

jh180069-NAH

GW space-time コードの大規模な有機-金属界面への適用に向けた高効率化

柳澤 将 (琉球大学)

概要 *GW* space-time コードによる大規模な表面・界面系の効率的な計算を可能にするため、今年度は、主に複素誘電行列の、LU 分解による逆行列化計算のノード内タスク並列化を検討した。結果、スレッド数の増加に伴う並列化の効果が認められ、行列サイズの増加に応じてその効果が大きくなることも認められた。表面などの二次元系の電子状態に向けた計算法についても、非周期的な方向でクーロンポテンシャルを打ち切る方法の実装を進めており、表面・界面系への適用が可能になってきている。これまでの共同研究によるプログラムの改善により、表面・界面系への適用も含め、以前よりも適用可能な問題・サイズが増え、応用計算の成果も出てきている。

1. 共同研究に関する情報

(1) 共同研究を実施した拠点名

東北大学サイバーサイエンスセンター

(2) 共同研究分野

- 超大規模数値計算系応用分野
- 超大規模データ処理系応用分野
- 超大容量ネットワーク技術分野
- 超大規模情報システム関連研究分野

(3) 参加研究者の役割分担

- ・柳澤 将 (琉球大学・代表) プログラム開発・テストおよび、実物質系の大規模計算の実行。
- ・江川 隆輔 (東北大学・副代表) プログラム効率化のための実装および助言。

2. 研究の目的と意義

本研究では、有機電子材料と関連のある物質系の基礎的な電子物性の理解に資するため、量子力学や電磁気学などをもとに物質の電子物性を予測・再現する第一原理電子状態計算のプログラムの整備・開発を進める。特に、有機分子膜と金属電極との接触する領域 (界面) でのキャリア (電子・正孔) 注入障壁の正確な予測に向けた第一原理計算 (*GW* 近似) を可能にするため、大規模な並列計算を効率的に実行することを目的に、MPI 並列と OpenMP とのハイブリッド並列の実装を行ってプログラムの効率向上を進める。

GW 近似の計算は最近、国内外の物性の理論研究で広く実行される一方、日本国内では、大規

模な計算資源を活用し有機電子材料系に適用した例や、表面・界面系に適用した例は少ない。本研究計画の達成により、その状況を改善し、大規模な *GW* 近似の計算による国内の大型計算機資源の使用を促すことにつなげたいと考える。

GW 近似は物質中のグリーン関数に基づく方法で、一体有効ポテンシャルではなく、電子間相互作用をより精度良く考慮した電子準位 (純粒子エネルギー) を計算し、半導体のバンドギャップなどの予測の信頼性が高い。しかし、一般に、空間中の任意の二点間の非局所的な多体電子間相互作用の記述が必要で、バンド計算で頻出の密度汎関数理論 (DFT) の計算ほどプログラムコードの効率の良い並列化が容易ではなく、計算に時間がかかる。本研究計画の達成で、*GW* 近似をはじめとする、DFT 計算よりも高度な第一原理計算法の大規模計算に向けた実装を促すことにもつながると考えられる。

また、有機半導体の電子物性の実験研究は日本では歴史と実績があり、現在でも分野を先導する最先端の成果が得られている。その一方、実験的観測の難しい電子状態に関する理論的解釈でも先端の成果を得て、総合的に有機半導体の研究分野を国際的にリードする体制の構築にもつなげることが必要である。本研究では、コード開発の進展とともに、大規模な計算機資源の利用により、実験事実の理論計算による洞察・解釈の深化や、新規の成果も目指す。

このような目的・動機の下、昨年度から *GW* 近似の計算プログラムコード *GW space-time* (*GWST*) をもとに研究を開始した。昨年度は、複素誘電行列の逆行列化に使用する *ScaLapack* のライブラリの使用をやめ、*Lapack* のライブラリを使用してノードごとに独立して逆行列化を行うようコードを改良し、一定の成果を得た。

本年度はその成果に続き、ライブラリの逆行列化コードをノード内でタスク並列実行することから検討を開始し、それ以外の箇所のノード内タスク並列化についても検証・実装を進めている。本報告書では、逆行列化コードとして *Lapack* のライブラリの LU 分解を用いた逆行列化コードでの、ノード内タスク並列のテスト結果を報告する。

また、*GWST* プログラムによる比較的規模の大きい応用として、*GWST* コードの適用目標である表面系について必要な実装・テストを開始したので、その成果も報告する。

3. 当拠点公募型共同研究として実施した意義

2016 年度まで、東北大学サイバーサイエンスセンター（東北大学 CC）と共同研究を進め、*GW* 近似プログラムの整備・開発や、それを用いた有機半導体の電子物性計算で少なからず成果を収めてきた。特に、通常の PC クラスタよりも 1 コアあたりのメモリの多い *SX-ACE* を使用することで、比較的大きなサイズの問題にもプログラムを適用できる利点がある。また、平面波基底関数に基づく第一原理計算コードであるため、ベクトル化で得られる利点も大きい。

昨年度より、東北大学 CC からプログラム開発の支援を受けながら、効率のよい大規模な並列計算を可能にし、計算対象をバルク結晶だけでなく表面・界面などの並進対称性のない系にまで拡大するため、計算機資源の使用料金で大きなメリットのある貴拠点公募型共同研究として進めている。特に、今年度からは東北大学 CC の関係者との定期的なディスカッションも開始しており、より綿密な共同研究の体制を確立し

ている。また、物性の計算科学コミュニティにも成果を広めることで、同センターでの開発プログラムの使用を広めることができる点も意義があると言える。

4. 前年度までに得られた研究成果の概要

固体・結晶系の準粒子エネルギーを正確に予測できる *GW* 近似の計算プログラム *GWST* を、有機半導体系の表面・界面の問題にも適用できるよう、MPI 並列とノード内タスク並列 (*OpenMP*) とのハイブリッド並列の導入によって、並列化効率の向上を試みた。プログラム中の複素誘電行列の逆行列化計算において、*ScaLapack* のライブラリの使用をやめ、並列度が増すにつれて通信時間の増大で発生するボトルネックの解消を進めた。計算データは、ノード・プロセス間では独立な変数についてノード間に分散させ、さらにノード内でタスク並列化を行うようにした。結果、現在テスト可能な 224 程度の並列度で、計算時間の短縮と並列化効率の改善が見られた [29 年度 JHPCN 最終報告書および 29 年度 JHPCN 最終報告会]。

表面の周期的スラブの *GW* 近似による取り扱い、はまだ実行できてないが、表面での有機分子の配向に依存する静電ポテンシャルの効果と、有機半導体の結晶内部での注入電荷への誘起分極の効果とを切り分けて扱う工夫を行なった。そのような計算によって、有機半導体の表面の構造に依存した電荷注入準位の正確な決定が可能となることが分かった[S. Yanagisawa, AIP Conf. Procs. **1906**, 030014-1-4 (2017); K. Yamada, S. Yanagisawa et al., Phys. Rev. B **97**, 245206-1-8 (2018)]。

5. 今年度の研究成果の詳細

A. *OpenMP* 導入による、MPI 並列とのハイブリッド並列の導入: LU 分解による複素行列の逆行列化計算のタスク並列

昨年度の、逆行列化計算の部分の MPI 並列化での改善に続き、ハイブリッド並列の実装に向

け、逆行列化計算のノード内タスク並列の検討を行った。今年度前期は、29 年度最終成果報告会でも報告した、8000 から 20000 次元程度の複素行列の逆行列化について、タスク並列化のテストを進めた。

これまでの MPI 並列化コードで用いた Lapack のインターフェースを変えずとも、新たにコンパイルオプションを追加するだけで Mathkeisan によるタスク並列化が起動できるので、その性能を検討した。

GWST コードでは計算される倍精度複素誘電行列がエルミート行列なので、LU 分解による逆行列化計算には、メモリ節約を主な理由に、上三角・下三角行列を扱う Lapack のライブラリのコード (ZHPTRF および ZHPTRI) を使用してきた。しかし、これらのコードをタスク並列化しても、並列化の効果が非常に小さいことが確認された。その主な理由は、下位で並列化のためにコールされる ParBLAS の処理よりも、LU 分解・逆行列化のコード自身の処理のコストの方が非常に大きいため、と考えている。

そこで、LU 分解と逆行列化のコードを、それぞれ、二次元配列をより高速に処理可能な ZHETRF・ZHETRI に変更した。二次元配列を扱うために書き換えるコードの箇所はそれほど多くなかったが、使用メモリの増大が懸念された。コードの各箇所の計算精度や計算量の有効数字を考慮した結果、一部の一時データの精度を落

表 1-1. SX-ACE 上での Lapack の library (ZHETRF/ZHETRI: Mathkeisan) の計測結果。タスク並列なし、またはタスク並列 (2 または 4 スレッド) による 8279 次元の複素誘電行列の LU 分解 (ZHETRF)・逆行列化 (ZHETRI) での経過時間。カッコ内は、それぞれの経過時間でタスク並列化なしの経過時間を割ったもの。

	ZHETRF (秒)	ZHETRI (秒)
no OpenMP	22.0 (1.00)	42.8 (1.00)
OpenMP (2)	13.3 (1.65)	24.4 (1.75)
OpenMP (4)	8.8 (2.50)	21.9 (1.95)

表 1-2. SX-ACE 上での Lapack の library (ZHETRF/ZHETRI: Mathkeisan) の計測結果で、14965 次元の複素行列の場合。それ以外の条件は、表 1-1 と同じ。

	ZHETRF (秒)	ZHETRI (秒)
no OpenMP	76.1 (1.00)	150.6 (1.00)
OpenMP (2)	42.7 (1.78)	134.2 (1.12)
OpenMP (4)	25.6 (2.97)	79.8 (1.89)

表 1-3. SX-ACE 上での Lapack の library (ZHETRF/ZHETRI: Mathkeisan) の計測結果で、20121 次元の複素行列の場合。それ以外の条件は、表 1-1 と同じ。

	ZHETRF (秒)	ZHETRI (秒)
no OpenMP	181.3 (1.00)	361.1 (1.00)
OpenMP (2)	99.2 (1.83)	234.5 (1.54)
OpenMP (4)	57.5 (3.15)	194.9 (1.85)

としても問題はないことが分かり、結果、それほど使用メモリは増加しない。表 1-1 から 3 に、測定されたタスク並列化の効果を示す。経過時間の計測には、簡易性能解析機能 ftrace を使用し、行列の次元は 8279、14965、および 20121 の三種類を検討した (対象物質はナフタレン: C₁₀H₈)。

まず OpenMP なしでの単体性能で、これまでの ZHPTRF・ZHPTRI のコードと比べて 1.5 倍以上の経過時間の軽減があることを確かめた。

OpenMP によるタスク並列化は、SX-ACE の最大値の 4 スレッドまで検討したところ、スレッド数の増加に伴う効率化が認められた。特に、LU 分解 (ZHETRF) については、行列の次元の増加とともに、並列化の効率が増す様子が伺える。たとえば、4 スレッドのタスク並列化によって、8279 次元ではタスク並列化しない場合の 2.5 倍の高速化だったのが、20121 次元では 3.2 倍の高速化となった。次元の増加とともに、演算処理時間の増加により、それ以外のオーバーヘッドの占める割合が減っていると考えられる。さらに次元が増すと、理想的な 4 倍の高速化に近づ

くと期待される。

しかしながら、それに続く逆行列化 (ZHETRI) では、LU 分解のように行列サイズの増加にともなう並列化の効率化が、必ずしも認められないのが分かる。基本的にいずれの行列サイズでもスレッド数の増加にともない並列化が効率化されるのは認められるが、たとえば 14965 次元でスレッド数が 2 の場合に、タスク並列なしと比較すると高速化の効果はわずか 1.1 にしかならず、これは他の行列次元の場合よりも低い。

表 2. SX-ACE 上での Lapack の library (ZHETRF/ZHETRI: Mathkeisan) の計測結果で、22997 次元の複素行列の場合 (パーフルオロペンタセン: C₂₂F₁₄)。

	ZHETRF (秒)	ZHETRI (秒)
no OpenMP	269.3 (1.00)	540.2 (1.00)
OpenMP (2)	144.7 (1.86)	324.1 (1.67)
OpenMP (4)	82.4 (3.27)	290.0 (1.86)

これまで述べたタスク並列化での特徴が、計算対象の物質や原子種に依存する傾向であるかどうかを確認するため、本年度後期には、ナフタレン単結晶以外の有機分子電子材料もいくつか検討した。その例の一つ (パーフルオロペンタセン: C₂₂F₁₄) の結果を表 2 に示す。ここでは行列の次元が 22997 で、表 1-3 よりもやや大きな次元だが、全体的な傾向は表 1-3 に近い傾向にあり、次元数がやや大きくなったことでスレッド数の増加に対する効率化が全体的に少し増している。これ以外の物質についても、並列の効率化の特徴は大体同じ傾向にあるので、本研究で見出された特徴は、物質によらず、誘電行列のサイズで概ね決まっているようである。

LU 分解と逆行列化のうち、逆行列化のコードで並列化の効果が比較的小さい上、その効果も LU 分解に比べて安定しない理由は、現時点では不明であるが、より大きなサイズのアプリケーションに向け、さらに大きな次元での実行・テストも試していくとともに、東北大学 CC の関係者

ともその原因を議論し、今後のさらなるタスク並列化の導入に生かしたいと考える。

B. 非周期系でのクーロン相互作用カットオフの実装: 表面系や、孤立分子系への適用

一般に固体・結晶系の電子状態・バンドエネルギーの第一原理計算では、周期的ポテンシャル下の電子波 (ブロッホ波) を効率よく計算するために、周期的境界条件の下で平面波基底関数を用いる。

表面などの二次元周期系では、表面垂直方向に周期性がないため、周期的繰り返し単位 (ユニットセル) を、その方向に沿って大きめに取り、隣接セル同士が相互作用しないように工夫する (スーパーセルの方法; 図 1 を参照)。スーパーセルの方法は、自由電子の振る舞いに近い典型金属などの表面では電子間相互作用が短距離的で小さいため、有効である。しかし、一般に固体中の多体電子間のクーロン相互作用は、誘起分極で遮蔽されても長距離的である。人為的なセル間の相互作用を避けるには、100 Å 以上のセル長が必要と見積られるが、そのようなサイズのスーパーセルの計算は、今日の計算機環境でも実行は容易ではない。

その状況の改善のため、2006 年に、本来相互作用すべきでない非周期的な方向のクーロン相互作用を、適当な距離 (カットオフ) で打ち切る方法が提案された [C. A. Rozzi et al., Phys. Rev. B **73**, 205119 (2006); S. Ismail-Beigi, *ibid.* **73**, 233103 (2006)]。この方法では、たとえば二次元シート物質の場合、面に垂直な方向のクーロン相互作用が、シートからの距離 z について、カットオフ z_c を超えるとゼロになる (図 1)。本年度の前期・後期では、後述の二次元系のクーロン相互作用カットオフの表式の組み込みと、その動作テストを、典型的な二次元シート物質について行なった。

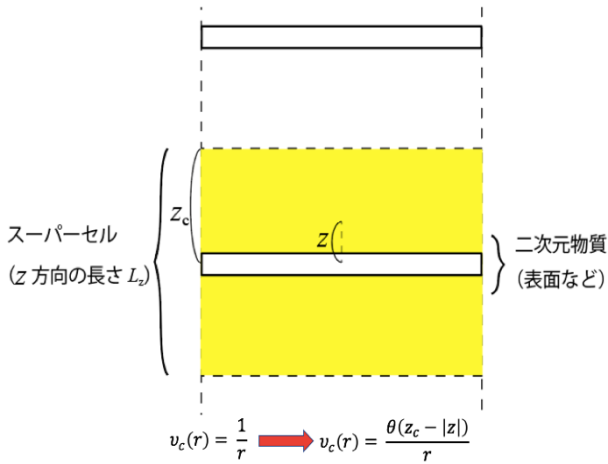


図 1. 二次元系でのクーロン相互作用打ち切りの概念図。表面からの距離 $|z|$ が z_c を超えると、クーロン相互作用はゼロとなる。

周期的ポテンシャルの下でのブロッホ波の計算では、実空間での表式よりも、逆格子空間の表式で計算の方が数値計算の収束に有利である。二次元系の、カットオフ z_c のクーロン相互作用の表式は、フーリエ変換により波数 \mathbf{k} の関数として、逆格子基底 \mathbf{G} の下で次の表式となる。

$$v_c(\mathbf{k} + \mathbf{G}) = \frac{4\pi}{|\mathbf{k} + \mathbf{G}|^2} \left\{ 1 - e^{-(\mathbf{k} + \mathbf{G})_{xy} z_c} \cos[(\mathbf{k} + \mathbf{G})_z z_c] \right\}$$

ここで $(\mathbf{k} + \mathbf{G})_{xy}$ は、逆格子ベクトル $\mathbf{k} + \mathbf{G}$ の xy 平面への正射影のノルム、 $(\mathbf{k} + \mathbf{G})_z$ は z 成分である。また、カットオフ z_c をスーパーセルの面垂直方向の長さ L_z の半分と設定することで表式が簡略化されているが、一般性は失われていない。

この表式の数値的扱いで注意すべき点は、表式の分母に現れるベクトルのノルムが数値的にゼロになり発散しうることである。そこで数値的発散を数値テクニックで回避し、クーロン相互作用カットオフの計算を実装する必要があるが、GWST コードではその実装が行われていなかった。

本研究では、GWST コードの次の特長[C. Freysoldt et al., Comput. Phys. Commun. **176**, 1 (2007)]を生かし、以下のようにクーロン相互作用カットオフの実装を試みた。

- i) ガンマ点 ($\mathbf{k} = \mathbf{0}$) における複素誘電行列 $\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{k} = \mathbf{0})$ の $\epsilon_{1,1}$ 要素 (“head”) と、 $\epsilon_{1,2..N}$ および $\epsilon_{2..N,1}$ 要素 (“wing”) を分離し、残りの $(N-1) \times (N-1)$ の行列 (“body”) だけで逆行列化計算を行う。
 - ii) $\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{k} = \mathbf{0})$ とクーロンポテンシャルの積から遮蔽ポテンシャル $W_{\mathbf{G}\mathbf{G}'}(\mathbf{k} = \mathbf{0})$ を得る。“body” 部分の計算はそのまま行うが、“wing” 部分は数値的に無視しても問題ないことが分かっており、ゼロと近似する。“head” 部分は、ここでは計算しない。
 - iii) 次に $W_{\mathbf{G}\mathbf{G}'}$ を逆フーリエ変換し、 $W_{\mathbf{r}\mathbf{r}'}$ を得る。誘電関数の “head” 部分は、遮蔽されないクーロンポテンシャルからの寄与を与えるので、その解析的表式を足し合わせるだけで、遮蔽相互作用の計算が完了する。
- このようにして、コード中での逆空間表示のクーロンポテンシャルの計算における数値的発散を回避することができた。

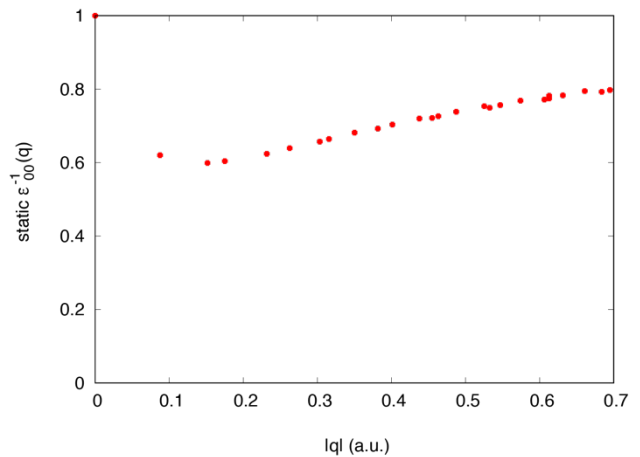


図 2. GaN 二次元シートにおける、静的な巨視的誘電関数の波数 q への依存性。

この実装のテストとして、二次元系のテスト計算を行った。GaN の二次元シートにおいて、計算された静的な巨視的誘電関数 ($\epsilon_{00}^{-1}(\mathbf{q}, \omega = 0)$) を図 2 に示す。また、ガンマ点 ($\mathbf{k} = \mathbf{0}$) での直接バンドギャップは 5.4 eV となった。これらの結果は、以前、他の GW 近似計算コードで得られた結果[S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006)]

と数値的によく一致している。この結果を見る限り、実装が成功していることが示唆される。

一方、新たなテスト対象として、半導体材料として注目されるリンの二次元シート（フォスフォレン）を検討した。本研究でのクーロンカットオフの導入の仕方に近い方法 [F. A. Rasmussen et al., Phys. Rev. B **94**, 155406 (2016)] でのバンドギャップの計算値 (1.8 eV) が最近報告された。本研究でも 1.9 eV と計算され、まずまず一致していると言える。

しかし、本研究での場合、2次元ブリルアンゾーン内の \mathbf{k} 点サンプル数を増やして電子間相互作用のおよぶ範囲を大きくすると、バンドギャップが数値的には無視できないほど減少する傾向が見られた。もしクーロンカットオフが正しく導入されていると、上述の 2016 年の論文でも見られるように、 \mathbf{k} 点サンプル数を増やしてもバンドギャップはほぼ全く変わらないのが正しい。その問題の理由を検討した結果、複素誘電行列 $\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{k})$ のブリルアンゾーン内の波数 \mathbf{k} への依存性、特に、ガンマ点 ($\mathbf{k}=\mathbf{0}$) 近傍での振る舞いは正しく記述できているが、その逆行列にクーロンポテンシャルをかけて遮蔽ポテンシャル $W_{\mathbf{G}\mathbf{G}'}$ を構築する際には、 $\mathbf{k}\rightarrow\mathbf{0}$ の極限でのガンマ点近傍の数値的扱いを当初予期していたよりも、注意深く行う必要があることがわかった。それでも、そのための改善策はほぼ分かっているため、来年度の継続課題を開始後間もなくには、より信頼性が向上したクーロンカットオフの取り扱いができることが期待できる。

C. その他の実績

昨年度からのプログラムコードの改良により、これまでよりも計算が速くなって、比較的大きな単位胞の有機分子結晶でも、計算結果の収束のための数値パラメータのチェックに要する時間を、以前よりもかなり短縮できることが分かった。このことを生かして、以前出版した論文で、計算時間の制約により詳細に検討できなかった有機半導体結晶での GW 近似によるバンド構造・

バンドギャップを正確に再現できたため、その成果を盛り込み、Book chapter として出版することができた (「7. 研究成果リスト (5) その他」、を参照)。

6. 今年度の進捗状況と今後の展望

今年度は、GWST プログラムコードのハイブリッド並列化に向けた取り組みとして、主に、SX-ACE (1 ノード内 4 コア) 上での LU 分解による逆行列化計算の、ノード内タスク並列化の検討を行った。その目的に向けて、これまでよりも高性能なライブラリのルーチンを使用し、ノード・コア単体での性能向上と、SX-ACE の最大スレッド数まで並列度を上げることによる並列化の両方を達成できた。

しかしながら、LU 分解の箇所に比べて、それに続く逆行列化のタスク並列化では、スレッド数の増加に対して並列化の効率が安定に増加しない場合がある、などの問題が見られた。この問題については、東北大学 CC の関係者との議論や、場合によっては計算機ベンダーの協力も得ながら、その原因と解決策を考えていきたい。

今後は、東北大学 CC の関係者とのディスカッションの機会を得ながら、計算プログラムの他の箇所の効率の解析も依頼し、共同研究関係をさらに深めることでプログラムコードの効率化を進めていく。

本研究計画での GW 近似の計算による達成目標の一つとして、表面系への適用を可能にする取り組みも行った。ここでは、従来提案された、クーロン相互作用カットオフの実装を GWST コードに初めて導入した。その際には、数値的な発散を回避するための工夫も行った。この作業により、表面系についても、信頼のできる取り扱いが可能になりつつある。一部、数値的な修正が必要な点も改善策はわかってきており、来年度には、実際に表面や界面での電荷注入準位のシミュレーションが可能になると予期される。

計算プログラムが以前より速くなったことにより、表面・界面のシミュレーションも含め、新

たな応用計算が射程に入ってきており、そのような成果も昨年度から引き続き得られている（「7. 研究成果リスト(5) その他」を参照）。継続課題となる 2019 年度も、共同研究利用のために提供される大規模な計算機資源を活用し、多様な成果の創出を図りたい。

また、「5. 今年度の研究成果の詳細」の項目 A, B それぞれについて、ご協力をいただいた東北大学サイバーサイエンスセンターの山下毅氏、および、ドイツ・バーダーボルン大学の Arno Schindlmayer 教授に、この場を借りて感謝申し上げます。

7. 研究成果リスト

(1) 学術論文

(2) 国際会議プロシーディングス

(3) 国際会議発表

S. Yanagisawa, “Enhancement of the *GW* space-time program code for accurate prediction of the electronic properties at surfaces and interfaces in organic electronics materials”, 29th Workshop on Sustained Simulation Performance, Cyberscience Center, Tohoku University, March 19-20, 2019 (Invited Talk).

(4) 国内会議発表

柳澤将, 「有機半導体の結晶構造と電子状態: 分子間ファン・デル・ワールス力の記述と、多体摂動論によるアプローチ」, 第二回琉球大学計算科学シンポジウム, 2018 年 10 月 19~21 日, 琉球大学工学部 2 号館.

(5) その他（特許, プレス発表, 著書等）

S. Yanagisawa and I. Hamada, A Chapter in *Theoretical Chemistry for Advanced Nanomaterials - Functional Analysis by Computation and Experiment* (Springer Nature, 2019), "Nano-scale first-principles electronic structure simulations of materials relevant to organic electronics"