Quasinewton Methods

For all methods for all runs in this part, **stopping criterion:** $\|\mathbf{g}\| < 10^{-6}$ or max iter $= 100$

BFGS:

Run 1: (scale = 0.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -1.0809 & -2.4388 & -0.9058 & -6.6190 & 4.5479 & 2.3531 \\ 5.7014 & 4.0227 & 4.6624 & -3.3264 & -0.7440 & -1.4663 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 13.175773

$\boldsymbol{\ell_k} = 1.008635$
$\boldsymbol{q_k} = 0.077213$

Runtime: 3.2498 seconds

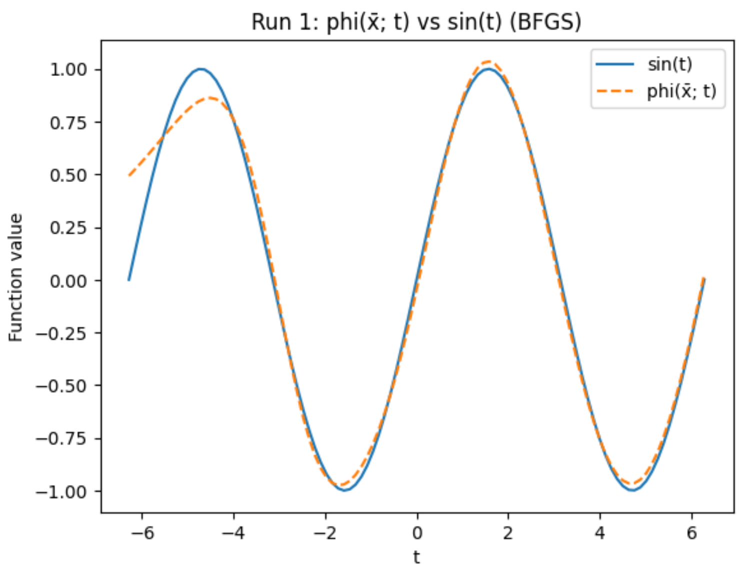

Figure 1: Convergence behavior of Run 1 (scale = 0.5)

Run 2: (scale = 0.75)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -0.9411 & -1.2555 & -0.6493 & 0.9411 & 1.2555 & 0.6493 \\ -0.5261 & -2.2526 & 0.4086 & 0.5261 & 2.2526 & -0.4086 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 5.205394

$\boldsymbol{\ell_k} = 1.000000$
$\boldsymbol{q_k} = 0.192108$

Runtime: 1.9939 seconds

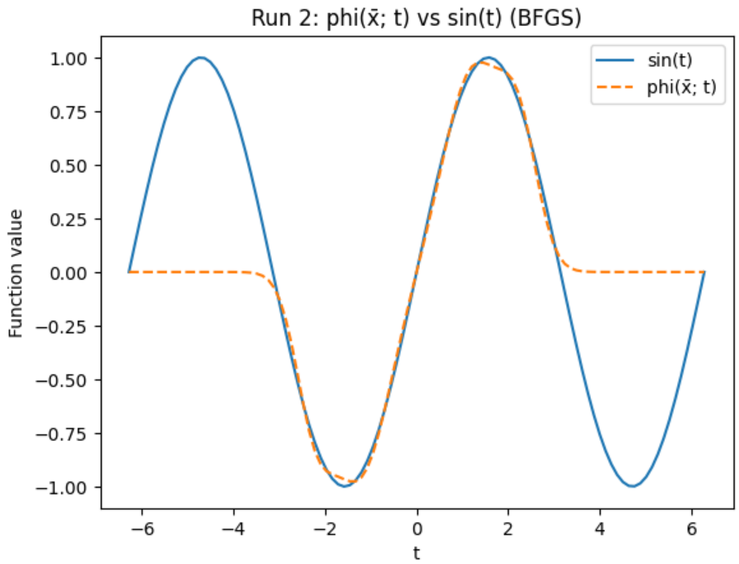

Figure 2: Convergence behavior of Run 2 (scale = 0.75)

Run 3: (scale = 1.0)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -15.3170 & -3.1996 & 1.6094 & -16.3998 & -1.1521 & 1.5963 \\ -5.1207 & -19.8600 & 13.1906 & 27.2052 & -2.2500 & 2.2417 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 43.367214

$\boldsymbol{\ell_k} = 0.998854$
$\boldsymbol{q_k} = 0.023006$

Runtime: 3.1850 seconds

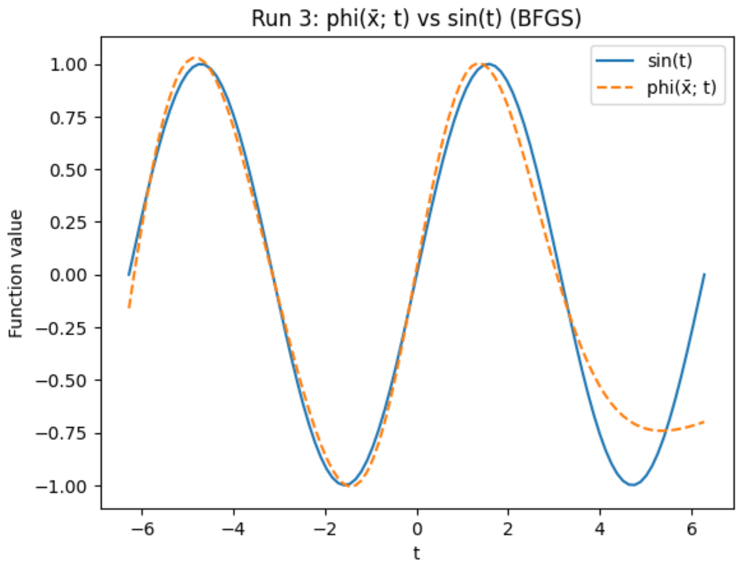

Figure 3: Convergence behavior of Run 3 (scale = 1.0)

Run 4: (scale = 1.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.8902 & -5.0129 & -0.6499 & 8.7286 & 1.4997 & 1.7231 \\ -7.6670 & 1.4735 & -2.1226 & 0.7368 & -3.7933 & 0.6935 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 13.790752

$\boldsymbol{\ell_k} = 1.008206$
$\boldsymbol{q_k} = 0.073707$

Runtime: 3.1941 seconds

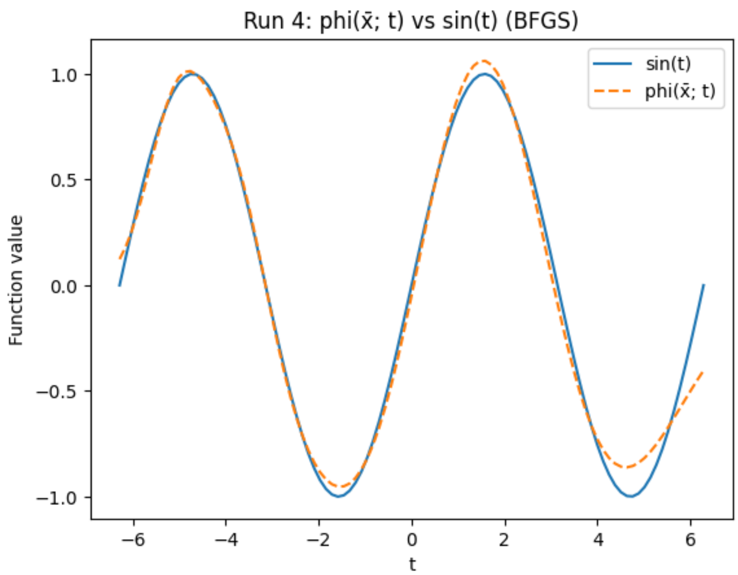

Figure 4: Convergence behavior of Run 4 (scale = 1.5)

Run 5: (scale = 2.0)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 9.1294 & -4.5336 & 2.2928 & -10.5100 & -2.6251 & 4.2798 \\ 10.7937 & 2.5347 & -3.4781 & -7.5699 & 4.3296 & 2.0809 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 21.915950

$\boldsymbol{\ell_k} = 1.011843$
$\boldsymbol{q_k} = 0.046716$

Runtime: 3.1972 seconds

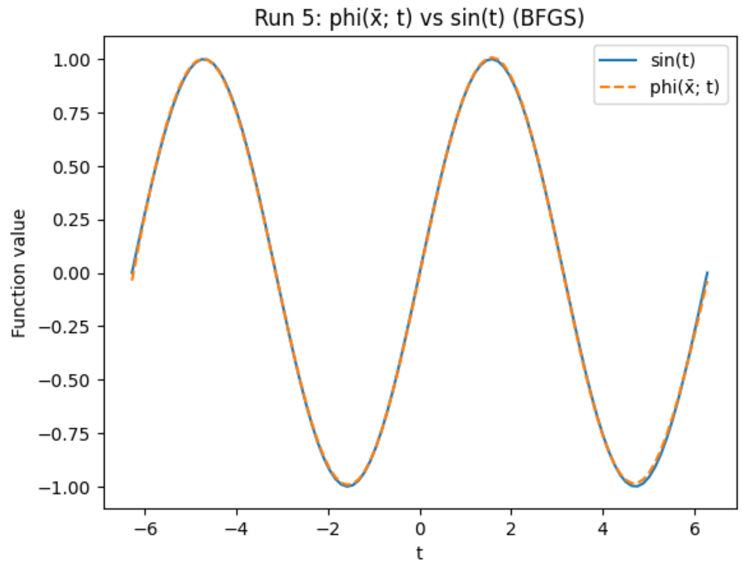

Figure 5: Convergence behavior of Run 5 (scale = 2.0)

Comments: The BFGS method exhibited high variability in performance, with a strong dependence on the initial scaling. Run 2 was the most stable, converging to within a distance of 5.21 from $\mathbf{x}^*$, with $\ell_k = 1.000$ and $q_k = 0.192$. This suggests near-linear convergence, though not superlinear as expected from BFGS in ideal conditions.

It fits so well to sin(t) likely because it is biased as $\bar{\mathbf{x}}$ was estimated using BFGS for methods 3-5.

DFP:

Run 1: (scale = 0.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -1.0781 & -1.5253 & 0.7567 & 0.5476 & 2.2466 & 0.4112 \\ 1.4494 & 0.6951 & 0.0009 & 0.9409 & 1.2440 & -0.6370 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 4.374717

$\boldsymbol{\ell_k}$ = 1.000000
$\boldsymbol{q_k}$ = 0.228586

Runtime: 2.0782 seconds

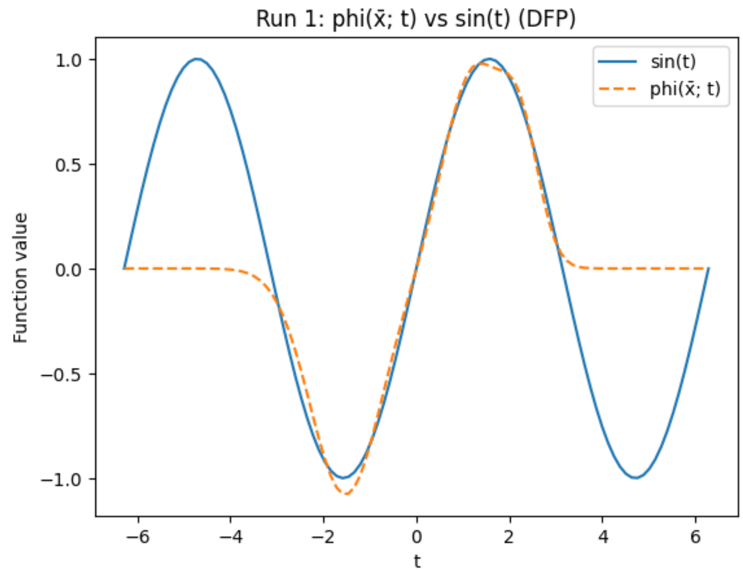

Figure 6: Convergence behavior of Run 1 (scale = 0.5)

Run 2: (scale = 0.75)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -1.1607 & 4.7025 & -0.8191 & -1.6038 & 3.0196 & 0.9323 \\ -1.0901 & -1.4623 & 0.7985 & 1.9986 & 2.5155 & -1.2523 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 7.218019

$\boldsymbol{\ell_k} = 1.000643$
$\boldsymbol{q_k} = 0.138720$

Runtime: 3.1931 seconds

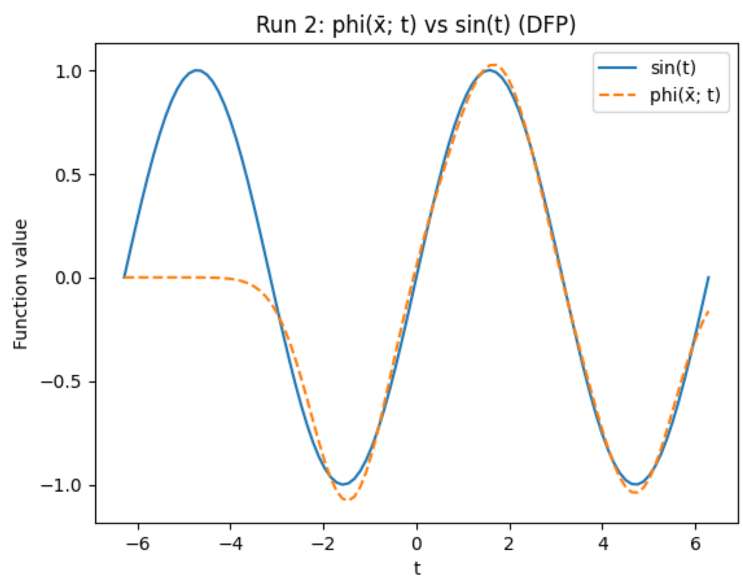

Figure 7: Convergence behavior of Run 2 (scale = 0.75)

Run 3: (scale = 1.0)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -1.0255 & -1.6297 & 0.7017 & -0.0642 & -0.1665 & 0.4478 \\ 1.0808 & 1.4850 & -0.7929 & -0.1356 & -0.6414 & -0.7589 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 4.093776

$\boldsymbol{\ell_k} = 0.996106$
$\boldsymbol{q_k} = 0.242375$

Runtime: 0.5208 seconds

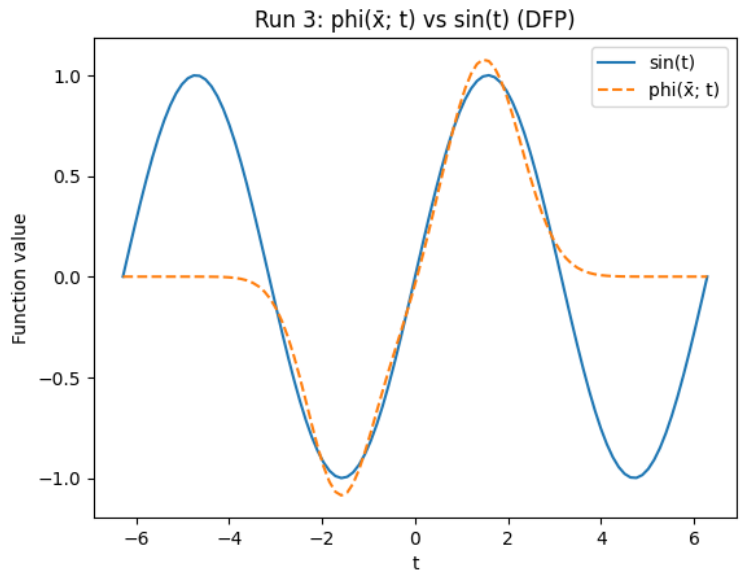

Figure 8: Convergence behavior of Run 3 (scale = 1.0)

Run 4: (scale = 1.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.4442 & 0.3916 & 0.0076 & -10.4362 & -0.3196 & 2.4098 \\ 8.2042 & 0.7796 & -1.9122 & 4.4001 & -2.4851 & 2.4265 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 14.920831

$\boldsymbol{\ell_k} = 1.006463$
$\boldsymbol{q_k} = 0.067890$

Runtime: 3.2800 seconds

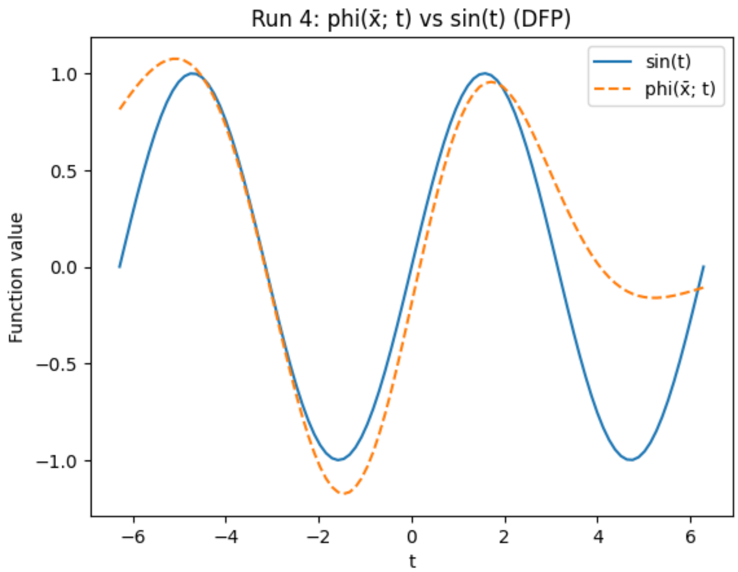

Figure 9: Convergence behavior of Run 4 (scale = 1.5)

Run 5: (scale = 2.0)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -6.9806 & 1.6608 & -2.5250 & 0.2264 & 3.8056 & -1.1233 \\ 1.0060 & 4.7691 & 30.0997 & 6.9964 & 1.5981 & -1.7032 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 31.871335

$\boldsymbol{\ell_k}$ = 1.003897
$\boldsymbol{q_k}$ = 0.031621

Runtime: 3.1991 seconds

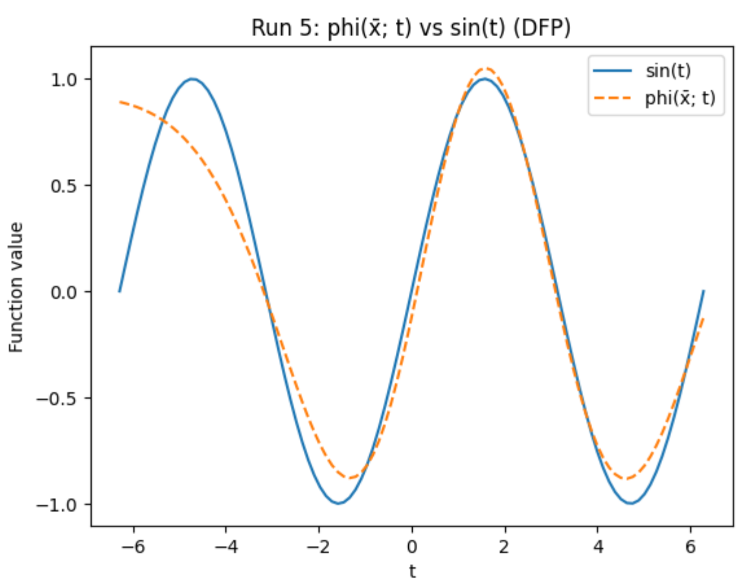

Figure 10: Convergence behavior of Run 5 (scale = 2.0)

Comments: DFP displayed highly erratic performance across the tested scales. Runs 1 and 3 achieved relatively moderate distances to $\mathbf{x}^*$ (4.37 and 4.09, respectively), but exhibited ℓ_k values extremely close to 1, and q_k values around 0.23–0.24, indicative of at best *linear convergence*. In both cases, curvature updates were occasionally skipped due to near-zero inner products ($y^T s \approx 0$), undermining Hessian updates and contributing to stagnation. However, the resulting $\varphi(\bar{\mathbf{x}}; t)$ in these runs still visually approximates $\sin(t)$ reasonably well, suggesting that even imperfect convergence can yield functionally useful results.

SR1 Results

Run 1 (scale = 0.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 1.0918 & 1.5568 & 0.7157 & -1.0687 & -1.6807 & 0.6749 \\ -0.3680 & -0.6154 & -0.2197 & -0.0035 & -2.3843 & 0.6564 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 4.091887

$\ell_k = 1.002692$
$q_k = 0.245704$

Runtime: 2.0340 seconds

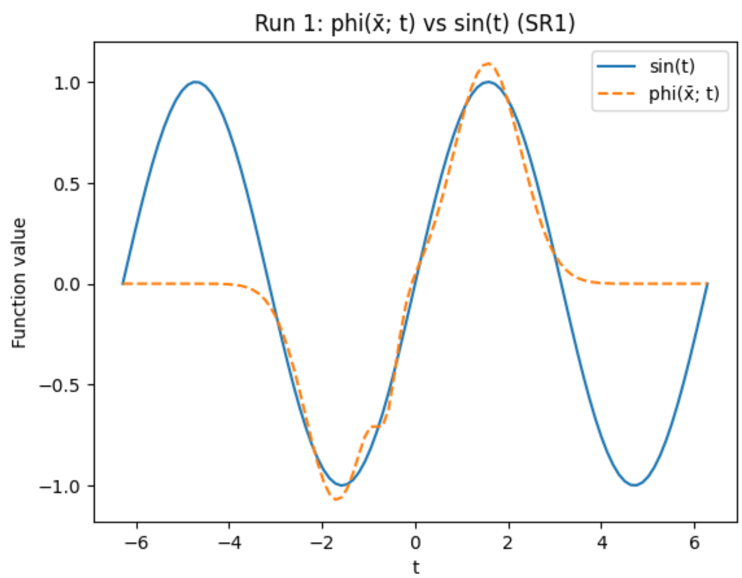

Figure 11: Convergence behavior of Run 1 (scale = 0.5)

Run 2 (scale = 0.75)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} -0.7811 & -0.9754 & -0.6688 & 0.9402 & 1.7955 & 0.6204 \\ 0.4541 & 0.6732 & -0.6821 & -0.7157 & -2.0541 & -0.5081 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 4.514975

$\ell_k = 1.026155$
$q_k = 0.233223$

Runtime: 2.0378 seconds

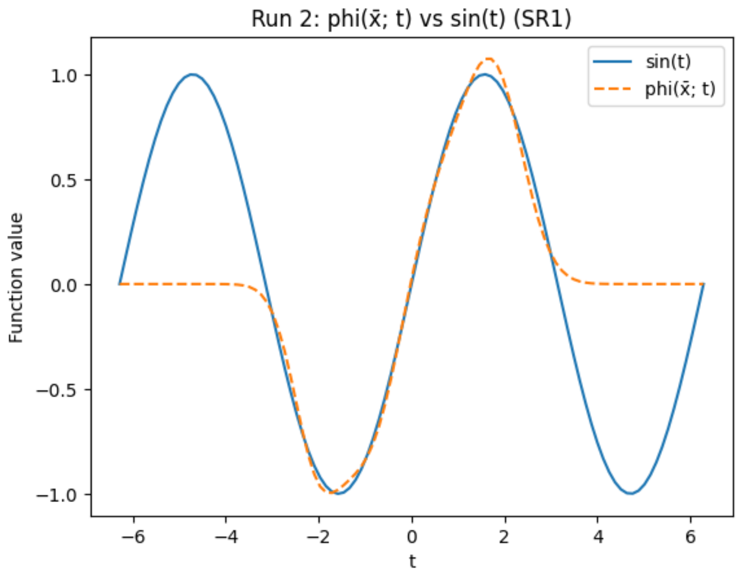

Figure 12: Convergence behavior of Run 2 (scale = 0.75)

Run 3 (scale = 1.0)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.5261 & 2.2526 & -0.4086 & -0.5261 & -2.2526 & -0.4085 \\ 0.9411 & 1.2555 & -0.6493 & -0.9411 & -1.2555 & -0.6493 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 4.457463

$\ell_k = 1.000000$
$q_k = 0.224343$

Runtime: 2.0291 seconds

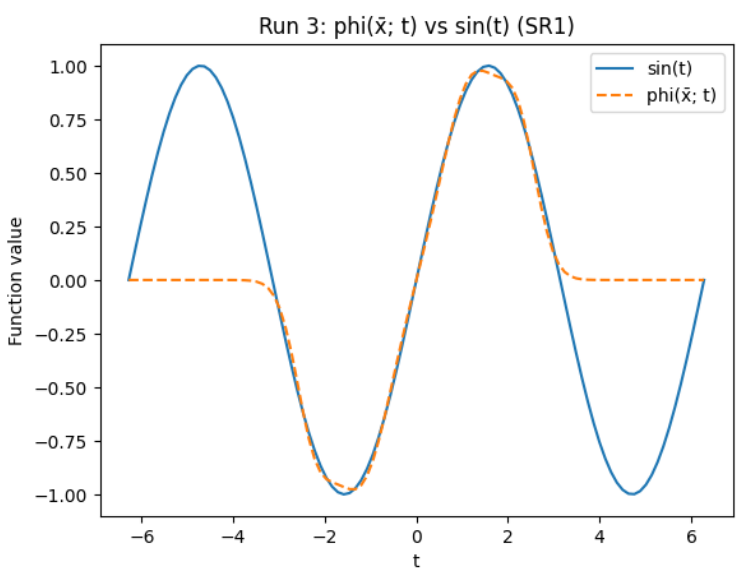

Figure 13: Convergence behavior of Run 3 (scale = 1.0)

Run 4 (scale = 1.5)

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.7247 & 0.9862 & 0.0200 & -0.4729 & -6.7756 & 0.5235 \\ -3.6530 & -1.7786 & 1.3066 & 2.4309 & -1.9672 & 2.2817 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 8.936805

$\ell_k = 0.942776$
$q_k = 0.099457$

Runtime: 2.0902 seconds

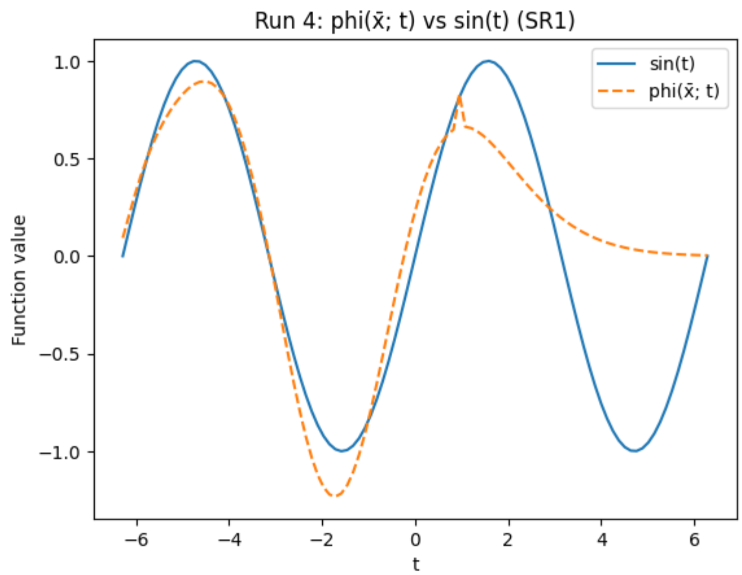

Figure 14: Convergence behavior of Run 4 (scale = 1.5)

Run 5 (scale = 2.0)

Stopping criterion: $\|g\| < 1 \times 10^{-6}$ or max_iter = 100

Final iterate:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.7866 & -0.7551 & -0.0173 & -1.4401 & -1.5910 & 0.8463 \\ 0.0306 & -2.6672 & 8.6018 & 0.5695 & -22.2815 & 18.4644 \end{bmatrix}$$

Distance to $\mathbf{x}^*$: 29.854298

$\ell_k = 1.034099$
$q_k = 0.035819$

Runtime: 2.0566 seconds

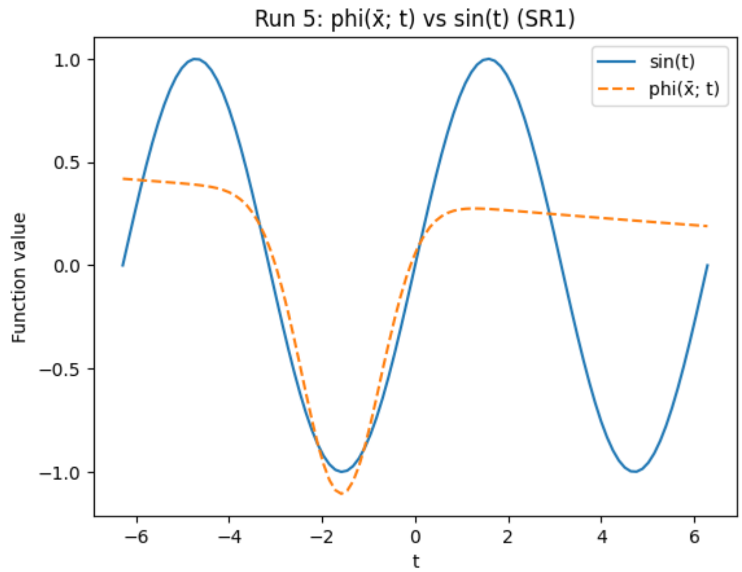

Figure 15: Convergence behavior of Run 5 (scale = 2.0)

Comments: SR1 generally performed with more stability than DFP, although it too showed signs of divergence at higher scales. Runs 1 through 3 consistently yielded distances in the range of 4.09 to 4.51, with $\ell_k \approx 1.00$ and q_k around 0.22–0.25. These values imply *steady but non-accelerating convergence*, suggestive of near-linear behavior. However, despite this steady convergence, SR1's final approximations of $\sin(t)$ were less accurate compared to those achieved by DFP and BFGS.