



ILJS-24-078 (SPECIAL EDITION)

## Comparison of Computational Efficiency of Conjugate Gradient Methods

**Bamigbola O. M.**

Department of Mathematics, University of Ilorin, Ilorin, Nigeria.

### Abstract

Based on the defects in the two existing metrics, namely, the performance profiling, and a device which adapted the descent condition to the Cauchy-Schwartz inequality, another metric, the coefficient in the sufficient descent condition, is proposed for utilization. The proposed metric can be obtained by the use of practical computations with the CG algorithms.

Keywords: Conjugate gradient method, Efficiency measure, Optimization problems, Performance profiling, Empirical data.

### 1. Introduction

Advances in science and technology has given rise to integrated or large-scale systems for the enhancement of human life. An important class of these systems evolve as large-scale nonlinear optimization problems which do arise in many sectors such as the telecommunication, energy and manufacturing segments, just to name a few. As a result of the dominant position occupied by these sectors in the modern society, there is an intensive search for accurate, reliable and fast methods of solving the problems arising therefrom.

The Conjugate Gradient Method (CGM) is preferred to other methods in solving large-scale nonlinear optimization problems as a result of the simplicity of its iterations and its very low memory requirements, Wei et al. (2006), Shi and Guo (2008) and Dong et al. (2015a). Even with the CGM, the quest for improved methods is unabated as a result of rapid evolution of more sophisticated systems. As a result, the CGM has been a subject of intensive study in an attempt to develop methods that are very efficient. Thus, there is a need for appropriate tools to assess efficiency of CGMs, especially the new ones.

There are in use two efficiency measures for CGMs, namely, the performance profiling software due to Dolan and Moré (2002), and a metric devised by Hager and Zhang (2005) adapting the descent condition to the Cauchy-Schwartz inequality.

Assessing computational performance of CGMs by profiling and the Cauchy-Schwartz inequality-based techniques are fraught with some anomalies. The former suffers from being dependent on some features of the test problems, see Bamigbola (2020). In the same vein, the latter lacks the desired consistency and reliability, as it can yield different efficiency metric values even for the same CGM, see Okundalaye (2015). There is an efficiency metric in-built in the CGM known as the sufficient descent condition. This device is designed to determine the degree of reduction in the value of an objective function, see Babaie-Kafaki and Reza (2014), and as well measure the convergence of the method. In order to take care of the deficiencies associated with performance profiling and the analytical device of Hager and Zhang, the coefficient in the sufficient descent condition is proposed as a metric for assessment of computational performance of CGMs. The parameter in the sufficient descent condition is the most appropriate metric for efficiency of CGMs, because the CGM itself is an iterative procedure for which convergence is of utmost concern. Furthermore, the parameter in the sufficient descent condition can easily be determined by using the practical computations obtained through the use of the algorithm of the CGM.

### 2. Conjugate Gradient Method

The CG method is a computational scheme for solving the unconstrained minimization problem:

$$\text{Minimize } f(x), x \in R^n, \quad \dots \quad (1)$$

where  $R^n$  is an  $n$ -dimensional Euclidean space and  $f: R^n \rightarrow R$  is assumed to be in  $C^1(\mathbb{R}^n)$ .

The CGM was originally devised by Hestenes and Stiefel (1952) for the iterative solution of linear algebraic equations. An extension into the Hilbert space setting was given in Hayes (1954). A new dimension was introduced by Fletcher and Reeves (1964) by its application to nonlinear equations. The doctoral thesis of Daniel (1965) threw more light into further theoretical applicability of the method. Shamamskii (1962)

considered some of the basic convergence theorems for the method. Presently, the method has become an area of intense research. The following publications are cited to illustrate the significant contributions made in the area: Guoyin et al. (2007), Andrei (2008a, 2013), Dong et al. (2015b) and Baluch (2017).

When the CGM is used for minimizing non-quadratic objective functions, the method is known as a nonlinear CGM (Sanmitias and Vercher, 1988). A nonlinear CGM generates a sequence  $\{x_k\}$  using an iterative scheme:

$$x_{k+1} = x_k + \alpha_k d_k, k = 0, 1, 2, \dots \quad \dots \quad (2)$$

where  $x_0 \in R^n$  is an initial point,  $\alpha_k$ , the step size at iteration  $k$ , is such that

$$\alpha_k = \operatorname{argmin} f(x_k + \alpha d_k), \quad \alpha \geq 0 \quad \dots \quad (3)$$

$d_k$  is the search direction given by

$$d_k \begin{cases} -g_k & \text{if } k = 0 \\ -g_k + \beta_k d_{k-1} & \text{if } k \geq 1 \end{cases} \quad \dots \quad (4)$$

in which  $g_k = \nabla f_k$  and  $\beta_k$  is a scalar known as the CG update parameter. Different choices of the update parameter result in different CG algorithms (Zhang and Zhou, 2008).

In a survey of the CG method, Andrei (2008b) identified as many as 40 CG algorithms defined by different expressions of  $\beta_k$  that are used for nonlinear optimization. With new variants of the classical and hybrid CG algorithms appearing very frequently in literature, the number of CG algorithms available by now may have quadrupled the initial figure.

### 3. CGM Efficiency Metrics

From the available literature, there are two major approaches used for assessing the efficiency of CGMs. These are as discussed below.

#### 3.1 Performance Profile

The only approach appearing in the literature in the context of the ensuing discussion is the performance profiling devised by Dolan and Moré (2002) as a tool for benchmarking and comparing optimization software using distribution functions. However, the tool finds ready use in assessing performance of computational methods using the number of test problems a computational method can solve for the metric. The performance profiling procedure is specified as follows: Given a set  $\mathcal{P}$  of  $n_p$  problems and a set  $\mathcal{S}$  of  $n_s$  methods, denote by  $t_{p,s}$ , the number of iterations, function/gradient evaluations or CPU time required to solve problem  $p \in \mathcal{P}$  by method  $s \in \mathcal{S}$ . With this, a performance between the methods is defined based on the ratio

$$r_{p,s} = \frac{t_{p,s}}{\min \{t_{p,i} : 1 \leq i \leq n_s\}}, \text{ and a performance profile for each method } s \text{ is defined by the probability}$$

$$\omega_s(\tau) = \frac{1}{n_p} \operatorname{size}\{p : 1 \leq p \leq n_p, \ln(r_{p,s}) \leq \tau, \tau \geq 0\}.$$

For each method, the fraction  $P(r_{p,s} \leq \tau)$  of the problems for which the method is within a factor  $\tau$  of the best time was plotted. The left side of the profile (figure) gives the percentage of the test problems for which a method is the fastest; the right side gives the percentage of the test problems successfully solved by each of the methods. The top curve is the method that solved the most problems in a time that is within a factor  $\tau$  of the best time.

Dolan and Moré's performance profiling is dependent on a number of extraneous factors such as the number and level of difficulty of the test problems used, as well as the tolerance level set for convergence of the problems, see Table 1 in the Appendix.

Ideally, all available problems on an aspect of the computation should be in the set of test problems. But this is not so most of the times, as the sample of the test problems utilized for the computational exercise is often not representative of the population of all the problems. In this case, the outcome of the efficiency assessment could be said to be not statistically valid.

The efficiency measures provided by the Dolan and Moré's software can be a true measure of computational performance if the sampling procedure for the test problems is statistically valid, and the generated performance profile can as well offer useful information for convergence analysis of the method.

#### 3.2 Analytical Measure of CGM Efficiency

The means of efficiency determination proposed in Hager and Zhang (2005, pg. 172) through the use of the Cauchy-Schwartz inequality is herein referred to as an analytical measuring device. Precisely, this efficiency metric is determined from the expression for the descent property, i.e.,

$$g_k^T d_k = -\|g_k\|^2 + \beta_k g_k^T d_{k-1}, \quad \dots \quad (5)$$

(with  $\beta_k$  being the update parameter), using the Cauchy-Schwartz inequality

$$u^T v \leq \frac{1}{2} (\|u\|^2 + \|v\|^2), \quad \dots \quad (6)$$

where  $u$  and  $v$  are vectors.

This measure could have been the most preferred metric for assessing efficiency of CGMs, but for its defects, in that different values are obtainable even for the same CGM, see Okundalaye (2015, pp. 30-31). In addition, the values of the efficiency metric do not tally with the proportion of test problems solved as indicated in the performance profile of the CGM.

### 3.3 Sufficient Descent Condition

A CGM is said to satisfy the descent condition if

$$g_k^T d_k < 0, \forall k = 1, 2, \dots \quad (7)$$

This property assures of reduction in the value of the objective function as the computation progresses. A feature more beneficial is that of sufficient descent condition, i.e.,

$$g_k^T d_k \leq -c g_k^T g_k, \quad c > 0. \quad (8)$$

Equation (8) does not only determine the degree of reduction in the value of the objective function, but is an indicator of convergence of the CGM. This accounts for why (8) is very fundamental in the proof of global convergence of CGMs, see Ahmed and Taher (2021), Masmali, Salleh and Alhawarat (2021) and Dehmiry and Kargarfard (2024).

The coefficient  $c$ , in particular, is a measure of the rate of convergence of the CGM. Thus, the coefficient in the sufficient descent condition is a good metric for efficiency, and it can be determined from practical computations obtained during the implementation of CG algorithms. The efficiency measure obtained this way will be useful for analysis and qualitative comparison of CGMs.

### 4. Concluding Remarks

The essence of this publication is to expose researchers in the area of CGMs to an important aspect of ensuring the need to assess the efficiency of their new methods, as well as to keep them inform of possible enhancements in the prescription of another device for carrying out performance evaluation and validation. The following are remarks on the essence of the discussion so far:

- i. In addition to the two existing devices for measuring performance of CGMs, namely, the performance profiling software of Donlan & More, and Hager and Zhang's Cauchy-Schwartz inequality-based approach, another efficiency metric is being proposed with the use of the coefficient in the sufficient descent condition. As far as we know from the available literature, this is the first time the coefficient in the sufficient descent condition is being suggested for use as a CGM efficiency measure.
- ii. Both performance profiling and parameter in the sufficient descent condition, which make use of practical computations, are more suitable for the CGM, which is computational in nature, than Hager & Zhang device which presents itself as an analytical tool.
- iii. The major drawback in the use of performance profiling as an efficiency metric is difficult to overcome as the population of test problems keeps increasing with the inclusion of sophisticated and complex problems characteristics of the challenges of technology are unabated.
- iv. As a way of correcting the defects in the analytical metric of Hager & Zhang, the existing procedure should be reviewed to bring about uniqueness in the coefficients of vectors  $u$  and  $v$ .
- v. The format for the output of the three (3) metrics can be unified for ease of comparison by normalizing the sufficient condition as

$$g_i^T d_i \leq -\bar{c} g_i^T g_i, \quad 0 < \bar{c} = \frac{c}{\max_i c} \leq 1. \quad (9)$$

### References

- Ahmed, C. & Taher, B. (2021). Global convergence of new conjugate gradient method with inexact line search. *International Journal of Electrical and Computer Engineering* 11(2), 1469-1475.
- Andrei, N. (2008a). A Dai–Yuan conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization. *Appl. Math. Lett.* **21**(2), 165–171.
- Andrei, N. (2008b). 40 conjugate gradient algorithms for unconstrained optimization: A survey on their definition. ICI Technical Report No. 13/08 of Research Institute for Informatics, Center for Advanced Modeling and Optimization, Bucharest, Romania.
- Andrei, N. (2013). Another conjugate gradient algorithm with guaranteed descent and the conjugacy conditions for large-scaled unconstrained optimization. *J. Optim. Theory Appl.* **159**(3), 159–182.
- Babaie-Kafaki, S. and Reza, G. (2014). The Dai–Liao nonlinear conjugate gradient method with optimal parameter choices. *Euro. J. Operat. Res.* **234**(3), 625–630.
- Baluch, B., Salleh, Z., Alhawarat, A. & Roslam, U.A.M. (2017). A new modified three-term conjugate gradient method with sufficient descent property and its global convergence. *Hindawi J. Math.* **2017**, 1 – 12.

- Bamigbola, O.M. (2020). Shared links on Results of Computational Experiments with the CGM shared. Unpublished Report in the Department of Mathematics, University of Ilorin, Nigeria.
- Daniel, J. W. (1965). The conjugate gradient method for linear and non-linear operator equations. Unpublished Ph.D. Dissertation, Stanford University, Stanford, California.
- Dehmiry, A. H., & Kargarfard, M. (2024). Global convergence of three-term conjugate gradient methods on general functions under a new inexact line search strategy. *Engineering Optimization*, 56(2), 240–251.
- Dong, X.L., Liu, H.W., He, Y.B. and Yang, X.M. (2015a). A modified Hestenes-Stiefel conjugate gradient method with sufficient descent condition and conjugacy. *J. Comp. & Appl. Math.* 281, 239 – 249.
- Dong X.L. Liu H. and He Y. H. (2015b). A self-adjusting conjugate gradient method with sufficient descent condition and conjugacy condition. *J. Optim. Theory Appl.* 165(1), 225–241.
- Dolan, E.D. & Moré, J.J. (2002). Benchmarking optimization software with performance profiles, *Mathematical Programming*, 91, 201-213.
- Fletcher, R. and Reeves, G. M. (1964): Function minimization by conjugate gradients. *Computer J.* 7, 149-154.
- Guoyin, L., Chunming, T. and Zengxin, W. (2007). New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. *J. Comp. Appl. Math.* 202(2), 523-539.
- Hager, W.W. and Zhang, H. (2005). A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J. Optim.* 16(1), 170–192.
- Hayes, R. M. (1954): Iterative methods for solving linear problems in Hilbert space. *Nat. Bur. Standards, Applied Mathematics*, 39.
- Hestenes, M. and Stiefel, E. (1952). Method of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards* 49, 409-436.
- Masmali, I.A. Salleh, Z. & Alhawarat, A. (2021). A decent three term conjugate gradient method with global convergence properties for large scale unconstrained optimization problems. *AIMS Mathematics* 6(10), 10742-10764
- Okundalaye, O.O. (2015). Improved hybrid BFGS-CG method for solving unconstrained optimization problems. Unpublished M. Sc. Dissertation, Department of Mathematics, University of Ilorin, Ilorin, Nigeria.
- Sanmtias, S. and Vercher, E. (1988). A generalized conjugate gradient algorithm. *Journal of Optimization Theory and Applications* 98, 489-502.
- Shamanskii, V. (1962): On some computational schemes for iterative processes. *Ukrain. Mat. Zh.* 14, 100 – 109.
- Shi, Z. J. and Guo, J. (2008) A new algorithm of nonlinear conjugate gradient method with strong convergence. *Math. Appl. comput.* 27(1), 1 -16.
- Wei, Z., Li, G. and Oi, L. (2006). New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems. *Applied Mathematics and Computation* 179, 467- 480.
- Zhang, L. and Zhou, W. (2008). Two descent hybrid conjugate gradient methods for optimization. *Journal of Computational and Applied Mathematics* 216, 251-26.

APPENDIX

Table 1: Sample of convergence status at the end of computations for selected CGMs

T	Selected CGMs											T	Selected CGMs											T	Selected CGMs										
	Dim	BA	F	P	H	C	D	L	H	MM	DL		P	Dim	BA	F	P	H	C	D	L	H	MM		D	P	Dim	BA	F	P	H	C	D	L	H
1	50	Y	D	C	Y	Y	D	Y	D	Y	D	5	50	C	C	C	C	D	C	C	C	D	C	9	50	C	C	Y	C	D	C	C	Y	D	C
	100	Y	D	C	Y	Y	D	Y	D	Y	D		100	C	Y	C	C	D	Y	C	Y	D	C		100	D	C	Y	D	D	C	C	Y	D	C
	500	C	D	Y	C	Y	D	D	D	D	D		500	Y	C	C	Y	D	C	C	Y	D	C		500	D	C	Y	D	D	C	C	Y	D	D
	1000	C	D	Y	C	D	D	D	D	D	D		1000	Y	C	C	Y	D	C	C	C	D	C		1000	D	C	Y	D	D	Y	C	Y	D	C
	5000	Y	D	Y	Y	D	D	D	D	D	D		5000	Y	C	C	Y	D	C	C	Y	D	C		5000	Y	D	Y	D	D	Y	Y	Y	D	D
	10000	Y	D	Y	Y	D	D	D	D	D	D		10000	Y	C	C	Y	D	C	C	Y	D	C		10000	Y	D	Y	D	D	Y	Y	Y	D	D
	0	Y	D	C	Y	D	D	D	D	D	D		0	Y	C	C	Y	D	C	C	Y	D	C		0	Y	D	Y	D	D	Y	Y	Y	D	D
	50	D	D	D	D	D	D	D	D	D	D		60	C	C	C	C	Y	C	C	C	D	C		50	D	D	D	D	D	D	D	D	D	D
	100	D	D	D	D	D	D	D	D	D	D		120	C	C	C	C	Y	C	C	C	D	C		100	D	D	Y	D	D	D	D	D	D	D
	500	D	D	D	D	D	D	D	D	D	D		600	C	C	C	C	D	C	C	C	D	C		500	D	D	D	D	D	D	D	D	D	D
2	1000	D	D	D	D	D	D	D	D	D	1200	C	C	C	C	D	C	C	C	D	C	1000	D	D	D	D	D	D	D	D	D	D			
	5000	D	D	D	D	D	D	D	D	D	6000	C	C	C	C	D	C	C	C	D	C	5000	D	D	D	D	D	D	D	D	D	D			
	10000	D	D	D	D	D	D	D	D	D	12000	C	C	C	C	D	C	C	C	D	C	10000	D	D	D	D	D	D	D	D	D	D			
	0	D	C	D	D	D	D	D	D	D	0	C	C	C	C	D	C	C	C	D	C	0	D	D	D	D	D	D	D	D	D	D			
	50	D	D	D	D	D	D	D	D	D	60	D	C	C	C	Y	C	C	C	D	C	50	D	C	C	D	Y	C	C	C	D	D			
	100	D	D	D	D	D	D	D	D	D	120	C	C	C	C	Y	D	C	C	D	C	100	D	C	C	D	Y	C	C	C	Y	C			
	500	D	D	D	D	D	D	D	D	D	600	C	C	C	C	Y	C	C	C	D	C	500	C	C	C	D	Y	C	C	C	D	D			
	1000	D	D	D	D	D	D	D	D	D	1200	C	C	C	C	Y	C	C	C	D	C	1000	D	C	C	D	Y	C	C	C	D	C			
	5000	D	D	D	D	D	D	D	D	D	6000	C	D	C	C	Y	C	C	C	Y	C	5000	D	C	C	C	D	C	C	C	D	D			
	10000	D	D	D	D	D	D	D	D	D	12000	C	D	C	C	Y	C	C	C	Y	C	10000	D	C	C	C	D	C	C	C	D	D			
3	0	D	D	D	D	D	D	D	D	D	7	0	D	D	C	C	D	D	C	C	D	C	1	0	C	C	C	D	D	C	C	C	D	C	
	50	C	C	C	C	D	C	C	C	D		C	50	C	Y	C	C	D	C	C	C	C		C	50	D	C	C	C	D	C	C	C	D	D
	100	C	C	C	C	Y	C	C	C	Y		C	100	C	Y	C	C	D	C	C	C	C		C	100	D	C	C	C	D	C	C	C	D	C
	500	C	C	C	C	D	C	C	C	Y		C	500	C	Y	C	C	D	Y	C	C	C		C	500	D	C	C	C	D	C	C	Y	D	C
	1000	C	C	C	C	D	C	C	C	Y		C	1000	C	Y	C	C	D	Y	C	C	C		C	1000	D	C	C	Y	D	Y	C	Y	D	C
4	5000	C	C	C	C	Y	C	C	C	D	C	8	5000	C	Y	C	C	D	Y	C	C	Y	C	2	5000	D	Y	Y	D	D	Y	C	Y	D	C



1000	Y	Y	Y	Y	D	Y	Y	Y	D	Y	1000	D	D	C	C	D	C	C	C	D	C	1000	C	C	C	C	Y	C	C	C	D	C	
5000	Y	Y	Y	Y	D	Y	Y	Y	D	Y	5000	D	C	C	C	D	C	C	D	D	C	5000	Y	Y	C	C	Y	Y	C	C	Y	C	
1000											1000											1000											
0	D	Y	Y	Y	D	Y	Y	Y	D	Y	0	D	C	C	D	D	D	C	C	D	C	0	Y	D	C	C	Y	Y	C	C	Y	C	
50	C	C	C	D	C	C	C	C	C	D	50	D	D	D	D	D	D	D	D	D	D	50	D	D	C	C	D	C	C	C	D	D	
100	C	C	C	D	C	C	C	C	C	D	100	D	D	D	D	D	D	D	D	D	D	100	D	C	C	D	D	C	C	C	D	D	
500	C	C	C	D	C	C	C	C	C	D	500	D	D	D	D	D	D	D	D	D	D	500	Y	D	C	C	Y	D	C	C	Y	D	
1000	C	C	C	D	C	C	C	C	C	D	1000	D	D	D	D	D	D	D	D	D	D	1000	D	D	C	D	D	D	C	C	D	C	
5000	C	C	C	D	C	C	C	C	C	D	5000	D	D	D	D	D	D	D	D	D	D	5000	D	D	C	C	D	D	C	C	D	C	
1											2											3											
0	0	C	C	C	D	C	C	C	C	D	4	0	D	D	D	D	D	D	D	D	D	8	0	D	D	C	C	D	D	C	C	D	D
	50	D	C	C	D	C	C	C	Y	D	C	50	D	D	D	D	D	D	D	D	D	50	D	D	C	D	D	D	D	D	D	D	
	100	C	C	C	D	C	C	C	Y	D	C	100	D	D	D	D	D	D	D	D	D	100	D	D	C	D	D	D	D	D	D	D	
	500	C	C	C	D	C	C	C	Y	D	C	500	D	D	D	D	D	D	D	D	D	500	D	D	C	D	D	D	D	C	D	D	
	1000	C	C	C	D	C	C	C	Y	D	C	1000	D	D	D	D	D	D	D	D	D	1000	D	D	C	D	D	D	D	D	D	D	
	5000	C	C	C	D	C	C	C	Y	D	C	5000	D	D	D	D	D	D	D	D	D	5000	D	D	C	D	D	D	D	C	D	D	
1											2											3											
1	0	C	C	C	D	C	C	C	Y	D	C	5	0	D	D	D	D	D	D	D	D	9	0	D	D	D	D	D	D	D	C	D	D
	50	D	D	D	D	D	D	D	D	D	D	60	D	D	D	D	D	D	D	D	D	50	C	C	C	C	D	C	Y	C	D	C	
	100	D	D	D	D	D	D	D	D	D	D	120	D	D	D	D	D	D	D	D	D	100	C	C	C	C	D	C	C	C	D	C	
	500	D	D	D	D	D	D	D	D	D	D	600	D	D	D	D	D	D	D	D	D	500	C	C	C	C	D	C	C	C	D	C	
	1000	D	D	D	D	D	D	D	D	D	D	1200	D	D	D	D	D	D	D	D	D	1000	Y	C	C	Y	D	C	C	Y	D	C	
	5000	D	D	D	D	D	D	D	D	D	D	6000	D	D	D	D	D	D	D	D	D	5000	Y	C	Y	Y	D	C	C	Y	D	C	
1											2											4											
2	0	D	D	D	D	D	D	D	D	D	6	0	D	D	D	D	D	D	D	D	D	0	0	Y	Y	Y	Y	D	Y	C	Y	D	C
	50	D	C	C	D	C	C	C	C	Y	C	50	D	D	D	D	D	D	D	D	C	50	C	C	C	C	Y	C	C	C	D	C	
	100	D	C	C	D	C	C	C	C	D	C	100	D	D	D	D	D	D	D	D	D	100	C	C	C	C	Y	C	C	C	D	C	
	500	D	C	C	D	C	C	C	C	D	C	500	D	D	D	D	D	D	D	D	D	500	C	C	C	C	Y	C	C	C	Y	C	
	1000	D	C	C	D	C	C	C	C	D	C	1000	D	D	D	D	D	D	D	D	C	1000	C	Y	C	C	Y	C	C	C	Y	C	
	5000	D	C	C	D	C	C	C	Y	D	C	5000	D	D	D	D	D	D	D	D	D	5000	C	Y	C	C	Y	C	C	Y	Y	C	
1											2											4											
3	0	D	C	C	D	C	C	C	Y	D	C	7	0	D	D	D	D	D	D	D	D	1	0	C	Y	C	C	Y	C	C	Y	Y	C

NB: C denotes Convergence D represents Divergence

Y stands for Yet to converge.