

yskit VASP 分段能带 workflow 使用与实测

鸢算科技 · 计算 workflow 使用文档

更新日期：2026年10月20日。软件环境：VASP 6.6.1、VASPKIT 1.5.1。本轮默认20点实测使用P100，历史回归另含V100。

当完整能带路径占用显存过多时，可以把路径分成若干次计算，再按原顺序合并。yskit 的 **菜单4「VASP 分段能带」** 已把路径准备、顺序运行、断点续算、结果检查和合并串成一个 workflow。最终保留 VASPKIT 可直接提取的 EIGENVAL、KPOINTS 等原生格式文件。

当前 workflow 按完整路径分段，保留VASPKIT默认的**每段20点**。本轮用6条路径、共120个路径点重新验证：128原子的P100单卡完整计算显存不足，逐段计算成功；双卡每卡采样峰值从 **10313 MiB降至3869 MiB，降低62.5%**。普通和HSE能带均完成拼接及VASPKIT提取对照。

1 功能怎样工作

图1 从结构准备到能带绘图数据



图1 workflow 示意。 所有分段共用同一份SCF参考，程序在同一个Slurm作业内依次启动独立的VASP进程。生成 workflow 时不提交计算。

对于LDA/GGA等半局域方法，workflow 读取收敛SCF的CHGCAR，以 `ICHARG=11` 固定密度计算各段。各段使用相同晶胞、势和物理设置，绘图统一采用原SCF的费米能。路径上的采样点不能可靠地重新确定费米能，因此不会对每段分别归零。VASP官方DFT能带流程给出了这一固定密度方法和费米能处理要求。

杂化泛函依赖规则网格上的轨道，不能只读取CHGCAR并设 `ICHARG=11`。workflow 保留同一组IBZKPT规则网格和WAVECAR，每次只附加当前路径的零权重点，使用 `ICHARG=1`。这个分段方式及其收敛注意事项见VASP官方杂化能带流程。

功能	实际行为
路径准备	隔离调用VASPKIT或vasptool生成KPATH，不覆盖原结构
按完整路径分段	$\Gamma \rightarrow X$ 算一次， $X \rightarrow M$ 再算一次；直接保留KPATH中的每段采样点数，VASPKIT默认20点
保留原始路径	保留逻辑路径顺序、原有重复端点、标签和断开
顺序计算	在一次Slurm资源分配中逐段运行，不额外并发占卡
断点续算	重新提交同一run.sh，校验后跳过已完成段，失败段先留档
合并检查	核对正常结束、电子收敛、点序、带数、电子数及共享点能量差
结果输出	自动生成merged；由VASPKIT 211或252提取绘图数据

meta-GGA采用保留规则网格与轨道的方法；本次真实数值对照覆盖普通、自旋、SOC和HSE。当前合并的是**本征值能带**，不合并PROCAR投影、WAVECAR或vasprun.xml。

2 从准备结构到拿到绘图数据

整个过程按下面的顺序进行：

最终结构 → 生成能带路径 → 同晶胞静态SCF → 生成分段任务 → 提交计算 → 自动检查并拼接 → VASPKIT提取绘图数据。

生成路径和分段任务都通过菜单操作，只需要输入目录名、按编号选择资源。下面用一个新算例说明。

第一步 准备结构并打开yskit

把已经确定的最终结构命名为 `POSCAR`，放进算例总目录下的 `structure` 文件夹。登录集群后，在终端进入自己的算例总目录，输入：

```
yskit
```

选择 **4. VASP 分段能带 → 1. 第一步：生成能带路径。**

屏幕提示	本例怎样填写
路径生成工具	选择VASPKIT；使用vasptool时选择对应项
体系维度	按材料选择三维、二维或一维
含POSCAR的目录	输入 <code>structure</code>
新路径目录	直接回车，使用 <code>band-kpath</code>

完成后，`band-kpath` 内会得到 `KPATH.in` 和路径对应的晶胞文件；VASPKIT通常生成 `PRIMCELL.vasp`。程序会显示实际文件名。

第二步 使用路径对应的晶胞完成SCF

在算例总目录建立 `scf` 文件夹。核对 `band-kpath/PRIMCELL.vasp`，将它作为SCF的 `POSCAR`，再准备所需的 `INCAR`、`POTCAR`和规则网格`KPOINTS`。如果路径工具输出其他晶胞文件，以屏幕提示为准。

SCF晶胞必须与KPATH一致。 路径工具可能把输入结构标准化为更小的原胞，不能直接把该路径套到原来的超胞上。已有正确路径和同晶胞SCF的用户，可以跳过前面的路径生成步骤。[VASPKIT路径说明](#)

SCF需要保存电荷密度，设置 `LCHARG = .TRUE.`；使用杂化或meta-GGA时还需保存波函数，即 `LWAVE = .TRUE.`。按所选方法准备其他参数，在 `scf` 目录用yskit主菜单的 **1. 复制 Slurm 脚本** 取得对应脚本，检查后提交。**等静态SCF正常完成且电子收敛，再继续下一步。** yskit的分段功能会检查已有SCF，不会代替用户完成结构优化或SCF计算。

完成的SCF目录至少包含：

```
scf/
├─ INCAR
├─ POSCAR
├─ POTCAR
├─ KPOINTS
├─ CHGCAR
├─ OUTCAR
└─ DOSCAR
```

杂化或meta-GGA还需要`WAVECAR`和`IBZKPT`。参考必须为 `NSW=0` 的自治计算；不使用结构优化、中途终止或非自治路径结果。d/f元素应在原SCF阶段正确设置`LMAXMIX`，常见要求分别为4和6。

第三步 用菜单生成分段任务

回到算例总目录，再输入 `yskit`，选择 **4. VASP 分段能带 → 2. 第二步：用已完成的SCF生成分段任务。**

屏幕提示	本例怎样填写
已完成的静态SCF目录	输入 <code>scf</code> ；若屏幕默认已是 <code>scf</code> ，直接回车
能带路径文件	输入 <code>band-kpath/KPATH.in</code> ；默认相同时直接回车
用户类型	按自己的账户选择核时用户或套餐用户
计算平台	选择P100、V100或CPU
GPU数量	按需要选择，例如1张或2张；CPU不显示此项
新 workflow 目录	直接回车，使用 <code>bands-workflow</code>
最后预览	核对路径、每段点数和资源，再选择生成 workflow

不用选择每批多少点。 程序按KPATH中成对端点逐条计算，例如 $\Gamma \rightarrow X$ 、 $X \rightarrow M$ 、 $M \rightarrow \Gamma$ 。KPATH第二行的20表示每条路径共20个采样点，包含两个端点；三条路径就是60条K点记录。程序保留文件里的数值，不会为了降低显存而减少采样点。

生成后得到 `bands-workflow`，此时还没有提交作业。程序自动识别与SCF相同的VASP版本；SOC使用相应的非共线程序。GPU资源按每卡—MPI、—CPU设置。

第四步 提交一次作业 依次算完所有路径

进入新生成的工作流目录，提交：

```
cd bands-workflow
sbatch run.sh
```

程序在同一个Slurm作业内先算第一条路径，结束后释放VASP进程资源，再启动下一条。例如先算 $\Gamma \rightarrow X$ 的20个点，再算 $X \rightarrow M$ 的20个点，最后算 $M \rightarrow \Gamma$ 的20个点。每条路径保存在独立的任务目录，不需要用户逐条手工提交。

第五步 等待自动检查和拼接

全部路径正常结束后，程序检查电子收敛、K点顺序、带数，以及不同路径共享端点的能量是否一致。检查通过后，自动生成 `merged` 文件夹，将所有路径按原顺序拼接。原有重复端点和路径断开保留，不对每段做额外能量平移。

若任一段失败或端点误差超限，程序停止并保留日志，不生成未经检查的合并结果。修复后可在原工作流目录重新提交 `sbatch run.sh`，已验证完成的段会跳过。

第六步 提取绘图数据

看到 `merged` 已生成后，仍在 `bands-workflow` 目录执行：

```
bash extract.sh
```

脚本自动调用VASPKIT，普通能带使用211，杂化等需要规则网格的结果使用252。完成后在 `merged` 中查看 **BAND.dat**、**REFORMATTED_BAND.dat**、**KLABELS**和**KLINES.dat**，可以用来绘图或导入Origin。文件用途见第7节。

阶段结束时	应得到什么
路径生成完成	<code>band-kpath/KPATH.in</code> 及对应晶胞文件
静态SCF完成	<code>scf</code> 中的收敛结果和CHGCAR等参考文件
分段任务生成完成	<code>bands-workflow/run.sh</code> 和各段输入
分段计算与拼接完成	<code>bands-workflow/merged/EIGENVAL</code> 等完整路径结果
VASPKIT提取完成	<code>merged/BAND.dat</code> 等可绘图数据

3 实测例子一 默认20点的六段高对称路径

Si2算例采用VASPKIT 303实际生成的路径，保留KPATH第二行默认的20：

$\Gamma \rightarrow X \rightarrow U \quad | \quad K \rightarrow \Gamma \rightarrow L \rightarrow W \rightarrow X$

实际计算的六段依次是 $\Gamma \rightarrow X$ 、 $X \rightarrow U$ 、 $K \rightarrow \Gamma$ 、 $\Gamma \rightarrow L$ 、 $L \rightarrow W$ 、 $W \rightarrow X$ 。每段20点，共120条K点记录。竖线表示U和K之间断开，不增加U $\rightarrow$ K路径。 workflow顺序计算6次，再与一次计算全部120点的结果比较。

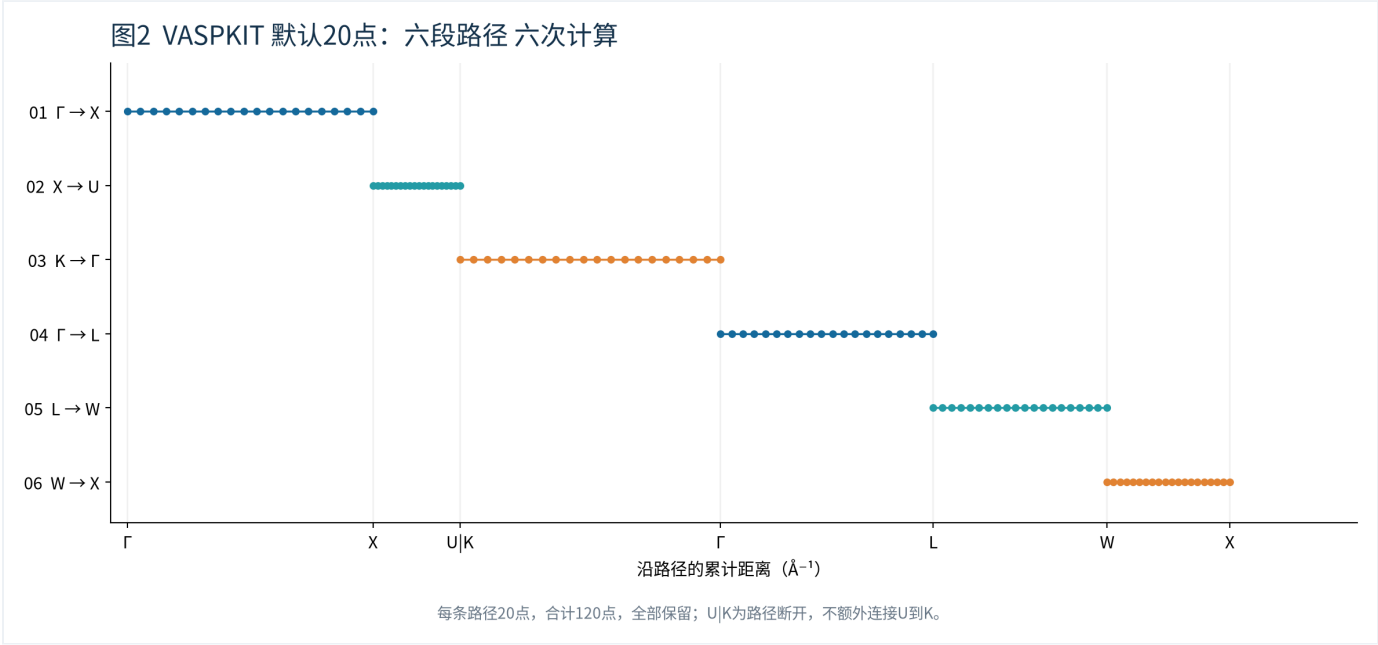


图2 每条路径对应一次完整计算。所有采样点、共享端点和路径断开均保留。

设置	本例数值
结构	Si2原胞，晶格常数5.43 Å
SCF网格	Γ中心2×2×2
路径采样	每段20点，六段共120点
主要参数	ENCUT=250 eV、EDIFF=10 ⁻⁸ eV、LREAL=FALSE、ALGO=Normal
分段方法	ICHARG=11，共同SCF密度
求解与输出	求解24带，输出前16带
最低迭代数	NELMIN=20
全部目标带最大差值	0.00000000 eV

图3 Si2 默认每段20点的完整与分段能带

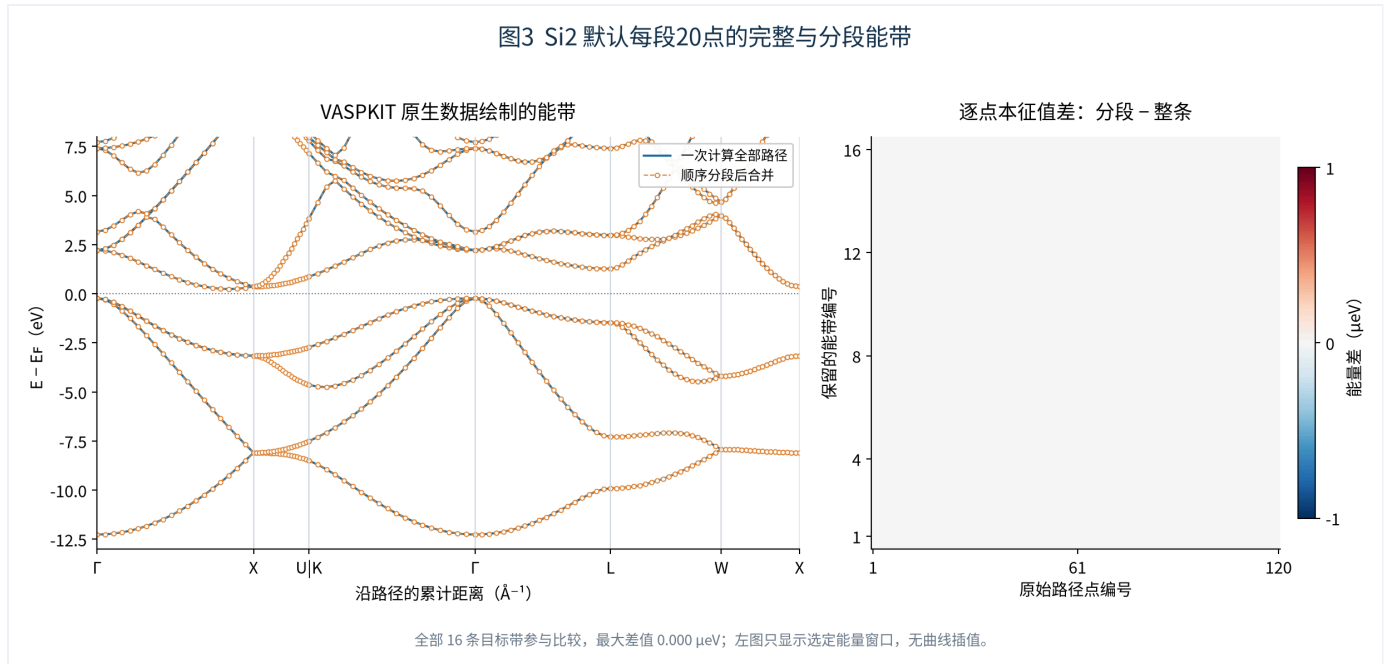


图3 蓝线为一次计算全部路径，橙色空心点为逐段计算后合并。右图比较全部120点、16条目标带；能量统一减去原SCF费米能。曲线不做平滑插值，断开处分别绘制。本例参数用于验证 workflow，材料的生产能带仍应检查SCF网格、截断能和路径密度收敛。

4 实测例子二 默认20点的HSE杂化能带

HSE也使用上面六条高对称路径，每段20点，共120个路径点。每次计算保留原SCF的3个正权重规则网格点，再附加当前路径的20个零权重点，所以**单次VASP实际包含23个K点**。六条路径仍对应六次计算。

设置	本例数值
结构与资源	Si2, P100单卡
杂化设置	LHFCALC=TRUE、AEXX=0.25、HFSCREEN=0.2 $\text{\AA}^{-1}$ 、HFRCUT=-1
算法	ALGO=Damped、TIME=0.4、EDIFF= 10^{-8} eV
重启方式	ISTART=1、ICHARG=1，保留共同规则网格和轨道
求解与输出	实际32条，输出前16条目标带
最低迭代数	NELMIN=40
与完整计算的最大差值	0.00000100 eV
VASPKIT提取入口	252

图4 HSE 默认每段20点的完整与分段能带

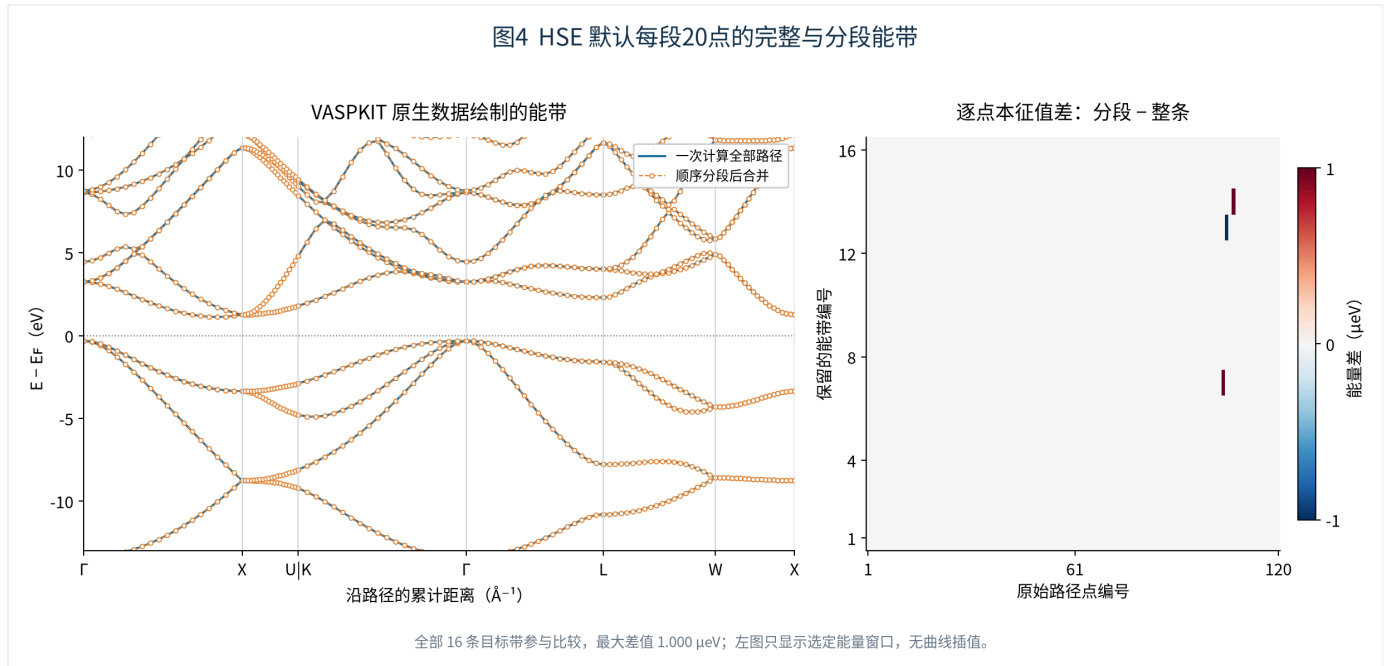


图4 HSE完整路径与分段结果。 左图统一减去原SCF费米能，右图比较全部120个路径点、16条目标带。合并EIGENVAL保留一份规则网格前缀；VASPKIT 252提取时将它与绘图路径区分。

5 实测例子三 128原子体系的显存不足与分段结果

Si128使用 $4 \times 4 \times 4$ 个Si原胞组成的FCC超胞，SCF为 Γ 单点。计算该超胞倒空间中的六段路径，每段20点，共120点；这是超胞折叠能带，不是原胞能带展开。两种计算方式使用同一份收敛CHGCAR、相同ENCUT=250 eV、EDIFF= 10^{-6} eV、LREAL=Auto，以及**实际求解352带、输出320带**。

完整计算一次放入全部120点；分段计算则依次运行六条完整路径，每次20点，既不减少总采样点，也不再切开单条路径。

图5 Si128：完整路径与逐条路径计算

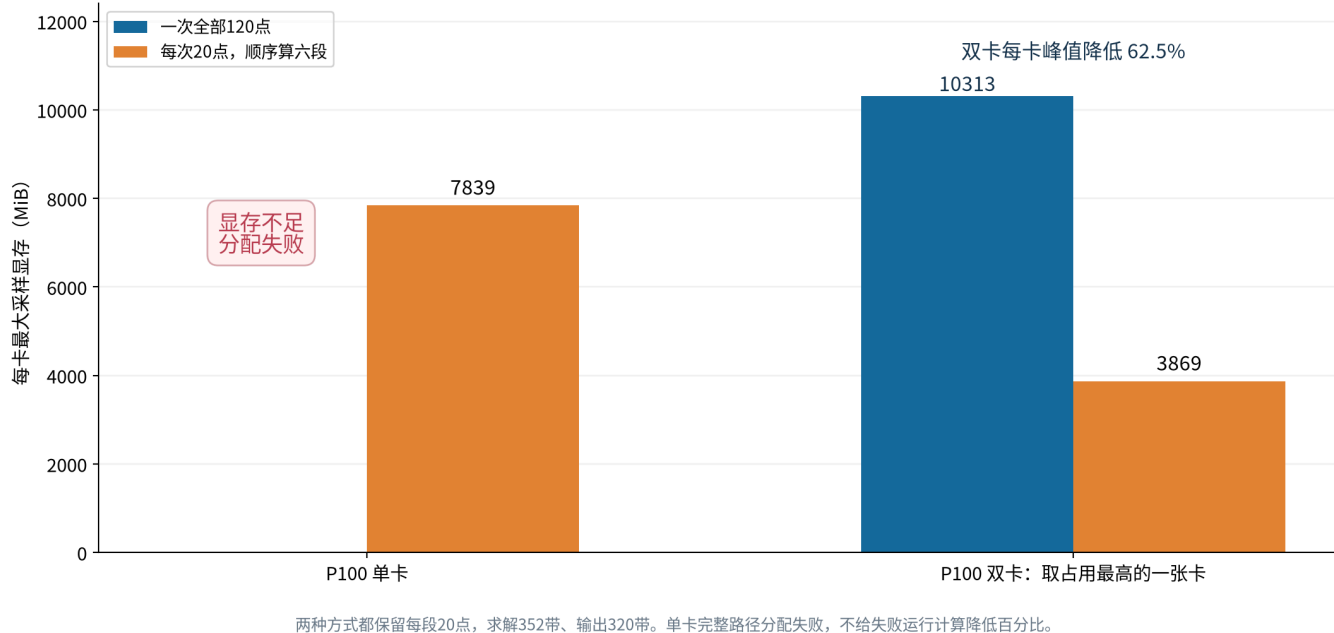


图5 同组输入下的实际结果。单卡完整计算在申请GPU内存时明确报错 CUDA_ERROR_OUT_OF_MEMORY，分段成功完成全部六段。双卡完整和分段都完成，以占用最高的一张卡比较峰值。

资源	一次全部120点	每次20点 顺序六段	显存结果
P100单卡	显存不足，未完成	完成，峰值7839 MiB	将本例从完整路径OOM变为可完成
P100双卡，取最高卡	完成，峰值10313 MiB	完成，峰值3869 MiB	每卡峰值降低62.5%

单卡完整计算没有成功分配所需显存，失败前观测到的占用不能当作“完整计算所需显存”，因此不给单卡计算降低百分比。单卡分段的全部320条目标带与成功的双卡完整路径结果比较，最大差值为0.00000000 eV；双卡分段与双卡完整计算的最大差值为0.00000000 eV。

图6 Si128 默认20点的原始显存曲线

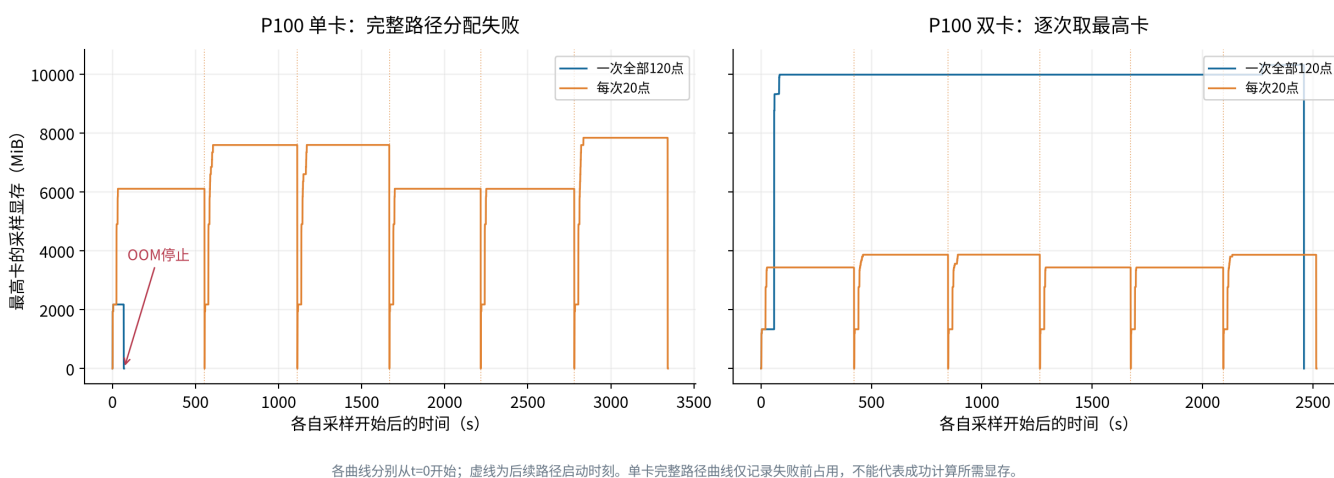


图6 原始显存采样曲线。各次计算从自身采样开始计时，虚线标出后续路径的启动时刻。单卡完整路径曲线在OOM时终止。目标采样间隔250 ms，实际间隔包含查询开销，可能漏掉短暂峰值。

分段降低了同时保留的路径相关数据，但每段仍有网格、轨道和求解工作区等开销。**完整的一条路径也必须能装入显存**。本例结果不能直接推广到所有100或200原子材料；元素、带数、截断能、真空层和泛函都会影响占用。分段还会增加进程启动和重复准备工作，不把省显存等同于加速。

6 为什么还要缓冲带和接缝检查

VASP达到总能量收敛条件时，最高几条空带仍可能没有充分收敛。本次测试观察到了重复k点能量不一致，程序因此拒绝合并。**VASP的NBANDS说明**也指出，接近所求最高态的本征矢通常收敛更慢。

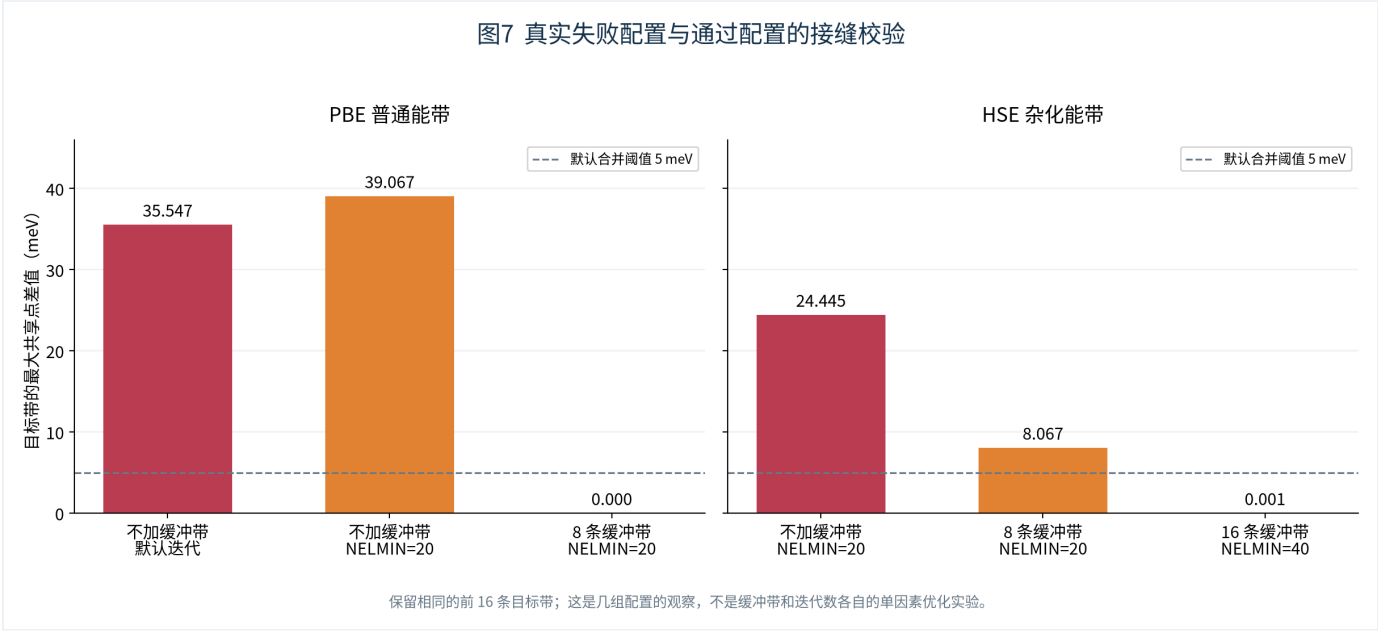


图7 保留的早期收敛测试结果。 此图用于说明共享点检查，不是本轮每段20点显存对照。普通方法仅提高最低迭代次数仍未通过；增加求解余量后，保留的16条目标带通过。HSE最终配置同时增加了缓冲带和最低迭代次数，因此不能把改善全部归因于某一个参数。

设置	普通固定密度	杂化或meta-GGA
默认额外缓冲带	至少8条或目标带数10%，取较大者	至少16条或目标带数20%，取较大者
默认最低迭代数	20	40
默认共享点合并阈值	0.005 eV	0.005 eV

实际求解带数还会按并行要求取整。缓冲带只参与求解，最终按明确声明的目标带数裁掉；目标能量行来自真实输出，不作逐段平移。

若接缝仍超差，先保留日志并检查SCF、目标空带和收敛设置，需要时增加求解余量后在新目录生成。默认输出带数不低于原SCF实际带数。不要通过逐段能量平移或随意放宽阈值掩盖误差。

7 最终文件怎样提取和画图

分段正常完成并通过检查后自动生成：

bands-workflow/	
├─ run.sh	提交与续算
├─ extract.sh	VASPKIT提取
├─ state.json	各段状态
├─ inputs/	共同SCF输入快照
├─ tasks/	各段真实输出
└─ merged/	
├─ EIGENVAL	合并后的目标能带
├─ KPOINTS	完整路径或规则网格加完整路径
├─ KPATH.in	完整路径及标签
├─ POSCAR	
├─ INCAR	
├─ POTCAR	
├─ DOSCAR	原SCF副本，用于共同费米能
└─ merge-report.json	来源与校验结果

在工作流目录执行：

```
bash extract.sh
```

脚本先检查输入与合并文件哈希，再加载VASPKIT 1.5.1并选择对应任务。也可以把merged整个目录复制到其他位置，手动提取：

```
cd merged
module load vaspkit/1.5.1
vaspkit
```

普通、自旋或SOC固定密度能带输入 **211**；杂化及规则网格加零权重路径输入 **252**。完成后输入0退出。

输出文件	用途
BAND.dat	每条带的路径距离和能量，可直接绘图
REFORMATTED_BAND.dat	按路径点展开的多列能量，便于导入Origin等工具
REFORMATTED_BAND_UP.dat和DW.dat	自旋极化时的两个通道重排数据
KLABELS	高对称点或测试点的标签与横轴位置
KLINES.dat	绘制路径分界线的坐标数据

DOSCAR在这里提供原SCF费米能，不代表重新完成了完整路径的DOS计算。整套结果是对各段真实本征值的聚合，不伪造一次完整VASP运行的OUTCAR或XML。

随附两组可以直接尝试VASPKIT提取的普通能带示例和杂化能带示例。它们是本次实际计算的后处理数据，已在新目录重新提取并确认主要绘图文件与原结果逐字节一致；操作方法见示例说明。

8 查询进度和断点续算

输入 `yskit`，选择 **4. VASP 分段能带** → **3. 查看已有 workflow 状态**，填写 workflow 目录，例如 `bands-workflow`。终端会显示总段数、完成段数及是否已经合并。

中断后在原 workflow 目录再次执行 `sbatch run.sh`。已完成段必须通过收敛、点序、带数和输出哈希检查，才会被跳过；失败段先留档再重新运行。重新提交时，终端会显示哪些段已验证并跳过，不需要用户另行拆分或拼接。

需要改变采样或参数时，在**新目录**生成 workflow。原 workflow 包含绝对路径和冻结输入，生成后不移动；可独立复制的是 merged 结果目录。

9 本轮验证范围与配套数据

实际测试	原子数	资源	每段点数与总点数	顺序任务	输出带数	最大能量差
普通能带	2	P100 1卡	20×6=120	6	16	0.00000000 eV
HSE杂化	2	P100 1卡	20×6=120	6	16	0.00000100 eV
较大体系单卡	128	P100 1卡	20×6=120	6	320	0.00000000 eV
较大体系双卡	128	P100 2卡	20×6=120	6	320	0.00000000 eV

四组均按VASPKIT默认每段20点、每条路径一个任务完成，检查了原生EIGENVAL、本征值、VASPKIT提取的距离和KLABELS。Si128单卡完整路径OOM，其分段数值以双卡完整路径为参考；其余各组与相同卡数的完整计算对照。“0 eV”表示输出记录精度内相同。

此前的自旋、SOC、V100等九组回归记录仍保留；它们的路径点数和任务划分属于各自历史测试，不混入上面的20点显存结果。CPU、其他VASP版本及meta-GGA尚未逐一完成本轮真实数值对照。

溶剂、固定占据等部分特殊输入不在当前自动流程范围。半局域SCF使用四面体法时，固定密度路径计算改用ISMEAR=0、SIGMA=0.05，保持SCF密度不变；杂化或meta-GGA的四面体输入会被拒绝，以免丢失网格连接信息。

完整数值见能带逐点对照CSV、显存采样CSV、验证汇总和数据说明。配套的PBE/HSE后处理示例也使用本轮每段20点的实际结果。