MEETING DISCUSSIONS

J.W. Bandler, Q.S. Cheng, M.A. Ismail, A.S. Mohamed,
F. Pedersen, J.E. Rayas-Sánchez and J. Søndergaard.

SOS-01-19-R

April 30 to May 18, 2001

People Present

Dr. John W. Bandler, Qingsha (Shasha) S. Cheng, Mostafa Ismail, Ahmed Mohamed, Frank Pedersen (until May 9), José Rayas-Sánchez and Jacob Søndergaard.

First Day (Monday April 30)

1. Notation

Symbol definition	McMaster University	Technical University of Denmark
Fine model point	$\boldsymbol{x}_f$	$\boldsymbol{x}$
Coarse model point	$\boldsymbol{x}_c$	$\boldsymbol{z}$
Fine model response	$\boldsymbol{R}_f$	$\boldsymbol{r}^{(f)}$
Coarse model response	$\boldsymbol{R}_c$	$\boldsymbol{r}^{(c)}$
Scalar objective function	U	H
Independent parameter	w	t
Residual between fine model response and specification	$\boldsymbol{R}_f - \boldsymbol{S}$	$\boldsymbol{f}$
Residual between coarse model response and specification	$\boldsymbol{R}_c - \boldsymbol{S}$	$\boldsymbol{c}$
Mapping	$\boldsymbol{P}$	$\boldsymbol{p}$

2. SM Based Neuromodeling (José)

Jose talked about exploiting Neural Networks (NN) in SM optimization. He gave a review of the previous work done using NN in device modeling. Then, he introduced his approach and verified it using some examples.

1

3. Improvement of the SM Optimization (Frank)

   Frank first presented his notation and defined the combined model problem. He introduced his new approach to update both the transition parameter and the trust region radius parameter. He verified his approach by doing a numerical experiment comparing the conventional updating strategy and his optimal updating strategy. He also presented a comparison of the different effects, like exact gradients, optimal parameters, semi and hard switch to direct optimization, Broyden update approximation, and his new approach.

Second Day (Tuesday May 1)

1. Improving on the SM Optimization, continued (Frank)

   Frank continued presenting his new approach. He introduced his suggested algorithm.

2. Expanded SM Design Framework Exploiting Preassigned Parameters (Mostafa)

   Mostafa introduced a new method for coarse model calibration. He used the non-optimizable parameters (preassigned parameters) to match the coarse model with the fine model. Then he established a mapping between some of the optimizable parameters and the preassigned parameters. Once this mapping is obtained it can be used in either optimization or yield analysis.

Third Day (Wednesday May 2)

1. SMX-A Novel Object-oriented Optimization (Shasha)

   Shasha spoke about SMX system as a general system to be used to exploit the SM concept in modeling and optimization. There are some comments about the formulation of the surrogate model, combined model.

2. SM from a mathematical viewpoint [1] (Jacob)

   Jacob presented the combined surrogate model approach between the mapped coarse model and linearized fine model. He discussed the usefulness of SM as a preprocessing technique. He also

discussed the necessary switch to a locally convergent optimization method in the final stages. He considered two examples use the suggested model and we had a discussion about the geometric meaning of both.

Fourth Day (Thursday May 3)

1. SM from a mathematical viewpoint, continued (Jacob)

   Jacob continued presenting the idea about the combined surrogate model and continued the discussion of the two examples.

2. Applications of SM Optimization Technology to Filter Design (Dr. Bandler)

   Dr. Bandler presented some selected milestones in CAD of microwave filters. He also presented the SM optimization exploiting surrogate model, SMX system, and the Expanded SM design using preassigned parameters.

Discussion

We had a discussion about the formulation of the surrogate model and Mostafa suggested the following algorithm

At the *i*th iteration

$$R_s(x_f, \omega) = \lambda R_c(x_c, \omega_c) + (1-\lambda)L(x_f, \omega) \tag{1}$$

$$\begin{bmatrix} x_c \\ \omega_c \end{bmatrix} = \begin{bmatrix} B & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} \Delta x_f \\ \omega \end{bmatrix} + \begin{bmatrix} x_c^{(i)} \\ \gamma \end{bmatrix} \tag{2}$$

$$L(x_f, \omega) = R_f(x_f^{(i)}, \omega) + J_f \Delta x_f \tag{3}$$

$$\Delta x_f = x_f - x_f^{(i)} \tag{4}$$

$$h^{(i)} = \arg\left\{ \min_{h} \ U(R_s(x_f^{(i)} + h)) \right\} \ , \|h\| \le \delta^{(i)} \tag{5}$$

$$\tilde{x}_f^{(i+1)} = x_f^{(i)} + h^{(i)} \tag{6}$$

3

$$x_f^{(i+1)} = \begin{cases} \tilde{x}_f^{(i+1)} & \text{if } U(R_f(\tilde{x}_f^{(i+1)})) < U(R_f(x_f^{(i)})) \\ x_f^{(i)} & \text{otherwise} \end{cases} \tag{7}$$

$$\left[\sigma^{(i+1)}, x_c^{(i+1)}, \gamma^{(i+1)}\right] = \arg\left\{ \min_{\sigma, x_c, \gamma} \; \left\| [e_1, e_2 \cdots e_{N_f}]^T \; \right\| \right\} \tag{8}$$

where $e_k^{(i+1)} = R_f(\tilde{x}_f^{(i+1)}, \omega_k) - R_c(x_c, \sigma\omega_k + \gamma)$, $k = 1...N_f$ \hfill (9)

update B, δ and J_f

update $B^{(i)}$

$$B^{(i+1)} = B^{(i)} + \frac{x_c^{(i+1)} - x_c^{(i)} - B^{(i)}h^{(i)}}{h^{(i)T}h^{(i)}} h^{(i)T} \tag{10}$$

$$h^{(i)} = \tilde{x}_f^{(i+1)} - x_f^{(i)}$$

update $\delta^{(i)}$

$$\delta_{i+1} = \begin{cases} 2\delta_i & \text{if } r > 0.75 \\ \delta_i/3 & \text{if } r < 0.1 \\ \delta_i & \text{otherwise} \end{cases} \tag{11}$$

$$r = \frac{U(R_f(x_f^{(i)})) - U(R_f(\tilde{x}^{(i+1)}))}{U(R_s(x_f^{(i)})) - U(R_s(\tilde{x}^{(i+1)}))}$$

update $J_f^{(i)}$

$$J_f^{(i+1)} = J_f^{(i)} + \frac{R_f^{(i+1)} - R_f^{(i)} - J_f^{(i)}h^{(i)}}{h^{(i)T}h^{(i)}} h^{(i)T} \tag{12}$$

$$h^{(i)} = \tilde{x}_f^{(i+1)} - x_f^{(i)}$$

update $\lambda^{(i)}$

$$\lambda^{(i+1)} = \begin{cases} 1 & \left\| \boldsymbol{E}_l^{(i)} \right\| > 2 \left\| \boldsymbol{E}_m^{(i)} \right\| \\ \dfrac{\left\| \boldsymbol{E}_l^{(i)} \right\|}{\left\| \boldsymbol{E}_l^{(i)} \right\| + \left\| \boldsymbol{E}_m^{(i)} \right\|} & \text{otherwise} \end{cases} \tag{13}$$

$$\boldsymbol{E}_m^{(i)} = \boldsymbol{R}_m^{(i)}(\boldsymbol{x}_f^{(i)} + \boldsymbol{h}^{(i)}) - \boldsymbol{R}_f(\boldsymbol{x}_f^{(i)} + \boldsymbol{h}^{(i)})$$
$$\boldsymbol{E}_l^{(i)} = \boldsymbol{R}_f(\boldsymbol{x}_f^{(i)}) + \boldsymbol{J}_f^{(i)}\boldsymbol{h}^{(i)} - \boldsymbol{R}_f(\boldsymbol{x}_f^{(i)} + \boldsymbol{h}^{(i)})$$

Comment Jacob suggested substituting (13) with an updating formula ensuring that $\lambda^{(i)} \to 0$ for $i \to \infty$.

Fifth Day (Friday May 4)

1. SM Technology with Applications in EM-Based Device Modeling and Statistical Design (Dr. Bandler and José)

   Dr. Bandler presented some concepts used SM approach like generalized SM tableau and Neural SM optimization. Jose explained his work about using NN in SM optimization. He also discussed the statistical analysis and yield optimization using neuromodels.

Sixth Day (Monday May 7)

1. Neural Inverse Space Mapping EM-Optimization (José)

   José introduced his new technique to obtain the inverse of the mapping using NN. First, he reviewed the other approaches used in microwave design, and then he presented his NISM approach for optimization. He used in his approach a statistical parameter extraction (PE) procedure. He also found that it is useful to obtain the extracted coarse model point by matching all responses available like S_{11}, and S_{21}. He applied his approach on some practical examples. He suggested that a linear inverse mapping might be sufficient in some problems therefore, it can be implemented without NN.

2. Some new concepts about SM (Jacob)

Jacob presented this equation to be solved

$$\min_{x} \frac{1}{2}\|f(x)\|_2^2 \tag{14}$$

The space-mapped solution can be obtained from the next formulation

$$\min_{x} \frac{1}{2}\|C(p(x))\|_2^2 \tag{15}$$

A sequence of points $\{x_k\}$ can be generated such that

$$\{x_k\} \quad : \quad f(x_{k+1}) - f(x_k) = C(p(x_{k+1})) + \varepsilon(x_{k+1}) - C(p(x_k)) - \varepsilon(x_k) \tag{16}$$

The fine model can be approximated through input and output mappings as follows

$$f(x) = g(C(p(x))) \tag{17}$$

taking the derivatives with respect to the fine model point x

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial g}{\partial C} \frac{\partial C}{\partial z} \frac{\partial z}{\partial x} \\ &\approx A_k \frac{\partial C}{\partial z} B_k \\ \text{where } & \frac{\partial C}{\partial z} = J_c \\ & J_c : \text{Jacobian of the coarse model} \end{aligned} \tag{18}$$

the fine model can be approximated by the following

$$f(x) \approx f(x_k) + A_k [C(p_k(x)) - C(p_k(x_k))] \quad ; A_k \in \Re^{m \times m} \tag{19}$$

A_k is updated through Broyden's formula. A_k can be stored efficiently using a limited memory update approach [2]. A_k could be also a diagonal matrix, so that the corresponding responses are scaled.

Different strategies to deal with the output mapping g are suggested:

6

a) Keep $A_k = I$ and apply regular space mapping until a number of uphill steps are found. Then

start updating A_k, which can be done while the mapped coarse model is kept frozen, or while the

mapped coarse model is also updated. (keep B_k constant or update it using Broyden's formula).

b) Start updating A_k and B_k from the beginning.

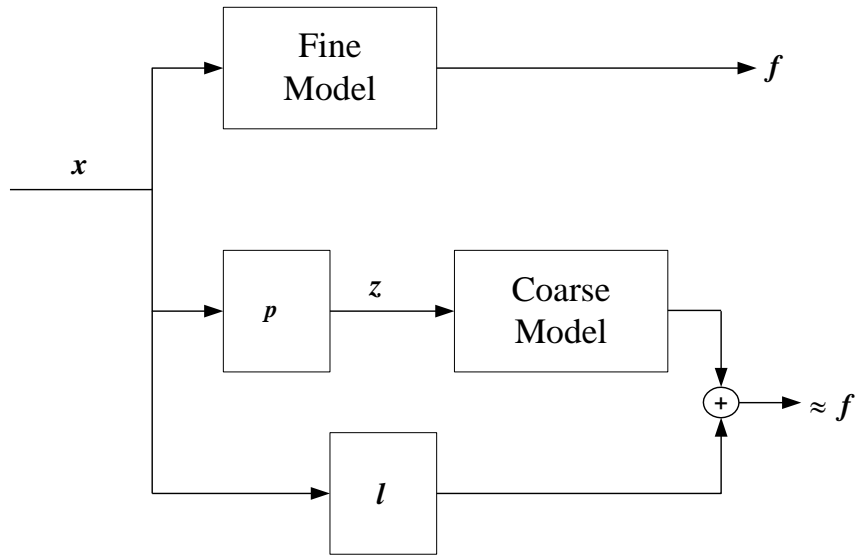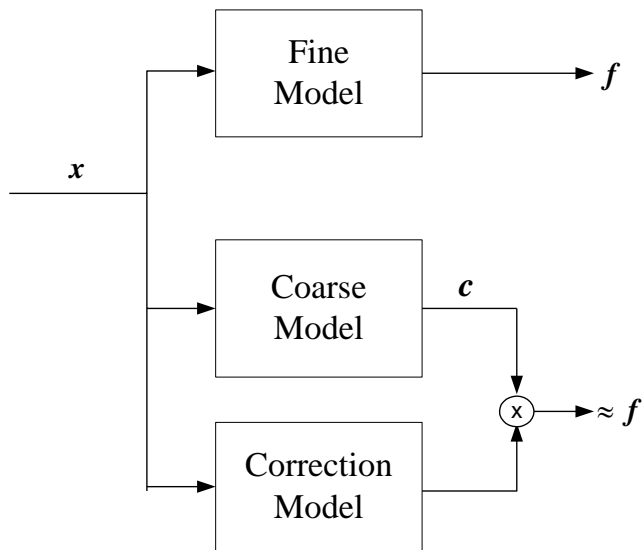José suggested some useful diagrams as shown below

Regular SM
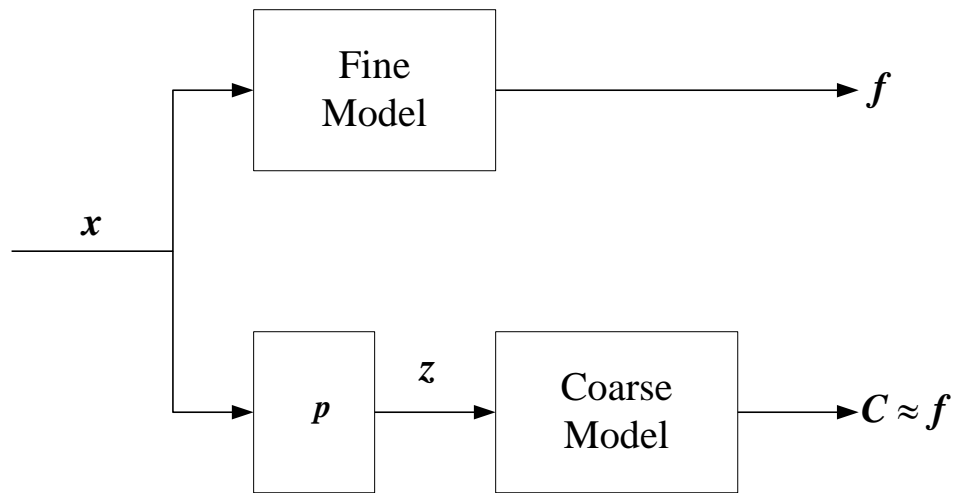
Input / output SM
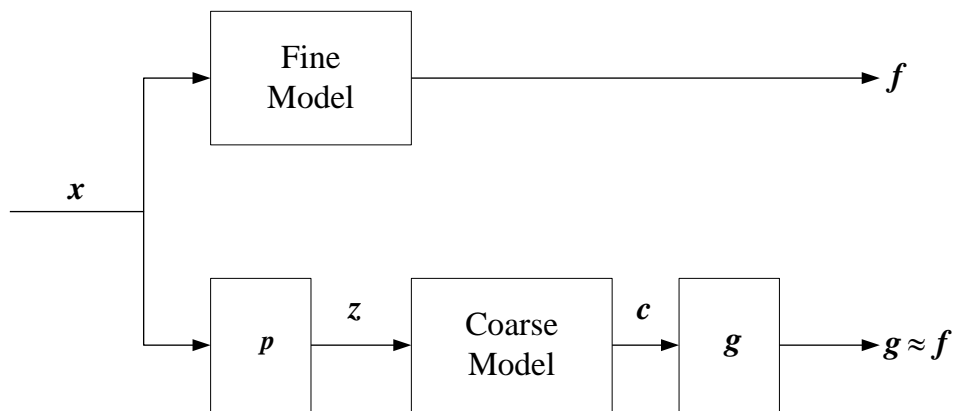
SM Surrogate "combined" model

SM Multiplicative Model

The idea of input/output space mapping was discussed after the SMSMEO workshop in Denmark in November 2000. We also discussed under what conditions the input mapping would be more effective than the output mapping from a mathematical point of view. José suggested that the mathematical formulation of this comparison, assuming the same complexity of p and g, could be described as

$$\left\| e_1^{iT} \quad e_2^{iT} \ldots e_N^{iT} \right\| \overset{?}{<} \left\| e_1^{oT} \quad e_2^{oT} \ldots e_N^{oT} \right\|$$

where

$$e_k^i = f(x_k) - c(p(x_k))$$

$$e_k^o = f(x_k) - g(c(x_k))$$

(20)

We agreed that some conditions should be defined about the relationship between the fine and coarse models.

Seventh Day (Wednesday May 9)

1. Illustration of Statistical PE procedure (José)

    José presented the idea of statistical PE by applying it on the 6th order BPF LC circuit. The "fine" model circuit is slightly different than the "coarse" model circuit by adding some series inductor and parallel capacitor in certain places. He used only one optimizable variable and plotted the objective function versus the design variable. He applied the statistical PE procedure on this circuit and showed the results graphically.

    Corollary Dr. Bandler suggested decomposing the circuits of both "coarse" and "fine" models. Then try to match the individual parts of both models.

2. STOGO, stochastic global optimization [3, 4] (Jacob)

    Jacob presented some ideas about using a stochastic approach for global optimization for solving the PE problem. This technique is discussed in more details in [3, 4].

9

3. Different approach for solving PE problem (Jacob)

    Jacob first define the set of points that minimizing the difference between the fine model and coarse

    model as follows

    $$S(\boldsymbol{x}) = \left\{ z \middle| z = \arg\min_{z} \|\boldsymbol{c}(z) - \boldsymbol{f}(\boldsymbol{x})\|_{\infty} \right\} \tag{21}$$

    the set of points that match the two models can be defined as

    $$S_e(\boldsymbol{x}) = \left\{ z \middle| \boldsymbol{c}(z) = \boldsymbol{f}(\boldsymbol{x}) \right\} \tag{22}$$

    which is a subset of the original one

    $$S_e(\boldsymbol{x}) \subseteq S(\boldsymbol{x}) \tag{23}$$

    Then, Jacob compares between the original PE problem

    $$p(\boldsymbol{x}) = \arg\min_{z} \|\boldsymbol{c}(z) - \boldsymbol{f}(\boldsymbol{x})\|_{\infty} \tag{24}$$

    and the proposed PE problem

    $$p(\boldsymbol{x}) = \arg\min_{z} \frac{1}{2} \|z - \boldsymbol{x}\|_2^2 \tag{25}$$
    $$s.t. \qquad z \in S(\boldsymbol{x})$$

4. Filter algorithm [5, 6]

    Jacob discussed an approach to solve (25). Using a filter algorithm developed by Fletcher [5, 6]. The

    algorithm contains two main steps Restoration step and Sequential Quadratic Programming (SQP)

    step [5, 6].

Eighth Day (Thursday May 10)

1. The seven-section TL impedance transformer (Ahmed)

    Ahmed presented the optimal coarse model values of the characteristic impedances of the seven-section TL impedance transformer. He made a comparison between the new results and the results obtained by M. Bakr for the same problem. The new optimal values give the exact equi-ripple response expected for the coarse model. The resulted coarse model response agrees with the analytical formulas that describe the seven-section TL transformer.

2. Interface to EM/ circuit simulators (Mostafa)

    Mostafa presented a suggested framework for interface to EM/ circuit simulators. The suggested flow diagram is shown.
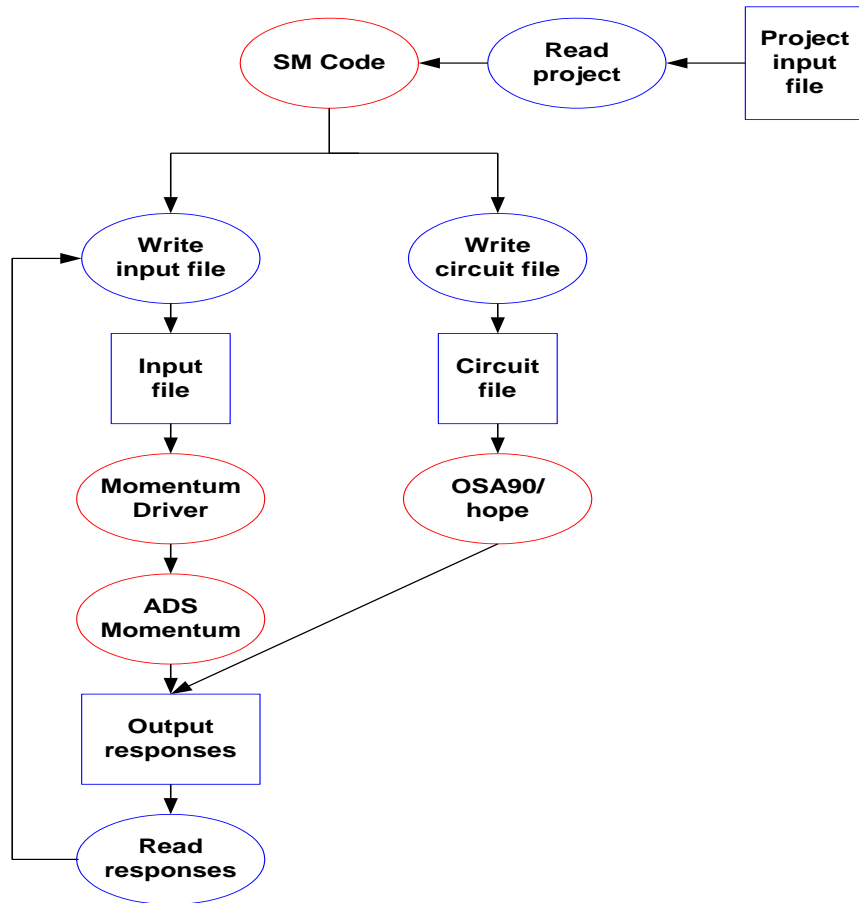
2. Drive OSA90 from MATLAB (José)

José illustrated a simple example for driving OSA90 from MATLAB. The MATLAB and macro files are attached.

Ninth Day (Monday May 14)

1. The capacitance effect of seven-section TL impedance transformer (Ahmed)

Ahmed illustrated the effect of increasing the capacitance values of the fine model of seven-section TL impedance transformer. As the capacitance values increase as the capacitively loaded TL fine model has a bad response and violates the specification. We agreed to keep the capacitance values at a lower level.

Tenth Day (Tuesday May 15)

1. SMX-A Novel Object-oriented Optimization (Shasha)

Shasha presented his last modified version of MTT conference posters about SMX system. He showed the final design architecture and some examples.

Comment Dr. Bandler mentioned a very useful and interesting comment about the idea behind the surrogate model. The surrogate model is a convex combination of both the space-mapped coarse model and the linearized fine model. In general, both of these models exist. In some times, when it is required to obtain a high accuracy of the final solution, the algorithm switches to the linearized fine model, and keep the space-mapped coarse model in the background. The space-mapped coarse model always exists and it is still useful when large steps are needed. One other interesting point that the value of the combination factor could be a function of the distance from the point of linearization. (See (26) and (27)).

Eleventh Day (Wednesday May 16)

1. Results update (Jacob)

   Jacob showed the results of his program using the new values of characteristic impedances. Dr. Bandler suggested that it could be useful to use the real and imaginary part of S_{11} in the PE process to match the two responses of fine and coarse models. We had a debate of using the real and imaginary part of S_{11} in both PE and optimization of the fine and coarse models.

2. Future research (Jacob)

   a. PE techniques, when to use what….

   b. Gather and categorize models from previous publications.

   c. Improve surrogate, e.g., $g(c_k(\boldsymbol{x}))$ or having $\omega_k(\boldsymbol{x})$

   $$S_k(x) = \omega_k(\boldsymbol{x}) c_k(\boldsymbol{x}) + (1 - \omega_k(\boldsymbol{x})) l_k(\boldsymbol{x}) \tag{26}$$

   where $c_k(\boldsymbol{x}) = c(p_k(\boldsymbol{x}))$

   $$
   \begin{aligned}
   \omega_k(\boldsymbol{x}) \to 1 \quad & \text{for} \quad \frac{M}{\|\boldsymbol{x} - \boldsymbol{x}_k\|} \to 0 \quad ; M : \text{is constant to be defined} \\
   \omega_k(\boldsymbol{x}) \to 0 \quad & \text{for} \quad \|\boldsymbol{x} - \boldsymbol{x}_k\| \to 0
   \end{aligned}
   \tag{27}
   $$

3. Direct Optimization (Mostafa)

   Mostafa made a comparison between the way the SMX system calculates gradients and the way that conventional optimization methods are used….

   a. SMX system uses Broyden update.

   b. Conventional optimization techniques use finite difference and interpolation.

4. Nested Mapping (Jacob)

   Dr. Bandler suggested the use of mappings of mappings. The idea is to enhance a coarse model using a mapping, then use this in space mapping optimization, or refine the enhanced coarse model further by a new mapping of the enhanced coarse model. Jacob showed how this approach could be done by nested mappings, hence

$$c_0 = c(\boldsymbol{I}\,\boldsymbol{x})$$
$$c_1 = c_0(p_1(\boldsymbol{x}))$$
$$c_2 = c_1(p_2(\boldsymbol{x})) \tag{28}$$
$$\vdots$$
$$c_k = c(\boldsymbol{I}\,p_1(p_2\cdots(p_k(\boldsymbol{x})))\cdots)$$

Twelfth Day (Friday May 18)

Mostafa and José made their final rehearsal for their presentations in MTT.

References

[1] M.H. Bakr, J.W. Bandler, K. Madsen and J. Søndergaard, "Space mapping from a mathematical viewpoint," *Optimization and Engineering*, 2001, submitted.

[2] R. Byrd, J. Nocedal, and R. Schnabel, "Representations of quasi-Newton matrices and their use in limited memory methods," *Math. Prog.*, vol. 63, 1994, pp. 129-156.

[3] K. Madsen and S. Zertchaninov, "A new branch-and-bound method for global optimization," *IMM-Report-1998-05*, 1998.

[4] S. Zertchaninov and K. Madsen, "A C++ program for global optimization," *IMM-Report-1998-04*, 1998.

[5] R. Fletcher, S. Leyffer and Ph.L. Toint, "On the global convergence of a filter-SQP algorithm," *Dundee Numerical Analysis Report NA/197*, 2000.
http://www.maths.dundee.ac.uk/~ftp/na-reports/NA197_RF_SL.ps.Z

[6] R. Fletcher, N. Gould, S. Leyffer and Ph.L. Toint, "On the global convergence of trust-region SQP-filter algorithms for general nonlinear programming," *Report 99/03*, Dept. of Mathematics, FUNDP, Namur (B), 1999.
ftp://thales.math.fundp.ac.be/pub/reports/TR99-03.ps