

## ТЕОРЕТИЧЕСКАЯ И МАТЕМАТИЧЕСКАЯ ФИЗИКА

УДК 537.533.7

О НАДЕЖНОСТИ ОДНОГО ВИДА ГЛАДКОЙ АППРОКСИМАЦИИ  
СТОХАСТИЧЕСКИХ ДАННЫХ

Н. Н. Негуляев

(кафедра компьютерных методов физики)

E-mail: bot@cmp.phys.msu.su

Исследована связь между погрешностью данных, получаемых в результате стохастического компьютерного эксперимента, и качеством их гладкой аппроксимации. Мерой близости (точностью аппроксимации) предложено считать надежность. В качестве примера рассмотрено описание функции плотности энерговыведения электрона, рассчитываемой с помощью имитационного моделирования методом Монте-Карло, суммой двух трехмерных гауссовых функций. Приведены результаты расчетов надежности в зависимости от параметров моделирования.

## Введение

Стохастическое моделирование является мощным средством исследования физических процессов. Например, при изучении явления взаимодействия электронного пучка с веществом, представляющего интерес для областей физики, связанных с электронной микроскопией и микроанализом, широкое распространение получил подход, основанный на имитационном моделировании процесса рассеяния электронов на базе метода статистических испытаний Монте-Карло [1].

При таком подходе энергия, выделяющаяся при взаимодействии электронов со средой, распределяется по ячейкам, количество и размеры которых определяются целями исследования. Функция пространственной плотности энерговыведения (ППЭ), рассчитанной на один электрон, определяемая подобным образом, представляет собой сложную стохастическую двух- или трехмерную кусочно-постоянную функцию. Часто оказывается, что знание ее значений в каждой ячейке является излишним и достаточно иметь такое аналитическое представление, которое состоит из простых функций и при этом сохраняет в себе все основные зависимости, полученные на основе моделирования.

Существует несколько подходов к решению проблемы сжатия данных при исследовании физических моделей поведения пучков частиц [2, 3]. Однако остается неисследованным вопрос о влиянии погрешности (шума) монте-карловского моделирования на качество аппроксимации. В данной работе в качестве меры близости предлагается использовать надежность и одновременно находить значения неизвестных параметров аппроксимирующих функций из условия максимальной надежности. В качестве иллюстрации рассмотрено описание ППЭ электрона суммой двух трехмерных гауссовых функций. Приводятся результаты расчетов надежности в зависимости от различных параметров моделирования.

## Аналитическое приближение функции ППЭ

Авторами работ [2, 3] предлагается следующий вид зависимости функции ППЭ  $G(x, z)$  в точке облучаемого образца с координатами  $x$  и  $z$ :

$$G(x, z) = \frac{d\varepsilon}{dzdx} = c_1 \frac{1}{\pi s_1 z^{3/2} q_1} \frac{2}{1 + \operatorname{erf}(d_1/q_1)} \exp\left(-\frac{x^2}{s_1^2 z^3} - \frac{(z-d_1)^2}{q_1^2}\right) + c_2 \frac{1}{\pi p_2 q_2} \frac{2}{1 + \operatorname{erf}(d_2/q_2)} \exp\left(-\frac{x^2}{p_2^2} - \frac{(z-d_2)^2}{q_2^2}\right), \quad (1)$$

здесь  $\varepsilon = E/E_{pr}$  — нормированная выделяющаяся энергия;  $E_{pr}$  — энергия первичного электрона;  $x$  — поперечное расстояние от точки падения пучка;  $z$  — глубина;  $s_1, p_2, q_1, q_2, d_1, d_2, c_1, c_2$  — подлежащий определению набор параметров,  $\operatorname{erf}(x)$  — функция ошибок. Результатом вычислительного эксперимента являются значения  $G(x, z)$  в конечном наборе точек  $(x_i, z_j)$ ,  $i = 1, \dots, NX$ ,  $j = 1, \dots, NZ$ , заданные с погрешностью, дисперсию которой можно вычислить. Двумерный массив  $G_{ij} = G(x_i, z_j)$  может быть преобразован в одномерный  $\xi_i$ , где  $i = 1, \dots, N$  и  $N = NX \cdot NZ$ . Считаем, что вектор  $\xi = (\xi_1, \dots, \xi_N) \in R_N$  наблюдается в эксперименте по схеме

$$\xi_k = c_1 \frac{g_1}{s_1 z^{3/2}} \exp\left(-\frac{x_k^2}{s_1^2 z^3} - \frac{(z_k - d_1)^2}{q_1^2}\right) + c_2 \frac{g_2}{p_2} \exp\left(-\frac{x_k^2}{p_2^2} - \frac{(z_k - d_2)^2}{q_2^2}\right) + \nu_k, \quad (2)$$

здесь  $g_i = 2/((\pi q_i)(1 + \operatorname{erf}(d_i/q_i)))$ ,  $i = 1, 2$ ,  $k = 1, \dots, N$ ,  $\nu_k$  — компонента вектора погрешности  $\nu$  (шум).

Схема измерений (2) может быть переписана в векторном виде:  $\xi = A(\lambda)\mathbf{f} + \nu$ , где использованы обозначения  $\mathbf{f} = (c_1, c_2) \in R_2$ ,  $\lambda =$

$= (s_1, p_2, q_1, q_2, d_1, d_2) \in R_6$ ,  $\nu = (\nu_1, \dots, \nu_N) \in R_N$  и введена матрица  $A(\lambda) \in R_2 \rightarrow R_N$ :

$$A(\lambda) = \left( \frac{g_1}{s_1 z_i^{3/2}} \exp \left( -\frac{x_i^2}{s_1^2 z_i^3} - \frac{(z_i - d_1)^2}{q_1^2} \right) + \frac{g_2}{p_2} \exp \left( -\frac{x_i^2}{p_2^2} - \frac{(z_i - d_2)^2}{q_2^2} \right) \right),$$

$i = 1, \dots, N$ . Вектор погрешности  $\nu$  в (2) есть сумма случайной погрешности метода Монте-Карло и систематической составляющей, возникающей из-за отличия функции ППЭ от ее аналитического представления (1).

Согласно методу максимальной надежности, для нахождения наилучшего приближения необходимо найти минимум функционала:

$$J(\lambda_1 \xi) = \left\| \left( I - S(\lambda_1) (S^*(\lambda_1) S(\lambda_1))^{-1} S^*(\lambda_1) \right) \xi \right\|^2,$$

где  $S(\lambda_1) = \Sigma^{-1/2} A(\lambda_1)$ ,  $\Sigma$  — ковариационная матрица погрешности  $\nu$  размерами  $N \times N$ ,  $I$  — единичная матрица таких же размеров [4]. Основной трудностью является расчет матрицы  $\Sigma^{-1/2}$ . Вектор  $\xi$  обычно содержит  $\geq 10^4$  компонент [3, 5]. Воспользовавшись тем, что корреляция между компонентами  $\xi_k$  слабая (при  $k$ , соответствующих ячейкам, являющихся в пространстве соседними, отношение  $|\Sigma_{ik}| / |\Sigma_{ii}|$  не превышает 0.1, а с увеличением расстояния между ними оно монотонно убывает до нуля), полагаем все недиагональные элементы матрицы  $\Sigma$  равными нулю.

Для минимизации (3) используем алгоритм, основанный на сочетании стохастических методов поиска минимума функционала (Монте-Карло и градиентного спуска) [6]. Значение функционала (3) в точке минимума определяет надежность модели, характеризующую отличие аппроксимации от истинной функции ППЭ в сравнении со случайной погрешностью  $\alpha = 1 - F_{\chi^2_{N-k}}(J(\lambda, \xi))$ , где  $F_{\chi^2_n}(x)$  — функция распределения Пирсона с  $n$  степенями свободы,  $k$  — размерность пространства  $R(A)$  значений матрицы  $A$  ( $k = 2$ ) [4].

### Результаты расчета

В качестве примера ниже рассмотрена аппроксимация функции ППЭ, возникающей при нормальном падении электронного пучка энергией  $E_{pr} = 30$  кэВ на однородный образец из кремния (рис. 1, 2). Для расчета элементов ковариационной матрицы  $\Sigma$  каждый раз проводилась серия из  $S = 2500$  испытаний. Число событий в одной ячейке в методе Монте-Карло имеет пуассоновское распределение. Это означает, что математическое ожидание и дисперсия числа событий растут линейно с ростом количества расчетных траекторий в одном испытании  $N_{trj}$  или с ростом объема ячейки  $V$ . Последнее утверждение справедливо, если характерные масштабы модели-

руемого события много больше размеров ячейки. При малом  $N_{trj}$ , когда дисперсия шума велика и аппроксимирующая функция везде лежит внутри шумового коридора, надежность аппроксимации  $\alpha$  близка к единице (рис. 3). Это означает, что функция (1) со значениями параметров  $s_1, p_2, q_1, q_2, d_1, d_2, c_1, c_2$ , найденными в результате минимизации функционала (3), не противоречит наблюдаемым данным и аналитическое описание ППЭ в виде (1) адекватно. С увеличением  $N_{trj}$  ширина шумового коридора уменьшается пропорционально  $N_{trj}^{-1/2}$ , а значение надежности  $\alpha$  монотонно убывает. При некотором  $N_{trj}$  аппроксимирующая кривая все еще остается внутри коридора, но в некоторых местах уже касается его границ. В этот момент  $\alpha$  начинает заметно отличаться от единицы. При дальнейшем увеличении величины  $N_{trj}$  (по мере исчезновения случайной составляющей погрешности метода Монте-Карло) кривая перестает попадать в шумовой коридор и надежность спадает до нуля.

Назовем критическим значением количества расчетных траекторий  $N_c$  при заданном фиксированном объеме ячейки разбиения  $V$  такое число траекторий

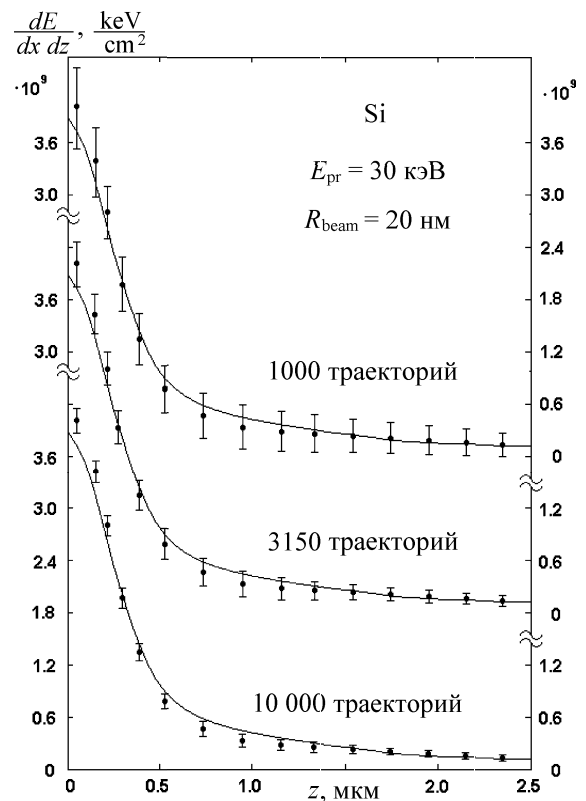


Рис. 1. Аппроксимация функции ППЭ  $G(x, z) = dE/(dx dz)$  при падении электронов с  $E_{pr} = 30$  кэВ в однородный кремниевый образец. Точки — данные, рассчитанные с помощью метода Монте-Карло [5], сплошные кривые — аппроксимирующие функции. Показано сечение вдоль оси пучка  $x = 0$ .  $s_1 = 0.290$ ;  $p_2 = 3.246$ ;  $q_1 = 1.708$ ;  $q_2 = 2.695$ ;  $d_1 = 0.810$ ;  $d_2 = 2.898$ ;  $c_1 = 0.195$ ;  $c_2 = 0.713$ ;  $V = 2 \cdot 10^{-12}$  см<sup>2</sup>. Небольшие различия в найденных значениях параметров функции  $G(x, z)$  по сравнению с данными работы [3] объясняются выбором несколько иной модели рассеяния электронов

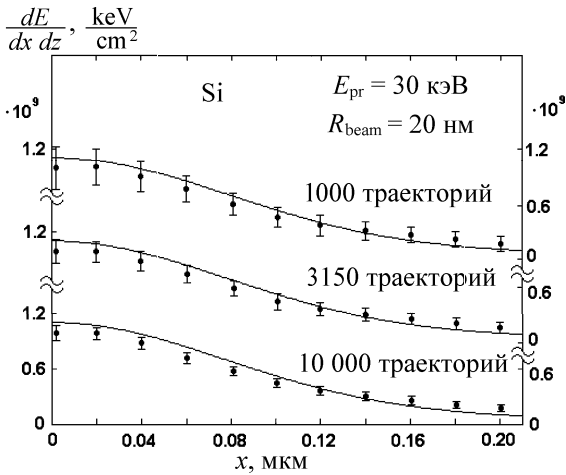


Рис. 2. Аппроксимация функции ППЭ  $G(x, z) = dE/(dx dz)$  при падении электронов с  $E_{pr} = 30$  кэВ в однородный кремниевый образец (сечение вдоль прямой  $z = 0.5$  мкм). Значения параметров те же, что и на рис. 2

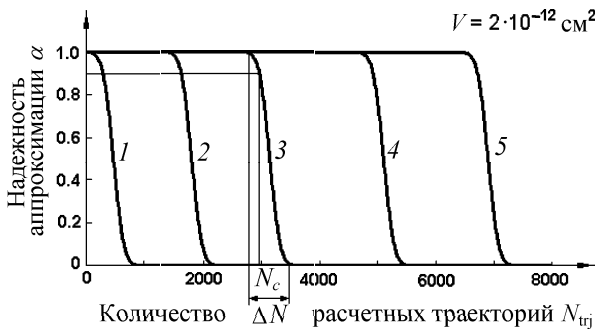


Рис. 3. Зависимость надежности аппроксимации  $\alpha$  от количества расчетных траекторий  $N_{trj}$  при различных  $E_{pr}$ , кэВ: 1 — 10, 2 — 20, 3 — 30, 4 — 40, 5 — 50. До тех пор пока аппроксимирующие кривые лежат внутри шумового коридора (рис. 1, 2),  $\alpha$  близка к единице. Как только они оказываются вне его,  $\alpha$  скачкообразно падает до нуля

в одном испытании, при котором величина надежности аппроксимации  $\alpha$  равна 0.9. Для случая, представленного на рис. 1, 2, значение  $N_c$  оказывается равным  $\approx 3150$ . К особенностям поведения функции  $\alpha(N_{trj})$  (рис. 3) относится то, что изменение значения надежности от единицы до нуля происходит скачком. Расчет показывает, что ширина интервала  $\Delta N$ , на котором оно происходит, мала по сравнению со значением критического числа расчетных траекторий  $N_c$  и по порядку величины составляет  $\approx 20 - 30$  траекторий. В связи с пуассоновским характером распределения числа событий при увеличении объема ячейки в  $m$  раз, значение  $N_c$  уменьшается в  $m$  раз, а  $\Delta N$  уменьшается в  $\sqrt{m}$  раз.

На рис. 4 представлены графики зависимости величины  $N_c$  от энергии первичного пучка  $E_{pr}$  при различных фиксированных значениях объема ячейки  $V$ . Видно, что  $N_c \sim E_{pr}^{5/3}$ . Происхождение такой зависимости может быть объяснено с помощью соотношения

$$s_0 = \left( E_{pr}^{5/3} A \right) / \left( 41 \rho Z^{4/5} \right), \quad (3)$$

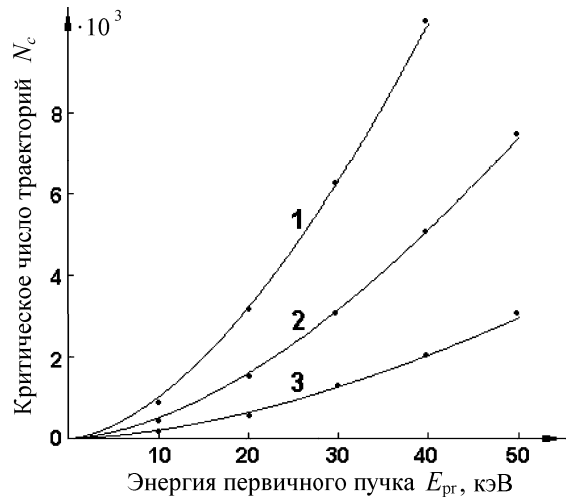


Рис. 4. Зависимости  $N_c$  от  $E_{pr}$  при различных фиксированных значениях объема ячейки разбиения  $V$ : 1 —  $V = 10^{-12}$  см<sup>2</sup>, 2 —  $V = 2 \cdot 10^{-12}$  см<sup>2</sup>, 3 —  $V = 5 \times 10^{-12}$  см<sup>2</sup>. Точками показаны значения  $N_c$ , найденные путем расчета, сплошными линиями — кривые  $N_c \sim E_{pr}^{5/3}$

где  $s_0$  — полный пробег электрона в мишени,  $E_{pr}$  — энергия первичного электрона,  $\rho$  — плотность вещества мишени,  $A$  — атомная масса,  $Z$  — атомный номер элемента [3]. Формула (3) показывает, что с ростом  $E_{pr}$  длина пути  $s_0$ , вдоль которого распадается энергия падающей частицы, возрастает по закону  $\sim E_{pr}^{5/3}$ . Следовательно, плотность энерговыделения вдоль траектории, нормированная на  $E_{pr}$ , убывает  $\sim E_{pr}^{-5/3}$ . Число событий (взаимодействий) одного электрона в ячейке напрямую связано с этой плотностью и подчиняется тому же закону  $\sim E_{pr}^{-5/3}$ , вследствие чего и возникает зависимость  $N_c \sim E_{pr}^{5/3}$ .

Расчет также показывает, что величина  $\Delta N$  интервала скачкообразного изменения надежности аппроксимации от 1 до 0 практически не зависит от энергии  $E_{pr}$  и для объема ячейки  $V$ , такого же, как на рис. 1 и 2, по порядку величины составляет 20–30 траекторий.

### Заключение

В работе с помощью метода максимальной надежности проведен анализ взаимосвязи погрешности функций, получаемых в стохастическом эксперименте, и качества их гладкой аппроксимации. В качестве меры близости предложено использовать надежность. Рассмотрен пример из электронной микроскопии: аппроксимация ППЭ электрона, рассчитываемой с помощью имитационного моделирования методом Монте-Карло, суммой двух трехмерных гауссовых функций. Проиллюстрировано, что выбор аналитического описания согласуется с наблюдаемыми данными только до тех пор, пока случайная компонента погрешности превосходит систематическую; значение надежности аппроксимации при этом близко к единице. По мере накопления статистики роль случайной составляющей погрешности уменьшается и надежность монотонно спадает до нуля.

Количество траекторий  $N_c$ , характеризующее верхнюю границу адекватности аналитического описания, обратно пропорционально объему ячейки разбиения  $V$  и подчиняется закону  $N_c \sim E_{\text{пр}}^{5/3}$ , где  $E_{\text{пр}}$  — энергия первичного электронного пучка.

В заключение автор выражает огромную благодарность профессору кафедры компьютерных методов физики Чуличкову Алексею Ивановичу за помощь в постановке задачи и обсуждении результатов.

#### Литература

1. *Валиев К.А., Раков А.В.* Физические основы субмикронной литографии в микроэлектронике. М., 1984.
2. *Михеев Н.Н., Петров В.И., Степович М.А.* // Изв. АН СССР. Сер. Физ. 1991. **55**, № 8. С. 1474.
3. *Werner U., Koch F., Olegart G.* // J. Phys. D: Appl. Phys. 1988. **21**. P. 116.
4. *Пытьев Ю.П., Сухорукова Г.В., Чуличков А.И.* // Мат. моделирование. 1994. **6**, № 11. С. 113.
5. *Борисов С.С., Грачев Е.А., Негуляев Н.Н.* и др. // Прикладная физика. 2004. № 1. С. 118.
6. *Калиткин Н.Н.* Численные методы. М., 1978.

Поступила в редакцию  
20.06.04