A Wide-Band Impulse-Noise Survey on Subscriber Lines and Inter-Office Trunks — Modeling and Simulation —

Werner Henkel, Thomas Kessler
Research Center of Deutsche Telekom
D-64276 Darmstadt, Germany
EMAIL: henkel@fz.telekom.de, kessler@...

Hong Y. Chung
AT&T Bell Laboratories
Holmdel, NJ 07733, USA
hychung@attmail.com

Abstract — Alternatives of a simulation model for impulse noise on subscriber lines are proposed and their statistical properties are outlined. Several pseudo-noise generators are combined with spectral shaping operations to approximate the desired statistics of the non-stationary disturbance. Although the considerations are based on the special statistical properties of impulse noise, the applied methods may as well be applicable to other modeling purposes. Especially, a proposal that allows to generate samples that fulfill the required amplitude statistics and bring forth a prescribed power density spectrum may be suitable elsewhere, too.

I. INTRODUCTION

In [1, 2, 3], three statistical properties of impulse noise on telephone lines are described: voltage, length, and inter-arrival time densities. Furthermore, the mean power density spectrum (PDS) and the typical group delay (phase difference) function are outlined. Exemplarily, the relation between the length and the energy of impulses is additionally presented here (Fig. 4). The goal of a simulation model is to meet as much of these properties as possible, while minimizing the realization effort. Principal simulation methods are:

1. representative impulse
2. data bank access with sampled impulses
3. Markov model
4. generator model

A *representative* impulse like the one by British Telecom (Cook, [4]) that has also been included in the ANSI and ETSI standards or even the one shown in our papers [2, 3] cannot represent any density functions. Thus, they are not very suitable for our modeling purposes. Often, they are also used with some chosen fixed inter-arrival time. Using a generator with the real inter-arrival time statistics would bring this method a little nearer to reality.

The *data bank access* with real measured impulses triggered by a stochastic inter-arrival time generator works quite well, but is dependent on the data bank, which may not be generally available. Furthermore, it demands for high storage capacities.

The *Markov model* is not very handy, because too many states would be needed, leading to a large number of free parameters, whose meaning may not be very obvious.

This paper describes a *generator model*, where the statistics are represented by pseudo-noise generators. Other components (filtering, etc.) ensure the spectral and phase properties. This approach corresponds directly to the determined statistics and the number of free parameters is not too high. No collection of data is needed. The disadvantage is that it is very hard to find a configuration that delivers samples that, at least, approximately comply with *all* the desired statistics. This is due to the quite large number of statistical properties that are considered. Therefore, based on one principal structure, alternatives will be described, that are in favor of some of the properties, while others cannot be obtained as well. The choice between them depends on the application.

Before describing the structure and its alternative realizations in some detail, generators for the three statistical properties, voltage, length, and inter-arrival time density, will be presented. Then, the inclusion of spectral properties in the form of the mean PDS are discussed. Finally, simulation schemes with and without the desired phase properties are compared.

II. GENERATOR REALIZATIONS FOR THE STATISTICS OF IMPULSE NOISE

A. *Generators with the described impulse-noise statistics*

We shall describe possible realizations of pseudo-noise generators for computer simulation that will follow the impulse-noise densities (voltage, inter-arrival time, length). The generators utilize well-known principal methods or variations thereof – 'transform method', 'rejection method', 'strip method'.

The *transform method* utilizes the density change when applying a function $g(x)$ to a random variable. The density is modified according to

$$f_y(y) = f_y(g(x)) = f_x(x) / \left| \frac{dg(x)}{dx} \right| . \qquad (1)$$

The **generator for the voltage samples of the impulses** with the desired density

$$f_i(u) = \frac{1}{240 u_0} e^{-|u/u_0|^{1/5}} , \qquad (2)$$

uses the *rejection method* (see [5], pp. 203-206 or [6], pp. 120-121). Herein, one selects a function $k \cdot f_\cap(x)$ according to the condition

$$k \cdot f_\cap(x) \geq f_i(x) , \quad k \in \mathbb{R} . \tag{3}$$

This function should be easily integrable. Additionally, $F_\cap(x) = \int_{-\infty}^{x} f_\cap(\eta) d\eta$ should also be easily invertible. The procedure works as follows: x is obtained by applying $x = F_\cap^{-1}(\xi)$ to some uniform (0,1) random number ξ. This can be seen as an application of the transform method, because $f_\cap(x) = f_\cap(F_\cap^{-1}(\xi)) = f_\xi(\xi) / \left| \frac{dF_\cap^{-1}(\xi)}{d\xi} \right| = 1 \cdot \left| \frac{dF_\cap(x)}{dx} \right|$. Then, $f_\cap(x)$ is multiplied with a constant k, such that condition (3) is fulfilled. A second uniform $(0, kf_\cap(x))$ generator determines, if x is taken as output or being rejected. Rejection takes place, when the second uniformly distributed random number is greater than $f_i(x)$. The exact value of k does not influence the out-coming distribution, but the speed of the random number generation. The farer $kf_\cap(x)$ is apart from the desired $f_i(x)$, the higher the percentage of rejected values. Figure 1a describes the procedure.

For the voltage-generator application, we selected $kf_\cap(x)$ and $kF_\cap(x)$ to be

$$\begin{aligned} kf_\cap(x) &= k\frac{c/\pi}{1+c^2x^2} \quad \text{and} \\ kF_\cap(x) &= k(\text{arc}\ \tan(cx) + \pi/2)/\pi . \end{aligned} \tag{4}$$

The constants k and c are chosen such that condition (3) is fulfilled in a certain range, $kf_\cap(0) = f_i(0)$, and $kf_\cap(x)$ is as small as possible.

The **generator for the inter-arrival times** with the density

$$f_d(x) = \frac{10^{a_1}}{\ln(10)} x^{a_4 - 1} 10^{-\frac{a_4}{\ln(a_2)} a_2^{(\log_{10}(x) - a_3)}} \tag{5}$$

($x = t/100$ ns) is based on a modified strip method. In its original form (see, e.g., [7], pp. 359-368, [6], pp. 118-122) the PDF is divided into N_S strips and the corresponding probabilities $p^{(\nu)}$, $0 \leq \nu \leq N_S - 1$, to obtain a value within a strip, are determined. $p^{(\nu)}$ equals the area of a strip ν. A special random number generator is used to select the strips according to the $p^{(\nu)}$ and afterwards any other procedure is taken to approximate the exact shape of the density within one strip.

For our application, we employ exponentially spaced boundaries (see, Fig. 1b) in order to divide the PDF into strips, which again are divided into rectangular regions (probability $p_\square^{(\nu)}$) always below the PDF and triangular ones (probability $p_\triangle^{(\nu)}$), featuring a piecewise linear approximation of the PDF. The strips (regions) are chosen by a random number generator due to Walker [8] (see, below). If a rectangular strip is met, a uniform number generator is used to generate the actual output between the strip boundaries, whereas in the case of a triangular strip, additionally, the transform method with $x(\xi) = \sqrt{\xi}$ is applied.

Maybe, a sketch of Walker's method should be given shortly, too. First, the desired discrete probabilities $(p^{(\nu)})$ are compared with a uniform distribution $(1/N_S)$. The greatest negative (C) and positive differences (D) between $p^{(\nu)}$ and $1/N_S$ are searched. For the position $\nu = \nu_C$ of the greatest negative difference, an 'alias' value ν_D is stored, which is output as an alternative to the original value that corresponds to a certain probability. The greatest negative difference is added to the positive one, thereby reducing this difference. Additionally, a threshold, defined by $1 - CN_S$, is stored. As $1 - CN_S = N(1/N_S - C)$, this threshold corresponds to the difference C. Now, the difference C at position ν_C is reduced to zero. The search for greatest negative (C) and positive differences (D) is repeated until all negative differences have been processed (at most $N_S - 1$ times).

Walker's generator itself is based on some sort of rejection method, in the sense that it accepts values from a uniformly distributed discrete generator, if a second uniformly distributed continuous generator delivers values less than the threshold. If the threshold is exceeded, the 'alias' value is output. For a Fortran routine, see [8].

The realization of the **generator for the impulse lengths** with the density

$$\begin{aligned} f_l(t) = \ & B\frac{1}{\sqrt{2\pi}s_1 t}\ e^{-\frac{1}{2s_1^2}\ln^2(t/t_1)} + \\ & (1-B)\frac{1}{\sqrt{2\pi}s_2 t}\ e^{-\frac{1}{2s_2^2}\ln^2(t/t_2)} \end{aligned} \tag{6}$$

is based on a normal (0,1) generator with the density $f_g(x)$ (e.g., polar method by Box, Muller et al.; see [6], pp. 117-118) together with an exponential function in the transform method:

$$\begin{aligned} f_g(x) &= \frac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}, \\ t &= e^{s_\nu x + \ln(t_\nu)}, \\ dt/dx &= s_\nu \cdot t, \end{aligned}$$

applying (1):

$$f_l^{(\nu)}(t) = \frac{1}{\sqrt{2\pi}s_\nu t}\ e^{-\frac{1}{2s_\nu^2}\ln^2(t/t_\nu)}, \quad \nu = 1, 2 . \tag{7}$$

The addition of two log-normal densities (Eqn. (6)) is taken into account by choosing the two parameter sets (t_ν, s_ν) randomly with the probabilities B and $1 - B$, respectively.

B. *Incorporating spectral properties*

A simple way to ensure spectral properties is to use the noise generators described in the previous section followed by an FIR filter, or equivalently, by an FFT and a multiplication in frequency domain. The coefficients of the FIR filter are determined from the square root of the mean PDS of a certain location or of an approximation thereof. About 200 coefficients are necessary.

Since the voltage density is changed by the FIR filter, a PDF $f_1(u)$ at the input must be searched that yields the desired PDF $f_i(u)$ at the output. The PDF $f_{out}(u)$ at the output of an FIR filter can be derived as follows:

A multiplication with one coefficient k_ν changes the density according to

$$f_1'(u) = \frac{f_1(u/k_\nu)}{|k_\nu|} \ . \tag{8}$$

The addition of n, e.g., $n = 200$, random processes results in the convolution of their densities, which would lead to $n-1$ such convolutions

$$f_{out}(u) = \frac{1}{\prod_{\nu=1}^{n} |k_\nu|} \int_{\eta_{n-1}} \cdots \int_{\eta_2} \int_{\eta_1}$$
$$f_1\left(\frac{\eta_{n-1}}{k_n}\right) f_1\left(\frac{\eta_{n-2}-\eta_{n-1}}{k_{n-1}}\right) \cdots f_1\left(\frac{\eta_j - \sum_{\mu=j+1}^{\mu=n-1} \eta_\mu}{k_{j+1}}\right) \cdots$$
$$\cdots f_1\left(\frac{u-\sum_{\mu=1}^{\mu=n-1} \eta_\mu}{k_1}\right) d\eta_1\, d\eta_2 \ldots d\eta_{n-1} \ . \tag{9}$$

This corresponds to a multiplication of the corresponding characteristic functions (the Fourier transforms) $F_1(k_\nu \omega)$

$$F_{out}(\omega) = \prod_{\nu=1}^{\nu=n} F_1(k_\nu \omega) \ . \tag{10}$$

Although (10) is a simplification of (9), it is not possible to construct any analytic expression for f_1, applying this equation. A feasible solution is the superposition of a Dirac onto the PDF of the input process. In our case, the Dirac approximately regenerates the desired PDF. The PDF at the input of the filter is changed according to

$$f_1(u) = \epsilon f_i(u) + (1-\epsilon)\delta(u) \qquad \epsilon \in [0.01, 0.1] \tag{11}$$

Note that the convolution of a Dirac with an arbitrary function yields the function itself. Assuming a dominant Dirac component as defined in (11), a convolution of two such functions (transformed with k_ν) results in a sum of four components: a Dirac multiplied with $(1-\epsilon)^2$, the two functions $\frac{f_i(u/k_\nu)}{|k_\nu|}$ themselves, each multiplied with $\epsilon(1-\epsilon)$, and a mixed convolution of the two functions $\frac{f_i(u/k_\nu)}{|k_\nu|}$ multiplied with ϵ^2. n such convolutions reduce the pure Dirac component by a factor of $(1-\epsilon)^n$. The characteristic function corresponding to (11) is given by

$$F_1(\omega) = \epsilon F_i(\omega) + (1-\epsilon) \ . \tag{12}$$

Inserting (12) into (10), yields

$$F_{out}(\omega) = \sum_{\nu=0}^{n} \left((1-\epsilon)^{n-\nu} \epsilon^{\nu} \sum_{(\zeta_1,\ldots,\zeta_\nu) \in C_\nu^n} \prod_{\mu=1}^{\nu} F_i(k_{\zeta_\mu} \omega) \right) , \tag{13}$$

where C_ν^n denotes the set of all combinations of ν elements out of a set of n elements. The pure Dirac component corresponds to the constant term for $\nu = 0$. Aside from the fact that the Dirac is reduced by the n-fold convolutions, it is quite difficult to prove that the shape of the density at the output of the filter $f_{out}(u)$ is nearly equal to the one of $f_i(u)$ after n convolutions. Only for small n, e.g., $n = 2, 3$, one still sees the dominance of the terms $f_1'(u)$ compared with convolution products thereof.

After all, the regeneration effect is due to the special shape of the density $f_i(x)$ and the special set of filter coefficients k_ν. Thus, the insertion of a Dirac into a density is not a generally applicable method.

Note that the coefficient 'u_0' will be different for the input and output densities. Hence, the corresponding coefficients at the input for a desired output density have to be determined in advance. Resulting frequency distributions will be shown in the next section.

For other principal approaches to include spectral properties, see [9].

Phase properties which have been described in [2, 3], can be ensured by setting the phase according to that function after an FFT and after the multiplication of the frequency response that corresponds to the mean PDS.

III. POSSIBLE SIMULATION SCHEMES

The basic structure of the simulation schemes is shown in Fig. 2. The inter-arrival time generator triggers the length generator[1] which activates the voltage generator during the length l. A further windowing with the same length l may be applied at the output. The voltage samples are transformed with an FFT. Then a multiplication with the square-root of the PDS ensures the spectral properties. Optionally, the resulting phase may be changed to the desired phase function. An IFFT yields the corresponding time function. In principal, different results are obtained if phase setting is applied or not.
The output windowing forces the output impulse to at most the length that is delivered from the length generator. If this windowing is used, one could also think of setting the activation of the voltage generator to some constant, quite long, time. Thus, six alternatives follow from Fig. 2:

	stochastic activation with windowing	stochastic activation without windowing	constant activation with windowing
without phase setting	sw	s	cw
with phase setting	spw	sp	cpw

The different possibilities are subsequently denoted with the given short-form notation.

All alternatives have in common that the mean PDS equals the desired PDS (slightly smoothed, if windowing is applied) and the voltage density can approximately be de-

[1] During the activation length, the length generator cannot be retriggered.

scribed by (2). The voltage frequency distribution for the case 's' is given in Fig. 3. Only the cases with windowing have a more dominant spike at zero, thereby reducing portions with low amplitudes near zero. This can be considered as a minor change.
More crucial are the length distribution and the length-energy relation.

Now, the influence of the phase setting will be outlined in some more detail. One effect is that the voltage histogram becomes slightly non-symmetric which can easily be avoided by alternating the polarity of the impulses. More important is that setting the phase to prescribed values means that all impulse events are also more or less forced to a certain pulse shape similar to the representative impulse shown in [2, 3]. This is due to the fact, that unlike real impulses, no variation in the multiplied PDS is provided (we use the mean PDS), because this would complicate the model significantly. In reality, also the group delay function has a slight variation around the typical function. This can be modeled by adding a Gaussian random variable to the samples of the group delay function. This results in some variations of the impulse shape and looks a little more realistic. However, without a phase setting the generated impulses do not look like real impulses. Nevertheless, as we shall see, they have advantages in some of the statistical properties. The effect of phase setting especially becomes clear when looking at the length-energy relation in Fig. 5 (case 'spw'). It reveals a limitation of the energy with increasing length. This corresponds to the fact that the impulse shape is concentrated to the length of the representative impulse, which means that a length change from a certain value (20 μs) on, does not have a significant effect on the impulse energy. Longer impulse events are also only encountered, if the inspection level of the impulse is rather low. This means that the length is determined by quite low amplitudes, not belonging to the main portion of the impulse. With such a low level, even the length distribution looks like the desired one (especially, when windowing is applied). However, the amplitudes that are output at positions around instant l after the beginning of the impulse are insignificantly small. Hence, a phase setting should only be applied, if the impulses should look like real impulses and the length distribution and the length-energy relation are not important.

Without phase setting a nearly linear length-energy relation is obtained (see, Fig. 5, case 'sw'). If additionally a length-dependent amplification is applied, even more complicated length-energy relations like the one in Fig. 4 can be realized. This may change the voltage density, though. The length-distribution, however, is still the most critical aspect. Due to the shape of the impulse response that corresponds to $\sqrt{\mathrm{PDS}}$, the length histogram has some discrete maxima at short lengths (see, e.g., Fig. 6, case 's'). The shape is slightly improved by additional windowing (Fig. 7, case 'sw'). For longer impulses the length histogram looks similar to the desired distribution. Although the length distribution cannot be regenerated exactly, one should bear in mind that anyway, in reality the length distribution showed the most variations dependent on the measurement location. Thus, the resulting length distribution is not totally unrealistic.

IV. CONCLUSIONS

A generator realization for non-stationary impulse-noise simulation has been proposed. It is based on distinctive pseudo-noise generators for the voltages, the lengths, and the inter-arrival times. Special measures have been taken to model spectral properties as well. Furthermore, an optional setting of the phase to a desired phase function was studied. We conclude that a phase setting is only suited, if the shape of impulses is of importance. Otherwise, in order to come to more realistic length-energy relations and length distributions, one should not include this operation. A good approximation was achieved for the voltage density, the mean power-density spectrum, and, of course, the inter-arrival time. Achieving the desired frequency distribution of the length remains critical, but nevertheless, the one obtained from a configuration with stochastic activation of the voltage generator together with windowing at the output seems to be acceptable.

References

[1] **Henkel, W., Kessler, T.:** Statistical Description and Modeling of Impulsive Noise on the German Telephone Network, *Electronics Letters*, Vol. 30, No. 12, June 9, 1994, pp. 935-936.

[2] **Henkel, W., Kessler, T.:** A Wideband Impulsive Noise Survey in the German Telephone Network: Statistical Description and Modeling, *AEÜ*, Vol. 48, No. 6, pp. 277-288, Nov./Dec. 1994.

[3] **Henkel, W., Kessler, T., Chung, H.Y.:** Coded 64-CAP ADSL in an impulse-noise environment — modeling of impulse noise and first simulation results —, *IEEE J. Sel. Ar. Comm.*, Dec. 1995.

[4] **Cook, J.W.:** Wideband Impulsive Noise Survey of the Access Network. *BT Technol. J.*, Vol. 11, No. 3, July 1993, pp. 155-162.

[5] **Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.:** Numerical Recipes, The Art of Scientific Computing, Cambridge University Press, Cambridge 1989, pp. 203-206.

[6] **Knuth, D.E.:** The Art of Computer Programming, Seminumerical Algorithms, Vol. 2, 2nd ed., Reading, Massachusetts, 1981.

[7] **Devroye, L.:** Non-Uniform Random Variate Generation. Springer 1986.

[8] **Walker, A.J.:** An Efficient Method for Generating Random Variables with General Distributions. *ACM Transactions on Mathematical Software*, Vol. 3, No. 3, Sept. 1977, pp. 253-256.

[9] **Johnson, G.E.:** Constructions of particular random processes. *Proceedings of the IEEE*, Vol. 82, 1994, pp. 270-285.

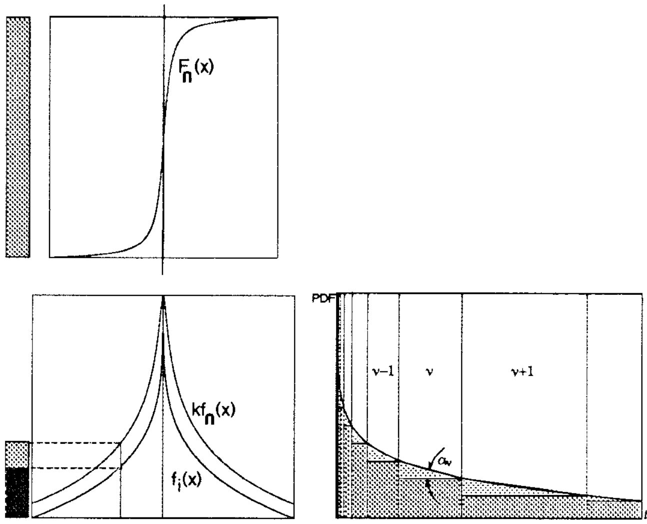

Figure 1: a) Rejection method b) Modified strip method

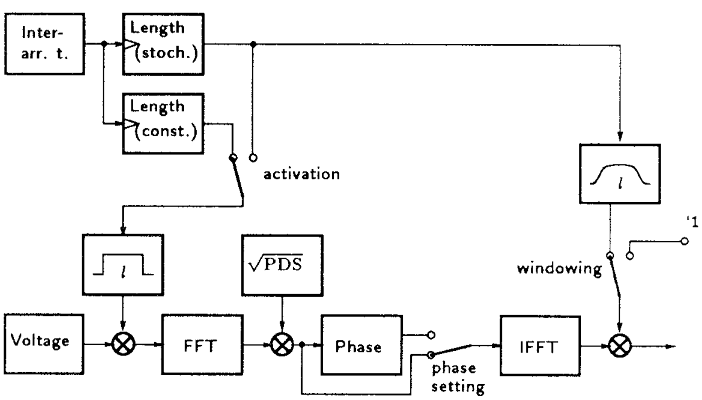

Figure 2: Block diagram of an impulse-noise generator (switches for different alternatives)

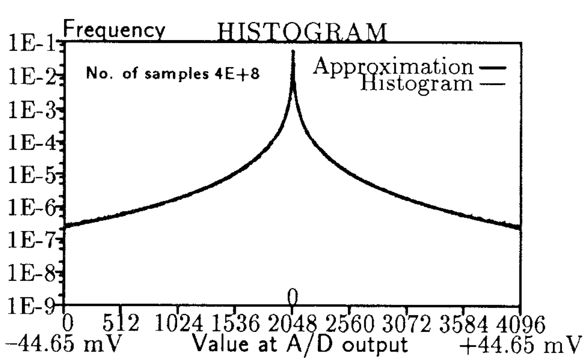

Figure 3: Voltage density, case 's'

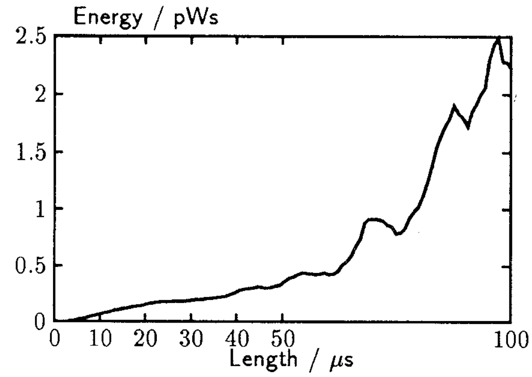

Figure 4: Length-energy relation of real measured impulses at one site

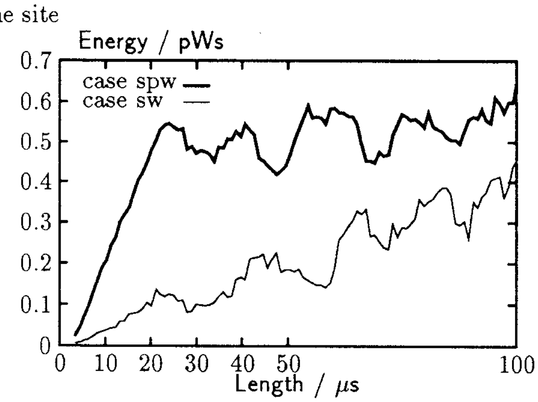

Figure 5: Length-energy relation, cases 'spw' and 'sw'

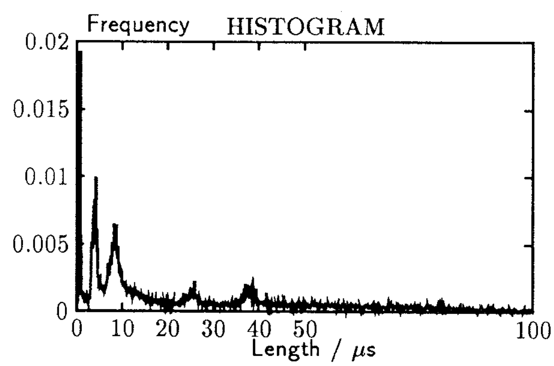

Figure 6: Impulse-length density, case 's'

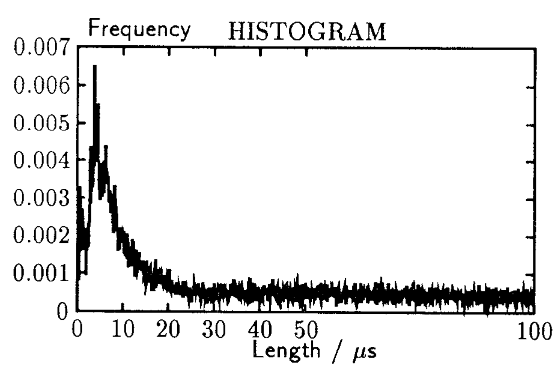

Figure 7: Impulse-length density, case 'sw'