

基于国产众核超级计算机的 6×10^5 核并行矩量法

顾宗静 吴昊翔 赵勋旺* 林中朝 张玉 张崎

(西安电子科技大学陕西省超大规模电磁计算重点实验室 西安 710071)

摘要: 为实现电磁计算的安全可靠和自主可控, 该文基于“天河二号”国产众核超级计算机平台, 开展大规模并行矩量法(MoM)的开发工作。为减轻大规模并行计算时计算机集群的通信压力以及加速矩量法积分方程求解, 通过分析矩量法电场积分方程离散生成的矩阵具有对角占优特性, 提出一种新型LU分解算法, 即对角块矩阵选主元LU分解(BDPLU)算法, 该算法减少了panel列分解的计算量, 更重要的是, 完全消除了选主元过程的MPI通信开销。利用BDPLU算法, 并行矩量法突破了 6×10^5 CPU核并行规模, 这是目前在国产超级计算平台上实现的最大规模的并行矩量法计算, 其矩阵求解并行效率可达51.95%。数值结果表明, 并行矩量法可准确高效地在国产超级计算平台上解决大规模电磁问题。

关键词: 矩量法; LU分解; 国产超级计算机; 6×10^5 核

中图分类号: TN820

文献标识码: A

文章编号: 1009-5896(2019)04-0845-06

DOI: 10.11999/JEIT180562

Parallel MoM Using the Six Hundred Thousand Cores on Domestically-made and Many-core Supercomputer

GU Zongjing WU Haoxiang ZHAO Xunwang LIN Zhongchao
ZHANG Yu ZHANG Qi

(Shaanxi Key Laboratory of Large Scale Electromagnetic Computing, Xidian University, Xi'an 710077, China)

Abstract: In order to realize safety, reliability and self-control of electromagnetic computing, the large-scale parallel MoM is studied based on domestically-made many-core supercomputer platform named “Tianhe-2”. A new LU decomposition algorithm named Block Diagonal matrix Pivoting LU decomposition (BDPLU) algorithm, is proposed by analyzing the diagonally dominant characteristics of the matrix generated through dispersing electric field integral equation of MoM, for the purpose of communication pressure reduction to computer cluster and solution acceleration to MoM integral equation during large-scale parallel computation. The BDPLU algorithm reduces the amount of calculation in the process of panel factorization. More importantly, the algorithm completely eliminates MPI communication when pivoting. Using BDPLU algorithm, the maximum number of CPU cores break through 6×10^5 CPU cores, which is the largest scale of parallel MoM computation in domestically-made and many-core supercomputing platform at present, and the parallel efficiency of solving matrix can reach 51.95%. Numerical results show that parallel MoM can accurately and efficiently solve large-scale electromagnetic field problems on domestic supercomputing platform.

Key words: Method of Moments (MoM); LU decomposition; Domestically-made supercomputer; 6×10^5 cores

1 引言

矩量法(Method of Moments, MoM)是电磁数值分析方法中的经典算法之一, 也是高精度的数值

算法之一, 在电磁仿真计算中得到了广泛的应用^[1]。频域矩量法需要建立并求解复数稠密矩阵方程, 为了保证求解精度, 采用基于高斯消元法的LU分解算法直接得到矩阵方程的解, 其存储复杂度和计算复杂度分别为 $O(N^2)$ 和 $O(N^3)$, 其中 N 为矩阵维数。当计算电大尺寸目标的电磁问题时会生成一个未知量十分庞大的复数稠密矩阵, 此时, 矩阵LU分解的时间占整个矩量法计算过程的90%以上^[2]。因而, 巨大的存储需求和过长的矩阵求解时间成为限制矩量法计算电大尺寸电磁问题的主要因素。突破

收稿日期: 2018-06-04; 改回日期: 2018-12-13; 网络出版: 2018-12-19

*通信作者: 赵勋旺 xwzhao@mail.xidian.edu.cn

基金项目: 国家重点研发计划(2017YFB0202102, 2016YFE0121600), 中国博士后科学基金(2017M613068)

Foundation Items: The National Key Research and Development Program of China (2017YFB0202102, 2016YFE0121600), The China Postdoctoral Science Foundation (2017M613068)

该瓶颈的常用方法之一是利用超级计算机集群进行大规模并行计算^[3,4]。

随着计算机技术的飞速发展,针对高度并行工作负载而设计的众核处理器,依靠其集成的大量计算单元,以更小的功耗,展现出比同期多核处理器更加强大的计算能力,特别适合处理矩阵LU分解这种数据密集型、计算密集型的高并发度任务。此前,矩量法已在国内超级计算机平台中的国产处理器^[5,6],通用处理器^[7,8]以及异构众核处理器^[9-12]上开展过大量的研究工作。但是,利用的众核处理器平台都是采用主机+设备(Host+Device)的异构模式,即Host端采用多核的CPU(Center Process Unit)控制任务的划分和数据传输,同时执行部分任务的计算。Device端的众核处理器以协处理器或加速卡的形式出现,执行大部分的高并发度任务,从而达到加速的效果^[13]。但是,该种异构计算模式的缺陷在于Host端和Device端的通信都是通过PCI-E接口实现的,通信速率较慢,限制了计算性能的提升。此外,在Host+Device异构超算平台进行算法的开发难度大、周期长,不利于应对复杂多变和迫切的科研需求。最为关键的是,国内大多数异构众核平台其加速卡或协处理器过分依赖国外,算法的性能提升和安全“受制于人”。

为保证科研工作安全,近年来,国内加快了对超级计算机特别是众核超级计算机的研制。“天河二号”超级计算平台,采用自主研发的FT2000+众核处理器,单核性能强大,是国内首款独立自主研发的同构众核超算平台,其同构计算系统使算法的研发和优化大为简化,在保证科研成果安全的同时,大大加快了科学研究的进度。矩量法在“天河二号”的研发,弥补了国产超算平台中电磁软件的不足,实现了电磁仿真的自主可控和安全可靠。

然而,当矩量法在超算平台进行大规模并行计算时,对于计算复杂、通信复杂的矩阵求解过程,进程间密集而频繁地通信对计算平台造成的通信压力和通信延迟是我们应该正视和亟待解决的问题。虽然,文献[14], [15]提出单向的避免通信矩阵LU分解(Communication Avoiding LU decomposition, CALU)算法,通过改变选主元策略,减少了panel列分解过程中的通信次数和通信量。而且,在通用处理器平台对CALU算法做了大量的测试工作和深入分析,其测试结果表明,CALU算法比公认的最快的商业数学库(Math Kernel Library, MKL)^[16]表现出更快的矩阵求解速度和更好的扩展性。但是,CALU算法增加了计算量,主元依然是在整个panel列中进行选取。

本文基于FT2000+同构众核处理器平台,通过分析矩量法生成的矩阵具有对角占优特性,提出一种新型的LU分解算法,即对角块矩阵选主元LU分解(Block Diagonal Pivoting LU decomposition, BDPLU)算法,该算法结合矩量法特性,改变了传统选主元策略,在未增加panel列分解计算量的前提下,完全消除了选主元过程的通信,进一步加速了矩阵求解过程,并验证了算法的正确性。

2 算法分析

2.1 并行矩量法

本文矩量法采用的基函数为RWG(Rao-Wilton-Glisson)基函数,是现今广泛使用的一种矩量法基函数,RWG基函数可模拟任意形状物体的表面电、磁流分布,采用伽略金方法可得到矩量法矩阵方程^[4]

$$[\mathbf{Z}_{mn}][\mathbf{I}_n] = [\mathbf{V}_m] \quad (1)$$

其中, $[\mathbf{Z}_{mn}]$ 表示大小为 $N \times N$ 的阻抗矩阵, $[\mathbf{I}_n]$ 和 $[\mathbf{V}_m]$ 分别表示电流列向量和激励电压列向量。为了保证计算精度,RWG基函数三角形剖分的网格边长一般在 $[\lambda/12, \lambda/12]$ 范围内^[4]。为了寻求更为高效地求解矩量法大型复数稠密矩阵方程的新方法,以下从矩量法产生的阻抗矩阵特性进行分析和推导。

2.2 矩阵特性分析

$$\begin{aligned} \mathbf{Z}_{mn}^{\pm\pm} &= \frac{jk\eta}{4\pi} \frac{l_m l_n}{4A_m A_n} \iint \left[(r_t - v_m^{\pm})(r_s' - v_n^{\pm}) - \frac{4}{k^2} \right] G(R) dS' dS \\ &= \frac{jk\eta}{4\pi} l_m l_n \left[\sum_{t=1}^7 \sum_{s=1}^7 W(t) W(s) (r_t r_s' - \frac{4}{k^2}) \right. \\ &\quad - v_n^{\pm} \sum_{t=1}^7 \sum_{s=1}^7 W(t) W(s) r_t \\ &\quad - v_m^{\pm} \sum_{t=1}^7 \sum_{s=1}^7 W(t) W(s) r_s' \\ &\quad \left. + v_m^{\pm} v_n^{\pm} \sum_{t=1}^7 \sum_{s=1}^7 W(t) W(s) \right] G(R) \quad (2) \end{aligned}$$

矩量法生成的阻抗矩阵 $[\mathbf{Z}_{mn}]$ 如公式(2)所示^[4],其中, r_t 和 r_s' 分别为场三角形和源三角形上的采样点, $W(t)$ 和 $W(s)$ 为采样点对应的权值, $v_m^{\pm}$ 和 $v_n^{\pm}$ 分别是第 m 和 n 条公共边对应的顶点坐标, $G(R) = e^{-jkR}/4\pi R$ 为自由空间格林函数, $R = |r_t - r_s'|$ 。分析式(2)可知,当 $r_t = r_s'$ 时,即场三角形和源三角形重合时, $[\mathbf{Z}_{mn}]$ 的值最大,因而可得MoM阻抗矩阵具有对角占优特性。为进一步证明该结论的正确

性,图1给出了积分方程离散生成的复数稠密矩阵元素分布。由图1对矩阵元素的描述可以看出,LU分解前,原始矩阵其较大值主要分布在对角线附近。当完成一次panel列分解,panel行更新、trailing更新后,矩阵的元素分布仍然具有对角占优特性,直到整个矩阵分解完成。对于矩量法生成的矩阵,主元主要分布在对角线附近,而CALU算法在整个panel列中进行选主元无疑产生了冗余的通信量和计算量。

2.3 BDPLU算法

BDPLU算法的前提是矩阵具有对角占优特性。算法原理如图2所示,假设图2(a)中矩阵的大小为 $N \times N$,分块矩阵大小为 $n_b \times n_b$,采用2维块循环分布方式分布在进程网格 $P_r \times P_c$ 中,当前panel列矩阵的对角分块矩阵 B_{ij} 分布在第 $P_{r(i)}$ 行进程和第 $P_{c(j)}$ 列进程中,主元只从 B_{ij} 分块矩阵中进行选取,参与选主元的进程仅为 $(P_{r(i)}, P_{c(j)})$,选主元过程完全避免了通信。而且行交换只有panel行中的进程参与执行,其他行进程不参与行交换,减少了行交换的通信。图2(b)中的 $(P_{r(i)}, P_{c(j)})$ 进程向下广播 U_{kk} 矩阵到其他行进程, $(P_{r(i+1)} \sim P_{r(i+n)}, P_{c(j)})$ 进程执行完图2(c)所示的消元与更新操作后,整个panel列分解计算完成。

以矩量法生成的双精度复数稠密矩阵LU分解

为例,分析CALU算法和BDPLU算法的通信时间。假设计算机发送1次消息的通信延迟为 α ,发送1个矩阵元素的时间为 β 。式(3)和式(4)分别给出了CALU算法和BDPLU算法panel列分解的通信时间 $T_{\text{CALU}}^{[4]}$ 和 T_{BPULU} 。可以得出,BDPLU算法其panel列分解的通信时间比CALU算法减少了 $\geq 1/3$,这对LU分解性能的提升至关重要。

$$\begin{aligned} T_{\text{BDPLU}} &= [(\alpha + \beta n_b^2) \times \log_2 P_r] \\ &= \alpha \log_2 P_r + \beta n_b^2 \log_2 P_r \end{aligned} \quad (3)$$

$$\begin{aligned} T_{\text{CALU}} &= [\alpha + \beta n_b^2] \times \log_2 P_r + (\alpha + \beta \times \frac{n_b^2}{2}) \\ &\quad \times \log_2 P_r = 2\alpha \log_2 P_r + 1.5\beta n_b^2 \log_2 P_r \end{aligned} \quad (4)$$

为进一步证明BDPLU算法的正确性和高效性。此处以飞机模型I的散射特性计算为例进行研究^[17,18]。飞机I仿真模型如图3(a)所示,平面波沿机头方向入射,沿Z轴方向极化,频率为500 MHz,矩量法所产生的未知量为58646,矩阵方程分别用BDPLU算法和CALU算法进行求解,飞机I的双站雷达散射截面(Radar-Cross Section, RCS)结果如图3(b)和图3(c)所示,并与商业软件FEKO计算结果完全吻合,证明了BDPLU算法和CALU算法在“天河二号”国产超算平台进行电磁散射计算的正确性。表1给出了BDPLU算法和CALU算法求解矩

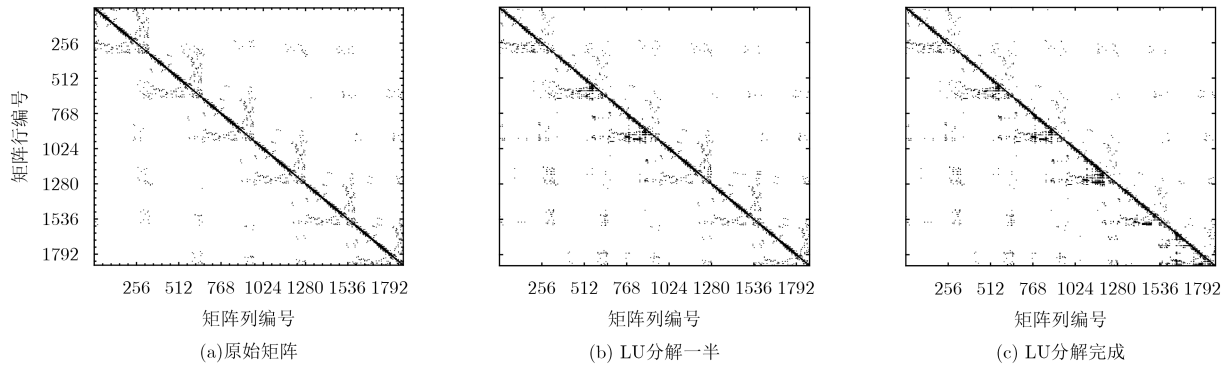


图1 LU分解过程矩阵特性分布

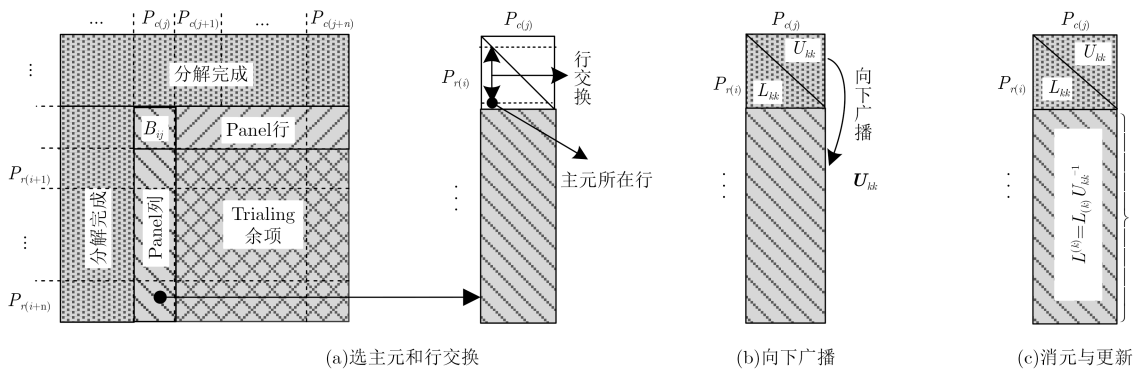


图2 BDPLU算法原理图

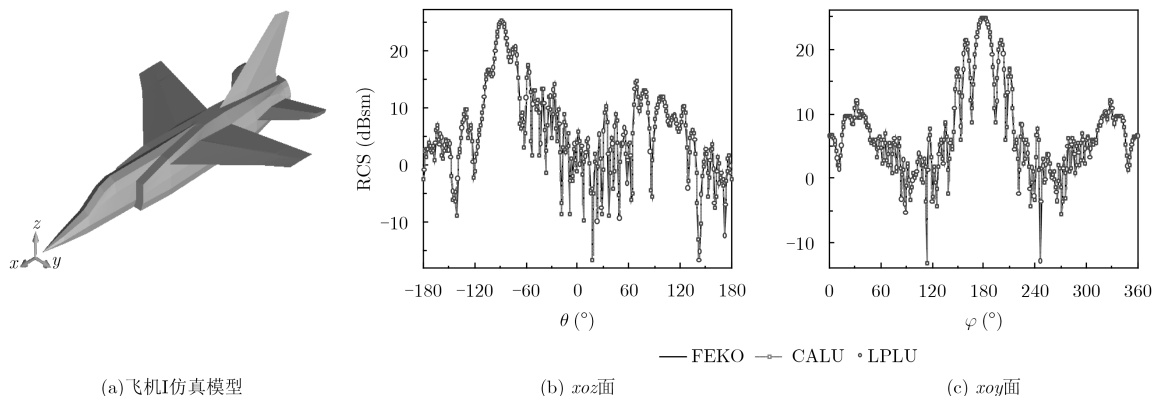


图3 飞机I仿真模型和双站RCS结果

阵的时间和效率,可以看出采用不同并行规模时, BDPLU算法总能表现出比CALU算法更快的矩阵求解速度和更好的扩展性。

表1 CALU算法与BDPLU算法矩阵求解时间对比

FT2000+核数	矩阵求解时间(s)		并行效率(%)	
	CALU	BDPLU	CALU	BDPLU
2000	796.54	742.57	100	100
3000	567.78	518.68	93.53	95.44
4000	463.93	421.12	85.85	88.17
5000	386.89	338.24	82.35	87.82
10000	226.57	187.83	70.31	79.07
15000	172.91	139.05	61.42	71.20
20000	133.97	118.38	59.46	62.73
40000	72.83	64.61	54.68	57.47

3 计算平台简介

本文使用的“天河二号”国产同构众核处理器超算平台,共有12个机柜,每个机柜有1440个FT2000+计算节点,全系统共17280个FT2000+计算节点,节点间采用25 G高速互连网络,全系统的峰值性能为10.17 PFlops (PetaFlops, 10^{15} 次浮点运算/s)。

FT2000+是国内首款自主研发的同构众核处理器,采用16 nm工艺,工作频率为2.3 GHz,每个处理器64核心,双精度峰值为588.8 GFlops,片内集成8个DDR4存储访问通道,Stream带宽初测为89 GB/S, DGEMM效率为92%, LINPACK效率为86%,操作系统为具有完全自主知识产权的“麒麟”(kylin)高性能操作系统^[19]。

4 数值算例

本文基于“天河二号”国产众核超算平台,利用并行矩量法,采用BDPLU算法求解矩阵方程,计算飞机模型II的散射特性,验证了并行矩量法可在“天河二号”国产众核超算平台进行大规模并行计算且能够有效解决电大尺寸目标电磁问题,并给出矩阵求解并行效率。

飞机II仿真模型如图4(a)所示,其几何尺寸为 $18.92\text{ m} \times 14.56\text{ m} \times 5.05\text{ m}$,平面波入射频率为400 MHz,相应的电尺寸为 $25.23\lambda \times 19.41\lambda \times 6.73\lambda$,平面波沿机头方向入射,极化方向为垂直极化,计算其双站RCS,在该频率下矩量法生成一个未知量为 190542×190542 的复数稠密矩阵。图4(b)、图4(c)给出了xoy面和xoz面的2维双站RCS结果^[17,18]。表2给出了求解该矩阵的时间和并行效率,60万核规

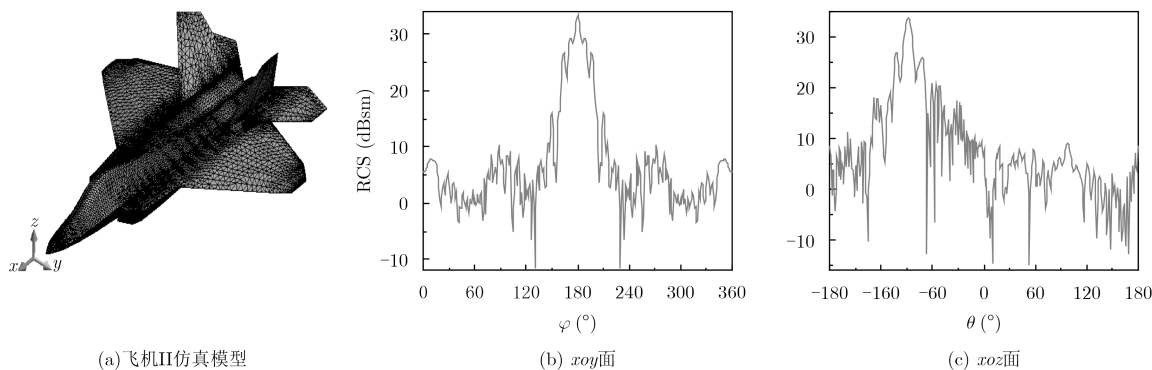


图4 飞机II双站RCS结果

表2 BDPLU算法求解矩阵的加速比和并行效率

FT2000+核数	矩阵求解时间(s)	加速比	并行效率(%)
9600	29183.33	1	100
48000	6336.53	4.61	92.11
96000	3501.73	8.33	83.34
192000	2035.35	14.34	71.69
240000	1764.53	16.54	66.16
336000	1328.13	21.97	62.78
384000	1227.91	23.77	59.42
432000	1133.64	25.74	57.21
480000	1043.45	27.97	55.94
504000	997.94	29.24	55.70
552000	937.82	31.12	54.12
600000	898.81	32.49	51.95

模是目前在国产超级计算平台中实现的最大规模的并行矩量法计算。

图5给出了该算法的并行性能曲线，以9600核为基准，当并行规模扩展到60万核时，并行效率仍

能达到51.95%，展现出了BDPLU算法良好的可扩展性。

5 结束语

本文基于“天河二号”国产众核超级计算平台，针对并行矩量法中通信复杂、计算复杂的矩阵求解过程，分析矩量法电场积分方程离散生成的矩阵具有对角占优特性，提出了BDPLU算法，完全消除了LU分解过程中panel列分解选主元的通信开销，加速了复数稠密矩阵的求解，缓解了大规模并行计算时计算集群的通信压力，实现了国产众核超算平台中并行矩量法的大规模并行计算。测试结果表明：BDPLU算法在保证结果精确的同时，比单向通信的CALU算法具有更好的性能。利用BDPLU算法，并行矩量法在“天河二号”国产众核超算平台率先突破了 6×10^5 CPU核的并行计算，并展现出了良好的并行性能。本文的研究工作为今后矩量法在完全自主研发的国产超算平台中解决更多具有实际工程意义的复杂电磁问题奠定了基础。

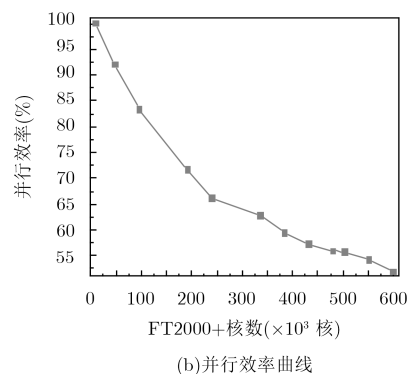
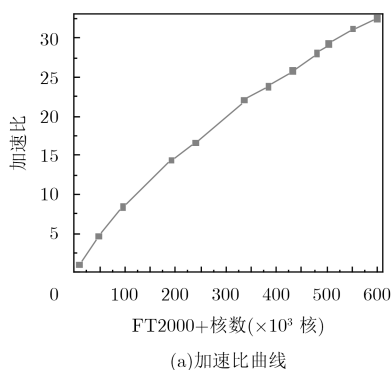


图5 加速比和并行效率

参考文献

- [1] HARRINGTON R F. Field Computation by Moment Methods[M]. New York: IEEE Press, 1993.
- [2] 王长清. 现代计算电磁学基础[M]. 北京: 北京大学出版社, 2005: 116–157.
WANG Changqing. Computational Advanced Electromagnetics[M]. Beijing: PeKing University Press, 2005: 116–157.
- [3] ZHANG Yu and SARKAR T K. Parallel Solution of Integral Equation Based EM Problems in the Frequency Domain [M]. Hoboken, USA: Wiley-IEEE, 2009: 107–136. doi: 10.1002/9780470495094.
- [4] 张玉, 赵勋旺, 陈岩, 等. 计算电磁学中的大规模并行矩量法[M]. 西安: 西安电子科技大学出版社, 2016: 11210. doi: 1002/9780470495094171.
ZHANG Yu, ZHAO Xunwang, CHEN Yan, et al. Massively Parallel Method of Moment in Computational Electromagnetics[M]. Xi'an: Xidian University Press, 2016: 11210. doi: 1002/9780470495094171.
- [5] 林中朝, 张爽, 王星, 等. 高阶矩量法在国产超级计算机上的并行性能[J]. 微波学报, 2014, 30(S1): 44–47.
LIN Zhongchao, ZHANG Shuang, WANG Xing, et al. Parallel performance of higher-order MoM on a domestically-made supercomputer[J]. Journal of Microwaves, 2014, 30(S1): 44–47.
- [6] 林中朝, 陈岩, 张玉, 等. 国产CPU平台中并行高阶矩量法研究[J]. 西安电子科技大学学报, 2015, 42(3): 43–47. doi: 10.3969/j.issn.1001-2400.2015.03.008.
LIN Zhongchao, CHEN Yan, ZHANG Yu, et al. Study of the parallel higher-order MoM on a domestically-made CPU platform[J]. Journal of Xidian University, 2015, 42(3): 43–47. doi: 10.3969/j.issn.1001-2400.2015.03.008.

- [7] ZHANG Yu, LIN Zhongchao, ZHAO Xunwang, *et al.* Performance of a massively parallel higher-order method of moment code using thousands of CPUs and its applications[J]. *IEEE Transactions on Antenna Propagation*, 2014, 62(12): 6317–6324. doi: 10.1109/TAP.2014.2361135.
- [8] ZHAO Xunwang, CHEN Yan, ZHANG Huanhuan, *et al.* A New Decomposition Solver for Complex Electromagnetic Problems[J]. *IEEE Antennas & Propagation Magazine*, 2017, 59(3): 131–140. doi: 10.1109/MAP.2017.2687119.
- [9] CHEN Yan, ZHANG Yu, ZHANG Guanghui, *et al.* Hybrid MIC/CPU parallel implementation of MoM on MIC cluster for electromagnetic problems[J]. *IEICE Transactions on Electronics*, 2016, 99(7): 735–743. doi: 10.1587/transele.E99.C.735.
- [10] CHEN Yan, ZHANG Guanghui, LIN Zhongchao, *et al.* Solution of EM problems using hybrid parallel MIC/CPU implementation of higher-order MoM[C]. IEEE International Symposium on Microwave, Antenna, Propagation, and Emc Technologies. Shanghai, China, 2016: 789–791. doi: 10.1109/MAPE.2015.7510434.
- [11] 左胜, 陈岩, 张玉, 等. 一种可扩展异构并行核外高阶矩量法[J]. 西安电子科技大学学报, 2017, 44(1): 146–151. doi: 10.3969/j.issn.1001-2400.2017.01.026.
- ZUO Sheng, CHEN Yan, ZHANG Yu, *et al.* Study of the scalable heterogeneous parallel out-of-core higher order method of moments[J]. *Journal of Xidian University*, 2017, 44(1): 146–151. doi: 10.3969/j.issn.1001-2400.2017.01.026.
- [12] CHEN Yan, ZUO Sheng, ZHANG Yu, *et al.* Large-scale parallel method of moments on CPU/MIC heterogeneous clusters[J]. *IEEE Transactions on Antennas & Propagation*, 2017, 65(7): 3782–3787. doi: 10.1109/TAP.2017.2700871.
- [13] TANG Min, ZHAO Jieyi, TONG Ruofeng, *et al.* GPU accelerated convex hull computation[J]. *Computers & Graphics*, 2012, 36(5): 498–506. doi: 10.1016/j.cag.2012.03.015.
- [14] 陈岩. 高性能矩量法及其在复杂目标电磁模拟中的应用[D]. [博士论文], 西安电子科技大学, 2017: 86–91.
- Chen Yan. High performance method of moments and its application in electromagnetic simulation of complex targets[D]. [Ph.D. dissertation], Xidian University, 2017: 86–91.
- [15] ZHANG Yu, CHEN Yan, ZHANG Guanghui, *et al.* A highly efficient communication avoiding LU algorithm for Methods of Moments[C]. IEEE International Symposium on Antennas and Propagation & Usnc/ursi National Radio Science Meeting, Vancouver, Canada, 2015: 1672–1673. doi: 10.1109/APS.2015.7305225.
- [16] Intel® Developer Zone: Intel® Math Kernel Library [OL]. <https://software.intel.com/en-us/forums/intel-math-kernel-library/>, 2018.
- [17] 徐晓飞, 曹祥玉, 高军, 等. 基于矩量法的电大目标RCS核外并行计算[J]. 电子与信息学报, 2011, 33(3): 758–762. doi: 10.3724/SP.J.1146.2010.00519.
- XU Xiaofei, CAO Xiangyu, GAO Jun, *et al.* Parallel out-of-core calculation of electrically large objects' RCS based on MoM[J]. *Journal of Electronics & Information Technology*, 2011, 33(3): 758–762. doi: 10.3724/SP.J.1146.2010.00519.
- [18] 马骥, 龚书喜, 王兴, 等. 一种快速计算目标宽带雷达截面的电磁算法[J]. 西安电子科技大学学报, 2012, 39(4): 98–102. doi: 10.3969/j.issn.1001-2400.2012.04.018.
- MA Ji, GONG Shuxi, WANG Xing, *et al.* Fast computation of the wide-band radar cross section of arbitrary objects[J]. *Journal of Xidian University*, 2012, 39(4): 98–102. doi: 10.3969/j.issn.1001-2400.2012.04.018.
- [19] 国家超级计算广州中心: 产品中心[OL]. <http://www.nscc-gz.cn/Product/HighPerformanceComputingService/ServiceCharacteristics.html>, 2018.6.
- 顾宗静: 男, 1989年生, 博士生, 研究方向为计算电磁学、大规模并行矩量法、区域分解算法.
- 吴昊翔: 男, 1995年生, 硕士生, 研究方向为计算电磁学、大规模并行矩量法.
- 赵勋旺: 男, 1983年生, 副教授, 研究方向为大型机载天线阵列分析.
- 林中朝: 男, 1988年生, 讲师, 研究方向为计算电磁学.
- 张 玉: 男, 1978年生, 教授, 研究方向为计算电磁学、大规模并行算法.