

بسم الله الرحمن الرحيم

مقدمه‌ای بر مونت کارلوی کوانتومی و روش‌های بهینه‌سازی

مهدی نیک عمل

دانشگاه تربیت دبیر شهید رجائی

فهرست

۱	مقدمه	۴
۲	روش‌های محاسباتی ساختار مواد	۶
۳	مفاهیم اصلی روش مونت کارلو	۸
۳-۱	انتگرال گیری‌های مونت کارلو	۹
۳-۱-۱	روش برخورد یا خطا	۹
۳-۱-۲	روش مقدار میانگین	۱۰
۳-۱-۳	نمونه‌گیری با اهمیت	۱۰
۳-۲	الگوریتم متروپولیس	۱۱
۴	روش‌های مونت کارلوی کوانتومی	۱۳
۵	مونت کارلوی کوانتومی وردشی (VMC)	۱۵
۵-۱	مقدمه	۱۵
۵-۲	اصل وردش	۱۵
۵-۳	مونت کارلوی کوانتومی وردشی	۱۶
۵-۳-۱	انرژی موضعی	۱۹
۶	روش مونت کارلوی کوانتومی پخشی	۲۰
۶-۱	روش مونت کارلو پخشی هدایت شده	۲۷
۶-۲	خطای روش مونت کارلو کوانتومی پخشی	۳۱
۷	اشاره‌ای به روش مونت کارلوی کوانتومی انتگرال مسیر	۳۲
۸	بررسی خطای محاسبات QMC	۳۷
۸-۱	خطاهای دستگاهی	۳۸
۹	توابع موج مونت کارلوی کوانتومی	۳۹
۹-۱	مقدمه	۳۹
۹-۲	تقارن	۳۹
۹-۳	تابع موج آزمایشی در محاسبات مونت کارلوی کوانتومی وردشی	۳۹
۹-۴	مجموعه پایه‌ها	۴۱
۹-۴-۱	اوربیتال‌های نوع نمایی	۴۱
۹-۴-۲	اوربیتال‌های نوع گوسی	۴۳
۹-۵	تابع جسترو	۴۴
۹-۵-۱	تابع جسترو دو ذره‌ای	۴۵
۹-۵-۲	تابع جسترو اسمیت - ماسکویتز	۴۵
۹-۶	شرایط کاسپ	۴۷
۹-۶-۱	شرایط کاسپ الکترون - الکترون	۴۷
۹-۶-۲	شرایط کاسپ الکترون - هسته	۴۸

۱۰ برخی مثال‌ها ۵۰

- ۵۰ ۱-۱۰ تابع موج اتم‌های بس الکترونی
- ۵۴ ۲-۱۰ تابع موج مولکول‌های H_2O, Li_2, H_2
- ۵۵ ۱-۲-۱۰ مولکول هیدروژن
- ۵۷ ۲-۲-۱۰ مولکول لیتیوم
- ۵۸ ۳-۲-۱۰ مولکول آب

۱۱ روش‌های بهینه‌سازی ۵۹

- ۵۹ ۱-۱۱ مقدمه
- ۵۹ ۲-۱۱ کمینه‌سازی واریانس انرژی
- ۶۰ ۳-۱۱ کمینه‌سازی انرژی
- ۶۱ ۴-۱۱ روش‌های بهینه‌سازی
- ۶۱ ۱-۴-۱۱ روش کاهش شیب (GD)
- ۶۳ ۲-۴-۱۱ روش نیوتن-رفسون
- ۶۴ ۳-۴-۱۱ گرادیان همبسته (CG)
- ۶۵ ۵-۱۱ روش شبیه‌سازی مونت کارلو سردسازی
- ۶۵ ۱-۵-۱۱ روش شبیه‌سازی مونت کارلو
- ۶۷ ۲-۵-۱۱ روش شبیه‌سازی سردسازی
- ۶۹ ۳-۵-۱۱ روش کمینه‌یابی مونت کارلو سردسازی
- ۷۰ ۶-۱۱ الگوریتم ژنتیک (GA)
- ۷۰ ۱-۶-۱۱ مفاهیم پایه
- ۷۱ ۲-۶-۱۱ جمعیت
- ۷۲ ۳-۶-۱۱ عملگر جهش
- ۷۳ ۴-۶-۱۱ عملگر ادغام
- ۷۴ ۵-۶-۱۱ عملگر گزینش

پیوست‌ها ۷۶

- ۷۶ پیوست الف: جمله سوق برای نمونه‌گیری با اهمیت
- ۷۸ پیوست ب: محاسبه‌ی گرادیان و لاپلاسی دترمینان‌های اسلیتر و عامل جسترو
- ۷۸ ب-۱) محاسبه‌ی گرادیان و لاپلاسی دترمینان‌های اسلیتر
- ۷۹ ب-۲) گرادیان تابع جسترو اسمیت و ماسکویتز
- ۸۰ ب-۳) لاپلاسی تابع جسترو اسمیت و ماسکویتز
- ۸۱ پیوست ج: محاسبه‌ی مشتق یکم و دوم دترمینان‌های اسلیتر نسبت به پارامترها
- ۸۱ ج-۱) مشتق یکم دترمینان‌های اسلیتر نسبت به پارامترها
- ۸۲ ج-۲) مشتق دوم دترمینان‌های اسلیتر نسبت به پارامترها
- ۸۳ ج-۳) محاسبه‌ی مشتق انرژی موضعی نسبت به پارامترهای وردش

۱ مقدمه

در چند دهه گذشته شاهد پیشرفت‌های چشمگیری در توانایی شبیه‌سازی سیستم‌های پیچیده فیزیکی با استفاده از کامپیوترها بوده‌ایم. این موضوع باعث شده که شبیه‌سازی کامپیوتری به عنوان راه‌سومی برای تحقیقات علمی باشد، که به آزمایش نزدیک‌تر است تا به نظریه و نظریه پردازی، اما مکملی برای هر دو آنها می‌باشد. شبیه‌سازی همانند آزمایش، اطلاعات نسبتاً بیشتری نسبت به نظریه تولید می‌کند.

در مطالعه خواص کوانتومی جامدات، دشواری از پیچیدگی معادلات ناشی می‌شود. به سختی ممکن به نظر می‌رسد بتوان معادله شرودینگر $3N$ بعدی را که N الکترون برهمکنشی در یک جامد را توصیف می‌کند به طور تحلیلی حل کرد. جواب دقیق همان چیزی بوده که روش‌های مونت کارلوی کوانتومی در ابتدا مدعی انجام آنها با خطای آماری اندک بوده است. با استفاده از کامپیوترهای موازی می‌توانیم سیستم‌های دوره‌ای دربرگیرنده بیش از هزار الکترون را شبیه‌سازی کنیم، که برای مدل‌سازی بسیاری از خواص جامدات با دقت بالا کافی می‌باشد. امروزه روش‌های مونت کارلوی کوانتومی به عنوان یک ابزار با ارزش برای برآورد خواص فیزیکی مواد حقیقی به رسمیت شناخته شده‌اند. در مورد عمومیت این روش می‌توان گفت نسبت به نظریه تابعی چگالی عمومیت کمتری دارند. این موضوع میتواند به دلایل مختلف صورت گرفته باشد از جمله طول عمر کوتاه، هزینه محاسباتی بالای آن و وابستگی عملی آن به روش‌های دیگر که اخیراً بیشتر نمود یافته است. امروزه مونت کارلوی کوانتومی به سمت و سویی می‌رود که به کمک خروجی‌های نظریه تابعی چگالی و... بتوان ورودی‌هایی را برای محاسبات مونت کارلوی کوانتومی ایجاد کرد و بالعکس.

همانطور که گفتیم معادله‌ی شرودینگر غیر نسبیتی را به عنوان اساسی‌ترین معادله-ی تابع موج مکانیک کوانتومی در حد غیر نسبیتی، در بیشتر موارد نمی‌توان به صورت تحلیلی حل کرد. از این‌رو باید به دنبال راهی برای حل عددی معادله‌ی شرودینگر در دستگاه‌های بس ذره‌ای مورد نظر باشیم. نخستین گام برای حل معادله‌ی شرودینگر در دستگاه‌های بس ذره‌ای، معادلات میدان خودسازگار هارتری-فوک^۱ بودند [۱۲]. مجموعه معادلات هارتری-فوک (HF) از یک پتانسیل میانگین در به دست آوردن نیروی وارد بر الکترون از طرف الکترون‌های دیگر دستگاه، استفاده می‌کند. در معادلات HF، به طور آشکار برهم‌کنش الکترون-الکترون در نظر گرفته نمی‌شود و از تقریب میدان متوسط استفاده می‌شود. تقریباً این ایده میدان متوسط در روش‌های دیگر محاسباتی نیز گنجانده شده است. این تقریب در مورد سیستم‌های با همبستگی قوی و بعضاً پیوندهای وادروالسی ضعیف با شکست روبرو شده و مثال‌های مختلفی برای جواب‌های نادرست نظریات مبتنی بر میدان متوسط وجود دارد.

¹Hartree-Fock

تنها راه برای اندازه‌گیری دقیق برهم‌کنش الکترون-الکترون دور شدن از تقریب میدان متوسط و در نظر گرفتن مختصات الکترونها در تابع موج دستگاه است. در نظر گرفتن سهم برهم‌کنش الکترون-الکترون در هامیلتونی دستگاه، محاسبه‌ی انتگرال‌هایی که منجر به حل معادله‌ی شرودینگر می‌شد را دشوار می‌ساخت. از این‌رو تلاش‌های تحلیلی نخستین در بررسی دستگاه‌های بس‌ذره‌ای به اتم‌های ساده (۲ تا ۴ الکترون) محدود شدند [۳].

یکی از موفقیت‌های اولیه و مشهور در مورد نقش مهمی که به وسیله روش‌های مونت کارلو کوانتومی در حوزه نظریه ساختار الکترونی ایفا می‌گردد، محاسبه انرژی تبادلی-همبستگی برای تقریب چگالی موضعی^۲ (عنصر اساسی سازنده نظریه تابعی چگالی^۳) به واسطه محاسبه همبستگی برای گاز الکترونی می‌باشد [۱]. اخیراً، مونت کارلو کوانتومی در جهت دستیابی به فهم بهتری از DFT به عنوان فراگیرترین روش محاسباتی ساختار الکترونی کمک مضاعفی بوده است، چه در بنا نهادن تخمینی از انرژی تبادلی-همبستگی یا از طریق تولید مقادیر دقیق تابعی چگالی [۲ و ۳]. علاوه بر این، پیشرفت قابل توجهی در کاربرد روش مونت کارلو کوانتومی در مسائل علوم مواد، از قبیل خواص بستگی خوشه‌ها^۴ [۴]، نقص‌ها در نیم‌رساناها [۵]، انرژی‌های ساختاری و واکنش‌ها در سطوح نیم‌رسانا [۶]، ترازهای برانگیخته مولکول‌ها [۷]، و جامدات [۸]، صورت گرفته است. (برای مروری جامع و با جزئیات در مورد کاربردهای QMC، مرجع [۴] را مشاهده کنید).

روش‌های مونت کارلو کوانتومی (QMC)^۵ از روش‌های پر قدرتی هستند که بسیار دقیق و قابل اطمینان می‌باشند و معمولاً دشوار و یا غیرممکن است که داده‌های حاصل از آنها را از هر روش محاسباتی دیگری بدست آورد.

از سویی دیگر روش‌های استاندارد انتگرال‌گیری (مانند روش ذوزنقه‌ای، سیمپسون و ...)، در محاسبه‌ی متوسط‌های ناشی از حل معادله‌ی شرودینگر زمان زیادی صرف می‌کنند و تنها برای دستگاه‌هایی با الکترون‌های کم سودمند هستند. روش‌های مونت کارلو، مجموعه روش‌هایی هستند که معمولاً با کمک فرایندهای کاتوره‌ای رفتار یک دستگاه را به صورت آماری بررسی می‌کنند. به بیان ساده تر روش‌های مونت کارلوی کوانتومی ایده‌ای شبیه همان روش مونت کارلو کلاسیک را می‌پروراند با این تفاوت که انتگرال‌های موجود برای حل انتگرال‌های ناشی از مقادیر چشم‌داشتی ظاهر شده در مکانیک کوانتومی می‌باشند.

²Local Density Approximation (LDA)

³Density Functional Theory (DFT)

⁴ Clusters

⁵ Quantum Monte Carlo Methods (QMC)

۲ روش‌های محاسباتی ساختار مواد

از آنجاکه با سیستم‌هایی با تعداد بسیار زیادی از ذرات سازنده (از مرتبه 10^{23}) روبرو هستیم، و خود این ذرات با هم در حال برهمکنش می‌باشند، عمده‌تاً دو روش اساسی برای بررسی سیستم بنا نهاده شده است:

(۱) روش کلاسیکی، در این روش با در نظر گرفتن یک پتانسیل بین اتمی مانند لندارد-جونز^۶ بین ذرات تشکیل دهنده یعنی اتم‌ها، خصوصیات مهمی بدست می‌آید. توجه داشته باشیم که ساخت چنین پتانسیلهایی خود درک کوانتومی نیاز دارد.

(۲) روش کوانتومی، این روش به محاسبات آغازین^۷ معروف می‌باشد، در واقع روش‌هایی که معادلاتشان شامل پارامترهای تجربی و نیمه تجربی نمی‌باشند و بدون دخالت داده‌های آزمایشگاهی مستقیماً از اصول تئوری حاصل می‌شوند، به طور کلی محاسبات آغازین نامیده می‌شوند و با یک معادله شرودینگر بس ذره‌ای با هامیلتونی به صورت زیر میتوان خصوصیات مهمی از سیستم را به دست آورد.

$$H\Psi(\{r_i\}, \{R_j\}) = E\Psi(\{r_i\}, \{R_j\}) \quad (1)$$

$$H = \sum_i \frac{-\hbar^2 \nabla_i^2}{2m_i} + \sum_j \frac{-\hbar^2 \nabla_j^2}{2M_j} + \sum_{ij} \frac{e^2 Z_j^2}{4\pi\epsilon_0 |\vec{r}_i - \vec{R}_j|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_{k \neq l} \frac{e^2 Z_k Z_l}{4\pi\epsilon_0 |\vec{R}_i - \vec{R}_j|} \quad (2)$$

که در این معادله $\vec{r}$ برداری است که محل الکترون را نشان می‌دهد و $\vec{R}$ موقعیت هسته را مشخص می‌کند.

این دو روش هر کدام معایب و مزایایی به همراه دارند. عیب عمده روش کلاسیکی این است که پتانسیل واحدی برای تمام مواد وجود ندارد و ناچاریم برای هر ماده و یا حداقل دسته‌ای از مواد پتانسیل خاص بسازیم. انتخاب پتانسیل هم به شدت به شرایط فیزیکی خاص حاکم بر سیستم وابسته می‌باشد، برای اینکه یک پتانسیل دقیق باشد و در شرایط مختلف برای مواد گوناگون قابل استفاده باشد نیازمند به استفاده از پارامترهای زیادی در عبارت پتانسیل خواهیم بود که این خود باعث افزایش حجم محاسبات می‌شود و اگر تعداد پارامترها را کم کنیم خطای پتانسیل افزایش می‌یابد. اما مزیت عمده آن سرعت بالای محاسبات و در نتیجه امکان محاسبه برای سیستم‌های بزرگ شامل هزاران ذره با رایانه‌ای با سرعت و حافظه معمولی می‌باشد. برخی از رفتارهای سیستم ناشی از خصوصیات ماکروسکوپی می‌باشد که بهره‌گیری از روش کلاسیک در اینجا بسیار سریعتر و ارزانتر می‌باشد.

⁶ Lenard-Jones

⁷ Ab-initio

مزیت روش کوانتومی در آن است که همانطور که قبلاً گفته شد خواص اساسی مواد بر پایه اصول مکانیک کوانتومی استوار است و در این روش نیز محاسبات بر همین اساس می‌باشد، به ویژه خواصی که منشأ کوانتومی دارند و در فرمول بندی کلاسیک ظاهر نمی‌شوند، نظیر ساختار نواری، چگونگی پیوند بین اتم‌ها تنها با روش‌های کوانتومی قابل بررسی می‌باشند.

تنگنای اصلی در روش‌های کوانتومی حجم زیاد محاسبات می‌باشد. حل دقیق معادله شرودینگر بس ذره‌ای از عهده بزرگترین رایانه‌ها برنمی‌آیند، بنابراین برای اینکه محاسبات به نتیجه برسد بایستی از کلیه اطلاعات موجود در مورد سیستم‌ها نظیر تقارن‌های ساختاری، تعداد اتم‌ها در سلول واحد، مکان‌های تقریبی اتم‌ها، به نحو هوشمندانه‌ای استفاده کنیم علی‌رغم این تمهیدات اعمال تقریب‌های متعدد اجتناب ناپذیر است. لذا همواره این سؤال مطرح است که نتایج حاصل تا چه حد قابل تطبیق با یک سیستم واقعی است به هر حال امروزه بسته به چگالی و دما و... کمیت قابل محاسبه از هر دو روش کلاسیک و کوانتومی بهره می‌گیرند.

آزمونی ساده برای تحقیق درستی فرضیه‌های کلاسیک، استفاده از طول‌موج گرمایی دوبروی،

$$\Lambda = \left(\frac{h^2}{2\pi m k_B T} \right)^{1/2}, \quad \text{اگر } \Lambda \text{ در مقایسه با ابعاد ظرف در برگیرنده دستگاه، خیلی کوچک}$$

باشد، یعنی $\Lambda^3 / V \ll 1$ ، آنگاه فرض کلاسیک بودن رفتار دستگاه، یک تقریب مناسب و قابل قبول است. همان‌طور که از رابطه بالا پیداست، برای دستگاه‌های خیلی سبک همچون H_2 ، He ، و Ne و همچنین در دماهای خیلی کم، تقریب کلاسیکی معتبر نیست و حتماً باید از روش‌های کوانتومی یا روش‌های کوانتومی میدان متوسط برای شبیه‌سازی دستگاه استفاده کنیم.

در روش‌های کوانتومی برای یافتن ساختار فیزیکی یک سیستم باید هامیلتونی مربوط به آن را با دقت بنویسیم و تلاش کنیم تا توابع موج و مقادیر ویژه آن را بدست آوریم. بنابراین در روش‌های محاسباتی ساختار مواد با دو مسئله اساسی روبرو خواهیم بود:

i. بازنویسی هامیلتونی سیستم

ii. حل معادله شرودینگر

در فصل دوم، روش‌های انتگرال‌گیری مونت کارلو، الگوریتم متروپلیس و مونت کارلوی کوانتومی وردشی بررسی می‌شود. گزینش تابع موج آزمایشی در روش VMC از اهمیت ویژه‌ای برخوردار است. در فصل سوم توابع موج مونت کارلوی کوانتومی و ویژگی‌های توابع موج در الگوریتم VMC را بررسی می‌کنیم. یافتن تابع موج حالت پایه‌ی یک دستگاه بس‌ذره‌ای با شمار زیادی از پارامترهای وردش تنها با به خدمت گرفتن روش‌های گوناگون بهینه‌سازی امکان‌پذیر است. امروزه ارائه‌ی الگوریتم‌های کارا در بهینه‌سازی تابع موج مونت کارلوی کوانتومی و یافتن نتایج بهتر مهم‌ترین چالش روش‌های QMC است [۱۲]. در فصل چهارم به بررسی سه روش بهینه‌سازی تندترین کاهش^۸، گرادیان همبسته^۹، مونت کارلو سرد سازی و الگوریتم ژنتیک^{۱۰} می‌پردازیم.

^۸Steepest Descent

۳ مفاهیم اصلی روش مونت کارلو

روش مونت کارلو، یک روش عددی برای حل مسائل ریاضی فیزیکی به وسیله شبیه‌سازی متغیرهای تصادفی می‌باشد. مهمترین کاربرد روشهای مونت کارلو محاسبه انتگرال‌های پیچیده است که با روش‌های دیگر به طور دقیق قابل حل نمی‌باشند. به طور عمومی تاریخ تولد پذیرفته شده روش منت-کارلو، سال ۱۹۴۹ می‌باشد، یعنی هنگامی که مقاله‌ای با نام روش مونت کارلو توسط متروپولیس^۹ و یولام^{۱۰} [۴۶] ارائه شد. ریاضی‌دانان آمریکایی جان وون نیومن^{۱۱} و استانیسلاو یولام^{۱۲} بنیانگذاران اصلی آن شمرده می‌شوند. پایه‌های نظری روش مدت‌ها قبل از اینکه مقاله وون نیومن - یولام منتشر شود، شناخته شده بود. از این گذشته پیش از سال ۱۹۴۹ برخی مسائل در آمار گاه بگاهی به وسیله نمونه‌گذاری تصادفی یعنی، در حقیقت، به وسیله روش مونت کارلو حل می‌شدند. به هر حال از آنجایی که شبیه‌سازی متغیرهای تصادفی با دست، کاری دشوار و پر زحمت می‌باشد، استفاده از روش مونت کارلو به عنوان یک تکنیک عددی همگانی تنها با ظهور کامپیوترها، به صورت کاربردی درآمد.

روش‌های مونت کارلو در حوزه گسترده‌ای کاربرد پیدا می‌کند، شامل بسیاری از زیر رشته‌های فیزیک، مانند فیزیک آماری یا فیزیک انرژی بالا، و گسترش می‌یابد به نواحی دیگر چون بیولوژی یا تحلیل بازار و دادوستدهای مالی.

غالباً مسأله اصلی تخمین انتگرال چند بعدی

$$I = \int dx f(x) \quad (3)$$

است، که برای این انتگرال پاسخ تحلیلی معلوم نمی‌باشد. برای اینکه دقیق‌تر و صریح‌تر باشیم، الگوریتمی را جستجو می‌کنیم که تخمینی عددی از انتگرال به همراه تخمینی از خطا ارائه دهد. علاوه بر این الگوریتم باید نتیجه را در زمان معقولی ارائه دهد. هیچ امتیازی در ایجاد الگوریتمی که نتیجه صحیحی را تحویل دهد، ولی تا ابد طول بکشد، وجود ندارد. همچنین هیچ الگوریتم کاملی وجود ندارد که برای همه مسائل مناسب باشد. هر چه بیشتر در مورد مسئله به خصوصی، آگاهی داشته باشیم، شانس بیشتری برای پیدا کردن الگوریتمی که مسئله را به طور کارآمدتری حل کند وجود دارد.

⁹Conjugate Gradient

¹⁰Genetic Algorithm

¹¹Metropolis

¹²Ulam

¹³John von Neumann

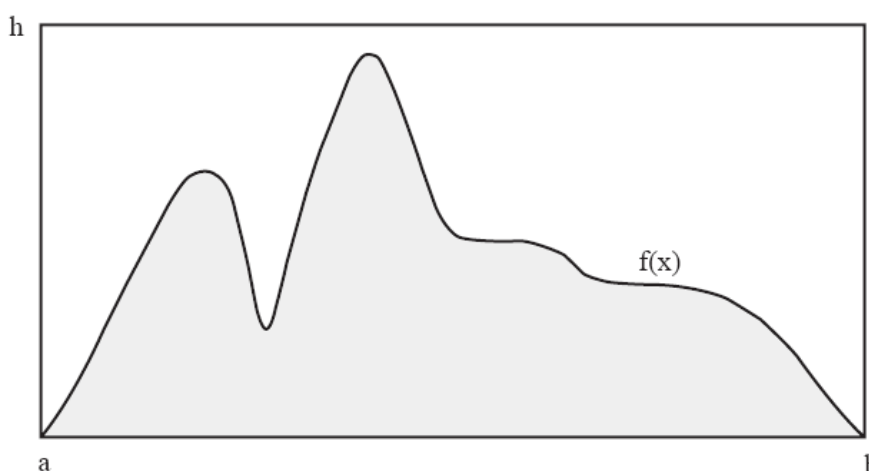
¹⁴Stanislav Ulam

۳-۱-۱ انتگرال‌گیری‌های مونت کارلو

روش‌های مونت کارلو به محاسبه انتگرال‌هایی می‌پردازد که محاسبه آنها به صورت تحلیلی امکان‌پذیر نیست. تمام روش‌های انتگرال‌گیری مونت کارلو بر پایه‌ی تولید اعداد کاتوره‌ای به محاسبه‌ی انتگرال می‌پردازد. در زیر به کوتاهی برخی از روش‌های انتگرال‌گیری را مرور می‌کنیم.

۳-۱-۱-۱ روش برخورد یا خطا

ساده‌ترین روش انتگرال‌گیری مونت کارلو روشی موسوم به روش برخورد یا خطا^{۱۵} است [۱۴]. شکل ۱-۱ را در نظر بگیرید.



شکل ۱) تابع $f(x)$ که در مستطیل با طول $(b-a)$ و عرض H قرار گرفته است. S ، سطح رنگ‌شده مساحت زیر نمودار و جواب انتگرال‌گیری از تابع $f(x)$ است.

در واقع حل انتگرال محاسبه مساحت زیر تابع $f(x)$ است. مستطیل با عرض H و طول $(b-a)$ طوری در نظر گرفته شده است که تابع داخل آن تعریف شده باشد. N زوج کاتوره‌ای (x_i, y_i) با شرط $a \leq x_i \leq b$ و $0 \leq y_i \leq H$ تولید می‌کنیم. کسری از نقاط (x_i, y_i) که $y_i \leq f(x)$ تخمینی از نسبت مساحت مستطیل به مساحت زیر تابع (انتگرال $f(x)$) می‌باشد. بنابراین مقدار تقریبی انتگرال را می‌توان به صورت زیر نوشت:

$$I = \int_a^b f(x) dx \cong S \frac{N'}{N} \quad (4)$$

که N' شمار نقاطی است که شرط $y_i \leq f(x)$ را برآورده می‌کند و S مساحت زیر نمودار را نشان می‌دهد. روش برخورد و خطا را به آسانی می‌توان به فضای چند بعدی گسترش داد و تخمینی از حجم یک جسم نیز به دست آورد.

۳-۱-۲ روش مقدار میانگین

با توجه به قضیه مقدار میانگین می‌توانیم انتگرال (۴) را به صورت زیر بنویسیم:

$$I_N = \frac{(b-a)}{N} \sum_{i=1}^N f(x_i) \quad (5)$$

که x_i ها اعداد کاتوره‌ای در بازه‌ی $[a, b]$ هستند. اگر شمار x_i ها، N ، به اندازه کافی بزرگ باشد، I_N به جواب دقیق انتگرال نزدیک‌تر می‌شود. این رابطه مشابه روش دوزنقه‌ای می‌باشد با این تفاوت که در روش دوزنقه‌ای x_i ها با فواصل یکسان گزیده می‌شوند. در روش‌های معمول انتگرال‌گیری با افزایش بعد انتگرال (d)، خطای محاسبات افزایش می‌یابد درحالی‌که در روش مونت کارلو خطا متناسب با $\frac{1}{\sqrt{d}}$ است [۱۵]. از این‌رو روش‌های انتگرال‌گیری مونت در محاسبه‌ی انتگرال‌های چندگانه نسبت به سایر روش‌های انتگرال‌گیری برتری قابل توجهی دارد.

۳-۱-۳ نمونه‌گیری با اهمیت

در بخش ***، خواهیم دید که خطای محاسبات مونت کارلو متناسب با σ (انحراف میانگین نمونه‌ها) و واریانس مجذور N (شمار نمونه‌گیری‌ها) می‌باشد. از این‌رو می‌توان با کاهش واریانس و یا افزایش شمار نقاط نمونه‌گیری خطای انتگرال‌گیری در روش مونت کارلو را کاهش داد. روش‌های نمونه‌گیری با اهمیت با کاهش σ ، کارایی محاسبات مونت کارلو را افزایش می‌دهند. در زیر این روش را توضیح می‌دهیم.

تابع مثبت $p(x)$ را در نظر بگیرید به طوری که

$$\int_a^b p(x) dx = 1 \quad (6)$$

می‌توان انتگرال ۴ را به صورت زیر بازنویسی کرد:

$$F = \int_a^b g(x) p(x) dx \quad (7)$$

که $g(x) = \frac{f(x)}{p(x)}$ است. انتگرال رابطه‌ی ۴، با N نمونه‌گیری از تابع توزیع $p(x)$ با محاسبه مجموع زیر به دست می‌آید:

$$F_n = \frac{1}{n} \sum_{i=1}^n g(x_i) \quad (8)$$

واضح است که با گزیدن $p(x) = \frac{1}{b-a}$ ، رابطه‌ی ۸، به رابطه‌ی ۵، تبدیل می‌شود. تابع $p(x)$ باید به گونه‌ای باشد که واریانس نسبت $[f(x)/p(x)]$ را کمینه کند. (بهترین گزینه برای تابع $p(x)$ می‌تواند به صورت $|f(x)/p(x)|$ باشد. اکنون نتیجه‌ی انتگرال‌گیری بسیار بهتر است، اگرچه میزان خطای مونت کارلو هنوز متناسب با $\frac{\sigma}{\sqrt{n}}$ می‌باشد.

۳-۲ الگوریتم متروپولیس

الگوریتم متروپولیس روشی برای تولید نقاط از یک تابع توزیع نایک‌نواخت است که توسط متروپولیس^{۱۶} و همکارانش در سال ۱۹۵۳ ارائه شد [۵]. در واقع الگوریتم متروپولیس یک فرایند نمونه‌گیری با اهمیت است که برای محاسبه‌ی میانگین‌هایی به شکل

$$\langle f \rangle = \frac{\int f(x) p(x) dx}{\int p(x) dx} \quad (9)$$

مفید است. در رابطه‌ی ۹، $p(x)$ تابعی است اختیاری، که لازم نیست بهنجار شده باشد. برای سادگی، الگوریتم متروپولیس را در حل یک انتگرال یک بعدی بررسی می‌کنیم. فرض کنیم که مایل هستیم از نمونه‌گیری با اهمیت برای تولید متغیرهای کاتوره‌ای با توزیع $p(x)$ استفاده کنیم. الگوریتم متروپولیس یک ولگشت کاتوره‌ای از نقاط $\{x_i\}$ با تابع توزیع احتمال $p(x)$ تولید می‌کند. یک ولگشت کاتوره‌ای با احتمال گذار $T(x_i \rightarrow x_j)$ از نقطه‌ی x_i به نقطه‌ی x_j طوری که توزیع نقاط $x_0, x_1, \dots$ به تابع $p(x)$ همگرا شود مشخص می‌شود. برای رسیدن به این امر تنها کافی است (بایسته نیست) که شرط تعادل جزئی^{۱۷} برآورده شود:

$$P(x_i)T(x_i \rightarrow x_j) = P(x_j)T(x_j \rightarrow x_i) \quad (10)$$

در رابطه‌ی ۱۰، $T(x_i \rightarrow x_j)$ یکتا نیست. یک گزینش ساده می‌تواند به صورت زیر باشد :

$$T(x_i \rightarrow x_j) = \min \left[1, \frac{P(x_j)}{P(x_i)} \right] \quad (11)$$

اگر در ولگشت در نقطه‌ی x_i باشیم، برای تولید نقطه‌ی x_{i+1} می‌توانیم با گزینش $T(x_i \rightarrow x_j)$ از رابطه‌ی (۱۱) مطابق الگوریتم زیر پیش برویم:

۱- گزینش موقعیت آزمایشی $x_{trial} = x_i + \delta_i$ که δ_i یک عدد کاتوره‌ای در بازه‌ی $[-\delta, \delta]$ است.

$$w = \frac{P(x_{trial})}{P(x_i)} \quad \text{۲- محاسبه‌ی}$$

۳- اگر $w \geq 1$ ، جابجایی اعمال شده را می‌پذیریم و $x_{i+1} = x_{trial}$ ، قرار می‌دهیم.

۴- اگر $w < 1$ ، عدد کاتوره‌ای r را تولید می‌کنیم.

۵- اگر $r \leq w$ ، دوباره این تغییر را می‌پذیریم و $x_{i+1} = x_{trial}$ ، قرار می‌دهیم.

۶- اگر در طی مراحل بالا تغییر آزمایشی پذیرفته نشد، آنگاه اجازه می‌دهیم تا $x_{i+1} = x_i$ باشد.

¹⁶Metropolis

¹⁷Detailed Balance

برای آنکه توزیع احتمال نقاط تولید شده به سوی $p(x)$ میل کند باید در دستگاه تعادل برقرار شود. بنابراین لازم است شمار زیادی نقطه کاتوره‌ای مطابق الگوریتم بالا تولید کنیم. اگر مقدار δ (گام مونت کارلو) در الگوریتم متروپلیس بسیار بزرگ باشد تنها درصد بسیار کمی از گام‌های آزمایشی پذیرفته می‌شود از این رو به نمونه‌گیری درستی از تابع $p(x)$ نمی‌رسیم. از طرف دیگر با گزینش δ بسیار کوچک درصد زیادی از گام‌های آزمایشی پذیرفته می‌شود اما این بار نیز نمونه‌گیری خوبی از تابع $p(x)$ صورت نگرفته است. یک معیار گزینش اندازه‌ی δ وقتی است که تقریباً یک‌سوم تا نصف گام‌های آزمایشی مورد پذیرش قرار گیرد. همچنین باید مقدار x_0 (نخستین نقطه آزمایشی) به گونه‌ای گزینش شود که توزیع $\{x_i\}$ تندتر به سوی توزیع مورد نظر میل کند. واضح است که بهترین گزینش برای x_0 ، این است که ولگشت در بیشینه مقدار تابع $p(x)$ باشد.

۴ روش‌های مونت کارلوی کوانتومی

معادله شرودینگر بس-الکترونی توصیف دقیقی از مواد را در سطح و تراز مکانیک کوانتومی در اختیار ما قرار می‌دهد، ولی یک معادله دیفرانسیل جزئی $3N+1$ بعدی است که تعداد الکترون‌ها، N ، می‌تواند خیلی بزرگ باشد. برای فائق آمدن بر مشکل ابعاد بالا، اکثر مطالعات مکانیک کوانتومی محاسباتی سیستم‌های بزرگ بر اساس نظریات ساده تک-الکترونی مانند نظریه تابعی چگالی کوهن-شام^{۱۸} و تقریب هارتری-فوک^{۱۹} بنا نهاده شده‌اند. در این نظریات برهمکنش‌های الکترون-الکترون به وسیله یک پتانسیل مؤثر جایگزین می‌گردد و بدین وسیله مسئله به یک دسته معادلات تک-الکترونی کاهش می‌یابد.

با وجود موفقیت‌های روش نظریه تابعی چگالی در تشریح ساختار الکترونی مولکول‌های پیچیده و جامدات، طرز رفتار همبستگی الکترونی در روش DFT تنها یک تقریب می‌باشد، که گاهی به نتایج نادرستی برای سیستم‌ها با همبستگی قوی و ضعیف منجر می‌شود. روش‌های مونت کارلو کوانتومی (QMC) در زمره موفق‌ترین روش‌های post-DFT است و محاسبات بسیار دقیقی را از ویژگی‌های همبسته برای مولکول‌های بزرگ یا سیستم‌های جامد ارائه می‌دهد، جایی که به کار بردن روش‌های شیمی کوانتومی متعارف به شدت مشکل می‌باشد. به خاطر همین دلایل، روش‌های مونت کارلو کوانتومی، خود را به عنوان ابزاری بی‌همتا برای کاوش همبستگی الکترونی در سیستم‌های مورد علاقه علوم مواد و برای دستیابی به جواب‌های قطعی در حالاتی که نظریه تابعی چگالی نابسند می‌باشد، پابرجا می‌کند.

مونت کارلو کوانتومی به عنوان روشی که از مشکلات سایر روش‌های تقریبی فارغ است از اوایل دهه ۱۹۸۰ معرفی شد. این روش اجازه می‌دهد تا با تابع موج بس-ذره ای کامل بطور مستقیم کار کنیم. با توجه به رهیافت‌های ریاضی و فیزیکی متفاوت چندین روش مونت کارلو کوانتومی وجود دارد که عبارتند از:

- ۱- مونت کارلو کوانتومی وردشی^{۲۰}
- ۲- مونت کارلو کوانتومی پخشی^{۲۱}
- ۳- مونت کارلو کوانتومی انتگرال مسیر^{۲۲}
- ۴- مونت کارلو کوانتومی تابع گرین^{۲۳}
- ۵- مونت کارلو کوانتومی میدان-کمکی^{۲۴}

¹⁸Kohn-Sham Density Functional Theory (DFT)

¹⁹ Hartree-Fock Approximation (HF)

²⁰Variational Quantum Monte Carlo (VQMC)

²¹Diffusion Quantum Monte Carlo (DQMC)

²²Path Integral Quantum Monte Carlo (PIQMC)

²³Green's Function Quantum Monte Carlo (GQMC)

²⁴Auxiliary-Field Quantum Monte Carlo (AFQMC)

مهمترین امتیاز روشهای مونت کارلو کوانتومی آن است که با آنها می‌توان مسائلی را بررسی کرد که توابع توزیع احتمالشان پیچیده است و با روش تحلیلی این توابع را نمی‌توان بدست آورد، مانند مسائل دو یا چند ذره‌ای. امتیاز دیگر این روش‌ها سادگی کاربرد آنها برای شرط مرزی‌های مختلف می‌باشد. از طرفی اشکال عمده این روش‌ها همانند سایر روشهای مونت کارلو وجود خطای آماری در محاسبات است که باید آن را تا حد ممکن کوچک کرد. اغلب این روشها هدفشان محاسبه انرژی و تابع موج حالت پایه سیستم است، به جز روش مونت کارلو کوانتومی انتگرال مسیر و مونت کارلو دمای محدود میدان کمکی که ماتریس چگالی را محاسبه می‌کنند، که ما در اینجا به این روش‌ها نمی‌پردازیم.

مونت کارلو کوانتومی وردشی کاربرد ساده‌ای از مونت کارلو مبتنی بر الگوریتم متروپولیس با این پیچیدگی که در آن ما عبارت تحلیلی تابع توزیع را نداریم تا نمونه‌گیری شود: توزیع برابر است با مربع تابع موجی که باید ساخته و بهینه شود تا یک تقریب خوب از تراز مورد نظر باشد. روش مونت کارلو کوانتومی پخشی یک راه ویژه از طرح ریزی حالت پایه از یک توزیع اولیه می‌باشد و به قدم تصادفی^{۲۵} داده شده بوسیله فرایندهای پخش^{۲۶} و شاخه‌زنی^{۲۷}، مربوط می‌شود.

آمار فرمی و شکست حل مربوط به حالت پایه سیستم بوزونی، برای مونت کارلو پخشی مشکل ایجاد می‌کند که در تقریب نقطه ثابت^{۲۸} از طریق قیودی که معرفی می‌شود، برطرف می‌گردد. پیش از آنکه به توضیح روش‌ها بپردازیم اجازه دهید که روش نوشتاری و نمادگذاری استفاده شده در این متن تعریف کنیم. ما فرض می‌کنیم که سیستمی غیر نسبیتی شامل N الکترون برهمکنشی داریم که به وسیله هامیلتونی زیر توصیف می‌شود:

$$H = T + V, \quad (12)$$

$$H = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N v_{ext}(\vec{r}_i) + \sum_{i < j}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i=1}^N \frac{1}{|\vec{r}_i - \vec{x}|}$$

$v_{ext}(\vec{r})$ می‌تواند یکپتانسیل خارجی ثابت و یا شبه پتانسیل باشد که در اینجا از آن صرف‌نظر می‌شود.

$3N$ مختصه ذرات با $\vec{R}$ مشخص می‌شود یعنی $\vec{R} = (\vec{r}, \vec{x})$.

در این نوشته از واحدهای اتمی استفاده شده است ($\hbar = m = e = 1$). واحد اتمی طول، شعاع

بوهر^{۲۹}: $\frac{\hbar^2}{me^2} = 0.529 \text{ \AA}$ ؛ واحد زمان: $\frac{\hbar^3}{me^4} = 2.419 \times 10^{-2} \text{ fs}$ ؛ و واحد انرژی (برای انرژی

حالت پایه و پتانسیل‌ها) هارتری^{۳۰} است یعنی: $\frac{me^4}{\hbar^2} = 27.25 \text{ eV}$.

²⁵Random Walk

²⁶Diffusion Process

²⁷Branching Process

²⁸Fixed-Node Algorithm

²⁹Bohr Radius

³⁰Hartree

۵ مونت کارلوی کوانتومی وردشی (VMC)

۱-۵ مقدمه

مونت کارلوی کوانتومی وردشی^{۳۱} (VMC)، یکی از مهم‌ترین کاربردهای مستقیم انتگرال‌گیری‌های مونت کارلو است. روش VMC بر پایه‌ی اصل وردش در مکانیک کوانتومی کران بالایی برای انرژی حالت پایه‌ی دستگاه‌های بس‌ذره‌ای می‌یابد. در این فصل نخست اصل وردش خطی در مکانیک کوانتومی را بیان می‌کنیم. در بخش‌های پسین به مطالعه‌ی هامیلتونی دستگاه‌های بس‌ذره‌ای و ساختار الگوریتم VMC می‌پردازیم.

۲-۵ اصل وردش

برای به دست آوردن اصل وردش ریلی-ریتز^{۳۲} در مکانیک کوانتومی، تابع موج آزمایشی بهنجار ψ_T را بر حسب ویژه‌حالات یک هامیلتونی به شکل زیر بسط می‌دهیم:

$$\psi_T = \sum_{i=0}^{\infty} c_i \psi_i \quad (13)$$

که c_i ها، ضرایب بسط هستند و به یک بهنجار شده‌اند:

$$\sum_{i=0}^{\infty} |c_i|^2 = 1 \quad (14)$$

مقدار چشم‌داشتی هامیلتونی یک دستگاه بس‌ذره‌ای، $\hat{H}$ ، برای تابع موج آزمایشی، ψ_T ، به صورت زیر به دست می‌آید:

$$\begin{aligned} \langle \psi_T | \hat{H} | \psi_T \rangle &= \left\langle \sum_i c_i \psi_i \left| \hat{H} \right| \sum_j c_j \psi_j \right\rangle \\ &= \sum_i \sum_j c_i^* c_j \langle \psi_i | \hat{H} | \psi_j \rangle \\ &= \sum_i |c_i|^2 \varepsilon_i \end{aligned} \quad (15)$$

که

$$\varepsilon_i = \langle \psi_i | \hat{H} | \psi_i \rangle \quad (16)$$

بنابراین مقدار چشم‌داشتی $\hat{H}$ ، برای یک تابع موج آزمایشی همواره بزرگتر یا برابر انرژی حالت پایه‌ی دستگاه است. شکل تابع موج آزمایشی در محاسبات وردشی اهمیت بسیاری دارد. با گزینش توابع موجی که ویژگی‌های تابع موج حالت پایه‌ی دستگاه را داشته باشند، رسیدن به انرژی حالت پایه‌ی دستگاه دور از دسترس نیست.

³¹Variational Quantum Monte Carlo (VMC)

³²Raleigh-Ritz

۵-۳ مونت کارلوی کوانتومی وردشی

مونت کارلوی کوانتومی وردشی (VMC) ساده‌ترین روش مونت کارلو کوانتومی است و تعمیمی از مونت کارلوی کلاسیکی ارائه می‌دهد. VMC برای مقایسه مقادیر انتظاری عملگرهای کوانتوم مکانیکی روی تابع موج آزمایشی داده شده به کار برده می‌شود و استفاده از هر تابع موج آزمایشی قابل محاسبه را، بدون اعمال قیود سختگیرانه روی ترکیب تابعی آن، مجاز می‌داند. این مسئله باید با روش‌های دیگر سنتی شیمیایی کوانتومی که تابع موج را بر حسب حاصل ضرب‌های مدارهای تک ذره برای انجام انتگرال‌های با ابعاد بالا در مقادیر انتظاری مختلف بیان می‌کنند، مغایرت داشته باشد. VMC اولین بار به وسیله مک میلان^{۳۳} [48] برای محاسبه ویژگی‌های حالت پایه 4He مایع استفاده شد، و بعد از آن توسط سپرلیو دیگران^{۳۴} برای سیستم‌های فرمیونی تعمیم داده شد [49]. مونت کارلو کوانتومی وردشی (VMC) از نام نظریه وردشی مشتق می‌شود، که این نظریه اظهار می‌دارد که برای یک تابع موج آزمایشی داده شده $\Psi(\vec{R})$ ، مقدار انتظاری هامیلتونی نسبت به این تابع موج آزمایشی حد بالایی برای انرژی حالت پایه دقیق E_0 می‌باشد:

$$E_V = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\int d\vec{R} \Psi^*(\vec{R}) H \Psi(\vec{R})}{\int d\vec{R} \Psi^*(\vec{R}) \Psi(\vec{R})} \geq E_0 \quad (17)$$

انرژی وردشی می‌تواند به صورت مقدار متوسط انرژی موضعی $E_L(\vec{R})$ روی توزیع $\rho(\vec{R})$ نوشته شود:

$$E_V = \int d\vec{R} \frac{H \Psi(\vec{R})}{\Psi(\vec{R})} \frac{|\Psi(\vec{R})|^2}{\int d\vec{R} |\Psi(\vec{R})|^2} = \int d\vec{R} E_L(\vec{R}) \rho(\vec{R}) = \langle E_L(\vec{R}) \rangle_\rho \quad (18)$$

که انرژی موضعی به صورت زیر تعریف می‌شود:

$$E_L(\vec{R}) = \frac{H \Psi(\vec{R})}{\Psi(\vec{R})}, \quad (19)$$

و توزیع $\rho(\vec{R})$ به صورت

$$\rho(\vec{R}) = \frac{|\Psi(\vec{R})|^2}{\int d\vec{R} |\Psi(\vec{R})|^2} \quad (20)$$

در واقع معادله ۱۸ مانند یک میانگین‌گیری وزنی با وزن چگالی احتمال بهنجار شده فوق می‌باشد.

در VMC، برای نمونه‌گرفتن از توزیع $\rho(\vec{R})$ از الگوریتم متروپولیس استفاده می‌شود و متوسط انرژی موضعی را روی M پیکربندی مستقل نمونه گرفته شده با احتمال $\rho(\vec{R})$ به صورت زیر حساب می‌شود:

³³McMillan

³⁴Ceperley et al.

$$E_V = \langle E_L(\vec{R}) \rangle_\rho \approx \frac{1}{M} \sum_{i=1}^M E_L(r_i), \quad (21)$$

که میانگین، یک تقریب است برای مقدار انتظاری و در حدود $M \rightarrow \infty$ به تساوی تبدیل می‌شود. بنابراین تخمین E_V توسط مونت کارلو با جمع‌بندی زیر تعیین می‌شود:

$$E_V = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M E_L(r_i), \quad (22)$$

M تعداد دفعاتی است که $\vec{R}$ از توزیع $\Psi^2(\vec{R})$ نمونه‌گیری می‌شود.

حال ببینیم که Ψ باید چه شرایطی را ارضاء کند که در روش مونت کارلو وردشی مورد استفاده قرار گیرد. تابع موج باید پیوسته، قابل بهنجارش باشد و تقارن مناسبی داشته باشد. به طور بسیار مهمی، برای اینکه خطای آماری روی انرژی E_V معنادار باشد، Ψ باید واریانس متناهی از انرژی σ_V^2 را عرضه کند،

$$\sigma_V^2 = \frac{\langle \Psi | (H - E_V)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle (E_L(\vec{R}) - E_V)^2 \rangle_\rho, \quad (23)$$

از آنجایی که خطا به عنوان تابعی از قدم‌های مونت کارلو به صورت زیر رفتار می‌کند:

$$err(E_V) \sim \frac{\sigma_V}{\sqrt{M}} \quad (24)$$

توجه داشته باشید که، هنگامی که تابع موج آزمایشی به ویژه حالت دقیق هامیلتونی H نزدیک شود، آنگاه انرژی موضعی ثابت و برابر با ویژه مقدار هامیلتونی روی آن تراز مورد نظر می‌شود. بنابراین در این حد، مقدار انتظاری انرژی واقعی می‌شود و واریانس به صفر میل می‌کند. این ویژگی مهم مونت کارلو کوانتومی وردشی به عنوان اصل واریانس صفر^{۳۵} مشهور است: همچنانکه تابع موج آزمایشی بهبود پیدا می‌کند، تخمین مونت کارلو از انرژی وردشی به طور سریع تری به صورت تابعی از تعداد پیکربندی‌های مونت کارلو (در حدی که Ψ ویژه‌حالت دقیق باشد، تنها یک پیکربندی در معادله ۲۲ برای محاسبه کردن انرژی لازم است) همگرا می‌شود. به خاطر ویژگی واریانس صفر، محاسبات مونت کارلو کوانتومی انرژی می‌تواند بسیار مؤثرتر و کارآمدتر از محاسبات مونت کارلو کلاسیکی باشد. هرچند که ما روی انرژی و محاسبه آن تمرکز کردیم، مونت کارلو کوانتومی وردشی می‌تواند برای محاسبه مقدار انتظاری هر عملگر χ که در نمایش $\vec{R}$ قطری است، مورد استفاده قرار گیرد. در عبارات قبلی، به سادگی هامیلتونی H را با χ و انرژی موضعی را با $X(\vec{R}) = \chi \Psi(\vec{R}) / \Psi(\vec{R})$ تعویض می‌کنیم.

نمونه‌ای از الگوریتم مونت کارلو کوانتومی وردشی می‌تواند به صورت زیر باشد:

۱- تابع موج آزمایشی دلخواه Ψ_T را که به یک پارامتر وردشی مثل λ وابسته باشد، انتخاب می‌کنیم

۲- یک جابجایی آزمایشی به صورت $r_{trial} = r_n + \delta_n$ انجام می‌دهیم، δ_n یک عدد تصادفی در بازه $[-\delta, \delta]$ است.

³⁵Zero Variance Principle

۳- نسبت $w = \frac{P(r_{trial})}{P(r_n)}$ را محاسبه می‌کنیم، که در آن $P(r) = \Psi^2(r)$ ، چگالی احتمال بهنجار شده است.

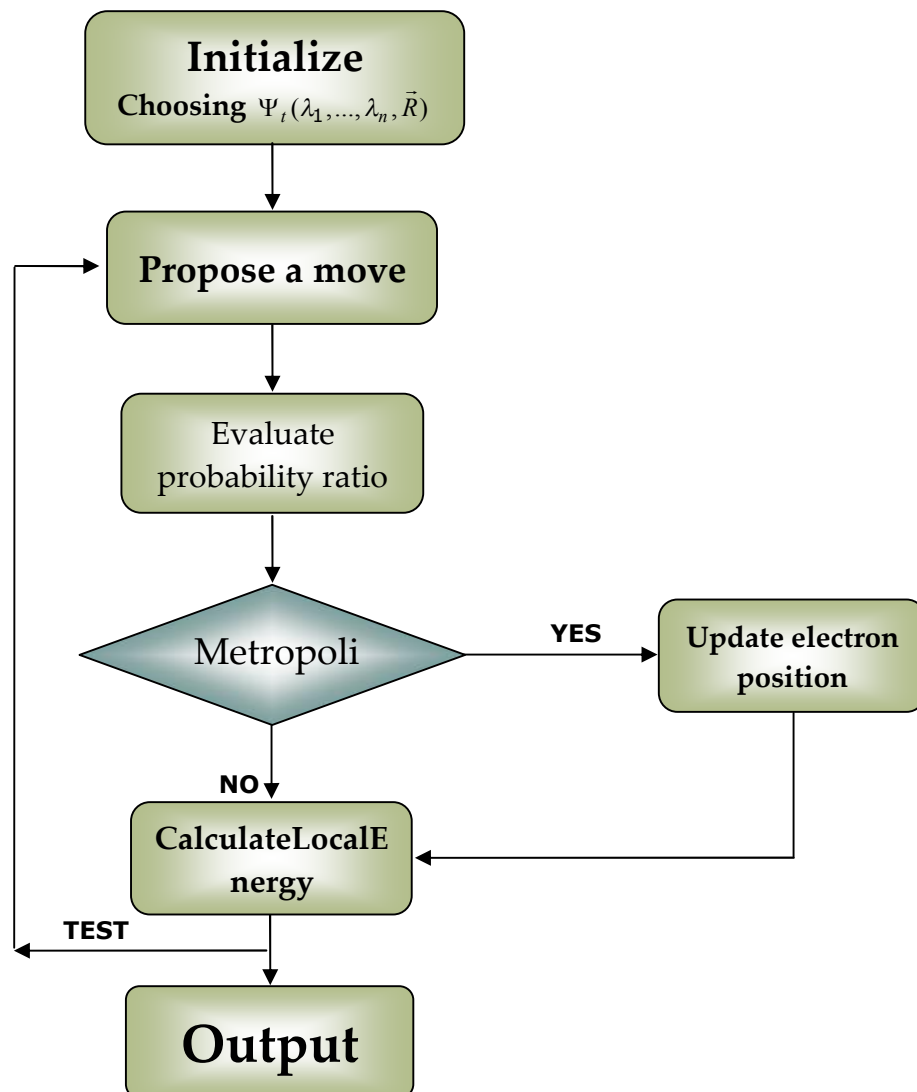
۴- اگر که $w \geq 1$ بود آنگاه جابجایی را می‌پذیریم و $r_{trial} = r_{n+1}$ در غیر این صورت، اگر $w < 1$ بود، یک عدد تصادفی مانند s را با w مقایسه می‌کنیم و در صورتیکه $s \leq w$ بود آنگاه r جدید را قبول می‌کنیم.

۵- در صورتیکه جابجایی قابل قبول نبود $r_{n+1} = r_n$.

۶- مقدار انرژی را با استفاده از r_n ها و روابط مربوط به E_L و E_V محاسبه می‌کنیم.

۷- مراحل ۲ تا ۶ را برای مقادیر مختلف پارامتر وردشی λ تکرار می‌کنیم تا $E[\Psi]$ یا همان E_V کمینه بدست آید.

در زیر فلوچارت روش مونت کارلو کوانتومی وردشی نمایش داده شده است:



شکل ۲- فلوچارت مربوط به روش مونت کارلو وردشی

۵-۳-۱ انرژی موضعی

انرژی موضعی، $E_L(\vec{R}) = \frac{H\Psi(\vec{R})}{\Psi(\vec{R})}$ ، یکی از کمیت‌های مهم در روش‌های مونت کارلوی کوانتومی است و ویژگی‌های این کمیت به بهینه‌سازی تابع موج آزمایشی کمک می‌کند. با توجه به معادله‌ی ۱۹، بدیهی است برای یک ویژه حالت هامیلتونی مقدار انرژی موضعی ثابت و برابر ویژه‌مقدار مربوطه است. اما برای یک تابع موج آزمایشی کلی انرژی موضعی ثابت نیست و واریانس آن بستگی به گزینش تابع موج آزمایشی دارد که تا چه حد یک ویژه حالت را تقریب می‌زند. از این‌رو میانگین‌گیری روی واریانس انرژی موضعی روشی سودمند در بهینه‌سازی تابع موج آزمایشی است و به نام روش کمینه‌سازی واریانس که در فصل چهارم، به آن پرداخته می‌شود، شناخته شده است.

۶ روش مونت کارلوی کوانتومی پخشی

مونت کارلو کوانتومی وردشی به‌خاطر مجموعه‌ای از دلایل روش قدرتمندی است، می‌توانیم از هر تابع موجی استفاده کنیم البته تا وقتی که در یک مقدار منطقی از زمان کامپیوتر قابل محاسبه باشد. و بالاخره اینکه در این روش هیچ مشکل علامتی وابسته به آمار فرمی وجود ندارد.

اشکالاتی هم در مورد VMC وجود دارد: نخست اینکه، همان‌طور که قبلاً ذکر شد، هیچ روش اتوماتیکی برای بهبود تابع موج در روند محاسبات وجود ندارد و انتخاب‌ها باید پیرامون فرم تابعی آن برای هر مسئله ویژه، انجام شوند. علاوه بر این، خواص دیگر به جز انرژی به طور قابل توجهی کم دقتتر هستند، چون خطای آنها روی تابع موج از مرتبه اول هستند، به جای اینکه همانند انرژی از مرتبه دوم باشند. به طور خلاصه، می‌توان گفت که هرآنچه برای تابع موج روی دهد، مستقیماً در کیفیت نتایج منعکس می‌شود.

مونت کارلو کوانتومی پخشی قدرتمندتر از مونت کارلو کوانتومی وردشی با پرداخت هزینه‌های تکنیکی بیشتر می‌باشد و تا حد زیادی تحت تاثیر قرار گرفتن نتایج از طرف تابع موج آزمایشی را از بین می‌برد.

روش مونت کارلو کوانتومی پخشی (DMC)، بر پایه بازنویسی معادله شرودینگر با زمان مجازی و انجام ول گشت در فضای فاز $3N$ بعدی است. این روش اولین بار توسط متروپولیس و یولام حین دنبال کردن پیشنهاد فرمی مبنی بر شباهت معادله شرودینگر وقتی که در زمان مجازی نوشته شود با یک معادله پخش و خلق/فنا، ارائه شد [46]. "جی بی اندرسون"^{۳۶} در سال ۱۹۷۵ روش مونت کارلو کوانتومی پخشی را برای سیستم الکترونی به کار برد [50]. اگر معادله شرودینگر وابسته به زمان را در نظر بگیریم:

$$i\hbar \frac{\partial \Psi(\vec{R}, t)}{\partial t} = \hat{H} \Psi(\vec{R}, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\vec{R}, t) + V(\vec{R}) \Psi(\vec{R}, t) \quad (25)$$

با تغییر متغیر $\tau = it$ و با استفاده از یکای اتمی این معادله تبدیل به معادله شرودینگر زمان مجازی می‌شود:

$$-\frac{\partial \Psi(\vec{R}, \tau)}{\partial \tau} = (\hat{H} - E_T) \Psi(\vec{R}, \tau) = -\frac{1}{2} \nabla^2 \Psi(\vec{R}, \tau) + (V(\vec{R}) - E_T) \Psi(\vec{R}, \tau) \quad (26)$$

E_T یک ثابت انرژی است که فقط مرجع انرژی را تغییر می‌دهد و تاثیری بر اصل مساله ندارد. میتوان معادله فوق را به صورت زیر نیز نوشت:

$$\frac{\partial \Psi(\vec{R}, \tau)}{\partial \tau} = D \nabla^2 \Psi(\vec{R}, \tau) - (V(\vec{R}) - E_T) \Psi(\vec{R}, \tau) \quad (27)$$

$$\frac{\partial \rho(\vec{R}, \tau)}{\partial \tau} = D \nabla^2 \rho(\vec{R}, \tau) - S(\vec{R}) \rho(\vec{R}, \tau) \quad (28)$$

معادله پخش^{۳۷} متداول با شاخه‌زنی برای یک سیستم از ذرات با چگالی $\rho(\vec{R}, t)$ در زیر معادله شرودینگر نوشته شده (معادله ۲۸) تا قیاس بین این دو معادله روشن تر انجام گیرد. در معادله شرودینگر جابجایی تنظیم‌پذیر، E_T ، برای راحتی در صفر انرژی معرفی شده است.

در معادله ۲۷، $D \equiv \hbar^2 / 2m = 1/2$ ضریب پخش این فرآیند محسوب می‌شود. ضریب پخش در هر مساله به میانگین مسافت پویش آزاد ذرات بستگی دارد. در این معادله که بیان کننده تحول زمانی و مکانی تابع موج Ψ می‌باشد، دو فرآیند پخش و خلق/فنا، یعنی به ترتیب جمله اول و جمله دوم در سمت راست معادله، وارد شده‌اند. در واقع این معادله همانند معادله پخش معمولی در فضای $3N$ بعدی با ثابت پخش $D \equiv 1/2$ و با پتانسیل می باشد، اگر از جمله $[V(\vec{R}) - E_T]$ در معادله شرودینگر صرف‌نظر شود، این معادله همان معادله سنتی پخش با ثابت پخش $D \equiv 1/2$ خواهد بود که به آسانی با انجام گام‌زن تصادفی در فضای فاز $3N$ بعدی قابل شبیه‌سازی و حل می‌باشد. بنابراین تابع موج بس - ذره ای $\Psi(\vec{R}, \tau)$ معادل چگالی ذرات در معادله پخش است. لازم به ذکر است که این ذرات واقعی نمی‌باشند بلکه ذراتی خیالی هستند که در فضای فاز $3N$ بعدی پخش و تحت شرایطی که در ادامه بیان خواهند شد تولید یا نابود می شوند تا شبیه سازی انجام شود. تعداد این گام‌زن‌ها در فضای فاز $3N$ بعدی در هر نقطه $\vec{R}$ بسته به چگالی گام‌زن‌ها و مقدار $[V(\vec{R}) - E_T]$ در آن نقطه تغییر می کند. با در نظر گرفتن جمله $[V(\vec{R}) - E_T]$ در سمت راست معادله شرودینگر این معادله به معادله نرخ تبدیل می‌شود که فرآیند شاخه‌زنی همچون واپاشی رادیو اکتیو، یا فرآیند نمایی خلق و فنا در جمعیت را تشریح می‌کند. چنین سیستمی نیز به آسانی به وسیله سیستمی از گام‌زن‌ها قابل شبیه‌سازی است. بنابراین معادله شرودینگر زمان مجازی به وسیله استفاده از ترکیبی از این فرآیندهای پخش و خلق/فنا که روی یک آنسامبلی از گام‌زن‌ها عمل می‌کند قابل شبیه سازی است. حل قراردادی معادله شرودینگر در زمان موهومی به صورت زیر می‌تواند نوشته شود:

$$\Psi(\vec{R}, \tau) = \exp(-\tau(\hat{H} - E_T))\Psi(\vec{R}, \tau). \quad (30)$$

که همانطور که قبلاً هم گفته شده، $\vec{R}(r_1, r_2, \dots, r_N)$ یک بردار $3N$ -بعدی است مختصات همه N الکترون را مشخص می‌کند. $\Psi(\vec{R}, \tau)$ می‌تواند به صورت بسط بر حسب ویژه‌توابع Φ_i هامیلتونیچند-جسمی $\hat{H}$ ، نوشت، که این ویژه توابع مطابق ویژه‌مقادیر مربوطه به دست آمده‌اند، به طوریکه $E_i \leq E_{i+1}$ ،

$$\Psi(\vec{R}, \tau) = \sum_{i=0}^{\infty} c_i \exp(-\tau(E_i - E_T))\Phi_i(\vec{R}). \quad (31)$$

همانطور که در بالا هم ذکر شد E_i ویژه مقدار انرژی متناسب با ویژه حالت $\Phi_i(\vec{R})$ می‌باشد، ضریب c_i وابسته به شرایط اولیه هستند.

بازاء زمانهای به اندازه کافی طولانی، یعنی چنانکه $\tau \rightarrow \infty$ ، عملگر $\exp(-\tau(E_i - E_T))$ پایین‌ترین ویژه‌حالت یعنی Φ_0 را ارائه می‌دهد، و تنها جمله با انرژی حداقل باقی می ماند، زیرا سایر

³⁷Diffusion Equation

جملات با انرژی‌های بالاتر همانطور که دیده می‌شود به صورت نمایی محو می‌شوند. بنابراین اگر $c_0 \neq 0$ باشد، آنگاه داریم :

$$\lim_{\tau \rightarrow \infty} \Psi(\vec{R}, \tau) = c_0 \exp(-\tau(E_0 - E_T)) \Phi_0(\vec{R}). \quad (32)$$

در اینجا نکات زیر قابل ذکر می‌باشند:

- i. اگر $E_0 = E_T$ (انرژی حالت پایه می‌باشد)، آنگاه برای زمان‌های به اندازه کافی بزرگ تمام Φ_i ها به جز Φ_0 ، صفر خواهند شد. $\lim_{\tau \rightarrow \infty} \Psi(\vec{R}, \tau) \cong \Phi_0(\vec{R})$.
 - ii. اگر $E_0 < E_T$ ، وزن لحظه‌ای کلی، که به صورت $I(\tau) = \int \Phi(\vec{R}, \tau) d\vec{R}$ تعریف می‌شود، به صورت نمایی با زمان بزرگ می‌شود.
 - iii. اگر $E_0 > E_T$ باشد، آنگاه $I(\tau) = \int \Phi(\vec{R}, \tau) d\vec{R}$ به صورت نمایی با زمان کاهش می‌یابد.
- با حل معادله ۲۷ به ازاء مقادیر مختلف E_T و پیدا کردن یک $\Psi(\vec{R}, \tau)$ که به ازاء آن $I(\tau)$ ثابت باشد، می‌توانیم مطمئن باشیم که $\Psi = \Psi_0$ و $E_0 = E_T$ است (ویژه مقدار متناظر با $\Psi(\vec{R}, \tau)$ می‌باشد).

برای تحلیل معادله ۲۷ معادله شرودینگر زمان مجازی را به صورت انتگرالی زیر می‌نویسیم:

$$\Psi(\vec{R}', \tau) = \int d\vec{R} G(\vec{R}', \vec{R}, \tau) \Psi(\vec{R}, 0), \quad (33)$$

که در آن تابع گرین به صورت زیر تعریف می‌شود:

$$G(\vec{R}', \vec{R}, \tau) = \langle \vec{R}' | \exp(-\tau(\hat{H} - E_T)) | \vec{R} \rangle \quad (34)$$

اگر ما بتوانیم تابع موج آزمایشی و تابع گرین را در معادله ۳۳ نمونه‌گیری کنیم، می‌توانیم این انتگرال با ابعاد بالا را به وسیله انتگرال‌گیری مونت کارلو انجام دهیم. برای فرمیون‌ها، از آنجایی که تابع موج باید نامتقارن باشد، نمی‌تواند به عنوان توزیع احتمال تفسیر شود. بنابراین، فعلاً فرض می‌کنیم که با سیستم بوزونی سروکار داریم که با تابع موج حالت پایه مثبت مشخص می‌شود، که برای سیستم‌های شامل حداکثر دو الکترون این فرض مشکلی ایجاد نمی‌کند زیرا اثرات اسپین قابل صرف‌نظر می‌باشد و تابع موج حالت پایه مثبت و متقارن می‌باشد.

از معادله ۳۳ می‌بینیم که $G(\vec{R}', \vec{R}, \tau)$ تابع موج سیستم Ψ را از زمان صفر به زمان τ منتشر می‌کند. به آسانی می‌توان نشان داد که $G(\vec{R}', \vec{R}, \tau) = \delta(\vec{R}' - \vec{R})$ و همچنین با گرفتن مشتق نسبت به زمان مجازی $(\partial/\partial\tau)$ و سپس با عمل دادن $(\hat{H} - E_T)$ روی دو طرف معادله انتگرالی می‌توان تحقیق کرد که تابع گرین $G(\vec{R}', \vec{R}, \tau)$ در معادله زیر صدق می‌کند

$$-\frac{\partial G(\vec{R}', \vec{R}, t)}{\partial \tau} = (\hat{H} - E_T) G(\vec{R}', \vec{R}, t). \quad (35)$$

که همانند معادله شرودینگر زمان مجازی می‌باشد. یادآور می‌شود که معادله شرودینگر زمان مجازی به وسیله جایگزین کردن $\tau = it$ در معادله قراردادی و رسمی شرودینگر بدست می‌آید. اگر هامیلتونی فقط شامل عملگر انرژی جنبشی باشد، معادله شرودینگر در زمان مجازی، همان معادله پخش در یک فضای $3N$ -بعدی است:

$$-\frac{\partial G_{diff}(\vec{R}', \vec{R}, \tau)}{\partial \tau} = -\frac{1}{2} \nabla^2 G_{diff}(\vec{R}', \vec{R}, \tau) \quad (36)$$

و تابع گرین به وسیله توزیع گاوسی داده می‌شود:

$$G_{diff}(\vec{R}', \vec{R}, \delta\tau) = (2\pi\delta\tau)^{3N/2} \exp\left[-\frac{(\vec{R}' - \vec{R})^2}{2\delta\tau}\right], \quad (37)$$

بنابراین می‌توان تابع گرین را به عنوان یک احتمال گذار تفسیر کرد و به‌سادگی نیز نمونه‌گیری می‌شود. از طرف دیگر، اگر هامیلتونی تنها شامل عملگر پتانسیل باشد، معادله شرودینگر همانند معادله نرخ می‌شود (همانطور که قبلاً نیز اشاره شده بود)

$$-\frac{\partial G_{rate}(\vec{R}', \vec{R}, \tau)}{\partial \tau} = (V(\vec{R}) - E_T) G_{rate}(\vec{R}', \vec{R}, \tau) \quad (38)$$

و تابع گرین بوسیله معادله زیر داده می‌شود:

$$G_{rate}(\vec{R}', \vec{R}, \delta\tau) = \exp\left[-\delta\tau \left(\frac{V(\vec{R}') - V(\vec{R})}{2} - E_T\right)\right] \delta(\vec{R} - \vec{R}'), \quad (39)$$

که باز هم مثبت است اما شرط بهنجارش را ارضا نمی‌کند، و عاملی را نمایش می‌دهد که در هر قدم، تابع توزیع لحظه‌ای $\Psi(\vec{R}, t)$ را در آن ضرب می‌کنیم. بنابراین می‌توانیم انتظار داشته باشیم که حل برای هامیلتونی کلی شامل هر دو جمله جنبشی و پتانسیل، ترکیبی از فرآیندهای پخش و شاخه‌زنی (خلق / فنا) را توصیف می‌کند. صریحاً، نظریه تروتر-سوزوکی^{۳۸} می‌گوید که در حد گام‌های زمانی کوچک، اجازه داریم که سهم انرژی پتانسیل و جنبشی را به طور جداگانه در نظر بگیریم، از آنجا که برای دو عملگر داده شده **A** و **B**، داریم

$$e^{(A+B)\tau} = e^{A\tau} e^{B\tau} + O(\tau^2). \quad (40)$$

بنابراین، می‌توان تابع گرین در گام‌های زمانی کوچک را به صورت زیر تقریب زد:

$$(41)$$

$$G(\vec{R}', \vec{R}, \delta\tau) = \langle \vec{R}' | e^{-\delta\tau \hat{H}} | \vec{R}_0 \rangle \approx \langle \vec{R}' | e^{-\delta\tau \hat{T}} e^{-\delta\tau \hat{V}} | \vec{R}_0 \rangle = \int d\vec{R}'' \langle \vec{R}' | e^{-\delta\tau \hat{T}} | \vec{R}'' \rangle \langle \vec{R}'' | e^{-\delta\tau \hat{V}} | \vec{R}_0 \rangle$$

$$= e^{-\delta\tau V(\vec{R}')} \langle \vec{R}' | e^{-\delta\tau \hat{T}} | \vec{R}'' \rangle e^{-\delta\tau V(\vec{R}_0)}, \quad (2.42)$$

معادله ۴۱ استفاده شده است، که با معادله ۳۸ ترکیب شده و تابع گرین در تقریب زمان کوچک^{۳۹} رینولد^{۴۰} در سال ۱۹۸۲ [24] داده می‌شود به صورت:

³⁸Trotter – Suzuki

³⁹Short-Time Approximation

⁴⁰Reynold et al. 1982

(۴۲)

$$G(\vec{R}', \vec{R}, \delta\tau) = (2\pi \delta\tau)^{3N/2} \exp\left[-\frac{(\vec{R}' - \vec{R})^2}{2\delta\tau}\right] \exp[-\delta\tau (V(\vec{R}) + V(\vec{R}') - 2E_T)/2] + O(\delta\tau^3). \quad (2.43)$$

عملگرهای $\hat{T}$ و $\hat{V}$ با توجه به معادله شرودینگر زمان مجازی برابرند با:

$$\hat{T} = -\frac{1}{2}\nabla^2, \quad (43)$$

$$\hat{V} = V(\vec{R}) - E_T$$

تکرار در معادله ۳۳ می‌تواند به صورت فرآیند مارکوف^{۴۱} تعبیر گردد با این تفاوت که تابع گرین بهنجار نمی‌باشد و قدم تصادفی شاخه‌زنی را بدست می‌آوریم: عامل نخست در رابطه تابع گرین زمان کوچک، تابع گرین برای پخش است، در حالی که جمله دوم تابع توزیع را در یک عدد مثبت ضرب می‌کند. از آنجایی که بیان زمان کوچک تابع گرین تنها در حدی معتبر می‌باشد که $\delta\tau$ به صفر نزدیک شود، محاسبات مونت کارلو کوانتومی پخشی باید برای مقادیر مختلف $\delta\tau$ انجام شود و نتیجه برای $\delta\tau$ ای که به صفر میل می‌کند، برون‌یابی شود.

حال همه اطلاعات لازم برای حل معادله شرودینگر زمان مجازی فراهم است. معادله ۳۹، نرخ شاخه زنی گام زنها را در هر مکانی که می‌روند، می‌دهد و معادله ۳۷، نحوه گام زدن آنها را معین می‌نماید. از معادله ۴۲ می‌توان این‌طور نتیجه گرفت که در هر گام زمانی، گام زنها با گام‌های تصادفی که از یک تابع وزن مانند ۳۹ پیروی می‌کنند، حرکت داده می‌شوند.

$$p = \exp[-\delta\tau (V(\vec{R}) + V(\vec{R}') - 2E_T)/2] \quad (44)$$

این فاکتور به مثابه بازبهنجارش وابسته به زمان (توزین مجدد) تابع گرین پخشی عمل می‌کند. این تغییر بهنجارش می‌تواند از طرق مختلفی وارد فرآیند سیر تکامل گام‌زن‌ها گردد. یک امکان این است که به هر گام‌زن یک وزن تخصیص دهیم و در حین انتشار حاصلضرب وزن‌ها را جمع کنیم. البته این گزینه کارآمدترین نمونه‌گیری را نمی‌دهد زیرا وزن گام‌زن‌ها به سرعت تغییر می‌کند و در حد زمان طولانی، یکی به صورت نمایی بر بقیه چیره می‌شود. راه بهتر استفاده از الگوریتم شاخه‌زنی یا خلق/فنا می‌باشد، که در آن p احتمال زنده ماندن در مرحله بعد را نشان می‌دهد و تعداد گام‌زن‌هایی که برای مرحله بعدی زنده می‌مانند را معین می‌کند [24].

در ساده‌ترین روش فرآیند به صورت زیر می‌باشد:

۱- اگر $p < 1$ گام‌زن با احتمال p سیر تکاملی خود را ادامه می‌دهد.

۲- اگر $p \geq 1$ گام‌زن سیر تکاملی خود را ادامه می‌دهد؛ علاوه بر این در همان مکان، با احتمال

$p-1$ یک گام‌زن جدید تولید می‌شود.

هر دوی این امکانات به راحتی می‌توانند به صورت یک دستور بیان شوند:

$$M_{new} = INT(p + \eta), \quad (45)$$

⁴¹Markov Process

که M_{new} تعداد گامزن‌هایی است که در مکان داده شده به مرحله بعدی می‌روند، INT قسمت صحیح یک عدد حقیقی را مشخص می‌کند، و η عددی تصادفی با توزیع یکنواخت در بازه $(0, 1)$ است. از عبارت داده شده برای p روشن است که در نواحی با انرژی پتانسیل بالا گامزن‌ها نابود می‌شوند، در حالیکه در نواحی با انرژی پتانسیل پایین تکثیر می‌شوند. بنابراین الگوریتم شاخه‌زنی، تراکم و فزونی وزن در نواحی انرژی پایین را به افزایش در چگالی گامزن‌ها در آن ناحیه تبدیل می‌کند. استفاده از جابجایی انرژی E_T در احتمال شاخه زنی کاملاً موجه است، چون E_T ویژه توابع $|\Psi_i\rangle$ هامیلتونی کل را تغییر نمی‌دهد. E_T برای کنترل جمعیت نهایی گامزن‌ها مورد استفاده قرار می‌گیرد. در حین انتشار E_T گهگاه تنظیم می‌شود بنابراین تعداد کلی گامزن‌ها حول مقدار متوسط مطلوب نوسان می‌کند.

اساس الگوریتم مونت کارلوی پخش معرفی شده در بالا نسبتاً ساده است و به صورت زیر بیان می‌شود:

۱. آنسامبلی از M_0 سیستم $R_1, \dots, R_{M_0}$ که دربرگیرنده گامزن‌ها هستند، تولید می‌شود به طوری که همپوشانی غیر صفر قابل توجهی بین توزیع آنها و تابع موج حالت پایه وجود داشته باشد. این تولید صفرم^{۴۲} است و تعداد پیکربندی‌ها، جمعیت تولید صفرم می‌باشد.

۲. هر گامزن را با توجه به رابطه $\vec{R}' = \vec{R} + \xi$ پخش می‌شود که در این رابطه ξ از توزیع گاوسی $g(\xi) = (2\pi\tau)^{3N/2} \exp(-\xi^2 / 2\tau)$ نمونه‌گیری می‌شود یا به عبارتی گامزن‌های زنده با اضافه شدن $\rho\sqrt{\Delta\tau}$ به مختصه‌هایشان حرکت می‌کنند، که در آن ρ یک عدد تصادفی است که از توزیع گاوسی با متوسط صفر و انحراف معیار یک، گرفته شده است.

۳. برای هر گامزن حاضر در آنسامبل، فاکتور p را محاسبه می‌کنیم

$$p = \exp[-\delta\tau (V(\vec{R}) + V(\vec{R}') - 2E_T) / 2] \quad (46)$$

اگر $p < 1$ باشد، گامزن با احتمال p زنده می‌ماند، در حالیکه اگر $p \geq 1$ باشد، گامزن زنده مانده و با احتمال $p - 1$ گامزن‌های جدیدی با مختصه‌های یکسان تولید می‌شوند. این مهم به وسیله تولید تعدادی کپی برابر با قسمت صحیح $p + \eta$ ، از گامزن حاضر، که در آن η یک عدد تصادفی در بازه $(0, 1)$ می‌باشد.

۴. E_T را طوری تنظیم می‌کنیم که جمعیت کلی حول مقدار مطلوب M_0 نوسان کند.

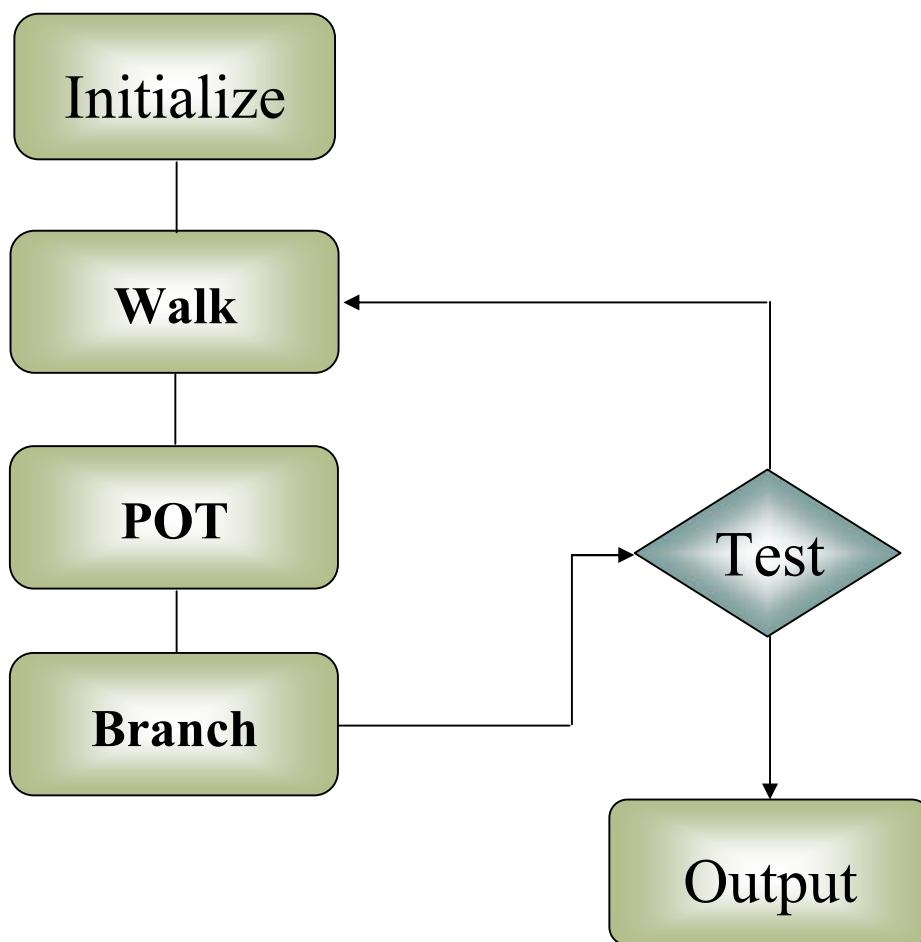
$$E_T = \langle V(\vec{R}') \rangle - \frac{M - M_0}{\delta\tau M_0}, \quad (47)$$

۴. پس از اینکه مراحل ذکر شده در بالا به تعداد کافی تکرار شدند، یعنی سیستم به حالت تعادل رسید، آنگاه تابع موج حالت پایه مستقل از آنسامبل اولیه می‌شود. حال هنگامی که تعداد سیستم‌های موجود در آنسامبل (M) همان تعداد مطلوب اولیه (M_0) باشد، متوسط انرژی پتانسیل تخمینی از انرژی حالت پایه می‌باشد.

⁴²Zero-th generation

روشی که توضیح داده شد، روش مونت کارلو پخشی نام دارد، در این روش، انرژی و تابع موج حالت پایه سیستم بدون داشتن تابع موج آزمایشی بدست می‌آید. در واقع خود الگوریتم، تابع موج حالت پایه را تخمین می‌زند که این مسئله یک امتیاز بزرگ روش مونت کارلو پخشی نسبت به روش مونت کارلو وردشی است که به مشخص کردن تابع موج آزمایشی و سپس تغییر متغیرهای وردشی احتیاج دارد.

در زیر فلوچارت ساده این الگوریتم نشان داده شده است:



شکل ۳- فلوچارت مربوط به روش مونت کارلو پخشی ساده

۶-۱ روش مونت کارلو پخشی هدایت شده

همان طور که تا اینجا توضیح داده شد روش مونت کارلو پخشی مبنی بر قیاس معادله شرودینگر/پخش، یک محدودیت مهم نیز دارد که هنوز به آن اشاره نشده است. نرخ شاخه‌زنی گام‌زن‌ها با پتانسیل متناسب است و این مسئله می‌تواند منجر به نوسانات بزرگ در جمعیت گام‌زن‌ها و کند شدن همگرایی تخمین‌زننده انرژی شود، به خصوص هنگامی که نوسانات پتانسیل بزرگ با بی‌نهایت هستند همانند مورد پتانسیل $1/r$. بنابراین احتیاج داریم که گام‌زن‌ها را به نواحی مهم آماری از فضای پیکربندی هدایت کنیم. این مشکل با استفاده از نمونه‌گیری با اهمیت همانند آنچه در مونت کارلو وردشی استفاده شد، قابل حل می‌باشد. در واقع نمونه‌گیری با اهمیت خطاهای تخمینی آماری ناشی از شاخه‌زنی مهار نشده را کاهش می‌دهد، و همزمان گام‌زن‌ها را دور از گره‌ها نگه می‌دارد.

حال در مورد این نمونه‌گیری با اهمیت بیشتر توضیح می‌دهیم :

این تکنیک به ما اجازه می‌دهد که مشخصه‌های توزیع احتمال نمونه‌گیری شده به وسیله قدم تصادفی در فضای پیکربندی $3N$ بعدی را تغییر دهیم [51 و 52]. معادله ۲۹ در یک تابع آزمایشی معلوم $\Psi_T(\vec{R})$ ضرب می‌شود، که این تابع آزمایشی باید رفتاری مشابه رفتار تابع موج حالت پایه حقیقی داشته باشد. یک انتخاب مناسب $\Psi_T(\vec{R})$ می‌تواند تابع آزمایشی استفاده شده در محاسبات مونت کارلو وردشی باشد. حال می‌توانیم معادله ۲۹ را برحسب جملات توزیع احتمال جدید $f(\vec{R}, \tau)$ بازنویسی کنیم، که $f(\vec{R}, \tau)$ به صورت زیر داده می‌شود:

$$f(\vec{R}, \tau) = \Psi(\vec{R}, \tau) \Psi_T(\vec{R}) \quad (48)$$

پس از مقداری دست‌کاری (ضمیمه A) معادله (30-2) به صورت زیر در می‌آید

$$\frac{\partial f(\vec{R}, \tau)}{\partial \tau} = D \nabla^2 f(\vec{R}, \tau) - (E_L(\vec{R}) - E_T) f(\vec{R}, \tau) - \vec{\nabla} \cdot (f \vec{F}(\vec{R})) \quad (49)$$

$$\frac{\partial \rho(\vec{R}, t)}{\partial \tau} = D \nabla^2 \rho(\vec{R}, t) - S(\vec{R}) \rho(\vec{R}, t) - \vec{\nabla} \cdot (\rho \mathbf{v}) \quad (50)$$

که در آن

$$E_L(\vec{R}) = \frac{\hat{H} \Psi_T(\vec{R})}{\Psi_T(\vec{R})} = \frac{-\nabla_{\vec{R}}^2 \Psi_T(\vec{R})}{2 \Psi_T(\vec{R})} + V(\vec{R}) \quad (51)$$

$$\vec{F}(\vec{R}) = D \nabla \ln |\Psi_T(\vec{R})|^2 = \frac{2 D \vec{\nabla} \Psi_T(\vec{R})}{\Psi_T(\vec{R})} \quad (52)$$

باز هم می‌توان مقایسه‌ای بین معادله شرودینگر ۴۹ و معادله پخش/شاخه‌زنی / سوق ۵۰ انجام داد. جملات پخش و شاخه‌گزینی هنوز هم وجود دارند (البته اصلاحاتی در جمله شاخه‌زنی به وجود آمده است)، اما معرفی تابع آزمایشی $\Psi_T(\vec{R})$ یک جمله به معادله شرودینگر زمان موهومی اضافه کرده است. اگر تنها این جمله در سمت چپ معادله وجود داشته باشد، آنگاه به معادله پیوستگی^{۴۳}

⁴³Continuity equation

می‌رسیم که بقای قدم‌زن‌ها درون حجمی از فضای $3N$ بعدی را نشان می‌دهد. قدم‌زن‌ها با سرعت سوق^{۴۴} حرکت می‌کنند که بوسیله $\vec{F}(\vec{R})$ داده شده است.

نمونه‌گیری با اهمیت به این خاطر مفید است که نرخ شاخه‌زنی توسط نوسانات انرژی موضعی E_L ، بهتر کنترل می‌شود. اگر تابع آزمایشی به طور منطقی دقیق باشد، نوسانات انرژی موضعی E_L بسیار کوچکتر از نوسانات پتانسیل $V(\vec{R})$ خواهد بود و جمعیت آنگونه که لازم است ثابت می‌شود. معادله انتگرالی معادل با معادله ۴۹ دقیقاً همانند معادله ۳۳ به نظر می‌رسد به جز اینکه تابع گرین شامل هامیلتونی جدیدی می‌باشد، که این هامیلتونی علاوه بر پخش و شاخه‌گزینی، فرآیند سوق را نیز دربر می‌گیرد.

$$f(\vec{R}, \tau + \Delta\tau) = \int d\vec{R}' G(\vec{R}, \vec{R}', \Delta\tau) f(\vec{R}', \tau). \quad (53)$$

حال باز هم از تقریب زمان کوچک استفاده می‌کنیم. یک تقریب مشهور برای تابع گرین دقیق تابع گرین زمان کوچک می‌باشد. با به کارگیری تقریب تروتور برای تابع گرین بدست می‌آوریم:

$$G(\vec{R}', \vec{R}, \Delta\tau) = \langle \vec{R}' | e^{-\Delta\tau \hat{H}} | \vec{R} \rangle \approx \langle \vec{R}' | e^{-\Delta\tau \hat{V}/2} e^{-\Delta\tau \hat{T}} e^{-\Delta\tau \hat{V}/2} | \vec{R} \rangle + O(\Delta\tau^3) \quad (54)$$

که جملات انرژی پتانسیل و جنبشی در رابطه بالا، از هامیلتونی به صورت زیر بدست می‌آیند

$$\begin{aligned} \hat{T} &= -D\nabla^2 + \nabla \cdot \vec{F}(\vec{R}), \\ \hat{V} &= E_L(\vec{R}) - E_T \end{aligned} \quad (55)$$

بنابراین می‌توان تابع گرین زمان کوچک را به صورت زیر نوشت

$$\begin{aligned} G(\vec{R}', \vec{R}; \Delta\tau) &\approx \exp[-\Delta\tau (E_L(\vec{R}') + E_L(\vec{R}) - 2E_T)/2] \langle \vec{R}' | \exp[-\Delta\tau \hat{T}] | \vec{R} \rangle \\ &= G_B(\vec{R}, \vec{R}'; \Delta\tau) G_D(\vec{R}, \vec{R}'; \Delta\tau) \end{aligned} \quad (56)$$

که در آن جمله نرخ و قسمت پخش/سوق به صورت زیر می‌باشند

$$G_B(\vec{R}', \vec{R}, \Delta\tau) = \exp[-\Delta\tau (E_L(\vec{R}') + E_L(\vec{R}) - 2E_T)/2], \quad (57)$$

$$G_D(\vec{R}', \vec{R}, \delta\tau) = \frac{1}{(4\pi D\Delta\tau)^{3N/2}} \exp\left[-\frac{(\vec{R}' - \vec{R} - \Delta\tau D\vec{F}(\vec{R}))^2}{4D\Delta\tau}\right], \quad (58)$$

خطا در تقریب برای تابع گرین از این حقیقت ناشی می‌شود که انرژی جنبشی و انرژی پتانسیل که در بالا تعریف شدند، عملگرهای غیر جابه‌جا شونده هستند و خطا همانطور که در معادله ۵۴ دیده می‌شود از مرتبه $\Delta\tau^3$ می‌باشد.

همانطور که پیش از این نیز اشاره شد، به کار بردن جابه‌جایی انرژی E_T در عبارت احتمال شاخه‌زنی $G_B(\vec{R}', \vec{R}, \Delta\tau)$ قانع کننده است زیرا E_T مجموعه ویژه‌حالات $|\Psi_i\rangle$ هامیلتونی کامل را

⁴⁴ Drift Velocity

تغییر نمی‌دهد. معادله (2-56) حاصلضرب $G_D(\vec{R}', \vec{R}, \delta\tau)$ ، جواب معادله پخش/سوق با ثابت پخش D برابر با $\frac{1}{2}$ ، و $G_B(\vec{R}', \vec{R}, \Delta\tau)$ ، جواب معادله شاخه‌زنی، می‌باشد. نرخ شاخه‌زنی قدم‌زن‌ها به وسیله انرژی موضعی E_L که در معادله ۵۱ تعریف شد، کنترل می‌شود.

حال همه اطلاعات لازم برای حل معادله شرودینگر زمان مجازی فراهم است، زیرا تابع گرین زمان کوچک، توانایی محاسبه احتمال گذار قدم‌زن‌ها را هنگامی که قدم تصادفی مونت کارلو را در فضای فاز $3N$ بعدی بازی می‌کنند، به ما می‌دهد. همچنانکه در زمان موهومی انتشار پیدا می‌کنند، انرژی موضعی در معادله ۵۱ در هر نقطه قدم‌زنی نمونه گیری می‌شود.

به طور خلاصه می‌توان گفت که در بسیاری از موارد الگوریتم مونت کارلو پخشی، می‌تواند توسط یک تابع موج آزمایشی مناسب که تابع راهنما نیز نامیده می‌شود، سریعتر و دقیقتر گردد. و این موضوع در مورد مسائلی که دارای پتانسیل تکینه هستند اهمیت پیدا می‌کند. بنابراین تابع موج آزمایشی طوری انتخاب می‌شود که هر گونه تکنیکی در پتانسیل را به طور دقیقی هموار کند. بنابراین الگوریتم مونت کارلو کوانتومی پخشی هدایت شده می‌توان به شرح زیر نوشت:

۱- مقدار دهی اولیه : در این مرحله آنسامبلی متشکل از تعداد مطلوب از قدم‌زن‌ها (M_0) ایجاد می‌کنیم و طول گام زمانی $\Delta\tau$ انتخاب می‌شود

۲- برای هر قدم‌زن یک جابه‌جایی آزمایشی به یک مکان جدید انجام می‌شود

$$\vec{R}' = \vec{R} + \frac{1}{2}\Delta\tau F(\vec{R}) + \xi\sqrt{\Delta\tau} \quad (59)$$

که ξ یک عدد تصادفی می‌باشد که از توزیع گاوسی با واریانس واحد و متوسط صفر انتخاب شده است.

۳- جابه‌جایی انجام شده قبول می‌شود اگر نسبت تست w ، که به صورت زیر تعریف می‌شود از یک عدد تصادفی یکنواخت کوچکتر باشد.

$$w = \frac{G(\vec{R}, \vec{R}'; \Delta\tau) \rho(\vec{R}')}{G(\vec{R}', \vec{R}; \Delta\tau) \rho(\vec{R})} \cong \frac{G(\vec{R}, \vec{R}'; \Delta\tau) \Psi_T^2(\vec{R}')}{G(\vec{R}', \vec{R}; \Delta\tau) \Psi_T^2(\vec{R})} \quad (60)$$

اگر تست متروپولیس موفق بود آنگاه به قسمت شاخه‌زنی می‌رویم.

۴- فاکتور شاخه‌گزینی را برای هر گام‌زن موجود در آنسامبل محاسبه می‌کنیم

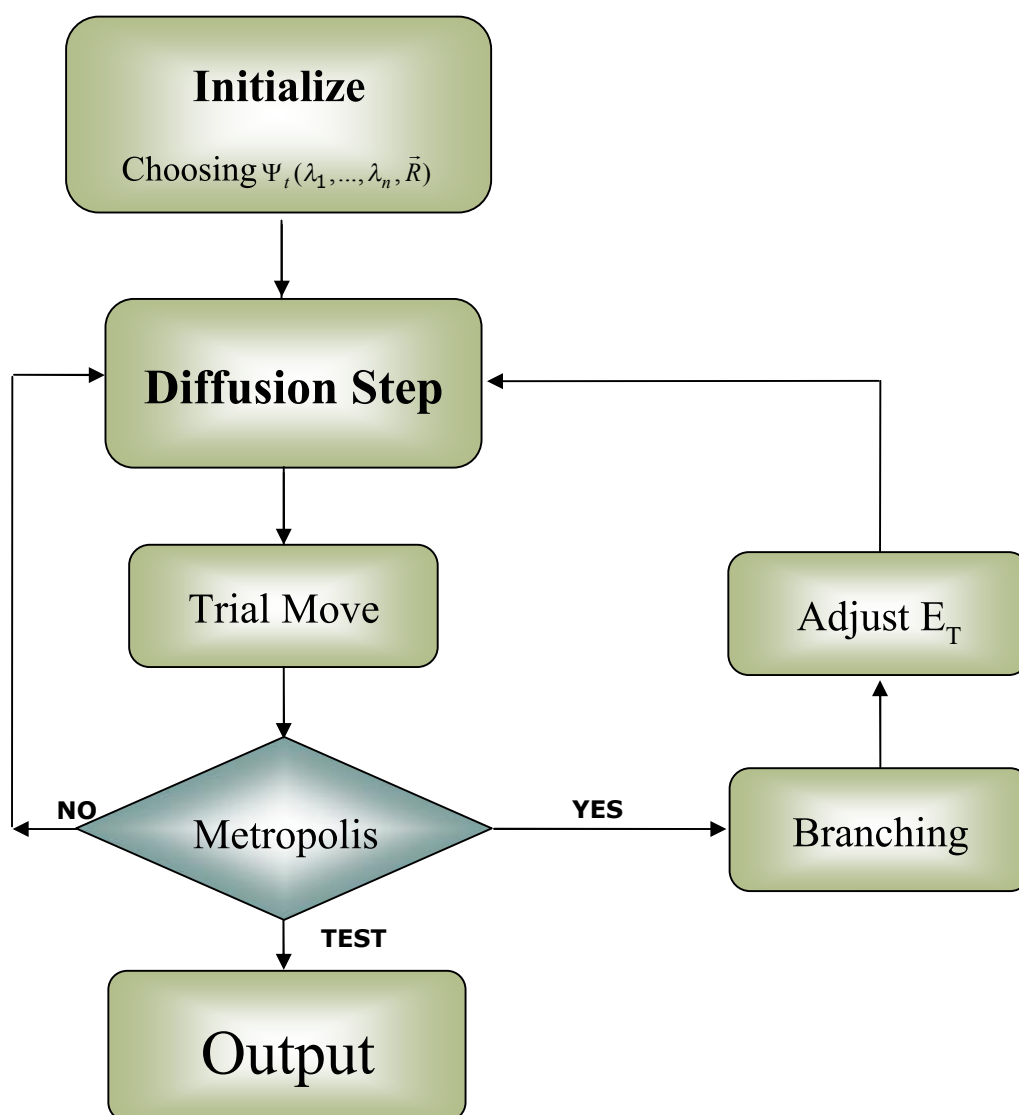
$$p = \exp[-\Delta\tau (E_L(\vec{R}') + E_L(\vec{R}) - 2E_T)/2] \quad (61)$$

همانند الگوریتم مونت کارلو پخشی ساده که قبلاً نوشته شد گام‌زن یا کشته می‌شود، یا نجات پیدا کرده و یا از آن چند کپی تهیه می‌شود.

۵- در این مرحله انرژی به صورت زیر تنظیم می‌شود این عمل برای اینکه تعداد گام‌زن‌ها در مقدار مطلوب باقی بماند انجام می‌شود

$$E_T \rightarrow E_T + \alpha \ln\left(\frac{M_T}{M_0}\right) \quad (62)$$

که در آن M_T تعداد گام‌زن‌های حاضر در این لحظه می‌باشد.
 ۶- انرژی حالت پایه، متوسط مقدار E_T روی گام‌های مونت کارلو می‌باشد.
 فلوچارت روش مونت کارلوی کوانتومی پخشی هدایت شده به صورت زیر می‌باشد:



شکل ۴- فلوچارت مربوط به روش مونت کارلو پخشی هدایت شده

۶-۲ خطای روش مونت کارلو کوانتومی پخشی

همانطور که دیده شد، در فرمول بندی و ساختار روش مونت کارلو کوانتومی از هیچ تقریبی مانند صرفنظر کردن از برخی جملات از مرتبه های خاص استفاده نشد. بنابراین می توان اینگونه عنوان کرد که این روش در ذات دقیق است. اما برای آنکه انجام محاسبات با رایانه ممکن شود ناچار به انجام تصحیحاتی هستیم که باعث وجود خطاهایی در محاسبات می شود. مهمترین منابع خطا در محاسبات مونت کارلو کوانتومی، جدای از خطای آشنای آماری، عبارتند از:

- بکار بردن گام های زمانی محدود به جای گام های زمانی بسیار بسیار کوچک که لازمه الگوریتم مونت کارلو پخشی است.
- خطای ناشی از متوسط گیری با استفاده از تعداد محدود اطلاعات برای استخراج نتایج محاسبات در انتهای شبیه سازی.
- خطاهای سیستماتیک که مربوط می شوند گرد کردن^{۴۵} در محاسبات، و ایده آل نبودن مولدهای اعداد تصادفی و از این قبیل.

در اینجا با ذکر چند نکته بخش روش محاسباتی یعنی روش های مونت کارلوی کوانتومی را به پایان می رسانیم. در این پژوهش از روش مونت کارلوی کوانتومی پخشی برای محاسبه انرژی حالت پایه و تابع موج حالت پایه اتم هیدروژن و اتم هلیوم استفاده شده است. همان طور که قبلاً نیز گفته شد از روش مونت کارلوی کوانتومی پخشی ساده در بسیاری از مسائل می توان سود جست و نتایج قابل قبول و دقیقی هم از آن بدست می آید، البته به کارگیری روش مونت کارلو کوانتومی هدایت شده دقیق تر و دارای خطای کمتری است ولی در برخی مسائل به کاربردن مونت کارلوی کوانتومی پخشی هدایت شده بسیار دشوار است.

بنابراین برای حل بخشی از مسائلی که در این پایان نامه مورد بررسی قرار گرفته اند از مونت کارلوی کوانتومی پخشی ساده و برای برخی از آنها به واسطه شرایط مرزی خاص از روش مونت کارلوی کوانتومی پخشی هدایت شده استفاده شده است، که البته مونت کارلوی کوانتومی پخشی ساده نیز نتیجه بخش بود ولی برای بالاتر بودن دقت نتایج ترجیحاً از روش هدایت شده استفاده شد.

⁴⁵Round off error

۷ اشاره‌ای به روش مونت کارلوی کوانتومی انتگرال مسیر

روشهای مونت کارلو را که تا به حال بررسی کرده ایم اغلب برای شبیه سازی و مطالعه حالت پایه مناسب هستند^{۴۶}. روش انتگرال مسیر را می توان از آن برای بررسی حالت پایه و خواص گرمایی (حالت‌های برانگیخته) دستگاههای کوانتومی استفاده کرد [2].

این روش کمی خارج از حوصله کتاب است و به مسائلی از انتگرالهای مسیر فاینمن، عملگر انتشارگر و .. مربوط می شود که احتمالاً دانشجویان فیزیک دوره های تحصیلات تکمیلی با این مفاهیم آشنایی دارند.

می دانیم که مکانیک کلاسیک را می توانیم با اصل حداقل کنش^{۴۷} فرمولبندی کنیم. این اصل بیان می کند که یک ذره کلاسیکی برای حرکت بین دو نقطه، مسیری را طی می کند که کنش S کمینه باشد.

$$S = \int_{x_0,0}^{x,t} L dt \quad (63)$$

در اینجا L لاگرانژین می باشد.

اما مکانیک کوانتومی را نیز با اصل حداقل کنش می توان فرمولبندی کرد که به آن فرمولبندی فاینمن یا انتگرال مسیر می گویند^{۴۸}.

در مکانیک کوانتومی تابع موج ذره در معادله شرودینگر صدق می کند و اگر تابع موج یک ذره را در زمان و مکان $(x_0, t_0 = 0)$ بدانیم با استفاده از عملگر انتشارگر^{۴۹} می توانیم تابع موج آن را در هر (x, t) ، با رابطه زیر تعیین کنیم.

$$\psi(x, t) = \int G(x, x_0, t) \psi(x_0, 0) dx_0, \quad t > 0 \quad (64)$$

عملگر انتشارگر G را می توان دامنه احتمال یافتن ذره ای در (x, t) که ابتدا در $(x_0, 0)$ بوده است، تعبیر کرد.

در فرمولبندی شرودینگر برای تعیین عملگر انتشارگر باید ابتدا ویژه توابع و ویژه مقادیر انرژی (هامیلتونی) را بدست آوریم و سپس انتشارگر را بر اساس آنها بیان کنیم. اما در فرمولبندی انتگرال مسیر عملگر انتشارگر را مستقیم بدست می آوریم.

بر اساس روش انتگرال مسیر برای تعیین عملگر انتشارگر باید تمام مسیرهای ممکن که $(x_0, 0)$ را به (x, t) متصل می کنند در نظر بگیریم و کنش مربوط به هر کدام را محاسبه کنیم^{۵۰}. سپس انتشارگر G بصورت زیر بدست می آید:

^{۴۶} البته با تلاش بیشتر می توان از این روشها برای بررسی چند حالت برانگیخته اول استفاده کرد.

^{۴۷} Principle Of Least Action

^{۴۸} - فرمولبندی رایج مکانیک کوانتومی فرمولبندی شرودینگر است که بر اساس مکانیک هامیلتونی می باشد.

^{۴۹} - $G = e^{\frac{-iHt}{\hbar}}$

$$G(x, x_0, t) = A \sum_{Paths} \exp\left(\frac{iS}{\hbar}\right) \quad (۶۵)$$

که A ضریب بهنجارش است. چون باید تمام مسیر هارا در نظر بگیریم جمع بندی بر روی تمام مسیره‌ای بین $(x_0, 0)$ و (x, t) می باشد.

همانطور که گفتیم G در معادله (۶۶-۳) صدق می کند. همچنین G و ψ هر دو یک معادله دیفرانسیل مشابه از x_0 و x را ا قناع می کنند. حال با توجه به اینکه ψ را می توان بر حسب ویژه توابع H_{op} یعنی $\phi_n(x)$ ها بسط داد G را بشکل زیر می نویسیم.

$$G(x, x_0, t) = \sum_n \phi_n(x) \phi_n(x_0) \exp\left(-\frac{iE_n t}{\hbar}\right) \quad (۶۶)$$

با تغییر متغیر $\tau = \frac{it}{\hbar}$ و انتخاب $\hbar = 1$ معادله (۶۷-۳) بشکل زیر نوشته می شود:

$$G(x, x_0, \tau) = \sum_n \phi_n(x) \phi_n(x_0) \exp(-\tau E_n) \quad (۶۷)$$

ابتدا حالت پایه را بررسی می کنیم.

در حد $\tau \rightarrow \infty$ داریم:

$$G(x, x_0, \tau) \rightarrow \phi_0(x_0)^2 \exp(-\tau E_0) \quad (\tau \rightarrow \infty) \quad (۶۸)$$

رابطه ۶۸ بر این تاکید می کند که حالت حدی جمله قابل توجه مربوط به تابع موج حالت پایه است و از باقی جملات می توان صرفنظر کرد.

همانطور که از رابطه ۶۸ دیده میشود برای بررسی حالت پایه باید انتشارگر G را داشته باشیم و از رابطه ۶۵ می بینیم که نیاز به محاسبه کنش S داریم. برای محاسبه S نیز باید لاگرانژین ۶۲ ذره را بدانیم.

با توجه به متغیر τ که برابر، لاگرانژین یک ذره به جرم واحد است، بصورت زیر نوشته می شود:

$$L = -\frac{1}{2} \left(\frac{dx}{d\tau}\right)^2 - V(x) = -E \quad (۶۹)$$

با تقسیم کردن بازه زمانی τ به N قسمت مساوی $\Delta\tau$ خواهیم داشت:

$$E(x_j, \tau_j) = \frac{1}{2} \frac{(x_{j+1} - x_j)^2}{(\Delta\tau)^2} + V(x_j) \quad (۷۰)$$

که در آن $\tau_j = j\Delta\tau$ و x_j جابجایی متناسب با آن می باشد.

^{۵۰} - مبحث انتگرال مسیر فاینمن (Feynman Path Integral) را می توان در اکثر کتابهای مکانیک کوانتوم یافت. از جمله کتاب Principle Of Quantum Mechanics, sec. ed., R. Shankar. فاینمن این موضوع را در مقاله ای تحت عنوان Space-Time Approach to Non-Relativistic Quantum Mechanics, Rev. Of Modern Phys., 20, 2, 1948 برای اولین بار مطرح کرد. همچنین این موضوع در کتاب Quantum Mechanics and Integrals, by R. P. Feynman and A.. R. Hibbs, McGraw-Hill (1965) به

با گسسته سازی زمان به نحوی که گفته شد، انتگرال گیری معادله (۶۳) به یک جمع بندی تبدیل می شود. در اینحالت بااستفاده از معادله (۷۰) و با کمک تقریب مستطیلی انتگرال، کنش بشکل زیرتعیین می گردد:

$$S = -i\Delta\tau \sum_{j=0}^{N-1} E(x_j, \tau_j) = -i\Delta\tau \left[\sum_{j=0}^{N-1} \frac{1}{2} \frac{(x_{j+1} - x_j)^2}{(\Delta\tau)^2} + V(x_j) \right] \quad (71)$$

و با توجه به رابطه بالا خواهیم داشت:

$$\exp(iS) = \exp(\Delta\tau \left[\sum_{j=0}^{N-1} \frac{1}{2} \frac{(x_{j+1} - x_j)^2}{(\Delta\tau)^2} + V(x_j) \right]) \quad (72)$$

می باشد.

اکنون عملگر انتشارگر G را از روی معادله (۷۲) و با استفاده از نتیجه فوق بصورت زیر تعیین می کنیم:

$$G(x, x_0, N\Delta\tau) = A \int dx_1 \dots dx_{N-1} \exp(\Delta\tau \left[\sum_{j=0}^{N-1} \frac{1}{2} \frac{(x_{j+1} - x_j)^2}{(\Delta\tau)^2} + V(x_j) \right]) \quad (73)$$

در رابطه (۷۳)، $x \equiv x_N$ می باشد و A ثابتی بی اهمیت است. همانطور که دیده می شود جمع بندی معادله (۶۵) به انتگرال گیری چند بعدی تبدیل شده است زیرا که هر مسیر بین x_0 و x در بازه زمانی $[0 - t_0]$ با تعداد زیادی عدد $\{x(t_0), \dots, x(t), \dots, x(t)\}$ ، که در واقع مقدار $x(t)$ در هر لحظه می باشند، تعیین می گردد؛ پس برای آنکه جمع بندی روی تمام مسیرها انجام شود باید بر روی تمام این متغیرها بجز $x(t_0)$ و $x(t_N)$ انتگرال گیری کنیم. بنابراین انتگرال چند بعدی بر روی تمام مسیرها می باشد.

بااستفاده از شرط تناوبی مرزی، $x_N = x_0$ ، در معادله (۷۳) می توانیم جمع بندی j را از ۱ تا N انجام دهیم.

$$G(x_0, x_0, N\Delta\tau) = A \int dx_1 \dots dx_{N-1} e^{-\Delta\tau \left[\sum_{j=1}^N \frac{1}{2} \frac{(x_j - x_{j-1})^2}{(\Delta\tau)^2} + V(x_j) \right]} \quad (74)$$

تحلیل فوق در واقع تبدیل مساله مکانیک کوانتومی تک ذره به یک مساله مکانیک آماری N ذره است، که در آن ذرات بر روی یک حلقه قرار دارند و بوسیله فنرهایی با ثابت $\frac{1}{(\Delta\tau)^2}$ به نزدیکترین همسایه های خود متصلند. j معرف مکان ذره بر روی حلقه و x_j جابجایی ذره j ام از محل تعادلش می باشد.

برای بررسی اینگونه مسائل مکانیک آماری از روش متروپلیس استفاده می کنیم. ذکر این نکته لازم است که این ذرات واقعی نیستند و مانند ولگردها در $DQMC$ فقط نتیجه تحلیل ماهستند.

نمونه ای از یک الگوریتم انتگرال مسیر می تواند بشکل زیر باشد.

الگوریتم PIQMC

۱- N و $\Delta\tau$ را طوری انتخاب می‌کنیم که $N\Delta\tau \gg 1$ باشد. (حد دمای صفر)
 δ ، حداکثر تغییر آزمایشی در جابجایی یک ذره و MCS، تعداد گام مونت کارلو بازاء هر ذره را انتخاب می‌کنیم.

۲- توزیع اولیه جابجایی‌های x_j را مشابه توزیع تقریبی دامنه احتمال حالت پایه، انتخاب می‌کنیم.
 ۳- ذره j ام را بطور تصادفی انتخاب و یک جابجایی آزمایشی بصورت 1- $x_j \rightarrow x_{trial}$ برای آن انجام می‌دهیم. (یک عدد تصادفی با توزیع یکنواخت در بازه $[0,1]$ است).
 سپس تغییر انرژی ΔE را محاسبه می‌کنیم.

$$\Delta E = \frac{1}{2} \left(\frac{x_{j+1} - x_{trial}}{\Delta\tau} \right)^2 + \frac{1}{2} \left(\frac{x_{trial} - x_{j-1}}{\Delta\tau} \right)^2 + V(x_{trial}) - \frac{1}{2} \left(\frac{x_{j+1} - x_j}{\Delta\tau} \right)^2 - \frac{1}{2} \left(\frac{x_j - x_{j-1}}{\Delta\tau} \right)^2 - V(x_j) \quad (75)$$

اگر $\Delta E < 0$ جابجایی را قبول می‌نماییم. در غیر اینصورت احتمال $p = \exp(-\Delta\tau\Delta E)$ را حساب کرده و با یک عدد تصادفی یک در بازه $[0,1]$ مقایسه می‌کنیم؛ اگر که $p \leq \xi$ ، جابجایی قبول خواهد شد. در غیر اینصورت این حرکت آزمایشی را کنار می‌گذاریم.

۴- امان $p(x_j)$ آرایه چگالی احتمال را به روز می‌کنیم. یعنی:
 $P(x = x_j) \rightarrow P(x = x_j) + 1$

(x_j جابجایی ذره انتخاب شده در مرحله سه بعد از اتمام آن مرحله است).
 به روز کردن را حتی وقتی حرکت آزمایشی مردود می‌شود انجام می‌دهیم. مقادیر ممکن x را به دسته‌های با پهنای یکسان Δx تقسیم بندی می‌نماییم.

۵- مراحل سه و چهار را به تعداد کافی از گامهای مونت کارلو بازاء هر ذره، تکرار می‌کنیم. (اطلاعات را تا وقتی که حافظه مسیر اولیه پاک نشده و دستگاه به تعادل نرسیده استخراج نباید کرد).
 چگالی احتمال $P(x)$ را با تقسیم کردن بر حاصلضرب mcs و N بهنجار کنید. پس از به تعادل رسیدن دستگاه انرژی حالت پایه با رابطه زیر تعیین می‌شود:

$$E_0 = \sum_x P(x) [T(x) + V(x)] \quad (76)$$

که در آن انرژی جنبشی $T(x)$:

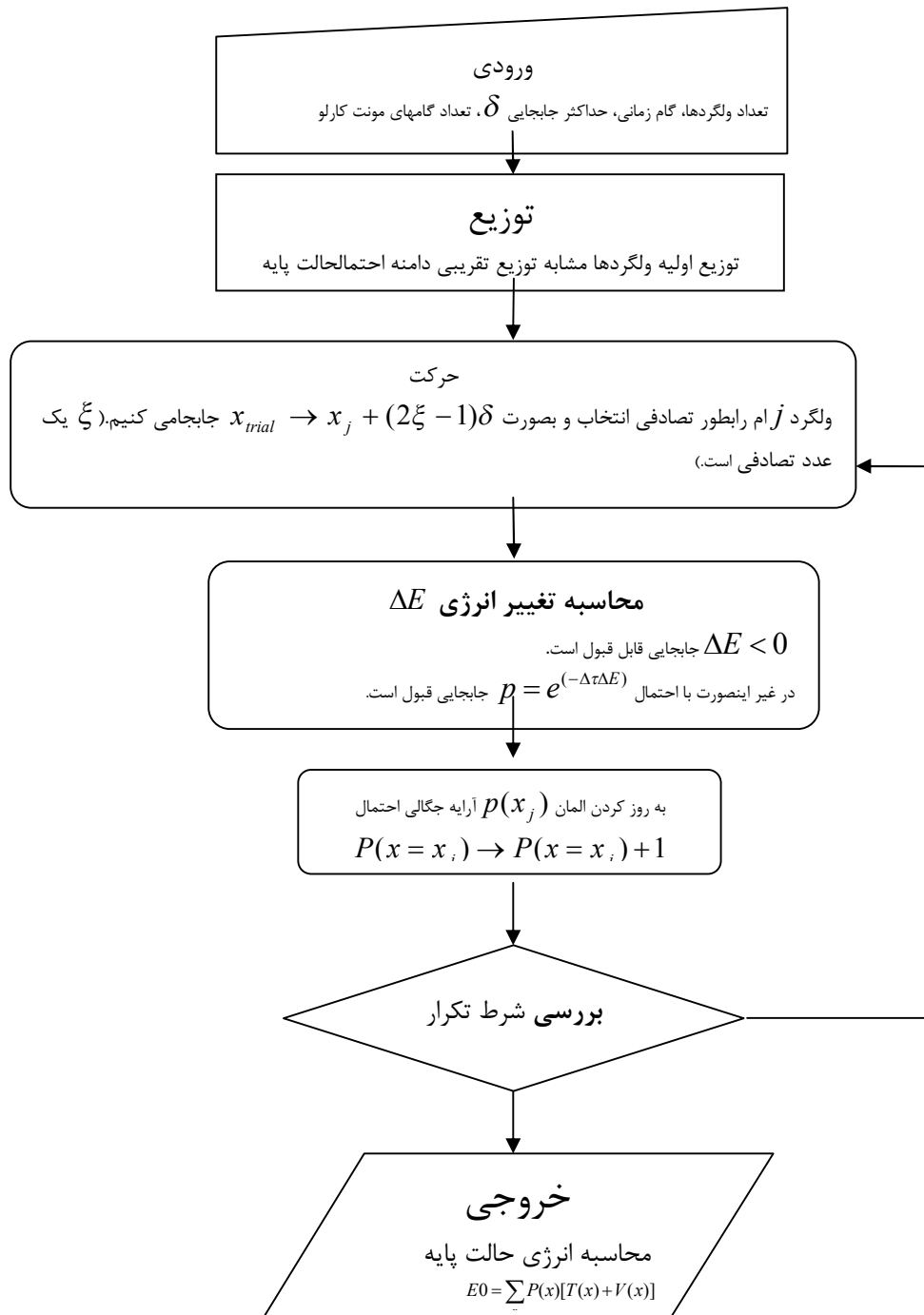
$$2T(x) = x \frac{dV}{dx} \quad (77)$$

می‌باشد.

تابع موج حالت پایه، $\phi(x)$ ، نیز از تقسیم کردن احتمال بهنجار شده $P(x)\Delta x$ بر Δx و محاسبه جذرش بدست می‌آید (بعبارت دیگر $\phi(x) = \sqrt{P(x)}$).

همانگونه که ذکر شد برای محاسبه حالت پایه، متناسب با حد دمای صفر ($\beta = \frac{1}{T} \gg 1$)، باید $N\Delta\tau$ خیلی بزرگ باشد. اما $\Delta\tau$ نیز باید تا حد ممکن کوچک باشد تا تقریب زمان پیوسته برقرار بماند. بنابراین در این حالت N را خیلی بزرگ انتخاب می‌کنیم.

با استفاده از روش مونت کارلو کوانتومی انتگرال مسیر می‌توانیم خصوصیات ترمودینامیکی ذره متصل به حمام گرمایی با دمای $T = \frac{1}{\beta}$ را، بادر نظر نگرفتن حد $\beta = N\Delta\tau \rightarrow \infty$ (محدود کردن $N\Delta\tau$) پیدا کنیم.^{۵۱} در شکل بعد می‌توان فلوچارت **PIQMC** را مشاهده کرد. کتاب مر بوط به مرجع [2] مثالهایی مفید از این روش ارائه داده است.



شکل ۵- الگوریتم روش انتگرال مسیر

^{۵۱} - این کار در واقع بررسی حالت برانگیخته ذره است.

۸ بررسی خطای محاسبات QMC

بررسی کارایی شبیه‌سازی‌های مونت کارلو و تحلیل اطلاعات به دست آمده از محاسبات QMC بسیار مهم است. در این بخش به بررسی تکنیک‌های آماری محاسبه‌ی میزان خطای مشاهده‌پذیرها می‌پردازیم. این تکنیک‌ها در تمام روش‌های مونت کارلو یکسان است. مقدار چشم‌داشتی هر مشاهده‌پذیر به صورت زیر به دست می‌آید:

$$\langle X \rangle = \frac{1}{N} \sum_{i=1}^N x_i \quad (78)$$

در تحلیل اطلاعات به دست آمده از شبیه‌سازی‌های QMC دو عامل بسیار مهم است:

۱- زمان تعادل مشاهده‌پذیر

۲- همبستگی^{۵۲} در نمونه‌های به دست آمده

در واقع جمع ۱-۲، پس از رسیدن نتایج محاسبات به تعادل روی نمونه‌هایی صورت می‌گیرد که مستقل از یکدیگر باشند. میزان خطای محاسبات QMC، به انحراف میانگین نمونه‌ها، σ ، بستگی دارد و برابر $\frac{\sigma}{\sqrt{N-1}} \approx \frac{\sigma}{\sqrt{N}}$ می‌باشد. با استفاده از قضیه حد مرکزی، σ در انتگرال‌گیری از تابع $f(x)$ به صورت زیر به دست می‌آید [۱۵]:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [f(x_i) - \langle f \rangle]^2} \quad (79)$$

که

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (80)$$

می‌توان رهیافت دیگری در محاسبه‌ی انحراف میانگین که بر پایه قسمت‌بندی نمونه‌هاست به کار برد. در این روش نخست نمونه‌ها در بلوک‌های $(1, 2, 4, \dots)$ تایی قسمت‌بندی می‌کنیم. انحراف میانگین از نمونه‌ها با محاسبه‌ی واریانس روی میانگین بلوک‌ها به دست می‌آید. برای هر بلوک می‌توان انحراف میانگین را به کمک رابطه زیر محاسبه کرد:

$$\sigma_b = \sqrt{\frac{1}{M_b-1} \sum_{j=1}^{M_b} (\langle x_j^2 \rangle - \langle x \rangle_b^2)} \quad (81)$$

که x_j میانگین نمونه‌های بلوک j ام است. $\langle x \rangle_b$ میانگین روی تمام بلوک‌هاست و M_b شمار بلوک‌هاست.

⁵²Correlation

۸-۱ خطاهای دستگاهی

در بخش پیش روش آماری برای تحلیل نمونه‌های QMC بررسی شد. این تکنیک‌ها توانایی تشخیص خطاهای دستگاهی که در شبیه‌سازی رخ می‌دهد را ندارند. برخی از این خطاها عبارتند از:

* جستجوی ناکافی در فضای نمونه‌گیری؛ در نتیجه‌ی کوتاه بودن زمان شبیه‌سازی (شمار اندک نمونه‌گیری‌ها) و گزینش نادرست گام زمانی مونت کارلو که درصد پذیرش جابجایی مونت کارلو را در الگوریتم متروپولیس محدود می‌کند (درصد پذیرش جابجایی‌ها معمولاً به گونه‌ای گزیده می‌شود که در محاسبات VMC تقریباً نیمی از جابجایی‌ها پذیرفته شود)، رخ می‌دهد.

* میانگین‌گیری روی نمونه‌های غیرتصادفی (هنگامی که دستگاه به تعادل نرسیده باشد) نتایج نامعقولی را به دست می‌دهد. بنابراین پس از رسیدن دستگاه به تعادل باید جمع ۷۸ را در محاسبه‌ی مقادیر چشم‌داشتی مشاهده‌پذیرها به کار گیریم.

۹ توابع موج مونت کارلوی کوانتومی

۹-۱ مقدمه

گزینش تابع موج آزمایشی یکی از مهم‌ترین بخش‌های محاسبات مونت کارلوی کوانتومی است. به دست‌آوردن نتایج دقیق در روش VMC به گزینش یک تابع موج آزمایشی خوب بستگی دارد. شکل تابع موج و شمار پارامترهای آن بسته به دستگاه مورد بررسی متفاوت است [۱۷]. در این فصل ویژگی‌های توابع موج مونت کارلوی کوانتومی در دستگاه‌های اتمی را بررسی می‌کنیم.

۹-۲ تقارن

تمام ذرات بنیادی را می‌توان به دو بخش فرمیون‌ها و بوزون‌ها تقسیم‌بندی کرد. بوزون‌ها (مانند فوتون‌ها و اتم‌های دوتریم) ذراتی با اسپین درست هستند، در حالیکه فرمیون‌ها (مانند الکترون‌ها و پروتون‌ها) اسپین نیم‌درست دارند. مکانیک کوانتومی بوزون‌ها و فرمیون‌ها کاملاً با یکدیگر متفاوت است. تابع موج کل یک دستگاه بوزونی باید متقارن باشد، از این رو جابجایی مکان‌های هر دو ذره‌ی یکسان شکل تابع موج را تغییر نمی‌دهد:

$$\Psi_{boson}(\dots, x_i, \dots, x_j, \dots) = \Psi_{boson}(\dots, x_j, \dots, x_i, \dots) \quad (۸۲)$$

از طرف دیگر تابع موج کل یک دستگاه فرمیونی پادمتقارن است و تعویض همزمان مکان‌ها و اسپین‌های هر دو ذره‌ی یکسان علامت تابع موج را تغییر می‌دهد:

$$\Psi_{fermion}(\dots, x_i, \dots, x_j, \dots) = -\Psi_{fermion}(\dots, x_j, \dots, x_i, \dots) \quad (۸۳)$$

اصل رانش پائولی (دو الکترون در یک دستگاه با ویژگی‌های یکسان نمی‌توانند در یک لحظه مکان یکسانی را اشغال کنند) یک نتیجه مستقیم از پادمتقارن بودن تابع موج دستگاه‌های فرمیونی است.

۹-۳ تابع موج آزمایشی در محاسبات مونت کارلوی کوانتومی وردشی

از آنجا که در این پایان‌نامه به بررسی حالت پایه‌ی دستگاه‌های اتمی، که بخش الکترونی آن فرمیونی است می‌پردازیم تابع موج آزمایشی باید پادمتقارن باشد. هر تابع موج پادمتقارن می‌تواند به عنوان تابع موج آزمایشی به کار رود اما نزدیک‌ترین تابع موج به حالت پایه‌ی دستگاه است که می‌تواند همگرایی تندتر در روش VMC و حل دقیقی برای حالت پایه‌ی دستگاه به دست دهد. بنابراین گزینش شکل تابع موج آزمایشی از مهم‌ترین بخش‌های روش VMC به شمار می‌آید. تابع موج زیر را می‌توان به عنوان یک تابع موج کلی خوب که اهداف روش VMC را برآورده می‌کند در نظر گرفت [۱۶]:

$$\Psi(\vec{R}) = D^\uparrow(\vec{r}_1, \dots, \vec{r}_{N_{up}}) D^\downarrow(\vec{r}_1, \dots, \vec{r}_{N_d}) e^J \quad (۸۴)$$

$D^\uparrow$ ، دترمینان اسلیتر بالا (شامل اوربیتال‌های تک‌ذره‌ای N_{up} الکترون با اسپین بالا ($\uparrow$)) و $D^\downarrow$ ، دترمینان اسلیتر پایین (شامل اوربیتال‌های تک‌ذره‌ای N_d الکترون با اسپین پایین ($\downarrow$)) می‌باشد. $(\vec{r}_1, \dots, \vec{r}_N)$ در دترمینان‌های اسلیتر فاصله‌ی هر الکترون را از هسته نشان می‌دهند. اوربیتال‌های تک‌ذره‌ای در دترمینان‌های اسلیتر به صورت زیر قرار می‌گیرند:

$$D^\uparrow = \begin{vmatrix} \varphi_1(\vec{r}_1) & \varphi_1(\vec{r}_2) & \cdots & \varphi_1(\vec{r}_{N_{up}}) \\ \varphi_2(\vec{r}_1) & \varphi_2(\vec{r}_2) & \cdots & \varphi_2(\vec{r}_{N_{up}}) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N_{up}}(\vec{r}_1) & \varphi_{N_{up}}(\vec{r}_2) & \cdots & \varphi_{N_{up}}(\vec{r}_{N_{up}}) \end{vmatrix}, \quad D^\downarrow = \begin{vmatrix} \varphi_1(\vec{r}_1) & \varphi_1(\vec{r}_2) & \cdots & \varphi_1(\vec{r}_{N_d}) \\ \varphi_2(\vec{r}_1) & \varphi_2(\vec{r}_2) & \cdots & \varphi_2(\vec{r}_{N_d}) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N_d}(\vec{r}_1) & \varphi_{N_d}(\vec{r}_2) & \cdots & \varphi_{N_d}(\vec{r}_{N_d}) \end{vmatrix} \quad (۸۵)$$

اگرچه استفاده از دترمینان‌های مجزا ($D^\downarrow$ و $D^\uparrow$) برای الکترون‌های با اسپین‌های مخالف پادمتقارن بودن تابع موج آزمایشی نسبت به تعویض مکان دو الکترون را نتیجه نمی‌دهد اما مقادیر چشم‌داشتی درستی نسبت به اثر عملگرهای نابسته از اسپین می‌دهد [۱۸].

در رابطه‌ی ۸۴، J ، عامل جسترو^{۵۳} نامیده می‌شود. تابع جسترو، e^J ، در بیشتر موارد تابعی از مکان همه‌ی الکترون‌ها و اتم‌هاست. توابع جسترو را در بخش پسین مورد بررسی قرار می‌دهیم. بخش دترمینانی تابع موج معمولاً از روش‌های خودسازگار^{۵۴} مانند LDA^{۵۵}، DFT^{۵۶} یا HF^{۵۷} [۲۰ و ۱۹] به دست می‌آید. در محاسبات انجام شده در این پژوهش اوربیتال‌های تک‌ذره‌ای، φ_i ، که در مجموعه‌ی پایه‌های متفاوت بسط داده می‌شوند، در دترمینان‌های اسلیتر بالا و پایین استفاده شده است، در بخش‌ها بعدی به بررسی این مجموعه پایه‌ها می‌پردازیم.

برای دستگاه‌هایی با گاف انرژی کوچک از دترمینان‌های چندگانه در تابع موج شامل ضرایب a_n به صورت زیر استفاده می‌شود:

$$\psi(\vec{R}) = \sum_n a_n D_n^\uparrow(\vec{r}_1, \dots, \vec{r}_N) D_n^\downarrow(\vec{r}_1, \dots, \vec{r}_N) e^J \quad (۸۶)$$

که a_n ها ثابت هستند. در واقع هر D_n ، اوربیتال‌های ویژه‌ی خودش را دارد.

اما در دستگاه‌های اتمی که در این پایان‌نامه بررسی می‌شوند و دستگاه‌هایی که گاف انرژی بزرگی دارند استفاده از دترمینان‌های چندگانه بهبود قابل توجهی در محاسبه‌ی انرژی حالت پایه ایجاد نمی‌کند [۱۶]. در بیشتر کاربردهای QMC می‌توان با دقت بسیار خوبی تنها از یک مجموعه دترمینان با اسپین بالا و پایین استفاده کرد.

⁵³Jastrow factor

⁵⁴Self-Consistent

⁵⁵Linear Density approximation

⁵⁶Density Functional Theory

⁵⁷Hartree Fock

۹-۴ مجموعه پایه‌ها

درواقع اوربیتال‌های φ_i ، در دترمینان‌های اسلیتر می‌توانند در مجموعه پایه‌هایی^{۵۸} که برای محاسبات QMC بهینه باشد بسط داده شوند، از این رو گزینش نوع پایه در بسط اوربیتال‌ها بسته به دستگاه کوانتومی مورد بررسی بسیار مهم است. در این بخش به بررسی انواع پایه‌ها می‌پردازیم.

۹-۴-۱ اوربیتال‌های نوع نمایی

برای نخستین بار استفاده از توابع ریاضی به شکل اوربیتال اتم هیدروژن در محاسبات اتمی پیشنهاد شد از آنجا که این پیشنهاد توسط اسلیتر^{۵۹} [۲۱] و زنر^{۶۰} [۲۲] ارائه شد، این‌گونه توابع به توابع (اسلیتر) مشهور شده‌اند که به صورت زیر تعریف می‌شوند:

$$\varphi(\vec{r}) = r^{n*-1} e^{-[(Z-s)/n*]r} \quad (۸۷)$$

که n^* عدد کوانتومی مؤثر، s ثابت پوششی و Z عدد اتمی است. توجه داریم این پارامترها به صورت وردشی مشخص نمی‌شوند و مقادیر آنها پیش از آغاز محاسبات جهت یافتن انرژی مشخص می‌گردد. پس از آن روتن^{۶۱} و باگوس^{۶۲} [۲۳] شکل زیر را معرفی کردند:

$$\varphi(\vec{r}) = r^{n-1} e^{-\zeta r} Y_{l,m}(\theta, \varphi) \quad (۸۸)$$

که n ، l و m به ترتیب اعداد کوانتومی اصلی و مغناطیسی هستند. $Y_{l,m}(\theta, \varphi)$ هماهنگ-های کروی‌اند که به l و m بستگی دارند. ζ در رابطه‌ی ۸۸، پارامتر وردش است که باید در فرآیند محاسبات بهینه گردد. به طور کلی توابعی به شکل رابطه‌ی ۸۸، اوربیتال‌های نوع اسلیتر (STO) نامیده می‌شوند که ویژگی مهم آن‌ها بستگی نمایی به r است.

توابع STO رفتار خوبی در نواحی نزدیک و دور از هسته از خود نشان می‌دهند و به خاطر رفتار نمایی‌شان نسبت به فاصله‌ی الکترون-هسته شرایط کاسپ^{۶۳} هسته‌ای را برآورده می‌کند (در بخش ۹-۶، به بررسی شرایط کاسپ پرداخته می‌شود). از این رو مجموعه پایه‌های STO بدون افزایش اعداد اتمی زاویه‌ای توصیف درستی برای نواحی نزدیک به هسته ارائه می‌دهند، از طرف دیگر در فاصله‌های زیاد درست شبیه به اوربیتال‌های اتمی در نواحی دور از هسته کاهش می‌یابند.

⁵⁸Basis set⁵⁹Slater⁶⁰Zener⁶¹Roothan⁶²Bagus⁶³Cusp Condition

نخست مجموعه پایه‌های توابع STO برای اتم‌های He تا Ne برای یافتن انرژی حالت پایه، توسط کلمنتی^{۶۴} بهینه شده‌اند [۲۴]. پس از آن کلمنتی و روتی^{۶۵} [۲۵] بهینه‌سازی توابع STO را تا اتم Xe و مک‌لین^{۶۶} تا اتم Re در جدول دوره‌ای عناصر انجام دادند. می‌توان پایه‌های STO را تنها با یک نما برای هر زیر پوسته به صورت زیر بسط داد [۲۶ و ۲۷]:

$$R_s(r) = e^{-\zeta_{1s}r} \sum_{i=1}^{k_{1s}} a_i^s r^{i-1} + e^{-\zeta_{2s}r} \sum_{j=1}^{k_{2s}} b_j^s r^{j-1} + \dots \quad (۸۹)$$

$$R_p(r) = re^{-\zeta_{2p}r} \sum_{i=1}^{k_{2p}} a_i^p r^{i-1} + \dots \quad (۹۰)$$

$R_s(r)$ و $R_p(r)$ قسمت شعاعی در رابطه‌ی ۸۸، را نشان می‌دهد. بسط پایه‌ها به صورت روابط ۸۹ و ۹۰ باعث می‌شود با کاهش پارامترهای ورودش سرعت محاسبات افزایش یابد. از این‌رو استفاده از این پایه‌ها (توابع اسلیتر با نمای یکتا^{۶۷} (SESF)) در دستگاه‌های اتمی و مولکولی سودمند است.

جدول ۱) پایه‌های SESF در بهینه‌سازی اتم کربن [۲۶ و ۲۷]، اعداد $\{k_{2p}\}$ و $\{k_{1s}, k_{2s}\}$ کران‌های بالا را در جمع روابط ۸۹ و ۹۰ نشان می‌دهد.

Basis Set	ζ_{1s}	ζ_{2s}	ζ_{2p}	$E(a.u.)$	Virial Ratio
{33} {4}	6.122253	1.458034	1.753768	-37.685803358	-2.0000000000
{33} {5}	6.115960	1.458734	1.942717	-37.687891980	-2.0000000001
{34} {4}	6.779258	1.728607	1.752483	-37.686225152	-2.0000000000
{34} {5}	6.771996	1.714643	1.942551	-37.688333474	-2.0000000000
{35} {5}	6.819651	1.716263	1.942562	-37.688333874	-2.0000000000
{44} {4}	6.715725	1.643143	1.752494	-37.686234034	-2.0000000000
{44} {5}	6.717072	1.627785	1.942140	-37.688342713	-2.0000000000
{45} {5}	6.635614	1.662539	1.942092	-37.688344748	-2.0000000000
{55} {5}	5.925759	1.568993	1.942018	-37.688346241	-2.0000000000

در واقع هر پایه در هر زیر پوسته‌ی اوربیتالی یک ترکیب خطی از پایه‌های STO با نماهای یکسان $\zeta_{1s}, \zeta_{2s}, \zeta_{2p}, \dots$ و اعداد کوانتومی m و l یکسان می‌باشد. جدول ۱ برخی نتایج به-کارگیرپایه‌های SESF را برای اتم کربن نشان می‌دهد.

⁶⁴Clementi

⁶⁵Rotti

⁶⁶McLean

⁶⁷Single Exponent Slater Function

به جز STO، از انواع دیگر پایه‌های نوع نمایی که در دستگاه‌های مولکولی کاربرد دارد توسط فیلتر^{۶۸} و استینبورن^{۶۹} [۲۸] موسوم به اوربیتال‌های نوع بسل (BTO) پیشنهاد شد و به صورت زیر تعریف می‌شوند:

$$\varphi(\vec{r}) = \hat{K}_{n-1/2}(\zeta r) r^l Y_{l,m}(\theta, \varphi) \quad (91)$$

که در آن

$$\hat{K}_{n-1/2}(\zeta r) = \sqrt{\frac{\pi}{2}} e^{-\zeta r} \sum_{i=0}^{n-1} \frac{(n-1+i)! (\zeta r)^{n-i-1}}{(n-1-i)! i! 2^i} \quad (92)$$

توابع بسل $\hat{K}_{n-1/2}(\zeta r)$ تبدیل فوریه‌ی ساده‌ای دارند. معادله‌ی ۹۱ نشان می‌دهد که پایه‌های BTO یک ترکیب خطی از پایه‌های STO هستند اما رفتار بهتری نسبت به STO در فاصله‌های نزدیک و بسیار دور از هسته از خود نشان می‌دهند.

۹-۴-۲ اوربیتال‌های نوع گوسی

از آنجا توابع نمایی و به ویژه توابع STO در ساختارهای الکترونیکی مولکولی مناسب نبودند توابع گوسی که به طور کلی اوربیتال‌های نوع گوسی^{۷۰} (GTO) نامیده می‌شوند جایگزین پایه‌های STO در محاسبات مولکولی شدند. برای نخستین بار توابع GTO کروی توسط بویز^{۷۱} و مکوینی^{۷۲} [۲۹] به شکل زیر پیشنهاد شد:

$$\varphi(\vec{r}) = r^{n-1} e^{-\alpha r^2} Y_{l,m}(\theta, \varphi) \quad (93)$$

α نمای تابع GTO است. l و m به ترتیب اعداد کوانتومی مداری و مغناطیسی می‌باشند و مقادیر n به صورت زیر تعریف می‌شود:

$$n = l+1, l+3, l+5, \dots \quad (94)$$

معمولاً تنها $n = l+1$ در روابط استفاده می‌شود. برای سادگی در حل انتگرال‌ها می‌توان در محاسبات مولکولی از توابع GTO دکارتی^{۷۳} استفاده کرد که البته دیگر تقارن کروی ندارند:

$$\varphi(\vec{r}) = x^i y^j z^k e^{-\alpha r^2} \quad (95)$$

که i ، j و k اعداد درست مثبت هستند. با گزینش $n = l+1$ ، توابع GTO تنها شامل اوربیتال 1s برای اوربیتال‌های اتمی نوع s ، $2p$ برای اوربیتال‌های p ، $3d$ برای اوربیتال‌های نوع d و ... می‌باشد. با این وجود هوزیناگا^{۷۴} ثابت کرد هنگامی که $n = l+1$ ،

⁶⁸Filter

⁶⁹Steinborn

⁷⁰Gaussian Type Orbitals

⁷¹Boys

⁷²McWeeny

⁷³Cartesian GTO

⁷⁴Huzinaga

لازم نیست که مقدار n در محاسبات افزایش یابد. مشتق یکم توابع GTO نسبت به r وقتی که r به سوی صفر میل می‌کند بسیار بزرگ می‌شود، در حالیکه توابع STO در $r \rightarrow 0$ این‌گونه نیستند. از این رو توابع GTO رفتار بدی در $r \rightarrow 0$ از خود نشان می‌دهند. در $r \rightarrow \infty$ توابع GTO در مقایسه با توابع STO بسیار تندتر کاهش می‌یابند. وجود این-گونه رفتارها در توابع GTO باعث می‌شود تا برای به دست آوردن انرژی دقیق نماهای غیر فیزیکی بزرگ لازم باشد تا در $r \rightarrow 0$ رفتار بهبود یابد. همچنین اعداد کوانتومی l را باید بزرگتر از اندازه‌ی l در اوربیتال‌های اتمی گزید. جدای از این رفتارها توابع GTO کاربرد زیادی در فیزیک محاسباتی دارند.

۹-۵ تابع جسترو

وجود تابع جسترو ابتدایی‌ترین روشی است که برهم‌کنش الکترونی را در محاسبات QMC در نظر می‌گیرد. شاید با توجه به معادلات هارتری-فوک که سهم تبادل الکترونی را به خوبی بیان می‌کنند، وجود تابع جسترو در محاسبات بایسته به نظر نرسد، اما معمولاً یک تابع موج هارتری-فوک نیز به تنهایی برای به دست آوردن جواب‌های دقیق کافی نیست. یک تابع جسترو خوب باید برهم‌کنش الکترونی را به درستی توصیف کند. همچنین شکل آن به گونه‌ای باشد که سرعت محاسبات را در فرایند بهینه‌سازی کند نکند. بنابراین تابع جسترو نقشی کلیدی در اثرات برهم‌کنش الکترونی در دستگاه‌های بس‌ذره‌ای ایفا می‌کند [۳۰].

در این بخش سه شکل عامل جسترو معرفی می‌شود. مطالعه‌ی کارایی شکل‌های متفاوت جسترو در دستگاه‌های گوناگون به شوند(دلیل) هزینه‌ی محاسباتی این امر، دشوار به نظر می‌آید. الکساندر^{۷۵} و کولدول^{۷۶} [۳۱]، ۱۱۸ توابع جسترو متفاوتی را برای اتم‌های هلیوم، لیتیم و برلیم بررسی کرده‌اند، اما برای دیگر اتم‌ها یا دستگاه‌های مولکولی و پیوسته بررسی انجام نگرفته است. اگرچه می‌توان به کارایی هر شکل از تابع جسترو با بررسی دقت انرژی که برای دستگاه مورد نظر به دست می‌آید پی برد. افزایش پارامترها در تابع جسترو معمولاً دقت محاسبات را افزایش می‌دهد. از طرفی با افزایش پارامترها زمان محاسبات نیز افزایش می‌یابد، از این رو در محاسبات VMC بسته به دستگاه مورد بررسی و دقت مورد انتظار می‌توان تابع جسترو را گزید.

⁷⁵Alexander

⁷⁶Coldwell

۹-۵-۱ تابع جسترو دو ذره‌ای

نخستین شکل تابع جسترو تنها به فاصله نسبی دو الکترون بستگی دارد و وابسته از موقعیت الکترون‌ها از هسته است. از این‌رو به تابع جسترو دو ذره‌ای موسوم است و توسط جسترو-پاده^{۷۷} [۳۲] در محاسبات QMC به صورت زیر مطرح شد:

$$J_2(r_1, \dots, r_N) = \exp\left(\sum_{i < j}^N u(r_{ij})\right) \quad (۹۶)$$

که $u(r_{ij})$ تنها به فاصله نسبی دو الکترون، $r_{ij} = |\vec{r}_i - \vec{r}_j|$ ، بستگی دارد و به صورت زیر مشخص می‌شود:

$$u(r_{ij}) = \frac{ar_{ij}}{1 + br_{ij}} \quad (۹۷)$$

پارامتر a به کمک شرایط کاسپ الکترون-الکترون (که در بخش ۹-۶-۲ بررسی می‌شود) تعیین می‌شود و برای الکترون‌های با اسپین همسو $\frac{1}{4}$ و برای اسپین‌های ناهمسو $\frac{1}{2}$ است. از این‌رو پارامتر b به عنوان تنها پارامتری است که باید در تابع جسترو وردش یابد.

۹-۵-۲ تابع جسترو اسمیت - ماسکویتز

یک شکل پرکاربرد دیگر برای تابع جسترو توسط اسمیت و ماسکویتز^{۷۸} پیشنهاد شد [۳۳] که افزون بر فاصله نسبی دو الکترون به فاصله هر الکترون از هسته نیز بستگی دارد. شکل این تابع به صورت زیر است:

$$\Psi_J = \exp\left(\sum_{a=1}^{N_{nuc}} \sum_{i < j}^{N_{ele}} U_{aij}\right) \quad (۹۸)$$

که

$$U_{aij} = \sum_{k=1}^{N_a} C_{ka} \left(\bar{r}_{ia}^{m_{ka}} \bar{r}_{ja}^{n_{ka}} + \bar{r}_{ja}^{m_{ka}} \bar{r}_{ia}^{n_{ka}} \right) \bar{r}_{ij}^{o_{ka}}, \quad (۹۹)$$

$$\bar{r}_{ia} = \frac{r_{ia}}{1 + r_{ia}}, \quad \bar{r}_{ij} = \frac{r_{ij}}{1 + r_{ij}}$$

در رابطه‌ی ۹۸، N_{nuc} و N_{ele} به ترتیب شمار الکترون‌ها و شمار هسته‌ها را نشان می‌دهد. در رابطه‌ی (۹۹)، C_{ka} پارامترهای وردش، r_{ij} فاصله نسبی دو الکترون و r_{ia} فاصله الکترون i از هسته‌ی a است. N_a شمار جملاتی است که بسته به دقت مورد نظر در بسط U_{aij} اطراف هسته به کار می‌رود. $\{m, n, o\}$ مجموعه اعداد درست سه‌تایی هستند که مقادیر زیر را می‌توانند داشته باشند:

^{۷۷}Pade - Jastrow

^{۷۸}Schmidt and Moskowitz

$$\{1,0,0\}, \{0,0,1\}, \{0,0,2\}, \{0,0,3\}, \{0,0,4\}, \\ \{2,0,0\}, \{3,0,0\}, \{4,0,0\}, \{2,2,0\}, \{2,0,2\} \quad (100)$$

با در نظر گرفتن بسط رابطه‌ی ۹۹ روی سه‌تایی‌های ممکن، این تابع معمولاً شامل چندین جمله می‌شود، بنابراین شمار پارامترهای زیادی برای وردش دارد. اگر از سه‌تایی $\{0,0,1\}$ در رابطه‌ی ۹۹ استفاده شود، شکل این تابع جسترو به شکل ساده‌ی جسترو دو ذره‌ای تبدیل می‌شود و در این حالت C_{ka} ها با توجه به شرایط کاسپ مربوط به جسترو دو ذره‌ای تعیین می‌گردند. این تابع جسترو همبستگی الکترونی را نسبت به تابع جسترو ساده-ی دو ذره‌ای (رابطه‌ی ۹۶) بسیار بهتر توصیف می‌کند. در تابع جسترو ۹۶، تنها در جمله‌ی مربوط به سه‌تایی $\{0,0,1\}$ شرایط کاسپ الکترون-الکترون اعمال می‌شود و در باقی جملات نیازی به اعمال این شرایط نیست.

۳-۵-۹ تابع جسترو میتاس

میتاس^{۷۹} و همکارانش شکل مناسب دیگری برای تابع جسترو پیشنهاد کردند و آن را در بهینه‌سازی تابع موج دستگاه‌های اتمی [۳۴]، کربن مولکولی [۳۵ و ۳۶] و سیلیکون [۳۷] به کار بردند که به صورت زیر است:

$$\Psi_J = \exp \left(\sum_I \sum_i \sum_{j \neq i}^N u(r_{il}, r_{jl}, r_{ij}) \right) \quad (101)$$

که

$$u(r_{il}, r_{jl}, r_{ij}) = -\frac{c}{\gamma} e^{-\gamma r_{ij}} + \sum_{k < l, m} \left[a_k(r_{il}) a_l(r_{jl}) + a_k(r_{jl}) a_l(r_{il}) \right] b_m(r_{ij}), \\ a_k(r) = \left(\frac{\alpha_k r}{1 + \alpha_k r} \right)^2, \quad \alpha_k = \frac{\alpha_0}{2^{k-1}}, \quad k > 0 \quad (102) \\ b_m(r) = \left(\frac{\beta_m r}{1 + \beta_m r} \right)^2, \quad \beta_m = \frac{\beta_0}{2^{m-1}}, \quad m > 0$$

N شمار الکترون‌ها، r_{il} فاصله‌ی میان i امین الکترون تا l امین هسته و r_{ij} فاصله‌ی میان دو الکترون i ام و j ام است.

در رابطه‌ی ۱۰۱، $a_0(r) = b_0(r) = 1$. این تابع جسترو شرایط کاسپ الکترون-الکترون را برآورده می‌کند و کاسپ الکترون-هسته ندارد. اگرچه کاربرد این تابع در برخی محاسبات موفقیت آمیز بوده است اما از آنجا که تابع همبستگی $u(r_{il}, r_{jl}, r_{ij})$ بلند برد است و شامل جمع روی تمام الکترون‌ها و یون‌ها است، در نتیجه با افزایش اندازه دستگاه زمان محاسبات نیز افزایش می‌یابد.

۹-۶ شرایط کاسپ

برای اینکه تابع موجی با واریانس بسیار کم از انرژی موضعی به دست آوریم باید توابع موج آزمایشی یک مجموعه شرایط موسوم به شرایط کاسپ^{۸۰} را برآورده کنند. معادله-ی شرودینگر وابسته به زمان برای یک دستگاه کولمبی N ذره‌ای را به صورت زیر در نظر می‌گیریم ($\hbar = 1$):

$$\left[-\frac{1}{2} \sum_{i=1}^N \frac{1}{m_i} \nabla_i^2 + \sum_{i<j}^N \frac{q_i q_j}{r_{ij}} \right] \Psi = E \Psi \quad (۱۰۳)$$

که m_i جرم ذره‌ی i ام و r_{ij} فاصله‌ی میان ذرات i ام و j ام است. q_i و q_j به ترتیب بارهای ذرات i ام و j ام می‌باشد. اگر دو ذره به سوی یکدیگر میل کنند ($r_{ij} \rightarrow 0$)، بخش مربوط به انرژی پتانسیل در معادله‌ی (۱۰۳) واگرا می‌شود. از این رو برای اینکه انرژی کل E ، بی‌نهایت نشود همگرایی در انرژی جنبشی باید دقیقاً همگرایی انرژی پتانسیل را خنثی کند. برآورده شدن شرایط کاسپ این امر را محقق می‌سازد. کاتو^{۸۱} [۳۸] برای نخستین بار این شرایط را برای دستگاه‌های کولمبی به صورت زیر به دست آورد:

$$\lim_{r_{ij} \rightarrow 0} \frac{\partial \bar{\Psi}}{\partial r_{ij}} \bigg|_{r_{ij}=0} = \frac{\mu_{ij} q_i q_j}{l+1} \lim_{r_{ij} \rightarrow 0} \Psi \quad (۱۰۴)$$

که $\bar{\Psi}$ ، میانگین توابع موج Ψ روی کره‌ی بی‌نهایت کوچک حول $r_{ij} = 0$ ، $\mu_{ij} = m_i m_j / (m_i + m_j)$ ، جرم کاهش یافته‌ی دو ذره‌ی i ام و j ام است و l از تقارن تابع موج ناشی می‌شود (تابع موج برای یک دستگاه فرمیونی پاد متقارن است). l برای فرمیون‌های یکسان 1 و برای دیگر فرمیون‌ها 0 است [۳۹].

۹-۶-۱ شرایط کاسپ الکترون - الکترون

در رابطه‌ی ۱۰۳، l برای فرمیون‌هایی که اسپین و مکان یکسان دارند یک و برای دیگر فرمیون‌ها صفر است [۳۸ و ۳۹]. از این رو شرایط کاسپ الکترون - الکترون برای الکترون‌های با اسپین همسو و الکترون‌های با اسپین پادهمسو متفاوت است. در تابع موج آزمایشی، $\Psi_{trial} = D^\dagger D^\downarrow e^J$ ، تابع جسترو، e^J ، سهم برهم‌کنش الکترونی را بر عهده دارد. بنابراین شرایط کاسپ الکترون - الکترون روی این تابع اعمال می‌شود و با استفاده از رابطه‌ی ۱۰۴ به صورت زیر به دست می‌آید:

⁸⁰Cusp Conditions⁸¹Kato

$$\left. \frac{\partial \exp(J)}{\partial r_{ij}} \right|_{r_{ij}=0} = \begin{cases} \frac{1}{4} & \text{for } \uparrow\uparrow \text{ or } \downarrow\downarrow \\ \frac{1}{2} & \text{for } \uparrow\downarrow \text{ or } \downarrow\uparrow \end{cases} \quad (۱۰۵)$$

در رابطه‌ی ۱۰۵، علامت‌های $\uparrow$ و $\downarrow$ مشخص‌کننده‌ی اسپین الکترون‌هاست. به عنوان نمونه اعمال شرایط کاسپ روی تابع جسترو دو ذره‌ای، رابطه‌ی ۹۶، پارامتر a را به صورت زیر مشخص می‌کند:

$$\left. \frac{\partial(e^J)}{\partial r_{ij}} \right|_{r_{ij}=0} = \left. \frac{\partial \left(\frac{ar_{ij}}{1+br_{ij}} \right)}{\partial r_{ij}} \right|_{r_{ij}=0} = a = \begin{cases} \frac{1}{4} & \text{for } \uparrow\uparrow \text{ or } \downarrow\downarrow \\ \frac{1}{2} & \text{for } \uparrow\downarrow \text{ or } \downarrow\uparrow \end{cases} \quad (۱۰۶)$$

۹-۶-۲ شرایط کاسپ الکترون- هسته

اینک با استفاده از معادله‌ی ۱۰۴ شرایط کاسپ الکترون- هسته را می‌توانیم به صورت زیر به دست آوریم:

$$\left. \frac{\partial \bar{\Psi}}{\partial r_{ij}} \right|_{r_{ij}=0} = \frac{m_e m_p}{m_e + m_p} \frac{-Ze^2}{l+1} \bar{\Psi}(r_{ij}=0) \quad (۱۰۷)$$

از آنجا که جرم پروتون بسیار بیشتر از جرم الکترون است ($m_p \gg m_e$)، در واحدهای اتمی ($e = \hbar = m = 1$)، رابطه‌ی ۱۰۷ به صورت زیر ساده می‌شود:

$$\left. \frac{\partial \bar{\Psi}}{\partial r_{ij}} \right|_{r_{ij}=0} = -Z \bar{\Psi}(r_{ij}=0) \quad (۱۰۸)$$

که Z عدد اتمی هسته است. در تابع موج محاسبات QMC دترمینان‌های اسلیتر سهم برهم‌کنش الکترون-هسته را نشان می‌دهند. از این‌رو در رابطه‌ی ۱۰۸، شرایط کاسپ الکترون-هسته تنها روی دترمینان‌های اسلیتر اعمال می‌شود:

$$\left. \frac{\partial D}{\partial r_i} \right|_{r_i=0} = -Z D(r_i=0) \quad (۱۰۹)$$

که r_i فاصله‌ی الکترون از هسته است. رابطه‌ی ۱۰۹ را می‌توان به صورت زیر نوشت:

$$\frac{\begin{vmatrix} \varphi_1(r_1) & \cdots & \frac{\partial \varphi_1(r_i)}{\partial r_i} & \cdots & \varphi_1(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_i(r_1) & \cdots & \frac{\partial \varphi_i(r_i)}{\partial r_i} & \cdots & \varphi_i(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_N(r_1) & \cdots & \frac{\partial \varphi_N(r_i)}{\partial r_i} & \cdots & \varphi_N(r_N) \\ \vdots & & \vdots & & \vdots \end{vmatrix}}{\begin{vmatrix} \varphi_1(r_1) & \cdots & \varphi_1(r_i) & \cdots & \varphi_1(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_i(r_1) & \cdots & \varphi_i(r_i) & \cdots & \varphi_i(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_N(r_1) & \cdots & \varphi_N(r_i) & \cdots & \varphi_N(r_N) \\ \vdots & & \vdots & & \vdots \end{vmatrix}}_{r_i=0} = -Z \quad (110)$$

وقتی $r_i \rightarrow 0$ ، دترمینان D نباید صفر شود. در رابطه‌ی بالا دترمینان‌های صورت و مخرج را می‌توان به صورت زیر باز کرد:

$$\frac{\sum_{i,j} \frac{\partial \varphi_{ij}(r_i=0)}{\partial r_i} M_{ij}(r_1, \dots, r_i=0, \dots, r_N)}{\sum_{i,j} \varphi_{ij}(r_i=0) M_{ij}(r_1, \dots, r_i=0, \dots, r_N)} = -Z \quad (111)$$

که M_{ij} ، ماتریس همسازهی درایه‌ی i ام و j ام است. رابطه‌ی ۱۱۱، تنها هنگامی برقرار است که هر اوربیتال اتمی در دترمینان‌های اسلیتر در $r_i \rightarrow 0$ ، به یک عدد ثابت میل کند:

$$\left. \frac{1}{\varphi_k} \frac{\partial \varphi_k}{\partial r_i} \right|_{r_i=0} = -Z \quad (112)$$

مگر اینکه کل دترمینان برابر صفر شود. با توجه به رابطه‌ی به دست‌آمده شرایط کاسپ الکترون-هسته روی اوربیتال‌های اتمی اعمال می‌شود. در واقع اعمال شرایط کاسپ الکترون-هسته در الگوریتم VMC قیدی روی پارامترهای ورودش اعمال می‌کند.

۱۰ برخی مثال‌ها

در مرحله نخست به ساخت تابع موج اتمها و ملکولهای ساده می پردازیم تا با استفاده از آنها به محاسبات مونت کارلوی کوآنتومی وردشی پردازیم. در مراحل بعدی روش کمینه سازی خود را برای این سیستمها امتحان خواهیم کرد.

۱-۱۰ تابع موج اتم‌های بس الکترونی

برای شروع محاسبات بهتر است با اتمها آغاز کنیم. در این صورت تعدادی الکترون اسپین بالا و بقیه اسپین پایین دارند. تابع جسترو هم فرم جفتی خود را دارد که خواهیم گفت.



$$\Psi(\vec{x}_1, \dots, \vec{x}_N) = \begin{vmatrix} \Phi 1s(\vec{x}_1 \uparrow) & \Phi 1s(\vec{x}_2 \uparrow) & \dots & \Phi 1s(\vec{x}_{Nep} \uparrow) & \Phi 1s(\vec{x}_{Nep+1} \downarrow) & \Phi 1s(\vec{x}_{Nep+2} \downarrow) & \dots & \Phi 1s(\vec{x}_N \downarrow) \\ \Phi 2s(\vec{x}_1 \uparrow) & \Phi 2s(\vec{x}_2 \uparrow) & \dots & \Phi 2s(\vec{x}_{Nep} \uparrow) & \Phi 2s(\vec{x}_{Nep+1} \downarrow) & \Phi 2s(\vec{x}_{Nep+2} \downarrow) & \dots & \Phi 2s(\vec{x}_N \downarrow) \\ \Phi px(\vec{x}_1 \uparrow) & \Phi px(\vec{x}_2 \uparrow) & \dots & \Phi px(\vec{x}_{Nep} \uparrow) & \Phi px(\vec{x}_{Nep+1} \downarrow) & \Phi px(\vec{x}_{Nep+2} \downarrow) & \dots & \Phi px(\vec{x}_N \downarrow) \\ \Phi py(\vec{x}_1 \uparrow) & \Phi py(\vec{x}_2 \uparrow) & \dots & \Phi py(\vec{x}_{Nep} \uparrow) & \Phi py(\vec{x}_{Nep+1} \downarrow) & \Phi py(\vec{x}_{Nep+2} \downarrow) & \dots & \Phi py(\vec{x}_N \downarrow) \\ \Phi pz(\vec{x}_1 \uparrow) & \Phi pz(\vec{x}_2 \uparrow) & \dots & \Phi pz(\vec{x}_{Nep} \uparrow) & \Phi pz(\vec{x}_{Nep+1} \downarrow) & \Phi pz(\vec{x}_{Nep+2} \downarrow) & \dots & \Phi pz(\vec{x}_N \downarrow) \\ \dots & & & & \dots & & & \dots \end{vmatrix} \exp(J(\{\vec{x}_{ij}\}))$$

where

$$J = \left(\sum_{i < j} u(r_{ij} \uparrow \uparrow \text{ or } \downarrow \downarrow) \right) + \left(\sum_{i < j} u(r_{ij} \uparrow \downarrow \text{ or } \downarrow \uparrow) \right)$$

$$u(r_{ij}) = \frac{ar_{ij}}{1 + br_{ij}} \quad a = \frac{1}{4} \uparrow \uparrow \text{ or } \downarrow \downarrow, \quad a = \frac{1}{2} \uparrow \downarrow \text{ or } \downarrow \uparrow$$

b is a free parameter usually is around 0.2.

تابع جسترو بالا ساده ترین نوع جفتی این نوع توابع است. بسیاری اوقات باید از سه ذره ای استفاده شود. هریک از اربیتالهای s به دو صورت و اربیتالهای p به یک صورت بتوانند باشند. دومی بیشتر مناسب سیستمهای با شبه پتانسیل می باشد که الکترون مغزی نداریم.

$$s - orbital \left\{ \begin{array}{l} \Phi n_s(r_i) = \left\{ \sum_{k=1}^M C_k r^{n_s-1} \exp(-\zeta_k r_i) \right\} \quad STO \\ or \\ \Phi s(r_i) = \left\{ \sum_{k=1}^M C_k \exp(-\zeta_k r_i^2) \right\} \quad GTO \end{array} \right.$$

n_s عدد کوانتومی اصلی است. در بسیاری موارد برای اربیتالهای با تقارن گروهی S مخصوصاً حالت $n=2$ تابعی به صورت زیر مناسب شناخته شده است.

$$\phi_s = \sum_{k=1}^{M1} C_{1k} \exp(-\zeta_{1k} r) + \sum_{k=1}^{M2} C_{2k} r \exp(-\zeta_{2k} r)$$

اگر از هارمونیکهای گروهی استفاده شود برای اربیتالهای p داریم با توجه به تقارن راستایی داریم:

$$p - orbital \left\{ \begin{array}{l} \Phi p x_j(\vec{x}_i) = \left\{ \sum_{k=1}^M C'_k x_j \exp(-\zeta'_k r_i) \right\} \quad STO \\ or \\ \Phi p x_j(\vec{x}_i) = \left\{ \sum_{k=1}^M C'_k x_j \exp(-\zeta'_k r_i^2) \right\} \quad GTO \end{array} \right. ; x_j = \{x, y, z\}$$

$$Y_n^m(\theta, \varphi) \equiv (-1)^m \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{im\varphi}$$

تعداد پارامترهای آزاد جسترو دو ذره ای همیشه ۱ بهترین بوده است. در حالتی که اتمی N الکترونی باشد (مثل هلیوم-بریلیوم-نئون-منیزیوم+آرگون....) تعداد کل پارامترها به کمک تعداد اربیتالهای غیر مشابه قابل محاسبه است. اکنون مشتقات زیر برای محاسبات مونت کارلوی کوانتومی مورد نیاز است:

$$\left. \begin{array}{l} T_i = -\frac{1}{4} \nabla_i^2 \ln \Psi \\ F_i = \frac{1}{\sqrt{2}} \nabla_i \ln \Psi \end{array} \right\} \quad KE_i = 2T_i - F_i^2 = -\frac{1}{2} \frac{\nabla_i^2 \Psi}{\Psi}$$

$$\left\{ \begin{array}{l} \nabla_i \ln \Psi = \nabla_i \ln D^\dagger + \nabla_i \ln D^\downarrow + \nabla_i \ln [\exp(J)] \\ = \frac{1}{D^\dagger} \nabla_i D^\dagger + \frac{1}{D^\downarrow} \nabla_i D^\downarrow + \nabla_i J, \\ \nabla_i^2 \ln \Psi = -\left(\frac{1}{D^\dagger} \nabla_i D^\dagger \right)^2 + \frac{1}{D^\dagger} \nabla_i^2 D^\dagger - \left(\frac{1}{D^\downarrow} \nabla_i D^\downarrow \right)^2 + \frac{1}{D^\downarrow} \nabla_i^2 D^\downarrow + \nabla_i^2 J \end{array} \right.$$

در واقع در مونت کارلوی وردشی ما دنبال محاسبه انرژی موضعی هستیم (Local Energy):

$$\frac{\mathcal{H}\Psi}{\Psi} = -\frac{1}{2} \sum_i \frac{\nabla_i^2 \Psi}{\Psi} + V \quad \text{و } V \text{ پتانسیل کولنی است (} V = \sum_i \frac{1}{r_i} + \sum_{i<j} \frac{1}{r_{ij}} \text{) که این جمله در}$$

پایان محاسبات خیلی ساده اضافه می شود. و نیازی به کار روی آن نداریم. مشتقات اصلی مورد نیاز به شکل زیر هستند.

$$\vec{\nabla}_i D \uparrow = \begin{vmatrix} \Phi 1s(\vec{x}_1 \uparrow) \dots \vec{\nabla}_i \Phi 1s(\vec{x}_i \uparrow) \dots \Phi 1s(\vec{x}_{Nep} \uparrow) \\ \Phi 2s(\vec{x}_1 \uparrow) \dots \vec{\nabla}_i \Phi 2s(\vec{x}_i \uparrow) \dots \Phi 2s(\vec{x}_{Nep} \uparrow) \\ \Phi px(\vec{x}_1 \uparrow) \dots \vec{\nabla}_i \Phi px(\vec{x}_i \uparrow) \dots \Phi px(\vec{x}_{Nep} \uparrow) \\ \Phi py(\vec{x}_1 \uparrow) \dots \vec{\nabla}_i \Phi py(\vec{x}_i \uparrow) \dots \Phi py(\vec{x}_{Nep} \uparrow) \\ \Phi pz(\vec{x}_1 \uparrow) \dots \vec{\nabla}_i \Phi pz(\vec{x}_i \uparrow) \dots \Phi pz(\vec{x}_{Nep} \uparrow) \\ \dots \end{vmatrix} \quad \nabla_i^2 D \uparrow = \begin{vmatrix} \Phi 1s(\vec{x}_1 \uparrow) \dots \nabla_i^2 \Phi 1s(\vec{x}_i \uparrow) \dots \Phi 1s(\vec{x}_{Nep} \uparrow) \\ \Phi 2s(\vec{x}_1 \uparrow) \dots \nabla_i^2 \Phi 2s(\vec{x}_i \uparrow) \dots \Phi 2s(\vec{x}_{Nep} \uparrow) \\ \Phi px(\vec{x}_1 \uparrow) \dots \nabla_i^2 \Phi px(\vec{x}_i \uparrow) \dots \Phi px(\vec{x}_{Nep} \uparrow) \\ \Phi py(\vec{x}_1 \uparrow) \dots \nabla_i^2 \Phi py(\vec{x}_i \uparrow) \dots \Phi py(\vec{x}_{Nep} \uparrow) \\ \Phi pz(\vec{x}_1 \uparrow) \dots \nabla_i^2 \Phi pz(\vec{x}_i \uparrow) \dots \Phi pz(\vec{x}_{Nep} \uparrow) \\ \dots \end{vmatrix}$$

برای اسپین پائین هم عیناً همینطور است. در مورد جمله جسترو نیز داریم:

$$\vec{\nabla}_k \exp(J) = \vec{\nabla}_k \left[\sum_{j>k} \exp[\mu(r_{kj})] \right] + \vec{\nabla}_k \left[\sum_{j=1}^{k-1} \exp[\mu(r_{jk})] \right]$$

$$\nabla_k^2 \exp(J) = \nabla_k^2 \left[\sum_{j>k} \exp[\mu(r_{kj})] \right] + \nabla_k^2 \left[\sum_{j=1}^{k-1} \exp[\mu(r_{jk})] \right]$$

در این مشتق‌ها توجه می کنیم که:

$$\left\{ \begin{aligned} \vec{\nabla}_k(\dots) &= \frac{\partial(\dots)}{\partial x_k} \hat{i} + \frac{\partial(\dots)}{\partial y_k} \hat{j} + \frac{\partial(\dots)}{\partial z_k} \hat{k} = \frac{\partial(\dots)}{\partial r_{kj}} \left[\frac{\partial r_{kj}}{\partial x_k} \hat{i} + \frac{\partial r_{kj}}{\partial y_k} \hat{j} + \frac{\partial r_{kj}}{\partial z_k} \hat{k} \right] = \frac{\partial(\dots)}{\partial r_{kj}} \hat{r}_{jk} \\ r_{kj} &= \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2 + (z_j - z_k)^2} \end{aligned} \right.$$

$$\text{for each term we have: } \vec{\nabla}_k [\exp(\mu(r_{kj}))] = \frac{\partial(\exp(\mu(r_{kj})))}{\partial r_{kj}} \hat{r}_{jk}$$

$$= \exp(\mu(r_{kj})) \frac{a}{(1 + br_{kj})^2} \hat{r}_{jk}$$

$$\left\{ \nabla_k^2(\dots) = \frac{1}{r_{kj}^2} \frac{\partial}{\partial r_{kj}} \left[r_{kj}^2 \frac{\partial(\dots)}{\partial r_{kj}} \right] = \frac{\partial^2(\dots)}{\partial r_{kj}^2} + \frac{2}{r_{kj}} \frac{\partial(\dots)}{\partial r_{kj}} \right.$$

$$\left. \nabla_k^2 [\exp(\mu(r_{kj}))] = \exp(\mu(r_{kj})) \left\{ \left[\frac{-2ab}{(1 + br_{kj})^3} \right] + \frac{2}{r_{kj}} \frac{a}{(1 + br_{kj})^2} \right\} \right.$$

برای مثال می‌توان نوشت:

$$\vec{\nabla}_j[\phi_s(r_j)] = \begin{cases} n=1 & \frac{\partial(\phi_s)}{\partial r_j} \hat{r}_j = \sum_{k=1}^M -\varsigma_k C_k r_j^{n_s-1} \exp(-\varsigma_k r_j) \frac{\vec{r}_j}{r_j} \quad STO \\ n=2 & \frac{\partial(\phi_s)}{\partial r_j} \hat{r}_j = \sum_{k=1}^M [1-\varsigma_k] r_j C_k \exp(-\varsigma_k r_j) \frac{\vec{r}_j}{r_j} \quad STO \\ & \frac{\partial(\phi_s)}{\partial r_j} \hat{r}_j = \sum_{k=1}^M -2r_j \varsigma_k C_k \exp(-\varsigma_k r_j^2) \frac{\vec{r}_j}{r_j} \quad GTO \end{cases}$$

شرایط کاسپ تنها برای اربیتالهای s با تقارن کروی اعمال می‌گردد زیرا تقارن کروی دارند و میتوان در مختصات کروی با آنها کار کرد:

$$-\frac{1}{2\mu_{ij}} \nabla^2 \Psi + \mathcal{V}(r) \sim -\frac{1}{2\mu_{ij}} \Psi'' - \frac{1}{\mu_{ij} r} \Psi' + \mathcal{V}(r) \\ \sim -\frac{1}{\mu_{ij} r} \Psi' + \mathcal{V}(r) \Bigg|$$

$$\mu_{ij} = m_i m_j / (m_i + m_j)$$

$$-\frac{1}{\mu_{ij}} \frac{1}{r} \Psi' + \mathcal{V}(r) = -\frac{1}{\mu_{ij}} \frac{1}{r} \Psi' + \frac{q_i q_j}{r} = \text{finite} \\ \left. \frac{\partial \hat{\Psi}}{\partial r_{ij}} \right|_{r_{ij}=0} = \mu_{ij} q_i q_j \Psi(r_{ij}=0) \quad \begin{matrix} e-N \\ \Theta-\Theta \end{matrix} \quad \left| \begin{matrix} \Psi' \\ \Psi \end{matrix} \right|_{r=0} = \begin{matrix} -Z \\ 1/2 \end{matrix}$$

$$n=1 \longrightarrow \left. \frac{\partial \phi_s}{\partial r} \right|_{r=0} = \left. \frac{\partial \left[\sum_{k=1}^M C_k \exp(-\varsigma_k r) \right]}{\partial r} \right|_{r=0} = -Z \phi_s(0)$$

Where $Z = \text{Atomic number}$

$$C_1 = \frac{\sum_{k=2}^M C_k (Z - \varsigma_k)}{\varsigma_1 - Z}$$

اگر تابع به صورت $\phi_s = \sum_{k=1}^{M/2} C_k \exp(-\zeta_k r) + \sum_{k=M/2+1}^M C_k r \exp(-\zeta_k r)$ تعریف می‌شد شرط فوق به

صورت روبرو بود

$$C_1 = \frac{\sum_{k=2}^{M-1} C_k (Z - \zeta_k) + \sum_{k=M/2+1}^M C_k}{\zeta_1 - Z}$$

شرط بهنجارش را با تقسیم سایر پارامترهای ضربی بر بزرگترین عدد به دست می‌آوریم:

$$C_i = \frac{C_i}{\text{Max}[C]} \quad i = 1, M$$

۱۰-۲ تابع موج مولکول‌های $\text{H}_2\text{O}, \text{Li}_2, \text{H}_2$

هنوز برخی از ملکولهای ساده دو اتمی با وجود روشهای دقیق محاسباتی طول پیوندهایشان در شک و تردید است. لذا برای شروع و امتحان کدهای محاسباتی خود برای چنین اتمهایی نمونه‌های ساده را می‌سازیم. مثلاً پیدا کردن طول پیوند ملکولهایی چون N_2O_2 یکی از نمونه‌هایی است که روشهای محاسباتی مختلف دقتهای گوناگونی را گزارش می‌کنند.

درمباحث زیر همه جا اندیس i, j برای الکترون‌ها m برای شماره اربیتالهای ملکولی و $\alpha, \beta, \dots$ برای هسته‌ها می‌باشند. هامیلتونی یک سیستم ملکولی به شکل زیر است:

$$\hat{H} = \sum_i -\frac{1}{2} \nabla_i^2 + \sum_i \sum_{\alpha} \frac{Z_{\alpha}}{|\vec{x}_i - \vec{d}_{\alpha}|} + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{|\vec{x}_i - \vec{x}_j|}$$

که Z_α بار هسته α ام است. و $\vec{x}_i$ محل الکترون i و $\vec{d}_\alpha$ محل هسته α است. اگر ملکول $N=N_{up}+N_{down}$ الکترون داشته باشد تابع موج به صورت زیر است:

$$\Psi(\vec{x}_1, \vec{x}_N, \vec{d}_1, \vec{d}_2, \dots) = \begin{vmatrix} \Phi 1(\vec{x}_1 \uparrow) & \Phi 1(\vec{x}_2 \uparrow) & \dots & \Phi 1(\vec{x}_{N_{up}} \uparrow) & \Phi 1(\vec{x}_{N_{up}+1} \uparrow) & \Phi 1(\vec{x}_{N_{up}+2} \uparrow) & \dots & \Phi 1(\vec{x}_N \uparrow) \\ \Phi 2(\vec{x}_1 \uparrow) & \Phi 2(\vec{x}_2 \uparrow) & \dots & & \Phi 2(\vec{x}_{N_{up}+1} \uparrow) & \Phi 2(\vec{x}_{N_{up}+2} \uparrow) & \dots & \\ \dots & & & & \dots & & & \\ \Phi m(\vec{x}_1 \uparrow) & \Phi m(\vec{x}_1 \uparrow) & \dots & & \Phi m(\vec{x}_{N_{up}+1} \uparrow) & \Phi m(\vec{x}_{N_{up}+2} \uparrow) & \dots & \\ \dots & & & & \dots & & & \end{vmatrix} \exp(J(\{\vec{x}_{ij}\}))$$

where

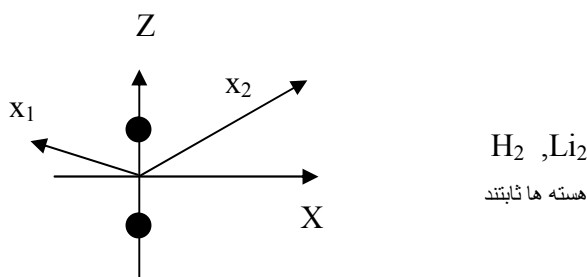
$$J = \left(\sum_{i < j} u(r_{ij} \uparrow \uparrow \text{ or } \downarrow \downarrow) \right) + \left(\sum_{i < j} u(r_{ij} \uparrow \downarrow \text{ or } \downarrow \uparrow) \right)$$

$$u(r_{ij}) = \frac{ar_{ij}}{1 + br_{ij}} \quad a = \frac{1}{4} \uparrow \uparrow \text{ or } \downarrow \downarrow, \quad a = \frac{1}{2} \uparrow \downarrow \text{ or } \downarrow \uparrow$$

در رابطه اخیر هر یک از اربیتالهای ملکولی به شکل زیر تعریف می شوند:

$$\Phi_m(i) = \Phi_m(\vec{x}_i, \{\vec{d}\}) = \sum_k \sum_\alpha C_{k,\alpha} r_{i,\alpha}^{n_{k,\alpha}-1} \exp(-\zeta_{k,\alpha} r_{i,\alpha}) Y_{l_{k,\alpha}, m_{k,\alpha}}(\hat{x}_{i,\alpha}) \quad \text{where} \quad r_{i,\alpha} = |\vec{x}_i - \vec{d}_\alpha|$$

که بسط روی یک مجموعه توابع پایه است یا بهتر همان اربیتالهای اتمی که در بحث اتمها ذکر شد. برای ملکول دو اتمی شکل زیر را در نظر میگیریم:



۱۰-۲-۱ مولکول هیدروژن

فاصله دو هسته $1/40$ و دو هسته روی محور Z باشند: $\begin{cases} \vec{d}_1 = (0, 0, 0.7) \\ \vec{d}_2 = (0, 0, -0.7) \end{cases}$ تعداد ۲ الکترون وجود یکی

اسپین بالا و دیگری اسپین پائین دارد. ما نیاز به یک اربیتال ملکولی داریم $\Phi_m(i) = \Phi_1(\vec{x}_i, \vec{d}_1, \vec{d}_2)$

چون اتمهای سازنده تنها اربیتال S داشته اند و در راستای Z اربیتالهای p_z با هم تداخل احتمالی خواهند داشت پس برای هر یک از دو الکترون داریم :

$$\Phi_m(i) = \Phi_1(\vec{x}_i, \vec{d}_1, \vec{d}_2) = \phi_{1s}(\vec{x}_i, \vec{d}_1) + \phi_{1s}(\vec{x}_i, \vec{d}_2) + \phi_{p_z}(\vec{x}_i, \vec{d}_1) + \phi_{p_z}(\vec{x}_i, \vec{d}_2)$$

چون حالت پایه اتمهای هیدروژن آزاد تنها با تعداد اندکی STO (اتم آزاد یک جمله کافیست) قابل بیانند ما برای اطمینان دو جمله در نظر میگیریم:

$$\phi_{1s}(\vec{x}_i, \vec{d}_\alpha) = C_{1,\alpha} \exp(-\zeta_1 r_{i,\alpha}) + C_{2,\alpha} \exp(-\zeta_2 r_{i,\alpha}) \quad \text{where} \quad r_{i,\alpha} = |\vec{x}_i - \vec{d}_\alpha|$$

و برای اربیتال دیگر تنها یکی کفایت می کند:

$$\phi_{p_z}(\vec{x}_i, \vec{d}_\alpha) = b_\alpha z_{i,\alpha} \exp(-\zeta' r_{i,\alpha}) \quad \text{where} \quad r_{i,\alpha} = |\vec{x}_i - \vec{d}_\alpha|$$

توجه کنید که در هر دو اربیتال بالا نما ها برای هر دو هسته یکسانند. اکنون تابع موج بسیار ساده است:

$$\Psi(\vec{x}_1, \vec{x}_2, \vec{d}_1, \vec{d}_2) = D(1) \uparrow D(2) \downarrow \exp(J)$$

$$D(\vec{x}_1, \vec{d}_1, \vec{d}_2) \uparrow = \Phi_m(1)$$

$$D(\vec{x}_2, \vec{d}_1, \vec{d}_2) \downarrow = \Phi_m(2)$$

$$J = \left(\frac{r_{12}}{2(1 + br_{12})} \right) \quad \text{where} \quad r_{12} = |\vec{x}_1 - \vec{x}_2|$$

$$\left. \begin{aligned} T_i &= -\frac{1}{4} \nabla_i^2 \ln \Psi \\ F_i &= \frac{1}{\sqrt{2}} \nabla_i \ln \Psi \end{aligned} \right\} \quad KE_i = 2T_i - F_i^2 = -\frac{1}{2} \frac{\nabla_i^2 \Psi}{\Psi}$$

$$\ln \Psi(\vec{x}_1, \vec{x}_2, \vec{d}_1, \vec{d}_2) = \ln[\Phi_m(1)] + \ln[\Phi_m(2)] + \frac{r_{12}}{2(1 + br_{12})}$$

$$\nabla_i \ln \Psi(\vec{x}_1, \vec{x}_2, \vec{d}_1, \vec{d}_2) = \frac{\nabla_i \Phi_m(i)}{\Phi_m(i)} \nabla_i \left[\frac{r_{12}}{2(1 + br_{12})} \right]$$

$$\nabla_i^2 \ln \Psi(\vec{x}_1, \vec{x}_2, \vec{d}_1, \vec{d}_2) = -\left[\frac{\nabla_i \Phi_m(i)}{\Phi_m(i)} \right]^2 + \frac{\nabla_i^2 \Phi_m(i)}{\Phi_m(i)} + \nabla_i^2 \left[\frac{r_{12}}{2(1 + br_{12})} \right]$$

شرایط کا بر روی هر دو هسته اعمال می شود.

۱۰-۲-۲ مولکول لیتیوم

اینجا ۶ الکترون داریم و به سه اربیتال ملکولی نیازمندیم در هر یک ۲ الکترون یکی اسپین بالا و دیگری پائین دارد. هر اربیتال ملکولی به صورت زیر است.

$$\Phi_m(i) = \Phi_m(\vec{x}_i, \vec{d}_1, \vec{d}_2) = \phi_{1s,m}(\vec{x}_i, \vec{d}_1) + \phi_{2s,m}(\vec{x}_i, \vec{d}_1) + \phi_{1s,m}(\vec{x}_i, \vec{d}_2) + \phi_{2s,m}(\vec{x}_i, \vec{d}_2) \\ + \phi_{p_z,m}(\vec{x}_i, \vec{d}_1) + \phi_{p_z,m}(\vec{x}_i, \vec{d}_2) \quad , \quad m = 1, 2, 3$$

$$\begin{cases} \phi_{1s,m}(\vec{x}_i, \vec{d}_\alpha) = C_{1,\alpha,m} \exp(-\zeta_{1,m} r_{i,\alpha}) \\ \phi_{2s,m}(\vec{x}_i, \vec{d}_\alpha) = C_{2,\alpha,m} r_{i,\alpha} \exp(-\zeta_{2,m} r_{i,\alpha}) \\ \phi_{p_z,m}(\vec{x}_i, \vec{d}_\alpha) = b_{m,\alpha} z_{i,\alpha} \exp(-\zeta_{p_z,m} r_{i,\alpha}) \end{cases} \quad \text{where} \quad r_{i,\alpha} = |\vec{x}_i - \vec{d}_\alpha|$$

دو باره نماها مستقل از شماره هسته اند. پس تابع موج خواهد بود:

$$\Psi(\vec{x}_1, \dots, \vec{x}_6, \vec{d}_1, \vec{d}_2) = D(1,2,3) \uparrow D(4,5,6) \downarrow \exp(J)$$

$$D \uparrow = \begin{vmatrix} \Phi_1(1) & \Phi_1(2) & \Phi_1(3) \\ \Phi_2(1) & \Phi_2(2) & \Phi_2(3) \\ \Phi_3(1) & \Phi_3(2) & \Phi_3(3) \end{vmatrix}$$

$$D \downarrow = \begin{vmatrix} \Phi_1(4) & \Phi_1(5) & \Phi_1(6) \\ \Phi_2(4) & \Phi_2(5) & \Phi_2(6) \\ \Phi_3(4) & \Phi_3(5) & \Phi_3(6) \end{vmatrix}$$

$$J = \left(\sum_{i < j} u(r_{ij}) \uparrow\uparrow \text{ or } \downarrow\downarrow \right) + \left(\sum_{i < j} u(r_{ij}) \uparrow\downarrow \text{ or } \downarrow\uparrow \right)$$

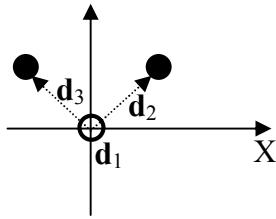
$$u(r_{ij}) = \frac{a r_{ij}}{1 + b r_{ij}} \quad a = \frac{1}{4} \uparrow\uparrow \text{ or } \downarrow\downarrow, \quad a = \frac{1}{2} \uparrow\downarrow \text{ or } \downarrow\uparrow \quad r_{ij} = |\vec{x}_i - \vec{x}_j|.$$

تمامی مشتقات مانند قبل:

$$\left. \begin{aligned} T_i &= -\frac{1}{4} \nabla_i^2 \ln \Psi \\ F_i &= \frac{1}{\sqrt{2}} \nabla_i \ln \Psi \end{aligned} \right\} \quad \text{KE}_i = 2T_i - F_i^2 = -\frac{1}{2} \frac{\nabla_i^2 \Psi}{\Psi}$$

$$\begin{cases} \nabla_i \ln \Psi = \nabla_i \ln D^\uparrow + \nabla_i \ln D^\downarrow + \nabla_i \ln [\exp(J)] \\ \quad = \frac{1}{D^\uparrow} \nabla_i D^\uparrow + \frac{1}{D^\downarrow} \nabla_i D^\downarrow + \nabla_i J, \\ \nabla_i^2 \ln \Psi = - \left(\frac{1}{D^\uparrow} \nabla_i D^\uparrow \right)^2 + \frac{1}{D^\uparrow} \nabla_i^2 D^\uparrow - \left(\frac{1}{D^\downarrow} \nabla_i D^\downarrow \right)^2 + \frac{1}{D^\downarrow} \nabla_i^2 D^\downarrow + \nabla_i^2 J \end{cases}$$

توجه شود همیشه، اگر الکترون i بالا یا پایین باشد تنها یکی از دو دسته مشتقات لازم است.



$$O : \vec{d}_1 = (0,0,0)$$

$$H : \vec{d}_2 = (+1.4325, 0, 1.10940)$$

$$H : \vec{d}_3 = (-1.4325, 0, 1.10940)$$

۱۰ تا الکترون ۵ تا اربیتال ملکولی می‌خواهیم:

$$\begin{aligned} \Phi_m(i) &= \Phi_m(\vec{x}_i, \vec{d}_1, \vec{d}_2, \vec{d}_3) = \\ &\phi_{1s,m}(\vec{x}_i, \vec{d}_2) + \phi_{1s,m}(\vec{x}_i, \vec{d}_3) + \phi_{1s,m}(\vec{x}_i, \vec{d}_1) + \phi_{2s,m}(\vec{x}_i, \vec{d}_1) \\ &+ \sum_{k=1}^3 \phi_{p_k,m}(\vec{x}_i, \vec{d}_1) \quad , \quad m=1,2,..5, \quad k=(1,2,3) \equiv (x,y,z) \end{aligned}$$

$$\begin{cases} \phi_{1s,m}(\vec{x}_i, \vec{d}_\alpha) = C_{1,\alpha,m} \exp(-\zeta_{1,m,\alpha} r_{i,\alpha}) \\ \phi_{2s,m}(\vec{x}_i, \vec{d}_\alpha) = C_{2,\alpha,m} r_{i,\alpha} \exp(-\zeta_{2,m,\alpha} r_{i,\alpha}) \\ \phi_{p_k}(\vec{x}_i, \vec{d}_1) = b_{1,m} x_{i,\alpha}^k \exp(-\zeta_{1,m,\alpha}' r_{i,\alpha}) + b_{2,m} x_{i,\alpha}^k \exp(-\zeta_{2,m,\alpha}'' r_{i,\alpha}) \quad k=(1,2,3) \equiv (x,y,z) \end{cases}$$

where $r_{i,\alpha} = |\vec{x}_i - \vec{d}_\alpha|$

قادریم تابع موج را بسازیم و مشتقات مثل قبل هستند.

$$\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{d}_1, \vec{d}_2, \dots) = D(1,2,..5) \uparrow D(6,..10) \downarrow \exp(J)$$

$$D \uparrow = \begin{vmatrix} \Phi_1(1) & \Phi_1(2) & \Phi_1(3) & \Phi_1(4) & \Phi_1(5) \\ \Phi_2(1) & \Phi_2(2) & \Phi_2(3) & \Phi_2(4) & \Phi_2(5) \\ \Phi_3(1) & \Phi_3(2) & \Phi_3(3) & \Phi_3(4) & \Phi_3(5) \\ \Phi_4(1) & \Phi_4(2) & \Phi_4(3) & \Phi_4(4) & \Phi_4(5) \\ \Phi_5(1) & \Phi_5(2) & \Phi_5(3) & \Phi_5(4) & \Phi_5(5) \end{vmatrix}$$

$$D \downarrow = \begin{vmatrix} \Phi_1(6) & \Phi_1(7) & \Phi_1(8) & \Phi_1(9) & \Phi_1(10) \\ \Phi_2(6) & \Phi_2(7) & \Phi_2(8) & \Phi_2(9) & \Phi_2(10) \\ \Phi_3(6) & \Phi_3(7) & \Phi_3(8) & \Phi_3(9) & \Phi_3(10) \\ \Phi_4(6) & \Phi_4(7) & \Phi_4(8) & \Phi_4(9) & \Phi_4(10) \\ \Phi_5(6) & \Phi_5(7) & \Phi_5(8) & \Phi_5(9) & \Phi_5(10) \end{vmatrix}$$

$$J = \left(\sum_{i < j} u(r_{ij} \uparrow \uparrow \text{ or } \downarrow \downarrow) \right) + \left(\sum_{i < j} u(r_{ij} \uparrow \downarrow \text{ or } \downarrow \uparrow) \right)$$

$$u(r_{ij}) = \frac{a r_{ij}}{1 + b r_{ij}} \quad a = \frac{1}{4} \uparrow \uparrow \text{ or } \downarrow \downarrow, \quad a = \frac{1}{2} \uparrow \downarrow \text{ or } \downarrow \uparrow \quad r_{ij} = |\vec{x}_i - \vec{x}_j|.$$

به همین طریق می‌توان تابع موج تمام ملکولهای دیگر را هم ساخت. کدی که ما آماده کرده ایم برای تمام ملکولهای تا ۴ اتم با هر تعداد الکترون به کار می‌رود.

۱۱ روش‌های بهینه‌سازی

۱-۱۱ مقدمه

پس از گزینش تابع موج آزمایشی که در گزارش اول ذکر آنها به میان رفت به بهینه‌سازی تابع موج خواهیم پرداخت به طوری که در این گزارش مروری بر روشهای بهینه سازی شده است.

همان‌گونه که بیان شد معمولاً تابع موج آزمایشی در محاسبات QMC به صورت حاصلضرب یک دترمینان پادمتقارن در تابع جسترو متقارن است. بخش پادمتقارن می‌تواند شامل یک یا چند دترمینان اسلیتر باشد، در حالی که تابع جسترو یک تابع بس‌ذره‌ای متقارن است و دربردارنده‌ی برهم‌کنش الکترونی در دستگاه است.

دو رهیافت کلی در بهینه‌سازی توابع موج وجود دارد؛ کمینه سازی انرژی (چشم‌داشتی) و کمینه سازی واریانس انرژی (چشم‌داشتی) که قبل از مرور بر روشهای بهینه سازی به آنها اشاره می‌کنیم:

۲-۱۱ کمینه سازی واریانس انرژی

اگر تابع موج آزمایشی، یک ویژه‌حالت هامیلتونی دستگاه باشد انرژی و مقادیر چشم‌داشتی عملگرهایی که با هامیلتونی جابجا می‌شوند از اصل واریانس صفر^{۸۲} پیروی می‌کنند. این روش به کمینه‌سازی واریانس انرژی موضعی روی N_c پیکربندی $\{\vec{R}_i\}$ (که از تابع توزیع $|\Psi(\vec{R}_i, \alpha^0)|^2$ نمونه‌گیری شده است)، می‌پردازد. $\Psi(\vec{R}_i, \alpha^0)$ تابع موج نخستین است (α^0 ها پارامترهای آغازین هستند) که پیش از آغاز فرایند بهینه‌سازی آن را مشخص می‌کنیم. واریانس انرژی را می‌توان به صورت زیر نوشت:

$$\sigma^2(\alpha) = \sum_{i=1}^{N_c} \left[\frac{H\Psi(\vec{R}_i, \alpha)}{\Psi(\vec{R}_i, \alpha)} - \bar{E} \right]^2 w(\vec{R}_i, \alpha) / \sum_{i=1}^{N_c} w(\vec{R}_i, \alpha) \quad (113)$$

که

$$\bar{E} = \sum_{i=1}^{N_c} \frac{H\Psi(\vec{R}_i, \alpha)}{\Psi(\vec{R}_i, \alpha)} w(\vec{R}_i, \alpha) / \sum_{i=1}^{N_c} w(\vec{R}_i, \alpha) \quad (114)$$

و

$$w(\vec{R}_i, \alpha) = \left| \frac{\Psi(\vec{R}_i, \alpha)}{\Psi(\vec{R}_i, \alpha^0)} \right|^2 \quad (115)$$

⁸²Zero variance principle

$\bar{E}$ ، انرژی میانگین روی پیکربندی‌هاست. توابع وزن $w(\vec{R}_i, \alpha)$ باعث می‌شود که نقاط گره در تابع موج در طی بهینه‌سازی آزادانه جابجا شوند و وزن درستی تولید کنند و نسبت هر یک از وزن‌ها به وزن میانگین از یک مقدار بیشینه اضافه‌تر نمی‌شود. از طرف دیگر با داشتن تعداد پیکربندی کم که از یک وزن بسیار بزرگ و انرژی موضعی نزدیک به $\bar{E}$ به دست آمده است می‌توانیم در طی فرآیند بهینه‌سازی به مقادیر کوچک $\sigma(\alpha)$ برسیم. معمولاً ۳۰۰۰ پیکربندی کافی است تا تابعی با ۵۰-۱۰۰ پارامتر را بهینه کنیم.

روش کمینه سازی واریانس روش کارآمدی است که در دو دهه‌ی گذشته کاربرد گسترده‌ای داشته است. در این روش با نزدیک شدن واریانس به صفر، انرژی به مقدار کمینه‌ی مورد انتظار می‌رسد و با این‌که تعداد پیکربندی‌ها کم است این کمینه یک کمینه-ی واقعی است (مقدار کمینه‌ی انرژی کمتر از مقدار انرژی تجربی به دست نمی‌آید)، در حالی‌که با کمینه‌سازی مستقیم انرژی، بهینه‌سازی تابع موج با تعداد کم پیکربندی‌ها، ضمانت رسیدن به یک کمینه‌ی واقعی را نمی‌دهد.

در روش کمینه‌سازی واریانس از یک تابع با پارامترهایی که بهینه نیستند آغاز می‌کنیم. در این روش در مسیر بهینه‌سازی نیازی به تولید دوباره‌ی پیکربندی نیست و تغییر پارامترها روی یک مجموعه پیکربندی ثابت صورت می‌گیرد. از این‌رو محاسبات QMC در روش کمینه‌سازی واریانس نسبت به روش‌های مستقیم کمینه‌سازی انرژی زمان بسیار کمتری صرف می‌کند. در عمل می‌توان در بهینه‌سازی تابع موج دو روش کمینه‌سازی واریانس انرژی و کمینه‌سازی انرژی را با هم به کار برد. در رابطه‌ی ۴ بالا معمولاً $\bar{E}$ را با E_{guess} جایگزینی کنند. E_{guess} را به گونه‌ای برمی‌گزینیم که اندازه‌ی آن از پایین‌ترین مقدار انرژی که معمولاً روش‌های QMC به دست می‌دهند کمتر باشد. در یک انتخاب اولیه می‌توان E_{guess} را برابر با E_{HF} (انرژی محاسبات هارتری-فوک) قرار داد.

۱۱-۳ کمینه‌سازی انرژی

از آنجا که در روش VMC به دنبال یافتن کمینه‌ی انرژی هستیم یک رهیافت مشخص کمینه‌سازی خود انرژی ($\bar{E}$) است. در این روش باید پارامترهای ورودش در تابع موج طوری تغییر یابند که ما را به سوی کمینه‌ی انرژی ببرند.

با داشتن مشتقات مرتبه اول و دوم انرژی نسبت به پارامترها می‌توانیم روش‌های بهینه‌سازی چون تندترین کاهش (بخش)، روش نیوتن^{۸۳} و ... را در کمینه‌سازی انرژی به کار گیریم. برای تابع موجی با پارامترهای زیاد به دست آوردن مشتقات انرژی نسبت به پارامترها به طور عددی دشوار است و از دقت نتایج نیز به طور چشم‌گیری کاسته می‌شود، مگر این‌که به دنبال روش‌هایی باشیم تا مشتقات انرژی نسبت به پارامترها را به طور تحلیلی و دقیق به دست آوریم.

⁸³Newton's method

۴-۱۱ روش‌های بهینه‌سازی

با داشتن روابط انرژی، واریانس انرژی و همچنین مشتقات انرژی نسبت به پارامترها اینک باید در پی الگوریتم‌هایی باشیم تا با کمک این روابط پارامترهای وردش را در سوی یافتن کمینه های مورد نظر تغییر دهند.

۱-۴-۱۱ روش کاهش شیب (GD)^{۸۴}

تابع حقیقی و مشتق پذیر $f(x)$ را در نظر می‌گیریم. به دنبال یافتن فرینه‌های این تابع هستیم، بدیهی است با داشتن ریشه‌های تابع $g(x)$

$$g(x) = \frac{d}{dx} f(x) \quad (۱۱۶)$$

به نقاط فرینه (کمینه و بیشینه‌ی) تابع $f(x)$ دست یافته‌ایم. یعنی باید الگوریتمی برای حل معادله‌ی غیرخطی $g(x) = 0$ ارائه کنیم. اگر نقطه‌ی x_k را در دامنه‌ی $f(x)$ برگزینیم و هر بار این نقطه را به اندازه‌ی Δx_k ، طوری تغییر می‌دهیم تا ما را به سوی یافتن نقطه‌ی فرینه هدایت کند معادله‌ی بازگشتی به صورت زیر خواهیم داشت:

$$x_{k+1} = x_k + \Delta x_k \quad (۱۱۷)$$

فرض کنیم به دنبال یافتن مقدار صفر کننده $g(x)$ هستیم. در مورد کمینه یابی، باید نقطه‌ی x_k را در جهتی جابجا کنیم که $f(x)$ به سرعت کاهش یابد. برای این کار اگر در رابطه‌ی فوق، Δx_k علامتی مخالف مشتق اول $f(x)$ در نقطه‌ی x_k ، $f'(x_k)$ ، داشته باشد رابطه‌ی بازگشتی را به سوی یافتن کمینه‌ی $f(x)$ هدایت کرده‌ایم (بالعکس برای یافتن بیشینه باید در جهت $+f'(x_k)$ حرکت کنیم). از این‌رو رابطه‌ی بازگشتی بالا، برای یافتن کمینه‌ی $f(x)$ به صورت زیر تغییر می‌کند

$$x_{k+1} = x_k - a_k f'(x_k), \quad k \geq 0 \quad (۱۱۸)$$

که a_k پارامتری مثبت، حقیقی، و کوچک است که در هر مرحله چنان تنظیم میشود تا به کمینه مقدار تابع رهنمون شویم. این عدد را در نزدیکی جواب باید کوچک و در نقاط دور از آن باید بزرگتر داشته باشیم. رابطه‌ی اخیر را می‌توان برای یک تابع چند متغیره به صورت زیر گسترش داد

$$\vec{x}_{k+1} = \vec{x}_k - a_k \vec{\nabla} f(x_k), \quad k \geq \quad (۱۱۹)$$

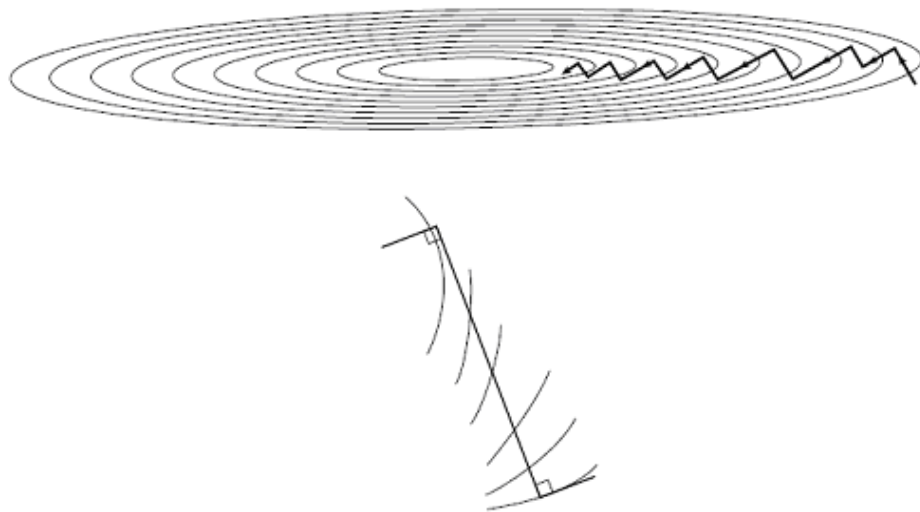
که $\vec{x} = (x_1, x_2, \dots, x_l)$ و $\vec{\nabla} f(\vec{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_l} \right)$ است. هر گام $\Delta \vec{x}_k$ در جهت $-\vec{\nabla} f(\vec{x}_k)$ تغییر می‌یابد و هر $\vec{x}_{k+1}$ جدید مجبور می‌شود در راستای تندترین کاهش جابجا شود)^{۸۵}:

⁸⁴Gradient Descent

^{۸۵}از ریاضیات مقدماتی می‌دانیم که جهت بیشینه تغییر یک تابع جهت گرادیان است و بالعکس.

$$f(x_{k+1}) < f(x_k) < \dots < f(x_0) \quad (۱۲۰)$$

برای تعیین a_k ، وقتی نقطه شروع $\vec{x}_0$ مشخص شده باشد، چون $f(x_1, a_0) = f(\vec{x}_0 - a_0 \vec{\nabla} f(x_0))$ تابعی از a_0 است می‌توان مقداری از آنرا که کمینه‌کننده $f(\vec{x}_1, a_0)$ است را از شرط $\frac{df(\vec{x}_1, a_0)}{da_0} = 0$ پیدا کرد. این شرط حاکی از آنست که گرادیان گام فعلی و گام بعدی متعامد باشند (برای اثبات از $\vec{\nabla} f(\vec{x}_1, a_0) \cdot \frac{d\vec{x}_1}{da_0} = \frac{df(\vec{x}_1, a_0)}{da_0}$ و مقدار $\vec{x}_1 = \vec{x}_0 - a_0 \vec{\nabla} f(x_0)$ استفاده می‌شود) یا عبارت دیگر مسیر حرکت در راستای پیدا کردن کمینه زیگزاگ مانند شکل ۶ باشد.



شکل ۶) روش کاهش شیب در یک دره‌ی باریک و بلند؛ هر گام در مسیر گرادیان موضعی، عمود بر خطوط منحنی یک مسیر مستقیم را طی می‌کند تا وقتی که به یک کمینه‌ی موضعی برسد، جایی که مسیر حرکت با خطوط منحنی همسو است. قسمت پایین شکل یک گام را نشان می‌دهد.

بنابراین اگر از یک نقطه نزدیک کمینه موضعی حرکت کنیم پس از گام متناهی به نزدیکترین کمینه موضعی به نقطه شروع می‌رسیم. برای مثال تابع $f(x) = x^4 - 3x^3 + 2$ با مقدار ثابت $a_k = 0.01$ و $x_0 = 6$ با ۷۰ گام به مقدار ۲/۲۴۹۹۶ می‌رسد. به عنوان تمرین تابع دو متغیره زیر را با شروع از (۰ و ۰) و با مقدار $a_k = 0.1$ کمینه یابی کنید.

$$f(x, y) = \sin\left(\frac{1}{2x^2} - \frac{1}{4y^2} - 3\right) \cos(2x + 1 - e^y) \quad (۱۲۱)$$

روش کاهش شیب برای فضای پارامتری بزرگ نیز کاراست. اما کاستی‌هایی نیز دارد؛ هنگامی که مسیر حرکت در سوی یافتن کمینه داخل دره‌ای باریک و بلند قرار می‌گیرد، حتی اگر دره شکل درجه دوم کاملی داشته باشد، باز هم گام‌های زیادی تا یافتن کمینه لازم است - یعنی این الگوریتم هزینه بردار است - الگوریتم کلی کار چنین است:

- Step 1. Estimate a starting design $\mathbf{x}^{(0)}$ and set the iteration counter $k = 0$. Select a convergence parameter $\varepsilon > 0$.
- Step 2. Calculate the gradient of $f(\mathbf{x})$ at the point $\mathbf{x}^{(k)}$ as $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$. Calculate $|\mathbf{c}| = \sqrt{\mathbf{c}^T \mathbf{c}}$. If $\|\mathbf{c}\| < \varepsilon$, then stop the iteration process as $\mathbf{x}^* = \mathbf{x}^{(k)}$ is a minimum point. Otherwise, go to Step 3.
- Step 3. Let the search direction at the current point $\mathbf{x}^{(k)}$ as $\mathbf{d}^{(k)} = -\mathbf{c}^{(k)}$.
- Step 4. Calculate a step size $\alpha^{(k)}$ to minimize $f(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)})$. A one-dimensional search is used to determine $\alpha^{(k)}$.
- Step 5. Update the design as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)}$. Set $k = k + 1$ and go to Step 2.

۱۱-۴-۲ روش نیوتن-رفسون

روش نیوتن^{۸۶} روشی معروف برای یافتن ریشه‌های تابعی چند متغیره است. اگر $g(x) = f'(x)$ باشد برای فرینه‌های تابع f ، باید ریشه‌های معادله‌ی $g(x) = 0$ را بیابیم. وقتی x_r ، یک ریشه‌ی این معادله باشد می‌توانیم معادله‌ی $g(x_r) = 0$ را حول نقطه‌ی x_r بسط تیلور تا مرتبه خطی دهیم:

$$g(x_r) \approx g(x) + (x_r - x)g'(x) + \dots = 0 \quad (122)$$

پس x می‌تواند یک نقطه‌ی آزمایشی در $(k+1)$ امین گام برای ریشه‌ی x_r در (k) امین گام باشد

$$g(x_{k+1}) = g(x_k) - (x_{k+1} - x_k)g'(x_k) \approx 0 \quad (123)$$

و x با رابطه‌ی بازگشتی زیر در هر گام تغییر می‌یابد

$$x_{k+1} = x_k - \Delta x_k = x_k - g(x_k)/g'(x_k) \quad (124)$$

رابطه‌ی اخیر بر حسب تابع $f(x)$ به صورت زیر تغییر می‌یابد

$$x_{k+1} = x_k - \Delta x_k = x_k - f'(x_k)/f''(x_k) \quad (125)$$

این به آن معناست که حول هر نقطه یک سهمی در نظر گرفته شده است و هر بار در جهت کمینه آن گام برداشته می‌شود. می‌توان معادله‌ی بالا را در یافتن کمینه‌ی تابع n متغیره‌ی $f(\vec{x})$ ، به صورت زیر گسترش داد

$$\vec{x}_{k+1} = \vec{x}_k - (H_k)^{-1} \vec{\nabla} f(x_k), \quad k \geq 0 \quad (126)$$

که $b = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$ و H ماتریس هسین^{۸۷} است (H^{-1} وارون ماتریس H است). ماتریس هسین باید مقارن باشد:

⁸⁶Newton's method

⁸⁷Hessian matrix

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix} \quad (۱۲۷)$$

روش نیوتن اصلاح شده را با اضافه کردن ضریب a_k به صورت زیر می نویسیم

$$\vec{x}_{k+1} = \vec{x}_k - a_k (H_k)^{-1} \vec{\nabla} f(x_k) \quad , k \geq 0 \quad (۱۲۷)$$

در مقایسه با روش تندترین شیب علاوه بر مشتق اول از مشتق دوم نیز در مسیر یابی به سوی کمینه استفاده می شود و این باعث لحاظ انحنای موضعی تابع و در نتیجه سرعت بالاتر الگوریتم به سمت کمینه موضعی می گردد.

دشواری‌هایی نیز وجود دارد: با توجه به رابطه‌ی بازگشتی، برای این‌که در هر گام در جهت کمینه‌ی تابع $f(\vec{x})$ حرکت کنیم نیاز به محاسبه‌ی مشتقات مرتبه‌ی اول و دوم تابع نسبت به متغیرها داریم. در عمل بسیاری مواقع محاسبه‌ی مشتقات تابع نسبت به متغیرها به صورت تحلیلی دشوار است. از طرفی به دست آوردن مشتقات به روش‌های عددی از دقت جواب‌ها می‌کاهد. در صورت صفر بودن دترمینان ماتریس هسین دیگر محاسبه‌ی وارون آن از روش‌های معمولی امکان‌پذیر نیست و باید به دنبال راهی برای زدودن تکینگی ماتریس هسین باشیم. یک راه این است که با تشکیل ماتریس جدید زیر می توان ویژه مقدارهای منفی ماتریس قدیم را با کمک ماتریس تصحیح B چنان تغییر داد تا به مقادیر مثبت برسند در عین حالی که ویژه بردارهای ماتریس جدید و ماتریس اصلی یکسان باشند.

$$H_{\text{new}} = H + B \quad , \forall k \geq 0 \quad (۱۲۸)$$

از طرفی به دلیل زمان بر بودن محاسبه وارون ماتریس از روشهایی مثل گرادیان همبسته برای حل معادله خطی ۱۲۶ استفاده می شود .

۱۱-۴-۳ گرادیان همبسته (CG)

در روش کاهش شیب فرض کنیم نخستین گام ما را به سوی نقطه‌ی کمینه‌ی تابع هدایت کند و حرکت در دومین گام در راستای محور طولی باشد، اما سوی گرادیان جدید در نقطه-ی کمینه عمود بر مسیر حرکت است از این رو به سویی حرکت کرده‌ایم که در حالت کلی ما را به نقطه‌ی کمینه نمی‌برد (شکل ۶). بنابراین باید فرآیندی بیابیم که به جای حرکت در جهت گرادیان در نقطه‌ی جدید در مسیری حرکت کنیم که به گرادیان نقطه‌ی پیشین همبسته باشد و این روش به گرادیان همبسته موسوم است.

در روش CG، در گام نخست تنها در جهت گرادیان نقطه‌ی آغازین حرکت نمی‌کنیم. بلکه جهت حرکت به گرادیان نقطه‌ی پیشین نیز بستگی دارد و این امر باعث می‌شود که بتوانیم مسیر بهینه‌ی بهتری در جهت رسیدن به مقدار کمینه انتخاب کنیم. برای یک تابع درجه‌ی دوم در N بُعد، با روش CG می‌توانیم پس از N گام مقدار کمینه را بیابیم. افزون بر این روش CG می‌تواند در بهینه‌سازی بسیاری از توابع پیچیده‌ی فیزیک (مانند معادلات کن-شم^{۸۸}) به کار رود و نتایج خوبی داشته باشد. فرض کنیم به دنبال یافتن کمینه‌ی تابع چند متغیره‌ی $f(\vec{x})$ هستیم. در روش CG، با بردار $\vec{g}_0$ که برابر منفی گرادیان $f(\vec{x})$ در نقطه‌ی $\vec{x}_0$ است، آغاز می‌کنیم و $\vec{h}_0$ را مساوی $\vec{g}_0$ در نظر می‌گیریم. حال می‌توان با کمک روابط بازگشتی زیر $\vec{g}_{i+1}$ ، $\vec{h}_{i+1}$ و $\vec{x}_{i+1}$ را بر حسب مقادیر پیشین‌شان ($\vec{g}_i$ و $\vec{h}_i$) به دست آورد:

$$\vec{g}_{i+1} = \vec{g}_i - \lambda_i H \vec{h}_i$$

$$\vec{h}_{i+1} = \vec{g}_{i+1} + \gamma_i \vec{h}_i \quad (129)$$

$$\vec{x}_{i+1} = \vec{x}_i + \lambda_i \vec{h}_i$$

که λ_i و γ_i به صورت زیر هستند:

$$\lambda_i = \frac{\vec{g}_i \cdot \vec{g}_i}{\vec{g}_i H \vec{g}_i} \quad (130)$$

$$\gamma_i = -\frac{\vec{g}_{i+1} \cdot \vec{g}_{i+1}}{\vec{h}_i H \vec{h}_i}$$

روش CG نسبت به روش SD روش بهینه‌سازی کاراتری است، به ویژه وقتی که تابع مورد نظر به شکل درجه‌ی دوم نزدیک‌تر باشد. اما به شونده نیاز به محاسبه‌ی مشتقات مرتبه‌ی دوم در ماتریس هسین پیچیدگی بیشتری دارد.

۱۱-۵ روش شبیه سازی مونت کارلو سردسازی^{۸۹}

بهتر است قبل از آنکه به الگوریتم سرد سازی بپردازیم، روش نمونه برداری با اهمیت مونت کارلو از فضای فاز سیستمی که به حمام گرمایی در دمای T متصل است، را به طور خلاصه مرور می‌کنیم.

۱۱-۵-۱ روش شبیه‌سازی مونت کارلو

روش مونت کارلو در ابتدا برای حل انتگرالهای چند بُعدی استفاده شدند. امروزه بیشتر برای نمونه برداری از بخشهای پراهمیت فضای فاز که مرتبط با حل انتگرالهای چند بُعدی موجود در

⁸⁸Kohn-Sham

⁸⁹Annealing monte carlo

ترمودینامیک و مکانیک آماری می‌شود، به کار می‌روند. برای حل انتگرالهای موجود در مکانیک کوآنتومی، یعنی انتگرالهای موجود در مقادیر انتظاری مشاهده پذیرها، نسخه‌های کوآنتومی این روش هم بوجود آمده‌اند. در حل مسائل فیزیکی مربوط به محاسبات ساختار الکترونی از نسخه کوآنتومی استفاده می‌کنند که روز به روز بر اهمیت و جایگاه آنها افزوده می‌شود. اولین و ساده‌ترین کاربردهای این روش در سیستمهای اسپینی - مثل مدل آیزینگ- است.

در نمونه برداری با اهمیت ترجیح می‌دهیم اعداد تصادفی (معادل متغیرهای تابع انتگرالده) را طوری انتخاب کنیم که منجر به مقدار بزرگتر تابع انتگرالده مورد انتگرال گیری بشوند. به بیان دیگر نمونه گیری ما بایستی وزن دار باشد تا مجبور به لحاظ سهمهای کم وزن انتگرالده نباشیم. در نمونه برداری با اهمیت مونت کارلو، میانگین هنگردی بندادی یک کمیت مثل A به صورت زیر نوشته می‌شود:

$$\langle A \rangle = \frac{1}{N_{trial}} \sum_{i=1}^{N_{trial}} A_i, \quad (131)$$

که جمع روی تعدادی از پیکربندیهاست و A_i مقدار کمیت مربوط به پیکربندی i ام تولید شده با یک فرایند کاتوره ای است، به نحوی که هر A_i با احتمال متناسب با $\exp(-\frac{U_i(\vec{r})}{k_B T})$ رخ می‌دهد. در رابطه اخیر U تابع انرژی سیستم مورد بررسی است^{۹۰}. توجه شود که در نمونه برداری با اهمیت مختصات تصادفی از یک توزیع غیریکنواخت انتخاب می‌شوند. این پیکربندی‌ها را می‌توان از یک زنجیر مارکوف^{۹۱} انتخاب کرد. در عمل برای نمونه برداری با اهمیت از الگوریتم متروپلیس^{۹۲} (ساده ترین جوابی که برای حل معادلات مربوط به مباحث نظری مونت کارلو پیشنهاد می‌شود) استفاده می‌کنیم:

دستگاه را در پیکربندی با انرژی U_i و دمای T در نظر بگیرید. یک جزء از دستگاه (معمولاً مختصه های فضایی یا زاویه ای) به صورت اتفاقی انتخاب می‌شود و به طور تصادفی به موقعیت جدید یا به جهت گیری جدید تغییر داده می‌شود. در این صورت، دستگاه در حالت جدید A_j با انرژی U_j می‌باشد. وقتی که $U_j \leq U_i$ باشد، در آن صورت حرکت با احتمال «یک» پذیرفته می‌شود ولی برای موردی که $U_j > U_i$ باشد حرکت با احتمال $\exp(-\frac{\Delta U_{ij}}{k_B T})$ پذیرفته می‌شود. عملاً مرحله دوم به این صورت انجام می‌شود که $\exp(-\frac{\Delta U_{ij}}{k_B T})$ با یک عدد تصادفی "ξ" بین ۰ و ۱ (این عدد تصادفی در دامنه ۰ و ۱ به طور یکنواخت تولید می‌شود) مقایسه می‌شود. اگر $\xi \leq \exp(-\frac{\Delta U_{ij}}{k_B T})$ باشد در آن

از U برای انرژی واقعی و از E برای پارامتر انرژی گونه استفاده می‌کنیم.

⁹¹Markov chain

⁹²لغت مونت کارلو توسط نیکولاس متروپولیس در سال ۱۹۴۹ انتخاب شد.

صورت حرکت پذیرفته می‌شود. در غیر این صورت حرکت رد می‌شود.^{۹۳} با تکرار این فرایند به اندازه کافی (تا جایی که سیستم حافظه خود را از دست بدهد یا اصطلاحاً ناهمبسته شود) ترتیبی از حالات تولید می‌شود (زنجیر مارکوف)، به این امید که از فضای فاز کاملاً نمونه برداری شده باشد. توجه کنیم که اعمال شرط دوم برای خلاصی از کمینه‌های موضعی دستگاه است.

۱۱-۵-۲ روش شبیه سازی سرد سازی

ایده سرد سازی در سال ۱۹۸۳ توسط Kirkpatrick و همکارانش معرفی شد. کمینه سازی سردسازی روشی عام (نه فقط برای کمینه یابی انرژی) و احتمالاتی برای پیدا کردن تقریبی کمینه اصلی تابعی است که فضای پارامترهای آن بزرگ است. مزیت اصلی این روش بر روشهایی که بر اساس مشتق گیری^{۹۴} از یک تابع تحلیلی به طور تعینی جلو می‌روند، آنست که در سرد سازی نگرانی از بابت عدم وجود مشتقات و گیر افتادن در کمینه‌های موضعی نیست.

شباهتی که بین یک مساله کمینه سازی سردسازی با این روش و سردشدن یک سیستم واقعی وجود دارد باعث تعریف پارامترهای دما و انرژی در هر مرحله سردسازی می‌شود به طوری که لزوماً این دو پارامتر معنای دما و انرژی داخلی آشنا در فیزیک را ندارند.

دمای مطروح صرفاً یک پارامتر برای حفظ شباهت با مساله سرد سازی واقعی تعریف می‌شود. پارامتر انرژی هم معیاری کنترلی است که نشان دهد تابع به حالت نهایی با کمینه مقدار (پارامتر انرژی) رسیده است.

روش کار به این صورت است که برای هر مقدار پارامتر دما T ، یک حالت r (که معمولاً در مساله های کمینه سازی انرژی یا ساختار مجموعه ای از پارامترها مثل مختصه های فضایی یک ترکیب ملکولی) و یک پارامتر انرژی $E(r)$ تعریف می‌کنیم، سپس با انتخاب تصادفی حالت‌های مجاور r' پارامتر انرژی $E(r')$ را حساب می‌کنیم و آنرا با $E(r)$ مقایسه کرده و با احتمال $p(E(r), E(r'), T)$ حالت r' را جایگزین r می‌کنیم. تابع p چنان انتخاب می‌شود^{۹۵} که حتی اگر $E(r') > E(r)$ بود عددی مثبت و غیر صفر باشد یعنی شانس بیشتر شدن پارامتر انرژی وجود داشته باشد تا این تضمینی برای رهایی از کمینه‌های موضعی باشد. همینطور برای پارامترهای T نزدیک صفر، تابع p به صفر میل کند تا بیشتر به پارامترهای $E(r)$ با مقادیر کمتر میل کنیم. همینطور تابع p چنان باشد که در $T=0$ فقط اگر $E(r') < E(r)$ بود حالت جدید پذیرفته شود. بنابراین در $T=0$ مقدار تابع p بایستی ۱ بشود. بهتر است تابع p چنان انتخاب شود که هر چه اختلاف $E(r') - E(r)$

^{۹۳}تلفیقی از این دو شرط را $(U_i \geq U_j \text{ or } U_j > U_i)$ به صورت $p = \min[1, \exp(-\Delta U / k_B T)]$ که به حل

مترو پولیس مشهور است، نشان می‌دهیم.

^{۹۴}Gradient methods

^{۹۵}این کمیت به طور معمول $p = e^{-\frac{\Delta E}{T}}$ تعیین می‌شود.

بیشتر شد، احتمال پذیرش کمتر شود تا همیشه به آهستگی به سمت کمینه اصلی نزول کنیم. در واقع در دماهای بالا تغییر انرژی‌های بالا و حتی مثبت هم ممکن هستند و این کمک می‌کند در دماهای بالا سیستم فضای فاز بزرگتری را جستجو کند و هنگامی که دما پایین می‌آید آرام آرام این فضای دسترسی به حوالی جواب محدود می‌شود. تعداد تغییر حالت‌های $r \rightarrow r'$ در T ثابت (ما این تعداد را تعداد گام‌های مونت کارلویی می‌نامیم) به این بستگی دارد که چقدر حالت‌های جدید به حالت‌های قبلی ناهمبسته باشند. همبستگی را به کمک تابع خود همبستگی $\tilde{E}(i) = E(i) - \bar{E}$ و $C(k) = \langle \tilde{E}(i) \tilde{E}(i+k) \rangle$ محاسبه می‌کنیم که در آن i و k حالت‌های مختلف در یک دما را نشان می‌دهند. اگر این تابع را برحسب k رسم کنیم طول همبستگی تعداد گام‌هایی است که $C(k)$ به $1/e$ مقدار اولیه خود برسد [1,65]. معمولاً برای تعداد گام‌های مونت کارلو چندین برابر طول همبستگی در T های بالا و در مرتبه طول همبستگی در T های پایین تخمین خوبی است. بعد از اطمینان از متعادل بودن در هر دمای خاص، مقدار دما $-T$ را کاهش می‌دهیم و تمام عملیات تکرار می‌شود. روش کاهش دما^{۹۶} بسته به سیستم مورد نظر می‌تواند پله ای یا نمایی باشد. در روش پله ای هر بار مقدار ثابتی از دمای فعلی کم می‌شود تا این که به صفر برسیم: $T_{i+1} = T_i - \Delta T$ در روش نمایی هر بار عدد مثبت کوچکتر از یک در دمای فعلی ضرب می‌شود: $T_{i+1} = \alpha T_i \rightarrow \Delta T = (\alpha - 1)T_i$. مزیت روش دوم آنست که در دماهای پایین تر حالت‌های r' به هم نزدیک تر بوده ($\Delta T \rightarrow 0$) و تحول سیستم کندتر صورت می‌گیرد و تعداد گام‌هایی که برای یک T ثابت (البته نزدیک صفر) طی می‌شود از نظر بزرگ یا کوچک بودن نسبت به طول همبستگی نگران کننده نخواهد بود. نمودار ۷-۱ فلوچارتی از شبیه سازی سرد سازی را نمایش می‌دهد. به طور خلاصه ۴ پارامتر اصلی باید در الگوریتم سرد سازی مشخص باشد:

۱. دمای اولیه: این دما باید چنان تعیین شود که نرخ پذیرش^{۹۷} حرکت‌های حلقه مونت کارلویی حدود ۸۰ درصد باشد.

دمای نهایی: دمایی که سردسازی باید پایان بپذیرد در حالت کلی به عدم پیشروی الگوریتم و کم شدن نرخ پذیرش باز می‌گردد. معمولاً بهتر است آنرا صفر در نظر بگیریم ولی با دنبال کردن مداوم نتایج حتی اگر بالاتر از مقدار صفر دما سیستم بسیار کند تغییر شد سرد سازی را متوقف می‌کنیم.

۲. تعداد گام‌های (گام‌های مونت کارلویی) لازم در هر دما: همانطور که گفتیم به همبستگی آماری نتایج وابسته است.

۳. روش کاهش دما: روش‌های پله ای یا نمایی که هر یک بسته به مساله مورد نظر مورد استفاده قرار می‌گیرند.

یک اظهار نظر کلی اینکه نمی‌توان یک قاعده عام برای تعیین پارامترها طرح کرد، معمولاً مسائل مختلف نیاز به گزینش‌های خاص خود دارند.

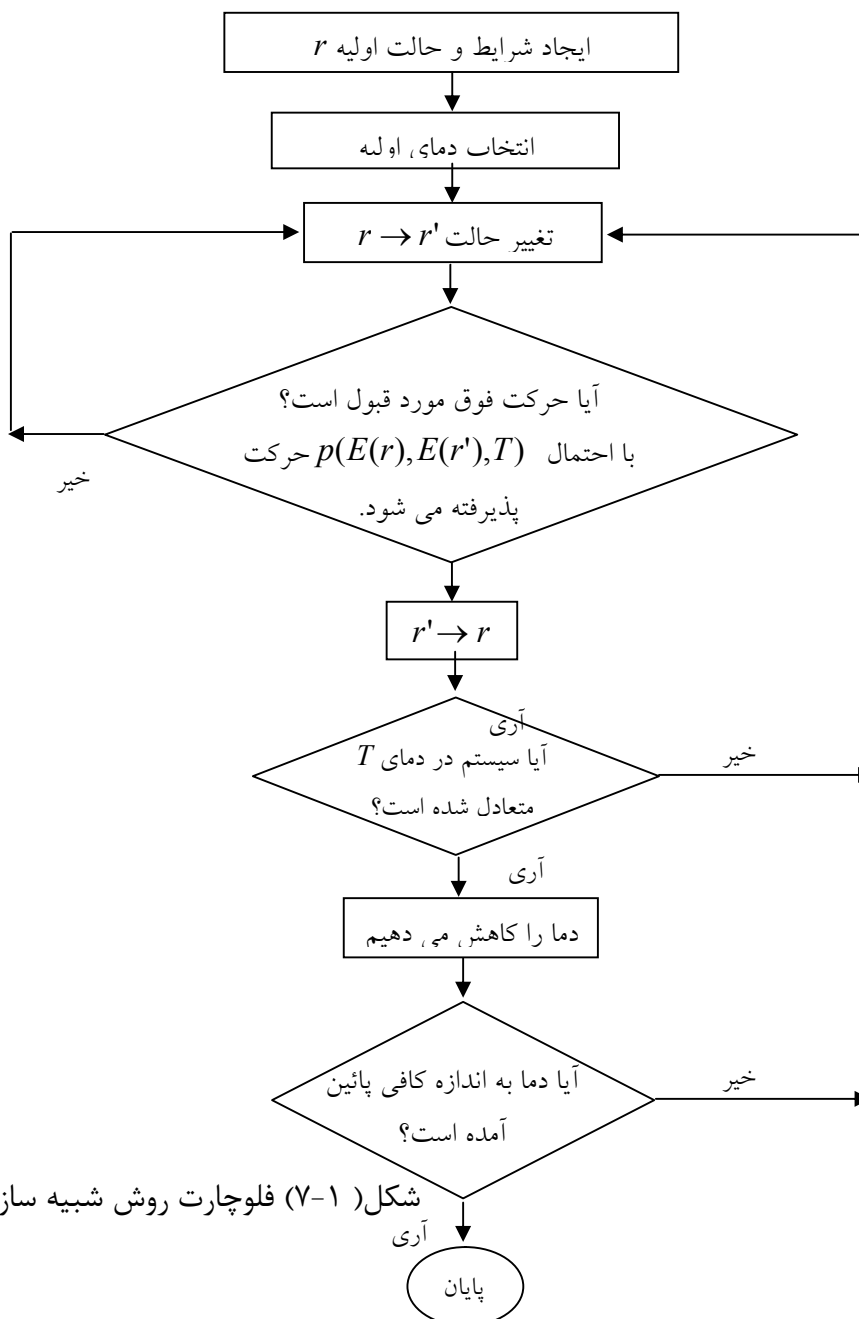
^{۹۶}Annealing schedule

^{۹۷}Acceptance rate

۱۱-۵-۳ روش کمینه‌یابی مونت کارلو سرد سازی

در کتاب حاضر هدف از سردسازی، پیدا کردن کمینه واقعی سیستم یعنی پیکربندی از MT ها با کمینه انرژی (بدون بُعد) و متعاقب آن پارامترهای این ساختار با کمینه انرژی است. از این رو به عنوان یک انتخاب، تابع $p = \min[1, \exp(-\Delta U / k_B T)]$ یعنی همان حل متروپولیس را به عنوان تابع پذیرش گذار بین حالتها انتخاب می‌کنیم. به دلیل همین انتخاب این نوع کمینه‌یابی‌ها را روش مونت کارلو سردسازی می‌نامند. در مدل مورد بررسی ما $E(r)$ انرژی ساختار N_s میکروتوبول است و حالات مربوطه، مجموعه متغیرهای x_i یعنی جابجایی هر PF نسبت به همسایه خود است: $[r = \{x_i\}, i = 1, 2, \dots, N]$. هم چنین تبدیل $r \rightarrow r'$ تغییری اندکی است (نسبت به δ اندازه یک هتروداپمر) که در جابجاییها بوجود می‌آوریم: $x'_k = x_k + \eta$. هم چنین برای هر دمای T و هر تعداد N ثابت تفاضل انرژی به صورت زیر تعریف می‌شود:

$$\Delta E = E\{x_1, x_2, \dots, x'_k, \dots, x_N\} - E\{x_1, x_2, \dots, x_k, \dots, x_N\} \quad (132)$$



شکل (۷-۱) فلوچارت روش شبیه سازی سرد سازی.

۱۱-۶ الگوریتم ژنتیک (GA)

ایده‌ی اصلی الگوریتم‌های تکاملی در سال ۱۹۶۰ توسط رکنبرگ^{۹۸} مطرح گردید [۴۵]. الگوریتم‌های ژنتیک^{۹۹} که منشعب از این نوع الگوریتم‌هاست، در حقیقت روش جستجوی کامپیوتری بر پایه‌ی الگوریتم‌های بهینه‌سازی بر اساس ساختار ژن‌ها و کروموزوم‌هاست که توسط هلاند^{۱۰۰} در دانشگاه میشیگان مطرح شد و پس از وی توسط جمعی از دانشجویانش گسترش یافت. این الگوریتم امروزه در بسیاری از دانش‌های گوناگون مانند زیست‌شناسی، دانش‌های فنی و مهندسی، دانش‌های پایه و ... کاربرد دارد.

۱۱-۶-۱ مفاهیم پایه

الگوریتم‌های ژنتیک یک روش جستجوی مؤثر در فضای بسیار وسیع و بزرگ است که در نهایت منجر به جهت‌گیری به سوی پیدا کردن یک پاسخ بهینه می‌گردد و شاید نتوان در مدت زمان زندگی یک فرد به آن پاسخ بهینه دست یافت. الگوریتم‌های ژنتیک تفاوت بسیار زیادی با روش‌های بهینه‌سازی قدیمی دارند. در GA با جمعیت یا مجموعه‌ای از نقاط در یک لحظه‌ی خاص کار می‌کنیم، در حالی که در روش‌های قدیمی بهینه‌سازی تنها برای یک نقطه‌ی خاص عمل می‌کردیم. این بدین معنی است که GA شمار زیادی از طرح‌ها را در یک زمان مورد پردازش قرار می‌دهد. اصول GA بر پردازش کاتوره‌ای یا به بیان درست‌تر پردازش کاتوره‌ای هدایت شده استوار است.

ایده‌ی اصلی GA، در واقع پیروی از فرایند تحوّل در نسل موجودات زنده است. موجوداتی که خود را با محیط پیرامون‌شان بهتر وفق می‌دهند احتمال زنده ماندن بیشتری دارند. بنابراین نسبت به موجودات ضعیف بیشتر می‌توانند به طور تناوبی ژن‌های خود را به نسل‌های پسین منتقل کنند. یعنی با گذشت زمان میانگین کیفیت جمعیت موجودات قوی بالا می‌رود. این ایده می‌تواند روشی برای بسیاری از مسئله‌های بهینه‌سازی ارائه کند. فرض کنیم به جای جستجوی بهترین موجودات، هدف ما یافتن کمینه یا بیشینه‌ی یک تابع است. بنابراین موجودات گوناگون به بردارهای متفاوت تبدیل می‌شوند و فرآیند تحوّل موجودات را می‌توان به آسانی به الگوریتمی برای یافتن بردارهای بهتر تا رسیدن به هدف تغییر داد. اصولاً برای استفاده از GA باید سه مفهوم مهم زیر مشخص شود:

۱- شناساندن تابع هدف

۲- شناساندن و پیاده‌سازی فضای ژنتیک

۳- شناساندن و پیاده‌سازی عملگرهای GA

⁹⁸Rechenberg

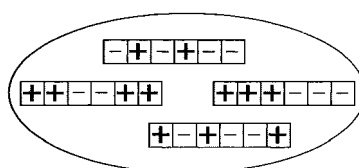
⁹⁹Genetic Algorithm

¹⁰⁰Holland

اگر این سه بخش به درستی بیان شوند، بی‌گمان GA به خوبی عمل خواهد کرد و سرانجام می‌توان با اعمال تغییراتی کارایی الگوریتم را افزایش داد. بدیهی است فضای جستجو و عملگرهای GA بسته به نوع مسئله‌ی بهینه‌سازی متفاوت است. اینک به بررسی بخش‌های گوناگون GA می‌پردازیم.

۱۱-۶-۲ جمعیت

نخست الگوریتم با مجموعه‌ای از جواب‌های کاتوره‌ای که به آنها جمعیت گفته می‌شود آغاز می‌گردد (شکل ۴-۲).



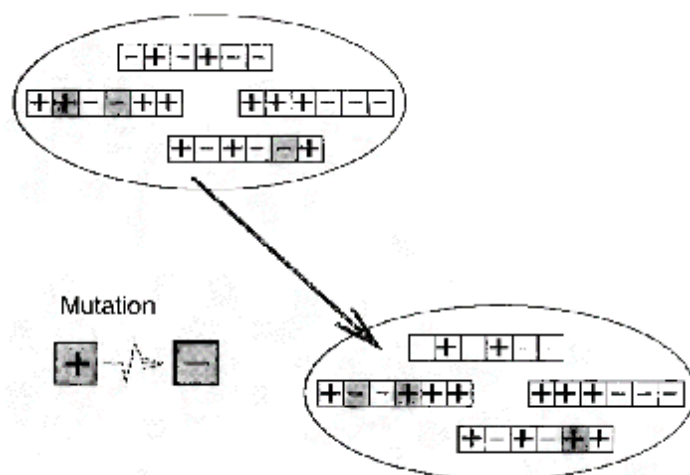
شکل ۷) جمعیتی از موجودات گوناگون؛ هر موجود شامل ویژگی‌های مخصوص به خود است (ژن، پارامتر یا مکان، در اینجا هر موجود با برداری از مقادیر $+/ -$ نشان داده شده است).

جمعیت در نظر گرفته شده و ویژگی‌های مربوط به هر موجود باید به صورتی بیان شود تا در کامپیوتر قابل کاربرد باشد. برای نمونه می‌توان جمعیت را به صورت رشته‌های بیتی نشان داد. برای این کار معمولاً از رشته‌هایی شامل بیت‌هایی از ۰ و ۱ استفاده می‌شود (گذردن مقادیر جمعیت در پایه‌ی ۲ و درازای رشته‌ها با توجه به دقت مسئله گزیده می‌شود).

پس از پیدایش جمعیت نخستین، اینک باید از این جمعیت برای ساخت جمعیت جدید پسین استفاده شود، به این امید که جمعیت‌های جدید بهتر از جمعیت‌های قدیم باشند. مفهوم بهتر بودن را بزودی توضیح خواهیم داد. در این میان عملگرهایی روی موجودات عمل می‌کنند تا بخش‌های گوناگون الگوریتم را بسازند و سرانجام دستیابی به بهترین جمعیت برآورده شود.

۱۱-۶-۳ عملگر جهش

جهش ژن‌ها در واقع تغییراتی است که به صورت کاتوره‌ای در یکی از موجودات اعمال می‌شود (جابجایی یک ذره یا تغییر یک پارامتر). شکل ۴-۳ اعمال عملگر جهش^{۱۱} را روی جمعیت که با یک احتمال کوچک رخ می‌دهد، نشان می‌دهد. بیت‌های یک رشته به صورت نابسته جهش می‌یابند، به این معنی که جهش یک بیت بر روی احتمال جهش سایر بیت‌ها تأثیر نمی‌گذارد. این عمل در یک GA ساده به منزله‌ی یک عملگر دومی و به منظور نگهداری اطلاعاتی که در حال از دست رفتن است تلقی می‌گردد. عمل جهش برای جلوگیری از همگرایی تند و کمک به الگوریتم جستجو برای فرار از به دام افتادن در کمینه‌های موضعی سودمند است.



شکل ۸) تأثیر عملگر جهش روی جمعیت؛ در اینجا موجودات که شامل بردارهایی از $+/-$ هستند. عملگر تحول با یک احتمال کوچک روی هر بیت در هر موجود اثر می‌کند و $+$ را به $-$ و برعکس تبدیل می‌کند. در بخش بالای شکل جمعیت نخستین و در بخش پایین جمعیت پس از اعمال جهش نشان داده شده است.

اعمال عملگر جهش در GA با احتمالی موسوم به نرخ جهش، P_m ، که براساس شمار بیت‌های جهش‌یافته به دست می‌آید مشخص می‌شود. P_m مقدار کوچکی است و معمولاً برای یک جمعیت با ۳۰ تا ۲۰۰ هموند (عضو)، میان 0.001 تا 0.5 خواهد بود [۴۷]. برای نمونه در رشته‌ی x از جمعیت، به درازای ۲۰ بیت، که در پایهی دوتایی به صورت زیر نشان داده شده است:

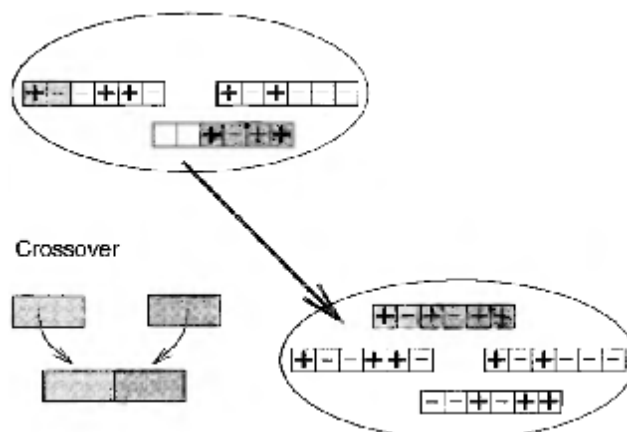
$$x = 00110111011101000011$$

و نرخ جهش $P_m = 0.2$ ، با اعمال عملگر جهش سه بیت شانس تغییر می‌یابند:

$$x = 00111111001101001011$$

۱۱-۶-۴ عملگر ادغام

جستجوی فضای پارامترها و تا حد امکان نگهداری اطلاعات نهفته در رشته‌ها در GA به کمک بخشی به نام تولید مثل^{۱۰۲} صورت می‌گیرد. در این بخش عملگری به نام ادغام^{۱۰۳} که روش‌های اعمال گوناگونی در الگوریتم‌های ژنتیک دارد، عمل تولید مثل را انجام می‌دهد. کلی‌ترین روش ادغام گزینش کاتوره‌ای یک یا چند موجود از جمعیت (والدین) و تقسیم‌کردن آنها از روش‌های گوناگون است تا موجودات جدیدی (فرزندان)، ایجاد شوند. در شکل ۴-۴ روش تولید فرزندان از والدین نشان داده شده است.



شکل ۹) عملگر ادغام، فرزندان با استفاده از قسمت‌های گوناگون والدین ایجاد می‌شوند. در اینجا برای تولید فرزندان تنها یک قسمت از بردار شامل $+/-$ با قسمت دیگری از بردار دوم ترکیب شده است.

با استفاده از روش زیر می‌توان دو فرزند c_1 و c_2 را از دو رشته‌ی x_i و x_j که
 $x_i = 0011011|1011101000011$ ، $x_j = 1011101|0010010100100$
تولید نمود:

یک نقطه‌ی ادغام، S ، که $S \leq P$ (طول رشته) به گونه‌ای کاتوره‌ای گزیده می‌شود. نخستین فرزند، c_1 ، شامل بیت‌های x_i^1 تا x_i^S از والد یکم و بیت‌های x_j^{S+1} تا x_j^P از والد دوم گزیده می‌شود و فرزند دوم، c_2 ، بیت‌های باقیمانده $(x_j^1, \dots, x_j^S, x_i^{S+1}, \dots, x_i^P)$ را بر می‌گزیند. اگر $S = 7$ ، که با خط عمودی در رابطه‌ی ۴-۱۹ مشخص شده است، c_1 و c_2 به صورت زیر هستند:

$$c_1 = 0011011|0010010100100 \quad , \quad c_2 = 1011101|1011101000011$$

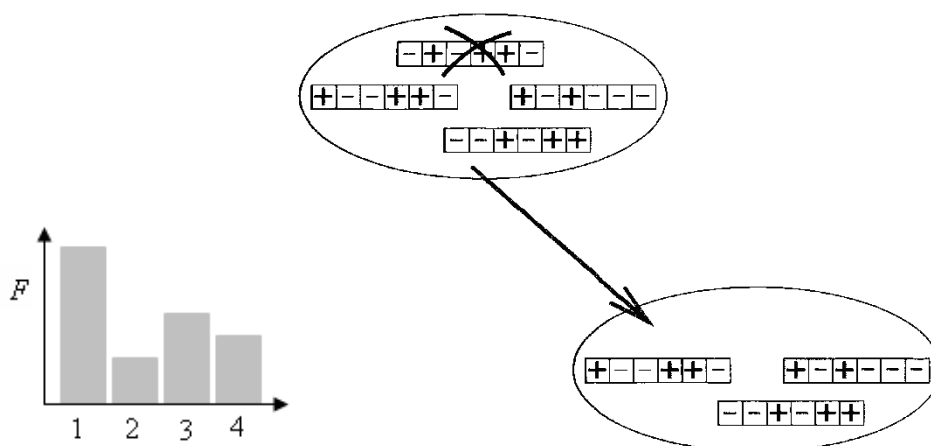
عمل ادغام در GA مانند عمل جهش با نرخی موسوم به نرخ ادغام، P_c ، که بیانگر احتمال ادغام است صورت می‌گیرد. نرخ P_c برای جمعیت‌هایی با اندازه‌ی ۳۰ تا ۲۰۰ موجود در محدوده‌ی ۰/۵ تا ۱ خواهد بود [۴۷].

¹⁰²Reproduction

¹⁰³Crossover

۱۱-۶-۵ عملگر گزینش

جمعیت‌های جدید که در هر گام اجرای GA ایجاد می‌شوند باید به سوی بهتر شدن پیش روند. عملگری که جمعیت‌های جدید را با توجه به مناسب بودن^{۱۰۴} آنها برمی‌گزیند عملگر گزینش^{۱۰۵} نامیده می‌شود. بنابراین نخست باید مناسب بودن هر یک از موجودات جمعیت را بسنجیم. در واقع موجوداتی را در جمعیت نگه داریم که می‌خواهیم ویژگی مورد علاقه‌ی ما (مانند کم بودن مقدار تابع در یافتن کمینه) را داشته باشند. شکل ۴-۵، روش گزینش موجودات و پدید آوردن جمعیت جدید را بر حسب مناسب بودن-شان نشان می‌دهد.



شکل ۹ اثر عملگر گزینش روی جمعیت؛ F مناسب بودن هر یک از موجودات را نشان می‌دهد. موجودات با F کمتر احتمال زنده ماندن کمتری دارند. همان‌گونه که در شکل مشخص است پس از اعمال عملگر گزینش موجود شماره‌ی ۲ از جمعیت کنار می‌رود.

به طور کلی مراحل الگوریتم ژنتیک را می‌توان به صورت زیر خلاصه کرد:

۱. گزینش جمعیت نخستین
۲. کد کردن مقادیر جمعیت در پایه‌ی ۲
۳. گزینش والدین به صورت کاتوره‌ای
۴. آفریدن فرزندان به کمک عملگر ادغام
۵. به کارگیری عملگر جهش
۶. محاسبه‌ی مقدار تابع هدف برای والدین و فرزندان
۷. مقایسه‌ی فرزندان با والدین‌شان بر اساس تابع هدف و گزینش موجوداتی که زنده می‌مانند.

¹⁰⁴Fitness

¹⁰⁵Selection operation

روش‌های گوناگونی برای بیان عملگرهای ادغام ، جهش و گزینش وجود دارد و آنچه در این پایان‌نامه بیان شد شکل بسیار ساده‌ای از ساختار یک الگوریتم GA بود. الگوریتم‌های ژنتیک در بسیاری از مسایل بهینه‌سازی موفق عمل می‌کنند و از این‌رو کاربرد GA در دو دهه‌ی گذشته بسیار افزایش یافته است.

پیوست‌ها

پیوست الف: جمله سوق برای نمونه‌گیری با اهمیت

به وسیله جایگذاری $f(\vec{R}, \tau) = \Psi(\vec{R}, \tau) \Psi_T(\vec{R})$ ، در معادله ۲۷ می‌توانیم یک فرم جدید از معادله پخش برای نمونه‌گیری با اهمیت را استخراج کنیم. Ψ_T یک تابع موج آزمایشی است که در روش مونت کارلو وردشی (VMC) استفاده شد و نقش تابع با اهمیت و یا هدایت کننده را در روش مونت کارلو پخشی (DMC) را بر عهده دارد. با معادله زیر شروع می‌کنیم

$$-D\nabla^2\left(\frac{f}{\Psi_T}\right) + \hat{V}\left(\frac{f}{\Psi_T}\right) = -\frac{1}{\Psi_T} \frac{\partial f}{\partial \tau} \quad (\text{الف-۱})$$

که در آن $D = 1/2$ ، جمله ∇^2 را می‌توان بسط داد و بدست می‌آید:

$$-D\nabla^2 f - 2D\Psi_T \nabla f \cdot \nabla \left(\frac{1}{\Psi_T}\right) - Df \Psi_T \nabla^2 \left(\frac{1}{\Psi_T}\right) + \hat{V}f = -\frac{\partial f}{\partial \tau} \quad (\text{الف-۲})$$

از نتایج زیر استفاده می‌کنیم

$$\nabla\left(\frac{1}{\Psi_T}\right) = -\frac{1}{\Psi_T^2} \nabla \Psi_T \quad (\text{الف-۳})$$

$$\nabla^2\left(\frac{1}{\Psi_T}\right) = -\frac{2}{\Psi_T^3} (\nabla \Psi_T)^2 - \frac{1}{\Psi_T^2} \nabla^2 \Psi_T \quad (\text{الف-۴})$$

در نتیجه می‌توان رابطه زیر را استخراج کرد

$$-D\nabla^2 f - \nabla f \cdot (2D \frac{\nabla \Psi_T}{\Psi_T}) - 2Df \left(\frac{\nabla \Psi_T}{\Psi_T}\right)^2 + Df \left(\frac{\nabla^2 \Psi_T}{\Psi_T}\right) + \hat{V}f = -\frac{\partial f}{\partial \tau} \quad (\text{الف-۵})$$

سرعت سوق را به صورت $\vec{F}(\vec{R}) = D \nabla \ln |\Psi_T(\vec{R})|^2 = \frac{2D \nabla \Psi_T(\vec{R})}{\Psi_T(\vec{R})}$ تعریف می‌شود، و از

$$\text{اتحاد } (\nabla f) \cdot F = \nabla \cdot (fF) - f \nabla \cdot F \quad (\text{الف-۶})$$

$$-D\nabla^2 f - \nabla \cdot (fF) - 2Df \nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T}\right) - 2Df \left(\frac{\nabla \Psi_T}{\Psi_T}\right)^2 + Df \left(\frac{\nabla^2 \Psi_T}{\Psi_T}\right) + \hat{V}f = -\frac{\partial f}{\partial \tau}$$

از معادله زیر، جایگزینی برای جمله $\nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T}\right)$ بدست می‌آوریم و در معادله قبلی جایگذاری

می‌کنیم که منجر به حذف شدن جملاتی و حاصل شدن معادله (۸-۴) می‌گردد.

$$\nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T}\right) = \frac{\nabla^2 \Psi_T}{\Psi_T} - \left(\frac{\nabla \Psi_T}{\Psi_T}\right)^2 \quad (\text{الف-۸})$$

$$-D\nabla^2 f + \nabla \cdot (fF) + [-D(\frac{\nabla^2 \Psi_T}{\Psi_T}) + \frac{\hat{V}\Psi_T}{\Psi_T}]f = -\frac{\partial f}{\partial \tau} \quad (\text{الف-۹})$$

با استفاده از عبارت مربوط به هامیلتونی $\hat{H}$ ، انرژی موضعی را به صورت زیر تعریف می‌کنیم:

$$\frac{\hat{H}\Psi_T}{\Psi_T} = -D\frac{\nabla^2 \Psi_T}{\Psi_T} + \frac{\hat{V}\Psi_T}{\Psi_T} = E_L$$

بنابراین به معادله ۴۹ می‌رسیم :

$$\frac{\partial f}{\partial \tau} = -D\nabla^2 f + E_L f + \bar{\nabla} \cdot (f \bar{F})$$

پیوست ب: محاسبه‌ی گرادیان و لاپلاسی دترمینان‌های اسلیتر و عامل جسترو

ب-۱) محاسبه‌ی گرادیان و لاپلاسی دترمینان‌های اسلیتر

اوربیتال‌های φ که به فاصله‌ی هر الکترون از هسته، r ، بستگی دارند به شکل زیر در دترمینان‌های اسلیتر قرار می‌گیرند:

$$D \equiv \begin{vmatrix} \varphi_1(r_1) & \cdots & \varphi_1(r_i) & \cdots & \varphi_1(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_i(r_1) & \cdots & \varphi_i(r_i) & \cdots & \varphi_i(r_N) \\ \vdots & & \vdots & & \vdots \\ \varphi_N(r_1) & \cdots & \varphi_N(r_i) & \cdots & \varphi_N(r_N) \end{vmatrix} \quad (\text{ب-۱})$$

برای محاسبه‌ی گرادیان و لاپلاسی دترمینان بالا نسبت به ذره‌ی i ام، آن را به صورت زیر باز می‌کنیم:

$$\sum_j \varphi_j(r_i) M_{ji}(r_1, \dots, r_N) \quad (\text{ب-۲})$$

که M_{ij} ، ماتریس همسازهی درایه‌ی i ام و j ام است و دیگر به r_i بستگی ندارد. از این‌رو عملگر گرادیان و لاپلاسی به صورت زیر در رابطه‌ی الف-۲، اثر می‌کند:

$$\sum_j M_{ji}(r_1, \dots, r_N) \vec{\nabla}_i \varphi_j(r_i), \quad (\text{ب-۳})$$

$$\sum_j M_{ji}(r_1, \dots, r_N) \nabla_i^2 \varphi_j(r_i)$$

در واقع می‌توان رابطه‌ی الف-۳ را برای گرادیان و لاپلاسی $D^\uparrow$ به صورت زیر نوشت:

$$\vec{\nabla}_i D^\uparrow = \begin{vmatrix} \varphi_{1s}(\vec{r}_1) & \cdots & \vec{\nabla}_i \varphi_{1s}(\vec{r}_i) & \cdots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \cdots & \vec{\nabla}_i \varphi_{2s}(\vec{r}_i) & \cdots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \varphi_{px}(\vec{r}_1) & \cdots & \vec{\nabla}_i \varphi_{px}(\vec{r}_i) & \cdots & \varphi_{px}(\vec{r}_{N_{up}}) \\ \varphi_{py}(\vec{r}_1) & \cdots & \vec{\nabla}_i \varphi_{py}(\vec{r}_i) & \cdots & \varphi_{py}(\vec{r}_{N_{up}}) \\ \varphi_{pz}(\vec{r}_1) & \cdots & \vec{\nabla}_i \varphi_{pz}(\vec{r}_i) & \cdots & \varphi_{pz}(\vec{r}_{N_{up}}) \\ \vdots & & \vdots & & \vdots \end{vmatrix}, \quad \nabla_i^2 D^\uparrow = \begin{vmatrix} \varphi_{1s}(\vec{r}_1) & \cdots & \nabla_i^2 \varphi_{1s}(\vec{r}_i) & \cdots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \cdots & \nabla_i^2 \varphi_{2s}(\vec{r}_i) & \cdots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \varphi_{px}(\vec{r}_1) & \cdots & \nabla_i^2 \varphi_{px}(\vec{r}_i) & \cdots & \varphi_{px}(\vec{r}_{N_{up}}) \\ \varphi_{py}(\vec{r}_1) & \cdots & \nabla_i^2 \varphi_{py}(\vec{r}_i) & \cdots & \varphi_{py}(\vec{r}_{N_{up}}) \\ \varphi_{pz}(\vec{r}_1) & \cdots & \nabla_i^2 \varphi_{pz}(\vec{r}_i) & \cdots & \varphi_{pz}(\vec{r}_{N_{up}}) \\ \vdots & & \vdots & & \vdots \end{vmatrix} \quad (\text{ب-۴})$$

گرادیان و لاپلاسی $D^\downarrow$ نیز به صورت بالا به دست می‌آید. با این تفاوت که اوربیتال‌ها در $D^\downarrow$ به مختصات $\vec{r}_{N_{up}+1}$ تا $\vec{r}_N$ بستگی دارند. بنابراین نخست باید گرادیان و لاپلاسی اوربیتال‌های اتمی را محاسبه کنیم و سپس گرادیان و لاپلاسی دترمینان‌های اسلیتر را برای هر ذره به دست آوریم.

ب-۲) گرادیان تابع جسترو اسمیت و ماسکویتز

در اینجا تنها به محاسبه‌ی گرادیان تابع جسترو اسمیت و ماسکویتز که شکل پیچیده‌تری دارد می‌پردازیم. با توجه به شکل این تابع جسترو، رابطه‌ی ۵-۴، عامل جسترو، J ، به صورت زیر است:

$$J = \sum_{i < j} \left\{ C_1 \left(\frac{r_{ij}}{1+r_{ij}} \right) + C_2 \left(\frac{r_{ij}}{1+r_{ij}} \right)^2 + C_3 \left(\frac{r_{ij}}{1+r_{ij}} \right)^3 + C_4 \left(\frac{r_{ij}}{1+r_{ij}} \right)^4 + \right. \\ \left. C_5 \left[\left(\frac{r_i}{1+r_i} \right)^2 + \left(\frac{r_j}{1+r_j} \right)^2 \right] + C_6 \left[\left(\frac{r_i}{1+r_i} \right)^3 + \left(\frac{r_j}{1+r_j} \right)^3 \right] + \right. \\ \left. C_7 \left[\left(\frac{r_i}{1+r_i} \right)^4 + \left(\frac{r_j}{1+r_j} \right)^4 \right] + 2C_8 \left[\left(\frac{r_i}{1+r_i} \right)^2 \left(\frac{r_j}{1+r_j} \right)^2 \right] + \right. \\ \left. C_9 \left[\left(\frac{r_i}{1+r_i} \right)^2 + \left(\frac{r_j}{1+r_j} \right)^2 \right] \left(\frac{r_{ij}}{1+r_{ij}} \right)^2 \right\} \quad (\text{ب-۵})$$

گرادیان نسبت به ذره‌ی k ام را می‌توان به صورت حاصل جمع دو جمله‌ی زیر نوشت:

$$\vec{\nabla}_k(J) = \vec{\nabla}_k \left(\sum_{k < j} \dots \right) + \vec{\nabla}_k \left(\sum_{k > j} \dots \right) \quad (\text{ب-۶})$$

گرادیان جملات وابسته به r_{kj} یا r_{jk} را می‌توان با استفاده از رابطه‌ی زیر به دست آورد:

$$\vec{\nabla}_k(\dots) = \frac{\partial(\dots)}{\partial x_k} \hat{i} + \frac{\partial(\dots)}{\partial y_k} \hat{j} + \frac{\partial(\dots)}{\partial z_k} \hat{k} \\ = \frac{\partial(\dots)}{\partial r_{kj}} \left[\frac{\partial r_{kj}}{\partial x_k} \hat{i} + \frac{\partial r_{kj}}{\partial y_k} \hat{j} + \frac{\partial r_{kj}}{\partial z_k} \hat{k} \right] \\ = \frac{\partial(\dots)}{\partial r_{kj}} \hat{r}_{jk} \quad (\text{ب-۷})$$

با انجام محاسبات گرادیان جمله‌ی $\vec{\nabla}_k \left(\sum_{k < j} \dots \right)$ در تابع جسترو به صورت زیر به دست می‌آید:

$$\vec{\nabla}_k \left(\sum_{k < j} \dots \right) = \sum_{k < j} \frac{\vec{r}_{kj}}{(1+r_{kj})^3} \left(\frac{C_1(1+r_{kj})}{r_{kj}} + 2C_2 + 3C_3 \frac{r_{kj}}{(1+r_{kj})} + 4C_4 \left(\frac{r_{kj}}{1+r_{kj}} \right)^2 + 2C_9 \left[\left(\frac{r_k}{1+r_k} \right)^2 + \left(\frac{r_j}{1+r_j} \right)^2 \right] \right) \\ + \frac{\vec{r}_{jk}}{(1+r_k)^3} \left(2C_5 + 3C_6 \frac{r_k}{(1+r_k)} + 4C_7 \left(\frac{r_k}{1+r_k} \right)^2 + 4C_8 \left(\frac{r_j}{1+r_j} \right)^2 + 2C_9 \left(\frac{r_{kj}}{1+r_{kj}} \right)^2 \right) \quad (\text{ب-۸})$$

جمله‌ی $\vec{\nabla}_k \left(\sum_{j > k} \dots \right)$ نیز به همین صورت به دست می‌آید، تنها باید تبدیل $\vec{r}_{jk} \rightarrow \vec{r}_{kj}$ را اعمال کنیم.

ب-۳) لاپلاسی تابع جسترو اسمیت و ماسکوینز

لاپلاسی J ، نسبت به ذره‌ی k ام را نیز به صورت جمع دو جمله‌ی زیر محاسبه می‌کنیم:

$$\nabla_k^2(J) = \nabla_k^2 \left(\sum_{k < j} \dots \right) + \nabla_k^2 \left(\sum_{j < k} \dots \right) \quad (\text{ب-۹})$$

با توجه به رابطه‌ی (الف-۴)، لاپلاسی جملات وابسته به r_{kj} یا r_{jk} مانند رابطه‌ی لاپلاسی در مختصات کروی به صورت زیر است:

$$\nabla_k^2(\dots) = \frac{\partial^2(\dots)}{\partial r_{kj}^2} + \frac{2}{r_{kj}} \frac{\partial(\dots)}{\partial r_{kj}} \quad (\text{ب-۱۰})$$

با انجام محاسبات لاپلاسی جمله‌ی $\nabla_k^2 \left(\sum_{k < j} \dots \right)$ به صورت زیر به دست می‌آید:

$$\nabla_k^2 \left(\sum_{k < j} \dots \right) = \sum_{k < j} \frac{1}{(1+r_{kj})^4} \left(\frac{2C_1(1+r_{kj})}{r_{kj}} + 6C_2 + 12C_3 \frac{r_{kj}}{(1+r_{kj})} + 20C_4 \left(\frac{r_{kj}}{1+r_{kj}} \right)^2 + 6C_9 \left[\left(\frac{r_k}{1+r_k} \right)^2 + \left(\frac{r_j}{1+r_j} \right)^2 \right] \right) + \\ \frac{1}{(1+r_k)^4} \left(6C_5 + 12C_6 \frac{r_k}{(1+r_k)} + 20C_7 \left(\frac{r_k}{1+r_k} \right)^2 + 12C_8 \left(\frac{r_j}{1+r_j} \right)^2 + 6C_9 \left(\frac{r_{kj}}{1+r_{kj}} \right)^2 \right) + \\ \frac{6C_9(r_k^2 - x_k x_j - y_k y_j - z_k z_j)}{(1+r_k)^3 (1+r_{kj})^3} \quad (\text{ب-۱۱})$$

جمله‌ی $\nabla_k^2 \left(\sum_{j < k} \dots \right)$ نیز به صورت رابطه‌ی الف-۱۱، به دست می‌آید.

پیوست (ج): محاسبه‌ی مشتق یکم و دوم دترمینان‌های اسلیتر نسبت به پارامترها

ج-۱) مشتق یکم دترمینان‌های اسلیتر نسبت به پارامترها

$D^\uparrow$ و $D^\downarrow$ به صورت زیر هستند:

$$D^\uparrow = \begin{vmatrix} \varphi_{1s}(\vec{r}_1) & \varphi_{1s}(\vec{r}_2) & \cdots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \varphi_{2s}(\vec{r}_2) & \cdots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \varphi_{px}(\vec{r}_1) & \varphi_{px}(\vec{r}_2) & \cdots & \varphi_{px}(\vec{r}_{N_{up}}) \\ \varphi_{py}(\vec{r}_1) & \varphi_{py}(\vec{r}_2) & \cdots & \varphi_{py}(\vec{r}_{N_{up}}) \\ \varphi_{pz}(\vec{r}_1) & \varphi_{pz}(\vec{r}_2) & \cdots & \varphi_{pz}(\vec{r}_{N_{up}}) \\ \vdots & & \ddots & \vdots \end{vmatrix} \quad (\text{ج-۱})$$

۹

$$D^\downarrow = \begin{vmatrix} \varphi_{1s}(\vec{r}_{N_{up}+1}) & \varphi_{1s}(\vec{r}_{N_{up}+2}) & \cdots & \varphi_{1s}(\vec{r}_N) \\ \varphi_{2s}(\vec{r}_{N_{up}+1}) & \varphi_{2s}(\vec{r}_{N_{up}+2}) & \cdots & \varphi_{2s}(\vec{r}_N) \\ \varphi_{px}(\vec{r}_{N_{up}+1}) & \varphi_{px}(\vec{r}_{N_{up}+2}) & \cdots & \varphi_{px}(\vec{r}_N) \\ \varphi_{py}(\vec{r}_{N_{up}+1}) & \varphi_{py}(\vec{r}_{N_{up}+2}) & \cdots & \varphi_{py}(\vec{r}_N) \\ \varphi_{pz}(\vec{r}_{N_{up}+1}) & \varphi_{pz}(\vec{r}_{N_{up}+2}) & \cdots & \varphi_{pz}(\vec{r}_N) \\ \vdots & & \ddots & \vdots \end{vmatrix} \quad (\text{ج-۲})$$

که $D^\uparrow$ الکترون‌های با اسپین بالا و $D^\downarrow$ الکترون‌های با اسپین پایین را شامل می‌شود. اینک مشتق یکم $D^\uparrow$ نسبت به پارامترها را می‌توانیم به صورت زیر بنویسیم:

$$\frac{\partial D^\uparrow}{\partial c_n} = \frac{\partial}{\partial c_n} \begin{vmatrix} \varphi_{1s}(\vec{r}_1) & \varphi_{1s}(\vec{r}_2) & \cdots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \varphi_{2s}(\vec{r}_2) & \cdots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \varphi_{px}(\vec{r}_1) & \varphi_{px}(\vec{r}_2) & \cdots & \varphi_{px}(\vec{r}_{N_{up}}) \\ \vdots & & \ddots & \vdots \end{vmatrix} \quad (\text{ج-۳})$$

$\frac{\partial}{\partial c_n}$ روی درایه‌های هر ستون دترمینان اثر می‌کند، زیرا تمام درایه‌ها به پارامترها بستگی دارند و رابطه‌ی ب-۳ به صورت حاصل جمع دترمینان‌های زیر به دست می‌آید:

$$\begin{aligned}
 & \left| \begin{array}{cccc} \frac{\partial \varphi_{1s}(\vec{r}_1)}{\partial c_n} & \varphi_{1s}(\vec{r}_2) & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \frac{\partial \varphi_{2s}(\vec{r}_1)}{\partial c_n} & \varphi_{2s}(\vec{r}_2) & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & & \vdots \end{array} \right| + \left| \begin{array}{cccc} \varphi_{1s}(\vec{r}_1) & \frac{\partial \varphi_{1s}(\vec{r}_2)}{\partial c_n} & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \frac{\partial \varphi_{2s}(\vec{r}_2)}{\partial c_n} & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & & \vdots \end{array} \right| + \dots \\
 & + \frac{\partial}{\partial c_m} \left| \begin{array}{cccc} \varphi_{1s}(\vec{r}_1) & \varphi_{1s}(\vec{r}_2) & \dots & \frac{\partial \varphi_{1s}(\vec{r}_{N_{up}})}{\partial c_n} \\ \varphi_{2s}(\vec{r}_1) & \varphi_{2s}(\vec{r}_2) & \dots & \frac{\partial \varphi_{2s}(\vec{r}_{N_{up}})}{\partial c_n} \\ \vdots & \vdots & & \vdots \end{array} \right|
 \end{aligned} \quad (ج-۴)$$

$\frac{\partial D^\downarrow}{\partial c_n}$ نیز به صورت بالا به دست می‌آید. توجه داریم که اوربیتال‌ها در $D^\downarrow$ به مختصات $\vec{r}_{N_{up}+1}$ تا $\vec{r}_N$ بستگی دارند.

ج-۲) مشتق دوم دترمینان‌های اسلیتر نسبت به پارامترها

مشتق دوم $D^\downarrow$ نسبت به پارامترهای ورودش، $\frac{\partial^2 D^\uparrow}{\partial c_m \partial c_n}$ ، با مشتق‌گیری از رابطه‌ی ج-۴ به دست می‌آید:

$$\begin{aligned}
 \frac{\partial^2 D^\uparrow}{\partial c_m \partial c_n} &= \frac{\partial}{\partial c_m} \frac{\partial D^\uparrow}{\partial c_n} \\
 &= \frac{\partial}{\partial c_m} \left| \begin{array}{cccc} \frac{\partial \varphi_{1s}(\vec{r}_1)}{\partial c_n} & \varphi_{1s}(\vec{r}_2) & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \frac{\partial \varphi_{2s}(\vec{r}_1)}{\partial c_n} & \varphi_{2s}(\vec{r}_2) & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & & \vdots \end{array} \right| + \frac{\partial}{\partial c_m} \left| \begin{array}{cccc} \varphi_{1s}(\vec{r}_1) & \frac{\partial \varphi_{1s}(\vec{r}_2)}{\partial c_n} & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \varphi_{2s}(\vec{r}_1) & \frac{\partial \varphi_{2s}(\vec{r}_2)}{\partial c_n} & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & & \vdots \end{array} \right| + \dots \\
 & + \frac{\partial}{\partial c_m} \left| \begin{array}{cccc} \varphi_{1s}(\vec{r}_1) & \varphi_{1s}(\vec{r}_2) & \dots & \frac{\partial \varphi_{1s}(\vec{r}_{N_{up}})}{\partial c_n} \\ \varphi_{2s}(\vec{r}_1) & \varphi_{2s}(\vec{r}_2) & \dots & \frac{\partial \varphi_{2s}(\vec{r}_{N_{up}})}{\partial c_n} \\ \vdots & \vdots & & \vdots \end{array} \right|
 \end{aligned} \quad (ج-۵)$$

حال باید الگوریتمی بسازیم که هر قسمت را در رابطه بالا محاسبه کند و در نهایت سهم $\frac{\partial}{\partial c_m}$ از تمام دترمینان‌ها را با هم جمع کند، به عنوان نمونه بخش نخست رابطه‌ی بالا بدین صورت به دست می‌آید:

$$\begin{aligned}
& \frac{\partial}{\partial c_m} \begin{vmatrix} \frac{\partial \varphi_{1s}(\vec{r}_1)}{\partial c_n} & \varphi_{1s}(\vec{r}_2) & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \frac{\partial \varphi_{2s}(\vec{r}_1)}{\partial c_n} & \varphi_{2s}(\vec{r}_2) & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} = \\
& \begin{vmatrix} \frac{\partial^2 \varphi_{1s}(\vec{r}_1)}{\partial c_m \partial c_n} & \varphi_{1s}(\vec{r}_2) & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \frac{\partial^2 \varphi_{2s}(\vec{r}_1)}{\partial c_m \partial c_n} & \varphi_{2s}(\vec{r}_2) & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} + \begin{vmatrix} \frac{\partial \varphi_{1s}(\vec{r}_1)}{\partial c_n} & \frac{\partial^2 \varphi_{1s}(\vec{r}_2)}{\partial c_m \partial c_n} & \dots & \varphi_{1s}(\vec{r}_{N_{up}}) \\ \frac{\partial \varphi_{2s}(\vec{r}_1)}{\partial c_n} & \frac{\partial^2 \varphi_{2s}(\vec{r}_2)}{\partial c_m \partial c_n} & \dots & \varphi_{2s}(\vec{r}_{N_{up}}) \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} + \dots \\
& + \begin{vmatrix} \frac{\partial \varphi_{1s}(\vec{r}_1)}{\partial c_n} & \varphi_{1s}(\vec{r}_2) & \dots & \frac{\partial^2 \varphi_{1s}(\vec{r}_{N_{up}})}{\partial c_m \partial c_n} \\ \frac{\partial \varphi_{2s}(\vec{r}_1)}{\partial c_n} & \varphi_{2s}(\vec{r}_2) & \dots & \frac{\partial^2 \varphi_{2s}(\vec{r}_{N_{up}})}{\partial c_m \partial c_n} \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} \\
& \text{روند محاسبه‌ی } \frac{\partial^2 D^\downarrow}{\partial c_m \partial c_n} \text{ نیز مانند } \frac{\partial^2 D^\uparrow}{\partial c_m \partial c_n} \text{ است.}
\end{aligned}$$

ج-۳) محاسبه‌ی مشتق انرژی موضعی نسبت به پارامترهای ورودش

با توجه به رابطه‌ی ۹۹ برای محاسبه‌ی $\frac{\partial E_L}{\partial c_n}$ ، باید $\frac{\partial \nabla_i^2 \ln \Psi}{\partial c_n}$ و $\frac{\partial \vec{\nabla}_i \ln \Psi}{\partial c_n}$ محاسبه کنیم. از آنجا که تابع موج QMC به شکل $\Psi = D^\uparrow D^\downarrow e^J$ است، از رابطه‌ی ۵-۱۴ داریم:

$$\begin{aligned}
\vec{\nabla}_i \ln \Psi &= \frac{1}{D^\uparrow} \vec{\nabla}_i D^\uparrow + \frac{1}{D^\downarrow} \vec{\nabla}_i D^\downarrow + \vec{\nabla}_i J, \\
\frac{\partial \vec{\nabla}_i \ln \Psi}{\partial c_n} &= -\frac{1}{(D^\uparrow)^2} \frac{\partial D^\uparrow}{\partial c_n} \vec{\nabla}_i D^\uparrow + \frac{1}{D^\uparrow} \frac{\partial \vec{\nabla}_i D^\uparrow}{\partial c_n} - \frac{1}{(D^\downarrow)^2} \frac{\partial D^\downarrow}{\partial c_n} \vec{\nabla}_i D^\downarrow + \frac{1}{D^\downarrow} \frac{\partial \vec{\nabla}_i D^\downarrow}{\partial c_n} + \frac{\partial \vec{\nabla}_i J}{\partial c_n}
\end{aligned}$$

(ب-۶)

و از رابطه‌ی ۱۰۶، برای $\nabla_i^2 \ln \Psi$ داریم:

$$\begin{aligned} \nabla_i^2 \ln \Psi = & -\left(\frac{1}{D^\uparrow} \vec{\nabla}_i D^\uparrow\right)^2 + \frac{1}{D^\uparrow} \nabla_i^2 D^\uparrow - \left(\frac{1}{D^\downarrow} \vec{\nabla}_i D^\downarrow\right)^2 + \frac{1}{D^\downarrow} \nabla_i^2 D^\downarrow + \nabla_i^2 J \\ \frac{\partial \nabla_i^2 \ln \Psi}{\partial c_n} = & \frac{2}{(D^\uparrow)^3} \frac{\partial D^\uparrow}{\partial c_n} |\vec{\nabla}_i D^\uparrow|^2 - \frac{2}{(D^\uparrow)^2} \left(\frac{\partial \vec{\nabla}_i D^\uparrow}{\partial c_n}\right) \cdot (\vec{\nabla}_i D^\uparrow) - \frac{1}{(D^\uparrow)^2} \frac{\partial D^\uparrow}{\partial c_n} \cdot (\nabla_i^2 D^\uparrow) + \frac{1}{D^\uparrow} \frac{\partial \nabla_i^2 D^\uparrow}{\partial c_n} \\ & + \frac{2}{(D^\downarrow)^3} \frac{\partial D^\downarrow}{\partial c_n} |\vec{\nabla}_i D^\downarrow|^2 - \frac{2}{(D^\downarrow)^2} \left(\frac{\partial \vec{\nabla}_i D^\downarrow}{\partial c_n}\right) \cdot (\vec{\nabla}_i D^\downarrow) - \frac{1}{(D^\downarrow)^2} \frac{\partial D^\downarrow}{\partial c_n} \cdot (\nabla_i^2 D^\downarrow) + \frac{1}{D^\downarrow} \frac{\partial \nabla_i^2 D^\downarrow}{\partial c_n} \\ & + \frac{\nabla_i^2 J}{\partial c_n} \end{aligned}$$

(ج-۷)

روش محاسبه‌ی مشتق دترمینان‌ها در بخش‌های ج-۱ و ج-۲، بیان شد. در محاسبه‌ی مشتقات $\frac{\partial \nabla_i^2 J}{\partial c_m}$ و $\frac{\partial \vec{\nabla}_i J}{\partial c_m}$ ، $\frac{\partial \nabla_i^2 D^{\uparrow,\downarrow}}{\partial c_m}$ ، $\frac{\partial \vec{\nabla}_i D^{\uparrow,\downarrow}}{\partial c_m}$ توجه داریم که نخست باید مؤلفه‌های ماتریس گرادیان نسبت به الکترون i ام را بسازیم سپس مشتق را نسبت به پارامترهای ورودش محاسبه می‌کنیم.