Explorations of Stellarator Configuration Space with the Control Matrix

H. Mynick, N. Pomphrey, PPPL

Varenna workshop on Confinement and Stability of Alternative Fusion Concepts
Varenna, Italy, Oct. 12–16, 2000

The space $\mathbf{Z} \equiv \{Z_{j=1,..N_z}\}$ of variables defining a stellarator configuration is large. To find attractive design points in this space, or to understand operational flexibility about a given design point, one needs insight into the topography in $\mathbf{Z}$-space of the physics figures of merit $\mathbf{P} = \{P_i(\mathbf{Z})\}(i = 1, ..M_p)$ which characterize the machine performance (*e.g.*, transport, kink stability, *etc.*). The control matrix (CM) method used in this work provides a mathematical means of obtaining this.[1]

Here, we apply the CM method in studying some candidate Quasi-Axisymmetric (QA) stellarator configurations considered by the NCSX design group. The associated exploration provides insight into results earlier found empirically, independent control over the P_is, and guidance on improving the operation of the automated optimizer, which has become a mainstay in such design efforts.

•CM-Method: Formulation:

•We consider 2 linearly–related 'configuration spaces' $\mathbf{Z}$ and $\mathbf{X}$ or $\mathbf{I}$ specifying a stellarator:
·The 'reduced-space' $\mathbf{Z}$ of combinations of those X_j or I_j which capture the most important physics:
$\mathbf{Z} \equiv \{Z_{j=1,..N_z}\}$, where $N \equiv N_z \leq N_x$.
·The 'full-space' of coefficients:
(I)Fixed-boundary application:
$\mathbf{X} \equiv \{X_{j=1,..N_x}\} = (R_{\mathbf{m}_1}, Z_{\mathbf{m}_1}, R_{\mathbf{m}_2}, ...Z_{\mathbf{m}_{N_x/2}})$
needed for a VMEC equilibrium specification of the boundary. Here, $\mathbf{m} \equiv (\tilde{n} \equiv n/N_{fp}, m)$, and $N_x \sim 70$.
(II)Free-boundary application:
$\mathbf{I} \equiv$ currents I_j in the j^{th} coil of a given coil set, plus, when desired, $\langle\beta\rangle$.

•Over $\mathbf{Z}$, we consider the behavior of $M \equiv M_p \sim 5$ physics figures of merit
$\mathbf{P} \equiv \{P_i(\mathbf{Z})\} = (\chi_1^2, \chi_2^2, W_1, W_2, \lambda)$, where:
·$P_5 \equiv \lambda =$ kink eigenvalue (from TERPSICHORE), and P_{1-4} are 4 measures of the ripple, hence of the level of nonaxisymmetric transport one might expect: $P_{1,2} \equiv \chi^2_{1,2} \equiv \chi^2(\psi_{1,2}) \equiv N_{\mathbf{m}}^{-1} \sum_{m,\tilde{n}\neq 0} B^2_{\mathbf{m}}/B_0^2$, with ψ the toroidal flux, $= \psi_a$ at the edge, and $\psi_1/\psi_a = 1/4, \psi_2/\psi_a = 1/2$.
·$P_{3,4} \equiv W_{1,2}$ is the 'water function'[2] at $\psi_{1,2}$, measuring how deep the ripple wells are over a flux surface.

•Expand $\mathbf{P}(\mathbf{Z} = \mathbf{Z}_0 + \mathbf{z}) = \mathbf{P}(\mathbf{Z}_0) + \mathbf{p}$ about $\mathbf{Z} = \mathbf{Z}_0$. One has (writing in component-form, with summation over repeated indices assumed)

$$p_i(\mathbf{Z}_0 + \mathbf{z}) = C_{ij}(\mathbf{Z}_0) z_j + \frac{1}{2} H_{ijk}(\mathbf{Z}_0) z_j z_k + (h.o.), \quad (1)$$

with $h.o.$ ≡ higher–order terms. For small enough $\mathbf{z}$, one has matrix equation

$$\mathbf{p} = \mathbf{C}_0 \cdot \mathbf{z}, \quad (2)$$

with $\mathbf{C}_0 \equiv \mathbf{C}(\mathbf{Z}_0)$ the $M \times N$ 'control matrix' at design point $\mathbf{Z}_0$. Using the SVD theorem
$\mathbf{C}_{M\times N} = \mathbf{U}_{M\times N} \cdot \mathbf{W}_{N\times N} \cdot \mathbf{V}^T_{N\times N}$,
(with $\mathbf{U}, \mathbf{V}$ unitary matrices, and $\mathbf{W}$ a diagonal matrix), one may obtain the pseudo-inverse $\mathbf{C}_0^+$ of $\mathbf{C}_0$.
·Taking the particular basis set $\boldsymbol{\pi}^{i=1,M}$ which has 1 in the i^{th} position and 0 elsewhere, one has the corresponding set $\boldsymbol{\xi}^i$ of displacements

$$\boldsymbol{\xi}^i \equiv \mathbf{C}_0^+ \cdot \boldsymbol{\pi}^i,$$

physically representing a set of displacements which vary a single physics parameter P_i, leaving the others unchanged. These span the 'range' of $\mathbf{C}$. The $(N - M)$ vectors $\mathbf{v}^{i=M+1,N}$ spanning the nullspace of $\mathbf{C}$ (change the configuration without modifying any of the P_i) also important for design flexibility.

(I)Fixed-boundary application:

•Topography of Z-space:

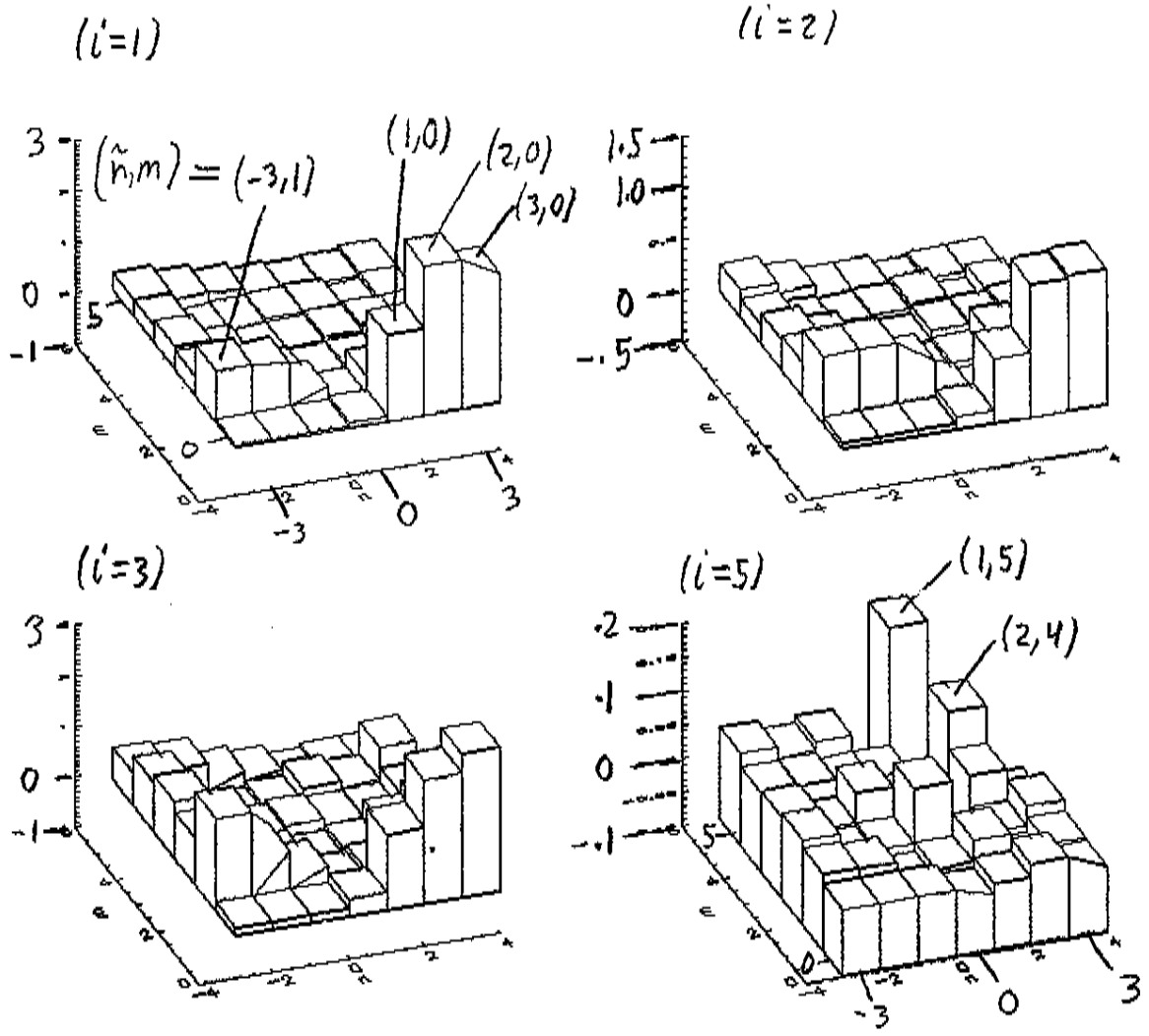
•The validity of Eqs. (1) or (2) depends on the typical scales of variation in **X**- or **Z**-space of the P_i. We have assessed this variation for all X_j for the P_i presently being used, in the vicinity of the $\mathbf{X}_0$ = c10–c82 family[4] of configurations.

•Reducing the dimensionality of Z:

•The dimensionality $N_x = N_z$ of the search space used up to now and $N_x - M_p$ of its null subspace are large (78 and 73, resp.). The smaller we can make these, the better. Here, we reduce N_z by 2 general methods:
(a)Removing the redundancy in the **X**-specification, by using only displacements Y_i normal to the plasma boundary, rather than $\delta R, \delta Z$ separately. This reduces N_z from $N_x = 78$ to $N_x/2 = 39$.
(b)Taking only the perturbations most effective in varying some P_i. Here, we choose the 4 most effective for P_1, and for P_5, thus further reducing N_z from 39 to 8.

•Reduced Model & 1st Calculation of C_{ij}, H_{ijk}:

•Sensitivity histograms in the $N_z = 39$ **Z**-space:



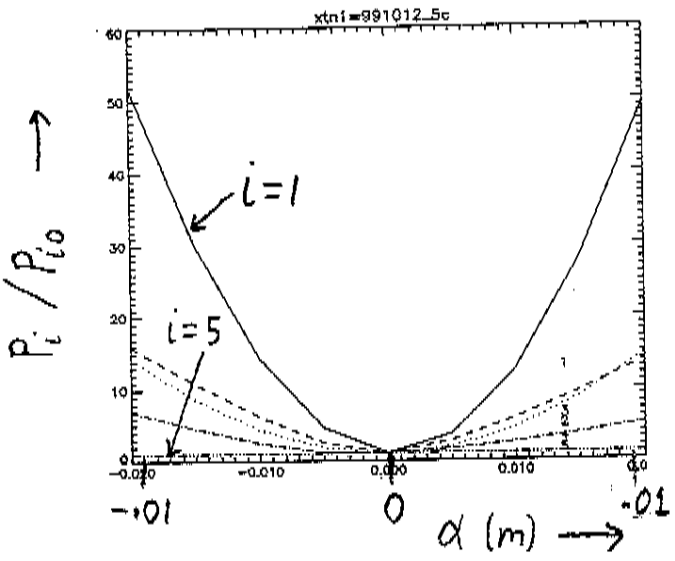
•Ranking $|P_i/P_{i0} - 1|$, select 4 most effective harmonics **m** for each of $i = 1, 5$:

•For P_1: $(\tilde{n}, m) = \{(1,0), (2,0), (3,0), (-3,1)\}$

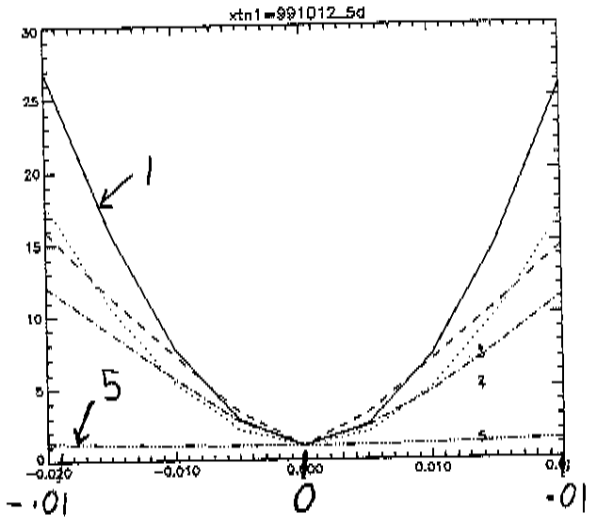
•For P_5: $(\tilde{n}, m) = \{(1,3), (1,4), (2,4), (1,5)\}$.

•Topography of reduced space:

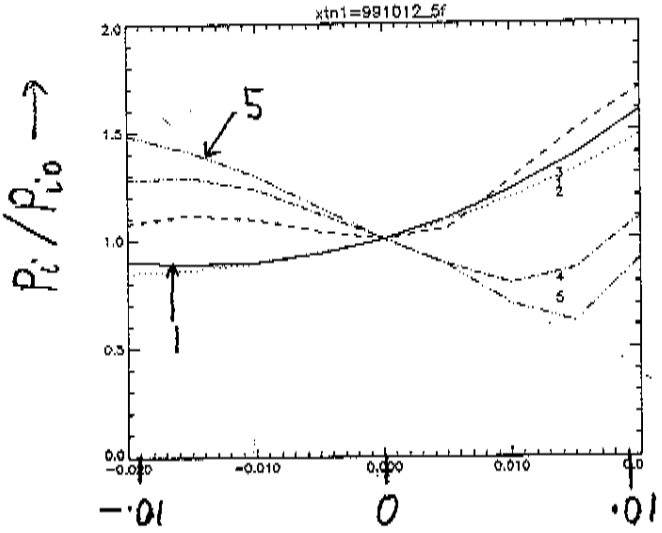
$j = 1$, $\underset{\sim}{m}_j = (1,0)$



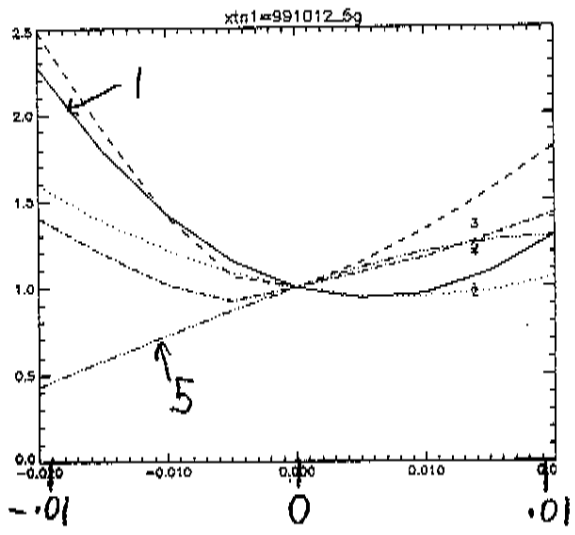
$j = 4$, $\underset{\sim}{m}_j = (-3,1)$



$j = 5$, $\underset{\sim}{m}_j = (1,3)$

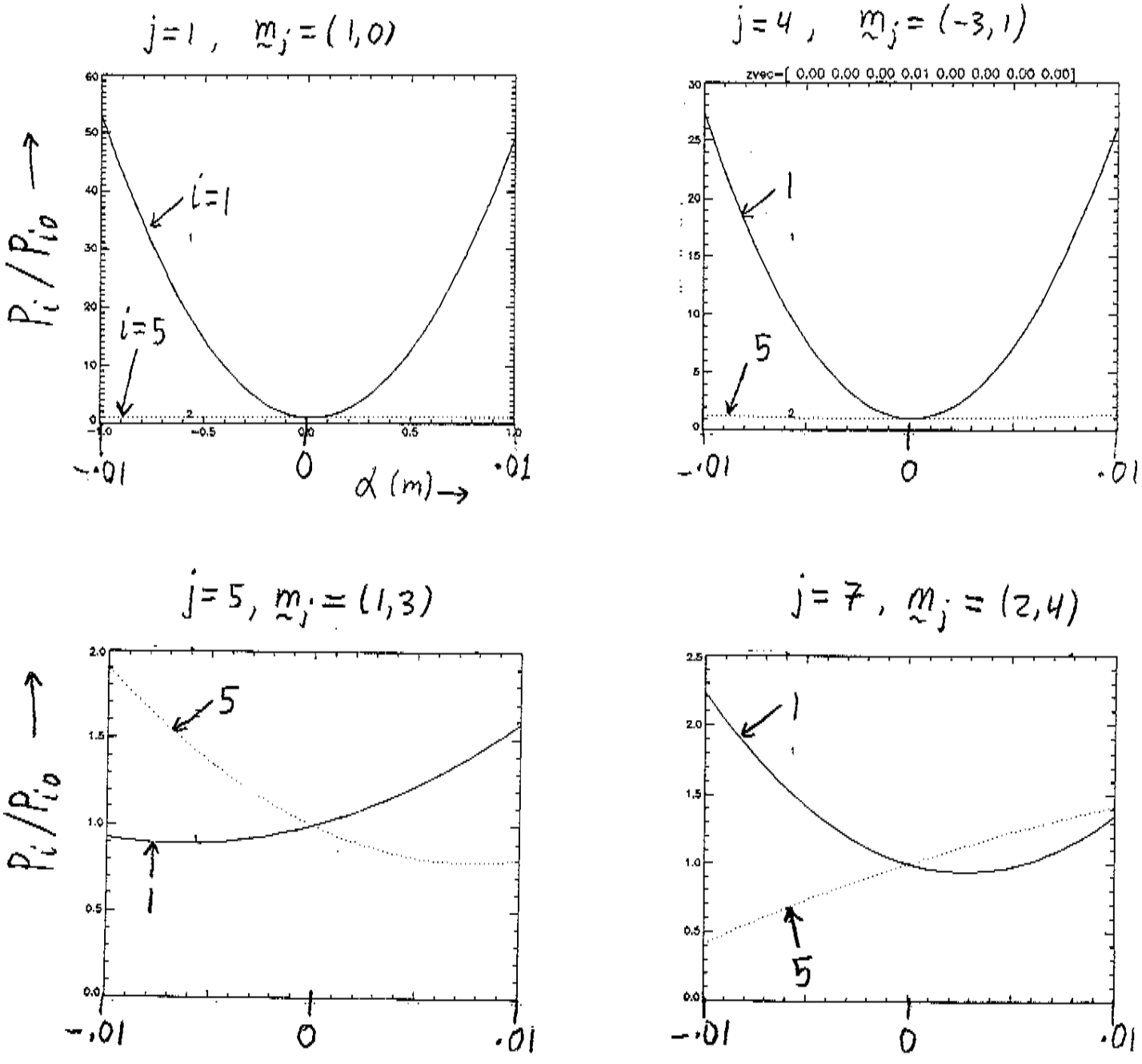


$j = 7$, $\underset{\sim}{m}_j = (2,4)$



- Distill $G_{ij}, H_{ijk} \Rightarrow$ 'Quadratic Model' for $N_z = 8$: (Requires $2N_z^2 = 128$ perturbed equilibria about $\mathbf{Z}_0$.)

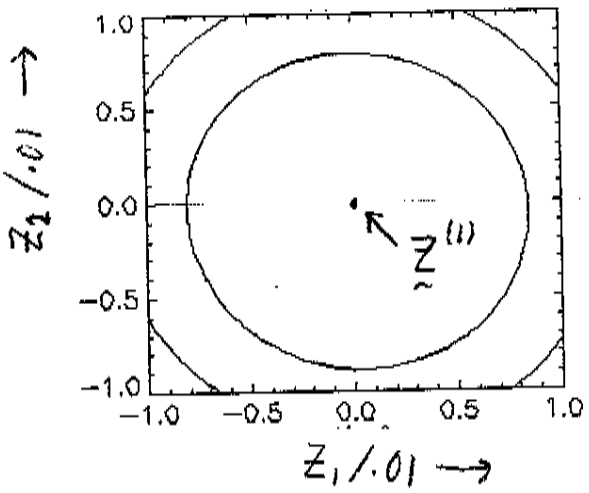
- Compare this semi-analytic model with numerical results just shown:



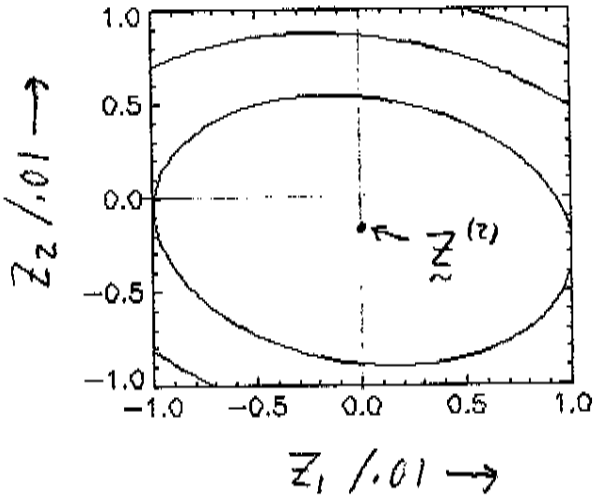
- Use Quadratic Model to analyse **Z**-space structure, compute $\boldsymbol{\xi}^{i}$'s, *etc.*.

  - Plot $P_{1,5}$ versus pairs (z_{j1}, z_{j2}):

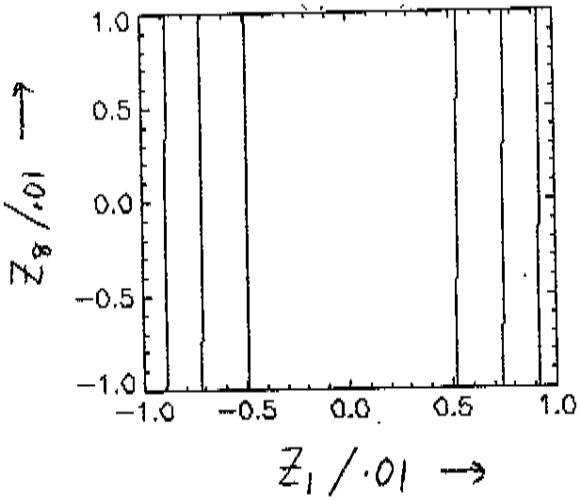
$P_1(z_1, z_2)$:



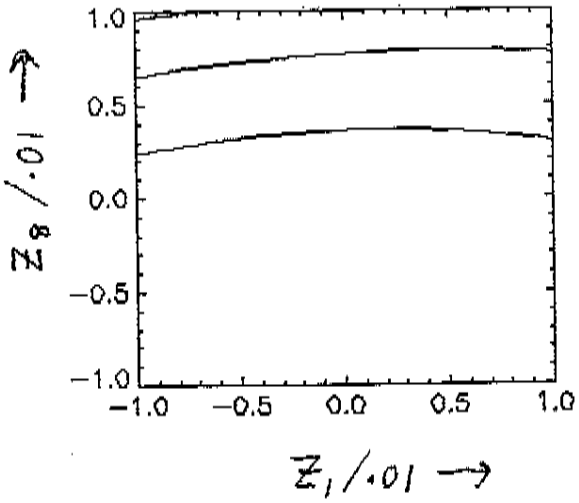
$P_5(z_1, z_2)$:



$P_1(z_1, z_8)$:



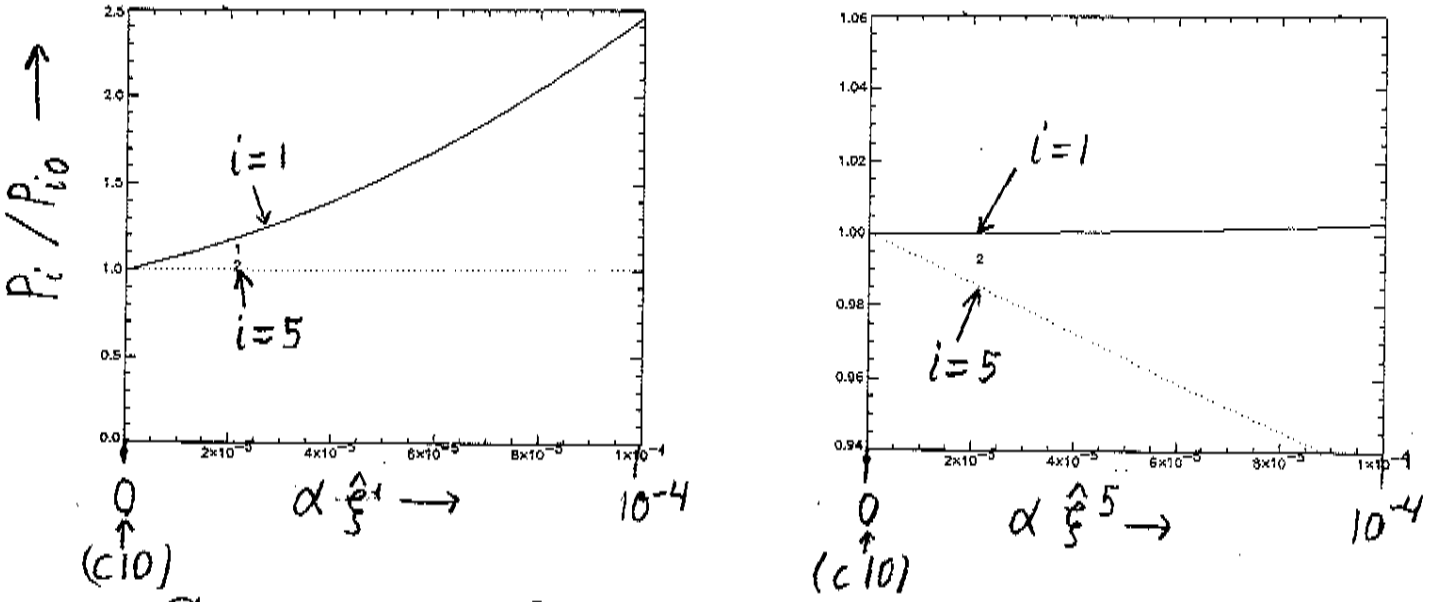
$P_5(z_1, z_8)$:



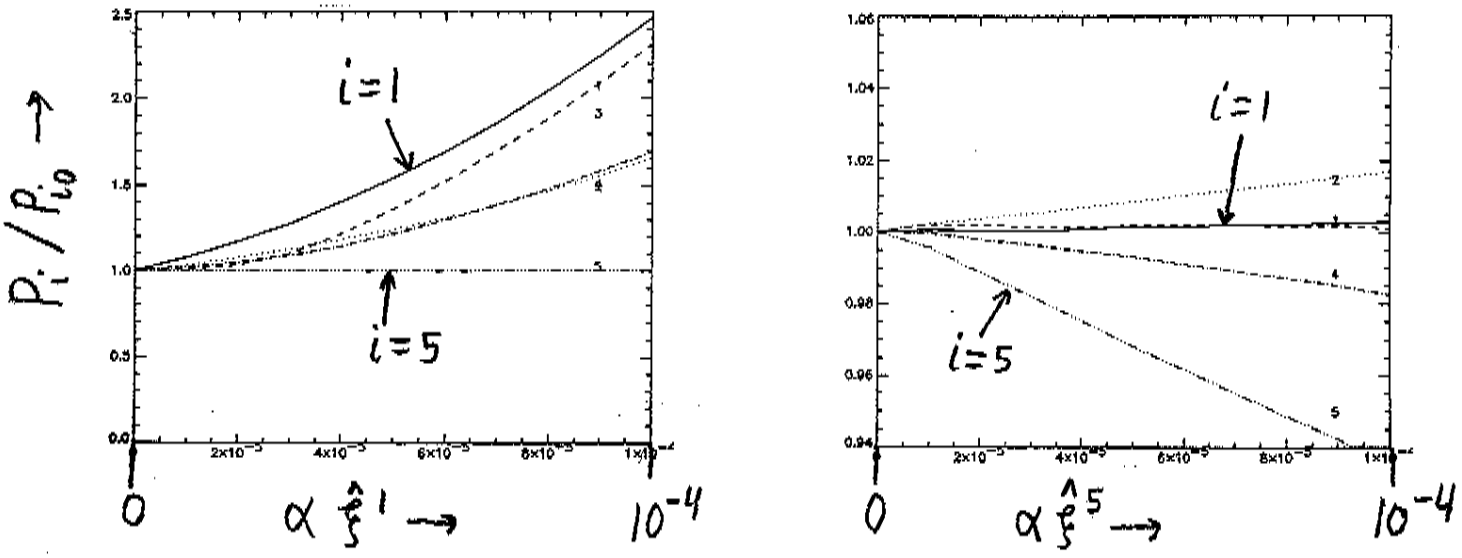
- (Note : z_1, z_8 almost perform job of $\underset{\sim}{\xi}^{1}, \underset{\sim}{\xi}^{5}$, viz, varying P_1, P_5 independently.)

- 'Proof of Principle' of CM method:

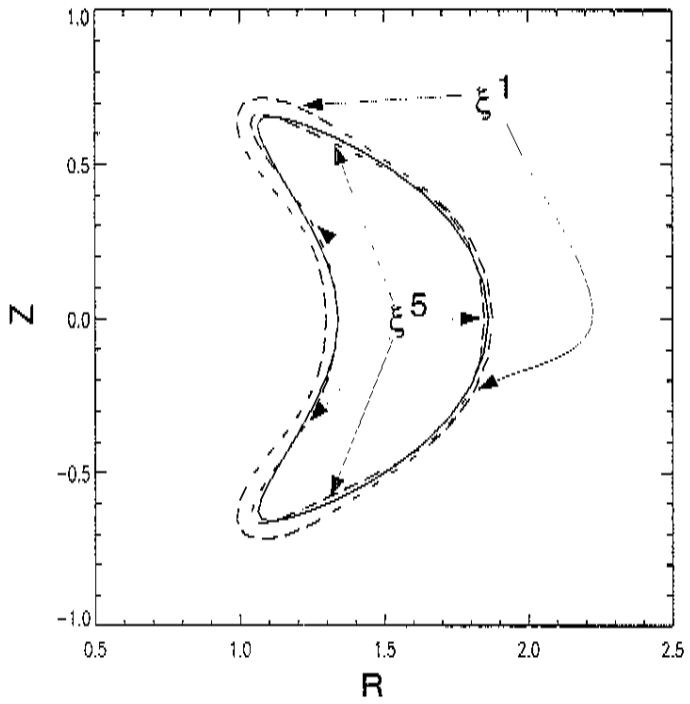
- Plot P_i versus $\alpha\xi^{j=1,5}$, verifying their independent control of $P_{1,5}$:

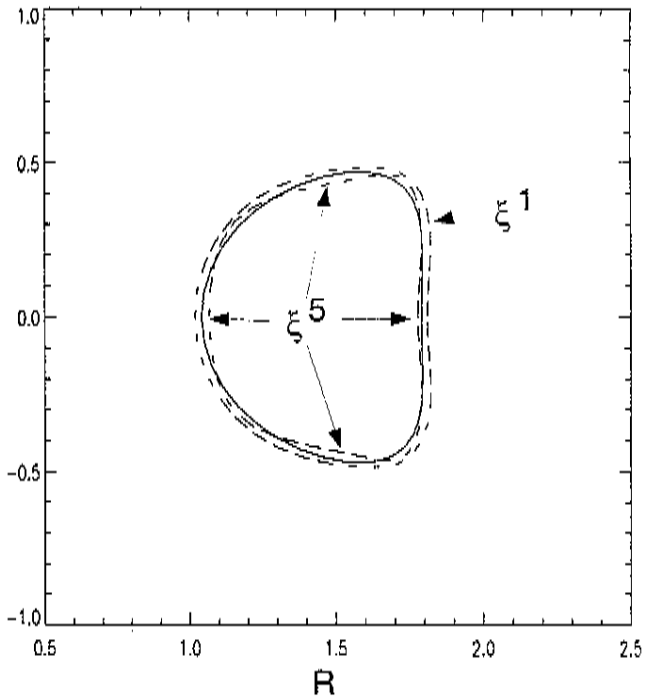


- Compare with numerical results from perturbed equilibria:



•For C10, $\boldsymbol{\xi}^5$ manifests the outboard indentation at $\tilde{\zeta} \equiv N_{fp}\zeta = \pi$ previously noted to stabilize the kink, enhancing C10's negative triangularity at $\tilde{\zeta} = \pi$.





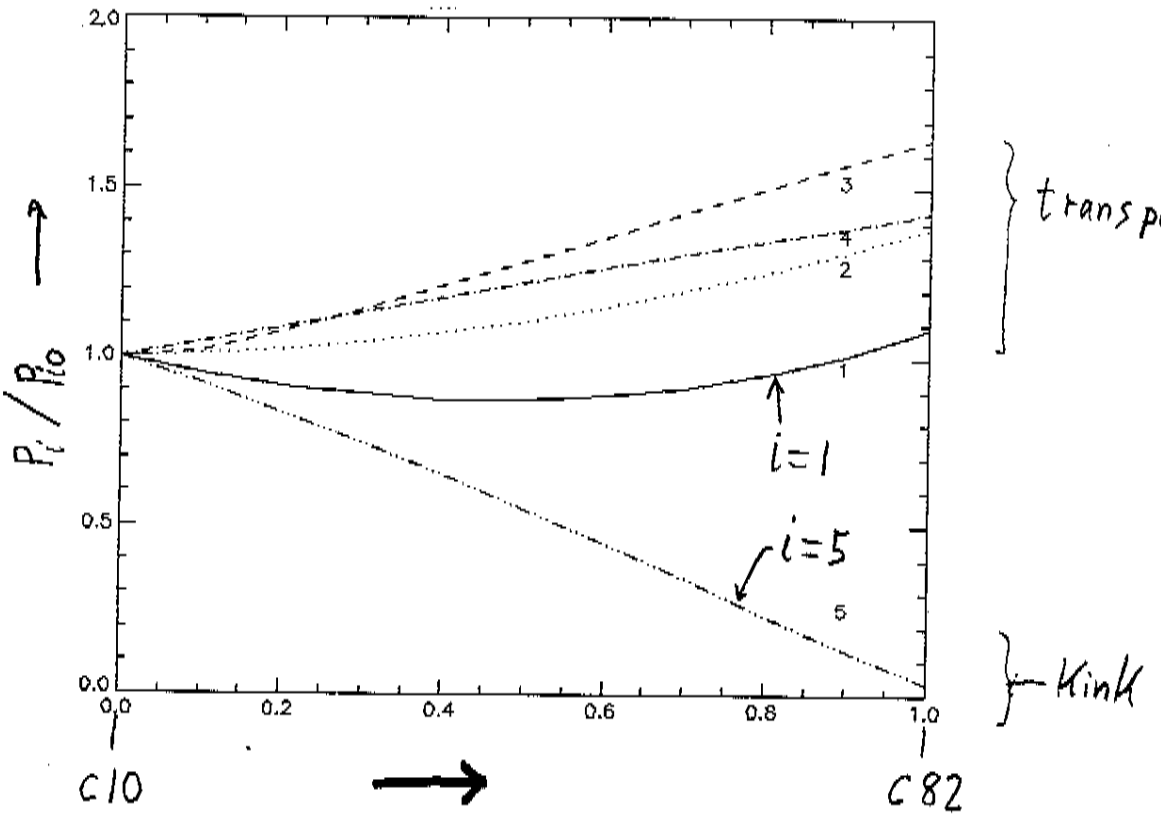
However, this behavior is not generic: for PG1, $\boldsymbol{\xi}^5$ enhances its *positive* triangularity, consistent with tokamak intuition on kink stabilization.

•Other QAS Design Points:

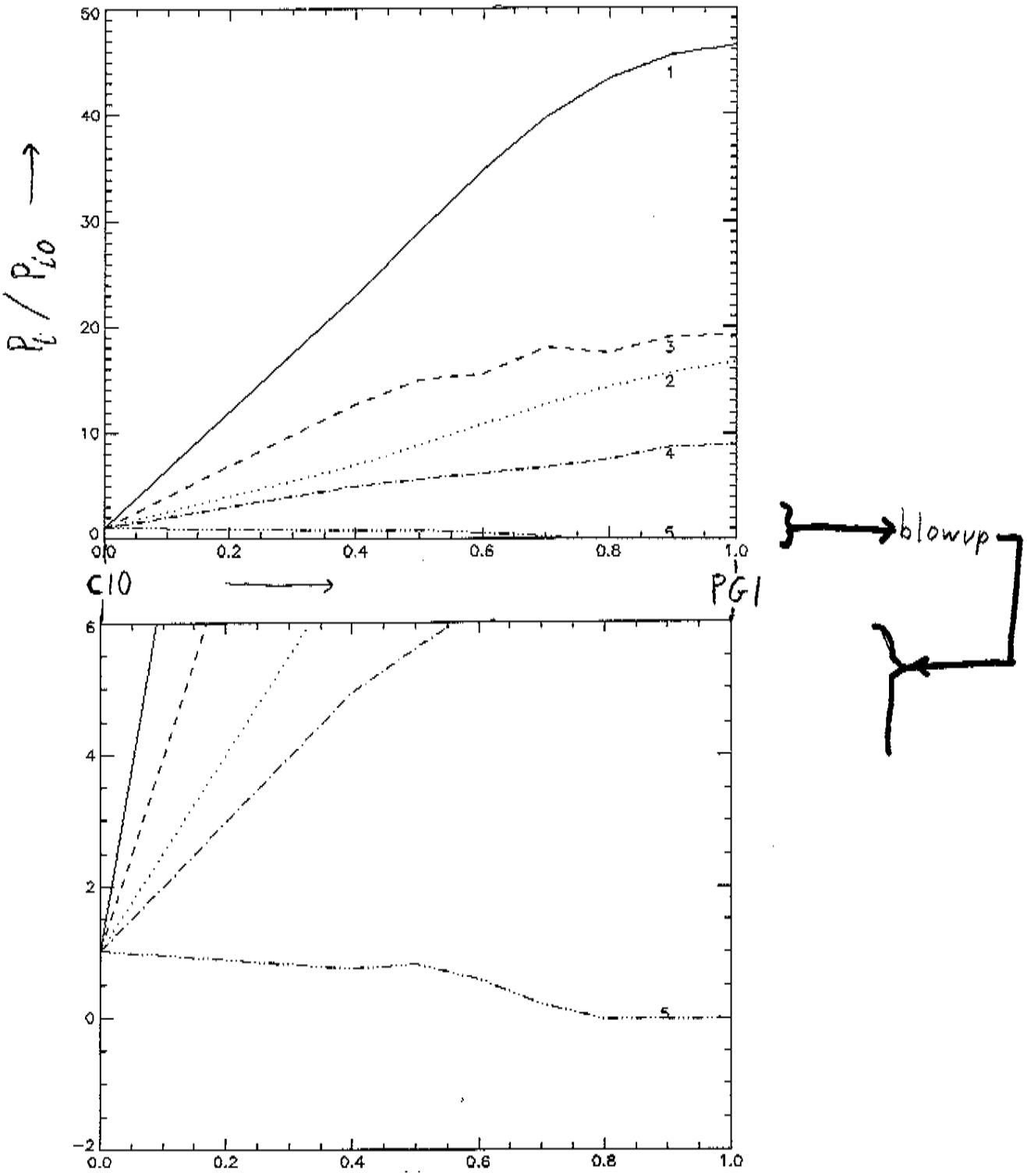
•c10 was arrived at along an involved path of human interaction with the optimizer, and it is unclear that other regions of **Z**-space, which would have been reached from different starting points, might not yield superior configurations. Thus, we are starting to study other proposed QAS configurations[2,3] with the same methods, and to consider the variation of the P_i as one moves from one such point $\mathbf{Z}_0$ to another.

•Warmup (**Z**'s nearby): The path c10 → c82:

•c82 was obtained from c10 in an effort to stabilize the kink. The level of QA-ness was slightly degraded in compensation. This borne out by the P_i's along a straight-line path in **Z**-space: $|\underset{\sim}{X}_{c82} - \underset{\sim}{X}_{c10}| \simeq .041$ m



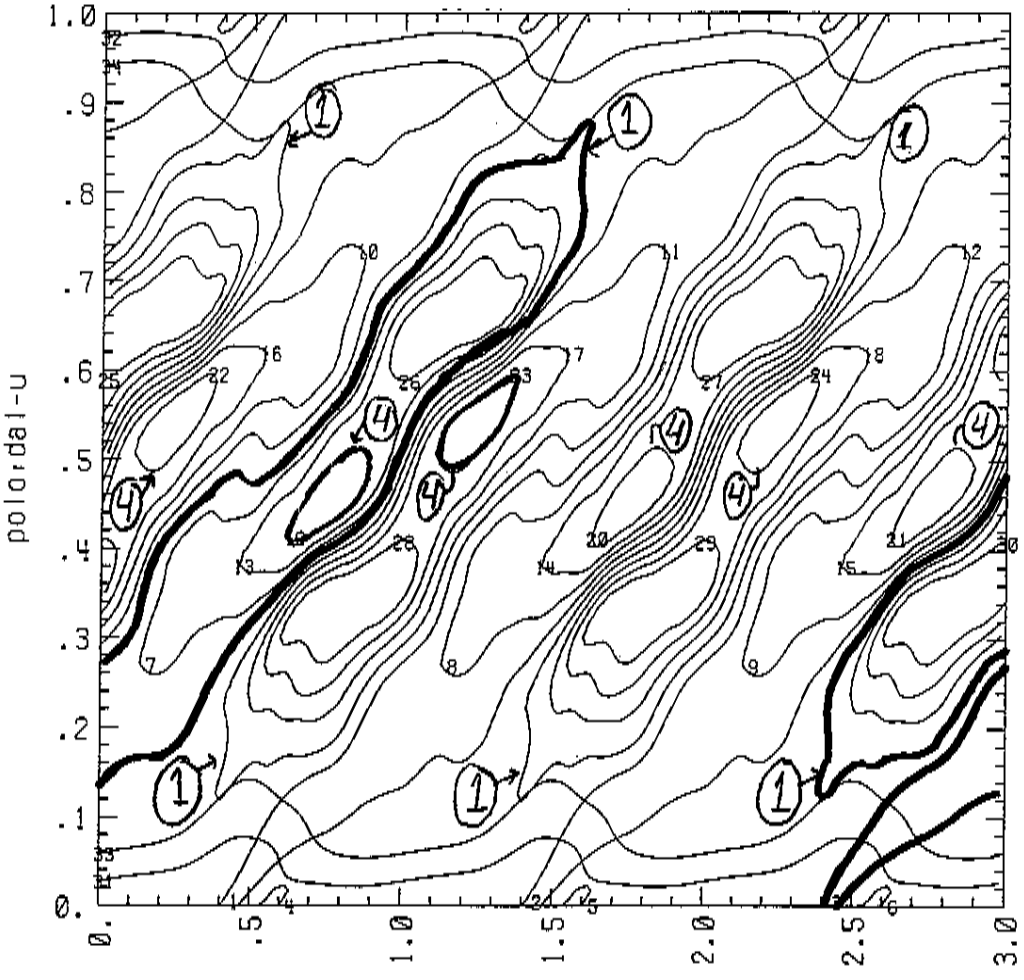
- The path c10 → PG1 (preliminary):
- Define norm $|\underset{\sim}{X}| \equiv [\sum_j X_j^2]^{1/2}$, $|\underset{\sim}{X}_{PG1} - \underset{\sim}{X}_{c10}| \simeq .228$ m
- PG1 characterized by much better kink stability, substantially worse QA-ness (mainly due to large mirror field $B_{m=0,\tilde{n}=1}$ present to enhance stability.

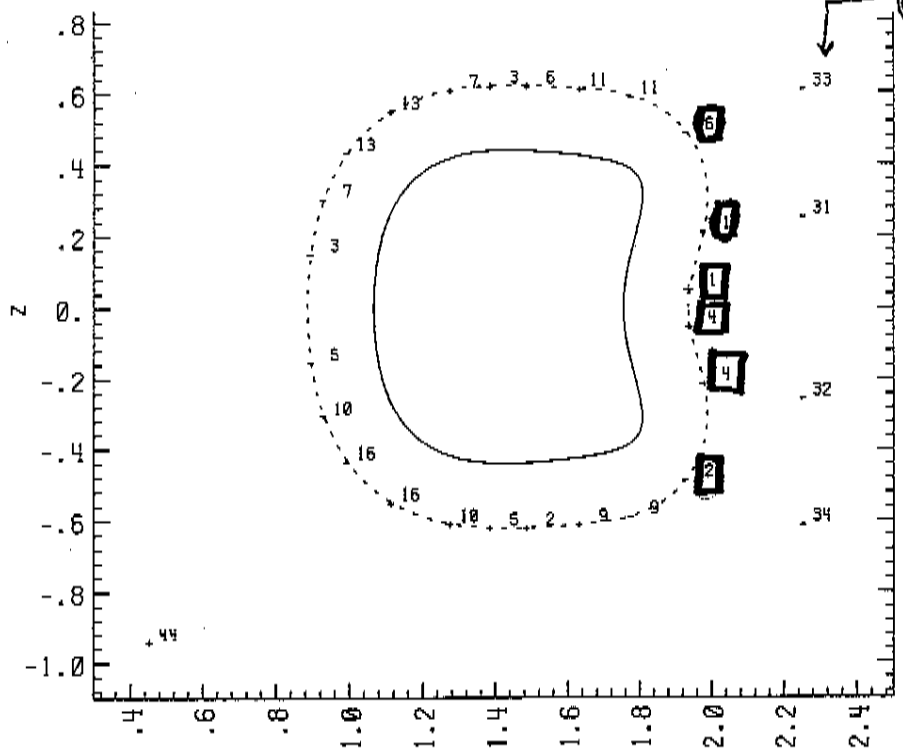


(II)Free-boundary application: $\mathbf{Z} = \mathbf{X} \rightarrow \mathbf{I}$.

•Now apply the CM machinery to a new configuration space $\mathbf{Z}$, where instead of the Z_j being the amplitudes X_j describing the plasma boundary, they are now the currents I_j in the external coils, plus sometimes the plasma $\langle\beta\rangle$, and parameters characterizing the profile shapes. Making this transition requires using free- instead of fixed–boundary VMEC.

•Coil set c82.4064_rev6: Neighborhood of C82b



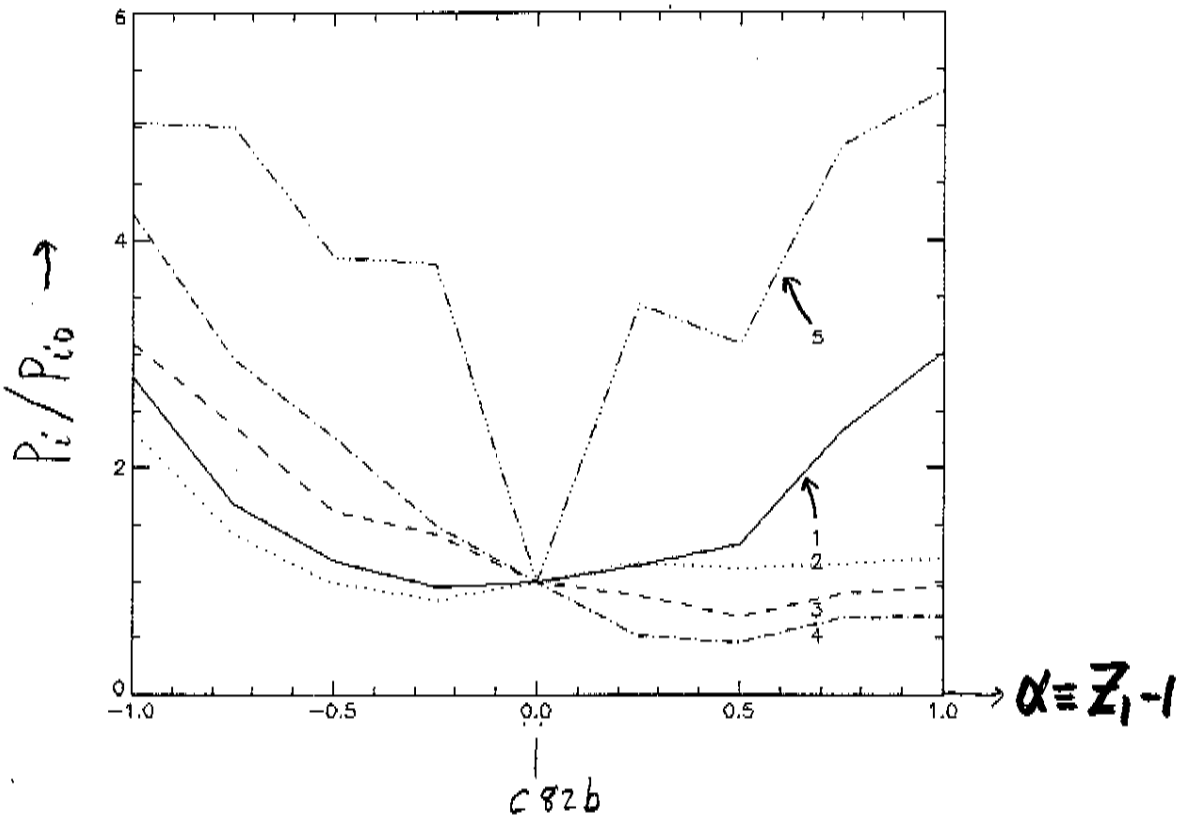


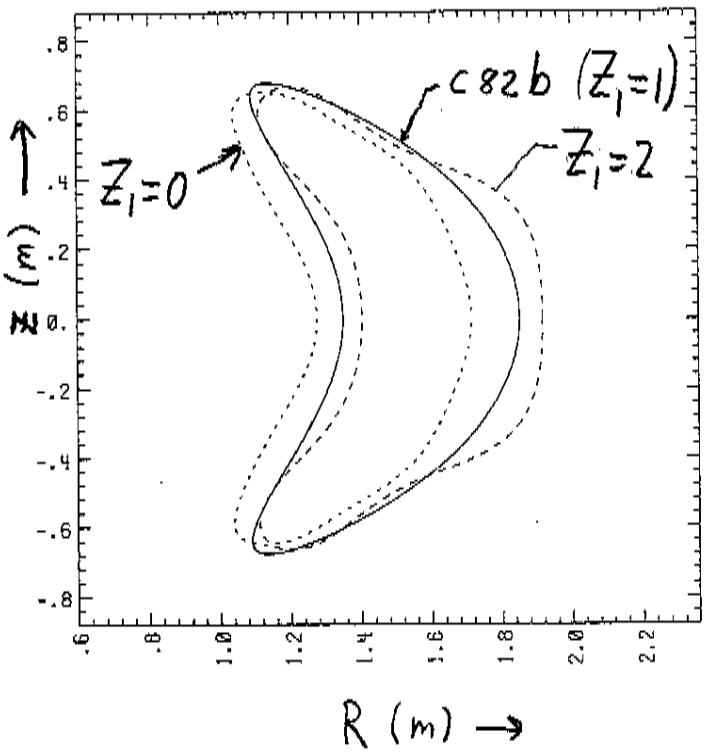
(6) = EF group

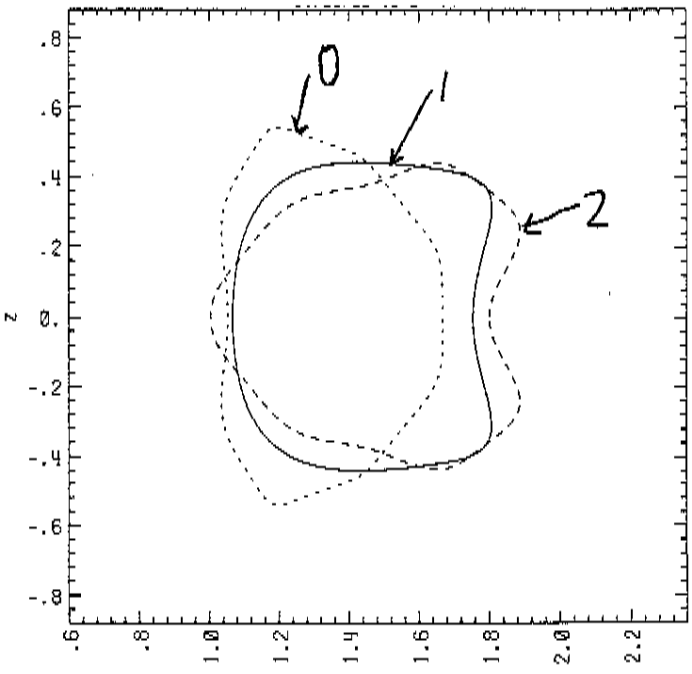
1-6 = group (1)

19-24 = group (4)

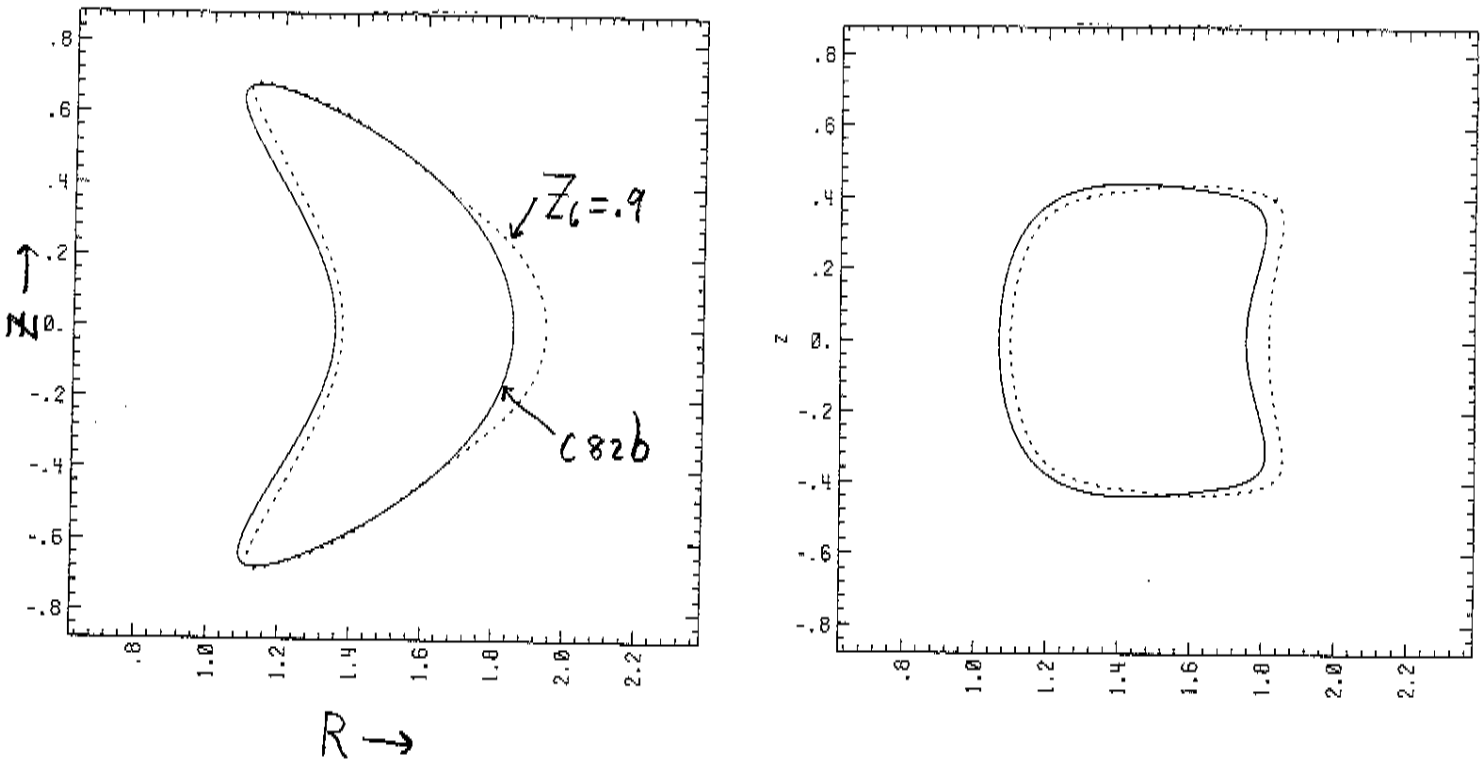
- Sweep of $Z_1 \equiv I_1/I_{10}, = 0, 0.25, .., 1.75, 2.0.$ Plot $P_{i=1-5}$ vs Z_j. (P_5 uses coec=coep=1.05):





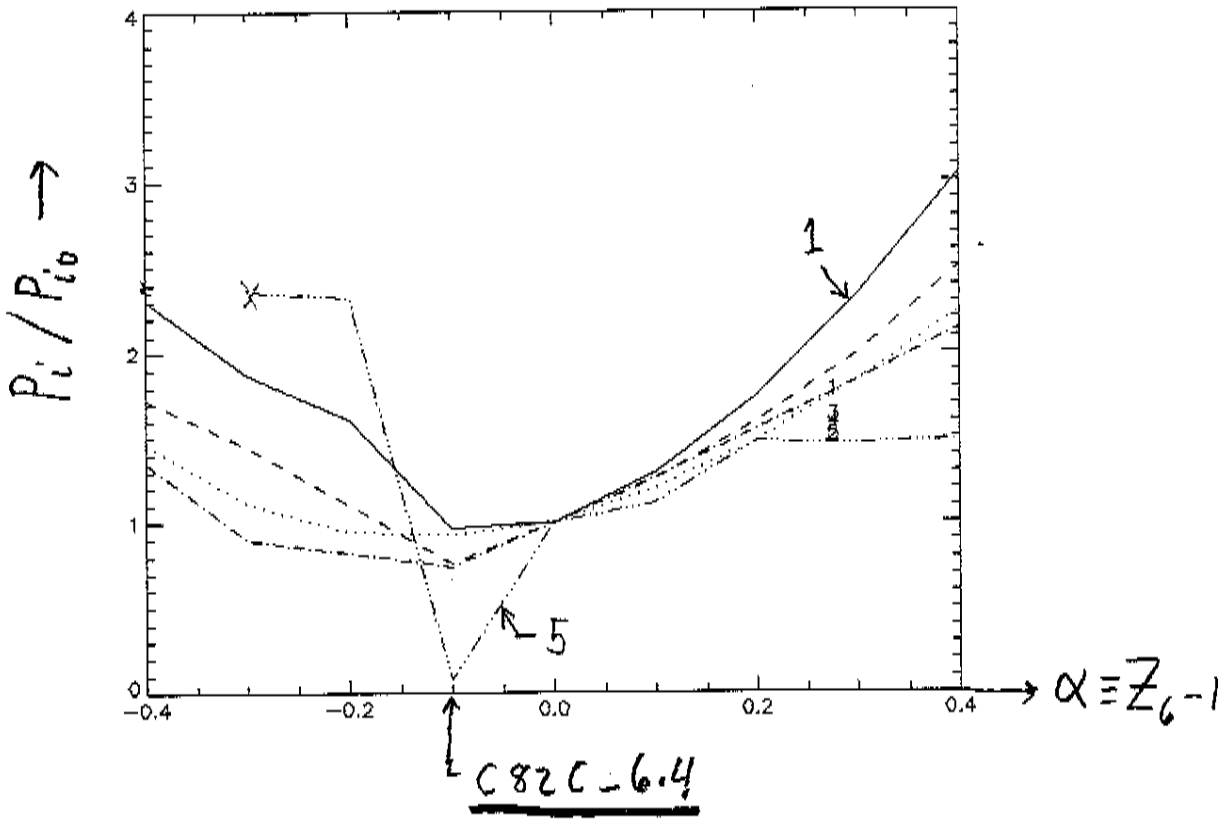


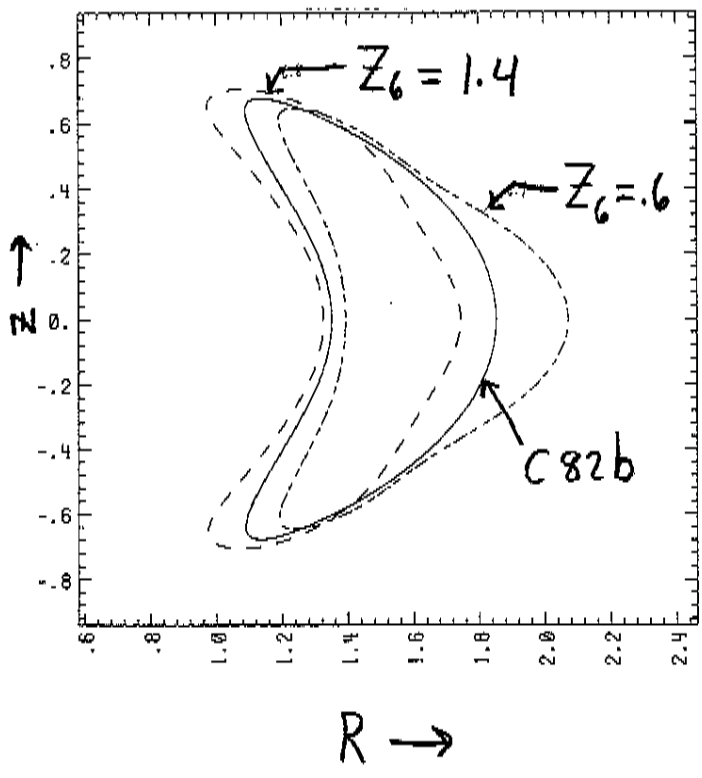
- Configuration c82c_6.4 : $Z_6 = 0.9$

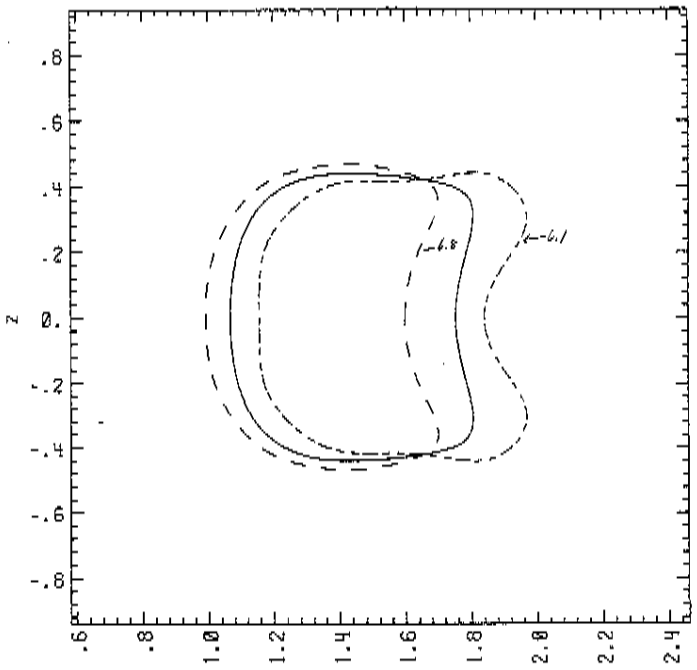


machine	P_1	P_5
c10a	1.4276E-04	-1.43303E-03
c82	1.5517E-04	-4.92465E-05
c82b	3.5605E-04	6.65416E-06
(with coec=coep=1.05):		
c82b2	3.5605E-04	-5.11848E-04
c82c_6.4	3.4273E-04	-4.25327E-05
(Brooks-Reiersen tilted-coil case121.5):		
c82BR.0	1.7708E-04	-1.61210E-03

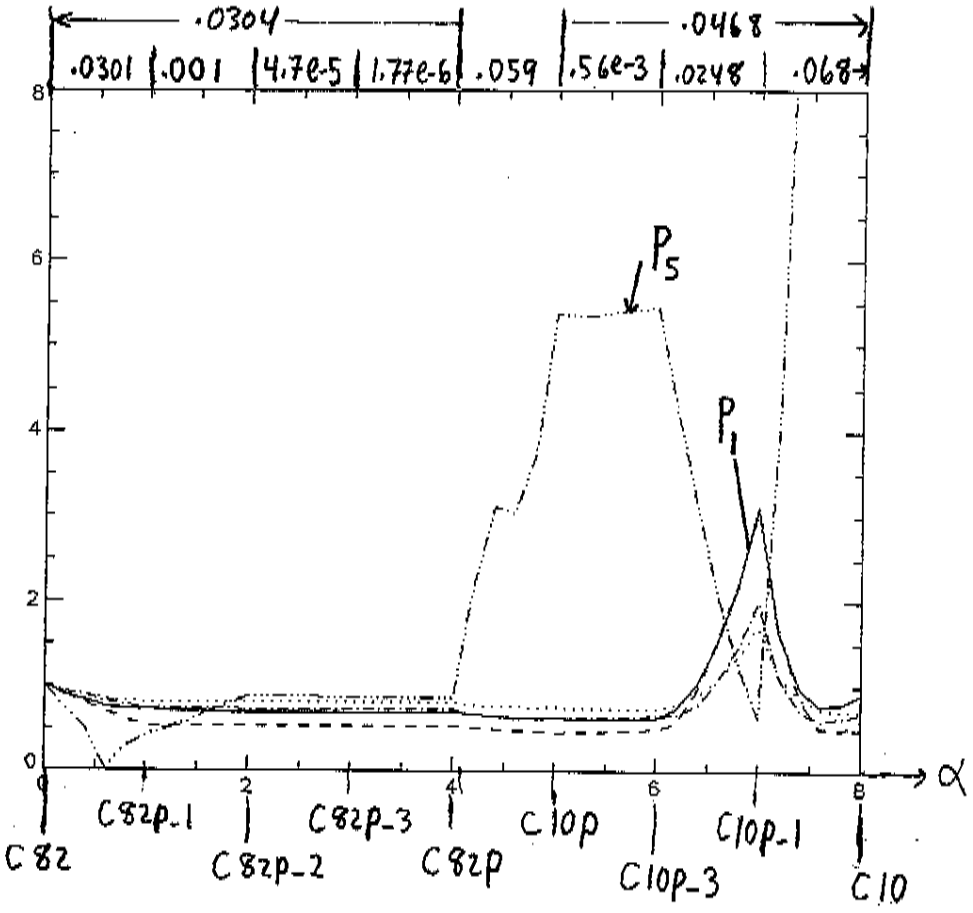
- Coil group 6=EF coils: P_i vs $Z_6, = 0.6, 0.7, ...1.4$:

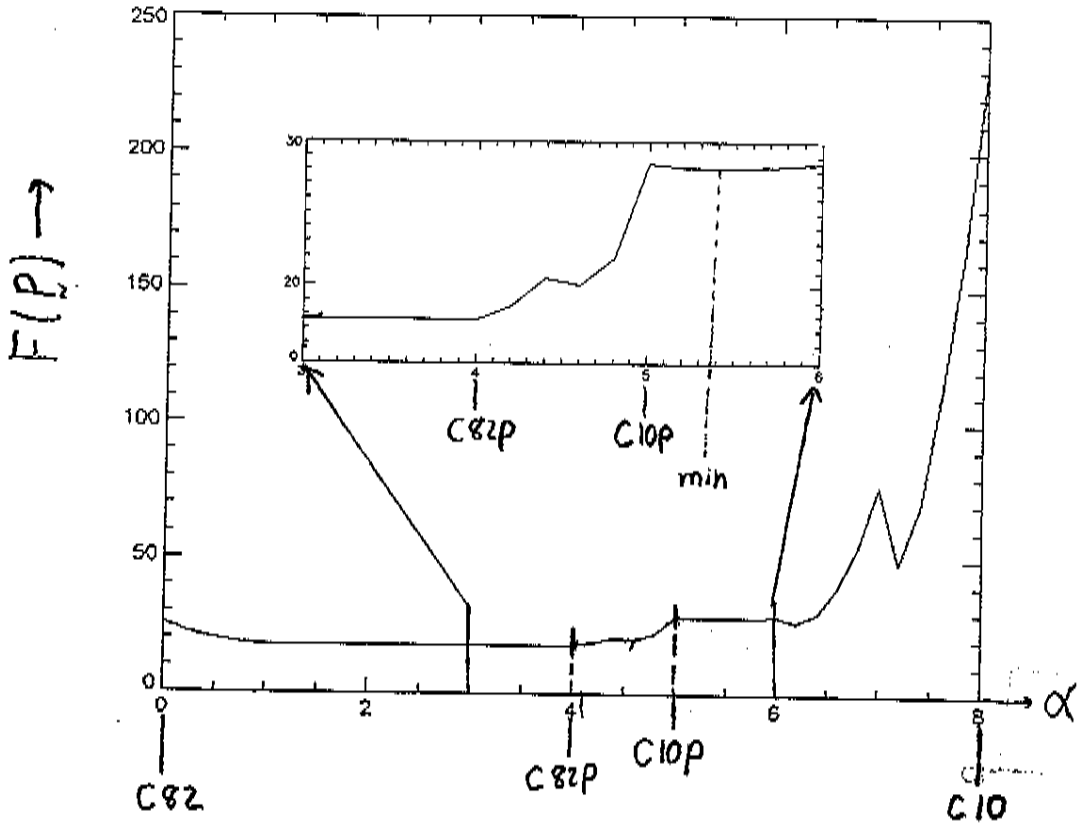






(III) Analyze Optimizer Behavior: Levenberg-Marquardt (LM):





•Differential Evolution (DE) algorithm[3] :

•Similar to genetic algorithm, but more natural to continuous spaces $\mathbf{z}$.

•Evolves an ensemble $\mathbf{z}_{i,G}, i = 0, 1, ...NP - 1$ of NP system points in D-dimensional space $\mathbf{z}$, through generations $G = 0, 1, ...$, similar to GA.

•Several variants. Basic method for getting new system points $\mathbf{z}_{i,G+1}$ from old ones:

(i)Compute

$\mathbf{v}_i = \mathbf{z}_{best,G} + (\mathbf{z}_{r_2,G} - \mathbf{z}_{r_3,G})F,$

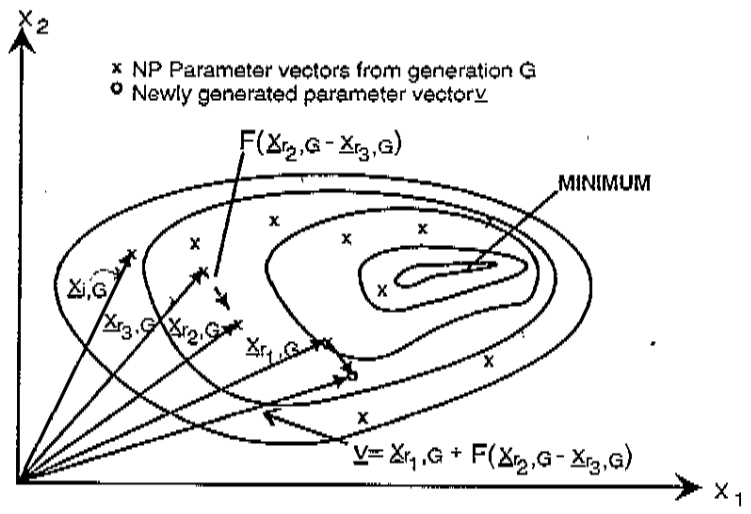
where $r_{1,2,3}$ random integers in $[0, NP - 1]$, and F is an adjustable parameter (0.9 typical).

(ii)Mutate a subgroup of L components v_j in $\mathbf{v}$ (similar to crossover in the GA) to produce trial point $\mathbf{u}$, according to:

$$u_j = \begin{cases} v_j & \text{for} j = n, n+1, ...\langle n + L - 1\rangle_D \\ (\mathbf{z}_{i,G})_j & \text{otherwise} \end{cases}$$

Here, $L \in [0, D - 1]$ chosen with probability $Pr(L = \nu) = (CR)^\nu$, and $\langle m\rangle = m$ modulo D.

(iii) If the cost function value $C(\mathbf{u})$ is smaller than that for a predetermined population member ($e.g.$, $\mathbf{z}_{i,G}$), it replaces that member: $\mathbf{z}_{i,G+1} = \mathbf{u}$.



•Storn-Price test results:

against Annealed Nelder&Mead strategy (ANM), Adaptive Simulated Annealing (ASA) (which is claimed to outperform GAs):

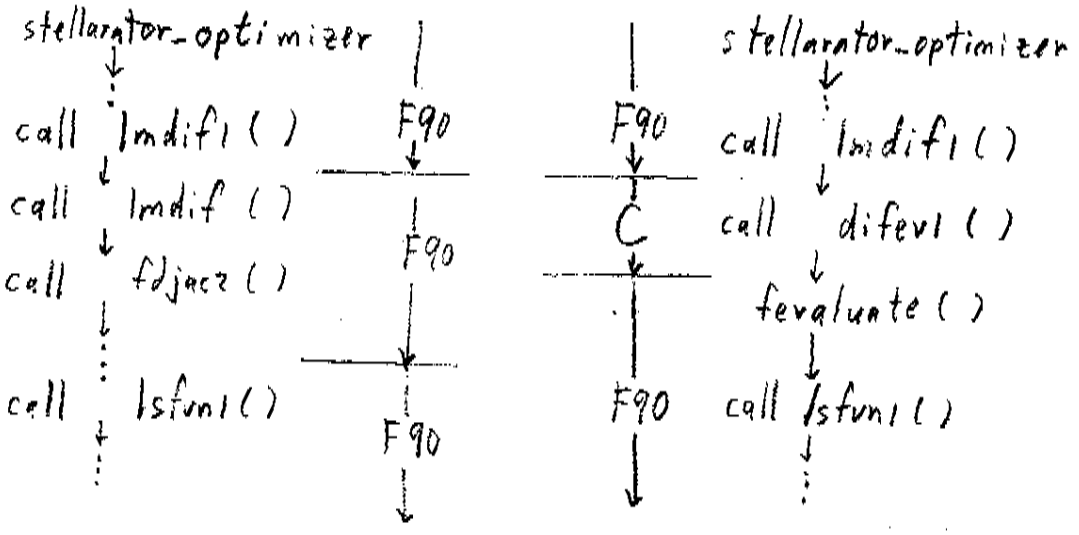
$f_i(\underline{x})$	ANM				ASA			DE1				DE2 (F=1)			
i	T	TF	NV	nfe	TRS	TAS	nfe	NP	F	CR	nfe	NP	λ	CR	nfe
1	0	n.a.	1	95	$1\cdot10^{-5}$	10	397	10	0.5	0.3	490	6	0.95	0.5	392
2	0	n.a.	1	106	$1\cdot10^{-5}$	10000	11275	6	0.95	0.5	746	6	0.95	0.5	615
3	300	0.99	20	90258	$1\cdot10^{-7}$	100	354	10	0.8	0.3	915	20	0.95	0.2	1300
4	300	0.98	30	-	$1\cdot10^{-5}$	100	4812	10	0.75	0.5	2378	10	0.95	0.2	2873*
5	3000	0.995	50	-	$1\cdot10^{-5}$	100	1379	15	0.9	0.3	735	20	0.95	0.2	826
6	$5\cdot10^6$	0.995	100	-	$1\cdot10^{-5}$	100	3581	10	0.4	0.2	834	10	0.9	0.2	1125
7	10	0.99	50	-	$1\cdot10^{-5}$	0.1	-	30	1.	0.3	22167	20	0.99	0.2	12804
8	5	0.95	5	2116	$1\cdot10^{-6}$	300	11864	10	0.8	0.5	1559	10	0.9	0.9	1076
9(k=4)	100	0.95	40	(391373)	$1\cdot10^{-6}$	1000	-	30	0.8	1	19434	30	0.6	1.0	14901
9(k=8)	$5\cdot10^4$	0.995	150	-	$1\cdot10^{-8}$	700	-	100	0.65	1	165680	80	0.6	1.0	254824

Table I: Averaged number of function evaluations (nfe) required for finding the global minimum. A hyphen indicates misconvergence and n.a. stands for "not applicable".

•DE characteristics:

- Provided $G = 0$ ensemble distributed over multiple wells, good at not getting trapped in local ones.
- Rule of thumb: $NP \simeq 7D$ provides adequate ensemble size.
- Once finds optimal local well, contracts rather quickly into it (hybrid of DE with hill-descending method unnecessary).
- For $D > 1$, with different scale-lengths in the different dimensions j, seems to contract first in dimensions with most opportunity for improvement, sequentially reducing the dimensionality of exploration.

•Rel'n of Stellopt-LM to Stellopt-DE:



•Estimate time needed for major global run:

$T = NP \times G_{max} \times T_1/N_{PE}$.

•Take $NP \simeq 7D$, $G_{max} \simeq 10D$

$\Rightarrow T \simeq 70D^2T_1/N_{PE}$.

•For $D \rightarrow 30, T_1 \rightarrow 10$ min,

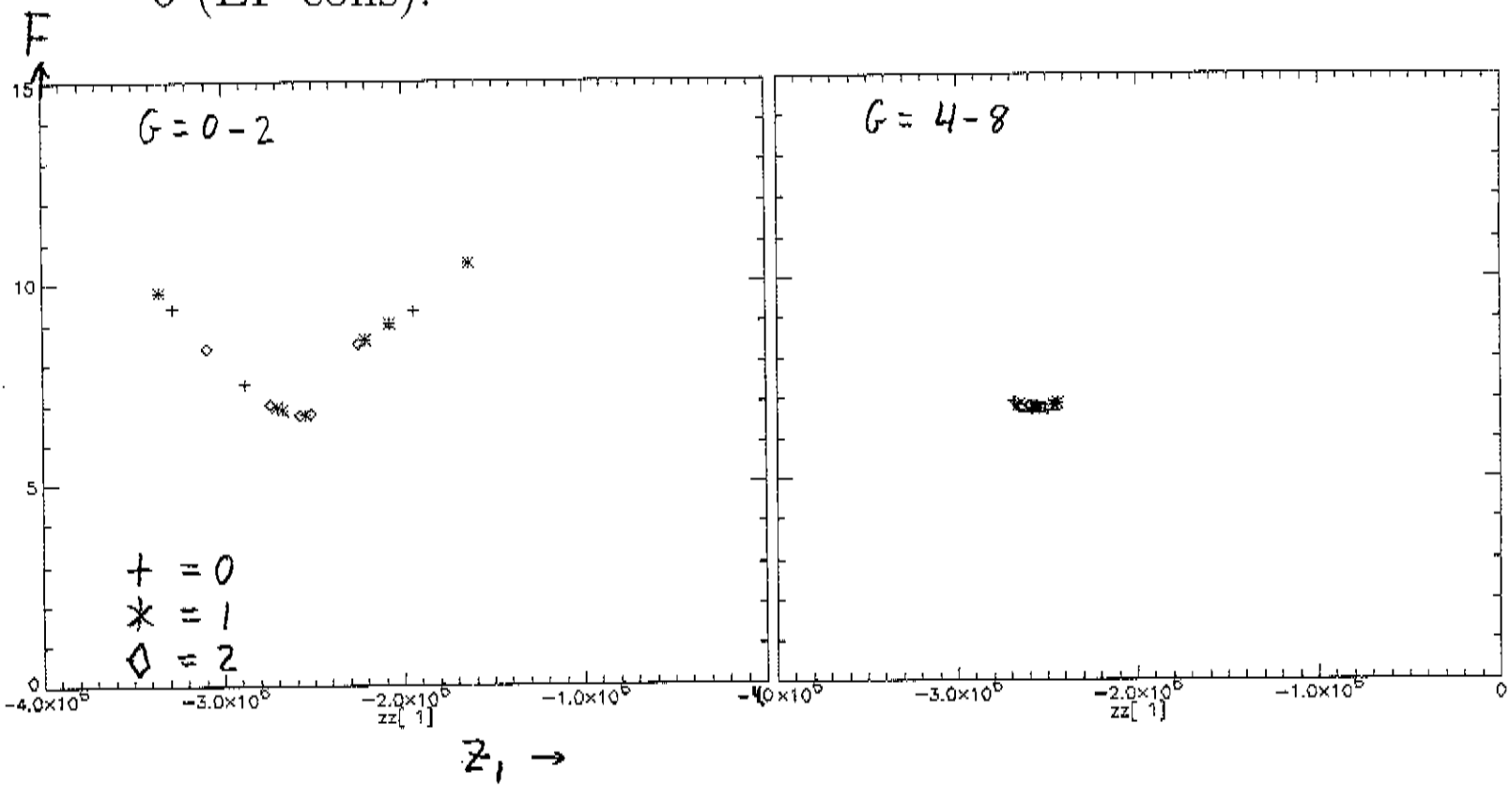
$\Rightarrow T \simeq 630 \times 10^3/N_{PE}$ min $\simeq 10^4/N_{PE}$ hr

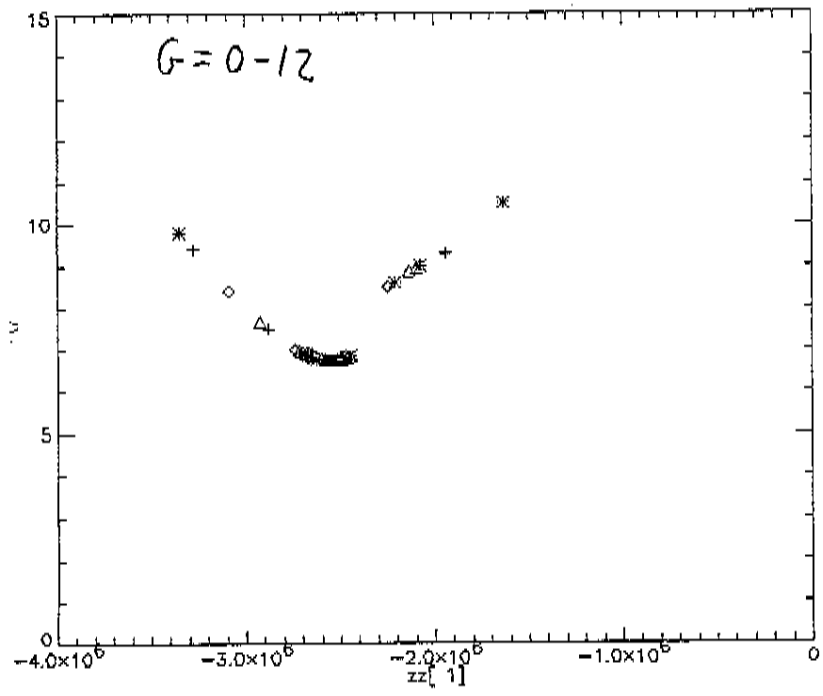
•Thus, having $N_{PE} \sim 100$ makes such calculations accessible.

•Experience thus far with Stellopt-DE:

[Parallelization now in final stages (S. Ethier). Presently testing with both the full lsfun1 and test functions to compute the fvec array for $F(\text{fvec}) = (\sum_{i=1}^{m} \text{fvec(i)}^2)^{1/2}$.]

1. $D = 1, NP = 7, G_{max} = 12$, lsfun1, free-bdy, perturbing about c82b (4064 coil set) with Z_6=coil-group 6 (EF coils).

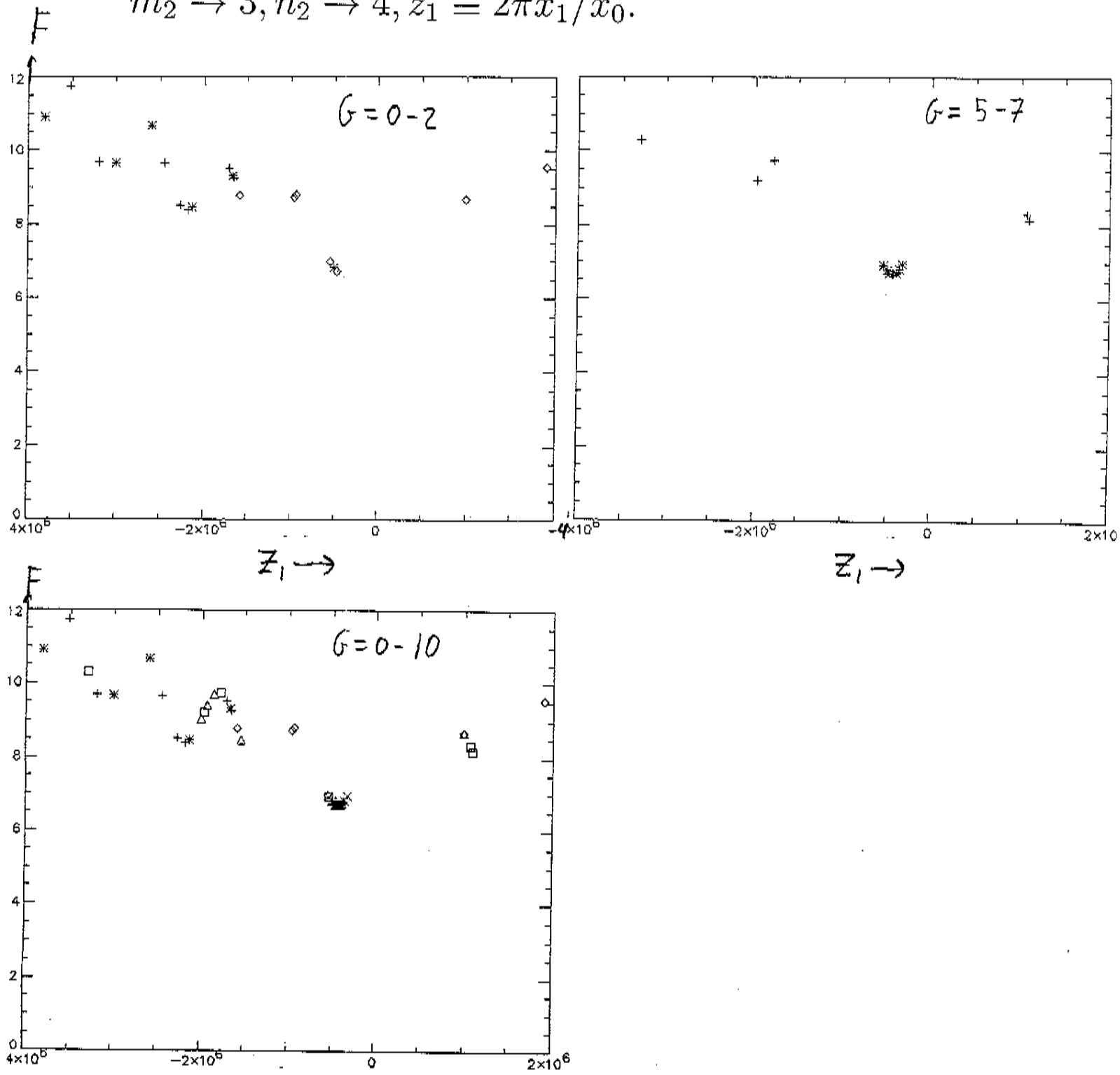




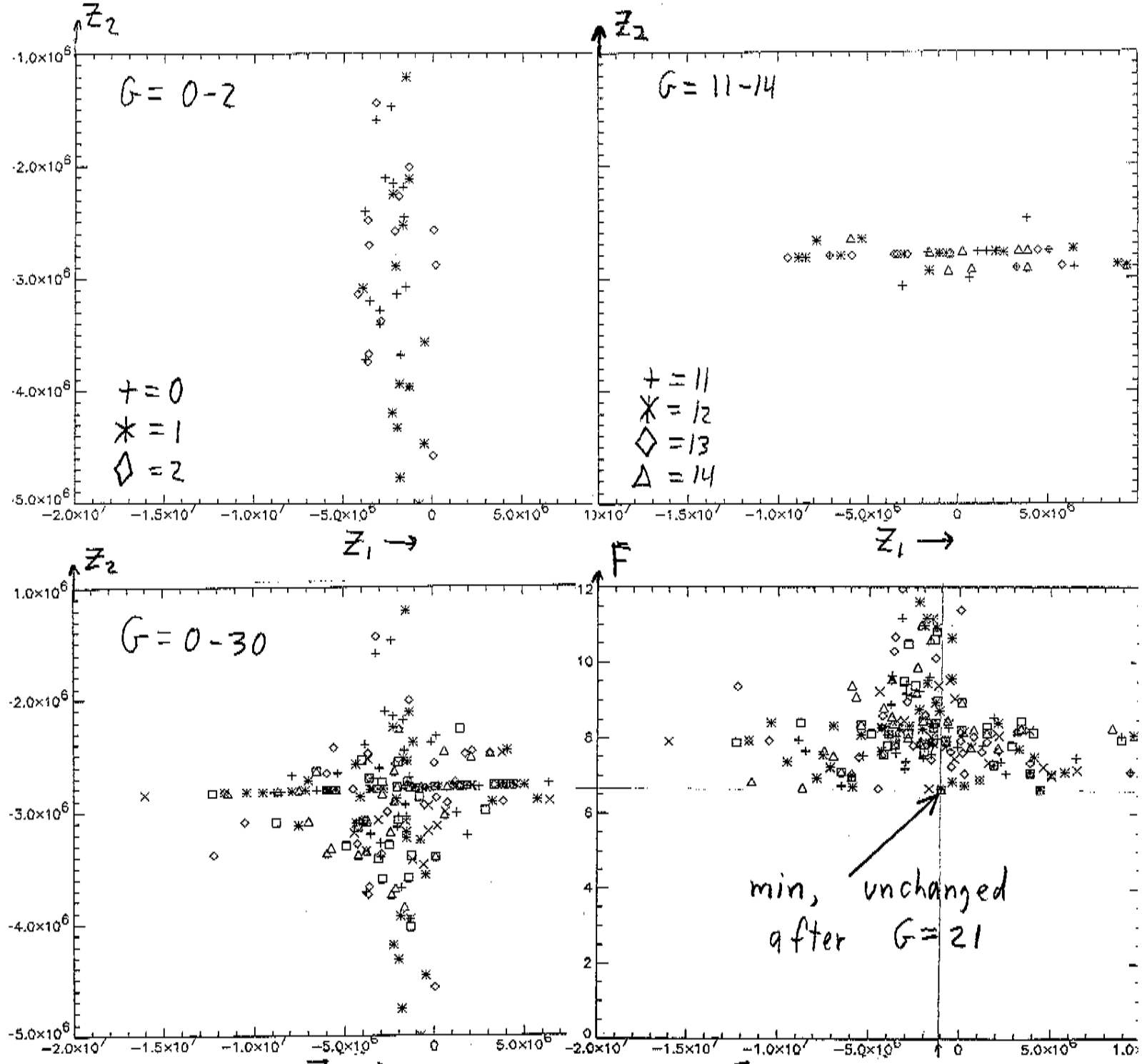
3. $D = 1, NP = 7, G_{max} = 10$, test_2_1D. Compare with behavior of LM-method.

∘Model fvec(i):

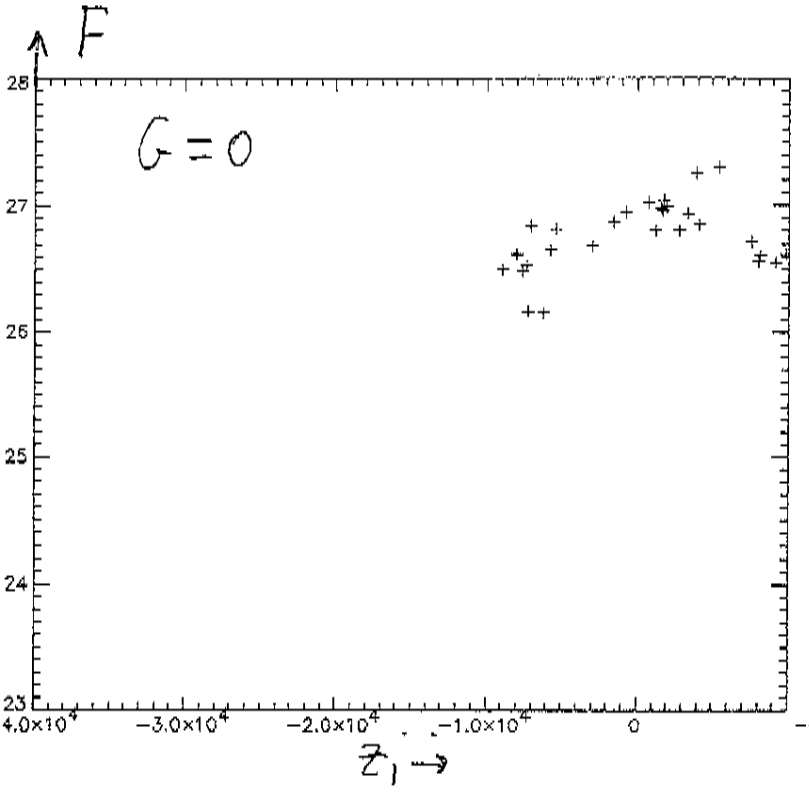
$F(\text{fvec}) \propto 1 - .2\cos(\iota z_1) + .1\cos[(m_2\iota - n_2)z_1]$, with $m_2 \to 3, n_2 \to 4, z_1 \equiv 2\pi x_1/x_0$.

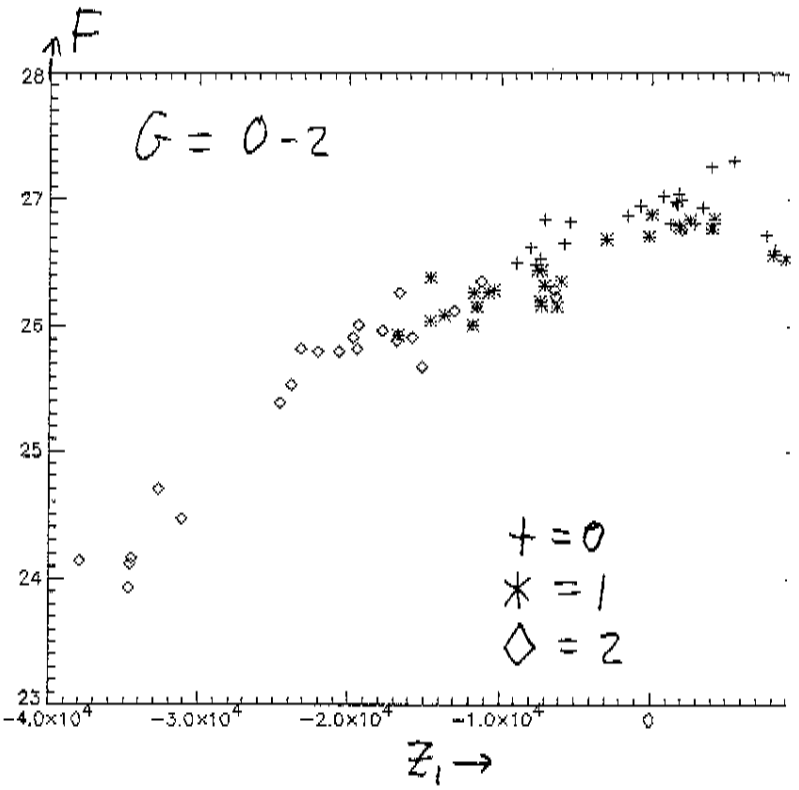


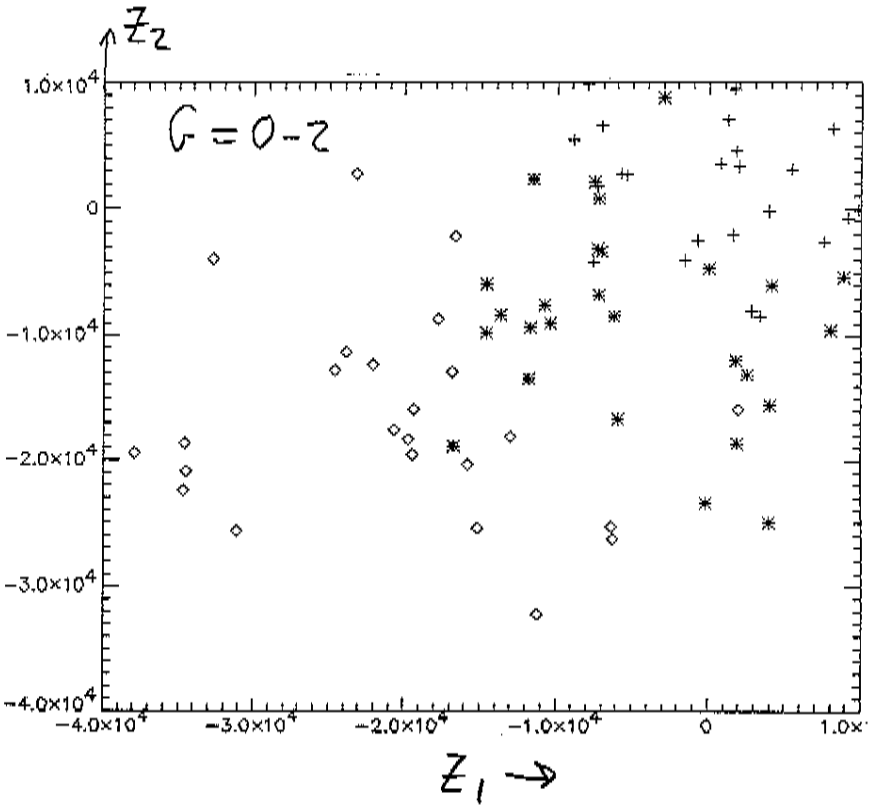
4. $D = 2, NP = 14, G_{max} = 30$, test_3_2D. Manifests dimensional contraction.
∘Model fvec(i):
$F(\text{fvec}) \propto 1 - .2\cos(z_2) + .1\cos[(m_2 z_2)\cos(n_2 z_1)]$, with $m_2 \to 3, n_2 \to 4, z_{1,2} \equiv 2\pi x_{1,2}/x_0$.



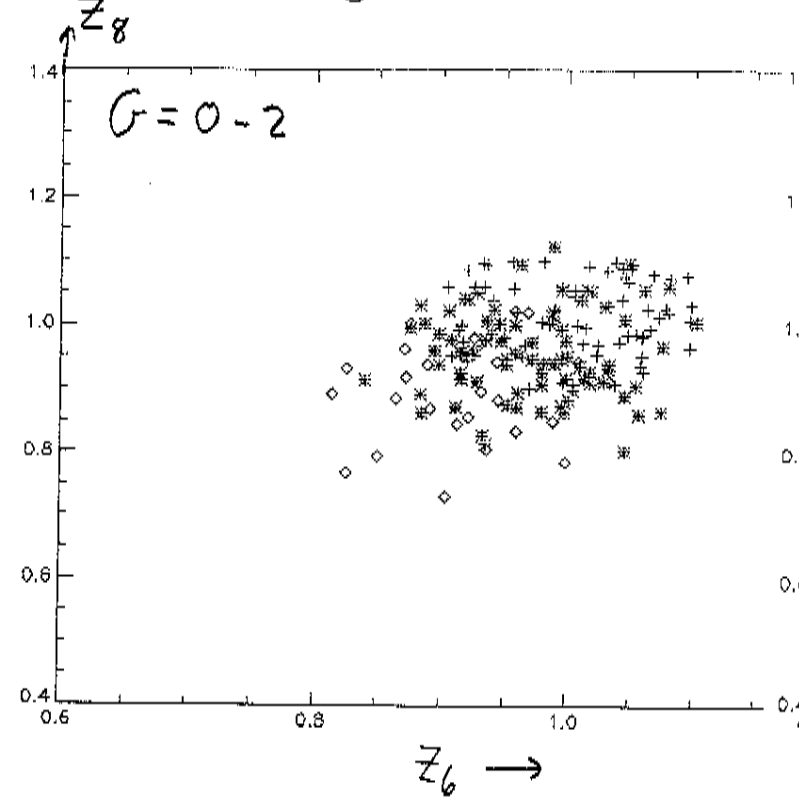
6. $D = 4, NP = 28, G_{max} = 2$, lsfun1, free-bdy run perturbing analytic multipole fields, where LM method fared poorly. After only $G = 2$ DE doing better than LM, and seems headed down a slope.

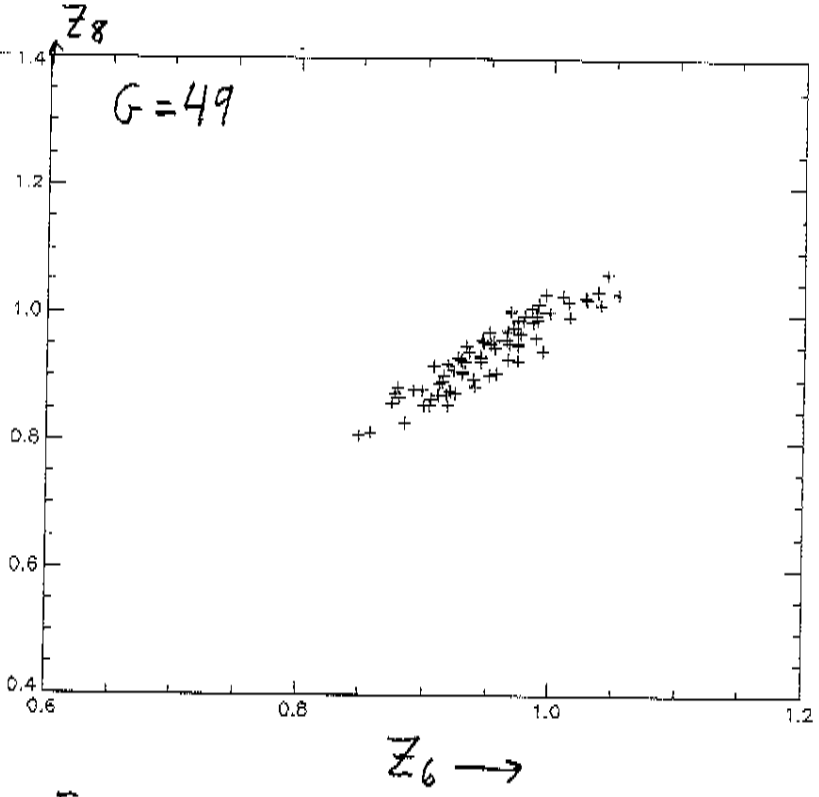


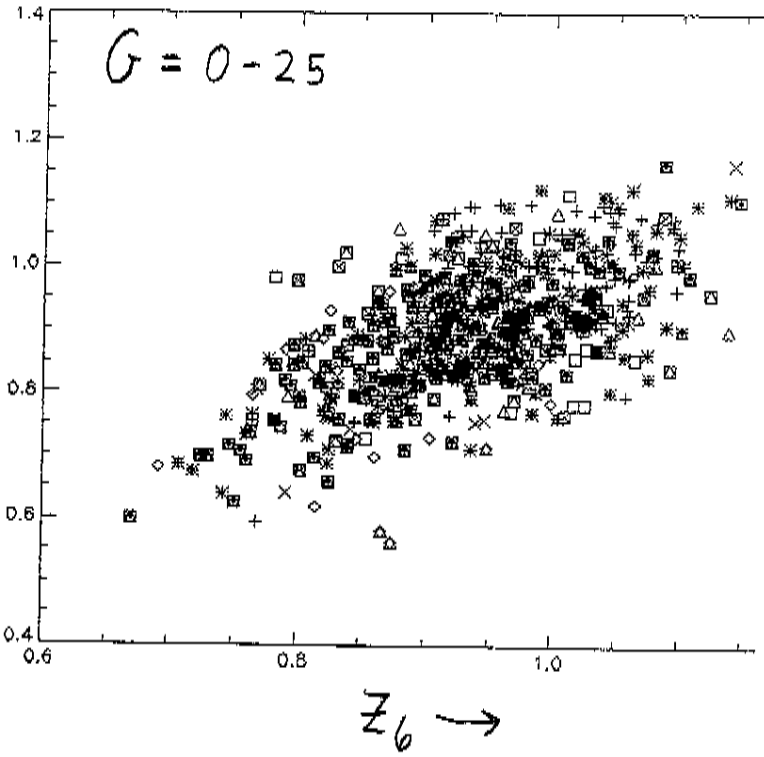


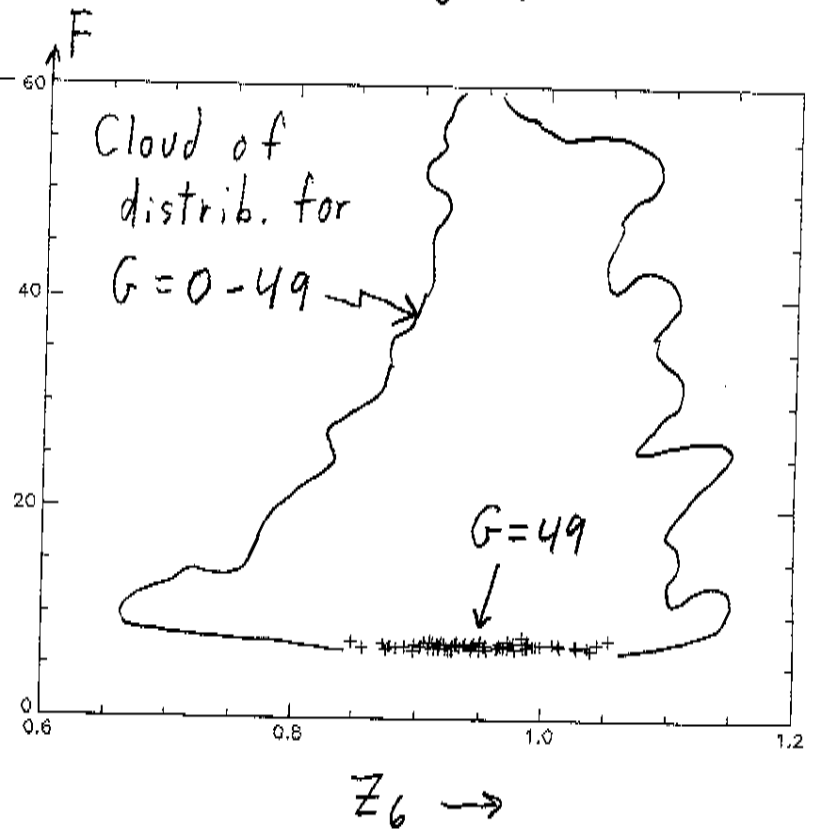


7. $D = 11, NP = 80, G_{max} = 49$, lsfun1, 1st large parallelized, free-bdy run for LI383 flexibility studies, showing dimensional contraction.









•Summary:

•The CM method has been used in exploring **Z**-space in the vicinity of some NCSX candidate configurations (c10,c82,PG1).

(I)Fixed-boundary (FxB) application (around c10,PG1), for study of configuration optimization:

•Have provided a 1st demonstration of the basic principle of the CM method, showing for c10 that the SVD-obtained perturbations $\boldsymbol{\xi}^i$ can in fact be used to vary the P_i independently.

•Have provided a 1st picture of the topography of the configuration space **X** or **Z** in which our searches for good stellarators are occurring. In an appreciable neighborhood of c10, the P_i may be modeled by a quadratic function of $\mathbf{z} = \mathbf{Z} - \mathbf{Z}_0$, and vary with little structure even over a scale comparable to the distance from c10 to PG1. We have constructed this quadratic representation about c10 for a reduced set ($N_z = 8$) of perpendicular displacements of the c10 boundary, computing both the CM G_{ij} and Hessian H_{ijk} for this set.

•The 4 different transport figures of merit produce

boundary displacements $\boldsymbol{\xi}^i(\theta,\zeta)$ similar in appearance. However, the G-matrix eigenvalues w_i show these are linearly independent, NOT nearly collinear.

•The $\boldsymbol{\xi}^{i=5}$ for kink stability differs in appearance from those for transport. For c10, $\boldsymbol{\xi}^5$ provides outboard indentation earlier seen to stabilize the kink, enhancing c10's negative triangularity at $N_{fp}\zeta = \pi$, while for PG1, $\boldsymbol{\xi}^5$ enhances its *positive* triangularity, consistent with tokamak intuition on kink stabilization.

•Have reduced the dimensionality N_z of the search space from $N_x = 78$ to $N_z = 8$ of the reduced model by 2 means:
(a)Removing the redundancy in the **X**-specification.
(b)Taking the perturbations most effective in varying P_i's of interest.

(II)Free-boundary (FrB) application (with 4064_rev6 coil set for c82), for study of operational flexibility.

•FrB sweeps permit much larger displacements ($\sim$ 10 cm) than in FxB studies ($\sim$ 1 cm) without causing pathological boundaries. These larger sweeps have revealed substantially more structure in the P_i topography (well beyond quadratic dependence).

•A large number of coil groups in the saddle coil set need not translate into a great deal of operational flexibility, because the coils all have about the same helicity. In particular, both the saddle and EF coil groups affect kink stability (P_5) about as in the FxB studies, but have far less effect on QA-ness (P_{1-4}). A coil set with low-m, larger-n, such as the TF set might provide, should reinstate this sensitivity.

•The EF current sweep found a new configuration (c82c_6.4), not yet optimized, which has much better kink stability and comparable QA-ness to c82.

(III)Study of optimizer performance.

•The LM-based Stellopt can often become trapped in small local wells, and not discover substantial improvements only a short distance away.

•We have implemented a more global DE-based Stellopt, from which initial results are promising.

References

[1]H.E. Mynick, N. Pomphrey, Princeton University Report PPPL-3434 (February, 2000) (to appear in *Phys. Plasmas*).

[2]M.Yu. Isaev, M.I. Mikhailov, D.A. Monticello, H.E. Mynick, A.A. Subbotin, L.P. Ku, A.H. Reiman, *Phys. Plasmas* **6** 3174 (1999).

[3]R. Storn, K. Price, U.C. Berkeley Technical Report TR-95-012, ICSI (March, 1995).

[4]G.H. Neilson, A.H. Reiman, M.C. Zarnstorff *et al.*, *Phys. Plasmas* **7**, 1911 (2000).

[5]P. Garabedian, L.P. Ku, *Phys. Plasmas* **6**, 645 (1999).