

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

EFEITOS DA VARIAÇÃO DA INTENSIDADE DE APRENDIZADO
LATERAL NO MODELO DE REDE NEURAL MULTICOLUNAR DO *TRION*

LUCAS FÜRSTENAU DE OLIVEIRA

Orientador

Prof. JORGE ALBERTO QUILLFELDT

Dissertação apresentada ao Departamento de Biofísica do Instituto de
Bióciências da Universidade Federal do Rio Grande do Sul para a obtenção do
título de Bacharel em Ciências Biológicas - Ênfase em Biofísica.

Porto Alegre, 1997.



Agradeço ...

ao Jorge, meu orientador, por muito ter me ensinado e por ter apostado em mim.

à minha mãe, por julgar o estudo algo fundamental.

à minha namorada, por ser tudo o que ela é.

aos amigos, simplesmente por serem amigos.

“Às vezes, é importante que a ciência saiba como esquecer-se das coisas de que tem mais certeza.”

Jean Rostand (1894 - 1977), biólogo francês

Abstract

VARIATIONS ON THE LATERAL LEARNING INTENSITY AND ITS EFFECTS ON THE *TRION* MULTICOLUMNAR NEURAL NET MODEL

The *Trion* structured multicolumnar neural net model is characterized by having two types of connections, internal and lateral, between its basic units, the *trions*, and by the fact that its states evolve synchronously. Each *trion* represents an idealized group of neurons, and a set of *trions* represents an idealized cortical column. Previous results have shown that each set of connections between the *trions* takes to a fixed repertoire of cyclic firing patterns (MP's). The evolution of each *trion* inside the MP's is influenced by its two previous firing states and also by its neighbors' two previous states. The model also allows learning by means of a hebbian algorithm, that effects small changes in the connections. In this work, we tried to verify the influence of various lateral learning intensities in the Cycle Probability (CP, the robustness of an MP) of different MP's with also different sets of connections.

The results obtained showed that increases in lateral learning intensity, with some exceptions, take to an increase in the CP of an MP, followed by stabilization and a plateau. Many MP's had a CP nearing 1.0. Also were observed scapes with multiple minimums and maximums for the CP depending on the ϵ lat's values until the plateau was reached. In a general manner, the CP for the out of phase MP's were higher than the one for the in phase MP's.

Índice

1. Introdução	1
1.1 Histórico do modelo do <i>Trion</i>	7
1.2 Princípios básicos incorporados no modelo do <i>Trion</i>	8
1.2.1 A sinapse de Hebb	8
1.2.2 A organização cortical multicolumnar de Mountcastle	9
1.2.3 O selecionismo de Edelman	10
1.2.4 Flutuações estatísticas	10
1.2.5 Balanço entre excitação e inibição	10
1.3 O modelo do <i>Trion</i>	11
1.3.1 O <i>Trion</i>	11
1.3.2 Colunas de <i>Trions</i>	11
1.3.3 Probabilidades para o estado de cada <i>trion</i>	14
1.3.3.1 O fator de degeneração g	15
1.3.3.2 A temperatura e a variável B	15
1.3.3.3 O Campo local M_i	15
1.3.4 Determinação final do estado	16
1.3.5 Padrões emergentes	16
1.3.5.1 MP's e estados iniciais da coluna	17
1.3.6 Probabilidade de Ciclagem dos MP's	18
1.3.7 Aprendizado Hebbiano	18
1.3.8 Multi-colunas	19
1.3.9 MP's em multi-colunas	20
1.3.10 Aprendizado lateral	21
1.4 Analogia com o sistema visual humano	22
2. Objetivo	24
3. Material e Métodos	25
3.1 A plataforma de trabalho	25
3.2 A programação	25
3.3 Parâmetros utilizados	26
3.3.1 Entrada dos MP's no computador	26
3.4 A simulação	27
3.5 Formatação dos resultados	27
4. Resultados	28
4.1 Observações gerais	28
4.2 As conexões	28
4.2.1 Conexão C1	28
4.2.2 Conexão C2	30
4.2.3 Conexão C3	30
4.2.4 Conexão C4	30
4.2.5 Conexão C5	31
4.3 Resumo dos resultados	32
4.3.1 Patamar	32
4.3.2 Probabilidade de Ciclagem máxima	32

4.3.2.1 Com relação ao sinal de ϵ_{lat}	32
4.3.2.2 Probabilidade de Ciclagem alta	33
4.3.2.3 Probabilidade de Ciclagem igual a 1.0	33
5. Resultados	35
6. Discussão	36
6.1 Efeitos do sinal da intensidade de aprendizado lateral ϵ_{lat}	36
6.2 Efeitos da fase do MP	38
6.3 Possível distinção entre as conexões C1 e C2 e as outras conexões	38
6.3.1 Auto-conexão <i>versus</i> conexões com vizinhos	39
6.3.2 Ausência de conexão forte	39
6.3.3 Competição excessiva favorecendo C1 e C2 com ϵ_{lat} inverso ?	40
6.3.4 PC's mais baixas para C1 e C2	40
6.4 Padrões semelhantes e seus resultados	41
6.5 O patamar	41
6.5.1 Possível desestruturação do modelo com aprendizado lateral muito intenso	42
6.6 PC máxima igual a 1.0	42
6.7 Utilidade do aprendizado lateral	43
6.8 Considerações finais	44
6.8.1 Para simulações futuras	44
6.8.2 Resultados mais significativos deste trabalhos	44
6.8.3 Modelos	44
7. Bibliografia	46
Lista de Figuras	51
Lista de Tabelas	52
Apêndice	53

1. Introdução

É impossível negar a importância que as máquinas possuem na história da sociedade humana moderna. Desde os primeiros modelos de motor a vapor de Savory e de Newcomen na virada do século XVI para XVII, passando pelos modelos aperfeiçoados por Watt, em 1769, fator primordial para a ocorrência da Revolução Industrial, as máquinas têm cada vez mais liberado o ser humano de tarefas que antes requeriam grande dispêndio de tempo e/ou energia. Permitiram também a realização de tarefas antes consideradas como do reino da magia (talvez seja esse um dos atrativos do cinema e dos efeitos especiais).

Um dos fatores principais para o desenvolvimento das primeiras máquinas, além da técnica de trabalho dos materiais necessários, entre outras coisas, foi emprego de mecanismos de controle dos processos executados pela máquina, invariavelmente baseados no princípio da retroalimentação. Um dos primeiros tipos funcionais de controlador foi o regulador, ou governador, centrífugo inventado por Watt, que era capaz de manter um motor a vapor girando sempre na mesma velocidade. Os mecanismos de controle e regulação, agora não mais identificados apenas em máquinas, mas também em seres vivos e outros sistemas, são campo de estudo da *cibernética*, termo cunhado por Norbert Wiener (1948).

O avanço da tecnologia desde a Revolução Industrial permitiu a construção de máquinas melhores, mais eficientes e complexas. Entretanto, para que tal fosse possível, foi necessário o desenvolvimento de mecanismos de controle também melhores, mais eficientes e complexos. Neste aspecto, a eletrônica foi importante, primeiro com a válvula, inventada por Lee De Forest em 1906, depois com o transistor em 1947 e, em 1964, com o circuito integrado, permitindo a construção dos primeiros computadores na década de 40. A eletrônica, através de computadores cada vez mais poderosos, é a responsável pela revolução nas comunicações que vem ocorrendo nos últimos anos. O desenvolvimento da Integração em Escala Muito Grande (VLSI), em meados dos anos 80, tem cada vez mais permitido o aumento no poder de processamento dos computadores.

Os computadores, entretanto, não são responsáveis apenas por mudanças nas formas de comunicação. Graças a sua grande versatilidade, entre outras tarefas, um computador pode ser programado para controlar os mais variados processos, de

maneira tão ou mais eficiente do que circuitos eletrônicos isolados. Programas cada vez mais sofisticados têm sido elaborados nos últimos anos com tal objetivo. Entretanto, à medida em que os processos controlados vão se tornando mais complexos, aumenta a dificuldade de se programar um computador para controlá-lo, ao ponto de que alguns processos só podem ser controlados por seres humanos. O computadores estão esbarrando neste limite de complexidade.

É normalmente dito que a principal característica que separa o homem dos outros animais é a sua inteligência. Apesar do ser humano ser risível em termos de diversidade biológica (é apenas uma espécie), estamos espalhados por todo o planeta (e não mais restritos a ele se levarmos em conta que fomos à Lua) e somos capazes de liquidar uma boa parcela da vida na Terra (nós, inclusive) com relativa facilidade (estimativas para 1990 eram de um total de 50000 ogivas nucleares nos arsenais de vários países). Mesmo fazendo-se ressalvas quanto à inteligência do homem quando se fala em armas nucleares, a capacidade de adaptação, aprendizado e comunicação que o cérebro humano possui certamente contribuiu para termos desenvolvido uma sociedade complexa e cosmopolita como a que temos atualmente. É essa capacidade que permite o controle de tarefas e resolução de problemas que ainda não são possíveis para computadores.

O conceito de variedade requerida de Ashby (1956) é bastante útil para entendermos as limitações dos computadores atuais como reguladores. Segundo esta lei, a Variedade (V_0) de um processo causada por um Desvio (D) das condições ideais só pode ser controlada por um Regulador (R). O controle adequado só ocorrerá se a variedade de R 's (V_r) existentes for compatível com a variedade de D 's (V_d) (Ashby, 1973). Muitos tipos de Desvios requerem muitos tipos de respostas do Regulador. Expressando logarithmicamente, $V_0 = V_d - V_r$. A diminuição no valor de V_0 para uma V_d constante só é possível com um aumento na V_r . Se analisarmos a questão como a transmissão de dados, D sendo uma perturbação ou um ruído, e R sendo o seu filtro, a possibilidade de uma transmissão correta é limitada pela capacidade de R como canal de comunicação (Ashby, 1973).

Nós, seres humanos, temos uma variedade de R 's enorme e somos capazes de aumentá-la, pois temos capacidade criativa. Temos então uma V_r bastante alta. Entretanto, somos limitados quanto a nossa velocidade, o que nos torna limitados como canais de comunicação. Os computadores atuais possuem pouca limitação

quanto a velocidade, mas o seu R é somente tão variado quanto for lhe programado. Se fossem capazes de aprender eficientemente, poderiam aumentar a sua V_r .

Mesmo antes da construção dos primeiros computadores, já havia teorias sobre como eles deveriam ser projetados e que tipo de tarefas seriam capazes de resolver. Em 1936, Alan Turing propôs o seu “computador universal”, baseado na teoria dos autômatos. Em 1943, McCulloch e Pitts, do MIT (Massachusetts Institute of Technology), desenvolveram a automação por redes neurais, um modelo com inspiração na estrutura do sistema nervoso. Em 1956, um grupo de pesquisadores, John McCarthy e Marvin Minsky do MIT entre eles, cunharam o termo inteligência artificial. Segundo McCarthy, seria “a ciência de fazer com que computadores realizem coisas que, se feitas por humanos, exigiriam inteligência” (Logsdon, D. & Logsdon, T., *The Software Toolworks Multimedia Encyclopedia*, ver. 1.5, 1992. Verbetes: “Automata, Theory of”). Uma visão mais atual seria a de que uma máquina com inteligência artificial deve “receber uma entrada de informações do meio, determinar uma ação ou resposta a ela e dar saída a esta resposta”.

A maneira como essas entradas de informação e saídas de respostas se daria inspirou pesquisas sobre a interpretação de padrões visuais, compreensão e utilização de linguagem natural e aprendizado. Vários tipos de algoritmos de computador têm sido desenvolvidos, mas nenhum consegue realizar tarefas com a mesma eficiência com que reconhecemos a letra “A” inequivocamente e com que entendemos a linguagem (há exceções para ambos os casos na espécie humana, infelizmente, resultado de uma variedade de deficiências cognitivas, entre elas, as afasias (Bear et al., 1996, Kandel et al., 1991)).

A dificuldade de se projetar um computador que alcance a inteligência artificial deriva do fato de que, apesar de reconhecermos a nossa própria inteligência, não conseguimos compreender o processo através do qual ela se dá. Somos capazes de obter medidas eletrofisiológicas bastante refinadas e até imagens do cérebro em atividade com *PET scan* e Ressonância Magnética Funcional, mas ainda não conseguimos saber o que o cérebro sendo avaliado está pensando naquele momento. Falta-nos o código para compreendermos a função cerebral.

Muito do que sabemos sobre o cérebro foi obtido através de pesquisa operacional. Este tipo de abordagem não busca inicialmente a compreensão do fenômeno sendo estudado, apenas o seu controle, suas respostas para um determinado

tipo de entrada (Ashby, 1973): é uma abordagem fenomenológica apenas, e o objeto estudado é considerado uma “caixa-preta”, seu interior sendo desconhecido. O estudo fenomenológico pode parecer pouco ambicioso quanto a seus objetivos, mas devemos ter sempre em mente o quanto o conhecimento puramente empírico é importante nas nossas vidas e foi na de nossos ancestrais. Para usar o exemplo de Ashby (1973), se esperássemos que a ciência clássica (a que procura o mecanismo dos fenômenos) descobrisse como fazer um pão comestível, ainda estaríamos sem poder comê-lo. Outro exemplo é o da farmacologia associada ao tratamento médico: quantos fármacos largamente utilizados, a cuja ação confiamos nossa saúde, operam por um mecanismo desconhecido. Sabemos apenas que funcionam, mas não como funcionam.

Existem, atualmente, tecnologias bastante poderosas para o estudo experimental dos mecanismos do cérebro. Entretanto, seus resultados até o momento não permitiram que se desvendasse o código do funcionamento cerebral. Talvez as tecnologias disponíveis ainda não sejam adequadas, mas é até possível que já exista a tecnologia adequada, e que, ou ela não deve estar sendo utilizada do modo correto, ou os resultados obtidos com ela não estão sendo analisados corretamente.

A pesquisa teórica, através de modelos, é limitada, como já se disse acima, pela nossa ignorância com respeito do funcionamento do cérebro. Entretanto, é possível com a pesquisa teórica o desenvolvimento de modelos que incorporem o maior número de características já reconhecidas como importantes para a função cerebral. Invariavelmente, este tipo de pesquisa leva a resultados surpreendentes, não só pela sua semelhança a alguns processos cerebrais, mas também por serem inesperados. Isso demonstra que freqüentemente os próprios modelos se portam como caixas-pretas.

A pesquisa de modelos de redes neurais passou por altos e baixos nas últimas décadas, e inúmeros modelos surgiram e caíram no esquecimento. Alguns pesquisadores, entretanto, estabeleceram as bases para o atual estado da pesquisa na área, estabelecendo conceitos bastante importantes.

McCulloch e Pitts (1943) postularam cinco princípios básicos para a representação de uma rede neural. São eles:

1. a atividade do neurônio é um processo de “tudo-ou-nada” (binário);

2. ativar um neurônio num dado instante de tempo, é necessária a atividade de um certo número de sinapses em um dado período de tempo, e este número é independente da atividade prévia ou da posição do neurônio;

3. o atraso sináptico é o único atraso significativo no sistema;

4. qualquer sinapse inibitória ativa pode impedir a ativação do neurônio num dado instante de tempo;

5. a estrutura da rede não muda com o tempo.

Os eventos que podem ser representados em uma rede deste tipo são aqueles que podem ser descritos como uma possibilidade dentro de um conjunto de entradas finito em um período de tempo também finito (Balaniuk, 1989). Von Neumann (1956) concluiu que o modelo de McCulloch e Pitts, na verdade uma materialização do computador de Turing (1936)) pode realizar qualquer função de transformação que possa ser definida logicamente sem ambigüidade em um número finito de passos.

Esses conceitos iniciais foram alterados com o passar do tempo, principalmente o número cinco. Segundo ele, a rede, não podendo ser alterada com o tempo, não pode aprender. Boa parte dos modelos atuais incorpora o aprendizado hebbiano. Hebb (1949) propôs o que posteriormente viria a ser chamado de “Regra de Hebb” ou “sinapse de Hebb”. Segundo o enunciado, “quando o axônio da célula A (...) excita a célula B e, repetida ou persistentemente, toma parte no ato de dispará-la, algum processo de crescimento ou mudança metabólica ocorrerá em uma ou em ambas as células de maneira que a eficiência de A como uma das células que dispara B será aumentada”. Hebb pressupôs que a eficiência de cada sinapse é uma característica dinâmica, podendo ser significativamente alterada com o passar do tempo.

Na virada para os anos 60, Rosenblatt desenvolveu uma classe de modelos conhecidos como *Perceptrons* (Rosenblatt, 1958, Block, 1962), incorporando conceitos de Hebb, McCulloch e Pitts. Várias formas foram testadas para os *perceptrons*, que alcançaram razoável grau de sucesso. Entretanto, Minsky e Papert (1969) levantaram uma série de críticas aos *perceptrons*, após uma detalhada análise

das suas deficiências. Tão contundentes foram as suas críticas que a área foi abandonada por uma série de pesquisadores (Balaniuk, 1989).

Apesar disso, alguns conceitos estabelecidos com os *perceptrons* sobreviveram. Entre eles está o de “retro-propagação” (“*back-propagation*”), que, ao incorporar o uso de vários níveis de conexão, resolveu as deficiências apontadas por Minsky e Papert (1969). Uma rede inicialmente aleatória é exposta a um estímulo a partir do qual há uma resposta esperada. A resposta da rede é comparada com a esperada e alterações nas conexões são feitas de modo a tornar a saída da rede mais próxima do esperado. Com esse tipo de modelo, é possível um certo grau de reconhecimento de consoantes pronunciadas, de imagens unidimensionais e até uma certa capacidade de compreensão e emissão de linguagem. Um dado interessante é que pequenos “danos” (alterações intencionais sem sentido nenhum) na rede geram saídas que lembram desordens humanas de linguagem, como a dislexia (North, 1987).

Um novo impulso na pesquisa em redes neurais veio com a publicação de um trabalho de Hopfield (1982), que deu origem a outros, como a máquina de Boltzmann (Hinton et al., 1984) e o ART 2, de Carpenter e Grossberg (1987). Outros modelos, menos relacionados ao trabalho de Hopfield, surgiram então.

Uma possível análise da perda de interesse na pesquisa com modelos que houve durante a década de 70 e da sua retomada nos anos 80 leva em consideração o apresentado por Thomas Kuhn, no seu “*The Structure of Scientific Revolutions*” (1969). Segundo Kuhn, quando um paradigma é estabelecido, surgem resultados que o corroboram. Entretanto, com o passar do tempo, são encontrados resultados que não se adequam ao paradigma, mas normalmente não recebem muita consideração. Se esses dados são realmente significativos e vão se acumulando, em um certo momento, percebe-se que o paradigma é, de fato, falho, pois aqueles resultados não são flutuações ou erros que possam ser desprezados. Um novo paradigma é então elaborado, mais amplo e geral, capaz de explicar todos os resultados anteriormente obtidos.

Sob essa ótica, o campo das redes neurais começou com muito entusiasmo nos anos 50 e 60, com o estabelecimento de alguns modelos como referenciais (com conceitos paradigmáticos). O trabalho de Minsky e Papert (1969) apontou então sérios problemas aos conceitos até o momento estabelecidos, e estes caíram. Foram

necessários mais alguns anos até que novos conceitos e abordagens começassem a surgir, e os modelos a se proliferar.

Dos vários modelos atuais de redes neurais, alguns foram mais bem recebidos pela comunidade científica, e têm se mostrado promissores em alguns aspectos. Osório (1992), falando sobre modelos de redes neurais em geral, diz que “A maioria desses modelos não abordados (no artigo dele) são variações dos modelos básicos apresentados neste artigo ou modelos que foram desenvolvidos e não possuem uma grande repercussão juntos aos principais grupos de pesquisa sobre redes neurais”. Seu comentário indica que, para a “descoberta” de um modelo correto quanto ao funcionamento do cérebro, não só é necessário o desenvolvimento deste modelo como também a sua aceitação “junto aos principais grupos de pesquisa sobre redes neurais”. É possível até que já tenham sido desenvolvidos modelos muito mais adequados que os mais conhecidos atualmente, mas que tenham sido publicado em algum periódico de pequena circulação ou que simplesmente não chamaram a atenção das “pessoas certas”.

O assunto dessa dissertação é um estudo realizado com um desses modelos menos conhecidos, o modelo do *Trion*, que incorpora uma série de conceitos gerais sobre o funcionamento do cérebro. Quem sabe no futuro ele não se revele um bom modelo para o funcionamento cortical.

1.1 Histórico do modelo do *Trion*

As origens do modelo do *Trion* estão no trabalho de Little (1974), que, sem saber das propostas semelhantes de Cragg e Temperley (1954), incorpora idéias da física de fenômenos cooperativos. Little, neste trabalho, estabeleceu uma analogia com o modelo de spin magnético da mecânica estatística idealizado por Ising (citado em Mello-e-Souza, 1993).

Segundo ambos os trabalhos, o comportamento de redes de neurônios seria semelhante ao de uma rede de magnetos, e o estado de cada unidade básica do modelo de rede neural estaria sujeito a duas variáveis importantes, temperatura (T) e energia (E). Essas duas variáveis influenciam a probabilidade de uma interação específica entre os magnetos, proporcional a $\exp(-E/kT)$.

O trabalho de Little (1974) estabelece uma analogia para E em uma rede neural. Essa energia seria comparável ao Campo Local de um neurônio. O conceito de Campo Local corresponde ao total de estímulos recebidos por um determinado neurônio em um determinado espaço de tempo. Ou seja, está relacionado à capacidade de um neurônio associar despolarizações e/ou hiperpolarizações na sua membrana, proporcionadas por células ligadas a ele, espacial e temporalmente próximas no momento em que “decide” se vai despolarizar completamente ou não.

Faltava ainda uma interpretação biológica para a variável temperatura dos modelos físicos. Uma interpretação foi sugerida por Shaw e Vasudevan (1974). Segundo estes autores, a temperatura representa as flutuações estatísticas na transmissão sináptica apontadas experimentalmente por Katz (1969).

Em 1978, Little e Shaw implementaram a capacidade de aprendizado hebbiano (ver seção 1.2.1) no modelo de Little (1974). Desse trabalho, duas linhagens tiveram origem.

A primeira delas, mais bem-recebida pelo público, foi a iniciada por Hopfield (1982), baseada em analogias de vidros de spin. A segunda, de menor impacto até agora, foi a do modelo do *Trion*, objeto de estudo deste trabalho, que incorporou conceitos do modelo ANNNI (*axial next-nearest neighbour Ising*) de Fisher e Selke (1980) juntamente com o princípio de organização colunar cortical de Mountcastle (1978).

1.2 Princípios básicos incorporados no modelo do *Trion*

De todos os trabalhos sobre o funcionamento cortical, alguns princípios têm se estabelecido como fundamentais. Alguns desses princípios estão em desacordo com os originalmente sugeridos por McCulloch e Pitts (ver acima), mas representam melhor o atual estado das pesquisas na área. Segundo Shaw et al. (1988), estes princípios seriam:

1.2.1 A sinapse de Hebb

Como foi referido anteriormente, em 1949, o psicólogo Donald Hebb postulou que deveria haver plasticidade sináptica entre dois neurônios para que houvesse aprendizado (ou memória).

1.2.2 A organização cortical multicolumnar de Mountcastle

Mountcastle (1978) propôs que a unidade básica de processamento do cérebro seria a coluna neuronal. Essa coluna, estabelecida como tendo aproximadamente 500 μm de diâmetro, compreende em torno de 10000 neurônios altamente interconectados, atravessando as seis camadas do córtex (Goldman-Rakic, 1984) e pode ser subdividida em subunidades de processamento, que seriam funcionalmente irreduzíveis (ver Shaw et al., 1982, Shaw e Pearson, 1986).

A atividade elétrica dos neurônios dentro da coluna leva à geração de padrões espaço-temporais complexos no conjunto, que, se propõe, seriam responsáveis pelo processamento cortical. Ressalta-se aqui a importância da escolha da correta escala espacial e temporal para a elaboração e análise de um modelo de redes neurais.

Boa parte dos modelos de redes neurais atuais toma como unidade básica o neurônio, normalmente abordado de forma binária (o “tudo ou nada” postulado por McCulloch e Pitts (1943). Se os dados eletrofisiológicos que indicam a ocorrência de padrões espaço-temporais de atividade no córtex na escala de centenas de microns realmente representarem a maneira como o cérebro “processa”, praticamente todos os modelos que trabalham na escala de “neurônio-como-unidade-funcional” estarão no caminho errado para a descoberta do código de funcionamento cortical.

A escala temporal também é importante para se desvendar o código (Shaw et al., 1988). Várias análises já foram realizadas com os disparos de um neurônio (Shaw e Pearson, 1986, Bach e Kruger, 1986), e alguns padrões claros e significativos até foram encontrados, mas nada que permitisse a identificação de um código. Uma escala temporal da ordem de dezenas de milissegundos foi proposta por Shaw e colaboradores (1985 e 1986, Silverman et al., 1986).

1.2.3 O selecionismo de Edelman

Edelman (1978 e 1987) propôs que uma rede de neurônios partiria de repertório inicial de padrões de atividade, e que o aprendizado se daria pela seleção hebbiana de alguns destes padrões específicos associados a estímulos também específicos. Essa proposição se opõe aos modelos que realizam aprendizado a partir de uma *tabula rasa* (Hopfield, 1982) ou que utilizam um supervisor (aprendizado supervisionado, típico de *back-propagation*) (North, 1987).

1.2.4 Flutuações estatísticas

Katz (1969) em conjunto com vários outros pesquisadores, demonstrou a existência de flutuações estatísticas tanto na taxa de liberação das vesículas sinápticas como na quantidade de neurotransmissor contido em cada uma dessas vesículas. Como disse Shaw e colaboradores (1988), essas flutuações podem não ser apenas “fatos da natureza com que o cérebro deve viver”, mas, sim, parte essencial do mecanismo de funcionamento cortical. Poderiam representar, talvez, o quanto “aquecido” está o sistema, quão predisposto a funcionar ele está.

1.2.5 Balanço entre excitação e inibição

Assim como as flutuações estatísticas, a existência de um balanço entre excitação e inibição tem sido mostrada como essencial para o funcionamento do sistema nervoso central. Um exemplo é o da inibição lateral nos campos receptivos de vários tipos de neurônios sensoriais (Shaw et al., 1988).

Levando-se em conta as escalas temporais dos potenciais excitatórios e inibitórios pós-sinápticos (PEPS e PIPS), pode-se perceber que os PIPS têm constantes temporais bem mais longas que os PEPS (Shaw et al., 1988). Essa característica pode ser representada em um modelo através de passos de tempo discretos em que a atividade de uma unidade é influenciada por pelo menos dois passos

de tempos anteriores, como no modelo ANNNI (Fisher e Selke, 1980). Na maior parte dos modelos de redes neurais, apenas um passo de tempo influencia a atividade das unidades.

1.3 O modelo do *Trion*

1.3.1 O *Trion*

A unidade básica do modelo do *Trion* é a que dá nome ao mesmo, *trion*. Essa unidade representa uma minicoluna idealizada de algumas centenas de neurônios altamente interconectados.

Um número arbitrário de estados de atividade é definido, sendo três estados o padrão (há, entretanto, trabalhos com cinco estados (Shaw et al., 1988)). O número de três foi escolhido como um compromisso entre “mais estados do que os normalmente utilizados (dois, modelando neurônios binários)” e “número suficientemente pequeno para permitir interpretação”, e é o utilizado neste trabalho. Esses três estados são invariavelmente denominados “alto” (+), “médio” (0) e “baixo”(-), representando, respectivamente, muitos, alguns e poucos neurônios na minicoluna disparando em um dado espaço de tempo. Um *trion* está representado na figura abaixo.

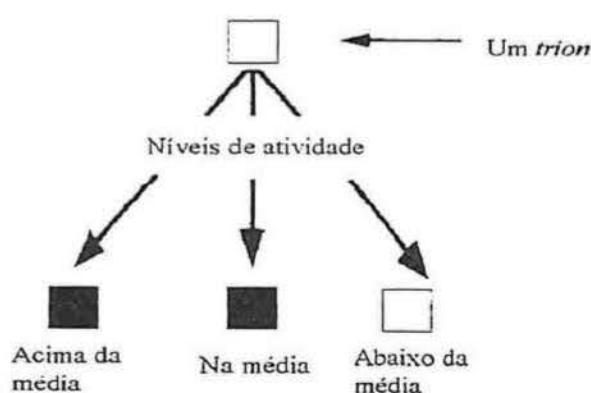


Figura I.1 Representação dos estados de atividade de um *trion* (reproduzido de Mello-e-Souza, 1993)

1.3.2 Colunas de *Trions*

Um conjunto de *trions* pode ser interligado, compondo uma coluna, análoga às colunas corticais propostas por Mountcastle (1978) referidas acima. Nos trabalhos anteriores, o número típico de *trions* em cada coluna variou entre quatro e oito, sendo seis o mais comum. Números maiores são possíveis, mas demandam maior tempo de computação.

Assim como as minicolunas de uma coluna cortical influenciam uma o estado das outras, o estado dos *trions* de uma mesma coluna influencia o estado dos outros *trions* da mesma coluna. Essa influência se dá através dos valores de “conexões” entre as unidades, representando uma analogia de axônios e feixes de axônios conectando uma minicoluna à outra. No modelo, os *trions* de uma mesma coluna estão agrupados em forma de anel, como na figura 1.2.

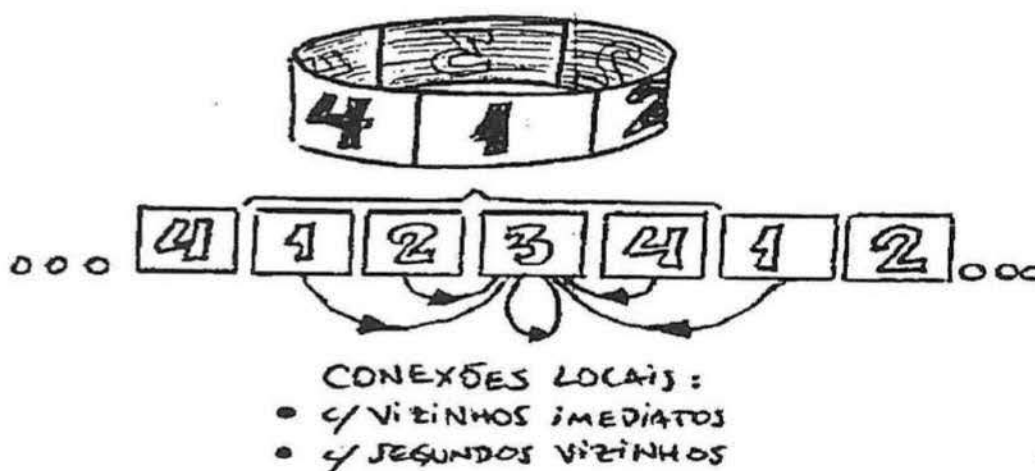


Figura 1.2 Representação esquemática de uma coluna de quatro *trions* e das conexões com seus vizinhos (reproduzido de Mello-e-Souza, 1993)

A estrutura de anel leva a conexões em que o estado de um determinado *trion* no passo de tempo τ em uma coluna é influenciado pelo seu próprio estado no passo de tempo anterior ($\tau-1$), assim como pelo estado dos seus vizinhos e dos seus segundos vizinhos também no passo de tempo anterior ($\tau-1$). A influência normalmente é maior para a auto-conexão, decrescendo a medida em que os *trions* ficam mais distantes (conexão com primeiros vizinhos mais forte que com segundos vizinhos).

Entretanto, não são só os estados do passo de tempo anterior que afetam o estado de um *trion*: como referido na seção 1.2.5, o modelo leva em conta a influência de dois passos de tempo, que no modelo ANNNI proporciona maior riqueza de estados. Sendo assim, além do primeiro passo de tempo anterior ($\tau-1$), também o segundo passo de tempo anterior ($\tau-2$) influencia o estado de um determinado *trion*. As conexões para $\tau-1$ são denominadas V, e as para $\tau-2$, W.

Os valores das conexões são tipicamente positivos para o primeiro passo de tempo anterior ($\tau-1$) e negativos para o segundo ($\tau-2$), normalmente com simetria para os primeiros e para os segundos vizinhos. O resultado dessa estrutura é que um determinado *trion* tende a ficar no estado em que ele e os seus vizinhos estavam em $\tau-1$ e ficar em um estado diferente do que ele e os seus vizinhos estavam em $\tau-2$. Um exemplo das conexões está abaixo.

<i>Trion</i>	<i>Seg. Viz.</i> <i>Esquerda</i>	<i>Prim. Viz.</i> <i>Esquerda</i>	<i>Auto-</i> <i>conexão</i>	<i>Prim. Viz.</i> <i>Direita</i>	<i>Seg. Viz.</i> <i>Direita</i>
V	0.5	1	2	1	0.5
W	-0.5	-1	-2	-1	-0.5

Tabela 1.1 Exemplo de conexões entre *trions*

A relação de um *trion* com os outros e o mapa de influências pode ser representado na figura abaixo:

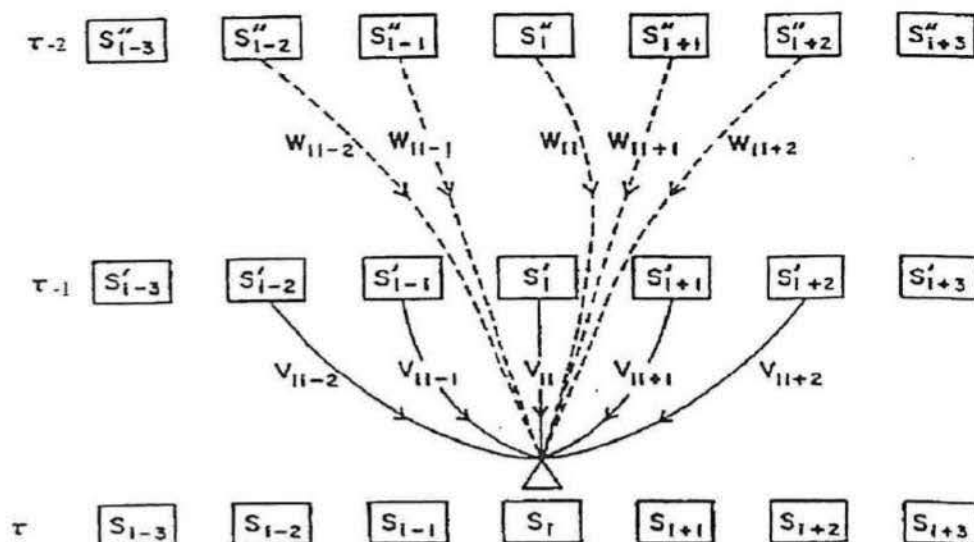


Figura 1.3 Mapa de conexões para um determinado *trion* (reproduzido de Mello-e-Souza, 1993)

Uma maneira econômica para se representar o conjunto de conexões de uma coluna é $V/W = 0 \ 1 \ 2 \ 1 \ 0 / 0 \ -1 \ -2 \ -1 \ 0$. Esse conjunto é o mesmo apresentado na tabela 1.1

1.3.3 Probabilidades para o estado de cada *trion*

Para a definição do estado (S) de um determinado *trion*, o primeiro passo é o cálculo da probabilidade de ocorrência de cada um dos três estados, alto, médio ou baixo. Isto é feito através da seguinte equação:

$$P(S_i) = \frac{g(S_i) \exp [B M_i S_i]}{\sum_s g(S) \exp [B M_i S]} \quad (\text{equação 1}).$$

O M_i da equação 1 é dado por:

$$M_i = \sum [V_{ij} S'_j + W_{ij} S''_j] - V_i^T \quad (\text{equação 2}).$$

1.3.3.1 O fator de degeneração g

Na equação 1, o g representa um fator de degeneração que aumenta a tendência de um *trion* estar no estado médio, o que seria esperado em uma minicoluna cortical. O g para o estado médio é muito maior que para os estado alto e ou baixo, ficando normalmente na ordem de 500:1:1.

1.3.3.2 A temperatura e a variável B

O B da equação 1 é inversamente proporcional à temperatura. Na equação física que deu origem ao modelo, B representava um conjunto de variáveis que tinha a temperatura T como um dos denominadores. No modelo do *trion*, T representa, segundo a interpretação dada por Shaw e Vasudevan (1974) a probabilidade de liberação acidental de vesículas sinápticas, Por conseqüência, quanto maior o valor de B , menor seria o número de vesículas sinápticas “acidentalmente liberadas” na simulação.

Trabalhos anteriores (Shenoy, 1988) demonstraram que uma ampla faixa de valores de B é interessante, indo de 2.0 até 8.0.

1.3.3.3 O Campo Local M_i

A variável M_i representa o Campo Local citado na seção 1.1. No modelo do *trion*, o Campo Local é o somatório das influências dos estados anteriores em $\tau-1$ e $\tau-2$ do *trion* sendo calculado, assim como dos estados dos seus primeiros e segundos vizinhos nestes mesmos passos de tempo. A influência individual de cada estado depende dos valores das conexões V e W usadas.

O V_i representa o limiar de disparo, fixado em zero no modelo do *Trion*, para que sempre haja atividade, mesmo que baixa, em cada *trion* (McGrann, com. pess.)

1.3.4 Determinação final do estado

Uma vez que as probabilidades de um determinado *trion* estar em cada um dos três estados estiverem calculadas, a determinação de qual deles efetivamente ocorrerá pode ser feita de duas maneiras: pelo método determinístico ou pelo protocolo Monte Carlo.

O método determinístico é relativamente simples: o estado com maior probabilidade de ocorrência é automaticamente escolhido. O protocolo Monte Carlo, entretanto, é um pouco mais complexo. Em vez de se escolher o maior, é feito um sorteio ponderado de acordo com as probabilidades de cada estado.

Em boa parte das vezes, um dos estados têm uma chance bem maior de ocorrer do que os outros (não necessariamente o estado médio, apesar do fator de degeneração g), e resultados semelhantes são obtidos pelos dois métodos. O método do protocolo Monte Carlo é, creio, mais representativo biologicamente e seria então mais interessante utilizá-lo. Entretanto, há uma característica do modelo, o surgimento de um repertório de padrões estáveis, que só é claramente observável e analisável quando se usa o método determinístico. Este, apresenta, então, algumas vantagens.

Uma vez que todos os cálculos tenham sido feitos para um determinado *trion*, e o seu estado tenha sido escolhido, o processo é repetido para os outros *trions* do mesmo passo de tempo um a um. Quando os estados de todos os *trions* do mesmo passo de tempo já foram calculados, passa-se para o passo de tempo seguinte.

1.3.5 Padrões emergentes

Quando são feitas simulações, uma característica logo observada é o surgimento de padrões cíclicos de atividade dos *trions*, identificados pela repetição dos estados de todos os *trions* no mesmo passo de tempo a cada n passos. Esse n é, na maioria das vezes, um múltiplo de seis, mas n 's iguais a três e a um também são comuns. O conjunto de padrões cíclicos que surgem para um determinado conjunto de conexões e uma determinada temperatura é chamado de repertório, e cada padrão é

chamado de MP (Padrão Mágico ou Mais Provável). O exemplo de um MP, obtido com as conexões 01210/0-1-2-1-0 e um B igual a 6.50, está abaixo:

Passo de tempo	Estado dos <i>trions</i>					
$\tau+0$	-	-	+	+	-	-
$\tau+1$	-	+	+	-	-	-
$\tau+2$	+	+	0	-	-	0
$\tau+3$	+	+	-	-	+	+
$\tau+4$	+	-	-	+	+	+
$\tau+5$	-	-	0	+	+	0
$\tau+6$	-	-	+	+	-	-
$\tau+7$	-	+	+	-	-	-

Figura 1.4 Exemplo de MP

Como se pode observar, o estado dos *trions* em $\tau+6$ é igual ao estado em τ , o mesmo valendo para $\tau+7$ e $\tau+1$. Se fosse continuada essa simulação, utilizando-se o método determinístico, $\tau+8$ seria igual a $\tau+2$, $\tau+9$ a $\tau+3$, e assim por diante; o padrão se repete.

1.3.5.1 MP's e estados iniciais da coluna

O estado inicial de uma coluna é a configuração dos *trions* para os dois primeiros passos de tempo, o mínimo para que a simulação possa ser feita, pois cada *trion* recebe influência de dois passos de tempo anteriores a ele. Se for usado o método determinístico, todos os possíveis estados iniciais de uma coluna acabarão entrando em um dos MP's do repertório daquela coluna, e não mais sairão deste MP. São padrões estáveis, portanto, característicos da rede (Shaw et al., 1988).

Utilizando-se o protocolo Monte Carlo, é possível se observar o surgimento dos mesmos MP's que aparecem com o método determinístico. Por vezes, entretanto, um MP sofre uma pequena alteração, decorrente da natureza aleatória do método, e passa a funcionar como um novo estado inicial, levando a coluna a sair do MP em que estava e entrar em outro. Os MP's são, então, pelo protocolo Monte Carlo, *quasi*-estáveis, e é comum o surgimento de vários MP's em uma mesma simulação, cada um durante alguns poucos ciclos e depois dando lugar a outro.

1.3.6 Probabilidade de Ciclagem dos MP's

Uma característica que pode ser obtida para um determinado MP é a sua Probabilidade de Ciclagem (PC). A PC, uma variável arbitrária mas útil assim mesmo representa a chance do MP ocorrer se estiver sendo usado o protocolo Monte Carlo. É, na verdade, a simples multiplicação das probabilidades de ocorrência dos estados escolhidos para cada *trion* em cada um dos passos de tempo do padrão. A PC é representada pela equação abaixo:

$$PC_{(MP)} = \prod_n \Pi_i(S_i(n\tau)) \quad (\text{equação 3}).$$

A PC é calculada mesmo quando se usa o método determinístico, pois ela é um índice da “robustez” de um MP. Uma PC alta indica uma grande “resistência” a eventuais perturbações, tais como o surgimento, se estiver valendo o protocolo Monte Carlo, de um *trion* com um estado diferente do que era esperado no MP. Quando é feita uma simulação com o protocolo Monte Carlo, os MP's com PC alta tendem a permanecer por mais ciclos seguidos do que os com PC baixa.

1.3.7 Aprendizado hebbiano

Para realizar-se o aprendizado hebbiano no modelo do *trion*, um determinado MP é escolhido para ser aprendido, e o aprendizado em si consiste em pequenas alterações nos valores das conexões entre os *trions*. Se uma determinada conexão colaborou para que um determinado *trion* assumisse os valores que ele assumiu no MP, essa conexão será intensificada. Se ela diminuiu as chances desse *trion* assumir esses valores, ela será enfraquecida.

O resultado final de um aprendizado normalmente é o aumento na Probabilidade de Ciclagem do MP sendo aprendido. Ou seja, este torna-se mais

“robusto”. Por vezes, entretanto, ocorre uma diminuição na PC, dependendo da intensidade do aprendizado.

O aprendizado é representado pelas equações abaixo, e a sua intensidade é determinada pela variável ε .

$$\Delta V_{ij} = \varepsilon \cdot \sum_{\text{ciclo}} S_i(n\tau) \cdot S_j((n-1)\tau) \quad (\text{equação 4})$$

$$\Delta W_{ij} = \varepsilon \cdot \sum_{\text{ciclo}} S_i(n\tau) \cdot S_j((n-2)\tau) \quad (\text{equação 5}).$$

A variável ε não possui necessariamente o mesmo valor para V e W, mas este geralmente é o caso. Quando os ε 's para V e W são diferentes, o mais comum é ambos possuírem o mesmo módulo, mas com sinal negativo para W.

Trabalhos anteriores (Leng et al., 1994) mostraram que para há valores ótimos de ε para o aprendizado de um MP. Os valores utilizados para ε normalmente ficam entre 0.05 e 0.5, mas outros também são possíveis.

1.3.8 Multi-colunas

Alguns anos após a elaboração do modelo do *trion*, algumas alterações foram sugeridas e incorporadas. Uma das mais interessantes é o uso de mais de uma coluna de *trions* lado a lado, que representaria um aumento na escala do modelo, passando a representar não apenas uma coluna cortical e as interações entre suas minicolunas, mas também duas ou mais colunas e as interações entre elas. Os trabalhos iniciais multi-colunas (Patera et al., 1989, Leng et al., 1994, Quillfeldt, 1994) utilizaram duas ou três colunas.

Para que as colunas interajam, é necessário que estejam conectadas de alguma maneira. No caso de três colunas, por exemplo, o primeiro *trion* de cada uma está conectado com o primeiro das outras duas, o segundo com os segundos, o terceiro com os terceiros, e assim por diante. Para duas colunas ou mais de três, o padrão se

repete, com o n -ésimo *trion* de cada uma conectado ao n -ésimo *trion* de todas as outras. Em todos os casos, a influência é recíproca.

Esse tipo de conexão é convencionalmente chamado de “conexão lateral”, e as conexões entre os *trions* de uma mesma coluna são chamadas “conexões internas”. Assim como com as conexões internas, as laterais também ligam os *trions* para os dois passos de tempo anteriores. As conexões laterais para $\tau-1$ são denominadas X , e as para $\tau-2$, Y .

Normalmente, as conexões laterais são tomadas como de uma ordem de grandeza abaixo do valor das internas (há um embasamento eletrofisiológico para tal (Leng et al., 1994)). Isso significa, obviamente, uma influência menor, o que é compatível com o que se conhece atualmente em termos biológicos: o número de conexões entre colunas corticais diferentes geralmente é menor do que dentro da mesma coluna cortical.

Em sistemas multi-colunas, o cálculo do Campo Local é um pouco diferente, pois também deve levar em conta as conexões laterais. A fórmula para M_i passa a ser então:

$$M_i = \Sigma (V_{ij} S_j^{\tau} + W_{ij} S_j^{\tau-1} + X_{ij} S_j^{\tau-2} + Y_{ij} S_j^{\tau-3}) - V_i^T \quad (\text{equação 6}).$$

A interação entre os *trions* de colunas diferentes está exemplificada na figura 1.5.

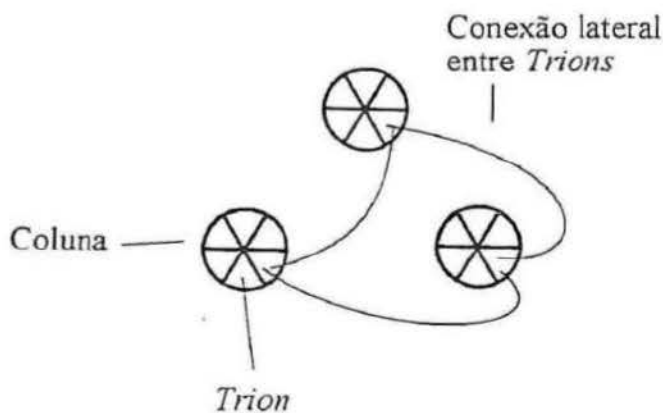


Figura 1.5 Exemplo de interação entre *trions* de diferentes colunas

1.3.9 MP's em multi-colunas

Normalmente as conexões internas de uma coluna são iguais às conexões internas das outras colunas. Isso significa que o repertório de MP's para todas as colunas deve ser igual, e, realmente, se todas as colunas possuírem o mesmo estado inicial, todas acabarão nos mesmo MP. O MP de cada coluna é chamado de "MP local", e o conjunto de MP's locais ocorrendo ao mesmo tempo no sistema multi-colunas é chamado de "MP global".

O MP global é normalmente composto pelo mesmo MP local em cada uma das colunas, mas pode haver um pequeno atraso de um passo de tempo de uma coluna em relação à outra. Se todas as colunas estão no mesmo passo de um MP, diz-se que estão em fase. Se uma está no quarto passo, outra no terceiro e ainda uma outra no segundo, diz-se, então, que estão fora de fase um passo. Um exemplo de MP global fora de fase está abaixo.

Passo de tempo	Coluna 1	Coluna 2	Coluna 3
$\tau+0$	--++--	-++---	++0--0
$\tau+1$	-++---	++0--0	++--++
$\tau+2$	++0--0	++--++	+-++++
$\tau+3$	++--++	+--+++	--0++0
$\tau+4$	+--+++	--0++0	--++--
$\tau+5$	--0++0	--++--	-++---
$\tau+6$	--++--	-++---	--++--
$\tau+7$	-++---	--++--	-++---

Figura 1.6 Exemplo de MP global composto pelo mesmo MP local em três colunas fora de fase

Trabalhos anteriores com o modelo (Leng et al., 1994) mostraram que a Probabilidade de Ciclagem de MP's fora de fase tendem a ser maiores do que se os mesmos MP's estão em fase.

1.3.10 Aprendizado lateral

Assim como o modelo permite aprendizado através de alterações nas conexões internas, ele permite o aprendizado através de alterações nas conexões laterais. Esse tipo de aprendizado é chamado "lateral", e deve resistir-se a tendência de chamá-lo "externo" (assim como chamar as conexões laterais de "externas"), para que não seja confundido com aprendizado supervisionado (este realmente externo, pois vem de fora do computador) que ocorre em alguns outros modelos, como os de *back-propagation*.

Para um sistema multi-colunas de *trions*, é possível, portanto, a realização de aprendizado interno, externo, ou ambos simultaneamente. A intensidade (ϵ) de ambos pode ser a mesma ou diferente, e o ϵ para X pode ser igual ou diferente ao ϵ para Y. Assim como com o aprendizado interno, quando os valores de ϵ diferem para X e Y, normalmente a diferença está apenas no sinal, e o módulo é o mesmo.

As variações nas conexões laterais são calculadas com as fórmulas abaixo.

$$\Delta X_{ij} = \epsilon \cdot \sum_{\text{ciclo}} S_i (n\tau) \cdot S_j ((n-1)\tau) \quad (\text{equação 7}) \text{ e}$$

$$\Delta Y_{ij} = \epsilon \cdot \sum_{\text{ciclo}} S_i (n\tau) \cdot S_j ((n-2)\tau) \quad (\text{equação 8}).$$

Trabalhos anteriores (Leng et al., 1994) mostraram que o aprendizado lateral também pode ser eficiente para promover o aumento na Probabilidade de Ciclagem de MP's.

1.4 Analogia com o sistema visual humano

Como uma maneira de facilitar a compreensão do modelo do *Trion*, é interessante traçar-se uma analogia entre o modelo e a memória visual do ser humano. A validade dessa analogia, pelo tanto que se sabe, não pode ser estendida com segurança ao real funcionamento do córtex visual, e muito cuidado deve ser tomado então para que extrapolações enganosas não ocorram.

Como foi referido acima, as colunas de *trions* representam colunas corticais, e os *trions* em si, minicolunas dentro das colunas. Um MP pode ser entendido como

sendo um rosto conhecido, e a Probabilidade de Ciclagem como sendo a intensidade da memória que o indivíduo tem desse rosto, a facilidade com que lembra-se dele. O conjunto de MP's pode representar o conjunto de rostos que o indivíduo conhece.

Um rosto cujo MP tem baixa Probabilidade de Ciclagem pode não ser reconhecido se uma pequena perturbação ocorrer (ver seção 1.3.6). No modelo, a perturbação pode ser um *trion* que assume um estado inesperado; na analogia, podem ser óculos escuros, cabelo pintado ou cortado, qualquer coisa que dificulte um pouco o reconhecimento.

O valor exato da Probabilidade de Ciclagem que indica a memória de um rosto pode ser dependente de características únicas possuídas pelo rosto. Olhos de uma cor incomum, cabelos diferentes do normal, uma boca muito grande ou dentes acavalados normalmente facilitam o reconhecimento. Um rosto com uma ou mais dessas características teria um MP de alta Probabilidade de Ciclagem.

À medida em que um rosto torna-se mais familiar e ocorre o aprendizado do mesmo, as conexões que aumentam a Probabilidade de Ciclagem do MP que o representa são levemente alteradas. O resultado é uma maior facilidade para o reconhecimento.

Essa analogia para o reconhecimento de rosto pode também ser feita para o reconhecimento de formas ou de outros padrões.

2. Objetivo

O objetivo deste trabalho é estudar os efeitos da variação da intensidade de aprendizado lateral na Probabilidade de Ciclagem de padrões em fase e fora de fase, com ε de mesmo sinal e de sinal inverso para o primeiro e o segundo passo de tempo, em um sistema com três colunas de seis *trions* cada.

3. Material e Métodos

3.1 A plataforma do trabalho

Todas as simulações foram rodadas no supercomputador Cray Y-MP 2E do Centro Nacional de Supercomputação (CESUP) da Universidade Federal do Rio Grande do Sul. O uso do supercomputador não se deveu, de início, à necessidade de grande poder de processamento, mas sim devido à prévia implementação de outros estudos com o modelo no CESUP (Mello-e-Souza, 1993).

3.2 A programação

O modelo foi programado na linguagem FORTRAN ANSI 77 e compilado com o compilador CF77, residente do supercomputador do CESUP. De maneira a facilitar sua compreensão e facilitar futuros trabalhos, o programa foi dividido em subrotinas, módulos responsáveis cada um pela execução de parte das tarefas do programa. Como exemplo, podemos citar a subrotina MCALC, responsável exclusivamente pelo cálculo do Campo Local (M_i) das equações 2 e 6, ou ainda a subrotina LATERAL, responsável pelo aprendizado lateral.

Algumas das subrotinas existiam anteriormente à elaboração deste programa, tendo sido enxertadas nele. Isso apresenta pelo menos três vantagens: poupa tempo de programação (como referido no parágrafo anterior), padroniza a subrotina e evita erros, pois não é necessária nem a elaboração de um novo algoritmo nem a digitação de um já estabelecido. A validade da construção do programa ao estilo "*Frankenstein*" foi comprovada, pois os erros encontrados durante os trabalhos residiam apenas na parte nova, recém-elaborada do programa.

O programa em si era apenas a implementação das equações 1 (cálculo da probabilidade de cada estado), 6 (cálculo do Campo Local para sistemas multi-colunas, 3 (Probabilidade de Ciclagem), e 7 e 8 (aprendizado lateral), acompanhado de outros algoritmos. Um deles realizava a leitura dos dados iniciais para os quais se calculariam os resultados; outro dava saída a estes resultados para um outro arquivo.

3.3 Parâmetros utilizados

Neste trabalho foram utilizadas três colunas de seis *trions* cada. Cinco tipos de conexão interna entre os *trions* foram testadas, e estão resumidas abaixo, junto com os nomes dados a elas. As conexões laterais eram inicialmente iguais a zero (ver seção 3.4).

Conexão	V	W
C1	0 1 1 1 0	0 -1 -1 -1 0
C2	0 1 0 1 0	-1 0 0 0 -1
C3	0 1 2 1 0	0 -1 -2 -1 0
C4	0 0 0 0 0	0 0 -1 0 0
C5	0 1 2 1 0	-1 0 -2 0 -1

Tabela 3.1 Conexões internas utilizadas neste trabalho

Para cada tipo de conexão, o repertório completo de MP's foi obtido, e, dentre estes, cinco foram escolhidos, perfazendo um total de 25 padrões no trabalho. É importante ressaltar que os MP's de um determinado tipo de conexão dificilmente são MP's de outro tipo de conexão.

Para cada um dos MP's, duas simulações foram executadas. A primeira tinha todas as colunas em fase. A segunda tinha uma coluna defasada um passo de tempo, e outra defasada dois passos.

Apenas o método determinístico de determinação do estado dos *trions* foi usado.

3.3.1 Entrada dos MP's no computador

A entrada do estado inicial que gera um MP em uma coluna foi feita a partir de um índice ternário unívoco, ou número de Identidade, (ID). O cálculo do ID de um MP é feito através da equação abaixo:

$$ID = \sum_{\tau=1} \sum_{NA=1} S_i(\tau) * (3^{((\tau-1)*(NA+NA-1))})$$

onde $S_i(\tau)$ é o estado do *trion*, no instante τ , e NA é igual ao número de *trions* na coluna.

No início da simulação, cada coluna recebia o seu ID, que geraria o MP desejado em fase ou fora de fase, de acordo com a simulação específica sendo realizada.

3.4 A simulação

Inicialmente, as conexões laterais tinham o valor de zero. Esse valor era modificado através do aprendizado hebbiano, e a Probabilidade de Ciclagem era calculada. Em seguida, as conexões laterais eram reinicializadas, e um novo aprendizado, com uma intensidade (ϵ) de módulo um pouco maior, era feito. A nova Probabilidade de Ciclagem era então calculada. Não foi realizado aprendizado interno.

Esse ciclo “conexões laterais zeradas - aprendizado - cálculo da PC” foi realizado trezentas e uma vezes para cada MP em fase e fora de fase, sempre se aumentando o módulo da intensidade do aprendizado em 0.01. O módulo inicial da intensidade era zero, e o final, três.

3.5 Formatação dos resultados

Os 301 valores de ϵ lateral foram plotados contra os valores obtidos para a Probabilidade de Ciclagem para cada um dos MP's estudados em cada uma das situações.

4. Resultados

4.1 Observações Gerais

A maior parte dos padrões estudados (escolhidos aleatoriamente dentro do repertório de MP's de cada tipo de conexão, como especificado na seção anterior) de cada uma das conexões comportou-se de maneira semelhante. Frequentemente, acima de determinada intensidade de aprendizado, a Probabilidade de Ciclagem (PC) estacionava, atingindo um patamar. A PC neste patamar foi também a PC máxima (a PC máxima é a maior Probabilidade de Ciclagem alcançada por cada MP em cada uma das condições testadas) para vários padrões. Para vários outros, entretanto, a PC não estava no patamar. De um modo geral, a PC máxima dos padrões fora de fase foi maior do que a dos em fase.

A maior parte dos resultados se baseia em comparações “em fase *versus* fora de fase” e “elat's de mesmo sinal *versus* elat's de sinal inverso”. É interessante ter isso em mente durante o exame dos resultados.

4.2 As conexões

4.2.1 Conexão C1: V 0 1 1 1 0

W 0 -1 -1 -1 0

A conexão C1 (ver tabela 1.1 para convenções) apresentou alguns resultados interessantes. De um modo geral, dos vinte resultados dessa conexão, em sete a PC não atingiu um patamar, mas em quatro destes parecia prestes a fazê-lo. Nas três ocasiões em que a PC não atingiu um patamar, os padrões estavam fora de fase e sinais inversos estavam sendo usados para os elat's.

Dos cinco padrões utilizados, em apenas um a PC máxima foi maior para a situação fora de fase, o que vai contra as observações feitas em trabalhos anteriores (Leng et al., 1994), nas quais geralmente os MP's fora de fase tinham PC's maiores que os em fase.

Dois MP's locais (como dito na seção 1, um MP global é o formado pelos MP's locais de cada coluna) utilizados eram bastante semelhantes e estão representados abaixo:

Passo	MP local 1	MP local 2	Passo
1	- + - 0 - +	- + - 0 - +	1
2	+ - + 0 + 0	+ - 0 - - -	2
3	- + 0 + + +	- + - 0 - 0	3
4	+ - + 0 + -	+ - + 0 + -	4
5	- + - 0 - 0	- + 0 + + +	5
6	+ - 0 - - -	+ - + 0 + 0	6

Tabela 4.1 Dois MP's semelhantes de C1.

A semelhança pode ser percebida se alterarmos a ordem dos passos de tempo de um deles. Alterando a ordem do MP 2, podemos obter:

Passo	MP 1	MP 2	Passo
1	- + - 0 - +	- + - 0 - +	1
2	+ - + 0 + 0	+ - + 0 + 0	6
3	- + 0 + + +	- + 0 + + +	5
4	+ - + 0 + -	+ - + 0 + -	4
5	- + - 0 - 0	- + - 0 - 0	3
6	+ - 0 - - -	+ - 0 - - -	2

Tabela 4.2 MP's rotados.

O MP 2 é claramente o MP 1 rodando na seqüência temporal contrária. Esse tipo de fenômeno já foi observado anteriormente por Bodner e colaboradores (1997). Os efeitos de tal fato são interessantes. Ambos padrões possuem comportamento semelhante se estiverem em fase e com elat's de mesmo sinal. A semelhança pode ser observada nos gráficos pelo traçado das curvas, (gráfico C1P3X e C1P9X do apêndice) e pelo fato da PC inicial (para elat igual a zero), a PC máxima e o elat da PC máxima serem exatamente os mesmos para ambos MP's. Houve, entretanto, diferenças: para todos os tipos de situação (fase ou fora de fase, elat's iguais ou inversos), os patamares eram diferentes, e, para elat de sinal inverso, os resultados gráficos eram bastante diferentes.

4.2.2 Conexão C2: V 0 1 0 1 0
W -1 0 0 0 -1

Os gráficos para a conexão C2 foram bastante mais complexos do que para C1, e nenhuma comparação pôde ser extraída facilmente deles. Pôde-se observar, entretanto, que seis dos vinte não atingiram um patamar, desempenho semelhante ao de C1.

Para esta conexão, a previsão de que os MP's fora de fase teriam uma PC máxima mais alta do que os mesmos MP's em fase se confirmou. Para todos os padrões, a PC máxima foi mais alta para a condição fora de fase.

4.2.3 Conexão C3: V 0 1 2 1 0
W 0 -1 -2 -1 0

A conexão C3 é a que temos mais estudado até hoje, e é provavelmente uma das mais relevantes biologicamente (Shaw et al., 1997). Os desempenhos de todos seus padrões foram bastante semelhantes, e as PC's máximas atingidas eram frequentemente próximas a 1, o máximo possível.

Para os MP's com elat's de mesmo sinal, a condição fora de fase apresentou PC's máximas mais altas. Entretanto, para elat's de sinal inverso, dois padrões em fase tiveram PC máxima mais alta do que os padrões correspondentes fora de fase.

Todos os MP's em todas as condições testadas atingiram algum patamar.

4.2.4 Conexão C4: V 0 0 1 0 0
W 0 0 -1 0 0

A conexão C4 é bastante simples, mas produziu curvas interessantes, que, sem exceção, atingiam um patamar. Alguns MP's chegaram a uma PC máxima igual ou próxima a 1. Para elat's de mesmo sinal, os MP's fora de fase obtiveram a maior parte

das PC's máximas., mas para ε lat's inversos, todas as PC's máximas eram dos MP's em fase, ao contrário do que geralmente é observado.

4.2.5 Conexão C5: V 0 1 2 1 0
W -1 0 2 0 -1

A conexão C5, apesar da sua semelhança com a C3, apresentou resultados um pouco diferentes dos desta. Uma observação inicial importante é que, após a execução das simulações, descobrimos que o primeiro e o terceiro MP's desta conexão eram de fato o mesmo MP: um é o "espelho" do outro. Este fato, assim como rotações espaciais (possível graças a estrutura anelar de cada coluna), não é detectado pelos algoritmos que obtêm o repertório da conexão e, portanto, permaneceu desconhecido até as simulações serem concluídas. A conexão C5 possui, portanto, apenas 4 MP's.

Um resultado desta conexão surgiu na comparação entre o terceiro e o quarto MP's, reproduzidos abaixo:

Passo	MP local 3	MP local 4	Passo
1	- + + - - -	0 + + - - -	1
2	- + + - - -	0 + + - - -	2
3	0 + + 0 - -	0 + + 0 - -	3
4	+ + + + - -	0 + + + - -	4
5	+ + + + - -	0 + + + - -	5
6	0 + + 0 - -	0 + + 0 - -	6

Tabela 4.3 Dois MP's semelhantes de C5.

Os MP's são relativamente semelhantes, e seus resultados para ε 's de mesmo sinal foram muito semelhantes, tanto visualmente no gráfico como em termos de valores de PC's máxima e inicial calculadas. Entretanto, para ε lat's de sinal inverso, os resultados são bastante diferentes.

Ainda sobre o quarto MP com ε lat's de sinal inverso, uma observação curiosa é o fato de que a PC inicial também era a PC máxima, pois a PC sofria uma queda logo no início e não se recuperava mais. Isso ocorreu tanto para o MP em fase como fora de fase, que tinham a mesma PC inicial, já que o fator "fase" não influencia a PC inicial

(as conexões laterais, responsáveis pelo “em fase-fora de fase”, são iguais a zero neste momento). A PC do patamar, entretanto, foi maior para o MP fora de fase do que o em fase.

Sobre as PC's máximas nas outras condições, para ϵ 's de mesmo sinal, três de quatro estavam com os MP's fora de fase. Para ϵ lat's de sinal inverso, a PC máxima ficou com a condição em fase para um dos MP's; para outros dois, ficou com o fora de fase. Para o quarto, a PC máxima da condição em fase ficou igual a da condição fora de fase (ver acima).

4.3 Resumo dos resultados

4.3.1 Patamar

Na tabela abaixo, estão indicadas as percentagens das simulações que atingiram um patamar.

	Em fase	Fora de fase
C1 ϵlat's de mesmo sinal	100	40
C1 ϵlat's de sinal inverso	100	20
C2 ϵlat's de mesmo sinal	100	60
C2 ϵlat's de sinal inverso	60	60
C3 ϵlat's de mesmo sinal	100	100
C3 ϵlat's de sinal inverso	100	100
C4 ϵlat's de mesmo sinal	100	100
C4 ϵlat's de sinal inverso	100	100
C5 ϵlat's de mesmo sinal	100	100
C6 ϵlat's de sinal inverso	100	100

Tabela 4.4 Percentagem dos MP's que atingiram um patamar, de acordo com as condições testadas.

4.3.2 Probabilidade de Ciclagem máxima

4.3.2.1 Com relação ao sinal de ϵ lat

A tabela abaixo indica a percentagem de PC's máximas que ficaram com os MP's fora de fase. A percentagem das PC's máximas com os MP's em fase é o complementar dos dados mostrados.

	elat's de mesmo sinal	elat's de sinal inverso
C1	20	80
C2	60	100
C3	100	60
C4	80	0
C5	75	25 *

Tabela 4.5 Percentagem dos MP's fora de fase que tiveram PC máxima maior do que suas versões em fase.

* - O quarto padrão teve a mesma PC máxima tanto em fase como fora de fase, mas o padrão em fase teve um pico secundário mais alto do que o fora de fase, o que foi levado em conta na tabela (ver gráficos C5P9Y e C5P10Y do apêndice)

4.3.2.2 Probabilidade de Ciclagem alta

Como uma maneira de perceber a efetividade de um tipo de conexão em produzir MP's de alta PC, abaixo está apresentada a percentagem de MP's (independente das condições) de cada conexão que obtiveram uma PC arbitrada alta, no caso, com valor acima de 0.9.

Conexão	Percentagem de MP's acima de 0.9
C1	35.00
C2	25.00
C3	65.00
C4	60.00
C5	56.25

Tabela 4.6 Percentagem de MP's acima de 0.9 por conexão.

4.3.2.3 Probabilidade de Ciclagem igual a 1.0

Houve bastante diferença entre a capacidade do elat's levarem a Probabilidades de Ciclagem iguais a 1.0. A percentagem de MP's que atingiram esta PC para cada conexão está representada na tabela abaixo.

Conexão	elat's mesmo sinal	elat's sinal inverso
C1	20	0
C2	0	0
C3	30	0
C4	30	10
C5	12.5	0

Tabela 4.7 Percentagem de MP's que atingiram PC igual a um por conexão e por tipo de elat.

5. Conclusões

As principais conclusões obtidas a partir dos dados foram:

i) Padrões fora de fase geralmente possuem uma Probabilidade de Ciclagem máxima mais alta do que os fora de fase. Isso é mais evidente quando se utilizam valores de ϵ_{lat} com mesmo sinal. (4.3.2.1);

ii) Um valor de auto-conexão com módulo menor ou igual às demais conexões internas parece diminuir as chances da PC chegar a um patamar, principalmente no caso de padrões fora de fase. (4.3.1);

iii) Tipos de conexões com um módulo de auto-conexão maior que o das conexões com os vizinhos parecem levar a PC's máximas de módulo também maior do que os tipos de conexões que não se encaixam neste padrão. (4.3.2.2);

iv) Padrões estruturalmente semelhantes apresentam resultados semelhantes apenas para ϵ_{lat} 's de mesmo sinal. (4.2.1 e 4.2.5);

v) Cada MP parece ter uma intensidade ótima de aprendizado lateral para aumentar sua Probabilidade de Ciclagem. (4.2);

vi) Alguns MP's, em determinadas condições, não apresentam nenhuma melhora com relação à PC inicial com o aumento da intensidade de aprendizado lateral. (4.2.5);

vii) ϵ_{lat} 's de sinal inverso levam a resultados diferentes do esperado segundo os trabalhos anteriores (Leng et al., 1994) e normalmente bastante erráticos.

6. Discussão

6.1 Efeitos do sinal da intensidade de aprendizado lateral ϵ_{lat}

A intensidade de aprendizado lateral ϵ_{lat} teve efeitos consideravelmente importantes no desempenho dos MP's testados se compararmos ϵ_{lat} 's de mesmo sinal e de sinal inverso. De um modo geral, quando os MP's eram rodados com ϵ_{lat} de sinal negativo para Y, seu desempenho eram bem inferior a quando eram rodados com ϵ_{lat} 's de mesmo sinal para X e Y, ou apresentavam resultados diferentes do esperado.

Segundo estudos anteriores (Leng et al., 1994), era esperado que os MP's tivessem Probabilidades de Ciclagem máximas maiores fora de fase do que em fase. Para simulações com ϵ_{lat} de mesmo sinal, isso ficou bastante evidente: ocorreu em quatro dos cinco tipos de conexões estudados. A conexão C1 foi a única que ficou abaixo de 50%, mas C2 ficou pouco acima deste valor (4.3.2.1). Para ϵ_{lat} 's de sinal inverso, C1 e C2 inesperadamente tiveram uma percentagem dos seus MP's com resultados melhores fora de fase, mas C3 desempenhou mal (60%) e C4 e C5 foram sofríveis em alcançar o esperado. C1 e C2, entretanto, possuem uma característica que pode os distinguir das outras conexões e que pode ter sido responsável, pelo menos em parte, por sua diferença de comportamento. Essa característica será discutida mais adiante (ver 6.3).

A utilização de ϵ_{lat} 's "inversos" (com sinal negativo para o segundo passo de tempo de interação entre os *trions*, como dito anteriormente) também dificultou aos MP's atingirem PC's iguais a 1.0, como pode ser visto na tabela 4.7 e discutido na seção 6.5.

Uma possível interpretação para essa diferença nos desempenhos de ϵ_{lat} 's de mesmo sinal e de sinal inverso reside em uma análise do algoritmo de aprendizado hebbiano. A função deste algoritmo é analisar a participação de cada *trion* em cada passo de tempo na formação do MP, o quanto cada *trion* específico colabora, através das suas conexões com os vizinhos, para que o MP seja alcançado. Se a influência de um determinado *trion* durante o cálculo do Campo Local M (eq. 6) colaborou para que um outro *trion* atingisse o estado que era necessário para a obtenção do MP, as conexões do primeiro serão intensificadas pelo algoritmo. Na situação contrária, se a

influência de um determinado *trion*, diminuiu as chances do outro chegar ao estado certo, as conexões do primeiro serão enfraquecidas.

Para o primeiro passo de tempo de influência, que nas nossas simulações sempre tem sinal positivo, o algoritmo funciona desta forma. Entretanto, para o segundo passo de tempo, aquele que variava o sinal de ϵ_{lat} , o fato de em algumas simulações ele estar com sinal diferente do ϵ_{lat} para o primeiro passo de tempo (sinal negativo, portanto), provavelmente levou o algoritmo a funcionar da maneira inversa à normal.

As conexões para o segundo passo de tempo são invariavelmente negativas. O algoritmo de aprendizado, ao verificar a sua eficiência na promoção do MP, tende a aumentar o valor absoluto das conexões colaboradoras se o ϵ_{lat} para Y for positivo. Efetivamente, está diminuindo a influência da conexão, ao invés de aumentá-la, pois seu módulo estará sendo diminuído. Entretanto, para ϵ_{lat} de sinal negativo para Y, o módulo da conexão eficiente tende a aumentar, reforçando a conexão. Por que, então, ϵ_{lat} 's de sinal inverso afastam o desempenho do MP do esperado e tornam seu comportamento errático em alguns casos?

Uma possível explicação é que, ao aumentar o módulo das conexões laterais de Y, o ϵ_{lat} negativo aumentou também a competição entre as influências do primeiro e do segundo passos de tempo. Por sua vez, o ϵ_{lat} positivo reduziu as influências do segundo passo de tempo, diminuindo a competição. Se essa interpretação é correta, uma conclusão possível é que o aprendizado lateral para o segundo passo só produz os resultados esperados se ele diminuir a eficiência das conexões Y, o contrário da intenção original, que era de aumento da eficiência. Então, se for realizado aprendizado lateral com ϵ_{lat} negativo para o segundo passo de tempo provavelmente surgirá uma competição lateral, entre as colunas.

É provável que essas conclusões sejam válidas também para o aprendizado interno no segundo passo de tempo, alterando as conexões W. Seria interessante testar essa hipótese, assim como também testar o comportamento dos MP's diante de ϵ 's com sinal negativo para o primeiro passo de tempo. Ocorreria um anti-aprendizado? E se ocorresse, como se apresentaria? Quedas na PC? Redução no número de estados iniciais que levam ao MP aprendido?

Uma sugestão que talvez possa ser de utilidade no futuro é esperar-se um comportamento mais errático para os MP's quando se utilizar ϵ 's de sinal inverso. Nesse caso, o esperado não seriam PC's mais altas para os MP's fora de fase, por exemplo. Até o fim deste trabalho, entretanto, continuaremos com a presunção de que o esperado para ϵ lat's inversos seria semelhante ao esperado para ϵ lat's de mesmo sinal.

6.2 Efeitos da fase do MP

Classicamente, MP's fora de fase são tidos como mais robustos do que MP's em fase: o atraso de uma rede em relação à outra traria, de alguma forma, estabilidade ao sistema. Isso se confirmou, como pode ser observado na tabela 4.5, mas com mais clareza para MP's com ϵ lat de mesmo sinal, como discutido na seção 6.1.

Houve, entretanto, um efeito inesperado: para C1 e C2, a condição fora de fase tornou mais difícil para os MP's atingirem um patamar, enquanto que, em fase, apenas para C2 com ϵ lat inverso os MP's não atingiram o patamar (ver tabela 4.4). Uma possível explicação está na característica anunciada em 6.1 que distingue C1 e C2 das outras conexões. Discutiremo-la agora, então.

6.3 Possível distinção entre as conexões C1 e C2 e as outras conexões

Para facilitar a comparação, todas as conexões estão reproduzidas na tabela abaixo.

Conexão	V	W
C1	0 1 1 1 0	0 -1 -1 -1 0
C2	0 1 0 1 0	-1 0 0 0 -1
C3	0 1 2 1 0	0 -1 -2 -1 0
C4	0 0 1 0 0	0 0 -1 0 0
C5	0 1 2 1 0	-1 0 2 0 -1

Tabela 6.1 Valores das conexões internas para os cinco padrões de conexão

A característica que distingue C1 e C2 das outras conexões pode não ser claramente percebida de início, mas uma análise cuidadosa mostrará que, tanto em C1 como em C2, o módulo da auto-conexão é menor ou igual aos módulos das conexões com os primeiros vizinhos. Em C3, C4 e C5, a auto-conexão sempre é mais intensa do que as conexões com os segundos vizinhos. Relacionado a essa observação também está o fato de que nem C1 nem C2 possuem uma conexão com intensidade 2. Surgem então duas possibilidades.

6.3.1 Auto-conexão *versus* conexões com vizinhos

É possível que, dado que a auto-conexão influencia tanto ou menos que os primeiros vizinhos, as influências de um *trion* para o outro tenham ficado difusas demais. Essa difusão permite que muitas conexões sejam capazes de influenciar significativamente o estado de cada *trion* individual, levando a uma competição excessiva, o que tornaria os resultados menos estáveis com o aumento da intensidade de aprendizado lateral e levaria a Probabilidades de Ciclagem máximas menores do que as alcançadas pelas outras conexões, como pode ser observado na tabela 4.6.

6.3.2 Ausência de conexão forte

É também possível que o comportamento de C1 e C2 tenha se devido a ausência de uma conexão com um valor mais forte. O fato de o módulo máximo das conexões C1 e C2 ser 1 pode tê-las tornado, assim como referido acima, difusas demais. Desse modo, nenhuma influência se sobressairia, e também haveria competição excessiva.

Há, entretanto, o caso da conexão 4, que também tinha um módulo máximo de 1 para suas conexões internas. Apesar desse fato, os resultados de C4 foram de acordo com o esperado (se não contarmos a influência perturbadora da condição de *elats* inversos, referida em 6.1). Se levarmos em consideração os resultados de C4, é mais provável que os resultados de C1 e C2 devam-se principalmente à sua reduzida auto-

conexão. Seria interessante, para uma verificação mais completa, testar-se o desempenho de conexões tipo 02220/0-2-2-20 ou 02020/-2000-2, modificações de C1 e C2.

6.3.3 Competição excessiva favorecendo C1 e C2 com ϵ lat inverso?

Os resultados expressos na tabela 4.5 chamam a atenção pelo fato de que, ao mesmo tempo em que C1 e C2 foram menos eficientes em obter PC's mais altas para os seus MP's fora de fase do que as outras conexões quando foram utilizados ϵ lat's de mesmo sinal, elas tiveram um desempenho excelente quando ϵ lat's de sinal inverso foram utilizados. Em 6.1 foi aventada a possibilidade do ϵ lat inverso aumentar a competição entre os *trions* de diferentes redes para influenciarem os outros; em 6.3.1 e 6.3.2 sugeriu-se que C1 e C2 teriam um considerável nível de competição interna. É interessante, portanto, que em condições nas quais se esperaria o máximo de competição, C1 e C2 com ϵ lat's inversos, os MP's desempenharam igual a situação em que, para as outras conexões, ocorre mínima competição.

Uma conclusão possível é que tanto a competição interna que naturalmente ocorreria em C1 e C2 como a competição lateral promovida pelos ϵ lat's de sinal inverso são prejudiciais para que os MP's desempenhem de acordo com o esperado segundo os estudos anteriores (Leng et al., 1994). Entretanto, a combinação das duas competições parece fazer com que os MP's comportem-se mais próximos do esperado, pelo em termos de "PC maior para MP fora de fase".

É importante ressaltar, entretanto, que mesmo com essa maior adequação ao esperado que surge da combinação C1 ou C2 com ϵ lat's inversos, as curvas de Probabilidade de Ciclagem são ainda bastante complexos. Se há um efeito de neutralização das competições quando se analisa a capacidade dos MP's dessas conexões cumprirem certos quesitos, como o referido no parágrafo anterior, pode-se também verificar um aumento na "erraticidade" das curvas.

6.3.4 PC's mais baixas para C1 e C2

Mesmo com ϵ lat's inversos parcialmente compensando a provável competição interna em C1 e C2, os MP's dessas duas conexões desempenharam pior do que as outras três conexões. Podemos perceber isso analisando-se a tabela 4.6. Nela, C1 e C2 tiveram poucos MP's com PC máxima acima de 0.9, o que leva a crer que seus MP's poderiam ser considerados menos robustos que os MP's das outras conexões. Este resultado pode ser também causado pela competição interna de C1 e C2.

6.4 Padrões semelhantes e seus resultados

Foi observado em 4.2.1 e 4.2.5 que padrões semelhantes tendem a apresentar resultados semelhantes quando é feito aprendizado com ϵ lat de mesmo sinal, mas resultados muito diferentes com ϵ lat de sinal inverso. A semelhança nos resultados era o esperado, pois não havia indícios do contrário, entretanto, não foi o que ocorreu. É provável que as diferenças causadas pelo uso de ϵ lat's inversos sejam originadas da competição que, sugere-se, este tipo de condição introduzir no sistema, mas não há certeza.

6.5 O patamar

O significado do patamar não é perfeitamente claro. O mais provável é que represente simplesmente o ponto onde não importa mais qual a intensidade de aprendizado lateral que se utiliza: qualquer aumento nas conexões laterais não terá mais reflexo na Probabilidade de Ciclagem. Há uma observação interessante: em nenhuma das simulações onde a PC atingiu um patamar este foi imediatamente precedido de um declive. Em todos os casos, a PC estava subindo e então estacionou. Sendo assim, o patamar representa sempre a mais alta PC que pode ser alcançada quando consideramos apenas os valores de ϵ lat próximos ao patamar.

6.5.1 Possível desestruturação do modelo com aprendizado lateral muito intenso

Uma das características mais importantes do modelo do *trion* é a sua estruturação em colunas mais conectadas internamente do que lateralmente, o que tem embasamento eletrofisiológico (Leng et al., 1994). Essa característica, aliada a influência de dois passos de tempo na evolução do sistema, garante o grande repertório de padrões que identifica o modelo. Como dito acima, é possível que um aprendizado lateral muito intenso venha a corromper essa estrutura, criando uma nova.

Quando é realizado aprendizado interno, conexões fortes, as internas, são alteradas, levando a um aumento na PC. A alteração pode tanto ser para cima ou para baixo, mas a flutuação em torno do valor inicial normalmente é pequena (nos trabalhos até agora realizados (Shaw et al., 1988). Por outro lado, com o aprendizado lateral, conexões inicialmente fracas (ou até nulas, como nesse trabalho), são intensificadas. Para baixas intensidades desse tipo de aprendizado, a estrutura do modelo se mantém, mas com a utilização de valores elevados de ϵ_{lat} , é possível que as conexões laterais atinjam valores até uma ordem de grandeza maiores do que as internas, mudando a estrutura do sistema. Se a estrutura era, por exemplo, três colunas de seis *trions* cada, pode acabar se tornando seis colunas de três *trions* cada.

Esse tipo de mudança estrutural é tão significativo que, além de provavelmente alterar todas as propriedades intrínsecas do modelo, torna-o comparável a modelos desestruturados, que partem de uma *tabula rasa* ou de conexões aleatórias, como boa parte dos modelos inspirados no de Hopfield (Hopfield, 1982).

É possível que o patamar atingido pela Probabilidade de Ciclagem represente o ponto onde a intensidade das conexões laterais, modificadas pelo aprendizado, altere a estrutura do modelo, alterando, assim, as suas propriedades. É interessante questionar se, no patamar (ou na desestruturação, se ambos não estiverem associados), ainda há outros MP's no repertório daquela conexão.

6.6 PC máxima igual a 1.0

Todas as simulações que atingiram Probabilidade de Ciclagem igual a 1.0, estabilizaram nela como patamar. Dessas 10 simulações, apenas uma foi obtida com elat de sinal inverso, o que confirma o caráter menos estável associado a esta condição. Uma Probabilidade de Ciclagem igual a 1.0 indica que, mesmo que estivesse sendo utilizado o protocolo Monte Carlo para definição do estado dos *trions*, o MP sempre seria mantido, sem chance de evolução para outros MP's, como freqüentemente ocorre com esse protocolo. Isto, entretanto, ainda precisa ser verificado.

Uma questão interessante de ser levantada é se, de maneira semelhante ao indagado na seção 6.5.1, há outros MP's ainda no repertório quando a PC 1.0 é atingida. Talvez a "absoluta robustez" do MP que atinja 1.0 leve à eliminação da chance de outros MP's também ocorrerem no repertório. Isto também precisa ser verificado.

6.7 Utilidade do aprendizado lateral

Em uma eventual utilização do modelo do *trion* como base para um aplicativo com capacidade de aprendizado, cada MP do repertório possivelmente codificará uma ação, uma informação ou uma seqüência das anteriores que deverá poder ser aprendida. É importante, então, que, ao realizar-se o aprendizado, a Probabilidade de Ciclagem efetivamente seja aumentada.

Esse desejado aumento, entretanto, não ocorreu consistentemente e de maneira utilizável com o uso de aprendizado lateral, mostrando a complexidade surgida a partir da competição entre dois níveis de arquitetura conexional em um sistema, outrossim, simples. A maior parte dos MP's, quando aprendidos, respondia diferentemente a um determinado valor de elat. Para alguns, um elat igual a 0.25 resultava em uma PC menor do que a inicial; para outros, a mesma intensidade de aprendizado resultava em um considerável aumento na PC. Se fosse feito um aplicativo tentando tirar proveito do aprendizado lateral, para cada ação codificada em um MP seria necessária a verificação de qual a elat mais eficiente para o aprendizado daquele MP, o que com certeza dificultaria não só a programação do aplicativo como também a sua velocidade de execução.

Sendo assim, é provável que o aprendizado lateral não seja muito útil para aplicativos baseados no modelo do *trion*. O aprendizado interno, pelo que se sabe até agora, parece um melhor candidato a sistema de aprendizado para utilização prática.

Não se deve, entretanto, descartar o uso de conexões laterais, apenas o aprendizado nelas é que deve ser questionado. Esse tipo de conexão é responsável pela definição de padrão “em fase” ou “fora de fase”, pois, sem seu uso, a noção de fase não existe, já que esta depende das interações entre colunas, determinadas pelas conexões laterais.

6.8 Considerações finais

6.8.1 Para simulações futuras

Apesar de, classicamente, MP's fora de fase serem mais estáveis do que MP's em fase, nesse trabalho foi observado que essa expectativa não se aplica quando é realizado aprendizado lateral com ϵ 's de sinais inversos para X e Y. É uma sugestão para trabalhos futuros, então, que sejam esperados valores mais altos de PC máxima para MP's em fase do que para fora de fase, ou que pelo menos tenha-se em mente essa possibilidade.

6.8.2 Resultados mais significativos deste trabalho

Dentre os resultados obtidos neste trabalho, o mais importante provavelmente é a observação de diferenças no comportamento do modelo com elat de mesmo sinal e de sinal inverso. Relacionado a isto está a possibilidade de competição introduzida não só pelo aprendizado como também pela utilização de conexões internas “pouco intensas”.

6.8.3 Modelos

É interessante observar como algo (um modelo) criado para facilitar a compreensão de um processo ou mecanismo indecifrado (uma caixa-preta), por vezes comporta-se ele próprio como uma caixa-preta, produzindo resultados inesperados para os pesquisadores. Entretanto, talvez seja este o objetivo: quando um modelo se comportar tão imprevisivelmente quanto o objeto modelado, é provável que o modelo seja uma representação, mesmo que um tanto simples, do objeto. Não percam as esperanças, pois.

7. Bibliografia

ASHBY, W.R. *Design for a Brain*, 2^a impr. Londres: Chappman & Hall, 1954.

ASHBY, W.R. *An Introduction to Cybernetics*. Londres: Chappman & Hall, 1956.

ASHBY, W.R. A variedade indispensável e suas implicações no controle de sistemas complexos. Trad. Leonidas Hegenberg e Octanny Silveira da Mota. In: EPSTEIN, I. *Cibernética e Comunicação*. São Paulo: Cultrix, 1973. Tradução de : Requisite varieties and its implications for the control for complex systems. In: *Cybernetica*, v. 1, n. 2. Namur: Association Internationale de Cybernétique, 1958.

BACH, M. e KRUGER, J. Correlated neuronal variability in monkey visual cortex revealed by a multi-microelectrode. *Exp. Brain Res.*, n. 61, p. 451-456, 1986.

BALANIUK, R. *Modelos de Redes Neurais: Conceitos Básicos e Evolução Histórica*. Porto Alegre: UFRGS. CPGCC, 1989.

BEAR, M.F., CONNORS, B.W. e PARADISO, M.A. *Neuroscience: Exploring the Brain*. Baltimore: Williams & Wilkins, 1996.

BLOCK, H.D. The Perceptron: a model for brain functioning. *Reviews of Modern Physics*, v. 34, p. 123-35, 1962.

BODNER, M., ZHOU, Y.-D., SHAW, G.L. & FUSTER, J.M. Symmetric temporal patterns in cortical spike trains during performance of a short-term memory task. No prelo, 1997.

CARPENTER, C.A. e GROSSBERG, S. ART 2: self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, v. 26:49/9 - 39, 1987.

CRAGG, B.G. e TEMPERLEY, H.N.V. The organization of neurones: a cooperative analogy. *EEG Clin. Neurophysiol.*, v. 6, p. 85-92, 1954.

EDELMAN, G.M. Group selection and phasic reentrant signaling: a theory of higher brain function. In: EDELMAN, G.M. e MOUNTCASTLE, V.B. (eds.). *The Mindful Brain*. Cambridge: MIT, p. 51-100, 1978.

EDELMAN, G.M. *Neural Darwinism: The Theory of Neuronal Group Selection*. Nova Iorque: Basic Books, 1987.

FISHER, M.E. e SELKE, W. Infinitely many commensurate phases in a simple Ising model. *Phys. Rev. Lett.*, v. 44, p. 1502-1505, 1980.

GOLDMAN-RAKIC, P.S. Modular organization of prefrontal cortex. *Trends in Neurosci.*, v. 7, p. 419-424, 1984.

HEBB, D.O. *The Organization of Behavior*. Nova Iorque: Wiley, 1949.

HINTON,, G., SEJNOWSKI, T. e ACKLEY, D. *Boltzmann Machines: constraint satisfaction networks that learn*. Carnegie-Mellon University, Tech. Report CMU-CS-84-119, 1984.

HOPFIELD, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. EUA*: v. 79, p. 2554-2558, 1982.

KANDEL, E.R., SCHWARTZ, J.H. e JESSELL, T.M. *Principles of Neural Science*, 3rd ed. EUA: Appleton & Lange, 1991.

KATZ, B. *The Release of Neural Transmitter Substances*. Springfield: Thomas, 1969.

KUHN, T.F. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962.

LENG, X., McGRANN, J.V., QUILLFELDT, J.A., SHAW, G.L. e SHENOY, K.V. Learning and memory processes and the modularity of the brain. In: DELACOUR, J. (ed.). *Neural Bases of Learning and Memory*. World Scientific, 1994.

LITTLE, W.A. Existence of persistent states in the brain. *Math. Biosci.*, v. 19, p. 101-120, 1974.

LITTLE, W.A. e SHAW, G.L. Analytic study of the storage capacity a neural network. *Math. Biosci.*, v. 39, p. 281-290, 1978.

McCULLOCH, W. & PITTS, W. A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, p. 115-133, 1943.

MELLO-E-SOUZA, T. *Estudos com o Modelo do Trion de Processamento Cortical de Informação*. Porto Alegre: UFRGS. Dissertação (Bacharelado em Ciências Biológicas, ênfase em Biofísica) - Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 1993.

MINSKY, M. e PAPERT, S. *Perceptrons: an Introduction to Computational Geometry*. Cambridge: MIT, 1969.

MOUNTCASTLE, V.B. An organizing principle for cerebral function: the unit module and the distributed system. In: EDELMAN, G.M. e MOUNTCASTLE, V.B. (eds.). *The Mindful Brain*. Cambridge: MIT, p. 51-100, 1978.

NEUMANN, J. von. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In: SHANNON, C.E. & McCARTHY, J. (eds.) *Automata Studies*. Princeton University, p. 43-98, 1956.

NORTH, G. A celebration of connectionism. *Nature*, v. 328, p. 107, 1987.

OSÓRIO, F.S. Simulação de redes artificiais de neurônios com aprendizado supervisionado. *Scientia*, v. 3, p. 45-66, 1992.

PATERA, J., SHAW, G.L., SLANSKY, R. e LENG, X. Model of Adaptive Temporal Development of Structured Finite Systems. *Phys. Rev.*, v. 40 (2), p. 1073-87, 1989.

QUILLFELDT, J.A. *O papel dos receptores glutamatérgicos do tipo AMPA na expressão da memória no córtex entorrinal e estruturas relacionadas*. Porto Alegre: UFRGS. Tese (Doutorado em Ciências Biológicas: Fisiologia) - Instituto de Biociências, UFRGS, 1994. Apêndice, p. 107-131. *O modelo do Trion de redes neurais*.

ROSENBLATT, F. The Perceptron, a probabilistic model for information storage and organization in the brain. *Psychobiological Review*, V. 65, p.386-408, 1958.

SHAW, G.L., HARTH, E. e SCHEIBEL, A.B. Cooperactivity in brain function: Assemblies of approximately 30 neurons. *Exper. Neuro.*, v. 77, p. 324-358, 1982.

SHAW, G.L. e PEARSON, J.C. Information processing in the cortex: the role of small assemblies of neurons. In: MINTZ, S.L. e PERLMUTTER, A. (eds.) *Information Processing in Biological Systems*. Nova Iorque: Plenum, p. 1-31, 1986.

SHAW, G.L., SILVERMAN, D.J. e PEARSON, J.C. Model of cortical organization embodying a basis for a theory of information processing and memory recall. *Proc. Natl. Acad. Sci.*, v. 82, p. 2364-2368, 1985.

SHAW, G.L., SILVERMAN, D.J. e PEARSON, J.C. Trion model of cortical organization: toward a theory of information processing and memory. In: PALM, G. e AERSTSEN, A. (eds.). *Brain Theory*. Berlin: Springer-Verlag, p. 177-191, 1986.

SHAW, G.L., SILVERMAN, D.J. e PEARSON, J.C. Trion model of cortical organization and the search for the code of short-term memory and information processing. In: DELACOUR, J. e LEVY, J.C.S. (eds.). *Systems with Learning and Memory Abilities*. North-Holland: Elsevier Science Publishers B.V., 1988.

SHAW, G.L. e VASUDEVAN, R. Persistent states of neural networks and the random nature of synaptic transmission. *Math. Biosci.*, v. 21, p. 207-218, 1974.

SHAW, G.L., SARNTHEIN, J., vonSTEIN, A., RAPPELSBERGER, P., PETSCHKE, H., RAUSCHER, F.H. Persistent patterns of brain activity: an EEG coherence study of the positive effect of music on spatial-temporal reasoning. *Neurological Research*, v. 19, p. 107-116, 1997.

SHENOY, K.V. *Monte Carlo Evolution of a Trion Model Repertoire*. Irvine: UCI, 1988. Dissertação (Graduação) - University of California, 1988.

SILVERMAN, D.J., SHAW, G.L. e PEARSON, J.C. Associative recall properties of the Trion model of cortical organization. *Biol. Cybern.*, v. 53, p. 259-271, 1986.

TURING, A. On computable numbers, with an application to the *entscheidungproblem*. *Proceeding of the London Mathematical Society*, 2nd series, v. 42, p. 230-265, 1936.

WIENER, N. *Cybernetics*. Cambridge: MIT, 1948.

Lista de figuras

Fig. 1.1 - Representação dos estados de atividades de um *trion* (reproduzido de Mello-e-Souza, 1993)

Fig. 1.2 - Representação esquemática de uma coluna de quatro *trions* e das conexões com seus vizinhos (reproduzido de Mello-e-Souza, 1993)

Fig. 1.3 - Mapa de conexões para um determinado *trion* (reproduzido de Mello-e-Souza, 1993)

Fig. 1.4 - Exemplo de MP

Fig. 1.5 - Exemplo de interação entre *trions* de diferentes colunas

Fig. 1.6 - Exemplo de MP global composto pelo mesmo MP local em três colunas fora de fase

Lista de tabelas

Tab. 1.1 - Exemplo de conexões entre *trions*

Tab. 3.1 - Conexões internas utilizadas neste trabalho

Tab. 4.1 - Dois MP's semelhantes de C1

Tab. 4.2 - MP's rotados

Tab. 4.3 - Dois MP's semelhantes de C5

Tab. 4.4 - Percentagem dos MP's que atingiram um patamar, de acordo com as condições testadas

Tab. 4.5 - Percentagem dos MP's fora de fase que tiveram PC máxima maior do que suas versões em fase

Tab. 4.6 - Percentagem de MP's acima de 0.9 por conexão

Tab. 4.7 - Percentagem de MP's que atingiram PC igual a um por conexão e por tipo de elat

Tab. 6.1 - Valores das conexões internas para os cinco padrões de conexão

Apêndice


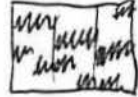
Como apêndice, aqui estão todos os gráficos plotados a partir dos dados obtidos nas simulações feitas para este trabalho.

Cada um dos cinco tipos de conexão está representado pelos dois primeiros caracteres do nome de cada gráfico: c1, c2, c3, c4 e c5.

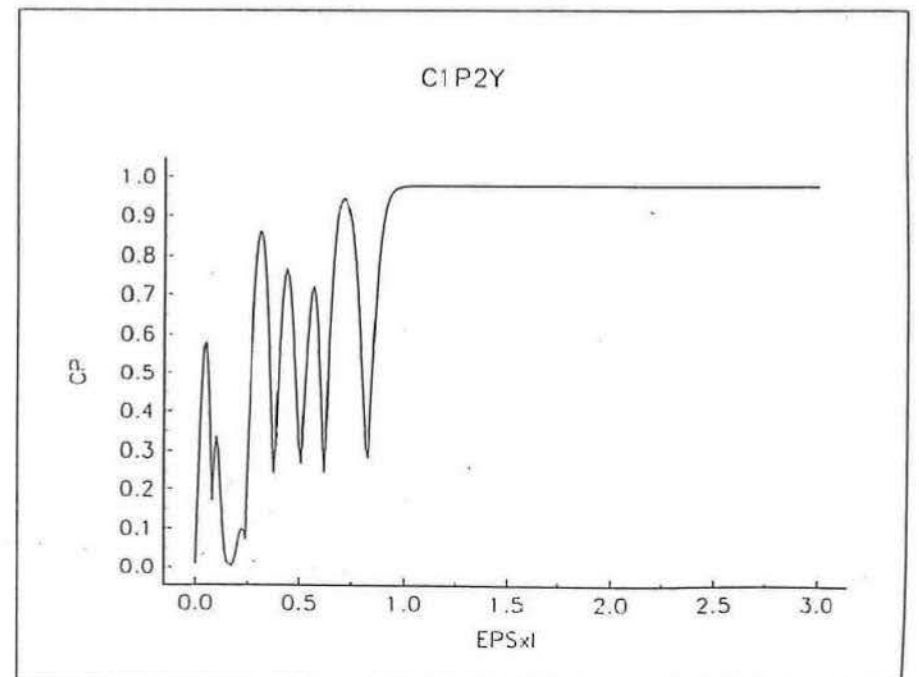
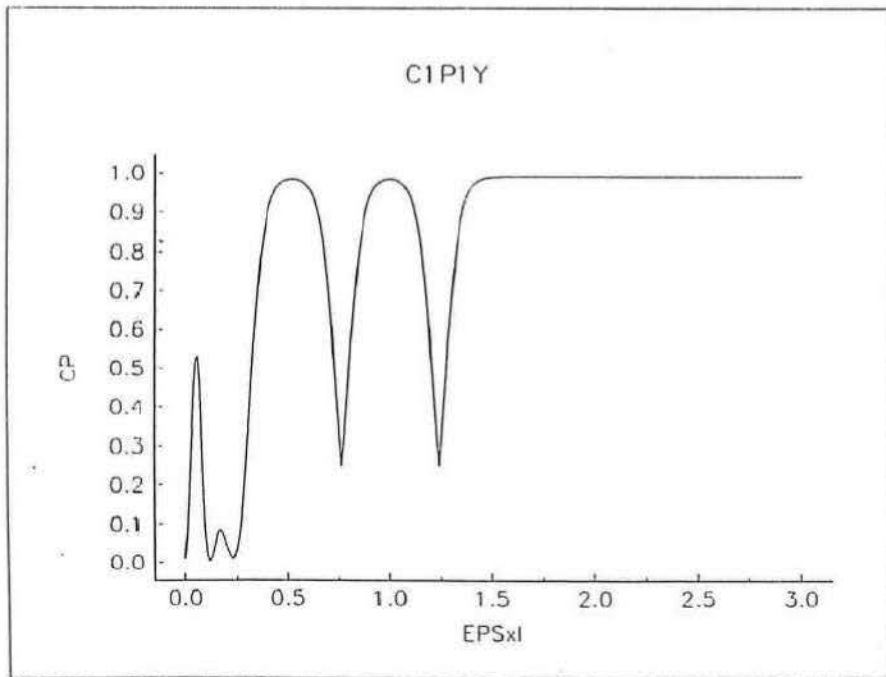
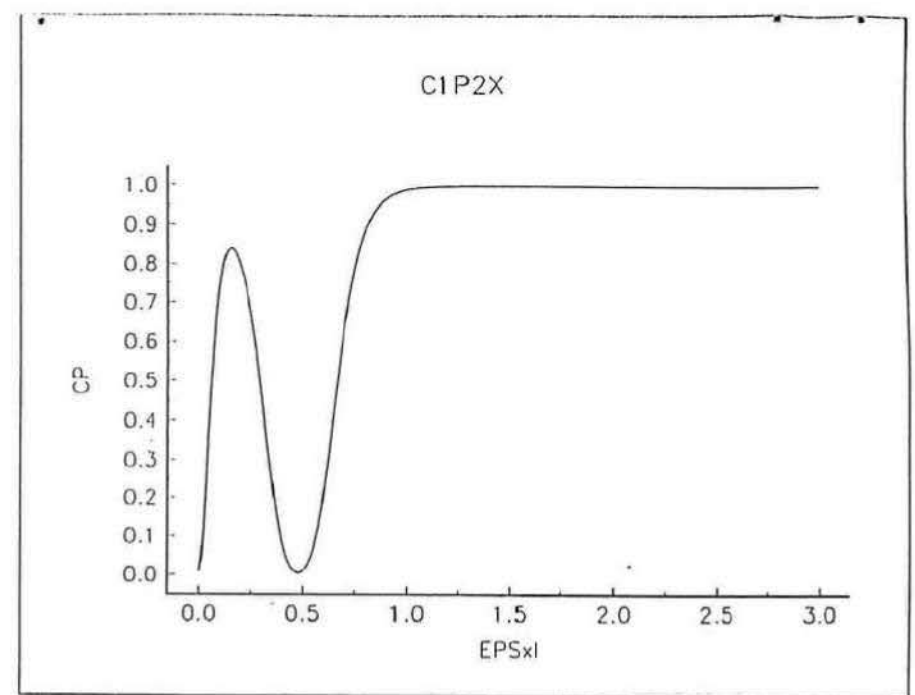
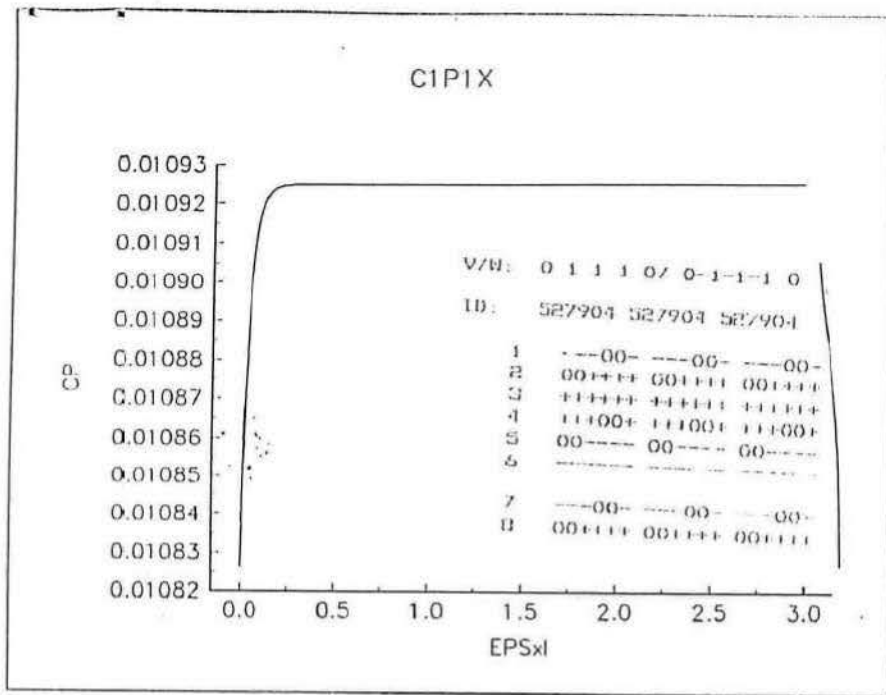
Em cada uma das folhas há os gráficos para um determinado MP nas quatro condições testadas. Os gráficos com "P" ímpar (p1, p3, p5, p7 e p9) representam o MP global com os MP's locais em fase. Os gráficos com "P" par (p2, p4, p6, p8 e p10) representam o MP global com os MP's locais fora de fase um passo de tempo.

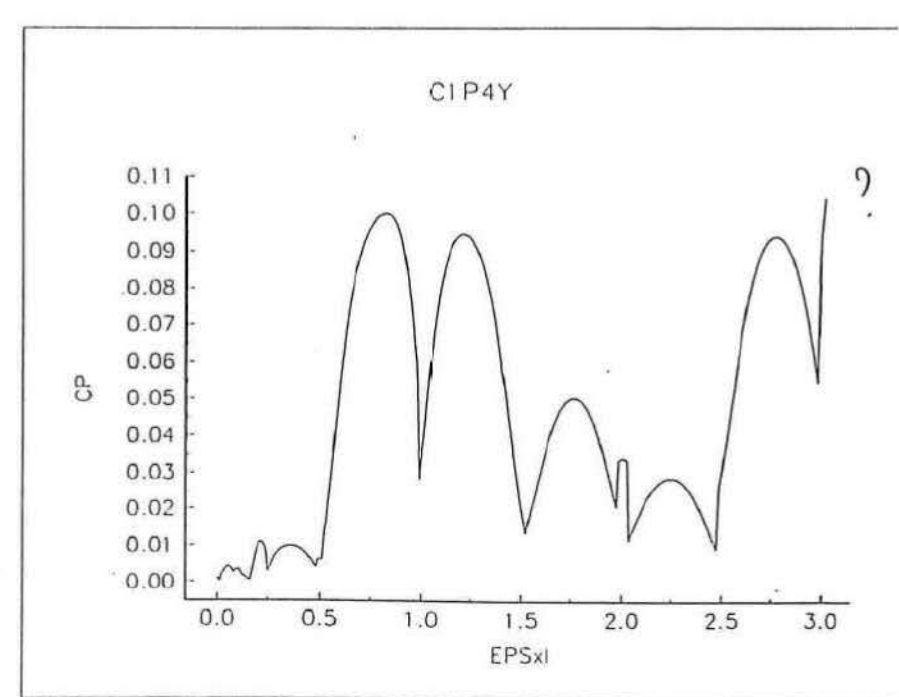
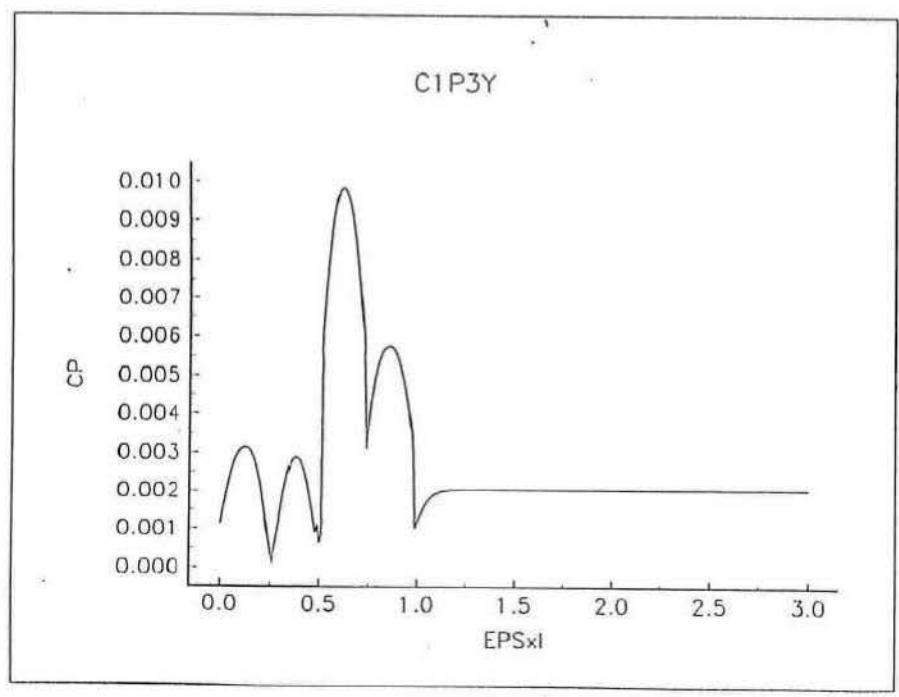
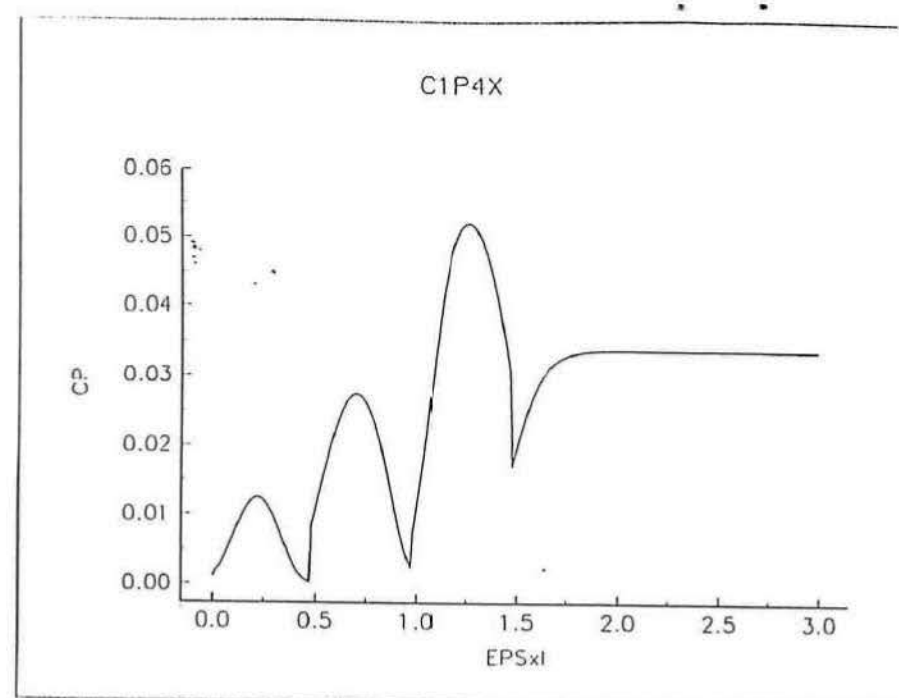
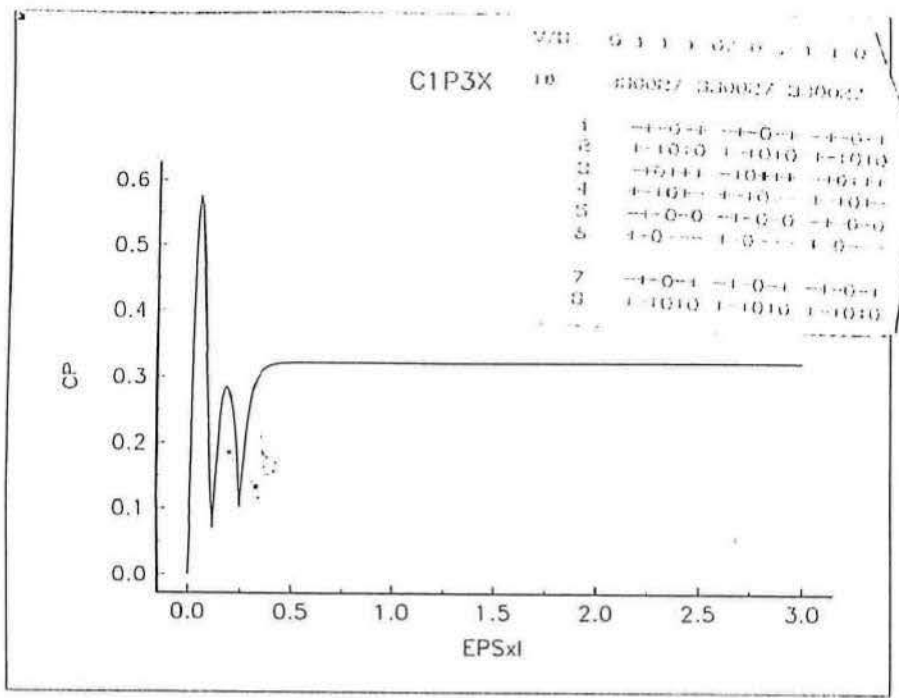
O último caractere dos nomes indica a condição do elat. O "X" indica elat's de mesmo sinal; o "Y", elat de valor inverso para o segundo passo de tempo.

Por exemplo, os gráficos do quarto MP do terceiro tipo de conexão são os denominados c3p7x, c3p8x, c3p7y e c3p8y.

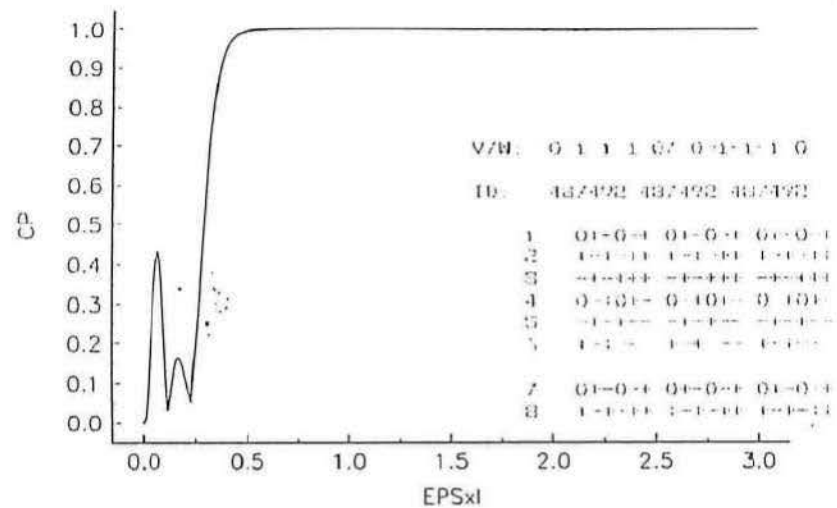
	MPs em fase	MPs fora de fase
elat \Rightarrow sinal invar	c1p7x 	c1p8x 
elat \Rightarrow sinal oposto	c1p7y 	c1p8y

CHAVE

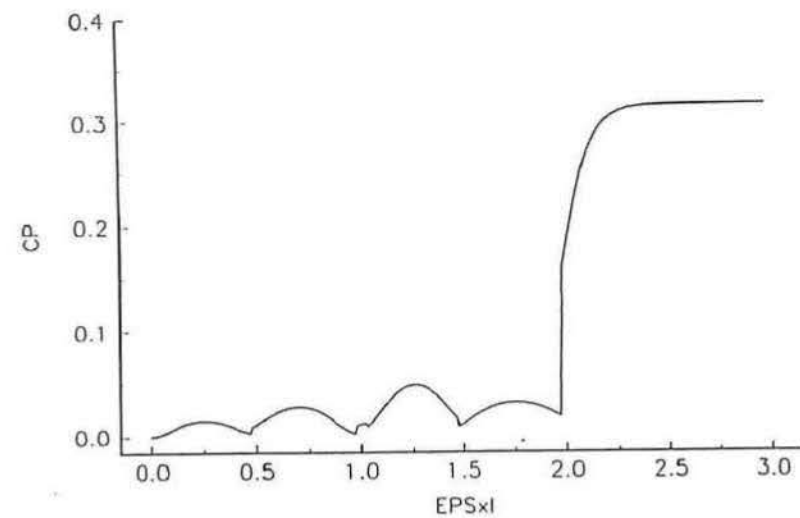




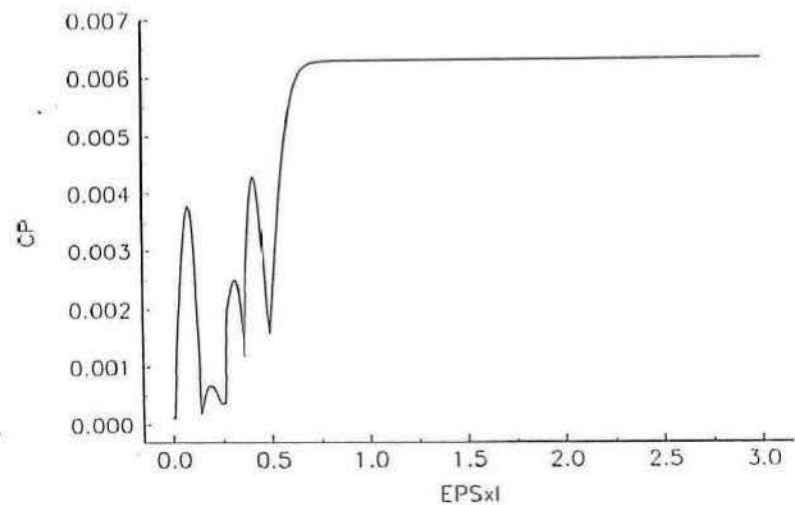
C1P5X



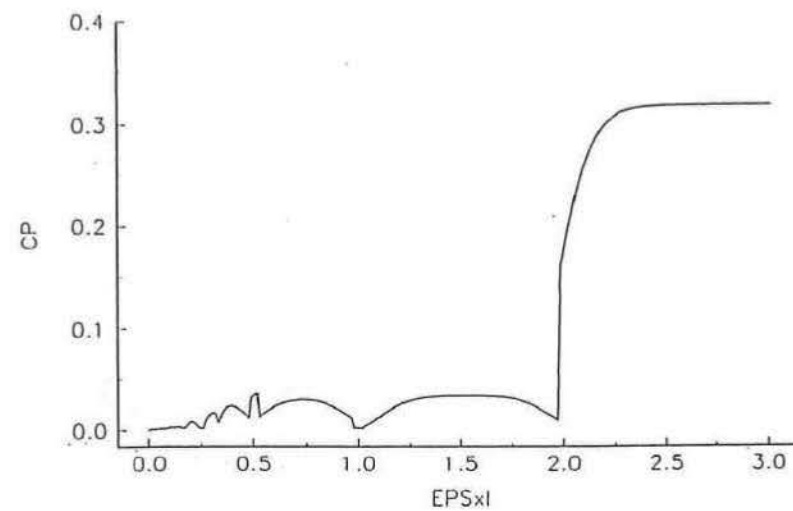
C1P6X



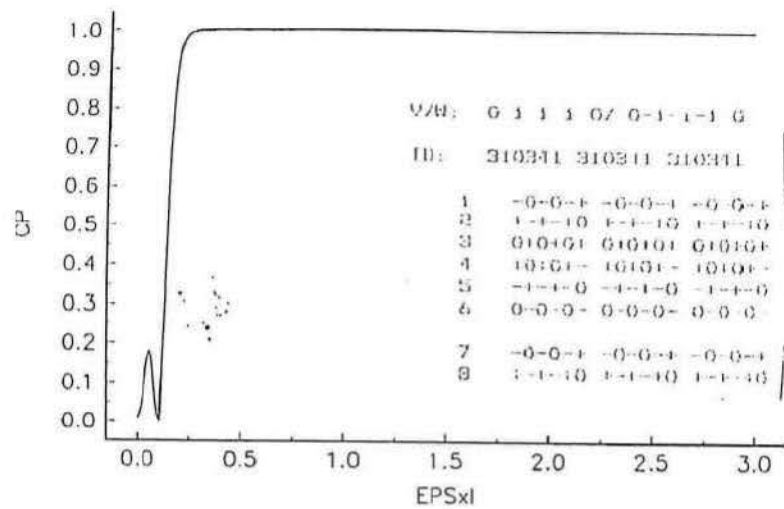
C1P5Y



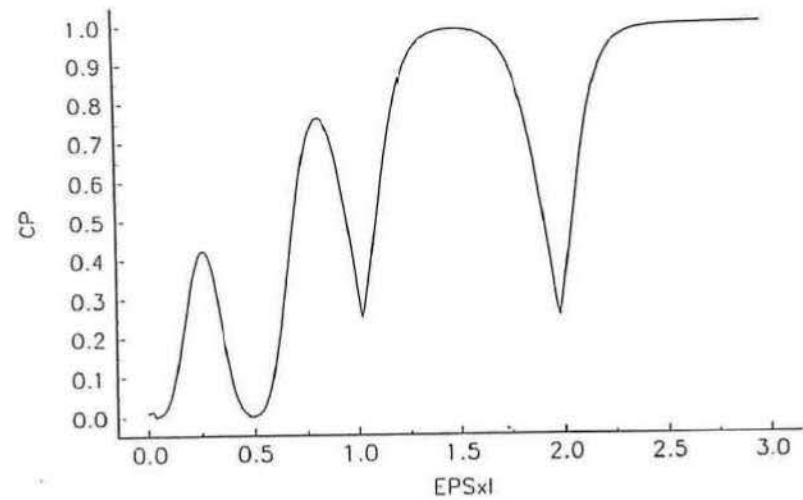
C1P6Y



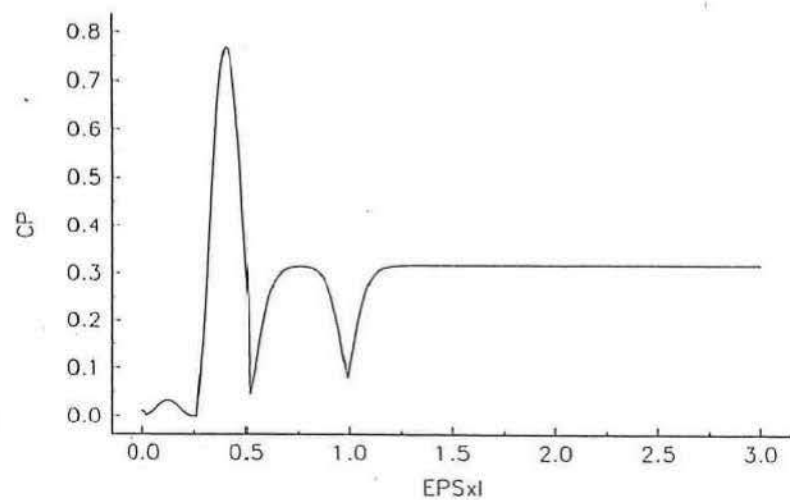
C1P7X



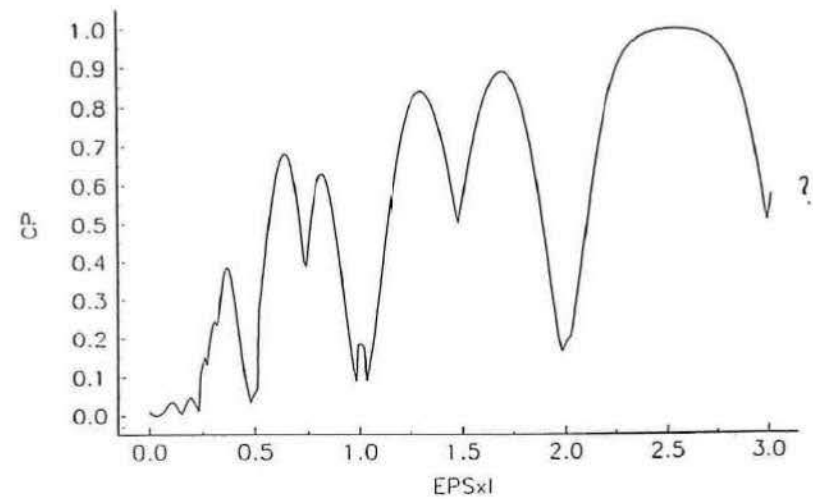
C1P8X



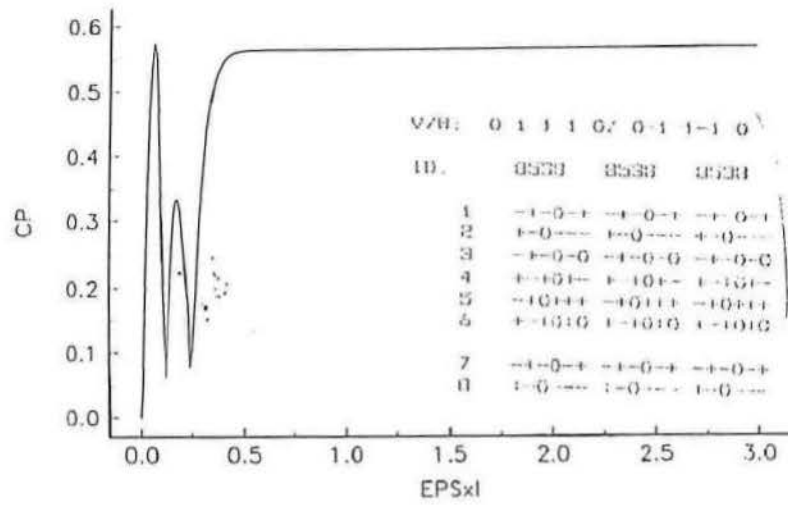
C1P7Y



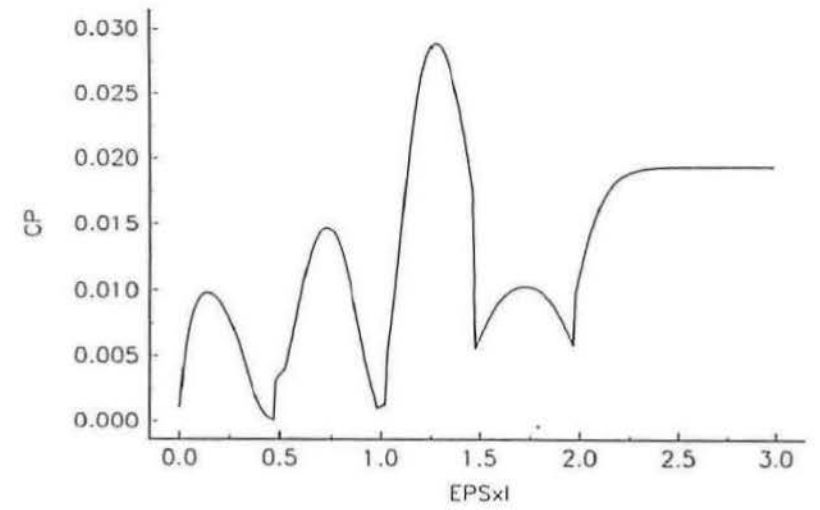
C1P8Y



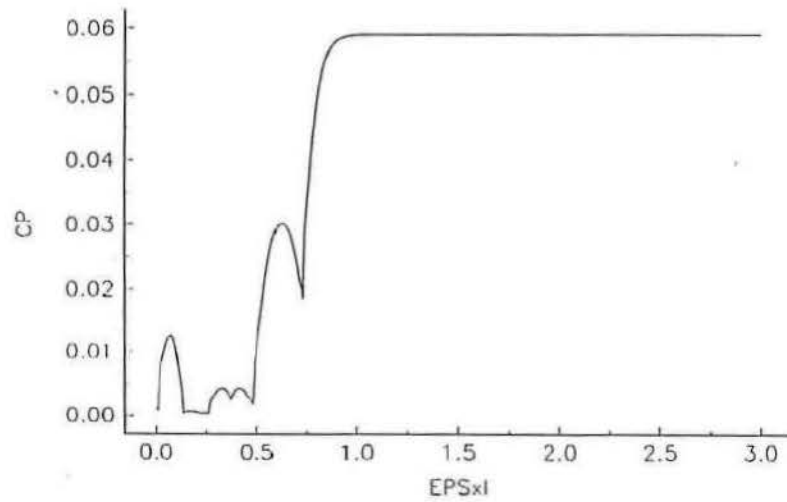
C1P9X



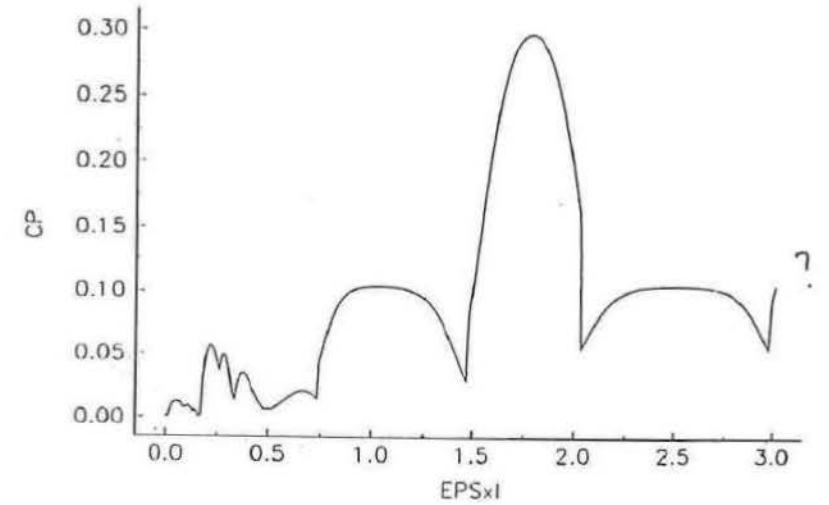
C1P10X



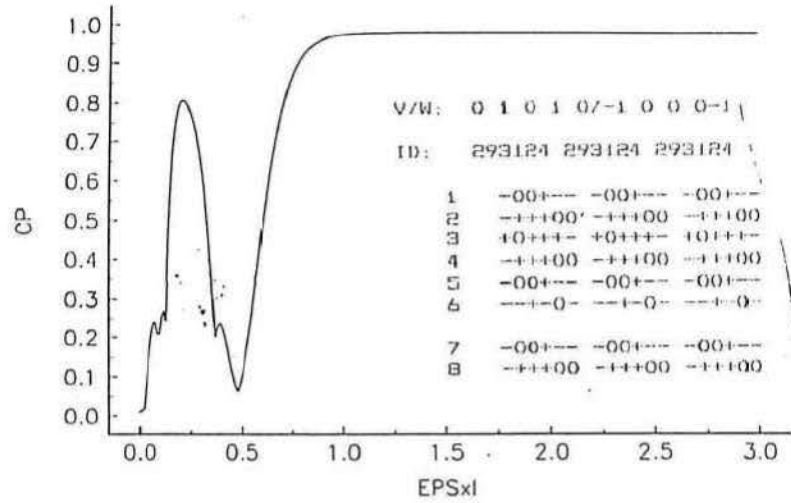
C1P9Y



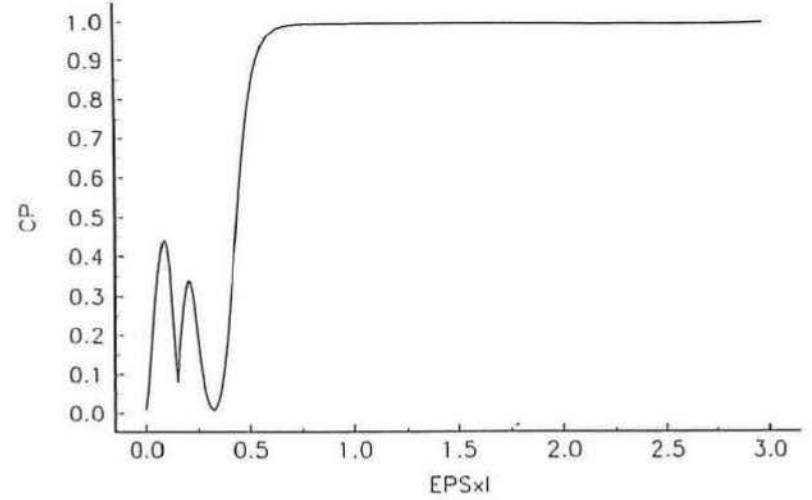
C1P10Y



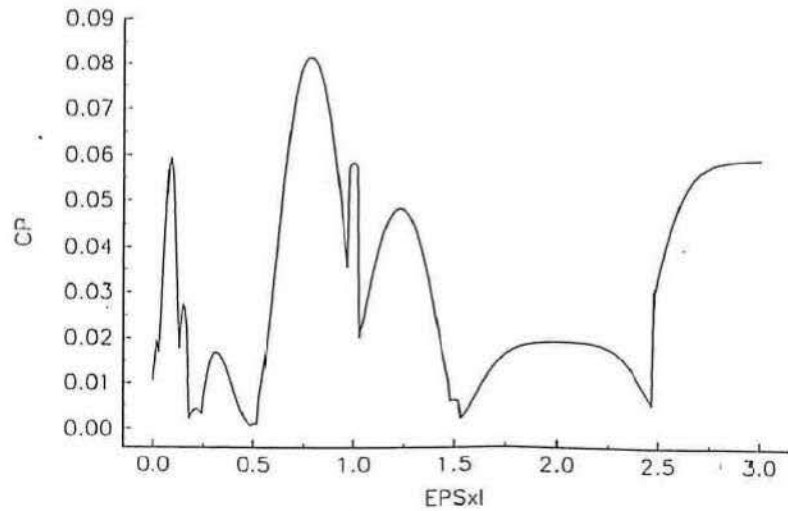
C2P3X



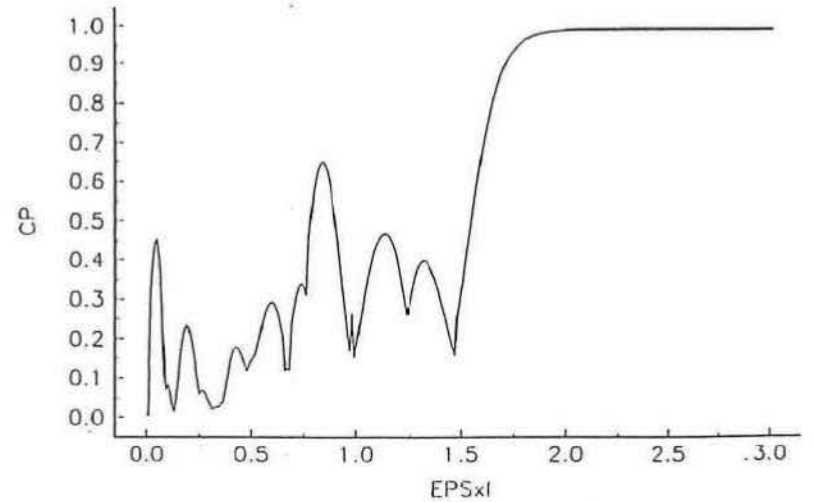
C2P4X



C2P3Y

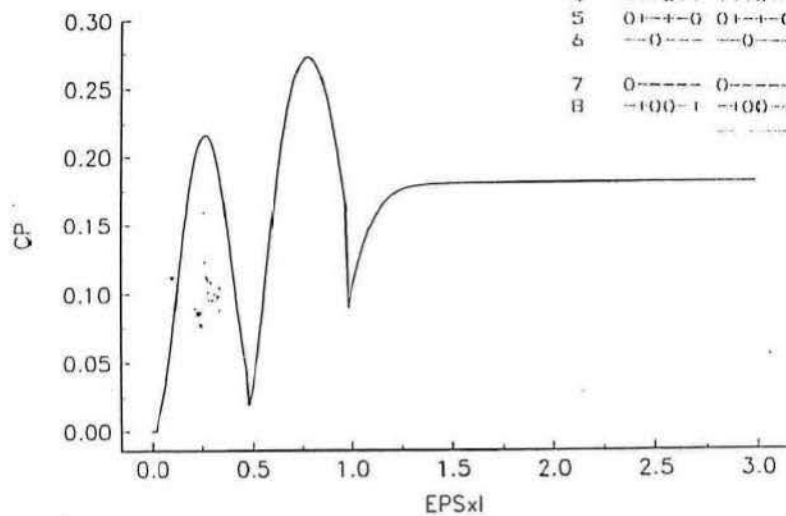


C2P4Y



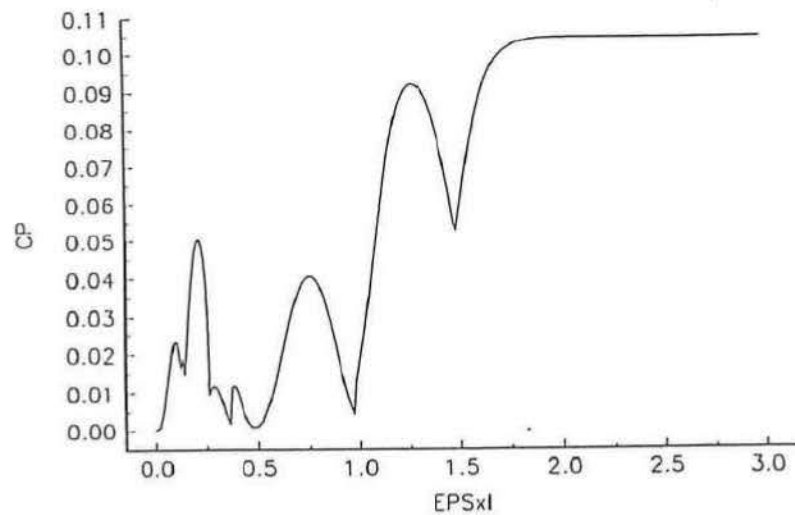
ID: 304913 304913 304913

C2P5X

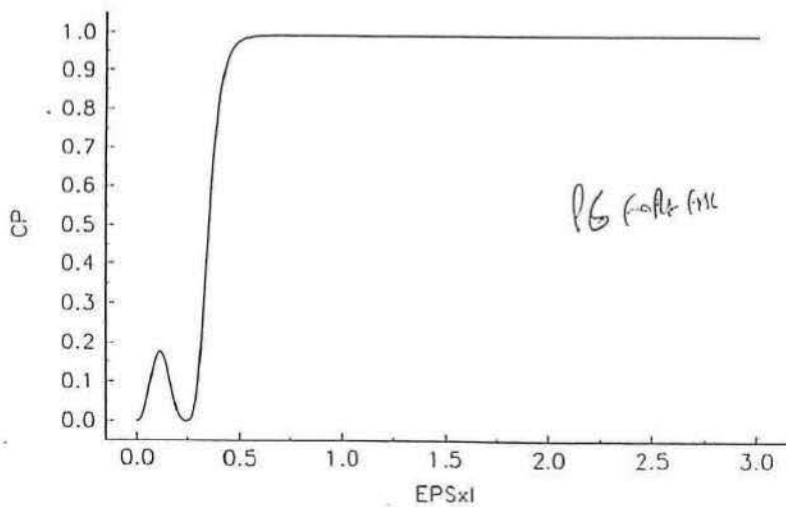


1	0-----0-----0-----
2	--+00--+--+00--+--+00--+
3	+++++0+++++0+++++0
4	+++0+++0+++0+++0+++0+++
5	0+--+0 0+--+0 0+--+0
6	--0-----0-----0-----
7	0-----0-----0-----
8	--+00--+--+00--+--+00--+

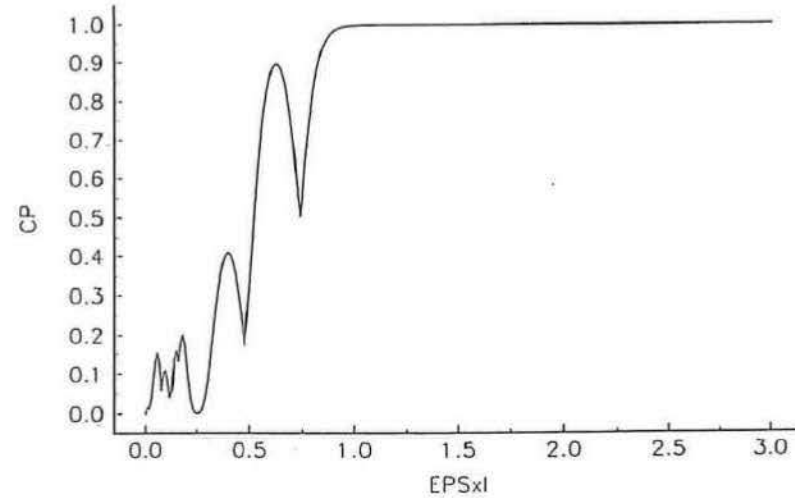
C2P5Y



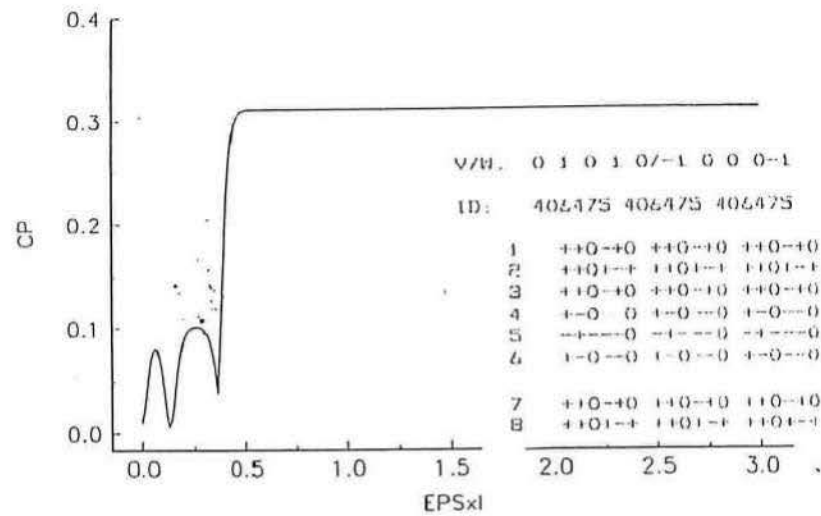
C2P6X



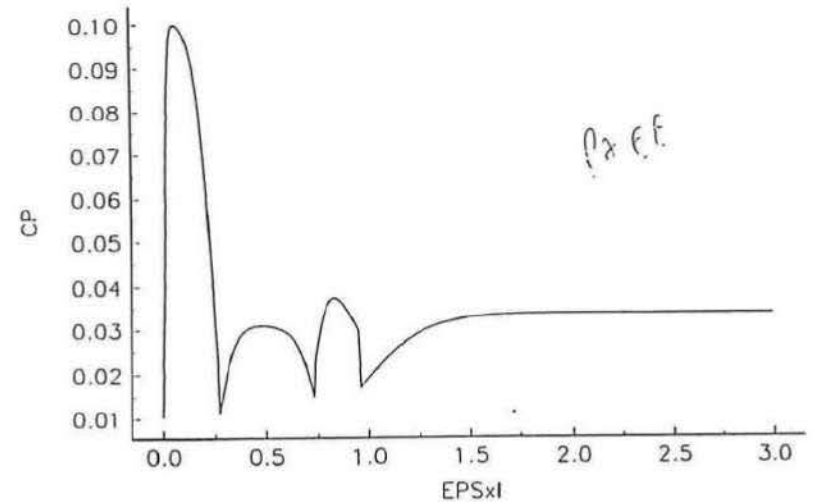
C2P6Y



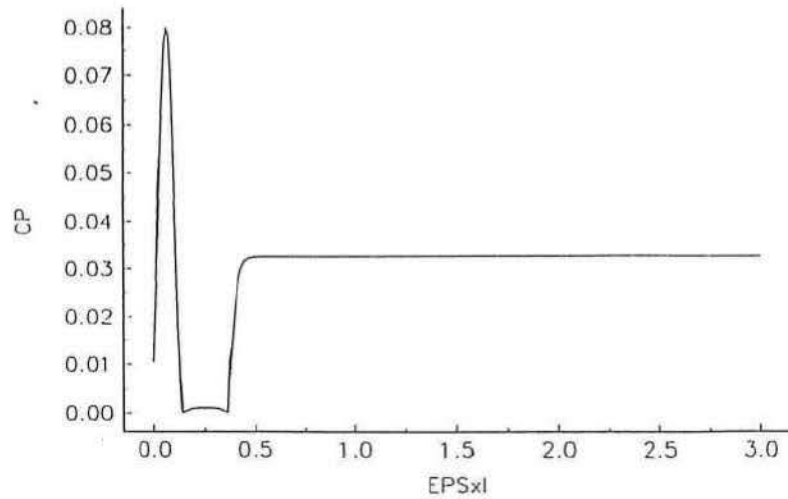
C2P7X



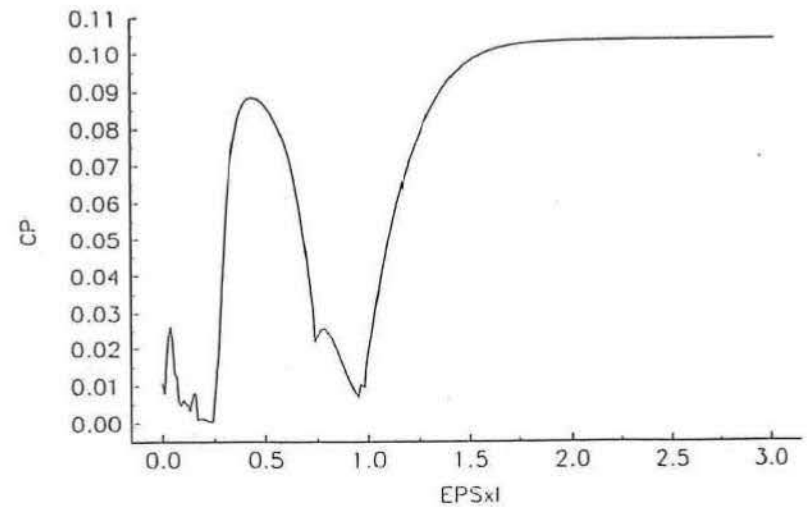
C2P8X

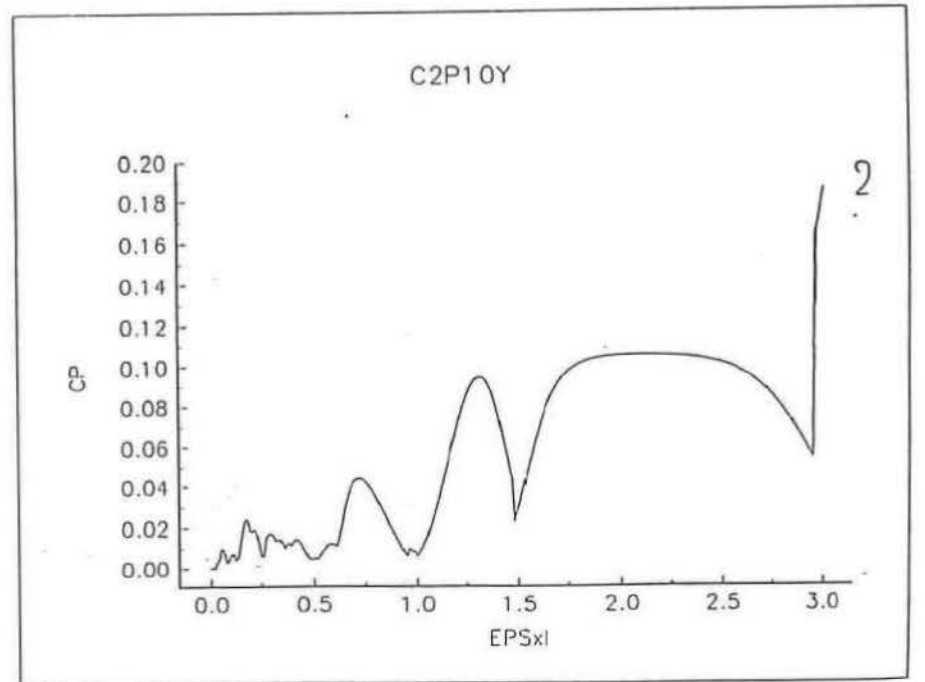
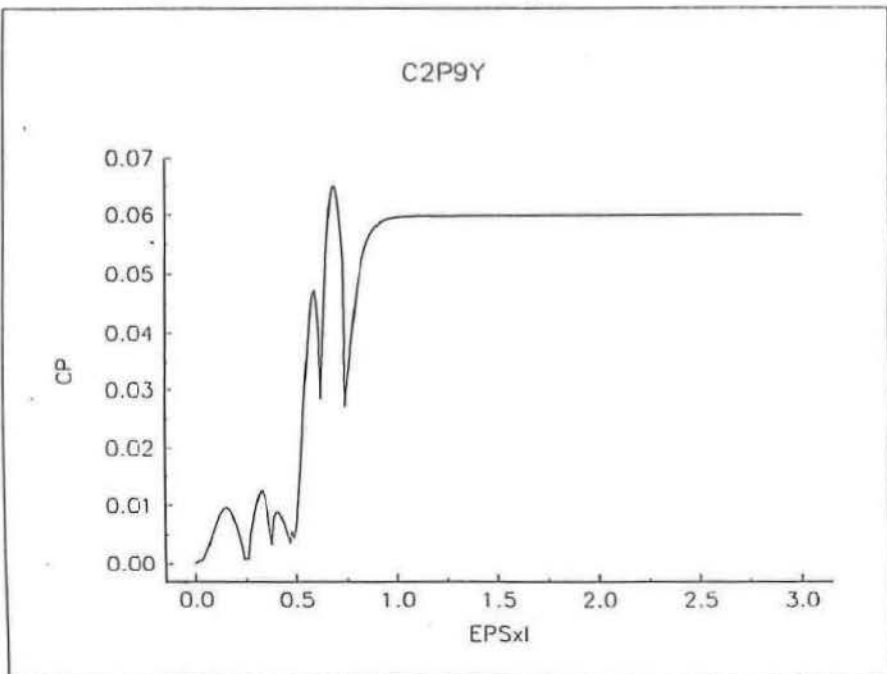
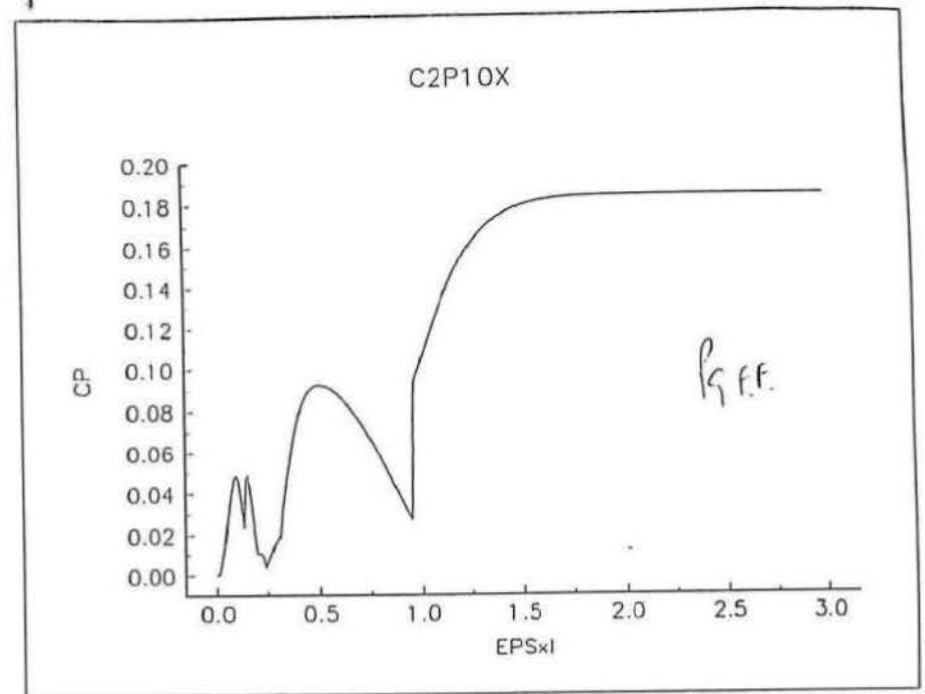
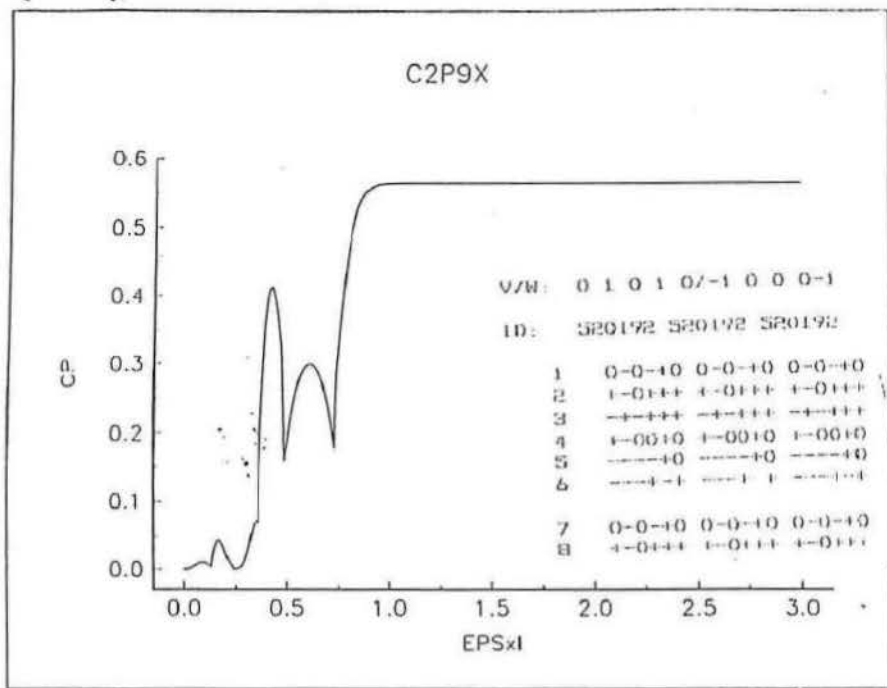


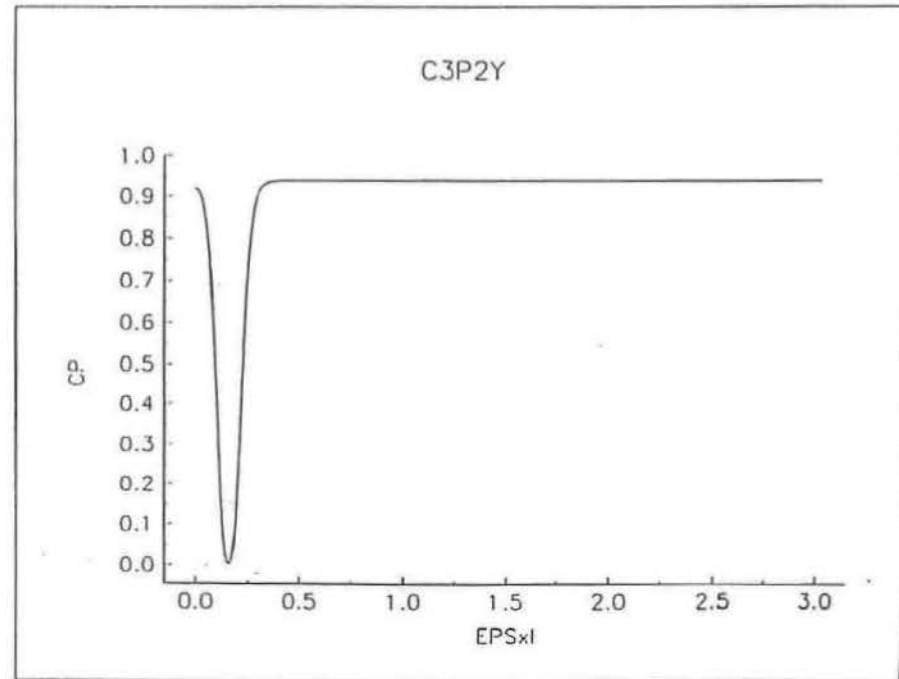
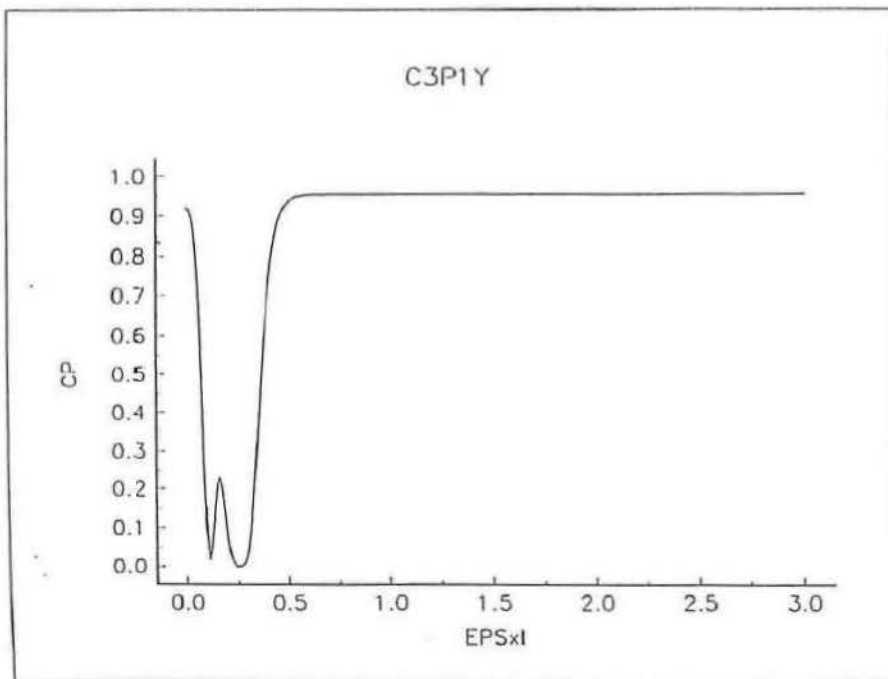
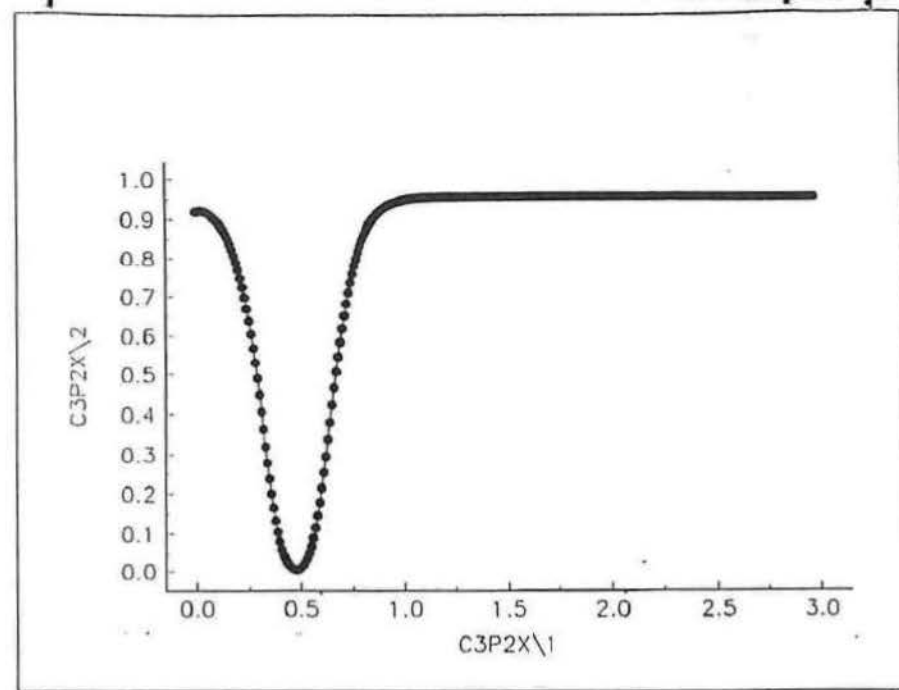
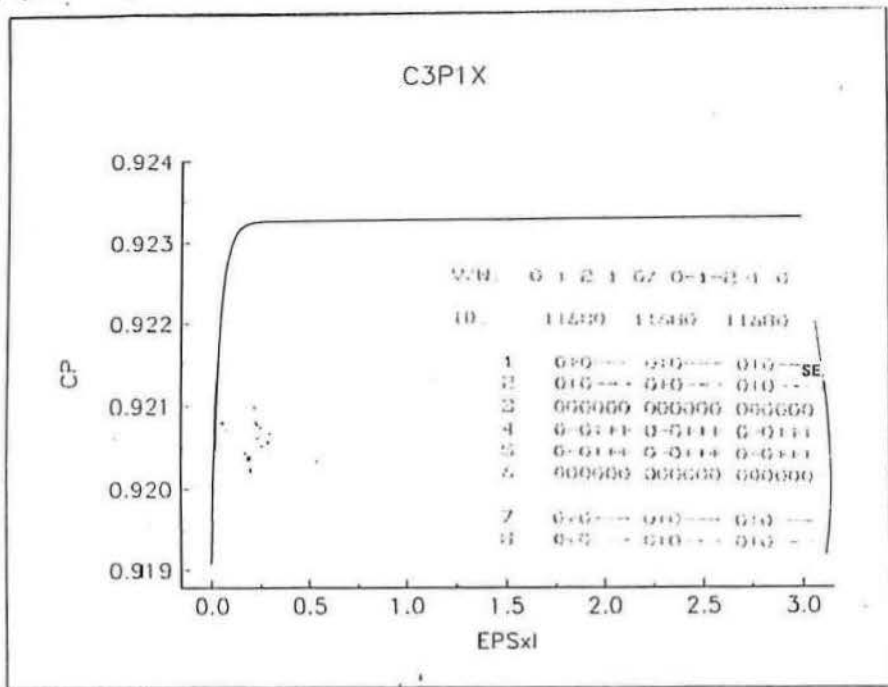
C2P7Y



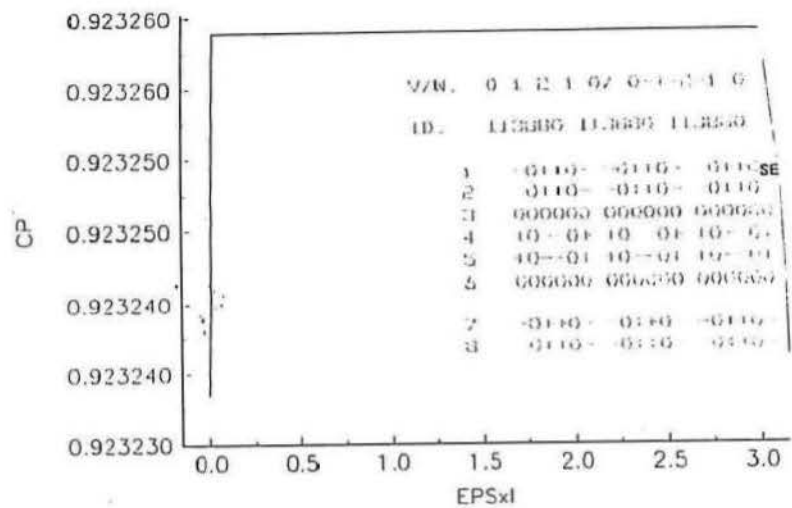
C2P8Y



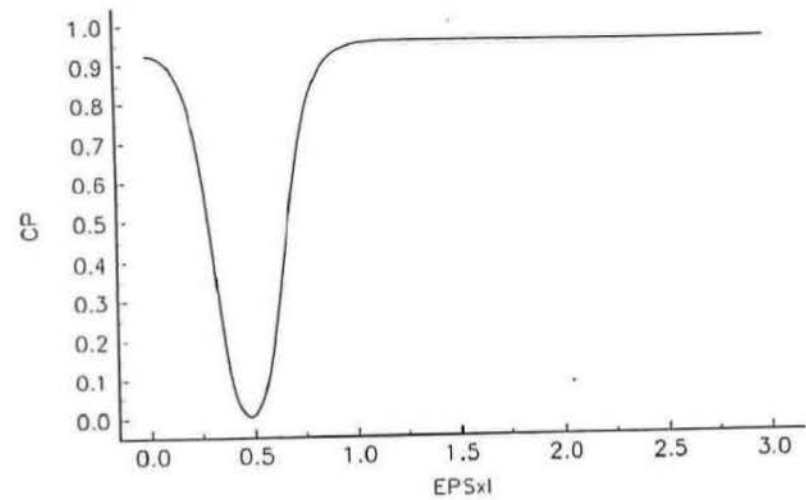




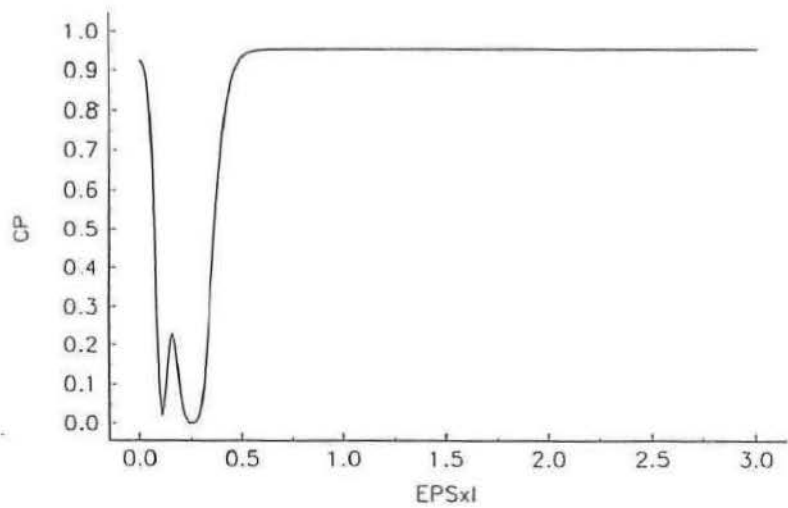
C3P3X



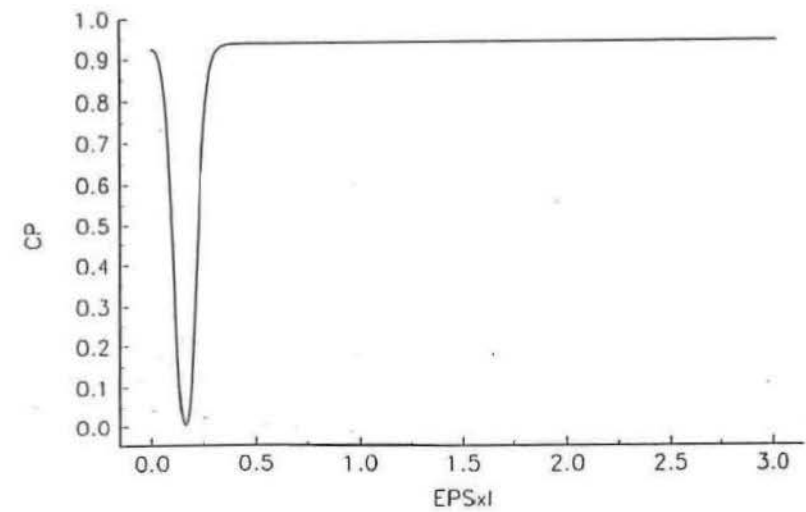
C3P4X



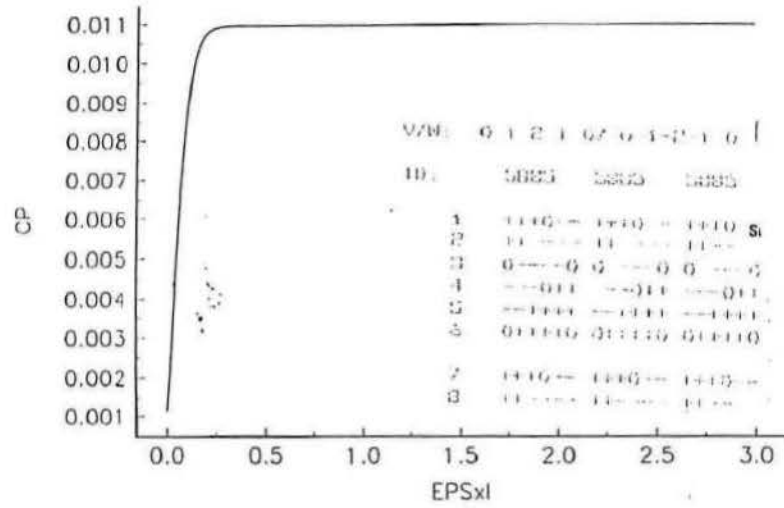
C3P3Y



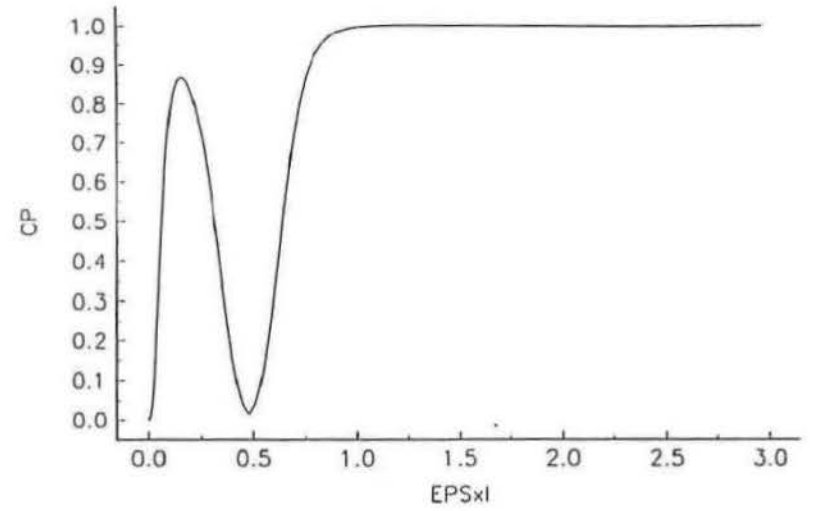
C3P4Y



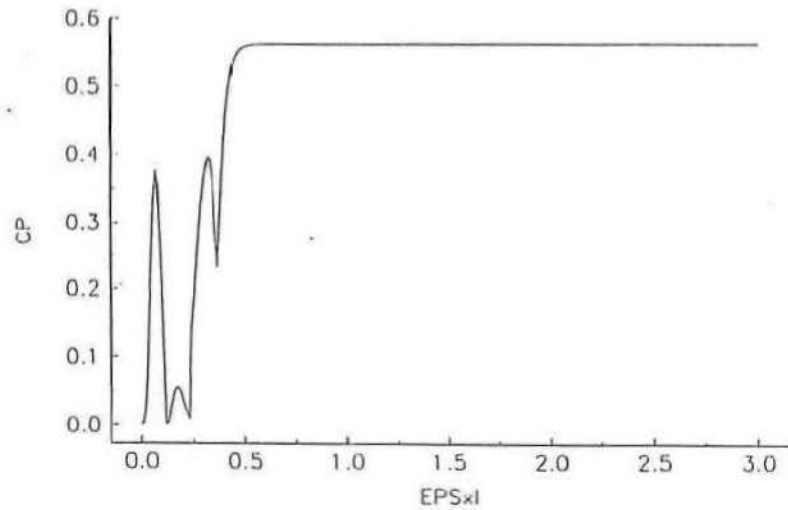
C3P5X



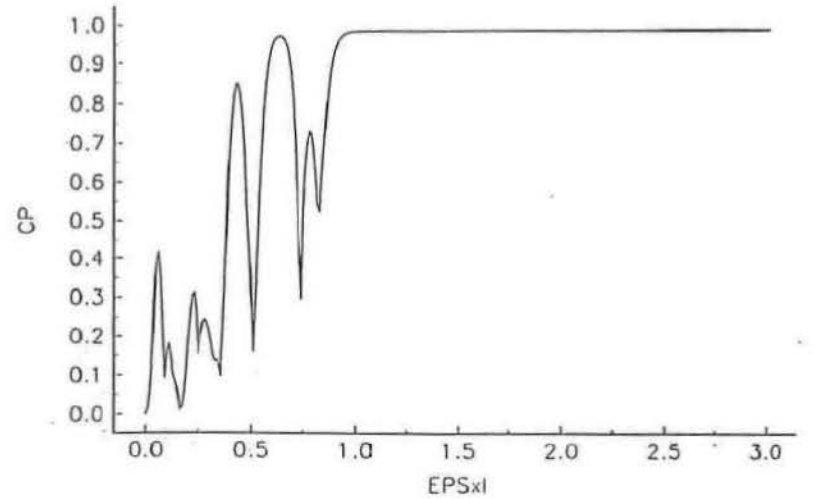
C3P6X



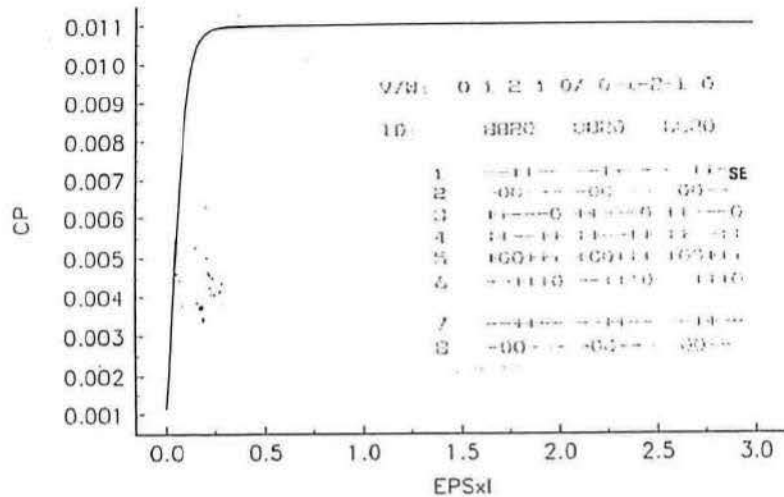
C3P5Y



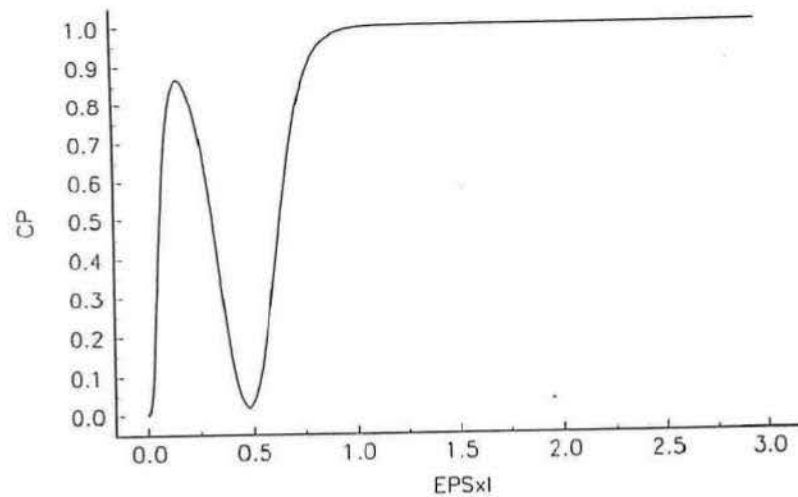
C3P6Y



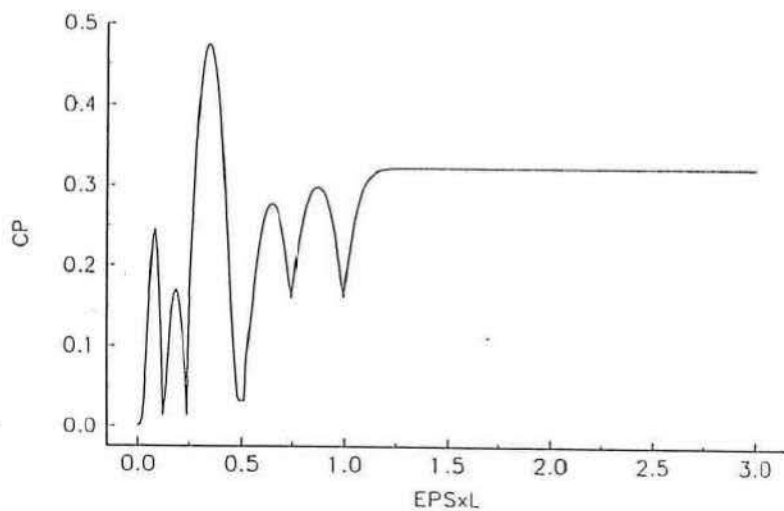
C3P7X



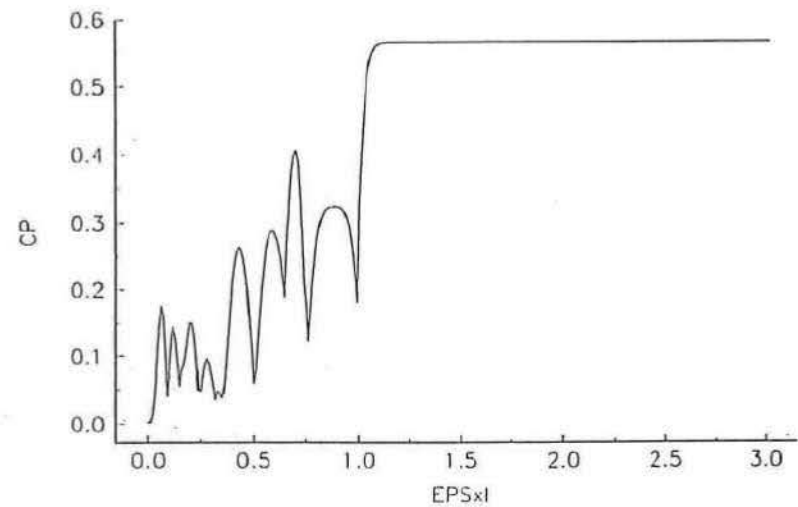
C3P8X

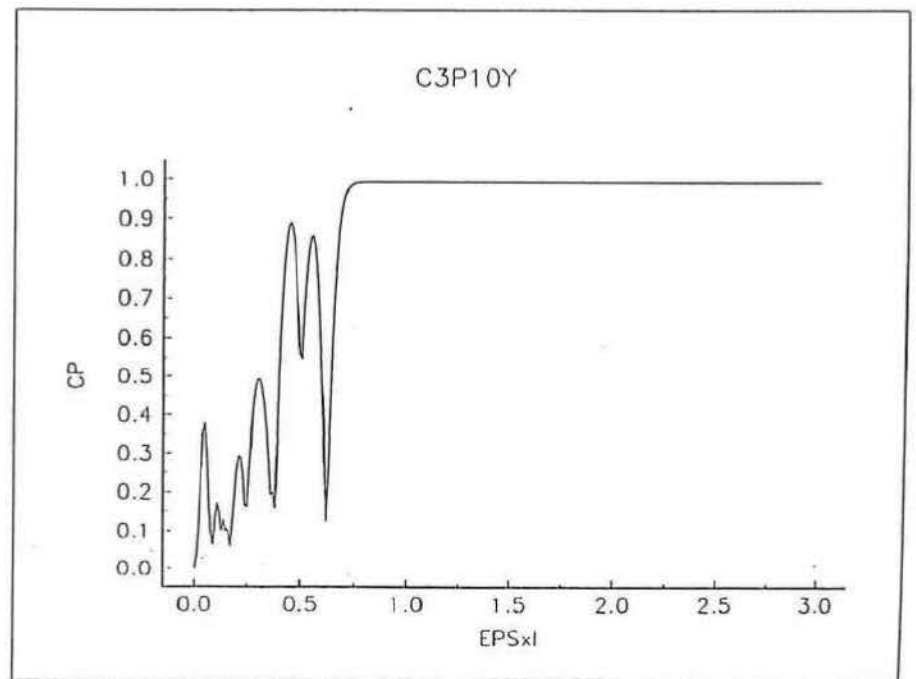
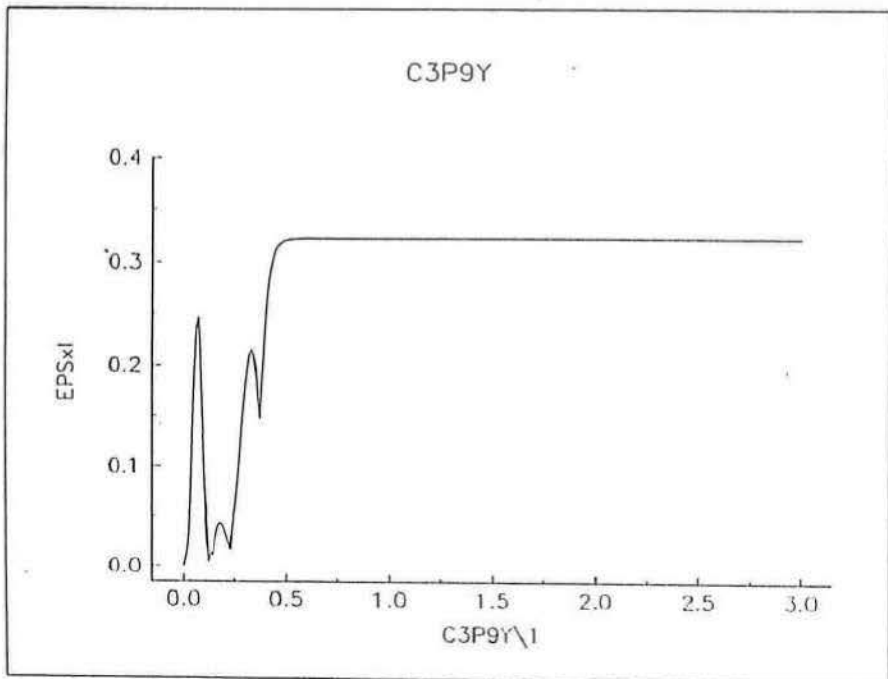
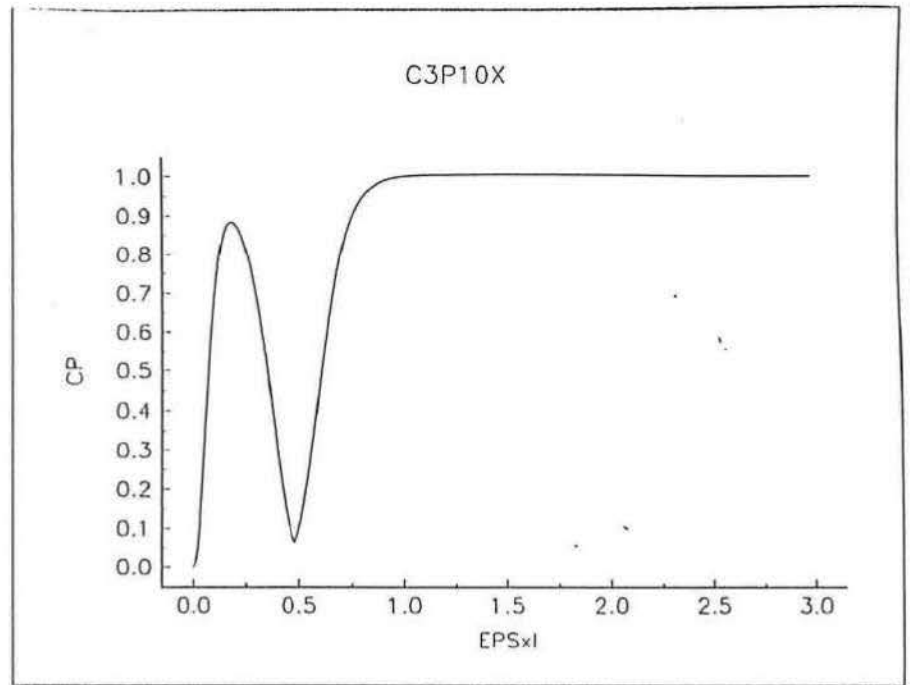
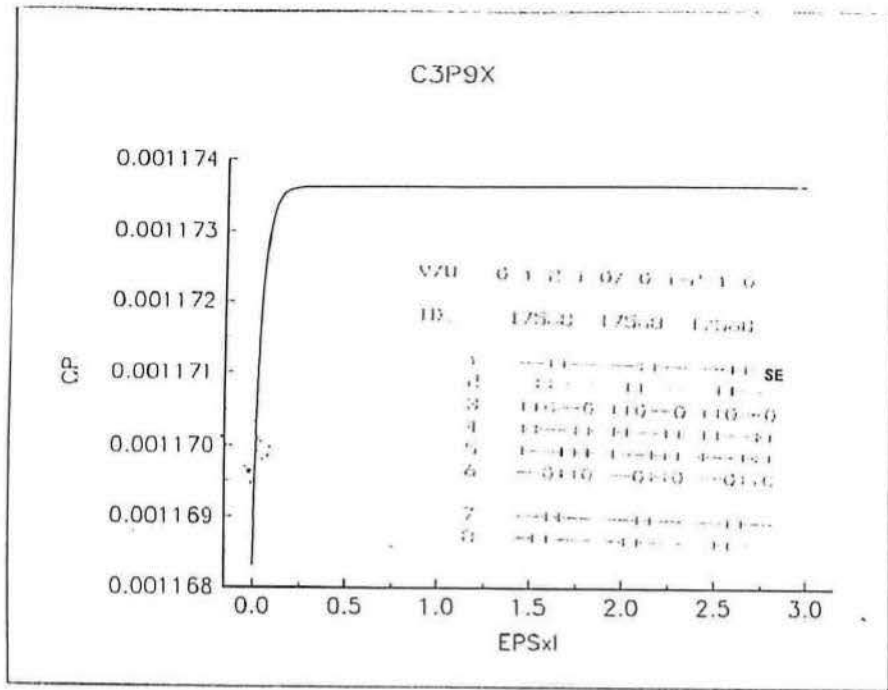


C3P7Y

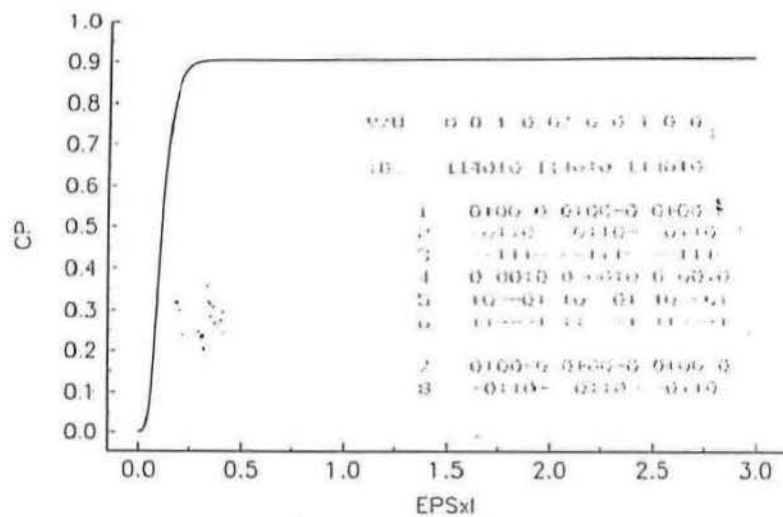


C3P8Y

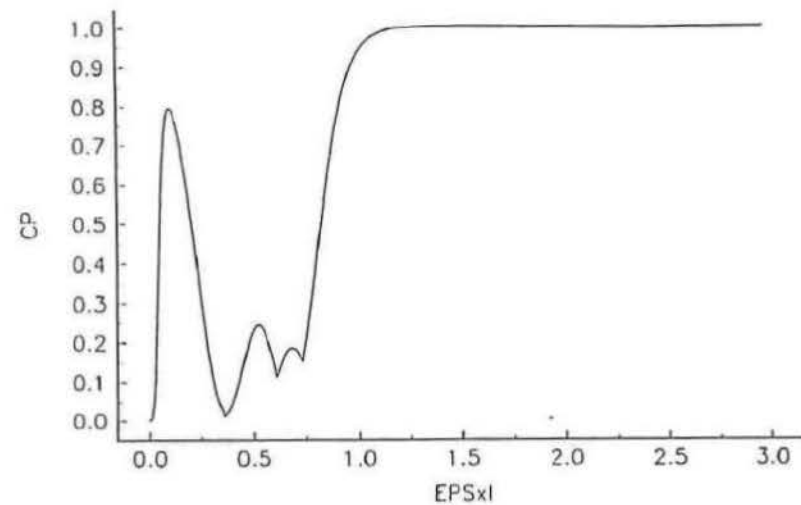




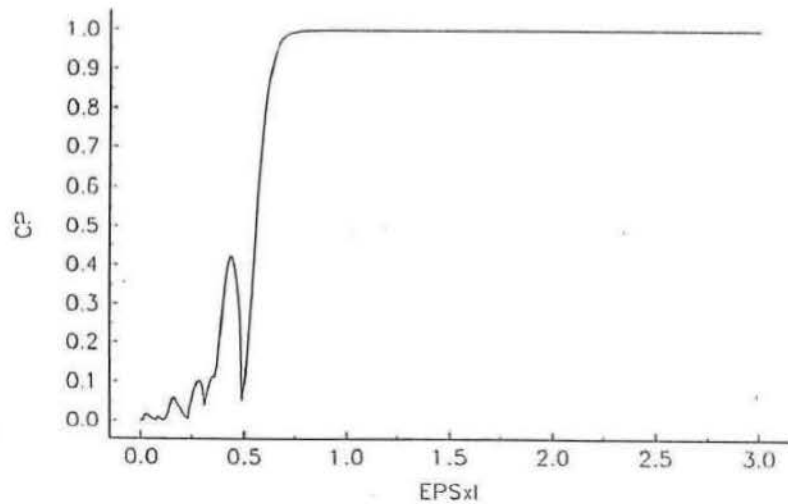
C4P1X



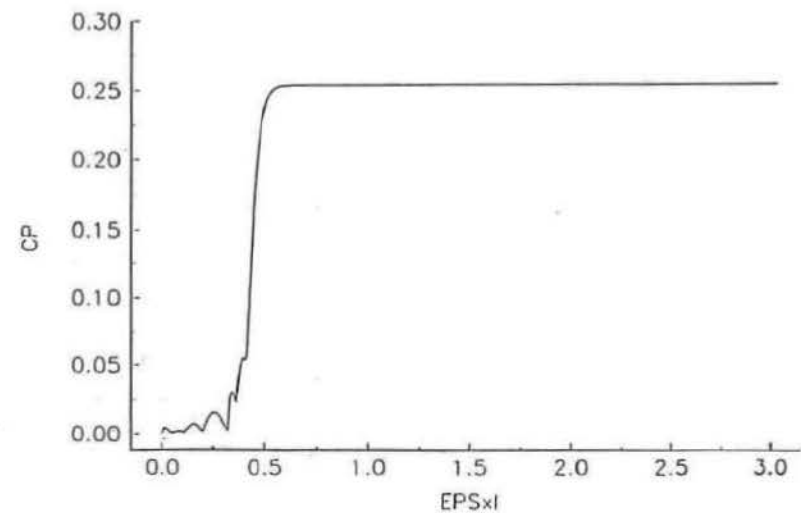
C4P2X

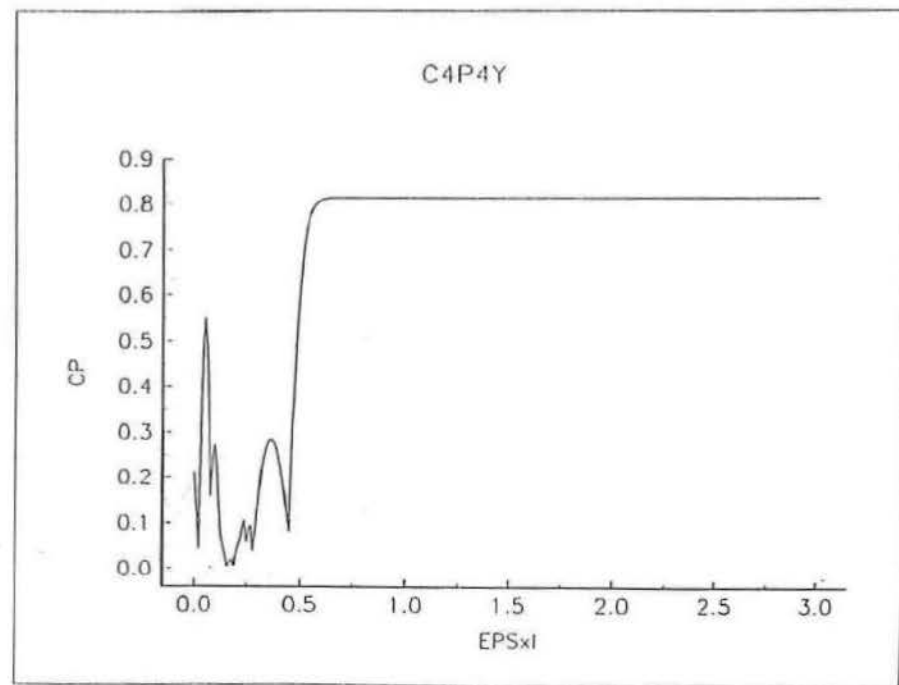
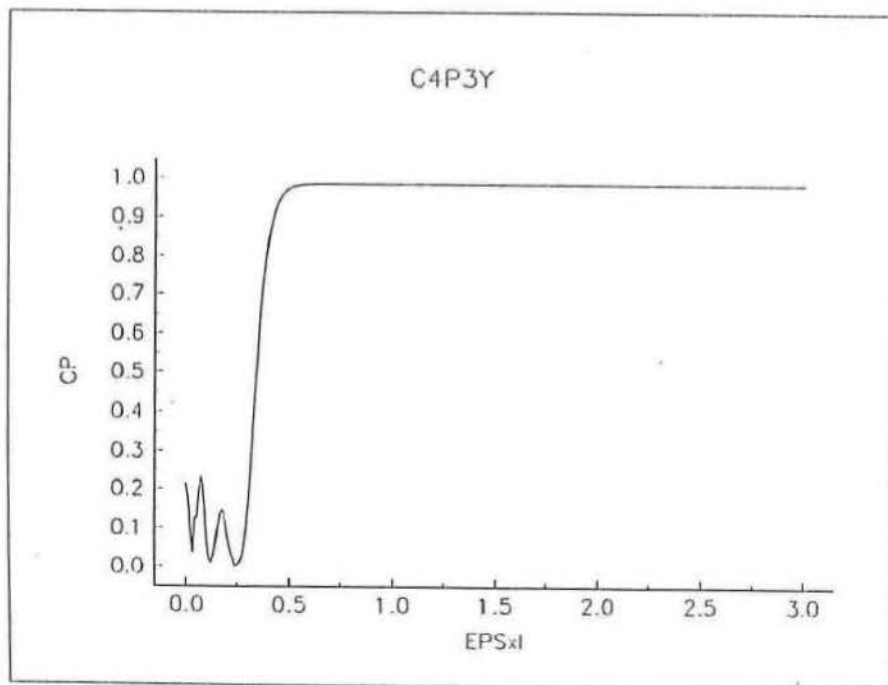
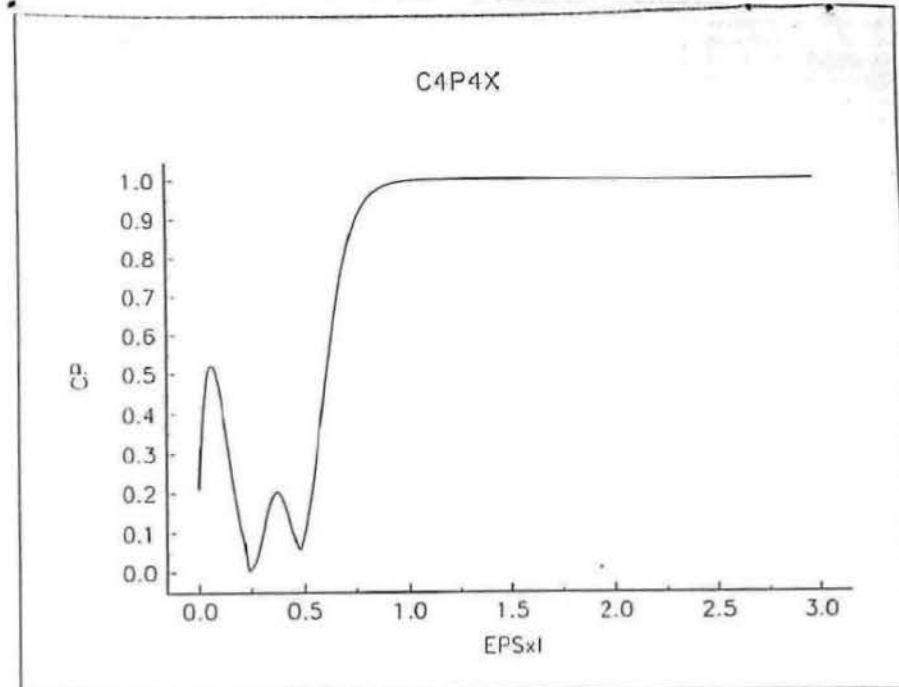
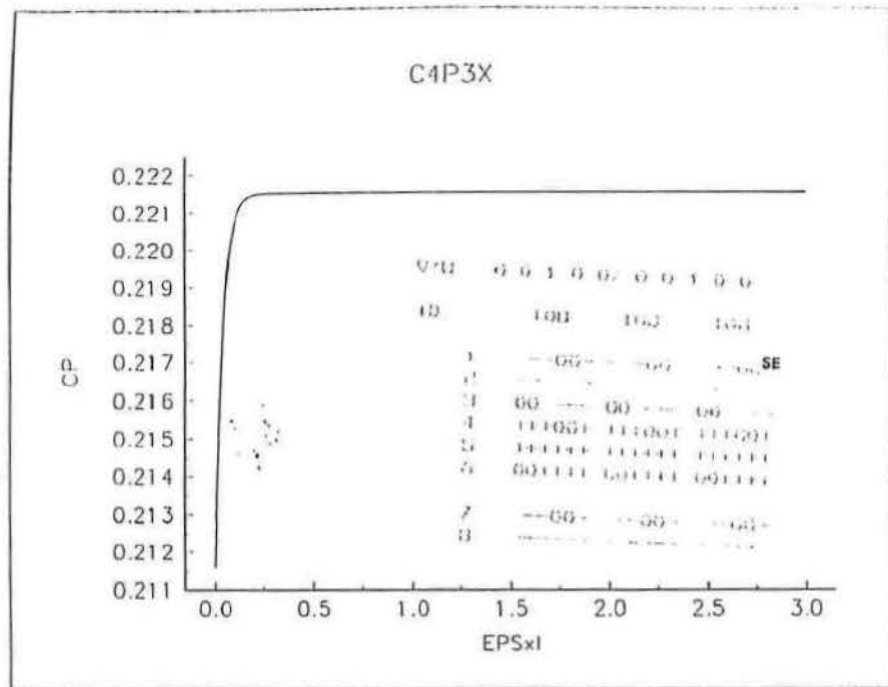


C4P1Y

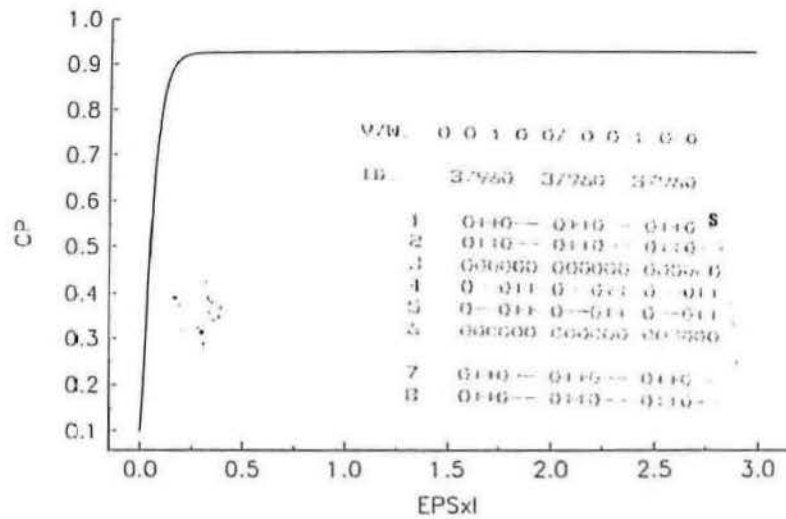


C4P2Y

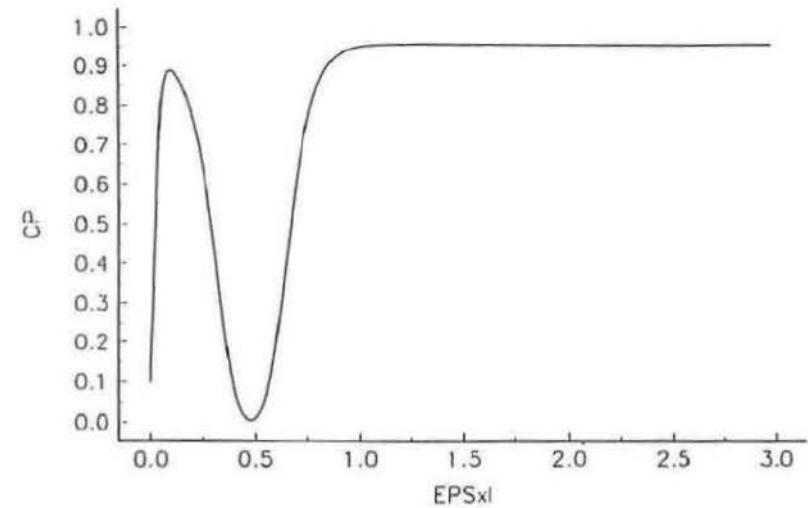




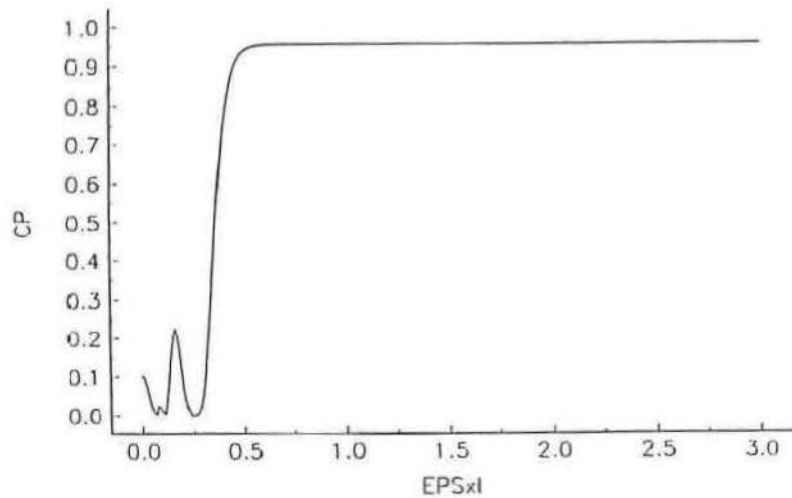
C4P5X



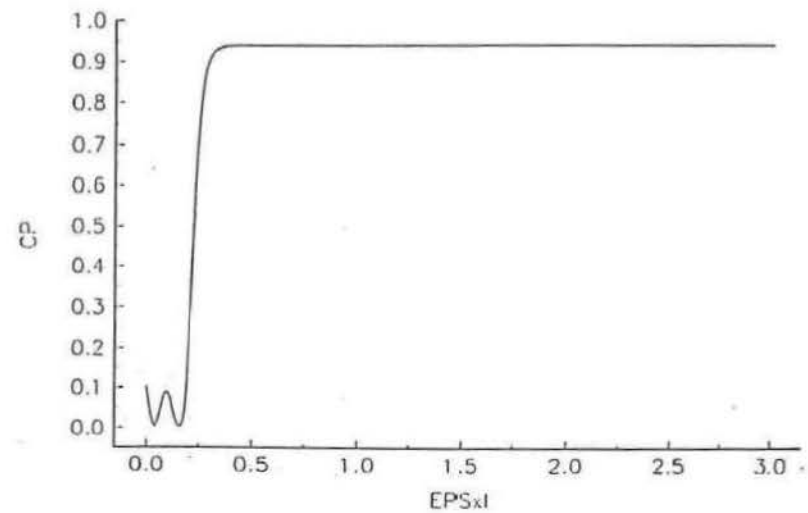
C4P6X

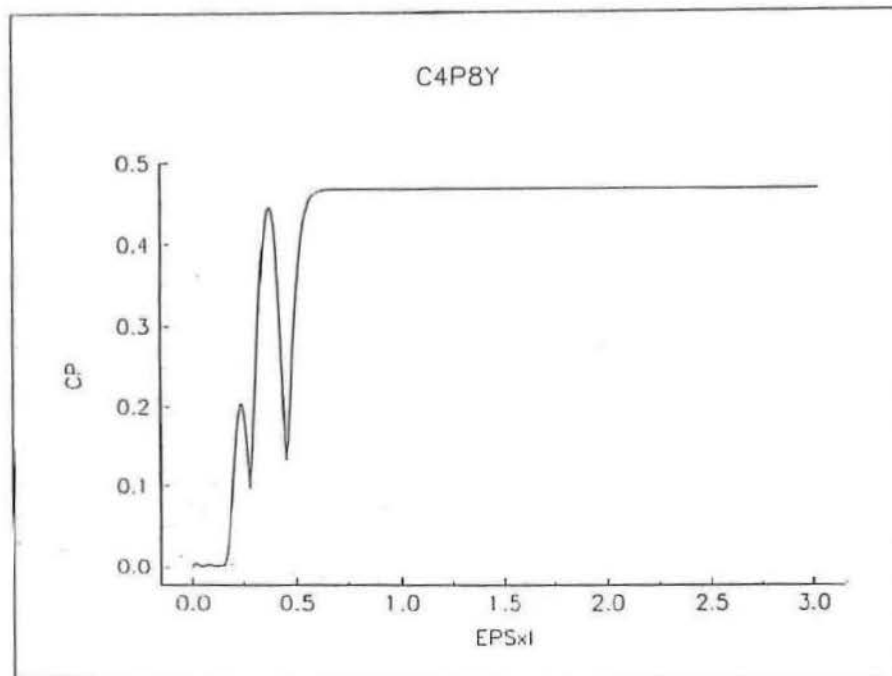
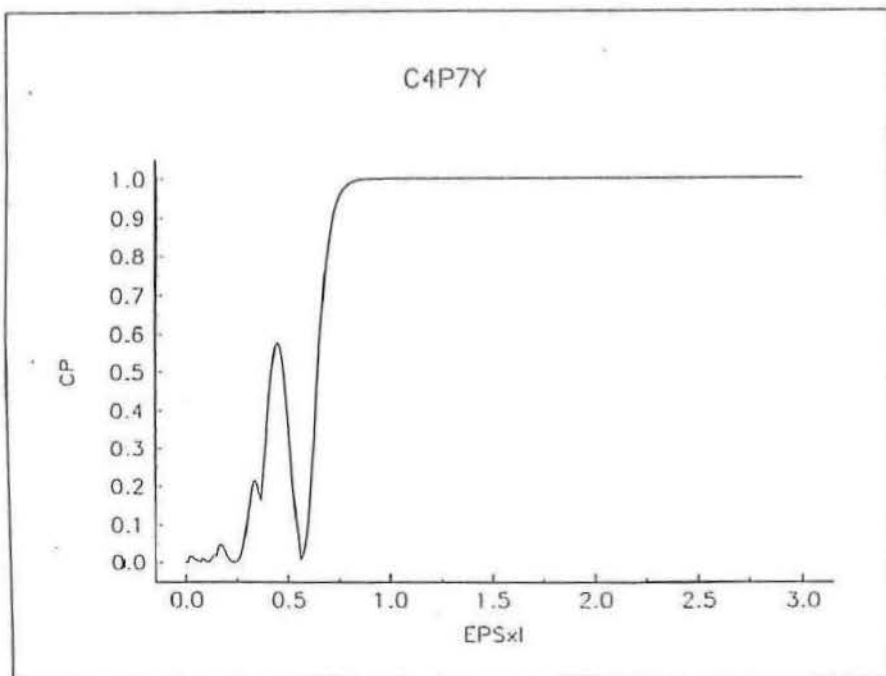
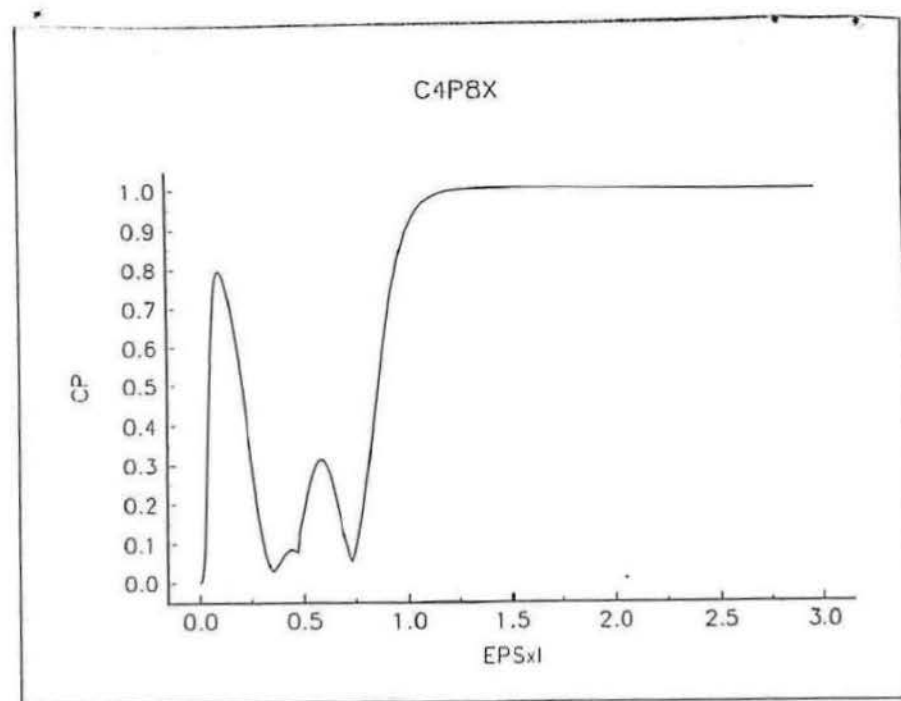
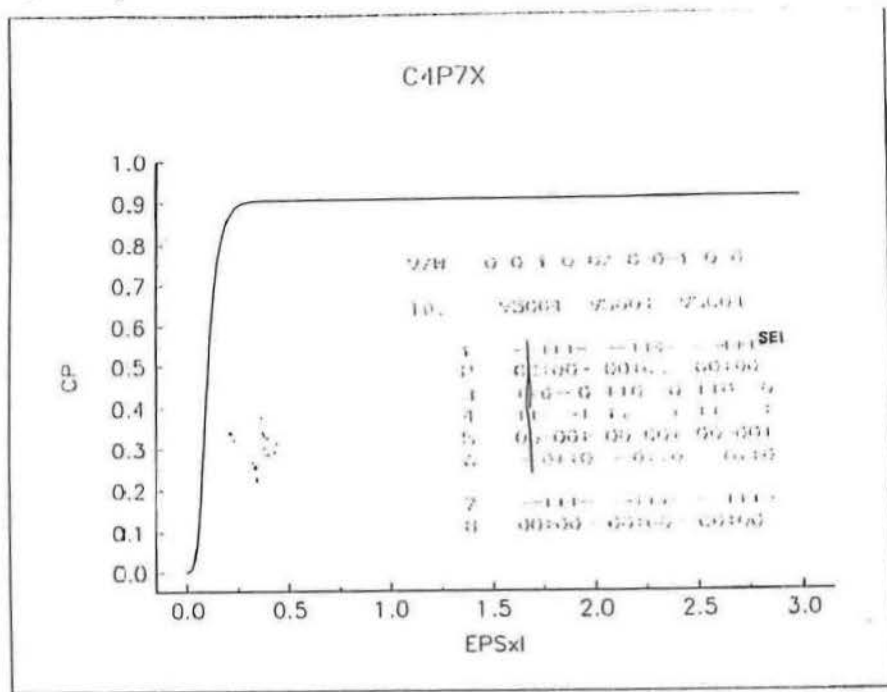


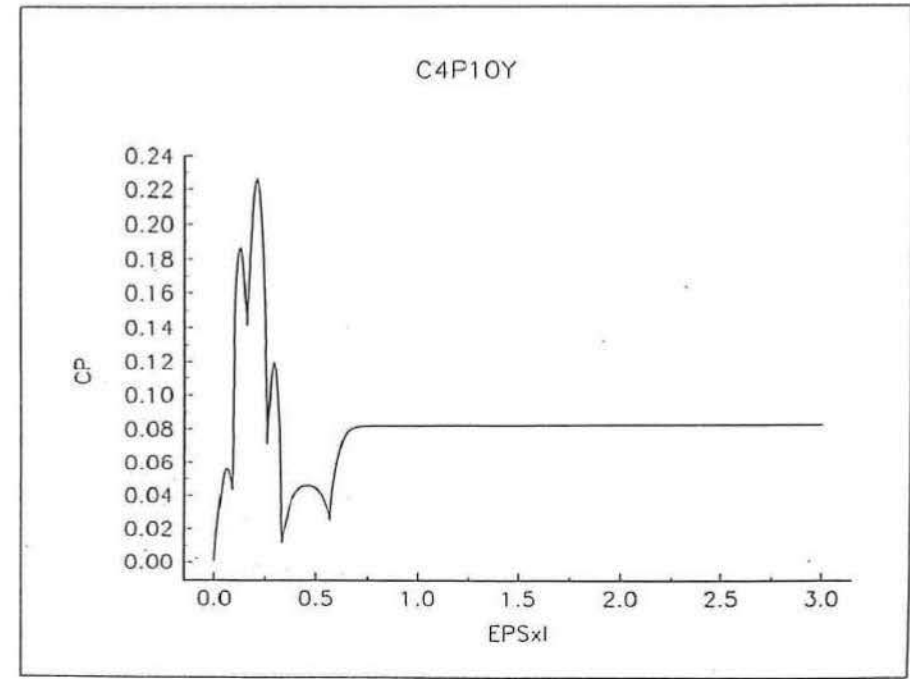
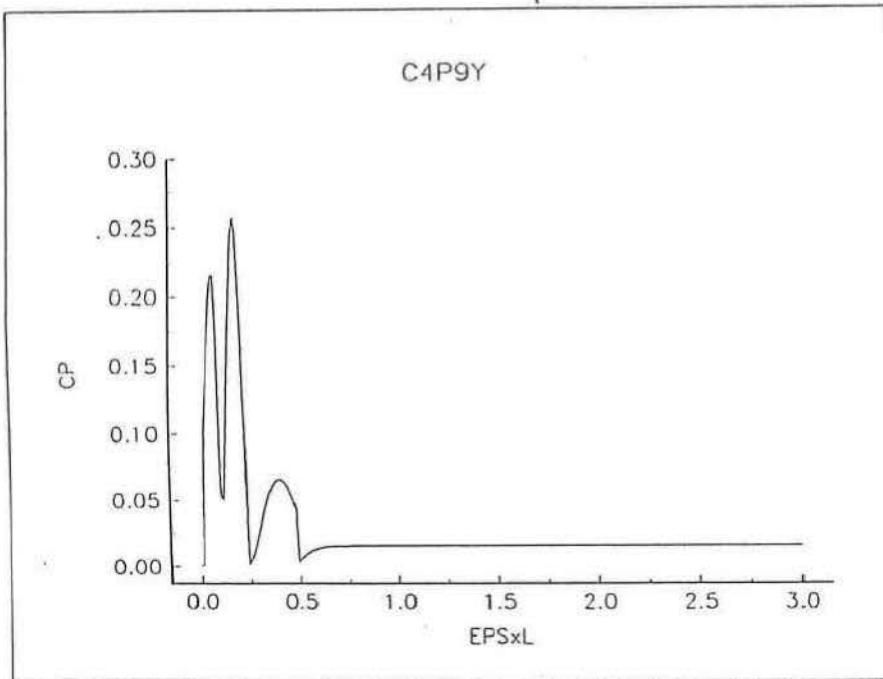
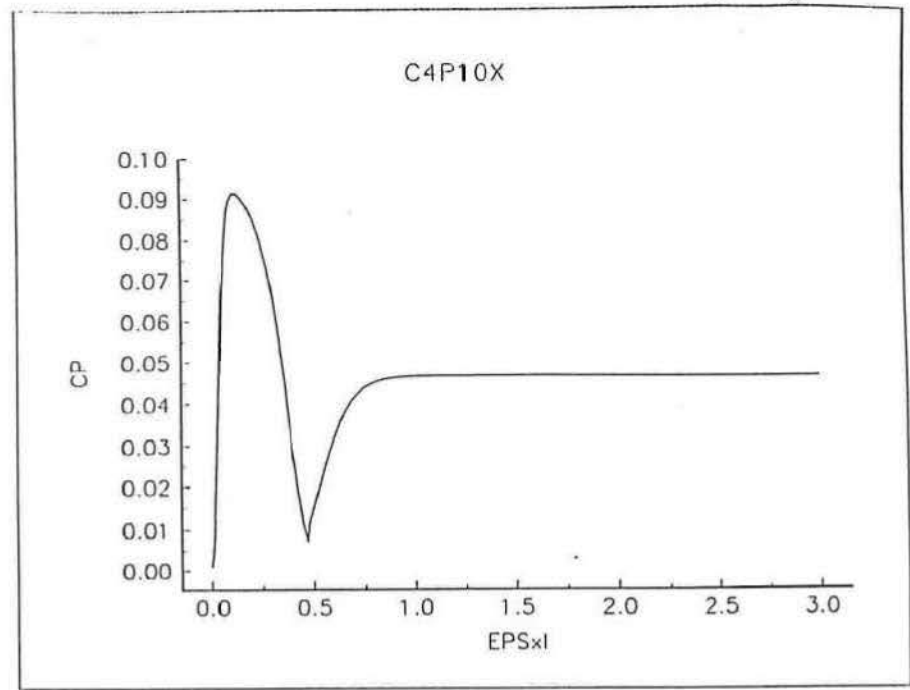
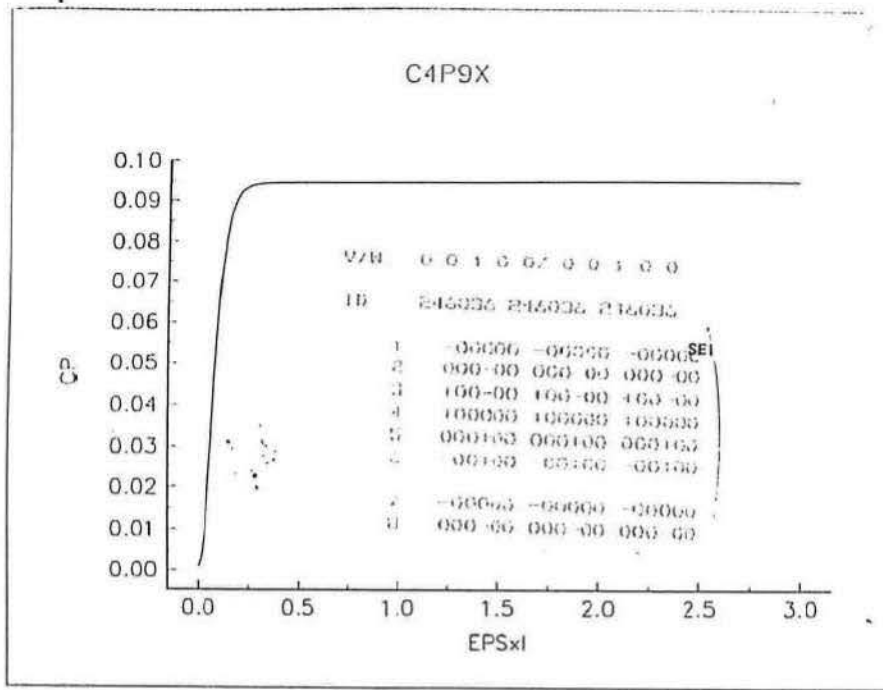
C4P5Y

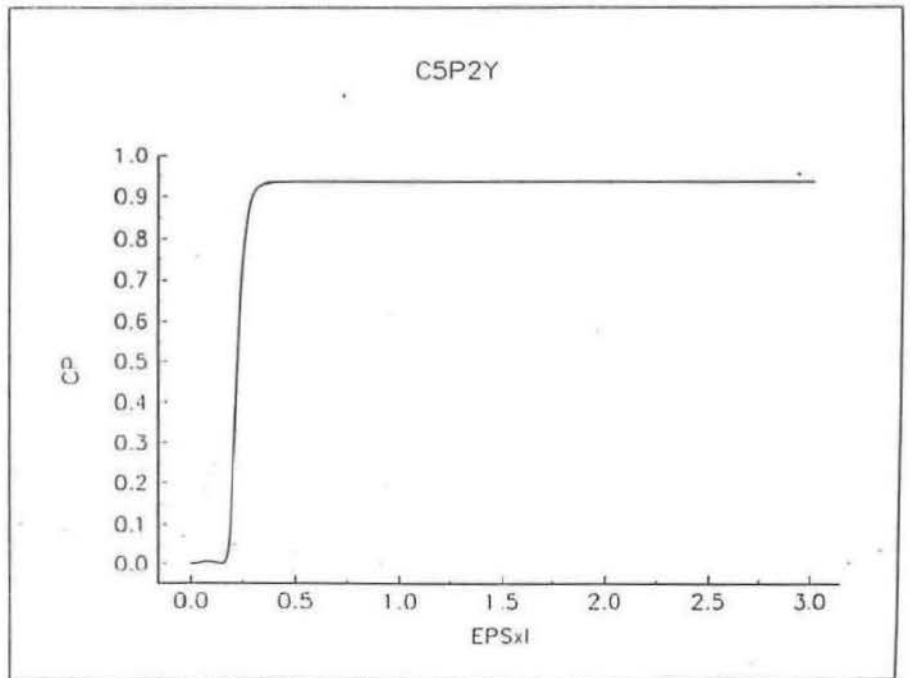
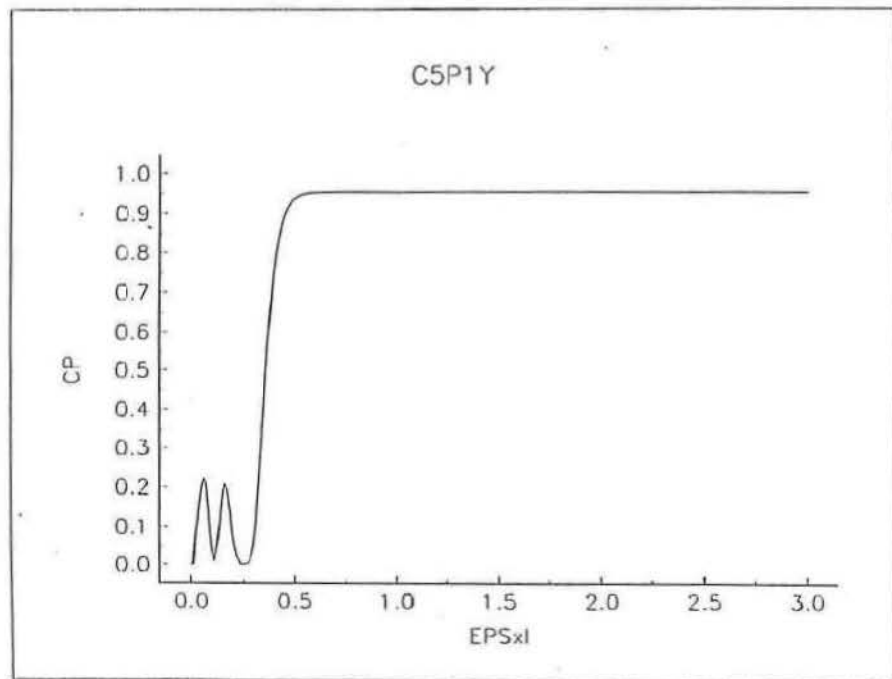
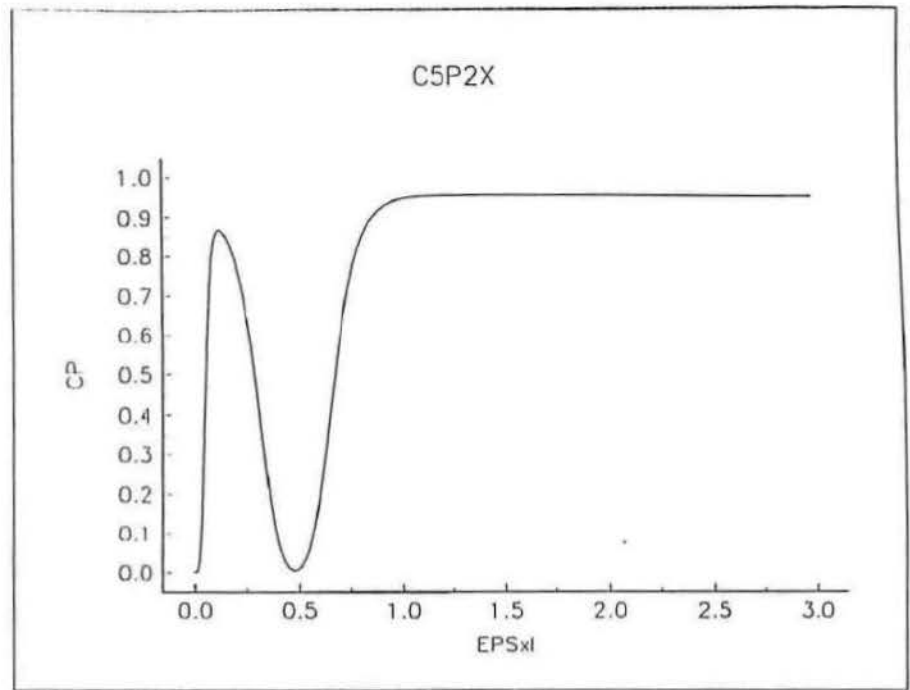
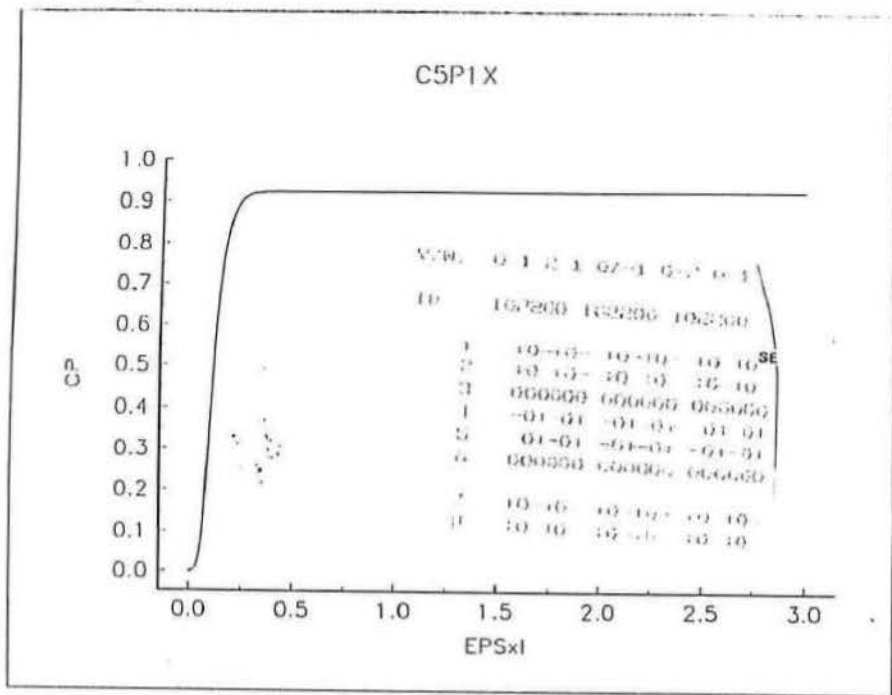


C4P6Y

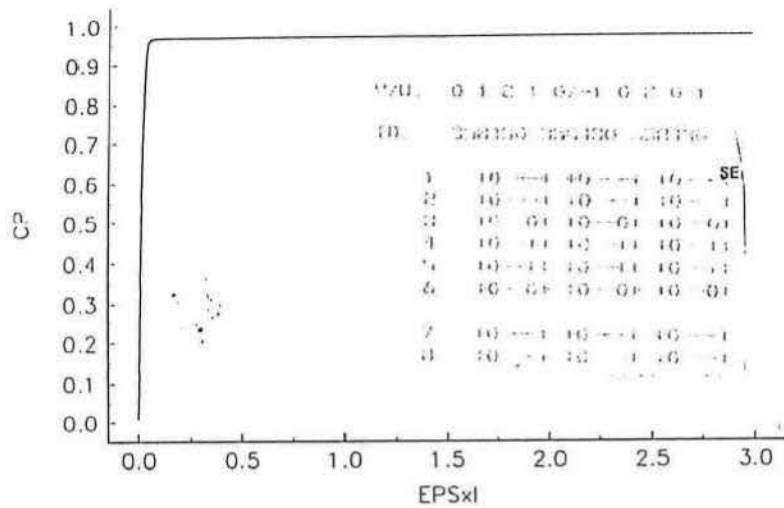




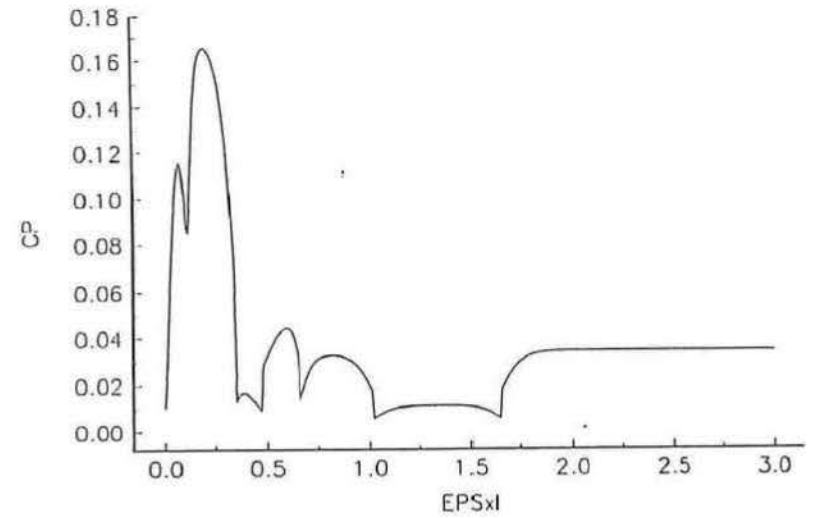




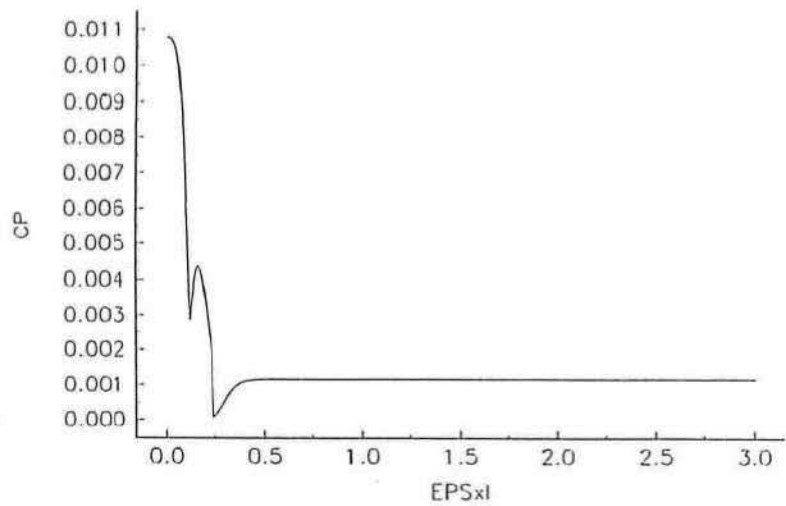
C5P3X



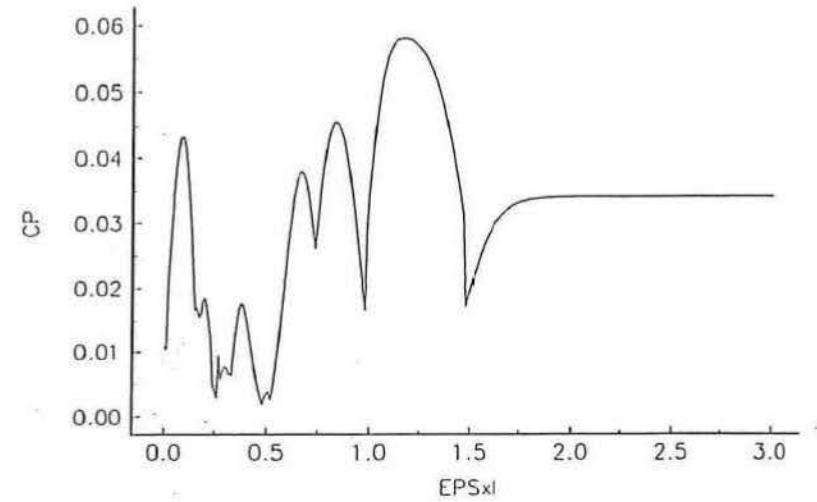
C5P4X

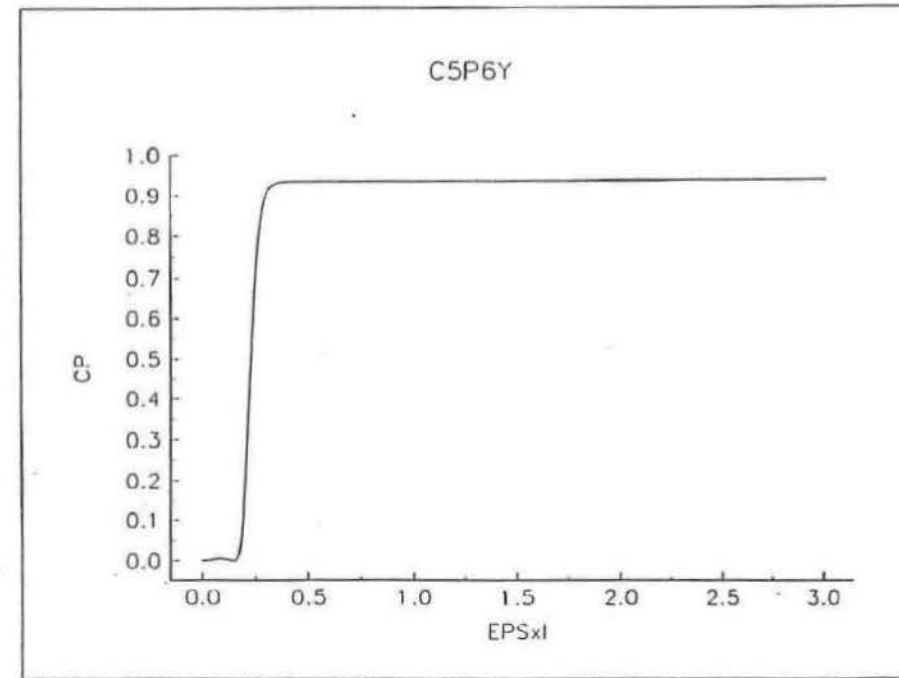
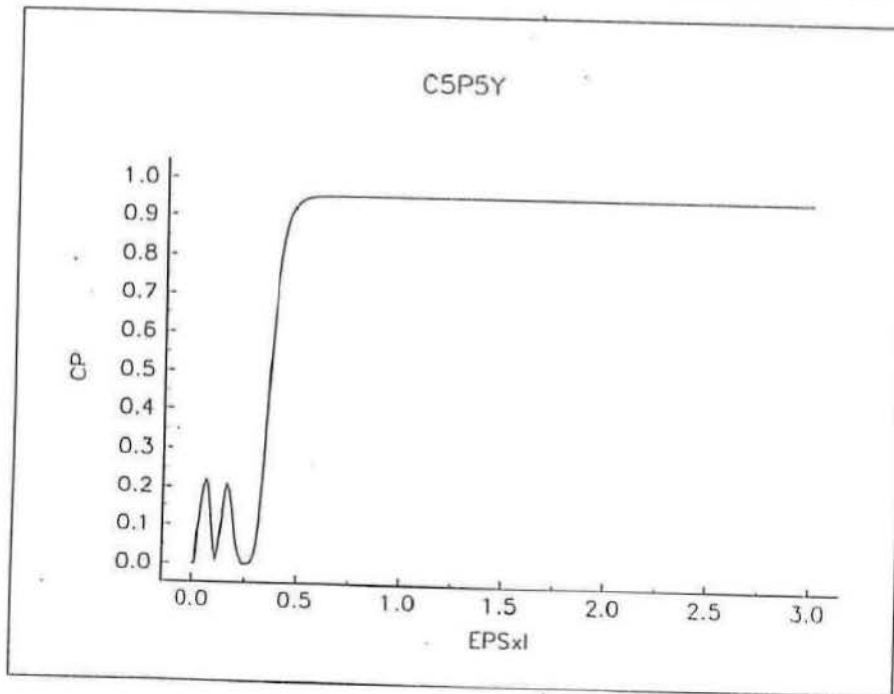
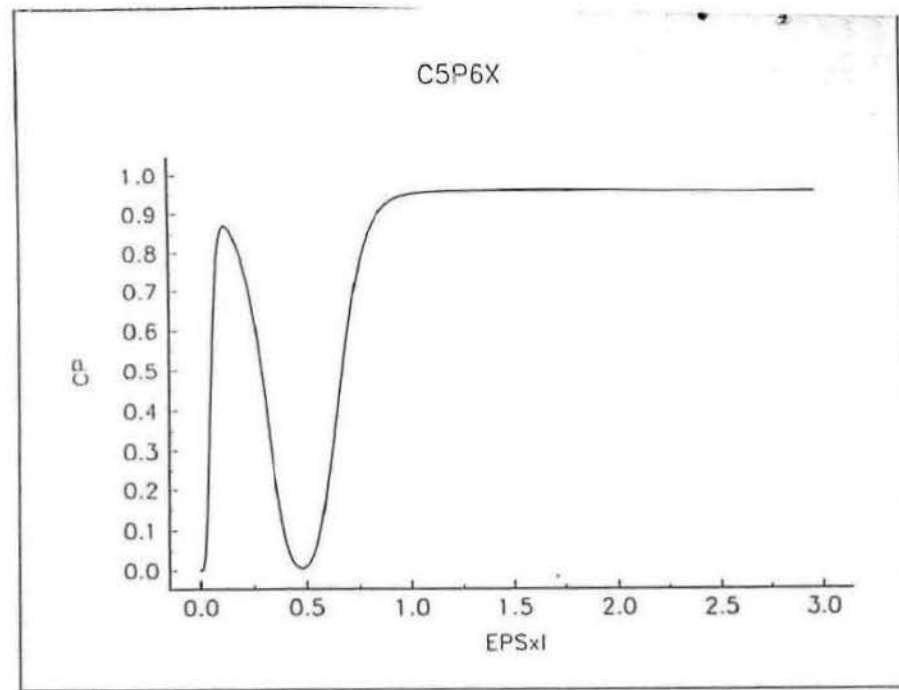
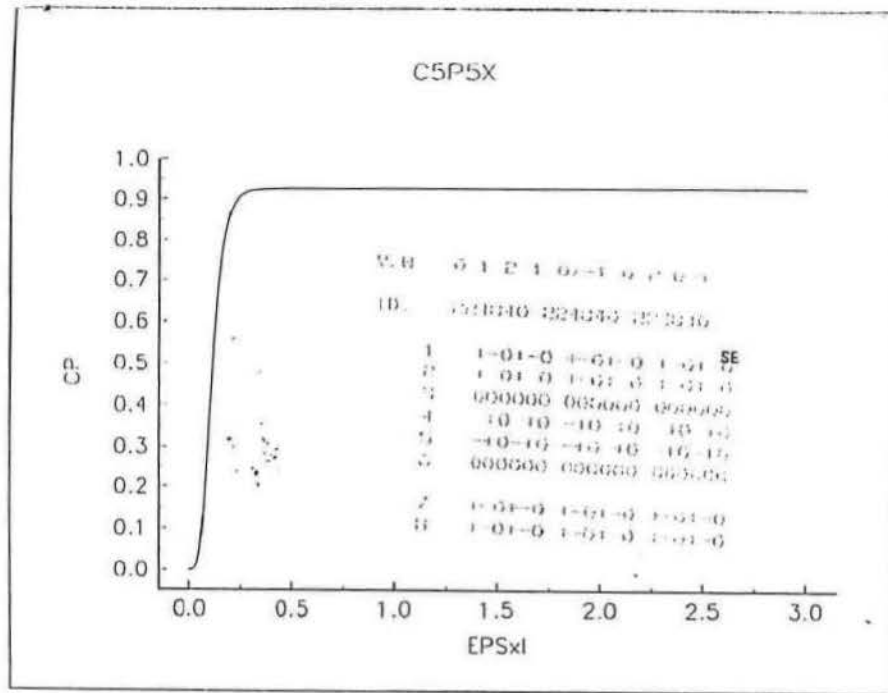


C5P3Y

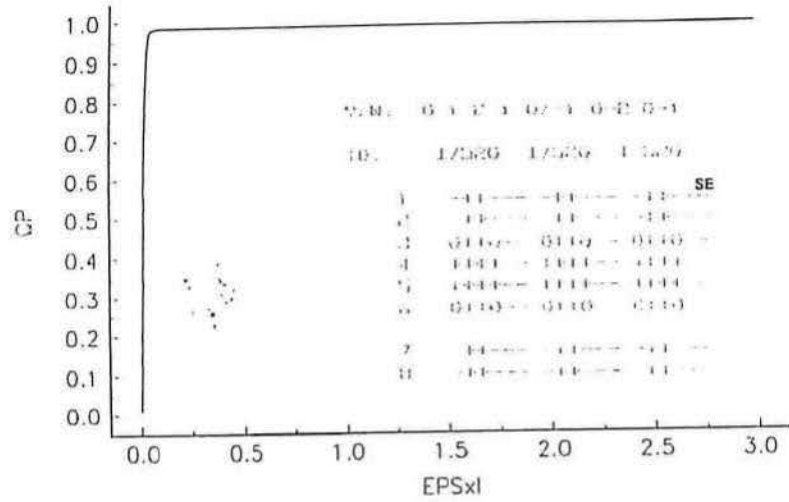


C5P4Y

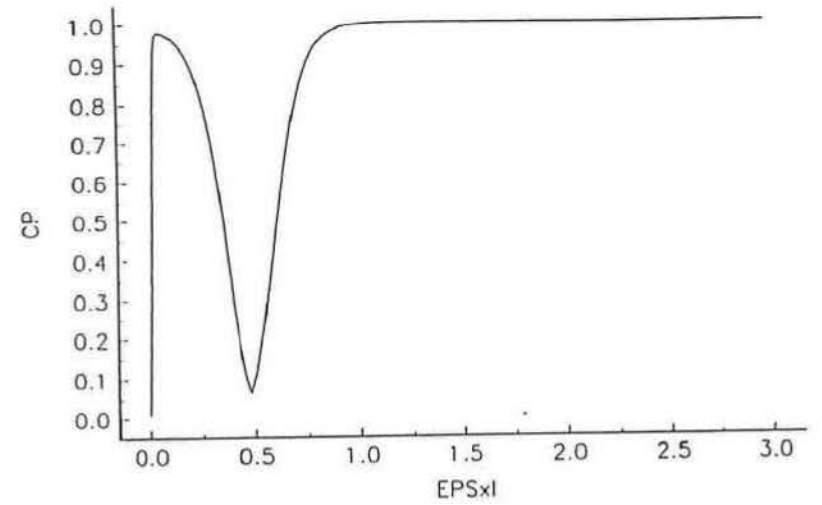




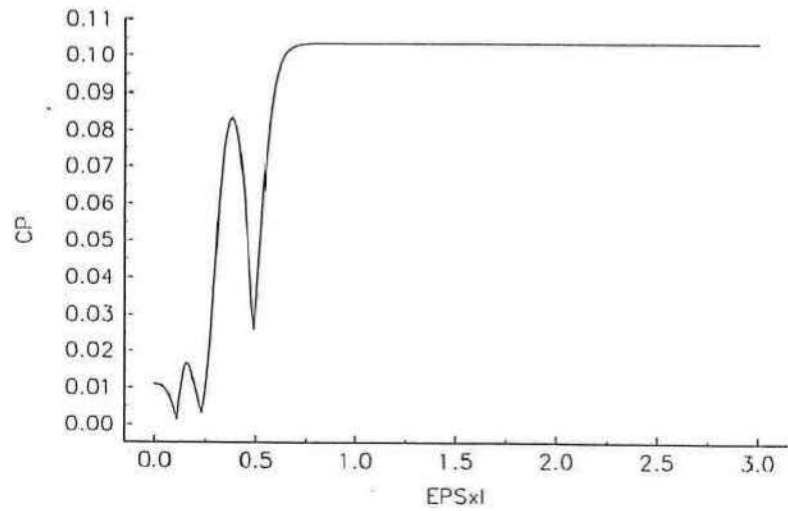
C5P7X



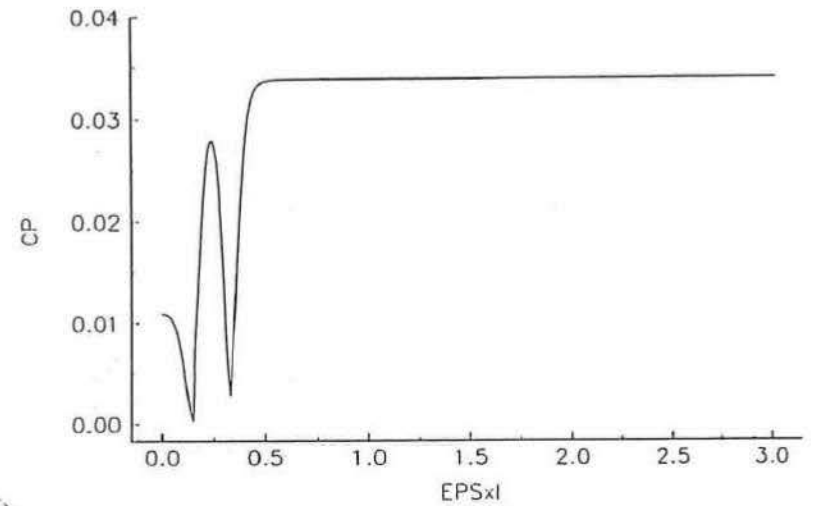
C5P8X



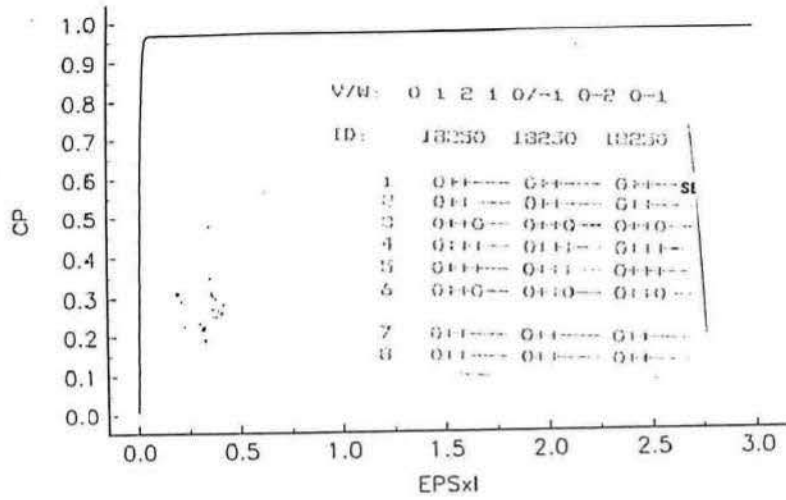
C5P7Y



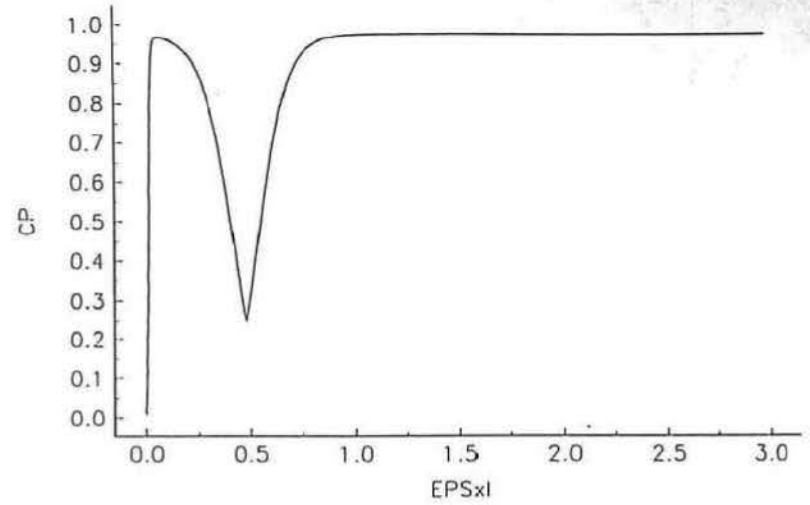
C5P8Y



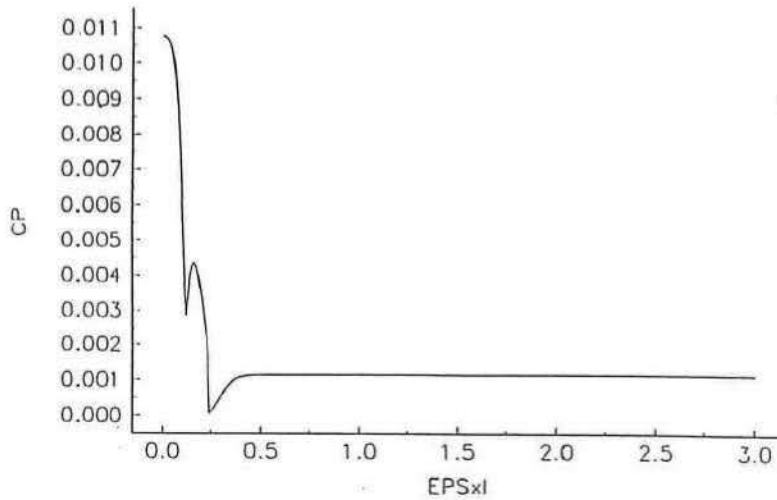
C5P9X



C5P10X



C5P9Y



C5P10Y

