

計算科学ロードマップ 2023

HPCI コンソーシアム計算科学フォーラム

2023 年 12 月 (第 0 版)

改定履歴

2023年12月26日 第0版

目次

第 1 章	序論	1
1.1	はじめに	2
第 2 章	各計算科学分野の課題	5
2.1	素粒子・原子核	6
2.2	ナノサイエンス・デバイス	18
2.3	エネルギー・材料	27
2.4	生命科学	31
2.5	脳・神経科学	50
2.6	創薬・医療	67
2.7	設計・製造	74
2.8	社会科学	90
2.9	地震・津波	99
2.10	気象・気候	108
2.11	宇宙・天文	120
2.12	人工知能	126
2.13	計算機要求性能まとめ	127
第 3 章	アプリケーションの分類	131
第 4 章	各課題の詳細	133
4.1	素粒子・原子核	134
4.2	ナノサイエンス・デバイス	176
4.3	エネルギー・材料	180
4.4	生命科学	209
4.5	脳・神経科学	232
4.6	創薬・医療	249
4.7	設計・製造	274
4.8	社会科学	306
4.9	地震・津波	307

4.10	気象・気候	308
4.11	宇宙・天文	355
付録 A	用語集	379
付録 B	執筆者一覧	433

第1章

序論

1.1 はじめに

現代の科学技術における知識の獲得、発見には、スーパーコンピュータは必須となっている。同時に、スーパーコンピュータは一般市民の毎日の生活を陰で支えてもいる。わが国は現在、新型コロナウィルス感染症、エネルギー問題、少子高齢化、不安定化する国際情勢、財政逼迫など山積する難題に直面している。スーパーコンピュータによる大規模シミュレーションは、科学技術を牽引するとともに、我々が直面しているこれらの困難な課題解決にも重要な役割を果たしている。日本社会を力強く支え、明日の時代を切り開くためにスーパーコンピュータは不可欠の基盤技術である。

大規模かつ高性能なスーパーコンピュータ (HPC) により得られる研究成果や研究手法は、科学技術の最先端をさらに伸ばし、次の時代には、企業自身がスーパーコンピュータを導入することによる産業活性化に展開し得る。また、医療の現場や気象予報などの現業におけるスキルの大幅な向上に繋がる可能性がある。HPC で培った技術が、最終的に産業や社会の現場で利用されること、すなわち計算科学の下方展開の重要性は今後ますます増加、シミュレーションや大規模データ処理の果たす役割はさらに拡大し、その結果は社会に大きな恩恵をもたらすであろう。

基盤技術としてのスーパーコンピュータが重要性を増す中で、2011年、HPCI 計画の推進にあたり国として今後の HPC 研究開発に必要な事項等を検討するため、文部科学省研究振興局長の諮問会議「HPCI 計画推進委員会」のもとに「今後の HPC 技術の研究開発のあり方を検討する WG」が設置された。そして、同 WG からの提言により「アプリケーション作業部会」と「コンピュータアーキテクチャ・コンパイラ・システムソフトウェア作業部会」が設置され、両者の緊密な連携のもと 2012 年 3 月に「計算科学ロードマップ白書」が取りまとめられた^{*1}。

その後、作業部会での議論の更なる精査をめざし、文部科学省委託研究「将来の HPCI のあり方の調査研究(アプリケーション分野)」(略称:「アプリ FS」)が 2012 年 7 月にスタートした。アプリ FS では、計算科学が貢献し得る社会的課題・科学的ブレークスルーの課題抽出が行われ、その成果として新たな「計算科学ロードマップ」がまとめられ、2014 年 3 月に公開された^{*2}。

2015 年には、作業部会・アプリ FS が築いてきた計算科学コミュニティを維持すべく、「計算科学ロードマップ」の各分野の執筆取りまとめ役を中心に「今後の HPCI を使った計算科学発展のための検討会」(略称:「計算科学検討会」)が発足した。計算科学検討会では、計算科学分野における横断的なコミュニティの育成・維持・発展、若手の発掘、情報交換・成果発表の場の創出、計算科学のための計算資源の確保を目指し、新規分野を含む「計算科学ロードマップ」の改定作業のための会合、分野横断的なオープンな学術的会合「計算科学フォーラム」が定期的に開催されてきた。その成果として、「計算科学ロードマップ」の改訂版である「計算科学ロードマップ 2017」が 2017 年 6 月に公開された^{*3}。計算科学検討会は 2014 年度から 2016 年度の間、理化学研究所の受託した文部科学省委託事業「HPCI の運営」の下で運営された。2017 年度からは HPCI コンソーシアムの下で「計算科学フォーラム」として再出発し、より広いコミュニティの中で多くの研究機関との連携を持ちながらの分野横断的な活動を継続している。

^{*1} <https://sites.google.com/view/ngaci/home/sdhpc>

^{*2} <https://cs-forum.github.io/hpci-aplfs>

^{*3} <https://cs-forum.github.io/kentoukai/>

本書は、計算科学フォーラムを核とする計算科学コミュニティでの議論を通じて抽出された、2027年から2032年頃に計算科学が解決すべき社会的課題・期待される科学的ブレークスルーと、そのために必要となる計算機システム性能等をとりまとめたものである。本書は、共用開始から2年半が経過したスーパーコンピュータ「富岳」により得られた新しい知見や予想に基づき、「計算科学ロードマップ2017」の内容を精査・更新し、計算科学の新しいシミュレーション手法や新しい分野を積極的に取り込むことを目指した。また近年、大規模言語モデルを代表とする、AI・機械学習の急速な発展により、スーパーコンピュータに要求される性能だけでなく、HPCや計算科学のあり方さえも大きく変わろうとしている。このような状況を鑑み、今回のロードマップでは新たな節を設け、計算科学の各分野で活用が進みつつあるAI技術について分野横断的に俯瞰し、HPCとデータサイエンスをどのように融合していくべきかについても議論を行った。

本書をまとめるにあたり、将来想定される計算機アーキテクチャーとその性能については、NGACI(次世代先端的計算基盤)により2020年にまとめられた「次世代先端的計算基盤に関する白書」を基本とした^{*4}。白書では、2028年時点でのスーパーコンピュータシステムのピーク性能は2EF～20EF程度と見積もられている。しかしながら、本書が示す大規模計算科学の可能性は、将来サイエンスとして何が必要かを純粋に議論して明らかにするものであり、決して最高性能のトップマシン(フラッグシップ)だけを意識したものではない。フラッグシップだけではなく、第二階層・第三階層システムを含む次世代計算基盤のアーキテクチャに関する提言に資する情報を提供できるものと考える。

なお、今回の版(2023年12月版)は暫定版であり、未完成の節・章が残っている。特に、2.8節は「計算科学ロードマップ2017」の2.7節を転記したものである。現在執筆を進めており、近々更新予定である。2.12節については、現在並行して理化学研究所計算科学研究センターを中心にAI-for-Science(AI技術の基礎科学研究への活用)の取りまとめが進んでおり、2024年3月をめどに本書と統合の予定である。さらに、3章については、2章の全ての節について計算機要求性能が出揃った時点で、演算速度、必要なメモリ量、メモリバンド幅、ノード間通信バンド幅、ノード間通信レイテンシなどの観点からアプリケーションの要求を概観し、アプリケーションの分類を行う予定である。更新版については計算科学フォーラムのWebページで順次公開する^{*5}。

^{*4} <https://sites.google.com/view/ngaci/home>

^{*5} <https://hpcic-kkf.com/>

第2章

各計算科学分野の課題

2.1 素粒子・原子核

2.1.1 分野の概要

我々が生きている自然界には、宇宙から極微の世界まで、多くの階層が存在する。素粒子分野では自然界の最小構成要素である素粒子の階層を探求し基礎法則の確立を目指している。原子核分野における原子核の階層の探求は、素粒子の世界から原子、元素の世界を説明することを可能にする。このようなミクロな階層の探求は宇宙の成り立ちの探求につながっている。また、素粒子・原子核分野での研究活動は歴史的に計算科学の発展を促してきた。現在においても以下で述べるように計算科学的手法を大いに活用する分野となっている。

素粒子分野

素粒子分野の目的はより深く素粒子の階層を探ることによってその世界を記述するより根源的かつ統一的な基本法則を探求するとともに、得られた基本法則を用いて宇宙の始まりから現在に至るまでの進化の過程を理解することにある。

素粒子の階層を含めた全ての自然界は強い力、電磁気力、弱い力、重力の4つの力によって支配されており、重力を除く3つの力は「素粒子標準理論」（以下、標準理論）と呼ばれる体系によって記述できることが知られている。しかしながら、標準理論に含まれる多くのパラメータの起源を説明できない問題や標準理論には量子重力が含まれていない問題など、標準理論の枠を超えた問題が存在している。これらを解決する新しい自然階層の存在が期待されており^{*1}、その手がかりを探す活発な研究が理論・実験の両面から世界的に進められている。この新しい階層を探す方法は大きく分けて、未知の領域を直接探索するアプローチと既知の物理現象の高精度検証によるアプローチの2つがある。2012年にヒッグス粒子を発見した欧州原子核研究機構（CERN）の大型ハドロン衝突型加速器（Large Hadron Collider : LHC）での粒子加速器実験は未知の領域を直接探索するアプローチの典型例であり、加速粒子の衝突エネルギーを上げることによって新しい粒子・物理現象の発見を目指している。一方、高エネルギー加速器研究機構（KEK）におけるBelle II実験やKEKと日本原子力研究開発機構（JAEA）が共同で建設したJ-PARCにおけるミュー粒子実験は、既知の物理現象を高精度に測定することを目的としている。このような高精度実験と標準理論の精密計算との比較から新しい階層の手がかりを得ることが高精度検証によるアプローチである。これら両アプローチにおいて計算科学は重要な役割を果たしている。実験におけるデータ解析では情報科学の手法や計算機科学の技術が利用されておりCERNのデータグリッド（Worldwide LHC Computing Grid : WLCG）はその一例である。理論においては新しい階層の理論候補の探索や実験精度に比する高精度の計算に計算科学的手法が必要不可欠となっている。

一方で新しい階層ではなく現在の標準理論から現在の宇宙の有り様を理論的に再現し宇宙の進化の過程を探る研究も活発である。すでに、現在の宇宙の構造や観測事実の説明にはダークエネルギーやダークマターの存在が必要となっており、これらは標準理論のみでは説明できない現象である。また、現在の宇宙

^{*1} 新しい階層を説明する理論の代表例としては、素粒子と重力を含む全ての相互作用を超対称性を持つ弦として考える超弦理論などがある。

を構成する元素の様相はミクロの階層の理論である標準理論から説明できなければならない。これらの現象を説明するためにはミクロからマクロの階層間の相互作用を取り入れる必要があり、宇宙物理学・原子核物理学・素粒子物理学を連携させる分野間連携も活発となっている。天体现象や宇宙の構造の観測が進んでいる現在、宇宙や天体のような広大な対象物を原子・原子核・素粒子のようなミクロの基本法則から定量的に取り扱うための連成計算が進んできている。ここにも計算科学の手法が利用されている。

以上のような標準理論を超える理論の探索や、標準理論から現在の宇宙の有り様を説明する研究、及び後述の原子核分野の研究で重要な鍵を握るのが、標準理論の一部である、強い力を記述する量子色力学 (Quantum Chromodynamics : QCD) である。特に上記研究に必要な、標準理論に含まれるパラメータの精密決定、宇宙における元素合成や核力による多様な原子核の性質の第一原理的理解には、QCD の第一原理計算である格子 QCD による大規模数値計算が不可欠となっている。近年では、格子 QCD の計算手法を標準理論を超える物理の候補に適応した研究も進んでいる。

原子核分野

強い力の物理の構成要素であるクォークやグルーオンは単独で取り出すことができないため、強い力の物理は必然的に多体問題の物理となる。原子核物理は QCD を指導原理として強い力が織りなす多体系の非常に多彩な様相を明らかにすることを目指した研究分野であり、その対象は有限多体系のハドロンや原子核だけでなく、初期宇宙に存在したと考えられるクォーク・グルーオンプラズマや中性子星など広範にわたる。宇宙の構成要素のうち、ダークエネルギーとダークマターを除いたものの大半をバリオンが占めており、強い力の物理を知ることは宇宙を知ることであるとも言える。

地上の世界においては、強い力の物理は核子と総称される陽子・中性子多体系の物理と捉えられる。この核子多体系の物理においては、地上にある安定同位体を構成する原子核のみならず、安定同位体に比べ中性子あるいは陽子が極めて過剰な原子核や、ニホニウムなどで知られる超重元素の原子核、中性子星の内部構造など、核子多体系のあらゆる存在形態およびそのダイナミクスを統一的に理解することを目指している。陽子/中性子過剰核は理化学研究所の RI ビームファクトリー (Radioactive Isotope Beam Factory : RIBF) や米国でつい最近運転を開始した Facility for Rare Isotope Beams (FRIB) などの大型加速器によってその性質が調べられ、これらの魔法数（原子核が相対的に安定となる核子数）が安定核とは異なることが明らかになるなど、多くの新たな知見が生まれている。こうした現象を可能な限り核力に基づいて理解することが原子核理論の重要な目標である。そのインプットとなる核力に関しては、核子 3 体力に対する実験的知見が限られていることから、格子 QCD による核力の構築に大きな期待が寄せられている。中性子過剰核の性質は天体における元素合成過程の解明に対して、また、原子核の崩壊過程の一つである二重ベータ崩壊はニュートリノの性質の解明に対して重要なインプットとなる。このように核子多体系の物理は、それ自身としての興味だけでなく、学際的応用へもそのフロンティアが広がっている。

初期宇宙や星の中心部等の高温・高密度の極限状況下での物質の性質を知ることは宇宙の進化を解き明かす上で重要である。初期宇宙のような高温下では核子は存在せず、物質はクォーク・グルーオンプラズマとして存在する。宇宙の進化とともに温度が下がるとクォークが核子内に閉じこめられたハドロンとなる。温度と密度とともにクォーク相からハドロン相へどう転移するかを示す、QCD 相図を決定することは QCD の物理が目指す最大の目標の一つである。高温・低密度での QCD に関しては米国の Relativistic Heavy Ion Collider (RHIC) や LHC などによって多くの知見が得られてきている。一

方、有限密度での性質については未解明な点が多く、ドイツで建設中の Facility for Antiproton and Ion Research (FAIR)、ロシアで建設中の Nuclotron-based Ion Collider fAcility (NICA)、J-PARC で提案されている重イオン衝突実験などで明らかにされることが期待される。

アップ・ダウンクォークの質量は単独では 5 MeV (メガ電子ボルト) 程度しかないにも関わらず、それら 3 個が核子として閉じ込められると合計 940 MeV 程度へと大きく増大する。このハドロン質量獲得起源も重要な問題である。このようなハドロンはクォークとグルーオンが複雑に絡み合ったものであり、その内部構造は米国で建設が決まった Electron-Ion Collider (EIC) によって解明されることが期待されている。核子は最も軽いバリオンの基底状態であるが、構成子クォークの種類や励起エネルギーの範囲を広げることで多種多様なハドロンが存在する。また、これまで構成子クォーク 3 個からなるバリオンと 2 個からなるメソン (中間子) がハドロンとして知られていたが、近年、4 個あるいは 5 個からなるエキゾチックハドロンと呼ばれる粒子と考えられる質量ピークが LHC 実験などで発見されている。このように、ハドロンの存在形態の全てを理解することも大きな目標である。ストレンジクォークを含むバリオンであるハイペロンが構成要素として含まれる原子核はハイパー核と呼ばれ、その性質は J-PARC などで精力的に調べられている。

コミュニティからの意見

素粒子・原子核分野における計算科学的手法の利用は必要不可欠であるが、安定的ポスト減少も一因となり、計算科学的手法を用いる若手研究者が不足している。若手人材の確保・育成のためには計算科学的手法と基礎科学の両輪を併せ持つ分野横断的教育拠点や、さらに裾野を広げた人材の循環体制構築が望まれている。また、素粒子・原子核分野で独自に開発・改良されてきた計算コードを分野の財産として公開し共有する取り組みの必要性も言及された。

素粒子・原子核分野は、各フラッグシップ計算機の開発・運用期間よりも長期に渡る研究が必要な分野であるため、上記のような取り組みを行いつつフラッグシップ計算機群を最大限活用した研究を推進するために、長期的視野に立ったサポートを行う組織作りの重要性も指摘された。

コミュニティからの意見は以下のオンライン会合にて伺った。

- 2021 年 12 月 14 日（火）オンラインミーティング「計算科学ロードマップ更新意見交換会」
- 2022 年 1 月 17 日（月）計算基礎科学連携拠点、「富岳」成果創出加速プログラム「シミュレーションで探る基礎科学：素粒子の基本法則から元素の生成まで」・「宇宙の構造形成と進化から惑星表層環境変動までの統一的描像の構築」主催、「富岳で加速する素粒子・原子核・宇宙・惑星」シンポジウム
- 2022 年 3 月 19 日（土）日本物理学会第 77 回年次大会（オンライン）、インフォーマルミーティング「計算科学ロードマップ素粒子・原子核パート更新意見交換会」
- 2022 年 9 月 16 日（金）計算基礎科学連携拠点主催、第 16 回 HPC-Phys 勉強会（オンライン）

2.1.2 長期目標と社会貢献

長期目標

最終的な目的は、宇宙の進化過程を解明するために必要なクォークから原子核にわたる複数階層をつなぐダイナミクスの QCD による統一的理解を進めること、標準理論に含まれているクォークの性質に関するパラメータの精密決定により新しい階層の物理、ひいてはいわゆる究極理論の手がかりを探ることである。素粒子・原子核分野における計算科学的手法が有効な長期目標は以下の通りである。

(a) 強い力が創り出す階層構造の非摂動的理

(a-1) QCD に基づくハドロン間相互作用の決定と応用

自然科学は実験による新たな階層の発見と、既存の物理法則を統合的に説明する新しい物理法則の構築により発展してきた。強い力の理解においても、原子核の階層 ⇒ 核子（陽子・中性子）の階層 ⇒ 素粒子であるクォークの階層の理解というプロセスをたどってきた。クォーク間の相互作用（強い力）は QCD で記述されている。QCD を第一原理とする格子 QCD 計算は、核子階層及び原子核階層での既知の物理法則を矛盾なく説明するとともに、それを超える未知の物理現象に対する理論的予言を与えると期待できる。これまで実験で決めるしかなかった核力を含むハドロン間相互作用を QCD に基づいて決定し、クォーク階層から原子核階層を説明することが目標となる。これらの計算で得られた知見は、素粒子の階層と原子核の階層にまたがる物理現象や多様な構造を持つ原子核階層の性質を理解する上で大きな役割を果たすとともに、クォーク・グルーオンの自由度から原子核を直接構成するアプローチの基礎としても重要である。

(a-2) 有限温度・有限密度 QCD の解明

クォークの階層についての理解を深めることも宇宙の進化解明にとって重要である。現在の宇宙の姿はビッグバンによる宇宙開闢後のミクロなスケールにおける物理現象の帰結であると考えられている。QCD は宇宙初期のような高温ではクォーク・グルーオンプラズマ相を、現在のような“冷えた”状態ではハドロン相を予言する。両者の間の相転移の次数はダークマターの候補に制限を与える。また、低温高密度における QCD の相構造解析は、超新星の爆発メカニズムの解明や、中性子星などの高密度天体の内部状態を知るうえで必須である。このようなクォークやグルーオンの自由度が重要となる超高温・低密度状態および超高温・超高密度状態における物質の相構造を QCD に基づいて解明することが目標となる。クォーク・グルーオンプラズマは、近年の重イオン衝突実験（米国の RHIC 実験、CERN の ALICE 実験）により研究は進められているが、温度と密度の関数として多様な相構造を持つことが予想される QCD は、その全貌を実験により解明することは難しく、格子 QCD 計算に大きな期待がかけられている。

(a-3) 第一原理計算による原子核構造・反応、状態方程式の記述

原子核では核子間の結合エネルギーが核子あたり数 MeV であることが知られている。これは核子自身の励起エネルギー（第一励起状態であるデルタ粒子は約 300 MeV）よりもはるかに小さいことから、核子多体系は核子を構成要素として記述することが非常によい近似となる。核子が励起する自由度は核子 3 体力（あるいはより多体力）として陰に取り込まれる。そのことから、核子多体系（一般にはバリオン多体系）では、核子 2 体力や核子 3 体力を出発点とし、現象論的補正なしに多体問題を解くことを第一原理

計算と位置づけており、この第一原理計算によって核子系の存在形態とダイナミクスを記述することがその重要な目標の一つとなっている。核力をそのまま用いてシュレーディンガー方程式を解く少数多体系計算では、手法および精度によるが、現在のところ核子数 5-10 程度の原子核の計算が可能であり、今後その限界を漸次拡大していく。より質量数の大きな原子核、さらには核子多体系の無限系である核物質の状態方程式（核子密度とエネルギー密度との間の関係式）を記述するためには、理論的・計算科学的観点からの発展が必要であり、より正確な解を得るために様々な手法が検討されている。こうした第一原理計算によって、格子 QCD の結果と核構造との間に整合性のある核子 3 体力（あるいは 4 体力以上の必要性）を決定する。さらに、 r 過程元素合成に関わる極めて中性子過剰な原子核の構造と反応を記述したり、逆に、ハイパー核の構造計算と実験データをもとにハイペロン間相互作用を引き出すなど、実験のアクセスが当面困難である現象に対し、理論的に答えを出すことを目指す。また、核物質の状態方程式を決定することによって、中性子星の構造を明らかにする。

(a-4) 適用範囲の広い有効模型の構築

強い相互作用に関わる物理全てを、クォーク多体系と捉えるべき系においては QCD から、また、核子多体系と捉えるべき系においては核力から第一原理的に記述することが分野の最終的な目標であるが、当面の間それは不可能であり、原子核物理における重要な課題の多くは有効模型によって記述される。有効模型あるいは有効理論とは、知りたい現象を本質的に重要な自由度によって記述する理論である。核子多体系については、数 MeV 程度の低エネルギー現象に対し、核力の長距離相関の自由度を取り入れた有効模型がこれまで大きな成功を収めてきた。有効模型計算の長期的目標は、模型が記述できる核種や現象を飛躍的に拡大することである。この中には未知核種や実験が困難な現象も含まれる。また、これまで計算能力の限界から、対象とする核種や現象に応じて出発点の異なった有効模型が使われてきたが、今後はこれらの有効模型の統一化を進めていく。さらに、有効模型計算の精度を上げるために、模型の有効相互作用を現象論的・経験的なものから核力に依拠した微視的なものへと置き換えていく。こうした統一的かつ定量的な原子核の記述は、天体核物理や原子力工学等への応用においても非常に重要な役割を果たす。

(b) 標準理論を超えた物理の探索

(b-1) 標準理論の高精度検証による新しい物理階層の探索

標準理論が記述する階層よりも更に一段深い階層の新しい物理を探るための手段の 1 つに、既知の物理現象に対する高精度の実験結果と標準理論による理論的予言（理論値）とのズレを見出すアプローチがある。このアプローチで重要なのが、ズレを明確に捕らえるために、理論・実験の双方が目標精度を達成することである。このアプローチの実験は、KEK の Belle II 実験と CERN の LHCb 実験によるボトムクォークに関する物理現象の精密測定や、米国のフェルミラボで実行及び日本の J-PARC で計画されているミュー粒子異常磁気能率の超高精度測定、将来建設が計画されている国際リニアコライダー (International Linear Collider : ILC) 実験が挙げられる。ボトムクォーク実験に対しては格子 QCD を用いたハドロン行列要素の精密計算、ミュー粒子異常磁気能率実験に対しては摂動論に基づく量子電磁力学 (QED) 高次補正計算と格子 QCD を用いたハドロン真空偏極計算、将来の ILC 計画のデータ解析にはファインマン図自動計算が不可欠であり、これらの全ての計算に対して実験精度の向上に合わせた精确な理論的値を与えることが目標となる。これらの高精度計算は大型計算機による計算が不可欠である。

(b-2) 超弦理論の量子ダイナミクスの解明

超弦理論は、宇宙に存在する全ての粒子と重力を含む4つの力を統一的に記述する究極理論の候補とされ、世界中で精力的な研究が展開されてきた。しかし、解析的計算だけでは理論の全体像を完全に把握することが難しく、様々な問題が未解明なまま取り残されている。宇宙創成の機構の解明、素粒子の運動を記述する標準理論の根源的な理解、時空が4次元である理由、さらに超弦理論から予測されるゲージ重力双対性を用いたブラックホールの情報喪失問題の解明や、QCD、超伝導、超流動のホログラフィックな記述の確立などである。このような多くの未解決問題を行列模型や超対称ゲージ理論の数値シミュレーションから解明することは、超弦理論の応用の広がりとともに、21世紀の物理学の中心的な課題の一つとなっている。

(b-3) 標準理論を超えた理論の探索とその性質解明

自然界には何故4つの力が存在するのか、ダークマターとは何か、素粒子の種類ごとに質量が大きく異なる理由は何か、このような多くの基本的な問題が依然として謎につつまれている。これらの謎は、標準理論を越えた階層に現れる新理論により解明されると期待されており、欧州のLHCにおいて、新理論の兆候を捕らえるための実験がすでに始まっている。新理論の候補となる理論は、超弦理論や大統一理論の基礎であり現在最も有力視されている超対称性理論など数多く提案されている。また、これらの理論を記述する枠組みである場の量子論においても θ 項を持つトポロジカルな理論など新しい知見が広がっており、その中から新理論の新しい候補が見つかる可能性がある。実験と相補的に、理論的な観点から新理論の候補を絞るため、これらの理論の力学的な性質を数値計算により明らかにすることが今後の課題となっている。この研究が進展すれば、将来世界的に推進される実験の方針決定に対して大きな貢献が期待できる。

社会貢献

素粒子・原子核物理学の最終目標は「自然界はどこからできているのか」「宇宙はどのように始まりどのように進化するのか」という、誰もが一度は考えたことのある基本的な問いに答えようとするものであり、その目標へ向けた研究で得られた知見は多くの人の知的好奇心を満たすものである。基礎科学である素粒子・原子核物理学は、社会への応用を目指していないがゆえに、その研究活動で得られる技術的・科学的知見は基礎的であり、そのものが新しい価値である。その知見を積み重ね、世代（個人の寿命）を超えて長い年月を経ても失われないよう、正しく後世へ伝えていくことが基礎科学の社会への貢献である。そのように蓄積された知見は、発見から数十年から百年経った後に思いがけないかたちで日常生活を豊かにし、社会の発展とともに生まれる問題の解決に貢献してきた歴史がある。

例えば、カーナビやスマートフォンに搭載されているGPSの理論的基礎となっている相対性理論^{*2}は、生まれてから百年を経て日常生活に欠かすことのできないものになっている。研究データの共有や研究コミュニティ形成のためにCERNで開発されたWorld Wide Web (WWW) もまた今日の日常生活では不可欠なものである。計算科学からの貢献の例としては、格子QCD分野で開発されたハイブリッドモンテカルロ法（ハミルトニアンモンテカルロ法）が分野を超えてデータサイエンスに取り入れられていることがあげられる。また、素粒子・原子核物理でも応用の始まった機械学習や量子計算を用いた研究で得

*2 相対性理論を用いた補正を考慮しないと、GPSの位置情報の精度は格段に悪くなる（日本物理学会誌 2005年9月号 741頁
「カーナビと相対論」中村卓史 <http://ci.nii.ac.jp/naid/110002079440>）

られる知見は、計算機だけでなくそれを活用する社会生活にも変革をもたらす可能性がある。

一方、原子核物理は、ハーン・シュトラスマン・マイトナー・フリッシュによる核分裂の発見以来、エネルギー利用とともに発展してきた側面がある。また、原子核の崩壊に伴う放射線は、医療や科学研究の道具として広く利用されている。近年、中学校・高校の教科書に掲載されたニホニウムは、新元素合成による学術的貢献というだけではなく、新しい一般教養の普及という社会貢献の側面もある。現在、狭義の原子力利用に関しては放射線廃棄物の処理法の確立、放射線利用については医療用同位体の安定供給体制の構築が喫緊の課題であり、これらの社会的要請に原子核物理の新知見が活かされる余地は多く残されている。

2.1.3 課題とその解決に必要な計算手法・アプリケーション

前述した長期目標を達成するためには様々な課題を計算科学的手法によって解決することが必要不可欠となってくる。本節では特に 2032 年ごろに解決すべき、もしくは解決できる課題について、それらの課題の概要とそのために必要な計算手法とアプリケーションについて計算規模が大規模なものと計算手法が特徴的なものを概説する。これらに加えて、近年発展の著しいテンソルネットワークを用いる手法や、機械学習や量子計算を取り入れた計算手法も議論されている。素粒子・原子核分野における、より詳細な課題の分類と課題内容、計算科学的解決方法については第 4.1 節にて説明する。

計算手法やアプリケーションは物理の基礎法則や考え方に基づいて記述されるため課題間で共通のものもある。課題概説には計算手法・アプリケーションの名称を列挙するにとどめ、名称に添字^(a)を付した。添字に対応する各計算手法・アプリケーションの説明は課題概説から分離し課題概説の後にまとめた。

課題概説

(i) 格子 QCD を用いたハドロン相互作用の決定 [長期目標 (a-1) の課題]

QCD に基づく第一原理計算である格子 QCD 計算を用いてハドロン間相互作用を決定することで、原子核やハイパー核（構成要素にハイペロンを含む原子核）、さらには中性子星までの統一的記述を目指す。今後 5-10 年で、2 体バリオン間のポテンシャル（S/P/D/F 波）や 3 体バリオン間のポテンシャルを決定し、またエキゾチックハドロンの解明に必要なメソン間、メソン・バリオン間の相互作用を決定する。得られたバリオン間相互作用はハイパー核や不安定核の性質を明らかにする上で不可欠であると共に、それを基にして高密度核物質の状態方程式の決定を行うことで、中性子星の構造や連星中性子星合体、超新星爆発、重元素合成などの解明を推し進める。また散乱理論と様々なハドロン間相互作用を組み合わせることで、近年の実験で報告されている多様なエキゾチックハドロンの構造解明を進める。課題詳細は課題詳細は 4.1.1 節にある。

必要な計算手法・アプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ウィルソン型フェルミオン作用^(b)、大規模疎行列反復解法^(b)、格子 QCD ハドロン多体波動関数計算^(c)

(ii) 格子 QCD を用いた QCD 相構造の理解 [長期目標 (a-2) の課題]

温度と密度を関数とするクォーク物質の相構造解析は、実験によって包括的に研究することは困難であるため、格子 QCD を用いた第一原理計算への期待は大きい。当面の課題は、相図におけるクォーク・グルーオンプラズマ相とハドロン相を隔てる相転移線が消える臨界終点の温度と密度の精密決定である。密度がゼロの時には、相転移の次数のクォーク質量依存性（コロンビアプロット）を明らかにするのが目標である。それと並行した課題に、低温中高密度領域での QCD 相構造解析と状態方程式の精密決定、及び重イオン衝突実験で注目されている保存電荷のゆらぎのガウス分布からのずれを表す高次キュムラントやクォーク・グルーオンプラズマの輸送係数の決定があげられる。課題詳細は 4.1.5 節、4.1.6 節にある。

必要な計算手法・アプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ウィルソン型フェルミオン作用^(b)、ドメインウォール型フェルミオン作用^(b)、大規模疎行列反復解法^(b)

(iii) 多核子力を含む核力の深い理解 [長期目標 (a-3) の課題]

2 核子間相互作用は核子散乱実験データによってかなりよくわかっているものの、核子が複合粒子であることに起因する多核子力（核子 3 体力など）の不定性は未だに大きい。核子多体系である原子核の構造・反応の精密な記述によって、その不定性を極力減らすことを目指す。そのためには、多核子力の影響が大きくなると考えられる、より核子数の大きな原子核へと核力に基づく第一原理計算のフロンティアを拡大することが必要である。少数多体系計算によって厳密解を担保するとともに、中重核に適用可能な方法論の開発も進めていく。同様のアプローチは、ハイパー核の構造・反応からハイペロンの相互作用を理解するのにも有用である。「少数多体系計算」、「QCD に立脚した原子核反応計算」、「配位間相互作用計算（殻模型計算）」の課題を含む。これらの課題詳細は 4.1.8 節、4.1.9 節、4.1.11 節にある。

必要な計算手法・アプリケーション

ガウス関数展開法^(e)、確率論的変分法^(e)、無芯殻模型計算^(e)、離散化連続チャネル結合法^(e)、グリーン関数モンテカルロ法、結合クラスター法

(iv) 普遍的有効核力の構築 [長期目標 (a-4) の課題]

有効模型では、模型であらわに取り扱われない自由度である核力の短距離成分を有効核力として繰り込む必要がある。有効核力は生の核力よりもさらに不定性が大きいものの、原子核の存在限界や未知核種の性質を決定づける極めて重要な要素である。微視的および現象論的観点から普遍的な有効核力の確立を目指す。特に核子 3 体力の最新の知見を取り入れた微視的理論を発展させるとともに、その物理的知見を現象論的有効核力に活かし、核構造のみならず核反応、状態方程式、核分裂のエネルギー面をも記述可能なり包括的な有効核力を構築する。そのためには、核構造、核反応、核分裂のいづれの現象においても、模型空間内の自由度を十分に取り込んだ高精度の多体系計算を広い質量領域で行うことが非常に重要である。「核力に基づく核物質の状態方程式」、「配位間相互作用計算（殻模型計算）」、「原子核密度汎関数計算」、「揺動散逸理論による重イオン反応・核分裂・新元素合成」、「微視的クラスター模型」の課題がある。これらの課題詳細は 4.1.10 節、4.1.11 節、4.1.12 節、4.1.13 節、4.1.14 節にある。

必要な計算手法・アプリケーション

密度汎関数法（乱雑位相近似や生成座標法なども含む）^(e)、殻模型計算^(e)、反対称化分子動力学計算^(e)、離散化連続チャネル結合法^(e)、ランジュバン方程式

(v) 新物理探索・解明のための精密格子 QCD シミュレーション [長期目標 (b-1) の課題]

主目的は、 B 中間子レプトニック崩壊の崩壊定数 (f_B, f_{B_s})、セミレプトニック崩壊 ($B \rightarrow \pi l\nu, B \rightarrow D^* l\nu, \Lambda_b \rightarrow p(\Lambda_c)l\nu$) の形状因子、 B 中間子混合行列要素を高精度で決定することである。同時に、標準理論に現れる行列要素だけでなく、新物理模型に現れる一般的な行列要素も包括的に計算する。これらの計算は Belle II 実験の測定精度に見合うように、理論誤差を 1–5% 以下に抑えることが目標となる。理論誤差を抑えるためには、①約 5GeV (ギガ電子ボルト) のボトムクォーク質量を直接扱える高詳細格子を用いた計算、②カイラル外挿の不定性を排除した物理点の計算、③新物理解明に重要な軽いハドロンの行列要素計算を行う。これらの計算では、計算の不定性を可能な限り抑えるため、カイラル対称性を保つフェルミオン作用を用いる。課題詳細は 4.1.15 節にある。

必要な計算手法・アプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ドメインウォール型フェルミオン作用^(b)、大規模疎行列反復解法^(b)

(vi) QED の高次補正計算 [長期目標 (b-1) の課題]

ミュー粒子の異常磁気能率は 0.00001% 程度の精度になると未知の重い粒子の影響が見え、新しい階層の手がかりが得られると期待されている。2023 年現在、まだこの精度には達していないものの実験値と理論値との差が問題になっている。電子の異常磁気能率は、現在より一桁精度が上がると新しい階層の新物理に様々な制約を与えることができる。新しい階層の手がかりを探すためには、実験および理論の双方が上記の精度を達成し、両者の差異を確定させることが必要である。そのために、QED 摂動計算の 10 次の展開係数の不定性を現在よりも 1/10 に削減することを目標とする。その後、将来実験の精度で有意な寄与として現れると推測される 12 次の展開係数をターゲットにし、電子については全ての項、ミュー粒子については主要項をピックアップして係数の目標精度を 10% 程度にし、数値計算を行う。課題詳細は 4.1.16 節にある。

必要な計算手法・アプリケーション

QED 高次摂動モンテカルロ積分計算^(d)、4 倍精度実数計算、高度に最適化された 4 倍精度ライブラリ

計算手法・アプリケーションの概説

上述の各課題で必要となる計算手法・アプリケーション群の概説を以下にまとめた。

(a) 格子 QCD ハイブリッドモンテカルロ法

格子 QCD は格子離散化されたユークリッド時空上で場の量子論に基づいて構築されており、物理量は膨大な自由度の重みつき積分で表現されている。この積分をそのままでは実行することはできないため、積分の重みを再現するようにアンサンブル（ゲージ配位）を生成し、統計平均から物理量の値を得る、モンテカルロ法が用いられている。現在の主流のゲージ配位生成法は、ハイブリッドモンテカルロ (Hybrid Monte Carlo: HMC) 法である。この HMC 法の中で計算コストの 8 割以上を占めるのが、次項目「(b) 格子 QCD フェルミオン作用・大規模疎行列反復解法」によるフェルミオン伝搬関数の計算である。ま

た、HMC の一部に機械学習の手法を取り入れたアルゴリズムも提案されている。

(b) 格子 QCD フェルミオン作用・大規模疎行列反復解法

フェルミオン伝搬関数の計算は、大規模疎行列連立方程式を解く問題に帰着できる。係数行列はクオークの運動方程式を格子離散化したものであり、4 次元構造格子上のステンシル計算になる。本節の課題では次の 2 つのものが使われる：① ウィルソン型作用、② ドメインウォール型作用。後者は内部自由度の構造から、5 次元構造格子上のステンシル計算とみなすこともある。これらの作用を用いたフェルミオン伝搬関数の計算（大規模疎行列連立方程式を解くこと）には、クリロフ（Krylov）部分空間反復法が用いられる。領域分割やマルチグリッド法など用いた前処理が組み合わされることもある。

(c) 格子 QCD ハドロン多体波動関数計算

ハドロン間相互作用ポテンシャルは、散乱位相差を記述する積分核として、ハドロン多体波動関数から定義・導出される。ハドロン多体波動関数の計算には、4 次元時空の異なる各点での多体ハドロン相関が必要であり、また相関関数の計算コストは相関関数に含まれるクオークの個数が増えるに従い急速に増大する。さらに計算には高速フーリエ変換（FFT）も必要となる。これらにより計算には従来の格子 QCD 計算よりも遙かに多くの演算と通信が必要になる。

(d) QED 高次摂動モンテカルロ積分計算

異常磁気能率に対する QED の高次摂動計算は、木下東一郎（コーネル大学）を中心とした日本人研究グループ独自の数値的アプローチによってのみ計算されてきた。摂動展開の係数は、ファインマン図と呼ばれる図形の和として表され、その図に現れるループが運動量空間における多次元積分に対応する。この積分を解析的に多次元パラメータ空間に変換した後、モンテカルロ法を用いて数値的に評価する。数値計算の特徴は、①被積分関数プログラムが 1 億行程度の長さになる、②発散の処理手続きの大きな桁落ちのため 4 倍精度の実数計算が必要、③モンテカルロ積分では通信をほとんど必要としない処理が可能な点が挙げられる。

(e) 核子系の多体計算

核子多体問題に対しては (iii) および (iv) 項に挙げたような様々な手法およびアプリケーションが用いられているが（代表的なものについては第 4.1 節に記述する）、数値計算の観点からは、その多くに共通する枠組が採用されている。それは、求めるべき多体状態を物理的に重要となる基底波動関数で展開するという方法である。そこでは主に、①基底波動関数の生成、②多体ハミルトニアン行列要素の計算、③多体ハミルトニアン行列の対角化、という 3 つのステップで構成される。ここでは基底波動関数の数が多体ハミルトニアン行列の次元となる。核子数が増えるにつれ、多体ハミルトニアン行列要素の計算時間がかかるとともに、必要な行列要素の数（すなわち行列の次元数）も増える傾向にある。

2.1.4 ロードマップ

前述した各課題および第 4.1 節で述べられている課題のロードマップを図 2.1 に記載した。各課題の詳細は第 4.1 節に記載されている。

課題＼年代	2022～	2024～	2026～	2028～	2030～	2032～
格子QCDを用いたパリオン相互作用の決定 長期目標(a-1)	S/D波/パリオン間相互作用の物理点計算	><			連続極限での決定	→
	P/F波/パリオン間相互作用の計算	><			物理点での決定	→
	3体/パリオン間相互作用の計算	><			クォーク質量を軽くしながら決定	→
	パリオン間相互作用計算に基づく原子核物理・天体物理への応用	<			→	
ハドロン階層の理解へ向けた格子QCD計算の進化 長期目標(a-1)	メンソン間力・メンソーパリオン間力の計算とエキゾチックハドロンへの応用	<			→	
	連続極限でのハドロン質量精密計算	><			→	
	ハドロン2体散乱、崩壊過程精密計算	><			QED効果を入れたハドロン質量精密計算 軽原子核束縛エネルギー直接計算	→
格子QCDを用いた核子構造の理解 長期目標(a-1)	アイソスピン対称性を課した連続極限での核子構造精密計算	<>	僅かなアイソスピン対称性の破れを含めた核子構造精密計算		→	
格子QCDを用いたQCD熱力学 長期目標(a-2)	状態方程式の解析	<	状態方程式の精密解析	><	キュムラントの解析	→
	輸送係数の解析	<	相関関数の精密解析	><	輸送係数の解析	→
	N=2+1 クォーク質量が現実の3倍 程度の相移動の確立	<>	物理点より軽い領域の探索			
格子QCDを用いたQCD相構造の理解 N=3相探索 大体積有限温度計算の加速 長期目標(a-2)	N=3相探索 細かい格子($a=0.08 \text{ fm}$)	><	細かい格子($a=0.13 \text{ fm}$)	><		
	計算手法開発	<		><	N=2+1, 3での開発した手法による発展研究	→
	RHIC・LHCの定量的な解析	<			低エネルギー領域の解析・ペイズ統計解析	→
高エネルギー重イオン衝突実験(クォーク・グルーオンプラスマの解明) 長期目標(a-2)	少数多体系計算 質量数5.6の軽い中性子過剰原子核の共鳴研究	<		><	質量数7以上のエキゾチックな原子核への応用	→
	QCDに立脚した原子核反応計算 大規模チャネル結合反応計算の実装	<		><	核構造モデルの結果を用いた原子核反応計算	→
	核物質状態方程式計算 高次の粒子間相関を考慮した多体変分法の構築	<		><	状態方程式の系統的解析と高密度天体への応用	→
	配位間相互作用計算 (殻模型計算) p殻核の第一原理核構造計算	<		><	sd殻核の第一原理核構造計算	→
普遍的有効核力の構築 長期目標(a-4)	配位間相互作用計算 (殻模型計算) 水銀・鉛領域の準粒子真空殻模型計算	<		><	ウラン領域、超重元素領域の準粒子真空殻模型計算	→
	原子核密度汎関数計算 全核種の線型応答計算による計算核データの構築	<		><	変形共存、核分裂の微視的記述	→
	揺動散逸理論による重イオン反応・核分裂・新元素合成 微視的クラスター模型 微視的輸送係数と多次元化の導入	<		><	全融合分裂過程を動力学モデルのみで計算するための改良	→
	現象論的模型による、 α 過程炭素燃焼過程の研究	<		><	アイソスピン対称性の破れを考慮した放射捕獲反応の研究 ビッグバン元素合成の第一原理計算	→
新物理探索・解明のための精密格子QCDシミュレーション 長期目標(b-1)	QCDの場の配位の生成	<>			B中間子のハドロン行列要素の計算	→
	10次の展開係数の精密計算	><				
量子電磁気学(QED)の高次補正計算 長期目標(b-1)	10次の展開係数の精密計算	<		><	12次の展開係数の計算	→
	ファインマン振幅自動計算 トップクォーク対生成電弱相互作用とQCD混在ループ計算	<>	2ループ図形の自動計算システム構築 トップクォーク対生成全電弱相互作用2ループ計算			→
超弦理論シミュレーションで探る時空の量子ダイナミクス 長期目標(b-2)	複素ランジュハン法に基づく数値計算による時空の生成の検証	><	レフシェツツ・シンプル法に基づく数値計算による時空の生成の検証	<	インフレーション仮説の検証、e-foldingの決定、密度揺らぎの測定 超弦理論から標準理論およびビッグバン宇宙の導出	→
	超弦理論から標準理論およびビッグバン宇宙の導出	<>		><	超弦理論の確立と新しい物理への応用	→
場の量子論の非摂動ダイナミクス トップダウン 長期目標(b-3)	ゲージ重力双対性的検証	<		>	超弦理論とM理論との双対性的検証	→
	ゲージ重力双対性的応用的研究	<		>		→
	有効理論の大まかな探索	<		>	超対称性の破れた真空の安定性解析	→
	新規アイデア 分野横断的新課題の検討 解析コード開発	<		>	スペクトル精密測定 新課題の数値計算	→
テンソル繰り込み群による素粒子物理学および物性物理学の研究 長期目標(a-2), (b-3)	4次元U(1)ゲージ理論 有限密度、θ項	<>	4次元SU(2)ゲージ理論 有限密度、θ項	><	4次元SU(3)ゲージ理論 有限密度、θ項	→
	(2+1)次元ハードモデル 反強磁性相・超伝導相の研究	<>		><	(3+1)次元ハードモデル 金属-絶縁体転移・反強磁性・超伝導相の研究	→

図2.1 素粒子・原子核分野ロードマップ

2.1.5 必要な計算機資源

2032年頃における課題解決のために必要な計算機資源の見積りを行い、前小節に挙げた代表的アプリケーションのうち以下のものについて、2.13節の表2.1に記載した。また、計算機資源の見積りの詳細は第4.1節に記載している。

- 課題(i) 格子QCDを用いたハドロン相互作用の決定 [長期目標(a-1)の課題]
 - アプリケーション(a) 格子QCDハイブリッドモンテカルロ法
 - アプリケーション(b) 格子QCD フェルミオン作用・大規模疎行列反復解法： ウィルソンフェルミオン作用による配位生成と測定
 - アプリケーション(c) 格子QCD ハドロン多体波動関数計算
- 課題(v) 新物理探索・解明のための精密格子QCDシミュレーション [長期目標(b-1)の課題]
 - アプリケーション(a) 格子QCDハイブリッドモンテカルロ法
 - アプリケーション(b) 格子QCD フェルミオン作用・大規模疎行列反復解法： ドメインウォールフェルミオン作用による配位生成と測定
- 課題(vi) QEDの高次補正計算 [長期目標(b-1)の課題]
 - アプリケーション(d) QED高次摂動モンテカルロ積分計算
- 課題(iv) 普遍的有効核力の構築 [長期目標(a-4)の課題]
 - アプリケーション(e) 核子系の多体計算: モンテカルロ殻模型計算

参考文献

- [1] 日本学術会議第三部拡大役員会・理学・工学系学協会連絡協議会. 理学・工学分野における科学・夢ロードマップ 2014. <http://www.scj.go.jp/ja/info/kohyo/kohyo-22-h201.html>, 2014.
- [2] 高エネルギー物理学研究者会議委員石野雅也他. 高エネルギー物理学将来計画検討小委員会答申. <http://www.jahep.org/>, 2017.
- [3] 中村卓史. カーナビと相対論. 日本物理学会誌, 2005年9月号, pp. 741–742, 2005. <http://ci.nii.ac.jp/naid/110002079440>.

2.2 ナノサイエンス・デバイス

2.2.1 分野の概要

半導体材料や超伝導物質など、20世紀の科学技術研究の中で生まれた物質群は、100種類ほどの元素の無限とも言える組み合わせの中から見出され、特異な機能や新しい現象の発現を通して、現代社会の産業基盤を形成してきた。これらの物質や材料をミクロな視点に立って研究する物質科学は、物性物理、統計物理を中心に、分子科学、材料科学などにもまたがり、基礎研究と応用研究をつなぐ役割をも担う、広大な学問分野である。

物質科学分野における大規模数値シミュレーションは、古くは Fermi-Pasta-Ulam の非線形励起・再帰現象、剛体球の Alder 転移などの概念革新への寄与に始まり、現代量子多体系では、分数量子ホール効果の数値検証、相転移と臨界現象の解明、高温超伝導の機構提案など、物質科学の基礎研究に欠かせぬものとなった。この分野では、電子状態を量子力学に基づいて第一原理的に計算する手法として、波動関数理論に基づいた量子モンテカルロ計算、密度汎関数理論に基づいたバンド計算がある。また、大規模な原子・分子系の集団運動を古典力学・統計力学に基づいて計算する古典分子動力学法、古典モンテカルロ法などのシミュレーション手法がある。

物質科学分野においては、現在「高性能半導体デバイス」、「光・電子融合デバイス」、「超伝導・新機能デバイス材料」、「高性能永久磁石・磁性材料」、「高信頼性構造材料」といった科学的・社会的に重要な課題群や、複合・マルチスケール問題を通じて極限の探求を目指す萌芽的課題などが、ポスト「京」コンピュータ時代に向けて重点的に進められており、国内外で超並列環境に最適化された大規模ソフトウェアの開発が進められている。一方で、基礎理論自体が多層構造をなしており、それぞれのレベルにおける方法論が非常に多岐にわたり、かつそれぞれが相補的に発展してきたことが、物質科学のもう一つの大きな特徴である。シミュレーションの大規模化・精密化を進めると同時に、全く新しい基礎理論・モデルの提唱、シミュレーション手法の開発・実装・高度化、最先端の計算機を使ったシミュレーションによる予言・検証と理論へのフィードバックというサイクルを効率よく進めていくことが、今後の計算物質科学分野の発展のために重要である。

(i) 半導体電子デバイス

半導体テクノロジーはポストスケーリング時代を迎え、ナノドット、ナノワイヤーなどの構造体が次世代デバイスに不可欠な要素となっている。それらナノ構造体の構造的安定性と電子機能についての高精度の予測を目指し、主に密度汎関数法に基づくシミュレーション技術が確立されてきた。海外では、ABINIT(白)・CASTEP(英)・CONQUEST(英日)・CP2K(欧)・CPMD(独)・QUANTUM ESPRESSO(伊)・SIESTA(西)・VASP(奥)・WIEN2K(奥)、また国内では、AkaiKKR・ASCOT・CMD・FEMTECK・OpenMX・PHASE・QMAS・RS-CPMD・RSDFT・RSPACE・TOMBO・UVSOR・xTAPPなど、それぞれ特色のあるコードが開発されている。国内ソフトウェアでは、超並列計算の実績のあるものも多く、代表的なものとしては、「京」コンピュータで 3PFLOPS(44% の実効効率) を達成し 2011 ACM Gordon-Bell Prize を受賞したプログラム RSDFT、初代地球シミュレータ 512 ノード(4096CPU)で

50% の実効効率を出したプログラム PHASE (12,288 原子系を「京」12,288 ノードで 23% の実効効率)などがある。これらの高度に最適化・高速化されたソフトウェアにより、1,000~200,000 原子規模のナノワイヤーの第一原理電子状態・伝導特性の計算、1,000,000 原子規模のゲルマニウム (Ge) ナノ構造の解明などが進んでいる。

(ii) 光・電子融合デバイス

一方、実在系ナノ構造体を対象とした光・電子融合デバイスの第一原理計算に基づく理論設計の試みは、現状では国内外ともにほぼ皆無と言ってよいが、わが国ではデバイス設計に不可避である光と物質の露わな相互作用を取り込んだナノ光学理論に基づく電子・電磁場ダイナミクス法プログラム SALMON が、「富岳」コンピュータ 27,648 ノードを用いた 13,632 原子からなる系に対する大規模計算に成功しており、十数ナノメートル程度の実在系ナノ構造体の電子・電磁場ダイナミクスが十分に計算可能な状態となっている。

(iii) 強相関量子物質

トポロジカル物質や電子間相互作用の強い強相関物質といった、古典的な粒子描像では表現できない“量子もつれ”により特異な性質が発現している強相関量子物質は、省エネルギー・デバイスや量子・デバイスといった新奇デバイスへの応用が期待されている材料として注目を浴びている。量子物質やそれらを単純化したモデルといった量子多体系を取り扱うために、厳密対角化法、量子モンテカルロ法、変分モンテカルロ法、密度行列くりこみ群法やテンソルネットワーク法、などの様々な数値計算手法がこれまでに提案してきた。近年では、こうした数値計算手法に機械学習分野の技術を融合させることで、利便性や汎用性、精度を高めた手法の開発が活発になっている。このような手法が実装されている国内ソフトウェアとして、厳密対角化法 HΦ、変分モンテカルロ法 mVMC、動的密度行列くりこみ群法 DDMRG、テンソルネットワーク法 TeNeS などがあげられ、それらを用いた大規模並列計算による研究成果が国内外で報告されている。手法やソフトウェアの発展に伴い、従来行われてきた量子物質で生じる特異な量子多体系現象の機構解明や新しい量子機能を持つ物質の探索を効率よく行えるようになっただけでなく、接合系・準結晶などの非周期性や光・磁場などの外場印加による量子機能の開拓・制御を目指した研究も進められている。

(iv) 永久磁石・スピントロニクス材料

現在使用されている永久磁石材料は、希土類元素と遷移金属元素から成る希土類磁石であり、国内産業基盤の強化に際し非常に重要な材料である。性能、レアメタルの資源リスク、価格などの観点から優れた永久磁石材料を開発するためには広大な材料空間探索が必須であり、第一原理計算を用いた電子論的解析だけでなく、デザイン主導の材料探索が積極的に行われてきた。永久磁石材料の第一原理的に取り扱う際には、元素配置の不規則性や f 電子の記述など、通常の材料では問題にならない困難な壁に直面する。現状では、コヒーレントポテンシャル近似やオープン・コア近似が実装されている AkaiKKR で、これらの問題を取り扱い網羅的探索が可能になっているが、材料探索の精度を向上させるには計算アルゴリズムやソフトウェアのさらなる発展が必要である。また、保磁力の改良には微細組織の最適化が重要となる。これらの課題には OpenMX や KKRnano などの O(N) 法を用いた主相・副相界面における微視的特性の

解析、そしてモンテカルロ法を組み合わせた多階層連結シミュレーションによる磁化反転機構の解明が重要な役割を担っている。

永久磁石材料に限らず磁性材料では電子状態、磁気特性、伝導特性などの多目的物理量の最適化が必要なため、インフォマティクス手法の活用に対して期待が大きい。特に、「富岳」コンピュータを利用した大規模磁性データベースの構築、そしてペイズ最適化・データ同化・頻出パターンマイニングなどを組み合わせたデータ駆動型マテリアルデザインが発展している。

(v) 構造材料

実用材料の飛躍的性能化にむけて、マルチスケールシミュレーションで材料組織を設計し、評価する試みが進んでいる。一例として、高精度の自由エネルギー計算による相図計算、各相の自由エネルギーや界面・粒界・欠陥の第一原理計算とフェーズフィールド法・分子動力学計算・モンテカルロ法などのメゾスケールでの粗視化シミュレーションとの連結が挙げられる。これらの連成計算では、各スケールの規模はそれほど大きくないものの、それらを効率的に組み合わせて実行する必要がある。下部計算となる第一原理計算では、ABINIT・QMAS・TOMBO・OpenMXなど国内外で開発されたさまざまなソフトウェアが対象に応じて使い分けられる。第一原理計算の結果を用いた上部計算(分子動力学計算)を行うFERAMなどのソフトウェアも開発され、広く使われている。また、機械学習や進化的アルゴリズムを用いた大規模複雑構造の探索の試みも精力的に進められている。

(vi) コミュニティーからの意見

物質科学分野では、これまで、討論会「エクサスケールコンピュータへの期待」(2012年7月13日、東大物性研)、計算物質科学シンポジウム(2012年10月22日、東大物性研)、CMSI研究会(2012年12月3日、分子研)、日本物理学会シンポジウム「エクサスケールに向けて歩み出す計算物理学」において、実験家、企業研究者も含めたコミュニティ全体に対しロードマップを紹介し、パネルディスカッションなどを通じて意見収集を継続的に行ってきている。具体的には、J-PARC、SPring-8、SACLAといった大型実験施設との連携強化、元素戦略(磁石、触媒・電池、電子材料、構造材料分野)への計算物質科学からの貢献への期待などの意見を得ることができた。また、最先端HPCだけでなく、非専門家がPCあるいはクラスターワークステーションでシミュレーションを実行できるよう計算物質科学コミュニティ全体でアプリケーション・ソフトウェアを整備することや、莫大なシミュレーション結果や実験結果を保存・公開する仕組みを整備することなどに対する強い要望を受け、計算物質科学のポータルMateriApps(<http://ma.cms-initiative.jp/>)の整備も進められている。

2.2.2 長期目標と社会貢献

物性科学は 10^{23} ほどの膨大な数の原子、分子多体系から成る自然を理解する営みを通じて、相転移とともに自発的対称性の破れ、集団運動励起やトポロジー励起、マクロ量子現象といった基礎科学を一新する普遍概念をもたらし、素粒子物理学から経済学まで広がるさまざまな学問分野に大きな影響を与えてきた。これらの成果は、20世紀以降の産業・先端技術革新を生み出す基盤となった。トランジスタ、トンネルダイオード、半導体レーザー、集積回路、巨大磁気抵抗素子、CCD(電荷結合素子)、有機ELな

どの革新デバイス、合成樹脂や導電性高分子などの新材料は、ノーベル賞の受賞対象ともなった物質科学の基礎研究が生んだ例である。同じく物質科学の精華である超伝導は、最先端の医療用MRIの超伝導マグネットに使われ、更に超伝導リニアモーターやエネルギー損失のない電力線として実用化されようとしている。20世紀の要素解明から21世紀には集団・階層解明と機能制御の時代に入ったと言われる現代科学の中核として、物質科学における基礎研究のフロンティアでは、量子ホール効果、トポロジー絶縁体、スピニ液体、量子臨界や脱閉じ込めといった新概念が次々に発見され、自然の新たな機構解明への挑戦が続けられている。概念の革新は次世代、次々世代の最先端技術へ展開する研究をますます活性化させていくが、この基礎研究から応用研究、更には産業応用への多段階リレーは、高度な蓄積を持つわが国を含むきわめて限られた国でのみ追求し得る。更に近年、高性能のスペコンを活用することで、近代までに確立された古典力学、量子力学、統計力学に基づいた、ナノ材料の物性のボトムアップ的な予測に期待が寄せられている。計算科学と実験・理論がタイアップし、次世代のデバイスの設計・開発に役立てることで、地球環境を守り、産業振興を助け、社会を豊かにすることにつなげることができる。

(i) 半導体電子デバイス

半導体電子デバイスの高速動作／高集積／低エネルギー消費の観点から、デバイス構造の微細化は依然として要求されている。しかし、微細化のみでは問題解決は困難であり、デバイスを構成する材料の変更、デバイス構造の立体化・複雑化なども行われている。また、半導体デバイスは、これまで広く使われてきた電子デバイス以外にも、スピントロニクスデバイス、熱電変換デバイス、光学デバイスなど、動作原理・用途を広げている。これらのデバイス機能予測・製造プロセス設計に計算科学的アプローチで貢献することが、当該分野に課せられた使命である。この使命を達成すべく、シミュレーションのモデルサイズ・時間スケールの拡大、予測可能な物理・化学現象の増加、計算精度の向上が目標である。

(ii) 光・電子融合デバイス

現有の半導体電子デバイスをベースとして更に高機能化したものは、次世代量子デバイスの有力候補の一つになると予想されるが、その一方で高機能化半導体電子デバイスの実現に際しては、微細化、高速化、大容量化、低消費電力化、熱対策等、解決すべき非常に大きな障壁があることも事実である。これらの問題に対して相補的な、あるいは理想的には根本的な解決策を与え、更には電子デバイスにはない高機能性を備えた次世代量子デバイスの候補として、光磁気効果を利用した光・電子スピニ機能性デバイスが考えられる。電子のスピニ自由度を不揮発性の情報キャリアとして用いるスピントロニクスと、発熱の無い情報伝送を可能にする光回路技術を融合することで、磁性ナノ構造体を活用した超低消費電力デバイスの実現が期待される。他の候補として、金属ナノ構造体を集積することで誘電率・透磁率の自由な制御を可能にするメタマテリアル技術を活用した3次元光メモリやナノアンテナなどの光デバイスも挙げられる。更に将来の極限的デバイスとして、アト秒科学の知見を応用し、ペタヘルツ高周波光で物質中の電子を駆動する超高速光駆動デバイスが想定されている。今後はこのような新しい動作原理に基づく光・電子融合デバイスの第一原理計算に基づく理論設計を積極的に進める。また、超微細レーザー加工はこれら新奇デバイスのオンデマンド、少量多品種生産技術の実現のために不可欠の技術であり、加工効率のさらなる向上のためには物質表面におけるレーザー加工過程を高精度にシミュレーションする必要がある。これらの計算のため、「富岳」コンピュータで開発された電子・電磁場ダイナミクス法プログラムの更なる超

並列化を進めるとともに、GPGPU 等の演算加速装置を積極的に活用することにより、デバイスの理論設計を加速させる。

(iii) 強相関量子物質

量子論の勃興期から現在に至るまで、分子や固体中の多体電子、液体ヘリウムや中性原子気体などの量子多体系が示す、超伝導あるいは超流動を始めとする古典力学からは予想できない量子力学的現象は、人類の好奇心を刺激し、産業の発展の基礎となってきた。中でも、強相関多体量子系は新しい量子力学的現象と概念の揺りかごとして、高温超伝導、トポロジーで分類される量子ホール相やトポロジカル絶縁体、スキルミオンを始めとする磁性体におけるトポロジカルな欠陥、多極子がもたらす特異な電気磁気応答などの物理を生み出すとともに、走査トンネル分光イメージングや角度分解光電子分光、非弾性中性子散乱および X 線散乱を始めとする観測手段の飛躍的発展を促し、遷移金属化合物、希土類化合物、有機導体などの強相関物質群や、グラフェンを始めとする原子層物質や半導体界面における微細構造、冷却中性原子などの様々な系の探索と理解へと人類を駆り立ててきた。

量子計算機の超越的な計算能力や量子暗号通信の安全性の源でもある“量子もつれ”は、これら新たな物理を統御する中心概念として物質科学においても近年、注目を集めている。新たな量子情報科学の時代の「資源」として研究が進むのみならず、物質科学の最前線の課題である高温超伝導の発現機構や、トポロジーによって分類される分数量子ホール状態および量子スピン液体の素性を理解する上で欠くことのできない概念となっている。さらには、生命科学においても特異な現象や機能、例えば、光合成の高いエネルギー変換効率を生み出す源として、また重力理論と量子多体系を結びつけるホログラフィー原理とも密接につながる性質として、量子力学の草創期から現在まで研究が続けられてきた。

今後の発展に向け、これまで研究が進められてきた強く相互作用する強相関多体量子系を第一原理的かつ高精度に取り扱える数値計算手法を拡張し、実験的な観測手法との定量比較と緊密な連携によって、量子もつれを測り、理解し、制御する方法を開拓していく必要がある。さらに、制御パラメータを突き止め物質設計を支援していくために、機械学習手法を用いて実験データから有効ハミルトニアンを抽出するなどの新機軸を持ち込み、トポロジカル量子計算を始めとする量子情報科学の時代の新たな産業を拓く物質探索手法として確立する。

(iv) 永久磁石・スピントロニクス材料

永久磁石材料は低炭素・脱炭素化、そして持続可能な環境調和型社会の実現に向け、必要不可欠な材料として注目度が増している。現在最強の磁石はネオジム磁石であり、その用途はハードディスク・家電製品・情報機器など多岐にわたるが、今後ますます電動車向け駆動モーター・風力発電機・ドローンなどでも需要の増加が見込まれる。我が国においては資源リスク・価格の観点から、省レアアース・省レアメタルの高性能永久磁石の開発を目指した元素戦略的研究は社会的急務である。

また、日本の情報通信 (ICT) 機器の電力使用量は、現在は全電力使用量の数 % 程度であるが、無策の場合には 2030 年に約 40%、2050 年には約 60% にも上ると予測される。電荷を制御する現代のエレクトロニクスが抱える課題を克服するスピントロニクス技術は、省エネルギーデバイスを実現できる可能性があり期待されている。巨大磁気抵抗効果 (GMR) やトンネル磁気抵抗効果 (TMR) を用いる金属スピントロニクスは一部実用化されたものの、半導体材料やデバイスにおいてスピノ自由度を利用する半導体スピ

ントロニクスは未だ実用の域には達していない。高精度・高速度の第一原理計算により、電子状態や磁性を系統的に調べ物理的機構を解明し、これらを定量化することで高性能化の指針を示すとともに、インフォマティクスの手法を用いて永久磁石・スピントロニクス材料開発速度を加速させ産業競争力を強化する必要がある。

(v) 構造材料

熱エネルギーを効率的に機械・電気エネルギーに変換する耐熱材料や低燃費・省エネルギーに寄与する高比強度軽量材料等の飛躍的高性能化が、エネルギー問題の解決に向けて求められている。こうした材料の強度・信頼性は、ミクロの原子間結合のみならず、メゾスケールの内部組織（析出相・粒界・異相界面・転移・点欠陥・不純物偏析等の集合体）に支配され、それらを系統的に設計・制御することが肝要である。そのために、第一原理による電子構造計算やエントロピーを考慮した自由エネルギー計算、更にミクロからメゾ、マクロをつなぐマルチスケール計算が必須となる。超並列大規模計算機環境を用いて、大規模第一原理計算・自由エネルギー計算を実行するとともに、フェーズフィールド法等により、ミクロ-メゾ-マクロをつなぐマルチスケール計算技術を確立し、内部組織の安定性、微細構造、強度、諸特性を解明する。

2.2.3 課題とその解決に必要な計算手法・アプリケーション

計算物質科学で使われるアプリケーションは非常に多岐にわたる。以下、計算物質科学分野における、現在の主要なアプリケーション・アルゴリズムの中から、凝縮系に対する第一原理計算、電子・電磁場ダイナミクス法、変分モンテカルロ法、厳密対角化、クラスターアルゴリズム量子モンテカルロ法、テンソルネットワーク法についてとりあげる。

(i) 半導体電子デバイス

半導体デバイスの性能向上は、微細化一辺倒ではなく構造や材料改良に拡大しつつある。構造や材料構成要素が複雑化したデバイスでは、少ない数の小規模計算モデルで結論を導くことは困難であるため、百万原子超の大規模モデルかつ多数のパターンを高速に処理する必要がある。演算加速器の有効活用を含め、大規模モデルに対する計算技術を発展させ、課題を解決する。また、電子デバイスのみならず、電子のスピントロニクスへの応用や熱電素子、光学素子などエネルギー変換デバイスへの応用が進んでいる。これらのデバイスの機能予測には、従来の原子構造・電子状態計算コードに、電子や熱などのキャリア伝導特性解析や光との相互作用解析機能追加が必須である。電子伝導については非平衡グリーン関数法や波動関数接合法で10万原子超の第一原理計算を実現している。デバイス開発の要求にこたえるべく、他の機能についても大規模モデル計算コードに実装していくことが必須である。さらに、半導体デバイスの構造や材料の多様化に伴い、第一原理シミュレーションは製造プロセス設計にも利用されつつある。化学反応を扱うプロセス設計には時間スケールの長い解析が必要なものもあるが、時間軸方向の並列化は容易でない。第一原理シミュレーションの高速化が重要であることは言うまでもないが、第一原理シミュレーションの結果を参考にした古典分子動力学計算や機械学習援用の高精度化の検討も必要である。半導体デバイスの機能予測精度向上のため、密度汎関数理論の局所密度近似を原因とするバンドギャップ過小評価を解消する要求は依然として大きい。局所密度近似を超える計算手法は夥しい計算量を

必要とするため、計算精度の向上と半導体デバイスを模せるようモデルの大規模化を両立する計算手法・コードの開発が必須である。

(ii) 光・電子融合デバイス

電気・光学・磁気特性を持った機能性材料は、エレクトロニクス、フォトニクス、スピントロニクス等のデバイスを構築する必須要素である。従来の機能性材料の主役は、シリコン等に代表される半導体ナノ構造体であった。一方、次世代の光・電子融合デバイスでは、光スピントロニクスデバイスにおける磁性ナノ構造体や、メタマテリアル光デバイスにおける金属ナノ構造体等、多種多様な新奇ナノ構造体が活用されていくと考えられる。また、機能についても、光円偏光と電子スピンの情報変換やペタヘルツ高周波によるアト秒での電子制御等、光と電子の複雑な相互作用が活用されていくと考えられる。そのため、従来と比較して複雑な系における複雑な光・電子相互作用の高精度予測が重要な課題となる。すなわち、次世代光・電子融合デバイス材料の機能発現機構解明やその合理的設計のためには、電気・光学・磁気特性単独の機能性だけではなく、ナノ領域において電子と光が露わに結合したような状態に起因する電子・光新奇機能性を統一的に予測できる高精度かつ大規模な第一原理計算手法の開発と実行が必須である。特に、磁性体のための電子スピン自由度を考慮した計算手法や、メタマテリアルに用いられる金属粒子等の量子効果を高速に計算できる手法が必要となる。次世代光・電子融合デバイスの理論設計を実現するためには、時間依存密度汎関数法によるスピン自由度を含めた電子ダイナミクスとマクスウェル方程式によるミクロスコピックな電磁場ダイナミクスが結合した状態を記述するナノ光学理論とその理論に基づく電子・電磁場ダイナミクス法が必須である。

(iii) 強相関量子物質

強相関量子物質の研究を推進し、トポロジカル量子計算の基盤となる物質やより高い転移温度を示す超伝導体を探査するためには摂動論や平均場近似を超えた量子もつれを正しく取り入れる必要がある。そのためには、厳密対角化や量子モンテカルロ法などの量子ゆらぎの効果を正しく取り込んだ第一原理的手法が不可欠であるが、電子数の指数関数で増大する計算規模や負符号によって、現実物質を定量的に扱うことは難しい。

一方で、様々な表現を用いた変分波動関数法が大きく発展し、1次元系や2次元・3次元の少数多体系において計算精度の一つの基準となってきた密度行列繰り込み群法(DMRG)に匹敵するより適用範囲の広い方法が開発され発展している。 $O(10^5)$ もの変分パラメータを最適化することで高精度を実現する多変数変分モンテカルロ法(mVMC)や、テンソルネットワーク法、さらには人工ニューラルネットワークを用いた変分波動関数について研究が進み、DMRGに迫る精度で、 $O(10^3)$ 電子あるいはそれ以上の系の基底状態について計算することが可能となっている。深層ニューラルネットワークを用いた波動関数も高精度を示すようになってきた。さらに実空間繰り込みによるテンソルネットワーク法や量子埋め込み理論を用いて無限系を解析する手法も発展している。

今後は、基底状態のみならず、有限温度や分光学スペクトルのシミュレーション、非線形非平衡状態への適用も進むと考えられる。統計力学的な量子多体模型のみならず、強相関第一原理電子状態計算で得られた低エネルギー有効ハミルトニアンの高精度解析も可能になってきており、既存の強相関量子物質の高精度な物性再現のみならず、物質設計指針の導出や、所望の物性を示す物質を探索するための仮想スク

リーニングへも適用可能性が広がっており、新材料・新デバイス開発の加速に貢献することが期待される。

(iv) 永久磁石・スピントロニクス材料

材料の特性を定量的に評価可能な第一原理計算は、永久磁石・磁性材料の材料探索においても期待が大きい。永久磁石材料に適用する際には、(1) 不規則系の配置平均、(2) 強相関電子系の記述、(3) 有限温度特性の評価が要求される。(1) に関しては、コヒーレントポテンシャル近似 (CPA) に代表されるシングルサイト近似が有効であるが、短距離秩序性が重要な系に対しては高速・高精度で非局所 CPA 計算が可能なアルゴリズムを開発する必要がある。(2) に関しては、自己相互作用補正・最適化ポテンシャル・非局所ポテンシャル・動的平均場近似など、多くの手法が提案されているが f 電子系を取り扱える決定的な交換相関ポテンシャルは未だに開発されていないため、さらなる発展が望まれる。(3) 現状の第一原理計算は絶対零度の物理量に対してのみ有効なため、有限温度特性の評価は容易でない。有力であるのはシングルサイト CPA と静的近似を用いて、フォノン・マグノンなどの有限温度における素励起を取り扱うことであり、実際に遷移金属合金や $\text{Nd}_2\text{Fe}_{14}\text{B}$ などに適用され有限温度磁性が議論され始めている。

スピントロニクスの分野においては、単純なバルク構造のみならず界面・超格子複合系、そしてスキルミオンやヘッジホッグ構造などの複雑なトポロジカルスピinn構造の計算需要が増加すると考えられる。そのため、全電子 $O(N)$ レベルでスピン軌道相互作用とノンコリニア磁性を取り扱える計算コードの開発と整備が必要である。また、スピントロニクスデバイスはスピン依存伝導特性の制御により実現される。それゆえ、今後は GMR や TMR などの素子構造を模した大規模系において、第一原理計算と久保公式に立脚し、パリスティック伝導特性だけでなく微視的な電子散乱過程（有限温度におけるフォノンやマグノンを含む）を考慮した伝導特性を、 $O(N)$ で計算可能な手法が必須となる。

(v) 構造材料

高精度自由エネルギー計算、第一原理計算、フェーズフィールド法などを連成したマルチスケール計算により、結晶相・化合物相、粒界・界面・欠陥の安定性・強度特性を第一原理から予測し、強度や耐久性、耐熱性を併せ持つ材料や軽量高強度の材料の開発を目指す。

フェーズフィールド法は、合金材料の組織形成をシミュレートする連続体モデルであり、実用合金の実プロセスにおいて問題となる組織形成過程の解析に応用されつつある手法である。合金凝固過程を対象とすると、流体の速度場、温度場、溶質拡散場および秩序変数の方程式を連立して解くことになる。フェーズフィールド法において、デンドライト組織形成を高精度にシミュレートするためには、 $10^{-8} \sim 10^{-7} \text{ m}$ 程度の空間分解能を要する。一方で、デンドライト集団の統計的挙動の議論において対象とする空間スケールは 10^{-3} m 程度のスケールである。したがって、凝固組織形成の高精度シミュレーションには、およそ $10^{13} \sim 10^{14}$ 規模のメッシュ数が必要となる。並列化による大規模計算における主たる演算は陽解法によるステンシル型の計算であり、袖領域のデータ通信が並列化における通信時間の主となる。

これらのシミュレーションには、典型的にはアレイジョブ、すなわちパラメータを変えた互いに独立で比較的小～中規模のジョブを複数（数十～数万）実行することが必要となる。小規模ジョブは 1 ノードで実行され、ノード間通信を必要としない。中規模ジョブは数ノード～100 ノード、大きくても数百ノードで実行される。1EFlops 級の計算機を用いることにより、数千～数万ジョブ規模の計算結果を一度で得ることができる。アレイジョブの結果からマクロな物理量を取り出し、次の計算へ素早くフィードバックす

るためにも、アレイジョブが一度に終了することは非常に重要である。

2.2.4 ロードマップ

年代 課題	2022～	2024～	2026～	2028～	2030～	2032～
半導体 デバイス	10～50万原子規模の電子・スピノ・熱電・光デバイス機能予測シミュレーション					
光・電子融合 デバイス	1万原子規模の数値ポンプ・プロープ実験	1万原子規模のレーザー加工シミュレーション	10～50万原子規模のハイスクループットシミュレーションによるデバイス創製プロセス設計	実在系の数値ポンプ・プロープ実験	実在系のレーザー加工シミュレーション	
強相関 量子物質	100原子系の分光学スペクトルのシミュレーション		1000原子系の分光学スペクトル計算	強相関量子物質における分光学測定による量子もつれ検出		

図2.2 ナノサイエンス・デバイス分野ロードマップ

2.2.5 必要な計算機資源

2032年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13節の表2.1に記載した。また、計算機資源の見積りの詳細は第4.2節に記載している。

参考文献

- [1] 押山淳, 天能精一郎, 杉野修, 大野かおる, 今田正俊, 高田康民. 計算と物質 (岩波講座計算科学第3巻). 岩波書店, 2012.
- [2] 日本学術会議第三部拡大役員会・理学・工学系学協会連絡協議会. 理学・工学分野における科学・夢ロードマップ. <http://www.scj.go.jp/ja/info/kohyo/kohyo-21-h132.html>, 2011.

2.3 エネルギー・材料

2.3.1 分野の概要

資源小国である我が国にとり、エネルギー問題の解決は差し迫った重要な課題である。当該分野においては、2015年から2019年にかけて、ポスト「京」（「富岳」）で取り組むべきプロジェクト「フラッグシップ2020重点課題5：エネルギーの高効率な創出、変換・貯蔵、利用の新規基盤技術の開発」が実施された。このプロジェクトでは、スーパーコンピュータ「京」を活用した数値シミュレーション研究が実施されており、「富岳」を活用するための基盤研究・開発が実施された。そこでは、以下の持続可能なエネルギー新規基盤技術を確立することを目標としていた。

- (i) 新エネルギー源の創出・確保—太陽光エネルギー太陽電池、人工光合成による新エネルギーの創出・確保
- (ii) エネルギーの変換・貯蔵—電気エネルギー燃料電池、二次電池によるエネルギーの変換・貯蔵
- (iii) エネルギー・資源の有効利用—化学エネルギーメタンやCO₂の分離・回収、貯蔵、触媒反応によるエネルギー・資源の有効利用

(i)–(iii)の課題においては、複雑で複合的な分子からなる構造体によってもたらされる化学・物理過程が中心的な役割を担っている。このような分子のミクロな挙動の理解は、新規材料の合理的な理論設計と機能制御の開発において最も重要なファクターである。しかしながら、そのような分子レベルの情報は実験のみからでは得ることは難しく、開発指針を経験から定めることも困難である。重点課題5で確立された分子論に基づく計算科学技術は、産業や社会問題と密接に関係しており、実験研究と相補的に用いることが可能である。

重点課題5で確立された基盤技術により、以下のような産業利用が期待される。(1) 安価・軽量な高効率太陽電池材料の開発、(2) 人工光合成系の実現による代替エネルギーの創出、(3) 自動車産業の根幹を変革する可能性を持つ低コスト電気自動車、燃料電池車の世界に先駆けた開発、(4) メタンハイドレートからのメタン資源の確保、(5) 高性能触媒の開発によるエネルギーを大量消費する工業過程からの転換、(6) CO₂による地球温暖化問題の解決と化石燃料の安定的利用の実現。

長期的な観点から、高効率かつ安価な太陽電池や二次電池は、発展途上国のエネルギー問題を改善する可能性を持ち、地球規模での資源・環境問題の解決によって国際貢献に繋がる。重点課題5で確立された技術は、単にエネルギー問題に限られることなく広く産業に応用され、我が国の産業の革新的発展に大きく貢献することが期待される。

2.3.2 長期目標と社会貢献

重点課題5では、分子のミクロな振る舞いが中心的役割を果たすエネルギー技術に対する計算科学基盤技術を確立することが目標として挙げられていた。そして、確立した基盤技術を活用することで、「富岳」運用開始5年後には、(i) 高効率・低コストの太陽電池や新規人工光合成系を設計・提案、(ii) 新しい燃料・二次電池に向けた材料探索や動的過程等の解析を進め、新規システムを設計、(iii) メタンやCO₂の

効率的かつ革新的な分離・回収、貯蔵、変換法の確立を行うことが目標とされた。また、「富岳」運用開始10年後には、安価かつ高効率な太陽電池、人工光合成系の実現、高性能電池の開発による電気自動車や燃料電池車の普及、メタンハイドレートのエネルギー資源としての利用が目指された。また、CO₂による地球温暖化問題の解決の見通しをつけ、化石燃料の安定的利用がなされるような環境を実現することも目標として挙げられていた。

この一連の目標達成に向けて、2020年度から開始されている「富岳」成果創出加速プログラムにおける重点課題5の後継にあたる課題として、下記の3つの課題が現在実施されている。

- (A) 「富岳」を活用した革新的光エネルギー変換材料の実現（代表: 中嶋隆人、理化学研究所）
- (B) 次世代二次電池・燃料電池開発によるET革命に向けた計算・データ材料科学研究（代表: 館山佳尚、物質・材料研究機構）
- (C) 環境適合型機能性化学品（代表: 松林伸幸、大阪大学）

各課題の具体的な目標と社会貢献は下記のとおりである。

(A) 「富岳」を活用した革新的光エネルギー変換材料の実現

「富岳」を活用したハイパフォーマンス材料シミュレーション・インフォマティクスを中心として、世界を牽引する実験グループ・企業組合と連携した革新的な光エネルギー変換材料の社会実装を行う。創出されるデータに基づき、光触媒による水素製造・ウィルス不活性化による感染症対策・高効率な太陽電池の産業レベルでの実現を目指す。

(B) 次世代二次電池・燃料電池開発によるET革命に向けた計算・データ材料科学研究

将来のET（エネルギー・環境技術）革命において中心的役割を果たす二次電池および燃料電池の次世代技術の開発や実用化に向けて、重要課題の微視的機構解明およびそれをもとにした材料探索・反応制御の指針提案を、「富岳」を用いた先端的計算・データ材料科学研究により推進し、熾烈な国際競争の中での我が国の産業競争力強化、カーボンニュートラル・脱炭素社会実現に貢献する。

(C) 環境適合型機能性化学品

付加価値額が全産業の約1割を占める化学産業において中核を担うポリマー材料の環境適合化を目指す。環境改善に資する分離膜の高度化、機能性材料における環境リスク成分の使用削減、および、バイオ由来樹脂の機能強化のために、全原子分子動力学(MD)シミュレーションと自由エネルギー計算を行い有用なポリマー構造を探索する。また、樹脂/金属界面の量子力学/古典力学(QM/MM)計算によって、異種材料からなる接着界面の劣化機構を分子レベルで解析する。

2.3.3 課題とその解決に必要な計算手法・アプリケーション

「富岳」成果創出加速プログラムにおける上記3課題は、実験研究者や企業研究者との連携に基づき、計算と実験が相互的な関係を持ち課題解決を目指すものである。これらの課題の各エネルギー新規基盤技術の実現において、新規材料の合理的な理論設計と機能制御の開発は最も重要なファクターとなっている。現状の計算科学研究では、理想的な要素に対するモデルの取り扱いにとどまっている。実在系の分子は、構成要素が複合して形成されており、直接的な分子過程の記述が不十分である。また、分子過程は、

複雑に絡みあった電子や分子の振る舞いが中心的な役割を担っている場合が多いので、全て構成要素を考慮した全系に対する大規模な数値シミュレーションが必要不可欠である。このような非常に複雑な現象の解明は学術的にも意義が高く、「富岳」あるいは富岳 NEXT なくしては不可能である。そのため、「富岳」を用いた超大規模並列シミュレーションにより、全系の計算を実行して、実験的に観測されている分子過程を世界に先駆けて解明することが必要である。さらに、富岳 NEXT 時代においては、実験に先駆けて化学現象を解明し、さらに分子設計や材料設計の設計指針を提案することが求められるであろう。計算科学者の立場から分子過程の知見を実験研究者や企業研究者に提供し、互いに協力しながらエネルギー問題を解決することが可能となる。以上のように、「富岳」のみならず富岳 NEXT の利用による明確な投資効果が見込まれる。

当該分野が対象とする分子・材料系や化学・物理現象は多岐に渡るため、その解決に必要な計算手法やアプリケーションソフトウェアも多彩であるが、ここでは「富岳」成果創出加速プログラム「「富岳」を活用した革新的光エネルギー変換材料の実現」を一例として、問題解決に必要な計算手法とアプリケーションソフトウェアについて述べることにする。

この課題では、再生可能エネルギーの主力化に必要な高効率な人工光合成と太陽電池に加え、感染症対策としての高機能を有する半導体材料の開発を、「富岳」を用いたハイパフォーマンスコンピューティングとインフォマティックスによるサイバー空間での物質探索を行い、企業・実験研究者等と緊密な連携を行うことによるサイバー・フィジカルシステムを構築することを目指している。産業力を強化する光触媒、感染症対策機能を有する光触媒、太陽電池の用途別に特化した革新的な光エネルギー変換材料の開発を行い、データマネージメントの構築にも同時に取り組まれている。この課題で進める光エネルギー変換材料に対するサイバー空間探索&サイバー・フィジカルシステムを図 2.3 に示す。サイバー空間において、「富岳」の大規模計算能力をフル活用することによって、人工光合成材料、感染症対策材料、太陽電池材料の大規模データ生成が行われている。この際、重点課題 5 で開発されてきた量子化学計算プログラム NTChem および GELLAN やバンド計算プログラム VASP および Quantum ESPRESSO による超並列計算によって網羅的な第一原理計算の大規模データが生成される。さらに、この大規模データを AI に学習させることで、材料の機能発現に不可欠なドーピングと元素欠損まで含めた材料探索を可能とするサイバーループが構築される。AI 学習においても、深層学習において産総研 ABCI の 15 倍以上の性能を持つ「富岳」を有効に活用している。さらには、連携機関として本テーマに関わる国内産学のトップクラス研究グループにて実施される材料合成・評価データを、理論計算データと統合しマルチモーダル学習や転移学習することで、AI によるフィジカル空間での材料探索も実施されている（サイバー・フィジカルループ：図 2.3）。この探索によって提案材料候補を連携機関にて実証した結果を新たなデータとして AI を成長させるとともに、物性種の追加（例えば、価電子帯や伝導帯領域の状態密度やバンド間のギャップなど）および計算精度・探索範囲のリファインメントとして大規模計算にフィードバックすることで新たなデータ生成に繋げることになる。このようにサイバー技術とリアル技術を融合させたサイバー・フィジカルループを実践することで、広域・外挿探索を可能とし、最終的には新材料候補の提案が目指されている。

まとめ

これらの課題解決により、孤立系や部分系における单一現象の科学から脱却サイバー・フィジカルループ均一性を有する電子、分子の複合現象を統合的に捉え得る新しい学術的視点を確立することができ、



図 2.3 サイバー・フィジカルループによる材料設計・探索

科学的にも大きな革新をもたらすことが期待される。現在当該分野において開発されている高精度量子化学計算、大規模第一原理計算、超大規模分子動力学計算などの方法論、アプリケーションは化学、物理、材料、生物分野等に適用できるものであり、エネルギー問題に限ることなく、広く他分野での展開も期待される。

2.3.4 ロードマップ

2.3.5 必要な計算機資源

2030年頃における課題解決のために必要な計算機資源を代表的な5つのアプリケーションについて、2.13節の表2.1に記載した。また、その他のアプリケーションも含め、計算機資源の見積りの詳細は第4.3節に記載しているので併せて参照されたい。

2.4 生命科学

2.4.1 分野の概要

本章では、生命科学に関する計算科学分野のうち、主に生体分子シミュレーションと、ゲノム解析・遺伝子ネットワーク解析に焦点を当てる。量子化学計算を含む創薬関連分子シミュレーション、脳科学、および細胞臓器レベルのシミュレーションについては 2.6 および 2.5 を参照してほしい。

生体分子シミュレーション

生体分子シミュレーションは、主に分子動力学（Molecular Dynamics (MD)）法を用いて、生体分子の挙動を分子レベルで計算する。対象はペプチドやタンパク質などの小分子から、細胞膜、ヌクレオソーム/クロマチン、リボゾーム、ウイルスなどの超分子まで多岐にわたる。生体分子のダイナミクスの時間スケールはナノ秒からミリ秒まで幅広いため、注目する現象に合わせて解像度（粒度）を選択する必要がある。生体分子シミュレーションは計算機演算能力の向上とともに、実験では解明できない生体分子の反応機構を分子レベルで解明したり、新薬の設計を支援したりするための重要な手法となっている。

世界的には、全原子モデル（2.4.3 を参照）を用いた MD シミュレーション（全原子 MD）が広く普及しており、生体分子の長時間（ミリ秒程度）のダイナミクスを再現するシミュレーションと、巨大分子複合体や複数の生体分子を含む大規模系（1～10 億原子程度）のシミュレーションの二つの潮流がある。長時間ダイナミクスについては、GPU や MD 専用計算機の発展が牽引している。日本においては MD 専用計算機「MDGRAPE」の開発が継続しており、MDGRAPE-3[1] が 2006 年に世界で初めて 1 PF/s の演算性能を達成するなどの先行した試みが進められてきた。その後米国で 2008 年に開発された専用計算機 ANTON によって、汎用機の 1,000 倍近い強スケーリング性能が達成され、全原子モデルを用いたタンパク質のダイナミクスをミリ秒規模で再現するなど、この分野に大きな飛躍をもたらした [2, 3]。これらの専用計算機は主に数十万原子程度の系をターゲットにしてきたが、2021 年に発表された ANTON-3 では、7 千万原子系のシミュレーションを一日で約 2 マイクロ秒進めることが可能になるなど、継続的な進歩を続けている [4]。日本でも ANTON-1 並みの性能を目指して MDGRAPE-4A が開発されたが、現状では 10 万原子系のシミュレーションを 1 日 1 マイクロ秒行える性能に留まっている [5]。またこのような時空間スケールの拡大に加えて、より幅広い構造空間を効率よく探索（サンプリング）したり、反応座標に沿った自由エネルギープロファイルを得るために様々なアルゴリズムの開発が進められている。このような計算は処理や通信が複雑になるため、ANTON 等の専用機ではユーザが直接実装することが難しい。こうしたアルゴリズムの開発と利用にあたっては、引き続き高速な汎用機が重要な役割を果たすと共に、今後専用機のユーザビリティを高めるアプローチが必要であり、専用アクセラレータと汎用機の高度な結合が求められる。

ゲノム解析・遺伝子ネットワーク解析

ゲノム解析や遺伝子ネットワーク解析は、次世代 DNA シークエンサやマイクロアレイなどの実験的データ（ビッグデータ）が大量に得られるようになった上で大きく発展した生命情報科学（バイオインフォマティクス）分野である。ゲノム解析は、次世代 DNA シークエンサから得られる大量の断片的配列

データを全長配列データに変換したり、遺伝的な多様性や変異を求めたり、さらに疾病に関する遺伝的な多様性や変異を同定したりする。遺伝子ネットワーク解析は、遺伝子発現データから複雑な計算によって遺伝子間の関係を予測・推定する。遺伝子発現データは、計測に用いるマイクロアレイや次世代シークエンサーの試薬のコストが下がり、またシングルセル RNA シークエンシング技術などの登場により、現在では数万サンプルを一度に取得・計測し解析することが可能になっている [6, 7]。得られた膨大なデータの解析には巨大な計算機リソースが必須であり、「富岳」を始めとするスーパーコンピュータや商用クラウドシステムは本分野における必須インフラとなっている。また、FPGA や GPU を活用した専用解析システムも登場しており [8, 9]、クリニカルシークエンスなどの現場で顕著な効果を発揮している。また少数サンプルや個別サンプルを用いた解析技術も登場し、これまで遺伝子ネットワーク解析に用いられなかったデータセットの利活用も可能になりつつある [10, 11]。さらに「富岳」などの計算リソースの登場により、解析対象を事前に絞らずに公開データベースに登録されている遺伝子発現データに対して網羅的にネットワーク推定ソフトウェアを適用するというアプローチもすでに行われている。今後はこうして得られた網羅的な遺伝子ネットワークに対して個別サンプル解析技術を組み合わせるという方法が可能になり必要な計算がデータセット数の増加分にとどまらず、組合せ爆発を起こすと予想される。さらにサンプル単位のアノテーション技術などによりデータセット横断的に新しいデータセットを定義することも可能で [12]、遺伝子ネットワーク解析対象となるデータは単に公共データベースにあるデータセット数にとどまらない。このような大規模な網羅的解析によって得られる情報は、病態の予測や、新規薬剤ターゲット遺伝子の探索、個人の体質に合わせた薬剤の選択などへの活用が期待されている。

AI・機械学習

近年、生命科学分野の技術革新に伴って、ゲノムや生体分子の構造・機能・ネットワークに関するデータが爆発的に増加しており、AlphaFold [13, 14] に代表される大規模データを AI・機械学習の技術によって解析・予測する手法が発展してきている。最近では、AI・機械学習は生体分子シミュレーションとの融合研究が進んでおり、いわゆる AI 力場に代表されるような分子シミュレーション自体を高精度・高速化する手法の開発や、分子シミュレーションを実験と捉えることで新たな機能を持った生体分子や薬候補を探索・デザインする研究が活発に行われている。AI・機械学習は、今後の生命科学分野における計算科学研究を一層加速するもので、計算による生命現象の理解を助け、また計算に基づく創薬や医療をより現実的なものにすると期待される。

他分野との連携

今後の生命分野の計算科学に大きな影響を与えるものとして、様々な構造生物学的実験手法によって得られたビッグデータが挙げられる。上記の次世代シークエンサーで高速に読み取られたゲノム情報の解析は、そのようなビッグデータ解析の代表例であるが、X 線自由電子レーザーによる高輝度パルスを用いた生体試料の時分割構造観測データや、クライオ電子顕微鏡を用いた単粒子解析により得られる生体分子の多型構造データ、電子線トモグラフィーにより得られる細胞内のメソスケールな構造体に関する顕微鏡画像などが重要な情報源となるであろう。また、新しく盛んになりつつある高速分子間力顕微鏡などによる実験からはこれまで得られなかつたリアルタイムの構造変化なども観測できるようになる。これらのビッグデータを元にした三次元イメージングだけでなく構造多様性の評価や時間軸も含めた 4 次元イメージングには、計算科学分野との連携が不可欠である。これらの実験データから得られる生体分子の構造情報は

シミュレーションを行う上で必要になる一方、解像度が十分ではない場合や観測で捕らえられない不安定構造が生じる場合がある。実験からは得られない情報をシミュレーションにより予測する相補的研究により一層、生体高分子の機能理解に役立つ研究を行うことができるようになるはずである。

2.4.2 長期目標と社会貢献

長期目標

生命現象の複雑さや多様さなどにより、これまでの生命科学の分野では異なる計算手法やモデルが用いられてきた。また、計算機の演算性能の限界により十分な時間およびサイズの計算が困難だった。今後、長期的研究ではその制約の多くが取り払われ、研究したい生命現象をまるごとシミュレーションできる日が来ると期待される。そして、理論・シミュレーションの専門家だけでなく、実験・医療関係者にもシミュレーションの有効性が明らかになっていくと思われる。そこで必要となる理論や手法の開発を行い、有効な実験とシミュレーションの共同研究を行っていく必要がある。更に、今後の生命系のシミュレーションでは、予測可能性という点が重視されると思われる。下記では、主に MD シミュレーション、ゲノム解析、遺伝子ネットワーク解析における長期的展望と目標について記述する。

(i) 生体分子シミュレーション

全原子 MD では、計算可能な時間スケールおよび空間スケールが更に増大していくが、研究対象はあくまで個々の研究の必要性によって多様性を持つと思われる。例えば、溶液中のタンパク質の折れたたみなどを解析するためには、10万原子程度の系で十分であるが、マイクロ秒から数ミリ秒に至る長時間の計算あるいはレプリカを用いたアンサンブル計算が必要である。分子モーターや膜輸送体などの構造変化を通して機能を発現する過程を見るためには、やはり数ミリ秒のシミュレーションが必要である。その場合も脂質膜も計算することを考慮すると、数十万から百万原子程度のサイズになる。更に大きなリボゾームやウイルス、ヌクレオソーム複合体などの生体超分子複合体、細胞環境を考慮したシミュレーションなどでは、数百万原子から1億原子に至る巨大な系のシミュレーションとなり、計算可能な時間スケールはやはり、その時代の計算機の限界に依存する。全原子 MD を古典分子力場を用いて計算する場合には、すでに手法が確立しているため、その手法自身をハードウェアで組んだ MD 計算専用計算機が今後とも計算速度はリードしていくと思われる。しかし QM 計算とのハイブリッドや、自由エネルギー計算を目的として、バイアスプロテナムを利用した計算法など、今後開発される新たな計算法では、汎用計算機の利用が重要な役割を果たす。GPU などの汎用の加速器は、いずれの計算においても既に大きな演算速度の加速を既に実現している。また、高並列な計算機で全原子 MD を行う場合には、多数のレプリカを疎結合して系の統計的性質を解析するアンサンブル計算の重要性が増していくと思われる。

一方粗視化 MD では、100nm から $10\mu\text{m}$ 規模で、ミリ秒から秒規模の現象が対象となる。具体的には、シグナル伝達系などのタンパク質の相互作用ネットワーク、転写や翻訳過程の構造機能計算、核内の遺伝子動態解析、膜や各種オルガネラを含む細胞系、ウイルス動態などの解明が重要な課題となる。課題としては、対象のシステムを適切に記述するモデルの導出が最初に挙げられる。モデルの導出には、全原子 MD や実験研究のデータを基に、機械学習や AI を利用して古典分子力場のパラメータ作成と精密化が行われることが一般的になりつつあり、今後もこれらの手法が発展するだろう。また、少ない粒子数（縮約

された情報量)で分子の性質や実験情報の再現を求められるため、粗視化モデルでは、分子や表したい性質によって力場の関数形をカスタマイズする必要がある。そのため、従来は粗視化モデルの開発者がMDソフトの開発にも関わり、モデルとMDソフトウェアが密接に結びつくことが多かった。しかし、近年では細胞内環境やLLPS(液液相分離)など異なる種類の分子が含まれるシステムが研究対象となり、異なるモデルを同時に利用でき、利用者により力場の導入・変更ができるようなプラットフォームを持つMDソフトウェアの構築が求められている。次に、全原子と粗視化という異なる解像度の分子モデルシームレスな接続も課題である。異なる解像度のモデルをつなぐ手法(バックマッピング)は広く研究されており、機械学習やAIを利用する手法も提唱されている。これらの課題を克服する方法論の進展に応じて、より高次の細胞生物学的な課題に適用されていくであろう。

(ii) AI, 機械学習、ベイズ統計

近年、次世代シークエンサやクライオ電子顕微鏡による単粒子構造解析等生命科学分野における技術革新に伴って、人類がアクセス可能な生命科学データは爆発的に拡大を続けている。一方、深層学習に代表されるような様々な機械学習・AIの技術も近年大幅に進展しており、特に画像や文章等の大規模データに対する学習を行うことで部分的には人間に近い、或いは人間超えるレベルでの学習・予測・生成が可能になってきている。この背景には、GPUやTPU(Tensor Processing Unit)等の機械学習に適したハードウェア開発があり、またそれらを世界中の開発者が有効活用できるようにしたソフトウェアの成熟及びオープンソース文化が存在する。近年、AI・機械学習を生命科学データに応用した研究が注目を集めており、AlphaFoldに代表されるように一部では計算によって実験に近い精度での予測が可能になっているものも存在する。AI・機械学習は分子シミュレーションにも応用され技術開発が進んでおり、大規模化するシミュレーションデータの解析、より高精度な力場の学習、データ同化によるシミュレーションと実験データの連携などが実施されている。また単にAI・機械学習によって、なんらかの現象や機能を予測するのみならず、新しい現象の発見、新しい機能を持った生体分子・薬候補分子の自動設計など、所謂データ駆動型サイエンスも研究開発が進んでいる。分子シミュレーションを含む計算とAI・機械学習による連携研究は生命科学分野における計算科学研究に不可欠なピースになりつつあり、今後より重要性を増すものと思われる。

(iii) 分子細胞シミュレーション

分子細胞シミュレーションについては、(I)混雑環境が基質(薬剤分子)の標的タンパク質への結合やタンパク質-タンパク質複合体形成などに与える影響を結合・解離定数(熱力学量)や速度定数(動力学量)の定量化によって明らかにすること、(II)細胞膜を含めた分子スケールでの細胞環境モデリング及びシミュレーションの実現、などが目標として挙げられる。これまで、細胞質を模倣したタンパク質混雑環境(数万～数百万原子程度)や細胞質(1億原子程度)の数百ナノ秒からマイクロ秒スケールの分子シミュレーションにより、タンパク質やタンパク質-基質間の非特異的相互作用が、タンパク質や基質の動態に与える影響に対する分子レベルでの知見が蓄積されている。そこからさらに、(I)で挙げた結合・解離に対する定量的な解析が実現すれば、混雑化に伴う薬効低下のメカニズムに対するより詳細な議論が可能になり、混雑環境を考慮した新規薬剤設計指針の構築への貢献が期待できる。これを達成するためには、従来以上に複数のタンパク質が作り出す不均一な環境をサンプリングすることが要求されるので、数十万～

数百万原子規模の系に対してマイクロ秒スケールのシミュレーションを数日レベルで多数実行できるような計算機環境及び計算プログラムの整備が必要となるだろう。同時に、シミュレーション結果から結合・解離定数や速度定数を求めるために従来用いられている、自由エネルギー摂動法やマルコフ状態モデル（MSM）法などについても、細胞質のような不均一環境においても適用できるように方法論及び計算プログラムの拡張を行うことが有効である。一方で、(II) に挙げたように、細胞膜を含めるなど細胞環境に対する分子解像度での記述をより現実系に近づけていくことも、細胞内動態の物理化学的理を深化させる上で重要と考えられる。現在、細胞膜を含めた細胞のマルチスケールモデリングを実現するプログラム群の開発が複数の研究チームから報告されており、細胞内で起きる生体現象をさまざまな解像度から明らかにしようという機運が高まっている。この目標を達成する上で、クライオ電子顕微鏡に代表される種々の実験データや粗視化モデルでのシミュレーション情報の取り込みにより分子レベルでの信頼度の高い細胞環境モデルを構築すること、そしてそのモデルをシミュレートする一連の方法論及び計算基盤を確立していくことが必要となるだろう。DNA がヒストンタンパク質に巻き付いたヌクレオソームの分子動力学シミュレーションによって、ヒストンタンパク質からの DNA 解離、ヌクレオソーム構造の形成・崩壊など、ヌクレオソーム単体のダイナミクスに関する理解はこれまでに大きく進んできた。現時点では、ヌクレオソームと小さなヌクレオソーム結合タンパク質の複合体（水分子を含めて約百万原子系）について、1 マイクロ秒の分子動力学シミュレーションを「富岳」32 ノードを用いて実施するには約 50 日かかる。将来は、ヌクレオソームが多数つながった多量体とヌクレオソーム結合タンパク質との超分子系（約千万原子系）の長時間分子動力学シミュレーションが必要となると予想されるので、これを実施するには「富岳」以上の計算機が不可欠である。また、ヌクレオソームにおいては、天然変性状態であるヒストンタンパク質の末端部位（ヒストンテール）のサンプリング不足が大きな問題となっており、拡張アンサンブルの開発が必要である。更には現在の力場は安定した立体構造を過剰に安定化する傾向があるので、天然変性状態を再現する力場の改良も必要である。以上から、今後のヌクレオソーム研究では、計算機性能の向上とともに、拡張アンサンブルや力場の改良も重要な課題となる。

(iv) ゲノム解析・バイオインフォマティクス

次世代 DNA シークエンサーによるゲノム解析については、より大規模な検体数の解析によるマッピング、遺伝的多様性や変異の同定が目標となるであろう。一方で、バイオインフォマティクスにおけるもう一つの代表的な研究として、遺伝子発現データの解析について述べる。現在では、ハイスループットなプロトームデータやメチロームデータなど、遺伝子発現（=トランスクリプトーム）データ以外にもさまざまなハイスループットデータが計測可能になっている。遺伝子発現データはこれらのうち最も過去から存在する定番的なものとなっているが、現在でも網羅性やコスト、データ量の点で圧倒しており解析対象として基本的な必須データとなっている。バイオインフォマティクスにおいて今後は、長期的にどのような観測技術がどの程度のコストで利用可能になるかは非常に流動的であり不明確な部分が多い。したがってここでは非常におおざっぱな議論をする。遺伝子発現データの解析、なかでもベイジアンネットワークを用いた遺伝子ネットワーク解析は、遺伝子発現データから複雑な計算によって遺伝子間の関係を予測・推定するもので、必要とする計算資源は他の発現データ解析やネットワーク解析手法と比べて圧倒的に多い。遺伝子発現データは、かつて計測に用いるマイクロアレイが非常に高価だったこともあり、数サンプルを用いて特定の疾患に特異的な遺伝子を発見するなどの研究が行われていた。しかし近年、ゲノム解

析技術のコスト低下と相まって次世代シークエンサーを用いた発現データ計測技術が普及し、またシングルセル RNA シークエンシング技術も登場し、2022 年時点では数万サンプルを一度に取得・計測し解析することも可能になっている [6, 7]。研究で計測された遺伝子発現データは、論文などの出版の際に公共データベース上に登録し公開することが一般的である。図 2.4 は、代表的な公共データベースである NCBI Gene Expression Omnibus(GEO) に登録されている遺伝子発現データのサンプル数と今後の予想である。現在およそ 500 万サンプル・2 万データセットが登録されており、2032 年頃には 1,300～1,500 万サンプルほどになると予想されている。遺伝子ネットワーク解析は、多くの計算資源を必要とし、大型計算機の利用が前提になっていることから、現状においてもまだごく一部の先駆的な研究でしか用いられない。ゲノム解析で大型計算機の利用がバイオインフォマティクスで一般化し、また「京」や「富岳」などの比較的の使いやすい大型計算機が登場した以降もこの状況はあまり変化していない。そのため「京」登場以降から、これまで行われたような研究者が興味のある疾患データや薬剤応答データなど特定の条件下で観測された比較的小規模なデータから遺伝子ネットワークを推定し解析をするということではなく、解析対象をあらかじめ絞らずに公共データベースに登録されているデータセットに対して網羅的にネットワーク推定ソフトウェアを適用するというアプローチが登場している [15]。現時点においても「富岳」の豊富な計算機リソースを用いることで GEO に登録されているヒトのデータセットをほぼ網羅するおよそ 20,000 のデータセット・2,000,000 サンプルを網羅的にネットワーク推定に適用しデータベース化を行うプロジェクトが進行中である。これには 10,000,000 ノード時間の計算が必要であるが、「富岳」全計算リソースの 1% を恒常に利用しても 260 日、現実的な 0.5% のリソースを仮定すると 520 日の計算で完了することができる。また前述の通り、公共データベースに登録される遺伝子発現データのサンプル数は線形オーダーを超える伸びを示しているが、2032 年以降に想定されるヒト以外も含めた予想サンプル数は、年間 1,000,000 サンプル程度の増加で、恒常にネットワーク推定を継続することができれば、現在の計算リソースに対して数倍程度の計算リソースがあればこの増加分全てを網羅的に計算可能と言える。一方、最近になって少数のサンプルまたは個別のサンプルに対して遺伝子ネットワーク解析を行う技術が登場している [10, 11]。これは任意の遺伝子発現データをすでに推定した遺伝子ネットワークを用いて評価することで個別の遺伝子ネットワーク解析を実現しているものであるが、網羅的に推定した遺伝子ネットワークに対して別の遺伝子発現データを適用し解析可能であることを意味している。つまり、遺伝子ネットワーク数 × サンプル数分の組み合わせを考えることが可能で、解析可能な個別サンプル数は単純に組み合わせ爆発を起こす。さらには、データセット横断的に特定条件のサンプルを収集し新たなデータセットを再定義することも可能である。つまり現時点で 500 万サンプル、将来は 1,500 万サンプルのデータセットから任意のサンプルを抽出しそれに対して遺伝子ネットワーク推定を考えると、推定対象となり得るデータセットは事実上無限に存在することになる。現実的には、意味のあるデータセット再定義自動化にはサンプル単位でアノテーションされたデータベースが必要であるが、その試みはすでに行われている [12]。単純にこれらを用いて 10 疾患を 5 条件で再定義し直すとデータセット数はそれだけで 50 倍となる。前述の通り公共データセットの網羅的ネットワーク推定に必要とされる計算リソースを考慮すると、2032 年頃には最低でも現状の 100 倍程度の計算リソースが遺伝子ネットワーク解析に必要と言える。また、サンプルの量が増えるだけでなく解析手法も多様化することが予想されるため、それらを見越した計算機資源の整備が研究推進に絶対不可欠であると言える。

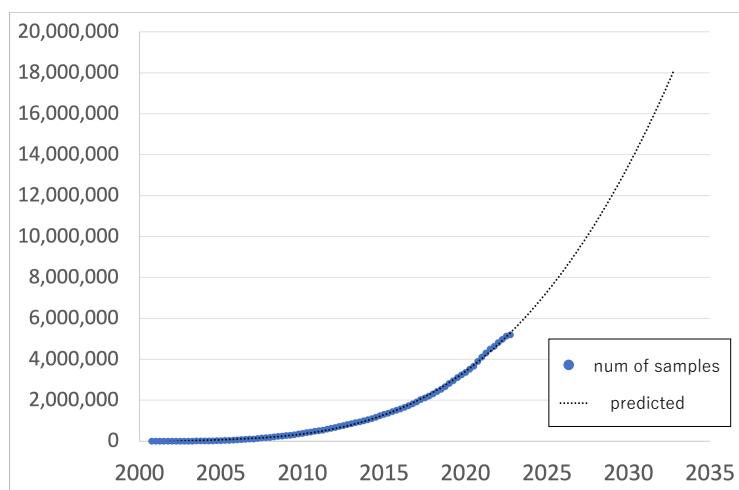


図 2.4 NCBI GEO に登録されている発現データサンプル数と今後の予測

社会的貢献 —健康で長寿な社会を目指して—

日本は世界に先んじて超高齢化社会を迎えており、国民の健康維持・増進による健康長寿社会の実現はきわめて重要な国家的課題である。健康長寿社会に資する創薬・医療技術の開発には、その基盤として人体における生命現象の理解と疾患発症メカニズムの解明が不可欠であるが、生命現象はあまりに多くの要素が絡み合って複雑に関係している現象であり、まさに、今後の HPC での計算能力の飛躍的増大が有効に活用される分野であると言える。

生物の遺伝情報の単位である遺伝子、および遺伝情報全体を意味するゲノムは、生命の設計図であるとよく言われる。21世紀に入ってすぐにヒトの全遺伝配列が決定されたが、それから20年以上経過し、遺伝子配列計測技術は飛躍的に進歩した。以前であれば数年かかった全ゲノム解読が、一人のゲノムについて数日程度で可能になっており、個人が自分自身のゲノム配列を知り、それに基づく医療（プレシジョン医療）を受けることができる個人ゲノム時代が既に到来している。このような超高速ゲノム解析を可能とする次世代 DNA シークエンサー（DNA を短く断片化し、並列に処理して高速に読み取る装置）では、膨大な観測データ（遺伝子配列データ）が日々算出されており、これら遺伝子配列の生物学的意味・臨床的意味を理解するためには莫大な計算が必要となる。例えば、がんゲノムにおいては、膨大な遺伝情報の中から、がんの種類に応じた、さらにはがん種横断的なゲノム上の特徴を見出す必要がある。更には、個々の遺伝子配列のみならず、異なる因子が複合的に関わる疾患では、複数の遺伝子がどのように関連しつつ働いているかを解明する遺伝子ネットワーク解析も、HPC の重要な応用分野となっている。

一方で、生体分子シミュレーションは、すでに新薬の設計やスクリーニングに広く用いられている。今後の HPC 分野においては、膨大な化合物対標的タンパク質の網羅的相互作用解析（2.6、4.6 で述べる）に加えて、複数のタンパク質から構成される細胞内における化合物の挙動をシミュレートする方法も創薬分野への貢献が期待できる。例えば細胞質を模した大規模シミュレーションでは、薬剤が実際に働く場である細胞中の分子混雑環境や競合的相互作用の影響が考慮されている。このような多対多の大規模シミュレーションを通じて、好ましくない副作用や、未知の相互作用を予測するための基盤技術を整備する事も

大きな社会貢献として目標にすべきことである。更に、近年では、ペプチド・抗体・核酸等に基づく医薬品（バイオ医薬品）の開発が創薬のニューモダリティーとして重視されていることから、低分子医薬のみならず中分子・高分子医薬を見据えたシミュレーション技術の開発も重要な課題である。

2.4.3 課題とその解決に必要な計算手法・アプリケーション

生体分子シミュレーションには、全原子モデルを用いた分子動力学（全原子MD）と粗視化モデルを用いた分子動力学（粗視化MD）がある。全原子MDを創薬に応用した創薬MD、酵素反応などを解析する量子化学/分子力学混合モデル（QM/MM）についても、2.6を参照してほしい。各シミュレーション法で方法論に小さな差異はあるが、経験的な分子間相互作用パラメータ（一般的に力場と呼ばれる。）を用いて、原子間に働く力を計算し、運動方程式を数値積分することで系の時間発展を得る点は共通している。

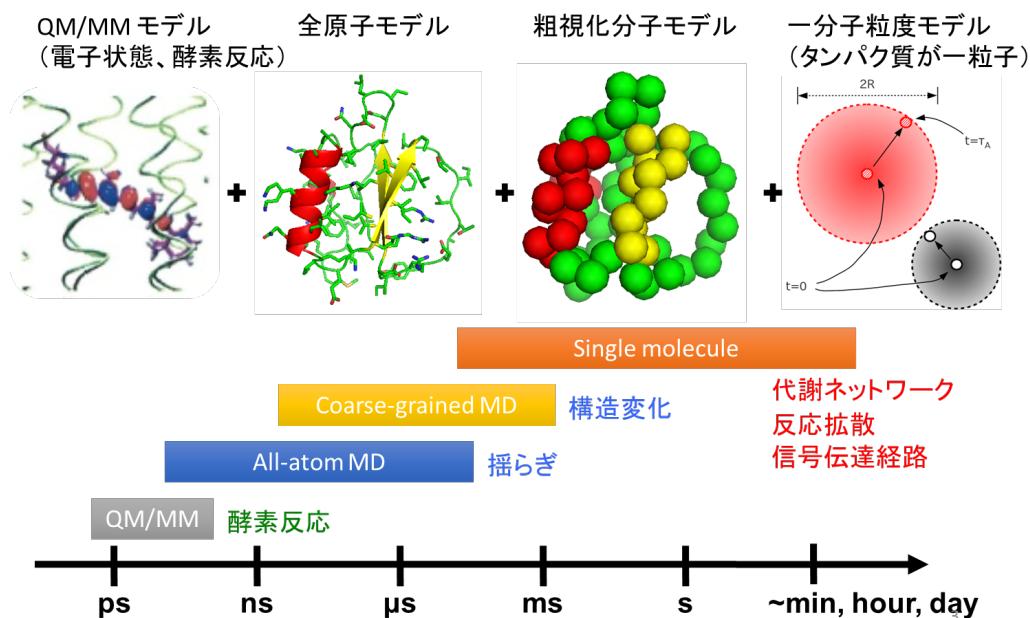


図 2.5 分子シミュレーションの異なるモデルと対象とする生命現象・時間スケール

(i) 全原子分子動力学法

全原子分子動力学法の代表的なアプリケーションは、既に世界的に普及しているものが多くある。代表的なものに、AMBER[16]（米）・CHARMM[17]（米）・Desmond[18]（米）・GROMACS[19]（蘭）・LAMMPS[20]（米）・NAMD[21]（米）・TINKER[22]（米）、TINKER-HP[23]（仏）、OpenMM[24]（米）国内では MARBLE[25]・MODYLAS[26]・GENESIS[27, 28]などがあり、非常に多岐にわたる。特に GENESIS では「富岳」や米国のスーパーコンピュータを用いて最大 10 億原子からなる細胞質[29] や染色体[30] のシミュレーションに成功している。

(ii) 拡張アンサンブル法

長時間の MD 計算だけを実行することで、タンパク質ダイナミクスを記述するために十分な構造空間を探索することは、MD 専用計算機等を十分に使うなど特殊な場合を除いて未だに困難である。この困難を克服するために、多くの効率の良いアルゴリズムが開発されている。拡張アンサンブル法はその一つであり、具体例として、レプリカ交換法、レプリカ交換アンブレラサンプリング法、レプリカ交換 Solute Tempering 法など多くの方法が提案されている。例えばレプリカ交換分子動力学 (REMD) 法 [31, 32] では、複数の MD 計算を並列に実行し、ある頻度で隣り合うレプリカ間で温度とエネルギーをメトロポリス判定で評価し、交換する。これにより温度空間の一次元酔歩を実現し、エネルギー極小状態に留まらない構造探索を実現するだけでなく、最重法によってカノニカル分布を再現し、広い温度範囲での物理量の予測を実現する。近年よく使われるようになってきた REST2 [33] や gREST[34] では、溶質と溶媒領域に分割することで大きな生体分子系のシミュレーションにも適用が可能になった。

(iii) マルコフ状態遷移モデル

マルコフ状態遷移モデル (MSM) によって、大量の短いトラジェクトリ群を統合することで、タンパク質の構造変化等の長時間ダイナミクス情報を得ることができる [35, 36]。MSM の入力であるトラジェクトリ群は、大量の短いシミュレーションから得られるため、並列に複数シミュレーションを実行できる計算環境に向いている。X 線結晶構造解析やクライオ電顕等によって解かれた複数状態の立体構造群、または、それらの状態間パスウェイ探索計算手法によって得た初期経路上の立体構造群を初期構造としてシミュレーションを実施する。特徴的なスナップショット構造を状態として定義し、トラジェクトリ中の一定時間間隔のスナップショットを状態に分類することで、状態間変化の遷移行列を求めることができる。状態間遷移行列から、各状態の出現頻度、状態間のつながり、及び、状態間遷移のカイネティクスの情報を得ることができる。例えば、米国では SARS-CoV-2 タンパク質について、Folding@home を使った大規模な並列シミュレーションを MSM によって統合し、構造変化の解析だけでなく、過渡的に現れるドラッガブルサイト探索が実施されている [37]。

(iv) 実験データとの連携

生命科学分野では古くから X 線結晶構造解析や NMR など様々な分光実験技術が生体内分子（タンパク質など）の構造解析・機能理解などに利用されてきた。近年、技術革新が進み、様々な実験手法が幅広く使われるようになってきている。クライオ電子顕微鏡 (Cryo-EM) により観測することのできる大きな高分子複合体の実験データから分子動力学シミュレーションを活用して構造の精密化を行う手法は GENESIS にも実装されデータの解析に使われるようになっている [38, 39]。高速 AFM による分子運動のリアルタイムでの観測画像からの構造モデリングの手法も開発されつつあり [40, 41]、今後シミュレーションを活用してより大きなタンパク質複合体について高解像度のモデリングを可能にする必要がある。また SACL A を用いた時分割構造解析も技術の開発が進みより多くの反応中間体構造が観測できるようになると期待され、これらを活用したシミュレーションによる機能解析はタンパク質の機能発現に関わる新しい情報を与えると期待される。また、Cryo-EM や SACL A データの解析で得られる構造情報よりも解像度は低くなるが、より分子の構造ダイナミクスに注目した解析として 1 分子計測時系列データと分子

動力学シミュレーションの融合(データ同化)が挙げられる。FRET や高速 AFM データ解析で既に応用されており、構造変化における中間構造の同定に成功している。時系列データの融合には、実験データとシミュレーションデータの時間スケールのギャップを埋めるために、MSM が用いられる。MSM の構築には、分子が取りうる構造をシミュレーションで網羅的にサンプルする必要があるが、このためにたくさんの独立なシミュレーションを流すアプローチが取られている。大規模サンプリング手法から得たタンパク質の構造変化モデルは、実験データとの比較解析によって妥当性が検証できる。小角 X 線散乱(SAXS) は、タンパク質溶液へ X 線を照射し、その散乱曲線データから溶液中のタンパク質概形を計測する手法である。原理的に低解像度ではあるが、クリスタルパッキングの影響のない、溶液中のペアなタンパク質の溶液構造概形を捉えることができる。MD-SAXS 法は、分子シミュレーションから得た高分解能な立体構造(アンサンブル)モデルから理論散乱曲線を求め、実験散乱曲線と比較解析を実施することで、実験中の溶液構造モデリングをする手法である[42]。SAXS 実験データとの比較によって、分子シミュレーションのサンプリング範囲についても、妥当性が検証できる。シミュレーションを使ったアンサンブル構造モデル探索手法として、立体構造の出現頻度を、実験 SAXS データから最適化する方法も複数提案されている。また、シミュレーションによるサンプリング過程に、SAXS 実験データをガイドとして組み込む方法も開発されている。

(v) 粗視化分子動力学法

粗視化分子動力学法では、全原子シミュレーションでは再現することが難しいような大規模な構造変化を計算することができる。代表的なものは Marrink らが開発する MARTINI 力場[43, 44](およびそれを用いたシミュレーションプログラム群など)であり、最初の生体膜モデルから拡張されてタンパク質系も取り込んで発展している。日本では CafeMol[45] が開発され、タンパク質や核酸の構造変化の研究に広く用いられている。粗視化モデルではアミノ酸 1 残基を 1-2 個程度の粒子で近似する場合が多く、全原子モデルに比べて粒子数が 10 分の 1 程度に減り、水分子もあらわに扱わないので粒子数が少數であるため、計算量をかなり減らすことができる。その一方で、全原子 MD で一般的な空間分割スキームを用いると粒子数の数の差が大きく、並列性能は得られにくい。粗視化された粒子間に働く相互作用はさまざまなものがあり、相互作用毎に異なるチューニングが必要になってくる。

(vi) AI・機械学習

生命科学データは増加しているものの、分子シミュレーションの計算対象と実験データには依然として乖離がある。実験データに対応するような生体分子の構造やダイナミクスのサンプリングは現状では非常にコストがかかり、AI・機械学習による効率的な計算手法の開発が望まれる。また、シミュレーション結果と実験データを、AI・機械学習で融合して扱うことで実験データの欠損を補いつつ、シミュレーションエラーの補正を実施する手法開発も有望である。コストの高い計算手法で計算したデータを学習データとして、よりコストの低い近似した枠組みで高精度な分子シミュレーションを行う手法(いわゆる AI 力場)は、学習データに依存する適用範囲の問題を解決することで今後より一般化していくものと期待される。また、創薬や新しい機能を持った生体分子の設計においては、シミュレーションを実験とみなし、ベイズ最適化やモンテカルロ木探索等の AI・機械学習手法と組み合わせることで計算に基づく分子のデザインが可能になる。今後、より実用的・高速な分子デザインのために新たな生成モデル・探索手法とシミュ

レーションとの高度な連携が必要になると考えられる。

(vii) 一分子粒度モデル

細胞スケールのシミュレーションの多くは、細胞内の分子が均一に分布していおることを仮定し、細胞内で起きる化学反応を化学マスター方程式（分子が衝突し、反応を起こす確率が均一であるとする）などで記述している。一方で、分子1つ1つの運動を考慮した細胞スケールシミュレータの開発も進められている。日本においては Spatiocyte [46] が代表的なアプリケーションとして使用されている。Spatiocyte は、タンパク質などの生体高分子を一つの粒子で記述（1分子粒度モデル）し、格子状に離散化された空間の中でモンテカルロ法によって運動を計算する。分子同士の反応は、Collins-Kimball 法および Gillespie 法に基づいて計算される。Spatiocyte は「京」などに向け、高並列化されたバージョン（pSpatiocyte）が開発され、HPCI 戦略プログラムのもと、細胞内シグナル伝達やリン酸化反応に及ぼす細胞環境の影響などが調査されている。細胞環境（特に分子混雑）の影響は全原子 MD 計算によっても活発に調査されており、近い将来に原子粒度の生体分子シミュレーションと pSpatiocyte などの細胞シミュレータが連携し、分子から細胞スケールまでの時空間階層を接続するマルチスケール・マルチフィジックスシミュレーションが実現する可能性が高いことを示唆している。1分子粒度モデル計算のアルゴリズムの詳細や、今後の課題、および要求性能などは、平成26年度版計算科学ロードマップを参照してほしい。

(viii) ゲノム解析

次世代シークエンサーから得られる大量の断片的配列データ（リード）の解析には、リードの標準 DNA 配列（リファレンス）に対する位置同定（マッピング）、ソート、PCR 増幅による重複リードの除去等の工程による配列情報再構成工程と、得られたデータから対象とする遺伝的多様性や遺伝子配列の変化（点変異、挿入・欠失、コピー数変化、遺伝子融合、構造変異）に焦点を合わせた候補部位抽出と統計的判定（シークエンス読み取りエラー等の偽陽性除去）が必要である。このような解析を大規模計算機を用いて高並列化したアプリケーションの代表的なものとして Genomon[47, 48] がある。次世代シークエンサーにより得られたリード配列をリファレンスとなるゲノム配列に対して“貼りつけ”（マッピング）るには、Burrows-Wheeler Transform (BWT) アルゴリズムが用いられる。

(ix) 遺伝子ネットワーク解析

DNA チップ（マイクロアレイ）や次世代シークエンサーなどで得られる遺伝子発現データ、タンパク発現（プロテオーム）データ、miRNA を含む noncoding RNA 発現データなどを用いて遺伝子間の関係（転写因子—被制御遺伝子の相互作用など）を明らかにする事を遺伝子ネットワーク解析という。広義には KEGG に代表される実験で得られた知識を集積したいわゆるパスウェイ・データベースに基づいたネットワーク解析も遺伝子ネットワーク解析に含めるがここではあくまで事前知識を仮定せずデータからのみに基づく、つまりデータ駆動型の遺伝子ネットワーク解析を前提とする。遺伝子ネットワーク解析では、薬剤や外的刺激を与えた後の遺伝子転写量の時系列データ（例：数時間おきに発現データを取る）や患者由来腫瘍サンプルの遺伝子発現データなどから遺伝子発現の因果関係（遺伝子間の相互作用）を明らかにし、それをネットワークとして推定する。SiGN-BN[49] は B-spline ノンパラメトリック回帰を用いたベイジアンネットワークによる遺伝子ネットワーク推定ソフトウェアで、元々「次世代生命体統合シ

「ミュレーションの研究開発」プロジェクトで研究開発されたソフトウェアの1つで「京」コンピュータなどの超並列型スーパーコンピュータの演算能力を最大限活用できるよう開発された遺伝子ネットワーク推定ソフトウェアである。現在でも開発・メンテナンスが続けられており「富岳」でも利用されている実績のある遺伝子ネットワーク推定ソフトウェアである。現在はその発展版である INGOR と呼ばれるベイジアンネットワーク推定ソフトウェアが開発されており、個別サンプル遺伝子ネットワーク解析手法が実装されている [10, 11]。SiGN-BN/INGOR にはネットワークサイズに応じた複数の構造推定アルゴリズムが実装されているが、ヒト全遺伝子に対応した Neighbor Node Sampling & Repeat アルゴリズム [50] が主として利用されている。

2.4.4 ロードマップ

ここでは、主に生体分子シミュレーション、遺伝子ネットワーク解析、ゲノム解析について 2032 年までのロードマップを表にした。

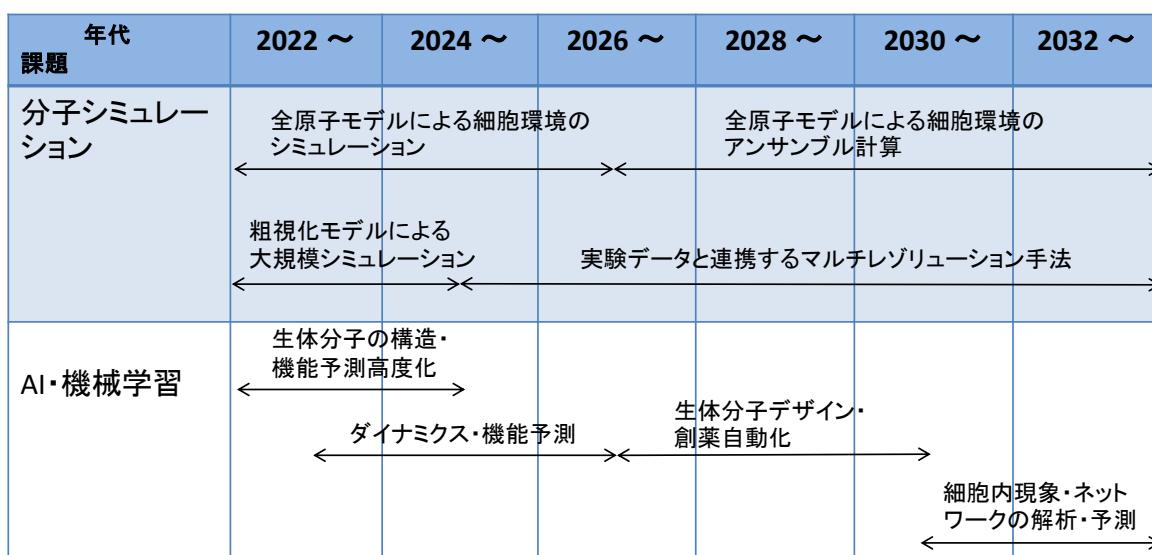


図 2.6 生命科学分野ロードマップ

2.4.5 必要な計算機資源

本分野は構成要素が非常に多いため、ここでは大まかな要求計算機資源を表にまとめた。要求資源の根拠や詳細については 4.4 を参照してほしい。

- 巨大分子系の分子動力学計算については、複数のタンパク質を含む多成分系の超大規模計算へ拡張し、10 億原子系の 1 マイクロ秒の MD シミュレーションを 10 日程度で完了することを目標にした。(詳細な見積もりは、4.4 を参照)。
- 個人ゲノム解読を行う次世代シークエンサー解析については、2020 年以降に行われると思われる 200,000 人規模の解析を想定した。

3. 遺伝子ネットワーク解析については、2032年には現在の100倍程度のネットワーク推定が量のデータが公開データベースにて入手可能になると予測され、そのデータ量を想定して解析を行う場合を想定した。
4. XFELの散乱データからのイメージングについては、 $10^7 \times 10^5$ イメージから構造の分類、時間軸の推定などを行い4次元イメージング（立体構造の時間変化）を実現する事を想定とした。

2032年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13節の表2.1に記載した。また、計算機資源の見積りの詳細は第4.4節に記載している。

参考文献

- [1] Tetsu Narumi, Yousuke Ohno, Noriaki Okimoto, Takahiro Koishi, Atsushi Suenaga, Noriyuki Futatsugi, Ryoko Yanai, Ryutaro Himeno, Shigenori Fujikawa, Makoto Taiji, and Mitsuru Ikei. A 55 tflops simulation of amyloid-forming peptides from yeast prion sup35 with the special-purpose computer system mdgrape-3, 2006.
- [2] David E. Shaw, Martin M. Deneroff, Ron O. Dror, Jeffrey S. Kuskin, Richard H. Larson, John K. Salmon, Cliff Young, Brannon Batson, Kevin J. Bowers, Jack C. Chao, Michael P. Eastwood, Joseph Gagliardo, J. P. Grossman, C. Richard Ho, Douglas J. Ierardi, István Kolossváry, John L. Klepeis, Timothy Layman, Christine McLeavey, Mark A. Moraes, Rolf Mueller, Edward C. Priest, Yibing Shan, Jochen Spengler, Michael Theobald, Brian Towles, and Stanley C. Wang. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM*, Vol. 51, No. 7, p. 91–97, 2008.
- [3] D. E. Shaw, J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L. S. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y. H. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, and C. Young. Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 41–53.
- [4] David E. Shaw, Peter J. Adams, Asaph Azaria, Joseph A. Bank, Brannon Batson, Alistair Bell, Michael Bergdorf, Jhanvi Bhatt, J. Adam Butts, Timothy Correia, Robert M. Dirks, Ron O. Dror, Michael P. Eastwood, Bruce Edwards, Amos Even, Peter Feldmann, Michael Fenn, Christopher H. Fenton, Anthony Forte, Joseph Gagliardo, Gennette Gill, Maria Gorlatova, Brian Greskamp, J. P. Grossman, Justin Gullingsrud, Anissa Harper, William Hasenplaugh, Mark Heily, Benjamin Colin Heshmat, Jeremy Hunt, Douglas J. Ierardi, Lev Iserovich, Bryan L. Jackson, Nick P. Johnson, Mollie M. Kirk, John L. Klepeis, Jeffrey S. Kuskin, Ken-

- neth M. Mackenzie, Roy J. Mader, Richard McGowen, Adam McLaughlin, Mark A. Moraes, Mohamed H. Nasr, Lawrence J. Nociolo, Lief O'Donnell, Andrew Parker, Jon L. Peticolas, Goran Pocina, Cristian Predescu, Terry Quan, John K. Salmon, Carl Schwink, Keun Sup Shim, Naseer Siddique, Jochen Spengler, Tamas Szalay, Raymond Tabladillo, Reinhard Tartler, Andrew G. Taube, Michael Theobald, Brian Towles, William Vick, Stanley C. Wang, Michael Wazlowski, Madeleine J. Weingarten, John M. Williams, and Kevin A. Yuh. Anton 3: twenty microseconds of molecular dynamics simulation before lunch, 2021.
- [5] Gentaro Morimoto, Yohei M. Koyama, Hao Zhang, Teruhisa S. Komatsu, Yousuke Ohno, Keigo Nishida, Itta Ohmura, Hiroshi Koyama, and Makoto Taiji. Hardware acceleration of tensor-structured multilevel ewald summation method on mdgrape-4a, a special-purpose computer system for molecular dynamics simulations, 2021.
- [6] Haiqing Xiong, Yingjie Luo, Qianhao Wang, Xianhong Yu, and Aibin He. Single-cell joint detection of chromatin occupancy and transcriptome enables higher-dimensional epigenomic reconstructions. *Nat Methods*, Vol. 18, pp. 652–660, 2021.
- [7] Benjamin Izar, Itay Tirosh, Elizabeth H Stover, Isaac Wakiro, Michael S Cuoco, Idan Alter, Christopher Rodman Rachel Leeson, Mei-Ju Su, Parin Shah, Marcin Iwanicki, Sarah R Walker, Abhay Kanodia, Johannes C Melms, Shaolin Mei, Jia-Ren Lin, Caroline B M Porter, Michal Slyper, Julia Waldman, Livnat Jerby-Arnon, Orr Ashenberg, Titus J Brinker, Caitlin Mills, Meri Rogava, Sébastien Vigneau, Peter K Sorger, Levi A Garraway, Panagiotis A Konstantinopoulos, Joyce F Liu, Ursula Matulonis, Bruce E Johnson, Orit Rozenblatt-Rosen, Asaf Rotem, and Aviv Regev. A single-cell landscape of high-grade serous ovarian cancer. *Nat Methods*, Vol. 26, No. 8, pp. 1271–1279, 2020.
- [8] N. A. Miller and et al. A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. *Genome Med.*, Vol. 7, p. 100, 2015.
- [9] Sam Friedman, Laura Gauthier, Yossi Farjoun, and Eric Banks. Lean and deep models for more accurate filtering of snp and indel variant calls. *Bioinformatics*, Vol. 36, No. 7, pp. 2060–2067, 2020.
- [10] Yoshihisa Tanaka, Yoshinori Tamada, Marie Ikeguchi, Fumiyo Yamashita, and Yasushi Okuno. System-based differential gene network analysis for characterizing a sample-specific subnetwork. *Biomolecules*, Vol. 10, No. 2, p. 306, 2020.
- [11] Yoshihisa Tanaka, Kako Higashihara, Mai Adachi Nakazawa, Fumiyo Yamashita and Yoshinori Tamada, and Yasushi Okuno. Dynamic changes in gene-to-gene regulatory networks in response to sars-cov-2 infection. *Sci Rep*, Vol. 11, No. 1, p. 11241, 2021.
- [12] Wang Z, Monteiro CD, Jagodnik KM, Fernandez NF, Gundersen GW, Rouillard AD, Jenkins SL, Feldmann AS, Hu KS, McDermott MG, Duan Q, Clark NR, Jones MR, Kou Y, Goff T, Woodland H, Amaral FMR, Szeto GL, Fuchs O, Schüssler-Fiorenza Rose SM, Sharma S, Schwartz U, Bausela XB, Szymkiewicz M, Maroulis V, Salykin A, Barra CM, Kruth CD, Bongio NJ, Mathur V, Todorovic RD, Rubin UE, Malatras A, Fulp CT, Galindo JA, Motiejunaite R,

- Jüschke C, Dishuck PC, Lahl K, Jafari M, Aibar S, Zaravinos A, Steenhuizen LH, Allison LR, Gamallo P, de Andres Segura F, Dae Devlin T, Pérez-García V, and Ma'ayan A. Extraction and analysis of signatures from the gene expression omnibus by the crowd. *Nat Commun*, Vol. 7, p. 12846, 2016.
- [13] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with alphafold. *Nature*, Vol. 596, No. 7873, pp. 583–589, 2021.
- [14] Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata Laydon, Augustin Žídek, Tim Green, Kathryn Tunyasuvunakool, Stig Petersen, John Jumper, Ellen Clancy, Richard Green, Ankur Vora, Mira Lutfi, Michael Figurnov, Andrew Cowie, Nicole Hobbs, Pushmeet Kohli, Gerard Kleywegt, Ewan Birney, Demis Hassabis, and Sameer Velankar. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, Vol. 50, No. D1, pp. D439–D444, 2021.
- [15] Rika Kasajima, Rui Yamaguchi, Eigo Shimizu, Yoshinori Tamada, Atsushi Niida, George Tremmel, Takeshi Kishida, Ichiro Aoki, Seiya Imoto, Satoru Miyano, Hiroji Uemura, and Yohei Miyagi. Variant analysis of prostate cancer in Japanese patients and a new attempt to predict related biological pathways. *Oncol Rep*, Vol. 43, No. 3, pp. 943–952, 2020.
- [16] D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, G.A. Cisneros III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, A. Rahnamoun, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman. Amber 2022, 2022.
- [17] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. Charmm: The biomolecular simulation

- program. *Journal of Computational Chemistry*, Vol. 30, No. 10, pp. 1545–1614, 2009.
- [18] K. J. Bowers, D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, and D. E. Shaw. Scalable algorithms for molecular dynamics simulations on commodity clusters. In *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, pp. 43–43, 2006.
- [19] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Pál, Jeremy C. Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, Vol. 1-2, pp. 19–25, 2015.
- [20] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. Lammps - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, Vol. 271, p. 108171, 2022.
- [21] James C. Phillips, David J. Hardy, Julio D. C. Maia, John E. Stone, João V. Ribeiro, Rafael C. Bernardi, Ronak Buch, Giacomo Fiorin, Jérôme Hénin, Wei Jiang, Ryan McGreevy, Marcelo C. R. Melo, Brian K. Radak, Robert D. Skeel, Abhishek Singharoy, Yi Wang, Benoît Roux, Aleksei Aksimentiev, Zaida Luthey-Schulten, Laxmikant V. Kalé, Klaus Schulten, Christophe Chipot, and Emad Tajkhорشید. Scalable molecular dynamics on cpu and gpu architectures with namd. *Journal of Chemical Physics*, Vol. 153, No. 4, p. 044130, 2020.
- [22] Joshua A. Rackers, Zhi Wang, Chao Lu, Marie L. Laury, Louis Lagardère, Michael J. Schnieders, Jean-Philip Piquemal, Pengyu Ren, and Jay W. Ponder. Tinker 8: Software tools for molecular design. *Journal of Chemical Theory and Computation*, Vol. 14, No. 10, pp. 5273–5289, 2018. doi: 10.1021/acs.jctc.8b00529.
- [23] Louis Lagardere, Luc-Henri Jolly, Filippo Lipparini, Félix Aviat, Benjamin Stamm, Zhifeng F. Jing, Matthew Harger, Hedieh Torabifard, G. Andrés Cisneros, Michael J. Schnieders, Nohad Gresh, Yvon Maday, Pengyu Y. Ren, Jay W. Ponder, and Jean-Philip Piquemal. Tinker-hp: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chemical Science*, Vol. 9, No. 4, pp. 956–972, 2018.
- [24] Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee-Ping Wang, Andrew C. Simmonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, and Vijay S. Pande. Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology*, Vol. 13, No. 7, p. e1005659, 2017.
- [25] Mitsunori Ikeguchi. Partial rigid-body dynamics in npt, npat and np γ t ensembles for proteins and membranes. *Journal of Computational Chemistry*, Vol. 25, No. 4, pp. 529–541, 2004.
- [26] Yoshimichi Andoh, Noriyuki Yoshii, Kazushi Fujimoto, Keisuke Mizutani, Hidekazu Kojima, Atsushi Yamada, Susumu Okazaki, Kazutomo Kawaguchi, Hidemi Nagao, Kensuke Iwahashi,

- Fumiyasu Mizutani, Kazuo Minami, Shin-ichi Ichikawa, Hidemi Komatsu, Shigeru Ishizuki, Yasuhiro Takeda, and Masao Fukushima. Modylas: A highly parallelized general-purpose molecular dynamics simulation program for large-scale systems with long-range forces calculated by fast multipole method (fmm) and highly scalable fine-grained new parallel processing algorithms. *Journal of Chemical Theory and Computation*, Vol. 9, No. 7, pp. 3201–3209, 2013. doi: 10.1021/ct400203a.
- [27] J. Jung, T. Mori, C. Kobayashi, Y. Matsunaga, T. Yoda, M. Feig, and Y. Sugita. Genesis: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, Vol. 5, No. 4, pp. 310–323, 2015.
- [28] C. Kobayashi, J. Jung, Y. Matsunaga, T. Mori, T. Ando, K. Tamura, M. Kamiya, and Y. Sugita. Genesis 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms. *Journal of Computational Chemistry*, Vol. 38, No. 25, pp. 2193–2206, 2017. Fd6cl Times Cited:6 Cited References Count:117.
- [29] Jaewoon Jung, Chigusa Kobayashi, Kento Kasahara, Cheng Tan, Akiyoshi Kuroda, Kazuo Minami, Shigeru Ishiduki, Tatsuo Nishiki, Hikaru Inoue, Yutaka Ishikawa, Michael Feig, and Yuji Sugita. New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems. *Journal of Computational Chemistry*, Vol. 42, No. 4, pp. 231–241, 2021.
- [30] J. Jung, W. Nishima, M. Daniels, G. Bascom, C. Kobayashi, A. Adedoyin, M. Wall, A. Lappala, D. Phillips, W. Fischer, C. S. Tung, T. Schlick, Y. Sugita, and K. Y. Sanbonmatsu. Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations. *Journal of Computational Chemistry*, Vol. 40, No. 21, pp. 1919–1930, 2019. Id6pn Times Cited:3 Cited References Count:34.
- [31] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, Vol. 314, No. 1, pp. 141–151, 1999.
- [32] Yuji Sugita, Akio Kitao, and Yuko Okamoto. Multidimensional replica-exchange method for free-energy calculations. *Journal of Chemical Physics*, Vol. 113, No. 15, pp. 6042–6051, 2000.
- [33] Tsuyoshi Terakawa, Tomoshi Kameda, and Shoji Takada. On easy implementation of a variant of the replica exchange with solute tempering in gromacs. *Journal of Computational Chemistry*, Vol. 32, No. 7, pp. 1228–1234, 2011.
- [34] M. Kamiya and Y. Sugita. Flexible selection of the solute region in replica exchange with solute tempering: Application to protein-folding simulations. *Journal of Chemical Physics*, Vol. 149, No. 7, p. 072304, 2018. 1089-7690 Kamiya, Motoshi Sugita, Yuji Journal Article United States 2018/08/24 J Chem Phys. 2018 Aug 21;149(7):072304. doi: 10.1063/1.5016222.
- [35] Brooke E. Husic and Vijay S. Pande. Markov state models: From an art to a science. *Journal of the American Chemical Society*, Vol. 140, No. 7, pp. 2386–2396, 2018. doi: 10.1021/jacs.7b12191.
- [36] Xiaowei Wang, Ilona Christy Unarta, Peter Pak-Hang Cheung, and Xuhui Huang. Elucidating molecular mechanisms of functional conformational changes of proteins via markov state models.

- Current Opinion in Structural Biology*, Vol. 67, pp. 69–77, 2021.
- [37] Maxwell I. Zimmerman, Justin R. Porter, Michael D. Ward, Sukrit Singh, Neha Vithani, Arthur Meller, Upasana L. Mallimadugula, Catherine E. Kuhn, Jonathan H. Borowsky, Rafal P. Wiewiora, Matthew F. D. Hurley, Aoife M. Harbison, Carl A. Fogarty, Joseph E. Coffland, Elisa Fadda, Vincent A. Voelz, John D. Chodera, and Gregory R. Bowman. Sars-cov-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. *Nature Chemistry*, Vol. 13, No. 7, pp. 651–659, 2021.
- [38] Osamu Miyashita, Chigusa Kobayashi, Takaharu Mori, Yuji Sugita, and Florence Tama. Flexible fitting to cryo-em density map using ensemble molecular dynamics simulations. *Journal of Computational Chemistry*, Vol. 38, No. 16, pp. 1447–1461, 2017.
- [39] Takaharu Mori, Marta Kulik, Osamu Miyashita, Jaewoon Jung, Florence Tama, and Yuji Sugita. Acceleration of cryo-em flexible fitting for large biomolecular systems by efficient space partitioning. *Structure*, Vol. 27, No. 1, pp. 161–174.e3, 2019.
- [40] B. Dasgupta, O. Miyashita, T. Uchihashi, and F. Tama. Reconstruction of three-dimensional conformations of bacterial clpb from high-speed atomic-force-microscopy images. *Front Mol Biosci*, Vol. 8, p. 704274, 2021. Dasgupta, Bhaskar Miyashita, Osamu Uchihashi, Takayuki Tama, Florence eng Switzerland 2021/08/24 Front Mol Biosci. 2021 Aug 5;8:704274. doi: 10.3389/fmlob.2021.704274. eCollection 2021.
- [41] S. Fuchigami and S. Takada. Inferring conformational state of myosin motor in an atomic force microscopy image via flexible fitting molecular simulations. *Front Mol Biosci*, Vol. 9, p. 882989, 2022. Fuchigami, Sotaro Takada, Shoji eng Switzerland 2022/05/17 Front Mol Biosci. 2022 Apr 29;9:882989. doi: 10.3389/fmlob.2022.882989. eCollection 2022.
- [42] Toru Ekimoto and Mitsunori Ikeguchi. *Hybrid Methods for Modeling Protein Structures Using Molecular Dynamics Simulations and Small-Angle X-Ray Scattering Data*, pp. 237–258. Springer Singapore, Singapore, 2018.
- [43] Siewert J. Marrink, Alex H. de Vries, and Alan E. Mark. Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, Vol. 108, No. 2, pp. 750–760, 2004. doi: 10.1021/jp036508g.
- [44] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The martini force field: Coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, Vol. 111, No. 27, pp. 7812–7824, 2007. doi: 10.1021/jp071097f.
- [45] Hiroo Kenzaki, Nobuyasu Koga, Naoto Hori, Ryo Kanada, Wenfei Li, Kei-ichi Okazaki, Xin-Qiu Yao, and Shoji Takada. Cafemol: A coarse-grained biomolecular simulator for simulating proteins at work. *Journal of Chemical Theory and Computation*, Vol. 7, No. 6, pp. 1979–1989, 2011. doi: 10.1021/ct2001045.
- [46] Spatiocyte. <http://spatiocyte.org/>.
- [47] <https://genomon.readthedocs.io/ja/latest/>.
- [48] 鈴木惣一朗, 伊東聰, 池田奈生, Balazs Gerofti, 三吉郁夫, 丸山直也, 滝澤真一朗, 村瀬洋介, 石川裕,

- 宮野悟. ヒトゲノム解析プログラム genomon-fusion の大規模スーパーコンピュータ向け最適化と性能モデル化, tuning and performance modeling of a human gene analysis program for large-scale supercomputers. ハイパフォーマンスコンピューティングと計算科学シンポジウム論文集, Vol. 2016, pp. 15–26, 2016.
- [49] Y. Tamada, T. Shimamura, R. Yamaguchi, S. Imoto, M. Nagasaki, , and S. Miyano. Sign: Large-scale gene network estimation environment for high performance computing. *Genome Informatics*, Vol. 25, No. 1, pp. 40–52, 2011.
- [50] Y. Tamada, S. Imoto, H. Araki, M. Nagasaki, C. Print, D.S. Charnock-Jones, , and S. Miyano. Estimating genome-wide gene networks using nonparametric bayesian network models on massively parallel computers. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 8, No. 3, pp. 683–697, 2011.
- [51] NGSAnalyzer. <http://www.csrp.riken.jp/islim/ngsanalyzer-j.html>.
- [52] SiGN. http://sign.hgc.jp/index_ja.html.

2.5 脳・神経科学

2.5.1 分野の概要

脳は呼吸・心拍のような生命維持から意識・思考といったいわゆる高次認知機能までを司る重要な器官であるが、そのためのシステムとしての脳の情報処理機構は未だに解明されていない。一方、脳の構成要素である神経細胞（ニューロン）の電気化学的応答は非常に良く研究されており、単一ニューロンの挙動を具体的に数式として記述する様々な数理モデルが提案されている。このような数式をニューロンの数だけ列挙し求解することで、脳活動を計算機上に再現し、情報処理機構や脳疾患を調べることが日々行われている。

脳科学、特に基礎神経科学においては、脳の微細な構造と機能に着目し、ボトムアップ的に脳機能の解明を目指す取り組みと、ニューロンのネットワークを抽象化・数理化したモデルで、トップダウン的に脳機能の解明を目指す取り組みが並行して行われている [1]。神経科学の知見を活かした工学応用である深層学習においては、課題によっては既にヒトの能力を凌駕する性能を達成している。この深層学習の電力消費の壁を超える性能向上を目指し、神経科学の知見をより取り入れたニューロモーフィックデバイスの開発が盛んに行われている。

過去 60 年間、脳科学と工学的応用である人工ニューラルネットは互いに影響を与え合いながら取り組まれている。近年、脳科学は各国で大型プロジェクトが推進されており [2]、深層学習はその飛躍的な発展によって、アカデミア・インダストリーを問わず驚異的な勢いで分野が拡大している [3]。それぞれの分野で HPC の利活用はますます重要視されつつあるが、分野毎に必要とされる計算法・計算性能・計算機アーキテクチャは異なる。本章では脳科学分野における代表的なアプリケーションである神経回路シミュレーションと、工学応用分野のニューロモーフィックデバイスについてそれぞれ検討する。

2.5.1.1 脳・神経回路シミュレーション

脳は、認知、思考、感情など、情報処理をつかさどる器官で、脳のニューロン同士がシナプスと呼ばれる構造を介して互いに結合し、主に活動電位と呼ばれる電気パルス状の信号で表象される信号を交換することで情報処理を行っている。動物の脳は種によって回路構造や規模は様々あり、例えば、無脊椎動物の線虫の脳は 302 個のニューロンと約 7000 本の結合からなるのに対して、人間の脳は約 860 億個のニューロンと 100 兆本以上の結合を持つ最大級の脳で、この世で最も複雑な構造物と言われている。神経科学は、脳の構成要素であるニューロンやシナプス結合の挙動、行動と神経活動の関係、脳機能マップ、脳結合マップなど、様々なことを明らかにしてきた。しかし、「全脳」の神経回路が協調動作して行う情報処理の機構については、線虫の脳ですらまだ解明されていない。ヒトの脳は神経細胞数でいえば、線虫の 1 億倍以上の規模の神経回路により、思考や言語、複雑な運動といった「高次脳機能」が実現しているが、その膨大な複雑さのため、全容解明はさらに遠く、現在も精力的に研究が行われている。

脳研究は学際領域であり様々な手法によるアプローチが行われている。計算機上に神経回路のモデルを構築し、数値シミュレーションによってその挙動を調べるアプローチには、大別すると 2 つの種類があ

る。1つは「計算論的神経科学」におけるトップダウンのアプローチであり、「脳は何をしているか?」という機能に関する仮説（計算論）を立て、「それはどのように実現されているか?」というアルゴリズムとデータ構造に関する考察を行い、「神経回路としてどのように実装するか?」という問題意識で、仮説を支持するコンパクトなモデルを構築し数値シミュレーションを行うものである。もう1つは「シミュレーション神経科学」というボトムアップ側からのアプローチで、解剖学・生理学データから実際の神経回路を忠実に構築・再現して数値シミュレーションを行うことにより主眼を置くものである。このボトムアップのアプローチはもともと計算論的神経科学の一部として行われていた。しかし、計算・計測可能な神経回路規模の急激な上昇とともに、観測される神経系の再現と予測を神経細胞レベルで対応させて行う機運が高まり、このボトムアップのアプローチをシミュレーション神経科学と分類する動きが近年起きており [4]。計算論的神経科学における仮説ベースのアプローチとシミュレーション神経科学の神経回路シミュレーションのボトムアップのアプローチによる双方向的な取り組みが求められる。本稿では、HPCで特に恩恵を受け発展するシミュレーション神経科学の神経回路シミュレーションに焦点を当て、検討を行う。シミュレーション神経科学の神経回路シミュレーションは、ニューロンとシナプスの挙動を微分方程式で記述し、ニューロン個数分・シナプス個数分の式を時間ステップ毎に解くことで、神経回路全体の挙動を数値的にシミュレートする。神経回路シミュレーションの始まりは、1952年の Hodgkin と Huxley によるイカの巨大軸索の活動電位のシミュレーションで、はじめは手回し計算機で数値計算が実行された [5]。それから 60 年以上の間、神経回路シミュレーションは発展し続けている。最近 15 年間の代表的な例として、100 万神経細胞からなる皮質一視床 [6]、16 億神経細胞からなる皮質一視床 [7]、17 億神経細胞からなる局所大脳皮質回路 [8]、いくつかの認知タスクを実行可能な神経回路 [9]、細胞形態やイオンチャネルを考慮する詳細なラット一次体性感覚皮質 [10] や一次視覚野のモデル [11]、ヒト 1/3 規模の大脳皮質 [12] やヒト規模小脳 [13]、エクサフロップス級スパコン向けシミュレータ開発 [14] など、様々な大規模神経回路シミュレーションに関わる報告がなされている。このような大規模モデルの実行には、大きいものではエクサフロップス以上の性能を持つ大規模並列型の電子計算機によって実行されている。2022 年現在は、積分発火型の神経細胞モデルを用いたヒト規模の大脳皮質、視床、小脳からなる神経回路モデルが実現している [15]。しかし、ニューロンの種類別の多様な発火応答特性や単一ニューロンの空間形状を利用した情報処理等はまだ取り入れられておらず、これらを考慮したヒト全脳シミュレーション、及び実時間実行を実現するには、後述するように、ゼタフロップス級の計算性能が必要と推定される。脳の大規模モデル構築には、脳の神経細胞間の結合や、神経細胞の電気生理学的な特徴について、全脳規模の膨大なデータが必要である。近年、世界中で行われている脳の大型プロジェクトは、全脳規模での結合計測、神経活動計測、脳データベースの拡充、大規模計算モデルの構築などについて取り組んでおり、大規模な脳のモデル研究をさらに後押ししている。ここでは代表的なものを紹介する。欧州の Human Brain Project(2013-2023) は、脳の情報処理と疾患の機構の解明・人工脳の開発、ヒト全脳シミュレーションを目指した取り組みを行っている [?]。米国の Brain Initiative (2016-2025) は線虫や昆虫からヒトまでの多くのモデル動物をスケーラブルに扱う上で全脳の機能マップ構築にも取り組み [16]、マウス運動皮質に焦点を当てた神経活動記録や空間トランスクリプトームが報告されている [17]。日本の Brain/MINDS(2014-2023 年) では小型の靈長類であるマーモセットの脳の構造と機能マップ解明を目指し、コネクトーム、大規模カルシウムイメージング、疾患動物モデルの開発が行われている [18]。中国の China Brain Project(2016-2030 年) は、脳と人工知能に関する取り組みを行い、特に大型の靈長類であ

るマカクザルを扱うことに特色としている [19]。これらのプロジェクトの恩恵もあり、シミュレーション規模、生理実験の計測規模と情報量の拡大は指数関数的に増加している [20][21]。シミュレーション神経科学と他分野の関わりにおいて代表的なものとして、1) 身体のシミュレーションと 2) 人工知能が挙げられる。1) 身体シミュレーションについては、脳は外部環境と相互作用するなかで情報処理を行うため、身体モデルと脳モデルの連成シミュレーションによる、運動情報処理や神経疾患に関する取り組みが行われている。例えば、前述の Human Brain Project の Neurorobotics Platform では、NEST などの神経回路シミュレータと身体シミュレータを統合したシステムの開発が行われている [22]。2) 人工知能については、脳の情報処理機構に基づく知能システムとして開発された深層学習の低消費電力化を目指し、人工ニューラルネットワークの発火神経回路への置き換え [23] やニューロモーフィックチップの開発と実装の取り組みが盛んに行われている [24]。これらのシミュレーション神経科学と他分野の連携による研究・技術開発の進展の可能性がより顕在化してきており、前述の各国の大型プロジェクトでも見られるよう に、連携が進んでいる。

2.5.1.2 ニューロモーフィックデバイス

Society 5.0 「超スマート社会」は、人類に真の幸福を与えることを目指す未来社会として定義され、人工知能は最重要インフラとして位置づけられている。サイバー空間の人工知能がビッグデータを解析し、そのためにフィジカル空間から各種情報が送信される。しかしながら、人工知能の従来システムは、ハイスペックなハードウェアで実行する複雑・長大なソフトウェアで、マシンサイズ・消費電力が甚大である。さらに、IoT からのデータ通信は大量で、現状のまま対策が無ければ、非常な大消費電力となる。この問題を解決する次世代のデバイスとして、ニューロモーフィックシステムが注目を集めている。生物の脳は、極めてコンパクト・ローエネルギー・個々の神経細胞が低速なわりに全体としてハイスピードで、その秀逸な特徴をハードウェアから模倣する計算基盤がニューロモーフィックシステムである。これまで、ニューロモーフィックシステムは、活発に研究開発が行なわれてきた [?]。米欧中でも研究開発は活発で、一部の用途は既に実用化されている。たとえば、アメリカでは、2008 年頃からアメリカ国防高等研究計画局 (DARPA) による研究開発が行われ、「SyNAPSE」というプロジェクトには IBM や複数の大学が参加していて、IBM は「TrueNorth」というレジェンド的な LSI チップを発表している [25]。最近では、Intel の Loihi・Loihi 2 や [26] [27]、BrainChip の Akida などは、既に市販されている。ヨーロッパでは、ヒューマンブレインプロジェクトがあり、SpiNNaker というシステムを発表している [28]。中国では、発表は少ないものの、特許出願は多く、国家の性格から、このような最先端技術には、膨大なリソースを投入していることは、想像に難くない。しかし、研究開発はいまだ不十分・本命技術が未確定・実用化も限定的で、さらに研究開発が求められている。たとえば、上述のニューロモーフィックシステムは、シリコン半導体デジタル回路を用いているが、後述の省エネルギー性能に優れたアナログ回路のニューロモーフィックシステムは、高度な機能を実現できる超高集積化には至っておらず、依然として課題が多い。また、アナログ回路のニューロモーフィックシステムでは、自律学習のしくみは、いまだ理論研究での課題が多く、ニューロモーフィックシステムへの実装例は非常に限られている。将来的に予測されるエネルギー消費を解決するため、超コンパクト・超ローパワーのニューロモーフィックシステムのための研究開発が必要とされている。脳を真似るだけでなく、「脳を超越」する超コンパクト・ローパワー・ハイ

スピード化と、その汎用人工知能（AGI）への搭載の可能性を追求することが求められる。その次世代のニューロモーフィックシステムを広範に搭載することによって、サイバー空間は超ローパワーのシステムに置換し、フィジカル空間とのインターフェイスのエッジは超コンパクトのシステムの搭載が可能でローカルな情報処理でデータ通信は必要量のみとなり、劇的な超低消費電力となることが期待できる。次世代のニューロモーフィックシステムは、Society5.0・SDGs・カーボンニュートラルを同時実現する未来社会に向け、2050年の情報関連の消費電力のゼロカーボンを実現する可能性を持っている。なお、英語では Neuromorphic System と呼びかたに疑問の余地は無いのに対して、日本語ではニューロモーフィックシステムとニューロモルフィックシステムという 2 つおりの呼びかたがあるが、Google 検索によると前者は約 452,000 件・後者は 18,500 件（2022 年 05 月 01 日）となるため、ここではニューロモーフィックシステムと呼ぶことにする。

2.5.2 長期目標と社会貢献

2.5.2.1 脳・神経回路シミュレーション

神経科学の最大の目標は、脳の情報処理機構の全容解明である。特に、人間の脳による高次脳機能の理解は、医学、産業、教育、芸術、哲学など社会における幅広い分野への影響があり、究極の目標といえよう。一方で、昆虫の脳は人間の脳に比べ遙かにスケールが小さいながら、昆虫は知能的行動と呼ばれる性質の大部分の要素を備えており、知的情報処理の一般的な性質を理解するのに扱いやすいモデルである。このような複雑度の違う異なる種の脳について、扱いやすさの違いから調べられることが異なるため、神経科学ではそれらの実験やシミュレーションが並行して行われてきた。今後も、複雑度の違う異なる種の脳の研究が行われていくことは間違いない。そこで、人間を含む哺乳類の脳と昆虫の脳のシミュレーションについて考えていく。哺乳類の脳に関しては、2000 年代以前、電気生理学における神経細胞レベルの活動記録は単一の脳領域の数十～百ニューロン程度に限られ、解剖学における結合記録はある領域内や領域間の結合の断片的な情報しか得られなかった。しかし、2000 年以降から現在に至る光遺伝学、細胞外記録技術、コネクトーム計測技術の発達により、記録可能な神経回路規模は飛躍的に拡大していて、複数領域の神経活動や全脳の包括的な結合情報を得る技術開発が進んでいる [20]。一方、計算機性能の指数関数的な向上により、シミュレーション可能な神経回路規模は全脳規模にまで拡大している [21]。これらの進展から、今後の神経回路シミュレーションは包括的な結合情報と神経活動の情報に基づく神経回路モデルによって、広範な領域間の相互作用を明らかにすることが中心的な目標となるとみられる。2022 年時点の生理学実験・シミュレーションでは、げっ歯類の全脳が当面の目標となっているが、今後 10-20 年の中長期で見れば、一般的に扱える研究対象として、小型霊長類及びヒトの全脳へと軸足が移っていくと予想される。長期的には、複数脳領域の協調で発生するヒトの思考に関する高次脳機能の理解、複数脳領域間で起きるてんかん、うつ病、パーキンソン病、統合失調症などの脳の疾患の理解や治療法などについて、全脳シミュレーションによる構成論的なアプローチが期待される。これらの様々な課題間で、神経回路モデルの記述レベルは異なり実現時期もそれに合わせて異なるだろう。例えば、病気に対する薬理学的效果を検討するためには、個々の神経細胞の構造や、細胞内の AMPAR, Ca²⁺, CaM, CaMKII, DARPP32, PKA, PP1, PP2B などに代表されるシグナル伝達系モデル（数 10～150 の分子種）の表現も必要になる

ことがある。このような詳細度の高いシミュレーションでは必要な演算量は数10倍から100倍増大するため、実現にはより時間がかかることになる。長期的な目標の中でも究極的なものといえるヒトの脳の神経回路シミュレーションによる理解は、いわば我々の精神世界の理解であり、ヒトとは何かといった根源的な問いに対してなにかしらのヒントを得ることになるかもしれない。さらに、ヒトの全脳シミュレーションを実時間、またはそれ以上の速度で実行できるようなら、人間の知性を持つシステムを様々な分野で応用できる可能性があり、人間の文明全体の発展を大いに加速するだろう。一方、小規模ながら一つの脳をよりボトムアップで構築していく方向もある。その対象として有力なのが昆虫である。たとえば、多くの昆虫は甘いショ糖刺激と特定の匂いを同時に与えられると特定の匂いに従って摂食行動をとるようになる。これは、昆虫には感覚を記憶する能力、別の種類の感覚を結びつける能力、欲求を満たすために行動する能力があることを示している。喧嘩を行うコオロギや、人間の顔を認識するハチの例もある。すなわち、人間が行う未解明な知能もその複雑さを度外視すればおおむね昆虫も持っている。その上で昆虫の全脳シミュレーションは、その規模の小ささのため、モデル構築のために必要な生理実験データと、モデルを実行するために必要な計算資源が、高等な動物の場合に比べ遙かに少なくてすむ利点がある。また、昆虫を機械とインターフェースで結合する試みや、遺伝子変容の容易さなどの特徴もある。そのため、実環境下における昆虫脳モデルを搭載したロボットを用いた脳機能の検証や、また、生理学実験で観測、操作された神経細胞の活動との比較を行い易いという利点がある。そのため、昆虫全脳シミュレーションによる知的情報処理の一般的な性質の研究は、哺乳類などの高等な動物の脳を扱うより早期に実現可能になるとみられる。また、その後の哺乳類、霊長類、人間の脳全脳シミュレーションにとってのテストケースとしても、役割を果たすだろう。

2.5.2.2 ニューロモーフィックデバイス

人工知能の当面の、そしていつ達成されるかは未だもって不明確な目標として、人間と同等に知的なふるまいをみせるコンピュータの実現が考えられる。未来学者のレイ・カーツワイルの予言によれば、人工知能が人間の知能を超える技術特異点が今世紀中にやってくるとみられている。一方で著者が知る限りにおいては、人工知能の諸分野の第一線の研究者で、そのような特異点がすぐに来ると考えている人物は少ない。

画像認識のようにデータセットが十分整備されていれば人間並みの精度で物体の名前がわかり、囲碁のようにゲームのすべての情報がプレイヤーにわかる場合であればトッププロ並みの腕前でプレイできるように学習できる。囲碁のように非常に難しいとされてきたゲームにおいてコンピュータが人間を越えたという点でセンセーショナルな成果ではあるが、基本的には画像による盤面の認識と強化学習の組み合わせただけである。視覚という感覚情報のみから物の名前を言い当てるだけでなく、人間並みに多様な感覚情報の情報から柔軟な環境認識を達成する、もしくは、報酬も不確定な環境下で、意思決定できるようなコンピュータが登場する…そのようなAIを実現するには、まだまだ多くの問題が残っている。

サイエンスの側面から考えられる社会貢献としては、知能とは何かという問いに構成論的なアプローチで答えることであろう。人間のように知的にふるまう計算機を目指すという分野の性質上、人間そのものにヒントを得た研究も多い。脳神経系に着想を得た機械学習手法としてのニューラルネットワークや深層学習、網膜や視覚野などを模した画像認識手法などがこれにあたる。一方で数学的に学習の最適化につい

て突き詰めたり、ナイーブな着想を実直に試行したりというようなところからブレークスルーが起こる場合もあり、深層学習の発展もどちらかといえばこうしたアプローチによっている。より人間に近い知的なシステムが達成されれば、知能とは何かという理学的な問いにフィードバックできる可能性がある。

また同時に—人間並みの知能を実現するにはまだ多くの時間がかかるといつても—その研究の過程の中で、社会に対してエンジニアリング的な側面から大きな影響を与えると考えられる。抽象的に述べれば、まずは単一の感覚情報で何らかの認識を行うような分野で精緻化が進み、次に多様な感覚情報を総合的に判断して認識するようなステージへと進む。最終的にはそのような認識結果として得られるシンボル情報から、より複雑な意思決定を計算機が行うようになるだろう。

医療画像からの病変の発見や防犯・監視カメラの自動処理といった部分が最初に達成されるはずである。このように限られた感覚情報のデータから対象となる事物を認識する作業でも、現在は人的コストを割いて行われている。かつてのエキスパートシステムでは、何に注目してどういう反応であればどういう認識結果を返すというルールをすべて明示的に与える必要があった。深層学習を中心とする機械学習の進歩によって、入力データと求める認識結果の正解を与えればこうしたルールを内部的に獲得してくれるようになった。

視覚のみから事物を認識するといった単一モーダルデータ認識の次に、より広範で多くの感覚情報を用いた認識へと進化すると考えられる。そこでは、仮にある感覚情報のデータが欠損したり信頼性の低いデータとなったりしていても、ほかの感覚情報から情報を補えると期待される。表情だけでなく声や身振り手振りなども含めて正確に人間の感情を推定できれば、人間の感情を考慮したロボットやコンピュータとのインタラクションが可能になり、接客やオペレーションの補助が可能にある。周囲の環境をさまざまなセンサ情報から頑健に把握できるようになれば、自動運転のように行動パターンの絞られたドメインにおいては意思決定システムが実現できる。

さらに将来的には、人間の知的活動そのものを支援するシステムが期待される。現在の状況を様々な感覚情報のデータから判断するだけではなく、その判断結果として帰着するシンボル情報を統合して次の行動を決定するような高度に知的な処理が自動化される。計算機の補助を受けながら指すアドバンスド・チェスのように、教育や経営判断といった知的作業にも計算機の補助が入るようになるとみられる。

2.5.3 課題とその解決に必要な計算手法・アプリケーション

2.5.3.1 脳・神経回路シミュレーション

神経回路シミュレーションは、回路の構成要素である個々のニューロンについて、(1) シナプス入力(入力)、(2) 膜電位(状態)、(3) 発火判定とスパイク伝播(出力)を、各時刻に全てのニューロンについて計算していく。それぞれは微分方程式(常微分・偏微分)で記述され、常微分方程式の場合は陽解法(e.g., ルンゲクッタ法)、偏微分方程式の場合は差分化し陰解法(例えばクランクニコルソン法)で計算される。陰解法の場合は連立一次方程式を解く必要があり、その際はガウスの消去法が主に用いられている。ニューロンの挙動を記述する微分方程式は、細胞の形状を無視し仮想的な点で表すモデル(点モデル)と、形状を考慮し細胞内をどのように信号が伝わるかまで表すモデル(空間モデル)の2種類に大別される。時間ステップは全ニューロンで固定にするのが一般的であるが、ニューロン毎に動的に可

変にすることもある。ニューロンのスパイク生成は数ミリ秒の非常に速い時間スケールで行われるが、それ以外ではニューロンはより長い時間スケールで振る舞うため、このような可変タイムステップは性能向上に大きく寄与する。2010年代からの第3次人工知能ブームでの深層学習の発達において、様々な科学技術分野で、物理シミュレーションを深層学習による代理モデルで置きかえ、計算量削減をはかる試みが盛んに行われている。ニューロンのシミュレーションにおいても、複雑な形態や複数のイオンチャネルを持つニューロンモデルの再現と計算高速化の報告がでてきていている。代理モデルは計算量の削減に有効であり、大規模神経回路シミュレーションにおける利用が今後必要であろう。神経回路シミュレーションは一般的に、個別のプログラムを書くというよりは、既に利用可能な汎用のシミュレータを用いて実施されることが多い。これは、大規模なモデルのプログラムを書くことが煩雑である一方、計算内容は比較的単純かつ明らかであり、性能を重視しなければ一から書く必然性が薄いからである。様々なシミュレータが存在し、シングルコンパートメントニューロンモデル向けの NEST [29]、マルチコンパートメントニューロンモデル向けの NEURON [30] などは、マルチコア CPU の大型並列計算機に対応していて、幅広く使われている。最近では、GPU や FPGA 等のアクセラレータを利用可能なものも登場している [31] [32] [33]。一方、ベクトル演算機能、マルチコア、キャッシュメモリ、通信機能などの性能を引き出すために、専用のプログラム開発も行われる。例えばスペコン「菖蒲」上で動作するネコ小脳のシミュレーションは専用のプログラムとして開発され [34]、NEURON シミュレータのコードが京向けにカスタマイズされた例などがある (Invertivrate Brain Platform: https://invbrain.neuroinf.jp/modules/htmldocs/IVBPF/IOSSIM/index.html?ml_lang=ja)。2022年時点における大型計算機を用いる神経回路シミュレーションの課題として大別すると、(1) 比較的単純なニューロンモデルを用いた霊長類全脳規模の巨大なネットワークシミュレーションと、(2) 小型無脊椎動物の全脳やげっ歯類脳の部分的な領域を対象とした、形態やイオンチャネル特性を含む非常に詳細なニューロンモデルで構築される、小一中規模のネットワークシミュレーションが挙げられる。これまでの(1) 単純な神経モデルによる大規模神経回路シミュレーションの例として、積分発火型モデルを用いたヒト全脳の半分に相当する 445 億個を持つ大脳皮質一小脳一視床神経回路シミュレーションが富岳で行われ、生物学的 1 秒を 20 秒程度で実行されている。神経回路規模拡大の可能性についてはある程度確認された状況である。これまでの(2) の詳細ニューロンモデルによる小一中規模神経回路シミュレーションの例としては、ポスト「京」萌芽的課題で行われた昆虫全脳リアルタイムシミュレーションや、ラットの体性感覚野 [10] や視覚野 [11] の神経回路シミュレーションがあり、無脊椎動物全脳や哺乳類の特定の脳領域のシミュレーションの例が増えてきている。これら(1),(2) 神経回路シミュレーションは、主にメモリバンド幅が実行効率を決める要因となっていて、キャッシュメモリによる効率化も有効である。今後、ムーアの法則による計算機性能向上が継続するかぎり、計算性能の向上が期待できる。(1) の単純な神経モデルによる大規模神経回路シミュレーションにおける今後の課題は、大型のスペコンにおける全脳シミュレーションの複数実行、相対的に規模の小さいスペコンによる全脳シミュレーションの実行、計算コストの高い Hodkin-Huxley 型などの複雑な神経細胞モデルを用いた全脳シミュレーション実行、急増する大規模神経活動データやコネクトームのデータの全脳モデルへの導入、実環境での機能検証のための実時間実行に向けた高速化などが想定される。神経科学における様々な課題に対して、全脳シミュレーションを利用可能にすることが必要であろう。(2) の詳細ニューロンモデルによる小一中規模神経回路シミュレーションの課題としては、無脊椎動物の同定されたニューロンの生理学的パラメータを、今後の向上す

る計算性能を活かしてデータ同化によって推定し、解剖学的構造だけでなく生理学的性質までを再構成する取り組みが考えられる。また、詳細ニューロンモデルを用いた神経回路モデルの規模の拡大によるげつ歯類や小型靈長類の全脳シミュレーションは、今後 10-20 年の計算性能向上により実現可能な対象になるとと考えられる。

2.5.3.2 ニューロモーフィックデバイス

生物の脳を模倣するのがニューロモーフィックシステムであるため、その構成要素も脳と同じでニューロンとシナプスである。ニューロンは、入力が閾値を超えるか否かにより、出力が発火・静止の 2 状態を生成する、非線形信号生成素子である。シナプスは、あるニューロンからの信号を重み付けしながらほかのニューロンへ伝達する、可変接続強度信号伝達素子である。ニューロモーフィックシステムの中でのシナプス素子の役割は、積和計算である。ニューロンの数に比べてシナプスの数は数百倍～数万倍と圧倒的に多いため、コンパクト化・ローパワー化・ハイスピード化への影響はシナプスが支配的であるので、ここではニューロンのシナプス素子について検討する。候補としては、シリコン半導体デジタル回路 [35] やメモリスタ [36] [37] [38] [39] [40]・メムキャパシタと [41] [42] といったメムデバイスなどがある。それぞれの特徴をまとめると、下表のとおりである。シナプス素子の特徴

シリコン半導体デジタル回路は、これまでの永年の技術蓄積があるため、技術成熟度は高く、制御性も良く、間違いない製品化が可能な素子であるが、アナログ回路に比べて、回路規模が大きく、コンパクト化・ローパワー化には劣り、また、積和計算に多くのステップを要し、ハイスピード化にも劣る。アナログ素子であるメモリスタは、電圧の印加やそれにより流れる電流の履歴により、コンダクタンスを変化させることのできる抵抗素子で、研究開発の歴史は浅いため、技術成熟度・制御性が不十分であるが、今後の研究開発で改善できる余地はある。デバイス構造が簡単であるため、コンパクト化に優れるが、定常電流が流れるため、ローパワー化には向きのところがある。また、積和計算は物理的な 1 ステップのみで実現できるため、ハイスピード化に優れる。材料としては、シリコン半導体ではなく、金属酸化物半導体の研究開発がさかんである。カーボンナノチューブや有機半導体などの黎明的な研究開発もある。メムキャパシタは、電圧の印加の履歴により、キャパシタンスを変化させることのできるキャパシタ素子で、メモリスタよりさらに研究開発の歴史は浅いため、やはり技術成熟度・制御性が不十分であるが、今後の研究開発でおおいに改善できる余地はある。メモリスタと同様に、デバイス構造が簡単であるため、コンパクト化に優れるとみられるが、キャパシタンスを確保するために、トレンチ構造のようなものが必要となりコンパクト化が限定的になる懸念はある。定常電流が流れず、充電に使われた電荷の再利用ができる可能性があり、究極のローパワー化の可能性がある。また、メモリスタと同様に、積和計算は物理的な 1 ステップのみで実現できるため、ハイスピード化に優れる。材料としては、各種の強誘電体材料が応用可能で、既にシリコン半導体技術とコンパチビリティのあるものもある。なお、人間の脳のような自律学習の機能として、スパイクタイミング依存可塑性 (Spike-Timing-Dependent Plasticity, STDP) のような、余分な制御回路を必要としない方式も必須である。しかし、まだまだ未解決の問題が多い。ここでは、ハイリスクではあるがハイリターンであるメムキャパシタを想定して、超コンパクト化・超ローパワー化・超ハイスピード化について、従来の人工知能システムと比べながら、考察してみる。超コンパクト化: 人工知能の従来システムは、ハイスペックなハードウェアで実行する複雑・長大なソフトウェア

で、マシンサイズは、たとえばスーパーコンピュータで、ひとつのフロア・ひとつのビル（あるいは冷蔵庫 100 台）を占有するほどのサーバー群である。現在メモリスタは、ReRAM としてではあるが既に数十 nm のディメンジョンでの実装が実証されていて、低消費電力による低熱発生により、積層可能性が高いとみられる。また、2022 年時点でのメムキャパシタは、やはり FeRAM としてではあるが既に数十 nm のディメンジョンでの実装が実用化されている。これらを踏まえ、ニューロモーフィックシステムにメムキャパシタを想定すると、たとえば将来的な微細化技術で各種デバイスのディメンジョンが $30 \times 30\text{nm}$ 程度は期待でき、より低い消費電力からも 3 次元集積化プロセス・デバイス技術で積層が 100 層程度は可能であると概算される。この場合、演算素子が人間の大脳皮質と同等の 100 兆個のニューロモルフィックシステムを構成するためには、チップ面積は、 $3 \times 3\text{cm}$ となると見積もられ、人間の大脳皮質のサイズの 2ℓ を大きく下回り、Society 5.0 のエッジ・IoT において十分に搭載が可能なサイズになると推測される。超ローパワー化: 現在の技術のまま、まったく省エネルギー対策がなされないと仮定すると、情報関連だけで、人工知能そのものの普及とデータ通信の増大のため、天文学的な消費電力が必要とされるという多くの予測がされている [2,3,4]。もちろんこれらの予測は現実的ではなく、逆にいえば、情報社会の進化を止めないために、消費電力の大幅な削減が必須であると言える。ニューロモーフィックシステムにメムキャパシタを想定すると、各種デバイスのディメンジョン $30 \times 30\text{nm}$ 、膜厚 15nm 、キャパシタンス 56aF (aF (アットファラド) = 10^{-18}F)、駆動電圧 0.2V 、実効周波数 100kHz とすると、充放電電力は、 $56\text{aF} \times (0.2\text{V})^2 \times 100\text{kHz} \simeq 200\text{fW}$ となる。人間の脳に相当する規模のシナプス 100 兆の超高集積化をすると、ニューロモーフィックチップの全体の消費電力は 20W となり、人間の脳の消費エネルギーである 20W と同等の超ローパワー化となる。そして、まず、人工知能そのものの消費電力を削減する。上表によれば、消費電力の削減効果は、たとえば 20W (超ローパワーのニューロモーフィックシステム) / 30kW (Loihi) = 概ね $1/1,000$ である。次に、データ通信の削減について推定する。エッジに人工知能が搭載されサイバー空間へは処理後データのみ送信されるとすると、たとえばデータ通信の削減の効果も $1/1,000$ になるようにするには、8K の画像 (データ量 50MB) の代わりにその画像認識の結果としてデータ量 $50\text{MB}/1,000 = 50\text{kB}$ が送信されるととして、 50kB ならば、その画像認識の結果を、さまざまな追加情報を含め、記述するのに十分なデータ量である。すなわち、人工知能そのものとデータ通信の両方で、消費電力の削減効果は $1/1,000$ 程度の実現が可能である。すると、2030 年には現在とほぼ変わらず年間 24PWh ・2050 年でも 29PWh が必要とされるだけであり、消費電力の課題は解決されることになる。超ハイスピード化: 超ローパワー化が具現化できるとして、処理スピードなどのパフォーマンスが犠牲にできるかというと、ユーザーは決して受け入れないため、パフォーマンスの維持のため、超ハイスピード化も要求される。人工知能の処理の大半を占める積和演算は、人工知能の従来システムではコンピュータのクロックレベルでは 100 ステップのオーダーの処理が必要であり、ニューロモルフィックシステムの従来システムでもシリコン半導体のデジタル処理では 10 ステップのオーダーの処理が必要とされると考えられる。本研究開発課題では、デバイスレベル・回路レベルのアナログ処理が可能であり、積和演算は基本的に物理レベルの 1 ステップで完了する。デバイスの動作速度が同一であれば、単純に 10 倍のスピード化が可能となる。さらに、超ローパワー化の恩恵を受けてマルチコア化も容易で、マルチコア化 10 並列を想定すれば、100 倍の超ハイスピード化が可能となるだろう。

2.5.4 ロードマップ

2.5.5 必要な計算機資源

2032年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13節の表2.1に記載した。また、計算機資源の見積りの詳細は第4.5節に記載している。

参考文献

- [1] 竹内郁雄（編）. AI 人工知能の軌跡と未来. 別冊日経サイエンス. 日経サイエンス社, 2016.
- [2] Katja Brose. Global neuroscience. *Neuron*, Vol. 92, No. 3, pp. 557–558, 2016.
- [3] 人工知能学会（編）. 深層学習. 近代科学社, 2015.
- [4] Xue Fan and Henry Markram. A brief history of simulation neuroscience. *Frontiers in Neuroinformatics*, Vol. 13, p. 32, 2019.
- [5] Christof J. Schwiening. A brief historical perspective: Hodgkin and Huxley. *The Journal of Physiology*, Vol. 590, pp. 2571–2575, 2012.
- [6] Eugene M. Izhikevich and Gerald M. Edelman. Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 105, No. 9, pp. 3593–3598, 2008.
- [7] Rajagopal Ananthanarayanan, Steven K. Esser, Horst D. Simon, and Dharmendra S. Modha. The cat is out of the bag: Cortical simulations with 10^9 neurons and 10^{13} synapses. In *Supercomputing 09: Proceedings of the ACM/IEEE SC2009*, 2009.
- [8] Kunkel S, Schmidt M, Eppler JM, Plessner HE, Masumoto G, Igarashi J, Ishii S, Fukai T, Morrison A, Diesmann M, and Helias M. Spiking network simulation code for petascale computers. *Front Neuroinform*, Vol. 8, p. 78, 2014.
- [9] Chris Eliasmith, Terrence C. Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Charlie Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *Science*, Vol. 338, No. 6111, pp. 1202–1205, 2012.
- [10] Henry Markram, Eilif Muller, Srikanth Ramaswamy, Michael W. Reimann, Marwan Abdellah, Carlos Aguado Sanchez, Anastasia Ailamaki, Lidia Alonso-Nanclares, Nicolas Antille, Selim Arsever, Guy Antoine Ateneckeng Kahou, Thomas K. Berger, Ahmet Bilgili, Nenad Buncic, Athanassia Chalimourda, Giuseppe Chindemi, Jean Denis Courcol, Fabien Delalondre, Vincent Delattre, Shaul Druckmann, Raphael Dumusc, James Dynes, Stefan Eilemann, Eyal Gal, Michael Emiel Gevaert, Jean Pierre Ghobril, Albert Gidon, Joe W. Graham, Anirudh Gupta, Valentin Haenel, Etay Hay, Thomas Heinis, Juan B. Hernando, Michael Hines, Lida Kanari, Daniel Keller, John Kenyon, Georges Khazen, Yihwa Kim, James G. King, Zoltan Kisvarday, Pramod Kumbhar, Sébastien Lasserre, Jean Vincent Le Bé, Bruno R.C. Magalhães, Angel Merchán-Pérez, Julie Meystre, Benjamin Roy Morrice, Jeffrey Muller, Alberto Muñoz-

- Céspedes, Shruti Muralidhar, Keerthan Muthurasa, Daniel Nachbaur, Taylor H. Newton, Max Nolte, Aleksandr Ovcharenko, Juan Palacios, Luis Pastor, Rodrigo Perin, Rajnish Ranjan, Imad Riachi, José Rodrigo Rodríguez, Juan Luis Riquelme, Christian Rössert, Konstantinos Sfyrakis, Ying Shi, Julian C. Shillcock, Gilad Silberberg, Ricardo Silva, Farhan Tauheed, Martin Telefont, Maria Toledo-Rodriguez, Thomas Tränkler, Werner Van Geit, Jafet Villafranca Díaz, Richard Walker, Yun Wang, Stefano M. Zaninetta, Javier Defelipe, Sean L. Hill, Idan Segev, and Felix Schürmann. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, Vol. 163, No. 2, pp. 456–492, 2015.
- [11] Yazan N Billeh, Binghuang Cai, Sergey L Gratiy, Christof Koch, Stefan Mihalas, and Anton Arkhipov. Systematic Integration of Structural and Functional Data into Multi-scale Models of Mouse Primary Visual Cortex. *Neuron*, Vol. 106, pp. 388–403.e18, 2020.
- [12] Jun Igarashi, Hiroshi Yamaura, and Tadashi Yamazaki. Large-Scale Simulation of a Layered Cortical Sheet of Spiking Network Model Using a Tile Partitioning Method. *Frontiers in Neuroinformatics*, Vol. 13, pp. 1–15, 2019.
- [13] Hiroshi Yamaura, Jun Igarashi, and Tadashi Yamazaki. Simulation of a Human-Scale Cerebellar Network Model on the K Computer. *Frontiers in Neuroinformatics*, Vol. 14, , 2020.
- [14] Jakob Jordan, Tammo Ippen, Moritz Helias, Itaru Kitayama, Mitsuhsisa Sato, Jun Igarashi, Markus Diesmann, and Susanne Kunkel. Extremely scalable spiking neuronal network simulation code: From laptops to exascale computers. *Frontiers in Neuroinformatics*, 2018.
- [15] Jun Igarashi, Tadashi Yamazaki, Hiroshi Yamaura, Kentaro Nomura, Sun Zhe, Carlos Enrique Gutierrez, and Kenji Doya. A spiking neural network simulation of layered sheet of cortico-cerebello-thalamic circuit at human-scale. *Proceedings of Society for Neuroscience annual meeting 2022*, 2022.
- [16] Lyric A. Jorgenson, Williamt Newsome, David J. Anderson, Cornelia I. Bargmann, Emery N. Brown, Karl Deisseroth, John P. Donoghue, Kathy L. Hudson, Geoffrey S.F. Ling, Peter R. Macleish, Eve Marder, Richard A. Normann, Joshua R. Sanes, Mark J. Schnitzer, Terrence J. Sejnowski, David W. Tank, Roger Y. Tsien, Kamil Ugurbil, and John C. Wingfield. The BRAIN Initiative: developing technology to catalyse neuroscience discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 370, No. 1668, 2015.
- [17] Callaway, et al. A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature*, Vol. 598, No. 7879, pp. 86–102, 2021.
- [18] Hideyuki Okano, Erika Sasaki, Tetsuo Yamamori, Atsushi Iriki, Tomomi Shimogori, Yoko Yamaguchi, Kiyoto Kasai, and Atsushi Miyawaki. Brain/MINDS: A Japanese National Brain Project for Marmoset Neuroscience. *Neuron*, Vol. 92, No. 3, pp. 582–590, 2016.
- [19] Mu ming Poo, Jiu lin Du, Nancy Y. Ip, Zhi Qi Xiong, Bo Xu, and Tieniu Tan. China Brain Project: Basic Neuroscience, Brain Diseases, and Brain-Inspired Computing. *Neuron*, Vol. 92, No. 3, pp. 591–596, 2016.
- [20] 平理一郎. 高次元神経活動から眺めるシステム神経科学. 神経回路学会誌, Vol. 27, pp. 152–164,

2020.

- [21] 五十嵐潤. 大型計算機と脳計測の技術動向から予測する哺乳類全脳シミュレーションの将来. 神経回路学会誌, Vol. 28, pp. 172–182, 2021.
- [22] Egidio Falotico, Lorenzo Vannucci, Alessandro Ambrosano, Ugo Albanese, Stefan Ulbrich, Juan Camilo Vasquez Tieck, Georg Hinkel, Jacques Kaiser, Igor Peric, Oliver Denninger, Nino Cauli, Murat Kirtay, Arne Roennau, Gudrun Klinker, Axel Von Arnim, Luc Guyot, Daniel Peppicelli, Pablo Mactinaz-Cañada, Eduardo Ros, Patrick Maier, Sandro Weber, Manuei Huber, David Plecher, Florian Röhrbein, Stefan Deser, Alina Roitberg, Patrick Van Der Smagt, Rüdiger Dillman, Paul Levi, Cecilia Laschi, Alois C. Knoll, and Marc Oliver Gewaltig. Connecting artificial brains to robots in a comprehensive simulation framework: The neurorobotics platform. *Frontiers in Neurorobotics*, Vol. 11, No. JAN, p. 2, 2017.
- [23] Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature Communications* 2020 11:1, Vol. 11, No. 1, pp. 1–15, 2020.
- [24] Steve Furber. Large-scale neuromorphic computing systems. *Journal of Neural Engineering*, Vol. 13, No. 5, p. 051001, 2016.
- [25] Paul A. Merolla, John V. Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, Jun Sawada, Filipp Akopyan, Bryan L. Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, Bernard Brezzo, Ivan Vo, Steven K. Esser, Rathinakumar Appuswamy, Brian Taba, Arnon Amir, Myron D. Flickner, William P. Risk, Rajit Manohar, and Dharmendra S. Modha. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, Vol. 345, pp. 668–673, 2014.
- [26] Mike Davies, Narayan Srinivasa, Tsung Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, Yuyun Liao, Chit Kwan Lin, Andrew Lines, Ruokun Liu, Deepak Mathaikutty, Steven McCoy, Arnab Paul, Jonathan Tse, Guruguhanathan Venkataramanan, Yi Hsin Weng, Andreas Wild, Yoonseok Yang, and Hong Wang. Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro*, Vol. 38, No. 1, pp. 82–99, 2018.
- [27] Mike Davies, Andreas Wild, Garrick Orchard, Yulia Sandamirskaya, Gabriel A. Fonseca Guerra, Prasad Joshi, Philipp Plank, and Sumedh R. Risbud. Advancing Neuromorphic Computing with Loihi: A Survey of Results and Outlook. *Proceedings of the IEEE*, Vol. 109, No. 5, pp. 911–934, 2021.
- [28] Steve B Furber, Francesco Galluppi, Steve Temple, and Luis A Plana. The spinnaker project. *Proc IEEE*, Vol. 102, pp. 652–665, 2014.
- [29] Marc-Oliver Gewaltig and Markus Diesmann. Nest (neural simulation tool). *Scholarpedia*, Vol. 2, No. 4, p. 1430, 2007.
- [30] Nicholas T. Carnevale and Michael L. Hines. *The NEURON Book*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.

- [31] Esin Yavuz, James Turner, and Thomas Nowotny. GeNN: a code generation framework for accelerated brain simulations. *Scientific Reports* 2016 6:1, Vol. 6, No. 1, pp. 1–14, 2016.
- [32] Pramod Kumbhar, Michael Hines, Jeremy Fouriaux, Aleksandr Ovcharenko, James King, Fabien Delalondre, and Felix Schürmann. CoreNEURON : An Optimized Compute Engine for the NEURON Simulator. *Frontiers in Neuroinformatics*, Vol. 13, p. 63, 2019.
- [33] Kit Cheung, Simon R. Schultz, and Wayne Luk. NeuroFlow: A general purpose spiking neural network simulation platform using customizable processors. *Frontiers in Neuroscience*, Vol. 9, p. 516, 2016.
- [34] Tadashi Yamazaki, Jun Igarashi, Junichiro Makino, and Toshikazu Ebisuzaki. Real-time simulation of a cat-scale artificial cerebellum on PEZY-SC processors. *International Journal of High Performance Computing Applications*, 2019.
- [35] 河野崇. 次世代人工知能へのアプローチ：シリコン神経ネットワーク. 情報管理, Vol. 60, No. 7, pp. 461–470, 2017.
- [36] Leon O. Chua. Memristor—The Missing Circuit Element. *IEEE Transactions on Circuit Theory*, Vol. 18, No. 5, pp. 507–519, 1971.
- [37] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* 2015 521:7550, Vol. 521, No. 7550, pp. 61–64, 2015.
- [38] M. Ishii, U. Shin, K. Hosokawa, M. Brightsky, W. Haensch, S. Kim, S. Lewis, A. Okazaki, J. Okazawa, M. Ito, M. Rasch, W. Kim, and A. Nomura. On-Chip Trainable 1.4M 6T2R PCM Synaptic Array with 1.6K Stochastic LIF Neurons for Spiking RBM. *Technical Digest - International Electron Devices Meeting, IEDM*, Vol. 2019-December, , 2019.
- [39] Vinod K. Sangwan, Hong Sub Lee, Hadallia Bergeron, Itamar Balla, Megan E. Beck, Kan Sheng Chen, and Mark C. Hersam. Multi-terminal memtransistors from polycrystalline monolayer molybdenum disulfide. *Nature* 2018 554:7693, Vol. 554, No. 7693, pp. 500–504, 2018.
- [40] Mutsumi Kimura, Yuki Shibayama, and Yasuhiko Nakashima. Neuromorphic chip integrated with a large-scale integration circuit and amorphous-metal-oxide semiconductor thin-film synapse devices. *Scientific Reports* 2022 12:1, Vol. 12, No. 1, pp. 1–6, 2022.
- [41] Kai Uwe Demasius, Aron Kirschen, and Stuart Parkin. Energy-efficient memcapacitor devices for neuromorphic computing. *Nature Electronics* 2021 4:10, Vol. 4, No. 10, pp. 748–756, 2021.
- [42] Mutsumi Kimura, Yuma Ishisaki, Yuta Miyabe, Homare Yoshida, Isato Ogawa, Tomoharu Yokoyama, Ken-ichi Haga, Eisuke Tokumitsu, and Yasuhiko Nakashima. Neuromorphic System Using Memcapacitors and Autonomous Local Learning. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–8, 2021.
- [43] Eric Richard Kandel, James Harris Schwartz, Thomas M. Jessell, and Sarah Mack, editors. *Principles of neural science*. McGraw-Hill Medical, 2013.
- [44] The Blue Brain Project, 2005. <http://bluebrain.epfl.ch/>.
- [45] Human Brain Project, 2013. <https://www.humanbrainproject.eu/>.

- [46] The BRAIN Initiative, 2014. <https://www.braininitiative.nih.gov/>.
- [47] 革新的技術による脳機能ネットワークの全容解明プロジェクト, 2014. <http://brainminds.jp/>.
- [48] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge, 2014.
- [49] JM Nageswaran, N Dutt, JL Krichmar, A Nicolau, and AV Veidenbaum. A configurable simulation environment for the efficient simulation of large-scale spiking neural networks on graphics processors. Vol. 22, No. 5-6, pp. 791–800, 2009.
- [50] J Igarashi, O Shouno, T Fukai, and H Tsujino. Real-time simulation of a spiking neural network model of the basal ganglia circuitry using general purpose computing on graphics processing units. *Neural Netw*, Vol. 24, No. 9, pp. 950–960, 2011.
- [51] Tadashi Yamazaki and Jun Igarashi. Realtime cerebellum: A large-scale spiking network model of the cerebellum that runs in realtime using a graphics processing unit. *Neural Netw*, Vol. 47, No. 11, pp. 103–111, 2013.
- [52] Masato Gosui and Tadashi Yamazaki. Real-world-time simulation of memory consolidation in a large-scale cerebellar model. *Front Neuroanat*, Vol. 10, p. 21, 2016.
- [53] Christof Koch and Idan Segev, editors. *Methods in Neuronal Modeling: From Ions to Networks*. MIT Press, 2nd edition, 1998.
- [54] Wulfram Gerstner and Werner Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. 2nd edition, 2002.
- [55] Andrew P Davison, et al. Pynn: A common interface for neuronal network simulators. *Frontiers in Neuroinformatics*, Vol. 2, p. 11, 2008.
- [56] James M. Bower and David Beeman. *The book of GENESIS - exploring realistic neural models with the GENeral NEural models SImulation system*. Springer, 1995.
- [57] Junwen Luo, Graeme Coapes, Terrence Mak, Tadashi Yamazaki, Chung Tin, and Patrick Degenaar. Real-time simulation of passage-of-time encoding in cerebellum using a scalable fpga-based system. *IEEE Transactions on Biomedical Circuits and Systems*, Vol. 10, No. 3, pp. 742–753, 2016.
- [58] Runchun Wang, Gregory Cohen, Klaus Stiefel, Tara Hamilton, Jonathan Tapson, and André van Schaik. An fpga implementation of a polychronous spiking neural network with delay adaptation. *Frontiers in Neuroscience*, Vol. 7, p. 14, 2013.
- [59] 日経テクノロジー online. スパコン上に猫の小脳を再現、リアルタイムで動作, 6月 13 日 2016. <http://techon.nikkeibp.co.jp/atcl/news/16/061302534/>.
- [60] マイナビニュース. ネコの小脳を shoubu スパコンで実現, 6月 20 日 2016. <http://news.mynavi.jp/series/numericalbrain/001/>.
- [61] Eugene M Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. MIT Press, 2007.
- [62] Suzana Herculano-Houzel. The human brain in numbers: a linearly scaled-up primate brain. *Front Human Neurosci*, Vol. 3, p. 31, 2009.

- [63] M Hawrylycz, C Anastassiou, A Arkhipov, J Berg, M Buice, N Cain, NW Gouwens, S Gratiy, R Iyer, JH Lee, S Mihalas, C Mitelut, S Olsen, RC Reid, C Teeter, S de Vries, J Waters, H Zeng, and C Koch. Inferring cortical function in the mouse visual system through large-scale systems neuroscience. *Proc Natl Acad Sci USA*, Vol. 113, No. 27, pp. 7337–7344, 2016.
- [64] Quentin J M Huys and Liam Paninski. Smoothing of, and parameter estimation from, noisy biophysical recordings. *PLoS Comput Biol*, Vol. 5, No. 5, p. e1000379, 2009.
- [65] Dimitrios V Vavoulis, Volkmar A Straub and John A D Aston, and Jianfeng Feng. A self-organizing state-space-model approach for parameter estimation in Hodgkin-Huxley-type models of single neuron. *PLoS Comput Biol*, Vol. 8, No. 3, p. e1002401, 2012.
- [66] Ryuta Mizutani, Rino Saiga, Akihisa Takeuchi, Kentaro Uesugi, and Yoshio Suzuki. Three-dimensional network of drosophila brain hemisphere. *Journal of structural biology*, Vol. 184, No. 2, pp. 271–279, 2013.
- [67] Hirotaka Sato and Michel M Maharbiz. Recent developments in the remote radio control of insect flight. *Frontiers in Neuroscience*, Vol. 4, p. 199, 2010.
- [68] Ryo Minegishi, Atsushi Takashima, Daisuke Kurabayashi, and Ryohei Kanzaki. Construction of a brain-machine hybrid system to evaluate adaptability of an insect. *Robotics and Autonomous Systems*, Vol. 60, No. 5, pp. 692–699, 2012.
- [69] EM Izhikevich. Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral cortex*, Vol. 17, No. 10, pp. 2443–2452, 2007.
- [70] P Orio and D Soudry. Simple, fast and accurate implementation of the diffusion approximation algorithm for stochastic ion channels with multiple states. *PloS one*, Vol. 7, p. e36670, 2012.
- [71] L Chittka and J Niven. Are bigger brains better? *Current Biology*, Vol. 19, No. 21, pp. R995–R1008, 2009.
- [72] Tetsuya Fukuda, Tomoki Kazawa, and Ryohei Kanzaki. Establishment of the estimation method of the neural network using cma-es for elucidating the neural mechanism of a silk-worm moth brain. In *9th INCF Congress in Neuroinformatics*, 2016. 3-4 September, Reading, UK.
- [73] 李夏栄, 加沢知毅, ハウプトステファン周一, 神崎亮平. 確率的手法によるカイコガ触角葉神経細胞の生物物理的膜特性の推定. Technical Report 382, 電子情報通信学会技術研究報告 ニューロコンピューティング (NC), 2014.
- [74] 加沢知毅, 宮本大輔, 後藤晃彦, 朴希原, 福田哲也, 神崎亮平. 昆虫全脳シミュレーションへむけて. 日本神経回路学会誌, Vol. 22, No. 3, pp. 89–102, 2015.
- [75] 宮本大輔, 加沢知毅, 神崎亮平. 昆虫嗅覚系全脳シミュレーションに向けて: スーパコンピュータによる大規模脳シミュレーションの現在とその展望. 人工知能: 人工知能学会誌, Vol. 30, No. 5, pp. 630–638, 2015.
- [76] Romain Brette, Michelle Rudolph, Ted Carnevale, Michael Hines, David Beeman, James M. Bower, Markus Diesmann, Abigail Morrison, Philip H. Goodman, Frederick C. Harris, Jr., Milind Zirpe, Thomas Natschläger, Dejan Pecevski, Bard Ermentrout, Mikael Djurfeldt, An-

- ders Lansner, Olivier Rochel, Thierry Vieville, Eilif Muller, Andrew P. Davison, Sami El Boustani, and Alain Destexhe. Simulation of networks of spiking neurons: A review of tools and strategies. *J Comput Neurosci*, Vol. 23, No. 3, pp. 349–398, 2007.
- [77] Esin Yavuz, James Turner, and Thomas Nowotny. Genn: a code generation framework for accelerated brain simulations. *Scientific Report*, Vol. 6, p. 18854, 2016.
- [78] Rittner P and Cleland TA. Myriad: a transparently parallel gpu-based simulator for densely integrated biophysical models. In *Society for Neuroscience Abstract*, 2014.
- [79] Hodgkin AL and Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, Vol. 117, No. 4, pp. 500—544, 1952.
- [80] HPCI 戦略プログラム分野 1「予測する生命科学・医療及び創薬基盤」. 課題 3 予測医療に向けた階層統合シミュレーション—循環器系および筋骨格系・神経系の階層統合シミュレーション—, 2011. http://www.scls.riken.jp/research/03_integration/index.html.
- [81] 文部科学省科学研究費・新学術領域研究(2016～2020年度). 人工知能と脳科学の対照と融合, 2016.
- [82] Sten Grillner, Nancy Ip, Christof Koch, Walter Koroshetz, Hideyuki Okano, Miri Polachek, Mu ming Poo, and Terrence J Sejnowski. Worldwide initiatives to advance brain research. *Nature Neuroscience*, Vol. 19, No. 9, pp. 1118–1122, 2016.
- [83] Terry Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical report, 1971.
- [84] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [85] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, Vol. 65, No. 6, pp. 386–408, 1958.
- [86] Herbert Gelernter. Realization of a geometry theorem proving machine. In *International Federation for Information Processing Congress*, pp. 273–281, 1959.
- [87] Raymond Kurzweil. *The Singularity Is Near: When Humans Transcend Biology*. Viking, 2006.
- [88] Bruce Buchanan, Georgia Sutherland, and Edward A Feigenbaum. *Heuristic DENDRAL: a program for generating explanatory hypotheses in organic chemistry*. Stanford University, 1968.
- [89] Edward H Shortliffe and Bruce G Buchanan. A model of inexact reasoning in medicine. *Mathematical biosciences*, Vol. 23, No. 3, pp. 351–379, 1975.
- [90] Thomas R Gruber. A translation approach to portable ontology specifications. *Knowledge acquisition*, Vol. 5, No. 2, pp. 199–220, 1993.
- [91] David Silver, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, Vol. 529, pp. 484–503,

2016.

- [92] Kunihiko Fukushima and Sei Miyake. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, Vol. 15, No. 6, pp. 455–469, 1982.
- [93] Thomas Serre, Lior Wolf, Stanley Bileschi, Maximilian Riesenhuber, and Tomaso Poggio. Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 29, No. 3, pp. 411–426, 2007.
- [94] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, Vol. 323, pp. 533–536, 1986.
- [95] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *International Conference on Machine Learning*, pp. 807–814, 2010.
- [96] Abdel-rahman Mohamed, George Dahl, and Geoffrey Hinton. Deep belief networks for phone recognition. In *NIPS Workshop on Deep Learning for Speech Recognition and Related Applications*, 2009.
- [97] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [98] Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf, and Kurt Keutzer. Firecaffe: Near-linear acceleration of deep neural network training on compute clusters. In *Computer Vision and Pattern Recognition*, pp. 2592–2600, 2016.
- [99] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.
- [100] Frank Seide, Gang Li, and Dong Yu. Conversational speech transcription using context-dependent deep neural networks. pp. 437–440, 2011.
- [101] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Computer Vision and Pattern Recognition*, pp. 1–9, 2015.
- [102] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

2.6 創薬・医療

2.6.1 分野の概要

—健康で長寿な社会を目指して—

日本は世界に先んじて超高齢化社会を迎えており、国民の健康維持・増進による健康長寿社会の実現はきわめて重要な国家的課題である。健康長寿社会に資する創薬・医療技術の開発には、その基盤として人体における生命現象の理解と疾患発症メカニズムの解明が不可欠であるが、生命現象はあまりに多くの要素が絡み合って複雑に関係している現象であり、まさに、今後の HPC での計算能力の飛躍的増大が有効に活用される分野であると言える。

画期的創薬・医療技術の創出には、生命科学分野の強力な進展を推し進める必要がある。生命現象の複雑さ故に、生命科学は経験的色彩の濃い分野であったが、HPC の進展により、経験的方法論から、モデル化、シミュレーションといった演繹的方法論への転換が図られようとしており、まさに時代の転換点にさしかかっていると言える。生命現象のモデル化においては、原子・分子レベルのミクロなスケールから、細胞・臓器・脳・全身スケールに至るまで、時空間スケールが非常に幅広く、しかも各スケールが階層的に密接に関係し合っているという特徴がある。そこで、HPC の発展とともに、幅広いスケールを統合的に扱おうとするマルチスケール・マルチフィジックスシミュレーション法（幅広い時空間にまたがる対象に対し、それぞれの階層での計算を連成させるシミュレーション法）の開発が進展し、新たな分野を切り開きつつある。マルチスケール・マルチフィジックスシミュレーションでは、従来、別な分野で独立に進められてきたさまざまなシミュレーションを統合することが求められる。分子レベルでは、ナノ分野で展開されてきた第一原理計算等の高精度の方法が生命科学分野に続々と応用され、大きな成果が期待される。また、細胞、細胞内小器官、ウイルス全体といった生命体の高次のレベルにまで、原子・分子のモデルを拡張したシミュレーション計算が可能になってくると思われる。また、そのような原子・分子レベルから、細胞レベル、臓器・全身スケールに至るまで、階層間をなめらかにつないでいくマルチスケール・マルチフィジックスシミュレーション技術を発展させることで、生命体の全体像が見えてくるであろう。

創薬分野コミュニティからの意見として、他の実験研究施設との連携についての指摘があった。創薬における分子レベルの計算では、薬の標的となるタンパク質など生体分子の立体構造が重要な役割を果たしているため、それを決定する放射光施設などの大型実験研究施設の連携を十分に考慮しつつ、進展させていくべきだという意見がでた。更に、今後発展が期待される、X 線自由電子レーザー（波の位相が揃ったレーザーの性質を持つ超高輝度の X 線を発生することのできる光源）の研究施設である SACLAC（国家基幹技術として兵庫県播磨に建設された X 線自由電子レーザー装置）や、国内の主要大学・研究機関に配備されている Cryo 電子顕微鏡装置や中性子回折装置とも十分に連携していくべきという意見が出た。更には、今後のスーパーコンピュータで想定される、膨大な CPU コア数や計算を加速するアクセラレータを活用する計算方法やソフトウェアの開発の強化についての意見も出た。

医療分野コミュニティからは、医学系の分野では、HPC になじみがある人はまだ少数であり、今後とも HPC 分野と医学系分野の連携を深めていき、分野間のコミュニケーションを充実させるべきであるという意見が出された。

また、次世代研究者育成について、HPC用プログラミング方法などの教育プログラムの充実や、生命分野でHPCの技術を持つ研究者の雇用先の拡充についての意見もあった。

(2012年10月19日バイオスペーコンピューティング研究会、2013年3月10日文部科学省科学研究費補助金新学術領域研究「システム的統合に基づくがんの先端的診断、治療、予防法の開発」プロジェクトの公開講演会、2013年6月12日-14日 第13回日本蛋白質科学会年会、2013年6月27日情報計算化学生物(CBI)学会研究講演会、2016年10月26日情報計算化学生物学会2016年大会パネルディスカッション「生体分子シミュレーション実用化への道」等にて)

以上のように、今後のHPCのもたらす莫大な計算能力が、さまざまな面で生命分野の発展に大きく資するのは間違いない、ひいては画期的創薬・医療技術開発の重要な科学基盤となり得る。

2.6.2 長期目標と社会貢献

2.4で述べられているゲノム解析により疾患に関わるメカニズムが明らかになってくると、それを制御する薬をいかに開発するかという課題に直面する。通常、薬は、開発から市販まで、10年以上の時間がかかるが、近年、より複雑なターゲットへの創薬が望まれており、開発期間が長期化する傾向にあり、新薬開発のためのコストが上昇している。実際、一つの新薬開発には数百億円規模の投資が必要と言われており、技術的革新を引き起こすブレークスルーが望まれている。そのような状況下で、近年では深層学習を創薬分野に応用したAI創薬が近年期待を集めている。AI創薬の利用例として、標的蛋白質探索および、標的蛋白質に対する候補化合物探索における探索時間の短縮が求められている。標的蛋白質は数万個、医薬品の候補となる化合物は10の60乗個もの膨大な数であるため、これをHPCにより効率よく予測することが今後行われてくるようになる。また、HPCを用いた計算創薬も引き続き期待されている。例えば、次世代のHPCで想定される膨大な計算資源を用いることで、量子化学のパラメータを用いて分子動力学計算を行うQM-MD法や、蛋白質をフラグメントに分割したFMO-MD法などがより巨大な系で実現可能となり、創薬分子の設計にさらに貢献できるようになるであろう。また、分子動力学計算においては、創薬に貢献可能な細胞環境やウイルス全体を対象とする計算がさらに現実的になるほか、現実の構造変化に近いタイムスケールでシミュレーションが可能となるため、それらの構造変化を考慮した創薬分子の設計などにも寄与できると期待できるであろう。

加えて、創薬のみならず、ナノテクノロジーと生命分野の境界領域に位置するナノバイオ分野において、生命分子と関連した新しいものづくりの技術革新への貢献が想定される。具体的には、タンパク質を用いた次世代デバイスやバイオセンサー、界面活性剤分子が会合(結合)したミセル(界面活性剤の集合体)や脂質膜を用いたドラッグデリバリーシステム(患部に薬剤を届ける仕組み)や生体親和性の高いインプラント(人造の骨や歯)、細菌汚染の足場となるバイオフィルム(細菌などの微生物の増殖の足場)形成を阻害する洗浄表面などである。これらは分子間相互作用レベルでの制御がなされており、多くの生命分子と関連した次世代ものづくりへの展開が期待される。

医療応用の分野では、分子レベルのミクロなスケールから細胞・臓器・脳・全身スケールに至るまで、幅広い時空間スケールを統合的に扱うマルチスケールシミュレーションが重要な役割を果たす。例えば、心筋梗塞や脳梗塞などにおいては、血流中における血栓形成の理解が重要であり、HPCを用いた血栓形成のシミュレーション法の確立が医療応用上の重要課題として積極的に進められている。血栓症のみなら

ず、糖尿病など、血液内のさまざまなイベントを通じて引き起こされる疾患は数多く、その予測に対する社会的要請は大きい。患者個別の情報を用いた血液中での薬効評価が可能になれば、患者個人の状態に合わせた医療を創出することにつながる。この達成には、臨床場で取得される大量のデータに対する機械学習アプローチとの併用も重要な視点である[1]。臓器レベルのシミュレーションでは、心臓シミュレーションの実用化も進んでおり、“心臓再同期療法の効果予測”や“先天性心疾患の術式決定と予後予測”といった実際の医療現場へシミュレーションの技術が適用されつつある。

また、医療機器開発におけるHPCの応用も重要である。患者のQuality of Life (QOL: 生活の質) を向上させる非侵襲治療法（皮膚などに傷をつけない治療法）の開発は重要な課題であるが、その一つである超音波治療機器の開発においても、シミュレーションの果たす役割は大きい。低侵襲治療によって患者のQOLが向上すれば、患者の社会復帰が容易になり社会が活性化するだけなく、医療費の低減につながることが期待される。現在の創薬・医療において、計算機による開発支援は非常に重要な課題である。



図 2.7 創薬・医療

2.6.3 課題とその解決に必要な計算手法・アプリケーション

上記のように、本分野は構成要素が非常に多いため、条件など詳細な個別の計算機需要は関連した箇所を参照していただくとして、ここでは代表的なもののみを述べる。

創薬の分野でよく用いられる分子動力学 (MD) シミュレーション（原子間相互作用に基づいて分子運動を計算する方法）では、短い時間間隔で繰り返し原子間相互作用を計算する必要があるため、演算性能

重視でかつ高いネットワーク性能（特に低レイテンシ）が求められる。現在、300 残基から成る標的タンパク質に対して 10,000 種類の化合物の結合強度を調べると、100 PFLOPS の計算機を使えば 7 日くらいで計算が終了するが、10 EFLOPS の計算パワーがあれば、同じ時間内に、候補化合物をある程度網羅できる 100 万種類の化合物について標的タンパク質との結合強度が評価できる。ここでは、各々の化合物に対する結合強度は密に通信する並列ジョブで計算されるが、その並列ジョブは各化合物別のアレイジョブとして実行されたとした。（MP-CAFE 法 [2]）

現在、FMO 創薬技術の実用化を目指した研究開発は、産官学連携の FMO 創薬コンソーシアム（FMODD）を中心に「富岳」などの HPCI を用いて行われている。これまでに、キナーゼ、核内受容体、プロテアーゼ、タンパク質-タンパク質間相互作用（PPI）、COVID-19 関連タンパク質等の重要な創薬ターゲットに着目した検討とともに、創薬のための方法論の開発を進めており、既に数万構造のタンパク質複合体の FMO 計算を実施している。「京」から「富岳」へのもともと大きな進展の 1 つは、古典 MD シミュレーションによるトラジェクトリからのサンプリング構造（10～1,000 構造）に対して FMO 計算を行い、水中での揺らぎ構造を平均化した相互作用エネルギーを評価できるようになったことである。手法としては MM-MD/FMO 連携であり、解析法としては Statistical Interaction Analysis や Dynamically Averaged FMO などと呼ばれる。豊かな計算力を前提として、従来の X 線結晶構造やクライオ電顕構造による結晶化した（あるいは凍結した）「静的な描像」から生体環境で構造的に揺らぐ「動的な描像」の解析へと進展することができたわけである。一例では、SARS-CoV-2 のメインプロテアーゼと阻害剤の解析があり、複合体の液滴モデル構造 1,000 個に対して一連の FMO 計算を行い、結晶構造中では数か所に偏在していた相互作用が動的に平均化され、結合ポケット中のアミノ酸残基と満遍なく相互作用している「生体環境を模したリアル描像」が得られた [3]。FMO 計算の時間は、1 構造あたり「富岳」の半ラック（192 ノード）で 0.6 時間であり、同時並行実行によって僅か 5 時間で 1,000 構造が完了している。ポスト「富岳」の計算環境では、こうした MM-MD/FMO 計算はセットアップ～実行～解析までが（ほぼ）自動化される。また、ウイルスの表面タンパク質などの数千残基のタンパク質の液滴モデル（フラグメント総数は数万）も容易に実行可能となろう。

次に、現時点での結合エネルギーの予測精度について例示する。ターゲットタンパク質に対するリガンド結合性予測として、Cyclin-dependent kinase-2（CDK2）のと 6 種の化合物に対して結合自由エネルギーの実測値と FMO 予測値の比較を行ったところ、静的構造では相関係数 $R^2 = 0.75$ であったが 100 個ずつサンプリングした動的構造では $R^2 = 0.99$ まで向上しており、このアプローチによって適切に生体内の揺らぎ構造が取り入れられると期待できる（下図） [4]。本アプローチは「京」時代から取り組み始めたが、計算力的な制限で実用的とは言えなかった。実のところ、後者の CDK2 の例は、「京」、TSUBAME3.0、および Oak-forest-PACS を利用してデータの取得に数か月を要した。しかし、「富岳」ではメインプロテアーゼの例で示したように、MM-MD/FMO 連携手法は完全に実用化された。今後、こうした構造揺らぎを考慮した FMO 計算による「動的な解析」が常態化するのは間違いない。

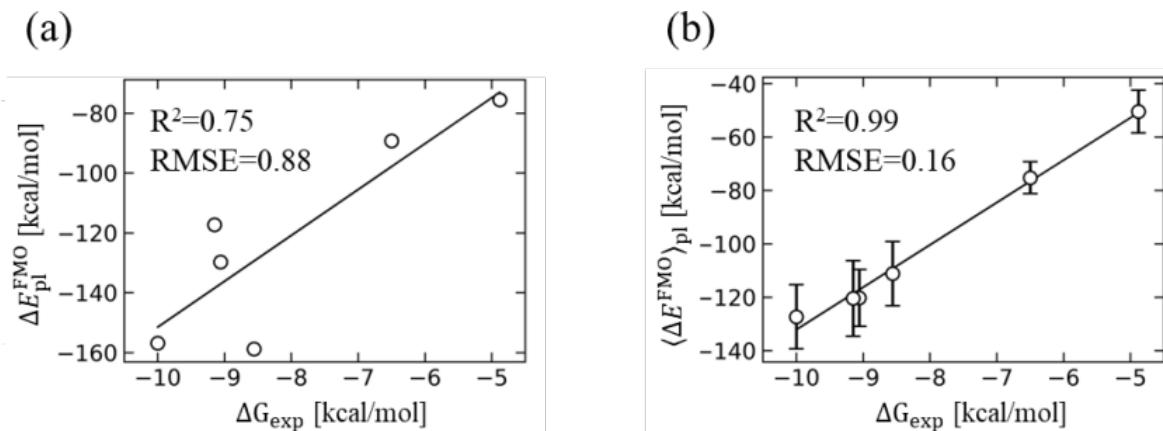


図 2.8 CDK2 と 6 化合物に対する FMO 結合エネルギー（計算値）と結合自由エネルギー（実験値）との比較。(a) X 線結晶構造に対する FMO 計算結果、(b) 古典 MD からのサンプリング構造に対して平均化された FMO 計算結果

医療応用の血流シミュレーションでは、連続体力学のステンシル計算と合わせて、必要に応じて反応方程式を用いられている。詳しくは 4.5 節を参照いただきたいが、血栓生成メカニズムを取り扱うには、長さ 100 mm、径 100 μm の血管中を流れる血球の変形挙動を取り扱い、0.1 μm の格子幅を用い、血栓生成の時間スケールである 10 秒の現象を計算するには、必要な主記憶容量は 1 PB 程度である。全演算量は 2.5×10^{23} FLOP 程度となり、実効性能が 40% とした場合、1 EFLOPS の計算機を用いると、約 174 時間の計算時間と見積もられる。ノード当たりの性能が 100 TFLOPS だとすると、理想的なキャッシュであれば B/F は 0.064 となるので、必要なメモリバンド幅は 6.4 TB/s と見積もられる。大規模な計算でモデルを精緻化することと併せ、現象の理解を促進する目的として、10 PFLOPS 程度の計算を 100 ケース程度並列させ、多くのパラメータスタディを行うことで医療に貢献するという視点もある。また、波動方程式やエネルギー方程式を FTDT で取り扱う超音波シミュレーションでは、頭部が表現できる 400 mm の立方体領域中の超音波の伝播が京の全ノードを占有した場合、実効性能が 20% 程度なため、1 日程度で計算できる。しかし、軟らかい生体組織の温度変化で生じる微妙な剛性の変化をとらえるためには、より高い解像度と長時間の時間積分が必要になり、最低でも 1000 倍の自由度の計算となり、1 EFLOPS の計算資源で 10 日程の計算時間が必要である。また、マイクロカプセルとの干渉によって生じる超音波音場の解析を 1 日で行うためには、計算コストが通常の超音波伝搬計算に比べて数十倍程度必要となるため、京の全ノードを占有した計算の 100 倍程度の計算処理速度が将来的に必要となると考えられる（詳しくは 4.6 節を参照いただきたい）。

2.6.4 ロードマップ

年代 課題	2022～	2024～	2026～	2028～	2030～	2032～
分子創薬		MD計算を用いた医薬品候補の探索 (10蛋白質×千化合物)			MD計算を用いた医薬品候補の探索 (10蛋白質×10万化合物)	
	<			><		
		ペプチド・中分子・核酸医薬品の探索・分子設計			高分子・抗体医薬品の探索・分子設計	
ゲノム医療		生体超分子複合体やウイルス環境における薬物動態シミュレーション		細胞環境における薬物動態シミュレーション		
	<		><			
		MD計算による意義不明変異の解明 (1000変異)			MD計算による意義不明変異の解明 (10万変異)	
	<			><		

図2.9 創薬・医療分野ロードマップ

2.6.5 必要な計算機資源

2032年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13節の表2.1に記載した。また、計算機資源の見積りの詳細は第4.6節に記載している。

参考文献

- [1] M. Alber, A. Buganza Tepole, W.R. Cannon, and et al. Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *npj Digit. Med.*, Vol. 2, p. 115, 2019.
- [2] Hideaki Fujitani, Yoshiaki Tanida, and Azuma Matsuura. Massively parallel computation of absolute binding free energy with well-equilibrated states. *Physical Review E*, Vol. 79, No. 2, p. 021914, 2009. PRE.
- [3] Ryo Hatada, Koji Okuwaki, Kazuki Akisawa, Yuji Mochizuki, Yuma Handa, Kaori Fukuzawa, Yuto Komeiji, Yoshio Okiyama, and Shigenori Tanaka. Statistical interaction analyses between sars-cov-2 main protease and inhibitor n3 by combining molecular dynamics simulation and fragment molecular orbital calculation. *Applied Physics Express*, Vol. 14, No. 2, p. 027003, 2021.
- [4] Kenichiro Takaba, Chiduru Watanabe, Atsushi Tokuhisa, Yoshinobu Akinaga, Biao Ma, Ryo Kanada, Mitsugu Araki, Yasushi Okuno, Yusuke Kawashima, Hirotomo Moriwaki, Norihito Kawashita, Teruki Honma, Kaori Fukuzawa, and Shigenori Tanaka. Protein–ligand binding affinity prediction of cyclin-dependent kinase-2 inhibitors by dynamically averaged fragment molec-

ular orbital-based interaction energy. *Journal of Computational Chemistry*, Vol. 43, No. 20, pp. 1362–1371, 2022.

2.7 設計・製造

2.7.1 分野の概要

これまで工業製品の開発・改良は、理論、実験ベースの設計手法で行われてきた。しかしながらその努力も限界に近づいており、今後、飛躍的な性能向上の実現には数値シミュレーションに基づく革新的設計・製造技術やそれを駆使した新たなる発想が必要である。そのため熱流体分野を始め、ものづくりの基本となる構造・騒音・振動解析、自動車などで重要な衝突解析、大規模プラントや建物・橋梁や高速道路などの耐震解析、更に生産過程、例えば塑性加工やプレス成型を再現するためのバーチャルマニュファクチャリングなど、多種多様なものづくりの設計・製造シーンでの数値シミュレーション技術の導入、すなわち CAE (Computer Aided Engineering) の普及が進んでいる。数値シミュレーションを設計・製造に適用することで、これまで未解明であった複雑な物理現象を解明し、物理メカニズムを把握したうえでの製品設計や、試行錯誤に基づいて決定していた種々の設計パラメータを理論的に最適化できる最適設計技術を活用することが可能となる。また、数値シミュレーションを用いることで、革新的な製品のアイディアやコンセプトを効果的に検証することが可能となり、これまでにない革新的な工業製品を実現することも可能となる。そして、これらを実現するためには、工業分野ごとにそれぞれの課題やブレークスルーが存在し、例えば、輸送機械などと深く関わる熱流体分野では、乱流輸送現象の解明がブレークスルーとなる。この乱流輸送現象の予測と制御は、他にも気象・気候分野、土木・建築分野、地震・津波分野、核融合分野、宇宙分野、物質科学分野などきわめて幅広い分野でも重要な課題となっている。大規模な計算科学によって乱流の基礎科学を推進、解明することはさまざまな乱流のより確かな予測と制御に向けての学際的共通基盤をより強固なものとするうえでも大いに意義のあることである。

この様に、ものづくり分野で重要性が増している数値シミュレーション技術ではあるが、これを実現するアプリケーションプログラムは、産業界が利用する場合、航空分野など一部の分野を除いて主に商用アプリケーションが用いられている。これらの商用アプリケーションソフトウェアの開発に関しては、わが国は欧米に対して大きく後れを取っている。現在、産業界で用いられている CAE 分野のアプリケーションソフトウェアは欧米製のものが大半であり、わが国で開発されたソフトウェアはほとんど用いられていない。例えば、熱流体解析では STAR-CCM や Ansys Fluent、燃焼計算では CHEMKIN、構造解析では NASTRAN、ABAQUS や LS-Dyna など欧米発の商用ソフトウェアが標準ソフトウェアとして広く使われている。

しかしながら、ここに来て新たな転機を迎えようとしている。現在のスーパーコンピュータは数万個以上の CPU や演算コアを搭載しており、その性能を発揮させるためにはこれらを同時並列的に、かつ、効率的に動作させる必要があるが、現在業界標準となっている商用アプリケーションソフトウェアの多くはこのような超並列計算には対応しておらず、これが逆に、産業界でのさらなる数値シミュレーション技術の導入を阻害している。また、ものづくりの設計・製造現場においては、計算プログラムのみならず、設計者の扱う CAD (Computer Aided Design) データから計算格子を自動的に生成する前処理プログラムや、膨大な量の計算結果から設計者が工学的に重要な情報を抽出し、その結果を設計にフィードバックすることを支援する可視化など後処理プログラムがあって初めて、ものづくりの現場でスーパーコンピュー

ティングの威力が発揮される。計算機ハードウェアの長足の進歩に伴い、アプリケーション・ソフトウェアやそれを利用した設計システムにもパラダイムシフトが起こりつつある [1]。以下、各分野を概観する。

2.7.1.1 航空機分野

ものづくりにおける低コスト化施策の費用対効果は上流に行くほど大きくなる（ものづくり白書等）。特に航空機開発では開発下流段階での航空機実機フライト試験など実機を製造し飛行試験を行わないと評価できない設計上の重要課題が取り除けておらず、設計プロセスに大きなリスクを抱えている。このことは、開発コスト増大のみならず、開発そのものを脅かす開発リスクや開発期間の大幅な遅延要素に直結している。それに伴い、民間航空機開発を推進している国内航空機開発メーカーの強いニーズとして、実機フライト試験への依存度を格段に低減する革新的設計技術により、航空機の『開発期間の短縮』、開発の『コスト低減』を図ると共に『開発リスク』を格段に低減することが望まれている。と同時に、『開発上流での最適化レベルを格段に向上』させることで、競合他機が容易にキャッチアップできない差別化技術を有した製品をタイムリーに市場提供することが望まれている。これらの革新的設計技術や差別化技術の開発は、日本の航空機産業の国際的優位性に直結する最重要課題である。

2.7.1.2 構造分野

FEM (Finite Element Method) ベースの構造解析ソルバーは、陽解法に基づく衝撃解析やほぼ非圧縮性状態の塑性加工解析のような特定の現象を解くための専用ソフトウェアも存在するものの、多様な解析機能を集積した汎用解析ソフトが一般的である。汎用解析ソフトは、機械・電気・化学・建築土木をはじめ、船舶海洋・航空宇宙・原子力・医療・環境などほとんどあらゆる分野で日常的に用いられている。また、大企業や研究所を中心に、解析モデルの大規模化詳細化と解析の高速化を目的として、PC クラスタや中規模スパコンなどの HPC 環境が導入されており、さらに HPC クラウド環境の活用も進んでいる。

一方、最新 HPC 環境での活用を前提に、超並列技術に対応した国産の構造解析ソフトの研究開発が進んでいる。1997 年～2002 年に実施された日本学術振興会未来開拓学術研究推進事業「計算科学」分野のプロジェクトにおいて開発され、2002 年に無料公開されたオープンソースの汎用並列 FEM 解析システム ADVENTURE と、それを基に開発された商用汎用構造解析ソフト ADVENTURECluster がその代表例である。ADVENTURE は 2002 年以降も我が国の計算科学に関する国家プロジェクトにおいて継続的に活用されており、ADVENTURECluster は我が国の幅広い産業分野において上述した欧米製ソフトと並んで標準的に使われるようになっている。また、科学技術庁の計算科学プロジェクトにおいて開発が開始され、順次名称が変更されながら、現在も研究開発が進められ産業界に使われている FEM ベースの並列構造解析ソフトには FrontISTR もある。エクサスケールのスパコンの登場と普及に伴い、HPC 環境で国産構造解析ソフトの研究開発がさらに進むことにより、最新 HPC 環境での構造解析は国産ソフトが牽引していくものと考えられる。

また、産業界においては、熱流体構造連成解析や流体構造連成解析などの連成解析、あるいはより一般的にはマルチフィジクス解析へのニーズが急速に高まってきており、汎用的連成解析機能を特色とする ANSYS や COMSOL などの商用解析ソフトもよく利用されている。連成解析は、流体解析の観点から語

られることが多いが、構造解析にとっても大変重要な意味を持つ。その理由は、現実の構造機器や構造物への入力荷重が、機械荷重のみならず、接する熱流体からの熱伝達・熱伝導によって生じる熱応力であったり、内部あるいは外部の複雑な流体場から受ける圧力や剪断力等の流体力などであるからである。また、先にも述べたように、大型化することにより相対的に薄肉となる機械や構造物では、構造機器や構造体の大きな変形や振動が、逆に内部あるいは外部流体場に効率低下や騒音発生等の重大な影響を与えることにもなる。航空機の翼構造や、高温流体を内包するエネルギー機器である圧力容器や配管、タービンブレード等では、このような連成解析が昔から注目されて、様々な研究開発が行われてきた。しかし、従来は対象形状が比較的単純な形状に近似できることや、変形があまり大きくならないように意図的に高剛性に設計されることから、高精度の計算力学シミュレーションを行うことよりも、半解析的なアプローチが研究開発され用いられてきた。それら以外の一般構造機器や構造体でも、従来は固く変形しにくく（高剛性に）作られることが普通であったことから、連成現象が問題となることは少なかった。しかしながら、最近の傾向として、機器や構造物の大型化や軽量化の進展によって、構造物の部材の板厚が相対的に薄く、細長くなることから、剛性が相対的に小さくなり、流体場との相互作用、すなわち連成現象がより顕在化する傾向にある。

2.7.1.3 核融合分野

持続可能な人類社会とその発展を支え得る、長期的・安定的に供給が可能な究極のエネルギー源の実現を目指し、核融合エネルギーの研究開発が世界規模で進められている。核融合エネルギーには、

- 燃料が偏在せず豊富である
- 高レベルの放射性廃棄物が生成されず環境負荷が低い
- 核融合反応は暴走せず原理的に高い安全性を有する
- 核拡散の制約を受けない

といった特徴があり、我が国にとっては国産エネルギーとなり得るばかりか、途上国を含めた地球規模での基幹エネルギー需要を長期的に代替し得る、革新的エネルギー源であると言える。従って、核融合エネルギーの実現に貢献することは、低炭素社会の実現に繋がるだけでなく、日本の新たなエネルギー戦略・産業の発展にも繋がる。磁場を用いた核融合炉では、約1億度に熱せられた高温の燃料プラズマを、トーラス状の磁場で効率的に閉じ込めることで、核融合エネルギーを取り出す。炉内では、核融合反応によって生成する高エネルギーのアルファ粒子によって燃料プラズマを自己加熱し、高温状態を維持しながら、炉心プラズマの核融合反応を持続させる。この核融合炉開発に関連するシミュレーション分野は、プラズマ、熱流体、中性子、材料と多岐にわたっているが、この中でもプラズマシミュレーションは最大規模の計算量を必要とする。これは、核融合プラズマの第一原理モデルが、異方性を持った強い磁場でプラズマが閉じ込められていることを利用して、粒子速度空間を3次元から2次元に簡約した5次元位相空間（3次元実空間+2次元速度空間）上の粒子分布関数の時間発展を記述する運動論モデル（ボルツマン方程式）で与えられているに基づく。ジャイロ運動論モデル[?]と呼ばれるこのモデルが、3次元実空間のみで記述される従来の流体モデルに比べて2次元速度空間の分だけ自由度が増大し、～1002倍大きい自由度を必要とするためである。プラズマシミュレーションが対象とする物理問題として、プラズマ乱流

に起因する熱輸送や燃料粒子輸送、あるいは、磁気流体现象に起因する高エネルギー粒子輸送が挙げられるが、このような物理現象は、高温プラズマの閉じ込め性能を左右するため、核融合炉の成立性に直結する最重要課題の一つとなっている。富岳やその他の大型計算機と数値シミュレーション技術の進歩によって、このようなシミュレーションの適用対象は、核融合プラズマ物理の理論的研究のみならず、実験データ解析、乱流輸送モデル開発、さらには、炉心プラズマ設計など、具体的な核融合炉を見据えた幅広い展開を見せるようになってきている [2, 3]。

これまでに JT-60U（日本）および JET（欧州）の 2 つの実験装置で、入力エネルギーと核融合出力が等しくなる臨界条件を達成しており、現在は、核融合炉の科学的・工学的可能性の実証を目的として、次代の実験装置である国際熱核融合実験炉（ITER）の建設が、日米欧露中韓印の世界 7 極の国際協力の下に、フランスのカダラッシュで進められている。2025 年 12 月に運転開始を予定している ITER では、人類初となる長時間の核燃焼プラズマ実験が計画されており、これまでの核燃焼のない模擬実験とは質的に異なる炉心プラズマが実現する。このため、核燃焼プラズマの性能予測や実験条件最適化において、計算機シミュレーションが大きな役割を果たすと期待されている。一方、計算機科学および計算科学から核融合研究を眺めると、欧米のフラッグシップ計算機において、核融合分野は継続的に主要な位置を占め続けており、特に米国においては、2016 年より開始された Exascale Computing Project (ECP) の主要テーマの一つに採択されている。米国 ECP では、核融合炉のまるごと解析 (Whole Device Modeling) を目指して、プラズマから材料までを包含するマルチスケール・マルチフィジックスの野心的なシミュレーションプロジェクトが進められている。国内においては、ポスト京重点課題⑥「革新的クリーンエネルギーの実用化」サブ課題 D 「核融合炉の炉心設計」（2015 年度～2019 年度）において、ITER の炉心プラズマ運転条件最適化に向けたエクサスケール核融合プラズマシミュレーションの開発と「京」を活用した開発コードの実証研究が進められた。これを受け 2020 年度からは、スーパーコンピュータ「富岳」成果創出加速プログラムの採択研究課題「核燃焼プラズマ閉じ込め物理の開拓」において、プラズマ乱流による熱・粒子・運動量と核融合反応で生成される高エネルギー粒子の輸送過程について、「富岳」を用いた超大規模の位相空間マルチスケール・シミュレーションの実現とデータ科学手法との融合により、その機構解明への取り組みが進められている。このように、国内外でエクサスケール核融合プラズマシミュレーションに向けた研究開発が展開している。

2.7.1.4 可視化・データ処理分野

可視化は、大量のデータから視覚的な情報を手掛かりに効率よく新しい科学的知見や洞察を得るために重要な基盤技術であり、自然科学分野、特に計算科学分野を中心に発展し、今では理学や工学だけでなく生命科学や社会科学など様々な学問分野に広く浸透している。現在、計算科学分野では、富岳コンピュータの登場によって幕を開けたエクサスケール時代の到来により、計算機環境の高効率・高性能化が進み、数値計算技術のさらなる高度化も相まって、時間当たりに出力されるデータ量がより一層大規模化している。また、近年では、ビッグデータ解析や AI の活用により、複雑化する科学的・社会的課題の解決のためにデータ科学と計算科学の融合に対する要求が高まっており、富岳に代表されるエクサスケール時代のハイパフォーマンスコンピューティング (HPC) 環境上で処理されるデータの多様化が進んでいくことが予想される。このような多種多様な大規模データから、人間が解釈できる科学的知見を獲得するための

可視化および分析技術は必要不可欠な技術であり、その重要性がますます高まっている。

2.7.2 長期目標と社会貢献

ものづくり分野で HPC が実際の製品の研究開発や設計の現場で活用されるためには、企業の研究者や設計者が容易にアクセスでき、また、長時間にわたり一定のリソースを占有できる計算機環境があることが必須となる。現在のスーパーコンピュータのコア性能は数十 GFLOPS のオーダーであり、コア数は数百万程度に達している。計算機の従来の価格性能比トレンドから推定すると、企業の研究者や設計者は世界最速のスーパーコンピュータの 1/10 程度の性能を持つ計算機を数年後に自由に使えるようになるものと予想される。したがって、2025 年頃には、数十万程度のコアを使用した超並列計算が先端的企業において、その有効性が実証されているものと推定される。ものづくり分野において、特に大規模数値シミュレーションは物理現象の解明のみならず、多数の設計評価を行うことに価値がある。この分野で扱う物理現象は、マルチスケール、マルチフィジックスを考慮する必要があり、対象とする物理現象を解明するためには第一原理的な高精度解析が必要となり、おのずと大規模解析が必要となる。一方、設計・製造においては単に解析ができるだけではなく、これらの解析が実際の設計で利用できることが重要となる。そのためには数百から数万ケースにも及ぶ計算を同時並行的に実施し、数値解析に基づく本格的な設計探査が実用化されものと期待されている（一般に大量解析と呼ばれるスーパーコンピュータの利用形態の一つである）。これを実現するためには更に多くの計算リソースが必要となると同時に、実用的に問題となる多くの目的関数がある場合にも適用し得る設計探査方法の研究開発と、計算効率の抜本的な向上による strong scaling（問題規模を一定にして並列数を大きくして計算速度の向上を図る）時の計算時間を飛躍的に短縮するための研究開発などが必要となる。また、HPC 環境で得られた高精度シミュレーションの結果に、人工知能 (Artificial Intelligence, AI) などのデータサイエンス的手法を適用することで、代理モデルや縮約モデルを構築し、シミュレーションの高速化を試みる研究開発も必要である。

従来から、工業製品の研究開発や設計・製造に数値シミュレーションは幅広く用いられていた。しかし、従来の計算機で予測できる現象や予測精度には限界があり、複雑なミクロ現象の解明や製品試作の数値シミュレーションによる代替は実現されていなかった。しかし、長足の進歩を遂げているスーパーコンピュータを駆使すれば、製品試作の一部あるいは全部をシミュレーションにより代替したり、試行錯誤に基づいて決定していた種々の設計パラメータを理論的に最適化したりすることが可能となる。

2.7.2.1 航空機分野

自動車などと比較して、航空機はレイノルズ数が 10 の 7 乗オーダーと非常に高いことからレイノルズ数条件を一致させた風洞試験による評価は困難であり、航空機開発における上記課題に対し、高度な計算科学（数値シミュレーション科学）による予測評価が期待されている。また計算科学による予測評価を用いて、実機を製造する前の設計初期段階で設計上の重要課題を評価可能とすることで、実機フライト試験を代替する新たな設計プロセス技術が強く望まれている。しかしながら現状、計算科学により予測評価できるのは、航空機周りの流れが定常的な巡航条件付近に限られており、離着陸性能を決める低速・高揚力条件や高速飛行性能を決める高速飛行時の遷音速バフェット（衝撃波による振動現象）境界の予測評価、

フリッター（流体と機体構造の連成現象）境界の予測評価、離着陸時の騒音の予測評価など、流れが強い非定常性を持つ条件に対する計算科学による予測評価は困難なのが現状である。

このような現状を打破するために、高度な数値シミュレーション科学、具体的には解析対象を準直接的に扱うことで流れの非定常性を高忠実に再現する数値シミュレーション手法（LES: Large-Eddy Simulation）の実用化が期待されている、航空機の飛行条件（フライトエンベロープ：航空機の飛行可能な速度や荷重や高度の範囲）に対して、上述した様に、これまで極一部の領域・現象のみしか予測評価が行えなかつたが、LES を用いることで、フライトエンベロープ全体への予測可能範囲の拡大が期待されている。またこのような期待は日本国内のみならず、世界的にも着目され、世界各国で熾烈な研究開発競争が進められている。具体的には、アメリカ航空宇宙局（NASA）は、「CFD Vision 2030: A Path to Revolutionary Computational Aerosciences」という 2030 に航空宇宙分野に革新を起こすビジョンのレポートをまとめており、高度な数値シミュレーション科学による航空機フライトエンベロープの予測可能範囲の拡大が 6 つの国家的プロジェクトの 1 つにあげられており、2022 年 6 月には本課題に関する国際ワークショップが開催される。加えて、航空機分野を含むより広い航空宇宙分野では、上述したような流れの非定常性を準直接的に扱う数値シミュレーションの基盤アプリが世界的に見ても確立されていないのが現状であり、航空宇宙分野の広い範囲において、学術研究から産業応用・実用までの使用に耐えうる高忠実な HPC 基盤アプリの構築も世界的な関心事項となっている。さらに航空機飛行条件の予測範囲の拡大が達成された先には、フリッター等の流体と構造の連成問題や、流体と機体運動の連成問題など、フライトエンベロープで起こりうる様々なシナリオに応えられる多角的な機体設計への取り組みも必要となる。

2.7.2.2 構造分野

構造解析は、流体解析と比べて解析モデルの規模は相対的に小さめであるが、流体解析よりも悪条件になりがちの係数マトリックスを効率的に解く反復法ソルバーが必須となる。さらに、パラメトリック解析や最適化等のために多数ケースの解析が必要となることから、エクサスケールのマシンを構造解析に活用するニーズは、Capability Computing よりも Capacity Computing にシフトしている。この結果、最新 HPC 環境をフル活用することにより生じる膨大な解析結果を適切かつ効率的に処理し、日々の設計や運用業務に活かす仕組みや支援ツールの開発が必須となる。

また、最近のニーズとして、実機と同程度の精度・時間応答をする解析モデル（デジタルツイン）を構築しておき、構造機器や構造物の運用の適正化を使うというニーズも増えつつあり、ここでも繰り返しの高速解析が求められる。このような産業界の最先端の現場ニーズについては、最新 HPC 環境の計算性能を最大限に活用して高精度解析を行えるだけでは不十分であり、そうした解析を短時間に繰り返し行えることが必要である。そのための方向性として、最新 HPC 環境での高精度シミュレーション結果を活用しつつ低次元での高精度モデル（サロゲートモデルあるいは ROM (Reduced Order Model) と呼ばれる）の構築法の研究開発がある。その代表例に AI を活用するデータサイエンス的アプローチや、POD（特異直交分解）などを活用する数理的なアプローチがある。

2.7.2.3 核融合分野

核融合プラズマの第一原理シミュレーションは、1990年代に米国で基本的な手法が確立し、それ以降、計算機性能の進展と共に急速に発展してきた。核融合プラズマでは、炉心サイズが～1m規模であるのに対し、乱流の最小スケールが、背景磁場中でサイクロトロン運動する荷電粒子の軌道幅程度となるため（イオン～5mm、電子～0.1mm）、直接数値計算（DNS）が主に用いられている。プラズマの数値計算法としては、歴史的に大きく2つの手法、すなわち、粒子法（Particle-In-Cell）と格子法（Eulerian、Semi-Lagrangian）が1970年代に開発されていた。1990年代後半に初めて実現した第一世代のトーラスプラズマ・シミュレーションでは、比較的少ないメモリでも計算が可能な粒子法が主流であった。第一世代シミュレーションの最大の成果は帶状流と呼ばれる乱流が励起するメゾスケールの流れ（平均流）が乱流輸送を抑制する効果を示したことであり、この成果は、ITER設計の妥当性を巡る議論に大きな影響を与えた。2000年代前半に開発された第二世代のプラズマシミュレーションでは、計算機性能、特にメモリ容量の増大によって5次元位相空間を格子で分解する見通しが得られ、格子法によるトーラスプラズマ・シミュレーションが開発された。格子法の計算は、粒子法で問題となる粒子ノイズの影響が無く、計算精度が安定している。このため、実験研究者等のシミュレーション研究者以外にも広く普及し、系統的な実験解析等に第一原理シミュレーションが活用されるようになっている。このような格子法のシミュレーションは、1本の磁力線の近傍をモデル化した局所モデルによって開発されたが、その後のさらなる計算機性能の進展に伴い、炉心全体を計算領域とする大域的モデル、あるいは、イオンスケールから電子スケールまでを局所的に取り扱うマルチスケール・モデルが開発されている。さらに、2000年代後半にはペタスケール規模の計算が視野に入り、より長時間スケールの数値実験を目的とする第三世代のプラズマシミュレーションが開発されるようになった。第三世代の大きな特徴は、第一・第二世代の計算が、プラズマ分布関数 f を背景プラズマ分布 f_0 と揺動成分 δf に分解したとき ($f = f_0 + \delta f$)、定常状態にある揺動成分のみを計算する δf モデルを採用していたのに対し、プラズマ分布と乱流成分と同じ第一原理モデルで同時に時間発展させ、非定常プラズマの過渡的変化を計算する full- f モデルが開発されたことである。これにより、 δf モデルで発見された帶状流に加えて、自己無撞着に発展するプラズマ分布と乱流の相互作用という新たなフィードバックループを含む数値実験が実現し、炉心出力に影響するプラズマ分布を第一原理モデルで計算することが可能になった。このような第三世代プラズマシミュレーションの単一イオン系乱流計算（～10ms程度の時間スケール）の計算量を図4.23に示す。ITER規模の準定常乱流計算を行うには10PFlops程度の計算が必要となる。このシミュレーションを多種イオンで構成される核燃焼プラズマに対応させるためには、さらに一桁以上大きな計算規模が要求される。これまでの実験は、水素もしくは重水素のみの模擬実験であったのに対し、ITERの核燃焼実験は燃料イオン（重水素、三重水素）、および、核融合反応で生成するヘリウムイオンを含む多種イオン系である。2025年に開始されるITERの実験で次々に蓄積される核燃焼プラズマの知見を炉心プラズマシミュレーションに反映させていくためにも多種イオン系でのシミュレーションは重要である。また、燃料粒子輸送や電子熱輸送をもたらす電子系乱流を計算するにはイオンに比べて数十倍速い電子の運動を追跡するため、核燃焼プラズマのシミュレーションはさらに計算コストが増大する。さらには、磁場の揺動までを含めた電磁乱流シミュレーションも視野に入れなければいけない。他にも、炉心プラズマだけでなく、周辺領域のプラズマをター

ゲットに据えていた別コードとのカップリングを含め、これらを第四世代として位置づけ、今後のエクサスケール規模の計算によって、ITER のエネルギー閉じ込め時間である数秒程度まで計算規模が拡張し、新たな炉心制御手法を数値実験で開発できるようになる。以上で述べた物理対象や計算規模に対する大規模化・高精度化に加え、これらの炉心プラズマ輸送の大規模シミュレーションに対して、その計算結果を機械学習やデータ科学を利用して高速かつ簡便に予測する手法開発も強く求められている。特に、燃焼プラズマの挙動に対して、大規模シミュレーションと同等の精度で、輸送レベルや温度・密度分布の予測が高速に行えるようにしておくことは、炉心性能の制御手法としても十分に確立しておくべき必須の課題である。近年、核融合研究分野においても、機械学習やデータ科学的手法の応用が進んでおり、国内でも精力的に研究が行われている。例えば、プラズマ乱流シミュレーションのデータを学習させ、乱流の波数空間パターンに基づき、プラズマの非線形飽和時間や輸送レベルの予測を行う研究や、データ科学的手法により乱流シミュレーションを実行すべき入力パラメータ領域をピンポイントで推測し実行する手法の開発[4]が進められている。

2.7.2.4 可視化・データ処理分野

計算科学分野における可視化は、SciVis として可視化黎明期から続く重要な研究分野の一つであり、その研究成果の多くが ParaView や VisIt といった汎用可視化ツールに実装されている。また、基本的な可視化技術は、数値計算処理のポスト処理機能の一部として、すでに多くのソルバーにも組み込まれている。単にデータを画像化するための基盤技術としての可視化は、計算科学研究者に対して広く提供されており、効率的な利活用のフェーズとなっている。しかし、昨今の HPC 環境の著しい発展に伴いデータの大規模化に伴い新たな問題が発生している。データの大規模化に伴い可視化の問題は、以前から存在しており、並列レンダリング技術や分散データ処理技術など様々な解決手段がすでに提供されている。しかし、シミュレーション計算終了後のデータ入出力や、可視化端末へのデータ転送が困難な場合、データを可視化することができない状況が発生する。この場合、高精度に計算された結果であるにも関わらず、可視化のためにデータの空間的・時間的間引きが必要となり、場合によっては知見獲得につながる重要な現象が欠落する可能性がある。近年、可視化研究分野では、このような問題を解決するための手法として、シミュレーション計算を行いながら、同時に可視化まで行う in-situ 可視化[5]に関する研究が注目されている。In-situ 可視化は、シミュレーションから可視化まで想定した計算において、時空間解像度の向上と膨大な計算能力の活用が可能であることから、新しい計算パラダイムとなることが期待されている。初期の研究開発においては、HPC 環境におけるシミュレーション処理と可視化処理の効率的な連携[5]について活発に議論され、その一部が汎用可視化ソフトウェアの ParaView[6] や VisIt[7] などに組み込まれている[8, 9]。また、GPU を前提として実装されたレンダリング手法を CPU 環境下で実行するための基盤ソフトウェア[10]も整備され、レイトレーシングによる高品位レンダリング基盤[11, 12]の In-situ 可視化への活用も進んでいる。さらに、ParaView や VisIt の可視化基盤ソフトウェアである VTK[13]のデータモデルを採用し、マルチコア・メニーコア向けに最適化された可視化基盤 VTK-m[14]の開発も進められている。その他、In-situ 可視化向けの統一的なデータモデルと API を提供する SENSEI[15] や、シミュレーションと可視化の効率的なデータ連携を実現する基盤である ADIOS[16]、基本可視化機能を含む軽量な In-situ 可視化基盤である Ascent[17] など、さまざまな In-situ 可視化ソフトウェアがオープ

ンソースで公開されている。以上のように、基盤環境としての In-situ 可視化は、多くの HPC 環境でその整備が進められており、応用のフェーズに入っている状況であるが、エクサスケール時代に突入するにあたり、2018 年にドイツのダグスツールで開催されたワークショップにおいて計算科学向け In-situ 可視化の未解決問題についてまとめられている [18]。以下に、それらの課題を簡単にまとめて列挙する。

1. データの質と圧縮： 事後のデータ探索（データ解析）処理に影響しない、In-situ 処理における質の高いデータ削減・圧縮手法
2. ワークフロー仕様： 様々な分野への In-situ 処理の適用において、想定されるタスクの高度な抽象化と再利用性の高いワークフロー仕様
3. ワークフロー実行： ワークフローを構成する複雑なタスクを、動的に変化するリソースを考慮して柔軟かつ高効率に実行するための基盤環境
4. エクサスケールシステム： 計算速度とディスクストレージ帯域幅のギャップが拡大するエクサスケールシステムにおいて、In-situ 可視化の更なる高度化とその性能評価
5. アルゴリズム： エコシステム化が進む In-situ 環境下において、制限されたリソース内での異種データ連携を前提した並列度の高い新しい可視化および解析アルゴリズム
6. 新しいユースケース： 単一の大規模シミュレーション向けの探索的分析だけでなく、不確実性を定量的に評価するための複数の計算を実行するアンサンブル計算、計測データを取り込み計算するデータ同化、アルゴリズムの収束やシステムの性能などの別データの分析など、In-situ 可視化が可能にする新しいユースケース
7. エクサスケールデータ： エクサスケールシステムでのシミュレーションによって生成される大規模・複雑データ向けに、シミュレーション・可視化・解析のすべてにおいて最適化されたデータモデルとそれを効率的かつ安定的に処理するための実行基盤
8. 性能評価モデル： In-situ 処理全体の性能最適化を目的とした、アルゴリズム実行性能の予測モデル構築とその検証
9. HPC とビッグデータの融合： In-situ 可視化と分析のための HPC とビッグデータの融合
10. ソフトウェアの複雑性： 利用者が In-situ 処理を導入する際に障壁となるソフトウェアの複雑さや異機種混在環境となる計算資源への対処

2.7.3 課題とその解決に必要な計算手法・アプリケーション

単にアプリケーションプログラムだけあっても、製品設計・製造に活かすことはできない。超大規模な数値シミュレーションをものづくりの設計・製造に活かすためには、設計データを用いて迅速に計算を実行することを可能にする前処理プロセスや、計算結果を効率的に設計にフィードバックするための後処理プロセス、更に、設計の最適化を支援する機能などを有した次世代の CAE システムを構築する必要がある。近い将来、格子点数は数十兆点規模になるものと予想されるが、このような大規模計算のための前後処理を従来の考え方の延長で実現することは不可能である。つまり、前後処理を含めた解析システムとしてのパラダイムシフトが必要となり、これを実現するための研究開発が必要である。今後の計算機の発達を考えると、さまざまな分野で大規模、高精度な解析が可能となることが期待されるが、計算機能力の向

上に任せておけば物理モデルの高度化が不要となるわけではない。複雑な物理現象、例えば流体であればキャビテーション流れや燃焼流れのように、乱流の渦スケールよりもはるかに小さな空間スケール、時間スケールを持ったマルチスケール現象に関しては、計算機性能が向上するだけでは分子スケールの運動を直接解析することは不可能であり、これらの物理現象のより定量的な予測のためには物理モデルの更なる高度化が必要である。また、流体・構造・音響現象などのマルチフィジックス現象に関しては、それぞれの現象の解析手法は確立されているが、連成現象の高精度で効率的な数値解析手法の研究開発も必要である。更に、これから計算機アーキテクチャに合致したシミュレーションの方法論やそれを具現化する計算アルゴリズムの研究開発も欠かしてはならない。また、大規模な連立一次方程式に対する、前処理も含めた高収束性で低コストの計算アルゴリズムの研究開発や輻射計算などで必須となる密行列の超並列解析アルゴリズムなどの研究開発も必要となる。

2.7.3.1 航空機分野

現在推進されている「富岳」成果創出加速プログラム「航空機フライト試験を代替する近未来型設計技術の先導的実証研究」では、これまで予測評価が困難であった実際の飛行レイノルズ数（10の7乗オーダー）における離着陸性能を決める低速・高揚力条件の予測評価が可能となってきているが、これは航空機飛行条件のある一端に留まる。次世代に解決されるべき課題は、高度な計算科学による予測可能範囲を航空機の全飛行領域、すなわち航空機フライトエンベロープ全体に拡大することである。それを実現するためには、特に未だ解決されていない重要課題、高速飛行性能を決める高速飛行条件における遷音速バフェット境界ならびにフラッター境界の予測評価、更には離着陸時の機体騒音の予測評価が不可欠となる。これらの課題を解決することにより、高度な計算科学が実機フライト試験の代替となり、デジタルトランスフォーメーションによる航空機設計、すなわち開発下流段階での実機フライト試験など実機を製造ならびに飛行試験を行わないと評価できない設計課題を事前に評価することで、開発リスクの根本的な解決が期待できる。さらには、航空機フライトエンベロープ全体の予測評価が可能となった先には、航空機の飛行条件全体において性能や安全性に関する要求を満足するべく、これらを統合的に評価する多目的・多設計変数の機体設計が期待される。

上記の課題の解決を実現するためのアプリケーションに必要な学術的課題は大きく以下の4つとなる。
① 航空機の細かい複雑形状ならびにそれらの形状が移動・変形する場合にも対応可能な解析、② 高レイノルズ数条件の準直接解析手法、③ 衝撃波を含む圧縮性流れの高忠実解析を実現する計算スキーム、および④ 多目的・多設計変数最適化問題で大域的最適解を高速に探索するアルゴリズムである。高忠実・圧縮性流体準直接解析基盤アプリ FFVHC-ACE では、①に対しては形状 STL データのみで細かい複雑形状に対しても自動かつ高速に格子生成を可能とする階層型等間隔直交格子法を採用している。また今後の開発項目となるが、物体境界との高速な交差判定を可能とすることで移動・変形を含む高速な解析を可能とすることを想定する。②では高レイノルズ数条件の準直接解析を可能とする壁面モデル LES (壁面近傍内層乱流のモデル化) を独自に開発し、実装している。物体の移動・変形に関しては今後の学術的開発項目となるが、物体境界での保存性を向上させることで、既存のシャープインターフェイス法で問題となる圧力振動問題を解決することを想定する。③ではこれまでの圧縮性流体解析の計算手法で問題となっていた数値散逸を用いずとも安定かつ高忠実な解析を可能とする計算手法 (KEEP スキーム) を用

いる。本手法は、有限差分法ベースで対流項の物理現象を高忠実に再現することで「安定性」と「非散逸性」を同時に満足する計算手法である。衝撃波のある流れへの適用が今後の学術的開発項目となる。④では設計変数（形状パラメータ）に対する目的関数（機体性能や安全性）のブラックボックスな構造を限られた評価データから代数式近似し、近似モデルベースで最適解の在処を瞬時に探索する。また今後の開発項目となるが、機械学習や圧縮センシングなどにより設計変数空間を次元縮約することで、多設計変数問題に対する近似モデルの適合性と正確性を上げることを想定する。

2.7.3.2 構造分野

大規模構造解析のニーズとしては、すでに産業界の広い分野において様々なものが顕在化しつつあるが、一例として、原子力プラントの耐震解析を取り上げる。巨大構造物の地震応答現象では、地盤、建屋、機器などのすべてのものが一度に揺すられ、それらは少なからず連成するので、各コンポーネントを個別に解くことは解析精度という観点から必ずしも適切ではない。原子力プラントでは、鉄筋コンクリート壁、配管、圧力容器、燃料集合体などすべてがソリッド要素で表現される。解析規模としては、原子炉圧力容器単体でも十数億自由度を要し、配管や発電機器を含むプラント全体と付近の地盤を合わせて数十億自由度と見積もられる。また、解析は弾塑性の動解析であり波動伝搬と比べて比較的ゆっくりと揺すられることから、時間方向および非線形計算の繰り返しからなる数万から数十万ステップの計算が必要となる。

ペタスケールコンピュータをもってしても、自由度や計算ステップに関して各々これらの10分の1程度のものを扱うのが限界となっており、現時点では、圧力容器や少数の配管群など、特定箇所にフォーカスした解析を行うか、あるいはアドホックなモデル化や線形解析に頼らざるを得ない。今後のHPC技術の進展により、プラント全体の統合計算が可能になるのみならず、地球物理・大気海洋分野で行われている地震シミュレーションや津波シミュレーションとの連携により、巨大構造物の地震や津波応答現象のより厳密かつ高精度なシミュレーションが可能となる。

次に、巨大な構造体を含む連成解析について述べる。これまであまり顕在化されていなかった構造体の例として、巨大な発電用風車、特に今後我が国でも大規模に設置が進む洋上風力発電施設を取り上げる。洋上風力発電施設は、発電効率向上のために、広大なサイトに巨大な風車(10MW以上)を多数(十数基から数十基)建設して運用される。個々の風車は、高さは東京都庁(243m)を超え、重量は大型航空機(ボーイング777-200は191トン)を超え、巨大なローター(直径は150~200m)が日本近海の厳しい風況・海況環境の中で20年超の長期にわたって日夜回転する動的機器であり、かつ台風、地震や波浪等の複雑で厳しい荷重の影響を受けるという、これまでのものづくりでは経験したことのない過酷な環境で稼働する人工物となる。しかも、風車を通過した乱れた流れ(後流[wake])を介して巨大風車同士が相互に影響し、発電効率低下や疲労寿命低下などの悪影響を受けることになる。このため、①サイト選定、設計・建設上の課題克服はもとより、運用開始後もアクセスが制限される、②点検が洋上のため工夫が必要、③故障や破壊事故が起こった時の原因究明・修理・取り換え等の対応策、④海況や風況などの公的データの利活用、が必要不可欠となる。したがって、日々の自然環境変化や機器のモニタリングデータの取得・活用や、最新シミュレーション技術に基づくデジタルツインの利活用など、あらゆる場面でデータ駆動型の技術革新を進めていくことが必須となる。

もう一つの大規模な連成解析事例として、次世代火力発電の機器の解析がある。発電や水素製造に用いる大型の石炭ガス化炉は、日本では 2013 年に初めて商用運転が開始された新しい技術であり、必ずしもガス化炉内部の現象を正確に理解した上で最適な設計がなされたものではない。ガス化炉内部は、微粉砕された石炭（微粉炭）が乱流中を分散するとともに、熱分解による揮発分放出、揮発分の気相反応、および固体炭素分のガス化（固気）反応が同時に進行する極めて複雑な乱流反応場である。また石炭中の灰分は高温に晒されて溶融し、炉壁や炉底に溶融スラグ流れを形成する。すなわち、ガス化炉内は高压容器に囲まれた固気液三相流となっているため、内部の現象を計測し、正確に把握するのは極めて困難である。さらに、炉内の状況は用いる石炭の種類にも大きく影響を受ける。加えて、高温高压に晒される炉容器構造体についても、実用化されて間もないことから、その健全性や耐久性についての知見もほとんどない。一方、石炭ガス化は微粉炭燃焼と現象が似ているものの、その LES 研究の数は限られている。これは、そもそもチャー粒子の固気反応過程において複数の活性化学種が混在しモデル化の観点から格段に複雑になることに加えて、圧力容器内の流れ場に対する計測が極めて困難で直接観察による現象理解が進まないことや、高圧化により乱流渦スケールが微細化し、より高い空間解像度が要求されること等によると考えられる。また、数値シミュレーションを実機設計に活用するためには、反応容器内の燃焼特性を評価することに加えて、炉容器構造体の健全性や耐久性についても評価する必要がある。本質的に非定常で複雑な現象である燃焼と構造の連成解析を実現した例は極めて限られている。

2.7.3.3 核融合分野

超高温の核融合プラズマは、

- 粒子衝突効果による散逸が非常に小さい（無衝突プラズマ），
- 乱流場の振幅が非常に小さい（平衡分布の 1% 程度）
- 電磁場の計算が電荷や電流の保存精度に敏感

という特徴がある。このため、主要な計算手法においては数値的な散逸や保存誤差を低減しつつ数値的な安定性を保証するための工夫が施されている。粒子法においては摂動分布関数 δf のみを粒子で表現することによってノイズを低減する δf -PIC 法が主流となっているのに対し、格子法のプラズマシミュレーションにはオイラー法（差分法、有限体積法など）、セミラグランジュ法（CIP 法など）、スペクトル法、あるいは、それらを組み合わせたコードが開発されてきている。また、高性能な閉じ込めプラズマの実現には、炉壁周辺領域におけるプラズマの密度・温度分布の急峻な勾配形成が必須となる。これは、プラズマ閉じ込め状態の遷移現象として知られており、この現象の解明には、装置中心部での 1 億度以上の超高温プラズマとともに、温度が数十分の 1 となる周辺プラズマを同時に取り扱うことが必須となる。この問題は、プラズマ乱流による構造形成の問題だけでなく、ITER の性能予測やダイバータへの熱負荷予測において取り組むべき重要課題である。このようなシミュレーションの実現には、中性粒子との相互作用を含む周辺・ダイバータ領域のプラズマ物理モデルと中心領域を扱う運動論的シミュレーションとのハイブリッド解析手法の確立が必須である。一方で、機械学習やデータ科学などに基づき、大規模プラズマ乱流シミュレーションの効率的な実行と豊富で巨大な計算データの有効的な解析手法の開発も求められている。

2.7.3.4 可視化・データ処理分野

国内の計算科学分野ではまだ十分な議論が行われていないが、HPC環境を想定した可視化とその関連技術が想定する課題は、ダグツールレポート[18]に挙げた10の未解決問題に集約されている。ここでは、国内外の可視化研究の動向、特に、国内の計算科学向け可視化研究の現状を踏まえて、優先して取り組むべき課題を以下にまとめる。

- In-situ 可視化の高度化
- 大規模データ向け視覚的解析環境の構築
- 可視化による HPC システムの高機能・高効率化

2.7.4 ロードマップ

年代 課題	2026～	2028～	2030～	2032～	2034～	2036～
航空機	100～1000億格子規模の壁面モデルLES解析 (遷音速バフェット、騒音解析)					

← →

--	--	--	--	--	--	--

← →

--	--	--	--	--	--	--

← →

--	--	--	--	--	--	--

図 2.10 設計・製造分野ロードマップ(航空機分野)

年代 課題	2026～	2028～	2030～	2032～	2034～	2036～
構造解析	大規模アセンブリ構造物の丸ごと解析技術の開発 大規模アセンブリ構造物の丸ごと連成解析技術の開発 多数の大規模アセンブリ構造物の丸ごと解析結果から 高精度低次元解析モデルを構築する技術の開発					

← →

--	--	--	--	--	--	--

← →

--	--	--	--	--	--	--

図 2.11 設計・製造分野ロードマップ(構造解析分野)

年代 課題	2016 ~	2020 ~	2024 ~	2028 ~	2032 ~	
核融合 ・コード開発	ITER規模の準定常乱流計算 (~10ms) マルチスケール電子系乱流計算 核燃焼プラズマシミュレーション	磁場擾動を含む電磁的モデル拡張 周辺・ダイバータプラズマモデルとのハイブリッド解析				
・実証・実用化 研究	高温プラズマ実験に対する実証実験 核燃焼プラズマの性能予測研究 ITER規模の過渡乱流輸送解析 (~1s)					
・実験計画／ 炉設計ロード マップ	大型ヘリカル装置 (LHD, ~2023) JT-60SA (日欧) ITER (日米欧韓中印中露) 原型炉設計					

図 2.12 設計・製造分野ロードマップ (核融合分野)

年代 課題	2026 ~	2028 ~	2030 ~	2032 ~	2034 ~	2036 ~
可視化 ・In-situ可視化 の高度化	基盤構築 多視点In-situ可視化					
・視覚的分析 環境の構築	データモデルと ワークフロー開発 In-situ分析 In-situ学習					
・HPCシステム の高機能・高 効率化	データベース化 ログデータ解析 対話的解析 資源利用の高度最適化					

図 2.13 設計・製造分野ロードマップ (可視化・データ処理分野)

2.7.5 必要な計算機資源

2032年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13節の表2.1、及び、表2.2に記載した。また、計算機資源の見積りの詳細は第4.7節に記載している。

参考文献

- [1] 加藤千幸. ものづくりと次世代スーパーコンピュータ－期待と課題－. 日本の科学者, Vol. 45, No. 12, pp. 28–33, 2010.
- [2] X Garbet, Y Idomura, L Villard, and T H Watanabe. Gyrokinetic simulations of turbulent transport. *Nucl. Fusion*, Vol. 50, No. 4, p. 043002, April 2010.
- [3] 洲鎌英雄, 渡邊智彦. ジャイロ運動論による磁化プラズマ乱流の研究：核融合から宇宙まで（解説）. 日本物理学会誌, Vol. 68, No. 5, pp. 296–304, May 2013.
- [4] M Nunami, S Toda, M Nakata, and H Sugama. Improved prediction scheme for turbulent transport by combining machine learning and first-principle simulation.
- [5] Hank Childs, Kwan-Liu Ma, Hongfeng Yu, Brad Whitlock, Jeremy Meredith, Jean Favre, and Scott Klasky. In situ processing. In E W Betherl, H Childs, and C Hansen, editors, *High Performance Visualization: Enabling Extreme Scale Scientific Insight*, pp. 171–198. Chapman and Hall, 2012.
- [6] Kitware, Inc. ParaView. <https://www.paraview.org/>. Accessed: 2023-12-3.
- [7] LLNL. VisIt. <https://sd.llnl.gov/simulation/computer-codes/visit>. Accessed: 2023-12-3.
- [8] Utkarsh Ayachit, Andrew Bauer, Berk Geveci, Patrick O’Leary, Kenneth Moreland, Nathan Fabian, and Jeffrey Mauldin. ParaView catalyst: Enabling in situ data analysis and visualization. In *Proceedings of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization*, ISAV2015, pp. 25–29, New York, NY, USA, November 2015. Association for Computing Machinery.
- [9] Brad Whitlock, Jean M Favre, and Jeremy S Meredith. Parallel in situ coupling of simulation with a fully featured visualization system. In *Proceedings of the 11th Eurographics conference on Parallel Graphics and Visualization*, EGPGV ’11, pp. 101–109, Goslar, DEU, April 2011. Eurographics Association.
- [10] Mesa Project. The mesa 3D graphics library. <https://www.mesa3d.org/>. Accessed: 2023-12-3.
- [11] Ingo Wald, Sven Woop, Carsten Benthin, Gregory S Johnson, and Manfred Ernst. Embree: a kernel framework for efficient CPU ray tracing. *ACM Trans. Graph.*, Vol. 33, No. 4, pp. 1–8, July 2014.
- [12] I Wald, G P Johnson, J Amstutz, C Brownlee, A Knoll, J Jeffers, J Günther, and P Navratil.

- OSPRay - a CPU ray tracing framework for scientific visualization. *IEEE Trans. Vis. Comput. Graph.*, Vol. 23, No. 1, pp. 931–940, January 2017.
- [13] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. Pearson Education, Incorporated, 2006.
- [14] Kenneth Moreland, Christopher Sewell, William Usher, Li-Ta Lo, Jeremy Meredith, David Pugmire, James Kress, Hendrik Schroots, Kwan-Liu Ma, Hank Childs, Matthew Larsen, Chun-Ming Chen, Robert Maynard, and Berk Geveci. VTK-m: Accelerating the visualization toolkit for massively threaded architectures. *IEEE Comput. Graph. Appl.*, Vol. 36, No. 3, pp. 48–58, 2016.
- [15] Utkarsh Ayachit, Brad Whitlock, Matthew Wolf, Burlen Loring, Berk Geveci, David Lonie, and E Wes Bethel. The SENSEI generic in situ interface. In *2016 Second Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV)*, pp. 40–44. IEEE, November 2016.
- [16] David A Boyuka, Sriram Lakshminarasimham, Xiaocheng Zou, Zhenhuan Gong, John Jenkins, Eric R Schendel, Norbert Podhorszki, Qing Liu, Scott Klasky, and Nagiza F Samatova. Transparent in situ data transformations in ADIOS. In *2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 256–266. IEEE, May 2014.
- [17] Matthew Larsen, James Ahrens, Utkarsh Ayachit, Eric Brugger, Hank Childs, Berk Geveci, and Cyrus Harrison. The ALPINE in situ infrastructure: Ascending from the ashes of strawman. In *Proceedings of the In Situ Infrastructures on Enabling Extreme-Scale Analysis and Visualization*, ISAV’17, pp. 42–46, New York, NY, USA, November 2017. Association for Computing Machinery.
- [18] Janine C Bennett, Hank Childs, Christoph Garth, and Bernd Hentschel. In situ visualization for computational science (dagstuhl seminar 18271). 2019.

2.8 社会科学

本節は、「計算科学ロードマップ 2017」の2.7節を転記したものである。現在、2030年頃における計算科学が解決すべき社会的課題・期待される科学的ブレークスルー、そのために必要となる計算機能性について執筆中であり、近々更新予定である。

2.8.1 分野の概要

社会科学・社会工学分野では古くからコンピュータが活用されてきた。19世紀末に現れたセンサデータの取得・分析のためのパンチカードシステムが今日のコンピュータのルーツの1つであることはよく知られており、その後もコンピュータは通信技術の発達と相俟って、今日に続く「ビッグデータ科学」に基づく社会科学を先導し続けている。他にも経済政策の立案・遂行、金融取引、交通管制、経営、会計などで古くから活用されており、選挙運動や戦争でもコンピュータシミュレーションは不可欠なものであった。そして現代にいたっては社会生活のすみずみまでコンピュータが使われている。さらに今後、コンピュータシミュレーションにより社会現象をより精緻に予測し制御することに対して人々の期待が大きい反面、不安と戸惑いも見過ごすことはできない。本節では、社会科学分野でのコンピュータシミュレーション、特にスーパーコンピュータの成長を想定したシミュレーションを中心に展望する。

社会現象を記述する社会科学的モデルは、自然現象とは異なり、記述する側とされる側との同一性が高い。そして、個々の構成要素である人や集団自身がおののの行動を決定することができる。自然現象の構成要素、例えば原子・分子や流体の運動・ふるまいは一定の法則に従っており変化することはないことは対照的である。このため社会科学的モデルは、このような行動の元でどのような社会になるかという自然科学的モデルと同様のベクトルのみならず、望む社会にするために個々がどのように行動すべきかという逆のベクトルも有する。このことは、モデル構築の上で普遍的に信頼できる原理、いわゆる第一原理の欠如を意味するかもしれない。しかし現在の社会は、個々人の気まぐれでは基本的には大変動はしない。社会全体のふるまいを個々が予測し、個々の行動を修正した結果、個々の行動の変化も織り込まれた（くりこまれた）社会となっていると考えることもできる。そこで現実の現象を再現するもっとも簡潔なモデルをつくり、これに基づいてシミュレートしするという物理科学の手法を使うことが期待される。

こうした研究は社会物理学と呼ばれ、19世紀前半に始まる分野であるが、特に近年のコンピュータの成長に伴って発展が著しい。1980年代以降は人流・自動車交通流の、90年代以降は経済の、そして2000年代以降、社会関係の物理科学的研究が進んでいる。この3分野こそは現在、社会シミュレーション研究の柱である。こうした研究の結果、種々の社会現象の普遍性・個別性についての分類が深まり、合わせて社会を構成する個々の要素から普遍性・個別性を再現する統計物理学的エージェントモデルの開発が進められている。しかしながら、社会現象は典型的な複雑・複合系であり、ある局面で確立したモデルが異なる状況では有効性が限られることが多い。このため、実社会データのマイニング・現象の特徴付け・モデルの開発・シミュレーションの応用などの研究開発の諸段階を、目標ごとに同時進行することになる。逆に個々のふるまいによりどのような社会ができるかをシミュレーションにより予測することにより、社会経済制度や政策・倫理方針・宗教指針の設計・制御を行うことも期待される。そもそも、現在の社会がど

のようなものであるかをわかりやすく提示する、すなわち「可視化」すること自体、データマイニングやシミュレーションの役割として認知されている。

1つの例として自動車交通のモデルについて挙げよう。1990年代半ば以降の研究により、1車線の1本道で発生する渋滞の原因が解明され、個々の自動車の運動に基づいたエージェントモデルにより再現・予測に途をつけた。しかしながら、分岐路や道路ネットワークにおける交通ダイナミクスの解明には未だ至っておらず、目的ごとに模索する状況が続いている。さらに自動車交通の目的は、関係する当事者により異なることが普通である。個々の自動車は、目的地に速く、あるいは低燃費で、あるいは予定時刻どおりに到着したいと考える一方、道路のある地域社会・警察は事故や騒音・大気汚染を無くしたいと考えるであろう。また、道端のお店はそこを通る自動車の数を多くしたいと考えるであろう。これらの諸目的は同時に達成されることは限らないし、おそらく互いに相入れない目的であることが多いであろう。すなわち自動車交通という社会的な問題は、単一の最適化問題として解決することはできないと推測される。どの立場での問題解決かに割り切るか、あるいは関係当事者すべてが同じ評価関数を共有するような交通ルール・交通経済を設計するところまで立ち戻る覚悟をする必要がある。同じ評価関数を共有することを目指している自由市場経済において、好ましからざる均衡点を回避する「見えざる神の手」の実現にもエージェントシミュレーションが活用されよう。

このような社会現象をシミュレートするモデルには、シミュレーションパラメータによって複雑多岐のふるまいがつきものである。またパラメータの数も多くなりがちで、現象を再現するパラメータの同定 자체、困難であることは日常的である。数個程度の少数パラメータであれば経験や網羅的シミュレーションで同定することもできようが、数十個程度以上のパラメータ、それも解析的ではない離散的なパラメータとなると難しい。例えば、人口100万程度の都市の道路網には数万か所の交差点があるが、経路探索や交差点での信号制御や右折ルールの設定を網羅的に探索することは不可能である。そこで何らかの知能的探索が必要となり、当面はいわゆる人工知能的な手法の活用が課題となろう。個々のパラメータに対するふるまいをシミュレートするにもそれなりの計算能力が必要となるため、こうしたパラメータ探索研究でもスーパーコンピュータが潜在的には競争力の源となると期待される。

社会現象のシミュレーションでは、自然現象のシミュレーションと比べて特別な性能をコンピュータに要求するわけではない。あえて言えば、道路ほかの地図、人間関係や企業間の取引、お金の流れといった社会的な関係を表わすグラフ構造の扱いが特徴である。例えば、自動車が走行する道路は交差点や終点をノードとし、ノード間を結ぶ道路をエッジとするグラフとして表現することができる。自動車交通シミュレーションはグラフ上に自動車のエージェントモデルを移動させることであり、走行経路はグラフ上の2点間の経路探索問題として表現される。また、道路の新設・廃止は道路を表すグラフの運動・変化として表現される。もう1例、マクロ経済・企業間取引のシミュレーションでは、企業間の取引関係を表すグラフの上でお金や原料・製品・サービスが移動し、また企業間競争・起業・倒産や災害・事故などに伴い取引関係を表すグラフも変化する。このように社会シミュレーションの多くは、グラフ上の運動・変化とグラフ自身の運動・変化として実行される。すなわちグラフシミュレーションなのである。

自然科学シミュレーションもグラフシミュレーションであることが多い。粒子・分子動力学法も含む離散要素法、有限要素法、格子模型などで扱う構成要素間の相互作用も一種のグラフシミュレーションである。こうした問題で扱うグラフは、ランダムグラフや完全グラフ、格子構造などの比較的単純なグラフである場合が多い。一方、社会シミュレーションでは、スモールワールド、スケールフリーといった複雑な

グラフを扱うことが多いことが特徴である。このためデータベースやファイルシステム研究とも共通する問題を含んでいる。

とはいって一般には、個々の実行時間自体は、社会シミュレーションは自然科学シミュレーションと同程度かむしろ短い問題を扱うことが多い。動力学シミュレーションでは1兆粒子を必要とすることがあるが、人間は現状、100億人もいないし、自動車の数はもっと少ないのである。しかし先に記述したように、最適化関数の多義性や離散的なパラメータ依存性に由来するパラメータ探索の必要性から、個々の課題の解決に必要なシミュレーション数が相当に多くなることが特徴である。この特徴は、物質設計・創薬・ものづくりほかとも共通するものであり、連携も期待される。

2.8.2 長期目標と社会貢献

社会科学の長期的目標は、ミクロからマクロまで、刹那から長期間までさまざまな社会現象を解明することである。応用面から簡潔に言うならば安心・安全で幸福な社会の実現ということになるが、これは社会科学に限らず、あらゆる科学的研究・学術研究、さらにはあらゆる人間活動の重要かつ根源的な目標でもあるこうした高邁な目標が掲げる一方、なにをもって安心・安全・幸福とするかの基準も多様であり、社会の構成要素の間で互いに矛盾することも多い。地域・時代によっても変化する。このため諸研究分野では各分野での基礎・応用研究を、その時々の社会が想定している安心・安全・幸福な社会の実現とのバランスを取りながら進めている状況といえよう。

社会科学には、諸分野と同様に安心・安全・幸福を所与の前提とするという面がある一方、諸分野とは異なる面もある。それは社会自身を設計・制御の対象とするため、なにをもって安心・安全・幸福とするか、あるいはどのような安心・安全・幸福を実現するかを決めるという面である。特に社会シミュレーションでは、構成要素や構成要件が変わった時に社会全体と個々人の生活とがどのように変化するかを予測することが期待される。これはシミュレーションモデルに依存した一面的なものではあるが、長期的にはモデルの改良により高度な予測ができるようになると期待される。例えば、経済政策の選択によりGDPがどのように推移するか、道路整備により物流がどのように変化するか、新しい製品が市場をどのように変えるか、新しい技術が人々の生活をどのように変えるかといったように、諸分野での研究成果や行動決定がどのような社会的影響を及ぼすかを予測し可視化することが期待される。また社会規則や政治制度を変更すると、社会がどのように変化するかを明らかとすることが期待される。構成要素スケールから社会全体をシミュレートする技術は、ひとりひとりの立場・状況から社会全体の姿までをシームレスに解像するわけであるから、ひとりひとりに寄り添いながらの全体展望を実現する。これには社会がはらむ矛盾や利害対立が先鋭化するリスクがある一方、先鋭化した議論や戦いからその解決策・妥協策の実現までも社会シミュレーションを活用することになろう。社会事象のはらむリスクを社会シミュレーションにより精査したり、微調整では到達できない変革を実現する駆動力として、戦争・革命・クーデターに置き換わることさえも長期的には想像されよう。

この長期目標に向けた研究開発は、現在、以下の3つに大別されよう：

- (1) 社会システムを構成する1サブシステムのシミュレーション技術の開発と応用
- (2) 社会システムの複数のサブシステムのシミュレーション技術の開発と応用
- (3) シミュレーションモデルの評価およびパラメータ探索技術の開発と応用

社会全体は、種々のサブシステムが密接に連携して機能している。まずもっとも基本的なサブシステムのシミュレーション技術の開発・応用が(1)の段階である。たとえば自動車交通では自動車が道路と交通ルールとに則って移動するというシステムが基本であり、こうした移動をシミュレートする技術の開発・応用がこの段階に当たる。市場経済では取引ルールに則った自由な売買が基本であり、こうした取引のシミュレーション技術である。

しかし移動・物流は自動車交通だけでは成立せず、他に歩行・鉄道・船舶・航空というサブシステムと連携する必要がある。また、通信・電力・ガス・水道他のサブシステムとも密接にかかわっている。これらのサブシステムのシミュレーション技術の開発・応用も進める必要がある。もう一つの例に挙げた経済も取引だけでは成立せず、生産活動と金融システムほかのサブシステムとの連携が不可欠である。さらに移動・物流と経済との連携も不可欠である。(2)の段階は、(1)で開発したサブシステムシミュレーションを連携してさらに大きな括りでのサブシステムシミュレーションを実現する技術の開発・応用である。複数のサブシステムを連携する際には時間的・空間的あるいは社会的なスケールが異なるサブシステムを連携させる方法を開発する必要があり、また必要となる計算資源も大きく増大するため、第2段階の研究課題とした。

サブシステムを連携し社会全体に向かう際にもう一つ問題となるのが、増えるモデルパラメータをどのように統御するかという課題である。数十パラメータですらそのすべてを計算機の能力で尽くすことは、一般には難しい。社会シミュレーションでは、(1)の段階のサブシステムですら莫大な数のパラメータを持ち、試すべきシナリオが想定される。モデルのふるまい自体、解析的でないどころか連続的でならないことが多い。これが(3)の課題である。そのために社会現象・社会シミュレーションのためのデータマイニング・データ同化技術や人工知能技術の開発も必要となる。また単にモデルパラメータを探索して統御するだけでなく、社会の円滑な運営と非常事態への安心・安全な対応とを実現できるような社会の実現も期待される。

これら3段階は、必ずしもステップバイステップというわけではなく、まず(1)に着手し、次に(2)、(3)にもというように進行している。京コンピュータの段階で既に、最も基本的なサブシステムと考えられた自動車交通・鉄道交通・歩行者集団・市場取引・金融システム・企業間取引・人間関係などの研究がある程度進展している。そして現在、ポスト京コンピュータを見据えてこれらのサブシステムの連携およびモデル評価の研究が始まっていると同時に、社会を構成する他のサブシステムに対する(1)の段階の研究も進んでいる。たとえば、航空交通である。さらに社会のサブシステムではあるが、既にそれぞれの分野での研究が進んでいる電力システム・通信システムほかのシミュレーションや、社会に大きくかかわる気候・環境・地震・津波・噴火などのシミュレーションとの連携も段階(2)の課題となろう。

2.8.3 課題とその解決に必要な計算手法・アプリケーション

上では長期目標・社会貢献の実現に向けた社会シミュレーションの研究開発の3段階を掲げた。これに沿って現在進められている具体的な課題として、交通・経済・社会関係の3種類のサブシステムとその連携とを以下に挙げる。

1. 交通の設計・制御および最適化

これまでに研究が進んでいる自動車交通、歩行者集団のシミュレーションをベースとして、種々の交通モード・交通政策をも加味した交通シミュレーションを実現し、現実の交通・物流への応用が期待される。まずここでは交通・物流は平時の社会を念頭に置いているが、災害はテロリズム・ミサイル攻撃や軍事進攻といった非常時における避難や物流のシミュレーションの実現と応用も期待される。前者では交通システムの最適設計や効率運営が、後者では交通インフラが部分的に破壊された際の最適輸送が主な課題となる。

こうした研究をさらに進め、交通と経済活動との連携シミュレーションが期待される。平時の効率化と非常時での緊急避難とが問題となる点では同様である。が、さらに費用対効果を経済活動の観点から、たとえば GDP インデックスによる定量化へと途を拓くと期待され、地域行政や政策判断への応用につながる可能性が期待される。

交通シミュレーションでは、個々の自動車や歩行者の動きを模したエージェントモデルを道路や街路・建物の内部などを模したシミュレーションスペース内に配置する。個々のエージェントモデルとしては、シミュレートする現象で本質的となる集団運動を再現する限り、できるだけ簡単なルールに従うものを使う。

例えば自動車であれば、先を走る自動車に衝突しないこと、交通ルールを守ること、および経路を決定することが必要最小限である。その際、運転手の操作や自動車が走る機械工学的・流体力学的シミュレーションは現段階では必要ではない。前を走る自動車との間隔と制限速度とから、数演算からせいぜい数十演算の簡単な関数で走行速度を決めて移動させるシミュレーションである。シミュレーションスペースである道路ネットワークは、交差点をノード、交差点を結ぶ道路セグメントをエッジとするグラフとして表現される。各エッジは車線構造を持ち、自動車は各車線上、時には車線変更をしながら走る。各ノードは、つながっている各エッジの各車線についてどこからどこに走行できるのかの構造を持つ。日本全土に道路は 130 万 Km 程度あり、600 万程度の交差点と 800 万程度の道路セグメントを持つグラフとして表現できる。日本の登録自動車数は 8000 万台程度である。また全世界の道路は 3,000 万 Km 程度、交差点 8,000 万程度、道路セグメント 1 億程度で、10 億台程度の自動車があると見積もられる。すべての自動車が同時に走行しているわけではないので、仮に 10 日本全体では実時間の十分の一程度、全世界でも実時間程度の経過時間で実行できる。これは極簡単な自動車エージェントモデルを使った場合の結果ではあるが、目安となる。

歩行者による群衆シミュレーションでは、歩道に列をなして歩くエージェントモデルから広い空間を自由に歩くモデルまで、目的に応じて計算量の異なるモデルが使われる。前者は上述の自動車シミュレーションに近く、計算量も少ない。各ステップで各エージェントの移動に数十演算程度である。後者ではより多くのエージェントとの相互作用を扱うため、前者に比べると演算量は数倍から数十倍程度多くなる。整然と避難する場合は前者に近く、スクランブル交差点やショッピングモール・公園・広場では後者に近いモデルを使う。

2. 経済の安定的成長の実現

自由市場経済では種々のものが自由に売買され、これを通して資源の最適配分と適正な生産活動が維持されるように取引ルールが決められている。とはいっても千分の一秒単位の為替取引から 10 年単位の資源開発・工場建設まで優に 12 枠の時間スケールや、1 円単位の商品から 1 兆円単位のプロ

ジェクトまでやはり 12 枠の金額スケールがどのように連携しているのかは経験則の域を出ない。現に株式取引に伴う株価の変動は 1980 年代までは素朴にガウス分布として扱われていたが、1990 年代のビッグデータ解析からまったく異質なべき分布であることが明らかとなった。と同時に、ブラックマンデー、リーマンショックといった大暴落も絶えず、個々の株価が他の株価と相関をもって変動することも明らかである。

こうした株価のダイナミクスをシミュレーションにより解明しようとする研究がすすめられた結果、株価の連鎖変動、ショックの連動といったふるまいが再現されている。また株式取引のルールが取引に与える影響についての研究も進められている。こうしたシミュレーションでは、株式を売買するエージェントモデル多数にモデル市場で取引をさせる。売買が成立するためには、同じ価格で売るエージェントと買うエージェントとがいないといけないため、数種類のエージェントを混在させる必要がある。基本的なエージェントは、株価の上下につられて売買するトレンドフォロワー、株価基準を持って高ければ売り安ければ買うファンダメンタリスト、売りが溜まれば買い、買いが溜まれば売るマーケットメイカーである。各エージェントの売買の意思決定自体は数十演算程度であるが、複数銘柄取引ではポートフォリオ管理をするエージェントを使うため最適ポートフォリオ計算のために演算量は桁違いに多くなる。目的により数百エージェントから数万エージェントあるいはさらに多くのエージェントをモデル市場で取引させることになる。例えば日本証券取引所では 1500 銘柄程度を扱うが、この規模を年単位相当でシミュレートするには京コンピュータ程度の性能が必要となる。

現在の株式取引のシステム自体にも連鎖安ショックを起こすダイナミクスがあることが明らかとなってきた。一方、連鎖安は株式を多く保有する金融機関が構築する世界規模の金融ネットワークのリスク管理とも関係していると考えられる。国際金融機関の自己資本比率を一定以上に規制するバーゼル規制のため、保有株の価格が下がった場合に株を売って規制を満たさなくなるリスクを回避するとする行動が起こると考えられる。また自己資本が尽きてしまい破綻する金融が現れると、金融ネットワークを介して破綻リスクが拡大する。こうした現象を念頭に、金融市場取引シミュレーションを実現し、安全な金融制度の設計および金融リスク予測制御への応用が進められている。

同様のシミュレーションは取引先が倒産することによる連鎖倒産リスクが企業間取引ネットワーク上に拡大する現象にも応用できる。災害などにより企業の倒産や減産が生じると、その企業にとどまらず取引企業にも影響が拡大する。この様子も企業間取引ネットワークを使ってシミュレートし、個々の経済イベントにより GDP がどれだけ増減するかを推定しようとする、金融と GDP の連携シミュレーションも進められている。

企業活動を担っている労働者と生産物を最終的に消費する消費者の動向は、内需経済が大きい日本では GDP の動向に大きな影響を与えると考えられる。地域社会の高齢化や過疎化により労働者人口や消費者人口が減少した場合の影響が懸念される現在、経済成長への影響評価は重要である。これを実現する企業活動シミュレーションと GDP シミュレーションおよびそれらの政策・地域運営への応用も期待される。

3. 社会の安定と安心・安全の維持

個々人はどのような人間関係・社会集団を形成するかは社会構造の基礎といえる。家族・血縁からはじまり、周りの人々と関係を構築してゆく様子は社会心理学分野で長い研究があり、いわゆるスマートワールドネットワーク構造の発見という大きな成果も挙げられてきた。今世紀に入り、携帯電話の通話記録やフェイスブック・ツイッターなどのSNSのデータマイニングから、人間関係とそのダイナミクスとに関わる実証データが増大し、研究が一段と深化している。これまでの研究と実証データとを合わせて、人間関係のダイナミクスを記述するモデルの開発も進んでいる。例えば人間関係のモデルの嚆矢となったクンプラモデルは、人は単にだれかと友人となるだけでなく、相手の友人も親しくなるという確率的なグラフダイナミクスモデルである。これを個々人の社会的な属性に拡張したものなど、種々のモデル開発が進められ、社会関係のシミュレーションの実現が期待されている。こうしたシミュレーションは、例えばマーケティングや情報伝達他、多様な応用が期待されている。

現代社会では人間集団の多様性の増進が活力の源泉であるとも考えられており、多様さの重要性は常々指摘されている。しかしここでいう重要な多様さとは何を意味するのか、例えば乱雑さとは違うのかも明確ではなかった。そもそも多様性の意義自体、生物の生態系進化の研究によるところが大きいため、この課題解決の糸口は生物進化や生態系のダイナミクスにあるものと考えられていた。これに対してはダーウィンの進化論以来、多くの議論と研究があり、特に遺伝子としてのDNAが発見されて以来、分子生物学・遺伝子工学分野での研究も相俟って精力的な研究が続けられていたが、簡潔な答えには至らなかった。この問題に対し、古生物学の研究のビッグデータ解析を契機として、今世紀に入って答えが出始めている。生物種をノードとし、生態系での種間の関係をエッジとするグラフのダイナミクスモデルとして、生態系進化の統計的特徴が引き延ばされた指数関数 $\exp(-\sqrt{t/\tau})$ であるとすることが提唱されている。このモデルを使ったシミュレーションは、多様性の維持と創出をシミュレートすることができ、環境保全への応用が期待される。さらにこの引き延ばされた指数関数はコンビニエンスストアの品揃えのビッグデータ解析からも発見されており、多様性の実態として社会・経済への応用も期待されている。

社会関係のシミュレーション技術は始まったばかりで、先の交通・経済と比べて技術展望に未知の部分が多く、具体的なロードマップは今後の研究に待たなくてはならない。とはいえ交通・経済と社会関係の連携シミュレーションは、社会シミュレーションを単なる社会のサブシステムのシミュレーションから、現実の社会の複雑さに匹敵するシミュレーションへと進化させる可能性をもつとも考えられる。

2.8.4 ロードマップ

2.8.5 必要な計算機資源

先述したように社会シミュレーションでは、問題解決のために多数のシミュレーションを実行しなくてはならないことが多い。このため必要な計算機資源を、キャパシティコンピューティングとケイパビリティコンピューティングとの2軸で図にまとめたものを、交通・経済・避難シミュレーションについて示

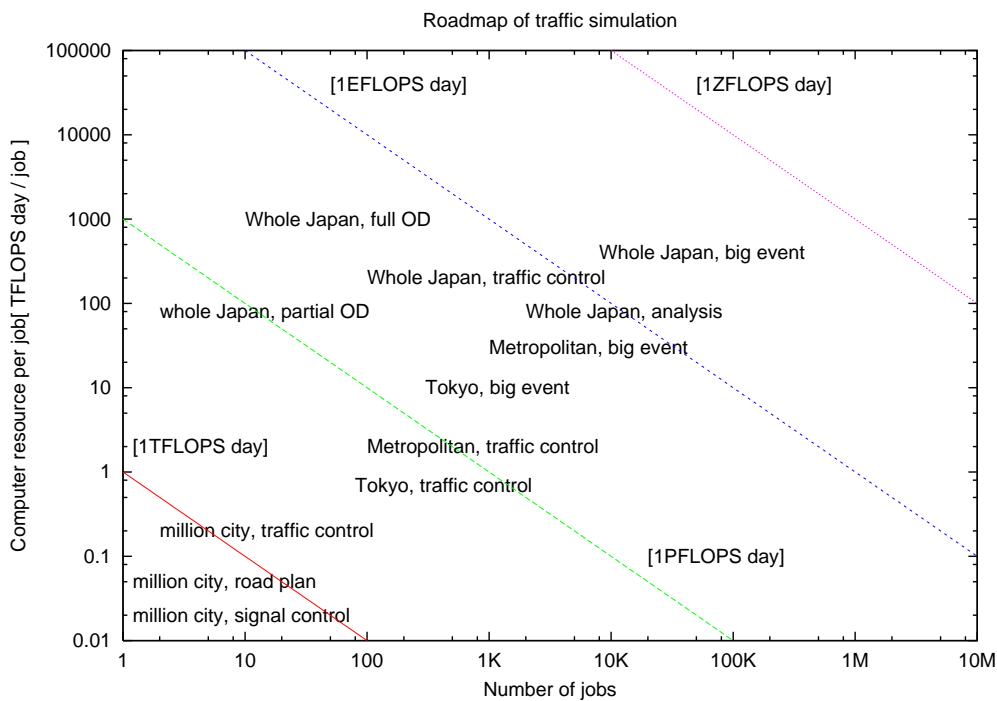


図 2.14 交通課題解決のために必要な推定計算量。

す^{*3}。いずれも横軸に問題解決のために必要と予想されるジョブ数を、縦軸に1ジョブあたりの計算量をTFLOPS・日を単位として示した両対数図である。図中の右下がりの直線が、計算量が等しい問題を表す。図2.14が交通シミュレーション、図2.15が経済、図2.16が避難である。

^{*3} これらの図は Itsuki Noda, Nobuyasu Ito, Tomohisa Yamashita, Hideki Mizuta, Tomio Kamada, Yohsuke Murase, Sachiko Yoshihama and Hiromitsu Hattori, "Roadmap for Multiagent Social Simulation on HPC", Proceedings of 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Singapore, December 6-9, 2015, vol.3, pp.22-25, (DOI:10.1109/WI-IAT.2015.178) の図を基に作成した。

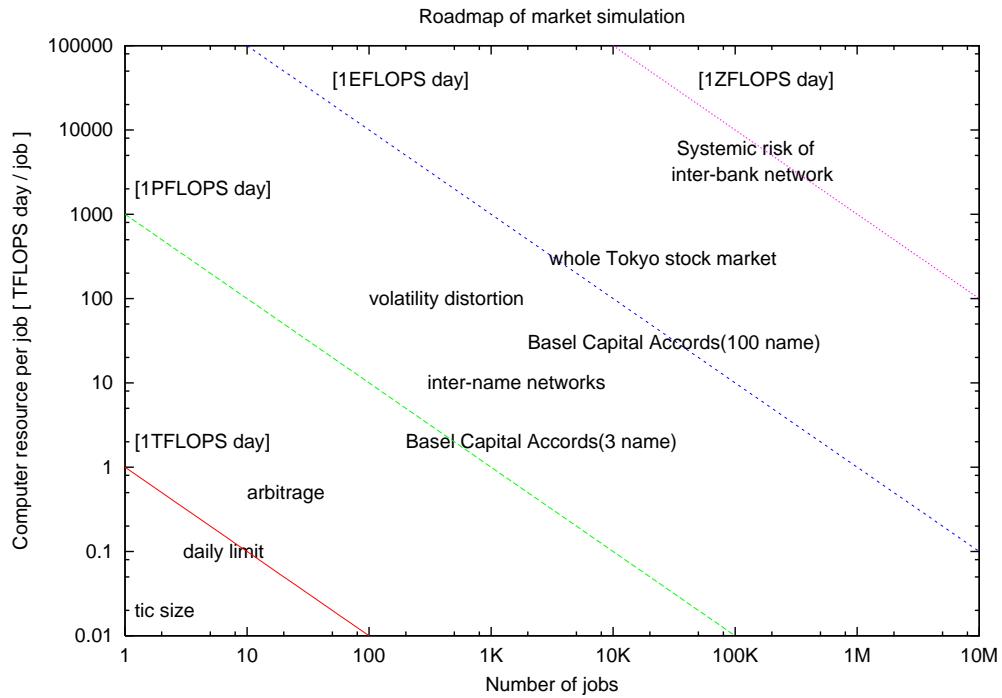


図 2.15 経済取引システム課題解決のために必要な推定計算量。

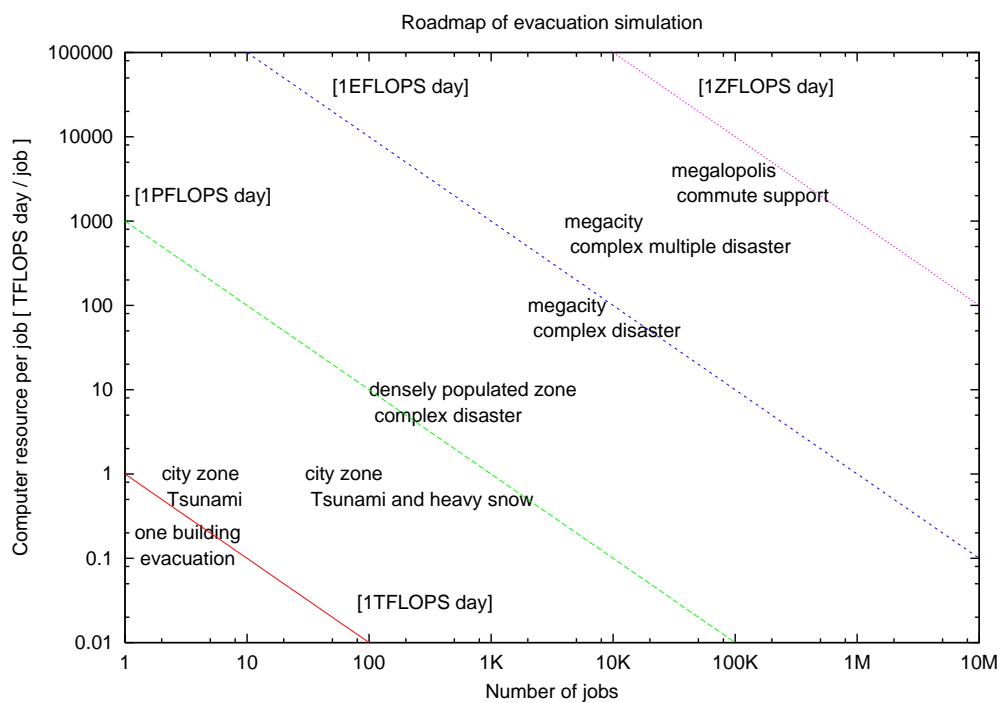


図 2.16 避難問題解決のために必要な推定計算量。

2.9 地震・津波

2.9.1 社会的貢献 —科学的知見に基づく災害予測のシステム化を目指して—

南海トラフの巨大地震や首都直下地震等、大地震の備えを不斷に行うことはわが国に課せられた宿命である。合理的に備えるためには、様々な要因によって生じる被害想定のばらつきを定量的に扱える必要があるが、現状では国レベルでの被害想定は、距離に応じた揺れの減衰の式と、揺れと被害の関係の経験式の組み合わせという単純なものであり、ばらつきが大きいだけでなく、それを小さくするために何をすべきかを合理的に検討することが困難である。より科学的・合理的な地震災害の想定・予測が必要とされており、大規模数値計算を使った地震・津波とそれにともなう災害のシミュレーションはこのための切り札となり得る。

被害想定においては、2011年東日本大震災が突き付けた「想定外」の轍を踏まないためにも、人間の先入観を排除し、観測されるデータと物理モデルから予測される多様なシナリオに対して、その結果、どのような被害が生じ得るかを考慮することが不可欠である。地震のシナリオを考えるにあたっては、地震の発生過程を左右する震源域の上限・下限・水平方向の広がり・破壊開始点・仮定する摩擦則をそれぞれ4～5通り考慮するだけでも、その組み合わせは1,000を超える。こうしたシナリオを、観測事実との整合性等を考慮して数十程度に絞り込むことになる。そうしたとしても、揺れを評価する個々の建物直下の地盤情報や建物の劣化状況、津波遡上に影響する港湾や河川毎の詳細地形の曖昧さを考慮したり、広域に被害をもたらす巨大地震で同時に複数の都市を計算しようとすると、結局1,000ケース程度の地盤震動や津波遡上を計算することになる。そのような多数の地震シナリオを想定し、地下や海底の情報の不確実性も考慮した上で、将来起こり得る被害の多様性を見据えて防災対策(以下では、災害を少しでも減らす「減災」の対策も含めた意味で用いる)を進めなければならない。これは現在のたかだか数例の地震シナリオの想定に基づく防災対策とは一線を画す。また、信頼度の高い被害想定のためには、合理的な法則に基づく構造物と都市のシミュレーションが不可欠である。具体的には、震源からの弾性波動伝播、地盤と構造物の震動、津波の伝播と遡上などのシミュレーションである。更に、揺れや遡上を受けて、どのような経路で避難するのがよいか、また避難誘導を効果的に行うには誘導する人がどこにどれだけ必要かといったことのための避難シミュレーションも必要となる。

また、地震発生から津波遡上に至る一連のシミュレーションは、地震発生前の防災対策に資する災害予測としてだけではなく、地震直後の初動対応のための災害予測としても重要である。地震直後は、特に広域災害で、かつ夜間などの場合、被害状況を推定することは困難を極める。そのような状況において、数時間以内に、発生した地震や津波のデータと、一連のシミュレーションを基に災害予測を行うことができれば、被害状況の推定にとって有益な情報が得られると期待される。

一方、東日本大震災での津波による被災過程から明らかになったように、巨大津波が港湾施設と市街地の隣接した地域を襲う際の被害の波及は、単なる海水の侵入による施設破壊に留まらない。例を挙げると、宮城県気仙沼市では津波の来襲にともなうタンクや船の漂流にともなう油の流出と引火により大規模延焼火災が発生し、被害は甚大なものとなった。このように、災害の外力が複合的・連鎖的に被害を拡大させる過程を「複合災害」と呼ぶが、その発生メカニズムや被害の波及過程を予測できる技術はいまだ発

展途上である。複合災害の被害想定には、地震波の伝播や津波の遡上といった連続体の計算だけでなく、長時間の震動による液状化、津波による構造物破壊、漂流物の移動やその衝突による構造物破壊、更には火災といった現象を扱う必要があり、さまざまな物理法則に支配される現象を予測する高精度なシミュレーションモデルを高度に連成させたマルチフィジックスシミュレーションを行う必要がある。

こうしたシミュレーションを行うことができれば、例えば防潮堤・防波堤等の海岸施設において、以下のようなことが期待される。まず、施設が破壊・倒壊するまでの時間を少しでも長くする、すなわち施設が完全に流失する「全壊」に至る可能性を少しでも減らすような、防災効果を目指した構造上の工夫への要件を明らかにすることができます。また、津波に対する構造物の安定性を評価する技術基盤を構築したり、沿岸部の津波被害(複合被害)拡大リスクを減少させるための評価技術を確立したりすることができます。

一方で、グローバリゼーションが進む経済活動の発展を考えると、将来の地震災害は、構造物や都市の被害という直接的な被害に加えて、この被害がもたらす都市や地域の経済活動の低下といった間接的な被害をより深刻なものとすることが指摘されている。したがって、直接被害の推定の信頼度を持続的に向上するシミュレーションの高度化とともに、地震発生直後の経済活動の低下と被害の復旧の進捗による経済活動の回復を解析できる、間接被害のシミュレーション手法の開発も重要な課題と言える。例えば、電力流通復旧過程において、どれだけの人数の作業員と指示員が、どこからどういう順番に復旧をしていくのがよいかといったことを分析するためには、複数のエージェントの相互作用による振る舞いを分析することができ、避難シミュレーションでも使われているマルチエージェントシミュレーションが応用できる(この場合は作業員がエージェント)。これを、都市や地域の経済の回復と関連させることができ、間接被害のシミュレーション手法開発の第一歩となるであろう。

以上のような地震発生シナリオから被害想定に至る一連のシミュレーションは、日本にとって重要なものである。その一方で、極端に言えば自然の中の複雑な人間活動すべてを計算機で追跡する極めて複雑なものであるため、その時代の最先端のHPC上で、継続的に解析手法・モデルを高度化しつつ開発を行うことが必要である。それと同時に、5年、10年単位でのHPCの高速化に応じて、一世代前にHPC上で開発したシミュレーションツールを、個別の地域を対象とする研究機関や大学での研究レベルや実務レベルに順次下方展開していくことも重要である。これにより、国レベルでの大規模災害に対する想定業務はもとより、各地方自治体や企業レベルでの想定業務が、継続的に高度化され、信頼性を高めていく仕組みを構築することができる。そのためには、継続的に最先端のHPC、それに続く研究レベルや実務レベルのHPCを開発・製造することが必要になる。これは、最新の計算科学に基づいた最良の被害評価手法を標準化したものとして社会で共有し、防災に結びつけることであり、この課題の社会的意義もある。

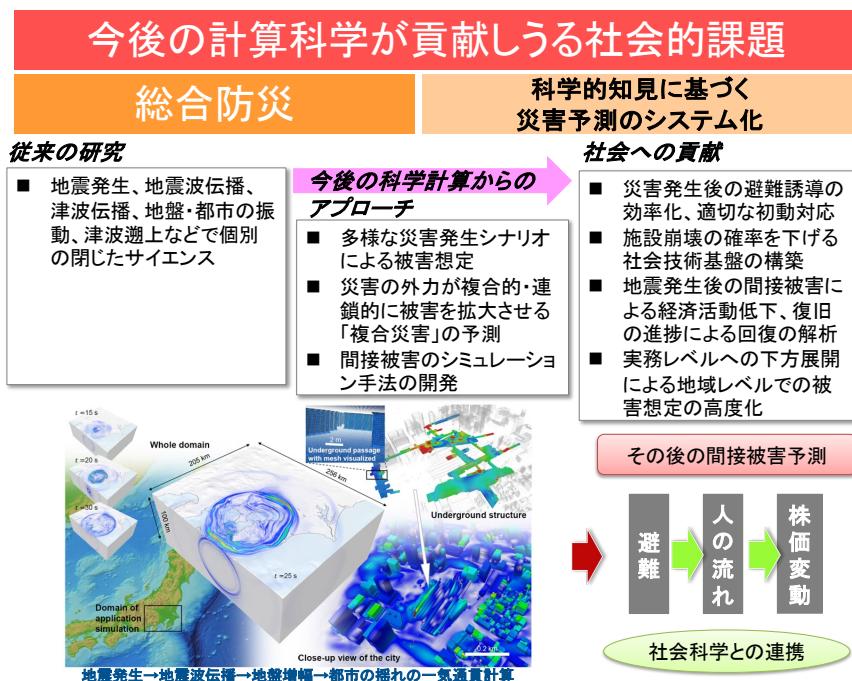


図 2.17 総合防災（地震・津波防災）

2.9.2 サイエンスの質的变化

地震・津波分野では、従来は、地震発生・地震波伝播・津波伝播・地盤や都市の振動・津波遡上といった各現象を独立して扱い、それぞれが閉じたサイエンスとして発達してきた。例えば、地震発生であれば、地震発生に関わる仮説を提案し、それに基づいて断層運動の結果としての地殻変動や地震波形をシミュレーションで計算し、観測データとの比較によって仮説を検証することが行われてきた。しかし、東日本大震災をもたらしたマグニチュード（以下では M）9 クラスの地震は、その繰り返し間隔が数百年以上に及ぶため、地殻変動や地震波形の観測データは、最近発生した地震に対してしか得られない（そこでは同様の地震は当面起こらない）。これに対して、地震発生から都市の振動や津波遡上までをつなぐシミュレーションが実現すれば、これまでとは質的に全く異なる状況となる。各現象には、それに応じた観測データがさまざまな時期にさまざまな精度で残されている。例えば、津波堆積物は 1000 年以上にも遡ってデータが存在する。そのため、ある一つの地震発生シナリオに基づいて津波遡上まで計算すれば、それが数百年前に津波堆積物をもたらした現象と整合するかどうかを、誤差を考慮したうえで科学的根拠に基づいて評価することが可能となる。それを非常に多数のシナリオに対して行えば、それらのシナリオがさまざまな時空間スケール、そしてさまざまな精度で散在している貴重な観測データに対して、どの程度整合しているかを評価できることになる。同様なことを、津波堆積物だけでなく、地震発生から建物の振動や津波遡上に至る過程に対しても適用していくことで、質の異なる多様なデータ（情報量の多さとしてビッグデータでもある）を統合的な知見として集約していくことができ、これまで個々に発展してきたサイエンスが、分野の垣根を越えて総体として発展していく道筋ができると期待される。

また、分野の垣根を越えるという意味では、津波に関わる課題において、これまでの各分野で独立に先鋭化された地震学、計算工学、流体力学、構造工学といった学術分野を専門とする研究者が連携して問題を取り組むようになりつつある。具体的には、(1) 沖合の津波観測情報のリアルタイム処理と波源の推定、(2) リアルタイム浸水予測を含む津波ハザード解析手法の高度化、(3) 津波氾濫にともなう土砂移動・地形変化の解析手法の高度化、(4) 固体・液体の連成解析手法の確立、(5) 避難行動解析を含む人的被害予測手法の高度化などが課題として挙げられる。なかでも、津波に強い町づくりに資する重要なテーマとして、「固体・液体の連成解析手法」の取り組みを更に発展させ、津波による複合災害の予測と防災対策への利用・普及を目的としたマルチフィジックスシミュレーションモデルを開発するといった目標を達成するため、以下に示すようなテーマに取り組むことで、今後より一層の連携が進むはずである。

1. 巨大地震津波の近地・遠地の波動場の高精度予測（全球シミュレーション）
2. 構造物の破壊・被害拡大過程のシミュレーション（水位・流速・波力・破壊・漂流物・火災等）
3. 流体・構造物・地盤の相互作用の解明
4. 「粘り強い構造」の性能と設計照査のための予測技術体系
5. 津波防災地域づくりへの貢献

更に、マルチエージェントシミュレーションを活用した間接被害のシミュレーション手法の開発は、計算科学と数理経済学を結ぶ新たな境界領域となることが期待され、これは、計算科学の社会科学的課題への貢献にもつながるものである。

2017年以降のサイエンスの質的な変化としては、AI・データ駆動科学の地震・津波研究への導入やSociety5.0の実現に向けた都市丸ごとのデジタルツインの構築の流れがある。前者については、2.9.4で後述するように大規模シミュレーションへの活用がすでに進んでいる。一方、後者は防災・減災分野におけるSociety 5.0実現を目指して、都市を対象とした防災デジタルツインの構築に向けた取組が進んでおり、将来的に大規模シミュレーションが求められることになると期待される。

2.9.3 コミュニティからの意見

一般論として、災害対策の実務において、シミュレーション技術の発展が社会の問題の具体的解決にどのように役立つかといった意見がある。具体的には、予測結果はどの程度正確なのか、結果のばらつき・不確実性はどのように考えるのか、逆に、災害対応の現場において果たしてそこまで高度な予測結果が必要なのか、といったことである。例えば、シミュレーション結果を津波に強い町づくりに向けてどの局面で活用していくか、といった課題がある。これらは、研究成果の社会実装を考えるうえで取り組むべき重要な課題であり、これらの疑問への回答を得ることを念頭において、シミュレーション結果の検証やばらつき・不確実性へ考え方も含めて課題に取り組む必要がある。なお、ばらつきや不確実性は、現象の本質やデータが不可避的に不十分であることに起因する面が大きいので、HPCを活用した多数シナリオの導入によってそれらを定量化したうえで、実務に関わる人達と一緒に、実務への生かし方を検討することが不可欠である。そのなかで、災害対応の現場でどのように高度な予測結果が生かせるかを示して行く必要がある。一方、企業が個別の領域での地震による揺れの計算（サイト地震動計算）を行う上では、国が行う被害想定のために行う地震動計算との共通化（震源断層のシナリオや三次元不均質の地下構造のモ

デル、計算コード等) の必要性がニーズとして指摘された。

また、平成 25 年からスタートした日本地球惑星科学連合合同大会の特別セッション「ハイパフォーマンスコンピューティングが拓く固体地球科学の未来」においては、工学分野からの意見として、災害予測や被害想定をシミュレーションベースで行う方向性について全面的に賛成する一方で、本当のニーズをいかに捉えて、それに応えるものにしていくかの重要性が指摘された。ニーズに応えるといつても、受け手からの一方的な要求に応えるのではなく、災害予測や被害想定の実力に見合った範囲で、かつ受け手が満足できるところのコンセンサスを得ることが重要である。そのためには「ご意見伺い」ではなく、ユーザを取り込み、分野を超えた共同研究を行うことが必要になる。また、これからの方針として、オンデマンドで、個人や個々の団体のニーズに合わせた防災に資する情報を、Web やスマートフォン等も活用して個別に発信していくことが複数の講演者から提案された。更に、地球科学は、真理を探求する認識科学から、観測データとシミュレーションを両輪として問題解決のための予測能力を高める設計科学への質的転換が必要であるとの指摘もなされた。

その後も同セッションは毎年開かれており、フォワード計算だけではなく、HPC をデータ解析にも活用する観点で講演・議論が行われている。その中では、フォワード計算や逆解析を行っている研究者らが現状の手法の抱える問題を提起するとともに、計算機科学・計算科学の発達についても紹介され、これらを固体地球科学分野の様々なフォワード計算や逆解析にフルに活用して解決していく方向性について議論があった。また、従来の構造を単純化したモデルの限界が示されるとともに、高密度化する観測データをフルに活用した高詳細なモデルでの地震波伝播や地殻変動・津波伝播の計算、多数回計算を行うことによる物性や構造の誤差の影響の検討、誤差にガウス分布以外の一般的な分布を許した断層すべりの逆解析などが行われつつある現状が紹介され、その活用の必要性・重要性が確認された。

2.9.4 前回のロードマップ以降の進展

前回のロードマップで取り上げたマルチフィジックスのシミュレーションについては、大学や研究機関の第二階層計算機を活用して、その後も着実に研究開発が進んでいる。一方、より大規模な計算が必要となる地震発生や地震動・地盤增幅等のシミュレーションについては、ポスト「京」重点課題や「富岳」成果創出加速プログラムにおいて、曖昧さや AI・データサイエンスとの融合を含んだ研究開発を進めてきた。ここでは、そのようなより大規模な非構造格子有限要素法による粘弾性地殻変動解析・非線形動的地震応答解析シミュレーションに関わる進展を取り上げる。

計算機科学・計算科学の中心的な国際会議である SC で、開発した大規模高効率有限要素コードについて発表し、2018 年に Gordon Bell 賞のファイナリストに選出された (Ichimura et al., 2018)。これまでの多くの大規模 3 次元詳細シミュレーションにおいては地盤や構造物の地震応答を分けてシミュレーションしていたところ、都市の地盤と構造物を一体として非構造の有限要素モデルを生成し、高速計算アルゴリズムを開発することで富岳及び Summit で計算コストの高い地盤・構造物の 3 次元詳細連成解析を実現した。高速アルゴリズム開発においては、ここ数年で各分野で進む AI for Science に先行的に取り組み、AI および精度混合演算 (FP16-FP21-FP32-FP64) を計算の前処理に活用することで Summit の 4096 ノードでピーク性能の 14.7% (=28.3 PFLOPS) という高い演算効率を得、計算精度を劣化させることなく従来手法の 25 倍の高速化を実現した。2019-2020 年には、AI 用に開発された加速演算器であ

る Tensor Core により地殻変動解析用の有限要素法ソルバーを効率化することで 1.5×10^{10} 自由度のモデルの地殻変動解析を Summit の 4544 ノードで演算効率 416 PFLOPS で実現し、3 次元不均質媒質中の応力評価のために必要な高解像度で M9 クラスの震源域を含む広域モデル ($1944 \text{ km} \times 2646 \text{ km} \times 480 \text{ km}$) の地殻変動解析を可能とした (Ichimura et al., SC19 Poster; Yamaguchi et al., PASC'20)。また、2021 年には人工知能によりグリーン関数を介して物理シミュレーションを高速化する方法を富岳上で開発するなど (Ichimura et al., SC20 SCALA)、AI を用いることで計算精度を劣化させることなく地震シミュレーションを高速化する手法を開発している。

2017 年のロードマップにおいて掲示された、地震発生・地殻中の地震波動伝播・都市の応答を含む統括的地震応答シミュレーションの実現に向けた研究開発も着実に進んでいる。過去時間ステップのデータを保持しつつ毎ステップ学習し将来ステップの初期解を予測するデータ駆動型手法によりさらに動的非線形有限要素解析を高速化することで、従来分けて実施していた広域の地震動伝播と都市内の地盤增幅・構造物応答シミュレーションを連成して一体解析することができるようになり、断層から都市までを単一の有限要素モデルにてモデル化し地殻中の波動伝播から地表付近での地盤增幅、構造物の応答までを高解像度で連成して解く世界で初めてのシミュレーションを実現した (Ichimura et al., HPC Asia 2022 Best Paper 賞を受賞)。なお、このデータ駆動型の初期解予測手法は粘弾性地殻変動解析においても有効で富岳上で高速計算が可能となっており (Fujita et al., SC22 ScalAH)、本ロードマップにおいて掲げている地殻変動と地震動の観測結果の双方に整合するモデルの構築に向け、着実にシミュレーション能力の高度化が進められている。

2017 年のロードマップにおいて提示されたもう一つの課題であるシミュレーションの曖昧さの評価のうち、地下構造における物性の曖昧さを考慮したシミュレーションを実現する面でも大きな進展があった。これまでの高詳細シミュレーションでは 1 ケースずつの計算を効率化したり、まとめて数ケース計算することで計算を高速化する方法が開発されてきたが、2022 年においては高詳細 3 次元地殻構造モデルを対象に地殻の物性の曖昧さを考慮した地殻変動解析を効率的に計算可能とする、富岳全系までスケールする高効率確率有限要素解析手法を開発した。通常、確率応答問題の数値解の信頼性を担保するためには空間方向に高詳細なメッシュを使うだけでなく確率応答分布の収束のため 10000 回オーダーのモンテカルロシミュレーションが必要となり膨大な計算コストが必要となるが、本手法を使うことで 1 回の計算で確率応答分布を計算でき、従来手法比で 100 倍以上の計算速度向上を実現した (Ichimura et al., Gordon Bell Prize Finalist, SC22)。本手法を用いることで、想定南海地震の震源域を含む $2496 \text{ km} \times 2496 \text{ km} \times 1101 \text{ km}$ 程度の領域で、3 次元不均質地殻モデルにおける媒質の物性の曖昧さを考慮して断層すべりによる地表の変位を計算することが可能となり、前回のロードマップに掲げた不確実性や多様性の考慮という課題の解決に向けて着実に進展したと言える。

これらの開発により、半無限均質媒質な地殻構造に対して行ってきた地震発生シナリオの計算を、3 次元不均質媒質に対して有限要素法を用いて行える道が拓けるとともに、地震発生から波動伝播、地盤增幅・都市応答まで、共通のソルバーを持つ非構造格子有限要素法コードで計算できることになる。また、モデルの曖昧さを考慮した応答計算も進みつつあり、さらに、津波シミュレーションについても、上記の有限要素法による地殻変動計算結果にもとづく初期波形を用いることで、詳細な地形や地下の構造不均質の影響を考慮した計算が可能となっている。そこで、富岳成果創出プログラムの 2 期目（2023-2025 年度）では、基礎的なサイエンスの進展に対して超大規模・高速シミュレーションによって貢献する方向の

研究開発も進めている。具体的には、地震発生予測のために必要となるM8.9の巨大地震震源域の詳細な3次元地下構造やその曖昧さを考慮した断層固着・すべり状態の時空間変化にともなう地殻変動や応力変化の高精度評価、あるいは全球スケールでの大地震や地殻活動の相互作用を考慮できる粘弾性応答シミュレーションなどの実現に向けた開発である。ただし、これらの詳細な3次元不均質構造やその曖昧さを考慮した計算がすべて「富岳」で実施できる訳ではなく、より大規模・高速な次世代計算機が必要となる。その点については、次節2.9.5で述べる。

一方、2014年・2017年のロードマップで共通の目標として述べていた社会的貢献の方向性として「最新の計算科学に基づいた最良の被害評価手法を標準化したものとして社会で共有し、防災に結びつけること」がある。これを受け、ポスト京重点課題（2015-2019年度）でのHPC上でのシミュレーションツールの開発、その成果を受けた富岳成果創出加速プログラムの1期目（2020-2022年度）によって、国レベルでの大規模災害に対する想定業務のうち、高層ビルや大規模構造物の評価にとって重要な、長周期地震動のシミュレーションツールの社会実装を実現した。具体的には、内閣府政策統括官（防災担当）が、「相模トラフ沿いの巨大地震に伴う長周期地震動による被害予測の高度化」の課題のため、ポスト京重点課題で開発した大規模有限要素プログラムをベースとして成果創出加速プログラムで整備したE-wave FEMを用いた長周期地震動の計算を「富岳」の政策枠で実施した。

また、コミュニティからのニーズにもとづいて、国が用いるのと同等なシミュレーションツールを、建築系の企業が実務に有用な独自の問題設定で、「富岳」を用いて大規模計算を実施できるような計算実施環境とサポート体制を整えた。これにより、複数の企業が、関東地震や千島海溝巨大地震を対象として、「富岳」を用いた長周期地震動大規模シミュレーションや多数回シミュレーションを実施し、成果発表に至るなど、ロードマップで掲げた社会的貢献の方向性の取組を着実に達成した。さらに、実務向けの計算で、「富岳」までは必要とならない規模の計算を、第二階層である地球シミュレータ（ES4）でも、同様に計算ができる環境を整えており、「京」や「富岳」向けに開発した大規模計算コードの下方展開についても実現したと言える。富岳成果創出プログラムの2期目（2023-2025年度）では、この社会実装をさらに進めている。

2.9.5 ロードマップ

年代 課題	2024～	2026～	2028～	2030～	2032～
全球スケール 地殻変動解析	全球スケールでの計算実現に向けたアルゴリズムの高度化				全球スケールの地殻変動解析
		全球スケールでの粘弾性モデルの構築			全球モデル改良
特定領域 地殻変動解析・ 地震波動解析	日本列島3次元弾性＆粘弾性地下構造モデル更新システムの構築と稼働				
	南海トラフ～相模トラフ域の解析	千島海溝～日本海溝域の解析	南西諸島、伊豆・小笠原海溝域の解析	多数解析とモデルの改良	多数解析とモデルの改良

図2.18 地震・津波分野ロードマップ

2.9.6 計算機要求

日本列島には、地震波形観測のための Hi-net や地表面変形観測のための GEONET など、それぞれ 1,000 点前後の観測点がある観測網をはじめとして、多くの観測網が設置されている。観測データと大規模シミュレーションの融合は、地震分野における様々なサイエンスへの貢献が期待されており、例えば、日本列島島弧の挙動予測、地震活動度推定、活動帯の高解像地殻変動・地震動を介したメカニズム解明、津波評価に必要な地殻変動推定、さらには、重要かつ困難な課題である地震発生予測をも視野に入れるインパクトのある成果が期待される。これらの解析を行う上で、不確定性を考慮するため、地殻構造や初期応力の異なるアンサンブル計算が必要となる。さらに、連続リアルタイムデータに対して適用し続けるために超大規模計算を高速に行うことが求められるため、次世代計算基盤への強い期待がある。全球スケールの大規模マクロ地殻変動シミュレータと特定地域に絞った高解像度ミクロ地殻変動・地震動のシミュレータを組み合わせたマルチスケールの 3 次元地殻変動・地震動シミュレータを開発し、日本列島における地震波形データと地殻変動データの双方に整合するマルチスケールモデルを構築することで、一定規模の地震が発生した後、周囲での地震発生の推移を予測できるようになると期待される。

上記のような 3 次元複雑形状・非均質構造を求解対象とする地震分野においては、陰解法の低次非構造格子有限要素法が主要な計算手法の一つとなる。低次構造格子有限要素法に基づく地殻変動解析(静弾性解析・粘弾性解析)・地震動解析(動的な線形・非線形波動場解析)において基礎アルゴリズム・カーネルは共通であり、反復法ベースのソルバー内の疎行列ベクトル積が主要カーネルとなる。このカーネルにおいてはランダムアクセス・データレカレンスが卓越し、メモリバンド幅ネックのカーネルとなる。また、疎行列ベクトル積における 1 対 1 通信と反復法ソルバーにおける 1 要素の Allreduce が主要な通信パターンとなる。

上記を踏まえ、2030 年ごろに想定される計算における要求計算規模について記述する。全球スケールの大規模マクロ地殻変動解析においては、マントルの底までを対象とした重力下での粘弾性応答計算・粘弾性完全緩和のため長期間の応答計算における必要計算資源量はこれまで富岳全系で実施してきた規模の問題の 30 倍程度(演算数 30 倍、メモリ容量 5-10 倍、メモリ帯域 5-10 倍)と見積もっている。特定領域のミクロ地殻変動・地震動解析においては、南海トラフなど、 $1,000 \text{ km} \times 1,000 \text{ km} \times 500 \text{ km}$ オーダーの解析における 1 ケースあたりの必要計算資源量は現在富岳で実際されている問題と同等だが、多数ケースを実施することとなる。次世代計算基盤において上記の計算を実施するには、計算加速器(accelerator)において演算性能だけでなくメモリ容量・メモリ帯域も確保される構成が望ましく(演算性能あたりメモリ容量・帯域が富岳と同等から $1/2$ 程度までに保たれることが理想的)、また、疎行列ベクトル積の計算においては足しこみ競合が発生するため、高速な hardware accelerated atomic add が実装されていることが性能上重要となる。また、1 対 1 通信においてはインターフェクトのバンド幅だけでなく計算加速器と CPU 間のバンド幅が律速になることが多いことから、計算加速器から CPU を介さずに通信するなど計算加速器からの通信性能が最大化される構成が望ましい。

以下、特に計算規模が大きい全球スケールのマクロ地殻変動解析の計算規模をまとめると。

- 問題規模：全球スケール、100 m メッシュ、100 兆自由度・1000 時間ステップの粘弾性地殻応答解析
- ケース数：10

- 計算時間：1 時間/ケース
- 演算量：10,000 EFLOP
- メモリ量：25PB
- メモリ転送量/ケース：5,000 EB
- ストレージ量/ケース：1PB
- 要求ファイル I/O 性能：10TB/s
- 想定するノード数：20,000（全系・ファットノードを想定）

2032 年頃における課題解決のために必要な計算機資源の見積りについては、2.13 節の表 2.1 も参照いただきたい。

2.10 気象・気候

2.10.1 分野の概要

近年、毎年のように、世界各地でさまざまな気象災害の発生が報告されている。2019年から2021年にかけて、インドや中国、ヨーロッパ中部などで記録的な大雨が観測され、死者や行方不明者が多数発生した^{*4}。その一方で、2021年は、ヨーロッパ南部や米国南西部を含む北半球各地で顕著な高温が観測され、米国南西部では「50年に1回以下」の極端な乾燥に見舞われた。国内においては、2018年7月猛暑、2019年の台風19号（東日本台風）や令和2年（2020年）7月豪雨、2021年7月の東海地方・関東地方南部を中心とした大雨に伴う日本各地での河川の氾濫や熱海での土砂災害は、記憶に新しいところである。これらの気象災害リスクは、地球温暖化に伴う気候変動により、より一層高まることが懸念されている。実際、過去のいくつかの事例については、事後解析により、地球温暖化が発生リスクを高めたことが指摘されている[1]。

気象災害や将来の気候変化を事前に予測し、リスク回避のための情報を提供することは、本分野の重要な役割の一つである。日々の天気予報から数十年先の気候予測まで、幅広い時間スケールでの予報・予測精度の向上が求められる。予報予測技術の発展のためには、理論、観測、実験に基づいた、様々な時空間スケールでの大気現象の理解をより一層高めていくことが不可欠である。数値モデルの使途は様々であるが、予報・予測の観点では、1日から数ヶ月の時間スケールを対象とした数値天気予報と過去から未来の長い時間スケールを対象とした気候計算の2つに大別することができる。以下では、数値天気予報から始まった当分野での数値シミュレーションの歴史とともに、最近の世界的な取り組みや動向について述べる。

大気の数値シミュレーションの歴史は古く、Richardson (1922)[2] の手動計算による数値天気予報の試みまでさかのぼる。Richardson (1922) の試み自体は、天気予報にとって重要ではない高周波のノイズが原因で失敗に終わったが、初期値にノイズを含まない客観解析、およびノイズをダンプするシミュレーションモデルの必要性が主張された[3]。現代においても、数値天気予報の精度はシミュレーションモデルとその初期値を提供する客観解析の二つの精度に依存する。大型計算機の発展とともに、数値計算手法や初期値を作成する客観解析技術の向上、水平解像度や鉛直層数の増強、アンサンブルメンバー数増強などが行われ、数値天気予報システムの性能はこれまで継続して向上してきた[4]。

一方、気候モデルの歴史は、1次元モデルを用いて、大気の放射バランスを計算したManabe and Wetherald (1967)[5] にさかのぼる。真鍋らはこの論文の中で、数値モデルを用いて地球のエネルギーバランスを解き、大気の気温の鉛直分布を再現した。また、同じ論文の中で、大気中の二酸化炭素濃度を2倍に変化させると地表平均気温が2.36度上昇することを示し、二酸化炭素の気候影響について、定量的に評価した。その後、真鍋らは、3次元の全球モデルの開発も行い、この際に用いられた3次元モデルが、現在広く使われている地球温暖化予測モデルの原型となった[6]。これらの功績により真鍋らは2021年のノーベル物理学賞に選ばれた。最初の気候モデルは大気と放射のみの計算であったが、その後、雲、乱流、陸面、海洋などが加わり、最新の地球システムモデル（Earth system model, ESM）では、化学物

^{*4} https://www.data.jma.go.jp/gmd/cpd/monitor/annual/annual_2020.html

質やエアロゾル、炭素循環も組み込まれている (e.g., [7], [8])。

全球気候モデルの開発は世界各地で行われており、社会経済シナリオや温室効果ガスシナリオなど、統一した条件で計算した結果を比較するプロジェクト (結合モデル相互比較計画; CMIP: Coupled Model Inter-comparison Project) が 1995 年より行われている [9]。これは、観測データの乏しい過去や将来気候の計算結果を異なるモデル間で比較することで、信頼性の高い情報を得るだけでなく、予測情報に含まれる不確実性や気候モデルに対する理解を深める目的で行われている。CMIP に提出された気候データは広く公開され、それを解析して論文として出版された知見は、気候変動に関する政府間パネル (IPCC: Intergovernmental Panel on Climate Change) 評価報告書や特別報告書の根拠となっている。^{*5} これらの活動は今後も計画されており、さらなる解像度の向上や、発生頻度の低い極端事象を捉えるためのアンサンブル数の増強、海洋の予測データの拡充などが望まれている。

これまでの気象予報や気候予測をさらに発展させた取り組みとして、近年、地球のデジタルツインである “Digital Earth” の構築が提唱され [10]、欧州において大型プロジェクト (Destination Earth) が開始されるとともに、WCRP (World Climate Research Programme; 世界気候研究計画) において今後 10 年を見据えた Digital Earth Lighthouse Activity が開始されている。Digital Earth は利用可能なあらゆる観測データを統合し、数値モデルやデータ科学の力を借りながら、過去・現在の地球の気候状態を仮想世界上に可能な限り忠実に再現し、さらに未来を予測しようという試みである。未来についてはユーザーがシナリオを自在に選択することで政策の効果を定量的に検証することが可能となる。Digital Earth は精緻な気象・気候モデルを前提としており、近年急速に実用化が進んでいる高解像度の全球雲解像モデル [11] と密接に関連している。全球雲解像モデルの気候モデル化を求める声も高まっている [12] が、膨大な計算コストとともに、膨大な出力データの取り扱いが研究のボトルネックとなっている。シミュレーションから解析までを同じプラットフォームで行う “Digital Earth” は、この問題に対しても有効と期待される。

気候計算などのより長期なシミュレーションでは、初期値の精度よりも、大気とともに気候システムを形成する海洋や陸域、生物圏などのモデル内での扱いが予測結果において重要な要素となってくる。それ故に、各要素の気候への影響についての理解を深め、各モデル要素の精緻化を進めるとともに、気候影響の定量的な評価を行なっていくことも重要な課題である。また地域スケールでの予測情報の活用も進んでいる。工学分野で発展してきた建物解像モデルを用いた都市域内での熱環境評価、土砂を含む河川モデルを用いた大雨による土砂災害・洪水被害などの複合災害評価、社会システムモデルやエージェントモデルによる気象災害による人的被害や経済リスクの評価、日本近海の海流情報を用いたスマート水産業や海上発電などさまざまな分野との連携も進んでいる。気候変動適応策の観点からも、分野の垣根を超えた研究が今後ますます発展していくものと期待される。

2.10.2 長期目標と社会貢献

気候気象分野の社会貢献の一つに、日々の天気予報と防災情報の提供、及び、将来の気候変化に対する適応策や影響評価のための将来気候予測情報の提供がある。予測精度の向上に向けて、新たな数値スキー

^{*5} 直近では、気候変動やその影響、また気候変動の緩和についてまとめた IPCC 第 6 次評価報告書 (AR6) が 2021 年から 2022 年にかけて発表されたところである。

ムの提案や高度化、その根拠となる各現象の理解を高める努力が日々行われている。以下では、数値天気予報、将来気候予測、モデルの精緻化の視点から、長期目標と社会貢献について述べる。

近年、極端な気象現象による自然災害が激甚化しており、今後更に災害が深刻化するおそれがある懸念されている。特に、近年の大暴雨に伴う災害では線状降水帯の寄与が大きく、気象庁は「線状降水帯の発生・停滞の予測精度向上により、集中豪雨の可能性を高い確度で予測する。特に、明るいうちからの避難など、早期の警戒と避難を可能にする」という目標を掲げている[13]。しかし、最も解像度の高い局地モデル(LFM: Local Forecast Model)であっても、線状降水帯を半日程度前から時間と場所を絞って予測することは依然困難である。予測の精度向上には、個々の積乱雲を表現するための高解像度化、高い解像度に適した物理過程の開発、海上の水蒸気量や風などの予報モデルの初期状態の精度向上などが課題である[13]。

台風防災も気象庁の2030年に向けた主な開発項目^{*6}の一つであり、モデルの精緻化とともに観測データの充実が重要な課題となっている。地球観測技術は年々高度化しており、数多くの人工衛星や地上観測装置が開発されている。そのデータ量は1世代前のものに比べて数十～百倍と、まさにビッグデータである。それ故、膨大な観測データを効率的にシミュレーションに取り込むデータ同化システムの開発も必要である。

一方、将来気候予測においては、世界の機関と協力して気候変動問題に貢献していくことが、気象気候分野に課された重要な課題である。2021年8月に公表されたIPCC第一作業部会(WG I)による第6次評価報告書(AR6)では、「人間の影響が大気、海洋及び陸域を温暖化させてきたことには疑う余地がない」という強い表現で人間活動の気候への影響が断言された[14]。このような背景のもと、全球気候モデルは、気候変動予測に対する科学的理解の増進に加え、温暖化の進行そのものを止める緩和抑制策の立案や温暖化への対策立案に資するデータ創出への貢献という大きな役割を担っている。

気候モデルに生態系の働きなどを組み入れた地球システムモデル(ESM)では、温室効果ガスの排出による気候変化のみならず、二酸化炭素の海洋吸収による海洋酸性化、農作物や生物燃料栽培、あるいは都市化に伴う土地利用変化の環境影響など、様々な問題を包括的に取り扱うことができる。こうした問題は伝統的な気候科学の範疇では取り扱えず、農学や社会経済など多様な分野の研究者との協働が必要とされる。さらに、社会問題としての地球環境問題に対処するためには一般社会における利害関係者の意見も聞きながら問題設定を行う必要がある。専門性を高め取り扱う問題を厳密に定義することで知見を集積し発展してきた従来型の科学とは異なるアプローチが必要になる。こうした状況に対し、社会経済分野で発達してきた統合評価モデル(将来想定シナリオの策定などに用いられる)と、ESMとを結合した地球システム-社会経済結合モデルを開発することで、有用なツールを提供できる。完成したツールを通じ、多様な分野、セクターを巻き込んだコミュニケーションを促進して、連携を深めることができると期待される。

また、温暖化適応策のための気候予測情報においては、全球を対象とした気候モデルでは計算機資源の制約により、現在において、最も高解像度のモデルでも格子間隔は20km程度である。この解像度では、細かな地形の影響を受ける降水や地表面状態の影響を直接受ける気温などについて、高い再現精度が望めない。そこで、日本周辺など対象となる特定領域のみを高解像度の領域モデルで計算したり、統計的なモ

^{*6} https://www.jma.go.jp/jma/kishou/know/tecdev/nwp_strategic_plan_towards_2030.html

モデルを用いて高解像度化（ダウンスケーリング）したりすることで、適応策のための領域気象情報を作成している。ダウンスケーリングで予測される気候情報の信頼性向上のためには、全球モデル及び領域モデルの両方の再現性や予測精度を向上させていく必要がある。それと同時に、ダウンスケールによって計算された地域気候予測結果に含まれる不確実性の要因の理解と定量的評価も気候情報の信頼性に不可欠である [15, 16, 17]。

気象災害につながるような極端現象など、出現率の低い現象の発生確率や将来変化の予測においては、予測結果の信頼性を統計的に評価するのに十分なサンプル数を確保する必要がある。国内においては、世界に先駆けて力学的ダウンスケーリングとアンサンブル実験によって温暖化の影響評価のための大規模データセット d4PDF[18] を作成し、様々な分野での影響評価に役立てている [19]。今後は、モデルの精緻化・高解像度化と気候予測のアンサンブル数の増加のどちらも向上させていく必要がある。

上記の数値天気予報、および、気候変動予測で用いる気象・気候モデルの再現や予測性能に対する信頼度を現状よりも高めていくためには、モデルそのものの高度化・精緻化が必要である。気象・気候モデルの高度化・精緻化は、2つの方向がある。1つは、対象とする領域を覆うメッシュを細かくする「高解像度化」、もう1つは、物理プロセスの精緻化である。気象・気候モデルにおける物理プロセスとは、格子サイズで解像可能なスケール以上の大気の流れを計算する力学過程以外の計算を扱うプロセスのことを指す。また、物理プロセスの精緻化には、高解像度化や物理モデルの精緻化が進むにつれて、従来の手法の適応限界を超えるため、新たに物理モデルを開発する意味での精緻化と、従来の気象・気候モデルでは計算資源の制約から考慮に入れることができなかった物理過程を取り入れる意味での精緻化がある。

個々の現象を物理的に正しく再現するという観点からは、モデルの高解像度化はストレートなアプローチである。地球大気は全球規模の大スケール（～10,000km）から雲・乱流といった小スケール（10m以下）の現象が相互に作用する複雑な系である。特に水平解像度を数km程度まで高めると、対流圈全層に渡る深い鉛直対流がモデル格子で直接解像され始め、雲・降水システムの表現が質的に大きく向上する [20, 21]。このような雲解像（cloud-resolving）モデルでは、雲・循環・放射の相互作用を物理的に整合した方法で表現できるため、現在の気候モデルでは不確定性が大きい雲の気候感度の見積りに加え、対流活動が活発な熱帯の気象現象の理解向上や台風などの予測の向上につながることが期待できる。さらに水平・鉛直解像度を100m程度以下まで微細化できれば乱流の一部が解像されはじめる。モデルで乱流を陽に扱えるようになれば（LES: Large-eddy simulation）、大気境界層の表現が飛躍的に向上し、現在では直接的な再現が難しい、下層雲を含む全球の雲の高さや光学的厚さの分布について、自励的に再現されることが期待できる。このような全球LESが気候スケールで実用的になれば、雲と気候の相互作用に対する理解が飛躍的に進展すると予想される。さらに、2030年頃に実用化が期待される技術として、雲が持つ電荷や雷を直接計算し、雷頻度を直接予報する「気象雷モデル [22, 23]」がある。これが実用化されれば、現在は雷警報や雷活動度といった抽象的な物理量による予報しか存在していない雷の予測を雷頻度や電荷の分布の予報に繋げることができるだけでなく、雷による電気設備への被害や航空機への被雷被害低減という点で、社会に対して大きな貢献となる。

気候モデルによる長期気候変動の予測の向上においては、大気モデルだけでなく、海洋モデルや化学モデルの高精度化も必須である。気候モデルの海洋コンポメントモデルで、全球をほぼ同じサイズの格子で覆う場合、CMIP6に参加している我が国のモデルの格子幅は100km程度、海洋高解像度版でも25km程度と、世界で最も細かいモデルの格子サイズ（約10km）より粗い。将来のCMIPを見据えて、5年後

を目標に、全球 10km 解像度の海洋モデルを組み込んだ気候モデルの構築を目指す。海洋モデルの高解像度化により、黒潮やメキシコ湾流などの西岸境界流の表現が大きく向上することが期待される。さらには、海洋モデルの精度向上により、気候予測だけでなく、実用面から社会的に重要視されている北極域や日本近海の海流情報の提供も期待される。一方、大気微量成分の大気中での動態およびその気候的影響を正しく評価するためには、大気中で物質がどのように運ばれ消えていくのか、すなわち 1) 地表・海洋表面上などから大気への排出および植物等による吸収、2) 輸送過程、3) 変性過程、4) 沈着過程の諸過程と、大気中における放射過程となるべく精緻に評価する必要がある。モデルの高解像度化に加えて、諸過程への理解を深め、それらのモデルへの取り込みを継続して進めていく。

2.10.3 課題とその解決に必要な計算手法・アプリケーション

気象・気候モデルは、大気モデルを中心に、用途に応じて海洋・海水、陸地、雪水、化学・エアロゾル、陸域・海洋生態系、といった多くの要素モデルを含む（図 2.19）。大気モデルは力学コアと物理スキームから構成され、力学コアでは運動方程式や熱力学の式を用いて各格子点における物理量を予報する。物理スキームでは、力学コアの方程式の適用対象外、もしくはモデル格子以下（サブグリッド）の現象がもたらす熱・運動量・水物質の輸送や相変化を計算する。特に放射、雲・降水、乱流プロセスが重要である。天気予報のような短期積分では初期値が重要であり、モデルの改良とともに観測データを利用してモデルの第一推定値を修正するデータ同化システムが重要である。一方、気候予測のような長期積分では、海洋・海氷をはじめとする要素モデルの重要性が相対的に増す。

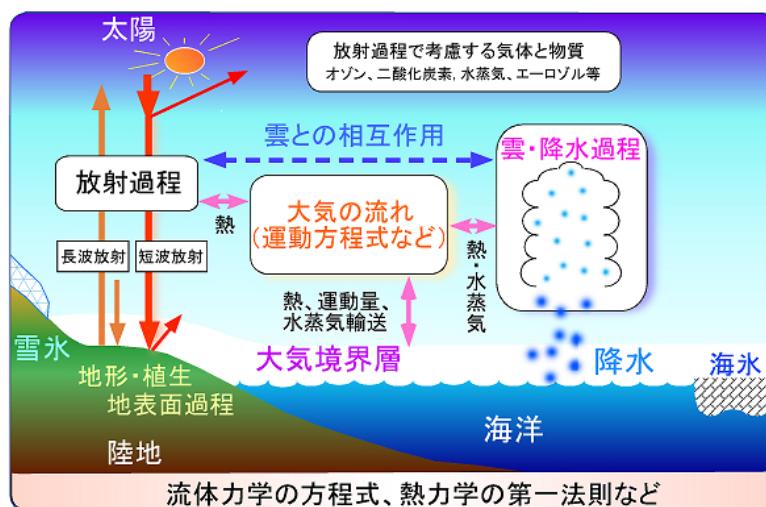


図 2.19 気象・気候モデルで考慮されているプロセスの模式図。気象庁ホームページ (<https://www.jma.go.jp/jma/kishou/know/whitep/1-3-1.html> ; 2022年7月7日閲覧) より。

気象・気候モデルに共通する特徴として、ソースコード（多くの場合は Fortran）の総行数が $O(10^5 \sim 10^6)$ 程度と大規模であるとともに、演算密度の低い (=高 B/F 比の) コンポーネントが多数あり、特に力学コアや 3 次元物理スキーム・要素モデルでは時間ステップごとに全通信や袖通信を必要とする。また、どこかのコンポーネントに浮動小数点演算量と処理時間が偏っていることがない (=Flat Profile)。

その結果、データを移動させる時間や最適化の甘い計算ループがソースコードの至る所でボトルネックとなっている。対策としては、浮動小数点演算の低精度化やデータの局所化といったデータ移動を削減する高速化施策が積極的に進められてきた。また、出力データの増大（ペタバイトクラス）も大きな課題であり、シミュレーションから解析までを同じプラットフォーム上で行うことも検討されている。

気象・気候モデルが目指す方向性として、大アンサンブル化、高解像度化、長期積分化、複雑化という軸がある。この中で大アンサンブル化は capacity computing であることから達成は比較的容易であり、今後も急速な発展が見込まれる。アンサンブル数に比例して増大する出力データサイズおよびファイル数への対応や、解析・可視化における並列化が課題である。高解像度化については基本的には capability computing であり、格子ステンシル計算のような演算が局所化されたモデルであれば、格子数に比例してノード数を増加させる（weak-scaling 問題）ことで技術的には比較的容易に達成できる。ただし、一般的に高解像度化によって時間刻みを小さくする必要があるため、計算時間が増加することに注意する必要がある。また、高解像度化に応じて利用する数値解法や物理スキームを再検討する必要がある（後述）。長期積分については現在、水平格子サイズが $O(10^2 \text{ km})$ の気候モデルを利用する事が一般的であり、高解像度モデルの長期積分は計算時間がかかりすぎるという問題がある。全体の問題サイズを変えずにノード数を増加させる（strong-scaling 問題）必要があるが、これにより並列性確保が困難になるとともに通信時間が増大するため、計算時間の削減は限定的である。複雑化については、予報変数の増加によるメモリ使用量増大への対処や、異なるタイプのモデルを連成させるためのカップラ技術の開発が課題である。

気象・気候モデルは、扱う対象に応じて全球モデルと領域モデルに分類される。全球モデルは天気予報から気候予測まで幅広く用いられ、領域モデルは天気予報や全球モデルからの力学的ダウンスケールなどの用途で用いられる。また、高解像度化による現象の忠実な再現を重視するか、実用的な気候予測情報の創出を重視するか、日々の天気予報の改善を重視するか、といった観点から、気象・気候分野では複数のタイプのモデルが開発されてきた経緯がある。以下では便宜上、(1) 全球高解像度モデル、(2) 全球気候モデル、(3) 領域気候モデル、(4) 現業数値予報モデル、に分けてモデルの特徴と課題を述べる。

(1) 全球高解像度モデル（4.10.1.1 節）は、数 km もしくはそれ以下の格子サイズを用いた全球モデルである。力学コアとして格子ステンシル計算を採用し、高い weak-scaling 性能を確保している。また、積雲スキームといった任意性を生む物理スキーム（パラメタリゼーション）を極力排除し、物理法則に沿ってスケール間相互作用をシームレスに取り扱うことが強みである。国内では NICAM、MSSG、SCALE-GM などが開発されており、ここでは代表アプリとして NICAM を取り上げる。NICAM は各時代のフラグシップマシンにおいて世界最高解像度のシミュレーションを達成し、富岳では全球 200m シミュレーションに挑戦中である。また、 $O(10^3)$ クラスの大アンサンブルシミュレーションによる延長予報・季節予測研究も活発である。気候モデルとしての利用は発展途上であるが、特に欧米ではモデルの全面的な書き換えを含む高速化の取り組みが急速に進んでいる。

(2) 全球気候モデル（4.10.1.2 節）は、気候の再現および予測を重視した全球モデルである。積雲対流スキームをはじめとする半経験則に基づく物理スキームを多用し、地球のエネルギーバランスを保つようパラメータ・チューニングされている。これにより、海洋モデルと結合して 100 年以上の積分を安定的に実行することが可能である。IPCC 報告書のベースとなる気候予測データの創出やイベント・アトリビューション研究、領域モデルの親モデルとしての利用など、温暖化の影響評価や適応策の研究において幅広く用いられている。さらに、生態系の働きなどを組み込んだ地球システムモデルを温暖化の緩和抑制

策へ用いる動きも活発である。このタイプのモデルは MIROC、および MRI-ESM が国内主要モデルである。これらのモデルは一般にスペクトル法を用いており、特に高解像度化した場合に集団通信にかかる時間の占める割合が増大する。

(3) 領域気候モデル（4.10.1.3 節）は、日本域など領域を限定して総演算量を削減し、その代わりに空間解像度を高めて、精度の高いシミュレーションを行うモデルである。モデルとしての特徴は(1)と同様であり、格子ステンシル計算であることから weak scaling 性能は高い。解像度は、現在は数 km 程度であるが、空間詳細な地形や土地利用を表現したり、積雲を格子で解像するためには数百 m 以下が、都市域の熱環境評価ではさらに細かい数十 m 以下が必要とされており、水平・鉛直分割によるノード増加が有力な高速化手法である。高解像度における領域気候のパフォーマンスへの影響調査や適切なスキームの検討が今後の課題である。また、現在、領域モデルは、将来気候変化に対する適応策検討への貢献が大変期待されている。モデルで計算された気候予測情報の信頼性向上のため、予測結果に含まれるシナリオやモデルエラーなどに由来する不確実性の定量的な理解を進めていくことも重要である。代表アプリとして取り上げる SCALE-RM の他、NHM やその後継の asuca ((4) を参照)、WRF、CReSS などが領域モデルとして研究コミュニティでは広く用いられている。

(4) 現業数値予報モデル（4.10.1.4 節）は、実際の天気予報システムに組み込まれている気象モデルである。信頼性および迅速性を最も重視し、日々の予報サイクルを通じてモデルの再現性・予測性能は厳しくチェックされている。初期値が重要であることから、客観解析の基盤技術であるデータ同化（後述）の重要性が高い。確率分布の非ガウス性を考慮した $O(10^5)$ スケールのアンサンブル同化による顕著現象予測や、水平 100m 格子以下の超高解像度化による現象の再現性向上などが次のマイルストーンである。フラッグシップマシンは 10 年先の現業マシンと同程度の性能であり、フラッグシップマシンで得られた知見を現業に生かす Research to Operation (R2O) や、現業システム・データの研究現場への提供といった Operation to Research (O2R) が近年ますます重視されてきている。代表アプリとして領域モデル asuca を取り上げる。

以上の気象・気候モデルを構成する力学コアおよび物理スキームの高度化は、モデル全体の性能を決める重要な課題である。特に気象・気候モデルの高解像度化がさらに進んだ場合、各物理プロセスの前提となる仮定が崩れることで従来の手法の適応限界を超えてしまう場合が生じる。例えば、力学コアとして多くのモデルで採用されている 2 次精度の空間差分式では、数値誤差が LES におけるサブグリッドスケールの渦粘性項の寄与を上回ってしまうことが指摘されている [24]。また、物理スキームについては特に乱流、放射、および雲微物理過程の高度化を検討する必要がある（4.10.2.1 節）。乱流過程の計算については、境界層内の対流や乱流の取り扱いに関するグレーゾーン [25] に相当する水平メッシュサイズ 100m～1km が今後主流となってくるため、このスケールでの乱流の取り扱いについて検証や知見を積み重ねる必要がある。さらに解像度 100m スケール以下の LES を気象気候に適用することを見据え、湿潤過程を考慮した理論の再構築が必要である。

放射過程に関しては、気象・気候モデルにおいて広く用いられている放射伝達モデルは、平行平面大気を仮定して、光の伝達方向を上下方向のみに近似する 2 流近似を用いているものが多いが、LES モデルが対象とする高解像度スケールにおいては、平行平面大気の仮定が成り立たず、斜めに光が入る効果を考慮する必要がある。LES スケールを念頭に、3 次元的に放射伝達を計算する 3 次元放射伝達モデルの開発をするとともに、どの程度の解像度から 3 次元的な放射を考慮しなければならないかを明確にすることに

も取り組んで行く必要がある。また、現状の気象モデルでは、雲の中における氷相の雲粒の微物理特性（雲粒の大きさや形状、密度）を経験的に仮定しているのが現状であり、このことが計算の不確実性を生む原因となっている。氷相の微物理特性を直接計算する雲モデル（Habit 予測型ビン法 [26] や、超水滴法 [27, 28] の開発を進めていく必要がある。さらに、雲に関連して気象雷モデルの実用化も今後進展すると見込まれる。これらはいずれも予報変数すなわち必要メモリ量の増大を招くとともに、物理スキームの3次元化に伴う全通信が計算速度のボトルネックとなりうる。ターゲットとなるスケールよりもさらに解像度の高い超高解像度実験や詳細で非常に計算量の多い物理過程スキームによる実験結果をベースに新しい理論の構築をしたり、それらを教師データとしてスキームそのものを機械学習で置き換えるという試みも想定される。

大気化学モデル（4.10.2.2 節）や海洋モデル（4.10.2.3 節）も気象・気候モデルを構成する重要な要素モデルである。大気化学モデルはエアロゾルを含む様々な化学物質の動態を計算するコンポーネントである。特に気候・化学相互作用や大気汚染の観点から重要であるとともに、最近は COVID-19 や原発事故に対する高い即応性も求められている。国内では MIROC-ESM および MRI-ESM が主要モデルであり、ここでは代表アプリとして MIROC-ESM に含まれる CHASER (MIROC-CHASER) を取り上げる。予報化学種のトレーサー数に応じて計算負荷が高まり、特に移流過程と光化学反応過程についてアルゴリズム改良が今後の課題である。

海洋は大気に比べて大きな熱容量を持つとともに、CO₂ やメタンを吸収するため、長期の気候変動を扱う場合に海洋モデルは不可欠な要素モデルである。特に、顕著な南北熱輸送を担う黒潮やメキシコ湾流・北大西洋海流を正確に表現するためには水平解像度 10km 程度が必要とされ、最新の気候モデルにおいても未だ実用化されていない。また、温暖化の影響を強く受ける北極海や日本近海の海流情報に対する産業界からのニーズの高まりを受け、北極海だけを高解像度化した海洋ネスティングモデルや、日本近海を 500m 程度まで高解像度化としたモデルの開発も進められている。海洋 100m 格子モデルについては非静力学コアの導入を検討する必要がある。

データ同化による初期値の推定は、数値天気予報における最も重要な技術の一つである。特に、災害をもたらす顕著現象を精度よく予測するためには、モデルおよび観測の高い時空間解像度、およびアンサンブル同化システムの大アンサンブル化が必要となる。また、確率分布の非ガウス性を考慮した新しい手法（粒子フィルタ、機械学習）の検討が進められている。さらに、計算コスト削減のため単精度演算や超解像技術の活用、モデル物理過程の機械学習による置き換え、機械学習を用いたデータ同化手法の改善、といった点も検討課題である。ここでは代表アプリとして NICAM-LETKF を取り上げる。アンサンブルデータ同化は行列演算であり、アンサンブル数の増大とともに演算量は劇的に増大する。また、データ量だけでなくファイル数の増大にも耐える必要がある。データ同化・データサイエンスについては 4.10.3 節で詳述する。

以上のような課題は、気象・気候モデル開発者だけでは解決は難しく、技術トレンドとのギャップを橋渡しする必要もある。従って、コデザインを通じた気象・気候モデル開発者とハード開発者の協調設計は、これまで以上に重要になってくるだろう。技術トレンドと気象・気候モデルの対応については 4.10.4 節に詳述する。また、コデザインのプロセスに携わることができる気象・気候モデル開発者が不足しているという問題もあり、大学や学会において人材育成に向けた取り組みが行われている。

2.10.4 ロードマップ

気象・気候分野における課題のロードマップを図 2.20 に記載した。各課題の詳細は第 4.10 節に記載されている。

2.10.5 必要な計算機資源

2032 年頃における課題解決のために必要な計算機資源の見積りを行い、代表的アプリケーションについて、2.13 節の表 2.2 に記載した。また、計算機資源の見積りの詳細は第 4.10 節に記載している。

参考文献

- [1] Imada, Y., M. Watanabe, H. Kawase, H. Shiogama, and M. Arai. The July 2018 high temperature event in Japan could not have happened without human-induced global warming. *SOLA*, Vol. 15A, pp. 8–12, 2019. <https://doi.org/10.2151/sola.15A-002>.
- [2] Richardson, L. F. *Weather prediction by numerical process*. Cambridge University Press, 1922.
- [3] Charney, J. G., R. Fjörtoft, and J. Neumann. Numerical integration of the barotropic vorticity equation. *Tellus*, Vol. 2, No. 4, pp. 237–254, 1950.
- [4] 気象庁. 数値予報 60 年誌. 60 pp., 2019. https://www.jma.go.jp/jma/kishou/know/whitep/doc_1-3-2-1/all.pdf.
- [5] Manabe, S. and R. T. Wetherald. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.*, Vol. 24, No. 3, pp. 241–259, 1967. [https://doi.org/10.1175/1520-0469\(1967\)024<0241:TEOTAW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2).
- [6] Manabe, S. and R. T. Wetherald. The effects of doubling the CO₂ concentration on the climate of a general circulation model. *J. Atmos. Sci.*, Vol. 32, No. 1, pp. 3–15, 1975. [https://doi.org/10.1175/1520-0469\(1975\)032<0003:TEODTC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2).
- [7] T. Hajima, M. Watanabe, A. Yamamoto, H. Tatebe, M. A. Noguchi, M. Abe, R. Ohgaito, A. Ito, D. Yamazaki, H. Okajima, A. Ito, K. Takata, K. Ogochi, S. Watanabe, and M. Kawamiya. Development of the MIROC-ES2L earth system model and the evaluation of biogeochemical processes and feedbacks. *Geosci. Model Dev.*, Vol. 13, No. 5, pp. 2197–2244, 2020.
- [8] M. Kawamiya, T. Hajima, K. Tachiiri, S. Watanabe, and T. Yokohata. Two decades of earth system modeling with an emphasis on model for interdisciplinary research on climate (MIROC). *Prog. Earth Planet. Sci.*, Vol. 7, No. 64, 2020. <https://doi.org/10.1186/s40645-020-00369-5>.
- [9] Meehl, G. A., G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer. The coupled model intercomparison (CMIP). *Bull. Amer. Meteorol. Soc.*, Vol. 81, No. 2, pp. 313–318, 2000.
- [10] Bauer, P., B. Stevens, and W. Hazeleger. A digital twin of Earth for the green transition. *Nat. Climate Change*, Vol. 11, No. 2, pp. 80–83, 2021. <https://doi.org/10.1038/s41558-021-00986-y>.
- [11] Stevens, B. and Coauthors. DYAMOND: the DYnamics of the Atmospheric general circula-

- tion Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.*, Vol. 6, No. 61, 2019. <https://doi.org/10.1186/s40645-019-0304-z>.
- [12] Slingo, J. and Coauthors. Ambitious partnership needed for reliable climate prediction. *Nat. Climate Change*, Vol. 12, No. 6, pp. 499–503, 2022. <https://doi.org/10.1038/s41558-022-01384-8>.
- [13] 気象庁. 2030 年に向けた数値予報技術開発重点計画. 41 pp., 2018. https://www.jma.go.jp/jma/press/1810/04b/nwp_strategic_plan_towards_2030_2body.pdf.
- [14] IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Vol. In Press. Cambridge University Press, Cambridge, United Kingdom and New York, NY, US, 2021.
- [15] Wilby, R. L. and S. Dessai. Robust adaptation to climate change. *Wea.*, Vol. 65, No. 7, pp. 180–185, 2010. <https://doi.org/10.1002/wea.543>.
- [16] Hawkins, E. and R. Sutton. The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, Vol. 90, No. 8, pp. 1095–1107, 2009. <https://doi.org/10.1175/2009BAMS2607.1>.
- [17] Hawkins, E. and R. Sutton. The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dyn.*, Vol. 37, pp. 407–418, 2011. <https://doi.org/10.1007/s00382-010-0810-6>.
- [18] Mizuta, R. and Coauthors. Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteorol. Soc.*, Vol. 98, No. 7, pp. 1383–1398, 2017. <https://doi.org/10.1175/BAMS-D-16-0099.1>.
- [19] Ishii, M. and N. Mori. d4PDF: large-ensemble and high-resolution climate simulations for global warming risk assessment. *Prog. Earth Planet. Sci.*, Vol. 7, No. 58, 2020. <https://doi.org/10.1186/s40645-020-00367-7>.
- [20] Miyamoto, Y., Y. Kajikawa, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita. Deep moist atmospheric convection in a subkilometer global simulation. *Geophys. Res. Lett.*, Vol. 40, No. 18, pp. 4922–4926, 2013. <https://doi.org/10.1002/grl.50944>.
- [21] Prein, A. F., R. M. Rasmussen, D. Wang, and S. E. Giangrande. Sensitivity of organized convective storms to model grid spacing in current and future climates. *Philos. Trans. R. Soc., A*, Vol. 379, No. 2195, p. 20190546, 2021. <http://doi.org/10.1098/rsta.2019.0546>.
- [22] Sato, Y., Y. Miyamoto, and H. Tomita. Large dependency of charge distribution in a tropical cyclone inner core upon aerosol number concentration. *Prog. Earth Planet. Sci.*, Vol. 6, No. 62, 2019. <https://doi.org/10.1186/s40645-019-0309-7>.
- [23] Sato, Y., S. Hayashi, and A. Hashimoto. Difference in the lightning frequency between the July 2018 heavy rainfall event over central Japan and the 2017 northern Kyushu heavy rainfall event in Japan. *Atmos. Sci. Lett.*, Vol. 23, No. 1, p. e1067, 2022. <https://doi.org/10.1002/asl.1067>.
- [24] Kawai, Y. and H. Tomita. Numerical accuracy of advection scheme necessary for large-eddy simulation of planetary boundary layer turbulence. *Mon. Wea. Rev.*, Vol. 149, No. 9, pp. 2993–3012, 2021. <https://doi.org/10.1175/MWR-D-20-0362.1>.

- [25] Wyngaard, J. C. Toward numerical modeling in the “Terra Incognita”. *J. Atmos. Sci.*, Vol. 61, No. 14, pp. 1816–1826, 2004. [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNTMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNTMITT>2.0.CO;2).
- [26] Hashino, T. and G. J. Tripoli. The Spectral Ice Habit Prediction System (SHIPS). part III: Description of the ice particle model and the habit-dependent aggregation model. *J. Atmos. Sci.*, Vol. 68, No. 6, pp. 1125–1141, 2011. <https://doi.org/10.1175/2011jas3666.1>.
- [27] Shima, S., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara. The super-droplet method for the numerical simulation of clouds and precipitation: a particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Quart. J. Roy. Meteor. Soc.*, Vol. 135, No. 642, pp. 1307–1320, 2009. <https://doi.org/10.1002/qj.441>.
- [28] Shima, S., Y. Sato, A. Hashimoto, and R. Misumi. Predicting the morphology of ice particles in deep convection using the super-droplet method: Development and evaluation of SCALE-SDM 0.2.5-2.2.0, -2.2.1, and -2.2.2. *Geosci. Model Dev.*, Vol. 13, No. 9, pp. 4107–4157, 2020. <https://doi.org/10.5194/gmd-13-4107-2020>.

年代 課題	2022 ~	2024 ~	2026 ~	2028 ~	2030 ~	2032 ~
全球規模での高解像度予測	全球LESモデル開発/全球200m実験 O(10 ³)アンサンブル予報実験 全球雲解像10年気候実験 全球雲解像地球システムモデル開発	全球100m実験(LES) O(10 ³)アンサンブル全球雲解像予報実験 全球雲解像100年気候実験 全球雲解像地球システム実験				
全球気候シミュレーション	適応データセット公開 ESMによる10年規模予測シス テム開発 地球システム・社会経済結合モ デル開発	適応データセット更新用モデル 高解像度化 ESMによる10年規模予測シス テム実験 地球システム・社会経済結合モ デル開発		次期適応データセット仕様検討、モ デル開発 ESMによる10年規模予測シス テム実験 地球システム・社会経済結合モ デル開発		
気象予報	気象庁システム	1km狭領域LFM試験運用ハイブ リッドデータ同化導入 LFM予報時間延長 様々な観測データの利用拡大・利 用手法改良(断続的に実施)	LFM高解像度化 LEPS運用開始	LFM改良 LEPS改良 アンサンブルデータ同化導入		
	超高解像度シミュレーション	水平解像度<100m級LESの利 用		水平解像度<50m級LESの利 用		
	アンサンブルデータ同化	asuka-based EnVar, EnKFの開 発	全現業データ同化(5km) 高解像度化(~1km)	非ガウス性10万メンバ ー同化		
領域気候予測		O(10 ³ m)でアンサンブル気候計算 O(10 ² m)で現在気候再現実験	O(10 ² m)で気候実験		O(10 ² m)でアンサンブル 気候計算	
物理モデル	3次元放射伝達モデル (3D-RD)	3D-RDの開発 気象モデルへの実装	結合モデルを用いた 小スケールでの数値実験	実問題での数値実験		
	精緻な雲モデルの開発	精緻な雲モデルの開発と実事例での実行に向けた高度化		実問題での数値実験		
	気象雷モデルの実用化	ハインドキャスト実験と高速化 モデルの検証と改良	特定の地点を対象とした実証実験	日本全体を対象とした予測 全球モデルへの拡張		
大気化学		エアロゾルモデルの高度化	高解像度衛星データ用いた高解像度輸送検証	精緻化されたモデルでの O(1km)でのイベント実験	精緻化されたモデルでの 長期計算	
海洋 (長期気候予測)		海洋全球10kmの気候モデル 北極3-5km/全球25kmの 気候モデル	左記モデルのアンサンブル実験	日本近海500m(多重ネスト)の海洋モデル 日本近海100mの海洋非静力モデル		
データ同化 機械学習		地球システムモデルによる 強結合データ同化システムの開発 高解像度データ同化手法の開発 (解像度100m以下) データ同化とAIの融合研究		地球システムモデルによる 強結合データ同化実験 高解像度データ同化手法の開発 (解像度10m以下) AIを使った新しいデータ同化手法の開発		

図 2.20 気象・気候分野のロードマップ

2.11 宇宙・天文

2.11.1 分野の概要

宇宙・天文の分野では地球の外で起こる物理現象全てを研究対象とし、多様な時間と空間スケールでの物質の進化を取り扱うことになる。その空間のレンジは隕石等、地球よりも小さいものから、銀河といったような地球より圧倒的に大きなもの、最大のものとしては我々と因果関係を持ちうる時間・空間領域全体として定義される我々の宇宙そのものまで及ぶ。このような長い時間スケール、大きな空間スケールの現象は実験室で再現することが不可能であるので、数値シミュレーションを行うことが必須となる。

数値シミュレーションが正しいかどうか確かめるためには、望遠鏡で天体を観測し、数値計算の結果と比較する必要がある。望遠鏡は可視光だけを見るのではなく、電波、赤外線、紫外線、X線、ガンマ線などのさまざまな波長を使って観測が行われてる（多波長天文学）。現在活躍している望遠鏡を例として挙げると ALMA、JWST、すばる、XRISM、CTAなどがある。次の大型計画として TMT の建築が予定されている他、すばる2、ALMA2などアップグレードも行われる。近年では、ニュートリノや重力波も用いたマルチメッセンジャー天文学が注目を集めている。Super-Kamiokande、IceCube、KAGRA が稼働している。さらにニュートリノの検出能力を上げた Hyper-Kamiokande の建築が予定されている。IceCube は IceCube-Gen2 にアップグレードされる。さらに、ダークマターの起源と考えられる未発見の粒子が起こす天体现象も熱心に探索されている。

宇宙の大きな空間スケールでは長距離に働く引力である重力が支配的になる。他の力のうち、強い力と弱い力は短距離にしか働くかず、最後に残った電磁気力は引力と斥力が両方働くため、天体のような大きな塊では片方の力が卓越しない。結果として天体の運動は重力に支配されることになる。このような理由から、宇宙の研究は重力の研究と馴染み深い。複数の粒子や広がった密度分布が作る重力中の運動は解析的には解けないため、重力が作る構造・密度分布を予想するためには、一般に数値シミュレーションを行う必要があり、多くの未解決な問題がある。

星など、天体の内部では温度が高く、イオンが電離している。電子がほぼ同数存在し、全体としては荷電中性に保たれているが、こうした環境でも磁場は物質の運動に影響を与える。そのような磁気流体力学が現れるのは天体物理学の一つの特徴である。

宇宙・天文の分野でもう一つの重要なファクターは光である。例えば光のスペクトラムを見れば放射源の温度が分かるように、光がどのように放射され、伝搬するのかを理解することができれば、天体の様子を深く知ることができる。ただし、光は輻射輸送方程式にしたがって物質と相互作用しながら外部に放出されるため、一般にシミュレーションの計算コストが高い。計算するためには様々な近似手法が用いられている。ニュートリノも光と同様に計算することができるものの、ニュートリノ振動によって種類が変わる効果を追加で取り入れる必要がある。

本節では例として5課題を紹介するが、宇宙物理は本来は非常に幅広い分野であり、全てを網羅的に紹介するのは難しい。より広範な分野の説明は4.11節に譲る。

2.11.2 長期目標と社会貢献

古来より天文学・宇宙物理学は二つの側面から社会に貢献してきた。一つは天動説対地動説のような我々の世界観の理解という哲学的側面であり、もう一つは農耕や航海術に使われてきた実用的な側面である。ここではより具体的に天文学・宇宙物理学が社会に貢献する理由を4つに大きくわけ、その内容や目標を紹介したい。4つとはそれぞれ(1)宇宙史の探求、(2)新たな物理法則の探求、(3)宇宙における生命存在の意味の探求、(4)宇宙環境学の推進である。それぞれ詳しく見ていく。

(1) 宇宙史の探求 宇宙史は、宇宙がたどってきた歴史を研究するものであるが、現在の宇宙論は観測に支えられた精密科学となっている。なぜなら、遠くの宇宙をみることで過去の宇宙を直接見ることができるからである。例えば10万光年遠く(光年は距離の単位である)の天体から出た光は10万年かけて地球の望遠鏡に届き、10万年前のその天体の姿を知らせてくれる。

こうして築き上げられた宇宙史は、人類の自然観の醸成に貢献し、貴重な知的財産となる。例えばビッグバンやブラックホールという概念は、数十年にわたる学術的論争を経て確立したものであるが、その言葉は、今や子供から大人まで誰もが知るものとなっている。これらは、科学への興味をかき立てる原動力になっている。

現代の宇宙像と宇宙史を図2.21に示す。現在の宇宙年齢は137億年である。宇宙誕生から38万年後にあたる時刻の情報である宇宙マイクロ波背景放射の精密な観測により、初期の宇宙は非常に熱い火の玉(ビッグバン)であったとわかった。このときの宇宙の密度は比較的一様だが、その中でも密度が高いところが重力相互作用で収縮し、大規模な構造を作っていく。宇宙誕生から数億年経つと銀河が形成され、その中には $10^7\text{--}10^{14}$ 個の星が内包されている。このように宇宙は星や銀河など小さい構造の複雑な相互作用で大きな構造ができているが、現在のスーパーコンピューターでは銀河や星を一つ一つ解像しながら、大きな構造の進化を追うシミュレーションは不可能である。計算の規模を大きくし、小さなスケールの構造をなるべく解像しながら大きなスケールの構造の時間発展を追うのが長期的な課題となる。

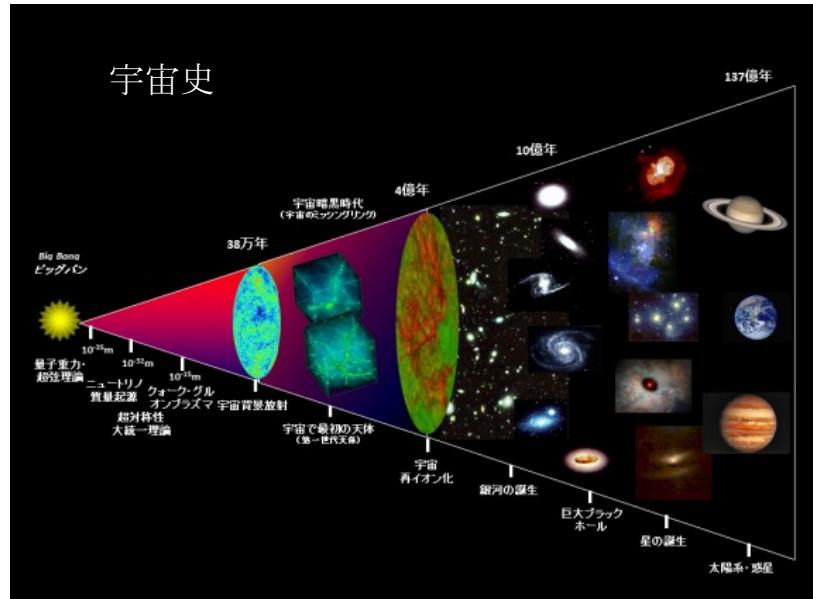


図2.21 現在の宇宙観

(2) 新たな物理法則の探求 宇宙の環境は地球の環境とはかけ離れている。そのため、地上の実験では到達できないエネルギーや空間スケールを支配する未知の物理法則を探求する実験場として優れている。物理学の基本法則の探究は、短期的な社会貢献が評価しづらい側面があるが、長期的に見れば、人間生活に大きな貢献をする可能性を持っている。例えば、一般相対性理論は、発表当時は宇宙の現象においてのみ重要な理論であり、我々の生活には無縁のものと考えられていた。しかしながら、科学技術の精度が高まった今日、地球重力場においても、一般相対論の効果が顕わになっている。GPSは、高度約2万kmを周回する4つの衛星によって機能しているが、GPS衛星では一般相対論の効果によって地上よりも1日45マイクロ秒早く時間が進む。この時間の補正をしないと、位置情報は1日で11kmもの誤差を生じてしまう。宇宙の基礎研究が長期的に見て、人間生活に大きな貢献をすることは今後も期待できることである。

本領域での長期的な課題は、ブラックホール、ニュートリノ、超高エネルギー宇宙線、ダークマターおよびダークエネルギーに関してその正体や発生メカニズムを特定することである。最新の実験や観測に基づいた既知の物理法則を数値シミュレーションに全て反映させること、そしてそのシミュレーション結果を観測と比較することでこれらの謎に挑んでいく。ここではブラックホールと超高エネルギー宇宙線について取り上げる。

ブラックホールはアインシュタインの提唱した一般相対性理論の予測する強い重力を持つ天体である。ある質量の星を圧縮していくと、その質量により決定されるクリティカルな半径（シュヴァルツシルト半径）より小さくなるぐらい星を圧縮すると、光すらも逃げ出せない強重力な領域が表てくる。それがブラックホールである。提案当初は単なる理論的な存在と考えられていたが、大質量星の進化の理論の発達とともに、そ

の存在可能性が現実味を帯び、現在では宇宙に多く存在する普遍的な天体であると考えられている。ブラックホールはその定義からして、本体が光を発することはないが、ブラックホールの周りに他の星が回っているとき、その星から流れてきたガスが円盤状にブラックホールを囲み、光を発する。その光は数多く観測されている。ブラックホールの存在の直接的な証明は、2015年の重力波の直接検出によりなされた。重力波は質量の大きな天体が高速で運動する際に生じる時空のさざ波であり、やはり一般相対性論が予想するものである。観測された重力波は二つのブラックホールが互いに回りつつ合体したケースの理論予測と一致している。図2.22にその様子を示した。この歴史的なイベントを経て、今後は重力波を用いて天体现象を観測する重力天文学が可能になる。中性子星連星の合体および超新星爆発はそれらを発する重要な候補天体である。ブラックホールはその質量によって分類され、それぞれの名称は太陽の3倍か



図2.22 ブラックホール連星合体の様子 ©SXS COLLABORATION

ら数十倍重い恒星質量ブラックホール、100倍から1000倍重い中間質量ブラックホール、 $10^6\text{--}10^9$ 倍重い超大質量ブラックホールとなっている。

宇宙に漂う高エネルギー粒子、”宇宙線”の存在は、粒子と大気との衝突から生じるエアシャワーや地球磁気圏のその場観測により、確認されているが、その生成メカニズムには依然として多くの謎が残っている。なぜ、エネルギー等分配から外れて一部の粒子が選択的に超高エネルギーまで加速されるのだろうか。地上で観測される宇宙線のうち 10^{15}eV (eV はエネルギーの単位) までは我々の銀河系内の超新星爆発の後にできた超新星残骸という天体で作られていると考えられているが、その具体的な加速メカニズムは未だに解決していない。更に 10^{15}eV を超える高エネルギーの領域になると、メカニズムだけでなく何処で加速されているのかについてもわかっていない。これらの高エネルギー粒子の起源は活動天体のダイナミックスの理解と密接に関連しており、非熱的粒子を生成する粒子加速の問題は、宇宙物理学・天文学・プラズマ物理学・宇宙線物理学などの学際的研究となっている。

(3) 宇宙における生命存在の意味の探求 「此処は何処で、我々は何者なのか?」、古くからの人類共通の問い合わせであるこの問題に迫ることが、本領域が担う知の探求の究極の目標である。その問い合わせにきちんと答えるには、地球の成立を追うだけではなく、第二の地球（太陽系以外にある地球と環境が似た惑星）を探査・検証する必要がある。地球以外に地球型の惑星は存在するのか、また地球以外の惑星に生命は存在するのか、これらは誰もが一度は疑問に思うことであり、その研究に対する社会的な期待は大きい。この研究は宇宙における生命存在の意味を再認識することへつながるばかりでなく、生命誕生過程の本質に迫るものである。本研究の長期的な課題は様々な惑星の形成・進化過程を解明し、我々の地球をそのシナリオの中に位置づけること、そして生命が生まれる環境の条件を精密化し、我々や我々以外の生命がどうして生まれて来たのか明らかにすることである。

現在標準だと考えられている、惑星の形成過程を図2.23に示した。上から下へ時間発展していく。惑星系が形成される母体となるのは原始惑星系円盤と言われるガスとダスト微粒子からなる円盤状構造である（上から一番目の状態）。これは中心星の周囲に形成される。そのダスト微粒子が衝突合体を繰り返し、微惑星（上から二番目の状態）、原始惑星（三番目の状態）、惑星（五番目の状態）とだんだん大きく育っていく。微惑星の状態ではガスとの相互作用を考える必要があり、微惑星が中心星に飲み込まれずに成長できるかは自明ではない。また原始惑星に育ったあの衝突合体も典型的な状況でのシミュレーションしか行われられておらず、ケーススタディが不十分である。上記のような複雑な形成過程を経て出来た惑星は質量、組成、中心星からの距離等がそれぞれ異なり、様々な個性を持っている。惑星形成後の時間発展を追うためには地球型惑星、木星型惑星等の違いを考慮にいれながら、天体内部および表層の構造分化と

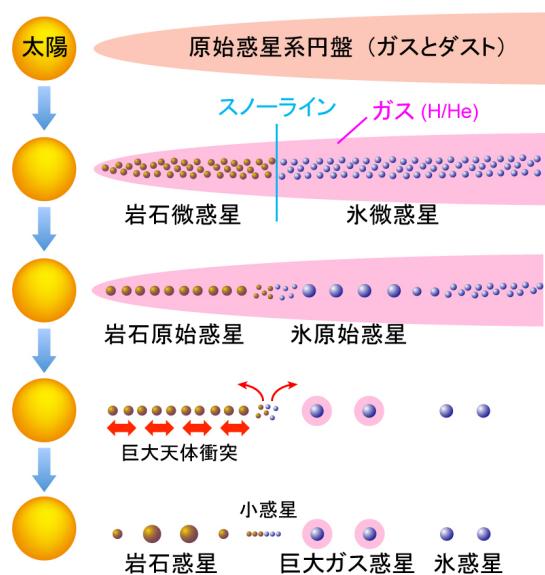


図 2.23 原始惑星系円盤から惑星形成までの時間進化

その時間発展を探求する必要がある。天体内部では対流運動や磁場を生み出すダイナモ作用、地球上の地形変形や地震等を生じさせるプレートテクトニクスなどが発現する。中心部の岩石の回りに気体をまとった天体では表層の研究も重要であり、中心星の活動度に対する、惑星気象・気候の多様性が熱心に研究されている。最後に惑星や衛星等で見られるであろうさまざまな表層環境における有機物化学を探求し生命の起源に迫る研究が続く。中心星の表面温度や活動度の違いがもたらす惑星表層へ入射する放射エネルギーの質の違い等も有機物の循環に大きな影響を与えるであろう。上記のようなシナリオの上に太陽～地球系での生命の発生を位置づけたい。

(4) 宇宙環境学の推進 ここで言う宇宙環境とは太陽系環境である。現在、人類の活動範囲は、地上から宇宙空間へと拡大している。人工衛星は、通信、テレビ、GPS 等で我々の生活に不可欠のものとなっている。宇宙環境は太陽活動によって大きな影響を受ける。太陽活動が活発になると、強い太陽風が生じ、現在我々の生活に欠かせない衛星の機能に大きな影響を与える。通信衛星の故障は、重要な通信機能の遮断になるだけでなく、制御不能に落ちいった衛星が地上に落下することで被害を及ぼすこともあり得る。また、太陽風によって生じた磁気嵐によって、変電所のコイルに大電流が流れ故障し、大停電が起きたこともある。したがって、太陽活動を予測できれば、これらを防災することができる。この目的で取り組みが進められているのが「宇宙天気予報」である。通信衛星、宇宙ステーション、惑星探査機等が、正確かつ安全に機能するためには、太陽系の環境科学の推進が必要である。

この分野の長期的な課題は太陽の内側、太陽の表面、そして地球の近傍の3つの領域にわたるプラズマの運動と磁気的な性質を明らかにすることである。近年、太陽黒点数と地球の気候変動との関係が注目されている。太陽黒点数は11年周期で変動しているが、2008年から増加に転じるはずであった黒点数が2009年後半まで増加せず、太陽活動に変調が生じている。17世紀には太陽黒点が長期に出現しなかった期間があり、今期の太陽活動周期の乱れがその前兆である可能性などが検討されている。太陽の様子は図2.24の左に表示した。黒点は磁場の強い場所に対応し、そこではフレアという爆発現象が起こる。太陽フレアの結果、高エネルギー粒子が大量に放射される。これが太陽風である。黒点活動を知るために、太陽における磁場の生成過程（太陽ダイナモ）を知る必要がある。また、太陽フレアでは、磁場に蓄えられたエネルギーが磁場のつなぎ換え（リコネクション）によって、プラズマに受け渡されて、高速の

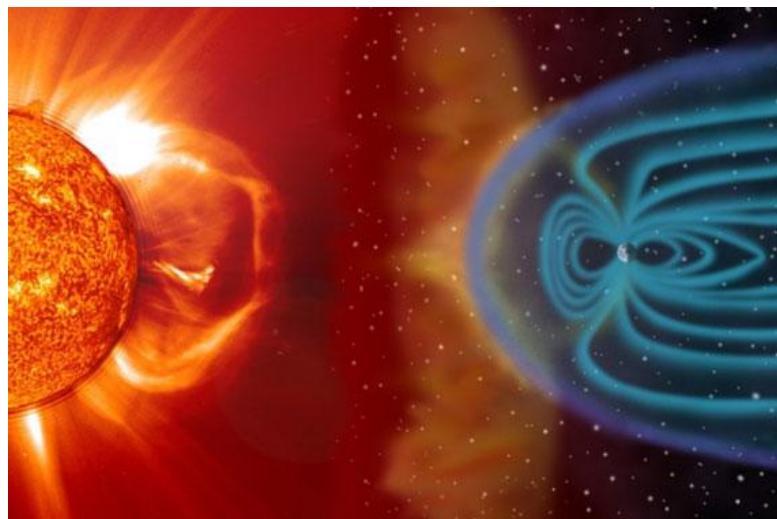


図 2.24 太陽（左）と地球磁気圏（右）。右の磁力線に囲まれた小さな球が地球を表している。©SOHO/LASCO/EIT NASA, ESA

粒子が生成されると考えられている。磁場のリコネクションは、まだ完全には解明されていない物理過程である。これは、磁気流体力学や電磁プラズマ粒子シミュレーションが必要であり、本格的な取り組みが始まったところである。また、太陽風が地球に接近すると、衝撃波が発生する。宇宙空間の密度は低いため、この場合の衝撃波は、地上で起こる衝撃波と異なり粒子同士の直接的な衝突ではなく、電磁場を介した運動量のやりとりで生じる衝撃波となる。これを無衝突衝撃波という。無衝突衝撃波のシミュレーションには、電磁プラズマ粒子シミュレーションが必要となる。この過程の様子は図 2.24 の右に表示した。

参考文献

- [1] Tomoaki Ishiyama, Francisco Prada, Anatoly A. Klypin, Manodeep Sinha, R. Benton Metcalf, Eric Jullo, Bruno Altieri, Sofía A. Cora, Darren Croton, Sylvain de la Torre, David E. Millán-Calero, Taira Oogi, José Ruedas, and Cristian A. Vega-Martínez. The Uchuu simulations: Data Release 1 and dark matter halo concentrations. *MNRAS*, Vol. 506, No. 3, pp. 4210–4231, September 2021.

2.12 人工知能

本節では、AI-for-Science (AI 技術の基礎科学研究への活用) について記載予定である。現在執筆・取りまとめを進めており、2024 年 3 月を目処にめどに本書と統合し、公開の予定である。

2.13 計算機要求性能まとめ

ここでは、2032年頃における課題解決のために必要な計算機資源見積りを、代表的アプリケーションに対する要求性能表としてまとめる。

第3章

アプリケーションの分類

第4章

各課題の詳細

4.1 素粒子・原子核

本節では素粒子・原子核分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。第 2.1 節で述べた素粒子・原子核分野における長期目標とこれらの課題は密接に関わっている。以下の各課題名には関係する長期目標を付与した。各課題間で共通の計算手法・アプリケーションを用いるものもあるが、計算規模やパラメータの違いから異なる計算機資源を要するものもある。GPU を利用する課題については課題ごとに触れている。格子 QCD については、GPU での実行もサポートするコードも広く使われており記述がなくとも多くの課題で GPU の利用が可能と考えられる。

2032 年ごろに各課題が必要とする計算機資源のうち総計算量 (EFLOP^{*1}) を表 4.1 にまとめた。基礎的な密行列線形代数の計算ライブラリ、ステンシル系の疎行列ベクトル計算、高精度（多倍長）浮動小数点数演算が必要なアプリケーション群となっている。また、格子 QCD バリオン 2 体波動関数計算^(c) では FFT を使用している。各課題の総計算量は広く分布しているが、大きいものは数十から数万エクサスケールの計算機を用いて半年から 1 年の計算量となっている。

また、直接的に長期目標や各課題に関連するものではないが、将来発展が期待される計算手法（機械学習、量子計算）と、本分野での格子 QCD 計算で利用されている基盤的な資源であるデータグリッド利用の動向について、それぞれ 4.1.21 節と 4.1.22 節にて説明した。

表 4.1: 素粒子・原子核分野における各課題の総計算量

長期目標	課題名	アプリケーション名（またはアルゴリズム名）	総計算量 (EFLOP)
(a-1)	格子 QCD を用いたハドロン相互作用の決定 (2.1.3(i) 節, 4.1.1 節)	格子 QCD ハイブリッドモンテカルロ法 ^(a) 、ウィルソン型フェルミオン作用 ^(b) 、大規模疎行列反復解法 ^(b) 、格子 QCD ハドロン多体波動関数計算 ^(c)	130,000,000 (2 体 S/D 波), 1,000,000,000 (2 体 P/F 波, 3 体それぞれ)
(a-1)	ハドロン階層の理解へ向けた格子 QCD 計算の進化 (4.1.2 節)	格子 QCD ハイブリッドモンテカルロ法 ^(a) 、ウィルソン型フェルミオン作用 ^(b) 、大規模疎行列反復解法 ^(b)	60,000,000
(a-1)	格子 QCD を用いた核子構造の理解 (4.1.3 節)	ウィルソン型フェルミオン作用 ^(b) 、大規模疎行列反復解法 ^(b)	14,100,000
(a-1)	深層学習を利用したハドロン散乱振幅の構築 (4.1.4 節)	部分波振幅をモデル非依存に求める深層学習プログラム	1.5
(a-2)	格子 QCD を用いた QCD 热力学 (2.1.3(ii) 節, 4.1.5 節)	格子 QCD ハイブリッドモンテカルロ法 ^(a) 、大規模疎行列反復解法 ^(b)	20,000
(a-2)	格子 QCD を用いた QCD 相構造の解明 (2.1.3(ii) 節, 4.1.6 節)	格子 QCD ハイブリッドモンテカルロ法 ^(a) 、ドメシイウォール型フェルミオン作用 ^(b) 、大規模疎行列反復解法 ^(b)	42,000,000
(a-2)	高エネルギー重イオン衝突実験（クォーク・グルーオンプラズマの解明）(4.1.7 節)	相対論的粘性流体方程式の数値解法リーマンソルバー、模型 (JAM, UrQMD など) によるイベントジェネレーター、ペイズ統計	60 – 600
(a-3)	少数多体系計算 (2.1.3(iii) 節, 4.1.8 節)	ガウス関数展開法 ^(e)	72 (6 体問題) – 151,200 (10 体問題)

*1 FLOP は浮動小数点演算回数。EFLOP は 10^{18} FLOP。フロップス (FLOPS:1 秒間に実行できる浮動小数点演算回数) との関係は $[FLOP]=[FLOPS \times 秒]$ 。1 年間で 1EFLOPS の性能を持つ計算機が実行できる理論的計算量は $1[\text{EFLOPS}] \times 365[\text{日}] \times 24[\text{時間}] \times 60[\text{分}] \times 60[\text{秒}] = 31,536,000 [\text{EFLOP}]$ 。

(a-3)	QCD に立脚した原子核反応計算 (2.1.3(iii) 節, 4.1.9 節)	離散化連続チャネル結合法 ^(e)	170 – 1,700
(a-4)	核力に基づく核物質の状態方程式 (2.1.3(iv) 節, 4.1.10 節)	クラスター変分法、連立積分微分方程式	150,000 – 15,000,000
(a-3)、(a-4)	配位間相互作用計算 (殻模型計算) (2.1.3(iv) 節, 4.1.11 節)	ランチョス法、モンテカルロ殻模型 ^(e)	厳密対角化法 (ランチョス法) : 50,000, モンテカルロ殻模型 ^(e) : 3,000,000
(a-4)	原子核密度汎関数計算 (2.1.3(iv) 節, 4.1.12 節)	密度汎関数法 ^(e) 、Hartree-Fock-Bogoliubov(HFB) 方程式、行列対角化	780,000
(a-4)	揺動散逸理論による重イオン反応・核分裂・新元素合成 (2.1.3(iv) 節, 4.1.13 節)	動力学模型、ランジュバン方程式	50,000 – 50,000,000
(a-4)	微視的クラスター模型 (2.1.3(iv) 節, 4.1.14 節)	反対称化分子動力学 ^(e)	360,000
(b-1)	新物理探索・解明のための精密格子 QCD シミュレーション (2.1.3(v) 節, 4.1.15 節)	格子 QCD ハイブリッドモンテカルロ法 ^(a) 、ドメシイウォール型フェルミオン作用 ^(b) 、大規模疎行列反復解法 ^(b)	88,000,000
(b-1)	QED の高次補正計算 (2.1.3(vi) 節, 4.1.16 節)	QED 高次摂動モンテカルロ積分計算 ^(d) 、4 倍精度実数計算、高度に最適化された 4 倍精度ライブラリ	160,000
(b-1)	ファインマン振幅自動計算 (4.1.17 節)	QED 高次摂動モンテカルロ積分計算 ^(d) 、4 倍精度実数計算、6 倍、8 倍精度実数計算、高度に最適化された 4 倍精度ライブラリ	1,200
(b-2)	超弦理論シミュレーションで探る時空の量子ダイナミクス (4.1.18 節)	ハイブリッドモンテカルロ法 ^(a) 、複素ランジュバン法、レフシェツ・シンプル法、密行列行列積、大規模疎行列反復解法 ^(b)	260,000
(b-3)	場の量子論の非摂動ダイナミクス (4.1.19 節)	複素ランジュバン法、レフシェツ・シンプル法、杉野格子作用、カプラン・カツ・ウンサル作用、大規模疎行列反復解法 ^(b) 、密行列行列積	105,000
(a-2)、(b-3)	テンソル繰り込み群による素粒子物理学および物性物理学の研究 (4.1.20 節)	特異値分解、行列行列積、HOTRG 法、ATRG 法	7,200,000

以下では、長期目標 (a) 強い力が作り出す階層構造の非摂動的理解の各課題、長期目標 (b) 標準理論を超えた物理の探索の各課題の順にそれらの詳細を述べる。

4.1.1 格子 QCD によるハドロン間相互作用の決定 [長期目標 (a-1) の課題]

課題概要と長期目標

陽子や中性子など物質を構成する粒子・ハドロンの間に働く力 (ハドロン間力) の決定は、物質の構造、そして宇宙の進化の歴史における元素の来し方行く末を解明する上で最重要課題の一つである。本課題の最終目標は、素粒子 (クォーク・グルーオン) の基礎理論である量子色力学 (QCD) に基づいた計算・格子 QCD 計算を行い、ハドロン間力を第一原理的に決定すると共に、素粒子・原子核・宇宙天文学という異なる物質階層のミッシングリンクを繋ぐことにある。

従来の原子核物理は主に実験情報を用いて現象論的に決定された核力に立脚しており、真の基礎理論である QCD との関係は不明瞭であった。またハイペロン力などそもそも実験的決定が困難であった相互作用も多い。本課題はこれらの状況に対して、QCD に直接基づいて原子核物理を構築するという、ブレイクスルーをもたらすものである。中でもハイペロン力や三体力の精密な決定は、高密度核物質の状態方

程式を決定する上でも必要不可欠であり、中性子星の中心部の構造や超新星爆発、元素合成のメカニズムを解明する上で重要な鍵を握っている。

ハドロン間力の解明は、世界各地の大規模実験・観測でも重要な課題となっている。なかでも J-PARC におけるストレンジネス核物理実験は、ハイペロン力の決定が最重要目標となっており、2030 年前後を目標とするハドロンホール拡張が実現すると、これまでアクセス不可能であったハイペロン力についても実験的情報が得られる見通しである。また 2017 年の重力波・電磁波による連星中性子星合体・キロノバ爆発・重元素合成の観測は、高密度状態方程式を支配するハドロン間力の重要性を改めて浮き彫りにした。今後 LIGO/Virgo/KAGRA 等による重力波観測により、桁違いの数の連星中性子星合体イベントが観測されると見込まれている。また近年の進展として、原子核衝突におけるハドロン相関測定実験によりハドロン間力の情報を取り出すことが可能となった。既に LHC ALICE 実験において格子 QCD の予言をサポートする実験データが得られており、今後 2030 年代初めまでの RUN3, RUN4 実験によりこれまでの 100 倍の統計量が得られる見込みである。さらに、Belle (II), BES III, LHCb などで、チャーム、ボトムクォークといった重いクォークを含むエキゾチックハドロン(共鳴)の発見が報告されている。これらの状態の理論的解明をするためにも、対応するハドロン間相互作用の解明が喫緊の課題となっている。

本課題では、これらハドロン間力の内、最も重要な成分である二体バリオン間相互作用 (S/D 波) について今後の計算の展望を与える。さらなる挑戦的課題は、二体バリオン間相互作用 (P/F 波)、そして三体バリオン間相互作用の決定である。また、エキゾチックハドロンの解明に対しては、様々なメソン間相互作用、メソンバリオン間相互作用を計算する必要がある。さらにより将来的な課題としては、これらのハドロン間相互作用について、格子 QCD + QED 計算によりアイソスピンの破れもフルに取り入れた計算を行うことが望ましいが、ここではアイソスピン対称極限での計算を考える。また、得られたハドロン間相互作用を基にした、原子核・ハドロンの構造計算や高密度状態方程式の導出も重要な課題であるが、これらのテーマは別課題でも扱われているのでここではその計算コストは議論しない。

解決に必要な計算手法とアプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ウィルソン型フェルミオン作用^(b)、大規模疎行列反復解法^(b)、格子 QCD ハドロン多体波動関数計算^(c)。上付き (a), (b), (c) については 2.1.3 節を参照。

計算量・計算資源

ハドロン間相互作用の決定のためには、物理的クォーク質量を用いて複数のハドロンを取り入れた格子 QCD 計算を行う必要があり、最も長距離の相互作用をもたらすパイオノン交換と比べて十分に大きい物理的空間体積 ($\sim (8\text{fm})^3$ 以上) が必要となる。また、格子 QCD の最終的な予言を行うには、格子間隔について、 $a \lesssim 0.1 \text{ fm}$ において 3 点以上の計算を行って連続極限を取ることが必要である。

計算資源の見積りは実績のあるセットアップ、すなわちウィルソン型フェルミオン作用、格子サイズ $L^3 \times T = 96^4$ 、パイ中間子質量 135 MeV、格子間隔 0.085 fm、物理体積 $(8.1\text{fm})^4$ を基に、格子間隔 0.064 fm, 0.043 fm を計算するとし、ゲージ配位数 2 万配位 (10^5 分子動力学時間) として見積もった。この計算は大きく分けて、1) ゲージ配位生成と 2) ハドロン多体波動関数の計算に分けられる。以下に、

それぞれについて詳細を述べる。これらの見積りの数値は 2.1.5 節にまとめている。なお本課題は相関関数計算におけるゲージ固定に GPU を活用しており、将来的には伝搬関数計算や縮約計算においても GPU による計算加速を検討している。

ゲージ配位生成

ゲージ配位生成は、主にクォーク伝搬関数の計算コストで占められる。クォーク伝搬関数の計算は、 $12 \times L^3 \times T$ 次元の連立方程式解法に帰着され、その計算は Krylov 部分空間反復法 (BiCGStab) により行う。この計算のコア部分は行列ベクトル積であり、その計算量は行列の次元、つまり格子サイズ $L^3 \times T$ に応じて増加する。解が収束するまでの反復回数は、行列の最小固有値に対応するクォーク質量の逆幕で多くなる。これを格子間隔 0.085 fm において配位生成した場合のコストは、1 分子動力学時間あたり 6 EFLOP であり、**ゲージ配位数 2 万配位 (10^5 分子動力学時間) 生成すると 6×10^5 EFLOP 必要となる**。物理的な体積を一定に保ったままで格子間隔を小さくしていくと、典型的には格子間隔の逆数の 7 乗の幕でコストが増大すると考えられており、**格子間隔 0.064 fm, 0.043 fm の計算に必要な計算量はそれぞれ 5×10^6 , 7×10^7 EFLOP となる**。

クォーク伝搬関数計算に必要な使用メモリ量は 4 次元体積 $L^3 \times T$ に応じて増加し、格子間隔に応じて **8TB, 25TB, 130TB** となる。生成するゲージ配位のファイル容量は 1 配位あたりそれぞれ 50 GB, 150 GB, 780 GB であり、2 万ゲージ配位および時空間 4 方向を回転してゲージ固定した配位を保存するには、ストレージ量としては **100 PB** が必要である。

ハドロン多体波動関数計算

ハドロン間相互作用を計算するには、ハドロンが異なる空間座標にある相関関数の計算を、時間方向に 20 点程度計算する必要があり、この計算には高速フーリエ変換 (FFT) が用いられている。また、クォーク場で書かれたハドロン演算子から相関関数を計算するには、クォーク伝搬関数の計算に加え、クォーク場の縮約計算を行う必要があるため、多くの計算量が必要になる。

ここではまず実績のあるバリオン 2 体系の S/D 波の波動関数計算を基に見積もる。格子間隔 0.085 fm, 格子サイズ 96^4 の 1 測定を行うのに必要な計算量は 0.6 EFLOP であり、計算量は測定数に比例して増加する。また、格子間隔を小さくしていくと、1 測定あたりの計算量は格子間隔の逆数の 4–5 乗の幕で増大することが予想される。配位あたり ($4 \times T$) 回の測定を行うとすると、**必要な計算量は、格子間隔毎に 5×10^6 , 2×10^7 , 1×10^8 EFLOP となる**。

使用メモリは格子サイズ 96^4 で 6 TB であり、空間方向の格子サイズと共に増大していく部分と、時空間方向の格子サイズと共に増大していく部分がある。ここでは時空間の 4 次元体積をもとに、格子間隔毎に、およそ **6TB, 20TB, 100TB と見積もる。必要ストレージ量は **280PB** である。** また、これらの計算では FFT を使用するため全対全ノード通信が必要になり、多くの通信量が必要となる。

これに対して、バリオン二体系の P/F 波の波動関数を計算するには、一般に all-to-all と呼ばれるクォーク伝搬関数を計算する必要があり、計算コストが増大する。対応して、クォーク場の縮約計算の計算コストも急速に増大する。最適な計算アルゴリズムは格子体積などに依存することが予想されるが、これまでの計算実績から考察すると、S/D 波と比べ 100–1000 倍以上の計算コストがかかると想定される。そこで当面は、 **$\mathcal{O}(1) \times 10^9$ EFLOP 程度の計算量**により、格子間隔 1 点で計算することが目標となる。

上記のようなバリオン二体力の導出を行った次のステップとしては、バリオン三体力の導出が重要な課

題となる。この場合、クォーク数が増えることから縮約計算の計算コストが非常に大きくなる。また、対象とする系によっては、all-to-all クォーク伝搬関数の計算と、対応する縮約計算が必要となる。それに加えて三体系は二体系と比べ統計誤差が大きくなるため、測定数も大幅に増大させる必要がある。二体系と比べた統計誤差の増大は、クォーク質量が軽くなるに従ってより厳しくなり、二体系 S/D 波計算と比べ少なくとも 1000 倍以上の計算コストがかかると想定される。そこで当面は、 $\mathcal{O}(1) \times 10^9$ EFLOP 程度の計算量により、格子間隔 1 点、クォーク質量を物理点よりやや重い点で計算することが目標となる。ただしハイペロン三体力はこれまでほとんど何も知られていないため、このような計算でも大きなインパクトをもたらすことができる。

エキゾチックハドロンに対応する計算としては、メソン間力、メソンバリオン間力も重要な課題となる。クォーク・反クォーク対生成消滅ダイアグラムを陽に含まない二体系については計算は比較的簡単であり、上記の二体バリオン間力計算よりも小さいコストで計算できると見込まれる。一方で対生成消滅を含む系や三体以上の結合を含む系の計算は挑戦的課題である。現在はクォーク質量が大きい領域において $I=1 \pi\pi$ ポテンシャルと ρ メソンの導出に成功した段階であるが、他のメソン間力・メソンバリオン間力に拡張し、またクォーク質量を軽くしていった場合の計算コストの振る舞いについては未解明な点が多く、今後の研究課題である。

4.1.2 ハドロン階層の理解へ向けた格子 QCD 計算の進化 [長期目標 (a-1) の課題]

課題概要と長期目標

素粒子標準理論の一部である QCD により強い力は記述されると考えられている。そうであれば、強い力で起こる全ての物理現象を QCD により説明できるはずである。格子 QCD の大きな目標の一つは、強い力で起こる物理現象を理解すること、つまり、既知の実験結果の再現である。この目標が達成されれば、強い力による未知の物理現象、例えば、新粒子探索などの研究に大きな貢献ができるようになるはずである。

近年の計算機性能の飛躍的な向上と格子 QCD 計算アルゴリズムの進展により、格子 QCD 計算の不定性は格段に小さくなり、格子 QCD は精密計算の時代へ入ろうとしている。特に基本的な物理量であるハドロン質量は誤差 1% レベルの結果が得られており、既知の実験値の厳密な再現のためには、ハドロン質量に対して 1% 程度の効果しかないと考えられている電磁気力の効果、及び、アップ・ダウンクォーク質量の違い、さらにストレンジクォークの次に重いチャームクォークの真空偏極効果を正しく評価する必要に迫られている。また、多くのハドロンは不安定粒子であり、その性質を正しく再現するためには崩壊先の終状態散乱を理解する必要がある。ハドロン質量計算と比べると、格子 QCD 分野でのハドロン崩壊・散乱過程の研究は進んでいないため、まだ精密計算の段階ではないが、近年の研究の進展により、2 体状態だけでなく 3 体状態を終状態に持つ崩壊過程の計算方法が確立されてきている。これらの方法を用いて、ハドロン崩壊過程の精密計算を行うのも重要な課題の一つである。その他にも、ハドロン階層の上の階層にある原子核の性質をクォーク・グルーオンから直接計算することも長期的な課題としてあげられる。

今後 10 年間の具体的な課題の一つに、アップ・ダウンクォーク質量の違いと電磁気力効果を取り入れ

た、ハドロン質量の精密決定がある。自然界では陽子が中性子よりも軽いため、陽子が安定に存在している。この事実は、電磁気力の効果のみでは説明できないことが知られており、陽子と中性子の質量差を精密に決定するためにはクォーク質量差と電磁気力の両方を取り入れることが重要である。また、ハドロン質量と同じく格子 QCD 計算では基本的な物理量である、 π 中間子と K 中間子崩壊定数の精密決定も重要な課題である。これらの崩壊定数の比は、標準理論とのズレが指摘されているクォーク混合行列（キャビボ-小林-益川行列）の行列要素 $|V_{us}|$ の決定に本質的な量であり、基本的な物理量精密測定を行うことで、標準理論を超えた物理探索研究にも貢献できる。

解決に必要な計算手法とアプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ウィルソン型フェルミオン作用^(b)、大規模疎行列反復解法^(b)。

上付き (a), (b) については 2.1.3 節を参照。

電磁気力の効果を取り入れたゲージ配位生成には、この他に QED を組み込んだハイブリッドモンテカルロ法が必要になる。

計算量・計算資源

ここでは、電磁気力の効果を取り入れた QCD モンテカルロシミュレーションを行うための計算量を見積もる。QED を取り入れた場合でも格子 QCD 計算のコア部分はクォーク伝搬関数の計算なので、格子 QCD 計算と同じ計算量の見積り方法を採用する。一方、電磁気力の効果により、有限体積効果が QCD のみの場合よりも非常に大きくなることが知られており、複数の体積を用いて無限体積極限への外挿が必要になる。さらに、現実的な素電荷の値で計算を行うと統計誤差を抑えることが難しいことも知られており、これに対応するため素電荷についての外挿が必要になる。この外挿のために、素電荷の値を現実よりも大きくした計算を複数点で行う必要である。これ以外にも、精密計算を実行するためには、アップ・ダウンクォークについて物理的な質量近傍でいくつかのパラメータで計算が必要である。まとめると、複数の体積・素電荷・クォーク質量でのモンテカルロシミュレーションが必要になる。

以下では、3 つの格子サイズ $256^4, 320^4, 512^4$ で、各々 5 パラメータの計算を実行するための計算量を見積もった。計算量見積りの基本にしたのは、QED を含まない現実的クォーク質量直上、格子サイズ 256^4 の格子 QCD モンテカルロシミュレーションである。このパラメータで 2,000 分子動力学時間のゲージ配位生成を行うと、計算量は 78,000 EFLOP である。格子 QCD モンテカルロシミュレーションの計算量は、典型的には一辺の長さの 5 乗および、分子動力学時間に比例するので、基礎とした計算量を各パラメータに合わせてスケールさせることで見積りを行った。さらに、アップクォークとダウンクォークを独立に扱う必要があるため、最終的な計算量を 2 倍にした。

各体積、パラメータで 2,000 分子動力学時間ゲージ配位を生成すると、全ての体積、パラメータでゲージ配位生成を実行するための計算量は 30,000,000 EFLOP である。物理量測定には、ゲージ配位生成と同程度の計算が必要だと仮定すると、総計算量は 60,000,000 EFLOP である。要求メモリ量は格子サイズに比例し、最大の格子サイズ 512^4 の場合、2 PB のメモリ量が必要になる。各パラメータで生成し

た 200 ゲージ配位を保存し、測定用のゲージ配位も全て保存した場合、**要求ストレージ量は 570 PB** である。現状ではクォーク伝搬関数計算に対して GPU の利用は可能であり、将来に向けてゲージ配位生成計算コード全体の GPU 化も検討している。

4.1.3 格子 QCD を用いた核子構造の理解 [長期目標 (a-1) の課題]

課題概要と長期目標

原子核を構成要素である陽子と中性子（2つを総称して核子）は、それ自体クォークとグルーオンにより構成されており、強い相互作用するクォーク・グルーオンの量子多体系である。核子の性質はこのクォークとグルーオンを自由度として記述される量子色力学（QCD）により決まる。核子の質量の約 99% は、ヒッグス機構由来ではなく QCD によって創られる複雑な真空構造によって創出されると理解されており、南部によるこのアイデア（南部理論）は QCD の第一原理計算（格子 QCD 計算）により厳密に検証されてきた。

核子の内部構造に関する研究は、1954 年のホフスタッターによる陽子形状因子の測定に始まり、QCD 理論の発見に至るマイルストーンであるが、60 年以上経過した現在においても、未だに核子の内部構造の全容理解には至っていない。特に最近では、核子の大きさ（平均二乗半径）が、電子-陽子弹性散乱の実験で見積もられたものと、ミュー粒子が電子の代わりに水素原子核に束縛したミュー水素原子の軌道半径を利用して評価したものと間に 4% ほどの有為な差異がある（核子の大きさのパズル）など、物質の構成要素の一つである核子の構造に関して未だ人智が及んでいない状況にある。核子の内部構造を反映しているのは、その構成要素であるクォークとグルーオンの力学「QCD」なのは当然である故、格子 QCD 計算によって、それらの問題が解明されるものと期待されている。

南部理論によれば π 中間子はクォーク質量が小さいことに伴うカイラル対称性の自発的破れにより出現する低エネルギー励起状態であり、そのため核子は π 中間子との強い結合を持ち、長波長の π 中間子を放出・吸収することができる。この“ π 中間子の雲”を纏うことで核子がより広がった構造を持つことになるが、格子 QCD 計算の格子空間サイズが十分大きくないと、長波長成分は著しく制限され、この効果が失われる。そのため物理点かつ巨大空間サイズでの格子 QCD 計算は必要不可欠であり、その条件を満たした格子 QCD 計算のみが、実験で指摘されている核子の大きさのパズルの解決に QCD の第一原理計算としてチャレンジする資格を有すると考えられる。

核子の大きさのパズル以外にも、核子構造に関連して幾つかの重要な課題がある。中性子が β 崩壊するまでの崩壊寿命は、従来から用いられている 2 種類の方法（崩壊後の陽子の数えるビーム法と崩壊せずに残った中性子を数える蓄積法）で実験的に測られているが、これらから得られる寿命の値が異なっている、中性子寿命の問題と呼ばれるパズルもある。中性子 β 崩壊は、クォークの素過程レベルではベクトル型 (V) と軸性ベクトル型 (A) の差、V-A 型の結合によって引き起こされ、核子軸性電荷 g_A はこの崩壊を特徴付ける物理量である。この核子軸性電荷 g_A は、最も精度が高く計算できるものとして、格子 QCD 計算の核子構造研究における計算精度の高さを測る指標として用いられており、現状、格子 QCD 計算において統計精度 1-2% での評価が可能となっている。しかし、実験の精度はこれよりも数十倍高く、残念ながら実験と比肩する精度の計算は現状ではまだ難しい。

一方、この中性子寿命の問題に関しては、今後の超冷中性子を利用した中性子崩壊の超精密測定実験により、通常の β 崩壊では現れない、「非 V-A 型」の中性子崩壊として、スカラー (S) 型とテンソル (T) 型の中性子崩壊の探索も計画されている。こういった標準模型を超えた新物理の探求には、標準模型の範囲内で決まる核子スカラー電荷 g_S や核子テンソル電荷 g_T のような核子構造を反映した物理量の情報が新たに必要となるが、それ自体は実験で直接測定することが困難な物理量でもある。そこで、標準模型の第一原理計算である格子 QCD 計算で、核子スカラー電荷 g_S や核子テンソル電荷 g_T を、核子軸性電荷 g_A と同程度の統計精度で評価することは、新しい実験結果から標準理論を超えた新物理を探るために一助になると期待されている。

さらに高精度なニュートリノ振動実験における主要な系統誤差となっているニュートリノ・原子核散乱の不定性を解決するため、T2K 実験などで新しいニュートリノ・核子散乱の実験が計画されているが、その実験解析に必要となる核子の内部構造の情報、特に軸性ベクトル型の核子形状因子の格子 QCD 計算による評価も重要である。また、核子をクォーク・グルーオンの量子多体系として扱うために必要な構造関数とそのモーメントに関する研究もさらなる核子の内部構造の理解には、欠くことができない。

解決に必要な計算手法とアプリケーション

ウィルソン型フェルミオン作用^(b)、大規模疎行列反復解法^(b)。

上付き (a), (b) については 2.1.3 節を参照。

クォーク伝搬関数の計算部分が全体の計算時間の 7 割以上を占める。計算アルゴリズムとしては、 $[8 \times 10^{11}] \times [8 \times 10^{11}]$ 要素をもつ大規模疎行列の逆解を反復法を用いて 10^{-8} 精度以下まで求めることになる。格子空間を計算ノードごとに分割することで、各ノードでは小行列計算に落としこみ、大規模並列計算を用いてこの逆解を $O(10^4)$ 回求める。領域分割法を用いた行列計算は効率化の面で有用である。さらに、AMA アルゴリズムと呼ぶ反復法における達成精度を効率的に変化させることで、全体の計算時間の大半を減らすことができる。

計算量・計算資源

核子構造研究のためには、これまでの格子 QCD 計算で無視されることが多かった QED 効果とクォーク非連結ダイアグラムを取り入れた精密計算が必要である。現実世界の電荷は小さいため、これをそのまま使った計算は難しい。そのため、複数の大きな電荷を用いた計算を行い、その結果を外挿することで現実的な電荷での結果を求める。また、核子構造計算で特に注意が必要な系統誤差は励起状態の寄与であり、この影響が十分抑えられていることを確認するために、複数の虚時間距離での計算が必要である。虚時間距離と共に増大する統計誤差を抑えるためには、大きな虚時間距離では非常に多くの計算が必要となる。

以上をふまえて、格子サイズ 256^4 の大体積での物理的アップ、ダウンドクォーク質量を用いた QCD+QED 計算の計算量を見積もる。計算量の大部分は、格子 QCD 計算と同じくクォーク伝搬関数の計算であるので、格子サイズ 160^4 での核子構造計算 QCD 計算を参考に見積りを行った。格子サイズ 256^4 で異なる虚時間間隔 3 点 (1.0, 1.2, 1.4 fm) の計算を行うと 820,000 EFLOP が必要である。こ

の計算を大きな QED 電荷 5 点で行うとすると 4,100,000 EFLOP となる。また、クォーク非連結ダイアグラム計算には、典型的に通常の計算の数十倍計算量がかかると予想されているため、最大の虚時間間隔の計算量の 20 倍としてクォーク非連結ダイアグラムの計算量を 10,000,000 EFLOP と見積もった。合計すると核子構造 QCD+QED 精密計算に必要な総計算量は **14,100,000 EFLOP** である。この計算の要求メモリ量は **500 TB** であり、要求ストレージ量は **12 PB** である。現在 GPU は利用していないが、今後 GPU 用計算コード移植を検討する。

4.1.4 深層学習を利用したハドロン散乱振幅の構築 [長期目標 (a-1) の課題]

課題概要と長期目標

ハドロン散乱実験データに見られる散乱振幅のピークなどの構造の起源、すなわち共鳴状態が形成されているのか、多数のハドロンチャンネルの結合が引き起こす闘効果（カスプ）なのか、などを特定することが求められている。多くの場合適当なモデルを仮定し、その枠内で散乱振幅を構築するため、結果は模型依存である。この難点を解決することを目指し、可能な限りモデル非依存の結果を得る試みが、最近になって深層学習の方法によってなされている。散乱振幅を複素エネルギーの関数と見た時、その特異点（極）がどのリーマン面のどの位置に存在するかによって、構造の起源を探ることができる。現在この手法は開発段階にあり、現段階では極がどのリーマン面に存在するか、その確からしさを求めるアルゴリズムが 2 チャンネル問題の場合に完成し、その有効性が示された。将来的にはこの手法をより多くのチャンネルに拡張し、さらに極の位置まで特定できることが望まれる。そのために、実験で与えられる 2 体散乱における断面積から、部分波振幅をモデル非依存に求めるための深層学習プログラムを構築するのに必要な、計算能力を概算してみる。

解決に必要な計算手法とアプリケーション

部分波振幅をモデル非依存に求める深層学習プログラム。

計算量・計算資源

断面積はエネルギーと角度の関数なので、そのデータ数はエネルギー点数と角度点数の積となる。SAID などの典型的な部分波解析ではエネルギー 50 点なので、ここでは角度で 180 点とることによって、 $K = 50 \times 180 = 9000$ 個のノードをインプット層に用意する。アウトプットでは、30 個程度の部分波振幅を各エネルギーで与える必要があるので、アウトプット層には $M = 30 \times 50 = 1500$ のノードを用意する。隠れた層のない最も単純な深層学習のトレーニングに必要な演算数は、これにバッチ数をかけることによって評価でき、ここでは 4096 を採用することにすると、(実際に必要な演算数を考慮して) $2 \times 1500 \times 4096 \times 9050 = 111 \times 10^9 = 111$ GFLOP となる。これは最も単純な場合の数で、より効果的なプログラムには隠れた層が必要となる。例えば 20000 ノードの層を用意すると、1 アップデート必要な演算数はおよそ 1.5 TFLOP となる。深層学習プログラムにはこれを繰り返し (エポック数) 完成することができる。この回数は 1000 と予想され、総演算数は **1.5 EFLOP** となる。現存する GPU、Nvidia

A100 80GB によれば 1 日で完了する計算である。より現実的な計算には（より多くの結合チャンネルの場合）さらに数 10-100 倍程度の性能を有することが望まれる。メモリ容量、ディスク容量については、特に大きなサイズを必要としない。

4.1.5 格子 QCD を用いた QCD 熱力学 [長期目標 (a-2) の課題]

課題概要と長期目標

ビッグバンによる開闢直後の初期宇宙や中性子星の中心部などでは、物質の温度や化学ポテンシャルが QCD スケールに達する超高温・高密度状態が実現する。このような極限的環境下では、QCD 真空の相転移であるカイラル相転移や非閉じ込め相転移などの多様な物性現象が起こることが知られている。これら超高温・高密度物質の物性的性質の研究においては、4.1.7 節で紹介する重イオン衝突実験と共に、格子 QCD に基づく物質の諸性質の数値解析が極めて重要な役割を果たす。

格子 QCD による有限温度・密度系の数値解析は、バリオン数密度がゼロの場合には通常のモンテカルロ法で遂行できる。この場合の数値計算については、最も基本的な熱力学量であるエネルギー密度と圧力（総称して状態方程式と呼ばれる）に関するスタッガード作用に基づく精密測定が 2014 年に終了した。ただし、これらの解析で使われるルートトリックと呼ばれる処方は連続極限に関して理論的な批判のある定式化であるため、連続極限が保証されたウィルソン作用を使った定式化に基づく数値計算が今後の重要な研究課題として残されている。ウィルソン作用に基づく熱力学量解析については、近年グラディエントフローを使った測定法の有用性が注目され、研究が進められている。

状態方程式以外の熱力学量の数値計算も興味深い研究課題である。特に保存電荷の高次キュムラントは、以下で述べるティラー展開法の係数であると共に重イオン衝突実験での観測量でもあるため広い分野から注目され、近年精力的に研究が行われている。キュムラントは次数が上がるに従い統計誤差が爆発的に増大するため、現状では高精度測定は 6 次程度までに留まっているが、より高次の精密測定を進めることは有限密度系の解析や重イオン衝突実験の理解などに大きな波及効果を及ぼす重要課題である。

一方、バリオン数密度が有限の場合には符号問題（作用の複素位相問題）が起こるため、通常のマルコフ連鎖モンテカルロ法が適用できない。この問題を解決し、有限密度系の数値解析を実現することは、有限温度密度 QCD 分野の重要な未解決問題である。有限密度系の研究に関しては、従来から行われているティラー展開法や再重み付け法などのモンテカルロ法に基づく研究に加え、最近は複素ランジュバン法やレフシェツ・シンプル法、別項で述べるテンソルネットワーク法といったモンテカルロ法を使わない全く新しい計算手法に関する研究が活発に行われており、今後の発展が期待される。

熱力学系の動的な性質の数値解析も、今後の重要な研究課題である。特に輸送係数（粘性係数など）は、重イオン衝突実験の理解に直結する重要な物理量だが、格子 QCD での数値解析は著しく困難で、信頼に足る計算結果は得られていない。この解析の実現には、二点相関関数の高精度測定が必要であることに加え、実時間への解析接続に伴う原理的な困難もある。相関関数の測定に関しては、グラディエントフロー法などの新しい解析法が使われる等の進展が見られる。これらの発展と、計算機能力の増大により、動的な物理量の数値解析も今後解決に向け前進することが期待される。

解決に必要な計算手法とアプリケーション

バリオン数密度がゼロの場合の格子 QCD 数値解析は、広く使われている格子 QCD ハイブリッドモンテカルロ法^(a)により遂行できる。この手法の詳細は 2.1 節の格子 QCD の解説に譲るが、計算の主要な過程を占めるのは巨大なベクトル空間（次元は 10^7 以上）上のベクトルと疎行列の線形計算（大規模疎行列反復解法^(b)）である。また、作用の局所性により、格子 QCD 数値解析は並列化に適している。格子 QCD のコードは世界各地のグループにより開発されており、様々な環境へのチューニングが進んでいるほか、一部のコードはオープン化されている。

上付き (a), (b) については 2.1.3 節を参照。

計算量・計算資源

有限温度系の数値解析では、異なる格子間隔での計算結果を使った連続極限への外挿に加え、空間体積を大きくする極限（熱力学極限）を取る必要がある。また、多くの熱力学量の計算で真空期待値の引き算が必要となるため、有限温度計算と同じパラメータでのゼロ温度計算を行わなくてはならない。このように、一つの温度に対する解析を実現するだけでも様々な格子間隔・格子点での計算が必要となることに加え、物理量の温度依存性を調べるためににはこのような解析を様々な温度に対して行う必要がある。

ウィルソン作用を用いた状態方程式の物理点シミュレーションについては、最近格子間隔 $a = 0.04, 0.06, 0.085 \text{ fm}$ での数値計算が試みられている。この計算では、固定された空間体積に対し、各格子間隔で約 10^3 のゲージ配位を生成し、物理量の測定が行われている。この解析を温度一点に対して行うのに必要な計算量は約 300 EFLOP である。このような解析を 3 種類の空間体積、温度 10 点に対して行うためには総演算量約 20,000 EFLOP が必要とされる。これらのゲージ配位を保存するのに必要なストレージは約 50 TB になる。必要メモリは 4 次元格子体積に比例するが、最も体積が大きい場合でも 0.6 TB 程度である。より測定の困難な高次キュムラントや、相関関数の精密解析を同様の設定で実現するには、更に 10 倍以上の統計量で測定を行う必要があり、より大きな計算資源が必要となる。

4.1.6 格子 QCD を用いた QCD 相構造の解明 [長期目標 (a-2) の課題]

課題概要と長期目標

我々の周りにある物質から宇宙の星に至るまで、その質量の起源は QCD の動力学に起因する。現在の宇宙ではクォークの振る舞いに現れるはずのカイラル対称性が、QCD の動力学により自発的に破れている状態にあり、クォークやグルーオンは陽子、中性子、パイオンなどのハドロンに閉じ込められる。宇宙の歴史を遡り、十分高温になるとカイラル対称性が回復し、ハドロンは溶解し、クォークやグルーオンが飛び回る世界となる。このような状態変化がある温度で急激に起こるか、あるいはゆるゆると起こるは現在の宇宙の成り立ちを理解するために重要な情報になる。最も極端なケースは相転移がある場合で、アップ (u)、ダウン (d) クォークの質量がゼロの極限では必ず起こることが理論的考察によりわかる。現実世界の u, d クォーク質量 (QCD 動力学の特徴的なスケールに比べとても小さい。現実世界のストレンジ

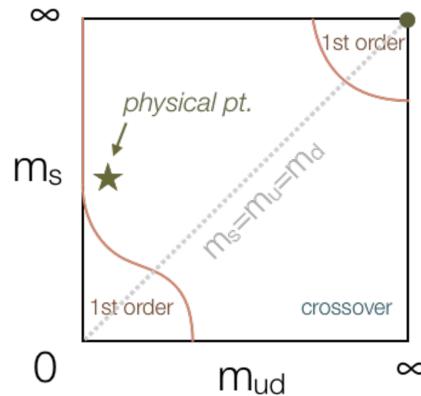


図4.1 コロンビアプロット。クォーク質量と相転移の有無の関係を図示している。

(s) クォーク点と併せて物理点とよぶ) や、その近傍での相転移の有無は、格子 QCD が答えうる重要な問題である。また、現実世界ではないが、クォーク質量ゼロの近傍の相転移の性質を明らかにすることで QCD が持つ対称性と相転移の深淵な関係の理解が進む事により、現実世界の転移温度近傍で起こりうる現象の解明にむすびつくことが期待される。また、この研究の一環として物理点の情報も詳細に調べることになる。これにより初期宇宙の(擬)相転移を規定する物理量や、ダークマター候補であるアクションの(擬)相転移近傍の性質も明らかになる。

クォーク質量と相転移の有無の関係を端的に示す図としてコロンビアプロット: 横軸に \$u, d\$ 縮退質量、縦軸に \$s\$ クォーク質量をとって一次転移とクロスオーバー(擬転移)の相転移領域を図示したもの(図 4.1)が用いられる。この相図の理解は近年大幅な変更を受けつつある。特に、\$s\$ クォーク質量を無限大に取って \$ud\$ クォーク質量がゼロの極限(カイラル極限)(\$N_f = 2\$)での性質、\$s\$ クォークの質量が物理点で \$ud\$ クォークの質量を物理点近傍からカイラル極限(\$N_f = 2+1\$)に取った時の性質、さらに \$uds\$ クォーク質量を等置した対角線上(\$N_f = 3\$)のカイラル極限領域の相構造の解明に多くの計算資源が充てられている。しかし、そこで重要なカイラル対称性を有限格子間隔でも具備する、所謂カイラルフェルミオンによる計算は限られている。次世代計算機の利用により、他と比べ高コストであるカイラルフェルミオンを用いたカイラル領域の計算による新しい知見の獲得が期待される。非カイラルな格子定式化でも、制御された連続極限によりカイラル対称性は回復することが期待される。さらに勾配流(グラディエントフロー)を用いて、連続極限に近い物理状況を実現する手法も研究されている。次世代計算機ではこのような手法の利用も考えられる。以下では最も理想に近い、最もコストの高いカイラルフェルミオンを前提として計算コストの算出を行う。

コロンビアプロットのカイラル領域とは逆の極限の計算も現在行われている。全てのクォークの質量が無限大では、グルーオンのみが動力学に関与する。この極限では格子 QCD 計算のコストは少なく、一次転移が起こることが確立されている。一方で、その一次転移がどこまで有限のクォーク質量領域までしみ出すかは、解明できていない。この領域では、クォークが重いが有限である事により、QCD の動力学スケールとは異なるスケールを扱う必要があるための困難がある。クォークが重いことを想定した ホッピングパラメタ展開などの手法を採用し研究が進められている。この手法の計算コストは軽いフェルミオンの計算に比べて低いため、コスト試算は省略する。

解決に必要な計算手法とアプリケーション

カイラルフェルミオンの実践的な手法としてドメインウォールフェルミオン (DWF) 型の作用^(b)を用いる。DWF は厳密な対称性を持つ究極の手法であるオーバーラップフェルミオン (OV) の良い近似であり、その近似の精度は連続極限に近づくほど良くなる。オーバーラップフェルミオンを用いたカイラル極限近傍の直接計算の現実的な方法がないため、中間格子間隔領域では、生成した DWF アンサンブルを再重み付け (reweighting) 法により OV アンサンブルに読み替える手法がしばしば用いられる。この方法は一定の成功を収めたが、より大きな体積、細かい格子では適用範囲が狭まり現実的な解決策にはならない。現状「富岳」で行われている計算の一部はさらに粗い格子を用いており、DWF の近似精度が悪く OV への再重み付けも効率が悪い（誤差の増加が著しい）。また、正しい連続極限の物理を得るために、より連続極限に近いシミュレーションが必要になるので、将来的な、精細格子、大体積シミュレーションでは、DWF を用いることになる。

DWF シミュレーションの実装は、質量が縮退した u, d クォークには通常のハイブリッドモンテカルロ (HMC) 法^(a)、残りの s クォークの寄与には HMC の一種である有理式ハリブリッドモンテカルロ (Rational Hybrid Monte Carlo, RHMC) 法を用いる。RHMC に類似の Exact One Flavor Algorithm (EOFA) の採用により若干の高速化が図れる可能性もあるが、既知の事例ではそれほどの差は認められないと、知見の蓄積がある RHMC の使用をここでは仮定する。RHCM 計算中、最も高コストなのが、フェルミオン行列の逆解きである。DWF のフェルミオン演算子の性質から CG 法に勝る逆解き手法が現時点では存在しない（大規模疎行列反復解法^(b)）。この逆解き解法には混合精度計算による高速化が有効であることが分かっている。次世代計算で用いる大体積では、さらにマルチグリッド法による高速化が期待されるが DWF ので実用化には至っていない。

上付き (a), (b), (c) については、2.1.3 節を参照。

計算量・計算資源

カイラルフェルミオンによる未知の領域探索であるため導入の必要な仮定が多いが、現時点での「富岳」で得られている知見と、他の計算コストが低い格子化手法の結果も踏まえて資源量を見積もる。

カイラル領域相構造探索では、クォーク質量の重い領域からはじめ、カイラル極限に攻めていく。十分重い領域では相転移が無く、クロスオーバーである。擬転移点を辿りながらクォーク質量を軽くしていくと、一次転移領域がある場合には臨界終点に、ない場合でもカイラル極限で臨界点に近づく。以下ではシミュレーション点には一次転移領域がない、つまり、極端な臨界減速が起こらず重いクォークの計算からのスケーリングが有効と仮定する。格子 QCD のシミュレーションコストはクォーク質量の減少と共に一般的に増加する傾向があり、これまでの擬相転移点付近での計算から経験的に

$$1 \text{ 分子動力学時間のコスト} \propto 1/m_q \quad (4.1)$$

と見積もる。計算コストの格子サイズ依存性は、分子動力学の時間刻みの調整も加味すると経験的に

$$1 \text{ 分子動力学時間のコスト} \propto (N_s^3 \times N_t)^{5/4} \quad (4.2)$$

となる。有限温度計算では時間と空間方向の格子サイズは非等方で、それぞれの格子点数を N_t, N_s ($N_s > N_t$) としている。自己相関長の増大の影響は、自己相関の知見の蓄積が不十分であること、考えるパラメータ領域では臨界減速が起こらないと仮定した試算であること、を踏まえて以下では無視する。

目標とするシミュレーションパラメータと計算コストは、 $N_f = 2 + 1$ を例に見積もる。物理点クォーク質量では、様々な結果から擬転移であることの示唆がある。そこで u, d クォーク質量は物理点質量から出発し、ここから軽くしてできるだけ臨界点に近づけるようにパラメータを選ぶ。ここでは以下の条件から目標パラメータ（最大コストケースのパラメータ）を算出する。

1. 格子間隔の逆数: $a^{-1} = 2.54 \text{ GeV}$ (既存のゼロ温度シミュレーションで用いられている格子間隔)
2. 最小クォーク質量: $m = m_{ud}^{phys}/3$ (s 質量の約 1/90), パイオン質量にして $m_\pi = 78 \text{ MeV}$.
3. 最大体積: そのパラメータでのゼロ温度のパイオン質量、崩壊定数の体積効果が高々 1% 程度

予想される擬転移温度^{*2} $T_{pc} \simeq 140 \text{ MeV}$ を実現するに必要な時間方向の格子点数（温度は時間方向の長さの逆数）は、 $a^{-1} = 2.54 \text{ GeV}$ では $N_t = 18$ である。体積効果を抑えるために必要な空間方向の格子点数は、ゼロ温度の体積効果の良い近似を与える有効理論 (Colangelo-Durr-Haefeli による有効理論) を用いて $N_s = 90$ になる。現時点で行われている擬相転移点付近でのシミュレーションパラメータ ([基準]) と上記の目標パラメータはそれぞれ

$$[\text{基準}] 32^3 \times 16 (\times 12(L_s)); \quad m = 3m_{ud}^{phys}, \quad (4.3)$$

$$[\text{目標}] 90^3 \times 18 (\times 12(L_s)); \quad m = m_{ud}^{phys}/3 \quad (4.4)$$

なので、式 (4.1), (4.2) により、1 分子動力学時間当たりのコスト比は $\text{Cost}[\text{目標}] \simeq 500 \times \text{Cost}[\text{基準}]$ となる。ここで、 L_s は DWF の仮想 5 次元サイズでカイラル対称性を制御するパラメータである。基準となる計算の演算数は、1 分子動力学時間あたり 0.06 EFLOP ^{*3} なので目標パラメータで 20,000 分子動力学時間の統計を得るために必要な演算量は

$$\text{目標パラメータ 1 点でのアンサンブル生成コスト} = 6 \times 10^5 \text{ EFLOP} \quad (4.5)$$

となる。生成配位を用いた各種解析には生成と同様なコストがかかる ($\times 2$) こと、最大体積で温度 3 点を確保すること ($\times 3$)、これより小さい体積、重いクォーク質量などで、同様なコストがかかるここと ($\times 2$) を考慮すると

$$(N_f = 2 + 1, L_s = 12 \text{ でのトータルコスト}) = 7.2 \times 10^6 \text{ EFLOP} \quad (4.6)$$

となる。ところで $L_s = 12$ DWF のカイラル対称性の破れは、実は $m_{ud}^{phys}/3$ と同程度である。この破れを完全に制御するため、この最大コスト点で、もう一つ $L_s = 24$ の計算を行う。最大コスト点での演算量 ((4.6) 式の半分) の 2 倍が追加になるため、 $N_f = 2 + 1$ での総コストは (4.6) 式の 2 倍になる

$$N_f = 2 + 1 \text{ でのトータルコスト} = 14 \times 10^6 \text{ EFLOP}. \quad (4.7)$$

^{*2} 他の格子フェルミオン作用を用いたシミュレーションから、カイラル極限 $m = 0$ での $T_c \simeq 130 \text{ MeV}$ 程度であることが知られており、 $m = m_{ud}^{phys}/3$ での T_{pc} はこれより若干大きくなる。DWF の「富岳」を用いた現時点でのカイラル凝縮からの結果は $T_{pc} \simeq 170 \text{ MeV}$ ($m = 3m_{ud}$) である。クォーク質量の減少と共に T_{pc} の減少も確認されており、上記の仮定を覆す発見は今のところない。

^{*3} 「富岳」2.2GHz(3.38 TFlops) で約 5 ノード時間、効率約 3% から試算。

ここでは、代表的な $N_f = 2 + 1$ の相探索を例に試算を行った。 $N_f = 2, 3$ では、詳細は異なるが、オーダーとしては同様なコストとなる。単純に 3 倍すると、総演算量は 42,000,000 EFLOP になる。メモリ使用量は概ね体積に比例する。現在の実測値をもとにスケールさせると、目標パラメータで 3TB のメモリが必要になる。ストレージは、10 分子動力学時間ごとに配位を保存するとして合計で 300 TB 程度必要になる。

4.1.7 高エネルギー重イオン衝突実験（クォーク・グルーオンプラズマの解明）[長期目標 (a-2) の課題]

課題概要と長期目標

クォーク・グルーオンプラズマ (QGP) の生成とその性質の定量的解明を目指し、高エネルギー重イオン衝突実験が世界各地で遂行されている。QGP 物理の困難さは、クォークやグルーオンを実験で直接観測できないことにあり、実験結果の理論的裏付けが重要である。そのため、計算機上で高エネルギー重イオン衝突実験の時間経過、すなわち QGP 生成、ハドロン化、終状態相互作用までを詳細に記述し、実験と理論の双方から QGP 物性を定量的に明らかにすることを目標にする。現在、そして少なくとも今後 10 年間にわたって、主流を担う現象論的模型は相対論的流体模型であると考えられる。この相対論的流体模型を中心として 2 つの発展が考えられる。一つは模型の枠組みそのものを吟味し改良していくことであり、もう一つはある程度確立した模型を用い高精度かつ高統計の計算を実行することである。

模型の枠組みの発展の点では、流体ゆらぎ、熱平衡化、流体化と行った新しい概念を取り入れた模型、低エネルギー領域、小さな系に適用可能な粒子描像と流体描像の両方を取り入れた模型といった発展がある。さらに磁場やカラー磁場を取り入れた本格的な数値計算も開始されている。これは現在の模型に新しい概念を付け加えることになるため、現在遂行されている程度の規模の計算であってもさらに計算資源を必要とすることになるであろう。

一方、実験データの統計 ($\sim O(10^7)$) 比べると理論計算の統計 ($O(10^3) \sim O(10^4)$) は圧倒的に小さいという現状がある。最近の実験のデータ解析の様々な発展に伴い、理論についても実験とほぼ同等の解析手法を行うことが求められるようになってきている。その際、理論計算の統計が小さいために実験結果との直接比較が難しいということも起こるようになってきた。さらに、ベイズ統計の手法を取り入れた解析も行われている。ただし、計算資源、計算時間の制約からある特定の物理量やパラメータに制限した解析であることが多い。ベイズ統計の手法こそ、実験結果を包括的に解析し QGP 物性を定量的に明らかにできる強力な方法であるがその利点が存分に生かされていない。

解決に必要な計算手法とアプリケーション

相対論的流体模型の枠組みの発展については、現在使用されている相対論的粘性流体方程式の数値解法を基本とし、注目する物理描像に応じてその拡張を行う。例えば、QGP の粘性を明らかにする場合には、小さな人工粘性でも安定で衝撃波を扱えられる、あるいは人工粘性の振る舞いを把握できているリーマンソルバーなどに基づくコードを使用する。流体描像と粒子描像の双方を取り入れた模型については粒子描像については多くの公開コードがあるため、それらを使用することが可能である。最近になって提案され

たベイズ統計の手法についても公開コードが存在し、それらを部分的、あるいは全面的に使用することが可能となっている。

計算量・計算資源

計算量は実行する計算の次元、ターゲットとする物理量の計算に求める精度、エネルギー領域などに依存する。最近では 3+1 次元の計算が標準になりつつあり、理論性能 64 GFLOPS 程度のシングルコアで 1800 – 18000 秒程度、必要メモリは 2 – 10GB と見積もられる。計算時間のほとんどは流体部分に費やされている。また、現在のベイズ統計の計算では模型のパラメーター数 (~ 10) \times 観測量 (~ 10) \times 必要統計量 (~ 1000 – 10000) $\sim 10^6$ 程度の計算量が必要とされている。計算量は統計量に比例することから、今後、ベイズ統計を採用し、実験データなみの統計量へ近づけることを目標にすると、 $\sim 10^7$ 程度の計算量を要し、現状と同じ時間で計算を遂行するには $\sim 10^7$ コア程度 (640 PFLOPS 程度) の計算資源が必要であると見積もられる。同じ時間の 1800 – 18000 秒程度で計算するので、実効性能が理論性能の 5% と仮定した場合、総演算量は 60 – 600 EFLOP となる。ディスク容量については、特に大きなサイズを必要としない。

4.1.8 少数多体系計算 [長期目標 (a-3) の課題]

課題概要と長期目標

物理学における最近の重要な課題の一つは、量子力学の極小レベルの様々な 3 体以上の少数多体系のシュレディンガー方程式を「厳密」に解くことである。このことにより、物理的な知見を深め、新しい発見や予言に至ることがしばしばある。そのためには、シュレディンガー方程式を少数多体系問題に基づいて精密に解く、適用の広い計算法の開発が不可欠である。これまで世界的に見て、ファデーフ法、no-core シェルモデル、確率的変分法、グリーン関数モンテカルロ法、ガウス展開法など様々な少数多体系計算法が考案され、成功をおさめてきた。これらの計算法を使用した目標は、(1) 精密に何体計算を行うことが可能であるのか？この際、現実的バリオン間相互作用を使用し、実験値と比較・検討し、2 体間バリオン間相互作用、および、3 体力の情報を暴き出すこと、(2) これらの計算手法の応用分野は広いので、陽子・中性子で構成される原子核分野のみならず、ハイパー核（核子とハイペロンから成る新しい原子核）、不安定核、ハドロン多体系、宇宙核物理、エキゾチック原子分子、冷却原子のような分野に適用し、それぞれの分野の固有の重要課題を克服することである。

解決に必要な計算手法とアプリケーション

ここでは、主に変分法を使用した少数多体系計算法について述べる。これにはガウス関数展開法を用いる。少数多体系問題を精密に解くために、ヤコビ座標を用意する。例えば、3 体問題の場合は、2 つのヤコビ座標、4 体問題の場合は、3 つのヤコビ座標を導入する。このように 1 つずつ粒子が増えるたびに、座標が一つずつ増える。この座標ごとに基底関数を張る。変分法であるので、「良い」基底関数を選ぶ必要がある。経験的に、動径部分は、ガウス関数、角度部分は球面調和関数で表す基底関数を使用し、特に、

ガウス関数のレンジは、等比級数に取ることがよいとされている。これらの基底関数を使用してハミルトニアンの行列要素を計算した後、一般化固有値問題を解くことによって、エネルギー固有値と波動関数を求める。

計算量・計算資源

現在のところ、ガウス関数展開法という変分法では、5体問題まで開発が完了している。5体問題の典型的なケースでは、約40,000次元の密行列の対角化問題を解くことに帰着され、理論演算性能350 GFLOPS程度のノード1つを用いて約2時間で計算が完了する。密行列の対角化は一般に高効率なので、簡単のために5体問題を解くのに必要な演算量を理論演算性能で近似する。今後、6体、7体、…と粒子を1個増やすごとに必要な行列の次元数は6倍ほど増加すると見込まれる。1つの行列要素の計算時間は5体の場合からほとんど増加しないとした場合、全行列要素を得るのに必要な計算時間は次元数の2乗に比例し、一般化固有値問題を完全に解くために必要な計算時間は次元数の3乗に比例する。今後は全ての固有値を計算しない対角化アルゴリズム（Krylov部分空間法を利用したものなど）を採用することによって対角化計算時間が相対的に小さくなるとすると、現在と同様の時間で計算を遂行するためには、6体問題、7体問題、8体問題、9体問題、10体問題ではそれぞれ10 TFLOPS程度、450 TFLOPS程度、16 PFLOPS程度、600 PFLOPS程度、21 EFLOPS程度の計算能力が必要となる。総演算量は**72 – 151,200 EFLOP**の範囲である。ハミルトニアン行列を保存するのに必要なメモリは、5体問題では1.3 GBほどであるが、6体問題、7体問題、8体問題、9体問題、10体問題ではそれぞれ50 GB、1.7 TB、60TB、2 PB、80 PB程度まで増加する。ディスク容量については、特に大きなサイズを必要としない。密行列の対角化については、GPU等アクセラレータの利用が可能であると考えられる。

4.1.9 QCDに立脚した原子核反応計算 [長期目標(a-3)の課題]

課題概要と長期目標

原子核は、強く相互作用する核子で形成された有限量子多体系であり、その構造の記述は一般に困難であるが、数値計算技術の向上と計算資源の拡大により、殻模型をはじめとする様々な手法を用いた定量的な核構造計算が可能になりつつある。他方、原子核反応は、そのような原子核同士の反応・衝突によって起きる動的なプロセスを研究対象とする分野である。その目的は、反応現象によってはじめて浮かび上がる、複数の原子核からなる開放系の性質を定量的に解明することである。今世紀以降、現象論から脱却し、核子間相互作用と原子核構造計算で得られた波動関数に基づいて原子核反応を記述するアプローチが目覚ましい進展を遂げている。特に近年注目されているのが、QCDの低エネルギー有効理論として知られるカイラル有効場理論(χ EFT)で得られた2体力・3体力に立脚した反応計算である。現在までに、核子と原子核の反応は、微視的な記述(予言)が確立していると言ってよい。しかし原子核-原子核反応に関しては、未だ道半ばである。その主な理由は、原子核-原子核反応では原子核の励起状態との複雑なチャネル結合が活性化するためである。この問題を解決するには、まず、様々な原子核の基底状態、励起状態(連続状態を含む)の波動関数のセットが必須である。これは、他の課題の成果として得られるものを活用

する。その上で、これらの波動関数のセットを取り入れた大規模チャネル結合反応計算を遂行し、原子核-原子核反応に対する微視的理解を確立することが、本研究課題の目的である。

解決に必要な計算手法とアプリケーション

離散化連続チャネル結合法の計算を行う。反応計算は大きく2つの部分に分かれる。1つはチャネル結合ポテンシャルの計算である。例えば反応に関与する2つの原子核それぞれについて 10^2 個の状態を用意する場合、多重極度も考慮すると、概算で 10^8 個のポテンシャルが必要である。これを、原子核間距離の分点($\sim 10^2$)ごとに用意する。計算に必要な5重積分はモンテカルロ法によって遂行する。もう1つがチャネル結合反応計算である。これには予測子-修正子型メロフ法を使用する。一般に閉じたチャネルが結合する場合には数値解が不安定化するので、独立解を適宜再直文化して解の安定化を図る。反応を解く際に生じる軌道角運動量の分岐を考慮すると、結合するチャネル数は 10^5 程度になると考えられる(現在の計算規模の 10^2 倍程度)。これを部分波($10^2 - 10^3$ 個程度)ごとに解く必要がある。

計算量・計算資源

大阪大学核物理研究センター(RCNP)の高性能演算システム(Xeon 1コア)上で上記の計算を実行する場合、1つの部分波あたり 8×10^7 秒程度の計算時間が必要となる。1つの部分波あたりの演算量は 1.7×10^9 GFLOPである。**必要なメモリは 100 GB 程度**である。計算は部分波ごとに独立しているため並列計算が可能である(並列化効率100%)。各部分波の計算については 10^5 個の独立解を並列的に処理することが可能であるが、再直文化が必要な場合にはその都度同期を取る必要があるため効率はある程度下がると考えられる。必要な部分波の数は $10^2 - 10^3$ 個であるので、**総演算量は 170 – 1,700 EFLOP**である。ディスク容量については、特に大きなサイズを必要としない。

4.1.10 核力に基づく核物質の状態方程式 [長期目標(a-4)の課題]

課題概要と長期目標

無数の核子で構成された無限に大きい核物質の状態方程式は、重力崩壊型超新星爆発・中性子星・ブラックホール形成等の様々な高密度天体现象と密接に関連している。これらの高密度天体の中心部に現れる核物質は、通常の原子核密度の数倍以上にも及ぶ極限状態であるため、その状態方程式を実験的に決定することは困難であり、理論計算が不可欠となる。特に、高密度では核力の近距離斥力による効果が支配的となるため、通常の原子核構造研究に用いられる現象論的な有効相互作用ではなく、2核子散乱実験を良く再現する現実的核力から出発した微視的多体計算によって、信頼性の高い状態方程式を導出することが望ましい。

このような生の核力から出発する多体計算手法として、摂動論に基づく Brueckner-Hartree-Fock 計算、変分法の一種である Fermi Hypernetted Chain (FHNC) 計算やクラスター変分法、さらに量子モンテカルロ計算などがこれまでに提案され、核物質状態方程式の計算に適用されてきた。しかしながら、これらの多体計算結果を比較するとその違いは未だに大きく、核物質状態方程式には不定性が残されてい

る。この不定性の要因は、主に「核力の不定性」と「量子多体計算手法の不定性」とに大別でき、前者の例としては、原子核の地上実験では決定が難しい3体核力の斥力成分が挙げられる。一方で、実験的に良くわかっている2体核力であっても、その複雑な状態依存性のために、多体計算手法の違いによる不定性は収束していない。これらの寄与を精度良く取り込み、信頼性の高い状態方程式を構築するためには、高精度な数値計算と膨大な計算時間が必要となる。

近年、中性子星を始めとする高密度天体の観測研究が飛躍的な発展を遂げており、今後も期待される同様の観測から原子核物理の情報を精度良く引き出し、核力の不定性を絞り込むためには、量子多体計算精度の高い手法による核物質状態方程式が待ち望まれる。このような現状を踏まえ、本研究課題では、核力の複雑な状態依存性を精度良く取り扱うことができる量子多体計算手法を確立し、高密度天体现象に適用可能な核物質状態方程式を構築することを目的とする。

解決に必要な計算手法とアプリケーション

Jastrow型波動関数を用いたクラスター変分法によって計算を行う。核物質に対する多体変分法として代表的なクラスター変分法では、Jastrow型波動関数によるハミルトニアン期待値を粒子間相関で級数展開（クラスター展開）して、核物質の1核子あたりのエネルギーを計算する。現在までの研究では、高次クラスター項を無視した最低次の近似である2体クラスター近似を採用し、適当な拘束条件を課すことでの、FHNC計算結果と定量的な一致を示す状態方程式を構築してきた。この研究を発展させ、高次のクラスター項まで系統的に考慮した多体変分計算を実行するためには、状態依存性を持つ複数の相関関数に対して、連立積分微分方程式を数値的に解き、自己無撞着な解を得るまで反復する必要がある。この数値計算によって、1つの熱力学状態（密度・粒子組成・温度を固定した1状態）における核物質の自由エネルギーが求まり、同様の計算を様々な密度・粒子組成・温度の組に対して行うことにより、膨大な核物質の熱力学量を完備した高密度天体研究用の状態方程式が完成する。

計算量・計算資源

従来の2体クラスター近似を用いた変分法の場合、1つの熱力学状態における核物質の自由エネルギーを得るためにには、約80 GFLOPSの性能のコアで10分程度の時間を要した。高密度天体への応用のためには、この演算を様々な密度・温度・粒子組成の組に対して行う必要がある。従来の高密度天体研究用の核物質状態方程式では、原子核の現れる非一様相まで自己無頓着に扱うため、およそ 3×10^7 組の熱力学状態について同様の計算を行い、その総演算量は約1,500 EFLOPであった。今後の研究で行う高次クラスター項まで考慮した多体変分計算の場合、連立積分微分方程式の自己無撞着な解を反復法で探す必要がある。3体クラスター項の主要部分までを取り入れた先行研究においては、十分安定な解を得るまでに $10^2 - 10^4$ 回ほどの反復を要しており、今後の研究では、1つの熱力学状態における自由エネルギーの計算において、5–500 PFLOPの演算量が必要になる。そのため、最終的な高密度天体研究用核物質状態方程式の完成までに、総演算量は**150,000 – 15,000,000 EFLOP**になることが予想される。メモリ容量、ディスク容量については、特に大きなサイズを必要としない。

4.1.11 配位間相互作用計算（殻模型計算）[長期目標(a-3)の課題][長期目標(a-4)の課題]

課題概要と長期目標

原子核では、原子と同様に、個々の核子はそれぞれの軌道を占有するという独立粒子描像がよく成り立つことが知られている。しかし、残留相互作用（有効相互作用）の強さが隣り合う軌道間のエネルギー差と近いことから、現実の多体状態はその影響を取り入れた複雑なものとなる。配位間相互作用計算とは、与えられた少数個の軌道からなる空間（模型空間あるいはバレンス殻）の多体問題を厳密に解く手法であり、原子核物理では殻模型計算と呼ばれている。これは与えられた模型空間内の多体問題を正確に解くものであるため、十分に多くの軌道を採ることによって第一原理核構造計算にも応用することができる。殻模型計算は与えられたハミルトニアンに非常に忠実な多体状態を与え、特に原子核の分光学的性質の解明に威力を発揮している。

現在のところ、ランチヨス法を用いた数値対角化の手法により、基底状態近傍については 10^{11} 次元程度までのハミルトニアン行列の固有状態を得ることが可能となっている。それによって、 pf 殻（陽子あるいは中性子数が 20 から 40 まで）をバレンス殻と仮定した系のハミルトニアン行列の完全対角化が可能であり、質量数 70 程度までの中性子過剰核を含む多くの原子核の性質が調べられている。より重い原子核、あるいは軽い原子核でも大きな変形を持つ状態については、さらに大きな模型空間が必要であり、厳密対角化の限界を超える重要な課題は多く残されている。その限界を超えるため、少数の精選された多体基底によって原子核の波動関数を構成する、モンテカルロ殻模型およびその後継の準粒子真空殻模型が提唱されている。これらの新しい手法によって、現在では核子数 150 程度までの原子核の構造研究が行われているが、将来的にはウラン領域、さらには超重元素領域までそのフロンティアを拡大することを目指す。また、モンテカルロ殻模型によって、 sd 殻（陽子あるいは中性子数が 20 まで）の原子核構造の第一原理計算を遂行することを今後 10 年ほどの目標としている。

解決に必要な計算手法とアプリケーション

これらの計算は殻模型計算に該当する。大きく分けて二つの計算手法が採られている。一つはランチヨス法による厳密対角化計算であり、模型空間に含まれる全ての多体基底を構成し、それらで張られるハミルトニアン行列を数値的に対角化する手法である。もう一つはモンテカルロ殻模型および準粒子真空殻模型であり、これらは全ての多体基底を構成するのではなく、着目する多体状態に重要な精選された 100 本程度の多体基底（量子数射影を施した相関基底）を採用することによって良い精度で核構造を記述しようというものである。後者の手法では基底の精選において確率的過程および変分的過程が取り入れられ、模型空間の次元数に比べてはるかに少数の基底によって厳密解に近い波動関数を表現することができる。厳密対角化計算とモンテカルロ殻模型計算は相補的な役割を担っているため、当面は両者が共存するとみられる。

計算量・計算資源

厳密対角化計算については、相互作用が2体力の範囲内の場合、ハミルトニアン行列の非対角要素の数が 10^3 オーダーであるケースが多い（ハミルトニアン行列の次元数に対する依存性は小さい）。中重核の興味ある問題をある程度カバーできる、ハミルトニアン行列が 10^{16} 次元程度の基底状態を得るために、ランチヨス法の反復回数を50として、 $10^{16} \times 10^3 \times 50 = 5 \times 10^{20}$ 回の浮動小数点演算が必要となる。100ケースの計算に必要な総演算量は**50,000 EFLOP**である。必要なメモリ容量、ディスク容量は波動関数のサイズで決まる。ハミルトニアン行列が 10^{16} 次元程度の問題の場合、波動関数のサイズは 8×10^{16} B= 8×10 PBである。メモリにはランチヨス反復過程で生成されるベクトルを保持する必要があり、ディスクには遷移行列要素を計算するなどの解析のため、得られた波動関数を保管するスペースが必要である。なお、ハミルトニアン行列要素はオンザフライ生成法により必要に応じて生成するため、そのためのメモリ容量は不要である。ランチヨスベクトル、保管する波動関数とともに典型的に100本のベクトルであるとすると、必要な総メモリ容量、ディスク容量はともに 8×10 PB $\times 100 = 8$ EBとなる。現在のところ、この計算でGPUなどのアクセラレータを使うことは難しい。

モンテカルロ殻模型計算については、第2.1.5節に計算資源の見積もり量を記述している。演算量は、[(a) 1つのハミルトニアンカーネルを得るのに必要な演算回数]×[(b) 量子数射影に必要な数値積分の分点数]×[(c) 収束した波動関数が得られるまで必要な試行回数]によって決まる。*sd*殻核の第一原理計算を行うために最低限必要なバレンス殻として、調和振動子基底8主殻を採用した場合、(a)はおよそ 10^9 FLOP=1GFLOPであることがわかっている。(b)として、オイラー角、パリティ反転の分点数 $2 \times 30^2 \times 20 = 3.6 \times 10^4$ 、(c)として厳選された400基底を得るのに必要な経験的回数である 8×10^7 を採ると、**1000ケースに必要な総演算量は $1\text{ GFLOP} \times 3.6 \times 10^4 \times 8 \times 10^7 \times 1000 \approx 3,000,000\text{ EFLOP}$** となる。メモリについては、2体相互作用行列要素など、全プロセスで同じ内容を保持しているものが大半を占める。したがって必要な総メモリ容量は、プロセスあたりのメモリ容量（これはノード数に依存しない）とノード数の積となる。8主殻を採用した計算を行う場合、プロセスあたりのメモリ使用量の実測値34GBに典型的なノード数5,000をかけ、**必要な総メモリ容量は0.2PBと見積もられる**。モンテカルロ殻模型計算はLevel3BLASとなる密行列積が主であり、これはByte/FLOP値が0.1でも性能が出るすれば、ケースあたりのメモリ転送量は、 $3,000\text{ EFLOP} \times 0.1\text{ Byte/FLOP} = 300\text{ EB}$ となる。ディスク容量については、特に大きなサイズを必要としない。通信は波動関数を最適化する際に必要なグラジエントベクトルが主となる。これは波動関数（スレーター行列式）のサイズに $(2I+1)^2$ （ I は求める状態の角運動量量子数）をかけたデータ量となる。8主殻に陽子、中性子とともに10個配置したスレーター行列式を表現するには、倍精度複素数を用いた場合、 $16 \times (240 \times 10 + 240 \times 10) = 76800$ Bのデータサイズとなることから、 $I=8$ の状態を計算するためのグラジエントベクトルのサイズは $76800\text{ B} \times (2 \times 8 + 1)^2 \approx 22\text{ MB}$ となる。通信回数は1基底を最適化するのに50回必要であると仮定すると、400基底の最適化を5,000ノードで行った場合、グラジエントベクトルの通信サイズの合計は、 $22\text{ MB} \times 50 \times 400 \times 5000 = 2200\text{ TB}$ となる。ノードあたりの要求バンド幅は、通信にかけてもよい時間を全体の10%、ケースあたりの計算時間を12時間とした場合、 $22\text{ MB} \times 50 \times 400 / (12 \times 3600\text{ s} \times 0.1) \approx 100\text{ MB/s}$ となる。現状ではアクセラレータを積極的に使ってはいないが、過去にGPUが利用可能であることを確かめたことがある[1]。

4.1.12 原子核密度汎関数計算 [長期目標 (a-4) の課題]

課題概要と長期目標

原子核の励起状態には、核子の独立粒子運動を反映した励起モードと並び、振動・回転といった多数の核子が運動に関与する集団励起モードが現れる。一核子状態（軌道）は、相互作用によって核子が自己無撞着に作り出す平衡平均場によって与えられる。また、多数の核子から成る集団運動は、平均場そのものが時間的に変化することによって与えられる。有効相互作用は短距離的であるので、平均場と核子密度は似たような空間分布を持つ。そこで、時間依存密度汎関数理論の考えに則って、核子系のエネルギー密度汎関数を出発点とし、一核子状態および集団運動を統一的に調べる微視的アプローチが、原子核密度汎関数法と呼ばれるものである。

エキゾチックな変形状態や、不安定核で見られる空間的に異常な核子密度分布をも表現できるよう、基本方程式 (Hartree–Fock–Bogoliubov (HFB) 方程式と呼ばれる) は実空間座標で解かれる。対角化すべきハミルトニアンは、3次元座標を用いる場合、各方向のメッシュ点の数を N とすると、 $4N^3$ 次元のエルミート行列である。原子核の質量数を A とすると、核半径は $A^{1/3}$ でしか大きくならないため、軽い原子核であっても重い原子核であっても計算量はほとんど変わらない。従って、原子核密度汎関数法は現在のところ、核図表上のすべての原子核を一つの枠組みで記述可能な唯一の方法であると期待されている。それでも多数の核種を対象とした系統的計算には、計算機資源が限られているため、軸対称を課した2次元計算が主流である。近年になり、軸対称性を破る変形状態やクランキング法による不安定核の集団的回転状態も記述できるよう、3次元コードの開発が進んでいる。直接対角化の方法のみならず、グリーン関数の方法を用いて直接対角化を回避する試みもなされている。いずれも大規模並列化によって計算が可能になってきた。また、集団的振動励起状態に関しては、非軸対称変形核に対する線型応答計算が可能になってきたところである。系統的計算にはまだ計算機資源が足りないが、低励起領域に現れる非線型振動や、複数の平衡平均場を跨ぐような大振幅集団運動を記述できる理論・計算法も提案されている。長期的目標は、核図表上の広い領域の原子核に対して非イラスト状態を含む低励起状態の系統的計算を行い、まだ見ぬエキゾチックな構造を予言し、その発現機構を理解することである。

解決に必要な計算手法とアプリケーション

密度汎関数法の計算である。3次元 HFB 計算に関しては、直接対角化とグリーン関数の方法が提案されている。密度分布（平均場）の自己無撞着な解を得るのに 100 回程度の反復が必要である。対角化には ScaLAPACK の汎用的なサブルーチンが用いられている。グリーン関数の方法では Krylov 部分空間法が用いられている。線型応答計算では、平均場に取り込めなかった残留相互作用の行列要素計算と、その行列の対角化が必要になる。更に大振幅集団運動の記述のためには、平衡平均場から離れた様々な配位での残留相互作用の行列要素計算並びに対角化が必要である。直接対角化を回避した有限振幅法が近年提案されており成功を収めている。

計算量・計算資源

現在の系統的な線型応答計算では、軸対称性及び反転対称性を課した計算が主流である。ある量子数に対する遷移強度分布を得るのに、二準粒子励起の組み合わせ数 (M) が 10^4 程度のとき、256 コアを用いて 1,000 秒程度の時間が必要である。このときの演算量は 0.45 PFLOP である。計算、格納すべき行列要素の数は M^2 であるが、それらは各 CPU に分散しているため 1 コア当たりのメモリは数 MB 程度である。

核分裂現象は大振幅集団運動の典型例であり、その微視的記述は最もチャレンジングな課題の一つである。核変形 5 次元空間上の 10^5 個の核変形に対して拘束付き HFB(CHFB) 計算にてポテンシャルエネルギーを求め、その配位での集団運動の質量パラメータを局所準粒子乱雑位相近似 (LQRPA) にて計算する。行列要素の演算量は M^2 に比例する。対角化については、全固有値・固有ベクトルを必要とするならば M^3 に比例するが、今の場合エネルギーの低い数個の固有状態が分かれば良い。さらに、実空間表示ではなく調和振動子基底で CHFB 方程式を解くことで $M = 10^6$ 程度の問題とする場合、演算量が M^2 に比例するベストケース (有限振幅法の理想値) で総演算量は **780,000 EFLOP** となる。総メモリ容量は **8 TB** 程度、ディスク容量については、特に大きなサイズを必要としない。

4.1.13 重イオン反応・核分裂・新元素合成 [長期目標 (a-4) の課題]

課題概要と長期目標

近年、重イオン衝突による多核子移行 (Multi-nucleon transfer: MNT) 反応が注目を集めている。核子移行反応によって得られる中性子過剰な高励起複合核の実験データは、原子炉において発生する放射性廃棄物の処理や、高速炉の設計などで利用されるためである。また、新元素合成の研究領域においても、中性子過剰核領域に存在すると予測されている安定な島（陽子数 114 個、中性子数 184 個）の中心に到達するには、多核子移行反応による中性子過剰核の生成が必要である。さらに 2017 年に中性子星合体の重力波の初観測によって話題になっている、宇宙元素合成の起源を探るための r プロセスの研究にも、中性子過剰核の研究が必要であり、多核子移行反応の利用が望まれている。このような背景のもと、日本原子力研究開発機構ではタンデム加速器を利用して、多核子移行反応による中性子過剰核の研究が活発に行われ、非常に多くの新しい実験データが取得されている。

超重元素の合成を成功に導くためには、入射核・標的核同士の非常に強いクーロン反発力を克服し、さらに融合に向けた運動エネルギーの強い散逸も考慮しなければいけない。このような状況下では、ほぼすべての事象は複合核形成には至らず、弾性散乱、非弾性散乱、または準核分裂を起こす。超重元素合成の可能性を議論するためには、すべての反応の中で融合に至る「非常に稀な事象」を扱う必要がある。すなわち平均値から遙か離れた稀有な事象を扱う模型が必要となる。

従来、重イオン融合反応の理論的記述は、原子核の形状を表すパラメータによって形成されるポテンシャルを計算し、その空間内を移動する軌道追跡によって、原子核の形状の時間発展を記述する方法が取られてきた。この手法は摩擦を考慮したポテンシャル上を移動する軌道をニュートン方程式によって記述することであり、言わば平均軌道の追跡にあたる。ところが、超重元素領域では、上に述べた理由によ

り、いくら高い入射エネルギーを投入しても、平均軌道は融合領域には到達することが出来ない。そこで、平均軌道からの揺らぎを考慮し、平均値から非常に離れた事象として、融合領域に到達できる事象を評価することが出来ないか検討されてきた。原子核の運動の中で、速度の遅い「形状を表す集団運動」と、それ以外の速度の速い運動（例えば核子の運動）の二つの運動に分類することで、前者をブラウン粒子、後者を熱浴とみなすことが出来る。すなわち、平均値からの揺らぎを原子核の形状の熱的揺らぎに対応させることで、散逸揺動理論に基づくブラウン運動の理論を、原子核の反応理論に適用しようとする試みである。具体的には、多次元ランジュバン方程式を採用し、揺動を入れた軌道を追跡し、融合領域に到達する確率を評価する手法を採用している。このような手法は統計模型に代表される静的模型に対し、「動力学模型」と呼ばれている。超重元素領域の融合確率を計算する際、有為な結果を得るためにには、用意するアンサンブル数として 10^4 以上は少なくとも必要であると考えられる。さらにこのような軌道計算で蒸発残留核確率まで求めようとすると、 $10^{10} - 10^{15}$ 程度のアンサンブルは必要である（必要なアンサンブル数はエネルギーに依存する）。また動力学模型では、原子核のポテンシャル（液滴模型と殻補正エネルギー）、摩擦テンソルと慣性質量テンソル（これらを輸送係数と呼ぶ）、ランダム力の強さが計算のインプットデータとして必要である。

現在は、原子核の形状を表すパラメータとして 5 変数（重心間距離、左右のフラグメントの変形度、質量対称度、ネックサイズ）を採用した計算が実現しており、研究が進められている。輸送係数は巨視的模型を用いて計算している。トリウムのアイソトープで見られるよう核分裂質量分布の三山構造を再現するような高精度のシミュレーションを行うためには、さらに原子核の形状を表す次元数を増やし、融合分裂過程における現実的な形状変化を記述可能にする必要がある。さらに核データで必要な独立収率を得るには、質量非対称度だけではなく、荷電非対称度も考慮する必要がある。融合過程において入射核と標的核が接触するまでの段階では、各々の原子核の回転効果も考慮に入れ（各々の原子核の回転核の 2 変数と相対角度の 1 変数の計 3 変数）、すなわち 9 次元以上のランジュバン方程式を解く必要があり、それに伴うポテンシャルや輸送係数テンソルが必要なため、非常に膨大な計算量が必要となる。現在は融合過程に動力学計算を行い、形成された原子核（すなわち複合核）の崩壊過程では統計模型を用いて計算を行っている。

上記は、前者の計算を 10^{-18} 秒まで行い、その後、統計模型にスイッチする方法である。しかし、統計模型には不定なパラメータも多いため、理想的には蒸発残留核形成まで含めた全融合分裂過程を動力学模型で解くことが望まれる。計算時間の大幅な増加、さらに上に述べたようにアンサンブルも 10^{15} 個程度必要であると考えられる。また超重元素合成では、融合過程においても核構造の影響が大きいと考えられ、そのような効果を取り入れた計算を行うことで、実験計画を提案する必要がある。したがって、微視的輸送係数の導入が不可欠であるが、それには準位密度計算が必要であり、動力学模型の計算を実行するために必要な空間でこれらのデータを準備するには、多大の計算時間を必要とする。また微視的輸送係数は原子核の温度に依存するために、それぞれの温度でのデータが必要となる。しかし、動力学計算において軌道の温度は時々刻々と変わるために、まともに解けばさらに膨大な計算時間がかかることが予測される。しかしながら、超重元素の研究では、生成確率という非常に微小な現象を扱うため、上記に述べたような効果が、ほんのわずかでもダイナミクスに影響を与えた場合、それに伴って微小な確率も影響を受け、絶対値自体が大きく変動する可能性がある。そのため、計算時間にかかる困難を乗り越え、超重元素合成の微小な生成確率を、高精度で算出することが必要である。

また日本原子力研究開発機構で得られた多核子移行反応による核分裂片質量分布の一連の実験データの解析や、 r プロセスのネットワーク計算に関する中性子過剰核の分裂片質量分布等、非常に多くの核種を系統的に取り扱う必要があり、大規模計算および計算の自動化が不可欠である。

解決に必要な計算手法とアプリケーション

超重元素領域における全融合分裂過程を、動力学模型を用いて計算するにはアンサンブルの増加、および計算時間の増大が回避できないため、計算手法を根本的に見直す必要がある。動力学模型の計算にはランジュバン方程式を用いている。ランジュバン方程式の数値計算は、並列化が非常に有効であり、使用する並列計算機に適したアルゴリズムを構築する必要がある。

また核子移行反応への適用やさまざまな実験データとの比較、さらに r プロセスネットワーク計算に用いる核分裂データベースの構築等、膨大な核種の計算が必要であるため、計算の自動化やデータ処理の最適化に向けた改良が必要である。

計算量・計算資源

新元素合成における蒸発残留核断面積の計算において、現在、数 pb(ピコバーン: $10^{-12} b = 10^{-36} \text{cm}^2$) の計算が活発に行われている。将来は、合成される元素の原子番号の増加に伴い、さらに 3 柄小さい数 fb(フェムトバーン: $10^{-15} b$) 程度のオーダーの断面積の計算が必要となってくる。このためには現在の計算機 (1 CPU(12 コア)、2.6 GHz) で 100 万秒程度 (11 日程度) の計算が必要である。これはひとつの入射エネルギーに対する計算であるが、励起関数を得るためにには、数ヶ月の計算が必要である。10 年後に挑むべき課題では、fb から ab(アトバーン: $10^{-18} b$) の断面積の評価が必要になる可能性がある。すなわち原子番号 124 番以上の元素合成が想定される。この場合、現在より $10^3 - 10^6$ 倍程度の計算量が必要となる。

さらに核子移行反応の研究に関連して核分裂データベースの構築のためには膨大な量の計算が必要となり、中性子過剰核領域を中心に 200 核子以上の計算が望まれる。この場合、現在より $10^6 - 10^9$ 倍程度の計算量を要し、総演算量は **50,000 – 50,000,000 EFLOP** となる。メモリ容量、ディスク容量については、特に大きなサイズを必要としない。

4.1.14 微視的クラスター模型 [長期目標 (a-4) の課題]

課題概要と長期目標

原子核の励起状態には、1 核子励起状態や集団運動状態に加えて、クラスター状態が存在する。クラスター状態にある原子核は、いくつかの部分系 (クラスター) に分かれた構造を持っている。例えば 3 つのヘリウム 4 原子核が弱く結合してつくられる「ホイル状態 (炭素 12 の第 2 励起状態)」はその代表例である。こうしたクラスター状態では、核子自由度はクラスター内部に閉じ込められており、核内でのクラスターの運動によって原子核の性質が説明される。つまりクラスター状態は、原子核の基本的自由度が核子から「クラスター」へと変化した状態と見ることができる。

また、クラスター状態は核反応チャンネルと連続的に繋がっているため、核反応と密接に関係する。特に、天体で起こる低エネルギー核融合では、クラスター状態が反応プロセスの中間状態として重要な役割を果たす。先ほどのホイル状態は、まさしくその代表例でもある。このようにクラスター状態は、核子多体系の動力学を考えるうえでも、天体现象を理解するうえでも非常に重要な研究対象である。

近年、計算機能力と微視的核模型の進歩によって、クラスター状態を正確に記述し、さらに散乱の境界条件に接続することが可能になった。その結果、重陽子の放射捕獲反応や炭素の核融合などの天体核反応を、微視的核模型に基づいて定量的に議論することが可能になりつつある。天体核反応の多くは、実験による測定が難しいため、こうした理論による反応率の精密計算は非常にインパクトの大きな課題になると期待される。そこで、様々な核反応チャンネルを統一的に扱える汎用的理論手法と計算プログラムを開発し、天体での核反応率を系統的に求める事が重要な長期的目標として挙げられる。

解決に必要な計算手法とアプリケーション

反応の中間状態として現れるクラスター状態を記述するプログラムと、それを散乱の境界条件に接続し、断面積を求めるプログラムが必要になる。いずれの場合も密行列の一般化固有値問題を解く必要があり、行列要素の計算と対角化が律速過程となる。

特に、共鳴状態を扱うためには計算手法とアプリケーション開発が必要となる。複素座標回転法では、非エルミート系を扱うため、複素対称行列の一般化固有値問題を高速に解くルーチンが必要となる。また、結合定数解析接続法 (analytic continuation of coupling constant) では、ハミルトニアンのパラメータをわずかに変えながらエルミート行列の一般化固有値問題を繰り返し解く必要があるため、固有値連続化 (eigenvalue continuation) の方法を適用することも考えられる。

計算量・計算資源

ひとつの行列要素を求めるのに必要な計算量は、ハミルトニアンに局所密度近似をした密度汎関数を用いる場合には質量数の 3 乗に、有限レンジの密度汎関数では質量数の 4 乗に比例する。また、回転角に関する数値積分が必要であり、計算量は扱う実空間のサイズに比例する。

求めるべき行列の次元は、1 チャンネル当たりおおよそ 1,000~10,000 程度であり、考慮するチャンネル数に比例する。ただし、各々の行列要素は完全に独立に計算できるため、ノード間通信なしで並列化できる。また、スレッドあたりのメモリ消費量は数 GB 程度である。

比較的コストの小さい、ヘリウムと酸素の放射捕獲反応をシングルチャンネルで計算する場合には、 $20 \times 20 \times 40 \text{ fm}^3$ (grid 数は約 20,000) 程度の空間が必要であり、行列の次元は 1,000 程度である。この場合の総演算量は 360PFLOP であり、そのほとんどを行列要素の計算が占める。より挑戦的な課題、例えば炭素原子核の核融合反応を full チャンネルで記述する場合を考えると、考慮すべきチャンネル数は 10 チャンネル程度、さらに 2 重角運動量射影が必要となるため計算量が $10^2 \times 10^4$ 倍程度となる。その結果、総演算量は **360,000 EFLOP** となる。メモリ容量、ディスク容量については、特に大きなサイズを必要としない。

4.1.15 新物理探索・解明のための精密格子 QCD シミュレーション [長期目標 (b-1) の課題]

課題概要と長期目標

標準理論が記述する階層よりも更に一段深い階層の新しい物理を探るためには、未知の高エネルギー領域を探索する実験と、既知の物理現象の精密検証実験の2種類のアプローチがある。KEKのSuperKEKB/Belle II実験とCERNのLHCb実験の目的は後者であり、既知の物理現象について高精度の実験結果と標準理論による精密な理論的予言（理論値）を比較することで、それらのズレを見出すこと、つまり標準理論では説明できない標準理論を超えた物理現象を捕らえることである。このズレを明確に捕らえるためには、高精度の実験結果と共に正確な理論的予言が要求されるが、どちらか一方が欠けても高精度検証は実行不可能になってしまうため、理論計算と実験結果の双方が要求精度を達成することが極めて重要である。

Belle II実験とLHCb実験で測定される物理量の理論計算には、ボトムクォークを含むハドロン行列要素が必要である。この量を最も精確に計算できるのは、ハドロン行列要素の第一原理計算が可能な格子QCD計算であると考えられている。しかし、現状では計算機性能の制約から、パイ中間子を構成する軽いクォークを再現しながら、1 fm程度のQCDの特徴的スケールと、その1/10程度のボトムクォークのスケールを同時に取り扱うことが困難な状況にある。このため、これまでには、軽いクォークの質量を現実世界よりも大きくするか、ボトムクォークを非物理的に軽くすることを余儀なくされてきた。近年の急速な計算機性能の向上により、この困難が解決される兆しが見え始めた。そこで今後の長期的な目標は、軽いクォークとボトムクォークを含む全てのクォークの質量を現実世界の値に固定し、クォーク質量についての近似を排してQCDの直接計算を行い、将来の実験プロジェクトで得られる精度に見合う理論的予言を与えることである。具体的には、 B 中間子レプトニック崩壊定数(f_B, f_{B_s})、セミレプトニック崩壊($B \rightarrow \pi l\nu, B \rightarrow D^* l\nu, \Lambda_b \rightarrow p(\Lambda_c) l\nu$)を記述する形状因子、 B 中間子混合行列要素(B_B)を1%精度で予言する。また、ボトムクォーク稀崩壊過程($B \rightarrow K^* ll, \Lambda_b \rightarrow \Lambda ll$)の形状因子については5%以下を目標にする。これらの目標精度は、2032年まで予定されているSuperKEKB/Belle II実験の測定精度に合わせたものになっている。

この格子計算は、QCDの非摂動論的性質が重要な興味深いテーマに幅広く応用できる。例えば、現実世界のクォーク質量と非常に細かい格子を用いるため、新しい物理階層探索で重要な軽いハドロンの行列要素計算($K \rightarrow \pi\pi$ 崩壊行列要素、中性子電気双極子能率、ミュー粒子異常磁気能率におけるハドロン真空偏極・光子-光子散乱)、及びILC実験計画で重要なQCDの基本的パラメータ(強結合定数やクォーク質量)を高精度で決定できる。また、近年、クォーク模型では記述できないエキゾチックハドロンの発見が報告されているが、このようなハドロンの多様な存在形態について理解を深めることができる。クォーク質量を変えると崩壊の閾値が変化してしまうため、エキゾチックハドロンのような不安定粒子の研究では、現実的なクォーク質量を用いることが重要である。これらの研究は、同じ目的の有効バリオン間相互作用の決定やQED高次補正計算と連携して推進していく。

解決に必要な計算手法とアプリケーション

格子 QCD ハイブリッドモンテカルロ法^(a)、ドメインウォール型フェルミオン作用^(b)、大規模疎行列反復解法^(b)。

上付き (a), (b) については 2.1.3 節を参照。

計算量・計算資源

格子上でカイラル対称性を近似的に保つ 5 次元メビウスドメインウォール型のフェルミオン作用を用い、アップ、ダウン、ストレンジ、チャーム、ボトムクォークに現実的な質量を用いた計算を行う。計算コストを削減するためにアップ、ダウンクォークは縮退していると近似し、その補正は格子 QCD 計算やカイラル摂動論により評価する。離散化誤差を抑制してボトムクォークの直接計算を行うため、現在の典型的な格子間隔 ($a = 0.1$ fm) よりも小さい $0.082 - 0.025$ fm の格子間隔を用いる。これらの計算の有限体積効果を抑えるため、各格子間隔で $M_\pi L \sim 4$ という条件を満たす、約 $(6 \text{ fm})^3$ の空間体積を用いる。本課題で用いる格子間隔 a と 5 次元格子サイズを表 4.2 にまとめた。

格子間隔 a [fm]	a^{-1} [GeV]	格子サイズ	計算量 [EFLOP]
0.082	2.4	$72^3 \times 144 \times 12$	53,000
0.055	3.6	$108^3 \times 216 \times 8$	600,000
0.044	4.5	$136^3 \times 272 \times 8$	2,300,000
0.033	6.0	$180^3 \times 360 \times 8$	12,900,000
0.025	8.0	$240^3 \times 480 \times 8$	72,300,000

表 4.2 新物理探索・解明のための精密格子 QCD シミュレーションに用いるパラメータと計算量

高統計精度を達成するため、格子間隔 0.044 fm、 $M_\pi = 135$ MeV で 11,000 分子動力学時間のゲージ配位生成を行う。その他のパラメータでは、自己相関長 $\propto 1/(aM_\pi^2)$ を仮定して、分子動力学時間をスケールさせる。配位生成に必要な計算量は、格子間隔 0.043 fm、 $M_\pi = 300$ MeV で既に行ったシミュレーションの実測値と自己相関長の式に基づき、次のスケーリング式により評価した。

$$960 \times \left(\frac{a^{-1}}{4.5 \text{ GeV}} \right)^6 \times \left(\frac{300 \text{ MeV}}{M_\pi} \right)^9 [\text{EFLOP}] \quad (4.8)$$

近年のシミュレーションの経験から、物理量計算にはゲージ配位生成の 2.5 倍の計算がかかると想定する。表 4.2 に、それぞれの格子間隔でのゲージ配位生成と物理量計算に必要な計算量をまとめた。よって、この課題全体に必要な総計算量は **88,000,000 EFLOP** である。

この課題の総計算量の約 8 割に相当する、格子間隔 0.025 fm の計算について、要求性能の詳細な内訳とそれらの数値を第 2.1.5 節に挙げている。

表4.2から、このパラメータでの総計算量は **72,300,000 EFLOP** である。ゲージ配位生成と物理量計算のコア部分は、クォーク伝搬関数の計算であり、ドメインウォールフェルミオン作用を用いたクォーク伝搬関数は、Krylov部分空間反復法である共役勾配法(CG)により計算される。

クォーク伝搬関数計算に必要なメモリ量は5次元格子サイズに比例し、一方、物理量測定に必要なメモリ量は4次元格子サイズに比例する。それらを合わせた、このパラメータでの**使用メモリ量は 500 TB**である。

パイ中間子質量135MeVでのクォーク伝搬関数を得る時のCGの反復回数は63,000回程度、ゲージ配位生成中のクォーク伝搬関数の計算は2,300,000回程度となるため、ゲージ配位生成に必要なCGの反復回数は 1.5×10^{11} 回である。物理量測定については、計算量と同じく、この2.5倍の反復回数が必要と見積もった。CGの計算に必要になる行列ベクトル積は、各ノードに分割された行列とベクトルを使って実行するため、隣接ノード間のメモリ転送が発生する。5次元方向は各ノードでメモリを確保していると仮定すると、その転送量はノード内の4次元の各方向の境界の大きさに比例する。ノード数を40,000とし、クォーク伝搬関数の計算に必要なCG反復回数を上の値とすると、ゲージ配位生成と物理量測定に必要な**メモリ転送量は 560,000 EB**になる。

生成されるゲージ配位は、4次元格子サイズの大きさで決まり、1ゲージ配位あたり3.8TBである。40分子動力学時間ごとに保存すると、このパラメータでは2,000ゲージ配位を保存することになり、計算に用いる中間ファイルの容量と合わせると**ストレージ量は 8 PB**が必要である。

4.1.16 量子電磁力学 (QED) の高次補正計算 [長期目標(b-1)の課題]

課題概要と長期目標

量子電磁力学 (Quantum Electrodynamics : QED) は、現在、実験的に最も高い精度で検証されている理論と言える。素粒子物理学的に重要な研究対象としては、特に電子とミュー粒子の異常磁気能率($g-2$)がある。QEDの $g-2$ への寄与を計算するには、電磁気力の強さを表す微細構造定数 α を展開のパラメタとして

$$(g-2)/2 = \alpha A_2 + \alpha^2 A_4 + \alpha^3 A_6 + \alpha^4 A_8 + \alpha^5 A_{10} + \alpha^6 A_{12} + (\alpha \text{の高次項}) \quad (4.9)$$

と摂動計算を行う。この各次の展開係数のうち、 $A_{2n}(n=1, 2, 3, 4)$ は、2017年までに電子及びミュー粒子の双方に対して、複数の計算方法(解析解、準解析計算値、数値計算値)によって値が計算され、完全に確立している。一方、 A_{10} は、ごく一部の寄与について解析解が知られているものの、ほとんどは数値計算によってのみ値が得られている。電子とミュー粒子の重要な違いは、その質量に起因する、 $g-2$ の値に対するQED以外の寄与の大きさである。

電子 $g-2$ の最高精度の実験はハーバード大学グループによる2008年の値で、その相対精度は 2.4×10^{-10} と12桁の数字列として知られている。今後、量子計測手法の開発、磁場の高精度化や安定化など、そして電子の反粒子である陽電子を用いた計測を通じて、さらに1~2桁の精度改良が可能で、新実験は順調に進捗している。

実験値を理論値と比較するためには、観測値としての微細構造定数 α の値が必要となる。2019年以降、相次いで原子波干渉計での実験成果が報告され、これらを用いて、電子 $g-2$ 実験の相対精度を上回る精

度で α の値が供給されることとなった。そこで電子 $g-2$ で理論と実験をほぼ同じ精度で比較することが可能になった。2020 年までの最も精度の良い 2 つの α の値は、 5.4σ の有意な差を示しているが、どちらの値を用いても、電子 $g-2$ の理論値と実験値は 2.4σ の範囲内で一致している。これらの原子干渉計実験もさらに一桁以上の改善を目指して準備が行われている。

電子は最も軽い荷電レプトン粒子であるため、素粒子標準理論を超えた現象への感度が低く、標準理論の破れあるいは新物理の探索には使えないといされていた。しかし、電子 $g-2$ 実験、 α の値、そして QED 計算の全てで、2022 年の時点で知られている値の精度から、さらにもう一桁精度が上がれば、新物理の模型構築の方針そのものや、模型の持つパラメタ領域に非常に強い制限を与えることが可能となる。

ミュー粒子は電子より約 200 倍重いため、質量比の 2 乗、40,000 倍程度、電子より高いエネルギーの現象に敏感である。そのため、ミュー粒子 $g-2$ では 10^{-7} 程度の相対精度を達成すれば、未知の重い粒子の影響による新物理の手がかりが見えるのではないかと期待されている。この発見のためには、実験および理論の双方がこの精度以下を達成し、両者の差異を確定させなくてはならない。

理論値としては、2021 年発表の素粒子標準理論としての世界的な合意値が広く知られている。QED だけでなく電弱相互作用の寄与、とりわけハドロンの寄与について、世界中の多くの関連研究者が慎重な議論と検討を経た上で、素粒子標準理論としての理論値とその不確かさを定めたものだ。実験については、現在、日本の J-PARC 及び米国のフェルミラボで別個に国際共同実験としてミュー粒子 $g-2$ 実験が進められている。2021 年には、まず、フェルミラボから第一報が出た。従来のブルックヘブン米国国立研究所での実験値が再確認され、理論値との差は 4.2σ にまで拡大した。その後、2023 年の追加実験の結果発表により、2020 年での標準模型での理論値との差は 5.0σ までに拡大している。一方で理論値サイドでも、2020 年の格子 QCD によるハドロンの寄与の評価や、2023 年の $\pi^+\pi^-$ 生成実験データを用いたハドロンの寄与の評価は、2023 年のミュー粒子 $g-2$ の実験結果と理論値の間に矛盾はないことを示唆している。これらの理論値と従来の理論値との間の矛盾は、格子 QCD の圧倒的大規模シミュレーションによるハドロン階層の計算（4.1.2 節）により解決できるはずである。綿密な理論と実験の双方の慎重極まる精査を経た上でも、既存物理学のブレークスルーへの期待が未だ存続し続けていること、これがミュー粒子 $g-2$ 実験への高い関心となっている。

QED $g-2$ 摂動計算の今後 5 年～7 年における目標の一つは、実験精度の飛躍的向上に対応すべく、10 次の展開係数の高精度計算を行い、この係数をできれば 5～6 衔で値を確定することである。同時並行として、更に高い次数の展開係数である 12 次の計算もターゲットとなる。12 次の展開係数は、電子の将来実験の精度と比較すると必須になると考えられ、2 衔程度の値は確定させたい。ミュー粒子では、新実験の精度では重要ではないものの、有意な寄与となることが、主要項の計算によってすでに示されている。

解決に必要な計算手法とアプリケーション

QED 高次摂動モンテカルロ積分計算^(d)、4 倍精度実数計算、計算機の仕様にあわせて高度に最適化された 4 倍精度および 8 倍精度実数計算ライブラリ。

上付き (d) については 2.1.3 節を参照。

計算量・計算資源

素粒子の反応は、視覚的に理解しやすいように图形化された「ファインマン図」で表される。展開係数 $A_{2n} (n \geq 1)$ は n 重のループを含む複数のファインマン図で表され、各ループが運動量空間における多次元積分に対応する。数値計算では、これをファインマン-パラメタ空間での多次元積分に変換することで積分の次元を下げたのち、モンテカルロ法による数値積分を行う。例えば、10次 A_{10} の場合は 5 ループを含む 1 万個以上という膨大な数のファインマン図を数値的に評価する必要がある。ファインマン図の数値計算における特徴として以下の点が挙げられる。

- 被積分関数が長大であること。10次ではすべてのファインマン図の寄与を合わせると 1 億行程度の長さになる。
- 紫外（運動量が大きい）および赤外（運動量が小さい）領域での発散の処理手続きにおいて大きな桁落ちが発生するため、擬似 4 倍精度、一部には擬似 8 倍精度の実数計算が必要である。
- モンテカルロ積分において、多次元運動量空間（8~13 次元）の各点における被積分関数の評価は独立であるため、通信をほとんど必要としない embarrassingly parallel 的な処理が可能である。
- 被積分関数の評価点の数で並列化可能であり、外部アクセラレーターを含めて、どのような階層性を持つ計算機の構成にも対応できる。

この計算に必要な要求性能の数値を第 2.1.5 節に挙げている。

10次の計算では、すでに過去の約 1 万個のファインマン図、全ての計算結果から 1 ~ 2 桁の数値が得られている。これにさらに 1 桁を積み増すためには、従来の約 100 倍のリソースを投入すれば良い。つまり、10次のファインマン図 1 個の値を、 1×10^{12} 個の評価点を擬似 4 倍精度実数を用いて計算すれば良い。計算量は、倍精度計算に換算して約 20 EFLOP である。ただし擬似 4 倍精度計算 1 回は倍精度計算 1 回のおよそ 20 倍とした。積分値の収束が遅い図形は 400 個程度あり、それらの平均評価回数を 10 とすると、10次の展開係数の総計算量は **80,000 EFLOP** と見積もられる。

一方、12次のファインマン図は、その数が約 20 万個となり、1 個の被積分関数の長さは 10 次の 10 倍程度である。しかし、精度の目標が 10% 程度なので、ほとんどの図は倍精度計算で評価を実行すれば十分だ。ミュー粒子 $g-2$ に対しては、主要寄与を与えるものだけを勘定すれば良く $\mathcal{O}(1000)$ 個で済むと見積もられる。電子 $g-2$ に関しては全図の計算が必要である。評価点の数を 1×10^9 程度とし、ほとんどを倍精度計算で実行するとすれば、10次の計算と同程度の計算量と見積もられる。従って、10次と 12次の計算を完了するために必要な総計算量は **160,000 EFLOP** である。

必要メモリ容量は並列化の実装の詳細によるが、コア当たり 1GB 程度を目安として見込んでおけばよい。メモリの多くは、データの保存ではなく、被積分関数を構成するための演算命令のストックに用いられる。1,000 コアを利用したとすると、**使用メモリ総量は 1 TB** である。また、計算に必要なストレージ容量は総計 **1TB** で十分である。

4.1.17 ファインマン振幅自動計算 [長期目標 (b-1) の課題]

課題概要と長期目標

大型加速器などによる素粒子反応実験のデータ解析では、理論から素粒子の反応確率や運動量分布などを計算して実験データと比較する。また、実験データには特徴を示すシグナルとなる現象とバックグラウンド現象が混ざっており、実験データを解析して理論計算を行い有意な結論を導いている。この素粒子反応の確率や分布の計算は相対論的量子場理論を摂動論的に計算する手法が定式化されている。摂動展開を数値的に評価する場合は展開係数をファインマン図で表し、それぞれの図形に含まれるループが表現する多次元積分を計算する方法が一般的である。摂動展開の次数が上がるごとに、ファインマン図に含まれるループの数が増え、かつファインマン図の数が指数関数的に増えていくため、次数が高い展開係数ほど計算が急激に難しくなっていく。前出の QED の高次補正計算もこの一例であるが、この課題では QED に限定せず、標準理論全体およびそれを超える新しい理論を計算対象とする。多種多様な素粒子反応を、この手法で数値計算するシステムをファインマン振幅自動計算と呼ぶ。

このようなシステムは、欧州でも開発されてきており、90 年代には CERN の LEP-II 実験、現在では LHC 実験などのデータ解析に貢献している。

将来建設が検討されている国際リニアコライダー (International Linear Collider : ILC) のような、素粒子の衝突エネルギーが高エネルギー・フロンティア型の実験では、重要なデータを得るために終状態が多粒子系となる場合が多い。また、反応を検知する測定器の感度が向上することにより、より精密で高精度な実験データが得られると予測されている。この 2 点に対応するため、理論計算においても複雑な終状態での精度の高い計算が要求される。また標準理論だけではなく超対称性を取り入れた新たな理論による計算も実行可能であり、シグナル現象を予測するなどして未知のエネルギー領域での新粒子・新現象の解明にも必要となる。日本のグループでは、1 ループを含むファインマン振幅自動計算により ILC のエネルギー領域では QCD と電弱理論の 1 ループの寄与が各々 10% 程度あることを数値的に示した。しかし、計画では ILC の実験精度は 1% 以下になるので、更なる高精度計算のために 2 ループ図形についての自動計算が必要でありこのシステムを完成させることが今後 5 – 10 年の課題となる。具体的には、2025 年ごろに電子・陽電子衝突でトップクォーク対生成（終状態 2 体）の電弱相互作用及び QCD の混在した 2 ループ図形計算、2030 年ごろに全電弱相互作用における 2 ループ図形計算を可能にする。これにより、将来のエネルギー・フロンティア領域での電子・陽電子衝突実験の精密測定に寄与することができると考えている。

解決に必要な計算手法とアプリケーション

基本的に QED の高次補正計算と同じ枠組みの計算 (2.1.3 節 (d) を参照) であるが、大きな違いとしては、理論が QED に限定されない点、反応の終状態が多岐にわたる点が挙げられる。これにより QED の高次補正計算よりも格段に複雑な計算となっている。また、レプトン、クォーク、ゲージ粒子、ヒッグスボソンという、質量が 6 衡も異なる粒子を扱うことや、様々な数値的相殺が発生することから、現在 4 倍精度計算が標準的であり、更なる高精度（6 倍、8 倍精度）を用いた計算も必要と考えられる。

ループ積分を実行する際の多次元積分方法として、二重指數関数型などの数値計算公式や、適応型積分公式を用いる。前者は SIMD 型で超並列計算可能であるが、適応型の場合には、並列度を高めるために並列化アルゴリズムが必要になる。

計算量・計算資源

ファインマン図の数は、ループの個数と共に指數関数的に増加していく。例えば、ILC のような電子・陽電子衝突実験における、トップクォーク対生成の電弱相互作用素粒子反応過程では、ループの無い図形の数は 4 個であるのに対して、ループ 1 個は 150 個、ループ 2 個は 31,193 個となる。

これらのファインマン図の計算を行うには、ループに対応する運動量積分を実行する必要があり、1 ループでは 3 次元積分、2 ループでは 6 次元積分が必要になる。またそれ以外に、素粒子反応の全確率を求めるためには、終状態の位相空間に対して積分が必要であり、積分の次元は終状態の粒子数に応じて増加していく。終状態が最も簡単な 2 体では 1 次元積分、3 体では 4 次元積分、4 体では 7 次元積分となる。

電子・陽電子衝突におけるトップクォーク対生成の場合に、電弱相互作用の 2 ループ図形で、運動量積分が最大の 6 次元の図形の数は 2,000 個である。1 次元あたりの評価点は 100 とし、2 体終状態の位相積分評価点は 10 とする。これらの図形 1 個の計算に、運動量積分と位相積分を 400 回程度実行する。それ以外の図形の計算は、この図形計算の 10 倍とすると、全体で 80×10^{18} 計算が必要である。この計算を擬似 4 倍精度で計算すると、倍精度計算の約 15 倍の計算が必要なので、この過程の 2 ループ計算の総計算量は **1,200 EFLOP** と見積もることができる。一方で、自動生成プログラムによるファインマン図の処理、積分アルゴリズムの改良によって、格段に計算量を減らすことも可能かもしれない。また、GPU による計算加速も可能である。使用メモリとストレージについては概ね第 4.1.16 節 QED の高次補正計算と同じでかつ使用量も少ないので、詳細な見積もりは省略する。

第 2.1.3(vi) 節および第 4.1.16 節にある QED 計算よりも複雑な計算が必要であるため、計算量は多くなるが、計算の枠組は同じであるため、計算機要求性能は共通する部分が多い。

- 高性能 4 倍精度浮動小数点計算が可能な計算機

4 倍精度の演算器を有する計算機、もしくは、倍精度演算器を利用して 4 倍精度計算を可能とする計算機

- 高度な SIMD 命令を生成できるコンパイラ及び#pragmaなどを用いたハードウェアに依存しないソフトウェアプログラミング環境
- 計算式が長大になるため、コードの簡約化などの技術
- 4 倍精度の高度なライブラリ
- 将来的には、多倍長計算が可能な計算機と高度なライブラリ

4.1.18 超弦理論シミュレーションで探る時空の量子ダイナミクス [長期目標 (b-2) の課題]

課題概要と長期目標

宇宙創成の機構の解明を理解するには、アインシュタインの一般相対論にミクロの世界で重要な量子論の効果を取り入れる必要がある。超弦理論は、プランク長さ (10^{-35}m) 程度に広がった弦を宇宙の最小構成要素と考えることで重力をうまく量子化し、20世紀の2大理論である相対論と量子論を統一的に記述することに成功している。超弦理論のダイナミクスを行列模型の数値計算から明らかにし、宇宙創成のシナリオや自然界の究極的法則の確立など、今後100年のうちに解決されるべき物理学の難問に挑戦する。

素粒子の標準理論とビッグバン宇宙論は、加速器実験や宇宙観測などにより精密に実証されており、これらを重力を含んだ微視的な理論から導くことは、21世紀の物理学の大きな目標の一つと言える。本課題では、重力を含んだ微視的な理論の最有力候補と考えられている超弦理論から、標準理論とビッグバン宇宙論を導ける可能性を明らかにする。そのために、超弦理論の非摂動的定式化として提唱されている行列模型の数値シミュレーションを実行し、宇宙の膨張の仕方や相関関数などを測定する。そこから、時空が4次元である理由を解明し、さらに、初期宇宙の密度揺らぎや時空を伝播する物質粒子の種類を同定することで、現実の宇宙がどのようにして生じるかのシナリオを確定する。

解決に必要な計算手法とアプリケーション

行列模型の数値シミュレーションにおいては、多変数積分における被積分関数が複素数値をもつため、通常のモンテカルロ積分が適用できない。この問題を回避するために、複素ランジュバン法やレフシェツ・シンブル法を呼ばれる方法が開発されてきた。これらの方法は、基本的には作用の微分を計算しながら行列の配位を更新していくという計算に基づいており、 10^3 次元程度の密行列に関する行列行列積を頻繁に評価する必要がある。

既に、複素ランジュバン法に関しては効率のよい計算コードが開発されており、実際の計算に使われている。但し、この方法を正当化するには一定の条件を満たす必要があり、これまでの結果から、行列模型のフェルミオン作用に変形項を導入しなければならないことが明らかになった。この変形項の影響を小さくするには限界があり、将来的に、その限界を越えた計算を行うには、レフシェツ・シンブル法を適用する必要があると考えている。まずは、複素ランジュバン法を用いた計算により、その限界までの計算を遂行し、それと併行して、レフシェツ・シンブル法に基づく計算コードを開発していく計画である。

また、上記の複素ランジュバン法のコードを用いたシミュレーションでは行列をブロックに分割して、各ブロックを1つのMPIプロセスが担当する並列計算が行われている。将来的には、より大きな 10^4 次元程度の行列のシミュレーションに対応するため、GPUを用いた計算用のコードを開発し、計算の高速化が図れるかを検討する必要がある。

計算量・計算資源

理論の構造を理解するためには、行列サイズが 1024 で少なくとも 10 個程度のパラメータについて計算することが望まれる。複素ランジュバン法のシミュレーションでは、充分な統計を得るために各パラメータに対して 10 ランジュバン時間の長さのシミュレーションを行う必要があり、10 個のパラメータについての総計算量は **260,000EFLOP** と見積もられる。1 個のパラメータについて使用メモリ量は **10TB** であり、各パラメータについて 1,0000 個の行列配位を保存するとすると **40TB** のストレージを必要とする。

4.1.19 場の量子論の非摂動ダイナミクス [長期目標 (b-3) の課題]

課題概要と長期目標

超対称性を有する場の量子論は、様々な観点から物理的に意義深い帰結をもたらすと期待されている。自然界の基本的な相互作用を統一的に記述する理論を模索する取り組みが続けられているが、これらはトップダウン的・ボトムアップ的な研究手法に大別される。本課題ではこの双方の方針に基づいて相互作用の統一的記述を与える究極理論の片鱗を掴むことを目指す。

トップダウン的な手法は、量子重力理論の有力候補とされる超弦理論の非摂動的な性質の理解が挙げられる。超弦理論の文脈で提唱されたゲージ重力双対性は、あるクラスの場の量子論と重力理論との間の等価性を予言する。対応関係が最も明確なのは極大超対称ゲージ理論と II 型超弦理論の双対性である。超対称ゲージ理論は格子正則化を用いて数値シミュレーションが可能なので、大規模数値シミュレーションによる超弦理論の非摂動的な性質の解明が期待できる。一般的に数値解析では、低次元模型の方が計算コストを含め多くの恩恵を受けやすいが、双対性の検証では BFSS 行列模型や BMN 行列模型等の低次元でありながら 4 次元理論に勝るとも劣らない物理的に興味深い対象が存在する。本課題での進展を通じて、超弦理論の背後に存在するとされる M 理論の性質の理解へつながる可能性も示唆されている。またゲージ重力双対性の拡張から、QCD、超伝導、超流動現象といった従来より解析計算が困難な物理の進展を数値的な側面からも後押しする。ここで、双対性の有効性が保証されるパラメータ領域は場の量子論の強結合領域に当たるため、格子正則化を用いた数値計算による研究の意義が大きいことを強調したい。

ボトムアップ的な研究の代表的な例として、標準模型を超える物理を軸とした究極理論の探索が挙げられる。素粒子標準模型を超える物理として長年期待されているのが超対称性を持つ理論である。超対称性の存在は、大統一理論の構築やヒッグス粒子の量子論的安定性、暗黒物質の正体の特定などの様々な物理を整合的に記述できる素地を持つ。この超対称性は我々の住む低エネルギー領域では自発的に破れていなければならぬが、破れた先の真空の安定性の解析は様々な困難が伴うことも知られており、数値的手法による超対称性の破れの機構や理論のダイナミクスの研究はその貢献度が高い。具体的な課題としては、**1. アフレック・ダイン・サイバーグモデルやイントリゲータ・サイバーグ・シィモデルといった種々の模型・理論的アイデアの系統的な解析による有力な模型の大まかな探索、** **2. 超対称性の破れた真空の安定性解析、** **3. 加速器実験等の観測量との比較に備えた有望な模型の低エネルギー領域における真空以外も**

含めたスペクトルの数値的な精密測定、などが挙げられる。これらの研究課題には格子 QCD で従来より用いられているテクニックの多くが活用でき、新規参画が比較的容易だと考えられる。多くの超対称模型の研究は超弦理論の研究と密接に関係し合いながら進展し続けていることから、前述のトップダウン的手法との連携も見込める。

この他にも、超対称性に限らず、場の量子論の持つ非摂動的な性質に着目することで、理論空間の持つ豊かな構造を明らかにし、超弦理論・素粒子現象論への展開を見据えた研究を展開することが重要である。従来より活発に調査されてきたウォーキングテクニカラー理論の非摂動的解析に加え、近年注目を集めている θ 項を持つトポロジカルな理論や非エルミート系・ \mathcal{PT} 対称な理論などの非摂動的な解析や、エンタングルメントエントロピーに代表される種々の情報量の計算など、場の量子論を介して素粒子理論・量子情報理論・物性理論等の複数の領域が関連する横断的な研究課題を推進することを目指す。

解決に必要な計算手法とアプリケーション

行列模型と場の理論のどちらも、物理量は多自由度の積分によって与えられるため、格子 QCD の場合と同様、ハイブリッドモンテカルロ (HMC) 法^(a) を用いた計算が可能である。具体例として、ゲージ群が $SU(N)$ の超対称ゲージ理論を考える。密行列に関する行列行列積を頻繁に評価する必要がある。超対称性を保つ計算をするためには、杉野格子作用或いはカプラン・カツ・ウンサル作用とよばれる格子超対称ゲージ作用を用いた計算を行うことが望ましい。この計算の主要部分は、格子 QCD で広く活用されている、大規模疎行列反復解法^(b) によるフェルミオン伝搬関数計算である。計算に現れる疎行列（ディラック演算子）の対角の帯部分には $N \times N$ の密行列を多数含み、そのサイズの行列行列積を複数回評価する必要がある。帯部分の $N \times N$ 行列以外は疎の構造を持つ行列であり、格子 QCD で用いられるのと同様の並列化手法が有効である。 N が大きい場合には、密行列の掛け算も並列化することが望ましい。計算にフェルミオンの複素位相の効果を取り入れる際には、近年確立してきた複素作用問題の解決策である複素ランジュバン法やレフシェツ・シンブル上の HMC 法等の適用が可能である。今後開発される計算機資源に適したアルゴリズムの開発・改良、および高速化を進めることも重要である。とくに低次元模型については、量子計算機の発展次第で今後 10 年以内に量子計算が可能になるかもしれない。古典計算アルゴリズムにとどまらず、量子計算のアルゴリズムの開発・改良も重要である。

上付き (a), (b) については 2.1.3 を参照。

計算量・計算資源

ここでは、超対称ゲージ理論の計算での計算量の見積りを述べる。この見積りは京コンピュータを用いた 2 次元超対称ゲージ理論の計算を元に行った。京コンピュータでの計算は、2 次元格子サイズ 16×64 でゲージ群の次元は 6、フェルミオンのスピノル次元は 16 で行い、128 ノードを用いた場合、1 パラメータの計算に 2000 時間かかった。このときの実行性能は 5% だとすると、5.9 EFLOP 用いたことになる。

将来計画している計算は、1 次元格子サイズ 128 でゲージ群の次元が最大で 256、スピノル次元は 16 を用いる。計算量はゲージ群のサイズ N の $5/2$ 乗、スピノル次元と格子サイズの $5/4$ 乗に比例すると仮定した。理論の性質を解明するためには、様々なパラメータでの計算が必要であり、全体で 20 パラメー

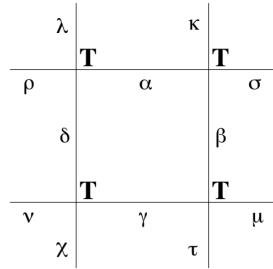


図 4.2 (4.10) 式が表すテンソルネットワーク。2次元の正方格子を仮定している。

タの計算を行うとすると、総計算量は **105,000 EFLOP** である。また、計算では各格子サイズで複素 256×256 行列を 300 個程度メモリに確保する必要があるため **40 GB** のメモリ量が必要になる。また、各パラメータに付き 1,000 個の配位を保存するとして **30 TB** のストレージが必要になる。

4.1.20 テンソル繰り込み群による素粒子物理学および物性物理学の研究 [長期目標 (a-2) の課題][長期目標 (b-3) の課題]

課題概要と長期目標

テンソルネットワーク (TN) スキームとは、多体問題を TN 形式によって定式化し、高精度解析を行う一群の理論的・計算手法的枠組みである。TN スキームには、大きく分けて、経路積分形式に基づくテンソル繰り込み群 (TRG) 法とハミルトニアン形式に基づくテンソルネットワーク状態法がある。TRG 法は既存の数値計算手法 (モンテカルロ法など) と異なり、(i) 原理的に符号問題・複素作用問題がないこと、(ii) 計算コストの体積依存性が対数的であること (一辺の長さ L の d 次元格子体積 L^d に対する計算コスト $\propto d \times \ln L$)、(iii) グラスマン数 (反可換性を持つ数) を直接扱えること、(iv) 物理量の期待値だけでなく分配関数 Z そのものを計算できること、などの魅力的な特徴を有している。中・長期的目標は、TN スキームを素粒子物理学分野および物性物理学分野に応用し、上記の利点を活用することによって、これまでの計算物理学が成し得なかった新たな物理研究の開拓を目指す。今後 10 年における具体的な目標としては、4 次元有限密度 QCD、 θ 項を持つ場の理論の非摂動的解析、(3+1) 次元および (2+1) 次元の Hubbard モデルや t-J モデル (経路積分形式) の相構造解析などが代表例として挙げられる。

解決に必要な計算手法とアプリケーション

まず、2007 年に Levin と Nave によって提案された TRG 法のアルゴリズムを 2 次元イジングモデルを例にとって説明する。TRG 法では、最初にモデルの分配関数を局所的 (格子点) に定義されたテンソルの積で書き表す (図 4.2)。

$$Z = \sum_{\text{all spin states}} \exp \left(-\frac{1}{k_B T} \mathcal{H} \right) = \sum_{\alpha, \beta, \gamma, \delta, \dots=1}^2 T_{\alpha \lambda \rho \delta} T_{\sigma \kappa \alpha \beta} T_{\mu \beta \gamma \tau} T_{\gamma \delta \nu \chi} \cdots \quad (4.10)$$

ここで、 $\alpha, \beta, \gamma, \delta, \dots$ は隣接するテンソル間のボンドを表す添字である。この表式は 2 次元正方格子を

仮定しているが、それ以外の場合でも、相互作用が局所的であれば必ず上式のような TN 形式で表せることが知られている。もちろん、このテンソル積の添字に関する縮約をすべて実行してしまえば厳密な分配関数 Z の値が得られるが、それでは自由度が巨大過ぎて、たとえ最先端のスーパーコンピュータであっても計算可能な格子サイズは非常に小さなものに限られる。そのため、Levin と Nave は、特異値分解に基づいた重要度の高い自由度の選択とブロック変換の一種による疎視化を組み合わせた手続きを反復することにより、分配関数 Z の値そのものを高精度で計算することが可能なアルゴリズムを提唱した。重要度の判断は、特異値の大きさに基づいて行う。具体的には、特異値の大きい順番に D_{cut} 個の自由度を残し、その他の自由度は棄ててしまう。TRG 法をコントロールするパラメータはこの D_{cut} のみである。なお、特異値分解とは行列の近似手法であり、画像データの圧縮など幅広い分野に応用されている。Levin と Nave が提唱したオリジナルな TRG 法は 2 次元スピニモデルにしか適用できないアルゴリズムであったが、その後高次元(3 次元以上)に適用可能な HOTRG(Higher Order TRG) 法が開発された。更に、HOTRG 法よりも計算コストを減らした(ただし同じ D_{cut} の場合、HOTRG よりも計算精度は劣化)ATRG(Anisotropic TRG) 法も開発されている。筑波大学を中心とするグループは、TRG 法の量子的場の理論への応用に取り組んでおり、TRG 法、HOTRG 法、ATRG 法をグラスマン数化するとともに、HOTRG 法と ATRG 法の計算コードの並列化を完了している。HOTRG 法、ATRG 法(グラスマン数化された場合を含む)における主要な演算はテンソルの縮約計算であり、この計算は行列行列積として実装可能である。行列行列積は演算律速であり、GEMM を用いることによってスーパーコンピュータおよび GPU クラスタの高い演算性能を十分に活用することが可能である。

計算量・計算資源

TRG 法では、モデル依存性は初期テンソルにのみ存在し、その後の繰り込み手続きはモデル依存性を持たないため、問題サイズ(D_{cut} の大きさにより決まる)と繰り込み回数(系の体積を決める)が同じであれば、モデルの詳細にかかわらず、計算コストも同程度と考えて良い。具体的に、最大 1024^4 の格子サイズでの 4 次元有限密度 QCD の相構造解析を想定してみる。問題サイズは 20^8 テンソル($D_{\text{cut}}=20$)で、繰り込み回数は 40 回とする。この場合、ジョブあたりの必要メモリサイズは 5TB であり、総計算量の見積もりは 7,200,000 EFLOP となる。ストレージはほとんど消費しないので、具体的な見積もりは省略する。

4.1.21 今後期待される計算手法(機械学習、量子計算)

課題概要と長期目標

多くの分野と同様に、素粒子・原子核物理学においても近年の機械学習の目覚ましい発展を踏まえた新しい計算手法が提案されている。また、量子計算機向けのアルゴリズムも盛んに議論されている。これらの手法は、これまでの項目で挙げている演算量を劇的に減らす可能性を秘めている。サイエンスではなく手法を軸にした切り口になるが、重要性を鑑み、今後 5-10 年で素粒子・原子核物理学に期待される応用の概要を述べる。機械学習を用いた研究課題は 4.1.4 節でも解説している。

解決に必要な計算手法とアプリケーション

機械学習を用いた計算手法は、詳細な計算抜きに何らかの物理量を推定することと、パラメータチューニングに大きく分けられる。前者では、相図や状態方程式の推定や、格子 QCD において配位から（厳密な計算をせずに）物理量の値を推定するなどが考えられる。ただ、推定の妥当性を検証できないまま物理の結果として用いることはためらわれる。その中で、格子 QCD の配位生成に機械学習アルゴリズムを組み込む自己学習ハイブリッド・モンテカルロ法 (self-learning hybrid Monte-Carlo method) は、機械学習による推定を用いつつ、厳密なアルゴリズムになっている。この手法が大規模な格子 QCD シミュレーションでも有効であれば、計算時間を大幅に減らすことができるかもしれない。符号問題のある系では、符号問題の影響が少ないように問題を変形することがある。機械学習を用いてこの変形に関するパラメータを決める手法が提案されている。加えて、ベイズ最適化は一般に計算アルゴリズム上のパラメータの選択にも利用できる。シミュレーションの設定に応じて、計算時間が短くなるようにアルゴリズム上のパラメータを調整する必要が生じることがあるが、これを勘と経験によらず機械学習に基づいて調整すれば、従来より効率よくパラメータを決められるようになると期待できる。

量子計算の利点は、符号問題のある系にも適応可能な点である。符号問題のある系は既に前項まで挙がっているが、有限密度系、実時間発展、 θ 項つきのトポロジカル理論などがある。これらの系をハミルトン形式で記述し、適切な状態での期待値を計算することになる。ノイズの影響を受ける NISQ デバイス上では、変分法を用いた計算が挙げられる。これはエネルギーの期待値を量子計算で行い、変分パラメータの最適化を古典計算機で行うハイブリッドな計算が可能である。量子誤り訂正付きの qubit が多数利用できるようになると応用は更に広がる。今後 5-10 年で利用可能な qubit 数でも 1+1 あるいは 2+1 の低次元系の量子シミュレーションが可能になると思われる。低次元系であっても、とくに実時間発展の量子シミュレーションが可能になれば散乱問題を直接扱えるようになる。また、ホログラフィー原理を通じて、量子重力の記述に用いる提案もなされている。

計算量・計算資源

既に記載してた課題の計算を加速させる手法なので、機械学習についての見積もりは省略する。量子計算については、今後 5-10 年で取り組みたい課題に必要な量子計算機の qubit 数について見積もる。符号問題のある系について現在の古典計算では難しい新規の計算結果を得るには、量子誤り訂正付きで数十ないし $O(100)$ qubit が必要になる^{*4}。これは N qubit の量子シミュレーションを古典計算機で模倣するには、ランクが 2^N の行列の対角化と同程度の演算量が必要との仮定に基づく。2+1 次元での場の量子論の時間発展を扱うには、少なくとも量子誤り訂正付きで $O(100)$ qubit が必要になる。

^{*4} テンソルネットワーク法を用いた計算ができる場合には、 $O(1000)$ qubit 相当まで古典計算機で計算可能。テンソルネットワーク法が適用できない系・物理量の計算で量子計算が利点が活きる。また、テンソルネットワーク法に NISQ デバイス上の量子計算を組み合わせる手法も議論されている。

4.1.22 データグリッド

課題概要と長期目標

格子 QCD 計算では、格子上の経路積分をモンテカルロ法を用いて計算している。このモンテカルロ法ではアンサンブルとしてゲージ配位（グルーオン場）データを格子 QCD の作用積分に基づく統計分布に従うように生成する。物理量の量子論的期待値はアンサンブルの各ゲージ配位上で計算された物理量の統計平均として計算することが出来る。現在主流のモンテカルロ法は真空中のクォークの仮想的対生成・対消滅の効果を取り入れることが出来る格子 QCD ハイブリッドモンテカルロ法^(a)である。この計算方法ではクォークの仮想的対生成・対消滅の効果の計算のためクォークの運動方程式を大規模疎行列反復解法^(b)を用いて大量に解く必要がある。ここには多くの計算資源を投入することになる。

このようにして生成されたゲージ配位のアンサンブルはデータとして長期保存され、さまざまな物理量の計算に使われる。またゲージ配位のアンサンブルを公開することで、ゲージ配位生成を行わない研究グループが物理量を計算することが出来るようになり、格子 QCD を用いた物理学研究の進展が促進される。

重力崩壊型超新星爆発のメカニズムの理解には流体とニュートリノ輻射の運動を同時に扱う一般相対論的輸送流体計算が必要不可欠である。特にニュートリノ輻射による加熱が超新星爆発には重要だと考えられている。ニュートリノ輻射輸送の計算では Boltzmann 方程式を解く。これはニュートリノの位相空間（座標と運動量）での 6 次元分布関数の時間発展をつかうものであり、高精度な計算のため位相空間の解像度を上げると急速に計算規模とデータサイズが大きくなる課題である。

このような一般相対論的輸送流体計算で生成されたデータを後解析や可視化することで重力崩壊型超新星爆発のメカニズムの理解につなげる必要がある。高解像度大規模データの後解析や可視化は、大規模計算を行った拠点以外の計算機で行うため、大規模データの共有や移動を効率的に行うことがニュートリノ輻射輸送計算では必要となってきている。

以上のように貴重な計算時間を費やして生成されたデータを長期保存^{*5}・公開・共有することは、研究を推進するために必要不可欠となっている。このような大規模データを用いる素粒子・原子核・宇宙分野の計算科学研究を円滑に進めるために 2008 年から Japan Lattice Data Grid (JLDG, <http://www.jldg.org/jldg/> [2]) という広域データグリッドが日本国内で運用されている（図 4.3 JLDG システム概念図）。2023 年において JLDG に参画している拠点として、筑波大学計算科学研究センター、高エネルギー加速器研究機構計算科学センター、大阪大学核物理研究センター、東京大学情報基盤センター（柏キャンパス）、京都大学基礎物理学研究所、理化学研究所仁科加速器科学研究センター、理化学研究所計算科学研究センターがある。JLDG では分散ファイルシステムである Gfarm システム [3] を用いて、これらの拠点を統括するように広域データグリッドを構築しており、さらに各拠点の HPC システムと連携している。また、JLDG システムは、2007 年から運用している国際格子 QCD データグリッド International Lattice Data Grid (ILDG, <https://hpc.desy.de/ildg/>) とも連携しており、JLDG で公

^{*5} 研究で使用・生成されたデータは論文執筆・出版後 10 年間保存することが次のように求められている。文部科学省による「研究活動における不正行為への対応等に関するガイドライン」を受けた、日本学術会議による「科学研究における健全性の向上について（回答）」に記載されている

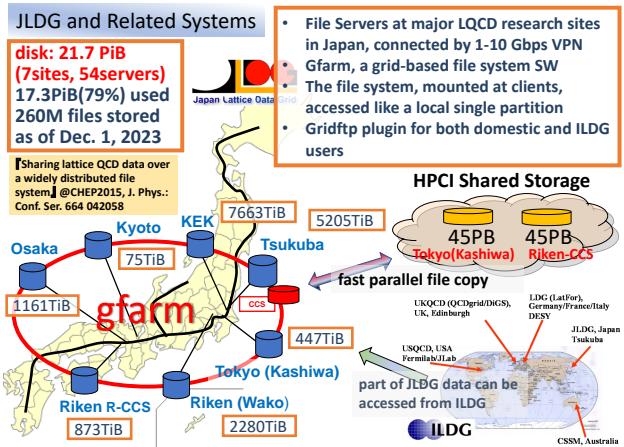


図 4.3 JLDG システムの概念図。

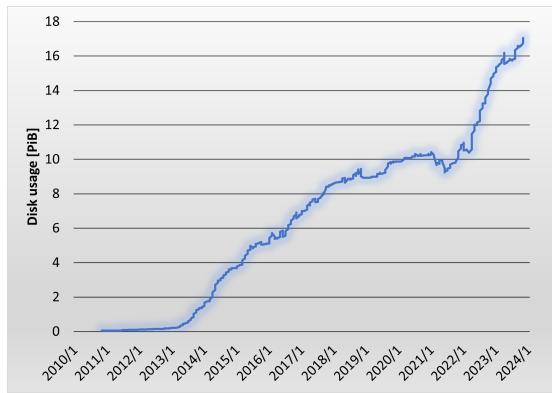


図 4.4 JLDG システムのディスク使用量の推移 (単位は PiB)。

開されているデータを ILDG へ供給することも出来るようになっている。

この広域分散グリッドシステムではデータの複製が行われており、万一どこかの拠点がダウンしてもデータが維持されることを保証している。また、複数の異なる研究拠点においてデータの解析やシミュレーションが分担されて行われることが日常となっており、広域分散グリッドシステムを用いることで拠点間でのデータの共有や移動を明示的に行う必要がなくなり利便性が向上している。さらに、2013年12月から HPCI 共用ファイルシステムと JLDG システムを連携する仕組みを運用開始しており、HPCI システムを利用して計算したデータを JLDG システムに移動することができ、各プロジェクト終了後もデータ保持や共有が出来るようになっている（図 4.3 の「HPCI Shared Storage」部）。日本の各地にいる研究者が近隣の JLDG 拠点にアクセスすることでデータ移動や共有の手間を省くことが出来る。

現在の JLDG の総データ容量は HPCI システムとの連携開始後急速に増加している状況（図 4.4）であり、2023 年 12 月時点で総容量約 22 PB のうち約 17 PB を使用している。現在は格子 QCD を用いた研究を行なっているグループによるデータ量が支配的であるが、超新星爆発に係るデータも増えてきている。2021 年にはスーパーコンピュータ「富岳」が稼働し、また、各大学・研究所にても「京」の性能を超える HPC システムがすでに稼働しており、今後も出力されるデータ量が増えていくことが予想される。2026 年頃のエクサスケールの HPC システムから生成されるデータの保持に必要な容量を確保していくことが必要である。JLDG システム自身は直接科学を行うことを目的としていないが、上述のような基礎物理学のための科学計算に必要不可欠なものである。

解決に必要な計算手法とアプリケーション

現在の JLDG の各拠点のファイルサーバは、国立情報学研究所が提供する SINET5 上の L3-VPN（レイヤ 3 バーチャルプライベートネットワーク）によって接続されている HEPnet-J/SC（高エネルギーネットワーク日本/スーパーコンピュータ、KEK:HEPnet-J の歴史:<http://www2.kek.jp/proffice/archives/3w/>）と呼ばれるネットワークにより相互に接続されている。ファイルサーバ群は、産業技術総合研究所・筑波大学で開発された Gfarm システム [3] により

束ねられており、ユーザーからはあたかも単一の（パーティションの区切りがない）ファイルシステム（JLDG ファイルシステム）に見えるようになっている。

2026 年以降も、基礎科学分野のアプリケーションから出力されるまとまったデータの容量の増大や複数の計算機リソース拠点間でのデータの移動や共有の必要性が見込まれる。また、日常一般でのクラウド利用の利便性が向上している昨今においては、グリッドファイルシステムにおいてもデータ操作の利便性向上も重要である。これらの要求に答えるためには、高容量のネットワーク、高容量のデータストレージ、洗練されたグリッドファイルシステムに対する研究開発の推進、および、それらを利用した JLDG グリッドファイルシステムの構築・維持が必要である。

計算量・計算資源

本節で述べられた課題のうち大きなストレージ容量を必要とするものに対応したストレージ量を用意することが必要となる。

参考文献

- [1] T. Togashi *et al.* *Procedia Computer Science*, Vol. 29, p. 1711, 2014.
- [2] T. Amagasa *et al.* Sharing lattice qcd data over a widely distributed file system. *21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015) IOP Publishing, J. Phys.: Conf. Ser.*, Vol. 664, p. 042058, 2015. [doi:10.1088/1742-6596/664/4/042058].
- [3] Gfarm File System. Project page. <http://sourceforge.net/projects/gfarm/>.

4.2 ナノサイエンス・デバイス

本節ではナノサイエンス・デバイス分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.2.1 課題を解決するために必要なアプリケーション群（要求性能）

(i) 電子・電磁場ダイナミクスの実時間第一原理計算

光と物質中の電子との相互作用による超高速現象を動作原理として活用する光・電子融合デバイスを理論的に設計するために、電子・電磁場ダイナミクスの第一原理計算が不可欠である。量子力学的な効果を詳細に解析するためには、デバイスを構成するナノ構造体と極超短パルス光の相互作用に関する微視的な時間発展計算が必要となる。対象とする系は、1 辺十数 nm から数十 nm 程度の実在系ナノ構造体である。典型的な原子数は数万原子、時間ステップ数は数万ステップである。光・電子スピニ機能性を持ったデバイスを対象とするため、スピニ軌道相互作用を考慮する。超高速な情報処理を実現するために、照射光はフェムト秒あるいはアト秒の極超短パルス光を想定する。また、超微細レーザー加工のシミュレーションのためには、破壊閾値に達する超高強度光も考慮できる必要がある。これらの条件を満たすアプリケーションは SALMON のみである。

【アルゴリズムの説明】 時間依存密度汎関数理論に基づき、電子系の時間依存 Kohn-Sham 方程式（方程式の形は時間依存シュレディンガー方程式と同一）を実時間・実空間差分法を用いて解く。実空間の 3 次元グリッドについては、x,y,z 方向の格子点数を各々 N_x, N_y, N_z とすると、総グリッド点数は $N_x \times N_y \times N_z$ となる。主要な演算は、時間発展計算の各時間ステップで行うハミルトニアンの軌道波動関数への作用である。この演算にはラプラシアンの作用と擬ポテンシャル非局所項の作用が含まれており、前者は差分法のステンシル演算、後者は疎行列・ベクトル積である。この他、ハートリーポテンシャルに関するポアソン方程式を解くための FFT 演算も計算コスト的に重要である。電磁場のマクスウェル方程式についても波動関数の時間発展と連立させ、実空間 3 次元グリッド上で差分法を用いて解くが、時間依存 Kohn-Sham 方程式に比べればその計算負荷は格段に低いので相対的に無視できる。計算コストのスケーリングについては、軌道波動関数の軌道本数と実空間グリッド点数の両方が系の大きさに比例するため、 $O(N^2)$ となる。また、時間発展計算では誤差の蓄積が問題となるため倍精度を用いる。

【想定する計算機】 性能を富岳の約 30 倍と仮定し、1 ノードあたり 100TFLOPS の計算機を想定する。

【想定する計算】 SALMON では 1,000 から 10 万原子規模の多様なナノ構造体について、様々な光パルス条件のもとで数値的なポンプ・プローブ実験やレーザー加工シミュレーションを行うことを想定する。光強度や偏光方向、位相、遅延時間などを変えながら多数の数値実験を行う必要があるため、試行回数は一つの系について 100 回程度となる。ここでは一例として、遷移金属ダイカルコゲナイト (TMDC) 薄膜の上に電極として金属ナノ構造体を付加した系を考える。TMDC は代表的なスピントロニクス材料であり、円偏光の電場によってバレー及びスピニ偏極した電流を誘起できる。これにナノ電極を付加することで典型的な光・電子融合デバイスを模擬する。Au のナノ構造体と TMDC 層を合わせて約 2 万原子、価電子数を約 20 万とし、実空間格子点数は $600 \times 600 \times 600 = 216M$ とする。スピニ軌道相互作用を

考慮すると軌道本数は40万となる。

富岳におけるアモルファス SiO₂の1万原子計算での実績をもとに計算コストを見積もる。格子点数 $216 \times 288 \times 480$ 、軌道本数 26,604、富岳 27,648 ノードの場合に、一つの時間ステップあたり約 1.2 秒かかった。このため、100TFLOPS/node の計算機を 10 万ノード用いれば、上記の課題を 1 ステップあたり約 1 秒で計算できることになる。一つの試行について典型的には約 5 万ステップの時間発展が必要なため、一回の計算時間は約 14 時間である。

【総演算量】 1 ステップ当たり 216M(総メッシュ数) \times 0.4M(総軌道数) \times 30 回(差分法演算回数) \times 5 回(テーラー展開 4 次 + 軌道エネルギー評価) 13P 回。

【総メモリ量】 計算グリッド上に波動関数をストアしておくためのメモリがほとんどを占める。216M(総メッシュ数) \times 0.4M(総軌道数) \times 16B(複素数) \times 3(作用前波動関数用配列、作用後波動関数用配列、テーラー展開での総和の波動関数用配列) 4PB。

【ネットワークバンド幅】 1 ステップ当たり 216M(総メッシュ数) \times 8B(実数) 1.7GB の電子密度用データの Allreduce による通信が発生する。通信を全実行時間の 1 割と考えるとノード当たり 17 GB/s 必要となる。

【メモリバンド幅】 1 ステップ当たりの総メモリ量を 1 秒でアクセスするために 4PB/s 必要となる。

【オンチップメモリ】 1 軌道当たりに必要とされるメモリ量は 216M(総メッシュ数) \times 16B(複素数) \times 3(波動関数用配列数) 10 GB。空間を 125 分割 ($5 \times 5 \times 5$ 分割) して、各ブロック内で差分法演算に使うデータをすべてオンチップ上に載せるためには約 0.1GB 必要になる。

【ストレージ容量】 波動関数を入力として読み込むために、216M(総メッシュ数) \times 0.4M(総軌道数) \times 16B(複素数) 1.4PB が最低必要となる。

(ii) 強相関量子物質の分光学スペクトル計算

量子論の勃興期から現在に至るまで、分子や固体中の多体電子を始めとする量子多体系が示す、超伝導を始めとする古典力学からは予想できない量子力学的現象は、科学者のみならず、多くの人々を興味を掻き立ててきた。中でも、量子計算機の超越的な計算能力や量子暗号通信の安全性の源となる量子もつれは、新たな量子情報科学の時代の資源として興味を集めてきただけでなく、物理や化学、生命科学においても特異な現象や機能、例えば、光合成の高いエネルギー変換効率を生み出す源として研究が続けられてきた。

強相関量子物質の代表例である高温超伝導体においては、量子力学的な波として干渉できる可干渉性(コヒーレンス)を持った個別の準粒子が失われてしまうことが実験的に明らかになってきている。一般にコヒーレンスが失われれば量子もつれも消滅してしまうだろう。しかしながら、高温超伝導体では、コヒーレントな準粒子が失われているにもかかわらず、巨視的スケールにわたる量子もつれの産物である超伝導が確かに生じている。さらに高温超伝導体研究の中で提案された共鳴原子価状態は、非自明な量子もつれを示す典型例として量子物質と量子計算の双方の分野で研究が進められてきた。

高温超伝導のメカニズムや量子物質に創発する量子もつれの素性を解明する上で、中性子/X線散乱スペクトルや光電子分光スペクトルに現れる量子もつれの兆候の探索が始まっている。しかし、 10^3 電子系の基底状態計算が可能となり、絶対零度における高温超伝導を定量評価ができるようになった現在の「富岳」をもってしても、高温超伝導体における分光学スペクトルと数値シミュレーションの定量比較は未だ

に困難である。そこで、本節では 2032 年ごろに利用可能と予想される計算機によって解決すべき問題として、銅酸化物高温超伝導体の分光学スペクトルにおける量子もつれの解析に必要となる計算資源について詳述する。

【アルゴリズムの説明】 実験で観測される中性子散乱スペクトルとの定量比較を行うため、変分波動関数法に基づく励起状態計算によって、 10^3 電子系の分光学スペクトルを計算する。分光学スペクトルを計算するためには、多数の励起状態を計算しなければならない。これを実行する上で、乗り越えなければならない数理的困難の一つは、一見単純な波動関数の線形結合をどう実装するかにある。先進的な量子多体系計算手法の多くは、変分波動関数を基礎として定式化されている。これらの変分波動関数に共通する困難は、変分波動関数の和 ($\phi_1 + \phi_2$) を、別の変分波動関数 ϕ_3 でどう表すか、ということにある。特殊な場合を除けば、mVMC や DMRG、tenso network で表現された波動関数の和を同種の変分波動関数で書くことは難しい。

この問題を解決するために用いられてきた方法の一つのが、原子核理論で研究されてきた生成座標理論や物性物理学で開発されてきた動的変分モンテカルロ法で採用された、仮想的に線形結合を実現する手法である。非直交基底系 $\{\phi_m\}$ を用いて、一般化固有値問題を構成し、これを解くことで、実質的に基底系の重ね合わせ状態および直交基底系を得る。

ここでは、変分波動関数法の中でも、高次元系や 10^3 電子系での基底状態計算で高精度を發揮してきた mVMC に基づく分光学スペクトル計算を対象に計算の規模と要求性能を見積もる。計算のボトルネックとなる演算が共通する現行の mVMC で得られた実測値とアルゴリズムからの推定を組み合わせて、Pfaffian の計算が計算時間の主な部分を占めると考えられる。また、本手法では、LOBCG 法によって一般化固有値問題の $m \simeq (100)$ 個の低エネルギー解を求める。

【問題規模】 $N_s = 32 \times 32 (=1024)$ 格子点の銅酸化物の有効ハミルトニアン 4 種類について、中性子散乱スペクトルの計算を 5 計算条件ずつ行う。非直交基底数は、既存の計算で用いられた式より $M = (2N_{\text{ET}}N_s^2 + 1) = 18874369$ ($N_{\text{ET}} = 9$) となる。各プロセスごとのモンテカルロサンプル数 $N_{\text{MCS}} = 100$ 、 $N_p = 10^5$ プロセスの計算を想定する。

【計算時間】 ノードあたり 100GB 程度のメモリ量を持つ「富岳」の 10 倍程度のノード数を有する計算機を想定する。本計算と類似の mVMC を用いた計算では、ノードあたり 100GFLOPS 以上が実測されているため、総計算時間は、メモリ確保のため 1 ノード 1 プロセス、 $20 \times N_p$ プロセスを全系で同時に走らせた場合、 $39600 \times 10^{18} \text{ FLOP} / 2 \times 10^6 \times 10^{11} \text{ FLOPS} = 55$ 時間となる。

【演算量】 1 計算条件あたりの演算量は $c \times N_s \times M \times N_{\text{MCS}} \times N_p \simeq 1980 \text{ EFLOP}$ 、総演算量は 4×5 ケースを合わせて 39600 EFLOP となる。これは Pfaffian 計算にかかる計算量であり、LOBCG の演算量は無視できる。ただし、比例係数 c は A64FX で実測された FLOPS と計算時間より $c \sim 4$ と見積もっている。

【メモリ量】 実用的なアルゴリズムでは、一般化固有値問題に現れる $M \times M$ の行列のランクが、プロセスごとのモンテカルロサンプル数となる低ランク性を用いることで、 M に比例する量の倍精度実数をメモリ上に格納して演算を行う。そこで、プロセス毎に $16 \times (3m \times M + N_{\text{MCS}} \times M) \simeq 120.8 \text{ GB}$ の倍精度実数をメモリ上に格納する必要がある。したがって、計算条件ごとの総メモリ量は、12.8PB となる。

【メモリ転送量】 メモリの転送量は、A64FX 上での類似の mVMC 計算から見積もられた FLOPS あたり 1.6B から、計算条件ごとにおよそ 3170EB となる。

- 136GFLOPS/node (ピーク性能比 6%、ピーク性能 3.3792TFLOPS) • メモリー要求: 2BG/node • B/F 値: 1.6 • 実測メモリスループット 215.7GB/s (ピーク性能比 21.1%) • ノード間通信: Reduction
 - 通信量: $1000 \times 900 \times 42 \times 16 = 0.56\text{GB}$ • IO 量: 0.2GB
- オンチップメモリ ボトルネックとなる計算 Pfaffian CG 法

参考文献

- [1] 著者 1, 著者 2, 著者 2. 文献名. 出版社, 2022.
- [2] 日本学術会議. ロードマップ. <http://xxx.yyy.zzz>, 2022.

4.3 エネルギー・材料

本節ではエネルギー・材料分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.3.1 現在行われている課題

(i) 大規模分子シミュレーションによるエネルギー・材料課題

近年、エネルギー・環境問題が地球規模で深刻化している。石油や石炭等の化石燃料を現状のペースで消費していくと仮定した場合、数十年後には枯渇してしまいエネルギー不足に陥ると危惧されている。また、化石燃料の大量消費によって、二酸化炭素が大量排出され、地球温暖化や大気汚染等の環境問題も引き起こされている。これらの問題を解決するべく、化石燃料に頼らない代替エネルギーの開発が世界中の大学・研究機関・企業で盛んに行われている。太陽光発電はその代替エネルギーの有力候補であり、二酸化炭素を排出しないクリーンな発電を可能にする。太陽光発電にはいくつかの方法があるが、その中でも有機分子化合物から構成される有機薄膜太陽電池が注目を集めている。この太陽電池は大部分が有機分子から構成されており、基板にその有機分子を塗布するだけで製造できるため、低環境負荷、低コストで大面积の薄膜太陽電池を作ることができる。さらに有機 EL デバイスのように軽く薄い、またフレキシブルな形状を持つという特徴があるので、様々な場所での自由な設置が可能となる。このような有機分子化合物を基礎とした新しい太陽光発電装置の開発は、その素材となる分子材料の設計が肝要である。優れた化学的特性を持つ分子を実験的に合成し、デバイスとして設計することは基本のプロセスであるが、高性能かつ高品質な有機薄膜太陽電池を実現するためには、候補となり得る分子の構造や特性を、分子スケールで詳細に調べることが必要不可欠となる。そのため実験的な測定のみならず、分子科学の基礎理論に基づく量子化学計算も重要なツールとなる。

量子化学計算は実験データや経験的なパラメータを殆ど用いずに、量子力学の基本方程式に基づいて分子や原子の電子状態を求め、その電子状態から物性値や化学反応を予測する手法である。その量子化学計算の基礎理論において、最も発展し現在も主流となっている理論といえば、1997 年にノーベル化学賞を受賞した Walter Kohn 達によって提唱された密度汎関数理論 (DFT) である。この DFT は、それまで主流であった Hartree–Fock (HF) 法と同等の計算量で、HF 法では計算が困難であった比較的大きな分子に対しても、一定の信頼精度で安価に計算することを可能にしてきた。そのため、現在、最も有用性のある量子化学計算理論としてその地位を保っている。

様々な分子サイズの物性値の予測で多くの成功を収めてきた DFT ではあるが、励起状態の記述が良くないことが知られている。特に量子化学計算によって求められる発光吸収スペクトルは、分子の励起状態に関連した物性値であり、その正確な記述は、太陽電池や光触媒における分子材料の設計に対して極めて重要である。DFT における励起状態の記述を改善する手法として、時間依存密度汎関数理論 (TDDFT) が提唱されている。TDDFT も分子の励起状態を正確かつ高速に記述出来る手法として、DFT 同様に現在幅広く利用されている。TDDFT の他に、励起状態の正確な記述に関しては、ポスト HF 法も有力な理論であり、実際に TDDFT よりも高精度な結果をもたらすことが多い。しかしながら、計算量は TDDFT

よりもずっと大きく、数十～数百原子数の分子サイズにしか適応出来ていないのが現状である。そのため幅広いサイズの分子系の光学特性を、適度な精度で見積もる手法としては TDDFT が最有力の手法であると言える。

大規模分子系のための電子状態計算法の開発も望まれる技術の一つである。例えば、生体で起きる現象や反応を理解する手法を確立することは、創薬などに活かされるだけでなく、新たな反応系・材料の開拓にとっても重要である。また、試験管で起こる化学反応であっても、現実の系では、微量成分や pH などの条件で反応は大きく変化するため、これらの影響を取り込んだ計算が望まれるケースが多い。量子化学計算手法におけるアルゴリズムの高速化も盛んに行われており、近年の計算機の進歩と並列化技術の向上も相まって、DFT や TDDFT に基づく量子化学計算の適応できる分子系のサイズは飛躍的に拡大している。また、「富岳」に代表される超並列環境を効率良く利用したハイパフォーマンスコンピューティングによって、不可能とされたウイルスや DNA、たんぱく質等の数万原子数以上に及ぶ大規模な分子系についても、量子化学計算が可能になると期待されている。しかしながら、 N 原子系の通常の電子状態計算に要する計算時間は、DFT では N^3 に、電子相関理論では N^5 よりも高い幕で増大する。そのため、たとえ計算機の性能が大幅に向上升しても、その分だけ大きなサイズの系を扱うことはできない。この問題を解決する方法として、計算時間が N に比例するような手法、すなわちオーダー N 法の開発が進められてきた。我が国ではフラグメント MO 法、分割統治法、エロンゲーション法といったフラグメント分割型計算法の他、密度行列最小化法などの先進的なオーダー N 法の開発が進められている。

(ii) 高精度分子シミュレーションによるエネルギー・材料課題

化学の世界は、Schrödinger 方程式や Dirac 方程式をはじめとする量子力学の原理に支配されており、これを正確に解くことができれば、物質科学の現象を正確に予測することができる。しかしながら、Schrödinger 方程式や Dirac 方程式を多電子系である原子・分子に対して厳密に解くことは長い間不可能であると考えられていて、何らかの近似を導入することで従来は解かれてきた。例えば、前述した DFT 法や Møller-Plesset 法に代表される摂動法、配置間相互作用 (CI) 法、coupled cluster 法は、計算の精度と信頼性を高める電子相関法として確立された量子化学計算法となっている。

多電子系の波動方程式を厳密に解くことは不可能とされていたが、最近、原子・分子の Schrödinger 方程式を正確に解く理論として、自由完員関数法 (Free Complement (FC) Theory) が提案された。量子化学分野では、FC 理論を実用的な予言的量子化学にまで発展させるプロジェクトを遂行しており、方法論・計算アルゴリズムと共にその実用化プログラム「プリンシピア (Principia)」の開発が進められている。電子間の強いクーロン相互作用と系の内部自由度より、特異な集団的量子効果を示す強相関電子系は、新しい物性の宝庫であり、基礎研究からその応用について国内外で精力的に研究が行われている。また強相関電子系は、高温超電導体など次世代のエネルギー・材料分野に対して非常に重要な分野の一つであると考えられる。強相関電子系の理論的研究は、バンド描像を超えた取り扱いが必要であり、厳密対角化法や量子モンテカルロ法、また密度行列繰り込み群法等による数値的研究が重要な役割を果たしている。しかしながら、強相関電子系の持つ電荷や спин、格子等による巨大な内部自由度は、その数値的取り扱いにおいて巨大な計算資源を要求する。また、強相関電子系の応用の立場からもこのような物質の動的性質や温度、外場等の様々な効果を取り入れた計算が必要となるため巨大な計算資源が必要になることが考えられる。

(iii) ダイナミクスシミュレーションによるエネルギー・材料課題

天然資源の乏しい日本ではエネルギー関連の技術革新は将来にわたる継続的な課題である。また工業の基盤となる材料について、他国では真似できないようなミクロレベルからの設計指針を開発・保持することは将来にわたる国際競争力を維持する上で重要である。

エネルギー問題について、廃棄物系・未利用系バイオマス利用技術や、日本近海に眠る唯一とも言える国産天然資源であるメタンハイドレートの利用技術の進歩のためには、熱力学過程や酵素反応過程の分子レベルからの解明が不可欠である。計算機能力の向上により、実験に並ぶ方法として計算科学的手法が重要な研究手段となっているが、現状では計算機上で基礎過程を再現することで精一杯である。今後5~10年の計算機の発展により、基礎過程の複合した現象、および周囲の環境との相互作用をフィードバックした形でのより現実的な計算機シミュレーションを行うことが可能になると期待される。例えば、周囲へのメタンガスの拡散過程を含めたハイドレートの融解過程の分子シミュレーションにより掘削時およびプラントでのガス化プロセスに関係するハイドレート気化過程を制御するための分子レベルの知見が得られ、より効率的なエネルギー利用が可能になると期待される。

また、燃料電池開発等では効率化のために電極界面での分子構造の解明が不可欠である。しかしながら電気化学で問題となる空間スケールは、原子、分子レベルのオングストロームオーダーからサブミクロンオーダーまで多岐にわたり、分子シミュレーションのみ、あるいは連続体理論のみで現象を説明することは大変難しい。これら両者の方法の溝を埋める手段として、大規模計算に期待が集まっている。

分子の自己組織化により形成されるナノスケールの分子集合体（ミセル・脂質二重層膜・タンパク質複合体・高分子）は総じて「ソフトマテリアル」と呼ばれる。これらソフトマテリアルは、高分子分離膜・食品・化粧品・生体模倣材料・ドラッグデリバリーシステムのキャリア・化学工学プラント設計・ガス分離・温暖化ガスの吸収・電池電解液・結晶成長といった広範な学術的・社会的ニーズに直結している。ソフトマテリアルの機能は分子集合体の微細構造やダイナミクスに応じて変わり、分子の種類や組み合わせを変えることで、機能を自在に制御・設計することができる。しかしながらそのスケールの微小性から実験からでは構造・ダイナミクスの詳細が不明なことが多く、また機能の発現には連続体モデルの適用が難しい界面が関係するため、ソフトマテリアルに対する原子分子レベルでの大規模シミュレーションを行うことでその詳細を解明、さらにはその知見を活かした新機能ソフトマテリアルを分子レベルから設計する技術の確立が期待できる。特に、生活に浸透した高分子材料について計算機能力の進展により革新的な分子設計が可能になると期待されている。高分子特有の粘弾性挙動を扱うために、現状高分子のシミュレーションは粗視化モデルをベースとした物質に依らない普遍的な物理現象を扱ったものが主流である。しかしながら、実用上重要な高分子材料の電気的性質・化学的性質を計算機上で再現するには原子分子レベルでの知見を直接ないしボトムアップ的に計算モデルに取り込む必要がある。

ソフトマテリアルについての革新的な分子設計を行うためには、階層化シミュレーション方法論の開発が必要である。さらに、燃料電池開発等で重要な電極界面での分子構造の解明においても、階層化シミュレーション法の開発は不可欠になる。原子分子の特性を反映した大規模かつ長時間のシミュレーションおよび数多くの試行が必要であり、計算機能力の向上が継続的に求められている。

(iv) 液体シミュレーションによるエネルギー・材料課題

液体の統計力学理論の分野では日本、カナダ、米国、欧州に積極的に開発を行っている研究グループがあるが、我が国はこの中でも多くの活発な研究グループを抱え、世界をリードしているといえる。とくに生体分子への応用を志向した手法(3D-RISM、MC-MOZ、XMOZ)などの開発・発展・応用において顕著な業績をあげている。

(v) 量子化学ソフトウェアの開発

化学結合エネルギー、電気エネルギー、太陽光、熱エネルギーの間の相互変換における物質の機能の役割を明確化し、エネルギー変換の大幅な高効率化につなげるための計算手法の開発、コード開発、シミュレーションが進められている。燃料電池やリチウムイオン二次電池における電気化学過程、有機系太陽電池の界面構造などの研究には、主として大規模な第一原理分子動力学計算が用いられている。また、水素・メタンハイドレートの熱力学過程の解明には古典分子動力学法などが、バイオマス利用における酵素反応の解明には有効媒質法(3D-RISM法)などが有効である。わが国においては、いずれのシミュレーション手法についても大規模並列環境への最適化・高度化が進んでおり、理論的な蓄積と合わせて国際的にも優位な立場にある。

量子化学計算については、ADF(蘭)・GAUSSIAN(米)・GAMESS(米)・NWChem(米)・MOLCAS(典)・MOLPRO(独)・Orca(独)・Q-Chem(米)・Turbomole(独)など欧米各国では一般ユーザーにも使いやすい汎用コードの開発が進んでいる。日本においても ABINIT-MP・GELLAN・NTChem・OPEN-FMO・PAICS・Protein-DF・SMASHなどが開発されている。更に、DFTで水クラスターのエネルギー計算を行い、大規模系の超高精度計算が実現可能になることを示した。

4.3.2 長期的目標と次世代に解決すべき課題

(i) 大規模分子シミュレーションによる解決目標

高効率の太陽電池や高効率熱電素子など、地球規模のエネルギー問題解決に向けた新しい概念に基づくデバイスも、物質科学の基礎研究に基づいて提案され始めている。燃料電池に用いられる白金触媒、色素増感太陽電池に用いられるルテニウム、透明電極に使われるインジウム、リチウムイオン電池材料のリチウムやコバルトなど供給量が希少、あるいは今後の需要増に応じて希少になると考えられる元素の代替材料を他国に先駆けて開発することが、わが国の産業競争力を高めるためにも必要である。計算科学と実験・理論がタイアップし、次世代の半導体デバイス、触媒材料、各種電池、薬剤、触媒などの設計・開発に役立てることで、地球環境を守り、産業振興を助け、社会を豊かにすることに繋げることができる。

(1) 分子機能と物質変換

物質変換技術は、基盤産業の基礎であるとともに、人類の生活を豊かにした科学技術という点で最も成功した技術であると言える。分子と化学反応の微視的理解に基づいて、さまざまなクロスカップリング反応のような精密有機合成反応、バイナップなどの使用による高立体選択的合成反応、重合反応や脱硝、脱硫反応など、新しい考え方による物質変換反応が次々に開発されてきた。衣類、家や車の構造物、更には液晶や有機ELのような電子材料、航空機の構造材料など、化学反応による物質変換の成果は枚挙にいと

まがない。今後のわが国の産業競争力を更に強化し、地球規模での資源問題を解決するためには、目的の物質を安価で汎用的な材料から効率よく作り上げることが不可欠である。複雑な構造を持つ新しい有機分子触媒、金属微粒子、金属（酸化物）表面、金属錯体などをまるごと精密に計算することにより、分子の相互作用と化学反応過程を解明し、新しい物質変換法を理論先導で開発することを目指す。

(2) エネルギー変換

エネルギー問題はわが国の最重要課題の一つであり、エネルギー資源小国であるわが国が世界の中でこれまで以上の繁栄を目指すためには、あらゆる科学技術を駆使してエネルギー創成・変換・利用技術を他国の追随を許さないレベルにまで高める必要がある。そのためには、世界最高水準のわが国のエネルギー関連技術を科学にまで高め、試行錯誤というレベルとは異なる次元から物質開発を行うための道筋をつけることが必須である。燃料電池や二次電池、太陽電池、あるいはバイオマス利用における電気化学過程、非平衡電気伝導、貴金属を用いない触媒反応などに対し、超並列大規模計算機を利用した第一原理に立脚したシミュレーション技術を、かつてない高いレベルに押し上げる。これにより、物質とエネルギーの関連を理論とシミュレーションから突き詰めて考え、次世代の高効率エネルギーおよび物質変換に寄与する材料開発に資する知見を獲得し、その成果を社会全体へと還元する。

また、数万～数十万原子規模の量子化学計算が可能となれば、幅広いサイズの分子材料の物性値を予測できるようになるため、理論的な分子材料設計の多様性が増大する。従って、大規模量子化学計算を実現し、長寿命でエネルギー変換効率の高い有機系太陽電池の素材となり得る分子を理論的に探索することは、昨今のエネルギー・環境問題の解決への道を切り開く可能性を与える。そのためには、効率の良い理論的予測が望まれるため、現在の「富岳」の性能を上回る富岳 NEXT を使った超並列計算は必要不可欠なものである。

(ii) 高精度分子シミュレーションによる解決目標

高精度分子軌道法では、複雑な電子状態をもつ系に対して構造や物性値を分光学的精度で計算すること、およびより大きな分子に適用することが恒久的な課題となっている。特に、生体内の複核金属酵素のように複数の金属原子と巨大な有機配位子が複雑に相互作用をした系や、ナノサイズの半導体光触媒のように光励起状態で反応が進行する系の高精度な計算には、巨大な分子系で、いわゆる動的電子相関と静的電子相関の両方を正確にとりこむ必要があり、方法論とアルゴリズムさらに計算機資源の発展が必要となっている。マンガン、鉄、コバルト、ニッケル、銅元素などの第一遷移周期元素は、比較的安価に入手するために触媒などの様々な応用分野での活用が期待されている一方で、これらの金属原子は電子相関が強く、構造や物性値を正確に算出するためには高精度な分子軌道法を適用する必要があることが知られている。また、有機分子からのみなる光電変換材料や光触媒は、分子設計の自由度がほとんど無限に近いために、高精度な分子軌道法の適用による物性値予測が大きく貢献できる対象であると言える。これらのことより、大規模計算機をもちいて電子相関を考慮した高精度な分子軌道法を適用することは、特に複雑な電子状態をもつ新規触媒や光が関与する機能性材料の設計に貢献するものである。

強相関電子系のエネルギー・材料分野への応用としては、銅酸化物超伝導体等の比較的高い超伝導転移温度を持つ高温超伝導体や、固体中の電子の持つ電荷の自由度のみではなくスピントン自由度も利用したスピントロニクス素子、複数の強的性質が絡み合ったマルチフェロイック材料など枚挙に暇がなく、エネル

ギーのライフサイクルの各ステージにおいて様々な応用が期待される。このような強相関電子系の理論的研究においては、電子間の強い相関を精密に取り扱う必要があるため、平均場近似に基づいた解析的理論による取り扱いが困難であり、数値的手法による取り扱いが重要な役割を持つ分野である。したがって、数値的研究による強相関電子系の理解とその応用研究の推進は、エネルギー・材料分野への応用、および当該分野における日本の国際競争力確保の立場からも非常に重要であると考えられる。

物質科学の現象を正確に予測するためには、あらゆる物質の変換原理を統べている量子力学的原理の深い理解と、それを使いこなす理論と方法、同時にその具体的な計算を短時間で遂行できる計算機システムの存在、が揃わなければ不可能である。そのため、この基礎方程式である Schrödinger 方程式と Dirac 方程式を化学精度で解くことのできる理論と、それに基づくシミュレーション技術を中心を持つことが絶対に必要である。当該分野の目標は、これらの基盤技術を整え、物質科学の計算シミュレーション分野に高度な信頼性と予言性を提供し、実験と並ぶ有用な研究手段として発展させることである。化学精度を満足する正確な Schrödinger 方程式解や化学反応ポテンシャルが記述できれば、計算主導で有用な物質設計や物性解明を行うことができる。例えば、溶液反応は、溶媒との弱い相互作用により気相反応とは反応機構も全く異なり、あらゆる相互作用を正確に記述でき正確な反応ポテンシャルを与える当該理論でしか研究できない。長期目標と社会貢献に関しては、これらの特徴を生かし、薬剤設計や機能性分子設計、気相と溶液中での反応設計などにより、実験や産業を先導するような高信頼性シミュレーションを実現することが期待されている。

(iii) 励起状態シミュレーションによる解決目標

光励起により電子移動の起こることは光化学における大きな関心事であり、高精度分光技術、量子化学理論の発達により、小さい規模に限定すると、高精度な知見が得られるようになっている。これは、計算機能力の発達に依るところが大きい。化学機能発現にとり現実的なサイズにより近づくには常にコストと精度バランスのとれた手法の開発と大規模応用に耐えうる計算機の開発、双方の発展と努力が望まれる。

電子状態の時間変化は電荷輸送特性に本質的に重要であり、一般に時間静的な計算に比べコストがかかる。また、より定量的な予測には、電子相関の考慮が不可欠であるが、これについてもやはり一般に多くの計算コストがかかる。関与する励起状態数が多くなり、また、構造が複雑になると、この事情は顕著になる。加えて、太陽光デバイス開発における電子エネルギー動力学の研究に必要とされる大規模な構造変化を伴う励起状態を記述するためには、単純なモデルでの計算は信頼性に欠ける。そこで、できる限り高精度電子状態理論に基づく微視的な機構の理解は非常に重要となり、実験との共同が部分的に必要になるが、これに応じて必要な計算規模も柔軟に設定する必要がある。同時に、定性的理解に向けた簡易模型開発の努力も必要である。

短時間過程については光電子分光技術の発達が近年著しい。様々な波長の短パルス光の複合利用やイオン-光電子相関観測等の技法を駆使して、エネルギー移動の解明研究が光デバイスに対して適用され、これまでに理論の側だけで予測されていたことの検証が可能になりつつある。これを受けて、計算科学側としても、より大規模で信頼性のある理論模型に基づく研究が必要となる。特に、始めと最後はわかっている実験研究に対し、その間の知見を得られる理論計算には大きなアドバンテージがある。若い計算物質科学研究者の興味を引き出す効果もあると予想される。国内外を問わず物質科学分野の維持発展にとり大事であり、計算科学はこれに貢献できるはずである。そのための計算機科学分野の発展は極めて重要である。

計算科学側からの太陽光エネルギー活用に向けたデバイス開発においては、光吸収と励起電子エネルギー輸送が電子外場環境としての分子構造運動にどのように左右されるかを計算機で解明することが求められる。特に、環境変動や電子的フラストレーションに伴う電子揺らぎ、電荷キャリアのトラップ、有機分子から成る集合系における電荷分離可能性、円錐交差による電子失活の抑制、揺らぎを逆利用した輸送特性の向上等の相関は、大規模計算科学を利用した研究ではじめて得られるものと考えられる。

複雑な励起状態の関わる光デバイス、反応系においては、原理的には時間に依存する電子の波動関数を追跡する電子動力学計算を行うことで、より微視的な情報を得ることができる。そのためにはやはり大型計算の重要性は疑いようがない。一方で大型の分子集合系における電子動力学の粗視化研究、基礎定式化研究も必要である。現時点での計算機レベルでは不可能な対象にも事前に粗くとも様々にあたりをつけておくことは、先端研究がしばしば陥りがちな分野の先細りを防ぎ、研究の多様性を維持する上で重要である。このことを踏まえて、大規模で柔軟性のある計算機が開発されることを望む。

(iv) 化学反応シミュレーションによる解決目標

人類は現在、環境問題・エネルギー問題・食糧問題など、多くの問題に直面している。古くは、ハーバーとボッシュがアンモニア合成触媒を開発して世界の人口問題を救ったように、目的反応を効率的に起こす触媒の開発は、これらの問題を解決するためにますます重要となっている。触媒開発による解決が望まれている環境問題としては、まず自動車排ガス浄化触媒の開発があり、現在実験と計算の研究者が一体となって研究が進められている。また、温暖化ガスである二酸化炭素の分離・回収技術の開発が進んでおり、火力発電所などの大型排出施設への設置が実現しつつあるが、回収した二酸化炭素を貯蔵するだけでなく、炭素源として利用する触媒の開発が期待されている。PM2.5に代表される粒子状物質の回収も、触媒開発による解決が望まれている。食糧問題では、リン資源の枯渇が重要な問題として認識され始めている。我が国は、肥料に利用されるリン酸を100%外国からの輸入に依存しているが、リンの希少価値が高まると輸入にかかるコストの大幅な増加が予測される。そこで、現在あるリン資源を回収し、有効利用する触媒の開発が望まれている。このように、新規触媒開発に集まる期待は大きいものの、その反応過程は非常に複雑であるため、理論・計算主導による設計はこれまで困難であった。しかし、近年、反応経路を自動的に探索する方法の開発が進み、複雑な触媒反応経路を人間の直感を極力排除して探り当てることが可能となってきた。

また、現実系のシミュレーションを行う手法として、対数スケールで効く条件まで考慮した希薄多成分溶液系の反応分子動力学計算法の開発が望まれる。このような反応系に対しては、オーダーの異なる原子数のシミュレーションをオーダーの異なる時間スケールで追うことが必要となっている。

(v) ダイナミクスシミュレーションによる解決目標

燃料電池や二次電池、太陽電池、あるいはバイオマス利用における電気化学過程、非平衡電気伝導などに対し、超並列大規模計算機を利用した第一原理に立脚したシミュレーション技術を、かつてない高いレベルに押し上げる。また、電気化学に関連して、新規溶媒として期待されるイオン液体といった新しい材料開発においても分子モデリングを含めた高いシミュレーション技術が求められており、計算科学的方法論を実験との相補的な手段としても確立する。

一方、界面活性剤、脂質、タンパク質、高分子などの多官能性の分子は、温度や塩・共溶媒濃度のよう

な外部パラメータによって、ミセル・膜・液晶・ナノカプセルといった多様なソフト構造体に自己組織化し、分子の認識、分配、分離、輸送の機能を発揮する。これらの機能は、環境の熱エネルギー程度の弱い相互作用の制御によって設計することができ、ミセル・ウィルス外殻（カプシド）を利用したドラッグデリバリーシステム、高分子分離膜（海水淡水化等）、食品・コスメティック、生体模倣材料、化学工学プラント設計、ガス分離、温暖化ガスの吸収、電池電解液、結晶成長といった広範な学術的・社会的ニーズに柔軟に対応した新規ソフトマテリアル設計が可能になる。

また、タンパク質分子に加えて脂質も結合した巨大複合体である HIV やインフルエンザウイルスなど病原性の高いウイルスについての大規模分子シミュレーションは、細胞との結合・解離といった機能の分子レベルでの解明とともに、ウイルス性疾患の予防・治療薬開発への展開も期待される。このようなタンパク質複合体の構造形成においては、溶媒や塩の効果もきわめて重要であること、更にはタンパク質と脂質膜との相互作用について未解決であることから、全原子の分子動力学シミュレーションによる生体機能の分子レベルでの解明が求められている。

これらにより、物質とエネルギーの関連を理論とシミュレーションから突き詰めて考え、次世代の高効率エネルギー変換に寄与する物質・材料開発に資する知見を獲得し、その成果を社会全体へと還元する。

4.3.3 課題を解決するために必要なアプリケーション群（要求性能）

(i) 大規模分子シミュレーションのためのアプリケーション群

必要な物質を汎用元素から効率よく合成する物質変換法の開発は、わが国の産業基盤の強化に止まらず、地球規模での持続可能な社会を構築するために不可欠である。このためには複雑な化学反応の微視的理解と予測・制御が必要であり、自由エネルギー面での理論・計算化学研究の遂行が不可欠である。研究対象としては、複雑な構造を持つ新しい有機分子触媒、金属微粒子、担持触媒、金属表面、金属酸化物表面、金属錯体などが挙げられ、1万から数十万原子系の電子状態計算、さらに周囲の溶媒などについては数百万原子を含めた実際の物質変換系について、主要部分を量子化学計算し、周囲を分子力場計算する QM/MM-MD 計算を実行するといったアプローチがある。

大規模量子化学計算を行うために使用するアプリケーションとして、国産の量子化学計算ソフトウェアの NTChem や SMASH の利用も可能である。これらの課題の解決のためには、現在の「富岳」以上の超並列環境において高効率に実行できるようなアルゴリズムの構築と、高度な並列プログラミングおよびチューニング、そしてメモリの分割が必要である。Hartree-Fock 法および密度汎関数法計算において最も計算時間要するのは二電子積分の見積りであり、この部分を高速化することが最も重要となる。二電子積分を高速化するための手法の一つとして、Pseudospectral (PS) 法がある。PS 法は二電子積分の一部を、解析積分から実空間グリッドによる数値積分に置き換えることによって、ある程度の計算精度を保ちつつ高速化を図る手法である。形式的に二電子積分は、 $O(N^4)$ (N : 基底関数の数) の計算量となっているが、PS 法を適用することによって、 $O(MN^2)$ (M : グリッド点の総数) まで計算量を下げることができる。また PS 法には、CPU コア数に対する並列化効率が非常に良い手法となっており、大規模な分子系であっても比較的高い並列性能を維持できる利点がある。従って、京のような大規模超並列環境に適した手法となっているため、PS 法による DFT 計算、即ち PS-DFT 計算が課題解決の計算手法になり得ると予想する。また計算する分子系のサイズが大きくなればなるほど、必要なメモリサイズは増大す

るため、各並列プロセスに対してのメモリ分割の処理も必要になる。

今後の大規模量子化学計算において必要となる計算機環境の1つとして、通信ミドルウェアの高性能化ならびに充実が挙げられる。現在、ほとんどの大規模量子化学計算プログラムでは、MPI通信ライブラリが利用されている。これは、並列プログラミングのモデルが現在の大規模分散メモリ型計算機環境に良く適合していると考えられるためである。しかしながら、例えば10万基底を超える大規模分子を対象にSCF計算をする際には、計算の中間データとして必須のFock行列のサイズが30GBを超えると試算される。今後の超並列コンピュータシステムの1プロセスあたりの使用可能メモリ量はハードウェア消費電力バジェットの分配の問題から数GB程度と予想されているため、1ノードのオンメモリにてSCFに必要な全行列要素を保存し対角化する現在のアルゴリズムが利用出来なくなると予想される。さらに、データを多ノードのメモリ上に保存した場合にも、計算中に他ノードのデータアクセスが必要となる場合には、そのノードのCPU処理を中断させないよう、RDMAの機構を利用する必要があり、現在多く利用されているメッセージパッシング型の通信ではなく、片側通信を利用することとなる。このため、RDMAをサポートするPGASやCoarray等の通信の仕組みが必要となり、さらにこの多ノード上に分割保存されたデータに対する対角化や連立方程式計算を効率よく行うミドルウェアが必要不可欠である。NWChemプログラムでは、多ノードにまたがる大規模なデータを保存可能なGlobal Arrays通信ライブラリを利用しておらず、また、このライブラリには対角化等の計算も同梱されているため、利用上便利な仕様となっている。また、MPI+PGASライブラリもMVAPICHグループにより開発されている。

しかしながら、現状のMPI実装のままではポストエクサ環境において省メモリかつ高性能な実装とすることは難しいことが知られており、Global Arraysは内部でMPIの初期化関数を利用していることもあり、今後省メモリ実装が必要不可欠となっている。これに対し、我が国で開発中のACP通信ライブラリはInfiniBandやTofu、Tofu2で利用可能なRDMA通信ライブラリである。MPIとの併用も可能であり、エクサ環境においてもメモリ使用量が1プロセス当たり数百MB程度となるように実装されており、有望であると考えられる。しかしながら、こちらは対角化等の計算アルゴリズムが未開発であり、今後開発が望まれる。

また、FMOなどの計算では対角化の遅延が積分計算の遅延と同等になるまで並列化されており、今後さらに大規模化するためには、この対角化遅延の隠蔽のため、モノマーの分子積分計算と対角化ルーチンを分離して、積分計算プロセスと対角化計算プロセスを独立に動作させる試みが必須となっている。この計算のためには、MPIの様なSPMD型計算モデルでは記述が困難である。MPI-Spawnの仕組みを利用することも考えられるが、ライブラリそのものではなく、複数のタスクを非同期的に利用可能なよりユーザーから便利に利用可能なフレームワークとしての通信ミドルウェアライブラリが必要となっている。

付録A1に大規模量子化学計算に向けた並列実装に対するNTChemでの取組みについて記載する。

(ii) 高精度分子シミュレーションのためのアプリケーション群

分子軌道法は、基本となるHartree-Fock(HF)法を出発点とし、摂動法、結合クラスター法、配置間相互作用法などのpost HF法により電子相関を取り込むことで、計算精度を系統的に引き上げられるという特徴がある。動的電子相関を考慮した高精度分子軌道法とは、一般にこれらの単参照のpost HF法を指すことがほとんどである。上記のpost HF法は、複数の電子配置から出発することで静的電子相関を考慮した多参照分子軌道法へと拡張することができ、参照状態にすべての重要な電子配置をとりこむこと

ができれば、分光学的精度を達成することができる。また、すべての電子配置を考慮した完全配置間相互作用法が適用できれば、与えた基底関数の範囲内では厳密な結果を得ることができるために、完全配置間相互作用量子モンテカルロ法などの確率論に基づいた近似手法やモデル空間量子モンテカルロ法のような確率論と決定論的な手法を組み合わせた計算手法も開発されてはいるが、大規模計算機で高効率に実行でき、かつ巨大系でも適用できるようにするために、方法論やアルゴリズムのうえでの更なる発展が必要である。

これらの計算理論の軸とは別に、分子軌道法では基底関数を適切に選択することも正確な計算のために必要な要素として挙げられる。Post HF 法は基底関数サイズに対する収束が遅いために、完全基底極限に近い結果を得るために最低でも一原子あたり 100 個程度の基底関数を用意しなければならない。露わに相關した電子状態理論は、基底関数サイズに対する遅い収束性の問題を解決でき、一原子あたり 30 個程度の基底関数で完全基底極限とほぼ同等の結果を得ることができる。

長期目標と社会貢献の項でふれたような生体内の金属酵素や半導体は、複雑な電子状態をもつ系のために、高精度な post HF 法に露わに相關した電子状態理論を組み合わせた計算が必要である。そのような post HF 法では、行列行列積、行列対角化、連立一次方程式計算、そして基底関数の軌道角運動量によって演算内容が異なる 1 電子・2 電子積分計算が主な演算で、計算対象分子を分割しない限り、どの方法でも全対全通信を行う必要がある。電子相關計算では、中間データ量が系の 3 乗もしくは 4 乗に比例して増加するため、各ノードに分散させてもノード当たりの必要メモリ量は多くなる。ここでは、電子相關の計算で必要になる 4 中心 2 電子分子軌道積分を、3 中心 1 電子分子軌道積分と求積点での分子軌道の値の積から数値求積法により計算するアルゴリズムについて検討を行う。この方法は、ノード間に求積点を分散させることでプロセス間の負荷分散が容易であるという利点と、求積点に関する総和が行列行列積になるために高い単体性能が得られるという利点がある。多電子積分計算についても同様に数値求積法で行う。最も重い演算は、倍精度実数の行列行列積であるため、DGEMM の性能がアプリケーションのピーク性能になると想ってもよい。ただし、3 中心 1 電子分子軌道積分に関しては、必要なメモリサイズが巨大なものになる。また、各ノードに分散されている求積点ごとのデータを集めるため、通信は Allreduce、Gatherv が大半であり、これらの通信性能も重要である。この課題を解決するために、FS2020 プロジェクトで NTChem の RI-MP2 計算部に対してコデザイン高度化を実施した。詳細について付録 A2 に記載する。

大規模並列第一原理密度行列繰り込み群法のアプリケーションとして、理化学研究所計算科学研究中心において「paraDMRG」が開発されている。第一原理密度行列繰り込み群法は、物性物理分野において特に一次元強相関量子系の計算に用いられていた密度行列繰り込み群法を、量子化学分野における強相関電子系の取り扱いへと発展させた数値的手法である。強相関電子系の取り扱いでは、系の電子相關を精密に取り扱う必要があるため、個々の電子の自由度を顕わに取り扱う電子の量子多体問題を解く必要がある。そのため、強相関電子系の取り扱いにおける系の内部自由度は取り扱う軌道数に対して指数関数的に増大する。したがって、強相関電子系の取り扱いについて、従来行われてきた厳密対角化法では、計算規模の面から取り扱う軌道数に対して非常に強い制限がかかる。このような問題に対して、密度行列繰り込み群法による取り扱いでは厳密対角化法における軌道数の制限を超えた取り扱いが可能である。さらに厳密対角化法に匹敵する高い計算精度の結果を得ることが可能となる。このような密度行列繰り込み群法が従来の厳密対角化法の限界を超えた取り扱いが可能となる理由は、密度行列繰り込み群法が計算の目的

とする特定の状態に焦点を絞った計算を行っているためである。これは密度行列繰り込み群法が計算の目的とする状態を最適に表現する基底を、計算規模や計算資源等により決定される任意の数 m の制限の中で与えるためである。

また、原子・分子の Schrödinger 方程式と Dirac 方程式を正確に解く一般理論である FC 理論に基づき、これを実用化するために様々な理論・計算手法が提案されている。FC 理論が生成する正確な波動関数に収束する完員関数群は、積分計算が困難な関数も含むため、積分フリーな局所 Schrödinger 方程式理論 (Local Schrödinger Equation (LSE) 理論) が提案された。LSE 法は、サンプリング手法の 1 つであるため、大規模な並列計算に適している。また、大きな分子系で最も計算負荷の大きい反対称化則（パウリ原理）の計算については、行列式ベースの Nk 理論や、化学の局所性に基づき遠く不要な電子交換を自然に排除する inter Exchange (iExg) 理論が提案された。これらのことにより、Schrödinger 方程式の一般的な解法でありながら大規模系にも展開可能な理論体系が確立されつつあり、その実用化プログラム「プリンシピア」の開発が進められている。

(iii) 古典ダイナミクスシミュレーションのためのアプリケーション群

分子科学は、化学反応の理解とそれに基づく新しい分子・分子集合体の創製を通じて、物質科学研究に大きな展開をもたらしてきた。

ハイドレートから得られたメタンを始めとした各種ガス（水素、二酸化炭素など）の貯蔵および分離、蓄熱などに利用する技術の開発が継続的に行われている。そこで問題は過冷却状態において不均一核生成によりハイドレード結晶核を如何に効率よく発生させるかにある。ある種の水溶液気液界面ないし固液界面はハイドレート結晶核生成を促進させる効果があることが知られている。その分子論を大規模分子シミュレーションによって解明し、さらには結晶化を促進する界面の分子設計に応用することができれば、エネルギー貯蔵コストを劇的に減少させることができる。さらにハイドレート結晶粒界の効果、バルクへのガス放出といった因子を大規模シミュレーションによって取り込むことで、液化天然ガスに代わる新形態としてエネルギー輸送のブレークスルーに貢献する。

ソフトマテリアルを構成する界面活性剤や脂質・高分子などの多官能性の分子は、温度や塩・共溶媒濃度のような外部パラメータによって、ミセル・膜・液晶といった多様な構造体に自己組織化し、分子の認識、分配、分離、輸送の機能を発揮する。イオン液体や超臨界流体は、溶媒条件選択による物理的・化学的性質のチューニング幅が大きく、特異的な反応選択性や溶解性・電気特性を発現する。更に、水のような「通常」の分子集団も、制限空間やクラストレート内では、分子間相互作用を原子レベルで反映して、構造や動的性質を大きく変える。一般的な高分子材料も、分子配向を高度に制御することにより、従来にない高機能・高強度化を示す。上記は、規則性とランダム性を兼ね備えた原子・分子集団としての働きであり、ドラッグデリバリーシステム (DDS)、分離膜（海水淡水化等）、食品・コスメティック、生体模倣材料、ガス分離、温暖化ガスの吸収、電池電解液、結晶成長、機能性高分子材料といった広範な学術的・社会的ニーズに直結する。

ソフト分子集団の機能の解析・設計には、原子・分子レベルにおける相互作用の理解と大域的な集団形状の記述の両方が必要である。数桁に及ぶ空間・時間スケールの解析には、量子化学計算・全原子分子動力学シミュレーション・粗視化分子動力学シミュレーション・散逸粒子動力学法・連続体モデルを融合したマルチスケール手法の開発が必須である。特に、実地応用につなげるためには、熱エネルギー程度の分

子間相互作用の効果を精度よく取り扱い、化学的個性を取り入れる必要がある。量子化学計算と分子シミュレーションの大規模化に加え、自由エネルギーを中心とする統計力学理論やレアイベント探索アルゴリズムの高度化が必要である。

特に界面系では、凸凹があり不純物も存在する現実系の取り扱いが可能になりつつある。気液界面でのガス吸収、分子素子やナノテクノロジー・トライボロジーの新展開の場であり、環境科学との関連も深い。また界面系での電気化学の基礎的問題として、水界面は酸性であるか塩基性であるかの長い議論があるが、未だに統一的見解は得られていない。この原因は、おそらく「界面」と定義される領域の空間スケールが問題によって異なることが原因の1つと考えられる。今後、多成分系を含めた実在の界面の解明のためには、界面構造の動的なゆらぎや電気二重層構造のイオン分布を分子シミュレーションで扱ったうえでマクロな理論に繋げる必要がある。

(iv) 古典・量子混合ダイナミクスシミュレーションのためのアプリケーション群

この節では、光エネルギー吸収後の過程を解析する手法「非断熱分子動力学シミュレーション」について述べる。この手法の究極的な目標は、「分子が光を吸収すると何が起こるのか」との問い合わせに対して、反応時間や反応機構をリアルタイムの分子シミュレーションで明らかにすることである。分子が光を吸収すると、電子はエネルギー準位が高い位置へと移動し、電子励起状態が生成される。励起状態からの緩和過程として、蛍光や燐光のような発光だけでなく、光を放出しない無輻射遷移が考えられる。エネルギー差が小さな領域に複数の電子励起状態が存在するとき、後者は特に顕著であり、超高速な緩和過程に大きく寄与する。非断熱分子動力学シミュレーションでは、状態間の遷移を取り扱う必要があるが、非断熱カップリングから確率論的な手法で導入している。したがって、励起状態の緩和時間や緩和過程を評価するうえで非常に重要な方法である。

非断熱分子動力学シミュレーションを実現するためには、(1) 非断熱遷移の理論的な取り扱い、および(2) 励起状態を記述できる高精度かつ計算コストが小さな電子状態計算が課題となる。現在、surface hopping に代表される非断熱分子動力学シミュレーションのアルゴリズムの開発や改良はめざましく、(a) 古典力学では取り扱いが困難な (de)coherence や (b) スピン許容およびスピン禁制遷移の統一的な取り扱いといった点が挙げられる。しかしながら後者一大規模な分子系の電子励起状態の計算コストは依然として解決されていない。半経験的手法を除いて、時間依存密度汎関数法 (TDDFT) に限られているのが実情である。TDDFT 法は計算コストが小さく、並列化も容易であるため、非断熱分子動力学シミュレーションと組み合わせた手法が採用されつつある。解析的な非断熱カップリングの計算は、通常の分子動力学における力 (エネルギー微分) とほぼ同じコストであるが、すべての状態対に対して計算する必要がある。すなわち、計算コストはほぼ状態数の 2 乗に比例する。計算コストを劇的に削減するため、「波動関数」の重なりを利用したスカラー版の非断熱カップリングを導入されることが多い。

将来の目標としては、太陽電池の分子設計があげられる。有機薄膜タイプの太陽電池は、ドナーの有機ポリマー分子とアクセプターのフラーレン分子から構成される分子系である。ドナーあるいはアクセプター単体の計算も分子設計においては有益であり、さまざまな有機分子が候補として提案されている。その一方で、ドナー・アクセプター間の電子移動のように分子対として成立する物性の理解や、移動度といった動的過程の解析には、非断熱分子シミュレーションが欠かせない。先に述べた TDDFT 法による surface hopping を適用するうえで問題となるのは、やはり電子状態計算の高速化である。特に、TDDFT

法によるエネルギー計算が律速となることが予想される。求める電子状態の数に対応した行列固有値と固有ベクトルが必要となるが、ベクトルの次元は(占有電子数)×(基底関数)で与えられるため、複雑な分子集合体では容易に 10^{10} に達する。大規模な並列計算機においても、メモリ分割などデータ保持をうまく行うアルゴリズムだけでなく、ノード間のデータ通信などハード面での貢献も欠かせないであろう。

(v) 量子ダイナミクスシミュレーションのためのアプリケーション群

化学反応力学は、電子状態変化と構造変化に特徴付けられる。電子励起状態とのエネルギー差が分子の運動エネルギーに比べ十分に大きく電子基底状態で進行する熱化学過程においては、電子状態が位相を除き原子核配置に対し瞬時に追随する描像が自然であり、多くの分子理論を支えている。この近似描像の妥当性と限界については数理的によく調べられており、軽い電子と重い原子核の質量比とハミルトニアン演算子の構造に依る。原子核配置の動きは電子にとっての外場変動である。一般に、電子運動にとって緩やかに変動する外場、例えば低い周波数成分からなるレーザー場が印加された場合にも同様の断熱描像を考えることができる。

多様な化学変化を伴う電子エネルギー輸送、中間捕縛、電子揺らぎ、電荷分離、化学反応場等、太陽光デバイスを特徴付ける化学機能発現の契機となる電子励起を伴う光化学過程においては、関与する複数の電子状態間に準位近接が生じ、断熱仮定が破綻する。厳密には、あらゆる分子構造においてこの破れは生じており、強度の差こそあれ、常に電子状態間の遷移が起きている。有意な強度がある場合、対比する化学過程との時間スケール比に応じ、破れへの対処法を選択することになる。例えば、分子運動スケールに対し十分長く、少しづつ電子状態遷移がおきる過程を特徴付けるために、無輻射遷移速度を摂動論に基づく黄金則表式を用いて求める例を考えるとわかり易い。

「状態近接領域で顕著に非断熱性が大きくなる」ことに言及したのは、動力学過程において電子遷移が状態エネルギー近接箇所で極めて強くなり、非断熱遷移項に対して非摂動的な扱いが必要になったり、離れている為に摂動的でよかつたり、或は極めてエネルギー差があることと対応して無視できる、といった様相が時間に沿い変化することを強調したかったためである。原子核間の配置構造の変化に電子状態が瞬時に追随できず電子状態間を移りあったり、断熱電子状態の重ね合わせ(=電子波束)が出現したりする。これを総称し分子の非断熱動力学過程と呼ぶ。電子の内部自由度と電子の運動の結合に由来する спин・軌道相互作用を介した異なるスピン多重度間の系間交差を含む一般的な場合については、更に計算コストが大きくなり、計算科学としては大きなチャレンジとなる。大型の計算資源の割当が切に望まれる。

非断熱遷移後、或は相互作用領域を通過した後、関与する電子断熱状態のポテンシャル面の勾配は一般に異なる為に、多くの場合、長時間後には分子波束は複数の電子状態面に枝分かれ(非断熱分岐)していく。サーフェスホップ法や非断熱分岐法に対して、初期状態に対する適切なサンプリングや、干渉破壊補正、軌道相関補正、原子核の量子効果補正等の工夫を施すことにより、記述の定量的および定性的精度をあげができる。改良近似手法は多数存在し、記述精度と目的に応じ選択或は開発するのが望ましい。統一的な記述としては、非断熱古典軌道に沿い時間大域的に評価した電子の非断熱性を有効作用に繰り込む、Pechukas流の経路積分に基づく手法が理論形式的に整っている。実時間経路積分は虚時間発展と比べ収束が著しく難しく、適切な和の取り方に関する計算科学に基づく基礎研究と、大型数値計算を用いた電子動力学への応用展開研究は学理的にもデバイス応用研究の面でも重要な課題であり、大型計算資源の確保が望まれる。

多くの非断熱化学過程の研究においては、電子遷移に伴う状態占有率の変化、電子緩和の時間スケールに注目することが多い。実験計測で得たデータと比較することにより、理論で仮定したシナリオを査定する形で、光化学過程における電子失活の機構解明が行われてきた。近年、こうした電子失活時定数の研究を超えて、光短パルス生成を含む分光観測技術の発達と実時間測定理論の進展に伴い、サブフェムト時間スケールでの電子状態変化の追跡が可能になりつつある。種々のエネルギー領域の光を用い、運動スケールを多角的に計測することで、化学変化やデバイスにおける、電子状態の動的変化の源に迫る重要な基礎研究である。計算科学にはこれらをサポートする能力があり、大型の計算資源の確保が求められる。

電子波束の時間変化は微視的な意味で化学反応動力学と対応づくこと、励起状態の重ね合わせからは多彩な電子状態が発現し得ることを考え合わせれば、光照射下での化学動力学における非断熱電子波束の時間発展をより基礎的な立場から記述、追跡そして解読する研究は、励起状態が多種多様に相關することで化学機能に結びついている複雑な励起化学過程の理解と制御を目指す上でも、電子波束動力学を基礎とした新しい化学を一般展開するという意味でも今後益々発展させたい分野である。太陽光エネルギーのデバイスを通じた加工、特に、制御可能なエネルギー形態への転換に向けた基礎研究としての貢献が可能である。

原子核座標微分重なりとしての電子状態間非断熱結合と分子運動速度の積に比例する強度で、複数の状態からなる部分空間を電子波束は変遷していく。特に、電子の非断熱性は、励起状態が密集する高擬縮重励起状態が関与する系で極めて顕著である。異なる特性を持つ状態の重ね合わせや、干渉破壊後に出現する状態の統計的混合と、結合変化を促す反応場や内部励起エネルギー移動の関わりについて理解を得ることは、エネルギー変換デバイスをより微視的なレベルからマクロなレベルへと連結する上で欠かせない。その為には、動的な非断熱電子波束と構造動力学を研究することが必要であり、これに伴って大型の計算機資源が必要となる。擬縮重励起状態では、エネルギー的に近接する電子状態が、電子にとっての外場を通じ容易に相互作用する。ここで、外場は、原子核配置の時間変化や輻射場を指す。エネルギー変換デバイスではこれらが質的に高い電子エネルギーの効率的な輸送特性に強く影響を及ぼす。

金属を含む化合物は一般に柔軟で多様な電子反応場を提供するが、多くの場合生成コストがかかることが多い。有機分子や周期表低位置の元素による代替が望まれている。例えば、低い周期表位置の元素からなり、励起状態に擬縮重性が発現し易い単体原子クラスターの例として、ホウ素クラスターがある。ホウ素は価電子軌道空間を埋める電子の数が少ないとから、集合しクラスター化することで、価電子自由度を主成分とした擬縮重励起状態群を生み出し易い。同様に価電子空間に豊富な空軌道の存在する遷移金属原子と異なり、比較的小さな軌道角運動量で軌道広がりも小さく、基底、励起状態ともに、僅かな分子運動に対し化学結合と電荷局在様式がフラストレーションを起こす。こうした特性は太陽光を吸収した後の物質媒体中におけるエネルギー輸送の制御にとって、一つの可能性を提供すると想像できる。要点を抑えつつ、できる限り第一原理的なアプローチ、例えばDFT、TDDFT、EOM-CCSD、MSCASPT2、MRCI、post-DMRGなどの高精度電子状態計算理論を活用して上記に向けた研究を行えれば素晴らしい。その為には、従来よりも大型の系を取り扱うことも含め、転送速度の点で高度な機能をもつメモリ資源と巨大な計算機環境が必要となる。

(vi) 溶液シミュレーションのためのアプリケーション群

酵素反応・ドラッグデザインなどさまざまな分野において溶媒環境を含めた計算が必須のものとなっており、分子シミュレーションのみならず液体の統計力学理論によるアプローチが不可欠の技術となっている。とくに、生体分子の溶媒和を志向した手法（3D-RISM、MC-MOZ、XMOZ 法）を基盤とした開発が有効と考えられる。これら手法の高速化・高精度化とともに量子化学計算手法や分子シミュレーションと行った手法との組み合わせによるマルチスケール手法の開発・高度化が不可欠である。

4.3.3.1 要求性能

これらの研究対象における要求をもとに必要なアプリケーションソフトウェアの要求性能を以下に記載する。

(1) NTChem

基底状態

大規模分子系（数万原子）に対して、混成密度汎関数法による構造最適化を行う。ダブルゼータの基底関数、合計 10 万を用いる。単一ノード演算性能が 10 TFLOPS のマシンで 1024 ノードを使用して行う場合を想定する。

演算量 1 回約 30 分の計算を 50 回繰り返すため、合計で 25 時間かかる。約 900 EFLOPS の計算が推定される。

メモリ 疎行列と積分グリッド点をメモリ上に確保する。1 ノードあたり 48 GB、合計で 48 TB になると推定される。

ストレージ チェックポイントと後の解析のために、ハミルトニアンと密度行列を MPI ファイル I/O により保存する。100 GB が必要だと推定される

通信性能 通信コストを計算時間の 10% に抑えるにはバンド幅として 15 GB/s が必要である。

MP2 摂動法

数万～数十万基底の大規模系に対して、2 次の Møller-Plesset 摂動法 (MP2) に基づく第一原理計算を行う。Resolution of the Identity (RI) 近似を適用する。単一ノード性能が 10 TFLOPS のマシンで 2 万ノードを使用して行う場合を想定する。

演算量 RI-MP2 法において最も演算量が多く、コストが高いのは 4 中心積分計算における倍精度浮動小数点数の行列行列積 DGEMM である。計算にかかる時間は 1500 秒で、15 EFLOPS と推定される。

メモリ 1 ノード 64 GB として、合計 128 TB が推定されるが、行列演算を適切にメモリ分割することで軽減することは可能である。

ストレージ 分子軌道係数をファイルに保存するのに 2.9 GB 必要である。

通信性能 計算時間の 10% で通信コストを抑えるためにはノード当たり 15 GB/s のバンド幅が必要になる。

励起状態計算 (TDDFT)

大規模分子系（数万原子）に対して、混成密度半関数法による時間依存密度汎関数法に基づく励起状態計算を行う。単一ノード演算性能が 10 TFLOPS のマシンで 1024 ノードを使用して行う場合を想定する。

演算量 Davidson アルゴリズムを用いた反復計算によって、励起状態の解を求める。系の種類や大きさに依存するが、20 個の励起状態の解を求めるためには、10 回程度の反復計算が必要となる。1 回の反復計算に状態あたり約 30 分の計算を繰り返すため、合計で 100 時間かかる。約 3600 EFLOPS の計算が推定される。

メモリ 各状態の試行ベクトルとシグマベクトルに対する疎行列と積分グリッド点をメモリ上に確保する。1 ノードあたり 960 GB、合計で 960 TB になると推定される。

ストレージ 計算途中の試行ベクトルなどのデータはメモリに主に保存するが、いくつかはストレージに中間ファイルとして保存する。チェックポイントと後の解析のために、収束したと励起波動関数とハミルトニアンを MPI ファイル I/O により保存する。960 TB が必要だと推定される。

通信性能 通信コストを計算時間の 10% に抑えるにはバンド幅として 15 GB/s が必要である。

APPENDIX

4.3.4 超並列環境での大規模分子計算に向けた NTChem の取組み

4.3.4.1 ハイブリッド型密度汎関数法に対する線形スケーリング計算技法

ここでは、局所的な Gauss 型基底関数に基づいたハイブリッド型密度汎関数法の Hamiltonian (Kohn–Sham Hamiltonian) に含まれる Coulomb 積分と HF 交換積分の計算について焦点を当てる。密度汎関数法は、低成本の平均場近似のもとで電子相関を考慮した計算が可能であるので、大規模な多原子分子系に対する電子状態計算にもっとも適した計算手法である。中でも密度汎関数の交換汎関数に Hartree–Fock (HF) 交換項を含めたハイブリッド型密度汎関数法は、分子の電子状態や物性値を精度良く記述することができるため、有望な電子状態理論である。Coulomb 行列 \mathbf{J} および HF 交換行列 \mathbf{K} の行列要素 (それぞれ、Coulomb 積分および HF 交換積分) は、密度行列 \mathbf{D} を使って以下の式からそれぞれ計算される。

$$J_{pq} = \sum_{rs} D_{rs}(pq|rs), \quad (4.11)$$

$$K_{pq} = - \sum_{rs} D_{rs}(ps|rq). \quad (4.12)$$

ここで、 p 、 q 、 r 、 s は原子軌道に関するインデックスであり、 $(pq|rs)$ は Gauss 型基底関数 $\chi(\mathbf{r})$ を使うと解析的に計算することができる 2 電子原子軌道積分を表す。

$$(pq|rs) = \iint d\mathbf{r}_1 d\mathbf{r}_2 \chi_p^*(\mathbf{r}_1) \chi_q(\mathbf{r}_1) \frac{1}{r_{12}} \chi_r^*(\mathbf{r}_2) \chi_s(\mathbf{r}_2). \quad (4.13)$$

Coulomb 積分と HF 交換積分の計算コストは、形式的には 2 電子積分における 4 つのインデックスに関するループに起因して、系の大きさの 4 乗に比例する。しかしながら、基底関数の局所性を有効に利

用して、寄与の小さい解析的積分を計算しないことで、式(4.11)と式(4.12)の計算スケーリングを減少させることが可能である。計算する必要がある解析的積分の数を減らすための有効なツールの1つは、Cauchy-Schwartz のスクリーニング法 [1] である。Cauchy-Schwartz スクリーニングでは、2電子積分に対し次の関係が成り立つことを使ってスクリーニングする。

$$|(pq|rs)| \leq C_{pq}C_{rs}, \quad (4.14)$$

$$C_{pq} = \sqrt{(pq|pq)}. \quad (4.15)$$

このスクリーニングを用いると、スケーリングがほぼ2乗に減少することがわかっている。スクリーニングの際に密度行列をあらわに考慮することで、計算する積分の数をさらに減らすことが可能である。これは、Gauss型関数のような局所的な基底関数を使用した場合、大規模な絶縁体や有限温度の金属に対して、空間的に離れた密度行列要素が大きく減衰するためである (Kohnの近視則)。次の関係を使って必要な積分をスクリーニングする。

$$\sum_{rs} D_{rs} |(pq|rs)| \leq \sum_{rs} |D_{rs}| C_{pq} C_{rs}. \quad (4.16)$$

これらの技法により計算する積分の数を減らすことができるが、系のサイズが大きくなるにつれて、スクリーニングに費やされる時間が積分計算に必要な時間を超える可能性がある。KussmanとOchsenfeldは積分のプレスクリーニングを行うことで、この問題を解決している [2]。この方法では、HF交換積分が次のような上限を持つことを使う。

$$K_{pq} \leq [\mathbf{C} \times |\mathbf{D}| \times \mathbf{C}]_{pq}. \quad (4.17)$$

彼らは HF 交換積分に対してのみ積分のプレスクリーニングを導入したが、Coulomb 積分についても同様の上限を導入することができる。

$$J_{pq} \leq [\text{trace}(\mathbf{C} \times |\mathbf{D}|) \times \mathbf{C}]_{pq}. \quad (4.18)$$

スクリーニングにおける演算数のさらなる削減は、密度重みをつけた積分の推定値を大きさ順にソートし、あるしきい値に達すると積分の計算をスキップするようにすることで達成することができる。

4.3.4.2 スクリーニング技法の並列化実装に伴う問題点

4.3.4.1 では、ハイブリッド型密度汎関数法計算に必要な Coulomb 積分と HF 交換積分に対する線形スクリーニング技法を紹介した。ここでは、スクリーニング技法の並列化実装に伴う問題点について指摘しよう。考慮すべき事項は2点ある。ひとつはメモリの制約である。式(4.11)と式(4.12)からわかるように、Coulomb 行列と HF 交換行列の行列要素 J_{pq}/K_{pq} を計算するためには、密度行列 \mathbf{D} のすべての行列要素が必要である。このため、多くの古い量子化学計算コードでは密度行列 \mathbf{D} のすべての要素を各 MPI プロセスに複製する必要があり、計算できる系のサイズが制限される。Mironovらは、このアプローチに基づいてハイブリッド MPI + OpenMP 並列を施し、Theta スーパーコンピュータ上のアクセラレータベースのファットノードを用いることで数千原子系の量子化学計算を行っている。[3] しかし、このよう

なアプローチでは、さらに大規模な分子系の量子化学計算は不可能である。2点目の考慮すべき事項は、大規模な行列計算を実行するためグローバルに行列を分散することによる通信コストの問題である。

これらの問題を解決するために、いくつかのアプローチが提案されている。量子化学計算ソフトウェアのひとつである GAMESS の過去の実装では、行列を一次元的に分散し、高度な負荷分散技術と組み合わせることで、低性能なネットワークでも有効な並列化を実現している [4, 5]。しかし、我々の知る限りでは、最近のスーパーコンピュータに対して同様のアプローチは報告されていない。量子化学計算ソフトウェア NWChem で使われている区分化大域アドレス空間 (Partitioned Global Address Space, PGAS) アプローチ [6, 7, 8, 9] は問題解決のためのひとつのアプローチである。PGAS アプローチではライブラリを使用することで、プロセス間で行列を分散し、リモートの行列要素を非同期的にフェッチする。しかし、通信パターンが複雑になるため、通信コストが高くなり、通信に起因するロードインバランスが生じる。解析的積分計算の場合は、ネットワーク通信だけでなく、データをより低次のキャッシュに移動したり、アクセラレータに移動したりする際にも問題が生じる。

4.3.4.3 三次元分散に基づいた 2 電子積分計算部の並列実装

これらの通信の問題を軽減するために、我々はネットワークアーキテクチャを有効に利用したアルゴリズムを設計している。今日のスーパーコンピュータでは、さまざまなネットワークアーキテクチャが存在する。例えば、「京」では Tofu ネットワークアーキテクチャが用いられている。Tofu のような高次元トーラスネットワークは、様々な計算カーネルに対して十分なパフォーマンスを実現しており、特に線形代数演算に対しては非常に高いパフォーマンスを達成している。トーラスアーキテクチャにヒントを得たアプローチは、また、平面波基底・擬ポテンシャルを用いたハイブリッド型密度汎関数法にも適用されている [10]。

Gauss 型基底関数に基づいた Coulomb 積分と HF 交換積分の計算においても同様のアルゴリズムを開発することが可能である。密度行列と Kohn–Sham 行列 (Kohn–Sham Hamiltonian の行列表現) のデータ分散は、必要メモリ、reduction コスト、およびデータ gathering コストに影響を与える。これらのコストは、うまくデータ分散させることで制御することができる。特に、データを二次元に分散し、三次元目に複製する三次元分散を使用することで高いパフォーマンスを実現できる。このデータ構造を用いたアルゴリズムを図 4.5 に示す。まず、密度行列と Kohn–Sham 行列を X 軸方向と Y 軸方向に分割し、 Z 軸方向に複製し各プロセスに分散する。図 4.5 では X 、 Y 軸方向にそれぞれ 2 ずつ分割し、 Z 方向に 1 つ複製している。これらの分割はシェルペア単位（原子軌道の角運動量や短縮の組）で行われ、単位内の積分計算に必要なすべての密度行列を各プロセスが持つようにする。各プロセスの Kohn–Sham 行列要素（ピンク色）を計算する際には密度行列の総和を計算するため、密度行列のすべての要素が必要となる。そこでまず、同じ Y 軸のプロセス間で密度行列に対し集団通信（黒矢印）による収集を行い、 Z 軸で分割された密度行列の列要素（黄色）を収集する。その後、プロセス行のループに入る。各ステップで、同じ行のプロセス間での密度行列要素（図 4.5 では黄色）の収集が実行される。データが受信されると、ローカルで 2 電子積分が計算され、集められた密度行列と掛け合わせて局所的な Kohn – Sham 行列（赤色）に足し合わされる。最後に、 Z 軸方向に沿って集団通信で総和が実行され Kohn–Sham 行列（茶色）となる。このように各軸のコミュニケーション内で通信を伴う操作を別々に行うことで、メモリサイ

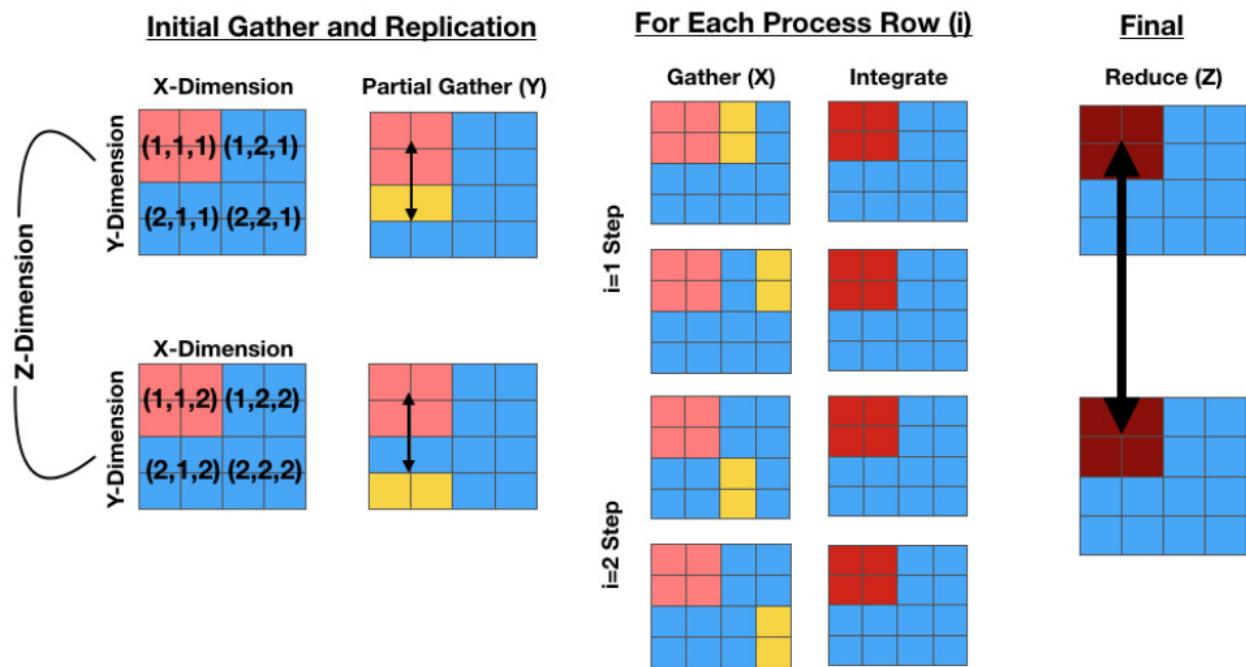


図 4.5 Coulomb 積分と HF 交換積分の三次元分散並列計算アルゴリズム。

ズと通信オーバーヘッドを減らすことができる。また、このアプローチでは、ループ中でプロセス行にわたる通信と計算をオーバーラップさせることができる。収集を行った後、いくつかのプロセスからのデータが受信される。次のデータセットに対してノンブロッキング収集が開始され、積分ブロックに関するループが始まる。1つのブロックが完了した後、send が調べられ、通信が積分計算より長くかかる場合にのみ barrier が必要になる。合算ステップも非同期に実行され、さまざまなタイプの通信と計算をオーバーラップさせることができる。

Z 軸方向の数はユーザーが入力として任意に選択することができる。使用ノードが増えるのにしたがって、行および列において収集を実行するコストが増加する。このコストを削減するために、ユーザーは Z 軸方向の数を増やすことができる。 Z 軸方向の数が大きくなるほど、データの複製により、あるノード上のメモリ占有量が大きくなる。さらに、総和の過程はより高いコストを持つため、これはトレードオフである。行列を構築するために使用される Z 軸方向の数の選択は、計算の他の部分と必ずしも同じにする必要はない。各ステップを調整するために、データリマッピングルーチンを使用してプロセスグリッドを調整することが可能である。

4.3.4.4 計算結果

開発したアルゴリズムのパフォーマンスを示す。計算には、NTChem の開発バージョンを用いた。このバージョンでは上記のアルゴリズムを採用し、さらに我々が開発している高並列疎行列ライブラリ NTPoly [11] を使って疎行列演算することで大規模な密度汎関数法計算を実現している。ノードあたり 4 つの MPI タスクと 12 つの OpenMP スレッド（コアごとに 1 つのスレッド）を使用した計算を「富

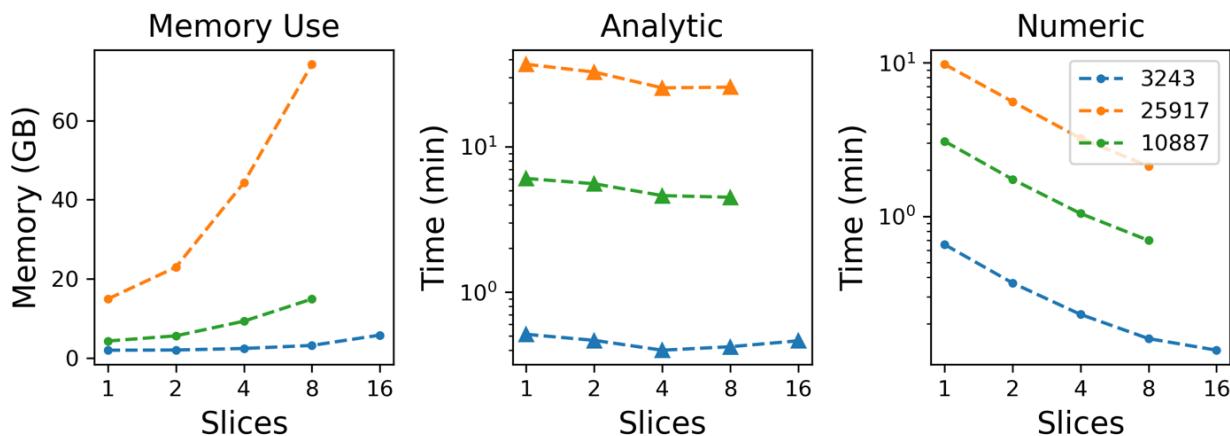


図 4.6 使用したノード数に対する 2 電子積分計算部の計算時間（横軸：原子数、縦軸：計算時間）。

岳」で実行した。密度行列を計算するとき、 1×10^{-6} 未満の行列要素をゼロとして、エネルギー計算での $O(N^3)$ スケーリングコストを持つ行列対角化を避けるために、二次トレース補正法に基づく密度行列 purification 法を使用している。ベンチマークの分子系として、三次元型水クラスターを使った。計算には MINI 基底関数を使用し、ハイブリッド型交換相関汎関数である PBE0 汎関数を使っている。

スライスサイズを変えたときのメモリ消費量とタイミングを図 4.6 に示す。スライスサイズを増やすと計算時間は短縮されるが、メモリ消費が増えるため、調整する必要がある。また、そのスケーリングを図 4.7 に示す。「富岳」は 1 ノードあたり 32 GiB のメモリしか持っていないことに注意してほしい。それにもかかわらず、本データ並列アルゴリズムにより、非常に大きな分子系を計算することが可能である。

ここで報告した計算例は非常に大規模な分子系であり、得られた計測時間はこれらの計算が標準ユーザーに対して許可された計算時間を超える可能性があることを示している。

4.3.5 RI-MP2 の高速化

2 次の Møller–Plesset 摂動法 (MP2) は、電子相関効果を考慮した波動関数ベースの理論的アプローチの中では最も単純なものである。しかしながら、MP2 の計算コストは基底関数の数 N に対して $O(N^5)$ で増大するため、実際に計算されるのは中程度の大きさの分子に限られる。これを解決するのが resolution-of-the-identity (RI) 法で、補助基底関数を用いて 4 中心電子反発積分 (electron repulsion integral, ERI) を 2 中心と 3 中心の ERI で近似することにより、化学的に妥当な精度を維持しながら計算時間、メモリ使用量、ディスクサイズを削減する。

以下では NTChem/RI-MP2 の高速化手法について [12] より改変して引用する。富岳ではターゲット問題実行時に京性能倍率 70 倍が達成された。

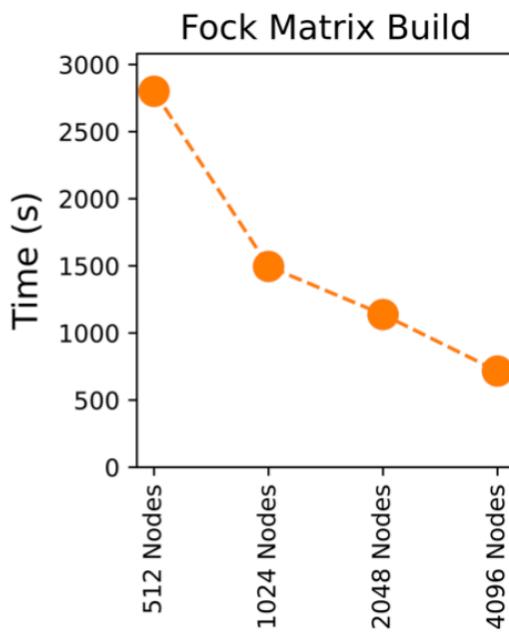


図4.7 使用したノード数に対する2電子積分計算部のストロングスケーリング（横軸：ノード数、縦軸：計算時間）。

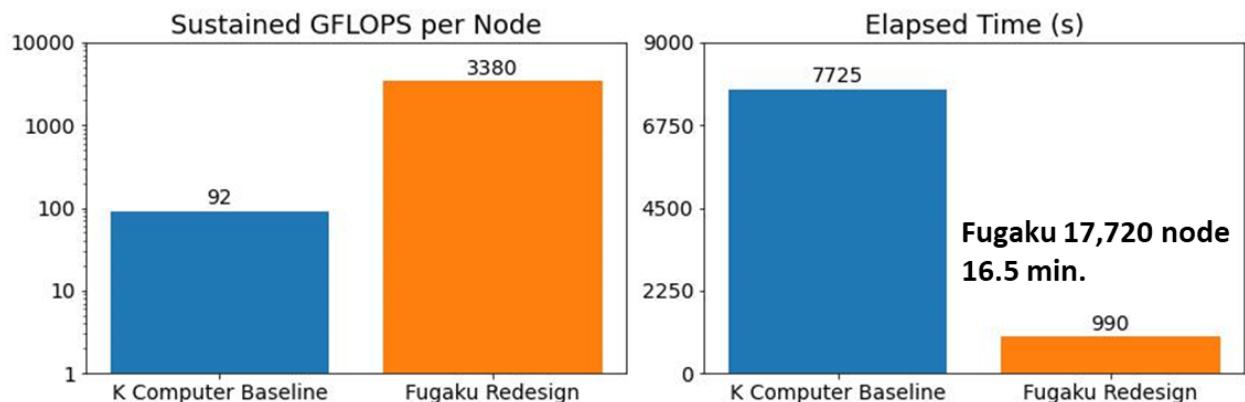


図4.8 RI-MP2法による計算効率。

4.3.5.1 RI-MP2計算の超並列実装

NTChem/RI-MP2モジュールにおけるMPI-OpenMPハイブリッド並列処理においては、計算のボトルネックとなる4中心ERI計算の並列化を行う際に、仮想分子軌道のインデックスをプロセス分割していた[13]。これにより占有分子軌道のインデックスをプロセス分割する従来のアルゴリズムよりも並列度が高く、京において数千ノード規模での効率的な並列計算の実行を可能となった。しかしながら、計算に用いることが可能なプロセス数が仮想分子軌道の次元数に制約されるため、数万プロセスを用いた超並列計算は不可能であることが問題となっていた。

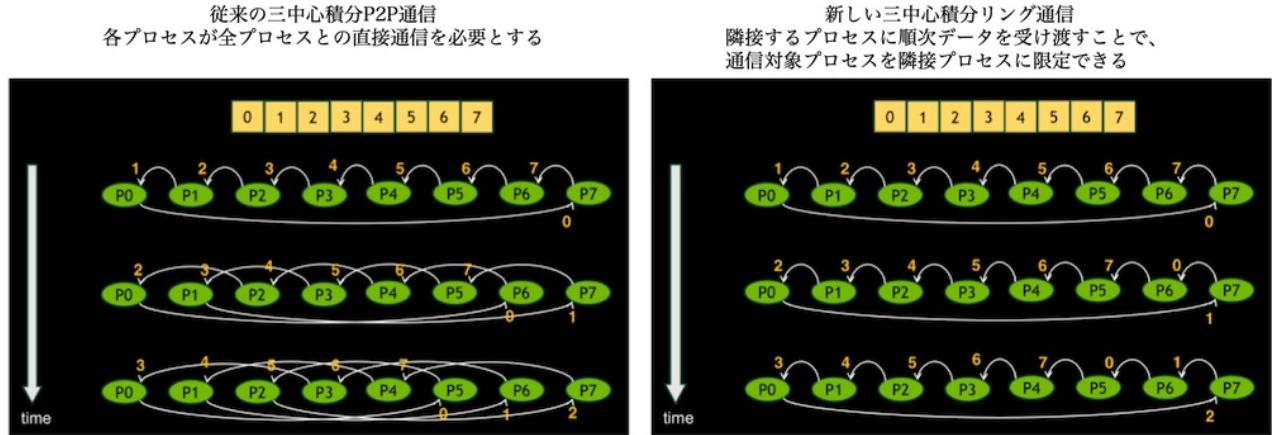


図 4.9 3 中心積分データ通信の P2P 通信法からリング通信法への変更。[12] より転載。

これを解決するため、局所的な MPI communicator を活用し 2 次元的な MPI 階層並列化を導入したアルゴリズムを開発した [14]。この改善では、最外ループとなる仮想軌道インデックスの第 1 軸の MPI 並列化に加え、このループ内で実行される積分計算と行列演算に関わる行列をブロック分割し第 2 軸の並列化を行った。

この 2 次元 MPI 階層並列化では、ノード間通信パターンを変更し 3 中心 ERI の 1 対 1 通信効率の向上作業も実施した。従来の実装では 4 中心 ERI 計算に必要な 3 中心積分データについて、データを保持するプロセスからデータを必要とする複数のプロセスへそれぞれ通信する P2P 通信法を採用していた(図 4.9 左)。この方法では、ネットワーク・トポロジーの存在するシステムでは通信相手の数が増大した場合に、ホップ数の違うプロセスとの通信で性能が低下することが、京でのベースラインプロファイルの結果より明らかとなった。通信されるデータの動きの詳細な解析を行った結果、データを隣接ノードに順次 1 対 1 通信で受け渡していくリング通信法(図 4.9 右)を採用することで、局所的なネットワーク・トポロジーを有効に活用した通信が可能であることが明らかとなったため、リング通信法の実装を新たに行なった。

4.3.5.2 RI-MP2 計算のメモリ分割実装

4.3.5.1 節で示した超並列実装を行ったコードでは、一部の行列演算を行う際に行列データをレプリカとしてメモリに保持して処理を行うアルゴリズムを採用している。しかしながら、本コードを用いて富岳におけるターゲット問題の計算を行うのに必要なメモリ量の見積もりを行った所、1 プロセスあたりで 31.6 GB のメモリが必要となる結果となり、本コードをそのまま用いた計算が不可能であることが明らかとなった。

基本設計時では、1 CMG あたりの使用メモリ量を低減するために、行列データの分散化を進めて、各プロセスがレプリカを排した計算が可能となるようなメモリ分割アルゴリズムの検討を行った。検討したメモリ分散アルゴリズムを想定して、後述する富岳ターゲット問題規模の計算を行う際に必要なメモリ量の概算見積りを行ったところ、富岳想定アーキテクチャでの計算が可能であるとの見通しを得た。

1. 2中心2電子積分行列計算のメモリ分割実装
2. 2中心2電子積分行列計算のコレスキーフィルタリングおよび分解して得られた上三角行列の逆行列の計算のメモリ分割実装
3. 3中心積分計算のメモリ分割実装
4. 行列を用いる一連の演算について行成分を一度に取り扱わず、分割して複数回に分けて行い、一時的配列についても任意にサイズを確保可能とする
5. 混合精度化アルゴリズムを導入し、倍精度データにかえて単精度データをメモリに格納することで、4中心積分計算区間で確保する配列のメモリサイズを半減化
6. 従来全プロセスが保持していた分子軌道係数の全データ 3.1 GB の格納方法を変更し、RI-MP2 計算中は 0.3 GB の占有軌道に関する係数は全プロセスが保持し、仮想軌道の係数は各プロセスが担当する部分（データ量は $(2.8/N\text{ProcsMat}) \text{ GB}$ ）のみを保持する。

これらの実装では 2 電子積分計算を基底関数の全角運動量に依存したシェルブロック単位で行うため、行列の不均一な列ブロック分割に基づいた実装を行った。

(1) の 2 中心 2 電子積分行列計算のメモリ分割実装については、列方向のシェルをブロック分割し各ブロック行列要素を MPI 並列化して計算する実装とした。シェルと基底関数の対応関係の都合上、ブロック分割した行列の次元に相当する基底関数の次元は不均一となっている。

(2) の行列のコレスキーフィルタリングおよび上三角行列の逆行列計算のメモリ分割実装では、(1) で得られたブロック行列毎に MPI 並列化し演算を行う実装とし、演算の実装にはコレスキーフィルタリングルーチン DPOTRF と逆行列計算ルーチン DTRTRI、BLAS のレベル 3 の行列–行列積 DGEMM を用いた。計算に必要なブロック行列データの通信についてはサブコミュニケーション MPI_COMM_MAT 内で MPI_Bcast 通信を行い、行方向にデータを送信する実装とした（図 4.10、図 4.11）。

(3) の 3 中心積分計算のメモリ分割実装では、3 中心積分計算時に列方向のシェルをブロック分割し各ブロックの計算を MPI 並列化する実装とした。(1) の 2 中心積分計算と同様にブロック分割した列の次元は不均一となっている。3 中心積分計算後に係数行列を作成する際の行列–行列積演算はそれぞれのブロックごとに BLAS の DGEMM を用いて行い、得られた行列データの部分和を必要に応じて MPI_COMM_MAT 内で MPI_Allreduce で収集する実装とした。

(1)–(3) の削減案を全て追加実装し、その効果を確認するために、京の上でターゲット問題の計算を実行した。当初案のプロセス数 82,928 と二次元分割パラメタ 16 を用いた場合、プロセスあたりのメモリ使用量は 7.5 GB で、プロセス数 71,280 と二次元分割パラメタ 8 では、6.5 GB となった。富岳において 1 CMG に 1 プロセス（1 ノードに 4 プロセス）を割り当てる想定すると、メモリ使用量は削減されたが、上限に近い値となることが分かった。そのため、さらに 4–6 を行った。

4.3.5.3 RI-MP2 計算の単精度・倍精度の混合精度化アルゴリズム

RI-MP2 計算において最も演算量が多い区間は 4 中心積分計算区間で、倍精度密行列の行列–行列積で BLAS Level3 の DGEMM を用いて処理を行っていた。一方で、MP2 法やより高精度な結合クラスター理論において実際に計算で得られる電子相関エネルギーは、HF や DFT のエネルギーよりも 4–5 倍程度小

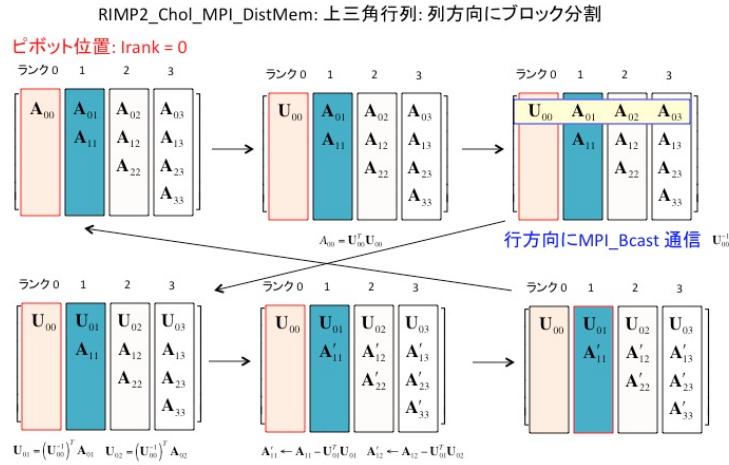


図 4.10 行列の Cholesky 分解のメモリ分割並列実装。[12] より転載。

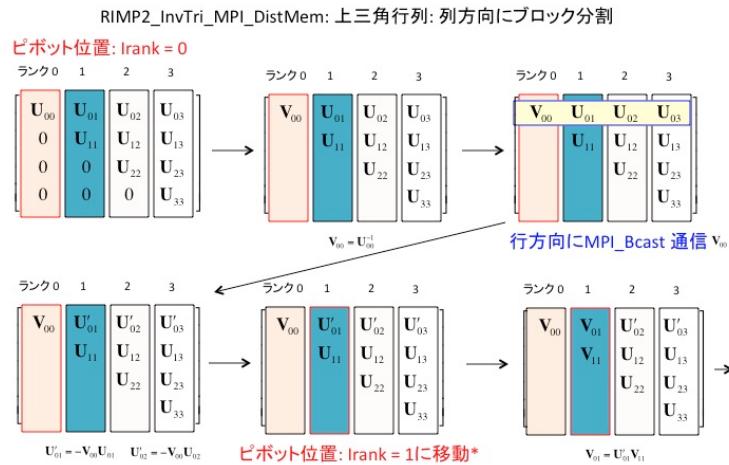


図 4.11 上三角行列逆行列計算のメモリ分割並列実装。[12] より転載。

さい値であり、実アプリで許容される有効数字として小数点以下 3 桁 (原子単位で 10^{-3} hartree のオーダー) 程度となる。

一方、京では单精度と倍精度で浮動小数点演算のコスト (演算スループット) が同じであるが、富岳においては单精度 SIMD 演算を倍精度 SIMD 演算の 2 倍の効率で実行できるマイクロアーキテクチャが採用され、計算性能・メモリ要求量の両観点から单精度計算の利点について認識が広まった。单純に計算全体を单精度で実行すると RI-MP2 計算に求められる精度を確保できないことは明らかであるため、電子相関エネルギー計算の実際の利用に耐えうる精度をアプリレベルで確保した上で計算を高速化することが可能な混合精度化アルゴリズムの検討を行った。

まず計算に必要な行列あるいはテンソル要素のうち、絶対値の小さなものについては単精度で保持して計算を進めるアプローチで検討を実施した。

- 倍精度密行列データ \mathbf{A} を、ある閾値 δ で単精度の密行列 \mathbf{A}^S と倍精度の疎行列 \mathbf{A}^D (compressed sparse row, CSR 形式) に分割、 $\mathbf{A} = \mathbf{A}^S + \mathbf{A}^D$, $|a_{ij}^S| < \delta$
- 単精度密行列同士の積 $\mathbf{A}^S \mathbf{B}^S$ を BLAS SGEMM で実施
- 倍精度疎行列と単精度密行列の積、 $\mathbf{A}^S \mathbf{B}^D$ 、 $\mathbf{A}^D \mathbf{B}^S$ 、および倍精度疎行列同士の積、 $\mathbf{A}^D \mathbf{B}^D$ 、を密ベクトル-疎ベクトルの内積として疎行列ライブラリで実施

行列要素の単精度の割合とハードウェアアーキテクチャのスペックをパラメータとした性能推定モデル式を構築したところ、富岳の場合全行列要素のうち倍精度要素が 5% 以下ならば、高速化が期待されることが分かった。

まずプロトタイプ実装として、4 中心積分計算の際の行列-行列積を BLAS の DGEMM を用いた倍精度密行列計算を SGEMM を用いた単精度密行列計算へ置換えた。代表的なテスト分子を対象に混合精度化の効果の評価を行ったところ、いずれのテストケースにおいても、元の倍精度版を用いた際と比較して混合精度化による誤差の絶対値が 10^{-6} hartree 以内となり、実アプリケーションの際に影響が及ばない範囲となった。また、混合精度版を用いた際の実行時間は倍精度版の結果と比較して 1.5–2.0 倍高速化する結果となり、演算と通信の単精度化による高速化の効果が確認された。

本プロトタイプ実装を用いた性能評価の結果、十分な精度を保つつ高速化が達成される結果となった。そこで、さらに実装を進めて、4 中心積分計算の区間に加えて、3 中心積分の変換の区間に対しても同様に、DGEMM を SGEMM へ置換えた実装を行った。またその際には計算の下流の工程である 4 中心積分計算から遡った 3 中心積分の変換を下流から順に単精度化した 5 種類の計算精度モデルを検討した。

それぞれの混合精度モデルの数値誤差の影響と高速化の効果を調べるために、ナノグラフェンの RI-MP2/cc-pVTZ 計算を実施した。その結果、4 中心積分のみを単精度化すると 10^{-6} hartree の精度を満足しながら、1.4 倍以上の高速化が確認されたが、それに加えて 3 中心積分の変換も単精度で計算すると、これ以上の高速化が達成されないものの数値誤差が増大し精度が不十分になることが分かった。

MP2 法の特長のひとつは弱い分子間相互作用を正しく記述できることで、これは平均場近似である HF 法では不可能である。そこで、以下の複合系データセットを対象に RI-MP2/cc-pVTZ で分子間相互作用を計算し、単精度計算の数値検証を行った。

1. 小さな分子系 66 種類について分子間距離を 8 通り変化させるデータセット S66 × 8[15]
2. 様々な水分子の集合分子系 38 種類を取り扱った Water Cluster データセット [16]
3. 高周期のハロゲン原子を含む分子系 40 種類を取り扱った X40 データセット [17]
4. 比較的大きい分子系を取り扱った L7 データセット [18]

その結果、1–3 のデータセットについても分子間相互作用エネルギーにおける倍精度計算と単精度計算の平均絶対誤差は 0.1 cal/mol 以下となり、これは実用的な計算における精度 0.1 kcal/mol を達成していることから、単精度計算を用いても全く問題ないことが確認された。しかしながら 4 では、4 中心積分のみを単精度化すると誤差が 2 cal/mol であったものの、4 中心と 3 中心積分を単精度化すると誤差が 1 kcal/mol 以上になった。一連の検証結果から、分子サイズを大きくするに従って誤差が増大し、比較

表 4.3 ターゲット問題の RI-MP2 相関エネルギー (a.u.) と倍精度 (DP) との誤差 (kcal/mol) 計算
精度モデル 相関エネルギー (a.u.) DP との誤差 (kcal/mol)

的大きな分子系の分子間相互作用計算においては、単精度計算の置換えに注意を払う必要があることが明らかとなった。

以上から、大規模系で精度を維持して混合精度計算を行う際は、4 中心積分のみを単精度化するのがよいとの知見が得られた。

4.3.5.4 RI-MP2 計算における密行列データ圧縮によるネットワーク通信量削減の検討

4.3.5.3 の浮動小数点演算の混合精度化を検討する中で、計算に必要な行列データも単精度の密行列と倍精度の疎行列 (CSR 形式) に分割することにより、倍精度密行列のみを用いる時よりもデータ量を削減できることが明らかとなった。この圧縮されたデータ構造を用いた場合、通信データ量の削減も期待される。そこで性能推定モデル式を構築し、富岳での性能を推定したところ、倍精度行列要素が 80% 以下ならばデータ圧縮による高速化が期待される結果を得た。次に前述のテスト計算結果を用いて通信量削減効果を見積もったところ、ほぼ全ての行列要素を単精度として扱えるため、データの単精度化による通信性能削減効果が期待できることがわかった。

4.3.5.5 RI-MP2 計算における 3 中心積分計算の高速化

NTChem/RI-MP2 を用いたターゲット問題の計算時間全体に占める 3 中心 ERI 区間の計算時間の割合は低いが、他の分子モデルの計算においては相対的に 3 中心 ERI の計算時間の割合が高くなる場合があるため、高速化を行った。NTChem/RI-MP2 の当初の実装においては McMurchie–Davidson (McD) アルゴリズム [19] に基づく独自の 3 中心積分計算ルーチンを採用していたが、改めて 3 中心積分計算方法に関する調査研究を行い、分子科学研究所の石村和也研究員が開発している SMASH 量子化学計算プログラム [20] の積分計算ライブラリが高速計算を実現するためのアルゴリズムを有し、NTChem/RI-MP2 計算プログラムに組み込み可能なプログラム構造を有しているとの評価結果を得た。この評価をうけ、SMASH の積分計算ライブラリを NTChem/RI-MP2 にルーチンとして組み込み、従来の独自ルーチンと SMASH ルーチンとを実行時に選択して計算することが可能となる実装を行った。3 中心積分計算では特に高い角運動量関数を基底とする積分の計算コストが高くなる。SMASH の積分カーネルコードは角運動量の低い基底の低い積分と高い積分を異なるアルゴリズムで計算することにより、計算コストの削減に成功している。角運動量の低い積分では Pople–Hehre (PH) アルゴリズム [21] と McD アルゴリズムを組み合わせた石村研究員開発のアルゴリズム (IN) [22] を、角運動量の高い積分では Rys 求積法を用いたアルゴリズム (DRK) [23] を採用している。

PH 法は基底関数中心間の二つのベクトルを座標変換することによって、積分を計算するための演算量を減らし、積分計算を高速化している。また、McD アルゴリズムはエルミートガウス関数と漸化関係式を導入することにより、計算コストを大幅に削減している。IN 法ではこの二つのアルゴリズムを組み合

わせることで高速な計算を実現している。不完全ガンマ関数(誤差関数)を区分求積法で算出し、その不完全ガンマ関数からエルミートガウス関数を求めた上で、電子反発積分を計算する。

DRK 法は数値積分を用いることにより、計算コストの高い不完全ガンマ関数の計算を回避する手法で、特に角運動量が高い基底の積分計算を行うときに威力を発揮する。Rys 求積法から出発して、漸化式を用いて中間積分を計算し、それら中間積分から電子反発積分を算出する。

導入した SMASH ライブラリルーチンを用いてターゲット問題の 3 中心積分区間の計算時間を FX100 システム上で試験的に短時間評価した結果、当初のプログラムでは 1.91 秒かかっていたのが、SMASH ライブラリルーチン版では 0.49 秒となり大幅な時間短縮効果が得られた。

この結果をもとに、以降の NTChem/RI-MP2 コデザインにおいては 3 中心積分計算を SMASH ライブラリルーチンを用いて進めることとした。公開されているバージョンの SMASH ソースプログラムを ASIS の状態で NTChem/RI-MP2 に組み込んで、FX100 システム上で 3 中心積分区間を詳細プロファイル分析を実施した結果、以下の改善がさらに望まれることが判明した。

- 演算待ち時間削減のための命令スケジューリング改善
- 整数・浮動小数点 L1D アクセス待ち時間削減のための命令スケジューリング改善
- 整数・浮動小数点 L2 アクセス待ち時間削減のためのデータアクセスの改善

SMASH ライブラリルーチンはそれ自体が相当量のソースプログラム行数を含み(1万行以上)、ライブラリの中でさらに計算負荷特性が異なる複数のサブルーチンが呼び出されるため、今後の富岳性能評価に向けて、SMASH ライブラリの内部をさらに細かく区間化し、中でも経過時間が長くかかる数区間を改めてカーネル化して、性能シミュレータによる性能評価・最適化を行う方針とした。

4.3.5.6 入力ファイル読み込みの高速化と選択切り替え

ターゲット問題の分子軌道係数ファイルサイズは、デフォルトのテキストデータでは 9.4 GB であり、京 1 ノードで読み込み時間を測定すると 588.8 秒であった。分子軌道係数ファイルのフォーマットをバイナリに変更することにより、ファイルサイズは 2.9 GB まで削減され、さらに書式変換の処理が不要となるため、京 1 ノードでの読み込みは 82.6 秒へ短縮できた。このため富岳コデザイン版では分子軌道係数の読み込みはテキストデータでもバイナリデータでも可能とし、プログラム実行時に選択できるようにプログラム仕様と入力仕様を変更した。

4.3.5.7 計算と通信のオーバーラップの明示的な促進

アシスタントコアを用いたノンブロッキング MPI 通信時間と計算コアを用いたの演算時間との積極的なオーバーラップを行なって、通信時間の見かけ上の待ち時間を短縮するために、富岳で利用可能な MPI ノンブロッキング通信促進 API を NTChem/RI-MP2 へ追加した。具体的には富岳における OpenMPI 4.0 に対する拡張ルーチンである FJMPI_Progress_start() および FJMPI_Progress_stop() をソースプログラム中に追加した。通信と演算のオーバーラップによる時間短縮効果は、小規模データを用いた測定では目立たないが通信時間・演算時間ともに長いターゲット問題規模のジョブでは効果が非常に大き

く、ターゲットジョブの測定結果で示されるように、計算時間全体における通信の待ち時間は効果的に削減された。

参考文献

- [1] Marco Häser and Reinhart Ahlrichs. Improvements on the direct scf method. *Journal of Computational Chemistry*, Vol. 10, No. 1, pp. 104–111, 1989.
- [2] Jörg Kussmann and Christian Ochsenfeld. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *The Journal of Chemical Physics*, Vol. 138, No. 13, p. 134114, 2013.
- [3] Vladimir Mironov, Yuri Alexeev, Kristopher Keipert, Michael D'mello, Alexander Moskovsky, and Mark S. Gordon. An efficient mpi/openmp parallelization of the hartree-fock method for the second generation of intel® xeon phi™ processor. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '17, New York, NY, USA, 2017. Association for Computing Machinery.
- [4] Yuri Alexeev, Ricky A. Kendall, and Mark S. Gordon. The distributed data scf. *Computer Physics Communications*, Vol. 143, No. 1, pp. 69–82, 2002.
- [5] Yuri Alexeev, Michael W. Schmidt, Theresa L. Windus, and Mark S. Gordon. A parallel distributed data cphf algorithm for analytic hessians. *Journal of Computational Chemistry*, Vol. 28, No. 10, pp. 1685–1694, 2007.
- [6] Jaroslaw Nieplocha, Robert J. Harrison, and Richard J. Littlefield. Global arrays: A nonuniform memory access programming model for high-performance computers. *The Journal of Supercomputing*, Vol. 10, pp. 169–189.
- [7] M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T.L. Windus, and W.A. de Jong. Nwchem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, Vol. 181, No. 9, pp. 1477–1489, 2010.
- [8] Xing Liu, Aftab Patel, and Edmond Chow. A new scalable parallel algorithm for fock matrix construction. In *Proceedings of the 2014 IEEE 28th International Parallel and Distributed Processing Symposium*, IPDPS '14, p. 902–914, USA, 2014. IEEE Computer Society.
- [9] David Ozog, Amir Kamil, Yili Zheng, Paul Hargrove, Jeff R. Hammond, Allen Malony, Wibe de Jong, and Kathy Yelick. A hartree-fock application using upc++ and the new darray library. In *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pp. 453–462, 2016.
- [10] William Dawson and Francois Gygi. Optimized scheduling strategies for hybrid density functional theory electronic structure calculations. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '14, p. 685–692. IEEE Press, 2014.

- [11] William Dawson and Takahito Nakajima. Massively parallel sparse matrix function calculations with ntpoly. *Computer Physics Communications*, Vol. 225, pp. 154–165, 2018.
- [12] 三上和徳, 河東田道夫. 富岳コデザイン・レポート (2022年3月公開) — 理化学研究所 計算科学研究センター (r-ccs). <https://www.r-ccs.riken.jp/fugaku/history/codesign-report/>.
- [13] Michio Katouda and Takahito Nakajima. Mpi/openmp hybrid parallel algorithm of resolution of identity second-order möller–plesset perturbation calculation for massively parallel multicore supercomputers. *Journal of Chemical Theory and Computation*, Vol. 9, No. 12, pp. 5373–5380, 2013.
- [14] Michio Katouda, Akira Naruse, Yukihiko Hirano, and Takahito Nakajima. Massively parallel algorithm and implementation of ri-mp2 energy calculation for peta-scale many-core supercomputers. *Journal of Computational Chemistry*, Vol. 37, No. 30, pp. 2623–2633, 2016.
- [15] Jan Řezáč, Kevin E. Riley, and Pavel Hobza. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *Journal of Chemical Theory and Computation*, Vol. 7, No. 8, pp. 2427–2438, 2011. PMID: 21836824.
- [16] Berhane Temelso, Kaye A. Archer, and George C. Shields. Benchmark structures and binding energies of small water clusters with anharmonicity corrections. *The Journal of Physical Chemistry A*, Vol. 115, No. 43, pp. 12034–12046, 2011. PMID: 21910428.
- [17] Jan Řezáč, Kevin E. Riley, and Pavel Hobza. Benchmark calculations of noncovalent interactions of halogenated molecules. *Journal of Chemical Theory and Computation*, Vol. 8, No. 11, pp. 4285–4292, 2012. PMID: 26605592.
- [18] Robert Sedlak, Tomasz Janowski, Michal Pitoňák, Jan Řezáč, Peter Pulay, and Pavel Hobza. Accuracy of quantum chemical methods for large noncovalent complexes. *Journal of Chemical Theory and Computation*, Vol. 9, No. 8, pp. 3364–3374, 2013. PMID: 24098094.
- [19] Larry E McMurchie and Ernest R Davidson. One- and two-electron integrals over cartesian gaussian functions. *Journal of Computational Physics*, Vol. 26, No. 2, pp. 218–231, 1978.
- [20] cmsi/smash: Massively parallel software for quantum chemistry calculations. <https://github.com/cmsi/smash>.
- [21] John A Pople and Warren J Hehre. Computation of electron repulsion integrals involving contracted gaussian basis functions. *Journal of Computational Physics*, Vol. 27, No. 2, pp. 161–168, 1978.
- [22] Kazuya Ishimura and Shigeru Nagase. A new algorithm of two-electron repulsion integral calculations: a combination of pople–hehre and mcmurchie–davidson methods. *Theoretical Chemistry Accounts*, Vol. 120, pp. 185–189.
- [23] Michel Dupuis, John Rys, and Harry F. King. Evaluation of molecular integrals over Gaussian basis functions. *The Journal of Chemical Physics*, Vol. 65, No. 1, pp. 111–116, 08 2008.

4.4 生命科学

本節では生命科学分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.4.1 現在行われている課題

生命科学分野でのシミュレーションの学術的意義が、今後ますます重要になっていくことは疑う余地がない。タンパク質立体構造を結晶構造解析やクライオ電子顕微鏡などの実験によって原子座標を解き明かすことの重要性は変わらないが、AlphaFold 2 に代表される機械学習の予測精度は急速に向上了ってきた。しかし、細胞内でタンパク質や酵素などの分子機械が働く姿を知ることが生命科学における重要な目的であり、そのためにはどうしても計算科学により実験科学で得られない情報を補完する必要がある。

生命科学研究分野におけるシミュレーション方法の特殊性は、時間スケールおよび空間解像度の異なるさまざまな現象を解析する必要があることである。そのため、研究の必要性や研究者の興味・関心となる現象に対してさまざまなモデルや方法論が利用される。方法論的には、おおまかに、生体分子を対象とする生体分子シミュレーション、ゲノム情報などの大量に蓄積されたデータを対象とするゲノム解析・遺伝子ネットワーク解析に分けられるので、その分類に沿って紹介する。細胞や臓器を対象とする細胞・臓器シミュレーションについては 2.6、4.6 節で、脳神経ネットワークのシミュレーションなどについては 2.5、4.5 節で説明する。

(i) 生体分子シミュレーション

生体分子シミュレーションには、全原子モデルを用いた分子動力学（全原子 MD）、全原子 MD を創薬に応用した創薬 MD、酵素反応などを解析する量子化学/分子力学混合モデル（QM/MM）、粗視化モデルを用いた分子動力学（粗視化 MD）がある。全原子 MD においても従来からよく用いられている点電荷モデルに加えて、分極モデルの利用も徐々に増えてきた。

全原子モデルを用いた分子動力学である全原子 MD では、小タンパク質の折れたたみ過程から、膜タンパク質、DNA タンパク質相互作用、リボゾームやウイルス全体などが研究対象としてよく用いられる [1, 2, 3]。原子数としては数十万から数百万原子、計算時間としては、数百ナノ秒から数十マイクロ秒程度の研究が標準的ではあるが、スパコン等の大規模な利用によって、原子数として十億原子規模の巨大なシミュレーションも行われている。特に、新型コロナウイルス 丸ごとを含む飛沫の計算ではシミュレーションする系に含まれる原子数は容易に十億を超える。このような大規模系に関しては数十マイクロ秒に至る長時間ダイナミクスは未だ困難であり、中規模から小規模系については、いかに述べる分子動力学専用スパコンを用いてミリ秒計算を実施した例が少数存在する。

海外では、米国 D. E. Shaw Research が開発した分子動力学専用スパコン ANTON[4] が特筆すべき性能を示している。汎用機の 1,000 倍近い強スケーリング性能により、1 ミリ秒を超える計算を実現し、80 残基程度までのタンパク質の折れたたみや薬剤の自発的結合を再現できている [2]。2021 年に開発された ANTON-3[5] では、7 千万程度の大規模な系への適用も報告されている。しかし、専用スパコンは計算速

度に関しては秀でているが、計算法の柔軟性・プログラマビリティに欠けており、一定の計算の範囲を超えると適用困難になる点が課題である。一方汎用スパコンは、さまざまな計算法の工夫、例えば、上記のマルチスケール・マルチレゾリューション法の導入や、複数のレプリカを用いてレプリカ間を疎結合で結ぶ手法で分子の統計量（自由エネルギーなど）を計算する手法には向いている。QM/MMについては4.6（創薬・医療）で説明する。

一方、巨大な生体超分子や細胞スケールの分子集合体に対しても、全原子MDの適用範囲が拡大しつつある。例えば、バクテリア（マイコプラズマ・ジェニタリウム）の細胞質中のほぼ全ての生体高分子と代謝物を原子レベルでモデリングし^[6]、現実的な濃度で集積した大規模細胞質モデル（総原子数約1億）が作成され、「京」とGENESIS^[7]（理研・計算科学研究機構（現：計算科学研究センター）が開発）による大規模MDから、細胞内分子ダイナミクスの微視的描像が得られた^[8]。更に米国のスーパーコンピュータを用いた10億原子の染色体システムの計算も行われた。^[9]しかし、系が1億原子を超えるサイズになる場合、全原子MD計算から得られる分子ダイナミクスの時間スケールは0.1-1μ秒程度であり、蛋白質間相互作用や代謝物と酵素の自発的結合を十分な回数観察するレベルには時間スケールが足りないのが現状である。

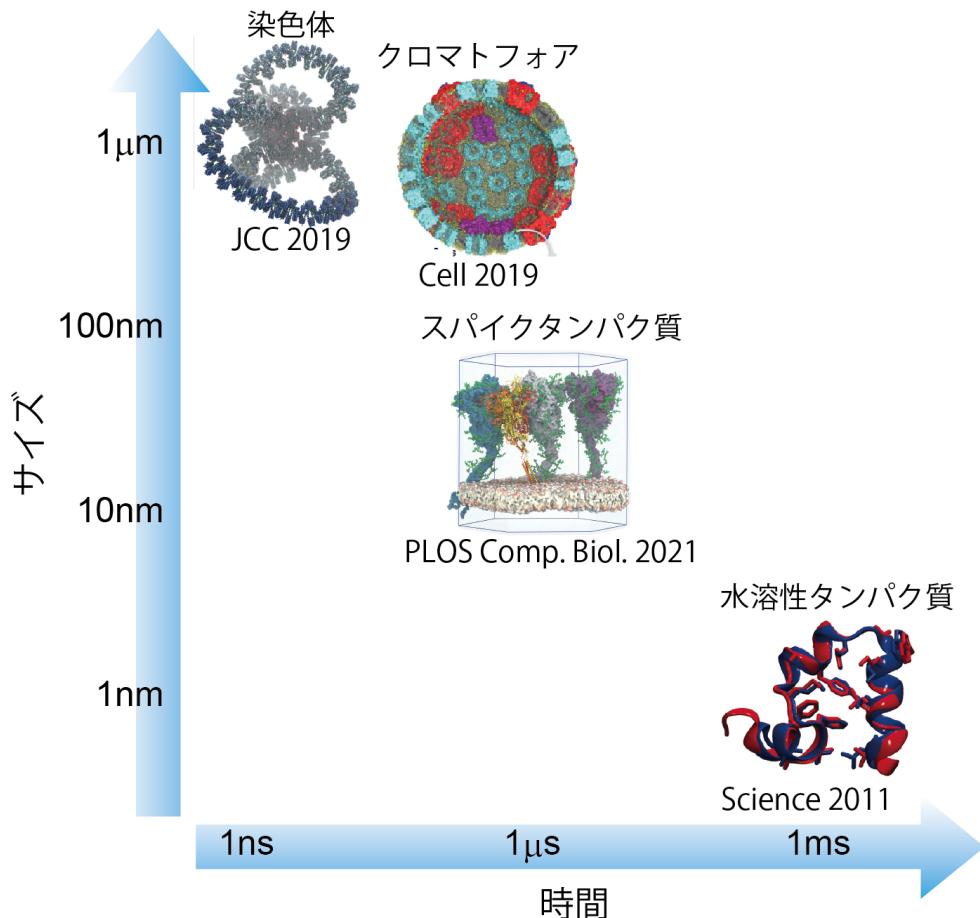


図4.12 代表的なシミュレーション [9, 10, 11, 2]

このように、大規模生体分子の全原子MD計算には、時空間スケールに限界があり、それを突破する

試みとして、粗視化モデルを用いた分子動力学である粗視化MDが盛んに行われている[12]。タンパク質折れたみ研究で開発された郷ポテンシャル（天然構造の相互作用のみを導入する）が、分子モーターや膜タンパク質などの構造解析にも適用可能なことが示され、世界的にもユニークな研究が数多く発表されている。更に近年では安定な天然構造を持たないタンパク質(IDP)、膜分子、DNA/RNAなどの核酸、更にはオルガネラまで様々な粗視化モデル構築と、モデル導出の方法論の開発が進み、広く応用されている。これらのモデルは、上記の郷ポテンシャルの流れとは異なり、様々な分子種の分子・原子間の相互作用の記述に力点を置いている。欧州の中心は、Marrinkらが開発する MARTINI 力場[13, 14]であり、最初の生体膜モデルから拡張されてタンパク質系も取り込んで発展している。方法論では米国の中 Vothらが精力的に理論展開をしており、それに基づいて巨大系の計算が行われている。更に米国ではスーパーコンピュータの性能を活かした粗視化モデルや数値モデルを用いたモデリングも盛んにおこなわれており、2020年以降に細胞システム[10]、エンベロープ[15]、コロナウィルスを含む飛沫[16]のモデリングが立て続けに発表されており、これらの大規模なシステムの全原子MD計算が現実的になりつつある。

(ii) 分子細胞シミュレーション

細胞質に代表される生体内混雑環境に対する分子シミュレーションは、従来行われてきた希薄水溶液に対する生体分子シミュレーションと実際の細胞内環境の間にあるギャップを埋める上で重要な試みである。2016年に報告されたマイコプラズマジェニタリウムの細胞質の分子シミュレーション（約1億原子）[8]では、タンパク質や核酸などの生体高分子や代謝物質の間の非特異的な相互作用によるタンパク質安定構造の変化や代謝物質の不均一分布など、分子レベルでの細胞内動態に関する新しい知見が得られた。細胞混雑環境を模倣したタンパク質混雑環境での一連のシミューションからは、過渡的なタンパク質複合体形成によるタンパク質の拡散速度低下のメカニズムや、混雑環境が脂質膜形態に及ぼす影響の解析など、系統立てた理解が着実に得られてきている。また、薬剤分子の標的タンパク質への結合過程に関するシミュレーションでは、混雑化に伴う標的タンパク質の局所的な構造平衡シフトによる結合経路の変化や標的タンパク質周囲の薬剤分子濃度の低下など、混雑環境効果による薬効低下メカニズムの一端が明らかにされた。分子細胞シミュレーションに関係した近年の取り組みとしては、分子細胞モデリングの精緻化が挙げられる。例えば、さまざまな解像度のモデルを経由することで細胞膜やミトコンドリアなどの細胞小器官の分子モデルを構築する方法論及びプログラム群が複数の研究チームによって報告されている。これらのツールや種々の実験データに基づく精緻な分子細胞モデリングが実現すれば、これまで以上に実際の細胞環境に近い分子シミュレーションが可能になり、細胞内動態に関する分子描像理解の飛躍的向上が期待できる。ヒトゲノムDNAは60億塩基対である。現在、100 kbpをひとつの粒子とみたポリマーシミュレーションによってゲノム全体をシミュレーションされ、染色体のテリトリーが生まれる機序が分かりつつある。染色体は、ヌクレオソームが数珠つなぎになった構造をしている。ひとつのヌクレオソームあたり200 bpと換算すると、3千万粒子である。現時点でのこの程度の規模のシミュレーションは可能である。しかし、この規模の分子のダイナミクスを調べるには、ミリ秒から秒のオーダーのシミュレーションが必要である。さらに、ヌクレオソームを全原子モデルで表現すると、約3万倍の粒子数、 1×10^{12} の規模の計算が必要となる。同時に、DNAと相互作用する分子、溶媒、イオンも考慮する必要が出てくる。

(iii) 実験データとの連携、機械学習

実験データからの構造モデリングにも MD 計算が活用される。クライオ電子顕微鏡を用いた単粒子解析により、溶液中での生体分子の立体像を近原子分解能で決定できるようになり、従来結晶化が難しかった膜タンパク質やスクレオソームなどの巨大生体分子複合体の立体構造が近年続々と解かれている。電顕像から生体分子の立体構造をモデリングしたり最適化したりする際、MD 計算に基づくフレキシブル・フィッティングが広く用いられ、NAMD, GROMACS, GENESIS などの MD 計算専用ソフトにもフレキシブル・フィッティングが搭載されるようになった。特に巨大生体分子複合体を対象とした場合、精度の高い計算だけでなく、高速な計算が要求されるため、粗視化モデルや拡張アンサンブル法を用いた方法 [17]、高速並列計算アルゴリズムの開発が行われている [18]。近年、クライオ電子線トモグラフィーを用いた細胞内のメソスケールな構造体を解析する実験も進み、オルガネラあるいは細胞スケールの分子システムを MD 計算でモデリングしようとする試みが始まっている [19]。また、近年盛んに生物物理学的研究で使われるようになっている高速 AFM の実験データからの構造モデリングの研究も進みつつある。粗視化モデルを用いた構造ダイナミクスシミュレーションを行うことで AFM の画像実験データから観測されている立体構造を推定する手法が提案されている [20, 21]。これまで AFM データの解像度や観測されている構造変化の大きさからモデリングも粗視化レベルで行われているが、今後、全原子 MD を用いたより高解像度のモデリングを行うことが望まれる。また更に、FRET などの 1 分子計測時系列データや、高速 AFM の動画とシミュレーショントラジェクトリをマルコフ状態モデルを介して融合(データ同化)させて構造の状態遷移を推定する技術も開発されており、機能発現に関わる重要な動態情報を与えると期待される。

(iv) 細胞シミュレーション

細胞シミュレーションについては、E-Cell [22]、Cell Illustrator [23]、米国産の Neuron [24] など、連続体的なアプローチに基づいた細胞内の反応シミュレータの開発が国内外で活発に進められてきた。細胞レベルの中核的な生化学機能は代謝、信号伝達、そして遺伝子発現の三つである。下記では、連続体細胞シミュレーションによる代謝系の計算および 1 分子粒度シミュレーションによる信号伝達経路のシミュレーションを中心に述べる。これら HPC への展開が比較的見通されている 2 方方に加え、今後の展望を考えるうえで取り上げるべき最近の顕著な動向として、スタンフォード大学が最近発表したマイコプラズマ菌や大腸菌の全ゲノムシミュレーションがある [25]。この全ゲノムシミュレーションでは、反応ネットワークモデルを採用しており、技術的にはわが国の E-Cell プロジェクトが 1999 年に発表した E-Cell 1 全細胞モデル [26] から大幅な飛躍は見られないが、ゲノムアノテーション(ゲノム中の遺伝子の機能などの注釈) やマルチオミクスデータ(遺伝子発現量などの網羅的データ) の蓄積が進んだことで E-Cell 1 全細胞モデルのように一部の遺伝子セットを選択するのではなく、特定のゲノムの全遺伝子を組み込んだモデルの構築に成功している点が特徴である。現状ではモデルにアドホックな部分が散見され、多数の経験的パラメータをどう決定していくかといった未解決の問題も多いが、今後継続的に改良を繰り返せば、多くの技術的な蓄積が進むと考えられる。統合的な細胞シミュレーションの実現は実用面でのインパクトが大きく、システム生物学の発祥の地の一つであるわが国としても長期的な取り組みを一層深く考えるべき時期に来ていると思われる。

一方、レーザー顕微鏡などを用いた高解像度でのライブイメージング技術が急速に発展したことにより、更に解像度を高めた1分子粒度での細胞シミュレーションへの取り組みが要請されている。1分子細胞シミュレータの代表的なものにChemCell（米サンディア国立研）[27]やMcell（米Salk Institute、コーネル大学、ピッツバーグスーパーコンピューティングセンター）[28]、Smoldyn（米国Fred Hutchinson Cancer Research CenterおよびThe Molecular Sciences Institute）[29]、GFRD（理研および蘭AMOLF）[30]、Spatiocyte（理研および慶應大学）[31]などがあるが、計算対象は大腸菌などの単純な原核細胞の信号伝達経路といった小規模かつ少数分子の系が多い。HPCへの展開はこれまで比較的ゆるやかであったが、1分子粒度の真核生物への展開には高並列化は不可避である。HPCI戦略プロジェクトにより、「京」を利用してSpatiocyteの大規模並列化が実装され[32]、理研がGFRDの並列版の開発を進めているほか、またSmoldynのGPUへの実装[33]などより大規模で複雑な系のシミュレーションの実現に向けて急速に展開している。長期的目標として、一分子粒度細胞シミュレーションにおいてはまず、巨大分子混雑を反映したより現実的な細胞環境のもとで、原核細胞よりも巨大な真核細胞（10μm以上）内の複雑な信号伝達経路モデルにを、転写因子が応答する時間スケールである1,000秒程度まで計算可能にすることが大きな課題となる。シミュレーション技術面では、分子構造シミュレーションの結果を取り入れ、分子の構造多型や重合脱重合化など、分子レベルの状態変化の表現が可能な計算手法の開発が期待される。更に、1細胞または少数の細胞のモデルから、多数の細胞から構成される集団および組織の計算が可能になると期待される。加えて、細胞の形、大きさの決定、また分化や発生過程のモデル構築のためには、化学反応と力学構造とのカップリングを取り込んだ計算が必要になり、この方面的研究開発が期待される。中長期的には遺伝子発現系や代謝系との連成により10,000秒以上の細胞周期を基準とする時間スケールでの予測性獲得が重要課題である。このためには、前述のE-Cell 1全細胞モデルやスタンフォード大学のマイコプラズマ/大腸菌モデルのようなゲノムスケールモデルとの連携、連成、また次に述べる連続体的アプローチによる代謝系モデルとの連携、連成が必要となる。

一方、レーザー顕微鏡などを用いた高解像度でのライブイメージング技術が急速に発展したことにより、更に解像度を高めた1分子粒度での細胞シミュレーションへの取り組みが要請されている。1分子細胞シミュレータの代表的なものにChemCell（米サンディア国立研）[27]やMcell（米Salk Institute、コーネル大学、ピッツバーグスーパーコンピューティングセンター）[28]、Smoldyn（米国Fred Hutchinson Cancer Research CenterおよびThe Molecular Sciences Institute）[29]、GFRD（理研および蘭AMOLF）[30]、Spatiocyte（理研および慶應大学）[31]などがあるが、計算対象は大腸菌などの単純な原核細胞の信号伝達経路といった小規模かつ少数分子の系が多い。HPCへの展開はこれまで比較的ゆるやかであったが、1分子粒度の真核生物への展開には高並列化は不可避である。HPCI戦略プロジェクトにより、「京」を利用してSpatiocyteの大規模並列化が実装され[32]、理研がGFRDの並列版の開発を進めているほか、またSmoldynのGPUへの実装[33]などより大規模で複雑な系のシミュレーションの実現に向けて急速に展開している。

(v) ゲノム解析・遺伝子ネットワーク解析

バイオインフォマティクス分野では、近年観測技術が大幅に進展し、次世代DNAシークエンサーによるDNA配列データやRNA配列データ解析技術による遺伝子発現データ、さらにはシングルセル解析技術や空間トランск립トーム計測技術などに代表されるような、今まで以上に短時間で大量に出力され

るデータ、いわゆるハイスループットデータが急速に蓄積されるようになってきている。米国 NCBI が運営するデータベースには 2019 年 10 月の時点で 240 万サンプルを超える [34]、2022 年 8 月の時点では 500 万サンプルを超える。

DNA シークエンシング技術を用いた遺伝子発現データだけでなく、近年ではプロテオームデータも網羅性を増しており、今後増大が見込まれる。またメチロームデータやメタボロームデータなど、生体分子ネットワークのさまざまな階層における「オミクス」データが網羅的に計測可能に計測可能になりつつある。これらの膨大なデータからの大規模な計算リソースを活用した生命科学、とりわけ医学や医療の発展が期待されている。特に、ヒトには 2 万数千の遺伝子があるとされ、そこから産生されるタンパクは 10 万種とも言われている。バイオインフォマティクスやシステム生物学と呼ばれる分野では、このような膨大な因子間の関係を大規模な観測データから統計科学や情報科学に基づき推定・予測する研究が盛んに行われている。遺伝子発現パターンの比較解析や、発現データに基づくクラスタリング、共発現解析や因果推論などの複雑な計算を必要とする遺伝子ネットワーク解析などがその代表的な例である。「京」を用いるグランドチャレンジアプリケーション研究開発プロジェクトや HPCI 戦略プログラム、その後の「富岳」重点課題においても、遺伝子発現データからの遺伝子ネットワーク推定・解析ソフトウェアの研究開発が進められており、ペタフロップス級の大規模計算による観測データ解析が当たり前になりつつある。膨大な計算リソースを前提とした新しいアルゴリズムも登場しており、「京」時代ではこれまで難しかった全遺伝子を含む遺伝子ネットワークの推定も現実となった。

また一方で、DNA 配列に代表されるゲノム配列解析では、ショットガン法による DNA 配列決定法の登場により大規模な計算機の利用が一般的になった。2008 年頃に登場した次世代 DNA シークエンサーではこれまで数年かかった全ゲノム解読を数日（最新のものでは 2 日以内）のレベルにまで短縮することが可能になり、個人一人一人が自分自身のゲノム配列を知りそれに基づく医療を受けることができる個人ゲノム時代がすでに実現している。次世代 DNA シークエンサーでは、1 回の計測で 1TB ほどのデータが発生し、そこから意味のあるデータに処理するために膨大な計算が必要となる。現在では、次世代シークエンサーを用いたシークエンスサービスを提供する専門の会社が多数設立され、研究者からの要望に応えている。また、試薬価格の低下とシークエンス速度向上の結果、米国 Broad Institute 等では全ゲノム解析がワークフローとして確立・採用されており、1 年でペタバイトを超える膨大なシークエンスデータを生成・蓄積している。医療の現場でも同様に、DRAGEN などの高速解析システムと合わせてクリニカルシークエンスが実施されている。HPCI 戰略プログラムにおいても次世代 DNA シークエンサーを用いたメタゲノムおよびがんゲノム解析が行われた。今後、解析サンプル数は急激に伸びることが予想されるため、これらのデータ解析に必要な計算機の導入が、今後の研究の進展に必要不可欠である。また個人の配列データが蓄積されるようになれば、全ゲノム関連解析（GWAS）もより大規模化・複雑化し、またゲノム配列データと発現データを組み合わせた解析もより複雑度を増していくことが容易に予想され、必要な計算リソースもそれに応じて膨大になると予想される。このように取り扱うデータが膨大であるため、生命科学におけるデータ解析での計算における一番のボトルネックはファイルシステムの性能となる。また一般に、多種多様なソフトウェアが組み合わせて利用され、また観測技術の進展・変化が非常に速いため、常に新しい計算・解析手法が提案されており一つのソフトウェアの寿命（利用期間）が短いのも特徴である。統計科学やシステム生物学における数値計算から、生物ネットワーク解析に必要なグラフ探索、DNA 配列解析における離散的な計算、あるいはそれぞれを複雑に組み合わせた計算など多種多様であり、

前述の次世代 DNA シークエンサー解析ソフトウェアもそのうちの一つに過ぎない。観測データから計算・解析手法を開発し、実際に計算し、計算結果を生物学的な実験にフィードバックし、最終的に生物学的な発見を論文としてまとめ上げるまでの時間の短縮が当該分野で求められていることである。逆に、すでに確立した一つの方程式をいかに高速に計算するか、という点は当分野ではあまり求められてはいない。また並列計算という点からは、一般にバイオインフォマティクスやシステム生物学ではデータ並列性があるため、計算の独立性が高く並列性能を出しやすい。したがって、複雑な並列アルゴリズムの構築を経なくても、数万コアから数十万コアに達するような超並列計算機の恩恵を即座に受けやすいと言える。一方、分割されたデータに対する処理時間に著しい開きがあることに注意が必要である。数百～数千倍の大きな処理時間差が生じることが頻繁に発生するため、解析全体での高性能化にはロードバランスをいかにとるかに留意する必要がある。

4.4.2 次世代に解決すべき課題

(i) 生体分子シミュレーション

全原子 MD では、構造生物学的に重要なタンパク質が次々と計算対象になるだろう。X 線結晶解析やクライオ電顕などで構造が決定されたタンパク質の機能を解析するために、全原子 MD を実行することがルーチンになっていくと思われる。また、AlphaFold 2 などの機械学習で予測された構造は MD 計算の始状態として十分な精度を持っている場合も多い。研究対象に従って、数多くの課題に対応する計算が行われることになるが、数万から数百万原子を含む小規模-中規模系が、多くのユーザが選択する研究対象となると思われる。そのなかでもやや大きめの数十万から百万原子規模は、分子モーターや膜輸送体などの分子機械において生体膜環境などを露わに取り込んだモデルに対応する。一方、数百万原子系のような大規模系では、単一のタンパク質というよりはリボソームに見られるようなタンパク質・核酸などを含む生体超分子複合体が主要な研究対象となるであろう。更に、細胞環境における生体膜だけでなく、細胞質やオルガネラなどの細胞内小器官も研究対象になるだろうが、そのような巨大系（数百万原子から 1 億原子系）のシミュレーションを行うためには、粗視化 MD などとの連成が必要となる。

(ii) 拡張アンサンブル法

MD 計算だけでなく、多くのアプリケーションが実行可能な汎用スパコンを用いて先端的な MD シミュレーションを行うためには、多くの CPU を並列に使って効率を上げるアルゴリズムが有効である。例えば、200 レプリカを用いた gREST/REUS 計算で、それぞれのレプリカを 1 マイクロ秒実行したならば、計算の総量として 200 マイクロ秒に到達する。これは「富岳」で既に実現している計算であり、それぞれの MD 計算を GPU などの汎用加速器等で高速化することでさらに短期間で計算を完了することができる。拡張アンサンブル法による高速化による効果を考慮すると、問題解決に必要な計算時間を遙かに短縮して高精度の結果を得ることができると期待される。一方で、多くのレプリカを用いて計算する拡張アンサンブル法ではシミュレーション後に得られるデータ量が莫大になり、解析するために必要な時間が増大する。MD シミュレーションに必要な 3 つのステップ（モデリング・シミュレーション・解析）の総時間として、研究の効率化を図るならばビッグデータ解析に必要な解析時間を短くすることも重要である。そのためには、MD 計算途中である程度のデータを計算する On-the-fly のアルゴリズムの導入や、

機械学習などを用いてトラジェクトリ解析の可能性を半自動的に調べることも考慮すべきだろう。

(iii) マルコフ状態遷移モデル

大量のシミュレーション群を統合して長時間ダイナミクスを推定できる MSM の特性を活かし、10,000 レプリカ系程度の超並列 MD シミュレーションと MSM を組み合わせた、ミリ秒オーダー構造変化の計算プロトコル開発に挑むべきである。この超並列サンプリング計算が実行できれば、單一ドメインタンパク質であれば数日、複数ドメインから成る全長タンパク質では数週間程度で、ミリ秒オーダーの構造変化を解析できると期待できる。これは、2022 年時点で、MD が最速の MD 専用スパコンでの性能や、大規模計算資源を使った SARS-CoV-2 研究での計算量に匹敵する。例えば、1 本の MD 実行が最も高速な専用スパコン Anton3 のベンチマークによると、64 ノード利用時に、單一ドメインタンパク質全原子溶液系（約 2 万～9 万原子系）で約 150～200 マイクロ秒/日のサンプリングが可能である [5]。Anton3 は 512 ノードなので、8 本の MD が並列に実行できたとすると、アグリゲーションで約 1,200～1,600 マイクロ秒/日のサンプリングが可能である。一方、富岳では、約 5 万原子溶液系の單一ドメインタンパク質のベンチマークを実施したところ、1 本の MD は約 100～200 ナノ秒/日程度であり、Anton3 の 64 ノード利用時の約 1/100 であった。並列に実行する MD の本数を考慮すると、富岳で標準的に利用可能な 4,000 ノードの約 500 レプリカ系で約 88 マイクロ秒/日、仮に、全ノード（約 15 万 9 千ノード）を利用して約 19,800 レプリカ系を実行できれば約 3,400 マイクロ秒/日の Anton3 全ノード利用時の約 2 倍のサンプリングが可能となる。従って、次世代スパコンでは、約 10,000 レプリカ系の計算が日常的に実行できれば、2022 年時点の Anton3 全ノード利用時と同程度のサンプリングが数日で可能となる。また、米国で実施された Folding@home を利用した SARS-CoV-2 タンパク質の大規模計算によると、約 5 万原子溶液系タンパク質では、アグリゲーションで約 8,900 マイクロ秒のサンプリングを実施している [35]。仮に、富岳で標準的に利用可能な 4,000 ノードの約 500 レプリカ系並列サンプリングで、同様のサンプリングを実施するには約 100 日かかる。もし、10,000 レプリカ系が実行できるプロトコルが実現できれば、約 5 日に短縮されるため、創薬サイクルに耐えうると期待される。また、クライオ電顕等の発展により、複数ドメインから成る全長タンパク質の立体構造が複数状態で解かれると予想される。Anton3 のベンチマークによると、複数ドメインタンパク質溶液系（約 32 万～約 200 万原子系）は 64 ノード利用時に約 25～約 90 マイクロ秒/日であり、全ノード利用時（512 ノード）で、約 90～160 マイクロ秒/日までスケールされる [5]。次世代スパコンでは、2022 年時点での Anton3 全台利用時の性能である、100 マイクロ秒/日程度のサンプリングを超並列 MD と MSM を組み合わせて目指すべきである。もし、100 マイクロ秒/日のサンプリングが可能となれば、米国の SARS-CoV-2 研究で実施された、複合体タンパク質（約 20 万原子溶液系）の約 4,000 マイクロ秒サンプリングや、Spike タンパク質三量体（約 44 万原子溶液系）の約 1,000 マイクロ秒サンプリングが、次世代スパコンでは数週間で実行可能となると期待される。

(iv) 粗視化 MD

粗視化 MD では、その高速性を生かして、細胞スケールの生命現象に挑むべきである。遺伝子動態、シグナル伝達系、膜・細胞骨格系、ウイルス動態などが挑戦的課題である。例えば、真核細胞における遺伝子の本体であるクロマチンの構造解析や、さまざまな転写因子の作用機序、DNA 複製機構など、構造情報に基づいて研究すべき課題が多い。そのうえで、老化やがんなどの疾患とかかわるエピジェネティク

スの理論計算は大きな挑戦である。シグナル伝達系は、分子構造に基づく MD と 1 分子粒度シミュレーションやシステム生物的な研究の接点になる可能性が高い。シグナル伝達を支えている足場タンパク質複合体などの構造基盤とその可塑性は、システム生物学に直接つながる粗視化 MD の課題となる。粗視化 MD における構造探索スピードの高速化などのメリットを生かしつつ、生命現象を詳細に理解するためには、粗視化・全原子連成アルゴリズムの開発が必要である。すなわち、粗視化モデルで得られたスナップショットからただちに（そして自動的に）全原子モデルを構築してすぐに高い解像度で計算を続ける必要がある。実際には、この変換を行う際に、多くの問題（キラリティエラー、リングペネットレーション、エンタングルメントなど）が生じる。これらを自動的に解決するモデリングツールが必要である。

これらの課題に対して MD そのものの計算能力向上が大きく寄与する事は疑いないが、実験技術の進歩によって得られる質の高い観測データとの連携を深めることも重要である。2.4 で述べた X 線自由電子レーザーに加え、近年は特にクライオ電子顕微鏡（Cryo-EM）が構造生物学に大きなブレークスルーをもたらしている。Cryo-EM を用いた実験では、分子スケールから細胞スケールまでの様々な空間解像度の画像データが大量に得られる。また、高速 AFM による分子運動のリアルタイムでの観測画像からの構造モデリングの手法も開発されつつあり、今後シミュレーションを活用してより大きなタンパク質複合体について高解像度のモデリングを可能にする必要がある。さらに SACLA を用いた時分割構造解析も技術の開発が進みより多くの反応中間体構造が観測できるようになると期待され、これらを活用したシミュレーションによる機能解析はタンパク質の機能発現に関わる新しい情報を与えると期待される。このようなビッグデータを迅速に解析するためには、高速な画像処理と構造モデリングの連成解析だけでなく、マルチスケールな解析法やソフトウェアプラットフォームの開発が重要になるだろう。他方、MD 計算からは座標のトラジェクトリーデータが大量に出力されるため、効率的な解析のためには、機械学習などの手法を用いて特徴的な変化を自動的に抽出するような新しい技術開発も必要となる。計算グループと実験グループが互いにビッグデータを効率的に共有し、試行計算や解析結果の議論を即座に行うためのインフラの整備も重要な課題となるだろう。生体分子が中間体をどのように遷移しているかを MD によって解明することも、機能を構造だけではなくダイナミクスから理解するうえで重要な課題となる。このような問題に対しては、実験で得られた各中間体構造から出発した複数の独立した MD によるトラジェクトリを、MSM（マルコフ状態モデル）を用いて連結する計算手法の開発が近年活発化している。

(v) 実験データとの連携

高速 AFM による分子運動のリアルタイムでの観測画像からの構造モデリングの手法も開発されつつあり、今後シミュレーションを活用してより大きなタンパク質複合体について高解像度のモデリングを可能にする必要がある。また SACLA を用いた時分割構造解析も技術の開発が進みより多くの反応中間体構造が観測できるようになると期待され、これらを活用したシミュレーションによる機能解析はタンパク質の機能発現に関わる新しい情報を与えると期待される。

(vi) 細胞シミュレーション

1 分子粒度細胞シミュレーションでは、1 細胞での 1 分子粒度での計算を実現、精密化することで、細胞の分化、癌化、増殖などの機序に信号伝達経路のダイナミクスからアプローチすることを目指す。計算を更に大規模化することにより、取り組む事が期待される問題として応答不均一性がある。ヒト細胞を用

いた1分子観察実験および免疫染色実験では、同一の遺伝情報を持つ均一な細胞集団に対して上皮成長因子や神経成長因子などの同一の因子を同一の条件で与えても細胞の応答には一定の不均一性が観察されることが確認されており、これが発生などの高次の生命現象に持つ意味が議論されている。このばらつきが生じる要因には、分子の構造多型によるとする説、信号伝達ネットワークにおける内因的ノイズの影響によるとする説、転写制御ネットワークなどの遺伝子発現系におけるノイズ、エピジェネティクスによるとする説など複数あり決着していないが、実際にはこれらの要因が複合的に働いている可能性が強い。このうち、構造多型や信号伝達系の内因的ノイズに関しては、現在の1分子粒度計算の大規模化でアプローチできるが、遺伝子発現系やエピジェネティクスなどの要因に迫るには、10,000秒以上の計算でも妥当性を失わない計算モデルが必要である。そのためには、長期的目標の項で述べたようにゲノムや代謝系シミュレーションとの連成が必要であるが、この方面では計算量の問題以前に大規模モデル構築における総合的、長期的な取り組みと技術の蓄積が必要とされ、長期目標の達成に至る筋道として微生物などの原核細胞などの比較的単純なモデル生物を用いた先行的なプロトタイプによって、精緻化を推し進めた1分子粒度シミュレーションと網羅性を目指した全細胞シミュレーションの融合を実現することが求められる。また、時間スケールの延長と同時に大規模なアンサンブル・シミュレーションにより、階層的なモデリングや予測精度の向上、また近年進展著しいAI・自動実験技術と連携した仮説生成への展開が期待される。

(vii) ゲノム解析・遺伝子ネットワーク解析

次世代DNAシークエンサーデータ解析

現在よく使われている型の次世代DNAシークエンサーは、DNA配列を50から数百塩基程度の長さに断片化し、それを同時に読み込むことによって超高速にDNA配列を読み取る装置である。したがって、得られるデータは大量の断片化されたDNA配列であり、ヒトの場合およそ2,000万個程度、ファイルサイズにして1TBになる。連続したDNA配列を得るためにこれらの断片を、次世代DNAシークエンサー解析ソフトウェアを用いて染色体ごとに1本につなげる必要がある。現在、主に行われている方法は、配列が既知である参照元となるDNA配列（いわゆるリファレンスゲノム）との部分に最も一致するか各断片1個1個について探索し一致させる計算である。その際、個人差や変異などにより完全に一致することはないため、ミスマッチを許す文字列探索法を用いている。次世代には個人個人が複数の組織のすべてのゲノムをシークエンスすることが当たり前になると予想され、それに必要な計算リソースもまた膨大となる。ゲノム解析では個人のDNA配列を得るだけでは、その個人が既知の多様性や変異を持っているかなどの診断は可能であるが、未知の原因までは見つけることはできない。したがってある程度まとまった量のデータが必要になり、そこから統計的解析を行う必要がある。特にがんゲノム解析では、がんはゲノムの異常であることから、がんに特異的な変異の同定だけでなく、DNAのコピー数解析なども非常に重要である。新しいシークエンス技術としてシングルセル解析[36]とロングリードシークエンスがある。シングルセル解析は文字通り、1細胞ごとにシークエンスする技術である。これまでのシークエンスでは、ターゲットとする組織（例えば腫瘍組織）から数百μg～数十mgの細胞塊を取得し、そこに含まれる細胞すべてのDNAを1つの集団として区別せずにシークエンスを行っていた。しかし、DNAは細胞ごとに異なる変異が入ること、また、腫瘍組織内にも正常細胞が入っていることなどから、細胞集団全体を一つとして取り扱う既存のシークエンス解析には精度上の限界が存在した。シングルセル解析では、読み取った塩基配列がどの細胞由来かを識別可能なため、1細胞ごとの変異の違いを確認することが

できるため、検出精度を格段に上げることが可能になる。

遺伝子ネットワーク解析

遺伝子ネットワーク推定ソフトウェアは、RNA シークエンシングなどによる遺伝子発現データから統計科学的・情報科学的な方法を用いて遺伝子間の制御の予測・推定を行うソフトウェアである。遺伝子ネットワーク推定の方法としてはさまざまなモデル・推定法が提案されているが、それぞれ一長一短があり、今後も解析対象・目的に合わせたさまざまな遺伝子ネットワーク推定法が現れると思われる。また前述のとおり、現在は「富岳」などの計算リソースの登場により、解析対象を絞らずに数万程度の全データセットに対して網羅的に遺伝子ネットワーク推定ソフトウェアを適用するというプロジェクトが進行中である。「京」時代は 2013 年頃に 256 データセット・30,517 アレイ（サンプル）での特定条件での限定的な網羅的計算を行っているが、「富岳」時代は全データセットを対象とすることが可能になった。これまでのデータの増加ペースを考えると 2032 年頃には 1,300～1,500 万サンプルほどになると予想される。また現在ではサンプル単位での遺伝子ネットワーク解析や、サンプルアノテーション技術などが登場しており、データセット横断的に新しいデータセットを再定義可能となり、また遺伝子ネットワークとサンプル単位での評価に任意の組み合わせが可能なことにより、恒常的に全データをこれらの観点から解析しようとすると 2032 年には現在の 100 倍程度の計算資源が必要となる。また解析手法が多様化することが予想されるため実際に必要な計算資源は更に大きくなる。公共データベースにある発現データをまるごと計算対象にすることが可能になれば、膨大な遺伝子ネットワークが蓄積可能となり、そこから更に解析・知識発見を行う、というもう一段上の研究ステージに上がることができる。

ゲノム解析・バイオインフォマティクスではこれまでの HPC での計算にない特徴的な部分がいくつかかる。ゲノム解析・バイオインフォマティクス分野で特異的に必要となる定性的スペックを挙げておきたい。当該分野ではこれまで主に PC クラスターを用いて解析が行われてきたが、「富岳」のようなフラグシップマシンでは十分な対応がなされていないのが現状である。これらの大部分は PC クラスターでは当たり前に実現されていることであり、将来のフラグシップシステムでも同等以上の利用環境になることが当該分野では必須事項である。当該分野の HPC における特徴や要望などを下記に列挙する。

- さまざまな既存アプリを利用するため、GCC を前提にしている一般的なオープンソースアプリをコンパイルし動作させる必要がある。Java/R/Python/Ruby などの高級言語、あるいはミドルウェアが高速で動作することが生命科学のデータ解析では求められている。あるいは、こういったミドルウェアレベルでの並列プログラミングサポートが必要である。また、一歩進んでスクリプト言語やミドルウェアレベルでの並列化のサポートが望ましい。
- 複数のアプリを組み合わせて順番に利用することがあり、ジョブ管理システムの API がスクリプト言語や C/C++ などから利用できることが望ましい。現状、計算ジョブを監視・制御する別の計算ジョブを走らせる、ということで実現しているがプログラミング言語レベルで対応されれば解析に必要な計算手順の構築が容易になり研究全体がより効率的になる。
- 前項のとおりさまざまな既存アプリが使われるため、よく使われるものについては、あらかじめシステムにインストールされていることが望ましい。例えばヒトゲノム解析センター（HGC）に導入されているスパコンには、バイオ系アプリに限っても 100 種のアプリがプリインストールされユーザに使われている。

- 計算の背景に膨大なデータを利用しており、高速なデータとのアクセスが可能であることが望ましい。特にバイオインフォマティクスでは、システム内に公開データベースのミラーデータがあり高速にアクセスできることが必要である。具体的には、NCBI のアレイデータおよび配列データなどである。HGC のスパコン SHIROKANE では、実際に各種データベースのミラーリングを提供している。
- データ解析では、ジョブ実行時間が長期間に及ぶこともある。伝統的な HPC の世界では長くても数日が最大のジョブ実行時間として運用されていると思われるが、SHIROKANE では数週間あるいは数ヶ月継続実行可能なキューが用意され、実際に使われている。
- ゲノム解析やバイオインフォマティクスでは、データ並列性が高いためアプリ自体は高並列化せずパラメータスイープ的な大量（具体的には数万から数十万）のアレイジョブによる並列実行が多い。
- ゲノム解析では DNA 配列という究極の個人情報を扱うためセキュリティにも配慮が必要である。

4.4.3 課題を解決するために必要な計算機性能

(i) 生体分子シミュレーション

MD 計算は、運動方程式を短い時間幅で数値積分することで、分子内の粒子の運動を計算する手法である。

全原子 MD では、最小単位は原子であり、対象は分子や原子の運動となる。演算性能が重視され、高いネットワーク性能（特に低レイテンシ）が必要となる。計算したい対象（タンパク質など）の多くは数万～百万原子数程度の規模であり、必要な計算量は $O(N \log N)$ または $O(N)$ であまり多くない。このため、決まった大きさの系に対して並列性能を上げる Strong Scaling 性能は、多くの場合中規模な並列化においてすら限界となっており、この向上と限界の見積もりが重要である。

粒子間にかかる相互作用は、近距離の粒子間に働く相互作用と、遠距離の粒子間に働く相互作用に分けられ、それぞれ計算コストのあり方がかなり異なる。近距離間相互作用は、空間的に系を分割し粒子を計算ノードに割り当てれば、隣接ノード間通信だけで計算に必要な情報はカバーできる。一方の遠距離間相互作用では、情報を大域的に共有する必要があるため、全体ノード間通信が発生する。遠距離間相互作用計算では、高速フーリエ変換 (FFT) を利用する計算アルゴリズムが普及している [37]。さらに遠方の相互作用力の数値積分に対してより長い時間幅で計算する reversible reference system propagator algorithms (RESPA) [38] を用いることで遠距離相互作用の計算頻度を落とす計算が広く利用されている。さらに、水素原子と共有結合する重原子の質量を変化（水素原子の質量を増し、その分の質量を重原子から差し引く）する Hydrogen Mass Repartitioning [39] の利用や、改良された数値積分法 [40, 41, 42, 43] を組みわせて利用することにより、現在では近距離相互作用の数値積分では 3.5 fs、遠距離相互作用では 7 fs の時間幅を利用することも可能となっている。

分子レベルでの重要な生命現象の時間スケールは、マイクロ秒から秒ぐらいであるため、単体のシミュレーションで到達する場合、例えば 1 日に 100 マイクロ秒、10 日で 1 ミリ秒の計算性能が必要である。この性能は、2020 年に米国で開発された MD 専用機である Anton3 の 512 ノードで約 100 万原子 (121.1 マイクロ秒/日)-220 万原子 (93.2 マイクロ秒/日) の計算に相当する。[5] ここからも、全原子 MD シミュ

レーションを単体の長時間実行する場合は、2030年においてもMD専用機が現実的であると考える。

ただし、実際の生物中では、分子は多数存在しアンサンブルとしてその機能を発揮しているので、一分子だけでなく、多数分子の統計量（自由エネルギー等）を計算すべきという考えがある。それに基づいた手法としては、対象系の複数のレプリカを計算機中に置き、それぞれのレプリカでシミュレーションを行う手法が存在する。幅広いサンプリングを実現するレプリカ交換MD[44, 45]や、比較的短い時間のMDシミュレーションを大量に実行し、その後状態間遷移確率を求めることで長時間時間ダイナミクスを再構築するマルコフ状態モデルが代表例である。これらの方法では、複数の条件（温度・反応座標・力場パラメータ・構造）を組み合わせた計算を行った後、ターゲットとなる条件を理論的に再構築することで単独のシミュレーションでは不可能なサンプリングを達成できるという利点がある。また、計算機科学からの視点においても、各レプリカは疎結合、もしくは、完全に平行計算のため、近年のスーパーコンピュータで行われる大規模並列計算に適しているという特徴がある。2022年現在では、ターゲットとなる生体分子（1-数個）と周囲の溶媒分子（水、膜分子など）からなる50万原子までのシミュレーションが主流だが、細胞質環境やスクレオソーム等の大規模システムの計算も多くなってきている。2022年では10億原子のシミュレーションが最大である[16, 46]。本稿では中規模システム（約158万原子）と大規模システム（約10億原子）での見積もりを考える。

中規模システムの計算では、現在はレプリカ交換MDやMSMを用いた場合は、各システムでのシミュレーション長はサブマイクロ秒が主流となっている。しかし、2マイクロ秒として考える。また、レプリカ・シミュレーション数は手法や原子数、集団座標によって異なる。最も多いMSMではシミュレーション数は数100のオーダーであり、シミュレーション数は500を考える。（シミュレーション全体で1ミリ秒となる）中規模システムの性能はGENESISのウェブサイトで提示されているベンチマーク[47]から、富岳1,024ノード(49,152コア)で1日約0.13マイクロ秒である。現状では15.4日で2マイクロ秒の計算が可能である。富岳と同規模のコア数(7,630,848コア)を持つ計算機だと、500回のシミュレーションは全系計算を4(3.22)回分に相当し、全部を計算するまでに約61.6日かかる。これを10日で行うには、速度向上×コア数増加が6以上になる必要がある。（ただし、コア数の增加分は3倍以下である必要がある。）コア数が2倍となった場合、MD実行速度を3倍以上の向上が必要である。そこでは、浮動小数点演算性能の大幅な向上が必要であるが、それは同時に演算時のデータ（メモリ、キャッシュ、レジスタ）のロード・ストア性能の向上が必須となっている。富岳におけるGENESISの短距離相互作用力の計算の性能プロファイルを見ると、演算やキャッシュロードの待ち時間が6割を占める事からもわかる[46]。更に富岳では高SIMD計算を行うと待ち時間が長くなり、むしろ実行時間が長くなる現象が散見される。これはレジスタ数やキャッシュの不足であり、コンパイラやアルゴリズムの改善では限界がある。ハードウェアで十分な性能を与える必要があると考える。

次世代先端的計算基盤に関する白書からは、GPU等の加速器について述べられている。現在ほとんどのMDソフトウェアはGPUに対応しており、高い計算速度の向上にはGPUや加速器の導入が望ましいと考える。一方で、メモリ転送やインターフェクトの問題はより深刻となる。GromacsのGPU利用のアルゴリズムについての論文でも同じ問題提起がなされている[48]。そのためにはPCI/eによる通信ではなく、特別なインターフェクトにより加速器とホスト間の転送が十分に早いことは必須である。

ストレージは、10ピコ秒ごとに単精度の座標履歴をとったとして、158万原子の1ミリ秒の計算では1.9PB必要である。一つのシミュレーションでは6.7秒に1回19MBしか出力されないため、単純な

速度としては現状の SSD の速度で十分である。ただし、100-500 のシミュレーションから同一のファイルシステムに同時にデータの出力が起きても安定して実行されることがより重要である。

10 億原子を超える大規模システムの計算では、単体の全原子モデルだけでミリ秒の過程を追うことは現実的ではない。しかし、分子間の相互作用や構造、集団運動を分子論的に理解するためには、解像度が高いモデルの計算が重要である。粗視化 MD から得られた特徴的な構造の全原子モデル化、粗視化 MD のパラメータ精製の参照データ取得を目的とした実行が想定される。2020 年ではこのサイズのシミュレーションは数ナノ秒程度であるが、1 マイクロ秒の計算について見積もる。大規模システムの性能は GENESIS のウェブサイトで提示されているベンチマーク [47] から、富岳 16,386 ノード (786,528 コア) で 1 日約 0.0012 マイクロ秒である。84 日かかることになる。これを 10 日程度で終わらせるには、8.5 倍の性能向上が必要となる。例えば 3 倍のコア数が利用できるとしても 3 倍程度の実行性能の向上が必要である。

メモリは、空間分割スキームを用いると 1 ノード当たりの必要量は減らすことができる。10 億原子システムは富岳 4,096 ノードから計算が可能であり、1 ノード 11 GB 程度を利用しているので、つまり全体で 0.05 PB 程度あれば計算可能である。ただし、このようなサイズのシミュレーションでは、セットアップ時に原子や力場パラメータを空間分割スキームに適用、各プロセスに分配するために膨大なメモリ (> 1 TB) を必要とする。現在の GENESIS では、セットアップ部分を別途 1 TB 以上のメモリを持つ計算機で行っていることは注意されたい。

ストレージは、10 ピコ秒ごとに单精度の座標履歴をとったとして、10 億原子の 1 マイクロ秒の計算では 1.2 PB 必要である。10 ピコ計算するのに数十秒かかるので、単純な書き込み速度としては現状の SSD の速度で十分である。

更に原子数が増えて、21 億を超える原子数を扱う場合には、原子のインデックスを 4 バイト整数では表すことが出来ず、ソフトウェアの大きな改変が必要になる可能性が高い。

粗視化 MD でも、運動方程式を数値積分する点では全原子 MD と共通である。一方で、粗視化モデルでは、溶媒を利用しないなど、分子種により大胆な縮約を行うため、演算量を減らすことができる。更に原子内や分子内の振動・運動を考慮しないと、時間幅を大きく取ることができ、その点でも演算量を大幅に減らすことができる。しかし、溶媒を利用しない粗視化モデルでは、空間による粒子の密度差が大きく、全原子 MD で用いられる空間分割はロードインバランスが大きくなり、並列性能は悪くなる。また、時間幅が大きいことは、一回の数値積分による粒子の運動量が大きいことにもつながり、粒子の密度が時間を進めると大きく変化し、ダイナミックロードバランシング等を必要となり並列スキームは更に難しいものとなる。前述した様に、粗視化モデルの構築・修正には実験データや全原子モデルとの連携が必要となる。近年は機械学習や AI フレームワークが汎用化されており、既存モジュールの利用はもちろん、新規手法の開発もフレームワーク上で行われることが想定される。そのため、python や Julia を用いた既存フレームワークを高速に動くシステムであることが必要である。また、高速で安定なデータ処理、例えば数 100-1,000 プロセスから同一データファイルへの同時読み込みがあっても極端な遅延やシステムの問題が起きない堅固なシステムであることが求められる。また、様々な外部クラウドストレージとのネットワークで接続する必要があり、データ転送の高速化が重要となる。

電頭実験データからフレキシブル・フィッティングを用いてリボソーム級の生体分子複合体やそれらの集合体構造を精密にモデリングするためには、通常の大規模系に対する MD 計算 (100 万-1 億原子系) に

加えて、実験データへのバイアスの計算、および、複数のレプリカを用いた拡張アンサンブル法に基づく効率の良い構造サンプリングが必要である。実験データが画像の場合、それらは膨大なピクセル数（数千万以上）、あるいは立体像の場合は膨大なボクセル数（数十億以上）から構成されるため、バイアスの計算にも大きなコストを必要とする。実際、電顕バイアスの計算は、PME の逆空間の計算と同程度の計算時間を必要とする。したがって、このような 2 次元あるいは 3 次元データを高速に解析するためには単純な CPU 計算ではなく GPU 計算が必要である。また、レプリカを用いた計算では、1 レプリカ内での演算性能と高いネットワーク性能が求められる。現在、富岳を用いてリボソーム級 100 万原子系のフレキシブル・フィッティングを行うと、十分な構造サンプリングとして 1 マイクロ秒の計算に約 20 日を要するが、GPU を利用した計算ができればバイアス計算がそのまま GPU に乗ることになり、同じ計算を 10 日に短縮できる。構造生物学においては迅速な解析が必要であり、MD 計算に基づく実験画像データ解析には、CPU と GPU のハイブリッド計算がより重要性を増すと考えられる。また、同様のモデリングのための計算は今後 AFM や XFEL による実験データについても行うことが必要になってくる。

(ii) AI・機械学習とシミュレーション連携による生体分子ダイナミクス・機能予測

概要と計算手法: 生体分子の分子動力学シミュレーションで得られた構造を学習データとし、ニューラルネットワークによって学習・予測することを想定する。学習・予測することによって、ニューラルネットベースのベイズ最適化等のシミュレーションと AI・機械学習の高度な連携を目的とする。問題規模として学習時間等が既知である AlphaFold 2 の学習と同程度の問題を考える（構造数 10 万個）。ただし、ベイズ最適化等への応用にあたってはこの学習を一度だけでなく何回も適用・アップデートすることを想定している。問題規模: 10 万生体分子構造のシミュレーションデータとそのニューラルネットによる学習備考: AlphaFold 2(2 つの attention 機構付きの transformer) では 10 万超の PDB のタンパク質構造データと配列データを扱って、配列データから構造データの予測に関する損失関数の最小化を行っている。学習においては、TPU を 128 個用いて実時間としておよそ 11 日間かかっている（予測にかかる時間はほとんど無視できる）。AlphaFold 2 と同様のネットワーク構造を用いて学習を行うことを想定すると、モデルの調整等を考えてその 10 倍程度は必要と予想される。更に、TPU が GPU と比べて大まかに 10-30 倍程度高速だとすれば、ニューラルネットの一回の学習で GPU を 2,000 台程度用いて 11 日間程度の資源が必要だと見積もられる。更に、ベイズ最適化等により分子動力学シミュレーションと連携させてニューラルネットの学習を繰り返し行う場合には、この 10 倍から 100 倍の資源が必要と見積られる。ニューラルネットの学習を実行するには、ソフトウェア面としては既存のフレームワーク（PyTorch や JAX ベースのフレームワーク）が性能を損なわずにそのまま動くことと、ハード面としては加速器（TPU や GPU）が動くことが重要な点となる。

(iii) 細胞シミュレーション

一分子粒度細胞シミュレーションでは、格子法と粒子法の 2 種類の計算手法がよく用いられる。どちらも基礎方程式として Langevin 方程式の過減衰極限である Smoluchowski の反応拡散方程式を用いる。ペタ世代では、大規模並列化で格子法が先行しており、主なアプリケーションとなっている。格子法は格子気体法を拡張したアルゴリズム（例えば Spatiocyte 法）が用いられ、0.1 から 1 PF の能力があれば、1 細胞でのレセプターへのリガンド結合から転写制御因子活性化までの一連の応答反応が計算できる。1

EFであれば、1,000から10,000細胞で構成される細胞集団、あるいは組織における応答不均一性の計算が行える。粒子法では、通常のブラウン動力学ではステップ幅がナノ秒を切り、生理学的な時スケールである1,000秒に到達するためには最低10の12乗のステップ数が必要であり、実用的ではない。このため、グリーン関数を用いたステップ幅の伸長(GFRD:Green's Function Reaction Dynamics法)や経験的パラメータによる近似解法(Smoldyn法)などの工夫が用いられる。近似を用いない正確な解法として現在最も性能がよい改良GFRD法(eGFRD法)はイベント駆動型であるため高並列化が困難であり、1,000分子程度の系までが適用限界であったが。これに対し、実用的な百万分子の系の計算を目標に、階層的時ステップを用いて高並列化を試みる新世代の手法であるpGFRD法の開発が理研で進んでいる。この分野のハードウェアへの要求の特徴として、まず第一に、格子法、粒子法の双方において反応ネットワークや境界条件などが複雑であり、イベント処理や例外処理が多いことがある。このため、SIMDも含めたベクトル系の命令の利用が十分に性能に反映されない場合が多く、整数系の性能が高いほうが有利である。特に、離散事象法においてはIntel Haswellアーキテクチャ以降一般的になると考えられるハードウェアトランザクショナルメモリーの有効性およびそのSimultaneous Multi-Threadingとの親和性が指摘されている。第二に、長時間のダイナミクスを対象とするため、必要なステップ数が多く、ネットワークのレイテンシが容易に律速となることがある。以下に、具体的な例を挙げての計算機要求の見積もりを示す。まず、格子法で10,000細胞のシミュレーションを10ケース程度行う場合を考える。細胞の体積が10 pL (10^{-12} L)、含まれるタンパク質すべての濃度が $10\mu M$ とすると、1細胞当たりおよそ109個のタンパク質分子が存在する。格子法における典型的な時ステップサイズは $1\mu s$ であり、3,600秒のシミュレーションを行う場合、総ステップ数は 3.6×10^9 となる。1粒子当たりの演算量はおよそ10 FLOPであり、その時の総演算量は、 3.6×10^5 EFLOPとなり、これを10ケース行う。10日間でシミュレーションする場合、必要な演算性能は、0.42 EFLOPSと見積もられる。同様に、粒子法で、1細胞の信号伝達経路のシミュレーションを行う場合を考える。粒子法での典型的な時ステップサイズは、1ナノ秒であり、3,600秒のシミュレーションを行う場合、総時ステップ数は、 3.6×10^{11} となる。1粒子当たりの演算量はおよそ1,000 FLOPであり、総演算量は、 3.6×10^5 EFLOPとなり、これを100ケース行う。格子法と同様に、10日間でシミュレーションする場合、必要な演算性能は0.42 EFLOPSと見積もられる。

(iv) ゲノム解析・バイオインフォマティクス

次世代DNAシークエンサーデータ解析

次世代シークエンサー解析に必要な計算スペックを以下に示す。DNAシークエンサー技術の今後についてはかなり流動的であるため一部あいまいな表現を残している。例えば想定しているのは解析対象となる人数であり、一人当たり何種の細胞のサンプルが取得できるかはコストや技術的な制約により可変であること、シングルセル解析、ロングリードシークエンス等の発展途上の技術に関しては技術の進展の予測が困難であることから、ある程度幅を持たせる意味で明示していない。2032年頃に1か月で1,000,000人分規模の解析(例えば、がんゲノム解析1,000,000人分のマッピングおよび変異同定)が必要であると想定して必要スペックを記す。なお、ここに挙げた計算スペックの見積は計算資源の総和として必要な物であり、10拠点ほどに分割されていても問題ない。また、複数台のシークエンサーからのデータを独立に処理可能であることから、1システムで実現が必要なスペックは導入されるシークエンサーの台数分で分割可能である。効率的に解析を行うためには、ゲノム解析に用いられる計算機はDNAシークエンサーに併

設されている必要がある。また、必要演算性能については、ゲノム解析では整数演算が中心であるため、FLOPS ではなく Operation 数で記述している。

【ネットワークバンド幅、レイテンシ】 2032 年:~2 TB/s (シークエンサー=ストレージ間の総和) (100 台のシークエンサーで分散させることを仮定すると~12 GB/s/台)

データ並列性があるため、計算機の計算ノード間の通信バンド幅およびレイテンシはあまり重要ではない。その代わりに、シークエンサー=ストレージ間のネットワーク帯域がボトルネックの一つとなり得る。1 検体分データ容量の想定は 5 TB であり、想定している 1,000,000 人分のデータを 30 日間 (2,592,000 秒) でシークエンサーから転送することを考えると最大で $5 \text{ TB} \times 1,000,000 = 5 \text{ EB}$ の転送が必要とされ、必要な帯域は実効でそれぞれ 1.9 TB/s 必要である。これは総計であるので、10 拠点に分散されれば必要な帯域はこの 1/10 となる。また 100 台のシークエンサーに分散されれば、1 台当たりは 1/100 である 19 GB/s ほどになる。実際には、シークエンスを専門業者に外注することも多く、その場合、データの受け取り方を含めた転送方法を十分に検討する必要がある。

【1 ケースあたりの総メモリ容量】 2032 年:1.46~5.84 PB

Genomon (ヒトゲノム解析プログラム) で最もメモリ消費の大きいマッピングアルゴリズムについて見積もる。1,000 万リード程度を 1 プロセスに割り当てることが実用的なインプットサイズで、その場合の容量は 8~16 GB の範囲に収まる。1 サンプルが 6 億リードと仮定した場合、 $6 \text{ 億}/1000 \text{ 万} \times 8 \text{ GB} = 4.8 \sim 9.6 \text{ TB}$ がサンプル当たりの必要メモリ量となる。ここから、スループットの想定も合わせて、スパコン全体での必要メモリ総量を見積もる。100 万検体/月であるならば、3 万検体/日の処理量になる。1 検体の解析時間が 24 時間ならば 3 万倍であるが、数ターンの解析が回ると考えるのが妥当である。3 ターン (1 時間 × 24 セット) で想定すると 1,250 倍になるので、 $(4.8 \sim 9.6) \text{ TB} \times 1,250 = 6 \sim 12 \text{ PB}$ が、1 ケースあたりの総メモリ容量となる。

【ストレージ容量】 2020 年:2~20 EB

単純に 1 人当たりの必要データ量に想定される人数を掛け合った値である。例えば、第 4 世代機による 1 サンプル 100 TB の場合、200,000 人では 20 EB 必要となる。

【ストレージ速度】 2032 年:5 EB (分割処理可能)

現行システムにおける 1 ファイル 80 MB の入力データに対するプロファイル結果は IO 量が 11.204 GB である。これを 1 サンプル 5 TB に外挿すると IO 量は 700,250 GB となる。想定している 2032 年での 1,000,000 人 (サンプル) 規模の解析を 30 日間で行うと仮定すると 1 サンプルの実行時間は 2.592 秒となる。以上より、必要な IO 性能は 270 TB/s となる。2015 年頃のデータ量はこの 1/10 であるので単純に IO 性能も 1/10 である。またこれは入力データが分割可能なことから同一システムで実現されなくとも総計として必要な IO 性能であって、システム (解析拠点) の数で分割可能である。

【CPU 速度】 Genomon での OPS 測定結果がないため、富岳での値を確認中。

遺伝子ネットワーク推定

遺伝子ネットワーク推定については、現在「富岳」Society5.0 推進利用課題として「「富岳」を基軸とした創薬 DX プラットフォーム構築」プロジェクトが採択され、そのプロジェクトにおいておよそ 20,000

データセット・2,000,000 サンプルを想定した網羅的遺伝子ネットワーク推定が SiGN-BN (INGOR) NNSR アルゴリズムを用いて進められている。このプロジェクトでは 1 つのデータセットに対して富岳を用いて 200NH 程度の計算時間を想定している。FLOP 換算すると 2.2EFLOP である。また富岳の 0.5% 程度の計算リソースを恒常に使用しこれを 2 年間で完了することを予定している。これは CPU 速度に換算すると 2.1PF となり、10 年前に想定していた値とほぼ等しい結果となっている。2032 年には年間 10,000 データセットの新規データ及び、サンプルの組み合わせによるデータセット定義により現在の 100 倍規模の解析、つまり 3,000,000 ケース程度の規模を想定すると、以下のような計算スペックが 1 ケースあたり必要となる。

【CPU 速度】 2032 年:~220 PFLOPS

2032 年には現在の 100 倍の規模である 1,000,000 データセットの計算を想定しており、1 セット分の計算を分割可能な 1 ケースとする。通常 1 データセットあたり 3 回の計算を仮定しているため 3,000,000 ケースを 1 年間で処理すると仮定すると 1 秒あたり 0.1 ケースの処理が必要である。1 ケース 2.2 EF の計算が必要であるため、恒常に必要な計算性能は 220 PFLOPS である。

【メモリ量】 2032 年:~1.0 PiB

現状 1 ケースあたり 300 ノード程度を利用している。したがって 1 ケースあたり $32 \text{ GiB} \times 300 = 10 \text{ TiB}$ である。単純に 100 倍のメモリを計算時に使用すると仮定する。従って $10 \text{ TiB} \times 100 = 1000 \text{ TiB} = 1.0 \text{ PB}$ となる。

【ネットワーク性能】 2032 年:~110PB/s

アルゴリズム自体は変更がないため過去の想定から求める。プロファイル結果から求めた一番多い計算ノードでの通信量は 897,000,000 B でこの計算は 120 秒であった。従っておよそ 1 GB を 120 秒で通信しているため 1 ノードあたり 8 MB/s 程度である。実データではこの 100 倍程度を仮定する。これを 140,000 ノード程度での計算を仮定すると $800 \text{ MB/s} \times 140,000 = 112,000,000 \text{ MB/s} = 112 \text{ TB/s}$ で十分と思われる。そこからさらに将来 10 倍の複雑度になると仮定しても $112 \text{ TB/s} \times 10 = 1.1 \text{ PB/s}$ である。2032 年にはさらにその 100 倍の想定となるため 110 PB/s となる。

【ストレージ容量】 2032 年:~1 TB

現行の網羅的ネットワーク推定で想定している 1 ケースあたりのストレージサイズは 300 MB の想定である。全ケースを想定した総ストレージ容量は 1 PB 程度である。計算が複雑化したとしても 1 ケースあたりでは 1 TB で十分といえる。

4.4.4 公的な意見交換とそのサマリー

2012 年 9 月 23 日の日本生物物理学会におけるシンポジウム「スーパーコンピューティング: 分子ネットワークと細胞内ダイナミクス」や、2012 年 10 月 9 日の計算物質科学イニシアティブ (CMSI) 計算分子科学研究拠点第 3 回研究会においては、生命分野のサイエンスロードマップの紹介が行われた。

2012 年 10 月 19 日のバイオスーパーコンピューティング研究会において、生命分野のサイエンスロードマップの紹介および意見交換が行われた。そこでは、放射光施設や X 線自由電子レーザーなどの大型

実験施設との連携を考慮して進めるようにとの意見が産業界から出された。また、医学分野との連携をより強く行うように医学界から意見が出された。

2013年6月12日～14日の第13回日本蛋白質科学会年会にて、「エクサフロップス時代の計算蛋白質科学」と題した生命科学・物質科学・ものづくりの各分野と連携したワークショップが開催され、意見交換が行われた。

2013年6月27日に、製薬企業関係者が多数集まる情報計算化学生物（CBI）学会の研究講演会にて、計算創薬の展望についての意見交換が行われた。放射光施設など大型実験施設との連携や、更なる並列数增加への対応、ものづくり分野との連携についての意見が出された。

2014年1月23日～24日のバイオスーパーコンピューティング研究会「ウィンタースクール2014」において、バイオスーパーコンピューティングにおけるポスト「京」の展望について意見交換が行われた。

2016年11月8日の理研シンポジウム「クライオ電子顕微鏡が拓く未来—構造生物学研究のパラダイムシフト—」において、MDをはじめとする計算科学との連携について意見交換が行われた。

2019年6月13日の「個別化医療・創薬基盤に衝撃-スペコン「富岳」が起こす地殻変動-Personalized and Preventive Medicine, and Drug Discovery Infrastructureにおいて、生命科学におけるポスト「京」の利活用に関して意見交換が行われた。

参考文献

- [1] Ron O. Dror, Albert C. Pan, Daniel H. Arlow, David W. Borhani, Paul Maragakis, Yibing Shan, Huafeng Xu, and David E. Shaw. Pathway and mechanism of drug binding to g-protein-coupled receptors. *Proceedings of the National Academy of Sciences*, Vol. 108, No. 32, pp. 13118–13123, 2011.
- [2] Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, and David E. Shaw. How fast-folding proteins fold. *Science*, Vol. 334, No. 6055, pp. 517–520, 2011.
- [3] Gongpu Zhao, Juan R. Perilla, Ernest L. Yufenyuy, Xin Meng, Bo Chen, Jiying Ning, Jinwoo Ahn, Angela M. Gronenborn, Klaus Schulten, Christopher Aiken, and Peijun Zhang. Mature hiv-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, Vol. 497, No. 7451, pp. 643–646, 2013.
- [4] David E. Shaw, Martin M. Deneroff, Ron O. Dror, Jeffrey S. Kuskin, Richard H. Larson, John K. Salmon, Cliff Young, Brannon Batson, Kevin J. Bowers, Jack C. Chao, Michael P. Eastwood, Joseph Gagliardo, J. P. Grossman, C. Richard Ho, Douglas J. Ierardi, István Kolossváry, John L. Klepeis, Timothy Layman, Christine McLeavey, Mark A. Moraes, Rolf Mueller, Edward C. Priest, Yibing Shan, Jochen Spengler, Michael Theobald, Brian Towles, and Stanley C. Wang. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM*, Vol. 51, No. 7, p. 91–97, 2008.
- [5] David E. Shaw, Peter J. Adams, Asaph Azaria, Joseph A. Bank, Brannon Batson, Alistair Bell, Michael Bergdorf, Jhanvi Bhatt, J. Adam Butts, Timothy Correia, Robert M. Dirks, Ron O. Dror, Michael P. Eastwood, Bruce Edwards, Amos Even, Peter Feldmann, Michael

- Fenn, Christopher H. Fenton, Anthony Forte, Joseph Gagliardo, Gennette Gill, Maria Gorlatova, Brian Greskamp, J. P. Grossman, Justin Gullingsrud, Anissa Harper, William Hasenplaugh, Mark Heily, Benjamin Colin Heshmat, Jeremy Hunt, Douglas J. Ierardi, Lev Iserovich, Bryan L. Jackson, Nick P. Johnson, Mollie M. Kirk, John L. Klepeis, Jeffrey S. Kuskin, Kenneth M. Mackenzie, Roy J. Mader, Richard McGowen, Adam McLaughlin, Mark A. Moraes, Mohamed H. Nasr, Lawrence J. Nociolo, Lief O'Donnell, Andrew Parker, Jon L. Peticolas, Goran Pocina, Cristian Predescu, Terry Quan, John K. Salmon, Carl Schwink, Keun Sup Shim, Naseer Siddique, Jochen Spengler, Tamas Szalay, Raymond Tabladillo, Reinhard Tartler, Andrew G. Taube, Michael Theobald, Brian Towles, William Vick, Stanley C. Wang, Michael Wazlowski, Madeleine J. Weingarten, John M. Williams, and Kevin A. Yuh. Anton 3: twenty microseconds of molecular dynamics simulation before lunch, 2021.
- [6] Michael Feig, Ryuhei Harada, Takaharu Mori, Isseki Yu, Koichi Takahashi, and Yuji Sugita. Complete atomistic model of a bacterial cytoplasm for integrating physics, biochemistry, and systems biology. *Journal of Molecular Graphics and Modelling*, Vol. 58, pp. 1–9, 2015.
 - [7] J. Jung, T. Mori, C. Kobayashi, Y. Matsunaga, T. Yoda, M. Feig, and Y. Sugita. Genesis: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, Vol. 5, No. 4, pp. 310–323, 2015.
 - [8] Isseki Yu, Takaharu Mori, Tadashi Ando, Ryuhei Harada, Jaewoon Jung, Yuji Sugita, and Michael Feig. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife*, Vol. 5, p. e19274, 2016.
 - [9] J. Jung, W. Nishima, M. Daniels, G. Bascom, C. Kobayashi, A. Adedoyin, M. Wall, A. Lappala, D. Phillips, W. Fischer, C. S. Tung, T. Schlick, Y. Sugita, and K. Y. Sanbonmatsu. Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations. *Journal of Computational Chemistry*, Vol. 40, No. 21, pp. 1919–1930, 2019. Id6pn Times Cited:3 Cited References Count:34.
 - [10] Abhishek Singharoy, Christopher Maffeo, Karelia H. Delgado-Magnero, David J. K. Swainsbury, Melih Sener, Ulrich Kleinekathöfer, John W. Vant, Jonathan Nguyen, Andrew Hitchcock, Barry Isralewitz, Ivan Teo, Danielle E. Chandler, John E. Stone, James C. Phillips, Taras V. Pogorelov, M. Ilaria Mallus, Christophe Chipot, Zaida Luthey-Schulten, D. Peter Tieleman, C. Neil Hunter, Emad Tajkhorshid, Aleksei Aksimentiev, and Klaus Schulten. Atoms to phenotypes: Molecular design principles of cellular energy metabolism. *Cell*, Vol. 179, No. 5, pp. 1098–1111.e23, 2019. doi: 10.1016/j.cell.2019.10.021.
 - [11] Mateusz Sikora, Sören von Bülow, Florian E. C. Blanc, Michael Gecht, Roberto Covino, and Gerhard Hummer. Computational epitope map of sars-cov-2 spike protein. *PLOS Computational Biology*, Vol. 17, No. 4, p. e1008790, 2021.
 - [12] Gregory A. Voth. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. CRC Press., 2008.

- [13] Siewert J. Marrink, Alex H. de Vries, and Alan E. Mark. Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, Vol. 108, No. 2, pp. 750–760, 2004. doi: 10.1021/jp036508g.
- [14] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The martini force field: Coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, Vol. 111, No. 27, pp. 7812–7824, 2007. doi: 10.1021/jp071097f.
- [15] Josh V. Vermaas, Christopher G. Mayne, Eric Shinn, and Emad Tajkhorshid. Assembly and analysis of cell-scale membrane envelopes. *Journal of Chemical Information and Modeling*, Vol. 62, No. 3, pp. 602–617, 2022. doi: 10.1021/acs.jcim.1c01050.
- [16] Abigail Dommer, Lorenzo Casalino, Fiona Kearns, Mia Rosenfeld, Nicholas Wauer, Surl-Hee Ahn, John Russo, Sofia Oliveira, Clare Morris, Anthony Bogetti, Anda Trifan, Alexander Brace, Terra Sztain, Austin Clyde, Heng Ma, Chakra Chennubhotla, Hyungro Lee, Matteo Turilli, Syma Khalid, Teresa Tamayo-Mendoza, Matthew Welborn, Anders Christensen, Daniel G. A. Smith, Zhioran Qiao, Sai Krishna Sirumalla, Michael O' Connor, Frederick Manby, Anima Anandkumar, David Hardy, James Phillips, Abraham Stern, Josh Romero, David Clark, Mitchell Dorrell, Tom Maiden, Lei Huang, John McCalpin, Christopher Woods, Alan Gray, Matt Williams, Bryan Barker, Harinda Rajapaksha, Richard Pitts, Tom Gibbs, John Stone, Daniel Zuckerman, Adrian Mulholland, Thomas Miller, Shantenu Jha, Arvind Ramanathan, Lillian Chong, Rommie Amaro. #covidisairborne: Ai-enabled multiscale computational microscopy of delta sars-cov-2 in a respiratory aerosol. *bioRxiv*, p. 2021.11.12.468428, 2021.
- [17] Osamu Miyashita, Chigusa Kobayashi, Takaharu Mori, Yuji Sugita, and Florence Tama. Flexible fitting to cryo-em density map using ensemble molecular dynamics simulations. *Journal of Computational Chemistry*, Vol. 38, No. 16, pp. 1447–1461, 2017.
- [18] Takaharu Mori, Marta Kulik, Osamu Miyashita, Jaewoon Jung, Florence Tama, and Yuji Sugita. Acceleration of cryo-em flexible fitting for large biomolecular systems by efficient space partitioning. *Structure*, Vol. 27, No. 1, pp. 161–174.e3, 2019.
- [19] Zane R. Thornburg, David M. Bianchi, Troy A. Brier, Benjamin R. Gilbert, Tyler M. Earnest, Marcelo C. R. Melo, Nataliya Safranova, James P. Sáenz, András T. Cook, Kim S. Wise, Clyde A. Hutchison, Hamilton O. Smith, John I. Glass, and Zaida Luthey-Schulten. Fundamental behaviors emerge from simulations of a living minimal cell. *Cell*, Vol. 185, No. 2, pp. 345–360.e28, 2022.
- [20] B. Dasgupta, O. Miyashita, T. Uchihashi, and F. Tama. Reconstruction of three-dimensional conformations of bacterial clpb from high-speed atomic-force-microscopy images. *Front Mol Biosci*, Vol. 8, p. 704274, 2021. Dasgupta, Bhaskar Miyashita, Osamu Uchihashi, Takayuki Tama, Florence eng Switzerland 2021/08/24 Front Mol Biosci. 2021 Aug 5;8:704274. doi: 10.3389/fmolb.2021.704274. eCollection 2021.
- [21] S. Fuchigami and S. Takada. Inferring conformational state of myosin motor in an atomic force microscopy image via flexible fitting molecular simulations. *Front Mol Biosci*, Vol. 9, p. 882989,

2022. Fuchigami, Sotaro Takada, Shoji eng Switzerland 2022/05/17 Front Mol Biosci. 2022 Apr 29;9:882989. doi: 10.3389/fmolsb.2022.882989. eCollection 2022.
- [22] E-Cell. <http://www.e-cell.org/>.
- [23] Cell Illustrator. <http://www.cellillustrator.com/>.
- [24] Neuron. <http://neuron.duke.edu/>.
- [25] J. R. Karr, et al. A whole-cell computational model predicts phenotype from genotype. *Cell*, Vol. 150, No. 2, pp. 389–401, 2012.
- [26] M. Tomita, et al. E-CELL: software environment for whole-cell simulation. *Bioinformatics*, Vol. 15, No. 1, p. 72, 1999.
- [27] ChemCell. <http://chemcell.sandia.gov/>.
- [28] Mcell. <http://www.mcell.cnl.salk.edu/>.
- [29] Smoldyn. <http://www.smoldyn.org/>.
- [30] GFRD. <http://gfrd.org/>.
- [31] Spatiocyte. <http://spatiocyte.org/>.
- [32] S.N. Arjunan, A. Miyauchi, K. Iwamoto, and et al. pspatiocyte: a high-performance simulator for intracellular reaction-diffusion systems. *BMC Bioinformatics*, Vol. 21, No. 33, 2020.
- [33] L. Dematte. Smoldyn on graphics processing units: Massively parallel brownian dynamics simulations. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 9, No. 3, pp. 655–667, 2012.
- [34] https://ftp.ncbi.nlm.nih.gov/pub/factsheets/ASHG2019/ASHG2019_PlatformTalk_dbGaP_Freq.pdf.
- [35] Maxwell I. Zimmerman, Justin R. Porter, Michael D. Ward, Sukrit Singh, Neha Vithani, Arthur Meller, Upasana L. Mallimadugula, Catherine E. Kuhn, Jonathan H. Borowsky, Rafal P. Wiewiora, Matthew F. D. Hurley, Aoife M. Harbison, Carl A. Fogarty, Joseph E. Coffland, Elisa Fadda, Vincent A. Voelz, John D. Chodera, and Gregory R. Bowman. Sars-cov-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. *Nature Chemistry*, Vol. 13, No. 7, pp. 651–659, 2021.
- [36] Method of the year 2019: Single-cell multimodal omics. *Nat Methods*, Vol. 17, No. 1, 2020.
- [37] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An $n \log(n)$ method for ewald sums in large systems. *The Journal of Chemical Physics*, Vol. 98, No. 12, pp. 10089–10092, 1993.
- [38] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, Vol. 97, No. 3, pp. 1990–2001, 1992.
- [39] K. Anton Feenstra, Berk Hess, and Herman J. C. Berendsen. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *Journal of Computational Chemistry*, Vol. 20, No. 8, pp. 786–798, 1999.
- [40] J. Jung, C. Kobayashi, and Y. Sugita. Kinetic energy definition in velocity verlet integration for accurate pressure evaluation. *Journal of Chemical Physics*, Vol. 148, No. 16, 2018. Ge5vn

Times Cited:0 Cited References Count:28.

- [41] J. Jung, C. Kobayashi, and Y. Sugita. Optimal temperature evaluation in molecular dynamics simulations with a large time step. *Journal of Chemical Theory and Computation*, Vol. 15, No. 1, pp. 84–94, 2019. Hh2oo Times Cited:2 Cited References Count:42.
- [42] Jaewoon Jung and Yuji Sugita. Group-based evaluation of temperature and pressure for molecular dynamics simulation with a large time step. *The Journal of Chemical Physics*, Vol. 153, No. 23, p. 234115, 2020.
- [43] Jaewoon Jung, Kento Kasahara, Chigusa Kobayashi, Hiraku Oshima, Takaharu Mori, and Yuji Sugita. Optimized hydrogen mass repartitioning scheme combined with accurate temperature-/pressure evaluations for thermodynamic and kinetic properties of biological systems. *Journal of Chemical Theory and Computation*, Vol. 17, No. 8, pp. 5312–5321, 2021.
- [44] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, Vol. 314, No. 1, pp. 141–151, 1999.
- [45] Yuji Sugita, Akio Kitao, and Yuko Okamoto. Multidimensional replica-exchange method for free-energy calculations. *Journal of Chemical Physics*, Vol. 113, No. 15, pp. 6042–6051, 2000.
- [46] Jaewoon Jung, Chigusa Kobayashi, Kento Kasahara, Cheng Tan, Akiyoshi Kuroda, Kazuo Minami, Shigeru Ishiduki, Tatsuo Nishiki, Hikaru Inoue, Yutaka Ishikawa, Michael Feig, and Yuji Sugita. New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems. *Journal of Computational Chemistry*, Vol. 42, No. 4, pp. 231–241, 2021.
- [47] RIKEN. Genesis benchmark. https://www.r-ccs.riken.jp/labs/cbrt/benchmark-2020/#Input_and_control_files_of_the_benchmark.
- [48] Szilárd Páll, Artem Zhmurov, Paul Bauer, Mark Abraham, Magnus Lundborg, Alan Gray, Berk Hess, and Erik Lindahl. Heterogeneous parallelization and acceleration of molecular dynamics simulations in gromacs. *The Journal of Chemical Physics*, Vol. 153, No. 13, p. 134110, 2020.

4.5 脳・神経科学

本節では脳・神経科学分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.5.1 脳・神経回路シミュレーション

本章では、

- 脳神経回路シミュレーションの概要
- 脳部位・ニューロンモデルの違いによる計算量・計算手法の違い
- アクセラレータや専用ハードウェアの利用
- 脳神経回路シミュレーションにおける代表的なアプリケーション
- 計算量の見積もり

を説明する。

4.5.1.1 脳・神経回路シミュレーションの概要

神経回路シミュレーションはプログラム 4.1 で行われる [1]。

ソースコード 4.1 神経回路シミュレーションの擬似コード

```

1 for t : 時間にに関するループ {
2   for i : ニューロンに関するループ {
3     ニューロンiへのシナップス入力の計算();
4   }
5   for i : ニューロンに関するループ {
6     ニューロンi の膜電位の計算();
7     ニューロンi が発火したかどうかの判定();
8     スパイクの伝播();
9   }
10 }
```

ニューロン毎に、(1) シナップス入力の計算、(2) 膜電位の計算、(3) 発火判定、(4) スパイク伝播、を順番に行う。

(1) シナップス入力の計算

ニューロンは別のニューロンからシナップスという構造を介して信号を受け取る。入力を送るニューロン（シナップス前細胞）の軸索終末が、入力を受け取るニューロン（シナップス後細胞）の樹状突起に結合^{*6}し、シナップス前細胞からは化学物質が分泌され、シナップス後細胞の受容体がそれを受け取る。

^{*6} 実際には、物理的に結合しているのではなく隙間がある（シナップス間隙）。

前者の計算は具体的には式 (4.19) に従う。時刻 t におけるニューロン i へのシナプス入力は、

$$I_{\text{syn}i}(t) = g_{\text{syn}i}(t)(v_i(t) - E_{\text{syn}}) \quad (4.19)$$

と書ける。ここで $I_{\text{syn}i}(t)$ はシナプス入力によって流れる電流、 $g_{\text{syn}i}(t)$ はシナプスコンダクタンス、 $v_i(t)$ は膜電位、 E_{syn} は反転電位と呼ばれる定数である。興奮性シナプスの場合 E_{syn} は通常 0 mV、抑制性シナプスの場合は $-75 \sim -80$ mV の値が用いられる。シナプスコンダクタンスは以下のように計算される。

$$g_{\text{syn}i}(t) = \bar{g}_{\text{syn}} \sum_j w_{ij} \sum_{f_j} \kappa(t - t^{(f_j)}) \Theta(t - t^{(f_j)}) \quad (4.20)$$

ここで \bar{g}_{syn} はピークコンダクタンス、ここで j はニューロン i に対するプレ側のニューロンの添字、 w_{ij} はシナプス結合の強度、 f_j はニューロン j のスパイク、 $t^{(f_j)}$ はそのスパイクがニューロン i に到達した時刻、 $\kappa(t)$ はスパイクを電位に変換するためのカーネル、 $\Theta(t)$ はヘビサイド関数である。 $\kappa(t)$ は最も簡単な場合は時定数 τ を持つ指数関数:

$$\kappa(t) = \exp(t/\tau) \quad (4.21)$$

が用いられる。より一般的には、2つの指数関数を用いて立ち上がりと減衰を記述する

$$\kappa(t) = \exp(t/\tau_{\text{decay}}) - \exp(t/\tau_{\text{rise}}), \quad (\tau_{\text{rise}} < \tau_{\text{decay}}) \quad (4.22)$$

や、 $\tau_{\text{rise}} = \tau_{\text{decay}} = \tau$ とした

$$\kappa(t) = (t/\tau) \exp(t/\tau) \quad (4.23)$$

が用いられる。

(2) 膜電位の計算

外部からの入力を得て、各ニューロンの状態を表す膜電位をニューロン毎に独立に計算する。モデル化の抽象度に応じて様々な記述が可能であるが、代表的なものを以下に記載する。

(a) Hodgkin-Huxley モデル

Hodgkin-Huxley モデルはヤリイカの巨大軸索のデータにもとづいて作成された最初のニューロンモデルであり、イオン電流の流れを具体的に記述したモデルである。式 (4.24) の 4 式からなる常微分方程式に従う。

$$\begin{aligned} C \frac{dV}{dt} &= -\bar{g}_{\text{leak}}(V(t) - E_{\text{leak}}) - \bar{g}_{\text{Na}}m^3h(V(t) - E_{\text{Na}}) - \bar{g}_{\text{K}}n^4(V(t) - E_{\text{K}}) + I(t) \\ \frac{dm}{dt} &= \alpha_m(V(t))(1 - m(t)) - \beta_m(V(t))m(t) \\ \frac{dh}{dt} &= \alpha_h(V(t))(1 - h(t)) - \beta_h(V(t))h(t) \\ \frac{dn}{dt} &= \alpha_n(V(t))(1 - n(t)) - \beta_n(V(t))n(t) \end{aligned} \quad (4.24)$$

ここで、 C は膜のキャパシタンス、 \bar{g}_{xxx} はイオン電流のコンダクタンス、 E_{xxx} は反転電位、 $I(t)$ はシナプス入力を含む外部入力の電流、 m, n, h はイオンの流入・流出を制御するゲート変数である。

ヤリイカの巨大軸索でのパラメータは $C = 1\mu F/cm^2$, $g_{\text{leak}} = 0.3mS/cm^2$, $E_{\text{leak}} = 10.6mV$, $g_{\text{Na}} = 120mS/cm^2$, $E_{\text{Na}} = 115mV$, $g_{\text{K}} = 36mS/cm^2$, $E_{\text{K}} = -12mV$ である [1]。 $\alpha_{xxx}(t)$, $\beta_{xxx}(t)$ はそれぞれゲートの開閉を表す式であり、同様にヤリイカの巨大軸索の場合は

$$\begin{aligned}\alpha_m(v) &= \frac{2.5 - 0.1v}{\exp(2.5 - 0.1v) - 1} \\ \beta_m(v) &= 4 \exp(-v/18) \\ \alpha_h(v) &= 0.07 \exp(-v/20) \\ \beta_h(v) &= \frac{1}{\exp(3 - 0.1v) + 1} \\ \alpha_n(v) &= \frac{0.1 - 0.01v}{\exp(1 - 0.1v) - 1} \\ \beta_n(v) &= 0.125 \exp(-v/80)\end{aligned}\tag{4.25}$$

である [1]。

Hodgkin-Huxley モデルはスパイクを生成するためのイオン電流である Na, K 電流を記述しており、スパイクの波形を忠実に再現することが可能である。スパイクの生成は 1 ミリ秒程度の非常に短時間で行われるため、時間刻みを数十マイクロ秒程度に小さくする必要がある。

一方、スパイク生成以外の部分については膜電位の変化は比較的遅いため、1 ミリ秒程度の時間刻みでもうまく動作する場合が多い。このため、時間刻みを動的に調整する手法が用いられる。あるいは、次に記載する積分発火型モデルを利用する。

(b) 積分発火型モデル

Hodgkin-Huxley モデルはスパイク生成を再現するため、時間刻みを小さく取る必要がある。一方、スパイク生成以外の部分については膜電位の変化は比較的遅いため、1 ミリ秒程度の時間刻みでもうまく動作する場合が多い。このため、スパイクが生成される詳細な機構と言うよりは、例えば大規模ネットワークにおいてスパイクがどのように交換されネットワークとして機能するのかといった問題を取り扱う場合は、スパイク生成の部分を簡略化した積分発火型モデルが用いられることが多い。

最も簡単な積分発火型モデルは以下の式で記述される。

$$\begin{aligned}C \frac{dv}{dt} &= -g_{\text{leak}}(v(t) - E_{\text{leak}}) + I(t) \\ v(t) > \theta &\Rightarrow \text{スパイク発射}, v(t) = v_{\text{reset}}\end{aligned}\tag{4.26}$$

ここで、Hodgkin-Huxley 方程式と同様に、 C は膜のキャパシタンス、 g_{leak} はリークコンダクタンス、 $v(t)$ は時刻 t での膜電位、 E_{leak} はリークの反転電位、 $I(t)$ は外部入力であり、新たに閾値と呼ばれる定数 θ 、リセット電位と呼ばれる定数 v_{reset} を導入する。膜電位が閾値を超えると、その時刻でニューロンはスパイクを発射したと見なし、かつ膜電位をリセット電位に強制的にセットする。これにより、スパイク発射とその後の過分極を模擬する。

両辺を g_{leak} で割り $\tau = C/g_{\text{leak}}$ と置くと、元の式を

$$\tau \frac{dv}{dt} = -v(t) + E_{\text{leak}} + I(t)/g_{\text{leak}}\tag{4.27}$$

と書くことができる。これは時定数が τ である通常の一次遅れ系に対応する。

最も簡単な積分発火型モデルは、一定の外部入力に対しては一定の間隔でスパイクを発射し続けるだけのダイナミクスしか持たないが、実際のニューロンは様々な種類のイオンチャネルや Ca^{2+} 依存型のチャネルを持ち、非常に複雑な発火パターンを示すことが知られている。Hodgkin-Huxley モデルでもそのようなチャネルを導入することで再現可能である。積分発火型モデルでは、閾値を時間依存にする、微分方程式ではなく積分方程式に書き直し適当な非線形カーネルを導入する (Spike Response Model, SRM) 等のアドホックな手法により再現可能である。一方、Hodgkin-Huxley モデルよりは簡単であるが、単一ニューロンの多様なダイナミクスを自然に再現できるモデルとして、次の Izhikevich モデル [2] が存在する。

(c) Izhikevich モデル

Izhikevich モデル [2] は膜電位の他にもう一つ内部パラメータを持つ、以下の式で記述される。

$$\begin{aligned} C \frac{dv}{dt} &= k(v(t) - v_r)(v(t) - v_t) - u(t) + I(t) \\ \frac{du}{dt} &= a(b(v(t) - v_r) - u(t)) \\ v(t) \leq v_{\text{peak}} &\Rightarrow v(t) = c, \quad u(t) = u(t) + d \end{aligned} \quad (4.28)$$

ここで、 $u(t)$ が内部パラメータであり、時刻 t での値を表す。 k, a, b, c, d は定数、 $v_r, v_t, v_{\text{peak}}$ も定数である。 v_r, v_t が膜電位の上限・下限を設定し、 v_{peak} が閾値を設定する。 k, a, b, c, d を変更すると様々なダイナミクスを再現することができる (図 4.13)。

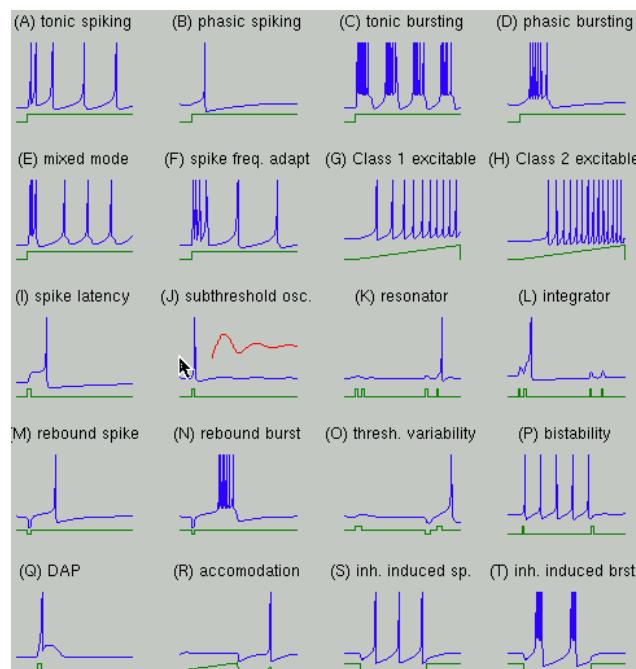


図 4.13 Izhikevich モデルのダイナミクスの例。画像は Simulation Platform (http://sim.neuroinf.jp/modules/xoonips/detail.php?item_id=67) から入手した。

(d) マルチコンパートメントモデル

上述の Hodgkin-Huxley モデル、積分発火型モデル、Izhikevich モデルは全て、ニューロンの形状を無視し、仮想的な点と考える。一方、実際のニューロンは複雑に広がる樹状突起や長距離を伸展する軸索など、特徴的な空間形状を有している。このような空間形状を考慮したモデルをマルチコンパートメントモデルと呼ぶ [3]。

マルチコンパートメントモデルは、(1) ニューロンの形状を十分短く電気的に等価なケーブルに分割し (コンパートメント)、(2) 隣接するコンパートメントを抵抗で接続したモデルである (図 4.14)。必要に応じて各コンパートメントにイオンチャネルやシナプスを配置する。例えば樹状突起には興奮性シナプスを、細胞体には抑制性シナプスを、軸索にはスパイク生成に必要な Na, K チャネルを割り当てる。

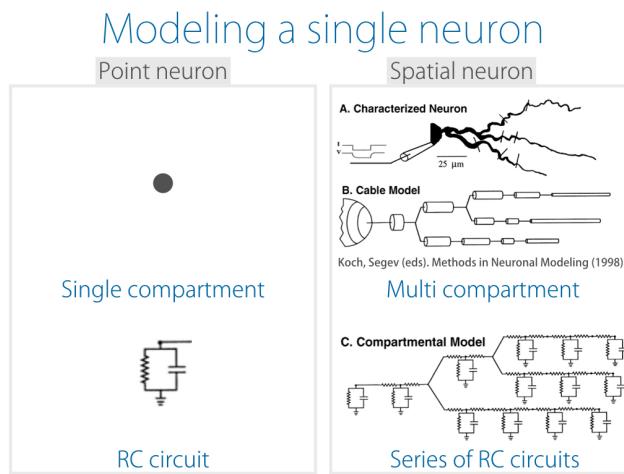


図 4.14 点モデル (左) とマルチコンパートメントモデル (右) の差異。

定式化は以下のように行う。電気的に均一な一次元のケーブルを考えると、ケーブル上の点 x での電位は以下の拡散方程式 (ケーブル方程式) で記述される。

$$C \frac{\partial v}{\partial t} = -g_{\text{leak}} (v(t, x) - E_{\text{leak}}) + \frac{a}{2R} \frac{\partial^2 v}{\partial x^2} + I(t, x) \quad (4.29)$$

ここで C は膜のキャパシタンス、 g_{leak} はリークコンダクタンス、 $v(t, x)$ は時刻 t 、位置 x における膜電位、 E_{leak} 、 a はケーブルの直径、 R は長軸方向の抵抗、 $I(t, x)$ はその場所への外部電流である。時間について後退差分、空間について 2 次の中心差分を取ると、

$$-\sigma v(t+\Delta t, x-\Delta x) + (1+2\sigma+\gamma)v(t+\Delta t, x) - \sigma v(t+\Delta t, x+\Delta x) - g_{\text{leak}}E_{\text{leak}} = -v(t, x) + I(t, x) \quad (4.30)$$

となる。ここで $\sigma = a\Delta t / 2RC(\Delta x)^2$ 、 $\gamma = g_{\text{leak}}\Delta t / C$ である。これを陰解法で求解していく。分岐がある場合も同様の定式化を行えば良い。

(3) ニューロンの発火の判定

膜電位の値がある閾値に達したことをもって、そのニューロンは発火したと考える。

(4) スパイク伝播

ニューロンが発射したスパイクは、通常 1 ミリから数十ミリ秒の遅延を持って他のニューロンに伝播

する。通常膜電位の計算に用いる時間刻みはそれより短いため、ニューロンが発火したことを持って直ちにスパイクを伝播しなくてもよく、その遅延に応じた時間スケールでまとめてスパイクを伝播すればよい。このように複数の時間スケールを考慮することで、通信と膜電位の計算をオーバーラップし、計算時間を削減することが可能となる。

マルチコンパートメントモデルは、細胞内でどのように電流が流れるのかをシミュレートするときに主に用いられる。例えば、樹上突起上の異なる場所についているシナプスの電流がどのように統合され細胞体へと流れるのかといった詳細なシミュレーションを行う際に必要となる。

点モデルは細胞内の詳細なダイナミクスというよりは、細胞同士がシナプスを介して結合し、ネットワークとしてどのように振る舞うのかをシミュレートする際に主に用いられる。常微分方程式の数値計算は通常のルンゲクッタ法等を用いて行われる。高度な数値解法が用いられる場合は稀である。

マルチコンパートメントモデルでネットワークを構成することも当然可能であるが、計算量が膨大になる。特に、偏微分方程式を解くことになるため、この場合は数値不安定性を回避するために陰解法、例えば陰的オイラー法やクランクニコルソン法が用いられる。その際に連立方程式を解く必要があり、NEURON [4] や GENESIS [5] はガウスの消去法を用いている。ニューロンの計算では係数行列は必ず正定値対称かつ疎行列になるため、新興のシミュレータでは共役勾配法を用いるものもある。

4.5.1.2 アクセラレータや専用ハードウェアの利用

グラフィクスプロセッシングユニット (GPU) は計算を加速するためのデバイスであり、神経回路シミュレーションにおいても利用されている。この分野のスタンダードな神経回路シミュレータである NEURON[4] は実験的に GPU をサポートしており、新興の神経回路シミュレータはいずれも GPU サポートをうたっている [6]。個々のニューロンを GPU の個々のスレッドに割り当て、スレッド毎にニューロンの計算を行う並列化を行っている。

汎用のシミュレータだけでなく、単一の脳部位のシミュレーションをおこなうためのプログラムも GPU で加速されている。例えば大脳基底核 [7] や小脳 [8, 9] のモデルは GPU で加速され、解剖学・生理学に比較的忠実なシミュレーションがリアルタイム^{*7}で実行可能である。後者は複数枚の GPU を利用し、100 万ニューロンからなる非常に大規模なシミュレーションを実現している。

Field-Programmable Gate Array (FPGA) も同様に用いられている。一般的なスパイキングネットワークのシミュレーション [10] や、小脳の神経回路モデルの実装 [11] がある。

PEZY Computing 社が開発している PEZY-SC プロセッサも利用されている。理化学研究所 情報基盤センターの「菖蒲」(図 4.15) は 320 ノード 1280 PEZY-SC チップからなる液浸冷却のスペーコンであるが、全ノードの 5 分の 4、1008 チップを利用して、10 億ニューロンからなる小脳神経回路のリアルタイムシミュレーションが既にできている [12, 13]。これはネコ一匹分の全小脳に相当する規模である。

^{*7} ここでは 1 秒間分の神経活動のシミュレーションが現実世界の 1 秒以内で完了することと定義する。



図 4.15 理研情報基盤センターの「菖蒲」。画像は http://www.riken.jp/pr/topics/2016/20160620_2/ より取得。

神経回路シミュレーションを行うための専用チップ(ニューロモルフィックチップ)の開発も進められている。

SpiNNaker [14] はマンチェスター大が開発しているニューロチップである(図 4.16)。1 個の SpiNNaker チップは 18 個の ARM968 コア、128MB の SDRAM、プロセッサ間でデータを交換するためのルータで構成されており、各コアには 64KB のデータメモリが搭載されている。このチップを 48 個並べて 1 枚の SpiNNaker ボードを構成し、さらに SpiNNaker ボードを多数並べて SpiNNaker システムを構成する。ボードのネットワークインターフェースは 100Mb/s の Ethernet である。1 SpiNNaker コアで約千個のスパイキングニューロン、約百万個のシナプスからなる神経回路をシミュレートできる。SpiNNaker ボードを 1200 枚からなる 19 インチラック 10 台で構成したシステムでは約 10 億ニューロンのシミュレーションができる、その際の消費電力は約 75KW である。プログラムは PyNN [15] で実装する。ニューロンモデルは Izhkevich モデルまたは積分発火型モデルが利用できる。空間形状を持つニューロンモデルは利用できない。

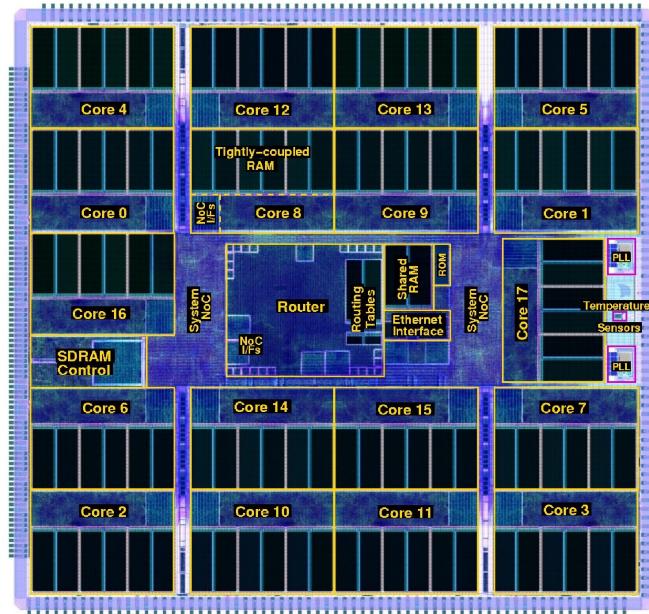


図 4.16 SpiNNaker チップ。画像は <http://apt.cs.manchester.ac.uk/projects/SpiNNaker/SpiNNchip/> より取得。

TrueNorth [16] は IBM が開発しているニューロチップである。1 チップは 4096 個のコアからなり、各コアには約 10KB のローカルメモリを備えている。1 コアで 256 個のスパイキングニューロンと 65536 個のシナプスをシミュレートし、コア間でスパイクをルーティングする。ニューロンモデルは積分発火型モデルを単純化したものであり、シナプスの計算も単純化されている。時間ステップは 1 ミリ秒であり 1kHz で動作する。TrueNorth の本質的な制約として、シナプス結合を動的に変更できない点ある。つまり学習ができない。学習は別のシステムで行う必要がある。

4.5.1.3 脳・神経回路シミュレーションにおける代表的なアプリケーション

代表的なアプリケーションとして、(1) シングルコンパートメントモデルのためのシミュレータ NEST、(2) マルチコンパートメントモデルのためのシミュレータ NEURON、(3) 小脳神経回路に特化したシミュレータ Realtime Cerebellum、をそれぞれ紹介する。

1. NEST [17]

ドイツのユーリッヒ研究所のマーカス・ディーズマン氏とアビゲイル・モリソン氏らの主導するグループによって、1993 年から開発が行われている神経回路シミュレータ。デスクトップコンピューターからスーパーコンピュータまで利用できる神経回路シミュレーターで、ソースコードが公開されていて、誰でも自由に利用することができる。数十種類の神経細胞モデルやシナプスモデルが利用できるようになっており、ユーザーによるモデルの拡張にも対応している。自動並列化に対応しており、MPI プロセス並列と、OpenMP、pthread などのスレッド並列を利用可能で、それらを組み合わせたハイブリッド並列にも対応している。神経細胞やシナプスのデータ構造のメモ

リ消費効率化がなされており、京コンピュータ上で NEST を用いて、17 億神経細胞と、10 兆個のシナプス結合からなる、大規模実行な皮質神経回路モデルの 1 秒分のシミュレーションを、約 40 分で実現した実績を持っている。

2. NEURON [4]

NEURON はデューク大学 Michael Hines, John W. Moore, and Ted Carnevale らによって 1980 年代に開発が始まり、現在も拡張されつづけている神経細胞シミュレータである。その核は形態をもつ神経細胞内の電気伝導を cable 方程式で計算することであり、MODL と呼ばれる形式でチャネルの動態を記述できるようになることで汎用的な神経シミュレータになった。ソースコードは公開されていて、膨大な量のイオンチャネルモデル、神経回路モデル、神経細胞モデルが ModelDB などで公開されている。2005 年より Blue BRAIN Project が始まって以来いくつかの並列化実装が試みられて次第に大規模並列が可能になってきており、日本で開発された京向けのチューニング済みの実装、NEURON K+ [18] では 20 万コアを用いた 80 万ニューロンの神経回路に対して 187TFLOPS の性能が達成されている。

3. Realtime Cerebellum [8]

アクセラレータを利用する代表的なアプリとして、上述の人工小脳を挙げる。CPU 版、マルチ GPU 版、FPGA 版、PEZY-SC 版と様々なバージョンが存在しており、アクセラレータの種類によらず加速に成功している。特に最新の PEZY-SC 版では理研情報基盤センターの「菖蒲」を利用し、252 ノード、1008 PEZY-SC チップを用いて 10 億ニューロンからなる人工小脳のリアルタイムシミュレーションに成功している。その際の実行効率は 2.6%、約 70 TFlops となっている。

4.5.1.4 計算量の見積もり

脳神経系シミュレーションにおいては、まず長期的な目標として、詳細な神経細胞モデルを用いたヒトの全脳規模の神経回路シミュレーションについての必要な性能の見積もりを行う。次に、2020 年以降に実現可能な二つのモデルとしての単純な神経細胞モデルを用いたヒトの全脳規模の神経回路シミュレーションと、詳細なモデルを用いた昆虫（カイコガ）脳の全脳規模の神経回路シミュレーションについて性能の見積もりを行う。ヒトの脳の場合、大脳皮質と小脳が 99% の神経細胞数を占め、全脳シミュレーションに必要な計算リソースについても、大脳皮質と小脳がそのほとんどを占める。大脳皮質と小脳皮質では、神経回路構造、神経細胞の性質、結合数などで大きな違いがあり、シミュレーションにおける要求性能が異なる。そこで、全脳規模のシミュレーションのより正確なプロファイルを取るため、大脳皮質と小脳の評価をそれぞれ行い、それらをもとに全脳規模のシミュレーションの見積もりを行う。

単純な神経細胞モデルとシナプスモデル

単純モデルとして神経細胞モデルに Leaky Integrate-and-Fire (LIF) neuron モデルを用いた場合を考える。LIF は、1 ステップで 5 FLOP [19] かかる。シナプスには指数減衰型のモデルを用いて、1 ステップで減衰の計算に更に 1FLOP かかる。シナプス可塑性は詳細なモデルと同様とする。

単純なモデルを用いた、ヒトの全脳規模の神経回路シミュレーションでは、神経細胞数やシナプス数などの設定は詳細なモデルを用いたヒトの全脳規模モデルと同様で、神経細胞とシナプスモデルに関

する演算とデータ量が異なる。生物学的 1 秒分の計算に必要な演算量は、大脳皮質は 0.5PFLOP、小脳は 2PFLOP、全脳の演算量は 2.5PFLOP となる。BF 値は～1.、ネットワークバンド幅の BF 値は $8\text{TB}/2.5\text{PFLOP} = 3 \times 10^{-3}$ となる。

詳細モデルな神経細胞モデルとシナプスモデル

詳細モデルにおいて神経細胞の膜電位変化のダイナミクスを十分に再現するには、Hodgkin-Huxley モデルの Na と K チャネルの他に、Nap、K(AHP)、Ca(high)、Ca(low)、H などのチャネルを追加することが必要であり、各コンパートメントは、Hodgkin-Huxley モデルの 5 倍程度の複雑度を持たせるとして、20 変数、600FLOP/ステップとする。これは Hodgkin-Huxley モデル 1 ステップの計算には、120FLOP かかるという Izhikevich の見積もり [19] に基づく。

コンダクタンスベースのシナプスモデルの計算は、Matrix exponential 法 [4] を用いると、1 ステップに 5FLOP かかる。シナプス可塑性は、時間スケールの異なるものがいくつかあり、短期可塑性を 2 種類、長期可塑性として STDP (Spike timing dependent plasticity)、短期と長期をつなぐ時間スケールの可塑性を 1 種類想定する。発火 1 回に対する指數関数の積和演算 13FLOP (STDP は 26FLOP) を可塑性の演算量として見積もる。

詳細なヒト全大脳皮質モデル

大脳皮質は様々な異なる処理を行う領域に分かれているが、各領域は同様の構造をしていて、層間の結合パターンについては、他の領域間と類似した規則性を持っている。また、大脳皮質が形成する結合は、約半数が同皮質領域内で形成され、残りの半数は外部の脳領域との間で形成されると言われている。これらの性質は、大脳皮質領域全体の神経回路モデルを並列化する場合、ロードバランス、通信量の削減、メモリ消費量の削減を実現するうえで重要である。例えば、大脳皮質の同領域内の神経細胞は、並列計算機の近傍の計算ノードに配置することで、発火情報の通信の回数やレイテンシを削減することができる。また、大脳皮質全体の構造である、2 次元シート状の構造のトポロジーを保つようにして、並列計算機上にマップすることで自然に実現できる。

ヒト大脳皮質は 160 億個の神経細胞を持ち、1 個の神経細胞あたり 1000～10000 個のシナプスを持つ。以下では、5000 コンパートメントのマルチコンパートメントモデルを用い、1 神経細胞あたり 10000 個のシナプスを持つとして要求性能の見積もりを行う。1 神経細胞とその細胞へのシナプスの 1 ステップあたりの計算量は、約 3 MFLOP ($600 \times 5000\text{FLOP} + 5 \times 5000 \times 4\text{FLOP} + 650\text{ FLOP}$) である。計算ステップを 0.1ms、神経細胞の発火頻度は平均 10Hz とするとき、ヒト大脳皮質全体では 53 PFLOP となる。生物学的 1 秒間の神経活動のシミュレーションは 530 EFLOP である。データ型として倍精度を用いたとき、必要なデータの主記憶メモリからの読み込みと書き出しの量は、合計で約 3.4 MB (神経細胞変数 : 800 KB、シナプスコンダクタンス : 160 KB、シナプス可塑性 : 320 KB の読み出し/書き出しと、神経細胞定数 : 800KB の読み出し) になる。Byte/FLOP (BF) 値は、1.1 となる。ただし、キャッシュメモリを利用して、最小シナプス伝搬遅延時間 1ms の間、主記憶メモリへのアクセスなしで計算を行う場合、BF 値はステップ数だけ減少し、0.11 になる。

計算ノード間のネットワークのデータ通信は、発火情報 (神経細胞の ID と発火時刻) の転送になる。平均発火頻度を 10Hz とし、最小シナプス伝搬遅延時間 1ms ごとにデータ転送を行うとする。脳内の結合パターンは、同一領域内で密、領域間では疎であるという性質がある。そこで、1 領域からデータを転送する領域の数は 10 程度として、転送量を見積もる。1 回の通信で発生する通信データ量は $2.56\text{TB} (1.6 \times 10^{10})$

細胞 \times 10Hz \times 10/10000 \times 16Byte \times 100 計算ノード) となる。ネットワークバンド幅の BF 値は、 $2.56 \text{ TB}/500 \text{ PFLOP} (10 \text{ ステップ}) = 5 \times 10^{-6}$ となる。

神経細胞、シナプスの変数や定数を格納するのに必要な主記憶メモリ量は、39PB と見積もれる。

10 EFLOPS の計算機での実行をする場合、生物学的 1 秒間 (530 EFLOP)、実効効率 10

詳細なヒト全小脳モデル

ヒト小脳の顆粒細胞は 700 億個で、そのほぼ全ては顆粒細胞と呼ばれる非常に小さなニューロンであり、小脳シミュレーションの計算でほとんどの計算リソースが顆粒細胞で消費される。以下では、顆粒細胞を 1000 コンパートメントのマルチコンパートメントモデルでモデル化し、1 神経細胞あたり 10 個のシナプスを持つとして要求性能の見積もりを行う。計算ステップを 0.1ms としたとき、1 頂粒細胞とその細胞へのシナプスの 1 ステップあたりの計算量は、約 0.3MFLOP ($600 \times 500 \text{ FLOP} + 5 \times 500 \times 4 \text{ FLOP} + 650 \text{ FLOP}$) で、ヒト小脳全体では 21PFLOP となる。生物学的 1 秒間の神経活動のシミュレーションは 210 EFLOP となる。データ型として倍精度を用いたとき、必要なデータの主記憶メモリからの読み込みと書き出しの量は、合計で約 0.3 MB (神経細胞変数 : 80 KB、シナプスコンダクタンス : 16 KB、シナプス可塑性 : 32 KB の読み出し/書き出しと、神経細胞定数 : 80KB の読み出し) になる。BF 値は 1.1 となる。ただし、大脳皮質同様に、キャッシュメモリを利用して、最小シナプス伝搬遅延時間 1ms の間、主記憶メモリへのアクセスなしで計算を行う場合、BF 値はステップ数分減少し、0.11 になる。

計算ノード間のネットワークのデータ通信は、大脳皮質同様、小脳領域内は密で、小脳と他領間では疎になる。そこで、小脳からデータを転送する領域の数は 10 度として、転送量を見積もる。また、頂粒細胞の発火頻度を 50Hz として見積もる。1 回の通信で発生する通信データ量は $5.6 \text{ TB} (7 \times 10^{10} \text{ 細胞} \times 50 \text{ Hz} \times 10/10000 \times 16 \text{ Byte} \times 100 \text{ 計算ノード})$ となる。ネットワークバンド幅の BF 値は $5.6 \text{ TB}/210 \text{ PFLOP} = 2.5 \times 10^{-5}$ となる。

神経細胞、シナプスの変数や定数を格納するのに必要な主記憶メモリ量は、22PB になる。

10 EFLOPS の計算機での実行をする場合、生物学的 1 秒間 (410 EFLOP)、実効効率 10

詳細なヒト全脳モデル

上述のヒト大脳皮質とヒト小脳の性能を合算して、ヒト全脳モデルの要求性能とする。計算ステップあたりの演算量は、74 PFLOP、生物学的 1 秒分のシミュレーションの場合、740 EFLOP となる。BF 値は 1.1、ネットワークバンド幅は 1×10^{-5} となる。神経細胞、シナプスの変数や定数を格納するのに必要な主記憶メモリ量は、50PB になる。

10 EFLOPS の計算機での実行をする場合、生物学的 1 秒間 (740 EFLOP)、実効効率 10

カイコガ全脳モデルリアルタイムシミュレーション 自然環境の中での脳の行動決定の振る舞いをみるには入力から出力までの一連の情報処理を取り扱う必要があり、これがリカレントなネットワークである以上最終的には全脳のシミュレーションが必要となる。この意味では昆虫脳のマルチコンパートメントはその神経数とシナプス数で京や富岳のレベルの計算規模としてフィットしており、よい対象といえる。時空間的な情報を取り扱うにはマルチコンパートメント Hodgkin-Huxley モデルが必要であり、1000 ニューロン 1000 コンパートメントであればリアルタイムが京コンピュータ上で達成されている。富岳においてはすべての昆虫の全脳規模である 100 万ニューロンのリアルタイム化が期待される。また、昆虫脳においては同定ニューロンといわれる 1 つまたは少数のニューロンが重要に役割をもち、それが多数の匿名ニューロンと結びつくような構造になっている。このような 神経回路の性質を再構築するには推定器が

重要になる。非線形で複雑な微分方程式の時間発展を推定するのは進化的アルゴリズムを基盤としたものが有力である [20]。また、typical な出力だけではなくて、様々な環境に対して確率的に行われることもある情報処理をシミュレーションにのせるなら統計的な振る舞いをもシミュレーションで行うことが必要である。すなわち stochastic multi-compartment Hodgkin-Huxley 型シミュレーションである。これは例えばコンパートメントごとに動的な荷重をかけた乱数項を付加することによってなされる [21]。このような確率的要素を含む詳細シミュレーションを 積み上げることによって匿名性と identified neuron が併存する神経回路の統計的な振る舞いを調べることができるようになる。最終的には、このような取り扱いにより神経回路のリアルタイムデータ同化すら可能となる [22, 23]。このように富岳時代の昆虫に代表される中規模の神経回路のシミュレーションを睨めば、昆虫全脳レベルのシミュレーションの高速化のみならずシミュレーションを使用した推定器や、分子過程の確率的な性質を考慮した神経回路の統計的な振る舞いが重要となる。

昆虫脳をモデルにしたマルチコンパートメントモデルの大規模神経ネットワークシミュレーションができるだけ早く効率的に行いたいという立場に立つとき、そのポイントはネットワーク性能よりコアの単体性能に存在する。現状では 5000 コンパートメントを仮定すると 50GFLOPS 程度の計算速度がリアルタイムシミュレーションには必要で、実行効率は 10% を大きく超えることは困難である。細胞を分割して計算することは可能だが、現状では計算遅延が問題になって、並列度は極めて低い。1 チップでの性能が 1TFLOPS 程度あれば、おおまかに 1 細胞は 1 チップで計算可能で、リアルタイムが可能となる（ネットワーク BF を考えてホスト CPU をもつアクセラレータタイプではないことを前提としている）。2020 年以降の高性能 CPU の普及に並行して、そのようなマルチコンパートメントタイプの高速神経回路シミュレーションが次第に行われていくと考えられる。また、さらに 10 倍の単体性能があると 1 チップで stochastic multi-compartment simulation が行うことが可能となる。並列性能に関しては单一ニューロパイルである数百ニューロンのネットワークでは本質的には（ただし通信を含めたロードバランシングの問題は存在する）京や富岳レベルのネットワークを考える限り問題は生じないが、1 万-100 万ニューロンの規模ではメモリを含めて様々な問題が生じる。したがって、サイエンス的には < 1000 ニューロンの神経回路 のシミュレーションが先に行われ、100 万ニューロンについてはごく限られた場合にだけ行われることになるだろう。後述の推定器の問題もある。昆虫の触角葉（一次感覚中枢）、LAL（全運動中枢）などはこのレベルの神経回路で詳細にシミュレーションすることが可能である。これらをシミュレーションと実験・データベースと連携して行うことによって、生物脳の物理的性質を関連づけた形で感覚受容や運動パターン生成の理解が進むことが期待できる。また、次世代計算機では 10 万-100 万ニューロンのシミュレーションが日常的に行われることが期待できる。ここでは行動と関連づけた脳全体のループが詳細モデルで再現できるので、基本的な知能のメカニズムの仕組みが物理的詳細モデルのレベルで理解できる世界が訪れるだろう。

推定器について

進化的アルゴリズムによる 1 細胞の性質推定にはだいたい 100 世代 100 遺伝子もあれば足りる。京世代において、細胞分割は 5-10 ノード間程度くらいが適当で、そのなかで上記のジョブが走ることは様々な細胞の性質を実験データと調和させて行うために可能となってきた。このような進化的アルゴリズムを用いた推定は HPC を用いて今後よく行われるようになっていくと考えられる。また、同様なアルゴリズムを用いた回路推定は初期的見積もりでは N の細胞結合強度に対して $N \log N$ 程度の遺伝子数が必要で

ある。500 細胞の神経回路で位置細胞の相手が 5 更に位置を 1 次元で表象して 10000 パラメータとしたときに 40000 遺伝子が必要な計算になる。並列度は 1/100000 のオーダーなのでこれはなんとかなる範囲であるが計算量が爆発する。統計的な手法などを用いて全体のパラメータ分布に制限をかけて自由パラメータは 100 パラメータ程度には押さえ込む必要がある。このためには様々な試行をする必要がある、FRINSTON のフリーエナジー原理や、標準脳マッピングの統計的解釈などにもとづき、統計的手法による研究の進展が期待される。

2020 年代以降を展望するならこのようなパラメータ推定を昆虫脳の全脳レベルで実験と平行しておこなう技術が考えられる。シングルコンパートメントの 1 細胞レベルでは、すでに実現しているが [22, 23]、このような技術が神経回路上で適用できることによって、詳細シミュレーションにとって必要なパラメータがより詳細に明らかになり、さらに深い脳の理解とシミュレーションにつながるという正のフィードバックが期待できる。このようなことが可能なのは物理的詳細モデルを考えながら全脳のリアルタイムが可能であると考えれる昆虫脳に向いている。

参考文献

- [1] Wulfram Gerstner and Werner Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. 2nd edition, 2002.
- [2] Eugene M Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. MIT Press, 2007.
- [3] Christof Koch and Idan Segev, editors. *Methods in Neuronal Modleing: From Ions to Networks*. MIT Press, 2nd edition, 1998.
- [4] Nicholas T. Carnevale and Michael L. Hines. *The NEURON Book*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.
- [5] James M. Bower and David Beeman. *The book of GENESIS - exploring realistic neural models with the GEneral NEural models SImulation system*. Springer, 1995.
- [6] JM Nageswaran, N Dutt, JL Krichmar, A Nicolau, and AV Veidenbaum. A configurable simulation environment for the efficient simulation of large-scale spiking neural networks on graphics processors. Vol. 22, No. 5-6, pp. 791–800, 2009.
- [7] J Igarashi, O Shouno, T Fukai, and H Tsujino. Real-time simulation of a spiking neural network model of the basal ganglia circuitry using general purpose computing on graphics processing units. *Neural Netw*, Vol. 24, No. 9, pp. 950–960, 2011.
- [8] Tadashi Yamazaki and Jun Igarashi. Realtime cerebellum: A large-scale spiking network model of the cerebellum that runs in realtime using a graphics processing unit. *Neural Netw*, Vol. 47, No. 11, pp. 103–111, 2013.
- [9] Masato Gosui and Tadashi Yamazaki. Real-world-time simulation of memory consolidation in a large-scale cerebellar model. *Front Neuroanat*, Vol. 10, p. 21, 2016.
- [10] Runchun Wang, Gregory Cohen, Klaus Stiefel, Tara Hamilton, Jonathan Tapson, and André van Schaik. An fpga implementation of a polychronous spiking neural network with delay

- adaptation. *Frontiers in Neuroscience*, Vol. 7, p. 14, 2013.
- [11] Junwen Luo, Graeme Coapes, Terrence Mak, Tadashi Yamazaki, Chung Tin, and Patrick Degehaar. Real-time simulation of passage-of-time encoding in cerebellum using a scalable fpga-based system. *IEEE Transactions on Biomedical Circuits and Systems*, Vol. 10, No. 3, pp. 742–753, 2016.
- [12] 日経テクノロジー online. スパコン上に猫の小脳を再現、リアルタイムで動作, 6月13日 2016.
<http://techon.nikkeibp.co.jp/atcl/news/16/061302534/>.
- [13] マイナビニュース. ネコの小脳を shoubu スパコンで実現, 6月20日 2016.
<http://news.mynavi.jp/series/numericalbrain/001/>.
- [14] Steve B Furber, Francesco Galluppi, Steve Temple, and Luis A Plana. The spinnaker project. *Proc IEEE*, Vol. 102, pp. 652–665, 2014.
- [15] Andrew P Davison, et al. Pynn: A common interface for neuronal network simulators. *Frontiers in Neuroinformatics*, Vol. 2, p. 11, 2008.
- [16] Paul A. Merolla, John V. Arthur, Rodrigo Alvarez-Icaza, Andrew S. Cassidy, Jun Sawada, Filipp Akopyan, Bryan L. Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, Bernard Brezzo, Ivan Vo, Steven K. Esser, Rathinakumar Appuswamy, Brian Taba, Arnon Amir, Myron D. Flickner, William P. Risk, Rajit Manohar, and Dharmendra S. Modha. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, Vol. 345, pp. 668–673, 2014.
- [17] Marc-Oliver Gewaltig and Markus Diesmann. Nest (neural simulation tool). *Scholarpedia*, Vol. 2, No. 4, p. 1430, 2007.
- [18] 宮本大輔, 加沢知毅, 神崎亮平. 昆虫嗅覚系全脳シミュレーションに向けて: スーパコンピュータによる大規模脳シミュレーションの現在とその展望. 人工知能: 人工知能学会誌, Vol. 30, No. 5, pp. 630–638, 2015.
- [19] EM Izhikevich. Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral cortex*, Vol. 17, No. 10, pp. 2443–2452, 2007.
- [20] 加沢知毅, 宮本大輔, 後藤晃彦, 朴希原, 福田哲也, 神崎亮平. 昆虫全脳シミュレーションへむけて. 日本神経回路学会誌, Vol. 22, No. 3, pp. 89–102, 2015.
- [21] P Orio and D Soudry. Simple, fast and accurate implementation of the diffusion approximation algorithm for stochastic ion channels with multiple states. *PloS one*, Vol. 7, p. e36670, 2012.
- [22] Dimitrios V Vavoulis, Volkmar Straub an John A D Aston, and Jianfeng Feng. A self-organizing state-space-model approach for parameter estimation in hodgkin-huxley-type models of single neuron. *PLoS Comput Biol*, Vol. 8, No. 3, p. e1002401, 2012.
- [23] 李夏栄, 加沢知毅, ハウプトステファン周一, 神崎亮平. 確率的手法によるカイコガ触角葉神経細胞の生物物理的膜特性の推定. Technical Report 382, 電子情報通信学会技術研究報告 ニューロコンピューティング (NC), 2014.
- [24] Eric Richard Kandel, James Harris Schwartz, Thomas M. Jessell, and Sarah Mack, editors. *Principles of neural science*. McGraw-Hill Medical, 2013.

- [25] The Blue Brain Project, 2005. <http://bluebrain.epfl.ch/>.
- [26] Human Brain Project, 2013. <https://www.humanbrainproject.eu/>.
- [27] The BRAIN Initiative, 2014. <https://www.braininitiative.nih.gov/>.
- [28] 革新的技術による脳機能ネットワークの全容解明プロジェクト, 2014. <http://brainminds.jp/>.
- [29] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge, 2014.
- [30] Suzana Herculano-Houzel. The human brain in numbers: a linearly scaled-up primate brain. *Front Human Neurosci*, Vol. 3, p. 31, 2009.
- [31] M Hawrylycz, C Anastassiou, A Arkhipov, J Berg, M Buice, N Cain, NW Gouwens, S Gratiy, R Iyer, JH Lee, S Mihalas, C Mitelut, S Olsen, RC Reid, C Teeter, S de Vries, J Waters, H Zeng, and C Koch. Inferring cortical function in the mouse visual system through large-scale systems neuroscience. *Proc Natl Acad Sci USA*, Vol. 113, No. 27, pp. 7337–7344, 2016.
- [32] Chris Eliasmith, Terrence C Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *Science*, Vol. 338, No. 6111, pp. 1202–1205, 2012.
- [33] Quentin J M Huys and Liam Paninski. Smoothing of, and parameter estimation from, noisy biophysical recordings. *PLoS Comput Biol*, Vol. 5, No. 5, p. e1000379, 2009.
- [34] Ryuta Mizutani, Rino Saiga, Akihisa Takeuchi, Kentaro Uesugi, and Yoshio Suzuki. Three-dimensional network of drosophila brain hemisphere. *Journal of structural biology*, Vol. 184, No. 2, pp. 271–279, 2013.
- [35] Hirotaka Sato and Michel M Maharbiz. Recent developments in the remote radio control of insect flight. *Frontiers in Neuroscience*, Vol. 4, p. 199, 2010.
- [36] Ryo Minegishi, Atsushi Takashima, Daisuke Kurabayashi, and Ryohei Kanzaki. Construction of a brain-machine hybrid system to evaluate adaptability of an insect. *Robotics and Autonomous Systems*, Vol. 60, No. 5, pp. 692–699, 2012.
- [37] Kunkel S, Schmidt M, Eppler JM, Plessner HE, Masumoto G, Igarashi J, Ishii S, Fukai T, Morrison A, Diesmann M, and Helias M. Spiking network simulation code for petascale computers. *Front Neuroinform*, Vol. 8, p. 78, 2014.
- [38] L Chittka and J Niven. Are bigger brains better? *Current Biology*, Vol. 19, No. 21, pp. R995–R1008, 2009.
- [39] Tetsuya Fukuda, Tomoki Kazawa, and Ryohei Kanzaki. Establishment of the estimation method of the neural network using cma-es for elucidating the neural mechanism of a silkworm moth brain. In *9th INCf Congress in Neuroinformatics*, 2016. 3-4 September, Reading, UK.
- [40] Romain Brette, Michelle Rudolph, Ted Carnevale, Michael Hines, David Beeman, James M. Bower, Markus Diesmann, Abigail Morrison, Philip H. Goodman, Frederick C. Harris, Jr., Milind Zirpe, Thomas Natschläger, Dejan Pecevski, Bard Ermentrout, Mikael Djurfeldt, Anders Lansner, Olivier Rochel, Thierry Vieville, Eilif Muller, Andrew P. Davison, Sami El Boustani, and Alain Destexhe. Simulation of networks of spiking neurons: A review of tools and strategies.

- J Comput Neurosci*, Vol. 23, No. 3, pp. 349–398, 2007.
- [41] Esin Yavuz, James Turner, and Thomas Nowotny. Genn: a code generation framework for accelerated brain simulations. *Scientific Report*, Vol. 6, p. 18854, 2016.
 - [42] Rittner P and Cleland TA. Myriad: a transparently parallel gpu-based simulator for densely integrated biophysical models. In *Society for Neuroscience Abstract*, 2014.
 - [43] Henry Markram, et al. Reconstruction and simulation of neocortical microcircuitry. *Cell*, Vol. 163, No. 2, pp. 456–492, 2015.
 - [44] Hodgkin AL and Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, Vol. 117, No. 4, pp. 500—544, 1952.
 - [45] Rajagopal Ananthanarayanan, Steven K. Esser, Horst D. Simon, and Dharmendra S. Modha. The cat is out of the bag: Cortical simulations with 10^9 neurons and 10^{13} synapses. In *Supercomputing 09: Proceedings of the ACM/IEEE SC2009*, 2009.
 - [46] HPCI 戦略プログラム分野1「予測する生命科学・医療及び創薬基盤」. 課題3 予測医療に向けた階層統合シミュレーション—循環器系および筋骨格系・神経系の階層統合シミュレーション—, 2011. http://www.scls.riken.jp/research/03_integration/index.html.
 - [47] 文部科学省科学研究費・新学術領域研究(2016～2020年度). 人工知能と脳科学の対照と融合, 2016.
 - [48] Sten Grillner, Nancy Ip, Christof Koch, Walter Koroshetz, Hideyuki Okano, Miri Polachek, Mu ming Poo, and Terrence J Sejnowski. Worldwide initiatives to advance brain research. *Nature Neuroscience*, Vol. 19, No. 9, pp. 1118–1122, 2016.
 - [49] Terry Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical report, 1971.
 - [50] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
 - [51] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, Vol. 65, No. 6, pp. 386–408, 1958.
 - [52] Herbert Gelernter. Realization of a geometry theorem proving machine. In *International Federation for Information Processing Congress*, pp. 273–281, 1959.
 - [53] Raymond Kurzweil. *The Singularity Is Near: When Humans Transcend Biology*. Viking, 2006.
 - [54] Bruce Buchanan, Georgia Sutherland, and Edward A Feigenbaum. *Heuristic DENDRAL: a program for generating explanatory hypotheses in organic chemistry*. Stanford University, 1968.
 - [55] Edward H Shortliffe and Bruce G Buchanan. A model of inexact reasoning in medicine. *Mathematical biosciences*, Vol. 23, No. 3, pp. 351–379, 1975.
 - [56] Thomas R Gruber. A translation approach to portable ontology specifications. *Knowledge acquisition*, Vol. 5, No. 2, pp. 199–220, 1993.
 - [57] David Silver, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy

- Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, Vol. 529, pp. 484–503, 2016.
- [58] Kunihiko Fukushima and Sei Miyake. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, Vol. 15, No. 6, pp. 455–469, 1982.
- [59] Thomas Serre, Lior Wolf, Stanley Bileschi, Maximilian Riesenhuber, and Tomaso Poggio. Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 29, No. 3, pp. 411–426, 2007.
- [60] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, Vol. 323, pp. 533–536, 1986.
- [61] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *International Conference on Machine Learning*, pp. 807–814, 2010.
- [62] Abdel-rahman Mohamed, George Dahl, and Geoffrey Hinton. Deep belief networks for phone recognition. In *NIPS Workshop on Deep Learning for Speech Recognition and Related Applications*, 2009.
- [63] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [64] Forrest N. Iandola, Matthew W. Moskewicz, Khalid Ashraf, and Kurt Keutzer. Firecaffe: Near-linear acceleration of deep neural network training on compute clusters. In *Computer Vision and Pattern Recognition*, pp. 2592–2600, 2016.
- [65] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.
- [66] Frank Seide, Gang Li, and Dong Yu. Conversational speech transcription using context-dependent deep neural networks. pp. 437–440, 2011.
- [67] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Computer Vision and Pattern Recognition*, pp. 1–9, 2015.
- [68] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

4.6 創薬・医療

本節では創薬・医療分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.6.1 現在行われている課題

「京」を用いて行ってほしい研究を一般の方々にアンケートすると、「創薬や医療に役立つ計算」という意見が非常に多かった。実際、ゲノム配列の次世代シークエンサーの発達によって、個人ゲノムの同定が日常的ななされるようになると、その次の段階として、病気などの予防のために、人によるゲノムの違いが与える影響を分子レベルで解明する必要が出てくる。従来は、実験による特徴づけが中心であったが、「京」が稼働してからは、タンパク質の結晶構造を初期値とした分子動力学シミュレーションを通じて、タンパク質自身の構造ダイナミクスや薬剤の結合安定性に与える遺伝子多型・変異の影響を高い精度で推定できるようになってきた。しかしながら、臨床現場では、患者のゲノム解析によって新たな変異が次々と検出されていることから、これらの変異の機能的意義をシミュレーションを用いて迅速かつ正確に推定することが求められている。また、医薬品開発では、熾烈な国家規模の競争が行われている。計算機を用いて医薬品開発を加速することは、国内医薬品産業の国際競争力を高めることにつながる。複数細胞ネットワークの解明、疾患の解明、がん細胞のネットワーク形成、および詳細な血中薬物動態、血管形状と血液内の物質移動など、実際に医療現場に役立つものまで、数値シミュレーションに対する期待は大きい。血栓症や糖尿病など、血液内のさまざまなイベントを通じて引き起こされる疾患は数多く、その予測に対する社会的要請は大きい。疫学的に知識が構築されてきた研究分野において、数値シミュレーションがもたらす学術的な影響は大きく、例えば、患者個別の血管幾何形状を利用した直接シミュレーションと、疫学的データを照らし合わせることにより、これまでの事実を裏づけ、更には、新たな因果法則を与えることも可能となる。また、患者個別の情報を用いた血流中の薬効評価が可能となれば、有効な医療を創出することにつながる。超音波治療シミュレーションについて、低侵襲治療によって患者の Quality of Life (QOL: 生活の質) が向上すれば、患者の社会復帰が容易になり社会が活性化する効果も期待できる。医療現場への適用という点では、心臓シミュレーションの実用化が期待されている。現在、「京」で研究開発された心臓シミュレーションは、”心臓再同期療法の効果予測”や”先天性心疾患の術式決定と予後予測 [1, 2]”に向けて臨床試験が進められている。この心臓シミュレーションは、「富岳」を利用し肥大型心筋症など指定難病の発症メカニズムの解明に向けて研究が進められている。

生命科学研究分野におけるシミュレーション方法の特殊性は、時間スケールおよび空間解像度の異なるさまざまな現象を解析する必要があることである。そのため、研究の必要性や研究者の興味・関心となる現象に対してさまざまなモデルや方法論が利用される。方法論的には、おおまかに、生体分子を対象とする生体分子シミュレーション、細胞や臓器を対象とする細胞・臓器シミュレーション、ゲノム情報などの大量に蓄積されたデータを対象とするゲノム解析・バイオインフォマティクスに分けられるので、その分類に沿って紹介する。

(i) 生体分子シミュレーション

生体分子シミュレーションには、全原子モデルを用いた分子動力学（全原子MD）、全原子MDを創薬に応用した創薬MD、酵素反応などを解析する量子化学/分子力学混合モデル（QM/MM）、フラグメント分子軌道法（FMO）、粗視化モデルを用いた分子動力学（粗視化MD）がある。

従来、創薬プロセスにおいては発見的側面が大きく試行錯誤にかける時間的経済的コストが高いものであった。近年、病気の原因となるタンパク質の構造が原子レベルで解かれるようになり、分子間相互作用を見ながら分子を設計するSBDD（Structure-Based Drug Design）という開発戦略が取られるようになってきた。（図4.17）しかしながら、数多くの医薬品候補化合物を探索することを目的として計算精度の低い創薬計算手法が主に用いられており、計算科学で創薬プロセスをリードできる状態には至っていない。典型的な創薬計算手法として用いられてきた分子ドッキングではタンパク質の動きは固定されている上に、水分子の影響は無視されており、こうした大胆な近似が結合親和性の予測を困難にしていたと考えられる。「富岳」を用いてこのタイプのドッキング計算を実行した場合、10億種類程度の化合物を現実的な時間で評価可能である。これに対して、ペタスケールの京速計算が可能になった現在、タンパク質の運動性や水分子との相互作用を考慮した全原子MDを活用した医薬品開発が可能になりつつある。

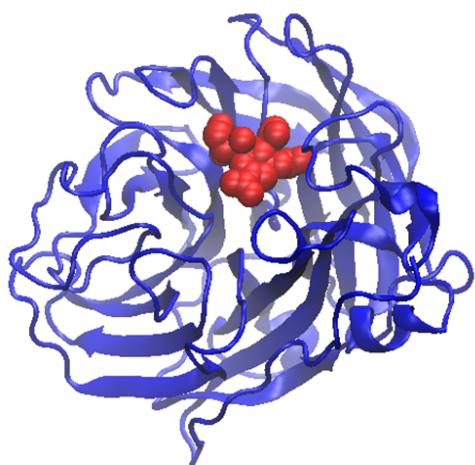


図4.17 医薬品（リレンザ、赤）とインフルエンザウィルスのタンパク質（青）が結合している様子

実際、全原子MDに基づいた標的タンパク質と化合物のドッキングシミュレーション（ColDock法、[3]）を用いて、標的タンパク質に安定に結合する医薬品候補を探索する取り組みがなされており、「富岳」を用いると数千種の化合物を調べることができるようになってきた（図4.18）。更に、標的タンパク質と化合物の結合親和性を計算するMP-CAFE法[4]を用いて、薬としての結合活性を精密に推定するという試みも行われている[5]。D.E. Shaw Researchも製薬会社と協力して医薬品開発を始めており、今後は計算機を用いた創薬支援は実践的なものに発展していくと期待されている。こうした流れを受けて、世界的にもスパコンによる合理的かつ効率的な医薬品開発を目指す風潮ができつつある。

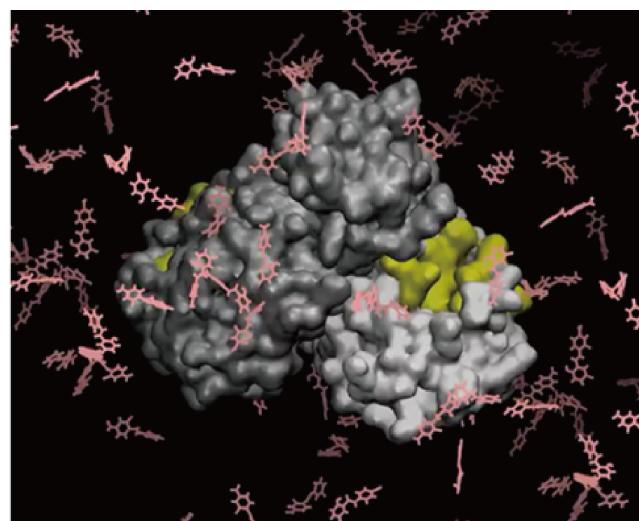


図 4.18 MDに基づいたドッキングシミュレーション(CoDock 法)。標的タンパク質 (surface モデル) の周囲に化合物 (stick モデル) を配置した状態から MD を実行することで、タンパク質ポケット (黄色) に対する結合プロセスを観測する

分子レベルのシミュレーション研究における QM/MM 法は、酵素活性中心などの分子の電子状態を求める必要がある部分空間を量子化学 (QM) 計算で扱い、その触媒活性に重要な影響を及ぼすタンパク質環境を分子力学 (MM) に基づく分子力場で記述するハイブリッド法である。非常に効率的に反応性を解析できるため、現在では、酵素反応解析の標準的な計算手法として広く用いられている。一方、QM/MM 法では、近年、実験的に重要性が認識されつつある酵素・基質の遅い揺らぎの効果を取り入れることが技術的な課題である。QM/MM 計算では、QM 計算は MM 計算よりも桁違いに負荷が大きく、また並列計算も困難なため、QM/MM 法に基づく MD 計算 (QM/MM-MD 計算) で達成できるシミュレーション時間は、通常の MM に基づく MD 計算より数桁短い。この問題を解決する研究が進んでいる。1 つ目は、DFTB (Density Functional Tight Binding) 法、GFN-xTBx (Geometries, Frequencies, and Non-covalent interactions Tight Binding) 法、EVB (Empirical Valence Bond) 法、Car-Parrinello 法など、多少近似が荒いものの高速な QM 手法を用いる方法である。この方法により、現在、ナノ秒オーダーの QM/MM-MD シミュレーションが可能であるが、QM 手法の精度と妥当性を注意深く検討する必要がある。2 つ目は、QM/MM 自由エネルギー構造最適化法である。この方法は、MM 領域の分布で定義される自由エネルギー曲面上で QM 構造の最安定構造を決定する手法である。自由エネルギー計算では、QM 原子を固定電荷におきかえた通常の全原子 MD 計算を実施するため、長時間の MD シミュレーションが可能である。また、QM 領域の構造最適化に高精度な QM 計算を用いることができ、酵素反応と相關するタンパク質の構造変化を見いだすことができる。一方、QM/MM 自由エネルギー構造最適化法では、QM 構造の統計サンプリングが含まれないことに注意が必要である。QM 領域を大きくすることで、QM 計算は正確になるが、自由エネルギー計算は損をするため、適切なバランスが求められる。3 つ目は、QM 計算を機械学習ポテンシャル (Machine learning potential: MLP) に置き換える方法である。この方法では、まず代表的な構造で QM 計算 (あるいは QM/MM 計算) を実行し、得られたデータを学習すること MLP 関数を生成し、これを MD 計算に用いるものである。MLP は無機材料探索では大

いに利用されているが、原子種が多く動きもダイナミックな生体分子系では、MLP 生成が技術的に難しい。現時点での利用はまだ限定的であるが、課題克服のための研究が活発に行われている。

別記の MM-MD/FMO を行う際、MM つまり古典力場の部分電荷パラメータを改良することでシミュレーションの信頼性を向上させる可能性が検討されている。通常の MM の仕組みでは、各原子の原子種や結合状態によって原子タイプを設定し、各原子タイプの部分電荷は、周辺環境の変化に関わらず不变の値を用いている。しかし、実際には共有結合を形成していない部分構造がどのように周辺に分布しているかによって環境静電ポテンシャル (ESP) は異なり、それに対応して各原子の部分電荷は異なる値を用いることが望ましい。FMO の ESP 由来の電荷は計算可能 [6] であるが、コストが高いため簡便な MM での構造最適化や MD の最中でも利用できる効率的な計算手法が求められていた。最近、中心原子の周辺環境を適切に表現できる AFCS 記述子を利用した深層学習による FMO 電荷予測の AI 手法が開発され [7]、今後の改良と応用展開が期待されている。下図は、大きさの異なる PolyQ10, Trp-Cage, BRD2 の 3 種類の系で 100 ナノ秒の MD シミュレーションを行い、1 万個のスナップショットをサンプリングして学習と検証を行った結果で、良好な決定係数が得られている。

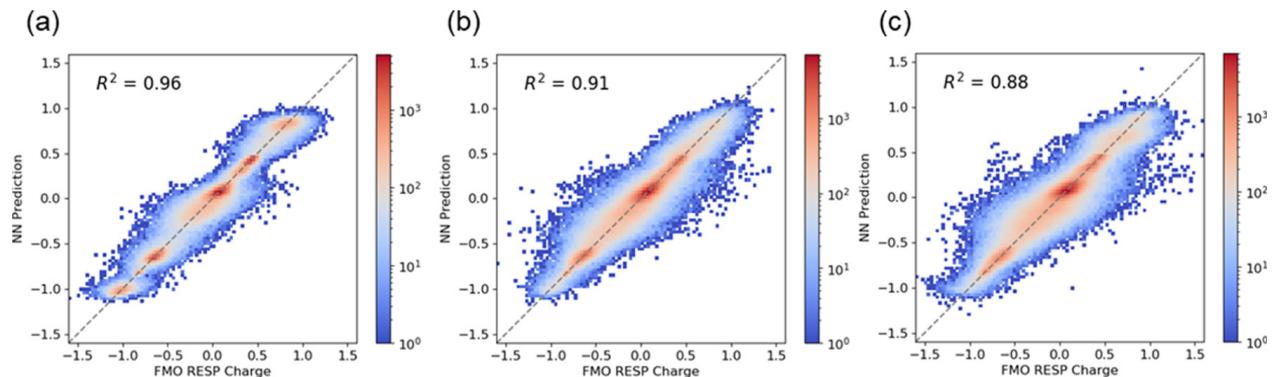


図 4.19 FMO-ESP 電荷の予測精度のプロットと決定係数。(a) PolyQ10, (b) Trp-Cage, (c) BRD2 の例

次に、相互作用解析の成分分解 (PIEDA) を用いた活性予測に関して述べる。タンパク質と医薬候補品等の低分子リガンドの間の結合親和性予測を行う際、FMO 法による相互作用エネルギー値は大変有用ではあるが、溶媒効果やエントロピーに関するエネルギーの寄与を考慮することが課題となっていた。一つの解決方法として、ポアソン-ボルツマン表面領域 (PBSA) による溶媒モデルのエネルギー値を加えることが既に提案されており、良好な結果を示している [8]。もう一つのアプローチとしては、AI による補正が考えられる。フラグメント単位及びエネルギー種別ごとに算出される PIEDA 値は記述子として有用であり、Aurola kinase A の 76 個の阻害剤のドッキング結果に対する PIEDA 値を用いて SVM による機械学習を行ったところ、クロスバリデーションの結果で 0.72 の決定係数を示した（図 4.20）。この値は、PIEDA 値を単純に合計した AI を使わない方法 (PIEDA sum) や分子力場によるエネルギー値を用いる AI (MM+SVM+NSGA-II) と比べて有意に高い予測性能を示した。FMO 計算からは離れるが、MD と機械学習を組み合わせて、リガンドの結合親和性を高効率で予測する新しい技術が国内で開発されている [9] ので触れておきたい。ポイントは、リガンドの結合タンパク質の構造揺らぎの類似性に注目した点で、比較的短い軌跡から相互作用に重要なアミノ酸残基を同定することが可能である。MM-MD/FMO 連携シミュレーションの文脈での進展も期待される。

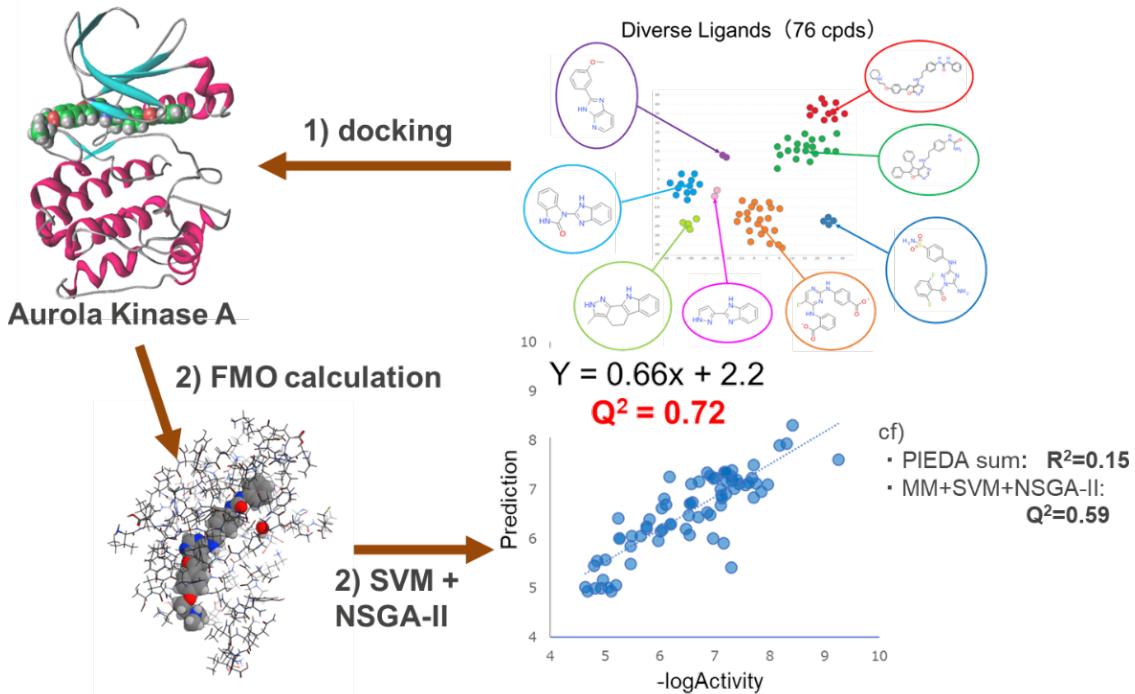


図 4.20 Aurora kinase A の 76 個の阻害剤のドッキングの SVM による評価結果

(ii) 細胞・臓器シミュレーション

細胞レベルから組織レベルを扱うことが可能な連続体細胞シミュレーションでは、物質量をモル濃度で細胞内反応を Michaelis-Menten 式や質量保存式などで記述し、細胞内の物質の濃度変化が計算される。これまでに、連続体細胞シミュレーションはシステム生物学の分野で細胞内の反応や物質相互作用のネットワークの検討に利用してきた。しかしながら、創薬や疾患を検討するうえで重要なのは、個体レベルでの挙動である。個体を理解するためには血液による物質移動や物質循環、相互作用を無視することができない。それを捉えるには、反応ネットワーク的な視点を超えて、細胞・臓器の持つ空間的な広がりや、血流との連成を考慮した解析が必要になる。米国の The Virtual Cell プロジェクト [10] や、EU での Virtual Physiome プロジェクト [11] では、細胞モデルの多次元化が進められ、特に心筋細胞、血管内皮細胞の刺激に対する反応について現在も検証が行われている。国内では、RICS プロジェクト [12]において、超大規模並列計算機の利用を見据え、細胞内構造の不均一性をボクセルデータで表現し、複数細胞内の数百の代謝物に対する数百の酵素反応を捉えようとする新たな取り組みがなされている。

臓器レベルのシミュレーション研究の一つとして、血流解析の例を取り上げる。血流解析シミュレーション技術の開発には、高齢化社会において深刻な問題となっている循環器系疾患の治療法の開発、腫瘍治療などの高度化を進めるうえで重要な意味がある（例えば [13] 参照）。先に挙げた細胞シミュレーションとの連成を進めることにより、創薬を支援することに直結する。そのために必要不可欠なのは、方法論の確立である。血流中における単一の赤血球あるいは血小板など生体膜の力学挙動に関するシミュレーションはここ十数年、数多く研究してきた。特に米国では、応用数学分野との連携により、ストークス力学、境界埋込法による解析に適した膜表面要素やバネ要素を用いる膜のモデル化が進められてき

た [14, 15, 16]。計算機性能の向上にともない、国内外で、複数の赤血球と血流を連成した3次元シミュレーションが報告されるようになってきた。このような適用可能範囲の急速な広がりにより、現在、血流シミュレーションに対する計算力学（とりわけ超大規模計算）の分野における研究者の注目が集まっている。赤血球を含む流体シミュレーションが2010年、2011年のGordon Bell賞を受賞したことは記憶に新しい [17, 18, 19]。国内では、「京」での超大規模並列計算を意識した（あるいは親和性のよい）血流解析手法が開発され、血管径0.2 mm程度の血流解析が進められている [20, 21, 22, 23]。こうした動向を受け、国内では血流中における血栓形成シミュレータの開発が、医療応用上の重要課題として積極的に進められている [13]。血栓形成の初期過程である血管壁への血小板接着のモデル化は、実験と平行して十数年前から行われており、現在单一の血小板接着シミュレーションに関して方法論が確立されつつある状況にある。「京」の計算資源は、数百 μm径の血管内に含まれる血球の「群」としての振る舞いを捉えるのに適していた。「京」を利用する研究課題には、血小板血栓の生成と、それに起因する生化学反応の数値シミュレーションが挙げられる。また、戦略分野では、物性値の変化、凝固・溶解過程のモデル化が進められている。これらの至近の課題に続き、疾患の再現と薬効の評価、治療法の検討などが長期的な研究対象として発展していくものと考えられる。

また、血流解析以外の臓器レベルのシミュレーション研究として、心臓シミュレーションの例を取り上げる。心臓シミュレーションUT-Heartは、有限要素法により心臓をモデル化し、分子レベルから臓器レベルまでの多階層（マルチスケール）を対象に様々な物理現象（マルチスケール）を絡めて再現することを特徴とした、マルチスケール・マルチフィジックスのシミュレーションである。「京」の研究開発では、心筋細胞に含まれるサルコメア（収縮タンパク質分子）の動きを統計力学の法則に基づくバネモデルを利用し再現していた [24, 25]。その後、サルコメアの動きを大きな時間刻み幅を適用できる準陰解法的粗視化分子シミュレーションにて再現し [26, 27]、計算科学の歴史上初となる分子シミュレーションと連成するマルチスケール心臓シミュレーションを開発した。これによりアミノ酸レベルの動きを再現することが可能となり、アミノ酸残基の変異に起因する心筋症の発症メカニズムの解明に向けて研究が進められている。

「京」を用いた他の血流解析として、全脳規模での三次元血流解析が進められてきた（ポスト「京」重点課題2、サブ課題B、責任者・和田成生（大阪大学、教授）、<http://postk.hgc.jp/research/subb>）。医用画像と数理モデルを組み合わせた大脳表層を覆う動脈・静脈経路モデルを用い [28]、計算性能に優れた格子ボルツマン法に基づく三次元流体解析を全「京」規模で実施することで、大脳表層血管内を巡る血流場の数値再現に成功している [29]。「富岳」においては、数百 μm立方領域での脳微小循環を対象とし、赤血球運動と酸素輸送を連成した三次元解析が進められており、これまで実験観察主体となってきた血流と脳機能の関連解明について、物理メカニズムを介した理解が進められている（成果創出加速プログラム、全脳血液循環シミュレーションデータ科学に基づく個別化医療支援技術の開発、課題代表・和田成生（大阪大学、教授）、2020年度2022年度）。

医療へのシミュレーションの応用では、低侵襲治療の一つである超音波治療の実現に向けて、現在、機器開発の段階で厳しい競争が繰り広げられている。その数値的研究は現在進行中の「ものづくり」的課題や、市場の動向と密接な関係がある。近年、世界的な流れとして、患者のQuality of Life（QOL）向上への要求とともに、低侵襲治療への関心が高まっている。なかでもエネルギー治療の一つである超音波治療が注目を集めている [30, 31, 32, 33, 34]。超音波治療の機器開発は、欧米と中国で盛んに進められて

おり、日本はその後塵を拝している。欧米では、前立腺肥大や子宮筋腫などに対する超音波治療装置がすでに実用化され、認可もされている。更に最近では、MRI を併用した MR ガイドの超音波治療装置の開発が進められ、より深部の腫瘍に対する超音波治療の臨床試験も進められている [35, 36]。それに対し日本では、超音波診断装置や MRI などの診断機器の技術開発のポテンシャルが高く、国内需要も高いため、積極的に新手法や新装置が研究開発されてきた。しかしながら肝心の治療器については、景気後退、健保財政の悪化、医療機器審査の高いハードルなどの複雑な国内事情により、製品化が困難であり消極的であった。このような消極性が、1980 年代以降、医療用具全体の貿易収支における赤字の膨大、更には、国民医療負担の増大をもたらした原因の一つとして指摘されており、わが国は、戦略的な治療器開発の必要に迫られている。その中で、低侵襲治療を実現する HIFU (High Intensity Focused Ultrasound) 装置の開発が進められていた [37]。特に装置の制御や生体へのリスク評価において、数値シミュレーションの果たす役割が大きく位置づけられている。海外ベンダーの状況に関して、精力的に開発が進められていることは伝わってくるものの、正確な計算機利用情報は、企業秘密の部分が多く不明である。だが、制御法に関しては 1990 年代に欧米で提案され、Finite-Difference Time-Domain 法との親和性の高い Time Shift 法 [38] や Time Reversal 法 [39, 40] などが方法論として確立しており、これらの開発競争においては計算機性能が鍵を握る。わが国では、これまで、HIFU シミュレータの開発が進み、脳内腫瘍を対象とした実寸スケールの超大規模計算が実現可能となっており、世界的に高い競争力がある [41, 42]。「京」の計算資源は、ある条件の計算を一つ行うことは能力的に余裕があるので、パラメータの探索を効率よく進めることができ、開発競争における戦略上、非常に魅力的である。そのため、「京」のハードウェア構成に合わせたソフトウェア開発が積極的に進められてきた。現在は、実験により各種試験が行われている状況である。一方で、超音波診断装置の向上を目的とした AI 向けの学習データ作成に超音波シミュレーションが利用されており、多くの詳細な三次元シミュレーションが必要となり、より高度な計算機が必要となる。

4.6.2 長期的目標

すでに述べてきたように、生命現象の複雑さや多様さなどにより、これまでの生命科学の分野では異なる計算手法やモデルが用いられてきた。また、計算機の演算性能の限界により十分な時間およびサイズの計算が困難だった。今後、長期的研究ではその制約の多くが取り払われ、研究したい生命現象をまるごとシミュレーションできる日が来ると期待される。そして、理論・シミュレーションの専門家だけでなく、実験・医療関係者にもシミュレーションの有効性が明らかになっていくと思われる。そこで必要となる理論や手法の開発を行い、有効な実験とシミュレーションの共同研究を行っていく必要がある。更に、今後の生命系のシミュレーションでは、予測可能性という点が重視されると思われる。社会応用という観点からも、創薬や医療機器開発においてシミュレーションが盛んに利用されるようになると予想される。下記では、各スケール・方法における今後の発展について記述する。

(i) 生体分子シミュレーション

全原子 MD では MD 計算専用計算機が今後とも計算速度はリードしていくと思われる。しかし QM 計算とのハイブリッドや、自由エネルギー計算を目的として、バイアスポテンシャルを利用する計算法な

ど、今後とも開発される新たな計算法では、汎用計算機の利用が重要な役割を果たす。また、高並列な計算機で全原子MDを行う場合には、多数のレプリカを疎結合して系の統計的性質を解析するアンサンブル計算の重要性が増していくと思われる。

創薬MDにおいても、計算機パワーの向上は、より大規模な化合物探索（ポスト「富岳」を利用した場合、数十万種から数百万種）を可能にすると予想される。高精度な全原子モデルによる大規模な化合物探索が可能になることで、有望な医薬品候補を発見できる可能性は格段に向上し、計算による創薬の有効性はますます高まるだろう。また、膜タンパク質のような複雑な系に対しても計算機で薬を設計することが可能になりつつある。現在は、標的タンパク質と化合物の結合強度に焦点が当たっているが、副作用を抑えるという観点から、標的ターゲット以外のタンパク質と化合物との相互作用を調べることは重要である。計算機の進歩により、こうした医薬品候補の特異性を予測する試みもシミュレーションに基づき行われると予想される。

今後、SACLAをはじめとする実験技術も大きく進展し、創薬ターゲットになり得るタンパク質の立体構造情報は指数関数的に増えていくと予想される。こうしたタンパク質立体構造情報をすぐに計算に持ち込み、医薬品候補を探し出すというスピード感も重要な側面となる。そのためには、実験との連携を十分に高めていくことは当然ながら、数多くの標的タンパク質に同時並行で対応できる計算機環境が重要である。

QM/MM法は、酵素反応活性解析のみならず、新規酵素活性設計にも適用されていくと考えらえる。特に、緑色蛍光タンパク質（GFP）やチャネルロドプシンのような、機能性タンパク質ツールの改変設計に大きな力を発揮していくと考えられる。生物学と化学のシームレスな記述を生かし、バイオミメティックな化学的触媒活性化合物の設計にも応用されていくと期待される。創薬研究においては、これまで問題であった高精度な相互作用の記述とタンパク質構造変化の考慮の両立が可能となる。

方法論的には、 $\mu\text{-m}$ 秒のタンパク質構造変化と高精度なQM計算を両立できるQM/MM自由エネルギー構造最適化法の適用が進むと見込まれる。これまで、実験的に直接追うことができていない酵素に特徴的な可塑的触媒場の新たな概念が明らかになる可能性がある。これは、上記の酵素活性設計や薬剤設計に、概念的にも、実際の設計スキームを構築するうえでも、決定的に重要となるであろう。一方、QM/MM-MD計算は、シミュレーション時間は短くなるものの、サブナノ-ナノ秒オーダまで伸びてくると予想される。これまでQM計算の並列化効率向上が困難であったが、並列化を前提としてスクラッチからQM計算プログラムを作り直す動きが始まっており、これを導入することで、DFTレベルのQM/MM-MD計算がサブナノ秒オーダで可能になる。さらに、レプリカを用いた高い並列度の全原子MDと組み合わせ、GPGPUや超並列計算機を用いることで、ナノ秒オーダのシミュレーションが可能になる。また、計算精度に気をつける必要はあるものの、DFTB法やGFN-xTB法を用いることで、シミュレーション時間を1-2桁伸ばすことが可能になる。MLPは、近年のNeural Networkの発展を受け、今後のさらなる進展が期待される。MLPでは、できることとできないことの明確化、on-the-flyでデータ生成と学習をする場合、データ生成も含めた計算コストの評価と再現性、MLPの汎用的な実装と高速化、などが課題である。これらの課題に対する研究は急速に進んでおり、MLP/MM-MDは強力な方法へと成熟することが予想される。

また、並列数の増加に従い、問題を大規模化・高精度化するようなQM計算の場合には、QM計算においてよくスケールする手法がいくつか開発されており、QM/MM法の適用が進むと考えられる。特に、

大規模分子全体にわたる電子波動関数の振る舞いが重要となる、電子移動系の解析には大きな力を発揮すると思われる。

ハートリーフォック (HF) と 2 次摂動論 (MP2) などの電子相関法による通常の FMO 計算とは別に、強結合近似による簡易密度汎関数法 (DFTB) による FMO-DFTB [43] が GAMESS-US に実装されており、計算コストの低さから構造最適化を含めて実問題に用いられるケースも出てきている。課題点を挙げるとすれば、タンパク質のように強度に分極された系での（信頼性の高い）強結合パラメータの設定であるが、今後、AI 技法の援用によって改善される可能性はある。一方、DFTB では分割-統合によるアプローチ (DC-DFTB) もあり、超並列処理対応の MD プログラム [44] も既に整備され、凝集系での反応ダイナミクスまで扱える点で先行していることは書き添える必要がある。

次に、定量的な相互作用解析に強みを持つ FMO 計算の特徴を活かし、創薬分野での公益情報基盤を構築する活動に言及したい。これは、FMO 創薬コンソーシアム (FMODDB) を中心に実施したタンパク質複合体の計算結果をデータベース (FMODDB) に蓄積し、世界中の誰もが利活用できるようにしているもので、2018 年から公開されている [45, 46]。2020 年からは、COVID-19 の創薬研究に貢献するために COVID-19 関連タンパク質の FMO 計算を系統的に実施し、特集ページにて公開している [47]。2022 年 7 月現在、2,956 種類のタンパク質の 14,924 個の FMO 計算データが登録され、そのうち 830 個は COVID-19 関連タンパク質である。FMODDB 内のこうしたデータは、別記の FMO 電荷予測 AI 等の機械学習系解析の元データにもなっている。FMODDB の構築及び拡張には、FMO 計算を半自動的に実施するプロトコル [48] に加え、「京」、TSUBAME、「富岳」などの HPCI リソースが要な役割を果たしている。このプロトコルについては、タンパク質の前処理の部分で市販ソフトウェアに一部依存しているが、近々に無償の（あるいは安価な）ソフトウェアのみで稼働できるようにする予定である。これにより、FMO 計算の熟練者ではない構造生物学者など他分野の研究者に対して FMO 計算による相互作用解析研究が普及すると期待される。併せて、FMO 計算のオンデマンドサービスの開発も視野に入っている。このサービスでは、座標データを Web を介して登録してリクエストを出すと、上記の自動プロトコルによって計算された結果が返される想定だが、その前提が「富岳」のような豊かな計算資源の活用であることは自明である。ポスト「富岳」でも、こうした FMO 計算の Web サービスは有力なアプリケーションの形態になり得るだろう。

FMO をマルチスケールシミュレーションに用いる試みも発展しつつあるので、記しておきたい。粗視化シミュレーションはメゾスケールの計算手法であり、粗視化 MD (CG-MD) や散逸粒子力学 (DPD) などがあり、ナノスケールの MD に比せば大きな配置空間を効率よく探索することができ、高分子の相分離挙動の解析などに用いられている。粗視化粒子間の有効相互作用は実験値に基づいてフィットするか、あるいは経験的な MM によって算定されるが、分極や電荷分離の大きな系、あるいは界面が複雑な系では上手くいかないケースがある。そのため、FMO 計算によって非経験的に DPD の有効相互作用 (χ) パラメータを求める準自動化されたプロトコル [49] とワークフローシステムが開発され、マルチスケールな FMO-DPD シミュレーションとして、脂質膜の構造再現 [50]、mRNA の DDS 用ベシクルの構造評価 [51]、ペプチドの畳み込み [52] などに用いられている。 χ パラメータ算定では、粗視化粒子の種類が増えると小規模な FMO ジョブが数十万のオーダーで必要となるが、「富岳」では待ち時間を除けば効率よく処理される (capacity computing)。現在、機械学習による粒子間相互作用エネルギーの推定を導入して FMO ジョブの実行数を低減してターンアラウンド時間を改善する改良が進められている。また、

FMO-DPD のナノ→メゾの向きの逆、すなわちメゾ→ナノのスケールを繋ぐツールの開発も行われている。これがリバースマッピングと呼ばれる技術であり、DPD 構造をアフィン変換した後に MD によって緩和させる。こうして得られた構造に対して通常の MD シミュレーションを始めたり、FMO 計算による詳細な相互作用解析を実施することもできる。

以上の分野は、物質科学的側面もあるので、物質科学分野、ものづくり分野との連携も重要である。さらに、機械学習あるいは広義の人工知能といった情報科学分野のコミュニティとの共同作業も必要になってくる。

(ii) 細胞・臓器シミュレーション

連続体細胞シミュレーションの適用可能な系が、計算機性能の向上にともない、均質な場から不均質な場へと広がってきた経緯を振り返ると、今後は、細胞の構造・形態のダイナミクスの捕捉を目指す方向へと発展するであろう。5年以内には、一番小さい組織である肺ランゲルハンス氏島（5,000～1万細胞）や血小板による血栓形成のシミュレーションが可能になると考えられる。より長期的視野から展望すると、細胞スケールの現象は、分子スケールから臓器スケールまでの異なるスケールの現象を渡す役割を果たすため、今後は、粒子系、流体・構造力学系の研究と連携、統合して研究が進むだろう。細胞の形態変化の仕組みは、実験観測による定性的な知見はあるものの、計測に限界があるため、それをもたらす力学的機構については定量的に不明な点が多い。この状況の打開には、粒子系の数値シミュレーションにより、ボトムアップ的なモデルの構築が期待される。また、細胞内反応を基とした組織・臓器・器官の生化学血流連成シミュレーションにより、疾患解明や創薬につなげる試みも進んでいくと思われる。

臓器レベルのシミュレーションにおける血流解析シミュレーションについて、これまで流体・構造・膜の力学的連成解析法の開発が大きな目的であったが、現在は方法論が確立しつつある。したがって今後の発展は、計算機性能の向上にともなう計算対象の大規模化に向かうものと考えられる。数年～十数年先においては、数百 μm ～数 mm 径の血管内の血流解析が、計算機で捕捉可能になると考えられる。ただし、生命研究として意義のある結果を得るには、依然問題が山積みである。血栓生成は数秒～数分の時間スケールの現象であり、数値シミュレーションで有意な結果を得るには、生体膜の力学挙動を捉えるための離散時間幅 1 マイクロ秒に対し、 $O(10^6)$ ～ $O(10^8)$ 回もの時間積分を行う必要がある。そのため、数値的安定性の向上および数値誤差蓄積の低減が不可欠である。また、超大規模並列化に耐えうる並列化アルゴリズム、あるいは方法論そのものの見直しも必要となる。上記の問題を解決する方策として、粗視化法の構築も課題となる。そのためには、現象の本質と関わる血流の多重スケール性を踏まえたモデル化が重要となる。更に、血栓シミュレーションにおいては、上述のように、細胞の生化学反応をモデリングしシミュレーションに取り入れる必要があるが、個別要素の積み重ねから得られる知見が重要であり、モデル構築においても詳細な解析が必要不可欠である。

心臓シミュレーションは、分子シミュレーションを融合しアミノ酸レベルの解像度を持つことで、アミノ酸残基の変異に起因する肥大型心筋症や拡張型心筋症を引き起こすメカニズム、および心筋に及ぼす影響の解明に向けて研究を進めている。「富岳」では、左心室の動きを対象とした”左心室モデル”を採用しているが、より忠実な現象を再現するためには心臓全体を再現して観測する必要がある。ポスト「富岳」の時代には、計算対象となる心臓部位の拡大し、さらにはサルコメア内に含まれる分子を三次元でモデル化し実際の数値に近づけることで、心臓の拍動、血圧、血流などを忠実に再現する。

医療への応用における超音波治療シミュレーションの今後は、現在と同様、製品化に至るまでの治療器の開発状況や、市場の動向と密接な関係を持つはずである。長期的な予測は不確定要素が多く困難であるが、これまでの超音波治療を取り巻く状況を踏まえて、今後の発展について具体例を挙げていく。MR ガイドの超音波治療装置は低侵襲治療を実現することが実証されており、すでに、実際の治療に用いられつつある。しかし、装置導入に掛かる高コストが、普及の妨げとなっている。そこで、治療部位を特定する際に、MR ガイドの代わりに低コストの超音波ガイドを利用した、超音波診断治療器の開発が検討されている（例えば [53, 54]）。超音波ガイドの実現には、軟らかい生体組織の温度変化で生じる微妙な剛性の変化をいかに精度よく捉えられるかが鍵となる。そしてその開発においては、剛性の温度依存性を捉える大規模数値シミュレーションによる支援が大きく期待されている。また近年では、薬物の体内動態を制御し、目標とする患部に必要な量の薬物を投与するドラッグデリバリーシステム（DDS:Drug Delivery System）において、超音波照射により治療効果を向上させる試みもなされている [30, 31, 32, 33, 34]。海外の超音波治療器メーカーは、薬剤メーカーと提携してその開発に注力し始めている。DDS で用いられる数ミクロンのマイクロカプセルが、血管中に 3 次元的に分布した生体中では、干渉によって超音波音場が大きく影響を受けることが知られている。マイクロカプセルの動的挙動には強い非線形性があるため、そのような場の予測は、数値シミュレーションによってのみ可能となる [55, 56]。よって、超音波診断治療装置を用いた DDS の実用化に向けた効率化や安全性の評価において、数値シミュレーションが果たすべき役割は大きい。超音波診断装置については、エコー画像の更なる向上を目的に AI の利用が進められており、AI の学習データの作成に超音波伝播シミュレーションが活用されている。現在は 2 次元のシミュレーションに留まっているが、今後は 3 次元シミュレーションに拡張することが想定される。その場合、AI 处理も含めた大規模なデータのハンドリングが必要になり、HPC システムへの課題となりうるとも考えられる。高精度で大規模な超音波シミュレーションによって超音波診断装置における画像再構成の高度化がすすめられると一緒に画像診断の精度も向上して、将来の医療の進歩につながることが期待できる。

4.6.3 次世代に解決すべき課題

(i) 生体分子シミュレーション

創薬 MD では、近年、可溶性タンパク質だけでなく、細胞膜の力場の発展に伴って膜タンパク質の分子動力学シミュレーションも簡易的に実行できるようになってきた。従って、様々な種類の標的タンパク質に対して実践的な創薬 MD を行うことが可能となりつつある。そこで、数万種から数十万種の化合物に対する結合シミュレーションを同時に実行して、標的タンパク質に対して結合強度の高い医薬品候補を選び出すため、創薬シミュレーションの自動化を行う必要がある。更に、副作用を軽減するために、オフターゲットとの結合強度も同時に計算し、医薬品候補の結合特異性を評価する必要がある。今後は実験技術が進歩し、解かれる標的タンパク質の 3 次元構造は指数関数的に増加していくと予想される。これに応じて、計算分野でも、即時に数多くの標的タンパク質に対応する体制が求められると予想される。更に、低分子医薬品だけでなく、ペプチド・核酸・抗体といった中分子・高分子医薬品の探索・分子設計を可能にする創薬シミュレーション技術の開発も求められる。

QM/MM 法では、まずは分子モータータンパク質や膜輸送体・G タンパク質共役受容体などの分子機

能解析がターゲットとなるであろう。これらは、タンパク質機能において化学反応とタンパク質構造変化が強く相関しており、生体分子機能のよいモデル系となる。また、実験的研究も進んでおり、計算科学的アプローチの有効性の検証に必須である緊密な共同研究も期待できる。また、酵素反応性解析で得られた知見に基づく酵素設計も行われる。酵素のアミノ酸配列の大幅な改変に基づく新規機能設計は、非常に挑戦的で、バイオインフォマティクスや全原子MD計算などと組み合わせた高度なスキームの洗練が必要となるであろう。一方、小規模改変による設計は、実験的な要求に応え得るもののが十分期待される時期に来ている。実験研究との連携を通して実績・経験を積みながら、高い仕様要求を満たす新規機能性タンパク質設計へのスキーム構築を目指すとともに、バイオミメティック触媒設計への応用を行う。

ここで、FMOの創薬応用とAIやデータ科学の関係について改めて記したい。MDやFMOなどの分子科学系シミュレーションとAI技術は相互補完的な関係であり、前者は計算コストは大きいものの、立体構造の情報さえあれば時系列変化や精密な相互作用エネルギーを予測することができ、阻害剤の種類による適用範囲の制限も少ない。一方、AIは学習データ依存性がきわめて強く、阻害剤の活性予測等について、既知データに近い化合物では高い精度で予測可能であるが、既知データから遠い化合物では急速に予測精度が下がる傾向がある。2010年代までは、シミュレーションとAIは別個に利用されることが多かったが、最近では両者を連携させる試みが増えている。FMOに関しても、別記している電荷予測や活性予測がある他、学習終了後には構造入力だけでアミノ酸残基間の相互作用エネルギーを高い信頼度で予測するサロゲートモデルも開発されている。他方、MDの長時間積分による統計量算定を敵対的生成ネットワーク(GAN)によって効率化するアプローチもある[57]。データの解析でも、主成分分解(PCA)、特異値分解(SVD)、テンソル分解などを用いるケースが出てきている。実際、MD軌跡のPCAはよく行われている。また、MM-MD/FMO計算の結果解析でPCAとSVDが使われた例[58]、スパイクタンパク質の3本鎖の残基間相互作用エネルギーの3階テンソルをCP分解して重要箇所を特定した例[59]もある。変わった事例では、タンパク質の相互作用エネルギーを可視化した上で深層学習で判定させ、構造的な特徴を自動判定させた試み[60]もある。「富岳」の時代になり、膨大な計算結果が簡単に得られるようになったが、解析による有用データの抽出は「人力ではやりきれないレベル」に達しているとも言え、AI・データ科学的な解析手法を援用することはもはや必然である。ポスト「富岳」の時代には、扱うべきデータ量はさらに増えるため、こうした解析手法を今から周到に整備しておくべきで、データ科学の専門家との共同作業も必要になる。

全原子MD及び粗視化MDに関する課題については4.2章を参考。

(ii) 細胞・臓器シミュレーション

細胞・血流シミュレーションについて、対象とする問題は多数考えられるが、例として、血栓性血小板減少性紫斑病[61]に関して説明する。これは、細動脈(血管径100~200μm)において過剰に血小板が積み重なることで、血流全体の血小板数が減少し、それに伴い出血(皮下出血による紫斑)が見られる疾患である。脳の細動脈で形成されると神経障害、腎の細動脈では腎障害を引き起こし、放置した際の致死率は90%以上と知られている。この疾患の原因是、血漿中に存在するADAMTS13と呼ばれるフォンビルブランド因子の切断酵素の欠損であり、通常、このADAMTS13は血中の超高分子量フォンビルブランド因子多量体を適度な大きさに断片化することにより、過度な血小板凝集を抑制している。近年では、ADAMTS13の立体構造と血流のずり応力との関係が指摘されており、徐々にその原因因子が明らかに

なってきているが、創薬の観点ではまだ研究の初期段階である。今後、分子動力学法あるいは実験的アプローチにより、単一の分子に対する創薬開発が期待できるが、実際の人体においては、血流と細胞の代謝反応が連成して起こり、通常、そのような代謝反応は正のフィードバック機構が働くため、血流も踏まえたシミュレーションによる薬効評価が重要となると考えられる。このため、対象はある程度限定しつつ、解析アプローチを確立することが研究の対象となる。「富岳」を対象としたプロジェクトでは、連続体力学に基づくシミュレーションと実験観察・医用計測情報を組み合わせた物理法則に基づくデータ同化手法の開発が進められている。例えば、非侵襲的に取得した患者個別の血流情報を用いデータ同化血流解析を行うことで、患者個別の血行動態・血行力学因子を時空間方向に高解像で評価し、疾病の機序解明・早期診断・手術支援などに繋げる試みがなされている。また、発話時の舌形状を用いデータ同化変形解析を行うことで、舌筋線維の収縮力および神経活動を推定し発話調整機構の機序解明や治療診断に繋げる試みがなされている。これら一連の解析では、物理法則やシミュレーションモデルを介し、観察される情報以外の物理量を評価することが可能となるため、エビデンスベースでの現象本質の解明や新たな医工学技術に繋がることが期待される。今後、大規模な時空間計測データを用いた同化解析に向けた大規模シミュレーションモデルの確立が必要とされる。マルチスケール解析においては、生命現象の本質である非平衡性を再現する数理的枠組みの確立が必要不可欠である。これは、それぞれのスケールで導入される物理モデル間の時空間差をどのように埋めるかがカギとなるが、非平衡統計力学など理論的枠組みの確立に加え、機械学習を用いた代理モデル構築やモデル縮約など計算科学・データ科学的な側面からのアプローチの確立も重要となる。

心臓シミュレーションについては前述したようにサルコメア内に含まれる分子の再現に分子シミュレーションを融合している。分子シミュレーションとして粗視化分子モデル CafeMol を利用し、計算対象となるミオシンの構造変化とその仕事に関するエネルギー的議論を可能とするために、粗視化力場を改良し、合わせて化学反応に伴う力場のスイッチについても変更を加えることで心筋細胞の動きをより忠実に再現することを可能とした。またミクロな分子シミュレーションとマクロな有限要素法による連続体シミュレーションは互いに因果関係があり、2 レベルの計算を交互に実行する必要があるため、マクロな計算を行う間は大多数のノードで待ち時間が生じる。このような待ち時間を最短にするためにマクロな計算はより高性能な小集団のノードに割り振り計算できることが望ましい。このような観点からマルチスケール解析では小集団の高性能計算ノードを混合したヘテロ計算機が望まれる。また心臓のマルチスケール解析の通信は、ミクロ-マクロ間のやり取りが主体となるため、高性能ノード集団と他の個々のノードとの通信が効率よく実施されることが望ましい。ミクロな計算を行う分子シミュレーションでは、心筋細胞を構成する有限要素ごとに含まれるタンパク質となるミオシン分子を対象としている。「富岳」では1つの有限要素内のミオシン分子数は16 個を基準に計算を進めていたが、ポスト「富岳」では4倍を超える75 個を螺旋状に配置して計算することにより忠実な心臓の拍動を再現する。分子シミュレーションは有限要素内の分子単位で並行して計算することが可能であり、大規模な並列化の適用が課題となっている。また、「富岳」では臓の拍動1秒分を再現するために10日の計算を要しており、上記で述べた並列化に加え、ハードウェア面、ソフトウェア面共に演算性能の向上が必要不可欠となっている。

超音波伝播シミュレーションについて、これまで生体内の縦波の伝播のみが扱われてきたが、エクサスケール級計算機では、縦波に比べて2桁程度遅い伝播速度を持つ横波の影響 [62] を考慮した解析を行う。これにより、治療において報告されている皮膚の火傷や骨の痛み等の生体組織界面における横波によ

る加熱の影響を加味することができ、超音波治療のリスク評価の高精度化や、温度上昇による剛性の変化を利用した超音波による治療部位特定手法の開発が可能になる。更に、ドラッグデリバリーにおけるマイクロカプセルと超音波が干渉した音場の解析や、超音波を援用した場合のドラッグの拡散に関するシミュレーションも可能になると考えられる。HIFU の研究について、最近は強度がより高く、ショートパルスな超音波照射による組織破碎を対象にした研究が増えてきている。このキャビテーションを応用した強力集束超音波治療において、キャビテーションの音響的、熱的また力学的特性の効果的な利用が新たな技術開発の鍵になっている。そのため、集束超音波キャビテーションに対するマルチスケール・マルチフィジックス解析によって、生体へのキャビテーションの力学的作用に関する知見を得る必要がある。このためには、音響的な非線形だけでなく、気泡の非線形体積振動を考慮した解析が必要になるため、現状の HIFU シミュレーションよりも大規模な計算が必要になる。また、計算の課題としては、気泡の非線形体積振動の時間スケールが非常に小さく、振動状態によって並列計算時のロードインバランスを引き起こすため、これの対策が必要になる。超音波診断装置の高度化に向けた利用では、より高性能な計算機環境が必要になるが、大規模データを手軽に扱えるシステムとのバランスが次世代の計算機には求められると考えられる。

4.6.4 課題を解決するために必要なアプリケーション群（要求性能）

(i) 生体分子シミュレーション

分子動力学計算は、運動方程式を時間について数値積分することで、分子を構成する粒子（原子）の動きのシミュレーションを行う計算手法である。計算する相互作用については、近距離の粒子間に働く相互作用と、遠距離の粒子間に働く相互作用があり、それぞれ計算コストのあり方がかなり異なる。近距離間相互作用は、空間的に系を分割し粒子を計算ノードに割り当てるか、隣接ノード間通信だけで計算に必要な情報はカバーできる。一方の遠距離間相互作用では、情報を大域的に共有する必要があるため、遠距離ノード間通信が発生する。現時点では、遠距離間相互作用計算では、FFT を利用する計算アルゴリズムが普及しているが、通信量を削減するため FFT を利用しないアルゴリズムの開発も求められる。

一方、実際の生物中では、分子は多数存在しアンサンブルとしてその機能を発揮しているので、一分子だけでなく、多数分子の統計量（自由エネルギー等）を計算すべきという考えがある。それに基づいた手法としては、対象系の複数のレプリカを計算機中に置き、それぞれのレプリカでシミュレーションを行いつつ、そのレプリカ間で情報交換を行いながら計算を進めていく方法がある。幅広いサンプリングを実現するレプリカ交換や、タンパク質の機能発現にともなう構造変化を扱うレプリカパス法が代表例だが、これらの方法では、レプリカ間は疎結合ですむので、大規模並列計算に適しているという特徴がある。

創薬への応用において、標的タンパク質と化合物との結合強度（結合自由エネルギー）を定量的に求めることが、薬としての効力を予測するうえで重要である。結合自由エネルギーを精密に求める方法（例えば MP-CAFE 法）は、全原子分子動力学計算に基づいているため、計算手法の改良によって高速化を図っていくことが重要な課題となる。また、大規模な化合物探索に向けて、化合物の力場の高精度化・及び一連の計算プロトコールの自動化を行っていく必要がある。また、実験チームと連携することで、シミュレーションによる予測結果を実験的に検証し、その結果を計算チームにフィードバックできる体制を整えていくことが重要である。

創薬 MD では、標的タンパク質に対して 10,000 種類の化合物を MP-CAFE 法で結合強度を調べる場合、100 PFLOPS マシンを使えば 7 日くらいで計算が終了する。従って、10 EFLOPS の計算パワーがあれば、同じ期間で 100 万種類の化合物を評価できると見積もられる。ここでは、各々の化合物に対する結合強度は密に通信する並列化ジョブで計算されるが、その並列化ジョブは各化合物別のアレイジョブとして実行されたとした。全原子 MD に基づいた標的タンパク質と化合物のドッキングシミュレーション (ColDock 法) においても、同じ規模の化合物を評価できると期待される。MP-CAFE 法や ColDock 法といった高精度計算によって評価できる化合物が飛躍的に増加し、更に複数の標的タンパク質に対して同様の計算を同時実行できるようになると、医薬品開発を大きく加速できる可能性があり、産業応用にとっても大きな魅力である。

QM/MM 法は、QM 計算と MM 計算を用いたハイブリッド法である。MM 部分の計算については全原子 MD と同等の計算である。QM 計算は用いている近似法に基づきさまざまな手法がある。一般的に用いられているのが HF 法や DFT 法などの SCF に基づく手法である。DFT 法では、原子基底と平面波基底を用いたアプローチに大別される。誤解を恐れずに大雑把な表現をするのであれば、前者は HF 交換相互作用を用いたハイブリッド汎関数に基づく高精度な計算が可能であり、後者は高い計算効率を志向している。さらに、HF 法で得られた波動関数を出発点とし、電子相関の効果を考慮する Møller-Plesset (MP) 法や coupled cluster 法などの post-SCF 法がある。また電子励起状態を計算する TDDFT や CASPT2 法がある。このアプローチは、計算コストが高いが、非常に高精度な計算が可能である。SCF 法では、大きく分けて、行列要素の計算と、その対角化という二つの複雑な計算プロセスの繰り返しがある。したがって、strong scaling するアルゴリズムの開発が非常に困難である。近年開発されている DFTB 法や GFN-xTB 法は、前者の行列要素の計算を経験的パラメータを用いて非常に簡略化することにより、高速化を図っている（精度とパラメータ決定の困難さがトレードオフとなる）。また、後者に関しては、Car-Parrinello 法により、対角化を避けて時間発展を計算する手法がよく用いられている（短いステップ長がトレードオフとなる）。これらを組み合わせることにより、並列化による大幅な計算速度の向上は見込めるが、汎用性や精度に関しては更なる検討を要する段階である。一方、weak scaling を目指す計算（大規模系や高精度）に関しては、分子科学や物性物理などの分野で SCF 法および post-SCF 法の両者に対して、領域分割や実空間積分などに基づくさまざまな方法及びアルゴリズムの開発が行われている。したがって、それらを開発している分子・物性などの研究チームとの密接な共同研究が重要なとなるであろう。QM/MM 計算によるエネルギーおよび力の計算においては、ほぼ QM 計算が律速になり、QM-MM 相互作用も QM 計算のアルゴリズムに基づくことになる。したがって、QM/MM-MD 計算は、QM 計算の効率如何となる。一方、QM/MM 自由エネルギー構造最適化法では、QM/MM 計算と MD 計算が完全に分離される。QM/MM 計算では、従来の QM/MM 計算に加えて、MD 計算により得られた大量のトラジェクトリスナップショット構造における QM-MM 間の平均場 MM 相互作用の計算が SCF の繰り返し毎に必要となる。

後者に関しては、スナップショット構造ごとに分割した計算が行えるので、超並列化に対して非常に高いスケーリングが期待され、長時間および大規模 MM 系に対する拡張は容易であると思われる（実効性能に関しては検討の余地あり）。前者に関しては、Strong Scaling が困難であるため、後者との同時実行を少数のプロセッサを用いて行うなどのアルゴリズム開発が必要であろう。QM/MM 計算は、QM および全原子 MD 計算のハイブリッド法のため、要求計算性能は、QM 計算および全原子 MD に準じる。

FMO ベースのマルチスケールシミュレーションである FMO-DPD に関しては別記しているが、ここではその発展の方向性と課題について述べたい。現在、「富岳」を用いて全アミノ酸を対象とする χ パラメータのライブラリを整備する作業が進められている。その延長として核酸もカバーできれば汎用性が高められるが、静電相互作用の DPD フレームワークでの扱いに関する改良も必要になろう。また、生体分子の DPD では拘束条件を課すことが多いが、該当のパラメータについては未だ非経験化がなされていない。この点も要改良である。膜タンパク質などの大型系の DPD シミュレーションではセルサイズの拡大も求められるが、DPD では分割されたセル間での通信量が比較的少なくなるので混成並列化によって直截に進められると思われる。DPD は大域の配置空間の探索が得手であり、リバースマッピングにより MD や FMO の構造生成の手段ともなり得、新奇な構造での相互作用解析や新種の力場作成などにも貢献できる可能性がある。一方、自由エネルギーの評価は不得手であり、MD や FMO による情報の補完が重要である。ポスト「富岳」の時代までには、さらにメゾスケールからマクロスケールへの接続が格子ボルツマン法 (LBM) などを利用して図られるかもしれない。こうしたマルチスケール手法の研究開発では、狭い計算化学の分野に閉じず、計算工学のコミュニティとのコラボレーションも必要になってくる。

既述のように、「富岳」の時代になって MM-MD/FMO 連携シミュレーションが実用化した。FMO プログラムの高速化、大規模化系への対応、機能の強化が計算機科学の専門家とのコラボレーションを通じて ABINIT-MP などで進められており、こうした努力が 2030 年までに結実しているとすれば、アミノ酸残基の総数で数千～1 万程度（ウイルスタンパク質、膜タンパク質複合体など）、水や対イオンを含んで 5～10 万フラグメントの液滴モデルの処理がソフトウェア的には可能になっていると思われる。ポスト「富岳」の（汎用計算部の）演算性能が「富岳」の数倍～10 倍とすれば、こうした超巨大系の実行も十分に可能である。ただ、消費電力の観点などからノードあたりのメモリ量が現状から大きく改善されない（数十 GB 程度）とすれば、フラグメント間の相互作用に関する情報を収納する（本質的にフラグメント数の自乗性のサイズを持つ）配列をノード間で分散して保持管理する仕組みの導入が不可避である。多数構造のサンプル数は、系のサイズにも拠るが現在より多く、数百から数千となる（より長時間の MD 軌跡が扱える）。肝心のフラグメント間相互作用の計算結果については、巨大系/多数サンプルとなるため、有意な箇所を選択的に出力して保存する工夫も必要となる。こうした対応を取っても、データ量は一つの計算対象で容易に TB オーダーになるため、複数の対象を同時に手掛ける HPCI 課題のプロジェクトとしては PB オーダーに達する可能性がある。計算結果に対して機械学習・データ科学的な解析を行うとすると IO 効率的に SSD の利用が望ましく、処理も GPU を使う方が有利とも思われる。この意味で、名大の「不老」のように、汎用計算用部とデータ解析用部を併せ持ち、ファイル領域が共有されている複合システムもポスト「富岳」のあり方として検討する価値が十分にある。また、FMO 計算部と解析部をプロセス間通信で繋いで IO フリー (on-the-fly 処理) とするような実行形態もあり得るが、この場合には東大の「Wisteria」の複合システムが現時点での参考になる。なお、別記の FMODB のように、創薬分野ではデータベース化も重要であるため、計算結果からの重要データの抽出と蓄積もデータハンドリングの中で意識しておかなくてはいけない。複合システムについて上述したが、FMO 計算の結果データの処理以外の観点からも追記してみたい。まず、計算対象とするタンパク質のモデリングにおける AlphaFold2 [63] 等の AI 援用の構造推定システムの活用から記す。例えば、新規パンデミックの発生時、ウイルスタンパク質のアミノ酸残基のシークエンス (FASTA) のみが分かれば立体構造が構築でき、MM-MD/FMO のワークフローに素早く投入可能となり、得られた計算結果を対策づくりに供せられる。

現行の AlphaFold2 さえ、多重配列アライメント (MSA) 段階で TB オーダーのデータベース検索を行うため、円滑な実行には SSD の利用が必須となっている。今後、AlphaFold2 は抗原-抗体などのタンパク質複合体の予測にも用いられるようになると思われ、データベース量も増える可能性が高い。また、残基数の拡大が想定されるため GPU に随伴するメモリの容量も現 NVIDIA V100 の 10 倍程度は望まれる。AlphaFold2 で構築された構造を MM-MD シミュレーションにかけて緩和させる際、GPU による加速の恩恵は論を待たないであろう。汎用計算のメインシステムでの FMO 計算のための MD による構造サンプル群の調製に、GPU 系のサブシステムを用いるやり方もあり得る。GPU サブシステムを設定するもう一つの布石は、量子コンピュータのシミュレーション環境の構築にある。現時点、量子コンピュータ実機の開発はノイズの多い NISQ からエラー補正を伴う FTQC への移行段階にあるが、材料系の応用では量子化学が有力とされており、量子シミュレータベースでのアルゴリズム・ソフトウェア開発が世界的に活発に進んでいる。幸い、この分野のアプリケーションでは日本は一定の存在感を持っている。量子シミュレータでは、NVIDIA の cuQuantum [64] が注目を集めており、テンソル積の GPU での大きな加速と量子ビット数の増大が可能となる。FMO 計算を量子コンピュータ・シミュレータのみで行うか（/行い得る）どうかは現時点では未知ではあるが、FeMoco に象徴される金属クラスターを活性中心に持つ金属酵素の扱いでは混成アプローチで量子コンピュータによる一部計算は有効と思われる（主な FMO 計算は汎用のメインシステムで行い、配置数爆発が起こる活性中心は量子計算で扱う）。また、量子コンピュータによるタンパク質の畳み込みも未だ格子モデルのレベルではあるが既に提案 [65] されており、今後アルゴリズムの改良と有効相互作用パラメータの整備によって、リガンドとの結合のモデル化も含めて実用性が増す可能性がある。さらに、量子機械学習の研究開発も行われており、今後の応用が期待される。こうしたことから、2030 年頃の量子コンピュータ実機の状況に因らず、シミュレータ環境という点で GPU のサブシステムには利があると言える。

(ii) 細胞・臓器シミュレーション

これまでに連続体細胞、臓器レベルで取り上げてきた応用例では、いずれも有限差分法に基づき、カーテシアン座標系で立方体格子を用いて基礎方程式を離散化する。このような方法は、並列計算において均等に領域分割することが容易であり、台数効果が高く、計算プロセッサ数の増加にともなう計算機性能の向上の恩恵を直接受けてきた。細胞・血流シミュレーションでは、連続体に対する移流・拡散・反応式を扱い、時間積分によって生化学反応、流体・構造・膜の動力学挙動を再現する。圧力の更新には、擬似圧縮性の解法を用いることで、圧力ポアソン式の求解の際に不可欠な反復をともなう隣接通信を避ける。時間刻みを、クーラン条件を満たすように移流速度と格子解像度から決めると、粘性項と拡散項の一部には陰解法を適用する必要があるものの、計算コストのかかる反応項、移流項、構造・膜の運動学構成式には陽的な時間積分の適用が可能である。この意味で、格子ボルツマン法を用いた流体解析手法も有用である。カーテシアン格子ベースで、（ほぼ）陽解法を採用する細胞・血流シミュレーションでは、格子点当たりの演算量が多いため、今後、B/F が下がるとしても、それを捉える程度にキャッシュサイズが増えるならば、実行性能を出しやすい。一方、コードの書き方の選択肢が少なく、アルゴリズムの工夫で格段に性能向上することは期待できない。長さ 100 mm、径 100~200 μm の血管を対象とし、流速を 10^{-2} m/s と考える。血球の変形挙動までを取り扱うため、 $0.1 \mu\text{m}$ 程度の格子幅を用いるとすると、格子点数は $10^6 \times 10^3 \times 10^3 = 10^{12}$ である。クーラン数を 0.1 とすると、時間解像度は 1us と見積られ、これは

膜の動的挙動と捉えるための時間スケールと同程度である。血栓生成の時間スケールを考え、10秒の現象を扱おうとすると、全計算ステップ数は 10^7 程度である。格子点上で定義される変数は速度ベクトル、圧力、VOF関数、バルクひずみテンソル、膜ひずみテンソル、参照曲率、反応種であり、合わせて 10^2 個とするならば、必要な主記憶容量は、 8×10^2 （変数） $\times 10^2$ （自由度） $\sim 10^{15}$ Byte = 1PByteである。例えば、ノード当たりの性能、計算規模が、それぞれ、100 TFLOPS、 10^8 格子 $\times 10^2$ 変数、実効性能が40%だとすると、1ステップ当たりの計算時間は 10^{-2} s程度と見積られる。 10^2 個の変数を持つ1格子点、1ステップ当たりの演算量は約25,000であり、 10^{12} の格子点に対する107ステップ数の計算の場合、全演算数は 2.5×10^{23} 程度となる。「京」の全ノードでの実効性能が40%の性能を有しているため、1 EFLOPSの計算機を用いた場合、実効性能が400 PFLOPSとなり、約174時間の計算時間と見積もられる。理想的なキャッシュであればB/Fは $2 \times 8 \times 10^2$ （変数）/25,000 = 0.064と見積もられる。したがって、必要なメモリバンド幅は 100 （TF/s） $\times 0.064$ （B/F）= 6.4 TB/sと見積もられる。血液運動と組織領域の間質液運動および酸素輸送を考慮した脳微小循環解析に関しては、マウス脳血管を対象とした場合、大脳皮質内の動静脈ユニットを扱うためには1mm立方程度の解析領域が望ましい。膜変形を含む血球運動を捕捉するためには、最低でも $0.2\text{ }\mu\text{m}$ 程度の格子解像度が必要であることから、直交格子を用いた場合、 $5000 \times 5000 \times 5000$ 1000億程度の計算格子が必要となる。また、膜変形および酸素輸送を安定に解析するための時間刻みは1マイクロ秒程度である。必要な主記憶容量として、1格子あたりの変数がおよそ100とし、倍精度浮動小数点を用いると、およそ0.1 PByteと見積もられる。「富岳」においては、ノード当たり $100 \times 100 \times 100$ 格子での実行性能が2%程度の時に1ステップ当たりの計算時間をおよそ0.5秒と仮定し、 $50 \times 50 \times 50 = 12,500$ ノードを用いた並列計算において並列性能を85%とすると、心拍の時間スケールである1秒程度の解析に100万ステップ必要となるため160時間程度の計算時間をする。ヒトを対象とした場合、サイズの違いから解析領域を2mm立方領域とするのが望ましいため、三次元の各軸方向に倍に拡張することになり、「富岳」と同程度の単体・並列性能を持ちさらに $2 \times 2 \times 2 = 8$ 倍のノード数を持つシステムが、実用範囲の時間で解析を実施するために今後求められる。

心臓シミュレーションでは、より高精度な心臓を再現するために有限要素数の拡大とサルコメアに含まれる分子数の拡大が必要となる。MPIプロセスは分子ごとに配置するため、有限要素数と有限要素内（サルコメア内）の分子数を積算した数が総数となる。ポスト「富岳」で想定される有限要素数は約50,000要素、1要素あたり分子数は75個であることから、プロセス数は3,750,000となる。また、1分子ごとに計算する分子シミュレーションは、OpenMPによるスレッド並列を適用しており、メモリが共有利用できる範囲でスケーリングすることが確認できている。「富岳」では1CMG内12コアを利用してスレッド並列を適用し、1ノード内に4プロセスを配置している。ポスト「富岳」でも1ノード4CMGと仮定すると、937,500ノードが必要となる。CPU単体性能の面から考えると、CafeMolの計算部分は、粗視化MDの手法を用いているためオンキャッシュの計算となっているが、キャッシュのデータの再利用ができないため、良好な性能が得にくいタイプの計算となっている。プログラムコードの改善は実施しているが、キャッシュ性能の強化やコンパイラの強化による性能改善も期待したい。また、ソルバーの部分は伝統的なリストアクセスによるベクトル参照と行列積の計算が主体であるため、このような処理の性能の向上が期待できる、ハードウェアとコンパイラが必要である。次にメモリ性能について述べる。UT-Heartではメモリバンド幅への要求は低く、1プロセスあたりの最大メモリ使用量についても「富岳」の実績から9.6 GiB程度を想定しており、「富岳」と同等以上の性能が保持できれば問題ない。最後にファイル

I/Oについて述べる。UT-Heartではシミュレーションの継続性を保証する目的で時間ステップごとに各プロセスが分子の状態をファイル出力する。そのため、I/Oを制御するノード（「富岳」ではSIOノード）にプロセス数分の負荷が同時にかかるため、ポスト「富岳」では「富岳」以上にI/O負荷に対する耐性が必要となる。またファイルを格納するディスク容量については心臓1秒分の拍動を計算した場合、1プロセスが約240MBのファイルを生成するため、全プロセスで総計すると約225TBの容量が必要となる。また「富岳」より進めている”インシリコ心疾患データベースの作成と公開”についてもデータの拡充を進めるため、多くのディスク容量を必要とする。「富岳」の100倍規模となる400万ケース規模のデータベースを想定し、1ケースあたり平均値25GBで試算すると最大100PBのディスク容量が必要となる。データサイズについてはシミュレーション側でも削減に向けた取り組みが必要不可欠となるが、合わせてデータを別媒体に転送するための接続性も検討する必要がある。

超音波治療に用いられるコードの「京」での実効性能は20%程度である。したがって、超音波治療で用いられる周波数1~2MHzの超音波の軟組織中の波長が、0.75~1.5mmであり、これを解像するのに0.05mm程度の格子幅を用いるとすると、「京」では、 $10,000^3$ 格子点(5.12×10^{11} 自由度)の計算が可能であり、400mmの立方体領域中の超音波の伝播が1日程度で計算できる。しかし、軟らかい生体組織の温度変化で生じる微妙な剛性の変化を捉えるためには、音波である縦波に比べて2桁程度遅い伝播速度を持つ横波を再現する必要があり、より高い解像度と長時間の時間積分が必要になり、最低でも「京」での計算の1,000倍の自由度の計算が求められ、「京」の100倍(1EFLOPS)の計算機を利用した場合、10日程度で計算ができる見積もりとなる。また、マイクロカプセルとの干渉によって生じる超音波音場の解析を行うためには、時間スケールが小さいマイクロカプセルの時間積分を行うためには計算コストが通常の超音波伝播計算に比べて数十倍程度必要となるため、現状の100倍程度の計算処理速度が将来的に必要となると考えられる。超音波診断装置では、超音波治療で用いられる周波数1~2MHzよりも高い3~8MHzの周波数が使用される。より周波数の高い超音波の伝播シミュレーションを行うためには空間解像度を上げる必要があり、例えば、周波数が4倍高くなると、その計算規模は256倍以上の規模になる。また、AIによる学習を行うためには、この規模のシミュレーションを膨大に実施し、その計算結果データをもとにした学習処理が必要となる。

4.6.5 公的な意見交換とそのサマリー

2017年9月19日~21日の日本生物物理学会第55回年会における共催シンポジウムにおいて、ポスト「京」重点課題1 生体分子システムの機能制御による革新的創薬基盤の構築では、「ハイパフォーマンス・コンピューティング(HPC)による次世代創薬計算技術」について、重点課題1のサブ課題A(ポスト「京」でのMD高度化とアルゴリズム深化)、サブ課題B(次世代創薬計算技術の開発)、サブ課題C(創薬ビッグデータ統合システムの開発)からそれぞれ研究成果について紹介が行われた。

2018年5月23日~26日の第59回日本神経学会学術集会シンポジウムにおいて、「ゲノムから孤発性神経疾患が解明できるか?」をテーマに意見交換が行われ、奥野恭史は、データベースの観点から「臨床ゲノム情報統合データベース“MGeND”」について報告した。

2019年10月22日~10月24日のCBI学会2019年大会では、プレナリー講演で「次世代の構造インフォマティクスに向けた取り組み」について、奥野恭史、福澤 薫、池口 満徳が現状と今後の取り組みに

について議論があった。

2020年5月15日には、「富岳」の特長とSociety5.0への貢献 / 「富岳」を利用した新型コロナウイルス研究をテーマに理化学研究所主催の記者勉強会が開催され、奥野恭史は、「「富岳」による新型コロナウイルスの治療薬候補同定」について「富岳」がいかにして貢献しうるのかを解説し、具体的な内容と進捗状況を報告した。

2021年9月27日の日本オミックス医学会シンポジウムでは、「AI創薬の展開」をテーマに奥野恭史が、「AIによる薬効・ADMET予測と企業間連合学習」について薬効・ADMET予測のためのAI開発を例にAI技術を紹介し、企業間でのデータ共有について議論が行われた。また、本間光貴が企業内データを活用した統合創薬AIプラットフォームの構築について産学連携によるライフインテリジェンスコンソーシアム（LINC）の活動について紹介した。

4.6.6 ロードマップ

参考文献

- [1] Taro Kariya, Takumi Washio, Jun ichi Okada, Machiko Nakagawa, Masahiro Watanabe, Yoshimasa Kadooka, Shunji Sano, Ryozo Nagai, Seiryo Sugiura, and Toshiaki Hisada. Personalized perioperative multi-scale, multi-physics heart simulation of double outlet right ventricle. *Ann Biomed Eng*, Vol. 48, No. 6, pp. 1740–1750, 2020.
- [2] Kazunori Yoneda, Jun ichi Okada, Masahiro Watanabe, Seiryo Sugiura, Toshiaki Hisada, and Takumi Washio. A multiple step active stiffness integration scheme to couple a stochastic cross-bridge model and continuum mechanics for uses in both basic research and clinical applications of heart simulation. *Front Physiol.*, Vol. 12, p. 712816, 2021.
- [3] Kazuhiro Takemura, Chika Sato, and Akio Kitao. Coldock: Concentrated ligand docking with all-atom molecular dynamics simulation. *The Journal of Physical Chemistry B*, Vol. 122, No. 29, pp. 7191–7200, 2018. doi: 10.1021/acs.jpcb.8b02756.
- [4] Hideaki Fujitani, Yoshiaki Tanida, and Azuma Matsuura. Massively parallel computation of absolute binding free energy with well-equilibrated states. *Physical Review E*, Vol. 79, No. 2, p. 021914, 2009. PRE.
- [5] Mitsugu Araki, Narutoshi Kamiya, Miwa Sato, Masahiko Nakatsui, Takatsugu Hirokawa, and Yasushi Okuno. The effect of conformational flexibility on binding free energy estimation between kinases and their inhibitors. *Journal of Chemical Information and Modeling*, Vol. 56, No. 12, pp. 2445–2456, 2016. doi: 10.1021/acs.jcim.6b00398.
- [6] Yoshio Okiyama, Hirofumi Watanabe, Kaori Fukuzawa, Tatsuya Nakano, Yuji Mochizuki, Takeshi Ishikawa, Kuniyoshi Ebina, and Shigenori Tanaka. Application of the fragment molecular orbital method for determination of atomic charges on polypeptides. ii. towards an improvement of force fields used for classical molecular dynamics simulations. *Chemical Physics Letters*, Vol. 467, No. 4, pp. 417–423, 2009.

- [7] Koichiro Kato, Tomohide Masuda, Chiduru Watanabe, Naoki Miyagawa, Hideo Mizouchi, Shumpei Nagase, Kikuko Kamisaka, Kanji Oshima, Satoshi Ono, Hiroshi Ueda, Atsushi Tokuhisa, Ryo Kanada, Masateru Ohta, Mitsunori Ikeguchi, Yasushi Okuno, Kaori Fukuzawa, and Teruki Honma. High-precision atomic charge prediction for protein systems using fragment molecular orbital calculation and machine learning. *Journal of Chemical Information and Modeling*, Vol. 60, No. 7, pp. 3361–3368, 2020. doi: 10.1021/acs.jcim.0c00273.
- [8] Chiduru Watanabe, Hirofumi Watanabe, Kaori Fukuzawa, Lorien J. Parker, Yoshio Okiyama, Hitomi Yuki, Shigeyuki Yokoyama, Hirofumi Nakano, Shigenori Tanaka, and Teruki Honma. Theoretical analysis of activity cliffs among benzofuranone-class pim1 inhibitors using the fragment molecular orbital method with molecular mechanics poisson–boltzmann surface area (fmo+mm-pbsa) approach. *Journal of Chemical Information and Modeling*, Vol. 57, No. 12, pp. 2996–3010, 2017. doi: 10.1021/acs.jcim.7b00110.
- [9] Ikki Yasuda, Katsuhiro Endo, Eiji Yamamoto, Yoshinori Hirano, and Kenji Yasuoka. Differences in ligand-induced protein dynamics extracted from an unsupervised deep learning approach correlate with protein–ligand binding affinities. *Communications Biology*, Vol. 5, No. 1, p. 481, 2022.
- [10] The Virtual Cell. <http://www.nrcam.uchc.edu/>.
- [11] Virtual Physiome Project. <http://www.vph-noe.eu/>.
- [12] RICS プロジェクト. http://cellsim.riken.jp/research_e.html.
- [13] 高木周. 生命と科学, 第5章「人体シミュレーション」. 岩波書店, 2012.
- [14] C. Pozrikidis. *Modeling and simulation of capsules and biological cells*. CRC Press, 2003.
- [15] H.-J. Bungartz and M. Schäfer. *Fluid-structure interaction: modelling, simulation, optimisation*, Vol. 53 of *Lecture Notes in Comput. Sci. Engrg*. Springer Science & Business Media, 2006.
- [16] C. Pozrikidis, editor. *Computational hydrodynamics of capsules and biological cells*. CRC Press, 2010.
- [17] A. Rahimian, et al. Petascale direct numerical simulation of blood flow on 200k cores and heterogeneous architectures. In *Proc. of International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–11, 2010.
- [18] M. Bernaschi, et al. Petaflop biofluidics simulations on a two million-core system. In *Proc. of International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 4:1–4:12, 2011.
- [19] L. Grinberg, et al. A new computational paradigm in multiscale simulations: Application to brain blood flow. In *Proc. of International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–12, 2011.
- [20] K. Sugiyama, et al. A full eulerian finite difference approach for solving fluid–structure coupling problems. *Journal of Computational Physics*, Vol. 230, No. 3, pp. 596–627, 2011.
- [21] S. Ii, et al. A full eulerian fluid-membrane coupling method with a smoothed volume-of-fluid approach. *Communications in Computational Physics*, Vol. 12, No. 2, p. 544–576, 2012.

- [22] S. Takagi, et al. A review of full eulerian methods for fluid structure interaction problems. *J. Appl. Mech.*, Vol. 79, No. 1, p. 010911, 2012.
- [23] Sugiyama K., et al. Development of explicit eulerian finite difference solver for large-scale fluid-structure interaction systems. ハイパフォーマンスコンピューティングと計算科学シンポジウム論文集, 第2012巻, pp. 153–162, jan 2012.
- [24] Takumi Washio, Jun ichi Okada, Akihito Takahashi, Kazunori Yoneda, Yoshimasa Kadooka, Seiryo Sugiura, and Toshiaki Hisada. Multiscale heart simulation with cooperative stochastic cross-bridge dynamics and cellular structures. *Multiscale Model. Simul.*, Vol. 11, p. 965–999, 2013.
- [25] Takumi Washio, Seiryo Sugiura, Ryo Kanada, Jun ichi Okada, and Toshiaki Hisada. Coupling langevin dynamics with continuum mechanics: Exposing the role of sarcomere stretch activation mechanisms to cardiac function. *Front Physiol.*, Vol. 9, p. 333, 2018.
- [26] Takumi Washio, Ryo Kanada, Xiaoke Cui, Jun ichi Okada, Seiryo Sugiura, Shoji Takada, and Toshiaki Hisada. Semi-implicit time integration with hessian eigenvalue corrections for a larger time step in molecular dynamics simulations. *J Chem Theory Comput.*, Vol. 17, No. 9, pp. 5792–5804, 2021.
- [27] Takumi Washio, Xiaoke Cui, Ryo Kanada, Jun ichi Okada, Seiryo Sugiura, Yasushi Okuno, Shoji Takada, and Toshiaki Hisada. Using incomplete cholesky factorization to increase the time step in molecular dynamics simulations. *J. Comput. Appl. Math.*, Vol. 415, p. 114519, 2022.
- [28] Satoshi Ii, Hiroki Kitade, Shunichi Ishida, Yohsuke Imai, Yoshiyuki Watanabe, and Shigeo Wada. Multiscale modeling of human cerebrovasculature: A hybrid approach using image-based geometry and a mathematical algorithm. *PLoS computational biology*, Vol. 16, No. 6, p. e1007943, 2020.
- [29] Shunichi Ishida, Hiroki Kitade, Satoshi Ii, Naoki Takeishi, Yoshiyuki Watanabe, Yohsuke Imai, and Shigeo Wada. Development of a numerical model of whole-scale cerebral circulation. *Journal of Cerebral Blood Flow & Metabolism*, Vol. 39, No. 6, p. 615, 2019.
- [30] F. A. Duck, et al. *Ultrasound in medicine*. CRC Press, 1998.
- [31] J. E. Kennedy, et al. High intensity focused ultrasound: surgery of the future? *Br. J. Radiol.*, 2014.
- [32] F. Wu, et al. Circulating tumor cells in patients with solid malignancy treated by high-intensity focused ultrasound. *Ultrasound in Med. Biol.*, Vol. 30, No. 4, pp. 511–517, 2004.
- [33] J. E. Kennedy. High-intensity focused ultrasound in the treatment of solid tumours. *Nat. Rev. Cancer*, Vol. 5, No. 4, pp. 321–327, 2005.
- [34] G. K. Hesley, et al. A clinical review of focused ultrasound ablation with magnetic resonance guidance: an option for treating uterine fibroids. *Ultrasound Q.*, Vol. 24, No. 2, pp. 131–139, 2008.
- [35] <http://www.healthcare.philips.com/main/products/mri/systems/sonalleve/>.

- [36] http://japan.gehealthcare.com/cwcjapan/static/rad/mri/MRgFUS_ExAblate2000/.
- [37] 佐々木明. バイオスーパーコンピューティングが拓くライフサイエンスの未来：超音波治療の推進および治療機器開発に欠かせない生体の音響的シミュレーション研究. *BioSupercomputing Newsletter*, Vol. 3, pp. 4–5, 2010.
- [38] J. Sun and K. Hynynen. Focusing of therapeutic ultrasound through a human skull: a numerical study. *J. Acoust. Soc. Am.*, Vol. 104, No. 3, pp. 1705–1715, 1998.
- [39] M. Fink. Time reversal of ultrasonic fields. i. basic principles. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, Vol. 39, No. 5, pp. 555–566, 1992.
- [40] M. Fink, et al. Time-reversal acoustics in biomedical engineering. *Annu. Rev. Biomed. Engrg.*, Vol. 5, No. 1, pp. 465–497, 2003.
- [41] K. Okita, et al. Numerical simulation of the tissue ablation in high-intensity focused ultrasound therapy with array transducer. *Int. J. Numer. Meth. Fluids*, Vol. 64, No. 10-12, pp. 1395–1411, 2010.
- [42] K. Okita, et al. Development of high intensity focused ultrasound simulator for large-scale computing. *Int. J. Numer. Meth. Fluids*, Vol. 65, No. 1-3, pp. 43–66, 2011.
- [43] Yoshio Nishimoto, Dmitri G. Fedorov, and Stephan Irle. Density-functional tight-binding combined with the fragment molecular orbital method. *Journal of Chemical Theory and Computation*, Vol. 10, No. 11, pp. 4801–4812, 2014. doi: 10.1021/ct500489d.
- [44] Yoshifumi Nishimura and Hiromi Nakai. Hierarchical parallelization of divide-and-conquer density functional tight-binding molecular dynamics and metadynamics simulations. *Journal of Computational Chemistry*, Vol. 41, No. 19, pp. 1759–1772, 2020.
- [45] FMODB.
- [46] Daisuke Takaya, Chiduru Watanabe, Shunpei Nagase, Kikuko Kamisaka, Yoshio Okiyama, Hirotomo Moriwaki, Hitomi Yuki, Tomohiro Sato, Noriyuki Kurita, Yoichiro Yagi, Tatsuya Tagagi, Norihito Kawashita, Kenichiro Takaba, Tomonaga Ozawa, Midori Takimoto-Kamimura, Shigenori Tanaka, Kaori Fukuzawa, and Teruki Honma. Fmodb: The world's first database of quantum mechanical calculations for biomacromolecules based on the fragment molecular orbital method. *Journal of Chemical Information and Modeling*, Vol. 61, No. 2, pp. 777–794, 2021. doi: 10.1021/acs.jcim.0c01062.
- [47] Kaori Fukuzawa, Koichiro Kato, Chiduru Watanabe, Yusuke Kawashima, Yuma Handa, Ami Yamamoto, Kazuki Watanabe, Tatsuya Ohyama, Kikuko Kamisaka, Daisuke Takaya, and Teruki Honma. Special features of covid-19 in the fmodb: Fragment molecular orbital calculations and interaction energy analysis of sars-cov-2-related proteins. *Journal of Chemical Information and Modeling*, Vol. 61, No. 9, pp. 4594–4612, 2021. doi: 10.1021/acs.jcim.1c00694.
- [48] Chiduru Watanabe, Hirofumi Watanabe, Yoshio Okiyama, Daisuke Takaya, Kaori Fukuzawa, Shigenori Tanaka, and Teruki Honma. Development of an automated fragment molecular orbital (fmo) calculation protocol toward construction of quantum mechanical calculation database for large biomolecules. *Chem-Bio Informatics Journal*, Vol. 19, pp. 5–18, 2019.

- [49] Koji Okuwaki, Yuji Mochizuki, Hideo Doi, and Taku Ozawa. Fragment molecular orbital based parametrization procedure for mesoscopic structure prediction of polymeric materials. *The Journal of Physical Chemistry B*, Vol. 122, No. 1, pp. 338–347, 2018. doi: 10.1021/acs.jpcb.7b08461.
- [50] Hideo Doi, Koji Okuwaki, Yuji Mochizuki, Taku Ozawa, and Kenji Yasuoka. Dissipative particle dynamics (dpd) simulations with fragment molecular orbital (fmo) based effective parameters for 1-palmitoyl-2-oleoyl phosphatidyl choline (popc) membrane. *Chemical Physics Letters*, Vol. 684, pp. 427–432, 2017.
- [51] Hiroki Tanaka, Tatsunari Takahashi, Manami Konishi, Nae Takata, Masaki Gomi, Daiki Shirane, Ryo Miyama, Shinya Hagiwara, Yuki Yamasaki, Yu Sakurai, Keisuke Ueda, Kenjirou Higashi, Kunikazu Moribe, Eiji Shinsho, Ruka Nishida, Kaori Fukuzawa, Etsuo Yonemochi, Koji Okuwaki, Yuji Mochizuki, Yuta Nakai, Kota Tange, Hiroki Yoshioka, Shinya Tamagawa, and Hidetaka Akita. Self-degradable lipid-like materials based on “hydrolysis accelerated by the intra-particle enrichment of reactant (hyper)” for messenger rna delivery. *Advanced Functional Materials*, Vol. 30, No. 34, p. 1910575, 2020.
- [52] Koji Okuwaki, Hideo Doi, Kaori Fukuzawa, and Yuji Mochizuki. Folding simulation of small proteins by dissipative particle dynamics (dpd) with non-empirical interaction parameters based on fragment molecular orbital calculations. *Applied Physics Express*, Vol. 13, No. 1, p. 017002, 2019.
- [53] K. Sasaki, et al. Effect of echo-guided high-intensity focused ultrasound ablation on localized experimental tumors. *J. Vater. Med. Sci.*, Vol. 68, No. 10, pp. 1069–1074, 2006.
- [54] 中村弘文ほか. 超音波 CT を用いた超音波診断治療の基礎研究. 第 11 回日本超音波治療研究会, 2012.
- [55] Y. Liu, et al. Numerical study on the shape oscillation of an encapsulated microbubble in ultrasound field. *Phys. Fluids*, Vol. 23, No. 4, p. 041904, 2011.
- [56] 杉山和靖ほか. 超音波音場における膜被包性マイロバブルの表面安定性の数値解析. 超音波テクノ, Vol. 24, pp. 90–94.
- [57] 克浩遠藤, 大輔湯原, 顯治泰岡. Md-gan : 機械学習による分子動力学シミュレーションの高速化. アンサンブル, Vol. 21, No. 1, pp. 39–44, 2019.
- [58] Shigenori Tanaka, Shusuke Tokutomi, Ryo Hatada, Koji Okuwaki, Kazuki Akisawa, Kaori Fukuzawa, Yuto Komeiji, Yoshio Okiyama, and Yuji Mochizuki. Dynamic cooperativity of ligand–residue interactions evaluated with the fragment molecular orbital method. *The Journal of Physical Chemistry B*, Vol. 125, No. 24, pp. 6501–6512, 2021. doi: 10.1021/acs.jpcb.1c03043.
- [59] Koji Okuwaki, Kazuki Akisawa, Ryo Hatada, Yuji Mochizuki, Kaori Fukuzawa, Yuto Komeiji, and Shigenori Tanaka. Collective residue interactions in trimer complexes of sars-cov-2 spike proteins analyzed by fragment molecular orbital method. *Applied Physics Express*, Vol. 15, No. 1, p. 017001, 2021.
- [60] Sona Saitou, Jun Iijima, Mayu Fujimoto, Yuji Mochizuki, Koji Okuwaki, Hideo Doi, and Yuto Komeiji. Application of tensorflow to recognition of visualized results of fragment molecular

- orbital (fmo) calculations; b_c. *Chem-Bio Informatics Journal*, Vol. 18, pp. 58–69, 2018.
- [61] J. E. Sadler. Von Willebrand factor, ADAMTS13, and thrombotic thrombocytopenic purpura. *Blood*, Vol. 112, No. 1, pp. 11–18, 2008.
- [62] K. Sugiyama, et al. A priori modeling of the acoustic boundary layer effect on the heat source in ultrasound. *J. Biomech. Engrg. Sci.*, Vol. 7, No. 1, pp. 84–101, 2012.
- [63] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michał Zielinski, Martin Steinegger, Michalina Pacholska, Tamás Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with alphafold. *Nature*, Vol. 596, No. 7873, pp. 583–589, 2021.
- [64] NVIDIA.
- [65] Alejandro Perdomo-Ortiz, Neil Dickson, Marshall Drew-Brook, Geordie Rose, and Alán Aspuru-Guzik. Finding low-energy conformations of lattice protein models by quantum annealing. *Scientific Reports*, Vol. 2, No. 1, p. 571, 2012.
- [66] S. Tanaka, et al. Electron-correlated fragment-molecular-orbital calculations for biomolecular and nano systems. *Phys. Chem. Chem. Phys.*, Vol. 16, No. 22, pp. 10310–10344, 2014.
- [67] D. G. Fedorov, et al. Exploring chemistry with the fragment molecular orbital method. *Phys. Chem. Chem. Phys.*, Vol. 14, No. 21, pp. 7562–7577, 2012.

4.7 設計・製造

本節では設計・製造分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.7.1 航空機分野

4.7.1.1 分野の概要

ものづくりにおける低コスト化施策の費用対効果は上流に行くほど大きくなる（ものづくり白書等）。特に航空機開発では開発下流段階での航空機実機フライト試験など実機を製造し飛行試験を行わないと評価できない設計上の重要課題が取り除けておらず、設計プロセスに大きなリスクを抱えている。このことは、開発コスト増大のみならず、開発そのものを脅かす開発リスクや開発期間の大幅な遅延要素に直結している。それに伴い、民間航空機開発を推進している国内航空機開発メーカーの強いニーズとして、実機フライト試験への依存度を格段に低減する革新的設計技術により、航空機の『開発期間の短縮』、開発の『コスト低減』を図ると共に『開発リスク』を格段に低減することが望まれている。と同時に、『開発上流での最適化レベルを格段に向上』させることで、競合他機が容易にキャッチアップできない差別化技術を有した製品をタイムリーに市場提供することが望まれている。これらの革新的設計技術や差別化技術の開発は、日本の航空機産業の国際的優位性に直結する最重要課題である。

自動車などと比較して、航空機はレイノルズ数が 10 の 7 乗オーダーと非常に高いことからレイノルズ数条件を一致させた風洞試験による評価は困難であり、航空機開発における上記課題に対し、高度な計算科学（数値シミュレーション科学）による予測評価が期待されている。また計算科学による予測評価を用いて、実機を製造する前の設計初期段階で設計上の重要課題を評価可能とすることで、実機フライト試験を代替する新たな設計プロセス技術が強く望まれている。しかしながら現状、計算科学により予測評価できるのは、航空機周りの流れが定常的な巡航条件付近に限られており、離着陸性能を決める低速・高揚力条件や高速飛行性能を決める高速飛行時の遷音速バフェット（衝撃波による振動現象）境界の予測評価、フラッター（流体と機体構造の連成現象）境界の予測評価、離着陸時の騒音の予測評価など、流れが強い非定常性を持つ条件に対する計算科学による予測評価は困難なのが現状である。

このような現状を打破するために、高度な数値シミュレーション科学、具体的には解析対象を準直接的に扱うことで流れの非定常性を高忠実に再現する数値シミュレーション手法（LES: Large-Eddy Simulation）の実用化が期待されている、航空機の飛行条件（フライトエンベロープ：航空機の飛行可能な速度や荷重や高度の範囲）に対して、上述した様に、これまで極一部の領域・現象のみしか予測評価が行えなかつたが、LES を用いることで、フライトエンベロープ全体への予測可能範囲の拡大が期待されている。またこのような期待は日本国内のみならず、世界的にも着目され、世界各国で熾烈な研究開発競争が進められている。具体的には、アメリカ航空宇宙局（NASA）は、「CFD Vision 2030: A Path to Revolutionary Computational Aerosciences」という 2030 年に航空宇宙分野に革新を起こすビジョンのレポートをまとめており、高度な数値シミュレーション科学による航空機フライトエンベロープの予測可能範囲の拡大が 6 つの国家的プロジェクトの 1 つにあげられており、2022 年 6 月には本課題に関する国際

ワークショップが開催される。加えて、航空機分野を含むより広い航空宇宙分野では、上述したような流れの非定常性を準直接的に扱う数値シミュレーションの基盤アプリが世界的に見ても確立されていないのが現状であり、航空宇宙分野の広い範囲において、学術研究から産業応用・実用までの使用に耐えうる高忠実なHPC基盤アプリの構築も世界的な関心事項となっている。さらに航空機飛行条件の予測範囲の拡大が達成された先には、フラッター等の流体と構造の連成問題や、流体と機体運動の連成問題など、フライトエンベロープで起こりうる様々なシナリオに応えられる多角的な機体設計への取り組みも必要となる。

4.7.1.2 次世代に解決すべき課題とその解決に必要な計算手法・アプリケーション

現在推進されている「富岳」成果創出加速プログラム「航空機フライト試験を代替する近未来型設計技術の先導的実証研究」では、これまで予測評価が困難であった実際の飛行レイノルズ数（ 10^7 乗オーダー）における離着陸性能を決める低速・高揚力条件の予測評価が可能となってきているが、これは航空機飛行条件のある一端に留まる。次世代に解決すべき課題は、高度な計算科学による予測可能範囲を航空機の全飛行領域、すなわち航空機フライトエンベロープ全体に拡大することである。それを実現するためには、特に未だ解決されていない重要課題、高速飛行性能を決める高速飛行条件における遷音速バフェット境界ならびにフラッター境界の予測評価、更には離着陸時の機体騒音の予測評価が不可欠となる。これらの課題を解決することにより、高度な計算科学が実機フライト試験の代替となり、デジタルトランスフォーメーションによる航空機設計、すなわち開発下流段階での実機フライト試験など実機を製造ならびに飛行試験を行わないと評価できない設計課題を事前に評価することで、開発リスクの根本的な解決が期待できる。さらには、航空機フライトエンベロープ全体の予測評価が可能となった先には、航空機の飛行条件全体において性能や安全性に関する要求を満足するべく、これらを統合的に評価する多目的・多設計変数の機体設計が期待される。

上記の課題の解決を実現するためのアプリケーションに必要な学術的課題は大きく以下の4つとなる。
①航空機の細かい複雑形状ならびにそれらの形状が移動・変形する場合にも対応可能な解析、②高レイノルズ数条件の準直接解析手法、③衝撃波を含む圧縮性流れの高忠実解析を実現する計算スキーム、および④多目的・多設計変数最適化問題で大域的最適解を高速に探索するアルゴリズムである。高忠実・圧縮性流体準直接解析基盤アプリFFVHC-ACEでは、①に対しては形状STLデータのみで細かい複雑形状に対しても自動かつ高速に格子生成を可能とする階層型等間隔直交格子法を採用している。また今後の開発項目となるが、物体境界との高速な交差判定を可能とすることで移動・変形を含む高速な解析を可能とすることを想定する。②では高レイノルズ数条件の準直接解析を可能とする壁面モデルLES(壁面近傍内層乱流のモデル化)を独自に開発し、実装している。物体の移動・変形に関しては今後の学術的開発項目となるが、物体境界での保存性を向上させることで、既存のシャープインターフェイス法で問題となる圧力振動問題を解決することを想定する。③ではこれまでの圧縮性流体解析の計算手法で問題となっていた数値散逸を用いずとも安定かつ高忠実な解析を可能とする計算手法(KEEPスキーム)を用いる。本手法は、有限差分法ベースで対流項の物理現象を高忠実に再現することで「安定性」と「非散逸性」を同時に満足する計算手法である。衝撃波のある流れへの適用が今後の学術的開発項目となる。④では設計変数(形状パラメータ)に対する目的関数(機体性能や安全性)のブラックボックスな構造を限られた

評価データから代数式近似し、近似モデルベースで最適解の在処を瞬時に探索する。また今後の開発項目となるが、機械学習や圧縮センシングなどにより設計変数空間を次元縮約することで、多設計変数問題に対する近似モデルの適合性と正確性を上げることを想定する。

上記の次世代に解決すべき課題（遷音速バフェット境界の予測評価、フラッター境界の予測評価、機体騒音の予測評価）の解析で必要となる格子点数はおよそ 1.5 兆点、時間ステップ数は 2000 万ステップとなる。航空機全機複雑形状の準直接解析（壁面モデル LES）を考え、10 の 7 乗オーダーとなる航空機実機飛行レイノルズ数（現在「富岳」で計算できている風洞試験スケールに対して 10 倍）で境界層厚みが約半分となる。さらに翼前縁の薄い境界層を壁面モデル LES として解像するとし、現状の「富岳」での解析規模（1000 億点規模）の 15 倍程度の 1.5 兆点程度の格子点数が必要となる。さらには上記の次世代に解決すべき課題は、乱流境界層現象の時間スケールと現象の時間スケールに差が大きい時定数の長い問題であるため、現状の 10 倍程度の物理時間の解析が必要となる。さらに格子細分化により時間ステップ幅が現状の約半分になることを考え、現状の「富岳」での時間ステップ数（100 万ステップ程度）の 20 倍を見込む。これらの解析に必要となる演算性能規模は、想定される格子点数（1.5 兆点）に対して現状の「富岳」での計算（1000 億点、13000 ノード程度）から外挿して算出すると、約 600PFLOPS（ピーク演算性能、倍精度）となる。これは次世代機のピーク演算性能が 2-20EFLOPS となることを想定すると、次世代機全系の約 3-30% 程度になると概算でき、次世代に解決すべき課題の解析が実現可能な範囲だと推算される。また必要となる総メモリサイズは、現状の「富岳」での計算から外挿すると 6.4PByte（「富岳」20 万ノード分に相当）となる。ストレージ量は、1 ケース当たり最低限約 500TByte となる（Restart 用のスナップショットデータを 2 つ（120TB）、乱流統計量データ、格子データ、その他出力データなどで約 500TB）。また非定常準直接流体解析を実施するので、少なくとも 1 ケースの非定常流体解析が望まれる。その場合は物体の周りの 3 次元非定常流れデータの保存として別途 2PByte 程度の利用が必要となる（20TB × 100 スナップショット）。ケース数は、最低 32 ケース以上の解析が必要である。内訳は、遷音速バフェット境界の予測評価では 4 マッハ数条件 × 3 迎角=12 ケース、機体騒音の予測評価では 2 マッハ数条件 × 4 迎角=8 ケース、フラッターの境界の予測評価では 4 マッハ数条件 × 3 Flutter Speed Index=12 ケースである。計算時間は、1 ケース 1000 時間程度の概算である。これは時間ステップ数が現状の「富岳」での計算 20 倍に増えているため、現状（50-100 時間/ケース）の 20 倍の演算時間を要すると概算した。ただしこの数字は演算性能（ノード割り当て数）を増やすことによって短縮可能である。メモリアクセスがボトルネックの計算となるので、最低限 B/F 比を現状の 0.4 程度を維持、できれば 1 程度が望ましい。メモリ容量/F（現 32GByte/3TFLOPS）も少なくとも現状維持、もしくは増加が望ましい。並列計算における通信としては、我々のアプリでは隣接通信が中心のため、ネットワークの通信性能とノードの演算性能の比（B/F）の性能バランスも重要であり、「京」並が望ましい。また次世代の開発上流での最適化レベルの格段の向上では、準直接解析を用いた多目的・多設計変数最適化として、近似モデルの構築に必要とされる評価データ（形状を変えた解析ケース）数を、現状（100 変数問題で 100-1000 ケース）の 10 倍程度（1000 変数問題で 1000-10000 ケース）と概算している。現状の「富岳」での中位ケース（100 億点規模で 1 ケース 1300 ノード）を想定すると、1 ケースあたり約 4PFLOPS（ピーク演算性能、倍精度）、1000-10000 ケースで合計 400-4000PFLOPS と「富岳」の演算性能レベルを超えるため、次世代機レベルの演算性能が必須である。

4.7.1.3 ロードマップ

航空機分野におけるロードマップを図 4.21 に示す。

年代 課題	2026 ~	2028 ~	2030 ~	2032 ~	2034 ~	2036 ~
航空機	100～1000億格子規模の壁面モデルLES解析 (遷音速バフェット、騒音解析)					

← →

1～10兆格子規模の壁面モデルLES解析
(飛行領域全体解析(フラッター、運動解析を含む))

← →

100億格子規模・1000-10000個体の壁面モデルLESを用いた
飛行領域全体を考慮した統合的多目的・多設計変数設計最適化

← →

図 4.21 ロードマップ (航空機分野)

4.7.2 構造分野

4.7.2.1 分野の概要

現在、産業界において製品およびその生産や維持工程を設計し評価する際に、シミュレーション技術が必要不可欠なものとなっている。例えば自動車の衝突解析や騒音・振動解析、航空機や船舶の構造健全性評価、大規模発電プラントや超高層建築物、橋梁、高速道路などの耐震解析、さらにものづくりにおける設計・生産過程、特に塑性加工やプレス成型、3D プリンティングを再現するためのバーチャルマニュファクチャリングなど、多種多様なものづくりや設計のシーンでシミュレーション技術が活用されている。このようなシミュレーション技術は CAE とも呼ばれる。有限要素法 (FEM : Finite Element Method)[1] に基づく構造解析ソルバーは、人工物や構造物の力学的特性や応答性、安全性・健全性などを評価するために用いられ、CAE の最重要コンポーネントでもある。

FEM ベースの構造解析ソルバーは、陽解法に基づく衝撃解析やほぼ非圧縮性状態の塑性加工解析のような特定の現象を解くための専用ソフトウェアも存在するものの、多様な解析機能を集積した汎用解析ソフトが一般的である。汎用解析ソフトは、機械・電気・化学・建築土木をはじめ、船舶海洋・航空宇宙・原子力・医療・環境などほとんどあらゆる分野で日常的に用いられている。また、大企業や研究所を中心に、解析モデルの大規模化・詳細化と解析の高速化を目的として、PC クラスタや中規模スパコンなどの HPC 環境が導入されており、さらに HPC クラウド環境の活用も進んでいる。

我が国の構造解析分野では、1960 年代後半～1970 年代より使われ出した NASTRAN、ABAQUS、ANSYS や LS-Dyna など欧米発の商用ソフトが、継続的に改良を重ねられ現在に至るまで幅広く用いられてきた。そして、産業界においても解析規模の大規模化、解析の高精度化・高速化への強いニーズを

背景に、HPC環境上で構造解析を行う機会が増えてきている。しかし、6.4.2.2項で述べるようにFEMベースの構造解析ソルバーの並列化については技術的な難易度が高く、上述した欧米製の汎用構造解析ソフトでは数万コアオーダーの超並列環境への対応が遅れており、産業界におけるさらなるHPC技術導入の阻害要因となっている。

一方、最新HPC環境での活用を前提に、超並列技術に対応した国産の構造解析ソフトの研究開発が進んでいる。1997年～2002年に実施された日本学術振興会未来開拓学術研究推進事業「計算科学」分野のプロジェクトにおいて開発され、2002年に無料公開されたオープンソースの汎用並列FEM解析システムADVENTUREと、それを基に開発された商用汎用構造解析ソフトADVENTUREClusterがその代表例である。ADVENTUREは2002年以降も我が国の計算科学に関する国家プロジェクトにおいて継続的に活用されており、ADVENTUREClusterは我が国の幅広い産業分野において上述した欧米製ソフトと並んで標準的に使われるようになっている。また、科学技術庁の計算科学プロジェクトにおいて開発が開始され、順次名称が変更されながら、現在も研究開発が進められ産業界に使われているFEMベースの並列構造解析ソフトにはFrontISTRもある。エクサスケールのスペコンの登場と普及に伴い、HPC環境で国産構造解析ソフトの研究開発がさらに進むことにより、最新HPC環境での構造解析は国産ソフトが牽引していくものと考えられる。

ここで改めて、流体解析と比較した大規模構造解析の特徴をまとめておく。第一に、主な産業分野における部品数は、一般的な家電製品で数千、自動車や重機では数万、そして航空宇宙や大規模プラント業界では百万のオーダーとなる。従来は部品単位での構造解析が主流だったものが、近年の傾向として、それらをアセンブリした機械や構造物全体の丸ごと解析が可能となりつつある。このことは、単に解析規模が増大するだけでなく、部品接続部分に導入される多点拘束機能(Multiple Point Constraint:MPC)や接触要素が、係数マトリクスの条件数を大幅に悪化する要因となり、通常反復法ソルバー[2]が必須となる大規模構造解析の難度を引き上げることになる。

第二に、構造解析は、多くの場合に構造体で囲まれた空間を、流体のような内容物の保持に活用したり人々の居住空間として活用したりするために、構造物が巨大になっても相対的には薄肉構造となり構造体そのものの体積の増加のオーダーは表面積に比例、すなわち寸法の2乗に比例する。一方、流体解析では、流体場の体積は寸法の3乗で増加する。この特性は、構造解析モデルの自由度の増加という観点では抑止する方向に作用するものの、構造解析モデルが相対的に薄肉構造となることにより、係数マトリクスの条件数が急速に悪化するという効果をもたらす。第一の特徴と第二の特徴は、大規模並列解析に不可避の反復法ソルバーがなかなか収束しないという大規模構造解析特有の課題を生み出す。この観点については、4.6.2.2項で改めて述べる。

第三に、流体解析では、乱流等の非線形性が生ずると、解析領域全体に非線形場が拡大するのに対して、構造体では、非線形領域は応力集中部のような局所領域に留まることも多く、線形領域には様々な簡易構造モデルを部分的に活用する可能性が存在する。ただし、この場合には、解析モデルの自由度の増大を抑制する効果があるものの、係数マトリックスの条件数を悪化させる傾向があり、やはり反復法ソルバーの収束性とのバランスが必要である。構造解析の第三の特徴に対する例外は、地盤中を伝わる地震動のような波動伝搬問題である。固体中の波動伝搬問題の解析では、要素サイズは場所によらずほぼ均等であることが望ましく、要素寸法は波動伝搬速度に反比例して細かくなる。

第四に、構造物の設計には、解析モデルの精度・信頼性以外にも、内在する様々な因子の不確実性を考

慮するために、設計時に安全係数を乗ずる文化があり、構造解析の計算精度のあくなき追及は、必ずしも構造設計のコンセプトとは相容れない。さらに、構造設計プロセスにおいても運用プロセスにおいても、多数ケースの解析が必要になる。

以上を要するに、構造解析は、流体解析と比べて解析モデルの規模は相対的に小さめであるが、流体解析よりも悪条件になりがちの係数マトリックスを効率的に解く反復法ソルバーが必須となる。さらに、パラメトリック解析や最適化等のために多数ケースの解析が必要となることから、エクサスケールのマシンを構造解析に活用するニーズは、Capability Computing よりも Capacity Computing にシフトしている。この結果、最新 HPC 環境をフル活用することにより生じる膨大な解析結果を適切かつ効率的に処理し、日々の設計や運用業務に活かす仕組みや支援ツールの開発が必須となる。

さて、産業界においては、熱流体構造連成解析や流体構造連成解析などの連成解析、あるいはより一般的にはマルチフィジクス解析へのニーズが急速に高まってきており、汎用的連成解析機能を特色とする ANSYS や COMSOL などの商用解析ソフトもよく利用されている。連成解析は、流体解析の観点から語られることも多いが、構造解析にとっても大変重要な意味を持つ。その理由は、現実の構造機器や構造物への入力荷重が、機械荷重のみならず、接する熱流体からの熱伝達・熱伝導によって生じる熱応力であったり、内部あるいは外部の複雑な流体場から受ける圧力や剪断力等の流体力などであるからである。また、先にも述べたように、大型化することにより相対的に薄肉となる機械や構造物では、構造機器や構造体の大きな変形や振動が、逆に内部あるいは外部流体場に効率低下や騒音発生等の重大な影響を与えることにもなる。航空機の翼構造や、高温流体を内包するエネルギー機器である圧力容器や配管、タービンブレード等では、このような連成解析が昔から注目されて、様々な研究開発が行われてきた。しかし、従来は対象形状が比較的単純な形状に近似できることや、変形があまり大きくならないように意図的に高剛性に設計されることから、高精度の計算力学シミュレーションを行うことよりも、半解析的なアプローチが研究開発され用いられてきた。それら以外の一般構造機器や構造体でも、従来は固く変形しにくく（高剛性に）作られることが普通であったことから、連成現象が問題となることは少なかった。しかしながら、最近の傾向として、機器や構造物の大型化や軽量化の進展によって、構造物の部材の板厚が相対的に薄く、細長くなることから、剛性が相対的に小さくなり、流体場との相互作用、すなわち連成現象がより顕在化する傾向にある。

現在産業界でよく用いられている商用連成解析ソフトでは、複数の单一現象解析ソフト間の連成解析用入出力データのやり取りをファイルベースで、しかし人手を介さずに自動的に行えるようにすることにより、連成解析プロセスをユーザーの目から隠ぺいした解析システムとなっている。これらのソフトは連成解析を解析ユーザーの身边にしたという点で大きな役割を果たしているが、その延長上では、最新 HPC 環境で大規模な連成解析を実現することはできない。このような連成解析手法は昔からよく知られている弱連成と呼ばれる手法であり、強い連成問題では精度が出ないことが知られている。また、大規模解析においては HPC 環境の活用が必須であるが、個別の解析ソフトについても、解析ソフト間のデータのやり取りについても並列化が不十分であり、本格的な大規模連成解析に適用することは困難であった。

強い連成問題においては、個別現象の方程式系と連成解析条件を一体型のマトリクス方程式に集約し解析するモノリシック解法（一体型解法、昔から強連成解法として知られる）があり、モノリシック解法でないと精度よく解けない問題（付加質量効果の大きな問題）をターゲットとして様々な研究開発が進められている。しかしながら、モノリシック解法の課題は、汎用的な解法を構成することが難しい点であり、

特定現象を対象とした専用ソフトに留まることが多い。また、モノリシック解法から導出される一体型のマトリクス方程式は条件数が悪化するため、最新HPC環境で解く際にも様々な工夫が必要となる。

計算力学分野の最新研究動向に目を転じると、連成解析アルゴリズムや大規模あるいは強い連成解析に関する研究開発が盛んに行われてきている。しかしながら、世界的にみても最新HPC環境で活用できる汎用連成解析ソフトはほぼ皆無といつてもよい。

唯一の例外が、JST CREST研究や文科省の計算科学プロジェクトにおいて、吉村らが研究開発した並列連成カプラー ADVENTURE_Coupler と REVOCAP_Coupler を用いた分離型解法に基づく汎用並列連成解析システムである。この解析システムは、現在富岳上で実機クリーンエネルギーシステムのマルチフィジクス・マルチスケール統合シミュレーションを対象とした研究開発が進められている。この並列連成解析システムでは、独立して開発された、構造解析やLES乱流解析、燃焼流解析向けの汎用並列解析コード (ADVENTURE、FFB、FFR-Comb) を、汎用並列連成カプラー (REVOCAP_Coupler、ADVENTURE_Coupler) を介して連携活用し、各ソルバーの計算が完全に並列化されるのみならず、連成データの並列送受信を行うこともでき、連成解析の全過程がすべて並列化される。さらに、分離型解法を用いているため、一方向連成解析、互い違い型の双方向連成解析、反復型双方向連成解析、を自在に選択して行うことができる。反復型双方向連成解析では、双方向連成問題を非線形問題として定式化し、各時間ステップにおいて非線形反復法を用いて収束解を得ながら解析を進めるので、分離型でありながら、モノリシック解法に基づく強連成解析と同等の解析精度を得ることができる。さらに、並列連成カプラーと並列ソルバーが送受信を行うためには、各並列ソルバーのI/O部分に通信ライブラリーを追加実装するだけでよいので、並列ソルバー改良の負担はほとんどないという利点もある。この解法は、今後のエクサスケール環境での実機大規模連成解析、マルチフィジクス解析の標準的なプラットフォームになると期待される。

本項の最後に、もう一つ重要な視点について述べる。構造解析は、構造機器や構造物の形や内部構造の設計（デザイン）の一環としてしばしば使われる。そのため、通常このプロセスにおいて多数回の計算が行われる。また、最近のニーズとして、実機と同程度の精度・時間応答をする解析モデル（デジタルツイン）を構築しておき、構造機器や構造物の運用の適正化に使うというニーズも増えつつあり、ここでも繰り返しの高速解析が求められる。このような産業界の最先端の現場ニーズについては、最新HPC環境の計算性能を最大限に活用して高精度解析を行えるだけでは不十分であり、こうした解析を短時間に繰り返し行えることが必要である。そのための方向性として、最新HPC環境での高精度シミュレーション結果を活用しつつ低次元での高精度モデル（サロゲートモデルあるいはROM（Reduced Order Model）と呼ばれる）の構築法の研究開発がある。その代表例にAIを活用するデータサイエンス的アプローチや、POD（特異値直交分解）などを活用する数理的なアプローチがある。

4.7.2.2 次世代に解決すべき課題とその解決に必要な計算手法・アプリケーション

a. 次世代における構造解析課題例

大規模構造解析のニーズとしては、すでに産業界の広い分野において様々なものが顕在化しつつあるが、一例として、原子力プラントの耐震解析を取り上げる。巨大構造物の地震応答現象では、地盤、建屋、機器などのすべてのものが一度に揺すられ、それらは少なからず連成するので、各コンポーネントを

個別に解くことは解析精度という観点から必ずしも適切ではない。原子力プラントでは、鉄筋コンクリート壁、配管、圧力容器、燃料集合体などすべてがソリッド要素で表現される。解析規模としては、原子炉圧力容器単体でも十数億自由度を要し、配管や発電機器を含むプラント全体と付近の地盤を合わせて数十億自由度と見積もられる。また、解析は弾塑性の動解析であり波動伝搬と比べて比較的ゆっくりと揺すられることから、時間方向および非線形計算の繰り返しからなる数万から数十万ステップの計算が必要となる。

ペタスケールコンピュータをもってしても、自由度や計算ステップに関して各々これらの 10 分の 1 程度のものを扱うのが限界となっており、現時点では、圧力容器や少数の配管群など、特定箇所にフォーカスした解析を行うか、あるいはアドホックなモデル化や線形解析に頼らざるを得ない。今後の HPC 技術の進展により、プラント全体の統合計算が可能になるのみならず、地球物理・大気海洋分野で行われている地震シミュレーションや津波シミュレーションとの連携により、巨大構造物の地震や津波応答現象のより厳密かつ高精度なシミュレーションが可能となる。

次に、巨大な構造体を含む連成解析について述べる。これまであまり顕在化されていなかった構造体の例として、巨大な発電用風車、特に今後我が国でも大規模に設置が進む洋上風力発電施設を取り上げる。洋上風力発電施設は、発電効率向上のために、広大なサイトに巨大な風車（10MW 以上）を多数（十数基から数十基）建設して運用される。個々の風車は、高さは東京都庁（243m）を超え、重量は大型航空機（ボーイング 777-200 は 191 トン）、を超える、巨大なローター（直径は 150～200m）が日本近海の厳しい風況・海況環境の中で 20 年超の長期にわたって日夜回転する動的機器であり、かつ台風、地震や波浪等の複雑で厳しい荷重の影響を受けるという、これまでのものづくりでは経験したことのないような過酷な環境で稼働する人工物となる。しかも、風車を通過した乱れた流れ（後流 [wake]）を介して巨大風車同士が相互に影響し、発電効率低下や疲労寿命低下などの悪影響を受けることになる。このため、①サイト選定、設計・建設上の課題克服はもとより、運用開始後もアクセスが制限される、②点検が洋上のため工夫が必要、③故障や破壊事故が起こった時の原因究明・修理・取り換え等の対応策、④海況や風況などの公的データの利活用、が必要不可欠となる。したがって、日々の自然環境変化や機器のモニタリングデータの取得・活用や、最新シミュレーション技術に基づくデジタルツインの利活用など、あらゆる場面でデータ駆動型の技術革新を進めていくことが必須となる。

もう一つの大規模な連成解析事例として、次世代火力発電の機器の解析がある。発電や水素製造に用いる大型の石炭ガス化炉は、日本では 2013 年に初めて商用運転が開始された新しい技術であり、必ずしもガス化炉内部の現象を正確に理解した上で最適な設計がなされたものではない。ガス化炉内部は、微粉砕された石炭（微粉炭）が乱流中を分散するとともに、熱分解による揮発分放出、揮発分の気相反応、および固体炭素分のガス化（固気）反応が同時に進行する極めて複雑な乱流反応場である。また石炭中の灰分は高温に晒されて溶融し、炉壁や炉底に溶融スラグ流れを形成する。すなわち、ガス化炉内は高压容器に囲まれた固気液三相流となっているため、内部の現象を計測し、正確に把握するのは極めて困難である。さらに、炉内の状況は用いる石炭の種類にも大きく影響を受ける。加えて、高温高压に晒される炉容器構造体についても、実用化されて間もないことから、その健全性や耐久性についての知見もほとんどない。一方、石炭ガス化は微粉炭燃焼と現象が似ているものの、その LES 研究の数は限られている。これは、そもそもチャー粒子の固気反応過程において複数の活性化学種が混在しモデル化の観点から格段に複雑になることに加えて、圧力容器内の流れ場に対する計測が極めて困難で直接観察による現象理解が進まない

ことや、高圧化により乱流渦スケールが微細化し、より高い空間解像度が要求されること等によると考えられる。また、数値シミュレーションを実機設計に活用するためには、反応容器内の燃焼特性を評価することに加えて、炉容器構造体の健全性や耐久性についても評価する必要がある。本質的に非定常で複雑な現象である燃焼と構造の連成解析を実現した例は極めて限られている。

b. 課題を解決するために必要なアプリケーション群

FEM を用いた構造解析の計算手法や解析の分類は、線形弾性解析、弾塑性解析、振動解析、熱応力解析、衝撃解析、塑性加工解析など多岐にわたる。ここではこれらを、HPC 技術との関連という見地から、連立一次方程式を解かないもの（陽解法系）と、連立一次方程式を解くもの（陰解法系）の 2 つに分類して議論する。これらはどちらも重要であり、個々の問題の特性や要求精度に合わせて使い分けられる。

前者の場合、各有限要素の単位で得られる情報だけを用いた行列ベクトル積が主要な演算パターンとなり、計算時間の大半は要素剛性行列評価や応力積分などの要素単位での計算処理に割かれる。一方、後者の場合、まず各要素の要素剛性行列を用いて、全体剛性行列と呼ばれる比較的大きなサイズの疎行列を組み上げる。非定常問題における陰解法ソルバーの場合、これに関する連立一次方程式を時間ステップあるいは非線形ステップごとに解くことになる。この手順は、静解析、固有値解析を含め、解析ソルバー内に連立一次方程式の解法が組み込まれているもの全般についてあてはまる。

数億から数十億自由度クラスの大規模解析の場合には、もっぱら反復型解法の系統が用いられる。その際、大規模な連立一次方程式を解くための超並列アーキテクチャ向けアルゴリズムに関して、並列反復法ソルバーおよび領域分割法の 2 つがある [3, 4, 5]。連立一次方程式を反復法ではなく直接法で扱う場合、並列化されたスパースダイレクトソルバーを利用することにより、ある程度の大規模化が可能である。とはいえ、反復法に比べ並列化のレベルは低く、現状では PC クラスタを用いて数千万自由度を扱っている事例が若干数存在するが、数億から数十億自由度クラスについてはいまだに困難な状態となっている。

以下、エクサスケール級計算によって可能となる超大規模問題を想定したうえで、エクサスケール級の HPC アーキテクチャ上で高い性能を発揮できると思われる計算スキームについていくつか取り上げる。

(i) 連立一次方程式を解かない場合（陽解法系）

陽解法は、衝突解析や塑性加工解析のように、非定常でしかもきわめて短時間に生じる現象を表現する場合に用いられる。例えば、商用ソフトで言えば LS-Dyna や PAM-CRASH などが有名である。本手法に基づくソルバーの場合、もっぱら個々の有限要素レベルでの比較的小規模な行列やテンソルに関する演算を数百万ステップ以上繰り返すことになる。FEM では一般にこのような演算パターンを Element-by-element(EBE) 型オペレーションと呼ぶ。コアごとの性能に関して、EBE 演算では要素ごとの限られた情報だけを用いて比較的大量の演算を行うため、キャッシュやレジスタを有効に利用することが可能であり、コンパイラによっては非常に高い性能を発揮する。B/F(Byte/FLOP) 値が低い計算機でも有効である。並列性に関しては、要素単位での計算が主なのでその並列化は容易である。ただし、衝撃解析や塑性加工解析では要素間の接触を扱うことが多く、接触判定を行うため検索処理の並列化に注意する必要がある。大規模アセンブリ構造物の取り扱いでは部品間の結合が重要であり、この処理方法に関し、活発な研究開発が行われている。

(ii) 連立一次方程式を解く場合（陰解法）：並列反復法ソルバー

並列反復法は、商用コードではあまり例がないが、国産フリーソフトの FrontISTR などをはじめいく

つかのコードで採用されている。並列反復法ソルバーの場合、対称正定値であれば CG 法、非対称であれば Bi-CGSTAB あるいは GMRES 法などにより、全体剛性行列に関する連立一次方程式を解く。構造解析では一部の非線形問題を除き、対称正定値行列を扱うことが多い。一方、構造物の大型化あるいは軽量化により、梁や板状の薄い形状を扱うことも多く、この場合係数行列の条件数は悪化する。したがって、強力な前処理が必要とされる。近年では代数マルチグリッドソルバーの利用なども進んでいる。並列反復法ソルバーでは、行列ベクトル積演算が繰り返される。FEM の場合、これは 3 種類の方法により実装することができる。まず、係数行列である全体剛性行列を要素剛性行列から組み上げ、その非ゼロ成分をメモリ上に保存する非ゼロソルバーがある。次に、係数行列の全体化は行わず、要素剛性行列の形でメモリ上に保存する EBE ソルバーがある。最後に、要素剛性行列すらメモリ上に保存せず、行列ベクトル積が必要とされる都度に評価する EBE-MSF(Matrix Storage-Free) ソルバーがある。以下、係数行列を対称とし、CG ソルバーにおいてこの 3 種類のソルバー実装の性能見積もりを行う。なお、行列ベクトル積部分の並列化は比較的容易であることから、通信パターンは隣接間が主なものとなる。

非ゼロソルバーの場合、全体剛性行列とベクトルとの積であることから、メモリアクセスは主に行列成分のみとなる。行列の 1 成分をメモリから読み出すごとに、対称性を考慮して 4 個の浮動小数点演算が必要となる。ただし、行列は疎でしかも非構造格子からのものであるため、メモリアクセスは不規則でベクトル化の効率は低く、しばしば節点単位でのレジスタブロッキングが不可欠となる。また CPU アーキテクチャおよびネットワーク通信の双方におけるハードウェア支援のギャザースキッタ機構は性能向上に役立つ。

EBE ソルバーの場合、メモリ上に記憶された個々の要素剛性行列とベクトルとの積の結果を組み上げていくことになる。要素剛性行列は密で対称であるため、非ゼロソルバーと同様、行列の 1 成分ごとに 4 個の浮動小数点演算が必要となる。EBE ソルバーは非ゼロソルバーより記憶量、演算量とも多いが、メモリアクセスパターンは比較的規則的で性能を出しやすい。

最後に、EBE-MSF ソルバーの場合、これは陽解法ソルバーと同様、EBE 演算がメインとなる。B/F 値の低い計算機でも有効である。もし B/F 値が極端に低い場合、本ソルバーは主にメッシュデータのみをメモリから読み出すため省メモリであり、しかも他のソルバーよりメモリアクセス量が少なくなるために結果的に最速ソルバーとなる可能性がある。

(iii) 連立一次方程式を解く場合(陰解法): 領域分割法

領域分割法(DDM:Domain Decomposition Method)は、国産フリーソフト ADVENTURE などで採用されている。先に述べた線形代数レベルでの並列化を行った並列反復法ソルバーと異なり、領域分割法は偏微分方程式レベルでの並列化手法である。まず全体領域を複数の部分領域に分割する。各部分領域に適当な境界条件を与えて独立に解き、領域間境界条件が落ち着いてくるまで反復的にこれを繰り返す。したがって、この方法は反復解法の一種でもある。構造問題においては、並列反復法ソルバーの場合と同様、強力な前処理が不可欠となる。DDM の場合、FETI(Finite Element Tearing and Interconnecting) や BDD(Balancing Domain Decomposition) などのマルチグリッド型前処理が多用される。

DDM の主要な演算パターンは部分領域ごとの FEM 解析(ローカルソルバー)である。DDM の並列化はローカルソルバーを単位として、その並列化は容易であり、通信パターンは隣接間がメインとなる。ローカルソルバー実装にはいくつかの方法が考えられる。一つ目は、部分領域ごとの剛性行列について、これをまず直接法ソルバーを用いて行列分解し、スカイライン形式などでメモリ上に保存しておく。その

上で DDM 反復ごとに前進後退代入を行うものである。二つ目は、部分領域ごとに反復法ソルバーを用いて部分領域問題を解くものである。このとき、メモリ上にはほとんどデータを記憶しない。三つ目は、部分領域ごとに対応する密行列をあらかじめ作成しておき、これらとベクトルとの積を繰り返すというものである。これは並列反復法ソルバーにおける EBE ソルバーのように、ローカル Schur 補元行列と呼ばれる部分領域ごとの剛性行列を、静的縮約により陽に導出してメモリ上に保存しておくものである。

最初の直接法ソルバーアプローチでは、行列分解されたデータに対する前進消去と後退代入演算が主である。この場合、行列の 1 成分を読み出すごとに 2 個の浮動小数点演算が必要となる。二つ目の反復法ソルバーアプローチでは、メモリからメッシュ情報を読んでキャッシュ上で反復法ソルバーを起動する。大量の演算がキャッシュ上で行われることになる。並列反復法ソルバーで説明した非ゼロ、EBE、あるいは EBE-MSF ソルバーのいずれも利用可能である。なお、前処理の導入により反復回数を削減したい場合、前処理行列をあらかじめ作成してメモリ上にストアしておくことも考えられる。最後のローカル Schur 補元行列保存の場合、EBE ソルバーと同様、対称な密行列とベクトルの積であり、1 成分ごとに 4 浮動小数点演算を要する。なお、部分領域ごとに静的縮約によりローカル Schur 補元行列を作成する際に、行列行列積を中心とした大量の浮動小数点演算が行われる。

(iv) 分離型解法に基づく並列連成カプラ

多くの研究者らが精度向上、問題規模の拡大、安定性、ロバスト性、効率の向上を目指して、流体構造連成解析法に関する研究開発を行ってきており、これまでに大きな進展を遂げてきた。しかし、依然として、現実世界の複雑な形状を有する機械や構造物の連成解析をターゲットに、計算時間の短縮、問題規模のさらなる拡大、多様な連成解析の汎用的な解析を目指した研究開発に対する強い要望がある。

一般に、連成現象は一括型 (monolithic) 解法かあるいは分離型 (partitioned) 解法によって解かれる。一括型解法では、流体と構造の支配方程式と連成界面の条件を一つのマトリクス方程式として定式化し解く。一方、分離型解法では、それらの方程式はそれぞれ個別に解く。

吉村・山田は、形状が大変複雑で大規模なメッシュを用いた並列解析が必須となる実機の高精度連成解析を、富岳のような最先端の並列計算機上で解くことを念頭に置き、次に述べる連成解析戦略を採用した。すなわち、①流体と構造それぞれに適した既開発のかつ並列計算機用に十分にチューニングされた方程式解法・ソルバーを採用できるように分離型解法を採用する。その上で、②分離型解法と非線形反復解法（固定点反復解法）を組み合わせる。分離反復 (partitioned iterative) 解法は、単純な分離型解法（互い違い解法）と比べて、解析精度やロバスト性が遙かに向上する。また、③並列計算機上で効率的に連成解析を行うには、各ソルバーが並列計算機用に十分にチューニングされているだけでは不十分であり、ソルバー間の通信プロセスや、連成解析のための反復解法も十分に並列化向けにチューニングされている必要がある。さらに、非線形反復法を用いる際に、Jacobi マトリクスを導出して解く、ニュートン法や疑似ニュートン法では、既存のソルバーを活用することが困難となる。そこで、④非線形反復法には、Gauss-seidel 法、マトリクスフリーニュートンクリオフ法、準ニュートン法などを採用する。これらを実装したソフトウェアツールとして、ADVENTURE_Coupler / REVOCAP_Coupler と名付けた並列カプラが開発された。なお、ADVENTURE_Coupler と REVOCAP_Coupler は基本的な並列連成解析機能は同一であり、後者は、加えて初期メッシュを 8 倍づつ細分化して大規模メッシュを構築（リファイン機能）し解析する機能を実装している。

(v) 次元低減化法

計算機性能の向上により大規模自由度モデルを用いた数値解析が現実的な計算時間内で可能になってきている。しかし、人工物の設計にはパラメトリック解析や最適化が重要であり、設計変数を様々に変え何度も解析を実行することが必要であり、さらなる計算の高速化が求められる。このような背景から、Reduced-order modeling (ROM) と呼ばれる、現象の特徴を十分に捉えつつ計算コストが低い代理モデルの構築手法が注目されている。ROM の代表的な手法として固有直交分解 (Proper orthogonal decomposition, POD) を用いた手法 (以後、POD-ROM という) があり、様々な分野の数値解析に応用されている。POD-ROM では、まず次元削減前の高自由度解析システム HPC 環境を活用してサンプリング解析を行い、Snapshot と呼ばれるトレーニングデータを収集する。次に、収集したデータから特異値分解を用いて POD 基底を抽出し、そのうちの少数の基底を用いて未知数を近似する。そして、ガラーキン法と組み合わせてこの近似を用いることで解析システムの低次元化がなされ、自由度を大幅に削減することができる。

4.7.2.3 ロードマップ

構造解析分野におけるロードマップを図 4.22 に示す。

年代 課題	2026 ~	2028 ~	2030 ~	2032 ~	2034 ~	2036 ~
構造解析	大規模アセンブリ構造物の丸ごと解析技術の開発 大規模アセンブリ構造物の丸ごと連成解析技術の開発 多数の大規模アセンブリ構造物の丸ごと解析結果から高精度低次元解析モデルを構築する技術の開発					

↔

大規模アセンブリ構造物の解析技術とフィールドからの観測データを連携活用したデジタルツインの構築と、設計・運用の最適化の実活用

↔

図 4.22 ロードマップ (構造解析分野)

4.7.3 核融合分野

4.7.3.1 分野の概要

核融合研究分野における計算科学の利用・普及状況

持続可能な人類社会とその発展を支え得る、長期的・安定的に供給が可能な究極のエネルギー源の実現を目指し、核融合エネルギーの研究開発が世界規模で進められている。核融合エネルギーには、

- 燃料が偏在せず豊富である
- 高レベルの放射性廃棄物が生成されず環境負荷が低い

- 核融合反応は暴走せず原理的に高い安全性を有する
- 核拡散の制約を受けない

といった特徴があり、我が国にとっては国産エネルギーとなり得るばかりか、途上国を含めた地球規模での基幹エネルギー需要を長期的に代替し得る、革新的エネルギー源であると言える。従って、核融合エネルギーの実現に貢献することは、低炭素社会の実現に繋がるだけでなく、日本の新たなエネルギー戦略・産業の発展にも繋がる。磁場を用いた核融合炉では、約1億度に熱せられた高温の燃料プラズマを、トーラス状の磁場で効率的に閉じ込めて、核融合エネルギーを取り出す。炉内では、核融合反応によって生成する高エネルギーのアルファ粒子によって燃料プラズマを自己加熱し、高温状態を維持しながら、炉心プラズマの核融合反応を持続させる。この核融合炉開発に関連するシミュレーション分野は、プラズマ、熱流体、中性子、材料と多岐にわたっているが、この中でもプラズマシミュレーションは最大規模の計算量を必要とする。これは、核融合プラズマの第一原理モデルが、異方性を持った強い磁場でプラズマが閉じ込められていることを利用して、粒子速度空間を3次元から2次元に簡約した5次元位相空間（3次元実空間+2次元速度空間）上の粒子分布関数の時間発展を記述する運動論モデル（ボルツマン方程式）で与えられていることに基づく。ジャイロ運動論モデル[6]と呼ばれるこのモデルが、3次元実空間のみで記述される従来の流体モデルに比べて2次元速度空間の分だけ自由度が増大し、～1002倍大きい自由度を必要とするためである。プラズマシミュレーションが対象とする物理問題として、プラズマ乱流に起因する熱輸送や燃料粒子輸送、あるいは、磁気流体现象に起因する高エネルギー粒子輸送が挙げられるが、このような物理現象は、高温プラズマの閉じ込め性能を左右するため、核融合炉の成立性に直結する最重要課題の一つとなっている。富岳やその他の大型計算機と数値シミュレーション技術の進歩によって、このようなシミュレーションの適用対象は、核融合プラズマ物理の理論的研究のみならず、実験データ解析、乱流輸送モデル開発、さらには、炉心プラズマ設計など、具体的な核融合炉を見据えた幅広い展開を見せるようになってきている[7, 8]。

現状と海外の情勢

これまでに JT-60U（日本）および JET（欧州）の2つの実験装置で、入力エネルギーと核融合出力が等しくなる臨界条件を達成しており、現在は、核融合炉の科学的・工学的可能性の実証を目的として、次代の実験装置である国際熱核融合実験炉（ITER）の建設が、日米欧露中韓印の世界7極の国際協力の下に、フランスのカダラッシュで進められている。2025年12月に運転開始を予定しているITERでは、人類初となる長時間の核燃焼プラズマ実験が計画されており、これまでの核燃焼のない模擬実験とは質的に異なる炉心プラズマが実現する。このため、核燃焼プラズマの性能予測や実験条件最適化において、計算機シミュレーションが大きな役割を果たすと期待されている。一方、計算機科学および計算科学から核融合研究を眺めると、欧米のフラッグシップ計算機において、核融合分野は継続的に主要な位置を占め続けており、特に米国においては、2016年より開始された Exascale Computing Project (ECP) の主要テーマの一つに採択されている。米国 ECP では、核融合炉のまるごと解析 (Whole Device Modeling) を目指して、プラズマから材料までを包含するマルチスケール・マルチフィジックスの野心的なシミュレーションプロジェクトが進められている。国内においては、ポスト京重点課題⑥「革新的クリーンエネルギーの実用化」サブ課題 D 「核融合炉の炉心設計」(2015年度～2019年度)において、ITERの炉心プラズマ運転条件最適化に向けたエクサスケール核融合プラズマシミュレーションの開発と「京」を活用

した開発コードの実証研究が進められた。これを受け 2020 年度からは、スーパーコンピュータ「富岳」成果創出加速プログラムの採択研究課題「核燃焼プラズマ閉じ込め物理の開拓」において、プラズマ乱流による熱・粒子・運動量と核融合反応で生成される高エネルギー粒子の輸送過程について、「富岳」を用いた超大規模の位相空間マルチスケール・シミュレーションの実現とデータ科学手法との融合により、その機構解明への取り組みが進められている。このように、国内外でエクサスケール核融合プラズマシミュレーションに向けた研究開発が展開している。

核融合プラズマシミュレーションの進展

核融合プラズマの第一原理シミュレーションは、1990 年代に米国で基本的な手法が確立し、それ以降、計算機性能の進展と共に急速に発展してきた。核融合プラズマでは、炉心サイズが～1m 規模であるのに對して、乱流の最小スケールが、背景磁場中でサイクロトロン運動する荷電粒子の軌道幅程度となるため（イオン～5mm、電子～0.1mm）、直接数値計算（DNS）が主に用いられている。プラズマの数値計算法としては、歴史的に大きく 2 つの手法、すなわち、粒子法（Particle-In-Cell）と格子法（Eulerian、Semi-Lagrangian）が 1970 年代に開発されていた。1990 年代後半に初めて実現した第一世代のトーラスプラズマ・シミュレーションでは、比較的少ないメモリでも計算が可能な粒子法が主流であった。第一世代シミュレーションの最大の成果は帶状流と呼ばれる乱流が励起するメゾスケールの流れ（平均流）が乱流輸送を抑制する効果を示したことであり、この成果は、ITER 設計の妥当性を巡る議論に大きな影響を与えた。2000 年代前半に開発された第二世代のプラズマシミュレーションでは、計算機性能、特にメモリ容量の増大によって 5 次元位相空間を格子で分解する見通しが得られ、格子法によるトーラスプラズマ・シミュレーションが開発された。格子法の計算は、粒子法で問題となる粒子ノイズの影響が無く、計算精度が安定している。このため、実験研究者等のシミュレーション研究者以外にも広く普及し、系統的な実験解析等に第一原理シミュレーションが活用されるようになっている。このような格子法のシミュレーションは、1 本の磁力線の近傍をモデル化した局所モデルによって開発されたが、その後のさらなる計算機性能の進展に伴い、炉心全体を計算領域とする大域的モデル、あるいは、イオンスケールから電子スケールまでを局所的に取り扱うマルチスケール・モデルが開発されている。さらに、2000 年代後半にはペタスケール規模の計算が視野に入り、より長時間スケールの数値実験を目的とする第三世代のプラズマシミュレーションが開発されるようになった。第三世代の大きな特徴は、第一・第二世代の計算が、プラズマ分布関数 f を背景プラズマ分布 f_0 と揺動成分 δf に分解したとき ($f = f_0 + \delta f$)、定常状態にある揺動成分のみを計算する δf モデルを採用していたのに対し、プラズマ分布と乱流成分を同じ第一原理モデルで同時に時間発展させ、非定常プラズマの過渡的変化を計算する full- f モデルが開発されたことである。これにより、 δf モデルで発見された帶状流に加えて、自己無撞着に発展するプラズマ分布と乱流の相互作用という新たなフィードバックループを含む数値実験が実現し、炉心出力に影響するプラズマ分布を第一原理モデルで計算することが可能になった。このような第三世代プラズマシミュレーションの単一イオン系乱流計算（～10ms 程度の時間スケール）の計算量を図 4.23 に示す。ITER 規模の準定常乱流計算を行うには 10PFlops 程度の計算が必要となる。このシミュレーションを多種イオンで構成される核燃焼プラズマに対応させるためには、さらに一桁以上大きな計算規模が要求される。これまでの実験は、水素もしくは重水素のみの模擬実験であったのに対し、ITER の核燃焼実験は燃料イオン（重水素、三重水素）、および、核融合反応で生成するヘリウムイオンを含む多種イオン系である。2025 年に開始される

ITER の実験で次々に蓄積される核燃焼プラズマの知見を炉心プラズマシミュレーションに反映させていくためにも多種イオン系でのシミュレーションは重要である。また、燃料粒子輸送や電子熱輸送をもたらす電子系乱流を計算するにはイオンに比べて数十倍速い電子の運動を追跡するため、核燃焼プラズマのシミュレーションはさらに計算コストが増大する。さらには、磁場の揺動までを含めた電磁乱流シミュレーションも視野に入れなければいけない。他にも、炉心プラズマだけでなく、周辺領域のプラズマをターゲットに据えていた別コードとのカップリングを含め、これらを第四世代として位置づけ、今後のエクサスケール規模の計算によって、ITER のエネルギー閉じ込め時間である数秒程度まで計算規模が拡張し、新たな炉心制御手法を数値実験で開発できるようになる。以上で述べた物理対象や計算規模に対する大規模化・高精度化に加え、これらの炉心プラズマ輸送の大規模シミュレーションに対して、その計算結果を機械学習やデータ科学を利用して高速かつ簡便に予測する手法開発も強く求められている。特に、燃焼プラズマの挙動に対して、大規模シミュレーションと同等の精度で、輸送レベルや温度・密度分布の予測が高速に行えるようにしておくことは、炉心性能の制御手法としても十分に確立しておくべき必須の課題である。近年、核融合研究分野においても、機械学習やデータ科学的手法の応用が進んでおり、国内でも精力的に研究が行われている。例えば、プラズマ乱流シミュレーションのデータを学習させ、乱流の波数空間パターンに基づき、プラズマの非線形飽和時間や輸送レベルの予測を行う研究や、データ科学的手法により乱流シミュレーションを実行すべき入力パラメータ領域をピンポイントで推測し実行する手法の開発[9]が進められている。

4.7.3.2 次世代に解決すべき課題とその解決に必要な計算手法・アプリケーション

超高温の核融合プラズマは、

- 粒子衝突効果による散逸が非常に小さい（無衝突プラズマ）,
- 乱流場の振幅が非常に小さい（平衡分布の 1% 程度）
- 電磁場の計算が電荷や電流の保存精度に敏感

という特徴がある。このため、主要な計算手法においては数値的な散逸や保存誤差を低減しつつ数値的な安定性を保証するための工夫が施されている。粒子法においては摂動分布関数 δf のみを粒子で表現することによってノイズを低減する δf -PIC 法が主流となっているのに対し、格子法のプラズマシミュレーションにはオイラー法（差分法、有限体積法など）、セミラグランジュ法（CIP 法など）、スペクトル法、あるいは、それらを組み合わせたコードが開発されてきている。また、高性能な閉じ込めプラズマの実現には、炉壁周辺領域におけるプラズマの密度・温度分布の急峻な勾配形成が必須となる。これは、プラズマ閉じ込め状態の遷移現象として知られており、この現象の解明には、装置中心部での 1 億度以上の超高温プラズマと同時に、温度が数十分の 1 となる周辺プラズマを同時に取り扱うことが必須となる。この問題は、プラズマ乱流による構造形成の問題だけでなく、ITER の性能予測やダイバータへの熱負荷予測において取り組むべき重要課題である。このようなシミュレーションの実現には、中性粒子との相互作用を含む周辺・ダイバータ領域のプラズマ物理モデルと中心領域を扱う運動論的シミュレーションとのハイブリッド解析手法の確立が必須である。一方で、機械学習やデータ科学などに基づき、大規模プラズマ乱流シミュレーションの効率的な実行と豊富で巨大な計算データの有効的な解析手法の開発も求められてい

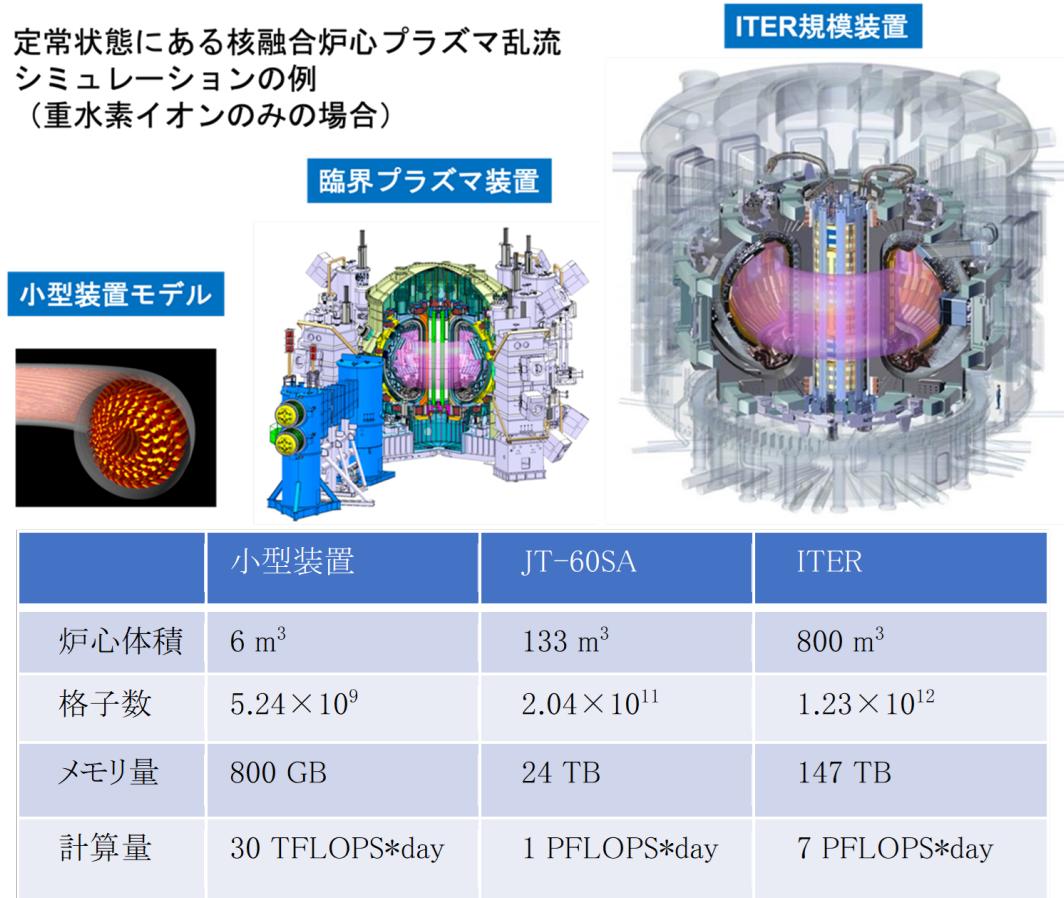


図 4.23 核融合炉心プラズマ乱流シミュレーションの計算規模（重水素イオンのみの場合）

る。以下では、現在、国内の核融合研究において最も大規模なシミュレーション研究が進められている代表的なコードとして、GT5D コード（大域的乱流コード）[10]、MEGA コード（粒子・流体ハイブリッドコード）[11]、GKV（局所的乱流コード）[12]、GKNET（大域的乱流コード）[13] を例に挙げ、次世代課題の対象や問題規模、ロードマップについて述べる。

a. 次世代の課題の目的や対象

GT5D コード：

GT5D コードは炉心全体を計算領域として大域的 full- f モデルの長時間シミュレーションによって炉心プラズマ性能を評価する。現在の大域的 full- f シミュレーションは乱流輸送で支配的な密度揺動に起因する乱流成分のみを記述する静電的モデルとなっている。これを磁場揺動を含む電磁的モデルに拡張することで、プラズマ対向材の工学設計に重要なプラズマ崩壊現象や周辺局在モードといった磁気流体现象まで含めた第一原理計算が可能となる。また、現在は定常プラズマに対する数 10~100 ミリ秒程度の数値実験を ITER のエネルギー閉じ込め時間として想定される数秒程度まで伸長できれば、輸送障壁形成等の乱流遷移現象、中性粒子ビームや電磁波加熱等の外部入力に対するプラズマの過渡的応答を解析できるようになり、新たな炉心制御手法を数値実験で開発できるようになる。また、中性粒子との相互作用を含む周

辺・ダイバータ領域のプラズマモデルとのハイブリッド解析も次世代の課題となる。

MEGA コード：

本コードの目的は磁場閉じ込め核融合プラズマにおける高エネルギー粒子の閉じ込めを解析することである。核融合炉において、核融合反応によって生じる高いエネルギーを持った粒子を如何に閉じ込めるかは、核融合研究における最重要課題の一つである。これは、国際協力によりフランス・カダラッシュで建設中の ITER が目指す世界初の核融合燃焼プラズマの実現の鍵となる重要な課題である。MEGA コードでは、高エネルギー粒子を粒子法により、背景プラズマを格子法により扱うハイブリッド・シミュレーションにより、高エネルギー粒子の閉じ込め問題に挑んでいる。次世代機で計算対象は、核燃焼プラズマを想定し、5 種類の粒子種、3 次元格子点数 ($1024 \times 1024 \times 1024$)、合計粒子数 1600 億個の規模が想定される。

GKV コード：

プラズマ分布関数の熱平衡成分と摂動成分に分けた場合の摂動成分 δf の時間発展を追う δf モデルに基づくコードである。一本の磁力線に沿った局所領域（ブラックス・チューブ領域）に注目し、プラズマの乱流およびプラズマ分布関数の時間発展を解き進める。GT5D コードや GKNET コードなど大域的な手法に基づくプラズマ乱流コードと違い、平衡分布の時間発展を記述することはできないが、磁場揺動や極微細スケール乱流を定量的に扱えるため、大型ヘリカル装置 LHD や JT-60U など、実際の大型プラズマ実験装置を対象にした実験結果の再現や、輸送予測計算が行われてきた。特に、電子スケールからイオンスケールに跨る広範囲な時空間スケールに渡った乱流シミュレーションが実現し、プラズマ乱流のマルチスケール性が明らかになった。しかし、マルチスケール性を考慮しながら、且つ、LHD やヴェンデル・シュタイン 7-X（独）といった複雑な 3 次元形状を有する系における乱流計算の実現は、富岳以降の次世代機を待つ必要がある。特に磁場配位の 3 次元性は ITER などトカマク系においても重要性が増しており、3 次元性を考慮したマルチスケール・プラズマシミュレーションにより、具体的で定量的な炉心設計が可能になる。また、データ科学に基づく手法により、巨大な計算データの解析と大規模シミュレーションの効率的な実行方法の開発も進める必要がある。

GKNET コード：

GKNET コードは、GT5D と同様の大域的な full- f 乱流シミュレーションコードであるが、特に、乱流を抑制することで断熱層/粒子遮蔽層として作用する内部輸送障壁の形成機構の問題などに対して、電子の運動論効果を導入した第一原理シミュレーションを実現し、挑んでいる。次世代機での計算対象として、これを ITER サイズに拡張し、核融合実験炉サイズでの大域的な乱流輸送シミュレーションを想定している。

b. アプリで使われている計算手法やアルゴリズムの詳細

GT5D コード：

プラズマの粒子分布を記述する 5 次元ジャイロ運動論方程式（簡約化ボルツマン方程式）、粒子間のクーロン衝突を記述する線形フォッカー・プランク衝突演算子、および乱流揺動による電場を決定するポアソン方程式（および、磁場を決定するアンペール則）を計算する。5 次元ジャイロ運動論方程式は 5 次元差分（無散逸保存型差分スキーム）によって離散化し、半陰解法時間積分によって磁力線方向の高速な移流を含む線形移流項のみを陰解法で計算する。この差分陰解法ソルバとして、省通信アルゴリズムや混

合精度演算に基づく先進的なクリオフ部分空間法ソルバを用いている。衝突演算子は中心差分によって離散化し、半陰解法時間積分を用いて速度空間の移流・拡散を計算する。ポアソン方程式については、トーラス断面上の有限要素法とトーラス方向のスペクトル法を組み合わせて計算する。上記3つのソルバでは各方程式の対称性に依存して異なる並列化軸を採用しており、ソルバ間でデータ転置の集団通信を行う。

MEGA コード：

MEGA コードは、プラズマ粒子に粒子シミュレーションを適用して各粒子種の密度、磁場平行速度、圧力を計算し、電磁流体方程式に基づいて電磁場を計算するハイブリッドシミュレーションコードである。電磁流体方程式は有限差分法と Runge-Kutta 時間積分で計算される。

GKV コード：

空間離散化として、磁力線垂直方向の2次元実空間に対してはフーリエ・スペクトル法（従って、2次元の波数空間）、磁力線平行方向の1次元実空間および2次元速度空間については、4次または5次精度差分法を採用し、時間積分には4次精度陽的 Runge-Kutta-Gill 法のほか、隠的解法も部分的に採用する。5次元位相空間に加え、プラズマ粒子種を区別するためにさらに1次元を加えた6次元配列を用意し、これに対して MPI を利用した5次元領域分割を行う。この際、並列2次元 FFT のための転置通信、差分演算のための1対1通信、速度空間・粒子種積分のための総和通信を利用している。

GKNET コード：

5次元ジャイロ運動論方程式については、位相空間5次元を有限差分法によって離散化し、時間方向を陽的 Runge-Kutta 法によって時間積分している。またポアソン方程式、およびアンペール則については有限差分法と FFT を併用して解いている。並列化に関しては、位相空間5次元のうち3方向と粒子種に対して領域分割し、MPI 並列を利用して計算する。近年、GPU 並列の導入を進めており、本期間の早い段階で両者の併用を想定している。

c. 必要な計算資源と見積もり

GT5D コード：

計算規模は、 $(R, \zeta, Z, v_{\parallel}, v_{\perp}) = (720, 128, 1080, 96, 24) \sim 2 \times 1012$ 格子上で、電子、重水素、三重水素、ヘリウム、不純物イオン等を含めた計5粒子種に対する総計~1013格子以上が想定される。1ケースあたりの演算量：508,000 EFLOP、メモリ量：0.4 PB、ストレージ量：8.3 PB、計算時間：336時間である。総演算量として、50,800,000 EFLOP、要求性能として、420 PFLOPS（ケース数100）となる。（参考）周辺プラズマとのハイブリッド解析での周辺・ダイバータ粒子シミュレーションでは、[不純物粒子：3種] × [テスト粒子数：1億5000万個] = [総計粒子数：4億5000万個] の規模を想定する。

MEGA コード：

想定規模は、 $1024 \times 1024 \times 1024$ の格子点、1600億個の粒子（5粒子種）である。1ケースあたりの演算量は17,700 EFLOP、メモリ量2.88 PB、ストレージ量1.05 PB、1ケースあたりの計算時間が50時間である。総演算量として、1,770,000 EFLOP、要求性能として、98.3 PFLOPS（ケース数100）となる。

GKV コード：

マルチスケール3次元プラズマシミュレーションで想定される格子点数は、 $(Nx, Ny, Nz, Nv, Nm, Ns) = (1024, 1024, 1024, 128, 64, 3) \sim 2,64 \times 1013$ 個である。富岳12,288ノードでの実績では、格子点数： $(Nx, Ny, Nz, Nv, Nm, Ns) = (1024, 1024, 48, 128, 64, 3) \sim 1.24 \times 1013$ 個（上記の次世代想

定ケースの約 4.6

GKNET コード：

$(Nr, N \theta, N \zeta, Nv, Nm) = (512, 1024, 128, 128, 32) \sim 2.75 \times 10^{11}$ 程度の格子点数を用いて位相空間 5 次元を離散化した系をターゲットとする。粒子種としては重水素イオンと電子を想定している。この場合、1 ケースあたりの演算量：50,000 EFLOP, メモリ量：0.15 PB, ストレージ量：2.4 PB, 1 ケースあたりの計算時間：240 時間である。総演算量として、2,500,000 EFLOP, 要求性能として、57.8 PFLOPS（ケース数 50）となる。

4.7.3.3 ロードマップ

核融合プラズマシミュレーションのロードマップとしては、図 4.23 に示すように～100PF 規模までは装置規模の拡大に伴う弱スケーリングを想定し、それ以降は、時間スケールの拡大に伴う強スケーリングを想定する。研究のロードマップとしては、図 4.24 に示すように

- (核燃焼のない) 超高温プラズマ実験に対する実証研究
- 核燃焼プラズマの性能予測
- ITER の実験条件最適化や運転手法の高度化

等を主要なターゲットとする研究開発の展開が想定される。既存実験に対する実証研究は JET、JT-60SA に代表される 2 次元形状トーラス磁場のトカマク型装置、あるいは、国内の大型ヘリカル装置 (LHD) や欧州のヴェンデルシュタイン 7-X (独) に代表される 3 次元形状トーラス磁場のステラレータ装置を対象とするが、後者の計算では荷電粒子軌道が複雑化するため計算コストが一桁～二桁以上増大する。このような 3 次元形状磁場効果の研究は ITER における制御磁場や炉内機器の炉心プラズマへの影響を評価する上で有用である。以下では、格子法に基づく局所的プラズマ乱流コード GKV と、大域的プラズマ流体解析コード GT5D に基づいてロードマップを説明する。

局所コード GKV の現状では、富岳において、2 次元系の燃焼プラズマにおける電子スケールからイオンスケールまでのマルチスケール計算が実現できており、12,288 ノードを利用して演算性能 3.1PFLOPS で 106 ステップのプロダクションランを 1263 万ノード時間で達成している。次世代では、このマルチスケール計算の対象を 2 次元系から 3 次元系に拡張するため、沿磁力線方向の解像度を 20 倍以上、高精度化する必要がある。また、Fat ノード構成により磁力線垂直方向の 2 次元空間をノード内に収めてノード内高速通信で all-to-all の 100 倍程度の加速を目指す。この高速化により、3 次元系マルチスケール乱流計算が多数実行でき、機械学習に用いるマルチスケール乱流解析データベースが生成できるようになる。

GT5D コードの現状としては、富岳において ITER 規模の数値実験が実現しており、処理速度としては数千ノードを用いて～10 秒/ステップが達成されている。この処理速度は～10 ミリ秒程度の定常解析を約 10 日間で処理できる性能である。次世代機ではこれをさらに 100 倍の閉じ込め時間スケール規模～1 秒まで伸長することが目標となる。これを同程度の処理時間で達成するためには～0.1 秒/ステップ程度まで処理を加速する必要がある。GT5D で富岳向けに開発した省通信型行列解法は～1 秒/ステップ程度（行列解法の中では～1 ミリ秒/反復）の処理速度を念頭に置いて解法を設計してきたが、次世代機ではさらに通信レイテンシや負荷バランスをきめ細かく制御、最適化していく必要がある。この問題の解決に向

けて GPU のような演算加速装置を多数搭載し、ノード間通信に比べて一桁以上高速なノード内通信機構を備えた Fat ノード構成に向けた最適化技術の開発を想定している。具体的には Fat ノード採用によるノード数の削減によってノード間通信のレイテンシを最小化し、さらにノード内外を意識した階層的並列化とノード内におけるきめ細かい負荷バランス制御によってレイテンシを一桁以上削減する。また、省通信型行列解法に関しても混合精度演算を活用した前処理技術によって収束特性を向上し、計算機科学、応用数学の手法を駆使して 100 倍の処理速度向上を目指す。

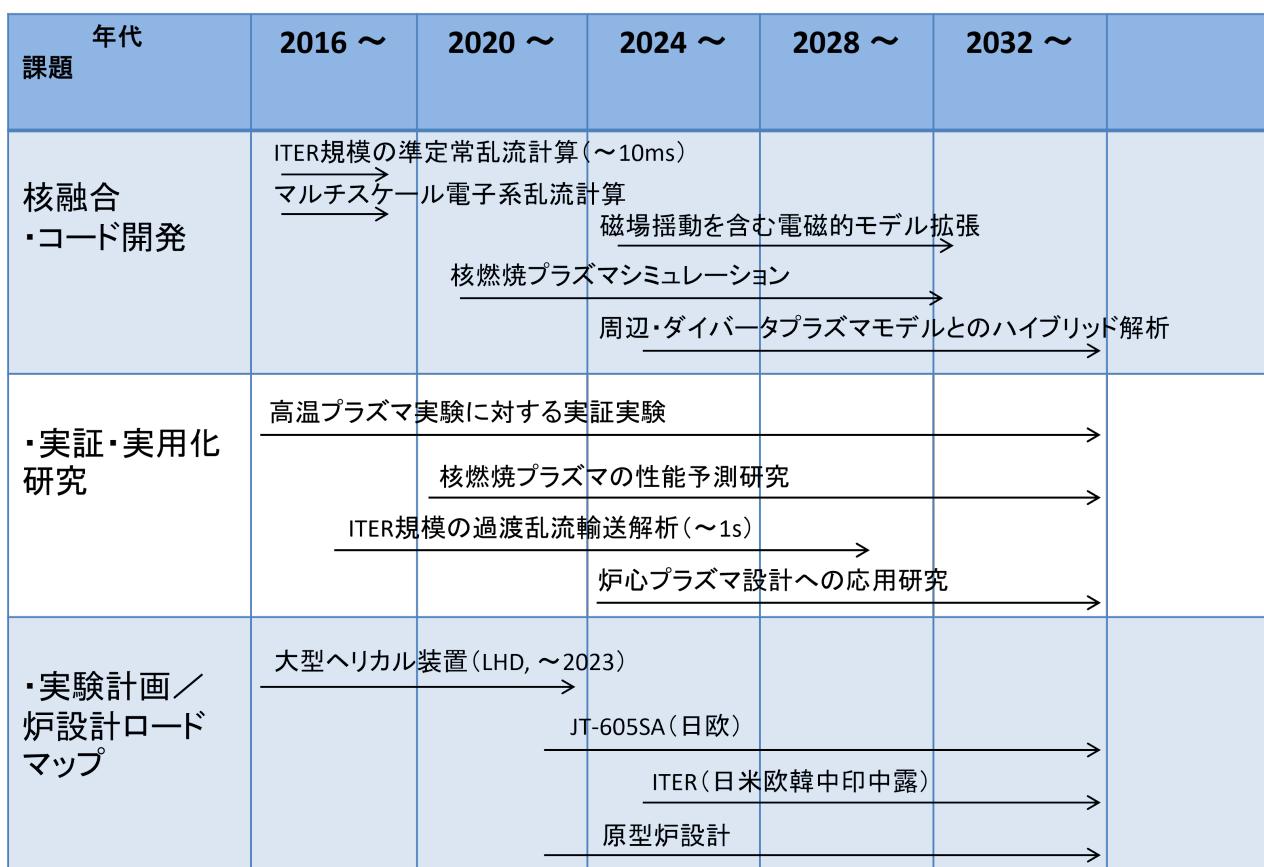


図 4.24 ロードマップ (核融合分野)

4.7.4 可視化・データ処理分野

4.7.4.1 分野の概要

可視化は、大量のデータから視覚的な情報を手掛かりに効率よく新しい科学的知見や洞察を得るために重要な基盤技術であり、自然科学分野、特に計算科学分野を中心に発展し、今では理学や工学だけでなく生命科学や社会科学など様々な学問分野に広く浸透している。現在、計算科学分野では、富岳コンピュータの登場によって幕を開けたエクサスケール時代の到来により、計算機環境の高効率・高性能化が進み、数値計算技術のさらなる高度化も相まって、時間当たりに出力されるデータ量がより一層大規模化している。また、近年では、ビッグデータ解析や AI の活用により、複雑化する科学的・社会的課題の解決のた

めにデータ科学と計算科学の融合に対する要求が高まっており、富岳に代表されるエクサスケール時代のハイパフォーマンスコンピューティング（HPC）環境上で処理されるデータの多様化が進んでいくことが予想される。このような多種多様な大規模データから、人間が解釈できる科学的知見を獲得するための可視化および分析技術は必要不可欠な技術であり、その重要性がますます高まっている。

可視化研究の動向

コンピュータを使った可視化研究は、1987年に米国 ACM より出版された特集記事” Visualization in Scientific Computing”（通称、ViSC レポート）[14] が起源とされる。この ViSC レポートでは、可視化は” Visualization is a method of computing” と定義されており、計算科学分野において、計算結果を単に処理する方法ではなく、それ自身が計算の一つの方法であるとされ、可視化の研究領域、研究への勧告、さらには産業競争力にいたる幅広い視点で可視化研究の重要性がまとめられている。このレポートの出版を機に可視化研究に関する議論が活発化し、1990年には米国で国際会議 IEEE Visualization (Vis) が創設され、可視化研究における最も権威のある国際会議の一つとして認知されている。その後、可視化研究の議論の場として、欧州を中心とする国際会議 EuroVis や環太平洋地域を中心とする PacificVis が創設され、国際的に活発な議論が進められている。ViSC レポート出版当初の可視化は、主に自然科学分野の数値シミュレーション向けの科学的可視化（Scientific Visualization=SciVis）として中心に発展した。1990年代中頃からは、自然科学に限らないより一般的なデータを扱う情報可視化（Information Visualization=InfoVis）に関する研究が盛んに進められてきた。その間、情報処理技術の急速な発展に伴い可視化技術も急速に進展した。その状況をうけて、ViSC レポート出版から約 20 年後の 2006 年に米国 NIH より” Visualization Research Challenges”（通称、VRC レポート）[15] が出版される。この VRC レポートでは、これまで明確化されていなかった可視化技術の利用者の位置付けを明確化し、可視化の出力は画像ではなく知識であるとし、データから可視化を通して知識を獲得するまでの一連のプロセスを可視化発見プロセスとして定義している（図 4.1）。このレポートの出版以降、可視化研究分野では、視覚情報を通じてデータから効率よく知見を獲得するための新しい可視化パラダイムとして、データ分析、対話操作、認知などの各理論を統合した視覚的分析論（Visual Analytics=VA）に関する研究開発が活発化している。

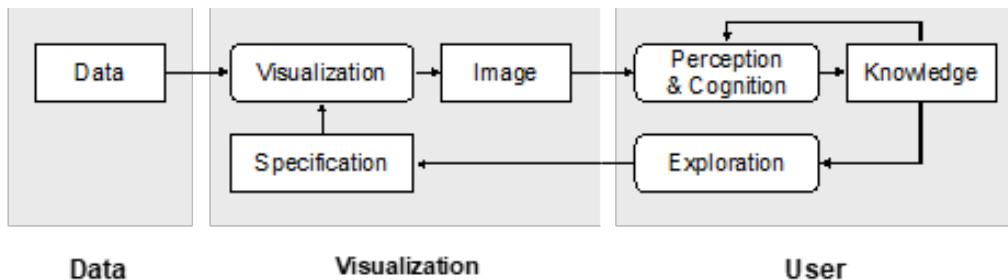


図 4.25 可視化発見プロセス

可視化研究分野では、これまで、SciVis, InfoVis, VA の 3 つの研究分野に分かれて議論が進められてきた。国際会議 IEEE Vis では、この 3 つの研究分野は 2007 年まではそれぞれ別の会議として議論されてきたが、2008 年からはこれらの会議が統合され VisWeek と呼ばれるようになり、2012 年からは VA,

InfoVis, SciVis の頭文字をとった VIS という名称で統合的な議論が展開されていた。さらに、近年の可視化研究では、統計的推論や深層学習などに関する議論も活発化し、データ科学における中核的な学問分野として発展している。そのため、2016 年に分野再編が議論され、可視化分野を次の 6 つのエリアモデルに分類する案 [16] が提出され、2021 年からはこのエリアモデルをベースに研究領域が設定されている。

- Area 1: 可視化の理論と実践 (Theoretical & Empirical)
可視化の基礎を確立するための理論的・実証的な研究
- Area 2: 可視化応用 (Applications)
特定領域の可視化の応用に焦点をあてた研究
- Area 3: 可視化システムとレンダリング (Systems & Rendering)
システム構築やレンダリングアルゴリズムに関する研究
- Area 4: 可視化における表現と対話 (Representations & Interactions)
タスクに応じたデータの視覚的表現法やインタラクション技術に関する研究
- Area 5: データ変換 (Data Transformations)
視覚的表現に適した効果的かつ効率的なデータ変換アルゴリズムに関する研究
- Area 6: データ解析と意思決定 (Analytics & Decisions)
データ解析、知識発見、意思決定、機械学習などの知的処理に関する研究

計算科学分野における可視化

計算科学分野における可視化は、SciVis として可視化黎明期から続く重要な研究分野の一つであり、その研究成果の多くが ParaView や VisIt といった汎用可視化ツールに実装されている。また、基本的な可視化技術は、数値計算処理のポスト処理機能の一部として、すでに多くのソルバーにも組み込まれている。単にデータを画像化するための基盤技術としての可視化は、計算科学研究者に対して広く提供されており、効率的な利活用のフェーズとなっている。しかし、昨今の HPC 環境の著しい発展に伴いデータの大規模化に伴い新たな問題が発生している。データの大規模化に伴い可視化の問題は、以前から存在しており、並列レンダリング技術や分散データ処理技術など様々な解決手段がすでに提供されている。しかし、シミュレーション計算終了後のデータ入出力や、可視化端末へのデータ転送が困難な場合、データを可視化することができない状況が発生する。この場合、高精度に計算された結果であるにも関わらず、可視化のためにデータの空間的・時間的間引きが必要となり、場合によっては知見獲得につながる重要な現象が欠落する可能性がある。近年、可視化研究分野では、このような問題を解決するための手法として、シミュレーション計算を行いながら、同時に可視化まで行う in-situ 可視化 [17] に関する研究が注目されている。In-situ 可視化は、シミュレーションから可視化まで想定した計算において、時空間解像度の向上と膨大な計算能力の活用が可能であることから、新しい計算パラダイムとなることが期待されている。初期の研究開発においては、HPC 環境におけるシミュレーション処理と可視化処理の効率的な連携 [17] について活発に議論され、その一部が汎用可視化ソフトウェアの ParaView[18] や VisIt[19] などに組み込まれている [20, 21]。また、GPU を前提として実装されたレンダリング手法を CPU 環境下で実行するための基盤ソフトウェア [22] も整備され、レイトレーシングによる高品位レンダリング基盤 [23, 24] の In-situ 可視化への活用も進んでいる。さらに、ParaView や VisIt の可視化基盤ソフトウェアである VTK[25] の

データモデルを採用し、マルチコア・メニーコア向けに最適化された可視化基盤 VTK-m[26] の開発も進められている。その他、In-situ 可視化向けの統一的なデータモデルと API を提供する SENSEI[27] や、シミュレーションと可視化の効率的なデータ連携を実現する基盤である ADIOS[28]、基本可視化機能を含む軽量な In-situ 可視化基盤である Ascent[29] など、さまざまな In-situ 可視化ソフトウェアがオープンソースで公開されている。以上のように、基盤環境としての In-situ 可視化は、多くの HPC 環境でその整備が進められており、応用のフェーズに入っている状況であるが、エクサスケール時代に突入するにあたり、2018 年にドイツのダグスツールで開催されたワークショップにおいて計算科学向け In-situ 可視化の未解決問題についてまとめられている [30]。以下に、それらの課題を簡単にまとめて列挙する。

1. データの質と圧縮： 事後のデータ探索（データ解析）処理に影響しない、In-situ 処理における質の高いデータ削減・圧縮手法
2. ワークフロー仕様： 様々な分野への In-situ 処理の適用において、想定されるタスクの高度な抽象化と再利用性の高いワークフロー仕様
3. ワークフロー実行： ワークフローを構成する複雑なタスクを、動的に変化するリソースを考慮して柔軟かつ高効率に実行するための基盤環境
4. エクサスケールシステム： 計算速度とディスクストレージ帯域幅のギャップが拡大するエクサスケールシステムにおいて、In-situ 可視化の更なる高度化とその性能評価
5. アルゴリズム： エコシステム化が進む In-situ 環境下において、制限されたリソース内での異種データ連携を前提した並列度の高い新しい可視化および解析アルゴリズム
6. 新しいユースケース： 単一の大規模シミュレーション向けの探索的分析だけでなく、不確実性を定量的に評価するための複数の計算を実行するアンサンブル計算、計測データを取り込み計算するデータ同化、アルゴリズムの収束やシステムの性能などの別データの分析など、In-situ 可視化が可能にする新しいユースケース
7. エクサスケールデータ： エクサスケールシステムでのシミュレーションによって生成される大規模・複雑データ向けに、シミュレーション・可視化・解析のすべてにおいて最適化されたデータモデルとそれを効率的かつ安定的に処理するための実行基盤
8. 性能評価モデル： In-situ 処理全体の性能最適化を目的とした、アルゴリズム実行性能の予測モデル構築とその検証
9. HPC とビッグデータの融合： In-situ 可視化と分析のための HPC とビッグデータの融合
10. ソフトウェアの複雑性： 利用者が In-situ 処理を導入する際に障壁となるソフトウェアの複雑さや異機種混在環境となる計算資源への対処

国内の動向

計算科学分野向けの可視化研究は、可視化情報学会、日本シミュレーション学会、計算工学会などの研究会および全国大会において議論されているが、数値シミュレーション向けの可視化である SciVis に関する研究は少なく、国内においては適切な研究発表の場も少ないので現状である。しかし、世界最高水準の HPC 環境を継続的に有している我が国において、大規模データの可視化に関する研究は重要な課題の一つである。特に、大規模ボリュームデータ向けの粒子ベースレンダリング手法 [31]（図 4.26）、粒子

ベースレンダリングを使った対話的 In-situ 可視化手法 [32]、ノード数の制約を排除した並列画像重畠手法 [33]、In-situ 可視化向け多視点全方位レンダリング法 [34]、In-situ 可視化向け適応的時間サンプリング法 [35] など独自の可視化手法に関して現在も発展的に議論が継続している。また、天体観測データ向け時空間特徴解析のための視覚的分析手法 [36]、空間線量分布データ向け視覚的多次元時系列解析手法 [37] など、特定分野のデータ解析を目的とした新しい可視化手法に関する研究も進められている。最近では、HPC システムの計算の安定性や信頼性の向上を目的とした視覚的システムログデータ解析に関する研究 [38]（図 4.28）など、計算科学分野向け可視化の新たな取り組みが始まっている。さらに、ニューラルネットワークを使ったデータ駆動型の偏微分方程式推定 [39] など、計算科学とデータ科学を融合した解析手法に関する研究も進められている。計算科学に限定しない可視化研究では、生態系データ向け視覚的因果解析システム [40]、多次元時系列データ向け多重次元削減法 [41]、階層型グラフデータ向けレイアウト評価基準 [42] など、可視化研究分野のトップジャーナルに掲載される研究も多く、国内の可視化研究は活発かつ発展的に議論が継続されている。

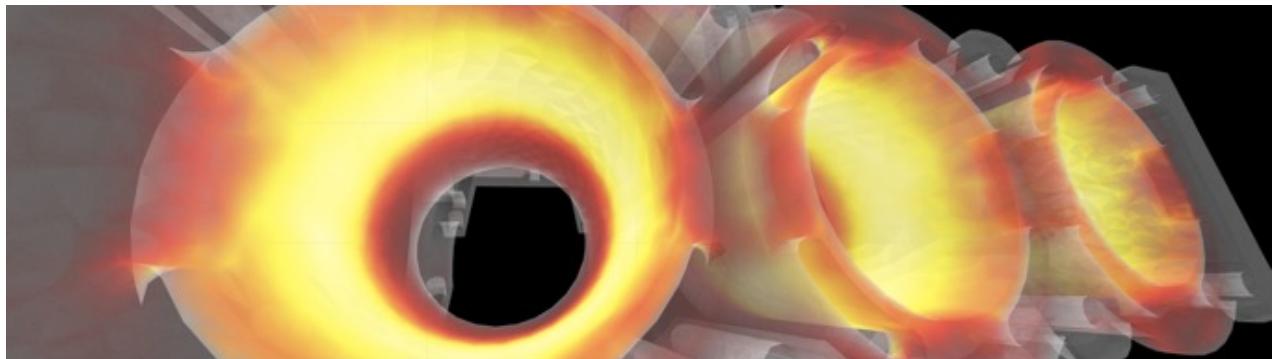


図 4.26 大規模構造解析データの高解像度リアルタイムレンダリング

国内における可視化研究に関しては、2017 年に日本学術会議・総合工学委員会（第 23 期）から提言「科学的知見創出に資する可視化に向けて」[43] が公表され、可視化のあるべき姿に関する議論が進められている。この提言では、ビッグデータ時代において人工知能とともに科学的知見を得るために可視化が果たすべき役割について、科学的方法に着目していくつかの問題点を確認し、それに対する解決案がまとめられている。特に、計算科学分野での可視化については、その必要性を確認しつつも、ポストペタスケールの高性能計算において可視化が重要課題と認識されていない状況を鑑み、科学的方法における観察・問題設定段階での超大規模データの俯瞰的可視化の重要性と、計算時間（Time To Solution, TTS）だけでなく可視化を介した知見獲得までの時間（Time To Discover, TTD）の短縮を目指した先進的な可視化技術に関する研究開発の必要が強く訴えられている。現在、この提言を機に第 24 期総合工学委員会において科学的知見の創出に資する可視化分科会が設置され、提言の具体的実践と日本初の可視化パラダイム策定に向けて、分野横断的な議論が展開されている。

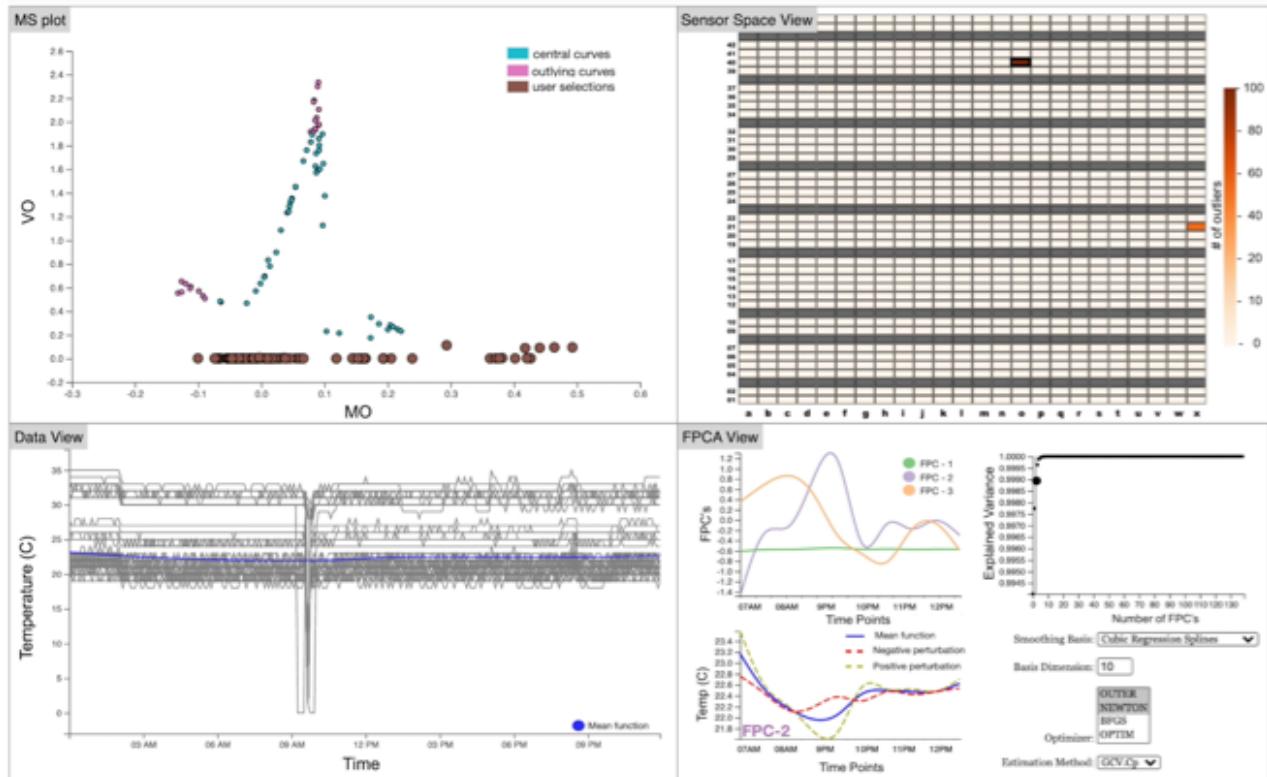


図 4.27 可視化発見プロセス

図 4.28 スパコン（京コンピュータ）ログデータの視覚的解析システム

4.7.4.2 次世代に解決すべき課題とその解決に必要な計算手法・アプリケーション

国内の計算科学分野ではまだ十分な議論が行われていないが、HPC 環境を想定した可視化とその関連技術が想定する課題は、ダグツールレポート [30] に挙げた 10 の未解決問題に集約されている。ここでは、国内外の可視化研究の動向、特に、国内の計算科学向け可視化研究の現状を踏まえて、優先して取り組むべき課題を以下にまとめる。

(1) In-situ 可視化の高度化

シミュレーションと可視化を同時にすることでデータ I/O コストを大幅に削減することができる In-situ 可視化は、エクサスケール時代の大規模データ向け可視化において最も重要かつ必要不可欠な技術のひとつである。In-situ 可視化の基盤となる可視化技術は、1987 年の ViSC レポート出版以降、SciVis 分野の重要課題として数多くの研究成果が発表され、すでに計算科学の諸分野に広く浸透し、ツールとしても整備され成熟の域に達している。In-situ 可視化については、現在、活発な議論が継続されているが、基盤システムとしての基本ソフトウェアも多く提供されており、それらを使った研究成果も増加傾向にある。特に、海外の HPC 環境においては、In-situ 可視化の基盤環境は、標準的なソフトウェア環境の一つとして事前にセットアップされており、研究者と専門のエンジニアが協力してサポートする体制が整

備されている場合が多い。しかし、国内の HPC 環境では、事前に In-situ 可視化の基盤システムがセットアップされていることは少なく、十分なサポート体制が構築されていない場合も多いため、このような状況が In-situ 可視化の利用の大きな妨げになっている。この問題は、研究開発における課題とは異なるが、提言 [43] においても、HPC 環境での永続的な可視化技術研究開発組織の構築の必要性が指摘されており、エクサ時代の計算科学の発展を考える上で、早急に解決しなくてはいけない最も重要な課題の一つである。In-situ 可視化の技術的な課題の一つがデータ探索性能の欠如である。一般的な In-situ 可視化はバッチジョブ方式によって実行されるため、事前に可視化パラメータを設定しておく必要があり、処理終了後は出力結果である画像を再生することしかできない。そのため、画像再生時の対話的なパラメータ変更ができず、探索的な操作ができないといった問題がある。この問題に対しては、事前に複数の視点や可視化パラメータを設定しておくことで画像データベースを構築し、処理終了後は設定した視点およびパラメータの範囲内でデータベースから対話的に画像を再生することで探索性能を高めるアプローチ [44] が一般化されつつある。現在は、視点配置方法の改善や、全方位画像レンダリングによる視線方向に対する対話性の向上 [34] などの改善がはかられており、今後、多視点 In-situ 可視化アプローチによるさらなるデータ探索性能の向上が期待される。一方、最近の可視化研究の傾向では、効率の良い画像化よりも、画像からの効率の良い可知化（知見獲得）にフォーカスした研究が増えている。その傾向に従えば、計算時間（TTS）の短縮が課題となる画像化技術を中心とした多視点 In-situ 可視化は、比較的速やかにその基盤技術が浸透すると考えられる。今後は、知見獲得の効率化を目指した In-situ 可視化に対する取り組みにシフトし加速化すると予想される。現在、知見獲得までの時間（TTD）の短縮を目指したスマート In-situ 可視化 [35, 45] に対する新しい In-situ 可視化の研究開発が始まっている、可視化する最適な時空間領域を自動的に推定する方法などの知的データ処理手法およびその実行基盤の開発が課題となる。すでに、富岳上でもこの基盤環境の構築が進められており [46, 47]、今後の発展が期待される。

(2) 大規模データ向け視覚的解析環境の構築

国内の計算科学分野における可視化は、数値シミュレーションデータ向けの画像化ツールという印象が根強く残っている。一方で、可視化研究分野では、画像化ツールとして可視化が議論された SciVis は、研究領域の抜本的な見直しを経て、その垣根がすでに取り払われており、特に、データ分析技術との連携による視覚的分析技術に関する研究開発が活発化している。計算科学分野では、大規模・高性能化が進む HPC 環境下での大規模データ可視化は解決されることのない問題として存在し続けており、効率の良い画像化が今もなお要求されている。先に述べた In-situ 可視化の発展は、この状況を打破する可能性を秘めており、HPC 向け視覚的解析技術に向けた取り組みの大きな進展が期待される。視覚的分析に関する研究開発においては、可視化研究分野の 6 エリアモデルが示すように、データ科学の関連分野との密接な連携が重要である。そのため、HPC 環境の利活用においては、シミュレーションと可視化の連携に加え、データ分析処理も考慮した統合的なシステム開発が要求される。一般に、各分野の最先端の研究開発では、方法の新規性や特殊性から、ツールとして提供されるソフトウェアがしばしば複雑化する。そのため、研究者やエンジニアは汎用的なツールよりも独自に開発した特殊なツールを利用することが多い。各分野内では、データモデルやワークフローの汎用化や統一が進められるが、それらの連携が必要となった場合、様々な問題が生じる。そのため、分野ごとに異なるコード間のデータ変換や高度に抽象化された新しいデータモデルが必要となる。また、異なるシステムの統合においてはタスクの構成が複雑化するた

め、高い汎用性に加え分野固有の抽象化も想定した新しいワークフローが必要となる。そして、これらのデータモデルとワークフローをサポートする実行基盤の開発が課題となる。また、分析まで想定した計算の In-situ 化 (In-situ 分析) が必要となる。In-situ 可視化による画像データベースの構築は、指定されるタスクと結果画像の学習によるワークフロー構築の自動化や、可視化パイプラインの推定およびそれを構成する各データ処理の高速化など様々な場面で機械学習の適用可能性を高めることが期待される。また、機械学習によるデータ駆動型偏微分方程式推定 [41] や物理知識をもとにした機械学習による偏微分方程式解法 [48] など、これまでには考えられていなかった全く新しい計算の方法が提案されており、シミュレーションと機械学習の結びつきがより一層強まる予測される。実際に、機械学習を使って物理シミュレーションを代替えするサロゲートモデルに関する研究が CAE 方面を中心に注目されており、可視化研究分野でもサロゲートモデルを使って視覚的にパラメータ探索を効率よく行うことができる手法 [49, 50] なども提案され始めている。計算コストが高い機械学習のトレーニング処理は In-situ 化 (In-situ 学習) が必要であろう。これらの研究成果は、数値シミュレーションを視覚的分析基盤における試行錯誤ループのいちモジュールとして機能させることができる構成を秘めている。計算精度の改善は必要だが、用途を限定した特定分野に対して実用化が期待できる。

(3) 可視化による HPC システムの高機能・高効率化

HPC システムにおける可視化は、大規模データの画像化技術からデータの解析技術へと発展している一方で、システムの挙動理解への活用が進められており、間接的に計算科学の進展に貢献することが期待できる。データ解析までを想定した HPC 環境を想定すると、システムアーキテクチャの発展や構成の高度化に対して、大規模な計算を効率よく安定的に実行するためには、システムの状態や性能を監視し、耐障害性や資源利用の最適化を実現する必要がある。この目的を達成するために、HPC システムの実行ログデータのデータベース化とそれを活用した可視化およびデータ分析が重要となる。一般にログデータは、多次元時系列データとして表現されることが多く、そのように表現されるデータからの特徴構造を抽出技術 [41] や、機械学習による異常状態の自動検知技術、電力消費量を考慮した動的なタスク割り当てなどの応用が容易に想定され、さらに可視化技術を活用することで、運用者のシステム挙動の正しい理解につながることが期待される。そして、HPC システム上での計算処理の実行が効率化され、結果として、知識発見時間の短縮に貢献することが期待される。

4.7.4.3 ロードマップ

可視化・データ処理分野のロードマップを図 4.29 に示す。

4.7.4.4 必要な計算資源

可視化は、一般的に、データ駆動型の処理となるため、対象とするデータスケールに応じて必要な計算資源も大きく変わるためにそのコストを見積ることは困難である。In-situ 可視化においては、計算結果をストレージに出力する必要がないため、データ I/O コストが大幅に削減されるが、分析まで想定した In-situ 可視化においては、コストモデルとその利用自体が未解決問題の一つとされている [30]。

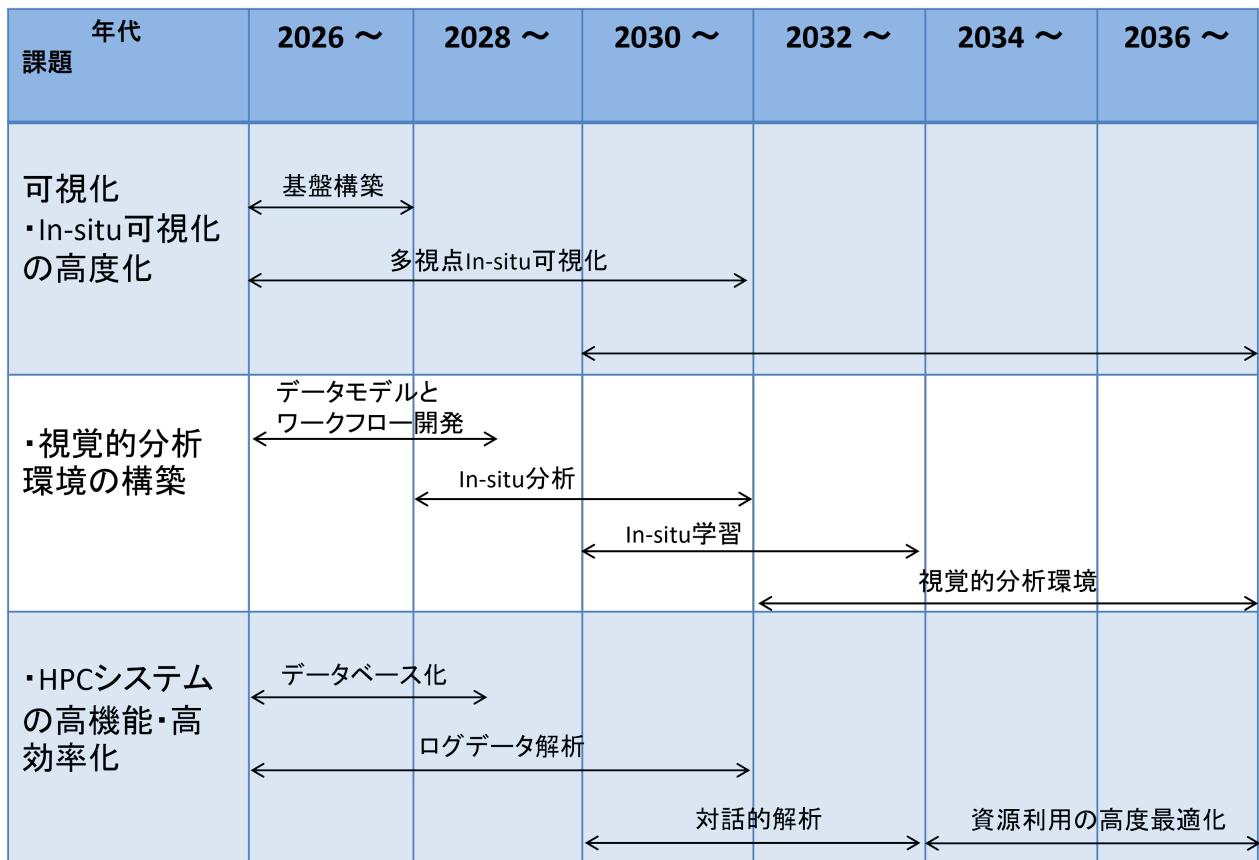


図 4.29 ロードマップ (可視化・データ処理分野)

GPU を前提した現在の可視化環境では、中小規模データであればリアルタイム処理が実現されており、In-situ 可視化による大規模データ可視化の場合でも、数値シミュレーションの計算コストと比較しても計算資源を圧迫するような状況にはならない。しかし、多視点 In-situ 可視化では視点数に応じた可視化処理コストが増加し、対話的なデータ探索を可能とするような高度な In-situ 可視化においては、単純な描画処理以外のデータ処理のコストが増加する傾向にある。また、データ分析処理や機械学習処理の In-situ 化が進むと、さらに大きな計算資源が必要となる。エクサスケール時代の高度な可視化・データ処理を実現するためには、シミュレーションと同等またはそれ以上の計算資源を想定しておく必要がある。

参考文献

- [1] O C Zienkiewicz and Robert Leroy Taylor. *The Finite Element Method: Solid mechanics*. Butterworth-Heinemann, 2000.
- [2] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, January 2003.
- [3] Barry Smith, Petter Bjorstad, and William Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, USA, June 1996.

- [4] Andrea Toselli and Olof B Widlund. *Domain Decomposition Methods - Algorithms and Theory*. Springer Berlin Heidelberg.
- [5] Tarek Poonithara Abraham Mathew. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Springer Berlin Heidelberg.
- [6] A J Brizard and T S Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, Vol. 79, No. 2, pp. 421–468, April 2007.
- [7] X Garbet, Y Idomura, L Villard, and T H Watanabe. Gyrokinetic simulations of turbulent transport. *Nucl. Fusion*, Vol. 50, No. 4, p. 043002, April 2010.
- [8] 洲鎌英雄, 渡邊智彦. ジャイロ運動論による磁化プラズマ乱流の研究: 核融合から宇宙まで(解説). 日本物理学会誌, Vol. 68, No. 5, pp. 296–304, May 2013.
- [9] M Nunami, S Toda, M Nakata, and H Sugama. Improved prediction scheme for turbulent transport by combining machine learning and first-principle simulation.
- [10] Yasuhiro Idomura, Masato Ida, Takuma Kano, Nobuyuki Aiba, and Shinji Tokuda. Conservative global gyrokinetic toroidal full-f five-dimensional vlasov simulation. *Comput. Phys. Commun.*, Vol. 179, No. 6, pp. 391–403, September 2008.
- [11] Y Todo and T Sato. Linear and nonlinear particle-magnetohydrodynamic simulations of the toroidal alfvén eigenmode. *Phys. Plasmas*, Vol. 5, No. 5, pp. 1321–1327, May 1998.
- [12] T-H Watanabe and H Sugama. Velocity-space structures of distribution function in toroidal ion temperature gradient turbulence. *Nucl. Fusion*, Vol. 46, No. 1, p. 24, December 2005.
- [13] Kenji Imadera and Yasuaki Kishimoto. ITB formation in gyrokinetic flux-driven ITG/TEM turbulence. *Plasma Phys. Controlled Fusion*, Vol. 65, No. 2, p. 024003, December 2022.
- [14] B H McCormick, T A DeFanti, and M D Brown. Visualization in scientific computing and computer graphics. *ACM SIGGR. Symp. Comput. Animat.*
- [15] Tamara Munzner, Chris Johnson, Robert Moorhead, Hanspeter Pfister, Penny Rheingans, and Terry S Yoo. NIH-NSF visualization research challenges report summary. *IEEE Comput. Graph. Appl.*, Vol. 26, No. 2, pp. 20–24, 2006.
- [16] Hans Hagen, Daniel Keim, Tamara Munzner, Stephen North, and Hanspeter Pfister. VIS restructuring recommendations (fall 2018), 2018.
- [17] Hank Childs, Kwan-Liu Ma, Hongfeng Yu, Brad Whitlock, Jeremy Meredith, Jean Favre, and Scott Klasky. In situ processing. In E W Betherl, H Childs, and C Hansen, editors, *High Performance Visualization: Enabling Extreme Scale Scientific Insight*, pp. 171–198. Chapman and Hall, 2012.
- [18] Kitware, Inc. ParaView. <https://www.paraview.org/>. Accessed: 2023-12-3.
- [19] LLNL. VisIt. <https://sd.llnl.gov/simulation/computer-codes/visit>. Accessed: 2023-12-3.
- [20] Utkarsh Ayachit, Andrew Bauer, Berk Geveci, Patrick O’Leary, Kenneth Moreland, Nathan Fabian, and Jeffrey Mauldin. ParaView catalyst: Enabling in situ data analysis and visualization. In *Proceedings of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale*

- Analysis and Visualization*, ISAV2015, pp. 25–29, New York, NY, USA, November 2015. Association for Computing Machinery.
- [21] Brad Whitlock, Jean M Favre, and Jeremy S Meredith. Parallel in situ coupling of simulation with a fully featured visualization system. In *Proceedings of the 11th Eurographics conference on Parallel Graphics and Visualization*, EGPGV '11, pp. 101–109, Goslar, DEU, April 2011. Eurographics Association.
 - [22] Mesa Project. The mesa 3D graphics library. <https://www.mesa3d.org/>. Accessed: 2023-12-3.
 - [23] Ingo Wald, Sven Woop, Carsten Benthin, Gregory S Johnson, and Manfred Ernst. Embree: a kernel framework for efficient CPU ray tracing. *ACM Trans. Graph.*, Vol. 33, No. 4, pp. 1–8, July 2014.
 - [24] I Wald, G P Johnson, J Amstutz, C Brownlee, A Knoll, J Jeffers, J Günther, and P Navratil. OSPRay - a CPU ray tracing framework for scientific visualization. *IEEE Trans. Vis. Comput. Graph.*, Vol. 23, No. 1, pp. 931–940, January 2017.
 - [25] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. Pearson Education, Incorporated, 2006.
 - [26] Kenneth Moreland, Christopher Sewell, William Usher, Li-Ta Lo, Jeremy Meredith, David Pugmire, James Kress, Hendrik Schroots, Kwan-Liu Ma, Hank Childs, Matthew Larsen, Chun-Ming Chen, Robert Maynard, and Berk Geveci. VTK-m: Accelerating the visualization toolkit for massively threaded architectures. *IEEE Comput. Graph. Appl.*, Vol. 36, No. 3, pp. 48–58, 2016.
 - [27] Utkarsh Ayachit, Brad Whitlock, Matthew Wolf, Burlen Loring, Berk Geveci, David Lonie, and E Wes Bethel. The SENSEI generic in situ interface. In *2016 Second Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV)*, pp. 40–44. IEEE, November 2016.
 - [28] David A Boyuka, Sriram Lakshminarasimham, Xiaocheng Zou, Zhenhuan Gong, John Jenkins, Eric R Schendel, Norbert Podhorszki, Qing Liu, Scott Klasky, and Nagiza F Samatova. Transparent in situ data transformations in ADIOS. In *2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 256–266. IEEE, May 2014.
 - [29] Matthew Larsen, James Ahrens, Utkarsh Ayachit, Eric Brugger, Hank Childs, Berk Geveci, and Cyrus Harrison. The ALPINE in situ infrastructure: Ascending from the ashes of strawman. In *Proceedings of the In Situ Infrastructures on Enabling Extreme-Scale Analysis and Visualization*, ISAV'17, pp. 42–46, New York, NY, USA, November 2017. Association for Computing Machinery.
 - [30] Janine C Bennett, Hank Childs, Christoph Garth, and Bernd Hentschel. In situ visualization for computational science (dagstuhl seminar 18271). 2019.
 - [31] Naohisa Sakamoto and Koji Koyamada. Stoachastic approach for integrated rendering of volumes and semi-transparent surfaces. In *2012 SC Companion: High Performance Computing, Networking Storage and Analysis*, pp. 176–185. IEEE, November 2012.

- [32] Takuma Kawamura, Tomoyuki Noda, and Yasuhiro Idomura. In-Situ visual exploration of multivariate volume data based on particle based volume rendering. In *2016 Second Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV)*, pp. 18–22. IEEE, November 2016.
- [33] Jorji Nonaka, Kenji Ono, and Masahiro Fujita. 234compositor: A flexible parallel image compositing framework for massively parallel visualization environments. *Future Gener. Comput. Syst.*, Vol. 82, pp. 647–655, May 2018.
- [34] Akira Kageyama and Naohisa Sakamoto. 4D street view: a video-based visualization method. *PeerJ Comput. Sci.*, Vol. 6, p. e305, November 2020.
- [35] Yoshiaki Yamaoka, Kengo Hayashi, Naohisa Sakamoto, and Jorji Nonaka. In situ adaptive timestep control and visualization based on the spatio-temporal variations of the simulation results. In *Proceedings of the Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization*, ISAV ’19, pp. 12–16, New York, NY, USA, November 2019. Association for Computing Machinery.
- [36] Naoko Sawada, Makoto Uemura, Johanna Beyer, Hanspeter Pfister, and Issei Fujishiro. Time-TubesX: A query-driven visual exploration of observable, photometric, and polarimetric behaviors of blazars. *IEEE Trans. Vis. Comput. Graph.*, Vol. 28, No. 4, pp. 1917–1929, April 2022.
- [37] Shigeo Takahashi, Daisuke Sakurai, Miyuki Sasaki, Hiroko N Miyamura, and Yukihisa Sanada. Visual analysis of geospatial multivariate data for investigating radioactive deposition processes. *Vis. Comput.*, Vol. 37, No. 12, pp. 3039–3050, December 2021.
- [38] Shilpika, Takanori Fujiwara, Naohisa Sakamoto, Jorji Nonaka, and Kwan-Liu Ma. A visual analytics approach for hardware system monitoring with streaming functional data analysis. *IEEE Trans. Vis. Comput. Graph.*, Vol. 28, No. 6, pp. 2338–2349, June 2022.
- [39] Koji Koyamada, Yu Long, Takuma Kawamura, and Katsumi Konishi. Data-driven derivation of partial differential equations using neural network model. *Int. J. Model. Simul. Sci. Comput.*, Vol. 12, No. 02, p. 2140001, April 2021.
- [40] Hiroaki Natsukawa, Ethan R Deyle, Gerald M Pao, Koji Koyamada, and George Sugihara. A visual analytics approach for ecosystem dynamics based on empirical dynamic modeling. *IEEE Trans. Vis. Comput. Graph.*, Vol. 27, No. 2, pp. 506–516, February 2021.
- [41] Fujiwara, Shilpika, Sakamoto, Nonaka, Yamamoto, and Ma. A visual analytics framework for reviewing multivariate Time-Series data with dimensionality reduction. *IEEE Trans. Vis. Comput. Graph.*, Vol. 27, pp. 1601–1611, February 2021.
- [42] Zipeng Liu, Takayuki Itoh, Jessica Q Dawson, and Tamara Munzner. The sprawlter graph readability metric: Combining sprawl and Area-Aware clutter. *IEEE Trans. Vis. Comput. Graph.*, Vol. 26, No. 6, pp. 2180–2191, June 2020.
- [43] 日本学術会議 総合工学委員会 機械工学委員会合同計算科学シミュレーションと工学設計分科会. 科学的知見の創出に資する可視化に向けて, 2017.

- [44] James Ahrens, Sebastien Jourdain, Patrick O'Leary, John Patchett, David H Rogers, and Mark Petersen. An Image-Based approach to extreme scale in situ visualization and analysis. In *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 424–434. IEEE, November 2014.
- [45] Naohisa Sakamoto. Smart in-situ visualization for large-scale numerical simulations aiming at efficient knowledge acquisition. 2021.
- [46] Ken Iwata, Go Tamura, Naohisa Sakamoto, and Jorji Nonaka. Smart in-situ visualization framework on the fugaku environment. https://sighpc.ipsj.or.jp/HPCAsia2022/poster/128_abstract.pdf, 2022. Accessed: 2023-12-3.
- [47] J Nonaka, Go Tamura, K Iwata, K Watanabe, Y Maejima, and N Sakamoto. Developments of visualization and analysis oriented In-Situ processing on the fugaku environment. 2022.
- [48] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, Vol. 3, No. 6, pp. 422–440, May 2021.
- [49] Wenbin He, Junpeng Wang, Hanqi Guo, Ko-Chih Wang, Han-Wei Shen, Mukund Raj, Youssef S G Nashed, and Tom Peterka. InSituNet: Deep image synthesis for parameter space exploration of ensemble simulations. *IEEE Trans. Vis. Comput. Graph.*, Vol. 26, No. 1, pp. 23–33, January 2020.
- [50] Neng Shi, Jiayi Xu, Skylar W Wurster, Hanqi Guo, Jonathan Woodring, Luke P Van Roekel, and Han-Wei Shen. GNN-Surrogate: A hierarchical and adaptive graph neural network for parameter space exploration of Unstructured-Mesh ocean simulations. *IEEE Trans. Vis. Comput. Graph.*, Vol. 28, No. 6, pp. 2301–2313, June 2022.

4.8 社会科学

本節では社会科学分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.9 地震・津波

本節では地震・津波分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

参考文献

- [1] 著者 1, 著者 2, 著者 2. 文献名. 出版社, 2022.
- [2] 日本学術会議. ロードマップ. <http://xxx.yyy.zzz>, 2022.

4.10 気象・気候

本節では気象・気候分野における2032年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。2.10.3節で述べたように、気象・気候モデルは、扱う対象や目的に応じて複数のタイプのモデルが開発されてきた。4.10.1節ではそれぞれのタイプのモデルを用いたシミュレーションについて詳述する。これらのモデルに含まれる物理プロセスや海洋や化学といった要素モデルの改良も重要な課題であり、4.10.2節で取り上げる。発展が著しいAI（人工知能）の活用を含むデータ同化・データサイエンスについては4.10.3節で述べる。近年の気象・気候モデル開発に大きな影響を与えていたりハードウェアの動向について4.10.4節で解説する。

4.10.1 気象・気候シミュレーション

4.10.1.1 全球高解像度シミュレーション

背景

全球大気を水平格子サイズ数km以下で解く数値モデルは、大循環から深い対流までをシームレスに解き、マルチスケール相互作用が重要な役割を果たす多くの気象現象の再現と予測にとって有用なツールである。特に台風は、現象全体の水平スケールは数百kmスケールであるが、眼の構造は10～数10km程度のスケールを持つ。また、台風の発生は数千kmスケールの環境場による影響が大きいことから、台風の理解と予測において全球高解像度モデルが果たす役割は大きい。また、台風は日本を含む東・東南アジアや北米において甚大な人的・経済的損失をもたらしており、台風の動向や予測に対する社会的な関心は高い。日本国内では、2019年台風19号（東日本台風）（死者・行方不明者108人）、2011年台風12号（同98人）、2004年台風23号（同98人）のように、近年においても100人規模の死者・行方不明者が発生している^{*8}。海外に目を向けると、2008年のサイクロン・ナルギス（ミャンマー：同13万人）、2013年の台風30号（フィリピン：同6千人）のように、特に東南アジアの発展途上国において台風は未だに膨大な数の人命を奪っている。

防災・減災や温暖化適応策の観点から、台風をはじめ豪雨や暴風といった極端現象について、様々な時間スケールの予測を改善する努力が続けられてきた。例えば、1週間程度先までの決定論的な台風予報については、進路予報の精度は長期的には向上している^{*9}が、近年は頭打ちの兆しがある。台風強度の予報精度向上はさらに難しく、観測データの充実とともにモデルの精緻化が重要な課題となっている。気象庁は2030年に向けた数値予報技術開発重点計画の主な開発項目^{*10}の一つとして台風防災を掲げており、台風予報研究は気象庁と研究者の連携強化のもとで重点的に進められている。週を超える台風予報や月・季節スケールの台風発生傾向の予測については研究段階であり、モンスーンや偏東風波動、季節内振動、上層の寒冷渦といった台風発生と関連が深い現象の理解と予測が重視されている。また、季節予測では海洋予測の重要性が増してくる。気候スケールでは、地球温暖化によって強い台風の割合が増加する傾向が最

*8 令和3年版防災白書（内閣府）：<http://www.bousai.go.jp/kaigirep/hakusho/r3.html>

*9 https://www.data.jma.go.jp/fcd/yoho/typ_kensho/typ_hyoka_top.html

*10 https://www.jma.go.jp/jma/kishou/know/tecdev/nwp_strategic_plan_towards_2030.html

新の IPCC 第 6 次評価報告書 (AR6)においても予測されている [1]。しかし、海盆毎、地域毎の台風予測については不確定性が大きい。近年の観測に基づく研究においても、東アジアに上陸する台風の強度の増加 [2]、東京に接近する台風の増加 [3]、台風移動速度の鈍化 [4] などの傾向が既にみられると指摘されており、温暖化適応策の観点からも信頼性の高い将来気候予測が求められている。

近年、地球のデジタルツインである “Digital Earth” の構築が提唱され [5]、欧州において大型プロジェクト^{*11}が開始されるとともに、WCRP（世界気候研究計画）において今後 10 年を見据えた Digital Earths Lighthouse Activity^{*12} が開始されるなど、Digital Earth は 2020 年代のトレンドになりつつある。Digital Earth は利用可能なあらゆる観測データを統合し、数値モデルやデータ科学の力を借りながら、過去・現在の地球の気候状態を仮想世界上に可能な限り忠実に再現し、さらに将来をできるだけ高い信頼度で予測しようという試みである。将来予測についてはユーザがシナリオを自在に選択でき、政策の効果を定量的に検証することも可能である。このような Digital Earth 構築の大前提として、できる限り多くのプロセスを取り入れた精緻な気象・気候モデルが必要となる。

以上のような社会的課題を解決するためには、用いる気象・気候モデルの再現・予測性能に対する信頼度を現状よりも高めていく必要がある。さらに学術的課題の解決までを見据えると、モデル性能の改善を着実に進めるアプローチとして、地球大気をできるだけ高解像度で計算することが有効である。地球大気は全球規模の大スケール ($\sim 10,000 \text{ km}$) から雲・乱流といった小スケール (10 m 以下) の現象が相互に作用する複雑な系である。特に水平解像度を数 km 程度まで高めると、対流圏全層に渡る深い対流がモデル格子で直接解像され始め、雲・降水システムの表現が質的に大きく向上する [6, 7]。このような全球雲解像 (cloud-resolving) モデル^{*13}では、雲・循環・放射の相互作用を物理的に整合した方法で表現できるため、雲フィードバックへの理解が進展し、現在の気候モデルでは不確定性が大きい気候感度の制約に結びつく可能性がある。また、対流の自己組織化や季節内振動、モンスーン、日変化といった熱帯気象の中心的な課題に対して特に強力なツールであるとともに、台風の基本的な構造が表現されることで最大風速などの予測の向上につながることが期待できる。さらに水平・鉛直解像度を 100 m 程度以下まで微細化できれば乱流の一部が解像されはじめる。このような全球を対象とする Large-eddy シミュレーション (LES) では大気境界層の表現が飛躍的に向上し、現在では直接的な再現が難しい、下層雲を含む全球の雲の高さや光学的厚さの分布について、自励的に再現されることが期待できる。このような全球 LES が気候スケールで実用的になれば、雲と気候の相互作用に対する理解が飛躍的に進展すると予想される。また、自励的に発生した台風について微細構造 [8] が再現できるようになり、疑似観測データとして活用することで台風発達の鍵を握る境界層スキームの改善につながることも期待できる。

全球雲解像実験に向けた新しいモデル開発の取り組みは、2000 年代初頭に日本 (NICAM) およびドイツ (ICON) において開始された [9]。特に NICAM は、世界初の全球雲解像実験 [10]、及び、マッデン・ジュリアン振動 (MJO) の再現実験 [11] に成功して以来、多くの実績を残してきた。米国では 2 次元の雲解像モデルを粗い全球モデルの格子に埋め込む super-parameterization[12, 13] が利用されてきた。

*11 Destination Earth : <https://digital-strategy.ec.europa.eu/en/policies/destination-earth>

*12 <https://www.wcrp-climate.org/digital-earths>

*13 類似の呼称として、全球嵐解像 (storm-resolving; $dx < \sim 5\text{km}$) モデル、全球対流許容 (convection-permitting) モデルといった呼称も広く使われている。呼称には議論があるが、数 km 程度の水平格子サイズで雲微物理スキームを用いる全球非静力学モデル、という意味では実質的な違いはあまりないと考えられる。

計算機の高速化を背景に、2010年代後半になると多くのモデルグループにおいて全球雲解像モデルの実用化が進み、日本およびドイツの主導によってモデル相互比較プロジェクト DYAMOND[14] が開始された。ここ数年は全球雲解像モデルを気候モデルとして実用化する動きが大型プロジェクト^{*14}によって急速に進められている。全球LESについては現在の計算機の能力では実現のハードルが高いが、NICAM、ICONのグループが実現に向けた取り組みを進めている。計算資源の削減のため、3次元LESモデルを親モデルのグリッドに埋め込む方法[15] や、特定の物理スキームのみ鉛直層を増強する方法[16] なども提案されている。いずれにせよ、ここまで述べてきたような高解像度全球シミュレーションの出力は膨大であり、データの移動や解析は今後ますます困難になると予想される。これに対応して ECMWF では、クラウド上でシミュレーション・解析を行う “Digital Earth” システムのプロトタイプを 2023 年頃にリリースする見込みであり、他機関への波及も予想される^{*15}。

学術的・技術的課題、目標

社会的・学術的課題を解決するため、高解像度全球モデルを用いた研究としてはさらなる高解像度化、大アンサンブル化、長期積分化、精緻化の方向性が考えられる。それについて、5~10年程度先の代表的な目標をロードマップとして示す。特に全球LESの実現は全球雲解像実験以来のマイルストーンであり、技術トレンドを考慮すると2020年代後半の実現が想定される。大アンサンブル化、長期積分化、精緻化については、水平解像度3.5km程度の全球雲解像モデルをベースにした研究が主流になると予想される。富岳での研究により $O(10^3)$ アンサンブルの有用性が少しずつ明らかになってきている。気候スケールの研究については、全球雲解像モデルを用いた100年気候実験や10年アンサンブル気候実験が視野に入るだろう。また、Digital Earth構築に向けた必然的な流れとして全球雲解像モデルの精緻化が加速し、地球システムモデルとしての利用可能性が模索されていくと予想される。

DYAMOND モデルに代表される多くの高解像度全球モデルでは、水平格子として準一様格子を用いている。一般的な全球モデルでは水平離散化法として球面調和関数を用いたスペクトル法が主流であるが、時間ステップ毎に大域通信を必要とする上、球面展開の演算量が水平解像度に対して非線形に増大する。従って、高解像度全球モデルではスペクトル法は回避される傾向にあるが、IFS (ECMWF) は水平解像度5kmモデルにおいてもスペクトル法を利用している。演算量の削減という観点では二重フーリエ法の有効性も検討されている[17]。準一様格子における離散化手法や山岳の取り扱いについても課題が残っている。

一般的な全球モデルでは、対流・雲スキームを用いることで格子スケール未満の対流・雲が格子スケールへ与える影響を表現している。しかし、このような診断型のスキームは半経験則に基づくため定式化やパラメータの任意性が大きく、特に将来気候予測において不確定性をもたらす大きな要因とされている。水平数kmスケールの全球雲解像モデルでは、雲や雨の質量などを予報する雲微物理スキームを用いるのが一般的であるが、追加的に対流・雲スキームを導入するかどうかはモデルによって判断が分かれている

^{*14} 海外における代表事例としては、NextGEMS（欧州）、E3SM（米国）、EarthWorks（米国）が挙げられる。国内では DNA 気候学 (<https://dna-climate.org/>) および富岳成果創出加速課題 (<https://cesd.aori.u-tokyo.ac.jp/fugaku/>) が全球雲解像気候実験に向けた研究を推進している。

^{*15} 最近のトレンドについては WMO が公開予定の白書 (WMO Research Board Concept Note on Exascale Computing and Data) も参照のこと。

[14]。水平数 km スケールでは深い対流は解像され始めるがまだ不十分であり、水平スケールが小さく背の低い浅い対流に至っては解像できない。放射への影響を考慮して多くのモデルでは部分雲診断スキームを導入しているが、雲・放射・対流の整合性という観点では理想的には望ましくない。また、雲微物理スキームで用いられているバルク法の精緻化や超液滴法 [18] のようなラグランジュ法の全球モデルへの導入も今後 5~10 年の検討課題になるだろう。水平数 km スケールというモデルの格子幅は衛星観測のフットプリントに近いことから、雲微物理スキームにおける未知パラメータについては、雲・降水過程を観測可能な衛星データから制約する試みも行われている。しかし、スキームが前提とする空間スケールとモデル格子幅との間には大きな乖離があるため、適切な気候場が再現されるという保証はない。超高解像度実験や詳細で非常に計算量の多い物理過程スキームによる実験を行って結果の収束性を探ることも、フラグシップマシンを用いた研究に求められる役割であろう。それらのチャンピオンデータは、データ同化や AI を用いたパラメータ・チューニングにおいても教師データとして有効に活用されることが期待される。

全球 LES に至るまでの水平メッシュサイズ 100 m~1km では、境界層内の対流や乱流におけるグレーゾーン問題に直面する [19]。気象モデルで用いられてきた予報式を診断式に簡略化したクロージャ境界層モデルの仮定が成り立つ範囲と、慣性小領域と呼ばれる領域より大きなスケールの渦を直接解像する LES での仮定が成り立つ範囲のちょうど狭間に位置するため、乱流スキームの扱いには注意を要する。さらに水平・鉛直解像度を 100 m スケール以下にまで微細化した LES は、多くの場合は乾燥大気を仮定して理論構築がされているため、気象に適応する場合には湿潤過程を考慮した理論の再構築が必要である。また、多くの全球モデルでは力学コアにおいて 2 次精度の空間差分式を採用しているが、これによる数値誤差は LES におけるサブグリッドスケールの渦粘性項の寄与を上回ってしまう。気象モデルにおける高次精度スキームの検討 [20] が進められるとともに、局所化と高精度化を両立した不連続ガラーキン法といった新しいスキームが次世代全球 LES の力学コアとして検討されている。

シミュレーションの大アンサンブル化については、これまでのモデルの延長線上で比較的容易に達成できる。特に社会的要請の大きい週スケールの台風進路予測や週を越えた台風の発生・進路や豪雨発生の確率予測といった研究が今後も発展するだろう。また、季節スケールの天候予測は農産業や水産業にとって重要であるが、初期値問題であるとともに境界条件（陸面・海面条件や大気組成等の外部条件など）も重要な時間スケールであるため、予測可能性の有無を含めてさらなる理解が必要である。このような時間スケールでは季節内振動やエルニーニョ現象といった月から季節スケールのメモリを持った現象の再現が重要であり、海洋と結合した全球雲解像モデルを用いたアンサンブル実験が多用されるであろう。並行して、大気・海洋の初期条件を作成するためのデータ同化手法も検討する必要がある。

気候予測という観点では、ストロング・スケール問題により全球雲解像モデルを数十年以上にわたって連続積分することは現時点では困難であり、モデルのさらなる高速化が技術的な課題である。台風や豪雨といった極端現象を統計的に取り扱うため、全球雲解像モデルを用いたアンサンブル気候実験の活用も進むであろう。

目標達成に向けて必要な取り組み

モデルの高解像度化を達成するためには、計算ノードの増加に対してウィーク・スケーリング性能が維持されることが重要である。演算が局所化された全球モデルであれば、理論上はモデル本体部分のウィーク・スケーリング性能の維持はそれほど難しくない。但し、時間刻みを解像度とともに小さくする必要が

年代 課題	2022～	2024～	2026～	2028～	2030～	2032～
全球規模での高解像度予測	全球LESモデル開発/全球200m実験 \leftarrow $O(10^3)$ アンサンブル予報実験 \leftarrow 全球雲解像10年気候実験 \leftarrow 全球雲解像地球システムモデル開発 \leftarrow	\rightarrow \rightarrow \rightarrow \rightarrow	\leftarrow $O(10^3)$ アンサンブル全球雲解像予報実験 \leftarrow 全球雲解像100年気候実験 \leftarrow 全球雲解像地球システム実験 \leftarrow	\rightarrow \rightarrow \rightarrow \rightarrow		

図 4.30 気象・気候分野、全球高解像度シミュレーションのロードマップ

あり、同じ積分期間に対する実行時間は解像度とともに遅くなってしまう。また、メッシュの微細化によるノードインバランスの拡大、ノード数増加による計算・IO エラー処理、といった課題が顕在化する可能性もある。また、NICAM（恐らく多くの高解像度モデル）では水平方向に MPI 並列、鉛直方向にスレッド並列を採用しているが、鉛直層数の増加によるノードあたりの必要メモリ量の増大にも対応する必要がある。

全球 LES を実現するためには、理論上は富岳（全球 200m 実験が可能）に対して 16 倍（空間格子数 8 倍、時間ステップ数 2 倍）程度の実行性能が必要である。これを現在の物理モデルだけの枠組みで 5~10 年後に達成するためには、モデル全体の演算加速器への最適化、低精度化、もしくは現在よりも高い B/F 比のマシンが不可欠であると予想される。代理モデルの活用までを考慮すると、広い意味での全球 LES の実現は間近の可能性もある^{*16}。

大アンサンブル化については capacity computing であることから、現在のモデルフレームワークのままでも比較的容易に達成できる可能性が高い。アンサンブル数に比例して出力データサイズが増大することから、ワークフロー全体を考慮すると解析や可視化までも含めた並列化を行う必要性が高まるであろう。

長期積分を現実的な実時間で実行するためにはストロング・スケーリング性能が重要になってくるが、ウィーク・スケーリング性能に比べて大規模ノードでの確保は極めて難しい。その要因として、ループ長短縮によるメモリアクセス・SIMD 演算・パイプライン処理の非効率化、単位演算あたりのノード間通信量の増大、ノードインバランスの顕在化、といった多くの点が考えられる上、これらはターゲットマシンの仕様に強く依存するため、長期的な性能向上は容易ではない。

アプリケーションと計算量・計算資源

アプリケーション

- モデル名：NICAM (Non-hydrostatic Icosahedral Atmospheric Model) [22, 23, 24, 25]
- 力学コア：正二十面体構造格子、有限体積法、完全圧縮、非静水圧近似のナビエストークス方程式
- 物理過程：雲微物理（シングルもしくはダブルモーメントバルク法）、大気放射、乱流、陸面過程

*¹⁶ DNS の効果を代理モデルによって LES に取り込んだ例：Kochkov et al. (2021)[21]

など

- 海洋モデル COCO[26] と結合して、大気海洋結合モデルとして実行することも可能

計算量・計算資源

- 問題規模
 - 格子点数：17兆（水平解像度 110m、鉛直 400 層： $400 \times 10 \times 4^{16}$ ）
 - ステップ数：70万（ $dt=0.25$ 秒、2日積分）
- 総メモリ量：51PByte（格子点数に比例すると仮定し、前回ロードマップより推定： $12/94 \times 400$ ）
- 総演算量：340 ZFLOP（格子点数・ステップ数に比例すると仮定し、前回ロードマップより推定： $1,216 \times (400/96) \times (70/1040)$ ）
- 実行時間：2週間（10 EFLOPS、実行性能 3% を仮定）
- アプリケーション B/F 比（演算密度の逆数）：2.1
- 総メモリ転送量：総演算量 × アプリケーション B/F 比 → $340\text{ZFLOP} \times 2.1 = 714\text{ZByte}$
 - 要求メモリ転送性能： $1/(10\text{EFLOPS} \times 3\%) \times 2.1 = 0.63\text{EByte/s}$
- 総ストレージ使用量：29PByte（実行時間中に 4 バイトの 3 次元変数換算で 10 変数を 1 時間毎に出力すると仮定； $4 \times 10 \times 17\text{兆} \times 24 \times 2$ ）
- 要求ファイル IO 性能：0.6 TB/s ($29\text{PByte}/14/3600$)

4.10.1.2 全球気候シミュレーション

背景

地球温暖化の影響が顕在化している、と考える専門家が多い。実際、2021年8月に公表された「気候変動に関する政府間パネル」(IPCC) の第一作業部会(WGI)による第6次評価報告書(AR6)で、「人間の影響が大気、海洋及び陸域を温暖化させてきたことには疑う余地がない」という強い表現で人間活動の気候への影響が断言されたことは、メディアでも盛んに報じられた[1]。温暖化が人間活動によるものと断定されることを受け、科学的理 解の増進に加え、温暖化への対策立案に資するデータの創出が気候モデルの大きな役割として注目されてきている。

こうした状況のもと、気候モデルを高解像度化し、今後避けられない温暖化に社会が適応するための政策(適応策)立案に資するよう、地域ごとに詳細な情報を発信する動きが盛んになってきている[27]。この場合、全球を対象とした気候モデルでは計算機資源の面から高解像度化に限界があり、最も高解像度のモデルでも格子間隔 20km 程度である。この解像度では、細かな地形の影響を受ける降水分布などについて、精度の高い再現性が望めないため、日本周辺など特定の領域を対象とした領域モデルが用いられることが多い。ただし、対象領域周辺に関する情報(境界条件)は全球モデルによる予測結果から与える必要があるため、全球モデルと領域モデル両方の再現性や予測精度を向上させていく必要がある。

このように、全球モデルで得られた予測データの一部を境界条件として領域モデルに与え、対象領域に関する予測データを高解像度化する手法を、力学的ダウンスケーリングと呼ぶ(他に、統計的なモデルを用いて高解像度化を行う統計的ダウンスケーリングという手法もある)。実際、温暖化の影響評価のため

に作成されたデータセット d4PDF[28] では、力学的ダウンスケーリングとアンサンブル実験によって大規模なデータセットを作成し、様々な分野での影響評価に役立っている [29]。

領域ごとの詳細な情報提供は、日本だけでなく世界的にも大きな趨勢になっている。例えば、前出の IPCC AR6(WGI) では、地域的な空間規模における温暖化予測について、全体 12 章のうち 3 章が充てられ、全球規模変化との関連、台風などの極端現象の変化、災害影響などについて評価がなされている。

また、将来予測だけでなく、現在発生している異常高温や集中豪雨に関し、温暖化がどの程度寄与しているかを評価する、「イベント・アトリビューション」と呼ばれる研究も盛んになってきている。例えば今田ら [30] は、2018 年に発生した日本の猛暑は、温暖化の寄与がなければほぼあり得ないレベルの高温であったことを明らかにしている。国際的にも、アメリカ気象学会誌 (BAMS) が毎年 1 回イベント・アトリビューションの特集を組むことにするなど、この種の研究の発展は著しい。極端現象発生に対する温暖化の寄与の評価は、適応策を立案するうえでも重要な情報となる。

一方、適応策だけではなく、温暖化の進行そのものを止める緩和抑制策の立案にも、気候モデルによる予測は貢献している。こうした場合、二酸化炭素の「大気中濃度」ではなく、人間活動による排出量を直接入力データとして与えられるよう、気候モデルに生態系の働きなどを組み入れた地球システムモデル (ESM) が用いられることが多い。ESM を用いた研究を通じた緩和抑制策への貢献として代表的な成果の一つに、人間活動で排出された CO₂ の累積量と、その時点までの昇温とのよい比例関係を示したことがあげられる [31]。このときの比例定数は「排出に対する過渡気候応答」(TCRE) と呼ばれる。国際交渉でしばしば言及される 2°C 目標や、2015 年に締結されたパリ協定で強調された 1.5°C 目標の達成のために、今後の排出量をどの程度にとどめておくべきか、その上限を決定する因子の一つである。TCRE 等に基づいて評価された今後の二酸化炭素排出許容量はカーボンバジエット (CB) と呼ばれる。また特に、CB の総量から現在までの排出量を差し引いた値を残余 (remaining) カーボンバジエット (RCB) と呼び、今後の排出削減を検討するうえで決定的に重要な量となる。

上述の AR6 では、1.5°C 目標達成のための RCB は、産業革命以前からの累積値で 500GtCO₂ と見積もられている。これは現在の年間の人為起源二酸化炭素排出量約 40GtCO₂ の 10 数年分ほどの量にあたる。世界各国が協調して取り組み、直線的に排出を削減して 2050 年までに排出ゼロを達成すれば、それまでの排出量は 1.5°C に対応する RCB とほぼ等しくなる。このことが、日本を含むいくつかの国が宣言している 2050 年カーボンニュートラルの背景になっていると見られ、ESM が社会的に有用な情報を提供しうる一つの例証と言えよう。他に ESM による研究から社会に発信できる可能性のある情報として、様々な社会経済シナリオに付随する土地利用変化の気候影響評価、温室効果気体の年々変動の収支の精確な把握、などが考えられる。

学術的・技術的課題、目標

・ 気候モデル・ESM による 10 年規模予測

近年、季節予報の手法をより長い時間スケールに応用し 10 年規模の予測情報を社会に発信するための研究が盛んにおこなわれている [32]。気象予報用のモデルのみならず、気候変動予測用のモデルをこうした時間スケールへ適用する試みも活発になってきている [33]。さらに、同様の手法を ESM にも応用した、炭素循環 10 年規模予測の研究も各国研究機関が取り組み始めた。炭素循環 10 年規模予測システ

ムは、今後進むであろう CO₂ 排出削減努力の可視化などに有効となることが期待できる。炭素以外の物質循環にも同様の手法を拡張することで、種々の温室効果気体の年々変動の収支の精確な把握にも寄与することが可能であろう。

- 適応データセット高解像度化

文部科学省、気象庁、環境省などが協力して、適応策立案のための予測データセットを 2022 年度までに整備し、以降、概ね 5 年毎に更新するという方針が気候変動適応法に基づき打ち立てられている。2022 年公表予定の「データセット 2022」^{*17}では、既出の d4PDF を基盤としてデータ配信システムなどの整備が図られる見込みであるが、その後の更新ではさらなる解像度の向上や、まれな極端事象の発生も捉えるためのアンサンブル数の増強、現状では緒に就いたばかりの海洋の予測データの拡充などが望まれる。

- 複合問題としての地球環境問題

ESM では、温室効果気体の排出による気候変化のみならず、二酸化炭素の海洋吸収による海洋酸性化、農作物や生物燃料栽培、あるいは都市化に伴う土地利用変化の環境影響など、様々な問題を包括的に取り扱うことができる。一方で、こうした問題は伝統的な気候科学の範疇では取り扱えず、農学や社会経済など多様な分野の研究者との協働が必要とされる。さらに、社会問題としての地球環境問題に対処するためには一般社会における利害関係者の意見も聞きながら問題設定を行う必要があろう。専門性を高め取り扱う問題を厳密に定義することで知見を集積し発展してきた従来型の科学とは異なるアプローチが必要になる。こうした状況に対し、社会経済分野で発達してきた統合評価モデル（将来想定シナリオの策定などに用いられる）と、ESM とを結合した地球システム－社会経済結合モデルを開発することで、有用なツールを提供できる。完成したツールを通じ、多様な分野、セクターを巻き込んだコミュニケーションを促進して、連携を深めることが可能になろう。

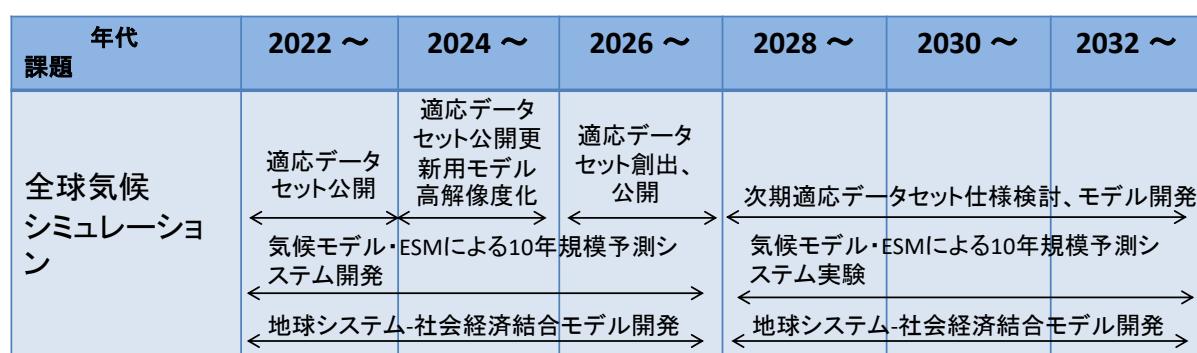


図 4.31 気象・気候分野、全球気候シミュレーションのロードマップ

目標達成に向けて必要な取り組み

- 気候モデル・ESM 精度向上

気候モデル・ESM 開発の今後の大きな方向性として、高解像度化、アンサンブル数増大、多くのプロ

^{*17} 既に公開済: <https://diasjp.net/ds2022/> (2023 年 1 月 28 日時点)

セスを取り込む複雑化の3つが挙げられる。高解像度化は、細かな空間スケールでの大気の対流活動や地形の変化が重要な役割を果たす降雨過程などについて、モデルの再現性を向上させるために必須である。全球の気候変動予測に用いられるモデルは、計算期間の長さと空間規模の大きさがネックとなり水平解像度は100km前後のものが多い。今後、理想的には、大気側では、日本周辺の気象もある程度詳しく再現できる20km、海洋側では中規模渦の再現が可能な10km程度の解像度で予測を行うことが望ましい。また、100年や1,000年に一度と言った、非常にまれな集中豪雨、熱波、渇水の発生確率の変化を検出するためには、従来高々100程度が普通であったアンサンブル数を1桁は増やす必要がある。

さらに、対流圏オゾンや各種エアロゾルなど、大気中の滞留時間は短いものの放射収支に有意な影響を与える短寿命気候強制因子(SLCF)の取り扱いには、大気中の化学反応などについて、これまでモデルに含まれていなかったプロセスも導入する必要がある(4.10.2.2節参照)。これらの方向性はいずれも計算負荷の増大を伴うことになるため、今後大型計算機の性能が順調に向上することがこの分野の発展にとって決定的に重要である。

- ・計算機の発展とモデルの高度化に伴うプログラムコードの複雑化への対応

GPU等の高速化技術を活用するためには、ESMのプログラムコードを書き換える必要がある。また、上述の通りモデルの高度化のため多くのプロセスが加わる傾向にあり、これらが相まってプログラムコードが著しく複雑化してきている。バージョン管理ツールなどのソフトウェア導入で対応を図ってはいるが、コードに初めて触れる若手研究者などに対する障壁となる可能性がある。プログラム構造を容易に把握できるツールの開発など、モデル開発環境の一層の整備が望まれる。

アプリケーションと計算量・計算資源

アプリケーション

- ・モデル名：MIROC (Model for Interdisciplinary Research on Climate) [34]
- ・力学コア：スペクトル法
- ・格子点数：大気 1.7×10^7 (水平解像度約60km, 鉛直層数81), 海洋 1.2×10^8 (水平解像度約30km, 鉛直層数63)
- ・演算量：800 PFLOP (1年積分あたり)
- ・メモリ使用量：1,000 GB
- ・出力データ量：400 GB (1年積分あたり)
- ・要求B/F：2.5 (大気), 3.5 (海洋)

実験規模（一例）

- ・積分年数：250年 (ヒストリカル+21世紀予測；1850～2100年)
- ・アンサンブル数：100程度
- ・想定実行時間：3カ月
- ・総演算量： $20\text{ZFLOP} = 800\text{PFLOP} \times 250 \times 100$

- 総ファイル出力量 : $10\text{PB} = 400\text{GB} \times 250 \times 100$

計算資源

- 総演算量 : システム全体で約 1.7 EFLOPS (システムの全ノードの 5% を使用し (数カ月の実験実施中にジョブの待ち時間があまり発生しないようにするため)、理論ピーク性能の 3% の実行性能を仮定。)

参考文献: 研究開発戦略センター (2019) 研究開発の俯瞰報告書 環境・エネルギー分野^{*18}

(気候モデリング分野でのトレンドや将来的な方向性について概説されており、本文中で直接の引用はないが、本稿執筆者が作成に協力している。本稿の執筆にあたり参照した。)

4.10.1.3 領域気候予測

背景

毎年、世界各地で台風や大雨、高温などによる気象災害が発生している。日本においても、ここ数年は、大雨や台風、高温などの気象災害が毎年どこかの地域で発生しており、災害への備えは喫緊の課題となっている。これらの気象イベントの一部の発生には、イベント・アトリビューションによって地球温暖化の影響が無視できないことが報告されている [30]。地球温暖化に伴う気候変動の地域気候への影響は、将来に懸念される事象ではなく、すでに実際の影響として顕在化しつつある。今後も続くと予測される気候変動に対し、原因となる温室効果ガス排出量削減といった地球温暖化の緩和への取り組みとともに、気候変動への適応も進めていく必要がある。2020 年には気候変動適応法が施行され、自治体ごとに気候変動への適応を推進するための拠点（地域気候変動適応センター、LCCAC）の設置が努力目標となった。LCCAC は、2023 年 1 月時点で 54 の自治体に設置済みである^{*19}。

領域気候に関する研究は、従来、地域特有の気象・気候現象の理解を目的に、観測、理論、数値計算を用いて行う研究の総称と言えるが、ここ 10 年ほどは、将来の領域気候予測のための研究に人的リソースと計算リソースの多くを投入している。これは、上記でも述べた通り、気候変化への地域スケールでの影響評価や適応策の検討への需要が高まっているためである。これまでの 10 年は、気候学者や影響評価研究者の間で、地域気候の将来予測とそれに伴う農業等各分野への影響評価研究がメインであったが、今後 10 年は、よりステークホルダーとの連携を深め、具体的な施策への貢献が求められるようになると推測される。より一層、気候研究者、影響評価研究者、自治体職員、民間企業、市民が一体となって、気候変動対策に取り組んでいく必要があろう。さらには、得られた将来予測や影響評価をどのように施策に反映させるかなど、社会学系の研究者とも連携していくことも重要である。

学術的・技術的課題、目標

地域の空間詳細な気候情報の見積もりには、力学的ダウンスケール手法が用いられる。全球気候モデルで計算された空間解像度の粗い気候情報を、高解像度の領域モデルを用いて計算することで、空間解像度

*18 <https://www.jst.go.jp/crds/report/CRDS-FY2018-FR-01.html> (2022 年 5 月 4 日閲覧)

*19 A-PLAT; <https://adaptation-platform.nies.go.jp/local/lccac/list.html>

を高める手法である。ただし、単に粗い解像度のデータを細かくするだけでなく、空間詳細な地形や土地利用をモデル内で表現することで、全球モデルでは表現できなかった地形性の降水や詳細な気温分布などの付加価値（Added value）が加わることも力学的ダウンスケールを行う一つの利点である。力学的ダウンスケールの他に、粗い気候情報と観測データとの関係から構築した統計モデルを用いて、空間解像度の高い気象データを見積もる統計的ダウンスケールという手法もある。統計的ダウンスケールは、力学的ダウンスケールに比べて、計算負荷がかからないことが利点であり、大型計算機を使用することは稀であるため、以下では、力学的ダウンスケールのみを取り上げる。

力学的ダウンスケールによる気候変動適応策のための気候予測と、予測に基づく影響評価を行う上で課題となるのが、(1) 気候予測情報の時空間解像度の向上と (2) 予測結果に含まれる不確実性への対応である。まず、時空間解像度のうち、地域の気候はその地域が位置する気候帯だけでなく、海からの距離や標高といった地形や土地利用の影響も受けるため、それらの地域特性をある程度再現可能な空間解像度が必要である。必要な空間解像度は評価対象によって異なるが、数 km 程度あれば、大雨や豪雨などの降水イベントや地域の気温分布、局地循環などある程度精度良く再現できると言われている。一方、気候データを農業や生態系などの影響評価に利用する場合には、さらに細かな数 100m 以下の空間解像度、街区内の熱環境評価となると、さらに 1 オーダー小さい 100m 以下の解像度が求められることがある。しかしながら、数 100m 以下の空間解像度の計算においては、計算リソース以外にも、課題は多い。気象気候モデルの高解像度化のための学術的課題については、他の節に詳しい記載があるため (4.10.1.1 節等) ここでは詳しく述べないが、一般的には高解像度化により、地形を含む地表面状態や乱流、積雲対流の表現の向上などが期待される。一方で、それらの表現向上による、領域気候のパフォーマンスへの影響については、十分な知見が得られていない。また、現在使われている数値スキームが、どの解像度まで適用可能であるかについても、検討すべき課題が残っている。次の 5~10 年でターゲットとなる O(100m) の空間スケールにおけるモデルパフォーマンスへの影響とスキームの検討については、早急に進めていく必要がある。一方で、影響評価や適応策検討をする際に、本当に現在よりも 1 オーダー 2 オーダー小さい解像度の気候情報が必要かどうか、という観点での検証も同時に必要である。時間解像度については、一般的に、必要とされる時間解像度よりも、モデルの計算間隔の方が十分に短いため、出力間隔の調整が可能であり、あまり問題になることはない。ただし、影響評価者や適応策を検討する自治体研究者は、大規模データの取り扱いの経験が少ないと多く、アスキーデータでの受け渡しを望むことが多い。ユーザー側が大規模データの取り扱いに慣れても、必要な地域のデータが入手可能なシステムの導入などが必要と考えられる。

力学的ダウンスケールは、温室効果ガス排出量や土地利用などの社会経済シナリオ、境界値データを提供する全球気候モデルのモデルエラー、そして、力学的ダウンスケールに使用する領域気候モデルのモデルエラーなどの不確実性の影響を受ける [35, 36]。不確実性は、予測を行う過程の中でカスケードし、一般に、予測過程の下流になるほど、不確実性は大きくなると考えられる [35]。地域スケールでの影響評価は、予測過程の中でも下流にあたるため、不確実性の影響が避けられない。このため、予測の幅（不確実性）や極値など出現率の低い現象においてはその発生確率を統計的に示すことなどにより、予測結果に対する信頼性情報を提供することも、予測と同様に重要である。予測の幅の見積りには、親モデル（全球気候モデル）や物理スキームなどの計算条件を変えたアンサンブル実験が、極値の予測にはサンプル数を増やす目的でのアンサンブル実験が有効である。

領域気候予測の精度向上のためには、高解像度化とアンサンブル数の増加のどちらも重要であるが、どちらにどれくらい計算リソースを投入するかは、使用可能な計算リソースから逆算して決定しているのが現状である。確率的予測のための大規模領域気候アンサンブル実験として、全球 d4PDF (水平解像度 60km) を力学的ダウンスケールした領域 d4PDF 実験 (水平解像度 20km) が実施されている [28, 29]。現在 (2022 年度) 実施中の、日本全域を対象とした領域 d4PDF からの更なるダウンスケールのうち、解像度が高く最もアンサンブル数が多いものは、水平格子サイズ 5km、アンサンブル数は約 1500 年分である。次の 5 年では、全球結合モデルによる新たな d4PDF 結果を基に、日本全国を水平格子サイズ 5km で 3000 年、2km で 500 年積分、及び、1km 格子サイズでの気候計算 (数十年程度) が目標である。気候モデルの格子サイズが 1km 程度になると、モデルで表現される地表面状態や地形の表現が向上し、気温や地形性降水、積雪の再現性向上などが期待される。また、さらに 5 年先の目標である $O(100\text{m})$ の格子サイズで LES 計算が可能となれば、乱流や雲の表現の改善によりモデルパフォーマンスの向上が期待されるが、この点については、次の 5 年の間に理解を深めておく必要がある。

将来の領域気候情報は、河川モデルや農業モデルなどの入力値として使用され、降水による災害リスクや農作物の生育など、各方面への影響の見積りが行われる [27, 37, 38]。これまでに、農業分野や水文分野などで、多数の影響評価研究が実施されており、今後は、他分野にもさらに広がると推測される。例えば、夏季の熱中症対策として、高解像度数値流体力学 (computational fluid dynamics; CFD) モデルや建物解像大渦解像 (LES) モデルを用いた都市域の熱環境評価 [39] は、研究段階から実用段階へとシフトしつつある。これらのモデルは主に、建築分野などの工学分野で発展してきたモデルであり、気象モデルと建物解像モデルの融合も、今後、研究が進展していくことが期待される。

年代 課題	2022 ~	2024 ~	2026 ~	2028 ~	2030 ~	2032 ~
領域気候予測	$O(10^3\text{m})$ でアンサンブル気候計算 $O(10^2\text{m})$ で現在気候再現実験			$O(10^2\text{m})$ で気候実験		$O(10^2\text{m})$ でアンサンブル気候計算

図 4.32 気象・気候分野、領域気候予測のロードマップ

必要な計算手法

高解像度シミュレーションについては、気象計算のように短時間の積分期間でよいものについては計算機の規模に合わせて実現可能な最高解像度を選択 (ケイパビティ・コンピューティング) すればよいが、気候計算では統計情報を得るために長期積分 (多くの場合、30 年程度が基準。アンサンブル計算ではそれを複数本実施する) が必要であるため、現実的な実時間内に計算を実行するためには、実現可能な最高解像度よりも解像度を落としたり、積分期間を複数個に分けて並行して実施 (キャパシティー・コンピューティング) したりする必要がある。実際には、その両方を適用することが多い。また、不確実性の問題に対して有効な気候アンサンブルを実施する際にも、キャパシティー・コンピューティング的な使用がメインとなる。

一般的に、計算量は格子サイズ比の 3~4 乗で増加する。よって、現在行なっている実験の格子サイズ

を $1/3$ にすると計算量は 27~81 倍、格子サイズを $1/10$ にすると計算量は 1,000 倍以上となる。増えた計算量は、計算機性能の向上と並列数の増加、アプリ側の高速化でカバーする必要がある。

領域気候モデルにおいては、ノード間通信は隣接通信が主であり、一般に week scale 性能は高い。従って、高解像度化に伴う計算量増加の対策として、使用ノード数を増やすことが有効である。そのためには、現状のほとんどの領域気候モデルで採用されている水平 2 次元のプロセス空間分割を鉛直方向も含めた 3 次元分割に拡張するなどが考えられる。また、現状における力学過程計算部分のボトルネックはメモリバンド幅であるが、将来の計算機の B/F 比はより小さくなることが予想されるため、計算機の効率的利用のためには B/F 比を下げるアルゴリズムの検討が課題である。

アプリケーションと計算量・計算資源

アプリケーション

- モデル名：SCALE(Scalable Computing for Advanced Library and Environment)[40, 41]
- 力学コア：3 次元完全圧縮非静力学方程式系、有限体積法

計算量・計算資源

- 問題規模：日本全域、1km
 - 格子点数： 2750×3775
 - 鉛直層数：60 層
 - 積分時間：60 年（現在、将来 2 度上昇、将来 4 度上昇）
 - アンサンブル数：各 30 メンバー
 - ステップ数：3 億回
- 総メモリ数：2TB
- 総演算量：780ZFLOP
- 実行時間：1 ケース当たり 6 時間
- アプリケーション B/F 比（演算密度の逆数）：0.8
- 総メモリ転送量：610ZB
 - 要求メモリ転送性能：315PB/s
- 総ストレージ使用量：2150PB（20 変数を 1 時間毎に出力）
 - 要求ファイル IO 性能：4TB/s

4.10.1.4 数値天気予報

背景

数値天気予報では、大気（天気）のシミュレーション計算を、現在の大気状態に基づいて未来に向かって行い、その未来の状態を客観的に予測する。毎日の暮らしを支える天気予報や防災気象情報は数値天気予報の予測計算結果をもとに発表されている。これらの情報に基づいた防災・減災行動を求めるためには

予測に対する信頼性を担保する必要があり、そのために不断の努力が為されている。このように国民生活の基盤となる情報であることから、数値天気予報は気象学による知見の最終的な出口と言って良く、実験や観測あるいは理論的考察からシミュレーションモデルへの実装に至る研究は、数値天気予報システムの精度向上を支えている。

最初の数値天気予報は Richardson (1922)[42] の手動計算によって試みられ、大型計算機の開発とともに栄光の道を歩み始めた [43]。ただし Richardson (1922) は初期値を観測データの内挿によって作成したためノイズを含み、かつ方程式系にこのノイズをダンプするスキームが備わっていないために非現実的な予測を算出するに至り、Charney et al. (1950)[43] はバロトロピックモデルを採用することで大気重力波をダンプすることに成功したものの、そもそもノイズを含まない初期値を作成する客観解析の必要性を訴えた。このように数値天気予報の精度はシミュレーションモデルとその初期値を提供する客観解析の二つの精度に依存している。

気象庁は 1959 年に世界で 3 番目に大型計算機を用いた数値天気予報を開始し [44]、同年の伊勢湾台風の予測にも用いられた [45]。現在では気象庁のスーパーコンピュータシステムは第 10 世代数値解析予報システム（理論性能 18.16PFlops）に進化し、全球モデル (GSM)、メソモデル (MSM)、全球アンサンブル予報システム (GEPS)、局地モデル (LFM) 及びメソアンサンブル予報システム (MEPS) が運用されるとともに、水平解像度や鉛直層数の増強、アンサンブルメンバー数増強、物理過程等の計算手法の改良などにより数値天気予報システムの性能を向上させてきた [44]。

また、客観解析手法についても、世界の数値予報センターで広く利用されていた最適内挿法から 3 次元変分法及び 4 次元変分法などの高度なデータ同化手法への移行が進み、さらにアンサンブル情報を用いて背景誤差共分散行列の大気状態への依存性を考慮するハイブリッド手法の導入に至っている。変分法の導入に伴い、衛星による放射観測など予報変数ではない物理量の同化ができるようになり、これによる利用できる観測データの拡大も客観解析の精度向上に貢献してきた。

このように気象庁では、客観解析を含む数値予報システムの機能向上を継続的に実施し、システムを高精度化、複雑化させてきた。このような不断の努力により、現在の数値予報の 3 日予報の精度は 1980 年代の 1 日予報の精度で行えるまでに至っている。

現業機関である気象庁のシステムが高度化していることから、数値天気予報技術の発展のためには、その高性能なシステムを研究者へ提供し、ここへ研究者による最先端の技術を取り込むことでさらなる進化を図る Operation to Research (O2R)、Research to Operation (R2O) と呼ばれるサイクルの必要性が認識されてきている。

R2O の一例として、天気予報のために開発された気象モデルを極めて高い解像度に設定したシミュレーションが、特定事例の詳細解析やパラメタリゼーションの開発といった気象学の基礎研究として利用されている。例えば、甚大な被害をもたらした 2013 年大島豪雨や 2014 年広島豪雨について数 100m の水平解像度のシミュレーションを行い、従来の計算より現実的な極端降水が再現されている [46, 47]。また、台風の内部において突風をもたらす微細構造を明らかにしている [8]。高解像度化により、大気を計算機で扱うために必須な離散化に伴う誤差が低減できるため、極端気象などの再現に関して、より精緻な再現が可能になると期待できる。

さらに O2R、R2O 双方向に利益をもたらしている成功例として、アンサンブルシステムの開発がある。上の気象庁システムに挙げられているハイブリッド同化や MEPS に使用されているアンサンブルシ

システムは研究サイドの開発が現業に用いられた例であり、また、研究サイドで用いられているアンサンブルデータ同化システムには気象庁で開発されている気象モデルや観測データが利用されている。

本節では (A) 気象庁による高解像度の現業数値天気予報システム、及び、(B) 超高解像度シミュレーション研究、(C) アンサンブルデータ同化研究について述べる。

学術的・技術的課題、目標

近年、顕著な気象による自然災害が激甚化しており、今後更に災害が深刻化するおそれがあることから、気象災害防止と軽減のため、数値予報精度向上はさらに強く求められている。このうち豪雨防災では「線状降水帯の発生・停滞の予測精度向上により、集中豪雨の可能性を高い確度で予測する。特に、明るいうちからの避難など、早期の警戒と避難を可能にする」との目標が掲げられている[48]。

(A) 気象庁数値天気予報システム

気象庁(2018)[48]によれば、最も解像度の高い LFM でも線状降水帯を半日程度前から時間と場所を絞って予測することは依然困難であり予測の不確実性が高いこと、個々の積乱雲を表現するには水平 2km の格子間隔では解像度が不足していること、また高い解像度に適した物理過程が必要なこと、海上の水蒸気量や風についての初期状態の精度が不十分であることなどが課題として述べられている。これらの課題を解決し、上記豪雨防災目標を達成しうる現業数値予報システムを運用することが気象庁数値天気予報システムの目標となる。また、ここでは触れないが LFM に境界値を与える MSM、さらに MSM に境界値を与える GSM の精度向上も重要な課題である。さらに、数値天気予報が現在の大気状態を速やかに解析し、それを初期値として予測計算を行った結果をリアルタイムに活用するため、日々決められた時間までにこれらの計算を確実かつ速やかに行う必要があるという極めて重要な要件があることに留意が必要である。

(B) 超高解像度シミュレーション

超高解像度シミュレーションで用いる力学・物理過程等の多くの部分は、天気予報のために開発された気象モデルを大幅に変更することなく活用できる。ただし一部の過程は解像度に応じた変更が必要である。様々な過程のうち、より単純化する方向の変更が必要な過程（対流・乱流・力学など）とより複雑する方向の過程（放射・接地境界層など）の両者があり、それぞれの過程の適応策の研究開発が必要である。この研究開発自体のためにより高解像度のシミュレーションが必要になる場合がある。

(C) アンサンブルデータ同化システム

アンサンブルデータ同化システムは、気象庁が現業運用している 4 次元変分法（4D-Var）データ同化システムと比べて開発が容易であるために、世界的にも研究コミュニティで広く用いられている。また、現業機関においても変分法データ同化システムとのハイブリッドシステムが普及しており、天気予報に欠かせない技術となってきた。

アンサンブル同化システムにおける本質的な課題はサンプリングエラーと誤差共分散行列におけるランク不足の克服である。前者については概ね 1,000 メンバーのアンサンブルによって実用上は問題ないことが過去研究で示されている。例えば Duc et al. (2021)[49] は 1,000 メンバーによる線状降水帯アンサン

ブル予測を実行し、アンサンブル平均解析場を初期値とする決定論予測精度、確率予測とともに非常に良い結果が得られた。後者については、アンサンブル変分法(EnVar)ではハイブリッドで対処可能であるが、アンサンブルカルマンフィルタ(EnKF)では有効な手法についてまだ研究段階にある。

サンプリングエラーに対して実用上、1,000 メンバーでは問題ないとしても、完全に無視できるレベルではなく、水平局所化と呼ばれる技術によって対応が必要である。また、ハイブリッド手法を用いずにランク不足にならないアンサンブルサイズは同化する観測データ数と近いオーダーが必要である。従って、現業メソ同化システムと同等の性能をアンサンブルデータ同化システムが持つためには、およそ 10 万メンバーが必要と推定され、これがポスト富岳において目指したいサイズとなっている。

また、研究コミュニティで広く使われている EnKF では、現業変分法データ同化システムで使われる変分 BC(バイアス補正)、変分 QC(品質管理)、ノイズ抑制項などが自然に実装できず、まだまだ開発の余地がある。一方、EnVar においては摂動生成手法に工夫の余地が多く、研究段階にある。

今後 10 年程度の目標としては、現業で同化可能な観測データはすべて同化し、かつ現業より優れた精度を持つアンサンブル同化システムの開発である。同時に積乱雲や線状降水帯といった極端現象の予測には高解像度のシステムが必要であり、ここでは従来のデータ同化理論の基礎であるガウス仮定を超えて非ガウス性を導入していく必要がある。

年代 課題		2022 ~	2024 ~	2026 ~	2028 ~	2030 ~	2032 ~
気象予報	(A) 気象庁システム	1km狭領域 LFM試験運用ハイブリッドデータ同化導入	LFM 予報時間延長	LFM 高解像度化 LEPS 運用開始		LFM改良 LEPS改良 アンサンブルデータ同化導入	
	(B) 超高解像度シミュレーション	水平解像度<100m級LESの利用			水平解像度<50m級LESの利用		
	(C) アンサンブルデータ同化	asuca-based EnVar, EnKFの開発(水平15km)		全現業データ同化(5km) 高解像度化(~1km)		非ガウス性10万メンバー同化	

図 4.33 気象・気候分野、数値天気予報のロードマップ

目標達成に向けて必要な取り組み

(A) 気象庁数値天気予報システム

気象庁(2018)[48]によれば、LFMを1km以下に高解像度化すること、積乱雲に関する諸過程を改良すること、集中豪雨の予測が持つ不確実性を適切に補足可能なアンサンブル予報システムを開発すること、衛星、航空機、レーダー、船舶 GNSS、IoT 化した多様なセンサなど最新の様々な観測による時間的、空間的に高密度な観測ビッグデータを更に活用すること、衛星データについて全天候利用手法を開発すること、複数の手法を組み合わせた最先端の同化技術であるハイブリッド同化を導入すること、豪雨事例等のメカニズム研究による最新の科学的知見に基づく改良をすることなどが開発の方向性として述べられて

いる。また、このような開発計画を実施していくためには、十分な能力を持つスーパーコンピュータや通信インフラを、継続的かつ計画的に確保していくことは必須であり、さらにそのインフラを有効に活用するため、プログラムの高速化（最適化）を実施することが必要とされている。

(B) 超高解像度シミュレーション

一般的にシミュレーションの計算精度は、解像度に従い向上することが期待される。しかし科学的に意義がある結果を得るために、計算コストが高くなってしまうため、高解像度化は必ずしも最善とは限らない。注目する現象の解像度依存性を把握しながら、適切なターゲットに対し超高解像度シミュレーションを戦略的に実行する必要がある。超高解像度の計算時に適したパラメタリゼーション等を確立させ、シミュレーションに実装する必要がある。また技術的な側面として、前処理やモデル本体も含めて、莫大な計算格子数による実行に様々な問題が生じうる。計算の実行および実行時の問題解決の円滑化も望まれる。

(C) アンサンブルデータ同化システム

現業システムと研究用システムの親和性を高めるために、気象庁メソモデル用の EnVar、EnKF の開発が必要である。新開発のシステムを用いて、衛星輝度温度、レーダー反射強度など全現業データ同化など、現行の研究用同化システムである NHM-LETKF で同化できない観測データ同化に取り組む。また、同時に 1,000 を超える大アンサンブル計算を行いつつ、高解像度化を図り、最終的に 10 万メンバー同化を目指す。さらに非ガウス性導入を行い、強非線形な積乱雲や線状降水帯といった顕著現象の予測を実現する。

アプリケーションと計算量・計算資源

(A) 気象庁数値天気予報システム

ここでは、気象庁の数値天気予報で最も高解像度で予測を行う局地数値予報システムの 2022 年 7 月現在の仕様について述べる。局地数値予報システムは、初期値を作成する局地解析 LA と局地モデル LFM により構成されている。LA では、MSM の予測値を第一推定値として対象時刻の前 3 時間を 1 時間に、様々な観測データの同化と予測計算を繰り返す手法を用いて対象時刻の初期値解析を行っている。観測データの同化は asuca-Var を用い、MEPS により出力される予報値を元に計算した流れ依存背景誤差を利用するハイブリッド 3 次元変分法により行っている [50]。なお LA におけるデータ同化は 5km の水平解像度で実施している。LFM では水平解像度 2km、鉛直 76 層の非静力学領域モデル asuca により 10 時間予報を実施している。この LA と LFM による計算を 1 日 24 回（毎正時）実施し、その結果は防災気象情報や航空気象情報、降水短時間予報等の作成支援に役立てられている。asuca 及び asuca-Var の詳細については気象庁予報部 (2014)[51] を参考願いたい。なお、LFM の側面境界値は解像度 5km の MSM により与えられており、さらに MSM の側面境界値は解像度 20km の GSM により与えられている。これらの親モデルの精度向上も LFM の精度向上に大きく影響する。

将来的には LFM の予報時間延長と、高解像度化及び LFM をベースとしたアンサンブル予報システムである局地アンサンブル予報システム LEPS の運用を実施する予定である。LEPS のメンバー数につい

てはリアルタイム運用可能な計算資源の制約により決定される。また、LA については、データ同化手法自体にアンサンブル手法(EDA)を導入することも今後の検討課題である。将来的には入力する観測データ量についても、静止気象衛星の赤外サウンダのデータが入ることなどで大幅に増加すると考えられる。

○ 現行 LFM の仕様

- 水平解像度 2km、水平 1581×1301 格子、鉛直 76 層 ($\sim 21.8\text{km}$)、積分時間間隔 12 秒
- 約 1,000 秒で 10 時間予測を実施（現行気象庁スパコンシステム（CrayXC50）138 ノード）
 - ※ 1 時間（300 ステップ当たり）約 100 秒（1 ステップ当たり 0.3~0.4 秒）
XC50 の演算性能はノード当たり約 3.2TFlops
- 入力データ量：初期値 0.6GB + 境界値 0.6GB × 予報時間
- 1 時間当たりの出力データ量：約 1.2GB (12GB/10 時間予測)

● 将来 LFM の仕様想定

- 水平解像度 1km、水平 3161×2601 格子、鉛直 76 層 ($\sim 21.8\text{km}$)、積分時間間隔 6 秒
- 約 1,000 秒で 18 時間予測を実施
- 入力データ量※：初期値 0.6GB + 境界値 0.6GB × 予報時間
 - ※初期値解析の解像度や親モデルの解像度が現状維持だった場合
- 1 時間当たりの出力データ量：5GB (90GB/18 時間予測)

● 将来 LEPS の仕様想定

LFM (1km) をメンバー数 100 で実施すると想定した場合。

なお個々のメンバーの仕様は上記将来の LFM の仕様想定に準拠する。

出力データ量はトータルで 9PB/18 時間予測 (90GB/18 時間予報 × 100) と推定できる。

※実際にはリアルタイム運用できる計算機資源の規模により解像度及びメンバー数の想定は変わってくる。

○ 現行 LA (ハイブリッド LA) における asuca-Var の仕様

- 水平解像度（インナーモデル）5km、水平 633×521 格子、鉛直 48 層
 - 約 60 秒でデータ同化を実施（現行気象庁スパコンシステム（CrayXC50）120 ノード）
 - 入力データ量 350GB
 - 出力データ量 0.6GB
- ※出力データ量はアウターモデル (LFM) 解像度準拠

● 将来の LA における asuca-Var の仕様想定 (50 メンバー EDA)

- 水平解像度（インナーモデル）5km、水平 633×521 格子、鉛直 48 層、50 メンバー
- 約 60 秒でデータ同化を実施
- 入力データ量 370GB

※衛星観測データ量については、現行で最大 $50\text{MB} \times 4$ 時間 = 200MB 程度である。

100 倍程度と見積もることで 20GB と想定する。

- 出力データ量 120GB

※出力データ量はアウターモデル (LFM) 解像度準拠、

$$0.6\text{GB} \times 4 \text{ (アウターモデル水平解像度 2 倍 (1km 化))} = 2.4\text{GB} \times 50 \text{ メンバー}$$

(B) 超高解像度シミュレーション

高解像度のシミュレーションのために高い解像度を設定しながら、広い計算領域をカバーするためには膨大な格子数が必要になる。その一方、計算時間に関する制限は厳しくない。実現できる計算規模は、直接的にはメモリやストレージによって制約される。

気象シミュレーションでは鉛直方向の成層を表現するため、従来、鉛直方向には比較的高い解像度を設定されている。水平解像度が 100m 程度に達するまでは鉛直解像度は大幅に増加させなくても合理的であった。しかし、水平解像度が 100 m を切ると鉛直解像度も水平解像度に比例させ向上させる必要がある。これは将来 LES の仕様を想定する場合は考慮しなければならない。

○ 現行 LES の仕様

- 水平解像度 100m、水平 $20,000 \times 20,000$ 格子、鉛直 60 層 ($\sim 23\text{km}$)、積分時間間隔 0.8 秒、計算時間 20 時間で 4 時間分の計算が進行 (富岳 2,000 ノード)
- メモリ使用量 : 60TB
- 入力データ量 : 初期値 1TB (リスタートの場合 17TB)
- 1 計算あたりの出力データ量 : 1TB (1 時間毎の計算結果) + 17TB (リスタート用)

● 将来 LES の仕様想定

- 水平解像度 33m、水平 $50,000 \times 50,000$ 格子、鉛直 200 層 ($\sim 21.8\text{km}$)、積分時間間隔 0.1 秒
- メモリ使用量 : 1.5PB
- 入力データ量 : 初期値 18TB (リスタートの場合 300TB)
- 1 計算あたりの出力データ量 : 27TB (1 時間毎の計算結果) + 300TB (リスタート用)

(C) アンサンブルデータ同化

大アンサンブルデータ同化では、現在、数年前まで気象庁で現業モデルとして使用されていた NHM と組み合わせて、EnKF の一種である LETKF を富岳で運用している。アンサンブルサイズは 1,000、同化システムの水平解像度は 15km で、5km 解像度の NHM をアンサンブルシミュレーションに用いている。アンサンブルサイズは 1,000 で、計算領域は気象庁メソ解析と同じ、同化ウィンドウは 3 時間である。

このスペックで 4 日間にわたって計算した場合の計算資源を図 4.34 に示した。今後、高解像度化によって 1,000 倍程度、メンバー数增加によって 100 倍程度の計算資源増加が見込まれている。グランドチャレンジによって試みたい。

4.10.2 気象・気候モデルにおける物理コンポーネントの精緻化

4.10.2.1 大気モデルの物理プロセス（雲・放射・雷コンポーネント）

背景

近年、計算機の性能の向上によって、気象・気候モデルは高度化・精緻化が劇的に進んでいる。具体的には、気象・気候モデルは 2 つの方向での高度化・精緻化を続けている。1 つ目は対象とする領域を覆うメッシュが細かくなる「高解像度化」、2 つ目は「物理プロセスの精緻化」であり、従来の気象・気候モ

	モデル名（略称）	モデル名	文献	計算手法	問題規模	総演算量	要求演算性能	想定ノード数	メモリ使用量
(A) 将来LFM	asuca	Asuca is a System based on a Unified Concept for Atmosphere	気象庁予報部 (2014)	有限体積法完全圧縮非静力学モデル	水平解像度1km、水平3161×2601格子、鉛直76層、積分時間間隔6秒、18時間予報、毎時	6.4EFlop /Member 毎日毎時実行	6.4PFlops /Member	2070ノード /Member	
(B) 将来LES	asuca-LES	Large eddy simulation by asuca	Ito et al. (2017)	有限体積法完全圧縮非静力学モデル	水平解像度30m、水平50000×50000格子、鉛直200層、積分時間間隔0.1秒、ステップ数500万	2.3ZFlops			1.5PB
(C) アンサンブルデータ同化	NHM-LETKF	Nohydrostatic Model initialized by Local Ensemble Transform Kalman Filter	Duc et al. (2021)	NHM: 完全圧縮非静力学モデル、雲物理過程、Mellar-Yamada Level2.5 LETKF: 局所アンサンブルカルマンフィルタ	NHM: 格子サイズ 5km、格子数 817x661x50、ステップ数 5760、アンサンブル数 1000、実行期間 4日 LETKF: 格子サイズ 15km、格子数 273x221x40、水平局所化半径 700km、鉛直局所化なし、5日間で20回	NHM: 172800 EFLOP LETKF: 388800 EFLOP		NHM: 10000 ノード LETKF: 1000 ノード	NHM: 320TB/cycle LETKF: 32TB/cycle

図 4.34 各実験の計算設定と計算量

ルでは計算資源の制約から考慮に入れることができなかった物理過程を取り入れた計算がその代表的な例である。精緻化に関しては、高解像度化などが進むにつれて、従来の手法の適応限界を超えるため、新たに物理モデルを開発するといった精緻化も含まれる。なお、気象・気候モデルの物理プロセスとは、図 2.19 に示される気象・気候モデルで考慮されるプロセスのうち、「大気の流れ」と表記されたプロセス以外のプロセスのことを指す。

このような高度化は、IPCC 第 6 次評価報告書 [1] をはじめとして報告されている「気候予測の不確実性の低減」、日々の天気予報をはじめとする「気象予測の予測精度の向上」、「気象予測においてこれまでに提供されてこなかった新たな付加価値の提供」といった社会的な要請に応える形で進められてきた背景がある。また前回のロードマップが作成された 2010 年代前半以降は、雲一つ一つを表現した計算が可能になったことで、気象予測の精度が向上するとともに、これらの高解像度な計算は衛星観測アルゴリズムの性能向上に大きく貢献してきた。

しかしながら、高度化が進んでもなお、雲の中の微物理特性に関する数値モデルの不確実性（特に氷を含んだ雲粒の微物理特性に関する不確実性）が高いのが現状であり [52]、この部分の高度化が気候予測・気象予測・さらには観測アルゴリズムの高度化のために求められている。また同時に、富岳をはじめとした現在の最先端の計算機によって開発が進み、2030 年ごろに汎用的な利用や、実用化が見込まれるようになってきた技術もある。

モデルの開発・高度化は防災や気候変動への対応の基盤となる技術であり、モデルの高度化なくしては、気象に関連した自然災害や気候変動に対する適応策や緩和策を立てることは困難である。それらの 2030 年頃に向けたモデルの高度化・技術の実用化につなげるために取り組むべき課題について本章で記載する。なお、1 つ目の項目である「高解像度化」については 4.10.1.1 節や 4.10.1.4 節にその説明を譲るとして、本節では 2 つ目の物理プロセスの精緻化と 2031 年前後までに実現が可能と見込まれ、現時点で取り組むべき点について記述する。物理プロセスの高度化という観点では、図 2.19 の全ての物理プロセスに対して高度化が進むことが考えられる。エーロゾルや大気化学に関しては 4.10.2.2 節、海洋モデルに

関しては4.10.2.3節で議論されており、本節では、これら以外で高度化が必要と考えられる物理プロセス（具体的には、雲・大気放射・雷）について議論する。

学術的・技術的課題、目標

気象・気候モデルの高解像度化がさらに進み、従来の手法の適応限界を超えてしまう可能性が高く、気象・気候モデルのコミュニティが解決しなければならない課題を抱えているのは、図2.19の大気境界層と書かれる「乱流過程」と「放射過程」である。また雲・降水に関連して、上記で述べたように「雲の中に含まれる氷の微物理過程の高度化」が必要である。

乱流過程の計算については、気象・気候モデルにおいて広く用いられているレイノルズ平均モデル(RANSモデル)はメッシュが1kmを切るような高解像度の計算に用いることは適切ではない。2030年に最先端の計算機を用いた計算は、ラージ・エディ・シミュレーション(LES)モデルでの計算が一般的になると考えられ、LESモデルを用いた計算を積み重ねる必要がある。これについては4.10.1.1節に説明がされている。

放射過程に関しては、気象・気候モデルにおいて広く用いられている放射伝達モデルは、平行平面大気を仮定して、光の伝達方向を上下方向のみに近似する2流近似を用いているものが多いが、これらの仮定はモデルの水平格子幅が十分に大きい場合には良い仮定であったが、LESモデルで計算をすることが適切な高解像度(メッシュの幅が数10m～数100m)の計算においては、斜めに光が入る効果を考慮せねばならず、仮定が崩れる。この問題に対応するために、3次元的に放射伝達を計算する3次元放射伝達モデルの開発が必要であるが、気象モデルと結合した形で3次元放射伝達モデルの開発を進め、この問題に対応する取り組みをしている国は、世界的にみても欧州の一部の国を除いて存在しない。そのため、この対応を世界に先駆けてできれば、日本が世界をリードできる可能性がある。

また、現状の気象モデルでの雲粒の扱いにおいては、雲の中における氷相の雲粒の微物理特性(雲粒の大きさや形状、密度)を経験的に仮定しているのが現状であり、このことが計算の不確実性を生む原因となっている。この問題に対応するために、氷相の微物理特性を直接計算する雲モデル(Habit予測型ビン法[53]や、超水滴法[18, 54]の開発が進められており、これらを利用することで気象・気候予測だけでなく、衛星観測の観測精度向上にも寄与できることが期待される。さらに、雲に関連した現象であり、2030年実用的に利用されることが期待される技術として、雲が持つ電荷や雷を直接計算し、雷頻度を直接予報する「気象雷モデル[55, 56]」がある。これが実用化されれば、現在は雷警報や雷活動度といった抽象的な物理量による予報しか存在していない雷の予測を雷頻度や電荷の分布の予報に繋げることができるだけでなく、雷による電気設備への被害や航空機への被雷被害低減という点で、社会に対して大きな貢献となる。

以上を踏まえると、2030年までに達成することが見込まれる目標は、「1.LESモデルを用いた計算の汎用化」「2.気象モデルと結合できる3次元放射伝達モデルを開発と、気象モデルと結合させた実験の実施」「3.精緻な雲モデル(氷の形状を直接計算するモデル)を用いた現実事例での数値実験」「4.気象雷モデルによる雷の予測の実用に向けた開発・検証」の4点である。これらの達成目標のロードマップについて表4.35に示す。なお、1については4.10.1.1節、4.10.1.3節、4.10.1.4節に説明がされているため、この章では割愛する。

年代 課題		2022～	2024～	2026～	2028～	2030～	2032～
物理モデル	3次元放射伝達モデル(3D-RD)	3D-RDの開発 気象モデルへの実装		結合モデルを用いた 小スケールでの数値実験		実問題での数値実験	
	精緻な雲モデルの開発		精緻な雲モデルの開発と実事例での実行に向けた高度化			実問題での数値実験	
	気象雷モデルの実用化	ハインドキャスト実験と高速化 モデルの検証と改良		特定の地点を対象とした実証実験		日本全体を対象とした予測 全球モデルへの拡張	

図 4.35 気象・気候分野、大気モデルの物理プロセスのロードマップ

目標達成に向けて必要な取り組み

図 4.35 の目標達成に対して問題となるのは、どの目標を達成するためにも、既存の気象・気候モデルに比べて、予報変数が増えるため、全てのモデルが従来のモデルに対して大きなメモリを必要とする点と、計算速度のボトルネックとなる全通信を必要とする計算が含まれるという点である。また、このような要求されるメモリの増加に対して、近年における最先端の大型計算機のメモリが小さくなり続けており、理論性能値としては高いが、上記のような目標達成の障害となることが想定される。

気象学のコミュニティにおいて、物理モデルを開発するユーザーの多くは計算機の知見をあまり持ち合っていないことが多い、実際に最先端の大型計算機が利用できる環境があったとしても、目標達成のために開発したモデルの性能が出ないという問題に直面する。そういう状態に陥った時に多くの物理モデルの開発者は最先端の大型計算機での利用を諦める傾向がある。

これらを解決するために、具体的には以下の 2 点の取り組みが必要となると考えられる。1 点目は、計算機科学と気象学の共同によるモデル開発といった、既に研究者として活躍している者同士の共同開発機関の設置である。計算機を使いこなすためのノウハウの提供とは“言うは易し行うは難し”という事項の典型例である。気象学者と計算機科学者は、お互いの研究上の興味は一致しないことがほとんどであるため、研究を共同で行うのは難しい。これを解決するには常日頃から気象学者と計算機科学者が顔を合わせる状況を作り、いざ必要な時に意見交換ができる状態を保っておくくらいしか方法は思いつかない。

2 点目は、計算機科学の知見も持ち合わせた気象学の人材育成である。気象学者が必ずしも計算機の知見を持ち合っていないのは、教育カリキュラムにそのような知見を獲得する機会がないためである。気象学者の中で計算機の知見を持ち合っている者は、計算機に興味を持って自分で調べて知見を獲得したものか、研究遂行上必要なため自分で調べて知見を獲得した者に限られる。そのため、多くの気象学者は大型計算機がなければ研究を進められないにもかかわらず、計算機の知見を持ち合っていないというものが現状である。

これら 2 点を同時に解決する方法として、大学などの教育機関において、気象学者と計算機科学者が所属するモデリング研究センターの設置が挙げられる。若手育成は、大学が主導するべきであり、大学にモデリング研究センターが設置され、そこに学生が配属されるような環境を作れば、人材育成は進み、2 点目の取り組みそのものになるであろう。同時に、既に活躍している研究者がそのようなモデリング研究セ

ンターに所属することで、1点目の取り組みにもつながるであろう。

最後に、欧米や中国をはじめとして、世界的にも物理モデルの高度化は日々進んでおり、日本でも開発を続けていかなければ世界に取り残される。その際に、欧米のように大型計算機の利用者として物理モデルの開発者を取り込む制度の設計が必要になるであろう。

アプリケーションと計算量・計算資源

上記の目標のために、精緻な雲モデル（超水滴方法、Habit予測型ビン法）、雷モデルは既に気象ライブラリSCALEに実装が完了しており、3次元放射伝達モデルについてはSCALEへの実装を見据えて開発に着手している。このうちSCALEと結合した気象雷モデル[56]について取り上げる。

水平解像度500mで、計算領域は5,000km四方（日本全域を想定）、鉛直90層で雷の予報計算を行う。その際雷をもたらす日時に絞って計算を行うため年間の2/3程度の200事例を対象とする。必要な計算量は図4.36の通りである。

モデル名 (略称)	モデル名	文献	計算手法	問題規模	総演算量	備考
SCALE-LT	Scalable Computing for Advanced Library and Environment with Lightning (LT) component	Nishizawa et al. (2015), Sato et al. (2015), Sato et al. (2019)	力学過程： 有限体積法、構造格子、非静力学完全圧縮流体方程式 物理過程： 雲微物理、大気放射、乱流、陸面過程、大気電気計算（SCALE本体にはない機能としてBi-CGSTABによる線形ソルバーあり）	格子サイズ：500m 問題サイズ： 10000x100000x90(90億) ステップ数：約50万ステップ (dt=0.5s、72時間) ケース数：200ケース	200000EFLOP (2000EFLOP/ケース)	SCALEのデフォルトでの時間に対して計算時間が2~5倍程度の計算量

図4.36 各実験の計算設定と計算量

4.10.2.2 大気化学

背景

人間活動によるさまざまな化学物質の大気中への放出は、放射過程や雲微物理過程等を介して大気および地表面の物理化学過程に大きな影響を与える。具体的には、二酸化炭素をはじめとする温室効果気体や、光化学オキシダントやPM2.5(微小粒子状物質)などの大気汚染物質は、さまざまな人為的および自然起源放出源から大気中に排出され、太陽からの短波放射や地表からの長波放射を吸収もしくは散乱するとともに、雲の生成・成長過程等に影響を与えることにより地球温暖化等に大きく影響している。また、高濃度の大気汚染物質に長期間曝露されることにより呼吸器系疾病の発病率や植物の枯死、農作物の収穫量減少など、地表動植物の生態系にも広く影響を及ぼすと考えられる。さらに、黄砂をはじめとする土壤粒子中に含まれる鉄分や降水を介した栄養塩供給が海洋生態系にも大きく影響を及ぼしている。

このような大気微量成分の大気中での動態や、気候・気象等との多岐にわたる相互作用について、地球表層を含めた循環メカニズムを詳細にシミュレートすることは、現在の計算機資源を以ってしても十全とは言い難いものの、気候モデルシミュレーションの黎明期からいくつかのキープロセスに着目したシミュレーション実験が行われてきた。二酸化炭素の気候影響については、真鍋らが1次元モデルおよび3次

元モデルを用いて評価し、二酸化炭素の大気中濃度を 2 倍に変化させると地表平均気温が 2.36 度上昇することを初めて示した [57, 58]。この際に用いられた 3 次元モデルが、現在広く使われている地球温暖化予測モデルの原型となっているといえる。また、1980 年代からは成層圏オゾン層破壊、いわゆるオゾンホールとその要因としてのフロンの人為起源排出が広く認識されてきたが、要因となる化学物質の排出量の長期的な規制シナリオとしてモントリオール議定書が 1987 年に採択された。この排出シナリオに基づく成層圏オゾン層の回復傾向を予測する際にも、2 次元および 3 次元の全球化学輸送モデルを用いた長期シミュレーション実験が行われており、2000 年代に入ってからはベクトル計算機を利用した 3 次元モデル計算が日本国内でも行われている [59, 60]。

二酸化炭素をはじめとする気候-化学相互作用の重要性が広く認識されるようになってから後は、地球表層の大気、海洋、陸域といった気候システムを構成するサブシステムとしての大気モデルが地球システムモデルのなかに要素モデルとして組み込まれ、要素モデル間での変数交換を行なながら大気中濃度等の時間発展をシミュレートすることが広く行われている。IPCC 第 6 次報告書の科学的知見の基盤となる CMIP6 の中では AerChemMIP がそれら化学輸送モデルの相互比較プロジェクトであるが、世界的には 10 のモデルグループ、うち日本からは 2 つのモデルグループが参加している。

学術的・技術的課題、目標

大気微量成分の大気中での動態およびその気候的影響を正しく評価するためには、大気中で物質がどのように運ばれ消えていくのか、すなわち 1) 地表・海洋表面上などから大気への排出および植物等による吸収、2) 輸送過程、3) 変性過程、4) 沈着過程の諸過程と、5) 大気中における放射過程となるべく精緻に評価する必要がある。そのため、この諸過程において考えられる課題について概観したい。

1) 排出・吸収過程

地球温暖化に最も影響を及ぼす大気成分である二酸化炭素は、人間活動における化石燃料消費や土地利用変化に伴う排出と、海洋や陸域生態系による吸収が重要な変動源であり、これらのプロセス理解の深化と交換量の精緻化が重要な課題である。大気モデルや人工衛星を用いたトップダウン推定のほかに陸域モデルや統計データを用いたボトムアップ推定がある。これまでこれら推定手法間の差異が大きかったものの、近年では観測データの拡充やモデル高度化に伴い、比較解析が進みつつある。たとえばグローバルカーボンプロジェクトでは、毎年大気モデルを用いたトップダウン推定を含めた二酸化炭素排出量の推定結果を取りまとめており、2020 年には COVID-19 によるロックダウン等の影響で 5.4% 減少したと推定されている [61]。近年は人工衛星のセンサー性能の向上に伴い、衛星データの高解像度化が進みつつある。たとえば ESA(欧州宇宙機関) が 2017 年に打ち上げた Sentinel-5P に搭載された TROPOMI は $7 \times 7\text{km}$ の解像度であり、JAXA(宇宙航空研究開発機構) が 2023 年に打ち上げ予定の GOSAT-GW は $3 \times 3\text{km}$ の解像度を想定している。これらの高解像度衛星に近い空間分解能で、かつ COVID-19 や福島第一原発事故のようなイベント的な変動に対しても即応性の高い対応が行えるようにすることなどが当面の課題として挙げられる。

エアロゾルに関しては気候変動に関連した地表生態系の応答に関する理解がまだ十分でないと考えられるため、これらのプロセス理解の深化が求められる。具体的には、気温上昇に伴う植生起源揮発性有機化合物 (BVOC) の排出量増大と、それに伴う 2 次有機エアロゾルの濃度増大、および二酸化炭素濃度増加

に伴う植物のイソプレン排出量の増加が挙げられる。また、信頼性の高い統計データを得づらい森林火災由来の排出量についても精緻化が求められる。とくに衛星から観測しづらい泥炭火災や、火災由来の対流による鉛直輸送などの精緻化については、土中の炭素含有量や燃焼時の熱量などにも大きく影響されるので、気候モデルにおける陸面過程の精緻化がこれらのプロセスの表現向上につながる可能性がある。

2) 輸送過程

大気微量成分の大気中の輸送およびその過程での混合等については、水平および鉛直方向の解像度が大きく影響している。例えば、八代らは水平解像度を2~3km程度まで高めることにより、インドネシアにおける降水の日変化が低解像度の場合と比較して大きく振る舞いを変えることを指摘している[62]。また、佐藤らは3.5km程度の高解像度実験を行うことにより、極向きのエアロゾル輸送が低解像度の際よりもより適切に表現されることを示した[63]。前線内の収束帶や降雨の構造をより適切に表現するためには、この程度の解像度が必要となると考えられるが、以降の変性過程や沈着過程などの精緻化と並行して、気象の諸プロセスが適切に表現できる水平・鉛直解像度で実験が行えるよう、気象モデル開発者・研究者と連携をとりつつ開発を進めることが求められる。

3) 変性過程

たとえば、光化学オキシダントの主成分である対流圏オゾンは地表から直接排出されず、大気中における光化学反応によって窒素化合物や炭化水素などから生成される。このように、とくに短寿命な大気微量成分にとっては大気中における変性過程が非常に重要である。一例として、MIROC-CHASERでは90以上の化学種の250種以上の化学反応を考慮している[64]。ただし衛星によるホルムアルデヒドやイソプレンなどの観測結果などから、都市域や森林周辺での未解明のVOC(炭化水素)とOH(水酸基)ラジカルの反応がありうるともいわれている。また、海洋上の遠隔地においてはオゾン濃度が10ppb以下になるような極低濃度が観測されているが、ハロゲン等との反応も指摘されており、海洋起源のヨウ素や臭素およびその大気中での化学反応を導入することによりモデルの予報精度が改善される可能性がある。より長期的には、今後の高解像度化および化学反応系の複雑化とともに、準定常状態(QSSA: quasi-steady-state approximation)を仮定するのではなく、より効率的な化学反応モジュールの導入が望ましいと思われる(cf. Shen et al., 2020[65])。

4) 沈着過程

大気中を漂う微小粒子であるエアロゾルは、降雨、降雪に捕らえられたり凝集核として雲に取り込まれたりするなどして大気中から除去される。PM2.5の成分の一つであるブラックカーボン(黒色炭素粒子)は、森林火災等での不完全燃焼の際に排出される「すす」である。排出時には水に溶けにくい疎水性粒子として排出されるが、輸送途上において徐々に表面に水滴が付着し、親水性を獲得していく(cf. Kanaya et al., 2016 [66])。このような過程を適切に表現するためには、エアロゾルの種類だけでなく、粒径やその表面に付着した他の物質等も表現しうる2次元の粒径分布を考慮する必要がある。現在日本国内においても複数の研究機関でそのような詳細なエアロゾル状態を表現できるモデルが開発されつつあり、今後も継続的な改良が期待できる(cf. Kajino et al., 2019 [67]; Matsui & Liu, 2021 [68])。

また、日本の山岳域などでは地表付近の霧による大気中からの除去が重要であるといわれており、関西

域での実観測等に基づくパラメタリゼーションが一部の領域化学モデルに導入されている [69, 70]。

5) 放射過程

沈着過程の項目で触れたようなエアロゾル表面上への水滴の被覆量は、エアロゾルの大気中からの除去されやすさだけでなく、表面の水滴がレンズのように太陽光等を屈折させてエアロゾルの大気加熱効果を増幅させる役割があり、この観点からも詳細なエアロゾルモデルは重要な役割を持っている。粒径や被覆量を適切に考慮することによって、ブラックカーボンエアロゾルの地球温暖化への影響が既存の粒子表現と比べて7倍近い幅を持ち、排出時の粒径分布や輸送経路上での親水性獲得を適切に評価する必要があると指摘されている [71]。

年代 課題	2022～	2024～	2026～	2028～	2030～	2032～
大気化学	エアロゾル モデルの 高度化	高解像度衛 星データを用 いた高解像 度輸送検証	化学反応系 の精緻化	精緻化され たモデルで のO(1km) でのイベント実 験	精緻化され たモデルでの 長期計算	

図 4.37 気象・気候分野、大気化学のロードマップ

目標達成に向けて必要な取り組み

二酸化炭素を含む気候-化学相互作用シミュレーションについては、30年以上の長期積分が一般的であるが、メタンなどの SLCF(短寿命気候因子) のシミュレーションについても光化学的なフィードバックを考慮した場合、成層圏の水蒸気生成等まで含めると 10 年程度の積分期間が必要となる。対流圏における全球的な汚染気塊の輸送シミュレーションとしては数週間程度と比較的短い積分期間での事例解析で良い場合もあるが、気候応答を考える場合は年単位のシミュレーションが必要となる。また、森林火災のように年々変動の影響が大きいものもあるため、長期トレンドと年々変動の適切な評価のためには複数年のシミュレーションを行う必要がある。

全球化学輸送モデル MIROC-CHASER においては数十種類のトレーサーの移流計算が必要になるため、気候モデルよりもノード間通信は隣接通信の比重が大きい。このため高解像度化に伴う計算量増加の対策として、使用ノード数を増やすことが有効である。現在 MIROC では緯度方向でのノード分割と鉛直方向のノード内並列を考慮しており、プロセス空間の分割についてはより最適化する必要がある。

また、現状における力学過程計算部分は予報化学種のトレーサー数に応じて計算負荷が高くなっているため、計算機の効率的な利用のためにはとくに移流過程と光化学反応過程のアルゴリズム改良が必要である。

アプリケーションと計算量・計算資源

全球スケールおよび領域スケールの化学輸送モデルとしては国内研究機関において開発・運用されているものが複数あるが、ここでは一例として MIROC-CHASER を取り上げる。以下の例では水平解像度 80km での 10 年積分だが、解像度を 2 倍にすると必要資源量は約 8 倍に増加する。

アプリケーション

- モデル名: MIROC-CHASER (Chemical AGCM for study of atmospheric environment and radiative forcing)
- 力学コア: スペクトル法
- 移流スキーム: フラックスセミラグランジュ法
- 物理過程: 雲微物理、大気放射、乱流、陸面過程など
- 化学過程: 光乖離、気相反応、エアロゾル、雷 NO_x 生成、湿性・乾性沈着、生物起源放出など

問題規模

- 格子点数: 1200 万 (水平解像度 80km, 鉛直 60 層)
- ステップ数: 530 万 (dt=1 分, 10 年積分)

要求性能精査の根拠等 (過程) などの説明

- メモリ量: 40PB
- 総演算量: 14,000 EFLOPS

4.10.2.3 海洋

背景

海洋は大気に比べて流れは緩やかであるが熱容量が大きいため、特に長期の気候変動に対して重要な役割を果たしている。また、CO₂ やメタンの吸収などの海洋による物質循環を通じて、地球温暖化にも大きな影響を持つと考えられる。すなわち、気候モデルを用いた長期気候変動予測のために、海洋は欠かすことのできない要素である。

全球規模の気候変動において、北極海は最も温暖化の影響が大きい海域で、特にここ数十年間での海氷面積の減少とそれに伴う海上気温の上昇は最も顕著である。近年は、北極海のバレンツ・カラ海での海氷減少に伴ってシベリア高気圧が強化し、我が国を含む北半球中緯度の冬季気候に対して遠隔的に影響しているとの指摘もある。さらに、北極海の海氷減少に伴って、我が国と米国東海岸・欧州を結ぶ航路として北極海を通る航路（北極海航路）の利用可能性に関する調査も進んでおり、数週間から数十年スケールの予測が航路選択や造船計画に利活用されつつある。このように、北極海の海洋-海氷変動システムの解明に対する学術界・産業界での重要性が近年高まっている。

一方、北大西洋のメキシコ湾流・北大西洋海流、北太平洋の黒潮は、低緯度にある高温な海水を高緯度に輸送することで顕著な南北熱輸送を担う代表的な海流である。しかしながら、これらの海流の正確な表現は、CMIP6 に参画するような最新の気候モデルでも困難である。特に我が国の南岸を流れる黒潮は、全球規模の気候に対してはもちろんのこと、日本の気候・気象に対しても大きな影響を持つことが知られている。近年はさらに、スマート水産業や洋上発電など、日本近海の海流情報の有効な活用法が提唱されつつあり、これら産業界のニーズを満たすためにも、詳細な海流の予測やそのメカニズム解明の重要性は

増している。

学術的・技術的課題、目標

気候モデルを用いた長期気候変動の予測の向上において海洋モデルの高精度化は必須である。気候モデルの海洋コンポネントモデルにおいて全球をほぼ同じサイズの格子で覆う場合、CMIP6に結果を提出している世界で最も細かいモデルの格子サイズは約10kmである。一方、CMIP6に参加している我が国の2つの気候モデルの海洋コンポネントの格子幅は100km程度であり、海洋高解像度版でも格子幅は25km程度である。この25kmと10kmの差によって、前述の黒潮やメキシコ湾流などの西岸境界流の正確な表現にも差が出てくると考えられる。格子サイズを変更するためには、新たなチューニングに加え、複数の感度実験と多くのアンサンブル実験を行わなければならない事情もあり、次期CMIPに向けた計算では格子幅25km程度に据え置かれている。さらに5~10年後に想定される後継のCMIPでは、多くのモデルで、10km程度の水平格子幅が標準的になる可能性が高い。国内のモデルグループも、後継CMIPに向けて、5年後を目標に全球10km海洋モデルを組み込んだ気候モデルの構築を目指す必要がある。

一方、上記のように全球をほぼ同じサイズの格子で覆うのではなく、一部だけを高解像度としたネスティング海洋モデルを気候モデルのコンポーネントにする試みも進めていくべきである。まず、温暖化の影響が最も大きく長期予測が学術的・社会的に重要視されている北極域を高解像度としたネスティング全球海洋モデルに対して、大気モデルを結合した気候モデルの構築・実行を目指す。

上記は5年以内に着手可能なものであるが、10年先を見据えてかつ日本近海の海流情報に対するニーズの高まりを考慮すると、日本近海をさらに高解像度としたモデルについても着手する必要があると考えている。この場合、最低でも日本近海の水平解像度を500m程度とする海洋モデルの構築を目指すべきだと考えている。また、海洋循環の応用分野をより広げるためにはさらなる高解像度モデルが必要で、日本近海のみ水平格子幅100m程度にしたモデルの構築も可能になるのではないかと期待している。

年代 課題	2022 ~	2024 ~	2026 ~	2028 ~	2030 ~	2032 ~
海洋 (長期気候予測)	海洋全球10kmの気候モデル ←→ 北極3-5km/全球25kmの 気候モデル ←→	左記モデル のアンサン ブル実験	日本近海500m(多重ネスト) の 海洋モデル ←→	日本近海100mの海洋非静力モデル ←→		

図4.38 気象・気候分野、海洋のロードマップ

目標達成に向けて必要な取り組み

まず、水平格子幅10kmの全球海洋気候モデルの開発の現状について述べる。我々のグループで開発している海洋モデルCOCOは、水平格子幅10kmの海洋単体モデルとして、現在、研究などに活用されており、問題なく稼働している。さらに、大気非静力モデルNICAMと海洋モデルCOCOを結合したモデルNICOCOにおいても、水平格子幅10kmの海洋モデルの結合に既に着手しており、それを用いた黒潮域における大気海洋相互作用に関する研究を進めている。このように、気候モデルにおける海洋モデル高解像度化に関する技術的な課題はこれらの経験を踏まえて克服されており、あとは計算機資源の不足が唯

一の問題である。

また、北極海高解像度の海洋ネスティングモデルは、既に海洋単体でのテストランを完了している。さらに、黒潮域を含む日本近海を高解像度とした海洋ネスティングモデルを組み込んだ気候モデルによる研究例も既に存在し、これらの経験が北極海高解像度気候モデルの構築にそのまま生かされる。すなわち、長期気候変動をターゲットとして海洋ネスティング気候モデルを構築・実行する上での技術的課題はおおよそ解決済みであり、既に計算機資源の充実が待たれる段階にある。

日本近海 500m の海洋モデルを組み込んだ気候モデルについては、多重ネスティング手法によって構築した海洋単体モデルの実行・研究経験がある一方、大気モデルを結合した気候モデルの開発は未着手であり、今後取り組むべき技術的課題であると考えている。また、日本近海のみ水平格子幅 100m 程度にした海洋単体モデルは、従来のモデルで用いられてきた静力学仮定が満たされない現象も陽に表現できるようになるため、非静力海洋モデルの導入も検討するべきだと考えている。

アプリケーションと計算量・計算資源

全球 10km 海洋モデル COCO は、格子数が $3,600 \times 3,000 \times 62$ である。タイムステップは 6 分、長期気候変動をターゲットとするために 200 年の積分を考えると、タイムステップ数は 1.3×10^8 である。東京大学 Wisteria (1 ノードあたり 48 コア/CPU で理論値 3.3TFLOPS) を 500 ノード (理論値 1.65PFLOPS) 使用して 1 年間の積分に約 5 時間半を要するので、200 年積分 (1 ケース) に 1 か月半かかる。上記は海洋単体モデルであり、気候モデル MIROC において大気モデルも同程度の計算量を割り当てる場合、1 ケース辺り 3 カ月要する。数ケースのシナリオの実行を想定すると、1.65PFLOPS を 1 年間使用する計算になる。さらにアンサンブル数を 5 度として、それを数年間で実行すると考えると、必要な総演算資源量は 3PFLOPS 程度である。

海洋モデル COCO の北極海高解像度ネスト版は、外モデルは全球 25km 海洋モデルで格子数が $1,440 \times 1,280 \times 62$ 、北極海モデル (解像度 3km) は $7,200 \times 1,280 \times 62$ である。タイムステップは 3 分、長期気候変動をターゲットとするために 200 年の積分を考えると、タイムステップ数は 2.6×10^8 である。上記の全球 10km モデルに比べて、実行内容は同じで海洋部分の総演算量は約 2 倍、大気モデルは前述の海洋 10km モデルと共にすると気候モデル MIROC としては 1.5 倍となる。そのため、必要な総演算資源量は 4.5PFLOPS 程度である。

日本近海 500m の海洋モデルは 4 重ネストで全球モデルから段階的に細かくする。最も解像度の高い最内モデルの格子数が $4,500 \times 3,600 \times 62$ である (より解像度の粗い外側のモデルの総計算量は最内モデルの 20% 程度だと考えられるので以下の見積もりでは割愛する)。積分期間は長くとも 1 年程度であり、タイムステップを格子幅からの推定で 1 分とすると、タイムステップ数は 5.3×10^6 である。ただし、これほどの高い解像度では潮汐の影響を考慮に入れなければならず、流速の大きい短周期の潮流が生じるため、タイムステップは 1 分より短くなる可能性もある。このような高解像度シミュレーションにおいては、アンサンブル数を確保する必要があり、それを 20 度とすると必要な総演算資源量は 81TFLOPS である。これよりさらに細かい日本近海 100m 格子モデルを用いて同様の計算を行うと、500m 格子モデルの 125 倍だとすると必要な総演算資源量は 10PFLOPS である。なお、この水平解像度では静水圧近似が成り立たない現象が顕著になる可能性もあり、日本近海 100m 格子モデルとして非静力海洋モデル KINACO の使用も検討する。

詳細な問題規模と総演算量については、図 4.39 を参照いただきたい。

モデル名（略称）	モデル名	文献	計算手法	問題規模	総演算量
MIROC	Model for Interdisciplinary Research on Climate	Tatebe et al. (2019)	(大気モデル) 力学過程：スペクトル法、静力学圧縮流体方程式 物理過程：積雲対流スキーム、大気放射、陸面過程など (海洋モデル) 一般直交全球三極座標・静力学ブシネスク流体方程式・海面混合層スキーム・海底境界層スキーム・高次移流スキーム	格子サイズ：大気はT213L40(60km)、海 洋は3600×3000×62 ステップ数：1.3億 (dt=6分、200年積分) ケース数：3 ケース アンサンブル数：5	3PFLOPS
MIROC	Model for Interdisciplinary Research on Climate	Tatebe et al. (2019)	(大気モデル) 力学過程：スペクトル法、静力学圧縮流体方程式 物理過程：積雲対流スキーム、大気放射、陸面過程など (海洋モデル) 一般直交全球三極座標・ネスティング手 法・静力学ブシネスク流体方程式・海面混合層スキ ーム・海底境界層スキーム・高次移流スキーム	格子サイズ：大気はT213L40(60km)、海 洋は全球1440×1280×62+北極海7200× 1280×62 ステップ数：2.6億 (dt=3分、200年積分) ケース数：3 ケース アンサンブル数：5	4.5PFLOPS
COCO	CCSR Ocean Component Model	Hasumi (2006)	(海洋モデル) 一般直交全球三極座標・ネスティング手 法・静力学ブシネスク流体方程式・海面混合層スキ ーム・海底境界層スキーム・高次移流スキーム	格子サイズ：日本近海4500×3600 (500m) ×62と全球モデルを段階的多重ネストモデ ルで連結 ステップ数：530万 (dt=1分、1年積分) アンサンブル数：20	81TFLOPS
COCO or KINACO	CCSR Ocean Component Model or KINACO is Not A COCO	Hasumi (2006) or Matsumura and Hasumi (2008)	(海洋モデル) 一般直交全球三極座標・ネスティング手 法・静力学or非静力学ブシネスク流体方程式・海面混合 層スキーム・海底境界層スキーム・高次移流スキーム	格子サイズ：日本近海22500×18000 (100m)×62と全球モデルを段階的多重ネ ストモデルで連結 ステップ数：2650万 (dt=12秒、1年積 分) アンサンブル数：20	10PFLOPS

図 4.39 各実験の計算設定と計算量

4.10.3 データ同化・データサイエンス（AI, 機械学習）

背景

データ同化は、数値予報モデルによる第一推定値を、観測データを利用して修正することで最適な初期値（解析値）を推定する方法である。これまで様々なデータ同化手法が多くの研究によって開発されてきた。現在では、4次元変分法（4D-VAR [72, 73]）及びアンサンブルカルマンフィルタ（EnKF [74, 75]）が主流となっている。4D-VAR、EnKF ともに流れに依存した誤差成分を考慮する高度なデータ同化手法で、理想的な条件のもとでは同等の精度が得られることが知られている [76]。また、両手法のメリットを生かすハイブリッドデータ同化手法の開発が盛んに進められている。さらに、これまでのデータ同化手法が確率分布のガウス性を仮定しているのに対し、非ガウス性を考慮することが可能な粒子フィルタ法を気象モデルに適用する研究も進んでおり [77]、降水過程などの非ガウス性の強い現象に対して効果を発揮することが期待される。また、様々な分野で研究開発が進められている機械学習をデータ同化で活用する取り組みも盛んである。

降水、強風、落雷といった顕著現象をより精度良く予測するには、時空間的には数 100m～数 km、数分～数時間スケールの気象情報が必要であり、数値モデルの計算のメッシュとしては数 10m～数 100m、時間刻みとしては 1 秒以下という高い時空間解像度が求められる。竜巻のようなマイクロスケールの顕著

現象の実用的な予測のためには、時空間的にさらなる高解像度が必要となる可能性がある。また、高い空間解像度は、数値モデルだけでなく、データ同化のための観測データにも求められる。

学術的・技術的課題、目標

これまで、数値天気予報におけるデータ同化は、大気、海洋などそれぞれにおいて行われてきた。近年では、大気と海洋のデータ同化を同時に行う強結合データ同化の開発が進められている。数値予報モデル開発においては、エアロゾルモデルや化学モデル、植生モデルなど様々なコンポーネントをつなぐ地球システムモデルの開発も進められている。地球システムモデルの包括的データ同化は数値予報の精度向上に重要な役割を果たすと考えられる。また、近年では、様々な分野で研究開発が進められている機械学習をデータ同化で活用する動きも見られる。

これまでにも、気象レーダや地上観測、人工衛星観測などの膨大なデータが用いられてきたが、近年はフェーズドアレイ気象レーダや気象衛星ひまわり8号などの登場により、従来と比べて一桁以上大きな観測ビッグデータが利用可能になってきた。高解像度高頻度の観測データを有効活用し予測精度を向上するには、数値モデルのさらなる高解像度化や膨大な観測データを効率的にシミュレーションに取り込むデータ同化システムの開発が求められている。線状降水帯のような極端現象は非線形性が強く、非ガウスを考慮するデータ同化手法の開発が必要となる。これを達成するためには、大規模並列計算に適した数値予報モデルとデータ同化手法が必要となる。例えば、数値予報モデルの解像度や観測データに含まれる誤差相関などの影響で、多くの観測データを有效地に利用できず、捨てざるを得ない。今後、データ同化の更なる研究開発により、観測誤差相関を直接考慮する方法などを開発して、より多くの観測データを効率よく利用して予報精度が向上することが期待される。

気象予測において、データ同化を用いて多くの観測情報を取り込んで精度の良い初期値を作成することは必要不可欠である。地球観測技術は年々高度化しており、数多くの人工衛星や地上観測装置が開発されている。そのデータ量は1世代前のものと比べると桁違いの数十～百倍となり、まさにビッグデータである。こうした人工衛星や、レーダ等の地上観測網の整備には多額の資金が必要であり、その価値を余すことなく利用することは我々に課せられた重要な責務である。近年では測器の設計前にその有効性をシミュレーションして評価する「観測システムシミュレーション実験 (Observing Systems Simulation Experiment: OSSE)」の重要性が高まっている。OSSEにより、どのような測器をどう配置・運用すれば効果的か、どのようなデータ同化システムを用いればどの程度予測の精度向上につながるかを事前に定量的に評価することができる。社会基盤としての気象予測システムの維持発展のためにはこのような研究も欠かすことができない。

実際の気象予測においては、フラッグシップ計算機のような計算資源を投入することは出来ないが、フラッグシップ計算機による知見を低コストで反映するために機械学習手法を応用することが考えられる。例として大規模計算の結果を学習し、低解像度の計算から高解像度予測を作成する超解像技術が挙げられる。その他のデータ駆動手法の応用としては、物理過程のパラメタリゼーションの機械学習による置き換え、モデル誤差の修正、観測演算子の機械学習による構築、機械学習手法を用いたデータ同化手法の改善、モデル出力の統計補正などが挙げられる。

目標達成に向けて必要な取り組み



図 4.40 気象・気候分野、データ同化・データサイエンス (AI, 機械学習) のロードマップ

現在気象分野で用いられているデータ同化手法には、大きく分けて変分法とアンサンブルカルマンフィルタがある。将来的な方向として、アンサンブルカルマンフィルタによって得られる「その日の流れ場」を考慮した背景誤差共分散を4次元変分法の中で利用することによって精度向上を図る、ハイブリッド手法があり、近年盛んに研究されている。また、これらのデータ同化手法が確率分布のガウス性を仮定しているのに対し、非ガウス性を考慮することが可能な局所粒子フィルタ法を気象モデルに適用する研究も進んでおり、降水過程などの非ガウス性の強い現象に対して効果を発揮することが期待される。同様に、非線形性を考慮するために機械学習手法を用いる研究もある。

アンサンブルカルマンフィルタにおける主要な演算は、固有値分解や逆行列などの行列演算である。これまでアンサンブル数はせいぜい数十から百程度であったが、今後は少ないアンサンブル数によるサンプリングノイズを減らすために数百以上のアンサンブル数を確保することが求められるようになるだろう。しかし、行列の演算量は行列の大きさに対してその2乗や3乗で増加するため、アンサンブル数や観測データ数によって演算量は劇的に増加する。数値予報モデルでは演算効率を上げるために、単精度実数の利用が進められている。データ同化においても単精度実数の利用を検討する必要があるが、行列演算の精度に影響を与える可能性がある。富岳でのグランドチャレンジでは、数値予報モデル、データ同化両方において積極的に単精度の利用を試み、解析精度の精度悪化は最小限であることを確認した。このような問題を含め、演算効率を向上させるために単精度実数の利用やアルゴリズムの改良を含めたさらなる開発が必要となるだろう。また、アンサンブルカルマンフィルタでは、アンサンブル数だけのファイル入出力を伴うため、ストレージの量だけではなく巨大になるファイル数に対応したファイルアクセス速度の向上も非常に重要である。気象学やデータ同化のアプリケーション開発者のみでこれらの問題を解決することは難しい。富岳の開発では、アプリケーション開発者と計算機開発者による協調設計「コデザイン」によって、アプリケーションの高度化、様々なタイプのアプリケーションにとって使いやすい計算機開発に成功した。次世代のスーパーコンピュータ開発においても「コデザイン」が非常に重要となる。

アプリケーションと計算量・計算資源

(A) データ同化 (NICAM-LETKF)

NICAM-LETKF は、局所変換アンサンブルカルマンフィルタ (LETKF) を全球非静力学大気モデル (NICAM) に適用したものである。LETKF は格子毎に計算が完全に独立しているため並列性に非常に

優れている。

問題規模

- 実行形態：1時間毎に新たな観測データを取り込んで初期値をアップデートする（1時間サイクル）。このため、大気モデルのシミュレーションを1.5時間分行い、1.5時間積分した時刻の前後30分の10分毎のアンサンブル出力値を用いてデータ同化を実行する。これを1,440サイクル繰り返し2ヶ月間のシミュレーションを行う。
- 格子数：全球3.5km、鉛直94層（約43億格子）
- アンサンブルメンバー数：1,000
- ケース数：2
- 総演算量：896,000 EFLOP
- 総メモリ使用量：10万ノードを使った実行で1.4PB
- 総ストレージ容量：15PB

大気モデルの演算量は解像度 n 倍に対して n^3 倍に増える一方、アンサンブルデータ同化では、解像度に対して n^2 倍、アンサンブルメンバー数 m に対して m^2 倍から m^3 倍で増加する。この実験を1ケースあたり24日で行うとする場合、必要となる計算機性能は演算性能で220PFLOPS、メモリバンド幅性能で270PB/sである。このうち、アンサンブルデータ同化は行列演算が大きな部分を占めるので、大気モデルと比較してより高い演算性能を要求する傾向にある。ネットワーク通信については、大気モデルでは主に隣接通信を行い、アンサンブルデータ同化では数回の大域通信を行う。どちらも現在のネットワーク速度で十分と見積もられる。ファイルI/Oでは、大気モデルシミュレーションとアンサンブルデータ同化にかかる時間が1:1であるとすると、大気モデルでは実時間で20~30秒ごとにノードあたり3GB程度のファイル出力をを行うと見積もられ、300MB/s程度のファイルI/O性能が必要である。これに対しアンサンブルデータ同化は書き出された大気モデルの結果を一気に読み込んで処理するため、ストレージ帯域全体で80TB/s程度の性能があると望ましい。大気モデル及びデータ同化における通信は、富岳での実測値からの推定で総経過時間に対する割合は高くないことがわかっている。上記の問題サイズにおける、1サイクルでのデータ同化での通信量は約3PBと見積もられる。データ同化1サイクルを約1時間で終わらせ、通信にかけても良い時間を全体の2%と仮定すると、要求されるネットワークバンド幅はノード当たり高々0.5GB/sとなる。

富岳のコデザインでの取り組みによって、アルゴリズムの改良やチューニングによって演算性能が向上した。また、現在気象モデルでは倍精度実数を用いるのが一般的であるが、演算性能を向上させるために単精度化への取り組みが始まっている。これにより演算性能が上がることが実証され、さらに物理性能もおおよそ保証されることが示されている。

(B) データ同化(SCALE-LETKF)

SCALE-LETKFは大規模並列計算に向けて設計された領域気象モデルSCALE-RMと、並列実行性能に優れたデータ同化手法である局所アンサンブル変換カルマンフィルタLETKFを組み合わせたシステムである。ここではSCALE-LETKFを念頭に、将来の領域気象予測に必要となる計算量の見積もり

を行う。

問題規模

竜巻スケールの現象を解像できる数値天気予報の実現可能性について考える。解析領域としてフェーズドアレイ気象レーダの探知範囲 120km 四方を設定する。50m 解像度、100 アンサンブルの領域気象モデルの時間積分を 30 秒先まで行い、30 秒毎にデータ同化を行う。観測データとして、300 方位、600 レンジビン、110 仰角のフェーズドアレイ気象レーダの観測データを同化する。得られた解析アンサンブルを初期値として、100 メンバーの 30 分先までのアンサンブル予測計算を行う。

- 格子数：50m 解像度、鉛直 50 層（2.88 億格子）
- アンサンブルメンバー数：100
- 総演算量：2EFLOP（30 秒あたり）

以上を 30 秒以内に処理するために必要な計算機性能は、約 10^{17} FLOPS=100PFLOPS。実行効率を現在と同じレベルの 10 %と仮定すると、計算機の理論性能としては 1EFLOPS が求められる。主記憶は合計 300TB、ファイル IO は 1 サイクル辺り 12TB 必要となる。

4.10.4 世界の動向とハードウェアに求める計算機性能

本項では気象・気候アプリケーションがハードウェアに求める性能要求について、各アプリケーションに見られる共通項を抽出し、俯瞰的な視点でまとめる。しかし現在のアプリケーションが置かれている立場は、ハードウェアに自由な要求を行う側にいるというよりも、電力性能向上や計算機の市場トレンドに駆動されたハードウェアの変化に起因し、さまざまな制限を受けている側にいるといえる。そのため、ハードウェア側への要求と同時に、アプリケーション側でどのような変革を行うべきかという視点に立って述べることとする。加えて本項では、世界における気象・気候アプリケーションの開発動向についても触れるが、多くのプロジェクトはハードウェアから受ける制約を如何に打破するかという視点で行われているという特徴がある。将来実現を目指す科学的な挑戦を進めるために克服すべき技術的挑戦の比重は増大しており、ハードウェアとシステムソフトウェア、そしてアプリケーションが互いに連携したコデザイン活動の重要性は一層増している。

並列性の確保の問題

ここ 30 年のコンピュータの性能は、ムーアの法則に従い 1 年間に 1.6 倍の浮動小数点演算性能の向上を達成してきた。しかし 2000 年代以降、CPU の動作周波数向上は電力消費量の増加を理由に頭打ちとなり、演算コアの数を増加させる方法に切り替えることで、性能向上が維持されることとなった。近年における CPU 内でメインメモリ・キャッシュを共有したコア数の増加は顕著であり、Intel Xeon Phi (Knights Landing 世代) や AMD EPYC (Millan 世代) のように、1CPU に 50 コア以上搭載されるのも珍しくない。また、CPU1 コア内においても、1 回の命令で複数のデータに対して同じ演算を行う方法 (Single Instruction Multi Data: SIMD) が強化される方向にあり、例えば京コンピュータでは倍精度・単精度演算で同時 2 つであったのに対して、富岳では倍精度・単精度・半精度の演算でそれぞれ 8、16、32

個のデータを処理する[78]。GPUにおいても、例えばNVIDIA TESLA A100は約7,000のStreaming Processor(SP, NVIDIAの場合はCUDAコア)を備え、CPUがコア数×SIMDで並列処理するデータをSPで処理する。この時、CPUもGPUも演算に対して相対的に速度の遅いデータ転送を隠蔽するために、演算とデータ転送がオーバーラップされる形で実行する。CPUではこの手法をソフトウェアパイプラインングを用いて実現しており、性能が向上するために必要なデータ数はSIMD実行数の数倍になる。そのため、実際にハードウェアが同時処理するデータ数よりもさらに多くのデータ数を用意する必要があり、例えば富岳では数百個のデータを同じ命令で同時処理する形になりがちである。スーパーコンピュータでは、このようなCPU・GPUをさらに数千・数万と並べてネットワーク接続することで最大性能を実現している。シミュレーション側はこのようなハードウェアの階層的な並列処理機構に合わせて、計算を処理する必要がある。最も重要なのは同じアルゴリズムで同時処理できるデータを用意することであり、これを並列性の確保という。気象・気候計算の科学的な要請とマッチしつつ並列性を確保するのに有効なのは、空間解像度(格子点数)の向上とアンサンブル実行数の向上である。特にアンサンブル数の向上は、後述する空間解像度の向上と比較して困難が少なく、シミュレーションモデルそのものに加える改良も少ないため、向こう10年の計算機の変化の中でも非常に有効な手段であるといえる。

一方で空間解像度を向上させた場合、一般的にCFL条件により時間刻みも小さくしなければならず、同じ時間分をシミュレーションするために必要な演算量が増大する。気象・気候モデルはほとんどの場合、空間方向にプロセス分割して並列化するため、高解像度化に伴う格子点数の増加はプロセスあたりの格子点数を変えずにプロセス数を増加(ウィーク・スケーリング)させることによって対応できる。しかしそのままでは、時間刻み幅が小さくなり計算ステップ数が増大したことに対応できないため、高解像度シミュレーションは気候シミュレーションと同程度の計算ステップ数では数日から数ヶ月の短期間計算にとどまらざるを得ない。現業の気象予報や長期気候計算で必要となる、計算をより短い時間で達成する(Time-to-Solution)という観点においては、ステップ数の増大に対応するためにプロセスあたりの問題サイズ(=格子点数)を減らしながら利用プロセス数を増加(ストロング・スケーリング)させる必要があり、並列性の確保や通信時間が占める比率の相対的な増加により、性能向上はかなり困難である。今後のモデル開発では、1) 1時刻ステップの計算を高速化する、2) 1時刻ステップの時間刻み幅を長くする、3) プロセス間通信にかかる時間を短縮する、といった点が重要視されると考えられる。ハードウェア・システムソフトウェア側に期待する機能を以下に挙げる。

- SIMD並列実行の柔軟性の向上：長いSIMDのビット幅を使い切れないような計算の場合も、ループ内の複数の計算式をまとめ上げて計算可能にするようなコンパイラの機能や、ハードウェア機構が追加されることが望ましい。
- 通信速度・レイテンシ性能に比重をおいた設計：相対的に比重の増すプロセス間通信について、できる限り通信時間を削減するような設計が行われることが望ましい。特に加速機構へオフロードしたデータをノード間をまたいで通信するケースについて、より通信時間を削減するための研究開発が求められる。

一方で、アプリケーション側に求められる変革としては、以下のようなものが挙げられる。

- モデルの時間発展に着目したアルゴリズム研究開発：1時刻ステップの時間刻み幅を長く取れるよ

うなスキームの選択や、計算安定性を向上させるための開発が求められる。一方で、空間解像度と同様に時間解像度もシミュレーションの表現力を向上するために重要な要素であり、敢えて短い時間刻み幅を前提としてアルゴリズムの簡略化を図り、1時刻ステップの計算を高速化するという戦略も考えうる。

- 新たな方法による計算高速化：解きたいアルゴリズムを計算機に最適化する方法は限界にきており、後述する低精度計算の活用や代替モデルの活用など、これまでのシミュレーション精度を犠牲にするリスクを伴う変革が必要とされる。
- 通信スキームの改良：ハードウェア側へ要求する通信速度向上と連動して、アプリケーション側も通信部分の再実装を必要に応じて進め、コデザインを行うことが想定される。

メモリ性能と容量の問題

ムーアの法則に従って性能が向上しているのは浮動小数点数演算性能であり、メモリ性能の向上はより緩やかにしか進展しない。その速度は概ね1年で1.16倍と見積もられており、10年で4.4倍しか高速化しないことになる。気象・気候モデルは演算量に対して用いる変数の数が多く、データ転送量が多いメモリ律速型のプログラムであり、ハードウェアのメモリ性能に頼った形の性能向上を期待していくには、いつまでも計算可能な問題を拡大することは出来ない。メモリ容量の問題もある。高速なメモリはコストが高く小容量であり、安価で大容量のメモリは低速であるというトレードオフがある。そのため、格子点を一度にたくさんメモリ上に準備して並列性を確保することと、高性能のメモリを準備して計算を高速化することの両立を図ることは困難である。また、より多くの鉛直層数やより多くの予報変数を用いて計算したいという科学的要請から、利用メモリは増大する傾向にあり、容量の不足により今後10年で計算したい問題がこれ以上大きくできない可能性も高い。

この問題を解決する方法として、低精度浮動小数点実数の活用と不揮発性メモリの利用が挙げられる。前者はアプリケーション側で行う施策であり、精度を落とすことにより、さまざまな計算上の恩恵を受けようとするものである。例えば、メモリ使用量の減少、メモリからの転送データ量の削減による高速化、キャッシュに留まるデータ数の増大、SIMD演算で同時に処理できるデータ数の増大、保存するファイルサイズの削減、ファイル圧縮率の向上等が利点として挙げられる。一方でシミュレーション結果の改悪が起こることもあるため、解きたい問題に合わせた再現性能の検証が必要となる。国内では大気モデルNICAMの力学コアについて、理想実験を用いた単精度シミュレーション結果の精度検証が行われ、その差は十分に小さいことが示された[79]。また欧州ECMWFは、2021年5月より現業予報を単精度で行うよう変更し、その代わりに鉛直層を増加させることによって、ほぼ同じ計算リソースで総合的なシミュレーション精度を向上させたと報告している[80]。

後者に挙げた不揮発性メモリは、ファイルストレージとメインメモリの中間に位置するような大容量のメモリである。この容量を活かして、より大きなメモリ空間を必要とする問題を計算することが期待されるが、前述の通り容量の大きいメモリは低速であったり、レイテンシが長いという特徴も併せ備えている。演算装置からキャッシュ、メインメモリに連なるデータ格納場所の階層構造がもう一段深くなり、さらに複雑化するという点も技術的課題として考えられる。これらの難点を打破し、高速な計算を実現するため、システムソフトウェアの研究開発の進展が望まれる。その際には、気象・気候モデルの開発者が協調して研究を支援し、アプリケーションにとって有効かつ望ましいハードウェアの進化が行われるべきで

ある。

プログラム・性能の移植性の問題

向こう 10 年のスーパーコンピュータのトレンドとして、GPU を用いた計算機が大勢を占めることは想像に難くない。近年の高性能計算の新たな需要は機械学習が牽引しており、多くの機械学習ライブラリが活用する GPU への研究開発投資や生産設備投資が増大している。結果として、計算機の調達コストは GPU を利用した方が抑えられる。今後利用可能な計算資源の中で GPU の占める割合が大きくなる場合、GPU 上で計算を行うことの可能なソフトウェアに有利に働く場面が多くなると予想される。しかし、気象・気候モデルを GPU に対応させるのは一般的に困難である。これは、データ転送律速型のアプリケーションであることと同時に、特定の区間が計算時間の大部分を占める「ホットスポット」が存在しないため、数十万行にわたるモデルコードの全体を GPU 上で計算しないといけないためである（例えば Yashiro et al., 2016[81], Lawrence et al., 2018[82]）。モデル開発者はハードウェアテクノロジーの変革のたびに大移植作業を行いたくないと考えており、そのような要請から、Fortran や C 言語でのソフトウェア開発をやめ、ドメイン特化言語（Domain Specific Language, DSL）や新たな設計で作られたプログラム言語を用いて、アルゴリズムの記述と計算機への最適化を分離するような開発の方向性が提案された。海外では既に、このような可搬性と高性能を両立させるプログラミング環境への移行が試験的なフェーズを終え、実際にプロダクションランを行う域に達している。例を挙げると、米国で開発されている E3SM は C++ のテンプレートフレームワークである kokkos [83] を利用して力学コアと主要な物理課程スキームを置き換え、世界トップクラスの GPU スパコンを用いた計算性能評価を進めている（例えば Bertagna et al., 2019 [84]）。また、NOAA の新現業予報システム UFS の中核となる大気モデル FV3 は、地球物理流体力学研究所（GFDL）とアレン人工知能研究所の研究者らの手によって、スイスのスーパーコンピュータセンターが開発した GridTools および GT4Py を用いた記述に置き換えられつつある [85, 86]。GridTools は C++ のテンプレートフレームワークであり、GT4Py はアルゴリズム開発者が Python で記述したコードを変換し、GridTools フレームワークに適用させるライブラリである。一方、カリフォルニア工科大学、マサチューセッツ工科大学、NASA ジェット推進研究所による気候モデリングアライアンス CliMA [87] は、変数の型変換や GPU の利用が容易な新しいプログラミング言語である Julia を用いて、ClimateMachine.jl をはじめとする新たな気候モデルソフトウェア群を整備している（例えば Sridhar et al., 2021 [88]）。欧州では、ECMWF が気候モデルの格子生成、離散化、プロセス並列化などを容易に切り替えられる DSL 言語 Atlas[89] を開発し、従来のスペクトルモデルから格子モデルへの移行や、マシンアーキテクチャの移行をスムーズに行おうとしている。物理過程についても、前述の GridTools/GT4Py を用いた実装が進められている。英国気象局は次世代モデル開発基盤として PSyclone というドメイン特化コンパイラを開発し、DSL で記述した新たな大気・海洋モデルの力学コアのプログラムコードをマシンに合わせた形で生成する機能のほか、既存の Fortran コードを解析して変換する機能を実現した [90]。また、EU の次世代地球システムモデル開発のための国際共同プロジェクト NextGEMS(<https://nextgems-h2020.eu/>) に関連して、ドイツ気象局とマックスプランク気象研究所が共同で開発した気候モデル ICON を、GridTools/GT4Py を用いて置き換えるプロジェクト (<https://exclaim.ethz.ch/the-project.html>) が進行している。

人工知能技術の活用

ディープニューラルネットワークをはじめとする機械学習技術の急速な進展は、気象・気候分野においても無視できないものとなった。この技術はまず気象・気候予測データから意味のある現象を抽出する試みから開始された。NICAM の計算結果から台風の「卵」を検出する試みは機械学習研究のコミュニティで広く行われている公開コンペの形で開催され、気象学と情報科学分野の新たな共同研究の繋がりを形成した [91]。気候データセットを用いた極端現象の抽出は 2018 年のゴードン・ベル賞に選出され [92]、その DNN アルゴリズムは現在機械学習のベンチマークである MLperf の一部として広く利用されている [93]。低解像度のシミュレーション結果から、高解像度の気象場分布を推定・生成する「超解像」という技術についても、ダウンスケーリングを用いた影響評価に用いられつつある（例えば Baño-Medina, 2019[94], Serifi et al., 2021[95], Vaughan et al., 2022[96]）。

機械学習技術をモデルシミュレーションの高速化に利用するという考え方も台頭してきている。物理モデルの結果を用いてニューラルネットワークを学習させ、「代理モデル」として利用することにより、計算時間を 100-1,000 倍高速化させようというものである。長時間の計算には向かない超高解像度計算や、爆発的に増加する観測データを用いて、まずは物理コンポーネントの中でも雲過程 [97, 98] や放射過程 [99] などのデータ駆動型モデルの構築が研究されている。また、粗い空間解像度の気候モデルの各格子点に高解像度 2 次元大気モデルを埋め込む「スーパーパラメタリゼーション」という手法の延長として、埋め込まれた子モデルをデータ駆動型モデルで代替する試みも進められている [100, 101]。

新たな計算原理等に基づくコンピュータの利用

量子コンピュータやニューロモーフィックコンピュータなどの、ノイマン型とは異なる原理に基づくコンピュータが台頭してきている。これらのコンピュータは今のところ特定のアルゴリズムの計算を劇的に高速化するが、気象・気候モデルが現状利用しているアルゴリズムは多岐に渡り、その全てを新計算原理によってカバーできるわけではないため、まだ応用について検討が始められたばかりである。しかし、当初は特定の用途のために開発された GPU が世代を重ねることでより汎用化され、現在のように気象モデルの大半のコンポーネントを GPU 上で計算できるようになったという歴史を鑑みても、さまざまなハードウェア・ソフトウェアの研究開発により実用的な利用が進む可能性はある。また、特定の物理課程については新計算原理を活用したパラメタリゼーション研究やモデリング研究が行われると考えられる。

近年利用が活発化している計算機コンポーネントに、FPGA がある。FPGA はプログラミングされたアルゴリズムを回路で再現するものであり、気象・気候モデルを含むデータ転送律速型のアプリケーションの計算を高速化、省エネ化する。現状における FPGA 利用の課題はチップサイズであり、巨大な気象モデルのソースコードをまるまるチップ上に再現するには回路面積が足りない。高位合成と呼ばれる回路作成についても時間がかかり過ぎて実用的ではないという問題がある。しかし FPGA についても、今後の技術進展によって気象モデルの一部または全部を FPGA で計算し高速化するといったことが現実に可能となることが期待され（例えば Ashworth et al., 2019[102]）、モデル側の研究開発についても、積極的な検討を行うべきである。

参考文献

- [1] IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Vol. In Press. Cambridge University Press, Cambridge, United Kingdom and New York, NY, US, 2021.
- [2] Mei, W. and S.-P. Xie. Intensification of landfalling typhoons over the northwest Pacific since the late 1970s. *Nat. Geosci.*, Vol. 9, No. 10, pp. 753–757, 2016. <https://doi.org/10.1038/ngeo2792>.
- [3] Yamaguchi, M. and S. Maeda. Increase in the number of tropical cyclones approaching Tokyo since 1980. *J. Meteorol. Soc. Japan Ser. II*, Vol. 98, No. 4, pp. 775–786, 2020. <https://doi.org/10.2151/jmsj.2020-039>.
- [4] Kossin, J. P. A global slowdown of tropical-cyclone translation speed. *Nature*, Vol. 558, No. 7708, pp. 104–107, 2018. <https://doi.org/10.1038/s41586-018-0158-3>.
- [5] Bauer, P., B. Stevens, and W. Hazeleger. A digital twin of Earth for the green transition. *Nat. Climate Change*, Vol. 11, No. 2, pp. 80–83, 2021. <https://doi.org/10.1038/s41558-021-00986-y>.
- [6] Miyamoto, Y., Y. Kajikawa, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita. Deep moist atmospheric convection in a subkilometer global simulation. *Geophys. Res. Lett.*, Vol. 40, No. 18, pp. 4922–4926, 2013. <https://doi.org/10.1002/grl.50944>.
- [7] Prein, A. F., R. M. Rasmussen, D. Wang, and S. E. Giangrande. Sensitivity of organized convective storms to model grid spacing in current and future climates. *Philos. Trans. R. Soc., A*, Vol. 379, No. 2195, p. 20190546, 2021. <http://doi.org/10.1098/rsta.2019.0546>.
- [8] Ito, J., T. Oizumi, and H. Niino. Near-surface coherent structures explored by large eddy simulation of entire tropical cyclones. *Sci. Rep.*, Vol. 7, No. 3798, 2017. <https://doi.org/10.1038/s41598-017-03848-w>.
- [9] Satoh, M., B. Stevens, F. Judt, M. Khairoutdinov, S.-J. Lin, W. M. Putman, and P. Düben. Global cloud-resolving models. *Curr. Climate Change Rep.*, Vol. 5, No. 3, pp. 172–184, 2019. <https://doi.org/10.1007/s40641-019-00131-0>.
- [10] Tomita, H., H. Miura, S. Iga, T. Nasuno, and M. Satoh. A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophys. Res. Lett.*, Vol. 32, No. 8, p. L08805, 2005. <https://doi.org/10.1029/2005GL022459>.
- [11] Miura, H., M. Satoh, T. Nasuno, A. T. Noda, and K. Oouchi. A Madden-Julian oscillation event realistically simulated by a global cloud-resolving model. *Science*, Vol. 318, No. 5857, pp. 1763–1765, 2007. <http://doi.org/10.1126/science.1148443>.
- [12] Grabowski, W. W. and P. K. Smolarkiewicz. CRCP: A cloud resolving convection parameterization for modeling the tropical convecting atmosphere. *Physica D: Nonlinear Phenomena*, Vol. 133, No. 1–4, pp. 171–178, 1999. [https://doi.org/10.1016/S0167-2789\(99\)00104-9](https://doi.org/10.1016/S0167-2789(99)00104-9).

- [13] Grabowski, W. W. Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.*, Vol. 58, No. 9, pp. 978–997, 2001. [https://doi.org/10.1175/1520-0469\(2001\)058<0978:CCPWTL>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<0978:CCPWTL>2.0.CO;2).
- [14] Stevens, B. and Coauthors. DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.*, Vol. 6, No. 61, 2019. <https://doi.org/10.1186/s40645-019-0304-z>.
- [15] Grabowski, W. W. Towards global large eddy simulation: Super-parameterization revisited. *J. Meteorol. Soc. Japan. Ser. II*, Vol. 94, No. 4, pp. 327–344, 2016. <https://doi.org/10.2151/jmsj.2016-017>.
- [16] Yamaguchi, T., G. Feingold, and V. E. Larson. Framework for improvement by vertical enhancement: A simple approach to improve representation of low and high-level clouds in large-scale models. *J. Adv. Model. Earth Syst.*, Vol. 9, No. 1, pp. 627–646, 2017. <https://doi.org/10.1002/2016ms000815>.
- [17] Yoshimura, H. Development of a nonhydrostatic global spectral atmospheric model using double Fourier series. In A. Zadra, editor, *Research Activities in Atmospheric and Oceanic Modeling*, Vol. 42 of *CAS/JSC WGNE*, pp. 3.05–3.06. WMO, 2012. https://wgne.net/bluebook/uploads/2012/individual-articles/03_Yoshimura_Hiromasa_WGNE_Yoshimura_1203.pdf.
- [18] Shima, S., K. Kusano, A. Kawano, T. Sugiyama, and S. Kawahara. The super-droplet method for the numerical simulation of clouds and precipitation: a particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Quart. J. Roy. Meteor. Soc.*, Vol. 135, No. 642, pp. 1307–1320, 2009. <https://doi.org/10.1002/qj.441>.
- [19] Wyngaard, J. C. Toward numerical modeling in the “Terra Incognita”. *J. Atmos. Sci.*, Vol. 61, No. 14, pp. 1816–1826, 2004. [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNTMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNTMITT>2.0.CO;2).
- [20] Kawai, Y. and H. Tomita. Numerical accuracy of advection scheme necessary for large-eddy simulation of planetary boundary layer turbulence. *Mon. Wea. Rev.*, Vol. 149, No. 9, pp. 2993–3012, 2021. <https://doi.org/10.1175/MWR-D-20-0362.1>.
- [21] Kochkov, D., J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer. Machine learning-accelerated computational fluid dynamics. *Proc. Natl Acad. Sci.*, Vol. 118, No. 21, p. e2101784118, 2021. <https://doi.org/10.1073/pnas.2101784118>.
- [22] Tomita, H. and M. Satoh. A new dynamical framework of nonhydrostatic global model using the icosahedral grid. *Fluid Dyn. Res.*, Vol. 34, No. 6, p. 357, 2004. <https://doi.org/10.1016/j.fluiddyn.2004.03.003>.
- [23] Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga. Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *J. Comput. Phys.*, Vol. 227, No. 7, pp. 3486–3514, 2008. <https://doi.org/10.1016/j.jcp.2007.02.006>.

- [24] Satoh, M. and Coauthors. The Non-hydrostatic Icosahedral Atmospheric Model: Description and development. *Prog. Earth Planet. Sci.*, Vol. 1, No. 18, 2014. <https://doi.org/10.1186/s40645-014-0018-1>.
- [25] Kodama, C. and Coauthors. The Nonhydrostatic ICosaHedral Atmospheric Model for CMIP6 HighResMIP simulations (NICAM16-S): experimental design, model description, and impacts of model updates. *Geosci. Model Dev.*, Vol. 14, No. 2, pp. 795–820, 2021. <https://doi.org/10.5194/gmd-14-795-2021>.
- [26] Hasumi, H. CCSR Ocean Component Model (COCO) version 4.0. Vol. 25 of *CCSR Rep.*, pp. 1–103. Center for Climate System Research, The University of Tokyo, 2006. <https://repository.exst.jaxa.jp/dspace/handle/a-is/42981>.
- [27] 気象庁. IPCC の RCP8.5 シナリオを用いた非静力学地域気候モデルによる日本の気候変化予測. 地球温暖化予測情報第 9 卷, 79 pp., 2017. <https://www.data.jma.go.jp/cpdinfo/GWP/>.
- [28] Mizuta, R. and Coauthors. Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteorol. Soc.*, Vol. 98, No. 7, pp. 1383–1398, 2017. <https://doi.org/10.1175/BAMS-D-16-0099.1>.
- [29] Ishii, M. and N. Mori. d4PDF: large-ensemble and high-resolution climate simulations for global warming risk assessment. *Prog. Earth Planet. Sci.*, Vol. 7, No. 58, 2020. <https://doi.org/10.1186/s40645-020-00367-7>.
- [30] Imada, Y., M. Watanabe, H. Kawase, H. Shiogama, and M. Arai. The July 2018 high temperature event in Japan could not have happened without human-induced global warming. *SOLA*, Vol. 15A, pp. 8–12, 2019. <https://doi.org/10.2151/sola.15A-002>.
- [31] IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. 1535 pp.
- [32] Smith, D. M., R. Eade, A. A. Scaife, and Coauthors. Robust skill of decadal climate predictions. *npj Clim. Atmos. Sci.*, Vol. 2, No. 13, 2019. <https://doi.org/10.1038/s41612-019-0071-y>.
- [33] Boer, G. J. and Coauthors. The decadal climate prediction project (DCPP) contribution to CMIP6. *Geosci. Model Dev.*, Vol. 9, No. 10, pp. 3751–3777, 2016. <https://doi.org/10.5194/gmd-9-3751-2016>.
- [34] Tatebe, H. and Coauthors. Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geosci. Model Dev.*, Vol. 12, No. 7, pp. 2727–2765, 2019. <https://doi.org/10.5194/gmd-12-2727-2019>.
- [35] Wilby, R. L. and S. Dessai. Robust adaptation to climate change. *Wea.*, Vol. 65, No. 7, pp. 180–185, 2010. <https://doi.org/10.1002/wea.543>.
- [36] Hawkins, E. and R. Sutton. The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, Vol. 90, No. 8, pp. 1095–1107, 2009. <https://doi.org/10.1175/2009BAMS2607.1>.
- [37] 気象庁. 日本の気候変動 2020 –大気と陸・海洋に関する観測・予測評価報告書-, 2020.

- https://www.data.jma.go.jp/cpdinfo/ccj/index.html.
- [38] 環境省. 気候変動影響評価報告書（総説）, 2020. <https://www.env.go.jp/earth/气候変動影響評価報告書（総説）.pdf>.
- [39] Ashie, Y. and T. Kono. Urban-scale CFD analysis in support of a climate-sensitive design for the Tokyo Bay area. *Int. J. Climatol.*, Vol. 31, No. 2, pp. 174–188, 2010. <https://doi.org/10.1002/joc.2226>.
- [40] Nishizawa, S., H. Yashiro, Y. Sato, Y. Miyamoto, and H. Tomita. Influence of grid aspect ratio on planetary boundary layer turbulence in large-eddy simulations. *Geosci. Model Dev.*, Vol. 8, No. 10, pp. 3393–3419, 2015. <https://doi.org/10.5194/gmd-8-3393-2015>.
- [41] Sato, Y., S. Nishizawa, H. Yashiro, Y. Miyamoto, Y. Kajikawa, and H. Tomita. Impacts of cloud microphysics on trade wind cumulus: which cloud microphysics processes contribute to the diversity in a large eddy simulation? *Prog. Earth Planet. Sci.*, Vol. 2, No. 23, 2015. <https://doi.org/10.1186/s40645-015-0053-6>.
- [42] Richardson, L. F. *Weather prediction by numerical process*. Cambridge University Press, 1922.
- [43] Charney, J. G., R. Fjörtoft, and J. Neumann. Numerical integration of the barotropic vorticity equation. *Tellus*, Vol. 2, No. 4, pp. 237–254, 1950.
- [44] 気象庁. 数値予報 60 年誌. 60 pp., 2019. https://www.jma.go.jp/jma/kishou/know/whitep/doc_1-3-2-1/all.pdf.
- [45] Kawabata, T., M. Kunii, K. Bessho, T. Nakazawa, N. Kohno, Y. Honda, and K. Sawada. Reanalysis and reforecast of Typhoon Vera (1959) using a mesoscale four-dimensional variational assimilation system. *J. Meteorol. Soc. Japan. Ser. II*, Vol. 90, No. 4, pp. 467–491, 2012. <https://doi.org/10.2151/jmsj.2012-403>.
- [46] Oizumi, T., K. Saito, J. Ito, T. Kuroda, and L. Duc. Ultra-high-resolution numerical weather prediction with a large domain using the K computer: A case study of the Izu Oshima heavy rainfall event on October 15–16, 2013. *J. Meteorol. Soc. Japan. Ser. II*, Vol. 96, No. 1, pp. 25–54, 2018. <https://doi.org/10.2151/jmsj.2018-006>.
- [47] Oizumi, T., K. Saito, L. Duc, and J. Ito. Ultra-high resolution numerical weather prediction with a large domain using the K computer. Part 2: The case of the Hiroshima heavy rainfall event on August 2014 and dependency of simulated convective cells on model resolutions. *J. Meteorol. Soc. Japan. Ser. II*, Vol. 98, No. 6, pp. 1163–1182, 2020. <https://doi.org/10.2151/jmsj.2020-060>.
- [48] 気象庁. 2030 年に向けた数値予報技術開発重点計画. 41 pp., 2018. https://www.jma.go.jp/jma/press/1810/04b/nwp_strategic_plan_towards_2030_2body.pdf.
- [49] Duc, L., T. Kawabata, K. Saito, and T. Oizumi. Forecasts of the July 2020 Kyushu heavy rain using a 1000-member ensemble Kalman filter. *SOLA*, Vol. 17, pp. 41–47, 2021. <https://doi.org/10.2151/sola.2021-007>.
- [50] 気象庁 数値予報開発センター. 局地解析へのハイブリッド同化手法の導入. 数値予報開発センター年報, pp. 102–105, 2022.

- https://www.jma.go.jp/jma/kishou/books/npdc/r03/npdc_annual_report_r03.pdf.
- [51] 気象庁予報部. 次世代非静力学モデル asuca. 数値予報課報告・別冊第 60 号, 151 pp., 2014.
https://www.jma.go.jp/jma/kishou/books/nwpreport/60/No60_all.pdf.
- [52] Morrison, H. and Coauthors. Confronting the challenge of modeling cloud and precipitation microphysics. *J. Adv. Model Earth Syst.*, Vol. 12, No. 8, p. e2019MS001689, 2020.
https://doi.org/10.1029/2019MS001689.
- [53] Hashino, T. and G. J. Tripoli. The Spectral Ice Habit Prediction System (SHIPS). part III: Description of the ice particle model and the habit-dependent aggregation model. *J. Atmos. Sci.*, Vol. 68, No. 6, pp. 1125–1141, 2011. https://doi.org/10.1175/2011jas3666.1.
- [54] Shima, S., Y. Sato, A. Hashimoto, and R. Misumi. Predicting the morphology of ice particles in deep convection using the super-droplet method: Development and evaluation of SCALE-SDM 0.2.5-2.2.0, -2.2.1, and -2.2.2. *Geosci. Model Dev.*, Vol. 13, No. 9, pp. 4107–4157, 2020.
https://doi.org/10.5194/gmd-13-4107-2020.
- [55] Sato, Y., Y. Miyamoto, and H. Tomita. Large dependency of charge distribution in a tropical cyclone inner core upon aerosol number concentration. *Prog. Earth Planet. Sci.*, Vol. 6, No. 62, 2019. https://doi.org/10.1186/s40645-019-0309-7.
- [56] Sato, Y., S. Hayashi, and A. Hashimoto. Difference in the lightning frequency between the July 2018 heavy rainfall event over central Japan and the 2017 northern Kyushu heavy rainfall event in Japan. *Atmos. Sci. Lett.*, Vol. 23, No. 1, p. e1067, 2022. https://doi.org/10.1002/asl.1067.
- [57] Manabe, S. and R. T. Wetherald. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.*, Vol. 24, No. 3, pp. 241–259, 1967.
https://doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.
- [58] Manabe, S. and R. T. Wetherald. The effects of doubling the CO₂ concentration on the climate of a general circulation model. *J. Atmos. Sci.*, Vol. 32, No. 1, pp. 3–15, 1975.
https://doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2.
- [59] Akiyoshi, H., T. Sugita, H. Kanzawa, and N. Kawamoto. Ozone perturbations in the Arctic summer lower stratosphere as a reflection of NO_x chemistry and planetary scale wave activity. *J. Geophys. Res.*, Vol. 109, No. D3304, 2004. https://doi.org/10.1029/2003JD003632.
- [60] Shibata, K., M. Deushi, T. T. Sekiyama, and H. Yoshimura. Development of an MRI chemical transport model for the study of stratospheric chemistry. *Pap. Meteorol. and Geophys.*, Vol. 55, No. 3+4, pp. 75–119, 2005. https://doi.org/10.2467/mripapers.55.75.
- [61] Friedlingstein, P. and Coauthors. Global carbon budget 2021. *Earth Syst. Sci. Data*, Vol. 14, No. 4, pp. 1917–2005, 2022. https://doi.org/10.5194/essd-14-1917-2022.
- [62] Yashiro, H., Y. Kajikawa, Y. Miyamoto, T. Yamaura, R. Yoshida, and H. Tomita. Resolution dependence of the diurnal cycle of precipitation simulated by a global cloud-system resolving model. *SOLA*, Vol. 12, pp. 272–276, 2016. https://doi.org/10.2151/sola.2016-053.
- [63] Sato, Y., H. Miura, H. Yashiro, D. Goto, T. Takemura, H. Tomita, and T. Nakajima. Unrealistically pristine air in the Arctic produced by current global scale models. *Sci. Rep.*, Vol. 6,

- No. 26561, 2016. <https://doi.org/10.1038/srep26561>.
- [64] Lamarque, J.-F. and Coauthors. The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geosci. Model Dev.*, Vol. 6, No. 1, pp. 179–206, 2013. <https://doi.org/10.5194/gmd-6-179-2013>.
- [65] Shen, L., D. J. Jacob, M. Santillana, X. Wang, and W. Chen. An adaptive method for speeding up the numerical integration of chemical mechanisms in atmospheric chemistry models: application to GEOS-Chem version 12.0.0. *Geosci. Model Dev.*, Vol. 13, No. 5, pp. 2475–2486, 2020. <https://doi.org/10.5194/gmd-13-2475-2020>.
- [66] Kanaya, Y., X. Pan, T. Miyakawa, Y. Komazaki, F. Taketani, I. Uno, and Y. Kondo. Long-term observations of black carbon mass concentrations at Fukue Island, western Japan, during 2009–2015: constraining wet removal rates and emission strengths from East Asia. *Atmos. Chem. Phys.*, Vol. 16, No. 16, pp. 10689–10705, 2016. <https://doi.org/10.5194/acp-16-10689-2016>.
- [67] Kajino, M. and Coauthors. NHM-Chem, the Japan Meteorological Agency's regional meteorology – chemistry model: Model evaluations toward the consistent predictions of the chemical, physical, and optical properties of aerosols. *J. Meteorol. Soc. Japan. Ser. II*, Vol. 97, No. 2, pp. 337–374, 2019. <https://doi.org/10.2151/jmsj.2019-020>.
- [68] Matsui, H. and M. Liu. Importance of supersaturation in Arctic black carbon simulations. *J. Climate*, Vol. 34, No. 19, pp. 7843–7856, 2021. <https://doi.org/10.1175/JCLI-D-20-0994.1>.
- [69] Kajino, M. and Coauthors. Deposition and dispersion of radio-cesium released due to the Fukushima nuclear accident: 2. Sensitivity to aerosol microphysical properties of Cs-bearing microparticles (CsMPs). *J. Geophys. Res.*, Vol. 126, No. 1, p. e2020JD033460, 2021. <https://doi.org/10.1029/2020JD033460>.
- [70] Katata, G., M. Kajino, T. Hiraki, M. Aikawa, T. Kobayashi, and H. Nagai. A method for simple and accurate estimation of fog deposition in a mountain forest using a meteorological model. *J. Geophys. Res.*, Vol. 116, No. D20, p. D20102, 2011. <https://doi.org/10.1029/2010JD015552>.
- [71] Matsui, H., D. S. Hamilton, and N. M. Mahowald. Black carbon radiative effects highly sensitive to emitted particle size when resolving mixing-state diversity. *Nat. Commun.*, Vol. 9, No. 3446, 2018. <https://doi.org/10.1038/s41467-018-05635-1>.
- [72] Talagrand, O. and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, Vol. 113, No. 478, pp. 1311–1328, 1987. <https://doi.org/10.1002/qj.49711347812>.
- [73] Rabier, F. and Z. Liu. Variational data assimilation: theory and overview. *Proc. Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*. pp. 29–43, Reading, UK, 2003. ECMWF. <https://www.ecmwf.net/sites/default/files/elibrary/2003/11805-variational-data-assimiltion-theory-and-overview.pdf>.
- [74] Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte

- Carlo methods to forecast error statistics. *J. Geophys. Res.*, Vol. 99, No. C5, pp. 10143–10162, 1994. <https://doi.org/10.1029/94JC00572>.
- [75] Hunt, B. R., E. J. Kostelich, and I. Szunyogh. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, Vol. 230, No. 1–2, pp. 112–126, 2007. <https://doi.org/10.1016/j.physd.2006.11.008>.
- [76] Kalnay, E., H. Li, T. Miyoshi, S. C. Yang, and J. Ballabrera-Poy. 4-D-Var or ensemble Kalman filter? *Tellus A Dyn. Meteorol. Oceanogr.*, Vol. 59, No. 5, pp. 758–773, 2007. <https://doi.org/10.1111/j.1600-0870.2007.00261.x>.
- [77] R. Potthast, A. Walter, and A. Rhodin. A localized adaptive particle filter within an operational nwp framework. *Mon. Wea. Rev.*, Vol. 147, No. 1, pp. 345–362, 2019. <https://doi.org/10.1175/MWR-D-18-0028.1>.
- [78] Sato, M. and Coauthors. Co-design for A64FX manycore processor and "Fugaku". In *SC '20: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–15, online, 2020. IEEE. <https://doi.org/10.1109/SC41405.2020.00051>.
- [79] Nakano, M., H. Yashiro, C. Kodama, and H. Tomita. Single precision in the dynamical core of a nonhydrostatic global atmospheric model: Evaluation using a baroclinic wave test case. *Mon. Wea. Rev.*, Vol. 146, No. 2, pp. 409–416, 2018. <https://doi.org/10.1175/MWR-D-17-0257.1>.
- [80] Lang, S. T. K. and Coauthors. More accuracy with less precision. *Quart. J. Roy. Meteor. Soc.*, Vol. 147, No. 741, pp. 4358–4370, 2021. <https://doi.org/10.1002/qj.4181>.
- [81] Yashiro, H., M. Terai, R. Yoshida, S. Iga, K. Minami, and H. Tomita. Performance analysis and optimization of Nonhydrostatic ICosahedral Atmospheric Model (NICAM) on the K computer and TSUBAME2.5. In *PASC '16: Proceedings of the Platform for Advanced Scientific Computing Conference*, No. 3, pp. 1–8, Lausanne, Switzerland, 2016. Association for Computing Machinery. <https://doi.org/10.1145/2929908.2929911>.
- [82] Lawrence, B. N. and Coauthors. Crossing the chasm: how to develop weather and climate models for next generation computers? *Geosci. Model Dev.*, Vol. 11, No. 5, pp. 1799–1821, 2018. <https://doi.org/10.5194/gmd-11-1799-2018>.
- [83] Edwards, H. C., C. R. Trott, and D. Sunderland. Kokkos: Enabling manycore performance portability through polymorphic memory access patterns. *J. Parallel Distrib. Comput.*, Vol. 74, No. 12, pp. 3202–3216, 2014. <https://doi.org/10.1016/j.jpdc.2014.07.003>.
- [84] Bertagna, L., M. Deakin, O. Guba, D. Sunderland, A. M. Bradley, I. K. Tezaur, M. A. Taylor, and A. G. Salinger. HOMEXX 1.0: a performance-portable atmospheric dynamical core for the Energy Exascale Earth System Model. *Geosci. Model Dev.*, Vol. 12, No. 4, pp. 1423–1441, 2019. <https://doi.org/10.5194/gmd-12-1423-2019>.
- [85] Fuhrer, O. and Coauthors. Enabling performance portability for FV3/xSHIELD using a Python-based domain-specific language. *ESCAPE-2 Final Dissemination Workshop*. online, 2021. https://www.hpc-escape2.eu/sites/default/files/2021-09/20210703_fuo_ESiWACEDisseminationWorkshop.pdf.

- [86] Ben-Nun, T. and Coauthors. Productive performance engineering for weather and climate modeling with Python. arXiv, 2022. <https://doi.org/10.48550/ARXIV.2205.04148>.
- [87] Schneider, T., N. Jeevanjee, and R. Socolow. Accelerating progress in climate science. *Phys. Today*, Vol. 74, No. 6, pp. 44–51, 2021. <https://doi.org/10.1063/PT.3.4772>.
- [88] Sridhar, A. and Coauthors. Large-eddy simulations with ClimateMachine v0.2.0: a new open-source code for atmospheric simulations on GPUs and CPUs. *Geosci. Model Dev.*, Vol. 15, No. 15, pp. 6259–6284, 2022. <https://doi.org/10.5194/gmd-15-6259-2022>.
- [89] Deconinck, W. and Coauthors. Atlas : A library for numerical weather prediction and climate modelling. *Comput. Phys. Commun.*, Vol. 220, pp. 188–204, 2017. <https://doi.org/10.1016/j.cpc.2017.07.006>.
- [90] Adams, S. V. and Coauthors. LFRic: Meeting the challenges of scalability and performance portability in Weather and Climate models. *J. Parallel Distrib. Comput.*, Vol. 132, pp. 383–396, 2019. <https://doi.org/10.1016/j.jpdc.2019.02.007>.
- [91] Matsuoka, D., M. Nakano, D. Sugiyama, and S. Uchida. Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model. *Prog. Earth Planet. Sci.*, Vol. 5, No. 80, 2018. <https://doi.org/10.1186/s40645-018-0245-y>.
- [92] Kurth, T. and Coauthors. Exascale deep learning for climate analytics. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC ’18)*, Piscataway, NJ, 2018. IEEE Press. <https://doi.org/10.1109/SC.2018.00054>.
- [93] Farrell, S. and Coauthors. MLPerfTM HPC: A holistic benchmark suite for scientific machine learning on HPC systems. In *2021 IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC)*, pp. 33–45, St. Louis, MO, USA, 2021. IEEE. <https://doi.org/10.48550/arXiv.2110.11466>.
- [94] Baño-Medina, J., R. Manzanas, and J. M. Gutiérrez. Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geosci. Model Dev.*, Vol. 13, No. 4, pp. 2109–2124, 2020. <https://doi.org/10.5194/gmd-13-2109-2020>.
- [95] Serifi, A., T. Günther, and N. Ban. Spatio-temporal downscaling of climate data using convolutional and error-predicting neural networks. In G. Wang, editor, *Front. Climate*, Sec. Predictions and Projections, pp. 1–15. april 2021. <https://doi.org/10.3389/fclim.2021.656479>.
- [96] Vaughan, A., W. Tebbutt, J. S. Hosking, and R. E. Turner. Convolutional conditional neural processes for local climate downscaling. *Geosci. Model Dev.*, Vol. 15, No. 1, pp. 251–268, 2022. <https://doi.org/10.5194/gmd-15-251-2022>.
- [97] Seifert, A. and S. Rasp. Potential and limitations of machine learning for modeling warm-rain cloud microphysical processes. *J. Adv. Model. Earth Syst.*, Vol. 12, No. 12, p. e2020MS002301, 2020. <https://doi.org/10.1029/2020MS002301>.
- [98] Grundner, A., T. Beucler, P. Gentile, F. Iglesias-Suarez, M. A. Giorgetta, and V. Eyring. Deep learning based cloud cover parameterization for ICON. arXiv, 2021.

<https://doi.org/10.48550/ARXIV.2112.11317>.

- [99] Chevallier, F., J.-J. Morcrette, F. Chéruy, and N. A. Scott. Use of a neural-network-based long-wave radiative-transfer scheme in the ECMWF atmospheric model. *Quart. J. Roy. Meteor. Soc.*, Vol. 126, No. 563, pp. 761–776, 2000. <https://doi.org/10.1002/qj.49712656318>.
- [100] Rasp, S., M. S. Pritchard, and P. Gentine. Deep learning to represent subgrid processes in climate models. *Proc. Natl Acad. Sci.*, Vol. 115, No. 39, pp. 9684–9689, 2018. <https://doi.org/10.1073/pnas.1810286115>.
- [101] Yuval, J., P. A. O’Gorman, and C. N. Hill. Use of neural networks for stable, accurate and physically consistent parameterization of subgrid atmospheric processes with good performance at reduced precision. *Geophys. Res. Lett.*, Vol. 48, No. 6, p. e2020GL091363, 2021. <https://doi.org/10.1029/2020GL091363>.
- [102] Ashworth, M., G. D. Riley, A. Attwood, and J. Mawer. First steps in porting the LFRic weather and climate model to the FPGAs of the EuroExa architecture. *Sci. Program.*, Vol. 2019, No. 7807860, 2019. <https://doi.org/10.1155/2019/7807860>.

4.11 宇宙・天文

本節では宇宙・天文分野における 2032 年頃に解決すべき科学的課題についてその課題の内容とアプリケーション・計算機資源について詳細を述べる。

4.11.1 自己重力 N 体シミュレーションによる宇宙構造形成の解明

研究概要 2021 年 12 月 24 日、ジェイムズ・ウェップ宇宙望遠鏡がアメリカ航空宇宙局により打ち上げられた。日本においても、国立天文台のすばる望遠鏡に搭載された超広視野主焦点カメラ、Hyper Suprime-Cam によって一辺約 100 億光年にわたる領域に存在する大規模測光天体観測データが得られつつある。2023 年にすばる望遠鏡に搭載され、運用が開始される予定の主焦点超広視野分光観測器 Prime Focus Spectrograph では同様の領域の大規模分光観測が計画されており、ジェイムズ・ウェップ宇宙望遠鏡とのさまざまなシナジーが期待されている。2020 年代後半までに、これらの大規模観測によって特に宇宙最初の星や銀河（初代星、初代銀河）の観測が飛躍的に進むと期待されている。過去 20 年にわたり、宇宙論的な枠組みの中でこれら天体の形成や進化過程は理論的に研究してきたが、これら天体がこれまで観測された直接的な証拠ではなく、その全貌は謎に包まれている。宇宙最初の星や銀河はどこでどのように生まれ、どのように現在の銀河へと進化し、宇宙全体の進化にどのような役割を担ってきたのか？これらは解決するべき課題として残されている。

そのため、大規模自己重力 N 体シミュレーションをベースにしたこれらの大規模観測データと直接比較可能な天体の大規模模擬カタログが必要である。シミュレーションに使用できる粒子数は有限であり、

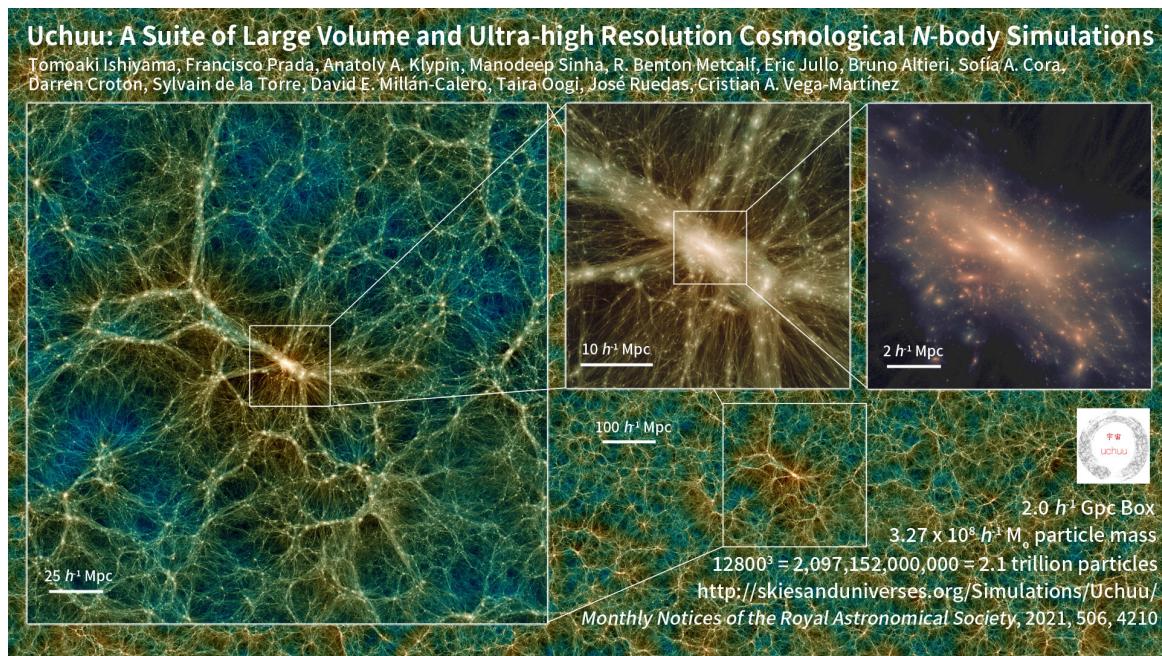


図 4.41 宇宙のダークマター階層構造 [1]。

用いるシミュレーションコードの性能やスーパーコンピュータによって制限される。最先端のシミュレーションにおいても、初代星や初代銀河のような小さな天体（銀河系と比較して、約10万から1000万分の1の質量）を解像する場合は、宇宙空間の極一部の空間体積のみしか表現できていない。また大規模観測と比較できる広大な空間体積を表現する場合は、小さな天体を解像できず、銀河の形成はある程度成長してからスタートするとしてモデル化されており、大きな不定性が存在していた。銀河は合体を繰り返して成長するため、本質的には最小の天体を解像してシミュレーションする必要がある。宇宙初期の星形成スケールから、現在の大規模構造スケールを同時に分解するには、銀河ひとつを10億粒子程度で分解したうえで、 $10^5 \sim 10^6$ 個の銀河形成を追う必要がある。したがって1000兆粒子以上の計算が必要である。現状、「富岳」では計算機能力が不足しており、次のフラグシップマシンで実現されると考えられる。

計算科学的特徴と要求性能 計算量 $O(N \log N)$ の自己重力 N 体計算である。粒子間の相互作用演算が計算の主要な部分であり、「富岳」では30～50%を占める。要求性能表に出した数字は、「富岳」上でチューニングし、フルシステム上で10%以上の実行効率で動作するコード、GREEMの詳細プロファイラによる性能解析に基づいている。

4.11.2 自己重力流体シミュレーションによる銀河形成過程の解明

銀河の形成と進化は、重力、流体力学、輻射、星形成、超新星爆発など恒星進化に伴うエネルギーフィードバックや、それに付随する元素合成による化学進化など、さまざまな物理及び過程が複雑に絡み合って進む。また、恒星スケール ($\sim pc$) から潮汐相互作用スケール (数10Mpc) までの幅広いスケールがその形成と進化に影響を与える。このように、銀河の形成と進化は、本質的に多階層・多物理の非線形現象であり、主に数値シミュレーションで行われる極めてチャレンジングな研究テーマである。

1990年代初頭にはじまった3次元の銀河形成シミュレーションでは、ダークマターの位相空間分布をサンプリングして N 体粒子で表現し、またバリオンの分布を同様にサンプリングしてガス粒子として表現し、それらの自己重力流体としての進化を解いた。最初のシミュレーションから程なくして、星形成や超新星爆発などを実際の恒星のスケールを分解せずに扱うサブグリッドモデルが構築され、それらがシミュレーションに取り込まれた。しかしながら、このようなシミュレーションでは当初円盤銀河の形成に成功しなかった。その理由は、このようなサブグリッドモデルでは超新星爆発で星形成領域に充分高温のガスを生み出すことができなかったからであった。この影響を解決するべく、サブグリッドモデルの範囲で超新星爆発によるシェル構造発達のシミュレーション結果を膨張モデルに取り込むことなどにより、その結果観測的に得られている銀河のハロー質量-バリオン比に代表される統計的性質を再現できるようになり、現在ではさまざまな形態の銀河の形成と進化の研究が行われている（図4.42）。また、とくにこの10年ぐらいは、数値計算方法の洗練化にも大きな労力が払われてきており、重力多体系および圧縮性流体の計算方法にも多数の進展があった。

このようなシミュレーションは一定の成功を収めているが、銀河形成の本質的な部分であるエネルギーフィードバックをすべてサブグリッドモデルに押し込んでいる。そのため、銀河形成の本質的な理解には、さらなるシミュレーションの進化が必要である。従来のシミュレーションの本質的な問題は恒星スケールを直接分解していないからである。しかし、我々の住む天の川銀河は数千億個の恒星からなる系で

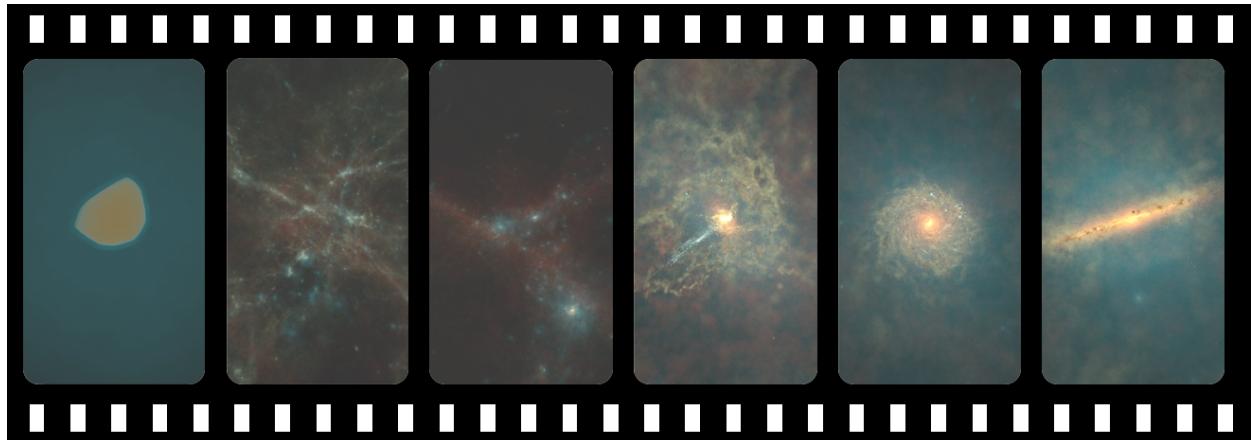


図 4.42 銀河形成シミュレーションの例。左から右に向けて現在に近づいていく。現在の我々の宇宙では、銀河より小さな天体が最初に形成され、それらが合体成長を繰り返して銀河を形成する。このシミュレーションでは宇宙空間の一部を切り出し、宇宙が現在のサイズのおよそ $1/100$ だったときを初期条件にシミュレーションを行っている。事前の荒い分解能のシミュレーションで銀河ができるところに当たりをつけてその領域だけにバリオンをおいている。それ以外の領域の物質はすべて暗黒物質で代表させている。青く淡い分布は暗黒物質の分布を、黄色い分布は星形成の材料となる冷たいガス、赤い分布は高温のガスを表す。白い点状の分布は恒星粒子を表す。動画は <https://www.youtube.com/watch?v=Rdd9KAUcvgQ>

あり、これは今までの最高分解能のシミュレーションでも恒星スケールを直接分解するのは極めて困難である。

近年では、欧州の Gaia 衛星が、そして日本で計画されている JASMINE 衛星が銀河を構成する恒星の膨大かつ精密な位相空間分布を、そして APOGEE、GALAH などが恒星の化学進化の情報をもたらすことで、天の川銀河の円盤の形成と進化の様子が読み解かれつつある。これらと直接比較し、恒星系としての銀河の成り立ちを理解するためにも、また、私達の住む太陽系の起源を理解する上でも、銀河形成シミュレーションで恒星スケールの直接分解することには大きな意義がある。

このような背景の元、我々は富岳の高い演算性能を用いて、恒星スケールまで直接分解した銀河形成シミュレーションを実現するための取り組みを行っている。同様の恒星スケールを分解しようという取り組みは、海外のグループでも取り組まれているが、矮小銀河程度が限界であり、まだどこも天の川銀河スケールでの実現はできていない。おそらくすぐにこの目標は達成できると考えられるが、このような計算は一つ実行するために世界的なスーパーコンピュータを数週間-数ヶ月使う必要がある。

今はまだこのスケールの計算を行うのは限界に近いが、今後 5 年、10 年後には、これらのスケールのシミュレーションが当たり前のものとなり、それによって、多数のパラメータスタディが可能になるだろう。また、孤立した銀河だけではなく、いすれば銀河群や銀河団など銀河同士の相互作用が重要になるスケールでも、このような大規模シミュレーションが実現可能になるだろう。これによって、宇宙に広がる様々な形態を持つ銀河がどのような物理過程により形成されてきたかという本質を理解することができるようになると期待される。

銀河形成のシミュレーション方法には、粒子法、適合格子を用いたメッシュ法、移動メッシュ法の三つ方法が用いられているが、ここでは粒子法に基づき計算資源などを見積もることとする。計算の主要な部

分は、 N 体法による自己重力相互作用、および流体力学相互作用部分になるため、これらの部分についてだけ見積もる。重力相互作用には Tree 法を、流体相互作用には、従来よく用いられてきた SPH 法の場合と、空間高次精度の圧縮性流体解法の一つである CPHSF 法を用いた場合についても見積もりを行う。

質量分解能として $1 M_{\odot}$ (流体要素が $1 M_{\odot}$ 、ダークマターが $\sim 10 M_{\odot}$)、空間分解能として 1 pc を仮定する。天の川銀河程度の銀河のスケールと、銀河群、銀河団での 1 ステップあたりの演算量(浮動小数点演算数)を表 4.4 に示す。ここで重力相互作用の 1 相互作用あたりの演算数を 30、相互作用リストの長さを 10^4 に取った。また、SPH 法による流体相互作用では 1 相互作用あたりの演算数を 300 とし、相互作用リストの長さを 100 に取った。CPHSF 法においては流体相互作用を空間精度が 2-4 次程度のところでの 1 相互作用あたりの演算数 150 と、相互作用リストの長さ 150 を採用した。

表 4.4 演算量

	銀河	銀河群	銀河団
天体質量	$\sim 10^{12} M_{\odot}$	$\sim 10^{13} M_{\odot}$	$\sim 10^{15} M_{\odot}$
粒子数	$\sim 10^{11}$	$\sim 10^{12}$	$\sim 10^{14}$
重力演算量 (N 体)/step	60 Peta	600 Peta	60 Exa
流体演算量 (SPH)/step	3 Peta	30 Peta	3 Exa
流体演算量 (CPHSF)/step	10 Peta	100 Peta	10 Exa
総演算量 (N 体/SPH)/step	63 Peta	630 Peta	63 Exa
総演算量 (N 体/CPHSF)/step	70 Peta	700 Peta	70 Exa

銀河形成では、宇宙年齢 138 億年分のシミュレーションを行いたい。時間刻み幅は基本的に超新星爆発領域の高温高密度領域の CFL 条件で決まる。音速を 1000 km s^{-1} 、スケールを 1 pc 、そして CFL 係数として $1/3$ を採用すると典型的な時間刻み幅は

$$dt = \frac{1}{3} \frac{1 \text{ pc}}{1000 \text{ km s}^{-1}} \sim 1000 \text{ yr} \quad (4.31)$$

になる。したがって、シミュレーションを終了までに 10^7 ステップが必要になる。

計算の効率としてピーク性能の 10% を仮定する。我々の富岳での取り組みではカーネルレベルでピーク性能の 20-30% まで出せたが、全体性能は 5% 程度だったため、この程度の計算効率を採用した。また、計算の性質上、いわゆるメモリリミットではないが、計算効率を上げるためにには、より多くの粒子を計算ノード内に持たせる必要がある。後述するが、富岳のノードあたりのメモリ実装量 (32GB/ノード) では高効率を実現するには不十分だと感じた。

さて、以上を持って宇宙年齢の計算を行ったときに必要となる計算資源量を見積もると表 4.5 のようになる(大差ないので CPHSF の場合のみを示す)。銀河形成シミュレーションで総演算量は 7 Yotta (10^{24}) 程度と見積もられる。ここで、 $2-20 \text{ E}(10^{18}) \text{ flops}$ のシステムを想定すると、年間の総演算数は $6 \times 10^{25} - 6 \times 10^{26}$ になるため、全系を用いた場合およそ 1 ヶ月から数日、10% のシステムを用いた場合には 1 年から 1 ヶ月を要することとなる。計算効率の向上、計算アルゴリズムのさらなる洗練化により、シミュレーションにかかる時間を更に短くすることは十分可能であると考えられる。しかしながら、個別の恒星を分解した銀河群、銀河団のシミュレーションの実現は当面不可能であろう。やや質量分解能

を落とした計算が可能だと思われる。

銀河形成シミュレーションでは一般に B/F 値への要求は低い。これは相対的に利用するデータ量に対して演算数が十分大きくなるからである。1000 粒子程度で相互作用粒子リストを共有すると、典型的な重力計算の場合の B/F 値は 10^{-3} 程度、流体計算の場合の BF は $10^{-3} \sim 10^{-4}$ 程度と見積もられる。

一方でストレージに対する要求は大きい。銀河の形成が計算領域のなかでどのように進むか事前に予測はできない。そのため、頻繁に出力してポストプロセスで調べることが必要になる。星形成の母体となる分子雲の力学的な時間スケールを分解するためにはおよそ 10^5 年おきに出力する必要があり、1 粒子あたり倍精度変数 100 程度出力するとすると、一回あたりあたり 100 TB、全部で 10 Exabyte 程度の出力になる。計算時間の 1% を出力に使うとすると、およそ 10 – 100 TB/s のストレージ性能が必要となる。出力頻度の調整、出力データの絞り込みをしてもポストプロセスに必要なデータは 1 Exabyte ぐらいは必要になるだろう。

表 4.5 総演算量

	銀河	銀河群	銀河団
総演算量 (N 体/CPHSF)	7 Yotta	70 Yotta	7000 Yotta

ここからは少し我々の富岳での取り組むについて述べよう。最小時間刻み幅の 1000 年で 10^7 ステップの計算を行うのは、不可能ではないとはいえるが計算量が多すぎる。このような領域は計算領域全体のごく一部であり、平均的な構造の進化を解くのに必要な時間刻み幅は 10^5 年程度である。天文シミュレーションではしばしば独立時間刻み法が用いられるが、この方法では $\mathcal{O}(N)$ の処理が大半を占めることになり、また、全体通信が発生するためそれほど効率的にはならない。そこで、計算中で最も短い時間刻み幅で進化するだろう超新星爆発の影響を受ける領域をハミルトン分割の手法を応用して切り出し、その小規模領域を全体の相互作用の計算と並列で計算を進めるという手法を開発した。実装においては MPI_Comm_Split を利用し全体計算を受け持つプロセスと、小規模で多ステップ計算する領域を受け持つ部分に分けている。また、爆発の影響を受ける領域の推定には深層学習を用いる。

この実装に成功したが、ノード数がかつてないほど多い富岳で実行するのにあたって困難に直面した。富岳全系は 15.8 ノードで構成されている。各ノードはメモリが 32 GB 搭載されていて、これを 48 コアで共有する。全系利用を試みると、ノードあたりのアプリケーション利用可能メモリが非常に少なく、我々の実装で性能を出すのが困難であった。富岳ポータルで提供されている“ユーザープログラムが利用可能なメモリ量の見積り”を元に、実際にユーザが利用可能なノードあたりのメモリ量を示したのが図 4.43 である。ノードあたり 32 GB 搭載されているが、5 GB 程をシステムが利用し、残りを MPI とユーザプログラムで分け合うことになる。通常の MPI 並列の場合（図 4.43 左パネル）、全系利用時には 1 ノードあたり 21 GB 程度しか使えない（青と緑の線が交わるところ）。この手の計算では性能を出すためにはノードあたりに可能な限り多くの粒子を乗せることが必要であり、これは非常に不利である。この状況は MPI_Comm_Split を用いて MPI プロセスの分割を行った場合には更に悲惨になる（図 4.43 右パネル）。MPI プロセス数が内部で倍になるため、ユーザプログラムが利用可能メモリ量はおよそ 4 GB になってしまう。そのため最終的に計算効率を犠牲にして（実行性能をおよそ半分にして）ノードあたり 1 MPI プロセスで実行している。新しいシミュレーションモデルの実装にはある程度メモリ的猶予が必要

である。

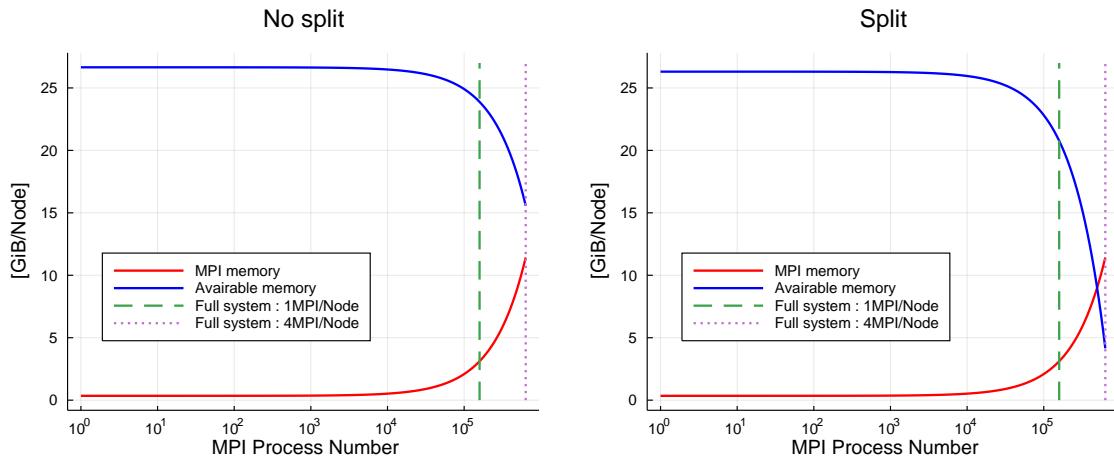


図 4.43

ノードあたりの利用可能メモリ量。1 ノードあたり 32GB だが、5GB はシステムが利用するので、残りの 27GB の配分について。横軸は MPI プロセス数で、縦軸はノードあたりのメモリ (GB)。左のパネルは MPI_Comm_split を用いない場合、右のパネルは MPI_Comm_split を用いて MPI プロセスを内部で 2 つに分解した場合。MPI の利用するメモリが赤の実線、ユーザが利用可能なメモリが青の実線で描かれている。緑の縦のダッシュ線との交わるところが富岳全系で 1 ノードあたり 1 MPI プロセスのとき、また、紫の縦の点線と交わるところが 1 ノードあたり 4 MPI プロセスのときの数字になる。

4.11.3 球状星団におけるコンパクト連星の形成

2015 年に人類史上初めて重力波の検出が達成された。その重力波は 2 つのブラックホール（連星ブラックホール）の合体によって放たれたものであった。これは連星ブラックホールの合体の初発見でもあった。その発見は「連星ブラックホールの起源はなにか？」という新たな問題を生み出し、その起源が世界的な論争となっている。有力な説は、太陽より数十倍質量の大きい恒星がブラックホールとなり、そのようなブラックホールが「なんらかの機構」によってもう 1 つのブラックホールと距離の小さい連星を成し、その連星ブラックホールが重力波放出によって軌道エネルギーを失って合体する、というものである。この「なんらかの機構」として提案されている説が 2 つある。1 つの説は、星形成のときから連星であった 2 つの恒星が、連星内相互作用の末、そのまま連星ブラックホールとなった、というものである（孤立連星起源）。もう 1 つの説は、球状星団内の異なる場所で形成された 2 つのブラックホールが、お互いを力学的に捕獲し合い、連星ブラックホールとなった、というものである（球状星団起源）。現在、この 2 つの説が拮抗しており、未だ決着がついていない。本研究の目的は、後者の説を作業仮説とし、球状星団内で形成される連星ブラックホールを理論的に予測して観測と比較し、後者の説の是非を明らかにすることである。

球状星団内の連星ブラックホール形成を理論的に扱うのには、星団内にあるすべての恒星の運動を追う必要がある。その手法として最も単純であるが、最も強力な方法は、重力 N 体シミュレーションと呼ばれる方法である。この方法では、星団内のすべての恒星の軌道をニュートンの万有引力の法則に従う運

動方程式に基いて追跡する。この重力 N 体シミュレーションの基礎方程式は以下のように表すことができる。

$$\frac{d^2\vec{r}_i}{dt^2} = \sum_{j \neq i}^N \frac{Gm_j}{|\vec{r}_j - \vec{r}_i|^3} (\vec{r}_j - \vec{r}_i) \quad (4.32)$$

ここで、 G は重力定数、 \vec{r}_i と \vec{r}_j はそれぞれ i と j 番目の恒星の位置ベクトル、 m_j は j 番目の恒星の質量である。この方程式を直接用いる場合、計算量は $O(N^2)$ である。実際には、遠くの複数の恒星を 1 つの恒星と見なして重力計算を行う、いわゆるツリー法を用いて重力計算を行うため、計算量は $O(N \log N)$ となる。従来、ツリー法を球状星団の重力 N 体シミュレーションに適用しつつ、恒星の軌道計算の精度を保つことは困難であった。しかし、球状星団用の重力 N 体シミュレーションコード PeTar がこの困難を解決した。後述する必要資源の見積りは PeTar の使用実績を基にしたものである。

球状星団は、恒星数 10^6 個、恒星質量密度 $10^3 M_\odot \text{pc}^{-3}$ (M_\odot は太陽質量)、年齢 100 億年という性質を持つ。しかしこれらの性質を揃えた現実的な球状星団の重力 N 体シミュレーションは未だ達成されていない。既存のシミュレーションは、恒星数が少ない、恒星数密度が小さい、追跡している時間が 100 億年より短かい、というものだけである。現実的な球状星団の重力 N 体シミュレーションが未だ達成されていないのは、そのような重力 N 体シミュレーションが大きな計算量を必要とするからである。同時に扱う恒星数が 10^6 個のため、系の規模としては大きくなり。それにもかかわらず、必要とする計算量が大きいのは、追跡するべきタイムステップ数が多いからである。球状星団内の恒星の自由落下時間がおよそ 10^5 年なので、100 億年を追跡するには 10^5 自由落下時間を追う必要があり、必要とされるタイムステップ数が非常に多くなる。

球状星団の重力 N 体シミュレーションは、系の規模が小さく計算の並列性が高くないため、計算資源が増加しても、1つ1つの計算時間は短かくなりにくい。しかし、以下の 2 点において、計算資源の増加による恩恵を受ける。1つ目は、球状星団の重力 N 体シミュレーションが必要とする計算資源が相対的に小さくなることである。すなわち、計算の待時間がより少なくてすむようになることである。これはシミュレーションの CPU 時間ではなく、柱時計時間に直結する。従って、現実的な時間で現実的な球状星団の重力 N 体シミュレーションを実行できるようになりうる。未だ達成されていないことができるようになるため、これは大きなインパクトである。2つ目は、パラメータサーベイができるようになることである。宇宙には多くの球状星団が存在し、それぞれ、金属量、銀河潮汐場の強さなどが異なる。特に金属量は恒星が残すブラックホールの質量に強く影響するため、様々な金属量について調べる必要がある。

いくつかのパラメータにおいて、現実的な球状星団の重力 N 体シミュレーションを実行して、形成される連星ブラックホールを理論的に導出できれば、現在稼動中の地上重力波望遠鏡 (LIGO, Virgo, KAGRA など) によって観測されている連星ブラックホールと直接比較することが可能となる。それは連星ブラックホールの起源を明らかにすることにつながるであろう。また、2030 年代に稼動する予定の地上重力波望遠鏡 (Einstein Telescope, Cosmic Explorer) や宇宙重力波望遠鏡 (LISA, TianQin, Taiji, DECIGO) によって発見される連星ブラックホールを予測することにもつながるであろう。特に、地上重力波望遠鏡と宇宙重力波望遠鏡による同一の連星ブラックホールの観測は、連星ブラックホールの離心率を決定できる。球状星団起源の連星ブラックホールが有限の離心率を持つ一方で、孤立連星起源の連星ブラックホールは離心率ゼロである。そのため、球状星団起源の連星ブラックホールの離心率分布を予言しておくこと

は、宇宙重力波望遠鏡時代における連星ブラックホールの起源解明に決定的な役割を果すはずである。

上述のように球状星団の重力 N 体シミュレーションは、恒星数 10^6 個と扱うデータ量が小さい一方で、1 タイムステップ当たりの計算量と必要なタイムステップ数が多いために、必要とする計算量が大きい。従って、メモリアクセスの速さが全体の計算の律速にはなりにくい問題である。

扱う球状星団は恒星数 10^6 個でおよそ $5 \times 10^5 M_{\odot}$ 、平均質量密度が $10^3 M_{\odot} \text{pc}^{-3}$ なので、恒星の自由落下時間は 2.3×10^5 年である。球状星団の年齢はおよそ 100 億年であり、自由落下時間当たりに必要なタイムステップ数は 1.3×10^5 タイムステップで、全部で 5.7×10^9 タイムステップ必要となる。タイムステップ当たりの粒子間相互作用数は 2.7×10^{10} で 1 相互作用当たりの演算数は 40 であるので、必要とする総演算数は 6200EFLOP となる。想定する計算資源は 100 ノードで理論ピーク性能 1PFLOPS である。これはノード当たりの理論ピーク性能が 10TFLOPS であることを意味し、富岳のノード当たりの理論ピーク性能が 2.5TFLOPS なので、将来的には十分ありうる値である。これまでの富岳での使用実績によると、PeTar の実行性能は理論ピーク性能のおよそ 10% である。従って、1 ケースあたりの計算時間は 17000 時間となる。ケース数は、金属量が 1, 0.1, 0.01 太陽金属量の 3 通り、銀河潮汐場の強さは銀河中心からの距離を 3kpc, 10kpc の 2 通りとすることで、全部で 6 通りとなる。この計算は多くのメモリ量、メモリバンド幅、ストレージ量を必要としないので、ここでは記述しない。

4.11.4 Beyond-Boltzmann: ニュートリノ輻射流体計算による超新星爆発機構の解明

重力崩壊型超新星爆発とは、太陽の約 10 倍以上重たい星がその一生の最期に起こす爆発的な天体现象であり、銀河に匹敵するほどに輝く宇宙で最も明るい現象の一つである。天体现象としての興味深さに加えて、様々な重い星がどのような超新星となり、どのような重元素を作り出し周辺に撒き散らすのか、中性子星あるいはブラックホールなどの極限天体を生み出すのかを突き止めることは、銀河・宇宙スケールにわたる物質階層の形成や進化を探るうえでも欠かせない。この爆発現象は一般相対性理論や流体力学などの巨視的物理過程と強い相互作用や弱い相互作用などの微視的物理過程とが複雑に絡まり合う、高度に非線形な現象であるため、メカニズムを解明するためには数値シミュレーションが必須である。爆発のエンジンとなる中心コア付近は非常に高温かつ高密度状態になっており、ニュートリノによるエネルギー輸送と組成の変化がその爆発メカニズムにおいて重要な役割を果たす。それゆえに、我々のグループはこれまで、「京」や「富岳」をはじめとする国内の大規模並列計算機を用いて、多次元ボルツマンニュートリノ輻射流体力学に基づく超新星爆発シミュレーションを実行してきた。これらの大規模シミュレーションを用いた研究によって爆発メカニズムの理解が深まってきた一方で、超新星のような高密度環境でのみ起こるニュートリノの量子的な効果によって、爆発メカニズムやその後の元素合成、観測されるニュートリノ信号が影響を受ける可能性が近年指摘されはじめてきている。これらの現象を探究するための数値シミュレーションは、量子効果を取り扱うために数値計算コストが従来に比べて格段に増えるが、今後の計算機の発展によってこれを取り扱うことが現実的になってくる。このような、古典運動論であるボルツマン方程式を超えた量子運動論的な方程式を取り扱った、Beyond-Boltzmann 超新星シミュレーションを行い、超新星爆発のメカニズムを明らかにしていくことが主たる目標である。

研究背景 超新星爆発のメカニズムを数値シミュレーションで検証することは天体物理学において 60 年近い歴史を持つ難問題である。その爆発メカニズムとして有力な仮説がニュートリノ加熱メカニズムである。すなわち、進化した大質量星の鉄コアが重力崩壊を起こして潰れ、中心が原子核密度になるとバウンス衝撃波を形成し、一度エネルギー損失によって停滞した衝撃波が中心にできる原始中性子星から照射されるニュートリノに加熱されることで復活し、星の外層を吹き飛ばすというシナリオである。この問題の複雑さは様々な物理過程を複合した非線形問題であることが原因だが、その中でも特に高温・高密度環境におけるニュートリノの反応と輸送を扱う多次元ニュートリノ輻射方程式を解くことは長年の課題となっている。これは従来的にはボルツマン方程式によって記述され、量子効果はニュートリノ反応にのみ組み込まれているが、古典論的な輸送であっても大きな計算資源が必要であり、歴史的には近似計算法を導入し簡単化したモデルから始まって徐々にその精密さと複雑さを増していった。すなわち、これまでのスーパーコンピュータの世代ごとに、その時点での計算機資源を最大限使い切るような計算によって段階的にモデルが精密化していき、そのたびに爆発メカニズムの新たな様相が発見されてきた。21 世紀初頭になって球対称における第一原理計算が初めて可能となり球対称では超新星爆発が起きないことが明白となった。これは、重力崩壊のうち中心部が跳ね返ることにより打ち上げられた衝撃波が、物質の降着により妨げられて常に停滞してしまうためである。その後、近似手法に基づくニュートリノ輻射輸送による空間 2 次元（軸対称）計算が可能となり、流体力学的不安定性とニュートリノ加熱（ニュートリノ吸収による熱エネルギー増加）の組み合わせが働いて、停滞した衝撃波が復活して爆発するシナリオが有力となった。さらに、京コンピュータの登場でようやく空間 3 次元計算が行われるようになり、近似手法のもとではあるが、軸対称の制限により得られる爆発とは異なる 3 次元的ダイナミクスを通じた爆発メカニズムがついに明らかになった。

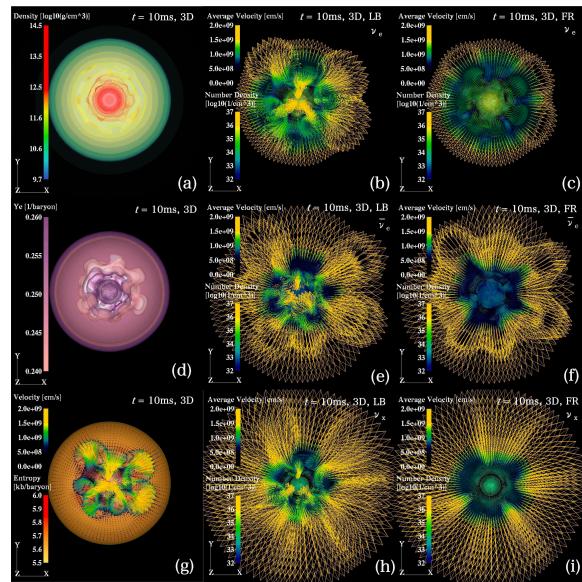


図 4.44 空間 3 次元での第一原理的シミュレーションにおける物質密度、電子分率、速度、エントロピー、およびニュートリノ平均速度の分布 [2]。

一方で、我々は緻密な計算による高信頼度を追求し第一原理的シミュレーションを行うことを目指して

いる。中心コアの上層部から物質が降着するなかで、停滞した衝撃波が復活に転ずるか否かは閾値付近にあり精密計算が欠かせない。これまで広く行われてきた多次元数値計算は近似手法によるものであるため、その信頼度が不明である。また、球対称計算の研究段階において近似計算では決着がつかず、爆発の可否についての結論が第一原理的シミュレーションから得られたことから、多次元計算においても第一原理的なシミュレーションが必要なのである。このために、我々は一般相対論的時空のなかでボルツマン方程式を解くことを目指しており、特に独自に開発したニュートリノ輸送のためのボルツマン方程式ソルバーは成功をおさめ、「京」や「富岳」を用いた空間2次元軸対称超新星シミュレーションによって数々のマイルストーン的な成果を創出してきた。近年では空間3次元での第一原理的シミュレーションの実現にも成功し、現在は「富岳」で長時間計算を行っている。一般相対論的效果や微視的素過程などの面でのモデル精密化も進んでおり、第一原理計算に最も近い超新星の理論モデルを構築している。これは緻密度において世界で最も進んだ取り組みである。本プロジェクトで得られる成果は、系統的なモデル計算へと拡げることで、超新星ダイナミクスや中性子星/ブラックホール形成といった他の天体现象を含む包括的な天体物理学の理解を進め、ニュートリノ、重力波、電磁波、重元素生成など、マルチメッセンジャー天文学の中心を担う。さらに、開発する計算コードはガンマ線バーストやマグネター形成等への拡張も視野にいれた天体シミュレーションの分野を横断しうるものであり、大規模疎行列の解法など計算科学全般との関わりも強く、多くの分野との共同研究が進んでいる。

こうして、多次元ボルツマンニュートリノ輸送に基づいた超新星の数値モデルが多数創出され、超新星爆発コア内のニュートリノ輻射場の特徴が徐々に明らかになってきた一方で、これまでの古典論的な枠組みでは取り扱いきれない量子論的效果の重要性も増してきた。我々の第一原理的な計算によってのみ得られるニュートリノ分布関数のビックデータを解析したところ、新規の量子物理現象の可能性が発見されたのである。超新星のようなニュートリノ密度の非常に高い環境においては、ニュートリノ同士の相互作用によってニュートリノ固有状態の分散関係が変調を受け、特殊なタイプのニュートリノ振動、すなわちニュートリノ集団振動を引き起こす。これは量子的な効果だが、ニュートリノ集団振動が起りうるかどうかは古典的な計算の結果をもとに、量子効果を摂動論的に考えることで判断することができる。以前から超新星内でニュートリノ集団振動が起こる可能性は指摘されていたが、実際に起こるかどうかは古典的な計算、すなわちボルツマン方程式を解いた結果として得られる、運動量空間内のニュートリノ輻射場の分布が必要となる。これまでには近似的なニュートリノ輻射輸送ばかりだったために必要な情報は不足しており、その詳細な解析は進んでいなかったが、我々は、ボルツマンニュートリノ輸送により、ニュートリノ運動量空間の詳細な解析を行い、ニュートリノ集団振動が超新星コア内で一般的に起こることをはじめて明らかにした。これは超新星爆発の理論モデルの構築において重大な発見であり、爆発メカニズムや元素合成の結果、さらに超新星ニュートリノの観測シグナルにも直接的、かつ大きな影響を与える可能性がある。それゆえに、世界中の多くの研究グループの注目を集めており、今後超新星におけるニュートリノ集団振動の研究は飛躍的に発展すると期待される。加えて、ニュートリノ集団振動は連星中性子星合体現象でも一般的に起こることが近年示されたこともあり、素粒子物理学や高エネルギー天体物理学分野の研究者たち全体が、その理論モデルの構築に取り組み始めている。

ニュートリノ集団振動を取り入れるには、古典的運動論を超えた、量子運動論方程式 (Quantum Kinetic Equation, QKE) によってニュートリノ輸送を記述する Beyond-Boltzmann 超新星シミュレーションを行う必要がある。また、これは超新星爆発のみならず、上述のように連星中性子星合体の理論モデルを

構築する上で非常に重要な要素の一つであり、ニュートリノ・重力波・電磁放射の観測を通じたマルチメッセンジャー天文学の発展著しい現在において、本課題を推進する優先度は非常に高い。一方で、QKE の数値計算はボルツマン輻射輸送計算と比較しても桁違いに高い計算コストが必要となる。それゆえ、QKE に基づくニュートリノ輻射流体シミュレーションの実装には、計算機資源の拡充を前提とした長期的な展望が必要となる。

研究手法 超新星爆発を記述するためには、流体力学と重力、ニュートリノ輸送を同時に解く必要がある。従来の研究ではこれらをニュートン的に扱い、連続の式とオイラー方程式、エネルギー方程式で流体運動を扱い、重力のためにポアソン方程式を解き、ニュートリノ輸送のためにボルツマン方程式を解いていた。流体運動の式は高精度衝撃波捕捉スキームを用いることで衝撃波を含む流れを精度よく解きつつ、ポアソン方程式は予め修正不完全コレスキーフ分解前処理付き共役勾配法で解いておいた逆行列をメモリに保存しておくことで解く。ボルツマン方程式は位相空間を直接離散化する discrete ordinate 法 (S_N 法) を利用しており、安定化双共役勾配法によって半陰的に解いている。また、ボルツマン方程式を一般的な計量テンソルのもとで定式化することで曲がった時空や加速度系に対応できるようにし、中心の原始中性子星に固定した座標系と流体静止系との行き來を前提にした特殊な座標離散化を用いることで、相対論的な効果を完全に取り入れることに成功している。また、今後の計画のために流体力学ソルバーと重力ソルバーも一般相対論化を進めており、近く古典論の範囲では第一原理的な超新星シミュレーションが可能となる。本課題においては、ニュートリノ輸送をさらに QKE で取り扱うことを目指す。

Beyond-Boltzmann 超新星シミュレーションを行うために実装する QKE は、ボルツマン方程式と良く似た方程式であるが、量子的な効果を取り入れるために、主に 2 つの違いがある。1 つ目は解くべき従属変数が古典的な位相空間分布関数から、量子的ないわゆる密度行列へと拡張される。これによって異なるフレーバーのニュートリノの間の量子的な重なりや混合状態を記述することで、集団振動を含むニュートリノ輸送が取り扱える。2 つ目は、ニュートリノ振動の直接の原因となる Hamiltonian potential の項が付け加わる。Hamiltonian potential には質量項、物質との相互作用項、そして自己相互作用項の三種類の寄与があるが、特に最後の項が集団振動を引き起こす上で重要な役割を担う。これらにより、解くべき方程式の数と方程式間の密な結合が増えことで、計算量と扱うデータ量は飛躍的に増大する。我々のグループでは QKE を解く数値解法については、これまで本課題のボルツマン輻射輸送計算で用いていた S_N 法を基礎として開発を進めている。一方で、モンテカルロ法などの粒子法による QKE ソルバーの開発も進めており、こちらも着実に成果を上げている。それぞれのメリットを活かしながら、相補的にコード開発を進める予定である。

QKE によるニュートリノ輻射輸送は非常に計算コストが高いため、機械学習技術の応用も視野に入れている。ニュートリノ振動の波長スケールは天体现象の巨視的なスケールと比べて桁違いに小さく、特に超新星の場合は 7 桁以上も差がある。それゆえ、これを数値計算において解像することは非常に難しく、たとえエクサスケールの計算機を完全に専有できたとしても、超新星のようなグローバル計算はそのままでは不可能である。現在、我々は局所的な QKE 計算の結果に基づいて、サブグリッドモデルの構築を目指しており、これを用いることで巨視的なシミュレーションの実装に取り組んでいる。一方で、ニュートリノ集団振動は強い非線形現象であるため、サブグリッドモデルの構築において解析的な取り扱いでは限界がある。そこで現在では、機械学習、特に深層ニューラルネットワークの強力な非線形関数再現能力を

応用することを模索している。これまでに多次元ニュートリノ分布関数のビックデータ解析の機械学習による研究実績があり、この経験を活かして、粗いグリッドの結果を入力として局所的な QKE 計算の結果を出力する深層ニューラルネットワークを構成し、これをサブグリッドモデルとして巨視的シミュレーションに結合させることで、小スケール非線形現象を超新星のグローバルなダイナミクスに反映させることを目指す。

ロードマップ 現在及び至近の計算機資源を利用して、古典論的な運動論に基づく超新星シミュレーションを第一原理的なものに大きく近づける。現在までに、「京」コンピュータによって S_N 法による 2 次元(軸対称)空間におけるボルツマン輻射流体計算が可能になり、「富岳」コンピュータによってさらに 3 次元空間におけるシミュレーションを実行中である。また、流体力学ソルバーとボルツマン方程式ソルバーの一般相対論化を完了しており、一般相対論的な重力を取り扱う数値相対論ソルバーも開発完了が近い。2020 年代前半においては、これらの一般相対論的效果を取り入れた古典論第一原理計算を行いながら、計算機資源の増大とともに段階的に 3 次元計算のモデル数を増やしていく。

ここからさらに QKE を用いた Beyond-Boltzmann 超新星シミュレーションを実装していくためには、ニュートリノの自由度増加と相互作用の拡充に対応するため、将来的な計算資源の飛躍的な拡充が必要不可欠である。しかしながら、ニュートリノ振動の波長を解像するような第一原理計算は、現在の 3 次元ボルツマン計算よりも 10 衍以上高い計算負荷がかかる見積もりのため、単純にスケールしたのではエクサスケールの計算機資源でも太刀打ちできない。そのため、本質的な部分をえぐり出して緻密度を失わないように方程式系を扱っていく必要がある。我々が上述のように計画しているサブグリッドモデルの構築はまさにその必要性に応えるものであり、2020 年代の後半から 2030 年代前半にかけてはこのサブグリッドモデルを結合させた一般相対論的な 3 次元シミュレーションの実装を目指す。サブグリッドモデルの活用により、現在の計算コストの 100 倍程度であれば科学的な成果を創出できる見込みである。

加えて、本プロジェクトの QKE 実装に向けた取り組みと並行して、核物質状態方程式、弱い相互作用、磁気流体などといったその他の物理素過程を精密化し、第一原理的な超新星の理論モデル構築を目指していく。これを完遂できれば、本計算によって爆発メカニズムを確立するとともに、ニュートリノや重力波の理論モデルを超新星分野やマルチメッセンジャー天文学分野周辺の研究者に提供することで、高エネルギー宇宙物理学分野全体の理解を飛躍的に深めていくことができる。さらに、数値シミュレーションで得られる密度行列としてのニュートリノ輻射場のビッグデータは素粒子物理学分野等との共同研究へ活用していく予定である。

計算科学的特徴 最も計算資源が必要となる量子的ニュートリノ輻射輸送の部分は、6 次元位相空間上で 3 フレーバー \times 3 フレーバーの密度行列の振る舞いを粒子と反粒子の二つに分けて解く巨大な計算である。特にニュートリノ反応に関わる項は、古典的なボルツマン方程式の時と同様にエネルギーごとのタイムスケールが大きく異なり、また高温高密度環境での反応率が理論的に求めづらい、取り扱いの難しいものである。これらを解くための方程式を線形化したのちの計算問題は、ニュートリノ反応に起因する密ブロック行列を対角上に持ち、ニュートリノの空間移流による非ゼロ要素が線状に現れるパターンを持つ大規模疎行列の解法に帰着する。空間 3 次元のメッシュ数を N_r 、 N_θ 、 N_ϕ 、ニュートリノ運動量空間のエネルギー・角度メッシュ数を M_ϵ 、 M_{θ_ν} 、 M_{ϕ_ν} とすると、密ブロック行列はサイズ $M_\nu \times M_\nu$ ($M_\nu = M_\epsilon M_{\theta_\nu} M_{\phi_\nu}$) であり、これらが空間グリッド数 $N_s = N_r N_\theta N_\phi$ 個並ぶパターンにより、全体の行

列サイズは $N_{\text{total}} \times N_{\text{total}}$ ($N_{\text{total}} = N_s M_\nu$) となる。少なく見積もって空間グリッド数を 100^3 、ニュートリノ空間グリッド数を 10^3 としてもベクトルは 10^9 次元あり、疎行列は $10^9 \times 10^9$ となる。また、密度行列は全部で 18 個の成分を持つので、方程式を 18 セット解くことになる。扱う行列・ベクトルのサイズが巨大であるため、行列を反復法によりいかに効率よく解くかが課題である。これらを分散してメモリ保持しつつ、行列要素の計算、行列解法、ベクトルのアップデートを行うことを超大規模並列において実現しなければならない。この大規模疎行列解法の技術は京や富岳において確立しているが、これを量子力学的記述へと拡張することがチャレンジとなる。

また、ニュートリノ振動を取り扱うためのサブグリッドモデルの構築の際には局所化した QKE を解くが、ここでは振動スケールが物質との相互作用の平均自由行程に比べて桁違いに小さいため、陽的な時間発展を行う。この問題の並列化は難しいものではないが、位相空間全体の高次精度化が必要な計算であり、どのように演算性能を高めていけばいいかが今後の課題である。加えて、サブグリッドモデルを深層ニューラルネットワークを用いて構築する場合、この学習には GPU を利用する。一度学習したあとの feed forward にはそれほどの計算量は必要ないので、サブグリッドモデルを超新星シミュレーションに組み込む際には GPU 計算で学習した結果を CPU に読み込ませ、CPU 上で feed forward を行う。この機械学習サブグリッドモデルを上述の QKE ソルバーに組み込む手法は新奇なものため、こちらも開発上の課題となっている。

要求性能表の算出根拠 まず、これまで多く計算して性能評価の確立している「京」コンピュータでのボルツマン超新星シミュレーションの性能とスケーリングを評価し、それに基づいて QKE 超新星シミュレーションの要求性能を推定する。

「京」コンピュータにおける空間 2 次元での計算量は、解像度が空間 384×128 、ニュートリノ空間 $20 \times 10 \times 6$ でニュートリノ種類数 $N_f = 3$ の場合、バウンス後 400 ms までの時間発展 ($\sim 10^7$ ステップ) に必要な資源は 1 モデルあたり 3M ノード時間であった。実行性能はピーク性能の 10% 程度であり、全浮動小数点演算量は $1.4 \times 10^{20} \text{ flop}$ 、1 時間ステップあたりでは $4 \times 10^{13} \text{ flop}$ である。1 ステップあたりの演算量の内訳は、行列解法の計算と反応率の計算でおよそ半々である。前者については密ブロック行列の反復法による解法を想定して演算量は $M_{\theta_\nu}^2 M_{\phi_\nu}^2 M_\epsilon N_s N_f n_{\text{op}} n_{\text{ite}}$ である。後者のニュートリノ反応の計算において、エネルギーと角度が変わる反応については演算量は 1 回あたり $M_{\theta_\nu}^2 M_{\phi_\nu}^2 M_\epsilon N_s m_{\text{op}}$ であり、 n_{op} 、 n_{ite} 、 m_{op} はそれぞれ行列解法の演算回数、反復回数、反応率計算での演算回数を表すパラメータである。「富岳」コンピュータを利用した 3 次元計算においては、空間次元をひとつ増やしてグリッド数が約 100 倍増え、演算量も 100 倍増えたものを実行中である。

本計画の最終目標段階においては、配位空間 $384 \times 128 \times 256$ 、運動量空間 $20 \times 20 \times 12$ グリッドの解像度を目指す。また、QKE 計算においては 18 成分の密度行列を取り扱うので、上記の評価からニュートリノ種類数 $N_f = 3$ の代わりに $N_f = 18$ とみなして評価すればよい。さらに、Hamiltonian potential の計算にもある程度の負荷があるので、これらを総合して現行の計算と同じ解像度にスケールした場合の約 10 倍程度の計算負荷がかかると見積もれる。このことから、1 ステップの演算量は 3 次元計算で $1.6 \times 10^{18} \text{ flop}$ (1.6 Eflop) であり、超新星シミュレーション 1 モデルにつき合計 $\sim 2 \times 10^7 \text{ Eflop}$ の演算量が必要となる。一方で、本課題においては 1 モデルあたり 1 年程度かけて計算し、初期条件を様々に変えた 3 モデルを行うことを想定した値を総演算量とした。メモリ量は、密ブロック行列要素を格納するの

に必要な量として評価した。必要なメモリ転送量は演算量のほぼ半分を占める行列ベクトル積によって決まり、メモリ量/演算量=4を仮定して 4×10^7 EBとなる。ストレージ量は、6次元密度行列を保持する量が1スナップショットで9TBであり、 10^4 スナップショット分で90PBである。データ書き出し速度は 10^3 ステップにつき1スナップショットの出力で、演算時間の0.03%程度とした。通信については、ノード数を 10^5 にした時に3次元空間での境界面のグリッドにおいて隣接のニュートリノ分布をやり取りするデータ量を見積もり、通信時間は演算時間に対して0.3%程度を要求した。

相対論的粒子計算による超高エネルギー現象と粒子加速機構の探究

研究背景：研究の目標、これまでの研究 宇宙から飛来する非熱的な高エネルギー粒子の起源は、その発見から100年以上経ったいまもなお、宇宙物理学の大きな謎のひとつとして挙げられる。X線観測による超新星残骸(SNR)衝撃波でのシンクロトロン放射の発見は、衝撃波近傍において電子が相対論的エネルギーまで加速されていることを強く示唆しており、現在では無衝突衝撃波は高エネルギー粒子の加速源として広く認識されている。地球に降り注ぐ宇宙線の大部分を構成する陽子については、近年、IceCubeプロジェクトによって高エネルギーニュートリノの起源がその他の電磁波観測とともに同定された。相対論的ジェットを伴うブレーザーが高エネルギーニュートリノ源、すなわち宇宙線陽子の有力な加速源であることが世界で初めて示された。また、近年の γ 線観測によって銀河系内の超新星残骸衝撃波においてもPeVまで陽子が加速されうることが示された。近年の観測的進展から、宇宙の爆発的現象に伴って荷電粒子(陽子、電子)が超高エネルギーを獲得していることが示唆されている。

宇宙にあまねく荷電粒子がどのような過程を経てエネルギーを獲得するかを理解することを目的に、プラズマの第一原理に基づくプラズマ粒子シミュレーションが行われてきた。スーパーコンピューターを用いて衝撃波や磁気リコネクションなどの時間発展を第一原理的に解くもので、電磁流体シミュレーションでは記述できない、非熱的粒子の生成や分布関数の形状などの議論が可能である。第一原理計算を行うことで、どのような物理過程で超高エネルギー粒子が生成されるかを理解し、得られる分布関数と観測を比較することで、その起源に迫ることを目標にしている。

今回の研究の手法: 基礎方程式とアルゴリズム 無衝突プラズマの第一原理であるブラソフ方程式(無衝突ボルツマン方程式)とマクスウェル方程式を基礎方程式としている。ブラソフ方程式で記述される分布関数の6次元位相空間内の時間発展は、その位相空間要素を代表する超粒子の運動をラグランジュ的に解くことで記述される。空間内に張られた格子上に定義される電磁場の時間発展は、マクスウェル方程式を有限差分法で解くことで記述される。粒子の空間分布に基づいて重み付けされた格子上の電流密度を計算する、particle-in-cell(PIC)法によって粒子と場が自己無同着にカップルされる。これらに加え、マクスウェル方程式の解法として陰解法を採用しており、その解法にCG法が使われる。電荷密度の時間発展と電流密度の発散が矛盾しないように、電流計算にアルゴリズム上の工夫が必要になっている。並列化は領域分割したMPIとOpenMPのハイブリッド並列。また、超並列計算に必要なロードバランス維持のためのロードバランサの実装がされている。

ロードマップ PICシミュレーションでは、現実の陽子・電子質量比(1836)を多次元計算で採用することが難しく、通常は数100程度の質量比の下で計算を行い、質量比に注意しながら物理的エッセンスを議

論する。「京」では SNR 衝撃波を想定した、非相対論的高マッハ数衝撃波の 3 次元 PIC シミュレーションが初めて実施可能になった(図 4.45)。「富岳」では、より現実的な質量比(400 程度)、幅広い衝撃波速度(光速の数%からローレンツ因子 100 程度まで)の状況下で、長時間計算を行うことで電子の長時間加速を追尾することが可能になってきている。安定した電子エネルギースペクトルを得ることで、多波長の電磁波観測のスペクトルと直接比較することが可能な時代になってきた。また、2 次元計算に限定するれば、さらなる長時間計算を行うことで、陽子の初期加速を議論できる。「富岳」後継機では、この陽子加速の長時間発展をも追尾することで、宇宙線陽子加速

問題の解決を視野に入れる。陽子加速機構を明らかにすることで、SNR 衝撃波近傍でのガンマ線観測、高エネルギーニュートリノ観測(IceCube-Gen2)などの観測に対して、エネルギースペクトルなどの情報を提供できる。また近年、米国の NIF などに代表される高出力レーザー施設を用いた実験室宇宙物理分野が成果を上げている。レーザーを使ってプラズマの噴出をコントロールすることで、室内で衝撃波や磁気リコネクションが再現されるようになってきた。実験室宇宙物理分野と大規模 PIC シミュレーション研究の連携も今後の大きな柱となる。

計算科学的な特徴 PIC シミュレーションは、ラグランジュ的手法とオイラー的手法のアルゴリズムが組み合わさった手法である。計算コストの大部分は粒子の空間分布からグリッド上での電流密度を計算する部分にある。電磁場配列への局所的なアクセスのため、粒子ソート、キャッシュチューニングが必要である。これら工夫が電流計算ルーチンの SIMD 化の促進を決めるため、計算実行効率に大きく影響する。これら最適化は「京」で有効であったが、「富岳」においてはレジスタスピルが大量に発生するため、そのためのデータのメモリ間の移動がボトルネックとなっている。その結果、実行性能が「京」で 15 %程度あったものが、「富岳」では数%までに低下した。ループを分割するなどの工夫によって 6 %程度まで改善できたが、キャッシュを有効に使えないためか、それ以上の向上が見込めなかった。レジスタが少ないという設計が一般ユーザーに対して最適化を困難にしている。

必要資源の見積もり 陽子加速を追尾するために、計算規模としては、格子数が $100000 \times 1000 \times 1000$ 点数、粒子数として 10^{13} 個が必要になる。総メモリ量は 1PB 程度。ストレージは 1 スナップショット 1PB / ケース程度で、5 ケースで 50PB。1 ケース 10 日(240 時間)で 100 万時間ステップ / ケース。これらを実現するための要求性能は、富岳での実績から外装して、実行効率 10 %程度を仮定とすると、要求演算性能として 12PFLOPS、メモリバンド幅として 6PB/s を必要とする。

4.11.5 銀河構造を考慮した高解像度星間媒質進化シミュレーションによる星形成過程の解明

研究背景：研究の目標、これまでの研究 銀河の基本構成要素である星がどのように形成されたのかは、銀河進化を司る物質循環を理解する上で解決すべき問い合わせである。星は星間媒質から誕生し、その質量が十分大きい場合は、短い寿命内に強い紫外線放射や星風を放ち、さらに寿命を終えるときに超新星爆発を起こし、周囲の星間媒質に強い擾乱を与えつつ星間媒質に還っていく。また、惑星は星形成に伴って作られるので、近年の観測技術の発達により発見されている多種多様な特徴をもつ系外惑星系が、どのような環境下で作られるのかを明らかにする上でも、星形成過程の理解が不可欠である。このように星形成は銀河

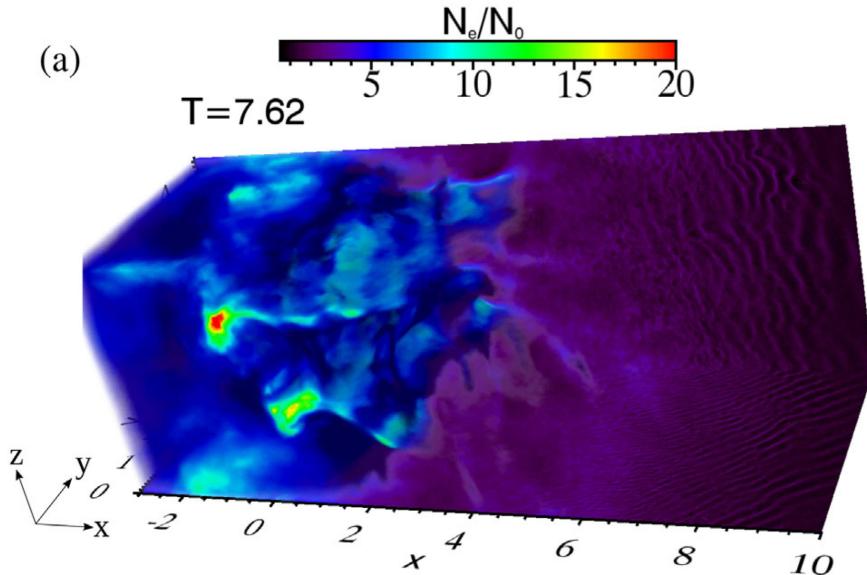


図 4.45 非相対論的高マッハ数衝撃波の 3 次元 PIC シミュレーション。色は電子密度分布を表す。[3]

における物質循環の駆動力であるとともに、銀河から惑星形成ひいては生命誕生をつなぐ、極めて重要な物理過程である。

星形成の場である分子雲は、絶対温度 10 度で高密度なガス雲であり、より温かく低密度な中性水素原子で構成される原子ガスが集積し、熱的不安定性を経て暴走冷却・凝縮することで形成される。この原子ガス集積の主要なメカニズムの一つは超新星爆発である。形成された分子雲の中で生まれた大質量星は、星風や紫外線放射・超新星爆発といったフィードバックを引き起こし、母体の分子雲や近傍の分子雲を破壊するとともに、時間が経ち弱まったフィードバックは、原子ガスを集積して新たな分子雲を作り新たな星形成を引き起こす。

このように大スケール (\sim 数 100 pc) の分子雲の形成・破壊と、小スケール (< 0.1 pc) で起こる星形成は、互いが互いの駆動源となっているため、両者を適切に扱ったシミュレーションが不可欠である。長さスケールで 1000 倍以上という大きなダイナミックレンジをもつ現象であるとともに、流体力学と自己重力を始めとして、磁場や輻射輸送・化学反応など多様な過程が複雑に絡み合う非常に複雑な過程であるため、数値流体シミュレーションが有用な研究手段として用いられてきた。しかしこれまでは、計算資源の問題から、いくつかのスケールごとの問題に分離してシミュレーションがおこなわれてきた。また、適合格子細分化法を用いて広いダイナミックレンジをもつシミュレーションも行われてはいるが、計算量を減らすために自己重力が効く高密度部に集中して細かい格子を配置しているため、低密度部が適切に分解できないという問題があった。分子雲形成は、大雑把には乱流状態にあるガスが冷却して密度が上昇していく過程なので、密度が低い段階から十分に空間分解することが極めて重要である。

今回の研究の手法：基礎方程式とアルゴリズム 米国 Princeton 大学および東北大学を中心とする国際協力で開発中の公開磁気流体シミュレーションコード Athena++ を用いる。本コードは宇宙物理学向けの多様な物理過程をサポートした最先端のコードであり、世界中の宇宙物理学シミュレーションで多数の利用実績がある。また、ポスト京萌芽課題と富岳での利用実績もある。

本研究でおこなう銀河局所シミュレーションでは、銀河内のある場所を中心に、曲率を無視して局所直交系を張る。動径方向を x 軸に、方位角方向を y 軸に、鉛直方向を z 軸に取る。 x 方向には銀河回転による速度シアの考慮に入れた境界条件を課す(シアリング・ボックス近似)。

ほとんどの計算時間を占める磁気流体計算と自己重力計算は、それぞれ時間空間 2 次精度近似リーマン解法(HLLD 法)と多重格子法を用いる。分子雲形成を含む星間媒質の熱進化を追うためには、ガスの加熱・冷却過程を考慮する必要がある。演算量を削減するため、詳細な熱・化学平衡計算の結果をテーブル化してエネルギー方程式のソース項として入れる。観測と直接比較する上では化学反応を流体力学と同時に解くことが重要であるが、ある程度信頼の置ける化学反応を採用した場合には、磁気流体計算の 10 倍以上の時間がかかる一方で、流体の熱力学への影響は限定的であることが知られている。そのため、今回はダイナミクスに注目するために化学反応は解かない。

分子雲コアが重力収縮して原始星を形成する時間・空間スケールは、本研究でのシミュレーションにおいても大きく乖離しているので、分子雲コアの重力収縮過程は追わずに、シンク粒子という流体の吸い込み口を実装する。また星間媒質の相転移を駆動するために、星によるフィードバック(星風と電離領域・超新星爆発)の実装が必須である。このうち超新星爆発についてはすでに実装している。星風と電離領域はすでに先行研究でモデル化されているので、適切なモデルを採用する。

ロードマップ 星は、幾何学的に細長いフィラメント状の分子雲で形成されるということが、*Herschel*衛星による近傍分子雲の観測によって示唆され、フィラメントでの星形成パラダイムが、星形成研究の中で世界的な潮流となっている。近年では遠方の大質量星形成領域や、代表的な系外銀河である大小ラゼラン雲でもフィラメントが星形成において極めて重要な役割を果たしていることが明らかとなっている。このフィラメントの太さは 0.1 pc またはそれ以下と非常に細いことが知られており、フィラメント形成を明らかにするためには大きなダイナミックレンジをもったシミュレーションが必須である。これまで幾つかのフィラメント形成メカニズムが提案してきたが、それらは理想化された問題設定上のシミュレーションや理論モデルがベースとなっており、現実的な条件下でどのフィラメント形成過程が有効なのかは不明である。フィラメントを分解できる解像度で現実的な分子雲形成・進化を明らかにする本研究によって、フィラメント形成メカニズム、そして星形成におけるフィラメント状分子雲の役割への知見が得られると期待される。また、これまで観測と理論シミュレーションの比較は、その多くが事例研究にとどまっているが、計算領域内に形成される無数のフィラメントを使って、これまで十分なサンプル数でできなかった統計的な議論が可能となる。

残念ながら本研究の分解能 0.03 pc は、星形成の直接の母体である分子雲コア内部の乱流などの微細な構造を分解するのには十分ではないが、本研究のシミュレーションデータを元に、多数の興味ある領域をさらに高解像度で再計算することが可能である(ズーム・イン・シミュレーション)。これにより、分子雲コアの統計的性質が明らかにできれば、星の初期質量関数の理解という星形成の最終目標への大きな知見が得られるとともに、連星形成過程や惑星形成の初期条件に関する知見が得られる。

計算科学的な特徴 大部分の計算時間を占める磁気流体計算は、局所的な隣接する格子のデータを参照すればよい。「富岳」では磁気流体計算は少なくとも数 10 万ノードまで完璧なウィークスケーリングが得られている。

多重格子法(自己重力解法)では、格子幅を順次変更していく、あるサイクルを実行することで解の収

束効率を最適化する。ノードが受け持つ領域サイズより格子幅が小さいサイクルでは、磁気流体計算と同様に局所的なデータを参照する。ノード領域サイズより格子幅が大きいサイクルでは、全体通信が必要となるが、全格子の情報を通信するわけではない。

「富岳」では、HBM が高いバンド幅を持っている反面、キャッシュのレイテンシが大きく、バンド幅も大きくはないため、メモリ律速の磁気流体計算では性能を出すのが難しい。また、レジスタ数が少ないため命令レベルでの並列性が低いのも性能低下に影響している。さらに、コンパイラの最適化性能が悪く、SIMD を有効にするために手動でのオンライン展開やループ成形を余儀なくされている。

必要資源の見積もり 今回採用する一様格子は格子分布が単純であるため、超並列計算に適している。ここでは、XC50 Skylake でのプロファイル情報を元に、大部分の計算時間を占める磁気流体計算を元に必要資源を見積もる。ここでは演算量については磁気流体力学の性能を単純に 2 倍して自己重力の寄与を考慮する。

- 格子数の見積もり

まず必要な格子数を見積もる。計算領域サイズは、 x 方向と y 方向には同じ長さ 1 kpc をとし、鉛直方向の成層構造と質量流出を解くために z 方向には x, y 方向より 4 倍大きな領域をとる。つまり体積は $(0.5 \text{ kpc}^2 \times 4 \text{ kpc})$ である。 $z > 0.5 \text{ kpc}$ は上空にいくに従い粗くする。格子数としては上空は無視できるので、以下では、見積もりでは $(512 \text{ pc})^3$ の立方体として見積もる。分子雲形成において数値的に収束した結果を得るために、暴走冷却するガスの冷却長を分解する必要がある。そのために必要な格子数は $N_{\text{tot}} = 4.4 \times 10^{12} (\Delta x / 0.03 \text{ pc})^{-3}$ である。

- 時間ステップ数の見積もり

必要な時間幅は、局所計算の経験から冷却を陽的に解く場合には高密度部の冷却時間で決まり、おおよそ $\Delta t \sim 10^{-4} \text{ Myr}$ 程度である。超新星爆発によって生じる高温電離ガス (10^7 K) の音速がセル幅 (0.03 pc) を横切る時間に CFL 数をかけた値も同等である。粗い解像度で行われた先行研究によると銀河円盤が定常に達するのは、おおよそ 100 Myr なので、それから 10 倍の時間計算する場合、必要な総時間ステップ数は $N_{\text{step}} = 1000 \text{ Myr} / \Delta t \sim 10^7$ である。

- 演算量とデータ転送量

磁気流体シミュレーションでは、近似リーマン解法 HLLD を用いた場合、1 格子 1 時間ステップあたり約 3500 演算が必要である。したがって、全浮動小数点演算数 F_{tot} は、 $F_{\text{tot}} = 2 \times 3500 \times N_{\text{step}} \times N_{\text{cells}} = 3 \times 10^5 \text{ Eflop} (\Delta x / 0.03 \text{ pc})^{-3}$ ここで、ファクタ 2 は自己重力の寄与である。

- 必要なメモリ量

1 格子あたり磁気流体変数の基本量と保存量、そして重力ポテンシャルを保持する必要があり、また各方向の数値流束の保持も必要である。2 段の時間積分に必要な一時的な変数を加味すると、1 セルあたり 1000 Byte 程度必要である。したがって、必要なメモリ量は $M_{\text{tot}} = 10^3 \text{ Byte} \times N_{\text{tot}} = 4.4 \text{ PByte} (\Delta x / 0.03 \text{ pc})^{-3}$ である。

自己重力計算においては、格子幅を変えていくサイクル毎の通信は隣接するノードとの間だけでおこない、磁気流体力学部の通信量よりも小さい。

- ノード数の見積もり

必要なノード数の最小値はノード当たりのメモリ量に依存する。「京」で 16GB、「富岳」で 8GB なので、この傾向が続くとして、10GB として見積もる。必要なメモリ量は約 4.4 PByte なので、 $N_{\text{node}} \sim 4.4 \times 10^5$ である。

- 要求演算性能

1 モデルを 1 ヶ月で終えるとすると、 $F_{\text{tot}}/(1 \text{ month}) \sim 10^2 \text{ Pflops}$ で演算する必要がある。この値が理論性能の 10% だとすると、要求理論性能は 10^3 Pflops である。ノード当たりでの要求演算性能は 2.3 TFlops である。

- 要求バンド幅

磁気流体計算は基本的にメモリ律速で、Athena++ の場合には(計算機に合わせて特別な最適化をしていない場合)、Xeon Skylake において B/F が 10 bytes/flop 程度である。ここでは、B/F を 10 bytes/flop として、演算性能 10^2 Pflops を妨げないとすると、要求される最低限のバンド幅は $\underline{10 \times 10^2 = 10^3 \text{ PByte s}^{-1}}$ である。ノード当たりでの要求演算性能は 2.3 TFlops である。

- ノード間メモリ転送量の見積もり

適合格子細分化法を 64^3 のブロック単位でおこなう。全ブロック数は $2^{24} \sim 1.7 \times 10^7$ で、 2^{18} 個のノードで分割すると、ノードあたり平均的に 64 個のブロックを保持する。ノードにおいてブロックを $4 \times 4 \times 4$ の立方体で保持する場合に、ノード間通信量が最小となる。

磁気流体計算の場合は立方体の糊代を隣接するブロックを担当するノードに通信すればよい。面を共有するノードへの通信量が最も大きく、空間二次精度で厚み 2 格子で面積 $((64 \times 4)^2)$ の平板 6 枚を通信するとすると、ノード当たり双方向通信で $\sim 800 \text{ MByte}$ で、シミュレーション全体では $\sim 350 \text{ TByte}$ である。

自己重力計算の場合は、多重格子法での各サイクルでは、隣接ノードとの通信だけでよい。通信量は最も細かい格子のサイクルが最も多く、重力ポテンシャルのみを通信すればよいので磁気流体計算にくらべて無視できる。一方、格子がノードサイズより大きくなったときに集団通信(MPI_Allgather)が必要である。ノード間通信量を減らすため、ノード内でデータを集約した上で、ノード間通信することにすると、 $N_{\text{node}} = 4.4 \times 10^5$ なので、 $\sim 800 \text{ GByte}$ である。

- IO 性能の見積もり

特にリスタートファイルは計算を再開始するために必要な倍精度データが全て含まれ、出力するファイルの中で最も大きな容量をもち、データ量は $\sim 4 \times 10^3 \text{ TB}$ である。これが 30 分で出力できる場合、 $\sim 2 \text{ TB s}^{-1}$ となる。この性能であれば、单精度での 1 スナップショットの出力に約 1.5 分を要する計算になる。

4.11.6 輻射磁気流体計算による太陽恒星ダイナモの探求

研究の目標 本研究の目標は太陽黒点数 11 年周期の謎を解明することである。太陽表面にしばしば現れる強磁場領域である黒点は、11 年の周期を持って変動するが、この周期維持のための物理機構は未だ明らかになっておらず太陽物理学最古、そして最難の問題となっている。この問題は新たな物理の探求・宇宙環境学の推進の二つの側面がある。11 年周期は、乱流と磁場、回転の相互作用の結果として発現すると

考えられている。しかし、なぜ11年になるのか、また、それにともなう蝶形図やヘールの法則といった首尾一貫した法則が出現する理由は理解されていない。特に首尾一貫したルールを司る未知の物理があるのであろう。また、熱対流の速度について観測と数値シミュレーションに隔たりがあり、未だ太陽内部の理解には大きく改善の余地がある。非常に大きなレイノルズ数の磁気乱流という宇宙で頻繁に実現される現象でありながら、地上実験では実証不可能な物理課程のプロトタイプとして広い汎用性をもつ。また、太陽11年周期に伴いフレア・コロナ質量放出と言った爆発現象は、地球を含んだ太陽系環境に直接影響する。11年周期の解明は、これらの現象の根源的理説、長期予測にも関連し、宇宙環境学の推進への貢献も大きい。

これまでの研究 太陽の11年周期を理解するためには、磁場の生成、つまり太陽内部のダイナモを理解する必要がある。太陽内部の外側30%は乱流的な熱対流に覆われており、この乱流と磁場の相互作用により11年周期が維持されていると考えられている。これまでに音速を無限大と仮定するアネラスティック近似 ($\nabla \cdot (\rho_0 \mathbf{v}) = 0$, ρ_0 は背景密度、 \mathbf{v} は流体速度) が広く使われてきたが、全体通信を多く必要とするために解像度の向上に限界が来ていた。そこで、実効的な音速を抑制する音速抑制法では、隣接通信だけで計算を実行することが可能で、少なくとも弱いスケーリングは現代の大きな計算機でも達成することができる。例えば富岳では図4.46に示されるような大規模計算が達成されており、ポスト富岳でもさらなる大規模化が期待される。

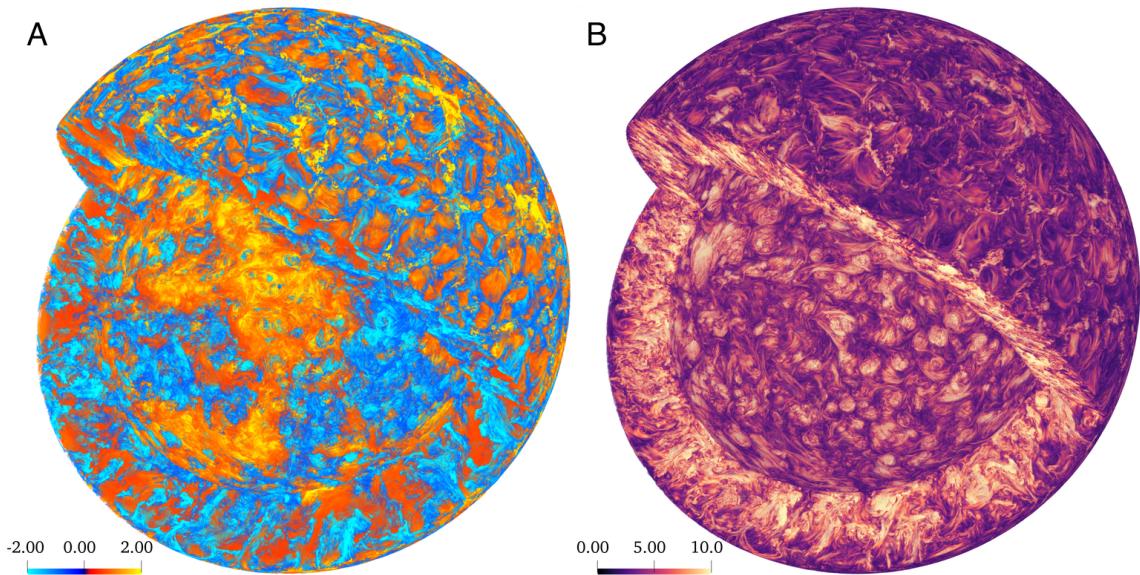


図4.46 富岳で達成した太陽内部熱対流の様子。A. 各高さで規格化したエントロピー B. 磁場強度 [G] を表す [4]。

基礎方程式とアルゴリズム 基礎方程式は、音速抑制法を用いた圧縮性輻射磁気流体方程式である。太陽の自転は遅いのでコリオリ力のみを取り入れれば良い。輻射については、太陽深部では光学的に厚いので拡散近似、太陽表面では、動径方向一次元の輻射輸送方程式を解いて輻射による冷却を見積もる。太陽表面では、OPALなどを用いてテーブル化した状態方程式を利用する。太陽深部では、線形近似をした状態方程式に切り替えることで、非常に小さな擾乱も数値的に正確に扱えるようにする。太陽全球を包括する

ためにインヤン格子を用いる。

ロードマップ 計算資源の増加により

1. 熱対流速度・磁場強度の数値的収束
2. 太陽深部・太陽表面の包括計算

が達成可能となる。それぞれについて簡単に説明する。これまでの計算では、解像度上げれば上げるほど磁場強度は上がり、乱流・平均的流れ場への影響は大きく変わってきた。どこまで小さいスケールを解像すれば現実の太陽の状況を再現できるのか、現時点では確証はないが、数値的収束が一つ目標となるであろう。太陽物理学で目指す数値的収束は、陽的な粘性・磁気拡散を仮定しない状況で、解像度を二倍にしても結果が変わらないということである。実際の太陽の最小の乱流スケールは cm のオーダーと考えられ、100 年スケールの計算機の発展を持ってしてもこのスケールを分解できるとは考えられない。数値的収束に辿り着くことで一つ、「信じられる」数値シミュレーションを達成できると考えられる。また、これまでの富岳の計算で、数値的収束に辿り着くような兆しも見えてきており、ポスト富岳の計算では達成できる可能性が大いにあると考えられる。

一方、小スケールの乱流の分解と同様に重要なのが異なる層の接続である。熱対流は太陽深部では一ヶ月、20 万 km ほど、太陽表面では数分、1000 km と極端に異なった時間スケールを持っている。これまで京や富岳を用いた計算でこの二つの層を包括する努力が続けられており、局所計算では成功している。ポスト富岳の数値資源があれば、全球を包括しながらこの二つの層を接続することができるようになると考えられる。この計算が成功すれば太陽内部での磁場生成から太陽表面での黒点形成までを一つの計算の中で追うことができるようになり、太陽黒点数 11 年周期の謎解明に大きく近づくことが期待される。

計算科学的な特徴 太陽内部物理は、正確な輻射輸送計算や電離を考慮した状態方程式を取り扱う必要があるものの、これらは高速化の手法が提案されている。一方で、一般相対論や自己重力などが必要ないという意味で比較的物理がシンプルで、計算機の大規模化が直接成果につながる面があり、新しいフラグシップを用いて早期に成果の出しやすい分野である。

現代の計算機を使うまでの困難もある。富岳のように少ないレジスタ・比較的小さい L1、L2 キャッシュのシステムでは磁気流体力学のシミュレーションは性能を出すことが難しい傾向にある。磁気流体力学では単純な流体力学やブラソフ方程式などよりも取り扱う変数が多いことがネックになっている。小さなキャッシュには、多くの変数の三次元データを格納しておくことは困難であるが、磁気流体力学の方程式では様々な方程式に複雑に変数が登場するためにメモリのアクセスを適切に整理する必要がある。ソフトウェアパイプラインを使って、キャッシュアクセスを隠蔽することが肝要であるが、少ないレジスタのために一筋縄ではいかない。富岳のようなシステムでの磁気流体力学方程式のプログラミング方法を業界全体として考案していくことは重要であると考えられる。

必要資源の見積もり 太陽全球を包括しつつ、対流層の底から太陽表面までを取り入れるために、インヤン格子で $1024 \times 6144 \times 18432 \times 2$ 格子点ほどが必要となる。太陽表面までを計算すると時間幅は 1 秒ほどとなるので、太陽内部で磁場を作ると考えられる 20 年を計算するには 6 億ステップほどが必要となる。1 格子点を 1 時間ステップアップデートするのに 8000 演算ほどが必要になるとすると計算をするの

に 1×10^{23} 演算ほどが必要となる。富岳の 100 倍ほどの 50EFLOPS のマシンが完成し、その 10% の資源を使い、実行性能はピーク性能の 10% ほどを想定すると 23 日ほどで計算が終了する。1 格子点あたり必要なメモリは 4 KB ほどであるとすると $B/F=0.5$ となりメモリバンド幅は 250 PB/s となる。必要メモリは 1.5 PB である。

プラズマ計算による太陽圏・宇宙空間無衝突衝撃波の研究

研究背景：研究の目標、これまでの研究 太陽フレアやコロナ質量放出などの太陽活動に伴い生成される高エネルギー荷電粒子は地球周辺の宇宙環境を大きく変動させ、人類の宇宙活動に大きな影響を与えることが知られている。宇宙空間の高エネルギー粒子環境変動を予測するには、粒子加速機構として最も有力視されている衝撃波の理解が必須である。宇宙空間のような衝突がほとんど起こらない無衝突プラズマ中の衝撃波は無衝突衝撃波と呼ばれ、プラズマ波動と荷電粒子の非線形相互作用によって実効的な粘性散逸がもたらされる極度に非線形性の強い系になっている。無衝突衝撃波は直接観測データが豊富な太陽圏内において理解が進められているが、超音速の流れが一般的な宇宙空間では普遍的に見られるため、宇宙物理学・天文学においても様々な応用が考えられている。特に無衝突衝撃波における粒子加速は宇宙線加速の標準理論とされており、無衝突衝撃波の理解無くして宇宙線加速を理解することは不可能である。このように無衝突衝撃波は自然の粒子加速器として重要であるだけでなく、非線形プラズマ物理学の観点からも非常に興味深い対象である。

無衝突衝撃波の研究の究極的な目標は非線形なエネルギーと運動量の輸送メカニズムを理解し、それに伴う荷電粒子の加熱・加速効率を明らかにすることである。理論や観測・地上実験による研究も勢力的に進められてきたが、自己無撞着な数値シミュレーションを用いた研究は本質的に重要な役割を果たしてきたと言えるだろう。しかし、異なる時空間スケールの現象が密接にかつ非線形に結合した無衝突衝撃波の物理の完全な理解には至っていないのが現状である。

今回の研究の手法: 基礎方程式とアルゴリズム 無衝突衝撃波の物理過程を理解するには無衝突プラズマの第一原理計算手法である Particle-in-Cell(PIC) 法が最も有効な手法である。これは Vlasov 方程式(無衝突 Boltzmann 方程式)を有限個の粒子に離散化して解く手法である。一方で電磁場については空間格子点状で定義する。荷電粒子の運動方程式を解く際には格子点で定義された電磁場から補間して求めた電磁場を用い、荷電粒子が作る電荷密度・電流密度は格子点に配置し、有限差分法等で Maxwell 方程式を解くことによって格子点上の電磁場を更新する。プラズマを構成する電子・陽子の両者を運動論的に扱うため、電子スケール、イオンスケール、さらには流体力学的スケールまでの全ての物理過程を近似を導入せずに解き進めるのが特徴である。

ロードマップ 計算資源の増加や技術の向上により可能になること。現在観測機器、将来の観測計画との関係を述べる。PIC 法では、現実的なパラメータを用いて無衝突衝撃波の 3 次元数値シミュレーションを行うことは事実上不可能であり、人工的な物理パラメータを採用せざるを得ない。このうち代表的なものが電子と陽子の質量比 m_i/m_e 、光速と Alfvén 速度の比 V_A/c である。また計算時間の観点から衝撃波の

性質を大きく左右する衝撃波角（衝撃波上流の磁場配位角） θ_{Bn} にも制限がついている。現状では、これらとして数値的に扱いやすい値 ($m_i/m_e \sim 100$, $V_A/c \sim 0.01$, $\theta_{Bn} \sim 90^\circ$) を用いた 2 次元数値シミュレーションでは比較的大きな領域サイズを採用した長時間計算ができ始めている。一方で、このような条件のもとでも 3 次元数値シミュレーションの計算時間は大きく制限されている。また、これらの値を妥協しない場合には 2 次元であっても領域サイズや計算時間に対する制約が大きい。

これまでの研究では、現実的な質量比を用いた大規模計算ではマルチスケールのプラズマ波動が見られ、粒子加速に重要な役割を果たすことが明らかになっている。このような特徴は NASA の MMS(Magnetospheric Multiscale) 衛星による地球前面衝撃波の観測とも整合的である。また、理論的には衝撃波におけるマルチスケールプラズマ波動が電子の加速、特に宇宙線加速の標準メカニズムである Fermi 加速への注入に重要であることが示唆されている。さらなる大規模計算によって、3 次元や、より現実的な V_A/c を用いた数値シミュレーションを実施し、粒子加速の効率を調べることによって天体衝撃波を含めてより広範な応用が可能になる。また、PIC 法シミュレーションによって得られる超熱的電子 (10-100 keV) は非熱的制動放射によって硬 X 線を放射するため、将来の XRISM や Athena による超新星残骸衝撃波の高感度 X 線観測によって検証ができる可能性がある。

計算科学的な特徴、特にメモリアクセスの特徴 位相空間を連続的に運動する粒子と格子点上で定義された電磁場が相互に情報をやり取りするため、必然的にランダムなメモリアクセスが必要になる。粒子配列を空間座標でソートすることによってある程度は問題を軽減できるが、ソート時間とメモリアクセスの高速化とのトレードオフを考慮しなくてはならない。また、並列化においては粒子数が少数のプロセスに集中しないような動的な負荷分散が大規模数値シミュレーションでは必須である。現在用いているアルゴリズムの高速化・洗練化、またアーキテクチャに合わせた最適化などが今後必要になってくる可能性が高い。

必要資源の見積もり 理論ピーク性能、メモリ量、メモリバンド幅、ストレージ量、計算時間/ケース、総計算時間、解像度などについて述べる。2022 年現在、富岳では限られたパラメータでしか実行できない計算規模として質量比 $m_i/m_e = 400$ 、Alfven 速度 $V_A/c \sim 2 \times 10^{-3}$ 、Alfven マッハ数 $M_A \sim 5$ 、程度を採用した計算を富岳の 10 倍程度の性能を持つシステムで実施することを想定する。空間解像度として電子慣性長の 1/10、 x 方向（衝撃波法線方向）に 200 イオン慣性長、 y および z 方向には 5 イオン慣性長 ($40000 \times 1000 \times 1000$)、総粒子数 10^{12} 個程度、100 万ステップの計算が 1 ケースあたりにかかる総演算量を 500EFLOP 程度と見積もった。この計算規模で Alfven マッハ数 M_A および衝撃波角 θ_{Bn} を変えた 16 ケース程度のパラメータ空間探索にかかる総演算量は 8000EFLOP 程度であり、1 ケースあたりの計算時間を ~10 時間程度とすると、実効性能として ~10PFLOPS 程度が必要となる。

参考文献

- [1] Tomoaki Ishiyama, Francisco Prada, Anatoly A. Klypin, Manodeep Sinha, R. Benton Metcalf, Eric Jullo, Bruno Altieri, Sofía A. Cora, Darren Croton, Sylvain de la Torre, David E. Millán-Calero, Taira Oogi, José Ruedas, and Cristian A. Vega-Martínez. The Uchuu simulations: Data Release 1 and dark matter halo concentrations. *MNRAS*, Vol. 506, No. 3, pp. 4210–4231, September 2021.

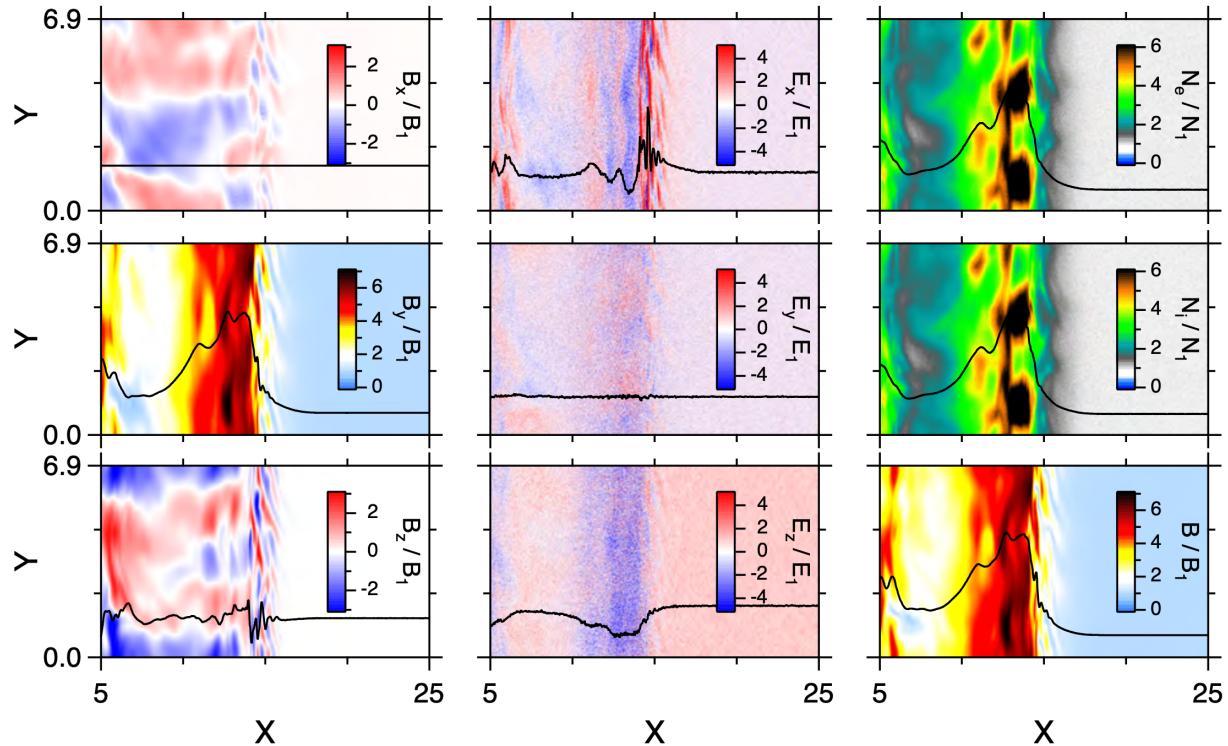


図 4.47 現実的な電子・陽子質量比 ($m_i/m_e = 1836$) を用いた 2 次元の PIC 法による無衝突衝撃波の数値シミュレーション例。磁場 3 成分 (左)、電場 3 成分 (中央)、密度および磁場強度 (右) を示す。複雑なマルチスケールのプラズマ波動が同定されている。[Matsukiyo & Matsumoto (2015)]

- [2] Wakana Iwakami, Hirotada Okawa, Hiroki Nagakura, Akira Harada, Shun Furusawa, Kosuke Sumiyoshi, Hideo Matsufuru, and Shoichi Yamada. Simulations of the Early Postbounce Phase of Core-collapse Supernovae in Three-dimensional Space with Full Boltzmann Neutrino Transport. *ApJ*, Vol. 903, No. 2, p. 82, November 2020.
- [3] Yosuke Matsumoto, Takanobu Amano, Tsunehiko N. Kato, and Masahiro Hoshino. Electron Surfing and Drift Accelerations in a Weibel-Dominated High-Mach-Number Shock. *Physical Review Letters*, Vol. 119, No. 10, p. 105101, September 2017.
- [4] H. Hotta and K. Kusano. Solar differential rotation reproduced with high-resolution simulation. *Nature Astronomy*, Vol. 5, pp. 1100–1102, September 2021.

付録 A

用語集

用語	解説
0	
0+ 状態	原子核の基底状態や励起状態は、角運動量 J とパリティ π で識別することができる。0+ 状態とは、 $J = 0$ でパリティが $\pi = +$ の状態。
1	
1 磁場散逸時間	磁場を作っている電流が電気抵抗により熱に変わること（ジュール散逸）によって、磁場が指数関数的に減少する典型的時間。
2	
2 電子クーロン反発積分	電子の振る舞いを調べるにはシュレディンガー方程式を解く必要があるが、その中で 2 個以上の電子を扱うためには電子間の反発を表すクーロン項を取り扱う必要がある。数値的計算では、2 つの電子間の反発を積分表現を用いて表す。このときに必要となる積分を 2 電子クーロン反発積分と呼ぶ。量子化学計算では、2 電子クーロン反発積分は数値計算の律速となるために、その取り扱いが重要となっている。
3	
3 体力, 三核子力	「核力」の項目を参照。
4	
4 中心 2 電子分子軌道積分	2 電子クーロン反発積分のうち 4 つの分子軌道中心を持つ 2 電子クーロン反発積分のこと。
A	
AFO	FMO 計算を 2 体展開の範囲内で固体系に適用するための技法の 1 つ。
Alder 転移	剛体球の密度を上げると、ある密度を境に液体から固体（結晶）に相転移すること。
Allreduce QR 法	QR 分解の手法の一種。高並列環境で高い並列性能を実現するために近年提案された新しい手法。

(次ページに続く)

用語	解説
allreduce, allgather, alltoall	計算ノード間での集団通信の様式。接頭辞の all はすべてのノードに渡って結果を共有する事を示す。allreduce は総和などリダクション型の演算を計算ノードにまたがって行うこと。allgather はすべてのノードから部分データを集めて全体データを作ること。alltoall は全体全通信によって、データの転置を行うこと。
AU	天文単位 (Astronomical Unit) の略称。天文単位とは長さの単位であり、地球と太陽間の距離に由来する。今日では 149,597,870,700 メートルと定義される。
B	
Belle II 実験	B 中間子の寿命や B 中間子がどのような粒子にどのような割合で崩壊するかを精密に測定する実験。高エネルギー加速器研究機構で行われている。2021 年頃に高精度データが取得できるよう計画されている。B 中間子に含まれるボトム夸克は、トップ夸克に次ぐ質量を持つ重い夸克であり、ボトム夸克の性質を詳細に調べることで、素粒子標準理論に内在する階層性の起源や素粒子標準理論にない新しい物理を明らかにできる。
Bi-CGSTAB 法	疎行列の連立一次方程式の解法の一つ。反復解法の一つである双共益勾配法 (BiCG 法) の残差を減少させ、安定化した手法。比較的高速・安定とされる手法で、偏微分方程式を解く多くの物理問題（熱流体・構造・電磁場など）の核となるソルバとして広く使用されている。
Bisection BW	バイセクションネットワークバンド幅のこと
BLAS	Basic Linear Algebra Subprograms (BLAS)。ベクトルと行列に関する基礎的な線形代数演算のサブプログラム集のこと。線形代数演算ライブラリ API のデファクトスタンダードでもあり、高度に最適化された実装がインテルなどの各ハードウェアベンダーなどから提供されている。
B 中間子	ボトム夸克を一つ含む中間子の総称。
B 中間子混合行列要素	B 中間子のうち中性の B 中間子は B^0 と \bar{B}^0 の 2 種類が存在する。これらの中性 B 中間子はお互いに粒子反粒子の関係にある。これらの粒子は弱い相互作用の影響で互いに変化することができる。この変化頻度のうち、強い相互作用に関する部分を因子化したものを B 中間子混合行列要素 B_B とよぶ。実験値での変化割合は強い相互作用と弱い相互作用を含む値であるが、理論計算による B_B を用いることで実験値から弱い相互作用による影響を引き出すことができる。
C	

(次ページに続く)

用語	解説
Convolutional Neural Network	入力データの1部との相関を計算するようなフィルターを多数用意し、データ上でフィルターを走査した結果を次の層へ入力するような畠み込み層をもつニューラルネットワーク。
CP 対称性	C は荷電共役変換（粒子 \leftrightarrow 反粒子）、P はパリティ変換（鏡映変換）を表す。これらの変換の組み合わせによって理論が不变であるとき、その理論は CP 対称性を持つという。
D	
DGEMM	倍精度汎用行列乗算のための BLAS のサブルーチンのひとつ。LINPACK ベンチマーク内で多用されているため、DGEMM 実装の性能はベンチマーク結果に大きな影響を与える。
DZP 基底関数	double zeta polarization 基底関数。1s, 2p などの原子基底の各成分を、二つの短縮ガウス型関数を用いて表現した基底を DZ(double zeta) 基底と言い、それに分極関数を加えたもの。
E	
EDA 標準技術	電子情報技術産業協会 EDA 標準技術専門委員会配下の EDA 標準化小委員会において推進している国際標準化機構の活動に対応した EDA (Electronic Design Automation) の標準化活動のこと。
ESR	電子スピノン共鳴の略。開殻系の電子状態に関する情報が得られる。
F	
Fermi-Pasta-Ulam の非線形励起	非線形なバネで互いにつながれた多数の粒子の運動において孤立した波が生じる現象。ソリトンと呼ばれる
Field Programmable Gated Array	プログラム可能な論理ゲート回路。略称は FPGA。
Finite-difference Time-domain 法	電磁場解析等で用いられる計算手法の一つ。空間を差分近似し陽的な時間進行法を用いる。
FMO	フラグメント分子軌道法 (FMO 法: Fragment Molecular Orbital Method) は、北浦和夫教授（現神戸大学）によって開発された量子化学理論。分子全体を小さなフラグメントに分割して計算をするため、通常の量子化学計算では不可能なタンパク質のような大規模分子系の量子化学計算が実行可能。また、分割した小規模のフラグメントごとに並列計算を実行することが可能なため、非常に効率よく並列計算を実行可能。
Fock 行列	多数の電子の振る舞いを平均化されたポテンシャル中を動く、一電子のシュレディンガー方程式を行列表現した行列のこと。
G	

(次ページに続く)

用語	解説
gatherv	MPI のデータ転送関数 (MPI_GATHERV)。全てのプロセスから一つの宛先プロセスにメッセージを転送する。
Gauss 関数の局所性	中心からの距離が大きくなるにつれて急速に値が減衰する Gauss 関数の性質。2 電子クーロン反発積分などの計算で、これを利用したカットオフ、演算削減は大きな効果がある。
GCC	GNU Compiler Collection。自由に使える C/C++ 言語などのプログラム言語のコンパイラ。
GMRES 法	疎行列の連立一次方程式の解法の一つ。反復解法の一種で、比較的ロバストなクリロフ部分空間法の一つとして知られている解法。同時に使用する前処理法や計算条件によって、並列計算性能が高く、高速に収束解が得られるため、偏微分方程式を解く物理問題の一部で使用されている。
GPU	グラフィクスプロセッシングユニットを参照。
G タンパク質共役受容体	細胞外の神経伝達物質やホルモンを受容してそのシグナルを細胞内に伝える受容体。その際 G タンパク質と呼ばれるタンパク質を介してシグナル伝達が行われる。多くの薬剤のターゲットになっている。
H	
HF 交換相互作用	密度汎関数法において、交換相互作用を表現する汎関数に Hartree-Fock(HF) 法の交換相互作用を使ったもの
High-radix 型	ある計算ノードからもう一つの計算ノードへの通信が、その他の計算ノード同士の間の通信と同時に実行しやすいネットワーク。すなわち他のノードに妨害されずに通信できる一ノードあたりのノード数が多いネットワーク。
Hodgkin-Huxley formalism	イカ巨大軸索を対象に神経細胞における活動電位の発生メカニズムをゲート（後に実体としてこれに相当するイオンチャネルがあることが明らかにされた）の協同性を用いて電気回路として記述したモデルが Hodgkin-Huxley モデルである。この形式を Hodgkin-Huxley formalism と呼び、多くの種類のイオンチャネルやマルチコンパートメントモデルに対しても用いられる。
HPCI 戰略分野	スーパーコンピュータ「京」を中心とした HPCI (High Performance Computing Infrastructure) を最大限に活用することによって、戦略的に取り組むべき 5 つの研究分野
I	
IACM	International Association for Computational Mechanics。国際計算力学連合。計算力学に関するいくつかの国際学術講演会を運営する。

(次ページに続く)

用語	解説
Izhikevich モデル	力学系の研究者であり理論神経科学者の Eugine Izhikevich が考案したニューロンの点モデル。膜電位の他に内部パラメータを持ち、パラメータを変更することで実際のニューロンが示す様々な発火パターンを再現することができる。
J	
J-PARC	J-PARC (Japan Proton Accelerator Research Complex) は大強度陽子加速器を中心とする実験施設で茨城県東海村にある施設。世界最高クラスの強度の陽子ビームを使って生成される、中性子、ミューオン、K 中間子、などの 2 次ビームを用いたさまざまな実験が行われている。
K	
Krylov (クリロフ) 部分空間反復法	線形連立方程式を数値的に解く計算方法 (アルゴリズム)。特に大規模な疎行列を係数とする連立方程式を効率よく解くことができる。共役勾配法 (Conjugate Gradient algorithm, CG) が良く知られている。
K 中間子	ハドロンのうち、クォーク 2 個 (クォーク・反クォーク対) からなるものが中間子 (meson) であるが、一つのクォークがストレンジクォークであり、もう一つのクォークがアップまたはダウンクォークのものを K 中間子と呼ぶ。これらのクォークの組み合わせと粒子反粒子の違いにより、 K^+, K^-, K^0, \bar{K}^0 の 4 種類の K 中間子が存在する。
L	
L1, L2 キャッシュ	CPU にはメモリとのデータ転送を節約するためのデータの一時的な保管場所があり、それをキャッシュという。キャッシュは演算装置とメモリとの間に多階層に配置されており、演算装置に近い順に L1, L2 と言う。
L1 正則化法	影響を与える因子の数を抑えることができる機械学習法
Langevin 方程式	ブラウン運動を記述する確率微分方程式
Leaky Integrate-and-Fire neuron モデル	積分発火モデルの中でリーク電流をつけたモデル
Lennard-Jones 型ポテンシャル	原子間の相互作用を記述する最も簡単なポテンシャルで、斥力の項と引力の項の 2 つによって記述される。
LHC (Large hadron Collider)LHCb	大型ハドロン衝突型加速器 (Large Hadron Collider)。欧州原子核研究機構 (CERN) で稼働中の加速器の名称。ヒッグス粒子の発見と超対称性粒子などの新しい物理の探索を目指している。
logP	化合物の脂溶性を表す量
LPB	LSI Package Board の略。

(次ページに続く)

用語	解説
M	
MIPS ピーク性能	MIPS (ミップス) 値、あるいは MIPS ピーク性能は、100 万命令毎秒 (million instructions per second) の略で、コンピュータの性能指標の一つ。
N	
NMR 分光	原子核の磁気を測定する手法。分子構造に関するデータが得られる。
NP 完全問題	クラス NP に属する問題でかつ、クラス NP すべての問題から多項式時間帰着可能な問題。このクラスに属する問題は多項式時間で解を見つけるアルゴリズムが存在しないと予想されている。(P≠NP 予想)
O	
on the fly	「実行中に」を意味し、プログラム中で繰り返し必要となるデータを、その度ごとに計算して用いるアルゴリズムを指す。これと対極にあるのは「あらかじめ計算して保存しておいたデータを、必要になる度に記憶装置から参照して用いる」やり方である。
ONIOM 法	ONIOM 法は諸熊啓治教授（現京都大学）により考案された QM/MM 計算の代表的な方法。生体高分子などの巨大分子をいくつかのレイヤーに分け、レイヤーごとに量子化学計算や分子力学計算を行うことで、巨大分子の電子状態や分子構造の評価や反応機構の解析を行うことが可能。
P	
PDB 構造	Protein Data Bank(PDB) に登録されている NMR 解析や X 線構造解析などの実験的手法によって得られた蛋白質の構造。
pKa	化合物の酸性度を表す量
Q	
QCD の第一原理計算	QCD の第一原理計算方法として格子 QCD による計算がある。近似を使わずに計算する。
QED 摂動計算	QED の計算では電荷を帯びた粒子同士の相互作用の強さ（電荷 e）が比較的小さい。物理量を電荷をゼロとした場合の解析的計算から e のべき級数として表すことができ、べき級数を打ち切った場合の近似誤差を制御できる。このばあい各次数の前の係数を解析的に計算することができる。このようなべき級数による計算を摂動計算という。
QueryDriven	データに対するクエリー（質問）を行いながら、対話的にデータを調べていく手法。
R	

用語	解説
Recurrent Neural Network	内部に閉路をもつニューラルネットワーク。この閉路上で前回までの入力に対応するメモリを保持しており、現在の入力への出力計算に対してこのメモリの情報も用いる。自然言語や音声、画像（画素列）といった系列データによく用いられる。
RHIC, RHIC 実験	RHIC (Relativistiv Heavy Ion Collider, 相対論的重イオン衝突型加速器) は、米国ブルックヘブン国立研究所 (BNL) にある衝突型加速器。この加速器を用いて重イオン同士をぶつけたて高温・高エネルギーの状態を作り出す実験を行っている。重イオンとしては金原子核や銅原子核などを用いている。大きな原子核同士を高エネルギーで衝突加熱することでクォーク・グルーオンプラズマ相の性質を調べることができる。
RI (放射性同位元素) ビームファクトリー (RIBF)	原子核は 6000 種以上存在すると考えられている（「核図表」の項目参照）が、天然に存在する安定同位体（「安定同位体」の項目参照）は 300 種程度しかない。RIBF は天然に存在しない不安定核種を人工的に合成するための加速器であり、理化学研究所にある。陽子・中性子過剰核（「陽子・中性子過剰核」の項目参照）を合成する能力において、現在、世界でトップを走る。RIBF によって、陽子・中性子過剰核の知られざる性質や、元素合成に関する新たな知見が得られている。
r 過程	宇宙における重元素生成過程は主に、星の内部で安定線上を時間をかけて進む中性子捕獲反応 (s 過程) と、わずか数秒間の爆発的な過程で安定線から離れた原子核を作る r 過程に分けられる。図 4.5.3.1 も参照。
S	
SPICE モデル	SPICE (Simulation Program with Integrated Circuit Emphasis) はカリフォルニア大学バークレー校で開発された回路シミュレータであり、SPICE モデルとはこの回路シミュレータで使用される、受動素子（抵抗、インダクタ、コンデンサ等）と能動素子（トランジスタ等）の等価回路モデルのこと。
Spike-timing dependent plasticity	あるシナプスにおいて、シナプスを形成する 2 つの細胞の発火のタイミングに応じてその結合の強さを決定するシナプス可塑性。略語は STDP。
STM	走査型トンネル顕微鏡。短針と固体側とのトンネル電流の観測により、表面構造や電子状態を知ることが可能。原理は異なるが、原子間力顕微鏡 (AFM) なども表面解析に用いられる。
Structure-based drug design (SBDD)	タンパク質などの立体構造をもとにして薬剤のデザインをする方法

(次ページに続く)

用語	解説
SU(3) 群, SU(3) ゲージ理論	SU(3) 群とは数学の群の性質を持つもので 3×3 の特殊ユニタリー行列が構成する群である。この群の構造をゲージ対称性として持つ場の理論を SU(3) ゲージ理論とよぶ。QCD は SU(3) ゲージ理論の一つである。
SuperKEKB/BelleII 実験計画	B 中間子の寿命や B 中間子がどのような粒子にどのような割合で崩壊するかを精密に測定する実験。高エネルギー加速器研究機構で行われている。2021 年頃に高精度データが取得できるよう計画されている。B 中間子に含まれるボトムクォークは、トップクォークに次ぐ質量を持つ重いクォークであり、ボトムクォークの性質を詳細に調べることで、素粒子標準理論に内在する階層性の起源や素粒子標準理論にない新しい物理を明らかにできる。
T	
tagSNP	ゲノムの特定の領域においてその領域中の他の SNP (一塩基多型) の代表となりうる SNP
Thin node	少数の演算器、メモリによって構成される計算ノード。
U	
uds ハドロン	6 種類のクォークのうち、軽い u,d,s の 3 種類のクォークから構成されるハドロン。
union-find アルゴリズム	グラフ構造の中から、連結クラスター (互いに辺でつながっている頂点の集合) を見つけ出す際に用いられるアルゴリズム。
V	
VOF 関数	Volume Of Fluid の略。空間を計算要素に分割した際、その計算要素に占める流体の体積比率を用いる手法を VOF 法と呼ぶ。このとき使用する全計算要素の VOF 値の事を VOF 関数と呼ぶ。
X	
XFEL	X 線自由電子レーザー (XFEL) は、波の位相がきれいにそろったレーザーの性質を持つ超高輝度の X 線を発生させることのできる光源
あ	
アクセラレータ	計算の一部を肩代わりする外部演算装置。GPU や FPGA はアクセラレータである。
足場タンパク質複合体	細胞内情報伝達系において、複数の情報伝達タンパク質と結合して複合体を形成する足場となるタンパク質の複合体
アパタイト	リン酸カルシウム (燐灰石) のことだが、生体では水酸基が入ったヒドロキシアパタイトとして歯や骨の主要構成要素となっている。
アンサンブルシミュレーション	沢山のシミュレーションを行い、その統計的性質を研究する計算手法

(次ページに続く)

用語	解説
アンジュレータ	加速された電子の直線軌道上に沿って、多数の N、S の磁極からなる磁石列を上下に配置して、その間を通り抜ける電子を周期的に小さく蛇行させて、明るく特定の波長を持った光を作り出す装置。
アーティフィシャルニューラルネットワークモデル	複数の同種神経の平均としての活動量関数とシナプス伝達関数を定義してネットワークを形成させるモデル。広義の McCulloch-Pitts Model
い	
位相空間	燃料プラズマ粒子の 3 次元位置と 3 次元速度を座標とする 6 次元空間。粒子間の衝突効果が十分に大きければ、局所的な熱力学的平衡を仮定して 3 次元流体モデルでプラズマを記述できるが、衝突効果が小さい高温プラズマに対しては 6 次元位相空間の粒子分布を記述する運動論モデルが必要になる。ただし、磁場閉じ込め核融合プラズマのような強磁場中の運動論モデルは 5 次元位相空間に簡約化できる。
位相空間密度	位相空間における密度。位相空間とは位置と速度（または運動量）を座標とした空間のことである。例えば、我々の世界ではそれぞれ 3 次元で合せて 6 次元の空間。
位相骨格	データを変化点の接続情報（スケルトン・骨格）により表し、大規模なデータを非常に小さなデータサイズで特徴付けることができる。
一般相対性理論	AINシュタインによって提案された重力の理論。物質のエネルギーが時空の幾何学を決定する理論。時空の幾何学を重力とみなす。星の重力を決定するだけでなく、宇宙全体の幾何学をも決定でき、宇宙物理学における基礎となる理論。量子力学が重要となるミクロの世界での重力の振る舞いについては記述できない。
遺伝子プロモータ	特定の遺伝子の発現を促すタンパクなどの細胞内物質
イベント駆動型	現象や手順を有限の数の瞬間に起きた事象（イベント）の連続として扱うやり方
陰解法	時間積分の一つ。時間微分の離散化において後退差分（現在と過去の値を使って離散化する）を用いて離散化を行う手法。元の偏微分方程式は未知変数の連立一次方程式にと離散化され、この連立一次方程式を解くことになる。
インフレーション	宇宙誕生直後における宇宙の指数関数的膨張のこと。
う	
ウィーク・スケール	並列単位当たりの問題サイズを一定にして、並列数を増やしていく場合（つまり、問題サイズが並列数に比例して大きくなる）での、計算時間の変化

用語	解説
え	
エキゾチックハドロン	中間子(メソン)はクォークと反クォークから構成され、重粒子(バリオン)は3個のクォークから構成されると考える単純なクォークモデルからは予測できない異種のハドロン。
液体論	液体は気体に比べて原子、または分子間の相互作用が強く、また固体とは違いこれらの粒子が動き回るためその取扱いは容易ではない。液体そのものから溶媒としての性質等についても議論がされており、数値シミュレーションによる研究も盛んである。
エネルギー階層性	素粒子標準理論のパラメータには以下のような階層性がある。(1)弱い力の媒介粒子の質量が重い。(陽子の約80倍と約90倍) クォークやレプトンの質量はバラバラであるが、(2)トップクォークは特に重い。(陽子の約170倍) (3)ニュートリノの質量が直接測定不可能なぐらい軽い。質量はエネルギーと等価であるので、エネルギー階層性という。このような階層性の起源を明らかにすることは現在の素粒子物理学の課題である。
エネルギーバンド	結晶といった固体物質の中に分布する電子について、量子力学的なエネルギー準位をあらわしたもの。
エネルギー分散外挿法	通常、変分計算によって得られたエネルギー期待値は、真のエネルギー期待値の上限しか与えることができない。変分空間を徐々に広げて、エネルギー期待値をエネルギー分散期待値の関数として外挿することによって、精度よく真のエネルギー期待値を見積もある方法。
エネルギー密度	1kgあたりどのくらいエネルギーを蓄えられるのか、電池の性能をあらわした数値。
エピジェネティクス	DNA塩基配列の変化を伴わないが、細胞分裂後も継承される遺伝子発現あるいは細胞表現型を研究する学問領域。
円偏光	光を始めとした電磁波は、進行方向と垂直に電場と磁場が振動する横波である。円偏光では、進行方向と直交する平面上で電場もしくは磁場の向きが円運動を描く。進行方向を手前に取って時計回り、反時計回りのものが存在する。
エージェント	外界からの情報を知覚し、自らの基準やルールに基づいて行動を選択できるような意思決定主体のこと。社会シミュレーションでは、社会現象の構成要素の個人や企業などのエージェントを模倣したエージェントモデルを使うことが多い。
お	

(次ページに続く)

用語	解説
オントロジー	事物の概念間を is-a や has-a といった関係で記述して知識を体系立てたデータ。近年では、クイズの人間チャンピオンを破った IBM 社のワトソンなどがオントロジーを用いた質問回答システムとして話題となった。
オーダリング	主に、メモリ空間でのデータの連続性を改善するために、数値データの格納順序を入れ替え、計算機による処理性能の向上を図ることを指す。
か	
階層的時ステップ	要素により計算時ステップ幅に幅がある場合に例えば 2 の整数乗などのあらかじめ決められた規則に沿って時ステップ幅を決定する事で同期を容易にするやり方
カイラル外挿	質量を持つフェルミオンで計算した物理量を質量がゼロとなる極限へ外装すること。質量がゼロのフェルミオンはカイラル対称性を持つためこのような言い方をする。
カイラル凝縮	クォーク・反クォーク対が凝縮し、真空期待値を持つこと。
カイラル対称性	質量ゼロのフェルミオンが持つ対称性の一つ。光速で運動するフェルミオンはそのスピンが運動量に対して平行な場合（右巻き）と反平行な場合（左巻き）の 2 つの独立な自由度に分かれる。理論が、右巻き粒子だけで（または左巻き粒子だけで）、粒子の入れ替え操作に対し不变である場合にカイラル対称性があるという。標準理論は質量ゼロのフェルミオンによって構成されている。
カイラルなゲージ対称性	フェルミオンの右巻き成分と左巻き成分が異なるゲージ対称性を持つ場合の対称性のこと
カイラル有効場理論	低エネルギー領域における物理現象を記述するために必要な自由度だけを取り入れた有効核力理論。多体力を自然な形で含むことができる事がその大きな特徴の一つである。
核子	原子核を構成する陽子と中性子の総称。大きさはおよそ 10^{-15} m。核子は 3 個のクォークが強い力で結合した粒子である。
核子多体系	複数の核子（陽子・中性子）が互いに相互作用しあっている物理系のこと。原子核もその一つ。
核図表	原子核は陽子と中性子から構成されている。陽子数・中性子数をそれぞれ縦軸・横軸にとって、強い力にて束縛しうる原子核の種類を平面的に図示したものを核図表 (nuclear chart) と呼ぶ。未発見の原子核を含め、6000 種以上の原子核が存在すると考えられている。
核変換テクノロジー	原子炉の廃棄物処理の一つとして、長い寿命をもつ放射性同位元素や特に毒性の強く危険なものを、核反応を利用して短い寿命のものに変換させ消滅させるために必要な技術、方法、基礎知識等。

用語	解説
殻模型計算 (原子核殻模型計算)	原子核の構造を計算する手法の一つで、量子化学における配置間相互作用計算と類似した手法。バレンス殻（「バレンス殻」の項目を参照）内の量子多体問題を正確に解く方法であり、その計算は大次元行列の固有値問題に帰着する。
核力	核子やバリオンの間に働く力。陽子と中性子を結び付けて原子核を形作る。湯川秀樹博士は核力をパイ中間子の交換による作用であると提唱し、実際にパイ中間子が発見された。核力は基礎的な力でなく強い力による副次的な力であり、複雑な様相を呈する。たとえば、3つの核子の間に働く核力（3体力あるいは3核子力）は2つの核子間に働く核力（2体力）の単純な重ね合わせではないことが挙げられる。核力の性質の理解には、強い力の深い理解が必要である。
過減衰極限	Langevin 方程式において慣性力を無視できるとした場合の特殊ケース
火成活動	マグマの発生や移動に伴って生じる諸現象の総称。
可塑的触媒場	タンパク質の構造変化などにより変化可能な触媒を行う環境
カットオフ半径	短距離の成分のみを取り扱う場合に、どの程度の長さまで扱うかという距離。この距離より離れた成分は0と考える。
カラーグラス凝縮	高エネルギーのハドロンで、量子色力学で「色」をもつグルーランが大量に生成されて高密度に凝縮した状態。
軽い原子核、重い原子核	陽子数と中性子数の和を質量数と呼び、原子核の質量はほぼ質量数に比例する。「軽い」「重い」とは、この質量数の大きさを指している。明確な線引きはできないが、質量数が10程度以下のものは軽い原子核、100に近くなると重い原子核と呼ばれる。重い原子核の中でも特に、質量数が100-150程度の原子核を中重核、200程度以上のものを重原子核と呼ぶことが多い。
カルシウムイメージング	カルシウム感受性蛍光色素を標的細胞に導入して、蛍光観察を行う方法。一般にカルシウムの配位結合によるセンサー分子の立体構造変化は大きく蛍光変化も大きい。そのせいか蛍光プローブを使った神経活動観察法としてはカルシウムイメージングは主流でありつづけている。脳組織内の多点同時観察を見据えると蛍光プローブの導入法が重要で特定神経組織へのローカルインジェクションや特定の遺伝子プロモータを標的としたカルシウムセンサータンパク質の遺伝子導入が2000年代になって多く行われている。
カーティアン座標系	直交座標系の事。空間の位置を示すのに互いに直交する座標系を用いる。

(次ページに続く)

用語	解説
カーネル最適化技術	プログラムにおいて主要なコストを占める逐次演算処理をカーネルと呼ぶ。プロセッサのアーキテクチャに依存して逐次演算処理の最適化方法は異なるため、特に、メニーコアプロセッサを効率的に利用するには新たな最適化技術の開発が必要となる。
カーパリネロ法	電子状態計算により原子にかかる力を直接見積もりながら、分子動力学計算を行う手法の1つ。電子状態に時間発展方程式を導入し、計算の高速化を実現している。
き	
機械学習	定式化された問題を基に、データの背後に存在する規則や知識を見つけるデータ処理技術。教師データの有無によって教師あり学習や教師なし学習、または半教師あり学習などが存在する。また、教師データを与える代わりに報酬を定義する強化学習も存在する。
季節内振動現象	中緯度帯にみられる高低気圧等の総観規模現象に比べて長く、季節変化より短い時間スケール(おおよそ10日～90日周期)の現象を総称して季節内振動現象と呼ぶ。有名な季節内振動現象として、地球規模の活発な積雲活動域が熱帯を東進していくMadden-Julian振動や、アジアにおけるモンスーン活動が知られており、中長期予報を行う際の重要な現象と考えられている。
基底重なり	電子雲を表現するために用いられる局在基底もしくは平面波基底間の空間的なオーバラップ(重なり)のこと。異なった平面波基底間の重なりは全空間で積分をするとゼロとなるが、局在基底間では重なり積分はゼロでない場合がある。
基底関数極限	無限に多くの基底関数を用い、基底関数展開による誤差がなくなる極限。デジタルカメラの画素数が上がり、アナログ写真との差がなくなった極限のような概念。
基底波動関数	多体波動関数を記述するために用意された波動関数のセットを基底波動関数と呼ぶ。これをどう選ぶかが多体問題を解く上で重要となる。
キャビテーション	液体の流れの中で局所的に圧力が変化することにより短時間に泡の発生と消滅が起きる物理現象であり空洞現象とも言われる。キャビテーションの発生は、発生する気泡により、ポンプなどの流体機器における振動・騒音の発生や性能低下の原因となる。また同時に発生する圧力波がこれらの機器表面のエロージョン(壊食)を起こして、効率を下げたり破壊することがある。
キャリア	電荷を運ぶ自由な粒子。例えば、電子やイオンがあげられる。半導体においては、伝導電子と正孔(ホール)のことを指す。

用語	解説
球面調和関数展開	球面調和関数は完全性をもち、球面上の任意の連続関数を一意に展開できる。このため、球面上のスカラー場の表現に用いられる。
境界埋込法	流体の運動と構造体の変形を同時（連成問題）にシミュレーションするときに用いる手法。流体をオイラー座標系で表現し、構造物をラグランジュ座標系で表現する。
境界層	大気層のうち、地表面に最も近い大気の層。正式には大気境界層という。地表面の影響を熱や摩擦応力などを介して強く受ける。境界層の厚さは日変化や季節変化、場所による変化があるが、熱対流が卓越しているときは、高度 1000m～2000m に達する。大気境界層内では、乱流による熱、運動量、物質の輸送が卓越している。
強化学習	教師あり学習のように出力として望ましい値を陽に与えることはできないが、出力に対して何らかの報酬が計算できる場合に用いられる機械学習の枠組み。近年話題となった AlphaGo は強化学習と探索を用いたシステムである。
教師あり学習	入力データに対して、クラスラベルや推定値など出力として望ましい値（教師情報）が与えられる機械学習の枠組み。識別や回帰などは通常教師あり学習として解かれる。一部のデータにのみこうした教師情報が付与されているような場合には、半教師あり学習と呼ばれる。
教師なし学習	入力データに対して、クラスラベルや推定値など出力として望ましい値が存在しない機械学習の枠組み。クラスタリングなどは通常教師なし学習として解かれる。一部のデータにのみこうした教師情報が付与されているような場合には、半教師あり学習と呼ばれる。
強震動	明確な定義を持つ言葉ではないが、一般に、建築・土木構造物の被害に直接関与するような地表面での強い地震動のことをいう。
強相関電子系	物質の中でも電子-電子間に働く有効な静電相互作用が強いものの総称。電子がもつ電荷自由度とスピン自由度が複雑に絡み合い、特異な物性を示すことがある。
共発現解析	ある遺伝子の発現と相關の高い遺伝子を同定し特定の生物学的現象に互いに関係のある遺伝子群の機能などを解析する方法
共役勾配法	連立一次方程式を解くため、または制限付きの 2 次形式の極値を求めるための反復的アルゴリズムの一つ。
共溶媒濃度	溶液中の溶質および主なる溶媒のほかに含まれる第二の溶媒成分の濃度

(次ページに続く)

用語	解説
局在基底	量子力学的には電子は点ではなく雲のように広がっている。この広がりを表し電子の雲の状態を記述するために用いられる関数のこと。電子は周囲の環境により、電子雲の広がりかたの度合いはことなるが、特にその広がりが強くない場合に用いられる関数のことを局在基底と呼ぶ。分子・原子では、電子の広がりは限定的であるために、原子・分子中の電子雲の状態を記述するために局在基底はしばしば用いられる。
局在軌道	特定の原子あるいは結合領域に、空間的に局在した分子軌道のこと。分子の量子化学計算で得られる分子軌道は、通常、分子全体に広がった（非局在化）形状をしているが、これらの非局在分子軌道に特定のユニタリー変換を施すことによって、局在軌道に変換することが出来る。空間的に離れた局在軌道どうしの積は無視できるほど小さくなることを利用して、計算コストの軽減をはかることができるほか、計算結果の物理化学的な解釈を手助けする目的にも用いられる。
局所準粒子乱雑位相近似	量子多体系において、非平衡状態の規準モード（近似的に独立な運動）を決定する方法。
虚時間軸	ある温度における統計力学的な平衡状態をあらわす式が、見た目上、通常の量子力学的な時間発展の式の「時間」のところに純虚数の値を入れた形になっており、「虚時間」と呼ばれます。単に見た目の問題というだけではなく、実時間 \leftrightarrow 虚時間の対応を考えることにより理論的にも見通しが良くなることが多い。
巨大応答	外部からの磁場や電場や光照射などの刺激によって物質中の集団秩序を変化させ、抵抗値などを劇的に変化させること。
均質化法	マルチスケール解析手法の一つ。材料の詳細ミクロ構造をマクロ解析に反映させるために、ミクロとマクロの連成解析を行う。
近似ベイズ計算	尤度を直接計算せずに、近似事後分布からサンプリングする方法。事前分布からパラメーターをサンプルしモデルのシミュレーションを行い、実際のデータとシミュレーション結果が十分に一致する場合のみパラメーターのサンプルを採択する。この計算を繰り返すことで、採択されたサンプルの集合は、近似事後分布からのサンプルであると見做すことができる。
金属原子拡散	燃料極における多孔質構造を変化させる金属原子の移動。三相界面長さの減少を通じて反応性を低下させる。
金属誘起ギャップ状態	半導体と金属の界面において金属の電子状態が半導体にしみ出すことで半導体ギャップ中に生成される新たな電子状態・準位。
ぎ	

用語	解説
逆引き用分割テーブル	配列の添字から対応する電子・スピンの状態を求めるためのテーブル。部分系に分割したテーブルを組み合わせて用いることで、そのサイズを大幅に小さくすることが可能となる。
ギャザー・キャッタ機構	配列に対する間接インデックス参照を効率的に行うためのハードウェア組み込み機構。
凝集物質(凝集系)	多数の原子や電子が集まり互いに強い影響をおよぼしあっている状態、あるいはそのような状態にある物質。固体や液体は典型的な凝集物質である
行列模型	弦理論の非摂動的定式化の一つ。
く	
クォーク	アップ、ダウン、チャーム、ストレンジ、トップ、ボトム、と名付けられた質量の異なる6種類のフェルミオンの族名。名前の違いはフレーバーと呼ばれる。電磁気力、弱い力、強い力を受ける。アップクォークとダウンクォークは強い力により束縛しあい、陽子や中性子、中間子などの粒子を形成する。クォークの名前の違いは質量によって決まっており、質量の軽い順にクォークを並べると、アップ、ダウン、ストレンジ、チャーム、ボトム、トップとなる。全てのクォークはそれぞれがカラーと呼ばれる内部自由度をもつ。クォークの間に働く強い相互作用は、この内部自由度の性質に基づいて記述されているQCDによって理解される。
クォーク作用	クォークに対する作用。作用から運動方程式などが得られる。
クォーク・グルーオンプロセス	クォークはハドロンの中に閉じ込められておらず、自由に動き回れる状態。
クォーク・グルーオン・プラズマ	通常、クォークはハドロンの中に閉じ込められているが、高エネルギー状態では自由に動き回れるようになる。クォークとグルーオンが電離したプラズマ状態。
クラウドソーシング	webを通じて不特定多数の人に仕事を依頼するサービスやプロセス。とくに作業量が多いタスクを細分化して多くの人に廉価で遂行してもらう場合がおく、機械学習においては大量のデータに教師情報を付与する際などによく使われる。たとえば、画像認識において深層学習が流行するきっかけとなったImageNetデータセットは、Amazon Mechanical Turkというクラウドソーシングサービスを用いて作成されている。
クラスターモデル	非周期条件の下で、固体側を有限の原子数のクラスターとして表現するモデル化。適宜水素終端化処理した上で、分子の吸着などを計算する。

(次ページに続く)

用語	解説
クラスレート	結晶構造中に異分子が共有結合をすることなく内包されたもの。包摂化合物。メタン分子が氷状結晶中に内包されたメタンハイドレートなどが知られる。
クリロフ部分空間解法	連立一次方程式の解を求める際に使用される行列解法の一つ。行列積を直接計算する代わりにベクトルを利用した解法の総称で、ロシアの数学者にちなんで名づけられた。現在最も主流の行列解法であり、具体例として Bi-CGSTAB 法、GMRES 法などがある。
クリープ構成則	主として金属材料が高温状態にさらされた際に呈する非線形挙動を応力とひずみの関係として記述したもの。各種金属に固有の温度を超えると、荷重が一定でもひずみが時々刻々変化する、いわゆるクリープ変形が顕著になる。その挙動を応力一ひずみ関係として記述したもの。
クーロン力	時間変化を無視できる場合に電荷をもった粒子の間に働く電気力の事。
ぐ	
クラスレートハイドレート	複数の水分子で作るかご型の構造中に気体分子が取り込まれた結晶。気体分子と水の混合物を加圧することにより生成する水和物
グラフィックスプロセッsingユニット	画像処理・画面表示用のハードウェア。多数の小さな計算コアを有しており、これを並列計算に用いる手法がある。
グラフェン	六角形二次元平面に周期的に配置された格子構造を持つ炭素結晶。
グラムシュミット直交化	直交化とは、いくつかの「線形独立だが互いに非直交なベクトル（または関数）の組」を、「互いに直交するベクトル（または関数）の組」に変換する操作を指す。直交化を施することで、数学的表現が簡素になって取扱い易くなる。グラムシュミット直交化は、いくつか存在する直交化法の中でも概念的に最もシンプルなもの。
グルーオン	強い力を媒介する粒子。
グルーボール、ハイブリッド粒子	グルーオンが複数個結合した複合粒子がグルーボール。これにクォークもからむとハイブリッド粒子と呼ばれる。
グローバルビュー	通常並列計算機は複数の計算機から構成される複合システムであり、個々の計算機間は別個のビュー、すなわち実行の状態（メモリ）をもつ。グローバルビューは特別なソフトウェアもしくはハードウェアにより並列計算機全体で单一のビューを共有する方式であり、これによって並列計算機のプログラミングが大幅に簡略化される。
グローバルファイルシステム	並列計算機のすべての計算ノードから参照可能な共有ファイルシステム。一般に利用者の恒久的なファイル置き場として使われ、ローカルファイルシステムと比較して大容量かつ安定性を重視した構成となっている一方、読み書きの速度は限定的である。

用語	解説
け	
形態学	細胞の形状と組織の広がりなどを調べる方法
結合強度	分子同士が結合している時の強度
結合クラスター展開	無限次の摂動論に相当する電子相關理論。複雑なテンソル積和処理を伴う繰り返し計算が必要で、2次に比して精度は高まるが計算コストは高い。
げ	
原子核殻模型計算	原子核の構造を計算する手法の一つ。陽子と中性子の多体系である原子核を、適切な1粒子状態を基礎にして核力に忠実に、多体相関を含みつつ量子力学的に計算する。量子化学における配置換相互作用計算と類似した手法である。計算は大次元行列の固有値問題に帰着する。その解法としては行列の対角化に基づく従来型の方法と、重要な多体状態の基底を探す方法の2種類がある。
原子軌道基底	分子軌道を表現するための関数群。原子軌道を表す関数の線形結合で分子軌道を表現。
元素合成	宇宙はビックバンにより始まったと考えられている。現在の宇宙に存在する元素は、ビックバン直後には存在しなく、ビックバン後、宇宙が広がり温度が下がるにつれてまず、水素とヘリウムが生成されたと考えられている。一方それ以外の元素は星の形成などによる核融合反応や超新星爆発などに伴う中性子捕獲によって合成されたと考えられている。このように元素が宇宙進化とともにどのように合成されてきたかを考え、明らかにすることは現代の物理学の課題の一つである。
元素の起源	現在の宇宙の元素の組成は、ほぼ水素とヘリウムで構成されおり、そのほかの元素の量は無視できるほどである。宇宙誕生後の物質進化の過程を追うことで、さまざまな元素の組成比を理解することが元素の起源を探ることである。
ゲージ群	力学系の作用が余分な自由度をもち、その自由度に対して変数変換しても作用が不変な場合がある。このような変換をゲージ変換といい、これは一般に群をなす。これをゲージ群という。ゲージ変換のもとで不变な理論をゲージ理論と呼ぶ。この場合余分な自由度は観測にからない。
ゲート	Hodgkin-Huxley モデルの中で電流の開閉を司るスイッチの役割を果たす仮想概念
こ	

(次ページに続く)

用語	解説
格子 QCD (格子量子色力学)	QCD はクォークとグルーオンの強い力の力学であるが、解析的に解くことはできていない。数値的に QCD を取り扱うことができるよう、4 次元時空を格子に差分化した理論が格子 QCD である。100TFlops クラスの計算機が登場した 2008~2009 年に、クォークの複合粒子である陽子や中性子などの性質 (質量や спинなど) を計算で再現できるようになった。
格子気体法	流体問題を空間と流体の両方を離散化して解く計算手法
高次キュムラント	この文脈では重イオン衝突実験で生成される荷電粒子の位相空間における分布のガウス分布からのズレを表す量。
拘束付平均場	ある量が決まった値になるように条件を付けながら計算をする平均場理論。
構造緩和	最初に仮定した物質の構造 (=原子の配置) を原子に働く力が小さくなる方向に原子を動かすことでもっとも安定な構造に近づけること
構造多型	タンパク質などの巨大分子が複数の安定な構造を持つ性質
構造ゆらぎ	タンパク質分子が機能を発現させるためにその構造を変化させること
光分解	原子核が光 (ガンマ線) を吸収して分解する反応。
高立体選択的合成反応	複数の立体異性体 (配位子の付き方が立体的に異なる分子) の生成が考えられる化学反応において、触媒の利用などにより特定の立体異性体を選択的に多く作り出す反応のこと。
呼吸鎖	細胞の呼吸 (ATP の生成) に関わるタンパク質群
骨格振動	2 重結合や芳香環などの分子構造に起因する特徴的な振動。赤外やラマンで分光測定することにより、対象分子系の分子構造を推定できる。
混雑物	分子混雑環境において溶存するタンパク質、DNA、RNA、糖をはじめとする様々な分子
コンダクタンス	電気伝導度。すなわち抵抗の逆数
ご	
誤差逆伝播法	ニューラルネットの学習法。ニューラルネットの出力と教師信号との誤差を、ネットワークの上から下に伝播させ学習する手法。
さ	
再帰現象	相互作用する多数の粒子の運動において以前と同じ状態が準周期的に現れる現象
細胞環境	細胞内分子にとっての環境。分子が溶液中にあるときの環境と異なり、多くの分子で混み合っている。

(次ページに続く)

用語	解説
サブボリューム	並列計算において、1プロセッサが担当する部分領域。なお、シミュレーションセルを空間分割して個々の並列プロセッサに割り当てる手法を領域分割法と呼ぶ。
差分法	微分方程式を数値的に解く際に用いられる離散化手法のひとつ。ある関数が2つの変数値に対してとる値の差を差分といい、この差分を変数値の差で割って得られる商を差分商と言い、この差分商を用いてもとの微分の近似値として偏微分方程式の離散化を実現する。
酸化還元電位	物質の電子の放出しやすさ、あるいは受け取りやすさを表す数値。
参照曲率	計算要素内で形状を表現する時に用いるパラメータの一つ。形状の曲率の事。
三相界面	燃料極と固体電解質、空気極の三相が接する境界面。その長さが燃料電池の反応性を左右する。
散乱理論	量子力学において、非束縛状態を取り扱う理論を一般に散乱理論という。原子核物理では、核反応理論の基礎となるとともに、核子間相互作用や共鳴状態（極めて短時間で崩壊する準束縛状態）の解析にも用いられる。
散乱・束縛状態	2粒子以上の系において、各々の粒子の運動が有限の範囲に限定されるものを束縛状態、無限遠方まで許されるものを散乱状態という。
ざ	
材料強度発現機構	材料の破壊を発生・進行させるメカニズム。その破壊挙動は、主に材料内の力学場と材料固有の強度との相関により決定される。
残基	タンパク質、核酸、多糖類などの重合体を構成している単量体
し	
システムリスク	個別の金融機関の支払不能等や、特定の市場または決済システム等の機能不全が、他の金融機関、他の市場、または金融システム全体に波及するリスクのこと。（日本銀行サイト https://www.boj.or.jp/announcements/education/oshiete/kess/i06.htm/ より）
システムインパッケージ	英語で system in a package のことで、1つの Package の中に複数の半導体チップを集積することにより、システムレベルの高度な機能を実現して、実装密度の向上とコストダウンを実現する技術。
システム生物学	生命現象をシステムとして理解することを目指す学問分野
質量異常次元	エネルギー・スケールの変化に対する質量の振る舞いを記述し、相互作用による効果を表す。
シナプス	ニューロン同士が結合する部分。
シナプス可塑性	シナプスの結合の強さは、シナプスを形成する2つの細胞の活動に応じて動的に変化する。その機構。

(次ページに続く)

用語	解説
シナプス遅延	シナプス前末端でカルシウム濃度が閾値を超えてからシナプス後膜でシナプス後電位が発生するまでの遅延
シフト型通信	各プロセスが隣接する他プロセスに対して、一斉に一定方向のデータ送信をする通信形態をいう。
シミュレーションセル	シミュレーションの中で考慮する空間領域
社会ネットワーク	社会経済現象をその構成要素の間の関係性に注目して構成したグラフのこと。格子などの規則的な構造や密結合の完全グラフとは大きく異なった構造をもつことが多い。
主殻	調和振動子ポテンシャルによる一粒子軌道によって空間を展開した際に、縮退した一粒子軌道の集合を指す。
出力密度	1 kgあたりどのくらいの出力を出せるのか、電池の性能を表した数値。
触媒反応	化学反応の際に、それ自身は変化せず、他の物質の反応速度に影響する働きをする物質(触媒)により引き起こされる反応。触媒反応には、触媒が反応によって変化する物質に混ざり合わない不均一触媒反応と、変化する物質に混ざる均一触媒反応がある。いずれの反応の場合も化学反応の平衡を変えることなく、正逆の反応速度をそれぞれ同じだけ変える。
ショットガン法	ゲノムDNAを断片化し読み取りそれを計算機を用いてつなぎ合わせることにより染色体の連続したDNAを読み取る方法
シリコン結晶型太陽電池	普及が最も進んでいる太陽電池であり、シリコン結晶を材料とする。結晶シリコンのうち単結晶系は、効率は高いがコストも高い。最も使用されているのは、多結晶系(小さな結晶がモザイク状に集まつたもの)。効率は低くなるが、生産コストをおさえることができる。
真空偏極	真空における粒子・反粒子の対生成・対消滅過程。
神経成長因子	特定の細胞の神経細胞への分化を促進する因子となる分子
震源過程	地震は、発生源で断層が破壊されることによって生じる。この断層の破壊過程を、震源過程という。
信号情報処理のマルコフ過程	一個一個のイオンチャネルの挙動やレセプタとリガンドの結合はリガンド濃度や電圧などに対して確率的に挙動する
深層学習	多層ニューラルネットワークを用いた機械学習。隠れ層を持たないペーセプトロンや、1層の隠れ層をもつニューラルネットワークとの違いを表すためにこの言葉を用いることが多い。Convolutional Neural NetworkやRecurrent Neural Networkの他、確率関数で層同士の関係を記述するボルツマンマシンなどがある。
じ	

(次ページに続く)

用語	解説
磁気回転不安定	差動回転(天体の各部分で異なる角速度を持つ回転)する磁気流体に起かる不安定性。通常の天体では内側の物質ほど角速度が大きい。内側の物質は角速度が大きいため、外側の物質に先行する。しかし、磁場を通して内側の物質と外側の物質はお互いを引っ張りあう。すると内側の物質は一旦減速し、外側の物質は一旦加速する。内側の物質は減速すると、天体の重力に引っ張られてさらに内側に落下する。内側ほど角速度は大きいため、内側の物質は結局減速前よりも大きい角速度を持つことになる。外側の物質はこれとは逆に加速前よりも小さい角速度を持つことになる。すなわちこの不安定は、内側と外側の物質の角速度差がどんどん大きくなる不安定である。
自己相関時間	系を時間発展させてサンプリングする際、ある時刻でのサンプルと、それとは独立と考えられる次のサンプルを採取するまでに要する時間。
自己無撞着	セルフコンシスティント(self-consistent)。
次世代シークエンサー	DNA を 100 塩基程度と非常に短く断片化し、それを並列に処理することにより高速に読み取ることのできる装置。読み込んだ DNA は断片であるため部位の特定のため計算機を用いた参照配列との照合に多量の計算が必要である。
実験分光データ	分光法で得られたスペクトルデータ。実験で得た分光データの既知試料との比較などで未知試料の同定を行う。
自発的対称性の破れ	系が本来持つ対称性の一部が自ずと破れて、より対称性の低い状態に系全体として落ち込むこと。この概念は相転移と密接に関連しており、たとえば水(液体)から水(固体)への変化は水分子の併進対称性が失われることとして理解される。
重イオン	陽子、ヘリウムなどの軽い原子核を除く、重い原子核のことを指す。電子をはぎ取った原子なのでイオンと呼ぶ。
重イオン衝突実験	RHIC や LHC において重元素の原子核同士を互いに衝突させて、高エネルギー・高温度の状態を作る実験。原子核としては、金原子核や銅、鉛原子核などが用いられている。
重合脱重合化	同種の分子が結合してより大きな構造を取ったり結合を解くこと
充放電曲線	縦軸に電池電圧、横軸に充電状態 0 % (完全放電状態) から 100 % (完全充電状態) をとって、同じ電流での充電と放電の状態推移をグラフ化したもの。充電や放電のエネルギーおよび両者の差であるエネルギー損失を見積もることが可能。
重陽子	陽子と中性子の束縛状態。二重水素の原子核。

(次ページに続く)

用語	解説
重力の量子化（量子重力）	素粒子標準理論の中の相互作用を記述する部分は、量子力学の原理に則り量子化され、ミクロな世界での物理を矛盾なく記述できている。一方で、重力理論であるアインシュタインの一般相対性理論を量子力学の原理に則り量子化しようとすると、うまくいかない。一般相対性理論や何らかの重力の理論を量子力学と矛盾なく量子化すること。宇宙そのものの誕生時を理解するためには、量子力学が必要なミクロな世界での重力を理解する必要があるため、重力の量子化は理論物理学の長年の夢であるがまだ実現していない。超弦理論がその候補とされている。
準粒子	相互作用している多体系を、近似的に自由に運動するある種の「粒子」の集まりとして記述することができるとき、この「粒子」を準粒子とよぶ。
状態空間モデル	時系列観測データのモデル化の方法の一つでデータを状態モデルと観測モデルに分離し記述する
状態方程式	物質の温度、圧力、エネルギー、密度、体積などの間に成り立つ関係式。
助触媒	触媒の効果を高める添加剤。触媒には化学反応を促進あるいは抑制する効果があるが、助触媒は単独で存在してもそのような効果を持たない。
人工光合成	光合成を人為的に行う技術。自然界での光合成は、水・二酸化炭素と、太陽光などの光エネルギーから化学エネルギーとして炭水化物などを合成するものであるが、広義の人工光合成には太陽電池を含むことがある。現在、光合成のような多段階電子移動については部分的に成功しているが、天然光合成のように電荷分離状態をいかに効率よく生成し、一方で逆電子移動による失活を防ぐかという問題が残されている。
人工市場	コンピュータ上に、人工的に作り出された架空の市場。計算機上に仮想的なディーラーの役割をするコンピュータプログラムを作り、これらのエージェントが自由に取引を行う仮想市場において実験を行う。(和泉潔「人工市場 市場分析の複雑系アプローチ」森北出版、2003 より)
す	
水平乱流	流れが乱れた状態（流体の粘性力に対して流れの慣性力が大きい状態）を乱流と呼ぶ。二次元（水平）乱流とは、大気のように成層が強い場で、鉛直方向の運動が制限され、水平方向の運動が卓越する状態を指す。水平乱流場においては、物質は水平方向に拡散される。
数値求積法	非解析的、近似的に積分値を求める手法。ガウス求積などの積分区間を区切る手法や、乱数を用いるモンテカルロ積分などがある。方法によって求積点数と誤差の関係が異なる。
スカイライン形式	疎行列に対するメモリ格納形式の一つで、バンド形式をより精緻化し、境界の輪郭線を行単位で正確になぞるようにしたもの。

用語	解説
スケール間相互作用	気象や気候現象に存在する複数の様々な時空間スケール(例えば、全球スケールや温帯高低気圧のスケールなど)の現象が相互に影響を及ぼしあっていること。
スタッガード型	格子上で定義されたクォーク作用の一つ。
ステップスケーリング	エネルギーを s 倍(典型的には $s=2$)ずつ不連続に変化させながら、結合定数などのエネルギー依存性を調べる数値計算手法。
ストレンジクォーク	標準模型に含まれる素粒子には6つの質量の異なるクォークがある。粒子質量の軽い順からアップ、ダウン、ストレンジ、チャーム、ボトム、トップと名前が付けられている。標準模型では質量以外の性質は同じである。ストレンジクォークは3番目に軽いクォークである。
ストレンジネス	ストレンジクォークが関与する量子数。正確には、ストレンジクォークの数とその反粒子の数の差。
ストロング・スケール	並列化の指標。計算量とcpu数が両方増えて行く時の計算効率。
ストークス力学	流れの状態を示すレイノルズ数が小さな場合に、流れを近似方程式で示す事が出来、これをストークス方程式と呼ぶ。近似方程式では非線形項である対流項を無視している。
スパイク	デルタ関数状の時間的に短い電気パルス。ニューロンが十分強く活動(脱分極)すると軸索から発射される。神経回路における情報のキャリアであると考えられる。
スパイク列	複数の活動電位が連続して出る様
スピン	素粒子または素粒子で構成される量子力学的系の有する基本的な量の一つでその静止系での系のもつ角運動量。原子スペクトルおよび元素の周期律を説明するため、1924年にW.パウリが初めて導入した。
スピニ液体	量子力学的な揺らぎや幾何学的フラストレーションの効果により、磁気モーメント間の集団的な秩序化が絶対零度まで妨げられた状態。
スピントロニクス	エレクトロニクスが物質中の電子が持つ電荷自由度だけを利用していたのに対し、スピン自由度も工学的に応用する技術。
スペクトル法	物理現象を表す偏微分方程式の時間積分法の一つで、物理変数の時間変化を直接計算するのではなく、周波数空間に置き換えて計算する手法。一般に高精度な解が得られるため基礎的な物理計算によく用いられるが、複雑な問題には対応が難しいとされている。
スーパーBファクトリー	電子と陽電子を高頻度で衝突させることによってボトムクォークを含むハドロンを大量に生成し、その崩壊を詳細に調べることを目的とした加速器。従来のBファクトリーの数十倍のルミノシティを目指す。

(次ページに続く)

用語	解説
スーパーセル	結晶中にとる事のできる周期セルのうち、基本セルよりも大きい物。基本セルよりも大きな空間スケールの構造揺らぎの表現に用いる。
せ	
正準化変換	Hartree-Fock 方程式を解く際に非直交基底関数の組を変換して規格直交系を作る手法のひとつ。
静的縮約	連立一次方程式において、自由度の一部を削除することで、係数行列のサイズを縮小する方法。スタティック・コンデンセーション。
世界線表示	量子力学に従う系は空間次元に加えてもう一つ虚時間と呼ばれる軸を導入することで、計算機が扱いやすい複素数での計算が可能となる。その際、系の状態が虚時間方向でどのように発展するかをグラフ的に表現することを世界線表示と呼ぶ。
積分発火モデル	細胞外に抵抗と容量で接続された点として考え、シナプス後電流が複数の別の入力細胞からはいったとき、その時空間的統合としての細胞電位が閾値を超えたときに活動電位が起り、結果過分極側に一定量電位がシフトすると考えるモデル。英語は Integrated-and-Fire model
セグメント	高分子を構成する基本単位のこと。
摂動的 QED(量子電気力学)	QED を摂動論を用いて解析すること。
摂動マッチング	有効理論に基づく計算では、まず、有効理論のパラメータを固定しなくてはいけない。理論的には、基礎理論に基づく計算と有効理論に基づく計算が有効理論が妥当な条件のもとで一致するように有効理論のパラメータを固定する。このような過程をマッチングという。量子場の理論ではこのようなマッチングを摂動理論を用いて行うため、摂動マッチングと呼ぶ。
セパレータ(電池セパレーター)	電池の中で正極と負極を隔離し、かつ電解液を保持して正極と負極との間のイオン伝導性を確保する重要な材料。
セミレプトニック崩壊の形状因子	ハドロンが崩壊する場合にハドロンとレプトンに壊れる場合をセミレプトニック崩壊という。崩壊時の娘粒子の軌跡やエネルギー依存性を因子化したものを形状因子という。
線型応答理論	熱平衡状態にある系に、磁場や電場などの外場が加わった時、その外場による系の状態の変化(応答)を扱う理論。
そ	

(次ページに続く)

用語	解説
相対性理論	アインシュタインによって提案された時空の幾何学。特殊相対性論と一般相対性理論の両方かどちらかを指す。特殊相対性論は異なる慣性系の間の関係を時間空間の幾何学で説明する。一般相対性理論は重力の理論であり物質のエネルギーが時空の幾何学を決定する理論。時空の幾何学を重力とみなす。星の重力を決定するだけでなく、宇宙全体の幾何学をも決定でき、宇宙物理学における基礎となる理論。量子力学が重要となるミクロの世界での重力の振る舞いについては記述できない。
相対論的流体	相対性理論の枠組で扱う必要がある流体。速度が光速近くに達する流体や、中性子星のよう強重力場中の流体などがこれに対応する。
相変態	ここでは、固体電解質材における結晶構造の変化。
素過程	複雑な自然現象は、様々な物理(電磁気学、熱力学、流体力学等)が絡み合って生じている。しかし、少くない現象においては、関わる物理をいくつかの構成要素に分割し、その要素間の相互作用として記述することが可能である。そのような構成要素のうち、特に基本的な物理で比較的単純に数学的に表現することができるものを素過程という。たとえば、流体力学で支配される移流(力学)過程、放射伝達方程式で支配される放射過程などがそのような素過程である。
阻害活性	化合物が標的タンパク質の機能を阻害する性質
粗視化分子動力学法	複数原子からなる集団を一粒子とみなしその群としての運動をシミュレートする手法。計算量の減少を通じて大規模で長時間の分子シミュレーションを可能とする。
粗視化モデル	原子のグループをまとめて、一つの相互作用点として表し、相互作用数を大幅に減らしたモデル。たとえば、タンパク質のアミノ酸を一つの相互作用点を近似する粗視化モデルなどがある。
塑性加工解析	金属部品の成型プロセスにおける材料加工処理のシミュレーション。この際に大変形弾塑性解析を行う必要がある。
袖領域	差分法等のステンシル計算では隣接する要素、格子上のデータを参照する。このため、計算領域を分割して並列処理を行う際に、隣接ノードの境界データを保持する。この境界データを袖領域という。
素反応	大部分の化学反応は途中に反応中間体を生じる多段階の反応である。このような反応は複合反応と呼ばれ、いくつかの単純反応の組み合わせとして記述できる。この複合反応の各段階を構成する単純反応を素反応という。

(次ページに続く)

用語	解説
ソフトウェアパイプラインング機能	コンパイラの最適化機能の一つ。ループ内で繰り返される一連の CPU の処理命令を 1 サイクルに 1 つずつ実行するのではなく、複数の処理命令を並列実行することで処理速度を向上させる。
ソリッド要素	構造解析において、連続体をそのまま表現するための有限要素。形状としては、四面体あるいは六面体などの形を有する。これとは別に、梁やシェルなどを表現するための構造要素がある。
素粒子標準理論（標準理論）	自然界の物質を構成する素粒子の運動と、素粒子間の相互作用を記述する法則をまとめた理論。素粒子としては、クォークと呼ばれる 6 種類のスピン 1/2 のフェルミオンと、レプトンと呼ばれる 6 種類のスピン 1/2 のフェルミオンが含まれる。相互作用は電磁気力、弱い力、強い力の 3 つの相互作用を媒介する 4 種類のスピン 1 を持つボソンが含まれている。電磁気力と弱い力を分化させ、素粒子に質量を与えるヒッグス粒子と呼ばれるスピン 0 のボソンを含む。量子力学と矛盾しないように作られている。実験との比較でしか決まらない 18 (+ α) 個の独立パラメータが含まれる。重力はここには含まれない。
素粒子標準理論に内在するエネルギーースケールの階層性	素粒子標準理論のパラメータには以下のような階層性がある。(1) 弱い力の媒介粒子の質量が重い。(陽子の約 80 倍と約 90 倍) クォークやレプトンの質量はバラバラであるが、(2) トップクォークは特に重い。(陽子の約 170 倍) (3) ニュートリノの質量が直接測定不可能なぐらい軽い。質量はエネルギーと等価であるので、エネルギーースケールの階層性という。このような階層性の起源を明らかにすることは現在の素粒子物理学の課題である。
た	
大域構造	原子同士が直接触れ合うような短い距離でみられる構造ではなく、多数の原子の集団同士の関係が作り出す長い距離で特徴づけられる物質の構造のこと。
対称正定値	行列が対称かつ、その固有値がすべて正值であること。この性質を有する行列はより効率的に扱うことができる。

(次ページに続く)

用語	解説
対称性の破れの機構	基礎理論には対称性があるが、現実の系ではそれが破れているように見えることがしばしばある。たとえば、磁石は高温では磁石となっていないが低温に冷やすと一つの方向に磁場をもつ。巨大で一様な磁石があつた時、低温では磁場の方向が特殊な方向であり、高温では磁場がないので特別な方向は何もない。これは高温での3次元の回転対称性が低温では磁場の方向の周りの回転対称性である2次元の回転対称性に対称性が減ったことになる。このように対称性が減る場合を対称性が破れたといふことがある。磁石の場合の対称性の破れは、ミクロな世界における原子の磁化が原子間でどのように相互作用しているかで記述されるものである。このようにミクロな世界での物理機構から、マクロな対称性の破れを議論することが、対称性の破れの機構を調べることとなる。
太陽電池	光起電力効果を利用し、光エネルギーを電力に変換する電力機器。光電池（Photovoltaic cell）とも呼ばれる。一般的な一次電池や二次電池のように電力を蓄える蓄電池ではなく、光起電力効果によって光を即時に電力に変換して出力する発電機である。タイプとしては、シリコン太陽電池の他、様々な化合物半導体などを素材にしたもののが実用化されている。最近では、低い変換効率ではあるものの柔軟で生産コストの低い有機薄膜型、高変換効率を見込めるペロブスカイト型や、量子ドット型の太陽電池の開発も進んでいる。
タイルドディスプレイ	高解像度の表示領域を確保するため、複数のモニタを並べて配置したデバイス。通常、クラスタシステムなどで動作する。
多参照理論	電子の波動関数を表すために、複数の電子配置の重ね合わせを用いる理論。分子の解離状態などでは、単一のSlater行列式では良い波動関数が表現できず、多参照理論が必要となる。
多次元効果	対称性（球対称や軸対称など）を仮定し次元を落としたシミュレーションでは現れない現象。例としては対流などがある。
多重格子法	ポアソン方程式を格子で離散化して反復法で解くような場合には、基本的に格子サイズ程度の短い波長の誤差が効率良く減衰する一方で、長波長の誤差はなかなか減衰せず、これが反復回数増大の原因となる。多重格子法は、格子サイズの異なる複数の格子を用意し、各波長の誤差を一様に減衰させることで反復回数の増大をおさえる数値解法である。
多体問題	複数の粒子や複数の物体の間の力学を解く問題。

(次ページに続く)

用語	解説
探索	迷路を解く経路や最適な組み合わせなどを求める際に用いられるアルゴリズムの総称。経路の探索で古典的なものとして、ダイクストラ法やA*などが知られている。統計的機械翻訳における出力単語列の探索などではマルチスタッキーム探索、囲碁などのゲームにおける手順の探索においてはモンテカルロ木探索がしばしば用いられる。
タンパク質の折れたたみ	タンパク質がある一定の立体構造をとる過程
だ	
第0近似的	実際の現象を細部まで捉えられてはいないが本質は捉えられている様子をいう。
第一原理計算	電子シュレディンガー方程式を（半）経験的パラメータによる積分の近似を用いないで数値的に解く計算手法。化学では、非経験的計算とも呼ばれる。
第一原理ダウンフォールディング法	第一原理計算を用いて対象とする物質の個性を残しつつ注目するエネルギー・スケールに応じた有効模型を構築すること。得られた模型をより精緻な計算手法で解析することで非経験的かつ高精度な物性値の計算が可能となる。
第一原理電子状態計算	電子シュレディンガー方程式を（半）経験的パラメータによる積分の近似を用いないで数値的に解く計算手法。化学では、非経験的計算とも呼ばれる。
大規模疎行列連立方程式 (反復解法)	ここでは変数の数が数千万から数十億程度の連立線形方程式を想定している。このような方程式を数値的に解く方法としてKrylov部分空間反復法が用いられる。
大規模連立線形方程式	ここでは変数の数が数千万から数十億程度の連立線形方程式を想定している。
大振幅集団運動	多数の核子が一斉にある秩序を持って運動することを集団運動とよぶが、特にその運動の振幅が大きく、物理学で良く使われる調和近似などが適用できない集団運動を大振幅集団運動とよぶ。
大統一理論	自然界の4つの基本的な力である電磁力・弱い力・強い力・重力のうち、電磁力と弱い力の統合(電弱統一理論)に加えて強い力をも統合する理論。
ダイナミカル行列	結晶内の原子の相互作用を記述した行列。
ダイナミクス	一般に力学過程や力学の仕組みのこと。

(次ページに続く)

用語	解説
タイリング	計算機上で、大規模なデータを配列の添え字ごとに細かく区切り、小さな部分配列 (=タイル) の集合として扱うこと。行列のような2次元のデータ配列をタイリングすると、四角形のタイルを敷き詰めたようなイメージになることから。
脱閉じ込め臨界現象	相転移でありながら、ランダウが提唱し相転移の標準的な起源として知られる「自発的対称性の破れ」の範疇に入らず、実在すれば教科書を書き換える発見になるとして注目されている、新しいタイプの臨界現象。
弾塑性構成則	主として金属材料の挙動を、応力とひずみの関係から記述する際の関係式を指す。金属材料は変形初期の段階では応力とひずみに線形関係がある、いわゆる線形弾性体であり、ある限界を超えると非線形な塑性挙動を呈するようになる。その限界値と、非線形挙動を応力ーひずみ関係として記述したものである。
ダークエネルギー	宇宙の大規模構造（銀河や物質の分布の様子）や運動（宇宙は現在膨張している）の記述は一般相対性理論にもとづいたビックバン理論で説明されている。今現在の宇宙は膨張しているがその膨張速度が少しづつ早くなっていることが我々の近傍の銀河の運動の詳細からわかっている。この加速膨張はあたかも反発する重力によって引き起こされているように見える。このような反発する重力を説明するために導入されたのがダークエネルギーの考え方である。AINSHUTAIN方程式に宇宙項として導入されたものもダークエネルギーの一つの候補である。
ダークマター	暗黒物質とも呼ばれる電磁気力と強い力が作用しない仮説上の物質。電磁相互作用しないので、地上実験や天文観測では直接検出できない。ダークマターはエネルギーを持ち重力に影響を及ぼすことから、ダークマターによる重力レンズ効果や、銀河の回転運動の検証などで間接的にその存在が推定されている。シミュレーションにより、宇宙の大規模構造の生成にも重要な役割をしていることが分かっている。近年の WMAP (Wilkinson Microwave Anisotropy Probe) 衛星による観測から、ダークマターは宇宙全体のエネルギーの内、約 20% を占めていると考えられている。素粒子標準理論にはダークマターに該当する粒子はない。
ち	
チェックポイントファイル	計算の途中の状態を保存するファイル。万一計算が計算機の故障で中断した場合、このファイルから計算を継続実行できる。

(次ページに続く)

用語	解説
地球温暖化	気候変動の一種で、温室効果ガスによって地球表面の大気や海洋の平均温度が長期的に上昇する現象。ここで温室効果とは、大気圏を有する惑星の表面から発せられる放射（電磁波により伝達されるエネルギー）が大気圏外に届く前にその一部が大気中の物質に吸収されることで、そのエネルギーが大気圏より内側に滞留し、結果として大気圏内部の気温が上昇する現象。温室効果ガスとは、大気圏にあり、地表から放射された赤外線の一部を吸収することにより温室効果をもたらす気体の総称であり、対流圏オゾン、二酸化炭素、メタンなどが該当する。
蓄電・水素エネルギー社会	二酸化炭素を排出しないエネルギー源である水素を活用する社会システム。水素や生成した電気エネルギーを蓄電することで、必要に応じてエネルギーを適宜柔軟に使用できる形態が理想的とされる。
地磁気異常の縞模様	海洋底の地磁気を調べて標準より強く帯磁している所を黒く塗ると海嶺と平行な縞模様が海嶺から両側に全く対称的に現れる。この縞模様は、海洋底が海嶺から湧き出して冷却する時に記憶する地球磁場が、その当時の地球磁場を反映して反転を繰り返しているためと説明される。
チャネル結合	核反応において、終状態（あるいは始状態）における放出粒子と残留粒子（あるいは入射粒子と標的粒子）の組の一つ一つをチャネルと呼ぶ。チャネル結合とは、各チャネルを基底にとった連立一次方程式のシュレーディンガー方程式として核反応の基礎方程式を表す方法で、チャネル間の遷移を非摂動的に取り扱うことができる。
チャネルロードプシン	緑藻植物のクラミドモナスなどがもつ色素たんぱく質で、光が当たるとイオンを透過する。
中間子	パイ中間子やオメガ中間子などがある。1つのクォークと1つの反クォークが強い力で結合してきた粒子の総称。中間子にはいろいろな種類があるが、それらは2個のクォークの組合せによる違いや内部の状態の違いで理解されている。
中重核	「軽い原子核、重い原子核」の項目を参照。
中性子過剰核	陽子数に比べて過剰に多い中性子を含む原子核。不安定であり、安定な原子核になるまで中性子から陽子への β 崩壊を繰り返す。
超音波治療機器	超音波の集束などにより、熱エネルギーに変換し幹部に直接治療を行う機器

(次ページに続く)

用語	解説
超弦理論	素粒子は点であるとして考えているが標準理論には多数の素粒子とパラメータが含まれており複雑である。一方素粒子を点と考えず紐（＝弦）と考え紐の振動の様子の違いで複数の素粒子を一つの弦で統一的に記述することができないかという考え方から弦理論というものが作られている。とくに超対称性を持つ弦理論を超弦理論とよぶ。標準模型に現れる粒子群に加え重力を含む超弦理論も構成できるため万物の理論の候補と考えられているが、非摂動的計算手法や理論の安定性など不明なところも多い。
超弦理論のゲージ重力 双対性予想	超対称性を持つゲージ理論と重力理論が等価関係にあるとする予想。この予想を使って、ゲージ理論の計算から重力理論の性質を調べる、またはその逆を調べる研究が行なわれている。
超新星爆発	大質量恒星の進化（一生）の最後に起こる爆発的現象。太陽質量の10倍より重い質量の恒星は、熱核融合反応により恒星の中心部に鉄の芯が形成される。鉄は熱核融合を起こさないため重力による収縮が起り、鉄コアの温度が上昇していく。ある温度で鉄原子核はヘリウムや核子に分解する吸熱反応を起こし、恒星外層部の物質が中心に向かって急速に落下（重力崩壊）し中性子の芯が形成される。外層部からさらに物質が中性子の芯へ落下ってきて中性子の芯に跳ね返され衝撃波が生じる。この衝撃波が恒星外層部を吹き飛ばし、超新星爆発を引き起こすと考えられている。計算機シミュレーションによる超新星爆発の再現は爆発機構の解明につながる重要な課題である。
超重元素の合成	原子番号92のウランを超える元素は天然に存在しないが、原子炉や加速器によって人工的に合成することは可能である。こうした超重元素と呼ばれる極めて原子番号の大きな新元素を作り出すと、元素の命名権が与えられるため、世界的に競争が激化している。
超対称性	ボゾンは整数スピンを持つ粒子であり、フェルミオンは半整数スピンを持つ粒子である。それらを入れ替えるような操作を超対称性変換と呼び、その変換に対して理論が不变であるとき、その理論は超対称性を持つという。
超対称性粒子/超対称性理論	超対称性理論では、標準理論に登場するすべての素粒子に対してペアとなる超対称性粒子を置く。標準理論のボーズ粒子に対してはフェルミ粒子、逆に標準理論のフェルミ粒子に対してはボーズ粒子が追加される。
超対称ヤンミルズ理論	ヤンミルズ理論はゲージ理論のこと。超対称ヤンミルズ理論はそれを超対称化したものである。

(次ページに続く)

用語	解説
超大規模並列	並列計算機の中で特に規模が大きいもの（莫大な数の CPU をもつもの）をいう。多数のノードを用いて並列実行が可能なため、計算速度の向上が見込まれる。
超流動核	核子（陽子・中性子）がクーパー対を作ることでボーズ凝縮し、低温の液体ヘリウムのように超流動性を示す原子核。低温の金属における超伝導と類似した現象。
超流動固体状態	固体秩序と超流動秩序が共存した状態
調和振動子	kx^2 のポテンシャルの中で運動する振り子、または量子。
つ	
対超流動	2つのボーズ粒子のペアからなるボーズ粒子の示す超流動現象。
通信マスク手法	非同期通信や通信用スレッドの実装によって演算処理の背後で通信処理を同時に実行する手法。
強い力	素粒子標準理論では、すべてのクォーク間に平等に働く力。電磁気力に比べ 100 倍強い。クォークの間でグルーオンと呼ばれるボソンが交換されることで力が作用しあうと考える。強い力では、3 種類のクォークを強固に一つにまとめる場合と、1 種類のクォークと 1 種類の反クォークを強固に一つにまとめる場合がある。陽子や中性子は 3 種類のクォークからなる複合粒子であり、パイ中間子はクォークと反クォークからなる複合粒子である。
て	
低侵襲治療	手術などに伴う痛み、発熱、出血などをできるだけ少なくする医療
低次元構造体	一次元もしくは二次元の周期的な結晶格子構造を持つ原子構造体。一次元の例としてはナノワイヤー、二次元の例としてはグラフェン等のナノシートが挙げられる。
低レイテンシ	通信の際に、データ転送などを要求してから、実際に送られてくるまでの遅延時間のことをレイテンシ（遅延）と呼ぶが、その遅延時間が短いこと
テクニカラー理論（ウォーキングテクニカラーラー）	標準理論を超えたモデルの一つで、ヒッグス粒子を複合粒子として考える。このモデルが妥当であるためには QCD に似た性質を持ちつつも相互作用の強さの性質が QCD と違った特徴を持つ必要がある。
テクニ中間子	テクニカラー理論において予言される複合粒子の一種。

(次ページに続く)

用語	解説
鉄よりも重い重元素の起源	恒星内部での熱核融合反応では、水素から始まる核融合反応は発熱反応であり、水素よりも安定な重い元素を合成する方向に進む。十分重い恒星では重力収縮と熱核融合反応の連鎖により、鉄原子核でできた恒星芯が形成される。しかし、鉄原子核まで合成が進むと鉄は最も安定な原子核であるので、発熱反応が終わり恒星芯での熱核融合反応は終了する。鉄よりも重い元素は恒星内部での熱核融合反応による元素合成では生成されず、中性子捕獲反応で生成されたと考えられる。超新星爆発は金やプラチナなどの、鉄よりも重い重元素の起源の一つと考えられている。
転位動力学法	結晶中の線状欠陥である転位の運動をシミュレートする手法。塑性を支配する、結晶すべり挙動の解析に用いられる。転位に働く力をモデル化することで、古典分子動力学法に比して大規模かつ長時間のシミュレーションが可能となる。
展開係数	ある関数を基底関数の線形結合で表した際の、それぞれの基底関数のもつ重み。
点欠陥	結晶中の不純物または空孔。
転写因子	DNA に特異的に結合し、DNA の遺伝情報の RNA への転写を促進、あるいは逆に抑制するタンパク質
テンソル縮約	二つのテンソル量を掛け合わせて、新しいテンソル量を得る操作。行列どうしの掛け算も、テンソル縮約の一つである。
テンソル力、3体力	核子(陽子・中性子)間に働く相互作用(核力)は、核子間の相対距離だけに依存する中心力とそれ以外の非中心力に分類できる。非中心力の代表がテンソル力(tensor force)。2核子のスピンの向きと配置の向きに依存する。これら2核子間の相互作用に加えて、原子核の定量的な記述には、3核子に働く3体力が不可欠であることが知られている。
天体降着流・噴出流	天体に向って落下する物質の流れと、天体から放出される物質の流れ。一般に中心天体に向かって円盤状に降着し、回転軸方向にビーム状に放出される。
転置転送	行列の転置操作に用いられるデータ転送パターン。
で	
電解質	溶媒中に溶解した際に、陽イオンと陰イオンに電離する物質。溶媒とは、他の物質を溶かす物質の呼称。
電荷移動型ポテンシャル	分子動力学法で用いる、電荷の局所的な移動を考慮した原子間ポテンシャル。原子における電荷の偏りを取り入れることで、電気陰性度の異なる異種原子間の結合を高精度に表現できる。
電荷分離	光照射によって半導体中に生じた電子と正孔が空間的に分離すること。

用語	解説
電気生理	細胞の電位を測定することによって生理的な性質を知ろうとする体系
電極	真空または誘電体中に電場をつくる為、或は電気伝導体中に電流を流す為に設けられた導体。
電源グラウンドバウンスノイズ	LSI 内部の回路動作に伴う電流の時間変化に起因して発生する LSI 内の電源及びグラウンド配線部分の電圧ノイズ。
電子供与体・受容体	他の分子に電子を渡しやすい性質をもった分子（供与体）、あるいは他の分子から電子を受け取りやすい性質をもった分子（受容体）。
電子相関	ある電子が他の電子の位置のゆらぎにより受ける効果
電子捕獲	原子核が電子を捕獲することで、陽子を中性子に変換させる反応。ニュートリノが放出される。
電子ボルト	エネルギーの単位。約 1.6×10^{-19} J. これは真空中において 1 ボルトの電圧差で生じる一様電場の中で電子を加速した場合に電子の得るエネルギーに相当する。
電磁気力	電気力や磁気力による相互作用の総称。素粒子標準理論では、電荷を持つ粒子の間で光子（フォトン）が交換されることで力が作用しあうと考える。
データセット	何らかの機械学習手法を適用・評価するためのデータベース。入力データや教師情報などが含まれており、通常はこれを訓練データ、検証データ、テストデータに分けて用いる。訓練データは学習させたいパラメータを獲得する為に用いられるもので、人手で決めなければならないハイパーアーマータを手動で調整する際には別途検証データを用いる。検証データで有望そうなハイパーアーマータを決定し、それに基づいて訓練データでパラメータを学習し、テストデータに学習結果を適用してその精度を評価する。テストデータはパラメータおよびハイパーアーマータを学習・調整する際には用いてはならない。
データ転置	多次元データの並列処理において並列化軸を切替える際に発生するデータ転送処理。
と	
統一的量子論	この文脈では、標準理論と重力を含むような量子理論のこと。
トカマク装置	磁場閉じ込め核融合炉において最も有望な方式の一つ。トーラス状の磁場を発生するのに、トーラスに沿って並べたコイルとプラズマ中の電流を用いる。プラズマ中の電流を電磁誘導で駆動する場合にはパルス運転となるが、中性粒子ビーム等が誘起する電流を用いる定常運転も提案されている。トーラス断面形状が一様となる特徴（軸対称性）があり、核融合反応で発生する高エネルギー α 粒子の閉じ込めに優れる。

用語	解説
トポロジー励起	複数の縮退した量子多体系の基底状態が局所的な情報だけからは互いに区別がつかず、巻きつき数などのトポロジカルな量によってのみ区別できる場合を、トポロジカルな状態という。また、このトポロジカルな量を変えるような非局所的な励起をトポロジー励起と呼ぶ。
トラジェクトリスナップショット構造	分子動力学シミュレーションでの履歴（トラジェクトリ）中の1構造のこと
トレーサ	一般的に大気や海洋の流れの影響を受けて、モデル内を移動する微量気体成分、溶存成分や個体のことを（パッシブ）トレーサーと呼ぶ。海洋生態系モデルにおいては栄養塩、植物・動物プランクトンや魚類がトレーサにあたる。
トンネル効果	量子がそのエネルギーより高いポテンシャルの山を越える、もしくはトンネルを抜けるようにくぐること。
トーラス状	ドーナツ型の幾何形状。磁場閉じ込め核融合炉ではプラズマを磁力線で覆い、かつ、端や磁場のゼロ点をもたないトーラス状の磁場を用いて高温の燃料プラズマを保持する。
ど	
同位体、安定同位体	元素を特定すると原子番号が一意にきまり、その原子核の陽子の数は原子番号と一致する。一方中性子の数は決まった数とはならず、複数の可能性がある。陽子の数は同じだが中性子の数の異なる原子核をもつ原子を互いに同位体の関係にあるという。特に、他の原子核へと崩壊しない、安定な原子核を持つ同位体を安定同位体という。不安定な同位体では、原子核がアルファ線やベータ線を放出したり核分裂することによって他の原子核へと崩壊する。
同時スイッチングノイズ	LSI の複数の外部出力信号がハイレベルからローレベルまたはその逆方向にほぼ同一のタイミングで変化する時に LSI 内部の出力回路用の電源及びグラウンド配線に発生する電圧ノイズ。
ドラッグデリバリーシステム (Drug Delivery System: DDS)	必要な薬物を必要な時間に必要な部位へと作用させるために、薬物の体内分布を制御し、患部に薬剤を届ける仕組み
に	
二次高調波	入射電磁波と相互作用する物質中において、非線形光学効果により発生する入射電磁波の2倍の周波数を持つ電磁波のこと。
2次摂動論	HF 計算では考慮されない電子相関（平均場からのずれ）を摂動論に従って取り込む Post-HF 法の中で、2電子励起だけを考慮する基本的なアプローチ。強相関系には適用できない。

用語	解説
二次電池	蓄電池(充電式電池)のこと。
ニホニウム	原子番号 113 の元素の名前。日本で初めて命名権を獲得した元素であり、理化学研究所にて合成された。元素記号は Nh。
ニュートリノ	素粒子でレプトン族の一種。電子ニュートリノ、ミューニュートリノ、タウニュートリノの 3 種類が確認されている。電磁気力を受けない中性粒子で、弱い力と重力が作用する。このため検出は難しい。太陽中心付近での熱核融合反応で電子ニュートリノが発生し、地球には 1 平方センチ当たり毎秒 660 億個やってきているが、ほぼ地球をすり抜ける。近年、質量がゼロではないことが分かったが、ほぼゼロであり詳しい質量は不明である。超新星爆発では中心部に中性子の芯が形成される際に大量にニュートリノが発生し外部に放出される。1987 年に大マゼラン星雲で超新星爆発(SN1987A)が起り、超新星爆発由来のニュートリノが初めて地球上で観測された。日本の「カミオカンデ」ニュートリノ観測施設では詳細な観測が行われ、超新星爆発機構の理解に貢献した。
ニューロインフォマティックス	神経データベースや情報理論のような情報学的な手法を神経科学で手用させる学問領域
ニューロン	神経細胞。脳の構成要素。
ぬ	
ヌクレオソーム	真核生物の核における DNA とタンパク質の複合体であるクロマチンの構成単位
ね	
熱電変換	2 種類の異なる金属または半導体を接合して両端に温度差を生じさせると起電力が生じることを利用し、熱を電力に変換すること。
熱搖らぎ	原子・分子程度の微小な粒子の熱運動に由来する運動エネルギー程度のエネルギー
燃料電池	電気化学反応によって燃料の化学エネルギーから電力を取り出す電池。燃料には、水素、炭化水素、アルコールなどが用いられる。
の	
ノルム保存型擬ポテンシャル	結晶内の電子の波動関数を平面波で展開するために真のクーロン型ポテンシャルの代わりに用いられるのが擬ポテンシャル。そのうちカットオフ半径内の電価(ノルム)を変えないのがノルム保存型擬ポテンシャル。
ノンブロッキング通信	並列計算における通信方法の一つ。データの送受信を行う際に、送受信の完了を待たず、他の処理を開始する通信方法。
は	

(次ページに続く)

用語	解説
ハイパー核	ストレンジクォークを含むバリオンをハイペロンと呼ぶ。ハイパー核とはハイペロンを含む原子核の総称。
ハイブリッド汎関数	実験値をより良く再現するために、従来の汎関数に HF 交換相互作用の要素を取り込んだ汎関数
ハイペロン	ストレンジ (s) クォークを含むバリオンはハイペロンと総称され、ラムダ粒子、シグマ粒子、オメガ粒子などがある。
ハイペロン間力	ハドロンの間に働く強い力のうちハイペロンの間に働く強い力をハイペロン間力と呼ぶ。
発火	スパイク様の活動電位が発生する様
ハドロン	強い力で結びついたクォークの複合粒子の総称。ハドロンはクォーク 3 個からなるバリオンとクォーク・反クォーク対からなるメソン（中間子）に分類される。陽子や中性子はバリオンの一種である。陽子と中性子は通常の原子核を構成することから総称して核子とも呼ばれる。
ハドロン共鳴	強い相互作用により様々なハドロンが形成されるが、その多くは短時間で崩壊するため、共鳴状態と呼ばれる。
ハドロン行列要素	相互作用を記述する演算子をハドロン状態で挟んだ行列要素。
ハドロン相	クォークはハドロン中に閉じ込められており、単独では取り出すことができない状態。
ハミルトニアン行列	物理系のエネルギーはハミルトニアンと呼ばれる。物理系の運動はハミルトニアンから記述される運動方程式に従う。量子力学ではハミルトニアンは演算子または行列として表され、その行列の性質（固有値・固有ベクトル）から系の性質を知ることができる。
半導体	電気伝導性の良い金属などの導体（良導体）と電気抵抗率の大きい絶縁体の中間的な抵抗率をもつ物質。
反応経路	ある化学反応において出発物質がどのような過程を経て最終生成物に変化していくかを指す。反応機構とも呼ばれる。
ハートリー項	2 つ以上の電子が存在するときに、電子と電子との間にはクーロン的な反発する力が働くが、それに関係するポテンシャル（位置エネルギー）もしくはエネルギーのことを指す。
ハートリーフォック (HF) 計算	電子間の反発を平均場近似の下で記述し、系の分子軌道を変分的に求める手法。
ハートリーポテンシャル	電子密度の空間分布で決まる静電ポテンシャル
ば	

(次ページに続く)

用語	解説
バイアスпотенシャル	自然状態では滅多に起こらないが重要な化学反応を人工的に高い頻度でシミュレーション上発生させるために加える原子間ポテンシャルのこと。Metadynamics 法はバイアスпотенシャルを生成する。
バイオインフォマティクス	生物学的な問題をハイスクープデータなどとアルゴリズムを組み合わせて計算機を用いて解決する研究手法
バイオセンサー	生体の認識システムを利用した化学的センサーのこと
バイオフィルム	細菌などの微生物の増殖の足場。カテーテル等の医療機器に形成されたバイオフィルム内の細菌は抵抗性が高くなり問題となる。
バイオミネラリゼーション	生物が結晶や無機鉱物を産生すること。骨や歯、貝殻などが身近な例。
バイオミメティック	生物が持つ優れた機能を人工の物質で実現しようとする化学
バイセクションネットワークバンド幅	通信網の性能の指標の一つ。通信網の中の計算ノードを 2 等分し、その分割された部分同士の間で単位時間あたりに通信できるデータ量の事。
バイナップ	立体選択的合成反応において広く利用されている配位子。バイナップ-ルテニウム触媒を用いた不斉水素化反応を開発した野依良治は 2001 年のノーベル化学賞を受賞した。
バタフライ演算	高速フーリエ変換などにあらわれる演算および通信パターン。
バリオン	ハドロンの中で、クォーク 3 個からなるものがバリオンと呼ばれる。陽子・中性子もバリオンの一種で 2 種類のクォーク (u,d) を 3 つ組み合わせて構成されている。ストレンジ (s) クォークを含むバリオンはハイペロンと総称され、ラムダ粒子、シグマ粒子、オメガ粒子などがある。
バルクナノメタル	一般の金属よりも小さな結晶粒からなる金属材料。強度、延性、靄性等の機械的性質の向上が見込まれる。
バルクひずみ	複数の材料が混ざった状態での計算要素内のトータルのひずみ量
バレンス殻	原子核の中の核子のうち、その構造の決定に特に重要な影響を与えるものをバレンス核子(粒子)と呼び、バレンス核子が運動する空間をバレンス空間、あるいはバレンス殻と呼ぶ。有限のバレンス核子、バレンス殻を設定することによって多体シュレーディンガー方程式を正確に解くことが現実的に可能となる。
バレンス粒子、バレンス空間	原子核の中の核子のうち、その構造の決定に特に重要な影響を与えるものをバレンス核子(粒子)と呼ぶ。近似的に、これらのバレンス粒子だけを取り扱った計算ができ、量子力学の計算ではバレンス粒子の運動を記述するヒルベルト空間を扱うため、これをバレンス空間と呼ぶ。
バンドギャップ	電子が占有出来ない禁止帯のこと。最高占有軌道準位と最低非占有軌道準位とのエネルギー差に対応する。

用語	解説
ぱ	
パラメータ化, パラメタリゼーション	連続流体の性質を持つ大気を数値モデル化する際、有限個の格子や波数で表現する必要がある。このとき、モデル格子で表現できない格子サイズよりも小さなスケールの現象が格子スケールの現象に及ぼす影響が無視できることがある。そのとき、格子の物理量を用いて、格子サイズよりも小さな現象が格子スケールに及ぼす影響を表現することをパラメタ化といい、その手法のことをパラメタリゼーションという。
パラメータスキャン	入力パラメータを変更して多数のシミュレーションを実行し、設計パラメータ等に対する性能や機能の依存性を検証すること。
parity	空間の反転に対して系が対称性をもつときの量子数。
ひ	
光関係タンパク質	光を刺激として受けるタンパク質。
光格子	対向するレーザー光を用いてその中を運動する粒子に対して周期的なポテンシャルを作り出す。その結果、粒子は結晶格子点に閉じ込められた粒子のように振舞い、そのような系のことを光格子と呼ぶ。
光触媒	光を当てることで触媒作用を発現する物質の総称。
光生理学	化学的あるいは遺伝子光学的な光センサ-分子や光による刺激素子の発展を背景に蛍光顕微鏡のような光学的な方法で生体の活動をしる学問体系
光電荷移動	光を吸収することで誘起される、ある分子が持っている電荷を別の分子と授受する化学反応。
非局所擬ポテンシャル	内殻電子などの及ぼす影響をポテンシャルに置き換えたもののうち、位置以外の要素(角運動量など)に依存するもの。
非侵襲治療法	生体に力学的影響を及ぼさない治療法
歪速度テンソル	速度場の空間的な変化を表す速度勾配テンソルから、回転を表す反対称成分を除いた対称成分で、変形の速度を表す。
非静水圧	大気や海洋の支配方程式を考える際、水平方向に十分大きな現象(気象では数十キロ以上)に着目する場合は、重力と鉛直方向の気圧傾度力が釣り合っていると近似(静水圧近似または静水圧近似と呼ぶ)することができる。全球を対象とした多くの大気・海洋モデルでは、静水圧近似した方程式が用いられている。一方、より細かな現象に着目する場合などは、静水圧近似が成り立たず、鉛直方向の運動方程式を陽に考慮する必要がある。このような方程式を、非静水圧の方程式と呼ぶ。
非摂動ダイナミクス	摂動的手法では解析が難しく、その本質を理解するためには非摂動的手法を必要とする力学現象。

(次ページに続く)

用語	解説
非線形光学応答核磁気共鳴	核スピン間の非線形相互作用に伴い生じる高調波を利用する核磁気共鳴法
非線形振動子系	複数の非線形振動子が結合されたシステム。非線形振動子とは、運動が初期値に比例しない振動子（ばねのように振動する要素）のこと。カオスや同期など、様々な興味深い現象を示すことが知られている。
ヒッグス場	素粒子が質量を持つ仕組みを説明する理論であるヒッグス機構において導入されるスカラー場。
ヒッグス粒子、ヒッグスボソン	標準理論において電弱相互作用から弱い相互作用と電磁相互作用を分化させ、クォークやレプトンなどに質量を与える重要な役割を担っている。
非熱的分布	するようになる（宇宙開闢後約 10^{-6} 秒）。その後、温度が下がると、いくつかの陽子と中性子は結びつき、一部がヘリウム原子核などを形成する元素合成が始まる（宇宙開闢後約 3 分から約 20 分の間）。さらに温度が冷えると元素合成は終了し、軽原子核は安定な原子核に崩壊し元素比率が固定される。この過程では無視できる量のリチウム 7 までの元素と水素 1 とヘリウム 4 が生成される。宇宙開闢後約 3 分から 20 分の間で生成された原子核を理論的に計算することができ、現在の宇宙の元素質量の割合が（水素 1 が約 75 %、ヘリウム 4 が約 25 %）であることを説明する。一方でビッグバン原子核合成ではわれわれにじみ深い炭素、鉄、金、銀などのリチウムより重い元素は全く生成できない。
被膜	電極を覆い包んでいる膜。電池の性能を左右するため、安定的な制御が求められる。
標準脳座標系	個体差を補償するように作られた脳内の標準座標系
標的タンパク質	薬が作用する標的となるタンパク質。
ヒルベルト空間、模型空間	量子力学では、系の状態は抽象的なヒルベルト空間の中のベクトルに対応している。この空間は無限次元であるが、実際の数値計算ではこれを有限の大きさの次元、しかもなるべく小さい次元の空間にする必要がある。このようにして計算に適した形に抜き出された空間を模型空間と呼ぶ。
び	
ビッグバン	現在広く受け入れられている学説によれば、宇宙は約 137 億年前に大きな爆発（ビッグバン）のように膨張して現在に至ったとされる。

(次ページに続く)

用語	解説
ビッグバン原子核合成	ビッグバン宇宙誕生直後に起こった原子核の合成を指す。宇宙誕生直後、宇宙全体は超高温高密度であった。宇宙誕生後ごく初期には物質はクォークの状態であったが、宇宙が膨張し冷えるとともにクォーク同士が結合し陽子や中性子を構成するようになる（宇宙開闢後約 10^{-6} 秒）。その後、温度が下がると、いくつかの陽子と中性子は結びつき、一部がヘリウム原子核などを形成する元素合成が始まる（宇宙開闢後約 3 分から約 20 分の間）。さらに温度が冷えると元素合成は終了し、軽原子核は安定な原子核に崩壊し元素比率が固定される。この過程では無視できる量のリチウム 7 までの元素と水素 1 とヘリウム 4 が生成される。宇宙開闢後約 3 分から 20 分の間で生成された原子核を理論的に計算することができ、現在の宇宙の元素質量の割合が（水素 1 が約 75 %、ヘリウム 4 が約 25 %）であることを説明する。一方でビッグバン原子核合成ではわれわれになじみ深い炭素、鉄、金、銀などのリチウムより重い元素は全く生成できない。
微物理過程	大気中で雲を構成する水滴・氷晶（雲粒）が、発生してから、雨・雪などの降水現象として地表面に落下する、もしくは蒸発により消滅するまでの一連の成長・消滅過程をさす。雲粒同士が大気中で衝突して併合する過程、雲粒が凍結・融解する過程などがある。
ふ	
ファインマン振幅	量子力学に基づいて素粒子反応の確率を計算する場合、絶対値の 2 乗が反応確率となる不变散乱振幅というものを計算する。通常普遍散乱振幅を解析的に厳密に計算することは困難であるため、摂動理論を用いて近似的に計算していく。摂動論では次数ごとにファインマン図形に基づく計算を行なう。この様な摂動計算による不变散乱振幅をファインマン振幅という。これにより素粒子反応の散乱断面積（反応確率）を求めることができる。
ファインマン図	相互作用を表す項が小さい場合に、その項をゼロにした計算から逐次近似で相互作用の効果を取り込んでいくことができる。このような計算手法を摂動計算というが、量子場の理論において摂動計算に必要な計算式を構成するためにグラフを用いて計算式を構成することができる。この際に用いるグラフをファインマン図という。量子場の理論においてグラフを用いた摂動計算の手法を R.P.Feynman が考案したためこのように呼ばれる。

(次ページに続く)

用語	解説
フィラメントワインディング	主に炭素繊維強化複合材料製の高圧容器を作製する際に用いられる製法。炭素繊維を数万本束ねた炭素繊維束を、ライナーと呼ばれる内容器に巻き付けて成型する方法。炭素繊維強化複合材料製高圧容器は、燃料電池自動車用高圧水素容器として使用され、高信頼性と軽量化の両立が求められている。
フェムトスケール	ハドロンや原子核の大きさ程度のミクロな世界。fm(フェムトメートル)は 10^{-15} m。
フェムト秒	1,000兆分の1秒が1フェムト秒。1フェムト秒は、光の速さ(秒速約30万キロメートル)でも0.3ミクロンしか進むことができないほどの極短時間。
フェルミオン	フェルミ粒子。スピン角運動量が半整数倍である。フェルミオンには、夸arks、電子、ニュートリノ、陽子、中性子などがある。
フェルミオン作用	フェルミオンの運動方程式を与える作用積分の事。作用積分はラグランジアン密度の時間空間積分である。格子場の理論では時空はとびとびの格子で近似される。この際フェルミオンの作用は格子上で近似されたラグランジアン密度の格子上での総和になる。格子化の方法には任意性があるため格子上のフェルミオン作用にはいくつかの種類がある。本文中ではウィルソン型とドメインウォール型が言及されている。
フェルミオン伝搬関数	フェルミオンが時空間を量子的に伝搬する確率振幅を表している関数。絶対値がフェルミオンの伝搬確率となる。
フォノン分散関係	結晶の格子振動を量子化したのがフォノン。そのエネルギーと波数との関係が分散関係。
フォルトトレランス機構	1ノードが故障したとしても補完により計算が止まらない仕組み
フォンビルブランド因子	血中にある凝固因子のひとつ
不純物偏析	結晶中の不純物が表面や欠陥など何らかの構造の周辺に集まること。
不定性	この文脈では、実験データの不足により、相互作用のポテンシャルを実験データから良く決めることができないこと。
フラグメント	フラグメント分子軌道法計算を行うために分子全体を部分系に分割した際の構成単位のこと。
フラグメント探索	化合物設計プロセスにおいて化合物の部品(フラグメント)を探し出すこと
フレーバー	夸arksとレプトンの種類を表す。たとえば、夸arksにはアップ、ダウントなどの種類があり、レプトンには電子、ミュー粒子などがある。

用語	解説
フーリエモード展開	時間微分を含む偏微分方程式を、正弦波の重ね合わせであるフーリエ級数に変換すること。複雑な波動を単純な波の重ね合わせとして表現することができる。
ぶ	
ブシネスク近似	流体を非圧縮性とし、圧力の変化に伴う密度変化は無視するが、温度変化に伴う密度変化は考慮する近似手法。
部分空間対角化	占有電子軌道など注目している一体電子軌道を基底とした空間でハミルトニアンを表現しそれを対角化すること
ブロッキング	配列のデータ処理をする際にデータ転送速度が高速なキャッシュやメモリに保持可能なデータサイズを考慮して配列を区分することで、処理性能の向上を図る性能チューニング手法。
分割統治法	分割統治法は大規模な問題を効率的に解くアルゴリズムの一つで、そのままでは解決することが難しい大きな問題をいくつかの小さな問題に分割して個別に解決していくことで最終的に大きな問題を解決する方式。量子化学計算のための分割統治法は Waitao Yang 教授（現デューク大学）より考案された。
分子混雑環境	細胞内のように、タンパク質をはじめとする様々な分子が高密度で存在する込み合った環境。分子は溶媒中における孤立した環境下とは異なった性質を示す。
分子シミュレーション	分子動力学などに則り、分子の動きを計算すること。物理・化学的性質を求めるために行われる。
分子動力学シミュレーション	原子間力に基づき、運動方程式を数値的に解き、分子の運動をシミュレーションする計算方法
分子動力学法	原子間力に基づき、運動方程式を数値的に解き、分子の運動をシミュレーションする計算方法。
分子モデリング	分子の立体構造を、計算機中で構築すること
分子モーター	生体内で ATP などのエネルギーを機械的な動きに変換する分子
分数量子ホール効果	半導体のヘテロ接合面等において実現される 2 次元電子系に強い磁場をかけると、低温でホール抵抗の値が量子化される現象が起こる。この値は e を電子の素電荷、 h をプランク定数とすると $(p/q) \cdot (e^2/h)$ と表される。ここで、 p と q は整数であり、 q が 3 以上の奇数で p/q が整数とならない場合を分数量子ホール効果と呼ぶ。これは物質中において分数電荷を持つ新たな素励起が生じるために起こる現象であり、発見者の Tsui, Stormer, Laughlin は 1998 年にノーベル物理学賞を受賞した。
ふ	

用語	解説
プラットホームシミュレータ	その中で独自のスクリプト言語を持つことにより多数の異なった現象が扱われるようになったシミュレータ
ヘ	
閉殻・開殻配位	希ガスの原子が特別に安定化するように、陽子・中性子数がある決まった数(魔法数)になると原子核も安定化する。このような原子核の核子は閉殻配位に対応すると称される。逆に陽子・中性子の数が魔法数からずれたものを開殻配位と呼ぶ。
平均場、平衡平均場	原子における電子と同様、原子核における核子は各々がポテンシャル場に支配されて独立に運動していると近似することができる。そのポテンシャルの源は、核子の相互作用自身が生み出す平均的な場であり、それを平均場と呼ぶ。特に時間に依存しない平均場を平衡平均場と呼ぶ。
平面波基底	分子軌道を表現するための関数群。平面波を表す関数の線形結合で分子軌道を表現。
ヘテロな構成のCPU	機能の異なるコアを組み合わせたCPU
ヘム	鉄イオンを含む化合物。しばしば、タンパク質に含まれ機能の発現に重要な寄与をする。
変形核	形状が球形からずれて変形した原子核。
ベ	
ベイジアンネットワーク	統計的因果モデルの一つで因子間の因果関係を点と有向枝からなるネットワークで表現したもの。バイオインフォマティクスでは遺伝子発現制御ネットワークの推定・モデル化で用いられる
ベイジアンフィルタ	ベイズ統計に基づくデータの学習・分類法
ペ	
ペロブスカイト太陽電池	ペロブスカイト結晶を用いた太陽電池。従来のシリコン太陽電池と比較して、光エネルギーから電気エネルギーへの変換効率が高いことが知られており、将来が有望であるとされている。
ほ	
ホールセルプランプ	細胞にパッチ電極を接続して全体を電位固定できる状態にして測定する方法
ぼ	
ボクセルデータ	形に沿った線や面で形状を表現するのではなく、空間を直方体で分割しその直方体内部の分布情報で形状を表現する方法。二次元の映像を示すピクセル(Pixel)に対して三次元(Volume)を表すボクセル(Voxel)
ボソン	ボース粒子。スピン角運動量が整数倍である。ボソンには、素粒子間の相互作用を媒介する粒子である、光子やグルーオンなどがある。

用語	解説
ボゾン系	構成粒子がボーズ粒子である量子系
ボトムクォーク, b クォーク、B クォーク	標準模型には 6 フレーバーのクォークが含まれている。そのうち第 3 世代のダウントイプのクォークがボトムクォーク（または b クォーク、ビューティクォーク）となぞけられている。6 フレーバーのクォークのうち b クォークは 2 番目に質量が重く、約 4.5 GeV の質量を持つ。b クォークは素粒子であるにも関わらず、複合粒子である陽子の質量の約 4.5 倍の質量を持っておりかなり重い粒子である。
ボリュームレンダリング	ボリュームデータに対する画像生成手法の一つ。データの内部構造や全体の様子を透過的なイメージで表現することができる。
ボーズ系モット転移	ボーズ粒子が互いの間に働く斥力相互作用によって絶縁体化すること。
ボーズ粒子	量子統計力学的な性質から分類した、粒子の種類のひとつ。多くの粒子が同じ量子状態を占めることができるという性質を持つ。典型的なボーズ粒子として光子やヘリウム 4 などがある
ボーズ・AINシューイング凝縮	多数のボース粒子が一つの量子状態を占めることで現れる物質の状態。
ぼ	
ポイントベースレンダリング	点群を基本とした画像生成手法。並列処理に向いた方法で、画像の品質を点の数により調整できる。
ポストスケーリング時代	大規模集積回路は、スケーリング（比例縮小）にもとづきトランジスタの微細化により高性能化と高集積化を同時に実現してきたが、今後は発熱や消費電力により困難となると予想されている。スケーリングの限界以降（ポストスケーリング）では、全く新しい指導原理が必要とされる。
ポロイダル・トロイダル展開	任意のソレノイダル場は、トロイダルポテンシャルとポロイダルポテンシャルの二つのスカラー場で表現された二つの項の和として一意に分解できる。
ま	
マイクロカプセル	極小のカプセル内に薬剤等を内包した物
マイコプラズマ	真正細菌の一種でゲノムサイズが小さく、細胞サイズも小さい
マイナー・アクチナイト核	ウラン・トリウムに代表される重元素をアクチナイトと呼ぶが、自然界に存在する安定な（寿命が非常に長い）ものの他に、原子炉などでは寿命の短いアイソトープが作られており、これらをマイナー・アクチナイトと呼ぶ。
膜電位	ニューロンの内部状態を表すパラメータ。ニューロンは電気的素子であり、膜の性質により細胞内外で電位差を生じる。その差のこと。
膜輸送体	生体膜を貫通し、膜を通して物質の輸送をするタンパク質の総称

(次ページに続く)

用語	解説
魔法数	原子核は陽子と中性子から構成されているが、ある特定の数の陽子または中性子を含むとき原子核は特にエネルギー的に安定となる。この数のことを魔法数と呼ぶ。古くからよく知られている魔法数として 2 (ヘリウム: ${}^4\text{He}$)、8 (酸素: ${}^{16}\text{O}$)、20 (カルシウム: ${}^{40}\text{Ca}$) などがある。
マルコフ連鎖	離散的な時系列を生成するための確率過程の一種で、ある時刻での状態は直前の時刻での状態のみに依存して決まり、それ以前の履歴と無関係である性質 (マルコフ性) を持つ。
マルチグリッド型前処理	連立一次方程式の解法を使用する際、行列の収束性を向上するために導入される前処理法の一種。疎・密の計算格子に対して順に解を求め、反復計算において早く収束解が得られるようにした手法。構造型の格子を直接用いる方法、代数的に疎格子を表現する方法など、種々存在する。
マルチコンパートメント	神経線維を多数のシリンドー様のコンパートメントの連なりと考えるモデル
マルチスケールシミュレーション	異なる時空間スケールのシミュレーションを組み合わせて取り扱うシミュレーション。
マルチスケール物質	原子スケールからはじまり、何桁も異なる長さスケールの効果が顕著な物質。あるいはそのような効果を取り入れなければ理論的に説明のできないような物質
マルチスケール・マルチレゾリューション法	幅広い時空間にまたがる対象に対し、それぞれの階層・解像度での計算を連成させるシミュレーション法
マルチフィジックスシミュレーション	異なる物理現象を組み合わせて取り扱うシミュレーション
マルチフェロイクス	強磁性と強誘電性など二つ以上の秩序状態が物質中に共存し、互いに関係を持つ状態。これにより磁場 (電場) をかけることで誘電性 (磁性) を制御することなどが可能となる。
み	
ミセル	疎水基と親水基を併せ持つ界面活性剤分子が、溶媒中において球状や棒状に自発的に会合した分子集合体
密度行列くりこみ群	強相関系の数値的計算手法のひとつ。特に 1 次元、または 2 次元的な電子構造を持つ低次元強相関系の研究に用いられる。
密度汎関数 (DFT) 法	系の電子エネルギーが電子密度の汎関数で与えられるコーンシャム方程式に基づき固体系や凝集系の電子状態を計算する手法。汎関数のバリエーションは多数あるが、物理分野では BLYP がよく用いられる。

(次ページに続く)

用語	解説
ミューオン異常磁気能率	ミューオンはレプトン族のうち 2 番目に重い粒子。質量以外の性質は電子と同じ。質量は電子の約 200 倍である。ミューオンはスピン 1/2 で自転しているため小さな磁石となっている。磁石の強さを磁気能率（磁気モーメント）という。磁気能率は量子力学に基づく計算と量子力学を使わない計算で違いが生じるため、その差を異常磁気能率と呼ぶ。ミューオンの磁気能率は高精度（相対誤差約 0.5×10^{-6} ）で計測されている。素粒子標準理論を用いた理論計算が可能である。2012 年現在、理論計算と実験値は相対的に約 21×10^{-6} ずれている。ずれの原因は、理論計算に含まれる精度不足である可能性と新しい物理の兆候である可能性がある。
ミュー粒子（ミューオン）	ミュー粒子（またはミューオン）はレプトンの一種でミュー粒子は標準模型では第二世代に属する。質量以外は電子と同じ物理的性質を持つ。質量は電子の約 200 倍である約 106MeV をもつ。
ミュー粒子異常磁気能率	ミュー粒子はレプトン族のうち 2 番目に重い粒子。質量以外の性質は電子と同じ。質量は電子の約 200 倍である。ミュー粒子はスpin 1/2 で自転しているため小さな磁石となっている。磁石の強さを磁気能率（磁気モーメント）という。磁気能率は量子力学に基づく計算と量子力学を使わない計算で違いが生じるため、その差を異常磁気能率と呼ぶ。ミュー粒子の磁気能率は高精度（相対誤差約 0.5×10^{-6} ）で計測されている。素粒子標準理論を用いた理論計算が可能である。2012 年現在、理論計算と実験値は相対的に約 21×10^{-6} ずれている。ずれの原因は、理論計算に含まれる精度不足である可能性と新しい物理の兆候である可能性がある。
ミュー粒子実験	ミュー粒子の性質を詳細に調べることで標準理論からのズレを見つけることができると期待されている。J-PARC においてもミュー粒子を用いた実験が行われている。ミュー粒子異常磁気能率を精密に測定する実験である。
め	
メソ降水系	水平スケールが 100km 程度（メソスケール）の積乱雲の集合体である。単純な集合体ではなく、上昇・下降流域といった構造を持つ「系」であるため、单一の積乱雲に比べて寿命が長い（6 時間以上）。大気の状態や地域特性によって形態を変え、停滞すると同じ場所に多量の降水をもたらす。
メソスケール	ナノスケール（1 ナノメートル = 10^{-9} メートル）と現実世界との中間スケールをさす。ナノスケールよりも数桁大きいサイズの物質を扱う。

(次ページに続く)

用語	解説
メタゲノム	特定の環境中の微生物群など單一種毎のゲノム解析が難しい場合に、その生物群内全体のゲノムの集合をひとつのゲノムとしてとらえる考え方
メタマテリアル	自然界では見られない性質を示す人工的に作られた物質一般を指す言葉であるが、特に負の屈折率を持つ物質を指すことが多い。光の波長よりも小さな物質で特殊な高次構造を作ることによって実現できる。その極めて特殊な光学的性質を利用した応用科学的研究も盛んに行われている。
メタン	最も単純な構造の炭化水素で、1個の炭素原子に4個の水素原子が結合した分子である。分子式は CH ₄ 。和名は沼氣（しょうき）。
メタンハイドレード	低温かつ高圧の条件下でメタン分子が水分子に囲まれた、網状の結晶構造をもつ包接水和物の固体。およその比重は 0.9 g/cm ³ であり、堆積物に固着して海底に大量に埋蔵されている。
メッシュ/トーラス	計算ノード間の通信ネットワークの形態の一つ。多次元の格子状のもの。格子の端を周期的に結合した物はトーラスと言う。
メモリバランス型	エクサスケールシステム構成例 の一つ。演算性能 100PFLOPS/ メモリ帯域 100PB/s メモリ量 100PB がめどの構成
も	
模型空間	量子力学では、系の状態は抽象的なヒルベルト空間の中のベクトルに対応している。この空間は無限次元であるが、実際の数値計算ではこれを有限の大きさの次元、しかもなるべく小さい次元の空間にする必要がある。このようにして計算に適した形に抜き出された空間を模型空間と呼ぶ。
モデル脊椎動物	線虫（神経数 300）・ショウジョウバエ・カイコ等の昆虫（神経数 10 万）は遺伝子が同定され、ある程度生理実験も可能な無脊椎系のモデル生物であるが、同様な意味で、脊椎動物においては・ゼブラフィッシュ（神経数 100 万）・マウス（神経数 1 億）などが世代が短く遺伝子が同定されており、かつ生理実験も可能な比較的単純なシステムを持つモデル脊椎生物といえる。
モルフォロジー	分子集合体の空間分布構造。有機薄膜型の太陽電池においては、高分子膜質を指す。
モンテカルロ法	数値計算において乱数を利用した計算手法を一般にモンテカルロ法と呼ぶ。計算アルゴリズムのその一部に乱数を利用して効率的に計算ができるようになっている場合「モンテカルロ法」という語が用いられる。本文でもさまざまな計算アルゴリズムに対して乱数を利用し高性能化したアルゴリズムが用いられている。
ゆ	

用語	解説
有機薄膜太陽電池	導電性高分子やフラー・レン等の電子授与性、授受性を持つ有機分子の組み合わせで構成される有機薄膜半導体を用いた太陽電池。
有限温度	非ゼロの温度を持つ物理系。
有限格子間隔効果	格子間隔が有限であることから生じる系統誤差
有限フェルミ多体系	核子（陽子と中性子）はフェルミ粒子であり、地球上に存在する原子核は最大でも数百個の核子から成っている。無限に近い粒子数の多体系と区別するため、有限多体系とよび、数値的にも多くの特有の困難がある。
有限密度	非ゼロの密度を持つ物理系。
有効媒質法	溶液分子の周りの溶媒の分布確率を表す分布関数を求める理論。様々な種類の分布関数理論があるが、特に 3D-RISM 法はタンパク質やナノチューブといった大きな分子の溶媒和を扱うことができる。
有効模型	一般には、与えられた基礎方程式に対し、着目しているエネルギー領域において有効となるよう自由度を落とした模型のこと。原子核物理では、特異性が強い核力を直接扱わずに、低運動量成分のみの扱いやすい相互作用を用いて MeV 領域の低エネルギー現象を記述する模型一般を指す。上記の「カイラル有効場理論」も同じ概念に基づく。
有効模型的アプローチ	特異性が強い核力を直接扱わずに、数値的に扱いやすい核力（有効相互作用）に変換する方法。上記の「カイラル有効場理論」と似た概念で、特定のエネルギー領域、制限されたヒルベルト空間における原子核多体系問題で用いられる。
輸送係数	一般に、外場の強さに対する系の応答の強さを与える比例係数のこと。クォーク・グルーオン・プラズマ (QGP) の物理においては、粘性係数などを指し、QGP の物性を特徴づける量として重要なものである。また、ランジュバン方程式に基づく重イオン反応の計算では、摩擦や慣性を表すインプットとして与えられる量である。
ユークリッド時空	ユークリッド幾何学が成り立つ 4 次元時空。時間方向と空間方向の区別はない。
よ	
陽子・中性子過剰核	陽子数に比べて過剰に多い中性子を含む原子核。不安定であり、安定な原子核になるまで中性子から陽子への β 崩壊を繰り返す。
溶媒和エネルギー	孤立状態の溶質分子が溶媒中へと移行することに伴って変化する自由エネルギー量
弱い力	不安定原子核が β 崩壊する際に働く力。素粒子標準理論では、すべてのフェルミオンの間で W ボソンや Z ボソンという粒子を交換されることで力が作用しあうと考える。

(次ページに続く)

用語	解説
ら	
ラジカル	電子が対になつてないことで不安定になっている化学物質
ランチョス法	エルミート行列を三重対角化する手法。数値計算において再帰計算による効率的な演算が可能であることから、固有方程式の解法等でよく用いられる。
乱流境界層	乱流で構成された境界層（粘性を有する流体中において粘性の影響を強く受ける領域で、一般には物体表面に見られる）乱流境界層では流体の渦運動により運動量やエネルギーの交換が強く行われる。このため、壁面近傍の流体へ運動量が供給されるので層流境界層よりも剥離しにくく、壁面付近で急激に減少する速度分布を持つため摩擦抗力が大きい。
乱流スケール	乱流における渦の大きさ
り	
リオーダリング	計算順序の並べ替えのこと。計算処理を速くしたり、並列処理ができるよう依存関係をなくすために行われる。
リガンド結合	受容体に特異的に結合する物質（リガンド）が結合すること
リチウムイオン二次電池	正極と負極の間をリチウムイオンが移動することで充電や放電を行う二次電池。（二次電池：充電式電池のこと）
リボゾーム	細胞内の構造体で、遺伝情報からタンパク質へと変換する機構である翻訳が行われる場である。
量子色力学 (QCD: Quantum Chromodynamics)	クォークとグルーオンの力学（強い力の力学）を量子力学的に記述する理論。素粒子標準理論の一部をなす。
量子化学計算	原子と電子の振る舞いから、物性や反応性を計算する。計算手法として、第一原理計算法が使用される。
量子数射影	量子多体計算において、変分計算によって得られた波動関数はハミルトニアンがもともと持っている対称性を自発的に破っていることが多い。この波動関数に射影演算子を作用させることによって、本来保つべき対称性を回復させる方法。
量子多体系（量子多体现象）	凝集系のうち特に量子力学的効果が顕著な物質を指す。特に極低温において、古典力学では説明できない様々な現象を示す
量子電気力学、量子電磁力学 (QED: Quantum Electrodynamics)	荷電粒子と光子の力学（電磁気学）を量子力学的に記述する理論。素粒子標準理論の一部をなす。
リラクサー	特殊な強誘電体。誘電率の周波数依存性に特徴がある。

(次ページに続く)

用語	解説
臨界終点	相図において一次相転移が終結し、熱力学変数が連続的に変化するようになる（クロスオーバー）へと移行する点
隣接通信	並列計算で領域分割法を用いる時に、隣り合う分割された領域間でデータの授受を行う通信の事。
る	
ルシフェラーゼ	螢の蛍光タンパク質
ルミノシティ	ビーム衝突型加速器実験において、ルミノシティ=単位時間あたりに起こる反応の回数÷断面積で定義される。
ループ	ファインマン図形において現れるループ構造。量子補正が高次になるにつれループの数が増える。
れ	
レアイベント探索アルゴリズム	通常のシミュレーションでは滅多に発生しないが、科学的に重要な事象を探索するためのアルゴリズム。例えば分子動力学シミュレーションでは高い活性化障壁をもった化学反応はなかなか発生しない。
零点振動	量子力学的に絶対零度でも不可避の量子の振動
レイトレーシング	光線追跡法。コンピュータグラフィクスの画像生成手法の一つで、光が反射屈折する物理現象を模倣し、画像を作成する。
レオロジー	物質の流動と変形を取り扱う学問。
レジスタブロッキング	CPU 内のレジスタ上になるべくデータを集められるようにするためのコーディングテクニック。これにより命令実行効率が向上する。
レプトニック崩壊の崩壊定数	ハドロンの崩壊のうちレプトンへ崩壊する場合をレプトニック崩壊という。ハドロンの崩壊の頻度やハドロンの寿命のうち強い相互作用に基づく寄与を因子化したものを崩壊定数と呼ぶ。
レプリカ法	同じ原子から構成されるシステムを複数用意し、それぞれシミュレーションの条件を変えながら、シミュレーションを行う方法。条件パラメータをある一定の法則に従って交換しながら実行するレプリカ交換法などがある。
ろ	
ローカルインジェクション	外部からガラス細管などで分子を脳の中の特定の領域に注入することさらに電圧を同時に付加することで特定領域の細胞内に分子を注入するローカルエレクトロポレーションなども存在する。

(次ページに続く)

用語	解説
ローカルファイルシステム	並列計算機の各計算ノードから独立して参照されるファイルシステム。各計算ノードで個別に使われるファイルを一時的に保存する場所として使われ、他の計算ノードからは参照できないためグローバルファイルシステムと比較して利便性に欠ける。一方でグローバルファイルシステムと比較して特に大規模なシステムにおいて高性能を達成しやすい構成である。

付録 B

執筆者一覧

全体エディタ		
藤堂 真治	東京大学 大学院理学系研究科	教授
足立 幸穂	理化学研究所 計算科学研究センター	研究員
素粒子・原子核		
執筆者		
宇都野 穣	日本原子力研究開発機構 先端基礎研究センター	研究主席
金森 逸作	理化学研究所 計算科学研究センター	研究員
山崎 剛	筑波大学 数理物質系	准教授
執筆協力者		
青木 保道	理化学研究所 計算科学研究センター	チームリーダー
有友 嘉浩	近畿大学 理工学部	教授
石川 健一	広島大学	准教授
石川 正	高エネルギー加速器研究機構 共通基盤研究施設計算科学センター	ダイヤモンドフェロー
江尻 信司	新潟大学 理学部	准教授
緒方 一介	九州大学 大学院理学研究院	教授
金児 隆志	高エネルギー加速器研究機構 素粒子原子核研究所	准教授
北沢 正清	京都大学 基礎物理学研究所	講師
木村 真明	理化学研究所 仁科加速器科学研究センター	室長
藏増 嘉伸	筑波大学 計算科学研究センター	教授
佐々木 勝一	東北大学 大学院理学研究科	准教授
土屋 麻人	静岡大学 理学部	教授
土井 琢身	理化学研究所 数理創造プログラム	専任研究員
富樫 甫	大邱大学校	研究員

(次ページに続く)

富谷 昭夫	大阪国際工科専門職大学 情報工学科	助教
仁尾 真紀子	理化学研究所 仁科加速器研究科学センター	上級研究員
西村 淳	高エネルギー加速器研究機構 素粒子原子核研究所	教授
野中 千穂	広島大学 大学院先進理工系科学研究科	教授
花田 政範	ロンドン大学クインメアリー校 数学科	上級講師
肥山 詠美子	東北大学 大学院理学研究科	教授
保坂 淳	大阪大学 核物理研究センター	教授
本多 正純	理化学研究所 数理構造プログラム	上級研究員
吉田 賢市	大阪大学 核物理研究センター	准教授
渡辺 展正	京都大学 基礎物理学研究所	特定研究員
ナノサイエンス・デバイス		
執筆者		
井戸 康太	東京大学 物性研究所	助教
小野 倫也	神戸大学 大学院工学研究科	教授
藤堂 真治	東京大学 大学院理学系研究科	教授
福島 鉄也	産業技術総合研究所 材料・化学領域	研究チーム長
星 健夫	核融合科学研究所 プラズマ量子プロセスユニット	教授
松下 雄一郎	東京工業大学 科学技術創成研究院	特任准教授
山田 俊介	量子科学技術研究開発機構 量子応用光学研究部	主任研究員
山地 洋平	物質・材料研究機構 ナノアーキテクトニクス材料研究センター	グループリーダー
エネルギー・材料		
執筆者		
神谷 宗明	岐阜大学	
川嶋 英佑	理化学研究所	
中嶋 隆人	理化学研究所	
三上 和徳	理化学研究所	
William Dawson	理化学研究所	
生命科学		
執筆者		
荒木 望嗣	京都大学大学院医学研究科人間健康科学系専攻	特定准教授
伊東 聰	東京医科歯科大学 M&D データ科学センター	助教
海津 一成	理化学研究所	上級研究員
奥野 恭史	京都大学大学院 医学研究科	教授
笠原 健人	大阪大学大学院 基礎工学研究科物質創成専攻	助教

(次ページに続く)

河野 秀俊	量子科学技術研究開発機構	プロジェクトディレクター
小林 千草	理化学研究所	技師
杉田 有治	理化学研究所	主任研究員
泰地 真弘人	理化学研究所	チームリーダー
玉田 嘉紀	弘前大学 健康未来イノベーションセンター	センター長・教授
寺山 慧	横浜市立大学	准教授
松永 康佑	埼玉大学大学院 理工学研究科	准教授
宮下 治	理化学研究所 計算科学研究センター 計算構造生物学研究チーム	上級研究員
宮野 悟	東京医科歯科大学 M&D データ科学センター	特任教授・センター長
森 貴治	理化学研究所 開拓研究本部 杉田理論分子科学研究室	専任研究員
浴本 亨	横浜市立大学	助教
脳・神経科学		
執筆者		
五十嵐 潤	理化学研究所	上級研究員
加沢 知毅	東京大学	
木村 瞳	龍谷大学	教授
山崎 匠	電気通信大学	准教授
創薬・医療		
執筆者		
荒木 望嗣	京都大学大学院医学研究科人間健康科学系専攻	特定准教授
沖田 浩平	日本大学生産工学部	教授
奥野 恭史	京都大学大学院 医学研究科	教授
川下 理日人	近畿大学理工学部	准教授
須永 泰弘	高度情報科学技術研究機構	課長
千葉 修一	ジャパンメディカルデバイス株式会社	CTO(最高技術責任者)
野田 茂穂	理化学研究所	上級研究員
福澤 薫	大阪大学大学院 薬学研究科	教授
本間 光貴	理研創薬化学 AI アプリケーションユニット	ユニットリーダー
望月 祐志	立教大学	教授
八木 清	理化学研究所	専任研究員
設計・製造		
執筆者		

(次ページに続く)

大西 順也	理化学研究所 計算科学研究センター 複雑現象統一的解法研究チーム	研究員
河合 宗司	東北大学大学院工学研究科 航空宇宙工学専攻	教授
坂本 尚久	神戸大学大学院システム情報学研究科 計算科学専攻	准教授
高木 亮治	宇宙航空研究開発機構 宇宙科学研究所	准教授
沼波 政倫	自然科学研究機構核融合科学研究所 研究部	教授
吉村 忍	東京大学大学院工学系研究科 システム創成学専攻	教授
執筆協力者		
石澤 明宏	京都大学大学院エネルギー科学研究所 エネルギー基礎科学専攻	教授
井戸村 泰宏	日本原子力研究開発機構 システム計算科学センター	室長
今寺 賢志	京都大学大学院エネルギー科学研究所 エネルギー基礎科学専攻	准教授
藤堂 泰	自然科学研究機構核融合科学研究所 ヘリカル研究部	教授
前山 伸也	自然科学研究機構核融合科学研究所 メタ階層ダイナミクスユニット	准教授
渡邊 智彦	名古屋大学大学院理学研究科 素粒子宇宙物理学専攻	教授
社会科学		
地震・津波		
執筆者		
堀 高峰	海洋研究開発機構 地震津波予測研究開発センター	センター長
藤田 航平	東京大学 地震研究所 計算地球科学研究センター	准教授
執筆協力者		
市村 強	東京大学 地震研究所 計算地球科学研究センター	センター長
気象・気候		
執筆者		
足立 幸穂	理化学研究所 計算科学研究センター 複合系気候科学研究チーム	研究員
伊藤 純至	東北大学 理学研究科	准教授
大越智 幸司	海洋研究開発機構 地球環境部門 環境変動予測研究センター 地球システムモデル開発応用グループ	准研究副主任
大塚 成徳	理化学研究所 計算科学研究センター データ同化研究チーム	研究員
川崎 高雄	東京大学 大気海洋研究所 気候システムモデリング研究部門	特任助教
川畑 拓矢	気象研究所 気象観測研究部 第4研究室	室長
河宮 未知生	海洋研究開発機構 地球環境部門 環境変動予測研究センター	センター長
小玉 知央	海洋研究開発機構 地球環境部門 環境変動予測研究センター 雲解像モデル開発応用グループ	副主任研究員
佐藤 陽祐	北海道大学 大学院理学研究院	准教授

(次ページに続く)

佐藤 芳昭	気象庁 情報基盤部 数値予報課 数値予報モデル基盤技術開発室	室長
滝川 雅之	海洋研究開発機構 北極環境変動総合研究センター 北極化学物質研究グループ	グループリーダー
寺崎 康児	理化学研究所 計算科学研究センター データ同化研究チーム（現 気象研究所 気象観測研究部 第4研究室）	研究員
西澤 誠也	理化学研究所 計算科学研究センター 複合系気候科学研究チーム	研究員
八代 尚	国立環境研究所 地球システム領域 衛星観測研究室	主任研究員
三好 建正	理化学研究所 計算科学研究センター データ同化研究チーム	チームリーダー
執筆協力者		
大野 知紀	海洋研究開発機構 地球環境部門 環境変動予測研究センター 雲解像モデル開発応用グループ	特任研究員
川瀬 宏明	気象研究所 応用気象研究部	主任研究員
富田 浩文	理化学研究所 計算科学研究センター 複合系気候科学研究チーム	チームリーダー
Le Duc	東京大学 工学系研究科 附属総合研究機構	助教
編集協力者		
佐賀 真理子	理化学研究所 計算科学研究センター 複合系気候科学研究チーム	研究パートタイマー
宇宙・天文		
執筆者		
滝脇 知也	国立天文台	准教授
執筆協力者		
天野 孝伸	東京大学	准教授
石山 智明	千葉大学	准教授
岩崎 一成	国立天文台	助教
岩本 昌倫	京都大学	特定研究員
斎藤 貴之	神戸大学	准教授
住吉 光介	沼津工業高等専門学校	教授
長倉 洋樹	国立天文台	国立天文台フェロー
原田 了	理化学研究所	基礎科学特別研究員
堀田 英之	名古屋大学	教授
松清 修一	九州大学	教授
松本 洋介	千葉大学	准教授

計算科学ロードマップ 2023
2023 年 12 月 26 日 (第 0 版)

HPCI コンソーシアム
計算科学フォーラム 編

※ 最新版は下記に掲載予定
<https://hpcic-kkf.com>