

Russian Supercomputing Days'26

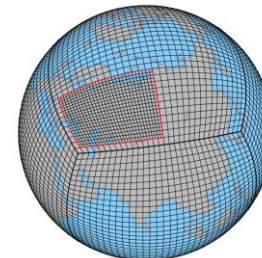
Реализация многошагового варианта метода сопряженных
градиентов в глобальной модели атмосферы на сетке
кубическая сфера

Г.С. Гойман, В.В. Шашкин

ИВМ РАН, Гидрометцентр России, МФТИ



Москва, 28.09.2026



Содержание

- Введение
- Явно-неявные аппроксимации по времени/методы решения СЛАУ
- Постановка модельной задачи
- Многоступенчатый метод сопряженных градиентов
- Численные эксперименты
 - Сходимость
 - Масштабируемость



Неявные методы интегрирования по времени

Гидротермодинамика атмосферы

- **Жесткая система**, характерные временные масштабы от секунд до месяцев
- Сильные ограничения на шаг по времени **явных** методов

Неявно-явные (IMEX) методы:

$$\frac{\partial \psi}{\partial t} = A\psi = A_{ex}\psi + A_{im}\psi$$

- Позволяет использовать бóльшие шаги интегрирования по времени
- Требуется решения **эллиптического уравнения Гельмгольца** на каждом шаге
- **Масштабируемый** и **вычислительно эффективный** метод решения линейных систем уравнений



Алгоритмы решения СЛАУ

- Параллельная масштабируемость и вычислительная эффективность

Наиболее распространенные подходы:

- Геометрический многосеточный метод:
 - + быстрая сходимость
 - низкая параллельная эффективность на грубых уровнях сетки
 - неуниверсальность относительно пространственной аппроксимации
- Методы Крыловского типа (CG, BiCGstab, GMRES, ...)
 - + приемлемая скорость сходимости
 - + универсальность
 - глобальные коммуникации на каждой итерации

Альтернативный вариант **s-step communication avoiding Krylov subspace methods**

- каждая итерация – расширение Крыловского пространства на s векторов
- алгебраически эквивалентны стандартным формулировкам
- ухудшение сходимости и устойчивости при больших s

Цель работы: исследовать применимость s -шаговых методов Крыловского типа в задачах численного моделирования атмосферы

Пример. Линеаризованные уравнения Эйлера

Линеаризованные уравнения Эйлера в (x-z)-плоскости:

$$\begin{aligned}\frac{\partial u'}{\partial t} + u_0 \frac{\partial u'}{\partial x} &= -c_p \theta_0 \frac{\partial \pi'}{\partial x}, \\ \frac{\partial w'}{\partial t} + u_0 \frac{\partial w'}{\partial x} &= -c_p \theta_0 \frac{\partial \pi'}{\partial z}, \\ \frac{\partial \pi'}{\partial t} + u_0 \frac{\partial \pi'}{\partial x} &= -\frac{R_d \pi_0}{c_v} \left(\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} \right)\end{aligned}$$

Применяя некоторую неявно-явную аппроксимацию получаем:

$$\begin{aligned}u'^{n+1} &= f_u^n - \gamma \Delta t c_p \theta_0 \frac{\partial \pi'^{n+1}}{\partial x}, \\ w'^{n+1} &= f_w^n - \gamma \Delta t c_p \theta_0 \frac{\partial \pi'^{n+1}}{\partial z}, \\ \pi'^{n+1} &= f_\pi^n - \gamma \Delta t \frac{R_d \pi_0}{c_v} \left(\frac{\partial u'^{n+1}}{\partial x} + \frac{\partial w'^{n+1}}{\partial z} \right)\end{aligned}$$

Исключим u'^{n+1}, w'^{n+1} из уравнения на π'^{n+1} и обозначим $\omega = \gamma \Delta t c_s$, где $c_s = \left(\frac{c_p \theta_0 R_d \pi_0}{c_v} \right)^{1/2}$ – скорость звука

$$\pi'^{n+1} - \omega^2 \left(\frac{\partial^2 \pi'^{n+1}}{\partial x^2} + \frac{\partial^2 \pi'^{n+1}}{\partial z^2} \right) = f^n$$

Обозначения:

$u = u_0 + u', w = 0 + w'$ – компоненты скорости
 $\pi = \pi_0 + \pi'$ – функция давления Экснера
 θ_0 – фоновая потенциальная температура
 γ – константа, зависящая от схемы по времени

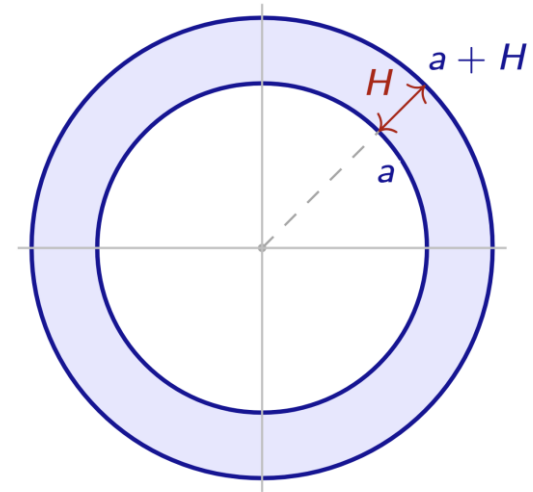
c_p, c_v – удельная теплоемкость газа при постоянном давлении/объеме
 R_d – универсальная газовая постоянная сухого воздуха

Модельное уравнение на сфере

По аналогии с 2D случаем рассматриваем модельное уравнение в сферической оболочке

$$\psi - \omega^2 \left(\Delta_S \psi + \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) \right) = f, \quad S \times [a, a + H]$$

$$\left. \frac{\partial \psi}{\partial r} \right|_{r=a} = 0, \quad \left. \frac{\partial \psi}{\partial r} \right|_{r=a+H} = 0$$



Сферическая оболочка
 $S \times [a, a + H]$

Учитываются ключевые особенности эллиптических задач в моделях атмосферы:

- сферическая геометрия
- вертикальная анизотропия ($H \ll a$)
- зависимость от шага по времени через параметр ω^2

Обозначения:

S – поверхность сферы

a – радиус Земли

H – толщина атмосферы

r – радиальная координата

Δ_S – горизонтальный оператор Лапласа

Модельное уравнение. Пространственная аппроксимация

Гномоническая равноугольная сетка кубическая сфера

$$\vec{r} = \frac{1}{\sqrt{1 + \tan^2 \alpha + \tan^2 \beta}} (\tan \alpha \quad \tan \beta \quad 1)^T \quad \alpha, \beta \in [-\pi/4, \pi/4]$$

$$\nabla \cdot \vec{v} = \frac{1}{J} \left(\frac{\partial J v^\alpha}{\partial \alpha} + \frac{\partial J v^\beta}{\partial \beta} \right) \quad \nabla \psi = \frac{\partial \psi}{\partial \alpha} \vec{a}^\alpha + \frac{\partial \psi}{\partial \beta} \vec{a}^\beta$$

Пространственная аппроксимация

По горизонтали: конечные разности со свойством суммирования по частям (SBP FD) 2го/4го порядка

По вертикали: конечные разности второго порядка

СЛАУ

$$(E - \omega^2(L_h + L_z))\psi = f$$

$$A = E - \omega^2(L_h + L_z)$$

L_h - горизонтальная часть оператора

L_z - вертикальная часть оператора

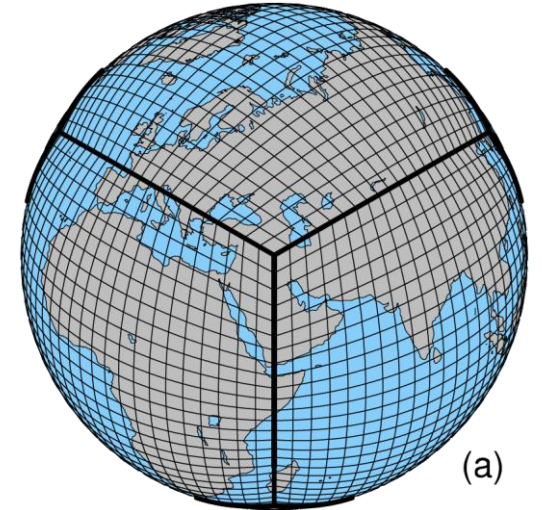
A – SPD матрица

Предобуславливатель

$$\beta \approx \left(\frac{\Delta x}{\Delta z} \right)^2 \gg 1$$

$$M = E - \omega^2 L_z$$

Применение предобуславливателя – решение набора трехдиагональных систем методом прогонки



PCG и s-step PCG (sPCG)

Решаем систему $Ax = b$ с предобуславливателем M

A, M - SPD матрицы размера n

PCG

$$\begin{aligned}\alpha^{(i)} &\leftarrow \frac{(r^{(i)})^T u^{(i)}}{(p^{(i)})^T A p^{(i)}} \\ x^{(i+1)} &\leftarrow x^{(i)} + \alpha^{(i)} p^{(i)} \\ r^{(i+1)} &\leftarrow r^{(i)} - \alpha^{(i)} A p^{(i)} \\ u^{(i+1)} &\leftarrow M^{-1} r^{(i+1)} \\ \beta^{(i+1)} &\leftarrow \frac{(r^{(i+1)})^T u^{(i+1)}}{(r^{(i)})^T u^{(i)}} \\ p^{(i+1)} &\leftarrow u^{(i+1)} + \beta^{(i+1)} p^{(i)}\end{aligned}$$

s итераций PCG:

- s умножений на A и применений M^{-1}
- $2s$ глобальных редукций
- $8sn$ локальных операций

sPCG (Mayer, Gansterer 2025)

$$S^{(k)} = \left[\phi_0(AM^{-1}) r^{(k)}, \phi_1(AM^{-1}) r^{(k)}, \dots, \phi_s(AM^{-1}) r^{(k)} \right] \in \mathbb{R}^{n \times (s+1)},$$

$$U^{(k)} = \left[\phi_0(M^{-1}A) u^{(k)}, \phi_1(M^{-1}A) u^{(k)}, \dots, \phi_{s-1}(M^{-1}A) u^{(k)} \right] \in \mathbb{R}^{n \times s},$$

$$\begin{aligned}P^{(k)} &\leftarrow U^{(k)} + P^{(k-1)} B^{(k)} \\ AP^{(k)} &\leftarrow AU^{(k)} + AP^{(k-1)} B^{(k)} \\ x^{(k+1)} &\leftarrow x^{(k)} + P^{(k)} a^{(k)} \\ r^{(k+1)} &\leftarrow r^{(k)} - AP^{(k)} a^{(k)}\end{aligned}$$

$\phi_k(A)$ – матричный полином степени k . Простейший вариант $\phi_k(A) = A^k$

1 итерация sPCG:

- s умножений на A и применений M^{-1}
- 1 глобальная редукция
- $(6s^2 + 6s)n$ локальных операций

sPCG – меньше глобальных коммуникаций за счет:

- больше локальной работы: $8sn$ против $(6s^2 + 6s)n$
- чувствительность к погрешностям округлений

Параллельная реализация

Реализация в рамках программного комплекса перспективной модели ИВМ РАН-Гидрометцентра России

MPI распараллеливание по горизонтали, геометрическая декомпозиция области

Обмены гало-зонами



Разбиение кубической сферы на 192 подобласти

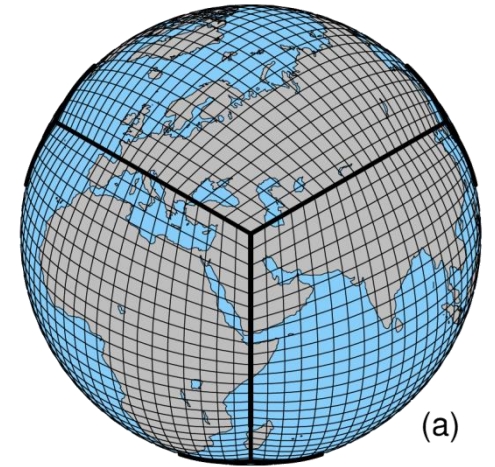
Расчеты на Cray XC40 (ГВЦ Росгидромета), Cray ARIES interconnect, Intel Xeon E2697v4 18-core CPUs, 36 CPU на узле.

Численные эксперименты

Модельное уравнение $(E - \omega^2(L_h + L_z))\psi = f$

Варируется:

- Пространственное разрешение. $C\{N_h\}L\{N_z\}$, N_h - количество узлов вдоль ребра грани, N_z - количество вертикальных уровней (C90L32)
- Горизонтальная аппроксимация SBP21/SBP42
- Характерное число Куранта звуковых волн
Режим **EU** - $\omega = 1.5\Delta x$, режим **SL** - $\omega = 7.5\Delta x$

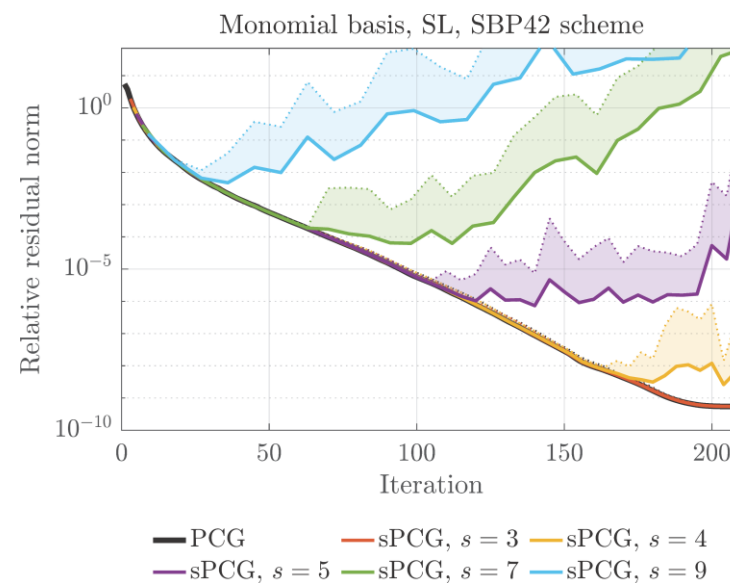
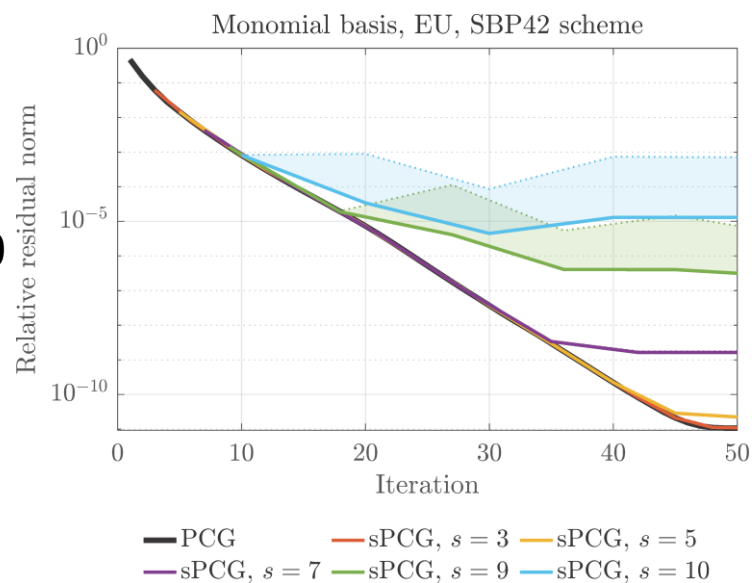
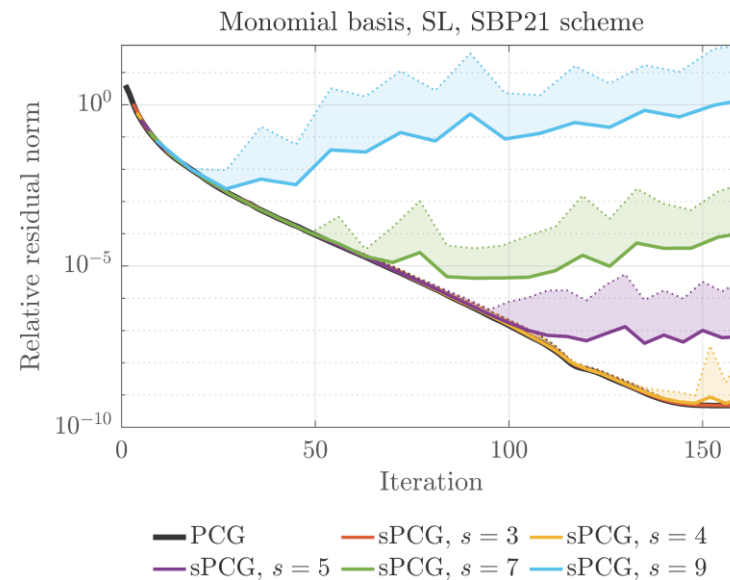
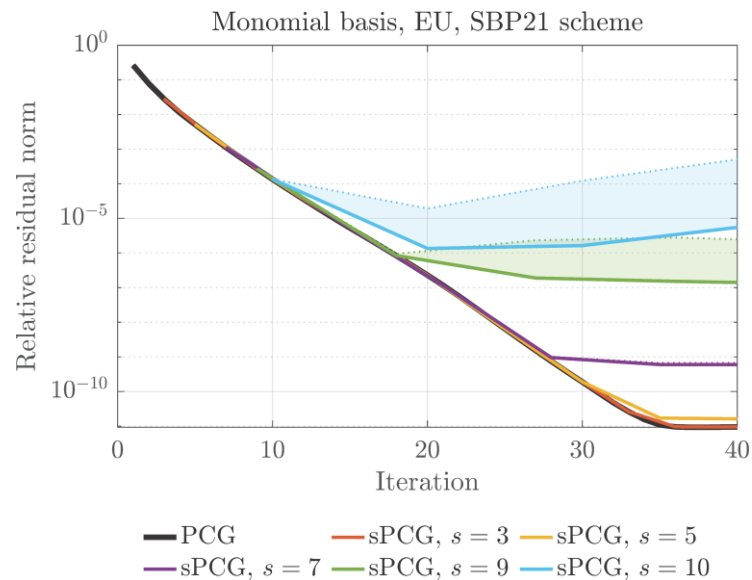


Численные эксперименты. Сходимость

- Пространственное разрешение $C64L32$
- Случайные правые части, 100 запусков, оценивается эволюция $\|b - Ax^k\|/\|b\|$
- Сплошные линии – среднее по 100 запускам
- Пунктирные линии – худший случай за 100 запусков

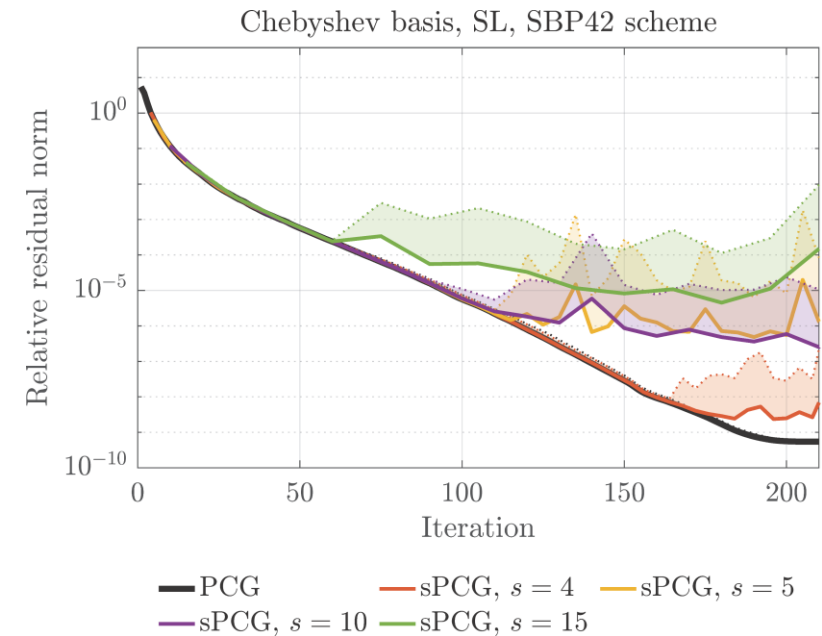
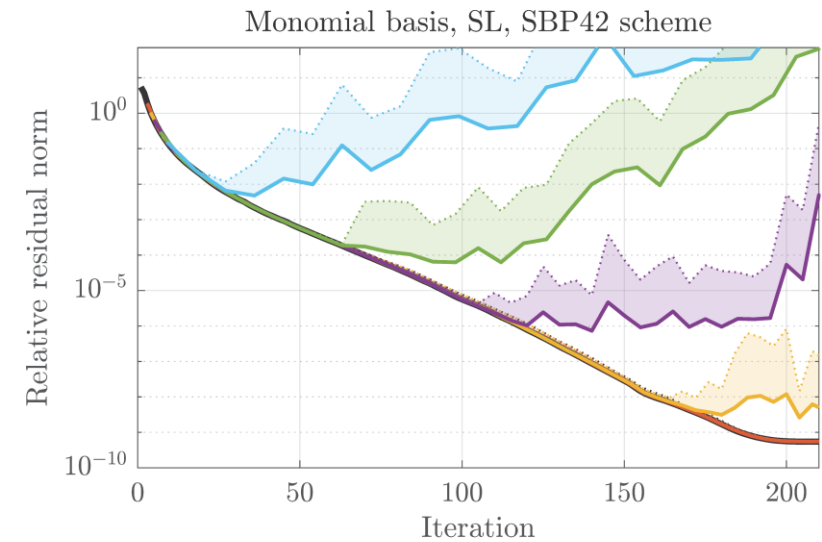
Режим EU. Допустимые значения $s \approx 9 - 10$

Режим SL. Допустимые значения $s \approx 5$



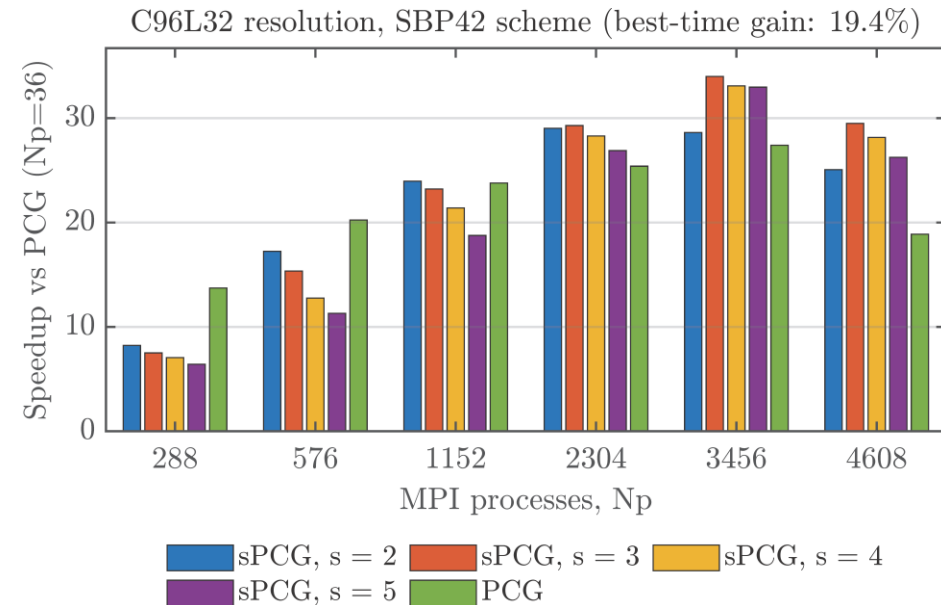
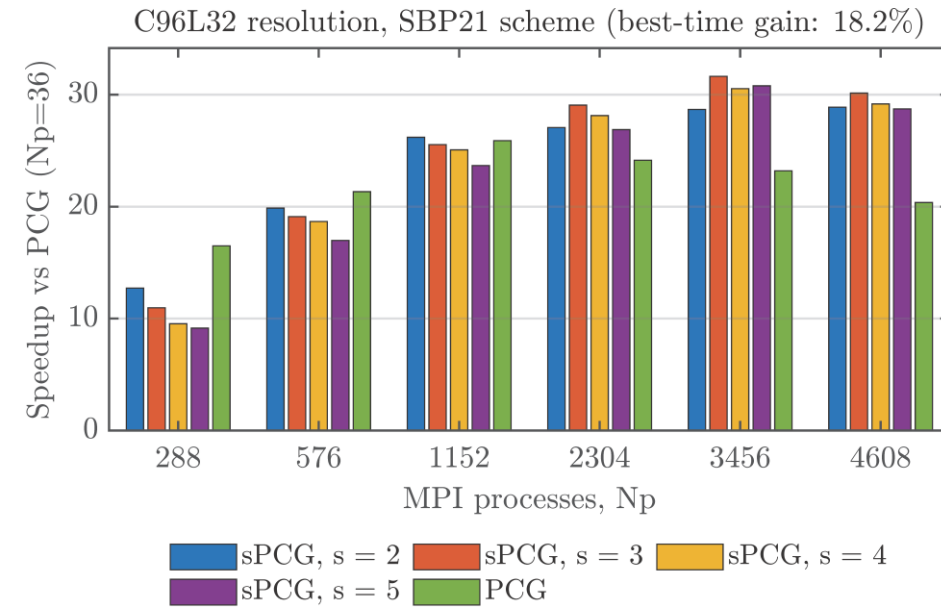
Численные эксперименты. Сходимость 2

- Пространственное разрешение $C64L32$
- Случайные правые части, 100 запусков, оценивается эволюция $\|b - Ax^k\|/\|b\|$
- Сплошные линии – среднее по 100 запускам
- Пунктирные линии – худший случай за 100 запусков
- Применение базиса на основе полиномов Чебышева ~ в два раза большие значения s



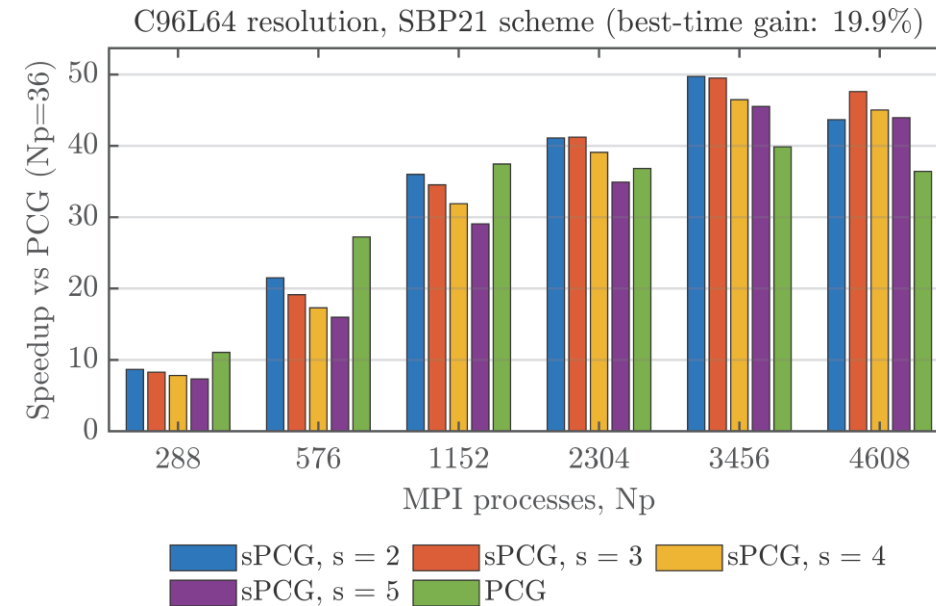
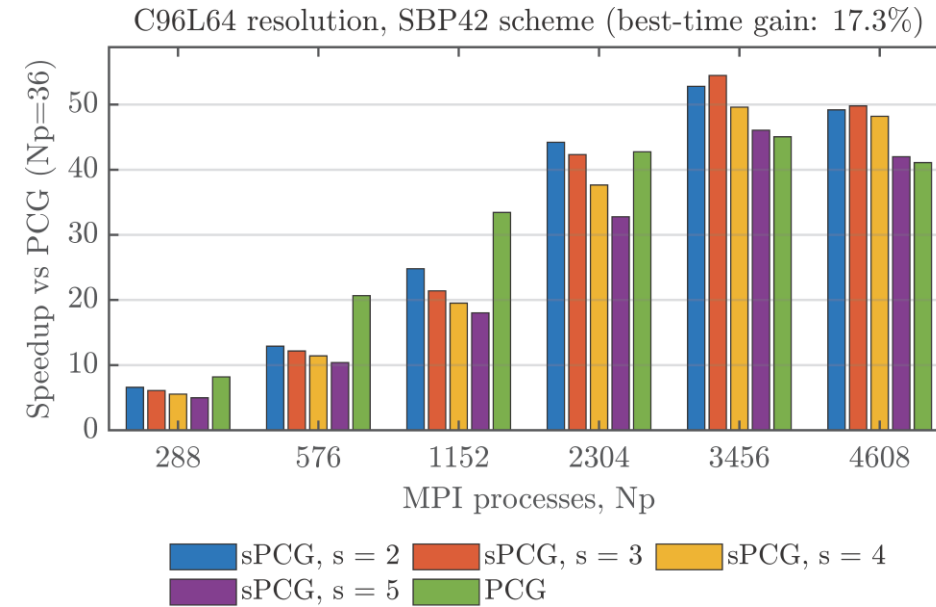
Численные эксперименты. Масштабируемость. C96L32

- Пространственное разрешение $C96L32$
- Ускорение относительно PCG с $N_p=36$
- sPCG, $s=2\dots 5$
- Лучшее время решения – снижение на ~20%
- Лучшее ускорение при фиксированном $N_p \sim 1.5$ раз



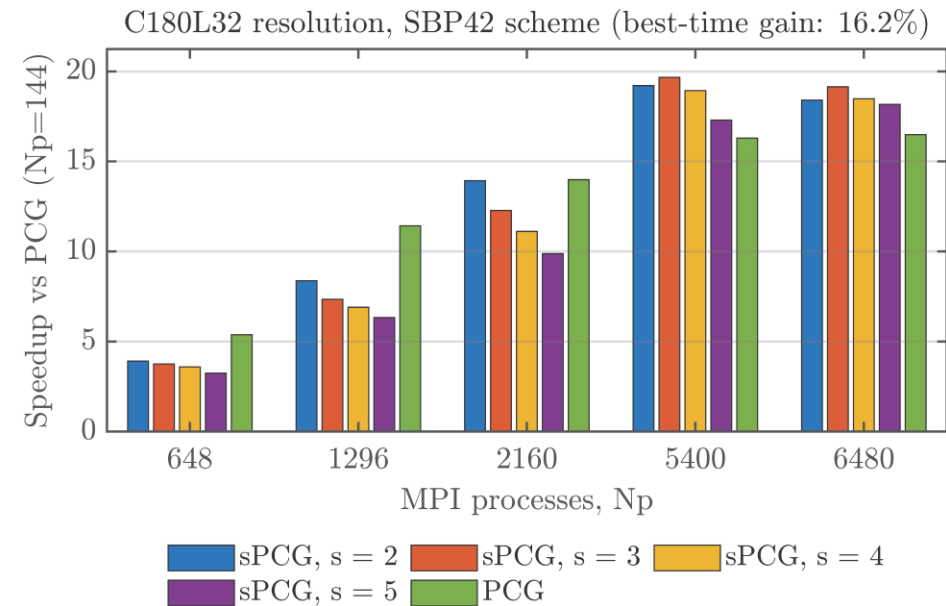
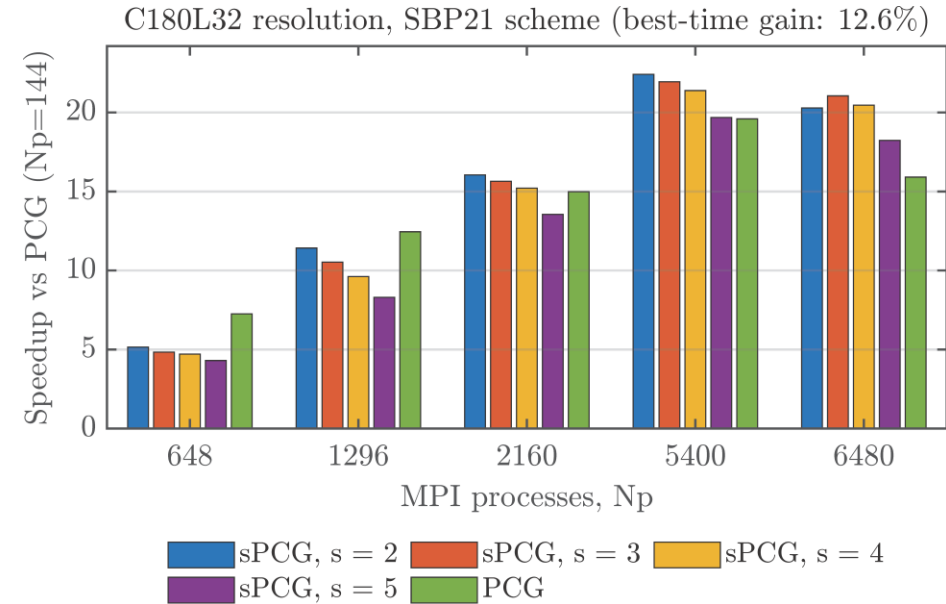
Численные эксперименты. Масштабируемость. C96L64

- Пространственное разрешение $C96L64$
- Ускорение относительно PCG с $N_p=36$
- sPCG, $s=2\dots 5$
- Лучшее время решения – снижение на ~20%
- Лучшее ускорение при фиксированном $N_p \sim 1.25$ раз



Численные эксперименты. Масштабируемость. C180L32

- Пространственное разрешение C180L32
- Ускорение относительно PCG с $N_p=36$
- sPCG, $s=2\dots 5$



Профиль Nr=2304, C96L32

- Среднее время выполнения отдельных блоков решателей в расчете на одну эквивалентную итерацию. Количество процессов 2304, разрешение C96L32

Блок	MPI_Allreduce	MatVec	Плотная лин. алгебра	Total
PCG	0.1921 ms (59.6%)	0.0967 ms (30.0%)	0.0122 ms (3.8%)	0.3222 ms (100%)
sPCG, s=2	0.0975 ms (41.0%)	0.0947 ms (39.9%)	0.0314 ms (13.2%)	0.2377 ms (100%)
sPCG, s=3	0.0719 ms (32.8%)	0.0954 ms (43.5%)	0.0391 ms (17.8%)	0.2194 ms (100%)
sPCG, s=5	0.0651 ms (27.5%)	0.1051 ms (44.4%)	0.0544 ms (23.0%)	0.2367 ms (100%)

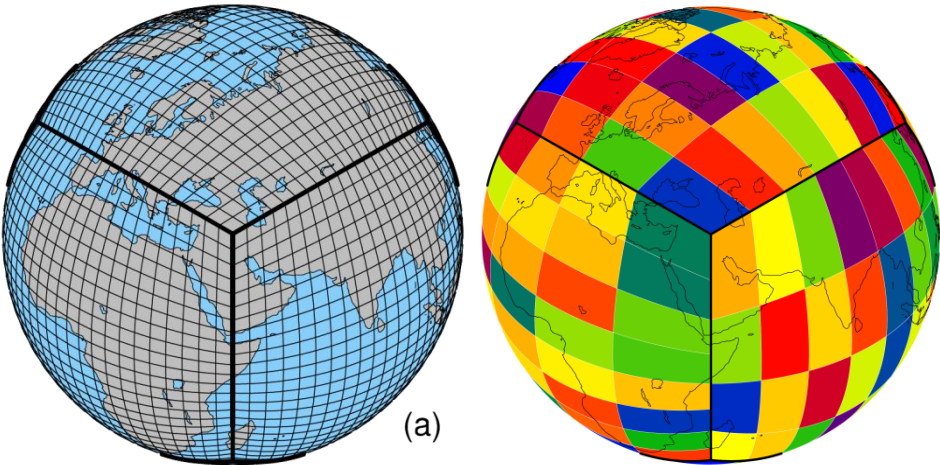
- PCG – 60% времени – mpi_allreduce
- sPCG – mpi_allreduce 60->32%, но плотная линейная алгебра 4->18%

Профиль Nr=2304, C96L32, matvec+mpi_barrier

- Среднее время выполнения отдельных блоков решателей в расчете на одну эквивалентную итерацию. Количество процессов 2304, разрешение C96L32

Блок	MPI_Allreduce	MatVec	MPI_Barrier	Плотная лин. алгебра	Total
PCG	0.0952 ms (27.4%)	0.0973 ms (27.9%)	0.1202 ms (34.5%)	0.0123 ms (3.5%)	0.3482 ms (100%)
sPCG, s=2	0.0294 ms (9.4%)	0.0967 ms (31.0%)	0.1382 ms (44.2%)	0.0316 ms (10.1%)	0.3125 ms (100%)
sPCG, s=3	0.0242 ms (8.6%)	0.0949 ms (33.9%)	0.1063 ms (38.0%)	0.0392 ms (14.0%)	0.2802 ms (100%)
sPCG, s=5	0.0242 ms (7.8%)	0.1000 ms (32.1%)	0.1177 ms (37.8%)	0.0544 ms (17.5%)	0.3111 ms (100%)

- PCG – mpi_allreduce 27.4%, matvec+barrier ~ 60%
- sPCG – mpi_allreduce ~ 9%, но плотная линейная алгебра 4->18%



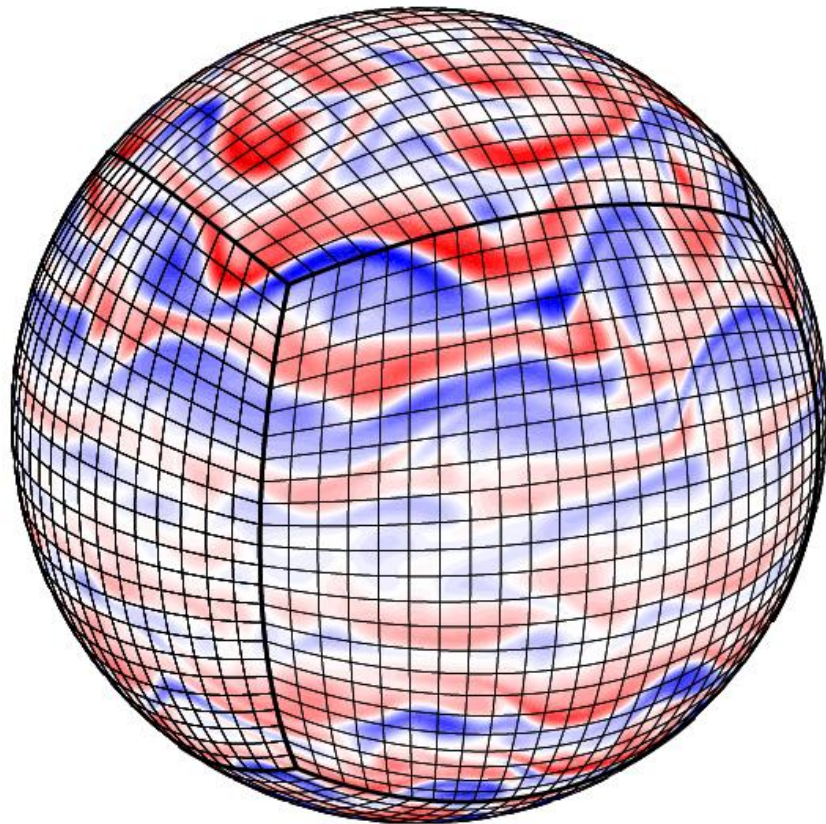
Выводы

- Применение sPCG позволяет ускорить относительно PCG решение модельных СЛАУ
- Для размера блока $s \leq 5$ достаточно использовать базис мономов -> black box решатель
- Применение полиномов Чебышева позволяет использовать $s \sim 10$ для самой сложной конфигурации модельной задачи

Дальнейшие направления работы:

- Тестирование в рамках модели атмосферы
- Тестирование других вариантов многоступенчатых Крыловских методов (GMRES, BiCGstab, ...)
- Применение одинарной/смешанной точности
- Тестирование на GPU

Спасибо за внимание!



Работы выполнены при поддержке гранта РФФ
№24-71-00123 (<https://rscf.ru/project/24-71-00123/>)

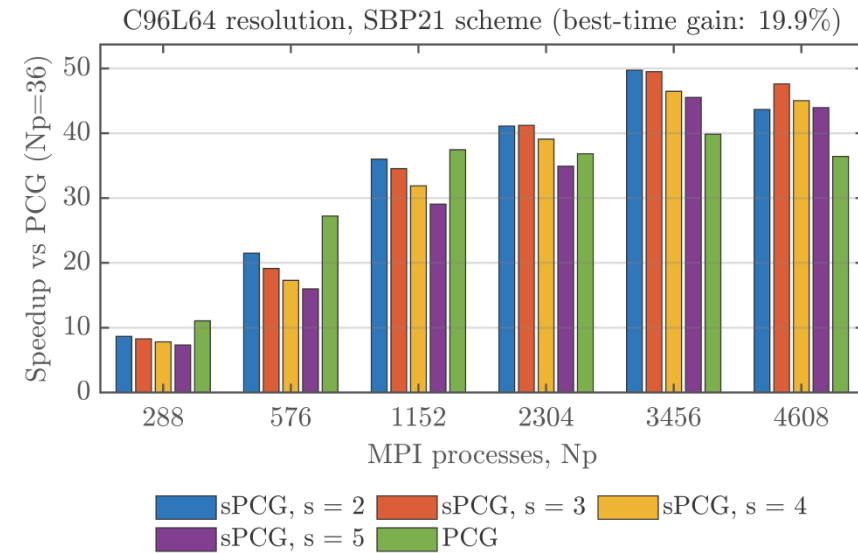
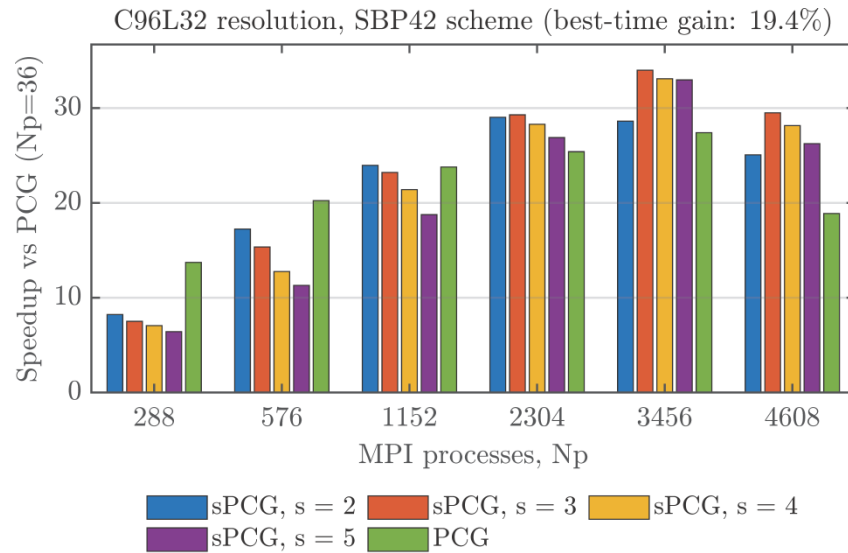
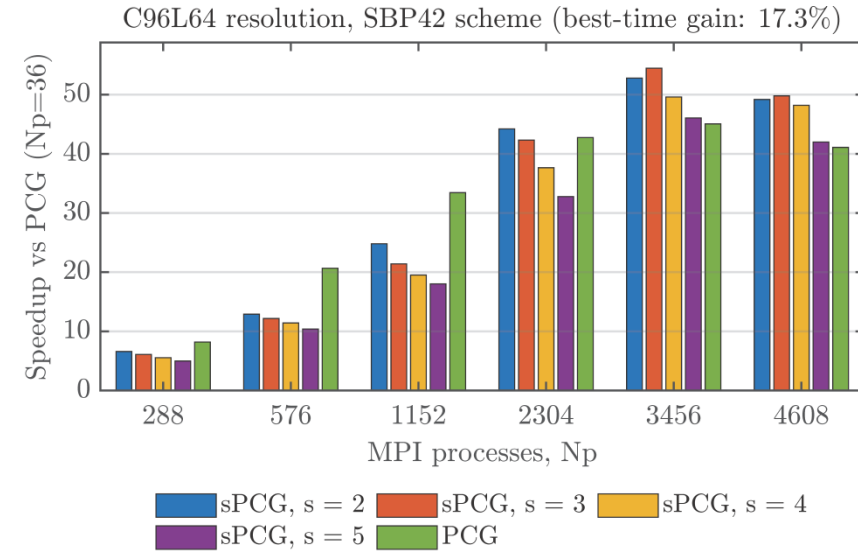
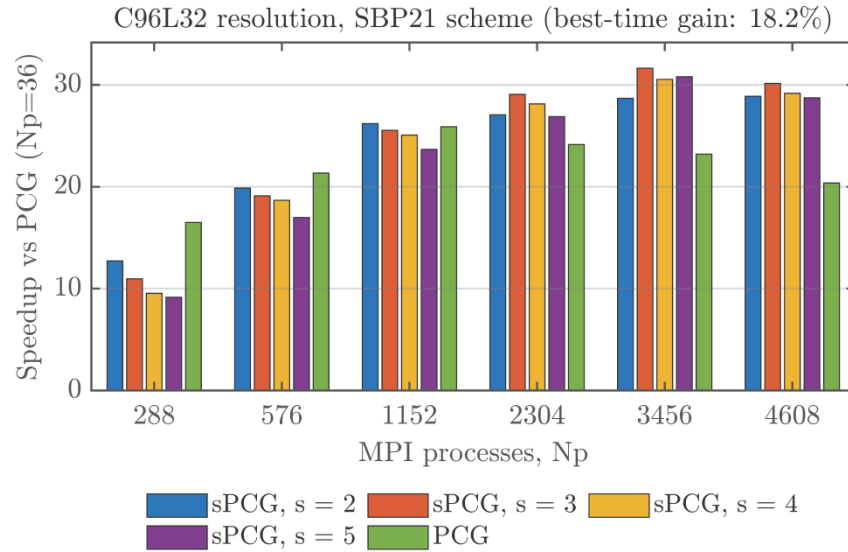
Гордей Гойман
(gordeygoyman@gmail.com)

<https://gitlab.inm.ras.ru/vshashkin/ParCS>



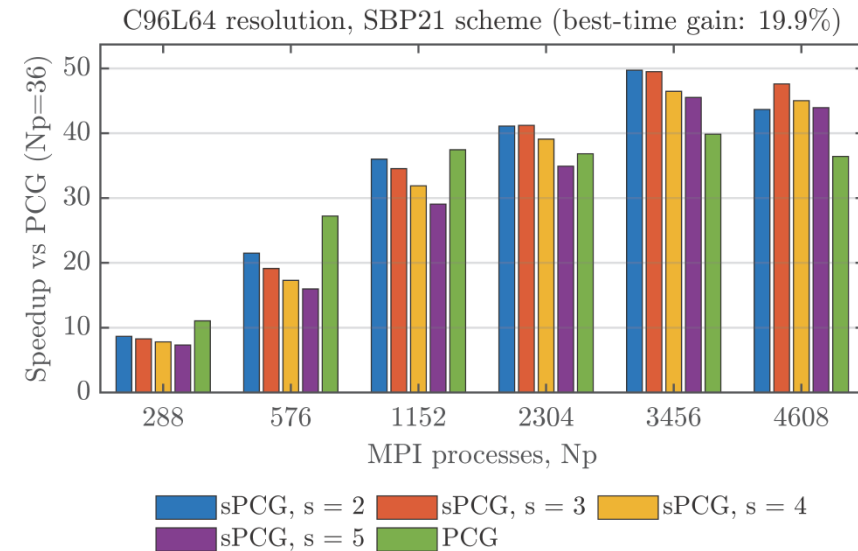
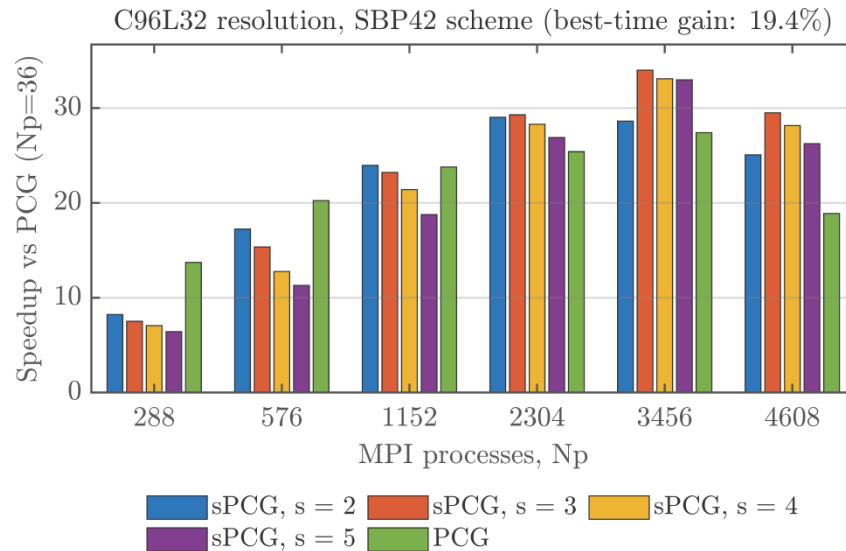
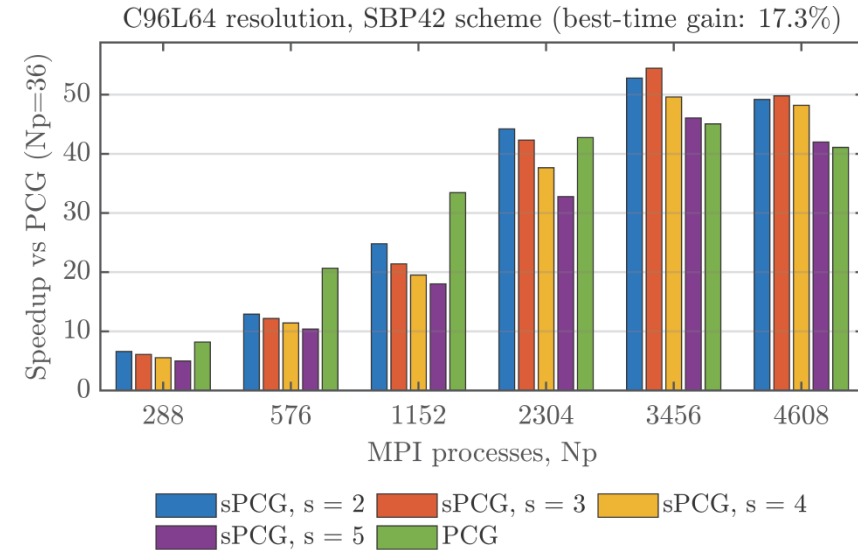
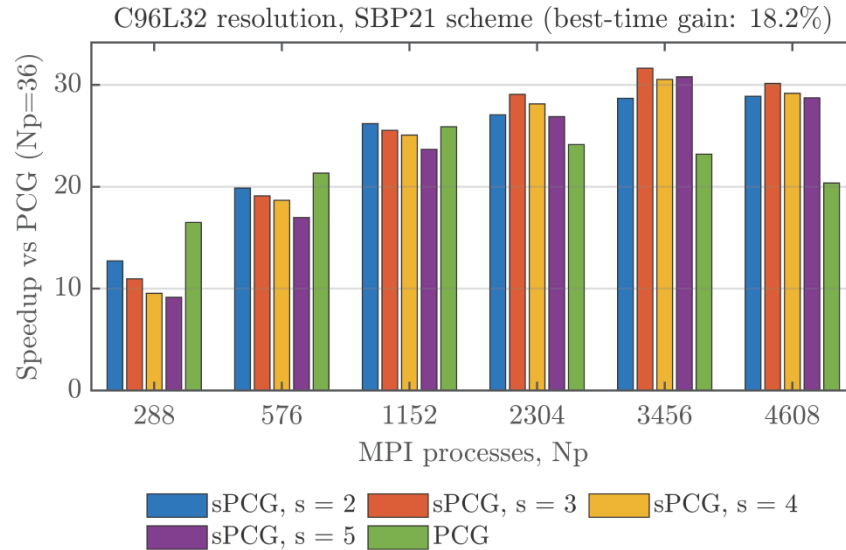
Численные эксперименты. Масштабируемость

- Пространственное разрешение $C96L32/C96L64$
- Ускорение относительно PCG с $N_p=36$
- sPCG, $s=2\dots5$
- Лучшее время решения – снижение на ~20%
- Лучшее ускорение при фиксированном $N_p \sim 1.6$ раз



Численные эксперименты. Масштабируемость

- Пространственное разрешение $C96L32/C96L64$
- Ускорение относительно PCG с $N_p=36$
- sPCG, $s=2\dots5$
- Лучшее время решения – снижение на ~20%
- Лучшее ускорение при фиксированном $N_p \sim 1.6$ раз



Численные эксперименты. Масштабируемость

- Пространственное разрешение $C96L32/C96L64$
- Ускорение относительно PCG с $Np=36$
- sPCG, $s=2\dots5$
- Лучшее время решения – снижение на ~7-16%
- Лучшее ускорение при фиксированном $Np \sim 1.4$ раза

