

## Chapter 4

# Finite Difference Gradient Information

In the masters thesis Thoresson (2003), the feasibility of using gradient-based approximation methods for the optimisation of the spring and damper characteristics of an off-road vehicle, for both ride comfort and handling, was investigated. The Sequential Quadratic Programming (SQP) algorithm and the locally developed Dynamic-Q method were the two successive approximation methods used for the optimisation. The determination of the objective function value is performed using computationally expensive numerical simulations that exhibit severe inherent numerical noise. The use of forward finite differences and central finite differences for the determination of the gradients of the objective function within Dynamic-Q is also investigated. The results of this study, presented here, proved that the use of central finite differencing for gradient information improved the optimisation convergence history, and helped to reduce the difficulties associated with noise in the objective and constraint functions.

This chapter presents the feasibility investigation of using gradient-based successive approximation methods to overcome the problems of poor gradient information due to severe numerical noise. The two approximation methods applied here are the locally developed Dynamic-Q optimisation algorithm of Snyman and Hay (2002) and the well known Sequential Quadratic



Programming (SQP) algorithm, the MATLAB implementation being used for this research (Mathworks 2000b). This chapter aims to provide the reader with more information regarding the Dynamic-Q successive approximation algorithm, that may be used as an alternative to the more established SQP method. Both algorithms are evaluated for effectiveness and efficiency in determining optimal spring and damper characteristics for both ride comfort and handling of a vehicle.

The initial vehicle model was used in this part of the optimisation investigation. This vehicle is fitted with the  $4S_4$  suspension system, and up to four design variables are considered in the optimisation of the suspension characteristics. The details of which were discussed in detail in Chapter 3, Section 3.1.

It is found that both optimisation algorithms perform well in optimising handling. However, difficulties are encountered in obtaining improvements in the design process when ride comfort is considered. This is attributed to the very noisy nature of the ride comfort objective function, which incorporates computed vertical accelerations. Nevertheless, meaningful design configurations are still achievable through the proposed optimisation process, at a relatively low cost in terms of the number of simulations that have to be performed.

## 4.1 Optimisation Algorithms

The following optimisation algorithms are evaluated in this chapter:

- The Dynamic-Q method (see Section 2.3) which constructs a sequence of simple spherical quadratic approximations to the original problem, and successively solves these sub-problems via the LFOPC (leapfrog) algorithm (Snyman 2000). The gradients used by the algorithm are evaluated by forward finite differences, or by central finite differences, with sufficiently large steps to smooth out the numerical noise.



- The Sequential Quadratic Programming (SQP) method (see Section 2.2) with Broyden- Fletcher-Goldfarb-Shanno approximations to the Hessian matrix and one-dimensional minimization by quadratic polynomial interpolation in the search directions. This method is invoked by the MATLAB m-file `fmincon`, which also uses forward finite difference approximations for the gradients.

## 4.2 Gradient Approximation Methods

Most gradient-based optimisation algorithms require the determination of the first and/or second order gradient information of the objective and constraint functions with respect to the design variables. In most engineering optimisation problems this gradient information is not analytically available. The only information available to the designer is the values of objective and constraint functions obtained via expensive simulations. This research investigated the use of forward and central finite differences in the Dynamic-Q optimisation algorithm, for the determination of the first order gradient information. The formulation of forward finite difference gradient information was presented in Section 2.4.1, and central finite difference gradient determination described in Section 2.4.2.

## 4.3 Optimisation

The vehicle model used is the initial vehicle model that was modelled in ADAMS View 12 and described in Section 3.1. This model made use of the 521 tyre model, which was later found to be not sufficiently suited for the desired correlation with measured data. In this initial study the optimisation parameters were defined as in the following sections.

### 4.3.1 Design Variables

In choosing the design variables for optimisation, the assumption is made that the left hand and right hand suspension settings will be the same, but that front and rear settings may differ. The design variables chosen for



optimisation are the static gas volume (Figure 4.1), and damper force scale factor (Figure 4.2), on both the front and rear axles. Thus there are two variables per axle.

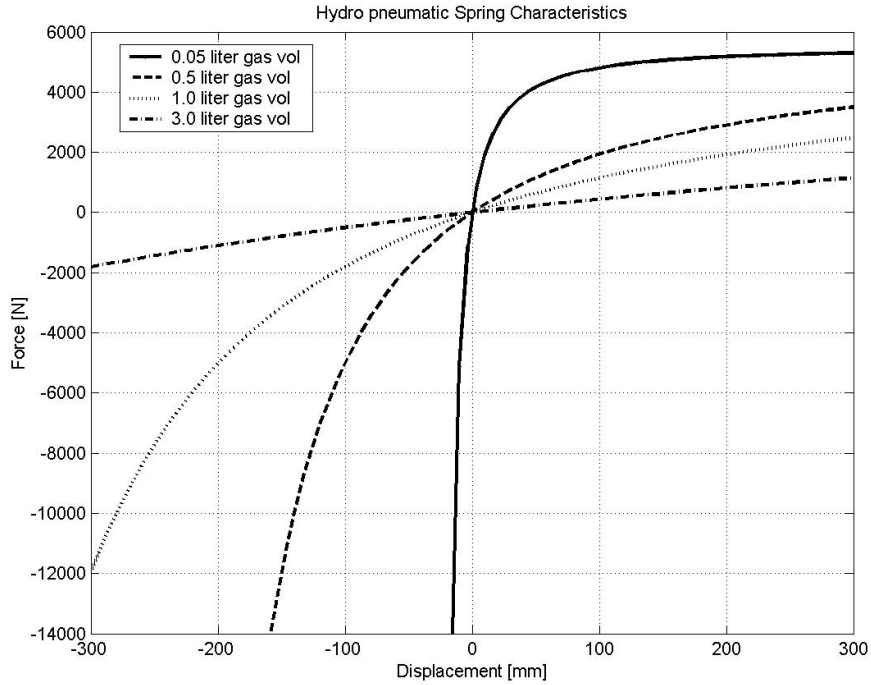


Figure 4.1: Definition of spring characteristics for various gas volumes

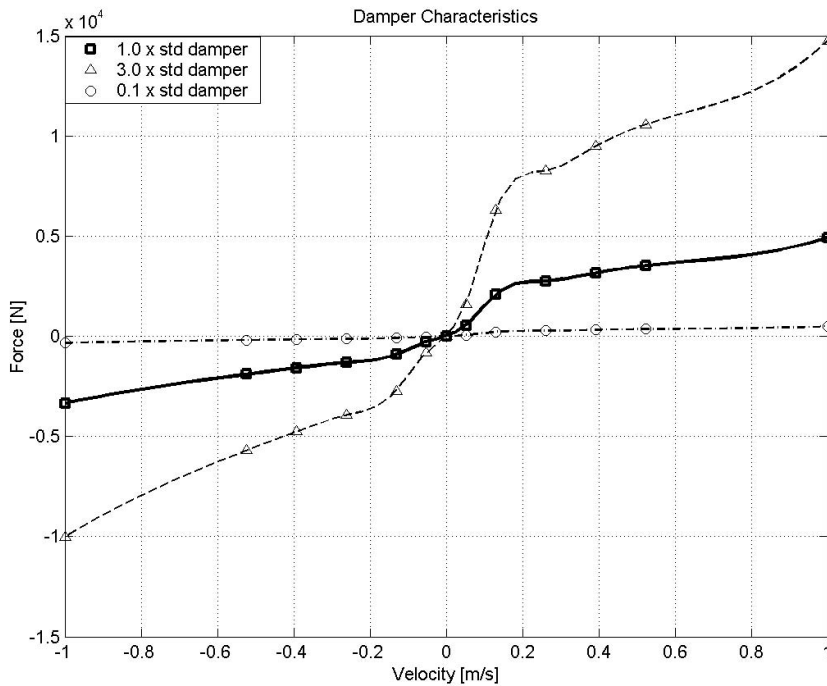


Figure 4.2: Definition of damper characteristics for various damper scale factors



For this initial study the standard damper force characteristic is multiplied by a factor which constitutes the damping design variable (Figure 4.2). The general shape and switch velocities of the damper are thus kept the same. This chapter only considers the cases of two and four design variables, which respectively corresponds to the case where the spring and damper characteristics are identical for the front and rear axles (two design variables), and where they may differ for front and rear (four design variables).

### 4.3.2 Two Variable Case

The two design variable study is an important starting point in the optimisation procedure as it gives the necessary insight into the problem. For this two design variable study, it was decided to use the same design variables as those considered by Els and Uys (2003) in their preliminary study, namely the static gas volume and the damper force scale factor. Figure 4.1 illustrates the spring characteristics for various static gas volumes. Figure 4.2 illustrates the damper characteristics for various damper scale factors.

The static gas volume is denoted by  $gvol$ , and the damper force scale factor by  $dpsf$ . These variables are allowed to range from 0.05 to 3 in magnitude, which are accordingly chosen as upper and lower bounds. The design variables are explicitly defined as follows:

$$x_1 = gvol, \quad x_2 = dpsf \quad (4.1)$$

with bounds

$$0.05 \leq x_i \leq 3, \quad i = 1, 2 \quad (4.2)$$

### 4.3.3 Four Variable Case

For the four design variable problem the front and rear settings are uncoupled. This means that there are separate front and rear damper scale factors and front and rear spring static gas volumes. This results in two design variables describing the front and two describing the rear, giving four design variables



in total.

The front damper scale factor is denoted by  $dpsff$ , the front static gas volume by  $gvolf$ , the rear damper scale factor by  $dpsfr$ , and the rear static gas volume by  $gvolr$ . These variables are also allowed to range from 0.05 to 3 in magnitude. Thus the design variables are defined explicitly as follows:

$$\begin{aligned} x_1 &= dpsff, & x_2 &= gvolf, \\ x_3 &= dpsfr, & x_4 &= gvolr \end{aligned} \quad (4.3)$$

with bounds

$$0.05 \leq x_i \leq 3, \quad i = 1, \dots, 4 \quad (4.4)$$

#### 4.3.4 Definition of Objective Functions

For ride comfort, the motion of the vehicle is simulated for travelling in a straight line over the Belgian paving (Gerotek 2006) and the sum of driver and rear passengers British Standard (BS6841 1987) weighted root mean square (RMS) vertical accelerations are used for the objective function. The Belgian paving test track used, is located at the Gerotek Test Facilities (Gerotek 2006), and has a ISO8608 (1995) roughness coefficient  $G_{do}$  of  $1 \times 10^{-4} \text{ m}^2/(\text{cycles}/\text{m})$ , and a terrain index  $\omega$  of 4 (Thoresson 2003). In a study performed by Els (2005), it was found that the BS6841 weighed RMS vertical acceleration corresponds well with subjective responses of ride comfort in off-road vehicles. For this reason the weighted RMS vertical accelerations will be used for the objective function, when considering ride comfort optimisation. The motion sickness component was ignored as it requires long run times and the Belgian paving test track is not long enough to evaluate motion sickness. No additional measures were used for the ride comfort objective function, despite numerous studies (Alleyne and Hedrick 1995, Kim and Ro 1998, Pilbeam and Sharp 1996, Miller 1998) where the tyre deflection or force is used as a measure of road holding, when considering a quarter car model. This was ignored as the suspension system has the ability to switch to the handling setting should a handling condition be detected. Also handling



lateral and roll degrees of freedom cannot be simulated with a quarter car model. Kim and Ro (2001) also suggest that tyre deflection is insufficient for evaluating handling parameters. This afforded the optimisation for ride comfort and handling to be done separately.

For handling, the vehicle performs a ISO3888-1 (1999) double lane change manoeuvre at 80 km/h and the maximum body roll angle at the first peak (Els and Uys 2003) is used as the objective function. For this initial investigation only roll angle was used as a measure of handling as suggested by Uys et al. (2006a).

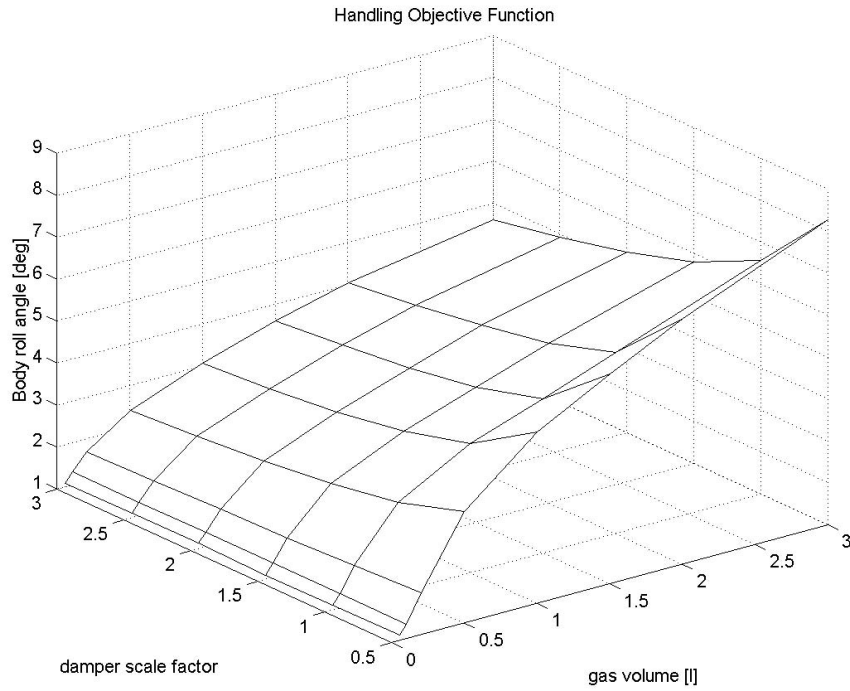
### 4.3.5 Design Space

For the two design variable optimisation, surface plots of the objective function over the complete design space were generated. However, with an increasing number of variables added this is not possible. These objective function surfaces were generated for the optimisation of handling (Figure 4.3) and ride comfort (Figure 4.4) separately. From the figures it can be seen that for excellent handling capability we require high damping and high spring stiffness, however, the damping does not really contribute to the improvement if the spring stiffness is high (a small gas volume in Figure 4.3).

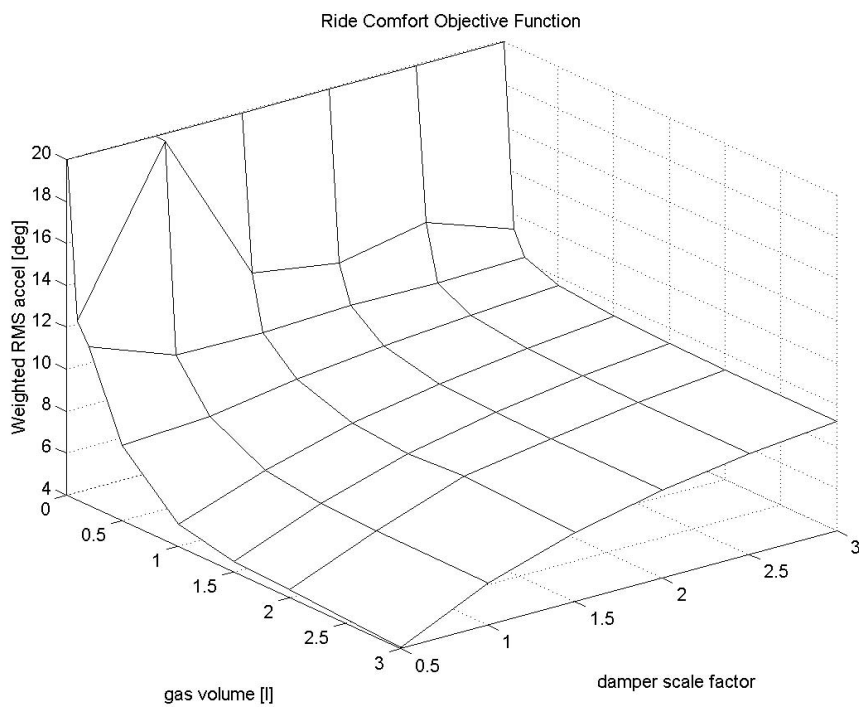
However, for ride comfort (Figure 4.4) we find that the opposite holds. The lowest spring stiffness and low damping is required, but a medium spring stiffness and low damping results in a minimal decrease in ride comfort compared to the optimum. The damper scale factor has a more noticeable effect on the ride comfort, as established previously by Els and Uys (2003) for the heavier version of this vehicle.

### 4.3.6 Handling Results

No significant problems were encountered in applying the algorithms to the optimisation of handling. For handling optimisation with two design variables both algorithms converged to an optimum without difficulty. Figure



**Figure 4.3:** Vehicle roll angle, double lane change at 80 km/h for the two variable design space



**Figure 4.4:** Vehicle Ride comfort, Belgian paving at 60 km/h for the two variable design space





4.5, depicting the convergence histories for each algorithm, plots the objective function value against number of function evaluations. The number of function evaluations is used for the x-axis, instead of iterations, as the cost of the optimisation depends on the number of function evaluations performed, and not the number of iterations, due to the computationally expensive nature of the simulation model. Note that each successive marker on the graphs denotes a new iteration, and that each iteration does not necessarily require the same number of function evaluations. The SQP convergence history for handling optimisation (Figure 4.5) indicates two local minimum solution sets with the same objective function value. This is observed when comparing the two distinct values of design variable  $x(2)$  (at 15 and 30 function evaluations for example) that result in little or no change in the objective function value. Because of the cost of the function evaluations, the objective function values are plotted against cumulative number of function evaluations at the iteration point. The use of Dynamic-Q with 10 % move limit (Figure 4.5) re-iterates the fact that design variable two (damper multiplication factor) has a limited effect on the objective function value as has already been established in Figure 4.3. Using a 20 % move limit (Figure 4.5) Dynamic-Q progresses faster to a minimum. Because of the excellent performance of the forward finite difference method the use of central finite differences at additional cost was not necessary.

The handling optimisation results for four design variables (Figure 4.7) were not really different to that for two variables. This can be expected as the dynamics of the system has not changed substantially. It is interesting to note that a move limit of as big as 30 % of the variable's range may be used in Dynamic-Q using forward finite differences. It can also be seen from Figure 4.6 that the optimisation histories are very well behaved. Figure 4.6 again indicates the definite existence of more than one local minimum with the same objective function value, but significantly different design variable values. This is attributed to the 'flat' region in the design space, where the objective function is relatively insensitive to the design variables. The SQP algorithm performed almost similarly to Dynamic-Q, and also found two different local minima, with the same objective function value. SQP converged in 9

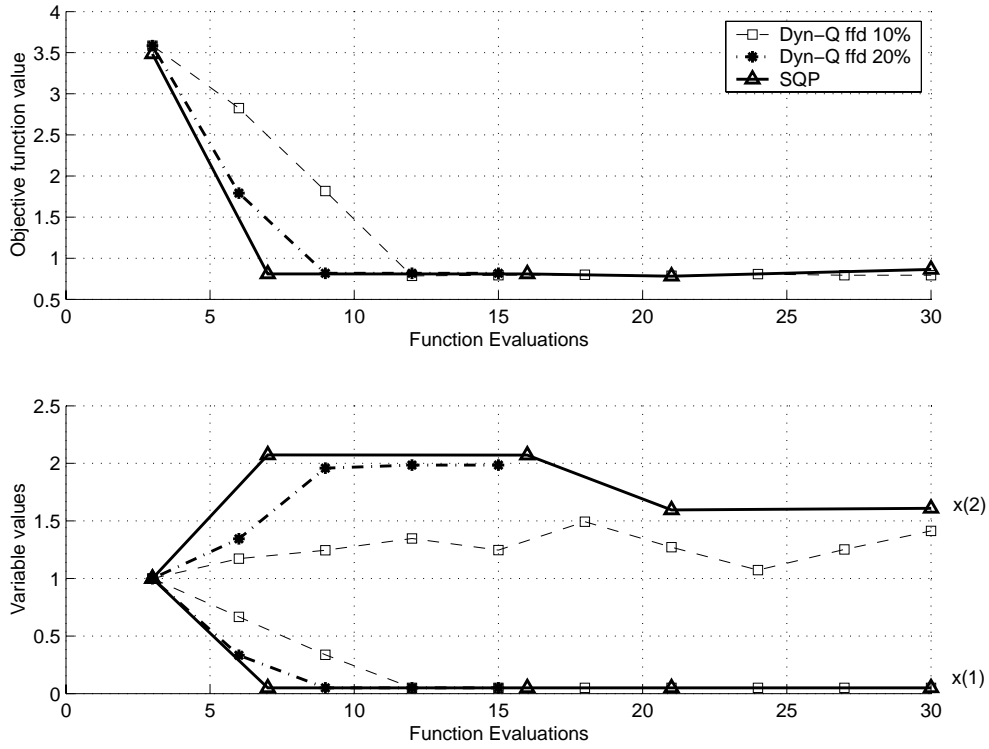


Figure 4.5: Handling optimisation, 2 design variables

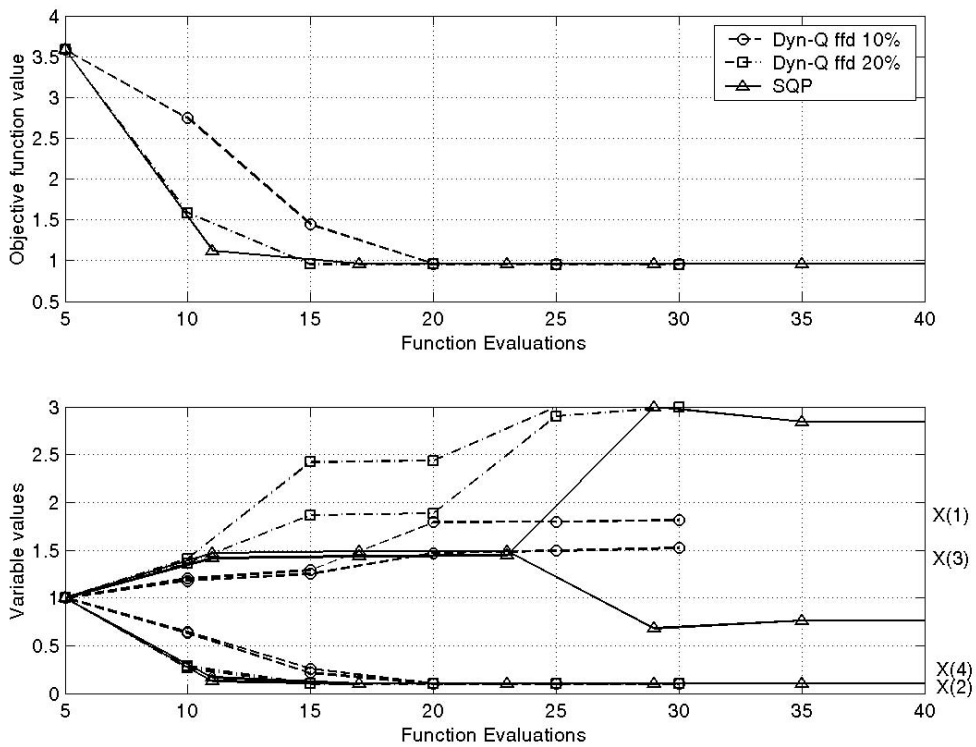


Figure 4.6: Handling optimisation, 4 design variables



iterations (49 function evaluations), and Dynamic-Q in 5 iterations (25 function evaluations). The two different optima correspond to the cases where the damper scale factors are respectively the same and different at the front and rear. Large differences in performance of the optimisation algorithms are only expected when the number of design variables increases. This reinforces our initial conclusion from the results for two variables: that the damper scale factor has negligible effect on the vehicle's handling performance (body roll angle) through the double lane change manoeuvre at the optimum (stiff) spring rate. It can also be observed from the results that the spring gas volume (two design variable optimisation, variable  $x_1$ , four design variable optimisation, variables  $x_2$  and  $x_4$ ) runs to the lowest bound, corresponding to the maximum possible spring stiffness (smallest possible gas volume).

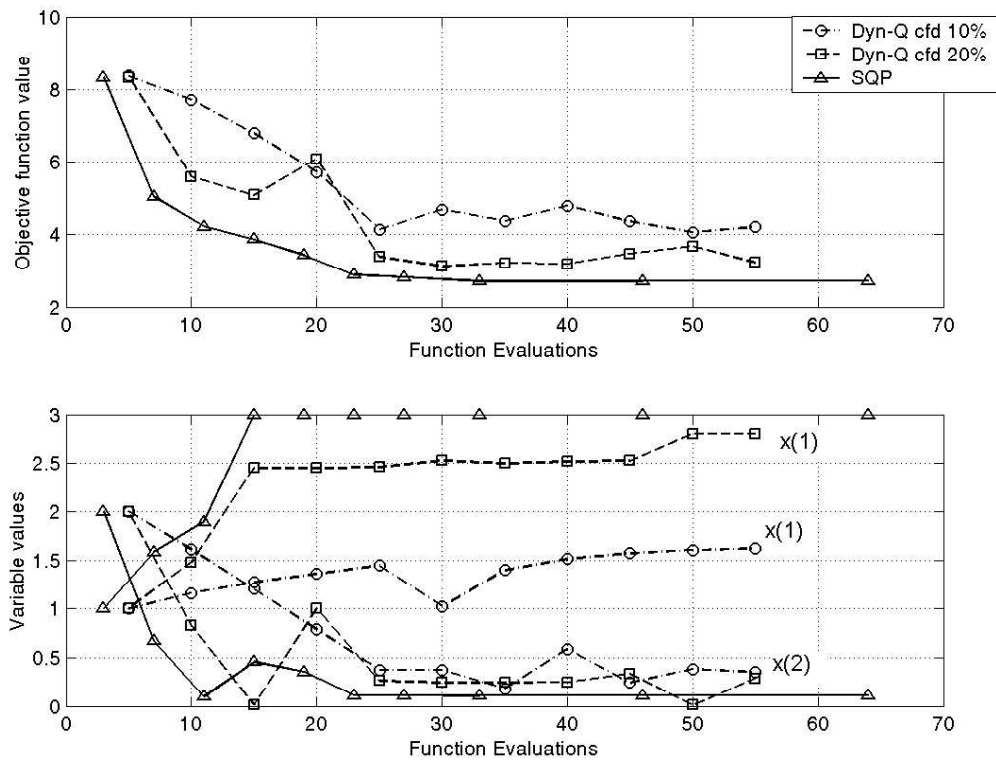
### 4.3.7 Ride Comfort Results

A summary of the results of the optimisation for ride comfort is shown in Table 4.1. Two design variable ride comfort optimisation encountered the problems associated with a noisy objective function. It is postulated that the severe noise present in the ride comfort objective function, as opposed to the smooth nature of the handling objective function, is related to the fact that for ride comfort, accelerations are used, while for handling, angular displacement is used for the objective function measure, however, the filtering and use of the RMS value is a traditional smoothing effect, which is not evident from the results though. Although not immediately apparent at this stage, especially when considering the course mesh of Figure 4.4 used to get a feel of the form of the optimisation problem, it will be shown that the objective function exhibits severe noise.

The SQP method (Figure 4.7) took 8 iterations (33 function evaluations) to stabilise on a minimum, corresponding to the lowest possible damping and stiffness, as expected from Figure 4.4. The Dynamic-Q method experienced greater difficulties in obtaining a stable minimum. For this reason, the



central finite difference method for determining the gradient, was introduced to obtain stability in the optimisation process. The Dynamic-Q method with central finite differences, with a 10 % move limit (Figure 4.7) took 9 iterations (50 function evaluations) to find a minimum, with inspection showing that this minimum is effectively reached after only 4 iterations (25 function evaluations). The vertical acceleration at this point is, however, significantly higher than that found with SQP indicating the existence of a separate interior local minimum. A 20 % move limit (Figure 4.7) took 6 iterations (30 function evaluations), finding a local minimum not far off the SQP minimum. The Dynamic-Q minimum design variable values are not at the extrema found by the SQP method, reinforcing the fact that the ride comfort design space has a flat plateau of local minima.



**Figure 4.7:** Ride comfort optimisation, 2 design variables

For the four design variable optimisation Dynamic-Q was modified so that the move limit for each iteration is 90 % of the move limit of the previous iteration. This was done so as to stabilise the convergence behaviour of the

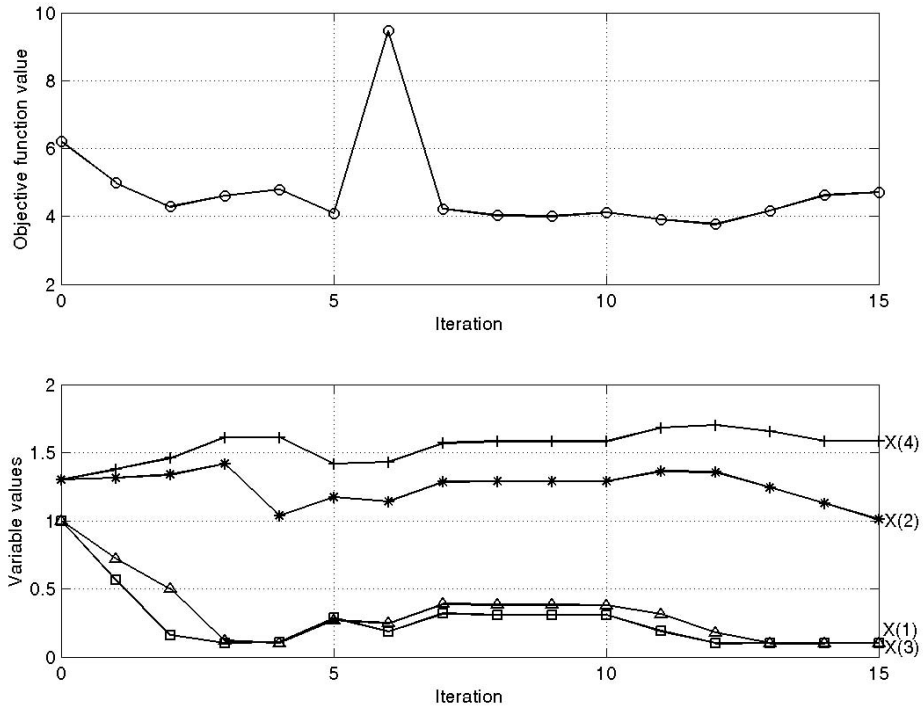


Figure 4.8: Dynamic-Q ffd ride comfort, 4 design variables, 10 % move limit

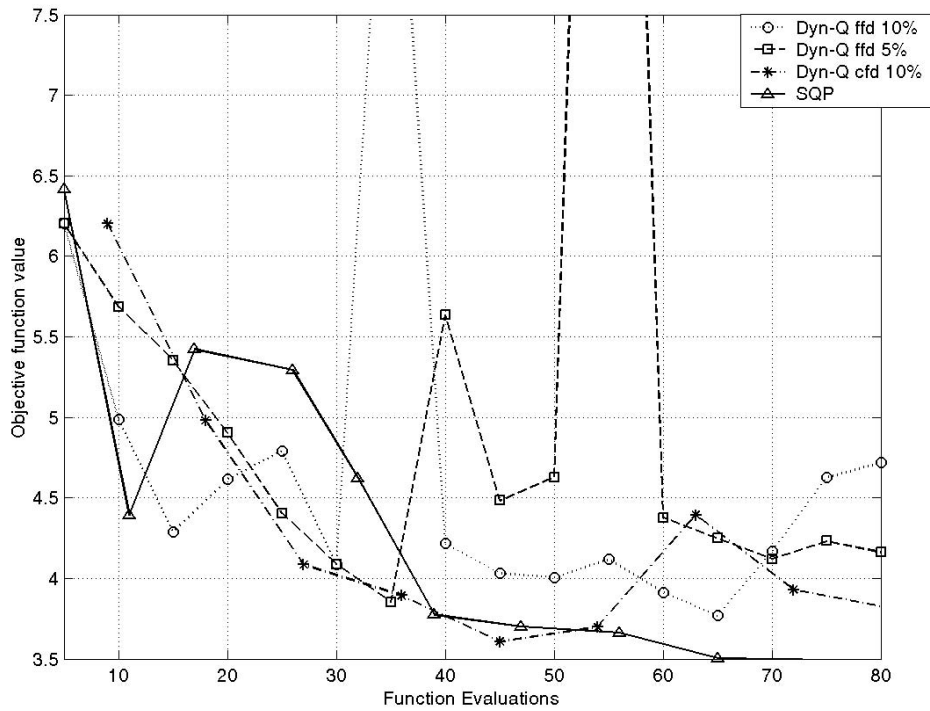


Figure 4.9: Ride comfort optimisation, 4 design variables

**Table 4.1:** Ride Comfort Optimisation Results

Algorithm	Move Limit	Figure Number	Iterations	Function Evaluations	Optimum value
Two Design Variables					
SQP	-	4.7	8	33	2.7
Dynamic-Q central	10 %	4.7	9 (4)	50 (25)	4.1
finite differences	20 %	4.7	6	30	3.1
Four Design Variables					
Dynamic-Q forward	10 %	4.8	12	65	3.8
finite differences	5 %	4.9	6	35	3.9
Dynamic-Q cfd	10 %	4.9	4	45	3.6
SQP	-	4.9	8	65	3.5

algorithm and to try and prevent high spikes in the optimisation process. These spikes are caused by a poor approximation to the objective function close to the minimum, resulting in the LFOPC algorithm finding a minimum of the approximated problem on the slope of the steep valley close to the actual minimum. However, Dynamic-Q quickly recovers within a single iteration (5 function evaluations) as can be seen in Figure 4.8 iteration 7. The results of the optimisation are presented for both central finite differences and forward finite differences used for the gradient approximations in Figure 4.9.

An alternate explanation for the spiky nature of the optimisation convergence histories is the numerical noise of the objective function. This becomes apparent when evaluating Figure 4.7 where it can be observed that for a relatively small change in the design variable values (Dyn-Q cfd 10 %) there is a relatively large change in the objective function value. This is also observed in Figure 4.8 iterations 5 and 6.

From Figures 4.8 and 4.9 for the forward finite difference Dynamic-Q implementation, it can be seen that the smaller move limit of 5 % is more stable reaching a minimum within 6 iterations (35 function evaluations),

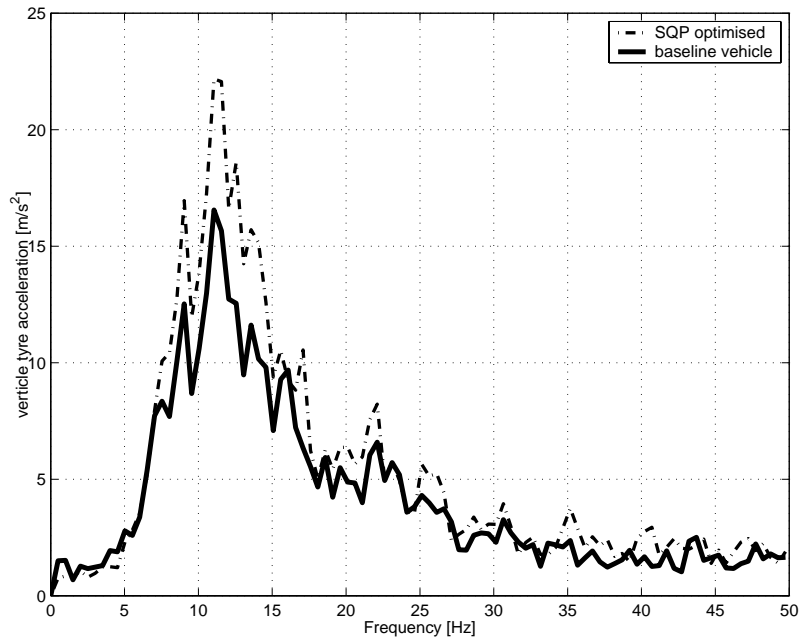


while 10 % move limit takes 12 iterations (65 function evaluations). The algorithm however does not converge due to the noisy objective function with steep valley. The convergence behaviour for central finite differences coupled to Dynamic-Q is shown in Figure 4.9 requiring 4 iterations (45 function evaluations). Again it has been determined that the smaller move limit is beneficial to finding the minimum. The central finite difference gradient evaluation builds into the system a level of robustness. From the results it can be seen that around 1.5 liter gas volume and limited damping returns the best results. The central finite difference results show that by increasing the rear gas volume with minimal damping, a better overall ride can be achieved (Figure 4.9).

SQP also found similar good results within 8 iterations (65 function evaluations) (Figure 4.9). From Table 1 it is concluded that Dynamic-Q with forward finite differences does not reach the same minimum as Dynamic-Q with central finite differences. Dynamic-Q with central finite differences is also comparatively economical to SQP, finding an minimum within 5 % of the SQP minimum objective function value.

The levels of tyre vertical acceleration associated with the obtained optimum design conditions were evaluated. It was found that compared to the baseline vehicle (Figure 4.10), the tyre does experience high levels of acceleration, which is associated with tyre hop, when driving in a straight line over the Belgian paving. However, these accelerations are not transmitted to the vehicle body, so that the objective function value calculated is indeed an optimal value. Due to the presence of tyre hop, it is suggested that even when considering ride comfort and handling separately, a measure of the tyre hop (vertical force, deflection or acceleration) should also be considered when optimising the vehicle's suspension for ride comfort, this is later included in the work performed in Chapter 5.





**Figure 4.10:** Tyre hop investigation. Vertical tyre accelerations for SQP optimised suspension compared to baseline vehicle.

## 4.4 Conclusion

The feasibility of using gradient-based approximation methods for the optimal design of a vehicle's suspension was investigated. An industry-standard version of the SQP method, and the in-house Dynamic-Q method, were evaluated. The determination of the objective function was performed using a full multi-body vehicle simulation model that was both computationally expensive, and exhibited severe inherent numerical noise when considering, in particular, ride comfort. The goal was to determine the vehicle's optimal spring and damper characteristics for both ride comfort and handling.

It is concluded that both optimisation algorithms work exceptionally well when optimising for handling. However, it is felt that the damper should play a role in the optimum suspension settings and thus a damper dependent measure should also be considered in the definition of the handling objective function.





Although difficulties were experienced in the ride comfort optimisation due to the severe inherent noise in the objective function, both algorithms sufficiently overcame this problem by yielding locally optimal feasible solutions. It was found that for the Dynamic-Q algorithm the use of central finite differences for the gradient approximations, at the cost of  $2n + 1$  function evaluations per iteration, achieved meaningful optima at a lower cost in terms of total number of function evaluations (simulations), than the SQP method. This can be attributed to the inherent stability that the central finite differencing technique introduces by considering information ahead and behind the current iteration point. Dynamic-Q can thus be strongly recommended for applications in vehicle suspension optimisation. The increase in tyre hop over the baseline vehicle suggests that tyre hop should be added as a constraint when optimising ride comfort, and cannot be neglected, as previously postulated.

The gradient-based approximation methods considered here prove to be feasible optimisation methods when noisy objective functions are to be optimised. These methods have the distinct advantage of requiring relatively few function evaluations, each of which corresponds to an expensive numerical simulation, before reaching an optimal design. It is concluded from this work that other means of eliminating the negative effects of the numerical noise should be investigated. Also a means of decreasing the total number of expensive numerical simulations even further, should be investigated as this method will still be prohibitively expensive when more design variables are considered. It is also concluded that a move limit of 10 % is a good general value to be used in Dynamic-Q. The Dynamic-Q algorithm will now be used further as it is an in-house code making access to the source code easy, for the implementation of the ideas of the rest of this research.