

УДК 519.254

ВЕЙВЛЕТ-АНАЛИЗ: КРИТЕРИЙ ДОСТОВЕРНОСТИ ВЫДЕЛЕНИЯ СИГНАЛА

М.С.Хвастунов

Предложен критерий достоверности выделения сигнала в условиях большой фоновой загрузки.

Работа выполнена в Лаборатории высоких энергий ОИЯИ.

Wavelet-Analysis: Criterion of Reliable Signal Selection

M.S.Khvastunov

Criterion is presented to select the Gaussian signal in a wide interval of the effect/background ratio.

The investigation has been performed at the Laboratory of High Energies, JINR.

1. ВВЕДЕНИЕ

Вейвлет-анализ является сравнительно новым математическим аппаратом [1]. Он предназначен для анализа информации, содержащей сигналы в виде локализованных пиков. Часто именно такого рода данные получают в экспериментах по физике высоких энергий.

Целью данной работы являлась разработка критерия достоверности выделения сигнала при вейвлет-анализе в условиях большой фоновой загрузки.

В данной работе используются гауссовы вейвлеты (производные функции Гаусса). Анализируемые сигналы задавались в гауссовой форме. Поставленная задача решалась путем моделирования экспериментальной ситуации методом Монте-Карло.

2. ВЕЙВЛЕТ-ОБРАЗ АНАЛИЗИРУЕМОГО СИГНАЛА

В предыдущей работе автора [2] была обнаружена высокая стабильность площади вейвлет-образа сигнала в широком интервале отношения эффект — фон. Под площадью здесь понимается интеграл по модулю от вейвлет-образа сигнала. Стабильность площади вейвлет-образа сигнала — следствие фильтрационного свойства вейвлет-преобразования.

Величина площади $SW_{g\phi_1}$ вейвлет-образа 1-го порядка сигнала в присутствии фона наиболее отличается от площади SW_{ϕ_1} вейвлет-образа 1-го порядка фонового сигнала в

сравнении с площадями вейвлет-образов более высоких порядков [2]. Поэтому в данной работе использовались вейвлет-образы 1-го порядка.

Для численной характеристики отношения эффект/фон вводился параметр

$$k = w \cdot N_g / (wN_g + N_{g\phi}), \quad (1)$$

где $N_{g\phi}$ — число фоновых событий в тех ячейках гистограммы, в которых присутствуют wN_g событий анализируемого сигнала; w — доля от N_g событий гауссова сигнала, попадающих в анализируемый участок гистограммы v . Например, при длине анализируемого участка гистограммы $v = 4,46\sigma(\pm 2,23\sigma)$ величина w , определяемая по таблице интеграла вероятностей, равна 0,9743. Соответствующий величине v интервал гистограммы вейвлет-образа сигнала определяется выражением

$$v_w/\sigma = 2\sqrt{(v/2\sigma)^2 - 2 \cdot \ln(1-w)}. \quad (2)$$

Моделирование гауссова сигнала производилось на компьютере с помощью генератора случайных чисел, распределенных в соответствии с функцией Гаусса. Центр модельного гауссова сигнала помещался в среднюю ячейку гистограммы. Параметру k придавались значения: $k_i = 1,0; 0,9; 0,85; 0,8; 0,7; 0,6; 0,5; 0,4; 0,3; 0,2$. При каждом значении k_i числам N_g придавались значения $N_{gj} = 20, 30, 40, \dots, 110, 120$. При этом числа фоновых (равномерно распределенных) событий $N_{g\phi}$ определялись выражением

$$N_{g\phi}(i, j) = av \cdot w \cdot (1/k_i - 1) \cdot N_{gj}, \quad (3)$$

где $av = v_{\max}/v$, v_{\max} — максимальное число ячеек в гистограмме анализируемого сигнала. Числа $N_{g\phi}(i, j)$ получались с помощью генератора равномерно распределенных случайных чисел. Информация о модельных сигналах (гауссовом и фоновом) суммировалась в одной гистограмме. Для каждой пары чисел (N_{gj}, k_i) выполнялось по 100 статистически независимых численных экспериментов на компьютере. Далее вычислялись площади $SW_{g\phi 1}(i)$ вейвлет-образов модельных гауссовых сигналов, в той или иной мере (определяемой параметром k_i) "подпорченных" равномерным фоном. Для каждого значения k_i по 100 значениям площади $SW_{g\phi 1}(i)$ вычислялись средние значения этой площади и среднеквадратическая ошибка отдельного измерения этой площади. Зависимость этих величин (средней площади $SW_{g\phi 1}$ и ошибки $\Delta SW_{g\phi 1}$) от N_g для каждого значения k_i хорошо аппроксимируется полиномами. В табл. 1 приведены значения коэффициентов этих полиномов.

Площадь $SW_{g\phi 1}$ вейвлет-образа линейно зависит от N_g [2]:

$$SW_{g\phi 1} = S_1 + S_2 \cdot N_g, \quad (4)$$

где коэффициенты S_1 и S_2 — из табл. 1, а $N_g = 20, 30, 40, \dots, 110, 120$. Параметр наклона S_2 слабо зависит от примеси фона в сигнале (от параметра k_i). Для интервала k_i от 0,3 до 1,0 среднее значение этого параметра равно $\langle S_2 \rangle = 0,4078 \pm 0,0035$ и среднее значение параметра $\langle S_1 \rangle = 1,1502 \pm 0,6193$ (здесь приведены среднеквадратические ошибки отдельного измерения).

Среднее значение параметра наклона S_2 несколько ниже приведенного в работе [2], поскольку здесь использовано меньшее значение вероятности w .

Таблица 1. Коэффициенты S_1, S_2 и p_1, p_2, p_3 полиномов $SW_{g\phi 1} = S_1 + S_2 \cdot N_g$ и $\Delta SW_{g\phi 1} = p_1 + p_2 \cdot N_g + p_3 \cdot N_g^2$

k_i	S_1	S_2	p_1	p_2	p_3
1,00	0,56609	0,40984	0,61556	$1,9755 \cdot 10^{-2}$	$-3,7935 \cdot 10^{-5}$
0,90	0,61265	0,41174	0,47773	$3,6085 \cdot 10^{-2}$	$-1,4665 \cdot 10^{-4}$
0,85	0,80228	0,40646	0,85261	$2,0815 \cdot 10^{-2}$	$-1,3189 \cdot 10^{-5}$
0,80	0,66546	0,41029	0,95480	$2,7895 \cdot 10^{-2}$	$-5,4807 \cdot 10^{-5}$
0,70	1,0344	0,40624	0,90228	$3,7558 \cdot 10^{-2}$	$-8,5336 \cdot 10^{-5}$
0,60	0,82157	0,40871	1,1955	$4,1772 \cdot 10^{-2}$	$-6,9662 \cdot 10^{-5}$
0,50	1,5670	0,40446	1,1870	$5,5469 \cdot 10^{-2}$	$-1,8017 \cdot 10^{-4}$
0,40	2,0940	0,40104	0,96759	$7,2269 \cdot 10^{-2}$	$-2,0780 \cdot 10^{-4}$
0,30	2,1185	0,41120	0,9420	$1,0040 \cdot 10^{-1}$	$-4,0188 \cdot 10^{-4}$
0,20	4,1138	0,39629	1,4338	$1,1533 \cdot 10^{-1}$	$-3,8713 \cdot 10^{-4}$

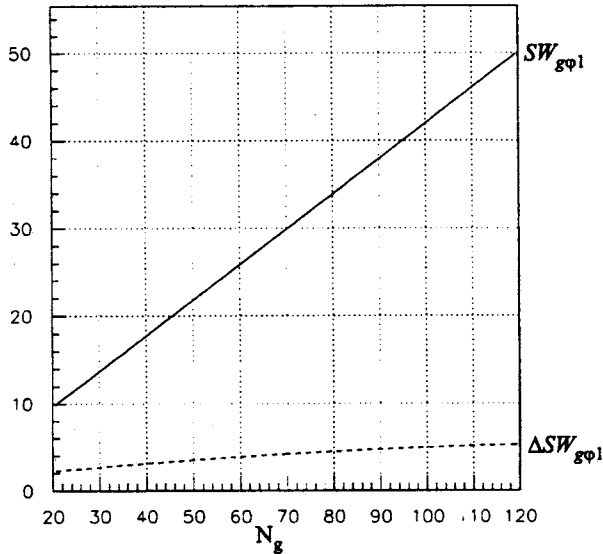


Рис. 1. Зависимость от N_g — числа событий в гауссовом сигнале — площади $SW_{g\phi 1}$ вейвлет-образа сигнала и ошибки $\Delta SW_{g\phi 1}$ этой площади при $k_i = 0,5$

Ошибка площади вейвлет-образа сигнала вычислялась по формуле

$$\Delta SW_{g\phi 1} = p_1 + p_2 \cdot N_g + p_3 \cdot N_g^2, \quad (5)$$

где коэффициенты берутся из табл. 1.

На рис.1 приведены в качестве примера зависимости (от числа N_g событий в гауссовом сигнале) площади $SW_{g\phi_1}$ вейвлет-образа сигнала и ошибки $\Delta SW_{g\phi_1}$ этой площади при $k_i = 0,5$.

3. ВЕЙВЛЕТ-ОБРАЗ ФОНОВОГО СИГНАЛА

Моделирование фонового сигнала производилось на компьютере с помощью генератора равномерно распределенных случайных чисел. Числа N_ϕ фоновых событий выбирались в соответствии с условием: в пределах $\pm v/2$ гистограммы сигнала числа $N_{\phi v} = w \cdot N_g + N_{g\phi}$ (см. выражение (1)), а в пределах $\pm v_{\max}/2$:

$$N_{\phi s}(i, j) = av \cdot w \cdot N_{gj}/k_i. \quad (6)$$

Аналогично изложенному в предыдущем разделе для каждой пары чисел (N_{gj}, k_i) выполнялось по 100 статистически независимых численных экспериментов на компьютере. Вычислялись площади $SW_{\phi_1}(i)$ вейвлет-образов фоновых сигналов. Для каждого значения k_i по 100 значениям площади $SW_{\phi_1}(i)$ вычислялись средние значения этой площади и среднеквадратическая ошибка $\Delta SW_{\phi_1}(i)$ отдельного измерения площади $SW_{\phi_1}(i)$. Зависимость этих величин ($SW_{\phi_1}(i)$ и $\Delta SW_{\phi_1}(i)$) от N_g для каждого значения k_i аппроксимировалась полиномами. В табл. 2 приведены соответствующие значения коэффициентов этих полиномов.

Таблица 2. Коэффициенты q_1, q_2, q_3 и r_1, r_2, r_3 полиномов $SW_{\phi_1} = q_1 + q_2 \cdot N_g + q_3 \cdot N_g^2$ и $\Delta SW_{\phi_1} = r_1 + r_2 \cdot N_g + r_3 \cdot N_g^2$

k_i	q_1	q_2	q_3	r_1	r_2	r_3
1,00	2,1495	$0,93716 \cdot 10^{-1}$	$-2,2369 \cdot 10^{-4}$	0,58959	$3,2027 \cdot 10^{-2}$	$-7,4229 \cdot 10^{-5}$
0,90	2,3669	$0,95507 \cdot 10^{-1}$	$-2,3332 \cdot 10^{-4}$	0,86912	$2,5605 \cdot 10^{-2}$	$-3,0490 \cdot 10^{-5}$
0,85	2,5340	$0,97644 \cdot 10^{-1}$	$-2,3609 \cdot 10^{-4}$	0,87950	$3,1298 \cdot 10^{-2}$	$-8,5636 \cdot 10^{-5}$
0,80	1,9970	0,12701	$-4,5922 \cdot 10^{-4}$	0,59002	$4,4591 \cdot 10^{-2}$	$-1,7541 \cdot 10^{-4}$
0,70	2,2048	0,13552	$-4,9518 \cdot 10^{-4}$	0,72375	$4,5018 \cdot 10^{-2}$	$-1,8756 \cdot 10^{-4}$
0,60	2,6949	0,13063	$-3,9413 \cdot 10^{-4}$	0,62504	$5,0250 \cdot 10^{-2}$	$-1,8583 \cdot 10^{-4}$
0,50	3,0336	0,13256	$-3,5148 \cdot 10^{-4}$	0,90117	$4,6703 \cdot 10^{-2}$	$-1,5120 \cdot 10^{-4}$
0,40	3,3558	0,15964	$-4,9178 \cdot 10^{-4}$	0,90276	$5,1379 \cdot 10^{-2}$	$-1,2819 \cdot 10^{-4}$
0,30	3,2883	0,19582	$-6,3315 \cdot 10^{-4}$	0,99721	$6,8826 \cdot 10^{-2}$	$-2,6190 \cdot 10^{-4}$
0,20	5,1338	0,20196	$-4,6287 \cdot 10^{-4}$	1,9489	$6,2345 \cdot 10^{-2}$	$-1,7505 \cdot 10^{-4}$

Площадь вейвлет-образа фонового сигнала вычисляется по формуле

$$SW_{\phi_1} = q_1 + q_2 \cdot N_g + q_3 \cdot N_g^2, \quad (7)$$

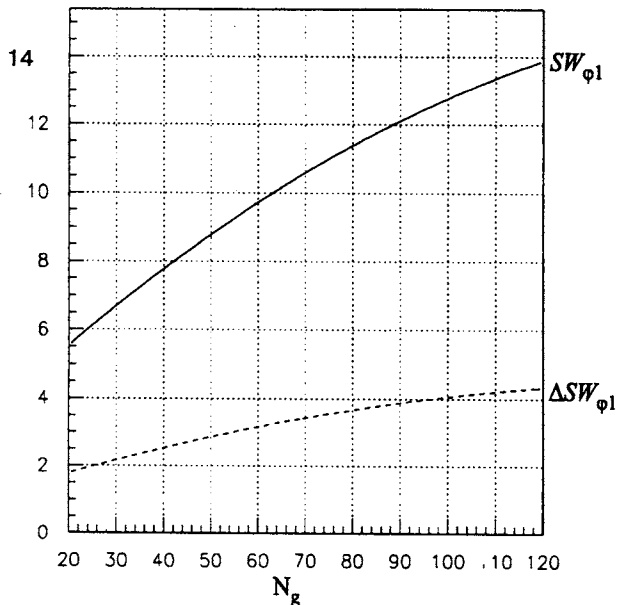


Рис. 2. Зависимость от N_g — числа событий в гауссовом сигнале — площади $SW_{\phi 1}$ вейвлет-образа фонового сигнала и ошибки $\Delta SW_{\phi 1}$ этой площади при $k_i = 0,5$

и ошибка этой площади —

$$\Delta SW_{\phi 1} = r_1 + r_2 \cdot N_g + r_3 \cdot N_g^2, \quad (8)$$

где коэффициенты $q_i, r_i, i = 1, 2, 3$ берутся из табл. 2.

На рис.2 приведены в качестве примера кривые $SW_{\phi 1}(i)$ и $\Delta SW_{\phi 1}(i)$, рассчитанные по формулам (7) и (8) для $k_i = 0,5$.

4. КРИТЕРИЙ ДОСТОВЕРНОСТИ ВЫДЕЛЕНИЯ СИГНАЛА

Теперь у нас достаточно информации для конструирования параметра, характеризующего вероятность имитации гауссова сигнала фоном. Для этой цели используем разность $d = SW_{g\phi 1} - SW_{\phi 1}$ и среднеквадратическую ошибку этой разности $\sigma_d = \sqrt{(\Delta SW_{g\phi 1})^2 + (\Delta SW_{\phi 1})^2}$. На рис.3 представлены кривые зависимостей отношения d/σ_d от N_g для k_i от 0,2 до 1,0.

Отношения d/σ_d определяют вероятность имитации гауссова сигнала фоном. Например, при $d/\sigma_d = 3,0$ эта вероятность равна 0,0027. Напомним, что величинам $SW_{g\phi 1}$ и $SW_{\phi 1}$ соответствуют одинаковые числа событий: $N_{g\phi}(\pm v_{\max}/2)$ событий (гауссова сигнала вместе с фоновым сигналом) при вычислении $SW_{g\phi 1}$ и $N_{\phi}(\pm v_{\max}/2) = N_{g\phi}(\pm v_{\max}/2)$ фоновых событий при вычислении $SW_{\phi 1}$.

В табл. 3 приведены для разных фоновых условий (различных значений k_i) и для различных отношений d/σ_d минимальные числа N_g событий в гауссовом сигнале. Так,

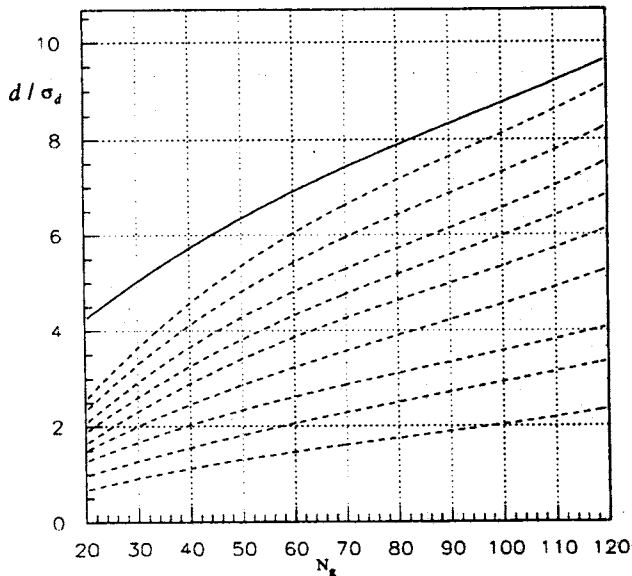


Рис. 3. Зависимость d/σ_d от N_g — числа событий в гауссовом сигнале для различных значений k_i (пунктирные кривые: снизу вверх — соответствуют $k_i = 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,85; 0,9$; верхняя сплошная кривая — $k_i = 1,0$)

Таблица 3. Числа N_g событий в гауссовом сигнале при разных фоновых условиях (различные значения k_i) и разных величинах d/σ_d

d/σ_d	k_i									
	1,0	0,9	0,85	0,8	0,7	0,6	0,5	0,4	0,3	0,2
2,0						25	30	40	58	98
3,0		24	26	32	36	42	54	75	104	
4,0		34	38	45	54	64	84	116		
5,0	30	45	53	64	76	90	114			
6,0	44	60	70	86	100	116				

для $d/\sigma_d = 3,0$: $N_g \geq 104(k_i = 0,3)$; $N_g \geq 42(k_i = 0,6)$; $N_g \geq 26(k_i = 0,85)$ и для $d/\sigma_d = 5,0$: $N_g \geq 114(k_i = 0,5)$; $N_g \geq 53(k_i = 0,85)$ и т. д.

Число N_g событий в гауссовом сигнале однозначно определяется площадью $SW_{g\phi 1}$ вейвлет-образа этого сигнала (см. формулу (4)):

$$N_g = (SW_{g\phi 1} - S_1)/S_2 \simeq SW_{g\phi 1}/S_2. \quad (9)$$

Параметр S_2 (как уже отмечалось выше) практически постоянен в широком интервале значений k_i . Поэтому ошибку в N_g можно представить в виде

$$\Delta N_g/N_g \simeq \Delta SW_{g\phi 1}/SW_{g\phi 1}. \quad (10)$$

Таблица 4. Параметры $\Delta N_g/N_g$ в зависимости от d/σ_d и k_i

k_i	d/σ_d			
	3,0	4,0	5,0	6,0
0,30	0,157/104			
0,40	0,162/75	0,135/116		
0,50	0,156/54	0,129/84	0,108/114	
0,60	0,157/42	0,132/64	0,116/90	0,106/116
0,70	0,137/36	0,122/54	0,102/76	0,091/100
0,80	0,130/32	0,110/45	0,094/64	0,082/86
0,85	0,122/26	0,100/38	0,086/53	0,077/70
0,90	0,120/24	0,105/34	0,094/45	0,070/60

В табл. 4 приведены эти ошибки и соответствующие значения N_g в зависимости от d/σ_d и k_i . Например, при $d/\sigma_d=5,0$ и $k_i=0,5$ $\Delta N_g/N_g \approx 0,108$ и $N_g=104$.

5. ПРИМЕНЕНИЕ КРИТЕРИЯ ДОСТОВЕРНОСТИ ВЫДЕЛЕНИЯ СИГНАЛА В РЕАЛЬНОМ ЭКСПЕРИМЕНТЕ

В реальном эксперименте мы знаем лишь площадь $SW_{g\phi_1}(e)$ вейвлет-образа зарегистрированного сигнала. Естественно возникает вопрос — какова вероятность имитации этого сигнала флуктуациями равномерного фона? Как нам представляется, на этот вопрос можно ответить, используя изложенное в предыдущих разделах данной работы.

Вычисление площадей вейвлет-образов экспериментальных сигналов должно производиться так же, как модельных (при тех же значениях параметров: dm/σ — отношения ширины ячейки гистограммы сигнала к σ — и при тех же значениях v , w , v_w , и v_{max}). При этом можно воспользоваться формулой (4) для вычисления числа N_{ge} событий в экспериментальном сигнале:

$$N_{ge} = (SW_{g\phi_1}(e) - S_1)/S_2, \quad (11)$$

где S_1 и S_2 — средние значения параметров, приведенные в разделе 2. Пусть полное число событий в гистограмме экспериментального сигнала в пределах $\pm v/2$ равно N_{se} . Среди этого числа событий содержится N_{ge} событий возможного гауссова сигнала. Экспериментальное значение параметра k_e определяется отношением $k_e = N_{ge}/N_{se}$. Среди набора модельных параметров k_i (1,0; 0,9; 0,85; 0,8; 0,7; 0,6; 0,5; 0,4; 0,3; 0,2) выбираем то значение k_{ei} , которое наиболее близко к параметру k_e . Пользуясь этим значением k_{ei} , выбираем из табл. 1 и 2 соответствующие коэффициенты полиномов: p_i , q_i , r_i , $i = 1, 2, 3$. Пользуясь этими коэффициентами, по формулам (5,7,8) вычисляем значения $\Delta SW_{g\phi_1}(e)$, $SW_{\phi_1}(e)$ и $\Delta SW_{\phi_1}(e)$. Далее вычисляем значения параметров:

$$d_e = SW_{g\phi_1}(e) - SW_{\phi_1}(e), \quad \sigma_{de} = \sqrt{(\Delta SW_{g\phi_1}(e))^2 + (\Delta SW_{\phi_1}(e))^2} \quad (12)$$

и находим их отношение d_e/σ_{de} .

Напомним, что величины $SW_{\phi_1}(e)$ и $\Delta SW_{\phi_1}(e)$ вычислены в предположении, что N_{se} событий экспериментального сигнала представляют собой равномерный фон. Величина d_e/σ_{de} определяет вероятность имитации экспериментального пика флуктуациями равномерного фона.

6. ЗАКЛЮЧЕНИЕ

Фильтрационное свойство вейвлет-преобразования дает возможность выработки критерия достоверности выделения гауссова сигнала в условиях большой фоновой загрузки.

Автор благодарен В.В.Глаголеву за обсуждения, стимулировавшие разработку данного критерия.

Литература

1. Астафьева Н.М. — УФН, 1996, т.166, №11, с. 1145.
2. Хвастунов М.С. — Краткие сообщения ОИЯИ, 1998, №6[92]-98, с.63.