APPENDIX

1 The Number of Mathematical Symbol Spaces

Theorem 1. *Suppose that in a population, in order to ensure that the similarity between two individuals is less than or equal to ε, the lower bound of k is $\frac{-\ln(\varepsilon)}{\ln(\max(d(M,N),1))}$.*

Proof. According to Equation (**??**) and Theorem 2. We have the Equation (1) as follows.

$$(\frac{1}{\max(d(M,N),1)})^k \leq \varepsilon \tag{1}$$

Take logarithms ln on both sides, we get

$$\ln((\frac{1}{\max(d(M,N),1)})^k) \leq \ln(\varepsilon) \tag{2}$$

Simplify, we get

$$-k\ln(\max(d(M,N),1)) \leq \ln(\varepsilon) \tag{3}$$

Then, we have

$$k \geq \frac{-\ln(\varepsilon)}{\ln(\max(d(M,N),1))} \tag{4}$$

So, finally, we can get the lower bound of the number of mathematical symbol spaces k is $\frac{-\ln(\varepsilon)}{\ln(\max(d(M,N),1))}$.

Theorem 2. *Assuming that the probability $P(\omega_i)$ of subspace ω_i being selected satisfies the Zipf distribution and that $P(\omega_i)$ cannot be less than λ to ensure that every subspace has a chance of being selected, then the upper bound of the number k of subspaces is $e^{\frac{1}{\lambda}-\gamma}$.*

Proof. According to the Zipf distribution probability distribution function, the formula for the number of subspaces k can be obtained as follows:

$$P(X,\alpha,k) = \frac{1}{X^\alpha \sum_{i=1}^{k} \frac{1}{i^\alpha}} \tag{5}$$

Since we want to ensure that each subspace is selected with probability at least λ, we can obtain:

$$\frac{1}{X^\alpha \sum_{i=1}^{k} \frac{1}{i^\alpha}} \geq \lambda \tag{6}$$

Because in the SFEV-GP algorithm, the probability of subspace selection is not uniformly distributed, so take α=1 here and substitute it into equation (6), we can obtain:

$$\frac{1}{X\sum_{i=1}^{k}\frac{1}{i}} \geq \lambda \tag{7}$$

Because the $\alpha = 1$, subspace selected probability is inversely proportional to the X, $P(X = 1, \alpha, k)$ or $P(X, \alpha, k)$, and $P(X, \alpha, k)$ or greater λ, so $P(X = 1, \alpha, k)$ or greater λ, can be obtained:

$$\frac{1}{\sum_{i=1}^{k}\frac{1}{i}} \geq \lambda \tag{8}$$

where $\sum_{i=1}^{k}\frac{1}{i}$ is harmonic progression, its approximation is $\ln(k)$, so we can obtain:

$$\frac{1}{\ln k + \gamma} \geq \lambda \tag{9}$$

By simplifying equation (9), the supremum of k can be obtained as follows:

$$k \leq e^{\frac{1}{\lambda}-\gamma} \tag{10}$$

Therefore, the upper bound of the number of subspaces k is $e^{\frac{1}{\lambda}-\gamma}$.

Based on the above analysis, it can be concluded that the reasonable range of the number of subspaces k is $\left[\frac{-\ln(\varepsilon)}{\ln\left(\max(d(M,N),1)\right)}, e^{\frac{1}{\lambda}-\gamma}\right]$.

Table 1: Classical Symbolic Regression Benchmarks(SRB).

FileNumber	FileName	Object Function	Data Set
F1	Keijzer-1	$0.3x * \sin(2\pi x)$	E[-1,1,0.1]
F2	Keijzer-2	$0.3x * \sin(2\pi x)$	E[-2,2,0.1]
F3	Keijzer-3	$0.3x * \sin(2\pi x)$	E[-3,3,0.1]
F4	Keijzer-4	$x^3 * e^{(-x)} \cos(x) \sin(x)(\sin^2(x) \cos(x) - 1)$	E[0,10,0.05]
F5	Keijzer-7	$\ln(x)$	E[1,100,1]
F6	Keijzer-8	$\sqrt{x}$	E[0,100,1]
F7	Keijzer-10	x^y	U[0,1,100]
F8	Keijzer-11	$x * y + \sin((x-1)(y-1))$	U[-3,3,20]
F9	Keijzer-12	$x^4 - x^3 + \frac{y^2}{2} - y$	U[-3,3,20]
F10	Keijzer-13	$6 \sin(x) * \cos(y)$	U[-3,3,20]
F11	Keijzer-14	$\frac{8}{2+x*x+y*y}$	U[-3,3,20]
F12	Keijzer-15	$\frac{x^3}{5} + \frac{y^3}{2} - y - x$	U[-3,3,20]
F13	Keijzer-5	$\frac{30x*z}{(x-10)*y^2}$	x, z:U[-1,1,1000], y:U[1,2,1000]
F14	Korns-1	$1.57 + 24.3 * v$	U[-50, 50, 200]
F15	Korns-4	$-2.3 + 0.13 * \sin(z)$	U[-50, 50, 200]
F16	Korns-5	$3 + 2.13 * \ln(w)$	U[-50, 50, 200]
F17	Korns-6	$1.3 + 0.13\sqrt{x}$	U[-50, 50, 200]
F18	Korns-11	$6.87 + 11 \cos(7.23x^3)$	U[-50, 50, 200]
F19	Korns-12	$2 - 2.1 \cos(9.8x) * \sin(1.3w)$	U[-50,50,200]
F20	Pagie-1	$\frac{1}{1+x^{-4}} + \frac{1}{1+y^{-4}}$	E[-5,5,0.4]

Table 1: Classical Symbolic Regression Benchmarks(SRB). (Continued)

FileNumber	FileName	Object Function	Data Set
F21	Korns-3	$-5.41 + 4.9\frac{(v-x+\frac{y}{w})}{3w}$	U[-50,50,200]
F22	Korns-14	$22 - 4.2(\cos(x) - \tan(y)) * \frac{\tan(z)}{\sin(v)}$	U[0,5,200]
F23	Korns-10	$0.81 + 24.3 * \frac{2y+3(z)^2}{4(v)^3+5(w)^4}$	U[-5,5,200]
F24	Nguyen-4	$x^6 + x^5 + x^4 + x^3 + x^2 + x$	U[-1,1,200]
F25	Nguyen-3	$x^5 + x^4 + x^3 + x^2 + x$	U[-1,1,200]
F26	Koza-1	$x^4 + x^3 + x^2 + x$	U[-1,1,200]
F27	Nguyen-1	$x^3 + x^2 + x$	U[-1,1,200]
F28	Koza-3	$x^6 - 2x^4 + x^2$	U[-1,1,200]
F29	Koza-2	$x^5 - 2x^3 + x$	U[-1,1,200]
F30	Nguyen-5	$\sin(x^2)\cos(x) - 1$	U[-1,1,200]
F31	Nguyen-6	$\sin(x) + \sin(x + x^2)$	U[-1,1,200]
F32	Nguyen-7	$\ln(x+1) + \ln(x^2+1)$	U[0,2,200]
F33	Nguyen-8	$\sqrt{x}$	U[0,4,200]
F34	Nguyen-9	$\sin(x) + \sin(y^2)$	U[-1,1,200]
F35	Nguyen-10	$2\sin(x) * \cos(y)$	U[-1,1,200]
F36	Vladislavleva-2	$e^{-x} * x^3(\cos(x) * \sin(x))(\cos(x) * \sin^2(x) - 1)$	E[0.05,10,0.1]
F37	Vladislavleva-1	$\frac{e^{-(x-1))^2}}{1.2+(y-2.5)^2}$	U[0.3,4,100]
F38	Vladislavleva-3	$e^{-x} * x^3(\cos(x)\sin(x))(\cos(x) * \sin^2(x) - 1)(y-5)$	x:E[0.05,10,0.1]
			y:E[0.05,10.05,2]
F39	Vladislavleva-6	$6\sin(x) * \cos(y)$	U[0.1,5.9,30]
F40	Vladislavleva-5	$30\frac{(x-1)(z-1)}{y^2(x-10)}$	x, z:U[0.05,2,300]
			y:U[1,2,300]
F41	Vladislavleva-4	$\frac{10}{5+(x-3)^2+(y-3)^2+(v-3)^2+(w-3)^2+(q-3)^2}$	U[0.05,6.05,1024]

Table 2: Feynman Symbolic Regression Benchmarks(FSRB).

FileNumber	FileName	Object Function	Data Set
F42	I.12.1	F = μ $\mathrm{N_n}$	μ, N_n : $U_{\log}(10^{-2}, 10^{0}, 8000)$
F43	I.12.5	$F = q_2 E$	$q_2 : U_{\log}(10^{-3}, 10^{-1}, 8000)$ $E : U_{\log}(10^{1}, 10^{3}, 8000)$
F44	I.14.3	$U = mgz$	$m, z : U_{\log}(10^{-2}, 10^{0}, 8000)$ $g : 9.807 \times 10^{0}$
F45	I.14.4	U = $\frac{k_{spring} x^2}{2}$	k_{spring}: $U_{\log}(10^{2}, 10^{4}, 8000)$ $x : U_{\log}(10^{-2}, 10^{0}, 8000)$
F46	I.18.12	$\tau = rF \sin\theta$	$r, F : U_{\log}(10^{-1}, 10^{1}, 8000)$ $\theta : U(0, 2\pi)$
F47	I.18.16	$L = mrv \sin\theta$	$m, r, v : U_{\log}(10^{-1}, 10^{1}, 8000)$ $\theta : U(0, 2\pi)$
F48	I.25.13	$V = \frac{q}{C}$	$q, C : U_{\log}(10^{-5}, 10^{-3}, 8000)$
F49	I.27.6	$f = \frac{1}{\frac{1}{d_1} + \frac{n}{d_2}}$	$d_1, d_2 : U_{\log}(10^{-3}, 10^{-1}, 8000)$ $n : U_{\log}(10^{-1}, 10^{1}, 8000)$
F50	I.30.5	$d = \frac{\lambda}{n \sin\theta}$	$\lambda : U_{\log}(10^{-11}, 10^{-9}, 8000)$ $n : U_{\log}(10^{0}, 10^{2}, 8000)$ $\theta : U(0, 2\pi)$
F51	I.43.16	v = μ q $\frac{V}{d}$	μ: $U_{\log}(10^{-6}, 10^{-4}, 8000)$ q: $U_{\log}(10^{-11}, 10^{-9}, 8000)$
			V: $U_{\log}(10^{-1}, 10^{1}, 8000)$ d: $U_{\log}(10^{-3}, 10^{-1}, 8000)$
F52	I.47.23	$c = \sqrt{\frac{\gamma P}{\rho}}$	$\gamma, \rho : U(1, 2, 8000)$ $P : U_{\log}(0.5 \times 10^{-5}, 1.5 \times 10^{-5}, 8000)$
F53	II.13.17	$B = \frac{1}{4\pi\epsilon c^2} \frac{2I}{r}$	$\epsilon : 8.854 \times 10^{-12}$ $c : 2.998 \times 10^{8}$ $I, r : U_{\log}(10^{-3}, 10^{-1}, 8000)$
F54	II.15.4	$U = -\mu B \cos\theta$	$\mu : U_{\log}(10^{-25}, 10^{-23}, 8000)$ $B : U_{\log}(10^{-3}, 10^{-1}, 8000)$ $\theta : U(0, 2\pi)$
F55	II.15.5	$U = -pE \cos\theta$	$p : U_{\log}(10^{-22}, 10^{-20}, 8000)$ $E : U_{\log}(10^{1}, 10^{3}, 8000)$ $\theta : U(0, 2\pi)$
F56	II.27.16	$S = \epsilon c E^2$	$\epsilon : 8.854 \times 10^{-12}$ $c : 2.998 \times 10^{8}$ $E : U_{\log}(10^{-1}, 10^{1}, 8000)$
F57	II.27.18	$u = \epsilon E^2$	$\epsilon : 8.854 \times 10^{-12}$ $E : U_{\log}(10^{-1}, 10^{1}, 8000)$
F58	II.34.29b	U = 2 π g μ B $\frac{J_z}{h}$	μ: 9.2740100783 $\times$ 10^{-24} B: $U_{\log}(10^{-3}, 10^{-1}, 8000)$
			g: U(-1, 1) J_z: $U_{\log}(10^{-26}, 10^{-22}, 8000)$ h: 6.626 $\times$ 10^{-34}
F59	II.38.14	$\mu = \frac{Y}{2(1+\sigma)}$	$Y : U_{\log}(10^{-1}, 10^{1}, 8000)$ $\sigma : U_{\log}(10^{-2}, 10^{0}, 8000)$
F60	II.38.3	F = Y A $\frac{\Delta l}{l}$	Y: $U_{\log}(10^{-1}, 10^{1}, 8000)$ A: $U_{\log}(10^{-4}, 10^{-2}, 8000)$
			Δ l: $U_{\log}(10^{-3}, 10^{-1}, 8000)$ l: $U_{\log}(10^{-2}, 10^{0}, 8000)$
F61	II.8.31	$u = \frac{\epsilon E^2}{2}$	$\epsilon : 8.854 \times 10^{-12}$ $E : U_{\log}(10^{1}, 10^{3}, 8000)$
F62	III.12.43	$J = \frac{mh}{2\pi}$	$m : U_{\log}(10^{0}, 10^{2}, 8000)$ $h : 6.626 \times 10^{-34}$
F63	I.10.7	$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$	$m_0 : U_{\log}(10^{-1}, 10^{1}, 8000)$ $v : U_{\log}(10^{5}, 10^{8}, 8000)$ $c : 2.998 \times 10^{8}$
F64	I.11.19	$A = x_1\ y_1 + x_2\ y_2 + x_3\ y_3$	x_1, y_1, x_2, y_2, x_3, y_3: $U_{\log}(10^{-1}, 10^{1}, 8000)$
F65	I.12.11	$F = q(E + Bvsin(\theta))$	$q, E, B, v : U_{\log}(10^{-1}, 10^{1}, 8000)$ $\theta : U(0, 2\pi)$
F66	I.13.12	U = G m_1 m_2 ($\frac{1}{r_2}$ - $\frac{1}{r_1}$)	G: 6.674 $\times$ 10^{-11} m_1, m_2, r_2, r_1: $U_{\log}(10^{-2}, 10^{0}, 8000)$

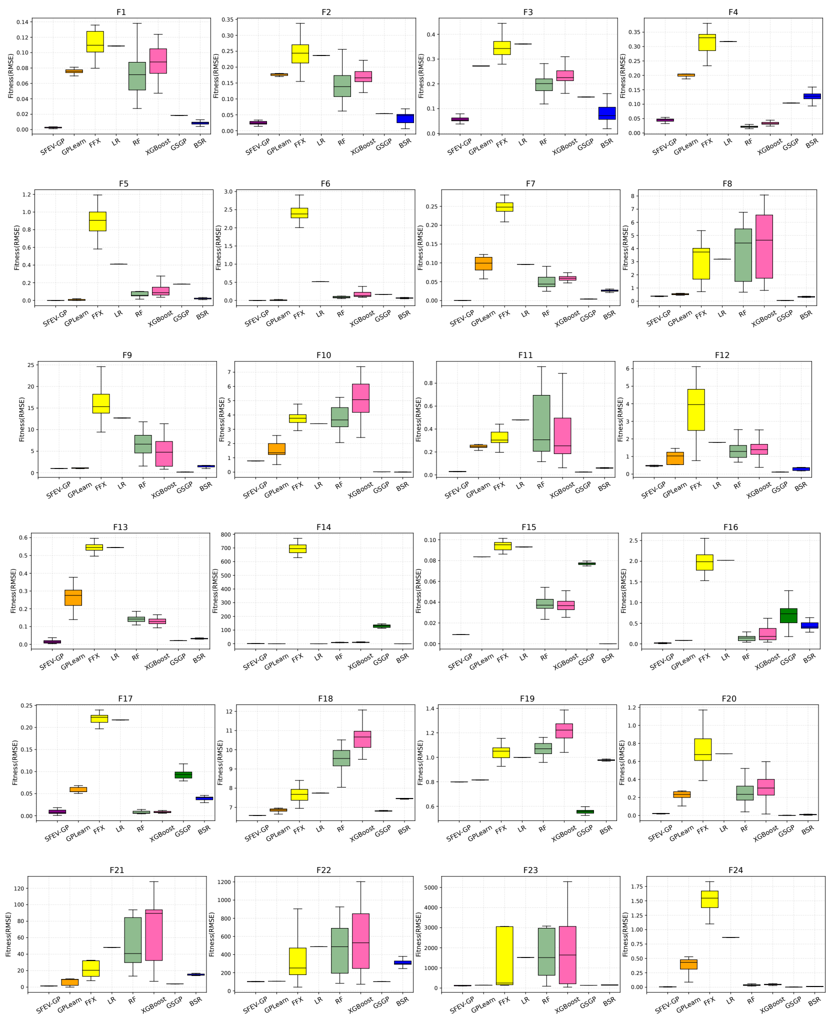

Fig. 1: Comparison of the RMSE fitness results on the benchmarks F1-F24.

Fig. 2: Comparison of the RMSE fitness results on the benchmarks F25-F48.

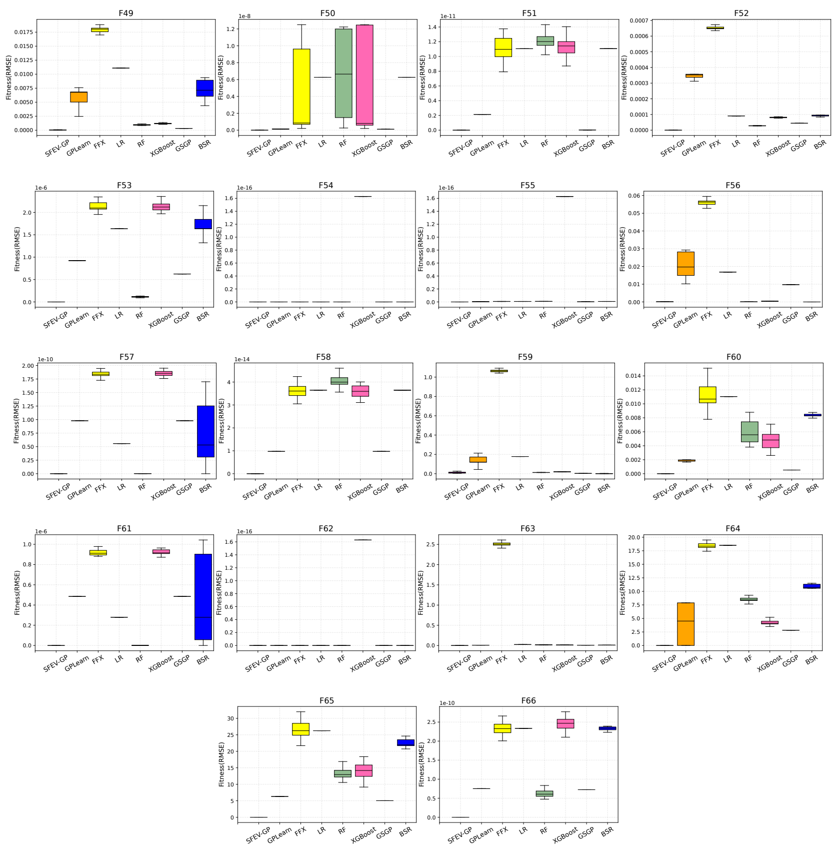

Fig. 3: Comparison of the RMSE fitness results on the benchmarks F49-F66.

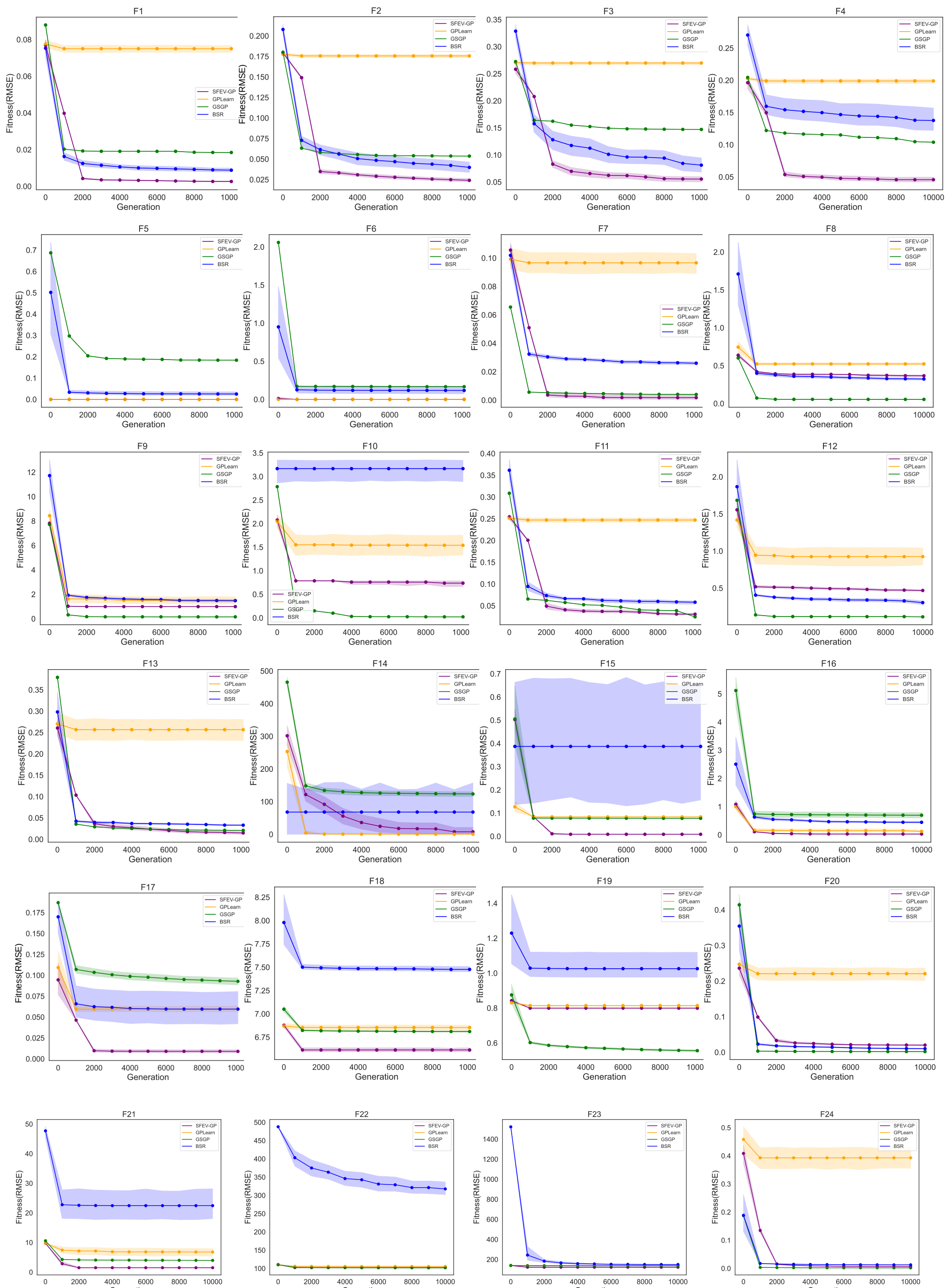

Fig. 4: Comparison of convergence on the benchmarks F1-F24.

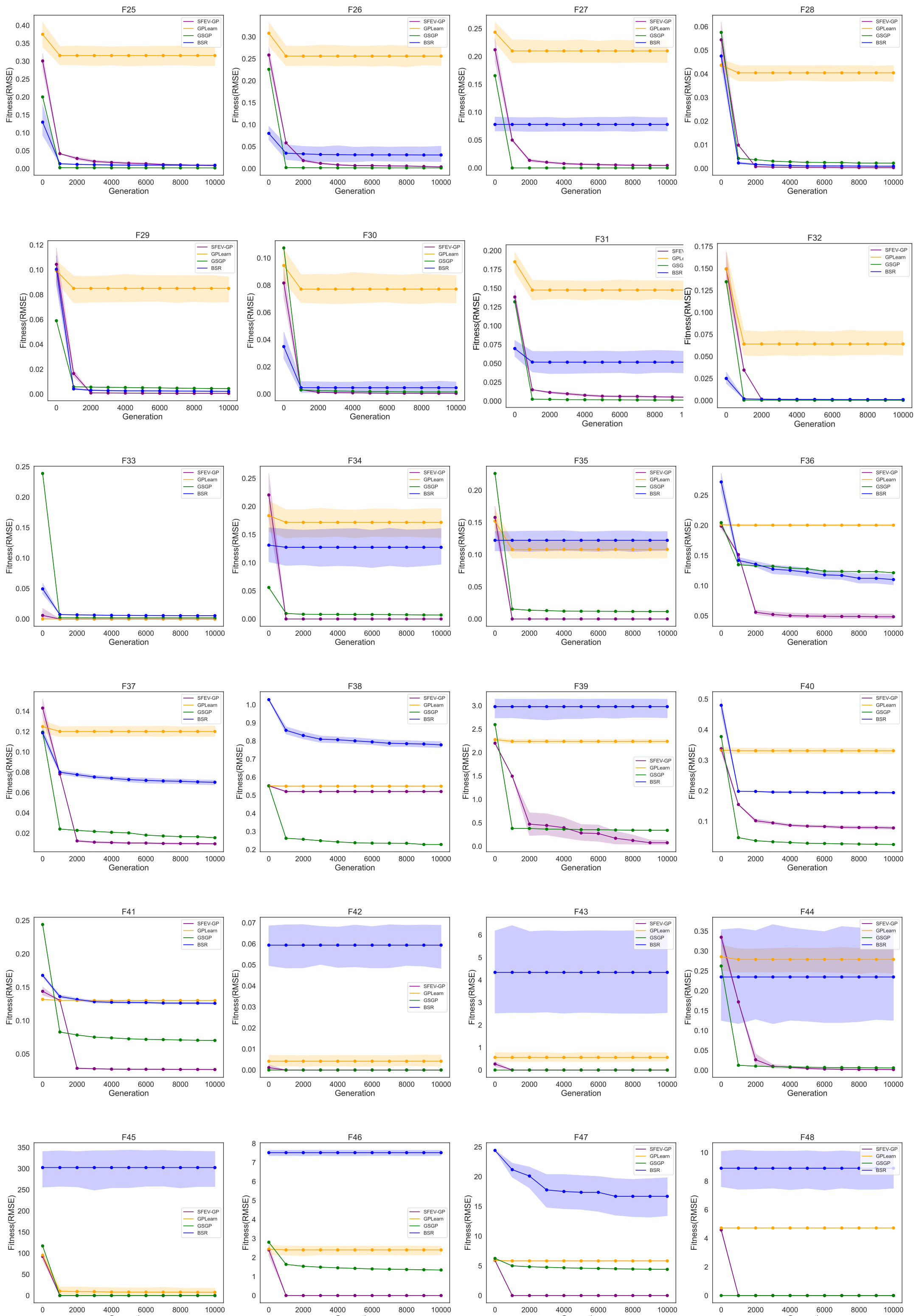

Fig. 5: Comparison of convergence on the benchmarks F25-F48.

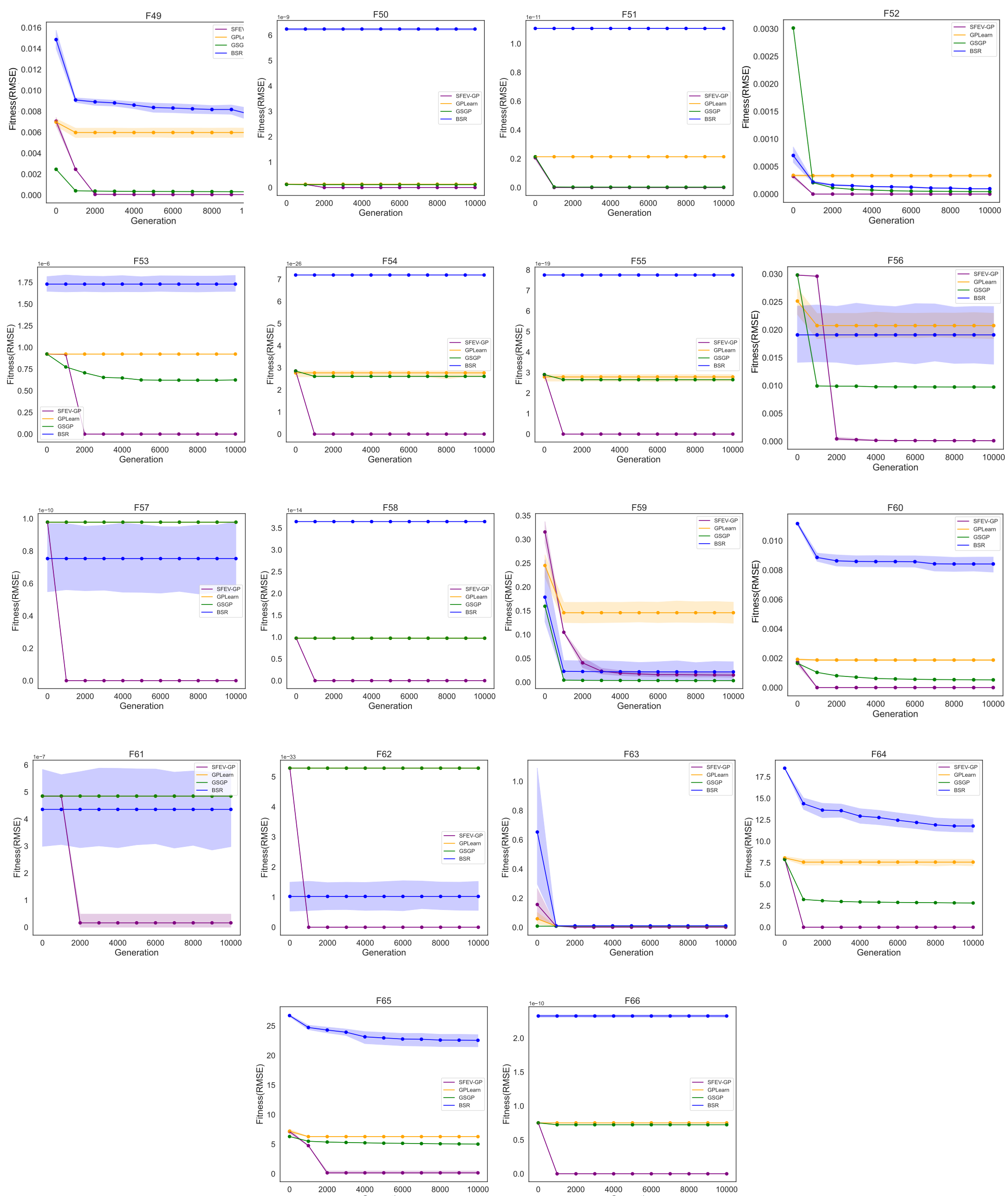

Fig. 6: Comparison of convergence on the benchmarks F49-F66.