Nelder-Mead Simplex Modifications For Simulation Optimization

Russell R. Barton, John S. Ivey Jr

Reporter: Yuchen Wan

Date: 8 November, 2021

Optimization problem

$$\text{minimize} \quad E(F(x)), \quad x \in R^n, \qquad (1)$$

$$F(x) = f(x) + \epsilon(x), \qquad (2)$$

F(x): response function
f(x): deterministic function
$\epsilon(x)$: stochastic function, $E(\epsilon(x)) = 0$ for all x

$$\text{minimize} \quad f(x), \quad x \in R^n. \qquad (3)$$

The Nelder-Mead Simplex Algorithm

- Origin: The Spendley, Hext, and Himsworth (SHN) algorithm (Spendley et al., 1962)
- For a function of n parameters
- Identify n+1 equally seperated extreme points in the parameter space >>> define a regular simplex in n dimensions
- Evaluate the function at each extreme point of the simplex
- The algorithm moves toward the optimum by reflecting the extreme point with the worst function value through the centroid (average) of the remaining n extreme points, to identify a new simplex adjacent to the previous one.

The Nelder-Mead Simplex Algorithm

- 1. Initialization: Same as SHN algotithm
- 2. Stopping criterion: standard deviation of F falls below a particular value or until the maximum number of function evaluations is reached

$$S_F \equiv [\sum (F(x_i) - \bar{F})^2/(n+1)]^{1/2}, \quad \bar{F} \equiv \Sigma F(x_i)/(n+1)$$

- 3. Reflect the worst point
- 4a. Accept reflection
- 4b. Attempt expansion
- 4c. Attempt contraction
- 4c'. Shrink

3. Reflect the worst point.

① Identify 3 points

P_{high} → F_{high}

P_{secht} → F_{secht}

P_{low} → F_{low}

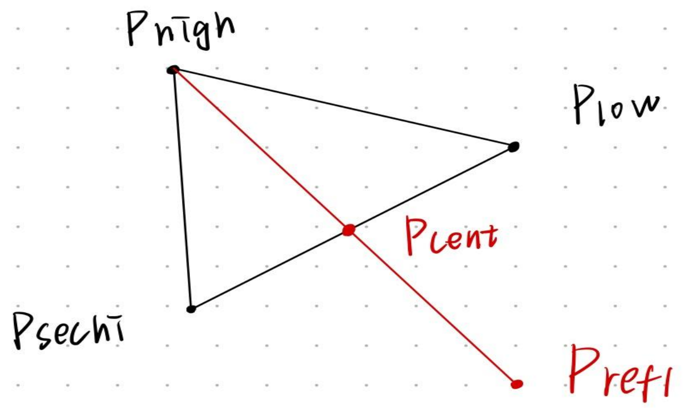
② Find the 4th point P_{cent} F_{cent}

↳ centroid of all vertices other than P_{high}

③ Generate a new vertex P_{refl} by reflecting P_{high} through P_{cent}

$$P_{refl} = (1+\alpha) P_{cent} - \alpha P_{high} \overset{\alpha=1}{=\!=} 2P_{cent} - P_{high}$$

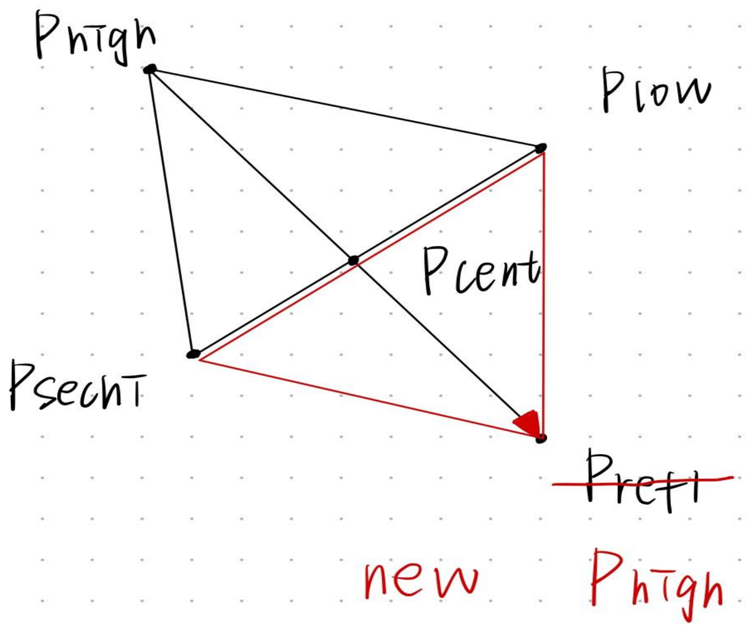
F_{refl}



4. Function Evaluation

a. $F_{low} \le F_{refl} \le F_{secht}$

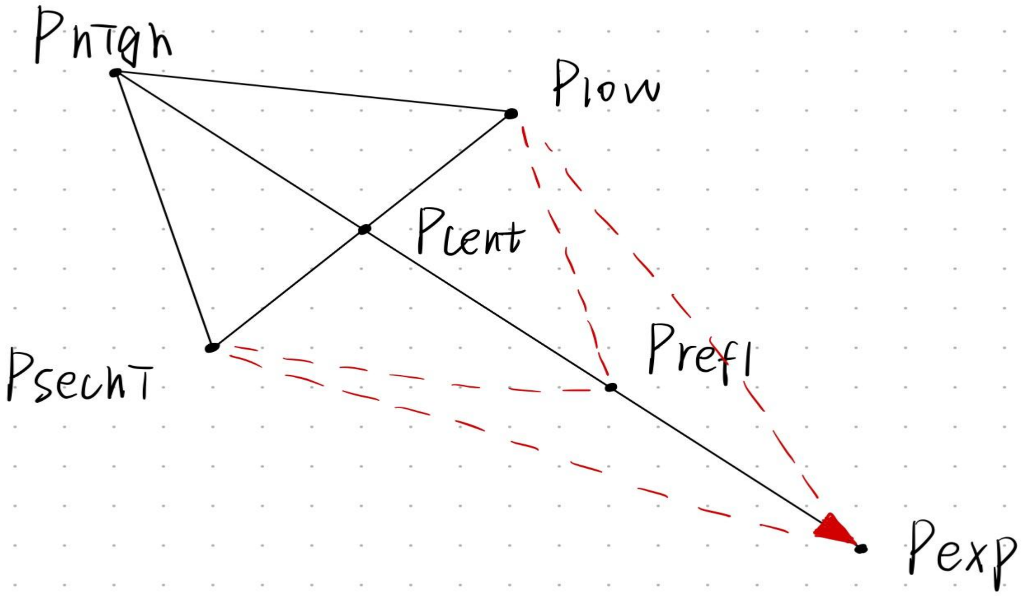
Accept Reflection. P_{refl} replaces P_{high}.



b. $F_{refl} < F_{low}$

Define a new point. $P_{exp} = r P_{refl} + (1-r) P_{cent}$ F_{exp}.

$$\underset{r=2}{=} 2P_{refl} - P_{cent}.$$



① $F_{exp} < F_{low}$ Accept Expansion. P_{exp} replaces P_{high}

② Otherwise Expansion is rejected and P_{refl} replaces P_{high}

C. $F_{refl} > F_{secht}$

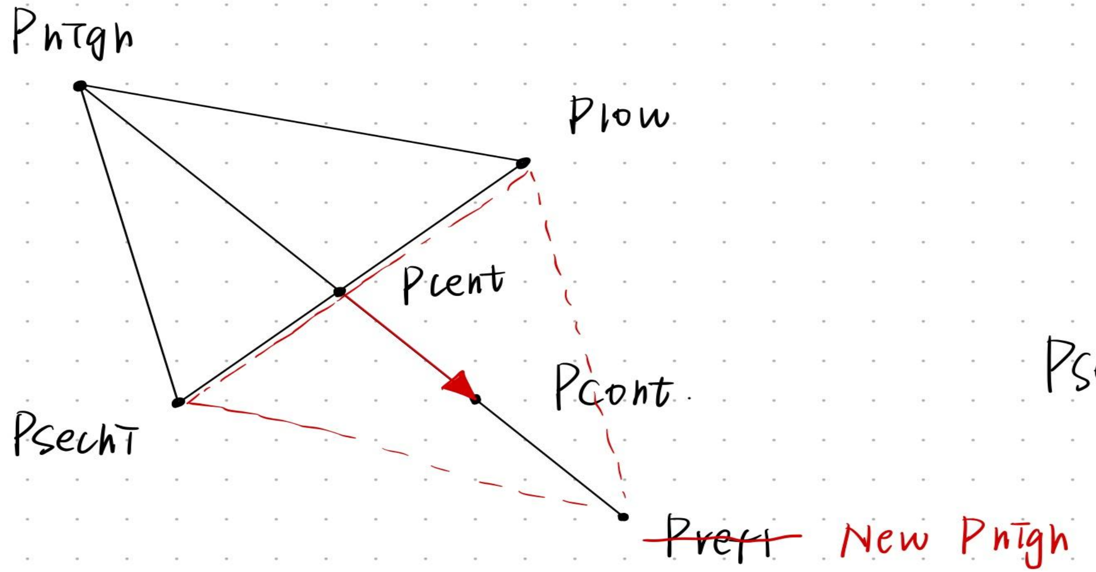
Accept Contraction before attempting contracting or shrinking

Find a new point P_{cont} F_{cont}

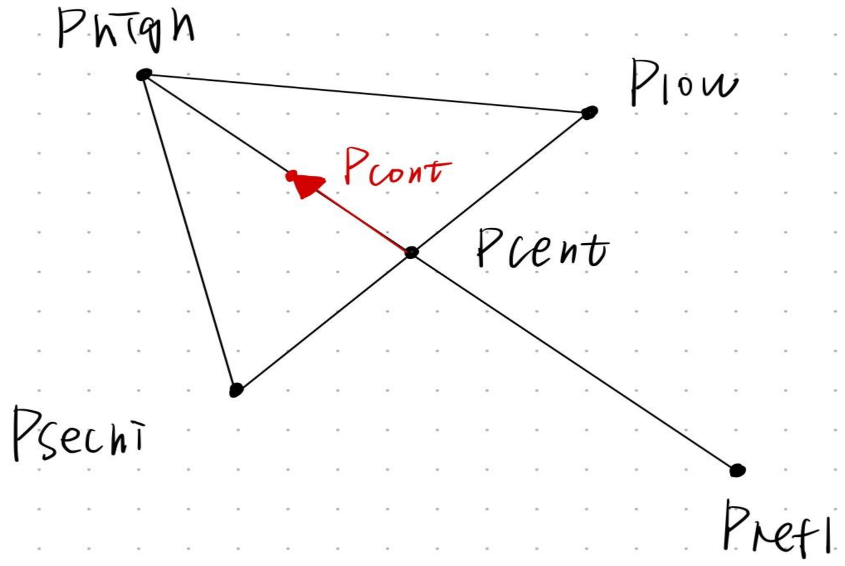
$$P_{cont} = \beta P_{high} + (1-\beta) P_{cent} \overset{\beta=0.5}{=\!=} 0.5 P_{high} + 0.5 P_{cent}$$

$F_{secht} < F_{refl} < F_{high}$

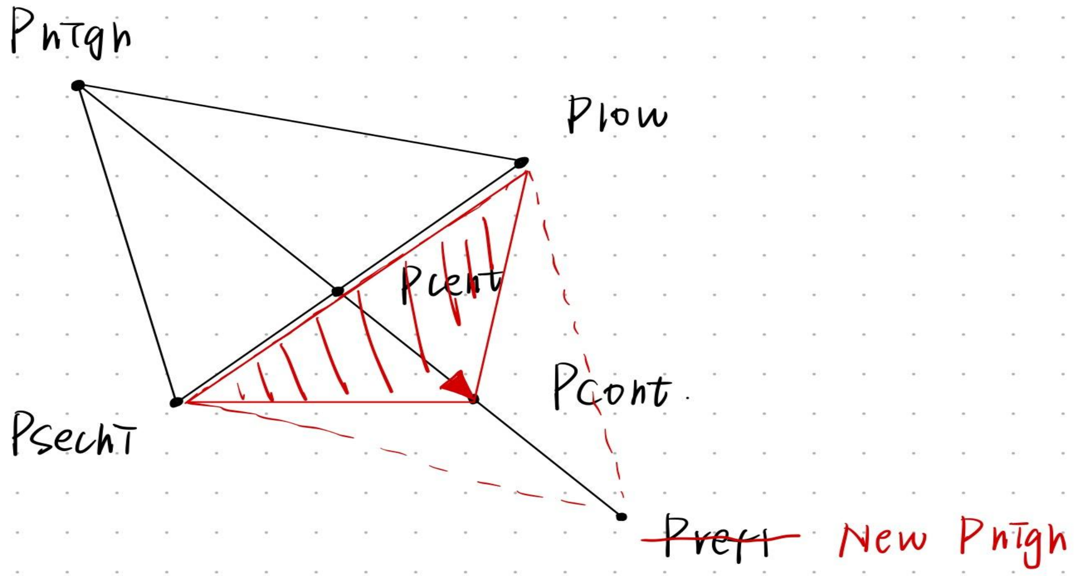
P_{refl} replaces P_{high}

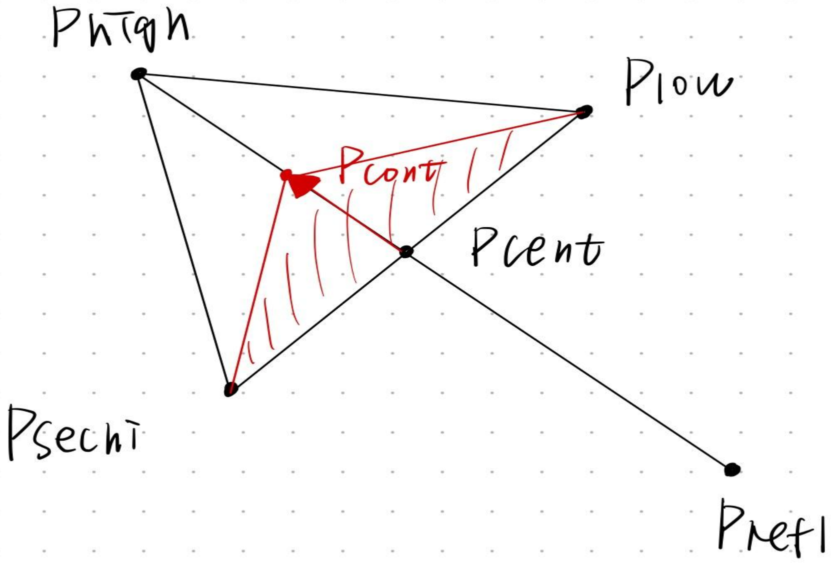


$F_{refl} > F_{high}$



① $F_{cont} \le F_{high}$ Contract



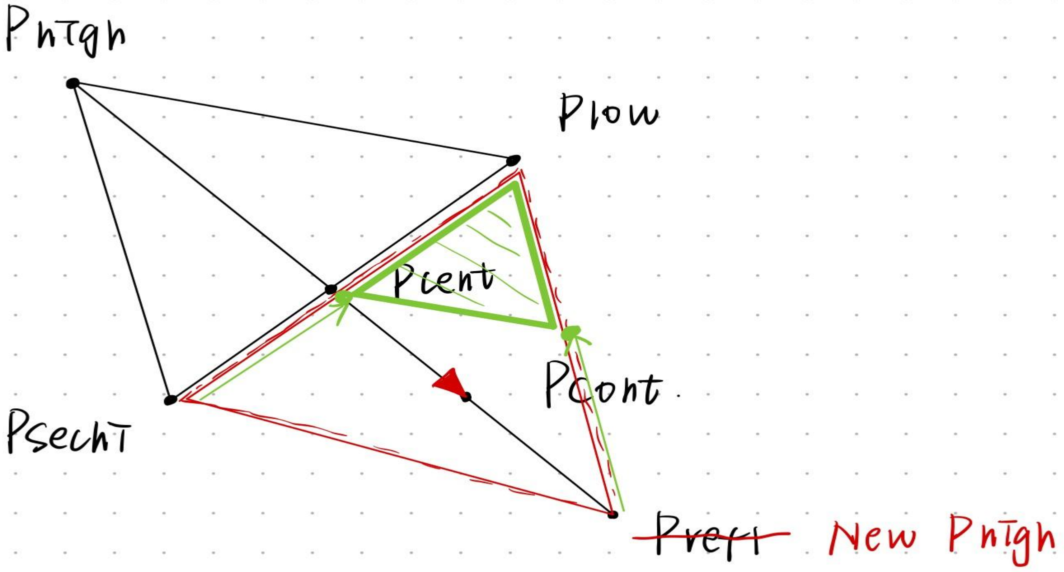


② $F_{cont} > F_{high}$ Shrink (except Plow)

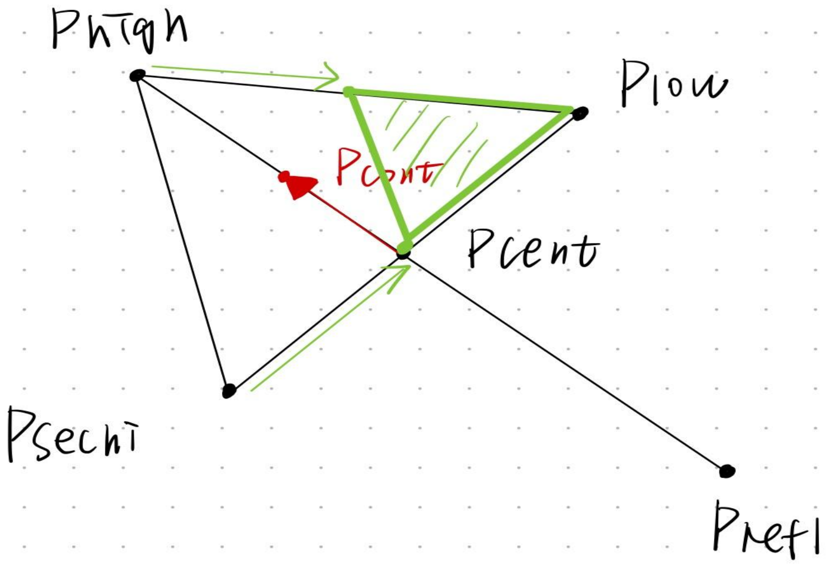
$P_t \longleftarrow \delta P_t + (1-\delta) P_{low}$

// $\delta = 0.5$

$0.5 P_t + 0.5 P_{low}$

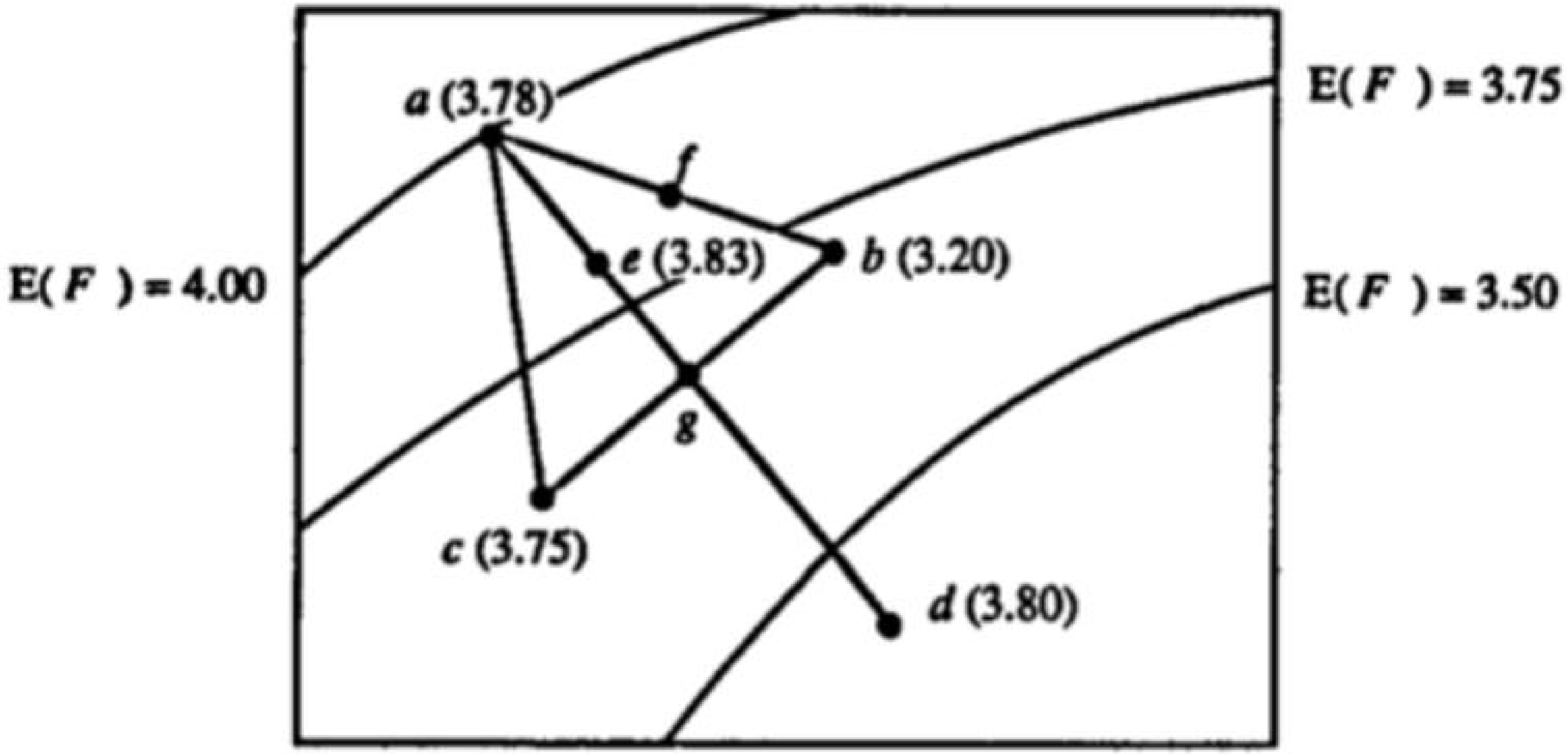


Phigh
Plow
Pcent
Pcont.
Psecht
~~Prefl~~
New Phigh



Phigh
Plow
Pcont
Pcent
Psecht
Prefl

Inappropriate Termination on Stochastic Function



Transition Probability when Noise Dominates

- First-iteration Transition Probability for a Function with Constant Expected Value

	Pr{Event}	
Event	$n = 2$	General n
$R \cap A_{ref}$	0.25	$(n-1)/(n+2)$
E	0.25	$1/(n+2)$
$E \cap A_{exp}$	0.10	$2/[(n+2)(n+3)]$
C	0.50	$2/(n+2)$
$C \cap A_{cont}$	0.30	$2(n+1)/[(n+2)(n+3)]$
$C \cap S$	0.20	$4/[(n+2)(n+3)]$

- *for $n = 2$*

$$\frac{P(E)}{P(C \text{ or } S)} = \frac{1}{5}$$

- *for general n*

$$\frac{P(E)}{P(C \text{ or } S)} = \frac{1}{n-3}$$

- bias toward contraction or shrinkage

Transition Probability when Noise Dominates

- Later iterations

Figure 5 **Moving Average Step Transition Probabilities, Two-Parameter Function**

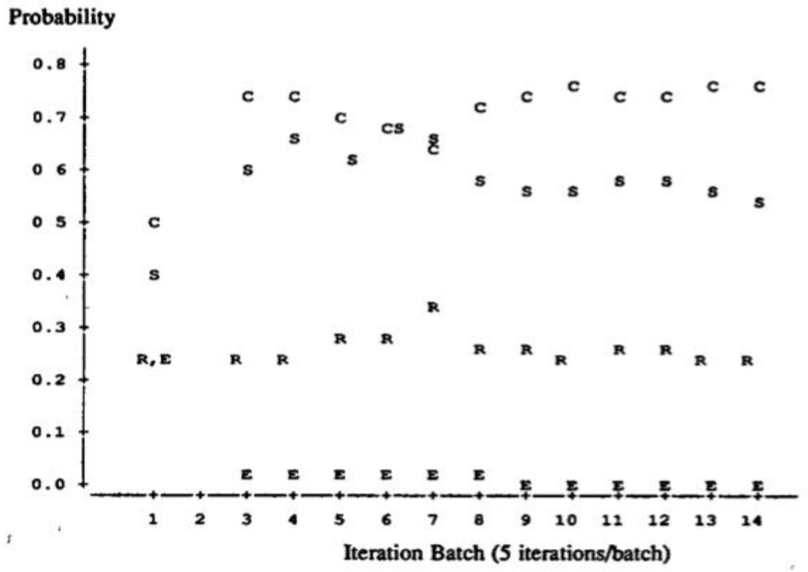
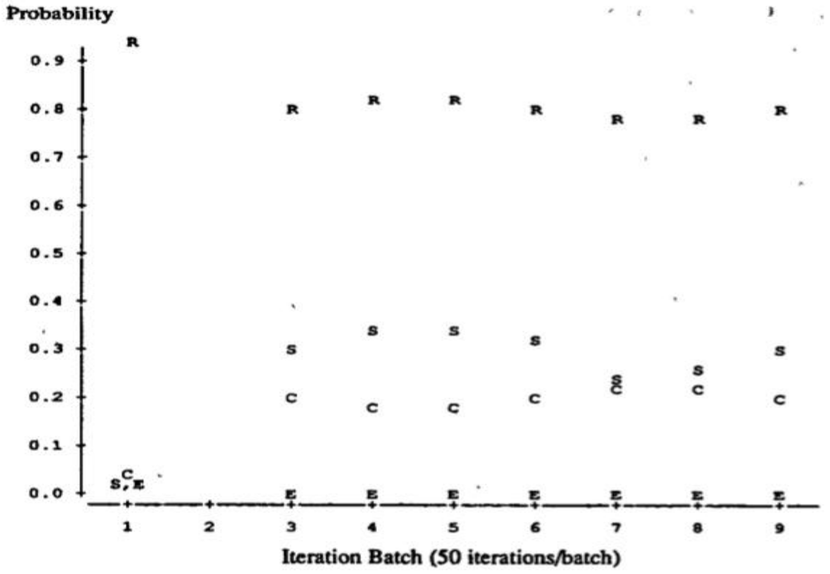


Figure 6 **Moving Average Step Transition Probabilities, Fifty-Parameter Function**



Existing Modifications to Reduce the Error at Termination

- TR: eliminate the use of a shrink step; a failed contraction be followed by a translation of the entire simplex such that the new simplex is centered about the location of the current best point (Ernst, 1968)

- NW: if the contracted point is the worst point of the new simplex, accept it and then reflect the second worst point of the new simplex (King, 1974)

- N3: controlling the retention time of good responses (Walters et al., 1991)

New Modifications

- S9: increase δ, the shrinkage coefficient, from 0.5 to 0.9, reducing the simplex by only 10% rather than 50%; at a cost of additional function evaluations

$$P_i \leftarrow \delta P_i + (1 - \delta)P_{low}.$$

- RS: reevaluate the best point after a shrink step before determining the next reflection, especially important when the simplex becomes small enough that random differences in the ovserved value of ϵ dominats differences in f
- PC: reevaluate P_{refl} and P_{sechi} and contract only then $F'_{refl} < F'_{sechi}$

Computational Experiment

- The expected response at the estimated optimal point obtained by RS+S9 had errors that averaged 15% less than at the original methos's estimated optimal point, at an average cost of three times as many function evaluations.

- Two existing modifications for stochastic response, the (n+3)-rule and the next-to-worst rule, were fount to be inferior to the new modification RS+S9.

Thanks For Listening!