Seismic signals hard clipping overcoming

Paula Olszowa[1],* and *Jakub* Sokolowski[2],**

[1]*Faculty of Pure and Applied Mathematics, Janiszewskiego 14a, Wroclaw University of Science and Technology, Poland*
[2]*Machinery Systems Division, Na Grobli 15, Wroclaw University of Science and Technology, Poland*

Abstract. In signal processing the clipping is understand as the phenomenon of limiting the signal beyond certain threshold. It is often related to overloading of a sensor. Two particular types of clipping are being recognized: soft and hard. Beyond the limiting value soft clipping reduces the signal real gain while the hard clipping stiffly sets the signal values at the limit. In both cases certain amount of signal information is lost. Obviously if one possess the model which describes the considered signal and the threshold value (which might be slightly more difficult to obtain in the soft clipping case), the attempt of restoring the signal can be made. Commonly it is assumed that the seismic signals take form of an impulse response of some specific system. This may lead to belief that the sine wave may be the most appropriate to fit in the clipping period. However, this should be tested. In this paper the possibility of overcoming the hard clipping in seismic signals originating from a geoseismic station belonging to an underground mine is considered. A set of raw signals will be hard-clipped manually and then couple different functions will be fitted and compared in terms of least squares. The results will be then analysed.

1 Introduction

In the data acquisition system the problem of good quality data collection is vital. The problem of data validation in the real world measurement system is renowned [1]. The quality of the data is poor mostly due to harsh environmental conditions (e.g. in monitoring of such machines as wind turbines [2] or mining machines [3–5]) or complexity of analysed process [6–8]. In the seismic signal processing one has to deal with plenty of difficulties. They include such problems as: small signal-to-noise ratio, missing data, signal distortions due to nonlinear effect of the data acquisition or transmission systems.

The signal clipping is one of the most common problems met in the signal processing field. The problem arises due to the limited range of values which the common sensors can register. The example of hard clipping can be seen in Figure 1

The negative impact of the clipping is obvious. The information of the signal is lost thus such values as the total energy which is travelling through the sensor is impossible to acquire [9].

One of the solutions implemented in order to overcome the clipping problem is the artificial division of signal values during the period of high values, thus producing the soft clipping. Such solution

*e-mail: 221708@student.pwr.edu.pl
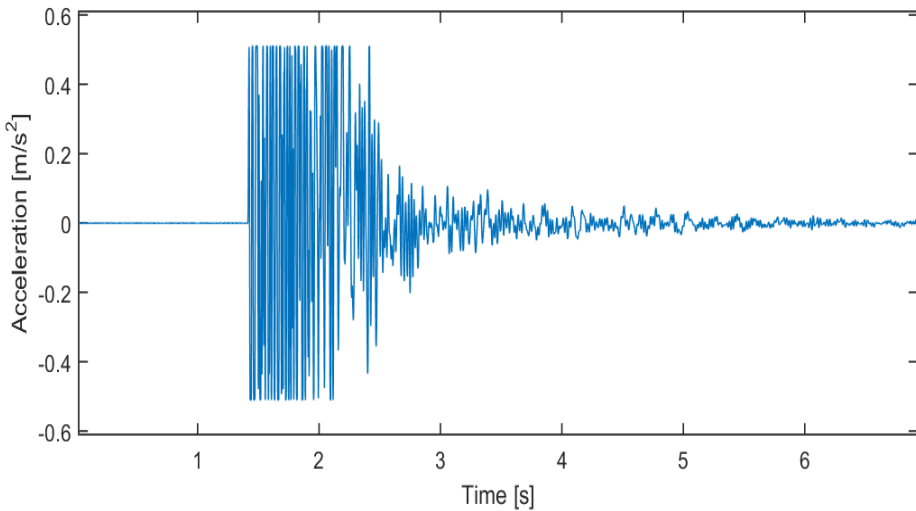**e-mail: jakub.sokolowski@pwr.edu.pl

Figure 1. The example of hard clipped signal

may be problematic as it requires to gather the information of the clipping periods. However, some sensors lack this kind of solution, thus some alternative is recommended.

One might suppose that the solution for real value evaluation during the hard clipping can be solved with use of the signal which is registered on parallel sensor (where the signal is not clipped). Unfortunately, due to different propagation paths these signals might be significantly diverse, which rejects such approach.

Due to the above mentioned fact, to solve the problem of choosing the proper function one has to work on the non-clipped signals, artificially clipping them afterwards and compare evaluated values to vanilla signals. It is commonly assumed that the seismic signals take the form of an impulse response of some specific system [10]. This may lead to belief that the sine wave may be the most appropriate to fit. However, it is recommended to be tested.

The main idea of this paper is to bring the signal artificially under hard clipping. Then basing on the signal samples which occur just before and after the clipping, the signal is estimated. With comparison of different functions fitted to artificially clipped periods, one can acquire information which function may estimate the clipping moments in the most accurate way.

Such results may provide the solution, which may be used for sensors which lack the advanced options.

2 Methodology

During the analysis of signals the artificial limit of clipping is estimated. It is done for non clipped signals and then the artificial limit is $\alpha = p \cdot max(x)$ for the upper clipping or $\alpha = p \cdot min(x)$ for the lower clipping, where $p \in (0, 1)$ and x is the signal. The main points needed to assign the estimation function are selected according to the following algorithm:

1. Find the first index of a sample, whose value is above the limit of clipping (the top clipping) or below the limit of clipping (the bottom clipping) and take this point as x_1 by changing its value to α.

2. Find the following point after x_1, which is below (the top clipping) or above (the bottom clipping) the straight line of clipping, then take the previous point and call it x_2, change its value to α.

3. Do points 1 and 2 for the whole signal.

The interval for which the clipping is estimated contains in $[x_1 - k, x_2 + k]$, where k is the number of points before or after the clipping.

To estimate the real signal, the basic functions are needed. They are described with the formulas:

- Two linear functions: $f(x) = a_1 \cdot x + b_1 \mathbb{1}_{[p1,c]} + a_2 \cdot x + b_2 \mathbb{1}_{(c,p2]}$,

- Quadratic function: $f(x) = (a \cdot x^2 + b \cdot x + c)\mathbb{1}_{[p1,p2]}$,

- Sine function: $f(x) = a \cdot sin(b \cdot x + c)\mathbb{1}_{[p1,p2]}$,

- Power function: $f(x) = (a \cdot |x - \bar{x}|^b + c)\mathbb{1}_{[p1,p2]}$,

- Exponential function: $f(x) = (e^{a \cdot x^2} + b)\mathbb{1}_{[p1,p2]}$,

where p_1, p_2 stands for left and right limit of single clipping. The main reason behind selection of these functions was their simplicity (before testing more complicated functions it is justified to check the simpler).

To make the function estimation more precise, it is necessary to take the proper number of points before and after clipping into account during the approximation. Depending on the taken p, the number k may be different. Additionally the considered observations are divided by length of clipping interval (lengths from 2 to 5 and lengths greater 5).

Sum of squared errors is defined as:

$$SSE = \sum_{i=1}^{n}(x_i - \widehat{x_i})^2, \tag{1}$$

Where n is length of the clipping period, x_i is the signal value before clipping, $\widehat{x_i}$ is estimated value. SSE is used to estimate the error with which the solution fits to the real signal.

3 Applications

In order to efficiently evaluate the individual function potential, one has to test it on the real data. In this section the details of analysed data and results of tested functions are presented.

The data-set on which the functions are tested consists of 38 real seismic signals acquired in an underground mine which were basically not clipped. In Figures 2-6 the results of proposed functions are contained. In order to compare their effectiveness the tested function is the same in each figure. The exemplary function clipping level is 0.25.

In Figure 2 one can see the example fitting of the linear function. It can be seen that the function has the tendency to fit values more extreme (for the absolute value of signal, the estimated values greater than the actual ones).

In Figure 3 one can see the second order polynomial fitted to the artificially clipped signal. The estimated values have the tendency to be closer to the mean value of the signal than the actual values.
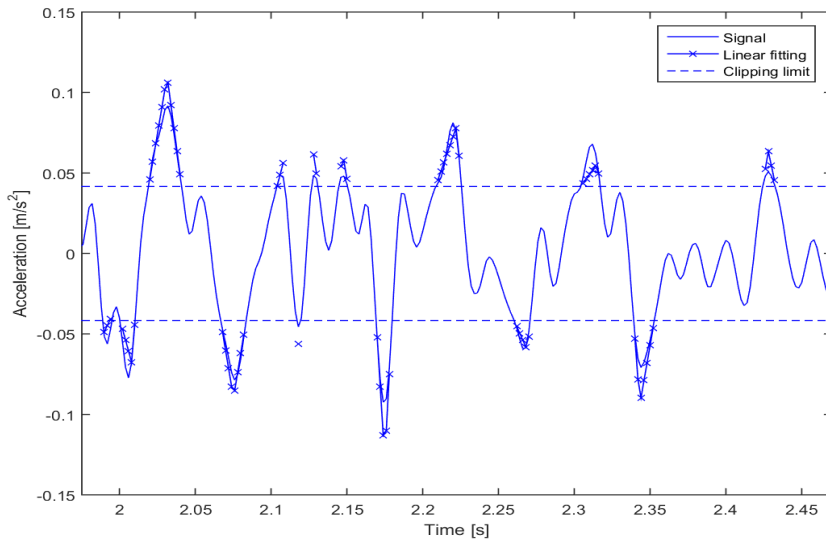
Figure 2. The example of linear function fitting used on clipped signal fragment
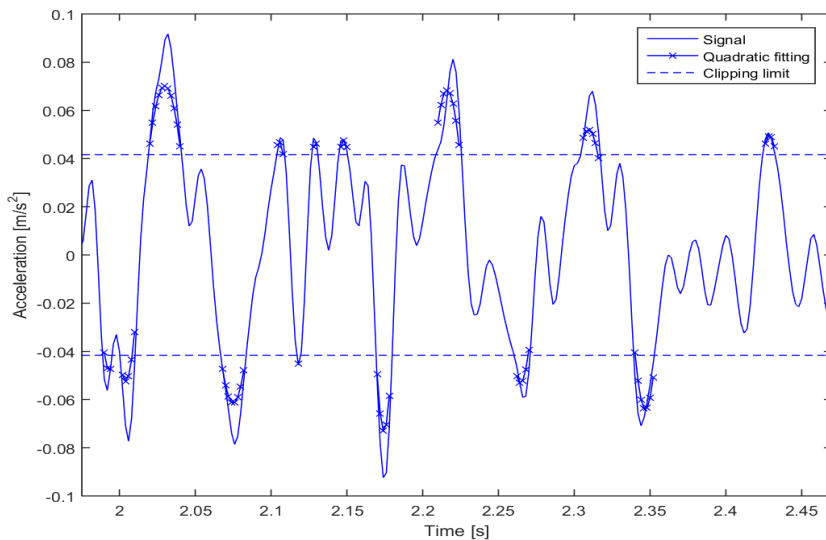

Figure 3. The example of second order polynomial function fitting used on clipped signal fragment

The fitting of the sine function can be seen in Figure 4. Clearly the values of the estimated function for the example signal are closer to actual values than for that in Figure 3.
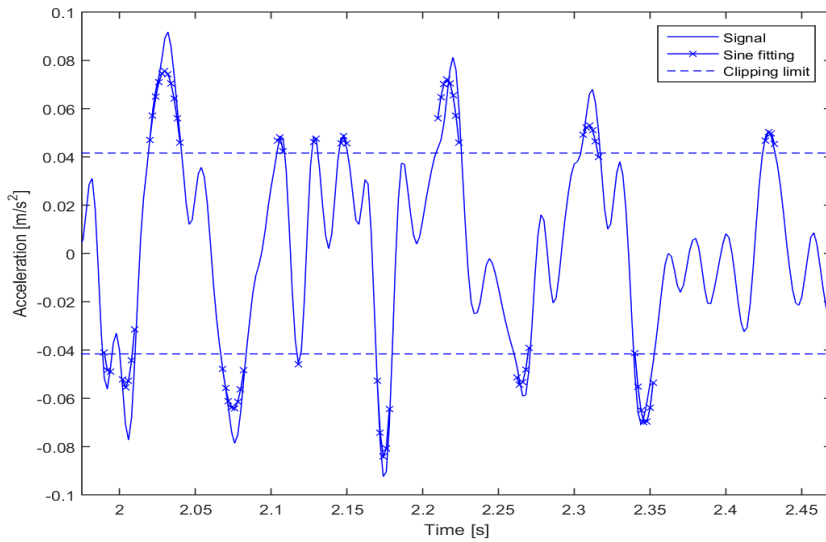
Figure 4. The example of sine function fitting used on clipped signal fragment
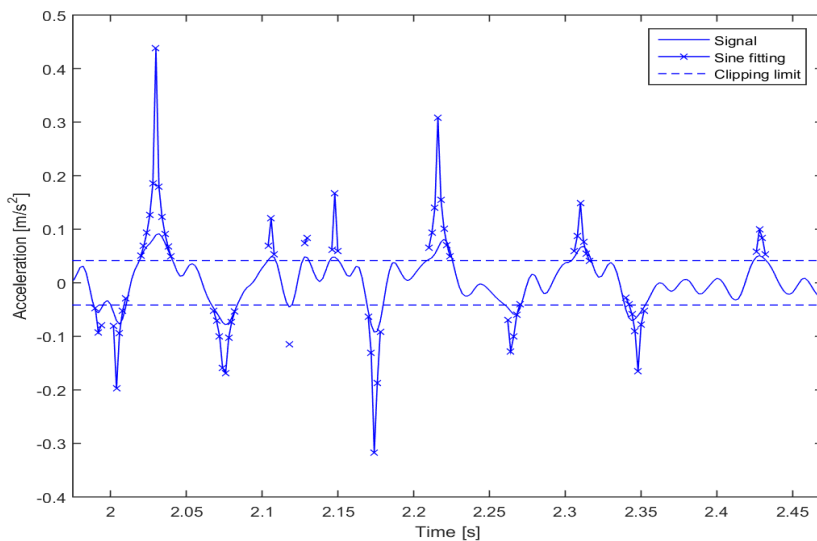

Figure 5. The example of power function fitting used on clipped signal fragment

In Figure 5 the power function fitted to the clipping is included. It can be seen, that the proposed function has the tendency to estimate values more extreme than the actual values.
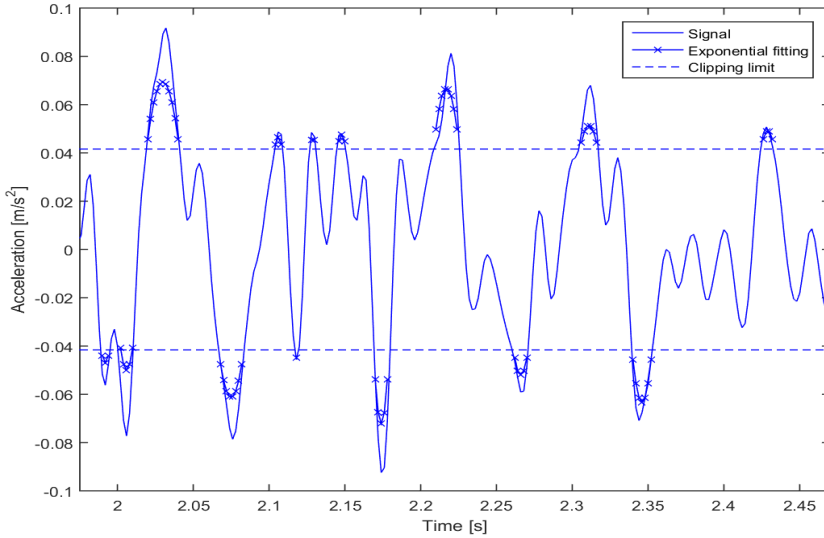
Figure 6. The example of exponential function fitting used on clipped signal fragment

The clipping with fitted exponential function can be seen in Figure 6. The results of this function seem to be similiar to those of the quadratic function (fitted function too smooth).

In Figure 7 the collective results for all different functions can be seen for one upper clipping (a), and one lower (b). Both examples show that the power and linear functions have the values steeper than the actual and the rest of the functions have the results considerably smoother (with considerably similar values). Moreover, the lack of the real signal symmetry can be seen (however, this can be due to the sampling frequency).

The summary of the functions fitting potential is presented in Table 1 for short clippings and in Table 2 for long ones.

Table 1. Clipping ≤ 5 samples

Clipping level	Linear	Quadratic	Sine	Power	Exponential
0.25	0.0157	0.0090	**0.0079**	0.0222	0.0101
0.40	**0.0156**	0.0221	0.0197	0.1980	0.0248
0.55	**0.0241**	0.0393	0.0364	0.0276	0.0421

In Table 1 the sum of squared errors between fitted functions and real signal for short clippings are included. It can be seen that for the clipping level 0.25 the sine function is the best fitted and the quadratic function is the second best. For clipping levels 0.40 and 0.55 the linear function is the best fitted function and the sine function is the second best for the 0.40 level, and power function for 0.55. The power function has the worst results for the 0.25 and 0.40 levels and the exponential function for 0.55. The good results of power function for the exemplary signal seem not to double for the rest of the signals.
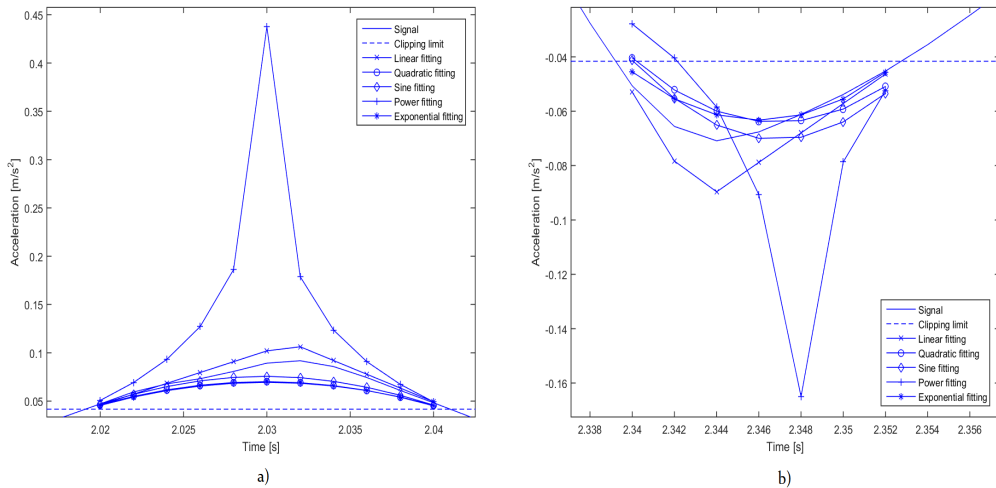
Figure 7. Comparison of all tested functions on a) upper and b) lower clipping

Table 2. Clipping > 5 samples

Clipping level	Linear	Quadratic	Sine	Power	Exponential
0.25	0.0091	0.0014	0.0015	**0.0011**	0.0013
0.40	0.0096	0.0051	**0.0042**	0.0054	0.0075
0.55	0.0490	0.0631	0.0458	0.0938	**0.0252**

In Table 2 the results for the longer clippings are included. In this case the power, sine and exponential functions have the best fitting for the respectively 0.25, 0.40, 0.55 clipping levels and the exponential, quadratic and sine functions are the second best. In case of longer clippings the power function is considerably more accurate for clipping level 0.25 and 0.40.

Unfortunately none of the considered functions proved to be undoubtedly better than the other. This may be due to the considerably small testing dataset.

4 Conclusions

In this paper a problem of seismic signals clipping overcoming has been shown. In order to find the best function which could be fitted to hard-clipped data, five different functions were considered. Four of them were symmetric, and one (linear) not.

The function were tested on a real non-clipped seismic signals originating from an underground mine. To exhaust the subject, the signals were clipped on different level and each of these levels were considered separately. Also the clippings were divided into short (≤ 5 samples) and long.

The results proved to be ambiguous, as for different clipping level and length of clipping, different functions proved to be the best fitted. It is suspected that this may due to relatively small testing dataset. In future work it is suggested to enlarge testing set in order to make the results clear. Also it is recommended to check fitting potential of other non-symmetric function.

References

[1] A. Ray, IEEE Control Systems **11/2**, 44-49 (1991)

[2] A. Jablosnki, T. Barszcz, M. Bielecka, Measurement **44/10**, 1954-1967 (2011)

[3] A. Jablosnki, T. Barszcz, Mechanical Systems and Signal Processing **38/1**, 248-263 (2013)

[4] P. Kepski, T. Barszcz, Diagnostyka **4/64**, 25-30 (2012)

[5] P. Kepski, T. Barszcz, in: T. Fakhfakh, W. Bartelmus, F. Chaari, R. Zimroz, M. Haddar, *Condition Monitoring of Machinery in Non-Stationary Operations, Part 4* (Springer, Berlin, 2012) 461-469

[6] R. Zimroz, J. Wodecki, R. Krol, M. Andrzejewski, P. Sliwinski, P. Stefaniak, in. C. Drebenstedt, R. Singhal, *Mine Planning and Equipment Selection* (Springer, Cham, 2014) 1285-1294

[7] P. Stefaniak, R. Zimroz, J. Obuchowski, P. Sliwinski, M. Andrzejewski, Procedia Earth and Planetary Science **15**, 797-805 (2015)

[8] P. Stefaniak, R. Zimroz, W. Bartelmus, M. Hardygora, Advances in Condition Monitoring of Machinery in Non-Stationary Operations **4**, 250-263 (2016)

[9] S.Q. Li, IEEE Transactions on Communications, **37/11**, 1192-1202 (1989)

[10] R. Hossa, R. Makowski, R. Zimroz, International Journal of Rock Mechanics and Mining Sciences, **91**, 29-39 (2017)