

# Small Current Fault Line Selection Based on Improved Empirical Mode Decomposition and Fractal Dimension Method

Chao Liu, Limei Yan\*, Yina Zhou

Dept. of Electrical Engineering and Automation, University NORTHEAST PETROLEUM of Daqing, China

\*Corresponding Author.

## Abstract

*In this paper, the fault steady state and transient characteristics of small current grounding system are analyzed, and the distribution of transient zero sequence current is introduced. A fault line selection based on EMD and fractal dimension method is proposed. After the parameter is determined, the problem is proposed and improved. Using the simulated annealing K-means algorithm to find the scale-free interval curve to get the line slope is the correlation dimension of the line. Finally, by comparing the size of the associated dimension, you can select the corresponding line of the faulty distribution network.*

**Keywords:** Small current grounding system; empirical mode decomposition; Improved EMD algorithm; G-P algorithm.

## I. Introduction

In 2003, Rilling proposed the method of mirroring continuation to extend the signal, which has a good effect on eliminating the phenomenon of endpoint effect[1]. At present, the effective methods to restrain the boundary flying wing of EMD include neural network prediction extension method, mirror extension method, symmetry extension method, extreme value extension method and so on. The mirror extension method can correctly separate the frequencies contained in the signal to some extent, but it still has some limitations[2-4]. Non-stationary nonlinear signal is an uncertain signal with no rules to follow, so some algorithms may be used for estimation, which will lead to errors, and the algorithm itself is not perfect[5,6]. The slope method can not solve to endpoint effect very well. In fractal theory, the k-means clustering algorithm that determines the scale-free interval may lead to the local optimal solution[7]. So simulated annealing algorithm and genetic algorithm can be used to make sure the scale-free intervals are more accurate when identification process.

The single-phase grounding fault line selection of small current grounding system in distribution network is studied in this paper. Since the zero-sequence current of each line in the distribution network is nonlinear and non-stationary when single-phase ground fault occurs in the small current grounding system, a method of fault line selection based on improved empirical mode decomposition (EMD) and improved G-P algorithm based on fractal theory is proposed.

## II. EMD Theory

Empirical modal decomposition is the decomposition of a complex signal into the sum of the terms of the natural modal functions composed of different frequencies. The components of the inherent mode function can be transient signals or steady signals. IMF components are obtained by adding up the frequencies from large to small, and the last term to decompose EMD is the average or monotonous trend term[8,9]. Although EMD has many advantages in signal analysis, this method relies on experience, the problem of endpoint effect in reality will bring errors to decomposition. Therefore, the research on the improvement of EMD algorithm is mainly to further improve the endpoint effect.

When the EMD method is used in the process of signal decomposition, the upper and lower envelope obtained by

fitting the maximum and minimum values of the initial original signal of the system is very critical. In this process, it is easy to appear the endpoint divergence, and this phenomenon will affect the whole decomposition process of empirical mode, then slowly affect the whole sequence of signal data, which is likely to distort the components obtained by EMD decomposition without any physical significance.

Aiming at the problems existing in the slope continuation method, the upper and lower envelope of the simulated signal through the improved slope continuation method is still not perfect. Therefore, in the slope of the endpoint border continuation method of operation, the mirror continuation method used in.

The steps of extension are as follows:

Firstly, the improved mirror method is used to extend a corresponding maximum or minimum value on both sides of the time series.

Secondly, by using the improved slope continuation method, the two extremum points which have been extended in the previous step are respectively extended to a corresponding minimum or maximum on both sides of the time series.

In Fig.1 it shows original signal with sampling point of 120. Fig. 2 is the result of the upper and lower envelope lines extended by the combination of the improved slope method and the mirror image method. In the graph  $MMin(0)$  and  $MMin(N + 1)$  are the extreme points on both left and right sides which are extended by the improved image method.  $SMax(0)$  and  $SMax(N + 1)$  are the extreme points of both left and right sides obtained by the improved slope continuation method.

It can be clearly seen from the Fig.2 that the upper and lower envelope lines obtained by the combination of the improved image method and the improved slope method do not have the problem of incomplete envelope caused by only using the slope method, which initially reflects the relative advantages of this method.

The simulation signals selected in this paper are as follows:

$$s(t) = \sin(200\pi t) + \sin(100\pi t) + \sin(40\pi t) + e^t \quad (1)$$

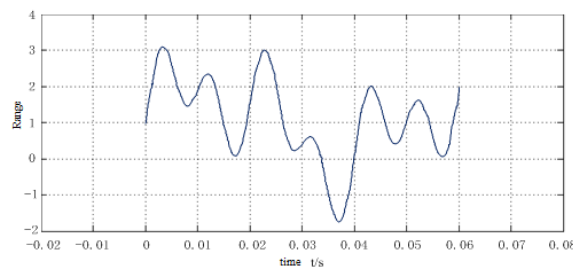


Fig 1: Original signal with sampling point of 120

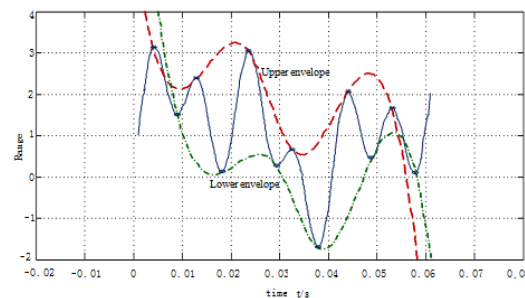


Fig 2: The upper and lower envelopes obtained without considering the end points

According to the endpoint effect method proposed in recent years, the accuracy of instantaneous frequency spectrum

is usually used as the qualitative index. Fig.3, Fig.4 and Fig.5 show the improved instantaneous frequency diagram.

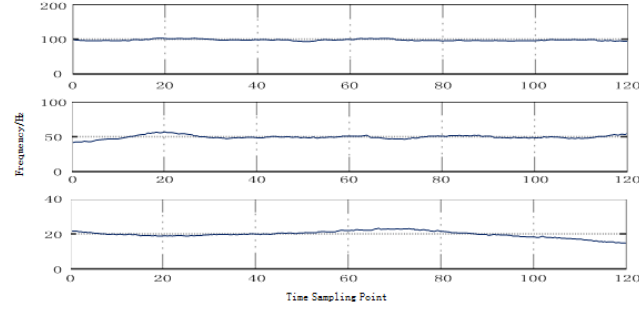


Fig 3: Instantaneous frequency map based on improved image extension method

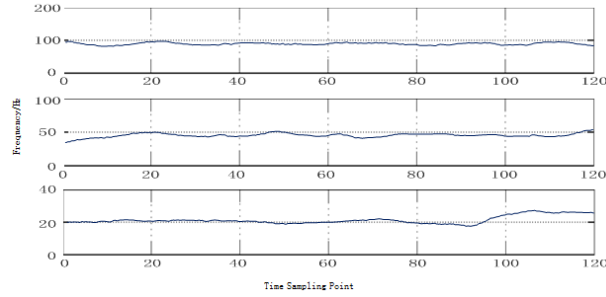


Fig 4: Instantaneous frequency diagram based on improved slope method

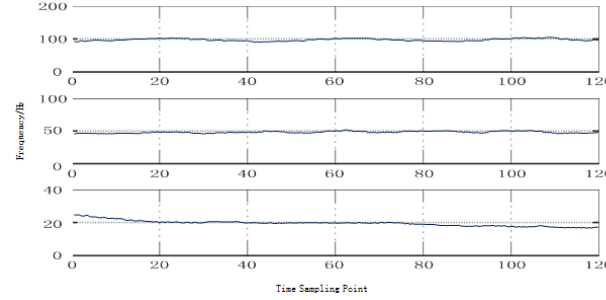


Fig.5: Instantaneous frequency based on combining improved image extension and slope method

Through comparison and analysis of figures 3-5, according to the stability of instantaneous frequency and the divergence of both ends of the data, the following conclusions can be drawn. The order of accuracy of the three methods from high to bottom is: combining improved image extension method and slope method, improved image extension method ,improved slope method. This shows that the method can better suppress the end effect of EMD.

### III. Correlation Dimension and G-P Algorithm

Correlation dimension is to judge the degree of association between each point in a set according to the property of correlation between any points in the set[10]. If the set A consists of M points, the time series of the points in the space is  $x(n), n = 1, 2, L, M$ . In space, the coordinates of two time series whose distance between two points is less than  $\delta$  are related. We can judge whether all the points in the space are related or not. The size of the associated coordinates in the whole space is called the correlation dimension:

$$C(\delta) = \frac{1}{M^2} \sum_{i,j=1}^M \theta(\delta - |x_i - x_j|) \quad (2)$$

$\theta(r)$  is Heabiside function.

$$\theta(r) = \begin{cases} 1, & r > 0 \\ 0, & r \leq 0 \end{cases} \quad (3)$$

The correlation dimension of chaotic time series is defined as follows:

$$D_2 = \lim_{\delta \rightarrow 0} \frac{\ln C(\delta)}{\ln \delta} \quad (4)$$

### 3.1 Shortcomings of G-P algorithm

In the analysis and Simulation of the G-P algorithm in Fig.6, we can see that the G-P algorithm also has the following shortcomings.

The selection of hypersphere radius  $r$  is based on people's subjective choice, and any real number can be set for the semi meridian of hypersphere. In this way, the selected hypersphere radius is too subjective.

The traditional scale-free interval judgment is too subjective, and whether the slope reaches saturation is not determined by objective rules. At present, the determination of scale-free interval can not rely on objective observation. Subjective judgment of its upper and lower envelope is located to determine the scale-free interval, and further artificial imagination is used to judge whether the slope of the segment in this interval reaches saturation.

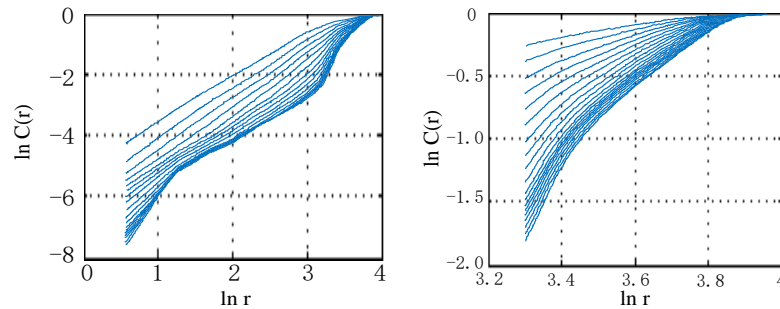


Fig.6: The relationship curves of different hypersphere radius intervals with the same data

### 3.2 Improvement of G-P algorithm

The original G-P algorithm is too subjective to determine the range of  $r$  value of hypersphere. In this paper, the minimum value of the range of  $r$  value is defined as the minimum value of the distance between any two points in chaotic space, similarly, the maximum value of the range of the scale  $r$  value is defined as the maximum value of the distance between any two phase points in the chaotic space. Expressed as:

$r : [\min(\|X_i - X_j\|_\infty), \max(\|X_i - X_j\|_\infty)]$ . The advantage of this method is that the range of  $r$  value can be adjusted according to the change of chaotic space.

The advantage of this method is that the range of  $r$  value can be adjusted adaptively according to the change of chaotic space. At the same time, it can also take into account all the phase points in the whole phase space without missing the time series, so it has better comprehensiveness and stability.

In order to solve the problem of subjective judgment in scale-free interval, this paper makes use of the characteristic that the second derivative of correlation integral curve fluctuates continuously around the value of 0 in scale-free interval. The corresponding points in the data with fluctuation of 0 value can judge the scale-free interval of time series through the simulated annealing genetic k-means clustering algorithm. In the classification and recognition, the scale-free interval can be well and comprehensively judged.

### 3.3 Simulated annealing genetic K-means clustering algorithm

Due to the strong ability of local search, annealing also has the powerful ability of comprehensive search[11,12]. Therefore, the combination of genetic algorithm and simulated annealing algorithm in K-means clustering algorithm will make the effect of clustering algorithm more effective and more quickly converge the time series to the optimal solution of the whole interval

The detailed steps of simulated annealing genetic K-means clustering algorithm are as follows:

Step1:Initialization parameters: Size of individual population, the maximum number of evolution MAXGEN, probability of crossover, probability of variation, initial annealing temperature T, temperature coefficient temperature of cooling, termination temperature.

Step2:Random initialization of C cluster centers, the initial population is established in the population, and the fitness value f of each cluster center position of time series for each data is calculated;

Step3:Set cycle count variable gen=0;

Step4:The population, mutation and crossover genetic selection for population manipulation is implemented to generate the central formula of each cluster using recalculation, fitness of each individual  $f'$ , if  $f' > f$ , the new individual is replaced by the new individual; if not, the new individual is accepted with probability and the only individual is discarded;

Step5:If  $gen < MAXGXN$ , then  $gen = gen + 1$ , go to Step 4. If not, go to Step6.

Step6:If  $T_i < T_{end}$ , then the algorithm has been completed successfully, and the global optimal solution is returned. Otherwise, the cooling operation will be carried out to step 3.

According to the experiment of the improved G-P algorithm described above in the fractal system simulation in MATLAB, the algorithm flow chart is shown in Figure 7, and the simulation results prove the effectiveness of the improved algorithm.

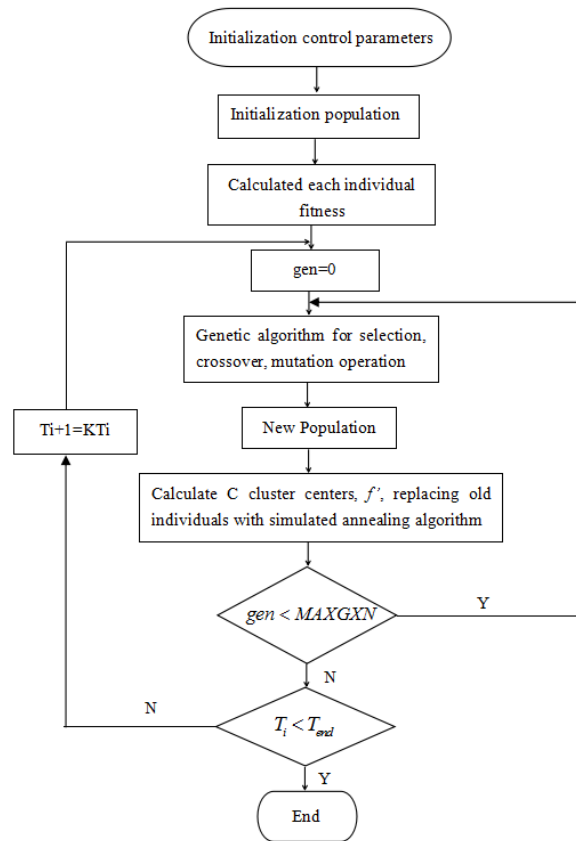


Fig.: Flow chart of simulated annealing genetic K-means clustering algorithm

Since Lorenz and Rossler systems are both typical fractal systems, take these two systems as examples in the following Table 1.

Table 1 Simulation of improved G-P algorithm

Model	Model Equation	Model Parameter	Sample	Embedding Dimension	Theoretical value of correlation dimension	Subjective identification method	K-means	genetic K-means	annealing K-means
Lorenz	$\frac{dx}{dt} = a(y-x)$ $\frac{dy}{dt} = cx - xz - y$ $\frac{dz}{dt} = xy - bz$	Initial value: [x,y,z]=[-1,0,1] [a,b,c]=[16,4,45.92] Integral step length:0.01 Integral interval:[0 1000]	X component [14001-15000]	7	$2.07 \pm 0.01$	$2.125 \pm 0.061$	2.401	2.084	2.072
Rossler	$\frac{dx}{dt} = y - z$ $\frac{dy}{dt} = by - x$ $\frac{dz}{dt} = c + z(x-a)$	Initial value:[x,y,z]=[1,1,1] [a,b,c]=[0.2,0.2,10] Integral step length:0.01 Integral interval:[0 1000]	X component [15001-16000]	16	$2.07 \pm 0.02$	$2.0094 \pm 0.056$	2.167	2.005	2.029

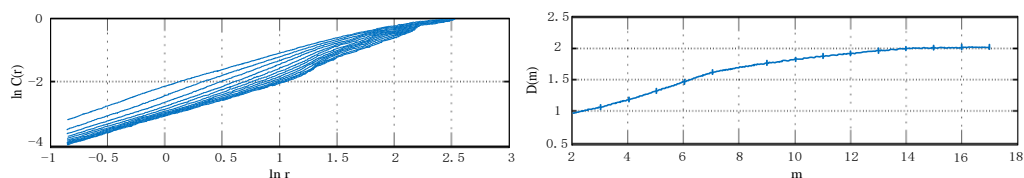


Fig.8: Correlation dimension of improved G-P algorithm in Rossler system

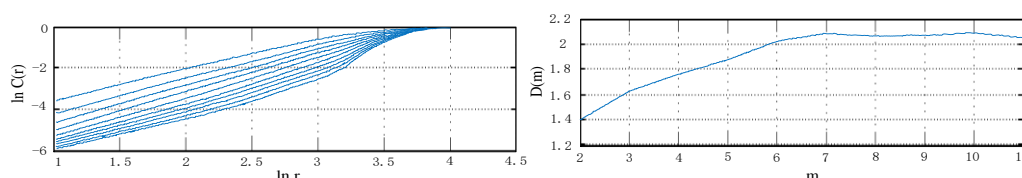


Fig.9: Correlation dimension of improved G-P algorithm in Lorenz system

It can be seen that the improved G-P algorithm is more accurate than the original G-P algorithm in Fig.8 and Fig.9.

#### IV. The Method of Line Selection based on Correlation Dimension

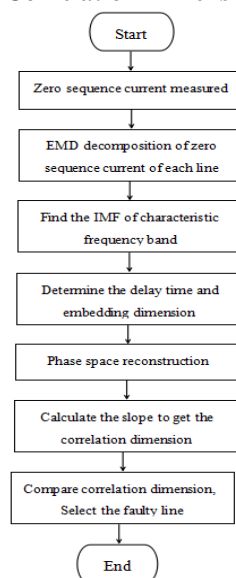


Fig 10: Flow chart of single phase grounding fault line selection

The process of line selection is shown in Figure 10. Firstly, the simulation software is used to set the single-phase ground fault, and the zero sequence current is detected. Then, the Improved EMD algorithm is used to decompose the zero sequence current of the four lines, and the IMF component of the characteristic frequency band can be obtained after the signal is decomposed; By reconstructing the phase space of the fractal set, the correlation integral curve of each line can be obtained by using the improved G-P algorithm. After finding the scale-free area of the correlation integral curve, the slope of the correlation integral curve is obtained by fitting the curve to obtain the correlation dimension. By comparing the correlation dimension, the fault line is selected.

In order to prove the effectiveness of the above method, the simulation line is established in the power system simulation software EMTP-ATP in Fig.11. The 110kV AC power supply is used. The transformer is connected by triangle and star. When the transformer is disconnected at the neutral point, the neutral point is grounded through the arc suppression coil. When the switch connected to the neutral point of the transformer is disconnected, the neutral point is non-ground system.

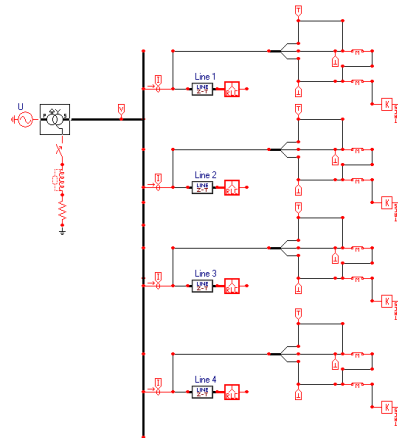
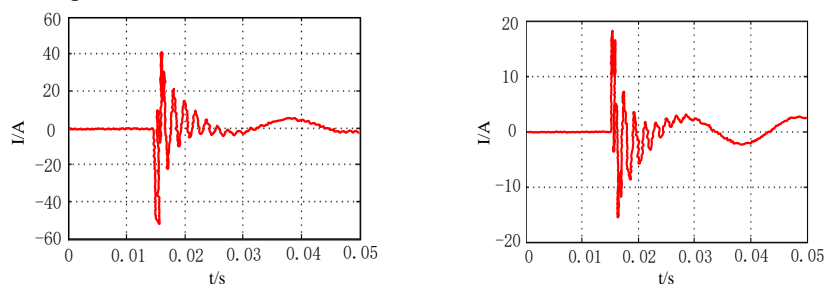
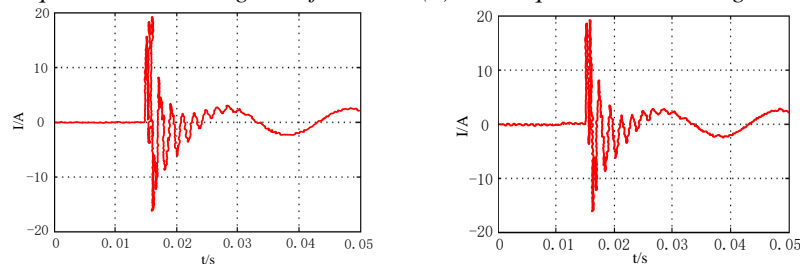


Fig.11: Simulation model of small current grounding

Fig.12 shows the zero sequence current of each line.



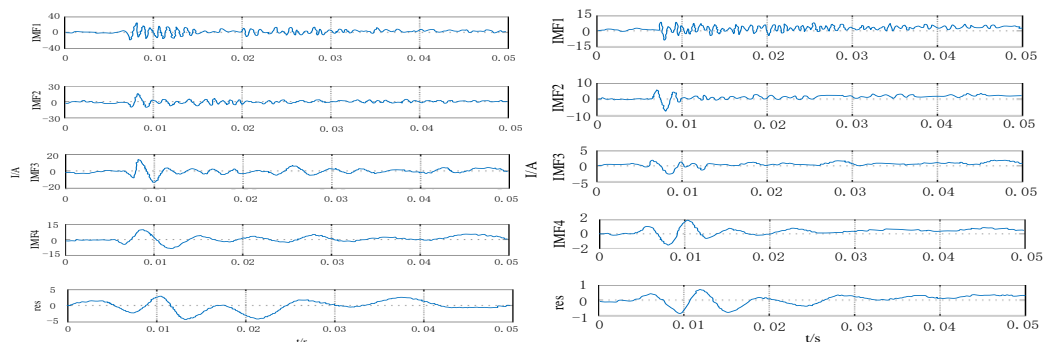
(1) Zero sequence current diagram of Line 1 (2) Zero sequence current diagram of Line 2



Zero sequence current diagram of Line 3 (4) Zero sequence current diagram of Line 4

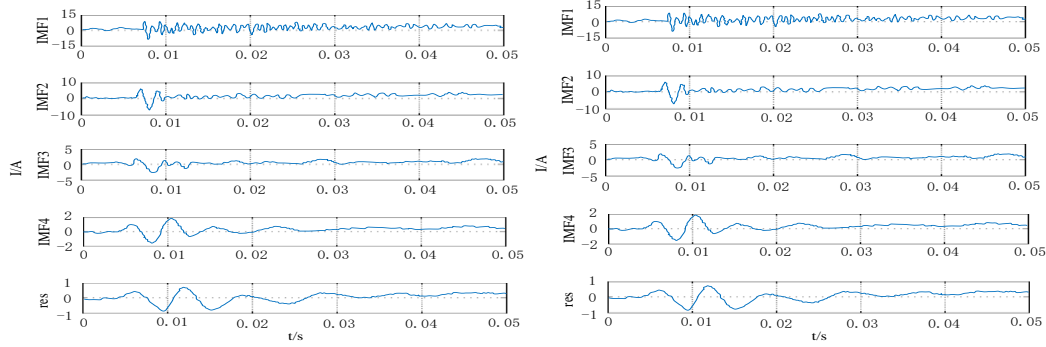
Fig.12: Zero sequence current in case of fault

For the zero sequence current signal of the above line, each zero sequence current signal is decomposed by the Improved EMD algorithm. The frequency of the current signal is very complex. Therefore, the four intrinsic mode functions and the residual are obtained, and the harmonic components are found through simulation. Each line is decomposed by the Improved EMD algorithm as shown in Fig.13.



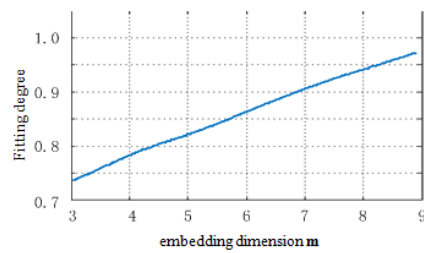
(1) Improved EMD decomposition of Line 1 (2) Improved EMD decomposition of Line 2



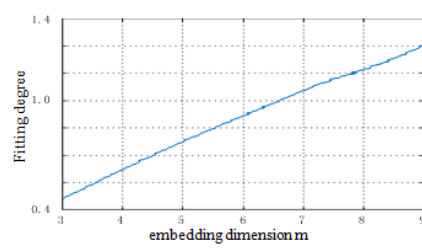


(3) Improved EMD decomposition of Line 3 (4) Improved EMD decomposition of Line 4  
Fig.13: Decomposition diagram of zero sequence current of transmission line by EMD algorithm

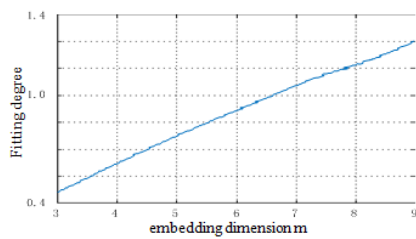
Through adaptive selection of neighborhood radius interval and simulated annealing genetic K-means clustering algorithm, the embedding dimension and scale-free interval are calculated. The correlation integral curve of each line is obtained by the improved G-P algorithm, and the approximate linear part is found in the interval of the correlation integral curve. The correlation integral fitting curve of each line is obtained by correlation integral fitting of this part, as shown in the figure below. The correlation integral curve of each line is obtained by the improved G-P algorithm, and the approximate linear part is found in the interval of the correlation integral curve. The correlation integral fitting curve of each line is obtained by correlation integral fitting of this part, as shown in the Fig.14.



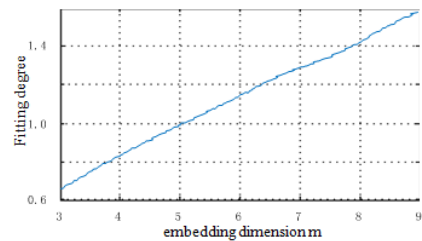
(1) Correlation fitting diagram of Line 1



(2) Correlation fitting diagram of Line 2



(3) Correlation fitting diagram of Line 3



(4) Correlation fitting diagram of Line 4

Fig.14: Correlation fitting curve of each line

The slope of the correlation fitting curve obtained from the above figures are the required correlation dimensions. Then the correlation dimensions can be calculated. The results are in the Table 2.

Table 2 Line correlation dimension

Line	Line1	Line2	Line3	Line4
Correlation Dimension	0.0413	0.1292	0.1426	0.1587

When calculating the correlation dimension of each line, we can find that the correlation dimension in the fault line is the smallest. Then the correctness of the method is proved.

## V. Conclusions

In this paper, by analyzing the difference between the Improved EMD algorithm and the G-P algorithm in fractal theory in dealing with nonlinear problems, a method of line selection for small current grounding fault in distribution network is proposed by combining the two methods. By analyzing the fault transient process of small current system, the Improved EMD algorithm decomposes the transient zero sequence signal to get the characteristic frequency band of the system, and then obtains the corresponding correlation dimension through the improved G-P algorithm. The fault line selection is determined by comparing the correlation dimension of each line.

## Acknowledgements

This work is supported by the Natural Science Foundation of Heilongjiang Province (Grants No. LH2019E016).

## References

- [1] G. Rilling, P. Flandrin and P. Goncalves, "On Empirical Mode Decomposition and Its Algorithms", in IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, 2003.
- [2] H. Liu, M. Zhang and J. Cheng, "Dealing with the End Issue of EMD based on Polynomial Fitting Algorithm", in Computer Engineering and Applications, no.16,2004,pp.84-86+100.
- [3] L. Xue, S. Yang, J. Ma and Y. Liu, "Improvement of Scale-invariant Feature Transform Algorithm Based on Polynomial Fitting", in Remote Sensing Information, vol.34, no.4, Aug.2019, pp. 54-61.
- [4] Y. G. Lei, Z. J. He and Y. H. Zi, "EEMD method and WNN for fault diagnosis of locomotive roller bearings", in Expert Systems with Applications, vol.38, no.6, 2011, pp.7334-7341.
- [5] Z. C. Ma and Q.S. Xiao, "The Application of Mirror Extension for the end Effects in EMD Method", in Science and Technology Vision, no.24, Aug. 2019, pp. 31-32+118.
- [6] P. P. Li, W. W. Feng, Y. L. Zhang and Q. Wang, "Study of Improved Mirror Extending EMD Method and Application", in Journal of Zhejiang Ocean University, vol.35, no.3,May. 2016,pp. 244-248.
- [7] Z. B. Li, H. L. Liu and Z. J. Ma, "Slope Matching Waveform Extension Method Suppressing the EMD end Effect", in Journal of Northwest University, vol.46, no.5,Oct. 2016,pp.673-679.
- [8] L.H. Tang, Y. L. Yang and J. Lu, "A hybrid prediction method based on empirical mode decomposition and multiple model fusion for chaotic time series", in Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena, vol.141, 2020.
- [9] H. Wang, Z.J. Hu, and Y.H.Zhang, "A Hybrid Model for Short-Term Wind Speed Forecasting Based on Ensemble Empirical Mode Decomposition and Least Squares Support Vector Machines", in Transactions of China Electrotechnical Society, vol.29, no.4, Apr. 2014, pp.237-245.
- [10] J. W. Dang and J. G. Huang, "Study of the Parameters Used in Calculating Correlative Dimension Based on G.P. Algorithm", in Pattern Recognition, no.1, 2004,pp.48-51.
- [11] J. Lee and D. Perkins, "A simulated annealing algorithm with a dual perturbation method for clustering", in Application Research of Computers, 2020.
- [12] Y. L. Si and L J. Zheng, "Transient Electromagnetic Nonlinear Inversion based on Genetic Annealing Evolutionary Algorithm", in Computer Applications and Software, vol.37, no.11, Nov, 2020,pp.172-177.