

INTRODUCING ADAPTED NELDER & MEAD'S DOWNHILL SIMPLEX METHOD TO A FULLY AUTOMATED ANALYSIS OF ECLIPSING BINARIES

A. Prša, T. Zwitter

University of Ljubljana, Dept. of Astrophysics, Jadranska 19, 1000 Ljubljana, Slovenia

ABSTRACT

Eclipsing binaries are extremely attractive objects because absolute physical parameters (masses, luminosities, radii) of both components may be determined from observations. Since most efforts to extract these parameters were based on dedicated observing programmes, existing modelling code is based on interactivity. Gaia will make a revolutionary advance in sheer number of observed eclipsing binaries and new methods for automatic handling must be introduced and thoroughly tested. This paper focuses on Nelder & Mead's downhill simplex method applied to a synthetically created test binary as it will be observed by Gaia.

Key words: Binaries: eclipsing; Methods: numerical, statistical; Missions: Gaia.

1. INTRODUCTION

Many assessments have already been done for Gaia's expected harvest of eclipsing binaries (EB) to $V < 15$, where both photometric and RV observations will be available (see Munari et al. 2001, Zwitter et al. 2003, Marrese et al. 2004 and others). Out of 50 million observed stars, roughly 100 000 will be double-lined eclipsing binaries. However, based on experience from Hipparcos, out of 1 billion stars observed to $V < 20$, there will be ~ 2 million (Eyer 2005) eclipsing binaries without spectroscopic observations, but with quite decent photometric accuracy. This study presents a new approach developed for automatic reduction of observed data along with an estimate of how much we may expect to obtain from them.

2. THE METHOD

Obtaining physical parameters from observations is an inverse problem solved numerically by a modelling programme. Affirmed and most widely used is the WD code (Wilson & Devinney 1971), which features Differential Corrections method (DC) powered by the Method

of Multiple Subsets (MMS) (Wilson 1993). The DC method has already been applied successfully to automatic parameter extraction (see e.g., Wyithe & Wilson 2001, 2002; Prša 2003), but its original philosophy is based on interactive monitoring of each convergence step. The algorithm is very fast and works well if the discrepancy between the observed and computed curves is relatively small, but it tends to diverge or give physically unmeaningful results if the discrepancy is large. As part of an effort to create a reliable and powerful package for EB analysis, a complementing minimization scheme is proposed.

2.1. Nelder & Mead's Downhill Simplex

There are two main deficiencies of the DC method that are especially striking. 1) Once a DC method converges to a minimum, there is no way of telling whether that minimum is local or global; even if it is local, the method is stuck and cannot escape. 2) The main source of divergence and the loss of accuracy in the DC algorithm is the computation of numerical derivatives.

To circumvent these two problems, Nelder & Mead's downhill Simplex¹ method (Nelder & Mead 1965), hereafter NMS, is implemented. Since NMS doesn't compute derivatives but relies only on function evaluations, it is a promising candidate for our purpose. The basic form of the NMS method along with WD implementation was first proposed by Kallrath & Linnell (1987). We take a step further and adapt the method specifically to EBs.

The NMS method acts in n -dimensional parameter hyperspace. It constructs n vectors \mathbf{p}_i from the vector of initial parameter values \mathbf{x} and the vector of step-sizes \mathbf{s} as follows:

$$\mathbf{p}_i = (x_0, x_1, \dots, x_i + s_i, \dots, x_n) \quad (1)$$

These vectors form $(n + 1)$ vertices of an n -dimensional simplex. During each iteration step, the algorithm tries to improve parameter vectors \mathbf{p}_i by modifying the vertex with the highest function value by simple geometrical

¹Nelder & Mead's downhill simplex is not to be confused with linear or non-linear programming algorithms, which are also referred to as Simplex methods.

transformations: reflection, reflection followed by expansion, contraction and multiple contraction (Galassi et al. 2003). Using these transformations, the simplex moves through parameter space towards the deepest minimum, where it contracts itself.

PHOEBE² is a software package built on top of the WD code that extends its basic functionality to encompass, among other extensions summarized in Prša & Zwitter (2005), the NMS method. It is especially suited for EBs: powered by heuristic scans, parameter kicking and conditional constraining, the method is able to efficiently escape from local minima.

2.2. Heuristic Scan

NMS is a robust method that always converges, but it can converge to a local minimum, particularly since parameter hyperspace in the vicinity of the global minimum is typically very flat, with lots of local minima. In addition, the global minimum may be shadowed by data noise and degeneracy.

The heuristic scan is a method by which a minimization algorithm selects a set of starting points in parameter hyperspace and starts the minimization from each such point. It then sorts all solutions by the cost function (the χ^2 , for example) and calculates parameter histograms and convergence tracers for given hyperspace cross-sections (specific examples are given in Section 4). The way of how the algorithm selects starting points is determined by the user: the points may be gridded, stochastically dispersed, distributed according to some probability distribution function (PDF) etc. The basic idea of the heuristic scan is to obtain decent statistics of adjusted parameter values from which a fair and realistic error estimate may be given.

2.3. Parameter Kicking

Another possible approach to detect and escape from local minima is to use some stochastic method like Simulated Annealing (SA). However, such methods are notoriously slow. Since the EB hyperspace is typically very flat, stochastic methods would be practical only in the vicinity of the global minimum. Thus instead of a full-featured SA scan, a simple new procedure has been developed that achieves the same effect as stochastic methods, but in significantly shorter time.

The idea is as follows: whenever NMS reaches a minimum within a given accuracy, the algorithm runs a globality assessment on that minimum. If we presume that standard deviations σ_k of observations are estimated properly and that they apply to all data points regardless of phase or nightly variations, we may use them for χ^2 weighting:

$$\chi_k^2 = \sum_{i=1}^M w_k (x_i - \bar{x})^2 = \frac{1}{\sigma_k^2} \sum_{i=1}^M (x_i - \bar{x})^2, \quad (2)$$

where index i runs over M measurements within a single data set and index k runs over N data sets (different photometric curves). Since the variance is given by:

$$s_k^2 = \frac{1}{N_k - 1} \sum_i (x_i - \bar{x})^2, \quad (3)$$

we may readily express χ_k^2 as:

$$\chi_k^2 = (N_k - 1) \frac{s_k^2}{\sigma_k^2} \quad (4)$$

and the overall χ^2 value as:

$$\chi^2 = \sum_k (N_k - 1) \left(\frac{s_k}{\sigma_k} \right)^2 \quad (5)$$

If σ_k are fair and all data sets contain approximately the same number of observations, then the ratio s_k/σ_k is of the order unity and χ^2 of the order NM . This we use for parametrizing χ^2 values:

$$\chi_0^2 = NM, \quad \lambda := (\chi^2/\chi_0^2) : \text{quantization.} \quad (6)$$

Parameter kicking is a way of knocking the obtained parameter-set out of the minimum: using the Gaussian PDF, the method randomly picks an offset for each parameter. The strength of the kick is determined by the Gaussian dispersion σ_{kick} , which depends on λ : if the value is high, then the kick should be strong, but if it is low, i.e., around $\lambda \sim 1$, then only subtle perturbations should be allowed. Experience shows that a simple expression such as:

$$\sigma_{\text{kick}} = \frac{0.5\lambda}{100} \quad (7)$$

works very reliably. This causes σ_{kick} to assume a value of 0.5 for 10σ offsets and 0.005 for 1σ offsets, being linear in between. Note that this σ_{kick} is *relative*, i.e., given by:

$$\sigma_{\text{kick}}^{\text{abs}} = x \sigma_{\text{kick}}^{\text{rel}}, \quad (8)$$

where x is the value of the given parameter.

2.4. Conditional Constraining

Having purely photometric (LC) observations, it is impossible to determine absolute physical parameters of the observed binary. However, if the distance to EB is measured independently or if additional assumptions about the EB are set, even purely photometric observations can yield absolute values of physical parameters. If additional constraints are imposed on the model from the outside, the model is referred to as *conditionally constrained* (CC'd).

Two CCs are immediately evident: 1) astrometric measurements on-board Gaia may be used to measure the distance with the accuracy of $\sim 11 \mu\text{as}$ at $V = 15$ to $\sim 165 \mu\text{as}$ at $V = 20$ (ESA 2000); 2) since a substantial number of stars are main-sequence stars, M-L, L-T and T-R relations may be adopted as constraints. See Prša & Zwitter (2005) for details on CC implementation in PHOEBE.

²<http://phoebe.fiz.uni-lj.si>

3. SIMULATION

To test the suitability of NMS for EBs, we built a partially-eclipsing synthetic main-sequence F8 V–G1 V binary using PHOEBE. Table 1 lists all of its principal parameters. The simulation presented in this paper is based exclusively on photometric data: two sets of observations corresponding to Johnson B and V filters are created, each with 82 points with Poissonian scatter $\sigma_{\text{obs}} = 0.02$ mag at $V = 20$, values typical to expect from Gaia (ESA 2000).

Simulation flow is the following: all physical parameters set for adjustment (a , i , T_1 , T_2 , Ω_1 , Ω_2 , L_1 and L_2) were displaced by $\sim 50\%$. The obtained set was used as the initial guess for the NMS. In the first part of the simulation the method converges to a solution using only heuristic scan and parameter kicking, which yields *relative* values of parameters. In the second part the simulation was additionally CC'd by the main-sequence constraint to obtain *absolute* values of parameters.

Table 1. Principal parameters of the simulated binary.

Parameter [units]	Binary	
	F8 V	G1 V
P_0 [days]		1.000
a [R_\odot]		5.524
$q = m_2/m_1$		0.831
i [$^\circ$]		85.000
T_{eff} [K]	6200	5860
L [L_\odot]	2.100	1.100
M [M_\odot]	1.236	1.028
R [R_\odot]	1.259	1.020
Ω [–] ^(a)	5.244	5.599

^(a)Unitless effective potentials defined by Wilson (1979).

4. RESULTS

An heuristic scan over all adjustable parameters has been performed to obtain accurate convergence statistics: over 2000 uniformly distributed starting points in parameter hyperspace were used during the simulation. We present the results of the overall minimization step-by-step.

a) Convergence assessment. Depending on the bumpiness of the hyperspace, an heuristic scan will generally yield different solutions from different starting points; it is our hope that only few of these solutions will account for most scans. To evaluate their quality, the globality assessment mechanism introduced in Section 2.3 is used to sort solutions by the depth of the reached minimum. Tests show that the NMS method itself is all-too-often stuck in local minima and only $\sim 15\%$ of all runs end up within one per cent of ideal λ ($\lambda=1$ in case of Figure 1). However, tests also show that parameter kicking *significantly* improves this percentage ($\sim 50\%$ after the first

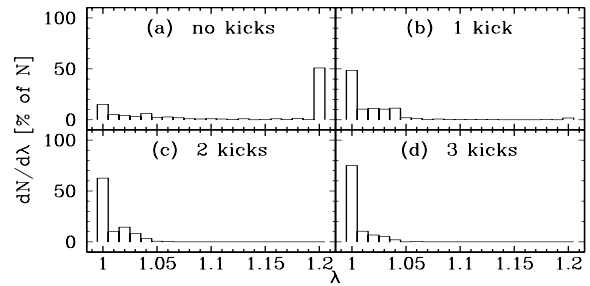


Figure 1. λ -histograms for 3 consecutive parameter kicks. Top-left figure demonstrates how numerous are local minima and how difficult it is for NMS to circumvent them. The other three figures show significant improvement by using parameter kicking. Histograms consist of 20 bins (single bin width is 0.01). The last bin encompasses all higher values of λ . Labels (a) through (d) are used consistently throughout the paper.

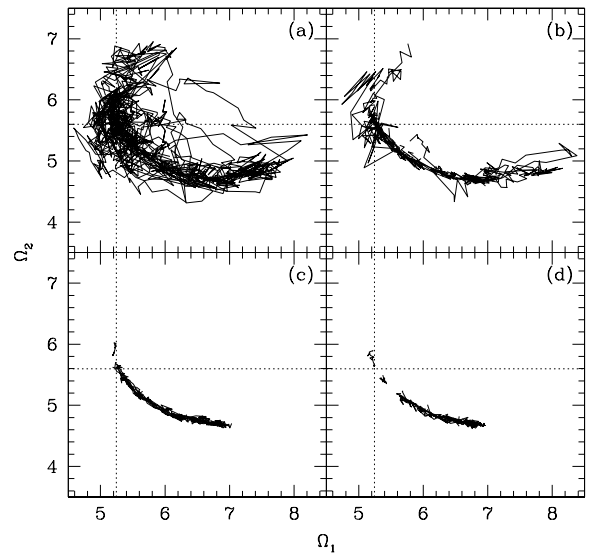


Figure 2. Convergence tracers for 3 consecutive parameter kicks. These plots trace convergence steps within 2D cross-sections of the hyperspace, revealing areas of minima and degeneracy. Cross-hairs mark the location of the global minimum.

kick, $\sim 63\%$ after the second and $\sim 75\%$ after the third kick); even more, parameter kicking also enhances convergence speed after each kick.

Although the values of λ may seem promising, they don't necessarily guarantee that the corresponding solution is optimal. Rather, additional assessments should be done. Figure 2 shows *convergence tracers*: 2D cross-sections of parameter hyperspace tracing convergence from each starting point of heuristic scan to the corresponding solution. Such tracers clearly show areas of minima and degeneracy. Since the location of the global minimum is not known in real life, extra care should be taken to *never*

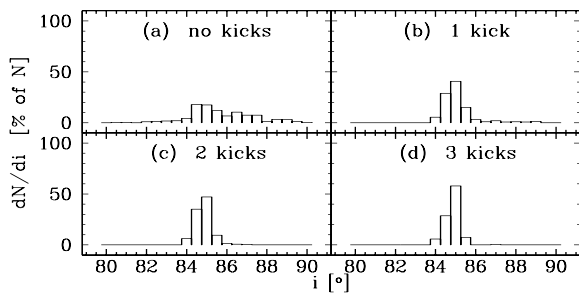


Figure 3. Histogram of the inclination for 3 consecutive parameter kicks. Histogram consists of 20 bins with 0.5° each. The solution is symmetric to $i = 90^\circ$, but we adopt $i < 90^\circ$ by convention.

blindly trust the statistics of such a degenerate problem.

b) Statistics of obtained parameters. The usual practice in the literature, when listing obtained parameters from the model, is to give their values with *formal* errors, i.e., standard deviations reported by the used numerical method. These errors are often too optimistic, since degeneracy and noise noticeably contribute to the overall error. NMS powered by heuristic scan and parameter kicking has the advantage of obtaining parameter errors *statistically*, independent of the method itself. Figure 3 shows an example of an obtained histogram for inclination i . It is evident that for the NMS without parameter kicking (and similarly for **any** other numerical method that cannot escape from local minima) any eclipsing system is a tie; for a fully minimized solution (bottom right plot on Figure 3) the error is simply standard deviation of the Gaussian being fitted over the histogram, yielding roughly 0.5° .

c) Conditional constraining. Inclination is the only intrinsic parameter that may be obtained in an absolute sense from photometric observations. Using CC this deficiency is removed: any particular CC adds one or more implicit parameter ties into the system. It basically introduces an intersection plane with the otherwise degenerate part of the hyperspace, thus eliminating degeneracy. Figure 4 demonstrates how a main-sequence constraint breaks the degeneracy for gravity potentials $\Omega_{1,2}$.

5. DISCUSSION

The idea behind the NMS implementation is *not* to replace the DC method but to *complement* it. DC is created for interactive usage and converges in discrete steps that need monitoring. NMS on the other hand aims to automate this process so that intermediate monitoring is no longer necessary, which is a key goal for Gaia. DC is one of the fastest methods (WD's DC in particular, since it is optimized for EBs), but may easily diverge. At the expense of speed, NMS is one of the most robust algorithms for solving non-linear problems and never diverges. Finally, both DC and NMS methods suffer from degeneracy and may be stuck in local minima. To overcome this,

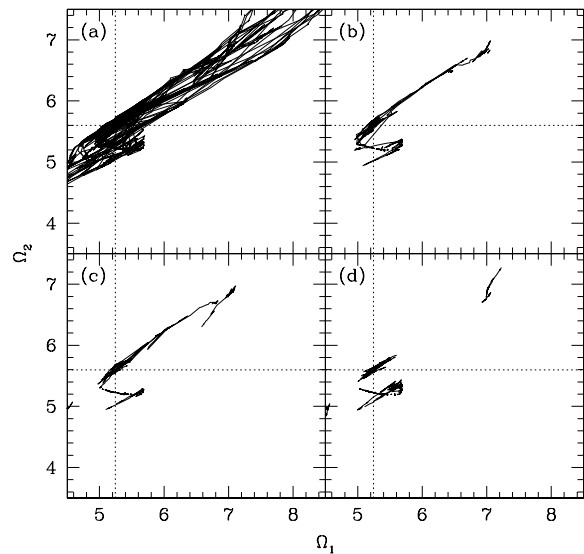


Figure 4. Main-sequence constrained convergence tracers for 3 consecutive parameter kicks. Comparing this result to Figure 2 clearly shows that the intersection of both areas indeed gives the right solution. This is expected, since our synthetic binary is in fact a main-sequence binary.

DC is complemented by the MMS and NMS is complemented by heuristic scan and parameter kicking. These differences in intent make a combination of both methods a powerful engine for solving the inverse problem.

REFERENCES

- ESA, 2000, 'Gaia Concept and Technology Study Report', ESA-SCI(2000)4
- Eyer, L., 2005, ESA SP-576, this volume
- Galassi, M., et al., 2003, ISBN: 0-9541617-3-4
- Kallrath, J., Linnell, A.P., 1987, ApJ 313, 346
- Marrese, P.M., Munari, U., Siviero, A., et al., 2004, A&A 413, 635
- Munari, U., Tomov, T., Zwitter, T., et al., 2001, A&A 378, 477
- Nelder, J.A., Mead, R., 1965, Comput. J. 7, 308
- Prša, A., 2003, ASP Conf. Ser. 298, 457
- Prša, A., Zwitter, T., 2005, submitted to PASP
- Wilson, R.E., Devinney, E.J., 1971, ApJ 166, 605
- Wilson, R.E., 1979, ApJ, 234, 1054
- Wilson, R.E., 1993, CASP, 91
- Wyithe, J.S.B., Wilson, R.E., 2001, ApJ 559, 260
- Wyithe, J.S.B., Wilson, R.E., 2002, ApJ 571, 293
- Zwitter, T., Munari, U., Marrese, P.M., et al., 2003, A&A 404, 333