



**UNIVERSIDADE FEDERAL DE PERNAMBUCO  
DEPARTAMENTO DE FÍSICA – CCEN  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA**

## **DISSERTAÇÃO DE MESTRADO**

### **MODELOS DA FÍSICA ESTATÍSTICA PARA AGRUPAMENTO NÃO SUPERVISIONADO DE DADOS**

por

**Rogério Mendes da Silva**

Dissertação apresentada ao Programa de Pós-Graduação em Física do Departamento de Física da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do título de Mestre em Física.

**Banca Examinadora:**

Prof. Sergio Galvão Coutinho (Orientador-UFPE)

Prof. Giovani Lopes Vasconcelos (UFPE)

Prof. Alejandro César Frery Orgambide (UFAL)

Recife - PE, Brasil  
Julho - 2004



Universidade Federal de Pernambuco  
Departamento de Física - CCEN  
Programa de Pós-Graduação em Física  
Cidade Universitária - 50670-901 Recife PE Brasil  
Fone ( ++ 55 81 ) 2126-8449/2126-8450 - Fax ( ++ 55 81 ) 3271-0359  
<http://www.df.ufpe.br/pg> e-mail: [posgrad@df.ufpe.br](mailto:posgrad@df.ufpe.br)

---

## Parecer da Banca Examinadora de Defesa de Dissertação de Mestrado

---

Rogério Mendes da Silva

### MODELOS DA FÍSICA ESTATÍSTICA PARA AGRUPAMENTO NÃO SUPERVISIONADO DE DADOS

A Banca Examinadora composta pelos Professores Sergio Galvão Coutinho (Presidente e Orientador), Giovani Lopes Vasconcelos, ambos da Universidade Federal de Pernambuco e Alejandro César Frery Orgambide da Universidade Federal de Alagoas consideram o candidato:

Aprovado com Distinção

Aprovado

Reprovado

Secretaria do Programa de Pós-Graduação em Física do Departamento de Física do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco aos vinte e três dias do mês de julho de 2004.

*Sergio G. Coutinho*

Prof. Sergio Galvão Coutinho  
Presidente e Orientador

*Giovani Lopes Vasconcelos*

Prof. Giovani Lopes Vasconcelos

*Alejandro C. Frery*

Prof. Alejandro César Frery Orgambide

## AGRADECIMENTOS

Agradeço a Deus por todas as oportunidades, por todo este percurso nestes dois anos que estive longe da minha família, a quem agradeço pelo apoio e incentivo. Agradeço também às famílias com quem convivi aqui em Recife, especialmente, à família de D. Santa (Sandra Maria Gomes).

Agradeço ao Prof. Sérgio Coutinho pela orientação.

Agradeço ao Prof. Mauro Copelli pelas várias discussões que facilitaram o andamento deste trabalho.

Agradeço à Ana Maria Caldeira, secretária da pós-graduação, pela agilidade e atenção com que trata os serviços burocráticos referentes a nós alunos.

Agradeço aos meus amigos do Laboratório de Física Teórica e Computacional (LFTC): Hallan, Pedro Hugo, Washington, Chico, Alex, Calitéia, ... pelas várias “dicas” que me pouparam bastante tempo.

Agradeço ao CNPq e à CAPES pelo suporte financeiro.

*Não é aos saltos que se sobe uma montanha mas a passos lentos.*

—SÃO GREGÓRIO MAGNO.

## RESUMO

Nesta dissertação, o problema do agrupamento não-supervisionado de dados ou padrões é investigado, com enfoque em dois recentes métodos fundamentados na Física Estatística. Basicamente, estes métodos associam os elementos do banco de dados aos elementos de um sistema físico, os quais se acoplam através de uma interação de alcance finito. O critério de similaridade entre os dados, necessário para a formação dos agrupamentos, está associado a uma grandeza física do sistema. O primeiro método, chamado agrupamento super-paramagnético, associa cada padrão (matematicamente definido como um ponto no espaço de atributos  $D$ -dimensional) a uma variável de spin do modelo de Potts ferromagnético não homogêneo. Neste caso, dois pontos pertencem ao mesmo *cluster* quando a correlação spin-spin, no equilíbrio termodinâmico a uma temperatura  $T$ , está acima de um certo limiar. No segundo método, o conjunto de dados é associado a um sistema de mapas caóticos acoplados e as classificações emergem da sincronia destes mapas. A fundamentação teórica e as implementações computacionais destes dois métodos são apresentadas e comparativamente discutidas. Os métodos são aplicados a três conjuntos de dados sintéticos que, por razões de visualização, são formados por padrões de pontos em um espaço bidimensional limitado. Os padrões são especificamente escolhidos para testar e comparar as potencialidades dos métodos empregados. O agrupamento super-paramagnético, por estar fundamentado em um sistema de spins ferromagnético, mostra-se mais robusto com a relação aos ruídos de fundo ou *background*, evitando a agregação aos *clusters* de possíveis "trilhas" (de caráter unidimensional) que se

formam no *background*. Em decorrência, o agrupamento super-paramagnético identifica os *clusters* através de suas densidades e correlações locais com maior precisão, sendo apropriado para identificar e delimitar formas mais compactas eliminando ruídos de fundo nas fronteiras. Por outro lado, o método por associação a mapas caóticos possui maior sensibilidade para identificar estruturas com caráter unidimensional subjacentes aos ruídos de fundo, por conta das conexões locais terem grande influência na transmissão das correlações, sendo, portanto mais eficiente na classificação de *clusters* com estruturas mais complexas. Computacionalmente, o agrupamento através de mapas caóticos mostrou-se ser  $\sim 265$  vezes mais rápido que o agrupamento super-paramagnético quando aplicado aos conjuntos de dados aqui selecionados, tendo em vista que o último carece de realizar simulações numéricas através do método Monte Carlo estabilizando termicamente o sistema para cada valor de temperatura considerado.

## ABSTRACT

In this dissertation the unsupervised data clustering problem is investigated. Two recent methods based on Statistical Mechanics are considered. These methods associate the data points to the elements of a physical system, which by their turn are coupled by a short-range interaction. The similarity criterion, which is required to define the clusters, is associated to a physical quantity of the system. Within the first method, so called *superparamagnetic clustering* (SPC), the data points defined in an appropriate D-dimensional attribute space are associated to the spin variables of the inhomogeneous ferromagnetic q-state Potts model. In this case two points belong to the same cluster whenever the spin correlation, in the thermodynamic equilibrium at a given temperature  $T$ , is above a certain threshold. Within the second method the data set is associated to a system of coupled chaotic maps (CM), the data clusters emerging from the synchrony between the maps at the steady state. The theoretical background as well as the computational implementation of both methods are presented and discussed. The methods are applied to a three sets of synthetic data built on a limited two-dimensional space for the sake of visualization. The data patterns are built on in such a way to test and compare the power and efficiency of the methods. The SPC method, which is based on a ferromagnetic spin system, is more robust with respect to the background noise avoiding the inclusion of eventual one-dimensional like structures to the surface of the clusters. Therefore it is more appropriated to identify and delimit more compact clusters being more efficient to eliminate the background noise at the frontiers. On the other hand, the

CM method is more sensitive at identifying one-dimensional like clusters hidden in the background due to the great influence of the local connection on the transmission of the correlations. As a consequence, this latter method is more efficient to classify and delimit clusters with more complex structure. From the computational viewpoint regarding the data sets considered in this work, the CM method runs of the order of  $\sim 265$  times faster than the SPC one, which spends long CPU time running Monte Carlo simulations up to the thermodynamical equilibrium state for each temperature value considered so far.

# SUMÁRIO

<b>Capítulo 1—Introdução</b>	1
1.1 Definições . . . . .	4
1.2 Representação de Padrões, Seleção e Extração de Atributos . . . . .	5
1.3 Medidas de Similaridade . . . . .	7
1.4 Métodos de Clustering . . . . .	9
<b>Capítulo 2—Os Métodos de agrupamento Super-paramagnético e por Mapas Caóticos</b>	14
2.1 Introdução . . . . .	14
2.2 O Modelo de Potts . . . . .	17
2.3 Mecânica Estatística da Classificação Não Paramétrica . . . . .	19
2.4 Diagrama de Fase de um Modelo de Campo Médio . . . . .	20
2.5 Definição de Interações Locais . . . . .	22
2.6 Cálculo de Grandezas Termodinâmicas . . . . .	23
2.7 Localizando a Fase super-paramagnética . . . . .	24
2.8 O Procedimento de Agrupamento . . . . .	25
2.9 Agrupamento por Mapas Caóticos . . . . .	25
2.9.1 O Método Computacional . . . . .	27
<b>Capítulo 3— Aplicações a Dados Artificiais</b>	29

---

3.1	Conjuntos de dados A . . . . .	30
3.2	Conjuntos de dados B . . . . .	33
3.3	Conjuntos de dados C . . . . .	35
<b>Capítulo 4—Conclusões</b>		<b>56</b>
<b>Apêndice A—Solução do Modelo de Potts com <math>q</math>-estados na Aproximação de Campo Médio</b>		<b>60</b>
A.1	Parâmetros de ordem e energia livre . . . . .	63
A.2	Diagrama de fase . . . . .	66
A.2.1	Transição Paramagnética - Super-Paramagnética . . . . .	66
A.2.2	Transição Super-Paramagnética - Ferromagnética . . . . .	67
A.2.3	Transição Paramagnética - Ferromagnética (A) . . . . .	72
A.2.4	Transição de ordem dos spins do background: fase ferromagnética (B) . . . . .	74
A.3	Funções de correlação . . . . .	76
A.4	Susceptibilidade . . . . .	78
<b>Apêndice B— O Método Monte Carlo</b>		<b>80</b>
B.1	Mecânica Estatística de Equilíbrio . . . . .	80
B.2	Processos de Markov, Ergodicidade e Balanceamento detalhado . . . . .	82
B.3	Razão de Aceitação . . . . .	83
B.4	O algoritmo Metropolis . . . . .	84
B.5	Os algoritmos de Wolff e de Swendsen-Wang . . . . .	85

## LISTA DE FIGURAS

1.1	(a) Exemplo de grupos de dados, (b) Identificação dos grupos. Os números indicam diferentes grupos. . . . .	2
1.2	Exemplo de similaridade com contexto. Na figura da esquerda, $B$ é mais similar a $A$ do que a $C$ , enquanto que na figura da direita $B$ é mais similar a $C$ do que a $A$ . . . . .	8
1.3	Exemplo de similaridade conceitual entre pontos, onde os grupos podem ser diferenciados pelo conceito geométrico de “retângulo” e “elipse”. . . . .	9
1.4	Exemplo de pontos agrupados segundo um nível de similaridade. . . .	10
1.5	Dendograma referente aos pontos da figura 1.4 [2]. . . . .	10
1.6	Três clusters gerados por uma distribuição gaussiana com pontos médios diferentes. . . . .	12
1.7	Classificação dos pontos da figura 1.6 pelo método <i>k-means</i> . . . . .	12
1.8	Clusters com formas mais complexas. . . . .	13
1.9	Classificação dos pontos da figura 1.8 pelo método <i>k-means</i> . . . . .	13
2.1	Representação esquemática de um modelo de Potts de campo médio, relevante para problemas de “ <i>clustering</i> ” [19]. Para maiores detalhes ver apêndice A. . . . .	20

- 2.2 Diagrama de fases obtido da solução da aproximação de campo médio de um modelo de Potts relevante para o problema de “*clustering*”. Este diagrama de fases foi obtido no limite termodinâmico  $\frac{N_b}{M} \rightarrow \infty$ , com  $J_1 = 1$  e  $J_2 = 0.05$ , e está em função de  $J_3$  e  $T^* = (2(q-1) \log(q-1)/(q-2))T$  [19]. Para maiores detalhes ver apêndice A. . . . . 21
- 2.3 Gráfico da susceptibilidade magnética obtido da solução de campo médio de um modelo de Potts relevante para o problema de “*clustering*” [19]. Para maiores detalhes ver apêndice A. . . . . 22
- 3.1 Conjunto de dados artificiais consistindo de dois grupos: um setor de círculo com 1600 e um círculo com 800 pontos, ambos com mesma densidade de pontos e gerados por uma distribuição de probabilidades uniforme, imersos em um background de 1050 pontos. . . . . 38
- 3.2 Gráfico da susceptibilidade referente ao conjunto de pontos da figura 3.1 ao se utilizar o agrupamento super-paramagnético. A região de flutuação (indicada pelas setas, entre 0.03 e 0.16, aproximadamente) em torno de um valor aproximadamente constante e diferente de zero, indica a fase super-paramagnética. . . . . 39
- 3.3 Gráfico da evolução do tamanho dos clusters em função da temperatura ( $\mathcal{T} = \frac{T}{K}$ ). A região entre  $\mathcal{T} \approx 0.03$  e  $\mathcal{T} \approx 0.08$  é caracterizada pela estabilidade dos dois maiores clusters. A seta indica o valor de  $\mathcal{T} = 0.05$  em que foi feita a classificação, apresentada na figura 3.4. . . . . 40
- 3.4 Clasificação dos pontos da figura 3.1 em uma temperatura  $\mathcal{T} = 0.05$  indicada pela seta no eixo horizontal do gráfico da figura 3.3. Os dois maiores agrupamentos, consistem de 1656 (em preto) e 816 (em vermelho) pontos, respectivamente. . . . . 40

- 
- 3.5 Classificação do conjunto de dados mostrado na figura 3.1 ao se aplicar o agrupamento por mapas caóticos em um limite de informação  $I = 0.12$ . Os dois maiores agrupamentos mais densos não são identificados corretamente. . . . . 41
- 3.6 Conjunto de pontos consistindo de dois agrupamentos com 1672 e 768 pontos, possuindo a mesma densidade uniforme de pontos e com tamanho e formato similares ao mostrado na figura 3.1, imersos em um background de 148 pontos. . . . . 42
- 3.7 Evolução do tamanho dos 4 maiores clusters em função da informação entre pares de mapas referentes aos pontos da figura 3.6. A seta indica o valor da informação que utilizamos para fazer a classificação mostrada na figura 3.8. . . . . 43
- 3.8 Classificação dos pontos da figura 3.6 em um limite de informação indicado pela seta no eixo horizontal do gráfico da figura 3.7. . . . . 43
- 3.9 Conjunto artificial de dados consistindo de dois agrupamentos de densidades não uniformes, cada um com 714 pontos, nas regiões:  $-0.7 < x < -0.3$  e  $0.3 < x < 0.7$  com  $-0.4 < y < 0.4$ , divididos longitudinalmente em subregiões com densidades  $\sigma_1 = 10\sigma_b$  e  $\sigma_2 = 5\sigma_b$ , respectivamente à esquerda e à direita, onde  $\sigma_b$  é a densidade do background (1000 pontos) no qual estão imersos. . . . . 44
- 3.10 Gráfico da susceptibilidade referente ao conjunto de dados da figura 3.9 ao se utilizar o agrupamento super-paramagnético. O intervalo assinalado indica a região de ocorrência dos agrupamentos. Note o decaimento mais suave da susceptibilidade devido ao gradiente de densidade dos agrupamentos. . . . . 45

---

3.11	Gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.9. . . . . .	46
3.12	Ampliação do gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.9, na região delimitada da figura 3.11. . . . . .	46
3.13	Gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.9. . . . .	47
3.14	Ampliação do gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.9, na região delimitada da figura 3.13. . . . . .	47
3.15	Classificação dos pontos (quatro maiores clusters) da figura 3.9 pelo agrupamento super-paramagnético em uma temperatura $\mathcal{T} = 0.07$ indicada pela primeira seta no eixo horizontal do gráfico da figura 3.11.	48
3.16	Classificação dos pontos (quatro maiores clusters) da figura 3.9 pelo agrupamento super-paramagnético em uma temperatura $\mathcal{T} = 0.0916$ indicada pela segunda seta no eixo horizontal do gráfico da figura 3.11.	48
3.17	Classificação dos pontos (oito maiores clusters) da figura 3.9 pelo agrupamento através de mapas caóticos em um limite de informação $I = 0.168$ indicado pela primeira seta no eixo horizontal do gráfico da figura 3.13. . . . . .	49

---

3.18	Classificação dos pontos (oito maiores clusters) da figura 3.9 pelo agrupamento através de mapas caóticos em um limite de informação $I = 0.197$ indicado pela segunda seta no eixo horizontal do gráfico da figura 3.13. . . . .	49
3.19	Conjunto artificial de dados consistindo de dois agrupamentos unidos por uma trajetória de pontos. Há 120 pontos no círculo menor à direita, 41 pontos na trajetória, 500 pontos no círculo maior à esquerda e 600 pontos no "aro" concêntrico ao círculo maior. . . . .	50
3.20	Gráfico da susceptibilidade referente ao conjunto de pontos da figura 3.19 ao se aplicar o agrupamento super-paramagnético. O menor pico à esquerda indica a fragmentação da trajetória de pontos e o início da fase super-paramagnética. . . . .	51
3.21	Gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.19. . . . .	52
3.22	Classificação dos pontos da figura 3.19 pelo agrupamento super-paramagnético na temperatura $\mathcal{T} = 0.01$ indicada pela primeira seta no eixo horizontal do gráfico da figura 3.21. Apenas um único cluster é identificado. . . . .	52
3.23	Classificação dos pontos da figura 3.19 pelo agrupamento super-paramagnético na temperatura $\mathcal{T} = 0.04$ indicada pela segunda seta no eixo horizontal do gráfico da figura 3.21. . . . .	53
3.24	Classificação dos pontos da figura 3.19 pelo método super-paramagnético na temperatura $\mathcal{T} = 0.055$ indicada pela terceira seta no eixo horizontal do gráfico da figura 3.21. . . . .	53
3.25	Gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.19. . . . .	54

---

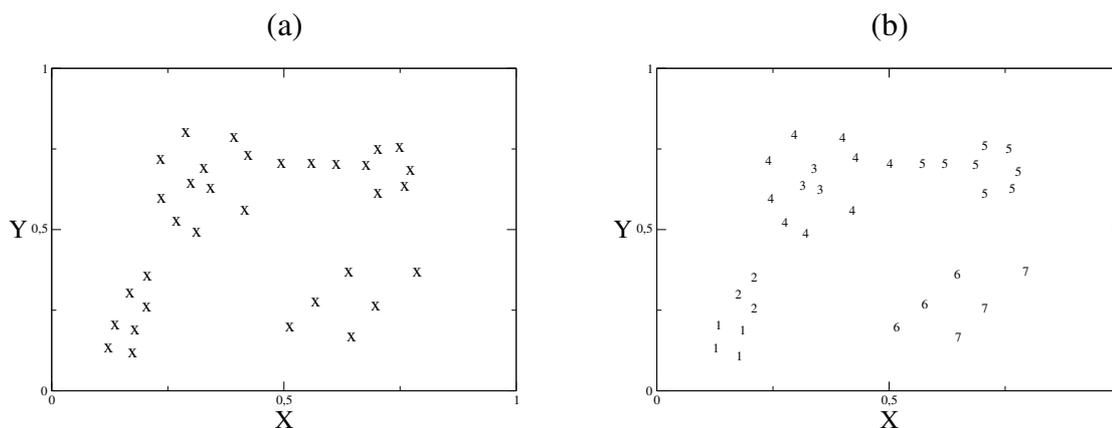
3.26	Classificação dos pontos da figura 3.19 no limite de informação $I = 0.02$ , indicado pelo primeira seta no eixo horizontal do gráfico da figura 3.25. . . . .	54
3.27	Classificação dos pontos da figura 3.19 no limite de informação $I = 0.1$ , indicado pela segunda seta no eixo horizontal do gráfico da figura 3.25. . . . .	55
3.28	Classificação dos pontos da figura 3.19 no limite de informação $I = 0.16$ , indicado pela terceira seta no eixo horizontal do gráfico da figura 3.25. . . . .	55
B.1	Gráfico da magnetização em função da temperatura para um sistema ferromagnético de $30 \times 30$ spins de Ising em uma rede quadrada, obtido através do algoritmo de Swendsen-Wang com 2000 passos de Monte Carlo sendo 1000 para a termalização. Note que as flutuações próximas a transição de fase são pequenas ao se aplicar o algoritmo de Swendsen-Wang. . . . .	88

# CAPÍTULO 1

## INTRODUÇÃO

Em todas as áreas, onde se necessita armazenar um grande número de informações ou dados, é imprescindível, seja em uma fase exploratória ou confirmatória, um agrupamento destes dados para se obter informações mais úteis. Áreas como biologia, psicologia, medicina, sensoriamento remoto e mercado financeiro, dentre outras [1, 2], utilizam-se de técnicas de agrupamento não supervisionado de dados, que doravante denominaremos por “*clustering*” (do inglês), sendo que cada área possui suas próprias terminologias e algoritmos específicos para este fim.

Em problemas de “*clustering*” o objetivo é agrupar dados ou padrões segundo suas similaridades, em classes naturais ou intrínsecas aos próprios dados ou padrões. Nestes problemas não existe uma solução “universal”. A solução procurada depende de um certo ponto de vista. Por exemplo, nós, humanos, possuímos uma excelente ferramenta de “*clustering*” em duas e três dimensões, que é a nossa visão. Todavia, diferentes indivíduos não necessariamente identificam os mesmos grupos em um certo conjunto de dados. As suas “medidas de similaridade” dependem de suas experiências educacionais e culturais, que se acumularam no decorrer da vida de cada um. Ou seja, um mesmo agrupamento de dados pode revelar diferentes informações úteis, dependendo de um dado contexto. Um médico pode se utilizar de dados para agrupar indivíduos, ressaltando atributos (relevantes para seu trabalho) como características sexuais ou raciais (nos referimos a fins éticos, é claro), um psicólogo exploraria características comportamentais, um cientista social estaria interessado em aspectos que ligam o indivíduo à sociedade, etc.



**Figura 1.1.** (a) Exemplo de grupos de dados, (b) Identificação dos grupos. Os números indicam diferentes grupos.

Em uma técnica de “*clustering*”, deve ser estabelecido um critério de similaridade consistente, para que esta técnica possa então ser “automatizada” e aplicada em larga escala, fornecendo resultados confiáveis. Olhando por este ângulo, podemos concluir que quanto mais o critério de similaridade refletir uma característica grupal, tanto mais abrangente será a técnica associada a esta medida.

Em física, muitos sistemas apresentam comportamentos coletivos, sendo estes comportamentos descritos por teorias bem estabelecidas. É de se esperar que a Física tenha muito a contribuir nesta área de agrupamento de dados, que já é bastante interdisciplinar. Nesta ótica, recentemente várias metodologias fundamentadas em modelos da Física foram propostas como “soluções” para este problema. Neste trabalho, apresentaremos e compararemos duas destas técnicas, aquela proposta por Domany *et al.* [3, 4, 5] fundamentada em um modelo ferromagnético de Potts, que exibe um comportamento coletivo segundo a densidade local de pontos, e outra proposta por De Carlo *et al.* [6], onde o comportamento coletivo emerge do acoplamento de mapas caóticos unidimensionais.

Esta dissertação está organizada como segue. Neste capítulo, daremos ênfase às definições de grupos (“*clusters*”) e começaremos a tratar o problema de uma forma mais matemática, definindo notações e representações utilizadas. Trataremos

de similaridade com alguns exemplos simples e, também, mostraremos uma das técnicas mais simples de agrupamento de dados, o método *k-means*.

No capítulo 2, trataremos dos métodos físicos ressaltando suas bases teóricas e mostrando suas implementações computacionais. No terceiro capítulo, aplicaremos estas técnicas a conjuntos de dados gerados artificialmente e apresentaremos algumas conclusões. No apêndice A, descreveremos a solução de campo médio para o modelo de Potts com  $q$ -estados relevante em problemas de “*clustering*” e no apêndice B, apresentaremos uma breve revisão sobre o método Monte Carlo.

## 1.1 DEFINIÇÕES

O objetivo de um método de “*clustering*” é classificar ou agrupar dados em classes naturais, intrínsecas, aos próprios dados. O problema de “*clustering*” pode ser colocado como segue.

Dados  $N$  padrões ou dados com  $D$  atributos ou propriedades, estes podem ser representados como pontos  $\vec{r}_i$ ,  $i = 1, 2, \dots$ , em um espaço  $D$ -dimensional. O problema de “*clustering*” é o de agrupar estes  $N$  pontos ou padrões em um número (não definido) de  $M$  grupos (“*clusters*” do inglês) definidos como:

**Definição 1:** Pontos pertencentes ao mesmo grupo, são pontos que possuem atributos ou propriedades mais similares entre si do que entre pontos pertencentes a outros grupos.

Para que esta definição seja precisa e útil é necessário definir o que significa ser “mais similar”.

Uma possível forma ou maneira de definir similaridade entre pontos pode ser obtida se for possível eleger um ponto como representativo de um grupo. Neste caso, então, um ponto estaria associado a este grupo dependendo de sua “distância” (faremos uma breve discussão sobre “distâncias” na próxima seção) a este ponto representativo. Este é um exemplo de uma abordagem de “*clustering*” chamada de *paramétrica*, cuja idéia é descrever os grupos através de alguns parâmetros previamente estabelecidos e utilizá-los para medir similaridade entre objetos. Neste caso o problema torna-se um problema de otimização onde são procuradas soluções *ótimas* para o problema de “*clustering*”.

Uma grande desvantagem destes métodos é que a estrutura (número de grupos, parâmetros dominantes, etc.) para os “*clusters*” é assumida previamente, e dependendo desta escolha, os resultados podem ser inócuos ou podem ocultar estruturas subjacentes. Então, quando não existe qualquer conhecimento prévio sobre

a estrutura dos dados é necessário adotar métodos *não-paramétricos*. Para isto é necessário estabelecer uma nova definição de grupos:

**Definição 2:** Os pontos que constituem um grupo, no espaço  $D$ -dimensional, encontram-se em uma região conectada deste espaço em que a densidade de pontos é alta. Dois grupos são separados por uma região em que a densidade de pontos é relativamente baixa [1].

Usando esta definição vemos que para se agrupar os pontos é necessário construir um contorno separando as regiões mais densas de regiões menos densas. Esta é a essência dos métodos não-paramétricos que, também, fornecem uma escala em que as densidades são estimadas e um critério para se definir entre “altas” e “baixas” densidades. No capítulo 2 deste trabalho, consideraremos dois recentes métodos de “*clustering*” não-paramétricos.

## 1.2 REPRESENTAÇÃO DE PADRÕES, SELEÇÃO E EXTRAÇÃO DE ATRIBUTOS

Um conjunto de  $N$  dados ou padrões pode ser representado por pontos em um espaço  $D$ -dimensional  $\varphi = \{\vec{r}_1, \dots, \vec{r}_N\}$  ou por uma matriz de “distâncias”  $N \times N$ , onde o conceito de “distância” não necessariamente é o de distância (euclidiana) do sentido usual. A dimensionalidade deste espaço é determinada pelo conjunto de atributos que compõe cada dado ou padrão  $\vec{r}_i = (x_{i1}, x_{i2}, \dots)$ , onde cada atributo representa um eixo de coordenadas deste espaço.

Um padrão pode caracterizar (ou descrever) objetos físicos ou abstratos e seus atributos podem ser classificados como:

- Características quantitativas: valores contínuos; valores discretos; intervalos de valores.

- Características qualitativas: nominais ou não-ordenadas; ordinais.

A seleção de atributos consiste no processo de identificar quais deles são relevantes no processo de “*clustering*”. A extração de atributos consiste em gerar novos atributos a partir de atributos já existentes. Quando necessário, tanto a seleção quanto a extração de atributos possibilitam uma melhor representação dos dados segundo um determinado contexto. Não existe regra geral para se escolher o melhor conjunto de atributos dentre outros, esta escolha depende do objetivo do processo de “*clustering*”. Um determinado processo de seleção ou extração, mesmo que simples, pode levar a uma melhora substancial no processo de “*clustering*”.

Como exemplo, consideremos o método da Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) [7]. Esta técnica tem como objetivo reduzir a dimensão  $D$  do espaço de atributos, seja para possibilitar uma visualização, ou para a aplicação de uma outra técnica.

Neste caso o espaço  $D$ -dimensional é projetado em um subespaço de menor dimensão e, para não haver uma perda significativa de informações, as correlações entre pontos devem ser “desacopladas”.

Para este fim calcula-se a matriz de covariância  $D \times D$ , entre os pontos  $\{\vec{r}_1, \dots, \vec{r}_n\}$

$$\begin{pmatrix} Cov(\mathbf{x}_1, \mathbf{x}_1) & Cov(\mathbf{x}_1, \mathbf{x}_2) & \dots & Cov(\mathbf{x}_1, \mathbf{x}_D) \\ Cov(\mathbf{x}_2, \mathbf{x}_1) & Cov(\mathbf{x}_2, \mathbf{x}_2) & \dots & Cov(\mathbf{x}_2, \mathbf{x}_D) \\ \vdots & \vdots & \dots & \vdots \\ Cov(\mathbf{x}_D, \mathbf{x}_1) & Cov(\mathbf{x}_D, \mathbf{x}_2) & \dots & Cov(\mathbf{x}_D, \mathbf{x}_D) \end{pmatrix}$$

onde

$$Cov(\mathbf{x}_i, \mathbf{x}_k) = \langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle) \cdot (\mathbf{x}_k - \langle \mathbf{x}_k \rangle) \rangle, \quad (1.1)$$

$\mathbf{x}_i$  é o vetor formado pelas  $i$ -ésimas componentes dos vetores  $\{\vec{r}_1, \dots, \vec{r}_n\}$  e  $\langle \dots \rangle$

refere-se à média dentre o universo de valores do  $i$ -ésimo atributo.

Esta matriz é hermitiana (real e simétrica), conseqüentemente podemos escrevê-la em uma forma diagonal com auto-valores reais. Assim, ao se projetar os pontos ou vetores na base de seus auto-vetores teremos pontos descorrelacionados. Para a redução do espaço escolhe-se os eixos cujos auto-vetores possuam os maiores auto-valores.

### 1.3 MEDIDAS DE SIMILARIDADE

Na definição de grupo o conceito de similaridade é fundamental. Devido as mais variadas formas de se definir similaridade, temos um grande número de técnicas de “*clustering*” disponíveis.

Em geral, as escalas em que são medidos os atributos são diferentes, então uma medida de “distância” deve ser cuidadosamente escolhida para que certos atributos não se sobreponham a outros. A medida mais comum é a distância euclidiana,

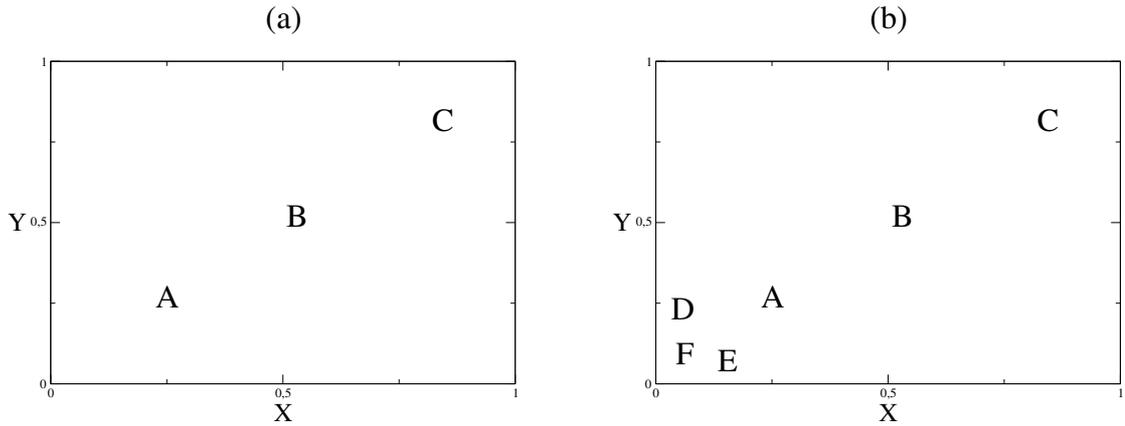
$$d(\vec{r}_i, \vec{r}_j) = |\vec{r}_i - \vec{r}_j| = \sqrt{\sum_{k=1}^D (x_{ik} - x_{jk})^2} \quad (1.2)$$

Esta métrica conduz a bons resultados quando o conjunto de dados formam clusters “compactos” ou “isolados”. Mas, como esta é uma medida local, em casos mais complexos os resultados não são satisfatórios.

Existem algumas medidas de distâncias que consideram os efeitos da vizinhança dos pontos, esta vizinhança é chamada de contexto. A similaridade entre dois pontos  $\vec{r}_i, \vec{r}_j$  em um dado contexto é definida por

$$s(\vec{r}_i, \vec{r}_j) = f(\vec{r}_i, \vec{r}_j, \mathcal{E}) \quad (1.3)$$

onde  $\mathcal{E}$  é o contexto.



**Figura 1.2.** Exemplo de similaridade com contexto. Na figura da esquerda,  $B$  é mais similar a  $A$  do que a  $C$ , enquanto que na figura da direita  $B$  é mais similar a  $C$  do que a  $A$ .

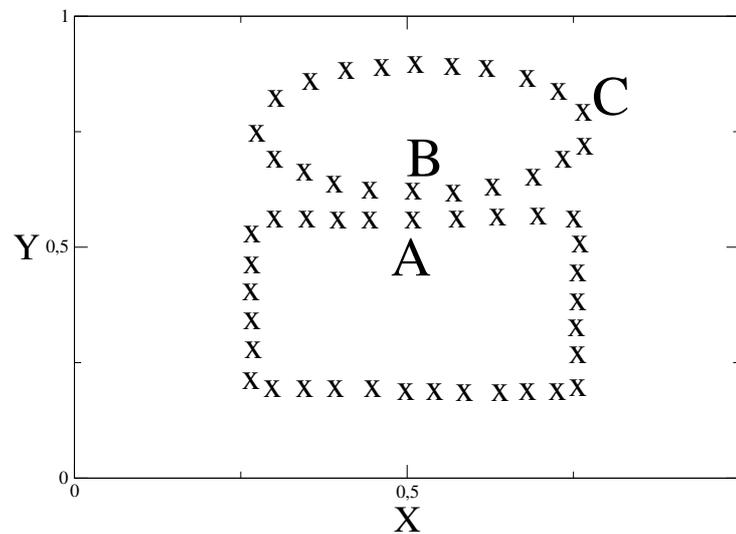
Como exemplo, a distância mútua entre vizinhos [8] (MND, do inglês *mutual neighbor distance*) é definida por

$$MND(\vec{r}_i, \vec{r}_j) = NN(\vec{r}_i, \vec{r}_j) + NN(\vec{r}_j, \vec{r}_i) \quad (1.4)$$

onde  $NN(\vec{r}_i, \vec{r}_j)$  é o número de pontos (vizinhos) pertencentes a uma região circular com centro em  $\vec{r}_j$  e raio  $|\vec{r}_i - \vec{r}_j|$ . Na figuras 1.2a e 1.2b apresentamos um exemplo. Na figura 1.2a o vizinho mais próximo de  $A$  é  $B$  e vice-versa. Então,  $NN(A, B) = NN(B, A) = 1$  e  $MND(A, B) = 2$ . Por sua vez, temos que  $NN(B, C) = 1$ , mas  $NN(C, B) = 2$  então  $MND(B, C) = 3$ . A figura 1.2b foi obtida da figura 1.2a adicionando-se três novos pontos  $D$ ,  $E$  e  $F$ . Agora  $MND(B, C) = 3$ , como antes, mas  $MND(A, B) = 5$ .  $MND(A, B)$  aumentou pela introdução dos pontos adicionais mesmo não se alterando as posições originais relativas entre os pontos  $A$  e  $B$ .

Existe ainda medidas de similaridade conceituais definidas como

$$s(\vec{r}_i, \vec{r}_j) = f(\vec{r}_i, \vec{r}_j, \mathcal{C}, \mathcal{E}) \quad (1.5)$$



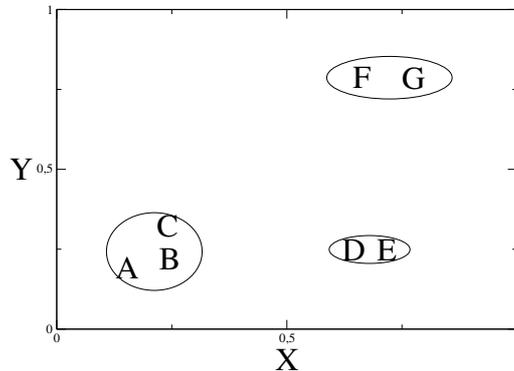
**Figura 1.3.** Exemplo de similaridade conceitual entre pontos, onde os grupos podem ser diferenciados pelo conceito geométrico de “retângulo” e “elipse”.

onde  $\mathcal{C}$  é um conjunto de conceitos pré-definidos. Esta noção é ilustrada na figura 1.3. Neste exemplo a distância euclidiana entre  $A$  e  $B$  é menor que a distância entre  $B$  e  $C$ . Entretanto  $B$  e  $C$  podem ser vistos como mais similares que  $A$  e  $B$  pois pertencem ao mesmo conceito “elipse” e  $A$  pertence a um diferente conceito “retângulo”.

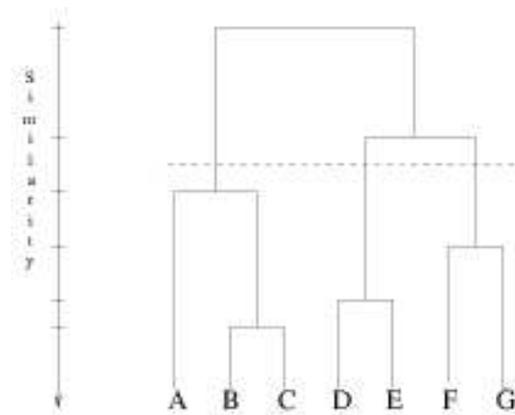
Trataremos neste trabalho de técnicas de “clustering”, inspiradas em modelos físicos sendo que a medida de similaridade emerge de um comportamento coletivo por isso a generalidade e a abrangência da aplicação destas técnicas é grande, indo desde a segmentação de imagens [9], até a análise de séries temporais [10, 11]. Estas técnicas serão discutidas nos capítulos posteriores.

## 1.4 MÉTODOS DE CLUSTERING

Os métodos de clustering podem ser classificados levando-se em consideração vários fatores tais como: se são aglomerativos ou divisivos, se são determinísticos



**Figura 1.4.** Exemplo de pontos agrupados segundo um nível de similaridade.



**Figura 1.5.** Dendrograma referente aos pontos da figura 1.4 [2].

ou estocásticos, etc. Mas, *grosso modo*, podemos classificar estes métodos como hierárquicos ou particionais.

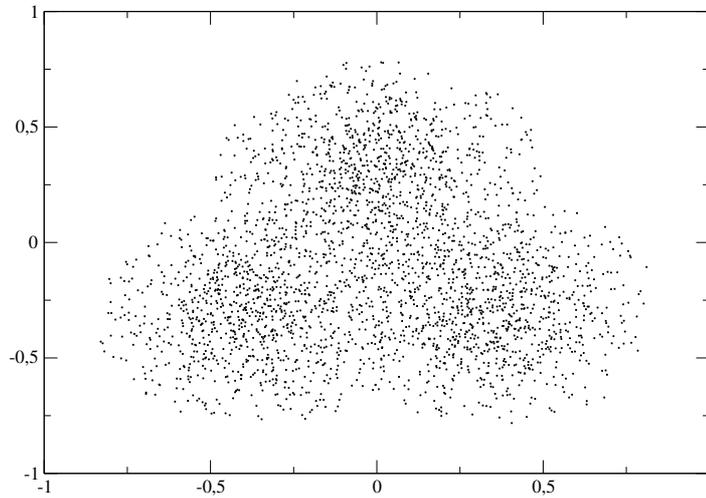
Os métodos hierárquicos produzem várias partições dos dados em série segundo um parâmetro que pode ser considerado como uma “resolução”. Uma técnica hierárquica produz um dendrograma que é a representação dos grupos formados segundo um nível de similaridade. Como exemplo mostramos na figura 1.5 um dendrograma referente à classificação dos pontos mostrados na figura 1.4, segundo uma técnica hierárquica. Os métodos que mostraremos nos capítulos posteriores são métodos hierárquicos onde um deles, o agrupamento super-paramagnético, utiliza propriedades de sistemas ferromagnéticos e tem a temperatura como parâmetro de resolução (como este método é estocástico, em alguns casos a partição dos dados pode não ser hierárquica). Em outra técnica que se utiliza do acoplamento de mapas caóticos unidimensionais, a medida de resolução é dada por um limiar da informação mútua entre pares de mapas. Os detalhes serão discutidos posteriormente.

Os métodos particionais são menos complexos que os métodos hierárquicos. Uma das poucas vantagens dos primeiros sobre os segundos é o menor custo computacional. Os métodos particionais são limitados, sendo que eles são mais empregados em conjuntos de dados que contém “clusters” isotrópicos sem muita complexidade.

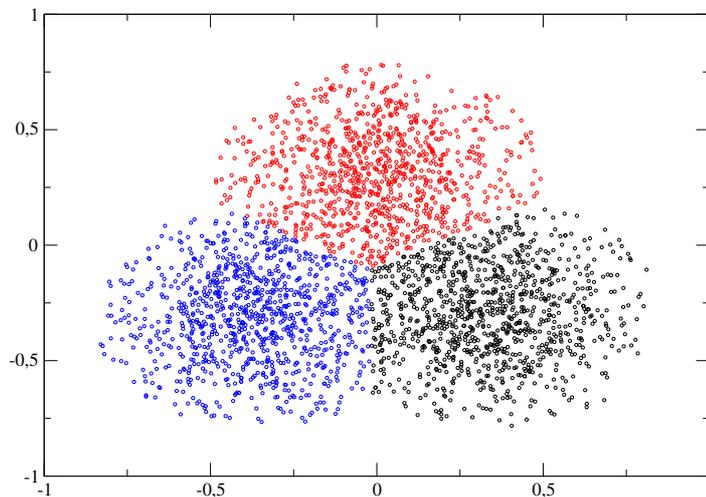
O critério mais intuitivo e mais utilizado em técnicas de “*clustering*” particionais é o critério do desvio quadrático, onde o exemplo mais simples é o algoritmo *k-means*. Este algoritmo tem o inconveniente de ser sensível às condições iniciais. O algoritmo *k-means* consiste em:

- i) Escolher aleatoriamente  $k$  centróides que podem ser padrões ou um ponto qualquer do hipervolume que contém os dados.
- ii) Associar cada padrão ao centróide mais próximo.
- iii) Calcular o novo centróide, a partir dos padrões recentemente associados ao cluster.
- iv) Se algum critério de convergência não foi alcançado retornar para o passo *ii*. Um possível critério de convergência pode ser a não associação de padrões a novos centróides.

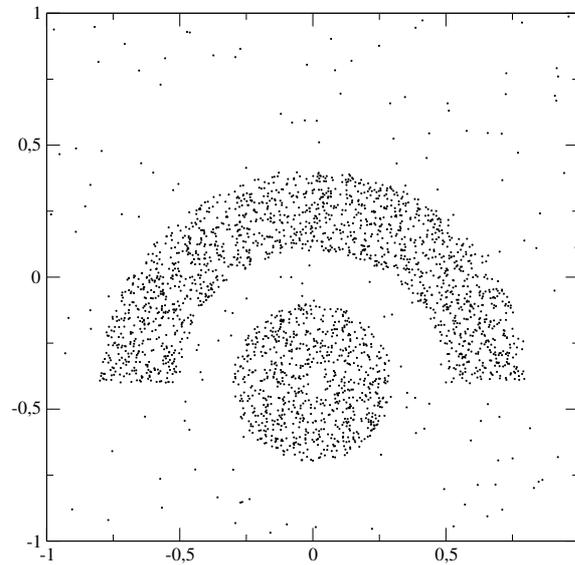
Implementamos o algoritmo *k-means* para dois conjuntos de dados artificiais. O primeiro conjunto de dados mostrado na figura 1.6 apresenta três clusters gerados por uma distribuição gaussiana com pontos médios diferentes. Na figura 1.7 apresentamos a classificação deste conjunto de dados utilizando o método *k-means*, onde três *clusters* de pontos mais densos são classificados. O segundo conjunto de dados mostrado na figura 1.8 apresenta conjuntos de pontos mais densos, porém com formas mais complexas. Na figura 1.9 apresentamos a classificação destes pontos pelo método *k-means*, onde os agrupamentos mais densos de pontos não foram identificados, como já se esperava.



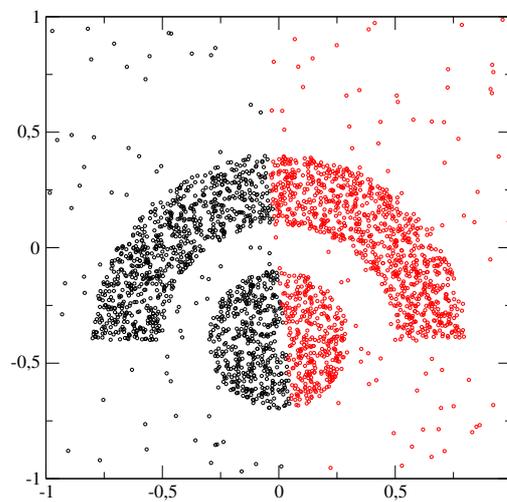
**Figura 1.6.** Três clusters gerados por uma distribuição gaussiana com pontos médios diferentes.



**Figura 1.7.** Classificação dos pontos da figura 1.6 pelo método  $k$ -means.



**Figura 1.8.** Clusters com formas mais complexas.



**Figura 1.9.** Classificação dos pontos da figura 1.8 pelo método *k-means*.

## CAPÍTULO 2

# OS MÉTODOS DE AGRUPAMENTO SUPER-PARAMAGNÉTICO E POR MAPAS CAÓTICOS

### 2.1 INTRODUÇÃO

Neste capítulo, trataremos de uma abordagem totalmente inovadora de agrupamento de dados. Este método, o agrupamento super-paramagnético, é baseado em um modelo físico, o modelo de Potts não-homogêneo ferromagnético e foi proposto por Domany *et al* [3, 4, 5] em 1996. Recentes aplicações deste método, citando somente algumas, para análise de séries temporais e agrupamento de expressões gênicas podem ser vistas nas referências [10] e [12, 13], respectivamente.

No agrupamento super-paramagnético, o critério de similaridade entre pontos, diferentemente da maioria das técnicas de agrupamento, não é diretamente relacionada à distâncias espaciais, mas emerge de um comportamento coletivo do sistema e é medido através de correlação spin-spin. Como consequência, em princípio, nenhum conhecimento prévio sobre a estrutura do conjunto de dados é necessário. A solução para o agrupamento dos dados é associada aos estados possíveis do modelo de Potts ferromagnético não-homogêneo. A grande vantagem é, que conjuntos de dados com estruturas relativamente complexas podem ser considerados, não se restringindo a utilização desta técnica apenas a casos particulares.

O método consiste basicamente em associar a cada dado (padrão em um

espaço  $D$ -dimensional), um spin de Potts de  $q$ -estados. Ao contrário do modelo de Potts ferromagnético padrão definido em uma rede regular com interações uniformes entre os spins de cada sítio [14], neste caso, temos uma rede não homogênea, onde define-se uma interação entre spins que decresce rapidamente com a distância entre pontos. Esta forma de interação é essencial para o método.

Definidas as interações entre pontos, simula-se, através do método Monte Carlo, utilizando a Mecânica Estatística de equilíbrio, a evolução deste sistema em contato com um reservatório térmico (à temperatura  $T$ ). O diagrama de fases deste modelo apresenta três regiões: fases ferromagnética, super-paramagnética e paramagnética. Na fase super-paramagnética emergem domínios de spins com o mesmo estado segundo a densidade local de pontos, e estes aglomerados são identificados com os grupos de dados.

A susceptibilidade magnética é utilizada para indicar a fase super-paramagnética, pois as flutuações na magnetização são muito maiores nesta fase do que nas outras fases. As correlações spin-spin, calculadas em uma certa temperatura da fase super-paramagnética, servem como critério de agrupamento, admitindo-se um certo limiar nesta correlação.

Como este método utiliza-se do método Monte Carlo para simular o sistema a cada temperatura, o agrupamento super-paramagnético pode demandar um razoável tempo computacional para conjuntos de dados relativamente extensos. Uma alternativa que se destaca por sua simplicidade e robustez é o agrupamento através de mapas caóticos [6] proposto por De Carlo *et al.* em 2000. Esta técnica é análoga ao método super-paramagnético onde a variável física associada ao elemento de dado é um mapa caótico unidimensional. O acoplamento entre mapas é feito com um termo de interação semelhante ao utilizado no agrupamento super-paramagnético. Ao se evoluir o sistema por um intervalo de tempo suficiente, o sistema alcança um regime estacionário, onde um comportamento coletivo emerge, sendo o comporta-

mento individual caótico. A medida de similaridade utilizada é a informação mútua média entre pares de mapas.

Como se tem somente iterações de mapas unidimensionais e estes evoluem uma única vez por um determinado intervalo de tempo, tem-se uma técnica computacionalmente rápida, e como veremos adiante produz resultados similares aos obtidos pelo agrupamento super-paramagnético, porém com algumas diferenças essenciais.

A organização deste capítulo é a seguinte: Na seção 2.2 é apresentada uma breve revisão sobre os modelos de Ising e de Potts. Na seção 2.4 é apresentado um diagrama de fases para um modelo de Potts na aproximação de campo médio que é usado para simular o problema de agrupamento de dados. Na seção 2.3 é mostrada uma justificativa formal para o método fundamentada na mecânica estatística. As seções 2.5, 2.6, 2.7 e 2.8, mostram como implementar o algoritmo de agrupamento super-paramagnético e a seção 2.9 introduz as idéias gerais do agrupamento por mapas caóticos.

## 2.2 O MODELO DE POTTS

O modelo de Ising é o modelo mais simples que reproduz as características básicas de um ferromagneto. É também o modelo mais pesquisado na Física Estatística.

O fenômeno do magnetismo em um material magnético não metálico é resultado da contribuição coletiva dos momentos magnéticos dos átomos que o constituem [15]. Estes momentos magnéticos acoplam-se via uma interação de troca, que é uma interação eletrostática de origem quântica resultante do caráter anti-simétrico da função de onda dos elétrons [16]. O modelo de Ising incorpora esta interação com um termo proporcional a  $\sigma_i\sigma_j$ , que é o produto dos spins.

No modelo de Ising os spins estão dispostos em uma rede, onde cada spin associado a um sítio  $i$  pode estar no estado “up” ( $\sigma_i = +1$ ) ou “down” ( $\sigma_i = -1$ ). Trata-se portanto de um modelo de dois níveis. No caso mais simples todas as interações entre spins têm o mesmo valor  $J$  e alcançam somente os primeiros vizinhos de um sítio  $i$ .

Um campo magnético externo  $B$  pode ser incluído no modelo resultando na Hamiltoniana,

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - B \sum_i \sigma_i \quad (2.1)$$

onde  $\langle i, j \rangle$  representa pares de spins vizinhos. Se  $J > 0$  o modelo é ferromagnético e os spins têm a tendência de se alinhar paralelamente. Se  $J < 0$  o modelo é anti-ferromagnético e os spins se alinham antiparalelamente.

Utilizando a mecânica estatística de equilíbrio, no ensemble canônico, para o modelo ferromagnético com  $B = 0$  e dimensões maiores ou igual a dois, é possível mostrar que existe um comportamento peculiar que se reflete na magnetização por spin definida como

$$m = \frac{1}{N} \langle \sum_i \sigma_i \rangle \quad (2.2)$$

onde  $N$  é o número de spins do sistema e  $\langle \dots \rangle$  indica a média termodinâmica sobre o ensemble canônico à temperatura absoluta  $T$ . Para valores de temperatura acima de  $T_c$  (temperatura crítica),  $m = 0$  e para valores de temperatura abaixo de  $T_c$ ,  $m \neq 0$ . Esta transição separa duas fases: a fase ferromagnética ( $m \neq 0$ ) e a fase paramagnética ( $m = 0$ ). O valor de  $m$  varia de uma forma abrupta, porém contínua, por isso esta transição de fase recebe o nome de segunda ordem ou contínua, em contraposição às transições ditas de primeira ordem que apresentam coexistência de fases e calor latente no ponto crítico.

O modelo de Potts é uma generalização do modelo Ising. Neste modelo os spins podem assumir  $q$ -estados com valores discretos,  $q = 1, 2, \dots$ , contendo, portanto, o caso Ising quando  $q = 2$ .

No modelo de Potts dois spins vizinhos contribuem com energia  $-J$  para a Hamiltoniana se eles possuem o mesmo valor ou contribuem com valor zero se possuem valores distintos, isto é,

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j} \quad (2.3)$$

onde  $\delta_{ij}$  é o delta de Kronecker.

No caso em que  $q = 2$  o modelo de Potts equivale ao modelo de Ising. Para valores maiores de  $q$ , o modelo de Potts tem um comportamento relativamente parecido com o modelo de Ising. Para  $J > 0$  (caso ferromagnético), o modelo de Potts possui  $q$ -estados fundamentais todos com os spins no mesmo estado. Quando a temperatura aumenta o suficiente, a ordem magnética de longo alcance é quebrada e o sistema exibe uma transição de fase onde os estados de spins ocorrem com a mesma frequência na rede. Entretanto existem algumas diferenças com relação ao modelo de Ising como por exemplo, a entropia no modelo de Potts com  $q > 2$  tende a ser maior que a entropia no modelo de Ising em uma mesma temperatura.

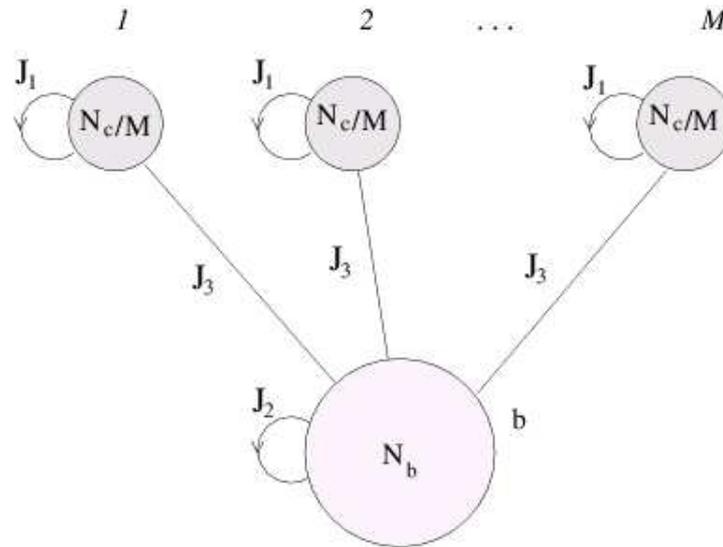
## 2.3 MECÂNICA ESTATÍSTICA DA CLASSIFICAÇÃO NÃO PARAMÉTRICA

O objetivo de um método de “*clustering*” não paramétrico é encontrar uma classificação “ideal”, sendo que esta classificação ideal pode variar drasticamente de técnica para técnica. Para um conjunto de dados representados por uma certa distribuição de pontos em um espaço de atributos  $D$ -dimensional, uma classificação, representada por  $\{s\}$ , pode ser construída associando-se a cada ponto  $\vec{r}_i$  da distribuição um “rótulo”  $s_i$  que pode assumir então um valor que varia de 1 a  $q$ . Definindo-se uma função custo que associa um valor negativo para pares de vizinhos com um mesmo rótulo e valor nulo em caso contrário, obtemos um problema de otimização para um dado conjunto de vínculos. Esta função, matematicamente, pode ser definida por,

$$\mathcal{H}[\{s\}] = - \sum_{\langle i,j \rangle} J_{ij} \delta_{s_i, s_j} \quad s_i = 1, \dots, q, \quad (2.4)$$

onde  $J_{ij}$  é uma função positiva que decresce com a distância  $|\vec{r}_i - \vec{r}_j|$ , e seu papel é de fazer com que pontos mais próximos sejam mais prováveis de pertencer ao mesmo grupo. A função 2.4, corresponde à Hamiltoniana do modelo de Potts ferromagnético não homogêneo [17]. Note que  $q$  não limita o número de agrupamentos, desde que é possível existir mais de um grupo com o mesmo valor de  $q$  em regiões diferentes do espaço  $D$ -dimensional.

Segundo a teoria da informação [18] devemos procurar uma distribuição de probabilidades que maximize a entropia para um dado vínculo  $\mathcal{H} = E$ . A distribuição resultante, corresponde à distribuição de probabilidades de um sistema de Potts em equilíbrio com um reservatório à temperatura absoluta  $T$ , com o inverso da temperatura  $\beta$  sendo o multiplicador de Lagrange correspondente ao vínculo  $E$ . Então, para o problema de agrupamento de dados caracterizados pela função 2.4, está disponível todo o arcabouço teórico fornecido pela mecânica estatística que



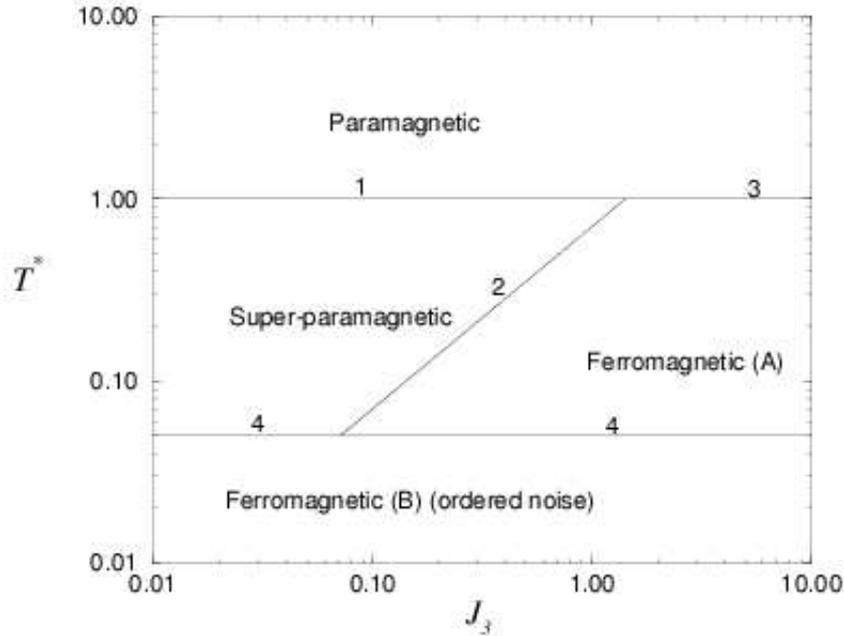
**Figura 2.1.** Representação esquemática de um modelo de Potts de campo médio, relevante para problemas de “clustering” [19]. Para maiores detalhes ver apêndice A.

fundamentará as simulações computacionais para este problema.

Nesta perspectiva, podemos prever que para altas temperaturas o sistema é desordenado (fase paramagnética). Quando a temperatura é decrescida ocorre uma transição de fase, onde spins de uma mesma região com densidade relativamente alta se alinham, formando os “super-spins”, sendo que os super-spins estão ainda descorrelacionados entre si. Esta é a fase super-paramagnética. Decrescendo ainda mais a temperatura, existe um alinhamento dos super-spins (fase ferromagnética *A*), e decrescendo ainda mais a temperatura, todo o sistema se alinha (fase ferromagnética *B*).

## 2.4 DIAGRAMA DE FASE DE UM MODELO DE CAMPO MÉDIO

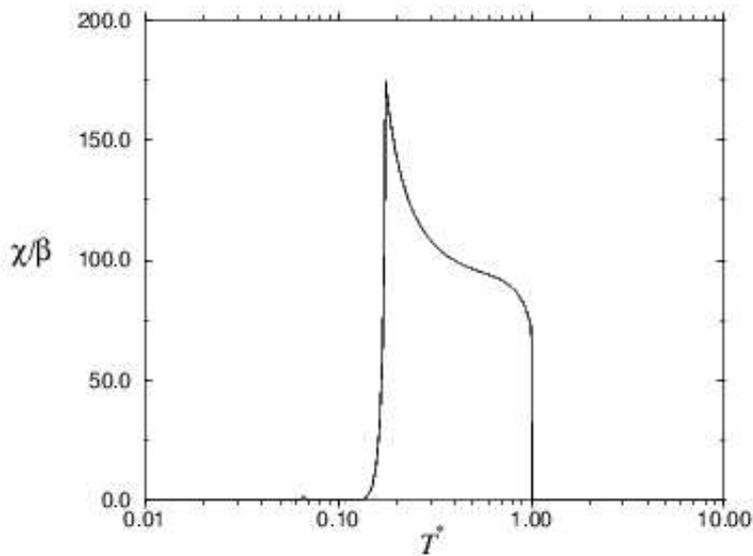
Em muitos problemas de agrupamento tem-se um conjunto de dados que formam grupos densos separados por regiões menos densas. Para simular esta situação Domany *et al.* apresentaram um modelo [19] que incorpora estas características,



**Figura 2.2.** Diagrama de fases obtido da solução da aproximação de campo médio de um modelo de Potts relevante para o problema de “clustering”. Este diagrama de fases foi obtido no limite termodinâmico  $\frac{N_b}{M} \rightarrow \infty$ , com  $J_1 = 1$  e  $J_2 = 0.05$ , e está em função de  $J_3$  e  $T^* = (2(q-1) \log(q-1) / (q-2)) T$  [19]. Para maiores detalhes ver apêndice A.

resolvendo-o analiticamente para vários limites, por meio da aproximação de campo médio. Os detalhes desta aproximação, para este modelo estão no apêndice A. O modelo consiste basicamente de  $N_c$  pontos divididos em  $M$  grupos sendo que a interação dentro de cada grupo é proporcional a  $J_1$ . Estes  $M$  grupos estão imersos em um “background” que contém  $N_b$  pontos, sendo que a interação dentro do “background” é proporcional a  $J_2$ . A interação entre spins dos grupos e spins do “background” é proporcional a  $J_3$ . Uma representação esquemática de tudo isto é mostrada na figura 2.1.

O diagrama de fases da figura 2.2, mostra os estados possíveis para este modelo, considerando-se que  $J_2 \ll J_1$ , o que é razoável, devido a forma das interações dos casos “reais”, que veremos mais adiante, onde as interações entre spins de regiões mais densas são maiores.



**Figura 2.3.** Gráfico da susceptibilidade magnética obtido da solução de campo médio de um modelo de Potts relevante para o problema de “clustering” [19]. Para maiores detalhes ver apêndice A.

A forma da susceptibilidade é mostrada na figura 2.3, onde o pico indica a transição ferromagnética ( $A$ ) - super-paramagnética. Para o valor de  $T^* = 1$  a susceptibilidade decresce abruptamente, indicando a fase paramagnética.

## 2.5 DEFINIÇÃO DE INTERAÇÕES LOCAIS

Por conveniência, no método super-paramagnético, são consideradas somente interações entre vizinhos. Como para problemas de agrupamento de dados temos que considerar redes sem simetria é necessário se definir vizinhança. Uma possível escolha é dividir a rede em celas de Voronoi [20] e considerar vizinhos, os pontos cujas celas possuem fronteiras em comum, mas para dimensões maiores que dois o custo computacional é elevado, como foi alegado em [19].

O critério de vizinhança que utilizamos é o critério de vizinhança mútua valor  $K$ , onde dois pontos  $\vec{r}_i$  e  $\vec{r}_j$  são vizinhos se  $\vec{r}_j$  é um dos  $K$  pontos mais próximos

(segundo a distância euclidiana) de  $\vec{r}_i$  e vice-versa. A partir desta definição é possível então se definir uma escala de comprimento local  $a$  que, em média, deve ser menor que as distâncias entre pontos vizinhos nas regiões mais densas e deve ser maior, em média, que as distâncias entre pontos vizinhos nas regiões menos densas. Uma possível escolha pode ser a distância média entre vizinhos.

Como foi mostrado em [19] o método super-paramagnético é pouco sensível em relação ao critério de vizinhança utilizado, bem como em relação à forma de interação  $J_{ij}$ , contanto que esta seja uma função positiva e rapidamente decrescente com a distância entre pontos. Uma possível escolha é,

$$J_{ij} = J_{ji} = \frac{1}{\hat{K}} \exp \left( -\frac{|\vec{r}_i - \vec{r}_j|^2}{2a^2} \right) \quad (2.5)$$

onde  $\hat{K}$  é o número médio de vizinhos por sítio.

## 2.6 CÁLCULO DE GRANDEZAS TERMODINÂMICAS

As grandezas que refletem as transições para fases mais ordenadas do sistema, são a susceptibilidade e a função de correlação. Uma vez escolhida a forma de interação  $J_{ij}$ , estas grandezas podem ser estimadas por métodos de simulação Monte Carlo. O algoritmo geralmente utilizado é o algoritmo de Swendsen-Wang [21, 22] (ver apêndice B). Este algoritmo divide toda a rede em clusters sendo que uma nova configuração é obtida associando-se probabilisticamente aos clusters uma nova variável de Potts, sendo todos os estados de Potts acessíveis. Isto é necessário, pois na fase super-paramagnética deseja-se evitar correlações acidentais entre spins. Este algoritmo também fornece uma estimativa para a função de correlação spin-spin

$$\langle \delta_{\sigma_i \sigma_j} \rangle = \frac{(q-1) \langle n_{ij} \rangle + 1}{q} \quad (2.6)$$

onde  $\langle n_{ij} \rangle$  é a probabilidade de  $\sigma_i$  e  $\sigma_j$  pertencerem ao mesmo cluster de Swendsen-Wang.

Recentemente, para processos de segmentação de imagens foi proposto por Feber *et al.* [9] um algoritmo que atualiza cada configuração, ou seja muda os estados dos elementos do cluster, de uma forma não independente levando-se em conta as interações de energia, ao contrário do algoritmo de Swendsen-Wang que muda os estados dos elementos do clusters de uma forma independente. Este algoritmo é o *ECU* ( do inglês *energy-sharing cluster algorithm*) e seu tempo de relaxação é cerca de *dez* vezes menor que o tempo de relaxação do algoritmo de Swendsen-Wang.

## 2.7 LOCALIZANDO A FASE SUPER-PARAMAGNÉTICA

O indicador da fase super-paramagnética é a susceptibilidade, que é proporcional à flutuação na magnetização  $m$ , isto é,

$$\chi = \frac{N}{T} (\langle m^2 \rangle - \langle m \rangle^2) \quad , \quad m = \frac{\left(\frac{N_{max}}{N}\right) q - 1}{q - 1} \quad (2.7)$$

onde  $N_{max} = \max \{N_1, N_2, \dots, N_q\}$  e  $N_\mu$  é o número total de spins com valor  $\mu$ .

A baixas temperaturas, as flutuações na magnetização são muito pequenas o que torna a susceptibilidade próxima de zero. Em  $T = T_{fs}$ , (temperatura de transição da fase ferromagnética para a fase super-paramagnética) a susceptibilidade apresenta um pico que indica o início da fase super-paramagnética. Na fase super-paramagnética as flutuações na magnetização dos clusters faz com que a susceptibilidade seja aproximadamente constante. Quando  $T = T_{ps}$ , ( $T_{ps} < T_{fs}$ ), a susceptibilidade decresce abruptamente tornando-se próxima de zero. Assim, entre a temperatura em que a susceptibilidade possui um máximo e a temperatura em que a susceptibilidade decresce abruptamente situa-se a fase super-paramagnética.

## 2.8 O PROCEDIMENTO DE AGRUPAMENTO

O método super-paramagnético consiste basicamente em identificar o intervalo de temperaturas onde ocorre a fase super-paramagnética, e em alguma temperatura neste intervalo calcular as correlações entre spins para identificar os clusters correspondentes. O procedimento pode ser resumido da seguinte forma:

- Associar a cada ponto  $\vec{r}_i$  um spin de Potts com  $q$ -estados.
- Encontrar os vizinhos deste ponto, segundo um critério de vizinhança.
- Calcular o valor das interações entre vizinhos utilizando a equação 2.5, ou outra interação similar.
- Realizar uma simulação Monte Carlo com a Hamiltoniana dada na equação 2.4 e calcular a susceptibilidade  $\chi$  para vários valores de temperatura.
- Identificar o intervalo de temperatura correspondente à fase super-paramagnética.
- Em uma certa temperatura nesta fase, estimar a função de correlação  $\langle \delta_{\sigma_i, \sigma_j} \rangle$ , para todos os pares de vizinhos  $\vec{r}_i$  e  $\vec{r}_j$ .
- Identificar os clusters da seguinte forma: se  $\langle \delta_{\sigma_i, \sigma_j} \rangle > \theta$ , os pontos  $\vec{r}_i$  e  $\vec{r}_j$  são definidos como similares. Os clusters são construídos segundo a regra da transitividade, ou seja, se  $\vec{r}_i$  é similar a  $\vec{r}_j$ , e  $\vec{r}_j$  é similar a  $\vec{r}_k$  então  $\vec{r}_i$  é similar a  $\vec{r}_k$ .

## 2.9 AGRUPAMENTO POR MAPAS CAÓTICOS

Seguindo a mesma “filosofia” do agrupamento super-paramagnético, surgiram outras técnicas de *clustering*, como a segmentação de imagens por sistemas neurais proposta por Neirotti *et al.* [23]. Nesta seção trataremos de uma técnica

proposta por De Carlo *et al.* [6] onde a classificação de dados emerge da sincronização entre mapas caóticos acoplados. A cada ponto da rede, é associada uma variável contínua  $x_i \in [-1, 1]$  e esta variável evolui caoticamente por meio de um mapa unidimensional acoplado ao dos seus vizinhos. O sistema, após um intervalo suficiente de tempo, alcança um regime estacionário, e agrupamentos de mapas com uma certa sincronia aparecem segundo a densidade local de pontos. As interações entre mapas é a mesma utilizada no agrupamento super-paramagnético, sendo  $J_{ij} = J_{ji} = \exp\left(-\frac{|\vec{r}_i - \vec{r}_j|^2}{2a^2}\right)$ , onde  $a$  novamente é uma escala local identificada como a distância média entre vizinhos.

A evolução do sistema é dada por,

$$x_i(t+1) = \frac{1}{C_i} \sum_{j \neq i} J_{ij} f(x_j(t)) \quad (2.8)$$

onde a atualização de cada configuração ocorre paralelamente,  $C_i = \sum_{j \neq i} J_{ij}$ , e o mapa escolhido é o mapa logístico  $f(x) = 1 - 2x^2$ .

O regime estacionário da equação 2.8 é um atrator macroscópico independente das condições iniciais. Após o sistema evoluir para este estado, é possível então serem medidas as correlações entre pares de mapas. A medida de sincronização utilizada originalmente por De Carlo *et al.* foi a *informação mútua média*  $I_{ij}$  entre pares de mapas cuja definição é a seguinte: Se o estado de um elemento  $x_i(t) > 0$  então associa-se a esta variável o bit  $S_i = 1$ , senão associa-se o bit  $S_i = 0$ . Em um determinado intervalo de tempo no regime estacionário, este procedimento gera uma sequência de bits que permite o cálculo da entropia de Boltzmann  $H_i$  para a variável  $x_i$ ,

$$H_i = - \sum_{S_i=0,1} P(S_i) \log P(S_i) \quad (2.9)$$

bem como permitem o cálculo da entropia conjunta entre as variáveis  $x_i$  e  $x_j$ ,

$$H_{ij} = - \sum_{S_i=0,1} \sum_{S_j=0,1} P(S_i, S_j) \log P(S_i, S_j) \quad (2.10)$$

A informação mútua média é então calculada de,

$$I_{ij} = H_i + H_j - H_{ij}. \quad (2.11)$$

Se os mapas evoluem independentemente  $I_{ij} = 0$ , e se os mapas estão completamente sincronizados  $I_{ij} = \log 2$ , devido a escolha de  $f$ .

### 2.9.1 O Método Computacional

Definidas a dinâmica e a medida de sincronia, podemos agora detalhar o método computacional. Primeiramente o valor de  $a$ , como já foi mencionado, é a distância média entre vizinhos. A vizinhança utilizada é a vizinhança mútua valor  $K$ , sendo o método bastante robusto quanto ao valor de  $K$  desde que um valor mínimo de  $K$  seja alcançado para que haja um nível mínimo de conexão na rede. No trabalho original, onde este método foi apresentado, foram consideradas somente interações com vizinhos  $\vec{r}_j$  que estivessem a menos de  $3a$  de distância de um ponto  $\vec{r}_i$ . Neste trabalho consideramos todos os vizinhos, assim como foi feito no método super-paramagnético.

Iniciando com uma configuração aleatória  $\{x\}$  (um valor no intervalo  $[-1, 1]$  é escolhido aleatoriamente como condição inicial para cada variável), os mapas são iterados até seu regime estacionário e, então, são medidas as informações mútuas médias entre pares de vizinhos. Os clusters são identificados da seguinte forma:

- Uma conexão é estabelecida entre um par de pontos, se  $I_{ij} > \theta$ , onde  $\theta$  é o limiar na informação mútua média.

- Os agrupamentos de dados são identificados como componentes conectadas do grafo obtido no passo anterior.

O valor de  $\theta$  controla a resolução com que os dados são agrupados e uma hierarquia de agrupamentos emerge ao se variar  $\theta$ .

No artigo onde o método foi apresentado [6] foram feitas comparações com o método de agrupamento super-paramagnético através do agrupamento de dados de imagens de satélite. A imagem consistia de 1489 pixels cada um com seis bandas espectrais. Os proponentes ressaltaram que o agrupamento por mapas caóticos classifica mais pontos, com a mesma pureza do método super-paramagnético, onde pureza significa a porcentagem de pontos classificados "corretamente".

Algumas recentes aplicações do método de agrupamento através de mapas caóticos são: detecção de minas através de imagens de infravermelho [24], diagnose do mal de Huntington por padrões encefalográficos [25] e aplicações em mercado financeiro [11]

## CAPÍTULO 3

# APLICAÇÕES A DADOS ARTIFICIAIS

Neste capítulo aplicaremos os métodos descritos no capítulo 2 a conjuntos de dados artificiais comparando a performance dos dois métodos. Tais conjuntos de dados são representados por padrões formados por conjuntos de pontos no espaço de atributos bidimensional, mais apropriados para a compreensão e visualização das potencialidades dos métodos. Apresentaremos gráficos de susceptibilidade magnética, classificações e gráficos que mostram a evolução do tamanho dos maiores clusters em função do parâmetro de resolução, onde “tamanho” não significa a “área” ocupada pelo cluster, mas o número de pontos que o compõe.

Apesar das duas metodologias serem similares veremos que cada uma tem suas peculiaridades, é claro, por estarem fundamentadas em diferentes sistemas físicos. Estas peculiaridades podem tornar um método mais apropriado do que o outro dependendo da aplicação.

Em todos os resultados obtidos utilizamos, no critério de vizinhança valor  $K$ , o valor de  $K$  igual a 20, para que haja uma conexão suficiente em toda a rede. Nas aplicações utilizando o agrupamento super-paramagnético utilizamos o limite de correlação spin-spin ( $\theta$ ), igual a 0.5 e o número de estados de Potts  $q$  igual a 10 (Uma discussão sobre a estabilidade do método de agrupamento super-paramagnético com relação ao limite de correlação spin-spin e o número de estados de Potts pode ser vista na referência [19]).

### 3.1 CONJUNTO DE DADOS A

O primeiro conjunto de dados ou padrões, está mostrado na figura 3.1. Consiste em uma região do espaço bidimensional com pontos cujas coordenadas estão no intervalo  $[-1, 1]$  contendo dois clusters mais densos, um círculo com 800 pontos e um setor circular, concêntrico ao círculo, com 1600 pontos. Estes dois grupos estão imersos em um fundo ou “background” de 1050 pontos. Tanto os clusters, quanto o *background* possuem densidades uniformes, porém diferentes. Os dois clusters são construídos de forma a terem a mesma densidade de pontos. Este conjunto de pontos foi escolhido de acordo com o da referência [6], com o *background* possuindo mais pontos.

Aplicando o agrupamento super-paramagnético, obtivemos um gráfico para a susceptibilidade, como mostra a figura 3.2. Em temperaturas  $T$  menores que 0,03, aproximadamente, temos a fase ferromagnética com quase todos os pontos formando um único cluster. Em  $T \approx 0.03$ , ocorre uma pseudo-transição de fase (pois este sistema é finito) indicada pelo maior pico na susceptibilidade. Entre  $T \approx 0.03$  e  $T \approx 0.15$ , identificamos a fase super-paramagnética onde as flutuações mostradas na figura 3.2 são causadas pelos super-spins. Para temperaturas acima de  $T \approx 0.15$ , tem-se a fase paramagnética onde os spins estão descorrelacionados.

A figura 3.3 mostra a evolução do tamanho dos clusters com a temperatura. É interessante ressaltar, que na região correspondente à fase super-paramagnética, é verificado que o tamanho dos maiores clusters é estável nesta fase, em um grande intervalo de temperatura (0,03 - 0,08), indicando a robustez do método. A seta no eixo horizontal da figura 3.3 indica a temperatura  $T = 0.05$  onde fizemos uma classificação, mostrada na figura 3.4. As cores diferentes indicam os quatro maiores agrupamentos encontrados pelo método. O tamanho destes agrupamentos são respectivamente, 1656, 816, 26, e 14 pontos. Observe que os dois primeiros se aproximam dos valores usados para a construção artificial dos clusters com erros de 3,5%

---

e 2%, respectivamente. É importante ressaltar que o terceiro e quarto maiores clusters são muito pequenos ( $\sim 10^2$ ) em relação aos dois primeiros e são decorrentes de flutuações na densidade espacial do *background*.

A aplicação do método dos mapas caóticos ao conjunto de pontos da figura 3.1 não produziu um resultado razoável, onde se observou que partes de ambos os agrupamentos mais densos pertencem ao mesmo cluster, como é mostrado na figura 3.5. Como será visto adiante em outro conjunto de dados, este método tem a tendência de formar “trilhas” ou “caminhos” presentes aleatoriamente no *background*.

Para se obter um resultado razoável através do agrupamento usando os mapas caóticos foi necessário diluir o *background* do conjunto de dados mostrados na figura 3.1. Os dados mostrados na figura 3.6 consistem em dois agrupamentos mais densos, com tamanhos e formas similares aos mostrados na figura 3.1. Neste caso os dois agrupamentos possuem 1672 e 768 pontos respectivamente e estão imersos em um *background* de 148 pontos.

O gráfico da figura 3.7, mostra a evolução do tamanho dos maiores clusters em função do limite de informação entre pares de mapas. Notemos que para  $I < 0.04$ , aproximadamente, todos os pontos formam um único cluster. Em  $0.04 < I < 0.18$  observamos que os tamanhos dos dois maiores clusters permanecem estáveis. Para  $I > 0.18$ , aproximadamente, os clusters são fragmentados até se tornarem clusters com um único ponto.

A seta no gráfico da figura 3.7, indica o valor da informação entre pares de mapas onde fizemos uma classificação. Esta classificação está mostrada na figura 3.8. Neste caso o tamanho dos quatro maiores clusters são: 1691, 774, 4 e 3 pontos, respectivamente. De modo análogo ao agrupamento super-paramagnético os dois maiores clusters reproduzem os dados artificiais com erros de 1,1% e 0,8%, respectivamente. Todavia este resultado seria esperado haja vista que a densidade

do *background* foi reduzida por um fator  $\sim 7$ .

---

## 3.2 CONJUNTO DE DADOS B

O conjunto de dados  $B$  mostrado na figura 3.9 consiste de dois grupos cada um com 714 pontos imersos em um *background* com 1000 pontos, e densidade  $\sigma_b$ . Os dois grupos não possuem densidades uniformes. Cada região está longitudinalmente localizada em:  $-0.7 < x < -0.3$  e  $0.3 < x < 0.7$  com  $-0.4 < y < 0.4$  e dividida em duas subregiões com densidades  $\sigma_1 = 10\sigma_b$  e  $\sigma_2 = 5\sigma_b$ , respectivamente à esquerda e à direita.

O objetivo deste conjunto de dados artificiais é testar a potencialidade das técnicas com relação à classificação de padrões com regiões de gradientes de densidade. A figura 3.10 mostra a susceptibilidade deste conjunto de pontos obtida aplicando-se o agrupamento super-paramagnético indicando o intervalo de temperaturas da fase super-paramagnética deste sistema entre  $\mathcal{T} \approx 0.05$  e  $\mathcal{T} \approx 0.17$ .

A figura 3.11 mostra a evolução dos tamanhos dos clusters em função da temperatura para o conjunto de dados  $B$  ao se utilizar o agrupamento super-paramagnético e a figura 3.12 mostra uma ampliação da região em destaque na figura 3.11. Na primeira classificação, mostrada na figura 3.15 os dois maiores grupos (retângulos com  $\sim 700$  pontos) são identificados em um grande intervalo de temperatura (0.05 – 0.09) indicando que este método é robusto nestes níveis de “resolução”. Na segunda classificação apresentada na figura 3.16, a região mais densa do agrupamento localizado à esquerda é quase que totalmente identificada, mas a região mais densa do agrupamento localizado à direita é fragmentada em duas partes devida a maior sensibilidade do método com relação às flutuações de densidade. O parâmetro  $a$  da equação 2.5 está relacionado com esta sensibilidade.

A figura 3.13 mostra a evolução dos tamanhos dos clusters em função da informação entre pares de mapas ao se utilizar o agrupamento através de mapas caóticos e a figura 3.14 mostra uma ampliação da região em destaque na figura 3.13. Na primeira classificação mostrada na figura 3.17, os dois maiores agrupamentos

possuem 817 e 696 pontos. Os dois retângulos mais densos em relação ao *background* são identificados com certa precisão, mas pontos do *background* são incorporados a estes agrupamentos. A segunda classificação apresentada na figura 3.18 mostra que este método não distinguiu os níveis de densidade presente nos retângulos.

Para efeito de comparação medimos o tempo de execução dos programas computacionais que geraram os dados das figuras 3.11 e 3.13. Para o agrupamento super-paramagnético que gerou os dados da figura 3.11 o tempo foi de aproximadamente 53.25h, sendo que para cada passo de temperatura foram feitas classificações. Para o agrupamento através de mapas caóticos que gerou os dados da figura 3.13 o tempo foi de 12.15min.

### 3.3 CONJUNTO DE DADOS C

No conjunto de dados mostrado na figura 3.19, substituímos o efeito do background por uma trajetória com 41 pontos conectando dois agrupamentos. Um destes agrupamentos consiste de um círculo envolto concentricamente por um “aro” circular. Há 500 pontos no círculo e 600 pontos no aro circular. O outro agrupamento consiste de um círculo menor com 120 pontos. O objetivo destas formas é mostrar como são transmitidas as correlações nos dois agrupamentos.

A figura 3.20 mostra a susceptibilidade obtida utilizando-se o agrupamento super-paramagnético. Observamos que esta forma para a susceptibilidade é diferente daquela mostrada para a solução da aproximação de campo médio (figura 2.3), pois naquele caso é considerado um background bem mais efetivo. Mas, mesmo assim a susceptibilidade ainda fornece um indicativo para as classificações. O primeiro pico no gráfico da figura 3.20 é devido a fragmentação da trajetória de pontos e está indicando o início da fase superparamagnética. Os maiores picos estão relacionados com a fragmentação dos dois agrupamentos.

A figura 3.21 mostra o gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura para os dados da figura 3.19 ao se aplicar o agrupamento super-paramagnético. Neste gráfico, existem três intervalos onde os agrupamentos são mais estáveis. Estes intervalos para a temperatura são:  $\mathcal{T} < 0.020$ ,  $0.020 < \mathcal{T} < 0.048$  e  $0.048 < \mathcal{T} < 0.064$ , aproximadamente. As setas indicam os pontos onde foram realizadas as classificações.

As figuras 3.22, 3.23 e 3.24 mostram classificações utilizando o agrupamento super-paramagnético em três temperaturas consecutivas. A primeira classificação (figura 3.22) foi obtida para uma temperatura  $\mathcal{T} = 0.01$ . Todos os pontos formam um único cluster. Na segunda classificação (figura 3.23),  $\mathcal{T} = 0.04$ , a linha que une os dois agrupamentos é fragmentada separando dois maiores agrupamentos com tamanhos de 1102 e 120. Os outros clusters classificados tiveram tamanho unitário.

---

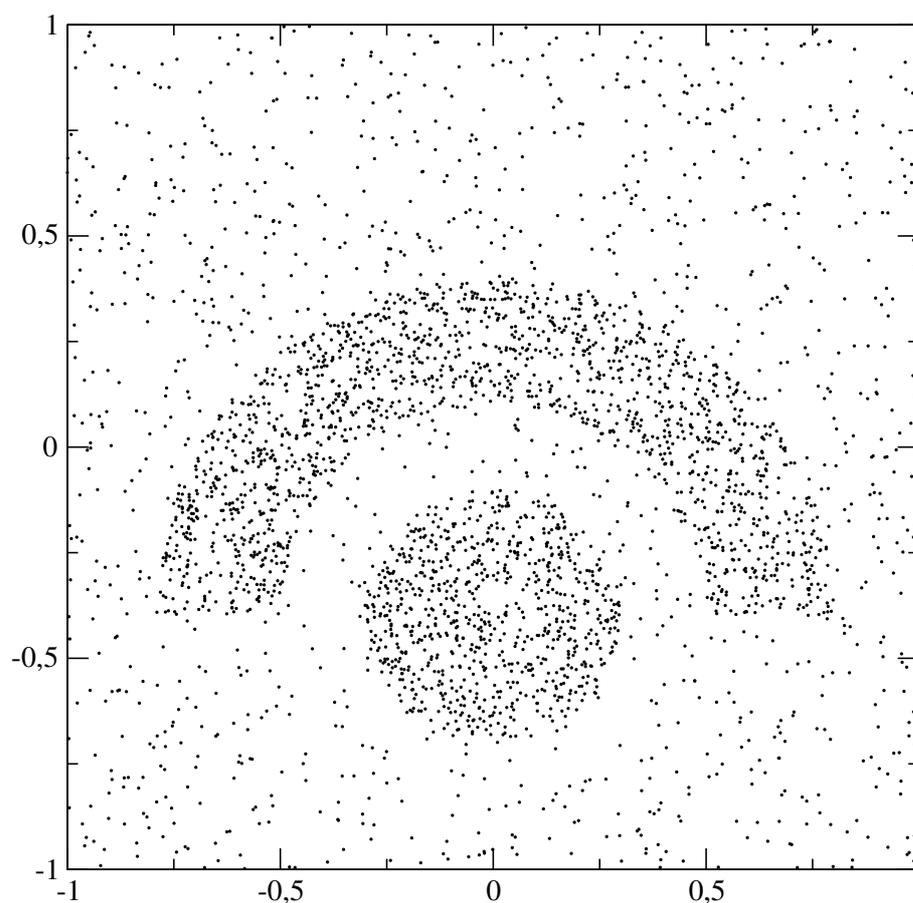
Na terceira classificação (figura 3.24),  $\mathcal{T} = 0.055$ , o agrupamento maior é finalmente dividido em um círculo, com 498 pontos, e em um "aro" de pontos, com 600 pontos. Os outros agrupamentos tiveram tamanhos de 120 e 2 pontos. É importante notar que, como estamos considerando um sistema ferromagnético unidimensional, a trajetória de pontos torna-se instável para pequenas flutuações na energia e os pontos tornam-se descorrelacionados. Este efeito é ilustrado na fragmentação da linha de pontos da segunda classificação (figura 3.23). Esta é a razão para que o método de agrupamento super-paramagnético não incorpora as "trilhas" ou "caminhos" unidimensionais do *background* aos clusters. Esse é um resultado importante porque a inexistência de ordem magnética de longo alcance em sistemas unidimensionais é um teorema da Física Estatística (teorema de Mermin) e portanto um resultado robusto, e não apenas do modelo de Potts. Qualquer outra técnica usando modelos de spin vai possuir esta característica. Portanto modelos desta classe são apropriados para eliminar *backgrounds* e ruídos de fundo.

A figura 3.25 mostra o gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas para os dados da figura 3.19 ao se aplicar o agrupamento através de mapas caóticos. Neste gráfico, existe três intervalos onde os agrupamentos são mas estáveis. Estes intervalos para a informação são:  $I < 0.053$ ,  $0.053 < I < 0.15$  e  $0.15 < I < 0.19$ , aproximadamente. As setas indicam os pontos onde foram realizadas as classificações.

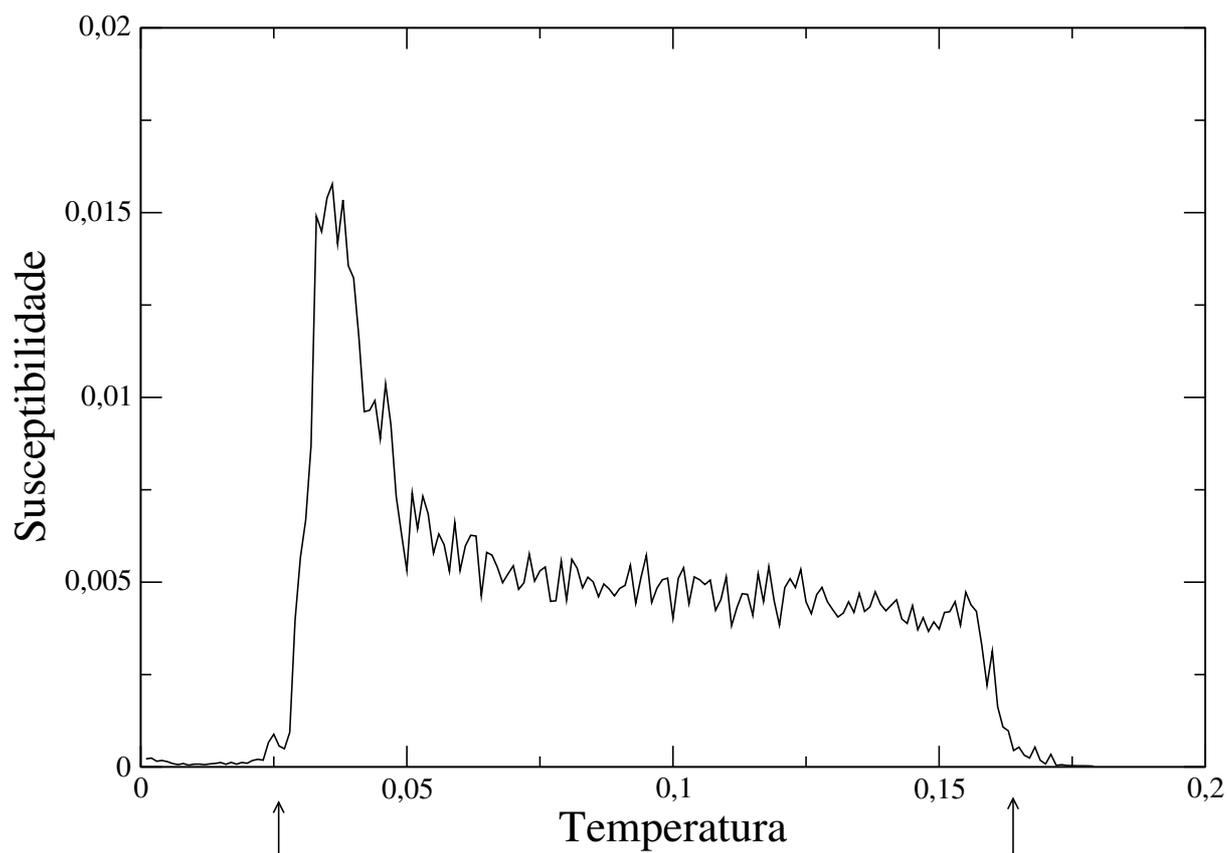
As figuras 3.26, 3.27 e 3.28 mostram classificações utilizando o agrupamento através de mapas caóticos. Na primeira classificação, (figura 3.26), todos os pontos pertencem a um único cluster. Na segunda classificação, ao contrário da segunda classificação apresentada para o agrupamento super-paramagnético (figura 3.23), a linha de pontos que une os agrupamentos é preservada sendo que o círculo interno do maior agrupamento se descorrelaciona de todo o grupo. O tamanho dos agrupamentos identificados pelo método são: 761 e 500 pontos. Nesta classificação o método

de agrupamento através de mapas caóticos conseguiu classificar um cluster de forma bastante complexa (“aro” de pontos conectado por uma trajetória a um círculo de pontos). Na terceira classificação a linha de pontos ainda se conserva, sendo agora que o “aro” de pontos se fragmenta em duas partes. Nesta classificação o tamanho dos agrupamentos encontrados pelo método são: 499, 412, 347 e 2 pontos.

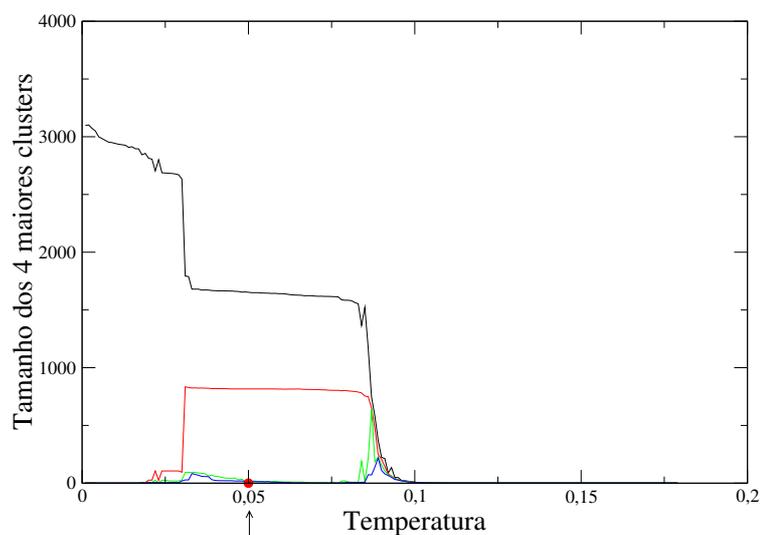
Os dois métodos produzem resultados diversos porque o método de agrupamento super-paramagnético, exclui os clusters unidimensionais. Este método estima densidades com maior precisão. No agrupamento através de mapas caóticos acoplados as conexões locais tem bastante influência na transmissão das correlações.



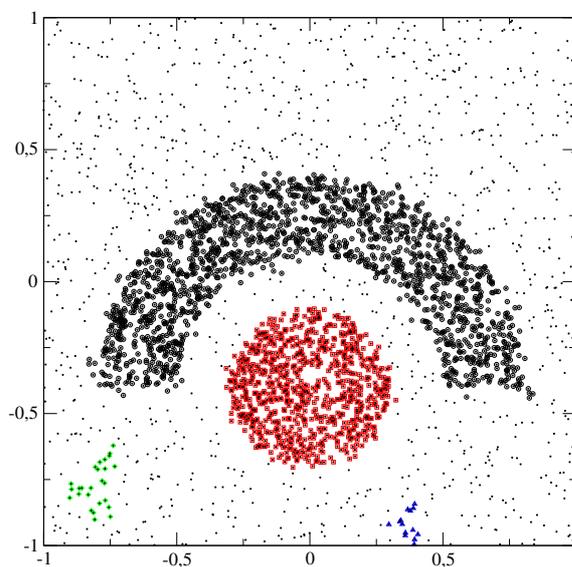
**Figura 3.1.** Conjunto de dados artificiais consistindo de dois grupos: um setor de círculo com 1600 e um círculo com 800 pontos, ambos com mesma densidade de pontos e gerados por uma distribuição de probabilidades uniforme, imersos em um background de 1050 pontos.



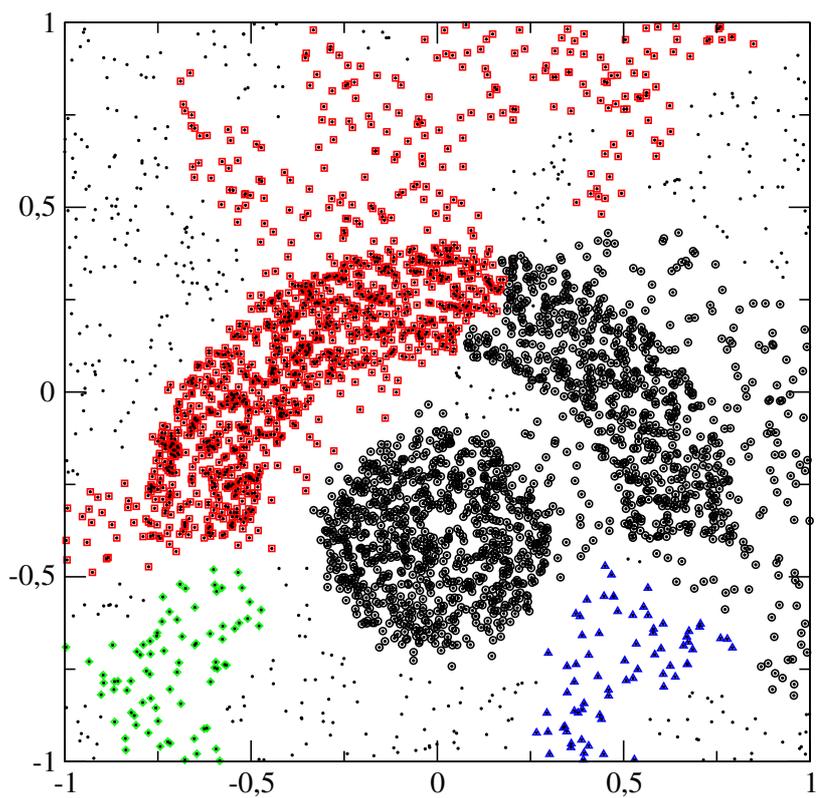
**Figura 3.2.** Gráfico da susceptibilidade referente ao conjunto de pontos da figura 3.1 ao se utilizar o agrupamento super-paramagnético. A região de flutuação (indicada pelas setas, entre 0.03 e 0.16, aproximadamente) em torno de um valor aproximadamente constante e diferente de zero, indica a fase super-paramagnética.



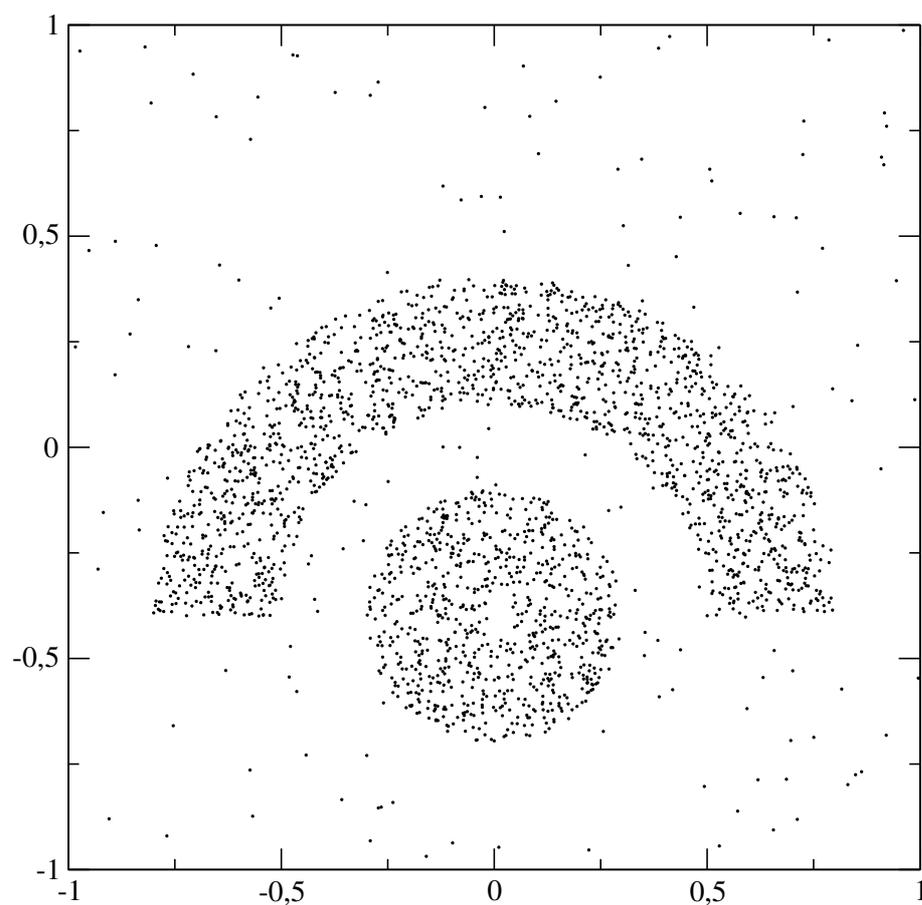
**Figura 3.3.** Gráfico da evolução do tamanho dos clusters em função da temperatura ( $\mathcal{T} = \frac{T}{K}$ ). A região entre  $\mathcal{T} \approx 0.03$  e  $\mathcal{T} \approx 0.08$  é caracterizada pela estabilidade dos dois maiores clusters. A seta indica o valor de  $\mathcal{T} = 0.05$  em que foi feita a classificação, apresentada na figura 3.4.



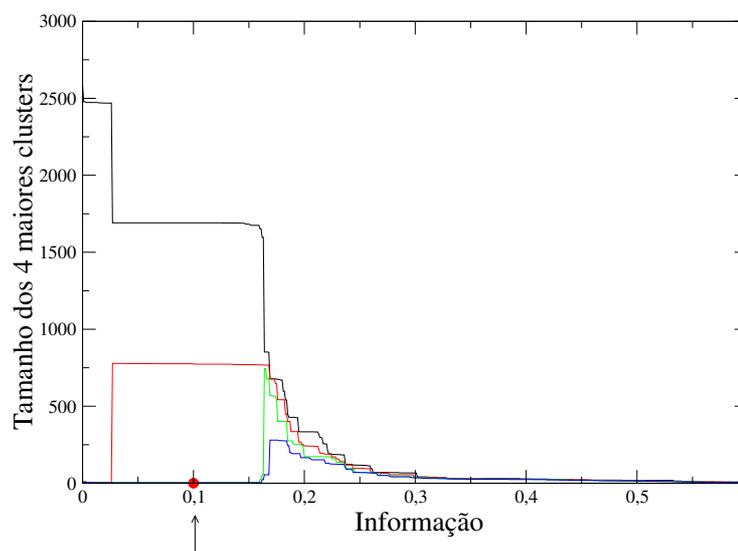
**Figura 3.4.** Classificação dos pontos da figura 3.1 em uma temperatura  $\mathcal{T} = 0.05$  indicada pela seta no eixo horizontal do gráfico da figura 3.3. Os dois maiores agrupamentos, consistem de 1656 (em preto) e 816 (em vermelho) pontos, respectivamente.



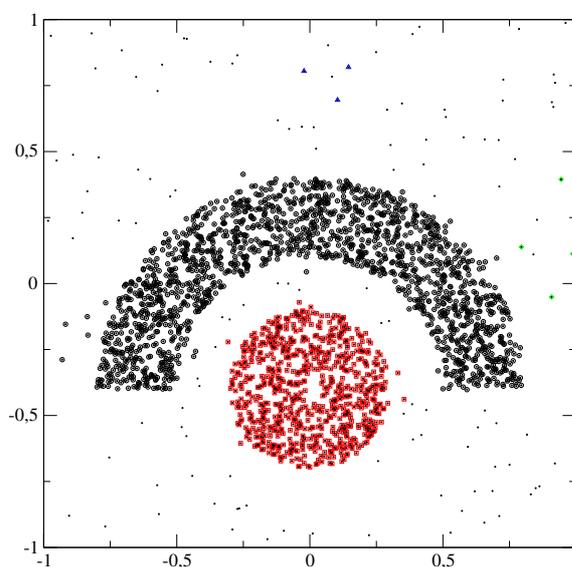
**Figura 3.5.** Classificação do conjunto de dados mostrado na figura 3.1 ao se aplicar o agrupamento por mapas caóticos em um limite de informação  $I = 0.12$ . Os dois maiores agrupamentos mais densos não são identificados corretamente.



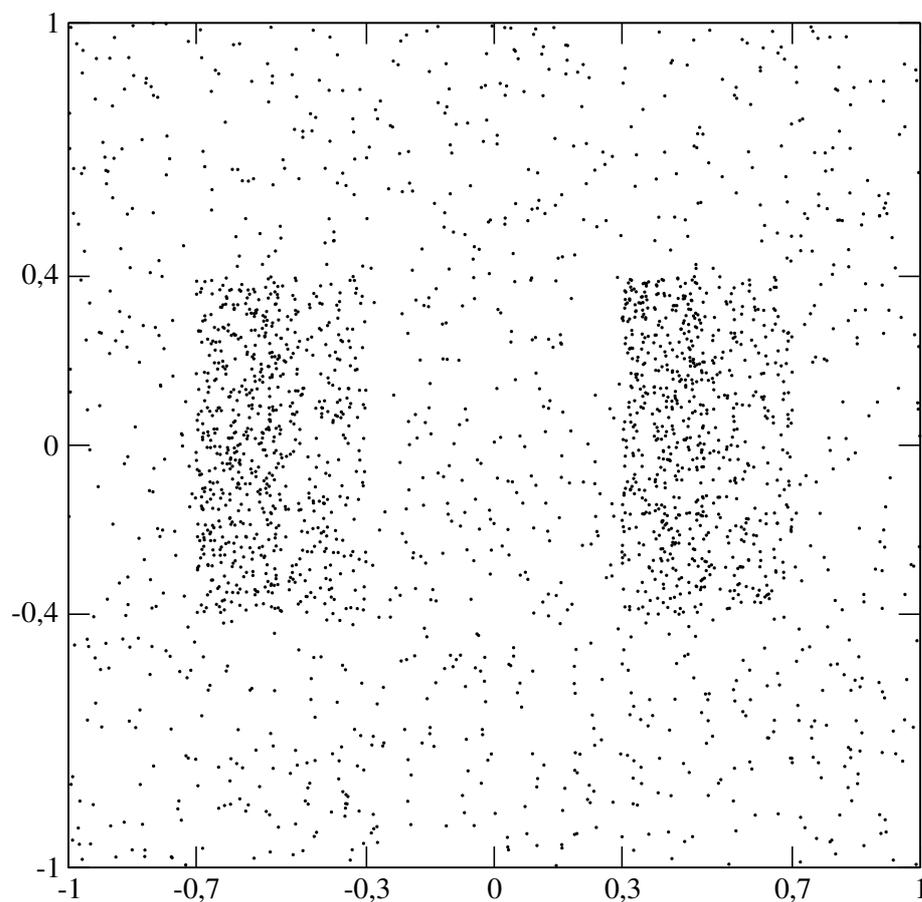
**Figura 3.6.** Conjunto de pontos consistindo de dois agrupamentos com 1672 e 768 pontos, possuindo a mesma densidade uniforme de pontos e com tamanho e formato similares ao mostrado na figura 3.1, imersos em um background de 148 pontos.



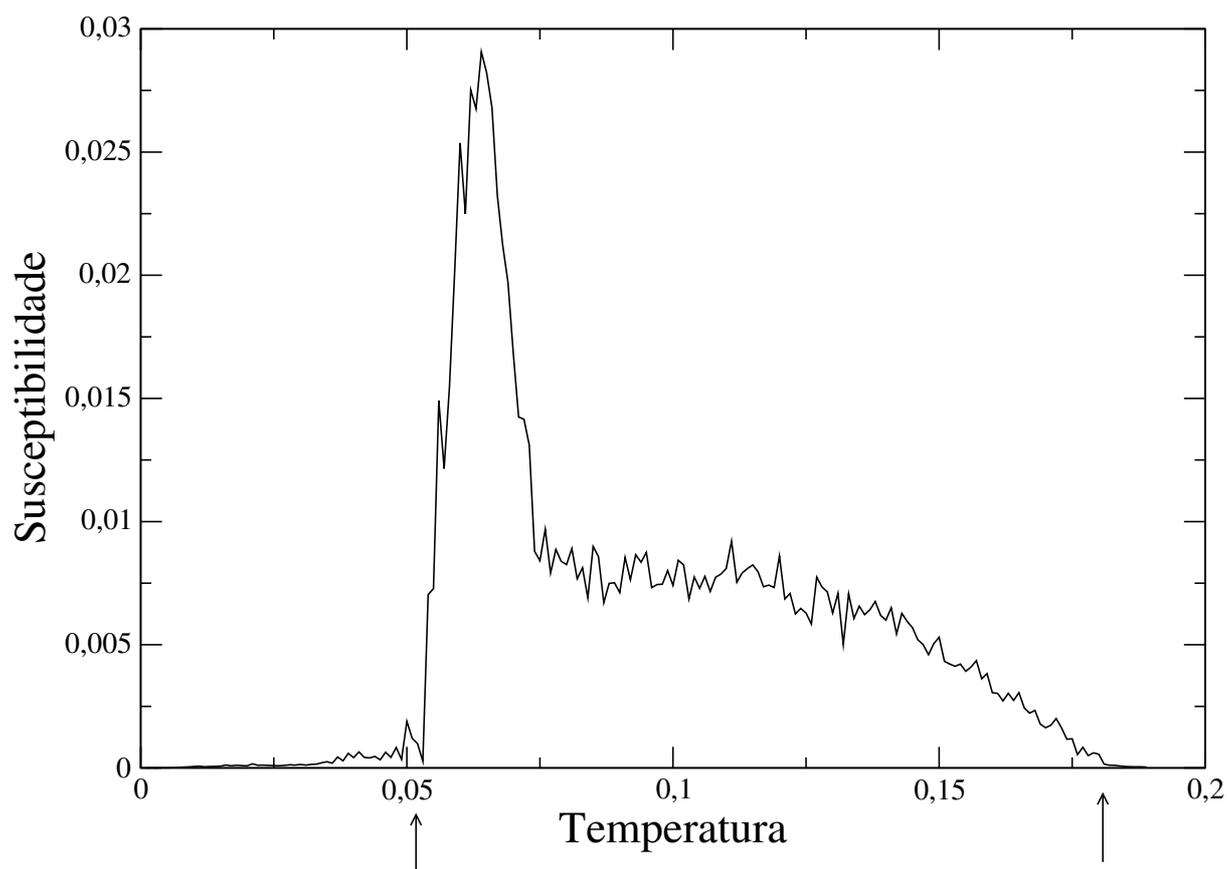
**Figura 3.7.** Evolução do tamanho dos 4 maiores clusters em função da informação entre pares de mapas referentes aos pontos da figura 3.6. A seta indica o valor da informação que utilizamos para fazer a classificação mostrada na figura 3.8.



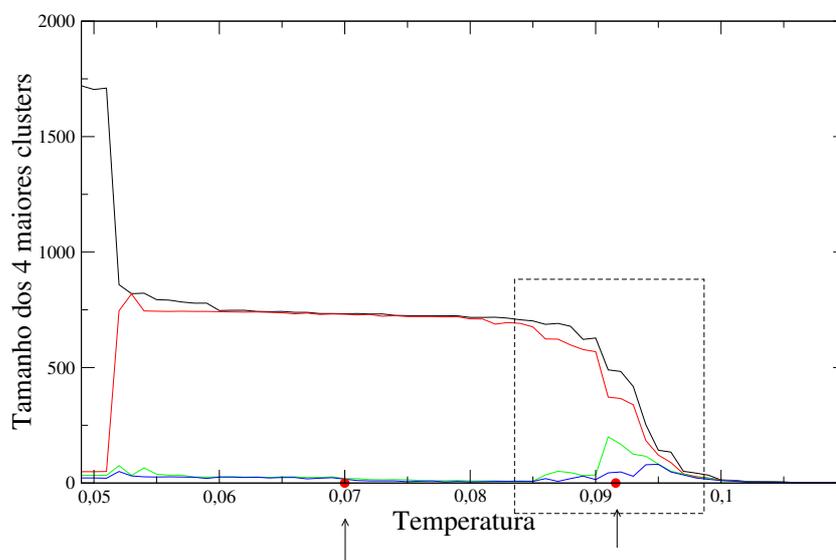
**Figura 3.8.** Classificação dos pontos da figura 3.6 em um limite de informação indicado pela seta no eixo horizontal do gráfico da figura 3.7.



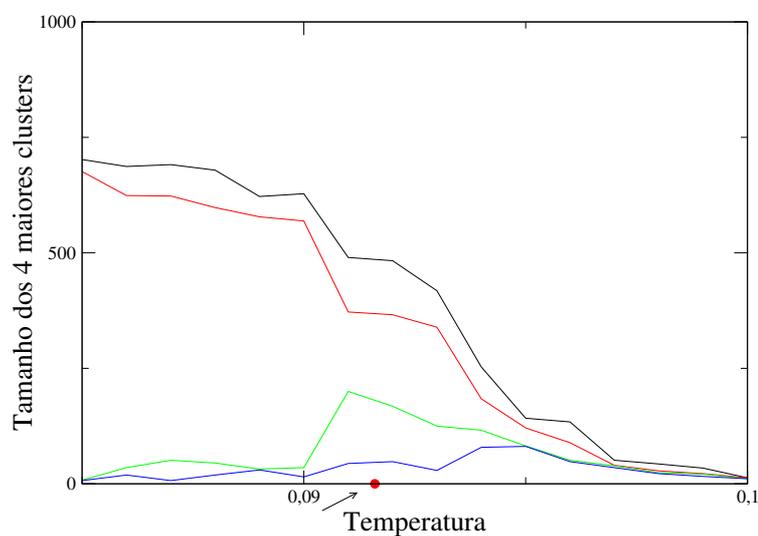
**Figura 3.9.** Conjunto artificial de dados consistindo de dois agrupamentos de densidades não uniformes, cada um com 714 pontos, nas regiões:  $-0.7 < x < -0.3$  e  $0.3 < x < 0.7$  com  $-0.4 < y < 0.4$ , divididos longitudinalmente em subregiões com densidades  $\sigma_1 = 10\sigma_b$  e  $\sigma_2 = 5\sigma_b$ , respectivamente à esquerda e à direita, onde  $\sigma_b$  é a densidade do background (1000 pontos) no qual estão imersos.



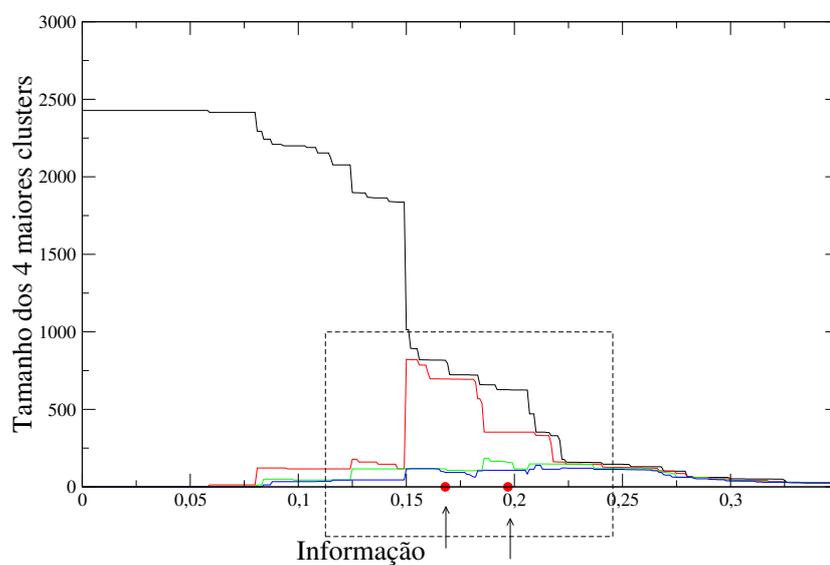
**Figura 3.10.** Gráfico da susceptibilidade referente ao conjunto de dados da figura 3.9 ao se utilizar o agrupamento super-paramagnético. O intervalo assinalado indica a região de ocorrência dos agrupamentos. Note o decaimento mais suave da susceptibilidade devido ao gradiente de densidade dos agrupamentos.



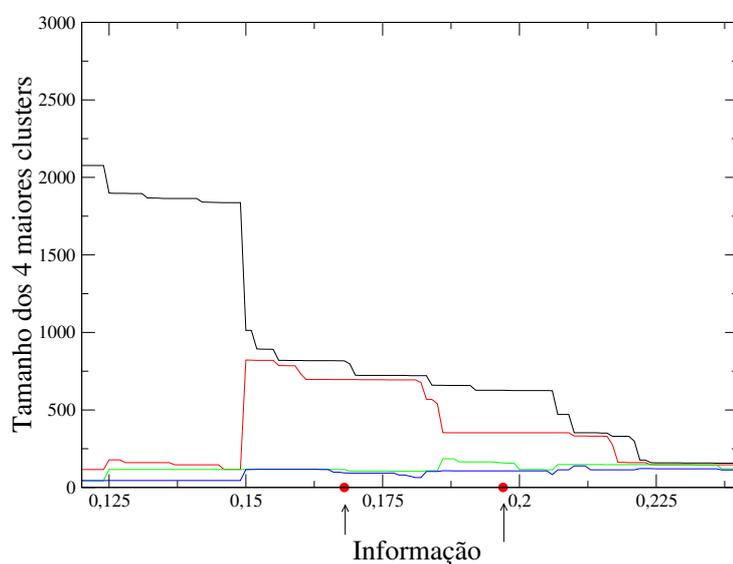
**Figura 3.11.** Gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.9.



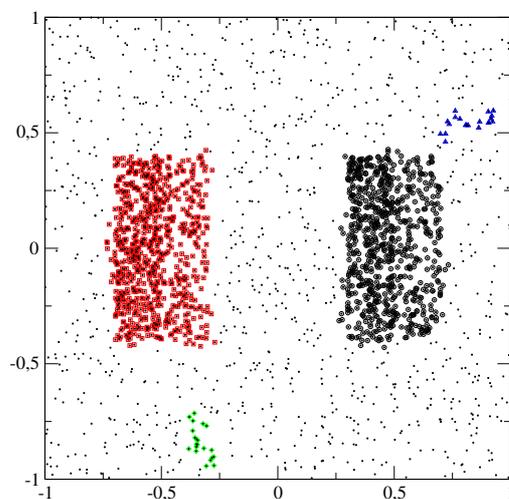
**Figura 3.12.** Ampliação do gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.9, na região delimitada da figura 3.11.



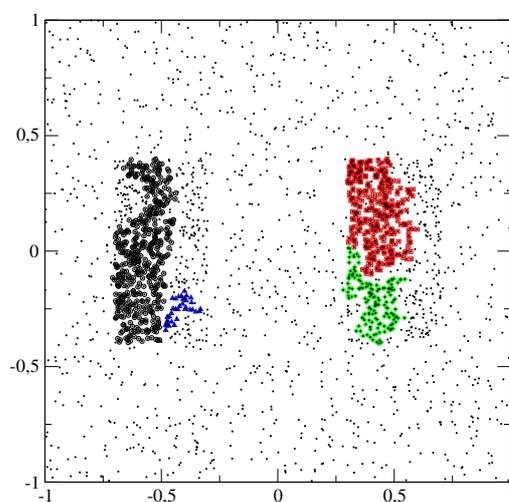
**Figura 3.13.** Gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.9.



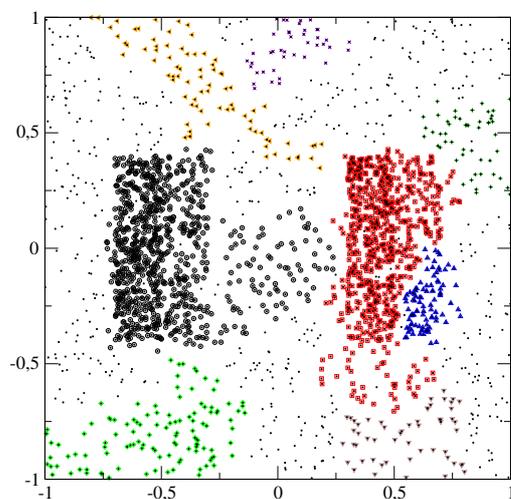
**Figura 3.14.** Ampliação do gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.9, na região delimitada da figura 3.13.



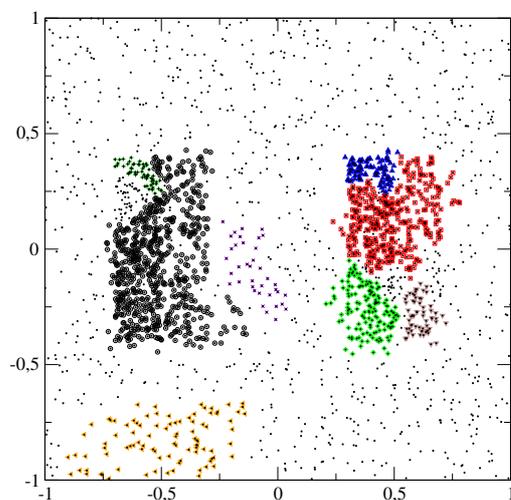
**Figura 3.15.** Classificação dos pontos (quatro maiores clusters) da figura 3.9 pelo agrupamento super-paramagnético em uma temperatura  $\mathcal{T} = 0.07$  indicada pela primeira seta no eixo horizontal do gráfico da figura 3.11.



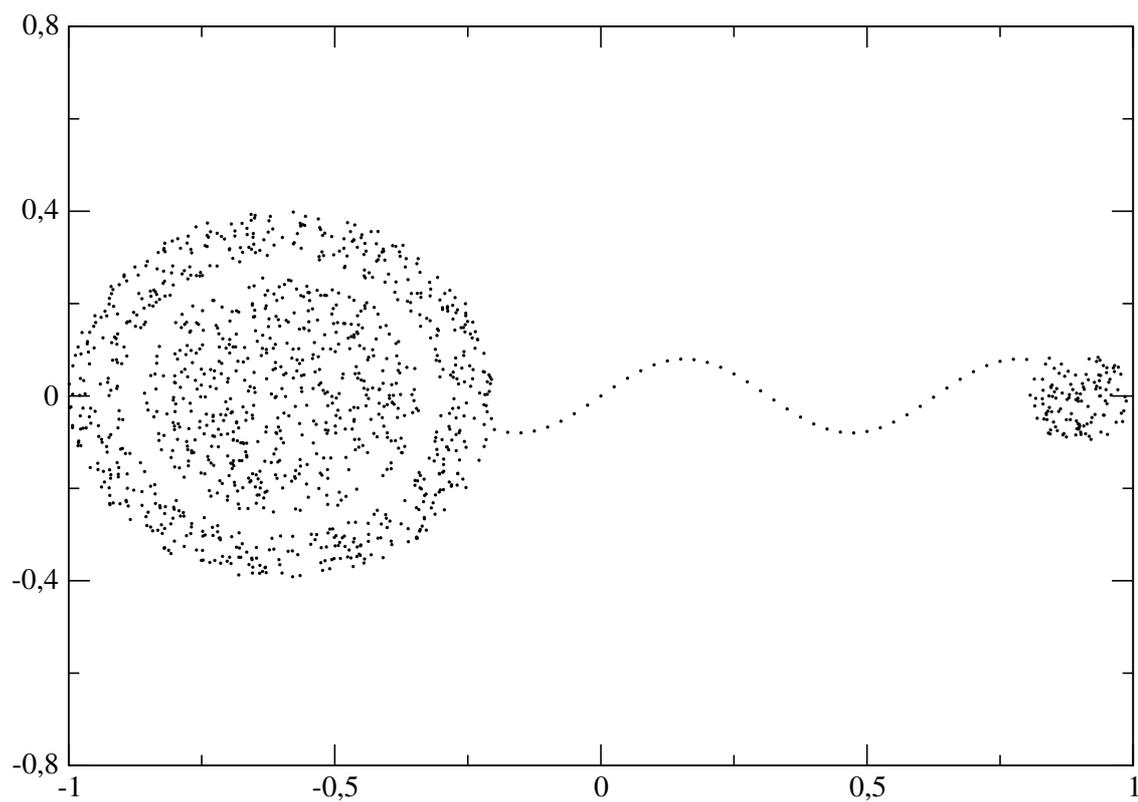
**Figura 3.16.** Classificação dos pontos (quatro maiores clusters) da figura 3.9 pelo agrupamento super-paramagnético em uma temperatura  $\mathcal{T} = 0.0916$  indicada pela segunda seta no eixo horizontal do gráfico da figura 3.11.



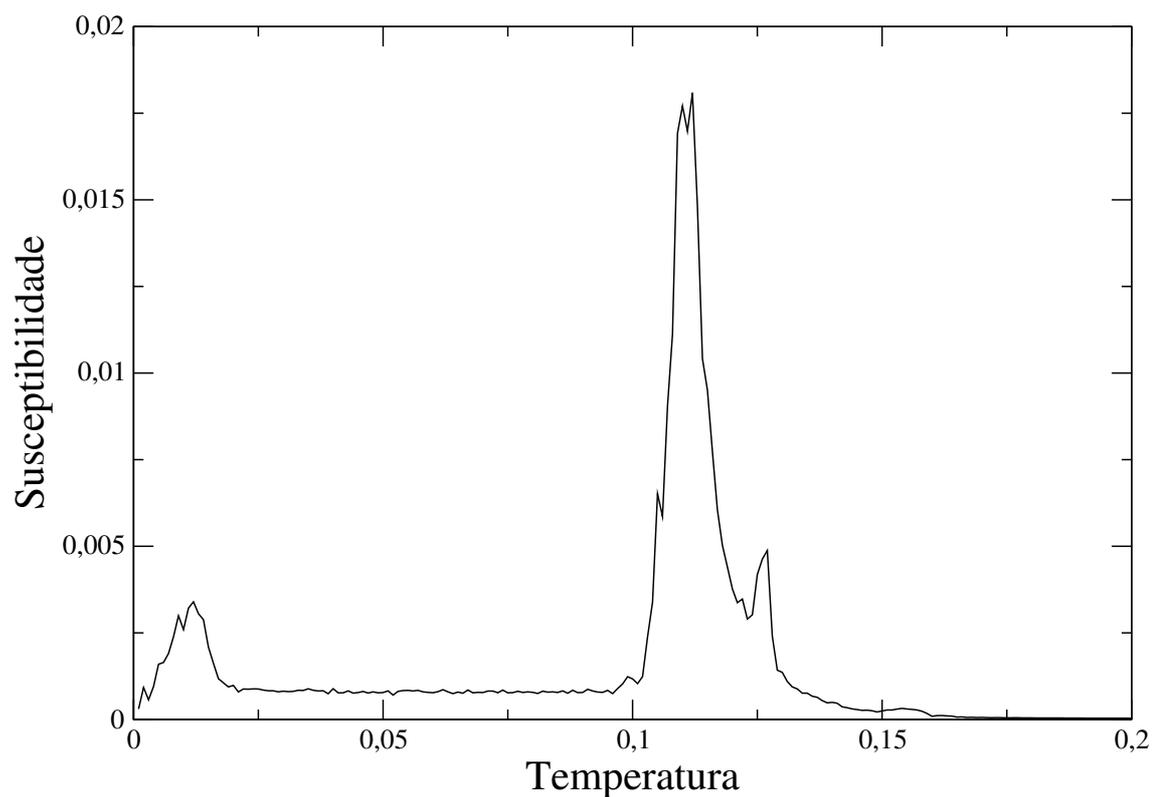
**Figura 3.17.** Classificação dos pontos (oito maiores clusters) da figura 3.9 pelo agrupamento através de mapas caóticos em um limite de informação  $I = 0.168$  indicado pela primeira seta no eixo horizontal do gráfico das figura 3.13.



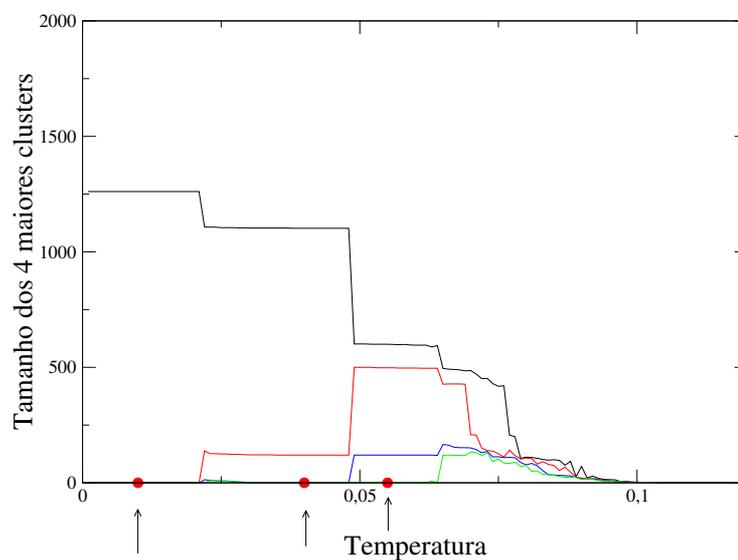
**Figura 3.18.** Classificação dos pontos (oito maiores clusters) da figura 3.9 pelo agrupamento através de mapas caóticos em um limite de informação  $I = 0.197$  indicado pela segunda seta no eixo horizontal do gráfico das figura 3.13.



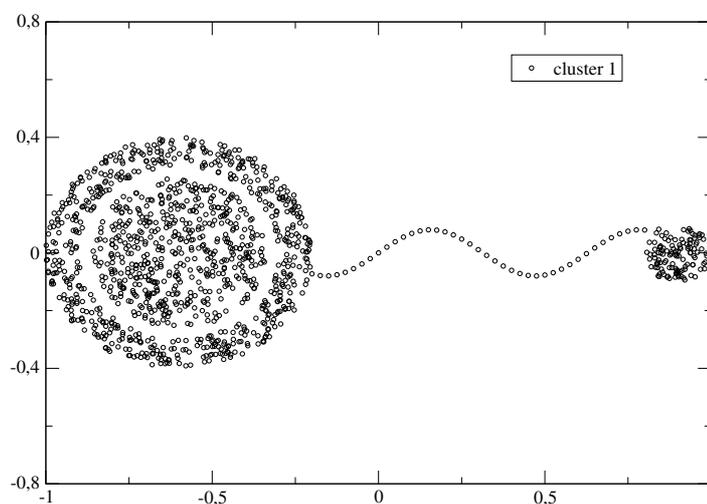
**Figura 3.19.** Conjunto artificial de dados consistindo de dois agrupamentos unidos por uma trajetória de pontos. Há 120 pontos no círculo menor à direita, 41 pontos na trajetória, 500 pontos no círculo maior à esquerda e 600 pontos no "aro" concêntrico ao círculo maior.



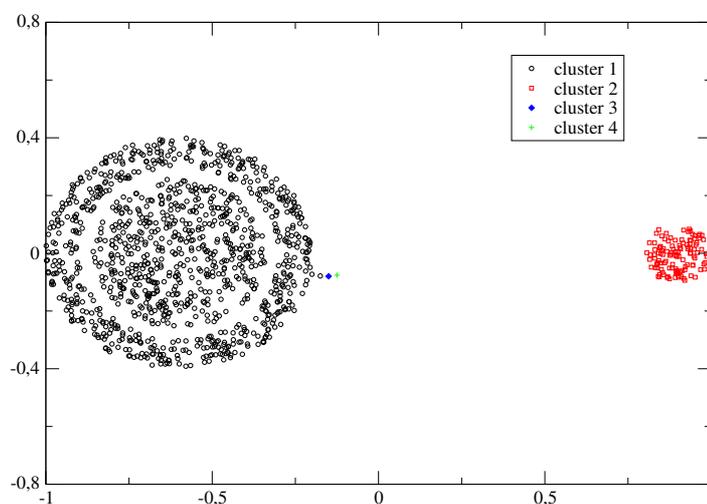
**Figura 3.20.** Gráfico da susceptibilidade referente ao conjunto de pontos da figura 3.19 ao se aplicar o agrupamento super-paramagnético. O menor pico à esquerda indica a fragmentação da trajetória de pontos e o início da fase super-paramagnética.



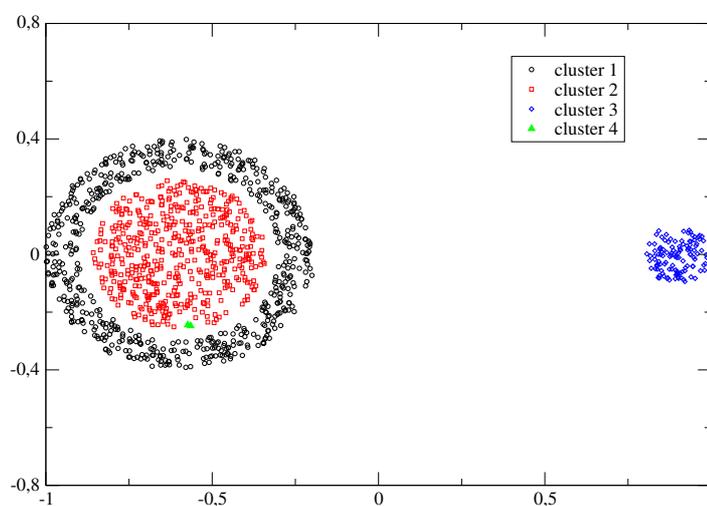
**Figura 3.21.** Gráfico da evolução do tamanho dos quatro maiores clusters em função da temperatura referente ao agrupamento super-paramagnético aplicado ao conjunto de pontos da figura 3.19.



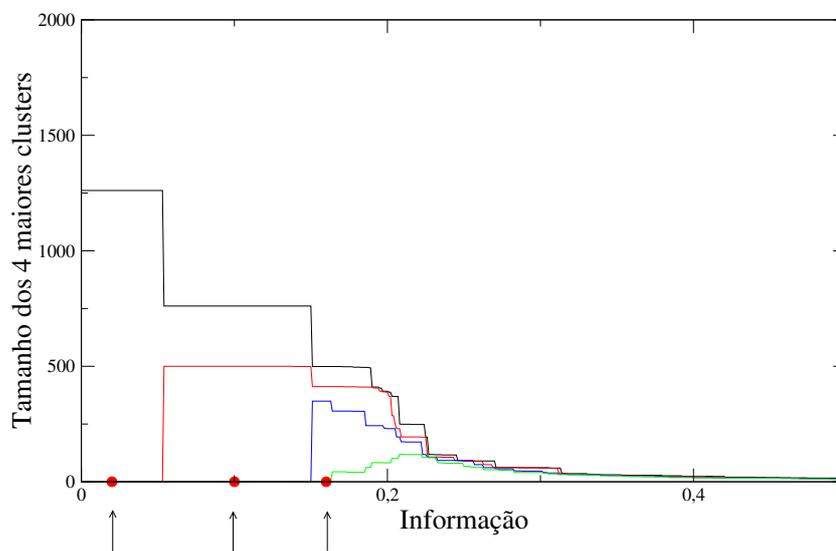
**Figura 3.22.** Classificação dos pontos da figura 3.19 pelo agrupamento super-paramagnético na temperatura  $T = 0.01$  indicada pela primeira seta no eixo horizontal do gráfico da figura 3.21. Apenas um único cluster é identificado.



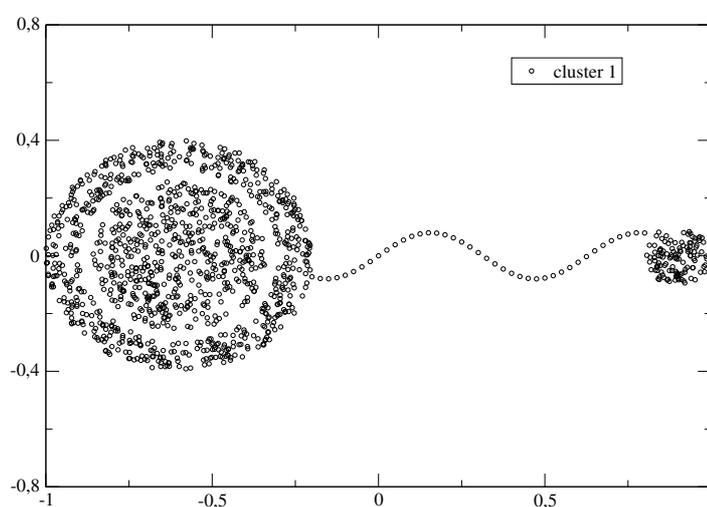
**Figura 3.23.** Classificação dos pontos da figura 3.19 pelo agrupamento super-paramagnético na temperatura  $\mathcal{T} = 0.04$  indicada pelo segunda seta no eixo horizontal do gráfico da figura 3.21.



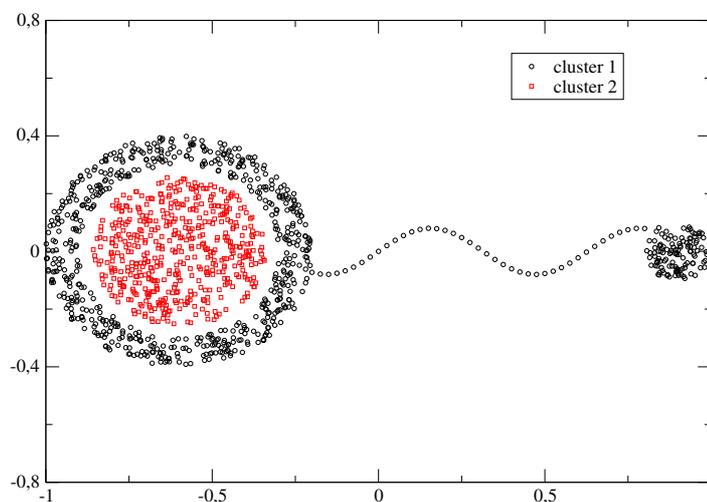
**Figura 3.24.** Classificação dos pontos da figura 3.19 pelo método super-paramagnético na temperatura  $\mathcal{T} = 0.055$  indicada pela terceira seta no eixo horizontal do gráfico da figura 3.21.



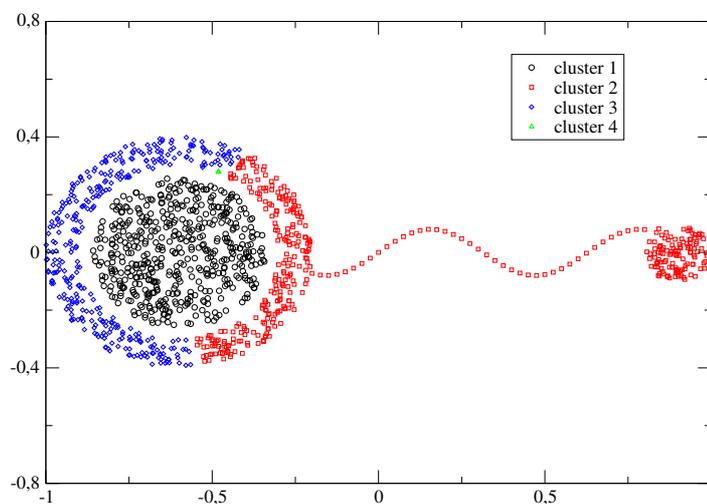
**Figura 3.25.** Gráfico da evolução do tamanho dos quatro maiores clusters em função da informação entre pares de mapas referente ao agrupamento através de mapas caóticos aplicado ao conjunto de pontos da figura 3.19.



**Figura 3.26.** Classificação dos pontos da figura 3.19 no limite de informação  $I = 0.02$ , indicado pelo primeira seta no eixo horizontal do gráfico da figura 3.25.



**Figura 3.27.** Classificação dos pontos da figura 3.19 no limite de informação  $I = 0.1$ , indicado pela segunda seta no eixo horizontal do gráfico da figura 3.25.



**Figura 3.28.** Classificação dos pontos da figura 3.19 no limite de informação  $I = 0.16$ , indicado pela terceira seta no eixo horizontal do gráfico da figura 3.25.

# CONCLUSÕES

Neste trabalho foram estudadas duas metodologias para agrupamento não supervisionado de dados ou padrões, dito *clustering*, um problema que consiste em agrupar dados ou padrões em classes *naturais* que emergem do próprio conjunto de dados. Trata-se de um problema cuja solução não é única e depende de vários fatores arbitrários, dentre eles a definição de *similaridade* entre os dados, requeridos para proceder ao agrupamento dos dados (*clusters*). As metodologias abordadas consistem basicamente em associar o conjunto de dados a sistemas físicos, e classificá-los segundo medidas de similaridade que são associadas a grandezas físicas.

O primeiro método considerado foi o do agrupamento super-paramagnético que associa a cada dado ou padrão um spin de um sistema de Potts de  $q$ -estados. Depois de definidas as interações entre spins, simulam-se através do método Monte Carlo, a evolução do sistema em contato com um reservatório térmico à temperatura  $T$ . Na fase super-paramagnética, que é identificada através da susceptibilidade magnética, emergem domínios de spins todos com o mesmo estado segundo a densidade local de pontos. A correspondente classificação dos dados é obtida a partir do seguinte critério de similaridade: dois dados pertencem ao mesmo *cluster* se a *correlação* entre seus spins correspondentes no sistema de Potts ultrapassar certo limiar, previamente fixado.

O segundo método que considerado, associa os dados a um conjunto de mapas caóticos unidimensionais acoplados. Quando o sistema de mapas caóticos evolui e atinge o regime estacionário emerge um comportamento coletivo, cujos padrões

de sincronia determinam a classificação buscada. Neste caso, o agrupamento dos dados é realizado a partir da sincronização entre os correspondentes mapas, isto é, dois dados pertencem ao mesmo *cluster* se a *informação mútua média* entre seus correspondentes mapas ultrapassar certo limiar pré-fixado.

Os dois métodos foram aplicados a conjuntos de dados artificiais constituídos de pontos em um espaço de atributos bidimensional, mais apropriados para a compreensão e visualização das potencialidades de cada método. O primeiro conjunto consistiu de dois agrupamentos densos, com formas previamente definidas, imersos em um fundo ou *background* e dispostos de uma forma não trivial. O agrupamento super-paramagnético classificou corretamente tais agrupamentos com erros da ordem de 3,5% e 2,0% para cada *cluster*, respectivamente, para certa temperatura. Por outro lado, o método de agrupamento por mapas caóticos aplicado a este conjunto de dados levou a um resultado com discrepância significativa com relação àquele obtido pelo agrupamento super-paramagnético, incorporando aos *clusters* estruturas unidimensionais ou “trilhas” presentes no *background*. O problema foi contornado com a diluição do *background* por um fator  $\sim 7$ , neste caso método dos mapas caóticos classificou corretamente os dois agrupamentos com erros de 1,1% e 0,8%, respectivamente.

O segundo conjunto de dados foi constituído de um *background* com dois agrupamentos retangulares, de mesmas dimensões e divididos longitudinalmente em duas regiões com densidades de 10 e 5 vezes maiores que a densidade do *background*. O objetivo deste conjunto de dados foi testar a potencialidade dos métodos com relação a classificação de padrões com regiões onde há gradientes de densidade. Os dois métodos distinguiram com sucesso os dois agrupamentos principais para uma certa “resolução”, mas as diferenças de densidades internas aos agrupamentos não foram claramente evidenciadas por nenhum dos métodos em face à baixa sensibilidade dos métodos às flutuações de densidades presentes aleatoriamente no conjunto

de dados. Para se explorar a potencialidade dos métodos, com relação à sensibilidade a flutuações de densidade, será necessário variar o fator  $a$  incluso nas interações entre os elementos do sistema físico e considerar padrões com grande número de pontos e escalas de densidade mais refinadas, o que demandará tempos de computação maiores.

No terceiro conjunto de dados foram definidos três agrupamentos de dados dispostos de forma conveniente sendo estes: um círculo envolto por um “aro” circular de pontos e um círculo menor. O *background* foi eliminado e substituído por uma trajetória de pontos conectando o “aro” ao círculo menor. O objetivo destas formas foi mostrar como são transmitidas as correlações entre os dois agrupamentos distanciados. Os resultados obtidos apresentaram diferenças essenciais entre os dois métodos por estarem fundamentados em diferentes sistemas físicos. Em uma classificação utilizando o agrupamento super-paramagnético a trajetória de pontos foi fragmentada separando os três agrupamentos, enquanto que o agrupamento através de mapas caóticos conservou esta trajetória para um mesmo nível de “resolução”, isto é, considerando o “aro”, o círculo menor e a trajetória com um único *cluster* (vide Figura 3.26). Todavia, ao se diminuir a resolução o método dos mapas acoplados separou o “aro” em duas partes, mantendo uma delas conectada ao círculo menor pela trajetória. Este resultado revela de forma clara a distinção entre os dois métodos.

O agrupamento super-paramagnético, por estar fundamentado em um sistema ferromagnético de spins, tem a característica de ser mais robusto com a relação aos ruídos de fundo ou *background*, justamente porque sistemas ferromagnéticos exibem fraca correlação em cadeias unidimensionais, como é o caso das “trilhas” que se formam no *background* ou na trajetória unindo os agrupamentos. Por esta razão, o método do agrupamento super-paramagnético é mais apropriado para identificar densidades com maior eficiência e precisão, e para eliminar o *background* e/ou ruídos

de fundo.

O agrupamento através de mapas caóticos acoplados possui maior sensibilidade quanto às estruturas unidimensionais presentes no *background* porque as conexões locais têm grande influência na transmissão da informação mútua média. Portanto, este método é apropriado para a classificação de *clusters* com formas complexas e ramificadas, quando a influência dos ruídos de fundo é relevante.

Com relação ao tempo computacional, o agrupamento através de mapas caóticos mostrou-se  $\sim 265$  vezes mais rápido que o agrupamento super-paramagnético, para os conjuntos de dados que consideramos, tendo em vista que no último é sempre necessário realizar simulações computacionais do sistema de Potts, para cada valor do intervalo de temperatura, até o sistema atingir o estado de equilíbrio termodinâmico. Todavia, é importante ressaltar que a grandeza física análoga à susceptibilidade, que é um parâmetro indicativo para as classificações no agrupamento superparamagnético, não existe no método de agrupamento através de mapas caóticos acoplados.

## APÊNDICE A

# SOLUÇÃO DO MODELO DE POTTS COM $Q$ -ESTADOS NA APROXIMAÇÃO DE CAMPO MÉDIO

Para uma maior compreensão do método de agrupamento super-paramagnético, apresentaremos a aproximação de campo médio de um modelo de Potts com  $q$ -estados que simula as regiões de alta densidade separadas por regiões de baixa densidade muitas vezes encontradas em problemas de agrupamento de dados, originalmente apresentada por Domany *et al.* [19]. O diagrama de fases do modelo foi obtido, bem como a forma da susceptibilidade que fornece um indicativo da fase super-paramagnética. Nesta seção, reproduziremos os resultados mais relevantes dando ênfase às transições Paramagnética - Super-paramagnética e Super-paramagnética - Ferromagnética.

A aproximação de campo médio, neste caso consiste em:

- Dividir  $N_c$  spins em  $M$  grupos.
- Spins dentro de um mesmo grupo interagem com uma interação do tipo  $\frac{M}{N_c} J_1$ .
- $N_b$  spins constituem o “background”; a interação entre dois destes spins é dada por  $\frac{1}{N_b} J_2$ .
- Todo spin que pertence a um dos  $M$  grupos interage com cada spin do *background* com uma interação  $\frac{1}{N_c} \sqrt{\frac{M}{N_b}} J_3$ .

A Hamiltoniana deste sistema de spins de Potts é dada então por:

$$\begin{aligned} \mathcal{H} = & -\frac{M}{N_c} J_1 \sum_{a=1}^M \sum_{i<j} \delta_{\sigma_i^a, \sigma_j^a} - \frac{1}{N_b} J_2 \sum_{i<j} \delta_{\sigma_i^b, \sigma_j^b} \\ & - \frac{1}{N_c} \sqrt{\frac{M}{N_b}} J_3 \sum_{a=1}^M \sum_{i,j} \delta_{\sigma_i^b, \sigma_j^a} \end{aligned} \quad (1.1)$$

onde  $\sigma_i^a$  denota o  $i$ -ésimo spin ( $i = 1, \dots, \frac{N_c}{M}$ ) do  $a$ -ésimo grupo ( $a = 1, \dots, M$ ),  $\sigma_i^b$  denota o  $i$ -ésimo spin ( $i = 1, \dots, N_b$ ) do background e  $\sigma_i^a, \sigma_i^b = 1, \dots, q$ .

Este modelo é relevante pois geralmente em problemas de agrupamento de dados tem-se uma alta dimensionalidade do espaço de padrões. Vemos deste modelo que quando os spins dentro de um grupo estão ordenados (no mesmo estado), um spin deste grupo está sujeito a um campo efetivo total da ordem de  $J_1$ . No background quando todos os spins estão ordenados, um spin pertencente ao mesmo está sujeito a um campo efetivo  $J_2$ . Como em problemas reais as interações decaem com a distância e o background é menos denso que os agrupamentos, consideramos que  $J_2 \ll J_1$ . O diagrama de fase é apresentado em função de  $J_3$ , para valores fixos de  $J_1$  e  $J_2$ .

Analisando o problema por meio da aproximação de campo médio consideramos que  $N_b, N_c \gg M \gg 1$ . O termo da energia referente às interações dentro do grupo pode ser obtido calculando-se a média de pares de spins do mesmo estado em um grupo  $a$ , que é dada por

$$\bar{N}_{\alpha\alpha}^a = \frac{1}{2} \left( \frac{N_c}{M} \right)^2 (m_{a\alpha})^2 \quad (1.2)$$

onde  $m_{a\alpha}$  é a fração de spins no  $a$ -ésimo grupo com valor  $\alpha$ .

O primeiro termo da Hamiltoniana torna-se então

$$-\frac{M}{N_c} J_1 \sum_{a=1}^M \sum_{\alpha=1}^q \left[ \frac{1}{2} \left( \frac{N_c}{M} \right)^2 (m_{a\alpha})^2 \right] = -\frac{N_c}{2M} J_1 \sum_{a,\alpha} (m_{a\alpha})^2 \quad (1.3)$$

Este cálculo pode ser estendido para os termos referentes a energia dentro do *background*, e interação *background-clusters* que resulta na energia total por spin

$$\begin{aligned} \frac{\mathcal{H}}{N_c} = E = & -\frac{J_1}{2M} \sum_{a,\alpha} m_{a\alpha}^2 - \frac{J_2 N_b}{2N_c} \sum_{\alpha} m_{b\alpha}^2 \\ & - \sqrt{\frac{N_b}{M} \frac{J_3}{N_c}} \sum_{a\alpha} m_{b\alpha} m_{a\alpha} \end{aligned} \quad (1.4)$$

A entropia pode ser calculada do número de configurações possíveis  $W$ , para um valor fixo da energia (ensemble canônico),

$$W = \frac{\left[ \left( \frac{N_c}{M} \right)! \right]^M}{\prod_{a,\alpha} \left( \frac{N_c}{M} m_{a\alpha} \right)!} \times \frac{(N_b)!}{\prod_{\alpha} (N_b m_{b\alpha})!} \quad (1.5)$$

$$\begin{aligned} \log W = & M \log \left[ \left( \frac{N_c}{M} \right)! \right] - \sum_{\alpha} \log \left[ \left( \frac{N_c}{M} m_{a\alpha} \right)! \right] + \\ & - \sum_{\alpha} \log [(N_b m_{b\alpha})!] + \log [N_b!] \end{aligned} \quad (1.6)$$

Utilizando a aproximação de Stirling  $\log(N!) \simeq N \log(N) - N$ , para  $N \gg 1$ , podemos fazer a seguinte aproximação,

$$\log W = -\frac{N_c}{M} \sum_{a,\alpha} m_{a\alpha} \log(m_{a\alpha}) - N_b \sum_{\alpha} m_{b\alpha} \log(m_{b\alpha}) \quad (1.7)$$

A entropia por spin é então dada por:

$$S = -\frac{1}{M} \sum_{a,\alpha} m_{a\alpha} \log(m_{a\alpha}) - \frac{N_b}{M} \sum_{\alpha} m_{b\alpha} \log(m_{b\alpha}) \quad (1.8)$$

## A.1 PARÂMETROS DE ORDEM E ENERGIA LIVRE

O estado do sistema pode ser representado por três parâmetros de ordem

$$\left. \begin{array}{l} \text{clusters} \quad \longrightarrow \quad s \\ \text{background} \quad \longrightarrow \quad t \\ \text{super-spins} \quad \longrightarrow \quad \hat{s} \end{array} \right\} \quad 0 \leq s, t, \hat{s} \leq 1 \quad (1.9)$$

onde o termo super-spins se refere ao estado dominante em cada cluster, sendo que todo este cluster pode ser representado por um único super-spin neste estado. Seja nos clusters, no background ou no grupo de super-spins é assumido que uma fração  $m_\alpha = m \geq \frac{1}{q}$  dos spins estão em um estado  $\alpha$  e os outros spins estão em um dos  $q - 1$  estados restantes com igual frequência. A variável,

$$P = \frac{\left(m - \frac{1}{q}\right)}{\left(1 - \frac{1}{q}\right)} \quad (1.10)$$

mede a ordem sendo esta identificada como  $s$ ,  $t$  ou  $\hat{s}$ , como mostra a relação 1.9 acima. Resolvendo para as frações  $m_\alpha$  temos

$$m_\alpha = \frac{(q-1)P + 1}{q} \quad ; \quad m_{\alpha \neq \gamma} = \frac{1-P}{q} \quad (1.11)$$

Para os clusters esta expressão pode ser posta em uma forma mais precisa

$$m_{\alpha\gamma} = \frac{1-s}{q} + s\delta_{\hat{\sigma}_a, \gamma} \quad (1.12)$$

onde  $\hat{\sigma}_a$  é o valor do super-spin. A fração dos super-spins com valor  $\gamma$  é denotada por  $m_\gamma = \frac{1}{M} \sum_a \delta_{\hat{\sigma}_a, \gamma}$ , onde  $\sum_a$  significa a soma em todos os grupos.

Em função dos parâmetros de ordem e utilizando-se as equações 1.11 e 1.12,

a energia por spin é dada por,

$$E = -\frac{J_1}{M} \left[ \frac{q-1}{q} s^2 + \frac{1}{q} \right] - \frac{J_2 N_b}{2N_c} \left[ \frac{q-1}{q} t^2 + \frac{1}{q} \right] + \\ -\sqrt{N_b M} \frac{J_3}{N_c} \left[ \frac{1-s}{q} + s \sum_{\gamma=1}^q m_{b\gamma} \hat{m}_\gamma \right] \quad (1.13)$$

Considerando que o estado dominante do background é o mesmo dos super-spins, temos para o último termo de E

$$-\sqrt{N_b M} \frac{J_3}{N_c} \left[ \frac{(q-1) st \hat{s}}{q} + \frac{1}{q} \right] \quad (1.14)$$

ou seja

$$E = -\frac{J_1}{M} \left[ \frac{q-1}{q} s^2 + \frac{1}{q} \right] - \frac{J_2 N_b}{2N_c} \left[ \frac{q-1}{q} t^2 + \frac{1}{q} \right] + \\ -\sqrt{N_b M} \frac{J_3}{N_c} \left[ \frac{(q-1) st \hat{s}}{q} + \frac{1}{q} \right] \quad (1.15)$$

A entropia pode ser calculada em função dos parâmetros de ordem como segue. A entropia deve ter a forma  $S \equiv S(s, t, \hat{s})$ , então, para um valor fixo de  $s$ ,  $t$  e  $\hat{s}$ , temos que o número de configurações possíveis relativas aos clusters, ao background e aos super-spins são respectivamente:

$$\frac{M! q}{\prod_{\gamma} (M m_{\gamma})} \quad (1.16)$$

$$\frac{\left[ \left( \frac{N_c}{M} \right)! \right]^M}{\left[ \prod_{\alpha} \left( \frac{N_c}{M} m_{1\alpha} \right)! \right]^M} \quad (1.17)$$

$$\frac{N_b! q}{\prod_{\alpha} (N_b m_{b\alpha})!} \quad (1.18)$$

A entropia é então, após algumas manipulações algébricas, dada a menos de

uma constante, por

$$S = -M \sum_{\gamma} \hat{m}_{\gamma} \log \hat{m}_{\gamma} - N_c \sum_{\alpha} m_{1\alpha} \log m_{1\alpha} + \\ - N_b \sum_{\alpha} m_{b\alpha} \log m_{b\alpha} \quad (1.19)$$

Em função dos parâmetros de ordem a entropia por spin é,

$$\frac{S}{N_c} = \mathcal{S}(s) + \frac{N_b}{N_c} \mathcal{S}(t) + \frac{M}{N_c} \mathcal{S}(\hat{s}) \quad (1.20)$$

onde,

$$\mathcal{S}(s) = -\frac{1+(q-1)s}{q} \log [1+(q-1)s] - \frac{(q-1)(1-s)}{q} \log [1-s]. \quad (1.21)$$

Utilizando as equações 1.15 e 1.20, podemos escrever a energia livre por spin como,

$$\beta f = -\beta \frac{J_1}{M} \left[ \frac{q-1}{q} s^2 + \frac{1}{q} \right] - \beta \frac{J_2 N_b}{2N_c} \left[ \frac{q-1}{q} t^2 + \frac{1}{q} \right] + \\ -\beta \sqrt{N_b M} \frac{J_3}{N_c} \left[ \frac{(q-1)st\hat{s}}{q} + \frac{1}{q} \right] \\ -\mathcal{S}(s) - \frac{N_b}{N_c} \mathcal{S}(t) - \frac{M}{N_c} \mathcal{S}(\hat{s}) \quad (1.22)$$

Onde  $\beta = \frac{1}{T}$  é o inverso da temperatura (em unidades de  $k_B = 1$ ).

## A.2 DIAGRAMA DE FASE

Os estados do sistema podem encontrados ao se minimizar a energia livre 1.22 com relação a  $s$ ,  $t$  e  $\hat{s}$ , que resulta em

$$\begin{aligned} \frac{q}{q-1} \frac{\partial \beta f}{\partial s} &= -\beta J_1 s - \beta \frac{\sqrt{MN_b}}{N_c} J_3 t \hat{s} \\ &+ \log [1 + (q-1)s] - \log [1 - s] = 0, \end{aligned} \quad (1.23)$$

$$\begin{aligned} \frac{q}{q-1} \frac{\partial \beta f}{\partial t} &= -\beta J_2 \frac{N_b}{N_c} t - \beta \frac{\sqrt{MN_b}}{N_c} J_3 s \hat{s} \\ &\frac{N_b}{N_c} \{ \log [1 + (q-1)t] - \log [1 - t] \} = 0, \end{aligned} \quad (1.24)$$

$$\begin{aligned} \frac{q}{q-1} \frac{\partial \beta f}{\partial \hat{s}} &= -\beta \frac{\sqrt{MN_b}}{N_c} J_3 s t \\ &-\frac{M}{N_c} \{ \log [1 + (q-1)\hat{s}] - \log [1 - \hat{s}] \} = 0, \end{aligned} \quad (1.25)$$

Analisando estas três equações constatamos que existem quatro fases possíveis,

- paramagnética ( $s = 0, t = 0, \hat{s} = 0$ )
- super-paramagnética ( $s \neq 0, t = 0, \hat{s} = 0$ )
- ferromagnética ( $s \neq 0, t \neq 0, \hat{s} \neq 0$ )

A quarta fase possível, onde ( $s = 0, t \neq 0, \hat{s} = 0$ ), não é de interesse pois ocorre somente para  $J_2 > J_1$ . Em algumas regiões do espaço dos acoplamentos ocorre uma transição adicional dentro da fase ferromagnética separando duas sub-fases, uma para valores pequenos de  $t$  e outra para valores grandes de  $t$ .

### A.2.1 Transição Paramagnética - Super-Paramagnética

Em altas temperaturas a energia livre é minimizada pela solução paramagnética das equações 1.23, 1.24 e 1.25,  $s = t = \hat{s} = 0$ . A transição para a fase super-

paramagnética é caracterizada pela igualdade da energia livre na transição. Utilizando 1.22 encontramos,

$$f(s, 0, 0) = f(0, 0, 0) \quad (1.26)$$

$$-\beta \frac{J_1}{2q} = -\beta \frac{J_1}{2} \left( \frac{q-1}{q} s^2 + \frac{1}{q} \right) - \mathcal{S}(s) \quad (1.27)$$

As equações 1.24, 1.25 são satisfeitas, mas  $s$  precisa ser uma solução de 1.23,

$$-\beta J_1 s + \log \left[ \frac{1 + (q-1)s}{1-s} \right] = 0 \quad (1.28)$$

As equações 1.27 e 1.28 são idênticas àquelas para o modelo de Potts homogêneo [14] cujas soluções são dadas por,

$$-\beta_{ps} J_1 = \frac{2(q-1)}{q-2} \log(q-1) \quad (1.29)$$

$$s_{ps} = \frac{q-2}{q-1} \quad (1.30)$$

O significado físico é que existe ordem nos clusters, mas não existe ordem entre os super-spins ( $\hat{s} = 0$ ), nem ordem no background ( $t = 0$ ).

### A.2.2 Transição Super-Paramagnética - Ferromagnética

Estando o sistema na fase super-paramagnética e diminuindo-se a temperatura uma transição ocorre à  $\beta_{fs}$  para a fase ferromagnética com  $s, t, \hat{s} \neq 0$ . Na transição o estado do sistema muda abruptamente de  $(s^-, 0, 0)$  para  $(s, t, \hat{s})$ . Os valores de  $s^-, s, t, \hat{s}$  e  $\beta_{fs}$  são encontrados a partir de cinco equações, sendo elas, a equação 1.23 para a fase super-paramagnética, a igualdade da energia livre na

transição e as equações 1.23,1.24 e 1.25 para a fase ferromagnética,

$$-\beta_{fs}J_1s^- + \log [1 + (q-1)s^-] - \log [1 - s^-] = 0 \quad (1.31)$$

$$-\beta_{fs}f(s^-, 0, 0) = -\beta_{fs}f(s, t, \hat{s})$$

$$\begin{aligned} & -\frac{\beta_{fs}J_1}{2} \left( \frac{q-1}{q}(s^-)^2 + \frac{1}{q} \right) - \beta_{fs} \frac{J_2N_b}{2qN_c} - \beta_{fs}J_3 \frac{\sqrt{MN_b}}{qN_c} - \mathcal{S}(s^-) = \\ = & -\frac{\beta_{fs}J_1}{2} \left( \frac{q-1}{q}(s^2)^2 + \frac{1}{q} \right) - \beta_{fs} \frac{J_2N_b}{2N_c} \left( \frac{q-1}{q}t^2 + \frac{1}{q} \right) - \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c} \left( \frac{q-1}{q}st\hat{s} + \frac{1}{q} \right) \\ & -S(s) - \frac{N_b}{N_c}S(t) - \frac{M}{N_c}S(\hat{s}) \end{aligned} \quad (1.32)$$

$$-\beta_{fs}J_1s - \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}t\hat{s} + \log [1 + (q-1)s] - \log [1 - s] = 0 \quad (1.33)$$

$$-\beta_{fs}J_2 \frac{N_b}{N_c}t - \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}s\hat{s} + \frac{N_b}{N_c} \{ \log [1 + (q-1)t] - \log [1 - t] \} = 0 \quad (1.34)$$

$$-\beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}st + \frac{M}{N_c} \{ \log [1 + (q-1)\hat{s}] - \log [1 - \hat{s}] \} = 0 \quad (1.35)$$

Considerando-se que  $1 - s^- \ll 1$  e  $\epsilon \equiv s - s^- \ll 1$  e subtraindo 1.33 de 1.31 encontramos,

$$\begin{aligned} & -\beta_{fs}J_1\epsilon + \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}ts + \log \left\{ \frac{[1+(q-1)s^-][1-s^-]}{[1+(q-1)s][1-s]} \right\} = 0 \\ & -\beta_{fs}J_1\epsilon + \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}ts + \log \left[ 1 - \frac{\epsilon qs^-}{1+s^-(q-2)-(s^-)^2(q-1)} \right] = \\ & -\beta_{fs}J_1\epsilon + \beta_{fs}J_3 \frac{\sqrt{MN_b}}{N_c}ts - \frac{\epsilon q}{1+s^-(q-2)-(s^-)^2(q-1)} \end{aligned} \quad (1.36)$$

Escrevendo 1.32 como função de  $\epsilon$  temos,

$$\begin{aligned} & \frac{-\beta J_1 (q-1)}{2} \frac{(q-1)}{q} (\epsilon^2 + 2\epsilon s^-) + \frac{N_b}{N_c} \left\{ \frac{\beta J_2}{2q} (q-1)t^2 + \mathcal{S}(t) \right\} + \\ & + \beta J_3 \frac{\sqrt{MN_b}}{N_c q} (q-1)st\hat{s} - \mathcal{S}(s^-) + \mathcal{S}(s) + \frac{M}{N_c} \mathcal{S}(\hat{s}) = 0 \end{aligned} \quad (1.37)$$

Expressando  $-\mathcal{S}(s^-) + \mathcal{S}(s)$  como sendo,

$$\begin{aligned} & \frac{q-1}{q} \epsilon \log \left\{ \frac{[1+(q-1)s^-](1-s)}{[1+(q-1)s](1-s^-)} \right\} - \frac{2+(q-1)(2s-\epsilon)}{2q} \log [1 + (q-1)s] + \\ & - \frac{(q-1)[2(1-s)+\epsilon]}{2q} \log (1-s) + \frac{2+(q-1)(2s^--\epsilon)}{2q} \log [1 + (q-1)s^-] + \\ & + \frac{(q-1)[2(1-s^-)+\epsilon]}{2q} \log (1-s^-) \end{aligned} \quad (1.38)$$

Aproximando o primeiro termo da expressão acima até segunda ordem em  $\epsilon$  e retornando à equação 1.32 encontramos,

$$\begin{aligned} & \beta J_1 \epsilon^2 \frac{(q-1)}{2q} - \frac{q-1}{2} \frac{\epsilon^2}{1+(q-2)s^--(q-1)(s^-)^2} + \beta J_1 \epsilon s^- \frac{(q-1)}{q} + \\ & + \frac{N_b}{N_c} \left\{ \left( \frac{\beta J_2}{2q} (q-1)t^2 + \mathcal{S}(t) \right) \right\} + \beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1)(s^- - \epsilon)t\hat{s} - \frac{M}{N_c} \mathcal{S}(\hat{s}) + \\ & + \frac{(q-1)[2(1-s^-)+\epsilon]}{2q} \log (1-s^-) + \frac{2+(q-1)(2s^--\epsilon)}{2q} \log [1 + (q-1)s^-] + \\ & - \frac{2+(q-1)(2s-\epsilon)}{2q} \log [1 + (q-1)s] - \frac{(q-1)[2(1-s)+\epsilon]}{2q} \log (1-s) = 0. \end{aligned} \quad (1.39)$$

Utilizando 1.36 com relação aos dois primeiros termos da expressão acima e 1.31 com relação ao terceiro termo, temos

$$\begin{aligned} & -\beta J_3 \frac{(q-1)}{2q} \frac{\sqrt{MN_b}}{N_c} t\hat{s}\epsilon + \frac{N_b}{N_c} \left[ \frac{\beta J_2}{2q} (q-1)t^2 + \mathcal{S}(t) \right] + \\ & + \beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1)t\hat{s}(s^- + \epsilon) + \frac{M}{N_c} \mathcal{S}(\hat{s}) + \left[ \frac{(q-1)[2(1-s^-)+\epsilon]}{2q} \log (1-s^-) + \right. \\ & - \frac{(q-1)}{q} \epsilon \log (1-s^-) + \frac{2+(q-1)(2s^--\epsilon)}{2q} \log [1 + (q-1)s^-] + \frac{(q-1)}{q} \epsilon \log [1 + (q-1)s^-] + \\ & \left. - \frac{(q-1)[2(1-s^-)+\epsilon]}{2q} \log (1-s) - \frac{2+(q-1)(2s-\epsilon)}{2q} \log [1 + (q-1)s] \right] = 0. \end{aligned} \quad (1.40)$$

Simplificando o último termo entre colchetes,

$$\begin{aligned} & \frac{(q-1)[2(1-s^-)-\epsilon]}{2q} \log (1-s^-) + \frac{2+(q-1)(2s^--\epsilon)}{2q} \log [1 + (q-1)s^-] + \\ & - \frac{2+(q-1)(2s-\epsilon)}{2q} \log [1 + (q-1)s] - \frac{(q-1)[2(1-s)+\epsilon]}{2q} \log (1-s) \end{aligned} \quad (1.41)$$

e fazendo a aproximação  $\log [1 + (q - 1) s^-] \approx \log [1 + (q - 1) s]$  e  $\log (1 - s^-) \approx \log (1 - s)$  pois  $\epsilon = s - s^- \ll 1$  e a variação da função  $\log$  é muito pequena, temos que a expressão acima anula-se, e 1.40 torna-se então,

$$\begin{aligned} -J_3 \beta \frac{\sqrt{MN_b}}{2qN_c} (q - 1) t \hat{s} \epsilon + \frac{N_b}{N_c} \left[ -\frac{\beta J_2}{2q} (q - 1) t - S(t) \right] + \\ -\beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q - 1) t \hat{s} (s - \epsilon) - \frac{M}{N_c} S(\hat{s}) = 0 \end{aligned} \quad (1.42)$$

Utilizando agora  $\epsilon \ll t \ll 1$  em primeira ordem em  $t$  na equação 1.34, encontramos,

$$\begin{aligned} -\beta J_2 \frac{N_b}{N_c} t - \beta J_3 \frac{\sqrt{MN_b}}{N_c} s \hat{s} + \frac{N_b}{N_c} [(q - 1) t - (-t)] \approx 0 \\ t \approx \sqrt{\frac{M}{N_c} \frac{\beta J_3 s \hat{s}}{(q - \beta J_2)}} \end{aligned} \quad (1.43)$$

Substituindo 1.43 em 1.35, resulta,

$$-\frac{M}{N_c} \frac{(\beta J_3 s^-)^2}{(q - \beta J_2)} \hat{s} + \frac{M}{N_c} \{ \log [1 + (q - 1) \hat{s}] - \log (1 - \hat{s}) \} = 0 \quad (1.44)$$

Expandindo  $\mathcal{S}(t)$  até segunda ordem em  $t$ , desprezando termos proporcionais a  $\beta J_3 \frac{\sqrt{MN_b}}{N_c} t \hat{s} \epsilon$  e usando 1.43, obtemos de 1.42,

$$\begin{aligned} \frac{N_b}{2qN_c} (q - 1) [q - \beta J_2] - \beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q - 1) t s s^- - \frac{M}{N_c} \mathcal{S}(\hat{s}) = 0 \\ -\frac{M}{2qN_c} \frac{(q-1)(\beta J_3 s^- \hat{s})^2}{(q - \beta J_2)} - \frac{M}{N_c} \mathcal{S}(\hat{s}) = 0 \end{aligned} \quad (1.45)$$

As equações 1.44 e 1.45 conduzem à soluções semelhantes às 1.27 e 1.28, ou seja ,

$$\frac{(\beta_{fs} J_3 s^-)^2}{(q - \beta_{fs} J_2)} = \frac{2(q-1)}{q-2} \log (q - 1) \quad (1.46)$$

$$\hat{s} = \frac{q-2}{q-1} \quad (1.47)$$

Assumindo agora  $\beta_{fs}J_2 \ll 1$ , 1.46 simplifica-se como,

$$\beta_{fs}J_3s^- = \sqrt{\frac{2q(q-1)}{q-2} \log(q-1)} \quad (1.48)$$

Aplicando 1.48 e 1.46 em 1.43 temos,

$$t = \sqrt{\frac{M}{N_b}} \sqrt{\frac{2(q-1) \log(q-1)}{q(q-2)}} \quad (1.49)$$

Para encontrarmos  $s^-$  e  $\epsilon$  assumamos que  $J_3 \ll 1$ , e que através de 1.48 implica em  $\beta_{fs} \gg 1$ .

Escrevendo  $s^-$  como sendo  $s^- = 1 - \delta$  onde  $\delta \ll 1$ , temos dos dois termos líderes de 1.31 e utilizando 1.48,

$$\begin{aligned} -\beta_{fs}J_1s^- - \log \delta &= 0 \\ -\frac{J_1}{J_3}c &= \log \delta \\ \delta &= \exp\left(-c\frac{J_1}{J_2}\right) \end{aligned} \quad (1.50)$$

onde

$$c = \sqrt{\frac{2q(q-1)}{q-2} \log(q-1)} \quad (1.51)$$

Calculando  $\epsilon$  a partir de 1.36 obtem-se:

$$\frac{1}{\delta} - \beta J_1 = \beta J_3 \frac{\sqrt{MN_b} t \hat{s}}{N_c \epsilon} \quad (1.52)$$

e usando 1.48 e 1.47, temos:

$$\epsilon \approx \beta J_3 \frac{\sqrt{MN_b}}{N_c} t s \delta = \frac{M}{N_c} \frac{2(q-2)}{q-1} \log(q-1) e^{-c\frac{J_1}{J_3}} \quad (1.53)$$

Em resumo, os valores finais dos parâmetros de ordem são,

$$\hat{s}_{fs} = \frac{q-2}{q-1} \quad (1.54)$$

$$t_{fs} = \sqrt{\frac{2(q-2)}{q(q-1)} \log(q-1)} \sqrt{\frac{M}{N_b}} \quad (1.55)$$

$$s^- = 1 - e^{-c \frac{J_1}{J_3}} \quad (1.56)$$

$$s = s^- - \epsilon \quad (1.57)$$

onde  $c$  é dado por 1.51,  $\epsilon$  é dado por 1.53, e o inverso da temperatura é dado por 1.48.

### A.2.3 Transição Paramagnética - Ferromagnética (A)

Para valores grandes de  $\frac{\sqrt{MN_b}}{N_c}$  temos que a interação entre os clusters e o background é tão grande que qualquer ordem nos grupos implica em certa ordem no background e conseqüentemente ordem de longo alcance entre grupos. A fase super-paramagnética, nestas condições, não é estável e a transição entre as fases paramagnética  $(0, 0, 0)$  e ferromagnética  $(s, t, \hat{s})$  ocorre diretamente. Neste caso para encontrarmos a temperatura de transição e os valores de  $s, t, \hat{s}$  necessitamos resolver quatro equações que são as três equações 1.23, 1.24 e 1.25, e a condição da igualdade da energia livre na transição,

$$\begin{aligned} f(s, t, \hat{s}) &= f(0, 0, 0) \\ -\frac{\beta J_1}{2q} (q-1) s^2 - \frac{\beta J_2}{2q N_c} (q-1) t^2 - \beta J_3 \frac{\sqrt{MN_b}}{q N_c} (q-1) st\hat{s} + \\ &- \mathcal{S}(s) - \frac{N_b}{N_c} \mathcal{S}(t) - \frac{M}{N_c} \mathcal{S}(\hat{s}) = 0 \end{aligned} \quad (1.58)$$

Consideraremos ao resolver estas equações, que  $J_2 \ll J_1 \ll J_3$  implicando que a temperatura de transição está bem acima da temperatura de ordenamento do background. Então, assumindo que  $t \ll 1$ , encontramos a partir de 1.24,

$$\begin{aligned} -\beta J_2 \frac{N_b}{N_c} t - \beta J_3 \frac{\sqrt{MN_b}}{N_c} s \hat{s} + \frac{N_b}{N_c} &= 0 \\ t = \beta \sqrt{\frac{M}{N_b}} \frac{J_3 s \hat{s}}{(q - \beta J_2)} & \end{aligned} \quad (1.59)$$

Este resultado é o mesmo de 1.43. Substituindo 1.59 em 1.23 e assumindo que  $\hat{s} = 1 - \delta$ , obtemos até ordem zero em  $\delta$ ,

$$-\beta J_1 s - \frac{M}{N_b} \frac{(\beta J_3)^2}{(q - \beta J_2)} s + \log [1 + (q - 1) s] - \log (1 - s) = 0 \quad (1.60)$$

Expandindo  $\mathcal{S}(t)$  até primeira ordem em  $t$  e  $\mathcal{S}(\hat{s})$  até ordem zero em  $\delta$  e utilizando 1.59 obtemos de 1.58,

$$\frac{(q - 1)}{2q} \left[ -\beta J_1 - \frac{M}{N_c} \frac{(\beta J_3)^2}{q - \beta J_2} \right] s^2 - S(s) + \frac{M}{N_c} \log q = 0 \quad (1.61)$$

Desprezando o termo  $\frac{M}{N_c} \log q$  ( $N_c \gg M$ ) temos que as equações 1.60 e 1.61 conduzem a soluções semelhantes às 1.29, 1.30 que são soluções para o modelo de Potts homogêneo,

$$\beta J_1 + \frac{M}{N_c} \frac{(\beta J_3)^2}{(q - \beta J_2)} = \frac{2(q-1)}{q-2} \log (q - 1) \quad (1.62)$$

$$s_{pf} = \frac{q-2}{q-1} \quad (1.63)$$

Utilizando  $\sqrt{\frac{M}{N_c}} J_3$ ,  $J_2 \ll 1$  encontramos a partir de 1.62 até ordem zero em  $\sqrt{\frac{M}{N_c}} J_3$ ,  $J_2$  que o valor do inverso da temperatura de transição é o mesmo daquele encontrado para a transição paramagnética - superparamagnética 1.29.

### A.2.4 Transição de ordem dos spins do background: fase ferromagnética (B)

Consideraremos agora  $J_3 \ll J_2$  ou seja a ordem no background não significa necessariamente ordem dos super-spins, então  $\hat{s} \ll 1$ . Os cálculos são semelhantes aos cálculos da seção A.2.2, sendo que agora a transição será do estado super-paramagnético para o estado ferromagnético (B). Começemos aproximando a equação 1.25 com a condição  $\hat{s} \ll 1$ ,

$$\begin{aligned} -\beta J_3 \frac{\sqrt{MN_b}}{N_c} st + \frac{M}{N_c} [(q-1)\hat{s} + \hat{s}] &= 0 \\ \hat{s} &= \frac{\beta J_3}{q} \sqrt{\frac{N_b}{M}} st. \end{aligned} \quad (1.64)$$

Substituindo 1.64 em 1.24 com a condição  $J_3 \ll J_2$ , conduz à

$$-\beta J_2 \frac{N_b}{N_c} t + \frac{N_b}{N_c} \{\log(1 + (q-1)t) - \log(1-t)\} = 0 \quad (1.65)$$

Aproximando 1.42 temos

$$\frac{N_b}{N_c} \left[ -\beta \frac{J_2}{2q} (q-1)t^2 - S(t) \right] - \beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1) \hat{s} t s^- - \frac{M}{N_c} \mathcal{S}(\hat{s}) = 0 \quad (1.66)$$

Podemos aproximar o segundo termo da equação acima como segue,

$$-\beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1) \hat{s} t s^- \approx -\frac{\sqrt{MN_b}}{N_c} (\beta J_3 t s) \hat{s}. \quad (1.67)$$

Utilizando 1.64, resulta,

$$-\beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1) \hat{s} t s^- \approx -\frac{M}{qN_c} (q-1) \hat{s}^2 \quad (1.68)$$

Como  $\hat{s}^2 \ll 1$  e  $\frac{M}{N_c} \ll 1$  podemos fazer a seguinte aproximação,

$$-\beta J_3 \frac{\sqrt{MN_b}}{qN_c} (q-1) \hat{s} t s^- \approx \mathcal{O} \left[ \left( \frac{M}{N_c} \right)^{\frac{1}{2}} \hat{s} \right] \approx 0 \quad (1.69)$$

Para o terceiro termo de 1.66 temos,

$$\frac{M}{N_c} \mathcal{S}(\hat{s}) \approx \frac{M}{N_c} \hat{s}^2 \approx 0 \quad (1.70)$$

então, 1.66 pode ser aproximado como,

$$-\beta \frac{J_2}{2q} (q-1) t^2 - \mathcal{S}(t) = 0. \quad (1.71)$$

Novamente, as equações 1.65 e 1.71 levam a solução semelhantes àquelas do sistema homogêneo de spins de Potts [14] ou seja,

$$\beta J_2 = \frac{2(q-1)}{q-2} \log(q-1), \quad (1.72)$$

$$t = \frac{q-2}{q-1}. \quad (1.73)$$

Substituindo 1.72 e 1.73 em 1.64 encontramos com  $s \approx 1$ ,

$$\hat{s} = \frac{2 \log(q-1)}{q} \sqrt{\frac{N_b}{M}} \frac{J_3}{J_2} \quad (1.74)$$

No caso  $J_3 \gg J_2$ , na fase ferromagnética (A) os super-spins agem como um “campo ordenante” para os spins do background. Aproximando  $s = 1$  e  $\hat{s} = 1$  na energia livre 1.22 temos um sistema de spins com interação  $J = J_2$  e campo externo  $h = \sqrt{\frac{M}{N_b}} J_3$ . O problema então reduz-se à transição de fase em um modelo de Potts em campo médio com campo externo. Na ausência de campo existe uma transição de primeira ordem que só será suprimida com um campo suficientemente forte. Então, existe uma linha de transição de primeira ordem com uma extremidade

em um ponto fixo que neste caso é

$$\beta_{ff}J = 4\frac{q-1}{q}, \quad \frac{h_c}{J} = \frac{q[\log(q-1) - 2(q-2)]}{4(q-1)}. \quad (1.75)$$

Para campos abaixo de  $h_c$  existe transição de primeira ordem. A condição para ocorrer uma transição dos spins do background é então,

$$\sqrt{\frac{M}{N_b}} \frac{J_3}{J_2} \leq \frac{h_c}{J} \quad (1.76)$$

Quando  $\frac{M}{N_b} \rightarrow 0$ , a transição ocorre para qualquer valor finito de  $J_3$  e  $J_2$  e o ponto crítico tende ao infinito. Isto é o que mostra a linha “4” da figura 2.2 (capítulo 2).

### A.3 FUNÇÕES DE CORRELAÇÃO

As funções de correlação podem ser obtidas da diferenciação da energia livre 1.22, com relação à  $J_1$ ,  $J_2$  e  $J_3$ ,

$$\langle \delta_{\sigma_i^a, \sigma_j^a} \rangle = \frac{q-1}{q} s^2 + \frac{1}{q} \quad (1.77)$$

$$\langle \delta_{\sigma_i^b, \sigma_j^b} \rangle = \frac{q-1}{q} t^2 + \frac{1}{q} \quad (1.78)$$

$$\langle \delta_{\sigma_i^a, \sigma_j^b} \rangle = \frac{q-1}{q} st\hat{s} + \frac{1}{q} \quad (1.79)$$

A correlação entre spins pertencentes a grupos diferentes  $a$  e  $a'$  pode ser calculada, adicionando-se uma interação infinitesimal  $J_4$  à Hamiltoniana,

$$\tilde{\mathcal{H}} = \mathcal{H} - \frac{M}{N_c^2} J_4 \sum_{a < a'} \sum_{i < j} \delta_{\sigma_i^a, \sigma_j^{a'}}. \quad (1.80)$$

Com a aproximação de campo médio, esta hamiltoniana conduz à seguinte função

de correlação,

$$\langle \delta_{\sigma_i^a, \sigma_j^{a'}} \rangle = \frac{q-1}{q} s^2 \hat{s}^2 + \frac{1}{q} \quad (1.81)$$

Utilizando os valores de  $s$ ,  $t$  e  $\hat{s}$  nas transições temos, para cada uma das fases:

- Na fase paramagnética,

$$\langle \delta_{\sigma_i^a, \sigma_j^a} \rangle = \langle \delta_{\sigma_i^b, \sigma_j^b} \rangle = \langle \delta_{\sigma_i^a, \sigma_j^b} \rangle = \langle \delta_{\sigma_i^a, \sigma_j^{a'}} \rangle = \frac{1}{q} \quad (1.82)$$

- Na fase super-paramagnética a correlação entre spins do mesmo grupo muda abruptamente enquanto que as outras correlações permanecem inalteradas,

$$\langle \delta_{\sigma_i^a, \sigma_j^a} \rangle \approx 1 - \frac{2}{q} + \mathcal{O}\left(\frac{1}{q^2}\right) \quad (1.83)$$

- Na fase ferromagnética,

$$\langle \delta_{\sigma_i^a, \sigma_j^a} \rangle \geq \frac{q-1}{q} \left[ 1 - 2 \exp\left(-\frac{cJ_1}{J_3}\right) \right] + \frac{1}{q} \approx 1 \quad (1.84)$$

$$\langle \delta_{\sigma_i^b, \sigma_j^b} \rangle > \left[ \frac{2(q-1)}{q^2} \log(q-1) \right] \frac{M}{N_b} + \frac{1}{q} \quad (1.85)$$

$$\langle \delta_{\sigma_i^a, \sigma_j^b} \rangle > \frac{(q-2)}{q} \sqrt{\frac{2(q-2)}{q(q-1)} \log(q-1)} \sqrt{\frac{M}{N_c}} + \frac{1}{q} \quad (1.86)$$

$$\langle \delta_{\sigma_i^a, \sigma_j^{a'}} \rangle > \frac{(q-2)^2}{q(q-1)} \left[ 1 - 2 \exp\left(-\frac{cJ_1}{J_3}\right) \right] + \frac{1}{q} \quad (1.87)$$

A mudança significativa é que spins pertencentes a diferentes grupos estão agora correlacionados.

## A.4 SUSCEPTIBILIDADE

A susceptibilidade do sistema pode ser calculada adicionando-se um termo de campo magnético na Hamiltoniana,

$$-h \left( \sum_{a,i} \delta_{\sigma_i^a,1} + \sum_i \delta_{\sigma_i^b,1} \right) \quad (1.88)$$

que resulta em um termo adicional na energia livre 1.22,

$$-\beta h \left[ \left( \frac{q-1}{q} s \hat{s} + \frac{1}{q} \right) + \frac{N_b}{N_c} \left( \frac{q-1}{q} t + \frac{1}{q} \right) \right] \quad (1.89)$$

Definindo  $\tilde{h} = \beta h$  a susceptibilidade por spin é dada por,

$$\frac{\chi}{\beta} = -\frac{d^2(\beta f)}{d\tilde{h}^2} = \frac{d}{d\tilde{h}} \left( \frac{d\beta f}{d\tilde{h}} \right). \quad (1.90)$$

Sendo  $d\beta f = \sum_i \frac{\partial \beta f}{\partial J_i} dJ_i + \frac{\partial \beta f}{\partial \tilde{h}} d\tilde{h}$  e  $J_i$  independente de  $\tilde{h}$ ,

$$\frac{d\beta f}{d\tilde{h}} = \frac{\partial \beta f}{\partial \tilde{h}}. \quad (1.91)$$

Substituindo 1.91 em 1.90 e utilizando 1.89, temos,

$$\frac{\chi}{\beta} = \frac{q-1}{q} \left( \hat{s} \frac{\partial s}{\partial \tilde{h}} + s \frac{\partial \hat{s}}{\partial \tilde{h}} + \frac{N_b}{N_c} \frac{\partial t}{\partial \tilde{h}} \right) \quad (1.92)$$

Na fase ferromagnética o único termo que contribui para a susceptibilidade é o terceiro termo de 1.92, logo,

$$\frac{\chi}{\beta} = \frac{N_a}{N_b} \frac{q-1}{q} \frac{1}{(q - \beta J_2)} \quad (1.93)$$

Na fase ferromagnética  $\chi \approx \mathcal{O}(1)$ . Na transição para a fase ferromagnética,

a suscetibilidade possui um valor da ordem de  $\chi \approx \mathcal{O}\left(\frac{N_c}{M}\right)$ ,

$$\frac{\chi}{\beta} = \frac{(q-2)^2}{q^2(q-1)} \frac{N_c}{M} \quad (1.94)$$

Na transição ferromagnética - super-paramagnética a susceptibilidade é aproximadamente,

$$\frac{\chi}{\beta} = \frac{(q-1)}{q^2} \frac{N_c}{M} \quad (1.95)$$

Na figura 2.3 (capítulo 2) tem-se o gráfico da susceptibilidade baseada na equação 1.92 com os seguintes parâmetros,  $T^* = \left(\frac{2(q-1)\log(q-1)}{(q-2)}\right) T$ ,  $J_3 = 0.2236$   $N_b = N_c = 10^6$ .

## APÊNDICE B

# O MÉTODO MONTE CARLO

Neste apêndice, mostraremos as ideias gerais que fundamentam o método Monte Carlo e mostraremos, também, três dos algoritmos mais utilizados em pesquisa atualmente.

### B.1 MECÂNICA ESTATÍSTICA DE EQUILÍBRIO

Diante da impossibilidade de tratamento exato de vários sistemas de muitos componentes, a Mecânica Estatística propõe um tratamento estatístico onde é estimada a probabilidade de um sistema se encontrar em um dado estado. Para sistemas termodinâmicos, apenas um número reduzido de estados são relevantes e contribuem efetivamente para o cálculo da média de grandezas termodinâmicas, dentre todos os estados possíveis, e isto garante a validade da mecânica estatística.

Em termos gerais temos que a dinâmica de um sistema pode ser representada por,

$$\frac{d\omega_\mu}{dt} = \sum_\nu [\omega_\nu(t) R(\nu \rightarrow \mu) - \omega_\mu(t) R(\mu \rightarrow \nu)], \quad (2.1)$$

onde  $\omega_\mu(t)$  é a probabilidade do sistema estar no estado  $\mu$  no tempo  $t$  e  $R(\nu \rightarrow \mu)$  é a taxa de transição entre os estados  $\nu$  e  $\mu$ . As probabilidades  $\omega_\mu(t)$  devem satisfazer à condição

$$\sum_\mu \omega_\mu = 1. \quad (2.2)$$

Resolvendo a equação 2.1 com o vínculo 2.2, temos uma distribuição de pro-

babilidades  $\omega_\mu$  que varia no tempo. As propriedades macroscópicas podem ser estimadas através de,

$$\langle Q \rangle = \sum_{\mu} Q_{\mu} \omega_{\mu}(t). \quad (2.3)$$

Considerando-se agora que o segundo membro da equação 2.1 anula-se, (ou seja, a taxa com que o sistema transita para o estado  $\mu$  é igual a taxa com que o sistema transita de  $\mu$  para outros estados), temos que  $\frac{d\omega_{\mu}}{dt} = 0$  e a distribuição de probabilidades é constante no tempo. O sistema é dito estar em *equilíbrio*.

Gibbs (1902) mostrou que para um sistema em equilíbrio térmico com um reservatório de calor à temperatura  $T$ , as probabilidades de ocupação dos estados  $\mu$  são,

$$p_{\mu} = \frac{1}{Z} \sum_{\mu} \exp(-E_{\mu}/kT), \quad (2.4)$$

onde  $E_{\mu}$  é a energia do estado  $\mu$ ,  $k$  é a constante de Boltzmann e  $Z$  é a função de partição,

$$Z = \sum_{\mu} \exp(-E_{\mu}/kT). \quad (2.5)$$

Conhecendo-se a função de partição é possível conhecer todo o comportamento macroscópico de um sistema no equilíbrio.

As simulações Monte Carlo geram uma trajetória de configurações segundo uma distribuição de probabilidades  $\omega_{\mu}$ . A essência do método é escolher um conjunto de taxas de transição  $R(\mu \rightarrow \nu)$  de um estado para outro de tal forma que a solução de equilíbrio de 2.1 seja a distribuição de Boltzmann dada na equação 2.4. Geradas estas configurações as grandezas de interesse são então estimadas.

Antes de discutirmos alguma dinâmica específica de Monte Carlo revisemos alguns conceitos mais gerais.

## B.2 PROCESSOS DE MARKOV, ERGODICIDADE E BALANCEAMENTO DETALHADO

Os processos de Markov, que fundamentam o método Monte Carlo, consistem em gerar uma trajetória de configurações tal que a nova configuração é escolhida de forma aleatória e depende somente da configuração imediatamente anterior. A probabilidade de se gerar um estado  $\nu$ , dado um estado  $\mu$  é chamada de probabilidade de transição  $P(\mu \rightarrow \nu)$  que depende somente dos estados  $\mu$  e  $\nu$ .

Partindo de um estado qualquer, o processo de Markov conduz o sistema para um estado de “equilíbrio”: uma sucessão de estados com probabilidades dadas pela distribuição de Boltzmann.

A condição de ergodicidade é que torna possível ao processo de Markov, partindo de um dado estado, alcançar qualquer outro estado em um intervalo de tempo suficiente para isto.

A condição de ergodicidade é necessária pois na distribuição de Boltzmann, todo estado  $\nu$  aparece com uma probabilidade diferente de zero, sendo assim se o estado  $\nu$  for inacessível a  $\mu$  a probabilidade de se encontrar  $\nu$  em uma cadeia de Markov é zero, não correspondendo ao peso de Boltzmann associado ao estado  $\nu$ .

A condição de balanceamento detalhado garante que a distribuição de Boltzmann é gerada no equilíbrio, e não outra distribuição. Da condição de equilíbrio temos,

$$\sum_{\nu} p_{\mu} P(\mu \rightarrow \nu) = \sum_{\nu} p_{\nu} P(\nu \rightarrow \mu). \quad (2.6)$$

Esta condição é essencial mas não suficiente para garantir que a distribuição de equilíbrio seja a distribuição de Boltzmann (ver detalhes em [26]). Uma condição adicional para as probabilidades de transição, deve ser considerada,

$$p_{\mu} P(\mu \rightarrow \nu) = p_{\nu} P(\nu \rightarrow \mu). \quad (2.7)$$

Esta é a condição de balanceamento detalhado. Esta condição satisfaz a equação 2.6 e nos mostra que na média, o sistema transita do estado  $\mu$  para o estado  $\nu$  na mesma taxa com que transita do estado  $\nu$  para o estado  $\mu$ , em uma certa temperatura  $T$ .

Escrevendo  $p_\mu$  e  $p_\nu$  satisfazendo à distribuição de Boltzmann encontramos de 2.7,

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{p_\nu}{p_\mu} = \exp[-\beta(E_\nu - E_\mu)] \quad (2.8)$$

Este é um vínculo para as probabilidades de transição  $P(\mu \rightarrow \nu)$ , que juntamente com a condição de ergodicidade garante que a distribuição de equilíbrio do processo de Markov é a distribuição de Boltzmann.

### B.3 RAZÃO DE ACEITAÇÃO

As condições descritas acima fundamentam o método Monte Carlo, mas especificamente, que processo utilizar para gerar a cadeia de Markov?

O processo utilizado no método Monte Carlo é o seguinte: gerar um estado  $\nu$  a partir de certo estado  $\mu$  e aceitá-lo com uma dada probabilidade. Matematicamente,

$$P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu) A(\mu \rightarrow \nu). \quad (2.9)$$

A quantidade  $g(\mu \rightarrow \nu)$  é a probabilidade de seleção e  $A(\mu \rightarrow \nu)$  é a razão de aceitação. Resolvendo o vínculo 2.8 temos,

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{g(\mu \rightarrow \nu) A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu) A(\nu \rightarrow \mu)} \quad (2.10)$$

Para um algoritmo ser eficiente é necessário se escolher razões de aceitação com valores próximos de  $um$ , o quanto possível. Caso contrário tem-se um algoritmo computacionalmente dispendioso, pois muitas configurações devem ser geradas para se escolher apenas uma destas configurações.

## B.4 O ALGORITMO METROPOLIS

O algoritmo Metropolis é o algoritmo de Monte Carlo mais conhecido e utilizado e foi introduzido por Metropolis *et al* em 1953 [27]. Este algoritmo utiliza a dinâmica de mudar o estado de um único spin, para gerar novas configurações, o que é razoável, pois no equilíbrio as flutuações de energia são pequenas, o que justifica esta escolha.

No algoritmo Metropolis, as probabilidades de seleção  $g(\mu \rightarrow \nu)$  para cada um dos possíveis estados  $\nu$  são iguais e a probabilidade de seleção dos outros estados é nula.

Existe então, partindo-se de um estado  $\mu$ ,  $N$  possíveis estados  $\nu$  onde  $N$  é o número de spins total, e as probabilidades de seleção não nulas são,

$$g(\mu \rightarrow \nu) = \frac{1}{N} \quad (2.11)$$

A condição de balanceamento detalhado 2.7 torna-se então,

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{g(\mu \rightarrow \nu) A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu) A(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\mu - E_\nu)} \quad (2.12)$$

Para termos o algoritmo o mais eficiente possível, a razão de aceitação deve ser mais próxima de  $um$ , o quanto possível. Para o algoritmo Metropolis esta razão é,

$$A(\mu \rightarrow \nu) = \begin{cases} e^{-\beta(E_\mu - E_\nu)} & , \text{ se } E_\mu - E_\nu > 0. \\ 1 & , \text{ se } E_\mu - E_\nu \leq 0. \end{cases}$$

ou seja, um estado selecionado é aceito se este possui uma energia menor que o estado presente. Se a energia do estado selecionado for maior, o estado selecionado é aceito com probabilidade  $e^{-\beta(E_\mu - E_\nu)}$ .

## B.5 OS ALGORITMOS DE WOLFF E DE SWENDSEN-WANG

O modelo físico que fundamentará a discussão desta seção é o modelo de Ising, discutido na seção 2.2.

Os algoritmos de Wolff [28] e de Swendsen-Wang [21] são algoritmos que atualizam cada configuração invertendo o estado dos spins de um cluster (algoritmo de Wolff) ou de todos os clusters com uma certa probabilidade (algoritmo de Swendsen-Wang). A grande vantagem destes algoritmos com relação ao algoritmo Metropolis é que na transição de fase, o sistema apresenta domínios de spins correlacionados, que para o algoritmo Metropolis são computacionalmente muito custosos de serem invertidos, pois este algoritmo atualiza cada configuração spin por spin.

A idéia do algoritmo de Wolff é inverter em um só passo de Monte Carlo, um cluster de spins correlacionados. Algoritmos deste tipo são chamados de *algoritmos de cluster*, e seus tempos de relaxação <sup>1</sup> são menores que o tempo de relaxação do algoritmo Metropolis.

A estratégia que o algoritmo de Wolff utiliza é escolher aleatoriamente um sítio, e a partir deste sítio construir um cluster, verificando se os vizinhos de cada sítio incorporado ao cluster estão no mesmo estado do spin situado no sítio original. Mas nem todo spin com mesmo estado é associado ao cluster. A adição de um novo spin deve ser feita com certa probabilidade. Por exemplo, a altas temperaturas os spins estão descorrelacionados e a probabilidade do algoritmo formar clusters deve ser muito pequena. Se houver dois spins vizinhos no mesmo estado a probabilidade deles estarem em um mesmo cluster deve ser próxima de zero, ou seja a probabilidade de se adicionar um spin ao cluster  $P_{add}$  deve ser próxima de zero. Esta probabilidade  $P_{add}$  depende da temperatura e deve ser tal que o algoritmo esteja de acordo com a condição de balanceamento detalhado (seção B.2) e que sua razão de aceitação seja

---

<sup>1</sup>Tempo (contado em unidades de passos de Monte Carlo) necessário para o sistema, a partir de uma configuração inicial, alcançar o equilíbrio.

mais próxima de  $um$  o quanto possível (B.3). Esta probabilidade é [26],

$$P_{add} = 1 - e^{-2\beta J} \quad (2.13)$$

Com esta escolha, a razão de aceitação é 1 e independe das propriedades dos estados, bem como independe da temperatura.

Notemos que quando  $T \rightarrow \infty$ ,  $P_{add} \rightarrow 0$  e quando  $T \rightarrow 0$ ,  $P_{add} \rightarrow 1$ , ou seja, o tamanho dos clusters tende a aumentar quando a temperatura decresce. O algoritmo de Wolff para o modelo de Ising pode ser implementado da seguinte forma:

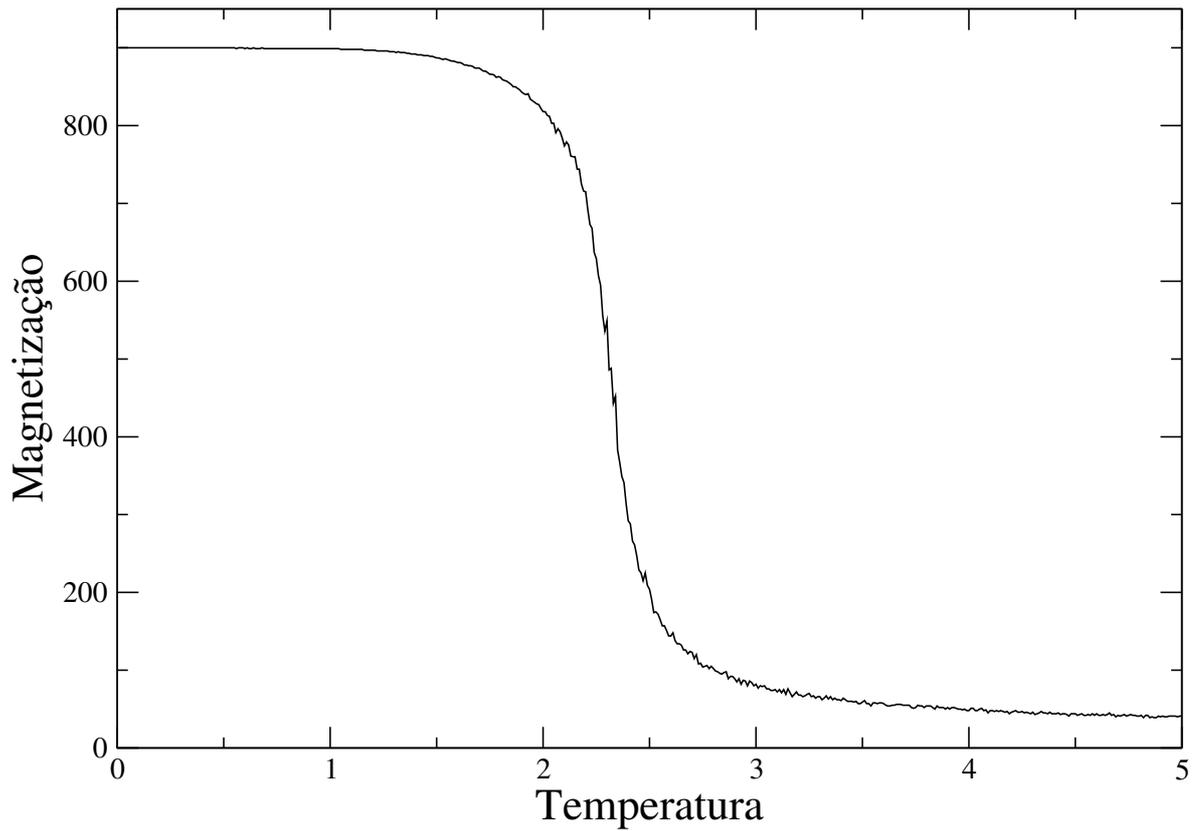
- i) Escolher um spin em um sítio, aleatoriamente na rede.
- ii) Verificar cada vizinho do spin escolhido inicialmente. Aqueles vizinhos que possuem o mesmo estado do spin original são adicionados ao cluster com probabilidade dada pela equação 2.13.
- iii) Para cada spin adicionado no passo anterior verificar seus vizinhos. Aqueles vizinhos que possuem o mesmo estado do spin original, são adicionados ao cluster com probabilidade dada por 2.13. Este passo é então repetido, *ad nauseum*, até não haver mais spins para serem adicionados.
- iv) Inverter o estado dos spins que formam o cluster.

O algoritmo de Swendsen-Wang é semelhante ao algoritmo de Wolff. Neste caso, toda a rede é dividida em clusters. Estes clusters são formados ao se atribuir ligações entre sítios vizinhos, que possuem o mesmo estado, com a mesma probabilidade 2.13 utilizada no algoritmo de Wolff. Os clusters são formados por spins pertencentes ao mesmo grafo.

No algoritmo de Swendsen-Wang, ao invés de se inverter o estado dos elementos de um único cluster, inverte-se o estado dos spins de todos os clusters com

---

probabilidade 0.5, ou seja para cada cluster é decidido independentemente se o estado de seus componentes deve ser invertido ou não. Como um exemplo da eficiência do algoritmo de Swendsen-Wang, em regiões próximas a transição de fase, implementamos o algoritmo de Swendsen-Wang para um sistema de  $30 \times 30$  spins de Ising em uma rede quadrada. Na figura B.1, mostramos a magnetização em função da temperatura para este sistema, utilizando-se o algoritmo de Swendsen-Wang. Observemos que as flutuações próximas a transição de fase são pequenas ao se aplicar o algoritmo de Swendsen-Wang.



**Figura B.1.** Gráfico da magnetização em função da temperatura para um sistema ferromagnético de  $30 \times 30$  spins de Ising em uma rede quadrada, obtido através do algoritmo de Swendsen-Wang com 2000 passos de Monte Carlo sendo 1000 para a termalização. Note que as flutuações próximas a transição de fase são pequenas ao se aplicar o algoritmo de Swendsen-Wang.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] A. K. Jain, R. C. Dubes, *Algoritms for Clustering Data* (Prentice Hall, Englewood Cliffs, NJ, 1988).
- [2] A. K. Jain, M. N. Murty, P. J. Flynn, *ACM Comput. Surveys*, **3**, vol. 31, 264 (1999)
- [3] M. Blatt, S. Wiseman, E. Domany, *Phys. Rev. Lett.* **76**, 3251 (1996).
- [4] M. Blatt, S. Wiseman, E. Domany, *Advances in Neural Information Processing Systems* **8**, 416 (1996).
- [5] E. Domany, *Progress in Statistical Physics, Proc. Choh Memorial Conference, Seoul (1997)*, World Scientific pag. 213 (1998) [Superparamagnetic Clustering of Data - The definitive solution of an ill-posed problem].
- [6] F. De Carlo , *et al.*, *Phys. Rev. Lett.* **85**, 554 (2000).
- [7] W. R. Dillon, M. Goldstein, *Multivariate Analysis: methods and applications* (John Wiley, New York, 1984).
- [8] K. C. Gowda, G. Krishna, *Pattern. Recogn.*, **10**, 105 (1977).
- [9] C. Von Ferber, F. Wörgötter, *Phys. Rev. E*, **62**, 1461 (2000).
- [10] L. Kullmann, J. Kertesz, R. N. Mantegna, *Phys. A*, **287**, 412 (2000).
- [11] N. Basalto, *et al.*, arXiv:cond-mat/040497 v1 (2004).

- 
- [12] M.Blatt, S. Wiseman, E. Domany, Phys. A, **279**, 457, (2000).
- [13] G. Getz, E. Levine, E. Domany, P. Natl. Acad. Sci. USA, **97**, 12079 (2000).
- [14] F. Y. Wu, Rev. Modern. Phys. **1**, vol. 54, 235, (1982).
- [15] D. C. Mattis, *The Theory of Magnetism* (Harper & Row, New York, 1965)
- [16] R. M. Write, *Quantum Theory of Magnetism* (McGraw-Hill, New York, 1970).
- [17] W. Janke, S. Kappler, Phys. Lett. A **197**(3), 227 (1995).
- [18] E. T. Jaynes, Phys. Rev., **4**(106), 620 (1957).
- [19] M.Blatt, S. Wiseman, E. Domany, Phys. Rev. E. **57**, 3767 (1998).
- [20] B. Joe, SIAM J. Sci. Comput., **14**, 1415, (1993).
- [21] R. H. Swendsen, J. S. Wang, Phys. Rev. Lett., **58**, 86 (1987).
- [22] R. H. Swendsen, J. S. Wang, Phys. A., **167**, 565 (1990).
- [23] J. P. Neirotti, S. M. Kurcbart, N. Caticha, Phys. Rev. E **68**, 031911 (2003)
- [24] L. Angellini, *et al.*, Opt. Eng. **40**, 2878 (2001)
- [25] R. Belloti, *et al.*, Phys. A, **334**, 222 (2004).
- [26] M. E. J. Newman, G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Claredon Press, Oxford, 1999).
- [27] N. Metropolis, *et al.*, J. Chem. Phys., **21**, 1087 (1953).
- [28] U. Wolff, Phys. Rev. Lett., **62**, 759 (1989).