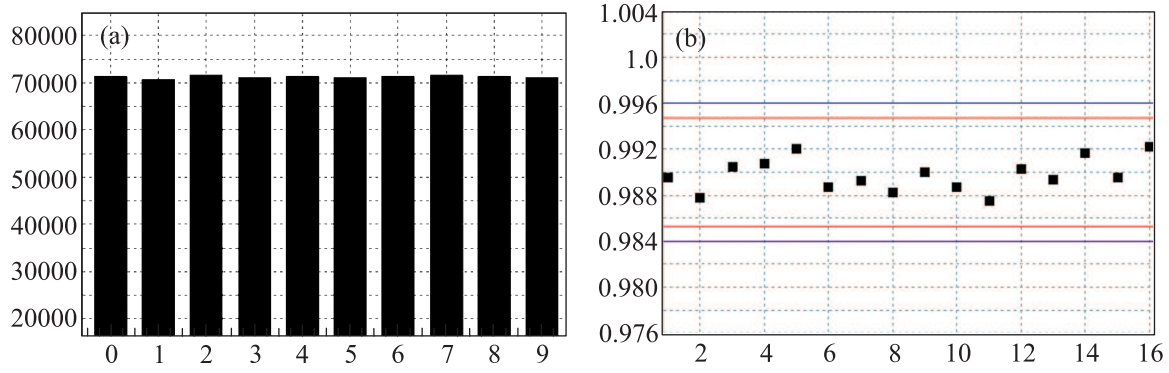

Рис. 3. (Цветной онлайн) (a) – Гистограмма значений ν_j (*P-value*) по всем тестам, попадающих в 10 интервалов $j = 0, ..., 9$. Интервалу [0.0,0.1] отвечает значение $j = 0$ по оси абсцисс, интервалу [0.1,0.2] отвечает значение $j = 1$ по оси абсцисс, и т.д. Гистограмма строилась для 4000 случайных последовательностей размером 10^6 бит каждая. (b) – По оси ординат – доля последовательностей, которые прошли тесты по значениям *P-value*. По оси абсцисс номера отвечают различным тестам по номенклатуре NIST [24]. Красными горизонтальными линиями (две внутренние) показаны верхняя и нижняя границы ("три сигма") в соответствии для всех тестов кроме *№12 – Random Excursions Test* и *№13 – Random Excursions Variant Test*. Для данных двух тестов число тестируемых последовательностей N не определено заранее, а определяется в процессе тестирования. Верхняя и нижняя границы ("три сигма") для данных двух тестов показаны синими горизонтальными линиями (две внешние линии). Для остальных тестов число тестируемых последовательностей равно $N = 4000$ каждая длиной 10^6 бит

На рис. 3 приведены результаты теста на равномерность *P-value* (рис. 3a). Финальное значение *P-value* от *P-value* равно $\hat{P} > 0.73743536$, которое должно быть больше критического значения $\hat{P}_c > 0.0001$. Это означает, что тест на равномерность *P-value* пройден с запасом. Доля последовательностей, прошедших тест, показана на рис. 3b для всех тестов (номера тестов NIST [24] приведены по горизонтали). Для всех тестов доля последовательностей, прошедших тесты по величинам *P-value*, лежит в пределах "три сигма", что означает успешное прохождение тестов – гипотеза о случайности последовательности принимается.

Выражаем благодарность коллегам из Академии Криптографии Российской Федерации за обсуждения и постоянную поддержку. Авторы также благодарят И.М. Арбекова, И.В. Радченко, К.С. Кравцова за обсуждения. Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации (проект # 03.G25.31.0254).

1. Д. Н. Клышко, *Фотоны и нелинейная оптика*, Наука, М. (1980).
2. Д. Н. Клышко, А. В. Масалов, УФН **165**, 1249 (1995).
3. M. Herrero-Collantes and J. Carlos Garcia-Escartin, Rev. Mod. Phys. **89**, 015004 (2017).
4. Y. Okawachi, M. Yu, K. Luke, D. O. Carvalho, M. Lipson, and A. L. Gaeta, Opt. Lett. **41**, 4194 (2016).
5. F.-X. Wang, C. Wang, W. Chen, S. Wang, F.-S. Lv, D.-Y. He, Z.-Q. Yin, H.-W. Li, G.-C. Guo, and Z.-F. Han, J. Light Wave Techn. **33**, 3319 (2015).
6. Q. Yan, B. Zhao, Z. Hua, Q. Liao, and H. Yang, Rev. Sci. Instrum. **86**, 073113 (2015).
7. Y.-Q. Nie, H.-F. Zhang, Z. Zhang, J. Wang, X. Ma, J. Zhang, and J.-W. Pan, Appl. Phys. Lett. **104**, 051110 (2014).
8. K. S. Kravtsov, I. V. Radchenko, S. P. Kulik, and S. N. Molotkov, J. Optical Soc. Amer. **32**, 1743 (2015); arXiv:1507.02059.
9. И. В. Радченко, *Приготовление и измерение квантовых состояний в протоколах квантовой коммуникации*, канд. дисс. Институт общей физики имени А.М. Прохорова РАН, М. (2015).
10. С. Н. Молотков, Письма в ЖЭТФ **105**, 374 (2017).
11. C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, and G. Leuchs, Nature Photonics Lett. **4** 771 (2010).
12. J. Yang, J. Liu, Q. Su, Z. Li, F. Fan, B. Xu, and H. Guo, Opt. Exp. **24**, 27475 (2016).
13. L. Trevisan, J. ACM **48**, 860 (2001).
14. W. Mauerer, C. Portmann, and V. B. Scholz, arXiv:1212.0520 (2012).
15. R. König, R. Renner, and C. Schaffner, IEEE Transactions on Information Theory **55**, 4337 (2009).
16. X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, Phys. Rev. A **87**, 062327 (2013).
17. D. Frauchiger, R. Renner, and M. Troyer, arXiv:1311.4547 (2013).

18. D. Stucki, S. Burri, E. Charbon, C. Chunnilall, A. Meneghetti, and F. Regazzoni, Proc. SPIE 8899, *Emerging Technologies in Security and Defence; and Quantum Security II; and Unmanned Sensor Systems X*, 88990R, October 29 (2013).
19. Q. Yan, B. Zhao, Q. Liao, and N. Zhou, Rev. Sci. Instr. **85**, 103116 (2014).
20. В. Ф.Бабкин, Проблемы передачи информации **VII**, 13 (1971).
21. D. E. Knuth, *The Art of Computer Programming*, v. 2, Addison Wesley, Cambridge (1981).
22. G. Marsaglia, *Die Hard: A battery of tests for random number generators*. [Electronic resours] http://stat.fsu.edu/pub/diehard (Date of downloads).
23. P. L'Ecuyer and R. Simard, *TestU01: A Software Library in ANSI C for Empirical Testing of Random Number Generators*, [Electronic resours] http://www.iro.umontreal.ca/ lecuyer (2002) (Date of downloads).
24. *A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications*. [Electronic resours] http://csrc.nist.gov/rng/SP800-22b.pdf (Date of downloads)
25. H. Cramer, *Mathematical Methods of Statistics*, Asia Publishing House (1946).
26. D. A. Kalashnikov, Si-Hui Tan, and L. A. Krivitsky, Opt. Exp. **20**, 5044 (2012).