CONVERGENT VARIANTS OF THE NELDER-MEAD ALGORITHM

A thesis
submitted in partial fulfilment
of the requirements for the
Degree of
Master of Science in Mathematics
at the
University of Canterbury
by

David Byatt

University of Canterbury
2000

Abstract

The Nelder-Mead algorithm for unconstrained optimisation has been used extensively to solve parameter estimation and other problems since its inception in 1965. Despite its age it is still the method of choice for many practitioners in the fields of statistics, engineering and the physical and medical sciences because it is easy to code and very easy to use. It belongs to the *direct search* class of methods which do not require derivatives and which are often claimed to be robust for problems with discontinuities or where the function values are affected by noise.

Recently (1998), M^c^Kinnon has shown that that the method can converge to non-solutions for certain classes of problems. Only very limited convergence results exist (Lagarias *et al*) for a restricted class of problems in one or two dimensions.

An overview of selected direct search methods for unconstrained optimisation is given and several variants of the Nelder-Mead algorithm are presented, culminating in a variant which is provably convergent in any number of dimensions, and which performs well in practice.

Contents

1 Introduction **1**

1.1 Historical development 2

1.2 Desirable features 3

1.3 Why choose the Nelder-Mead algorithm? 4

1.3.1 The presence of noise 6

1.4 Thesis overview 6

2 Methods of optimisation **9**

2.1 Direct search methods 9

2.1.1 Early methods 9

2.1.2 Alternating variable search 11

2.1.3 Hooke and Jeeves 12

2.1.4 Conjugate directions 15

2.1.5 Simplex based methods 15

2.2 Gradient methods 19

2.3 Stopping criteria 20

2.4 Rating algorithm performance 21

3 The Nelder-Mead algorithm **23**

3.1 Introduction 23

3.1.1 Outline of the Nelder-Mead algorithm 24

3.2 The original Nelder-Mead algorithm 28

3.2.1 Choice of parameters 31

3.2.2 Stopping criterion 31

3.3 The standard Nelder-Mead algorithm 32
3.3.1 Choice of parameters 33
3.3.2 Stopping criteria 33
3.4 Problems with Nelder-Mead 34
3.4.1 Choice of initial simplex 34
3.4.2 M^cKinnon's examples 35
3.4.3 The standard quadratic 39
3.4.4 No convergence proof 40

4 Frame based convergence **43**
4.1 Positive bases 44
4.2 Frames 45
4.3 A convergent frame based template 47
4.3.1 The framework 47
4.3.2 The algorithm 47
4.4 Convergence analysis 48

5 Variants of Nelder-Mead **53**
5.1 Failures of the Nelder-Mead algorithm 53
5.1.1 Collapse of the simplex 53
5.1.2 Failure to make sufficient progress 54
5.2 Requirements for a successful variant 54
5.2.1 Preventing simplex collapse 55
5.2.2 Sufficient descent condition 59
5.3 Ranking the performance of the variants 60
5.4 Variant one 61
5.4.1 What if there is insufficient descent? 61
5.4.2 Algorithm outline for variant one 65
5.4.3 Convergence of variant one 65
5.4.4 Performance of variant one 66
5.5 Variant two 67
5.5.1 Algorithm outline for variant two 69

5.5.2 Convergence of variant two . . . 70
5.5.3 Performance of variant two . . . 70
5.6 Variant three . . . 71
5.6.1 Performance of variant three . . . 72
5.7 Variant four . . . 73
5.7.1 Performance of variant four . . . 76
5.7.2 Parameter selection for variant four . . . 77
5.7.3 Numerical results . . . 79

6 Summary **83**
6.1 Further work . . . 84
6.2 Acknowledgements . . . 85

A The suite of test functions **87**

B Numerical results **91**
B.1 Results for the suite of test functions . . . 92
B.1.1 Low tolerance results . . . 92
B.1.2 High tolerance results . . . 93
B.2 Results for M^{c}Kinnon's example . . . 94
B.2.1 Low tolerance results . . . 94
B.2.2 High tolerance results . . . 94

C The standard quadratic **95**

D Schematic diagrams **101**
D.1 Variant one . . . 102
D.2 Variant two . . . 103
D.3 Variant three . . . 104
D.4 Variant four . . . 105

E Performance of the variants **107**
E.1 Variant one . . . 108
E.1.1 Low tolerance results . . . 108

E.1.2 High tolerance results 110
E.2 Variant two 112
E.2.1 Low tolerance results 112
E.2.2 High tolerance results 114
E.3 Variant three 116
E.3.1 Low tolerance results 116
E.3.2 High tolerance results 118
E.4 Variant four 120
E.4.1 Low tolerance results 120
E.4.2 High tolerance results 122

F Parameter summary for variant four **125**
F.1 Using the default parameters 126
F.1.1 Determinant parameter — δ 126
F.1.2 Frame reduction parameter — κ 127
F.1.3 Sufficient descent reduction parameter — ν 128
F.2 Using the new parameter information 129
F.2.1 Determinant parameter — δ 129
F.3 The most successful variants 131

G Choice of parameters **133**
G.1 Determinant parameter (δ) using ψ_3, κ_2, ν_1 134
G.1.1 Low tolerance results 134
G.1.2 High tolerance results 140
G.2 Frame size reduction parameter (κ) using ψ_3, δ_1, ν_1 146
G.2.1 Low tolerance results 146
G.2.2 High tolerance results 148
G.3 Frame size reduction parameter (κ) using ψ_3, δ_5, ν_1 150
G.3.1 Low tolerance results 150
G.3.2 High tolerance results 152
G.4 Epsilon reduction parameter (ν) using $\psi_3, \delta_1, \kappa_2$ 154
G.4.1 Low tolerance results 154

G.4.2 High tolerance results 160
G.5 Epsilon reduction parameter (ν) using $\psi_3, \delta_5, \kappa_1$ 166
G.5.1 Low tolerance results 166
G.5.2 High tolerance results 172
G.6 Determinant parameter (δ) using ψ_3, κ_1, ν_8 178
G.6.1 Low tolerance results 178
G.6.2 High tolerance results 184
G.7 Determinant parameter (δ) using ψ_3, κ_2, ν_8 190
G.7.1 Low tolerance results 190
G.7.2 High tolerance results 196

H MATLAB code **203**
H.1 `FMINSEARCH` 204
H.2 `NM4_`$\psi_3\delta_6\kappa_1\nu_8$ 211
H.2.1 `SIMPLEX` 220
H.2.2 `FRAME` 223
H.3 `HH` 225

Bibliography **227**

List of Figures

2.1 Spherical contours and the alternating variable method 11
2.2 Elongated contours and the alternating variable method . . . 12
2.3 Elongated contours and the method of Hooke and Jeeves . . . 14
2.4 The simplex method of Spendley, Hext and Himsworth. 17

3.1 Nelder-Mead reflect step 25
3.2 Nelder-Mead expand step 25
3.3 Nelder-Mead contract-outside step 26
3.4 Nelder-Mead contract-inside step 26
3.5 Nelder-Mead shrink step 27
3.6 All Nelder-Mead trial points 27
3.7 Simplex rotated by an angle ω. 35
3.8 Simplex collapse for the McKinnon example 36
3.9 Contour plot for the McKinnon example 38

5.1 Simplex side vectors 55
5.2 Reflect step from the ghost simplex 62
5.3 Expand step from the ghost simplex 63
5.4 An alternate search direction 63
5.5 All of the trial points for variant one. 64
5.6 Completing the frame 69
5.7 Failure of variant two on McKinnon's function 74
5.8 Close-up of frame for McKinnon's function 74
5.9 Simplex near collapse 75

C.1 Low tolerance quadratic minima with FMINSEARCH 96
C.2 Low tolerance function evaluations with FMINSEARCH 96
C.3 Low tolerance quadratic minima with NM4_$\psi_3\delta_6\kappa_1\nu_8$ 97
C.4 Low tolerance function evaluations with NM4_$\psi_3\delta_6\kappa_1\nu_8$ 97
C.5 High tolerance quadratic minima with FMINSEARCH 98
C.6 High tolerance function evaluations with FMINSEARCH 98
C.7 High tolerance quadratic minima with NM4_$\psi_3\delta_6\kappa_1\nu_8$ 99
C.8 High tolerance function evaluations with NM4_$\psi_3\delta_6\kappa_1\nu_8$ 99

List of Tables

Notation xv

5.1 Parameter options for variant four. 78
5.2 High tolerance performance of NM4_$\psi_3\delta_1\kappa_6\nu_8$ 80

A.1 The suite of test functions 89

B.1 Low tolerance performance of NM4_$\psi_3\delta_1\kappa_6\nu_8$ 92
B.2 High tolerance performance of NM4_$\psi_3\delta_1\kappa_6\nu_8$ 93
B.3 Low tolerance performance on M^{c}Kinnon's example 94
B.4 High tolerance performance on M^{c}Kinnon's example 94

E.1 Low tolerance results for variant one 109
E.2 High tolerance results for variant one 111
E.3 Low tolerance results for variant two 113
E.4 High tolerance results for variant two 115
E.5 Low tolerance results for variant three 117
E.6 High tolerance results for variant three 119
E.7 Low tolerance results for variant four 121
E.8 High tolerance results for variant four 123

F.1 Determinant parameter for NM4_$\psi_3\kappa_2\nu_1$. 126
F.2 Frame reduction parameter for NM4_$\psi_3\delta_1\nu_1$. 127
F.3 Frame reduction parameter for NM4_$\psi_3\delta_5\nu_1$. 127
F.4 Sufficient descent reduction parameter for NM4_$\psi_3\delta_1\kappa_2$. 128
F.5 Sufficient descent reduction parameter for NM4_$\psi_3\delta_5\kappa_1$. 128

F.6 Determinant parameter for NM4_$\psi_3\kappa_1\nu_8$. 129
F.7 Determinant parameter for NM4_$\psi_3\kappa_2\nu_8$. 130
F.8 Results of the most successful variants 131

G.1 Low tolerance results for NM4_ψ_3 δ_1–δ_3. 135
G.2 Low tolerance results for NM4_ψ_3 δ_4–δ_6. 137
G.3 Low tolerance results for NM4_ψ_3 δ_7–δ_9. 139
G.4 High tolerance results for NM4_ψ_3 δ_1–δ_3. 141
G.5 High tolerance results for NM4_ψ_3 δ_4–δ_6. 143
G.6 High tolerance results for NM4_ψ_3 δ_7–δ_9. 145
G.7 Low tolerance results for NM4_$\psi_3\delta_1$ κ_1–κ_3. 147
G.8 High tolerance results for NM4_$\psi_3\delta_1$ κ_1–κ_3. 149
G.9 Low tolerance results for NM4_$\psi_3\delta_5$ κ_1–κ_3. 151
G.10 High tolerance results for NM4_$\psi_3\delta_5$ κ_1–κ_3. 153
G.11 Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_1–ν_3. 155
G.12 Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_4–ν_6. 157
G.13 Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_7–ν_9. 159
G.14 High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_1–ν_3. 161
G.15 High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_4–ν_6. 163
G.16 High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_6–ν_9. 165
G.17 Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_1–ν_3. 167
G.18 Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_4–ν_6. 169
G.19 Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_7–ν_9. 171
G.20 High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_1–ν_3. 173
G.21 High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_4–ν_6. 175
G.22 High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_7–ν_9. 177
G.23 Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_1–δ_3. 179
G.24 Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_4–δ_6. 181
G.25 Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_7–δ_9. 183
G.26 High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_1–δ_3. 185
G.27 High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_4–δ_6. 187
G.28 High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_7–δ_9. 189

G.29 Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_1–δ_3. 191
G.30 Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_4–δ_6. 193
G.31 Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_7–δ_9. 195
G.32 High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_1–δ_3. 197
G.33 High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_4–δ_6. 199
G.34 High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_7–δ_9. 201

Notation

$\mathbf{b}_i$	The i^{th} base point
d_i	A small perturbation of the i^{th} independent variable x_i
$\mathrm{diam}(V)$	Diameter of the simplex V
E_i	Error in the Taylor series for f in the direction of the i^{th} positive basis vector
f	The objective function $f : \mathbf{R}^n \to \mathbf{R}$
$f(\mathbf{x})$	The value of the objective function at the point $\mathbf{x}$
$\bar{f}$	The average function value over the simplex
f_{tol}	Convergence criteria for the difference in function values
F	A frame in $\mathbf{R}^n$
$F(\mathbf{x}_0, P_+, h)$	A particular frame in $\mathbf{R}^n$
$\mathcal{F}$	A compact subset of $\mathbf{R}^n$
FE	The number of function evaluations
$\bar{\mathbf{g}}$	Centroid of the n best vertices of the ghost simplex
h	Frame size parameter
H	Worst vertex of the ghost simplex
k	Iteration counter
K	Upper bound for the length of positive basis vectors
L	Simplex size reduction parameter
M_i	Maximum absolute difference in the gradient of f along a line in the direction of the i^{th} positive basis vector
Mag	Difference in magnitude of numerical results
n	The number of independent variables

N	Constant associated with the sufficient descent parameter
$\mathbf{N}$	The natural numbers, $\mathbf{N} = \{1, 2, 3, \ldots\}$
`NM1-4`	Variants 1–4 of the Nelder-Mead algorithm
$\mathbf{p}_i$	Elements of a positive basis
$\mathbf{p}_i^{(\infty)}$	The limit of a sequence of positive basis vectors
P_+	A set of vectors which form a positive basis
$P_+^{(\infty)}$	The limit of a sequence of positive bases
Q	An upper bound for the number of random trial points
$\mathbf{R}$	The real numbers
$\mathbf{R}^n$	Euclidean n-space
$\mathbf{s}_i$	The i^{th} side vector of a simplex
s_i	The length of $\mathbf{s}_i$
s_{tol}	Convergence criteria based on step lengths
$\mathbf{u}$	A vector in $\mathbf{R}^n$
$\mathbf{u}^{(i)}$	New vertices introduced in M^{c}Kinnon's example
U	An open subset of $\mathbf{R}^n$
$\mathbf{v}$	A vector in $\mathbf{R}^n$
$\mathbf{v}_i$	The i^{th} vertex of a simplex
$\hat{\mathbf{v}}_i$	Position of the i^{th} vertex after a shrink step
$\bar{\mathbf{v}}$	The centroid of the n best vertices of a simplex
V	A simplex in $\mathbf{R}^n$
$\mathrm{vol}(V)$	The volume of the simplex V
x_i	The i^{th} independent variable
$\mathbf{x}$	A vector in $\mathbf{R}^n$
$\mathbf{x}^*$	A (local) minimiser of the objective function
$\mathbf{x}_0$	Central frame point
$\mathbf{x}_c$	Nelder-Mead contract-outside point
$\mathbf{x}_{cc}$	Nelder-Mead contract-inside point
$\mathbf{x}_e$	Nelder-Mead expansion point
$\mathbf{x}_p$	Pseudo expansion point
$\mathbf{x}_r$	Nelder-Mead reflection point

x_{tol}	Convergence criterion for the distance between successive iterates
$\mathbf{z}$	A vector in $\mathbf{R}^n$
α	Nelder-Mead reflection coefficient
β	Nelder-Mead contraction coefficient
γ	Nelder-Mead expansion coefficient
δ	Lower bound used for determinants
ϵ	Sufficient descent parameter
η	Positive constants used with the positive basis vectors
θ	Parameter used for M^{c}Kinnon's functions
κ	Reduction parameter for h
(λ_1, λ_2)	Coordinates of a vertex in M^{c}Kinnon's initial simplex
μ	Simplex volume scale-factor
ν	Reduction parameter for ϵ
σ	Nelder-Mead shrink coefficient
τ	Parameter used for M^{c}Kinnon's functions
Υ	The termination tolerance for the original Nelder-Mead algorithm
ϕ	Parameter used for M^{c}Kinnon's functions
ψ	Method of simplex reshape
ω	Simplex rotation angle
Ω	A subsequence of central frame points
$\square$	End of theorem marker
$\blacksquare$	End of proof marker
$[\mathbf{u}, \mathbf{v}]$	The distance from $\mathbf{u}$ to $\mathbf{v}$
$\lceil x \rceil$	The least integer greater than or equal to x
$\star$	Maximum number of function evaluations (100000) reached before the stopping criteria were met
$\dagger$	Algorithm failed to produce the correct solution
$\clubsuit$	Algorithm failed to produce the correct solution for at least one function in the test suite

Note on notation

Sometimes a superscript is used to refer to a particular iterate or iteration. For example, $X^{(k)}$ refers to the k^{th} iterate of X. Similarly, $X^{(0)}$ is the initial, original, or starting X. If no subscript is used then X refers to the current value of X, $X^{(+)}$ is the next X and $X^{(-)}$ is the previous X.

Chapter 1

Introduction

Optimisation may be described as the science of determining the "best" solutions to certain mathematically defined problems. These problems are often models of physical situations and so there is an extremely diverse range of practical applications. In fact the applicability of optimisation methods is so widespread it reaches into almost every activity in which numerical information is processed; science, engineering, mathematics, statistics, economics and commerce. In the preface to his 1987 book [11, p. ix] Fletcher writes

> The subject of optimization is a fascinating blend of heuristics and rigour, of theory and experiment. It can be studied as a branch of pure mathematics, and yet has applications in almost every branch of science and technology.

But what exactly is optimisation? The word optimal comes from the Latin word *optimus*, which means best. So optimisation involves finding an optimal, or best solution to a problem [12, p. ix]. The measure of "goodness" of the alternative choices is described by an *objective function* whose value depends on a set of independent variables or parameters $x_1, x_2, \ldots, x_n$ [1, p. 1]. Mathematically, optimisation means finding a minimum (or maximum) of a real-valued function of n variables. For convenience these variables can be represented by the vector $\mathbf{x}$ so that

$$f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n) \tag{1.1}$$

Often there are restrictions or *constraints* that define acceptable values for the variables [13, p. 1]. If there are no restrictions on the values that the variables can hold then the optimisation process is called *unconstrained optimisation.* This is the only type of optimisation that will be considered throughout the remainder of this thesis.

For completeness, formal definitions for the terms; *minimum*, *minimiser* and *optimisation*, are given below.

Definition 1.1 (Minimum) *Given a real-valued function* $f : \mathbf{R}^n \to \mathbf{R}$ *we say that* $f(\mathbf{x}^*)$ *is a (local) minimum of* f *if there is an open region* $U \subseteq \mathbf{R}^n$ *containing a point* $\mathbf{x}^*$ *in its interior such that*

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in U \tag{1.2}$$

Definition 1.2 (Minimiser) *A point* $\mathbf{x}^* \in \mathbf{R}^n$ *for which equation* (1.2) *holds is called a (local) minimiser of* f.

Definition 1.3 (Optimisation) *Optimisation is the process of finding a minimum and/or a minimiser for the objective function* f.

Since finding the maximiser of some function f is equivalent to finding the minimiser of $-f$, it is common practice for the words optimise and minimise to be synonymous.

In general it is only practicable to find local rather than global solutions to minimisation problems [11, p. 12] and, in practice, some conditions may need to be imposed, such as the continuity of the objective function, in order to guarantee the existence of a minimum.

1.1 Historical development

The existence of optimisation problems is probably as old as mathematics. The first systematic techniques for the solution of optimisation problems stem from the development of calculus and are associated with the names of

Newton, Lagrange, and Cauchy (who made the first application of a method of steepest descent in 1847). However before 1940 relatively little was known about methods for numerical optimisation of functions of several variables. Some "least squares" calculations had been carried out and "steepest decent" type methods had been applied to some physics problems. More sophisticated versions of Newton's method in many variables were also attempted. However any complex problem required armies of assistants operating desk calculating machines.

There is no doubt that the advent of the computer was paramount in the development of optimisation methods and numerical analysis [11, p. 3]. At around the same time demand for solutions to various decision problems increased dramatically, due in part, to World War II [19, p. 5]. More recently much of the stimulus for modern advances has come from large scale industries. In such expensive fields as the space and aeronautics industries, even relatively small savings of only one or two percent can give enormous cash savings. Considerations of this sort have led to considerable effort and energy being put into the related fields of optimisation and control [5, p. 1].

1.2 Desirable features

If an optimisation algorithm is to be useful it must meet some standard of performance. Typically the behaviour of an optimisation algorithm could be described as acceptable if successive approximations $\mathbf{x}^{(k)}$ move steadily towards the function minimiser $\mathbf{x}^*$ and then converge rapidly to the point $\mathbf{x}^*$ itself. Ideally the convergence should be quadratic, or a least super-linear near the solution. Some algorithms have other features, such as attempting to reduce the value of $f(\mathbf{x}^{(k)})$ at each iteration [10, p. 16], [11, p. 20].

In general terms, an optimisation algorithm is said to perform well if it produces an output quickly and *cheaply*, and the output produced is an accurate approximation to the solution. The *cost* of using a certain algorithm is often determined by the number of function evaluations required to produce the output. These ideas are expanded upon in later sections, particularly section 2.4.

As the main focus of this thesis is the development of a provably convergent variant of the Nelder-Mead algorithm, the following section discusses some of the reasons for choosing the Nelder-Mead algorithm to begin with.

1.3 Why choose the Nelder-Mead algorithm?

The simplex method of Nelder and Mead (1965) for unconstrained optimisation belongs to the *direct search* class of methods which do not require derivatives and are often claimed to be robust for problems where there are discontinuities or where the function values are affected by noise.

The initial popularity of the Nelder-Mead algorithm began to decline in favour of methods which use derivative information to locate the minima of functions. Some authors recommended that methods which did not make use of derivative information be avoided. For example, Gill *et al* [13, p. 93] write

> A method using function comparison should only be used if there is no other suitable method available. If a user decides to use a function comparison method only because of its simplicity and seeming generality, he may pay a severe price in speed and reliability. ... The substantial disadvantage of function comparison methods is that few (if any) guarantees can be made concerning convergence.

Many modern texts on optimisation only mention in passing the existence and use of derivative free methods — some authors ignoring this type of optimisation method altogether in favour of derivative based methods. Recently however, derivative free methods have become fashionable again [4, p. 14], [6], [29].

The main motivation for the development and use of derivative free methods is that there is a large number of practical problems where derivative information just is not available. If the function values are the result of some physical, chemical or econometrical measurement then they will be subject to noise and so derivative information may range from unreliable right through to totally unusable. The nature of the problems may mean that finite difference approximations to the derivatives are not feasible, either because the objective function is expensive to compute (either in real cost — as some expensive physical process, or in the time that such a computation requires), or because of the physical nature of the problem itself. For example, it may not be practical to alter the temperature in the chamber where a chemical reaction is taking place by one part in 10^6 and then to measure the effect that this has on the objective function, or, there may be a lag in the time from when a change is made to the result of the change. In such situations, calculating many finite difference approximations becomes impractical [6, p. 84]. The occurrence of problems of this nature is common in the industrial world.

The Nelder-Mead algorithm generally performs well, even for functions of high dimensionality or in the presence of noise [17, p. 63], [26]. Several authors suggest that this method becomes less competitive as the dimensionality of the objective function is increased, but Kowalik and Osborne [19, p. xi] make the following observation

> In our numerical experiments we have found the simplex method in particular to be surprisingly successful. It seems to us that appropriate implementation can offset to a certain extent at least the alleged decrease in efficiency of these methods as the dimensionality of the problem is increased. There seems to be room for much research and experiment here.

Despite its age, the Nelder-Mead simplex algorithm for unconstrained optimisation is still the method of choice for many practitioners across a wide range of fields.

1.3.1 The presence of noise

Most optimisation software is not designed to solve problems in which the computation of the function values are subject to noise. It is usually assumed that the objective function can be evaluated on a computer to full machine precision. A problem with the optimisation of noisy functions is that if the algorithm requires the estimation of derivatives by differences and if the difference parameter does not depend on the level of noise then incorrect derivative approximations are usually obtained. This invariably leads to the failure of the algorithm. Another possible problem is that the termination criteria of the algorithm must recognise when differences in the objective function values are only due to noise [22, p. 340]. Since the Nelder-Mead algorithm only ranks function values its performance should be relatively unaffected by noise provided the noise level does not alter the ranking of the function values.

1.4 Thesis overview

Despite its generally good performance, there are some "nice" functions for which the Nelder-Mead algorithm performs poorly. The main focus of this thesis is the development of a provably convergent variant of the Nelder-Mead algorithm for unconstrained optimisation that is both useful in practice and guaranteed to converge to a stationary point under mild conditions on the objective function.

Chapter 2 gives a general overview of some optimisation methods and introduces the Nelder-Mead algorithm. A more detailed discussion of the Nelder-Mead algorithm, along with some of the problems from which it suffers is given in chapter 3. The convergence framework used by the variants of the Nelder-Mead algorithm is covered in chapter 4. The variants themselves are introduced and discussed in chapter 5. Extensive tables of results are listed in the appendices. These will be of little interest to most readers! They are included for completeness, and because the data were used to fine-tune the performance of the successful Nelder-Mead variant. A more user-friendly summary of this data is included in appendix F.

When the ideas being discussed are illustrated graphically, two dimensional examples are used. This is mainly to avoid the complications of having to draw higher dimensional spaces, but also because, in every case the two dimensional examples provide all the necessary information, without the clutter.

This project is based heavily on experimentation and extensive use has been made of the computer software package MATLAB[1]. Due to its widespread use and availability, MATLAB's own implementation of the Nelder-Mead algorithm, `FMINSEARCH` was used as a control when comparing the performance of the Nelder-Mead variants on a suite of test functions [23]. These functions are listed in appendix A on page 87. All computations were carried out on a Sun Enterprise 450 machine with MATLAB R11.1.

[1]MATLAB is a registered trademark of The MathWorks, Inc.

Chapter 2

Methods of optimisation

2.1 Direct search methods

Definition 2.1 (Direct search methods) *Methods which use comparisons of the values of the objective function and do not require the evaluation of any derivatives are called direct search methods* [2, p. 7], [13, p. 94].

2.1.1 Early methods

The early methods of optimisation did not have strong theoretical backgrounds. They generally attempted to enclose the region in which the minimum existed by bisecting each of the variables in turn. These methods then used some systematic way of contracting the region to locate the minimum more accurately. Unfortunately the amount of effort required to implement these methods goes up rapidly (typically as 2^n) which caused their authors to coin the phrase *the curse of dimensionality* [11, p. 17].

Exhaustive search

This is the simplest (in concept) of the direct search methods. With this method the objective function is evaluated at every point on a grid (or node of a lattice) generated over the region of interest. Although simple to use, it is hard to over-estimate the inefficiency of this procedure. For example if the objective function depended on 10 variables and each dimension of the region of interest was divided into a modest 20 sub-intervals then evaluating the function at all of the grid points would require 10^{20} function evaluations. Even on a super-computer capable of performing 10^9 function evaluations per second it is likely that any research grant would have expired long before the task was completed. Despite this, the exhaustive search method has been used in practice, usually where the number of variables has not been too large and the function to be optimised has been *intractable* mathematically [9, p. 154].

Random search

Another early method of optimisation was to guess solutions at random in some fixed region. The final "solution" was then taken to be the guess which gave the best results over some large number of trials [32]. A probabilistic version of this simple method was proposed by Brookes [3] in 1958, in which the objective function is evaluated at Q uniformly distributed random points in some region. For this simple random search method a relationship exists between the number of points and the probability that the error in function value is less than an arbitrary given amount.

The type of methods that use either an exhaustive search or repeated random guesses are often referred to as *brute force* methods. These methods are used as the basis for some global optimisation techniques.

2.1.2 Alternating variable search

In this method each of the independent variables is considered in turn while the $n-1$ remaining variables are kept fixed. Each variable is altered until a minimum of the objective function is located. The current best point moves parallel to each of the co-ordinate axes in turn, moving in a new direction when a minimum in the current direction of search is reached.

As shown in figure 2.1, if the contours of the objective function are hyperspherical then this method will locate the minimum of a function of n independent variables in at most n searches, regardless of the initial point chosen. However, if the contours of the objective function are elongated in some direction then the direction of the elongation (or *major axis*) may not be parallel to any of the co-ordinate axes. In this situation the alternating variable search may only be able to take very small steps at each iteration and the search will become extremely inefficient. An example of this is shown

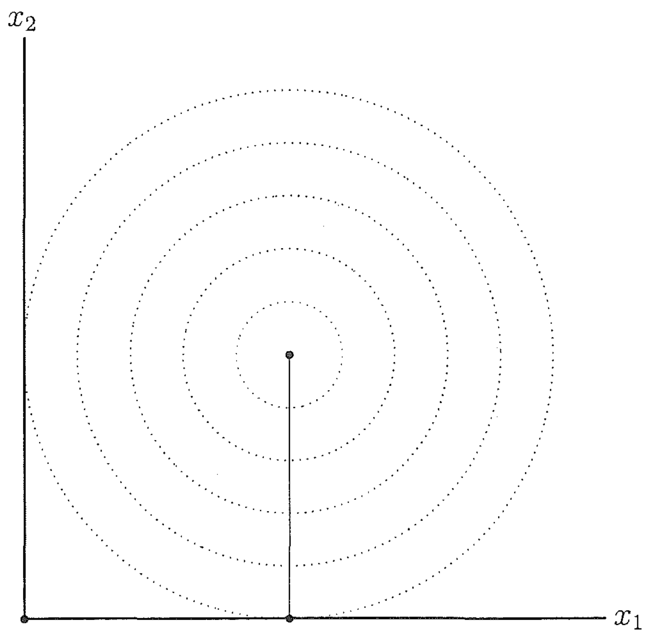


Figure 2.1: The alternating variable method and a function with hyperspherical contours.

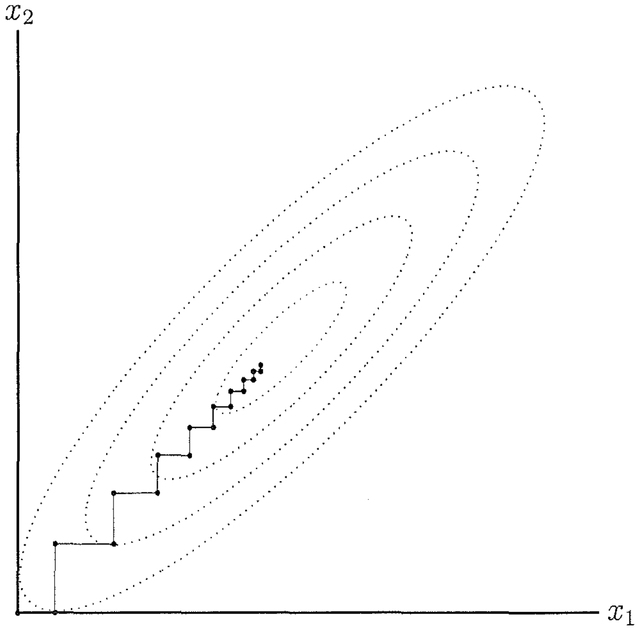


Figure 2.2: The alternating variable method and a function with an elongated major axis at an angle to the co-ordinate axes.

in figure 2.2. This inefficiency becomes greater as the number of variables is increased [2, p. 25].

Numerical examples have been presented [28] showing that a method which changes one variable at a time may cycle without calculating any point where $f(\mathbf{x})$ has been reduced. Such algorithms can fail if the search directions are always identical and chosen in constant order, and the objective function is non-uniformly bounded and strictly convex [31].

2.1.3 Hooke and Jeeves

The method of Hooke and Jeeves [16] is another direct search method. It is similar to the alternating variable search except that it attempts to align the search direction to the major axis of the objective function. The method, described below, alternately uses two types of moves, called *exploratory* and *pattern* moves.

Initialise: The initial point forms what is called the first *base point* $\mathbf{b}_1$. This is also the *current point.*

Exploratory moves: Each of the n variables is considered in turn with x_i being perturbed by some small amount d_i to obtain a *trial point.* If the function value at the trial point is lower than the function value at the current point then the trial point becomes the new current point and this procedure is repeated on the next independent variable.

If the function value at the trial point is not lower than the function value at the current point then the trial point is rejected. The sign of d_i is changed and a trial point is tried in the opposite direction. Again, only if the function value at this trial point is lower than the function value at the current point is the trial point accepted as the new current point. In either case the procedure is continued on the next independent variable.

The exploratory moves have been completed once all of the n variables have been considered in turn. The point obtained after the completion of all exploratory moves becomes the new base point $\mathbf{b}_2$.

Pattern move: The first set of exploratory moves results in the base point moving from $\mathbf{b}_1$ to $\mathbf{b}_2$. The idea of Hooke and Jeeves is to see if any further progress can be made in the direction $\mathbf{b}_2 - \mathbf{b}_1$ in the hope that this direction approximates the major axis of the objective function. A pattern move is to repeat the move made between the most recent pair of base points. The pattern move moves the current point from the base point $\mathbf{b}_2$ to the point $2\mathbf{b}_2 - \mathbf{b}_1$.

From the current point the exploratory moves are repeated for each of the n variables to get a new base point $\mathbf{b}_3$. If $f(\mathbf{b}_3) \leq f(\mathbf{b}_2)$ then the pattern move plus exploratory moves are considered a success and a new pattern move is made from the base point $\mathbf{b}_3$ to $2\mathbf{b}_3 - \mathbf{b}_2$ and the entire process is repeated. If $f(\mathbf{b}_3) > f(\mathbf{b}_2)$ then the pattern move plus exploratory moves are considered a failure and the sequence of exploratory moves is repeated from the base point $\mathbf{b}_2$.

These steps are repeated until all exploratory moves from a base point fail. When this occurs either the minimum has been located to the accuracy of the step sizes d_i, or the search has led the last base point into a steep, skew valley that cannot be negotiated using the current step sizes. In either case the step sizes d_i are reduced and the whole procedure is repeated from the last base point.

The stopping criterion for this method is reached when each of the step lengths have been reduced below some pre-determined minimum step size [2, pp. 25–26].

An illustration of the steps made by the method of Hooke and Jeeves is shown in figure 2.3. The exploratory moves are shown as solid lines, the pattern moves are dashed lines.

As shown in [7], the method of Hooke and Jeeves is provably convergent.

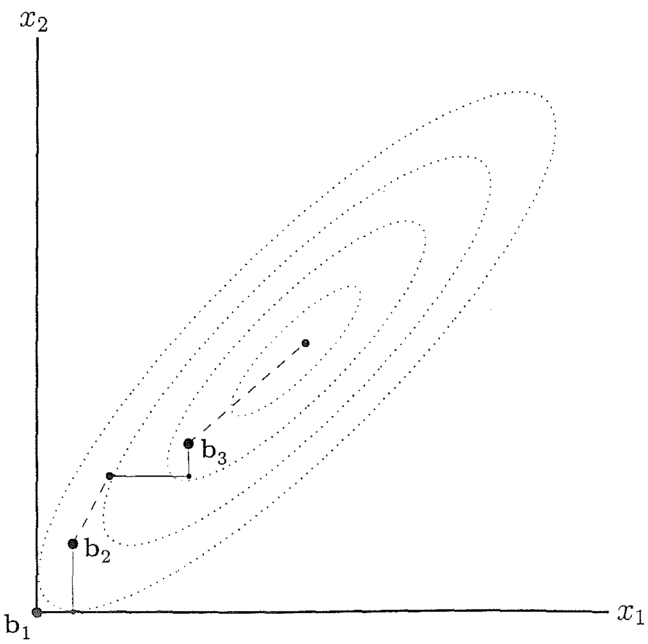


Figure 2.3: The Hooke and Jeeves method and a function with an elongated major axis at an angle to the co-ordinate axes.

2.1.4 Conjugate directions

Conjugate directions methods are a class of algorithm that assumes the continuity of all second derivatives. Since the other methods described in this thesis are not so restrictive, the conjugate directions class will not be discussed further. A full description of this type of method can be found in most modern texts on optimisation. See, for example, Fletcher [11], or Nocedal and Wright [25].

2.1.5 Simplex based methods

Definition 2.2 (Simplex) *A simplex in* $\mathbf{R}^n$ *is a set of* $n+1$ *points which do not lie in a hyperplane.*

If the vertices of a simplex are all mutually equidistant then the simplex is said to be *regular*. For example, in two dimensions, a regular simplex is an equilateral triangle, while in three dimensions a regular simplex is a regular tetrahedron.

Each iteration of a simplex based direct search method begins with a simplex specified by its $n+1$ vertices and their associated function values. One or more test points are computed along with their corresponding function values. The iteration terminates with a new (and different) simplex such that the function values at the vertices of the new simplex satisfy some form of descent condition when compared to the previous simplex [20, pp. 112–113].

Simplex methods are the most successful direct search methods [11, p. 17]. The descriptions of the simplex methods presented in this thesis make use of the following definitions.

Definition 2.3 (Best point) *The best point of a simplex is the vertex with the lowest function value.*

Definition 2.4 (Worst point) *Similarly, the worst point of a simplex is the vertex with the highest function value.*

Both of these definitions can be generalised slightly so that they apply to any finite set of points in $\mathbf{R}^n$. Throughout the remainder of this thesis, pairs of points are often compared in this way. For example, if $f(\mathbf{v}_i) \leq f(\mathbf{v}_j)$ then $\mathbf{v}_i$ is *better* than $\mathbf{v}_j$, or alternately, $\mathbf{v}_j$ is *worse* than $\mathbf{v}_i$.

Spendley, Hext and Himsworth

The simplex method of Spendley, Hext and Himsworth [30] uses a sequence of regular simplices to produce an approximation to the minimiser of the objective function. The method is started by setting up an initial, regular simplex in $\mathbf{R}^n$ and evaluating the objective function at each of its $n + 1$ vertices. A description of the method follows.

1. Reflect the vertex with the highest function value in the centroid of the n remaining vertices to form a new (also regular) simplex.

2. Evaluate the objective function at this new vertex and proceed from step 1.

   If the new vertex happens to be the vertex with the greatest function value in the new simplex then the above procedure will cease to make progress and merely oscillate between the last two simplices. To prevent this happening a new rule is introduced.

3. If at any stage the vertex with the highest function value selected by step 1 is the most recently introduced vertex of the current simplex, then reflect the vertex with the second highest function value in the centroid of the n remaining vertices instead.

   If the same vertex is in a large number of consecutive simplices then another rule is introduced.

4. If one vertex remains unchanged for more than L consecutive iterations then the simplex size is reduced by halving the distances of the remaining vertices from the vertex that has remained fixed. The entire search procedure is then repeated [2, pp. 20–21].

In general the value of L will depend on the number of variables. Spendley, Hext and Himsworth suggest that for best results L be set to

$$L = 1.65n + 0.05n^2 \tag{2.1}$$

A typical performance of the method for a function of two variables is illustrated in figure 2.4. The initial simplex is labelled a–b–c and the numbers indicate the order in which the new vertices were introduced. It will be noticed that step 3 has been used in the simplices made up of the vertices 8–9–10, 8–10–11 and 8–11–12. Since vertex 8 is close to the minimiser, each new simplex merely revolves about this point. The coincidence of vertices 6 and 13 is a peculiarity of regular simplices in $\mathbf{R}^2$ (equilateral triangles) and does not happen in higher dimensions.

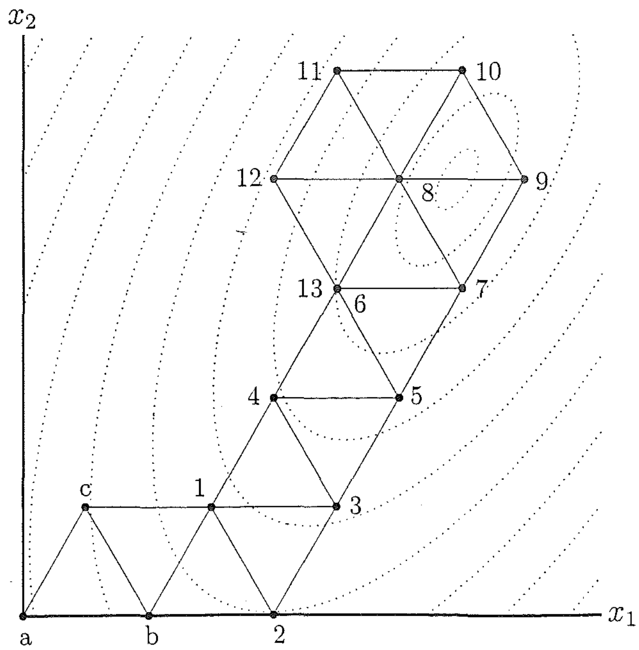


Figure 2.4: The simplex method of Spendley, Hext and Himsworth.

Introduction to the simplex method of Nelder and Mead

The method of Nelder and Mead [24] is another simplex based method which is descended from the method of Spendley, Hext and Himsworth. Nelder and Mead's method does not require successive simplices to be regular and so it offers greater flexibility than the simplex method of Spendley, Hext and Himsworth. With each iteration of the Nelder-Mead method the simplex shape can "evolve" so that successive simplices can adapt to the local contours of the objective function. Nelder and Mead [24, p. 311] make the comment

> Our method is highly opportunist, in that the least possible information is used at each stage and no account is kept of past positions. No assumptions are made about the surface except that it is continuous and has a unique minimum in the area of the search.

Function values are only used to rank the vertices of each simplex. The algorithm attempts to create a new simplex by replacing the worst vertex in the current simplex by a better point. A major disadvantage of Spendley, Hext and Himsworth's method is that the use of regular simplices limits the movement of the vertex with the lowest function value at each iteration. Once a favourable direction has been found it would be more efficient to move further in that direction than the regular simplex pattern allows. This obstacle can be overcome by permitting the use of non-regular simplices [34, p. 81]. Since the shape of each simplex in the Nelder-Mead algorithm does not remain fixed, successive simplices can alter to better fit the local nature of the objective function and contract to the minimiser [17, p. 14], [24]. A detailed description of this method is given in chapter 3.

2.2 Gradient methods

Definition 2.5 (Gradient methods) *Methods which make use of the derivatives of the objective function with respect to the independent variables, as well as the values of the objective function are called gradient methods* [2, p. 7].

In general, gradient methods perform better than any method which does not make use of derivative information. This is to be expected as the more information available about the objective function, the better the approximation to its minima. However the calculation of derivatives is restricted to certain well-behaved functions, and even then the algebraic problems that arise from calculating derivatives are frequently extremely formidable. Consequently the use of derivatives is not a sufficiently powerful tool to handle all realistic optimisation problems of practical importance [14, p. 4]. The use of exact derivative information can be by-passed by making finite difference approximations. However these approximations are sensitive to noise and round-off errors. As such, most convergence proofs which require exact derivatives are no longer valid when finite difference approximations are used.

In order to keep the limitations on the objective function as general as possible the availability of derivatives will not be assumed. Therefore depending on the objective function, gradient methods, for example; steepest descent, Newton's method, quasi-Newton methods and conjugate gradients, may not be applicable. As most modern texts on optimisation give full descriptions of these and other gradient methods (see, for example, Fletcher [11] or Nocedal and Wright [25]) they will not be discussed further.

2.3 Stopping criteria

For each of the above iterative methods, deciding whether the sequences

$$\{\mathbf{x}^{(k)}\}, \quad \{f(\mathbf{x}^{(k)})\} \qquad k = 1, 2, 3, \ldots \tag{2.2}$$

have converged to the minimiser $\mathbf{x}^*$ and minimum $f(\mathbf{x}^*)$ respectively, is very difficult [2, p. 8]. For many of the methods of optimisation is it difficult to base such deductions on the progress made during a set of iterations, since this progress can be exceedingly erratic [2, p. 8]. Typically there are two tests which determine the stopping criteria for an optimisation algorithm. The first is that the change in function values from one iteration to the next is sufficiently small

$$\left| \, f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k-1)}) \, \right| < f_{tol} \tag{2.3}$$

and the second is that successive estimates of the minimiser $\mathbf{x}^*$ must be sufficiently close

$$\left\| \, \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \, \right\| < x_{tol} \tag{2.4}$$

For simplex based algorithms equation (2.4) is replaced by a condition on the size, or *diameter*, of the current simplex, which is described below.

Definition 2.6 (Simplex diameter) *The diameter of a simplex is the maximum distance between any two vertices of the simplex.*

If V is a simplex in $\mathbf{R}^n$ and

$$S = \{1, 2, \ldots, n+1\} \tag{2.5}$$

then

$$\operatorname{diam}(V) = \max_{i,j \in S}\{ \, \| \, \mathbf{v}_i - \mathbf{v}_j \, \| \, \} \tag{2.6}$$

For simplex based methods equation (2.4) is replaced by

$$\operatorname{diam}(V) < x_{tol} \tag{2.7}$$

which requires the diameter of the simplex to be sufficiently small before the algorithm can terminate.

Note that these criteria could be satisfied away from the minimiser (especially near the vicinity of a stationary point), resulting in the termination of the algorithm before an accurate approximation to the function minimiser (or minimum) has been achieved [2, p. 9].

2.4 Rating algorithm performance

The practical usefulness of an optimisation algorithm can often be determined by the algorithm's performance on a selection of test problems. For a test problem to be useful it is best that the problem have a single minimum or at least a restricted number of minima. An algorithm cannot be expected to find the global minimum if a problem has more than one (local) minimum. However an algorithm should reach a local minimum to be considered successful, or at the very least, a stationary point. Locating the global minimum of a function is a separate problem, outside the scope of this thesis. Global optimisation methods frequently use local optimisation algorithms as subroutines.

It is generally accepted that the primary criterion in evaluating general purpose optimisation algorithms is whether or not the algorithm can solve most of the problems posed. That is, whether or not the algorithm is *robust.*

It is worth mentioning that solutions to test problems are known in advance, whereas it is unusual for the solution to a "real-life" problem to be known beforehand. Since good performance by an algorithm on a suite of test functions does not guarantee good performance on any particular real problem, theoretical convergence results may be necessary to gauge the practical usefulness of an algorithm. However convergence and rate of convergence results alone are not a guarantee of good performance when algorithms are implemented on a computer. Computer round-off error may make a crucial difference to the behaviour of the algorithm in practice, and so, the development of a successful optimisation algorithm often relies on experimentation. A successful algorithm must have acceptable behaviour on a variety of test functions. These test functions should be chosen to represent the different features or obstacles which arise in general (as much as this is possible and/or practicable anyway). Such experimentation can never guarantee that an algorithm will always perform well. However, in practice, a well chosen suite of test functions can give a reliable indication of the performance of an algorithm. The ideal situation is for an algorithm to have convergence and rate of convergence proofs backed up by a good selection of experimental test results [11, p. 20].

Often convergence guarantees come at some price, usually restrictions on the objective function which may not be easy to verify. In some cases (for example if the objective function is required to be strictly convex) these conditions may not be satisfied in practice.

Common criteria for measuring the effectiveness of an algorithm on test problems are the number of function evaluations required to reach the desired stopping criteria and the total computation time required for the algorithm to terminate [15, pp. 70–73].

Chapter 3

The Nelder-Mead algorithm

3.1 Introduction

The simplex method of Nelder and Mead [24] is a generalisation of the simplex method of Spendley, Hext and Himsworth.

Some authors have stated that Nelder and Mead's original paper is ambiguous, for example, Lagarias *et al* [20, p. 115] write

> The 1965 paper [24] contains several ambiguities about strictness of inequalities and tie-breaking that have led to differences in interpretation of the Nelder-Mead algorithm.

While it is true that the text description of the algorithm in Nelder and Mead's original paper does contain ambiguities the authors included a flowchart for the algorithm which removed any ambiguity that may have existed in the text.

The algorithm described in the paper by Nelder and Mead using the interpretation given by the flowchart [24, p. 309] will be referred to as the *original* Nelder-Mead algorithm. The *standard* Nelder-Mead algorithm will refer to the implementation by Lagarias *et al* as this differs from the original Nelder-Mead algorithm and it is used by MATLAB in the optimisation routine `FMINSEARCH`. Any reference to the Nelder-Mead algorithm will be valid for both the original and the standard implementations, both of which are described in the following sections.

3.1.1 Outline of the Nelder-Mead algorithm

There are four basic operations in the Nelder-Mead simplex algorithm. These are:

(a) Reflection,

(b) Expansion,

(c) Contraction:

  (i) outside,

  (ii) inside,

(d) Shrink.

The algorithm uses steps (a)–(c) to create a new simplex by attempting to replace the vertex with the highest function value with a better one. If the attempt to find a better point is unsuccessful then the current simplex is reduced in size using step (d), and the entire procedure is repeated. Each of the steps (a)–(d) are described below and illustrated in figures 3.1–3.5 where the new simplex is shown with a solid outline and the original simplex is shown with a dashed outline.

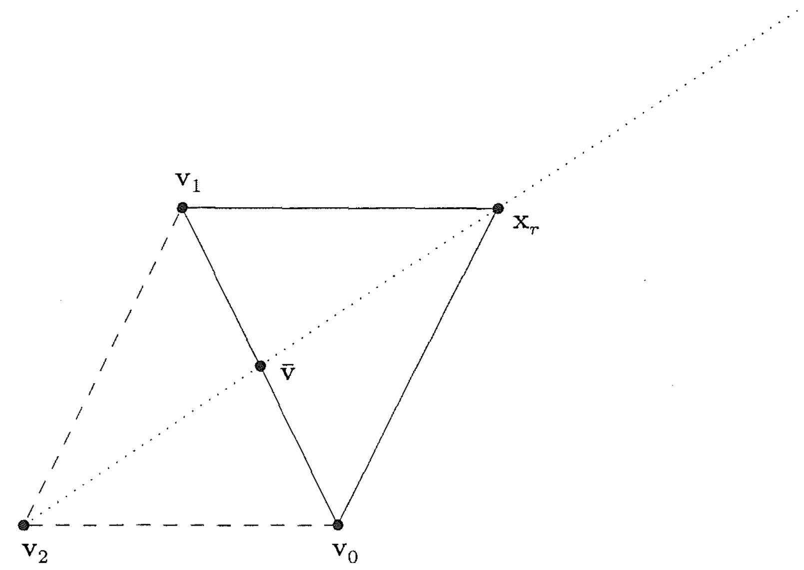


Figure 3.1: The reflection step for the Nelder-Mead algorithm.

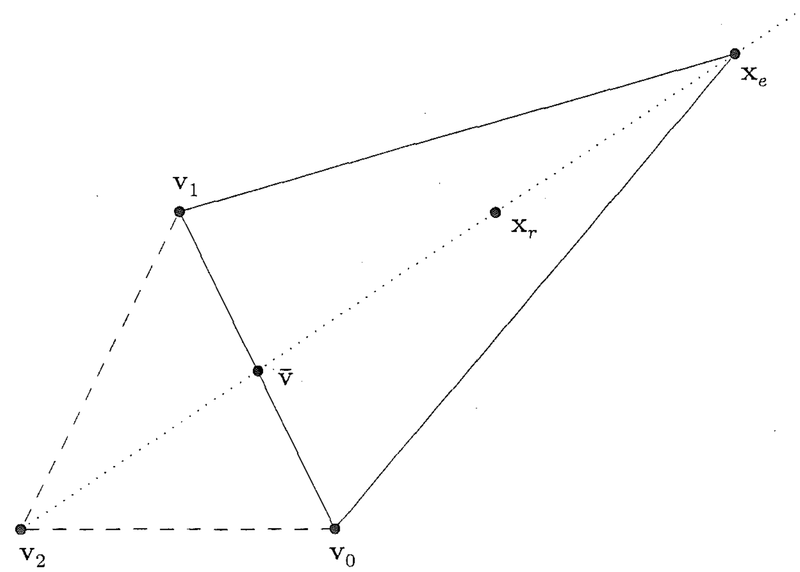


Figure 3.2: The expansion step for the Nelder-Mead algorithm.

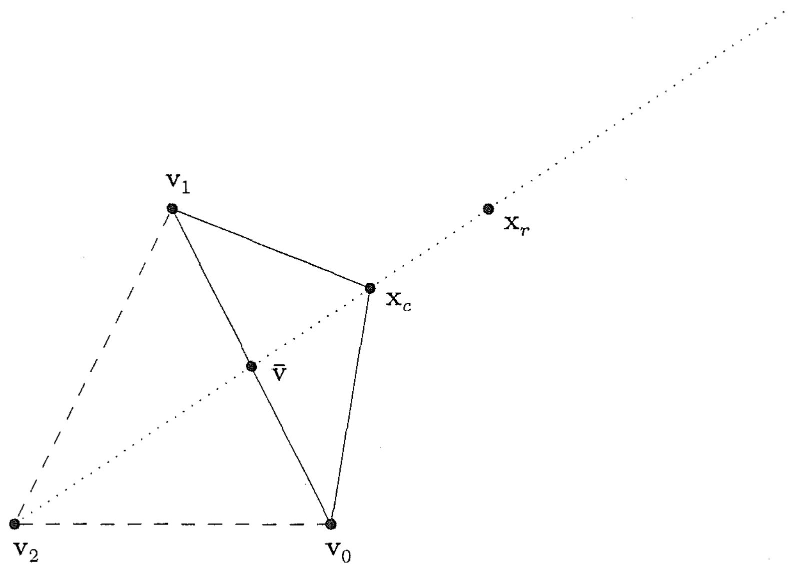


Figure 3.3: The contraction outside step for the Nelder-Mead algorithm.

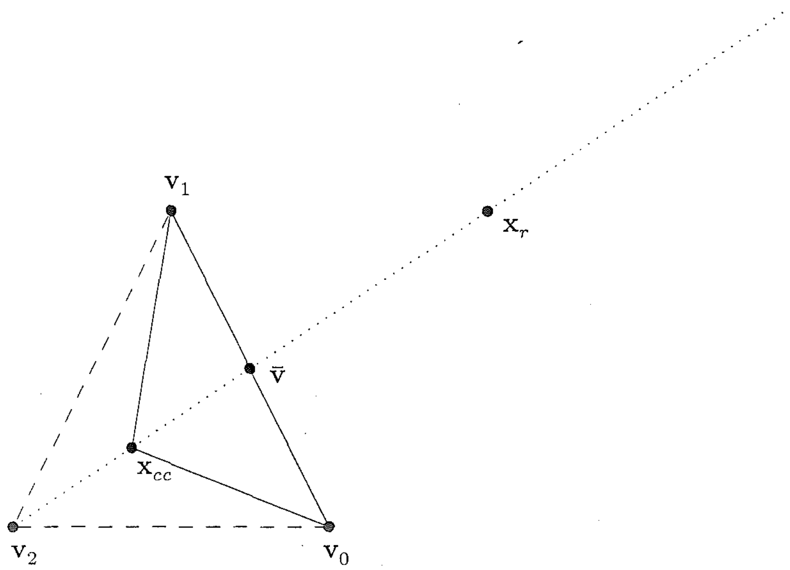


Figure 3.4: The contraction inside step for the Nelder-Mead algorithm.

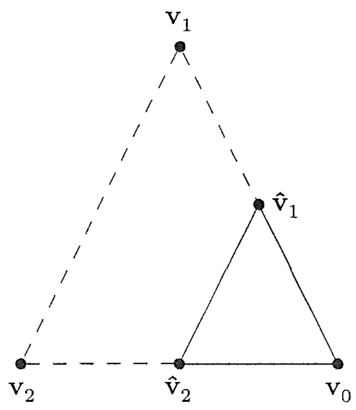


Figure 3.5: The shrink step for the Nelder-Mead algorithm.

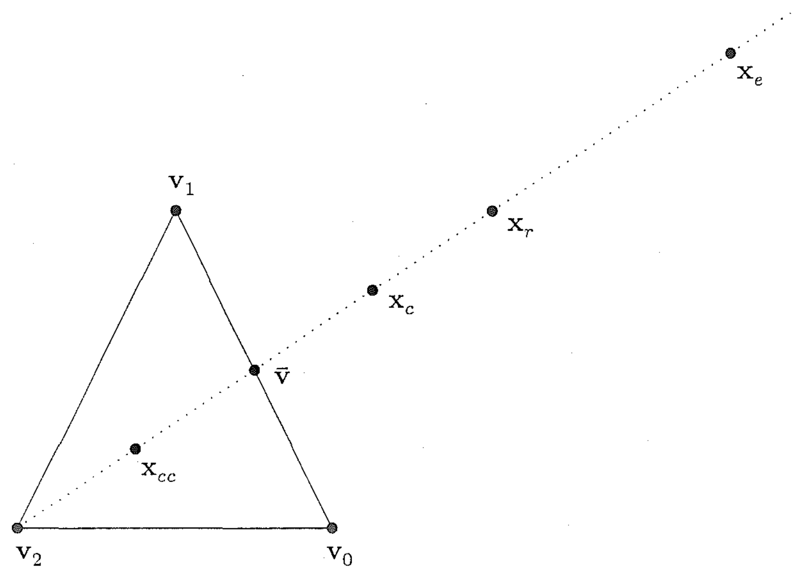


Figure 3.6: All of the trial points used by the Nelder-Mead algorithm.

3.2 The original Nelder-Mead algorithm

The Nelder-Mead algorithm initially proceeds like the simplex method of Spendley, Hext and Himsworth. The method is started by setting up an initial simplex in $\mathbf{R}^n$. Then the objective function is evaluated at each of the $n+1$ vertices of the simplex. The vertices are ordered $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n$ so that

$$f(\mathbf{v}_0) \leq f(\mathbf{v}_1) \leq \cdots \leq f(\mathbf{v}_n) \tag{3.1}$$

In the unlikely event of equality some tie-breaking rules will be required. No specific tie-breaking rules are given in Nelder and Mead's original paper [24], however, `FMINSEARCH` uses MATLAB's own `SORT` function to order the function values.

Reflection: The vertex with the highest function value $\mathbf{v}_n$ is reflected in the centroid $\bar{\mathbf{v}}$ of the n remaining vertices. Since the simplices no longer have to be regular, the size of the reflection step can be scaled by a positive constant $\alpha > 0$ called the *reflection coefficient.*

The *reflect point* $\mathbf{x}_r$ is given by

$$\begin{aligned} \mathbf{x}_r &= \bar{\mathbf{v}} + \alpha(\bar{\mathbf{v}} - \mathbf{v}_n) \\ &= (1+\alpha)\bar{\mathbf{v}} - \alpha\mathbf{v}_n \end{aligned} \tag{3.2}$$

where

$$\bar{\mathbf{v}} = \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{v}_i \tag{3.3}$$

and so α is equal to the ratio of $[\mathbf{x}_r, \bar{\mathbf{v}}]$ to $[\mathbf{v}_n, \bar{\mathbf{v}}]$, that is

$$\alpha = \frac{[\mathbf{x}_r, \bar{\mathbf{v}}]}{[\mathbf{v}_n, \bar{\mathbf{v}}]} \tag{3.4}$$

where $[\mathbf{u}, \mathbf{v}]$ denotes the distance between the points $\mathbf{u}$ and $\mathbf{v}$.

If $f(\mathbf{x}_r) < f(\mathbf{v}_0)$ then the algorithm attempts a further search in the same direction to see if a further reduction in the function value is possible. This is performed by the expansion step.

If $f(\mathbf{v}_{n-1}) < f(\mathbf{x}_r)$ then if the reflect point was accepted it would become the worst vertex in the new simplex and so the reflect point is rejected and a contraction step is tried.

Otherwise $f(\mathbf{v}_0) \leq f(\mathbf{x}_r) \leq f(\mathbf{v}_{n-1})$ and the reflect point is accepted. The new simplex is formed by replacing the worst vertex of the current simplex $\mathbf{v}_n$ by the reflect point $\mathbf{x}_r$.

Expansion: If $f(\mathbf{x}_r) < f(\mathbf{v}_0)$ then it is worth investigating whether a further step in the same direction would be successful. The *expansion coefficient* $\gamma > 1$ is used to generate the *expand point* $\mathbf{x}_e$ by extending the search in the same direction as the line from $\mathbf{v}_n$ to $\bar{\mathbf{v}}$ by

$$\begin{aligned} \mathbf{x}_e &= \bar{\mathbf{v}} + \gamma(\mathbf{x}_r - \bar{\mathbf{v}}) \\ &= \gamma\mathbf{x}_r + (1-\gamma)\bar{\mathbf{v}} \end{aligned} \tag{3.5}$$

where γ is equal to the ratio of $[\mathbf{x}_e, \bar{\mathbf{v}}]$ to $[\mathbf{x}_r, \bar{\mathbf{v}}]$, that is

$$\gamma = \frac{[\mathbf{x}_e, \bar{\mathbf{v}}]}{[\mathbf{x}_r, \bar{\mathbf{v}}]} \tag{3.6}$$

If $f(\mathbf{x}_e) < f(\mathbf{v}_0)$ then the expand step is accepted. The new simplex is formed by replacing the worst vertex of the simplex $\mathbf{v}_n$ by the expand point $\mathbf{x}_e$.

If $f(\mathbf{x}_e) \geq f(\mathbf{v}_0)$ then the expand point is rejected and the reflect point is accepted.

Contraction: If $f(\mathbf{x}_r) > f(\mathbf{v}_{n-1})$ then the reflect step has not been successful and so a contraction step using the *contraction coefficient* $0 < \beta < 1$ is attempted. The contraction step comes in two types. If $f(\mathbf{x}_r) > f(\mathbf{v}_n)$ then a contract-inside step is attempted, otherwise a contract-outside is attempted.

Contraction inside: The contract-inside point is given by

$$\begin{aligned}\mathbf{x}_{cc} &= \bar{\mathbf{v}} + \beta(\mathbf{v}_n - \bar{\mathbf{v}}) \\ &= (1-\beta)\bar{\mathbf{v}} + \beta\mathbf{v}_n \end{aligned} \tag{3.7}$$

where β is equal to the ratio of $[\mathbf{v}_{cc}, \bar{\mathbf{v}}]$ to $[\mathbf{v}_n, \bar{\mathbf{v}}]$, that is

$$\beta = \frac{[\mathbf{x}_{cc}, \bar{\mathbf{v}}]}{[\mathbf{x}_n, \bar{\mathbf{v}}]} \tag{3.8}$$

The contract-inside step is successful if $f(\mathbf{x}_{cc}) \leq f(\mathbf{v}_n)$

Contraction outside: The contract-outside point is given by

$$\begin{aligned}\mathbf{x}_c &= \bar{\mathbf{v}} + \beta(\mathbf{x}_r - \bar{\mathbf{v}}) \\ &= (1-\beta)\bar{\mathbf{v}} + \beta\mathbf{x}_r \end{aligned} \tag{3.9}$$

where β is equal to the ratio of $[\mathbf{x}_c, \bar{\mathbf{v}}]$ to $[\mathbf{x}_r, \bar{\mathbf{v}}]$, that is

$$\beta = \frac{[\mathbf{x}_c, \bar{\mathbf{v}}]}{[\mathbf{x}_r, \bar{\mathbf{v}}]} \tag{3.10}$$

The contract-outside step is successful if $f(\mathbf{x}_c) \leq f(\mathbf{v}_n)$

If the function value at the contract point is greater than $f(\mathbf{v}_n)$ then the contract step failed and a shrink is performed on the simplex.

Shrink: If none of the above steps have produced a point better than $\mathbf{v}_n$ then the current simplex is shrunk using the *shrink coefficient* $0 < \sigma < 1$. Each vertex $\mathbf{v}_i$ of the simplex is replaced by a new vertex $\hat{\mathbf{v}}_i$ given by

$$\begin{aligned}\hat{\mathbf{v}}_i &= \mathbf{v}_0 + \sigma(\mathbf{v}_i - \mathbf{v}_0) \\ &= (1-\sigma)\mathbf{v}_0 + \sigma\mathbf{v}_i \end{aligned} \tag{3.11}$$

The shrink coefficient represents the ratio of $[\hat{\mathbf{v}}_i, \mathbf{v}_0]$ to $[\mathbf{v}_i, \mathbf{v}_0]$, where $\hat{\mathbf{v}}_i$ is the new position of the i^{th} vertex after the shrink, so that

$$\sigma = \frac{[\hat{\mathbf{v}}_i, \mathbf{v}_0]}{[\mathbf{v}_i, \mathbf{v}_0]} \tag{3.12}$$

The coefficients α, β, γ are the factors by which the volume of the simplex is changed by the reflect, contract and expand steps respectively [24, p. 308]. A shrink step changes the volume of the simplex by σ^n. Although clear geometrically for the two dimensional case, for a proof see [20, p. 120].

3.2.1 Choice of parameters

After some experimentation Nelder and Mead chose the following values for the coefficients; $\alpha = 1, \quad \gamma = 2, \quad \beta = \frac{1}{2}, \quad \sigma = \frac{1}{2}$. These numbers correspond to a symmetric reflect step, an expansion step which is twice the length of the reflect step, and contraction steps which are half the length of the reflect step. Similarly the shrink step halves the distance from the best vertex to each of the other vertices in the simplex.

3.2.2 Stopping criterion

The stopping criterion suggested by Nelder and Mead is to terminate the algorithm whenever

$$\sqrt{\frac{1}{n+1}\sum_{i=0}^{n}\left[f(\mathrm{v}_i) - \bar{f}\right]^2} < \Upsilon \tag{3.13}$$

for some pre-determined termination value Υ, where

$$\bar{f} = \frac{1}{n+1}\sum_{i=0}^{n} f(\mathrm{v}_i) \tag{3.14}$$

is the average function value over the simplex.

The reasoning behind this choice of stopping criterion is that in statistical problems where one is concerned with finding the minimum of a negative likelihood surface (or of a sum-of-squares surface) the curvature near the minimum gives the information available on the unknown parameters. If the curvature is slight then the sampling variance of the estimates will be large and so there is no sense in finding the coordinates of the minimum very accurately. Whereas if the curvature is marked, then there is justification for pinning down the minimum more exactly [24, p. 309].

3.3 The standard Nelder-Mead algorithm

As mentioned above, because of its availability with the widespread use of the software package MATLAB, the standard Nelder-Mead algorithm refers to the implementation used by MATLAB in their algorithm `FMINSEARCH`. This implementation is discussed in the paper by Lagarius *et al* [20].

The overall structure of the standard Nelder-Mead algorithm is the same as the original Nelder-Mead algorithm. The main differences between the algorithms are the criteria for deciding when each of the reflect, expand or contract steps have been successful. These details are described below.

1. Create an initial simplex in $\mathbf{R}^n$. `FMINSEARCH` uses an initial estimate (guess) of the minimiser to produce the initial simplex. The initial estimate is used as one of the vertices of the simplex. The n remaining vertices of the initial simplex are found by perturbing, in turn, each of the coordinates of the initial guess. The perturbation used is:
   (a) If the coordinate is non-zero then set the perturbed coordinate to 105% of its current value,
   (b) If the coordinate is zero then set the perturbed coordinate to 0.00025.
2. Evaluate the objective function at each of the $n + 1$ vertices of the simplex.
3. Order the vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n$ of the current simplex so that
$$f(\mathbf{v}_0) \leq f(\mathbf{v}_1) \leq \cdots \leq f(\mathbf{v}_n) \tag{3.15}$$
4. Calculate the reflect point.
5. Accept the reflect point if $f(\mathbf{v}_0) \leq f(\mathbf{x}_r) < f(\mathbf{v}_{n-1})$

6. If $f(\mathbf{x}_r) < f(\mathbf{v}_0)$, calculate the expansion point.
   (a) Accept the expansion point if $f(\mathbf{x}_e) < f(\mathbf{x}_r)$
   (b) Otherwise accept the reflect point.
7. If $f(\mathbf{x}_r) \geq f(\mathbf{v}_{n-1})$ then perform a contraction.
   (a) If $f(\mathbf{v}_{n-1}) \leq f(\mathbf{x}_r) < f(\mathbf{v}_n)$ then contract outside. Accept the contract outside point if $f(\mathbf{x}_c) \leq f(\mathbf{x}_r)$, otherwise shrink the simplex.
   (b) If $f(\mathbf{v}_n) \leq f(\mathbf{x}_r)$ then contract inside. Accept the contract inside point if $f(\mathbf{x}_{cc}) < f(\mathbf{v}_n)$, otherwise shrink the simplex.

3.3.1 Choice of parameters

A small difference between the two algorithms is that the coefficients for reflection, expansion and contraction, have been renamed from α, γ and β in the original Nelder-Mead algorithm to ρ, χ and γ in the standard Nelder-Mead algorithm. Regardless of the names, the values for the reflection, expansion, contraction and shrink coefficients remain unchanged.

3.3.2 Stopping criteria

The stopping criteria used for the standard Nelder-Mead algorithm are the same as those previously mentioned in section 2.3 with the addition that it is the infinity norm which is used for equation (2.6).

3.4 Problems with Nelder-Mead

3.4.1 Choice of initial simplex

As Nelder and Mead observed in their paper [24]

> A difficulty encountered in testing the procedure was that the size and orientation of the initial simplex had an effect on the speed of convergence ...

Experiments [27, pp. 20–125] show that the success of the Nelder-Mead algorithm has a remarkable dependence on the choice of initial simplex. Interestingly, choosing the initial simplex so that the distance of the vertices from one central point is fairly constant (regardless of the actual position of these vertices) does not have much effect on the performance of the algorithm. However choosing an initial simplex where these side-lengths are varied may increase the number of function evaluations required to reach the same stopping criteria by over 300%. Varying the orientation of an initial simplex may also increase the number of function evaluations required to reach the stopping criteria by over 300%. These variations have been encountered across a wide range of test functions. In some experiments a change in the orientation of the initial simplex by only one degree varied the number of function evaluations required before the stopping criteria were met by ±45%.

Since even small changes in the orientation of an initial simplex can lead to such wide variations in the number of function evaluations required to reach the stopping criteria, it appears that the deliberate choice of an initial orientation holds an element of risk for all but very regular functions. As the choice of orientation parameters widens with an increase in dimension, these difficulties are compounded by raising the dimensionality of the objective function [27, pp. 20–125].

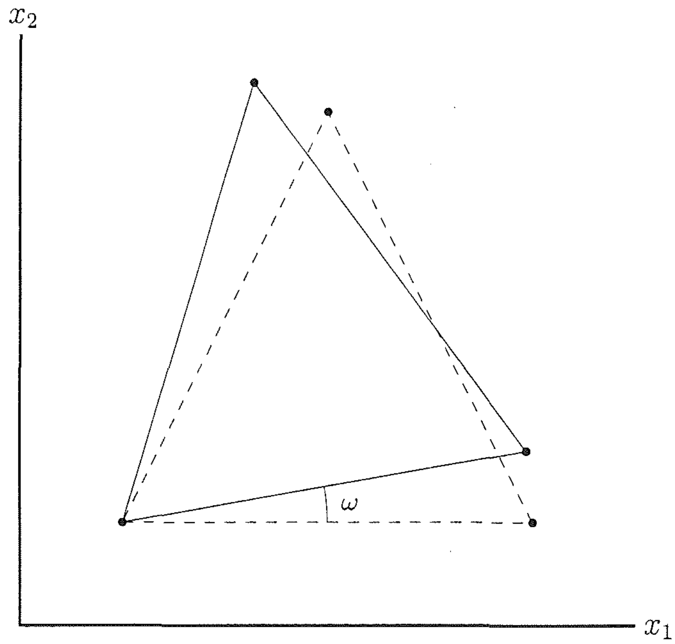


Figure 3.7: Simplex rotated by an angle ω.

3.4.2 M^{c}Kinnon's examples

M^{c}Kinnon [21] has produced a family of functions of two variables which cause the Nelder-Mead algorithm to converge to a non-stationary point. The members of the family are strictly convex with up to three continuous derivatives. With the appropriate initial simplex these functions cause the Nelder-Mead algorithm to perform the inside contraction step repeatedly while the best vertex remains fixed. The simplices tend to collapse to a straight line which is orthogonal to the steepest descent direction at the best vertex. M^{c}Kinnon has shown that this behaviour cannot occur for smoother functions. These examples are the best behaved functions currently known which cause the Nelder-Mead algorithm to converge to a non-stationary point [21, p. 157].

As mentioned in section 3.4.1, the choice of initial simplex can have a major effect of the performance of the Nelder-Mead algorithm. In order for the Nelder-Mead algorithm to fail on the M^{c}Kinnon examples a particular initial simplex (depending on the function) is required. As a result the standard Nelder-Mead algorithm with its built-in choice of initial simplex does not suffer from the convergence to a non-stationary point described in the M^{c}Kinnon paper. Although presumably, it would be possible to scale and rotate the M^{c}Kinnon functions so that this effect could be seen with the standard Nelder-Mead algorithm, with its own choice of initial simplex. However to observe this behaviour it is more convenient to manually set an appropriate initial simplex.

M^{c}Kinnon has shown that if the constants λ_1 and λ_2 are set to

$$\lambda_1 = \frac{1+\sqrt{33}}{8}, \qquad \lambda_2 = \frac{1-\sqrt{33}}{8} \tag{3.16}$$

then for an initial simplex with vertices $(0,0)$, (λ_1, λ_2) and $(1,1)$ each iteration the Nelder-Mead algorithm accepts the contract-inside point to construct the next simplex. The sequence of simplices is shown in figure 3.8. If the new vertex introduced by the k^{th} iteration of the Nelder-Mead algorithm is

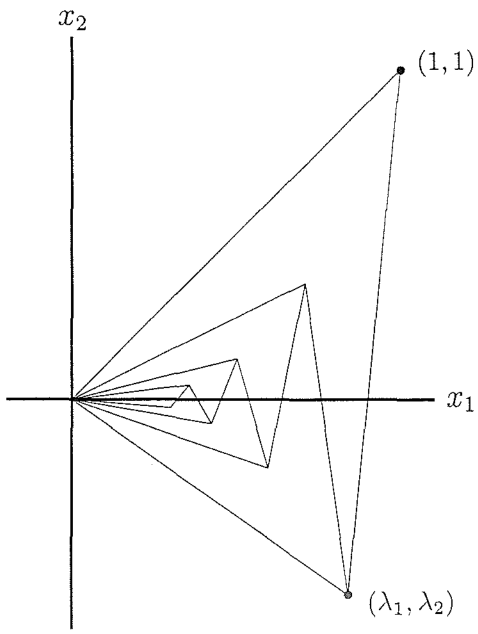


Figure 3.8: The sequence of successive simplices for M^{c}Kinnon's example showing the eventual collapse of the simplex.

labelled $\mathbf{u}^{(k)}$ then it can be shown [21, p. 151] that the coordinates for this vertex are given by

$$\mathbf{u}^{(k)} = (\lambda_1^{k+1}, \lambda_2^{k+1}) \tag{3.17}$$

Hence the vertices for the k^{th} simplex generated by the k^{th} iteration of the Nelder-Mead algorithm are $(0,0), (\lambda_1^k, \lambda_2^k)$ and $(\lambda_1^{k+1}, \lambda_2^{k+1})$. The vertical height of the k^{th} simplex is

$$\begin{aligned} height &= |\lambda_2^{k+1} - \lambda_2^k| \\ &= |\lambda_2^k|(1-\lambda_2) \\ &= \frac{7+\sqrt{33}}{8}|\lambda_2|^k \end{aligned} \tag{3.18}$$

and the horizontal length of the k^{th} simplex is

$$length = \lambda_1^k \tag{3.19}$$

so that the ratio of height to length is

$$\frac{height}{length} = \frac{7+\sqrt{33}}{8}\left(\frac{|\lambda_2|}{\lambda_1}\right)^k \tag{3.20}$$

and as $0 < |\lambda_2| < \lambda_1$ it follows that the ratio

$$\frac{height}{length} \longrightarrow 0 \qquad \text{as} \qquad k \to \infty \tag{3.21}$$

Hence successive simplices become more and more needle like, and in the limit, collapse.

Functions which cause this behaviour

Consider the function in two dimensions given by

$$f(\mathbf{x}) = \begin{cases} \theta\phi|x_1|^\tau + x_2 + x_2^2 & x \leq 0 \\ \theta x_1^\tau + x_2 + x_2^2 & x \geq 0 \end{cases} \tag{3.22}$$

where θ and ϕ are positive constants (see [21] for more details). Note that $(-1,0)$ is the descent direction from the origin and that f is strictly convex

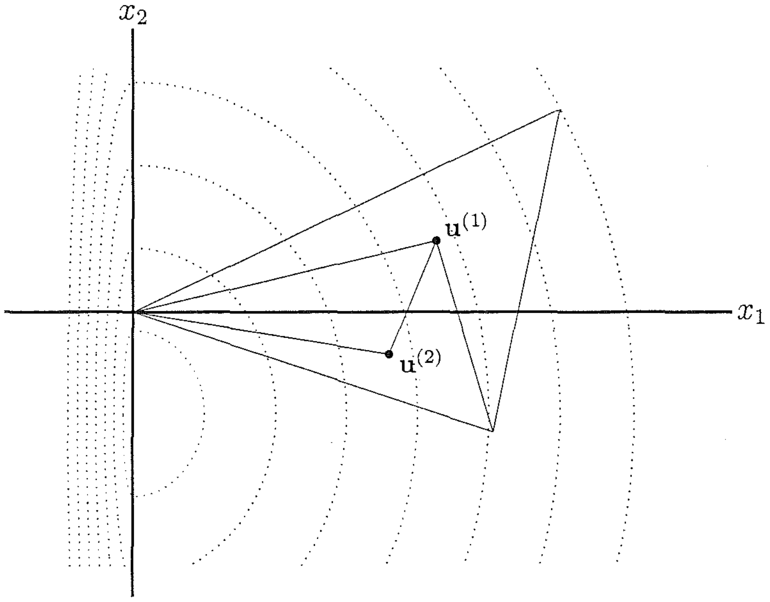


Figure 3.9: Contour plot for the M^{c}Kinnon example also showing the collapse of successive simplices.

provided $\tau > 1$. Figure 3.9 shows a contour plot and the first three simplices for this function with $\tau = 2$, $\theta = 6$ and $\phi = 60$. Any further references to either M^{c}Kinnon's function or M^{c}Kinnon's example, if referring to a particular function, will refer to the function described in equation (3.22) with these parameter values.

M^{c}Kinnon [21, p. 157] states

> These results highlight the need for variants of the Nelder-Mead method which have guaranteed convergence properties.

The next section also begins with a comment from M^{c}Kinnon and introduces another area of concern with the performance of the Nelder-Mead algorithm.

3.4.3 The standard quadratic

M^{c}Kinnon [21, p. 149] makes the comment

> However, it is not yet known even for the function x^2+y^2, the simplest strictly convex quadratic function of two variables, whether the method always converges to the minimiser, or indeed whether it always converges to a single point.

The simplest quadratic functions such as these, for arbitrary dimensions, could be described by

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{x} \tag{3.23}$$

This type of function will be referred to as the *standard quadratic*. Regardless of the dimension, the standard quadratic has a single minimum of zero and the minimiser is the origin. We would expect gradient or conjugate direction type optimisation methods to do very well on this type of problem. The behaviour of the Nelder-Mead algorithm for such problems is by no means clear. Appendix C gives a summary of the performance of `FMINSEARCH` on this function for dimensions which range from two through to 100.

Due to the symmetry of the standard quadratic, it would seem that the choice of the initial point is somewhat arbitrary. A quick check with a selection of different initial points all produced very similar results. The initial point for the standard quadratic used throughout this thesis is $(2, 1, 1, \ldots, 1)$.

It is interesting to note how badly `FMINSEARCH` performs on the 24-d quadratic with low tolerance stopping criteria and the 42-d quadratic with high tolerance stopping criteria. Even though `FMINSEARCH` produces a respectable solution for the 24-d quadratic with high tolerance stopping criteria over two hundred thousand function evaluations are required. Standard quadratics with dimensions $4, 8, 16$ and 24 have been included in the suite of test functions.

3.4.4 No convergence proof

In their paper [20], Lagarias *et al* give some quite limited convergence results for the Nelder-Mead algorithm when used on functions of one and two dimensions. They make the following comment [20, p. 113]

> Remarkably there has been no published theoretical analysis explicitly treating the Nelder-Mead algorithm in more than 30 years since its publication.

Published convergence analysis of simplex based direct search methods impose one or both of the following requirements:

(a) The edges of the simplex remain uniformly linearly independent at every iteration;

(b) A descent condition stronger than simple decrease is satisfied at every iteration.

The Nelder-Mead method fails, in general, to have either of these properties. The resulting difficulties in analysis may explain the long-standing lack of convergence results. At present there is no function in any dimension greater than one for which the Nelder-Mead algorithm has been proved to converge to a minimiser [20, pp. 113–114].

Lagarias *et al* [20, p. 114] also make the comment

> Given all the known inefficiencies and failures of the Nelder-Mead algorithm (see for example [33]), one might wonder why it is used *at all*, let alone why it is so extraordinarily popular. We offer three answers. First, in many applications, for example in industrial process control, one simply wants to find parameter values that improve some performance measure; the Nelder-Mead algorithm typically produces significant improvement for the first few iterations. Second, there are important applications where a function evaluation is enormously expensive or time consuming,

> but derivatives cannot be calculated. In such problems a method that requires at least n function evaluations at every iteration (which would be the case if using finite difference gradient approximations or one of the more popular pattern search methods) is too expensive or too slow. When it succeeds, the Nelder-Mead method tends to require substantially fewer function evaluations than these alternatives, and its relative "best-case efficiency" often outweighs the lack of convergence theory. Third, the Nelder-Mead method is appealing because its steps are easy to explain and simple to program.

McKinnon [21] has recently given a family of strictly convex functions with continuous derivatives and a starting configuration in two dimensions for which all vertices produced by the Nelder-Mead method converge to a non-stationary point. With the failure of the Nelder-Mead algorithm on such a "nice" class of function (in only two dimensions) it is little wonder that no convergence results have been produced — the Nelder-Mead algorithm is not convergent. In order to make the algorithm, as it stands, provably convergent, it would seem that quite harsh restrictions would need to be imposed on the functions for which the results would hold.

Chapter 4

Frame based convergence

The convergence proof used for the variants of the Nelder-Mead algorithm presented in this thesis is based on two research papers by Coope and Price [7], [8]. The relevant sections of these papers are reproduced here for both convenience and completeness.

Although no explicit use of derivatives is made by the Nelder-Mead algorithm or any of the variants discussed in this thesis, the objective function f is assumed to be continuously differentiable with bounded level sets. This restriction on f is a requirement of the convergence proof. The algorithms themselves are direct search methods and so do not make use of derivative information.

The general convergence proof presented here relies on the use of *frames* and *positive bases*. The convergence proof shows that an infinite sequence of central frame points will contain a convergent subsequence whose limit is a stationary point of the objective function. First however, it is necessary to introduce the new terms; positives bases and frames, which are described in the following sections.

4.1 Positive bases

Definition 4.1 (Positive basis) *A positive basis P_+ is a set of vectors $\{\mathbf{p}_i\}$ such that the following conditions hold:*

(a) *Every vector in $\mathbf{R}^n$ can be written as a non-negative combination of the vectors in the positive basis;*

(b) *No element of P_+ is expressible as a non-negative combination of the remaining elements of P_+.*

The variants of the Nelder-Mead algorithm discussed in this thesis all use positive bases with $n+1$ members of the form

$$P_+ = \left\{ \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n, -\frac{1}{n}\sum_{i=1}^{n} \mathbf{p}_i \right\} \tag{4.1}$$

and, in addition, these bases are required to satisfy

$$\left| \det \left(\, [\, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n \,] \right) \right| \geq \delta > 0 \tag{4.2}$$

and

$$\|\mathbf{p}_i\| \leq K \quad \forall i \in \{1, 2, \ldots, n+1\} \tag{4.3}$$

To simplify the notation later on, it will be assumed that the members of the positive basis are ordered. All positive bases mentioned from here on will be assumed to be *ordered* positive bases.

Definition 4.2 (Limit of positive basis) *Any set*

$$\left\{ \mathbf{p}_1^{(\infty)}, \mathbf{p}_2^{(\infty)}, \ldots, \mathbf{p}_{n+1}^{(\infty)} \right\} \tag{4.4}$$

is the limit of the sequence of positive bases $\left\{ P_+^{(k)} \right\}_{k=1}^{\infty}$ if and only if

$$\lim_{k \to \infty} \mathbf{p}_i^{(k)} = \mathbf{p}_i^{(\infty)} \qquad \forall i \in \{1, 2, \ldots, n+1\} \tag{4.5}$$

Equations (4.2) and (4.3) are required conditions for the convergence proof. Equation (4.2) prevents the span of the limit of a sequence of positive bases from collapsing to a subspace of $\mathbf{R}^n$.

Theorem 4.1 *If P_+ is a positive basis and $\mathbf{z} \in \mathbf{R}^n$ then*

$$\mathbf{z}^T\mathbf{p}_i \geq 0 \quad \forall \mathbf{p}_i \in P_+ \quad \Longrightarrow \quad \mathbf{z} = \mathbf{0} \tag{4.6}$$

□

Proof Let

$$-\mathbf{z} = \sum_{i=1}^{n+1} \eta_i \mathbf{p}_i \qquad \text{where} \quad \eta_i \geq 0 \quad \forall\, i \in \{1, 2, \ldots, n+1\} \tag{4.7}$$

then

$$0 \geq -\mathbf{z}^T\mathbf{z} = \sum_{i=1}^{n+1} \eta_i\, \mathbf{z}^T\mathbf{p}_i \geq 0 \tag{4.8}$$

so the only possibility is $\mathbf{z} = \mathbf{0}$ ■

4.2 Frames

Definition 4.3 (Frame) *A frame F consists of a set of $n+2$ points in $\mathbf{R}^n$. The frame $F = F(\mathbf{x}_0, P_+, h)$ is specified in terms of a central point $\mathbf{x}_0$, a positive basis P_+ and a frame size $h > 0$ so that*

$$F = \{\mathbf{x}_0\} \cup \{\mathbf{x}_0 + h\mathbf{p}_i : \mathbf{p}_i \in P_+\} \tag{4.9}$$

For a given positive basis and frame size it is possible to refer to the frame centred on $\mathbf{x}_0$, or the frame about $\mathbf{x}_0$.

The frame size h is adjusted from time to time in a manner that guarantees convergence under the appropriate conditions. Equation (4.3) guarantees that the size of the frame tends to zero (all the frame points tend to the central frame point) as h tends to zero.

Definition 4.4 (Minimal frame) *A frame F about $\mathbf{x}_0$ which satisfies*

$$f(\mathbf{x}_0 + h\mathbf{p}_i) \geq f(\mathbf{x}_0) \quad \forall \mathbf{p}_i \in P_+ \tag{4.10}$$

is a minimal frame about $\mathbf{x}_0$.

Definition 4.5 (Minimal point) *If the frame F about $\mathbf{x}_0$ is a minimal frame then $\mathbf{x}_0$ is called a minimal point.*

For convergence purposes it is more convenient to work with frames that are "nearly" minimal, or *ϵ-quasi-minimal*, where ϵ is a positive constant.

Definition 4.6 (ϵ-quasi-minimal frame) *A frame F about $\mathbf{x}_0$ which satisfies*

$$f(\mathbf{x}_0 + h\mathbf{p}_i) + \epsilon \geq f(\mathbf{x}_0) \quad \forall \mathbf{p}_i \in P_+ \tag{4.11}$$

is an ϵ-quasi-minimal frame.

Definition 4.7 (ϵ-quasi-minimal point) *If the frame F about $\mathbf{x}_0$ is an ϵ-quasi-minimal frame then $\mathbf{x}_0$ is called an ϵ-quasi-minimal point.*

When the value of the constant ϵ is not in doubt the shorter term *quasi-minimal* can be used. For the variants of the Nelder-Mead algorithm discussed in this thesis the value used for ϵ is

$$\epsilon = Nh^{\nu} \tag{4.12}$$

where N is a positive constant and $\nu > 1$. The parameter ϵ effectively acts as a measure of sufficient descent of the objective function over the frame. If there exists a frame point with function value more than ϵ below the function value at the central frame point then the frame is not ϵ-quasi-minimal. If a frame is not ϵ-quasi-minimal then the value of the objective function can be reduced by more than ϵ by moving from the central frame point to one of the other frame points. If it is not possible to reduce the value of the objective function by more than ϵ by moving from the central frame point to one of the other frame points then the central frame point is an ϵ-quasi-minimal point.

These definitions allow for the formation of a generic algorithm framework which is provably convergent and upon which the variants of the Nelder-Mead algorithm are based.

4.3 A convergent frame based template

4.3.1 The framework

The fundamental aim of the algorithm is to generate a sequence of central frame points $\{\mathbf{x}_0^{(k)}\}_{k=1}^{\infty}$ which contains an infinite subsequence $\{\mathbf{x}_0^{(k_j)}\}_{j=1}^{\infty}$ of quasi-minimal points. The convergence theory shows that this subsequence converges to a stationary point of the objective function under mild conditions. The sequence of function values $\{f(\mathbf{x}_0^{(k)})\}_{k=1}^{\infty}$ is monotonically decreasing and so all limit points of $\{\mathbf{x}_0^{(k)}\}$ take the same value.

4.3.2 The algorithm

The algorithm proceeds as follows:

1. Choose an initial starting point to use as the central frame point and values for the constants N, ν, h and κ, which is the scale-factor for h;

2. Select a positive basis. Use this basis and the frame size parameter h to complete the frame about the central point;

3. Evaluate the function at each of the frame points;

4. If choosing the frame point with the lowest function value gives sufficient descent (a reduction of more than ϵ in the value of the objective function) then make the frame point with the lowest function value the new central point and proceed from step 2. Otherwise there exists a quasi-minimal frame about the central point;

5. Reduce the frame size parameter h by the scale-factor κ and proceed from step 2.

The algorithm can be terminated whenever the stopping criteria have been met.

4.4 Convergence analysis

As proved by Coope and Price [8], any algorithm which follows the framework in section 4.3.1 in conjunction with mild conditions on the objective function produces a subsequence $\{\mathrm{x}_0^{(k_j)}\}_{j=1}^{\infty}$ of quasi-minimal points which converge to a stationary point of the objective function.

The convergence properties of the algorithm in section 4.3.2 are stated in two theorems. The first shows that there is an infinite subsequence of quasi-minimal points $\{\mathrm{x}_0^{(k_j)}\}_{j=1}^{\infty}$. Then second shows that this subsequence converges under mild conditions to a stationary point of the objective function.

Theorem 4.2 *Given:*

(a) *The sequence of points* $\{\mathrm{x}_0^{(k)}\}$ *is bounded;*

(b) f *is continuously differentiable;*

then the subsequence $\{\mathrm{x}_0^{(k_j)}\}$ *of quasi-minimal points is infinite.* □

Proof The proof is by contradiction. Assume the subsequence of quasi-minimal frames is finite and let $\mathrm{x}_0^{(k_m)}$ be the last quasi-minimal point in the sequence. Let Ω be the subsequence of $\{\mathrm{x}_0^{(k)}\}$ generated after the last quasi-minimal point is found.

The continuity of f and the boundedness of $\{\mathrm{x}_0^{(k)}\}$ imply that the sequence of function values $\{f(\mathrm{x}_0^{(k)})\}$ is also bounded.

Let $f_0^{(\infty)} = \inf\{f(\mathrm{x}_0^{(k)})\}$. For the current value of ϵ the maximum number of descent steps that can be made before another quasi-minimal frame is found is

$$\frac{f(\mathrm{x}_0^{(k_m)}) - f_0^{(\infty)}}{\epsilon} \tag{4.13}$$

Hence the subsequence Ω contains a quasi-minimal point, contradicting the assumption that $\{\mathrm{x}_0^{(k_m)}\}$ is the last quasi-minimal point in the sequence $\{\mathrm{x}_0^{(k)}\}$. Therefore the subsequence $\{\mathrm{x}_0^{(k_j)}\}$ of quasi-minimal points is an infinite sequence. ■

The next theorem shows that all limit points of the subsequence of quasi-minimal points are stationary points of f.

Theorem 4.3 *Given:*

(a) *The sequence of points* $\{\mathbf{x}_0^{(k)}\} \subseteq \mathcal{F}$ *a bounded, closed subset of* $\mathbf{R}^n$*;*

(b) f *is continuously differentiable;*

(c) *Each positive basis* $P_+^{(k)}$ *satisfies conditions* (4.2) *and* (4.3)*;*

(d) $h^{(k)} \to 0$ *as* $k \to \infty$*;*

(e) $\exists\, N_{max}$ *such that* $0 < N^{(k)} \le N_{max} \quad \forall k \in \mathbf{N}$*;*

then each limit point $\mathbf{x}_0^{(\infty)}$ *of the sequence of quasi-minimal points is a stationary point of* f. □

Proof Replace the subsequence $\{\mathbf{x}_0^{(k_j)}\}_{j=1}^{\infty}$ of quasi-minimal points with a subsequence of itself which converges to $\mathbf{x}_0^{(\infty)}$, and for which the subsequence $\{P_+^{(k_j)}\}$ converges to a unique limit $P_+^{(\infty)}$. Such a subsequence exists by conditions (a) and (c).

The bound on the determinant in (4.2) ensures that $P_+^{(\infty)}$ is a positive basis for $\mathbf{R}^n$.

Now let

$$E_i^{(k_j)} = \int_0^{h^{(k_j)}} \left[\mathbf{p}_i{}^T\Big(\, \nabla f(\mathbf{x}_0 + u\mathbf{p}_i) - \nabla f(\mathbf{x}_0)\,\Big)\right]^{(k_j)} du$$

and

$$M_i^{(k_j)} = \max\left\{ \left\| \,\nabla f(\mathbf{x}_0^{(k_j)} + u\mathbf{p}_i^{(k_j)}) - \nabla f(\mathbf{x}_0^{(k_j)})\, \right\| \, : 0 \le u \le h^{(k_j)}\right\}$$

Since $\mathcal{F}$ is a compact set ∇f is uniformly continuous on $\mathcal{F}$ and so $M_i^{(k_j)} \longrightarrow 0$ as $j \to \infty$ for each $i \in \{1, 2, \ldots, n+1\}$ by condition (d).

Now,

$$\begin{aligned}
\left|E_i^{(k_j)}\right| &\leq \int_0^{h^{(k_j)}} \left(\left\|\mathbf{p}_i^{(k_j)}\right\| \left\|\nabla f(\mathbf{x}_i^{(k_j)} + u\mathbf{p}_i^{(k_j)}) - \nabla f(\mathbf{x}_i^{(k_j)})\right\| \right) \mathrm{d}u \\
&\leq \int_0^{h^{(k_j)}} KM_i^{(k_j)} \mathrm{d}u \\
&\leq h^{(k_j)} KM_i^{(k_j)} \qquad (4.14)
\end{aligned}$$

so that

$$0 \leq \left| \left(\frac{E_i}{h}\right)^{(k_j)} \right| \leq KM_i^{(k_j)} \longrightarrow 0 \text{ as } j \to \infty \qquad (4.15)$$

for each $i \in \{1, 2, \ldots, n+1\}$.

Since the average slope of f in the direction of $\mathbf{p}_i$ in the interval $[\mathbf{x}_0, \mathbf{x}_0 + h\mathbf{p}_i]$ is

$$\frac{f(\mathbf{x}_0 + h\mathbf{p}_i) - f(\mathbf{x}_0)}{h} = \frac{1}{h} \int_0^h \left[\nabla f(\mathbf{x}_0 + u\mathbf{p}_i)\right]^T \mathbf{p}_i \, \mathrm{d}u \qquad (4.16)$$

it follows that

$$\begin{aligned}
&f(\mathbf{x}_0^{(k_j)} + h^{(k_j)}\mathbf{p}_i^{(k_j)}) \\
&\quad = f(\mathbf{x}_0^{(k_j)}) + \int_0^{h^{(k_j)}} \left[\nabla f(\mathbf{x}_0^{(k_j)} + u\mathbf{p}_i^{(k_j)})\right]^T \mathbf{p}_i^{(k_j)} \, \mathrm{d}u \\
&\quad = f(\mathbf{x}_0^{(k_j)}) + \\
&\qquad \int_0^{h^{(k_j)}} \left[\nabla f(\mathbf{x}_0^{(k_j)} + u\mathbf{p}_i^{(k_j)}) - \nabla f(\mathbf{x}_0^{(k_j)}) + \nabla f(\mathbf{x}_0^{(k_j)})\right]^T \mathbf{p}_i^{(k_j)} \, \mathrm{d}u \\
&\quad = f(\mathbf{x}_0^{(k_j)}) + h^{(k_j)} \left[\nabla f(\mathbf{x}_0^{(k_j)})\right]^T \mathbf{p}_i^{(k_j)} + E_i^{(k_j)} \qquad (4.17)
\end{aligned}$$

and from the definition of a quasi-minimal point we have

$$\begin{aligned}
f(\mathbf{x}_0^{(k_j)} + h^{(k_j)}\mathbf{p}_i^{(k_j)}) &\geq f(\mathbf{x}_0^{(k_j)}) - \epsilon^{(k_j)} \\
&\geq f(\mathbf{x}_0^{(k_j)}) - N_{max} \left(h^{(k_j)}\right)^\nu \qquad (4.18)
\end{aligned}$$

Combining equations (4.17) and (4.18) gives

$$\begin{aligned} h^{(k_j)} \left[\nabla f(\mathbf{x}_0^{(k_j)})\right]^T \mathbf{p}_i^{(k_j)} &\geq -N_{max} h^{(k_j)^\nu} - E_i^{(k_j)} \\ \left[\nabla f(\mathbf{x}_0^{(k_j)})\right]^T \mathbf{p}_i^{(k_j)} &\geq -N_{max} h^{(k_j)^{\nu-1}} - \left(\frac{E_i}{h}\right)^{(k_j)} \end{aligned} \qquad (4.19)$$

Equation (4.19) holds for each $i \in \{1, 2, \ldots, n+1\}$ so that as $j \to \infty$

$$\left[\nabla f(\mathbf{x}_0^{(\infty)})\right]^T \mathbf{p}_i^{(\infty)} \geq 0 \qquad \forall i \in \{1, 2, \ldots, n+1\} \qquad (4.20)$$

and so by Theorem 4.1 it follows that

$$\nabla f(\mathbf{x}_0^{(\infty)}) = 0 \qquad (4.21)$$

Since the limit point $\mathbf{x}_0^{(\infty)}$ was chosen arbitrarily, every limit point of the sequence of quasi-minimal points is a stationary point of the objective function. ■

The monotonicity of $\{f(\mathbf{x}_0^{(k)})\}$ means that the sequence $\{\mathbf{x}_0^{(k)}\}$ converges to a set of points on which f is constant.

In the usual case when $\{\mathbf{x}_0^{(k)}\}$ converges to a unique point, that limit point is a stationary point of the objective function.

The convergence theory shows that the constant N can be adjusted after each quasi-minimal iterate. Although not discussed further here, this should permit an algorithm to choose an appropriate value for N which is neither too lax nor too strict, and this should improve the performance of the algorithm.

The convergence framework also permits both increases and decreases to h on an iteration by iteration basis. Given that the Nelder-Mead algorithm can increase or decrease the size of a simplex, this flexibility in h is essential if frame based variants of the Nelder-Mead algorithm which capture the essential features of the Nelder-Mead algorithm are to be constructed. The choice of an appropriate h is important because if h is too small the algorithm will have to take many smaller steps requiring many function evaluations to make significant progress. On the other hand if h is too large a succession of identical quasi-minimal iterates may be generated before any movement of the central frame point is made.

Algorithms which rely on these frame based convergence results are generally easy to construct, understand and examine. If any variant of the Nelder-Mead algorithm completes, and then if necessary, reduces the size of a frame whenever insufficient progress is being made, the algorithm will fit into the above convergence framework and so will be provably convergent.

Provided the current simplex in the Nelder-Mead algorithm has not collapsed the vertices of the simplex can be used for $n + 1$ of the $n + 2$ points required to complete a frame about the best vertex. All that remains is to find a suitable final frame point to complete the frame. In fact, as we shall see, this "choice" of final frame point is made automatically due to the choice of positive basis made in equation (4.1).

Chapter 5

Variants of Nelder-Mead

As previously mentioned, the Nelder-Mead algorithm generally performs well in most situations. However, it has also been shown that there are some classes of quite nice functions (for example the 24-d and 42-d standard quadratics and M^{c}Kinnon's example) for which the performance of the Nelder-Mead algorithm ranges from poor to fails miserably. Perhaps it is best to start with an overview of what might be happening when the Nelder-Mead algorithm performs poorly.

5.1 Failures of the Nelder-Mead algorithm

5.1.1 Collapse of the simplex

As illustrated by M^{c}Kinnon's example, performing repeated contract-inside steps while the best vertex remains fixed can cause successive simplices to collapse. If this happens the Nelder-Mead algorithm will be trapped in some subspace of the search space. Since the Nelder-Mead algorithm only performs vector addition and scalar multiplication, once it is in a subspace it will not be able to generate a descent direction out of the subspace and escape. The best the Nelder-Mead algorithm can do in this situation is to find the minimum of the objective function restricted to the subspace.

5.1.2 Failure to make sufficient progress

Even if the sequence of successive simplices do not collapse, the trial points used by the Nelder-Mead algorithm may suffer from an ever diminishing descent condition. Eventually the algorithm will terminate, either because the simplex is sufficiently small or the maximum number of function evaluations has been reached. Possible solutions to these problems will be discussed in the following sections. Recently (1999), Kelley [18] proposed a sufficient descent test, which, if passed at every iteration, guarantees the convergence of the Nelder-Mead algorithm to a stationary point under mild conditions on the objective function. Furthermore, if insufficient progress is made, Kelley proposed an *oriented restart* using a difference approximation to the steepest descent direction from the best vertex. Although further discussion of Kelley's ideas is outside the scope of this thesis, it does help to illustrate how topical the Nelder-Mead algorithm is at the present time.

5.2 Requirements for a successful variant

It was decided at the beginning of this project that any successful variant of the Nelder-Mead algorithm should meet the following requirements:

(a) Any changes should keep the "spirit" of the Nelder-Mead algorithm. The changes should be based, as much as possible, on ideas already present in the Nelder-Mead algorithm;

(b) None of the changes should hinder the performance of the Nelder-Mead algorithm when it is performing well; and

(c) If the Nelder-Mead algorithm begins to perform poorly the variant should intervene and prevent further poor performance.

To assist with the descriptions of the variants presented in this thesis it is convenient to introduce some new notation and definitions.

Definition 5.1 (Side vector) *The i^{th} side vector $\mathbf{s}_i$ of a simplex V is*

$$\mathbf{s}_i = \mathbf{v}_i - \mathbf{v}_0 \qquad i \in \{1, 2, \dots, n\} \tag{5.1}$$

Definition 5.2 (Side length) *The length s_i of the i^{th} side vector is*

$$s_i = \| \mathbf{s}_i \| \tag{5.2}$$

Definition 5.3 (Unit side vector) *The i^{th} unit side vector $\hat{\mathbf{s}}_i$ is the vector with the same direction as $\mathbf{s}_i$ such that*

$$\| \hat{\mathbf{s}}_i \| = 1 \tag{5.3}$$

5.2.1 Preventing simplex collapse

One method of detecting whether the current simplex is near collapse is to calculate the determinant of the matrix whose columns are the side vectors of the simplex. These side vectors are illustrated in figure 5.1. In order to

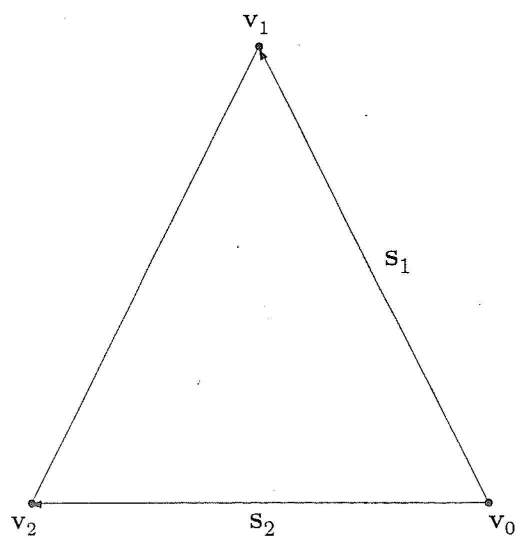


Figure 5.1: Simplex side vectors.

remove the scaling effects of different side lengths, the unit side vectors could be used. The value of this determinant is the volume of the parallelepiped whose edges are defined by the (unit) side vectors of the simplex. In the two dimensional case this represents the area of the parallelogram whose sides are the (unit) side vectors of the simplex. At this point it is useful to give a precise definition of what is meant by the collapse of a simplex.

Definition 5.4 (Simplex collapse) *The current simplex is deemed to have collapsed if*

$$\left| \det \left(\, [\, \hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \ldots, \hat{\mathbf{s}}_n \,] \, \right) \right| < \delta \tag{5.4}$$

for some lower bound δ on the determinant.

Calculating determinants

Since calculating a determinant directly would involve $O(n^3)$ operations it would be an advantage if this calculation could be done more cheaply. This can be achieved most easily if the initial simplex is orthogonal.

Definition 5.5 (Orthogonal simplex) *Given some initial point $\mathbf{v}_0$, the simplex V formed by $\mathbf{v}_i = \mathbf{v}_0 + \mathbf{s}_i, \quad i \in \{1, 2, \ldots, n\}$ is an orthogonal (and as yet unordered) simplex if and only if*

$$\mathbf{s}_i \perp \mathbf{s}_j \quad \textit{whenever} \quad i \neq j \tag{5.5}$$

If the initial simplex V is orthogonal then the determinant initially equals the product of the side lengths. That is

$$\left| \det \left([\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n] \right) \right| = \prod_{i=1}^{n} s_i \tag{5.6}$$

and the volume of the simplex V is

$$\mathrm{vol}(V) = \frac{1}{n!} \prod_{i=1}^{n} s_i \tag{5.7}$$

As mentioned in section 3.2 each of the algorithm operations; reflection, expansion, contraction and shrink, scale the volume of the simplex by the value of the coefficients α, γ, β and σ^n respectively. This allows the calculation of the determinant to be done indirectly by keeping track of the relative change in the volume of the current simplex compared to the volume of the initial simplex, with a volume scale-factor μ.

If the initial simplex $V^{(0)}$ is orthogonal then

$$\mathrm{vol}(V^{(0)}) = \frac{1}{n!} \prod_{i=1}^{n} s_i^{(0)} \tag{5.8}$$

If the current simplex is $V^{(k)}$ then its volume will be

$$\mathrm{vol}(V^{(k)}) = \frac{1}{n!} \det\left(\left[\hat{\mathbf{s}}_1^{(k)}, \hat{\mathbf{s}}_2^{(k)}, \hat{\mathbf{s}}_3^{(k)}, \ldots, \hat{\mathbf{s}}_n^{(k)}\right]\right) \prod_{i=1}^{n} s_i^{(k)} \tag{5.9}$$

but this will be equal to the original volume multiplied by the relative change in volume μ, and so

$$\mathrm{vol}(V^{(k)}) = \mu \ \mathrm{vol}(V^{(0)}) \tag{5.10}$$

hence

$$\det\left(\left[\hat{\mathbf{s}}_1^{(k)}, \hat{\mathbf{s}}_2^{(k)}, \hat{\mathbf{s}}_3^{(k)}, \ldots, \hat{\mathbf{s}}_n^{(k)}\right]\right) = \mu \frac{\prod_{i=1}^{n} s_i^{(0)}}{\prod_{i=1}^{n} s_i^{(k)}} \tag{5.11}$$

Reshaping the simplex

If the determinant calculation (either directly, or indirectly via the change in volume) reveals that the simplex has fallen below the acceptable bounds for non-collapse then it must be reshaped somehow. There are several possibilities. The ones considered in this thesis ($\psi_1 - \psi_4$) are presented below.

(ψ_1) An orthogonal simplex formed from the orthogonal decomposition of the current longest side about the best point. If $\mathbf{s}_j$ is the longest side, then

$$s_j \geq s_i \qquad \forall\, i \in \{1, 2, \ldots, n\} \tag{5.12}$$

The side vectors which form the new simplex $V^{(+)}$ are orthogonal, and

$$\sum_{i=1}^{n} \mathbf{s}_i^{(+)} = -\mathbf{s}_j \tag{5.13}$$

This orthogonal decomposition can be calculated efficiently with the use of *Householder* matrices. This method of simplex reshape will be referred to as the *HH* method of simplex reshape, or more briefly as the *HH* simplex. The MATLAB code used to generate this orthogonal decomposition is included in appendix H.3 on page 225.

(ψ_2) A regular orthogonal simplex about the current best vertex with sides parallel to the coordinate axes. Although regular and orthogonal individually may seem contradictory, when used together they are taken to mean an orthogonal simplex with

$$s_i = s_j \qquad \forall\, i, j \in \{1, 2, \ldots, n\} \tag{5.14}$$

Each side vector of the new simplex is parallel to one of the coordinate axes. The side lengths of the new simplex are equal to the length of the shortest side of the collapsed simplex. This method of simplex reshape will be referred to as the *regular IJK* method of simplex reshape using the shortest side length, or more briefly, the *IJK* simplex. Note that due to the computer's finite precision arithmetic, the shortest side length may actually be zero. In this case some minimum side length would be used.

(ψ_3) An orthogonal simplex about the best vertex which preserves the longest side vector of the collapsed simplex. The longest side vector is preserved because presumably this side of the simplex is (locally at least) parallel to the contours of the objective function, and so will lie along a valley floor. A QR decomposition can be used to efficiently achieve the orthogonalisation of the remaining side vectors. This method of simplex reshape will be referred to as the *QR* method of simplex reshape, or more briefly as the *QR* simplex.

(ψ_4) The simplex that FMINSEARCH would construct about the current best vertex (described in section 3.3). This method of simplex generation only takes into account the position of the best vertex. Some experimentation has shown that although this may be a good method of generating an initial simplex, it is not very good in practise if repeated simplex reshaping is required. This is probably because it does not make use of any other simplex information — such as the simplex orientation or size, as used in the other simplex reshaping methods discussed above. So although this method may produce a useful initial simplex it will not be considered further as a means of reshaping a collapsed simplex.

5.2.2 Sufficient descent condition

The sufficient descent condition measures whether sufficient progress is made by the algorithm at each iteration.

Definition 5.6 (Sufficient descent) *Sufficient descent is being made if the function value at the worst vertex of the simplex is reduced by a sufficient amount at each iteration.*

Sufficient descent can be measured by using a *sufficient descent parameter* ϵ, which can be altered from time to time in accordance with the convergence proof.

Perhaps the simplest way of implementing a sufficient descent condition is to set some minimum level of descent ϵ, which must be met before the Nelder-Mead move will be accepted. If the specified level of descent is not obtained then some intervention is required. If the Nelder-Mead algorithm is meeting the sufficient descent condition at each iteration then any variant should let the Nelder-Mead algorithm proceed unhindered.

5.3 Ranking the performance of the variants

The following ranking system was used in order to determine which of the variants performed the best.

(a) An algorithm that produces an accurate approximation to the solution for *every* function in the test suite is better than an algorithm that does not.

(b) An algorithm that requires fewer function evaluations to produce the approximations to the solutions of the functions in the test suite is better than one that requires more function evaluations.

(c) An algorithm that produces more accurate approximations to the solutions is better than an algorithm that is less accurate.

Since the minimum of most of the functions in the test suite is zero, the base ten logarithm was used to rank the relative accuracy of two different algorithms as outlined in (c). The lower the value of

$$\log_{10}(f(\mathbf{x}) - f(\mathbf{x}^*)) \tag{5.15}$$

where $f(\mathbf{x})$ is the approximation to a minimum of the objective function $f(\mathbf{x}^*)$, produced by the algorithm, the better the solution. This ranking system is ineffective at ranking the accuracy of the approximations to non-zero solutions as, typically, these solutions have only been given to a few significant digits. Rather than ranking the performance of the algorithms for non-zero solutions in this way, the algorithms were deemed to have either succeeded (that is, produced an accurate approximation to the solution) or failed.

There is some interplay between (b) and (c) since if an algorithm requires very few function evaluations to produce an approximation to the solution, but the approximation is poor, then it may be better to use an algorithm that requires more function evaluations but produces a more accurate approximation.

All of the variants were tried on the functions in the test-suite with two sets of tolerances. The low tolerance termination criteria were chosen to match the default settings for `FMINSEARCH` which are

$$x_{tol} = 10^{-4} \quad \text{and} \quad f_{tol} = 10^{-4} \tag{5.16}$$

and the high tolerance criteria are

$$x_{tol} = 10^{-8} \quad \text{and} \quad f_{tol} = 10^{-12} \tag{5.17}$$

Since some of the problems in the test suite are so badly scaled, the high tolerance termination criteria are required to prevent the algorithms from terminating prematurely. As such, only the high tolerance results were considered when ranking the performance of the algorithms. It is now appropriate to introduce the variants of the Nelder-Mead algorithm.

5.4 Variant one

Variant one (`NM1`) is the standard Nelder-Mead algorithm with a sufficient descent condition and a new strategy if insufficient progress has been made.

5.4.1 What if there is insufficient descent?

If insufficient progress has been made then some other trial points have to be found that hopefully will give sufficient descent. Since the choice of an alternate trial point is to keep the spirit of the Nelder-Mead algorithm, one approach is to take a step back for a moment and consider how the current simplex could have been created.

Experimental evidence suggests that, in general, the reflection step is the most common. So it is likely that the current simplex is the result of a reflection from some previous simplex. Since there is no way of telling for sure if this is what actually happened, this previous simplex will be referred to as the *ghost simplex.* An illustration of the ghost simplex is shown in figure 5.2, where H denotes the vertex of the ghost simplex with the highest function value. If the current simplex is the result of such a reflection then the reflect point has the lowest function value and so an expand step would also be attempted. This would generate a new point not on the usual Nelder-Mead search line for the current simplex. Since this expansion step uses the ghost simplex rather than the current simplex, it will be referred to as the *pseudo expansion* step and is illustrated in figure 5.3.

The Nelder-Mead algorithm attempts to find a point with function value lower than the worst vertex of the current simplex by looking along the line from the worst vertex through the centroid of the n remaining vertices. The pseudo expand move attempts to find a better point by looking along the line through the best vertex and the centroid of the n remaining vertices. The Nelder-Mead algorithm tries a search direction away from the worst vertex and `NM1` tries a search direction towards the best vertex.

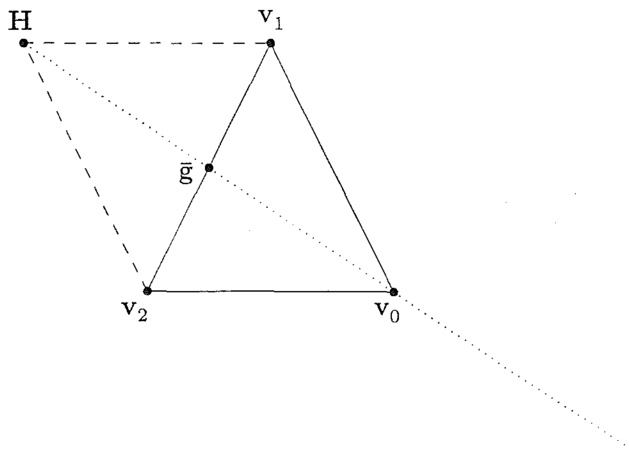


Figure 5.2: The reflection step from the ghost simplex.

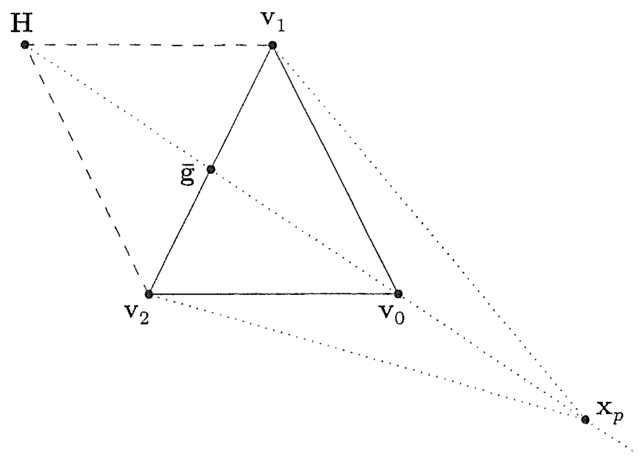


Figure 5.3: The expansion step from the ghost simplex which yields the pseudo expansion point.

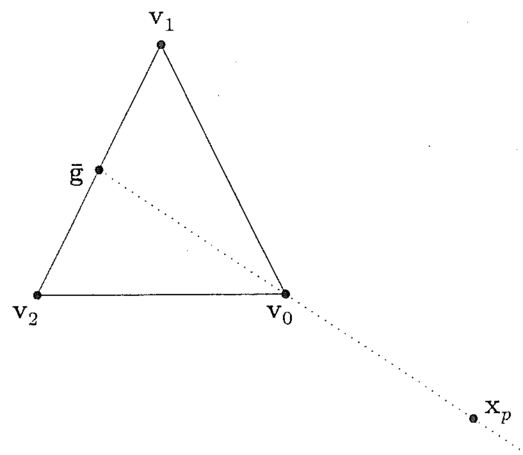


Figure 5.4: An alternate search direction for the current simplex.

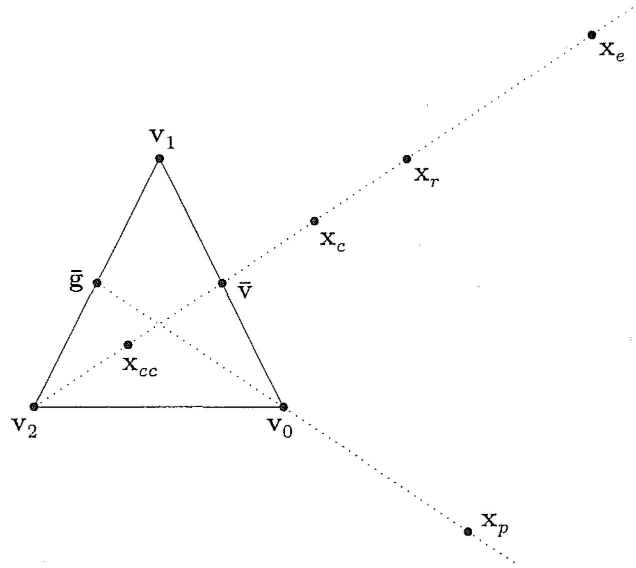


Figure 5.5: All of the trial points for variant one.

If the centroid for the ghost simplex is labelled $\bar{g}$ then reflection from the ghost simplex gives

$$v_0 = \bar{g} + \alpha(\bar{g} - v_2^{(-)}) \tag{5.18}$$

and the pseudo expand point x_p is given by

$$x_p = \bar{g} + \gamma(\bar{g} - v_2^{(-)}) \tag{5.19}$$

Clearly, from equation (5.18) $\alpha(\bar{g} - v_2^{(-)}) = v_0 - \bar{g}$ so that

$$\begin{aligned} x_p &= \bar{g} + \frac{\gamma}{\alpha}(v_0 - \bar{g}) \\ &= \frac{\gamma}{\alpha}v_0 - \left(\frac{\gamma - \alpha}{\alpha}\right)\bar{g} \end{aligned} \tag{5.20}$$

and using the usual values for the parameters ($\alpha = 1,\ \gamma = 2$) gives

$$x_p = 2v_0 - \bar{g} \tag{5.21}$$

If the Nelder-Mead algorithm failed to make sufficient progress with any of the usual trial points found by reflection, expansion or contraction, an alternative trial point could be generated by the pseudo expand step. Together these ideas seem like a suitable approach for finding a point that gives sufficient descent. If both of these fail to produce sufficient descent then the simplex can be shrunk as per usual in the Nelder-Mead algorithm.

5.4.2 Algorithm outline for variant one

1. Use the Nelder-Mead algorithm while it is producing sufficient descent.

2. If the Nelder-Mead algorithm fails to produce sufficient descent then try the pseudo expand point.

3. If the pseudo expand point produces sufficient descent then use this point to create the new simplex and continue using the Nelder-Mead algorithm from step 1.

4. If the pseudo expand point does not make sufficient progress check whether the simplex has collapsed, if it has, then reshape the simplex about the current best vertex, reduce the sufficient descent parameter ϵ and proceed from step 1.

5. If the simplex has not collapsed, shrink the simplex, reduce the sufficient descent parameter ϵ and proceed from step 1.

A schematic diagram for `NM1` is given in appendix D.1 on page 102.

5.4.3 Convergence of variant one

This algorithm is basically the Nelder-Mead algorithm with an extra search point added on. As such, it suffers from the same problems that the Nelder-Mead algorithm does as far as convergence results go, and so in general, there is no convergence proof.

5.4.4 Performance of variant one

A complete list of the results for NM1 on the suite of test functions is given in appendix E.1. As can be seen, each of the methods for reshaping a collapsed simplex ($\psi_1 - \psi_3$) were tried. Unfortunately none of them produced good results for every function in the test suite. For 30 of the 38 functions in the test suite the performance of NM1 was similar to that of FMINSEARCH. This seems reasonable since NM1 is basically FMINSEARCH with an extra trial point added if insufficient progress is made. Some of the results worth noting are discussed below.

(a) Although all three variations of NM1 produced an accurate approximation to the minimum of the Bard 3-d test function, the QR simplex reshape method required 91979 function evaluations to reach the termination criteria compared to 943 function evaluations required using the HH simplex reshape method and 6645 function evaluations required using the IJK simplex reshape method.

(b) Each variation of NM1 found the global minimum (zero) of the Biggs EXP6 test function, whereas FMINSEARCH located the stationary point with function value $5.65565\ldots \times 10^{-3}$. All of the algorithms terminated after a similar number of function evaluations.

(c) FMINSEARCH successfully reached the termination criteria but failed to produce accurate approximations to the function minima for each of the five higher dimension test problems: Extended Rosenbrock 8-d, Watson 9-d, Extended Rosenbrock 10-d, Penalty (1) 10-d, and, Penalty (2) 10-d, whereas, NM1 did produce accurate approximations to the minima (except that NM1 using the QR method of simplex reshape failed to terminate before 100000 function evaluations were made).

(d) FMINSEARCH required 4926 function evaluations to produced an accurate approximation to the minimum ($4.01377\ldots \times 10^{-2}$) of the Osbourne (2) 11-d test function. The NM1 algorithm using the IJK simplex reshape produced the same solution with 9760 function evaluations, whereas the HH and QR simplex reshape versions of NM1 required 15052 and 15659 function evaluations respectively to produce 3.29282×10^{-1} as the (poor) approximation to the solution.

(e) Finally, the HH, IJK and QR variations of the NM1 algorithm required 98685, 96384 and 32961 function evaluations respectively to produce accurate approximations (of the order of 10^{-8}) to the solution of the 24-d standard quadratic (zero). Whereas FMINSEARCH failed to terminate before 100000 function evaluations were reached and produced a poor approximation (5.04216×10^{-1}) to the minimum.

Overall, NM1 performed better than FMINSEARCH on the suite of test functions. However, there were several test functions for which each variation of NM1 performed poorly. This, and the lack of a convergence proof, shows the need for the development of a new algorithm. This new algorithm (variant two) was designed to be provably convergent right from the start and is discussed in the next section.

5.5 Variant two

The main difference between variant one and variant two (NM2) is that NM2 uses the frame based convergence proof to provide the framework for the new algorithm. As with variant one, if the Nelder-Mead algorithm is making sufficient progress then it is allowed to proceed unhindered. Whenever NM2 fails to make sufficient progress a frame is completed about the current best vertex. If the simplex has collapsed, then it is reshaped before the frame is completed.

To make use of the simplex vertices when completing the frame about the best vertex, the first n positive basis vectors are defined as

$$\mathbf{p}_i = \frac{\mathbf{v}_i - \mathbf{v}_0}{h} \qquad i \in \{1, 2 \ldots, n\} \tag{5.22}$$

That is, the basis vectors are the side vectors of the simplex scaled by a suitable amount so that $n+1$ of the frame points are the same as the vertices of the simplex. The frame points are defined as follows: the central frame point $\mathbf{x}_0$ is the best vertex of the simplex $\mathbf{v}_0$, and the remaining frame points are

$$\begin{aligned} \mathbf{x}_i &= \mathbf{x}_0 + h\mathbf{p}_i \qquad i \in \{1, 2, \ldots, n+1\} \\ &= \mathbf{v}_0 + h\mathbf{p}_i \qquad i \in \{1, 2, \ldots, n+1\} \end{aligned} \tag{5.23}$$

so that

$$\mathbf{x}_i = \mathbf{v}_i \qquad i \in \{0, 1, \ldots, n\} \tag{5.24}$$

From the choice of positive basis detailed in section 4.1 the final frame point is given by

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{x}_0 + h\,\mathbf{p}_{n+1} \\ &= \mathbf{v}_0 - \frac{h}{n}\sum_{i=1}^{n}\mathbf{p}_i \\ &= \mathbf{v}_0 - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{v}_i - \mathbf{v}_0) \\ &= \mathbf{v}_0 + \mathbf{v}_0 - \frac{1}{n}\sum_{i=1}^{n}\mathbf{v}_i \\ &= 2\mathbf{v}_0 - \bar{\mathbf{g}} \\ &= \mathbf{x}_p \end{aligned} \tag{5.25}$$

and so, as illustrated in figure 5.6, `NM2` creates a frame about the best vertex $\mathbf{v}_0$ by using the n remaining vertices of the simplex and the pseudo expand point introduced in `NM1`, to complete the frame. An algorithm outline for `NM2` follows, and a schematic diagram is given in appendix D.2 on page 103.

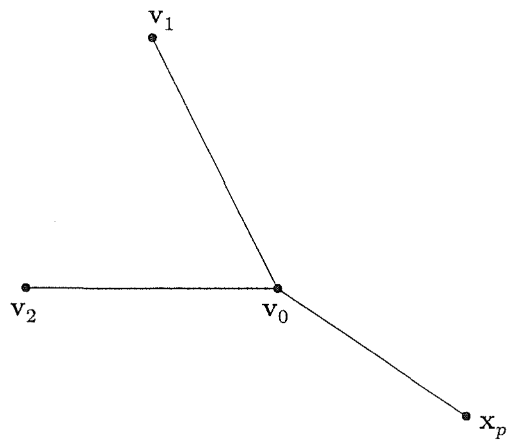


Figure 5.6: Creating the frame about the best vertex of the current simplex.

5.5.1 Algorithm outline for variant two

1. Use the Nelder-Mead algorithm while it is producing sufficient descent.
2. If the Nelder-Mead algorithm fails to produce sufficient descent check whether the simplex has collapsed and reshape if necessary.
3. Use the pseudo expand point to complete a frame about the best vertex.
4. If there is sufficient descent, then (if the pseudo expand point is better than $\mathbf{v}_0$ then replace $\mathbf{v}_0$ with the pseudo expand point), continue with the usual Nelder-Mead algorithm.
5. If there is an ϵ-quasi-minimal frame about the best vertex then reduce h, recalculate ϵ, complete the new (smaller) frame about the best vertex and calculate the function values at the new frame points. Repeat this step until either ϵ descent is possible, or the frame size is small enough that the terminating criteria are met and the algorithm can be stopped.

5.5.2 Convergence of variant two

Since this algorithm meets all of the requirements of the frame based convergence proof it is a provably convergent algorithm. The performance of NM2 on the suite of test functions is discussed in the next section.

5.5.3 Performance of variant two

Although NM2 is a provably convergent algorithm its performance on some of the problems from the suite of test functions was disappointing. In general NM2 required more function evaluations than FMINSEARCH to reach the terminating criteria, probably because when the frame based convergence part of the algorithm takes over, up to $n+1$ function evaluations are required at each iteration.

The poor performance of NM2 in practise is probably because the algorithm fails to make sufficient progress with a simplex that is near to, but not quite collapsed. In this situation the frame size parameter h may be reduced too quickly for the algorithm to make any significant progress.

A complete list of the results for NM2 on the suite of test functions is given in appendix E.2. Once again, each of the methods for reshaping a collapsed simplex ($\psi_1 - \psi_3$) were tried. As with NM1, none of these produced good results for every function in the test suite. Some of the results worth noting are discussed below.

(a) With each of the methods of reshaping a collapsed simplex ($\psi_1 - \psi_3$), NM2 returned zero as the approximation to the minimum of M^{c}Kinnon's function, rather than the true solution, -0.25.

(b) Using the QR simplex reshape method, NM2 required 26348 function evaluations to produce 2.42677×10^{-1} as the (poor) approximation to the minimum of Biggs EXP6 function, whereas the other NM2 variants required around 3000 function evaluations to locate the stationary point with function value 5.65565×10^{-3}. FMINSEARCH required only 1130 function evaluations to locate this stationary point.

(c) The IJK variant of NM2 required 54256 function evaluations to produce 2.94423×10^{-1}, as the (poor) approximation to the minimum (zero) of the Extended Rosenbrock 10-d function.

(d) The HH variant of NM2 required 2754 function evaluations to produce 1.91794×10^{1}, as the (poor) approximation to the minimum (zero) of the 24-d standard quadratic. Interestingly, the IJK variant of NM2 was particularly successful at producing accurate approximations to the minimum of the standard quadratic, producing solutions of around 10^{-16} with 890 and 1332 function evaluations for the 16-d and 24-d standard quadratics respectively.

Generally, NM2 produced accurate approximations to the minima of the functions in the test suite more often than either FMINSEARCH or NM1. However, for the functions which NM1, NM2, and FMINSEARCH all produced accurate approximations, NM2 tended to require more function evaluations to reach the terminating criteria. This is probably to be expected as, although completing a frame about the best vertex only requires one extra function evaluation, if the central frame point remains fixed for several iterations then $n+1$ function evaluations are required to complete the new frame every time h is reduced.

Overall, NM2 performed better on the suite of test functions than either NM1 or FMINSEARCH however there were still several test functions for which each variation of NM2 performed poorly. In particular, the failure of NM2 (with each of the simplex reshape methods) on M^{c}Kinnon's function highlights the need for the development of a better algorithm.

5.6 Variant three

In practise, the Nelder-Mead algorithm tends to work well in most situations, but it is not convergent. Since NM2 is provably convergent but does not do so well in practise it might be advantageous to allow the Nelder-Mead algorithm a greater opportunity to meet the sufficient descent condition. Variant three

(NM3) lets the Nelder-Mead algorithm have several attempts to meet the sufficient descent condition before the frame based part of the algorithm intervenes. Initially the Nelder-Mead algorithm was allowed, at most, n attempts to make sufficient descent. The performance of NM3 is discussed in the next section. As with NM2, this algorithm meets all of the criteria required by the frame based convergence proof and so it is a provably convergent algorithm. In fact any finite process could be used instead of letting the Nelder-Mead algorithm have several attempts at making sufficient descent and still be within the framework of the convergence proof.

A schematic diagram for NM3 is given in appendix D.3 on page 104.

5.6.1 Performance of variant three

Even though NM3 is provably convergent, and the Nelder-Mead part of the algorithm is given a greater opportunity to make sufficient progress, its performance on the suite of test functions is similar to NM2. A complete list of the results for NM3 on the suite of test functions is given in appendix E.3. Once again, each of the methods for reshaping a collapsed simplex ($\psi_1 - \psi_3$) were tried. As with NM2, none of these produced good results for every function in the test suite. Some of the results worth noting are discussed below.

(a) For each of the functions: Powell badly scaled 2-d, Brown badly-scaled 2-d, Bard 3-d, Meyer 3-d, Box 3-d, Penalty (1) 4-d, Penalty (1) 10-d, and Penalty (2) 10-d, at least one of the NM3 variants ($\psi_1 - \psi_3$) meets the terminating criteria with relatively few function evaluations but produces poor approximations to the function minima. In these cases, it seems as though the simplices have become trapped in some subspace of the solution space and so the algorithms are unable to produce more accurate approximations to the function minima. This may be because allowing the Nelder-Mead algorithm to proceed even though insufficient progress is being made leads into a subspace from which even the frame based part of the algorithm cannot, in practise, escape.

(b) Even trying a different numbers of attempts to make sufficient descent before resorting to the frame based part of the algorithm failed to produce notably better results. Only results for the case when at most n attempts were allowed have been included but other quantities tried were: $1, 2, 5, 10, 20, n$, $n + 1$, $2n, 3n, 4n$ and finally $\lceil \sqrt{n}+1 \rceil$. Typically, (and somewhat annoyingly) each alternative produced an algorithm that worked remarkable well on some of the problems in the test suite, but failed on one or two others.

So NM3 did not produce satisfactory results for every function in the test suite. Since an algorithm which is provably convergent, *and* works well in practise is desired, investigating why NM2 and NM3 do not perform well in practise may be beneficial. This is discussed further in the following section.

5.7 Variant four

Since none of the three simplex reshape options used with NM2 perform well on M^{c}Kinnon's function, understanding this failure may provide a key to improving the algorithm's performance. A possible reason for this failure is illustrated in figure 5.7. Suppose a frame is centred on the origin. The central frame point is the best point of the frame. However, it does not seem that any amount of shrinking the frame will produce a point with a lower function value. The frame proof guarantees that under such a condition the central frame point must be a stationary point. This is clearly not the case for M^{c}Kinnon's function. What is going on? The scale of the figure hides what is really happening. Since the first derivatives of the M^{c}Kinnon function are continuous, close enough to the origin the contour lines shown must cross one of the frame lines. A close-up of the origin is shown in figure 5.8 to illustrate the point. From figure 5.8 it is clear that if the frame is reduced enough, eventually a point with lower function value can be found. However, if this only happens when the frame size is very small, then convergence, although guaranteed, may now only be theoretical rather than practical.

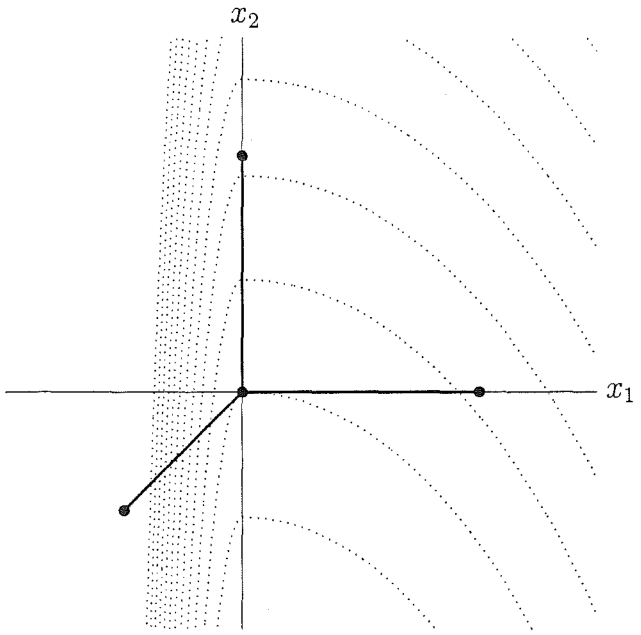


Figure 5.7: An orthogonal frame about the origin and McKinnon's function.

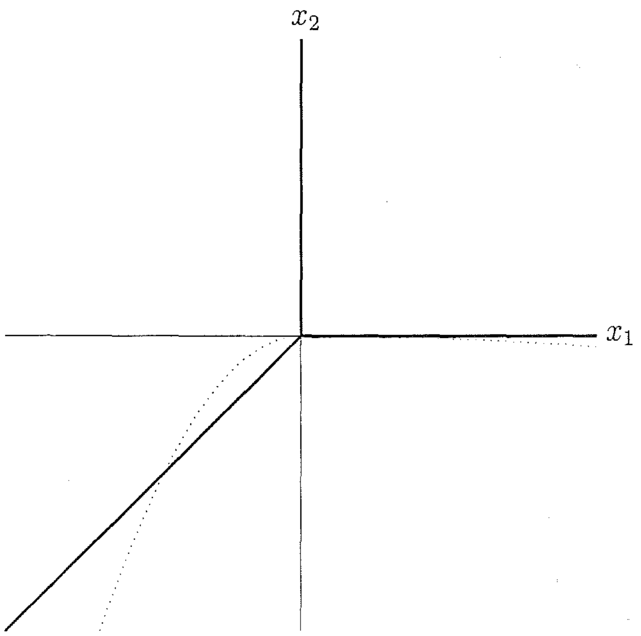


Figure 5.8: A close-up of the frame about the origin and McKinnon's function.

An easy way to overcome this type of poor performance is to reverse the direction of the positive basis vectors every time a frame is reduced in size. This should allow a descent direction to be found more easily. This approach is used in variant four (`NM4`), which is a refinement of `NM2`.

Also, if it is necessary to create a frame made up of new points, then it may as well be a "good" one, that is, one in which the first n positive basis vectors are orthogonal. After all, if the simplex has not collapsed, but has become badly mis-shaped, then even reversing the direction of the positive basis vectors will not help produce a descent direction without a big reduction in the frame size (exactly as in the McKinnon example above). This is illustrated in figure 5.9 where the near-collapse simplex is shown with a solid outline and the frame points with alternating positive basis vectors are shown as dashed lines.

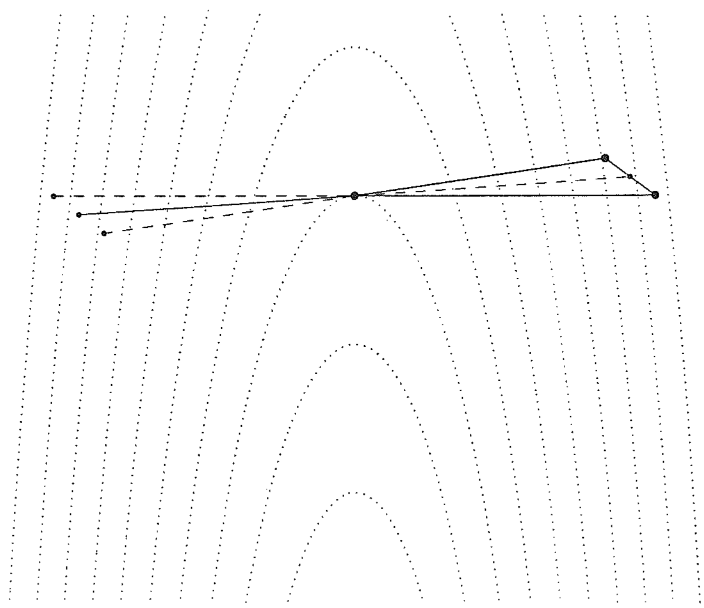


Figure 5.9: If a simplex is near collapse then even alternating the direction of the positive basis vectors may not produce a descent direction without a big reduction in the frame size.

It seems reasonable to expect these refinements will improve the performance of the algorithm, while still keeping within the requirements of the frame based convergence proof. A schematic diagram for NM4 is given in appendix D.4 on page 105 and its performance on the suite of test functions is discussed in the following section.

5.7.1 Performance of variant four

A complete list of the results for NM4 on the suite of test functions is given in appendix E.4. Once again, each of the methods for reshaping a collapsed simplex ($\psi_1 - \psi_3$) were tried. The performance of NM4 on the suite of test functions is discussed below.

(a) NM4 with either the HH or IJK simplex reshape methods, fails to produce accurate approximations to the minima of the Brown badly scaled 2-d function and the M^{c}Kinnon 2-d function.

(b) Using the IJK simplex reshape method NM4 required 83758 function evaluations to reach the terminating criteria, compared to 1288 function evaluations with the HH simplex reshape method and 1622 function evaluations with the QR reshape method.

(c) NM4 with the HH simplex reshape method fails to produce accurate approximations to the minima of the Meyer 3-d function and the Watson 9-d function.

(d) Using the IJK simplex reshape method NM4 fails to produce accurate approximations of the minima of the Penalty (1) 10-d function and the Penalty (2) 10-d function. In addition, the terminating criteria had not met after 100000 function evaluations.

(e) Using the HH simplex reshape method NM4 did not meet the terminating criteria after 100000 function evaluations with the Extended Powell 12-d function.

(f) NM4 with the HH simplex reshape method failed to produce an accurate approximation to the minimum of the 24-d standard quadratic.

(g) Using the QR simplex reshape method NM4 successfully reached the terminating criteria and produced accurate approximations to the minima for *every* function in the test suite.

Overall, with both the HH and IJK simplex reshape methods, NM4 generally performs quite well for most of the functions in the test suite. Unfortunately, there are some functions for which the algorithm fails to terminate before 100000 function evaluations have been calculated, or produces poor approximations to the function minima. However, using the QR simplex reshape method, NM4 successfully reached the terminating criteria well before 100000 function evaluations were reached and produced accurate approximations to the function minima for *every* function in the test suite.

5.7.2 Parameter selection for variant four

Apart from the method of reshaping a collapsed simplex, there are other parameters within NM4 that could also be varied in order to fine tune the performance of the algorithm. In particular, there are choices for: the determinant collapse parameter δ, the frame size reduction parameter κ and the sufficient descent reduction parameter ν.

Initially the default settings tried were:

$$\delta = 10^{-10}, \quad \kappa = 0.50, \quad \nu = 1.25 \tag{5.26}$$

These values were chosen simply by a best guess at what seemed a reasonable place to start — certainly not as the result of any detailed analysis! In order to decide the final values for these parameters, the performance of the algorithm with different parameter choices was investigated. All of the parameter values considered are listed in table 5.1. Appendix F summarises the performance results for each choice of parameter.

The most successful version of NM4 uses

$$\psi = \text{QR}, \quad \delta = 10^{-18}, \quad \kappa = 0.25, \quad \nu = 4.50 \tag{5.27}$$

In order to differentiate between the different algorithm versions, the following naming notation is used: the main algorithm is either NM1, NM2, NM3 or NM4 and then each of the parameter values is given after an underscore character. So the most successful variant of the Nelder-Mead algorithm produced, NM4_$\psi_3\delta_6\kappa_1\nu_8$ is the NM4 algorithm using the QR simplex reshape method (ψ_3), determinant parameter $\delta = 10^{-18}$, frame size reduction parameter $\kappa = 0.25$ and the sufficient descent reduction parameter $\nu = 4.50$.

The numerical results for the performance of this algorithm compared with FMINSEARCH using the high tolerance terminating criteria is reproduced in the next section. A complete set of comparison results is given in appendix B on page 91. Also, a comparison of the performance of NM4_$\psi_3\delta_6\kappa_1\nu_8$ and FMINSEARCH for M^{c}Kinnon's example, with the appropriate initial simplex is given in appendix B.2. Finally, a comparison of the performance of NM4_$\psi_3\delta_6\kappa_1\nu_8$ and FMINSEARCH for the standard quadratic with dimensions from two through to 100 is given in appendix C.

ψ	δ	κ	ν
$\psi_1 = \text{HH}$	$\delta_1 = 10^{-5}$	$\kappa_1 = 0.25$	$\nu_1 = 1.25$
$\psi_2 = \text{IJK}$	$\delta_2 = 10^{-8}$	$\kappa_2 = 0.50$	$\nu_2 = 1.50$
$\psi_3 = \text{QR}$	$\delta_3 = 10^{-10}$	$\kappa_3 = 0.75$	$\nu_3 = 2.00$
	$\delta_5 = 10^{-15}$		$\nu_5 = 3.00$
	$\delta_6 = 10^{-18}$		$\nu_6 = 3.50$
	$\delta_7 = 10^{-20}$		$\nu_7 = 4.00$
	$\delta_8 = 10^{-25}$		$\nu_8 = 4.50$
	$\delta_9 = 10^{-30}$		$\nu_9 = 5.00$

Table 5.1: Parameter options for variant four.

5.7.3 Numerical results

This section contains the numerical results for FMINSEARCH and the recommended variant NM4_$\psi_3\delta_6\kappa_1\nu_8$ on the suite of test functions using the high tolerance stopping criteria. The symbols (*) and (†) are used with the following meanings:

* Maximum number of function evaluations (100000) reached before the stopping criteria were met,

† Algorithm failed to produce an acceptable approximation to the solution.

The heading *FE* represents the number of function evaluations that were required to meet the stopping criteria.

	FMINSEARCH		$\texttt{NM4\_}\psi_3\delta_6\kappa_1\nu_8$	
Function	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	219	$1.09909e-18$	285	$1.39058e-17$
Freudenstein and Roth 2-d	172	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	754	$1.11069e-25$	969	$4.23980e-25$
Brown badly scaled 2-d	335	$7.03868e-18$	498	$7.99797e-17$
Beale 2-d	162	$6.11428e-18$	191	$2.07825e-18$
Jennrich and Sampson 2-d	133	$1.24362e+02$	157	$1.24362e+02$
M^{c}Kinnon 2-d	290	$-2.50000e-01$	426	$-2.50000e-01$
Helical valley 3-d	428	$4.78479e-17$	342	$9.83210e-16$
Bard 3-d	$^{\star}100004$	$1.74287e+01$	1134	$1.74287e+01$
Gaussian 3-d	216	$1.12793e-08$	194	$1.12793e-08$
Meyer 3-d	$^{\star}100004$	$8.79459e+01$	2801	$8.79459e+01$
Gulf research 3-d	687	$1.13899e-22$	529	$5.44511e-19$
Box 3-d	701	$3.05741e-22$	478	$8.70459e-21$
Powell singular 4-d	956	$3.56353e-28$	1045	$6.73509e-26$
Wood 4-d	572	$1.56392e-17$	656	$2.57400e-16$
Kowalik and Osbourne 4-d	398	$3.07506e-04$	653	$3.07506e-04$
Brown and Dennis 4-d	$^{\star}100001$	$8.58222e+04$	603	$8.58222e+04$
Quadratic 4-d	326	$4.52859e-17$	440	$2.15350e-17$
Penalty (1) 4-d	1371	$2.24998e-05$	1848	$2.24998e-05$
Penalty (2) 4-d	3730	$9.37629e-06$	4689	$9.37629e-06$
Osbourne (1) 5-d	1098	$5.46489e-05$	1488	$5.46489e-05$
Brown almost linear 5-d	782	$1.45905e-18$	648	$1.08728e-18$
Biggs EXP6 6-d	1130	$5.65565e-03$	4390	$1.16131e-20$
Extended Rosenbrock 6-d	7015	$2.79071e-17$	3110	$1.35844e-14$
Brown almost-linear 7-d	1819	$9.72059e-18$	1539	$1.51163e-17$
Quadratic 8-d	1519	$2.93256e-16$	1002	$8.07477e-17$
Extended Rosenbrock 8-d	5958	$^{\dagger}6.66424e-01$	5314	$3.27909e-17$
Variably dimensional 8-d	3780	$2.08479e-16$	2563	$1.24784e-15$
Extended Powell 8-d	2513	$^{\dagger}5.13165e-07$	7200	$6.43822e-24$
Watson 9-d	3229	$^{\dagger}3.98475e-03$	5256	$1.39976e-06$
Extended Rosenbrock 10-d	6684	$^{\dagger}9.72338e+00$	7629	$2.22125e-16$
Penalty (1) 10-d	5479	$^{\dagger}7.56754e-05$	9200	$7.08765e-05$
Penalty (2) 10-d	6783	$^{\dagger}2.97789e-04$	32768	$2.93661e-04$
Trigonometric 10-d	3105	$2.79506e-05$	2466	$2.79506e-05$
Osbourne (2) 11-d	4926	$4.01377e-02$	6416	$4.01377e-02$
Extended Powell 12-d	6607	$^{\dagger}5.52519e-06$	20076	$1.11105e-20$
Quadratic 16-d	8543	$7.70363e-16$	2352	$1.41547e-16$
Quadratic 24-d	$^{\star}100000$	$^{\dagger}5.04216e-01$	4766	$1.21730e-15$

Table 5.2: Comparison of the performance of FMINSEARCH and $\texttt{NM4\_}\psi_3\delta_1\kappa_6\nu_8$ with high tolerance stopping criteria.

Some results from table 5.2 worthy of note are discussed below.

(a) FMINSEARCH fails to produce accurate approximations to the minima of the following functions: Extended Rosenbrock 8-d, Watson 9-d, Extended Rosenbrock 10-d, Penalty (1) 10-d, Penalty (2) 10-d, Extended Powell 12-d, and the 24-d standard quadratic.

(b) FMINSEARCH fails to reach the terminating criteria in less than 100000 function evaluations for the functions: Bard 3-d, Meyer 3-d, Brown and Dennis 4-d, and the 24-d standard quadratic.

(c) Although FMINSEARCH requires 1130 function evaluations to terminate with the Biggs EXP6 6-d function compared to the 4390 function evaluations required by NM4_$\psi_3\delta_6\kappa_1\nu_8$, FMINSEARCH locates the stationary point with function value 5.65565×10^{-3} rather than the function minimum (zero), located by NM4_$\psi_3\delta_6\kappa_1\nu_8$.

(d) For the functions where both FMINSEARCH and NM4_$\psi_3\delta_6\kappa_1\nu_8$ reach the terminating criteria successfully and produce accurate approximations to the function minima, both algorithms require a comparable number of function evaluations.

Even with the appropriate choice of initial simplex, NM4_$\psi_3\delta_6\kappa_1\nu_8$ does not fail on M^{c}Kinnon's example as FMINSEARCH does. A comparison of the numerical results for NM4_$\psi_3\delta_6\kappa_1\nu_8$ and FMINSEARCH with M^{c}Kinnon's example can be found in appendix B.2.

Across all of the functions in the test suite, NM4_$\psi_3\delta_6\kappa_1\nu_8$ performs just as well as FMINSEARCH when FMINSEARCH is performing well, and considerably better than FMINSEARCH when FMINSEARCH is not performing well. In addition, NM4_$\psi_3\delta_6\kappa_1\nu_8$, is a provably convergent algorithm that successfully met the terminating criteria *and* produced accurate approximations to the minima of the test functions for *every* function in the test suite.

Chapter 6

Summary

A general introduction to optimisation has been presented along with a thorough discussion of both the original and standard Nelder-Mead algorithms. The Nelder-Mead algorithm has been in widespread use, often without thought to its convergence (or lack of) properties. Some of the failings of the Nelder-Mead algorithm have been exposed, and some functions have been presented that highlight the need for caution when using the Nelder-Mead algorithm.

Several variants of the Nelder-Mead algorithm have been presented, all of which keep the spirit of the Nelder-Mead algorithm. The recommended variant, NM4_$\psi_3\delta_6\kappa_1\nu_8$, is a provably convergent algorithm that performs well on every function in the test suite and so seems to be robust. It performs at least as well as FMINSEARCH in most situations and considerably better in some situations. Additionally, it avoids the failings of the Nelder-Mead algorithm (with the appropriate choice of initial simplex) for M^{c}Kinnon's example, and, performs vastly better than FMINSEARCH on the higher dimensional standard quadratic function.

In conclusion, the recommended variant performs well on all of the functions in the test suite and avoids the known problems of FMINSEARCH. In addition, the original aims of the project; to develop a provably convergent variant of the Nelder-Mead algorithm that maintains the good performance of the Nelder-Mead algorithm (when it performs well) and corrects the poor performance of the Nelder-Mead algorithm (when it does not perform well), all while keeping the spirit of the Nelder-Mead algorithm, have been met.

6.1 Further work

The results presented in this thesis are by no means exhaustive, or complete. There are many areas where further research may produce an algorithm that performs better. Some possibilities are listed below.

(a) Only three methods of reshaping a collapsed simplex have been investigated. It is possible that a thorough investigation may reveal a better method of reshaping a collapsed simplex.

(b) If a better choice of positive basis is made then the performance of the algorithm should improve. Perhaps using one fixed positive basis should be avoided.

(c) It may be beneficial to avoid a fixed criterion for determining whether the current simplex has collapsed or not. A variable collapse parameter may allow for improved performance by tuning the allowable simplex shapes to the function being optimised.

(d) As mentioned in chapter 4, it should be possible to develop an algorithm that uses a more dynamic method of adjusting the sufficient descent parameter. Presumably this could also be tuned to match the function being optimised.

(e) Perhaps an entirely different algorithm would be more suited to adaptation by the frame based template discussed in this thesis.

(f) Although the recommended variant presented in this thesis performs well on all of the functions in the test suite, it is still to some extent, a work in progress. At the very least, the MATLAB code for the algorithm could certainly be improved.

6.2 Acknowledgements

I would like to thank both of my research supervisors, Dr. Chris Price and Dr. Ian Coope for their guidance during this project.

Appendix A

The suite of test functions

This appendix lists the test functions that were used to determine the performance of the variants of the Nelder-Mead algorithm. Most of these functions were selected from the article by Moré, Garbow and Hillstrom [23]. In addition, the standard quadratic in dimensions 4, 8, 16 and 24, and M^{c}Kinnon's example have been included.

Function	*Initial point*	*Minimiser*	*Minimum*
Rosenbrock 2-d	$(-1.2, 0)$	$(1, 1)$	0
Freudenstein and Roth 2-d	$(0.5, -2)$	$(5, 4)$	0 [a]
Powell badly scaled 2-d	$(0, 1)$	$(1.098\ldots, 9.106\ldots)$	0
Brown badly scaled 2-d	$(1, 1)$	$(10^6, 2 \times 10^{-6})$	0
Beale 2-d	$(1, 1)$	$(3, 0.5)$	0
Jennrich and Sampson 2-d	$(0.3, 0.4)$		$124.362\ldots$
Helical valley 3-d	$(-1, 0, 0)$	$(1, 0, 0)$	0
Bard 3-d	$(1, 1, 1)$		$8.21487\ldots \times 10^{-3}$ [b]
Gaussian 3-d	$(0.4, 1, 0)$		$1.12793\ldots \times 10^{-8}$
Meyer 3-d	$(0.02, 4000, 250)$		$87.9458\ldots$
Gulf research 3-d	$(5, 2.5, 0.15)$	$(50, 25, 1.5)$	0
Box 3-d	$(0, 10, 20)$		0
Powell singular 4-d	$(3, -1, 0, 1)$	$(0, 0, 0, 0)$	0
Wood 4-d	$(-3, -1, -3, -1)$	$(1, 1, 1, 1)$	0
Kowalik and Osbourne 4-d	$(0.25, 0.39, 0.415, 0.39)$		$3.07505\ldots \times 10^{-4}$ [c]
Brown and Dennis 4-d	$(25, 5, -5, -1)$		$85822.2\ldots$
Osbourne (1) 5-d	$(0.5, 1.5, -1, 0.01, 0.02)$		$5.46489\ldots$
Biggs EXP6 6-d	$(1, 2, 1, 1, 1, 1)$	$(1, 10, 1, 5, 4, 3)$	0 [d]
Extended Rosenbrock 6-d	$(-1.2, 1, -1.2, 1, -1.2, 1)$	$(1, 1, \ldots, 1)$	0
Brown almost linear 7-d	$(0.5, 0.5, \ldots, 0.5)$		0
Quadratic 8-d	$(2, 1, 1, \ldots, 1)$	$(0, 0, \ldots, 0)$	0
Extended Rosenbrock 8-d	$(-1.2, 1, -1.2, 1, -1.2, 1, -1.2, 1)$	$(1, 1, \ldots, 1)$	0
Variably dimensional 8-d	$(\frac{7}{8}, \frac{6}{8}, \frac{5}{8}, \ldots, 0)$	$(1, 1, \ldots, 1)$	0
Extended Powell 8-d	$(3, -1, 0, 1, 3, -1, 0, 1)$	$(0, 0, \ldots, 0)$	0

[a] $f = 49.9842\ldots$ at $(11.41\ldots, -0.8968\ldots)$
[b] $f = 17.4286\ldots$ at $(0.8406\ldots, -\infty, -\infty)$
[c] $f = 1.02734\ldots \times 10^{-3}$ at $(+\infty, -14.07\ldots, -\infty, -\infty)$
[d] $f = 5.65565\ldots \times 10^{-3}$

Function	*Initial point*	*Minimiser*	*Minimum*
Watson 9-d	$(0, 0, \ldots, 0)$		$1.39976\ldots \times 10^{-6}$
Extended Rosenbrock 10-d	$(-1.2, 1, -1.2, 1, -1.2, 1, -1.2, 1, -1.2, 1)$	$(1, 1, \ldots, 1)$	0
Penalty (1) 10-d	$(1, 2, 3, \ldots, 10)$		$7.08765\ldots \times 10^{-5}$
Penalty (2) 10-d	$(0.5, 0.5, \ldots, 0.5)$		$2.93660\ldots \times 10^{-4}$
Trigonometric 10-d	$(0.1, 0.1, \ldots, 0.1)$		0
Osbourne (2) 11-d	$(1, 3, 0, 65, 0, 65, 0.7, 0.6, 3, 5, 7, 2, 4.5, 5.5)$		$4.01377\ldots \times 10^{-2}$
Extended Powell 12-d	$(3, -1, 0, 1, 3, -1, 0, 1, 3, -1, 0, 1)$	$(0, 0, \ldots, 0)$	0
Quadratic 16-d	$(2, 1, 1, \ldots, 1)$	$(0, 0, \ldots, 0)$	0
Quadratic 24-d	$(2, 1, 1, \ldots, 1)$	$(0, 0, \ldots, 0)$	0

Table A.1: The suite of test functions

Appendix B

Numerical results

This appendix contains the numerical results for FMINSEARCH and NM4_$\psi_3\delta_6\kappa_1\nu_8$ on the suite of test functions and M[c]Kinnon's example with the appropriate initial simplex. Both low and high tolerance results are included. The symbols (*) and (†) are used with the following meanings:

* Maximum number of function evaluations (100000) reached before the stopping criteria were met,

† Algorithm failed to produce an acceptable approximation to the solution

The heading *FE* represents the number of function evaluations that were required to meet the stopping criteria.

B.1 Results for the suite of test functions

B.1.1 Low tolerance results

	FMINSEARCH		NM4_$\psi_3\delta_6\kappa_1\nu_8$	
Function	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	159	$8.17766e-10$	218	$9.85723e-10$
Freudenstein and Roth 2-d	120	$4.89843e+01$	148	$4.89843e+01$
Powell badly scaled 2-d	700	$1.42227e-17$	770	$4.68385e-11$
Brown badly scaled 2-d	275	$2.00356e-09$	393	$5.14292e-09$
Beale 2-d	107	$1.39263e-10$	121	$1.70860e-10$
Jennrich and Sampson 2-d	72	$1.24362e+02$	92	$1.24362e+02$
M^{c}Kinnon 2-d	165	$-2.50000e-01$	23	$^\dagger-6.24961e-05$
Helical valley 3-d	153	$1.67150e-04$	137	$9.80281e-05$
Bard 3-d	*100004	$1.74287e+01$	1099	$1.74287e+01$
Gaussian 3-d	62	$^\dagger1.18892e-08$	61	$^\dagger1.29096e-08$
Meyer 3-d	1781	$8.79459e+01$	2632	$8.79459e+01$
Gulf research 3-d	578	$2.02284e-13$	437	$7.57613e-14$
Box 3-d	242	$5.71916e-04$	355	$2.18028e-12$
Powell singular 4-d	305	$1.39059e-06$	358	$1.67017e-11$
Wood 4-d	400	$3.80143e-09$	389	$2.58277e-07$
Kowalik and Osbourne 4-d	260	$3.07506e-04$	484	$3.07506e-04$
Brown and Dennis 4-d	333	$8.58222e+04$	418	$8.58222e+04$
Quadratic 4-d	204	$2.06937e-09$	283	$6.45079e-09$
Penalty (1) 4-d	583	$^\dagger2.35458e-05$	1652	$2.24998e-05$
Penalty (2) 4-d	2726	$^\dagger9.38054e-06$	187	$^\dagger1.02820e-05$
Osbourne (1) 5-d	904	$5.46489e-05$	215	$^\dagger7.22121e-05$
Brown almost linear 5-d	614	$3.55233e-10$	364	$1.92843e-10$
Biggs EXP6 6-d	916	$5.65565e-03$	3702	$3.31156e-12$
Extended Rosenbrock 6-d	2141	$^\dagger2.13141e+00$	2839	$2.95708e-08$
Brown almost-linear 7-d	808	$2.26780e-06$	623	$1.55045e-08$
Quadratic 8-d	1050	$1.64029e-08$	630	$5.34118e-09$
Extended Rosenbrock 8-d	3439	$^\dagger9.94743e-01$	4168	$3.92654e-09$
Variably dimensional 8-d	1786	$^\dagger1.54617e+00$	1365	$4.12240e-08$
Extended Powell 8-d	1006	$7.14391e-07$	2333	$2.24110e-10$
Watson 9-d	1766	$^\dagger7.90568e-03$	3233	$^\dagger9.24202e-06$
Extended Rosenbrock 10-d	5295	$^\dagger9.74269e+00$	5703	$7.95335e-09$
Penalty (1) 10-d	3909	$^\dagger7.57248e-05$	6754	$7.08775e-05$
Penalty (2) 10-d	4017	$^\dagger2.97871e-04$	1578	$^\dagger2.97345e-04$
Trigonometric 10-d	2243	$2.79608e-05$	1676	$2.79539e-05$
Osbourne (2) 11-d	3827	$4.01377e-02$	5153	$4.01378e-02$
Extended Powell 12-d	2791	$9.52302e-06$	5016	$1.35382e-06$
Quadratic 16-d	6244	$2.20200e-07$	1248	$2.26014e-08$
Quadratic 24-d	58526	$^\dagger5.14700e-01$	3094	$2.43870e-08$

Table B.1: Comparison of the performance of FMINSEARCH and NM4_$\psi_3\delta_1\kappa_6\nu_8$ with low tolerance stopping criteria.

B.1.2 High tolerance results

	FMINSEARCH		NM4_$\psi_3\delta_6\kappa_1\nu_8$	
Function	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	219	$1.09909e-18$	285	$1.39058e-17$
Freudenstein and Roth 2-d	172	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	754	$1.11069e-25$	969	$4.23980e-25$
Brown badly scaled 2-d	335	$7.03868e-18$	498	$7.99797e-17$
Beale 2-d	162	$6.11428e-18$	191	$2.07825e-18$
Jennrich and Sampson 2-d	133	$1.24362e+02$	157	$1.24362e+02$
M^{c}Kinnon 2-d	290	$-2.50000e-01$	426	$-2.50000e-01$
Helical valley 3-d	428	$4.78479e-17$	342	$9.83210e-16$
Bard 3-d	*100004	$1.74287e+01$	1134	$1.74287e+01$
Gaussian 3-d	216	$1.12793e-08$	194	$1.12793e-08$
Meyer 3-d	*100004	$8.79459e+01$	2801	$8.79459e+01$
Gulf research 3-d	687	$1.13899e-22$	529	$5.44511e-19$
Box 3-d	701	$3.05741e-22$	478	$8.70459e-21$
Powell singular 4-d	956	$3.56353e-28$	1045	$6.73509e-26$
Wood 4-d	572	$1.56392e-17$	656	$2.57400e-16$
Kowalik and Osbourne 4-d	398	$3.07506e-04$	653	$3.07506e-04$
Brown and Dennis 4-d	*100001	$8.58222e+04$	603	$8.58222e+04$
Quadratic 4-d	326	$4.52859e-17$	440	$2.15350e-17$
Penalty (1) 4-d	1371	$2.24998e-05$	1848	$2.24998e-05$
Penalty (2) 4-d	3730	$9.37629e-06$	4689	$9.37629e-06$
Osbourne (1) 5-d	1098	$5.46489e-05$	1488	$5.46489e-05$
Brown almost linear 5-d	782	$1.45905e-18$	648	$1.08728e-18$
Biggs EXP6 6-d	1130	$5.65565e-03$	4390	$1.16131e-20$
Extended Rosenbrock 6-d	7015	$2.79071e-17$	3110	$1.35844e-14$
Brown almost-linear 7-d	1819	$9.72059e-18$	1539	$1.51163e-17$
Quadratic 8-d	1519	$2.93256e-16$	1002	$8.07477e-17$
Extended Rosenbrock 8-d	5958	$^\dagger 6.66424e-01$	5314	$3.27909e-17$
Variably dimensional 8-d	3780	$2.08479e-16$	2563	$1.24784e-15$
Extended Powell 8-d	2513	$^\dagger 5.13165e-07$	7200	$6.43822e-24$
Watson 9-d	3229	$^\dagger 3.98475e-03$	5256	$1.39976e-06$
Extended Rosenbrock 10-d	6684	$^\dagger 9.72338e+00$	7629	$2.22125e-16$
Penalty (1) 10-d	5479	$^\dagger 7.56754e-05$	9200	$7.08765e-05$
Penalty (2) 10-d	6783	$^\dagger 2.97789e-04$	32768	$2.93661e-04$
Trigonometric 10-d	3105	$2.79506e-05$	2466	$2.79506e-05$
Osbourne (2) 11-d	4926	$4.01377e-02$	6416	$4.01377e-02$
Extended Powell 12-d	6607	$^\dagger 5.52519e-06$	20076	$1.11105e-20$
Quadratic 16-d	8543	$7.70363e-16$	2352	$1.41547e-16$
Quadratic 24-d	*100000	$^\dagger 5.04216e-01$	4766	$1.21730e-15$

Table B.2: Comparison of the performance of FMINSEARCH and NM4_$\psi_3\delta_1\kappa_6\nu_8$ with high tolerance stopping criteria.

B.2 Results for M^{c}Kinnon's example

The results presented here all use M^{c}Kinnon's initial simplex configuration.

B.2.1 Low tolerance results

	FMINSEARCH		NM4_$s_3\delta_6\kappa_1\beta_8$	
Function	*FE*	*Minimum*	*FE*	*Minimum*
M^{c}Kinnon 2-d	143	$^\dagger 0.00000e+00$	72	$^\dagger -2.53997e-05$

Table B.3: Comparison of the performance of FMINSEARCH and NM4_$\psi_3\delta_1\kappa_6\nu_8$ with low tolerance stopping criteria on M^{c}Kinnon's example.

B.2.2 High tolerance results

	FMINSEARCH		NM4_$\psi_3\delta_6\kappa_1\nu_8$	
Function	*FE*	*Minimum*	*FE*	*Minimum*
M^{c}Kinnon 2-d	359	$^\dagger 0.00000e+00$	351	$-2.50000e-01$

Table B.4: Comparison of the performance of FMINSEARCH and NM4_$\psi_3\delta_1\kappa_6\nu_8$ with high tolerance stopping criteria on M^{c}Kinnon's example.

Appendix C

The standard quadratic

This appendix contains a summary of the results for both the standard Nelder-Mead algorithm and the recommended variant NM4_$\psi_3\delta_6\kappa_1\nu_8$ on the function $f(\mathrm{x}) = \mathrm{x}^T\mathrm{x}$ for dimensions ranging from two through to 100 using both low and high tolerance stopping criteria. For both sets of tolerances the vertical scale for the graphs of the number of function evaluations has been kept the same for ease of comparison. Due to the differences in the results obtained for the values of the function minima the vertical scales for these graphs could not be kept the same. The results produced by FMINSEARCH are up to about 10^6 times worse than the results produced by NM4_$\psi_3\delta_6\kappa_1\nu_8$ with low tolerance stopping criteria and up to about 10^{14} times worse with high tolerance stopping criteria. For both sets of tolerances, FMINSEARCH required many more function evaluations than NM4_$\psi_3\delta_6\kappa_1\nu_8$ to reach the stopping criteria.

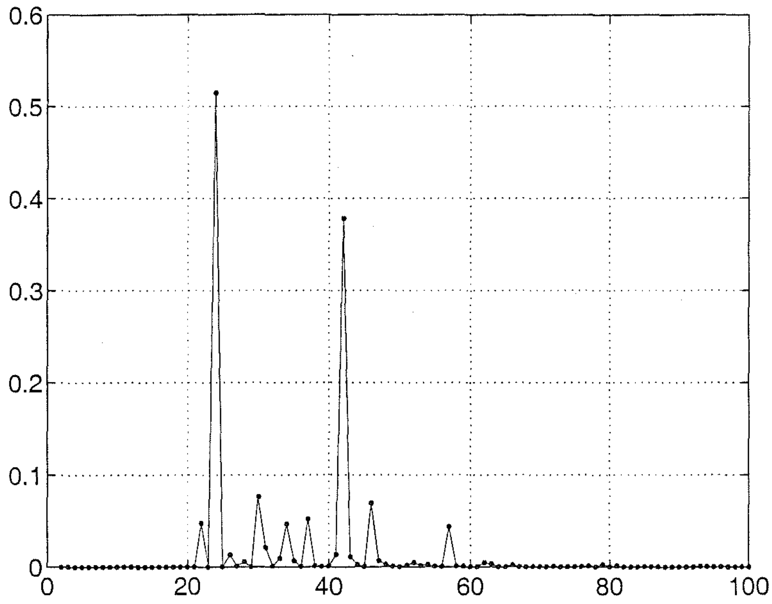


Figure C.1: The minima found by FMINSEARCH for the standard quadratic with low tolerance stopping criteria.

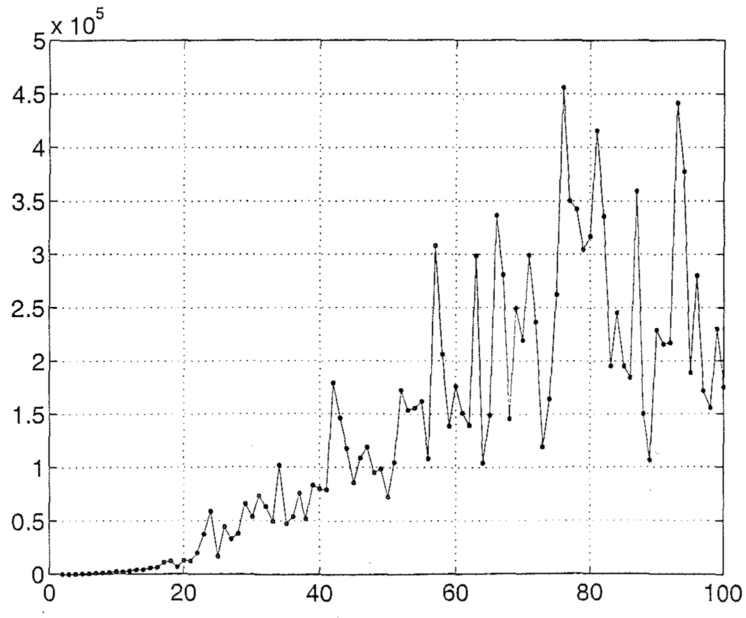


Figure C.2: The number of function evaluations required by FMINSEARCH before terminating with low tolerance stopping criteria.

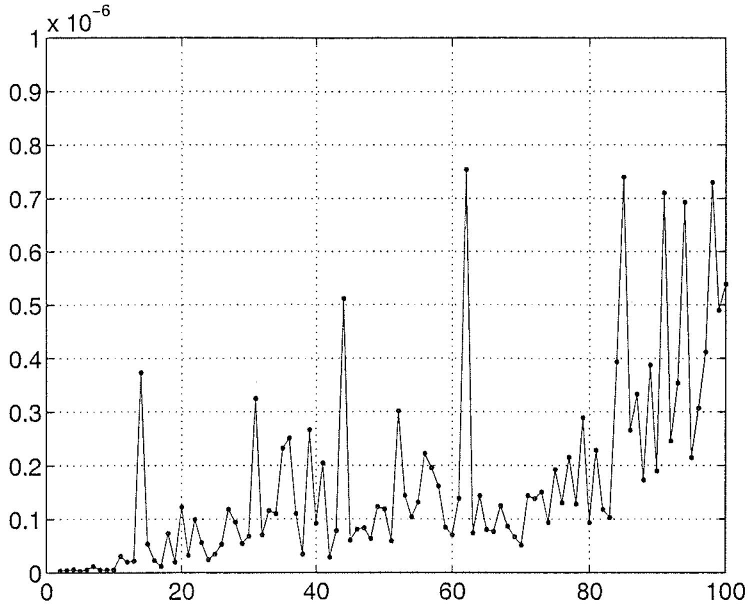


Figure C.3: The minima found by NM4_$\psi_3\delta_6\kappa_1\nu_8$ for the standard quadratic with low tolerance stopping criteria.

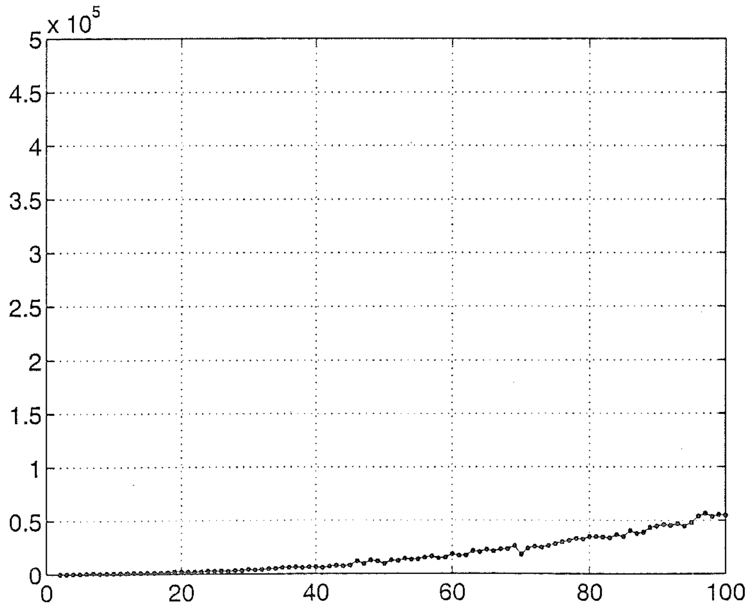


Figure C.4: The number of function evaluations required by NM4_$\psi_3\delta_6\kappa_1\nu_8$ before terminating with low tolerance stopping criteria. Note that the scale is the same as the graph in figure C.2.

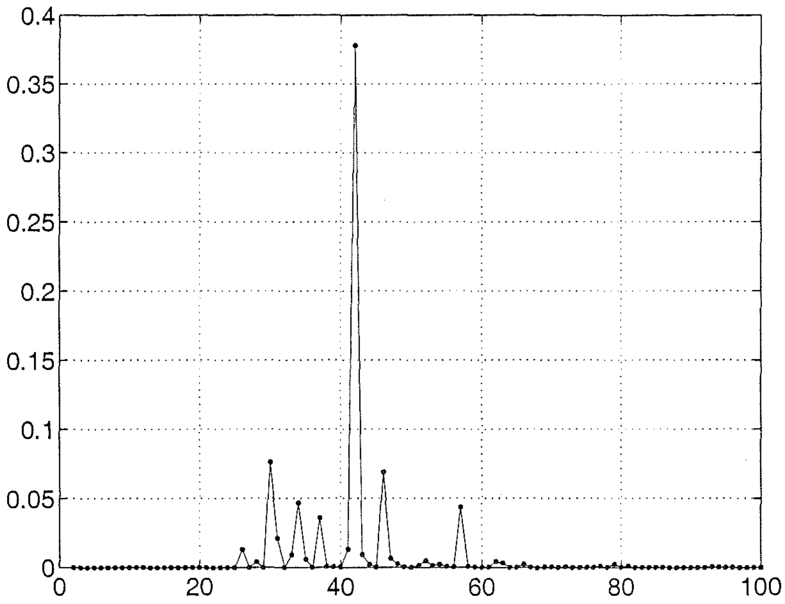


Figure C.5: The minima found by FMINSEARCH for the standard quadratic with high tolerance stopping criteria.

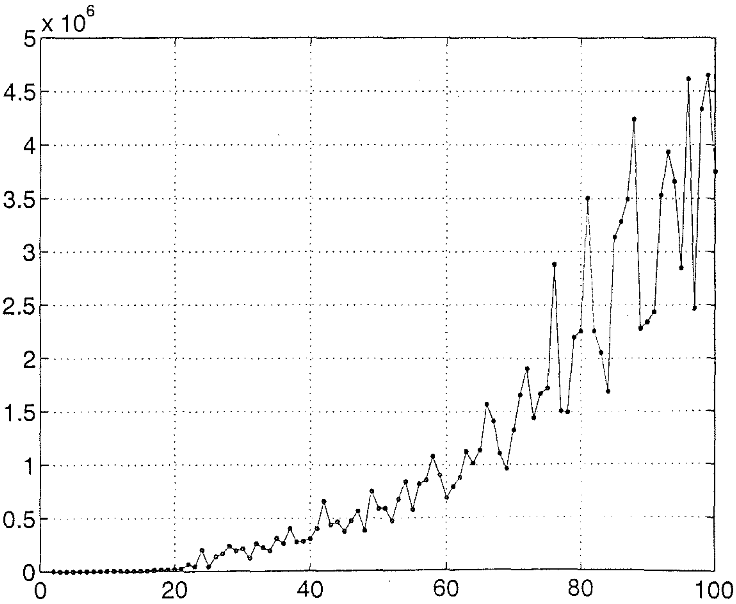


Figure C.6: The number of function evaluations required by FMINSEARCH before terminating with high tolerance stopping criteria.

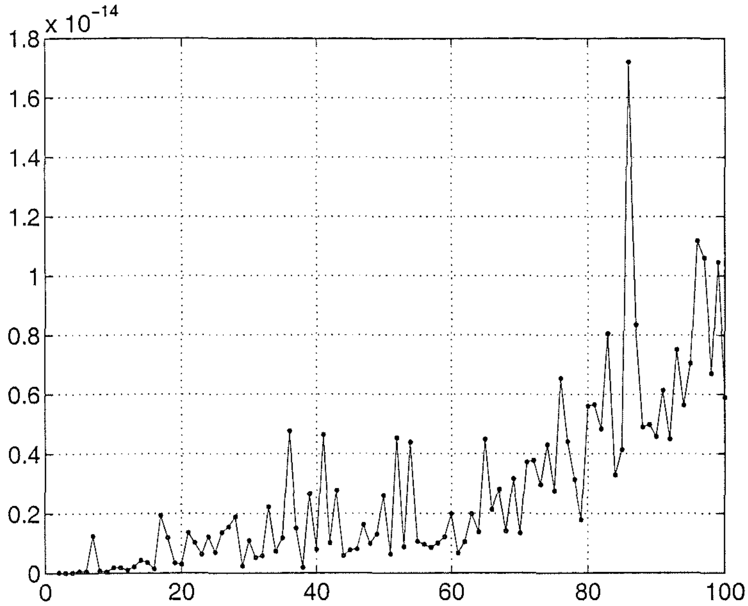


Figure C.7: The minima found by NM4_$\psi_3\delta_6\kappa_1\nu_8$ for the standard quadratic with high tolerance stopping criteria.

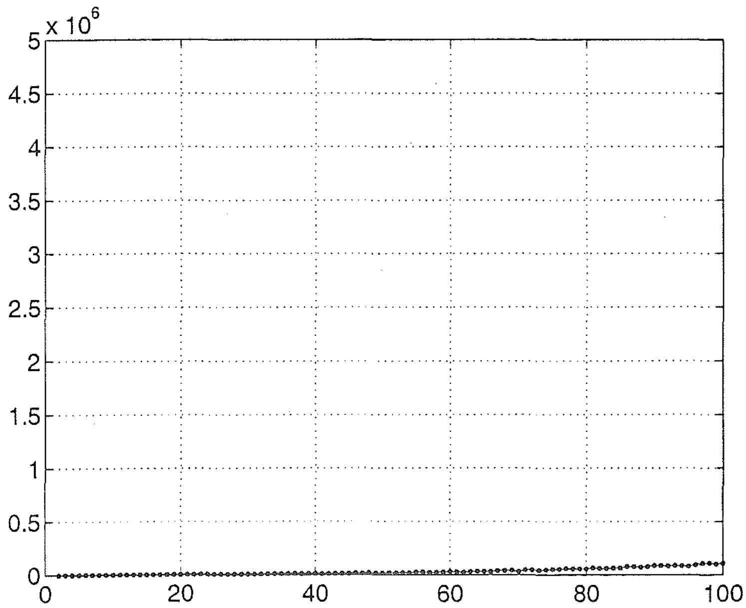


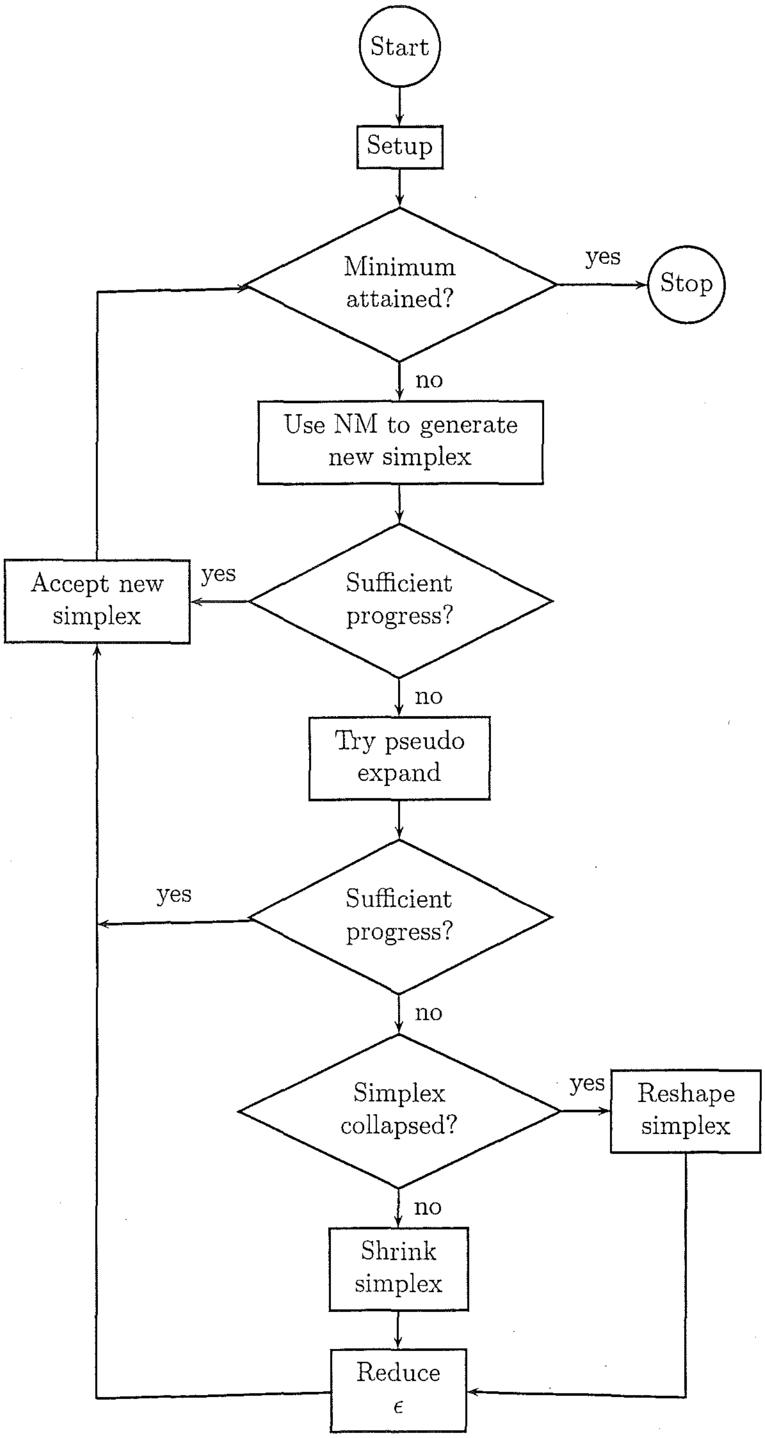
Figure C.8: The number of function evaluations required by NM4_$\psi_3\delta_6\kappa_1\nu_8$ before terminating with high tolerance stopping criteria. Note that the scale is the same as the graph in figure C.6.

Appendix D

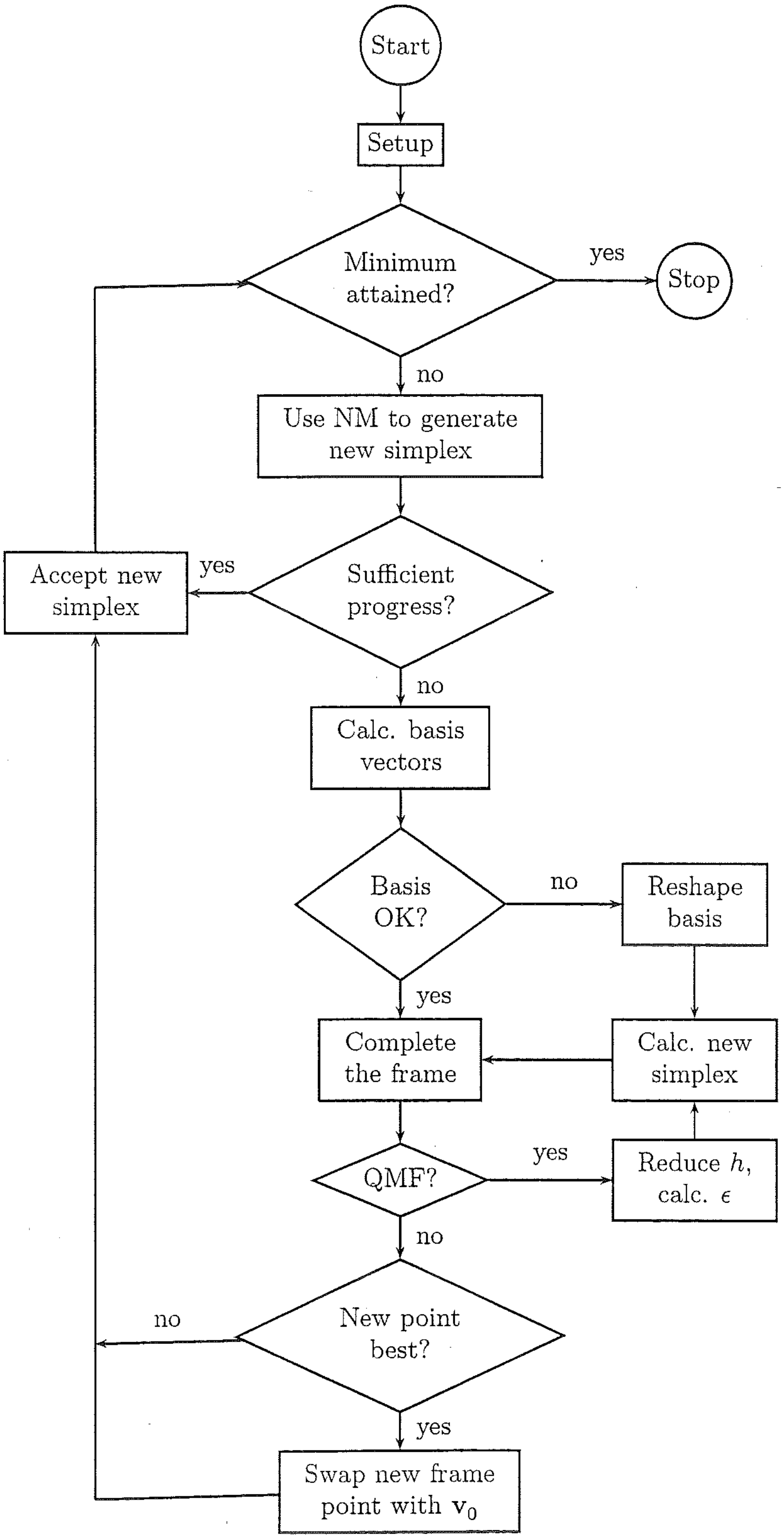
Schematic diagrams

This appendix contains schematic diagrams for each of the Nelder-Mead variants (`NM1-4`). They are included to aid understanding of the algorithms, rather than provide technical details.

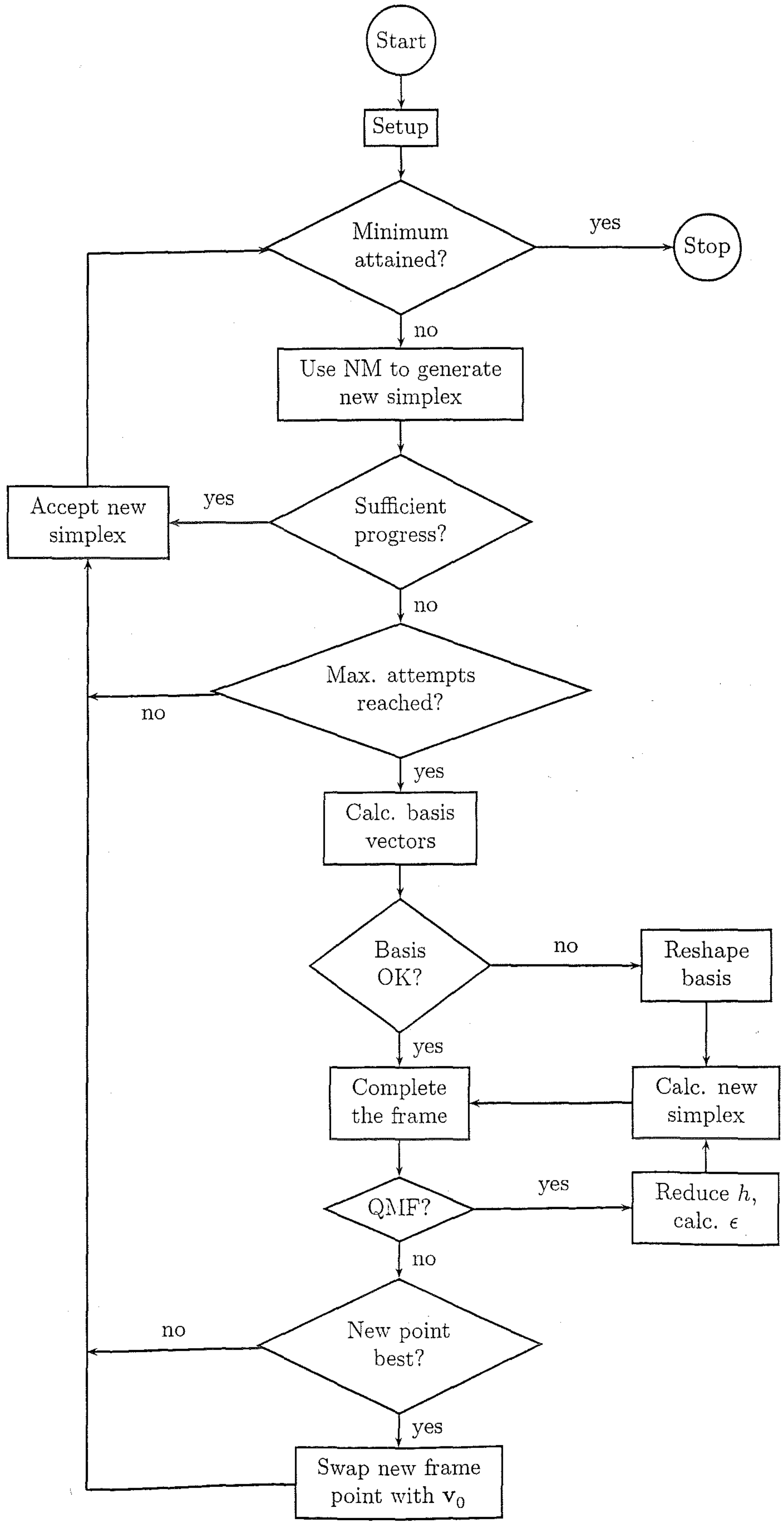
D.1 Schematic diagram for variant one



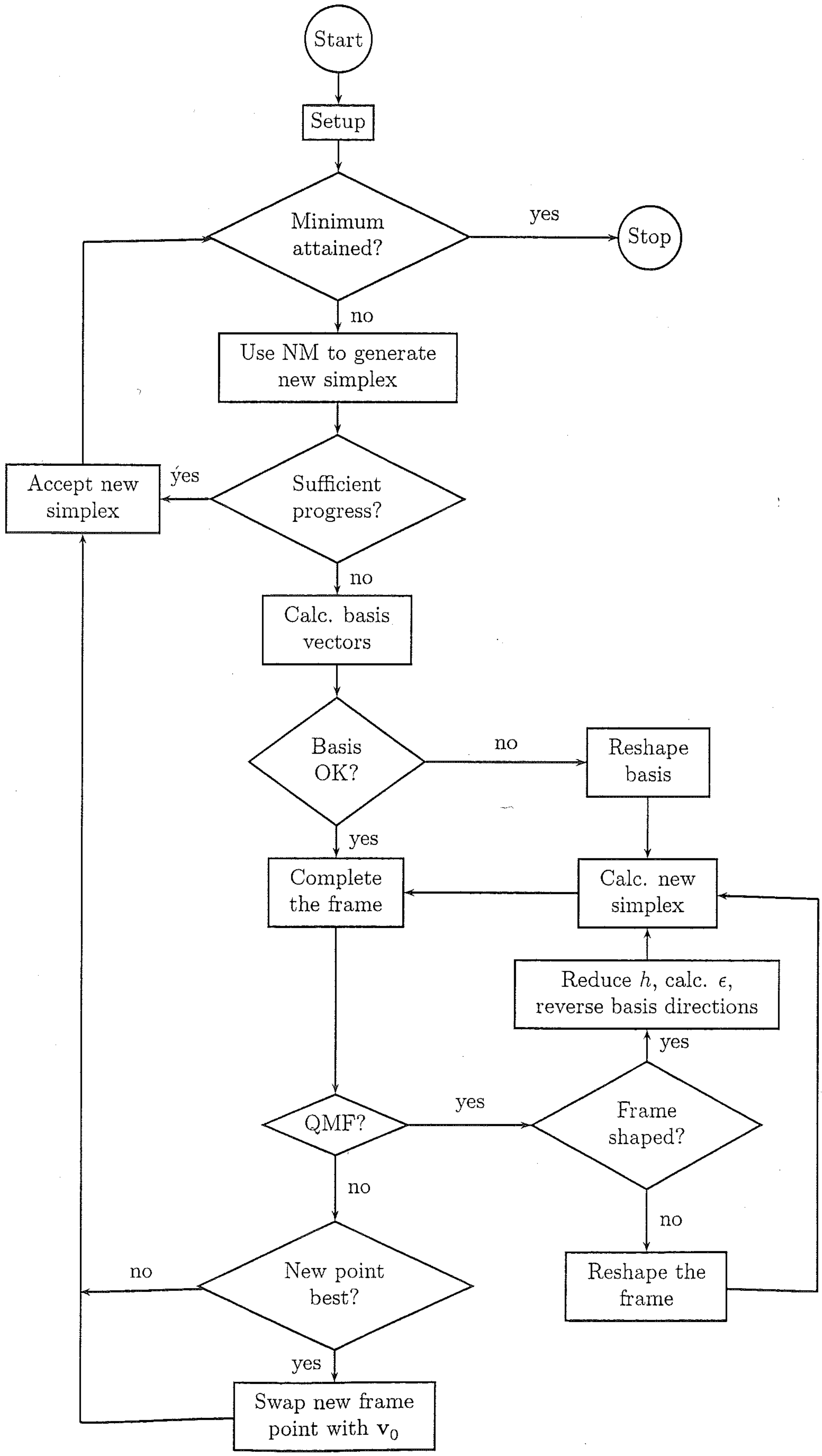
D.2 Schematic diagram for variant two



D.3 Schematic diagram for variant three



D.4 Schematic diagram for variant four



Appendix E

Performance of the variants

This appendix contains the numerical results for each of the Nelder-Mead variants (NM1-4) using the default parameters and each of the simplex reshape methods.

E.1 Variant one

E.1.1 Low tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	224	$7.43245e-10$	224	$7.43245e-10$	224	$7.43245e-10$
Freudenstein and Roth 2-d	135	$4.89843e+01$	135	$4.89843e+01$	135	$4.89843e+01$
Powell badly scaled 2-d	773	$1.88088e-17$	773	$1.88088e-17$	773	$1.88088e-17$
Brown badly scaled 2-d	295	$2.18815e-09$	295	$2.18815e-09$	295	$2.18815e-09$
Beale 2-d	111	$1.58601e-09$	111	$1.58601e-09$	111	$1.58601e-09$
Jennrich and Sampson 2-d	74	$1.24362e+02$	74	$1.24362e+02$	74	$1.24362e+02$
M^{c}Kinnon 2-d	33	$0.00000e+00$	33	$0.00000e+00$	33	$0.00000e+00$
Helical valley 3-d	168	$2.13050e-04$	168	$2.13050e-04$	168	$2.13050e-04$
Bard 3-d	869	$1.74287e+01$	6645	$1.74287e+01$	91979	$1.74287e+01$
Gaussian 3-d	56	$1.23004e-08$	56	$1.23004e-08$	56	$1.23004e-08$
Meyer 3-d	1757	$8.79459e+01$	1757	$8.79459e+01$	1757	$8.79459e+01$
Gulf research 3-d	550	$9.88737e-14$	550	$9.88737e-14$	550	$9.88737e-14$
Box 3-d	245	$5.71874e-04$	245	$5.71874e-04$	245	$5.71874e-04$
Powell singular 4-d	327	$1.38594e-07$	327	$1.38594e-07$	327	$1.38594e-07$
Wood 4-d	443	$3.64421e-09$	443	$3.64421e-09$	443	$3.64421e-09$
Kowalik and Osbourne 4-d	357	$4.32454e-04$	357	$4.32454e-04$	357	$4.32454e-04$
Brown and Dennis 4-d	326	$8.58222e+04$	326	$8.58222e+04$	326	$8.58222e+04$
Quadratic 4-d	251	$9.48272e-09$	251	$9.48272e-09$	251	$9.48272e-09$
Penalty (1) 4-d	990	$2.27028e-05$	990	$2.27028e-05$	990	$2.27028e-05$
Penalty (2) 4-d	196	$9.62687e-06$	196	$9.62687e-06$	196	$9.62687e-06$
Osbourne (1) 5-d	392	$7.12421e-05$	392	$7.12421e-05$	392	$7.12421e-05$
Brown almost linear 5-d	892	$2.97253e-10$	892	$2.97253e-10$	892	$2.97253e-10$
Biggs EXP6 6-d	767	$5.83603e-04$	767	$5.83603e-04$	767	$5.83603e-04$

Function	ψ_1 *FE*	ψ_1 *Minimum*	ψ_2 *FE*	ψ_2 *Minimum*	ψ_3 *FE*	ψ_3 *Minimum*
Extended Rosenbrock 6-d	2392	$2.07397e-09$	2339	$1.10809e-09$	1942	$2.64804e-03$
Brown almost-linear 7-d	1350	$1.27022e-03$	1350	$1.27022e-03$	1350	$1.27022e-03$
Quadratic 8-d	1660	$1.10073e-08$	1660	$1.10073e-08$	1660	$1.10073e-08$
Extended Rosenbrock 8-d	5270	$3.05156e-05$	6598	$3.22738e-09$	6597	$3.28019e-08$
Variably dimensional 8-d	1749	$3.46864e-08$	1577	$1.17666e-08$	1847	$6.61800e-01$
Extended Powell 8-d	857	$8.06237e-07$	843	$8.06758e-07$	1480	$1.15962e-10$
Watson 9-d	2677	$1.00325e-04$	2071	$4.70065e-05$	2693	$4.54656e-05$
Extended Rosenbrock 10-d	5214	$1.43961e-04$	9251	$2.78120e-08$	11945	$5.81612e-09$
Penalty (1) 10-d	8730	$7.09013e-05$	3556	$7.94711e-05$	8639	$7.24894e-05$
Penalty (2) 10-d	901	$2.98800e-04$	901	$2.98800e-04$	901	$2.98800e-04$
Trigonometric 10-d	2430	$4.47771e-07$	1730	$3.71101e-05$	2477	$1.68871e-04$
Osbourne (2) 11-d	9322	$3.29282e-01$	5363	$8.42018e-02$	11724	$3.29282e-01$
Extended Powell 12-d	3768	$1.64399e-07$	6248	$2.69872e-07$	8231	$7.05249e-07$
Quadratic 16-d	3067	$4.15995e-06$	5344	$3.62660e-03$	6605	$9.13077e-03$
Quadratic 24-d	11109	$3.90032e-06$	9874	$6.07956e-05$	9674	$7.95825e-07$

Table E.1: Low tolerance results for variant one

E.1.2 High tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	297	$3.59110e-18$	297	$3.59110e-18$	297	$3.59110e-18$
Freudenstein and Roth 2-d	206	$4.89843e+01$	206	$4.89843e+01$	206	$4.89843e+01$
Powell badly scaled 2-d	848	$3.28549e-26$	848	$3.28549e-26$	848	$3.28549e-26$
Brown badly scaled 2-d	365	$2.55082e-17$	365	$2.55082e-17$	365	$2.55082e-17$
Beale 2-d	179	$1.03538e-30$	179	$1.03538e-30$	179	$1.03538e-30$
Jennrich and Sampson 2-d	160	$1.24362e+02$	160	$1.24362e+02$	160	$1.24362e+02$
M^{c}Kinnon 2-d	484	$-2.50000e-01$	484	$-2.50000e-01$	484	$-2.50000e-01$
Helical valley 3-d	461	$2.40675e-16$	461	$2.40675e-16$	461	$2.40675e-16$
Bard 3-d	943	$1.74287e+01$	6645	$1.74287e+01$	91979	$1.74287e+01$
Gaussian 3-d	364	$1.12793e-08$	364	$1.12793e-08$	364	$1.12793e-08$
Meyer 3-d	2101	$8.79459e+01$	4941	$8.79459e+01$	4941	$8.79459e+01$
Gulf research 3-d	653	$4.28408e-22$	653	$4.28408e-22$	653	$4.28408e-22$
Box 3-d	751	$3.67126e-21$	751	$3.67126e-21$	751	$3.67126e-21$
Powell singular 4-d	1103	$9.15508e-26$	1126	$9.15508e-26$	1509	$2.51593e-28$
Wood 4-d	566	$4.91432e-17$	566	$4.91432e-17$	566	$4.91432e-17$
Kowalik and Osbourne 4-d	848	$3.07506e-04$	848	$3.07506e-04$	848	$3.07506e-04$
Brown and Dennis 4-d	638	$8.58222e+04$	638	$8.58222e+04$	638	$8.58222e+04$
Quadratic 4-d	371	$4.16618e-17$	371	$4.16618e-17$	371	$4.16618e-17$
Penalty (1) 4-d	1855	$2.24998e-05$	1855	$2.24998e-05$	1855	$2.24998e-05$
Penalty (2) 4-d	2916	$9.37629e-06$	2909	$9.37629e-06$	2729	$9.37629e-06$
Osbourne (1) 5-d	1340	$5.46489e-05$	1661	$5.46489e-05$	2135	$5.46489e-05$
Brown almost linear 5-d	1079	$1.21401e-17$	1079	$1.21401e-17$	1079	$1.21401e-17$
Biggs EXP6 6-d	3557	$1.03065e-20$	3916	$2.18806e-20$	3619	$1.12135e-20$
Extended Rosenbrock 6-d	2604	$4.71367e-17$	2567	$2.35905e-17$	2819	$1.13781e-17$
Brown almost-linear 7-d	2472	$4.45859e-18$	2472	$4.45859e-18$	2472	$4.45859e-18$

Function	ψ_1 FE	ψ_1 Minimum	ψ_2 FE	ψ_2 Minimum	ψ_3 FE	ψ_3 Minimum
Quadratic 8-d	2047	$9.07857e-16$	2047	$9.07857e-16$	2047	$9.07857e-16$
Extended Rosenbrock 8-d	6348	$4.46235e-17$	7111	$6.07486e-17$	7069	$7.16744e-17$
Variably dimensional 8-d	2359	$2.64464e-16$	2064	$5.97945e-17$	3413	$5.68830e-16$
Extended Powell 8-d	9216	$4.26368e-14$	2824	$3.49245e-13$	4756	$1.98094e-22$
Watson 9-d	16456	$1.39976e-06$	17746	$1.39976e-06$	7978	$1.39976e-06$
Extended Rosenbrock 10-d	9943	$1.67618e-12$	10086	$2.31194e-16$	12531	$1.37676e-16$
Penalty (1) 10-d	14050	$7.08765e-05$	9492	$7.08765e-05$	18858	$7.08765e-05$
Penalty (2) 10-d	38826	$2.93661e-04$	33183	$2.93661e-04$	100000	$2.93661e-04$
Trigonometric 10-d	3132	$4.47357e-07$	3685	$2.79506e-05$	10434	$2.97867e-05$
Osbourne (2) 11-d	15052	$3.29282e-01$	9760	$4.01377e-02$	15659	$3.29282e-01$
Extended Powell 12-d	11458	$1.54831e-12$	11413	$1.07786e-11$	40603	$2.21187e-23$
Quadratic 16-d	8929	$3.47784e-15$	15278	$5.93871e-15$	18987	$3.79038e-15$
Quadratic 24-d	98685	$8.67385e-09$	96384	$8.49944e-07$	32961	$2.98090e-10$

Table E.2: High tolerance results for variant one

E.2 Variant two

E.2.1 Low tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	298	$1.15624e-08$	298	$1.15624e-08$	298	$1.15624e-08$
Freudenstein and Roth 2-d	241	$4.89843e+01$	241	$4.89843e+01$	241	$4.89843e+01$
Powell badly scaled 2-d	879	$4.11100e-17$	879	$4.11100e-17$	879	$4.11100e-17$
Brown badly scaled 2-d	547	$6.50725e-10$	547	$6.50725e-10$	547	$6.50725e-10$
Beale 2-d	174	$3.26134e-08$	174	$3.26134e-08$	174	$3.26134e-08$
Jennrich and Sampson 2-d	145	$1.24370e+02$	145	$1.24370e+02$	145	$1.24370e+02$
M^{c}Kinnon 2-d	22	$0.00000e+00$	22	$0.00000e+00$	22	$0.00000e+00$
Helical valley 3-d	224	$3.45161e-06$	224	$3.45161e-06$	224	$3.45161e-06$
Bard 3-d	300	$1.74362e+01$	1434	$1.74287e+01$	302	$1.74362e+01$
Gaussian 3-d	55	$2.42522e-08$	55	$2.42522e-08$	55	$2.42522e-08$
Meyer 3-d	401	$8.59238e+04$	2969	$8.79459e+01$	3166	$8.79459e+01$
Gulf research 3-d	645	$1.71802e-10$	645	$1.71802e-10$	645	$1.71802e-10$
Box 3-d	93	$9.48272e-04$	93	$9.48272e-04$	93	$9.48272e-04$
Powell singular 4-d	436	$2.02573e-06$	576	$2.94747e-06$	418	$9.78737e+00$
Wood 4-d	526	$4.39477e-07$	526	$4.39477e-07$	526	$4.39477e-07$
Kowalik and Osbourne 4-d	478	$3.07506e-04$	478	$3.07506e-04$	478	$3.07506e-04$
Brown and Dennis 4-d	378	$9.96990e+04$	378	$9.96990e+04$	378	$9.96990e+04$
Quadratic 4-d	599	$1.43624e-03$	828	$2.56032e-09$	608	$2.70479e-09$
Penalty (1) 4-d	216	$2.84133e-05$	216	$2.84133e-05$	216	$2.84133e-05$
Penalty (2) 4-d	201	$9.92499e-06$	201	$9.92499e-06$	201	$9.92499e-06$
Osbourne (1) 5-d	307	$9.10858e-04$	307	$9.10858e-04$	307	$9.10858e-04$
Brown almost linear 5-d	329	$4.20431e-04$	388	$3.35245e-04$	401	$7.07527e-09$
Biggs EXP6 6-d	2319	$5.65565e-03$	1477	$6.47539e-03$	355	$3.05891e-01$

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Extended Rosenbrock 6-d	291	$1.28627e+01$	291	$1.28627e+01$	291	$1.28627e+01$
Brown almost-linear 7-d	251	$1.24236e-02$	1513	$1.61762e-06$	1688	$5.67590e-10$
Quadratic 8-d	405	$1.50850e-09$	585	$7.91524e-08$	2021	$2.27664e-08$
Extended Rosenbrock 8-d	4252	$2.48621e+00$	3722	$4.09137e-01$	4384	$4.60512e-03$
Variably dimensional 8-d	199	$9.44687e-01$	1949	$1.53458e-08$	229	$9.63874e+00$
Extended Powell 8-d	1403	$1.16131e-02$	1610	$4.16166e-05$	3688	$5.83429e-11$
Watson 9-d	1355	$9.06935e-05$	2096	$1.10001e-04$	1462	$9.42411e-04$
Extended Rosenbrock 10-d	9301	$2.44373e+00$	3125	$1.42859e+01$	8112	$2.15354e-04$
Penalty (1) 10-d	1530	$1.00402e-04$	3820	$7.09026e-05$	9563	$7.08848e-05$
Penalty (2) 10-d	1723	$2.99139e-04$	1742	$2.99516e-04$	2116	$3.05215e-04$
Trigonometric 10-d	2392	$3.93063e-05$	1667	$5.55892e-07$	2882	$4.73616e-05$
Osbourne (2) 11-d	6663	$4.01377e-02$	4025	$4.01377e-02$	4752	$4.01377e-02$
Extended Powell 12-d	2807	$5.70567e-06$	3895	$1.68075e-06$	7756	$1.76192e-06$
Quadratic 16-d	1060	$2.61004e-07$	621	$4.02846e-09$	5266	$1.63425e-06$
Quadratic 24-d	1421	$1.91707e+01$	640	$1.46351e-08$	5631	$1.18931e-06$

Table E.3: Low tolerance results for variant two

E.2.2 High tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	489	$7.29197e-16$	489	$7.29197e-16$	489	$7.29197e-16$
Freudenstein and Roth 2-d	374	$4.89843e+01$	374	$4.89843e+01$	374	$4.89843e+01$
Powell badly scaled 2-d	994	$4.39168e-24$	994	$4.39168e-24$	994	$4.39168e-24$
Brown badly scaled 2-d	609	$2.37891e-15$	609	$2.37891e-15$	609	$2.37891e-15$
Beale 2-d	335	$1.06969e-17$	335	$1.06969e-17$	335	$1.06969e-17$
Jennrich and Sampson 2-d	313	$1.24362e+02$	313	$1.24362e+02$	313	$1.24362e+02$
M^cKinnon 2-d	100	$0.00000e+00$	100	$0.00000e+00$	100	$0.00000e+00$
Helical valley 3-d	326	$2.44295e-06$	326	$2.44295e-06$	326	$2.44295e-06$
Bard 3-d	378	$1.74362e+01$	1470	$1.74287e+01$	1084	$1.74287e+01$
Gaussian 3-d	422	$1.61547e-08$	422	$1.61547e-08$	422	$1.61547e-08$
Meyer 3-d	3393	$8.79459e+01$	3183	$8.79459e+01$	3352	$8.79459e+01$
Gulf research 3-d	955	$5.43579e-19$	955	$5.43579e-19$	955	$5.43579e-19$
Box 3-d	949	$1.95373e-18$	909	$4.20949e-22$	870	$1.91517e-18$
Powell singular 4-d	1381	$1.79745e-17$	1316	$1.32409e-21$	1413	$3.09802e-13$
Wood 4-d	1174	$1.92730e-15$	1174	$1.92730e-15$	1174	$1.92730e-15$
Kowalik and Osbourne 4-d	744	$3.07506e-04$	744	$3.07506e-04$	744	$3.07506e-04$
Brown and Dennis 4-d	1383	$8.58222e+04$	1383	$8.58222e+04$	1383	$8.58222e+04$
Quadratic 4-d	1229	$1.26589e-17$	965	$3.60555e-17$	831	$7.45348e-13$
Penalty (1) 4-d	1846	$2.24998e-05$	2168	$2.24998e-05$	3013	$2.24998e-05$
Penalty (2) 4-d	5335	$9.37629e-06$	6212	$9.37629e-06$	3666	$9.37629e-06$
Osbourne (1) 5-d	2047	$5.46489e-05$	1911	$5.46490e-05$	1951	$5.46489e-05$
Brown almost linear 5-d	945	$1.06367e-17$	957	$1.76674e-18$	1254	$7.20644e-13$
Biggs EXP6 6-d	3133	$5.65565e-03$	3362	$5.65565e-03$	26348	$2.42677e-01$
Extended Rosenbrock 6-d	7397	$1.02920e-15$	6276	$1.84528e-15$	5693	$2.44357e-16$
Brown almost-linear 7-d	5008	$7.73161e-06$	2291	$2.68743e-18$	2658	$2.63521e-16$

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	1135	$4.71000e-17$	2581	$8.47611e-17$	3176	$7.45816e-12$
Extended Rosenbrock 8-d	24337	$4.79991e-16$	30675	$2.27225e-16$	7566	$2.25893e-13$
Variably dimensional 8-d	5628	$8.74450e-08$	3059	$1.08772e-15$	3777	$3.05907e-13$
Extended Powell 8-d	7825	$1.19667e-12$	4453	$1.72555e-11$	4496	$1.06060e-12$
Watson 9-d	11843	$1.39976e-06$	21818	$1.39976e-06$	9672	$1.39976e-06$
Extended Rosenbrock 10-d	37266	$2.73138e-12$	54256	$2.94423e-01$	11789	$8.90003e-17$
Penalty (1) 10-d	10585	$7.08765e-05$	8163	$7.08765e-05$	13712	$7.08765e-05$
Penalty (2) 10-d	37476	$2.93661e-04$	23801	$2.93661e-04$	29642	$2.93661e-04$
Trigonometric 10-d	13963	$2.79506e-05$	4937	$4.47357e-07$	23719	$4.47357e-07$
Osbourne (2) 11-d	7778	$4.01377e-02$	6613	$4.01377e-02$	7325	$4.01377e-02$
Extended Powell 12-d	13600	$4.35719e-11$	13389	$1.01569e-09$	21739	$2.08892e-14$
Quadratic 16-d	3048	$6.87698e-15$	890	$4.31030e-17$	29073	$5.66083e-14$
Quadratic 24-d	2754	$1.91704e+01$	1332	$2.03035e-16$	9293	$1.49379e-15$

Table E.4: High tolerance results for variant two

E.3 Variant three

E.3.1 Low tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	263	$1.96997e-08$	222	$1.30125e-05$	188	$8.12755e-05$
Freudenstein and Roth 2-d	146	$4.89850e+01$	237	$4.89843e+01$	180	$4.89843e+01$
Powell badly scaled 2-d	114	$3.18399e-03$	107	$3.34996e-03$	278	$4.70905e-06$
Brown badly scaled 2-d	297	$1.91988e-02$	293	$1.91529e-02$	280	$1.93925e-02$
Beale 2-d	134	$2.88691e-05$	116	$2.22680e-05$	102	$1.72585e-04$
Jennrich and Sampson 2-d	136	$1.24362e+02$	170	$1.24362e+02$	115	$1.24362e+02$
M^{c}Kinnon 2-d	149	$-2.49988e-01$	156	$-2.49996e-01$	145	$-2.49995e-01$
Helical valley 3-d	134	$4.09825e-04$	115	$9.64526e-04$	149	$6.97225e-04$
Bard 3-d	247	$1.74354e+01$	238	$1.74354e+01$	272	$1.74354e+01$
Gaussian 3-d	51	$5.48432e-08$	45	$1.80849e-08$	55	$7.93123e-08$
Meyer 3-d	433	$8.72520e+04$	456	$8.78606e+04$	580	$7.85053e+04$
Gulf research 3-d	291	$2.89569e-02$	338	$2.88980e-02$	285	$2.89661e-02$
Box 3-d	162	$8.67161e-04$	95	$4.08021e-03$	99	$4.07679e-03$
Powell singular 4-d	219	$1.26890e-04$	219	$3.05778e-03$	186	$5.97567e-03$
Wood 4-d	220	$1.91704e+00$	255	$1.91565e+00$	225	$1.91701e+00$
Kowalik and Osbourne 4-d	472	$3.07516e-04$	400	$3.07553e-04$	158	$3.61568e-04$
Brown and Dennis 4-d	499	$8.58222e+04$	684	$8.58222e+04$	625	$8.58222e+04$
Quadratic 4-d	173	$8.33422e-05$	178	$8.32037e-05$	245	$9.19734e-06$
Penalty (1) 4-d	211	$2.83347e-05$	195	$2.83420e-05$	195	$2.83514e-05$
Osbourne (1) 5-d	283	$7.95694e-05$	294	$7.27591e-05$	354	$7.62421e-05$
Brown almost linear 5-d	399	$2.15023e-04$	682	$6.12985e-09$	734	$9.07996e-07$
Biggs EXP6 6-d	1098	$5.65566e-03$	683	$5.69407e-03$	1009	$5.65574e-03$
Extended Rosenbrock 6-d	2775	$3.19242e-06$	2214	$9.29475e-06$	194	$1.26143e+01$

Function	ψ_1 FE	ψ_1 Minimum	ψ_2 FE	ψ_2 Minimum	ψ_3 FE	ψ_3 Minimum
Brown almost-linear 7-d	755	$3.37785e-06$	567	$7.64211e-05$	1210	$3.93733e-05$
Quadratic 8-d	796	$1.42928e-04$	805	$1.43456e-04$	919	$9.88489e-05$
Extended Rosenbrock 8-d	213	$1.68783e+01$	3490	$6.22711e-04$	302	$1.68769e+01$
Variably dimensional 8-d	241	$1.49528e+01$	2647	$2.15648e-06$	2520	$7.44367e-06$
Extended Powell 8-d	991	$2.75387e-04$	1434	$2.47927e-04$	2264	$8.43923e-07$
Watson 9-d	2577	$5.78557e-04$	977	$1.84574e-02$	1014	$1.84594e-02$
Extended Rosenbrock 10-d	4078	$3.44327e-04$	4953	$1.15144e-02$	11226	$4.93290e-07$
Penalty (1) 10-d	1230	$9.62821e-05$	1268	$8.62219e-05$	1870	$1.20163e-04$
Penalty (2) 10-d	540	$1.70861e-03$	711	$2.99587e-04$	940	$2.99701e-04$
Trigonometric 10-d	1837	$2.79932e-05$	1427	$2.79843e-05$	1438	$3.83899e-05$
Osbourne (2) 11-d	995	$1.60737e-01$	5671	$4.01378e-02$	2679	$4.01427e-02$
Extended Powell 12-d	1247	$6.56269e-04$	1231	$7.05250e-04$	2097	$3.67257e-04$
Quadratic 16-d	3892	$2.80212e-06$	3141	$4.77784e-05$	4325	$3.76601e-04$
Quadratic 24-d	15898	$1.61113e-05$	16082	$1.05066e-07$	23389	$8.64291e-08$

Table E.5: Low tolerance results for variant three

E.3.2 High tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	429	$2.69575e-14$	414	$1.59584e-12$	459	$1.52240e-17$
Freudenstein and Roth 2-d	311	$4.89843e+01$	289	$4.89843e+01$	351	$4.89843e+01$
Powell badly scaled 2-d	198	$3.18031e-03$	193	$3.26570e-03$	951	$3.00621e-12$
Brown badly scaled 2-d	401	$1.91419e-02$	368	$1.91407e-02$	539	$4.42587e-11$
Beale 2-d	365	$2.57827e-15$	309	$3.60438e-11$	300	$6.43370e-13$
Jennrich and Sampson 2-d	358	$1.24362e+02$	259	$1.24362e+02$	352	$1.24362e+02$
M^{c}Kinnon 2-d	326	$-2.49999e-01$	283	$-2.50000e-01$	293	$-2.49999e-01$
Helical valley 3-d	557	$4.86101e-13$	776	$2.14069e-14$	513	$2.75341e-07$
Bard 3-d	351	$1.74354e+01$	305	$1.74354e+01$	1123	$1.74287e+01$
Gaussian 3-d	455	$1.12793e-08$	111	$8.30802e-08$	425	$1.12793e-08$
Meyer 3-d	589	$8.72520e+04$	594	$8.78606e+04$	3400	$8.79459e+01$
Gulf research 3-d	932	$2.08401e-13$	1047	$3.25034e-13$	1027	$3.53756e-13$
Box 3-d	895	$4.73708e-17$	988	$4.03251e-13$	241	$9.30423e-04$
Powell singular 4-d	481	$1.95864e-09$	503	$4.10752e-07$	1296	$4.96603e-13$
Wood 4-d	1547	$4.26977e-10$	1540	$7.75890e-11$	1453	$8.04226e-13$
Kowalik and Osbourne 4-d	932	$3.07506e-04$	551	$3.07510e-04$	718	$3.07506e-04$
Brown and Dennis 4-d	916	$8.58222e+04$	1494	$8.58222e+04$	1266	$8.58222e+04$
Quadratic 4-d	685	$6.19190e-12$	467	$1.68895e-09$	867	$8.59560e-14$
Penalty (1) 4-d	265	$2.83554e-05$	270	$2.83554e-05$	296	$2.83553e-05$
Osbourne (1) 5-d	1784	$5.46489e-05$	1780	$5.46489e-05$	1841	$5.46489e-05$
Brown almost linear 5-d	1536	$1.69625e-14$	1243	$5.57365e-15$	1461	$1.38792e-15$
Biggs EXP6 6-d	2297	$5.65565e-03$	2285	$5.65565e-03$	1978	$5.65565e-03$
Extended Rosenbrock 6-d	2864	$1.82444e-12$	2891	$2.07694e-11$	5442	$5.44384e-14$
Brown almost-linear 7-d	2452	$1.02927e-15$	1500	$1.71444e-09$	3760	$2.25419e-13$
Quadratic 8-d	2123	$3.31318e-13$	2728	$5.84796e-15$	3238	$2.54126e-14$

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Extended Rosenbrock 8-d	8527	$1.52065e-12$	13144	$1.49696e-12$	14862	$7.29488e-16$
Variably dimensional 8-d	4884	$1.15090e-13$	3963	$1.07664e-14$	4836	$1.10802e-12$
Extended Powell 8-d	2553	$5.25957e-08$	3508	$2.48071e-10$	5198	$4.66812e-14$
Watson 9-d	19144	$1.39976e-06$	18437	$1.39976e-06$	10765	$1.39976e-06$
Extended Rosenbrock 10-d	7145	$8.93943e-12$	12647	$7.52405e-11$	15316	$1.18281e-11$
Penalty (1) 10-d	15072	$7.08765e-05$	50342	$7.14675e-05$	2024	$1.20164e-04$
Penalty (2) 10-d	1500	$2.97569e-04$	1431	$2.98788e-04$	25637	$2.93661e-04$
Trigonometric 10-d	3878	$2.79506e-05$	4088	$2.79506e-05$	6658	$2.79506e-05$
Osbourne (2) 11-d	10527	$4.01377e-02$	9553	$4.01377e-02$	7848	$4.01377e-02$
Extended Powell 12-d	10400	$1.40951e-10$	15716	$2.08084e-11$	22969	$2.96685e-13$
Quadratic 16-d	5250	$1.38564e-09$	5328	$4.81974e-13$	13111	$5.84285e-14$
Quadratic 24-d	29169	$3.22155e-14$	20088	$1.08547e-15$	33147	$2.70247e-15$

Table E.6: High tolerance results for variant three

E.4 Variant four

E.4.1 Low tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	267	$9.14807e-06$	289	$7.09679e-08$	296	$4.85725e-07$
Freudenstein and Roth 2-d	163	$4.89848e+01$	324	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	104	$1.21565e-01$	107	$1.18160e-01$	1217	$1.46409e-10$
Brown badly scaled 2-d	371	$2.33639e+02$	433	$6.49698e+01$	612	$3.07505e-05$
Beale 2-d	95	$2.96187e-04$	151	$2.06814e-06$	195	$2.25494e-10$
Jennrich and Sampson 2-d	255	$1.24362e+02$	134	$1.24362e+02$	178	$1.24363e+02$
M^{c}Kinnon 2-d	54	$-6.24961e-05$	30	$-6.06691e-05$	217	$-2.49999e-01$
Helical valley 3-d	139	$5.88697e-03$	215	$1.35335e-02$	160	$1.29542e-04$
Bard 3-d	291	$1.77864e+01$	82707	$1.79374e+01$	1462	$1.74287e+01$
Gaussian 3-d	62	$3.19326e-07$	49	$3.33819e-07$	64	$2.76991e-07$
Meyer 3-d	558	$8.64198e+04$	746	$9.11277e+04$	2910	$8.79459e+01$
Gulf research 3-d	487	$4.85057e-04$	410	$4.82234e-03$	1079	$1.41324e-10$
Box 3-d	311	$5.68514e-07$	218	$4.27029e-03$	566	$4.98516e-11$
Powell singular 4-d	802	$3.22025e-06$	272	$4.43294e-03$	668	$1.57430e-08$
Wood 4-d	548	$2.96755e-03$	1295	$1.07727e-04$	689	$9.52877e-05$
Kowalik and Osbourne 4-d	727	$3.07507e-04$	575	$3.07509e-04$	649	$3.07507e-04$
Brown and Dennis 4-d	753	$8.58222e+04$	508	$8.58222e+04$	872	$8.58222e+04$
Quadratic 4-d	228	$3.10826e-08$	262	$3.11175e-10$	256	$1.91309e-06$
Penalty (1) 4-d	272	$4.05736e-05$	264	$3.03302e-05$	287	$2.66985e-05$
Penalty (2) 4-d	213	$1.03728e-05$	245	$1.01092e-05$	236	$1.01174e-05$
Osbourne (1) 5-d	133	$2.01371e-04$	146	$3.92301e-03$	231	$5.52534e-04$
Brown almost linear 5-d	254	$3.04975e-04$	445	$1.17245e-08$	208	$8.68485e-04$
Biggs EXP6 6-d	1968	$7.91708e-05$	4350	$3.55731e-06$	6601	$1.31587e-07$

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Extended Rosenbrock 6-d	2916	$5.17117e-03$	2231	$6.29395e-04$	2686	$2.07218e-04$
Brown almost-linear 7-d	152	$5.17947e-03$	230	$2.33361e-09$	1660	$1.86730e-06$
Quadratic 8-d	446	$1.97159e-06$	303	$2.56738e-09$	924	$5.54485e-07$
Extended Rosenbrock 8-d	4371	$8.25807e-01$	3696	$5.02591e+00$	4102	$1.03419e-04$
Variably dimensional 8-d	208	$9.44704e-01$	3151	$3.42812e-06$	259	$9.63431e+00$
Extended Powell 8-d	946	$1.59976e-01$	2144	$2.72241e-06$	2156	$1.57513e-04$
Watson 9-d	649	$5.53723e-03$	1504	$5.00692e-05$	2398	$7.97759e-04$
Extended Rosenbrock 10-d	5693	$4.56265e-01$	6333	$1.02077e+00$	4051	$3.84036e-03$
Penalty (1) 10-d	551	$7.53285e-03$	745	$1.03666e-04$	1330	$9.78209e-05$
Penalty (2) 10-d	420	$1.96874e-03$	1237	$2.98331e-04$	1004	$2.99790e-04$
Trigonometric 10-d	508	$5.87076e-05$	1414	$2.79987e-05$	2418	$2.81070e-05$
Osbourne (2) 11-d	697	$2.62935e-01$	6415	$4.01381e-02$	5154	$4.01951e-02$
Extended Powell 12-d	497	$1.71595e+01$	5198	$5.61271e-05$	7669	$9.77144e-05$
Quadratic 16-d	737	$2.33345e-05$	772	$6.35516e-09$	2739	$5.02812e-08$
Quadratic 24-d	1413	$1.91721e+01$	977	$1.49394e-08$	4387	$1.69048e-07$

Table E.7: Low tolerance results for variant four

E.4.2 High tolerance results

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	502	$4.86131e-12$	350	$5.59925e-08$	484	$2.32840e-15$
Freudenstein and Roth 2-d	401	$4.89843e+01$	657	$4.89843e+01$	412	$4.89843e+01$
Powell badly scaled 2-d	890	$1.61238e-09$	7522	$9.12187e-15$	1793	$8.57271e-20$
Brown badly scaled 2-d	504	$2.33639e+02$	552	$6.49697e+01$	733	$1.50980e-09$
Beale 2-d	394	$6.05795e-15$	446	$1.33463e-15$	362	$2.47613e-17$
Jennrich and Sampson 2-d	477	$1.24362e+02$	312	$1.24362e+02$	399	$1.24362e+02$
M^{c}Kinnon 2-d	149	$-6.24961e-05$	206	$-6.90469e-05$	649	$-2.50000e-01$
Helical valley 3-d	857	$7.78057e-14$	616	$3.28527e-12$	808	$3.38803e-17$
Bard 3-d	1288	$1.74287e+01$	83758	$1.69027e+01$	1622	$1.74287e+01$
Gaussian 3-d	378	$1.12793e-08$	340	$1.12793e-08$	416	$1.12793e-08$
Meyer 3-d	736	$8.64198e+04$	4018	$8.79459e+01$	3252	$8.79459e+01$
Gulf research 3-d	977	$2.84803e-11$	873	$9.10624e-12$	1510	$1.06763e-16$
Box 3-d	673	$3.77860e-13$	979	$5.82960e-14$	796	$3.88191e-16$
Powell singular 4-d	1746	$1.57061e-14$	1712	$3.44583e-16$	1136	$2.24553e-11$
Wood 4-d	1578	$4.30914e-10$	2358	$2.11440e-12$	1729	$1.84091e-17$
Kowalik and Osbourne 4-d	1115	$3.07506e-04$	1262	$3.07506e-04$	1013	$3.07506e-04$
Brown and Dennis 4-d	1279	$8.58222e+04$	1070	$8.58222e+04$	1453	$8.58222e+04$
Quadratic 4-d	494	$3.71518e-15$	422	$8.18402e-17$	909	$7.63455e-18$
Penalty (1) 4-d	2268	$2.24998e-05$	2440	$2.24998e-05$	2117	$2.24998e-05$
Penalty (2) 4-d	7086	$9.37629e-06$	5247	$9.37629e-06$	5280	$9.37629e-06$
Osbourne (1) 5-d	2903	$5.46489e-05$	2173	$5.46490e-05$	2169	$5.46489e-05$
Brown almost linear 5-d	1301	$3.05012e-14$	782	$3.42996e-18$	1887	$5.81022e-15$
Biggs EXP6 6-d	4538	$4.36848e-11$	6253	$1.31946e-14$	8931	$3.33591e-19$
Extended Rosenbrock 6-d	5279	$1.89732e-10$	3806	$1.11839e-11$	5814	$6.32421e-14$
Brown almost-linear 7-d	4403	$6.92985e-14$	1179	$1.49580e-18$	4129	$2.57774e-15$

	ψ_1		ψ_2		ψ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	1202	$4.39412e-14$	638	$6.37957e-17$	2188	$1.64109e-16$
Extended Rosenbrock 8-d	25177	$5.32457e-12$	15366	$5.42900e-12$	9589	$6.97090e-12$
Variably dimensional 8-d	3618	$1.46469e-07$	4284	$1.35438e-14$	5566	$5.14570e-15$
Extended Powell 8-d	7932	$1.21807e-11$	5605	$1.27570e-11$	7889	$4.01552e-13$
Watson 9-d	3470	$7.55556e-05$	19613	$1.39976e-06$	15214	$1.39976e-06$
Extended Rosenbrock 10-d	38229	$8.92368e-09$	42888	$7.19274e-12$	12620	$5.23747e-12$
Penalty (1) 10-d	7250	$7.08783e-05$	100006	$8.82749e-05$	22520	$7.08765e-05$
Penalty (2) 10-d	934	$2.98332e-04$	100001	$2.97300e-04$	27651	$2.93661e-04$
Trigonometric 10-d	2100	$2.79554e-05$	5092	$2.79506e-05$	5993	$2.79506e-05$
Osbourne (2) 11-d	9922	$4.01377e-02$	41881	$4.01377e-02$	15880	$4.01377e-02$
Extended Powell 12-d	100000	$3.51576e-08$	13134	$2.39979e-11$	33123	$8.42361e-15$
Quadratic 16-d	2329	$2.36664e-11$	1457	$6.35956e-18$	5888	$7.63090e-16$
Quadratic 24-d	2328	$1.91721e+01$	1755	$4.00351e-17$	10213	$1.67792e-15$

Table E.8: High tolerance results for variant four

Appendix F

Summary of parameter choices

This appendix contains the performance summary for the NM4_ψ_3 algorithm with different values for each of the parameters δ, κ and ν. These results were used to decide the final choice of parameters for the recommended variant (NM4_$\psi_3\delta_6\kappa_1\nu_8$) of the Nelder-Mead algorithm.

The results summarised in this appendix were obtained directly from the complete and comprehensive list of tables in appendix G — all 70 pages of them! As mentioned previously, only the results for the high tolerance tests were used to decide the parameter values.

In the following tables, *FE* represents the total number of function evaluations required to produce approximations for all of the functions in the test-suite. A clubsuit symbol ($\clubsuit$) next to an entry indicates that this combination of parameters failed to find an accurate approximation of the solution for at least one of the functions in the test-suite. The table heading *Mag* represents a relative measure of the accuracy of the approximations produced (relative to the results produced by the parameter in the first row of each table). The higher the value, the greater the accuracy compared to the first combination of parameters. The most successful value of each parameter is circled. These values were then used to determine the remaining parameters.

F.1 Using the default parameters

F.1.1 Determinant parameter — δ

Summary for NM4_$\psi_3\kappa_2\nu_1$

Name	*Value*	*FE*	*Mag*
δ_1	10^{-5}	2.09×10^5	0
δ_2	10^{-8}	2.42×10^5	−19
δ_3	10^{-10}	2.24×10^5	−13
δ_4	10^{-12}	2.12×10^5	−19
δ_5	10^{-15}	2.05×10^5	−17
δ_6	10^{-18}	2.27×10^5	−12
δ_7	10^{-20}	2.08×10^5	−17
δ_8	10^{-25}	2.12×10^5	−16
δ_9	10^{-30}	2.09×10^5	−17

Table F.1: Determinant parameter for NM4_$\psi_3\kappa_2\nu_1$.

F.1.2 Frame reduction parameter — κ

Summary for NM4_$\psi_3\delta_1\nu_1$

Name	*Value*	*FE*	*Mag*
κ_1	0.25	♣1.95×10^5	0
κ_2	0.50	2.09×10^5	32
κ_3	0.75	2.55×10^5	16

Table F.2: Frame reduction parameter for NM4_$\psi_3\delta_1\nu_1$.

Summary for NM4_$\psi_3\delta_5\nu_1$

Name	*Value*	*FE*	*Mag*
κ_1	0.25	1.89×10^5	0
κ_2	0.50	2.05×10^5	−4
κ_3	0.75	♣2.65×10^5	−8

Table F.3: Frame reduction parameter for NM4_$\psi_3\delta_5\nu_1$.

F.1.3 Sufficient descent reduction parameter — ν

Summary for NM4_$\psi_3\delta_1\kappa_2$

Name	*Value*	*FE*	*Mag*
ν_1	1.25	2.09×10^5	0
ν_2	1.50	1.90×10^5	8
ν_3	2.00	1.97×10^5	30
ν_4	2.50	1.75×10^5	29
ν_5	3.00	$\clubsuit 1.43 \times 10^5$	10
ν_6	3.50	$\clubsuit 1.81 \times 10^5$	13
ν_7	4.00	1.52×10^5	39
ν_8	4.50	1.48×10^5	42
ν_9	5.00	1.59×10^5	57

Table F.4: Sufficient descent reduction parameter for NM4_$\psi_3\delta_1\kappa_2$.

Summary for NM4_$\psi_3\delta_5\kappa_1$

Name	*Value*	*FE*	*Mag*
ν_1	1.25	1.89×10^5	0
ν_2	1.50	$\clubsuit 1.95 \times 10^5$	4
ν_3	2.00	1.68×10^5	35
ν_4	2.50	$\clubsuit 1.54 \times 10^5$	20
ν_5	3.00	$\clubsuit 1.86 \times 10^5$	23
ν_6	3.50	1.68×10^5	46
ν_7	4.00	1.75×10^5	59
ν_8	4.50	1.52×10^5	68
ν_9	5.00	$\clubsuit 1.67 \times 10^5$	42

Table F.5: Sufficient descent reduction parameter for NM4_$\psi_3\delta_5\kappa_1$.

F.2 Using the new parameter information

Since ψ_3 and ν_8 appear in both of the most successful variants found so far, these values will be used to find the best determinant parameter δ for each of κ_1 and κ_2.

F.2.1 Determinant parameter — δ

Summary for NM4_$\psi_3\kappa_1\nu_8$

Name	*Value*	*FE*	*Mag*
δ_1	10^{-5}	1.71×10^5	0
δ_2	10^{-8}	1.58×10^5	10
δ_3	10^{-10}	1.55×10^5	23
δ_4	10^{-12}	1.62×10^5	10
δ_5	10^{-15}	1.52×10^5	24
δ_6	10^{-18}	1.36×10^5	24
δ_7	10^{-20}	1.43×10^5	19
δ_8	10^{-25}	1.43×10^5	19
δ_9	10^{-30}	1.68×10^5	19

Table F.6: Determinant parameter for NM4_$\psi_3\kappa_1\nu_8$.

Summary for NM4_$\psi_3\kappa_2\nu_8$

Name	*Value*	*FE*	*Mag*
δ_1	10^{-5}	1.48×10^5	0
δ_2	10^{-8}	1.53×10^5	15
δ_3	10^{-10}	1.67×10^5	9
δ_4	10^{-12}	1.45×10^5	7
δ_5	10^{-15}	1.64×10^5	14
δ_6	10^{-18}	1.47×10^5	8
δ_7	10^{-20}	1.65×10^5	14
δ_8	10^{-25}	1.62×10^5	13
δ_9	10^{-30}	1.62×10^5	13

Table F.7: Determinant parameter for NM4_$\psi_3\kappa_2\nu_8$.

F.3 The most successful variants

Name	*FE*	*Mag*
NM4_$\psi_3\delta_1\kappa_2\nu_8$	1.48×10^5	0
NM4_$\psi_3\delta_4\kappa_2\nu_6$	1.45×10^5	8
NM4_$\psi_3\delta_5\kappa_1\nu_8$	1.52×10^5	13
NM4_$\psi_3\delta_6\kappa_1\nu_8$	1.36×10^5	13

Table F.8: Results of the most successful parameter choices for variant four.

The most successful variant of the Nelder-Mead algorithm considered so far is NM4_$\psi_3\delta_6\kappa_1\nu_8$.

Appendix G

Choice of parameters

This section includes a comprehensive list of all experimental data obtained for deciding the values of the parameters for variant four. For completeness, both the low and high tolerance results are presented here. However, as mentioned previously, only the results of the high tolerance tests were used to decide the values of the parameters. A summary of this data is contained in appendix F.

G.1 Determinant parameter (δ) using ψ_3, κ_2, ν_1

G.1.1 Low tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	296	$4.85725e-07$	296	$4.85725e-07$	296	$4.85725e-07$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1282	$4.83924e-12$	1215	$1.46409e-10$	1217	$1.46409e-10$
Brown badly scaled 2-d	610	$3.07505e-05$	612	$3.07505e-05$	612	$3.07505e-05$
Beale 2-d	195	$2.25494e-10$	195	$2.25494e-10$	195	$2.25494e-10$
Jennrich and Sampson 2-d	178	$1.24363e+02$	178	$1.24363e+02$	178	$1.24363e+02$
M^{c}Kinnon 2-d	217	$-2.49999e-01$	217	$-2.49999e-01$	217	$-2.49999e-01$
Helical valley 3-d	160	$1.29542e-04$	160	$1.29542e-04$	160	$1.29542e-04$
Bard 3-d	1452	$1.74287e+01$	1457	$1.74287e+01$	1462	$1.74287e+01$
Gaussian 3-d	64	$2.76991e-07$	64	$2.76991e-07$	64	$2.76991e-07$
Meyer 3-d	3206	$8.79459e+01$	3016	$8.79459e+01$	2910	$8.79459e+01$
Gulf research 3-d	1077	$1.41324e-10$	1079	$1.41324e-10$	1079	$1.41324e-10$
Box 3-d	135	$4.67565e-04$	566	$4.98516e-11$	566	$4.98516e-11$
Powell singular 4-d	755	$3.27267e-06$	668	$1.57430e-08$	668	$1.57430e-08$
Wood 4-d	626	$3.80801e-05$	689	$9.52877e-05$	689	$9.52877e-05$
Kowalik and Osbourne 4-d	649	$3.07507e-04$	649	$3.07507e-04$	649	$3.07507e-04$
Brown and Dennis 4-d	871	$8.58222e+04$	872	$8.58222e+04$	872	$8.58222e+04$
Quadratic 4-d	256	$1.91309e-06$	256	$1.91309e-06$	256	$1.91309e-06$
Penalty (1) 4-d	288	$2.70355e-05$	287	$2.66985e-05$	287	$2.66985e-05$
Penalty (2) 4-d	235	$1.01174e-05$	236	$1.01174e-05$	236	$1.01174e-05$
Osbourne (1) 5-d	229	$5.52534e-04$	231	$5.52534e-04$	231	$5.52534e-04$
Brown almost linear 5-d	173	$1.72332e-03$	208	$8.68485e-04$	208	$8.68485e-04$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	4136	$1.15161e-10$	7349	$6.52699e-10$	6601	$1.31587e-07$
Extended Rosenbrock 6-d	3002	$2.07820e-08$	2685	$2.07218e-04$	2686	$2.07218e-04$
Brown almost-linear 7-d	261	$5.19471e-03$	985	$3.33064e-07$	1660	$1.86730e-06$
Quadratic 8-d	1088	$1.56330e-06$	1047	$8.23915e-08$	924	$5.54485e-07$
Extended Rosenbrock 8-d	3970	$6.35038e-05$	4274	$2.35953e-05$	4102	$1.03419e-04$
Variably dimensional 8-d	2537	$2.25052e-06$	224	$1.00346e+01$	259	$9.63431e+00$
Extended Powell 8-d	2347	$1.30161e-04$	3492	$1.21785e-05$	2156	$1.57513e-04$
Watson 9-d	2730	$8.69483e-04$	2491	$7.89410e-05$	2398	$7.97759e-04$
Extended Rosenbrock 10-d	4650	$1.29318e-03$	6445	$4.41289e-04$	4051	$3.84036e-03$
Penalty (1) 10-d	1392	$8.09398e-05$	8603	$7.09144e-05$	1330	$9.78209e-05$
Penalty (2) 10-d	1098	$2.99373e-04$	1194	$2.99902e-04$	1004	$2.99790e-04$
Trigonometric 10-d	2410	$2.81070e-05$	2418	$2.81070e-05$	2418	$2.81070e-05$
Osbourne (2) 11-d	6717	$4.01410e-02$	6398	$4.01778e-02$	5154	$4.01951e-02$
Extended Powell 12-d	7129	$9.60933e-05$	6732	$2.10920e-04$	7669	$9.77144e-05$
Quadratic 16-d	2183	$2.72998e-08$	2271	$5.82072e-08$	2739	$5.02812e-08$
Quadratic 24-d	4601	$5.95042e-08$	4376	$1.69048e-07$	4387	$1.69048e-07$

Table G.1: Low tolerance results for `NM4`_ψ_3 δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	296	$4.85725e-07$	296	$4.85725e-07$	296	$4.85725e-07$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1217	$1.46409e-10$	1217	$1.46409e-10$	1217	$1.46409e-10$
Brown badly scaled 2-d	612	$3.07505e-05$	612	$3.07505e-05$	612	$3.07505e-05$
Beale 2-d	195	$2.25494e-10$	195	$2.25494e-10$	195	$2.25494e-10$
Jennrich and Sampson 2-d	178	$1.24363e+02$	178	$1.24363e+02$	178	$1.24363e+02$
M^{c}Kinnon 2-d	217	$-2.49999e-01$	217	$-2.49999e-01$	217	$-2.49999e-01$
Helical valley 3-d	160	$1.29542e-04$	160	$1.29542e-04$	160	$1.29542e-04$
Bard 3-d	2074	$1.74287e+01$	1625	$1.74287e+01$	1627	$1.74287e+01$
Gaussian 3-d	64	$2.76991e-07$	64	$2.76991e-07$	64	$2.76991e-07$
Meyer 3-d	2910	$8.79459e+01$	2910	$8.79459e+01$	2910	$8.79459e+01$
Gulf research 3-d	1079	$1.41324e-10$	1079	$1.41324e-10$	1079	$1.41324e-10$
Box 3-d	566	$4.98516e-11$	566	$4.98516e-11$	566	$4.98516e-11$
Powell singular 4-d	668	$1.57430e-08$	668	$1.57430e-08$	668	$1.57430e-08$
Wood 4-d	689	$9.52877e-05$	689	$9.52877e-05$	689	$9.52877e-05$
Kowalik and Osbourne 4-d	649	$3.07507e-04$	649	$3.07507e-04$	649	$3.07507e-04$
Brown and Dennis 4-d	872	$8.58222e+04$	872	$8.58222e+04$	872	$8.58222e+04$
Quadratic 4-d	256	$1.91309e-06$	256	$1.91309e-06$	256	$1.91309e-06$
Penalty (1) 4-d	287	$2.66985e-05$	287	$2.66985e-05$	287	$2.66985e-05$
Penalty (2) 4-d	236	$1.01174e-05$	236	$1.01174e-05$	236	$1.01174e-05$
Osbourne (1) 5-d	231	$5.52534e-04$	231	$5.52534e-04$	231	$5.52534e-04$
Brown almost linear 5-d	208	$8.68485e-04$	208	$8.68485e-04$	208	$8.68485e-04$
Biggs EXP6 6-d	6863	$9.19311e-10$	6867	$9.19311e-10$	6867	$9.19311e-10$
Extended Rosenbrock 6-d	2686	$2.07218e-04$	2686	$2.07218e-04$	2686	$2.07218e-04$
Brown almost-linear 7-d	1662	$1.86730e-06$	1662	$1.86730e-06$	1662	$1.86730e-06$
Quadratic 8-d	925	$5.54485e-07$	925	$5.54485e-07$	925	$5.54485e-07$
Extended Rosenbrock 8-d	3941	$1.24884e-03$	3942	$1.24884e-03$	3942	$1.24884e-03$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	2417	$7.84288e-07$	276	$9.17632e+00$	2138	$7.48019e-07$
Extended Powell 8-d	3635	$3.05323e-06$	2880	$2.94498e-06$	2880	$2.94498e-06$
Watson 9-d	2400	$7.97759e-04$	2655	$1.03023e-03$	3046	$3.83630e-04$
Extended Rosenbrock 10-d	3386	$1.27839e-02$	3388	$1.27839e-02$	3388	$1.27839e-02$
Penalty (1) 10-d	1330	$9.78209e-05$	1330	$9.79506e-05$	1334	$9.79424e-05$
Penalty (2) 10-d	957	$2.99807e-04$	971	$2.99765e-04$	1024	$2.99755e-04$
Trigonometric 10-d	2418	$2.81070e-05$	2418	$2.81070e-05$	2418	$2.81070e-05$
Osbourne (2) 11-d	5608	$4.01389e-02$	5609	$4.01389e-02$	5609	$4.01389e-02$
Extended Powell 12-d	6951	$1.00453e-04$	6805	$1.94398e-04$	8134	$6.01118e-04$
Quadratic 16-d	2739	$5.02812e-08$	2770	$6.88791e-08$	1955	$3.94857e-07$
Quadratic 24-d	4439	$1.20597e-07$	5354	$5.43072e-08$	4533	$9.59420e-08$

Table G.2: Low tolerance results for NM4_ψ_3 δ_4–δ_6.

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	296	$4.85725e-07$	296	$4.85725e-07$	296	$4.85725e-07$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1217	$1.46409e-10$	1217	$1.46409e-10$	1217	$1.46409e-10$
Brown badly scaled 2-d	612	$3.07505e-05$	612	$3.07505e-05$	612	$3.07505e-05$
Beale 2-d	195	$2.25494e-10$	195	$2.25494e-10$	195	$2.25494e-10$
Jennrich and Sampson 2-d	178	$1.24363e+02$	178	$1.24363e+02$	178	$1.24363e+02$
$M^cKinnon$ 2-d	217	$-2.49999e-01$	217	$-2.49999e-01$	217	$-2.49999e-01$
Helical valley 3-d	160	$1.29542e-04$	160	$1.29542e-04$	160	$1.29542e-04$
Bard 3-d	1628	$1.74287e+01$	1628	$1.74287e+01$	1629	$1.74287e+01$
Gaussian 3-d	64	$2.76991e-07$	64	$2.76991e-07$	64	$2.76991e-07$
Meyer 3-d	2910	$8.79459e+01$	2910	$8.79459e+01$	2910	$8.79459e+01$
Gulf research 3-d	1079	$1.41324e-10$	1079	$1.41324e-10$	1079	$1.41324e-10$
Box 3-d	566	$4.98516e-11$	566	$4.98516e-11$	566	$4.98516e-11$
Powell singular 4-d	668	$1.57430e-08$	668	$1.57430e-08$	668	$1.57430e-08$
Wood 4-d	689	$9.52877e-05$	689	$9.52877e-05$	689	$9.52877e-05$
Kowalik and Osbourne 4-d	649	$3.07507e-04$	649	$3.07507e-04$	649	$3.07507e-04$
Brown and Dennis 4-d	872	$8.58222e+04$	872	$8.58222e+04$	872	$8.58222e+04$
Quadratic 4-d	256	$1.91309e-06$	256	$1.91309e-06$	256	$1.91309e-06$
Penalty (1) 4-d	287	$2.66985e-05$	287	$2.66985e-05$	287	$2.66985e-05$
Penalty (2) 4-d	236	$1.01174e-05$	236	$1.01174e-05$	236	$1.01174e-05$
Osbourne (1) 5-d	231	$5.52534e-04$	231	$5.52534e-04$	231	$5.52534e-04$
Brown almost linear 5-d	208	$8.68485e-04$	208	$8.68485e-04$	208	$8.68485e-04$
Biggs EXP6 6-d	6867	$9.19311e-10$	6867	$9.19311e-10$	6867	$9.19311e-10$
Extended Rosenbrock 6-d	2686	$2.07218e-04$	2686	$2.07218e-04$	2686	$2.07218e-04$
Brown almost-linear 7-d	1662	$1.86730e-06$	1662	$1.86730e-06$	1662	$1.86730e-06$
Quadratic 8-d	925	$5.54485e-07$	925	$5.54485e-07$	925	$5.54485e-07$
Extended Rosenbrock 8-d	3942	$1.24884e-03$	3942	$1.24884e-03$	3942	$1.24884e-03$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	2060	$6.04474e-07$	2061	$6.04474e-07$	2061	$6.04474e-07$
Extended Powell 8-d	2880	$2.94498e-06$	2880	$2.94498e-06$	2880	$2.94498e-06$
Watson 9-d	3182	$3.94112e-04$	3183	$3.94112e-04$	3183	$3.94112e-04$
Extended Rosenbrock 10-d	3388	$1.27839e-02$	3388	$1.27839e-02$	3388	$1.27839e-02$
Penalty (1) 10-d	1335	$9.79424e-05$	1336	$9.79424e-05$	1336	$9.79424e-05$
Penalty (2) 10-d	1026	$2.99755e-04$	1026	$2.99755e-04$	1026	$2.99755e-04$
Trigonometric 10-d	2418	$2.81070e-05$	2418	$2.81070e-05$	2418	$2.81070e-05$
Osbourne (2) 11-d	5609	$4.01389e-02$	5609	$4.01389e-02$	5609	$4.01389e-02$
Extended Powell 12-d	7572	$2.03968e-04$	7573	$2.03968e-04$	7573	$2.03968e-04$
Quadratic 16-d	2189	$6.37851e-07$	2190	$6.37851e-07$	2190	$6.37851e-07$
Quadratic 24-d	4731	$8.10657e-08$	5079	$4.51984e-08$	3736	$1.06930e-07$

Table G.3: Low tolerance results for NM4_ψ_3 δ_7–δ_9.

G.1.2 High tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	484	$2.32840e-15$	484	$2.32840e-15$	484	$2.32840e-15$
Freudenstein and Roth 2-d	412	$4.89843e+01$	412	$4.89843e+01$	412	$4.89843e+01$
Powell badly scaled 2-d	1951	$6.43174e-24$	1787	$8.57271e-20$	1793	$8.57271e-20$
Brown badly scaled 2-d	731	$1.50980e-09$	733	$1.50980e-09$	733	$1.50980e-09$
Beale 2-d	362	$2.47613e-17$	362	$2.47613e-17$	362	$2.47613e-17$
Jennrich and Sampson 2-d	399	$1.24362e+02$	399	$1.24362e+02$	399	$1.24362e+02$
M^{c}Kinnon 2-d	672	$-2.50000e-01$	714	$-2.50000e-01$	649	$-2.50000e-01$
Helical valley 3-d	808	$3.38803e-17$	808	$3.38803e-17$	808	$3.38803e-17$
Bard 3-d	1612	$1.74287e+01$	1617	$1.74287e+01$	1622	$1.74287e+01$
Gaussian 3-d	415	$1.12793e-08$	416	$1.12793e-08$	416	$1.12793e-08$
Meyer 3-d	3475	$8.79459e+01$	3406	$8.79459e+01$	3252	$8.79459e+01$
Gulf research 3-d	1508	$1.06763e-16$	1510	$1.06763e-16$	1510	$1.06763e-16$
Box 3-d	1016	$1.26933e-20$	796	$3.88191e-16$	796	$3.88191e-16$
Powell singular 4-d	1503	$1.97194e-16$	1136	$2.24553e-11$	1136	$2.24553e-11$
Wood 4-d	1385	$7.74702e-15$	1729	$1.84091e-17$	1729	$1.84091e-17$
Kowalik and Osbourne 4-d	1013	$3.07506e-04$	1013	$3.07506e-04$	1013	$3.07506e-04$
Brown and Dennis 4-d	1452	$8.58222e+04$	1453	$8.58222e+04$	1453	$8.58222e+04$
Quadratic 4-d	909	$7.63455e-18$	909	$7.63455e-18$	909	$7.63455e-18$
Penalty (1) 4-d	2346	$2.24998e-05$	2115	$2.24998e-05$	2117	$2.24998e-05$
Penalty (2) 4-d	5340	$9.37629e-06$	5345	$9.37629e-06$	5280	$9.37629e-06$
Osbourne (1) 5-d	2241	$5.46489e-05$	2167	$5.46489e-05$	2169	$5.46489e-05$
Brown almost linear 5-d	1694	$3.27276e-16$	1886	$5.81022e-15$	1887	$5.81022e-15$
Biggs EXP6 6-d	6377	$1.11532e-16$	9470	$6.32358e-16$	8931	$3.33591e-19$
Extended Rosenbrock 6-d	5283	$6.03960e-18$	5808	$6.32421e-14$	5814	$6.32421e-14$
Brown almost-linear 7-d	3344	$9.73250e-16$	3062	$4.06752e-15$	4129	$2.57774e-15$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	2376	$4.11882e-16$	2332	$1.73637e-15$	2188	$1.64109e-16$
Extended Rosenbrock 8-d	8850	$1.73540e-13$	9905	$1.73942e-13$	9589	$6.97090e-12$
Variably dimensional 8-d	4639	$1.12738e-14$	5415	$1.71538e-13$	5566	$5.14570e-15$
Extended Powell 8-d	8025	$5.01237e-12$	8155	$1.09875e-14$	7889	$4.01552e-13$
Watson 9-d	10928	$1.39976e-06$	14705	$1.39976e-06$	15214	$1.39976e-06$
Extended Rosenbrock 10-d	12979	$1.57300e-13$	13933	$6.22827e-12$	12620	$5.23747e-12$
Penalty (1) 10-d	17571	$7.08765e-05$	21743	$7.08765e-05$	22520	$7.08765e-05$
Penalty (2) 10-d	36125	$2.93661e-04$	32890	$2.93661e-04$	27651	$2.93661e-04$
Trigonometric 10-d	5123	$2.79506e-05$	5991	$2.79506e-05$	5993	$2.79506e-05$
Osbourne (2) 11-d	18071	$4.01377e-02$	18201	$4.01377e-02$	15880	$4.01377e-02$
Extended Powell 12-d	21995	$6.02124e-14$	44052	$1.59087e-14$	33123	$8.42361e-15$
Quadratic 16-d	5150	$6.31179e-16$	5571	$1.41355e-15$	5888	$7.63090e-16$
Quadratic 24-d	9979	$1.12022e-15$	9514	$2.90385e-15$	10213	$1.67792e-15$

Table G.4: High tolerance results for NM4_ψ_3 δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	484	$2.32840e-15$	484	$2.32840e-15$	484	$2.32840e-15$
Freudenstein and Roth 2-d	412	$4.89843e+01$	412	$4.89843e+01$	412	$4.89843e+01$
Powell badly scaled 2-d	1793	$8.57271e-20$	1793	$8.57271e-20$	1793	$8.57271e-20$
Brown badly scaled 2-d	733	$1.50980e-09$	733	$1.50980e-09$	733	$1.50980e-09$
Beale 2-d	362	$2.47613e-17$	362	$2.47613e-17$	362	$2.47613e-17$
Jennrich and Sampson 2-d	399	$1.24362e+02$	399	$1.24362e+02$	399	$1.24362e+02$
M^c^Kinnon 2-d	650	$-2.50000e-01$	650	$-2.50000e-01$	650	$-2.50000e-01$
Helical valley 3-d	808	$3.38803e-17$	808	$3.38803e-17$	808	$3.38803e-17$
Bard 3-d	2218	$1.74287e+01$	1799	$1.74287e+01$	1801	$1.74287e+01$
Gaussian 3-d	416	$1.12793e-08$	416	$1.12793e-08$	416	$1.12793e-08$
Meyer 3-d	3252	$8.79459e+01$	3252	$8.79459e+01$	3252	$8.79459e+01$
Gulf research 3-d	1510	$1.06763e-16$	1510	$1.06763e-16$	1510	$1.06763e-16$
Box 3-d	796	$3.88191e-16$	796	$3.88191e-16$	796	$3.88191e-16$
Powell singular 4-d	1136	$2.24553e-11$	1136	$2.24553e-11$	1136	$2.24553e-11$
Wood 4-d	1729	$1.84091e-17$	1729	$1.84091e-17$	1729	$1.84091e-17$
Kowalik and Osbourne 4-d	1013	$3.07506e-04$	1013	$3.07506e-04$	1013	$3.07506e-04$
Brown and Dennis 4-d	1453	$8.58222e+04$	1453	$8.58222e+04$	1453	$8.58222e+04$
Quadratic 4-d	909	$7.63455e-18$	909	$7.63455e-18$	909	$7.63455e-18$
Penalty (1) 4-d	2117	$2.24998e-05$	2117	$2.24998e-05$	2117	$2.24998e-05$
Penalty (2) 4-d	5281	$9.37629e-06$	5281	$9.37629e-06$	5281	$9.37629e-06$
Osbourne (1) 5-d	2169	$5.46489e-05$	2170	$5.46489e-05$	2170	$5.46489e-05$
Brown almost linear 5-d	1887	$5.81022e-15$	1887	$5.81022e-15$	1887	$5.81022e-15$
Biggs EXP6 6-d	8977	$4.79768e-16$	8981	$4.79768e-16$	8981	$4.79768e-16$
Extended Rosenbrock 6-d	5814	$6.32421e-14$	5814	$6.32421e-14$	5814	$6.32421e-14$
Brown almost-linear 7-d	4131	$2.57774e-15$	4131	$2.57774e-15$	4131	$2.57774e-15$
Quadratic 8-d	2189	$1.64109e-16$	2189	$1.64109e-16$	2189	$1.64109e-16$
Extended Rosenbrock 8-d	8932	$3.12675e-12$	8935	$3.12675e-12$	8935	$3.12675e-12$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	4978	$2.33073e-14$	4785	$3.67800e-14$	4523	$1.74669e-15$
Extended Powell 8-d	8635	$2.20960e-13$	7704	$8.35058e-15$	7706	$8.35058e-15$
Watson 9-d	12714	$1.39976e-06$	12049	$1.39976e-06$	13659	$1.39976e-06$
Extended Rosenbrock 10-d	13009	$1.57013e-11$	13021	$1.57013e-11$	13021	$1.57013e-11$
Penalty (1) 10-d	20662	$7.08765e-05$	19594	$7.08765e-05$	19478	$7.08765e-05$
Penalty (2) 10-d	27846	$2.93661e-04$	28011	$2.93661e-04$	30038	$2.93661e-04$
Trigonometric 10-d	5993	$2.79506e-05$	5993	$2.79506e-05$	5993	$2.79506e-05$
Osbourne (2) 11-d	13654	$4.01377e-02$	13658	$4.01377e-02$	13658	$4.01377e-02$
Extended Powell 12-d	27170	$8.02697e-14$	22096	$5.19968e-13$	42021	$5.37237e-17$
Quadratic 16-d	5892	$7.63090e-16$	6198	$2.85678e-16$	6089	$2.38569e-16$
Quadratic 24-d	10255	$1.35943e-15$	11005	$1.64832e-16$	10057	$7.27663e-16$

Table G.5: High tolerance results for NM4_ψ_3 δ_4–δ_6.

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	484	$2.32840e-15$	484	$2.32840e-15$	484	$2.32840e-15$
Freudenstein and Roth 2-d	412	$4.89843e+01$	412	$4.89843e+01$	412	$4.89843e+01$
Powell badly scaled 2-d	1793	$8.57271e-20$	1793	$8.57271e-20$	1793	$8.57271e-20$
Brown badly scaled 2-d	733	$1.50980e-09$	733	$1.50980e-09$	733	$1.50980e-09$
Beale 2-d	362	$2.47613e-17$	362	$2.47613e-17$	362	$2.47613e-17$
Jennrich and Sampson 2-d	399	$1.24362e+02$	399	$1.24362e+02$	399	$1.24362e+02$
M^{c}Kinnon 2-d	650	$-2.50000e-01$	650	$-2.50000e-01$	650	$-2.50000e-01$
Helical valley 3-d	808	$3.38803e-17$	808	$3.38803e-17$	808	$3.38803e-17$
Bard 3-d	1802	$1.74287e+01$	1802	$1.74287e+01$	1803	$1.74287e+01$
Gaussian 3-d	416	$1.12793e-08$	416	$1.12793e-08$	416	$1.12793e-08$
Meyer 3-d	3252	$8.79459e+01$	3252	$8.79459e+01$	3252	$8.79459e+01$
Gulf research 3-d	1510	$1.06763e-16$	1510	$1.06763e-16$	1510	$1.06763e-16$
Box 3-d	796	$3.88191e-16$	796	$3.88191e-16$	796	$3.88191e-16$
Powell singular 4-d	1136	$2.24553e-11$	1136	$2.24553e-11$	1136	$2.24553e-11$
Wood 4-d	1729	$1.84091e-17$	1729	$1.84091e-17$	1729	$1.84091e-17$
Kowalik and Osbourne 4-d	1013	$3.07506e-04$	1013	$3.07506e-04$	1013	$3.07506e-04$
Brown and Dennis 4-d	1453	$8.58222e+04$	1453	$8.58222e+04$	1453	$8.58222e+04$
Quadratic 4-d	909	$7.63455e-18$	909	$7.63455e-18$	909	$7.63455e-18$
Penalty (1) 4-d	2117	$2.24998e-05$	2117	$2.24998e-05$	2117	$2.24998e-05$
Penalty (2) 4-d	5281	$9.37629e-06$	5281	$9.37629e-06$	5281	$9.37629e-06$
Osbourne (1) 5-d	2170	$5.46489e-05$	2170	$5.46489e-05$	2170	$5.46489e-05$
Brown almost linear 5-d	1887	$5.81022e-15$	1887	$5.81022e-15$	1887	$5.81022e-15$
Biggs EXP6 6-d	8981	$4.79768e-16$	8981	$4.79768e-16$	8981	$4.79768e-16$
Extended Rosenbrock 6-d	5814	$6.32421e-14$	5814	$6.32421e-14$	5814	$6.32421e-14$
Brown almost-linear 7-d	4131	$2.57774e-15$	4131	$2.57774e-15$	4131	$2.57774e-15$
Quadratic 8-d	2189	$1.64109e-16$	2189	$1.64109e-16$	2189	$1.64109e-16$
Extended Rosenbrock 8-d	8935	$3.12675e-12$	8935	$3.12675e-12$	8935	$3.12675e-12$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	4431	$9.86163e-15$	4432	$9.86163e-15$	4432	$9.86163e-15$
Extended Powell 8-d	7707	$8.35058e-15$	7708	$8.35058e-15$	7708	$8.35058e-15$
Watson 9-d	13526	$1.39976e-06$	13528	$1.39976e-06$	13528	$1.39976e-06$
Extended Rosenbrock 10-d	13021	$1.57013e-11$	13021	$1.57013e-11$	13021	$1.57013e-11$
Penalty (1) 10-d	17971	$7.08765e-05$	21225	$7.08765e-05$	19741	$7.08765e-05$
Penalty (2) 10-d	28812	$2.93661e-04$	29054	$2.93661e-04$	29056	$2.93661e-04$
Trigonometric 10-d	5993	$2.79506e-05$	5993	$2.79506e-05$	5993	$2.79506e-05$
Osbourne (2) 11-d	13658	$4.01377e-02$	13658	$4.01377e-02$	13658	$4.01377e-02$
Extended Powell 12-d	25750	$2.03924e-13$	25760	$2.03924e-13$	25765	$2.03924e-13$
Quadratic 16-d	5722	$8.91285e-16$	5723	$8.91285e-16$	5723	$8.91285e-16$
Quadratic 24-d	10548	$1.58755e-15$	10878	$4.37501e-16$	9629	$1.25873e-15$

Table G.6: High tolerance results for `NM4`_ψ_3 δ_7–δ_9.

G.2 Frame size reduction parameter (κ) using ψ_3, δ_1, ν_1

G.2.1 Low tolerance results

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	257	$2.38228e-07$	296	$4.85725e-07$	378	$3.84598e-09$
Freudenstein and Roth 2-d	236	$4.89843e+01$	257	$4.89843e+01$	357	$4.89843e+01$
Powell badly scaled 2-d	80	$1.18649e-01$	1282	$4.83924e-12$	1541	$1.31087e-13$
Brown badly scaled 2-d	565	$6.16301e-06$	610	$3.07505e-05$	739	$2.18477e-05$
Beale 2-d	147	$1.48861e-08$	195	$2.25494e-10$	258	$1.23485e-09$
Jennrich and Sampson 2-d	148	$1.24362e+02$	178	$1.24363e+02$	210	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	217	$-2.49999e-01$	227	$-2.50000e-01$
Helical valley 3-d	133	$1.11500e-04$	160	$1.29542e-04$	227	$9.37180e-05$
Bard 3-d	1627	$1.74287e+01$	1452	$1.74287e+01$	3303	$1.74287e+01$
Gaussian 3-d	56	$7.72823e-08$	64	$2.76991e-07$	126	$9.53183e-08$
Meyer 3-d	494	$7.38471e+04$	3206	$8.79459e+01$	843	$5.72219e+04$
Gulf research 3-d	605	$5.91366e-13$	1077	$1.41324e-10$	780	$3.06676e-11$
Box 3-d	116	$4.66775e-04$	135	$4.67565e-04$	271	$4.58198e-04$
Powell singular 4-d	840	$7.48685e-09$	755	$3.27267e-06$	936	$8.25584e-09$
Wood 4-d	694	$8.82027e-07$	626	$3.80801e-05$	1130	$7.25170e-05$
Kowalik and Osbourne 4-d	556	$3.07506e-04$	649	$3.07507e-04$	639	$3.07506e-04$
Brown and Dennis 4-d	621	$8.58222e+04$	871	$8.58222e+04$	1003	$8.58222e+04$
Quadratic 4-d	357	$6.13604e-06$	256	$1.91309e-06$	494	$9.50114e-07$
Penalty (1) 4-d	1296	$2.25014e-05$	288	$2.70355e-05$	342	$3.05960e-05$
Penalty (2) 4-d	196	$1.03124e-05$	235	$1.01174e-05$	238	$1.13811e-05$
Osbourne (1) 5-d	186	$1.67604e-03$	229	$5.52534e-04$	339	$1.87391e-03$
Brown almost linear 5-d	748	$2.72595e-09$	173	$1.72332e-03$	853	$2.78773e-07$

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	17870	$4.47083e-03$	4136	$1.15161e-10$	2756	$4.24106e-08$
Extended Rosenbrock 6-d	2504	$2.13309e-04$	3002	$2.07820e-08$	2410	$8.66067e-04$
Brown almost-linear 7-d	1127	$8.71999e-06$	261	$5.19471e-03$	1710	$2.45488e-06$
Quadratic 8-d	748	$3.02333e-07$	1088	$1.56330e-06$	1427	$1.09461e-08$
Extended Rosenbrock 8-d	2699	$4.32538e-04$	3970	$6.35038e-05$	5550	$8.05238e-05$
Variably dimensional 8-d	1971	$4.66647e-06$	2537	$2.25052e-06$	501	$1.08232e+01$
Extended Powell 8-d	3037	$1.15153e-07$	2347	$1.30161e-04$	4066	$1.47782e-06$
Watson 9-d	2692	$7.20350e-04$	2730	$8.69483e-04$	4305	$7.37896e-04$
Extended Rosenbrock 10-d	6180	$3.37291e-04$	4650	$1.29318e-03$	6146	$1.78373e-03$
Penalty (1) 10-d	1116	$8.26326e-05$	1392	$8.09398e-05$	1743	$1.02111e-04$
Penalty (2) 10-d	811	$2.98290e-04$	1098	$2.99373e-04$	1248	$2.97603e-04$
Trigonometric 10-d	2189	$2.80721e-05$	2410	$2.81070e-05$	2306	$2.79830e-05$
Osbourne (2) 11-d	6420	$4.01393e-02$	6717	$4.01410e-02$	8761	$4.01466e-02$
Extended Powell 12-d	5029	$8.42281e-05$	7129	$9.60933e-05$	10937	$1.76545e-04$
Quadratic 16-d	2055	$1.35722e-07$	2183	$2.72998e-08$	3466	$4.30124e-08$
Quadratic 24-d	3936	$1.67122e-07$	4601	$5.95042e-08$	4696	$5.74040e-08$

Table G.7: Low tolerance results for NM4_$\psi_3\delta_1$ κ_1–κ_3.

G.2.2 High tolerance results

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	410	$1.22973e-16$	484	$2.32840e-15$	650	$1.93824e-16$
Freudenstein and Roth 2-d	342	$4.89843e+01$	412	$4.89843e+01$	622	$4.89843e+01$
Powell badly scaled 2-d	1715	$1.17083e-21$	1951	$6.43174e-24$	2017	$8.11188e-19$
Brown badly scaled 2-d	644	$5.35973e-06$	731	$1.50980e-09$	1479	$1.13843e-12$
Beale 2-d	319	$1.46430e-18$	362	$2.47613e-17$	534	$1.15831e-18$
Jennrich and Sampson 2-d	295	$1.24362e+02$	399	$1.24362e+02$	466	$1.24362e+02$
M^{c}Kinnon 2-d	319	$-2.50000e-01$	672	$-2.50000e-01$	792	$-2.50000e-01$
Helical valley 3-d	448	$1.13898e-11$	808	$3.38803e-17$	833	$1.16642e-15$
Bard 3-d	1683	$1.74287e+01$	1612	$1.74287e+01$	3303	$1.74287e+01$
Gaussian 3-d	288	$1.12793e-08$	415	$1.12793e-08$	638	$1.12793e-08$
Meyer 3-d	3417	$8.79459e+01$	3475	$8.79459e+01$	4035	$8.79459e+01$
Gulf research 3-d	872	$6.32883e-17$	1508	$1.06763e-16$	1237	$2.39499e-17$
Box 3-d	837	$7.51573e-20$	1016	$1.26933e-20$	1316	$2.59602e-20$
Powell singular 4-d	1519	$3.94715e-15$	1503	$1.97194e-16$	2059	$4.62060e-15$
Wood 4-d	1215	$3.41173e-13$	1385	$7.74702e-15$	2089	$1.06333e-14$
Kowalik and Osbourne 4-d	1004	$3.07506e-04$	1013	$3.07506e-04$	1440	$3.07506e-04$
Brown and Dennis 4-d	965	$8.58222e+04$	1452	$8.58222e+04$	1974	$8.58222e+04$
Quadratic 4-d	854	$6.93293e-17$	909	$7.63455e-18$	1223	$4.92206e-16$
Penalty (1) 4-d	2208	$2.24998e-05$	2346	$2.24998e-05$	2876	$2.24998e-05$
Penalty (2) 4-d	6046	$9.37629e-06$	5340	$9.37629e-06$	674	$1.02827e-05$
Osbourne (1) 5-d	2283	$5.46489e-05$	2241	$5.46489e-05$	3455	$5.46489e-05$
Brown almost linear 5-d	1499	$2.74539e-16$	1694	$3.27276e-16$	2451	$1.57551e-15$
Biggs EXP6 6-d	18057	$4.47083e-03$	6377	$1.11532e-16$	6330	$1.21614e-15$
Extended Rosenbrock 6-d	4346	$1.07628e-15$	5283	$6.03960e-18$	6186	$5.66552e-13$
Brown almost-linear 7-d	3175	$1.72750e-15$	3344	$9.73250e-16$	5467	$1.48620e-15$

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	1938	$2.55709e-15$	2376	$4.11882e-16$	2980	$5.91018e-16$
Extended Rosenbrock 8-d	7618	$1.97170e-12$	8850	$1.73540e-13$	10268	$8.01756e-13$
Variably dimensional 8-d	3857	$3.17992e-14$	4639	$1.12738e-14$	5728	$2.56280e-14$
Extended Powell 8-d	6930	$3.18056e-15$	8025	$5.01237e-12$	9459	$1.97487e-13$
Watson 9-d	9528	$1.39976e-06$	10928	$1.39976e-06$	15099	$1.39976e-06$
Extended Rosenbrock 10-d	12556	$2.53545e-12$	12979	$1.57300e-13$	16242	$2.83523e-11$
Penalty (1) 10-d	16268	$7.08765e-05$	17571	$7.08765e-05$	25890	$7.08765e-05$
Penalty (2) 10-d	31139	$2.93661e-04$	36125	$2.93661e-04$	26809	$2.93661e-04$
Trigonometric 10-d	4290	$2.79506e-05$	5123	$2.79506e-05$	6850	$2.79506e-05$
Osbourne (2) 11-d	15319	$4.01377e-02$	18071	$4.01377e-02$	25191	$4.01377e-02$
Extended Powell 12-d	14924	$3.16198e-12$	21995	$6.02124e-14$	37685	$2.83427e-13$
Quadratic 16-d	5617	$9.60340e-16$	5150	$6.31179e-16$	7866	$3.94351e-16$
Quadratic 24-d	9984	$4.63211e-15$	9979	$1.12022e-15$	11061	$2.32824e-15$

Table G.8: High tolerance results for NM4_$\psi_3\delta_1$ κ_1–κ_3.

G.3 Frame size reduction parameter (κ) using ψ_3, δ_5, ν_1

G.3.1 Low tolerance results

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	257	$2.38228e-07$	296	$4.85725e-07$	378	$3.84598e-09$
Freudenstein and Roth 2-d	236	$4.89843e+01$	257	$4.89843e+01$	357	$4.89843e+01$
Powell badly scaled 2-d	80	$1.18649e-01$	1217	$1.46409e-10$	1737	$1.63502e-17$
Brown badly scaled 2-d	567	$6.16301e-06$	612	$3.07505e-05$	745	$2.18477e-05$
Beale 2-d	147	$1.48861e-08$	195	$2.25494e-10$	258	$1.23485e-09$
Jennrich and Sampson 2-d	148	$1.24362e+02$	178	$1.24363e+02$	210	$1.24362e+02$
McKinnon 2-d	23	$-6.24961e-05$	217	$-2.49999e-01$	227	$-2.50000e-01$
Helical valley 3-d	133	$1.11500e-04$	160	$1.29542e-04$	227	$9.37180e-05$
Bard 3-d	538	$1.74289e+01$	1625	$1.74287e+01$	4586	$1.74287e+01$
Gaussian 3-d	56	$7.72823e-08$	64	$2.76991e-07$	126	$9.53183e-08$
Meyer 3-d	497	$7.46529e+04$	2910	$8.79459e+01$	3439	$8.79459e+01$
Gulf research 3-d	606	$5.91366e-13$	1079	$1.41324e-10$	780	$3.06676e-11$
Box 3-d	95	$1.69960e-03$	566	$4.98516e-11$	216	$5.46718e-03$
Powell singular 4-d	695	$5.54210e-08$	668	$1.57430e-08$	847	$7.42268e-07$
Wood 4-d	694	$8.82027e-07$	689	$9.52877e-05$	1130	$7.25170e-05$
Kowalik and Osbourne 4-d	556	$3.07506e-04$	649	$3.07507e-04$	639	$3.07506e-04$
Brown and Dennis 4-d	621	$8.58222e+04$	872	$8.58222e+04$	1003	$8.58222e+04$
Quadratic 4-d	357	$6.13604e-06$	256	$1.91309e-06$	494	$9.50114e-07$
Penalty (1) 4-d	257	$3.80543e-05$	287	$2.66985e-05$	343	$3.05960e-05$
Penalty (2) 4-d	196	$1.03124e-05$	236	$1.01174e-05$	238	$1.13811e-05$
Osbourne (1) 5-d	186	$1.67604e-03$	231	$5.52534e-04$	339	$1.87391e-03$
Brown almost linear 5-d	165	$8.76986e-04$	208	$8.68485e-04$	269	$8.56999e-04$

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	3114	$1.03165e-05$	6867	$9.19311e-10$	2611	$5.65565e-03$
Extended Rosenbrock 6-d	2003	$2.65649e-04$	2686	$2.07218e-04$	3040	$6.68353e-05$
Brown almost-linear 7-d	1065	$7.84278e-06$	1662	$1.86730e-06$	1733	$9.82514e-07$
Quadratic 8-d	761	$2.56975e-07$	925	$5.54485e-07$	881	$3.34258e-06$
Extended Rosenbrock 8-d	3689	$2.08893e-04$	3942	$1.24884e-03$	5270	$3.60709e-04$
Variably dimensional 8-d	171	$9.20511e+00$	276	$9.17632e+00$	2518	$2.80911e-07$
Extended Powell 8-d	2167	$2.30543e-05$	2880	$2.94498e-06$	2965	$8.37651e-05$
Watson 9-d	2914	$1.87392e-04$	2655	$1.03023e-03$	3984	$1.84769e-04$
Extended Rosenbrock 10-d	6399	$2.46853e-04$	3388	$1.27839e-02$	6174	$6.62915e-03$
Penalty (1) 10-d	1108	$9.15252e-05$	1330	$9.79506e-05$	1734	$9.73031e-05$
Penalty (2) 10-d	867	$2.97718e-04$	971	$2.99765e-04$	1454	$2.98197e-04$
Trigonometric 10-d	1823	$2.80466e-05$	2418	$2.81070e-05$	2078	$2.79855e-05$
Osbourne (2) 11-d	6129	$4.01413e-02$	5609	$4.01389e-02$	8170	$4.01440e-02$
Extended Powell 12-d	3483	$5.31657e-04$	6805	$1.94398e-04$	10282	$9.78284e-05$
Quadratic 16-d	2378	$9.30813e-08$	2770	$6.88791e-08$	3644	$1.82186e-08$
Quadratic 24-d	4765	$1.33484e-07$	5354	$5.43072e-08$	6025	$4.94540e-08$

Table G.9: Low tolerance results for NM4_$\psi_3\delta_5$ κ_1–κ_3.

G.3.2 High tolerance results

Function	κ_1 FE	κ_1 Minimum	κ_2 FE	κ_2 Minimum	κ_3 FE	κ_3 Minimum
Rosenbrock 2-d	410	$1.22973e-16$	484	$2.32840e-15$	650	$1.93824e-16$
Freudenstein and Roth 2-d	342	$4.89843e+01$	412	$4.89843e+01$	622	$4.89843e+01$
Powell badly scaled 2-d	1627	$1.69783e-21$	1793	$8.57271e-20$	1886	$1.39443e-17$
Brown badly scaled 2-d	646	$5.35973e-06$	733	$1.50980e-09$	1490	$1.13843e-12$
Beale 2-d	319	$1.46430e-18$	362	$2.47613e-17$	534	$1.15831e-18$
Jennrich and Sampson 2-d	295	$1.24362e+02$	399	$1.24362e+02$	466	$1.24362e+02$
M^{c}Kinnon 2-d	319	$-2.50000e-01$	650	$-2.50000e-01$	576	$-2.50000e-01$
Helical valley 3-d	448	$1.13898e-11$	808	$3.38803e-17$	833	$1.16642e-15$
Bard 3-d	585	$1.74289e+01$	1799	$1.74287e+01$	4586	$1.74287e+01$
Gaussian 3-d	288	$1.12793e-08$	416	$1.12793e-08$	638	$1.12793e-08$
Meyer 3-d	3366	$8.79459e+01$	3252	$8.79459e+01$	4203	$8.79459e+01$
Gulf research 3-d	828	$1.14677e-18$	1510	$1.06763e-16$	1417	$4.53545e-19$
Box 3-d	676	$1.24773e-15$	796	$3.88191e-16$	1042	$5.93516e-17$
Powell singular 4-d	1378	$2.18977e-15$	1136	$2.24553e-11$	2066	$1.85367e-15$
Wood 4-d	1217	$3.41173e-13$	1729	$1.84091e-17$	2133	$1.22313e-14$
Kowalik and Osbourne 4-d	1005	$3.07506e-04$	1013	$3.07506e-04$	1441	$3.07506e-04$
Brown and Dennis 4-d	965	$8.58222e+04$	1453	$8.58222e+04$	1974	$8.58222e+04$
Quadratic 4-d	854	$6.93293e-17$	909	$7.63455e-18$	1223	$4.92206e-16$
Penalty (1) 4-d	1981	$2.24998e-05$	2117	$2.24998e-05$	2761	$2.24998e-05$
Penalty (2) 4-d	6053	$9.37629e-06$	5281	$9.37629e-06$	5177	$9.37629e-06$
Osbourne (1) 5-d	1911	$5.46489e-05$	2170	$5.46489e-05$	3206	$5.46489e-05$
Brown almost linear 5-d	1193	$3.44755e-15$	1887	$5.81022e-15$	1876	$6.01802e-15$
Biggs EXP6 6-d	5965	$3.71641e-17$	8981	$4.79768e-16$	3683	$5.65565e-03$
Extended Rosenbrock 6-d	3555	$1.87813e-14$	5814	$6.32421e-14$	6290	$9.37360e-12$
Brown almost-linear 7-d	2456	$1.58620e-15$	4131	$2.57774e-15$	5547	$1.55546e-15$

	κ_1		κ_2		κ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	2034	$5.27897e-17$	2189	$1.64109e-16$	2810	$1.41475e-16$
Extended Rosenbrock 8-d	8341	$5.72152e-16$	8935	$3.12675e-12$	10806	$9.42212e-13$
Variably dimensional 8-d	4934	$4.39972e-13$	4785	$3.67800e-14$	5587	$3.00273e-14$
Extended Powell 8-d	10902	$2.26254e-18$	7704	$8.35058e-15$	15531	$9.46343e-19$
Watson 9-d	11317	$1.39976e-06$	12049	$1.39976e-06$	17298	$1.39976e-06$
Extended Rosenbrock 10-d	13790	$3.60446e-14$	13021	$1.57013e-11$	15417	$1.55563e-11$
Penalty (1) 10-d	13684	$7.08765e-05$	19594	$7.08765e-05$	23223	$7.08765e-05$
Penalty (2) 10-d	30660	$2.93661e-04$	28011	$2.93661e-04$	31535	$2.93661e-04$
Trigonometric 10-d	4285	$2.79506e-05$	5993	$2.79506e-05$	6261	$2.79506e-05$
Osbourne (2) 11-d	12111	$4.01377e-02$	13658	$4.01377e-02$	20665	$4.01377e-02$
Extended Powell 12-d	21468	$5.90556e-12$	22096	$5.19968e-13$	39635	$6.66801e-16$
Quadratic 16-d	5496	$3.20851e-16$	6198	$2.85678e-16$	7156	$7.32654e-16$
Quadratic 24-d	11507	$3.51310e-16$	11005	$1.64832e-16$	13118	$1.86017e-16$

Table G.10: High tolerance results for NM4_$\psi_3\delta_5$ κ_1–κ_3.

G.4 Epsilon reduction parameter (ν) using $\psi_3, \delta_1, \kappa_2$

G.4.1 Low tolerance results

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	296	4.85725e-07	223	8.51959e-07	242	2.30387e-09
Freudenstein and Roth 2-d	257	4.89843e+01	277	4.89843e+01	187	4.89843e+01
Powell badly scaled 2-d	1282	4.83924e-12	920	4.03699e-10	1148	7.64836e-17
Brown badly scaled 2-d	610	3.07505e-05	702	2.26370e-08	560	2.61629e-07
Beale 2-d	195	2.25494e-10	137	2.85559e-08	164	2.64241e-10
Jennrich and Sampson 2-d	178	1.24363e+02	178	1.24362e+02	140	1.24362e+02
McKinnon 2-d	217	-2.49999e-01	217	-2.49999e-01	198	-2.49999e-01
Helical valley 3-d	160	1.29542e-04	147	1.64096e-04	172	1.18251e-04
Bard 3-d	1452	1.74287e+01	2201	1.74287e+01	1731	1.74287e+01
Gaussian 3-d	64	2.76991e-07	75	1.27329e-08	85	1.30383e-08
Meyer 3-d	3206	8.79459e+01	757	7.06057e+04	2465	8.79459e+01
Gulf research 3-d	1077	1.41324e-10	1010	1.11929e-10	962	5.84050e-11
Box 3-d	135	4.67565e-04	155	4.60009e-04	397	6.97578e-10
Powell singular 4-d	755	3.27267e-06	803	1.21062e-09	855	1.67554e-10
Wood 4-d	626	3.80801e-05	667	3.73478e-05	610	1.77895e-08
Kowalik and Osbourne 4-d	649	3.07507e-04	593	3.07506e-04	633	3.07506e-04
Brown and Dennis 4-d	871	8.58222e+04	781	8.58222e+04	708	8.58222e+04
Quadratic 4-d	256	1.91309e-06	362	4.81968e-09	288	3.07821e-09
Penalty (1) 4-d	288	2.70355e-05	1326	2.25091e-05	1747	2.24998e-05
Penalty (2) 4-d	235	1.01174e-05	253	1.01694e-05	182	1.02464e-05
Osbourne (1) 5-d	229	5.52534e-04	546	7.73230e-05	580	7.78932e-05
Brown almost linear 5-d	173	1.72332e-03	764	4.25048e-09	715	1.03333e-07

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	4136	$1.15161e-10$	2057	$5.00666e-07$	5800	$1.39907e-09$
Extended Rosenbrock 6-d	3002	$2.07820e-08$	2215	$7.51073e-06$	2657	$1.06386e-05$
Brown almost-linear 7-d	261	$5.19471e-03$	1084	$1.13377e-07$	1588	$5.17279e-08$
Quadratic 8-d	1088	$1.56330e-06$	1148	$1.77146e-09$	734	$6.50866e-09$
Extended Rosenbrock 8-d	3970	$6.35038e-05$	4312	$1.91119e-06$	2509	$7.03382e-05$
Variably dimensional 8-d	2537	$2.25052e-06$	2236	$1.83113e-07$	2264	$7.02952e-08$
Extended Powell 8-d	2347	$1.30161e-04$	2596	$6.67461e-06$	3515	$1.27426e-06$
Watson 9-d	2730	$8.69483e-04$	4604	$6.16560e-05$	4012	$5.75290e-05$
Extended Rosenbrock 10-d	4650	$1.29318e-03$	5295	$3.09084e-04$	7604	$4.88782e-06$
Penalty (1) 10-d	1392	$8.09398e-05$	1112	$7.21273e-05$	5936	$7.09829e-05$
Penalty (2) 10-d	1098	$2.99373e-04$	644	$2.98717e-04$	758	$2.98831e-04$
Trigonometric 10-d	2410	$2.81070e-05$	1672	$2.79665e-05$	1447	$1.27858e-08$
Osbourne (2) 11-d	6717	$4.01410e-02$	5903	$4.01467e-02$	5858	$4.01389e-02$
Extended Powell 12-d	7129	$9.60933e-05$	5624	$9.27047e-05$	5313	$3.48377e-05$
Quadratic 16-d	2183	$2.72998e-08$	2151	$2.60938e-08$	1460	$7.95837e-08$
Quadratic 24-d	4601	$5.95042e-08$	3597	$1.74238e-07$	3711	$5.84648e-08$

Table G.11: Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_1–ν_3.

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	268	$1.52988e-10$	253	$6.48974e-11$	241	$7.85188e-10$
Freudenstein and Roth 2-d	191	$4.89843e+01$	243	$4.89843e+01$	198	$4.89843e+01$
Powell badly scaled 2-d	1083	$1.04017e-15$	371	$1.00121e-06$	956	$2.83937e-17$
Brown badly scaled 2-d	501	$1.37273e-07$	454	$1.38341e-06$	448	$8.91906e-07$
Beale 2-d	134	$6.39394e-10$	139	$1.90327e-10$	150	$5.37760e-10$
Jennrich and Sampson 2-d	153	$1.24362e+02$	126	$1.24362e+02$	129	$1.24362e+02$
M^{c}Kinnon 2-d	198	$-2.49999e-01$	196	$-2.49999e-01$	196	$-2.50000e-01$
Helical valley 3-d	166	$1.33035e-04$	168	$1.33141e-04$	162	$1.22152e-04$
Bard 3-d	1848	$1.74287e+01$	1707	$1.74287e+01$	1772	$1.74287e+01$
Gaussian 3-d	74	$1.79492e-08$	98	$1.33039e-08$	82	$1.47402e-08$
Meyer 3-d	2516	$8.79459e+01$	2171	$8.79459e+01$	2205	$8.79459e+01$
Gulf research 3-d	1029	$1.10374e-11$	873	$1.47940e-11$	970	$1.68163e-13$
Box 3-d	479	$5.37202e-14$	491	$5.25980e-14$	428	$1.92764e-13$
Powell singular 4-d	514	$2.42837e-07$	733	$1.81482e-11$	427	$3.40915e-07$
Wood 4-d	589	$1.15889e-06$	596	$7.68330e-07$	628	$5.04459e-09$
Kowalik and Osbourne 4-d	466	$3.07508e-04$	486	$3.07506e-04$	461	$3.07506e-04$
Brown and Dennis 4-d	507	$8.58222e+04$	559	$8.58222e+04$	480	$8.58222e+04$
Quadratic 4-d	304	$1.48528e-09$	277	$1.71843e-09$	251	$3.24529e-09$
Penalty (1) 4-d	1650	$2.24998e-05$	1150	$2.25842e-05$	2078	$2.24998e-05$
Penalty (2) 4-d	199	$1.02450e-05$	222	$1.03000e-05$	3180	$9.48972e-06$
Osbourne (1) 5-d	381	$7.00832e-05$	277	$8.18200e-04$	385	$6.95615e-05$
Brown almost linear 5-d	641	$5.84918e-10$	473	$2.10788e-08$	550	$1.18897e-08$
Biggs EXP6 6-d	3828	$2.72455e-12$	1761	$5.65565e-03$	41366	$4.46835e-03$
Extended Rosenbrock 6-d	2528	$9.75830e-06$	2391	$3.60650e-08$	1829	$2.43348e-07$
Brown almost-linear 7-d	1344	$1.78756e-08$	1526	$7.26466e-10$	1249	$5.09949e-09$
Quadratic 8-d	705	$1.13152e-08$	712	$3.17092e-09$	619	$1.91993e-08$
Extended Rosenbrock 8-d	3643	$1.00506e-04$	3906	$3.65924e-09$	2694	$1.43040e-07$

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1925	$6.37636e-08$	2032	$9.24811e-08$	1778	$9.64017e-07$
Extended Powell 8-d	2267	$2.13305e-06$	1968	$5.90042e-06$	2150	$1.79807e-08$
Watson 9-d	6849	$7.81264e-06$	4013	$1.00475e-04$	2619	$7.55580e-05$
Extended Rosenbrock 10-d	5079	$4.97163e-05$	6239	$1.09363e-05$	4862	$1.11875e-06$
Penalty (1) 10-d	6802	$7.08790e-05$	5818	$7.08809e-05$	7660	$7.08785e-05$
Penalty (2) 10-d	815	$2.98794e-04$	672	$2.98761e-04$	6089	$2.95240e-04$
Trigonometric 10-d	1771	$2.79536e-05$	1443	$2.79522e-05$	1521	$2.79524e-05$
Osbourne (2) 11-d	6476	$4.01381e-02$	5034	$4.01378e-02$	5387	$4.01378e-02$
Extended Powell 12-d	4644	$1.89127e-05$	3973	$2.21982e-06$	6038	$1.21519e-06$
Quadratic 16-d	1609	$2.13513e-08$	1602	$2.36483e-08$	1481	$1.04434e-08$
Quadratic 24-d	3388	$4.53230e-08$	3356	$5.85629e-08$	2946	$2.77000e-08$

Table G.12: Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_4–ν_6.

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	235	$3.97205e-10$	241	$1.13562e-09$	234	$4.86709e-10$
Freudenstein and Roth 2-d	159	$4.89843e+01$	160	$4.89843e+01$	159	$4.89843e+01$
Powell badly scaled 2-d	966	$1.87067e-16$	885	$2.65600e-15$	894	$2.34156e-17$
Brown badly scaled 2-d	443	$1.36614e-09$	464	$3.67093e-09$	436	$7.24521e-10$
Beale 2-d	147	$1.12248e-10$	128	$1.86459e-09$	136	$1.02850e-09$
Jennrich and Sampson 2-d	124	$1.24362e+02$	100	$1.24362e+02$	108	$1.24362e+02$
M^{c}Kinnon 2-d	195	$-2.50000e-01$	195	$-2.50000e-01$	195	$-2.50000e-01$
Helical valley 3-d	142	$1.26975e-04$	165	$1.11853e-04$	149	$1.10399e-04$
Bard 3-d	1226	$1.74287e+01$	1442	$1.74287e+01$	1373	$1.74287e+01$
Gaussian 3-d	80	$1.15947e-08$	77	$1.59743e-08$	68	$1.55967e-08$
Meyer 3-d	2205	$8.79459e+01$	2097	$8.79459e+01$	2070	$8.79459e+01$
Gulf research 3-d	998	$1.91065e-14$	956	$1.12798e-14$	948	$6.00756e-12$
Box 3-d	404	$2.46851e-14$	371	$1.49473e-13$	412	$3.07154e-13$
Powell singular 4-d	656	$1.14857e-11$	433	$2.72831e-08$	456	$3.88515e-09$
Wood 4-d	564	$1.27432e-08$	570	$1.38215e-08$	490	$1.15954e-09$
Kowalik and Osbourne 4-d	384	$3.07506e-04$	422	$3.07506e-04$	441	$3.07506e-04$
Brown and Dennis 4-d	476	$8.58222e+04$	492	$8.58222e+04$	419	$8.58222e+04$
Quadratic 4-d	261	$2.17860e-09$	257	$2.17267e-09$	245	$2.56693e-09$
Penalty (1) 4-d	1192	$2.25000e-05$	1246	$2.25412e-05$	1400	$2.24998e-05$
Penalty (2) 4-d	189	$1.02610e-05$	200	$1.02580e-05$	194	$1.02581e-05$
Osbourne (1) 5-d	287	$6.92059e-05$	291	$6.91831e-05$	284	$7.18492e-05$
Brown almost linear 5-d	492	$1.07936e-09$	501	$4.22838e-10$	475	$8.42071e-09$
Biggs EXP6 6-d	1363	$9.50171e-08$	2819	$3.03508e-11$	2100	$3.42136e-13$
Extended Rosenbrock 6-d	1907	$7.71469e-09$	2183	$3.59780e-07$	2102	$2.56407e-06$
Brown almost-linear 7-d	997	$4.46297e-08$	989	$1.68953e-09$	881	$1.07064e-08$
Quadratic 8-d	646	$3.13853e-09$	677	$2.10376e-09$	687	$7.77695e-09$
Extended Rosenbrock 8-d	3582	$5.29385e-06$	3303	$3.51475e-07$	4182	$1.00191e-07$

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1930	$2.80129e-07$	1579	$3.04271e-08$	1720	$1.58055e-07$
Extended Powell 8-d	3005	$6.41927e-09$	1856	$1.16813e-09$	1303	$3.75701e-07$
Watson 9-d	2018	$8.04480e-04$	2069	$7.97411e-04$	3910	$4.38811e-06$
Extended Rosenbrock 10-d	7759	$5.65980e-05$	5791	$3.63214e-04$	8400	$1.81936e-06$
Penalty (1) 10-d	8096	$7.08769e-05$	7316	$7.08767e-05$	5847	$7.08863e-05$
Penalty (2) 10-d	2663	$2.97603e-04$	2307	$2.97884e-04$	712	$2.99529e-04$
Trigonometric 10-d	1547	$2.79512e-05$	1170	$2.79905e-05$	1363	$2.79526e-05$
Osbourne (2) 11-d	5684	$4.01378e-02$	3744	$4.01560e-02$	3671	$4.01429e-02$
Extended Powell 12-d	5554	$4.15572e-07$	3930	$2.91125e-06$	2075	$1.60981e-05$
Quadratic 16-d	1654	$1.12534e-08$	1488	$6.09802e-09$	1426	$1.10148e-08$
Quadratic 24-d	2565	$1.57047e-08$	2795	$1.60197e-08$	2772	$7.85127e-08$

Table G.13: Low tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_7–ν_9.

G.4.2 High tolerance results

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	484	$2.32840e-15$	412	$4.53983e-16$	356	$1.13882e-17$
Freudenstein and Roth 2-d	412	$4.89843e+01$	394	$4.89843e+01$	314	$4.89843e+01$
Powell badly scaled 2-d	1951	$6.43174e-24$	1611	$3.82732e-24$	1339	$2.34004e-20$
Brown badly scaled 2-d	731	$1.50980e-09$	986	$2.89203e-15$	845	$3.32073e-15$
Beale 2-d	362	$2.47613e-17$	306	$2.46368e-19$	299	$1.13709e-19$
Jennrich and Sampson 2-d	399	$1.24362e+02$	304	$1.24362e+02$	267	$1.24362e+02$
M^{c}Kinnon 2-d	672	$-2.50000e-01$	786	$-2.50000e-01$	510	$-2.50000e-01$
Helical valley 3-d	808	$3.38803e-17$	672	$1.15570e-12$	651	$4.49923e-17$
Bard 3-d	1612	$1.74287e+01$	2253	$1.74287e+01$	1783	$1.74287e+01$
Gaussian 3-d	415	$1.12793e-08$	373	$1.12793e-08$	310	$1.12793e-08$
Meyer 3-d	3475	$8.79459e+01$	3506	$8.79459e+01$	2701	$8.79459e+01$
Gulf research 3-d	1508	$1.06763e-16$	1566	$7.68178e-21$	1302	$1.84026e-18$
Box 3-d	1016	$1.26933e-20$	948	$5.97472e-20$	798	$2.21622e-20$
Powell singular 4-d	1503	$1.97194e-16$	2358	$2.73588e-21$	1494	$5.68242e-19$
Wood 4-d	1385	$7.74702e-15$	1346	$5.95848e-15$	987	$7.83343e-13$
Kowalik and Osbourne 4-d	1013	$3.07506e-04$	1036	$3.07506e-04$	1146	$3.07506e-04$
Brown and Dennis 4-d	1452	$8.58222e+04$	1213	$8.58222e+04$	1088	$8.58222e+04$
Quadratic 4-d	909	$7.63455e-18$	789	$7.37100e-17$	512	$5.23441e-17$
Penalty (1) 4-d	2346	$2.24998e-05$	2242	$2.24998e-05$	2286	$2.24998e-05$
Penalty (2) 4-d	5340	$9.37629e-06$	5013	$9.37629e-06$	4032	$9.37629e-06$
Osbourne (1) 5-d	2241	$5.46489e-05$	2125	$5.46489e-05$	2319	$5.46489e-05$
Brown almost linear 5-d	1694	$3.27276e-16$	1695	$2.85522e-16$	1588	$6.85629e-17$
Biggs EXP6 6-d	6377	$1.11532e-16$	3751	$6.51359e-14$	7290	$1.70465e-18$
Extended Rosenbrock 6-d	5283	$6.03960e-18$	4764	$2.14444e-16$	4074	$4.44789e-17$
Brown almost-linear 7-d	3344	$9.73250e-16$	3193	$5.10119e-15$	2986	$6.72845e-16$

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	2376	$4.11882e-16$	1907	$1.75104e-16$	1371	$3.97634e-16$
Extended Rosenbrock 8-d	8850	$1.73540e-13$	7885	$2.84191e-13$	5552	$1.98091e-12$
Variably dimensional 8-d	4639	$1.12738e-14$	4312	$1.08539e-14$	4027	$5.84557e-15$
Extended Powell 8-d	8025	$5.01237e-12$	6822	$1.57694e-14$	11786	$1.61656e-22$
Watson 9-d	10928	$1.39976e-06$	11582	$1.39976e-06$	10744	$1.39976e-06$
Extended Rosenbrock 10-d	12979	$1.57300e-13$	12293	$1.22336e-11$	11369	$8.32038e-12$
Penalty (1) 10-d	17571	$7.08765e-05$	15159	$7.08765e-05$	14010	$7.08765e-05$
Penalty (2) 10-d	36125	$2.93661e-04$	31741	$2.93661e-04$	39519	$2.93661e-04$
Trigonometric 10-d	5123	$2.79506e-05$	4087	$2.79506e-05$	2930	$2.20935e-15$
Osbourne (2) 11-d	18071	$4.01377e-02$	14344	$4.01377e-02$	10523	$4.01377e-02$
Extended Powell 12-d	21995	$6.02124e-14$	22492	$3.74256e-13$	31654	$6.33881e-13$
Quadratic 16-d	5150	$6.31179e-16$	4793	$7.03944e-16$	3998	$9.95725e-17$
Quadratic 24-d	9979	$1.12022e-15$	8748	$5.73579e-16$	7858	$2.94461e-16$

Table G.14: High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_1–ν_3.

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	389	$2.31822e-17$	356	$1.65123e-17$	361	$1.37189e-18$
Freudenstein and Roth 2-d	285	$4.89843e+01$	339	$4.89843e+01$	301	$4.89843e+01$
Powell badly scaled 2-d	1396	$2.84953e-23$	1141	$1.57864e-24$	1060	$4.09299e-21$
Brown badly scaled 2-d	668	$1.29269e-15$	709	$2.03479e-16$	636	$3.25680e-15$
Beale 2-d	242	$2.65528e-18$	220	$7.51790e-19$	247	$5.04851e-18$
Jennrich and Sampson 2-d	245	$1.24362e+02$	237	$1.24362e+02$	233	$1.24362e+02$
M^{c}Kinnon 2-d	498	$-2.50000e-01$	600	$-2.50000e-01$	388	$-2.50000e-01$
Helical valley 3-d	472	$6.96323e-15$	447	$2.14733e-15$	445	$2.48901e-16$
Bard 3-d	1848	$1.74287e+01$	1759	$1.74287e+01$	1820	$1.74287e+01$
Gaussian 3-d	278	$1.12793e-08$	262	$1.12793e-08$	247	$1.12793e-08$
Meyer 3-d	2725	$8.79459e+01$	2528	$8.79459e+01$	2439	$8.79459e+01$
Gulf research 3-d	1256	$1.69336e-20$	1149	$1.17662e-22$	1148	$3.14930e-22$
Box 3-d	625	$4.30526e-21$	648	$3.40616e-22$	602	$3.64058e-21$
Powell singular 4-d	1703	$5.83488e-23$	1504	$1.16844e-20$	1102	$4.64151e-20$
Wood 4-d	934	$3.18660e-14$	945	$7.03763e-15$	893	$4.73838e-15$
Kowalik and Osbourne 4-d	1186	$3.07506e-04$	824	$3.07506e-04$	730	$3.07506e-04$
Brown and Dennis 4-d	890	$8.58222e+04$	1036	$8.58222e+04$	858	$8.58222e+04$
Quadratic 4-d	491	$5.26087e-17$	484	$1.74450e-17$	459	$1.88539e-17$
Penalty (1) 4-d	1928	$2.24998e-05$	1700	$2.24998e-05$	2491	$2.24998e-05$
Penalty (2) 4-d	4071	$9.37629e-06$	7562	$9.37629e-06$	5558	$9.37629e-06$
Osbourne (1) 5-d	1778	$5.46489e-05$	1923	$5.46489e-05$	1865	$5.46489e-05$
Brown almost linear 5-d	1039	$6.45813e-16$	1036	$2.00020e-17$	927	$2.91970e-18$
Biggs EXP6 6-d	4842	$3.98077e-19$	2102	$5.65565e-03$	41938	$4.46835e-03$
Extended Rosenbrock 6-d	4123	$5.20544e-15$	3187	$3.33797e-16$	2567	$2.13042e-14$
Brown almost-linear 7-d	2711	$3.25387e-17$	2301	$4.50557e-14$	2094	$7.18870e-16$
Quadratic 8-d	1177	$3.20943e-17$	1081	$2.32734e-17$	1050	$4.44875e-17$
Extended Rosenbrock 8-d	7491	$2.19369e-13$	4761	$1.28246e-11$	5563	$2.26846e-14$

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	3355	$3.04949e-14$	3251	$7.24861e-15$	3673	$6.50023e-16$
Extended Powell 8-d	7287	$1.04775e-17$	7716	$1.97821e-15$	3962	$1.11816e-14$
Watson 9-d	10334	$1.39976e-06$	8657	$1.39976e-06$	8219	$1.39976e-06$
Extended Rosenbrock 10-d	11585	$6.96586e-12$	11718	$3.30962e-14$	12357	$2.43371e-16$
Penalty (1) 10-d	13972	$7.08765e-05$	13828	$7.08765e-05$	12635	$7.08765e-05$
Penalty (2) 10-d	36132	$2.93661e-04$	23068	$2.93661e-04$	30898	$2.93661e-04$
Trigonometric 10-d	2913	$2.79506e-05$	2591	$2.79506e-05$	2605	$2.79506e-05$
Osbourne (2) 11-d	12991	$4.01377e-02$	8173	$4.01377e-02$	9902	$4.01377e-02$
Extended Powell 12-d	21629	$3.79328e-20$	14300	$4.83936e-15$	11807	$4.54209e-16$
Quadratic 16-d	2735	$2.52427e-16$	2672	$1.90226e-16$	2258	$1.38431e-16$
Quadratic 24-d	6539	$3.05187e-16$	6332	$3.72108e-16$	5039	$5.08518e-16$

Table G.15: High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_4–ν_6.

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	332	$8.36370e-18$	330	$2.90725e-17$	321	$1.01964e-17$
Freudenstein and Roth 2-d	250	$4.89843e+01$	257	$4.89843e+01$	256	$4.89843e+01$
Powell badly scaled 2-d	1221	$3.07064e-25$	1052	$5.81912e-26$	1069	$5.16903e-24$
Brown badly scaled 2-d	624	$1.12912e-17$	581	$4.31731e-17$	608	$1.80441e-18$
Beale 2-d	256	$2.12382e-18$	224	$3.95699e-18$	226	$1.13685e-17$
Jennrich and Sampson 2-d	198	$1.24362e+02$	207	$1.24362e+02$	201	$1.24362e+02$
M^{c}Kinnon 2-d	355	$-2.50000e-01$	411	$-2.50000e-01$	459	$-2.50000e-01$
Helical valley 3-d	507	$4.05819e-16$	447	$5.66748e-17$	399	$1.66988e-16$
Bard 3-d	1278	$1.74287e+01$	1498	$1.74287e+01$	1425	$1.74287e+01$
Gaussian 3-d	221	$1.12793e-08$	249	$1.12793e-08$	230	$1.12793e-08$
Meyer 3-d	2493	$8.79459e+01$	2471	$8.79459e+01$	2277	$8.79459e+01$
Gulf research 3-d	1170	$2.24305e-21$	1112	$7.11219e-22$	1135	$1.69454e-22$
Box 3-d	540	$1.77016e-21$	544	$9.80770e-22$	572	$5.34279e-22$
Powell singular 4-d	1545	$9.44044e-21$	823	$5.99302e-17$	1232	$1.41825e-28$
Wood 4-d	789	$3.71270e-17$	787	$2.68521e-16$	761	$1.35725e-16$
Kowalik and Osbourne 4-d	591	$3.07506e-04$	628	$3.07506e-04$	684	$3.07506e-04$
Brown and Dennis 4-d	814	$8.58222e+04$	801	$8.58222e+04$	770	$8.58222e+04$
Quadratic 4-d	435	$2.63137e-17$	438	$2.18326e-17$	410	$1.28774e-17$
Penalty (1) 4-d	1633	$2.24998e-05$	1718	$2.24998e-05$	1646	$2.24998e-05$
Penalty (2) 4-d	4644	$9.37629e-06$	5538	$9.37629e-06$	5160	$9.37629e-06$
Osbourne (1) 5-d	1665	$5.46489e-05$	1862	$5.46489e-05$	1573	$5.46489e-05$
Brown almost linear 5-d	986	$3.12588e-18$	865	$6.44921e-18$	842	$9.75922e-18$
Biggs EXP6 6-d	2844	$7.02480e-19$	3372	$2.93965e-16$	2514	$3.62766e-21$
Extended Rosenbrock 6-d	2657	$1.01827e-16$	3337	$2.46270e-17$	3319	$7.82495e-18$
Brown almost-linear 7-d	1891	$4.36519e-14$	1845	$1.63582e-17$	1491	$9.02801e-16$
Quadratic 8-d	1138	$4.14933e-17$	1026	$3.30772e-17$	1065	$4.82527e-17$
Extended Rosenbrock 8-d	4928	$1.88592e-14$	4923	$5.02697e-14$	6046	$4.42140e-14$

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	3309	$3.82904e-15$	2136	$7.76437e-16$	2815	$1.35441e-15$
Extended Powell 8-d	6514	$4.21808e-19$	7003	$1.01817e-23$	5880	$2.73547e-19$
Watson 9-d	8337	$1.39976e-06$	7637	$1.39976e-06$	6497	$1.39976e-06$
Extended Rosenbrock 10-d	13219	$4.18340e-14$	12273	$1.61152e-16$	10958	$9.78473e-15$
Penalty (1) 10-d	10966	$7.08765e-05$	11651	$7.08765e-05$	10982	$7.08765e-05$
Penalty (2) 10-d	27977	$2.93661e-04$	27122	$2.93661e-04$	41366	$2.93661e-04$
Trigonometric 10-d	2651	$2.79506e-05$	2380	$2.79506e-05$	2146	$2.79506e-05$
Osbourne (2) 11-d	11042	$4.01377e-02$	7342	$4.01377e-02$	7331	$4.01377e-02$
Extended Powell 12-d	24320	$6.10753e-18$	26246	$3.52429e-13$	26917	$1.80468e-21$
Quadratic 16-d	2956	$7.61630e-16$	2269	$1.89810e-16$	2451	$4.43509e-16$
Quadratic 24-d	4215	$2.07933e-16$	4330	$3.30148e-16$	4680	$1.31183e-16$

Table G.16: High tolerance results for NM4_$\psi_3\delta_1\kappa_2$ ν_6–ν_9.

G.5 Epsilon reduction parameter (ν) using $\psi_3, \delta_5, \kappa_1$

G.5.1 Low tolerance results

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	257	$2.38228e-07$	255	$9.01937e-10$	208	$1.03505e-09$
Freudenstein and Roth 2-d	236	$4.89843e+01$	135	$4.89843e+01$	153	$4.89843e+01$
Powell badly scaled 2-d	80	$1.18649e-01$	153	$1.52888e-02$	1129	$1.48362e-16$
Brown badly scaled 2-d	567	$6.16301e-06$	359	$8.38323e-04$	489	$2.71265e-08$
Beale 2-d	147	$1.48861e-08$	116	$4.58327e-08$	126	$3.47827e-10$
Jennrich and Sampson 2-d	148	$1.24362e+02$	116	$1.24362e+02$	123	$1.24362e+02$
McKinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	133	$1.11500e-04$	118	$2.15074e-04$	131	$8.80903e-05$
Bard 3-d	538	$1.74289e+01$	1337	$1.74287e+01$	1243	$1.74287e+01$
Gaussian 3-d	56	$7.72823e-08$	55	$1.92032e-08$	66	$1.35140e-08$
Meyer 3-d	497	$7.46529e+04$	3423	$8.79459e+01$	2761	$8.79459e+01$
Gulf research 3-d	606	$5.91366e-13$	530	$3.05112e-10$	487	$1.88644e-10$
Box 3-d	95	$1.69960e-03$	412	$1.65098e-10$	451	$4.94988e-14$
Powell singular 4-d	695	$5.54210e-08$	558	$2.72987e-07$	636	$3.68062e-09$
Wood 4-d	694	$8.82027e-07$	626	$4.37998e-06$	529	$2.86526e-09$
Kowalik and Osbourne 4-d	556	$3.07506e-04$	619	$3.07506e-04$	566	$3.07506e-04$
Brown and Dennis 4-d	621	$8.58222e+04$	660	$8.58222e+04$	511	$8.58222e+04$
Quadratic 4-d	357	$6.13604e-06$	320	$2.76388e-07$	309	$7.35070e-08$
Penalty (1) 4-d	257	$3.80543e-05$	2174	$2.24998e-05$	1176	$2.24998e-05$
Penalty (2) 4-d	196	$1.03124e-05$	197	$1.02890e-05$	220	$1.02819e-05$
Osbourne (1) 5-d	186	$1.67604e-03$	185	$1.65693e-03$	462	$7.35766e-05$
Brown almost linear 5-d	165	$8.76986e-04$	548	$1.74375e-09$	590	$2.26014e-08$

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	3114	$1.03165e-05$	26864	$4.46883e-03$	1611	$6.02893e-06$
Extended Rosenbrock 6-d	2003	$2.65649e-04$	2171	$1.82318e-07$	1975	$2.05968e-08$
Brown almost-linear 7-d	1065	$7.84278e-06$	905	$1.07244e-06$	1124	$2.11863e-07$
Quadratic 8-d	761	$2.56975e-07$	850	$1.67271e-08$	694	$8.34838e-08$
Extended Rosenbrock 8-d	3689	$2.08893e-04$	3139	$4.82162e-04$	4833	$4.89349e-06$
Variably dimensional 8-d	171	$9.20511e+00$	1596	$1.67600e-06$	1602	$3.49265e-07$
Extended Powell 8-d	2167	$2.30543e-05$	2050	$3.95467e-05$	1908	$2.49984e-06$
Watson 9-d	2914	$1.87392e-04$	2465	$4.08703e-05$	3583	$2.31365e-05$
Extended Rosenbrock 10-d	6399	$2.46853e-04$	5549	$4.85959e-05$	6936	$1.90485e-04$
Penalty (1) 10-d	1108	$9.15252e-05$	9495	$7.08781e-05$	5551	$7.09380e-05$
Penalty (2) 10-d	867	$2.97718e-04$	611	$2.98518e-04$	640	$2.98565e-04$
Trigonometric 10-d	1823	$2.80466e-05$	1175	$2.79751e-05$	1417	$2.79819e-05$
Osbourne (2) 11-d	6129	$4.01413e-02$	5716	$4.01473e-02$	6201	$4.01378e-02$
Extended Powell 12-d	3483	$5.31657e-04$	5282	$4.54185e-05$	5207	$2.82462e-05$
Quadratic 16-d	2378	$9.30813e-08$	2302	$3.09824e-07$	1905	$6.52876e-08$
Quadratic 24-d	4765	$1.33484e-07$	4335	$2.07085e-07$	3903	$1.35975e-07$

Table G.17: Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_1–ν_3.

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	237	$1.89190e-09$	221	$2.01550e-09$	222	$5.50116e-10$
Freudenstein and Roth 2-d	164	$4.89843e+01$	150	$4.89843e+01$	144	$4.89843e+01$
Powell badly scaled 2-d	877	$1.57901e-15$	813	$9.98991e-14$	881	$6.80758e-18$
Brown badly scaled 2-d	415	$3.50585e-08$	365	$1.15547e-09$	340	$1.25506e-05$
Beale 2-d	127	$3.80261e-10$	114	$7.68905e-10$	120	$2.92376e-10$
Jennrich and Sampson 2-d	113	$1.24362e+02$	108	$1.24362e+02$	101	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	129	$8.77705e-05$	131	$8.83408e-05$	141	$9.59747e-05$
Bard 3-d	1096	$1.74287e+01$	1109	$1.74287e+01$	961	$1.74287e+01$
Gaussian 3-d	62	$1.46188e-08$	72	$1.15828e-08$	78	$1.13854e-08$
Meyer 3-d	2526	$8.79460e+01$	2709	$8.79459e+01$	2545	$8.79459e+01$
Gulf research 3-d	486	$3.61830e-14$	506	$3.82246e-13$	401	$5.37308e-12$
Box 3-d	387	$4.55867e-13$	328	$7.50267e-13$	403	$1.05915e-13$
Powell singular 4-d	412	$1.07922e-10$	602	$1.32290e-10$	339	$3.12783e-09$
Wood 4-d	474	$5.33733e-07$	490	$3.13917e-08$	439	$4.71375e-07$
Kowalik and Osbourne 4-d	435	$3.07506e-04$	323	$3.07506e-04$	323	$3.07506e-04$
Brown and Dennis 4-d	415	$8.58222e+04$	381	$8.58222e+04$	340	$8.58222e+04$
Quadratic 4-d	319	$1.60115e-09$	309	$6.05717e-09$	302	$1.82193e-09$
Penalty (1) 4-d	1831	$2.24998e-05$	1751	$2.24998e-05$	1467	$2.24998e-05$
Penalty (2) 4-d	184	$1.02572e-05$	189	$1.02541e-05$	2170	$9.57794e-06$
Osbourne (1) 5-d	543	$7.35458e-05$	363	$7.10772e-05$	185	$7.25405e-05$
Brown almost linear 5-d	543	$1.68930e-09$	473	$6.29718e-09$	366	$2.35474e-09$
Biggs EXP6 6-d	1827	$5.65565e-03$	1647	$5.65565e-03$	2548	$8.13521e-14$
Extended Rosenbrock 6-d	2470	$1.12155e-07$	2265	$5.80201e-09$	1850	$1.00959e-07$
Brown almost-linear 7-d	909	$7.85919e-09$	1104	$2.42837e-10$	754	$1.76765e-08$
Quadratic 8-d	570	$5.10728e-08$	653	$1.12288e-08$	748	$4.10269e-09$
Extended Rosenbrock 8-d	4888	$2.62388e-09$	4823	$1.98466e-09$	5936	$2.07958e-06$

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1762	$3.02171e-07$	1547	$2.16943e-06$	1766	$1.48740e-07$
Extended Powell 8-d	2187	$8.25264e-08$	2240	$4.94741e-08$	1779	$1.59526e-08$
Watson 9-d	2136	$2.27743e-04$	1906	$1.17812e-04$	2449	$8.83957e-05$
Extended Rosenbrock 10-d	7747	$3.35214e-05$	8513	$2.74187e-05$	8720	$5.29574e-08$
Penalty (1) 10-d	8469	$7.08776e-05$	4030	$7.17947e-05$	7895	$7.08766e-05$
Penalty (2) 10-d	647	$2.98631e-04$	11162	$2.94969e-04$	4911	$2.94481e-04$
Trigonometric 10-d	1813	$2.79554e-05$	1399	$2.79626e-05$	1418	$2.79532e-05$
Osbourne (2) 11-d	5157	$4.01382e-02$	3974	$4.01396e-02$	4048	$4.01379e-02$
Extended Powell 12-d	4877	$2.62397e-06$	3909	$6.67566e-06$	4696	$1.45532e-06$
Quadratic 16-d	1689	$1.18124e-08$	1416	$6.49922e-08$	1467	$1.06886e-07$
Quadratic 24-d	3450	$5.26190e-08$	3195	$8.53100e-08$	3078	$3.18972e-08$

Table G.18: Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_4–ν_6.

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	218	$5.38087e-11$	218	$9.85723e-10$	217	$1.14475e-10$
Freudenstein and Roth 2-d	151	$4.89843e+01$	148	$4.89843e+01$	147	$4.89843e+01$
Powell badly scaled 2-d	899	$3.42702e-17$	770	$4.68385e-11$	762	$9.50687e-17$
Brown badly scaled 2-d	386	$6.46531e-08$	393	$5.14292e-09$	371	$9.91099e-10$
Beale 2-d	117	$6.65985e-10$	121	$1.70860e-10$	117	$7.70208e-13$
Jennrich and Sampson 2-d	96	$1.24362e+02$	92	$1.24362e+02$	84	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	137	$9.80281e-05$	137	$9.80281e-05$	128	$9.75795e-05$
Bard 3-d	1367	$1.74287e+01$	1099	$1.74287e+01$	787	$1.74287e+01$
Gaussian 3-d	72	$1.33833e-08$	61	$1.29096e-08$	61	$1.29096e-08$
Meyer 3-d	2605	$8.79459e+01$	2632	$8.79459e+01$	2472	$8.79459e+01$
Gulf research 3-d	486	$4.39801e-14$	437	$7.57613e-14$	437	$7.57613e-14$
Box 3-d	410	$9.53899e-14$	355	$2.18028e-12$	327	$7.35720e-14$
Powell singular 4-d	525	$3.40790e-13$	358	$1.67017e-11$	299	$1.41375e-08$
Wood 4-d	507	$4.69030e-09$	389	$2.58277e-07$	390	$5.81618e-09$
Kowalik and Osbourne 4-d	334	$3.07506e-04$	484	$3.07506e-04$	470	$3.07506e-04$
Brown and Dennis 4-d	373	$8.58222e+04$	418	$8.58222e+04$	364	$8.58222e+04$
Quadratic 4-d	306	$1.85490e-09$	283	$6.45079e-09$	300	$4.89311e-09$
Penalty (1) 4-d	1349	$2.25010e-05$	1652	$2.24998e-05$	1130	$2.26143e-05$
Penalty (2) 4-d	190	$1.02675e-05$	187	$1.02820e-05$	4106	$9.39957e-06$
Osbourne (1) 5-d	238	$7.22008e-05$	215	$7.22121e-05$	215	$7.22121e-05$
Brown almost linear 5-d	380	$1.31909e-10$	364	$1.92843e-10$	362	$1.23970e-10$
Biggs EXP6 6-d	3242	$2.24703e-12$	3702	$3.31156e-12$	1488	$5.65565e-03$
Extended Rosenbrock 6-d	2428	$4.49726e-08$	2839	$2.95708e-08$	1874	$1.57214e-09$
Brown almost-linear 7-d	1083	$6.58927e-09$	623	$1.55045e-08$	855	$5.52252e-10$
Quadratic 8-d	644	$4.85093e-09$	630	$5.34118e-09$	588	$4.11947e-09$
Extended Rosenbrock 8-d	3073	$3.65974e-06$	4168	$3.92654e-09$	3375	$5.20740e-07$

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1832	$5.74454e-09$	2000	$1.53907e-07$	1483	$8.80142e-09$
Extended Powell 8-d	1849	$2.33901e-07$	2333	$2.24110e-10$	1949	$8.09086e-08$
Watson 9-d	3074	$4.60161e-05$	3591	$4.33185e-06$	2423	$6.44712e-05$
Extended Rosenbrock 10-d	7680	$2.01904e-08$	5703	$7.95335e-09$	6857	$4.93001e-08$
Penalty (1) 10-d	7372	$7.08798e-05$	10363	$7.08793e-05$	7173	$7.08910e-05$
Penalty (2) 10-d	582	$2.98846e-04$	4112	$2.95590e-04$	4538	$2.96687e-04$
Trigonometric 10-d	1555	$2.79669e-05$	1676	$2.79539e-05$	1676	$2.79539e-05$
Osbourne (2) 11-d	4925	$4.01377e-02$	5153	$4.01378e-02$	4596	$4.01378e-02$
Extended Powell 12-d	6745	$1.98860e-08$	4855	$3.58141e-06$	2713	$1.96182e-06$
Quadratic 16-d	1621	$5.16870e-08$	1590	$2.15888e-08$	1628	$4.88574e-08$
Quadratic 24-d	3137	$9.31400e-08$	2963	$1.13355e-07$	3573	$6.08289e-08$

Table G.19: Low tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_7–ν_9.

G.5.2 High tolerance results

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	410	$1.22973e-16$	351	$5.12253e-16$	296	$5.40517e-19$
Freudenstein and Roth 2-d	342	$4.89843e+01$	299	$4.89843e+01$	242	$4.89843e+01$
Powell badly scaled 2-d	1627	$1.69783e-21$	1346	$1.73098e-21$	1409	$6.86006e-23$
Brown badly scaled 2-d	646	$5.35973e-06$	679	$9.45087e-14$	699	$7.67771e-17$
Beale 2-d	319	$1.46430e-18$	200	$1.42899e-14$	210	$1.57824e-17$
Jennrich and Sampson 2-d	295	$1.24362e+02$	237	$1.24362e+02$	213	$1.24362e+02$
M^{c}Kinnon 2-d	319	$-2.50000e-01$	547	$-2.50000e-01$	369	$-2.50000e-01$
Helical valley 3-d	448	$1.13898e-11$	509	$2.33784e-14$	413	$1.49519e-17$
Bard 3-d	585	$1.74289e+01$	1415	$1.74287e+01$	1271	$1.74287e+01$
Gaussian 3-d	288	$1.12793e-08$	333	$1.12793e-08$	262	$1.12793e-08$
Meyer 3-d	3366	$8.79459e+01$	3594	$8.79459e+01$	3066	$8.79459e+01$
Gulf research 3-d	828	$1.14677e-18$	847	$2.65467e-16$	778	$1.12987e-20$
Box 3-d	676	$1.24773e-15$	764	$1.37905e-19$	589	$9.34788e-20$
Powell singular 4-d	1378	$2.18977e-15$	1535	$4.85828e-20$	1441	$1.30909e-18$
Wood 4-d	1217	$3.41173e-13$	1131	$2.15167e-15$	674	$1.67414e-13$
Kowalik and Osbourne 4-d	1005	$3.07506e-04$	1074	$3.07506e-04$	841	$3.07506e-04$
Brown and Dennis 4-d	965	$8.58222e+04$	1013	$8.58222e+04$	779	$8.58222e+04$
Quadratic 4-d	854	$6.93293e-17$	620	$7.45672e-17$	574	$6.37987e-17$
Penalty (1) 4-d	1981	$2.24998e-05$	2883	$2.24998e-05$	1491	$2.24998e-05$
Penalty (2) 4-d	6053	$9.37629e-06$	3870	$9.37629e-06$	5063	$9.37629e-06$
Osbourne (1) 5-d	1911	$5.46489e-05$	2166	$5.46489e-05$	1846	$5.46489e-05$
Brown almost linear 5-d	1193	$3.44755e-15$	1078	$2.11387e-16$	1456	$3.05477e-17$
Biggs EXP6 6-d	5965	$3.71641e-17$	26956	$4.46883e-03$	4248	$9.92518e-22$
Extended Rosenbrock 6-d	3555	$1.87813e-14$	3334	$6.29238e-15$	3009	$1.65343e-16$
Brown almost-linear 7-d	2456	$1.58620e-15$	2738	$3.23866e-16$	2167	$4.56790e-14$

	ν_1		ν_2		ν_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	2034	$5.27897e-17$	1463	$4.50475e-15$	1322	$1.31060e-16$
Extended Rosenbrock 8-d	8341	$5.72152e-16$	7148	$1.46920e-13$	6734	$2.93615e-14$
Variably dimensional 8-d	4934	$4.39972e-13$	4138	$3.01454e-15$	3729	$8.51133e-16$
Extended Powell 8-d	10902	$2.26254e-18$	9929	$6.36422e-20$	6505	$3.88690e-17$
Watson 9-d	11317	$1.39976e-06$	9767	$1.39976e-06$	9798	$1.39976e-06$
Extended Rosenbrock 10-d	13790	$3.60446e-14$	9856	$1.99316e-13$	12357	$3.47022e-12$
Penalty (1) 10-d	13684	$7.08765e-05$	14929	$7.08765e-05$	12510	$7.08765e-05$
Penalty (2) 10-d	30660	$2.93661e-04$	30747	$2.93661e-04$	35981	$2.93661e-04$
Trigonometric 10-d	4285	$2.79506e-05$	3038	$2.79506e-05$	2506	$2.79506e-05$
Osbourne (2) 11-d	12111	$4.01377e-02$	10669	$4.01377e-02$	11393	$4.01377e-02$
Extended Powell 12-d	21468	$5.90556e-12$	18952	$2.14531e-13$	19827	$6.40170e-13$
Quadratic 16-d	5496	$3.20851e-16$	5000	$5.91622e-16$	3789	$3.91233e-16$
Quadratic 24-d	11507	$3.51310e-16$	9599	$5.44595e-16$	8065	$3.76712e-16$

Table G.20: High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_1–ν_3.

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	326	$9.51821e-17$	318	$5.01367e-18$	309	$5.92068e-18$
Freudenstein and Roth 2-d	217	$4.89843e+01$	225	$4.89843e+01$	213	$4.89843e+01$
Powell badly scaled 2-d	1134	$2.31374e-24$	1068	$1.68758e-25$	1052	$9.57956e-26$
Brown badly scaled 2-d	549	$9.44820e-15$	461	$4.36638e-18$	479	$1.15698e-14$
Beale 2-d	202	$2.40154e-17$	200	$2.21645e-17$	190	$3.56955e-18$
Jennrich and Sampson 2-d	205	$1.24362e+02$	180	$1.24362e+02$	177	$1.24362e+02$
M^{c}Kinnon 2-d	355	$-2.50000e-01$	395	$-2.50000e-01$	371	$-2.50000e-01$
Helical valley 3-d	390	$4.73426e-14$	455	$2.69167e-17$	377	$4.94383e-16$
Bard 3-d	1120	$1.74287e+01$	1137	$1.74287e+01$	999	$1.74287e+01$
Gaussian 3-d	268	$1.12793e-08$	216	$1.12793e-08$	181	$1.12793e-08$
Meyer 3-d	2880	$8.79459e+01$	2901	$8.79459e+01$	2728	$8.79459e+01$
Gulf research 3-d	643	$5.55925e-21$	647	$1.77339e-20$	632	$1.97719e-22$
Box 3-d	513	$2.21254e-19$	523	$4.57585e-20$	483	$3.76045e-18$
Powell singular 4-d	1144	$5.59743e-21$	1177	$1.40815e-18$	874	$7.89917e-21$
Wood 4-d	758	$6.50981e-16$	701	$1.39074e-14$	663	$6.27213e-15$
Kowalik and Osbourne 4-d	694	$3.07506e-04$	485	$3.07506e-04$	604	$3.07506e-04$
Brown and Dennis 4-d	612	$8.58222e+04$	650	$8.58222e+04$	560	$8.58222e+04$
Quadratic 4-d	472	$4.42602e-17$	450	$1.11744e-16$	442	$4.24161e-17$
Penalty (1) 4-d	2072	$2.24998e-05$	1993	$2.24998e-05$	1744	$2.24998e-05$
Penalty (2) 4-d	4717	$9.37629e-06$	4984	$9.37629e-06$	5066	$9.37629e-06$
Osbourne (1) 5-d	1956	$5.46489e-05$	1538	$5.46489e-05$	1625	$5.46489e-05$
Brown almost linear 5-d	937	$3.39796e-17$	877	$2.81098e-17$	818	$6.19501e-18$
Biggs EXP6 6-d	2079	$5.65565e-03$	1980	$5.65565e-03$	3020	$2.43327e-20$
Extended Rosenbrock 6-d	3452	$6.18763e-17$	2895	$3.05543e-17$	2359	$1.13252e-14$
Brown almost-linear 7-d	2172	$1.88311e-16$	1714	$2.85389e-17$	2237	$2.20196e-16$
Quadratic 8-d	1125	$3.14603e-17$	1066	$7.06420e-16$	1153	$4.70076e-17$
Extended Rosenbrock 8-d	6095	$4.25559e-13$	6459	$9.05932e-16$	7222	$3.08444e-14$

	ν_4		ν_5		ν_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	3013	$6.98138e-14$	3007	$1.61427e-14$	2743	$1.28223e-16$
Extended Powell 8-d	4917	$1.34821e-14$	4850	$7.82025e-15$	4340	$1.82333e-16$
Watson 9-d	7898	$1.39976e-06$	24324	$1.39976e-06$	7631	$1.39976e-06$
Extended Rosenbrock 10-d	11805	$4.43692e-14$	12693	$2.29228e-12$	12321	$1.69608e-14$
Penalty (1) 10-d	15010	$7.08765e-05$	13993	$7.08765e-05$	13562	$7.08765e-05$
Penalty (2) 10-d	26553	$2.93661e-04$	51824	$2.93661e-04$	51301	$2.93661e-04$
Trigonometric 10-d	2967	$2.79506e-05$	2344	$2.79506e-05$	2629	$2.79506e-05$
Osbourne (2) 11-d	10902	$4.01377e-02$	9830	$4.01377e-02$	7386	$4.01377e-02$
Extended Powell 12-d	24607	$9.86010e-17$	19344	$1.18995e-13$	21589	$6.50121e-18$
Quadratic 16-d	3024	$3.61229e-16$	2775	$7.78383e-16$	2665	$2.05308e-16$
Quadratic 24-d	5915	$1.04695e-15$	5366	$1.32264e-16$	5188	$1.97109e-15$

Table G.21: High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_4–ν_6.

Function	ν_7 FE	ν_7 Minimum	ν_8 FE	ν_8 Minimum	ν_9 FE	ν_9 Minimum
Rosenbrock 2-d	288	$1.47671e-17$	285	$1.39058e-17$	284	$5.00395e-18$
Freudenstein and Roth 2-d	229	$4.89843e+01$	217	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	1025	$1.20181e-25$	969	$4.23980e-25$	881	$2.83330e-25$
Brown badly scaled 2-d	497	$1.30223e-17$	498	$7.99797e-17$	470	$1.15666e-17$
Beale 2-d	185	$1.75420e-19$	191	$2.07825e-18$	191	$5.58508e-18$
Jennrich and Sampson 2-d	169	$1.24362e+02$	157	$1.24362e+02$	170	$1.24362e+02$
M^c^Kinnon 2-d	426	$-2.50000e-01$	426	$-2.50000e-01$	397	$-2.50000e-01$
Helical valley 3-d	373	$1.72765e-16$	342	$9.83210e-16$	325	$5.02187e-17$
Bard 3-d	1405	$1.74287e+01$	1134	$1.74287e+01$	863	$1.74287e+01$
Gaussian 3-d	211	$1.12793e-08$	194	$1.12793e-08$	189	$1.12793e-08$
Meyer 3-d	2796	$8.79459e+01$	2801	$8.79459e+01$	2635	$8.79459e+01$
Gulf research 3-d	658	$2.44925e-22$	529	$5.44511e-19$	550	$3.52255e-21$
Box 3-d	495	$1.03381e-19$	478	$8.70459e-21$	440	$2.12341e-21$
Powell singular 4-d	864	$1.28278e-18$	1045	$6.73509e-26$	1484	$9.23531e-28$
Wood 4-d	659	$9.68516e-17$	656	$2.57400e-16$	619	$4.41227e-17$
Kowalik and Osbourne 4-d	551	$3.07506e-04$	653	$3.07506e-04$	713	$3.07506e-04$
Brown and Dennis 4-d	638	$8.58222e+04$	603	$8.58222e+04$	617	$8.58222e+04$
Quadratic 4-d	458	$2.59084e-17$	440	$2.15350e-17$	456	$4.82633e-17$
Penalty (1) 4-d	1614	$2.24998e-05$	1848	$2.24998e-05$	1796	$2.24998e-05$
Penalty (2) 4-d	5427	$9.37629e-06$	4688	$9.37629e-06$	5597	$9.37629e-06$
Osbourne (1) 5-d	1580	$5.46489e-05$	1488	$5.46489e-05$	1466	$5.46489e-05$
Brown almost linear 5-d	711	$1.00627e-17$	648	$1.08728e-18$	581	$1.58304e-17$
Biggs EXP6 6-d	3704	$5.76638e-19$	4390	$1.16131e-20$	1905	$5.65565e-03$
Extended Rosenbrock 6-d	3243	$2.28125e-17$	3110	$1.35844e-14$	2334	$1.20515e-17$
Brown almost-linear 7-d	1798	$3.95949e-18$	1539	$1.51163e-17$	1266	$2.80733e-17$
Quadratic 8-d	1147	$3.22611e-17$	1002	$8.07477e-17$	1012	$1.71573e-16$
Extended Rosenbrock 8-d	4294	$2.36680e-12$	5314	$3.27909e-17$	4296	$7.08914e-14$

	ν_7		ν_8		ν_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	2387	$3.35756e-16$	2829	$6.09434e-15$	2042	$4.53560e-15$
Extended Powell 8-d	6809	$6.60298e-25$	7199	$6.43822e-24$	5384	$8.17073e-21$
Watson 9-d	7669	$1.39976e-06$	6986	$1.39976e-06$	6588	$1.39976e-06$
Extended Rosenbrock 10-d	9603	$7.25704e-14$	7629	$2.22125e-16$	8364	$3.89649e-14$
Penalty (1) 10-d	12267	$7.08765e-05$	13106	$7.08765e-05$	12975	$7.08765e-05$
Penalty (2) 10-d	54949	$2.93661e-04$	45181	$2.93661e-04$	61892	$2.93661e-04$
Trigonometric 10-d	2977	$2.79506e-05$	2466	$2.79506e-05$	2691	$2.79506e-05$
Osbourne (2) 11-d	8829	$4.01377e-02$	6416	$4.01377e-02$	8971	$4.01377e-02$
Extended Powell 12-d	25792	$1.23516e-20$	16827	$4.15827e-19$	17907	$1.43117e-11$
Quadratic 16-d	2977	$6.23699e-16$	2743	$5.16000e-16$	2622	$1.93155e-16$
Quadratic 24-d	5025	$3.42648e-16$	4632	$7.66866e-16$	5419	$3.08287e-16$

Table G.22: High tolerance results for NM4_$\psi_3\delta_5\kappa_1$ ν_7–ν_9.

G.6 Determinant parameter (δ) using ψ_3, κ_1, ν_8

G.6.1 Low tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	218	$9.85723e-10$	218	$9.85723e-10$	218	$9.85723e-10$
Freudenstein and Roth 2-d	148	$4.89843e+01$	148	$4.89843e+01$	148	$4.89843e+01$
Powell badly scaled 2-d	841	$2.44665e-17$	843	$2.44665e-17$	770	$4.68385e-11$
Brown badly scaled 2-d	392	$5.14292e-09$	393	$5.14292e-09$	393	$5.14292e-09$
Beale 2-d	121	$1.70860e-10$	121	$1.70860e-10$	121	$1.70860e-10$
Jennrich and Sampson 2-d	92	$1.24362e+02$	92	$1.24362e+02$	92	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	137	$9.80281e-05$	137	$9.80281e-05$	137	$9.80281e-05$
Bard 3-d	1097	$1.74287e+01$	1097	$1.74287e+01$	1097	$1.74287e+01$
Gaussian 3-d	61	$1.29096e-08$	61	$1.29096e-08$	61	$1.29096e-08$
Meyer 3-d	2628	$8.79459e+01$	2631	$8.79459e+01$	2632	$8.79459e+01$
Gulf research 3-d	437	$7.57613e-14$	437	$7.57613e-14$	437	$7.57613e-14$
Box 3-d	387	$7.85709e-12$	355	$2.18028e-12$	355	$2.18028e-12$
Powell singular 4-d	519	$5.48362e-12$	358	$1.67017e-11$	358	$1.67017e-11$
Wood 4-d	389	$2.58277e-07$	389	$2.58277e-07$	389	$2.58277e-07$
Kowalik and Osbourne 4-d	484	$3.07506e-04$	484	$3.07506e-04$	484	$3.07506e-04$
Brown and Dennis 4-d	418	$8.58222e+04$	418	$8.58222e+04$	418	$8.58222e+04$
Quadratic 4-d	283	$6.45079e-09$	283	$6.45079e-09$	283	$6.45079e-09$
Penalty (1) 4-d	1651	$2.24998e-05$	1652	$2.24998e-05$	1652	$2.24998e-05$
Penalty (2) 4-d	187	$1.02820e-05$	187	$1.02820e-05$	187	$1.02820e-05$
Osbourne (1) 5-d	215	$7.22121e-05$	215	$7.22121e-05$	215	$7.22121e-05$
Brown almost linear 5-d	426	$1.81523e-10$	364	$1.92843e-10$	364	$1.92843e-10$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	2453	$2.19939e-06$	3702	$3.31156e-12$	3702	$3.31156e-12$
Extended Rosenbrock 6-d	2837	$2.95708e-08$	2839	$2.95708e-08$	2839	$2.95708e-08$
Brown almost-linear 7-d	941	$3.52679e-08$	622	$1.55045e-08$	623	$1.55045e-08$
Quadratic 8-d	630	$6.66647e-09$	1057	$6.33124e-09$	629	$5.34118e-09$
Extended Rosenbrock 8-d	3835	$6.75330e-09$	4393	$1.66468e-05$	4166	$3.92654e-09$
Variably dimensional 8-d	1442	$6.38317e-07$	1486	$1.90415e-07$	1427	$3.55750e-08$
Extended Powell 8-d	2116	$2.05329e-10$	1331	$7.34791e-08$	2250	$3.95142e-07$
Watson 9-d	4016	$8.79992e-06$	3103	$9.68586e-06$	3060	$9.32248e-06$
Extended Rosenbrock 10-d	12384	$2.13003e-07$	5693	$7.95335e-09$	5698	$7.95335e-09$
Penalty (1) 10-d	8146	$7.08786e-05$	8022	$7.08814e-05$	7412	$7.08790e-05$
Penalty (2) 10-d	1876	$2.97726e-04$	1684	$2.97928e-04$	1837	$2.97401e-04$
Trigonometric 10-d	1675	$2.79539e-05$	1676	$2.79539e-05$	1676	$2.79539e-05$
Osbourne (2) 11-d	4685	$4.01378e-02$	5023	$4.01379e-02$	5822	$4.01377e-02$
Extended Powell 12-d	9409	$7.54988e-07$	5698	$5.35652e-07$	6182	$8.67443e-08$
Quadratic 16-d	1758	$5.58385e-08$	1585	$1.75059e-08$	1743	$2.13713e-08$
Quadratic 24-d	3005	$1.44362e-08$	3747	$1.75664e-08$	3747	$1.75664e-08$

Table G.23: Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	218	$9.85723e-10$	218	$9.85723e-10$	218	$9.85723e-10$
Freudenstein and Roth 2-d	148	$4.89843e+01$	148	$4.89843e+01$	148	$4.89843e+01$
Powell badly scaled 2-d	770	$4.68385e-11$	770	$4.68385e-11$	770	$4.68385e-11$
Brown badly scaled 2-d	393	$5.14292e-09$	393	$5.14292e-09$	393	$5.14292e-09$
Beale 2-d	121	$1.70860e-10$	121	$1.70860e-10$	121	$1.70860e-10$
Jennrich and Sampson 2-d	92	$1.24362e+02$	92	$1.24362e+02$	92	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	137	$9.80281e-05$	137	$9.80281e-05$	137	$9.80281e-05$
Bard 3-d	1098	$1.74287e+01$	1099	$1.74287e+01$	1099	$1.74287e+01$
Gaussian 3-d	61	$1.29096e-08$	61	$1.29096e-08$	61	$1.29096e-08$
Meyer 3-d	2632	$8.79459e+01$	2632	$8.79459e+01$	2632	$8.79459e+01$
Gulf research 3-d	437	$7.57613e-14$	437	$7.57613e-14$	437	$7.57613e-14$
Box 3-d	355	$2.18028e-12$	355	$2.18028e-12$	355	$2.18028e-12$
Powell singular 4-d	358	$1.67017e-11$	358	$1.67017e-11$	358	$1.67017e-11$
Wood 4-d	389	$2.58277e-07$	389	$2.58277e-07$	389	$2.58277e-07$
Kowalik and Osbourne 4-d	484	$3.07506e-04$	484	$3.07506e-04$	484	$3.07506e-04$
Brown and Dennis 4-d	418	$8.58222e+04$	418	$8.58222e+04$	418	$8.58222e+04$
Quadratic 4-d	283	$6.45079e-09$	283	$6.45079e-09$	283	$6.45079e-09$
Penalty (1) 4-d	1652	$2.24998e-05$	1652	$2.24998e-05$	1652	$2.24998e-05$
Penalty (2) 4-d	187	$1.02820e-05$	187	$1.02820e-05$	187	$1.02820e-05$
Osbourne (1) 5-d	215	$7.22121e-05$	215	$7.22121e-05$	215	$7.22121e-05$
Brown almost linear 5-d	364	$1.92843e-10$	364	$1.92843e-10$	364	$1.92843e-10$
Biggs EXP6 6-d	3702	$3.31156e-12$	3702	$3.31156e-12$	3702	$3.31156e-12$
Extended Rosenbrock 6-d	2839	$2.95708e-08$	2839	$2.95708e-08$	2839	$2.95708e-08$
Brown almost-linear 7-d	623	$1.55045e-08$	623	$1.55045e-08$	623	$1.55045e-08$
Quadratic 8-d	630	$5.34118e-09$	630	$5.34118e-09$	630	$5.34118e-09$
Extended Rosenbrock 8-d	4168	$3.92654e-09$	4168	$3.92654e-09$	4168	$3.92654e-09$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1812	$3.51367e-09$	2000	$1.53907e-07$	1365	$4.12240e-08$
Extended Powell 8-d	2917	$4.34701e-13$	2333	$2.24110e-10$	2333	$2.24110e-10$
Watson 9-d	3061	$9.32248e-06$	3591	$4.33185e-06$	3233	$9.24202e-06$
Extended Rosenbrock 10-d	5701	$7.95335e-09$	5703	$7.95335e-09$	5703	$7.95335e-09$
Penalty (1) 10-d	6361	$7.08917e-05$	10363	$7.08793e-05$	6754	$7.08775e-05$
Penalty (2) 10-d	2324	$2.97269e-04$	4112	$2.95590e-04$	1578	$2.97345e-04$
Trigonometric 10-d	1676	$2.79539e-05$	1676	$2.79539e-05$	1676	$2.79539e-05$
Osbourne (2) 11-d	5152	$4.01378e-02$	5153	$4.01378e-02$	5153	$4.01378e-02$
Extended Powell 12-d	4913	$7.29451e-07$	4855	$3.58141e-06$	5016	$1.35382e-06$
Quadratic 16-d	1743	$2.13713e-08$	1590	$2.15888e-08$	1248	$2.26014e-08$
Quadratic 24-d	3747	$1.75664e-08$	2963	$1.13355e-07$	3094	$2.43870e-08$

Table G.24: Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_4–δ_6.

Function	δ_7 FE	δ_7 Minimum	δ_8 FE	δ_8 Minimum	δ_9 FE	δ_9 Minimum
Rosenbrock 2-d	218	$9.85723e-10$	218	$9.85723e-10$	218	$9.85723e-10$
Freudenstein and Roth 2-d	148	$4.89843e+01$	148	$4.89843e+01$	148	$4.89843e+01$
Powell badly scaled 2-d	770	$4.68385e-11$	770	$4.68385e-11$	770	$4.68385e-11$
Brown badly scaled 2-d	393	$5.14292e-09$	393	$5.14292e-09$	393	$5.14292e-09$
Beale 2-d	121	$1.70860e-10$	121	$1.70860e-10$	121	$1.70860e-10$
Jennrich and Sampson 2-d	92	$1.24362e+02$	92	$1.24362e+02$	92	$1.24362e+02$
M^{c}Kinnon 2-d	23	$-6.24961e-05$	23	$-6.24961e-05$	23	$-6.24961e-05$
Helical valley 3-d	137	$9.80281e-05$	137	$9.80281e-05$	137	$9.80281e-05$
Bard 3-d	1099	$1.74287e+01$	1100	$1.74287e+01$	1101	$1.74287e+01$
Gaussian 3-d	61	$1.29096e-08$	61	$1.29096e-08$	61	$1.29096e-08$
Meyer 3-d	2632	$8.79459e+01$	2632	$8.79459e+01$	2632	$8.79459e+01$
Gulf research 3-d	437	$7.57613e-14$	437	$7.57613e-14$	437	$7.57613e-14$
Box 3-d	355	$2.18028e-12$	355	$2.18028e-12$	355	$2.18028e-12$
Powell singular 4-d	358	$1.67017e-11$	358	$1.67017e-11$	358	$1.67017e-11$
Wood 4-d	389	$2.58277e-07$	389	$2.58277e-07$	389	$2.58277e-07$
Kowalik and Osbourne 4-d	484	$3.07506e-04$	484	$3.07506e-04$	484	$3.07506e-04$
Brown and Dennis 4-d	418	$8.58222e+04$	418	$8.58222e+04$	418	$8.58222e+04$
Quadratic 4-d	283	$6.45079e-09$	283	$6.45079e-09$	283	$6.45079e-09$
Penalty (1) 4-d	1652	$2.24998e-05$	1652	$2.24998e-05$	1652	$2.24998e-05$
Penalty (2) 4-d	187	$1.02820e-05$	187	$1.02820e-05$	187	$1.02820e-05$
Osbourne (1) 5-d	215	$7.22121e-05$	215	$7.22121e-05$	215	$7.22121e-05$
Brown almost linear 5-d	364	$1.92843e-10$	364	$1.92843e-10$	364	$1.92843e-10$
Biggs EXP6 6-d	3702	$3.31156e-12$	3702	$3.31156e-12$	3702	$3.31156e-12$
Extended Rosenbrock 6-d	2839	$2.95708e-08$	2839	$2.95708e-08$	2839	$2.95708e-08$
Brown almost-linear 7-d	623	$1.55045e-08$	623	$1.55045e-08$	623	$1.55045e-08$
Quadratic 8-d	630	$5.34118e-09$	630	$5.34118e-09$	630	$5.34118e-09$
Extended Rosenbrock 8-d	4168	$3.92654e-09$	4168	$3.92654e-09$	4168	$3.92654e-09$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1365	$4.12240e-08$	1365	$4.12240e-08$	1365	$4.12240e-08$
Extended Powell 8-d	2333	$2.24110e-10$	2333	$2.24110e-10$	2333	$2.24110e-10$
Watson 9-d	3372	$6.04106e-06$	3373	$6.04106e-06$	3373	$6.04106e-06$
Extended Rosenbrock 10-d	5703	$7.95335e-09$	5703	$7.95335e-09$	5703	$7.95335e-09$
Penalty (1) 10-d	6759	$7.08775e-05$	6761	$7.08775e-05$	6761	$7.08775e-05$
Penalty (2) 10-d	1580	$2.97345e-04$	1580	$2.97345e-04$	1581	$2.97345e-04$
Trigonometric 10-d	1676	$2.79539e-05$	1676	$2.79539e-05$	1676	$2.79539e-05$
Osbourne (2) 11-d	5153	$4.01378e-02$	5153	$4.01378e-02$	5153	$4.01378e-02$
Extended Powell 12-d	4680	$1.68978e-07$	4681	$1.68978e-07$	4681	$1.68978e-07$
Quadratic 16-d	1701	$1.18070e-07$	1702	$1.18070e-07$	1702	$1.18070e-07$
Quadratic 24-d	3000	$2.66616e-08$	3130	$7.20153e-08$	2485	$4.33544e-08$

Table G.25: Low tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_7–δ_9.

G.6.2 High tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	285	$1.39058e-17$	285	$1.39058e-17$	285	$1.39058e-17$
Freudenstein and Roth 2-d	217	$4.89843e+01$	217	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	955	$2.49715e-19$	958	$2.49715e-19$	969	$4.23980e-25$
Brown badly scaled 2-d	496	$7.99797e-17$	498	$7.99797e-17$	498	$7.99797e-17$
Beale 2-d	191	$2.07825e-18$	191	$2.07825e-18$	191	$2.07825e-18$
Jennrich and Sampson 2-d	157	$1.24362e+02$	157	$1.24362e+02$	157	$1.24362e+02$
M^{c}Kinnon 2-d	418	$-2.50000e-01$	425	$-2.50000e-01$	426	$-2.50000e-01$
Helical valley 3-d	342	$9.83210e-16$	342	$9.83210e-16$	342	$9.83210e-16$
Bard 3-d	1132	$1.74287e+01$	1132	$1.74287e+01$	1132	$1.74287e+01$
Gaussian 3-d	194	$1.12793e-08$	194	$1.12793e-08$	194	$1.12793e-08$
Meyer 3-d	2795	$8.79459e+01$	2798	$8.79459e+01$	2801	$8.79459e+01$
Gulf research 3-d	529	$5.44511e-19$	529	$5.44511e-19$	529	$5.44511e-19$
Box 3-d	567	$5.70608e-22$	478	$8.70459e-21$	478	$8.70459e-21$
Powell singular 4-d	697	$3.34420e-14$	1043	$6.73509e-26$	1043	$6.73509e-26$
Wood 4-d	656	$2.57400e-16$	656	$2.57400e-16$	656	$2.57400e-16$
Kowalik and Osbourne 4-d	653	$3.07506e-04$	653	$3.07506e-04$	653	$3.07506e-04$
Brown and Dennis 4-d	603	$8.58222e+04$	603	$8.58222e+04$	603	$8.58222e+04$
Quadratic 4-d	440	$2.15350e-17$	440	$2.15350e-17$	440	$2.15350e-17$
Penalty (1) 4-d	1847	$2.24998e-05$	1848	$2.24998e-05$	1848	$2.24998e-05$
Penalty (2) 4-d	4687	$9.37629e-06$	4688	$9.37629e-06$	4688	$9.37629e-06$
Osbourne (1) 5-d	1486	$5.46489e-05$	1486	$5.46489e-05$	1486	$5.46489e-05$
Brown almost linear 5-d	797	$1.97092e-18$	648	$1.08728e-18$	648	$1.08728e-18$
Biggs EXP6 6-d	3748	$8.21578e-20$	4389	$1.16131e-20$	4390	$1.16131e-20$
Extended Rosenbrock 6-d	3108	$1.35844e-14$	3110	$1.35844e-14$	3110	$1.35844e-14$
Brown almost-linear 7-d	1828	$1.61693e-18$	1538	$1.51163e-17$	1539	$1.51163e-17$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	1073	$2.87237e-16$	1671	$3.06204e-17$	1001	$8.07477e-17$
Extended Rosenbrock 8-d	4757	$8.58244e-17$	6030	$2.74071e-17$	5312	$3.27909e-17$
Variably dimensional 8-d	3229	$1.55274e-14$	2432	$6.01131e-17$	2314	$9.21883e-16$
Extended Powell 8-d	5884	$2.62471e-23$	5800	$4.50184e-22$	8502	$1.52805e-24$
Watson 9-d	7141	$1.39976e-06$	7467	$1.39976e-06$	5810	$1.39976e-06$
Extended Rosenbrock 10-d	16493	$6.80618e-15$	7615	$2.22125e-16$	7623	$2.22125e-16$
Penalty (1) 10-d	14051	$7.08765e-05$	14654	$7.08765e-05$	11707	$7.08765e-05$
Penalty (2) 10-d	49480	$2.93661e-04$	36341	$2.93661e-04$	40374	$2.93661e-04$
Trigonometric 10-d	2465	$2.79506e-05$	2466	$2.79506e-05$	2466	$2.79506e-05$
Osbourne (2) 11-d	6684	$4.01377e-02$	8220	$4.01377e-02$	7131	$4.01377e-02$
Extended Powell 12-d	22528	$7.75024e-15$	28280	$2.16766e-11$	25501	$3.36752e-18$
Quadratic 16-d	2912	$1.83645e-16$	2471	$7.75129e-16$	2663	$2.34974e-15$
Quadratic 24-d	5176	$4.60463e-16$	5451	$9.06847e-16$	5451	$9.06847e-16$

Table G.26: High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	285	$1.39058e-17$	285	$1.39058e-17$	285	$1.39058e-17$
Freudenstein and Roth 2-d	217	$4.89843e+01$	217	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	969	$4.23980e-25$	969	$4.23980e-25$	969	$4.23980e-25$
Brown badly scaled 2-d	498	$7.99797e-17$	498	$7.99797e-17$	498	$7.99797e-17$
Beale 2-d	191	$2.07825e-18$	191	$2.07825e-18$	191	$2.07825e-18$
Jennrich and Sampson 2-d	157	$1.24362e+02$	157	$1.24362e+02$	157	$1.24362e+02$
M^{c}Kinnon 2-d	426	$-2.50000e-01$	426	$-2.50000e-01$	426	$-2.50000e-01$
Helical valley 3-d	342	$9.83210e-16$	342	$9.83210e-16$	342	$9.83210e-16$
Bard 3-d	1133	$1.74287e+01$	1134	$1.74287e+01$	1134	$1.74287e+01$
Gaussian 3-d	194	$1.12793e-08$	194	$1.12793e-08$	194	$1.12793e-08$
Meyer 3-d	2801	$8.79459e+01$	2801	$8.79459e+01$	2801	$8.79459e+01$
Gulf research 3-d	529	$5.44511e-19$	529	$5.44511e-19$	529	$5.44511e-19$
Box 3-d	478	$8.70459e-21$	478	$8.70459e-21$	478	$8.70459e-21$
Powell singular 4-d	1044	$6.73509e-26$	1045	$6.73509e-26$	1045	$6.73509e-26$
Wood 4-d	656	$2.57400e-16$	656	$2.57400e-16$	656	$2.57400e-16$
Kowalik and Osbourne 4-d	653	$3.07506e-04$	653	$3.07506e-04$	653	$3.07506e-04$
Brown and Dennis 4-d	603	$8.58222e+04$	603	$8.58222e+04$	603	$8.58222e+04$
Quadratic 4-d	440	$2.15350e-17$	440	$2.15350e-17$	440	$2.15350e-17$
Penalty (1) 4-d	1848	$2.24998e-05$	1848	$2.24998e-05$	1848	$2.24998e-05$
Penalty (2) 4-d	4688	$9.37629e-06$	4688	$9.37629e-06$	4689	$9.37629e-06$
Osbourne (1) 5-d	1487	$5.46489e-05$	1488	$5.46489e-05$	1488	$5.46489e-05$
Brown almost linear 5-d	648	$1.08728e-18$	648	$1.08728e-18$	648	$1.08728e-18$
Biggs EXP6 6-d	4390	$1.16131e-20$	4390	$1.16131e-20$	4390	$1.16131e-20$
Extended Rosenbrock 6-d	3110	$1.35844e-14$	3110	$1.35844e-14$	3110	$1.35844e-14$
Brown almost-linear 7-d	1539	$1.51163e-17$	1539	$1.51163e-17$	1539	$1.51163e-17$
Quadratic 8-d	1002	$8.07477e-17$	1002	$8.07477e-17$	1002	$8.07477e-17$
Extended Rosenbrock 8-d	5314	$3.27909e-17$	5314	$3.27909e-17$	5314	$3.27909e-17$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	2531	$1.78090e-16$	2829	$6.09434e-15$	2563	$1.24784e-15$
Extended Powell 8-d	3887	$3.23172e-18$	7199	$6.43822e-24$	7200	$6.43822e-24$
Watson 9-d	5811	$1.39976e-06$	6986	$1.39976e-06$	5256	$1.39976e-06$
Extended Rosenbrock 10-d	7627	$2.22125e-16$	7629	$2.22125e-16$	7629	$2.22125e-16$
Penalty (1) 10-d	10367	$7.08765e-05$	13106	$7.08765e-05$	9200	$7.08765e-05$
Penalty (2) 10-d	51001	$2.93661e-04$	45181	$2.93661e-04$	32768	$2.93661e-04$
Trigonometric 10-d	2466	$2.79506e-05$	2466	$2.79506e-05$	2466	$2.79506e-05$
Osbourne (2) 11-d	6415	$4.01377e-02$	6416	$4.01377e-02$	6416	$4.01377e-02$
Extended Powell 12-d	27641	$1.29835e-11$	16827	$4.15827e-19$	20076	$1.11105e-20$
Quadratic 16-d	2663	$2.34974e-15$	2743	$5.16000e-16$	2352	$1.41547e-16$
Quadratic 24-d	5451	$9.06847e-16$	4632	$7.66866e-16$	4766	$1.21730e-15$

Table G.27: High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_4–δ_6.

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	285	$1.39058e-17$	285	$1.39058e-17$	285	$1.39058e-17$
Freudenstein and Roth 2-d	217	$4.89843e+01$	217	$4.89843e+01$	217	$4.89843e+01$
Powell badly scaled 2-d	969	$4.23980e-25$	969	$4.23980e-25$	969	$4.23980e-25$
Brown badly scaled 2-d	498	$7.99797e-17$	498	$7.99797e-17$	498	$7.99797e-17$
Beale 2-d	191	$2.07825e-18$	191	$2.07825e-18$	191	$2.07825e-18$
Jennrich and Sampson 2-d	157	$1.24362e+02$	157	$1.24362e+02$	157	$1.24362e+02$
M^{c}Kinnon 2-d	426	$-2.50000e-01$	426	$-2.50000e-01$	426	$-2.50000e-01$
Helical valley 3-d	342	$9.83210e-16$	342	$9.83210e-16$	342	$9.83210e-16$
Bard 3-d	1134	$1.74287e+01$	1135	$1.74287e+01$	1136	$1.74287e+01$
Gaussian 3-d	194	$1.12793e-08$	194	$1.12793e-08$	194	$1.12793e-08$
Meyer 3-d	2801	$8.79459e+01$	2801	$8.79459e+01$	2801	$8.79459e+01$
Gulf research 3-d	529	$5.44511e-19$	529	$5.44511e-19$	529	$5.44511e-19$
Box 3-d	478	$8.70459e-21$	478	$8.70459e-21$	478	$8.70459e-21$
Powell singular 4-d	1045	$6.73509e-26$	1045	$6.73509e-26$	1045	$6.73509e-26$
Wood 4-d	656	$2.57400e-16$	656	$2.57400e-16$	656	$2.57400e-16$
Kowalik and Osbourne 4-d	653	$3.07506e-04$	653	$3.07506e-04$	653	$3.07506e-04$
Brown and Dennis 4-d	603	$8.58222e+04$	603	$8.58222e+04$	603	$8.58222e+04$
Quadratic 4-d	440	$2.15350e-17$	440	$2.15350e-17$	440	$2.15350e-17$
Penalty (1) 4-d	1848	$2.24998e-05$	1848	$2.24998e-05$	1848	$2.24998e-05$
Penalty (2) 4-d	4689	$9.37629e-06$	4689	$9.37629e-06$	4689	$9.37629e-06$
Osbourne (1) 5-d	1488	$5.46489e-05$	1488	$5.46489e-05$	1488	$5.46489e-05$
Brown almost linear 5-d	648	$1.08728e-18$	648	$1.08728e-18$	648	$1.08728e-18$
Biggs EXP6 6-d	4390	$1.16131e-20$	4390	$1.16131e-20$	4390	$1.16131e-20$
Extended Rosenbrock 6-d	3110	$1.35844e-14$	3110	$1.35844e-14$	3110	$1.35844e-14$
Brown almost-linear 7-d	1539	$1.51163e-17$	1539	$1.51163e-17$	1539	$1.51163e-17$
Quadratic 8-d	1002	$8.07477e-17$	1002	$8.07477e-17$	1002	$8.07477e-17$
Extended Rosenbrock 8-d	5314	$3.27909e-17$	5314	$3.27909e-17$	5314	$3.27909e-17$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	2563	$1.24784e-15$	2563	$1.24784e-15$	2563	$1.24784e-15$
Extended Powell 8-d	7200	$6.43822e-24$	7202	$6.43822e-24$	7202	$6.43822e-24$
Watson 9-d	7365	$1.39976e-06$	7368	$1.39976e-06$	7368	$1.39976e-06$
Extended Rosenbrock 10-d	7629	$2.22125e-16$	7629	$2.22125e-16$	7629	$2.22125e-16$
Penalty (1) 10-d	9205	$7.08765e-05$	9207	$7.08765e-05$	9207	$7.08765e-05$
Penalty (2) 10-d	35246	$2.93661e-04$	35643	$2.93661e-04$	61476	$2.93661e-04$
Trigonometric 10-d	2466	$2.79506e-05$	2466	$2.79506e-05$	2466	$2.79506e-05$
Osbourne (2) 11-d	6416	$4.01377e-02$	6416	$4.01377e-02$	6416	$4.01377e-02$
Extended Powell 12-d	21389	$2.69071e-15$	21394	$2.69071e-15$	21395	$2.69071e-15$
Quadratic 16-d	2707	$1.24348e-15$	2708	$1.24348e-15$	2708	$1.24348e-15$
Quadratic 24-d	4747	$1.34834e-16$	4868	$5.93331e-16$	4298	$7.17485e-16$

Table G.28: High tolerance results for NM4_$\psi_3\kappa_1\nu_8$ δ_7–δ_9.

G.7 Determinant parameter (δ) using ψ_3, κ_2, ν_8

G.7.1 Low tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	241	$1.13562e-09$	241	$1.13562e-09$	241	$1.13562e-09$
Freudenstein and Roth 2-d	160	$4.89843e+01$	160	$4.89843e+01$	160	$4.89843e+01$
Powell badly scaled 2-d	885	$2.65600e-15$	888	$2.65600e-15$	889	$2.65600e-15$
Brown badly scaled 2-d	464	$3.67093e-09$	468	$3.67093e-09$	468	$3.67093e-09$
Beale 2-d	128	$1.86459e-09$	128	$1.86459e-09$	128	$1.86459e-09$
Jennrich and Sampson 2-d	100	$1.24362e+02$	100	$1.24362e+02$	100	$1.24362e+02$
M^{c}Kinnon 2-d	195	$-2.50000e-01$	195	$-2.50000e-01$	195	$-2.50000e-01$
Helical valley 3-d	165	$1.11853e-04$	165	$1.11853e-04$	165	$1.11853e-04$
Bard 3-d	1442	$1.74287e+01$	1442	$1.74287e+01$	1690	$1.74287e+01$
Gaussian 3-d	77	$1.59743e-08$	77	$1.59743e-08$	77	$1.59743e-08$
Meyer 3-d	2097	$8.79459e+01$	2101	$8.79459e+01$	2102	$8.79459e+01$
Gulf research 3-d	956	$1.12798e-14$	957	$1.12798e-14$	957	$1.12798e-14$
Box 3-d	371	$1.49473e-13$	317	$1.77655e-12$	317	$1.77655e-12$
Powell singular 4-d	433	$2.72831e-08$	580	$8.81610e-12$	580	$8.81610e-12$
Wood 4-d	570	$1.38215e-08$	570	$1.38215e-08$	570	$1.38215e-08$
Kowalik and Osbourne 4-d	422	$3.07506e-04$	422	$3.07506e-04$	422	$3.07506e-04$
Brown and Dennis 4-d	492	$8.58222e+04$	492	$8.58222e+04$	492	$8.58222e+04$
Quadratic 4-d	257	$2.17267e-09$	257	$2.17267e-09$	257	$2.17267e-09$
Penalty (1) 4-d	1246	$2.25412e-05$	1247	$2.25412e-05$	1247	$2.25412e-05$
Penalty (2) 4-d	200	$1.02580e-05$	200	$1.02580e-05$	200	$1.02580e-05$
Osbourne (1) 5-d	291	$6.91831e-05$	293	$6.91831e-05$	293	$6.91831e-05$
Brown almost linear 5-d	501	$4.22838e-10$	374	$2.91817e-09$	374	$2.91817e-09$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Biggs EXP6 6-d	2819	$3.03508e-11$	2831	$2.84764e-12$	2833	$2.84764e-12$
Extended Rosenbrock 6-d	2183	$3.59780e-07$	2277	$1.85383e-09$	2277	$1.85383e-09$
Brown almost-linear 7-d	989	$1.68953e-09$	825	$8.53723e-09$	826	$8.53723e-09$
Quadratic 8-d	677	$2.10376e-09$	806	$1.47953e-08$	610	$1.19657e-08$
Extended Rosenbrock 8-d	3303	$3.51475e-07$	4327	$6.83087e-07$	3524	$9.49918e-07$
Variably dimensional 8-d	1579	$3.04271e-08$	1642	$1.50237e-07$	1512	$2.72739e-07$
Extended Powell 8-d	1856	$1.16813e-09$	2495	$9.76922e-11$	1367	$1.01634e-06$
Watson 9-d	2069	$7.97411e-04$	4091	$5.96920e-06$	2127	$6.67076e-05$
Extended Rosenbrock 10-d	5791	$3.63214e-04$	4927	$1.39619e-05$	6378	$3.87908e-07$
Penalty (1) 10-d	7316	$7.08767e-05$	6307	$7.08795e-05$	7230	$7.08766e-05$
Penalty (2) 10-d	2307	$2.97884e-04$	2704	$2.97435e-04$	1457	$2.99054e-04$
Trigonometric 10-d	1170	$2.79905e-05$	1173	$2.79905e-05$	1173	$2.79905e-05$
Osbourne (2) 11-d	3744	$4.01560e-02$	3752	$4.01560e-02$	4314	$4.01416e-02$
Extended Powell 12-d	3930	$2.91125e-06$	4243	$1.63474e-07$	4449	$2.24341e-06$
Quadratic 16-d	1488	$6.09802e-09$	1553	$8.33109e-09$	1612	$1.73938e-08$
Quadratic 24-d	2795	$1.60197e-08$	2857	$4.04801e-08$	2857	$4.04801e-08$

Table G.29: Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	241	$1.13562e-09$	241	$1.13562e-09$	241	$1.13562e-09$
Freudenstein and Roth 2-d	160	$4.89843e+01$	160	$4.89843e+01$	160	$4.89843e+01$
Powell badly scaled 2-d	889	$2.65600e-15$	889	$2.65600e-15$	889	$2.65600e-15$
Brown badly scaled 2-d	468	$3.67093e-09$	468	$3.67093e-09$	468	$3.67093e-09$
Beale 2-d	128	$1.86459e-09$	128	$1.86459e-09$	128	$1.86459e-09$
Jennrich and Sampson 2-d	100	$1.24362e+02$	100	$1.24362e+02$	100	$1.24362e+02$
M^{c}Kinnon 2-d	195	$-2.50000e-01$	195	$-2.50000e-01$	195	$-2.50000e-01$
Helical valley 3-d	165	$1.11853e-04$	165	$1.11853e-04$	165	$1.11853e-04$
Bard 3-d	1691	$1.74287e+01$	1691	$1.74287e+01$	1691	$1.74287e+01$
Gaussian 3-d	77	$1.59743e-08$	77	$1.59743e-08$	77	$1.59743e-08$
Meyer 3-d	2104	$8.79459e+01$	2104	$8.79459e+01$	2104	$8.79459e+01$
Gulf research 3-d	957	$1.12798e-14$	957	$1.12798e-14$	957	$1.12798e-14$
Box 3-d	317	$1.77655e-12$	317	$1.77655e-12$	317	$1.77655e-12$
Powell singular 4-d	580	$8.81610e-12$	580	$8.81610e-12$	580	$8.81610e-12$
Wood 4-d	570	$1.38215e-08$	570	$1.38215e-08$	570	$1.38215e-08$
Kowalik and Osbourne 4-d	422	$3.07506e-04$	422	$3.07506e-04$	422	$3.07506e-04$
Brown and Dennis 4-d	492	$8.58222e+04$	492	$8.58222e+04$	492	$8.58222e+04$
Quadratic 4-d	257	$2.17267e-09$	257	$2.17267e-09$	257	$2.17267e-09$
Penalty (1) 4-d	1247	$2.25412e-05$	1247	$2.25412e-05$	1247	$2.25412e-05$
Penalty (2) 4-d	200	$1.02580e-05$	200	$1.02580e-05$	200	$1.02580e-05$
Osbourne (1) 5-d	293	$6.91831e-05$	293	$6.91831e-05$	293	$6.91831e-05$
Brown almost linear 5-d	374	$2.91817e-09$	374	$2.91817e-09$	374	$2.91817e-09$
Biggs EXP6 6-d	2833	$2.84764e-12$	2833	$2.84764e-12$	2833	$2.84764e-12$
Extended Rosenbrock 6-d	2278	$1.85383e-09$	2278	$1.85383e-09$	2278	$1.85383e-09$
Brown almost-linear 7-d	826	$8.53723e-09$	826	$8.53723e-09$	826	$8.53723e-09$
Quadratic 8-d	611	$1.19657e-08$	611	$1.19657e-08$	611	$1.19657e-08$
Extended Rosenbrock 8-d	3528	$9.49918e-07$	3528	$9.49918e-07$	3528	$9.49918e-07$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1812	$4.25605e-08$	1645	$3.02718e-08$	1981	$1.59528e-07$
Extended Powell 8-d	2177	$4.95165e-09$	1731	$9.25915e-08$	1731	$9.25915e-08$
Watson 9-d	2129	$6.67076e-05$	2976	$4.41046e-05$	2664	$3.60670e-05$
Extended Rosenbrock 10-d	6384	$3.87908e-07$	6387	$3.87908e-07$	6387	$3.87908e-07$
Penalty (1) 10-d	7396	$7.08803e-05$	7520	$7.08768e-05$	8750	$7.08766e-05$
Penalty (2) 10-d	842	$2.99518e-04$	6583	$2.95426e-04$	10829	$2.94005e-04$
Trigonometric 10-d	1173	$2.79905e-05$	1173	$2.79905e-05$	1173	$2.79905e-05$
Osbourne (2) 11-d	5482	$4.01381e-02$	5484	$4.01381e-02$	5484	$4.01381e-02$
Extended Powell 12-d	4152	$1.61245e-06$	2897	$8.11245e-06$	4748	$1.06114e-07$
Quadratic 16-d	1612	$1.73938e-08$	1685	$8.63381e-09$	1243	$1.14330e-08$
Quadratic 24-d	2857	$4.04801e-08$	3189	$7.27939e-08$	2292	$2.23357e-08$

Table G.30: Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_4–δ_6.

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	241	$1.13562e-09$	241	$1.13562e-09$	241	$1.13562e-09$
Freudenstein and Roth 2-d	160	$4.89843e+01$	160	$4.89843e+01$	160	$4.89843e+01$
Powell badly scaled 2-d	889	$2.65600e-15$	889	$2.65600e-15$	889	$2.65600e-15$
Brown badly scaled 2-d	468	$3.67093e-09$	468	$3.67093e-09$	468	$3.67093e-09$
Beale 2-d	128	$1.86459e-09$	128	$1.86459e-09$	128	$1.86459e-09$
Jennrich and Sampson 2-d	100	$1.24362e+02$	100	$1.24362e+02$	100	$1.24362e+02$
M^{c}Kinnon 2-d	195	$-2.50000e-01$	195	$-2.50000e-01$	195	$-2.50000e-01$
Helical valley 3-d	165	$1.11853e-04$	165	$1.11853e-04$	165	$1.11853e-04$
Bard 3-d	1692	$1.74287e+01$	1572	$1.74287e+01$	1441	$1.74287e+01$
Gaussian 3-d	77	$1.59743e-08$	77	$1.59743e-08$	77	$1.59743e-08$
Meyer 3-d	2104	$8.79459e+01$	2104	$8.79459e+01$	2104	$8.79459e+01$
Gulf research 3-d	957	$1.12798e-14$	957	$1.12798e-14$	957	$1.12798e-14$
Box 3-d	317	$1.77655e-12$	317	$1.77655e-12$	317	$1.77655e-12$
Powell singular 4-d	580	$8.81610e-12$	580	$8.81610e-12$	580	$8.81610e-12$
Wood 4-d	570	$1.38215e-08$	570	$1.38215e-08$	570	$1.38215e-08$
Kowalik and Osbourne 4-d	422	$3.07506e-04$	422	$3.07506e-04$	422	$3.07506e-04$
Brown and Dennis 4-d	492	$8.58222e+04$	492	$8.58222e+04$	492	$8.58222e+04$
Quadratic 4-d	257	$2.17267e-09$	257	$2.17267e-09$	257	$2.17267e-09$
Penalty (1) 4-d	1247	$2.25412e-05$	1247	$2.25412e-05$	1247	$2.25412e-05$
Penalty (2) 4-d	200	$1.02580e-05$	200	$1.02580e-05$	200	$1.02580e-05$
Osbourne (1) 5-d	293	$6.91831e-05$	293	$6.91831e-05$	293	$6.91831e-05$
Brown almost linear 5-d	374	$2.91817e-09$	374	$2.91817e-09$	374	$2.91817e-09$
Biggs EXP6 6-d	2833	$2.84764e-12$	2833	$2.84764e-12$	2833	$2.84764e-12$
Extended Rosenbrock 6-d	2278	$1.85383e-09$	2278	$1.85383e-09$	2278	$1.85383e-09$
Brown almost-linear 7-d	826	$8.53723e-09$	826	$8.53723e-09$	826	$8.53723e-09$
Quadratic 8-d	611	$1.19657e-08$	611	$1.19657e-08$	611	$1.19657e-08$
Extended Rosenbrock 8-d	3528	$9.49918e-07$	3528	$9.49918e-07$	3528	$9.49918e-07$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	1981	$1.59528e-07$	1981	$1.59528e-07$	1981	$1.59528e-07$
Extended Powell 8-d	1731	$9.25915e-08$	1731	$9.25915e-08$	1731	$9.25915e-08$
Watson 9-d	3969	$1.20346e-05$	3970	$1.20346e-05$	3970	$1.20346e-05$
Extended Rosenbrock 10-d	6387	$3.87908e-07$	6387	$3.87908e-07$	6387	$3.87908e-07$
Penalty (1) 10-d	7972	$7.08808e-05$	8591	$7.08765e-05$	7033	$7.08781e-05$
Penalty (2) 10-d	2165	$2.98196e-04$	2167	$2.98196e-04$	2167	$2.98196e-04$
Trigonometric 10-d	1173	$2.79905e-05$	1173	$2.79905e-05$	1173	$2.79905e-05$
Osbourne (2) 11-d	5484	$4.01381e-02$	5484	$4.01381e-02$	5484	$4.01381e-02$
Extended Powell 12-d	5162	$6.56641e-05$	5163	$6.56641e-05$	5163	$6.56641e-05$
Quadratic 16-d	1556	$1.58485e-08$	1557	$1.58485e-08$	1557	$1.58485e-08$
Quadratic 24-d	2516	$1.32077e-08$	2367	$2.88596e-08$	2615	$1.51491e-07$

Table G.31: Low tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_7–δ_9.

G.7.2 High tolerance results

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	330	$2.90725e-17$	330	$2.90725e-17$	330	$2.90725e-17$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1052	$5.81912e-26$	1055	$5.81912e-26$	1056	$5.81912e-26$
Brown badly scaled 2-d	581	$4.31731e-17$	586	$4.31731e-17$	586	$4.31731e-17$
Beale 2-d	224	$3.95699e-18$	224	$3.95699e-18$	224	$3.95699e-18$
Jennrich and Sampson 2-d	207	$1.24362e+02$	207	$1.24362e+02$	207	$1.24362e+02$
M^{c}Kinnon 2-d	411	$-2.50000e-01$	397	$-2.50000e-01$	397	$-2.50000e-01$
Helical valley 3-d	447	$5.66748e-17$	447	$5.66748e-17$	447	$5.66748e-17$
Bard 3-d	1498	$1.74287e+01$	1498	$1.74287e+01$	1742	$1.74287e+01$
Gaussian 3-d	249	$1.12793e-08$	249	$1.12793e-08$	249	$1.12793e-08$
Meyer 3-d	2471	$8.79459e+01$	2476	$8.79459e+01$	2477	$8.79459e+01$
Gulf research 3-d	1112	$7.11219e-22$	1114	$7.11219e-22$	1114	$7.11219e-22$
Box 3-d	544	$9.80770e-22$	510	$5.11265e-21$	510	$5.11265e-21$
Powell singular 4-d	823	$5.99302e-17$	1347	$8.87082e-23$	1348	$8.87082e-23$
Wood 4-d	787	$2.68521e-16$	787	$2.68521e-16$	787	$2.68521e-16$
Kowalik and Osbourne 4-d	628	$3.07506e-04$	628	$3.07506e-04$	628	$3.07506e-04$
Brown and Dennis 4-d	801	$8.58222e+04$	801	$8.58222e+04$	801	$8.58222e+04$
Quadratic 4-d	438	$2.18326e-17$	438	$2.18326e-17$	438	$2.18326e-17$
Penalty (1) 4-d	1718	$2.24998e-05$	1719	$2.24998e-05$	1719	$2.24998e-05$
Penalty (2) 4-d	5538	$9.37629e-06$	10064	$9.37629e-06$	10065	$9.37629e-06$
Osbourne (1) 5-d	1862	$5.46489e-05$	1864	$5.46489e-05$	1866	$5.46489e-05$
Brown almost linear 5-d	865	$6.44921e-18$	753	$2.50863e-17$	753	$2.50863e-17$
Biggs EXP6 6-d	3372	$2.93965e-16$	3414	$6.00186e-20$	3416	$6.00186e-20$
Extended Rosenbrock 6-d	3337	$2.46270e-17$	2956	$9.92346e-17$	2956	$9.92346e-17$
Brown almost-linear 7-d	1845	$1.63582e-17$	1845	$3.52023e-17$	1846	$3.52023e-17$

	δ_1		δ_2		δ_3	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Quadratic 8-d	1026	$3.30772e-17$	1290	$5.13854e-16$	1126	$5.00766e-16$
Extended Rosenbrock 8-d	4923	$5.02697e-14$	5462	$7.65359e-17$	5825	$6.96963e-16$
Variably dimensional 8-d	2136	$7.76437e-16$	2645	$3.59159e-15$	2682	$3.56603e-16$
Extended Powell 8-d	7003	$1.01817e-23$	7028	$1.83747e-24$	6863	$4.01460e-21$
Watson 9-d	7637	$1.39976e-06$	6588	$1.39976e-06$	6838	$1.39976e-06$
Extended Rosenbrock 10-d	12273	$1.61152e-16$	8395	$1.60467e-13$	10637	$1.48964e-13$
Penalty (1) 10-d	11651	$7.08765e-05$	11414	$7.08765e-05$	10066	$7.08765e-05$
Penalty (2) 10-d	27122	$2.93661e-04$	33096	$2.93661e-04$	45675	$2.93661e-04$
Trigonometric 10-d	2380	$2.79506e-05$	2383	$2.79506e-05$	2383	$2.79506e-05$
Osbourne (2) 11-d	7342	$4.01377e-02$	7594	$4.01377e-02$	8049	$4.01377e-02$
Extended Powell 12-d	26246	$3.52429e-13$	24424	$1.23646e-21$	24503	$1.06023e-17$
Quadratic 16-d	2269	$1.89810e-16$	2511	$1.99537e-16$	2810	$9.16920e-17$
Quadratic 24-d	4330	$3.30148e-16$	4293	$1.39638e-16$	4294	$1.39638e-16$

Table G.32: High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_1–δ_3.

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	330	$2.90725e-17$	330	$2.90725e-17$	330	$2.90725e-17$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1056	$5.81912e-26$	1056	$5.81912e-26$	1056	$5.81912e-26$
Brown badly scaled 2-d	586	$4.31731e-17$	586	$4.31731e-17$	586	$4.31731e-17$
Beale 2-d	224	$3.95699e-18$	224	$3.95699e-18$	224	$3.95699e-18$
Jennrich and Sampson 2-d	207	$1.24362e+02$	207	$1.24362e+02$	207	$1.24362e+02$
M^{c}Kinnon 2-d	397	$-2.50000e-01$	397	$-2.50000e-01$	397	$-2.50000e-01$
Helical valley 3-d	447	$5.66748e-17$	447	$5.66748e-17$	447	$5.66748e-17$
Bard 3-d	1743	$1.74287e+01$	1743	$1.74287e+01$	1743	$1.74287e+01$
Gaussian 3-d	249	$1.12793e-08$	249	$1.12793e-08$	249	$1.12793e-08$
Meyer 3-d	2479	$8.79459e+01$	2479	$8.79459e+01$	2479	$8.79459e+01$
Gulf research 3-d	1114	$7.11219e-22$	1114	$7.11219e-22$	1114	$7.11219e-22$
Box 3-d	510	$5.11265e-21$	510	$5.11265e-21$	510	$5.11265e-21$
Powell singular 4-d	1347	$7.17793e-23$	1349	$7.17793e-23$	1349	$7.17793e-23$
Wood 4-d	787	$2.68521e-16$	787	$2.68521e-16$	787	$2.68521e-16$
Kowalik and Osbourne 4-d	628	$3.07506e-04$	628	$3.07506e-04$	628	$3.07506e-04$
Brown and Dennis 4-d	801	$8.58222e+04$	801	$8.58222e+04$	801	$8.58222e+04$
Quadratic 4-d	438	$2.18326e-17$	438	$2.18326e-17$	438	$2.18326e-17$
Penalty (1) 4-d	1719	$2.24998e-05$	1719	$2.24998e-05$	1719	$2.24998e-05$
Penalty (2) 4-d	10065	$9.37629e-06$	10065	$9.37629e-06$	10065	$9.37629e-06$
Osbourne (1) 5-d	1866	$5.46489e-05$	1868	$5.46489e-05$	1868	$5.46489e-05$
Brown almost linear 5-d	753	$2.50863e-17$	753	$2.50863e-17$	753	$2.50863e-17$
Biggs EXP6 6-d	3416	$6.00186e-20$	3416	$6.00186e-20$	3416	$6.00186e-20$
Extended Rosenbrock 6-d	2957	$9.92346e-17$	2957	$9.92346e-17$	2957	$9.92346e-17$
Brown almost-linear 7-d	1846	$3.52023e-17$	1846	$3.52023e-17$	1846	$3.52023e-17$
Quadratic 8-d	1127	$5.00766e-16$	1127	$5.00766e-16$	1127	$5.00766e-16$
Extended Rosenbrock 8-d	5829	$6.96963e-16$	5829	$6.96963e-16$	5829	$6.96963e-16$

	δ_4		δ_5		δ_6	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	3220	$1.46753e-15$	2245	$9.57800e-16$	3130	$1.20403e-15$
Extended Powell 8-d	5717	$6.71301e-21$	7386	$2.35556e-22$	7388	$2.35556e-22$
Watson 9-d	6840	$1.39976e-06$	7162	$1.39976e-06$	5802	$1.39976e-06$
Extended Rosenbrock 10-d	10419	$1.62014e-13$	10423	$1.62014e-13$	10423	$1.62014e-13$
Penalty (1) 10-d	13904	$7.08765e-05$	10327	$7.08765e-05$	12884	$7.08765e-05$
Penalty (2) 10-d	26503	$2.93661e-04$	25129	$2.93661e-04$	25746	$2.93661e-04$
Trigonometric 10-d	2383	$2.79506e-05$	2383	$2.79506e-05$	2383	$2.79506e-05$
Osbourne (2) 11-d	10010	$4.01377e-02$	10062	$4.01377e-02$	10063	$4.01377e-02$
Extended Powell 12-d	15987	$2.92816e-17$	37470	$6.47893e-22$	19570	$2.43058e-17$
Quadratic 16-d	2810	$9.16920e-17$	2670	$6.18697e-16$	2240	$1.32233e-16$
Quadratic 24-d	4294	$1.39638e-16$	5065	$4.76431e-16$	3702	$4.42648e-16$

Table G.33: High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_4–δ_6.

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Rosenbrock 2-d	330	$2.90725e-17$	330	$2.90725e-17$	330	$2.90725e-17$
Freudenstein and Roth 2-d	257	$4.89843e+01$	257	$4.89843e+01$	257	$4.89843e+01$
Powell badly scaled 2-d	1056	$5.81912e-26$	1056	$5.81912e-26$	1056	$5.81912e-26$
Brown badly scaled 2-d	586	$4.31731e-17$	586	$4.31731e-17$	586	$4.31731e-17$
Beale 2-d	224	$3.95699e-18$	224	$3.95699e-18$	224	$3.95699e-18$
Jennrich and Sampson 2-d	207	$1.24362e+02$	207	$1.24362e+02$	207	$1.24362e+02$
M^{c}Kinnon 2-d	397	$-2.50000e-01$	397	$-2.50000e-01$	397	$-2.50000e-01$
Helical valley 3-d	447	$5.66748e-17$	447	$5.66748e-17$	447	$5.66748e-17$
Bard 3-d	1744	$1.74287e+01$	1624	$1.74287e+01$	1441	$1.74287e+01$
Gaussian 3-d	249	$1.12793e-08$	249	$1.12793e-08$	249	$1.12793e-08$
Meyer 3-d	2479	$8.79459e+01$	2479	$8.79459e+01$	2479	$8.79459e+01$
Gulf research 3-d	1114	$7.11219e-22$	1114	$7.11219e-22$	1114	$7.11219e-22$
Box 3-d	510	$5.11265e-21$	510	$5.11265e-21$	510	$5.11265e-21$
Powell singular 4-d	1349	$7.17793e-23$	1349	$7.17793e-23$	1349	$7.17793e-23$
Wood 4-d	787	$2.68521e-16$	787	$2.68521e-16$	787	$2.68521e-16$
Kowalik and Osbourne 4-d	628	$3.07506e-04$	628	$3.07506e-04$	628	$3.07506e-04$
Brown and Dennis 4-d	801	$8.58222e+04$	801	$8.58222e+04$	801	$8.58222e+04$
Quadratic 4-d	438	$2.18326e-17$	438	$2.18326e-17$	438	$2.18326e-17$
Penalty (1) 4-d	1719	$2.24998e-05$	1719	$2.24998e-05$	1719	$2.24998e-05$
Penalty (2) 4-d	10065	$9.37629e-06$	10065	$9.37629e-06$	10066	$9.37629e-06$
Osbourne (1) 5-d	1868	$5.46489e-05$	1868	$5.46489e-05$	1868	$5.46489e-05$
Brown almost linear 5-d	753	$2.50863e-17$	753	$2.50863e-17$	753	$2.50863e-17$
Biggs EXP6 6-d	3416	$6.00186e-20$	3416	$6.00186e-20$	3416	$6.00186e-20$
Extended Rosenbrock 6-d	2957	$9.92346e-17$	2957	$9.92346e-17$	2957	$9.92346e-17$
Brown almost-linear 7-d	1846	$3.52023e-17$	1846	$3.52023e-17$	1846	$3.52023e-17$
Quadratic 8-d	1127	$5.00766e-16$	1127	$5.00766e-16$	1127	$5.00766e-16$
Extended Rosenbrock 8-d	5829	$6.96963e-16$	5829	$6.96963e-16$	5829	$6.96963e-16$

	δ_7		δ_8		δ_9	
Function	*FE*	*Minimum*	*FE*	*Minimum*	*FE*	*Minimum*
Variably dimensional 8-d	3130	$1.20403e-15$	3130	$1.20403e-15$	3130	$1.20403e-15$
Extended Powell 8-d	7389	$2.35556e-22$	7391	$2.35556e-22$	7391	$2.35556e-22$
Watson 9-d	6486	$1.39976e-06$	6489	$1.39976e-06$	6489	$1.39976e-06$
Extended Rosenbrock 10-d	10423	$1.62014e-13$	10423	$1.62014e-13$	10423	$1.62014e-13$
Penalty (1) 10-d	15445	$7.08765e-05$	12581	$7.08765e-05$	12153	$7.08765e-05$
Penalty (2) 10-d	36915	$2.93661e-04$	36922	$2.93661e-04$	36924	$2.93661e-04$
Trigonometric 10-d	2383	$2.79506e-05$	2383	$2.79506e-05$	2383	$2.79506e-05$
Osbourne (2) 11-d	10063	$4.01377e-02$	10063	$4.01377e-02$	10063	$4.01377e-02$
Extended Powell 12-d	23227	$8.08558e-21$	23233	$8.08558e-21$	23233	$8.08558e-21$
Quadratic 16-d	2504	$4.85337e-17$	2505	$4.85337e-17$	2505	$4.85337e-17$
Quadratic 24-d	3827	$6.74238e-17$	3898	$2.48074e-16$	4177	$2.71881e-16$

Table G.34: High tolerance results for NM4_$\psi_3\kappa_2\nu_8$ δ_7–δ_9.

Appendix H

MATLAB code

This appendix contains the MATLAB code for FMINSEARCH and the most successful variant NM4_$\psi_3\delta_6\kappa_1\nu_8$. FMINSEARCH is self contained, while NM4_$\psi_3\delta_6\kappa_1\nu_8$ makes calls to the functions SIMPLEX, FRAME and NM_STD. The code for SIMPLEX and FRAME is also included. The function NM_STD is the Nelder-Mead step of FMINSEARCH moved to a separate function and so is not listed separately.

All of the variants used FMINSEARCH as a "skeleton" and so in general retain the same basic code — all of FMINSEARCH's output options are maintained. Except for FMINSEARCH, the code presented here has been left, to some extent, in "debug mode". Any final implementation could be streamlined and some of the displayed messages removed.

For completeness the code used to generate the HH orthogonal simplex has also been included.

Some slight changes have been made to the code that is presented here. The changes only effect the "white space" of the code and have been made purely for layout reasons — in particular, to prevent a line of code from wrapping onto the next line.

H.1 FMINSEARCH

```
function [x,fval,exitflag,output] = fminsearch(funfcn,x,options,varargin)
%FMINSEARCH Multidimensional unconstrained nonlinear minimization
%   (Nelder-Mead).
%   X = FMINSEARCH(FUN,X0) returns a vector X that is a local
%   minimizer of the function that is described in FUN
%   (usually an M-file: FUN.M)
%   near the starting vector X0.  FUN should return a scalar function
%   value when called with feval: F=feval(FUN,X).  See below for more
%   options for FUN.
%
%   X = FMINSEARCH(FUN,X0,OPTIONS)  minimizes with the default
%   optimization parameters replaced by values in the structure
%   OPTIONS, created with the OPTIMSET function.
%   See OPTIMSET for details.  FMINSEARCH uses these options:
%   Display, TolX, TolFun, MaxFunEvals, and MaxIter.
%
%   X = FMINSEARCH(FUN,X0,OPTIONS,P1,P2,...) provides for additional
%   arguments which are passed to the objective function,
%   F=feval(FUN,X,P1,P2,...).
%   Pass an empty matrix for OPTIONS to use the default values.
%   (Use OPTIONS = [] as a place holder if no options are set.)
%
%   [X,FVAL]= FMINSEARCH(...) returns the value of the objective
%   function, described in FUN, at X.
%
%   [X,FVAL,EXITFLAG] = FMINSEARCH(...) returns a string EXITFLAG
%   that describes the exit condition of FMINSEARCH.
%   If EXITFLAG is:
%     1 then FMINSEARCH converged with a solution X.
%     0 then the maximum number of iterations was reached.
%
%   [X,FVAL,EXITFLAG,OUTPUT] = FMINSEARCH(...) returns a structure
%   OUTPUT with the number of iterations taken in OUTPUT.iterations.
%
%   The argument FUN can be an inline function:
%      f = inline('norm(x)');
%      x = fminsearch(f,[1;2;3]);
%
%   FMINSEARCH uses the Nelder-Mead simplex (direct search) method.
%
%   See also FMINBND, OPTIMSET, OPTIMGET.
%
%
```

```
%   Reference: Jeffrey C. Lagarias, James A. Reeds,
%   Margaret H. Wright, Paul E. Wright,
%   "Convergence Properties of the Nelder-Mead Simplex
%   Algorithm in Low Dimensions", May 1, 1997.
%   To appear in the SIAM Journal of Optimization.
%
%   Copyright (c) 1984-98 by The MathWorks, Inc.
%   $Revision: 1.8 $  $Date: 1998/10/23 20:52:22 $
%
defaultopt = optimset('display','final','maxiter',...
   '200*numberOfVariables','maxfunevals',...
   '200*numberOfVariables','TolX',1e-4,'TolFun',1e-4);
% If just 'defaults' passed in, return the
% default options in X
if nargin==1 & nargout <= 1 & isequal(funfcn,'defaults')
   x = defaultopt;
   return
end

if nargin<3, options = []; end
n = prod(size(x));
numberOfVariables = n;

options = optimset(defaultopt,options);
printtype = optimget(options,'display');
tolx = optimget(options,'tolx');
tolf = optimget(options,'tolfun');
maxfun = optimget(options,'maxfuneval');
maxiter = optimget(options,'maxiter');
% In case the defaults were gathered from
% calling: optimset('fminsearch'):
if ischar(maxfun)
   maxfun = eval(maxfun);
end
if ischar(maxiter)
   maxiter = eval(maxiter);
end

switch printtype
case {'none','off'}
   prnt = 0;
case 'iter'
   prnt = 2;
case 'final'
```

```
   prnt = 1;
case 'simplex'
   prnt = 3;
otherwise
   prnt = 1;
end

header = ' Iteration  Func-count  min f(x)  Procedure';

% Convert to inline function as needed.
funfcn = fcnchk(funfcn,length(varargin));

n = prod(size(x));

% Initialize parameters
rho = 1; chi = 2; psi = 0.5; sigma = 0.5;
onesn = ones(1,n);
two2np1 = 2:n+1;
one2n = 1:n;

% Set up a simplex near the initial guess.
xin = x(:); % Force xin to be a column vector
v = zeros(n,n+1); fv = zeros(1,n+1);
v = xin;       % Place input guess in the simplex!
               % (credit L.Pfeffer at Stanford)
x(:) = xin;    % Change x to the form expected by funfcn
fv = feval(funfcn,x,varargin{:});

% Following improvement suggested by L.Pfeffer at Stanford
usual_delta = 0.05;           % 5 percent deltas for non-zero
zero_term_delta = 0.00025;    % Even smaller delta for zero
                              % elements of x
for j = 1:n
   y = xin;
   if y(j) ~= 0
      y(j) = (1 + usual_delta)*y(j);
   else
      y(j) = zero_term_delta;
   end
   v(:,j+1) = y;
   x(:) = y; f = feval(funfcn,x,varargin{:});
   fv(1,j+1) = f;
end
```

```
% sort so v(1,:) has the lowest function value
[fv,j] = sort(fv);
v = v(:,j);

how = 'initial';
itercount = 1;
func_evals = n+1;
if prnt == 2
   disp(' ')
   disp(header)
   disp([sprintf(' %5.0f   %5.0f   %12.6g   ', ...
      itercount, func_evals, fv(1)), how])
elseif prnt == 3
   clc
   formatsave = get(0,{'format','formatspacing'});
   format compact
   format short e
   disp(' ')
   disp(how)
   v
   fv
   func_evals
end
exitflag = 1;

% Main algorithm
% Iterate until the diameter of the simplex is less than tolx
%   AND the function values differ from the min by less than tolf,
%   or the max function evaluations are exceeded.
%   (Cannot use OR instead of AND.)
while func_evals < maxfun & itercount < maxiter
   if max(max(abs(v(:,two2np1)-v(:,onesn)))) <= tolx & ...
         max(abs(fv(1)-fv(two2np1))) <= tolf
      break
   end
   how = '';

   % Compute the reflection point

   % xbar = average of the n (NOT n+1) best points
   xbar = sum(v(:,one2n), 2)/n;
   xr = (1 + rho)*xbar - rho*v(:,end);
   x(:) = xr; fxr = feval(funfcn,x,varargin{:});
   func_evals = func_evals+1;
```

```
if fxr < fv(:,1)
   % Calculate the expansion point
   xe = (1 + rho*chi)*xbar - rho*chi*v(:,end);
   x(:) = xe; fxe = feval(funfcn,x,varargin{:});
   func_evals = func_evals+1;
   if fxe < fxr
      v(:,end) = xe;
      fv(:,end) = fxe;
      how = 'expand';
   else
      v(:,end) = xr;
      fv(:,end) = fxr;
      how = 'reflect';
   end
else % fv(:,1) <= fxr
   if fxr < fv(:,n)
      v(:,end) = xr;
      fv(:,end) = fxr;
      how = 'reflect';
   else % fxr >= fv(:,n)
      % Perform contraction
      if fxr < fv(:,end)
         % Perform an outside contraction
         xc = (1 + psi*rho)*xbar - psi*rho*v(:,end);
         x(:) = xc; fxc = feval(funfcn,x,varargin{:});
         func_evals = func_evals+1;

         if fxc <= fxr
            v(:,end) = xc;
            fv(:,end) = fxc;
            how = 'contract outside';
         else
            % perform a shrink
            how = 'shrink';
         end
      else
         % Perform an inside contraction
         xcc = (1-psi)*xbar + psi*v(:,end);
         x(:) = xcc; fxcc = feval(funfcn,x,varargin{:});
         func_evals = func_evals+1;

         if fxcc < fv(:,end)
            v(:,end) = xcc;
```

```
            fv(:,end) = fxcc;
            how = 'contract inside';
         else
            % perform a shrink
            how = 'shrink';
         end
      end
      if strcmp(how,'shrink')
         for j=two2np1
            v(:,j)=v(:,1)+sigma*(v(:,j) - v(:,1));
            x(:) = v(:,j); fv(:,j) = feval(funfcn,x,varargin{:});
         end
         func_evals = func_evals + n;
      end
   end
end
[fv,j] = sort(fv);
v = v(:,j);
itercount = itercount + 1;
if prnt == 2
disp([sprintf(' %5.0f   %5.0f   %12.6g   ', ...
   itercount, func_evals, fv(1)), how])
elseif prnt == 3
   disp(' ')
   disp(how)
   v
   fv
   func_evals
end
end   % while


x(:) = v(:,1);
if prnt == 3,
   % reset format
   set(0,{'format','formatspacing'},formatsave);
end
output.iterations = itercount;
output.funcCount = func_evals;
output.algorithm = 'Nelder-Mead simplex direct search';

fval = min(fv);
if func_evals >= maxfun
   if prnt > 0
```

```
      disp(' ')
      disp('Exiting: ')
      disp('Maximum number of function evaluations has been exceeded')
      disp('         - increase MaxFunEvals option.')
      msg = sprintf('         Current function value: %f \n', fval);
      disp(msg)
   end
   exitflag = 0;
elseif itercount >= maxiter
   if prnt > 0
      disp(' ')
      disp('Exiting: Maximum number of iterations has been exceeded')
      disp('         - increase MaxIter option.')
      msg = sprintf('         Current function value: %f \n', fval);
      disp(msg)
   end
   exitflag = 0;
else
   if prnt > 0
      convmsg1 = sprintf([ ...
         '\nOptimization terminated successfully:\n',...
         ' the current x satisfies the termination criteria\n', ...
         ' using OPTIONS.TolX of %e \n', ...
         ' and F(X) satisfies the convergence criteria\n', ...
         ' using OPTIONS.TolFun of %e \n'], ...
         options.TolX, options.TolFun);
      disp(convmsg1)
      exitflag = 1;
   end
end
```

H.2 NM4_$\psi_3\delta_6\kappa_1\nu_8$

```
function [x,fval,exitflag,output] = NM4_s3d6k1n8(fun,x,options)
% NM4 is (basically) Matlab's FMINSEARCH with frame based
% convergence algorithm added to it and the NM step moved
% to a separate function.
%
% Other functions inside this one:  eval_func, vsort
%
% Other functions called by this one:
% simplex
% frame
% NM_std
%
% define global variables so the NM step can see which function
% to use and update the number of function evaluations
global funfcn func_evals
funfcn = fun;

% set up Matlab's fminsearch environment
defaultopt = optimset('display','final','maxiter', ...
   '200*numberOfVariables','maxfunevals',...
   '200*numberOfVariables','TolX',1e-4,'TolFun',1e-4);
if nargin<3, options = []; end
n = prod(size(x));
numberOfVariables = n;
options = optimset(defaultopt,options);
printtype = optimget(options,'display');
tolx = optimget(options,'tolx');
tolf = optimget(options,'tolfun');
maxfun = optimget(options,'maxfuneval');
maxiter = optimget(options,'maxiter');
switch printtype
case {'none','off'}
   prnt = 0;
case 'iter'
   prnt = 2;
case 'final'
   prnt = 1;
case 'simplex'
   prnt = 3;
otherwise
   prnt = 1;
end
```

```
header = ' Iteration   Func-count     min f(x)     Procedure';
header = [header blanks(15) 'QMF         '];
header = [header 'Det      Max length   Status'];

x = x(:);
rho = 1; chi = 2; psi = 0.5; sigma = 0.5;
onesn   = ones(1,n);
two2np1 = 2:n+1;
two2np2 = 2:n+2;
one2n   = 1:n;
one2np1 = 1:n+1;

% set up parameters for new algorithm
h = 1;  K = 1.0e+03;  N = 100;
nu = 4.50;             % shrink factor for epsilon
kappa = 0.25;          % shrink factor for h
delta = 1e-18;         % min for det. before collapse signalled

% keep track of numbers of each step performed
reflect = 0;
expand = 0;
cont_outside = 0;
cont_inside = 0;
shrink = 0;
frames = 0;
reshapes = 0;
total_qmf = 0;

% setup initial simplex using Matlab's fminsearch's method
v = simplex('mat',x);
fv = eval_func(v);
func_evals = n+1;
mu = prod(frame('l',v));
[v fv] = vsort(v,fv);
N = (fv(end) - fv(1))/(N * n * h^nu);
epsilon = N * h^nu;
itercount = 1;
how = 'initial';

if prnt == 2
   disp(' ')
   disp(header)
   disp([sprintf(' %5.0f   %10.0f  %15.6g   %s', ...
      itercount, func_evals, fv(1)), how])
```

```
elseif prnt == 3
   clc
   formatsave = get(0,{'format','formatspacing'});
   format compact
   format short e
   disp(' ')
   disp(how)
   v
   fv
   func_evals
end

exitflag = 1;

% Main algorithm
% Iterate until the diameter of the simplex is less than
% tolx AND the function values differ from the min by less
% than tolf, or the max function evaluations are exceeded.
% (Cannot use OR instead of AND.)
while func_evals < maxfun & itercount < maxiter
   if max(max(abs(v(:,two2np1)-v(:,onesn)))) <= tolx & ...
         max(abs(fv(1)-fv(two2np1))) <= tolf
      break
   end

   basis_status = 'ok';
   more_info = 0;
   qmf = 0;

   % generate new (sorted) simplex by NM step
   [vnm fnm how] = nm_std(v, fv);

   % new algorithm begins
   if fv(:,end) - fnm(:,end) >= epsilon
      % NM made sufficient progress, accept simplex
      v = vnm;
      fv = fnm;

      switch how
      case 'reflect'
         reflect = reflect + 1;
      case 'expand'
         expand = expand + 1;
         mu = mu * chi;
```

```
      case 'contract outside'
         cont_outside = cont_outside + 1;
         mu = mu * psi;
      case 'contract inside'
         cont_inside = cont_inside + 1;
         mu = mu * psi;
      case 'shrink'
         shrink = shrink + 1;
         mu = mu * sigma^n;
      end

   else
      % NM made insufficient progress

      more_info = 1;
      lengths = frame('l',v,h);
      sides_prod = h^n * prod(lengths);

      % calculate determinant
      if sides_prod > 0,
         basis_det = mu / sides_prod;
      else
         basis_det = 0;
         basis_status = 'basis too small';
      end

      max_length = max(lengths);

      % currently, don't do anything if the basis is too big
      if basis_det > 0 & max_length > K,
         basis_status = 'basis too big';
      end

      if basis_det < delta,
         % basis unacceptable

         if basis_det > 0,
            basis_status = 'det collapsed';
         end

         % reshape simplex
         v = simplex('qr',v);
         fv(two2np1) = eval_func(v(:,two2np1));
         func_evals = func_evals + n;
```

```
      mu  = prod(frame('l',v));
      reshape_flag = 1;
      reshapes = reshapes + 1;
      how = 'reshape';
   else
      reshape_flag = 0;
      frames = frames + 1;
      how = 'frame';
   end

   % given the simplex, complete the frame
   v_frame = frame('f',v);
   fv_frame = [fv eval_func(v_frame(:,end))];
   func_evals = func_evals + 1;

   % until there is epsilon descent, reduce the frame
   while max([fv_frame(1) - fv_frame(two2np2)]) <= epsilon
      % have we been in the qmf too long?
      if func_evals > maxfun | itercount > maxiter,
         basis_status = 'stuck in QMF';
         exitflag = 0;
         break
      end

      if reshape_flag == 0,
         % reshape the simplex
         v = simplex('qr',v_frame(:,1:n+1));
         fv(two2np1) = eval_func(v(:,two2np1));
         func_evals = func_evals + n;
         mu = prod(frame('l',v));

         % complete the frame
         v_frame = frame('f',v);
         fv_frame = [fv eval_func(v_frame(:,end))];
         func_evals = func_evals + 1;

         reshape_flag = 1;
         reshapes = reshapes + 1;
         how = [how ' + reshape'];
      else
         qmf = qmf + 1;
         kappa = -kappa;
         h = abs(h * kappa);
         epsilon = N * h^nu;
```

```
        % shrink the frame
        v_frame = frame('sh', v_frame, kappa);
        mu = mu * abs(kappa)^n;
        temp = v_frame(:,two2np1) - v_frame(:,onesn)
        simplex_diam = max(max(abs(temp)));

        % has the frame collapsed?
        if simplex_diam == 0,
           fv_frame(two2np2) = fv_frame(1);
           basis_status = [basis_status ' + frame zero'];
           break
        else
           fv_frame(two2np2) = eval_func(v_frame(:,two2np2));
           func_evals = func_evals + n+1;
        end

        % has required tolerance been reached?
        if simplex_diam <= tolx & ...
              max(abs(fv_frame(1) - fv_frame(two2np1))) <= tolf,
           break
        end
     end
  end

  total_qmf = total_qmf + qmf;

  % reduce frame to simplex again
  v  =  v_frame(:,one2np1);
  fv = fv_frame(:,one2np1);

  % check if new frame point is the lowest
  if fv_frame(end) < fv(1),
     v(:,1) = v_frame(:,end);
     fv(1) = fv_frame(end);
     if strcmp(how, 'reshape'),
        mu = mu * (n+1);
     else
        mu = mu * chi;
     end
     how = [how ' + swap'];
  end
end
```

```
   [v fv] = vsort(v, fv);

   % original algorithm continues on from here
   itercount = itercount + 1;
   if prnt == 2
      how = [how blanks(22 - length(how))];
      if more_info,
      disp([sprintf('%5.0f %12.0f %16.6g %25s %3.0f %12.4g %9.3g %s',...
            itercount, func_evals, fv(1), how, qmf, ...
            basis_det, max_length, basis_status) ])
      else
         disp([sprintf(' %5.0f        %5.0f     %12.6g     %18s', ...
               itercount, func_evals, fv(1), how) ])
      end

   elseif prnt == 3
      disp(' ')
      disp(how)
      v
      fv
      func_evals
   end
end   % while

x = v(:,1);

if prnt == 3,
   % reset format
   set(0,{'format','formatspacing'},formatsave);
end
output.iterations = itercount;
output.funcCount = func_evals;
output.algorithm = 'Frame based convergence algorithm';

fval = min(fv);
if func_evals >= maxfun
   if prnt > 0
      disp(' ')
      disp('Exiting: ')
      disp('Maximum number of function evaluations has been exceeded')
      disp('          - increase MaxFunEvals option.')
      msg = sprintf('         Current function value: %f \n', fval);
      disp(msg)
   end
```

```
   exitflag = 0;
elseif itercount >= maxiter
   if prnt > 0
      disp(' ')
      disp('Exiting: Maximum number of iterations has been exceeded')
      disp('         - increase MaxIter option.')
      msg = sprintf('         Current function value: %f \n', fval);
      disp(msg)
   end
   exitflag = 0;
else
   if prnt > 0
      convmsg1 = sprintf([ ...
         '\nOptimization terminated successfully:\n',...
         ' the current x satisfies the termination criteria\n', ...
         ' using OPTIONS.TolX of %e \n', ...
         ' and F(X) satisfies the convergence criteria\n', ...
         ' using OPTIONS.TolFun of %e \n'], ...
         options.TolX, options.TolFun);
      disp(convmsg1)
      exitflag = 1;
   end
end
output.Reflect = reflect;
output.Expand = expand;
output.Outside = cont_outside;
output.Inside = cont_inside;
output.Frames = frames;
output.Total_QMF = total_qmf;
output.Reshapes = reshapes;
```

```
%-----------------------------------------------------------
function fv = eval_func(v);
global funfcn
n = size(v,2);
fv = zeros(1,n);
% evaluate function values
for j = 1:n,
   fv(j) = feval(funfcn,v(:,j));
end

%-----------------------------------------------------------
function [newv,newfv] = vsort(v, fv)
% sort so v(1,:) has the lowest function value
[newfv,j] = sort(fv);
newv = v(:,j);

%-----------------------------------------------------------
```

H.2.1 SIMPLEX

```
function v = simplex(type,x,b)
% SIMPLEX generates a matrix whose columns are the vertices
% of an orthogonal simplex with root-vertex x.
%
% v = simplex(type, x, [b])
%
% Where type is either:
%
% mckinnon = proper starting simplex for mckinnon's example
%      mat = Matlab's method based on initial vertex x
%       qr = QR decomposition of basis b about initial vertex x
%      ijk = orth. reg. simplex about x with side lengths in b
%       hh = orth. simplex about x which is the
%            'orthogonal complement' of -b

switch type
case 'mckinnon'
   SQR33 = sqrt(33);
   lambda1 = (1 + SQR33)/8;
   lambda2 = (1 - SQR33)/8;
   v = [0 lambda1 1; 0 lambda2 1];
otherwise
   if nargin < 2,
      error('Insufficient information supplied')
   end

   % initialise
   n = size(x,1);

   if n == 1;
      x = x(:);
      n = length(x);
   end

   v = zeros(n,n+1);
   v(:,1) = x(:,1);

   switch type
   case 'mat'
      usual_delta = 0.05;
      zero_term_delta = 0.00025;
      for j = 1:n
         y = x;
```

```
      if y(j) ~= 0
         y(j) = (1 + usual_delta)*y(j);
      else
         y(j) = zero_term_delta;
      end
      v(:,j+1) = y;
   end

case 'ijk'
   if nargin ~= 3,
      error('Side length information has not been supplied'),
   end
   if length(b) == 1,
      % create a regular orthogonal simplex
      b = ones(1,n)*b;
   end
   % create an orthogonal simplex using side_lengths
   % check there are no zero lengths
   if any(b == 0)
      error('Side lengths must be non-zero')
   elseif length(b) < n,
      error('Not enough side lengths given')
   else
      for j = 1:n
         y = x;
         y(j) = y(j) + b(j);
         v(:,j+1) = y;
      end
   end

case 'hh'
   % create simplex about x(:,1) with orthogonal
   % decomposition from vector -b
   % sum of orthogonal decomposition vectors = -b

   % given simplex, find the longest basis vector
   [basis basis_lengths] = frame('pb',x);
   j = find(basis_lengths(1:n) == max(basis_lengths(1:n)));
   j = j(1);
   bl = basis(:,j);

   % calculate the orthogonal decomposition, vectors sum to -b
   basis_orth = hh(bl);
```

```
      % create the new simplex
      for j = 1:n,
         v(:,j+1) = v(:,1) + basis_orth(:,j);
      end

   case 'qr'
      % create simplex about x(:,1) using QR decomposition
      % of basis vectors for simplex x

      % get the basis vectors for the current simplex
      [basis basis_lengths] = frame('pb',x);

      % order basis vectors according to length of first
      % n basis vectors
      [sorted_lengths, j] = sort(basis_lengths(1:n));

      % get in descending order
      j = fliplr(j);
      basis = basis(:,j);

      % find QR decomposition of the ordered basis vectors
      [Q R] = qr(basis);

      % setup new length criteria
      d = diag(R);
      davg = sum(abs(d)) / n;
      sign_d = sign(0.5 + sign(d));
      d_new = sign_d .* max(abs(d), davg/10);
      D = diag(d_new);

      % calculate new basis vectors
      basis = Q*D;

      % create new simplex about x(;,1)
      % new simplex is x(:,1) and x(:,1) + basis(:,j) for j=1..n
      for j = 1:n,
         v(:,j+1) = v(:,1) + basis(:,j);
      end

   otherwise
      error('An unknown simplex type has been used')
   end
end
```

H.2.2 FRAME

```
function [F,G] = frame(type,v,h)
% frame(type,v,h)
%
% The output from frame is determined by the type of
% information required.
%
% v is a matrix whose columns represent the vertices of
% the simplex or frame, h is a scale factor and type is either:
%
% l  = returns the n side lengths for the simplex v, scaled by h
% f  = complete the frame for the current simplex where the length
%      of the new frame point is scaled by h
% pb = return the n+1 positive basis vectors and their lengths
% sh = shrink the current frame v towards v(:,1) by scale factor h

% DB 27 Jan 00

if nargin < 2,
   error('Incorrect number of input arguements'),
end

% dimension
n = size(v,1);

if nargin == 2,
      h = 1;
end

% initialise
vectors = zeros(n,n+2);
lengths = zeros(1,n+2);
G = [];

switch type
case 'f'
   % return the completed frame
   vectors(:,1:n+1) = v;
   vectors(:,n+2) = (1 + h)*v(:,1) - h/n * sum(v(:,2:n+1), 2);
   F = vectors;
case 'sh'
   if size(v,2) ~= n+2,
      error('The input frame has the wrong dimensions.');
   end
```

```
      % shrink the current frame towards v(:,1) by h
      vectors(:,1) = v(:,1);
      for j = 2:n+2,
         vectors(:,j) = v(:,1) + h*(v(:,j) - v(:,1));
      end
      if vectors == v,
         % changes are beyond machine precision
         vectors(:,2:n+2) = vectors(:,ones(1,n+1));
      end
   F = vectors;
otherwise
   for j = 1:n,
      vectors(:,j) = (v(:,j+1) - v(:,1)) / h;
      lengths(:,j) = norm(vectors(:,j));
   end
   switch type
   case 'pb'
      vectors(:,n+1) = -sum(vectors(:,1:n), 2) / n;
      lengths(:,n+1) = norm(vectors(:,n+1));
      F = vectors(:,1:n+1);
      G = lengths(:,1:n+1);
   case 'l'
      F = lengths(1:n);
   otherwise
      error('An unknown frame arguement has been used')
   end
end
```

H.3 HH

```
function B = HH(b)

% For a given vector b, B = HH(b) returns a matrix B
% whose columns are orthogonal and sum(B(:,i)) = -b

n = length(b);
t = norm(b)/sqrt(n)*sign(sign(b(1)+.5));
v = b + t;
B = t*(eye(n)-((2/(v'*v))*v)*v');
```

Bibliography

[1] E. M. L. Beale. *Introduction to optimization.* John Wiley & Sons, Chichester, 1988.

[2] M. J. Box, D. Davies, and W. H. Swann. Non-linear optimization techniques. Technical Report 5, Imperial Chemicals Industries Limited, Edinburgh, 1969.

[3] S. H. Brookes. Discussion of random methods for locating surface maxima. *Operations research*, 6:244–251, 1958.

[4] M. D. Buhmann and R. Fletcher. M. J. D. Powell's contribution to numerical mathematics. In M. D. Buhmann and A. Iserles, editors, *Approximation theory and optimization*, pages 1–30. Cambridge University Press, Cambridge, 1997.

[5] D. M. Burley. *Studies in optimization.* International Textbook Company, Leighton Buzzard, 1974.

[6] A. Conn, K. Scheinberg, and P. Toint. On the convergence of derivative-free methods for unconstrained optimization. In M. D. Buhmann and A. Iserles, editors, *Approximation theory and optimization*, pages 83–108. Cambridge University Press, Cambridge, 1997.

[7] I. D. Coope and C. J. Price. On the convergence of grid based methods for unconstrained optimization. Technical Report 180, Department of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand, 1999.

[8] I. D. Coope and C. J. Price. Frame based methods for unconstrained optimization. *Journal of Optimization Theory and Applications*, 107(2), November 2000. To appear.

[9] L. Cooper and D. Steinberg. *Introduction to methods of optimization.* W. B. Saunders Company, Philadelphia, 1970.

[10] R. Fletcher. *Practical methods of optimization.* John Wiley & Sons, New York, 1980.

[11] R. Fletcher. *Practical methods of optimization.* John Wiley & Sons, New York, second edition, 1987.

[12] L. R. Foulds. *Optimization techniques.* Springer-Verlag, New York, 1981.

[13] P. E. Gill, W. Murray, and M. H. Wright. *Practical optimization.* Academic Press, London, 1981.

[14] B. S. Gottfried and J. Weisman. *Introduction to optimization theory.* Prentice Hall, New Jersey, 1973.

[15] D. M. Himmelblau. A uniform evaluation of unconstrained optimization techniques. In F. A. Lootsma, editor, *Numerical methods for non-linear optimization*, pages 69–97. Academic Press, London, 1976.

[16] R. Hooke and T. A. Jeeves. Direct search solution of numerical and statistical problems. *Journal of the Association for Computing Machinery (ACM)*, 8:212–219, 1961.

[17] A. R. Jarvis. Minimising noisy functions. Master's thesis, University of Canterbury, Christchurch, New Zealand, 1997.

[18] C. T. Kelley. Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition. *SIAM Journal on Optimization*, 10(1):43–55, 1999.

[19] J. Kowalik and M. R. Osborne. *Methods for unconstrained optimization problems.* American Elsevier Publishing Company, New York, 1968.

[20] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147, 1998.

[21] K. I. M. McKinnon. Convergence of the Nelder-Mead simplex method to a nonstationary point. *SIAM Journal on Optimization*, 9(1):148–158, 1998.

[22] J. J. Moré. Notes on optimization software. In M. J. D. Powell, editor, *Non-linear optimization 1981*, pages 339–352. Academic Press, New York, 1982.

[23] J. J. Moré, B. S. Garbow, and K. E. Hillstrom. Testing unconstrained optimization software. *ACM Transactions on Mathematical Software*, 7(1):17–41, 1981.

[24] J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

[25] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer, New York, 1999.

[26] J. M. Parkinson and D. Hutchinson. A consideration of non-gradient algorithms for the unconstrained optimization of functions of high dimensionality. In F. A. Lootsma, editor, *Numerical methods for non-linear optimization*, pages 99–113. Academic Press, London, 1976.

[27] J. M. Parkinson and D. Hutchinson. An investigation into the efficiency of variants on the simplex method. In F. A. Lootsma, editor, *Numerical methods for non-linear optimization*, pages 115–135. Academic Press, London, 1976.

[28] M. J. D. Powell. On search direction for minimization algorithms. *Mathematical programming*, 4:193–201, 1973.

[29] M. J. D. Powell. Direct search algorithms for optimization calculations. *Acta Numerica*, 7:287–336, 1998.

[30] W. Spendley, G. R. Hext, and F. R. Himsworth. Sequential application of simplex designs in optimisation and evolutionary operation. *Technometrics*, 4:441–461, 1962.

[31] C. Sutti. Remarks on conjugate direction methods for minimization without derivatives. In L. C. W. Dixon and G. P. Szegö, editors, *Towards global optimization*, pages 290–304. North Holland, 1975.

[32] W. H. Swann. Direct search methods. In W. Murray, editor, *Numerical methods for unconstrained optimization*. Academic Press, London, 1972.

[33] V. Torczon. *Multi-directional search: A direct search algorithm for parallel machines*. PhD thesis, Rice University, Houston, Texas, 1989.

[34] G. R. Walsh. *Methods of optimization*. John Wiley & Sons, Chichester, 1979. Revised reprint.