

UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
CORSO DI LAUREA IN FISICA

COMPORTAMENTO CRITICO ED EFFETTI DINAMICI
IN UN GAS FORZATO
SU RETICOLO QUASI UNIDIMENSIONALE

RELATORE:
PROF. SERGIO CARACCILO
CORRELATORE:
DOTT. ANTONIO RAGO

CLAUDIO MANZATO
MATR. N° 554246
PACS 05.50.Q, 64.60

ANNO ACCADEMICO 2005/2006

alla mia famiglia
a Valeria

... the basic point is that the nonequilibrium system is *defined* by its dynamics without regard to any concept of energy and it is the dynamics which should seem reasonable or “physical”. This is distinct from an equilibrium system where the energy function should be “physical” and the dynamics is usually defined in an ad hoc way simply to guarantee that one obtains the Gibbs-Boltzmann weight with the specified energy.

M. R. Evans

*Phase transitions in one-dimensional
nonequilibrium systems*[19]

INDICE

Indice	i
1 Markov Chain Monte Carlo	1
<i>Dove incontreremo una famiglia particolare di processi stocastici, grazie alla quale è possibile generare campioni distribuiti con quasi ogni distribuzione voluta. Dove impareremo come sfruttarne le proprietà e dove scopriremo come questi processi divengano particolarmente interessanti quando procedono accoppiati.</i>	
1.1 Catene di Markov	2
1.2 Markov Chain Monte Carlo	13
1.3 Coupling	16
2 Il Modello	23
<i>Dove viene finalmente introdotto l'oggetto dei nostri studi. Non subito, però. Prima si parlerà del gas su reticolo. Quindi si illustrerà il gas forzato su reticolo bidimensionale dando risalto, in entrambi i casi, alla definizione della dinamica microscopica e alle proprietà possedute. Infine verrà esposto la varietà quasi unidimensionale del gas su reticolo.</i>	
2.1 Gas su reticolo	23
2.2 Gas forzato su reticolo	27
2.3 Proprietà	30
2.4 Gas forzato su reticolo quasi unidimensionale	35
3 Dalle Teorie d'Equilibrio al Driven Diffusive System	41
<i>Dove si renderà conto del tentativo di considerare un'infinità di particelle discrete come un continuo, e di descrivere il loro comportamento dimenticandosi della loro esistenza.</i>	
3.1 Sistemi all'Equilibrio	41
3.2 Sistemi Lontano dall'Equilibrio	47

4	Analisi del Modello	51
4.1	Setup Sperimentale	51
4.2	Conclusioni e Prospettive	66
4.A	Breve discussione sul metodo di analisi impiegato	68
4.B	Dati di coalescenza	71
	il multispin coding, nota storica	79
	<i>Dove si parla di come l'introduzione dei supercomputer abbia rivoluzionato lo studio numerico di modelli stocastici</i>	
	1970: Friedberg e Cameron	79
	1979-1982: La Multiple Spin Coding Technique	81
	1984: Il super-spin coding	83
	1985	84
	1986: Inizia il msc su reticoli in parallelo	88
	1987: Kikuchi e Okabe [39]	90
	1988: Roland e Grant [54]	90
	1989: Zhang [67]	90
	1991	91
	2006: Lee & Okabe, Reuse [41]	92
	Bibliografia	97

INTRODUZIONE

Tra la fine del diciannovesimo e l'inizio del ventesimo secolo, Boltzmann e Gibbs posero le basi della meccanica statistica d'equilibrio, della quale uno dei maggiori successi fu il collegamento tra interazioni microscopiche in sistemi a molti corpi e comportamento macroscopico all'equilibrio. Grazie alla loro descrizione è possibile, almeno in linea teorica, calcolare ogni osservabile macroscopica di interesse. A queste basi seguì una serie di continui progressi per mezzo dei quali si è giunti oggi a possedere una buona comprensione dei fenomeni critici d'equilibrio, in particolare grazie alle nozioni legate al gruppo di rinormalizzazione, tra le quali il concetto di classi d'universalità. Anche le proprietà dinamiche di tali sistemi sono comprese in modo abbastanza soddisfacente [30, 1, 68].

Il modello di Ising, pur poco considerato ai suoi albori, guadagnò nel tempo grande considerazione grazie soprattutto alla comprensione del fatto che le caratteristiche essenziali di sistemi cooperativi (specialmente al punto critico), non dipendono dai dettagli delle forze intermolecolari, ma dal meccanismo di propagazione di ordine a lungo raggio. Dunque grazie ad un modello ipersemplificato della realtà si ottenne la possibilità di avere accesso ad informazioni su sistemi strutturalmente ben più complessi. Naturalmente ad esso seguirono molte varianti, ognuna nel tentativo di racchiudere qualche caratteristica saliente di modelli reali.

In seguito tali modelli furono sfruttati anche dalle già citate teorie di campo ed ebbero nuova fortuna come rappresentazioni discrete di teorie al continuo, quali la teoria di Ginzburg-Landau ϕ^4 per il modello di Ising e l'associato modello B per la dinamica diffusiva (di Kawasaki), come nel caso del gas su reticolo introdotto da Lee e Yang nel 1952.

I metodi Monte Carlo (MC), che dovremmo chiamare più propriamente Markov Chain Monte Carlo (MCMC), sfruttano al contempo le proprietà dei sistemi all'equilibrio e delle catene di Markov per riuscire a simulare gli stati

stazionari dei primi. Infatti tali processi stocastici, se soddisfacenti opportune proprietà quali l'ergodicità, giungono ad una distribuzione stazionaria. Richiedendo inoltre anche la reversibilità si riesce a rendere tale distribuzione indipendente dalla specifica scelta della matrice di transizione, a condizione che il rapporto tra le probabilità di transizione tra i vari stati rimanga inalterata. In questo modo non solo si è in grado di poter simulare uno stato stazionario in base ad una distribuzione di probabilità nota, quale quella di Gibbs $e^{-\beta\mathcal{H}}$, ma si ha anche la possibilità di scegliere il modo più opportuno (ad esempio il più veloce) di giungere a tale distribuzione. In termini più fisici la distribuzione di equilibrio di un sistema viene simulata indipendentemente dalla dinamica che il sistema adotta per giungervi, se tale dinamica soddisfa opportune proprietà, tra cui quella di bilancio dettagliato. In tale caso si parla di "disaccoppiamento" della distribuzione d'equilibrio dalla dinamica.

Purtroppo in natura sistemi realmente all'equilibrio sono molto rari, se non inesistenti in quanto richiederebbero un isolamento totale dall'ambiente circostante. Molto spesso ciò che riteniamo all'equilibrio si trova in uno stato che varia molto lentamente rispetto al tipico tempo d'osservazione, come il diamante che impiega milioni di anni per ridursi a grafite, a temperatura ambiente. In altri casi ciò che vediamo si trova in quello che si dimostra essere uno stato stazionario di non-equilibrio, il più classico dei casi un filo percorso da corrente che si trova ad una certa temperatura stazionaria dovuta al bilanciamento tra la dissipazione termica e il calore prodotto dal flusso costante di elettroni all'interno di esso.

Questi due sistemi (il diamante ed il filo percorso da corrente), appartengono ognuno ad una delle due diverse classi in cui sono divisi i sistemi fuori dall'equilibrio:

1. sistemi cui una sorgente esterna (energia, particelle, etc. . .) impedisce di raggiungere l'equilibrio;
2. sistemi il cui non equilibrio è causato dalla dinamica e non ci sono perturbazioni esterne.

Esempi del primo caso sono la corrente nei fili, l'acqua nei tubi, etc. . . . Spesso questi sistemi si trovano in uno *stato stazionario di non equilibrio* (non-equilibrium steady state, NESS) caratterizzato da proprietà stazionarie e dinamica invariante per traslazioni temporali delle fluttuazioni.

Esempi del secondo caso sono invece tutti quei sistemi che posseggono

una dinamica molto lenta, cosicché il sistema, anche se appena perturbato, impiega moltissimo a rilassare verso l'equilibrio. E, per i tempi comuni d'osservazione, non vi arriva mai.

Un possibile modo per distinguere tra equilibrio e non equilibrio è studiare la condizione di reversibilità microscopica. Essa è fortemente legata all'idea di simmetria per inversione temporale delle quantità dinamiche nello stato stazionario.

La differenza tra l'equilibrio e il primo caso, entrambi dotati di simmetria per traslazioni temporali delle osservabili, è che all'equilibrio vale l'invarianza anche per inversione temporale, fuori equilibrio al contrario il sistema presenta una "direzione temporale privilegiata".

Nel secondo caso invece, dato che si ha un rilassamento infinito verso uno stato d'equilibrio che non viene mai raggiunto, l'invarianza temporale è rotta sia per inversione che per traslazione.

Ricollegandoci all'esempio del filo percorso da un flusso di elettroni, a livello sperimentale è proprio la presenza di una corrente in uno stato stazionario che viene sovente utilizzata come segnale della presenza di un NESS. Un altro fenomeno caratteristico è la mancanza di validità del bilancio dettagliato nella descrizione fisica microscopica. In conseguenza di questi fatti la distribuzione di Gibbs cessa in generale di essere la distribuzione stazionaria del sistema, e con essa cessa anche la possibilità di comprendere e spiegare i fenomeni osservati tramite il background fornito dalla meccanica statistica d'equilibrio.

All'inizio degli anni '80, Katz *et al.* [35, 36] introdussero, motivati anche dallo studio dei conduttori superionici, un modello stocastico per lo studio di uno stato stazionario di non equilibrio. Spinti dall'esigenza di poter studiare un modello abbastanza semplice, che ritenesse al contempo le caratteristiche salienti dei sistemi fisici di nonequilibrio, basarono il loro studio su di una variante del gas su reticolo di Lee e Yang [42], variante qui denominata DLG (Driven Lattice Gas).

Per giungere e rimanere in uno stato stazionario di non equilibrio, essi aggiunsero un termine forzante di campo alle probabilità di transizione microscopiche di Metropolis in modo da determinare un loro incremento (diminuzione) nella direzione equiversa (opposta) al campo, lasciando inalterate quelle ad esso trasverse. Al limite in cui tale campo diviene infinito, le mosse contrarie alla direzione del campo divengono proibite, mentre quelle a favore del campo sono ammesse con probabilità unitaria. Importanza pri-

maria rivestono le condizioni al bordo che, se periodiche, impediscono di poter scrivere il campo come potenziale aggiuntivo all'hamiltoniana del sistema. Possibilità al contrario presente nel caso di bordi chiusi almeno nella direzione del campo, nel qual caso il sistema diviene pseudo-gravitazionale e possiede una configurazione d'equilibrio.

Questi due fattori, il campo e le condizioni al contorno, rendono la catena di Markov sottostante alla realizzazione stocastica del modello non-reversibile, provocando così a livello macroscopico la perdita di validità del bilancio dettagliato. Tale particolare si rivela sufficiente a mantenere il sistema lontano dall'equilibrio, pur possedendo la catena di Markov, irriducibile ed aperiodica, uno stato stazionario.

Per valore nullo del campo il DLG si riduce al gas su reticolo, che presenta una transizione di fase in corrispondenza della ben nota temperatura di Onsager. Al di sotto di questa le particelle si aggregano formando una zona omogenea densa, separata da una zona omogenea vuota. Rispettando considerazioni puramente energetiche l'interfaccia tra le due zone tende a minimizzare la propria superficie, assumendo in due dimensioni per reticoli quadrati una forma circolare. Per sistemi con densità pari a 0.5 la transizione di fase diviene di secondo ordine. Similmente, anche il DLG (campo non nullo) presenta una transizione di fase del secondo ordine per sistemi pieni a metà, con una segregazione in due fasi particelle/vuoto. La principale differenza a livello macroscopico consiste nell'orientazione dell'interfaccia che in questo caso risulta allineata alla direzione del campo. Un altro effetto distintivo del modello è la presenza di una corrente netta di particelle attraverso il sistema.

Oltre a queste caratteristiche "visibili", il DLG presenta molte altre peculiarità che emergono dalle simulazioni numeriche, come la presenza di correlazioni a decadimento secondo legge di potenza anche al di fuori della criticità, una singolarità nell'origine per il fattore di struttura e una temperatura critica superiore a quella di equilibrio, nonostante il campo forzante immetta energia nel sistema.

A partire dalla descrizione microscopica del modello si è passati in seguito alla derivazione di una teoria al continuo dalla quale ottenere, attraverso una procedura di rinormalizzazione, informazioni sul comportamento critico. Generalmente tale compito viene assolto postulando un'opportuna teoria al continuo che contenga le caratteristiche ritenute rilevanti alla criticità.

Janssen e Schmittmann [33] e Leung e Cardy [43, 60] hanno proposto

una teoria al continuo basata su di un'equazione di Langevin descrivente l'evoluzione stocastica della densità locale di particelle. Essa contiene due ingredienti chiave: un termine forzante $\mathcal{E}\nabla_{\parallel}\phi^2$ — dove \mathcal{E} è il campo mesoscopico — che ha come effetto la rottura della simmetria up-down di Ising e la presenza di un alto grado di anisotropia, con due coefficienti di diffusione τ_{\parallel} e τ_{\perp} differenti nelle differenti direzioni. Per render conto di alcune discrepanze con le misure numeriche sono state in seguito introdotte delle correzioni di taglia finita, ottenendo così un accordo molto soddisfacente (si veda per esempio [10]) coi risultati numerici.

In contrasto con questa, Garrido, de los Santos e Muñoz [24, 16] hanno proposto un'altra teoria avente come punto di partenza la Master Equation del modello, derivandone un'equazione di Fokker-Planck e determinando l'equazione di Langevin stocasticamente equivalente. Ciò per tenere in conto i dettagli della dinamica microscopica che influenza il comportamento critico. Il modello da essa rappresentato è detto Anisotropic Diffusive System (ADS) in quanto fa dipendere il comportamento critico solamente dall'anisotropia, non contenendo alcun termine di corrente.

Le due teorie non sono compatibili tra di loro, prevedendo infatti differenti dimensioni critiche, differenti valori degli esponenti critici ed ovviamente differenti classi d'universalità. Sono state derivate partendo da "filosofie" differenti ed i risultati portano a diverse interpretazioni della fisica del modello. Non è nostra intenzione qui dimostrare il valore dell'una o dell'altra (in [56] Schmittmann *et al.* forniscono una concisa esposizione di entrambe), bensì evidenziare come non esista una evidente derivazione teorica univoca a partire da un modello microscopico di non equilibrio.

Anche l'analisi numerica presenta notevoli inconvenienti, rispetto ai modelli dotati di una distribuzione stazionaria d'equilibrio. Infatti lo stato stazionario di non equilibrio, la cui distribuzione in genere non è nota, è strettamente legato alla dinamica microscopica utilizzata nella definizione del modello. Come conseguenza principale di ciò si ha l'impossibilità di poter introdurre algoritmi particolarmente vantaggiosi dal punto di vista computazionale. L'utilizzo di dinamiche più veloci alla criticità, insieme ad algoritmi collettivi come quello dovuto a Swendsen-Wang [59], porterebbero alla definizione di un modello differente da quello di partenza. L'unica soluzione al rallentamento critico è ottenere implementazioni algoritmiche più efficienti della medesima dinamica stocastica. In questa direzione viene lo sviluppo, iniziato negli anni '70, degli algoritmi a multispin coding (MSC). L'idea alla base di tali algoritmi risiede nella capacità di poter immagazzi-

nare le informazioni relative allo stato di un singolo spin, che può assumere valori 0 ed 1, in un singolo bit. Se una “parola” dell’elaboratore contiene N bit, è possibile utilizzarla per contenere lo stato di N spin. In questo modo è possibile far evolvere contemporaneamente N reticoli di volume Λ , utilizzando Λ parole. Un notevole incremento della velocità computazionale si ottiene effettuando i calcoli sui reticoli tramite operazioni binarie (XOR e shift).

Durante il nostro lavoro di tesi abbiamo utilizzato, per l’evoluzione del modello, un algoritmo consistente in una variante del multispin coding in cui tra i vari sistemi è introdotto un accoppiamento utilizzando la stessa sequenza di numeri pseudocasuali sia per la determinazione della coppia di siti su cui effettuare il tentativo di mossa, sia per l’accettazione o meno di tale mossa. Tale accoppiamento, denominato coupling nella letteratura markoviana, fa sì che i reticoli dopo un certo periodo di tempo giungano in una medesima configurazione. Dato che la sequenza di numeri usata è identica, da questo istante in poi rimarranno uguali (da qui il nome coalescenza per questo fenomeno). Tra i molteplici utilizzi del coupling ricordiamo la stima del tempo di termalizzazione, nel caso in cui una delle distribuzioni accoppiate sia quella stazionaria, e recentemente la realizzazione di campioni distribuiti secondo la distribuzione stazionaria, grazie ad algoritmi come il “Coupling from the Past” di Propp e Wilson [52].

L’utilizzo del coupling nel nostro lavoro è basato sull’osservazione di alcune proprietà particolari, a nostra conoscenza non ancora utilizzate. In particolare la stretta dipendenza dell’algoritmo dalla dinamica dei sistemi simulati. Nel caso in esame la stretta non-reversibilità della catena di Markov alla base del modello porta ad una rapida coalescenza dei reticoli che risulta possedere un minimo nella distribuzione dei tempi alla temperatura critica. È risultato dunque un metodo in grado di fornire informazioni sulla dinamica di sistemi stocastici, in particolar modo quando tale dinamica è disaccoppiata dallo stato stazionario.

Uno

MARKOV CHAIN MONTE CARLO

Dove incontreremo un famiglia particolare di processi stocastici, grazie alla quale è possibile generare campioni distribuiti con quasi ogni distribuzione voluta. Dove impareremo come sfruttarne le proprietà e dove scopriremo come questi processi divengano particolarmente interessanti quando procedono accoppiati.

INTRODUZIONE

UN PROBLEMA RICORRENTE IN FISICA consiste nel generare campioni di un sistema con spazio delle configurazioni \mathcal{S} i cui stati sono distribuiti in accordo con una distribuzione π . In generale tale distribuzione è nota, per esempio perché ricavata sperimentalmente, ma in alcuni casi — e il nostro rientra in questi, come vedremo — è nota solamente la dinamica del sistema. Se la distribuzione è nota, potrebbe sembrare un problema banale, ma tale sensazione si rivelerebbe sbagliata. Molto spesso infatti tale compito, a causa della struttura particolarmente complessa della distribuzione, della sua elevata dimensionalità o anche dei limiti di algoritmizzazione, risulta impraticabile. Ancor più ostico risulta se la distribuzione con cui si vorrebbero distribuiti gli stati risulta sconosciuta.

Per risolvere tali problemi una soluzione quasi universalmente adottata è evitare di generare direttamente i campioni dalla distribuzione π , ma piuttosto generare un processo in grado di evolvere in modo stocastico in maniera tale che la propria dinamica lo porti in una distribuzione (detta, a volte impropriamente, d'equilibrio) che coincida con π . Inoltre dovremo essere capaci di raggiungere questa distribuzione a partire da qualunque stato iniziale. In pratica inventiamo un'evoluzione temporale stocastica fittizia in grado di far raggiungere al sistema la distribuzione voluta. A questo

punto potremo calcolare le medie delle osservabili di interesse come medie *temporali* sul processo che convergeranno alle medie su π .

Sebbene anche questo compito potrebbe apparire arduo, in realtà si rivela, almeno a livello algoritmico, abbastanza agevole. Tale praticabilità si fonda infatti sull'uso di una particolare classe di processi stocastici noti sotto il nome di catene di Markov¹.

Nel seguito del capitolo introdurremo prima i concetti ed i risultati necessari della teoria di Markov² e quindi spiegheremo come tali risultati possano venire applicati alla simulazione di sistemi fisici attraverso il "Markov Chain Monte Carlo".

1.1 CATENE DI MARKOV

La teoria di Markov può essere in generale ambientata in spazi continui e non limitati. Per le applicazioni di nostro interesse, tale approccio porterebbe solamente ad un appesantimento della notazione e ad un oscuramento dei collegamenti coi concetti "fisici", senza al contempo guadagnare nei contenuti. Per tale motivo la nostra esposizione si occuperà qui solamente di processi su tempi discreti con spazio degli stati \mathcal{S} discreto e finito³. Se non specificheremo altrimenti, si intenderà che il numero di elementi di \mathcal{S} è n . Assumeremo inoltre che tali processi siano temporalmente omogenei⁴.

È d'uso nella teoria delle probabilità la notazione (p_1, \dots, p_n) per i vettori colonna e $[p_1, \dots, p_n] = (p_1, \dots, p_n)^T$ per i vettori riga. Ci atterremo a tale notazione.

¹Markov, brillante matematico russo vissuto a cavallo dei secoli XIX e XX, ha sviluppato le basi di tale teoria solamente in dal punto di vista matematico in rapporto con la teoria delle probabilità. Nessuna applicazione alla fisica o ad altre scienze sperimentali vennero da lui proposte, se non quella di utilizzare i processi — chiamati in seguito *processi di Markov* — per ottenere in maniera automatica testi con struttura lessicale e sintattica quasi identica a quella di testi aventi significato semantico. Tale applicazione venne in seguito ripresa anche da Claude Shannon, padre della moderna teoria della comunicazione.

² Con questo capitolo non ci poniamo pretese di esaustività, né di completezza. Per un'analisi più dettagliata rimandiamo a [29] per le catene di Markov, [45, 52, 17, 48, 12] per il coupling a [6, 7, 58, 51] per il MCMC.

³Si dimostra che, se n è il numero dei suoi elementi, \mathcal{S} è isomorfo a $\{1, 2, \dots, n\} \subset \mathbb{Z}$. In tal caso definiremo $|\mathcal{S}| \equiv n$.

⁴Per una definizione di omogeneità si vede più avanti

Processi Stocastici

Un vettore \mathbf{p} è detto *stocastico* se ogni suo elemento è non negativo — $p_i \geq 0$, $\forall i$ — mentre la loro somma è pari all'unità — $\sum_i p_i = 1$ —. Tale vettore può essere utilizzato per definire una distribuzione di probabilità su di uno spazio degli stati \mathcal{S} finito, in cui la probabilità dello stato s_i è data dall' i -esima componente p_i di \mathbf{p} .

Similmente definiamo *matrice stocastica* una matrice di cui ogni vettore riga sia un vettore stocastico. Essa sarà inoltre detta non negativa, positiva o strettamente positiva a seconda che i suoi elementi siano rispettivamente tutti non negativi, non negativi e non tutti nulli, tutti positivi. Se una matrice stocastica possiede un autovettore sinistro (destra) con autovalore 1, diciamo che tale vettore riga (colonna) è un vettore riga (colonna) fisso.

Definiamo come *processo stocastico* a valori in uno spazio degli stati \mathcal{S} una sequenza di variabili aleatorie $\{X_0, X_1, \dots\}$ a valori in \mathcal{S} con spazio di probabilità Ω .

Fisicamente i vari stati di un processo stocastico possono essere interpretati come configurazioni successive di un sistema fisico. Successive sia in senso spaziale, come nel caso dei vari monomeri di un polimero, sia in senso temporale, intendendoli come rappresentanti l'evoluzione temporale di un medesimo sistema. Sarà quest'ultimo il caso cui si farà riferimento implicitamente lungo tutto il capitolo.

Iniziamo dunque col dire che un processo stocastico è (temporalmente) *omogeneo*, se la probabilità di transizione tra due stati non dipende dal tempo.

Proprietà di Markov: Un processo stocastico soddisfa la proprietà di Markov, se $\forall t \geq 0, \forall s_0, \dots, s_t \in \mathcal{S}$

$$\mathcal{P}(X_{t+1} | X_0 = s_0, \dots, X_{t-1} = s_{t-1}, X_t = s_t) = \mathcal{P}(X_{t+1} | X_t = s_t)$$

Questa proprietà viene sovente esposta con dicendo che in un processo che la soddisfa “il futuro dipende dal passato solo attraverso il presente”. Infatti afferma che la probabilità di passare da uno stato ad un altro dipende solamente dallo stato in cui il processo si trova e non da come vi è giunto. Per il fatto di non dipendere esplicitamente dalla storia del processo, tale proprietà viene anche detta “proprietà della mancanza di memoria” (memoryless property).

Catene di Markov

DEFINIZIONE 1.1.

Una catena di Markov (MC, Markov Chain), è un processo stocastico su di uno spazio degli stati $\mathcal{S} = \{s_1, s_2, \dots, s_n\}$ che soddisfa la proprietà di Markov. Ad ogni passo il processo si trova in uno stato s_i e si sposterà in un altro stato s_j con probabilità di transizione $p_{ij} = \mathcal{P}(X_{t+1} = s_j | X_t = s_i)$.

Un processo markoviano è completamente determinato per mezzo del vettore (stocastico) di distribuzione iniziale \mathbf{p}^0 e della matrice (stocastica) di transizione \mathbf{P} così definiti:

$$\mathbf{p}^0 = [p_1^0, \dots, p_n^0], \quad \text{dove } p_i^0 = \mathcal{P}(X_0 = s_i)$$

$$\mathbf{P} = \{p_{ij}\}, \quad \text{dove } p_{ij} = \mathcal{P}(X_t = s_j | X_{t-1} = s_i)$$

Vale infatti il seguente

TEOREMA 1.1.

Sia una MC con matrice di transizione \mathbf{P} e distribuzione iniziale \mathbf{p}^0 . Allora la distribuzione $\mathbf{p}^{(t)}$ al tempo $t \geq 0$ è data da

$$\mathbf{p}^{(t)} = \mathbf{p}^0 \mathbf{P}^t$$

Il teorema si può anche enunciare dicendo che la MC soddisfa le equazioni di Chapman-Kolmogorov

$$\mathcal{P}(X_{t+t'} = s_j | X_0 = s_i) = \sum_{l \in \mathcal{S}} \mathcal{P}_{ij} \mathcal{P}_{il}$$

$$\mathbf{p}^{(n+m)} = \mathbf{p}^{(n)} \mathbf{P}^{(m)} \quad \text{e, per induzione } \mathbf{P}^{(n)} = \mathbf{P}^n. \quad (1.1)$$

Per riferirci ad una catena di Markov, ci riferiremo alla terna $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ composta dallo spazio degli stati \mathcal{S} , dalla matrice di transizione \mathbf{P} e dalla distribuzione iniziale \mathbf{p}^0 .

Definizioni e Prime Proprietà

DEFINIZIONE 1.2.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC e $s_i \in \mathcal{S}$.

I tempi di ricorrenza dello stato s_i sono $N_i := \{t \geq 1 \mid \mathbf{P}_{ii}^t > 0\}$.

Il periodo $d(s_i)$ dello stato s_i è il massimo comun divisore dei tempi di ricorrenza.

Se $d(s_i) = 1$, lo stato è detto aperiodico.

Una MC è detta aperiodica se lo sono tutti i suoi stati. Altrimenti è detta periodica.

DEFINIZIONE 1.3.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC e $s_i, s_j \in \mathcal{S}$;
 s_j è raggiungibile da s_i , ($s_i \rightarrow s_j$), se $\exists t \geq 0$ t.c. $P_{ij}^t > 0$.
 s_i e s_j comunicano, $s_i \leftrightarrow s_j$, se $s_i \rightarrow s_j$ e $s_i \leftarrow s_j$.

DEFINIZIONE 1.4.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC
 Un insieme di stati $C \subseteq \mathcal{S}$, $C \neq \emptyset$ è invariante (chiuso) se $p_{ij} = 0 \forall s_i \in C, s_j \in C^c$
 Un singolo stato chiuso è detto assorbente (i.e. $p_{ii} = 1$).
 Una MC è detta assorbente se possiede almeno uno stato assorbente s_j e se tale stato è raggiungibile da ogni altro, $s_i \rightarrow s_j \quad \forall s_i \in \mathcal{S}$.

DEFINIZIONE 1.5.

Stati che non appartengono a nessun insieme invariante sono detti transienti

LEMMA 1.2.

Ogni insieme invariante contiene un sottinsieme minimale. La relazione \leftrightarrow è una relazione d'equivalenza su \mathcal{S} . Tutti gli insiemi invarianti minimali sono classi d'equivalenza rispetto a tale relazione.

TEOREMA 1.3.

$$s_i \leftrightarrow s_j \Rightarrow d(s_i) = d(s_j)$$

Gli elementi di una classe d'equivalenza possiedono lo stesso periodo.

DEFINIZIONE 1.6 (Definizioni equivalenti di irriducibilità).

Una MC $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ è detta irriducibile se:

1. $\forall s_i, s_j \in \mathcal{S} \quad \exists t \geq 0$ t.c. $P_{ij}^t > 0$.
2. $s_i \longleftrightarrow s_j \quad \forall s_i, s_j \in \mathcal{S}$.
3. \mathcal{S} è l'unico insieme chiuso minimale di stati.

Dalla terza definizione e dal teorema 1.3 segue immediatamente che tutti gli stati di una MC irriducibile possiedono lo stesso periodo.

Tecnicamente un processo irriducibile è tale che da uno stato si può raggiungere, prima o poi, ogni altro stato. Dunque può muoversi liberamente in tutto lo spazio degli stati. Dato che le catene considerate sono positive, comunque l'irriducibilità implica la ricorrenza, ovvero il processo passerà infinite volte da ciascuno stato. I casi in cui questo non avviene sono quelli

1. MARKOV CHAIN MONTE CARLO

in cui la catena di Markov presenta degli stati transienti e degli insiemi invarianti minimali (o degli stati assorbenti). In questo caso infatti, esistono degli stati in cui si può “entrare ma non uscire”. Il processo rimane quindi intrappolato in una frazione finita di stati e non raggiungerà mai più quelli transienti.

DEFINIZIONE 1.7.

Una MC $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ è detta regolare se

$$\forall s_i, s_j \in \mathcal{S} \quad \exists n_0 \text{ t.c. } p_{ij}^{(t)} \geq 0 \quad \forall t > t_0.$$

TEOREMA 1.4.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC aperiodica.

Allora $\exists t_0 < \infty$ t.c. $p_{ii}^{(t)} > 0$, $\forall s_i \in \mathcal{S}, \forall t > t_0$.

COROLLARIO 1.5.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC irriducibile ed aperiodica

Allora $\exists t_0 < \infty$ t.c. $p_{ij}^{(t)} > 0$, $\forall s_i, s_j \in \mathcal{S}, \forall t > t_0$.

Il corollario afferma che ogni MC aperiodica ed irriducibile è regolare. Bisogna comunque prestare attenzione alle proprietà soddisfatte, in quanto una MC regolare è irriducibile, mentre in generale una MC irriducibile non è regolare. Ogni qualvolta scriveremo “regolare”, intenderemo una MC irriducibile ed aperiodica.

DEFINIZIONE 1.8.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC e $s_i, s_j \in \mathcal{S}$

first hit probability per $s_i \rightarrow s_j, t > 0$:

$$f_{ij}^{(t)} := \mathcal{P}(X_1 \neq s_j, \dots, X_{t-1} \neq s_j, X_t = s_j | X_0 = s_i).$$

hitting probability per $s_i \rightarrow s_j$: $f_{ij}^* \equiv \sum_{t>0} f_{ij}^t$.

$$\text{expected hitting time per } s_i \rightarrow s_j: \mu_{ij} = \begin{cases} \sum_{t>0} t f_{ij}^t & \text{se } f_{ij}^* = 1 \\ \infty & \text{se } f_{ij}^* < 1 \end{cases}.$$

expected return time per s_i : $\mu_i := \mu_{ii}$.

DEFINIZIONE 1.9.

Uno stato $s_i \in \mathcal{S}$ è detto ricorrente se $f_{ii}^* = 1$. Altrimenti è detto transiente.

Forma Canonica di \mathbf{P}

Per ogni MC $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ esiste la possibilità di rinumerare gli stati per ottenere una matrice di transizione in cui le transizioni all'interno di una medesima

classe d'equivalenza risultino un blocco diagonale, mentre le transizioni tra classi distinte risultino blocchi fuori dalla diagonale principale. Una matrice di transizione \mathbf{P} si dice essere in forma canonica:

$$\mathbf{P} = \left(\begin{array}{ccc|c} \mathbf{I}_1 & & & \mathbf{0} \\ & \ddots & & \\ & & \mathbf{I}_r & \\ \hline & & \mathbf{R} & \mathbf{Q} \end{array} \right).$$

Dove le \mathbf{I}_i sono matrici quadrate rappresentanti le transizioni all'interno degli insiemi minimali (invarianti); mentre \mathbf{Q} fornisce quelle all'interno degli stati transienti. Infine \mathbf{R} contiene le transizioni dagli stati transienti agli stati ricorrenti.

Dato che la matrice \mathbf{Q} deve essere substocastica, cioè deve avere almeno una somma sulle righe minore di 1, allora il vettore distribuzione stazionaria per tale matrice deve essere nella forma

$$\boldsymbol{\pi} = [\pi_1, \dots, \pi_r, 0, \dots, 0].$$

Se non siamo interessati alle transizioni interne degli insiemi invarianti minimali \mathcal{I}_i ma solo all'evoluzione del sistema verso tali insiemi, possiamo rappresentare ogni insieme con un singolo stato assorbente, le cui probabilità di transizione sono date dalla somma delle probabilità sugli stati dell'insieme. Assumiamo per semplicità che esista un unico insieme minimale invariante \mathcal{J} e siano i gli stati invarianti e t un qualsiasi stato transiente. Chiamando s lo stato assorbente che vogliamo costruire, si ha:

$$\mathcal{P}(X_{t+1} = s | X_t = t) = \sum_{i \in \mathcal{J}} \mathcal{P}(X_{t+1} = i | X_t = t).$$

In termini della matrice di transizione, ciò equivale a sommare sulle righe e sulle colonne indicizzate dagli stati che si vogliono contrarre. Per definizione di stato assorbente, si ha che $p_{ii} = 1$. Ciò è naturalmente verificato, dato che a tale termine corrisponderà la somma sui termini interni all'insieme invariante che dà, per l'appunto, l'unità.

Questo ci serve per poter usare i risultati disponibili per processi markoviani con stati assorbenti, ricordandoci che valgono anche sostituendo ad uno stato assorbente un insieme minimale invariante.

Sia una MC assorbente $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ con r stati assorbenti e t transienti. Posso rinumerare gli stati per avere prima quelli transienti ottenendo la matrice di

1. MARKOV CHAIN MONTE CARLO

transizione \mathbf{P} in forma canonica:

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbb{I}_r \end{array} \right)$$

dove \mathbb{I}_r è la matrice identità $r \times r$, $\mathbf{0}$ la matrice nulla $r \times t$, \mathbf{Q} una matrice $t \times t$ rappresentante le transizioni tra elementi transienti e \mathbf{R} una matrice $t \times r$.

Effettuando la potenza, ottengo

$$\mathbf{P}^n = \left(\begin{array}{c|c} \mathbf{Q}^n & * \\ \hline \mathbf{0} & \mathbb{I}_r \end{array} \right).$$

La forma esatta di \mathbf{Q} qui non ci interessa, ma si può dimostrare che

$$\mathbf{Q}^n \xrightarrow{n \rightarrow \infty} \mathbf{0}.$$

TEOREMA 1.6.

Data una MC assorbente $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ la probabilità che il processo sia assorbito è 1.

Risultati sulla Probabilità di Assorbimento

TEOREMA 1.7.

Sia \mathbf{P} la matrice di transizione in forma canonica di una MC assorbente.

Allora $\mathbb{I} - \mathbf{Q}$ ammette inversa \mathbf{N} ed $\mathbf{N} = \mathbb{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots$.

$n_{ij} = \{\mathbf{N}\}_{ij}$ fornisce il numero medio di volte che la catena, partita da s_i , passa da s_j . Ovviamente sia s_i che s_j sono stati transienti.

DEFINIZIONE 1.10.

Sia \mathbf{P} matrice di transizione in forma canonica di una MC assorbente. La matrice $\mathbf{N} = (\mathbb{I} - \mathbf{Q})^{-1}$ è la matrice fondamentale per \mathbf{P} .

TEOREMA 1.8.

Sia t_i il numero medio di passi che compie una MC assorbente $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ partita da s_i prima di essere assorbita. Sia quindi $\mathbf{t} = (t_1, \dots, t_n)$.

Allora $\mathbf{t} = \mathbf{N}\mathbf{c}$, dove \mathbf{c} è un vettore di costante 1.

TEOREMA 1.9.

Sia b_{ij} la probabilità che ha una MC assorbente, partita da s_i , di essere assorbita in uno stato assorbente s_j . Sia quindi $\mathbf{B} = \{b_{ij}\}$ matrice $t \times r$.

Allora $\mathbf{B} = \mathbf{N}\mathbf{R}$, dove \mathbf{N} è la matrice fondamentale ed \mathbf{R} la forma canonica della matrice di transizione.

In generale se un processo markoviano possiede degli stati assorbenti e/o degli insiemi invarianti, non esiste un'unica distribuzione stazionaria. In presenza di stati assorbenti, per $t \rightarrow \infty$, il processo entrerà in uno di questi stati e vi rimarrà indefinitivamente. Un po' più complessa è la situazione nel caso vi fossero degli insiemi invarianti. Similmente al caso degli stati assorbenti, la catena rimarrà "intrappolata" all'interno di uno di tali insiemi, con una probabilità relativa che va ricercata nella struttura della matrice \mathbf{R} . Una volta lì giunta, il comportamento della catena può essere studiato prendendo in considerazione solamente la matrice di transizione \mathbf{I} interna a tale insieme. Appena più semplice la situazione in cui esiste un unico insieme invariante minimale, in quanto certamente il processo, prima o poi, vi giungerà. In tal caso è sufficiente scartare la matrice di transizione globale e studiare solamente \mathbf{I} , se non si è interessati al comportamento transiente.

Distribuzioni Stazionarie e Convergenza

Consideriamo una MC $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ qualunque ed assumiamo che esista una distribuzione limite $\boldsymbol{\pi}$, cioè una distribuzione tale che

$$\lim_{t \rightarrow \infty} \mathbf{p}^{(t)} = \boldsymbol{\pi}. \quad (1.2)$$

Da ciò segue che

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$$

e $\boldsymbol{\pi}$ è un autovettore sinistro di \mathbf{P} con autovalore 1 che può essere calcolato risolvendo tale equazione. Le soluzioni così trovate sono dette *distribuzioni stazionarie*. Il termine "stazionaria" deriva dal fatto che, prendendo $\boldsymbol{\pi}$ come probabilità iniziale, si ottiene per ogni t , $\boldsymbol{\pi}\mathbf{P}^t = \boldsymbol{\pi}$.

Però ciò significa anche che per generare stati distribuiti secondo $\boldsymbol{\pi}$, dobbiamo assumere proprio $\boldsymbol{\pi}$ come distribuzione iniziale, contro le nostre aspettative di poterlo raggiungere partendo da qualsiasi distribuzione iniziale, per esempio assumendo equiprobabili tutti gli stati. Fortunatamente, sotto opportune ipotesi, si riesce a provare la convergenza di una MC a $\boldsymbol{\pi}$ anche nel caso si abbia $\mathbf{p}^0 \neq \boldsymbol{\pi}$.

Ogni catena di Markov irriducibile ed aperiodica su uno spazio degli stati finito ammette sempre almeno una distribuzione stazionaria, ma non è garantito che essa sia unica o, se lo è, che il processo vi converga per $t \rightarrow \infty$.

1. MARKOV CHAIN MONTE CARLO

Fortunatamente si può dimostrare che, sotto la condizione di regolarità, ogni MC ammette un'unica distribuzione stazionaria π a cui converge per $t \rightarrow \infty$ ¹.

TEOREMA 1.10.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC regolare.

Allora, per $t \rightarrow \infty$, \mathbf{P}^t tende ad una matrice limite Π in cui tutte le righe sono identiche ed uguali ad un vettore di probabilità strettamente positivo π . ($\pi_i \geq 0 \forall s_i, s_j \in \mathcal{S}$, $\sum_i \pi_i = 1$).

TEOREMA 1.11.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC regolare, $\Pi = \lim_{t \rightarrow \infty} \mathbf{P}^t$ con riga comune π e \mathbf{c} un vettore colonna di costante 1. Allora

$$\pi \mathbf{P} = \pi.$$

Inoltre ogni vettore riga \mathbf{v} t.c. $\mathbf{v} \mathbf{P} = \mathbf{v}$ è un multiplo di π .

$$\mathbf{P} \mathbf{c} = \mathbf{c}.$$

Inoltre ogni vettore colonna \mathbf{x} t.c. $\mathbf{P} \mathbf{x} = \mathbf{x}$ è un multiplo di \mathbf{c} .

In pratica si afferma che π è l'unico vettore riga fisso che sia anche un vettore stocastico. Da qui l'unicità della distribuzione stazionaria.

TEOREMA 1.12.

Sia $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ una MC regolare e \mathbf{p}^0 un vettore di probabilità qualsiasi.

Allora $\lim_{t \rightarrow \infty} \mathbf{p}^0 \mathbf{P}^t = \pi$.

Reversibilità ed Equilibrio

Dato un processo stocastico con matrice di transizione \mathbf{P} , la matrice che rappresenta il processo con inversione temporale è \mathbf{P}^* , data da

$$\{\mathbf{P}^*\}_{ij} = \mathcal{P}(X_m = s_j | X_{m+1} = s_i) = \frac{\pi_j \{\mathbf{P}\}_{ij}}{\pi_i}.$$

DEFINIZIONE 1.11.

Un processo è detto reversibile se $\mathbf{P}^* = \mathbf{P}$.

Si arriva così a stabilire il seguente fondamentale

TEOREMA 1.13.

Una MC irriducibile è reversibile $\Leftrightarrow \mathbf{P}^* \equiv \mathbf{P}$

¹Per convergenza qui e oltre intenderemo convergenza della distribuzione $\mathcal{L}^{(t)}(X)$ al tempo t degli stati della MC nella *total variation distance* $\Delta_V^t(t) = \max_{A \subseteq \mathcal{S}} |\mathcal{L}^{(t)}(X) - \pi| = \frac{1}{2} \sum_{i \in \mathcal{S}} |\mathcal{L}_i^{(t)}(X) - \pi_i|$. Spesso tale distanza, su spazi continui, è indicata con $\|\mu - \pi\|$.

Un modo equivalente di definire la reversibilità riguarda la distribuzione stazionaria π :

DEFINIZIONE 1.12.

Una distribuzione di probabilità π su \mathcal{S} è reversibile per una MC $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$ se, $\forall s_i, s_j \in \mathcal{S}$,

$$\pi_i P_{ij} = \pi_j P_{ji}. \quad (1.3)$$

Tale definizione riveste particolare importanza, in quanto vale il seguente teorema:

TEOREMA 1.14.

Se una distribuzione π è reversibile per una MC, allora è anche una distribuzione stazionaria per tale processo.

Dunque il bilancio dettagliato è una condizione sufficiente per la stazionarietà di una distribuzione.

Distribuzioni stazionarie e reversibili vengono dette *distribuzioni d'equilibrio*.

Rileggendo in senso inverso quanto enunciato, si ottiene che, costruendo una catena di Markov irriducibile che renda reversibile la distribuzione d'interesse π , allora con certezza la catena convergerà a tale distribuzione.

Dal punto di vista fisico, la reversibilità di un sistema la cui dinamica microscopica è (simulata da) un processo markoviano è strettamente legata alle simmetrie temporali.

L'evoluzione temporale di tale sistema può essere definita per mezzo di una *Master Equation*¹

$$\frac{\partial}{\partial t} \mathcal{P}(C, t | C_0, t_0) = \sum_{C' \in \mathcal{C}} \{W[C' \rightarrow C] \mathcal{P}(C', t | C_0, t_0) - W[C \rightarrow C'] \mathcal{P}(C, t | C_0, t_0)\}; \quad (1.4)$$

dove, dato uno spazio delle configurazioni \mathcal{C} , $\mathcal{P}(C, t | C_0, t_0)$ rappresenta la probabilità del sistema di trovarsi nello stato C al tempo t se al tempo t_0 si trovava nello stato C_0 . $W[C' \rightarrow C]$ rappresenta invece la frequenza di transizione dallo stato C' allo stato C .

Consideriamo ora una traiettoria $\{C_t\}_{t \in [t_1, t_2]}$ nello spazio delle configurazioni \mathcal{C} , che ha una data probabilità di venire realizzata durante l'evoluzione

¹Per qualche istante passiamo a sistemi temporalmente continui, per meglio sottolineare il collegamento con la fisica.

stocastica dello stato stazionario. Se la simmetria per inversione temporale è valida, ci aspettiamo che la traiettoria inversa $\{C_t^*\} \equiv \{C_t\}_{-t+t_1+t_2}$ abbia la stessa probabilità di venire realizzata.

Questo a sua volta implica che

$$W[C' \rightarrow C]P^s(C') = W[C \rightarrow C']P^s(C), \quad \forall C, C' \in \mathcal{C}. \quad (1.5)$$

L'identità precedente è la condizione di *bilancio dettagliato* per la misura stazionaria P^s . Ritornando alla notazione precedente, vediamo la chiara identità tra le eqq. (1.5) e (2.3).

Mentre l'eq. (1.5) richiede la conoscenza della misura stazionaria P^s , esiste un'altra condizione equivalente che necessita della sola conoscenza delle probabilità di transizione. Infatti un processo soddisfa la condizione di bilancio dettagliato, se

$$\begin{aligned} W[C_1 \rightarrow C_2]W[C_2 \rightarrow C_3] \dots W[C_n \rightarrow C_1] = \\ W[C_1 \rightarrow C_n]W[C_n \rightarrow C_{n-1}] \dots W[C_2 \rightarrow C_1], \end{aligned} \quad (1.6)$$

per ogni sequenza $\{C_1 \dots C_n\} \in \mathcal{C}$.

In particolare questa seconda condizione fornisce una condizione per verificare la *non* soddisfazione della reversibilità temporale. A tal fine è sufficiente infatti trovare una sequenza di stati che non soddisfa (1.6).

Velocità di Convergenza

Sia (P, p^0, \mathcal{S}) una MC regolare e studiamone la convergenza ad una distribuzione π che può essere o meno quella stazionaria per tale catena.

In generale, siano π e ρ due distribuzioni di probabilità.

Relative pointwise distance:

$$d_{rp}^U(\rho, \pi) = \max_{j \in U} \frac{|\rho_j - \pi_j|}{\pi_j},$$

con $U \subseteq \mathcal{S}$.

Definiamo quindi la *convergence rate function* come

$$\Delta^U(t) = \max_{i \in U} d_{rp}^U(\pi^{(i,t)}, \pi) = \max_{j \in U} \frac{|p_{ij}^{(t)} - \pi_j|}{\pi_j} \quad \text{dove } \pi_j^{(i,t)} = p_{ij}^{(t)}.$$

Considerando la convergenza sull'intero spazio degli stati finalmente otteniamo

$$\Delta(t) = \Delta^{\mathcal{S}}(t).$$

DEFINIZIONE 1.13.

Definiamo il mixing time di una MC regolare come

$$\tau(\epsilon) = \min \{t | \Delta(t') \leq \epsilon, \forall t' \geq t\}.$$

DEFINIZIONE 1.14.

A volte una MC è determinata da qualche input x , nel qual caso scriviamo $\Delta_x(t)$ e $\tau_x(\epsilon)$.

Una (famiglia di) catene di Markov è rapidly mixing se

$$\tau_x(\epsilon) = \text{poly}(|x|, \log \frac{1}{\epsilon}).$$

1.2 MARKOV CHAIN MONTE CARLO

Introduzione

Assumiamo di dover generare campioni in accordo ad una distribuzione di probabilità π che sia troppo complicata da poter essere descritta esplicitamente (o addirittura sconosciuta). Una soluzione intelligente potrebbe essere costruire una catena di Markov che abbia π come propria distribuzione stazionaria, far evolvere per un certo periodo di tempo tale catena e quindi campionare gli stati generati dalla catena.

Markov Chain Monte Carlo è il nome di una classe di algoritmi iterativi utilizzati proprio quando il campionamento diretto dalla distribuzione di interesse non è praticabile. In sua vece viene fatta evolvere una MC opportuna con distribuzione stazionaria π . Partendo da uno stato iniziale arbitrario, l'evoluzione è protratta sino a quando si reputa di essere vicini all'equilibrio e lo stato finale è considerato distribuito come π . Le difficoltà più comuni vengono dal fatto che l'utilizzatore deve determinare quanto a lungo far evolvere la MCMC. In ogni caso, tenendo conto del fatto che la catena si fermerà dopo un tempo finito e che non è fatta partire con distribuzione π , il metodo produrrà *sempre* un campione approssimato.

Finora abbiamo mantenuto il discorso sulle proprietà matematiche delle catene di Markov, cui ci siamo riferiti con la terna $(\mathbf{P}, \mathbf{p}^0, \mathcal{S})$. Ora però il nostro interesse si sposta dalla probabilità che il processo si trovi in un certo stato, all'effettivo stato in cui la catena si trova istante per istante, Alla sua evoluzione temporale. Per tale motivo cambiando notazione da qui in avanti ci riferiremo esplicitamente ad una MC come ad una sequenza di variabili aleatorie $\{X_i\}_{i \in \mathcal{J}} = \{X_0, X_1, \dots\}$ lasciando implicita la dipendenza della

sequenza dai parametri considerati in precedenza. Etichetteremo genericamente con \mathcal{J} l'indice degli stati in $\{X_i\}_{\mathcal{J}}$ specificando eventualmente se tale indice è finito o infinito solamente dove necessario.

Realizzazione di una Markov Chain Monte Carlo

Le transizioni di una MC $\{X_i\}_{\mathcal{J}}$ possono essere descritte tramite una *funzione di update* ϕ che, a partire dallo stato attuale dalla catena e con l'ausilio di un numero casuale, fornisce lo stato successivo della catena in accordo con le probabilità di transizione date dalla matrice \mathbf{P} .

Formalmente,

DEFINIZIONE 1.15.

Una funzione di update (*deterministica*) è una funzione

$$\phi : \mathcal{S} \times [0, 1]^{\mathcal{L}} \rightarrow \mathcal{S} \quad (1.7)$$

$$(s_t, \mathbf{u}_{t+1}) \mapsto s_{t+1}, \quad (1.8)$$

dove $\{\mathbf{u}_i\}_{\mathcal{J}}$ è una sequenza di vettori di numeri casuali distribuiti uniformemente nell'intervallo $[0, 1]$. La precisa forma del vettore di lunghezza \mathcal{L} è determinata dal particolare algoritmo utilizzato.

In generale quindi un algoritmo MCMC genera configurazioni successive di un processo stocastico secondo la legge $X_{t+1} = \phi(X_t, \mathbf{u}_{t+1})$. I vari algoritmi MCMC variano tra di loro per la distribuzione obiettivo e per la particolare forma della funzione di update.

Metropolis MCMC

Come esempio di algoritmo MCMC introduciamo quello ideato da Metropolis e formalizzato da Hastings [46, 5]. Scegliamo tale algoritmo in quanto sarà utilizzato nel seguito.

Si assuma che nello spazio degli stati \mathcal{S} sia definita una nozione di *vicinanza* (nel caso di sistemi su reticolo, tale nozione è quella indotta dalla metrica del reticolo) e denotiamo con $N(i)$ l'insieme dei primi vicini dello stato s_i .

L'algoritmo Metropolis per la distribuzione π sullo spazio degli stati \mathcal{S} opera nel modo seguente:

1. si scelga uno stato iniziale X_0 uniformemente estratto da \mathcal{S} ;
2. dato lo stato $X_t = s_i$ per $t \geq 0$, lo stato X_{t+1} è ottenuto come segue:

- a) si scelga $s_j \in N(i)$ a caso con distribuzione uniforme;
 b) si accetti $X_{t+1} = s_j$ con probabilità $\min\{1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\}$;
 Altrimenti si lasci $X_{t+1} = s_i$.

Qui $q_{ij} = \{Q\}_{ij}$ è la matrice di transizione di un'arbitraria MC.

La matrice di transizione può essere scritta come

$$\{P\}_{ij} = \begin{cases} \frac{1}{q_{ji}} \min\{1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\}, & \text{se } s_j \in N(i); \\ 0, & \text{se } s_j \notin N(i), s_j \neq s_i; \\ 1 - \sum_{s_j \in N(i)} P_{ij}, & \text{se } s_j = s_i. \end{cases}$$

Per mostrare che π è stazionaria per questa MC, è sufficiente dimostrare che soddisfa la condizione di bilancio dettagliato:

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall s_i, s_j \in \mathcal{S}.$$

Se $s_j = s_i$ oppure se $s_j \notin N(i)$ la questione è banale. Il caso $s_j \in N(i)$ si divide in due parti:

$$\frac{\pi_j q_{ji}}{\pi_i q_{ij}} \geq 1 \quad \begin{cases} \pi_i p_{ij} = \pi_i \frac{1}{q_{ji}} 1 & = \frac{\pi_i}{q_{ji}} \\ \pi_j p_{ji} = \pi_j \frac{1}{q_{ij}} \frac{\pi_i q_{ij}}{\pi_j q_{ji}} & = \frac{\pi_i}{q_{ji}} \end{cases}$$

$$\frac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1 \quad \begin{cases} \pi_i p_{ij} = \pi_i \frac{1}{q_{ji}} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} & = \frac{\pi_j}{q_{ij}} \\ \pi_j p_{ji} = \pi_j \frac{1}{q_{ij}} 1 & = \frac{\pi_j}{q_{ij}} \end{cases}$$

Inoltre la catena è aperiodica se esiste almeno un $s_i \in \mathcal{S}$ per cui $\frac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1$.
 Ovvero se esiste almeno uno stato s_i per cui $p_{ii} > 0$.

La funzione di update $\phi(X_t, \mathbf{U}_{t+1})$ dell'algoritmo è implementata nel passo 2. In realtà tale funzione potrebbe essere a sua volta divisa in due sottofunzioni: una *funzione di proposta* che sceglie lo stato finale, ed una *funzione di scambio* che controlla se lo stato sarà effettivamente raggiunto. Nell'algoritmo Metropolis il vettore \mathbf{U}_t è composto da 2 numeri casuali.

Convergenze di Medie

Per quanto riguarda il valore di aspettazione di una qualche osservabile del sistema fisico simulato dalla MC, valgono i seguenti teoremi.

TEOREMA 1.15 (Convergenza delle somme).

Se la MC $\{X_i\}_j$ è ergodica (positiva, ricorrente, aperiodica) con distribuzione stazionaria π , allora per ogni stato iniziale

$$\frac{1}{n} \sum_{i=1}^n h(X_i) \xrightarrow{n \rightarrow \infty} \langle h(X) \rangle_{\pi},$$

se $\langle |h(X)| \rangle_{\pi} < \infty$

TEOREMA 1.16 (Teorema del Limite Centrale per MC Reversibili).

se la MC $\{X_i\}_j$ è irriducibile, aperiodica e reversibile con distribuzione stazionaria π allora, per ogni stato iniziale, per $n \rightarrow \infty$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n [h(x_i) - \langle h(X) \rangle_{\pi}] \rightarrow \mathcal{N}(0, \sigma^2),$$

se $0 < \sigma^2 \equiv \text{Var}(h(X_0)) + \sum_{i=1}^{\infty} \text{Cov}_{\pi}(h(X_0), h(X_i)) < \infty$

1.3 COUPLING

Nonostante i numerosi risultati teorici su cui sono basati gli algoritmi MCMC rimangono comunque numerose le difficoltà da affrontare per assicurare che i dati in possesso siano fedeli al sistema in esame. Tra le difficoltà maggiori ci sono l'inabilità di assicurare che le configurazioni generate siano scorrelate e che esse siano rappresentative della distribuzione stazionaria. Una tra le soluzioni più comuni per risolvere il primo problema consiste nel lasciar trascorrere un certo intervallo tra due successivi campionamenti dei dati¹. Rimane comunque il problema di assicurare la fedeltà dei campioni alla distribuzione desiderata. È inoltre estremamente difficile anche solo decidere se tale distribuzione sia stata raggiunta.

Recentemente (Propp e Wilson, 1996 [52]) è stata introdotta una procedura che garantisce che il campione generato ad un preciso istante sia distribuito esattamente con la distribuzione cercata. Pur non utilizzando tale procedura nel nostro studio, essa verrà brevemente accennata nel seguito, in quanto è probabilmente l'applicazione di maggior rilievo che faccia uso di una particolare classe di MCMC di cui parleremo in questa sezione.

¹Ovviamente supportati da considerazioni analitiche.

Introduzione

In generale, l'accoppiamento (*coupling*) di due di due processi stocastici $\{X_t\}$ e $\{Y_t\}$ è un processo stocastico $Z_t = (X'_t, Y'_t)$ che possiede $\{X_t\}$ e $\{Y_t\}$ come distribuzioni marginali. Nel caso di MC finite, ciò significa

$$\mathcal{P}(X'_{t+1} = s_k | X'_t = s_i, Y'_t = s_j) = \mathcal{P}(X_{t+1} = s_k | X_t = s_i) = \{P^X\}_{ik}, \quad (1.9)$$

$$\mathcal{P}(Y'_{t+1} = s_l | X'_t = s_i, Y'_t = s_j) = \mathcal{P}(Y_{t+1} = s_l | Y_t = s_j) = \{P^Y\}_{jl}. \quad (1.10)$$

Riferiamoci ora a qualcosa di più specifico.

Coupling di MC Regolari

Siano $\mathcal{M} = \{X_i\}_I$ e $\mathcal{N} = \{Y_j\}_J$ due MC *regolari* sullo stesso spazio degli stati \mathcal{S} e con la stessa matrice di transizione \mathbf{P} . Per l'unicità della distribuzione stazionaria, le due catene posseggono la stessa distribuzione stazionaria $\boldsymbol{\pi}$.

Consideriamo ora la catena di Markov $\mathcal{M} \times \mathcal{N}$ con variabili aleatorie $Z_t = (X_t, Y_t)$. Per le probabilità di transizione si ha:

$$\begin{aligned} P_{ij}^Z &= \mathcal{P}(Z_t = (s_k, s_l) | Z_{t-1} = (s_i, s_j)) \\ &= \mathcal{P}(X_t = s_k | X_{t-1} = s_i) \cdot \mathcal{P}(Y_t = s_l | Y_{t-1} = s_j) \\ &= P_{ik} P_{jl}. \end{aligned}$$

Dato che \mathcal{M} e \mathcal{N} sono regolari con distribuzione stazionaria $\boldsymbol{\pi}$, anche $\mathcal{M} \times \mathcal{N}$ è regolare e possiede la distribuzione stazionaria $\boldsymbol{\pi}^Z = \boldsymbol{\pi}^T \boldsymbol{\pi}$ (i.e. $\pi_{ij}^Z = \pi_i \pi_j$).

Vogliamo far notare che $\mathcal{M} \times \mathcal{N}$, marginalizzata sul primo o sul secondo componente, porta alla realizzazione dello stesso processo stocastico:

$$\begin{aligned} \mathcal{P}(Z_t = (k, \cdot) | Z_0 = (k_0, l_0)) &= \mathcal{P}(X_t = k | X_0 = k_0) \\ &= P_{k_0 k}^{(t)}, \quad \text{indipendentemente da } l_0. \\ \mathcal{P}(Z_t = (\cdot, l) | Z_0 = (k_0, l_0)) &= \mathcal{P}(Y_t = l | Y_0 = l_0) \\ &= P_{l_0 l}^{(t)}, \quad \text{indipendentemente da } k_0. \end{aligned}$$

Definiamo ora una variabile aleatoria T che, per ogni realizzazione di $\mathcal{M} \times \mathcal{N}$ indichi il primo tempo in cui X_t e Y_t hanno lo stesso valore:

$$T \equiv \inf_{t \geq 0} \{X_t \equiv Y_t\}.$$

Tale variabile è detta *tempo d'accoppiamento*.

Un accoppiamento di catene di Markov è quindi un'evoluzione parallela dello stesso processo markoviano, con stati iniziali differenti e, nel caso di una procedura MCMC, identica funzione di update ma sequenze di numeri casuali differenti. Si può anche modificare $\mathcal{M} \times \mathcal{N}$ in modo da far sì che dopo l'accoppiamento, le due componenti non siano solo distribuite allo stesso modo, ma siano identiche. È sufficiente definire

$$X'_t = \begin{cases} X_t, & t < T; \\ Y_t, & t \geq T. \end{cases}$$

Denotiamo la catena così ottenuta con $\mathcal{M}|\mathcal{N}$.

Nel caso di catene accoppiate dipendenti il tempo d'accoppiamento viene detto *tempo di coalescenza*, dato che da quel momento in poi i processi rimangono identici. Ora le proiezioni di $\mathcal{M}|\mathcal{N}$ sulle componenti non sono più indipendenti ma, viste come marginalizzazioni, hanno le stesse identiche proprietà stocastiche.

In particolare, in una catena accoppiata $\mathcal{M}|\mathcal{N}$, scegliamo uno stato iniziale arbitrario $X_0 = k_0$ per \mathcal{M} e similmente $Y_0 = l_0$ per \mathcal{N} . Denotando le distribuzioni al tempo generico t come $p^{(t)} = (p_{k_0 k}^{(t)})_k$ e $q^{(t)} = (p_{l_0 l}^{(t)})_l$ si ha:

$$p^{(t)}(A) = \mathcal{P}(X_t \in A) \geq [\dots] \geq q^{(t)}(A) - \mathcal{P}(t < T),$$

da cui si legge che $q^{(t)}(A) - p^{(t)}(A) \leq \mathcal{P}(t < T)$. Similmente si dimostra che $p^{(t)}(A) - q^{(t)}(A) \leq \mathcal{P}(t < T)$ e dunque

$$|p^{(t)}(A) - q^{(t)}(A)| \leq \mathcal{P}(t < T) \quad \forall A \subseteq \mathcal{S}.$$

Che implica un limite inferiore alla velocità con cui le due catene convergono l'una verso l'altra.

$$d_V(p^{(t)}, q^{(t)}) = \sup_{A \subseteq \mathcal{S}} |p^{(t)}(A) - q^{(t)}(A)| \leq \mathcal{P}(t < T).$$

Ora considereremo il mixing time rispetto alla total variation distance, cioè

$$\tau(\epsilon) = \tau^V(\epsilon) = \min_t \{d_V(p^{(t,s)}, \pi) \leq \epsilon \quad \forall s \geq t, \forall \text{ stato iniziale}\}$$

TEOREMA 1.17 (Lemma d'accoppiamento).

Sia \mathcal{M} una MC finita e regolare; sia $Z_t = (X_t, Y_t)$ un coupling di due copie di \mathcal{M} . Supponiamo inoltre che $t : (0, 1] \rightarrow \mathbb{N}$ sia una funzione t.c. per ogni $\epsilon \in (0, 1]$, vale $\mathcal{P}(X_t \neq Y_t) \leq \epsilon$ per ogni $t \geq t(\epsilon)$, per qualunque scelta dello stato iniziale (X_0, Y_0) .

Allora il mixing time $\tau(\epsilon)$ di \mathcal{M} è limitato superiormente da $t(\epsilon)$.

MCMC Coupling

Cosa fare per rendere $\{X_i\}_J$ indipendente dal valore iniziale? Una possibile risposta è lavorare con catene di Markov accoppiate in modo dipendente.

Supponiamo che vi siano n stati in \mathcal{S} ed iniziamo simultaneamente una catena in ognuno di questi al tempo $t = 0$. Queste catene parallele possono essere accoppiate tramite l'uso di una medesima funzione di update ϕ e di un'identica sequenza di numeri casuali $\{\mathbf{U}_t\}$.

Due o più MC accoppiate *coalescono*¹ al tempo t se a tal tempo sono uguali. Una volta coalite, esse rimarranno uguali per ogni $t' > t$.

TEOREMA 1.18.

Siano $n = |\mathcal{S}|$ MC accoppiate $\{X^{(1)}, \dots, X^{(n)}\}$, dove

1. $\{X^{(j)}\}$ inizia nello stato s_j (ogni stato una catena);
2. l'updating è fatto in accordo a $X_{t+1}^{(j)} = \phi(X_t^{(j)}, \mathbf{U}_{t+1})$, e gli \mathbf{U}_i sono mutuamente indipendenti.

Allora

1. Il tempo T di coalescenza è una variabile casuale che dipende solo dalla sequenza $\mathbf{U}_1, \mathbf{U}_2, \dots$;
2. La variabile casuale X_t , il valore delle catene alla coalescenza, è indipendente da ogni valore iniziale.

La seconda conclusione del teorema 1.18 dice che T è un tempo per cui l'influenza dello stato iniziale è stata eliminata. Uno potrebbe sperare che X_T sia distribuita secondo π . Questa speranza è purtroppo infondata. Ciò che risulta vero è che, se T^* è un tempo fissato e X_{T^*} è indipendente da X_0 , allora $X_{T^*} \sim \pi$. Sfortunatamente T è un tempo casuale e, in generale, $X_{T^*} \not\sim \pi$.

Coupling From The Past (CFTP)

Diamo qui una rapida esposizione del metodo di cui abbiamo accennato all'inizio della sezione, che fa appunto uso del coupling tra catene di Markov per generare stati distribuiti esattamente con la distribuzione stazionaria di tale MC. Per tale motivo tale campionamento fa parte della (ristretta) categoria dei cosiddetti "perfect sampler".

¹Nella letteratura (inglese) per tale fenomeno vengono utilizzati il verbo *to coalesce* ed il sostantivo *coalescence*. Sebbene tali termini siano di derivazione latina, la forma verbale non è sopravvissuta nell'italiano attuale. Per sopperire abbiamo scelto di utilizzare una "volgarizzazione" del latino *coalesco*, *-is*, *coalui*, *coalitum*, *-ère*.

Tale metodo, dovuto a Propp e Wilson [52], viene chiamato metodo di Propp-Wilson o, in relazione alla sua caratteristica principale, “coupling-from-the-past” (CFTP). L’idea alla base di tale procedura è di far evolvere una MC con coupling dipendente *non* avanti nel tempo a partire dal tempo zero, bensì partendo da un tempo $-t$ per giungere in $t = 0$, usando lo stesso algoritmo normalmente impiegato. Che differenza sussiste tra evolvere da $-t$ a 0 e da 0 a t ? La conoscenza della sequenza di numeri casuali utilizzati.

Prendiamo due configurazioni qualsiasi al tempo $-t$ e facciamole evolvere fino al tempo 0, usando la stessa sequenza di numeri casuali per entrambe. In qualche tempo intermedio t_c tra le due configurazioni potrebbe avvenire la coalescenza e, da qui per ogni tempo futuro fino a 0, rimarranno uguali. Questa è la struttura di base dell’algoritmo. La sua applicazione completa è la seguente, esposta per semplicità nel caso di un modello di Ising su reticolo bidimensionale con un algoritmo MCMC qualsiasi. Prendiamo *tutte* le 2^N configurazioni e partiamo dal tempo $t = -1$, facciamole evolvere fino a $t = 0$ e controlliamo se sono tutte coalite. In caso affermativo, poniamo $t_{PW} = -1$, altrimenti proseguiamo spostando il punto di partenza sempre più indietro nel passato finché non giungiamo ad un tempo per cui nel presente ho coalescenza. L’importante è utilizzare sempre la stessa sequenza di numeri casuali per i tempi già provati. Ovvero se per evolvere da $t = -13$ a $t = 0$ ho usato la sequenza \mathbf{U}_{13} , per partire da $t = -14$ dovrò generare solo i numeri casuali necessari per *uno* step temporale, e poi riutilizzare \mathbf{U}_{13} . E così via. . . . La configurazione ottenuta quando tutte le 2^N configurazioni sono coalite, è garantita essere un “campionamento esatto” della distribuzione stazionaria.

Perché? Immaginiamo di far evolvere un ensemble infinito di configurazioni dal tempo $t = -\infty$ fino al tempo $t = -t_{PW}$ usando differenti sequenze di numeri casuali per ognuna di esse. Dato che la catena di Markov è supposta infinitamente lunga, allora la distribuzione al tempo $t = -t_{PW}$ è quella stazionaria ed ogni configurazione iniziale sarà evoluta in una delle 2^N configurazioni ivi presenti. Ma quale precisa configurazione ci sia al tempo $t = -t_{PW}$ non interessa, dato che sappiamo che al tempo zero tutte saranno coalite in un’unica configurazione. In pratica si è persa, *per quella particolare sequenza di numeri casuali*, la dipendenza dalla configurazione iniziale e, date le ipotesi di partenza, la configurazione finale è distribuita come desiderato.

Un fatto disagiabile è l’esigenza di dover trattare simultaneamente 2^N configurazioni. Per determinati casi, tale richiesta si può notevolmente rilasciare. Se nello spazio degli stati è possibile introdurre una relazione d’ordine parziale e se la funzione di update conserva tale relazione (ovvero se è mono-

tona rispetto ad essa), allora è sufficiente far evolvere solo gli stati massimali per tale ordine.

L'algoritmo del CFTP è il seguente:

1. Si selezioni come valore iniziale del tempo (nel passato) $t_0 = -1$;
2. si generi un vettore di numeri casuali \mathbf{U}_{t_0+1} ;
3. si simulino n MC, una per ogni stato iniziale $s_i \in \mathcal{S}$, dal tempo t_0 e le si facciano evolvere secondo $X_{t+1}(t_0, X_i) = \phi(X_t(t_0, X_i), \mathbf{U}_{t+1})$ fino al tempo 0: $t = t_0, t_0 + 1, \dots, -1$;
4. si controlli se è avvenuta la coalescenza al tempo 0, ovvero se tutte le catene $X_0(t_0, X_i)$ occupano lo stesso stato, $\forall i$. In caso affermativo, si restituisca il comune valore X_0 .
Altrimenti si ponga $t_0 \leftarrow (t_0 - 1)$ e si riprenda dal passo 2.

TEOREMA 1.19.

- i: L'algoritmo CFTP restituisce un valore con probabilità 1 e
- ii: il valore restituito è la realizzazione di una variabile casuale distribuita in accordo con la distribuzione stazionaria della MC.

Risultati sui Tempi di Coalescenza

Sia $X^{t_2}(t_1, x_m)$ lo stato al tempo t_2 di una MC partita dallo stato x_m al tempo $t_1 < t_2$. Sia inoltre $X_j^{t_2}(t_1, x_m)$ la componente j -esima di tale vettore, $j = 1, \dots, N$

DEFINIZIONE 1.16.

Definiamo il tempo di coalescenza nel futuro T_{fc} come il tempo di coalescenza in una simulazione in avanti che sia stata fatta partire al tempo 0.

DEFINIZIONE 1.17.

Definiamo il tempo di coalescenza dal passato T_{cftp} come il minor tempo t t.c. partendo dal tempo $-t$ si ha coalescenza al tempo 0 evolvendo tramite il CFTP.

TEOREMA 1.20.

I tempi di coalescenza nel futuro e dal passato t_{fc} e t_{cftp} posseggono la medesima distribuzione.

Propp e Wilson (1996) hanno provato un limite superiore per il tempo di coalescenza medio. Sia $\tau = \min_{t>0} \{d(t) \leq 1/e\}$ la soglia di variazione di una MC con matrice di transizione \mathbf{p} , dove $d(t)$ è la massima variazione della distanza. Per una procedura di update che preserva l'ordine parziale e che compie l'update di ognuno degli N componenti di \mathbf{X} singolarmente, si ha:

$$E(T_{cftp}) \leq 2\tau(1 + \log N).$$

In breve, il tempo medio di coalescenza sarà piccolo se l'algoritmo sottostante converge velocemente nella total variation distance.

Un'Interpretazione Alternativa del Coupling

Oltre a quanto visto finora, è possibile fornire una differente interpretazione dell'accoppiamento di più catene di Markov che, pur non presentando novità alla comprensione del soggetto, rappresenta un differente modo di intenderlo.

Il tutto si basa sul considerare l'accoppiamento di n catene non come n processi stocastici separati, bensì come un unico processo. La differenza principale tra le due interpretazioni si rende esplicita in particolar modo nel caso di catene accoppiate in modo dipendente. In questo caso infatti, una volta che i singoli processi divengono identici per la prima volta, essi coalescono e rimangono uguali per ogni altro tempo. Questo significa che la catena di Markov "globale" perde l'irriducibilità in favore di un insieme chiuso minimale di stati, quelli che presentano le configurazioni singole identiche.

Dalla teoria precedentemente esposta sappiamo che la convergenza all'insieme chiuso è assicurata e che, una volta raggiunto, sarà possibile conoscere l'evoluzione della catena in base solamente alla sottomatrice invariante della matrice di transizione. Sappiamo altresì che anche lo studio della dinamica che conduce al raggiungimento della coalescenza si riduce allo studio della matrice fondamentale del processo, dopo aver opportunamente contratto all'unità scalare la matrice ricorrente.

Due

IL MODELLO

Dove viene finalmente introdotto l'oggetto dei nostri studi. Non subito, però. Prima si parlerà dal gas su reticolo. Quindi si illustrerà il gas forzato su reticolo bidimensionale dando risalto, in entrambi i casi, alla definizione della dinamica microscopica e alle proprietà possedute. Infine verrà esposto la varietà quasi unidimensionale del gas su reticolo.

LA STORIA DEL MODELLO DI ISING è il racconto, lungo ed affascinante, di una delle creature, dotata di vita autonoma, partorite dalla mente di un fisico...¹

2.1 GAS SU RETICOLO

Consideriamo un reticolo ipercubico bidimensionale $\Lambda \subset \mathbb{Z}^2$ di volume $|\Lambda|$ e ad ogni sito i del reticolo assegnamo una variabile $n_i \in \{0, 1\}$. Interpretando il valore 1 come rappresentante una particella ed il valore 0 come uno spazio vuoto, questo modello può essere utilizzato per simulare un gas su di un reticolo. Lo spazio delle configurazioni (o degli stati) del sistema è dato da $\mathcal{C} = \{0, 1\}^{|\Lambda|}$. Dato che i valori ammessi per un sito sono solamente due, il modello rappresenta un gas nell'approssimazione di volume escluso, ovvero un sito non potrà ospitare più di una particella².

L'interazione tra le particelle è attrattiva e a corto raggio. Più precisamente l'hamiltoniana del sistema è data dall'hamiltoniana di Ising

¹Per descrivere il modello di Ising [32] e le sue varianti servirebbe un'enciclopedia. Ricordiamo qui le argomentazioni euristiche di Peierls [50, 27] e la soluzione esatta di Onsager [49]. Per un breve resoconto storico, consigliamo [8]. Per quanto concerne le simulazioni MCMC del modello di Ising, ricordiamo Metropolis [46, 5], Glauber [26], Swendsen e Wang [59] e Kawasaki [38, 37].

²Tale modello, conosciuto per l'appunto col nome di *gas su reticolo*, venne introdotto inizialmente da Lee e Yang [42].

$\mathcal{H}_J[\boldsymbol{\sigma}] = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$ con la conversione spin/particelle $\sigma_i = 2n_i - 1$:

$$\mathcal{H}_\Lambda[\mathbf{n}] = -4J \sum_{\langle i,j \rangle} n_i n_j, \quad (2.1)$$

dove con $\langle i, j \rangle$ indichiamo che la somma corre sui primi vicini. Con l'usuale metrica euclidea su \mathbb{Z}^2 , due siti i e j sono considerati primi vicini se $|i-j| = 1$.

Dato che intendiamo utilizzarlo come modello di un gas, dobbiamo richiedere che la dinamica conservi il numero di particelle. Adotteremo perciò la dinamica diffusiva di Kawasaki [38, 37] su tempi discreti, in tal modo la densità di particelle $\rho_\Lambda = \frac{1}{|\Lambda|}$ rimarrà costante.

Tale modello, al limite termodinamico (ottenuto mandando all'infinito le due dimensioni in modo indipendente) subisce una transizione di fase che, per $\rho = 1/2$, diviene del secondo ordine con temperatura critica data dalla temperatura di Onsager $T_C = \frac{2}{\ln(\sqrt{2}+1)} \sim 2.269$. Al di sopra di tale temperatura, il sistema si presenta con una distribuzione eterogenea di particelle e spazi vuoti. Al decrescere della temperatura la natura attrattiva delle interazioni tende a creare dei cluster di particelle che, nella fase ordinata, si aggregano in una fase unica, divisa da una fase "vuota" da un'interfaccia che disposta in modo da minimizzare la propria lunghezza (sfericamente nel caso di reticoli quadrati o quasi, parallelamente al lato minore nel caso di reticoli con grandi differenze tra le dimensioni).

In fig. 2.1 sono riportate alcune configurazioni tipiche del DLG, sia nello stato ordinato (figg. 2.1a-2.1h) che in quello disordinato (figg. 2.1i e 2.1j).

Descrizione Stocastica

Diamo ora una descrizione del modello nei termini del processo markoviano utilizzato per definirlo. Tale identificazione del sistema fisico e della sua dinamica con una catena di Markov si basa su quanto esposto nel capitolo precedente. In particolare identificheremo la probabilità che, a partire da una data configurazione, una particella effettui un salto in un sito libero vicino, con la componente della matrice di transizione che collega questi due stati¹.

¹È possibile definire la distanza tra due stati come il più piccolo numero di salti di particelle necessario per passare da uno stato all'altro. Se adottiamo questa definizione, allora la matrice di transizione avrà tutti gli elementi nulli eccetto che per quelle componenti che collegano stati a distanza unitaria. Questo in linea di principio, dato che le effettive probabilità di transizione potrebbero essere nulle anche tra stati primi vicini, a seconda della specifica dinamica adottata.

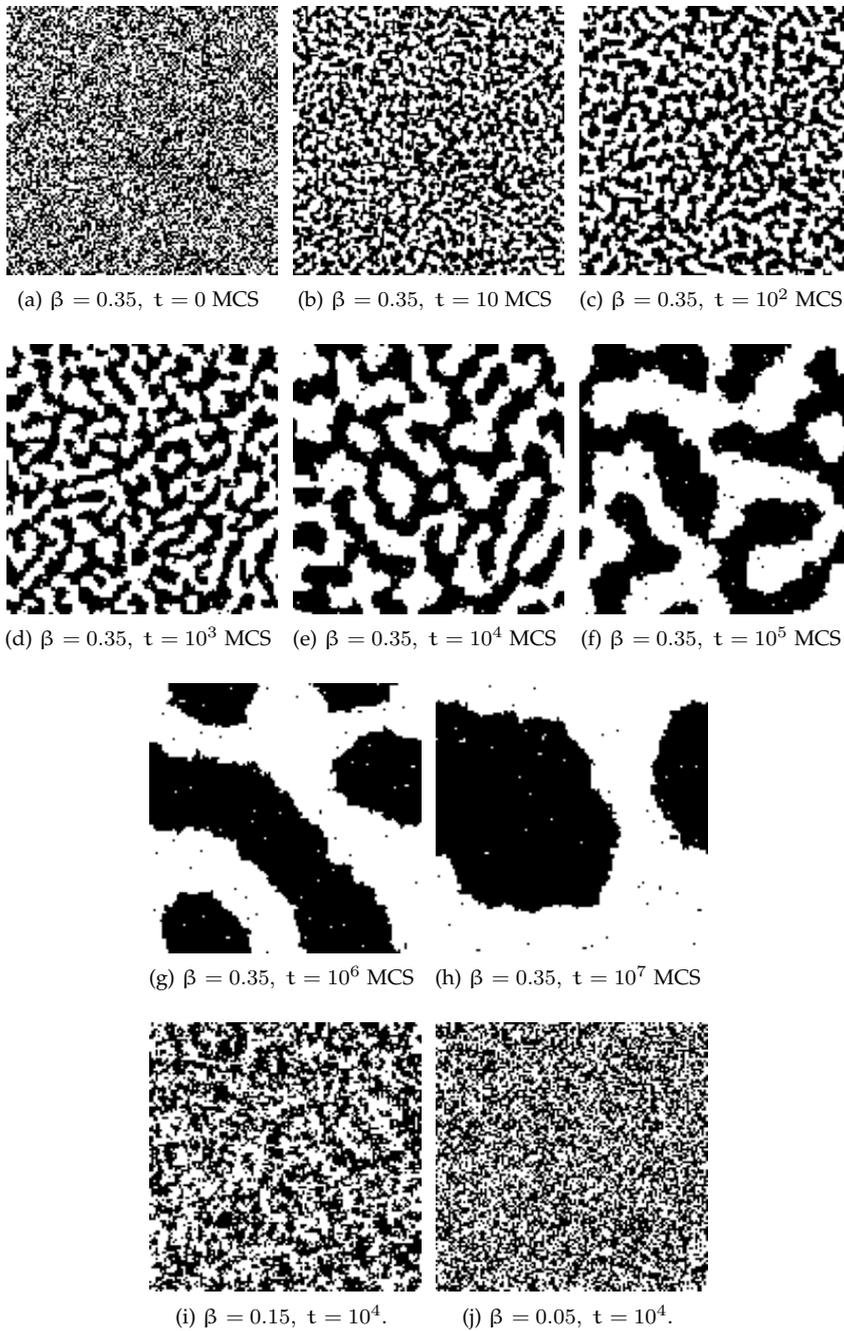


Figura 2.1: Configurazioni tipiche per un reticolo 128×128 nella fase ordinata ($\beta = 0.35$, figg. 2.1a–2.1h) e disordinata ($\beta = 0.15$, fig. 2.1i e $\beta = 0.05$, fig. 2.1j).

2. IL MODELLO

Sia \mathbf{p}^0 la distribuzione iniziale su $\mathcal{S} \equiv \mathcal{C}(\Lambda)$, e \mathbf{P} la matrice di transizione della catena di Markov corrispondente alla dinamica MCMC scelta. Sia quindi $\mathcal{P}_{(\mathbf{n}, \mathbf{n}_{ij})}$ la probabilità di transizione dallo stato \mathbf{n} allo stato \mathbf{n}_{ij} identico allo stato \mathbf{n} eccetto che per lo scambio degli spin nei siti i e j , primi vicini. Sia $\mathbf{p}_n^{(t)}$ la probabilità $\mathcal{P}(\mathbf{p}^0 \mathbf{P}^t = \mathbf{n})$ di avere lo stato \mathbf{n} al tempo t .

Allora $\mathbf{p}_n^{(t)}$ segue l'equazione di Fokker-Planck per tempi discreti

$$\Delta_t[\mathbf{p}_n^{(t)}] = \frac{1}{2} \sum_{i,j \in \Lambda} \left[\mathbf{p}_{\mathbf{n}_{ij}}^{(t)} \mathcal{P}_{(\mathbf{n}_{ij}, \mathbf{n})} - \mathbf{p}_n^{(t)} \mathcal{P}_{(\mathbf{n}, \mathbf{n}_{ij})} \right]. \quad (2.2)$$

Come abbiamo visto precedentemente, se la catena di Markov che utilizziamo per simulare la dinamica del modello è irriducibile, allora esiste un'unica distribuzione stazionaria $\boldsymbol{\pi}$ cui il processo converge.

Il modello in esame è stato scelto per rappresentare un gas su reticolo nell'ensemble canonico. In questo caso può essere pensato come un sistema termodinamico accoppiato ad un serbatoio di calore, con cui può scambiare energia. Questa interpretazione fisica del modello comporta due importanti condizioni per la MC:

- la distribuzione stazionaria deve essere data dalla misura di Gibbs per l'hamiltoniana di Ising, $\boldsymbol{\pi} = Z^{-1} e^{-\beta \mathcal{H}[\mathbf{n}]}$, dove β è la temperatura inversa del serbatoio di calore e svolge il ruolo di parametro;
- la dinamica deve essere invariante per inversione temporale. Dunque la catena di Markov deve essere reversibile.

La seconda condizione implica che la matrice di transizione \mathbf{P} deve soddisfare la proprietà di bilancio dettagliato

$$\pi_n \mathcal{P}_{\mathbf{n}, \mathbf{n}_{ij}} = \pi_{\mathbf{n}_{ij}} \mathcal{P}_{\mathbf{n}_{ij}, \mathbf{n}}, \quad \forall i, j \in \Lambda, \quad (2.3)$$

mentre la prima condizione è soddisfatta se

$$\frac{\mathcal{P}_{\mathbf{n}, \mathbf{n}_{ij}}}{\mathcal{P}_{\mathbf{n}_{ij}, \mathbf{n}}} = e^{\mathcal{H}[\mathbf{n}] - \mathcal{H}[\mathbf{n}_{ij}]}, \quad \forall i, j \in \Lambda. \quad (2.4)$$

L'equazione (2.4) può essere soddisfatta scegliendo \mathbf{P} tale che

$$\mathcal{P}_{\mathbf{n}, \mathbf{n}_{ij}} = w(\mathcal{H}[\mathbf{n}] - \mathcal{H}[\mathbf{n}_{ij}]), \quad (2.5)$$

dove w è qualsiasi funzione che soddisfi $w(-x) = e^x w(x)$.

La scelta di questa funzione determina la dinamica con cui il modello raggiungerà l'equilibrio, ma non influenza la precisa forma dello stato stazionario. Questa è una caratteristica importante nelle simulazioni numeriche

dei sistemi in equilibrio: grazie a questa libertà è possibile di volta in volta scegliere la dinamica più adatta allo studio che si intende effettuare. Per esempio si veda [40, 61], dove viene mostrato come la scelta tra le probabilità di transizione di Glauber, Metropolis e dell'heat-bath non modifichi il comportamento critico del modello (all'interno degli errori sperimentali). Nei sistemi fuori dall'equilibrio, tale libertà di scelta non è più concessa, in quanto il modello è definito proprio attraverso la dinamica. Una sua variazione, per quanto lieve, porterebbe in principio ad un modello differente. Nel modello che ci apprestiamo ad introdurre, gli studi di cui sopra hanno per esempio riportato una riduzione della temperatura critica T_C nel caso si usi l'algoritmo Metropolis piuttosto che gli altri due algoritmi, i quali portano a risultati simili. Tale differenza è dovuta alla maggior probabilità di accettazione degli algoritmi di Glauber e dell'heat-bath, che fa sì che il sistema si trovi ad una temperatura effettiva maggiore.

2.2 GAS FORZATO SU RETICOLO

Il gas forzato su reticolo (Driven Lattice Gas, DLG) è stato introdotto nel 1983 da Katz, Lebowitz e Spohn [35, 36], motivati dagli studi fisici sui conduttori superionici. Tale modello¹ è una generalizzazione del gas su reticolo in cui viene introdotto un campo di forza uniforme (nello spazio e nel tempo) diretto lungo uno degli assi del reticolo che per essere definiti, prendiamo corrispondente all'asse x : $\mathbf{E} = E\hat{x}$. L'effetto di tale campo è di favorire (sfavorire) i salti delle particelle nella direzione \hat{x} equiversa (opposta) al campo, lasciando inalterati quelli ad esso ortogonali.

Se imponiamo condizioni al bordo chiuse, il campo diviene conservativo e possiamo includere il suo effetto aggiungendo un termine di potenziale (lineare) all'hamiltoniana (2.1). Il sistema verrà a trovarsi ancora in uno stato d'equilibrio che possiamo descrivere come diviso in due fasi (quella vuota in alto e quella piena in basso, assumendo il campo diretto verso il basso) con un gradiente di densità tra le due zone.

Imponendo, invece, condizioni al bordo periodiche² la geometria risultante è quella di un toro ed in questo caso il campo cessa di essere con-

¹Per una introduzione generale al modello ed alle sue varianti si veda [57], mentre una storia dell'evoluzione degli studi teorici (al 2002) con referenze agli articoli originali si trova in [28, 23]. Per uno studio (recente) dettagliato anche dal punto di vista numerico si veda [10].

²In realtà sono sufficienti condizioni periodiche solo nella direzione del campo, per ottenere il comportamento critico.

servativo¹. Non esiste più un potenziale globale e nello stato stazionario è presente un flusso netto di particelle lungo la direzione di \mathbf{E} , violando così l'invarianza per inversione temporale.

Per $\mathbf{E} \neq 0$ e $\rho = 1/2$, il sistema subisce una transizione di fase continua per un valore finito di $\beta_C(\mathbf{E})$. Al variare di \mathbf{E} , $\beta_C(\mathbf{E})$ è decrescente e si hanno evidenze di una saturazione per $\beta_C(\infty) \approx 0.71\beta_C(0)$. Mentre per $\beta < \beta_C(\mathbf{E})$, le particelle sono distribuite omogeneamente nello spazio, per $\beta > \beta_C(\mathbf{E})$ si ha una separazione di fasi: troviamo la coesistenza di una fase ad alta densità, corrispondente a siti quasi tutti pieni, ed una fase a bassa densità, corrispondente a siti per lo più vuoti. Le due fasi dividono il reticolo in due regioni distinte, con l'interfaccia che le separa parallela alla direzione del campo \mathbf{E} .

In fig. 2.2 sono riportate alcune configurazioni tipiche del DLG, sia nello stato ordinato (figg. 2.2a–2.2h) che in quello disordinato (figg. 2.2i e 2.2j). È evidente dalla fig. 2.2i come l'influenza del campo sia evidente anche nella fase disordinata, per temperature non troppo elevate.

Descrizione Stocastica del Modello

Partendo dal modello del gas su reticolo uno stato stazionario, ma di non equilibrio, potrebbe essere raggiunto facendo cessare la validità dell'eq. (2.3) di bilancio dettagliato. La catena di Markov rimarrebbe regolare, garantendo così la convergenza ad un'unica distribuzione limite ma, non essendo più reversibile, tale distribuzione non sarebbe più una distribuzione d'equilibrio.

Fisicamente tale stato può essere prodotto accoppiando il gas con un campo di forza esterno \mathbf{E} che agisca su ogni sito del reticolo influenzando le frequenze di transizione tra i vari stati. È possibile sceglierlo in modo da far sì che, al posto dell'eq. (2.4), venga soddisfatta

$$\frac{P_{\mathbf{n}, \mathbf{n}_{ij}}}{P_{\mathbf{n}_{ij}, \mathbf{n}}} = e^{\mathcal{H}[\mathbf{n}] - \mathcal{H}[\mathbf{n}_{ij}] + \epsilon \beta |\mathbf{E}(i,j)|} \quad (2.6)$$

dove $\epsilon = \frac{i-j}{|i-j|} \cdot \frac{\mathbf{E}}{|\mathbf{E}|} \in \{-1, 0, +1\}$ per tentativi di salto di una particella nella direzione del campo, trasversi e contro la direzione del campo. Fisicamente, questa modifica serve a tener conto del lavoro compiuto localmente dal campo esterno sulle particelle.

¹In questo caso il sistema diviene anche molto difficile da riprodurre fisicamente. Sarebbe infatti necessario mantenere attivo un campo magnetico rapidamente crescente, fatto possibile solo per tempi molto limitati

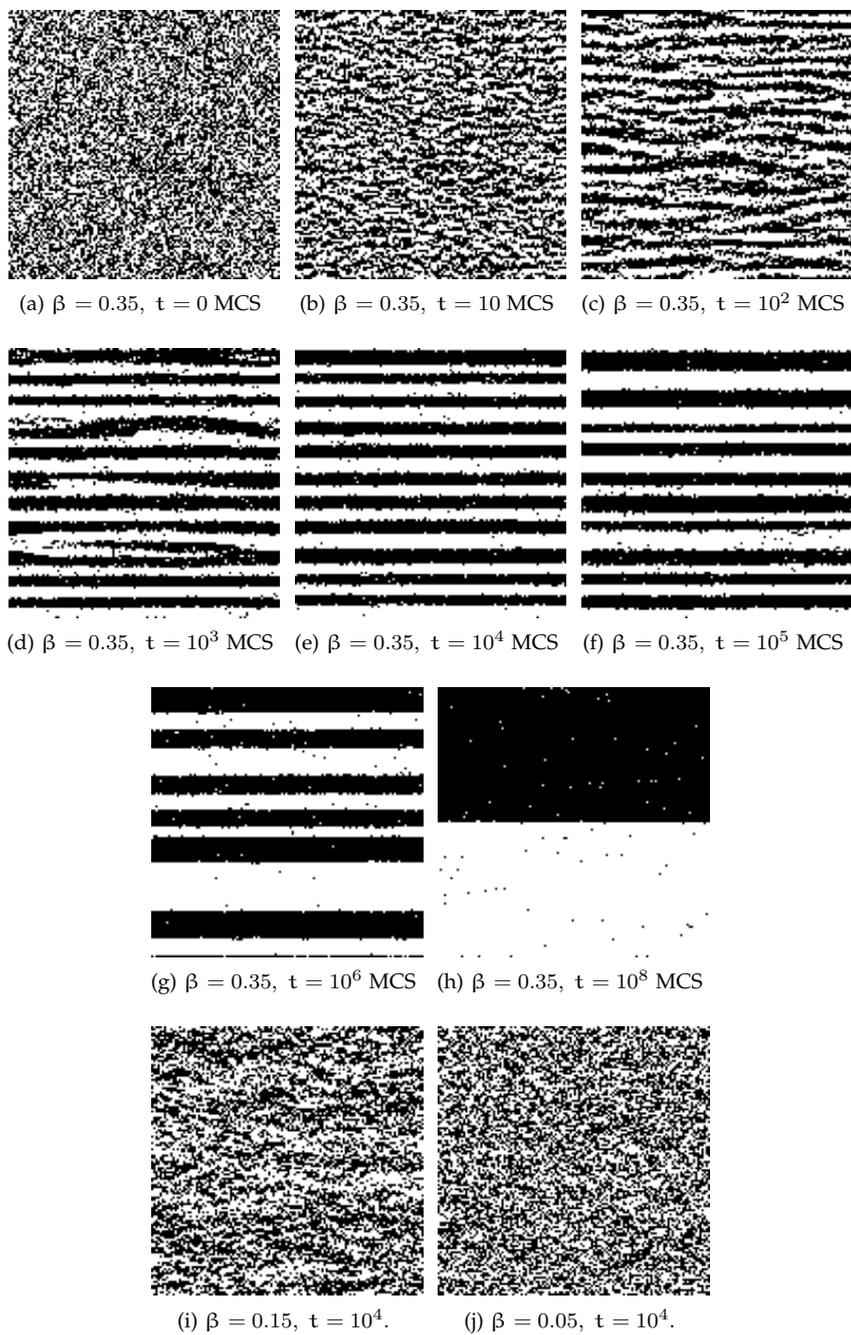


Figura 2.2: Configurazioni tipiche per un reticolo 128×128 con campo $\mathbf{E} = \infty$, nella fase ordinata ($\beta = 0.35$, figg. 2.2a–2.2h) e disordinata ($\beta = 0.15$, fig. 2.2i e $\beta = 0.05$, fig. 2.2j).

Una delle maggiori conseguenze di questi assunti è che abbiamo perso la conoscenza a priori della distribuzione stazionaria. Ciò ci impedisce di poter applicare gran parte dei risultati già stabiliti per i sistemi all'equilibrio.

2.3 PROPRIETÀ

Osservabili

Definiamo ora le osservabili che ci permettono di caratterizzare quantitativamente le transizioni di fase del DLG e che sono accessibili tramite simulazioni numeriche. Per comodità daremo tutte le definizioni con riferimento al caso bidimensionale in quanto coincide, almeno a livello di notazione, col caso quasi-unidimensionale. Sia dunque Λ un reticolo ipercubico di dimensioni $L_{\perp} \times L_{\parallel}$, $L_{\perp} = 8$, con condizioni periodiche al contorno.

Una caratteristica importante del modello è il fatto che la magnetizzazione totale è costante. In base a ciò non ci è permesso utilizzare, come in modelli magnetici, questa osservabile come parametro d'ordine. Il sorgere della transizione è evidenziato piuttosto dalle disomogeneità (fluttuazioni) di tale osservabile a grandi lunghezze d'onda. Utilizziamo le variabili di spin $\sigma_i \equiv 2n_i - 1$ con la relativa trasformata di Fourier

$$\phi(\mathbf{K}) \equiv \sum_{j \in \Lambda} e^{i\mathbf{k} \cdot \mathbf{j}} s_j; \quad (2.7)$$

dove i valori di \mathbf{k} permessi sono dati da

$$\mathbf{k}_{n,m} = \left(\frac{2\pi n}{L_{\parallel}}, \frac{2\pi m}{L_{\perp}} \right), \quad (n, m) \in \mathbb{Z}_{L_{\parallel}} \times \mathbb{Z}_{L_{\perp}}. \quad (2.8)$$

Consideriamo ora il modello con densità $\rho = 0.5$. Allora

$$\sum_{j \in \Lambda} s_j = 0, \quad \text{implica } \phi(\mathbf{k}_{0,0}) = 0. \quad (2.9)$$

Nella fase ordinata $|\phi|$ raggiunge il suo massimo per $\mathbf{k} = \mathbf{k}_{0,1}$ ed il valore di aspettazione sullo stato stazionario del suo modulo

$$m(\beta; L_{\perp}, L_{\parallel}) = \frac{1}{|\Lambda|} \langle |\phi(\mathbf{k}_{0,1})| \rangle \quad (2.10)$$

è un buon parametro d'ordine. Nel caso $8 \times L_{\parallel}$

$$\max_{\Lambda} m = \frac{1}{2} \sqrt{1 + \frac{\sqrt{2}}{2}} \simeq 0.653. \quad (2.11)$$

Nello spazio dei momenti, il fattore di struttura statico (definito come la trasformata di Fourier della funzione di correlazione a due punti)

$$\tilde{G}(\mathbf{k}; L_{\perp}, L_{\parallel}) = \frac{1}{|\Lambda|} \langle |\phi(\mathbf{k})|^2 \rangle \quad (2.12)$$

svanisce in $\mathbf{k}_{0,0}$ a causa dell'eq. (2.9) e raggiunge il proprio massimo in $\mathbf{k}_{0,1}$. Diviene naturale definire la suscettività come

$$\chi_{\perp}(\beta; L_{\perp}, L_{\parallel}) \equiv \tilde{G}(\mathbf{k}_{0,1}; L_{\perp}, L_{\parallel}). \quad (2.13)$$

Definiamo inoltre la funzione di correlazione connessa a 4 punti

$$\tilde{G}^{(4)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4; L_{\perp}, L_{\parallel}) = \frac{1}{|\Lambda|} \langle \phi(\mathbf{k}_1) \phi(\mathbf{k}_2) \phi(\mathbf{k}_3) \phi(\mathbf{k}_4) \rangle_{\text{conn}} \quad (2.14)$$

ed il relativo cumulante di Binder (trasverso) $g(\beta; L_{\perp}, L_{\parallel})$

$$g(\beta; L_{\perp}, L_{\parallel}) = 2 - \frac{\langle |\phi(\mathbf{k}_{01})|^4 \rangle}{\langle |\phi(\mathbf{k}_{01})|^2 \rangle^2} = - \frac{\tilde{G}^{(4)}(\mathbf{k}_{(0,1)}, \mathbf{k}_{(0,1)}, -\mathbf{k}_{(0,1)}, -\mathbf{k}_{(0,1)}; L_{\perp}, L_{\parallel})}{|\Lambda| [\tilde{G}^{(4)}(\mathbf{k}_{(0,1)}; L_{\perp}, L_{\parallel})]^2}. \quad (2.15)$$

Se la distribuzione di $\phi(\mathbf{k}_{0,1})$ è gaussiana nello stato stazionario, allora $\langle |\phi(\mathbf{k}_{01})|^4 \rangle = 2 \langle |\phi(\mathbf{k}_{01})|^2 \rangle^2$ e quindi $g = 0$. Questo dovrebbe valere nella fase disordinata. Al di sotto della temperatura critica, ci aspettiamo invece $\langle |\phi(\mathbf{k}_{01})|^n \rangle = 2 \langle |\phi(\mathbf{k}_{01})| \rangle^n$ e dunque $g = 1$.

Definire una lunghezza di correlazione è invece più difficile. Possiamo definirla in funzione del comportamento a grandi distanze della funzione a due punti, o del comportamento per piccoli momenti.

La funzione a due punti

$$G(\mathbf{x} - \mathbf{y}) = \langle n_{\mathbf{x}} n_{\mathbf{y}} \rangle = \frac{1}{L_{\perp} L_{\parallel}} \sum_{\mathbf{k}} \tilde{G}(\mathbf{k}) e^{-i\mathbf{k} \cdot (\mathbf{x} - \mathbf{y})} \quad (2.16)$$

è già stata studiata per mezzo di simulazioni MC ed approssimazioni analitiche: essa mostra un comportamento a legge di potenza

$$G(\mathbf{x}) \propto |\mathbf{x}|^{-d} \quad (2.17)$$

e non è definita positiva. Il decadimento algebrico a legge di potenza delle correlazioni, da vita ad una discontinuità nel fattore di struttura statico $\tilde{G}(\mathbf{k}) = \sum_{\mathbf{x} \in \Lambda} e^{i\mathbf{k} \cdot \mathbf{x}} G(\mathbf{x})$ per $\mathbf{k} = 0$. È comunque possibile dimostrare come il decadimento esponenziale venga ritrovato se si considera la correlazione media tra due punti sull'angolo solido.

In questo studio useremo per la lunghezza di correlazione una definizione alternativa basata sull'andamento della funzione a due punti per piccoli momenti (per la sua definizione ed uno studio dettagliato, vedi [?] e le referenze qui presenti):

$$\xi_{ij}(L_{\parallel}, L_{\perp}) \equiv \sqrt{\frac{1}{\hat{q}_j^2 - \hat{q}_i^2} \left(\frac{\tilde{G}_{\perp}(q_i; L_{\parallel}, L_{\perp})}{\tilde{G}_{\perp}(q_j; L_{\parallel}, L_{\perp})} - 1 \right)}, \quad (2.18)$$

dove $\tilde{G}_{\perp}(q; L_{\parallel}, L_{\perp}) \equiv \tilde{G}_{\perp}((0, q); L_{\parallel}, L_{\perp})$ è il fattore di struttura a momento longitudinale nullo in volume finito ($\tilde{G}_{\perp}(0; L_{\parallel}, L_{\perp}) = 0$) e $\hat{q}_n = 2 \sin(\pi n/L_{\perp})$ è il momento del reticolo. Dato che il sistema si ordina in un numero pari di strisce (contando sia quelle piene che quelle vuote) $\tilde{G}_{\perp}(q_i; L_{\parallel}, L_{\perp})$ è nullo per i pari se $\beta \rightarrow \infty$. quindi scegliamo i e j dispari. Dato che le correzioni di taglia finita aumentano all'aumentare di i, j, scegliamo $(i, j) = (1, 3)$ e quindi $\xi_{\perp} = \xi_{13}$.

La lunghezza di correlazione longitudinale viene definita allo stesso modo, con la variante che si ha $\tilde{G}_{\perp}(q; L_{\parallel}, L_{\perp}) \equiv \tilde{G}_{\perp}((q, 0); L_{\parallel}, L_{\perp})$ e che, potendo utilizzare anche i valori pari di i e di j, prenderemo $\xi_{\parallel} = \xi_{12}$.

Vicino ad un punto critico di transizione di fase, le quantità sopra definite mostrano una divergenza a legge di potenza. Definiamo gli esponenti critici trasversi, ponendo

$$\xi_{\perp} \sim t^{-\gamma_{\perp}} \quad (2.19)$$

$$\chi_{\perp} \sim t^{-\gamma'_{\perp}} \quad (2.20)$$

per $t \equiv \frac{\beta_c - \beta}{\beta_c} \rightarrow 0^+$.

Nella descrizione al continuo del DLG effettuata con la teorie di campo [33, 43] è possibile calcolare *esattamente* i valori degli esponenti critici, per $2 \leq d \leq 5$. Li riportiamo in tab. 2.1, insieme a quelli per il modello di Ising.

Anisotropia ed Esponenti Critici

La presenza di una forte anisotropia nel sistema è una delle proprietà più evidenti del DLG. In realtà la presenza di anisotropia viene riscontrata anche in sistemi all'equilibrio ma, come vedremo, si tratta di un fenomeno differente. Nei modelli fisici sono infatti possibili 2 tipi differenti di anisotropie. Il primo tipo è collegato al fatto che le funzioni di scala presentano ampiezze differenti lungo le varie direzioni spaziali; tuttavia continuano a possedere

	Ising	DLG ^a
α	0 (log)	
β	1/8	1/2
γ	7/4	1
δ	15	
η	1/4	0
ν	1	1/2
Δ	0	$\frac{1}{3}(8-d)$
z		4

^aI valori riportati si riferiscono agli esponenti *trasversi*. Sono comunque collegati a quelli longitudinali tramite il parametro d'anisotropia Δ . Se s è un generico esponente critico, $\alpha_{\parallel} = (1 + \Delta)\alpha_{\perp}$.

Tabella 2.1: *Esponenti critici per il modello di Ising bidimensionale e per il gas forzato su reticolo.*

lo stesso valore dell'esponente critico (per esempio il modello di Ising con costanti d'accoppiamento differenti lungo gli assi).

Il secondo tipo di anisotropia, spesso denominata "forte" mentre la prima è detta "debole", è presente quando sono gli esponenti stessi ad avere valori diversi nelle varie direzioni. Come conseguenza nel nostro caso si ha uno scaling fortemente anisotropico, con un sdoppiamento di tutti gli esponenti critici. L'assunzione fondamentale delle teorie di scala è che quando ci si avvicina al punto critico, l'unica scala di lunghezze rilevante sia quella di correlazione $\xi = \tau^{-\nu}$, dove $\tau \propto T - T_C$. Tutte le altre lunghezze, microscopiche nei confronti di ξ , non influenzano il comportamento critico. Per tale motivo tutte le funzioni termodinamiche e le osservabili macroscopicamente accessibili saranno funzioni (omogenee) di ξ .

Prendiamo in considerazione il *fattore di struttura dinamico* $S(\mathbf{k}, t, \tau)$, definito come la trasformata di Fourier della funzione di correlazione a due punti. Se indichiamo con μ la scala per i momenti, si ottiene per S che

$$S(\mathbf{k}_{\parallel}, \mathbf{k}_{\perp}, t, \tau) = \mu^{-2+\nu} S\left(\frac{\mathbf{k}_{\parallel}}{\mu^{1+\Delta}}, \frac{\mathbf{k}_{\perp}}{\mu}, t, \frac{\tau}{\mu^{1/\nu}}\right). \quad (2.21)$$

Per tener conto dell'anisotropia è necessaria l'introduzione di un ulteriore esponente, chiamato esponente di anisotropia e denotato con Δ . Per sistemi all'equilibrio $\Delta = 0$ in quanto le varie direzioni scalano con la stessa potenza della scala dei momenti. Come conseguenza dell'eq. (2.21), assumendo

2. IL MODELLO

$\mu = \tau^\nu$, si ottiene:

$$\mathbf{k}_\perp \sim \tau^\nu, \quad \mathbf{k}_\parallel \sim \tau^{\nu(1+\Delta)}, \quad (2.22)$$

Da questa equazione possiamo ricavare i due diversi andamenti dell'esponenti critico ν :

$$\nu_\perp = \nu, \quad \nu_\parallel = \nu(1 + \Delta). \quad (2.23)$$

Quanto sopra può essere interpretato come l'emergere di due diverse lunghezze di correlazione, una longitudinale e l'altra parallela alla direzione del campo,

$$\xi_\perp \sim \tau^{-\nu_\perp}, \quad \xi_\parallel \sim \tau^{-\nu_\parallel}. \quad (2.24)$$

Basandoci su $\mu = t^{-1/z}$, ricaviamo gli esponenti anisotropici corrispondenti a z sfruttando le proprietà a lungo termine:

$$\mathbf{k}_\perp \sim t^{-1/z}, \quad \mathbf{k}_\parallel \sim t^{-(1+\Delta)/z}, \quad (2.25)$$

da cui

$$z_\perp = z, \quad z_\parallel = z/(1 + \Delta). \quad (2.26)$$

Concentriamoci ora su η , definito come l'esponente critico che regola per $\tau = 0$ la dipendenza del fattore di struttura statico dai momenti.

$$S(\mathbf{k}_\parallel, \mathbf{k}_\perp) = \mathbf{k}_\perp^{-2+\nu_\perp \Sigma_\perp} \left(\frac{\mathbf{k}_\parallel}{\mathbf{k}_\perp^{1+\Delta}} \right) = \mathbf{k}_\perp^{-2+\eta_\parallel \Sigma_\parallel} \left(\frac{\mathbf{k}_\perp}{\mathbf{k}_\parallel^{1/(1+\delta)}} \right) \quad (2.27)$$

Confrontando (2.27) con (2.21), otteniamo

$$\eta_\perp = \eta, \quad \eta_\parallel = \frac{\eta + 2\Delta}{1 + \Delta}. \quad (2.28)$$

Per la suscettibilità il discorso è più complesso, dato che ci sono due possibili definizioni, tra cui quella usata solitamente per i sistemi all'equilibrio, legata al comportamento per bassi momenti della funzione di correlazione a due punti. Ciò non vale ovviamente fuori dall'equilibrio, ma è tuttavia utile definirla in questo modo.

Poniamo dunque

$$\begin{aligned} \chi_\perp(\tau) &= S(\mathbf{k}_\parallel = 0, \mathbf{k}_\perp \rightarrow 0, 0; \tau) \\ \chi_\parallel(\tau) &= S(\mathbf{k}_\parallel \rightarrow 0, \mathbf{k}_\perp = 0, 0; \tau) \end{aligned} \quad (2.29)$$

e definiamo altri due esponenti critici:

$$\chi_\parallel(\tau) \sim |\tau|^{-\gamma_\parallel}, \quad \chi_\perp(\tau) \sim |\tau|^{-\gamma_\perp}. \quad (2.30)$$

Riprendendo l'eq. (2.21) possiamo ottenere

$$\gamma_{\parallel} = \gamma_{\perp} = \nu(2 - \eta) = \gamma. \quad (2.31)$$

Ricordando infine che la relazione di Fisher è soddisfatta anche in presenza di anisotropie possiamo concludere con

$$\nu_{\parallel}(2 - \eta_{\parallel}) = \nu_{\perp}(1 + \Delta) \left(2 - \frac{\eta_{\perp} + 2\Delta}{1 + \Delta} \right) = \nu_{\perp}(2 + \eta_{\perp}); \quad (2.32)$$

da cui

$$\gamma_{\parallel} = \nu_{\parallel}(2 - \eta_{\parallel}) \quad \gamma_{\perp} = \nu_{\perp}(2 - \eta_{\perp}). \quad (2.33)$$

2.4 GAS FORZATO SU RETICOLO QUASI UNIDIMENSIONALE

Prendiamo ora in considerazione il modello che sarà oggetto del nostro studio. Tale modello consiste in una variante quasi unidimensionale del DLG, in cui "quasi unidimensionale" significa che il limite termodinamico viene effettuato mandando all'infinito solo la lunghezza del reticolo in una direzione, mentre l'altra viene mantenuta fissata. Nello specifico caso in esame la direzione che sarà libera di variare sarà quella parallela alla direzione del campo, mentre quella trasversa sarà fissata ad un valore di 8 passi reticolari.

Tale modello, tra le varianti del DLG, è forse quella che ha finora ricevuto l'attenzione minore, nonostante presenti caratteristiche di notevole interesse.

Una Breve Descrizione

La versione quasi unidimensionale del gas forzato su reticolo coincide con quella della sua versione strettamente bidimensionale, con la sola differenza che, data la geometria del reticolo ipercubico $L_{\perp} \times L_{\parallel}$, si mantiene sempre costante L_{\perp} (in questo studio, per motivi specificati nel seguito, $L_{\perp} = 8$) e si effettua il limite termodinamico $L_{\parallel} \rightarrow \infty$ solo sulla dimensione parallela al campo. In questo modo $\lim_{L_{\parallel} \rightarrow \infty} \frac{L_{\perp}}{L_{\parallel}} \rightarrow 0$.

Proprietà Note

A nostra conoscenza, la letteratura sull'argomento è abbastanza modesta:¹

¹La letteratura è più ricca per quanto concerne il sistema all'equilibrio, dove peraltro il modello viene chiamato sovente "Ising cylinder" o "Ising strip"; termine quest'ultimo che avrebbe

- una breve citazione in un articolo di Wang[62] sul finite size scaling anisotropico in cui si afferma, con molto pragmatismo, che “Per forme molto allungate (quando $L \rightarrow 0$ tenendo M fisso o viceversa) $T_C(L, M)$ decresce verso lo zero, dato che il sistema in questo limite è quasi-unidimensionale”.
- Un'altra breve comparsa in un articolo di Vallés e Marro relativo all'influenza delle frequenze di transizione sulle transizioni di fase, in cui gli autori hanno simulato sistemi di dimensione 2×10000 , individuando la presenza di una transizione di fase “diffusa” su di un intervallo $\Delta T_C \xrightarrow{L_\perp \rightarrow L_\parallel} T_C$ che si assottiglia all'avvicinarsi delle due dimensioni.
- Uno studio dettagliato del diagramma di fase in funzione di β ed E compiuto da Caravenna[11], che è alla base del nostro lavoro e al quale ci riferiremo implicitamente durante tutta questa sezione.

L'interesse verso questo modello è legato soprattutto alla presenza, in una regione definita dei parametri del sistema, di una transizione di fase molto simile a quella del caso bidimensionale. Tale fenomeno risulta notevole in quanto l'analogo sistema all'equilibrio non presenta nessuna transizione di fase.

Infatti generalmente i modelli quasi unidimensionali d'equilibrio sono assimilabili a quelli strettamente unidimensionali; in particolare in questi ultimi, se dotati di interazioni a corto raggio, non vi è la presenza di transizioni di fase [27]. In fisica vi molti sono esempi di sistemi unidimensionali, come l'ASEP (Asymmetric Simple Exclusion Process, processo di esclusione asimmetrico semplice), in cui tali transizioni sono presenti. Ci teniamo tuttavia a far notare che tali sistemi, ben noti nella letteratura, presentano notevoli differenze rispetto al DLG, principalmente per quanto concerne la dinamica (l'ASEP ha bordi aperti con transizioni speciali, e la transizione di fase è indotta proprio dalle frequenze di transizione ai bordi).

Nella scelta della dimensione trasversa, la scelta di un valore pari ad 8 passi reticolari è motivata dalla ricerca di un compromesso tra la minor dimensione possibile e la stabilità della striscia che si forma nella fase ordinata. Per valori fino a 6 passi reticolari infatti la striscia che si forma risulta

provocato non poca confusione, se fosse stato adottato anche per il DLG. Per il nostro caso gli articoli recenti sono: [31, 44, 18, 2, 55, 9], anche se va notato che nella maggior parte di essi, alla fine viene mandato al limite termodinamico anche il lato “corto”.

instabile contro le fluttuazioni termodinamiche e tende a rompersi in continuazione. Possiamo assumere che 8 sia la dimensione minore per cui la striscia è stabile rispetto a tali fluttuazioni.

Riportiamo nelle figg. 2.3 e 2.4 alcune configurazioni tipiche del modello quasi-unidimensionale, sia nella fase ordinata, che in quella disordinata.

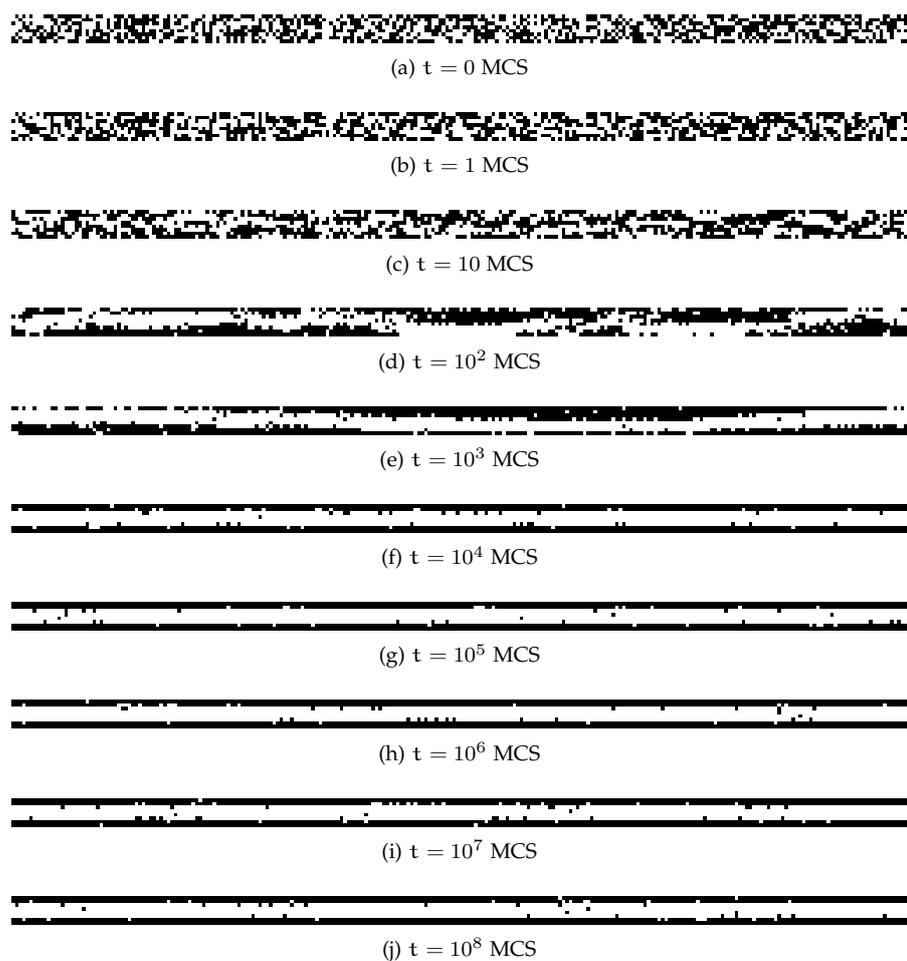


Figura 2.3: Configurazioni corrispondenti all'evoluzione temporale di un modello su reticolo di dimensione 8×256 inizializzata in maniera casuale. $\beta = 0.35$.

Diagramma di Fase La struttura del diagramma di fase è stata analizzata a fondo in [11], dove vengono evidenziate le diverse zone in cui si ripartisce lo spazio delle fasi parametrizzato da (β, E) .



Figura 2.4: Configurazione corrispondente alla stessa geometria della fig. 2.3 con $\beta = 0.15$, dopo 10^4 MCS.

Nei casi limite si ha:

- $\beta = 0$ Tutte le transizioni trasverse e nella direzione del campo sono sempre ammesse, mentre quelle contrarie al campo sono ammesse con probabilità e^{-E} . La misura risultante è quella stazionaria.
- $\beta = \infty$ Tutte le transizioni che comportano un aumento di energia sono soppresse e lo stato finale dipende fortemente dallo stato iniziale.
- $E = 0$ In questo caso il modello risultante è esattamente il modello di Ising con dinamica diffusiva che, nel limite termodinamico, non presenta transizione di fase per valori finiti di β .
- $E = \infty$ I salti contrari al campo sono proibiti *sempre*, quelli favorevoli ammessi *sempre*. Questo è il caso su cui ci concentreremo in questo studio.

Lo spazio delle fasi si divide principalmente in due zone: una ordinata O, ed una fase disordinata, che è divisa a sua volta in D_1 e D_2 . Inoltre è presente una terza zona B, compresa tra O e D_2 , che è caratterizzata da metastabilità. Queste zone sono identificate da una temperatura critica per la transizione (di secondo ordine) D_1 -O $\beta_c \simeq 0.7$ e dall'estensione della zona D_2 , che si estende tra D_2 ed O per valori di E grossomodo compresi tra 0.7 e 2.

Nel dettaglio:

- O Questa è la zona ordinata, caratterizzata da una suddivisione, analogamente al caso bidimensionale, del reticolo in una zona ad alta densità ed una zona a bassa densità, separate da un'interfaccia parallela alla direzione del campo.
- D_1 Questa è la zona disordinata, in cui le configurazioni tipiche presentano una distribuzione eterogenea di particelle e di buchi.
- D_2 Questa zona, appartenente alla macrozona ordinata D, è caratterizzata dalla presenza di aggregazioni di particelle in strisce perpendicolari al campo.

B In questa zona sono presenti fenomeni di metastabilità, nella forma di un alternarsi di strisce parallele/ortogonali alla direzione del campo. Non ci sono comunque evidenze numeriche che questa zona possa sopravvivere al limite termodinamico.

Noi ci occuperemo dello studio del modello esclusivamente per valori del campo $E = \infty$.

Tre

DALLE TEORIE D'EQUILIBRIO AL DRIVEN DIFFUSIVE SYSTEM

Dove si renderà conto del tentativo di considerare un'infinità di particelle discrete come un continuo, e di descrivere il loro comportamento dimenticandosi della loro esistenza.

GLI STRUMENTI DELLA TEORIA DI CAMPO ed i metodi del gruppo di ri-normalizzazione (RG) svolgono un ruolo fondamentale nella comprensione delle leggi di scala universali che sono presenti nelle vicinanze dei punti critici per transizioni di fase all'equilibrio, inclusa la dinamica dei sistemi critici.

Qui vogliamo dare una breve esposizione di ciò e del tentativo di descrivere allo stesso modo le proprietà di scala di sistemi forzati lontano dall'equilibrio, ma che subiscono una transizione di fase continua. In pratica il tentativo è quello di catturare la dinamica — stocastica — dei modi a grande lunghezza d'onda dei gradi di libertà lenti, come il parametro d'ordine e di quantità conservate, e di altre osservabili di interesse.

3.1 SISTEMI ALL'EQUILIBRIO

Rallentamento Critico in Prossimità di una Transizione di Fase

Nelle vicinanze di un punto critico sono presenti correlazioni e fluttuazioni in maniera nettamente maggiore rispetto ad altre zone del diagramma di fase. Il sistema si comporta quindi in maniera fortemente cooperativa e di conseguenza i metodi di approssimazione della meccanica statistica, basati sull'assunzione di interazioni deboli, falliscono.

Nel limite termodinamico, all'avvicinarsi ad una transizione di fase del secondo ordine le fluttuazioni del parametro d'ordine $S(\mathbf{x})$ (scelto in modo da annullarsi nella fase disordinata) sono governate da una scala di lunghezze che diverge come

$$\xi(\tau) \sim |\tau|^{-\nu} \quad (3.1)$$

per $\tau \ll 1^1$.

Qui ν è l'esponente critico che caratterizza il decadimento della lunghezza di correlazione, definita classicamente per mezzo del decadimento esponenziale della funzione di correlazione a due punti

$$C(\mathbf{x}) = \langle S(\mathbf{x})S(0) \rangle - \langle S \rangle^2 \sim e^{-\frac{|\mathbf{x}|}{\xi}}. \quad (3.2)$$

Per $\tau \rightarrow 0$, la divergenza della lunghezza di correlazione mostra l'assenza di una scala caratteristica per le fluttuazioni di S . Ci aspettiamo quindi che le correlazioni seguano un andamento a legge di potenza

$$C(\mathbf{x}) \sim |\mathbf{x}|^{-(d-2+\eta)}, \quad (3.3)$$

dove abbiamo definito l'esponente di Fisher η e d indica le dimensioni del reticolo. Vicino alla temperatura critica, vale invece la seguente legge di scala:

$$C(\tau, \mathbf{x}) \sim |\mathbf{x}|^{-(d-2+\eta)} \tilde{C}_{\pm}(\mathbf{x}/\xi), \quad (3.4)$$

con due distinte ampiezze per $T \gtrless T_C$.

Prendendone la trasformata di Fourier $C(\tau, \mathbf{q}) = \int d\mathbf{x} e^{-i\mathbf{q}\cdot\mathbf{x}} C(\tau, \mathbf{x})$, otteniamo

$$C(\tau, \mathbf{q}) = |\mathbf{q}|^{-2+\eta} \hat{C}_{\pm}(\mathbf{q}\xi) \quad (3.5)$$

con funzioni di scaling

$$\hat{C}_{\pm}(\mathbf{p}) = |\mathbf{p}| \int d^d \mathbf{y} e^{-i\mathbf{p}\cdot\mathbf{y}} |\mathbf{y}|^{d-2+\eta} \tilde{C}_{\pm}(\mathbf{y}), \quad (3.6)$$

dove $\mathbf{p} = \mathbf{q}\xi$ e $\mathbf{y} = \mathbf{x}/\xi$.

Dato che gli esponenti critici statici effettivamente indipendenti sono solamente due, è possibile utilizzare le precedenti equazioni per esprimere gli esponenti critici vicino ad una transizione di fase di secondo ordine in termini di ν e di η .

¹Useremo per la temperatura l'espressione ridotta $\tau \equiv (T - T_C)/T_C$, che fornisce una misura della distanza dal punto critico.

Nella fase ordinata ($\tau < 0$) ci si attende che S cresca come

$$\langle S \rangle \sim -T^\beta; \quad (3.7)$$

da cui ricaviamo la relazione d'iperscala:

$$\beta = \frac{\nu}{2}(d - 2 + \eta). \quad (3.8)$$

La suscettività isoterma statica χ_T all'equilibrio è data (teorema di fluttuazione-dissipazione) dall'integrale spaziale della funzione di correlazione $C(\tau, \mathbf{x})$:

$$\chi_\tau(\tau) = \beta \lim_{\mathbf{q} \rightarrow 0} C(\tau, \mathbf{q}), \quad (3.9)$$

da cui si ottiene

$$\chi_\tau(\tau) \sim \xi^{2-\eta} \sim |\tau|^{-\nu(2-\eta)}. \quad (3.10)$$

Definendo γ come l'esponente critico associato a χ ($\chi_\tau(\tau) \sim |\tau|^{-\gamma}$) si ottiene un'altra relazione di scala:

$$\gamma = \nu(2 - \eta). \quad (3.11)$$

All'equilibrio la probabilità di trovare un sistema ad una temperatura T in una configurazione $S(\mathbf{x})$, è data da

$$\mathcal{P}_{eq}[S] \propto \exp \left\{ -\frac{\mathcal{H}[S]}{k_B T} \right\}. \quad (3.12)$$

Da questo punto una strada che si può seguire è quella che porta alla classica descrizione statica dei sistemi critici all'equilibrio, per cui una teoria con un parametro d'ordine vettoriale con simmetria $O(n)$ $S^\alpha(\mathbf{x})$ viene descritta dall'equazione di Landau-Ginzburg-Wilson (LGW)

$$\mathcal{H}[S] = \int d^d \mathbf{x} \sum_\alpha \left[\frac{r}{2} [S^\alpha(\mathbf{x})]^2 + \frac{1}{2} [\nabla S^\alpha(\mathbf{x})]^2 + \frac{u}{4!} [S^\alpha(\mathbf{x})]^2 \sum_\beta [S^\beta(\mathbf{x})]^2 + \dots \right] \quad (3.13)$$

Sceglieremo invece di sviluppare una descrizione dinamica del sistema, nei termini di un'equazione di Langevin non lineare.

Ricordiamo che la regione correlata di dimensione ξ aumenta notevolmente di dimensioni nelle vicinanze della transizione di fase e che il tempo di rilassamento per tale regione cresce con il proprio volume. Per tali motivi è lecito attendersi che la scala temporale t_c per il rilassamento delle fluttuazioni del parametro d'ordine cresca anch'essa all'avvicinarsi di T_C :

$$t_c(\tau) \sim \xi(\tau)^z \sim |\tau|^{-z\nu}, \quad (3.14)$$

dove z è l'esponente critico *dinamico* che regola il rallentamento critico alla transizione. Generalmente $z \geq 1$.

Dato che le frequenze di rilassamento scalano come

$$\omega_c(\tau) = \frac{1}{t_c(\tau)} \sim |\tau|^{z\nu} \quad (3.15)$$

è possibile utilizzare $\mathbf{p} = \mathbf{q}\xi$ per generalizzare (3.14) e formulare una ipotesi dinamica di scala per la relazione di dispersione del parametro d'ordine:

$$\omega_T(\tau, \mathbf{q}) = |\mathbf{q}|^z \hat{\omega}_\pm(\mathbf{q}\xi). \quad (3.16)$$

Definendo le variabili $s = \frac{t}{t_c(\tau)}$ oppure $s = \frac{\omega}{\omega_c(\tau, \mathbf{q})}$, si possono scrivere le leggi di scala dinamiche:

$$\langle S(\tau, t) \rangle = |\tau|^\beta \hat{S}\left(\frac{t}{t_c(\tau)}\right), \quad \text{con} \quad \begin{cases} \hat{S}(s \rightarrow \infty) = \text{cost} \\ \hat{S}(s \rightarrow 0) \sim s^{-\frac{\beta}{z\nu}} \end{cases} \quad (3.17)$$

e, per $\tau = 0$,

$$\langle S(t) \rangle \sim t^{-\alpha}, \quad (3.18)$$

dove $\alpha = \frac{\beta}{z\nu} = \frac{1}{2z}(d - 2 + \eta)$.

Procedendo allo stesso modo si ottiene la relazione per la suscettività (dinamica) del parametro d'ordine (funzione di risposta):

$$\chi(\tau, \mathbf{q}, \omega) = |\mathbf{q}|^{-2+\eta} \hat{\chi}_\pm(\mathbf{q}\xi, \omega\xi^z). \quad (3.19)$$

Applicando il teorema di fluttuazione dissipazione, otteniamo ora la funzione di correlazione dinamica:

$$C(\tau, \mathbf{q}, \omega) = |\mathbf{q}|^{-z-2+\eta} \hat{C}_\pm(\mathbf{q}\xi, \omega\xi^z) \quad (3.20)$$

che, ritrasformata per spazio e tempo reali, è

$$C(\tau, \mathbf{x}, \omega) = |\mathbf{x}|^{-(d-2+\eta)} \tilde{C}_\pm\left(\frac{\mathbf{x}}{\xi}, \frac{t}{\xi^z}\right). \quad (3.21)$$

Da tale formula, nel caso $t = 0$, ritroviamo l'eq. (3.4).

Verso una Teoria di Campo

Il rallentamento critico vicino alla transizione di fase fornisce una separazione delle scale dei tempi. Effettuando l'ipotesi che non vi siano altre quantità conservate nel sistema (in modo che il parametro d'ordine sia l'unico modo

lento) possiamo dare una descrizione coarse-grained per grandi lunghezze d'onda e tempi lunghi, concentrandoci solamente sulla cinetica del parametro d'ordine e sommando via l'effetto di ogni altro grado di libertà più veloce in un termine di rumore stocastico.

Ciò porta ad un equazione di Langevin mesoscopica per le variabili lente $S^\alpha(\mathbf{x}, t)$:

$$\frac{\partial S^\alpha(\mathbf{x}, t)}{\partial t} = F^\alpha[S](\mathbf{x}, t) + \zeta^\alpha(\mathbf{x}, t). \quad (3.22)$$

Nel caso più semplice, la forza sistematica $F[S]$ è data dalla dinamica di rilassamento verso la configurazione d'equilibrio:

$$F^\alpha[S](\mathbf{x}, t) = -D \frac{\delta \mathcal{H}[S]}{\delta S^\alpha(\mathbf{x}, t)}, \quad (3.23)$$

dove D è il coefficiente di rilassamento e $\mathcal{H}[S]$ è l'hamiltoniana che governa la transizione di fase.

Per quanto riguarda le forze stocastiche, possiamo assumere la forma più conveniente: un rumore bianco gaussiano a media nulla con secondo momento all'equilibrio dato dalle relazioni di Einstein

$$\langle \zeta^\alpha(\mathbf{x}, t) \rangle = 0 \quad (3.24a)$$

$$\langle \zeta^\alpha(\mathbf{x}, t) \zeta^\beta(\mathbf{x}', t') \rangle = 2k_B T D \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \delta^{\alpha\beta}. \quad (3.24b)$$

La seconda delle equazioni precedenti in particolare garantisce che $\mathcal{P}[S, t \rightarrow \infty] \rightarrow \mathcal{P}_{eq}[S]$.

Le equazioni (3.22), con (3.23), l'hamiltoniana LGW (eq. (3.13)) ed la correlazione del rumore forniscono il modello A per il rilassamento di un campo $O(n)$ -simmetrico non conservato [30].

Modelli con Dinamica Diffusiva

Se il parametro d'ordine è conservato, dobbiamo considerare anche l'equazione di continuità

$$\partial_t S^\alpha + \nabla J^\alpha = 0 \quad (3.25)$$

in cui la corrente conservata è generalmente assunta data dal gradiente del campo S^α

$$J^\alpha = -D \nabla S^\alpha + \dots \quad (3.26)$$

In conseguenza di questo le fluttuazioni del parametro d'ordine rilassano diffusivamente verso l'equilibrio con coefficiente di diffusione D . Il relativo modello B per la dinamica critica di rilassamento di un parametro

d'ordine conservato può essere ottenuto con la sostituzione $D \rightsquigarrow D(i\nabla)^a$ in (3.23), (3.24) dove $a = 0$ nel caso non conservativo e $a = 2$ in quello conservativo.

Più esplicitamente

$$\begin{aligned} \frac{\partial S^\alpha(\mathbf{x}, t)}{\partial t} &= -D(i\nabla)^a \frac{\delta H[S]}{\delta S^\alpha(\mathbf{x}, t)} + \zeta^\alpha(\mathbf{x}, t) = \\ &- D(i\nabla)^a [r - \nabla^2 + \frac{u}{6} \sum_{\beta} [S^\beta(\mathbf{x})]^2] S^\alpha(\mathbf{x}, t) + D(i\nabla)^a h^\alpha(\mathbf{x}, t) + \zeta^\alpha(\mathbf{x}, t) \end{aligned} \quad (3.27)$$

con

$$\langle \zeta^\alpha(\mathbf{x}, t) \zeta^\beta(\mathbf{x}', t') \rangle = 2k_B T D(i\nabla)^a \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \delta^{\alpha\beta}. \quad (3.28)$$

Come si nota chiaramente, già la sola presenza di una legge di conservazione per il parametro d'ordine comporta una differente descrizione dinamica per sistemi comunque caratterizzati dallo stesso comportamento stocastico all'equilibrio.

Il Gas su Reticolo Bidimensionale

Soffermiamoci ora sul modello d'equilibrio relativo al DLG, ovvero il gas su reticolo bidimensionale. Per tale modello una possibile scelta per il parametro d'ordine, basata sulla letteratura esistente e supportata dalle simulazioni numeriche al riguardo, ricade sul campo scalare della densità di particelle $\rho = \rho(\mathbf{x}, t)$ che esprimeremo, con cambio di notazione $\phi(\mathbf{x}, t) = 2\rho(\mathbf{x}, t) - 1$, nel linguaggio dei sistemi magnetici.

Sulla base di ciò riscriviamo l'hamiltoniana di Landau–Ginzburg–Wilson (eq. (3.13)) come

$$\mathcal{H}[\phi] = \int d^d \mathbf{x} \left\{ \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} \tau \phi^2 + \frac{u}{4!} \phi^4 + \dots \right\} \quad (3.29)$$

dove $\tau \propto |T - T_C|$ misura la deviazione dalla criticità e il coefficiente (positivo) u è necessario per la stabilità nella fase condensata e descrive la magnetizzazione spontanea. Inoltre $\mathcal{H}[\phi]$ è pari in ϕ ed isotropica in \mathbf{x} per rispettare le simmetrie del modello microscopico.

Da qui la rispettiva equazione del moto

$$\frac{\partial}{\partial t} \phi = \nabla_i \left\{ \lambda \nabla_i \frac{\delta \mathcal{H}}{\delta \phi} + \xi_i \right\}. \quad (3.30)$$

Le equazioni precedenti definiscono completamente il modello e da esse possono essere estratte tutte le informazioni sulle proprietà critiche statiche e/o dinamiche, per esempio attraverso un'analisi col gruppo di rinormalizzazione intorno alla dimensione critica d_C , che per l'ILG è 4.

Simmetrie Microscopiche

Ricordiamo qui anche le simmetrie possedute dal modello microscopico che sono soddisfatte da tali equazioni. Esse sono importanti in quanto svolgono un ruolo di primo piano nel determinare il comportamento del sistema, e quindi le proprietà universali. Per tale motivo dovranno essere rispettate da ogni teoria mesoscopica, se vuole appartenere alla stessa classe d'universalità del modello microscopico.

Le transizioni del modello sono invarianti per le seguenti operazioni:

T: traslazioni, modulo la periodicità del reticolo;

C: coniugazione di carica ed inversione del campo, consistente nello scambiare le particelle coi buchi $n_i \leftrightarrow (1 - n_i)$ e nell'invertire il campo $\mathbf{E} \leftrightarrow -\mathbf{E}$;

R: riflessione lungo la direzione del campo x_{\parallel} ed inversione del campo, consistente in $x_{\parallel} \leftrightarrow -x_{\parallel}$ seguita da $\mathbf{E} \leftrightarrow -\mathbf{E}$.

Le ultime due simmetrie possono essere combinate ad ottenere l'invarianza per inversione di carica e riflessione sulla direzione del campo (CR).

Passiamo ora a vedere come sia possibile adattare la teoria sopra esposta nel caso di un sistema lontano dall'equilibrio. Per fare ciò ci soffermeremo sul caso in esame, in quanto una teoria così generale, come quella per i modelli dinamici A e B, purtroppo ancora non esiste.

3.2 SISTEMI LONTANI DALL'EQUILIBRIO

Contrariamente al caso precedente, nei sistemi con uno stato stazionario di non equilibrio le proprietà dinamiche non possono essere trascurate in quanto la realizzazione dello stato stazionario dipende fortemente proprio da esse.

Se aggiungiamo al gas su reticolo il campo forzante \mathbf{E} , uno degli effetti principali è la presenza di un termine addizionale $j_{\mathbf{E}}(\rho, \mathbf{E})$ alla corrente \mathbf{J} .

Tale contributo, che rende conto dell'aumento delle transizioni nella direzione del campo, andrebbe calcolato con una procedura di coarse-graining per la dinamica microscopica. In pratica si postula per tale termine la forma più semplice compatibile con le simmetrie presenti. Tale forma, assumendo una proporzionalità diretta col campo microscopico \mathbf{E} , è

$$\mathbf{j}_{\mathbf{E}} = \mathbf{j}_{\mathbf{E}}(\rho, \mathbf{E}) = \sigma(\phi)\mathbf{E} = \{\sigma_0 + \sigma_1\phi + \sigma_2\phi^2 + \dots\}\mathbf{E}.$$

In ogni caso può essere mostrato che per motivi di simmetria l'unico elemento valido sia quello con coefficiente σ_2 . Il vettore $\mathbf{j}_{\mathbf{E}}$ deve essere isotropico ed invariante sotto la trasformazione C . Quindi tale fattore dovrà essere pari sia in ϕ che in \mathbf{E} . Considerando che la corrente, anche localmente, è nulla per reticoli completamente vuoti o pieni di particelle ($\rho = 0, 1$), la forma più semplice è data da $\rho(1 - \rho)\mathcal{E}$, avendo definito come \mathcal{E} la versione — dispari — coarse-grained del campo microscopico \mathbf{E} .

La conoscenza della forma precisa dei parametri τ , u , and \mathcal{E} presenti in $\mathcal{H}[\phi]$ non è fortunatamente necessaria ai fini del calcolo delle proprietà critiche. Quindi, assorbendo in \mathcal{E} tutte le costanti, poniamo

$$\mathbf{j}_{\mathbf{E}} = 4\rho(1 - \rho)\mathcal{E} = (1 - \phi^2)\mathcal{E}. \quad (3.31)$$

Nella regione $T \gtrsim T_C$ possiamo tralasciare i termini di ϕ in potenze maggiori di 2 in quanto si dimostrano irrilevanti in senso del gruppo di rinormalizzazione.

Intuitivamente si potrebbe pensare di modificare (3.30) aggiungendo un termine $-\nabla \cdot \mathbf{j}_{\mathbf{E}} = \mathcal{E}\partial\phi^2$ ¹ che tenga conto della corrente $\mathbf{j}_{\mathbf{E}}$. Ma tale aggiunta può essere utile solo considerando gli effetti di \mathbf{E} limitatamente agli ordini minori, in quanto la forte anisotropia che presenta la dinamica microscopica deve sopravvivere anche alla procedura di coarse-graining. Per tale motivo è lecito estendere lo spazio dei parametri, consentendo alle costanti d'accoppiamento associate ai gradienti di possedere valori differenti nelle direzioni trasversa e longitudinale al campo. Inoltre anche la matrice di rumore non sarà più uguale all'unità. A causa della simmetria CR, potenze pari (dispari) di ∂ devono essere associate a potenze pari (dispari) di \mathcal{E} . Infine, le proprietà associate allo spazio trasverso $(d - 1)$ -dimensionale conservano la loro isotropia.

Denotando ora come ∇ e ∂ i gradienti nelle direzioni trasversa e longitudinale, e dividendo egualmente la corrente in ξ e ζ , otteniamo un'equazione

¹qui e per tutto il resto del libro ∂ denoterà il gradiente nella direzione del campo

di Langevin data da

$$\begin{aligned} \frac{\partial}{\partial t} = \lambda \left\{ (\tau_{\perp} - \nabla^2) \nabla^2 \phi + (\tau_{\parallel} - \alpha_{\parallel} \partial^2) \partial^2 \phi - 2\alpha_{\times} \partial^2 \nabla^2 \phi \right. \\ \left. + \frac{u}{3!} (\nabla^2 \phi^3 + \kappa \partial^2 \phi^3) + \varepsilon \partial \phi^2 \right\} - (\nabla \cdot \xi + \partial \zeta) \end{aligned} \quad (3.32)$$

dove i coefficienti di $\nabla^4 \phi$ sono stati assorbiti in λ e quelli di $\nabla^2 \phi^3$ in u .

Ora \mathbf{N}_{ij} è ancora diagonale, ma è isotropica solo nel sottospazio trasverso, con valori n_{\parallel} e n_{\perp} . Al posto di (3.24), ora abbiamo

$$\langle \nabla \cdot \xi(\mathbf{x}, t) \nabla' \cdot \xi(\mathbf{x}', t') \rangle = n_{\perp} (-\nabla^2) \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \quad (3.33a)$$

$$\langle \partial \zeta(\mathbf{x}, t) \partial \zeta(\mathbf{x}', t') \rangle = n_{\parallel} (-\partial^2) \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \quad (3.33b)$$

Tutti i parametri presenti in queste equazioni sono funzioni di J , E e T ; dato che è necessario che tale dipendenza venga esplicitata, daremo qui qualche proprietà di base.

Le equazioni (3.33) e (3.32) definiscono in realtà un sistema all'equilibrio (o ivi vicino) con costanti d'accoppiamento e/o coefficiente di trasporto anisotropici. Il FDT è comunque tuttora soddisfatto per cui, al meno, la matrice del rumore deve essere proporzionale alla matrice di diffusione e quindi vale $\frac{\tau_{\perp}}{\tau_{\parallel}} = \frac{n_{\perp}}{n_{\parallel}}$. A differenza del caso forzato, in cui il campo mantiene il sistema sempre lontano dall'equilibrio, sebbene esso possa raggiungere uno stato stazionario. Il FDT non sarà più rispettato ed avremo in generale

$$\frac{\tau_{\perp}}{\tau_{\parallel}} \neq \frac{n_{\perp}}{n_{\parallel}}. \quad (3.34)$$

All'equilibrio τ_{\parallel} e τ_{\perp} grazie al FDT sono entrambi proporzionali all'inverso della suscettività statica e, sebbene in generale con ampiezze differenti, entrambi si annullano per $T \simeq T_C$ con lo stesso esponente critico γ . Se non sussiste più l'appoggio dello FDT, si hanno tre possibili casi all'occorrenza della criticità:

- a) $\tau_{\perp} \rightarrow 0$ e $\tau_{\parallel} > 0$;
- b) $\tau_{\perp} > 0$ e $\tau_{\parallel} \rightarrow 0$;
- c) $\tau_{\perp}, \tau_{\parallel} \rightarrow 0$.

Per la scelta corretta bisogna farsi guidare dalle conoscenze sperimentali derivate dalle simulazioni, dove si sono osservate solo interfacce parallele al

campo. Dunque avremo

$$a) \tau_{\perp} \rightarrow 0 \text{ e } \tau_{\parallel} > 0. \quad (3.35)$$

D'altra parte, in quanto varianze della corrente di rumore, sia n_{\parallel} che n_{\perp} devono rimanere positivi. Dunque la violazione del FDT "cresce" all'avvicinarsi della criticità. Oltre tale soglia, per poter spiegare alcune proprietà, si farà uso della scelta $\tau_{\perp} < 0$, $\tau_{\parallel} > 0$. Si possono dare basi più solide a tale scelta notando per esempio che τ_{\parallel} e τ_{\perp} , in quanto coefficienti delle derivate del secondo ordine, possono essere considerati come i coefficienti di diffusione nudi nelle direzioni specifiche. Tali coefficienti vengono rinormalizzati sia dalle interazioni a primi vicini, sia dal campo microscopico E . Mentre l'effetto delle prime è quello di indebolire i coefficienti fino ad annullarli per T_C , il secondo indebolisce invece tali interazioni nel sottospazio longitudinale. L'effetto collettivo sarà un più rapido annullamento di τ_{\perp} al decrescere di T .

Le basi per uno studio analitico del DLG sono date dall'equazione di Langevin (3.32), dalle correlazioni del rumore (3.33) e dalla violazione del FDT (3.34).

Nella regione critica le cose sono difficili e serve uno studio con il gruppo di rinormalizzazione.

Sopra T_C le fluttuazioni del parametro d'ordine locale sono piccole ed il sistema si comporta in maniera tale da poter ignorare le non linearità. Qui la teoria diventa Gaussiana.

Sotto T_C , ignorando i termini di rumore, si ottiene una descrizione di campo medio.

Quattro

ANALISI DEL MODELLO

4.1 SETUP SPERIMENTALE

Per effettuare lo studio del DLG quasi unidimensionale ci siamo avvalsi di simulazione numeriche Monte Carlo, basate sui metodi esposti in precedenza. Considerata la definizione del modello, è stato d'obbligo utilizzare l'algoritmo Metropolis, adattato alla dinamica diffusiva Kawasaki.

In questo caso gli spin non sono "flippati" singolarmente, ma ne viene presa una coppia e, se differenti, vengono "flippati" entrambi, mantenendo così costante la magnetizzazione. La variazione di energia non sarà quindi più data dalla somma sugli spin vicini ad un sito, pari a 4, bensì sui vicini alla coppia di siti, che sono 6. Cambiano dunque anche i possibili valori di $\Delta\mathcal{H}$: invece di $\{-8J, -4J, 0, +4J, +8J\}$, avremo $\{-24J, -16J, -8J, 0, +8J, +16J, +24J\}$.

Per evidenziare la differente dinamica, possiamo pensare che, invece di prendere in considerazione un singolo sito per volta, stiamo prendendo in considerazione un "legame" tra due siti. Se il legame collega due siti con valori differenti, allora proveremo a invertire tali valori in base alla differenza di energia "attorno al legame".

È noto che, vicino al punto critico, il tempo di autocorrelazione (tempo di rilassamento) dei processi dinamici aumenta vertiginosamente (rallentamento critico). In pratica il tempo di autocorrelazione τ è perlopiù proporzionale a $\min\{L, \xi^z\}$, dove L è la dimensione lineare del sistema, ξ la lunghezza di correlazione di un sistema infinito alla stessa temperatura e z l'esponente critico dinamico (per Ising $z \approx 2$ in ogni dimensione).

Per tentare di risolvere tale problema sono solitamente utilizzati 2 approcci differenti: affrontarlo dal punto di vista computazionale o dal punto di vista fisico.

1. Mantenendo in considerazione l'algoritmo Metropolis abbiamo a disposizione tre metodi per tentare di velocizzare la simulazione:
 - a) possiamo usare macchine più veloci;
 - b) possiamo usare macchine costruite appositamente per tale compito (special-purpose machines);
 - c) possiamo riorganizzare la struttura dei dati e l'algoritmo per aumentarne l'efficienza.

I primi due sono legati all'hardware, il terzo al software.

2. Il secondo metodo è lavorare su sistemi fisici differenti che hanno un rallentamento minore. In tale categoria rientrano i sistemi ad update collettivo, come quello di Swendsen e Wang [59].

Come abbiamo già avuto occasione di notare in precedenza, a causa delle sue caratteristiche che vincolano la scelta della dinamica, non è possibile ricorrere ad altre dinamiche, che permetterebbero di evitare il rallentamento che occorre nelle simulazioni in un intorno della zona critica.

Non ci è possibile, essendo legati alle macchine a nostra disposizione, neppure tentare di sfruttare i punti 1a ed 1b.

L'unico modo per ottenere un aumento della velocità delle simulazioni, rimane dunque tentare di trovare un'implementazione della dinamica più efficiente. Per tale motivo la nostra scelta è ricaduta sulle tecniche di *multispin coding*¹.

Il Multi Spin Coding

Il multispin coding, abbreviato in msc, deve il suo nome e la sua fortuna al fatto di riuscire a gestire contemporaneamente più variabili a due valori (spesso associate ad uno spin up-down), riuscendo così in un certo qual modo a "parallelizzare" la simulazione di un sistema su reticolo, su macchine che non sono parallele.

¹In quanto segue, diamo una veloce descrizione del metodo. Se il lettore fosse curioso di conoscere le origini di tale algoritmo, nutriamo la speranza che possa soddisfarla consultando l'appendice.

Sia K il numero di bit di una parola di un elaboratore. Ognuno di questi bit può trovarsi in uno degli unici due stati binari possibili: 0 e 1. Risulta evidente la somiglianza tra di essi e le particelle del nostro modello, anch'esse associate all'insieme di valori $\{0, 1\}$. Come poter sfruttare questo fatto? Semplicemente utilizzando un singolo bit per tener conto dello stato di un singolo sito del reticolo. In questo modo, all'interno di una parola possono "trovar posto" K particelle. Già ora notiamo che così facendo, è possibile ottenere un notevole risparmio di memoria per codificare l'informazione: invece di usare una parola per bit, ne usiamo solo una frazione.

Sia ora $N = L_{\perp} \times L_{\parallel}$ il volume di un reticolo, ed assegniamo ad ogni sito del reticolo una differente parola. Immagazzinando l'informazione dei siti appartenenti ad uno stesso reticolo in parole diverse, ma nella stessa posizione all'interno della parola, saremo così in grado di far evolvere contemporaneamente K reticoli usando N parole.

Per essere più chiari, si immagini di avere parole di 4 bit e di dover effettuare una simulazione di un gas su un reticolo di volume 16. Allora, utilizzando 16 parole, riusciremo a fare evolvere 4 reticoli. Il primo bit di ogni parola sarà "dedicato" al primo reticolo, il secondo bit al secondo reticolo, e così via. Assegnando in un modo qualunque una numerazione da 1 a 16 ai siti del reticolo, e numerando in egual modo anche le parole, saremo in grado di poter tenere traccia di tutti i nostri sistemi.

Il DLG è affetto, come tutti i sistemi che subiscono una transizione di fase, da un notevole rallentamento della dinamica quando si avvicina alla temperatura critica. Come diretta conseguenza, il tempo impiegato dalle simulazione per raggiungere la distribuzione stazionaria¹ (per termalizzare, in gergo) aumenta a dismisura, andando presto a saturare le risorse computazionali o di tempo. Per ovviare a tale problema sono stati introdotti numerosi metodi numerici in grado di velocizzare questo passaggio. Purtroppo però nessuno di tali metodi è qui utilizzabile, in quanto cambierebbe la dinamica che definisce il nostro modello e, come conseguenza diretta, modificherebbe la stessa definizione del modello.

Quindi l'importanza del multispin coding risiede anche nella sua ininfluenza sulla dinamica, che lo rende applicabile anche nel nostro caso.

Sono possibili varie implementazioni del multispin coding: i vari reticoli possono essere posti tutti a temperature uguali o differenti, può essere utilizzata un'identica sequenza di numeri casuali per tutti i reticoli, o una

¹O meglio, per giungere in uno stato in cui *noi* pensiamo sia stato raggiunto lo stato stazionario.

differente per ognuno. Mentre scelte come la temperatura sono semplicemente soggette al gusto personale o alle specifiche misurazioni da effettuare, la scelta della sequenza di numeri casuali comporta una variazione di comportamento non indifferente.

Prima di parlare delle scelte fatte, diamo una descrizione del multispin coding all'interno della teoria delle catene di Markov.

Il multispin coding effettuato con sequenze di numeri casuali differenti per ogni reticolo può essere pensato come un coupling indipendente. È possibile che le catene si trovino più volte nella stessa configurazione, ma ogni volta ognuna proseguirà per conto proprio. La coalescenza non avrà influenza sugli osservabili, per esempio, dato che i reticoli non rimanendo uguali hanno una probabilità bassa di trovarsi numerosi in una medesima configurazione.

Se i numeri casuali sono gli stessi per tutti i reticoli, avremo invece un accoppiamento dipendente che porterà, con probabilità unitaria, alla coalescenza. Da qui in poi tutti i reticoli presenteranno la medesima configurazione. La conseguenza più immediata di tutto ciò è l'inutilità di tale algoritmo, in quanto si avrebbe solo uno spreco di risorse computazionali e, nel caso in cui uno non si accorgesse della coalescenza, una netta sottostima degli errori.

Generalmente comunque tale tipo di approccio non risente della coalescenza in maniera tale da influenzare le misure, in quanto i tempi di coalescenza sono talmente elevati, rispetto ai tempi tipici d'osservazione, che la loro influenza risulta molto ridotta, se non nulla.

In base a tali considerazioni abbiamo scelto di utilizzare per le simulazioni un algoritmo basato su tale tecnica¹, giudicando nettamente maggiori i vantaggi dovuti all'incremento di velocità derivante dal dover generare solamente una sequenza di numeri casuali, rispetto ai "rischi" di una coalescenza².

Per effettuare lo studio del DLG quasi-unidimensionale, abbiamo utilizzato i computer presenti nella facoltà, in particolare quelli appartenenti al cluster denominato "farm" e quelli presenti nelle aule di calcolo, su cui era da poco installato un programma (Condor) per la gestione di proces-

¹Il programma che abbiamo utilizzato è stato creato da Caracciolo e successivamente sviluppato da Gubinelli. È già stato utilizzato, per esempio, in [10].

²Ci teniamo a ricordare che, come vedremo più compiutamente in seguito, l'unico effetto derivante da tali rischi è una sottostima degli errori, e non la generazione di valori erronei.

si. Su tali macchine, dotate di architettura i586, il linguaggio C++ alloca una variabile "int" utilizzando 32 bit. Utilizzando tale variabile per contenere le informazioni sui singoli bit, siamo stati in grado di far evolvere contemporaneamente 32 reticoli.

Utilizzando un'unica sequenza di numeri casuali, identica per ogni reticolo, significa nel nostro caso guadagnare circa un fattore 30 come incremento di velocità computazionale.

Durante la raccolta dei dati, ci siamo però accorti che, contrariamente alle aspettative, molto spesso reticoli differenti appartenenti allo stesso gruppo di 32 fornivano dati identici, chiaro segno, prima ancora che di coalescenza, del fatto che qualcosa non funzionava come avremmo voluto. Abbiamo così inizialmente rianalizzato il codice alla ricerca di qualche possibile errore ma, non trovandolo, ci siamo dovuti arrendere all'evidenza che la coalescenza era avvenuta ben prima di quanto ci aspettassimo.

In qualche modo il campo introduce un fenomeno il cui effetto è la tendenza ad ordinare più rapidamente i vari reticoli. Tale effetto va ricercato nella dinamica microscopica ed in particolare nel fatto che le mosse equiverse al campo (opposte al campo) sono sempre accettate (rifiutate). Iniziamo col prendere un singolo reticolo e consideriamo un legame trasverso. Se il legame è saturo¹ la configurazione finale sarà banalmente identica a quella iniziale. Se invece il legame collega una particella ed un posto vuoto, allora sussiste una possibilità di salto. A seconda dei siti nell'intorno di tale legame, il salto sarà effettuato con una probabilità data dall'usuale frequenza di Metropolis $\min\{1, \exp(-\beta\Delta\mathcal{H})\}$. Introduciamo ora la *distanza di Hamming*² d_H tra due reticoli come il numero di siti differenti tra di essi. Tale distanza ci servirà per misurare la distanza dalla coalescenza completa, dove $d_H = 0$.

Consideriamo reticoli sovrapposti. Se almeno uno dei due legami è saturo, qualunque sia lo stato finale dell'altro dopo un tentativo di salto risulta ininfluenza ai fini della coalescenza, in quando $\Delta d_H = 0$. Se invece i legami sono entrambi non saturi, l'evoluzione temporale potrebbe portare ad un "avvicinamento" o ad un "allontanamento" dei reticoli. Consideriamo per

¹Intendiamo con *saturo* che i siti alle sue estremità sono entrambi vuoti o entrambi pieni, cosicché non possa avvenire nessuno scambio attraverso esso.

²La distanza di Hamming viene generalmente utilizzata in teoria dell'informazione ed in matematica discreta (in particolare nella teoria dei codici e dei disegni) proprio per definire la distanza tra oggetti composti da elementi binari.

4. ANALISI DEL MODELLO

primo il caso in cui le particelle sono disposte dalla stessa parte del legame, ovvero sia minima la distanza ristretta al legame stesso¹.

La probabilità che entrambe effettuino un salto, estraendo un unico numero casuale, è il minimo dei rispettivi coefficienti di Metropolis, mentre la probabilità che entrambe rimangano ferme, è data da 1 meno il massimo dei coefficienti. Per chiarire, ipotizzando che sia $\Delta_1\mathcal{H} \leq \Delta_2\mathcal{H}$ e denotando u il numero casuale estratto, si avrà:

numero casuale	particella 1	particella 2	Δd_H
$u < e^{-\beta\Delta_1\mathcal{H}}$	salta	salta	0
$e^{-\beta\Delta_1\mathcal{H}} < u < e^{-\beta\Delta_2\mathcal{H}}$	ferma	salta	± 2
$e^{-\beta\Delta_1\mathcal{H}} < u < 1$	ferma	ferma	0

Tabella 4.1: *Variazione della distanza di Hamming in funzione del numero casuale estratto.*

Con ovvia estensione di notazione, e di complessità, lo stesso procedimento può essere esteso a più reticoli. In questo caso, aumentando le combinazioni possibili presenti, l'effetto generale sarà quello di ottenere in media una variazione casuale di Δd_H .

Prendiamo ora in considerazione la medesima situazione, ma con il legame parallelo al campo. Cosa succede? Qualunque sia la configurazione di partenza, in presenza di un legame non saturo, alla fine del passo temporale la particella si troverà *sempre* dalla parte del legame che punta verso il campo. Se abbiamo 2 o più reticoli, alla fine del passo temporale avremo *sempre* una serie di legami in cui quelli non saturi sono esattamente nella stessa configurazione, indipendentemente dalle configurazioni dei vicini e senza bisogno di estrarre alcun numero casuale. Come si può ben immaginare tale situazione non può che portare verso una diminuzione della distanza di Hamming tra i reticoli, o al più lasciarla inalterata se è presente al più un solo legame non saturo.

Dunque la dinamica del gas forzato su reticolo contribuisce in maniera netta ad incrementare la velocità di coalescenza rispetto ad un sistema analogo privo di campo forzante. Ed è questo il motivo per cui essa è avvenuta così rapidamente nelle nostre simulazioni.

Un altro possibile modo di studiare il fenomeno, è guardare direttamente allo spazio delle configurazioni. Nel sistema forzato, le configurazioni

¹Nel caso di distanza ristretta ad un singolo legame, $d_H \in \{0, 2\}$ e $\Delta d_H \in \{-2, 0, +2\}$.

accessibili al sistema con una probabilità non troppo piccola sono molto inferiori in numero a quelle del caso non forzato. Essendo minore lo spazio in cui i processi sono costretti ad evolvere, minore sarà anche il tempo che impiegheranno per incontrarsi.

Possiamo riassumere il tutto affermando che:

Il campo agisce come un “accoppiamento effettivo” di forza infinita tra particelle appartenenti a reticoli differenti, se queste si muovono equiverse ad \mathbf{E} .

Dopo aver compreso le cause di una così veloce coalescenza, possiamo ora indagare le modalità con cui questa coalescenza avviene. Per fare ciò abbiamo studiato le distribuzioni dei tempi di coalescenza per diversi valori di β e di L_{\perp} , riportate in fig. 4.1. Abbiamo raccolto tali valori nel modo seguente¹. Per ogni valore di (L_{\parallel}, β) , sono state effettuate almeno N simulazioni, ognuna delle quali consisteva nel far evolvere i reticoli controllando ogni Δt MCS se era avvenuta la coalescenza. In tal caso la simulazione veniva terminata ed il valore di t attuale veniva registrato. Alla fine delle N simulazioni, abbiamo calcolato la media e l'errore². Basandoci poi sulle informazioni ricavate dall'elaborazione dei dati raccolti, abbiamo concentrato le misure nella zona in cui il tempo di coalescenza per L_{\parallel} fissato mostrava un minimo.

Come risulta evidente dalla figura, le distribuzioni per i tempi di coalescenza, riuniti per valori costanti di L_{\parallel} , presentano un evidente minimo che mostra la tendenza a crescere. In base alla letteratura già presente sul valore si β_c per questo modello, pari a circa 0.35, è apparsa naturale l'ipotesi che la temperatura in cui si ha il punto di minimo per il tempo di coalescenza possa corrispondere con la temperatura critica del modello, traslata in funzione di L_{\parallel} per l'effetto di taglia finita. A livello teorico tale ipotesi è supportata da alcune considerazioni sul legame che vi è tra la coalescenza, i cui meccanismi dinamici abbiamo descritto in precedenza, e la stessa dinamica

¹I dati raccolti sono riportati nella tab 4.B in appendice a questo capitolo.

²Per una discussione sugli errori e l'affidabilità dei dati, si veda l'appendice al termine del capitolo. Ci teniamo a sottolineare che, quando non esplicitamente indicato, **gli errori come indicati nei dati e nelle figure sono calcolati secondo**

$$\delta_t = \frac{1}{\sqrt{N}} \sqrt{\frac{\sum_{i=1}^N (t_i - \bar{t})^2}{N - 1}}. \quad (4.1)$$

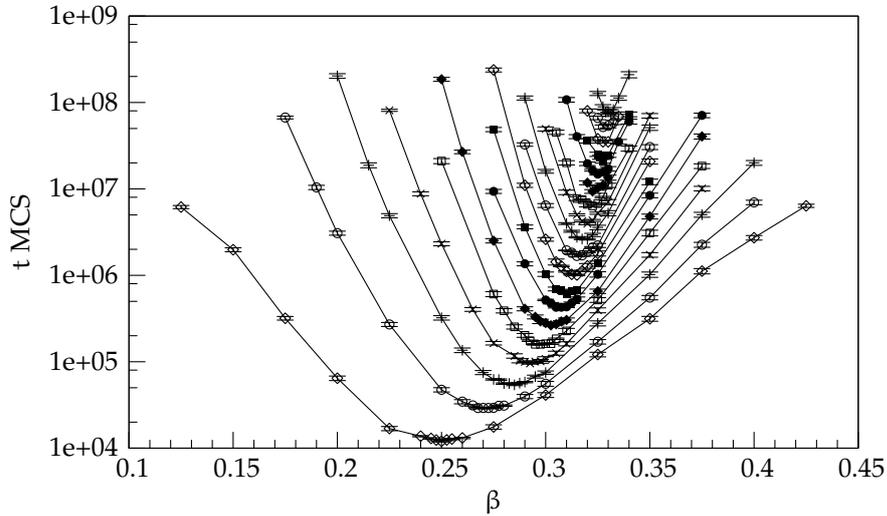


Figura 4.1: Tempi di coalescenza per reticoli $8 \times L_{\parallel}$ in funzione di β . Le varie curve collegano punti con lo stesso valore di L_{\parallel} , partendo dal basso $L_{\parallel} = 8, 12, \dots, 76, 80$. Le linee continue sono solamente di utilità visiva.

microscopica. Avendo identificato le mosse parallele al campo come principali responsabili dell'accoppiamento dei vari reticoli, e le mosse trasverse come rumore, ci chiediamo ora come queste possano essere influenzate dalle variazioni di β .

Partendo da valori di temperatura molto elevati, $\beta \ll 1$, si ha che anche le mosse in direzione trasversa vengono accettate con probabilità pressoché unitaria e dunque le frequenze relative agli spostamenti nelle varie direzioni (esclusa quella antiparallela al campo) divengono equivalenti. Tale fatto comporta un incremento del "rumore termico" che, a sua volta, rende più difficile l'accoppiamento delle particelle nei vari reticoli. Man mano che la temperatura decresce, pur rimanendo sopra la criticità, le mosse trasverse divengono sempre più improbabili. Dato che si ha una diminuzione di una causa del rallentamento nella coalescenza, quest'ultima diviene più rapida. Giunti in prossimità della temperatura critica si ha un cambio della fase del sistema e, di conseguenza, della sua dinamica microscopica. Una volta formatasi la striscia, infatti, le particelle perdono quasi ogni possibilità di potersi muovere nella direzione parallela al campo. In tale situazione la maggior parte del lavoro per portare i reticoli verso configurazioni identiche viene compiuto dalle mosse nelle direzioni trasverse. Notiamo anche

che, se le strisce si formano per reticoli differenti in posizioni reciprocamente traslate, bisognerà attendere che, per fluttuazione termica, tali strisce vengano a sovrapporsi almeno parzialmente. Da questo punto in poi, ogni ulteriore diminuzione della temperatura comporta un'ulteriore diminuzione della "mobilità" trasversale e, indirettamente, un allungamento del tempo di coalescenza.

Sorretti da queste considerazioni, abbiamo deciso di ricercare delle conferme sperimentali alla nostra identificazione della temperatura per cui si ha il minimo del tempo di coalescenza $t_{\min}(L_{\parallel})$ con la temperatura critica del sistema. Tali conferme possono essere condotte in due differenti modi. Il primo è lo studio dell'andamento di $t_{\min}(L_{\parallel})$ attraverso studi di taglia finita, per poter avere un confronto con dati noti nella letteratura. L'altro metodo, di carattere più generale, consiste nell'applicazione del multispin coding con coupling dipendente anche in altri modelli con dinamica simile al nostro. Affronteremo i due argomenti in questa sequenza.

Studio di taglia finita

Abbiamo verificato che $\beta(t_{\min}(L_{\parallel}))$ segue l'andamento di taglia finita adottato usualmente per β_C :

$$\beta_C(L_{\parallel}) = \beta_C(\infty) - AL_{\parallel}^{-1/\nu_{\parallel}}, \quad (4.2)$$

trovando un buon accordo, come riportato in fig. 4.2, per i valori

$$\begin{aligned} \beta_C(\infty) &= 0.3765, \\ \nu_{\parallel} &= 0.4298, \\ A &= 0.3065. \end{aligned} \quad (4.3)$$

Riportiamo (in fig. 4.3) per un confronto anche il fit effettuato fissando il valore di ν_{\parallel} su quello — esatto — previsto dalla teoria per il caso bidimensionale, $\nu_{\parallel} = 3/2$. In questo caso i valori dei parametri di scala che si adattano meglio ai dati risultano essere:

$$\begin{aligned} \beta_C(\infty) &= 0.356, \\ \nu_{\parallel} &= 3/2 \quad (\text{fissato}), \\ A &= 0.48. \end{aligned} \quad (4.4)$$

4. ANALISI DEL MODELLO

Tabella 4.2: Dati corrispondenti al valore di β che minimizza t . Qui N è il numero di campioni generati e Δt è il tempo (in Monte Carlo Step) che trascorre tra due successivi controlli per la coalescenza.

L_{\parallel}	$\beta(t_{\min})$	t_{\min}	δt_{\min}	N	Δt
8	0.50	1.251e+04	2.222e+02	1000	1e+03
12	0.545	2.918e+04	5.214e+02	1000	1e+03
16	0.57	5.395e+04	1.014e+03	1000	1e+03
20	0.585	9.559e+04	1.722e+03	1000	1e+03
24	0.595	1.5893e+05	2.994e+03	1000	1e+04
28	0.605	2.636e+05	4.835e+03	1000	1e+04
32	0.61	4.266e+05	7.745e+03	1000	1e+03
36	0.62	6.120e+05	1.163e+04	1000	1e+05
40	0.63	1.008e+06	1.897e+04	1000	1e+05
44	0.635	1.705e+06	4.529e+04	512	1e+05
48	0.635	2.612e+06	4.670e+04	1000	1e+05
52	0.64	3.958e+06	1.121e+05	500	1e+05
56	0.645	6.350e+06	1.192e+05	1005	1e+05
64	0.65	1.480e+07	3.558e+05	504	1e+05
68	0.655	2.093e+07	1.075e+06	100	1e+05
72	0.66	3.507e+07	1.166e+06	311	1e+05
80	0.66	7.859e+07	4.232e+06	100	1e+06

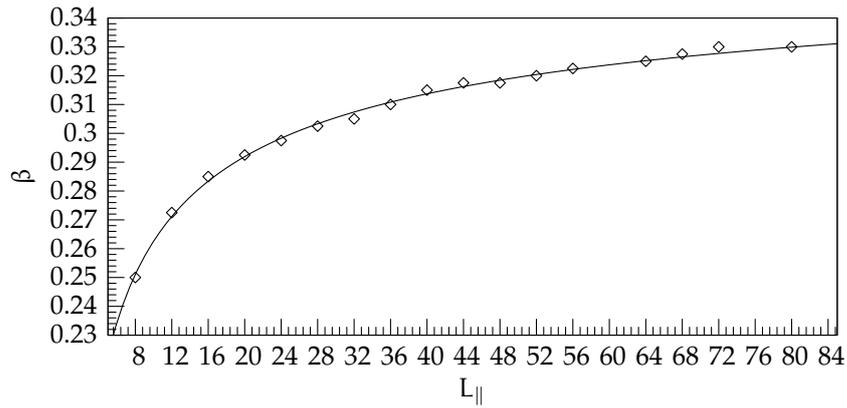


Figura 4.2: Valore di β per cui occorre il minimo del tempo di coalescenza, per reticoli di dimensioni $8 \times L_{\parallel}$, in funzione di L_{\parallel} . La curva raffigurata è la funzione di scala $\beta_C(L_{\parallel}) = \beta_C(\infty) - AL_{\parallel}^{-1/\nu_{\parallel}}$ per $\beta_C(\infty) = 0.3765$, $\nu_{\parallel} = 0.4298$ e $A = 0.3065$.

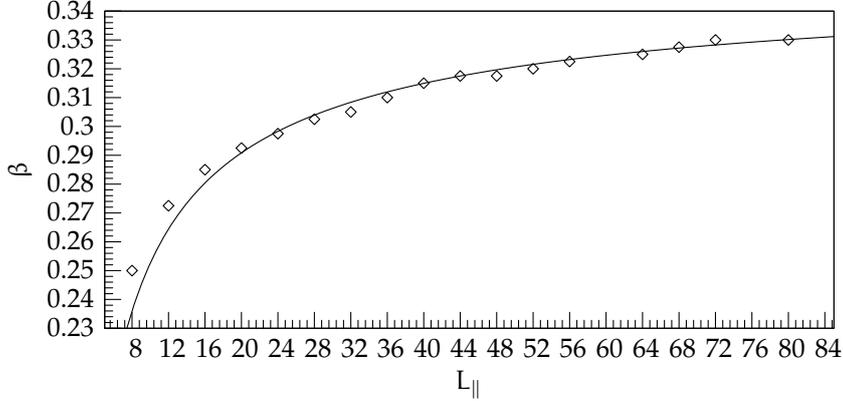


Figura 4.3: Valore di β per cui occorre il minimo del tempo di coalescenza, per reticoli di dimensioni $8 \times L_{\parallel}$, in funzione di L_{\parallel} . La curva raffigurata è la funzione di scala $\beta_c(L_{\parallel}) = \beta_c(\infty) - AL_{\parallel}^{-1/\nu_{\parallel}}$ per $\beta_c(\infty) = 0.356$, $\nu_{\parallel} = 3/2$ (esatto) e $A = 0.48$.

Da tale analisi si evidenzia che il valore ottenuto per $\beta_c(\infty)$ è consistente col valore riportato da Caravenna¹ in [11], pari a circa 0.35. Purtroppo, come risulta evidente dalla fig. 4.2, per bassi valori di L_{\parallel} , per cui un eventuale effetto di taglia finita è più forte, la zona in cui giace il minimo mostra un plateau abbastanza marcato. Per valori più alti di L_{\parallel} , la zona critica risulta più ristretta, ma non siamo stati in grado, per limiti di tempo e di potenza di calcolo, di raggiungere precisioni molto elevate. Riteniamo comunque di poter interpretare i risultati fin qui ottenuti come un'evidenza della bontà della nostra identificazione di $\beta(t_{\min}(L_{\parallel}))$ con $\beta_c(L_{\parallel})$, sebbene i risultati numerici fin qui ottenuti non consentano la determinazione di $\beta_c(\infty)$ e ν con precisione paragonabile ai risultati presenti in letteratura per il DLG².

Per completare l'analisi di taglia finita, abbiamo controllato l'andamento del minimo dal tempo di coalescenza in funzione della geometria del reticolo, per ottenere uno scaling completo.

L'accordo è stato raggiunto confrontando i dati sperimentali con la fun-

¹Tale studio è di carattere quasi esclusivamente qualitativo e rivolto soprattutto alla determinazione del diagramma di fase e non ad una valutazione precisa della temperatura critica.

²La precisione di tali risultati, ottenuti tramite l'utilizzo di supercalcolatori, non sarebbe stata comunque accessibile nemmeno con i metodi "tradizionali", date le risorse a nostra disposizione.

zione di scala

$$t_{\min}(L_{\parallel}) = C L_{\parallel}^{\alpha} e^{L_{\parallel}/L_0}, \quad (4.5)$$

per valori dei parametri corrispondenti a

$$\begin{aligned} \alpha &= 1, \\ C &= 800, \\ L_0 &= 11.15. \end{aligned} \quad (4.6)$$

Utilizzando tutte le informazioni in nostro possesso possiamo così ottenere una descrizione di scala che porta ad un collasso più che soddisfacente delle distribuzioni relative ai differenti valori di L_{\parallel} , sia in relazione a β , che al tempo t (*cfr.* fig. 4.5).

Applicazione al gas su reticolo

Per avere la certezza che quanto si è riscontrato grazie all'uso del coupling nel multispin coding tragga origine dalla dinamica microscopica dei reticoli sottostanti e non sia invece un fenomeno particolare seppur incredibilmente fortuito, presente solamente nel modello fin qui esaminato, un buon metodo di controllo è l'applicazione dello stesso procedimento ad altri modelli. In particolare abbiamo rivolto la nostra attenzione al gas su reticolo. Esso presenta una dinamica sufficientemente differente — per esempio non avviene il “congelamento” delle mosse relativamente ad un'unica direzione — ma

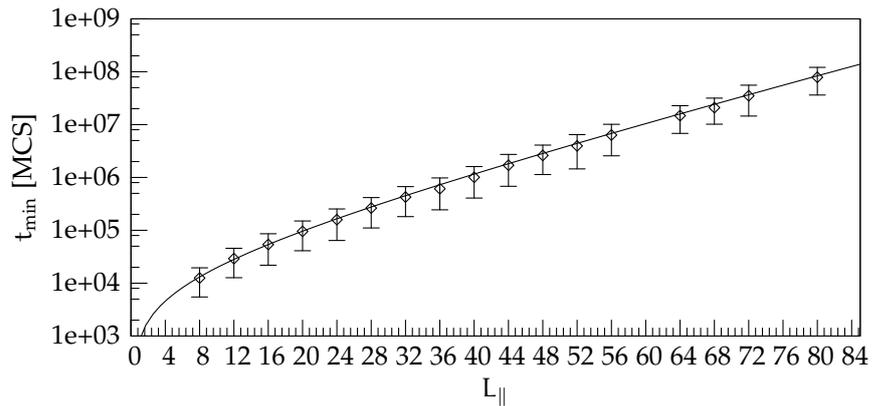


Figura 4.4: Minimo del tempo di coalescenza, per una data geometria $8 \times L_{\parallel}$ in funzione di L_{\parallel} . La curva riporta la funzione di scala $t_{\min}(L_{\parallel}) = C L_{\parallel}^{\alpha} e^{L_{\parallel}/L_0}$, con $\alpha = 1$, $C = 800$ e $L_0 = 11.15$.

abbastanza semplice da poter essere analizzata analogamente al caso precedente. Inoltre l'adattamento del programma non risulta eccessivamente laborioso.

La dinamica del gas su reticolo è assimilabile senza problemi a quella del DLG, se restringiamo l'analisi alle mosse trasverse al campo. Rimandiamo dunque a quella descrizione per il calcolo delle probabilità di transizione. Da tale analisi si evince che in questo caso, cessando l'effetto "d'accoppiamento" del campo E , la coalescenza dei vari reticoli è determinata quasi esclusivamente dalle fluttuazioni termiche. Per essere più precisi si può affermare che, interpretando l'evoluzione dei reticoli come una traiettoria nello spazio delle configurazioni del modello, non vi è più un'attrazione reciproca tra le varie traiettorie. Ogni traiettoria seguirà un percorso quasi¹ indipendente

¹L'uso del lemma *quasi* è dovuto alla presenza di un accoppiamento indotto dall'utilizzo della medesima sequenza di numeri casuali, che corrisponde ad effettuare i tentativi di salto negli stessi siti in tutti i reticoli. Utilizzando sequenze differenti per ogni reticolo, la coalescenza sarebbe esclusivamente dovuta alle fluttuazioni della posizione delle particelle ed avverrebbe in tempi ancora più elevati.

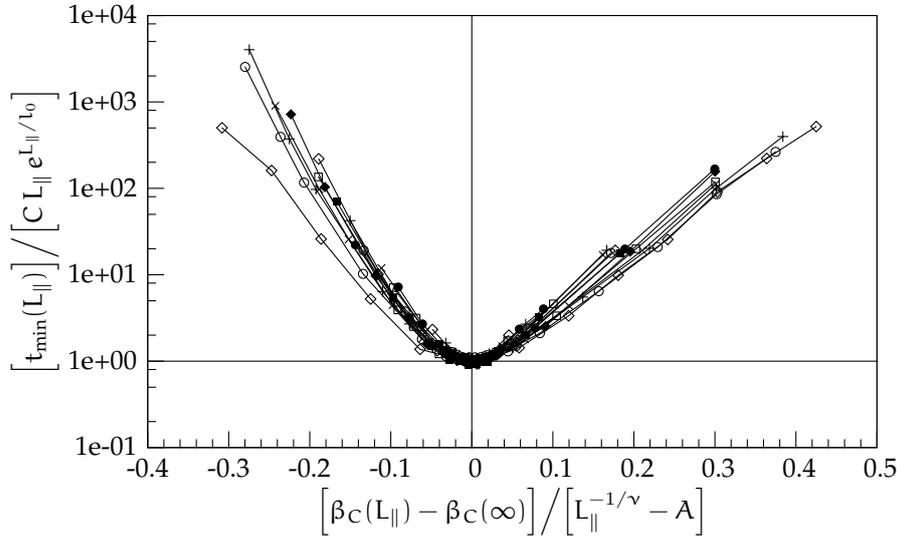


Figura 4.5: Tempi di coalescenza per reticoli $8 \times L_{\parallel}$ in funzione di β , riscaldati con le funzioni di scala (4.2) e (4.5). Le varie curve collegano punti con lo stesso valore di L_{\parallel} , da 8 a 80 ad intervalli $\Delta L_{\parallel} = 4$. Le linee continue sono solamente di utilità visiva. I valori dei parametri di scala sono, per le ascisse e le ordinate rispettivamente, $\beta_C(\infty) = 0.3765$, $A = 0.3065$ e $\nu_{\parallel} = 0.4298$; $C = 800$ e $L_0 = 11.15$.

finché non giungerà a contatto con la traiettoria di un altro reticolo. Da qui in poi le due traiettorie proseguiranno coalite.

Anche in questo caso la transizione di fase, che corrisponde ad una soppressione generale della mobilità delle particelle, svolge un ruolo di separazione tra due zone caratterizzate da andamenti distinti del tempo di coalescenza. Se la nostra interpretazione è corretta, dovremmo ritrovare, come nel caso precedente, una diminuzione del tempo di coalescenza al crescere di β fino a raggiungere la condensazione nella zona critica. da qui in avanti, ad un'ulteriore diminuzione di β , corrisponderà invece un aumento del tempo di coalescenza.

Tale interpretazione si rivela corretta anche in questo caso. Infatti dai dati raccolti si ottiene una distribuzione dei tempi di coalescenza come in fig. 4.6.

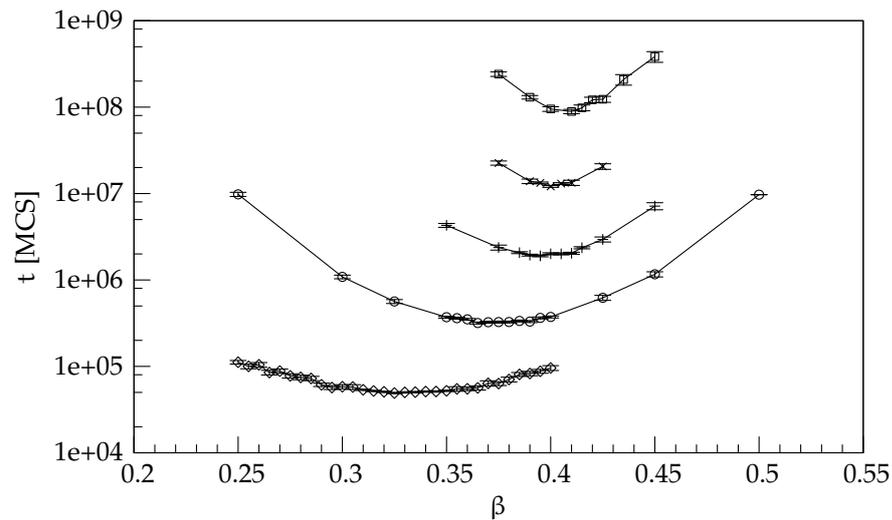


Figura 4.6: *Tempi di coalescenza per reticoli $L \times L$ in funzione di β . Le curve collegano valori con L costante; rispettivamente, partendo dal basso $L = 6, 8, 10, 12, 14$.*

Dal confronto con fig. 4.1, risulta evidente come i tempi di coalescenza siano aumentati notevolmente, circa di due ordini di grandezza, in accordo con la nostra analisi. Purtroppo questo fatto, unito al poco tempo¹ a nostra disposizione non ci ha permesso di simulare una varietà di reticoli paragonabile al caso precedente. Nonostante ciò siamo riusciti ad ottenere dei

¹A titolo puramente indicativo, il minimo tempo fisico necessario per ottenere un singolo dato per reticoli 16×16 è risultato pari a 36 ore.

risultati notevoli nell'ambito dell'analisi di taglia finita. In particolare si è trovato un accordo soddisfacente, all'interno delle incertezze, tra i dati sperimentali e lo scaling di $\beta_C(L)$, ottenuto utilizzando i valori esatti noti dalla letteratura. Riportiamo qui solo il grafico di questo accordo e con lo scaling totale dei tempi di coalescenza, insieme coi valori dei parametri utilizzati.

Tabella 4.3: Dati corrispondenti al valore di β che minimizza t . Qui N è il numero di campioni generati e Δt è il tempo (in Monte Carlo step) che trascorre tra due successivi controlli per la coalescenza.

L	β	t_{\min}	δt_{\min}	N	Δt	$\delta\beta$
6	0.66	5.004e+04	8.711e+02	1000	1e+03	0.03
8	0.75	3.260e+05	8.391e+03	500	1e+03	0.02
10	0.79	1.901e+06	4.808e+04	500	1e+04	0.02
12	0.80	1.213e+07	4.744e+05	539	1e+04	0.01
14	0.82	8.809e+07	6.438e+06	100	1e+04	0.02

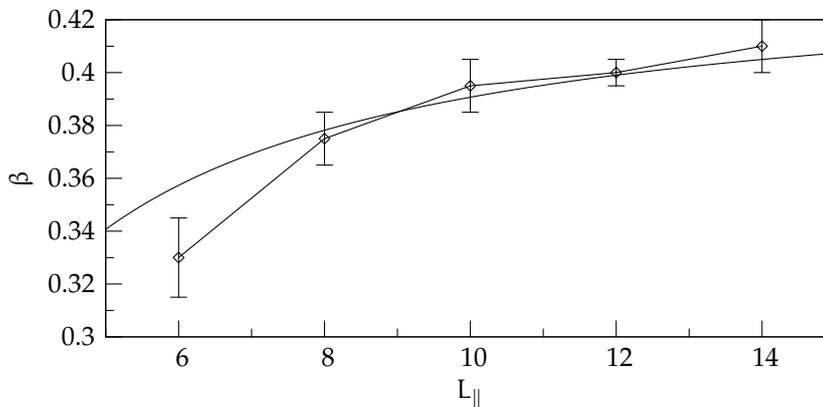


Figura 4.7: Valore di β per cui occorre il minimo del tempo di coalescenza, per una data geometria $L \times L$ in funzione di L . La curva rappresenta la funzione di scala $\beta_C(L) = \beta_C(\infty) - A L^{-1/\nu}$, coi parametri

L'unico scostamento dall'accordo risulta essere relativo ai dati del reticolo di dimensioni 6×6 . Data l'esiguità di tali dimensioni, è verosimile ritenerlo un effetto di taglia finita.

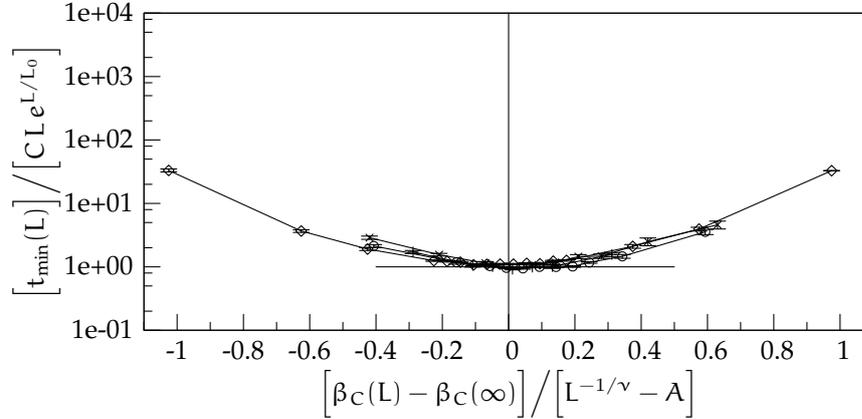


Figura 4.8: Tempi di coalescenza per reticoli $L \times L$ in funzione di β . Le varie curve uniscono punti con L costante, $L = 8, 10, 12, 14$. Le curve sono comunque solamente una guida visiva. Le linee orizzontali e verticali corrispondono alle rette d'equazione $x = 0$ e $y = 1$.

In base a quanto ottenuto con le conferme sperimentali, riteniamo di aver dimostrato la validità della nostra assunzione che la temperatura corrispondente al minimo del tempo di coalescenza coincide con la temperatura critica per la transizione di fase.

4.2 Conclusioni e Prospettive

Il nostro lavoro di tesi ha inizialmente riguardato lo studio del gas forzato su reticolo quasi-unidimensionale attraverso simulazioni Monte Carlo, per ottenere stime numeriche degli esponenti critici e di altri parametri, in modo da poter evidenziare eventuali differenze o similarità con l'analogo bidimensionale. Durante la fase di analisi dei dati raccolti ci siamo accorti che l'algoritmo multispin coding da noi utilizzato, contrariamente a quanto atteso, portava ad una coalescenza dei reticoli simulati in tempi molto brevi rispetto a quanto lecito aspettarsi.

Abbiamo dunque cercato di dare una spiegazione a tale fatto, basandoci soprattutto sulle strette connessioni che il tempo di coalescenza ha con la dinamica dei sistemi simulati. In base a tali considerazioni siamo giunti alla conclusione che la presenza del campo forzante induce un accoppiamento effettivo tra le particelle che tentano d'effettuare balzi nella direzione del campo.

Proseguendo l'analisi in dettaglio, grazie a simulazioni numeriche volte

a verificare alcuni aspetti dell'algoritmo, ci siamo resi conto che la distribuzione dei tempi di coalescenza presentava differenti minimi, per differenti valori di $L_{||}$, distribuiti in maniera regolare. In base a considerazioni sull'influenza della dinamica e della transizione di fase in relazione al processo di coalescenza dei reticoli, siamo giunti all'ipotesi che il valore di beta in cui tali minimi avvenivano corrispondesse al valore critico per il modello.

Abbiamo a questo punto svolto delle verifiche sperimentali per confermare tale ipotesi. L'analisi di taglia finita sui dati in nostro possesso ha restituito valori in accordo con la letteratura presente. Ancora più confortante è stata l'applicazione dell'algoritmo, opportunamente modificato, al gas su reticolo. In questo caso abbiamo ottenuto un ottimo accordo tra i dati ottenuti grazie al multispin coding non solo con la temperatura critica del modello (nota esattamente) ma anche con l'esponente critico relativo allo scaling della lunghezza di correlazione.

Riteniamo in base a ciò di aver dimostrato come un tale uso dell'algoritmo, non presente in letteratura, possa portare grazie alla stretta connessione con la dinamica all'identificazione della temperatura critica e degli esponenti critici ad essa connessi.

Riteniamo comunque che tale metodo d'indagine possa essere ulteriormente sviluppato ed affinato, al fine sia di aumentare la comprensione degli effetti dinamici che ne sono alla base, sia perfezionarne l'uso per ottenere risultati numerici con maggiore precisione. In particolare esso potrebbe trovare un uso specifico in quei casi in cui la dinamica del modello presenti delle variazioni molto nette, o anisotropie rispetto a determinate direzioni.

4.A Breve discussione sul metodo di analisi impiegato

In relazione ai dati raccolti durante l'analisi del tempo di coalescenza del modello, alcune tra le domande che ci si potrebbe (e dovrebbe) porre, ci sono: "Che valore possiedono questi dati? Quanto sono affidabili e di cosa sono rappresentativi?"

Gli errori Come già dichiarato nel testo, gli errori rappresentati in figura sono calcolati attraverso la formula

$$\delta t = \frac{1}{\sqrt{N}} \sqrt{\frac{\sum_{i=1}^N (t_i - \bar{t})^2}{N-1}}, \quad (4.7)$$

che rappresenta la deviazione media della media dal suo valore vero, se i dati sono distribuiti in modo gaussiano. Nel nostro caso tale assunzione non corrisponde completamente a verità. La distribuzione dei dati se ne discosta, ma non giudichiamo corretto ritenerli completamente indipendenti. Prima che di essi, dobbiamo però occuparci di *cosa* rappresentano effettivamente quei dati. Dato che noi stiamo simulando 32 reticoli, il valore da noi misurato è il tempo in cui avviene la coalescenza di tutti e 32 i reticoli. Possiamo rapportarlo a qualcosa di più generale?

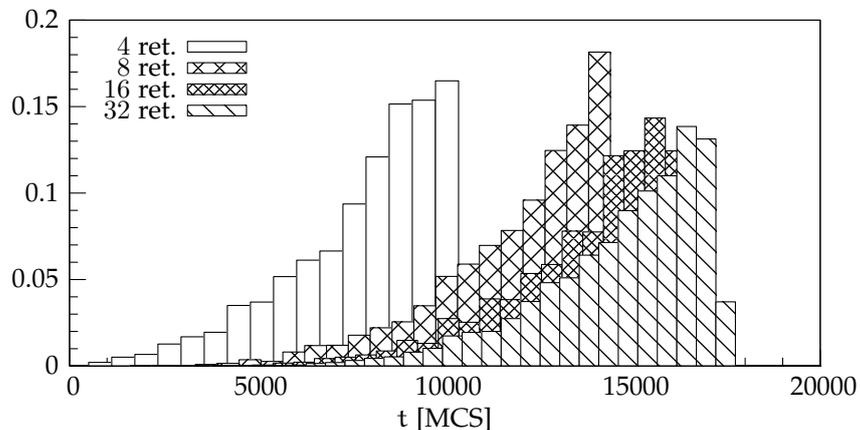


Figura 4.9: Distribuzione del tempo di coalescenza finale per msc con 4, 8, 16, 32 reticoli fatti evolvere contemporaneamente. Le simulazioni reticoli hanno interessato reticoli 8×12 con $\beta = 0.2725$. Per ogni reticolo, la distribuzione è basata su 10000 run effettuate con differenti sequenze di numeri casuali. Il range di valori di ogni singola distribuzione è stato diviso in 50 intervalli.

In fig. 4.A sono riportate le distribuzioni dei dati dei tempi di coalescenza, a $\beta = 0.2725$, per simulazioni multispin coding dipendenti effettuate con $r = 4, 8, 16, 32$ reticoli contenuti in una parola¹ dell'elaboratore. Possiamo al riguardo fare 2 osservazioni:

- la distribuzione dei tempi di coalescenza, per r fissato, pur non essendo gaussiana, non risulta nemmeno uniforme;
- la distribuzione dei tempi presenta una notevole asimmetria nel massimo che si sposta, al crescere di r , verso tempi maggiori.

In realtà non si sposta indefinitamente, ma tende ad un valore finito che corrisponde al tempo di coalescenza per tutti gli $|\Lambda|^2$ reticoli dello spazio delle configurazioni. Per il teorema 1.20 tale tempo è distribuito come il tempo di Propp-Wilson. Tale fatto si nota particolarmente bene effettuando il grafico della distribuzione dei tempi di coalescenza per r fissati, in funzione di β , come in fig. 4.A

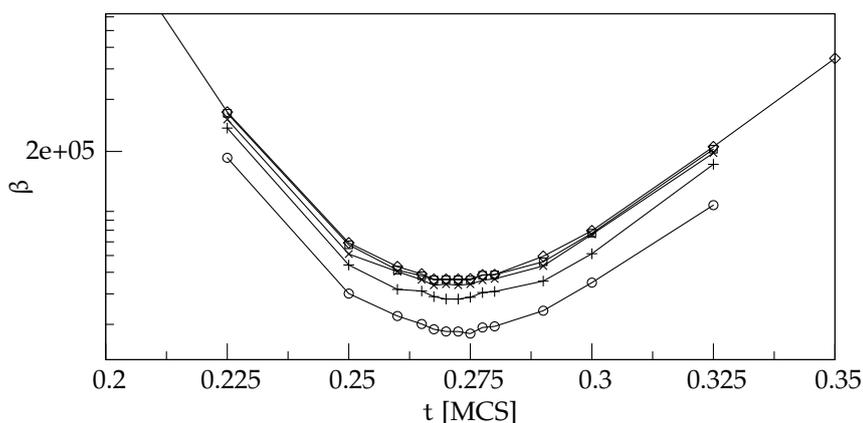


Figura 4.10: Distribuzione dei tempi di coalescenza totale per reticoli 8×12 simulati con MSC con 2, 4, 8, 16, 32 reticoli contemporaneamente. Rispettivamente dal basso verso l'alto. Nella zona centrale sono state effettuate 1000 run, di lato 100.

In realtà la cosa migliore[58, 6, 51] da fare sarebbe una stima degli errori tramite metodi Monte Carlo per il resampling, come il bootstrap o il rebinning. Ci riserviamo tale compito per il futuro.

¹In realtà i reticoli per parola rimangono 32. Per simulare n reticoli, abbiamo posto i $32 - n$ rimanenti uguali tra loro ed uguali al primo degli n . È stato il metodo più semplice che non richiedesse la riscrittura del programma.

4. ANALISI DEL MODELLO

Concludiamo con qualche osservazione sui tempi di coalescenza dei reticoli. Se noi simuliamo l'evoluzione di r reticoli contemporaneamente, in realtà in tutto avremo al più¹ $r - 1$ tempi di coalescenza, compreso quello finale che è quello che abbiamo studiato maggiormente. All'interno di un singolo insieme di simulazioni per fissati β ed r , si nota come le distribuzioni di maggior rilievo — sono più concentrate! — risultano essere le prime e le ultime. I tempi di coalescenza centrali sono “sparsi” in maniera considerevole.

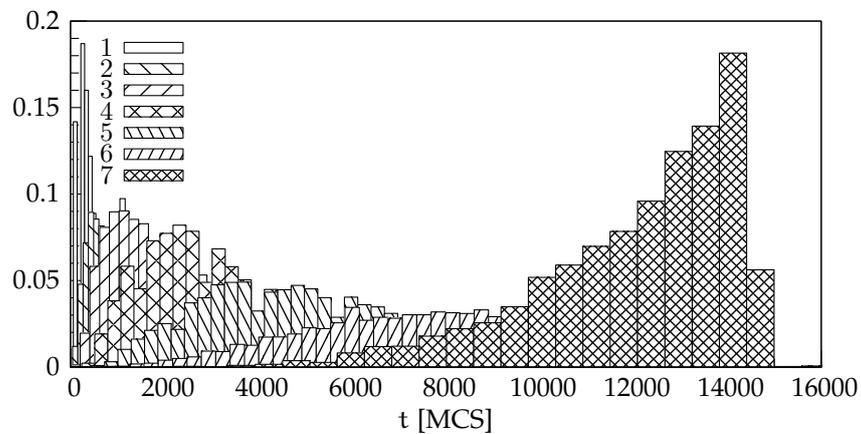


Figura 4.11: Distribuzione dei 7 tempi di coalescenza di un reticolo 8×12 , ad una temperatura inversa $\beta = 0.2725$. Le distribuzioni è basata su 10000 run effettuate con differenti sequenze di numeri casuali. Il range di valori di ogni singola distribuzione è stato diviso in 50 intervalli.

¹al più poiché potrebbe avvenire la coalescenza contemporanea di più reticoli

4.B Dati di coalescenza

Riportiamo nelle pagine seguenti i dati raccolti durante l'analisi dei tempi di coalescenza per reticoli $8 \times L_{\parallel}$, con L_{\parallel} variabile da 8 a 80 ad intervalli di 4. I dati sono riuniti per valori costanti di L_{\parallel} . Per ogni insieme di valori abbiamo concentrato il nostro studio nella zona in cui si presentava un minimo per il tempo di coalescenza t ; in questa zona N , che è il numero di simulazioni differenti effettuate è solitamente maggiore di un ordine di grandezza rispetto alle zone esterne. Δt è l'intervallo temporale (in Monte Carlo Step, MCS) con cui si è controllato se la coalescenza è avvenuta.

4. ANALISI DEL MODELLO

Tabella 4.4: dati di coalescenza

L_{\parallel}	β	t	δt	N	Δt
8	0.25	6.168e+06	2.546e+05	100	1e+03
8	0.30	1.978e+06	1.043e+05	100	1e+03
8	0.35	3.187e+05	1.533e+04	100	1e+03
8	0.40	6.443e+04	3.900e+03	100	1e+03
8	0.45	1.687e+04	9.396e+02	100	1e+03
8	0.48	1.376e+04	2.372e+02	1100	1e+03
8	0.49	1.296e+04	2.380e+02	1000	1e+03
8	0.495	1.242e+04	2.337e+02	1000	1e+03
8	0.50	1.201e+04	2.221e+02	1000	1e+03
8	0.505	1.245e+04	2.291e+02	1000	1e+03
8	0.51	1.277e+04	2.383e+02	1000	1e+03
8	0.52	1.313e+04	2.503e+02	1000	1e+03
8	0.55	1.761e+04	1.014e+03	100	1e+03
8	0.60	4.120e+04	2.619e+03	100	1e+03
8	0.65	1.213e+05	7.318e+03	100	1e+03
8	0.70	3.146e+05	2.025e+04	100	1e+03
8	0.75	1.118e+06	7.848e+04	100	1e+03
8	0.80	2.723e+06	1.649e+05	100	1e+03
8	0.85	6.375e+06	2.732e+05	100	1e+03
12	0.35	6.699e+07	2.476e+06	100	1e+03
12	0.38	1.040e+07	5.892e+05	100	1e+03
12	0.40	3.077e+06	1.653e+05	100	1e+03
12	0.45	2.703e+05	1.458e+04	100	1e+03
12	0.50	4.737e+04	2.667e+03	100	1e+03
12	0.52	3.451e+04	1.910e+03	100	1e+03
12	0.53	3.128e+04	5.697e+02	1000	1e+03
12	0.535	2.916e+04	5.150e+02	1000	1e+03
12	0.54	2.919e+04	5.443e+02	1000	1e+03
12	0.545	2.918e+04	5.214e+02	1000	1e+03
12	0.55	2.931e+04	5.238e+02	1000	1e+03
12	0.555	3.091e+04	5.631e+02	1000	1e+03
12	0.56	3.105e+04	5.831e+02	1000	1e+03
12	0.58	3.960e+04	2.220e+03	100	1e+03
12	0.60	5.565e+04	3.615e+03	100	1e+03
12	0.65	1.704e+05	1.021e+04	100	1e+03
12	0.70	5.524e+05	3.254e+04	100	1e+03
12	0.75	2.260e+06	1.501e+05	100	1e+03
12	0.80	6.977e+06	4.176e+05	100	1e+03

Tabella 4.4: dati di coalescenza, continua

$L_{ }$	β	t	δt	N	Δt
16	0.40	2.029e+08	1.243e+07	100	1e+03
16	0.43	1.879e+07	1.222e+06	100	1e+03
16	0.45	4.902e+06	2.684e+05	100	1e+03
16	0.50	3.225e+05	1.882e+04	100	1e+03
16	0.52	1.359e+05	8.202e+03	100	1e+03
16	0.54	7.504e+04	4.046e+03	100	1e+03
16	0.55	6.247e+04	1.138e+03	1000	1e+03
16	0.555	6.208e+04	1.182e+03	1000	1e+03
16	0.56	5.581e+04	1.006e+03	1000	1e+03
16	0.565	5.596e+04	1.012e+03	1000	1e+03
16	0.57	5.395e+04	1.014e+03	1000	1e+03
16	0.575	5.657e+04	1.061e+03	1000	1e+03
16	0.58	5.815e+04	1.057e+03	1100	1e+03
16	0.59	6.735e+04	2.645e+03	200	1e+03
16	0.60	7.489e+04	3.297e+03	200	1e+03
16	0.65	2.784e+05	1.886e+04	100	1e+03
16	0.70	1.020e+06	7.196e+04	100	1e+03
16	0.75	5.030e+06	3.261e+05	100	1e+03
16	0.80	2.002e+07	1.272e+06	100	1e+03
20	0.45	8.119e+07	2.279e+06	100	1e+03
20	0.48	8.788e+06	4.638e+05	100	1e+03
20	0.50	2.317e+06	1.213e+05	100	1e+03
20	0.53	4.030e+05	1.984e+04	100	1e+03
20	0.55	1.639e+05	7.640e+03	100	1e+03
20	0.57	1.168e+05	6.804e+03	100	1e+03
20	0.575	1.029e+05	5.885e+03	100	1e+03
20	0.58	9.975e+04	1.940e+03	1000	1e+03
20	0.585	9.559e+04	1.722e+03	1000	1e+03
20	0.59	1.005e+05	1.825e+03	1000	1e+03
20	0.595	1.011e+05	5.113e+03	100	1e+03
20	0.60	1.077e+05	6.775e+03	100	1e+03
20	0.61	1.256e+05	6.767e+03	100	1e+03
20	0.62	1.615e+05	8.847e+03	100	1e+03
20	0.65	3.943e+05	2.574e+04	100	1e+03
20	0.70	1.725e+06	1.067e+05	100	1e+03
20	0.75	1.010e+07	6.063e+05	100	1e+03

4. ANALISI DEL MODELLO

Tabella 4.4: dati di coalescenza, continua

$L_{ }$	β	t	δt	N	Δt
24	0.50	2.101e+07	1.434e+06	100	1e+04
24	0.55	6.069e+05	3.997e+04	100	1e+04
24	0.56	3.892e+05	2.238e+04	100	1e+04
24	0.57	2.534e+05	1.426e+04	100	1e+04
24	0.58	2.019e+05	1.029e+04	100	1e+04
24	0.585	1.749e+05	3.209e+03	1002	1e+04
24	0.59	1.5889e+05	2.906e+03	1005	1e+04
24	0.595	1.5893e+05	2.994e+03	1000	1e+04
24	0.60	1.620e+05	3.088e+03	1000	1e+04
24	0.605	1.618e+05	3.015e+03	1000	1e+04
24	0.61	1.844e+05	3.555e+03	1000	1e+04
24	0.62	2.275e+05	1.570e+04	100	1e+04
24	0.65	5.185e+05	3.264e+04	100	1e+04
24	0.70	3.093e+06	2.326e+05	200	1e+04
24	0.75	1.836e+07	1.147e+06	100	1e+04
28	0.50	1.856e+08	9.726e+06	100	1e+04
28	0.52	2.682e+07	1.303e+06	100	1e+04
28	0.55	2.514e+06	1.445e+05	100	1e+04
28	0.58	4.106e+05	2.210e+04	100	1e+04
28	0.59	3.305e+05	1.726e+04	100	1e+04
28	0.595	2.934e+05	1.698e+04	100	1e+04
28	0.60	2.809e+05	5.208e+03	1000	1e+04
28	0.605	2.636e+05	4.835e+03	1000	1e+04
28	0.61	2.718e+05	5.201e+03	1000	1e+04
28	0.615	2.926e+05	5.664e+03	1000	1e+04
28	0.62	3.044e+05	1.689e+04	100	1e+04
28	0.65	6.546e+05	3.769e+04	100	1e+04
28	0.70	4.795e+06	3.084e+05	100	1e+04
28	0.75	4.045e+07	2.679e+06	100	1e+04

Tabella 4.4: dati di coalescenza, continua

$L_{ }$	β	t	δt	N	Δt
32	0.55	9.367e+06	5.023e+05	100	1e+03
32	0.58	1.360e+06	7.243e+04	100	1e+03
32	0.60	5.187e+05	9.583e+03	1000	1e+03
32	0.605	4.757e+05	8.921e+03	1000	1e+03
32	0.61	4.266e+05	7.745e+03	1000	1e+03
32	0.615	4.274e+05	8.116e+03	1000	1e+03
32	0.62	4.312e+05	8.378e+03	1000	1e+03
32	0.625	4.812e+05	9.820e+03	1000	1e+03
32	0.63	5.312e+05	3.033e+04	100	1e+03
32	0.65	1.027e+06	7.336e+04	100	1e+03
32	0.70	8.407e+06	5.185e+05	100	1e+03
32	0.75	7.092e+07	4.303e+06	100	1e+05
36	0.55	4.867e+07	2.659e+06	100	1e+05
36	0.58	3.634e+06	1.834e+05	100	1e+05
36	0.60	1.032e+06	5.971e+04	100	1e+05
36	0.61	6.960e+05	3.535e+04	100	1e+05
36	0.615	6.764e+05	1.317e+04	1000	1e+05
36	0.62	6.120e+05	1.163e+04	1000	1e+05
36	0.625	6.611e+05	1.353e+04	1000	1e+05
36	0.63	6.670e+05	4.686e+04	100	1e+05
36	0.65	1.370e+06	9.484e+04	100	1e+05
36	0.70	1.208e+07	8.884e+05	100	1e+05
40	0.55	2.381e+08	1.268e+07	101	1e+05
40	0.58	1.101e+07	6.051e+05	104	1e+05
40	0.60	2.610e+06	1.274e+05	100	1e+05
40	0.61	1.430e+06	1.129e+05	100	1e+05
40	0.615	1.280e+06	2.433e+04	1000	1e+05
40	0.62	1.110e+06	2.074e+04	1000	1e+05
40	0.625	1.027e+06	1.995e+04	1000	1e+05
40	0.63	1.008e+06	1.897e+04	1000	1e+05
40	0.635	1.124e+06	2.230e+04	1000	1e+05
40	0.64	1.287e+06	8.200e+04	100	1e+05
40	0.65	1.793e+06	1.193e+05	100	1e+05
40	0.70	2.078e+07	1.413e+06	104	1e+05

Tabella 4.4: dati di coalescenza, continua

L_{\parallel}	β	t	δt	N	Δt
44	0.58	3.262e+07	1.501e+06	100	1e+05
44	0.60	6.427e+06	3.345e+05	100	1e+05
44	0.62	1.944e+06	1.043e+04	100	1e+05
44	0.625	1.826e+06	4.269e+04	582	1e+05
44	0.63	1.676e+06	4.370e+04	503	1e+05
44	0.635	1.705e+06	4.529e+04	512	1e+05
44	0.64	1.872e+06	4.652e+04	569	1e+05
44	0.645	2.097e+06	5.776e+04	508	1e+05
44	0.65	2.217e+06	1.177e+05	100	1e+05
44	0.70	3.051e+07	1.984e+06	100	1e+05
48	0.58	1.128e+08	6.051e+06	100	1e+05
48	0.60	1.600e+07	8.126e+05	100	1e+05
48	0.62	3.972e+06	1.003e+05	500	1e+05
48	0.625	3.234e+06	8.244e+04	500	1e+05
48	0.63	2.759e+06	4.903e+04	1000	1e+05
48	0.635	2.636e+06	4.763e+04	1000	1e+05
48	0.64	2.611e+06	5.099e+04	1000	1e+05
48	0.645	2.933e+06	1.258e+05	200	1e+05
48	0.65	3.586e+06	2.098e+05	100	1e+05
48	0.66	5.238e+06	3.612e+05	100	1e+05
48	0.70	5.116e+07	3.635e+06	100	1e+05
52	0.60	4.942e+07	2.741e+06	103	1e+05
52	0.62	9.132e+06	5.491e+05	100	1e+05
52	0.63	5.016e+06	1.262e+05	500	1e+05
52	0.635	4.240e+06	9.922e+04	500	1e+05
52	0.64	3.958e+06	1.121e+05	500	1e+05
52	0.645	4.320e+06	1.129e+05	500	1e+05
52	0.65	5.241e+06	1.559e+05	500	1e+05
52	0.66	7.048e+06	4.835e+05	100	1e+05
52	0.70	7.021e+07	4.280e+06	99	1e+05

Tabella 4.4: dati di coalescenza, continua

$L_{ }$	β	t	δt	N	Δt
56	0.61	4.523e+07	2.519e+06	100	1e+05
56	0.62	2.006e+07	1.143e+06	100	1e+05
56	0.63	7.728e+06	3.945e+05	100	1e+05
56	0.635	7.524e+06	1.382e+05	1000	1e+05
56	0.64	6.623e+06	4.301e+05	100	1e+05
56	0.645	6.350e+06	1.192e+05	1005	1e+05
56	0.65	6.762e+06	1.186e+05	1000	1e+05
56	0.655	8.052e+06	1.519e+05	1003	1e+05
56	0.66	1.101e+07	6.310e+05	107	1e+05
56	0.68	2.934e+07	1.959e+06	100	1e+05
60	0.64	1.178e+07	6.275e+05	100	1e+05
60	0.645	9.350e+06	4.687e+05	100	1e+05
60	0.65	1.026e+07	6.180e+05	100	1e+05
60	0.655	1.099e+07	6.561e+05	100	1e+05
60	0.66	1.352e+07	8.973e+05	100	1e+05
64	0.62	1.076e+08	6.540e+06	100	1e+05
64	0.63	4.029e+07	2.185e+06	100	1e+05
64	0.64	1.957e+07	9.623e+05	100	1e+05
64	0.645	1.646e+07	4.258e+05	500	1e+05
64	0.65	1.480e+07	3.558e+05	504	1e+05
64	0.655	1.568e+07	4.066e+05	480	1e+05
64	0.66	1.709e+07	1.034e+06	100	1e+05
64	0.67	3.511e+07	2.223e+06	100	1e+05
64	0.68	6.015e+07	3.911e+06	100	1e+05

4. ANALISI DEL MODELLO

Tabella 4.4: dati di coalescenza, finisce.

L_{\parallel}	β	t	δt	N	Δt
68	0.64	3.599e+07	2.031e+06	100	1e+05
68	0.65	2.517e+07	1.406e+06	100	1e+05
68	0.6525	2.325e+07	1.263e+06	100	1e*05
68	0.655	2.093e+07	1.075e+06	100	1e+05
68	0.6575	2.440e+07	1.337e+06	100	1e+05
68	0.66	2.455e+07	1.395e+06	100	1e+05
68	0.7	.294e+07	4.607e+06	100	1e+05
72	0.64	8.010e+07	4.067e+06	100	1e+05
72	0.65	3.796e+07	1.876e+06	100	1e+05
72	0.655	3.507e+07	1.095e+06	327	1e+05
72	0.66	3.507e+07	1.166e+06	311	1e+05
72	0.67	6.875e+07	4.175e+06	100	1e+05
76	0.65	6.632e+07	3.128e+06	100	1e+05
76	0.655	5.200e+07	2.884e+06	100	1e+05
76	0.66	5.369e+07	2.813e+06	102	1e+05
76	0.665	5.884e+07	3.391e+06	100	1e+05
80	0.65	1.273e+08	6.845e+06	100	1e+06
80	0.655	8.909e+07	5.217e+06	100	1e+06
80	0.6575	7.749e+07	4.326e+06	100	1e+06
80	0.66	7.859e+07	4.232e+06	100	1e+06
80	0.6625	7.250e+07	3.415e+06	123	1e+06
80	0.665	8.253e+07	4.885e+06	102	1e+06
80	0.67	1.127e+08	6.588e+06	100	1e+06
80	0.68	2.099e+08	1.712e+07	100	1e+06

IL MULTISPIN CODING, NOTA STORICA

Dove si parla di come l'introduzione dei supercomputer abbia rivoluzionato lo studio numerico di modelli stocastici

QUESTA APPENDICE intende fornire una breve storia del multispin coding, attraverso l'analisi degli articoli che lo hanno introdotto e sviluppato. In realtà tale ricerca è stata effettuata per scoprire se quanto abbiamo trovato nel corso dei nostri studi fosse stato già rilevato ed analizzato in precedenza. Da questo punto di vista la nostra ricerca ha avuto esito negativo, ma non ce la siamo sentiti di non raccontare questa breve storia che si è sviluppata in parallelo nell'informatica e nella fisica¹.

1970: Friedberg e Cameron

In un articolo del 1970, che purtroppo venne per molto tempo sottovalutato, Friedberg e Cameron [22] condussero un'analisi lucida e rigorosa dello stato dell'arte delle simulazioni Monte Carlo, introducendo ed applicando inoltre alcune intuizioni tecniche che dimostrarono la loro importanza solamente molti anni più tardi.

"In the present study the primary object of examination is the Monte Carlo method itself, rather than the systems to be simulated."

L'articolo, "Test of the Monte Carlo Method: Fast Simulation of a Small Ising Lattice", si proponeva di introdurre e/o analizzare alcune tecniche per rendere più veloci (secondo gli standard dell'epoca) le simulazioni MC; tale test era effettuato, per ottenere risultati quantitativi, su reticoli bidimensionali di dimensioni davvero ridotte: 4×4 . Diamo qui un veloce riassunto delle idee fondamentali.

¹A scanso di equivoci, ammettiamo che tale ricerca non è stata condotta seguendo un metodo storiografico, non in nostro possesso. Abbiamo piuttosto cercato di reperire tutte le informazioni possibili sullo sviluppo e sugli studi del multispin coding.

In una simulazione MCMC le possibili fonti di errore sono essenzialmente quattro:

1. La dimensione finita del campione. Tale caratteristica in realtà genera a sua volta 2 tipi di errori:
 - a) il “finite sampling error”, che è proporzionale alle fluttuazioni termodinamiche delle quantità di interesse, presente anche se i campioni fossero campionati perfettamente dalla distribuzione voluta;
 - b) una fluttuazione, imprevedibile, dovuta alle correlazioni tra stati generati successivi.
2. la dimensione finita del sistema, confrontata con sistemi fisici reali che contengono $O(10^{23})$ elementi;
3. la natura semplificata del sistema ideale — tale discrepanza appare però solo se il sistema simulato è un’idealizzazione di uno specifico sistema reale — ;
4. errori sistematici nell’implementazione. Tali errori appaiono se il metodo MCMC non è applicato correttamente, se ci sono sottigliezze che lo invalidano in tutto o in parte, se il generatore di numeri casuali ha qualche bias, se il programma contiene errori, se il computer è bacato, etc. . . .

Vogliamo concentrarci sull’errore 1b, eliminando gli altri 3: il tipo 1a è eliminato dalla teoria stessa; il tipo 3 viene eliminato confrontando i risultati con quelli di calcoli analitici sullo stesso sistema, e non col mondo reale; il secondo tipo viene eliminato usando soluzioni analitiche di sistemi finiti. Per separare 1 da 4 e studiarlo in dettaglio, è utile avere procedure MCMC molto rapide per poter produrre campioni molto grandi. In tale caso gli errori di tipo 4 persistono, mentre quelli di tipo 1a e 1b dovrebbero diminuire con il quadrato della dimensione dei campioni.

La velocità di una procedura MCMC è affetta sia dalla complessità del sistema studiato, dalla velocità del computer e dall’efficienza del programma.

la procedura

Dato che un singolo sito di un modello di Ising è descritto da un singolo bit, mentre un computer 7094 (utilizzato nelle simulazioni) usa parole di 36

bit, è possibile guadagnare velocità trattando contemporaneamente più siti come attivi. Il reticolo viene diviso in 2 sottoreticoli secondo una geometria a scacchiera, cosicchè nessuna coppia di spin primi-vicini appartenga allo stesso sottoreticolo. Allora tutti i siti di un sottoreticolo possono divenire attivi allo stesso tempo, ed i due sottoreticoli essere trattati alternatamente. Le probabilità di transizione utilizzate per lo spin flip sono date dall'algoritmo Metropolis, con una leggera variazione nel caso $\Delta E = 0$. Dato che il computer possiede parole di 36 bit e nella simulazione si utilizzano reticoli 4×4 , è possibile far evolvere 9 di tali reticoli indipendentemente allo stesso tempo. I 144 siti sono immagazzinati in 4 parole, cosicchè ogni parola contiene un sottoreticolo di 4 siti. La procedura stocastica viene realizzata con operazioni logiche su queste 4 parole; dunque ogni operazione-macchina si occupa di 36 distinti siti.

una piacevole sorpresa

In accordo con le idee di Fosdick [21] sono state fatte evolvere sequenze con 2 differenti configurazioni iniziali, così da studiare empiricamente il rate di scomparsa dell'influenza dell'inizializzazione (effettuando un confronto tra i valori medi di energia e magnetizzazione). Tale studio giunge alla conclusione che ad alte temperature l'influenza dello stato iniziale scompare rapidamente, mentre per temperature minori l'influenza decade più lentamente per quanto riguarda l'energia e molto più lentamente per quanto concerne la magnetizzazione.

1979-1982: Dieci Anni Dopo...

The Multi(ple) Spin Coding (Technique)

Analizziamo qui 5 articoli dovuti a Creutz *et al.* (1979) [14, 15], Rebbi e Swendsen (1980) [53], Zorn *et al.* (1981) [69], Kalle e Winkelmann (1982) [34].

Sfortunatamente dovettero passare quasi dieci anni prima che alcune delle idee di Friedberg e Cameron venissero riprese ed applicate. Ma da questo momento l'efficacia della multiple spin coding technique — questo uno dei suoi primi nomi, MSCT — diviene nota al grande pubblico (di specialisti) e non smetterà più di essere sviluppata ed applicata.

“The multi-spin-coding technique (MSCT) developed by Rebbi *et al.* [14, 15, 53] was first mentioned by Friedberg and Came-

ron [22] (§4.2). It was used to reduce memory space and execution time in Monte Carlo (MC) simulations of discrete variables and was successfully applied to gauge model [14] and the Potts model [15].

[...] We demonstrate that the static values obtained by the two methods¹ are identical and that the MSCT needs only about one third of the execution time required by the conventional method. Of course, it also requires less memory space.”

Zorn et al., 1981 [69]

Nella sua prima applicazione il MSC venne utilizzato per velocizzare e soprattutto per ridurre l’uso di memoria nella simulazione di modelli di Ising quadrati. Il computer in uso (CDC CYBER 76) possedeva parole di 60 bit che vennero usate per immagazzinare 20 spin. I 2 bit liberi tra uno spin e l’altro vennero usati per inserirvi i valori dell’energia (4 possibili valori). Un array SP0 conteneva gli spin di colonna dispari (odd) $SP0(i, 1) = s_{i,1}, s_{i,3}, \dots, s_{i,39}$, $SP0(i, 2) = s_{i,41}, s_{i,43}, \dots, s_{i,79}$ etc... Un array SPE conteneva gli spin di colonna pari (even) $SPE(i, 1) = s_{i,2}, s_{i,4}, \dots, s_{i,40}$ etc...

Le varie parole venivano poi “XORate” per calcolare l’energia d’interazione; quindi veniva estratto un numero casuale che veniva confrontato con la corretta probabilità contenuta in una tabella, e così via seguendo l’onnipresente algoritmo Metropolis.

In pratica l’uso principale è la riduzione di memoria — 20 spins-per-word invece di 1 spins-per-word — ma si ha anche una netta riduzione del tempo di esecuzione in quanto quasi tutte le istruzioni sono composte da operazioni logiche sui bit e da bit-shift.

“In CDC 6000 and 7000 series computer each memory unit (“word”) has 60 bits; in a (spin-1/2) Ising model, one bit is sufficient to represent a single spin. Multispin coding allows three bits for every spin and thus make it possible to count the number of antiparallel neighbor spin, from zero to six in the simple cubic lattice, with these three bits. Thus 20 spins (or 20 interaction energies) can be stored in one word and calculations for these can be made with one word-oriented instruction. Therefore, compared with conventional one-word-per-spin methods, the memory required is reduced by a factor of 20, the computer time

¹multispin coding technique e Monte Carlo, n.d.A.

by a factor of 2. Basically the program performs an exclusive-or operation (XOR) between a word representing 20 spins and their neighbors; the result gives in each of the 20 three-bits parcels the number of antiparallel neighbors, and thus the interaction energy. Thereafter each spin has to be flipped separately according to the probability with a random number distributed homogeneously between zero and one.”

Kalle e Winkelmann, 1982 [34]

1984: Il super-spin coding

Williams e Kalos [65]

È noto che, vicino al punto critico, il tempo di autocorrelazione (relaxation time) dei processi dinamici aumenta vertiginosamente (critical slowing down). In pratica il tempo di autocorrelazione τ è perlopiù proporzionale a $\min\{L, \xi^z\}$, dove L è la dimensione lineare del sistema, ξ la lunghezza di correlazione di un sistema infinito alla stessa temperatura e z l'esponente critico dinamico (per Ising $z \approx 2$ in ogni dimensione).

Ci sono 2 metodi differenti per tentare di risolvere tale problema: attaccarlo dal punto di vista computazionale o dal punto di vista fisico.

1. restando all'interno dell'algoritmo Metropolis a singolo spin-flip, ho tre metodi per tentare di velocizzare la simulazione:
 - a) posso usare macchine più veloci;
 - b) posso usare macchine costruite appositamente per tale compito (special-purpose machine);
 - c) posso riorganizzare la struttura dei dati e l'algoritmo per aumentare l'efficienza.

I primi due sono legati all'hardware, il terzo al software.

2. Il secondo metodo è lavorare su sistemi fisici differenti che hanno un rallentamento minore. In tale categoria rientrano i sistemi ad update collettivo.

Qui si provvede ad implementare un algoritmo msc di maggior efficacia. I msc presenti all'epoca sul mercato, sebbene trattassero più spin in parallelo, erano comunque rallentati dal fatto di dover effettuare il tentativo di flip

serialmente su di ogni singolo spin. Si effettua l'intera procedura msc in parallelo, usando solo operazioni booleane.

All'incirca il funzionamento è il seguente:

si prende il reticolo e si immagazzinano i siti lungo una direzione (e.g. le colonne) nello stesso posto in parole di N bit. Come al solito non vengono immagazzinati bit consecutivi sulla stessa parola dato che altrimenti non si potrebbe effettuare lo XOR per calcolare la variazione di energia. Dopodiché si prepara una striscia di bit in proporzione alle possibili frequenze di accettazione e le si usano, sempre bit-a-bit, per fare l'update simultaneo di tutti gli spin in una parola.

Qualora si utilizzassero macchine vettoriali, è possibile anche simulare più reticoli in parallelo.

Va a favore degli autori il rilevare la sottostima che ha avuto l'articolo di Friedberg e Cameron [22] del 1970.

Wansleben, Zabolitzky e Kalle [64]

In questo articolo viene utilizzato il multispin coding per svolgere un'indagine sul modello di Ising tridimensionale. Pur se di interesse fisico, l'articolo non presenta novità dal punto di vista dell'implementazione algoritmica del msc.

Diamo solo una rapida descrizione del loro utilizzo.

Una configurazione (su-giù, codificata da 0-1) di 20 spin è immagazzinata in una parola, usando 3 bit per spin. Questo permette l'aggiunta del valore dello XOR a tutti e sei i primi vicini per ognuno di questi spin simultaneamente durante il calcolo dell'energia d'interazione. Ciò obbliga a mettere spin p.v. in parole differenti. Allora la dimensione minima del sistema è 40^3 , dove ogni riga (diciamo nella direzione 1) è rappresentata da 2 parole. In questo modo il reticolo è diviso in 2 sottoreticoli ognuno contenente i siti pari/dispari

1985

Gawlinski, Grant, Gunton e Kaski [25]

In questo articolo gli autori cercano di dare una giustificazione teorica all'utilizzo del msc nella simulazione di modelli stocastici. Presumibilmente è il primo articolo che affronta una tale tematica. È un segnale che il multispin coding sta diventando maturo. . .

L'applicazione del metodo MCMC alla dinamica è un po' controversa. Il principale problema viene dal fatto che l'algoritmo che uno sceglie per l'update diviene una parte integrante del modello. Quindi, per differenti procedure impiegate, una nuova equazione dinamica è in effetti simulata.

Esaminiamo le relazioni tra la Master Equation standard al tempo discreto

$$\rho(\sigma, t+1) - \rho(\sigma, t) = - \sum_{i=1}^N [W_i(\sigma_i)\rho(\sigma_i, t) - W_i(-\sigma_i)\rho(-\sigma_i, t)] \quad (8)$$

e l'algoritmo MC

$$\rho(\sigma, t+1) - \rho(\sigma, t) = - \sum_{i=1}^N \left[W_i(\sigma_i)\rho\left(\sigma_i, t + \frac{i-1}{N}\right) - W_i(-\sigma_i)\rho\left(-\sigma_i, t + \frac{i-1}{N}\right) \right] \quad (9)$$

dove l'apice sulla sommatoria indica che gli spin σ_i sono valutati in un ordine particolare.

Le due equazioni sono diverse a causa della differente dipendenza temporale nel lato destro di (9). In (8) tutti gli spin flippano nello stesso ambiente senza un ordine particolare, mentre in (9) flippano in un ordine ben preciso: mentre lo spin σ_1 flippa nell'ambiente del t-esimo MCStep, lo spin σ_N flippa nell'ambiente del (t+1)-esimo MCS. Dunque, se l'algoritmo deve simulare la Master Equation standard, non deve esserci variazione significativa durante un MCS.

Meglio espresso: La quantità fisica $R(t)$ deve soddisfare

$$\left| \frac{R(t+1) - R(t)}{R(t)} \right| \ll 1 \quad (10)$$

per l'eq. (9), dove t è espresso in termini di MCS. Tale condizione è certamente necessaria, ma non sufficiente. Non esistono comunque, a conoscenza, dimostrazioni analitiche di limiti sull'uguaglianza delle due Master Equation.

La differenza è data dal modo in cui update gli spin nel reticolo. Si potrebbe fare in sequenza, ma gli esponenziali indicano che ciò viola (10). Un modo che invece si rivela rispettoso della (10) è update gli spin a caso; in questo caso tale algoritmo simula fedelmente la standard ME.

Notiamo che, effettuando gli update in maniera casuale, saranno girati spin largamente spazati. Per tale motivo uno può scegliere contemporaneamente più spin distanti tra loro ed effettuare l'update contemporaneamente. È ciò che fa il multispin coding. Dai test si trova che tale metodo soddisfa

anch'esso (10) ed è anzi indistinguibile dal classico update completamente casuale.

Williams [66]

L'articolo indaga sulla dinamica sul modello indotta dalla dinamica della simulazione e su quali conseguenze ciò possa portare alla classe d'universalità. Precisamente si chiede se la realizzazione multi- o single- spin-flip della dinamica del modello appartengono alla stessa classe d'universalità dinamica.

Qui per msc si intende l'implementazione à la Zorn, dove il modello di Ising è diviso in sottoreticoli in modo che primi-vicini non appartengano allo stesso sottoreticolo; gli spin di ogni sottoreticolo sono presi in considerazione collettivamente ed è effettuato l'update — via Metropolis/Glauber — sequenzialmente o parallelamente a seconda dell'architettura della macchina.

Uno spin-flip consiste di due passaggi:

1. la valutazione di $\omega_j[\sigma]$ per il sito j ;
2. la decisione di flippare o meno j .

Qui $\omega_j[\sigma]$ è un rate che può essere, tra gli altri, Metropolis o Glauber. Rispettivamente $\omega_j[\sigma] = \min\{1, e^{-2\sigma_j \mathcal{H}_j}\}$ e $\omega_j[\sigma] = \frac{1}{2}[1 - \sigma_j \tanh(\mathcal{H}_j)]$.

Il primo punto può essere tranquillamente essere implementato via msc dato che rende attivi solo spin non pv, la cui $\Delta\mathcal{H}$ dipende solo dai pv, che non sono attivi.

Per testare il tutto è stato utilizzato usato un reticolo 4×4 confrontando i risultati numerici (Glauber) con quelli analitici esatti di $\langle |M| \rangle$ e di $\langle \sigma_i \sigma_j \rangle$. Il buon accordo tra il risultato simulato e quello esatto suggeriscono che la procedura di update in parallelo funziona effettivamente per le proprietà statiche d'equilibrio.

Il modello di Ising non ha una dinamica intrinseca, ma può essere introdotta una dinamica stocastica che può essere interpretata, nel caso di Glauber, come l'accoppiamento ad un serbatoio di calore di capacità infinita che induce spin-flip casuali sugli spin.

Sia

$$\Omega[\sigma] = \sum_i \omega_i[\sigma]. \quad (11)$$

Allora la probabilità che per t' non accada nulla e nel dt successivo flippi uno spin è data da

$$p(t')dt' = \Omega[\sigma]e^{-\Omega[\sigma]t'}dt' \quad (12)$$

e la probabilità che lo spin flippato sia σ_i è

$$\frac{\omega_i[\sigma]}{\Omega[\sigma]}. \quad (13)$$

Possiamo introdurre due schemi per simulare la dinamica. Uno è detto "evento-per-evento" ed è una realizzazione fedele del processo al continuo; l'altro è detto "tempo-per-tempo" ed è una interpretazione al tempo discreto del metodo Monte Carlo standard.

Primo schema. Generiamo due numeri casuali uniformemente distribuiti tra 0 ed 1 per determinare quando t' e che σ_i spin flippare in accordo a (12) e (13). t' sarà dato da

$$t' = -\frac{(\log x)}{\Omega}. \quad (14)$$

Per determinare i , l'intervallo aperto unitario è diviso in N porzioni di lunghezza data da (13), ognuna assegnata ad uno spin. quello che flippa è individuato dalla porzione in cui cade il numero casuale.

Secondo schema. e' lo standard single-spin-flip MC, con il sito attivo determinato casualmente. Ricordiamo che la scala temporale è data da un Monte Carlo Step (MSC), che equivale a N flip-trial. Per valori elevati di N , l'intervallo δt è piccolo rispetto ad ogni scala fisica e nel limite termodinamico l'approssimazione diviene esatta.

Nel primo schema per ogni step elementare uno spin è updatato, motivo per cui viene detto evento-per-evento; nel secondo uno step elementare consiste di un flip-trial che ha luogo ad intervalli prefissati. Quindi è detto tempo-per-tempo. Ovviamente la corrispondenza tra tempo e numero di configurazioni generate è differente nei due casi. Calcoliamo il tempo medio τ in cui uno spin flippa, in entrambi i casi.

Nel primo caso

$$\tau = \int_0^{\infty} dt t \Omega e^{-\Omega t} = \frac{1}{\Omega} \quad (15)$$

Nel secondo caso, sia P_k la probabilità che il prossimo evento accada nello step k . Avremo

$$P_1 = \sum_i \frac{\omega_i}{N} = \frac{\Omega}{N} \quad (16)$$

$$P_k = \left(1 - \frac{\Omega}{N}\right)^{k-1} \frac{\Omega}{N}$$

e τ è dato da

$$\begin{aligned}\tau &= \sum_k k P_k = \frac{N}{\Omega} \quad \text{in steps} \\ &= \frac{1}{\Omega} \quad \text{in MCS/spin.}\end{aligned}\tag{17}$$

Quindi, una unità di tempo nel primo schema equivale a un MCS/spin nel secondo schema.

Il primo schema va bene per temperature basse, il secondo per temperature alte.

1986: Inizia il msc su reticoli in parallelo

Bhanot, Duke e Salvador [3, 4]

Il modello simulato in questo articolo è Ising 3D, con update sequenziali a scacchiera, probabilità Metropolis, su reticoli di dimensioni tra 8 e 44, su un CDC CYBER 205. La macchina ha parole di 64 bit, così si simula contemporaneamente l'evoluzione di 64 reticoli.

La richiesta di maggior velocità negli algoritmi per lo studio di modelli stocastici (e.g. il modello di Ising), è causata dal fatto che, nelle vicinanze del punto critico, la lunghezza di correlazione ξ diverge e con essa (solitamente come ξ^2) diverge il CPU-time necessario per produrre due configurazioni successive indipendenti. Il collo di bottiglia di tutte le simulazioni è l'esigenza, ogni volta che si deve effettuare una decisione d'update, di un confronto tra \mathcal{H} ed un numero random generato per l'occasione. La generazione di numeri casuali porta via troppo tempo. Allora facciamo sì che il confronto sia effettuato solo con una breve sequenza di operazioni logiche tra bit.

C'è l'accenno ad un altro metodo per evitare tale collo di bottiglia: il metodo sviluppato da Williams e Kalos [65], vedi §4.2.

Procedura:

1. uso una variazione del msc in modo da mettere in una parola n siti da reticoli differenti. Qui (CDC CYBER 205) $n = 64$;
2. codifichiamo l'informazione per ogni numero random in 2 bit;
3. ogni coppia di bit è usata esattamente una volta per ognuno dei 64 reticoli;
4. usiamo operazioni logiche per gli update. Quindi 64 spin sono trattati contemporaneamente.

In ogni simulazione Metropolis si usa una sequenza finita di numeri casuali da comparare a $\Delta\mathcal{H}$. In tal modo tale sequenza non è distribuita uniformemente in $(0,1)$, e ciò significa che β non ha il valore desiderato, ma piuttosto un qualche β_{eff} . Per le nostre simulazioni, tale valore è dato da:

$$\begin{aligned}n_1/n &= e^{-4\beta_{\text{eff},1}} \\n_2/n &= e^{-8\beta_{\text{eff},2}} \\n_3/n &= e^{-12\beta_{\text{eff},3}},\end{aligned}$$

dove n è il numero totale dei random usati, ed n_i il numero di volte che tali random sono stati minori di $e^{-4i\beta}$.

La ragione per preoccuparsi di tale errore è che nelle simulazioni abbiamo usato, per ognuno dei 64 reticoli, lo stesso set di random. Quindi ogni lattice ha lo stesso β_{eff} e gli shift si sommano coerentemente. Se invece i vari random fossero sempre diversi, gli effetti tenderebbero ad annullarsi statisticamente.

Michael [47]

Se ho W bit in una parola, posso immagazzinare in un'unica parola info su più siti. Un buon metodo è guardare ai siti come appartenenti a W reticoli differenti che voglio far evolvere in parallelo. Però un fattore che fa perdere tempo è il calcolo di W numeri casuali a virgola mobile, uno per sito; per ovviare a ciò devo trovare un modo per generare un unico random per ogni update parallelo di W siti. Qualcuno [3, 4], (vedi §4.2 *Bhanot, Duke e Salvador [3, 4]*) ha proposto di generare W numeri casuali, mischiarli ed usarli per W volte, ogni volta in ordine diverso. Potrebbe servire, ma può portare un errore sistematico dovuto a correlazioni tra i vari reticoli, soprattutto se sono tutti troppo caldi o troppo freddi.

Il metodo qui usato è quello descritto in §4.2, riportato da [41], implementato usando un array per i bit casuali di lunghezza 2^{15} . Ovviamente l'uso di tale metodo comporta la discretizzazione dei valori utilizzabili di β , tanto più fine, quanto la lunghezza dell'array cresce. La macchina usata era un IBM PC (probabilmente 8086) con 16 bits-per-word.

Ceccatto e Riera [13]

Le modifiche alla dinamica introdotte dal msc sono già state studiate in [66, 25], dove si sono ottenute conferme numeriche all'ipotesi di non influenza delle proprietà critiche all'equilibrio (e quindi della classe di universalità).

Qui si giunge alla stessa conclusione in maniera analitica studiando la Master Equation nel caso al continuo, con update ritardato, con update sequenziale dei sottoreticoli, con update casuale dei sottoreticoli.

1987: Kikuchi e Okabe [39]

Niente di nuovo: applicano semplicemente il msc a reticoli in parallelo di Bhanot et al. ([3, 4] e in §4.2 *Bhanot, Duke e Salvador [3, 4]*); l'unica differenza con l'originale è che, invece di far evolvere n reticoli alla stessa temperatura, qui vengono fatti evolvere con temperature differenti. Qui $n = 32$.

1988: Roland e Grant [54]

Qui trovo solo l'applicazione, su Cray X-MP del msc descritto in §4.2 *Bhanot, Duke e Salvador [3, 4]* [3]. Gli autori applicano il msc alla dinamica di scambio (Kawasaki).

1989: Zhang [67]

Da quando i supercomputer hanno iniziato ad essere usati per le simulazioni MC, i fisici hanno cercato algoritmi che fittassero al meglio tali architetture. Per l'ordinario modello di Ising, il primo passo fu fatto nel 1984 quando un msc altamente ottimizzato per un CDC scalare venne trasferito su di un CDC CYBER 205. Un po' di storia nota. . .

Ora la questione è se sia possibile estendere queste idee alla dinamica di scambio di Kawasaki.

Usando il msc (one-spin-one-bit con operazioni bitwise), il salto di particelle è implementato da uno XOR tra la coppia di parole, controllato da una parola di scambio: un bit 1 in tale parola causa lo scambio tra le due storage-word, mentre uno 0 lascia le cose inalterate. Quindi trovare un veloce algoritmo per il msc significa trovare un modo elegante ed efficiente di settare la parola di scambio in base alle probabilità di transizione.

Per sistemi all'equilibrio:
all'inizio tutti i bit della parola di scambio sono posti uguali ad 1, quindi

1. seleziona casualmente una coppia di primi-vicini;
2. calcola il numero di legami primi-vicini della coppia n ;
3. genera un numero casuale tra 0 e 1;

4. compara il numero casuale con le probabilità di transizione, che sono funzione di n ;
5. metti gli spin a 0 se $n \geq 5$;
6. metti gli spin a 0 se il numero casuale è maggiore della probabilità di transizione.

Tra le possibili frequenze di transizione, scegliamo quelle di Metropolis $\mathcal{P}(n) = \min\{1, e^{-4\beta J(n-5)}\}$.

If one try to apply this algorithm to nonequilibrium problem, more randomization may be needed. For example, we have applied it to a nonequilibrium stochastic lattice gas model in which the jumps in the z direction are biased by a constant external field E .

[...]

I believe that in any nonequilibrium simulations which do not obey the detailed balancing, one will encounter similar problems if one tries to apply vectorized multispin coding techniques.

1991

The ultrafast multispin coding

Questo è articolo a firma di Ferrenberg e Landau [20].

Sfortunatamente questo articolo non detiene l'ideazione del metodo, ma solamente la sua applicazione. Il merito è qui conferito a *Wansleben, Comput. Phys. Commun. 43, 315 (1987)*, che purtroppo non siamo riusciti ad ottenere.

Le simulazioni hanno qui riguardato un reticolo ipercubico tridimensionale $L \times L \times L$ con condizioni periodiche al contorno utilizzando questo efficiente multispin coding vettorizzato. Il sistema è stato decomposto in sottoreticoli compenetranti e sono stati immagazzinati in parole di 64 bit (in un CDC Cyber 205) fino a 64 spin per parola. Tutti gli spin di una parola sono flippati tutti in una volta l'intero processo è stato vettorizzato per essere ancora più veloce. Nello specifico, inoltre, è stato generato un nuovo numero casuale per ogni spin update.

Un metodo simile era già stato introdotto in [64], dove si confronta il msc scalare (su CDC Cyber 176, 60 bits-per-word) con una versione vettoriale (su CDC Cyber 205, 64 bits-per-word) su un modello di Ising tridimensionale.

La differenza implementativa è nelle istruzioni e nella maggior velocità di calcolo vettoriale, che raggiunge quella dei computer special-purpose. 3 bits-per-spin, sottoreticoli compenetranti e XORization per il calcolo dell'energia d'interazione. Qui, ogni volta che prendo una parola, flippo due spin per risparmiare tempo.

Wansleben e Landau [63]

L'articolo presenta lo studio del modello di Ising tridimensionale all'equilibrio con quanto già esposto in [64]. Praticamente nulla di nuovo. Ribadisce [66, 13] che la procedura mcs non affetta la cinetica a lungo termine del modello e che tale affermazione coincide con la loro esperienza in cui il numero di sottoreticoli usati non influenza il tempo di rilassamento. Tale tempo comunque dipende dalla specifica dinamica scelta attraverso le probabilità di transizione (cita: Kikuchi e Okabe; J. Phys. Soc. Jap. 55, 1359 (1986)). I risultati per gli esponenti critici restano comunque inalterati.

2006: Lee & Okabe, Reuse [41]

Secondo gli autori di questo articolo, gli ideatori del multispin coding sono stati Bhanot *et al.* [4] e Michael [47]. L'utilizzo per studi sul modello di Ising è stato praticato da Kikuchi e Okabe [39] e da Ito e Kanada (non recuperato). L'utilizzo su lattice gas (dinamica Kawasaki) da Gawlinski *et al.* [25], Roland e Grant [54] e da Zhang [67].

La tecnica è basata sul fatto che per sistemi con simmetrie discrete come Ising, è possibile sfruttare l'architettura del computer per velocizzare i tempi. Infatti l'informazione sullo stato di un singolo spin (-1,+1) è contenibile in un singolo bit; in base a tale osservazione si può sfruttare una parola come contenitore per più bit contemporaneamente. Per esempio gli spin di 64 sistemi possono essere stoccati in un'unica parola, se si ha un computer con parole a 64 bit (come il loro).

“The Monte Carlo spin flip process is executed with the logical commands, and 64 Ising systems are updated simultaneously with a single random number sequence.”

Nel caso di 64 sistemi, uno può impiegare sia sistemi con parametri (tipo la temperatura) differenti, sia sistemi con lo stesso parametro, ma con diverse sequenze di numeri casuali.

Il metodo è stato utilizzato anche per sistemi con dinamica Kawasaki [25, 54, 67].

L'articolo è dedicato, come dice il titolo, alla spiegazione dell'utilizzo del multispin coding col reweighting per sistemi fuori equilibrio. Il multispin è illustrato nel caso KLS ma può essere utilizzato per molti altri sistemi simili.

la struttura

Frequenze Metropolis: $T_{\beta, \epsilon}(\sigma'|\sigma) = \min\{1, e^{-\beta(\Delta\mathcal{H} - \epsilon E)}\}$

$$\text{Parametro d'ordine: } \rho = \frac{1}{L_x/2} \sum_{j=1}^{L_x} \left| \frac{1}{L_y} \sum_{i=1}^{L_y} n_{ij} - \frac{1}{2} \right|$$

$\Delta\mathcal{H} - \epsilon E$	Transition rate	$\mathbf{M} = (M_3, M_2, M_1, M_0)$
∞	0	0 0 0 0
12	$\exp(-12\beta)$	0 0 0 1
8	$\exp(-8\beta)$	0 0 1 0
4	$\exp(-4\beta)$	0 0 1 1
0	1	0 1 0 0
-4	1	0 1 0 1
-8	1	0 1 1 0
-12	1	0 1 1 1
$-\infty$	1	1 0 0 0

Tabella 1: *Mapping of $\Delta\mathcal{H} - \epsilon E$ to a four-bit pattern \mathbf{M} .*

The transition rates are also given.

Nella procedura Kawasaki prendo a caso una coppia di spin e , se sono opposti, li scambiano col rate (4.2). Come si può notare dalla Table 1 i possibili valori di $\Delta\mathcal{H} - \epsilon E$, nel caso $E = \infty$, sono $\infty, +12, +8, +4, 0, -4, -8, -12, -\infty$, dove $+12, +8, +4, 0, -4, -8, -12$ sono presenti nella direzione ad trasversa e $+\infty, -\infty$ nella direzione longitudinale. Tali 9 possibili valori sono mappabili in una parola di 4 bit $\mathbf{M} = (M_3, M_2, M_1, M_0)$.

Nel msc introduciamo una variabile ausiliaria a 2 bit p che prende valori da 0 a 3 per implementare la procedura di accettazione/rigetto delle mosse.

Genero ora 2 array $p0[D]$ e $p1[D]$ — dove la dimensione D è scelta dall'utente molto grande: qui 2^{24} — e li riempio in modo da rispettare i rapporti

largh	$D * e^{-12\beta}$	$D * (e^{-8\beta} - e^{-12\beta})$	$D * (e^{-4\beta} - e^{-8\beta})$	$D * (1 - e^{-4\beta})$
p0	1 ... 1	0 ... 0	1 ... 1	0 0
p1	1 ... 1	1 ... 1	0 ... 0	0 0

Tabella 2: $p0$ e $p1$.

tra le probabilità di salto, come in Table 2. Come si vede, ora, considerando la parola $p = (p1, p0)$ ho

- $\mathcal{P}((p1[\lambda], p0[\lambda]) \geq 0) = 1;$
- $\mathcal{P}((p1[\lambda], p0[\lambda]) \geq 1) = e^{-4\beta};$
- $\mathcal{P}((p1[\lambda], p0[\lambda]) \geq 2) = e^{-8\beta};$
- $\mathcal{P}((p1[\lambda], p0[\lambda]) \geq 3) = e^{-12\beta};$

se λ è un numero casuale estratto tra 0 e MAX-1 compresi.

Gli array $p1$ e $p0$ sono costruiti bit-a-bit e poi mischiati, stando attenti a cambiare di posto allo stesso modo i bit di entrambi.

la procedura

1. prendo due siti $n(i, j)$ e $n(i', j')$;
2. da $\Delta\mathcal{H} - \epsilon E$ calcolo $M = (M3, M2, M1, M0)$;
3. genero λ in $[0, D-1]$;
4.
 - se $8 * M3 + 4 * M2 + 2 * M1 + M0 + 2 * p1[\lambda] + p0[\lambda] \geq 4$ pongo il bit di accettazione $A=1$;
 - altrimenti pongo $A=0$.

il tutto può essere implementato come in Table 3;

5. scambio le particelle con uno XOR \oplus tra A e gli spin dei siti scelti:

$$n(i, j) = n(i, j) \oplus A$$

$$n(i', j') = n(i', j') \oplus A.$$

M	p1	p0	A
0000	0	0	0
0001	0	0	0
0010	0	0	0
0011	0	0	0
0100	0	0	1
0101	0	0	1
0110	0	0	1
0111	0	0	1
1111	0	0	1

M	p1	p0	A
0000	0	1	0
0001	0	1	0
0010	0	1	0
0011	0	1	1
0100	0	1	1
0101	0	1	1
0110	0	1	1
0111	0	1	1
1111	0	1	1

M	p1	p0	A
0000	1	0	0
0001	1	0	0
0010	1	0	1
0011	1	0	1
0100	1	0	1
0101	1	0	1
0110	1	0	1
0111	1	0	1
1111	1	0	1

M	p1	p0	A
0000	1	1	0
0001	1	1	1
0010	1	1	1
0011	1	1	1
0100	1	1	1
0101	1	1	1
0110	1	1	1
0111	1	1	1
1111	1	1	1

Tabella 3: *Tavola di verità per A in funzione di M, p1 e p0***Claudio's note**

Se genero 64 sequenze di 2^{24} bit e poi le mischio in maniera differente, per farlo devo comunque generare 64×2^{24} numeri casuali, tanto vale farlo di volta in volta. Inoltre non capisco perchè, una volta mischiati, devo generare un numero casuale per andare a prenderli... tanto vale farlo in sequenza. O no? In ogni caso l'idea è ottima; è l'applicazione che lascia a desiderare.

BIBLIOGRAFIA

- [1] Daniel J. Amit, *Field theory, the renormalization group, and critical phenomena*, second ed., World Scientific Publishing Co, 1979.
- [2] Helen Au-Yang and Michael E. Fisher, *Bounded and inhomogeneous Ising models. ii. specific-heat scaling function for a strip*, *Physical Review B* **11** (1975), no. 9, 3469–3487.
- [3] G. Bhanot, D. Duke, and R. Salvador, *A fast algorithm for the Cyber 205 to simulate the 3D Ising model*, *Journal of Statistical Physics* **44** (1986), no. 5-6, 985–1002.
- [4] ———, *Finite-size scaling and the three-dimensional Ising model*, *Physical Review B* **33** (1986), no. 11, 7841–7844.
- [5] Gyan Bhanot, *The Metropolis algorithm*, *Report on Progress in Physics* **51** (1988), 429–457.
- [6] K. Binder and D.W. Heermann, *Monte Carlo simulation in statistical physics*, Springer Series in Solid-State Sciences, vol. 80, Springer, 1988.
- [7] Kurt Binder, *Applications of Monte Carlo methods to statistical physics*, *Report on Progress in Physics* **60** (1997), 487–559.
- [8] Stephen G. Brush, *History of the Lenz-Ising model*, *Review of Modern Physics* **39** (1967), no. 4, 883–893.
- [9] Theodore W. Burkhardt and Bernard Derrida, *Moments of the total magnetization and conformal invariance in the finite two-dimensional Ising model*, *Physical Review B* **32** (1985), no. 11, 7273–7279.
- [10] Sergio Caracciolo, Andrea Gambassi, Massimiliano Gubinelli, and Andrea Pelissetto, *Finite-size scaling in the driven lattice gas*, *Journal of Statistical Physics* **115** (2004), no. 1 - 2, 281–322.

BIBLIOGRAFIA

- [11] Francesco Caravenna, *Diagramma di fase di un gas quasi unidimensionale fuori dall'equilibrio termodinamico*, Ph.D. thesis, Università degli Studi, Pisa, 2002.
- [12] George Casella, Michael Lavine, and Christian P. Robert, *Explaining the perfect sampler*, *The American Statistician* **55** (2001), no. 4, 299–305.
- [13] H. A. Ceccatto and J. A. Riera, *Effective dynamics in Monte Carlo simulations with multispin updates*, *Physical Review B* **34** (1986), no. 7, 4874–4877.
- [14] Michael Creutz, Laurence Jacobs, and Claudio Rebbi, *Experiments with a gauge-invariant Ising system*, *Physical Review Letters* **42** (1979), no. 21, 1390–1393.
- [15] ———, *Monte Carlo study of Abelian lattice gauge theories*, *Physical Review D* **20** (1979), no. 8, 1915–1922.
- [16] F. de los Santos and P. L. Garrido, *Continuum field model of driven lattice gases*, *Journal of Statistical Physics* **96** (1999), no. 1 - 2, 303–324.
- [17] Xenia K. Dimakos, *A guide to exact simulation*, *International Statistical Review* **69** (2001), no. 1, 27.
- [18] Andrzej Drzewinski and Jacek Wojtkiewicz, *Cumulant ratios and their scaling functions for Ising systems in a strip geometry*, *Physical Review E* **62** (2000), no. 3, 4397–4400.
- [19] M. R. Evans, *Phase transitions in one-dimensional nonequilibrium systems*, *Brazilian Journal of Physics* **30** (2000), 42.
- [20] Alan M. Ferrenberg and D. P. Landau, *Critical behavior of the three-dimensional Ising model: A high-resolution Monte Carlo study*, *Physical Review B* **44** (1991), no. 10, 5081–5091.
- [21] Lloyd D. Fosdick, *Calculation of order parameters in a binary alloy by the Monte Carlo method*, *Physical Review* **116** (1959), no. 3, 565–573.
- [22] R. Friedberg and J. E. Cameron, *Test of the Monte Carlo method: Fast simulation of a small Ising lattice*, *The Journal of Chemical Physics* **52** (1970), no. 12, 6049–6058.
- [23] Andrea Gambassi, *Dynamic critical behavior of non-equilibrium systems*, Ph.D. thesis, Scuola Normale Superiore, Pisa, 2002.

-
- [24] P. L. Garrido, F. de los Santos, and M. A. Muñoz, *Langevin equation for driven diffusive systems*, *Physical Review E* **57** (1998), no. 1, 752–755.
- [25] E. T. Gawłinski, Martin Grant, J. D. Gunton, and K. Kaski, *Growth of unstable domains in the two-dimensional Ising model*, *Physical Review B* **31** (1985), no. 1, 281–286.
- [26] Roy J. Glauber, *Time-dependent statistics of the Ising model*, *Journal of Mathematical Physics* **4** (1963), no. 2, 294–307, Glauber algorithm.
- [27] Robert B. Griffiths, *Peierls proof of spontaneous magnetization in a two-dimensional Ising ferromagnet*, *Physical Review* **136** (1964), no. 2A, A437–A439.
- [28] Massimiliano Gubinelli, *Finite-size scaling in non-equilibrium critical phenomena*, Ph.D. thesis, Università degli Studi, Pisa, 2002.
- [29] Olle Haggstrom, *Finite Markov chains and algorithmic application*, Tech. report, Goteborg Universitet, 2001.
- [30] P. C. Hohenberg and B. I. Halperin, *Theory of dynamic critical phenomena*, *Review of Modern Physics* **49** (1977), no. 3, 435–479.
- [31] Ming-Chang Huang, Tsong-Ming Liaw, Yu-Pin Luo, and Simon C. Lin, *Finite-size scaling of partition function zeros and first-order phase transition for infinitely long Ising cylinder*, arXiv **cond-mat/0407731** (2004), 0.
- [32] Ernst Ising, *Beitrag zur theorie des ferromagnetismus*, *Z. Physik* **31** (1925), 253–258.
- [33] H. K. Janssen and B. Schmittmann, *Field theory of critical behaviour in driven diffusive systems*, *Zeitschrift für Physik B Condensed Matter* **64** (1986), no. 4, 503–514.
- [34] C. Kalle and V. Winkelmann, *Monte Carlo technique for very large Ising models*, *Journal of Statistical Physics* **28** (1982), no. 4, 639–648.
- [35] Sheldon Katz, Joel L. Lebowitz, and H. Spohn, *Phase transitions in stationary nonequilibrium states of model lattice systems*, *Physical Review B* **28** (1983), no. 3, 1655–1658.
- [36] Sheldon Katz, Joel L. Lebowitz, and Herbert Spohn, *Nonequilibrium steady state of stochastic lattice gas models of fast ionic conductors*, *Journal of Statistical Physics* **34** (1984), no. 3-4, 497–537.

BIBLIOGRAFIA

- [37] K. Kawasaki, *Kinetics of Ising models*, Phase transition and critical phenomena (C. Domb and M.S. Green, eds.), Phase transition and critical phenomena, vol. 2, Academic Press, 1972, pp. 443–501.
- [38] Kyozi Kawasaki, *Diffusion constants near the critical point for time-dependent Ising Models. I*, Phys. Rev. **145** (1966), no. 1, 224–230.
- [39] Macoto Kikuchi and Yutaka Okabe, *Renormalization, self-similarity, and relaxation of order-parameter structure in critical phenomena*, Physical Review B **35** (1987), no. 10, 5282–5384.
- [40] W. Kwak, D. P. Landau, and Beate Schmittmann, *Driven diffusive systems: How steady states depend on dynamics*, Physical Review E **69** (2004), no. 6, 066134.
- [41] Hwee Kuan Lee and Yutaka Okabe, *Multispin coding technique for nonequilibrium reweighting*, International Journal of Modern Physics C **17** (2006), no. 2, 157–165.
- [42] T. D. Lee and C. N. Yang, *Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model*, Physical Review **87** (1952), no. 3, 410–419.
- [43] Kwan-tai Leung and John L. Cardy, *Field theory of critical behavior in a driven diffusive system*, Journal of Statistical Physics **44** (1986), no. 3 - 4, 567–588.
- [44] T. M. Liaw, M. C. Huang, S. C. Lin, and M. C. Wu, *Scaling functions of interfacial tensions for a class of Ising cylinders*, Physical Review B **60** (1999), no. 18, 12994–13005.
- [45] Torgny Lindvall, *Lectures on the coupling method*, Wiley Series in Probability & Mathematical Statistics, John Wiley & Sons, September 1992.
- [46] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller, *Equation of state calculation by fast computing machine*, The Journal of Chemical Physics **21** (1953), no. 7, 1087–1092.
- [47] C. Michael, *Fast heat-bath algorithm for the Ising model*, Physical Review B **33** (1986), no. 11, 7861–7862.

-
- [48] M. A. Novotny, *Introduction to the Propp-Wilson method of exact sampling for the Ising model*, Computer Simulation Studies in Condensed Matter Physics XII (D. P. Landau, S. P. Lewis, and H.-B. Schüttler, eds.), Springer Proceedings in Physics, vol. 85, Springer Verlag, 2000, p. 179.
- [49] Lars Onsager, *Crystal statistics I. A two-dimensional model with an order-disorder transition*, Physical Review **65** (1944), no. 3-4, 117–149.
- [50] R. Peierls, *On Ising's model of ferromagnetism*, Proc. Cambridge Phil. Soc. **32** (1936), 477–481.
- [51] A. Pelissetto, *Introduction to the Monte Carlo method*, Elementary particles, quantum fields and statistical mechanics. (M. Bonini, G. Marchesini, and E. Onofri, eds.), Lectures given at the summer school of theoretical physics, Parma 1991–1993, 1992, pp. 111–153.
- [52] James Gary Propp and David Bruce Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures and Algorithms **9** (1996), no. 1-2, 223–252.
- [53] Claudio Rebbi and Robert H. Swendsen, *Monte Carlo renormalization-group studies of q-state Potts models in two dimensions*, Physical Review B **21** (1980), no. 9, 4094–4107.
- [54] Christopher Roland and Martin Grant, *Monte Carlo renormalization-group study of the late-stage dynamics of spinodal decomposition*, Physical Review Letters **60** (1988), no. 25, 2657–2660.
- [55] Jesús Salas, *Exact finite-size-scaling corrections to the critical two-dimensional Ising model on a torus*, Journal of Physics A **34** (2001), 1311–1331.
- [56] B. Schmittmann, H. K. Janssen, U. C. Täuber, R. K. P. Zia, K.-T. Leung, and J. L. Cardy, *Viability of competing field theories for the driven lattice gas*, Physical Review E **61** (2000), 5977–5980.
- [57] Beate Schmittmann and Royce K. P. Zia, *Statistical mechanics of driven diffusive systems*, vol. 17, Academic, NY, 1995.
- [58] A. Sokal, *Monte Carlo methods in statistical mechanics: Foundations and new algorithms*, Cargese Summer School Lecture Notes, 1996.

BIBLIOGRAFIA

- [59] Robert H. Swendsen and Jian-Sheng Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, *Physical Review Letters* **58** (1987), no. 2, 86–88.
- [60] Kwan tai Leung and John L. Cardy, *Erratum: Field theory of critical behavior in a driven diffusive system*, *Journal of Statistical Physics* **45** (1986), 1087.
- [61] Hal Tasaki, *A remark on the choice of stochastic transition rates in driven nonequilibrium systems*, arXiv, 2004.
- [62] Jian-Sheng Wang, *Anisotropic finite-size scaling analysis of a two-dimensional driven diffusive system*, *Journal of Statistical Physics* **82** (1996), 1409–1427.
- [63] S. Wansleben and D. P. Landau, *Monte Carlo investigation of critical dynamics in the three-dimensional Ising model*, *Physical Review B* **43** (1991), no. 7, 6006–6014.
- [64] Stephan Wansleben, John G. Zabolitzky, and Claus Kalle, *Monte Carlo simulation of Ising models by multispin coding on a vector computer*, *Journal of Statistical Physics* **37** (1984), no. 3-4, 271–282.
- [65] G. O. Williams and M. H. Kalos, *A new multispin coding algorithm for Monte Carlo simulation of the Ising model*, *Journal of Statistical Physics* **37** (1984), no. 3-4, 283.
- [66] J. K. Williams, *Parallel processing: the Ising model and Monte Carlo dynamics*, *Journal of Physics A* **18** (1985), no. 10, 1781–1794.
- [67] M. Q. Zhang, *A fast vectorized multispin coding algorithm for 3D Monte Carlo simulations using Kawasaki spin-exchange dynamics*, *Journal of Statistical Physics* **56** (1989), no. 5-6, 939–950.
- [68] Jean Zinn-Justin, *Quantum field theory and critical phenomena*, The International Series of Monographs on Physics, vol. 113, Oxford Science Publications, 1996.
- [69] R. Zorn, H. J. Herrmann, and C. Rebbi, *Tests of the multi-spin-coding technique in Monte Carlo simulations of statistical systems*, *Computer Physics Communications* **23** (1981), 337–342.