

AtomPPGen：模守恒赝势的生成、KB 形式与可转移性验证

杨远青 学号：22300190015

2025 年 12 月 21 日

1 问题描述

平面波基组密度泛函理论（Density Functional Theory, DFT）在固体与材料计算中应用广泛，但其直接处理原子核附近的强库仑势会导致价电子波函数在近核区域强烈振荡，从而需要极高的平面波截断能量，计算代价昂贵。赝势（pseudopotential）的核心思想是将核附近的快速振荡用一个在近核区平滑的有效势替代，同时保持价电子在核外区域的散射性质与能级结构，从而显著降低平面波基组规模。

本作业围绕 AtomPPGen 项目展开：该项目实现了模守恒（Norm-Conserving, NC）赝势生成的一个教学级工作流，包括 Troullier – Martins (TM) 伪化、半局域势反演、Kleinman – Bylander (KB) 可分离非局域形式构造，以及可转移性验证工具（范数守恒、对数导数匹配、幽灵态检测）。项目在线文档：bud-primordium.github.io/AtomPPGen/。

本文系统说明：

1. 模守恒赝势的理论动机与基本条件；
2. TM 伪化、势反演与 KB 变换的数学形式与数值实现要点；
3. AtomPPGen 与其依赖 AtomSCF 的程序结构与数据流；
4. 以 Al (LDA) 为主的验证结果（范数误差、对数导数指标、幽灵态分类），并简述 Si、Na 的对比案例；
5. 与 NIST 原子数据的对比与主要误差来源讨论。

本文中所有能量均采用 Hartree 原子单位 (Ha)，径向坐标采用 Bohr 半径 (a_0)。

2 理论基础

2.1 模守恒赝势概述

2.1.1 为什么需要赝势

对原子核的库仑吸引势 $-Z/r$ 会使全电子 (All-Electron, AE) 价层波函数在近核区域具有很大曲率，并伴随正交化条件导致的节点结构。在平面波表示中，近核振荡对应大量高频分量，从而需要高截断能量 E_{cut} 才能准确描述。赝势方法用一个在 $r \leq r_c$ 内平滑的势 $V_{\text{PS}}(r)$ 替代真实势的核附近部分，使价电子波函数在核附近无节点、连续可微，同时在核外区域保持与 AE 解一致，从而降低基组需求。

2.1.2 模守恒 (Norm Conservation) 条件

对每个角动量通道 l ，给定一个参考价态（通常是该通道的最外层占据态）的 AE 径向规约波函数 $u_{\text{AE}}(r)$ 与能量 ε_l ，模守恒赝势要求构造伪径向函数 $u_{\text{PS}}(r)$ ，满足：

1. 核外一致性：在截断半径 r_c 外，伪波函数与 AE 波函数完全一致

$$u_{\text{PS}}(r) = u_{\text{AE}}(r), \quad r > r_c. \quad (1)$$

2. 核内平滑性：在 $0 \leq r \leq r_c$ 内无节点且足够光滑（高阶导数连续），避免非物理振荡。

3. 范数守恒：核内积分范数一致

$$\int_0^{r_c} |u_{\text{PS}}(r)|^2 dr = \int_0^{r_c} |u_{\text{AE}}(r)|^2 dr. \quad (2)$$

直观上，范数守恒确保核附近电荷分布的局域信息在伪化后保持一致；在散射理论中，它与相移对能量的一阶导数相关，因此有助于提高赝势在不同化学环境下的可转移性 (transferability)。

2.2 Troullier – Martins 伪化方法

2.2.1 径向方程与记号

在球对称势场下，三维薛定谔方程可分离变量。令 $u_{nl}(r) = rR_{nl}(r)$ 为径向规约波函数，则径向本征方程为

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V_l(r) \right] u_{nl}(r) = \varepsilon_{nl} u_{nl}(r), \quad (3)$$

其中 $V_l(r)$ 为角动量通道相关的半局域 (semilocal) 势（对 AE 解而言可理解为 Kohn – Sham 有效势，不含离心项的部分）。

2.2.2 TM 伪波函数形式

TM 方法在核内区域选取指数多项式形式以保证无节点与解析性。在 $r \leq r_c$ 内，取

$$u_l(r) = r^{l+1} \exp\left(\sum_{i=0}^N a_{2i} r^{2i}\right), \quad r \leq r_c. \quad (4)$$

其中 r^{l+1} 保障原点处 $u_l(0) = 0$ 的正确行为；仅使用偶次幂使所有阶导数在 $r = 0$ 处连续；指数形式 $\exp(\cdot)$ 保证 $u_l(r)$ 在核内正定，从而避免引入额外节点。

2.2.3 连续性约束与方程组

TM 伪化要求在 $r = r_c$ 处与 AE 波函数平滑拼接。一般可要求函数值及若干阶导数连续。在 AtomPPGen 的教学实现中，采用 $N = 2$ （求解 a_0, a_2, a_4, a_6 共四个系数），并施加四个约束：

1. $u(r_c)$ 匹配；

2. $u'(r_c)$ 匹配；

3. $u''(r_c)$ 匹配；

4. 核内范数守恒。

将这些约束写成残差方程 $F(\alpha) = 0$ ，可用非线性求解器得到系数 α 。求解完成后，定义分段伪波函数

$$u_{\text{PS}}(r) = \begin{cases} r^{l+1} \exp\left(\sum_{i=0}^N a_{2i} r^{2i}\right), & r \leq r_c, \\ u_{\text{AE}}(r), & r > r_c. \end{cases} \quad (5)$$

因此在核外区域，伪波函数与 AE 波函数严格重合；这也是后续图像中 $r > r_c$ 区域吻合的直接原因。

2.3 半局域势反演

在得到伪波函数 $u_{\text{PS}}(r)$ 及其对应能量 ε_l 后，可以将径向方程反过来解出半局域势：

$$V_l(r) = \varepsilon_l + \frac{1}{2} \frac{u''(r)}{u(r)} - \frac{l(l+1)}{2r^2}. \quad (6)$$

该式说明反演对二阶导数非常敏感，因此数值实现必须避免节点附近的 u''/u 发散。

2.3.1 内区解析导数

对 TM 内区形式，令 $p(r) = \sum_{i=0}^N a_{2i} r^{2i}$ ，则可解析得到

$$\frac{u''(r)}{u(r)} = \frac{l(l+1)}{r^2} + \frac{2(l+1)p'(r)}{r} + p'(r)^2 + p''(r). \quad (7)$$

代回式(6)，离心项相消后得到更稳定的表达

$$V_l(r) = \varepsilon_l + \frac{l+1}{r} p'(r) + \frac{1}{2} [p'(r)^2 + p''(r)], \quad r \leq r_c. \quad (8)$$

该形式完全避免了数值微分与除以 $u(r)$ 的放大效应，因此在核内更稳定。

2.3.2 外区样条导数与节点保护

在 $r > r_c$ 的核外区域， $u_{\text{PS}}(r) = u_{\text{AE}}(r)$ ，一般不再具有简单解析形式。AtomPPGen 采用三次样条（cubic spline）对 $u(r)$ 做局部插值求导，从而得到 $u''(r)$ 并代回式(6)。

当 $u(r)$ 在某些点接近零（节点附近）， $u''(r)/u(r)$ 可能数值发散。为此实现中引入节点阈值：对满足 $|u(r_i)| < \epsilon_{\text{tol}}$ 的点暂不直接计算势，而在邻近的非节点区域求势后对节点区插值填充。此外，为避免极端异常值影响后续步骤，对 $V_l(r)$ 可做裁剪

$$V_l(r) \in [-V_{\max}, V_{\max}], \quad (9)$$

其中 V_{\max} 为预设上限（教学实现中取 10^3 Ha 量级）。

2.4 Kleinman – Bylander (KB) 可分离形式

2.4.1 半局域势的平面波表示困难

半局域势算符可写成角动量投影算符之和

$$\hat{V}^{\text{SL}} = \sum_{l,m} |Y_{lm}\rangle \ V_l(r) \ \langle Y_{lm}|, \quad (10)$$

其中 Y_{lm} 为球谐函数。平面波基函数 $e^{iG \cdot r}$ 不携带固定 l, m ，因此直接作用半局域势需要做球谐展开并对通道求和，在实际平面波代码中代价较高。

2.4.2 KB 变换的结构

KB 的核心思想是：选择一个角动量通道 l^* 作为局域势

$$V_{\text{loc}}(r) = V_{l^*}(r), \quad \Delta V_l(r) = V_l(r) - V_{\text{loc}}(r), \quad (11)$$

并将剩余非局域部分写成有限秩（可分离）投影形式。对每个 $l \neq l^*$ ，取参考径向函数

$$\phi_l(r) = \frac{u_l(r)}{r}, \quad (12)$$

构造未归一化投影子

$$\chi_l(r) = \Delta V_l(r) \phi_l(r), \quad (13)$$

并定义

$$W_l = \langle \chi_l | \chi_l | \chi_l | \chi_l \rangle, \quad Z_l = \langle \phi_l | \Delta V_l | \phi_l | \phi_l | \Delta V_l | \phi_l \rangle. \quad (14)$$

则归一化投影子与耦合系数为

$$\beta_l(r) = \frac{\chi_l(r)}{\sqrt{W_l}}, \quad D_l = \frac{W_l}{Z_l}. \quad (15)$$

最终的 KB 非局域项写成

$$\hat{V}_{\text{NL}}^{\text{KB}} = \sum_{l \neq l^*} \sum_m |\beta_{lm}\rangle D_l \langle \beta_{lm}|, \quad (16)$$

其中 $\beta_{lm}(r) = \beta_l(r) Y_{lm}(\hat{r})$ 。在平面波计算中，矩阵元可写成重叠积分乘积的和，从而将非局域势作用的计算复杂度显著降低。

2.4.3 局域道选择的经验

在许多元素中，常用较高角动量、且对基态贡献较小的通道作为局域道，例如 Al 案例中选择 d 通道作为 V_{loc} ，并对 s/p 通道通过非局域投影进行校正。这样做的动机是：价电子主导通道需要更严格地匹配散射性质，而高角动量通道对基态较不敏感，适合作为背景局域势。

2.5 可转移性验证方法

赝势生成后必须验证其可转移性，避免在不同化学环境下出现散射性质偏差或非物理束缚态。AtomPPGen 实现了三类常用验证：范数守恒检验、对数导数匹配与幽灵态检测。

2.5.1 范数守恒检验

TM 伪化在构造过程中通过方程组强制满足范数守恒，但仍可在数值上输出相对误差

$$\Delta_{\text{norm}} = \frac{Q_{\text{PS}} - Q_{\text{AE}}}{Q_{\text{AE}}}, \quad Q = \int_0^{r_c} u(r)^2 dr. \quad (17)$$

通常的验收标准为 $|\Delta_{\text{norm}}| < 10^{-6}$ ；在高精度实现下可达到 10^{-13} 量级的数值极限。本文 Al (LDA) 案例的验证报告给出：

$$|\Delta_{\text{norm}}^{(s)}| \approx 1.6 \times 10^{-16}, \quad |\Delta_{\text{norm}}^{(p)}| \approx 1.1 \times 10^{-13}, \quad |\Delta_{\text{norm}}^{(d)}| \approx 1.0 \times 10^{-13}. \quad (18)$$

2.5.2 对数导数 (Logarithmic Derivative) 匹配

对数导数定义为

$$L(E, r) = r \frac{d \ln \psi(E, r)}{dr} = r \frac{\psi'(E, r)}{\psi(E, r)}. \quad (19)$$

当 $r = r_{\text{test}}$ 取在核外区域 (通常 $r_{\text{test}} \gtrsim r_c$) 时, 散射解可写成携带相移 $\delta_l(E)$ 的渐近形式

$$u_l(r) \propto \sin \left(kr - \frac{l\pi}{2} + \delta_l(E) \right), \quad k = \sqrt{2E}, \quad (20)$$

代入得

$$L_l(E, r) = kr \cot \left(kr - \frac{l\pi}{2} + \delta_l(E) \right). \quad (21)$$

因此在固定 r_{test} 下比较 AE 与 PS 的 $L(E)$ 曲线, 等价于比较它们的散射相移随能量的变化。

极点与尖峰的来源 当某个能量下 $\psi(E, r_{\text{test}}) = 0$ 时, $L(E, r_{\text{test}})$ 存在极点 (分母为零), 曲线会出现尖峰。这不是赝势必然“失败”的信号, 而是对数导数定义本身的数学特征。在数值评价中, AtomPPGen 会对 $|L|$ 过大的点做离群过滤 (例如 $|L| > 50$), 以避免极点附近的单点对 RMS 指标产生不成比例的影响。

评价指标与阈值 AtomPPGen 采用两类紧凑指标:

1. **零点偏差 (可选):** 若扫描窗口内 AE/PS 都出现足够多的过零点, 可比较零点能量的均方根偏差

$$\Delta E_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i^{\text{AE}} - E_i^{\text{PS}})^2}, \quad (22)$$

典型阈值为 $\Delta E_{\text{RMS}} < 0.025 \text{ Ha}$ 。

2. **曲线 RMS (主要):** 在价区窗口 (例如 $E \in [-0.05, 0.05] \text{ Ha}$) 计算

$$L_{\text{RMS}}^{\text{val}} = \sqrt{\frac{1}{M} \sum_{j=1}^M [L_{\text{AE}}(E_j) - L_{\text{PS}}(E_j)]^2}. \quad (23)$$

对金属元素 (Al、Na 等) 取阈值 $L_{\text{RMS}}^{\text{val}} < 16.0$, 对共价元素 (Si 等) 取更严格阈值 $L_{\text{RMS}}^{\text{val}} < 3.0$ 。

2.5.3 幽灵态 (Ghost State) 检测: 能量感知分类

幽灵态是赝势哈密顿量中出现的非物理深束缚态, 若其能量落在价电子附近, 会导致平面波计算产生错误的基态或能带结构。在径向框架下, 可对每个 l 通道构造径向哈密顿量

$$\hat{H}_l = -\frac{1}{2} \frac{d^2}{dr^2} + V_l^{\text{PS}}(r) + \frac{l(l+1)}{2r^2}, \quad (24)$$

在有限盒区间上离散后对角化，得到能量谱并筛选能量窗口内的束缚态候选。

有限盒截断会把连续散射谱离散化为一系列“盒态”(box states)，它们并非真正束缚态。AtomPPGen 使用尾部比例

$$\tau = \frac{|\psi(R_{\max})|}{\max_r |\psi(r)|} \quad (25)$$

对盒态做初步判别 (τ 较大表示边界未充分衰减)。

进一步地，为减少对金属体系的虚假误判，引入能量感知分类：设 ε_{val} 为该通道参考价态能量，取能量容差 $\delta = 0.01 \text{ Ha}$ ，则

1. $E > 0$: 正能散射态（盒态离散化），归为安全态；
2. $0 > E > \varepsilon_{\text{val}} - \delta$: Rydberg 激发态序列（高主量子数束缚态），对基态 DFT 影响可忽略；
3. $E < \varepsilon_{\text{val}} - \delta$: 潜在危险幽灵态，需要结合 τ 二次判别。

该分类的一个重要物理含义是：在能量接近电离阈值 ($E \rightarrow 0^-$) 时，赝势可能支持一系列浅束缚态或近阈值态，它们与价态能量相隔较远（高于 $\varepsilon_{\text{val}} - \delta$ ），因此对基态 DFT 的影响可忽略。能量感知分类的目的正是将这类“安全态”与可能影响价区的真幽灵态区分开来，避免虚假误判。

3 数值方法

3.1 TM 非线性方程组求解

以 $N = 2$ 为例，未知参数为 $\alpha = (a_0, a_2, a_4, a_6)$ 。在 $r = r_c$ 处需要计算 $u_{\text{AE}}(r_c)$ 、 $u'_{\text{AE}}(r_c)$ 、 $u''_{\text{AE}}(r_c)$ ，并计算 AE 核内范数 $Q_{\text{AE}} = \int_0^{r_c} u_{\text{AE}}(r)^2 dr$ 。将 TM 内区形式及其导数代入约束即可得到残差方程组。

AtomPPGen 采用 `scipy.optimize.fsolve` 求解该非线性方程组。合理的初值对收敛性非常关键，常用估计为

$$a_0 \approx \ln \left(\frac{u_{\text{AE}}(r_c)}{r_c^{l+1}} \right), \quad a_{2i} = 0 \ (i \geq 1), \quad (26)$$

即保证在 r_c 处函数值量级正确，其余高阶系数从零开始，由求解器迭代调整。

3.2 数值稳定性要点

3.2.1 节点符号处理（以 p 通道为例）

高角动量通道的径向波函数常带节点；若在读取或插值过程中丢失节点处的符号信息，则核内积分与连续性约束会被破坏，表现为范数误差显著增大。AtomPPGen 在 TM 伪化实现中对节点符号进行了显式处理，从而保证范数守恒与后续 KB 投影构造的稳定性。

3.2.2 指数溢出与对数形式计算

TM 内区含有 $\exp(p(r))$, 当 $p(r)$ 在数值上过大时可能溢出。稳健做法是在必要时以对数形式计算

$$\ln u(r) = (l + 1) \ln r + p(r), \quad (27)$$

并在最终需要 $u(r)$ 时再指数化（同时配合裁剪），以避免中间溢出导致的 NaN。

3.2.3 对数导数尖峰的离群过滤

如前所述, $L(E, r)$ 在 $u(E, r) = 0$ 附近有极点。为避免少数尖峰主导 RMS 指标, AtomPPGen 在对数导数评价前过滤 $|L|$ 过大的点。在绘图与写作时也应采用过滤后的曲线作为展示版本, 同时在文字中说明尖峰的数学来源。

3.3 径向网格与 AtomSCF 的变量变换方法

AtomPPGen 的 AE 参考解由 AtomSCF 产生。AtomSCF 支持多种径向网格：线性网格、对数网格以及指数变换网格 (exp_transformed)。其中指数变换网格形式为

$$r_j = R_p (e^{j\delta} - 1), \quad j = 0, 1, \dots, N - 1. \quad (28)$$

配合变量变换 $u(r) = v(r) \exp(-r/(2R_p))$ 并取 $R_p = 1/\delta$, 可在离散化后消除一阶导数项, 使哈密顿量矩阵保持对称, 从而在同等网格规模下提升本征值精度。该方法同时包含 $r = 0$ 点, 对原子近核区解析性更友好。

4 程序实现

4.1 整体工作流与模块关系

AtomPPGen 的核心工作流可概括为：

1. 由 AtomSCF 求解 AE 原子 (LDA 模式) 得到 $\{u_{\text{AE}}^l(r), \varepsilon_l\}$;
2. 对每个通道执行 TM 伪化, 得到 $u_{\text{PS}}^l(r)$ 与范数误差;
3. 由伪波函数反演半局域势 $V_l^{\text{PS}}(r)$;
4. 选择局域道并构造 KB 投影子 β_l 与系数 D_l ;
5. 运行验证: 范数守恒、对数导数匹配、幽灵态检测;
6. 导出 JSON/NPZ 数据, 供文档与后续平面波程序使用。

对应的主要模块文件为：

- **ae_atom.py**: 封装 AtomSCF，生成 AEAtomResult (网格、波函数、能量、密度与能量分解等)；
- **tm.py**: TM 伪化求解器，输出 TMResult (含伪波函数、系数、范数误差等)；
- **invert.py**: 半局域势反演，输出 InvertResult ($V_l(r)$ 及诊断信息)；
- **kb.py**: KB 变换，输出局域势 V_{loc} 、投影子 β_l 与系数 D_l ；
- **validate.py**: 验证工具 (范数/对数导数/幽灵态) 与汇总报告；
- **export.py**: 导出 JSON/NPZ (当前实现以数值数组为主，UPF 格式框架已预留)。

更详细的 API 说明、教程与示例脚本可参阅项目在线文档。

4.2 关键接口示例（最小工作流）

下面给出一个尽量精简的示例，展示从 AE 求解到验证与导出的主流程（省略异常处理与绘图）：

Listing 1: AtomPPGen 最小工作流示例（节选）

```

1  from atomppgen.ae_atom import solve_ae_atom
2  from atomppgen.tm import tm_pseudize
3  from atomppgen.invert import invert_semitlocal_potential
4  from atomppgen.kb import kb_transform
5  from atomppgen.validate import run_full_validation
6  from atomppgen.export import export_pseudopotential
7
8  ae = solve_ae_atom(Z=13, spin_mode="LDA", lmax=2)
9
10 rc_map = {0: 2.1, 1: 2.2, 2: 2.4}
11 tm_results = {}
12 inv_results = {}
13 for l in (0, 1, 2):
14     tm_results[l] = tm_pseudize(
15         r=ae.r, w=ae.w,
16         u_ae=ae.u_by_l[l][-1],
17         eps=ae.eps_by_l[l][-1],
18         l=l, rc=rc_map[l],
19         continuity_orders=2,
20     )
21     inv_results[l] = invert_semitlocal_potential(tm_results[l], ae.r)

```

```

22
23 u_by_l = {l: ae.u_by_l[l][-1] for l in (0, 1, 2)}
24 kb = kb_transform(inv_results, u_by_l, ae.r, ae.w, loc_channel=2)
25
26 report = run_full_validation(ae, tm_results, inv_results, r_test=2.9)
27 export_pseudopotential(ae, tm_results, inv_results, report,
28                         output_prefix="test_results/al_lda", formats=[
29                           "json", "npz"])

```

4.3 测试覆盖

AtomPPGen 在 `tests/` 目录中提供了单元测试与集成测试，用于覆盖 TM 伪化、势反演、KB 变换与验证模块的关键路径。项目当前包含约 60 个测试用例，作为赝势生成与验证正确性的回归保障。

5 计算结果

5.1 Al (LDA) 验证结果

5.1.1 参数选择

以 Al ($Z = 13$) 为主案例，采用 LDA (Perdew – Zunger 1981 参数化) 作为交换关联泛函。推荐的截断半径与局域道选择如表1。

表 1: Al (LDA) 蕴势生成的推荐参数 (Bohr)。

r_c^s	r_c^p	r_c^d	KB 局域道
2.1	2.2	2.4	d 通道

5.1.2 波函数对比 (AE vs PS)

图1 展示 Al 的 $3s/3p$ 价层参考态以及 d 通道参考函数在 AE 与 PS 情况下的对比。由于 TM 伪化在 $r > r_c$ 强制 $u_{\text{PS}}(r) = u_{\text{AE}}(r)$ ，因此核外区域两条曲线应严格重合；核内区域 PS 波函数更平滑，且不引入额外节点。

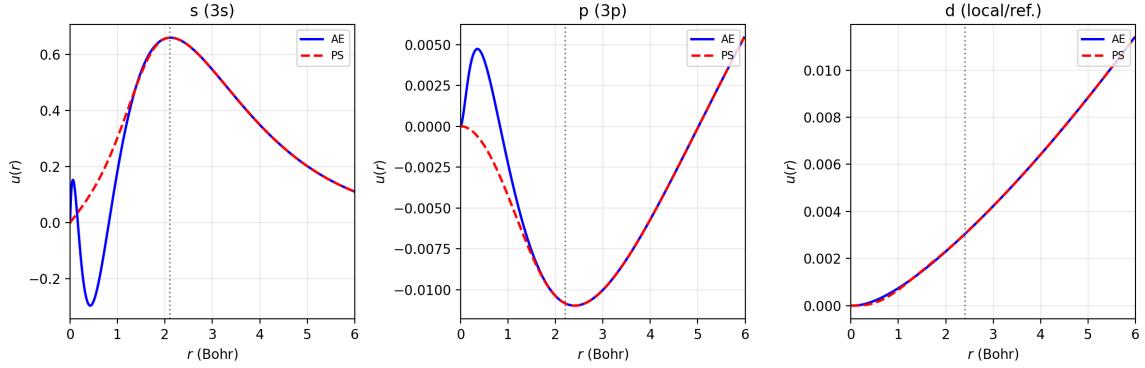


图 1: Al 各通道径向波函数对比: AE (实线) 与 PS (虚线)。垂直虚线为截断半径 r_c 。

半局域势的通道差异 在 TM 伪化与势反演后, 可得到三通道半局域赝势 $V_l^{\text{PS}}(r)$ 。图2给出 $s/p/d$ 通道势的对比以及对应的截断半径位置。该图用于直观展示: 不同 l 通道在核内的有效势形状存在差异; 在本 LDA 原子参考下, 远离原子核后势能逐渐衰减并趋于较小的量级, 从而核外区域的通道差异显著减弱。

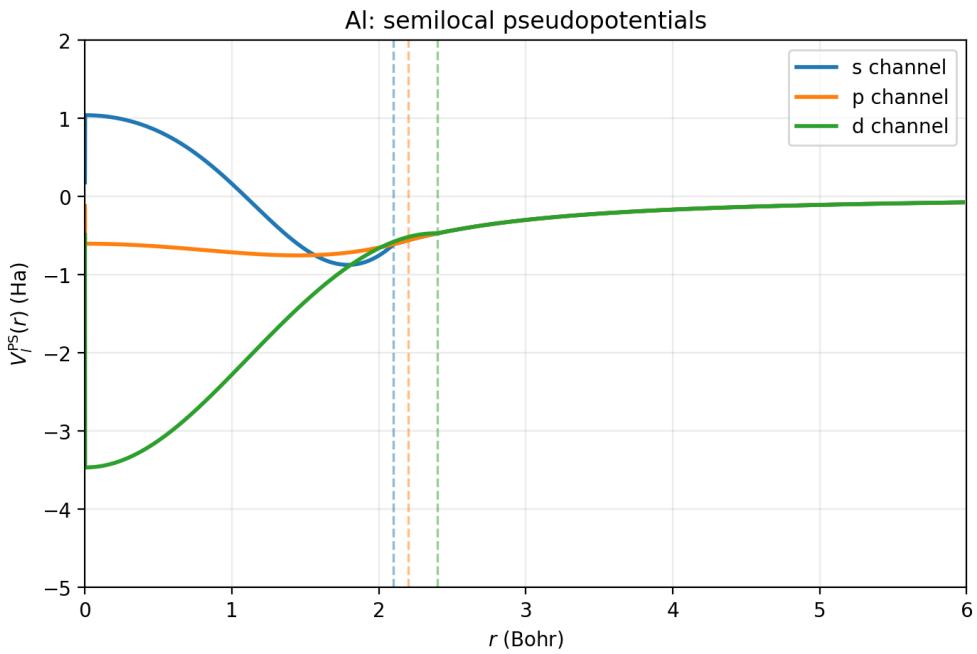


图 2: Al: $s/p/d$ 三通道半局域赝势 $V_l^{\text{PS}}(r)$ 对比 (虚线标注各通道截断半径)。

5.1.3 范数守恒精度

表2给出 Al 的 $s/p/d$ 通道范数误差数量级。验收标准为 $|\Delta_{\text{norm}}| < 10^{-6}$, 该实现能达到 10^{-13} 甚至 10^{-16} 量级, 属于数值极限范围。

表 2: Al (LDA) 范数守恒误差 (来自验证报告摘要)。

通道	最大 $ \Delta_{\text{norm}} $	标准要求	结论
s	1.58×10^{-16}	$< 10^{-6}$	PASS
p	1.08×10^{-13}	$< 10^{-6}$	PASS
d	1.04×10^{-13}	$< 10^{-6}$	PASS

5.1.4 对数导数匹配

对数导数验证在测试半径 r_{test} 处扫描能量窗口，并比较 AE 与 PS 的 $L(E)$ 曲线。对金属元素采用价区曲线 RMS 阈值 < 16.0 。Al 的推荐参数组在价区窗口的对数导数 RMS 如表3。

表 3: Al (LDA) 对数导数价区 RMS ($E \in [-0.05, 0.05]$ Ha)。

通道	$L_{\text{RMS}}^{\text{val}}$	阈值 (金属)	结论
s	8.70	< 16.0	PASS
p	0.60	< 16.0	PASS
d	0.10	< 16.0	PASS

图3 给出对数导数曲线示意。为保证测试点位于核外匹配区，本文采用

$$r_{\text{test}} = \max(r_c^l) + 0.5, \quad (29)$$

对本参数组有 $r_{\text{test}} = 2.9$ Bohr。需要强调： $L(E, r)$ 在 $\psi(E, r_{\text{test}}) = 0$ 附近会出现极点，因此曲线可能出现尖峰；在评价与绘图时应对 $|L|$ 过大的离群点做过滤，并在报告中注明该尖峰的数学来源。该取值用于本文的 Al 示例；Si、Na 的 r_{test} 取值以其验证文档的设置为准（见表6）。

为何金属元素的 s 通道差异更“显眼” 对金属元素（如 Al），在核外匹配区的有效势相对平缓，价电子散射解的相位对能量的变化往往较慢，从而使 $L_s(E, r_{\text{test}})$ 在价区呈现“缓变”特征。此时 AE 与 PS 若存在一定相位差，在 $L(E) = kr \cot(kr - l\pi/2 + \delta_l(E))$ 的映射下可能被放大，表现为两条曲线在纵轴上分离较明显。因此，金属元素的对数导数评价通常采用更宽松的阈值（本实现取 $L_{\text{RMS}}^{\text{val}} < 16.0$ ），并优先关注价区窗口 $E \in [-0.05, 0.05]$ Ha 内的 RMS 指标，而非要求整段能量范围内逐点重合。

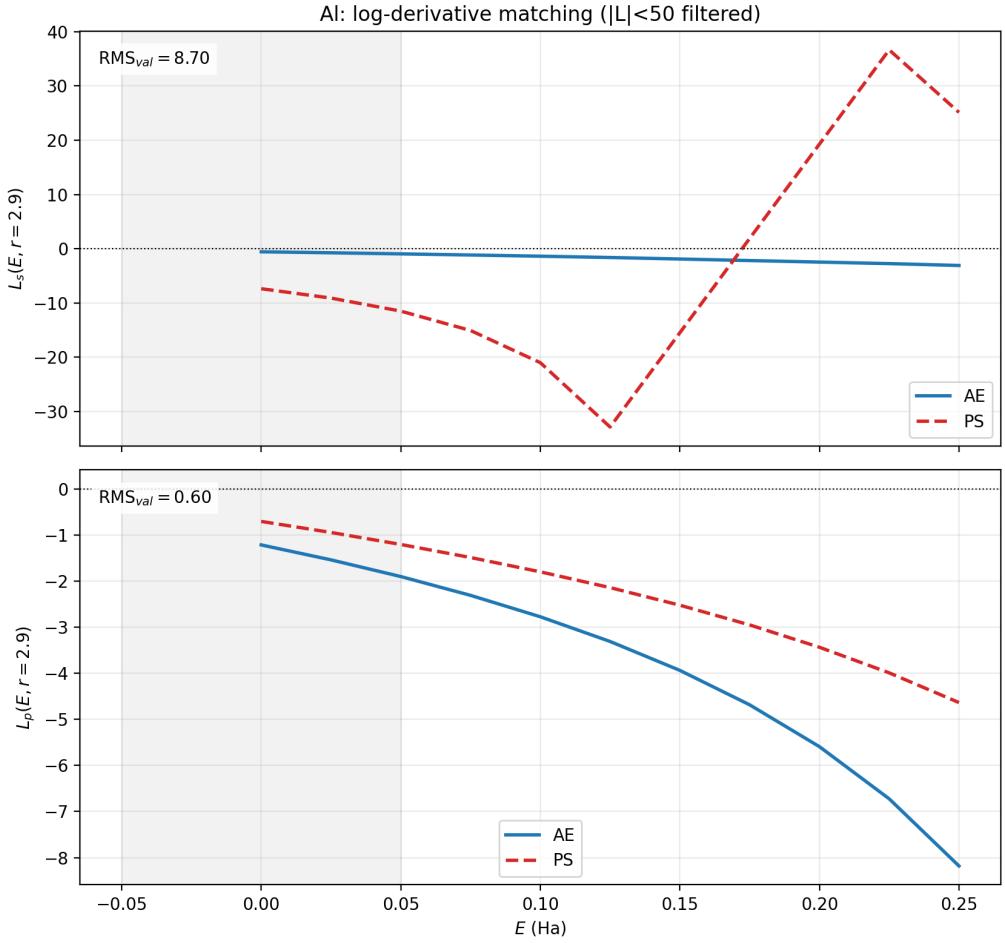


图 3: Al: 对数导数匹配曲线示意 (过滤 $|L| > 50$ 的离群点后绘制)。上: s 通道; 下: p 通道。

KB 投影子的径向形状 选择 d 通道作为局域道后, KB 非局域部分仅对 s/p 通道构造投影子 $\beta_l(r)$ 。图4 给出 $\beta_s(r)$ 与 $\beta_p(r)$ 的典型形状: 它们主要局域在核附近, 并在远离原子核后快速衰减, 这反映了 $\Delta V_l(r) = V_l(r) - V_{\text{loc}}(r)$ 的短程性质。

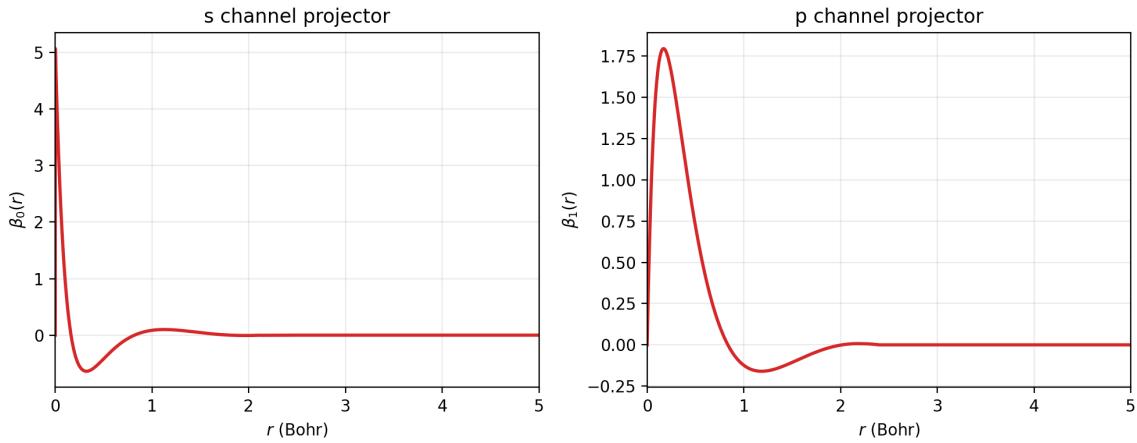


图 4: Al: KB 投影子曲线 (左: s 通道; 右: p 通道; d 通道作为局域道不构造投影子)。

5.1.5 幽灵态检测与能量感知分类

KB 形式下的幽灵态检测可视为对“是否出现额外束缚解”的数值排查：在有限径向边界上对径向哈密顿量对角化，得到一组离散本征值。由于有限边界会将连续散射谱离散化，同时 $E \rightarrow 0^-$ 附近存在一系列 Rydberg 激发态，因此需要结合能量位置与尾部指标进行分类，避免将“盒态/激发态”误判为真幽灵态。

本文以 s 通道为例，取价态能量 $\varepsilon_{3s} \approx -0.247$ Ha，并使用能量容差 $\delta = 0.01$ Ha： $E > 0$ 的点归为盒态/散射态离散化产物， $0 > E > \varepsilon_{3s} - \delta$ 的点归为 Rydberg 激发态，更深的 $E < \varepsilon_{3s} - \delta$ 才作为潜在幽灵态候选。对本参数组，在 $E \in [-0.15, 0.05]$ Ha 内 s 通道得到 3 个 Rydberg 态与 9 个盒态/散射态离散化产物，未检出真幽灵态。

图 5 将该能量窗口内的本征值按能量感知规则分类并可视化：绿色为 Rydberg 激发态，蓝色为盒态/散射态离散化产物，红色（若出现）为潜在幽灵态；同时标注了价态能量 ε_{3s} 及阈值 $\varepsilon_{3s} - 0.01$ Ha 的位置，便于直观理解分类边界。

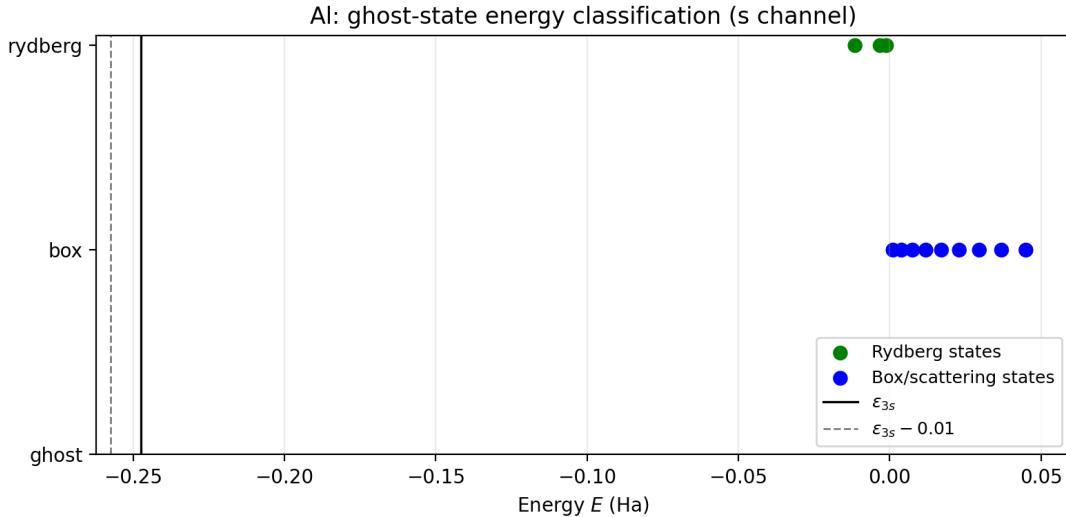


图 5: AI: 幽灵态能谱分类示意 (s 通道，能量感知分类)。黑实线为 ε_{3s} ，灰虚线为 $\varepsilon_{3s} - 0.01$ Ha。

表 4: AI (LDA) 幽灵态检测摘要（能量窗口 $E \in [-0.15, 0.05]$ Ha）。

通道	Rydberg 数	盒态/散射态数	真幽灵态数
s	3	9	0
p	-	-	1

其中 p 通道在本次验证中检出 1 个幽灵态计数；由于本文未给出其能谱分类图，这里仅作为风险提示，需要结合能量位置与尾部比例进一步判定其物理影响。

5.1.6 p 通道的节点符号处理影响

对 p 通道而言，若在伪化过程中未正确处理径向节点的符号信息，会导致范数守恒误差显著增大。表5给出不同 r_c^p 下，在未进行节点符号处理与进行节点符号处理两种情况下的范数误差对比。可以看到，加入节点符号处理后误差回到 10^{-13} 量级，且随 r_c^p 变化保持稳定。

表 5: p 通道节点符号处理对范数守恒的影响 (Al, LDA)。

r_c^p (Bohr)	未处理节点符号的范数误差	处理后范数误差
2.2	1.74×10^{-2}	1.08×10^{-13}
2.3	1.87×10^{-2}	1.31×10^{-16}
2.4	9.67×10^{-1}	4.51×10^{-15}

5.2 Si 与 Na 的对比案例（简述）

为展示“金属/共价”阈值策略的差异，AtomPPGen 还给出了 Si (共价) 与 Na (简单金属) 的教学示例验证结果。表6 汇总了典型参数与对数导数价区 RMS。

表 6: Si 与 Na 的示例验证摘要 (LDA)。

元素	(r_c^s, r_c^p, r_c^d) (Bohr)	r_{test} (Bohr)	通道	$L_{\text{RMS}}^{\text{val}}$	阈值
Si	(1.8, 2.0, 2.2)	6.0	s	2.31	< 3.0
			p	0.49	< 3.0
			d	0.31	< 3.0
Na	(2.6, 2.8, 3.0)	3.5	s	2.16	< 16.0
			p	0.12	< 16.0
			d	0.09	< 16.0

在这两组示例中，幽灵态检测均未发现真幽灵态；报告中出现的少量盒态为有限边界离散化产物，不影响通过判定。

6 误差分析

6.1 与 NIST 原子数据对比

AtomPPGen 的 AE 参考解依赖 AtomSCF 的非相对论 LDA/LSDA 实现。表7 给出以 Al 为例的能量对比 (单位 Ha)。

表 7: Al 原子能量与 NIST LSD 参考数据对比 (单位 Ha)。

项目	AtomSCF	NIST LSD	差异
总能量	-237.30	-241.32	$\sim 4 \text{ Ha} (\sim 1.7\%)$
1s 能级	-54.27	-55.15	$\sim 0.88 \text{ Ha}$
3s 能级	-0.25	-0.30	$\sim 0.05 \text{ Ha}$

需要指出: NIST 的参考值基于 LSD (局域自旋密度) 并可能包含特定的数值细节与边界处理; AtomSCF 的实现为教学用途, 在网格离散化、泛函参数化细节、以及收敛策略上可能与 NIST 使用的实现存在差异。对赝势生成而言, 更关键的是价层能级与散射性质的相对一致性, 因此价层 $3s/3p$ 的能量偏差 ($\sim 0.03 - 0.05 \text{ Ha}$) 通常可接受。

6.2 参数敏感性: 截断半径的影响

截断半径 r_c 决定赝势“软硬”程度:

- r_c 更小: 核内区域更接近 AE, 散射性质更易匹配, 但势更“硬”, 在平面波计算中需要更高的截断能量;
- r_c 更大: 蕴势更平滑更“软”, 平面波代价更低, 但对数导数匹配可能恶化, 且更易出现投影子数值不稳定或幽灵态风险。

因此实际应用中需在可转移性与计算效率之间折中。Al 的教学示例中通过参数扫描在多个组合中比较范数误差、对数导数 RMS 与幽灵态分类后, 选择表1 的组合作为稳定且便于演示的参数。

6.3 物理与方法局限性

本项目与本文档的讨论基于以下前提与近似:

1. **非相对论近似:** 未包含自旋轨道耦合与相对论修正, 对重元素需进一步扩展。
2. **局域/半局域 XC 泛函:** 以 LDA (PZ81) 为主, 未包含 GGA、meta-GGA 或杂化泛函效应。
3. **原子参考环境:** 蕴势以孤立原子参考态构造, 真实固体环境中的电荷转移与屏蔽效应需通过可转移性验证间接保证。
4. **幽灵态检测的盒效应:** 径向对角化在有限边界上进行, 必须结合尾部判据与能量分类排除盒态假信号。

7 总结

本文围绕 AtomPPGen 的模守恒赝势生成与验证流程，主要结论如下：

1. 采用 TM 伪化在核外强制匹配 AE 波函数，并通过连续性约束与范数守恒方程组得到核内平滑、无节点的伪波函数；
2. 通过势反演将伪波函数映射为半局域势，并在内区使用解析导数、外区使用样条导数，同时对节点附近做保护处理；
3. 通过 KB 变换将半局域势改写为“局域势 + 少量投影子”的可分离形式，便于平面波程序高效使用；
4. 验证方面，Al (LDA) 示例在范数守恒上达到 10^{-13} 量级精度；对数导数价区 RMS 满足金属阈值；幽灵态检测通过能量感知分类有效区分 Rydberg 激发态与真幽灵态；
5. Si (共价) 与 Na (金属) 的示例结果展示了不同元素类型下阈值策略的差异化处理。

参考文献

1. Troullier, N. & Martins, J. L. (1991). Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43, 1993 – 2006.
2. Kleinman, L. & Bylander, D. M. (1982). Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.*, 48, 1425 – 1428.
3. Giannozzi, P. (2019). Notes on pseudopotential generation. *Quantum ESPRESSO Documentation*.
4. Rappe, A. M. et al. (1990). Optimized pseudopotentials. *Phys. Rev. B*, 41, 1227 – 1230.
5. Gonze, X. et al. (2002). First-principles computation of material properties. *Comput. Mater. Sci.*, 25, 478 – 492.
6. Martin, R. M. (2004). *Electronic Structure: Basic Theory and Practical Methods*. Cambridge Univ. Press.
7. NIST Atomic Reference Data (LSD): nist.gov/pml/atomic-reference-data
8. AtomPPGen 项目仓库: github.com/bud-primordium/AtomPPGen
9. AtomSCF 项目仓库: github.com/bud-primordium/AtomSCF